

Molecular Aspects of Enzyme Catalysis



KODANSHA



Exclusive sales rights for

Japan, Taiwan, Hong Kong, Republic of Korea and People's Republic of China:

Kodansha Ltd., 12-21, Otowa 2-chome, Bunkyo-ku, Tokyo 112-01 (Japan)

Distribution for

Switzerland: VCH Verlags AG. P. O. Box, CH-4020 Basel (Switzerland)

Great Britain and Ireland: VCH Publishers (UK), Ltd., 8 Wellington Court, Wellington Street, Cambridge,

CB1 1HZ (Great Britain)

USA and Canada: VCH Publishers, 220 East 23rd Street, New York, NY 10010-4606 (USA)

for all other countries: VCH Verlagsgesellschaft, P. O. Box 101161, D-69451 Weinheim (Federal Republic of Germany)

ISBN3-527-30017-1 VCH Verlagsgesellschaft

ISBN1-56081-771-9 VCH Publishers

ISBN4-06-206497-9 Kodansha Ltd.

Molecular Aspects of Enzyme Catalysis

Edited by
Toshio Fukui
Kenji Soda



Tokyo

1994



Weinheim • New York
Basel • Cambridge • Tokyo

Toshio Fukui
Director and Professor
Institute of Scientific and
Industrial Research,
Osaka University
Ibaraki 567

Kenji Soda
Professor
Institute for Chemical
Research,
Kyoto University
Uji 611

Library of Congress Card No. applied for

A CIP catalogue record for this book is available from the British Library.

Die Deutsche Bibliothek-CIP-Einheitsaufnahme

Molecular aspects of enzyme catalysis/ed. by Toshio Fukui;

Kenji Soda.-Tokyo: Kodansha; Weinheim; New York;
Basel; Cambridge; Tokyo: VCH, 1994

ISBN 3-527-30017-1

NE: Fukui, Toshio [Hrsg.]

This book was carefully produced. Nevertheless, authors, editors and publishers do not warrant the information contained therein to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Published jointly by
Kodansha Ltd., Tokyo (Japan),
VCH Verlagsgesellschaft mbH, Weinheim (FRG) and
VCH Publishers Inc., New York, NY (USA)

Copyright ©Kodansha Ltd., Tokyo, 1994

All rights reserved. No part of this book may be reproduced in any form, by photostat, microfilm, retrieval system, or any other means, without the written permission of Kodansha Ltd.(except in the case of brief quotation for criticism or review).

Printed in Japan

To the late Professor Masanobu Tokushige



List of Contributors

Numbers in parentheses refer to the pages on which a contributor's paper begins.

Ahmed, Syed Ashrafuddin (127)

*National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health,
Building 8 Room 2A09, Bethesda, MD 20892, U.S.A*

Asada, Kozi (191)

The Research Institute for Food Science, Kyoto University, Uji, Kyoto 611, Japan

Futai, Masamitsu (211)

*Department of Organic Chemistry and Biochemistry, The Institute of Scientific and Industrial
Research, Osaka University, Ibaraki, Osaka 567, Japan*

Fukui, Toshio (73, 107)

The Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Osaka 567, Japan

Hyde, C. Craig (127)

*National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health
Building 8 Room 2A09, Bethesda, MD 20892, U.S.A.*

Iwamoto, Atsuko (211)

*Department of Organic Chemistry and Biochemistry, The Institute of Scientific and Industrial
Research, Osaka University, Ibaraki, Osaka 567, Japan*

Kagamiyama, Hiroyuki (15, 87)

Department of Medical Chemistry, Osaka Medical College, Takatsuki, Osaka 569, Japan

Kainosho, Masatsune (37)

Faculty of Science, Tokyo Metropolitan University, Hachioji, Tokyo 192-03, Japan

Kanematsu, Sumio (191)

*Department of Food Science and Technology, Minami-Kyushu University, Takanabe, Miyazaki
884, Japan*

Kawata, Yasushi (165)

Department of Biotechnology, Faculty of Engineering, Tottori University, Tottori 680, Japan

Kayastha, Arvind M. (127)

*National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health,
Building 8 Room 2A09, Bethesda, MD 20892, U.S.A.*

Lu, Zichun (127)

*National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health,
Building 8 Room 2A09, Bethesda, MD 20892, U.S.A.*

Maeda, Masatomo (211)

Department of Organic Chemistry and Biochemistry, The Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Osaka 567, Japan

Matsuzawa, Hiroshi (229)

Department of Agricultural Chemistry, The University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

Miles, Edith Wilson (127)

National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Building 8 Room 2A09, Bethesda, MD 20892, U.S.A.

Miyake, Yoko (37)

Faculty of Science, Tokyo Metropolitan University, Hachioji, Tokyo 192-03, Japan

Mori, Hiroyuki (107)

The Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Osaka 567, Japan

Ohta, Takahisa (229)

Department of Agricultural Chemistry, The University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

Ruvinov, Sergei B. (127)

National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Building 8 Room 2A09, Bethesda, MD 20892, U.S.A.

Soda, Kenji (1, 147)

Institute for Chemical Research, Kyoto University, Uji, Kyoto, 611, Japan

Tagaya, Mitsuo (73, 107)

The Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Osaka 567, Japan

Tanizawa, Katsuyuki (73, 107)

The Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Osaka 567, Japan

Torchinsky, Yuri M. (165)

The Institutes for Applied Research, Ben-Gurion University of the Negev, Beer-Sheva 84110, Israel

Ueno, Hiroshi (15, 87)

Department of Medical Chemistry, Osaka Medical College, Takatsuki, Osaka 569, Japan

Yang, Xiang-Jiao (127)

National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Building 8 Room 2A09, Bethesda, MD 20892, U.S.A.

Yonaha, Kazuo (1)

Department of Biotechnology, Faculty of Agriculture, University of The Ryukyus, Nishihara, Okinawa 903-01, Japan

Yoshimura, Tohru (147)

Institute for Chemical Research, Kyoto University, Uji, Kyoto, 611, Japan

Yumoto, Noboru (55)

Organic Material Department, Government Industrial Research Institute, Ikeda, Osaka 563, Japan

Contents

Dedication	v
List of Contributors	vii
Preface	xv
Frontispiece	xvii

Introduction—New Trends in Enzyme Studies 1

Enzyme Purification	1
Enzyme Structure	3
Enzyme Reaction Mechanisms	6
Novel Enzymes	8
Enzyme Applications	10
Perspectives	11
References	12

1. Advanced Protein Sequencing Techniques 15

1.1	Introduction	15
1.2	The Primary Structure of Aspartate Aminotransferase (AspAT)	17
1.3	Progress in Primary Structure Analysis	20
1.3.1	Direct Sequencing of Protein	20
1.3.2	A Sequencing Approach Based on cDNA	30
1.3.3	Application of Mass Spectrometry to Protein Primary Sequence Determination	32
1.3.4	Database	33
1.4	Conclusions	33
	References	34

2. Nuclear Magnetic Resonance Studies on *Streptomyces* Subtilisin Inhibitor and its Complexes with Proteinases 37

2.1	Introduction	37
2.2	Assignment of the Backbone Peptide Carbonyl Carbon ^{13}C NMR Signals for Larger Proteins	39
2.3	Structural Elucidation of the Stable Intermediate for the SSI-Subtilisin System by ^{13}C NMR Spectroscopy	41

2.4	Facile Restoration of the Hydrolyzed Scissile Bond of SSI* in the Subtilisin Complex	45
2.5	Global Conformational Changes of SSI Induced by Complex Formation with Subtilisin	47
2.6	Conformations of the Flexible Segments of Domain II and the Topology of Proteinase Surface are Interdependent	49
2.7	Summary and Conclusion	51
	References	52
3.	Subunit Assembly of Oligomeric Enzymes	55
3.1	Introduction	55
3.2	Functional Roles of Subunit Assembly and Disassembly	55
3.3	Determination of Composition and Geometric Arrangements of Subunits	57
3.4	Determination of Subunit Contact Regions and Strength of Subunit Interactions	58
3.4.1	Identification of Subunit Contact Regions	58
3.4.2	Evaluation of the Strength of Interactions Between Subunits	59
3.5	Experimental Approaches to Reveal the Relationship Between Oligomeric Structure and Biological Activity	61
3.5.1	Analysis of Dissociation and Association Processes in Oligomeric Proteins	61
3.5.2	Hybridization Experiments	61
3.5.3	Site-directed Mutagenesis at the Subunit Interface	66
3.5.4	Analysis of Subunit Association and Dissociation in the Cell	66
3.6	Proteins Involved in Subunit Assembly <i>in vivo</i>	66
3.6.1	Molecular Chaperones	66
3.6.2	Other Proteins Involved in Subunit Assembly	69
3.7	Conclusion	69
	References	70
4.	Affinity Labeling of Nucleotide-Binding Sites	73
4.1	Introduction	73
4.2	Nucleoside Polyphosphopyridoxals (NP _n -PL)	75
4.2.1	Synthesis	75
4.2.2	Inactivation Kinetics	75
4.2.3	Identification of Labeled Sites	76
4.3	Case Studies	76
4.3.1	Glycogen Synthase	76
4.3.2	Adenylate Kinase	77
4.3.3	UDP-Glucose Pyrophosphorylase	78
4.3.4	F ₁ -ATPase	81
4.3.5	Ion-Transporting ATPases	82
4.3.6	Protein Kinases	83

4.4	Concluding Remarks	84
	References	85
5.	Structure and Function of Aspartate Aminotransferase	87
5.1	Introduction	87
5.2	Background	87
5.3	Site-directed Mutagenesis on AspAT	93
5.4	X-ray Crystallographic Study on AspAT	93
5.5	Evaluation of the Active Site Residues in AspAT by <i>in vitro</i> Mutagenesis	94
5.5.1	Lys-258 and Tyr-70*	95
5.5.2	Tyr-225	97
5.5.3	Asp-222 and His-143	98
5.5.4	Arg-292* and Arg-386	98
5.5.5	Trp-140	99
5.5.6	Cys Residues	99
5.6	Substrate Binding Pocket	101
5.7	Future Prospects	104
	References	104
6.	Higher Plant Phosphorylases	107
6.1	Introduction	107
6.2	Phosphorylase Isozymes from Higher Plants	108
6.3	Structures of Phosphorylase Isozymes from Potato Tuber	109
6.3.1	Type-L Isozyme	109
6.3.2	Type-H Isozyme	111
6.3.3	Sequence Comparison and Stereostructure	113
6.3.4	Functional Role of the 78-Residue Insertion in Type-L Isozyme	119
6.4	Catalytic Mechanism	120
6.5	Evolution of Plant Phosphorylases	122
6.6	Concluding Remarks	123
	References	123
7.	Tryptophan Synthase	127
7.1	Introduction	127
7.2	Structure and Function of the α Subunit	128
7.2.1	Structure	128
7.2.2	Active Site Residues	129
7.2.3	Flexible Loop	130
7.2.4	Protein Folding and Stability	131

7.3	Structure and Function of the β Subunit	131
7.3.1	Structure and Domains	131
7.3.2	Pyridoxal Phosphate Intermediates	133
7.3.3	Active Site Residues	133
7.4	The Indole Tunnel and Indole Channeling	138
7.5	Interactions Between the α and β Subunits	140
7.6	Reciprocal Communication Between the α and β Subunits	141
7.7	Catalytic Mechanism of the Physiological $\alpha\beta$ Reaction	142
7.8	Unanswered Questions and Future Directions	142
	References	144
8.	Alanine Racemase: Structure and Functions	147
8.1	Physiological Function of Alanine Racemase	147
8.2	Genetical and Evolutional Aspects of Alanine Racemase	148
8.3	Structure of Thermostable Alanine Racemase of <i>Bacillus stearothermophilus</i>	150
8.3.1	Subunit Dissociation and Unfolding	150
8.3.2	Limited Proteolysis	152
8.3.3	Construction and Expression of Fragmentary Enzyme	152
8.3.4	Unfolding and Reconstitution of Fragmentary Enzyme	154
8.4	Reaction Mechanism of Alanine Racemase	155
8.5	Inhibitors of Alanine Racemase	159
	References	162
9.	Tryptophanase: Structural, Spectral, and Catalytic Properties	165
9.1	Introduction	165
9.2	Structural Properties	166
9.2.1	Primary and Quaternary Structure	166
9.2.2	Unfolding and Refolding	170
9.3	Spectral Properties of the Free Enzyme and Catalytic Intermediates	172
9.3.1	Absorption Spectra	172
9.3.2	Circular Dichroism	174
9.3.3	Fluorescence	177
9.3.4	NMR Spectra	178
9.4	Chemical Modification and Site-Directed Mutagenesis	179
9.5	Catalytic Properties	182
9.5.1	Activation by Cations	182
9.5.2	Coenzyme Reorientations in the Active Site	183
9.5.3	Catalytic Mechanism	186
	References	188

10 Superoxide Dismutase 191

- 10.1 Introduction 191
- 10.2 Distribution of SOD among Organisms 192
 - 10.2.1 Phylogenic Distribution 192
 - 10.2.2 Subcellular Localization 192
- 10.3 Physiological Functions 193
 - 10.3.1 Defense against Oxygen Toxicity 193
 - 10.3.2 Gene Regulation 194
- 10.4 Molecular Properties 195
 - 10.4.1 CuZn-SOD 195
 - 10.4.2 Fe-SOD and Mn-SOD 200
- 10.5 Properties of Superoxide Radical and Analytical Methods 201
 - 10.5.1 Superoxide Anion Radical as a Substrate 202
 - 10.5.2 Diffusion-controlled Reaction 202
 - 10.5.3 Assay of SOD 202
- 10.6 Kinetics and Reaction Mechanisms 204
 - 10.6.1 CuZn-SOD 204
 - 10.6.2 Fe-SOD and Mn-SOD 207
 - 10.6.3 Inactivation by H_2O_2 208
- References 209

11. H^+ -ATPase: Catalysis and H^+ Translocation 211

- 11.1 Introduction 211
- 11.2 Structure and Function of F-type ATPase 213
 - 11.2.1 Preparation of F_0F_1 213
 - 11.2.2 Assembly and Structure of F_0F_1 214
 - 11.2.3 Uni-site and Multi-site Catalysis by F_1 216
- 11.3 Catalytic Site of F-type ATPase 217
 - 11.3.1 Affinity Labeling of Catalytic Site of F_1 ATPase 217
 - 11.3.2 Mutations in the Glycine-rich Sequence 219
 - 11.3.3 Amino Acid Residues Interacting with Glycine-rich Sequence 220
- 11.4 Pathway and Regulation of H^+ Transport 221
 - 11.4.1 Assembly of F_0 221
 - 11.4.2 Proton Pathway in F_0 221
 - 11.4.3 Regulation of Proton Conduction 225
- 11.5 Conclusion 226
- References 226

12.	Stabilization of Proteases by an Engineered Disulfide Bond	229
12.1	Aqualysin I is a Thermostable Protease	230
12.2	Aqualysin I Belongs to a Family of Subtilisin-type Serine Proteases	230
12.3	Comparison of Known Three-dimensional Structures of Subtilisin-type Serine Proteases: Probable Structural Similarity Between Aqualysin I and Other Subtilisin-type Enzymes	232
12.4	Aqualysin I Contains Two Disulfide Bonds	232
12.5	Introduction of a Disulfide Bond into Subtilisin by Site-directed Mutagenesis	234
12.5.1	Mutant Subtilisin E Having a Cys61-Cys98 Disulfide Bond Gains Thermostability	234
12.5.2	Another Mutant Did Not Form Disulfide Linkage	237
12.6	Other Attempts to Increase the Thermostability of Subtilisin are Unsuccessful	237
12.7	Another Example, an Alkaline Protease from <i>Aspergillus oryzae</i>	237
12.8	Conclusions	238
	References	238
	Index	241

Preface

Enzymes play the central role in cellular functions of a living organism; most reactions in biological systems do not occur in the absence of enzymes. Because of their high selectivity and enormous power, enzymes are also used as biocatalysts in the syntheses of many useful compounds. Therefore, studies on enzyme catalysis and mechanism have been extensively performed by many biochemists for a long time. Yet, the precise relationship between the structure and function of enzymes remains to be elucidated. It is still difficult to answer even such simple questions on a molecular basis as: Why are enzymes so powerful? Why are enzyme reactions so selective?

Fortunately, however, the methodological advances in the field of protein science during the past decade have led to an explosive growth of knowledge of enzyme structures and methods for the design and manufacture of engineered proteins. The aim of this volume is to provide the latest information mainly from Japan concerning enzyme structure and function studies and to stimulate the redesign of more effective biocatalysts by various rapidly developing new methodologies.

The original idea for this volume was provided by the late Professor Masanobu Tokushige of Kyoto University who wrote us about the plan from his bed. After his sudden death at the early age of 62, we have succeeded in achieving his idea with the help of Mr. Ippei Ohta of Kodansha Scientific Ltd. whom we gratefully acknowledge. Professor Tokushige was not only an outstanding scientist but was very friendly to all persons including many from overseas. His early death is a great loss for everyone who knew him and liked him. We, with all the authors, would like to dedicate this book to our close friend, "Toku-san."

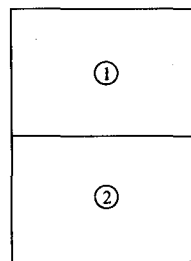
August 15, 1993

Toshio FUKUI
Kenji SODA

Legend for the color plates ① to ④

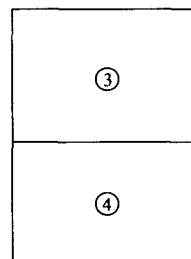
Schematic views based on the crystal structure of the tryptophan synthase $\alpha_2\beta_2$ complex (see Chapter for details).

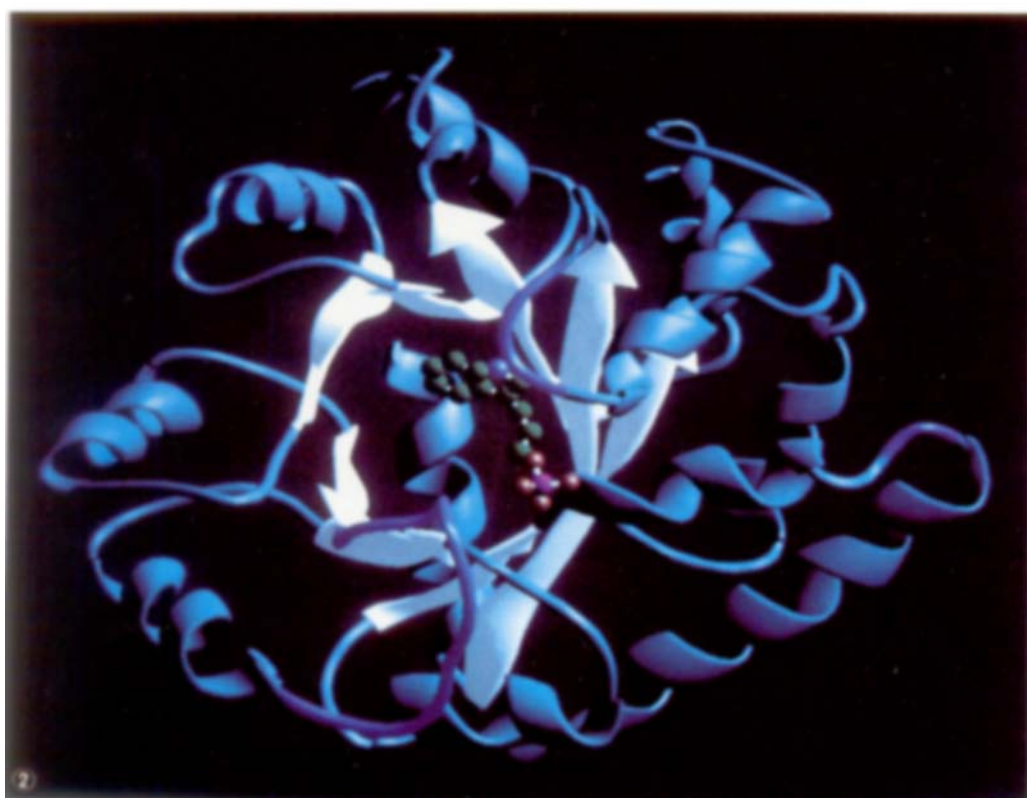
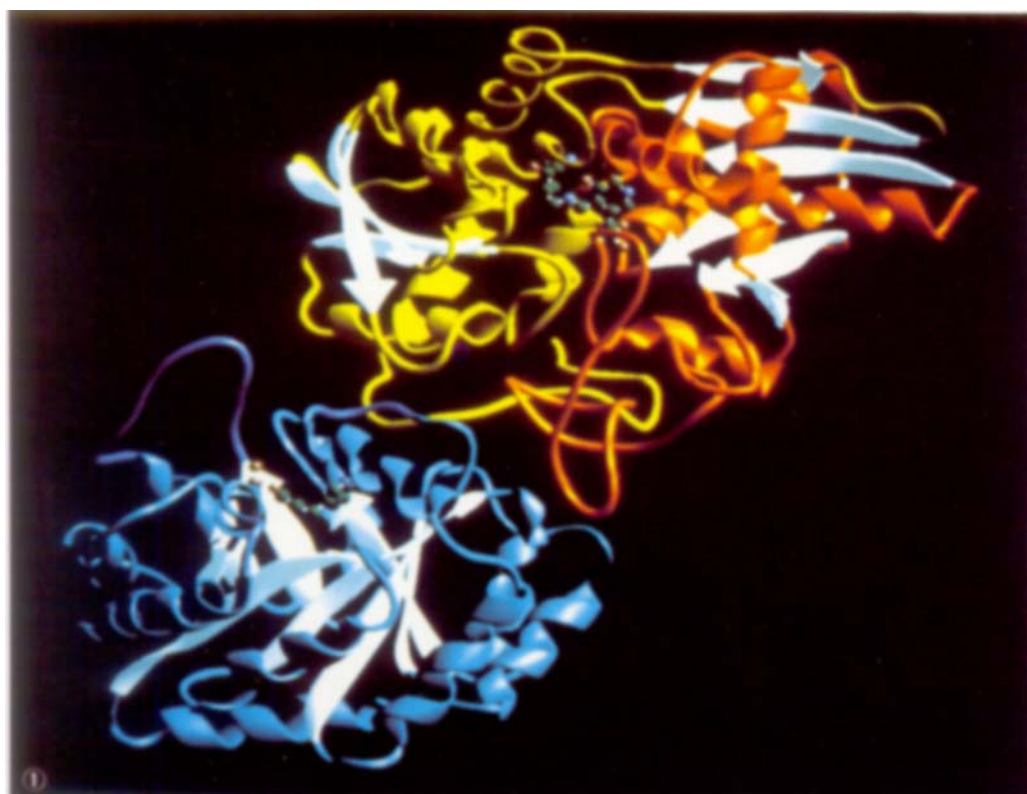
- ① View of an α/β pair. The α subunit (blue) is shown with the eight strands of the central barrel (white) surrounding the bound substrate analog (indole propanol phosphate) at the active center. Residues 57–59 (“disordered loop”) and 177–191 (“Flexible loop”) are shown in darker blue. The β subunit N-domain (yellow) and C-domain (orange) are shown surrounding the coenzyme pyridoxal phosphate at the active site. The central β -sheet strands of each domain are also shown in white. Note the region of the C-domain which does not have a well-defined secondary structure and which is in contact with the α subunit.
- ② This view of the α subunit shows that it has the overall peptide fold of an α/β barrel structure. In addition to the eight strands (white) and eight helices (blue) found in a typical α/β barrel structure, the α subunit contains at least three additional helices. An extra N-terminal helix can be seen at the bottom of the barrel beneath the bound indole propanol phosphate. This inhibitor is shown in a ball-and-stick representation with the phosphate in red. Figures were generated using the ribbons program by Mike Carson, University of Alabama, Birmingham, Alabama.

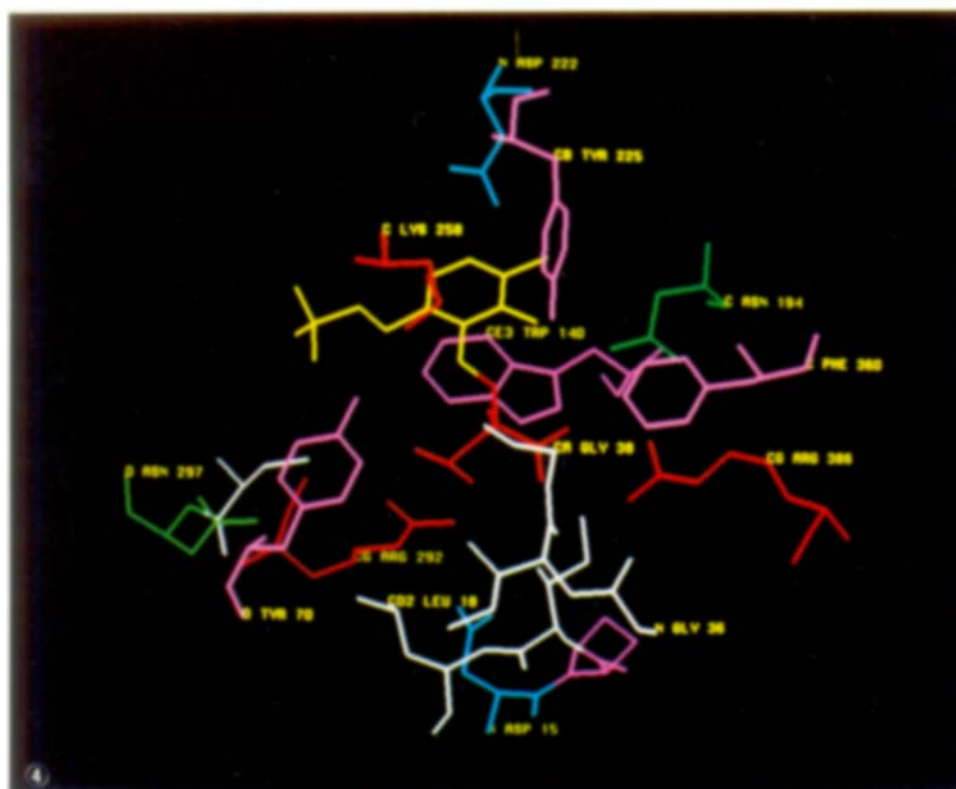
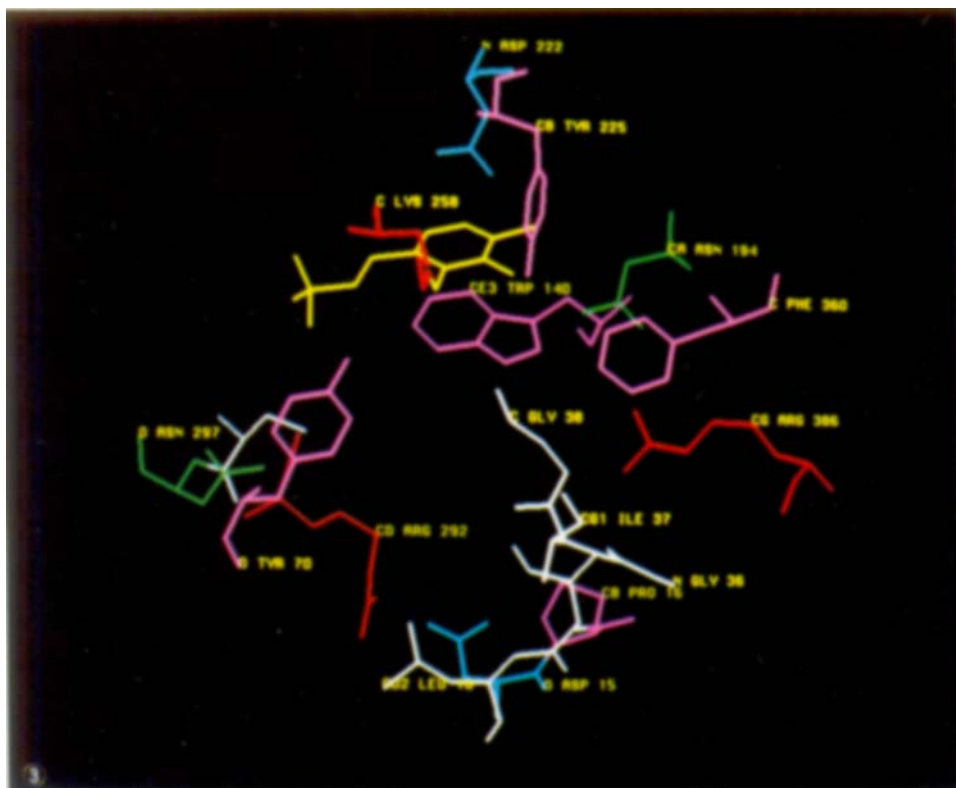


Active site structure of *Escherichia coli* aspartate aminotransferase (see Chapter 5 for details).

- ③ The coenzyme, pyridoxal 5'-phosphate (PLP), forms an aldimine linkage with the ϵ -amino group of Lys-258 via its aldehyde group. PLP also interacts with Asp-222, Tyr-225 and Asn-194 by forming hydrogen bonds; thus it maintains proper orientation at the active site. The enzyme is in the open conformation.
- ④ Conformational change upon binding of substrate: Enzyme-substrate complex with 2-methylaspartate (2-Me-Asp) is shown (red structure in the middle is 2-Me-Asp). PLP forms an external aldimine linkage with 2-Me-Asp; thus it frees the ϵ -amino group of Lys-258. Upon binding of 2-Me-Asp, the PLP ring tilts, the α -carboxyl and distal carboxyl groups of 2-Me-Asp form hydrogen bonds with the guanidinium group of Arg-386 and Arg-292, respectively, where the side chain position of Arg-292 shifts away from its original position. While Arg-292 has such motion, other residues in the small domain (white) and Asp-15 (blue), the latter of which was originally hydrogen bonded with Arg-292, have moved in such a way as to try to wrap 2-Me-Asp. The enzyme-substrate complex is in the closed conformation.







Introduction

New Trends in Enzyme Studies

Kazuo Yonaha*¹ and Kenji Soda*²

*¹ Department of Biotechnology, Faculty of Agriculture, University of The Ryukyus,
Nishihara, Okinawa 903-01, Japan

*² Institute for Chemical Research, Kyoto University,
Uji, Kyoto 611, Japan

Since the first clear recognition of an enzyme was made by Payen and Persoz in 1833¹⁾ when they found that an alcohol precipitate of malt extract contained a thermolabile substance which converted starch into sugar (named diastase, meaning separation in Greek and now known as amylase), there have been a number of enzymes discovered as described in voluminous treatises including *Methods in Enzymology* and *The Enzymes*. The term⁶⁾ "enzyme," meaning "in yeast," was introduced by Kühne in 1878. In 1898 Duclaux proposed the suffix "-ase" for use as a general term for enzyme to be attached to a root indicating the substrate on which the enzyme acts. This terminology provides the basis for the systematic nomenclature of enzymes which is still in use; a few names ending in "-in" have been employed for digestive enzymes, especially proteinases such as pepsin and trypsin.

The chemical nature of enzyme was controversial for a long time, until Buchner succeeded in isolating an enzyme system (zymase) from yeast in a cell-free extract in 1897.²⁾ Urease was then crystallized by Sumner in 1926,³⁾ followed by crystallization of several proteolytic enzymes by Northrop and his colleagues. At present the chemical nature of enzyme is defined as "a protein with catalytic activity based on the specific activation of its substrate." However, this definition has been somewhat open to debate since a catalytic RNA, ribozyme, was discovered in 1982.

In this section is an overview of some historical aspects of enzyme studies with special emphasis on new methods of purification, structure determination, and research on the reaction mechanism of an enzyme. Enzyme application for medical and industrial use and development of novel enzymes such as catalytic RNA (ribozyme) and catalytic antibody (abzymes) are also briefly described.

Enzyme Purification

In the early days of enzyme study attention was focused mainly on glycolysis in muscle, fermentation in yeast and oxidative phosphorylation, all of which provided the basis for vast and rapid development of biochemistry. Dixon and Webb estimated the number of enzymes to be about 650 in their book, *Enzyme*, in 1957. The number has grown to about 3000 at the present time and continues to increase (Fig. 1. 1).

Enzymes are extracted from cell or tissue by grinding with sea sand, size-controlled glass beads, or aluminium oxide, or by ultrasonic disruption of cell walls. The enzyme first

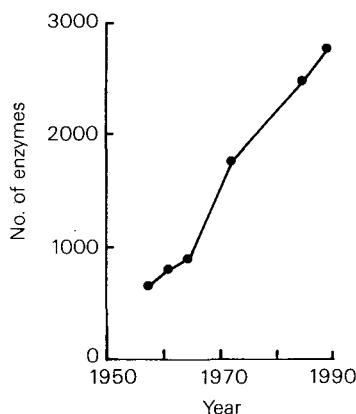


Fig. 1.1 Number of enzymes.

crystallized, urease, was isolated by fractionation with ammonium sulfate, although the first crystals were far from pure. This procedure, however, was followed by the purification of other enzymes.

Column chromatography, developed in the early 1950's, has become the most effective method of enzyme purification. The mechanism of separation depends on adsorption, ion exchange, specific affinity to immobilized ligands or molecular sieving effects. Various types of ion exchanger, anion or cation exchange resin (or cellulose), gel matrixes for molecular sieving chromatography such as dextran gels (Sephadex), polyacrylamide gels (Bio-gel), and agarose gels (Sephacrose) have been developed and employed successfully for enzyme purification. In affinity chromatography ligands can be substrates, coenzymes, inhibitors, activators, antibodies, or carbohydrates.⁴⁾ Group specific, dye-ligand affinity chromatography has also been developed for the purification of a group of enzymes having affinity to adenine nucleotide containing compounds such as NAD, FAD, and ATP.⁵⁾ Various dehydrogenases, kinase, albumins or interferon have been purified by this chromatography. A dye containing triazine, for example, Cibacron Blue (or Red), or Blue (or Red) Sepharose 4B, are widely used. There are two types of chromatography operating in a high concentration of salts, *i.e.* hydrophobic chromatography⁶⁾ and hydrogen bond chromatography.⁷⁾ In hydrophobic chromatography interaction occurs between a hydrophobic ligand (an alkyl or phenyl group) bound to the support matrix and the hydrophobic region of the enzyme protein. The interaction via hydrogen bonding between a hydrophilic matrix such as ion exchange cellulose and an enzyme protein molecule serves as the dominant force by blocking ionic interaction under a high concentration of salt.

Innovation has been brought to isolation techniques in biochemistry and molecular biology including enzyme purification. When high performance liquid chromatography (HPLC) system came into general use, the HPLC system greatly improved the efficiency of enzyme purification by reducing operation time and high resolution potential. It has become even more important due to the recent marked development in genetic engineering techniques, in which purification of a gene product (protein) is required for a gene cloning. Thus, the enzyme is first purified even in a small amount by HPLC, its partial primary

structure is determined, then genetic engineering technique is applied to clone the gene with an oligonucleotide synthesized according to the sequence as a probe to express efficiently the cloned gene in the host cell. After enhancement of enzyme productivity, either a conventional procedure or an HPLC system can be employed for the enzyme purification. Although capillary electrophoresis has higher resolution potential than HPLC, this method cannot be applied to the preparative scale of enzyme purification. HPLC system with a microbore column may be an alternative to this procedure.

Genetic engineering works elegantly in some cases of enzyme purification, especially those of thermostable enzymes from thermophilic microorganisms. The gene coding an enzyme from a thermophile is introduced by gene manipulation into mesophile cells such as *Escherichia coli*. After expression of the gene, the thermostable enzyme can be easily separated from usual thermolabile enzymes (or proteins) that are inherently produced by host cells by heating cell-free extract from mesophile host cells. The resulting precipitate containing denatured proteins is removed by centrifugation. In this manner L-leucine dehydrogenase from *Bacillus stearothermophilus* expressed in *E. coli* was purified by a single column chromatography with DEAE-cellulose after heat treatment of cell-free extract (Table I. 1).^{8,9)} Tryptophan synthase of *E. coli* purified by a direct crystallization of the enzyme from crude cell-free extract prepared from the cells enhanced the gene expression. Thus, it seems likely that the procedure of enzyme purification can be greatly improved by a combination of advanced separation technology and genetic engineering technique.

TABLE I.1 Purification of Leucine Dehydrogenase of *Bacillus stearothermophilus*

(A) From <i>B. stearothermophilus</i> ⁸⁾				
Steps	Total protein (mg)	Total activity	Specific activity	Yield (%)
Crude extract	5,500	825	0.15	100
Polyethyleneglycol	1,280	948	0.74	115
DEAE-cellulose	681	804	1.18	92
Ammonium sulfate	280	868	3.06	98
Sephadex G-150	139	577	4.15	66
1st Preparative electrophoresis	6.06	494	82.0	57
2nd Preparative electrophoresis	2.67	320	120	39
(B) From <i>Escherichia coli</i> MV1184 Carrying the Enzyme Gene ⁹⁾				
Steps	Total protein (mg)	Total activity	Specific activity	Yield (%)
Crude extract	170	12,000	71	100
Heat treatment	75	9,000	120	75

Enzyme Structure

Study of enzyme structure often means elucidation of protein structure. In the early stage of enzyme studies, protein was considered to be the heterogeneous macromolecule that could be never determined by methods available in those days. In 1926, Svedberg developed an ultracentrifugation technique to determine molecular weight of macro-

molecules besides physicochemical constants such as diffusion and sedimentation coefficients.¹¹⁾ Although the determination of molecular weight of protein by ultracentrifugation analysis had been theoretically established, the experimental procedure was complicated and an expensive instrument was required. In the 1960's a more convenient method for the determination of molecular weight of protein was established and is still in general use, although the process was rather empirical like gel-filtration¹²⁾ and SDS-polyacrylamide gel electrophoresis.¹³⁾ By a combination of these methods, molecular weight and subunit structure of enzyme are determined. Direct observation of enzyme molecules using an electron microscope was performed for a few enzymes, revealing subunit structures of enzymes such as fructose diphosphate aldolase¹⁴⁾ and glutamine synthetase.¹⁵⁾

In 1958, Spackman, Moore and Stein first constructed an automatic amino acid analyzer, which enabled us to characterize proteins in chemical terms. This was followed by their brilliant work in determining the primary structure of ribonuclease A.¹⁶⁾ Several hundreds micromoles of sample and almost one day were required for one analysis with their prototype machine. The sensitivity and analysis time, however, have been greatly improved by the introduction of size-controlled ion exchange resin and fluorescence detection system with fluorescamine or *o*-phthalaldehyde. The analysis of pre-column labeled amino acids with phenylisothiocyanate is a convenient procedure in the primary structure determination of proteins.¹⁷⁾ With these improved procedures amino acids can be determined at the picomole or even femtomole level within 20 to 30 min. It has been assumed that all enzyme proteins consist of only L-amino acids. Enantiomer analysis has not been carried out for amino acids of proteins. However, there have been some reports concerning the possible presence of D-amino acids in proteins,^{18,19)} and demand for enantiomer analysis of amino acids is increasing. An amino acid analyzer capable of determining protein-composing amino acid enantiospecifically has been developed (Fig. I. 2).²⁰⁾

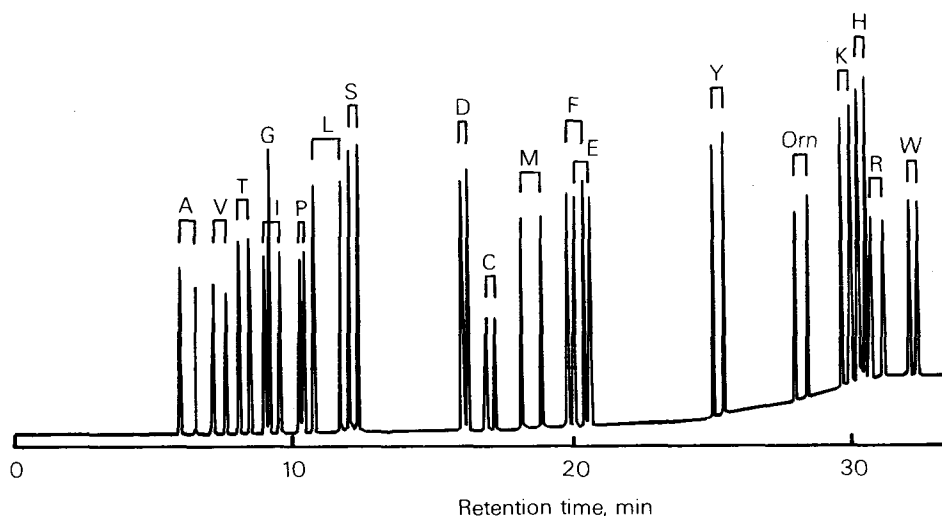


Fig. I. 2 Enantiomeric analysis of amino acids.
D-Isomer is eluted before L-isomer in each amino acid.

The primary structure of protein was not considered to be determined, until the structure of insulin was elucidated by Sanger and Thompson in 1953.²¹⁾ In their work determination of amino terminal amino acid of peptides derived from insulin by fragmentation with HCl played an important role. They used 2,4-dinitrofluorobenzene as amino terminal labeling reagent because of its high stability and color development of the derived amino acids. Several other chemicals have been developed for more precise determination of amino terminal amino acids; thiocyanate,²²⁾ dansyl chloride,²³⁾ 4-*N,N*-dimethylazobenzene-4'-isothiocyanate²⁴⁾ and phenylisothiocyanate.²⁵⁾ In 1967 Edman and Begg developed an automatic amino acid sequencer (sequenator) of protein (peptide) based on the reaction with phenylisothiocyanate—a liquid phase method.²⁶⁾ A solid phase method was developed by Laurensen based on the same principle, but the sample is immobilized on glass beads.²⁷⁾ The phenylthiocarbamate derivatives of amino acids (PTC-amino acids) are converted into phenylthiohydantoin of amino acids (PTH-amino acids) in a dilute acid or amino acids by back hydrolysis. These are analyzed on a thin layer chromatography or gas-liquid chromatography, and an amino acid analyzer, respectively. Although these procedures are efficient, their sensitivity and accuracy are not. Hunkapiller and Hood²⁸⁾ have improved these drawbacks by an exhaustive purification of chemicals employed for the Edman reaction, keeping the whole system under high vacuum (not more than a few microtolls), modifying the whole process to be completely automatic by developing an automatic conversion system, originally designed by Wittmann-Liebold, Graffunder and Kohls,²⁹⁾ and introducing HPLC to the PTH-amino acid analysis. With this improved sequencer Hunkapiller and Hood³⁰⁾ determined the amino acid sequence of interferon up to 20 amino acid residues with a picomole samples. The improvement of the Edman chemistry has led to innovation in the sequencing of protein: the development of a gas phase sequencer.³¹⁾ The gas phase machine has handed the sequencing work from a skillful and well-experienced researcher to the researcher less-experienced in Edman chemistry. At present most biochemists routinely determine the amino acid sequence of protein with a gas-phase sequencer, although skill in the Edman chemistry is still required for complete determination of the amino acid sequence of protein. Gene technology has revolutionized sequencing of the protein structure; the primary structure of protein can be deduced from the nucleotide sequence of the gene encoding a protein, although determination of a partial amino acid sequence of protein is required for determination of the open reading frame of the nucleotide of the gene. The number of amino acid sequences deduced from nucleotide sequence surpasses that of sequences determined by Edman reaction. Homology search between sequences accumulated in a data base such as GenBank or EMBL has revealed a common segment of sequence (a consensus sequence) essential for enzyme catalysis or physiological function of protein such as nucleotide binding site,³²⁾ a metal ion binding site³³⁾ or a cell surface recognition site,³⁴⁾ and also classified proteins into a family, a subfamily, or a superfamily in molecular evolution.³⁵⁾

In comparison with the marked advancement in sequence determination of the amino terminal region of protein, no good methods have been developed for the determination of the amino acid sequence of the carboxyl terminal region, although several enzymatic and chemical procedures are available. Exoproteinases such as carboxypeptidases A, B, C and Y, have been employed for enzymatic determination. These enzymes catalyze sequential liberation of an amino acid from a carboxyl terminal. However, the reaction efficiency, *i.e.* velocity and specificity, is not always consistent for all amino acids consisting of a

carboxy terminal and affected by a neighboring amino acid, which can lead to mistakes in the assignment of the amino acid liberated. The chemical methods widely employed are hydrazinolysis,³⁶⁾ a specific radio-labeling method of a carboxy terminal amino acid,³⁷⁾ and a consecutive liberation method of amino acid from a carboxy terminal as an amino acid hydantoin.³⁸⁾ Based on these chemical methods an automatic carboxy terminal amino acid sequencer has been constructed, but it is not as efficient as the sequencer for the amino terminal.³⁹⁾ A carboxy terminal amino acid sequence can be determined with a gas phase sequencer, after separation of carboxy terminal peptide specifically.⁴⁰⁾

The final goal in the research of enzyme structure is the determination of its three-dimensional structure. X-ray diffraction crystallography is the only method available at present for this purpose, although recent remarkable advance in nuclear magnetic resonance spectrometry (NMR) elucidates three-dimensional structures of proteins of lower molecular weight in solution. There are about 500 proteins in a data base for which the three-dimensional structures have been resolved at the atomic level. Preparation of a single crystal with appropriate sizes in three dimensions (xyz axes) is prerequisite for X-ray diffraction crystallography. The single crystal can be prepared by a dialysis method; the enzyme solution is dialyzed against a high concentration of salt such as ammonium sulfate or sodium sulfate, or by a hanging (or standing) drop method, where the enzyme in a drop is gradually concentrated by removing water via vapor diffusion into a high concentration of polyethyleneglycol.⁴¹⁾ The mechanism of protein crystallization is not fully understood, but factors such as protein concentration, pH, temperature and ion species have been suggested to affect the formation of enzyme crystal. Recently crystallization of protein has been carried out under microgravity using space aircraft.^{42,43)}

Protein X-ray crystallography goes back to Bernal's pioneering work on pepsin in 1934, followed by Perutz and Kendrew, who determined three-dimensional structures of hemoglobin in 1953 and myoglobin in 1957. Lysozyme was the first enzyme whose structure was determined at the atomic level.⁴⁴⁾ In those days a large and strong crystal was required for the analysis due to weak X-ray generated from X-ray sources. Although the construction of a strong X-ray generator and automatic four circle diffractometer dramatically improved efficiency of data collection, it still required days to collect one set of diffraction data. During the past decade, however, extremely strong and white X-rays with a continuous wavelength in a wide range has become available through synchrotron radiation. With this X-ray and use of an imaging plate instead of X-ray film as the detector, data collection can be performed in one hour, enabling the analysis of small (about 0.1 mm) and fragile crystals.⁴⁵⁾ Improvements in X-ray crystallography coupled with developments in computer graphics are very useful in the research of protein engineering.

Protein has a secondary structure: α -helix, β -structure or random chain. The contents of these components in the protein structure can be calculated on the basis of circular dichroism spectrum in the region of far-ultraviolet wavelength (around 220 nm),⁴⁶⁾ or amino acid sequences.⁴⁷⁾ Although these methods do not always reflect a secondary structure of protein, they are applicable to research on the structure of proteins, especially homologous proteins whose three-dimensional structures have not been shown.

Enzyme Reaction Mechanisms

One of the most important problems of enzymology is to elucidate the mechanism of action of enzymes, *i.e.* discover how to account for their high efficiency in catalysis in

chemical terms. None of the mechanisms of enzyme reactions has been completely clarified. Emil Fischer first showed substrate specificity and stereospecificity of enzymes. He studied enzyme actions on various compounds in detail and proposed his famous “lock and key” theory for the enzyme-substrate relationship in 1894.⁴⁸⁾ There are, however, a number of points which are not readily reconciled with the theory of a rigid template. For instance, it is difficult to explain the lack of activity of a smaller substrate which has appropriate combining groups and a similar affinity for the enzyme as that of the specific substrate. Koshland in 1959⁴⁹⁾ proposed the theory of “induced-fit” of the active site for the substrate with a conformational change of enzyme, although this idea had been presented by Laidler to explain the effect of pressure on enzyme activity.⁵⁰⁾ Koshland’s theory provides a reasonable picture of enzyme–substrate interaction, but it should be noted that a conformational change is not always an essential part of the activity of every enzyme, and many enzymes can combine with both substrate and product which may be different in structure and are activated by the enzyme. Enzymes, therefore, may not fit both their substrates and products, but the intermediates derived from both reactants by activation, in other words, distortion of the chemical structure in the enzyme active site has occurred, resulting in “a transition state intermediate.” This idea was suggested by Pauling in 1946.⁵¹⁾ Pauling also suggested that this strain in the structure of substrate (or product) may lead to enormous acceleration of enzyme reactions.

An enzyme reaction mechanism can be studied from at least four different standpoints: kinetic aspects, thermodynamic aspects, chemical aspects and structural aspects. The fundamentals of enzyme kinetics was established by Michaelis and Menten in 1913,⁵²⁾ followed by the steady-state theory of Briggs and Haldane in 1925.⁵³⁾ In the kinetic mechanism, the order of binding of substrates and dissociation of products, and kinetic parameters such as k_{cat} and K_m are determined. Furthermore, the rate constants of the elementary steps of enzyme reaction can be determined based on a pre-steady state kinetics studied by rapid flow-technique or relaxation technique developed by Chance and Eigen, respectively.⁵⁴⁾ The isotope effect method, in particular the multiple isotope effect method is, useful for identifying the sequence of the elementary steps of a reaction.⁵⁵⁾ NMR measurement is also effective in determining the order of chemical events in enzyme reactions as reviewed in the following chapter. In light of the transition state theory and equilibrium thermodynamics developed by Eyring and Evans-Polanyi, and kinetic parameters, the thermodynamic mechanism is determined, which shows a rate-determining step involved in the formation of transition states, which are derived from substrates, cofactors or products during the course of a reaction. These chemical species are detected by electronic spectrometry, and confirmed with a model compound. Based on the results the reaction process is described by the nuclear and electronic changes that the substrate (and cofactor) undergoes during the course leading to the product.

Structural mechanism study is based on identification of essential amino acid residues and spatial orientation of their atoms in the active site, thereby reaching an understanding of how an enzyme expresses substrate and reaction specificities, and high efficiency in the lowering of activation energy of the reaction. To elucidate these problems, we have been using a technique of chemical modification, affinity labeling, suicide substrate inactivation and site-directed mutagenesis in combination with X-ray crystallography in some cases. The details are described in the following chapters. Chemical modification is employed to show that a functional group is important in catalysis and occurs at an active site or its

vicinity. If a group is catalytically essential, the enzyme whose group is modified becomes inactivated. Although it is generally not easy to obtain chemical reagents which specifically modify a group, several reagents have been developed to react only with a specific group to inactivate the enzyme. The samples are shown in the classification of proteinases: serine proteinase inactivated with diisopropyl fluorophosphate or its analogues, cysteine proteinase reacting with sulfhydryl reagents such as iodoacetate and aspartate proteinase inhibited with diazoacetyl norleucinomethylester and 1,2-methoxy-3-(*p*-nitrophenoxy)propane.⁵⁶⁾ Affinity labeling reagents, namely a chemical reagent which is a substrate analogue or competitive inhibitor containing a group reacting specifically with an important functional group of enzyme, have been developed. These chemical reagents react selectively with a specific amino acid residue in the enzyme active site; example, azaserine, a structural analogue of glutamine, inactivates phosphoribosyl glycine amidine synthase (EC6.3.5.3)⁵⁷⁾ by reacting with a thiol group of the enzyme, and tosyl-phenylalanylchloromethylketone inactivates chymotrypsin through modification of an active site histidine.⁵⁸⁾ A suicide substrate is a structural analogue of substrate which is inherently not reactive but binds the active site and is activated by enzyme catalysis to react with a functional group of the active site to inactivate the enzyme. Since the chemical process of the inactivation is analogous to the inherent catalytic process, the modified residue is usually the one that participates in the enzyme catalysis. Alteration of a certain amino acid residue by a site-directed mutagenesis gives more direct evidence regarding the role of the residue, although a conformational change of the enzyme caused by mutagenesis may affect the catalytic process.

The chemical methods and mutagenesis mentioned above are useful for identifying amino acid residues in the active site that participate in catalysis, and we can assign essential amino acids in combination with their kinetic and thermodynamic analysis. It is not until the precise arrangement of atoms is determined by X-ray crystallography that a detailed enzyme reaction mechanism can be drawn. Structures of enzymes complexed to catalytically relevant ligands such as substrate, inhibitor (substrate analogue) or transition state analogue must be determined by X-ray crystallography. Comparing these structures may reveal the dynamic aspects of the enzyme reaction. Abler, Petsko and Tsernoglou⁵⁹⁾ have shown the formation of an acyl intermediate of elastase by cryo-X-ray crystallography, analyzing at -55°C to slow down the hydrolytic process of the intermediate. The binding of maltoheptose, a substrate, to glycogen phosphorylase b has been shown by X-ray crystallography with flow cell.⁶⁰⁾ Although some steps in the dynamic process of the enzyme reaction could be trapped by these experiments, the most promising approach to elucidate the dynamic process of enzyme catalysis may be analysis using time-resolved X-ray crystallography.⁶¹⁾ In this experiment the crystal containing substrate with a caged group such as nitrophenyl ethyl group may be analyzed by the Laue method. The caged compound in the crystal can be converted simultaneously to a normal substrate by photolysis with laser beam. Recently, Schlichting *et al.*⁶²⁾ analyzed the crystal of Ha-Ras p21 with caged GTP by this technique, although the detailed process of GTP hydrolysis by this protein could not be drawn due to several factors that require improvement.

Novel Enzymes

Site-directed mutagenesis can provide a novel enzyme altered in substrate specificity or stability and even a chimera enzyme in which the structural domains of two different

enzymes are fused. This section, however, will focus on two categories of novel enzyme: ribozymes and abzymes.

In the last decade RNA molecules that have the intrinsic ability to break and form covalent bonds have been demonstrated. Such RNAs were termed “ribozymes” by Cech and his colleagues in 1982.⁶³⁾ The first ribozyme to be shown was the Group I intervening sequence in the rRNA precursor of *Tetrahymena thermophila*.⁶³⁾ This ribozyme catalyzes its own excision in the presence of guanosine and is modified during the process with the compound covalently coupled to the 5'-end of the ribozyme itself. The ribozyme-catalyzed reaction cannot be considered to be truly enzymatic; ribozymes act in a quasi-catalytic manner. However, there have been shown ribozymes acting in *trans* in a truly catalytic manner with a turnover of more than one and without being modified: the RNA component of bacterial RNase P, the only naturally occurring RNA catalyzing a true enzyme reaction.⁶⁴⁾ This ribozyme cleaves precursor tRNA to generate mature tRNA with multiple turnover, and no modification on the ribozyme occurred in the course of catalysis. Eight nucleotides-shortened Group I intervening sequence of *Tetrahymena* also catalyzes a number of *trans*-esterification reactions in a truly enzymatic manner.⁶⁵⁾ Subsequently there have been small catalytic RNAs acting in *trans*, which are derived from or synthesized on the basis of the nucleotide sequence of a self-cleaving RNA found in plant and animal pathogenic RNAs and others.⁶⁶⁾ Zaug, Been and Cech⁶⁷⁾ showed a specific endoribonuclease activity of L-19IVS ribozyme (about 400 nucleotides in length) derived from the *Tetrahymena* precursor RNA. Key sequence variations in the ribozyme led to cleavage at different sites. With a much smaller ribozyme (about 50 nucleotides) a site-specific cleavage of different RNAs has been performed.⁶⁸⁾ Research along this line will provide an opportunity to design a unique, small ribozyme (of less than 20 nucleotides) with a site-specific endoribonuclease activity, an event which will stimulate RNA engineering to target a variety of pathogenic RNAs *in vivo* as demonstrated for potato leafroll virus.⁶⁹⁾

In 1946 Pauling proposed that the ability of an enzyme to accelerate a chemical reaction stemmed from the complementarity of the structure of the enzyme active site to the activated substrate complex, a transition-state complex.⁵¹⁾ This proposal attracted research to seek the structural requirements for binding a transition-state analogue, since a transition-state complex has a negligible lifetime. These analogues acted a potent inhibitor with high affinity, and furnished a convenient starting point for creating catalytic antibodies (abzymes) by immunizing with these analogues. Earlier attempts, however, failed because they used polyclonal rather than monoclonal antibodies. It was not until 1986 that catalytic antibodies were discovered.^{70,71)} The number and diversity of abzyme-catalyzed reactions has been growing rapidly. There are at the moment about 50 reactions catalyzed antibodies.⁷²⁾ Among them the catalysis of sequence-specific peptide bond cleavage, *i.e.* the Gly-Phe bond, was shown with the abzyme produced by immunizing with a Co(III) triethylenetetramine(trien)-peptide hapten.⁷³⁾ Although these antibodies can catalyze a truly enzymatic reaction, the rate acceleration is in most cases less than 0.1% of those catalyzed by enzymes. The catalytic efficiency of abzyme can be improved by a hapten design that introduces a suitable catalytic residue such as a carboxyl group of glutamate/aspartate and an imidazole group of histidine or modulates the hydrophilic/hydrophobic character of the antibody-binding site.⁷⁴⁾ These improvements may also be performed by genetic or chemical modification of the antibody-binding site. In addition the introduction of a potential cofactor binding site such as a nucleotide binding domain

and metal binding site would greatly improve the catalytic activity of abzyme. Furthermore, cloning of the immunological response into *E. coli* or yeast may facilitate screening of the desired abzyme.

Enzyme Applications

Enzymes are characterized by their high specificity to substrates and reactions, and also by their enormous acceleration of reaction rate. Although the specificities are nearly absolute or very high in most enzyme reactions, there are enzymes acting on broad structural types such as whole classes of molecules sharing a common structural denominator, or catalyzing various types of reactions apparently different from their inherent reactions. Based on a thorough understanding of the molecular mechanism of enzyme reactions, the application of enzymes as catalysts for the production of useful compounds in the chemical and pharmaceutical industries has been achieved (Table I. 2).⁷⁵⁻⁷⁷⁾

Enzymes generally demonstrate unerring and complete stereospecificity in catalysis when they act on bonds linked to a chiral center of a substrate or form chiral products from achiral substrates. In some cases enzymes show unnecessarily high stereospecificity from the viewpoint of inherent catalysis. For instance most amino acid decarboxylases catalyze decarboxylation of L-amino acids with retention of the substrate configuration; a proton is added to the carbon from the same direction from which the carboxyl group is removed, and in NAD(P)-dependent dehydrogenase reactions, only one of the two prochiral C-4 hydrogens of the nicotinamide ring of the coenzyme is transferred to the substrates.^{78,79)}

TABLE I.2 Enzymatic Synthesis of Useful Compounds in Industry⁷⁶⁾

Compound	Enzyme
S-Substituted L-homocysteines	L-Methionine γ -lyase
S-Substituted L-cysteines	Cysteine desulfhydrase
	Tryptophan synthase
Tryptophan analogues	Tryptophase
Selenium amino acids	L-Methionine γ -lyase
L-Tyrosine analogues	β -Tyrosinase
D-Amino acids	Hydantoinase
Purine nucleosides	Purine nucleoside phosphorylase
Epoxides	Halohydrin epoxidase
Amides	Nitrile hydratase
Antibiotics	Penicillin acylase
Peptides	Proteinases

TABLE I.3 Enzymatic Synthesis of Stereospecifically Labeled Compounds⁷⁵⁾

Compound	Enzyme
Amino acids	Pyridoxal 5'-phosphate dependent enzymes
NAD(P)H	Dehydrogenases
Pyridoxamine 5'-phosphate	Transaminases
Pyruvate	Pyridoxal 5'-phosphate dependent enzymes
Amines	Amino acid decarboxylases

As a whole, such high stereospecificity is difficult to attain by chemical catalysis, although asymmetric synthesis has developed markedly in recent years.⁸⁰⁾ Stereospecificity has been applied to the production of optically active compounds such as amino acids, and also to the synthesis of stereospecifically isotope-labeled compounds, the synthesis of which is not easily achieved by a chemical method (Table I. 3).⁷⁵⁾

Enzymes play a central role in biotechnological processes in the production of useful materials other than those presented above, and enzyme sensors in which enzyme reaction is coupled with an electrochemical monitoring system are used clinical, process and environmental analyses. It is, however, out of the scope of this section to review these topics and other reviews should be consulted.⁸¹⁾

Perspectives

In the early days of enzyme study research focused on the isolation of enzymes. The studies were laborious and time-consuming. This situation has been greatly improved by innovations such as the HPLC system and microcapillary electrophoresis, which have enabled the purification of a small amount of enzyme effectively. We in general can isolate and purify most enzymes by a combination of gene technology with these methods without difficulty.

Increase in sensitivity and efficiency of analysis in structural studies of enzymes with a gas phase sequencer have made it possible to determine the primary structure in a shorter period of time with a small amount of enzyme at the picomole or even femtomole level. In addition, thanks to the DNA sequencing technique the number of enzymes (or proteins) whose amino acid sequences are registered in a data base file has expanded explosively. The introduction of mass spectrometry on the primary structure determination of protein has stimulated the search for a new methodology other than Edman chemistry.

The strategy for the determination of three-dimensional structures of enzymes has also been improved by the introduction of synchrotron radiation as an X-ray source and the use of an imaging plate in the detector system, accompanied by the development of computer technology. The recent marked developments in NMR technology will give impetus to the structural studies of enzymes in solution, although it is at present limited to enzymes of low molecular weight. Although static structures, *i.e.* spatial arrangement of atoms, of enzymes have been resolved in detail, it is arguable that not a single enzyme is fully understood regarding its dynamic components of catalysis. Time-resolved X-ray crystallography with photolytic caged compounds is very useful for studies along these lines. When ways to determine the substrate and reaction specificities including stereospecificity at the molecular level are found, the molecular mechanism of the enzyme reaction will be elucidated in more detail.

Based on these mechanisms a new frontier in the application of enzymes to biotechnology, including the development of synthetic enzymes (synzymes)^{82,83)} will be exploited. Research on structure-function relationships between ribozymes and abzymes will lead to the development of a number of sequence specific catalysts, which will control expression of a specific gene or its products, and eventually to application as pathogen controls in agriculture and to clinical use. Sequence specific abzymes may also facilitate research on the primary structure determination of protein.

Although enzymes have been studied extensively and our understanding of them has deepened tremendously, there remains much to learn in nature. Screening of enzymes

showing novel catalytic activity is still important for understanding the molecular mechanism of enzymes. One old but important aspect of enzyme studies is the mechanism of construction of enzyme steric structure based on amino acid sequences. Experiments with molecular chaperones will shed light on this mechanism.^{84,85)}

REFERENCES

1. A. Payen, J.F. Persoz, *Ann. Chim. Phys.*, **53**, 73 (1833).
2. H. Buchner, *Munch. Med. Wschr.*, **44**, 299 (1897).
3. J.B. Sumner, *J. Biol. Chem.*, **69**, 435 (1926).
4. *Methods Enzymol.*, **34** (1974).
5. C.V. Stead, *J. Chem. Tech. Biotechnol.*, **37**, 55 (1987).
6. T. Nagasawa, H. Nagasawa-Fujimori, P.C. Heinrich, *Eur. J. Biochem.*, **94**, 31 (1979).
7. T. Fujita, Y. Suzuki, J. Yamauchi, I. Takahara, K. Fujii, J. Yamashita, T. Horio, *J. Biochem.*, **87**, 89 (1980).
8. T. Ohshima, S. Nagata, K. Soda, *Arch. Microbiol.*, **141**, 407 (1985).
9. M. Oka, Y.-S. Yang, S. Nagata, N. Esaki, H. Tanaka, K. Soda, *Biotech. Appld. Biochem.*, **11**, 307 (1989).
10. W.R. Drewe, Jr., M.F. Dunn, *Biochemistry*, **24**, 3977 (1985).
11. T. Svedberg, J.B. Nicols, *J. Amer. Chem. Soc.*, **45**, 2910 (1923).
12. P. Andrews, *Biochem. J.*, **95**, 222 (1964).
13. K. Weber, M. Osborn, *J. Biol. Chem.*, **244**, 4406 (1969).
14. E. Penhoet, M. Kochomann, R.C. Valentine, W.J. Rutter, *Biochemistry*, **6**, 2941 (1967).
15. R.C. Valentine, B.M. Shapiro, E.R. Stadtman, *Biochemistry*, **7**, 2143 (1968).
16. D.G. Smith, W.H. Stein, S. Moore, *J. Biol. Chem.*, **238**, 227 (1963).
17. R.L. Heinrikson, S.C. Meredith, *Anal. Biochem.*, **136**, 65 (1984).
18. S. Muraoka, N. Fujii, I. Tamao, K. Harada, *Biochem. Biophys. Res. Commun.*, **146**, 1432 (1987).
19. R. Shapira, G.F. Austin, S.S. Mirra, *J. Neurochem.*, **50**, 69 (1988).
20. K. Nokihara, R. Yamamoto, T. Nishine, J. Gerhardt, *Anal. Sci.*, **7**, 285 (1991).
21. F. Sanger, E.O.P. Thompson, *Biochem. J.*, **53**, 366 (1953).
22. G.R. Stark, D.G. Smyth, *J. Biol. Chem.*, **238**, 214 (1963).
23. B.S. Hartley, V. Massey, *Biochim. Biophys. Acta*, **21**, 58 (1956).
24. J.Y. Chang, E.H. Creser, K.W. Bentley, *Biochem. J.*, **153**, 607 (1976).
25. P. Edman, *Acta Chem. Scand.*, **4**, 283 (1950).
26. P. Edman, G. Begg, *Eur. J. Biochem.*, **1**, 80 (1969).
27. R.A. Laurensen, *Eur. J. Biochem.*, **20**, 89 (1971).
28. M.W. Hunkapiller, L.E. Hood, *Biochemistry*, **17**, 2124 (1978).
29. B. Wittman-Liebold, H. Graffunder, H. Kohls, *Anal. Biochem.*, **75**, 621 (1976).
30. M.W. Hunkapiller, L.E. Hood, *Science*, **207**, 523 (1980).
31. R.M. Hewick, M.W. Hunkapiller, L.E. Hood, W.J. Dreyer, *J. Biol. Chem.*, **256**, 7990 (1981).
32. R.K. Wierenga, W.G.J. Hol, *Nature*, **302**, 842 (1983).
33. H. Takeya, K. Oda, T. Miyata, T. Omori-Satoh, S. Iwanaga, *J. Biol. Chem.*, **265**, 16068 (1990).
34. E. Rouslahti, M.D. Pierschbacher, *Science*, **238**, 491 (1987).
35. J. Kyte, R.F. Doolittle, *J. Mol. Biol.*, **157**, 105 (1982).
36. S. Akabori, K. Ohno, K. Narita, *Bull. Chem. Soc. Jap.*, **25**, 214 (1952).
37. H. Matsuo, Y. Fujimoto, T. Tatsuno, *Biochem. Biophys. Res. Commun.*, **22**, 69 (1966).
38. G.E. Tarr, in: *Solid-Phase Methods in Protein Sequence Analysis*, Pierce Chemical Co., Rockford, IL (1975), p. 139.
39. V.L. Boyd, M.B. Bozzini, G. Zon, R.L. Noble, R.J. Mattaliano, *Anal. Biochem.*, **206**, 344 (1992).
40. T. Kumazaki, K. Terasawa, S. Ishii, *J. Biochem.*, **102**, 1539 (1987).
41. A. McPherson, in: *Preparation and Analysis of Protein Crystals*, John Wiley & Sons, New York, (1982).
42. W. Littke, C. Johan, *Science*, **225**, 203 (1984).
43. L.J. DeLucas, F.L. Suddath, R. Snyder, R. Naumann, M.B. Broom, M. Pusey, V. Yost, B. Herren, D. Carter, B. Nelson, E.J. Meahan, A. McPherson, C.E. Bugg, *J. Cryst. Growth*, **76**, 681 (1986).
44. D.C. Phillips, *Proc. Natl. Acad. Sci. USA*, **57**, 484 (1967).
45. N. Watanabe, K. Sakabe, N. Sakabe, T. Higashi, K. Sasaki, S. Aibara, Y. Morita, K. Yonaha, S. Toyama, H. Fukutani, *J. Biochem.*, **105**, 1 (1989).
46. A.J. Adler, N.J. Greenfield, G.D. Fasman, *Methods Enzymol.*, **27**, 675 (1973).
47. J. Garnier, D.J. Osguthorpe, B. Robson, *J. Mol. Biol.*, **120**, 97 (1978).

48. E. Fischer, *Ber. Dtsch. Chem. Ges.*, **27**, 2985 (1894).
49. D.E. Koshland Jr., in: *The Enzyme*, Vol. 1, Academic Press, New York, (1959) p. 305.
50. K.J. Laidler, *Arch. Biochem.*, **30**, 226 (1951).
51. L. Pauling, *Chem. Eng. News*, **24**, 1375 (1946).
52. L. Michaelis, M.L. Menten, *Biochem. Z.*, **49**, 333 (1913).
53. G.E. Briggs, J.B.S. Haldane, *Biochem. J.*, **19**, 338 (1925).
54. G.G. Hammes, P.R. Schimmel, in: *The Enzymes* Vol. 2, Academic Press, New York, (1970) p. 67.
55. M.H. O'Leary, *Annu. Rev. Biochem.*, **58**, 377 (1989).
56. K.A. Walsh, P.E. Wilcox, *Methods Enzymol.*, **19**, 31 (1970).
57. B. Levenberg, I. Melnick, J.M. Buchanan, *J. Biol. Chem.*, **225**, 163 (1957).
58. G. Schoellman, E. Shaw, *Biochemistry*, **2**, 252 (1963).
59. T. Abler, G.A. Petsko, D. Tsernoglou, *Nature*, **263**, 297 (1976).
60. J. Hajdu, P.A. Machin, J.W. Campbell, T.J. Greenhough, I.J. Clifton, S. Zurek, S. Gover, L.N. Johnson, M. Elder, *Nature*, **329**, 178 (1987).
61. K. Moffat, D. Szebery, D. Bilderback, *Science*, **223**, 1423 (1984).
62. I. Schlichting, S.C. Almo, G. Rapp, K. Wilson, K. Petratos, A. Lentfer, A. Wittinghofer, W. Kabsch, E.F. Pai, G.A. Petsko, R.S. Goody, *Nature*, **345**, 309 (1990).
63. K. Kruger, P.J. Grabowski, A.J. Zaug, J. Sands, D.E. Gottschling, T.R. Cech, *Cell*, **31**, 147 (1982).
64. C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace, S. Altman, *Cell*, **35**, 849 (1983).
65. T.R. Cech, *Annu. Rev. Biochem.*, **59**, 543 (1990).
66. R.H. Symons, *Annu. Rev. Biochem.*, **61**, 641 (1992).
67. A.J. Zaug, M.D. Been, R.R. Cech, *Nature*, **324**, 429 (1986).
68. A. C. Foster, C. Davies, C.J. Hutchins, R.H. Symons, *Methods Enzymol.*, **181**, 583 (1990).
69. J.W. Lamb, R.T. Hay, *J. Gen. Virol.*, **714**, 2257 (1990).
70. S.J. Pollack, J.W. Jacobs, P.G. Schulz, *Science*, **234**, 1570 (1986).
71. A. Tramontano, K.D. Janda, R.A. Lerner, *Science*, **234**, 1566 (1986).
72. R.A. Lerner, S.J. Benkovic, P.G. Schulz, *Science*, **252**, 659 (1991).
73. B.L. Iverson, R.A. Lerner, *Science*, **243**, 1184 (1989).
74. S.J. Benkovic, *Annu. Rev. Biochem.*, **61**, 29 (1992).
75. K. Yonaha, K. Soda, *Adv. Biochem. Eng./Biotechnol.*, **33**, 95 (1986).
76. K. Soda, K. Yonaha, in: *Biotechnology*, Vol. 7a, VCH Verlagsgesellschaft, Weinheim (1987), p. 605.
77. I. Chibata, T. Tosa, T. Sato, in: *Biotechnology*, Vol. 7a, VCH Verlagsgesellschaft, Weinheim (1987), p. 653.
78. J. Jeffery, in: *New Comprehensive Biochemistry*, Vol. 3, Elsevier Biomedical Press, Amsterdam (1982) p. 113.
79. H.G. Floss, J.C. Vederas, in: *New Comprehensive Biochemistry*, Vol. 3, Elsevier Biomedical Press, Amsterdam (1982) p. 161.
80. J.D. Morrison, *Asymmetric Synthesis*, Vol. 5, Academic Press, New York, (1982).
81. I. Kurabe, in: *Biotechnology*, Vol. 7a, VCH Verlagsgesellschaft, Weinheim (1987) p. 685.
82. P. Dunnill, A. Wisemann, N. Blakebrough, *Enzymic and Non-enzymic Catalysis*, Ellis Harweed Publishers, London (1980).
83. V.W. Pike, in: *Biotechnology*, Vol. 7a, VCH Verlagsgesellschaft, Weinheim (1987) p. 465.
84. M.T. Fisher, *Biochemistry*, **31**, 3955 (1992).
85. R.M. Wynn, J.R. Davie, R.P. Cox, D.T. Chuang, *J. Biol. Chem.*, **267**, 12400 (1992).

1

Advanced Protein Sequencing Techniques

Hiroyuki Kagamiyama and Hiroshi Ueno

Department of Medical Chemistry, Osaka Medical College, Takatsuki, Osaka 569, Japan

1.1 Introduction

Analyses of protein sequences have provided a significant chemical basis in the understanding protein structures and their roles in biology. Several fields of research have benefited from protein primary structure information, including the understanding of structure and function relationship of protein, molecular evolution of homologous proteins, post-synthetic modification, tertiary structure determination and determination of reading frame. In enzymology, information on an enzyme's primary structure provides ideas as to how an enzyme undergoes catalytic processes, *i.e.* provides information on active site and functional domain structures. Furthermore, primary structural information is indispensable for determining the tertiary structure of proteins, which is necessary for the complete understanding of the enzyme mechanism.

Nearly forty years have passed since Sanger and his colleagues determined the first primary structure of a protein, pancreatic insulin, in 1955,¹⁾ and almost thirty years have passed since Moore and Stein determined the first primary structure of an enzyme, pancreatic ribonuclease A, in 1960.²⁾ Sanger spent nearly ten years determining the 51 amino acid-long insulin structure. This work was aided by his own development of a DNP method and by the paper chromatographic method, which had just become available at that time. About the same time, Edman developed a sequence method which utilized phenylisothiocyanate (PITC) and analyzed proteins sequentially from the N-terminal end (1950).³⁾ Moore and Stein spent over five years sequencing ribonuclease A whose sequence was twice as long as that of insulin. They had utilized all available techniques for determining the sequence, peptide fragmentation with trypsin and/or chymotrypsin, ion exchange chromatography to separate amino acids and peptide fragments, and Edman degradation. The strategy of current sequencing technique is still an extension of that which Moore and Stein established years ago. However, considerable improvements have been made for the sensitivity of detecting proteins, peptides, amino acids and PTH-amino acids (phenylthiohydantoin derivative of amino acid) and for efficiency in separation time. These improvements allow us to sequence proteins formerly considered difficult to analyze because of their minute quantity. Recent developments which enable us to deduce amino acid sequence from cDNA has revolutionized sequence strategy. The number of

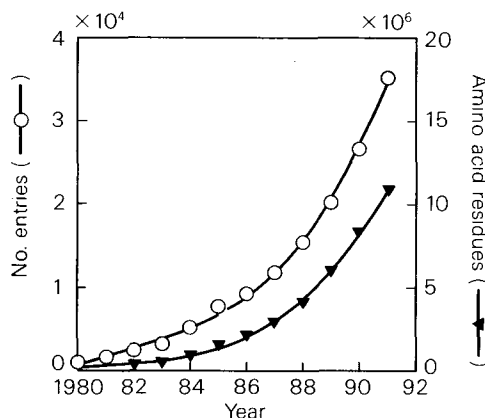


Fig. 1.1 Number of protein sequence data input into PRF/SEQDB. (Data from Dr. Ichiro Honda, Protein Research Foundation (Osaka, Japan)).

sequenced proteins has been increasing exponentially for the last few years as increasing amounts of cDNA sequence information and improved techniques become available (Fig. 1.1). Moreover, advances in computer sciences have made it possible to manage vast amounts of information. These days, the computer is an indispensable tool in studies of proteins. It allows us to design novel enzymes which possess new catalytic functions and to investigate the molecular evolution of enzymes based on the understanding of the structure and function relations of enzymes. In particular, when proteins from different species with the same functions are compared, computer analysis can provide us not only with information on protein functions, but also information on molecular evolution. It is often true that when such protein sequences are compared, the regions involved in the catalysis or function are preserved. Moreover, when proteins having different functions show homologous sequences, there is a strong possibility that these proteins share a common ancestor. Chemical modification techniques developed by Moore and Stein have been a powerful method for the understanding of the catalytic function of enzymes. The principle of chemical modification is to modify target amino acids at specific sites in a protein using a chemical reagent. The analysis of the modified protein can provide useful information on how the protein functions. In contrast, a recent advancement in gene manipulation technique allows us to induce a similar altering effect by amino acid substitution. This manipulation technique, called protein engineering, is becoming the main tool in the investigation of proteins. Protein engineering is considered to be the advanced stage in sequencing since it allows investigators to suggest the role of amino acid residues at the specific site found in the primary structure.

During the 1970s and 1980s, Kagamiyama *et al.* worked to determine the primary structure of aspartate aminotransferase (AspAT). Since then the studies in Kagamiyama's group have expanded to the determination of tertiary structure of AspAT and the application of protein engineering techniques to this enzyme. In this review, the development of methods and techniques of determining primary structure of protein will be described based on the studies on AspAT.

1.2 The Primary Structure of Aspartate Aminotransferase (AspAT)

A.E. Braunstein discovered AspAT in 1937.⁴⁻⁶⁾ AspAT is a PLP-dependent enzyme that catalyzes transamination between aspartate and α -ketoglutarate or glutamate and oxalacetate. In the early 1960s, AspAT from higher animal cells were separated into two forms by their characteristics in electrophoretic mobility.^{7,8)} The anionic form was found to be in cytosol and is called cAspAT, while the cationic form was found in mitochondria and called mAspAT. Similarities and differences between the two forms have been reported. The similarities are molecular weight of 90–100K; identical subunits to form a dimer; one active site per subunit; each subunit containing one coenzyme, pyridoxal 5'-phosphate (PLP); both forms showing similar optical spectra due to the coenzyme; and identical K_m values for substrates.⁹⁾ Some differences were also found in their isoelectric points, stability against temperature and denaturing agents,¹⁰⁾ and immunological reactivity.¹¹⁾ These studies have suggested that cAspAT and mAspAT are isozymes. cAspAT from pig heart was first purified to homogeneity in 1958. Since then, a number of studies particularly in the field of enzyme chemistry have been carried out on this enzyme as a typical example of PLP-dependent enzymes. Although those studies explained how PLP participates in the catalytic functions, they did not provide any information on the role of apoprotein, which enables the enzyme to be an efficient catalyst. The primary structures of both isozymes were thought to be different, but certain homology in their active site structures was expected because of the similarity in their catalytic actions. When analysis of the primary structure of AspAT was initiated at the beginning of 1970, it was assumed that if primary structures of both isozymes were revealed, the structural information necessary to assign active site residues and other structural requirements for the catalysis would be obtained.

The 1972 edition of *Atlas of Protein Sequence and Structure* (vol. 5) noted that the number of entries of primary structures of proteins was over 500, a majority of which were of molecular weight less than 20K.¹²⁾ Ovchinnikov *et al.* determined the amino acid sequence of pig heart cAspAT in 1973.¹³⁾ This enzyme has 412 amino acid residues and was the third largest protein sequenced at that time. Kagamiyama *et al.* determined the primary structure of pig heart mAspAT with 401 amino acid residues in 1977,¹⁴⁾ the first example in which primary structures of isozymes of mitochondrial and cytosolic origin were compared (Fig. 1.2). Despite of the differences in their gene locations on chromosomes, nearly 50% sequence homology between the two isozymes suggested that they share a common ancestor. Through these sequence studies, we anticipated obtaining some sequence information which would lead to the reaction mechanism of AspAT's or to a clue regarding cellular localization of AspAT's. Because sequence homology between the two AspAT's was too high, we could not assign meaningful sites as we had anticipated. After the completion of sequences on chicken c- and m-AspAT's^{15,16)} and *Escherichia coli* AspAT,¹⁷⁾ it became possible to compare enzyme structures from eukaryotes and prokaryotes (Fig. 1.2). Since the development of DNA recombinant technique, sequences of AspAT's from more than ten different species have been determined. On the animal isozymes, over 80% homology was found when sequences of one type of isozymes from different species were compared. In contrast, 50% homology was found between cytosolic and mitochondrial isozymes from the same species (Fig. 1.3). Christen estimated the time of isozyme

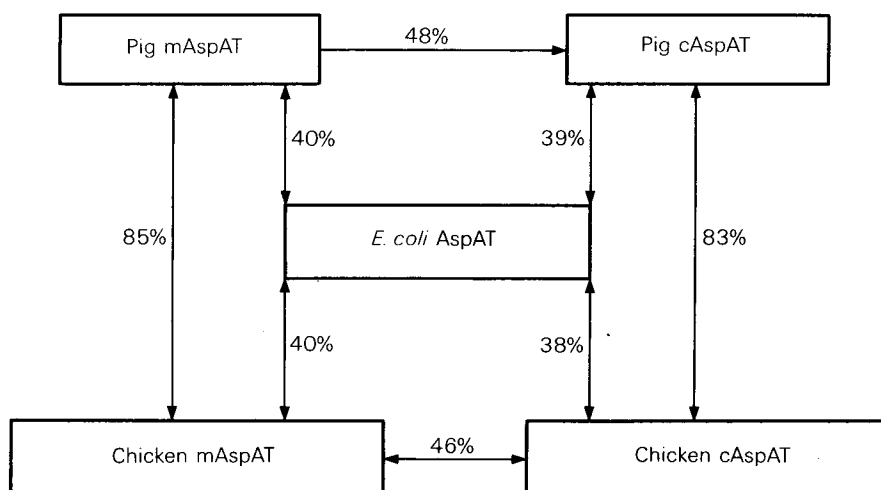


Fig. 1.3 Relationships in degrees of sequence identity.

appearance to be about one billion years ago based on the natural substitution rate for a single amino acid residue.¹⁸⁾ This happened about the same time eukaryotes appeared. *E. coli* AspAT with 396 amino acid residues exhibited about 40% sequence homology with either cAspAT or mAspAT (Fig. 1.3). This homology between *E. coli* and animal enzymes supported the idea of sharing a common ancestor gene. When the primary sequence of all AspAT's were compared, 25% of their amino acid residues were found to be in common. This implied that 75% of the amino acid residues had been replaced during the long evolutionary process. The conserved residues are expected to be involved in catalytic reactions and construction of tertiary structures common to transaminases. Such an idea is supported by X-ray crystallographic analyses on pig cytosolic,¹⁹⁾ chicken cytosolic,^{20,21)} chicken mitochondrial,²²⁾ and *E. coli*²³⁾ enzymes. Despite differences in primary structure, remarkable resemblance in three-dimensional structures of cytosolic, mitochondrial and *E. coli* AspAT's was observed (Chapter 5). Such analyses showed that all catalytically indispensable amino acid residues were conserved in the primary sequence of all species. In recent years, extensive *in vitro* mutagenesis studies on AspAT have been carried out (Chapter 5). This mutagenesis technique has become a powerful method for examining the exact role of amino acid residue, which is homologous among different AspAT's in catalysis.

Recently, the primary structure of *Thermophilic bacillus* AspAT was deduced from its cDNA sequence.²⁴⁾ However, there was no significant sequence homology found between thermophilic enzyme and other AspAT's. After careful comparison among primary structures of AspAT's, we found that catalytically essential amino acid residues located at the active site in AspAT were also assigned in the thermophilic enzyme (Fig. 1.2). Although thermophilic AspAT is presumed to have an entirely different tertiary structure from other AspAT's, catalytically essential amino acid residues may be conserved in its structure.

1.3 Progress in Primary Structure Analysis

Two approaches have been taken for determining the primary structure of protein. One is to determine the amino acid sequence directly from the isolated protein. The other is to deduce amino acid sequence from the cDNA base sequence. cDNA is complementary to mRNA in which the protein of interest is encoded. Recent advance in gene manipulation techniques have made the latter approach possible. cDNA analysis method is becoming popular because sequencing of proteins of large molecular weight and those existing in very minute amounts is not a difficult task once their cDNA's are obtained. It also provides pre-sequence information. Although cDNA analysis is becoming the popular choice for determining protein sequences, direct sequencing is still necessary particularly for the following situations: analysis of amino acid residues which receive post-synthetic modification²⁵⁾ (Table 1.1); determination of reading frame for the mature protein (including N- and C-terminal residue determinations); preparation of oligonucleotide probe necessary to isolate cDNA which encodes the protein of interest; and final confirmation of amino acid sequence deduced from cDNA. Both direct sequencing and cDNA analysis should complement each other for further refining protein sequencing techniques.

1.3.1 Direct Sequencing of Protein

The sequencing strategy follows the order: 1) preparation of pure protein, 2) selective cleavage of peptide bond 3) isolation of cleaved peptide fragments, 4) analysis of primary structure on amino acid sequencer, 5) determination of entire primary structure. Although this order has not changed since our primary structure determination work on AspAT carried out in early 1970, many improvements have been made, e.g. preparation of micro scale amount of sample and introduction of high sensitivity analytical instruments. These improvements have shortened analysis time and lowered the sample amount of required.

A. Protein Sample Preparation

Development of ion-exchange cellulose (1950), introduction of hydroxyapatite resin (1956), development of gel permeation chromatography on cross-linked dextrin (Sephadex) (1959), establishment of affinity chromatography (1968), and introduction of hydrophobic chromatography (1973) are some of the technical advances that have contributed to the improved purification of protein. Recent development of high-performance liquid chromatography (HPLC) and fast protein liquid chromatography (FPLC) have revolutionized separation techniques. In these systems, previously developed chromatographic conditions for ion-exchange, gel permeation, hydroxyapatite and hydrophobic chromatography are applicable without any major modification. These separation media and the newly developed reversed-phase chromatography on HPLC and FPLC have made it possible to purify a minute amount of protein in comparatively short separation time. For some proteins, an affinity chromatography or an immuno-affinity chromatography was reported to be highly effective. Gel electrophoresis provides convenient, fast and high resolution separation; thus it is often used for protein separations as well as for evaluation of protein purity. It became possible to perform sequence analysis at 50–100 pmol levels directly from SDS–PAGE gels. In this procedure, a specific protein

TABLE 1.1 Some Examples of Post-Translational Modification*

Modified amino acid residue	Modification	Modified amino acid residue	Modification
α -Amino terminal	Formyl	Histidine	$N^{\epsilon(n)}$ -Phosphorylhistidine
	Acetyl		N^{ϵ} -(Methyl)histidine
	Pyruvoyl		4-(iodo)histidine
	α -Ketobutyl		8 α -(histidyl)flavin
	Glucuronyl		ADP-ribosylation
	α -Amino acyl	Lysine	δ -Hydroxylysine
	Pyroglutamate		Methylation
	Murein		Phosphorylation
	Myristoyl		Acetylation
	Glycosylation		N^{ϵ} -(Diaminopimelyl)lysine
	Methylation		N^{ϵ} -Biotinyllysine
α -Carboxyl terminal	Unsubstituted and substituted amides		N^{ϵ} -Lipoyllysine
	ADP-ribosylation		N^{ϵ} -(Phosphopyridoxal) lysine aldimine
Arginine	Methylation	Phenylalanine	Glycosylation
	N^{ω} -ADP-ribosyl		O^{β} -glucosyl- β -hydroxyphenylalanine
	N^{ω} -Phosphoryl	Proline	Hydroxyproline
Asparagine	Glycosyl		O^4 -Arabinosyl (galactosyl) hydroxyproline
	N^{ϵ} -(β -Aspartyl)lysine	Serine	Phosphorylation
	Deamidation		O^{β} -(ADP-ribosyl-phospho)serine
Aspartate	β -Aspartylphosphate (O^4 -phosphoaspartate)		Glycosylated
	β -Methylaspartate		O^{β} -Palmitylserine
	Racemization		O^{β} -Methylserine
Cysteine	Disulfide formation	Threonine	O^{β} -Phosphothreonine
	Glycosylation		Glycosylation
	S-Cysteinyllheme	Tryptophan	Hydrotryptophan
	S-Cysteinyll phycocyanobilin		β -Ketotryptophan
	Cysteine-S-phosphate	Tyrosine	Chloro (iodo) tyrosine
	Cysteine-s-sulfonate		DOPA
	Seleno-cysteine glycerylation		(Dihydroxyphenylalanine)
Glutamate	γ -Carboxyglutamate		Tyrosine O^4 -sulfate
	O^{γ} -Methyl-glutamate		Tyrosine O^4 -phosphate
	O^{γ} -ADP-ribosylglutamate		O^{β} -glycosyl- β -hydroxytyrosine
Glutamine	Deamidation		
	N^5 -Methylglutamine		
	Pyroglutamate		

* For details see the review article written by F. Wold.²⁵⁾

band is punched out from the gel, transferred onto a membrane disk made of poly(vinylidene difluoride) (PVDF) or polybrene treated glass fiber filter, and placed on a gas phase protein sequencer.²⁶⁾

B. Fragmentation of Peptide

During the 1970s when manual Edman degradation was employed, fragmentation of

peptides was carried out by using a site-specific cleavage method, *i.e.*, endopeptidase or chemical cleavages. For endopeptidase cleavage, trypsin, chymotrypsin and thermolysin were often chosen. As chemical cleavage, CNBr which was efficient and site specific was used quite frequently to produce fewer peptide fragments of medium size. But it sometimes gave insoluble products which made purification difficult. Manual Edman degradation was the method of choice at the time of primary sequence determination of mAspAT in the 1970s when the automated sequencer was not yet in general use. Although there had been constant improvement, *i.e.*, purity of reagents, optimization of reaction conditions, thin layer chromatography (TLC) to separate PTH-amino acids, and detection method(s) in which a fluorescence light source was employed to visualize PTH-amino acids as dark spots on bright background, the manual Edman degradation had the limitation of not being able to sequence more than 10–15 residues. Because of this, small peptides had to be prepared for manual sequencing.

Pre-sequence preparation has changed since automated sequencers, especially gas-phase sequencer, became available, and larger sized peptide fragments routinely sequenced. A small number of medium size peptide fragments are prepared for the automated sequencer. Many highly specific commercially available endopeptidases are used for the fragmentation,²⁷⁾ *e.g.* lysylendopeptidase (also called *Achromobacter* protease I) from *Achromobacter lyticus* M497-1, Lys-C endoproteinase from *Lysobacter enzymogenes* and *Staphylococcus aureus* V8 protease. The use of metalloendopeptidase, which cleaves X–Lys bond and arginylendopeptidase [Arg-C endoproteinase from mouse submaxillary glands], or clostripain, which cleaves the Arg–X bond, was also reported. Lysylendopeptidase is popular, as it can cleave at all lysines except the lysine at the N-terminal and works in the presence of 4 M urea or 0.1% SDS. V8 protease can function in the presence of 0.2% SDS and cleaves the Glu–X bond when an ammonia-containing buffer is used, and it also cleaves the Asp–X bond when a phosphate buffer is used. In addition to CNBr, BNPS skatole is known to cleave at the Trp–X bond, hydroxyl amine is known to cleave the Asn–Gly bond, and acidic condition is also known to cleave the Asp–Pro bond. These chemical cleavages, however, have a tendency to give lower cleavage efficiency and inconsistency toward proteins. When cleavage efficiency is not consistent, one expects complicated products, making analysis difficult. For a few specific purposes, however, a chemical cleavage method is still useful.

C. Separation of Peptide Fragments

Peptide separation and purification are the two most important steps in protein sequence analysis. Since obtaining larger numbers of pure peptide fragments leads to successful sequencing results, high efficiency is a critical requirement for peptide isolation. Chromatographies on paper, ion-exchange column, and gel filtration, and high-voltage paper electrophoresis were popular methods for isolating small peptide fragments. Taking the process of mAspAT sequence determination as an example mAspAT was digested first with trypsin and separated on Sephadex G-50 into four major fractions (Fig. 1.4A). Fractions I through III, containing higher molecular weight materials, were applied to DEAE-cellulose chromatography (Fig. 1.4B). Fraction IV, containing small molecular weight materials, was applied to Dowex-50 ion-exchange chromatography. Fractions obtained from these chromatographies were further purified on either paper chromatography or high-voltage paper electrophoresis. It took overnight to complete a

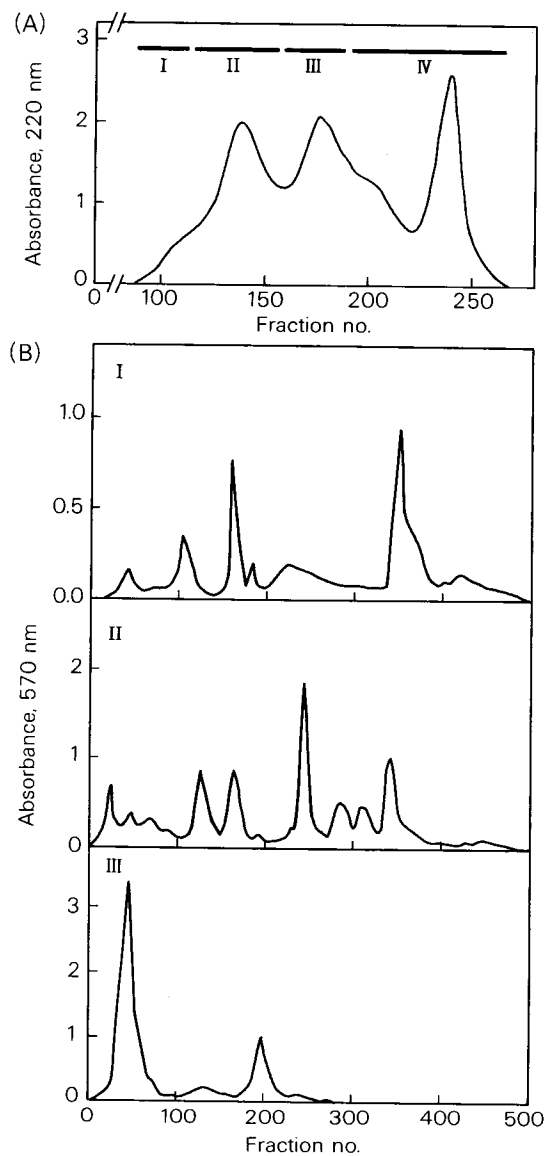


Fig. 1.4 Chromatographic separation of tryptic peptides from m-AspAT. A. Sephadex G-50 column chromatography. Fractions I-III were applied on DEAE-cellulose chromatography (B). B. DE-52 column chromatographies of fractions, I, II and III. Data are from Kagamiyama *et al.*²⁸⁾

paper chromatography and an additional overnight to elute the peptides from the paper. After these processes were repeated several times, 39 peptide fragments were isolated as pure forms and their sequences were determined by manual Edman degradation method,²⁸⁾ although theoretically 49 peptide fragments should have been isolated. Total amino acid residues recovered from the 39 peptide fragments were 80% that of the starting mAspAT. Then CNBr cleavage was carried out on mAspAT, which has 11 methionine residues to produce 12 CNBr fragments, in theory. Five major fractions (I–V in order of elution) were obtained after the Sephadex G75 column chromatography. Fraction I was found to be pure and to contain 114 amino acids, the largest molecular weight peptide. Other fractions containing impurities were applied for further purification steps. Fractions IV and V, containing small molecular weight peptides, were applied to paper chromatography and high-voltage paper electrophoresis. Fractions II and III were insoluble so were solubilized in urea then applied to phosphocellulose ion-exchange chromatography in the presence of 8 M urea. Twelve fragments were isolated in pure form as expected from the number of methionine residues.²⁹⁾ At that time, further fragmentation of these CNBr peptides by endopeptidases was required to establish their complete sequences.

HPLC brought far superior performance to separate peptide fragments and it became indispensable for sequencing. It has taken over the place of paper chromatography and high-voltage paper electrophoresis. Reversed phase ODS columns with superfine resins having high pressure tolerance have shown excellent performance in separations. The primary structure of *E. coli* AspAT was determined with a reversed phase HPLC employed in 1987.³⁰⁾ The use of HPLC made separation time shorter with much less sample. Meanwhile, a problem of sample loss associated with micro-scale sample preparation had become apparent, including nonspecific adsorption of peptides to glassware, formation of insoluble materials after the drying, and loss during the freeze-drying process. Nevertheless, HPLC has improved steadily, particularly with regard to pump capability to deliver precise flow, detectors with high sensitivity and stability, and microcomputers which allow sophisticated manipulation of HPLC results.

Capillary electrophoresis (CE) was developed recently and is a promising method for peptide separation. CE separates peptides by applying a high voltage (a few hundred volts per cm) with extremely high resolving power. Separation is carried out in a capillary tube with an inner diameter of 20 or 30 microns, eliminating Joules heat. Since CE provides superior sensitivity and resolution, only a minute amount of sample is required. Any substance with charge can be separated on CE regardless of molecular weight. Possible applications of CE in protein chemistry are in sample preparation, peptide mapping, separation of PTH-amino acids,³¹⁾ and primary structure determination in conjunction with mass spectrometry (see Section 1.3.3., p 32).

D. Amino Acid Composition

In 1944, about the time Sanger determined the primary structure of insulin, two-dimensional paper chromatography became available for analyzing amino acids of protein samples.³²⁾ This method allowed Sanger to analyze 20 amino acids in a single run with considerably less sample and time compared to the previous methods. The development of an automated amino acid analyzer in 1958 by Spackman, Stein and Moore had made further progress.³³⁾ This first amino acid analyzer performed an analysis with 1 μ mol of sample in 20 hours. Due to the continuous improvements made on amino acid analyzers,

routine analysis could be performed within 2 hours with 5–10 nmol of sample by the end of 1970. Now an analysis takes less than 60 min. Traditional ninhydrin post-column derivatization method developed by Moore–Stein is commonly used as the detection method with improved sensitivity of 50–100 pmol. Fluorescent post-column derivatization method with fluoescamine³⁴⁾ or *O*-phthalaldehyde (OPA)^{35–39)} gives one or two orders of magnitude higher sensitivity than the ninhydrin method. In the 1980s, derivatization of amino acids with reagents having absorption maxima in UV or VIS regions or having fluorescence were developed, e.g. phenylisothiocyanate (PITC),⁴⁰⁾ 4-dimethylamino-5-azobenzene-4'-sulfonyl chloride (Dabsyl-Cl),^{41,42)} Dansyl chloride (DNS-Cl)⁴³⁾ and 9-fluorenylmethoxycarbonyl chloride (FMOC-Cl).⁴⁴⁾ Derivatized amino acids were conveniently separated on reversed phase HPLC and gave superior sensitivity. Thus, the amino acid analyzer with pre-column derivatization method gave 10⁶–10⁷ sensitivity over the Moore–Stein system and made several picomol analysis possible in 1/20 the running time over the post-column derivatization method.

For amino acid composition analysis, with the exception of tryptophan, the sample must be hydrolyzed in constant boiling HCl (5.7 N) at 110°C for 24 hours. In general, HCl is removed by vacuum evaporation after the hydrolysis step and prior to the amino acid analysis. For tryptophan analysis, 4 *N*-methanesulfonic acid must be used instead of HCl,⁴⁵⁾ since acid hydrolysis in HCl destroys tryptophan residues. As methanesulfonic acid does not evaporate, it must be neutralized with a base prior to the amino acid analysis. Because methanesulfonic acid hydrolysis method requires this extra step, which causes high salt content, this method is not preferred for the amino acid analysis of a minute amount of sample or for pre-column derivatization. There are other hydrolysis conditions, *i.e.*, at higher temperature for shorter hydrolysis time.^{46–48)} Amino acid analysis at picomol levels requires extra caution since it can easily pick up amino acids from contaminated glassware or reagents. Such contamination is also present in the environment or can occur during sample handling, and lead to erroneous results. Even a fingerprint has enough amino acids to appear in picomol level analysis. In order to avoid such contamination, the vapor-phase hydrolysis method of HCl was developed.⁴⁹⁾

E. Amino-terminal Analysis

The amino-terminal sequence information plays a critical role 1) when the entire primary structure is ready to be constructed from many sequences assigned for small peptide fragments; 2) when protein processing is involved in which amino-terminal region of translated polypeptide receives site specific cleavage reaction to produce mature, functional protein; 3) when post-synthetic modification is involved in which amino-terminal end group receives some sort of modification, such as acetylation.

Sanger's reagent, 1-fluoro-2,4-dinitrobenzene (FDNB), which was used in the earlier days for the quantitation of primary amino groups by colorimetric determination, can also be used in the identification of amino-terminal residue, but not for sequencing. At the present time, N-terminal analysis is performed on a protein sequencer.

A free amino-terminal end is a requirement for the Edman degradation method; thus, those having amino-terminal ends blocked or masked cannot be analyzed in a sequencer which utilizes the Edman method. There has been significant progress in the analysis of blocked amino-terminal end and it is becoming reasonably easy to perform such analysis. The blocking group most often found is the acetyl group; the other blocking groups are

formyl, myristoyl, pyroglutamyl, glycosyl, pyruvoyl, α -ketobutyl, monomethyl, and dimethyl (Table 1.1). Analysis of the amino-terminal blocking group is performed essentially as follows: Peptide sample is digested with protease, which possesses relatively wide specificity, and the digested materials are applied on cation-exchange chromatography (preferably SP-Sephadex C-25) under strong acidic conditions. A small peptide with a modified amino-terminal is eluted off together with the unadsorbed fractions. If this peptide contains basic amino acid residues, it is adsorbed on the resin. The unadsorbed fractions may contain peptides with acidic properties. They are further chromatographed on reversed phase HPLC and each of the peptides isolated thus far has been analyzed for their amino-terminal. Aminopeptidase M digestion is one of the methods used to find blocked amino-terminal-containing peptide in peptide mixture. Aminopeptidase M digests amino acid sequentially from amino-terminal, but it does not work on a peptide whose amino-terminal is blocked. Thus, it is possible to determine the presence of a blocked amino acid by comparing HPLC patterns before and after the treatment with aminopeptidase M. Identification of the blocking group as well as of blocked amino-terminal residue is important once the presence of a blocked amino-terminal is suspected. Isolation of a peptide fragment containing an amino-terminal blocking group is usually the first step. Then, the amino acid composition of such peptid fragment is determined. The use of an FAB mass spectrometer should be considered to identify the blocking group. Advanced mass spectrometers, such as the tandem mass spectrometer (MS-MS) in which two mass spectrometers are connected in a row, have been able to determine both modified groups and amino acid sequence at the same time. If an acetyl group is expected to be the blocking group, an acyl amino acid-releasing enzyme should be used^{50,51)} as the released acyl-amino acid is applied to HPLC analysis or mass spectrometry for the assignment of the terminal-amino acid residue. The remaining polypeptide with a newly created free amino-terminal end can be sequenced on a protein sequencer. Since the acyl-amino acid-releasing enzyme does not work well on high molecular weight protein, the target protein should be digested with protease prior to the treatment with the acyl-amino acid releasing enzyme. Jones and Manning found that partial denaturation of high molecular weight protein with chemical modification, such as aminoethylation, could unfold the amino-terminal region so that this enzyme could work (Jones and Manning, personal communication). There is another method for identifying amino terminal residues without isolating acylated peptide. Free amino terminals resulted from the protease treatment of the blocked peptide are reacted with PITC, then further oxidized with hydrogen peroxide to produce phenylcarbamoyl moiety. By treating the oxidized product with the acyl-amino acid releasing enzyme, only the acetylated peptide can release free amino terminal end which is subjected to sequence analysis.⁵²⁾ Pyroglutamyl peptidase is used to remove the pyroglutamyl moiety.⁵³⁾ Pyroglutamyl peptidase does not work well on large size proteins, however. A formyl group could be removed from formylpeptide by the hydrolysis with 1 N HCl in methanol for 1.5 hours in ambient condition.⁵⁴⁾

F. C-terminal Analysis

Traditional methods of C-terminal end determination⁵⁵⁾ were hydrazine degradation, carboxypeptidase digestion, and tritium labeling. Unfortunately, no chemical methods analyzing amino acid sequence sequentially from the C-terminal end is available with reliability similar to that of Edman degradation. Carboxypeptidase digests a protein and

releases free amino acids sequentially from its C-terminal end. Amino acid analysis performed after the short exposure of protein with carboxypeptidase and removal of protein moiety by centricon separation can provide useful information to identify amino acids that are expected at the C-terminal region. The procedure is simple and useful for confirming the C-terminal region of an amino acid sequence obtained from cDNA. Among known carboxypeptidases, carboxypeptidase Y is recommended for the C-terminal analysis because of its low substrate specificity and its activity in the presence of 1% SDS.⁵⁶⁾ Recently, a method was developed to isolate a peptide fragment which originated from the C-terminal end by protease digestion.⁵⁷⁾ An affinity column packed with anhydrotrypsin-immobilized agarose beads is used for lysylendopeptidase- or trypsin-treated protein. Anhydrotrypsin is a genetically engineered trypsin mutant whose active site serine is substituted by alanine. Anhydrotrypsin binds to substrate peptide without catalytic activity. Since lysylendopeptidase or trypsin gives peptide fragments whose C-terminal ends are either lysine or arginine, these peptide fragments are retained in the affinity column. Only peptide fragments originating in the C-terminal end of the protein can pass through the column without retention, and are thus separated from the other fragments. When arginine or lysine is expected to be the C-terminal amino acid, a protease other than trypsin is used for the cleavage and the fraction bound to the affinity column is further separated on HPLC.

G. Edman Degradation

Edman degradation method has been the most extensively used technique for protein sequencing. It cleaves off amino acid as PTH-derivative from the free N-terminal of peptide/protein. By repeating the same cycle, the sequence of the peptide/protein can be established. In the 1960s and 1970s, Edman degradation was carried out manually and the PTH-amino acids routinely separated on paper chromatography or silica gel TLC. Iodine-azide reaction⁵⁸⁾ was employed for detecting PTH-amino acids on TLC plate, but the sensitivity was relatively poor (a few nmol). When quantitation was required, a subtractive Edman degradation was performed. In this method, after each cycle, a small aliquot was withdrawn and applied to amino acid analysis after HCl hydrolysis of the residual peptide. Any amino acid which disappeared since the previous cycle is the one located at the N-terminal end. A 10-residue-peptide was about the maximum size that can be sequenced by this method.

In the 1970s, many improvements were made on the Edman degradation techniques, *i.e.*, reaction conditions, reagent purity, and development of TLC plates with fluorescent dye. The fluorescent dye enhanced sensitivity for TLC analysis by detecting PTH-amino acids under UV lamp, enabling the determination of more than 10 residue sequences with only a 0.05–0.1 μmol sample. Better sensitivity was obtained by the Dansyl-Edman method,^{59,60)} in which DNS-Cl was applied to small aliquots at the end of each Edman degradation cycle. A peptide with a newly created N-terminal which reacted with DNS-Cl was acid hydrolyzed. The DNS-amino acid was then separated on 2-D TLC and detected under fluorescent lamp. Although detectability of the Dansyl-Edman method provided a picomol level analysis which was 10–100 times better than PHT-amino acid analysis on TLC, the method failed to qualitate asparagine, glutamine, and tryptophan, which were destroyed during acid hydrolysis. Another disadvantage of this method was sample loss, since small amounts had to be withdrawn for each Edman cycle.

At the time the structure of mAspAT was determined, PHT-amino acids were usually separated on fluorescent TLC plates. In some cases, newly introduced gas chromatography was used, which method was promising for PTH amino acid analysis.⁶¹⁾ But, because of the limited ability of the instrument to separate amino acids, gas chromatography has never become a popular choice. In the late 1970s, protein sequencing became popular when an automated sequencer developed by Edman and Begg in 1967⁶²⁾ and an improved HPLC method for PTH-amino acid analysis became available. Edman and Begg's automated device performed Edman degradation by utilizing a method called spinning cup or liquid-phase in which a high-speed spinning cup spread sample onto its inner wall as a thin film and, while maintaining the thin film, reactions and extractions were performed. The device yielded higher recovery after each Edman degradation cycle.⁶²⁾ Sixty amino acid residues from the amino-terminal end of myoglobin were determined with only 0.3 μ mol of sample.⁶²⁾ The early spinning cup model required a manual process for converting thiazoline-amino acids to PTH-amino acids. This manual conversion was eliminated when an auto converter device was later developed. The spinning cup sequencer was used for partial sequence determination of mAspAT. The sequences of the native enzyme (0.2 μ mol) and the longest CNBr peptide fragment (residue numbers 32–145) were determined effectively up to 48 and 32 residues, respectively.²⁹⁾ The most reliable, but time-consuming, way to determine Edman degradation products over 20 steps was to analyze them on an amino acid analyzer after thiazoline-amino acid derivatives were converted back to free amino acids by acid hydrolysis.^{63,64)} The spinning cup instrument tended to lose sample during the extraction cycles in which organic solvent flushed into the cup. Therefore, hydrophobic peptides were not applicable. This problem of losing sample was eliminated, although not completely, when polybrene, a quaternary amine having excellent binding ability with negatively charged compounds, was introduced. An additional problem was the low extraction recovery of thiazolinone derivatives of strongly *hydrophilic* amino acids. In 1971, Laursen introduced a solid phase sequencer in order to overcome the problems associated with the spinning cup system.⁶⁵⁾ The significance of this method was that a sample was covalently coupled to beads which were made of either glass or polystyrene resins prior to the Edman degradation. The major advantage of the solid phase sequencer was that sample loss was minimized during the extraction steps, even when hydrophobic samples were analyzed, making it was possible to complete sequence analysis on small peptide fragments. The hydrophilic products of Edman degradation, which usually gave poor recovery in the spinning cup method, were extracted with good efficiency and the sequence of phosphate-containing peptide was successfully determined.⁶⁶⁾ Several years after the completion of mAspAT primary structure determination, sequence work on *E. coli* AspAT was performed, in which the complete sequence of 6 of 7 CNBr fragments was determined solely on a solid phase sequencer.³⁰⁾ Each CNBr fragment contained from 7 to 41 residues and 20–200 nmol amounts were successfully analyzed with considerably high efficiency without further fragmentation of the peptide.³⁰⁾ Unfortunately the solid phase sequencer did not become a popular method for the following reasons: 1) Only three functional groups in a peptide fragment, amino group, carboxyl group and C-terminal homoserine lactone residue, could form a linkage with resin. There was no general approach to make a chemical bond during the coupling step between only the C-terminal amino acid residue and the solid support. 2) The coupling process was not quantitative, thus was not efficient. This inefficiency hampered the solid phase sequencer from being

able to analyze sub-nanomol level sequencing. 3) Those residues directly attached to the solid support via amino groups, such as amino-terminal amino acid and lysine, or via carboxyl groups, such as aspartate and glutamate residues, could not be derivatized as PTH-amino acid. Although the use of the solid phase concept had declined after the introduction of the gas phase sequencer in the 1980s, an updated version of the solid phase sequencer was recently introduced.⁶⁷⁾ This instrument immobilizes sample onto glass fiber filter or PVDF membrane to eliminate the above-mentioned disadvantage of the solid phase instrument; thus it has been reported able to sequence 50–100 residues at 30 min per cycle speed.

In the early 1980s, Hood and his associates developed a gas phase sequencer which primarily focused on analysis of very small samples.⁶⁸⁾ This instrument was designed to hold sample on a glass fiber filter disk coated with polybrene and carried out both pH control by trimethylamine and cleavage reaction with trifluoroacetic acid (TFA) by delivering these two reagents as the gaseous phase. The most recent model, however, uses liquid TFA for the cleavage reaction to provide better yield (liquid-pulse gas phase sequencer). The gas phase instrument has superior advantage of minimizing extraction loss due to the small volume of its reaction vessel and minimum consumption of solvents and sample. There is also a built-in autoconverter which converts thiazoline to PTH, and an on-line HPLC which analyzed PTH-amino acids. The gas phase sequencer had a strong impact in the area of protein sequencing, particularly after its success in sequencing the N-terminal portion of interferon.⁶⁹⁾ Due to its availability in extremely small quantities, the sequencing attempts for interferon had been unsuccessful by conventional methods.

Today's predominance of the gas phase sequencer for sequence analysis is partly due to the advancement of PTH-amino acid analysis on HPLC, in which the analysis takes only 10 min to complete using reversed-phase column.^{70–72)} Thus, a gas phase sequencer with an on-line PTH-analyzer can perform a single Edman degradation cycle within one hour. With the use of a microbore column and a data station which handles various data manipulations, a few pmol of PTH-amino acid can be quantitatively analyzed. Although the minimum detection level of amino acid derivatives has improved significantly in the last 10 years (10^3 – 10^4 -fold enhancement), further improvement is expected with the use of a fluorescent compound, *i.e.*, fluorescein isothiocyanate (FITC), and the laser detection technique.

H. Concluding Remarks

In general, the primary sequence of the protein is completed by aligning amino acid sequences of peptide fragments. Those peptide fragments are the results of different protease treatments or chemical cleavages of the native protein and should be able to provide overlap information. Remarkable improvements in peptide sequencing methods over the past 40 years have been described above. It is of interest to review the historical accomplishments in which the time scale for sequencing a single residue is compared. Two and a half months per residue was recorded for sequencing 51-residue insulin in 1955. Two weeks per residue was recorded for sequencing 124-residue RNaseA in 1962. Six and 2 days per residue were recorded for 401-residue mAspAT in 1977 and 389-residue *E. coli* AspAT in 1983, respectively. More recent data for 495-residue rat liver alanine aminotransferase showed that it took only 9 hr per residue.⁷³⁾ The entire primary structure

of this enzyme was determined within 6 months and a total amount of the enzyme used for the sequencing was merely 1–2 mg (K. Titani, personal communication), two orders of magnitude less than the amount required for *E. coli* AspAT.

1.3.2 A Sequencing Approach Based on cDNA

Recombinant DNA technology enabled determination of the primary structure of protein from cDNA information. In the 1980s, this method became the most popular method to determine protein sequence. There are several advantages in cloning cDNA encoding for a protein. (1) Its primary sequence can be easily deduced. (2) The structure of the precursor protein, whose information may not be provided from the sequence of the mature protein, can be deduced. (3) Proteins isolated in minute amounts or found in nature in limited quantities can be mass produced when cDNA encoding for a protein is inserted into microorganisms, particularly *E. coli*, or suitable tissue culture systems. (4) By performing a base substitution, a mutation site can be introduced into the protein sequence so the role of amino acid residue which is encoded at the mutation site can be studied in connection with protein structure and functional relationship.

A general method for cDNA cloning is as follows. (1) An amino acid sequence of any part of the protein is determined. (2) Oligonucleotide is prepared based on the amino acid sequence determined in step (1). DNA synthesizer is commonly used for this step. (3) cDNA, which hybridizes with an oligonucleotide prepared in step (2) as a probe, is obtained from the cDNA library. In the case of a prokaryote, restriction enzyme fragments of genomic DNA can be used in place of the cDNA library. It is practical to use at least two different probes which recognize separate parts of the structure. This eliminates the possibility of obtaining false positive results. When 15- to 20-base-long probes are designed, a mixed nucleotide which contains every codon degeneracy should be used. When 50-base-long or longer probes are designed, specific amino acid codons based on the preferred codon usage for the species from which the cDNA is derived are chosen to achieve better hybridization efficiency. One must pay attention to the particular amino acid at the region where the probe is planned to be designed. Arginine, serine or leucine are usually avoided, since there are multiple DNA codons designated for these amino acids. Methionine, which is assigned to a single DNA codon, is a preferred amino acid to be included into the probe design. Therefore, it is highly recommended to sequence as many peptide fragments as possible for cloning work. Besides use in probe design, this sequence information can be a useful proof for the cDNA once it is cloned. Amino acid sequence deduced from the cloned cDNA can be confirmed with the sequence data obtained from peptide fragments.

The primary structure of AspAT from *Thermophilic bacillus* was determined from cDNA sequence.²⁴⁾ Sequence information of an N-terminal portion of the native enzyme and 19 tryptic peptide fragments, recovered from HPLC, was obtained from gas phase sequencer analyses. Based on such sequence information, four oligonucleotide probes were prepared. cDNA encoding AspAT was cloned by screening restriction enzyme fragments from genomic cDNA of *Thermophilic bacillus* species YM-2. Amino acid sequence of *Thermophilic bacillus* AspAT deduced from cDNA was confirmed by the sequences made available by gas phase sequencer analysis.

There are structures in protein which cannot be identified by the information solely based on DNA sequence. Such structures are (1) post-synthetically modified amino acid

				-37	-36	-35	-34	-33	-32	-31
				CCC	ACT	GCT	CTC	TTG	CCG	CCC
-30	-29	-28	-27	-26	-25	-24	-23	-22	-21	
ACC	ATG	GCC	CTG	CTG	CAC	TCC	GGC	CGC	GTC	
	<i>Met</i>	<i>Ala</i>	<i>Leu</i>	<i>Leu</i>	<i>His</i>	<i>Ser</i>	<i>Gly</i>	<i>Arg</i>	<i>Val</i>	
-20	-19	-18	-17	-16	-15	-14	-13	-12	-11	
CTG	TCC	GGA	GTC	GCC	TCC	GCC	TTC	CAC	CCA	
<i>Leu</i>	<i>Ser</i>	<i>Gly</i>	<i>Val</i>	<i>Ala</i>	<i>Ser</i>	<i>Ala</i>	<i>Phe</i>	<i>His</i>	<i>Pro</i>	
-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	
GGC	CTC	GCT	GCT	GCA	GCA	TCT	GCC	AGA	GCC	
<i>Gly</i>	<i>Leu</i>	<i>Ala</i>	<i>Ala</i>	<i>Ala</i>	<i>Ala</i>	<i>Ser</i>	<i>Ala</i>	<i>Arg</i>	<i>Ala</i>	
1	2	3	4	5	6	7	8	9	10	
AGC	TCC	TGG	TGG	GCT	CAT	GTG	GAG	ATG	GGG	
<i>Ser</i>	<i>Ser</i>	<i>Trp</i>	<i>Trp</i>	<i>Ala</i>	<i>His</i>	<i>Val</i>	<i>Glu</i>	<i>Met</i>	<i>Gly</i>	
11	12	13	14	15	16	17	18	19	20	
CCC	CCA	GAT	CCC	ATC	CTG	GGA	GTC	ACA	GAA	
<i>Pro</i>	<i>Pro</i>	<i>Asp</i>	<i>Pro</i>	<i>Ile</i>	<i>Leu</i>	<i>Gly</i>	<i>Val</i>	<i>Thr</i>	<i>Glu</i>	
21	22	23	24	25	26	27	28	29	30	
GCC	TTT	AAG	AGA	GAC	ACC	AAC	AGC	AAA	AAG	
<i>Ala</i>	<i>Phe</i>	<i>Lys</i>	<i>Arg</i>	<i>Asp</i>	<i>Thr</i>	<i>Asn</i>	<i>Ser</i>	<i>Lys</i>	<i>Lys</i>	
. . .										

Fig. 1.5 Primary structure of the *n*-terminal portion of the pig mAspAT cDNA sequence. Numbers correspond to the amino acid sequence of the mature protein. Presequence region found in the precursor protein is italicized. Data from Joh *et al.*⁷⁶⁾

residues, (2) S-S bonds, and (3) N- and C-terminal amino acid residues. (3) is related to the problem of how to determine the reading frame and to assign the correct initiation codon, AUG = Met, for the protein expression. Amino acid sequence information on the amino-terminal region is required for this assignment. Mature protein does not necessarily contain Met as its amino-terminal residue.^{74,75)} A sequence determination from protein must be performed when the isolated protein does not have the amino-terminal Met or when multiple Met codons are found at the 5'-end of the coding region of cDNA. As an example, Fig. 1.5 shows a comparison of the beginning of the cDNA base sequence and the N-terminal portion of the amino acid sequence of the precursor form of pig mAspAT.⁷⁶⁾ A functional and mature protein found in cells has Ser for its amino-terminal end. If the amino-terminal sequencing has not been completed, it is possible to misjudge the correct initiation codon for Met since Met is also found at the 9th position and the -29th position from the N-terminal Ser. It was thought that the latter Met could be an initiation site based on the base sequence of the 5' noncoding region. Actually this region encodes for a signal sequence, which is required for the mAspAT precursor protein. mAspAT is translated on ribosomes in cytosol and must be transported into mitochondria.

We are now able to obtain not only the sequence information at the amino terminal region of the protein but also information concerning the inner structure of the protein from

only a minute amount (0.1–1 nmol) of sample. This is mainly due to HPLC improvements in separation power and detection sensitivity for peptide fragments and PHT-amino acids. Nonetheless, further improvements in efficiency and sensitivity in sequencing methods using DNA cloning technique are required.

1.3.3 Application of Mass Spectrometry to Protein Primary Sequence Determination

The mass spectrometer (MS) has become a useful tool for the study of protein and peptide structures as recent advancements in ionization techniques allow a protein or peptide to be ionized without being fragmented.^{77,78)} Such powerful techniques include plasma desorption (PD, also called ²⁵²Cf-PDMS), fast atom bombardment (FAB), electrospray ionization (ESI), and laser desorption (LD), all of which utilize both a high sensitivity ion detection system and a computerized data processing system. A typical MS is constructed with three separate components. The first component ionizes protein/peptide molecules. The second component guides ions to the mass analyzer while accelerating particles. Ionized particles are also separated according to their molecular mass as they travel to the third component where ions are detected according to their molecular mass. When a mass spectrometer is employed for protein sequence determination, protein digest is first ionized to release the highest mass fragments called parent ions for each corresponding peptide fragment. These parent ions are further fragmented into daughter ions. By analyzing masses of parent ions and daughter ions as well as mass differences between these ion fragments, a computer is able to reconstruct the entire amino acid sequence of the sample protein/peptide. When a protein of large molecular weight is analyzed, purification of its peptide fragments on HPLC, for example, is recommended prior to the mass analysis. A mass spectrometer can fragment a parent ion into daughter ions by colliding neutral atoms, such as He or Ar, or by irradiating a laser beam to break the peptide backbone, *i.e.*, collision induced dissociation (CID), and collision activated dissociation (CAD). Daughter ion mass is analyzed by either a linked scan mass analyzer or a tandem mass spectrometer (MS/MS) in which two mass spectrometers are connected in a row. The latter instrument is preferred for protein samples and is now commercially available.

In 1987, the primary structure of thioredoxin was determined solely by a tandem mass spectrometer.⁷⁹⁾ The first MS provided molecular weight of protease digests of thioredoxin by ionizing the digests with the FAB method and fractionated their daughter ions. The ionized peptides were further fragmented by the CID part of the instrument. The mass of ion fragments were detected by the second MS. A computer constructed the entire sequence by aligning the sequences found for each fragment. Mass spectrometers with capillary electrophoresis or HPLC have also been used for the structural analysis of peptide.⁷⁷⁾ Identification of a modified group in the protein structure can also be carried out on MS.^{77,80,81)} Structure determination of a 50K-dalton protein on mass spectrometer was successful with a 0.1–1 nmol sample, and the sequence of 30–70% of its entire structure was completed within 4–5 days. A recent report demonstrated that the structure of physiologically active peptide in fmol quantity was achieved by a combination of capillary electrophoresis and MS.⁸²⁾ Such achievement suggests the strong possibility that the peptide sequencer may be superseded by the mass spectrometer in the future.

1.3.4 Database

Computer database systems are employed to store vast amounts of sequence information. The following databases are currently available.

A. NBRF-PIR

National Biomedical Research Foundation specializes in providing a database for protein primary structure. This database contains all the information from the *Atlas of Protein Sequence and Structure* edited by M.O. Dayhoff. In this database proteins are categorized according to their super family grouping. In addition to the primary structure information, detailed descriptions of proteins, including active site, prosthetic group, etc., are included.

B. PRF/SEQDB

Protein Research Foundation (Japan) prepares this database from its parent database, PRF/LITDB (LITDB=literature database), which covers literature information on peptides and proteins. PRF/SEQDB (SEQDB=sequence database) is unique in that it allows duplicate entries when different authors report the sequence on the same protein.

C. GenInfo Backbone

National Library of Medicine started its database service in November, 1989. This contains sequence information on proteins and nucleotides appearing in MEDLINE.

D. SWISS-PROT

This database provides both protein sequences which are deduced from DNA sequences registered in the European Molecular Biology Laboratory (EMBL) and those from the NBRF/PIR database.

1.4 Conclusions

Today's sequence techniques have achieved 9–10 orders of magnitude higher detection limit for PTH-amino acid than Edman's first report, in which he described the sequence of Ala-Leu-Gly with 100 μ mol. Matsuo *et al.* succeeded in determining the structure of FSH-RH (LH-RH), a decapeptide hormone, in 1971 by using the Dansyl Edman method.⁸³⁾ This historical work used 200 nmol of FSH-RH purified from 165,000 porcine hypothalami and led a principal investigator, A. Schally, to win the Nobel Prize. Today, we would only need less than one hundred porcine hypothalami to complete the same job on a state-of-art sequencer. Those who plan the structural analysis of even barely detectable amounts of a peptide or protein are urged to start the plan; such study could open up a new area in biochemistry.

More protein sequences will be deduced from DNA sequences in the future. However, the sequencer utilizing the principle of Edman degradation will continue to be an important tool. Sensitivity should be enhanced with further development of fluorescent reagents. As enormous improvement has been made in recent years, mass spectrometry is quite promising for its role in sequence determination. Two-dimensional polyacrylamide gel electrophoresis and capillary electrophoresis show excellent resolution and detection sensitivity.

The demand for these methods in peptide and protein separations in micro-quantity will increase. A new device which combines electrophoretic apparatus and a mass spectrometer or a sequencer may be developed in the near future.

REFERENCES

1. A. P. Ryle, F. Sanger, L. F. Smith, R. Kitai, *Biochem. J.*, **60**, 541 (1955).
2. C. H. W. Hirs, S. Moore, W. H. Stein, *J. Biol. Chem.*, **235**, 633 (1960).
3. P. Edman, *Acta Chem. Scand.*, **4**, 283 (1950).
4. A. E. Braunstein, M. G. Kritzmman, *Enzymologia*, **2**, 129 (1937).
5. A. E. Braunstein, M. G. Kritzmman, *Biokhimiya*, **2**, 242 (1937).
6. A. E. Braunstein, M. G. Kritzmman, *Biokhimiya*, **2**, 859 (1937).
7. G. A. Fleisher, C. S. Potter, K. G. Wakim, *Proc. Soc. Expl. Biol. Med.*, **103**, 229 (1960).
8. N. Katsunuma, T. Matsuzawa, A. Huzino, *J. Vitaminol.*, **8**, 74 (1962).
9. H. Wada, Y. Morino, *Vitamins and Hormones*, **22**, 411 (1964).
10. H. Wada, H. Kagamiyama, T. Watanabe, In: *Pyridoxal Catalysis: Enzymes and Model Systems* (E. E. Snell, ed.) Wiley Interscience, New York, p. 111 (1966).
11. Y. Morino, H. Kagamiyama, H. Wada, *J. Biol. Chem.*, **239**, 943 (1964).
12. M. O. Dayhoff, "National Biomedical Research Foundation", Washington, D. C. (1972).
13. Y. A. Orchinnikov, C. A. Egorov, N. A. Aldanova, M. Y. Feigina, V. M. Lipkin, N. G. Abdulaev, E. V. Grishin, A. P. Kislev, N. N. Modyanov, A. E. Braunstein, D. L. Polyanovsky, V. V. Nosikov, *FEBS Lett.*, **29**, 31 (1973).
14. H. Kagamiyama, R. Sakakibara, H. Wada, S. Tanase, Y. Morino, *J. Biochem.*, **82**, 291 (1977).
15. S. V. Shlyapnikov, A. N. Myasnikov, E. S. Severin, M. A. Myagkova, Y. M. Torchinsky, A. E. Braunstein, *FEBS Lett.*, **106**, 385 (1979).
16. U. Graf-Hausner, K. J. Wilson, P. Christen, *J. Biol. Chem.*, **258**, 8813 (1983).
17. K. Kondo, S. Wakabayashi, T. Yagi, H. Kagamiyama, *Biochem. Biophys. Res. Commun.*, **122**, 62 (1984).
18. P. Christen, U. Graf-Hausner, F. Bossa, S. Doonan, In: *Transaminases* (P. Christen, D. E. Metzler, eds.) John Wiley & Sons, New York, p. 173 (1985).
19. A. Arnone, P. H. Rogers, C. C. Hyde, P. D. Briley, C. M. Metzler, D. E. Metzler, In: *Transaminases* (P. Christen, D. E. Metzler, eds.) John Wiley & Sons, New York, p. 138 (1985).
20. E. G. Harutyunyan, V. N. Malashkevich, V. M. Kochkina, Y. M. Torchinsky, In: *Transaminases* (P. Christen, D. E. Metzler, eds.) John Wiley & Sons, New York, p. 164 (1985).
21. V. V. Borisov, S. N. Borisova, G. S. Kachalova, N. I. Sosfenov, B. K. Vainshtein, In: *Transaminases* (P. Christen, D. E. Metzler, eds.) John Wiley & Sons, New York, p. 155 (1985).
22. J. N. Jansonius, G. Eichele, G. C. Ford, D. Picot, C. Thaller, M. G. Vincent, In: *Transaminases* (P. Christen, D. E. Metzler, eds.) John Wiley & Sons, New York, p. 110 (1985).
23. S. Kamitori, A. Okamoto, K. Hirotsu, T. Higuchi, S. Kuramitsu, H. Kagamiyama, Y. Matsuura, Y. Katsube, *J. Biochem.*, **108**, 175 (1990).
24. M.-H. Sung, K. Tanizawa, H. Tanaka, S. Kuramitsu, H. Kagamiyama, K. Hirotsu, A. Okamoto, T. Higuchi, K. Soda, *J. Biol. Chem.*, **266**, 2567 (1991).
25. F. Wold, *Ann. Rev. Biochem.*, **50**, 783 (1981).
26. H. Hirano, *J. Prot. Chem.*, **8**, 115 (1989).
27. P. Matsudaira, *Meth. Enzymol.*, **182**, 602 (1990).
28. H. Kagamiyama, K. Teranishi, S. Tanase, Y. Morino, R. Sakakibara, H. Wada, *J. Biol. Chem.*, **255**, 6138 (1980).
29. R. Sakakibara, H. Kagamiyama, S. Tanase, Y. Morino, H. Wada, *J. Biol. Chem.*, **255**, 6144 (1980).
30. K. Kondo, S. Wakabayashi, H. Kagamiyama, *J. Biol. Chem.*, **262**, 8648 (1987).
31. T. Bergman, In: *Methods in Protein Sequence Analysis* (K. Imahori, F. Sakiyama, eds.) Plenum Press, New York, p. 21 (1993).
32. R. Consden, A. H. Gordon, A. J. P. Martin, *Biochem. J.*, **38**, 224 (1944).
33. D. H. Spackman, W. H. Stein, S. Moore, *Anal. Chem.*, **30**, 1190 (1958).
34. S. Stein, S. Udenfriend, *Anal. Biochem.*, **136**, 7 (1984).
35. J. R. Benson, P. E. Hare, *Proc. Natl. Acad. Sci., USA*, **72**, 619 (1975).
36. A. Fiorino, G. Frigo, E. Cucchetti, *J. Chromatogr.*, **476**, 83 (1989).
37. M. Fujiwara, Y. Ishida, N. Nimura, A. Toyama, T. Kinoshita, *Anal. Biochem.*, **166**, 72 (1987).
38. H. W. Jarrett, K. D. Cooksy, B. Ellis, J. M. Anderson, *Anal. Biochem.*, **153**, 189 (1986).
39. D. W. Hill, F. H. Walters, T. D. Wilson, J. D. Stuart, *Anal. Chem.*, **51**, 1338 (1979).

40. B. A. Bidlingmeyer, S. A. Cohen, T. L. Tarvin, *J. Chromatogr.*, **336**, 93 (1984).
41. V. Stocchi, G. Piccoli, M. Magnani, F. Palma, B. Biagiarelli, L. Cucchiaroni, *Anal. Biochem.*, **178**, 107 (1989).
42. R. Knecht, J. -Y. Chang, *Anal. Chem.*, **58**, 2375 (1986).
43. V. T. Wiedmeier, S. P. Porterfield, C. E. Nendrich, *J. Chromatogr.*, **231**, 410 (1982).
44. R. Cunico, A. G. Mayer, C. T. Wehr, T. L. Sheehan, *Biochemistry*, **1**, 7 (1986).
45. R. J. Simpson, M. R. Neuberger, T. Y. Liu, *J. Biol. Chem.*, **251**, 1936 (1976).
46. F. Westall, W. Hesser, *Anal. Biochem.*, **61**, 610 (1974).
47. A. Tsugita, J. -J. Scheffler, *Eur. J. Biochem.*, **124**, 585 (1982).
48. S. -H. Chiou, K. -T. Wang, *J. Chromatogr.*, **491**, 424 (1989).
49. S. A. Cohen, D. J. Strydom, *Anal. Biochem.*, **174**, 1 (1988).
50. S. Tsunasawa, K. Narita, K. Ogata, *J. Biochem.*, **77**, 89 (1975).
51. A. Scaloni, W. M. Jones, D. Barra, M. Pospischil, S. Sassa, A. Popowicz, L. R. Manning, O. Schneewind, J. M. Manning, *J. Biol. Chem.*, **267**, 3811 (1992).
52. S. Tsunasawa, H. Takakura, F. Sakiyama, *J. Prot. Chem.*, **9**, 265 (1990).
53. D. N. Podell, G. N. Abraham, *Biochem. Biophys. Res. Commun.*, **81**, 176 (1978).
54. Y. Sugino, S. Tsunasawa, K. Yutani, K. Ogasahara, M. Suzuki, *J. Biochem.*, **87**, 351 (1980).
55. K. Narita, H. Matsuo, T. Nakajima, In: *Protein Sequence Determination* (S. B. Needleman, ed.), Springer-Verlag, New York, NY, p. 30 (1975).
56. R. Hayashi, S. Moore, W. H. Stein, *J. Biol. Chem.*, **248**, 2296 (1973).
57. T. Kumazaki, K. Terasawa, S. Ishii, *J. Biochem.*, **102**, 1539 (1987).
58. J. Sjoquist, *Acta Chem. Scand.*, **7**, 447 (1953).
59. W. R. Gray, *Meth. Enzymol.*, **25**, 121 (1972).
60. W. R. Gray, *Meth. Enzymol.*, **25**, 333 (1972).
61. J. J. Pisano, T. J. Bronzert, *J. Biol. Chem.*, **244**, 5597 (1969).
62. P. Edman, G. Begg, *Eur. J. Biochem.*, **1**, 80 (1967).
63. O. Smithies, D. Gibson, E. M. Fanning, R. M. Goodfliesth, J. G. Gilman, D. L. Ballantyne, *Biochemistry*, **10**, 4912 (1971).
64. E. Mendez, C. Y. Lai, *Anal. Biochem.*, **68**, 47 (1975).
65. R. A. Laursen, *Eur. J. Biochem.*, **20**, 89 (1971).
66. H. Kondo, Y. Baba, K. Takaki, K. Kondo, H. Kagamiyama, *Biochem. Biophys. Res. Commun.*, **142**, 155 (1987).
67. J. M. Coull, D. J. C. Pappin, J. Mark, R. Aebersold, H. Köster, *Anal. Biochem.*, **194**, 110 (1991).
68. R. M. Hewick, M. W. Hunkapiller, L. E. Hood, W. J. Dreyer, *J. Biol. Chem.*, **256**, 7990 (1981).
69. M. W. Hunkapiller, L. E. Hood, *Science*, **219**, 650 (1983).
70. C. L. Zimmerman, J. J. Pisano, *Meth. Enzymol.*, **47**, 45 (1977).
71. N. D. Johnson, M. W. Hunkapiller, L. E. Hood, *Anal. Biochem.*, **100**, 335 (1979).
72. S. Tsunasawa, J. Kondo, F. Sakiyama, *J. Biochem.*, **97**, 701 (1985).
73. M. Ishiguro, M. Suzuki, K. Takio, T. Matsuzawa, K. Titani, *Biochemistry*, **30**, 6048 (1991).
74. H. Kagamiyama, R. Sakakibara, S. Tanase, Y. Morino, H. Wada, *J. Biol. Chem.*, **255**, 6153 (1980).
75. Y. Inoue, Y. Kawasaki, T. Harada, K. Hatakeyama, H. Kagamiyama, *J. Biol. Chem.*, **266**, 20791 (1991).
76. T. Joh, H. Nomiyama, S. Maeda, K. Shimada, Y. Morino, *Proc. Natl. Acad. Sci., USA*, **82**, 6065 (1985).
77. D. F. Hunt, J. E. Alexander, A. L. McCormack, P. A. Martine, H. Michel, J. Shabanowitz, N. Sherman, In: *Techniques in Protein Chemistry II* (J. J. Villafranca, ed.) Academic Press, New York, p. 441 (1991).
78. B. T. Chait, S. B. H. Kent, *Science*, **257**, 1885 (1992).
79. R. S. Johnson, K. Biemann, *Biochemistry*, **26**, 1209 (1987).
80. S. A. Carr, K. Biemann, *Meth. Enzymol.*, **106**, 29 (1984).
81. Y. Fukada, T. Takao, H. Ohguro, T. Yoshizawa, T. Akino, Y. Shimonishi, *Nature*, **346**, 658 (1990).
82. M. A. Moseley, L. J. Deterding, K. B. Tower, J. W. Jorgenson, *Anal. Chem.*, **63**, 109 (1991).
83. H. Matsuo, Y. Baba, R. M. G. Nair, A. Arimura, A. V. Schally, *Biochem. Biophys. Res. Commun.*, **43**, 1334 (1971).

2

Nuclear Magnetic Resonance Studies on *Streptomyces* Subtilisin Inhibitor and its Complexes with Proteinases

Yoko Miyake and Masatsune Kainosho

Faculty of Science, Tokyo Metropolitan University,
1-1 Minamiohsawa, Hachioji, Tokyo 192-03, Japan

2.1 Introduction

Streptomyces subtilisin inhibitor (SSI) was isolated from the cultured broth of *Streptomyces albogriseolus* S-3253 by Murao and Sato in 1971 as a new class of protein proteinase inhibitors.¹⁾ Since then, this protein has been extensively used in Japan as a target protein for collaborative researches from many aspects. The earlier results obtained from these researches have been summarized in the monograph entitled, "*Protein Proteinase Inhibitor-The Case of Streptomyces Subtilisin Inhibitor (SSI)*".²⁾ SSI has two features making it distinctive from the other proteinase inhibitors: a dimeric structure and a wide inhibitory spectrum.^{3,4)} It has recently been found that SSI-like proteins are produced by a wide variety of *Streptomyces*,⁵⁾ leading one to assume that they may possess some essential biological functions for these microorganisms. Since these family proteins have quite conservative amino acid sequences, especially in the five-fold β -sheets region (Fig. 2.1), which is responsible for the subunit-subunit interface, the dimeric structure may have been conserved in order to maintain their primary biological functions, which are not yet firmly established. The ultimate goal of our study is to sort the NMR evidence to correlate the structural features of SSI with its inhibitory activity.

The molecular weight of SSI, which is 23 kDa as a homodimer, makes most of the NMR techniques difficult to apply.⁶⁾ The situation is even more serious for the SSI complexes with target proteinases, since each of the SSI subunits binds with a proteinase molecule without dissociating into monomer. The SSI-subtilisin BPN' complex, for example, is 78 kDa, which is much too large for any of the conventional NMR methodologies. We therefore had to develop a new NMR strategy suitable for acquiring structural information for larger protein complex.⁷⁻⁹⁾ Appropriate use of the stable isotopes may be the only way to overcome difficulties associated with such large proteins.¹⁰⁾ Most of the stable isotope-assisted NMR methods, however, focus on the side-chain moieties since the catalytic or functional properties of proteins are usually carried out by the side-chain functional groups in the active sites.¹¹⁾ In the case of protein proteinase inhibitors, on the other hand, the primary functional group is a backbone peptide bond, which is often called the "scissile bond."^{2,12)}

It is therefore not surprising that structural studies of various proteinase-inhibitor

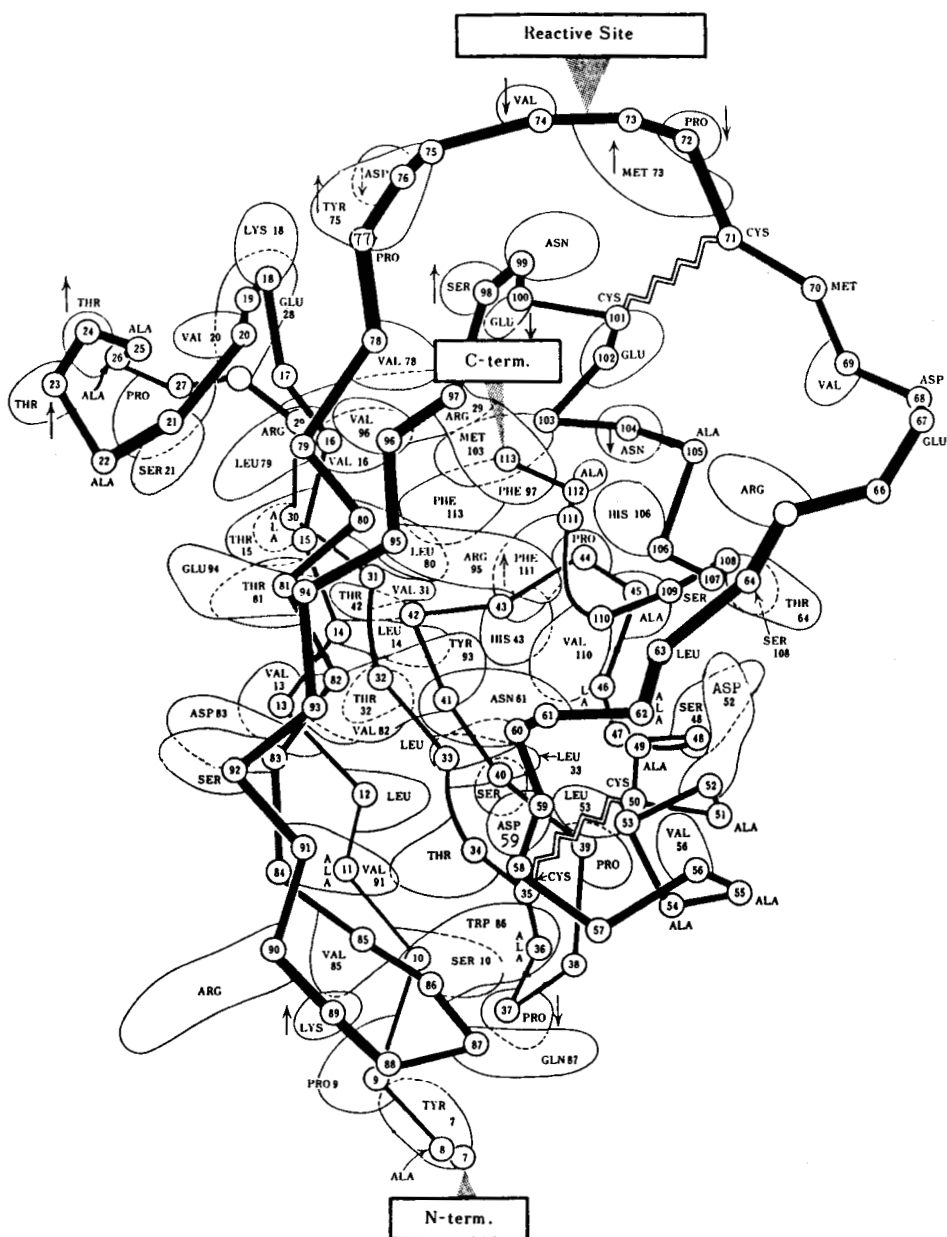


Fig. 2.1 Structure of SSI subunit. (Reproduced with permission from Y., Mitsui, *et al.*, *Protein Protease Inhibitor—The Case of Streptomyces Subtilisin Inhibitor (SSI)* (K. Hiromi *et al.* ed.), p. 174, Elsevier, Amsterdam (1985)).

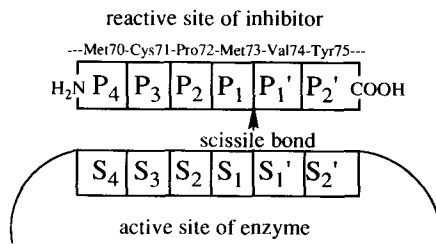


Fig. 2.2 Subsites of enzyme and inhibitor. Commonly used nomenclature for amino acid residues of the subsites (S_1 , etc.) of the active site of a proteinase and corresponding subsites of its inhibitor (P_1 , etc.). The arrow indicates the peptide bond to be split by the enzyme (the scissile bond). The corresponding residual number of the reactive site subsites of SSI are also indicated.

complexes have so far been directed toward the states of the scissile bonds in the complexes. X-ray analyses of various complexes have indicated that the scissile bonds between the P_1 and P_1' residues (Fig. 2.2) are always in direct contact with the catalytic residues of the target proteinases.¹³⁻¹⁶ Precise structural information, such as the planarity of the scissile bonds in the complexes, could barely be obtained even from accurate X-ray analyses.^{17,18} In principle, NMR spectroscopy should afford a better way to obtain more precise local information.

2.2 Assignment of the Backbone Peptide Carbonyl Carbon ^{13}C NMR Signals for Larger Proteins

Structural information of the scissile bond in an inhibitor-proteinase complex can be most directly obtained if one can selectively observe the ^{13}C NMR signal of the P_1 carbonyl carbon. It was, however, difficult to observe a single carbon resonance at a specific peptide bond in a large protein such as the SSI-subtilisin complex. We developed a stable isotope assisted NMR method for this particular purpose, now widely known as the ^{15}N , ^{13}C -double labeling method.⁷⁻⁹ This method, which utilizes the conformation-independent ^{15}N - ^{13}C one-bond coupling constants ($^1J_{\text{CN}}$), will be illustrated below taking the case of assignment of the three Met carbonyl signals in SSI as an example.⁷⁾

The Met carbonyl carbon signals of SSI could be selectively enhanced by a factor of 90 over the natural abundant (1.1%) ^{13}C signal level by feeding *Streptomyces albogriseolus* with $[1-^{13}\text{C}]\text{Met}$ (98% ^{13}C). The SSI obtained from the cultured fluid was found to be labeled with $[1-^{13}\text{C}]\text{Met}$. The labeled SSI, which we denote as $[\text{M}]\text{SSI}$, showed in fact three enhanced carbonyl carbon resonances corresponding to the three Met residues in the SSI subunit, i.e. Met⁷⁰ (P_4 residue), Met⁷³ (P_1 residue), and Met¹⁰³. Further assignment of these signals to the specific site in the sequence (Fig. 2.3) was achieved by labeling SSI simultaneously with $[1-^{13}\text{C}]\text{Met}$ and ^{15}N amino acid, which corresponds to each of the C-terminal neighbors of the three Met residues: Met⁷⁰-Cys¹⁰¹, Met⁷³-Val⁷⁴, and Met¹⁰³-Asn¹⁰⁴. As the carbonyl carbon of the i -th residue is directly bonded to the amide nitrogen of the $(i+1)$ -th residue, each of the doubly labeled SSIs must show a doublet signal for only one of the three Met carbonyl signals. If there is only a unique dipeptide pair, as in the case of the Met residues in SSI, one can assign the carbonyl carbon signals by a series of specific

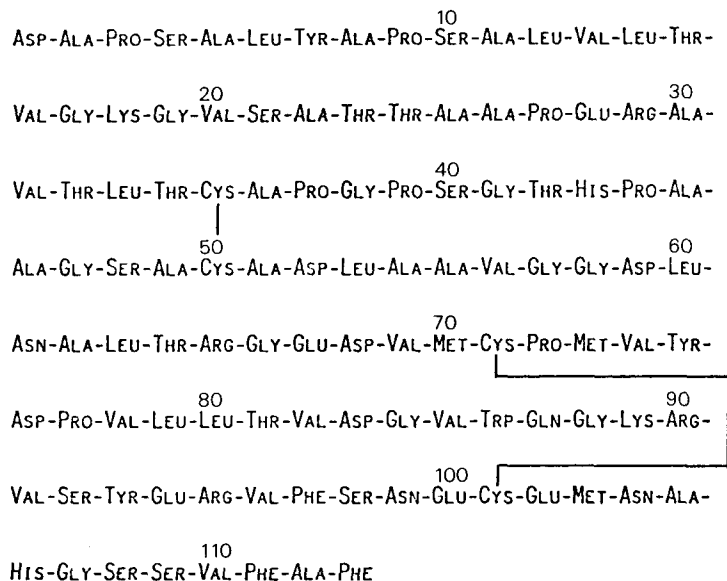


Fig. 2.3 Amino acid sequence of SSI subunit. (Reproduced with permission from T. Ikenaka, *et al.*, *J. Biochem.*, 76, 1191 (1974)).

double labeling experiments.

The carbonyl carbon regions of the ^{13}C NMR spectra of [M]SSI, [M,V]SSI, and [M,C]SSI are shown in Fig. 2.4. We designate the double labeled SSIs as [A,B]SSI, where the first and second letters indicate [$1\text{-}^{13}\text{C}$]- and [$^{15}\text{N}_\alpha$]amino acid fed to the microorganisms, respectively. The assignment shown in Fig. 2.4 was straightforward since each of the two specifically double labeled SSI showed a doublet pattern very clearly, although the center peak due to the Met carbonyl carbon adjacent to the ^{14}N neighbors overlapped the doublet signals. The ^{15}N enrichment of Val and Cys residues in [M,V]SSI and [M,C]SSI, respectively, was high enough for clear assignment. This may not always be the case for other amino acids where ^{15}N scrambling reactions due to transamination take place during the cultivation. The specific double labeling method may be supplemented by other assignment techniques, such as site-directed mutagenesis, if the dipeptide unit in question is not unique in the sequence or specific ^{15}N labeling is not attainable. We have hitherto established virtually all of the carbonyl carbon resonances of SSI by a combination of these assignment techniques.

The double labeling method can also be applied to proteins of much larger size, since the carbonyl carbons usually give narrower ^{13}C signals than those having directly bonded proton(s).¹⁹⁾ Arata and coworkers have successfully applied the method for intact IgG, whose molecular weight is 160 kDa. The credibility of the method may often be reduced, especially in the case of larger proteins, since the carbonyl carbon signals severely overlap even by specific labeling and broaden to the order of $^1J_{\text{CN}}$ (15 Hz or so). The ambiguity of the carbonyl carbon assignment by the double labeling method arises actually from the poor resolution among the ^{15}N satellite signals and ^{14}N center signal due to the increased line-

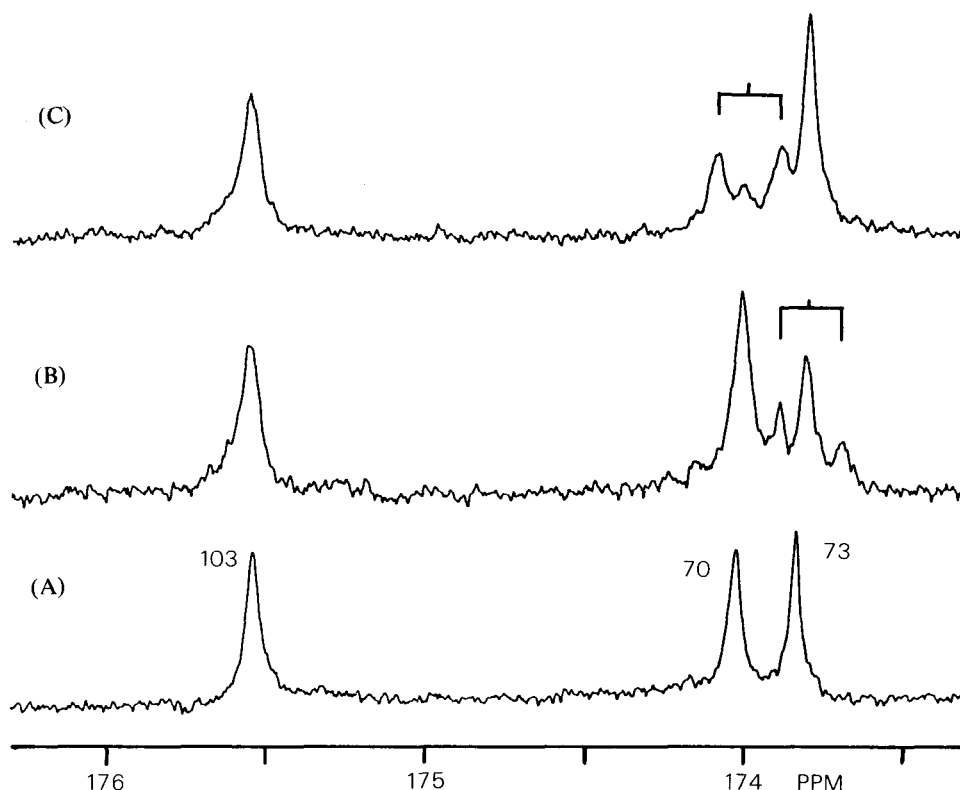


Fig. 2.4 Double labeling experiments to assign the Met carbonyl carbon resonances of SSI. A) [M]SSI; B) [M,V]SSI; C) [M,C]SSI; (5 mg/0.5 ml, 50 mM phosphate D_2O buffer, $\text{pD}=7.3$, 40°C , 75.4 MHz.) The signal which was decreased in intensity at the expense of the satellite peaks due to spin coupling with the amide ^{15}N of Val⁷⁴ in the spectrum of [M,V]SSI(B) could be unambiguously assigned to Met⁷³, and that in the spectrum of [M,C]SSI(C) could be assigned to Met⁷⁰. The result of the assignment of the three Met residues of SSI is indicated on (A). The chemical shifts were measured from the internal dioxane peak as 67.8 ppm.

widths for larger proteins. The assignment based on the standard double labeling experiments for the SSI-subtilisin complex, for example, was not as clear as that for a small protein such as free SSI. A modified version of ^{15}N , ^{13}C -double labeling method, which we denote as the reverse double labeling method, was found to work better for the complex. In the reverse double labeling method, all nitrogen atoms are uniformly labeled with ^{15}N , except for the amino acid type of the C-terminal neighbor of the ^{13}C -labeled site. In the reverse double labeling experiments, the assignment is made by detecting a sharp center singlet signal adjacent to the labeled ^{14}N atom. We will describe this alternative method later.

2.3 Structural Elucidation of the Stable Intermediate for the SSI-Subtilisin System by ^{13}C NMR Spectroscopy

The other benefit of the ^{15}N , ^{13}C -double labeling, which may in fact be more crucial for

the structural study of the proteinase-inhibitor complexes, is that the scissile bond itself can be "labeled" by double labeling. In the case of [M,V]SSI, one can use the presence or absence of $^1J_{\text{CN}}$ for the Met⁷³ carbonyl carbon signal to deduce whether the scissile bond is cleaved or not. With the spin coupling constant as the structural probe, one can differentiate among many possible intermediates which may exist in the equimolar mixture of SSI and subtilisin (Fig. 2.5A,B). Moreover, as will be described below, the $^1J_{\text{CN}}$ value also gives us an invaluable clue for assessing the hybridization states of the scissile peptide bond atoms.^{2,20,21)}

Despite the large molecular weight, the ^{13}C NMR spectrum of [M]SSI-subtilisin BPN' complex in an aqueous solution clearly showed three enhanced resonances in the carbonyl

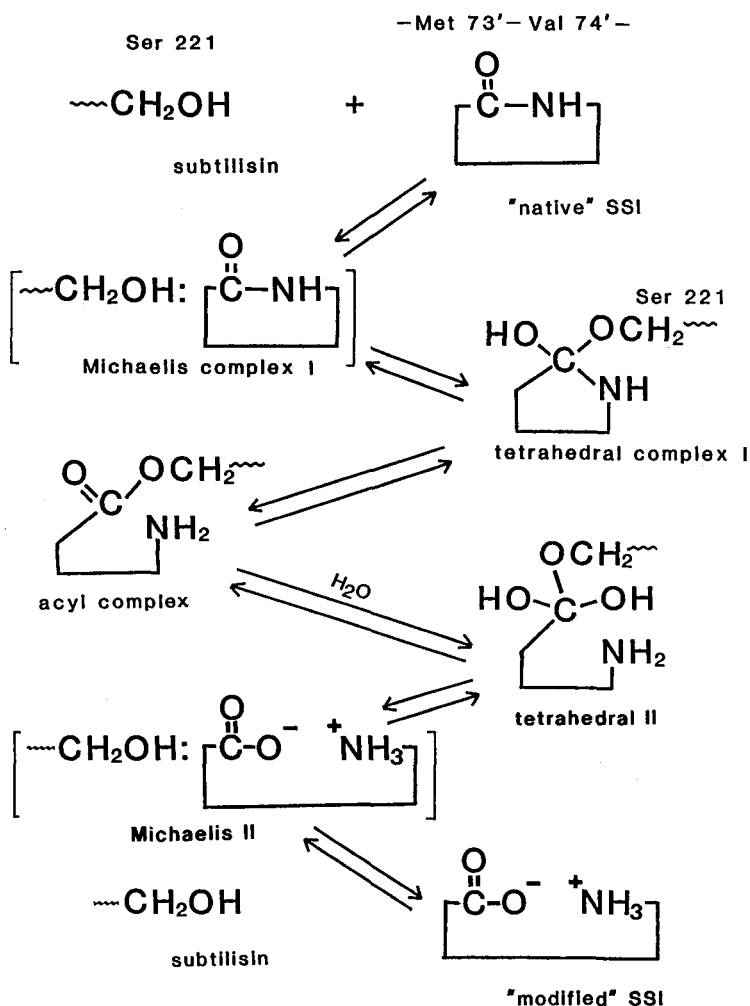


Fig. 2.5(A) Equilibria among intermediates for SSI-subtilisin mixture. $\sim$ indicated the rest of the subtilisin molecule except for CH_2OH of active site Ser-221.

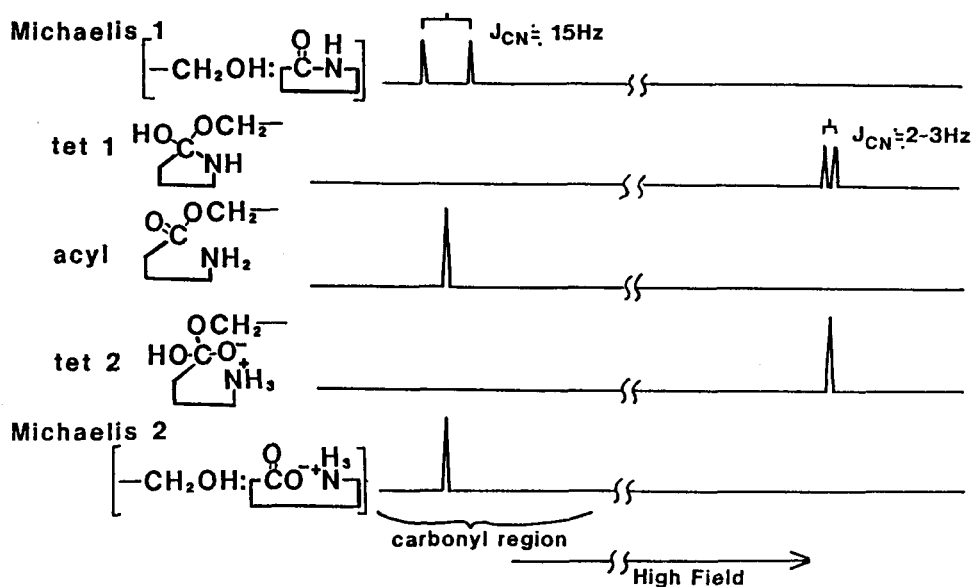


Fig. 2.5(B) Expected NMR patterns for each intermediate. The chemical shift and the $^1J_{\text{CN}}$ coupling of the carbonyl carbon signal of P_1 residue must be very sensitive to the state of the complex. In the Michaelis I intermediate, the P_1 carbonyl carbon should have the resonance in the carbonyl region and its coupling from ^{15}N of P_1' residue should be observed at about 15 Hz, since the scissile bond in this intermediate is normal peptide conformation. In the tetrahedral I intermediate, it should shift strongly upfield and the coupling constant should be decreased to 2–3 Hz, since the hybridization of the P_1 carbonyl carbon changes from sp^2 to sp^3 by the nucleophilic addition of the γ -O of active Ser of the proteinase. In the other three intermediates, $^1J_{\text{CN}}$ should be diminished, since the scissile bond would not be bonded.

carbon region. These three resonances are apparently due to three Met residues of the SSI subunit in the SSI-subtilisin BPN' complex. There were no other enhanced signals in the whole spectral region. The assignment of these signals, which appeared at quite different chemical shift positions compared to those for the free SSI, has been established by the reverse ^{14}N , ^{13}C -double labeling methods (Fig. 2.6). The carbonyl carbon signals of [M, rev ^{14}N -V]SSI, in which all nitrogen except Val residues are uniformly ^{15}N labeled, shows a sharp center peak only for the lower field resonance. On the other hand [M, rev ^{14}N -C]SSI shows a sharp center peak at the highest field. These center peaks are clearly due to the Met carbonyl bound with ^{14}N . This modified double labeling method was found to be more effective for assigning the carbonyl resonances of larger proteins.

The carbonyl carbon signal of Met⁷³ in the subtilisin complex (P_1 residue) was shifted low field by 3.67 ppm. The signal due to Met⁷⁰ (P_4 residue), on the other hand, shifted toward up field by 1.47 ppm. Very interestingly, the carbonyl carbon signals due to Met¹⁰³ also shifted by 0.45 ppm down field (see Table 2.1). Met¹⁰³ is in the α -helix of the C-terminal side of the polypeptide chain, which has been denoted as the " α_2 -helix,"³⁾ and is not in direct contact with the bound subtilisin. This unexpected result drove us to survey extensively the induced chemical shift changes for residues other than Met and led us to find the global

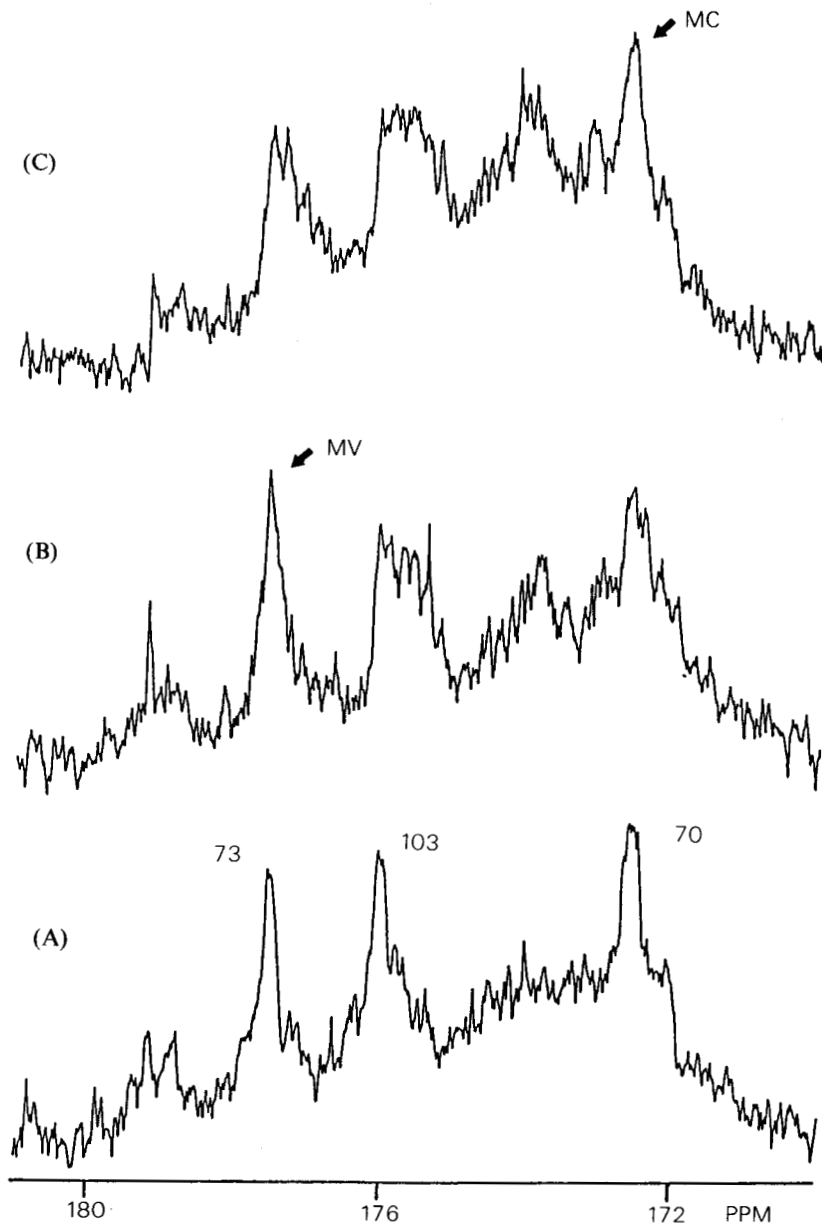


Fig. 2.6 Assignment of the Met carbonyl signals of the SSI-subtilisin BPN' complex by reverse double labeling methods. A) [M]SSI-subtilisin BPN' complex; B) [M, rev¹⁴N-V]SSI-subtilisin BPN' complex; C) [M, rev¹⁴N-C]SSI-subtilisin BPN' complex. (0.7 mM, 40°C, 50 mM phosphate D₂O buffer pD=7.3, 75.4 MHz). The result of the assignment of Met residues of [M]SSI-subtilisin BPN' complex by this reverse labeling method was indicated on the spectrum (A).

TABLE 2.1 List of the chemical shifts of [M]SSI in free and complexed state with subtilisin BPN' (40°C, pD=7.3). Complexation shift (c)=shift (b) – shift (a)

residue number		(a) free (ppm)	(b) complex (ppm)	(c) complexation shift (ppm)
Met	103	175.55	176.00	0.45
	70	174.03	172.55	-1.47
	73	173.84	177.51	3.67

conformational change of the SSI complexed with subtilisin.

We could immediately conclude from the spectra shown in Fig. 2.6 several important consequences: (1) there was only one major intermediate in the equimolar mixture of SSI and subtilisin BPN' in aqueous solution at pH 7.3; (2) the chemical shift of the P₁ residue Met⁷³ was in the carbonyl carbon region, therefore the possibility of a tetrahedral intermediate was eliminated; (3) a doublet signal for Met⁷³ observed in the spectrum of [M, *ul*-¹⁵N]SSI indicated that the intermediate was in fact the Michaelis complex with the scissile bond intact. Controversy regarding the planarity of the scissile bond in the Michaelis complex was also be resolved by NMR spectroscopy. As the spin coupling interactions between directly bonded nuclei are governed mainly by the Fermi contact term, which is proportional to the *s*-electron densities of the interacting nuclei, ¹*J*_{CN} values and the hybridization of the carbonyl carbon and amide nitrogen atoms may be expressed by the following empirical formula:

$$^1J_{CN}(\text{Hz}) = 80(\%S_C)(\%S_N)$$

where %*S*_C and %*S*_N are the fractional *s*-characters at the carbon and nitrogen atoms of a peptide bond, respectively.²²⁻²⁴⁾

The important implication of this formula is that ¹*J*_{CN} of the nonplanar peptide bond, which has less *s*-characters for both atoms, should decrease in value from those of the planar peptide bonds. We have measured the coupling constant between the carbonyl carbon of Met⁷³ and the amide nitrogen of Val⁷⁴ quite accurately by lowering the external field strength (4.6 tesla; 50 MHz for ¹³C NMR) where the line-widths of carbonyl carbon become considerably narrower compared to those measured at higher fields, because of the lower contribution of the chemical shift anisotropy term in the observed line-widths. The ¹*J*_{CN} value (15.2 Hz) for Met⁷³–Val⁷⁴ in the subtilisin complex was almost the same as that in the free SSI (15.4 Hz). These values are almost identical to the averaged ¹*J*_{CN} values (15.3±0.7 Hz) which have been determined for the peptide bonds in free SSI. The result indicates that there is no evidence to support any appreciable distortion of the scissile bond in the SSI-subtilisin complex from the standard peptide bond plane in a globular protein.

2.4 Facile Restoration of the Hydrolyzed Scissile Bond of SSI* in the Subtilisin Complex

It has been established that the complex between SSI and subtilisin BPN' exists in aqueous solution as the Michaelis intermediate with the intact undistorted scissile bond. This intermediate seems to be identical to that found in the solid state, although the precise

structure of the scissile bond, namely the planarity of the peptide bond plane, could barely be determined by the X-ray analyses.^{17,25)} A question remains whether the Michaelis complex is really a thermodynamically stable intermediate in the SSI-subtilisin mixture or not. This may not be the case if the next step from the Michaelis complex to the tetrahedral intermediate requires activation energy too high to overcome. In such a case, the Michaelis complex should be regarded as the kinetically controlled product rather than thermodynamically favorable one. These two alternative possibilities should be distinguished by comparing the subtilisin complex prepared from the scissile bond cleaved SSI, SSI* (the modified SSI), with that of the native SSI (the virgin SSI).

SSI*, which retains strong inhibitory activity, was prepared in good yield by incubating the acidic aqueous solution of the complex at low temperature.^{26,27)} The ¹³C NMR spectrum of [M]SSI*, which was derived from the [M]SSI-subtilisin complex, showed three enhanced Met carbonyl resonances. The newly formed carboxyl carbon of Met⁷³ of [M]SSI* appeared as a very sharp signal at 5.5₆ ppm down field from the original Met⁷³ position of [M]SSI. The signals due to Met⁷⁰ and Met¹⁰³ also shift a little, but not much. In order to assess the effect of the hydrolysis of the scissile bond on the overall structure of SSI, we have extensively compared the chemical shifts of the backbone carbonyl carbon resonances of the residues other than Met and found that the conformational change localized at the C-terminal half of the SSI subunit. The region is completely overlapped with that found for the conformational change induced by protease binding (*see below*).

We compared the ¹³C NMR spectra of the subtilisin complexes of [M]SSI and [M]SSI*. Rather surprisingly, the NMR spectrum taken within two hours after the preparation of [M]SSI*-subtilisin BPN' complex was absolutely identical to that of [M]SSI-subtilisin BPN'. A period of two hours was necessary to obtain the ¹³C NMR spectrum of the [M]SSI*-subtilisin complex with a sufficient signal-to-noise ratio. This included the time after the modified inhibitor was brought in contact with subtilisin BPN'. There were no extra signals detected other than those observed for the [M]SSI-subtilisin complex, indicating that the cleaved scissile bond in [M]SSI* can be rapidly restored in the complex. Since time-resolution of ¹³C NMR spectroscopy is rather limited by its inherent insensitivity, we are not able to tell exactly how fast this process is. Tonomura *et al.*, however, have recently found by a stopped flow technique an unknown kinetic process having a half-life time of two seconds for the SSI*-subtilisin system. Obviously this process should be the restoration process of the cleaved scissile bond of SSI* in the complex. Therefore, the hydrolyzed scissile bond could in fact be restored within several seconds! (private communication).

This extremely facile restoration process indicates that the environment of the scissile bond in the complex is quite hydrophobic in nature so the dehydration reaction, which is highly endergonic in aqueous environment, proceeds unusually rapidly. One can therefore conclude that the Michaelis intermediate with the intact scissile bond is indeed the thermodynamically stable form and not the kinetically controlled one. The subtilisin is being entrapped in the complex since it is excessively stable, hence its proteolytic activity is inhibited by SSI.

Why is the complex so stable? This seems to be a natural question but it cannot be answered by the NMR evidence obtained only for the three Met residues. An important clue was obtained, as we have pointed out already, from the chemical shift change observed for Met¹⁰³, which is located in the second of the two α -helices in the SSI subunit, and

designated as the α_2 helix. The Met¹⁰³ is not in direct contact with subtilisin, but the carbonyl carbon resonance shifts 0.4₅ ppm down-field in the complex. The conformational changes induced by the complex formation with proteinases were thus not limited to the SSI-proteinase interface, where Met⁷⁰ and Met⁷³ exist. We decided to determine the region of the SSI subunit where the conformational change is being induced by the proteinase binding.

2.5 Global Conformational Changes of SSI Induced by Complex Formation with Subtilisin

The fact that the carbonyl carbon chemical shifts of residues other than the scissile bond, such as Met⁷⁰ and Met¹⁰³, showed considerably large chemical shift changes induced by the complex formation with subtilisin BPN' led us to examine the distribution of the induced chemical shifts of these other residues. We prepared a number of SSI's selectively labeled with a single type of [¹³C]amino acids and independent assignments for the complexed and free SSI were often required. The elaborate experiments afforded, however, extremely interesting results.²¹⁾ By plotting the induced shift values against the SSI sequence, we immediately found that only the C-terminal half of the SSI subunit changes its conformation on the subtilisin binding (Fig. 2.7). This region, which contains the scissile bond, includes almost all the residues that showed large chemical shift changes, leading us to conclude that the region plays a role as the functional domain.

This region, which we denote domain II, behaves quite different from the N-terminal half of the SSI subunit (domain I). The structure of domain I is almost unchanged upon subtilisin binding as revealed by very small induced chemical shifts induced on subtilisin binding. By taking a little closer look at Fig. 2.7, however, one may notice that two β -strands, β_4 and β_5 , in domain II (residue number *ca.* 80–95) show considerably less chemical shift changes compared to the other residues in this domain. As these two β -strands are involved within the five-fold antiparallel β -sheets where the SSI subunits are associated with each other. Large conformational changes may not be possible for the interface residues of this subunit, if SSI needs to remain a dimer. Actually, the SSI dimer could be dissociated into monomers in the presence of a surfactant such as sodium dodesyl sulfate (SDS), but the monomeric SSI was in a denatured state and no longer had inhibitory activity. Note that the flexible part found for the subtilisin complex is almost the same as that observed for the scissile bond cleavage (see above). We can therefore speculate that the SSI molecule binds various proteinases by the concomitant adjustment of the conformation of the flexible polypeptide segments in domain II in such a way that the Michaelis intermediate becomes energetically the most favorable state. The flexible segment can be supported by anchoring onto the rather rigid subunit interface and domain I, and therefore is quite resistant to proteinase digestion.

The conformational flexible part found in domain II of SSI is in stark contrast with other protein proteinase inhibitors, such as BPTI. In the case of BPTI, the backbone conformation is found to be nearly identical in both the free and the bovine trypsin complex.²⁸⁾ The conformational rigidity of protein proteinase inhibitors has been considered for a long time to be a necessary condition to inhibit their target enzymes and to protect themselves from attack by other proteinases. It is generally recognized that the substrate-like protein proteinase inhibitors, such as BPTI, STI, and Ovomucoid domain 3,

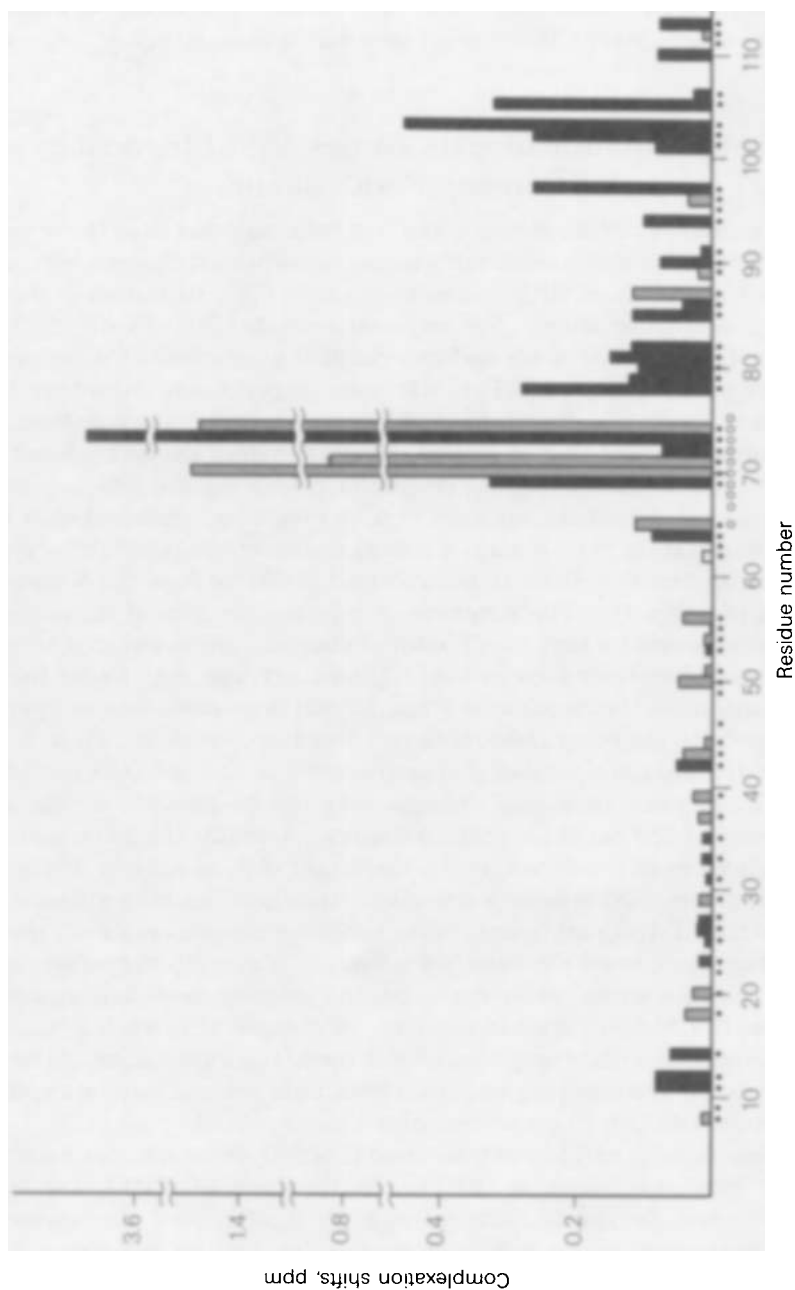


Fig. 2.7 Differences in the chemical shift of the SS1 backbone carbonyl ^{13}C resonances in free and complexed state with subtilisin BPN' (pH 7.3, 40°C). The dark and dotted bars show down- and up-field shifts by complexation, respectively. Small solid circles on the residue numbers (horizontal scale) represent the examined residues. The residues in direct contact with the complexed subtilisin are shown by small open circles. (Reproduced with permission from Y. Miyake, Thesis, Tokyo Metropolitan University (1992))

exhibit inhibitory activities because the activation energy from the Michaelis intermediate to the subsequent intermediate is too high. However, the SSI complexed with subtilisin is the thermodynamically most stable state. The conformational flexibility found for SSI seems to be closely related with its wide inhibitory spectrum,^{2,21)} although the binding constants are not enormously high for each enzyme. This hypothesis can be tested experimentally.

Very recently it was shown that the X-ray analysis of eglin c and its complex with subtilisin revealed a similar type of global structural changes induced by the subtilisin binding.²⁹⁾ Obviously SSI has no sequence or structural similarity with eglin c, which is a monomeric protein having 70 amino acid residues produced by the leech *Hirudo medicinalis*. It will thus be very interesting to compare the structural as well as biochemical characteristics of these two inhibitors.

2.6 Conformations of the Flexible Segments of Domain II and the Topology of Proteinase Surface are Interdependent

If the flexible segments in domain II are in fact responsible for stabilizing the proteinase complexes by adjusting their conformation, the induced chemical shifts of the residues in those segments should be different for various enzymes. Since there are a number of subtilisin variants having slightly different surface topology for the SSI-subtilisin interface, a comparison of the chemical shift changes among the SSI complexes with subtilisin variants would be interesting to discover how subtle conformational differences in the enzyme structures can be reflected in the flexible segments of SSI. The induced shifts of the complexes prepared from the three subtilisin variants *subtilisin BPN'*, *Carlsberg*, and *Bacillus amylosacchariticus* were compared.³⁰⁾

The overall features of the induced shift profiles of these three complexes are of course quite similar, although there are a few distinctive differences in the induced shift values (Fig. 2.8). We have observed especially noteworthy differences for the residues in the flexible segments which are not in direct contact with subtilisins, leading us to conclude that the conformations of the flexible segments containing those residues are in fact slightly different for the three subtilisin variants. This was the first evidence found for the induced fit type conformational changes in the flexible segments of domain II.

The above experiments imply that the flexible segments of SSI change conformation in order to accommodate the enzyme surface by induced fit and not *vice versa*. If this assumption is correct, conformational changes in the flexible segments of SSI induced by the amino acid replacement for the residues which are not in direct contact with the bound proteinases should be restored in the complex. Fortunately, SSI has been cloned by Miura *et al.* and the high expression system has been firmly established.³¹⁾ We have constructed various mutants which have different conformation in the flexible segments but no change in the inhibitory activity. Taking the subtilisin-induced shift found for the Met¹⁰³ residue in the α_2 -helix into account, we prepared mutants for the α_2 -helix residues, such as E¹⁰²Q, E¹⁰²H H¹⁰⁶E SSI.³²⁾ Since the α_2 -helix is not in direct contact with the bound subtilisin, all of these mutants have the same inhibitory activity as the WT SSI and are biologically silent mutants. The chemical shifts of the backbone carbonyl carbons of these mutants, on the other hand, were quite different, especially for the flexible segments in domain II. Therefore, the conformations of these mutants are considerably different from WT SSI, but such conformational changes are easily restored by the template effect of subtilisin (or

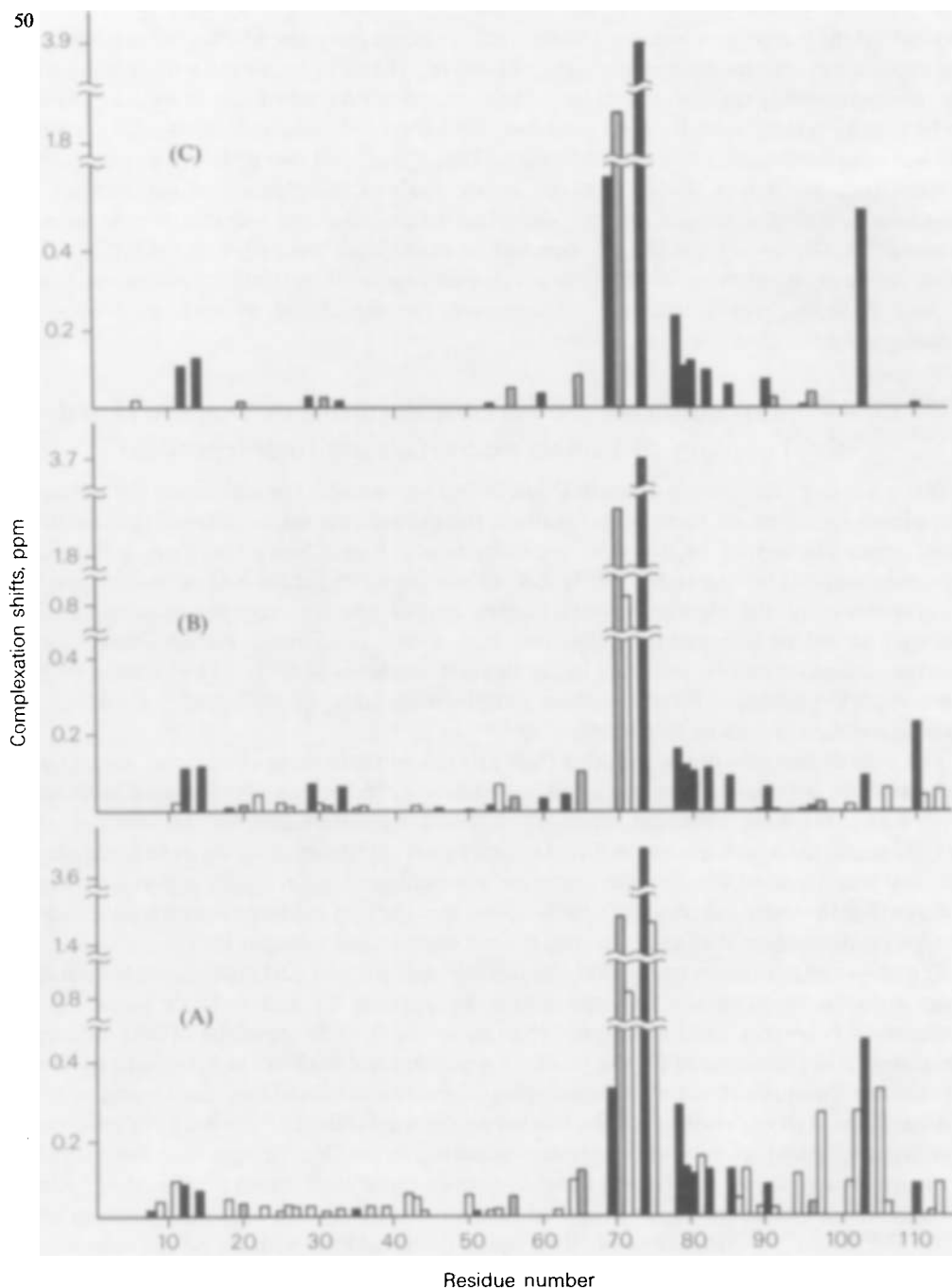


Fig. 2.8 Comparison of the complexation shifts of the SSI backbone carbonyl ^{13}C resonances with subtilisin BPM' (A), subtilisin Carlsberg (B) and subtilisin via *Bacillus amylosacchariticus* (C) (pH 7.3, 40°C). The dark and dotted bars show down- and up-field shifts by complexation, respectively. Only the bars of the residues which were examined in common with these three cases are shown as dark and dotted. (Source: Miyake, Y., Thesis, Tokyo Metropolitan University, (1992))

proteinases in general).

These three mutants showed very similar chemical shift change profiles, indicating that the local structural changes initiated on the α_2 -helix seem to propagate to the other flexible segments in domain II through common pathways. Fairly large chemical shift changes observed for the surface residues around the scissile bond indicate the existence of interdependent pathways between these two regions, which are in fact identical to what we called the flexible segments found by the subtilisin binding (see Fig. 2.7).

The results strongly suggest that there are numerous conformational substrates having very close free energy levels to each other for the flexible segments in domain II. Even with a minor structural perturbation such as a single amino acid replacement may shift one conformational substate(s) to the other. Note that such a facile change in the conformational substates distribution may not necessarily mean that the segments are in dynamically flexible states. On binding with a specific type of proteinase, one portion of the conformational substates is being stabilized more than the others and is thus more distributed. It may be argued that a proteinase can act as a template to control the distribution of the conformational substates of the flexible segments so as to stabilize the overall free energy of the complex. We have thus compared the carbonyl carbon chemical shifts of the WT and the mutant SSIs in their free and subtilisin BPN' complexed states. Very excitingly, large chemical shift differences observed for the mutants and WT SSI in their free form were almost lost in the subtilisin complexes. This may be another piece of evidence indicating that the flexible segments of SSI are responsible for recognizing the target proteinase.

We have explained so far several results indicating that the conformational flexible segments of domain II are responsible for its wide inhibitory activity. This hypothesis can be better supported by comparing the NMR chemical shifts among the complexes with more different structures, such as trypsins and subtilisins. Unfortunately, as the P₁ residue of WT SSI is Met, trypsin-like enzymes cannot be efficiently inhibited by WT SSI and such comparison was not possible. We have succeeded in making such comparisons, however, by using M⁷³R SSI which inhibits bovine trypsin as well without losing the inhibitory activity toward subtilisins.³³⁾

The conformation of M⁷³R SSI in the free state was nearly identical to that of the WT SSI since there was almost no chemical shift difference between these two inhibitors. The carbonyl carbon chemical shift profile of the M⁷³R SSI-subtilisin BPN' complex was also very similar to that of the WT SSI-subtilisin BPN' complex, showing that the conformations of these complexes are also nearly the same. The shift profile of the M⁷³R SSI-bovine trypsin, however, was quite different from those of the subtilisin complexes. This can be taken as a good indication that the conformational flexible segments are responsible for the wide inhibitory activity of SSI.

2.7 Summary and Conclusion

We have shown above, although in a very qualitative way, that various NMR evidence seems to indicate that two peculiar features of SSI, namely dimeric structure and wide inhibitory activity, which are rather unusual for protein proteinase inhibitors, are actually closely related. The molecular assembly of the SSI subunit can be conceptually divided into two subdomains having a similar topology: domain I, which is located at the N-

terminal side of the SSI subunit, and domain II at the C-terminal side. The proteinase binding site, which is the peptide segment around the scissile bond, is in domain II. SSI effectively inhibits the target proteinases by forming a stable Michaelis complex having an intact planar scissile bond. The excessive stability of the Michaelis complex seems to be closely related to the proteinase-induced global conformational changes found for the flexible segments in domain II. We found that exactly the same peptide segments in domain II change their conformation in a similar way either by hydrolyzing the scissile bond or by site-specific mutations for the residues in the α_2 -helix. These conformational changes, or even the cleaved scissile bond, were mostly restored in the subtilisin complex.

The flexible nature of the peptide segments in domain II, which seems to be stabilized against proteolysis by the rather rigid structural domain I and two β - sheets of domain II, through the subunit-subunit association. The existence of such flexible segments must be the origin of the rather broad inhibitory spectrum of SSI. It was not so simple to answer to the important question about the real biological roles of the SSI family proteins in *Streptomyces*. We just wanted to point out at this stage that the SSI family proteins are secreted in enormous amounts from the microbial cells, and thus are probably not functional intracellular regulatory proteins. As most *Streptomyces* have been found to produce SSI family proteins, they may be related to the defense mechanisms of these microorganisms. A wide inhibitory spectrum for the unknown exogenous proteinases, at the sacrifice of highly specific activity for a single proteinase, should be more versatile for that function.

Acknowledgments

The main part of this work is the graduate work of Dr. Yoko Miyake and included in her thesis (Tokyo Metropolitan University, 1992). The manuscripts will be submitted shortly for publication. We would like to thank those students who have contributed to this project for years, conducting painstaking work for the assignments by double labeling methods. We also wish to thank Prof. Y. Mitsui, Prof. K. Miura, Prof. Tonomura, and Prof. Momose for their collaboration and personal communication on their unpublished works. Last, but not least, MK would like to dedicate this article to the late professor M. Tokushige who encouraged him from the very early stages of this project. The work was supported in part by Special Coordinate Funds of the Science and Technology Agency, an NEDO Bioscience Grant, and also by various funds from the Ministry of Education.

REFERENCES

1. S. Murao, S. Sato, *Agr. Biol. Chem.*, **36**, 160 (1972); S. Murao, S. Sato, N. Muto, *Agr. Biol. Chem.*, **36**, 1737 (1972).
2. In: *Protein Protease Inhibitor—The Case of Streptomyces Subtilisin Inhibitor (SSI)*, (K. Hiromi, K. Akasaka, Y. Mitsui, B. Tonomura, S. Murao, eds.) Elsevier, Amsterdam, (1985).
3. Y. Mitsui, Y. Satow, Y. Watanabe, Y. Iitaka, *J. Mol. Biol.*, **131**, 697 (1979).
4. S. Ishii, Y. Mitsui, *Nippon Nogeikagaku Kaishi*, **62**, 1810 (1988) (in Japanese).
5. S. Taguchi, I. Kumagai, A. Nakayama, A. Suzuki, K. Miura, *Biotechnology*, **7**, 1063 (1991); S. Taguchi, S. Kojima, I. Kumagai, H. Ogawara, K. Niura, H. Momose, *FEMS Microbiology Letters*, **99**, 293 (1992).
6. K. Wüthrich, *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York (1986).

7. M. Kainosho, T. Tsuji, *Biochemistry*, **21**, 6273 (1982).
8. M. Kainosho, H. Nagao, Y. Imamura, K. Uchida, N. Tomonaga, Y. Nakamura, T. Tsuji, *J. Mol. Struct.*, **126**, 549 (1985); M. Kainosho, H. Nagao, T. Tsuji, *Biochemistry*, **26**, 1068 (1987).
9. M. Kainosho, Y. Miyake, *Tanpakushitsu III, koujikouzou*, Sinseikagaku-Jikken-Koza, 151–190, Tokyo Kagaku Dojin, Tokyo (1989) (in Japanese).
10. M. Kainosho, *NMR*, Seikagaku-Jikken-Koza V, 347, Maruzen, Tokyo (1990) (in Japanese).
11. O. Jardetzky, G.C.K. Roberts, *NMR in Molecular Biology*, Academic Press, London (1981).
12. M. Laskowski, Jr., R.W. Sealock, in: *The Enzymes*, (P.D. Boyer, ed.), Vol. 3, pp. 375–473, Academic Press, New York (1971).
13. S. Hirono, K.T. Nakamura, Y. Iitaka, Y. Mitsui, *J. Mol. Biol.*, **131**, 855 (1979); Y. Takeuchi, Y. Satow, K. Nakamura, Y. Mitsui, *J. Mol. Biol.*, **221**, 309 and references cited therein (1991).
14. A. Ruhlmann, D. Kukla, K. Schwager, K. Bartels, R. Huber, *J. Mol. Biol.*, **77**, 417 (1973).
15. R.M. Sweet, H.T. Wright, J. Janin, C.H. Chothia, D.M. Blow, *Biochemistry*, **13**, 4212 (1974).
16. M. Fujinaga, R.J. Read, A. Sielecki, W. Ardelt, M. Laskowski, Jr., M.N.G. James, *Proc. Natl. Acad. Sci. USA*, **79**, 4868 (1982).
17. M. Marquart, J. Walter, J. Deisenhofer, W. Bode, R. Huber, *Acta Cryst.*, **B39**, 480 (1983).
18. R.J. Read, M.N.G. James, in: *Proteinase Inhibitors* (A.J. Barret, G. Salvesen, eds.), pp. 301–336, Elsevier, Amsterdam; and references cited therein (1986).
19. A. Allerhand, *Methods in Enzymology*, **61**, 458 (1979).
20. M. Kainosho, T. Tsuji, J. Akagawa, Y. Mitsui, *Tanpakushitsu-Kakusan-Koso*, **27**, 1556 (1982) (in Japanese).
21. M. Kainosho, Y. Miyake, *Nippon Nogeikagaku Kaishi*, **62**, 1822 (1988) (in Japanese).
22. G. Binsch, J.B. Lambert, B.W. Roberts, J.D. Roberts, *J. Amer. Chem. Soc.*, **86**, 5564 (1964).
23. J.W. Schulman, T. Venanzi, *J. Amer. Chem. Soc.*, **98**, 4701; *ibid.*, 6739 (1976).
24. G.C. Levy, R.L. Lichter, *¹⁵N Nuclear Magnetic Resonance Spectroscopy*, John Wiley & Sons, New York (1979).
25. R.J. Read, M. Fujinaga, A.R. Sielecki, M.N.G. James, *Biochemistry*, **22**, 4420 (1983).
26. B. Tonomura, S. Matsumori, H. Iwanari, T. Toyono, K. Hiromi, *Seikagaku*, **55**, 827 (1983) (in Japanese).
27. H. Iwanari, *Masters Thesis*, (Kyoto Univ., 1984) (in Japanese).
28. R. Huber, D. Kukula, W. Bode, P. Schwager, K. Batels, J. Deisenhofer, W. Steigmann, *J. Mol. Biol.*, **89**, 73 (1974).
29. W. Bode, E. Papamokos, D. Musil, U. Seemueller, H. Frits, *The EMBO Journal*, **5**, 813 (1986).
30. D. Tsuru, *Abstract of Kosokin Symposium*, pp. 11–15 (1973) (in Japanese).
31. S. Obata, S. Taguchi, I. Kumagai, K. Miura, *J. Biochem.*, **105**, 372 (1989).
32. T. Nirasawa, *Masters Thesis* (Tokyo Metropolitan Univ., 1990) (in Japanese).
33. S. Kojima, S. Obata, I. Kumagai, and K. Miura, *Biotechnology*, **8**, 449 (1990); S. Kojima, Y. Nishiyama, I. Kumagai, K. Miura, *J. Biochem.*, **109**, 377 (1991).

3

Subunit Assembly of Oligomeric Enzymes

Noboru Yumoto

Organic Material Department, Government Industrial Research Institute, Ikeda, Osaka 563, Japan

3.1 Introduction

The term “quaternary structure” was proposed to supplement the terms “primary, secondary, and tertiary structure,”¹⁾ and refers to the spatial arrangement of noncovalently linked polypeptides which we shall call “subunits.” Generally, a subunit is defined as a tertiary structural unit composed of a single polypeptide, but the definition is somewhat ambiguous. In aspartate transcarbamoylase from *Escherichia coli*, for example, tertiary structural units composed of three catalytic polypeptide chains are called catalytic subunits and those composed of two regulatory chains are called regulatory subunits.²⁾

Since it was first recognized by T. Svedberg and coworkers 50 years ago that proteins might be constituted of subunits, a large number of functionally important proteins have been revealed to be oligomeric. For example, 19 of 21 enzymes in the glycolytic pathway and the citric acid cycle are oligomeric, and almost all the enzymes described in Chapters 5–12 of this volume also fit this description. However, the advantages of subunit assembly in oligomeric proteins are not yet fully understood.

In this review, we shall focus on new approaches with particular emphasis on the subunit assembly in the cell after a brief summary of the functional aspects of quaternary structure, because the comprehensive review of the quaternary structure written by Klotz, Darnall and Langerman³⁾ in 1975 is essentially valid.

3.2 Functional Roles of Subunit Assembly and Disassembly

The functional activities of oligomeric proteins display a complexity which is seldom seen in monomeric proteins, and are controlled and mediated via subunit-subunit interactions on the quaternary level. However, the individual mechanism by which the activity of a specific oligomeric protein is controlled varies greatly. The late Professor M. Tokushige, the original editor of this book, classified functional roles of subunit assembly and disassembly into six categories (Fig. 3.1).⁴⁾ We shall discuss functional aspects of quaternary structure by giving typical examples of each category.

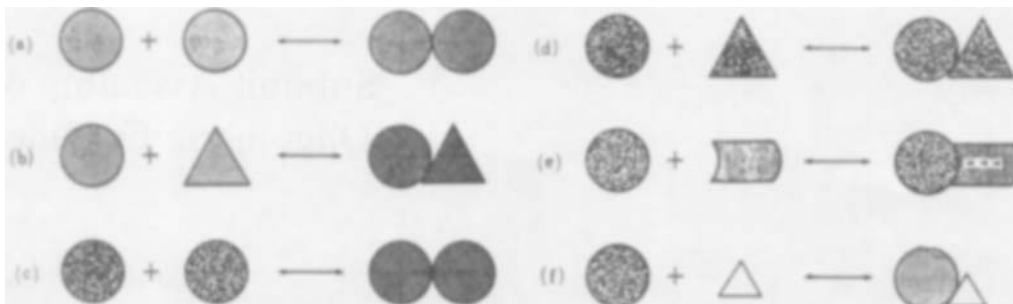


Fig. 3.1 Classification of functional roles of subunit assembly and disassembly.⁴⁾ a) The assembly of identical subunits is essential for the catalytic activity. b) The assembly of nonidentical subunits is essential for the catalytic activity. c) The assembly of active subunits enhances the catalytic activity. d) Sequential metabolic reactions are efficiently catalyzed by the assembly of subunits. e) The assembly of subunits is required for the expression of regulatory properties. f) The assembly of nonidentical subunits diminishes the catalytic activity. See text for details. (Reproduced with permission from S. Tokushige, *Kagaku Zokan*, 103, 41 (1984), in Japanese).

a. The assembly of identical subunits is essential for the catalytic activity.

Monomeric intermediates during the assembly of oligomeric proteins are usually inactive and the formation of native quaternary structures are often a prerequisite for catalytic activity. One clear reason for this is that the active sites of some enzymes are located at the interface between subunits and are formed by amino acid residues from different subunits. Such examples are aspartate transcarbamoylase from *E. coli*⁵⁾ and ribulose biphosphate carboxylase/oxygenase from *Rhodospirillum rubrum*.⁶⁾

b. The assembly of nonidentical subunits is essential for the catalytic activity.

H⁺-ATPase in membranes of mitochondria, chloroplasts and bacteria catalyzes ATP synthesis coupled with electron transport.⁷⁾ The catalytic portion of the enzyme F₁ is composed of five different subunits, denoted α through ϵ in order of decreasing molecular weight. The ATPase activity can be reconstituted from isolated α , β and γ subunits, each of which is devoid of catalytic activity in the separated state.⁷⁾ Interestingly, the γ subunit is indispensable for the functional reconstitution, although the catalytic site is suggested to be located on the β subunit or at the interface of the α and β subunits.⁷⁾

c. The assembly of active subunits enhances the catalytic activity.

Yeast phosphoglycerate mutase is composed of four identical subunits. Renaturation experiments of the denatured enzyme revealed that monomeric intermediates formed during the reconstitution of tetrameric structure are partially active, but that the assembly of subunits enhances the catalytic activity about three fold.⁸⁾

d. Sequential metabolic reactions are efficiently catalyzed by the assembly of subunits.

There are multienzyme complexes that efficiently catalyze sequential reactions in some metabolic pathways. The overall rate of a sequential reaction is greatly enhanced by the assembly of subunits possessing sequential metabolic activities, because the metabolites are transferred directly from one active site to another without diffusing in the solution.⁹⁾ Typical examples are the mammalian pyruvate dehydrogenase complex (three catalytic subunits)¹⁰⁾ and the bacterial tryptophan synthase (two catalytic subunits).¹¹⁾

e. The assembly of subunits is required for the expression of regulatory properties.

Some oligomeric enzymes acquire sensitivity to various effectors upon the assembly of

catalytic and regulatory subunits. For example, aspartate transcarbamoylase loses its sensitivity to CTP without loss of enzymatic activity, when the catalytic subunits are separated from the regulatory subunits.¹²⁾

f. The assembly of nonidentical subunits diminishes the catalytic activity.

cAMP-dependent protein kinase consists of 49-kDa regulatory subunits which can bind cAMP and 38-kDa catalytic subunits.¹³⁾ In the absence of cAMP, the two kinds of subunits form a heterotetrameric complex that is enzymatically inactive. The binding of two molecules of cAMP to each of the regulatory subunits leads to the dissociation of subunits, and then active catalytic subunits are liberated.

As demonstrated by the examples above, recent studies aimed at understanding the relationship between the quaternary structure and the physiological activity of oligomeric proteins have produced a considerable amount of information which explains many aspects of this phenomenon. In the process, a number of experimental methods have been used. The following sections provide an overview of new approaches which are generally applicable for elucidating the structure-function relationship of oligomeric proteins.

3.3 Determination of Composition and Geometric Arrangements of Subunits

To establish the quaternary structure of an oligomeric protein, stoichiometric constitution and geometric arrangements of subunits must be determined. From the difference in molecular weight between native and denatured states of a protein, the number of subunits is usually estimated. However, errors in the determination of molecular weight of a large protein, which are inevitable due to the shape of protein, surface charges, etc., sometimes lead to a wrong conclusion. Therefore, other methods such as cross-linking of subunits should also be used. Cross-linking experiments are also useful in determining the spatial arrangement of subunits as exemplified in the study on ribosomes.¹⁴⁾

Because the molecular weight of oligomeric proteins is generally large, their X-ray crystallographic analysis is laborious and the number of available high resolution data is limited. Therefore, alternative approaches have been developed to reveal the geometric arrangements of oligomeric proteins. One is the cross-linking analysis described above and the other is electron microscopy. Negatively stained or rotary shadowed molecules are analyzed by conventional transmission electron microscopy (TEM), whereas unstained "naked" molecules can be visualized by scanning transmission electron microscopy (STEM), which can also provide information about the molecular mass of visually observed molecules. For example, the dimeric butterfly-like structure of α_2 -macroglobulin was revealed by TEM, and its native molecular mass was determined to be 354 kDa by STEM.¹⁵⁾ By the aid of computer processing of TEM images, the cube-like structure of creatine kinase¹⁶⁾ and the unique trifoliolate structure of proline- β -naphthylamidase¹⁷⁾ (Fig. 3.2) were clearly observed. For evaluating a structural model proposed by electron microscopy, data obtained by small-angle X-ray scattering, for example the radius of gyration and the largest distance within the molecule, have been revealed to be effective as shown in the studies on the structure of hagfish protease.¹⁸⁾



Fig. 3.2 Unique trifoliolate structure of proline- β -naphthylamidase.¹⁷⁾ Electron micrographic images of negatively stained molecules are rotationary averaged after superimposition of images each with 120° rotation from the others. Scale bar, 3 nm. (Reproduced with permission from K. Takahashi, *FEBS Lett.*, **280**, 298 (1991)).

3.4 Determination of Subunit Contact Regions and Strength of Subunit Interactions

From the high-resolution crystal structures of 23 oligomeric proteins, generalizations concerning the subunit interfaces have been attempted.¹⁹⁾ The amino acid residues which compose the subunit contact sites are usually hydrophobic. A notable exception is the Arg residue, which is the second most common residue found in these regions after Leu. These Arg residues are postulated to be involved in hydrogen bond formation between subunits.¹⁹⁾ Contact areas vary from 670 Å² per subunit in superoxide dismutase, which corresponds to 9% of the subunit surface, to 10570 Å² in catalase, which is 40% of the subunit surface.¹⁹⁾ Although it is obvious that X-ray crystallography is necessary for a detailed discussion of subunit contact regions, general questions can be resolved by other approaches.

3.4.1 Identification of Subunit Contact Regions

Synthetic peptides corresponding to various portions of a protein sequence have been used as powerful tools with which to investigate interactions between subunits. For example, two contact sites on the α -subunit of human lutropin have been identified by this method.²⁰⁾ Peptides corresponding to the entire sequence of the α -subunit of lutropin were incubated in the presence of separated α - and β -subunits, and the reconstitution of heterodimeric structure was monitored by difference absorption spectroscopy, reverse phase HPLC and gel filtration (Table 3.1). After a 24-hour incubation, peptide 33–58

TABLE 3.1 Reconstitution of heterodimeric structure of lutropin in the presence of peptides corresponding to the entire sequence of the α -subunit²⁰⁾

peptide	time (h)	difference spectroscopy (%)	rpHPLC (%)	gel filtration (%)
none	24	100	100	100
1-15	24	82	82	86
11-27	24	91	76	75
22-39	24	92	84	80
33-58	24	0	11	0
	48	39	38	56
	72	76	67	ND [†]
51-65	24	61	29	10
	48	94	65	87
61-78	24	90	71	63
	48	92	74	90
73-92	24	88	86	88

† ND=not determined.

(Reproduced with permission from T. T. Anderson, *Biochemistry*, **30**, 1860 (1991)).

allowed only 0–11% reconstitution, and peptide 51–65 allowed 10–60% reconstitution, while the inhibitory effects of other peptides were very weak, suggesting that these two peptide regions are specifically involved in inter-subunit interactions. Similar experiments revealed that three separate regions in the β -subunit of follicle-stimulating hormone are involved in the interaction with α -subunit.²¹⁾

In other studies on the domain structure of oligomeric proteins, limited proteolysis, deletions, and site-directed mutagenesis have also successfully provided information about subunit contact regions. Functional domains containing subunit interaction sites of cAMP-dependent protein kinase²²⁾ and carbamoyl phosphate synthetase²³⁾ have been revealed by such experiments.

3.4.2 Evaluation of the Strength of Interactions Between Subunits

Quantitative evaluation of the strength of interactions between subunits is often difficult because of the multiplicity of the interactions. However, subunit exchange experiments can provide information about the relative interaction strength. In the study on aspartate transcarbamoylase,^{24,25)} the dissociation rates of mutant subunits were determined by subunit exchange leading to hybrid molecules. The formation of hybrid molecules composed of native and succinylated catalytic subunits could be easily monitored by polyacrylamide gel electrophoresis, and it was suggested that the trimeric structure of a mutant in which Gly-128 was replaced by Asp was much more stable than that of the wild-type.^{24,25)}

The first direct measurement of the strength of subunit interactions was attained in experiments on an actin filament.²⁶⁾ A single actin filament (about 7 nm in diameter) labeled with fluorescent phalloidin was caught by two glass microneedles under a fluorescence microscope, as shown in Fig. 3.3(A). The tensile strength of the filament was determined by pulling the stiff needle until the filament broke, and measuring the maximum bending of the flexible needle just before the filament broke (Fig. 3.3(A,B)). Using this method, the strength of subunit interaction in an actin filament was determined to be

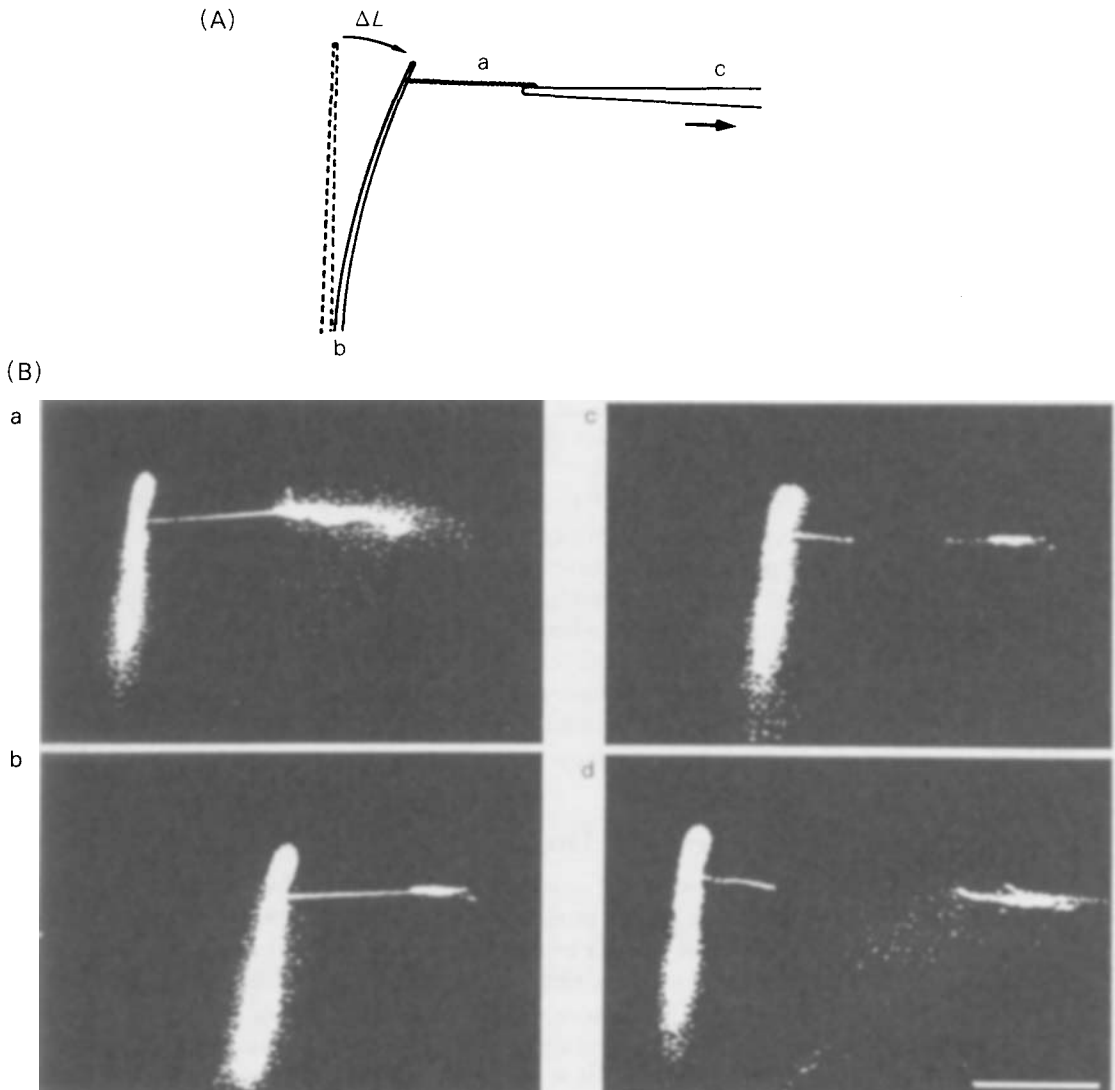


Fig. 3.3 Direct measurement of subunit-subunit bond strength in a single actin filament.²⁶⁾ (A) The principle of measuring tensile strength of a single actin filament. One end of a single filament (a) floating in solution was bound to a flexible needle (b) and the other was bound to a stiff needle (c). The tensile strength was determined from the displacement (ΔL) of the flexible needle required to break the filament. (B) Fluorescent images of a single actin filament labeled with fluorescent phalloidin. Both ends of the filament were attached to the two microneedles before stretching (a). The stiff needle was moved to the right just before the filament broke (b). After the filament was broken (c), the flexible needle returned to the original position (d). Bar, 20 μm . (Reproduced with permission from M.-J. Gething and J. Sambrook, *Nature*, 355, 35 (1992)).

108 pN, irrespective of filament length.²⁶⁾ This value was nearly ten times smaller than the value estimated theoretically from the equilibrium free energy of actin-actin bonds in the filament.²⁶⁾

3.5 Experimental Approaches to Reveal the Relationship Between Oligomeric Structure and Biological Activity

3.5.1 Analysis of Dissociation and Association Processes in Oligomeric Proteins

In oligomeric proteins composed of nonidentical subunits, it is possible to determine the functional role of each subunit, if dissociation and isolation of subunits are attained in their native states. Examples are tryptophan synthase, lactose synthetase, and aspartate transcarbamoylase.²⁵⁾

For many oligomeric proteins composed of identical subunits, it is not easy to reveal the functional role of each subunit. In most cases it is impossible to isolate monomers without disrupting the tertiary structure. Therefore, to determine whether monomers are active or not, a kinetic analysis of the renaturation process of unfolded monomers is often performed.²⁷⁾ In the case of creatine kinase, when the enzyme denatured in guanidine hydrochloride was renatured upon removal of the denaturant, the recovery of enzymatic activity preceded the formation of the native dimeric form, evidence that monomeric subunits of the enzyme are active.²⁸⁾ Monomeric intermediates found during the reconstitution of several other enzymes such as yeast phosphoglycerate mutase and rabbit muscle aldolase are also active.²⁷⁾ In most cases, however, the intermediates are inactive and native quaternary structures are a prerequisite to catalytic activity. As an example, the reactivation and reconstitution kinetics of tetrameric *E. coli* aspartase²⁹⁾ are shown in Fig. 3.4. When the enzyme denatured in 4 M guanidine hydrochloride was renatured by dilution, the yield of reactivation was dependent on the renaturation temperature and little increase in the activity was observed at 4°C (Fig. 3.4(A)). The enzyme renatured at temperatures from 7.5° to 28°C can be further reactivated by a second temperature shift to 30°C. HPLC analysis using a temperature-controlled gel filtration column revealed that the molecular assembly of the subunits was also dependent on the renaturation temperature (Fig. 3.4(B)). The elution profile in terms of enzyme activity suggested that only the native-like molecule was active and other species of smaller Stokes radii were inactive. The inactive species was converted to a native-like molecule by incubation at 30°C. These results indicate that tetramer formation is an essential prerequisite for expression of the aspartase activity, and that intermediates such as monomers, dimers, and trimers, in the assembly process are inactive. In a similar experiment on homoserine dehydrogenase from *Scaccharomyces cerevisiae*, it was indicated that formation of the native dimeric structure is required for activity expression.³⁰⁾

3.5.2 Hybridization Experiments

Hybridization of various kinds of subunits has long been used as a powerful approach in the study of structure and function of oligomeric proteins.²⁵⁾ The hybridization technique was originally used in the study of isozymes to determine the subunit composition. Because variants of a given oligomeric protein are not usually available, altered subunits are

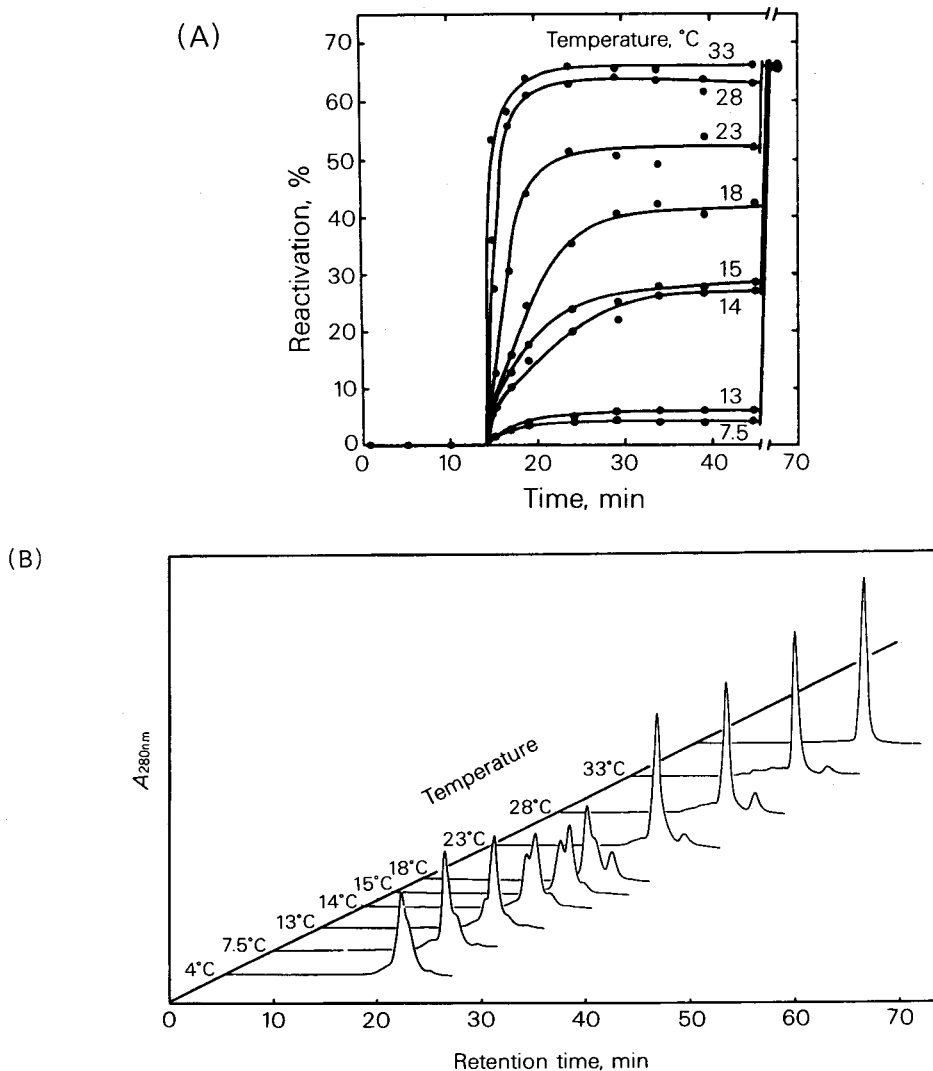


Fig. 3.4 Temperature-dependent reconstitution of tetrameric *E. coli* aspartase.²⁹⁾ A: Reactivation of denatured aspartase. The enzyme denatured in 4 M guanidine-HCl was renatured at 4°C by dilution. After 14 min, the temperature of each preparation was shifted up as indicated in the figure. The temperature of each preparation was further shifted up to 30°C after 45 min. B: HPLC analysis of intermediates in the renaturation process. Aspartase renatured at 4°C was incubated for 15 min at the indicated temperatures. An aliquot of each preparation was applied to a TSKgel G3000SWXL column (7.5×300 mm) and eluted with a flow rate of 0.5 ml/min. The temperature of the sample in the sample loop, elution buffer and the column was maintained constant. (From *Physiol. Chem. Phys. Med. NMR*, 21, 222:226 (1989)).

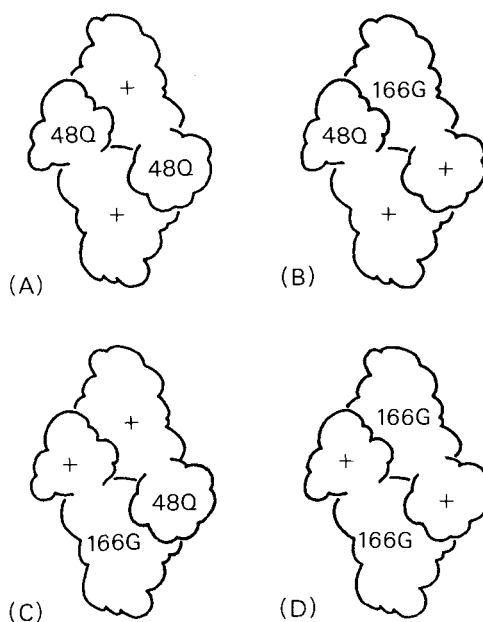


Fig. 3.5 *In vivo* hybridization of two site-directed mutants of ribulose biphosphate carboxylase/oxygenase from *R. rubrum*.⁶⁾ Two inactive mutants, in which Glu-48 was changed to Gln (A) and Lys-166 to Gly (D), were coexpressed in a cell. Heterodimers formed *in vivo* (B and C) were active, indicating that domains from each of the subunits interact to form a functional active site (+ adjacent to + in the figure). (Reproduced with permission from *J. Biol. Chem.*, **262**, 15328 (1987)).

produced by chemical or genetic modification to generate hybrid proteins. The formation of hybrid enzymes *in vitro* or *in vivo* revealed the intersubunit location of the active sites in aspartate transcarbamoylase from *E. coli*⁵⁾ and ribulose biphosphate carboxylase/oxygenase from *R. rubrum* (Fig. 3.5).⁶⁾ In hybridization experiments, however, a major problem is the inability to resolve the hybrid species from the parental forms. To generate separable hybrids, several methods have been reported, in which surface charges of subunits were changed by chemical modification or genetic engineering. The hybrids possessing different surface charges can easily be separated by ion-exchange chromatography or electrophoresis. This approach was used successfully with aspartate transcarbamoylase modified by succinylation or reversible tetrahydrophthaloylation,²⁵⁾ and with glutathione reductase, in which a penta-arginine "arm" followed by a unique protease (factor Xa) cleavage site was added to the N-terminus by insertion mutagenesis.³¹⁾

Another approach to separating hybrids is a method utilizing dye-ligand chromatography and affinity elution with specific ligands.^{32,33)} The anthranilate synthase-phosphoribosyl transferase complex from *Salmonella typhimurium* is a heterotetrameric (TrpE)₂(TrpD)₂ enzyme whose activities are feedback-regulated by L-tryptophan. The wild-type complex is bound tightly to Orange A agarose, and is eluted specifically by an addition of L-tryptophan as a result of an induced conformational change. The feedback

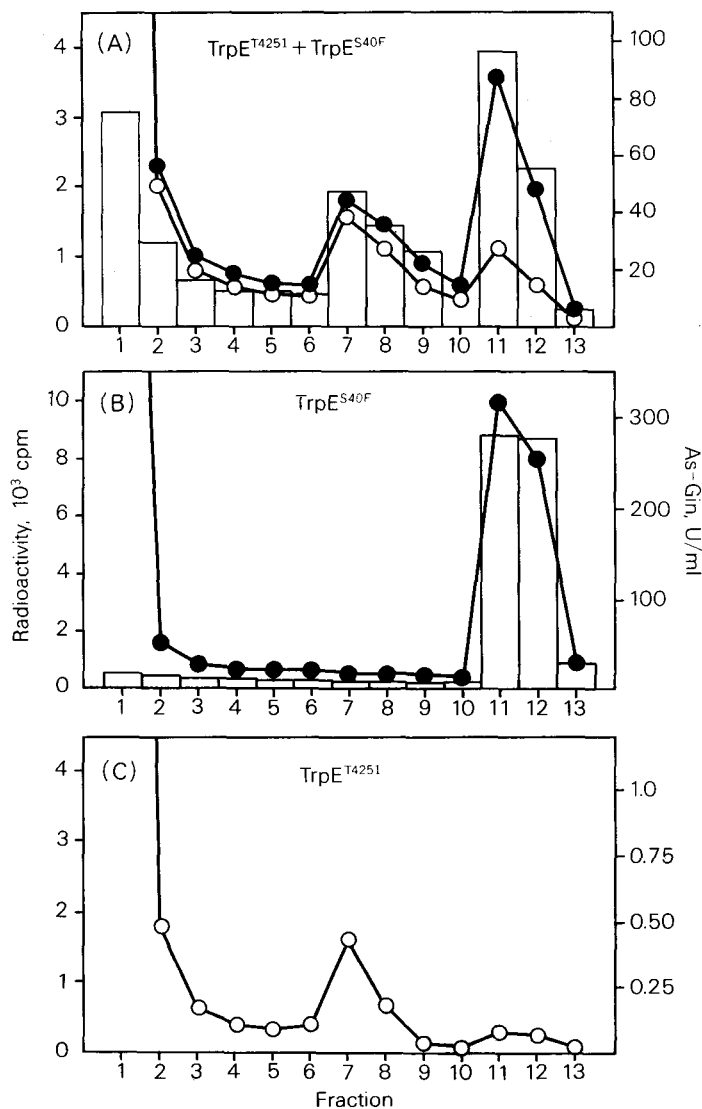


Fig. 3.6 Analysis of the hybrid of feedback-insensitive mutant ($\text{TrpE}^{\text{S40F}}$) and catalytically inactive mutant ($\text{TrpE}^{\text{T425I}}$) of anthranilate synthase complex with Orange A agarose chromatography.³³⁾ ^3H -labeled $\text{TrpE}^{\text{S40F}}$ and ^{14}C -labeled $\text{TrpE}^{\text{T425I}}$ were applied to an Orange A agarose column separately (B and C) or after *in vitro* hybridization of the two mutant subunits (A), and the elution was carried out as follows: Fractions 1 to 6, standard buffer; 7 to 10, standard buffer supplemented with L-Trp; 11 to 13, standard buffer supplemented with 1.5 M KCl. The peaks of ^3H (●) and ^{14}C (○) counts were coincided well the Trp-eluted peak of the enzymatic activity (open bars) in A, demonstrating the presence of hybrids containing both mutant subunits. (Reproduced with permission from R. Baerle, Science, **252**, 1846 (1991). Copyright 1991 by the AAAS).

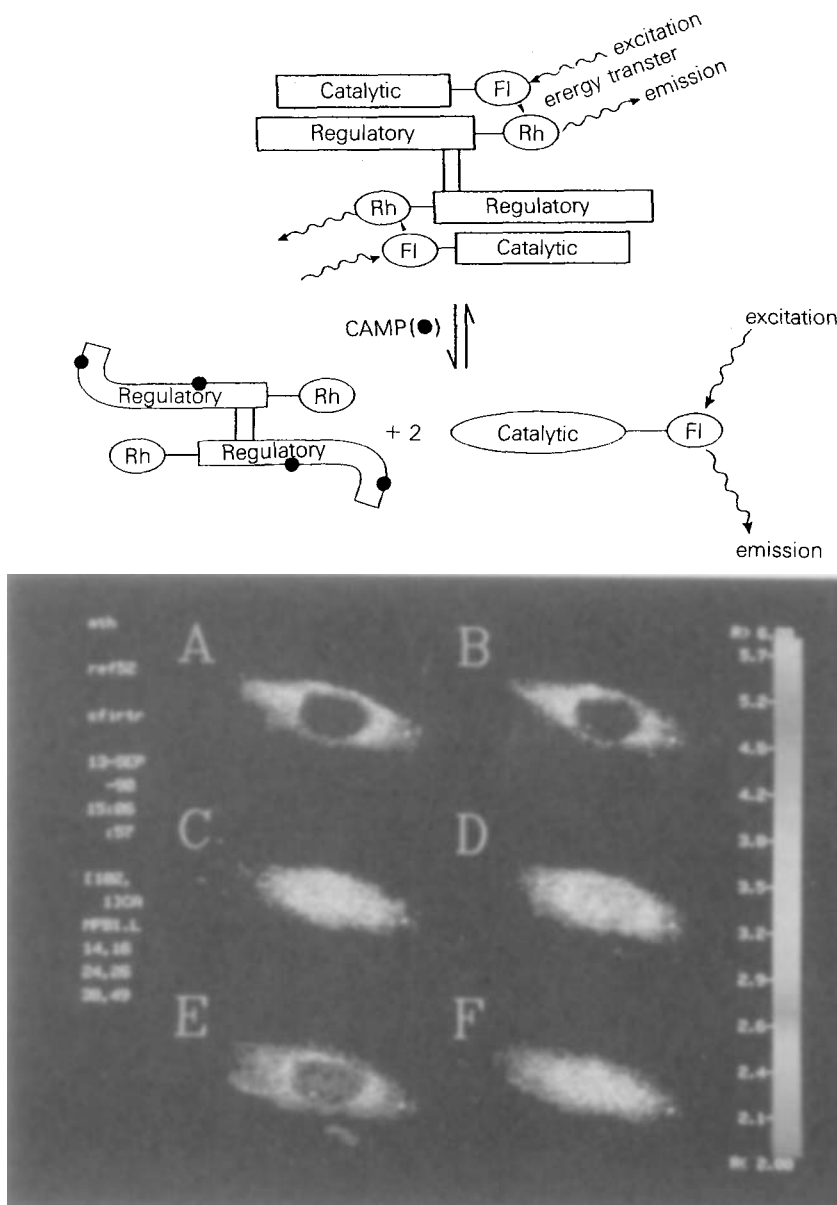


Fig. 3.7 A) Schematic diagram showing how to detect subunit assembly and disassembly of cAMP-dependent protein kinase using fluorescence energy transfer between two fluorophores. B) Subcellular localization of labeled kinase microinjected in a single REF-52 fibroblast.³⁷⁾ A): Fl, fluorescein; Rh, rhodamine. In the holoenzyme the two dyes are close enough so that excitation of the fluorescein donor at 480–495 nm results in detectable emission at 570–620 nm from the rhodamine as a result of fluorescence energy transfer. Upon addition of cAMP the dissociation of subunits eliminates the energy transfer, and excitation of the fluorescein gives brighter emission at 500–570 nm. The ratios of fluorescence emissions at 500–530 nm to >570 nm correlate well with the concentration of cAMP. B): Increasing ratios of 500–530 to >570 nm intensities were coded in pseudocolor hues ranging through the spectrum from blue to red in the original figure.³⁷⁾ A, before stimulation; B, 5.5 min after addition of prostaglandin E₁, which increases intracellular concentration of cAMP; C, 40.5 min after addition of the prostaglandin; D, 7 min after removal of the prostaglandin from the medium; E, 66 min after removal of the prostaglandin; F, 45 min after addition of forskolin, which also increases the cAMP concentration. (Reproduced with permission from R. Y. Tsien, *Nature*, **349**, 694:696 (1991)).

resistant $(\text{TrpE}^{\text{S40F}})_2 (\text{TrpD}^+)_2$ complex is bound also tightly to the gel, but is resistant to elution by L-tryptophan, while the catalytically inactive $(\text{TrpE}^{\text{T425I}})_2 (\text{TrpD}^+)_2$ complex is bound only poorly to the gel.³³⁾ When the two mutants were hybridized *in vitro*, one-third of the applied enzyme activity was eluted by L-tryptophan (Fig. 3.6), indicating that the $(\text{TrpE}^{\text{T425I}})(\text{TrpE}^{\text{S40F}})(\text{TrpD}^+)_2$ hybrid complex formed is not only catalytically active, but also responsive to the allosteric effector.³³⁾ The properties of the hybrid complex demonstrated that the binding of a single effector to one TrpE subunit is sufficient for the propagation of a conformational change that affects the active site of the companion subunit.³³⁾

3.5.3 Site-directed Mutagenesis at the Subunit Interface

In the case of oligomeric proteins in which subunit contact regions have been revealed by X-ray crystallography^{34,35)} or other methods described above,³⁶⁾ the equilibrium between oligomer and monomer can be changed by site-directed mutagenesis. For example, stable monomers of tyrosyl-tRNA synthetase were produced by a mutation of Phe-164 at the subunit interface to Asp, and it was revealed that the monomers are inactive and do not bind the substrate tyrosine.³⁴⁾ In the case of yeast triosephosphate isomerase, replacement of Asn-78 at the subunit interface did not cause dissociation of subunits under normal conditions.³⁵⁾ However, the stability of the enzyme was significantly lowered by the mutation, probably due to decreased subunit-subunit interaction.³⁵⁾

3.5.4 Analysis of Subunit Association and Dissociation in the Cell

Functions of oligomeric proteins are sometimes regulated by association and dissociation of subunits. A typical example is cAMP-dependent protein kinase as described above.¹³⁾ The subunit assembly and disassembly of the kinase can be visualized in a single living cell by microinjection of the enzyme in which the catalytic and regulatory subunits are each labeled with a different fluorescent dye capable of fluorescence energy transfer in the holoenzyme complex (Fig. 3.7(A)).³⁷⁾ cAMP changes the ratio of emission amplitudes at two wavelength bands, and it was found that the ratio correlates well to the concentration of cAMP.³⁷⁾ In a REF-52 fibroblast cell loaded with the labeled kinase, cytoplasm shows a low emission ratio (Fig. 3.7(B)), whereas the nucleus appears black because the holoenzyme is excluded. Elevation of intracellular cAMP with prostaglandin E_1 or forskolin first dissociates the enzyme and increases the emission ratio, whereupon the freed catalytic subunits gradually translocate to the nucleus. Upon removal of cAMP stimulation, the holoenzyme can be reconstituted in the cytoplasm, and the nucleus is cleared of catalytic subunits. This method will be very useful to analyze *in vivo* interactions between subunits of many important oligomeric proteins such as RNA polymerase³⁸⁾ and G-proteins,³⁹⁾ in which association and dissociation of specific subunits are required for transcription initiation and signal transduction, respectively.

3.6 Proteins Involved in Subunit Assembly *in vivo*

3.6.1 Molecular Chaperones

One of the greatest mysteries concerning the quaternary structure of oligomeric proteins has been the method by which these proteins fold to assume their active conformation in the

TABLE 3.2 Proteins that may be involved in subunit assembly in cells⁽⁴¹⁾

Organism/organelle	Enzymes			Chaperones		
	Protein family	PDI	Cyclophilin PPIase	FKBP PPIase	Hsp60 (Chaperonin-60)	Hsp70 (Stress-70) Hsp90 (Stress-90)
<i>E. coli</i>						
Cytosol	Thioredoxin		PPIase b		GroEL	DnaK HtpG (C62.5)
Periplasm			PPIase a (Rotamase)			
Yeast						
Cytosol			Cph1p (Cpr1p)	Fkb1p (Fkr1p) (Rbp1p)	Ssa1-4p	Hsp83 Hsc83
ER	PDI		yCyPB		Kar2p (BiP)	
	Eug1p				Ssc1p	
Mitochondria					Hsp60 (Mit4p)	
Drosophila						
Cytosol			CyP		Hsp68 Hsp70 Hsc1,2,4	Hsp83
ER	PDI		NinaA			
Mammals						
Cytosol			Cyclophilin (PPIase) (CyPA)	FKBP	Hsp70 (p73) Hsc70 (p72) (CUATPase) (Prp73)	Hsp90 (Hsp83) (Hsp87)
ER	PDI (ERp59) GSBP ERp72 ERp61		CyPB (rCyPLP)		BiP (Grp78)	Grp94 (ERp99) (endoplasmic)
Mitochondria						
Plants					Hsp60 (Hsp58)	Hsp70 (Grp75)
Cytosol						
ER	PDI				b70 (BiP)	
Chloroplasts					RuSBP	

(Reproduced with permission from M.-J. Gething and J. Sambrook, *Nature*, 355, 35 (1992)).

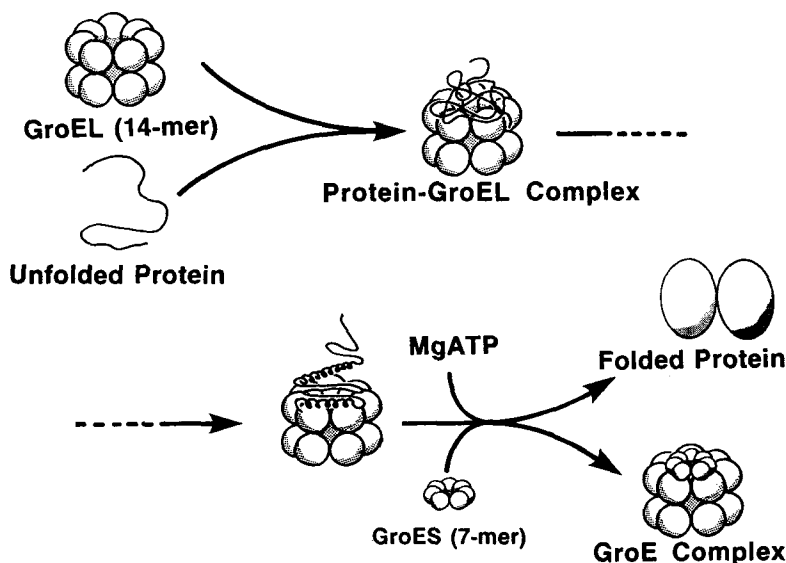


Fig. 3.8 Proposed molecular mechanism of GroEL-facilitated reconstitution of oligomeric proteins. Many aspects in this figure such as roles of GroES (chaperonin-10) and ATP remain unclear. See text for details.

cell. The classical work of C. B. Anfinsen established the hypothesis that all the information required for a polypeptide chain to assume its active conformation is contained in its amino acid sequence. This hypothesis holds true in many small monomeric proteins, while larger proteins, in particular oligomeric proteins have often failed to fold spontaneously into their active conformation from the unfolded state under *in vitro* conditions.⁴⁰⁾ The fact that folding *in vivo* is highly efficient indicates the existence of proteins mediating the correct folding of other proteins in the cell.⁴¹⁾ Recent studies have shown that such proteins do exist in the cell and actively assist the folding of other proteins. One class of such proteins is composed of the molecular chaperones which are members of the heat-shock protein family. Another class is composed of enzymes catalyzing specific steps in the folding process (Table 3.2). We first summarize studies on molecular chaperones, then review the enzymes involved in the folding.

Molecular chaperones are defined as a family of unrelated proteins found in all types of cells which mediate the correct assembly of other polypeptides, but are not components of the mature assembled structures.⁴²⁾ There are three major sequence-related families, chaperonin, hsp70 and hsp90 families among molecular chaperones (Table 3.2). Nucleoplasmins, signal recognition particle, trigger factor, Sec B protein, Pap D protein, and several other proteins are also included in this group.⁴²⁾ Among them, the chaperonin and hsp70 families are proposed to be involved in the assembly of oligomeric proteins (Fig. 3.8).⁴³⁾ In fact, evidence that the two families facilitate the *in vitro* assembly of oligomeric proteins is accumulating rapidly. For example, active dimeric ribulose biphosphate carboxylase can be reconstituted from an unfolded state by chaperonins,⁴⁴⁾ and heat-inactivated RNA polymerase can be reactivated by Dnak.⁴⁵⁾ Furthermore, *in vivo* assembly of hexadecameric ribulose biphosphate carboxylase was also facilitated by

co-expressed chaperonins.⁴⁶⁾ Interestingly, chaperonins have been found to be indispensable for their own subunit assembly *in vivo*.^{47,48)} However, many aspects of the effects of chaperonins remain unclear; for instance, the presence of chaperonin-10 is necessary for the reconstitution of some oligomeric proteins,^{44,46)} but not for others^{49,50)}; the hydrolysis of ATP is necessary in some cases,^{44,49)} but not in other cases.⁵⁰⁾ Very recently, the presence of chaperonin-10 and the hydrolysis of ATP have also been demonstrated to be unnecessary for reconstitution of tetrameric tryptophanase and dimeric enolase by Y. Kawata's group.^{51,52)} Previous refolding studies on *E. coli* tryptophanase revealed that the yield of refolding depends on the enzyme concentration and that the dependence is ascribable to a refolding intermediate with a high tendency toward irreversible aggregation.⁵³⁾ In the absence of nucleotides, *E. coli* GroEL (chaperonin-60) bound tightly to refolding intermediates of tryptophanase. Surprisingly, the addition of ADP as well as ATP and its non-hydrolyzable analogues resulted in an increase in refolding yield. This is the first time that chaperonin-facilitated refolding was observed in the presence of ADP. These results indicate that the GroEL-facilitated refolding of tryptophanase does not require energy from hydrolysis of ATP, rather that the conformational changes caused by the binding of specific nucleotides to GroEL are important.⁵¹⁾ Therefore, molecular chaperones seem to function not as catalysts of structure formation, but rather to recognize and stabilize intermediates during folding and assembly.

3.6.2 Other Proteins Involved in Subunit Assembly

Because polypeptide folding and subunit assembly should proceed in a properly coordinated manner for the successful formation of the quaternary structure of oligomeric proteins, it is likely that enzymes which catalyze rate-determining steps in protein folding are required for the *in vivo* assembly of some oligomeric proteins. At present, two such proteins have been found: one is protein disulfide isomerase (PDI)⁵⁴⁾ which catalyzes thiol/disulfide interchange reactions, and the other is peptidyl prolyl isomerase (PPIase),⁵⁵⁾ which catalyzes *cis-trans* isomerization of X-Pro peptide bonds (Table 3.2). PPIase has been a subject of intense study in recent years, because its activity is inhibited by clinically important immunosuppressive drugs. There are two families named after the drugs (Table 3.2); cyclophilin proteins bind cyclosporin A, whereas FK506-binding proteins (FKBP) bind FK506 and rapamycin.⁵⁶⁾

A third type of protein involved in subunit assembly is a dimerization cofactor of hepatocyte nuclear factor-1 α , which selectively stabilizes the dimeric structure of the transcription factor necessary for the binding to DNA.⁵⁷⁾ Although the transcription factor can dimerize on its own *in vitro*, *in vivo* assembly is suggested to be regulated by this tissue-specific 11-kDa cofactor protein.⁵⁷⁾

3.7 Conclusion

Quaternary structure has advantages in potential for flexibility in function and regulation. This seems to be the main reason why so many proteins are oligomeric. It is now generally accepted that oligomeric proteins play a vital role in the overall mechanism of a cell. However, studies on the structural mechanisms by which they fulfill that role were not developed until recently because of the complexity of each oligomeric protein. As described above, recent results have provided valuable insights on the structure-function

relationship of oligomeric proteins, and the new experimental methods guarantee a multitude of additional results in the near future. Studies on the *in vivo* characteristics of oligomeric proteins, exemplified by fluorescent label studies and the results on the role of specific proteins such as molecular chaperones, have been especially valuable in contributing to understanding the overall phenomenon of oligomeric proteins in the cell. Continued research along these lines will no doubt open a new era in this intriguing field.

Acknowledgments

I thank Dr. Yasushi Kawata (Tottori University) and Mr. Tomohiro Mizobata (Kyoto University) for valuable discussions and suggestions.

REFERENCES

1. J. D. Bernal, *Discuss. Faraday Soc.*, **25**, 7 (1958).
2. E. R. Kantrowitz, W. N. Lipscomb, *Trends Biochem. Sci.*, **15**, 53 (1990).
3. I. M. Klotz, D. W. Darnall, N. R. Langerman, in: *The Proteins* (3rd Ed.) Academic Press, Vol. I, p. 293 (1975).
4. M. Tokushige, *Kagakuzohkan (in Japanese)*, **103**, 29 (1984).
5. E. A. Robey, H. K. Schachman, *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 361 (1985).
6. F. W. Larimer, E. H. Lee, R. J. Mural, T. S. Soper, F. C. Hartman, *J. Biol. Chem.*, **262**, 15327 (1987).
7. M. Futai, T. Noumi, M. Maeda, *Annu. Rev. Biochem.*, **58**, 111 (1989); see also Chapter 11 in this volume.
8. R. Hermann, R. Jaenicke, N. C. Price, *Biochemistry*, **24**, 1817 (1985).
9. P. A. Srere, *Annu. Rev. Biochem.*, **56**, 89 (1987).
10. L. J. Reed, *Curr. Top. Cell. Reg.*, **18**, 95 (1981).
11. E. W. Miles, R. Bauerle, S. A. Ahmed, *Methods Enzymol.*, **142**, 398 (1987); see also Chapter 7 in this volume.
12. H. K. Schachman, *J. Biol. Chem.*, **263**, 18583 (1988).
13. S. S. Taylor, J. A. Buechler, W. Yonemoto, *Annu. Rev. Biochem.*, **59**, 971 (1990).
14. R. R. Traut, C. Casiano, N. Zecherle, in: *Protein Function: A Practical Approach*, IRL Press, Oxford, p. 101 (1989).
15. P. B. Armstrong, W. F. Mangel, J. S. Wall, J. F. Hainfield, K. E. Van Holde, A. Ikai, J. P. Quigley, *J. Biol. Chem.*, **266**, 2526 (1991).
16. T. Schnyder, A. Engel, A. Lustig, T. Wallimann, *J. Biol. Chem.*, **263**, 16954 (1988).
17. T. Takahashi, M. Nishigai, A. Ikai, K. Takahashi, *FEBS Lett.*, **280**, 297 (1991).
18. R. Österberg, B. Malmensten, A. Ikai, *Biochemistry*, **30**, 7873 (1991).
19. J. Janin, S. Miller, C. Chothia, *J. Mol. Biol.*, **204**, 155 (1988).
20. S. R. Krystek, Jr., J. A. Dias, T. T. Andersen, *Biochemistry*, **30**, 1858 (1991).
21. T. A. Santa-Coloma, L. E. Reichert, Jr., *J. Biol. Chem.*, **266**, 2759 (1991).
22. L. D. Saraswat, G. E. Ringheim, J. Bubis, S. S. Taylor, *J. Biol. Chem.*, **263**, 18241 (1988).
23. F. Guillo, S. D. Rubino, R. S. Markovitz, D. M. Kinney, C. J. Lusty, *Proc. Natl. Acad. Sci. U.S.A.*, **86**, 8304 (1989).
24. Y. R. Yang, H. K. Schachman, *Anal. Biochem.*, **163**, 188 (1987).
25. E. Eisenstein, H. K. Schachman, in: *Protein Function: A Practical Approach*, IRL Press, Oxford, p. 135 (1989).
26. A. Kishino, T. Yanagida, *Nature*, **334**, 74 (1988).
27. R. Jaenicke, R. Rudolph, *Methods Enzymol.*, **131**, 218 (1986).
28. S. H. Grossman, J. Pyle, R. J. Steiner, *Biochemistry*, **20**, 6122 (1981).
29. H. Imaishi, N. Yumoto, M. Tokushige, *Physiol. Chem. Phys. Med. NMR*, **21**, 221 (1989).
30. N. Yumoto, Y. Kawata, S. Noda, M. Tokushige, *Arch. Biochem. Biophys.*, **285**, 270 (1992).
32. H. Imaishi, N. Yumoto, M. Tokushige, *Biotechnol. Appl. Biochem.*, **12**, 196 (1990).
33. M. G. Caligiuri, R. Bauerle, *Science*, **252**, 1845 (1991).
34. D. H. Jones, A. J. McMillan, A. R. Fersht, G. Winter, *Biochemistry*, **24**, 5852 (1985).
35. J. I. Casal, T. J. Ahern, R. C. Davenport, G. A. Petsko, A. M. Klibanov, *Biochemistry*, **26**, 1258 (1987).
36. Y. J. Buechler, S. S. Taylor, *J. Biol. Chem.*, **266**, 3491 (1991).
37. S. R. Adams, A. T. Harootunian, Y. J. Buechler, S. S. Taylor, R. Y. Tsien, *Nature*, **349**, 694 (1991).

38. A. Ishihama, *Adv. Biophys.*, **14**, 1 (1981).
39. H. R. Bourne, D. A. Sanders, F. McCormick, *Nature*, **348**, 125 (1990).
40. R. Jaenicke, *Prog. Biophys. Mol. Biol.*, **49**, 117 (1987).
41. M. -J. Gething, J. Sambrook, *Nature*, **355**, 33 (1992).
42. R. J. Ellis, S. M. van der Vies, *Annu. Rev. Biochem.*, **60**, 321 (1991).
43. R. J. Ellis, *Science*, **250**, 954 (1990).
44. P. Goloubinoff, J. T. Christeller, A. A. Gatenby, G. H. Lorimer, *Nature*, **342**, 884 (1989).
45. D. Skowyra, C. Georgopoulos, M. Zylicz, *Cell*, **62**, 939 (1990).
46. P. Goloubinoff, A. A. Gatenby, G. H. Lorimer, *Nature*, **337**, 44 (1989).
47. N. M. Lissin, S. Y. Venyaminov, A. S. Girshovich, *Nature*, **348**, 339 (1990).
48. M. Y. Cheng, F. -U. Hartl, A. L. Horwich, *Nature*, **348**, 455 (1990).
49. J. Buchner, M. Schmidt, M. Fuchs, R. Jaenicke, R. Rudolph, F. X. Schmid, T. Kiefhaber, *Biochemistry*, **30**, 1586 (1991).
50. I. G. Badcoe, C. J. Smith, S. Wood, D. J. Halsall, J. J. Holbrook, P. Lund, A. R. Clarke, *Biochemistry*, **30**, 9195 (1991).
51. T. Mizobata, N. Yumoto, Y. Kawata, *J. Protein Chem.*, **11**, 413 (1992); T. Mizobata, Y. Akiyama, K. Ito, N. Yumoto, Y. Kawata, *J. Biol. Chem.*, **267**, 17773 (1992).
52. T. Kubo, T. Mizobata, Y. Kawata, *J. Protein Chem.*, **11**, 413 (1992); T. Kubo, T. Mizobata, Y. Kawata, *J. Biol. Chem.*, in press (1993).
53. Y. M. Torchinsky, Y. Kawata, See Chapter 9 in this volume.
54. R. B. Freedman, *Cell*, **57**, 1069 (1989).
55. G. Fischer, F. X. Schmid, *Biochemistry*, **29**, 2205 (1990).
56. S. L. Schreiber, *Science*, **251**, 283 (1991).
57. D. B. Mendel, P. A. Khavari, P. B. Conley, M. K. Graves, L. P. Hansen, A. Admon, G. R. Crabtree, *Science*, **254**, 1762 (1991).

4

Affinity Labeling of Nucleotide-Binding Sites

Mitsuo Tagaya, Katsuyuki Tanizawa, and Toshio Fukui

The Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Osaka 567, Japan

4.1 Introduction

The ultimate goal of enzymologists is to understand the mechanisms of enzyme reactions. A variety of physicochemical and biochemical techniques have been invented and used for identification of the amino acid residues involved in catalytic reactions and for estimation of their functional roles. Affinity labeling is one of the most useful techniques for these purposes since its appearance in the 1960s.^{1,2)} An affinity label consists of two functionally different groups: the affinity group that is structurally close to the substrates to facilitate the binding of the reagent to the active sites of enzymes and the reactive group that undertakes the covalent binding to amino acid residues in the regions where the label binds. Several criteria are applied to judge whether the modification reagent acts as an affinity label. For example, a hyperbolic curve (saturation kinetics) should be observed when the rate of inactivation (the pseudo-first order rate constant) is plotted against the concentration of the reagent. The enzyme should be more effectively inactivated by the affinity label than the simple modification reagent having the same reactive group as the affinity label, but without the affinity group. Inactivation should be protected by the substrates or competitive inhibitors.

Progress in peptide separation techniques, especially, the widespread distribution of high performance liquid chromatography equipments, has made it easy to identify the site(s) labeled by the affinity labeling reagent. In the last decade, a tremendous amount of information concerning the active sites of enzymes has been accumulated by the use of affinity labeling techniques with the identification of the labeled sites. One question that has been troubling enzymologists for a long time is whether the amino acid residue(s) labeled by the affinity label is essential for enzyme catalysis or not. Since the affinity label once bound to the substrate-binding site hinders the binding of substrates, it is usually impossible to evaluate the functional roles of the labeled amino acid residue(s). However, the appearance of site-directed mutagenesis dramatically changed the situation. Once the expression system for the cloned enzyme is constructed, it takes only a few months to clarify the role of the amino acid residue(s) modified by affinity label. Several misassignments from chemical modification studies have been corrected by the results of site-directed mutagenesis studies. However, this is not to deny the utility of affinity

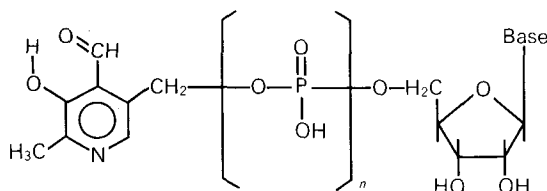


Fig. 4.1 General structure of nucleoside polyphosphopyridoxal (NP_n-PL).

labeling techniques. It is true that affinity labeling is one of the most convenient and fastest methods to determine the target amino acid residue(s) for site-directed mutagenesis studies. Affinity labeling will continue to be a good partner for site-directed mutagenesis in the next decade.

Several excellent reviews on affinity labeling are available, including general principles and practical considerations.¹⁻⁴⁾ This chapter focuses on affinity labeling of nucleotide-binding proteins, especially, with nucleotidylyl derivatives of pyridoxal 5'-phosphate (PLP) (Fig. 4.1) invented by our group⁵⁻¹⁷⁾ and the Cross group¹⁸⁻²¹⁾ independently. This type of reagent has several advantages as affinity labels for the nucleotide-binding sites of enzymes. They are easily soluble in water and stable at neutral pH if not irradiated by light. The reagents contain the whole structure of natural nucleotides, promising the binding of the reagents to the nucleotide-binding sites of proteins. The pyridoxal group, which may form a Schiff base with the ϵ -amino group of a lysyl residue, is located at the position corresponding to the β - or γ -phosphate of nucleotides. Because most nucleotide-binding enzymes catalyze the hydrolysis or transfer of the β - or γ -phosphate of nucleoside triphosphate, it is particularly important to reveal the structures of their triphosphate-binding loci for understanding the catalytic mechanisms. Nucleoside polyphosphopyridoxals (NP_n-PL) are expected to be useful for this purpose. Several other nucleotide affinity labels are also available. These include nucleotidylyl derivatives with an alkyl halide²²⁻²⁶⁾, benzylaldehyde²⁷⁻²⁹⁾, or fluorosulfonylbenzoyl^{30,31)} substituent at the 5'-position. Among these, fluorosulfonylbenzoyl nucleotides have been most frequently used for the identification of amino acids at the active sites. The fluorosulfonyl group is almost ideal as a reactive group because it is capable of forming covalent bonds with several amino acids including tyrosyl, lysyl, histidyl, cysteinyl, and seryl residues in proteins. The position of the fluorosulfonylbenzoyl group, if the reagent takes an extended conformation, is equivalent to the β - or γ -phosphate of nucleotide triphosphate. Although this type of reagent lacks the phosphoryl group, the carbonyl group adjacent to the 5'-position is structurally similar to the first phosphoryl group of nucleotide triphosphates. Since the synthesis of fluorosulfonylbenzoyl adenosine (FSBA) by Colman and her collaborators in 1975,³²⁾ many nucleotide-binding proteins have been labeled by this type of reagent, and the labeled sites have been identified.^{30,31)} Among these, several proteins were also labeled by NP_n-PL. Because the two types of affinity reagents have their reactive groups in a similar position, it may be valuable to compare the sites labeled by the two reagents when conducting modification studies on a particular nucleotide-binding protein in which one is interested.

4.2 Nucleoside Polyphosphopyridoxals (NP_n-PL)

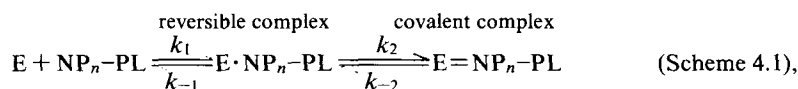
PLP has been used as a chemical modification reagent for lysyl residues of proteins.³³⁾ The specificity of modification with this reagent seems to be undertaken by its negatively charged phosphate group. Thus, this reagent was successfully used for the modification of enzymes having phosphate- and nucleotide-binding sites. In several cases, however, its specificity for directing the active sites of enzymes was not satisfactory. In these cases, a few or more lysyl residues in the regions apart from the active sites of enzymes were labeled concomitantly with inactivation of the enzymes.³⁴⁻³⁶⁾ In 1985, we first synthesized uridine diphosphopyridoxal (UP₂-PL) as a modification reagent for lysyl residue(s) of rabbit muscle glycogen synthase, which utilizes UDP-glucose as the substrate for the synthesis of glycogen.⁵⁾ This reagent bound to the enzyme with an affinity 1000-fold higher than PLP. This success prompted us to synthesize the adenine and guanine derivatives with different numbers of phosphate groups.^{6,12,17)} Subsequent studies showed that this type of reagent is useful for a wide variety of nucleotide-binding proteins.

4.2.1 Synthesis

Two different methods have been used for the synthesis of NP_n-PL. We applied the anion-exchange reaction of Michelson.³⁷⁾ For the synthesis of adenosine diphosphopyridoxal (AP₂-PL), for example, AMP was condensed with diphenyl phosphochloridate in the presence of tributylamine producing P¹-5'-adenosyl P²-diphenyl diphosphate, then PLP was added to exchange the diphenyl phosphate moiety of the compound. Adenosine triphosphopyridoxal (AP₃-PL) was also synthesized in a similar manner. Since P¹-5'-adenosyl P⁴-diphenyl tetraphosphate was not obtained by the reaction of ATP with diphenyl phosphochloridate, PLP was condensed with diphenyl phosphochloridate, and ATP was added to synthesize adenosine tetraphosphopyridoxal (AP₄-PL). In this case, AP₃-PL and presumably AP₂-PL were formed as byproducts. Probably, each phosphate group of ATP is capable of attacking the phosphorus of the phosphate group adjacent to pyridoxal moiety of P¹-5'-pyridoxal P²-diphenyl diphosphate. Routinely, we used to isolate AP₃-PL and AP₄-PL at the same time. Other NP_n-PL were synthesized by essentially the same method. In the same year we reported the synthesis of adenosine polyphosphopyridoxals⁶⁾ (AP_n-PL) following the synthesis of UP₂-PL, Tamura, Rakov and Cross¹⁸⁾ also reported the synthesis of AP₂-PL by using adenosine 5'-phosphomorpholidate and PLP. It seems possible to synthesize AP₃-PL by the same method if adenosine 5'-diphosphomorpholidate is used.

4.2.2 Inactivation Kinetics

The reaction of NP_n-PL with the enzyme protein (E) proceeds as shown in Scheme 4.1.³⁸⁾



where E·NP_n-PL and E=NP_n-PL represent reversible and covalent complexes, respectively. First, the reagent forms a reversible complex with the enzyme. Subsequently, the formyl group of the reagent reacts with the ε-amino group of lysyl

residue, forming a Schiff base. The degree of inactivation depends on the equilibrium between the two complexes. If this equilibrium leans completely toward the formation of the Schiff base, the degree of inactivation should be close to 100%. In this case, the inactivation kinetics is the same as that of the conventional affinity labels, and can be analyzed by Eq. 4.1.

$$\frac{1}{K_{\text{obs}}} = \frac{1}{k_2} + \frac{K_d}{k_2} \frac{1}{[\text{NP}_n\text{-PL}]} \quad (4.1)$$

where K_d represents the dissociation constant (k_{-1}/k_1). If the equilibrium does not lean toward the Schiff base, the inactivation kinetics can be analyzed by Eq. 4.2.

$$\frac{1}{[\text{NP}_n\text{-PL}]} = \frac{K_{\text{eq}}}{K_d} \frac{1}{(1-R)} - \frac{(1+K_{\text{eq}})}{K_d} \quad (4.2)$$

where R and K_{eq} represent the fractional residual activity at equilibrium and the equilibrium constant (k_2/k_{-2}).

4.2.3 Identification of Labeled Sites

Since the Schiff base formation is reversible, it should be reduced by sodium borohydride for the fixation of the label. The rate of the reduction of the Schiff base becomes slow as the number of the phosphate groups of the label increases. However, except for adenylate kinase, the $\text{NP}_n\text{-PL}$ bound to the proteins were easily fixed by borohydride reduction. After reductive fixation, labeled proteins are cleaved by appropriate methods. The labeled lysine is cleaved by neither trypsin nor lysyl endopeptidase. There are at least three ways to detect the labeled peptide during isolation: 1) use of radioactive reagent, 2) use of radioactive sodium borohydride for reduction of the Schiff base, and 3) use of fluorescence derived from the pyridoxyl moiety of the reagent (excitation at 295 nm and emission at 390 nm at acidic pH). The labeled lysyl residue is not positively identified in the amino acid sequence analysis. However, the presence of the label in the peptide isolated can be confirmed by the presence of pyridoxyl lysine in the amino acid analysis.

4.3 Case Studies

Table 4.1 summarizes the proteins modified by $\text{NP}_n\text{-PL}$. Several representative examples have been chosen and described below to demonstrate the utility of this type of reagent.

4.3.1 Glycogen Synthase^{5,15)}

$\text{UP}_2\text{-PL}$, the first example of $\text{NP}_n\text{-PL}$, was synthesized in 1985 as a new affinity labeling reagent for the UDP-glucose-binding site in rabbit muscle glycogen synthase. Incubation of the enzyme with this reagent resulted in an almost complete loss of activity in a time- and dose-dependent manner. The inactivation followed pseudo-first order kinetics, and the inactivated enzyme restored almost all the original activity by the addition of aminothiols compounds such as cysteamine. These results support the validity for analyzing the inactivation kinetics by Eq. 4.1, and a K_d of 25 μM and a maximal rate constant of 0.22 min^{-1} were calculated. The enzyme was almost completely protected from the inactivation by UDP or UDP-glucose, but not by the allosteric activator glucose 6-

TABLE 4.1 Proteins modified by NP_n-PL

Protein	Reference	Protein	Reference
Glycogen synthase	5,15	Adenylate kinase	8,9,18
UDP-glucose pyrophosphorylase	17	Pyruvate kinase	18
Callose synthase	96	Hexokinase	18,19
Lactate dehydrogenase	6	3-Phosphoglycerate kinase	18
Alcohol dehydrogenase	18	Phosphorylase kinase	11
Aldose reductase	97	Pyridoxal kinase	99,100†
F ₁ -ATPase	7,10,60,61	DNA-glyrase	21
Amyloplast inner membrane ATPase	98	Glutathione Synthetase	101
Ca ²⁺ -ATPase	13,14	Aminoacyl tRNA-synthetase	16
Na ⁺ ,K ⁺ -ATPase	73	Ribulose phosphate kinase	102
		Rho factor	20
		<i>ras</i> P21 protein	12

† AP₄-PL is a slow-binding inhibitor for pig liver pyridoxal kinase.

phosphate, suggesting binding of the reagent to the UDP-glucose (UDP)-binding site of the enzyme. Based on the increase in the absorbance at 425 nm, which is derived from the Schiff base formation, the stoichiometry of the bound reagent to the enzyme subunit was calculated to be 1 : 1. The results of the effect of pH on the inactivation showed that the p*K* value of the amino acid residue(s) involved in inactivation is 8.9. This value most likely reflects the p*K* value of the ε-amino group of the labeled lysine. Sequence study revealed that the structure of the labeled peptide to be Glu-Val-Ala-Asn-Labeled Lys-Val-Gly-Gly-Ile- (Thr). Later, the labeled lysine was identified as Lys-38 by Mahrenholz, Wang and Roach.³⁹⁾

Escherichia coli glycogen synthase has several different properties from the mammalian enzymes. It utilizes ADP-glucose in place of UDP-glucose as a glucosyl donor, and its subunit molecular weight is 45,000, about half the subunit size of the mammalian enzyme.⁴⁰⁾ Apparently, no significant sequence homology is observed between the *E. coli* and mammalian enzymes.^{40,41)} To compare the structure of the active sites of the two enzymes, we used AP₂-PL for the modification of the *E. coli* enzyme. AP₂-PL bound to the enzyme with the stoichiometry of the reagent to the enzyme subunit to be 1 : 1. ADP and ADP-glucose offered a protective effect on the inactivation. Sequence study revealed that Lys-15 is the labeled lysine. Comparison of the labeled amino acid lysine and its surroundings of the *E. coli* enzyme with those of the mammalian enzyme showed that the sequence of Lys*-X-Gly-Gly, where Lys* and X represent the labeled lysyl and unspecific amino acid residues, respectively, is conserved. More recently, it was found that a residue equivalent to Lys-38 of the rabbit muscle enzyme is an arginyl residue in yeast glycogen synthases.⁴²⁾ Therefore, the presence of a positive charge, not specific amino acid residue, in this sequence may be important for enzyme catalysis. Site-directed mutagenesis studies on this conserved region of the *E. coli* enzyme are now being carried out in our laboratory. Preliminary results suggest the involvement of this region in the binding of ADP-glucose (K. Furukawa *et al.*, unpublished results).

4.3.2 Adenylate Kinase^{8,9)}

Adenylate kinase was inactivated by incubation with AP₂-PL in a time-dependent

manner. Surprisingly, the reagent has such high affinity for the enzyme that almost all the reagent added to the inactivation mixture bound to the enzyme in a 1 : 1 stoichiometry. ADP, ATP, and MgATP, but not AMP, protected the enzyme from inactivation, suggesting the binding of AP₂-PL to the MgATP-binding site of the enzyme. The labeled lysine was identified as Lys-21. AP₃-PL and AP₄-PL modified the same lysyl residue as AP₂-PL. However, protein denaturing reagents such as urea are required for the reductive fixation of the Schiff bases between Lys-21 and the labeling reagents. In the absence of protein denaturing reagents, half the bound AP₃-PL and all of the bound AP₄-PL were resistant to sodium borohydride reduction. Except for this enzyme, AP₃-PL and AP₄-PL were fixed by the sodium borohydride treatment to the nucleotide-binding proteins so far examined. To explain why the three reagents bound to Lys-21 showed different susceptibilities of Schiff bases to sodium borohydride, we postulated that the ϵ -amino group of Lys-21 is mobile, fitting the position of the formyl group of each reagent. Thus, the environments of the Schiff bases are different from one another. It is known that a Schiff base between the coenzyme PLP and the lysyl residue in phosphorylase is resistant against mild borohydride treatment.⁴³⁾ The protein denaturing reagents are required for the reductive fixation.^{44,45)} This property of the Schiff base in phosphorylase was explained by the hydrophobic nature around the bound coenzyme.⁴⁶⁾ Because borohydride is hydrophilic, it cannot access the Schiff base located in the hydrophobic environment. The different resistances of the Schiff bases in adenylylate kinase against borohydride reduction may be explained in a similar manner.

Lys-21 of adenylylate kinase is in the consensus sequence for nucleotide-binding^{47,48)} (Gly-Gly-Pro-Gly-Ser-Gly-Lys-Gly, where the underlined amino acids are conserved). It is noteworthy that AP₃-PL and GP₂-PL modified the conserved lysyl residues in the consensus sequences of the β -subunit of *E. coli* F₁-ATPase¹⁰⁾ (as described below) and *ras* p21 protein,¹²⁾ respectively. Therefore, NP_{*n*}-PL are capable of probing the consensus sequence for nucleotide-binding. Results of site-directed mutagenesis suggest that the consensus sequence of adenylylate kinase plays an important role in the binding of the two nucleotides and enzyme catalysis. Replacement of the conserved lysyl residue by Glu or Met resulted in an almost complete loss of activity.^{49,50)} Individual replacement of the conserved glycyl residues by Ala resulted in remarkable increase in the apparent binding constants for the two nucleotides without significant change of $V_{\max}$.⁵¹⁾ X-ray crystallographic analysis of adenylylate kinase⁵²⁾ and *ras* p21 protein⁵³⁾ revealed that the consensus sequence is close to the phosphoryl group of the bound nucleotide.

4.3.3 UDP-Glucose Pyrophosphorylase^{17,54)}

UDP-glucose pyrophosphorylase catalyzes the following reversible reaction. The liver enzyme is a large protein consisting of eight subunits with a molecular weight of 60,000.⁵⁵⁾ By contrast, the potato tuber enzyme is monomeric with a molecular weight of 50,000.⁵⁶⁾ The simpler structure of the potato enzyme is suited for analysis of structure-function relationships by chemical and molecular biological methods.



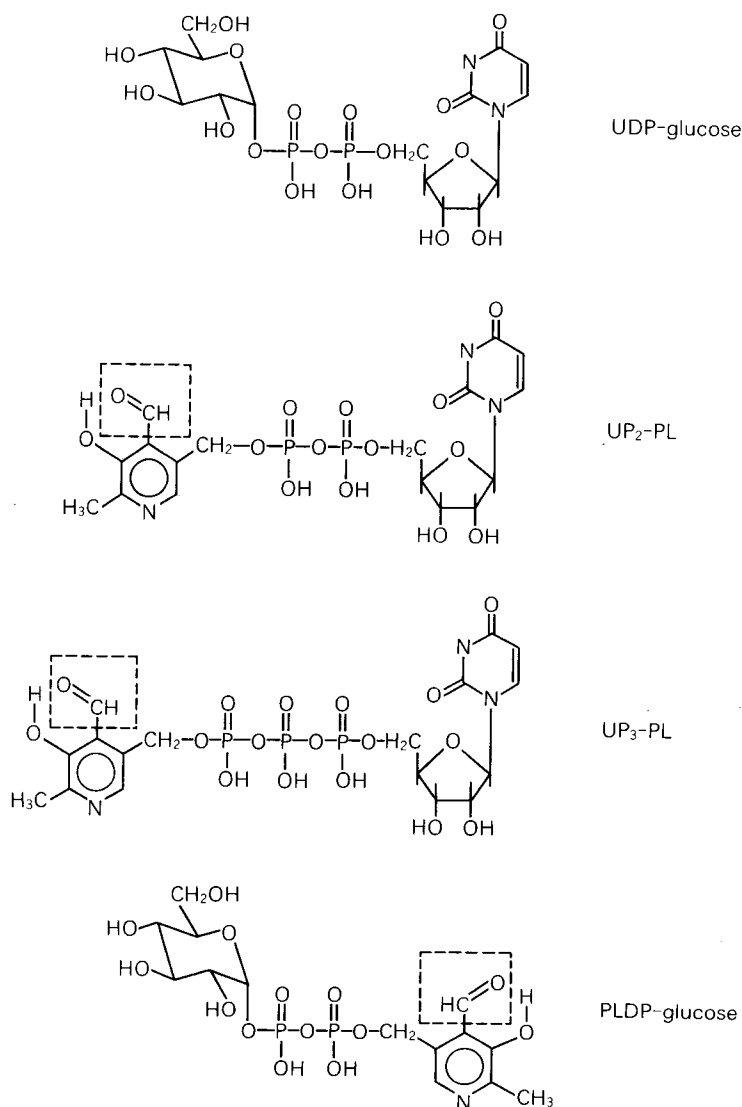


Fig. 4.2 Structures of UDP-glucose and its reactive analogues used in the comparison affinity labeling.⁵⁴⁾

Recently, we have purified the potato tuber enzyme,⁵⁶⁾ modified the enzyme with several types of PLP analogues,^{17,54)} cloned its cDNA,⁵⁷⁾ and constructed the expression system in *E. coli*.⁵⁸⁾ Fig. 4.2 shows the structures of the reagents used for the modification of the enzyme in comparison with that of the substrate. Although UP₂-PL and UP₃-PL bound to the enzyme in a 1 : 1 stoichiometry, five lysyl residues (Lys-263, Lys-329, Lys-367, Lys-409, and Lys-410) were labeled in varying degrees (Table 4.2). Pyridoxal-diphos-

TABLE 4.2 Comparative affinity labeling of five lysyl residues with three kinds of reactive UDP-glucose analogues³⁴⁾

Reagent	Residual activity (%)	Relative amounts of label (%)				
		K263	K329	K367	K409	K410
UP ₂ -PL	15	23	9	59	6	3
UP ₃ -PL	15	20	9	61	7	3
PLDP-Glc	20	17	18	32	11	22

(Reproduced with permission from *J. Biochem.*, **110**, 711 (1991)).

phoglucose (PLDP-glucose), containing the reactive pyridoxal moiety on the opposite side of the pyrophosphate linkage to those of UP₂-PL and UP₃-PL, also bound to the enzyme in a 1:1 stoichiometry and labeled the same five lysyl residues. However, the incorporation of PLDP-glucose into each lysyl residue was different from those of UP₂-PL and UP₃-PL (Table 4.2). The relative amounts of PLDP-glucose bound to Lys-329, Lys-409, and Lys-410 were larger than those of the other two reagents. By contrast, the relative amount of PLDP-glucose bound to Lys-367 was significantly lower than the amounts of the other two reagents.

Assuming that these differences in the pattern of labeling reflect the distance between the reactive group of the reagent and the ϵ -amino group of the lysyl residue, a hypothetical model of the location of the five lysyl residues at the active site of potato tuber UDP-glucose pyrophosphorylase was proposed (Fig. 4.3). In this model, the ϵ -amino group of Lys-367 is positioned most closely to the substrate and nearer to the glucose moiety than to the uridine moiety. Because Lys-367 is dominantly labeled by UP₂-PL and UP₃-PL, but not PLDP-glucose. The ϵ -amino groups of Lys-263 and Lys-329 are near the glucose and the uridine moieties, respectively. Lys-409 and Lys-410 are farther from the substrate than the other lysyl residues but by the side of the uridine moiety.

The results of site-directed mutagenesis studies are generally consistent with this model.⁵⁸⁾ Table 4.3 summarizes the kinetic parameters of the mutant enzymes where the five labeled lysyl residues are individually replaced by Gln. The Lys-367/Gln enzyme was almost completely inactive. The Lys-263/Gln enzyme had slight activity with decrease in affinity to pyrophosphate and glucose 1-phosphate. The Lys-329/Gln enzyme also had considerably lower affinity to the two substrates, but possessed a $V_{\max}$ value comparable to the wild-type enzyme. The replacements of Lys-409 and Lys-410 by Gln did not change the kinetic properties of the enzyme. The kinetic properties of the Lys-329/Gln enzyme suggest that Lys-329 interacts with pyrophosphate after UDP-glucose binds to the active site.

Recently, we tried to probe the pyrophosphate-binding site by pyridoxal diphosphate (PLDP)¹⁰³⁾. This reagent also bound to the enzyme in a 1:1 stoichiometry. Although UDP-glucose had a protective effect on inactivation by UP₂-PL and UP₃-PL, it rather enhanced the degree of inactivation by PLDP. This unusual effect on inactivation can be explained by the idea that an additional site for the binding of PLDP is created by the addition of UDP-glucose. Because the kinetics of UDP-glucose pyrophosphorylase is an ordered Bi-Bi mechanism where pyrophosphate binds to the enzyme following the binding of UDP-glucose, this new site is probably the pyrophosphate-binding site. This idea was confirmed by the fact that UDP-glucose did enhance the degree of inactivation by PLDP,

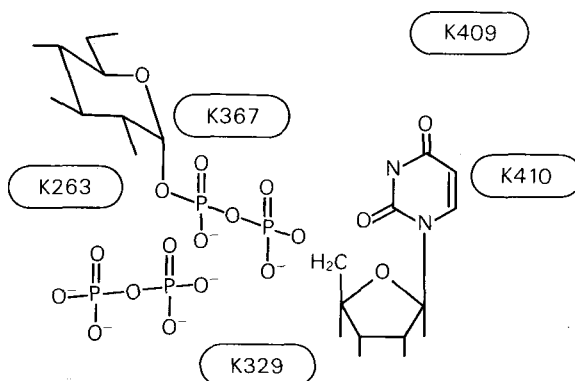


Fig. 4.3 A hypothetical model for the location of five lysyl residues around the substrates bound to potato tuber UDP-glucose pyrophosphorylase, constructed based on the results of the comparative affinity labeling.⁵⁴⁾ (Reproduced with permission from *J. Biochem.*, **110**, 713 (1991)).

TABLE 4.3 Kinetic parameters of wild-type and mutant enzymes of potato tuber UDP-glucose pyrophosphorylase⁵⁸⁾

Enzyme	$V_{\max}$		K_m			
	Degradation (unit/mg)	Synthesis (unit/mg)	UDP-Glc (mM)	PPi (mM)	G-1-P (mM)	UTP (mM)
Wild	1300	780	0.15	0.14	0.13	0.076
K263Q	40	10	0.066	14	1.4	0.074
K329Q	920	440	0.19	3.1	0.99	0.17
K367Q	0.15	0.38	0.12	2.0	8.0	0.33
K409Q	1000	720	0.17	0.17	0.19	0.11
K410Q	1200	950	0.22	0.16	0.23	0.10

(Reproduced with permission from *Biochemistry*, **30**, 8550 (1991)).

but by neither PLP nor pyridoxal triphosphate. In the absence of UDP-glucose, several lysyl residues were labeled by PLDP. In contrast, Lys-329 was almost exclusively labeled in the presence of UDP-glucose. These results strongly suggest that Lys-329 is involved in the binding of pyrophosphate.

4.3.4 F_1 -ATPase

A. Modification of *E. coli* F_1 -ATPase by AP_n -PL^{7,10)}

E. coli F_1 -ATPase consists of the structure $\alpha_3\beta_3\gamma\delta\epsilon$. The α -subunit and β -subunit each independently possesses a nucleotide-binding site and the catalytic site is believed to be on the β -subunit or at the interface between the α - and β -subunits (for reviews, see Ref. 59 and Chapter 11 of this volume). *E. coli* F_1 -ATPase was more effectively inactivated by AP_3 -PL and AP_4 -PL than AP_2 -PL. The binding of one mole of AP_3 -PL to one mole of F_1 resulted in an almost complete loss of activity; 70% of the label bound to F_1 to the α -subunit and the rest to the β -subunit. The sequence analysis of the F_1 modified by AP_3 -PL in the absence of Mg^{2+} revealed that Lys-201 of the α -subunit and Lys-155 of the β -subunit are the

labeled sites. Lys-201 of the α -subunit was also labeled when the isolated α -subunit was modified by a relatively high concentration of $\text{AP}_3\text{-PL}$.⁶⁰⁾ When Lys-201 of the α -subunit was replaced by Gln or deleted, $\text{AP}_3\text{-PL}$ bound exclusively to the β -subunit.⁶¹⁾ These results suggest that the two labeled lysyl residues are closely located and consist of the catalytic ATP-binding site, although Lys-201 of the α -subunit is not essential.

Mg^{2+} is essential for the hydrolysis of ATP by F_1 . Addition of Mg^{2+} decreased the concentration of $\text{AP}_3\text{-PL}$ required for half-maximum inactivation (10 μM and 2.5 μM in the absence and presence of Mg^{2+} , respectively) and changed the distribution pattern of bound $\text{AP}_3\text{-PL}$ without change in inactivation stoichiometry.⁶¹⁾ In the presence of Mg^{2+} , the β -subunit was almost exclusively labeled by the reagent. Sequence study identified Lys-155 and Lys-201 of the β -subunit as the labeled sites, suggesting that this residue as well as Lys-155 of the β -subunit and Lys-201 of the α -subunit are present at the catalytic site.

Lys-155 of the β -subunit is in the consensus sequence for the nucleotide-binding (Gly–Gly–Ala–Gly–Val–Gly–Lys–Thr, where the conserved residues are underlined). The results of site-directed mutagenesis studies of this sequence suggest its importance for ATP hydrolysis.^{62–64)}

B. Modification of Bovine Mitochondrial $\text{F}_1\text{-ATPase}$ by FSBA^{65–67)}, and Fluoro-sulfonylbenzoyl Inosine (FSBI)⁶⁸⁾

The inactivation of bovine $\text{F}_1\text{-ATPase}$ by FSBA exhibited biphasic kinetics. A double reciprocal plot of the inactivation rate for the fast phase against FSBA concentration gave a curved line, whereas the same type of plot for the slow phase yielded a straight line, giving a K_d value of 0.23 mM. The slow phase was diminished when the enzyme was inactivated in the presence of 0.2 M phosphate. On complete inactivation, about three moles of FSBA bound to one mole of bovine $\text{F}_1\text{-ATPase}$. Regardless of the presence of phosphate, His-427 of the β -subunit was dominantly modified at pH 6, whereas Tyr-368 of the β -subunit was dominantly modified at pH 8. At pH 7, the two residues were modified in a similar ratio.

The inactivation of bovine $\text{F}_1\text{-ATPase}$ by FSBI showed single phase kinetics. A double reciprocal plot of inactivation rate against FSBI concentration yielded a straight line, giving a K_d value of 0.5 mM. FSBI modified only one copy of the three β -subunits and the labeled site was identified as Tyr-345 of the β -subunit. When the sensitivities of the ITPase and ATPase activities of F_1 against FSBA and FSBI were compared, the ITPase activity diminished faster than the ATPase activity in the modification by FSBA, whereas both activities were lost at a similar rate by FSBI. These differences in the modifications by FSBA and FSBI were explained by the different binding sites for the two labels; FSBI binds to the catalytic ATP-binding site, whereas FSBA binds to the regulatory site. A similar conclusion was obtained by the modification by 2-azido ATP. This label binds to Tyr-368 (noncatalytic) and Tyr-345 (catalytic) of the β -subunit.⁶⁹⁾ Tyr-345 of the β -subunit was also modified by 3'-*O*-(4-benzoyl)-benzoyl ATP.⁷⁰⁾

4.3.5 Ion-Transporting ATPase

A. Modification of sarcoplasmic reticulum ATPase by $\text{AP}_n\text{-PL}$ ^{13,14)}

Sarcoplasmic reticulum Ca^{2+} -transporting ATPase (SR-ATPase) is a member of the ATPase family that couples ion translocation across cell membranes to hydrolysis. There

are significant conservations in the amino acid residues located at the ATP-binding site of this class of ATPases.^{71,72)} SR-ATPase was more effectively inactivated by AP₃-PL and AP₄-PL than AP₂-PL and PLP. Ca²⁺, a cation transported by this enzyme, markedly decreased the concentrations of AP_n-PL required for half-maximum inactivation. However, regardless of the presence Ca²⁺, AP₃-PL bound to the enzyme in a 1:1 stoichiometry. Sequence study revealed that Lys-684 is labeled exclusively by AP₃-PL in the presence of Ca²⁺, whereas Lys-492 and Lys-684 are labeled in an almost equal ratio in the absence of Ca²⁺. These results imply that the reactivities of Lys-492 and Lys-684 to AP₃-PL are changed by Ca²⁺. This change suggests a conformational change in the labeled amino acid residues by Ca²⁺. Both lysyl residues are conserved for the Na⁺, K⁺-ATPase and H⁺, K⁺-ATPase.^{71,72)} Interestingly, Lys-480 of Na⁺, K⁺-ATPase, equivalent to Lys-492 of SR-ATPase, was labeled by PLP, when the enzyme was modified in the presence of both Na⁺ and K⁺.⁷³⁾ AP₂-PL also bound to Lys-480 of Na⁺, K⁺-ATPase.⁷³⁾ Lys-497 of H⁺, K⁺-ATPase, equivalent to Lys-492 of SR-ATPase, was also labeled by PLP in the absence of K⁺.⁷⁴⁾ It seems that the reactivities of Lys-480 of Na⁺, K⁺-ATPase and Lys-497 of H⁺, K⁺-ATPase to PLP are higher than that of Lys-492 of SR-ATPase irrespective of the presence of cationic ions. Recently, site-directed mutagenesis showed that Lys-480 of Na⁺, K⁺-ATPase (Lys-492 of SR-ATPase) is not essential for the binding and hydrolysis of ATP.⁷⁵⁾

B. Modification of Na⁺, K⁺-ATPase by FSBA^{76,77)}

Na⁺, K⁺-ATPase was inactivated by FSBA in the presence of Mg²⁺. Although Na⁺ and K⁺ are not essential for inactivation, they accelerate the inactivation rate. In the absence of Mg²⁺, the enzyme was not inactivated by FSBA even in the presence of both Na⁺ and K⁺. Sequence study on the FSBA-modified dog kidney enzyme revealed that one of the major labeled sites is a lysyl residue equivalent to Lys-725 of the torpedo enzyme. The simultaneous replacement of both Lys-712 and Lys-713 by Met of SR-ATPase (The two lysyl residues are equivalent to Lys-725 and Lys-726 of the torpedo Na⁺, K⁺-ATPase) did not change the Ca²⁺-transporting activity or the formation of phosphoenzyme intermediate,⁷⁸⁾ indicating that these lysyl residues are not essential for enzyme activity.

4.3.6 Protein Kinases

Protein kinases catalyze the transfer of the γ -phosphate of ATP to a seryl or threonyl residue (Ser/Thr kinase) or a tyrosyl residue (Tyr-kinase). Their primary structures exhibit significant homologies to one another (for a review, *see* Ref. 79).

A. Modification of Phosphorylase Kinase by AP_n-PL¹¹⁾

Phosphorylase kinase, one of the Ser/Thr-kinases, is oligomeric with the stoichiometry ($\alpha\beta\gamma\delta$)₄.⁸⁰⁾ The γ -subunit of the enzyme is homologous to other protein kinases and possesses protein kinase activity when separated from other subunits.⁸¹⁾ Phosphorylase kinase was more effectively inactivated by AP₃-PL and AP₄-PL than PLP and AP₂-PL. Ca²⁺ and Mg²⁺, activators for this enzyme, enhanced the degree of inactivation by all the pyridoxal compounds. Inactivation by AP₃-PL and AP₄-PL was markedly protected by adenylyl, β,γ -imidodiphosphate, a nonhydrolyzable analogue of ATP, and ADP. Because the α - and β -subunits have regulatory ATP-binding sites and the kinase activity is regulated by these subunits,⁸²⁾ the incorporation of AP₃-PL into the γ -subunit and

inactivation were not correlated. However, the incorporation of AP₃-PL into the γ -subunit was significantly decreased by ATP and ADP. Sequence study revealed that Lys-151 is the major site labeled by AP₃-PL. This lysyl residue is conserved for all the Ser/Thr-kinases, but not for Tyr-kinase.⁷⁹⁾

The carboxyl terminal site of Lys-151 was postulated to be the substrate peptide recognition region for the following reasons. The structure of Lys¹⁵¹-Pro-Glu-Asn¹⁵⁴ is well conserved for Ser/Thr-kinases except for protein kinase C (Lys-Leu-Asp-Asn) and casein kinase II (Lys-Pro-His-Asn). The corresponding region of Tyr-kinase is Arg-Ala-Ala-Asn or Ala-Ala-Arg-Asn. Ser/Thr-protein kinases containing an acidic amino acid at the third position from the amino-terminus of this segment phosphorylate a seryl or threonyl residue in the basic regions,⁸³⁻⁸⁵⁾ whereas casein kinase II, containing a histidyl residue at the third position, recognizes a seryl or threonyl residue in the acidic regions.⁸⁶⁾ In addition, Tyr-protein kinase prefers a tyrosyl residue in the acidic regions.⁸⁷⁾ Therefore, the difference in these four amino acid residues on the carboxyl side of Lys-151 is correlated with the difference in the specificity for protein kinases.

Glu-170 of the catalytic subunit of cAMP-dependent protein kinase (protein kinase A), equivalent to Glu-153 of the γ -subunit of phosphorylase kinase, was labeled by 1-ethyl-3-[3-diethylaminopropyl]carbodiimide.⁸⁸⁾ The modification of this residue was well protected by ATP and a peptide inhibitor together, but not by each one alone, consistent with the idea that the region around Glu-153 is the site for the binding of both substrates. Genetic study also confirmed the important role of the region around Glu-153. The mutation of Leu-217 of yeast protein kinase A, equivalent to Leu-155 of the γ -subunit of phosphorylase kinase, resulted in loss of protein kinase activity.⁸⁹⁾

B. Modification of Protein Kinase A by FSBA^{90,91)}

The catalytic subunit of protein kinase A was inactivated by FSBA. Inactivation follows pseudo-first order kinetics, giving K_d and k_2 values of 0.057 mM and 0.0364 min⁻¹. Inactivation was protected by MgATP. Sequence study on the FSBA-modified porcine heart enzyme revealed that Lys-71 is the labeled site. This lysyl residue is completely conserved for both Ser/Thr- and Tyr-kinases.⁷⁹⁾ The equivalent lysyl residues of pp60src⁹²⁾ and epidermal growth factor receptor⁹³⁾ were also labeled by FSBA. The replacement of this lysyl residue of pp60src protein resulted in loss of transforming activity,^{94,95)} suggesting the importance of this conserved lysyl residue. Although the mutation of Lys-71 of protein kinase A has not yet been carried out, it is most likely that this lysyl residue is essential for protein kinase activity.

4.4 Concluding Remarks

NP_n-PL have been successfully used for the modification of various proteins containing nucleotide-binding sites. Although this type of reagent and fluorosulfonylbenzoyl nucleosides have the reactive groups in a similar position, the two types of reagents modified different amino acid residues of proteins so far examined. Thus, it is of worth to use both types of reagents for the identification of amino acids located at the phosphoryl-binding regions of nucleotide-binding proteins as well as for the determination of the target amino acids for site-directed mutagenesis.

REFERENCES

1. E. Show, *The Enzymes*, 3rd Ed. vol. 1, p. 91, Academic Press, New York (1970).
2. B.R. Baker, *Design of Active-Site-Directed Irreversible Enzyme Inhibitors*, John Wiley and Sons, Inc., New York, NY (1967).
3. B.V. Plapp, *Methods in Enzymology*, vol. 87, p. 469, Academic Press, New York (1982).
4. W.B. Jakoby, M. Wilchek(eds.) *Methods in Enzymology*, vol. 46, Academic Press, New York (1977).
5. M. Tagaya, K. Nakano, T. Fukui, *J. Biol. Chem.*, **260**, 6670 (1985).
6. M. Tagaya, T. Fukui, *Biochemistry*, **25**, 2958 (1986).
7. T. Noumi, M. Tagaya, K. Miki-Takeda, M. Maeda, T. Fukui, M. Futai, *J. Biol. Chem.*, **262**, 7686 (1987).
8. M. Tagaya, T. Yagami, T. Fukui, *J. Biol. Chem.*, **262**, 8257 (1987).
9. T. Yagami, M. Tagaya, T. Fukui, *FEBS Lett.*, **229**, 261 (1988).
10. M. Tagaya, T. Noumi, K. Nakano, M. Futai, T. Fukui, *FEBS Lett.*, **233**, 347 (1988).
11. M. Tagaya, S. Hayakawa, T. Fukui, *J. Biol. Chem.*, **263**, 10219 (1988).
12. N. Ohmi, M. Hoshino, M. Tagaya, T. Fukui, M. Kawakita, S. Hattori, *J. Biol. Chem.*, **263**, 14261 (1988).
13. H. Yamamoto, M. Tagaya, T. Fukui, M. Kawakita, *J. Biochem.*, **103**, 452 (1988).
14. H. Yamamoto, Y. Imamura, M. Tagaya, T. Fukui, M. Kawakita, *J. Biochem.*, **106**, 1121 (1989).
15. K. Furukawa, M. Tagaya, M. Inouye, J. Preiss, T. Fukui, *J. Biol. Chem.*, **265**, 2086 (1990).
16. C. Hountondji, J.-M. Schmitter, T. Fukui, M. Tagaya, S. Blanquet, *Biochemistry*, **29**, 11266 (1990).
17. Y. Kazuta, Y. Omura, M. Tagaya, K. Nakano, T. Fukui, *Biochemistry*, **30**, 8541 (1991).
18. J.K. Tamura, R.D. Rakov, R.L. Cross, *J. Biol. Chem.*, **261**, 4126 (1986).
19. J.K. Tamura, J.R. LaDine, R.L. Cross, *J. Biol. Chem.*, **263**, 7907 (1988).
20. A.J. Dombroski, J.R. LaDine, R.L. Cross, T. Platt, *J. Biol. Chem.*, **263**, 18810 (1988).
21. J.K. Tamura, M. Gellert, *J. Biol. Chem.*, **265**, 21342 (1990).
22. G.K. Kovaleva, L.L. Ivanov, I.A. Madoyan, O.O. Favorova, E.S. Severin, N.N. Gulyaev, L.A. Baranova, Z.A. Shabarova, N.I. Sokolova, L.L. Kiselev, *Biochemistry (Engl. Transl.)*, **43**, 525 (1978).
23. R. Krauspe, G.K. Kovaleva, N.N. Gulyaev, L.A. Baranova, M.B. Agalarova, E.S. Severin, N.I. Sokolova, Z.A. Shabarova, L.L. Kiselev, *Biochemistry (Engl. Transl.)*, **43**, 656 (1978).
24. L.K. Skolysheva, P.L. Vulfson, N.N. Gulyaev, E.S. Severin, *Biochemistry (Engl. Transl.)*, **43**, 1914 (1978).
25. L.I. Mikhailova, P.L. Vulfson, L.K. Skolysheva, M.B. Agalarova, E.S. Severin, *Biochemistry (Engl. Transl.)*, **43**, 2016 (1978).
26. I.A. Grivennikov, T.V. Bulargina, Y.V. Khropov, N.N. Gulyaev, E.S. Severin, *Biochemistry (Engl. Transl.)*, **44**, 771 (1979).
27. M.A. Grachev, G.R. Hartman, T.G. Maximova, A.A. Mustaev, A.R. Schaffner, H. Sieber, E.F. Zaychikov, *FEBS Lett.*, **200**, 287 (1986).
28. M.A. Grachev, T.I. Kolocheva, E.A. Lukhtanov, A.A. Mustaev, *Eur. J. Biochem.*, **163**, 113 (1987).
29. A.R. Schaffner, E.D. Jorgensen, W.T. McAllister, G.R. Hartman, *Nucleic Acids Res.*, **15**, 8773 (1987).
30. R.F. Colman, *Annu. Rev. Biochem.*, **52**, 67 (1983).
31. R.F. Colman, *The Enzymes*, 3rd Ed. vol. 19, p. 283, Academic Press, San Diego (1990).
32. P.K. Pal, W.J. Wechter, R.F. Colman, *J. Biol. Chem.*, **250**, 8140 (1975).
33. R.E. Feeney, G. Blandenhorn, H.B. Dixon, *Adv. Protein Chem.*, **29**, 135 (1975).
34. S.-S. Chen, P.C. Engel, *Biochem. J.*, **151**, 447 (1975).
35. K.G. Gould, P.C. Engel, *Biochem. J.*, **191**, 365 (1980).
36. K.G. Gould, P.C. Engel, *Arch. Biochem. Biophys.*, **215**, 498 (1982).
37. A.M. Michelson, *Biochim. Biophys. Acta*, **91**, 1 (1964).
38. S.-S. Chen, P.C. Engel, *Biochem. J.*, **149**, 627 (1975).
39. A.M. Mahrenholz, Y. Wang, P.J. Roach, *J. Biol. Chem.*, **263**, 10561 (1988).
40. A. Kumar, C.E. Larsen, J. Preiss, *J. Biol. Chem.*, **261**, 16256 (1986).
41. M.F. Browner, K. Nakano, A.G. Beng, R.J. Fletterick, *Proc. Natl. Acad. Sci. USA*, **86**, 1443 (1989).
42. I. Farkas, T.A. Hardy, M.G. Goebel, P.J. Roach, *J. Biol. Chem.*, **266**, 15602 (1991).
43. E.H. Fischer, A.B. Kent, E.R. Snyder, E.G. Krebs, *J. Amer. Chem. Soc.*, **80**, 2906 (1958).
44. K. Lerch, E.H. Fischer, *Biochemistry*, **14**, 2009 (1975).
45. K. Nakano, S. Wakabayashi, T. Hase, H. Matsubara, T. Fukui, *J. Biochem.*, **83**, 1085 (1978).
46. S. Shaltiel, M. Cortijo, *Biochem. Biophys. Res. Commun.*, **41**, 594 (1970).
47. J.E. Walker, M. Saraste, M.J. Runswick, N.J. Gay, *EMBO J.*, **1**, 945 (1982).
48. D.C. Fry, S.A. Kuby, A.S. Mildvan, *Proc. Natl. Acad. Sci. USA*, **83**, 907 (1986).
49. J. Reinstein, M. Brune, A. Wittinghofer, *Biochemistry*, **27**, 4712 (1988).
50. G. Tian, H. Yan, R.-T. Jiang, F. Kishi, A. Nakazawa, M.-D. Tsai, *Biochemistry*, **29**, 4296 (1990).

51. T. Yoneya, M. Tagaya, F. Kishi, A. Nakazawa, T. Fukui, *J. Biochem.*, **105**, 158 (1989).
52. D. Dreusicke, P.A. Karplus, G.E. Schulz, *J. Mol. Biol.*, **199**, 359 (1988).
53. E.F. Pai, W. Kabsch, U. Krenkel, K.C. Holmes, J. John, A. Wittinghofer, *Nature*, **341**, 209 (1989).
54. Y. Kazuta, K. Tanizawa, T. Fukui, *J. Biochem.*, **110**, 708 (1991).
55. S. Levine, T.A. Gillett, E. Hageman, R.G. Hansen, *J. Biol. Chem.*, **244**, 5729 (1969).
56. K. Nakano, Y. Omura, M. Tagaya, T. Fukui, *J. Biochem.*, **106**, 528 (1989).
57. T. Katsube, Y. Kazuta, H. Mori, K. Nakano, K. Tanizawa, T. Fukui, *J. Biochem.*, **108**, 321 (1990).
58. T. Katsube, Y. Kazuta, T. Tanizawa, T. Fukui, *Biochemistry*, **30**, 8546 (1991).
59. M. Futai, T. Noumi, M. Maeda, *Ann. Rev. Biochem.*, **58**, 111 (1989).
60. R. Rao, D. Cunningham, R.L. Cross, A.E. Senior, *J. Biol. Chem.*, **263**, 5640 (1988).
61. K. Ida, T. Noumi, M. Maeda, T. Fukui, M. Futai, *J. Biol. Chem.*, **266**, 5424 (1991).
62. D. Parsonage, S. Wilke-Mounts, A.E. Senior, *J. Biol. Chem.*, **262**, 8022 (1987).
63. D. Parsonage, M.K. Al-Shawi, A.E. Senior, *J. Biol. Chem.*, **263**, 4740 (1988).
64. M. Yohda, S. Ohta, T. Hisabori, Y. Kagawa, *Biophys. Biochim. Acta*, **933**, 156 (1988).
65. F.S. Esch, W.S. Allison, *J. Biol. Chem.*, **253**, 6100 (1978).
66. A. DiPietro, C. Godinot, J.-C. Martin, D.C. Gautheron, *Biochemistry*, **18**, 1738 (1979).
67. D.A. Bullough, W.S. Allison, *J. Biol. Chem.*, **261**, 5722 (1986).
68. D.A. Bullough, W.S. Allison, *J. Biol. Chem.*, **261**, 14171 (1986).
69. R.L. Cross, D. Cunningham, C.G. Miller, Z. Xue, J.-M. Zhou, P.D. Boyer, *Proc. Natl. Acad. Sci. USA*, **84**, 5715 (1987).
70. A. Admon, G.G. Hammes, *Biochemistry*, **26**, 3193 (1987).
71. R. Serrano, *Biochim. Biophys. Acta*, **947**, 1 (1988).
72. N.M. Green, *Biochem. Soc. Transact.*, **17**, 819 (1989).
73. H.R. Hinz, T.L. Kirley, *J. Biol. Chem.*, **265**, 10260 (1990).
74. S. Tamura, M. Tagaya, M. Maeda, M. Futai, *J. Biol. Chem.*, **264**, 8580 (1989).
75. K. Wang, R.A. Farley, *J. Biol. Chem.*, **267**, 3577 (1992).
76. J.B. Cooper, C.G. Winter, *J. Supramol. Struct.*, **3**, 165 (1980).
77. T. Ohta, K. Nagano, M. Yoshida, *Proc. Natl. Acad. Sci. USA*, **83**, 2071 (1986).
78. K. Maruyama, D.M. Clarke, J. Fujii, G. Inesi, T.W. Loo, D.H. MacLennan, *J. Biol. Chem.*, **264**, 13038 (1989).
79. T. Hunter, J.A. Cooper, in: *The Enzymes*, 3rd Ed., vol. 17, Academic Press, New York, pp. 191 (1986).
80. P. Cohen, A. Burchell, J.G. Foulkes, P.T.W. Cohen, T.C. Vanaman, A.C. Nairn, *FEBS Lett.*, **92**, 287 (1978).
81. S.M. Kee, D.J. Graves, *J. Biol. Chem.*, **261**, 4732 (1986).
82. H.K. Paudel, G.M. Carlson, *J. Biol. Chem.*, **262**, 11912 (1987).
83. K.-F.J. Chan, M.O. Hurst, D.J. Graves, *J. Biol. Chem.*, **257**, 3655 (1982).
84. C. House, R.E.H. Wettenhall, B.E. Kemp, *J. Biol. Chem.*, **262**, 772 (1987).
85. B.E. Kemp, R.B. Pearson, *J. Biol. Chem.*, **260**, 3355 (1985).
86. F. Meggio, F. Marchiori, G. Borin, G. Chessa, L.A. Pinna, *J. Biol. Chem.*, **259**, 14576 (1984).
87. T. Patschinsky, T. Hunter, F.S. Esch, J.A. Cooper, B.M. Sefton, *Proc. Natl. Acad. Sci. USA*, **79**, 973 (1982).
88. J.A. Buechler, S.S. Taylor, *Biochemistry*, **29**, 1937 (1990).
89. S. Cameron, L. Levin, M. Zoller, M. Wigler, *Cell*, **53**, 555 (1988).
90. M.J. Zoller, S.S. Taylor, *J. Biol. Chem.*, **254**, 8363 (1979).
91. M.J. Zoller, N.C. Nelson, S.S. Taylor, *J. Biol. Chem.*, **256**, 10837 (1981).
92. M.P. Kamps, S.S. Taylor, B.M. Sefton, *Nature*, **310**, 589 (1984).
93. M.W. Russo, T.J. Lukas, S. Cohen, J.V. Staros, *J. Biol. Chem.*, **260**, 5205 (1985).
94. M.A. Snyder, J.M. Bishop, J.P. McGrath, A.D. Levinson, *Mol. Cell Biol.*, **5**, 1772 (1985).
95. M.P. Kamps, B.M. Sefton, *Mol. Cell Biol.*, **6**, 751 (1986).
96. S.M. Read, D.P. Delmer, *Plant Physiol.*, **85**, 1008 (1987).
97. N.A. Morjana, C. Lyons, T.G. Flynn, *J. Biol. Chem.*, **264**, 2912 (1989).
98. P. Harinasut, T. Takabe, T. Akazawa, M. Tagaya, T. Fukui, *Plant Physiol.*, **88**, 119 (1988).
99. P. Dominici, G. Scholz, F. Kwok, J.E. Churchich, *J. Biol. Chem.*, **263**, 14712 (1988).
100. M. Tagaya, K. Yamano, T. Fukui, *Biochemistry*, **28**, 4670 (1989).
101. T. Hibi, H. Kato, T. Nishioka, J. Oda, H. Yamaguchi, Y. Katsube, T. Tanizawa, T. Fukui, *Biochemistry*, **32**, 1548 (1993).
102. H.M. Mizioroko, C.A. Brodt, T.J. Krieger, *J. Biol. Chem.*, **265**, 3642 (1990).
103. Y. Kazuta, M. Tagaya, K. Tanizawa, T. Fukui, *Protein Sci.*, **2**, 119 (1993).

5

Structure and Function of Aspartate Aminotransferase

Hiroshi Ueno and Hiroyuki Kagamiyama

*Department of Medical Chemistry, Osaka Medical College
Takatsuki, Osaka 569, Japan*

5.1 Introduction

What actually occurs during enzyme catalysis? How does the enzyme recognize its substrates and how can substrate binding transmit the message to the rest of the molecule to initiate catalytic action? Those of us who study enzymes have approached these problems with the available biochemical and analytical techniques. Recently, we have utilized a genetic engineering technique called site-directed (or *in vitro*) mutagenesis to define the role of specific amino acid residues in enzyme functions. New insights into the understanding of the enzyme catalysis of aspartate aminotransferase (AspAT) are reviewed in this chapter, particularly those provided by recent site-directed mutagenesis studies. AspAT, which catalyzes the transamination reaction between dicarboxylic amino acid and the corresponding ketoacid, has been the model enzyme for studying enzyme mechanisms. This is because AspAT has several advantages over the other enzymes. 1) Pyridoxal 5'-phosphate (PLP), a coenzyme in AspAT, is a highly sensitive reporter molecule. An optical absorption of this chromophore represents various ionic states which reflect the environmental changes at the active site of AspAT. 2) Both chemically and enzymologically, AspAT is the most extensively studied pyridoxal enzyme. 3) It is isolated from many different sources, *i.e.* cytosol and mitochondria of mammals, bacteria and other species. 4) The amino acid sequences of the isolated AspATs are known and are highly homologous, particularly those residues located at the active site and coenzyme binding site. 5) X-ray crystallographic data for each of the AspATs are now available for detailed evaluation in enzyme catalysis. 6) *In vitro* mutagenesis systems in both mammalian and bacterial AspAT have been established. The catalytic as well as substrate recognition mechanisms on AspAT are also discussed in this review (see Christen and Metzler¹⁾ for a general collection of reviews on various aminotransferases).

5.2 Background

AspAT is a PLP-dependent enzyme. This dimeric enzyme contains one PLP per subunit whose molecular weight is about 50,000.

Overall reaction catalyzed by AspAT is as follows:

	10	20	30	40	50	60	70	80
PcAsp	APPSVFAEVPQAQPVLFVKLIADREFDPDPKRVNLGVGAYTDDCQPWVLVVRKVVEQRIANDSSLNHEYLPIGLAEFFTCASR							
CcAsp	AASIFAAPVRAPPAVEFKLTADREFDGSRSKRVNLGVGAYTDEGQPWVLVVRKVVEQLIAGNSLNHEYLPIGLPEFRANASR							
HcAsp	TSPSIFVEVPQAQPVLFVKLTADREFDPDPKRVNLGVGAYTDDCQPWVLVVRKVVEQKIANNSSLNHEYLPIGLAEFFRSCASR							
RcAsp	APPSFFAQVPQAQPVLFVKLIADREFDPDPKRVNLGVGAYTDDSQPWVLVVRKVVEQKIANDSHSLNHEYLPIGLAEFFRSCASQ							
McAsp	TSPSIFVEVPQAQPVLFVKLTADREFDPDPKRVNLGVGAYTDDCQPWVLVVRKVVEQKIANNSSLNHEYLPIGLAEFFRSCASR							
PmAsp	SSWVAHVENGPPDPILGVTEAFKRDTSKKMNLGVGAYRDDNGKPYVLPVSRKAEAAQIAAKN-LDKEYLPIGGLAEFCOKASAE							
CmAsp	--SSWWSHVENGPPDPILGVTEAFKRDTSKKMNLGVGAYRDDNGKPYVLCNVCVRKAEAMIAAKK-MDKEYLPIAGLADFTFRASAE							
HumAsp	--SSWVAHVENGPPDPILGVTEAFKRDTSKKMNLGVGAYRDDNGKPYVLPVSRKAEAAQIAAKN-LDKEYLPIGGLAEFCOKASAE							
HmAsp	--SSWVAHVENGPPDPILGVTEAYKRDTSKKMNLGVGAYRDDNGKPYVLPVSRKAEAAQIAAKN-LDKEYLPIGGLAEFCOKASAE							
MmAsp	--SSWVAHVENGPPDPILGVTEAFKRDTSKKMNLGVGAYRDDNGKPYVLPVSRKAEAAQIAAKN-LDKEYLPIGGLAEFCOKASAE							
eAsp	MFENITAAPADPILGLADLFRADERPGKINLGIGVYKDETGKTPVLTSPVKAEQIYLLNE-TTKNYLIGIDGIPERFORCTQE							
tAsp	MKELLANRVKTLTPSTTLAITAKAKEMKAQGIDVIGLGAGEPDENTPQNIMDAADSMQOGYTKYTPSPSGGLPALKQAQIE							
	90	100	110	120	130	140	150	160
PcAsp	LALGDDSPALQEKRVGVSQSLGGTGALRIGAEFLARWYNGTNNKDTPVYVSSPTWENENGVFTTAGFKDIRSYRYWDTEKRGGLDL							
CcAsp	IALGDDSPAIAQKRVGVSQSLGGTGALRIGAEFL-RWYNGNNTATPVYVSSPTWENENSVFMDAGFKDIRTYRYWDAAKRGGLDL							
HcAsp	LALGDDSPALQEKRVGVSQSLGGTGALRIGAEFLSRWYNGTNNKNTPVYVSSPTWENENGVFSGAGFKDIRSYHYWDATKRGGLDL							
RcAsp	LVLGDNSPALRENRVGSQSLGGTGALRIGADFLRWYNGTNNKNTPVYVSSPTWENENGVFSAAAGFKDIRSYRYWDAEKRGGLDL							
McAsp	LALGDDSPALQEKRVGVSQSLGGTGALRIGAEFLSRWYNGTNNKNTPVYVSSPTWENENGVFSGAGFKDIRSYHYWDATKRGGLDL							
PmAsp	LALGENNEVLKSGRYVTQISGTGALRIGANFLQRFEFKES--RD--VELPKPSWGNBTPIFRDAGM-QLHSRYRYDPKTCGFDF							
CmAsp	LALGENSEAFKSGRYVTQISGTGSLRVGANFLQRFEFKES--RD--VYLKPKSWGNBTPIFRDAGL-ELQAYRYYDPKTCGLDF							
RmAsp	LALGENSEVLKSGRFVTQTIISGTGALRVGASFLQRFEFKES--RD--VELPKPSWGNBTEIARDAGM-QLEGYRYYDPKTCGFDF							
HumAsp	LALGENSEVLKSGRFVTQTIISGTGALRIGASFLQRFEFKES--RD--VELPKPTWGNBTPIFRDAGM-QLQGYRYYDPKTCGFDF							
HmAsp	LALGENSEALKSGRYVTQISGTGALRIGANFLQRFEFKES--RD--VELPKPSWGNBTPIFRDAGL-QLHAYRYYDPKTCGFDF							
MmAsp	LALGENNEVLKSGRFVTQTIISGTGALRVGASFLQRFEFKES--RD--VELPKPSWGNBTPIFRDAGM-QLQGYRYYDPKTCGFDF							
eAsp	LLFGKGSALINDKRARTAQTPGGTGALRVAADFLAKNTSVK--R--VWVSNPSWPNHKSVFNSAGL-EVREYAYYDAENHHTLDF							
tAsp	KFKRDNQLEYKP-----NEIIVGVGAKHVLYTLFQVILNEG--D-EVIIPIPYWVSYPEQVKLGG--VPVYIATSS--EQNYK							
	180	190	200	210	220	230	240	250
PcAsp	QGFLSDLENAPEFSIFVLHACAHNPTGTDPTPEQWKQIASVMKRRFLFPFEDSAYQGFASGNLEKDAWAIRYFVSEGFELFCAQS							
CcAsp	QGLLSDMEKAPESIFILHACAHNPTGTDPTPEQWKQIAAVMKRRCLFPFEDSAYQGFASGNLEKDAWAVRYFVSEGFELFCAQS							
HcAsp	QGFLNDLENAPEFSIFVLHACAHNPTGTDPTPEQWKQIASVMKRRFLFPFEDSAYQGFASGNLDRDAWAVRYFVSEGFELFCAQS							
RcAsp	QGFLNDLENAPEFSIFVLHACAHNPTGTDPTPEQWKQIAAVMKRRFLFPFEDSAYQGFASGNLDRDAWAVRYFVSEGFELFCAQS							
McAsp	QGFLNDLENAPEFSIFVLHACAHNPTGTDPTPEQWKQIASVMKRRFLFPFEDSAYQGFASGNLDRDAWAVRYFVSEGFELFCAQS							
PmAsp	TGALEDISKIPAQSVILLHACAHNPTGVDPRPEQWKEMATLVKNNLFAFDIMAYQGFASGDNKDAWAVRHFIEQGINVCLQS							
CmAsp	TGAMEDISKIPEKSIILLHACAHNPTGVDPREEQWKELASVVKNNLLAFDMAYQGFASGDNINRDAWALRHFIEQGINVCLQS							
RmAsp	SGALEDISKIPEQSVLLHACAHNPTGVDPRPEQWKEMATLVKNNLFAFDIMAYQGFASGDNKDAWAVRHFIEQGINVCLQS							
HumAsp	TGAVEDISKIPEQSVLLHACAHNPTGVDPRPEQWKEMATLVKNNLFAFDIMAYQGFASGDNKDAWAVRHFIEQGINVCLQS							
HmAsp	TGALEDISKIPEQSVLLHACAHNPTGVDPRPEQWKEMATLVKNNLFAFDIMAYQGFASGDNKDAWAVRHFIEQGINVCLQS							
MmAsp	SGALEDISKIPEQSVLLHACAHNPTGVDPRPEQWKEMATLVKNNLFAFDIMAYQGFASGDNKDAWAVRHFIEQGINVCLQS							
eAsp	DALINSLNEAQAGDVLLFHGCGCHNPTGIDPTLEQWQTLAQLSVKKGWLPULFDFAIQGFARG--LEEDAEGLRAPFAAMHKELIVASS							
tAsp	ITAEQLKNAITDKTKAVIINSPSNPTGMVYTRREELEDAIKALENNILIVSDEIYEKLLYN-GAEHFSIAQISEEVKQATIVING							

	260	270	280	290	300	310	320	330	340
PcAsp	FSKNFGLYNERVGNLT	TVVAK	EPDSILRVLSQMEKIV	RVTVTSNPPAQGARIV	ARTLSDP	ELFHEW	TGNVKT	MADRIL	SMRSEL
CcAsp	FSKNFGLYNERVGNLT	SVGKDEDNVQ	RVLSQMEKIV	RVTVTSNPPSQGARIV	ATLTSD	PGFAEK	WKNVKT	MADRIL	SMRSEL
HcAsp	FSKNFGLYNERVGNLT	TVVAK	EPDSILRVLSQMEKIV	RVTVTSNPPAQGARIV	AFVATL	SDPGLFKE	WKNVKT	MADRIL	SMRSEL
RcAsp	FSKNFGLYNERVGNLT	VVGEHDSV	RVLSQMEKIV	RVTVTSNPPAQGARIV	ATLTSL	NPELFKE	WKNVKT	MADRIL	SMRSEL
McAsp	FSKNFGLYNERVGNLT	TVVAK	EPDSILRVLSQMEKIV	RVTVTSNPPAQGARIV	AFVATL	SDPGLFKE	WKNVKT	MADRIL	SMRSEL
PmAsp	YAKNMGLYGERVGAFT	VVCKDAE	EAKRVESQ	LKILIRPMY	SNPPNGAR	IASTIL	SPDLRQ	QWLEVK	GMADRIIS
CmAsp	YAKNMGLYGERVGAFT	VVICRDAE	EAKRVESQ	LKILIRPMY	SNPPNGAR	IASTIL	SPDLRQ	QWLEVK	GMADRIIS
RmAsp	YAKNMGLYGERVGAFT	VVCKDAE	EAKRVESQ	LKILIRPLY	SNPLNGAR	IAATIL	SPDLRQ	QWLEVK	GMADRIIS
HumAsp	YAKNMGLYGERVGAFT	VMVCKDAE	EAKRVESQ	LKILIRPMY	SNPLNGAR	IAAAI	NTPD	LKQWLEVK	GMADRIIS
HmAsp	YAKNMGLYGERVGAFT	VMVCKDAE	EAKRVESQ	LKILIRPLY	SNPLNGAR	IASTIL	SPDLRQ	QWLEVK	GMADRIIS
MmAsp	YAKNMGLYGERVGAFT	VMVCKDAE	EAKRVESQ	LKILIRPLY	SNPLNGAR	IASTIL	SPDLRQ	QWLEVK	GMADRIIS
eAsp	YSKNFGLYNERVGAFT	LVADSETV	DRAFSQMKAIR	ANYSNPPAHGASV	VATILS	NDAIRAI	WEQELT	DMRQRI	QMRQLFVNT
tAsp	VSKSHMTGWRIGYAAG	NAD----	IINAMTDLASH	TSTNPTASQYAA	EAYNGPQDS	VEEMRK	AFESR	LETIY	PKLSAIPGFK
	350	360	370	380	390	400	410		
PcAsp	LEALKTPTGWNHIT	DQIGMF	SFTGLNPKQVE	--YLINEKH	IYLLPSG	-----	RINMCG	LTTKNLDY	VATSIHEAVTKIQ
CcAsp	LESLGTPGTWNHIT	DQIGMF	SFTGLNPKQVE	--YMIKEKH	IYLMASG	-----	RINMCG	LTTKNLDY	VAKSIHEAVTKIQ
HcAsp	LEALKTPTGWNHIT	EQIGMF	SFTGLNPKQVE	--YLVNQKH	IYLLPSG	-----	RINMCG	LTTKNLDY	VATSIHEAVTKFQ
RcAsp	LEALKTPTGWSHIT	EQIGMF	SFTGLNPKQVE	--YLVNEKH	IYLLPSG	-----	RINMCG	LTTKNLDY	VATSIHEAVTKFQ
McAsp	LEALKTPTGWNHIT	EQIGMF	SFTGLNPKQVE	--YLVNQKH	IYLLPSG	-----	RINMCG	LTTKNLDY	VATSIHEAVTKFQ
PmAsp	LKKEGSSHNWQHIT	DQIGMF	CFGTGKPEQVE	--RLTKEFS	IYMTKDG	-----	RI	SVAGVT	SGNVGYLAHAHQ-VTK
CmAsp	LKKEGSSHNWQHIT	DQIGMF	CFGTGKPEQVE	--RLTKEFS	IYMTKDG	-----	RI	SVAGVT	SGNVGYLAHAHQ-VTK
RmAsp	LKKEGSSHNWQHIT	DQIGMF	CFGTGKPEQVE	--RLTKEFS	IYMTKDG	-----	RI	SVAGVT	SGNVGYLAHAHQ-VTK
HumAsp	LKKEGSSHNWQHIT	DQIGMF	CFGTGKPEQVE	--RLTKEFS	IYMTKDG	-----	RI	SVAGVT	SGNVGYLAHAHQ-VTK
HmAsp	LKKEGSSHNWQHIT	DQIGMF	CFGTGKPEQVE	--RLTKEFS	IYMTKDG	-----	RI	SVAGVT	SGNVGYLAHAHQ-VTK
MmAsp	LKKEGSSHNWQHIT	DQIGMF	CFGTGKPEQVE	--RLTKEFS	IYMTKDG	-----	RI	SVAGVT	SGNVGYLAHAHQ-VTK
eAsp	LQEKGANRDFSTIK	QNGMFS	SGLTKEQVL	--RLREEF	GVAVASG	-----	RVN	VAGMT	PDNMAPLCEAIVA-VL
tAsp	VVKPQGAFYLLPDV	SEAAQKTGFA	--SVDE	FASALLTEAN	VIPSGFGAP	STIRISYA	--TSLN	LIE--	EAIERIDRFVK

Fig. 5.1 Primary structures of AspATs from various species. P, C, H, Hu, R and M denote for pig, chicken, hamster, human, rat and mouse, respectively. c and m stand for cytosolic and mitochondrial, respectively. eAspAT and tAspAT denote *Escherichia coli* AspAT and Thermotable AspAT from Thermophilic *Bacillus species*⁽⁹⁾, respectively. Letters in bold face type represent residues discussed in the text.



where Asp, α -KG, Glu and OA represent aspartate, α -ketoglutarate, glutamate and oxalacetate, respectively. This overall transamination reaction is divided into two half reactions:



where EL and EM stand for the PLP and PMP forms of the enzyme. These forms exhibit absorption maxima at 360 and 330 nm, respectively. Therefore, reactions shown in Eqs. (5.2) and (5.3) can be conveniently studied on a stopped-flow spectrometry.

In aminotransferases, PLP is usually attached to the enzyme via a Schiff base (an internal aldimine bond) with a Lys residue (Lys-258 for AspAT), and facilitates the reaction by the electron-withdrawing action of the pyridine moiety, labilizing the bonds around the α -carbon of the substrate through the conjugation system of the Schiff base. The overall reaction efficiency and specificity are determined by the functional groups of the protein. The pK_a of the Schiff base is very sensitive to its environment. The factors influencing Schiff base pK_a are neighboring amino acid residues, substrates, conformational change of the enzyme, *etc.* Therefore, considerable effort has been devoted to the estimation of the pK_a . Spectrophotometric titration of the enzyme showed pK_a of the Schiff base to be 6.3–6.7.²⁻⁴⁾ Both ^{13}C NMR spectroscopy⁵⁾ and 500-MHz ^1H NMR spectroscopy (the latter was measured in H_2O ⁶⁾) showed it to be 6.3. Enzyme crystals could also provide reliable pK_a . Metzler *et al.* titrated a single crystal and measured its polarized spectral shifts on pH to give a pK_a of 6.45.⁷⁾ This value agrees with pK_a values obtained by other methods.

AspAT is found in a wide variety of species, from microorganisms to higher animals and even in plants. In both animals and plants, this enzyme activity is found in cytosol and mitochondria. The amino acid sequences determined by Ovchinnikov *et al.*⁸⁾ for pig heart cytosolic AspAT (cAspAT) in 1973 and by Kagamiyama *et al.*⁹⁾ for pig heart mitochondrial AspAT (mAspAT) in 1977 revealed that both forms are constructed with over 400 amino acid residues and exhibits 40% sequence homology. Thus, cAspAT and mAspAT are isozymes. Amino acid sequence of *E. coli* AspAT (eAspAT) showed 40% homology with both mAspAT and cAspAT.¹⁰⁾ Three enzymes exhibited 25% sequence homology among them, and this 25% homology was equivalent to about 100 amino acid residues (see Chapter I of this Volume for details).

Primary sequences of AspAT from a variety of species are summarized by Mehta, Hale and Christen.¹¹⁾ Fig. 5.1 shows the results of the sequence alignment of 13 AspATs. Those amino acid residues conserved among all AspATs and known for their functions are listed in Table 5.1. Nuclear DNA encodes for both mAspAT and cAspAT. Entire genes of mAspAT and cAspAT of chicken recently sequenced by Juretic *et al.*¹²⁾ showed that nine exons encoded the mature enzyme (Fig. 5.2). Similar analysis was performed on mouse AspAT by Obaru *et al.*¹³⁾ Both Obaru *et al.* and Juretic *et al.* demonstrated the structural homology of genes encoding cAspAT and mAspAT, in which the five introns, out of 8 and 9 for cAspAT and mAspAT, respectively, were placed at the same positions in

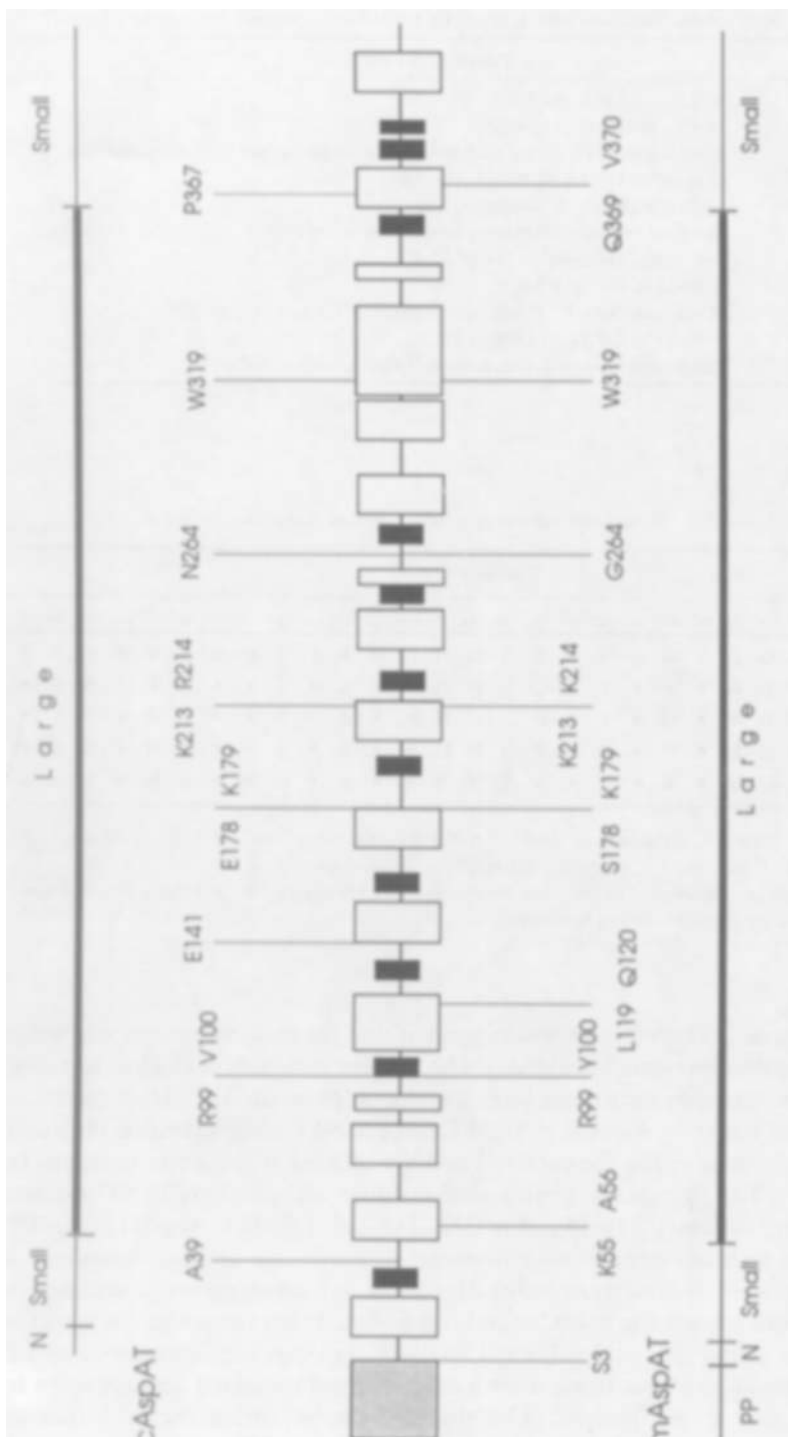


Fig. 5.2 Spatial orientation of secondary structure of AspAT exons. The shaded N-terminal box indicates the presequence region; PP, prepiece. Locations of introns are denoted by vertical lines and the exon/intron boundary amino acid residues are indicated in one-letter code. The numbering of the residues corresponds to those in pig cAspAT. α -Helices are represented by open boxes and β -sheets by black boxes. The domains of the proteins are also indicated: N, N-terminal; Small, small domain; Large, large domain. Results reproduced from Juretic *et al.*¹²⁾

TABLE 5.1 Roles of Invariant Residues with Known Functions in Aspartate Aminotransferase

Residue	Postulated Functions
Tyr-70*	H-bond to OP2 of PLP
Pro-138	<i>cis</i> -Pro involved in bending
Trp-140	Stacking with PLP ring and H-bond to distal carboxyl of the substrate
Asn-194	H-bond to O3' of PLP and Arg-386
Pro-195	<i>cis</i> -Pro involved in bending
Gly-197	Involved for loop formation at domain interface
Asp-222	Salt bridge/H-bond to N1 of PLP
Tyr-225	H-bond to O3' of PLP
Lys-258	Schiff base with PLP and participates in 1,3 prototropic shift
Arg-266	H-bond to OP2 and OP4 of PLP
Arg-386	Salt bridge/H-bond with α -carboxylate group of substrate

TABLE 5.2 Residues Forming the Active Site of AspATs

	Residue No. ^{†2}																																		
ATs ^{†1}	15	17	18	37	38	39	70*	107	108	109	140	141	142	143	193	194	222	224	225	255	258	266	292*	293*	296*	297*	360	386							
cAspAT	V	V	F	V	G	A	Y	G	G	T	W	E	N	H	H	N	D	A	Y	S	K	R	R	V	S	N	F	R							
mAspAT	D	I	L	V	G	A	•	S	•	T	•	G	N	H	H	•	•	A	•	S	•	•	R	P	•	N	F	•							
eAspAT	D	I	L	I	G	V	•	G	•	T	•	P	N	H	H	•	•	A	•	S	•	•	R	A	•	N	F	•							
eAroAT	D	I	L	I	G	L	•	G	•	S	•	E	N	H	H	•	•	A	•	S	•	•	R	R	•	S	F	•							
tAspAT	K	L	T	I	D	V	•	V	•	V	•	V	S	Y	S	•	•	I	•	G	•	•	S	H	•	N	A	•							

†1 Abbreviations used: cAspAT, pig cytosolic AspAT; mAspAT, pig mitochondrial AspAT; eAspAT, *E. coli* AspAT; eAroAT, *E. coli* AroAT; tAspAT, thermophilic *Bacillus* species.

†2 Conserved amino acid residues found at the same residue number are marked ●, and those residues with high degree of replacement are shown in bold face.

the amino acid sequence. Therefore, it was suggested that the introns were placed before gene divergence of the two isozymes occurred. The tertiary structure of chicken mAspAT indicated the presence of two domains, one large and the other small.^{14,15} High resolution X-ray crystallographic data on cAspAT, mAspAT and eAspAT have identified the amino acid residues which are involved in the catalysis and are located at the active center of the enzyme (Table 5.2). The important amino acid residues include Tyr-70* (* indicates residues from opposite subunit), Trp-140, Asp-222, Tyr-225, Lys-258, Arg-266, Arg-292* and Arg-386. These residues interact with substrate or coenzyme and are thought to be involved in the catalysis either directly or indirectly. The tertiary structures elucidated for cAspAT, mAspAT, and eAspAT are similar and it is believed that the active site structure among the AspATs is nearly identical. Recent analysis of cDNA sequences from over 10 species of AspATs has shown that those amino acid residues indicated above except for Arg-292* are conserved in all species. The studies have solidified the idea that the conserved amino acid residues are involved in the functionality of the enzyme.

5.3 Site-directed Mutagenesis on AspAT

Both mammalian and bacterial enzymes are expressed in *E. coli* system.¹⁶⁻²⁰⁾ However, the major part of the recent mutagenesis studies has been carried out on *E. coli* enzyme. eAspAT gene (asp C) is cloned into a high-copy expression vector which produces 500-fold more AspAT than the wild-type strain. A strain (TY 103) lacking an original AspAT gene has been prepared in order to avoid the contamination of the wild-type enzyme when mutant enzyme with low enzymic activity is expressed.²⁰⁾

5.4 X-ray Crystallographic Study on AspAT

Three-dimensional X-ray crystallographic structures of AspATs are available for pig heart cAspAT,^{21,22)} chicken cAspAT,²³⁻²⁵⁾ chicken mAspAT,^{14,15,26-29)} and eAspAT.^{20,30-32)} Highly refined structure of chicken heart mAspAT (1.9 Å resolution) in high pH form is now available together with two other forms, one in low pH form (2.3 Å resolution) and the other in PMP form (2.2 Å resolution).²⁷⁾ As expected from the high sequence homology between *E. coli* and animal AspATs, many structural features are shared among these species. The subunit structure of AspAT is made of a large domain, composed of residues 47-324, and a small domain, composed of 12-46 and 325-C-terminal.²⁹⁾

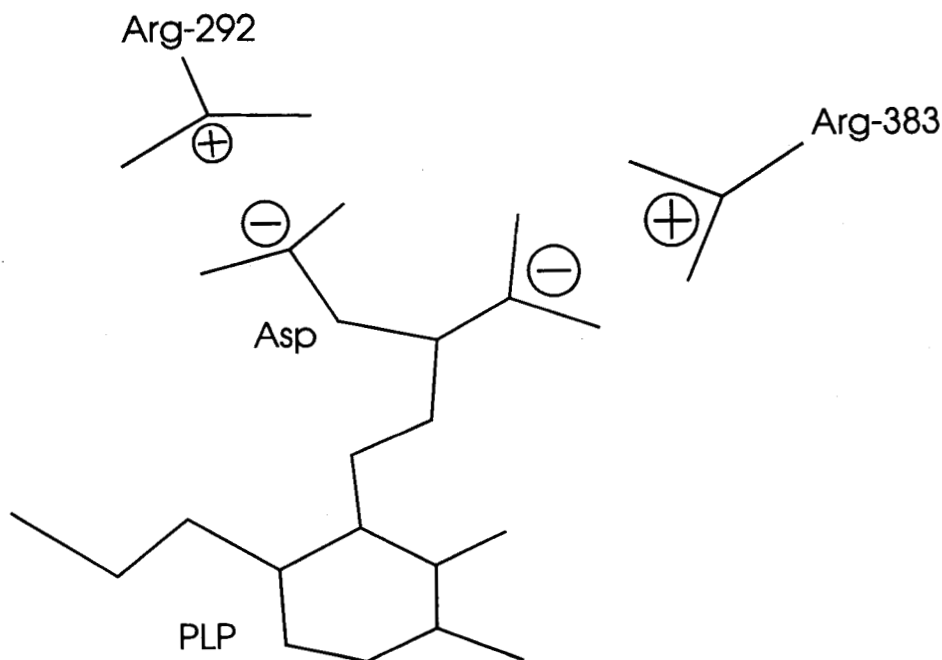


Fig. 5.3 Spatial orientation of substrate binding site of AspAT. A substrate, Asp, which forms an external aldimine complex with PLP interacts with Arg-292 and Arg-386 through ionic interactions.

About 10% of the β -sheet participates in the coenzyme binding sites. The crystal structure of α -methyl-aspartate (α -Me-Asp)·enzyme complex demonstrates a clear picture of how enzyme interacts with substrate. Both α - and distal carboxylate groups form ionic interactions with Arg-386 and Arg-292*, respectively (Fig. 5.3). PLP is stabilized by the hydrogen bonding with Tyr-225 and Asn-194 through 3'-OH, by the ionic interaction with Asp-222 via $N(1)$, and also by the ionic interaction with α -helix (108–122), Tyr-70 and Arg-266 via phosphate moiety. In the absence of substrate, PLP is also stabilized by the Schiff base linkage with the ϵ -amino group of Lys-258. The aromatic side chain of Trp-140 is stacked with a PLP ring to form a van der Waals interaction, determining the orientation of the PLP ring during the catalysis. It is also evident that the coenzyme ring tilts along the $N(1)$ and phosphate group upon the binding of substrate. This idea of ring tilt was originally proposed by Ivanov and Karpeisky in 1969.³³⁾

High resolution crystallographic studies have demonstrated that AspAT has open and closed conformations. Upon the binding of a substrate which favors the closed conformation, the small domain rotates so that the active site residues are brought closer to the substrate moiety. This movement of the small domain is important for initiation of the catalytic reaction. Therefore, the closed conformation is believed to represent the functional states of the enzyme.²⁹⁾ Three distinct crystal forms of mAspAT were analyzed, namely triclinic crystal for open conformation, monoclinic and orthorhombic crystals for closed conformation. The enzymes cocrystallized with substrate analogs seemed to have a tendency to assume the closed conformation.^{28,29)} The order of this tendency was shown to be apoenzyme[†] > PLP form of the enzyme > PMP form of the enzyme.

Although highly refined X-ray structures of AspAT are available, they do not provide an explanation for the subtle differences that occur when AspAT binds to various substrates. Such subtle differences have been distinguished by the kinetic analysis, however.

It is of interest to compare the tertiary structure of AspAT with that of other PLP-dependent enzymes. Some PLP enzymes whose primary structures are quite different from AspAT exhibit similar tertiary structures. Such enzymes are ω -amino acid:pyruvate aminotransferase,³⁴⁾ phosphoserine aminotransferase³⁵⁾ and tyrosine-phenol lyase³⁶⁾ (Phillips, R., personal communication). Similarity in tertiary structure among these PLP enzymes may lead to the idea that many PLP-dependent enzymes share the same ancestor protein. There are PLP enzymes belonging to its own category, such as glycogen phosphorylase and tryptophan synthase.^{37,38)} These enzymes do not share any similarities in either primary or tertiary structures with AspAT.

5.5 Evaluation of the Active Site Residues in AspAT by *in vitro* Mutagenesis

In vitro mutagenesis is used to evaluate how the individual residues are involved in enzyme catalysis; this is difficult to pursue with traditional methods, such as physico-chemical or chemical modification. Mutagenesis study has been successfully employed to define the role of active-site residues in AspAT. It was found that some of the active-site residues are involved in coenzyme binding, some are used for substrate recognition, and others are involved directly in catalysis.

[†] Apoenzyme has been shown to form cocrystals with substrate analogs.²⁹⁾

5.5.1 Lys-258 and Tyr-70*

Based on X-ray crystallographic results and other studies, the most important catalytic residues appear to be Lys-258 and Tyr-70*. This is because the α -proton of the substrate amino acid must be removed by the action of the catalytic base at the beginning of the catalytic process, and this step takes place stereospecifically at the *si*-plane of the PLP ring. Lys-258 and Tyr-70* are the only residues which are located within van der Waals contact distance from the α -proton and are able to participate in the proton abstraction. The results of various chemical modification and physico-chemical studies on wild-type enzyme, and high resolution X-ray crystallographic studies could not clearly define Lys-258 and Tyr-70* as a catalytic base. Therefore, mutagenesis studies on Lys-258 and Tyr-70* were used to define the roles of these residues.

The mutation at Lys-258 was carried out to replace Lys with Ala.^{16,39)} This mutant lost activity by more than 10^6 -fold compared to the wild-type enzyme. Other amino acids, such as Met and Arg, were also introduced into this position.¹⁶⁾ It was found that enzymic activity in these mutants diminished significantly. In contrast, spectroscopic studies showed that Lys-258 mutants incubated with Asp or β -hydroxy-Asp exhibited spectral changes almost identical with those of the wild-type enzyme incubated with a pseudo-substrate, α -Me-Asp. α -Me-Asp is known to give a Michaelis complex in which the catalytic process is seized prior to the α -proton abstraction step because of the absence of the α -proton in α -Me-Asp. This suggested that in the mutants the reaction was stopped at the step just prior to the α -proton removal. Both activity and spectrum studies supported the catalytic role of Lys-258.

The activity of K258A mutant was enhanced by adding exogenous amines which apparently replaced Lys-258.³⁹⁾ The results suggested that added exogenous amines appear to act as an acid/base catalyst to perform the 1,3-prototropic shift on the aldimine/ketimine interconversion.³⁹⁾ However, when a similar experiment conducted with an active-site mutant (K145Q) of the D-amino acid transaminase,⁴⁰⁾ it was shown that the rate of the external aldimine formation increased upon the addition of the exogenous amines while the transformation to the PMP form was unaffected. The studies on K145Q have indicated the important role of exogenous amines in raising the levels of external aldimines, thus enhancing the reaction rate. This enhancement, however, may not be caused by the action of amines as an acid/base catalyst. Recently Nishimura *et al.*⁴¹⁾ examined the effect of amines on the active-site lysyl mutants of D-amino acid transaminase and found that the effect of the exogenous amines was influenced by the type of amino acid residue at this position. A mutant having the basic amino acid substituted for Lys-145, K145R, showed no effects of exogenous amines on the formation of the external aldimine. On the other hand, Ala substitution mutant, K145A, clearly showed effects of exogenous amines. Thus, further functional studies are required to elucidate the role of exogenous amines.

Although K258A mutant is considered to be a "dead" enzyme, it has shown a small residual transaminase activity. The stereochemical fidelity of the K258A mutant enzyme was examined by measuring the residual transamination reaction in which the labilization of the pro-S C-4' hydrogen of PMP was monitored.⁴²⁾ During the normal transamination reaction, it was observed that the α -proton of the substrate was transferred to the pro-S C-4' position of PMP and the same proton was removed from the pro-S C-4' position then transferred to the second substrate, a part of the 1,3-prototropic shift in the transamination

reaction. In K258A mutant, this proton abstraction was not observed.⁴²⁾ The fact that K258A lost the ability to abstract the pro-S C-4' proton supports the idea that Lys-258 is the base participating in the 1,3-prototropic shift.

K258C-EA, which possessed an aminoethylated Cys residue in place of Lys-258, was created by a combination of site-directed mutagenesis and chemical modification.⁴³⁻⁴⁵⁾ The difference in size between the Lys side chain and the aminoethylated Cys was predicted to be very subtle. K258C-EA mutant prepared for the thermostable AspAT showed 25% activity of the wild-type enzyme and exhibited an absorption maximum at 375 nm.⁴³⁾ The results suggested the presence of the internal Schiff base in the mutant enzyme. When the same mutant was prepared for eAspAT, the catalytic activity was about 7% of the wild-type enzyme; however, this eAspAT K258C-EA mutant exhibited 10⁵ times more reactivity than K258C, from which K258C-EA mutant was derived. The pK_a of ϵ -NH₂ group of γ -thialysine was found to be shifted 1 pH unit lower than that of the wild-type enzyme.⁴⁵⁾ Thus it was reasoned that the almost 90% loss of the activity in K258C-EA mutant was due to this pK shift, which is considered to be a significant change for the residue acting as a catalytic base. Large kinetic isotope effects observed with K258C-EA suggested that the C α proton abstraction step was the rate-determining step for this mutant. These results strongly support the postulation that Lys-258 is the essential residue for the transamination catalysis.

Transaminase catalyzed reactions are known to be highly stereospecific. However, a racemization reaction catalyzed by AspAT was recently reported by Kochhar and Christen.⁴⁶⁾ They found that AspAT racemized dicarboxylic acid substrates very slowly in the presence of their cognate oxo acids. When Tyr, phe and Ala were used as substrates, racemization rates similar to that of the dicarboxylic natural substrates, Asp and Glu were observed. On the other hand, transamination reaction rates for these substrates were found to be 3-5 orders of magnitude slower than that for the dicarboxylic substrates. Kochhar and Christen suggested that the racemization reaction of AspAT was substrate-independent as they speculated that α -proton abstraction catalyzed by Lys-258 took place at the *si*-face and the proton return was carried out from the *re*-face by the water molecule. Similar asymmetric racemization reaction was recently reported for mandelate racemase, in which Lys and His share the catalytic role as proton acceptor and proton donor, respectively.⁴⁷⁾ Changes in the optical specificity during the transamination reaction were also observed on D-amino acid transaminase.⁴⁸⁾ This enzyme catalyzes a reversible transamination reaction between D-amino acid and α -ketoglutarate to form corresponding keto acid and D-Glu. The study showed that this enzyme could catalyze the slow transamination of L-Ala; the stereochemistry of this product, Glu, was identified to be the D-conformer. Therefore, the step adding the proton back to the substrate maintains stricter stereoselectivity than the step abstracting the proton.

Tyr-70* is conserved in all transaminases¹¹⁾; these transaminases utilize Glu $\cdot$ α -KG pair as substrates. X-ray studies have shown that Tyr-70* interacts with the phosphate part of the coenzyme. Y70*F mutant was prepared to examine whether or not Tyr-70* could be a catalytic base. The Y70F* mutant showed 15% of the activity of the wild-type enzyme; thus, it was quite unlikely that Tyr-70* was the catalytic base. However, Tyr-70* was involved in the interaction with the coenzyme, since the affinity towards the coenzyme was significantly reduced in Y70*F mutant.⁴⁹⁻⁵¹⁾ The equilibrium dissociation constants for the PMP $\cdot$ holoenzyme complex were 1.3 nM and 30 nM for the wild-type and mutant

enzymes, respectively. Also determined were the dissociation constants for the PLP·holoenzyme complex, which were 0.4 pM and 40 pM for the wild-type and mutant, respectively.

The Y70*F mutant also reduced the formation of quinonoid intermediate by 75% in the reaction with *erythro*- β -hydroxy-Asp.⁵¹⁾ No explanation was given for this observation. Similar destabilization of the quinonoid species was reported by the mutagenesis in the other part of the molecule, such as R386K mutant.⁵²⁾ Taylor, Metzler and Arnone, however, suggested possible hydrogen bonding between the phenolic oxygen atom of Tyr-70* and β -O of the *erythro*- β -hydroxy-Asp based on the molecular modeling.⁵³⁾ Therefore, Tyr-70* may play an important role in stabilizing the quinonoid structure.

A recent study by Inoue *et al.*⁵⁴⁾ showed that the OH group of Tyr-70* was involved in the stabilization of the transition state (about 2 kcal/mol) of the enzyme·substrate complex. Moreover, it was shown that the benzene ring of Tyr-70* was required for the recognition of five-carbon substrates, *i.e.* Glue and α -KG, rather than four-carbon substrates, *i.e.*, Asp and OA. Although, the involvement of Tyr-70* in the catalytic process has been questioned,⁵²⁾ it is suggested that in some cases Tyr-70* is involved in the catalytic event.

5.5.2 Tyr-225

Tyr-225 forms a hydrogen bond with 3'-O⁻ of PLP. Y225F, a Tyr-225 mutant prepared by Inoue *et al.*, was unable to form this hydrogen bond.⁵⁵⁾ This mutant showed identical structural conformation with the wild-type enzyme; however, the activity was decreased to 0.3% that of the wild-type enzyme. A large decrease in the k_{cat} and k_{max} values for Y225F mutant suggests that the hydrogen bond between 3'-O⁻ and Tyr-225 plays an important role in the catalysis. The $\text{p}K_{\text{a}}$ for the internal Schiff base of Y225F increased from 6.7 to 8.6 in the absence of the substrate.⁵⁵⁾ This increase in $\text{p}K_{\text{a}}$ suggests that Tyr-225 may lower the electron density on PLP by forming a hydrogen bond between 3'-O⁻ and Tyr-225. This hydrogen bonding is important to stabilize the negative charge on 3'-O⁻; it contributes to the stabilization of the conjugated structure of the external Schiff base as well as to the removal of the α -proton by reducing the electron density at the C_{α} position (Fig. 5.4). Thus, it is believed that the lack of the catalytic activity in the Y225F mutant is due to the inability to form this hydrogen bond. Goldberg *et al.*⁵⁶⁾ also studied the properties of the Y225F mutant. They found an increase in $\text{p}K_{\text{a}}$ from 7.1 (wild-type) to 8.4 by plotting a $k_{\text{cat}}/K_{\text{Asp}}$ vs. pH profile. This shift in $\text{p}K_{\text{a}}$ was similar to the one observed by Inoue *et al.*⁵⁵⁾ They note that the rate-determining step in Y225F mutant is the hydrolysis step of the

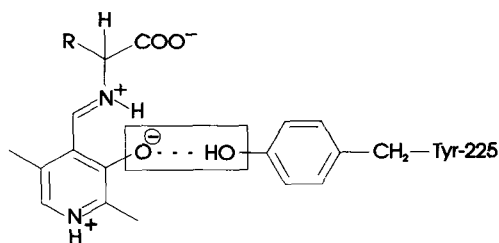


Fig. 5.4 Hydrogen bond formation between 3'(O) of PLP and Tyr-225.

ketimine intermediate rather than the α -proton abstraction process.⁵⁶⁾

5.5.3 Asp-222 and His-143

Resonance Raman, optical and other spectroscopic analyses have shown that the $N(1)$ position of PLP is always positively charged and the $O(3')$ position is mostly negatively charged.^{1,22)} In aqueous media, pK_a of $N(1)$ is usually around 5 to 6, which means that $N(1)$ pK_a has to be raised in order to remain protonated at the physiological pH. It is speculated that the hydrogen bond formation between Asp-222 and $N(1)$ can influence the pK_a of $N(1)$ position. The mutagenesis of Asp-222 to Ala or Asn diminished enzymic activity but D222E mutant exhibited 10% activity for the forward reaction with Asp as a substrate.⁵⁷⁾ The reverse reaction of the PMP form of D222E enzyme with α -KG showed even higher activity than the wild-type enzyme. It was also shown that the α -proton abstraction step was significantly disturbed in both D222A and D222N mutants. These results suggest that the negative charge on Asp-222 stabilizes the positive charge on $N(1)$ and also aids the abstraction of an α -proton from the substrate during the catalysis.

The binding affinity for PMP was greatly diminished (by 3 orders of magnitude) in D222A and D222N mutants. This suggests that Asp-222 is catalytically important not only for forming an electrostatic interaction with the $N(1)$ position but also for supporting the coenzyme ring in position.

X-ray crystallographic studies have shown that the His-143 forms a hydrogen bond with Asp-222. The role of the imidazole group on His-143 was thought to form a hydrogen bond network with Asp-222 and $N(1)$ and to stabilize the positive charge on the $N(1)$ position. The hydrogen bond between $N(1)$ and Asp-222 was assumed to be broken upon the binding of substrate, causing the coenzyme ring to tilt.²¹⁾ This ring tilt could create the motion at the $N(1)$ position of the coenzyme to move away from Asp-222; thus, hydrogen bond between Asp-222 and $N(1)$ broke and the adjacent hydrogen bond between Asp-222 and His-143 became tighter. Yano *et al.* have attempted a mutagenesis at His-143 to study the role of this residue.⁵⁸⁾ The mutant enzymes, H143A and H143N, maintained significant activity (above 60%) toward normal substrates. The effects of mutagenesis on k_{max}/K_d showed only 10-fold decrease, suggesting that His-143 did not participate directly in the catalysis. His-143 may be involved in assisting the formation of the external aldimine because the mutagenesis on His-143 does not effect the binding of succinate or glutarate, non-aldimine-forming analogs, whereas it does effect the binding of 2-methylaspartate, an external-aldimine-forming substrate analog.⁵⁸⁾ Nevertheless, it is now believed that the hydrogen bond between $N(1)$ of the coenzyme ring and Asp-222 does not break during the catalysis. This was also supported by the earlier work of Metzler and his associates on ^{19}F -NMR spectroscopy. Recently Metzler *et al.* have observed this hydrogen bond directly by ^1H -NMR spectroscopy measured in H_2O .^{6,59)} In their study, a nuclear Overhauser effect between these two protons, which are about 3.4 Å apart, helped identify the proton signals corresponding to $N(1)$ and His-143 protons. The existence of the hydrogen bond network among $N(1)$, Asp-222, and His-143 was proven by NMR.

5.5.4 Arg-292* and Arg-386

Both Arg-292* and Arg-386 are known to be involved in the recognition of dicarboxylic acid substrates. These residues are considered to be the critical residues for the substrate recognition mechanism in AspAT. X-ray crystallographic studies have shown that Arg-

386 interacts with the α -carboxyl group and Arg-292* interacts with the distal carboxyl group (Fig. 5.2).

The effect of the positive charge at Arg-292* was examined on an Arg-292* mutant enzyme, R292*K.⁶⁰⁾ The R292*K mutant showed a 10^3 -fold decrease in k -values and about a 10^2 -fold increase in k_d values. This indicates that both catalytic and substrate-binding processes are disturbed in R292*K mutant. Similar observation was made on Arg-386. This residue is found in all AspATs, and is primarily involved in the recognition of the α -carboxyl group of the substrates. The conservative mutagenesis of Arg-386 to Lys resulted in an extensive loss of enzyme activity, to 0.2–0.8% of the wild type and about a 10-fold-reduction in the binding ability toward its substrates.⁵²⁾ Although Lys residue maintained a positive charge similar to that of the arginyl residue, an ϵ -amino group of Lys was not structurally adequate for either Arg-292* or Arg-386 in maintaining efficient binding to the substrate carboxyl group. Instead, such efficient interaction was provided only with the guanidinium group of Arg residue.

Arg-292* was subjected to further study to examine the aromatic substrate binding nature of eAspAT. It is known that aromatic amino acids were relatively good substrates for this enzyme (Table 5.3). Series of mutants at Arg-292* were prepared as Arg being replaced by Lys, His, Gly, Ala, Ser, Gln, Val, Leu, Met, Tyr, Phe, Asp, and Glu.^{60–62)} Studies on these mutant enzymes demonstrated large decreases in catalytic efficiency (k_{cat}/K_m) for Asp or Glu, but increases for aromatic amino acids. In particular, R292*V mutant showed a 10^5 -fold reduction in k_{cat}/K_m values toward Asp and Glu. The same mutant, however, exhibited catalytic efficiency toward Trp ($k_{cat}/K_m = 10,100 \text{ M}^{-1} \cdot \text{s}^{-1}$) compatible with that of the wild-type enzyme toward Glu ($k_{cat}/K_m = 18,000 \text{ M}^{-1} \cdot \text{s}^{-1}$).⁶²⁾ It is of interest to note that no matter what amino acids are introduced to the Arg-292* position, there is no change in K_d values for Trp in any of these mutants. It seems reasonable to suggest that Arg-292* may not participate in accommodating the aromatic side chain as Sandmeier and Christen have previously noted.⁶³⁾

5.5.5 Trp-140

The indole ring of Trp-140 is located adjacent to the coenzyme ring in nearly a parallel manner. The role of this residue is to hold the coenzyme ring in place during the catalytic process. W140F mutant, which was prepared by a conservative mutagenesis on Trp-140, showed a significant decrease in catalytic efficiency: about 300-fold in the half reaction with Glu.⁶⁴⁾ Another mutant enzyme, W140G, showed further decrease in catalytic efficiency, about 10^4 -fold. The results support the original idea that the stacking interaction is present between the indole ring of Trp-140 and coenzyme ring.^{14,21)} In addition to the stacking effect, Trp-140 may be involved in the substrate binding process, since $N(1)$ of the indole ring of Trp-140 is located within a distance close enough to form a hydrogen bond with the distal carboxyl group of the substrate. When W140F and W140G mutants were examined, increases in K_d values, 80- and 20-fold for OA and α -KG, respectively, but not for Trp, were observed.⁶⁴⁾ This suggests the important role of the hydrogen bond between the indole ring and the carboxyl group of the dicarboxylic substrate.

5.5.6 Cys Residues

Recently Gloss, Planas and Kirsch addressed the role of Cys in eAspAT based on the assumption that some Cys residues may participate in the folding process.⁶⁵⁾ There are five

TABLE 5.3 Substrate Preference of Transaminases

Substrate	$K_{\max}/K_d$ ($M^{-1} s^{-1}$)				Ratio (<i>E. coli</i> AroAT:Pig cAspAT)
	Pig cAspAT	Pig mAspAT	<i>E. coli</i> AspAT	<i>E. coli</i> AroAT	
Asp	(100) †1	208	100	48.3	0.48
Glu	77.5	62.5	15	8.3	0.11
Trp	0.00026	0.002	0.73	425	1,600,000
Tyr	0.0019	0.014	0.33	917	483,000
Phe	0.00018	0.00225	0.17	1,667	9,300,000
His	0.00021	0.00022	0.011	0.558	2,700
Met	0.00092	0.00028	0.018	0.77	840
Leu	0.000036	0.000035	0.0020	0.375	10,400
Val	< 0.0000008	< 0.0000008	< 0.0000008	0.00014	180
Ile	< 0.000000008	< 0.000000008	< 0.0000008	< 0.0000008	10
Gly	< 0.000000008	< 0.000000008	< 0.0000008	< 0.0000008	10
Ser	0.000031	0.0000048	0.00004	0.000082	2.6
Thr	< 0.00000008	< 0.00000008	0.000007	0.000062	78
Ala	0.00012	0.000043	0.00064	0.0015	13
Asn	0.000041	0.00013	0.0012	0.015	370
Gln	0.000078	0.000028	0.00062	0.0021	27
Arg	< 0.00000008	< 0.00000008	0.000007	0.00043	540
Lys	< 0.00000008	< 0.00000008	0.00001	0.0004	500

†1 For comparison purposes, all numbers ($K_{\max}/K_d$) were normalized as pig cAspAT for Asp to be 100.

Cys residues in eAspAT. Only Cys-191 is conserved among mammalian and *E. coli* AspAT. The other four Cys residues in *E. coli* do not have any homology with the mammalian enzymes, except for Cys-82, which is found in pig cAspAT (Table 5.4). They mutated all five Cys residues individually and in combinations, and found that all Cys mutants showed the proper function and stability.⁶⁵⁾ Thus, it is unlikely that any of the five Cys residues are involved in the folding process. There are several Cys residues found in mAspAT. Among them, Cys-166 is the only thiol group exposed in the spatial structure of the protein. Giannattasio *et al.*⁶⁶⁾ have proposed a hypothesis that Cys-166 in mAspAT may be involved in the import mechanism of this isozyme which must be transported into mitochondria.

TABLE 5.4 Distribution of Cysteine Residues in Aspartate Aminotransferases

Isozyme	Residue Number														
	12	52	80	82	166	191	192	216	251	253	270	275	361	390	401
m Chicken	G	C	T	A	C	C	A	N	V	S	F	C	C	A	A
m Pig	G	S	C	A	C	C	A	N	C	C	F	C	C	A	A
m Human	C	S	C	A	C	C	A	N	C	C	F	C	C	A	A
c Chicken	A	V	R	N	R	C	A	C	F	A	L	G	S	C	A
c Pig	A	V	R	C	R	C	A	F	F	A	L	A	S	C	A
<i>E. coli</i>	A	S	G	C	H	C	C	G	I	A	C	A	S	A	C

So far no Cys residue has been demonstrated to be involved in AspAT functions. Is this a common phenomenon among transaminases? Merola *et al.*⁶⁷⁾ attempted to address the role of Cys residues in D-amino acid transaminase by site-directed mutagenesis and found that none of the Cys residues was essential for the catalysis. At the moment, it appears that no Cys residue is catalytically involved in pyridoxal-dependent transaminases.

5.6 Substrate Binding Pocket

AspAT prefers dicarboxylic amino acids as substrates. Animal AspATs have rather strict preference toward Asp and Glu as amino donor (Table 5.3). It has been shown that both Arg-292* and Arg-386 are involved in substrate recognition mechanism.¹⁵⁾ Of the two residues, Arg-292* plays a more critical role than Arg-386 in distinguishing the substrate, since it is the residue that recognizes the distal carboxyl group. Moreover, it is probably the residue that sends a triggering signal to the other parts of the protein so that it can initiate the shift from an open (less active) to a closed (more active) conformation. Picot *et al.*²⁹⁾ and more recently McPhalen *et al.*²⁸⁾ observed a crystalline form of mAspAT in the closed conformation. They found that the small domain underwent a rigid-body rotation of 13° for animal enzymes (6° for eAspAT³¹⁾) and such rotation closed the active site pocket. They also demonstrated the correlation of this movement in the small domain to the intensity changes of the covalently labeled fluorescent dye, monobromo-trimethyl-ammoniumbimane, at Cys-166. They have concluded that the conformational change that closes the active site is a ligand-induced event. This supports the notion that the movement of Arg-292*, which is the first residue to sense the incoming substrate, is the triggering mechanism for the conformational change. Kagamiyama and his associates have

addressed the role of Arg-292* in relation to this mechanism with the divergence in the substrate specificity of AspAT family.^{60,62} It has been shown that eAspAT has higher preference to neutral amino acids than other AspATs; although the values are small compared to Asp or Glu, they are quite significant (Table 5.3). Mutagenesis studies on Arg-292* described above suggest the presence of a second pocket designated for the neutral side chain recognition. *E. coli* aromatic amino acid transaminase (eAroAT), whose primary structure is similar to that of eAspAT, shows extremely high substrate preference towards aromatic amino acids (Table 5.3), and it exhibits AspAT activity; in other words, it binds to dicarboxylic substrates. Together with the amino acid sequence similarity, both AroAT and AspAT are considered to share the same ancestor protein. It is reasonable to assume that both enzymes possess two distinct binding pockets for substrates: one for a carboxyl side chain and other for a neutral side chain. There should be a mechanism for controlling the traffic so that the preferred side chain enters the correct pocket. In order to define the side-chain binding pocket, Kuramitsu examined the relationship between Gibbs free energy ($\Delta G_T^\ddagger$) of eAspAT (wild and R292V) and accessible surface area of the substrates⁶⁸ (Fig. 5.5). Fig. 5.5 gives three groups of straight lines: one for eAroAT having a steeper slope than the others for eAspAT's. Since the slope may represent a measure of the hydrophobic nature of the binding pocket which accommodates the side chain of a substrate, the binding pocket of AroAT was considered to be more hydrophobic (a larger slope) than that of eAspAT.

In 1988, before the crystal structure of eAspAT was available, Seville *et al.* postulated a potential binding pocket for aromatic side chain in both AspAT and AroAT by computer simulation, based on the atomic coordinate of mAspAT.⁶⁹ A possible pocket for Trp binding was porposed at the space between Trp-140 and Leu-18 side chains. They suggest

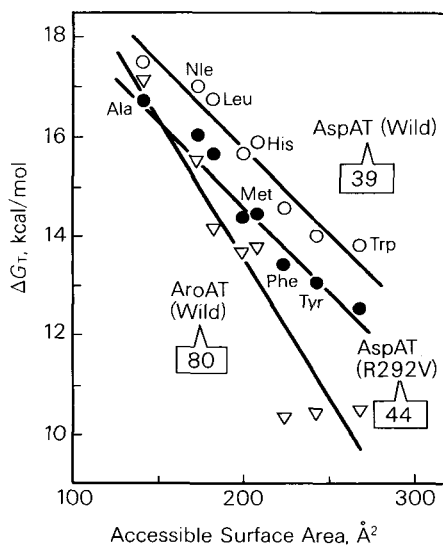


Fig. 5.5 Relationship between Gibbs free energy and accessible surface area of the substrates. Substrate amino acids are indicated as one-letter code. Boxed numbers show the slope of the straight line (cal/mol·Å²).

that Leu-18 and Trp-140 make walls so that the aromatic ring of the substrate can be sandwiched between the two walls. This Leu is commonly found among mAspAT, eAspAT and cAspAT, which are capable of catalyzing the transamination reaction with Trp. On the other hand, cAspAT, which is not able to catalyze the transamination reaction with Trp, has Phe at this position. This could explain the poor ability of cAspAT to act on aromatic substrates, where Phe is too bulky to allow the aromatic ring to fit into the space. In the above pocket, two residues at positions 37 and 39 were thought to be involved in substrate recognition, since different amino acids are found at these positions in different AspATs (Table 5.2, Fig. 5.6). At position 37, Val is found in cAspAT and mAspAT; Ile is found in eAspAT and eAroAT. At position 39, Ala is found in cAspAT and mAspAT, Val in eAspAT and Leu in eAroAT. There seem to be bulkier side chains in eAspAT and AroAT than in animal enzymes, so these bulky side chains may act as a guide for the incoming aromatic side chain. Although, the direct involvement of residues 37 and 39 in aromatic side-chain recognition has not been demonstrated, it is reasonable to assume that these positions play critical roles in the formation of the closed structure.

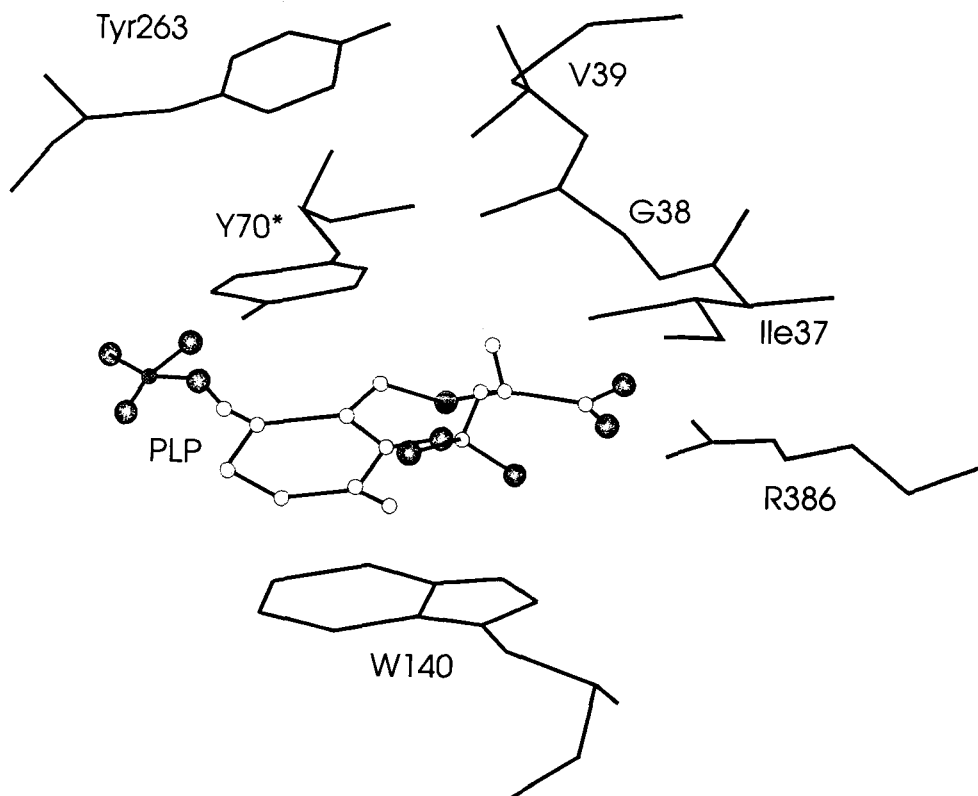


Fig. 5.6 Spatial orientation around G-39 at the active site of cAspAT.

5.7 Future Prospects

Our understanding of the enzymatic mechanism of AspAT has greatly increased since the introduction of site-directed mutagenesis. However, the mechanism is far from being understood completely. We have just begun to identify some critical residues at the active site of this enzyme. These residues are only a small portion of the 410 amino acid residues which comprise AspAT.

Studying enzymes, we tend to focus on an active site and not on other areas, assuming that the former is more important. Since protein is a dynamic molecule, it should be always treated as one unit. We should not neglect the rest of the molecule. In the future, mutagenesis studies should be extended to the whole enzyme molecule. The information gained from such mutagenesis studies as well as analytical approach to the whole enzyme should bring us closer to understanding how the enzyme works.

Acknowledgments

The authors wish to thank to Drs. K. Hirotsu, S. Tanase and A. Arnone for providing the X-ray crystallographic coordinates of AspATs, and to Dr. D. Yamamoto for assistance in molecular modeling, and to Y. Ueno for assistance in preparing the manuscript.

REFERENCES

1. P. Christen, D. E. Metzler, John Wiley-Interscience, New York, (1985).
2. W. T. Jenkins, L. D'Ari, *Biochemistry*, **5**, 2900 (1966).
3. R. G. Kallen, T. Korpela, A. E. Martell, Y. Matsushima, C. M. Metzler, D. E. Metzler, Y. M. Morozov, I. M. Ralston, F. A. Savin, Y. M. Torchinsky, H. Ueno, in: *Transaminases* (P. Christen, D. E. Metzler, eds.) J. Wiley & Sons, New York (1985) p. 37.
4. C. M. Metzler, D. E. Metzler, *Anal. Biochem.*, **166**, 313 (1987).
5. T. Higaki, S. Tanase, F. Nagashima, Y. Morino, A. I. Scott, H. J. Williams, N. J. Stolowich, *Biochemistry*, **30**, 2519 (1991).
6. A. Kintanar, C. M. Metzler, D. E. Metzler, R. D. Scott, *J. Biol. Chem.*, **266**, 17222 (1991).
7. C. M. Metzler, J. Mitra, D. E. Metzler, M. W. Makinen, C. C. Hyde, P. H. Rogers, A. Arnone, *J. Mol. Biol.*, **203**, 197 (1988).
8. Y. A. Ovchinnikov, C. A. Egorov, N. A. Aldanova, M. Y. Feigina, V. M. Kipkin, N. G. Abdulaev, E. V. Grishin, A. P. Kiselev, N. N. Modyanov, A. E. Braunstein, D. L. Polyanovsky, V. V. Nosikov, *FEBS Lett.*, **29**, 31 (1973).
9. H. Kagamiyama, R. Sakakibara, H. Wada, S. Tanase, Y. Morino, *J. Biochem.*, **82**, 291 (1977).
10. K. Kondo, S. Wakabayashi, H. Kagamiyama, *J. Biol. Chem.*, **262**, 8648 (1987).
11. P. K. Mehta, T. I. Hale, P. Christen, *Eur. J. Biochem.*, **186**, 249 (1989).
12. N. Juretic, U. Mattes, M. Ziak, P. Christen, R. Jaussi, *Eur. J. Biochem.*, **192**, 119 (1990).
13. K. Obaru, H. Nomiya, K. Shimada, F. Nagashima, Y. Morino, *J. Biol. Chem.*, **261**, 16876 (1986).
14. J. N. Jansonius, G. Eichele, G. C. Ford, D. Picot, C. Thaller, M. G. Vincent, in: *Transaminases*, (P. Christen, D. E. Metzler, eds.) J. Wiley & Sons, New York, (1985) p. 110.
15. G. C. Ford, G. Eichele, J. N. Jansonius, *Proc. Natl. Acad. Sci., USA*, **77**, 2559 (1980).
16. Y. Morino, K. Shimada, H. Kagamiyama, *Ann. NY Acad. Sci.*, **585**, 32 (1990).
17. T. Joh, H. Nomiya, S. Maeda, K. Shimada, Y. Morino, *Proc. Natl. Acad. Sci., USA*, **82**, 6065 (1985).
18. F. Nagashima, S. Tanase, Y. Fukumoto, T. Joh, H. Nomiya, T. Tsuzuki, K. Shimada, S. Kuramitsu, H. Kagamiyama, Y. Morino, *Biochemistry*, **28**, 1153 (1989).
19. M. -H. Sung, K. Tanizawa, H. Tanaka, S. Kuramitsu, H. Kagamiyama, K. Hirotsu, A. Okamoto, T. Higuchi, K. Soda, *J. Biol. Chem.*, **266**, 2567 (1991).

20. S. Kamitori, K. Hirotsu, T. Higuchi, K. Kondo, K. Inoue, S. Kuramitsu, H. Kagamiyama, Y. Higuchi, N. Yasuoka, M. Kusunoki, Y. Matsuura, *J. Biochem.*, **101**, 813 (1987).
21. A. Arnone, P. H. Rogers, C. C. Hyde, P. D. Briley, C. M. Metzler, D. E. Metzler, in: *Transaminases*, (P. Christen, D. E. Metzler, ed.) John Wiley & Sons, New York (1985) p. 138.
22. D. E. Metzler, C. M. Metzler, J. Mitra, R. D. Scott, A. Arnone, C. C. Hyde, P. Rogers, in: *Biochemistry of Vitamin B₆* (T. Korpela, P. Christen, eds.) Birkhaeuser Verlag, Basel (1987) p. 103.
23. V. V. Borisov, S. N. Borisova, G. S. Kachalova, N. I. Sosfenov, B. K. Vainshtein, Y. M. Torchinsky, A. E. Braunstein, *J. Mol. Biol.*, **125**, 275 (1978).
24. V. V. Borisov, S. N. Borisova, G. S. Kachalova, N. I. Sosfenov, B. K. Vainshtein, in: *Transaminases* (P. Christen, D. E. Metzler, eds.) John Wiley & Sons, New York (1985) p. 155.
25. E. G. Harutyunyan, V. N. Malashkevich, V. M. Kochkina, Y. M. Torchinsky, in: *Transaminases* (P. Christen, D. E. Metzler, eds.) John Wiley & Sons, New York (1985) p. 164.
26. C. A. McPhalen, M. G. Vincent, D. Picot, J. N. Jansonius, in: *Biochemistry of Vitamin B₆* (T. Korpela, P. Christen, eds.) Birkhaeuser Verlag, Basel (1987) p. 99.
27. C. A. McPhalen, M. G. Vincent, J. N. Jansonius, *J. Mol. Biol.*, **225**, 495 (1992).
28. C. A. McPhalen, M. G. Vincent, D. Picot, J. N. Jansonius, A. M. Lesk, C. Chothia, *J. Mol. Biol.*, **227**, 197 (1992).
29. D. Picot, E. Sandmeier, C. Thaller, M. G. Vincent, P. Christen, J. N. Jansonius, *Eur. J. Biochem.*, **196**, 329 (1991).
30. S. Kamitori, A. Okamoto, K. Hirotsu, T. Higuchi, S. Kuramitsu, H. Kagamiyama, Y. Matsuura, Y. Katsube, *J. Biochem.*, **108**, 175 (1990).
31. D. L. Smith, S. C. Almo, M. D. Toney, D. Ringe, *Biochemistry*, **28**, 8161 (1989).
32. F. Jager, E. Hohler, P. Tucker, U. Sauder, Z. Housley-Markovic, I. Fotheringham, M. Edwards, M. Hunter, K. Kirschner, J. N. Jansonius, *J. Mol. Biol.*, **209**, 499 (1989).
33. V. I. Ivanov, M. Y. Karpeisky, *Adv. Enzymol.*, **32**, 21 (1969).
34. N. Watanabe, K. Sakabe, N. Sakabe, T. Higashi, K. Sasaki, S. Aibara, Y. Morita, K. Yonaha, S. Toyama, H. Fukutani, *J. Biochem.*, **105**, 1 (1989).
35. W. Stark, J. Kallen, Z. Markovic-Housley, B. Fol. M. Kania, J. N. Jansonius, in: *Enzymes Dependent on Pyridoxal Phosphate and Other Carbonyl Compounds as Cofactors* (T. Fukui, H. Kagamiyama, K. Soda, H. Wada, eds.) Pergamon Press, Oxford, (1991) p. 656.
36. A. A. Antson, B. V. Strokopytov, G. N. Murshudov, M. N. Isupov, E. H. Harutyunyan, T. V. Demidkina, D. G. Vassilyev, Z. Dauter, H. Terry, K. S. Wilson, *FEBS Lett.*, **302**, 256 (1992).
37. N. B. Madsen, S. G. Withers, in: *Chemical, Biochemical, and Medical Aspects of Vitamin B₆, Pyridoxal Phosphate* (D. Dolphin, R. Poulson, O. Avramovic, eds.) John Wiley & Sons, New York, part B (1986) p. 355.
38. C. C. Hyde, S. A. Ahmed, E. A. Padlan, E. W. Miles, D. R. Davies, *J. Biol. Chem.*, **263**, 17857 (1988).
39. M. D. Toney, J. F. Kirsch, *Science*, **243**, 1485 (1989).
40. S. Futaki, H. Ueno, A. Martinez del Pozo, M. A. Pospischil, D. Ringe, B. Stoddard, K. Tanizawa, T. Yoshimura, K. Soda, *J. Biol. Chem.*, **265**, 22306 (1990).
41. K. Nishimura, K. Tanizawa, T. Yoshimura, N. Esaki, S. Futaki, J. M. Manning, K. Soda, *Biochemistry*, **30**, 4072 (1991).
42. S. Kochhar, W. L. Finlayson, J. F. Kirsch, P. Christen, *J. Biol. Chem.*, **262**, 11446 (1987).
43. T. Yoshimura, Y. Matsushima, K. Tanizawa, M. Sung, T. Yamauchi, M. Wakayama, N. Esaki, K. Soda, *J. Biochem.*, **108**, 699 (1990).
44. A. Planas, J. F. Kirsch, *Prot. Engineer.*, **3**, 625 (1990).
45. A. Planas, J. F. Kirsch, *Biochemistry*, **30**, 8268 (1991).
46. S. Kochhar, P. Christen, *Eur. J. Biochem.*, **203**, 563 (1992).
47. J. A. Landro, A. T. Kallarakal, S. C. Ransom, J. A. Gerlt, J. W. Kozarich, D. J. Neidhart, G. L. Kenyon, *Biochemistry*, **30**, 9274 (1991).
48. A. Martinez del Pozo, M. Merola, H. Ueno, J. M. Manning, K. Tanizawa, K. Nishimura, K. Soda, D. Ringe, *J. Biol. Chem.*, **264**, 17784 (1989).
49. M. D. Toney, J. F. Kirsch, *J. Biol. Chem.*, **262**, 12403 (1987).
50. M. D. Toney, J. F. Kirsch, *Biochemistry*, **30**, 7461 (1991).
51. M. D. Toney, J. F. Kirsch, *Biochemistry*, **30**, 7456 (1991).
52. Y. Inoue, S. Kuramitsu, K. Inoue, H. Kagamiyama, K. Hiromi, S. Tanase, Y. Morino, *J. Biol. Chem.*, **264**, 9673 (1989).
53. J. E. Taylor, D. E. Metzler, A. Arnone, *Ann. NY Acad. Sci.*, **585**, 58 (1990).
54. K. Inoue, S. Kuramitsu, A. Okamoto, K. Hirotsu, T. Higuchi, H. Kagamiyama, *Biochemistry*, **30**, 7796 (1991).
55. K. Inoue, K. Kuramitsu, A. Okamoto, K. Hirotsu, T. Higuchi, Y. Morino, H. Kagamiyama, *J. Biochem.*, **109**, 570 (1991).

56. J. M. Goldberg, R. V. Swanson, H. S. Goodman, J. F. Kirsch, *Biochemistry*, **30**, 305 (1991).
57. T. Yano, S. Kuramitsu, S. Tanase, Y. Morino, H. Kagamiyama, *Biochemistry*, **31**, 5878 (1992).
58. T. Yano, S. Kuramitsu, S. Tanase, Y. Morino, K. Hiromi, H. Kagamiyama, *J. Biol. Chem.*, **266**, 6079 (1991).
59. C. M. Metzler, D. E. Metzler, A. Kintanar, R. D. Scott, M. Marceau, *Biochem. Biophys. Res. Commun.*, **178**, 385 (1991).
60. H. Hayashi, Y. Inoue, S. Kuramitsu, H. Kagamiyama, in: *Enzymes Dependent on Pyridoxal Phosphate and Other Carbonyl Compounds as Cofactors* (T. Fukui, H. Kagamiyama, K. Soda, H. Wada, eds.), Pergamon Press, Oxford, (1991) p. 183.
61. C. N. Cronin, J. F. Kirsch, *Biochemistry*, **27**, 4572 (1988).
62. H. Hayashi, S. Kuramitsu, Y. Inoue, Y. Morino, H. Kagamiyama, *Biochem. Biophys. Res. Commun.*, **159**, 337 (1989).
63. E. Sandmeier, P. Christen, *J. Biol. Chem.*, **257**, 6745 (1982).
64. H. Hayashi, Y. Inoue, S. Kuramitsu, Y. Morino, H. Kagamiyama, *Biochem. Biophys. Res. Commun.*, **167**, 407 (1990).
65. L. M. Gloss, A. Planas, J. F. Kirsch, *Biochemistry*, **31**, 32 (1992).
66. S. Giannattasio, E. Marra, R. A. Vacca, G. Iannace, E. Quagliariello, *Arch. Biochem. Biophys.*, **298**, 532 (1992).
67. M. Merola, A. Martinez del Pozo, H. Ueno, P. Recsei, A. DiDonato, J. M. Manning, K. Tanizawa, Y. Masu, S. Asano, H. Tanaka, K. Soda, D. Ringe, G. Petsko, *Biochemistry*, **28**, 505 (1989).
68. S. Kuramitsu, *Vitamins (Japan)*, **64**, 633 (1990) (in Japanese).
69. M. Seville, M. G. Vincent, K. Hahn, *Biochemistry*, **27**, 8344 (1988).

6

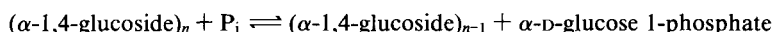
Higher Plant Phosphorylases

Katsuyuki Tanizawa, Hiroyuki Mori, Mitsuo Tagaya and Toshio Fukui

*The Institute of Scientific and Industrial Research, Osaka University,
Ibaraki, Osaka 567, Japan*

6.1 Introduction

α -Glucan phosphorylase (α -1,4-D-glucan:orthophosphate α -D-glucosyltransferase, EC 2.4.1.1) is a pyridoxal 5'-phosphate (pyridoxal-P)-dependent enzyme which catalyzes the phosphorolytic degradation of various α -1,4-glucan molecules at the nonreducing ends.



This enzyme occurs ubiquitously in nature and has been purified from a wide variety of organisms, including bacteria, fungi, yeasts, slime molds, plants, insects, and mammalian tissues such as skeletal muscle, liver, and brain.¹⁻³⁾ The phosphorylase reaction is reversible with a K_{eq} value of about 0.3 at pH 6.8, favoring the synthetic direction *in vitro*. However, the enzyme is believed to function *in vivo* mainly in the phosphorolysis direction due to the high ratio of P_i /glucose 1-phosphate inside the cells where glucose 1-phosphate is readily converted to glucose 6-phosphate by phosphoglucomutase and then metabolized through the glycolytic pathway. Hence, the physiological role of phosphorylase is to mobilize storage polysaccharides such as glycogen and starch, liberating the energy-conserving sugar phosphate. The enzyme existing as several isozymes in animal organs and tissues acts as the key enzyme in the glucose homeostasis.

One of the characteristics of mammalian phosphorylase is that the enzyme activity is controlled through complicated mechanisms. The activity of the rabbit skeletal muscle enzyme is regulated by both covalent and noncovalent modifications, *i.e.*, through the phosphorylation-dephosphorylation of the protein by phosphorylase kinase and phosphatase⁴⁻⁶⁾ and through the interaction with AMP.^{7,8)} Edwin G. Krebs and Edmond H. Fischer received the Nobel prize in 1992 by their pioneering works on this control mechanism of phosphorylase as the first example of cellular regulations by protein phosphorylation-dephosphorylation. In yeast, however, the enzyme activity is primarily controlled by the phosphorylation-dephosphorylation, and not by the interaction with AMP.⁹⁾ Maltodextrin phosphorylase from *Escherichia coli* is not phosphorylated but slightly activated by AMP.^{10,11)} Phosphorylases from higher plant tissues, for example potato tuber, are regulated neither by allosteric nor covalent modification and exist as a constantly active form.¹²⁻¹⁴⁾

Earlier studies on the properties of phosphorylases isolated from various sources have indicated that their subunits are similar in size with about 100,000 daltons.¹⁵⁻¹⁷⁾ The reaction proceeds in a rapid equilibrium random Bi-Bi mechanism as has been shown by kinetic studies with rabbit skeletal muscle phosphorylases *a*¹⁸⁻²⁰⁾ and *b*,^{21,22)} rabbit liver enzyme,²³⁾ potato tuber enzyme,²⁴⁾ and the enzyme from *E. coli*.²⁵⁾ In contrast, the substrate specificities for various glucans differ considerably depending on the enzyme sources. The rabbit muscle enzyme has high affinity for branched glucans such as glycogen and amylopectin but low affinity for amylose and maltodextrin.^{26,27)} The potato tuber enzyme can act on amylose, amylopectin, and maltodextrin but only poorly on glycogen,^{28,29)} while the *E. coli* enzyme shows high affinity for maltodextrin.¹⁰⁾

The enzyme from rabbit skeletal muscle is so far the most extensively studied phosphorylase. The primary structure of the rabbit muscle enzyme was determined by protein chemical techniques,¹⁶⁾ and deduced later from the nucleotide sequence of its cDNA³⁰⁾; an additional amino acid, isoleucine, had been introduced at position 308 in the amino acid sequence deduced from cDNA. Since the corrected numbering of the sequence of the rabbit muscle enzyme gives a better homology with amino acid sequences of phosphorylases from other sources,³⁰⁾ this numbering is employed hereafter when referring to the rabbit muscle enzyme. The X-ray crystallographic studies of both phosphorylases *a* and *b* have been completed,^{31,32)} and the structural bases for the catalytic mechanism and the control of activity by covalent and noncovalent modifications have already been reviewed extensively,³³⁻³⁵⁾ and are beyond the scope of this review. In this chapter, we will focus on the structure, function, and evolutionary aspect of phosphorylases from higher plants, in particular the two distinct phosphorylase isozymes from potato tuber that have been studied in this laboratory for more than twenty years.

6.2 Phosphorylase Isozymes from Higher Plants

In higher plants, many enzymes involved in the carbohydrate metabolism are distributed in both cytoplasmic and plastidic compartment.³⁶⁻³⁹⁾ These enzymes, mostly existing as isoforms or isozymes, include not only the enzymes in the glycolytic pathway^{40,41)} but also those acting on α -1,4/1,6-linked glucans, such as α -amylase,^{42,43)} β -amylase,⁴⁴⁾ isoamylase (debranching enzyme),⁴⁵⁾ and D-enzyme (disproportionating enzyme).^{42,45,46)} As for phosphorylase, at least two isozymes showing different intracellular localizations have also been reported, one occurring in the cytoplasm and the other in the plastid.⁴⁶⁻⁵¹⁾ Potato tuber contains two forms of phosphorylase, although the content of one form by far exceeds that of the other.⁵²⁾ The two types of phosphorylases are separable based on the difference in their affinity for glucans. The major form uses, as substrate, maltodextrin and amylose, as well as amylopectin, but not glycogen.⁵³⁾ On the other hand, the minor form has extraordinarily high affinity for various glucans, including glycogen,⁵⁴⁾ and is slightly different from the major form in molecular size. Immunological studies revealed the presence of phosphorylase (major form) in the amyloplast of potato tuber during starch accumulation.⁵⁵⁾ Recently, Brisson *et al.*⁵⁶⁾ have reported that the enzyme (presumably the major form) is localized inside the stroma of amyloplast in young potato tuber, whereas in mature tuber, it is found within the cytoplasm in the immediate vicinity of the plastid.

Phosphorylase activity in spinach leaf is also distributed in the chloroplastic and cytoplasmic fractions, but in comparable amounts.^{46,47)} The cytoplasmic enzyme has much higher affinity for

TABLE 6.1 Phosphorylase Isozymes in Higher Plants

	Potato Tuber		Potato Leaf	
	Type L	Type H	Type L	Type H
Subunit Molecular Weight	103,916	95,382	104,000	94,000
Km for Glycogen (mM)	2.4	0.001	1.4	0.004
Relative Contents	Major	Minor	Comparable	
Distribution	Amyloplast	?	Chloroplast	Cytoplasm

glucans than the chloroplastic enzyme.⁵⁰⁾ A similar pair of phosphorylases have been found in pea leaf,⁵⁷⁾ as well as in etiolated soybean seedling.⁵⁸⁾ Fukui *et al.*⁵⁹⁾ isolated the two phosphorylase isozymes from potato leaves and compared their enzymological properties with those of the two isozymes from the tuber of the same plant. From these comparisons, it was concluded that plant phosphorylases can be classified into types L and H, after their low and high affinity for glucans, respectively (Table 6.1). The type-L isozyme is localized within the plastids (chloroplast or amyloplast) and the type-H isozyme is localized in the cytoplasm, although the intracellular localization of the tuber minor isozyme (type H) has not been examined. The affinity for glycogen of the type-H isozyme is approximately three orders of magnitude higher than that of the type-L isozyme. The polypeptides of the two isozymes differ in size by about 10,000 daltons and in immunological properties, and are encoded by distinct chromosomal genes as described below.

In animal cells, glycogen is deposited in cytoplasm in granular forms. The α -1,4-glycosidic linkages in the glucan are synthesized and degraded by glycogen synthase and phosphorylase, respectively. In higher plant tissues, however, starch is reserved within plastids, either chloroplasts, or amyloplasts. Although it is believed that starch synthase and phosphorylase are respectively responsible for the synthesis and degradation of starch, as in animal cells,⁶⁰⁾ the physiological significance of the presence of the two plant phosphorylase isozymes and their roles in carbohydrate metabolism remain unclear.

6.3 Structures of Phosphorylase Isozymes from Potato Tuber

6.3.1 Type-L Isozyme

The complete amino acid sequence of the potato tuber type-L isozyme was determined in this laboratory by using only protein-chemical techniques.⁶¹⁻⁶⁵⁾ Given the marked progress in recombinant DNA technology over the past decade, this may well be the last example of protein-chemical sequence determination for such a large protein.

The polypeptide of type-L isozyme is composed of 916 amino acid residues. There are two covalently modified amino acids; about one-fourth of the amino-terminal threonines are blocked by an acetyl group, and a lysyl residue (Lys⁷⁶²) is linked to the cofactor pyridoxal-P. The partial amino-terminal acetylation appeared to be a natural feature of type-L isozyme, and both the blocked and unblocked forms of type-L phosphorylase may exist in potato tubers.

The subunit has a calculated molecular weight of 103,916 including the acetyl group and the cofactor. The sequence is two residues longer at the amino-terminus and, similar to the bacterial enzyme, nine residues shorter at the carboxyl-terminus than the muscle enzyme when the two sequences are aligned for maximum matching (see below). Furthermore, there is a 78-residue insertion in the middle of the polypeptide chain of the type-L isozyme that occurs between residues

414 and 415 in the rabbit muscle sequence and accounts for most of the molecular weight difference between the two enzymes. The large insertion is unique in the type-L isozyme of higher plant phosphorylases including that from sweet potato,⁶⁶⁾ even compared with the α -glucan phosphorylase sequences from animal and microorganism.²⁾ The structural features and a putative function of this 78-residue insertion are discussed in a later section. In agreement with the lack of regulation in the potato enzyme,⁶⁷⁾ the amino acid sequence of the amino-terminal region (about 80 amino acid residues) of the type-L enzyme is considerably different from that of the muscle enzyme.

Subsequently, Nakano, Mori and Fukui⁶⁸⁾ isolated a nearly full-length cDNA encoding the type-L isozyme from a λ gt10 phage library prepared from immature potato tuber. The clone contained a DNA insert of about 3.1 kilobase pairs, and its nucleotide sequence was determined and supplemented by a partial genomic clone. The transcription initiation site was identified by a primer extension experiment to be 43 bases upstream from the putative translation initiation ATG codon. A TATA-box-like sequence, AATAAAATAA, is located 30 nucleotides upstream from the transcription initiation site. The cDNA sequence around the initiation ATG codon is consistent with the plant consensus sequence proposed by Heidecker and Messing.⁶⁹⁾ Since the message also had a poly(A) sequence, the type-L isozyme is thought to be encoded in the nuclear genome, like many other chloroplastic proteins. A putative polyadenylation signal, AATAA, was found at position 3053, preceding the ends of the mature message by 30 nucleotides. There is a potential site that makes a hairpin structure around the polyadenylation signal from position 3053 to 3077. All these features indicate that the type-L phosphorylase isozyme cDNA isolated is a typical plant message.⁶⁹⁾

The coding region of the type-L isozyme is 2898 base pairs long. It corresponds to 966 amino acid residues with a molecular weight of 109,648, and consists of two regions: the amino-terminal extended sequence of 50 residues and the mature protein sequence of 916 residues. The amino acid sequence of the mature protein deduced from the cDNA structure is perfectly matched to the sequence chemically determined as described above. The stop codon, TAA, appeared at position 2942 in the cDNA sequence just after the carboxyl-terminal amino acid of the mature enzyme, indicating that no post-translational processing occurred at the carboxyl terminus.

The type-L isozymes purified from amyloplast of potato tuber and chloroplast of potato leaf are immunologically indistinguishable.⁵⁹⁾ Moreover, it has been assumed that the amyloplast and chloroplast themselves developed from common proplastids. In fact, interconversion between the two organelle has been observed in several plants including *Picea abies*.⁷⁰⁾ Also, there have been several reports showing that the amyloplast and chloroplast genomes from the same plant are indistinguishable.⁷¹⁻⁷³⁾ It is therefore reasonable to assume that the amino-terminal 50 amino acid extension found in the type-L isozyme cDNA functions as a transit peptide for the amyloplast, like those of many nuclear-encoded chloroplastic proteins.⁷⁴⁾ The transit peptides of chloroplastic proteins share some characteristics: rich in serine and threonine and positive in the net charge. Phenylalanine seems to be most prevalent among aromatic amino acids. Although there seems to be no consensus sequence among these peptides, the sequences for the same protein from different plant species show significant similarities. Among the transit peptides of glyceraldehyde 3-phosphate dehydrogenases from maize and two species of tobacco, for example, high contents of acidic residues and phenylalanine are prominent, and histidine, tyrosine, and tryptophan are of low abundance.⁷⁵⁻⁷⁷⁾ Furthermore, substantial regions with amino acid homology and conservation of charged residues have been noted in the transit peptides of the small subunit

of ribulose biphosphate carboxylase.⁷⁴⁾

The transit peptide of the type-L phosphorylase isozyme of potato amyloplast has the following amino acid composition and sequence:

R₅H₅K₃D₀E₀G₁S₇T₅C₁Y₁N₆Q₀F₆W₀A₃L₃I₁V₁M₁P₁
MAT ANG AHL FNH YSS NSR FIH FTS RNT SSK LFL TKT SHF RRP KRC FHV NN

There is no acidic amino acid in this transit peptide sequence. By contrast, 16% of the total 50 residues are arginines plus lysines, and, if histidines are included, the content of basic residues becomes 26%. The basic nature of the transit peptide is in contrast to the acidic nature of the mature protein; the content of aspartic and glutamic acids in the mature protein is much higher than the content of lysine, arginine, and histidine. It seems likely that the acidity of the mature protein molecule is neutralized with the excessive basicity of the transit peptide.

A similar relationship between a transit peptide and the corresponding mature protein can be seen in glutamine synthetase from pea.⁷⁵⁾ In glyceraldehyde 3-phosphate dehydrogenase from maize, whose transit peptide has a rather high content of acidic residues, the charges in the mature protein are nearly balanced.⁷⁶⁾ Smeekens *et al.*⁷⁷⁾ analyzed the functional significance of the transit peptide by constructing a chimeric protein with a positively charged sequence just after the transit sequence. The protein was more tightly bound to the chloroplast, although the protein was not imported into the organelle. They concluded that the mature protein structure, in addition to the transit peptide, may contribute to its import into the chloroplast, and that transit peptides may have evolved cooperatively with mature proteins for efficient transportation into the organelle. The features of the transit peptide of potato amyloplast type-L phosphorylase isozyme may support this hypothesis.

Another interesting feature of the transit peptide of the type-L phosphorylase isozyme is the high content of histidine; the peptide contains five histidyl residues in a total of 50 amino acid residues, in contrast to other transit peptides that contain no or little, if any, histidyl residues. It is not known if these residues have any special role other than increasing the positive charge of the peptide.

6.3.2 Type-H Isozyme

Although the content of the type-H isozyme in potato tuber is very low, its purification is not so difficult if one utilizes the glycogen-Sepharose affinity column.⁷⁸⁾ However, determination of its complete sequence by protein-chemical techniques is not practical because of the very large molecular size. Thus, using specific oligonucleotide probes synthesized on the basis of partial amino acid sequences, a cDNA encoding the type-H phosphorylase isozyme has been isolated from a cDNA library of immature potato tuber.⁷⁹⁾ The isolated cDNA clone contained an inserted DNA of 2665 base pairs, including an open reading frame and a poly(A) tail. The amino acid sequence deduced from the DNA sequence agreed with partial sequences of several tryptic peptides and the amino-terminal peptide determined chemically; the open reading frame found corresponded to the entire polypeptide of the type-H phosphorylase isozyme. The length of mRNA for the type-H isozyme of potato tuber is slightly shorter than that of mRNA for the type-L isozyme having about 3.1 kilobase pairs as described above. 3'-Downstream of the open reading frame,

there are two AATAA sequences at positions 2541 and 2577, which are characteristic of the polyadenylation signal in mRNAs of plants and animals. These features support the conclusion that the type-H isozyme is also encoded by a nuclear gene distinct from that for the type-L isozyme,⁶⁸⁾ and eliminate the possibility of the age-related differential compartmentation of the single enzyme species in tuber cells.⁵⁶⁾

The DNA sequence around the initiation ATG codon of the type-H isozyme cDNA is rather dissimilar to that of the type-L isozyme cDNA and the consensus sequence for plant genes proposed by Heidecker and Messing.⁶⁹⁾ The difference in the sequences of this region may be correlated with the difference in the expression levels of each isozyme in potato tuber.

The coding region of the type-H isozyme consists of 2514 base pairs and corresponds to 838 amino acid residues with a calculated molecular weight of 95,382, including the cofactor pyridoxal-P and the amino-terminal-blocking acetyl group. Since the molecular weight calculated from the amino acid sequence of the type-L isozyme subunit is 103,916,⁶⁵⁾ the difference in the subunit molecular weight of the two isozymes agrees with the difference observed on sodium lauryl sulfate-polyacrylamide gel electrophoresis.⁵⁹⁾

The amino-terminal Met of the type-H phosphorylase was found to be completely blocked by an acetyl group, which was identified by fast atom bombardment-mass spectroscopy of a blocked peptide in the amino-terminal region.⁷⁹⁾ Although the molecular mechanisms for amino-terminal processing and acetylation at the α -amino group are still poorly understood, a rule, in which the cleavage of the initiator methionine is prevented when the adjacent residue is a charged amino acid, has been proposed.⁸⁰⁾ Since the adjacent residue in the type-H isozyme is Glu, the rule applies to this case. However, the second residue in the mature enzyme of the type-L phosphorylase is Leu, although its amino-terminal Thr is partially acetylated.⁶⁴⁾ Thus, the "N-end" rule is inapplicable to the type-L isozyme due to the presence of a transit peptide that must be cleaved off after translation and during transportation into plastids.

To study the tissue- and age-specific expression of the two phosphorylase isozymes, Mori, Tanizawa and Fukui⁷⁹⁾ performed Northern blot analysis on poly(A) RNAs isolated from potato tubers of different ages using cDNA probes specific for each enzyme mRNA. The cDNA probe for the type-L isozyme hybridized to an mRNA of about 3100-nucleotide length, which was present in tubers of all the developmental stages tested. The type-H isozyme cDNA probe also hybridized to an mRNA, which appeared to be slightly shorter than that of the type-L message and was also present in all stages. No significant difference in the length or the amount of two phosphorylase isozyme mRNAs was observed among tubers of three different ages. However, the semi-quantitative analysis by dot blotting of the potato tuber cDNA library with the cDNA fragments of both isozymes indicated that the content of mRNA encoding the type-L isozyme was about 15 times that of the type-H isozyme. The mRNA ratio may correspond to the ratio of the amounts of each isozyme protein. Therefore, it is suggested that the contents of the proteins are mainly regulated at the transcriptional level. This is consistent with the previous observations that the type-L isozyme is two orders of magnitude more abundant than the type-H isozyme in potato tuber, although the two isozymes are present at a comparable level in the leaf.⁵⁹⁾

Sequence analysis of the 5'-upstream regions of the genomic DNAs of the type-H and -L isozymes (unpublished results) indicated that in the type-L isozyme gene there is a significantly homologous region with the 5'-upstream region of class II patatin,⁸¹⁾ which is

one of the storage proteins in potato tubers, but not in the type-H isozyme gene. This finding suggests that the type-L isozyme gene is expressed efficiently in tubers, whereas the type-H isozyme gene is expressed at constitutive levels in both leaves and tubers. In fact, transgenic potatoes constructed recently, carrying the 5'-upstream region of the type-L isozyme gene (about 1 kilobase pair) fused to the GUS (β -glucuronidase) reporter gene, expressed very high GUS activity in tubers (unpublished results). Identification of the promoter regions responsible for the differential expressions of the two isozymes is currently in progress in this laboratory to elucidate the regulatory mechanisms of expression of the higher plant phosphorylase genes.

6.3.3 Sequence Comparison and Stereostructure

The amino acid sequences of the type-H and -L phosphorylase isozymes are aligned by introducing gaps (hyphens) for maximum matching (Fig. 6.1). Two marked differences between the sequences are apparent: the presence in the type-L isozyme of the putative transit peptide of 50 residues and a lengthy insertion of 78 residues in the middle of its polypeptide chain, which is by no means similar to any part of the type-H isozyme sequence. The 78-residue insertion accounts for most of the difference in subunit molecular weight between the two phosphorylase isozymes in mature forms. The absence in the type-H isozyme sequence of a transit peptide region is compatible with its presumed cytoplasmic localization.⁷⁹⁾ When these two portions are excluded in the comparison, the sequences of the two potato isozymes show an amino acid identity as high as 63% after the introduction of only five single gaps. However, this homology score is even lower than those among the human phosphorylase isozymes from various tissues: the muscle enzyme *versus* the brain enzyme, 83%; the muscle enzyme *versus* the liver enzyme, 80%; and the brain enzyme *versus* the liver enzyme, 80%.²⁾

Figure 6.1 also includes the sequence comparison with the rabbit muscle phosphorylase.^{16,30)} The sequence homologies of the animal enzyme with the potato type-L and type-H phosphorylase isozymes are 38% and 47%, respectively. Further sequence comparison of the plant and animal enzymes with the enzyme from *E. coli*⁸²⁾ reveals an overall homology of at least 40%. These high similarities indicate that the phosphorylase family is one of the well-conserved protein groups.

Nakano and Fukui⁶⁵⁾ previously proposed the idea that the potato type-L isozyme and rabbit muscle phosphorylase have similar polypeptide chain folds on the basis of the considerable sequence similarity between them. Since the type-H isozyme sequence is more homologous than the type-L isozyme sequence with the rabbit muscle enzyme, it is appropriate to assume that the type-H isozyme has polypeptide chain folds also similar to the rabbit muscle enzyme. All the deletion and insertion of residues introduced for maximum sequence matching are located at the solvent accessible surface of the protein, and therefore the two potato isozymes can be accommodated without significant alterations in the chain folding, in the tertiary structure of the rabbit muscle phosphorylase reported.^{83,84)} The sequences in the region of the β -strands constituting the central core are highly conserved in all the phosphorylase sequences reported so far.²⁾ By contrast, the sequences in the regions of α -helices and coils that surround the central core are less conserved. In particular, two long sequences from both the type-L and -H isozymes in these regions show considerably low similarities of less than 20% with the rabbit enzyme sequence, although one of them (residue 717–800) is highly

```

H      1 MEGGAKSNDVSAAPIAQ
L      MATANGAHLFNHYSSNSRFIHFTSRNTSSKLFLTKTSHFRPRKRCFHVNNLTSEKIHHPITEQGGSDL$
R      MSRFLSDQEKRRKQISVRGL

H      20 40 60 80
L PLSEDPDIA$NIKYHAQYTPH$P$FKFEPLQAYYAATADSVDRDLIKQWNDTYLHYDKVNP$KQTYIYLSM
R SFAP AAS T S EF V ER LPK FF- Q S LIN A DI E L M A
R AGVENVTELKK FNR LHF -LVKDRNVATPRD F L HT H VGR IR QQ YEKD RI L

H      100 120 140
L FYLQGRALTN$VGNLDIHNAYADALNKL$GQOLEEVVEQEKDAALGN$GGLRLASCFLDSMATLNL$PANGY
R FYM T Q TMV ALE CDE TYQ LDM LE I E G A G A Y

H      160 180 200 220
L GLRIRYGLFKQLITKAGQEEVPE$DWLEK$FSPWEIVRHDVVFPIRFFGHVEVLPSGSRK$WVGGEVLQALAY
R I K R D A IG V N SY K Y K STGSD K Y I DIK V
R I EF I N K CGGW M EAD RYGN KA PEFTL VH Y R HTSQ A- DTQ VL MP

H      240 260 280
L DVPIPGYRTKNTN$LRLEWAK$SED$FNLF$FNDGQYDAA$QLHSRAQ$ICAVLYPGDATENGKLLRLKQ
R T V NNVV TM S -PN KD V G IQ VLDRNL EN SR N NFFE E

H      300 320 340 360
L QFFLCSASLQDIIARFKEREDGKGS---HOWSEFPKKVAIQ$NDHTPTLTIPELMRLLMDDEGLGWDESW
R YT S - RS DR----IK E F E V M C I I LK N N A
R EY VVA T R SSKF CRDPVRTN$FDA D S A V V L R D KA

H      380 400
L NITTRTIAYTNHTVLPEALEK$W$QAVMWKLLPRHMEIIEIDKRFVATI-----
R Q V YEL Q V A EEL HE VLKYG$MDLNKLEEKLT$TMR I
R EV VK C R PVHLEET LQ Y NQ LNRV-----

H      420
L LENFDLPSSVAELFIKPEISVDDDTETVEVHDKVEASDKVVTNDEDDTGKKT$SVKIEAAAEK IDK T V
R -----AAAF GDVDRLLR

H      440 460 480
L MRILDHNATKPVVHMANLCV$VSSH$TVNGVAQLHSDILKAELFADYVSVWPTK$FQNK$TNGITPRRWIRFC$
R SPEPAVIPP K- R GG A EI E V E V N FYEL E V
R SLVEEG V R-IN H IAG A RI E KTI K FYELE H LVL N

H      500 520 540 560
L PELSHIITKWLKTDQWVTNLELLANLREFADNSELHAEWESAKMANKORLAQYILHVTGVSIDPNSLFDI
R P A TG ED LKT K E QK ED QN REA RS IKVVSFLKEK Y VV DAM
R G AE AERI-GEYISD DQ RK LSYV DEAFIRDVAKV QE LKF A LEREYK H N V

H      580 600 620
L QVKRIHEYK$RQLLNILGVIYRYK$KLKGMSPEERKNT-TPRTVMIGGKAFATYTN$AKRIVKL$VTDVGDV$VN
R F IV M E TAA TNFV VCIF VQ FI ATI
R C H TL NRIK---K PN FV-V APG HM M I I AI

H      640 660 680 700
L SDPDVNDYLKVVFPV$NPNY$VSAEMLIPG$SELSQHISTAGMEASGTSNMKFALNGCLII$GTL$DGANVEIRE
R H EIG L D L A D E M IQ
R H V G R R I LE R L KV AAD EQ T G M A T M MA

H      720 740 760
L EIGEDNFFLFGATADEVPQLRK$DRENGLFKPDPRFEEAKQFIR-SGAFGTYDYNPLLESLEGN$SGYGRGD
R V E Q H IAG E AD K V E V E V - S N DD IG E F A
R A E I MRVED DR DQRGY AQEYI RIP LRQIIEQL SG FSPKQPD FKDIVNMLMH--

H      780 800 820 838
L YFLVGHD$F$PSYMDAQARVDEAYKDRKR$WIKMSILSTSGSGKFSSDRTISQYAKEIWNIAECRVP
R K IEC EK R Q TT N A SY HE D EAVEIA
R R K FA YEE VKC ER SAL NP$RE TR V RNIA$T AQ R GVEPSRQRLPAPDE

R      KIP
      840 842

```

Fig. 6.1 Amino acid sequence comparison of α -glucan phosphorylases. H, the potato type-H isozyme; L, the potato type-L isozyme; and R, rabbit muscle phosphorylase.⁷⁹⁾ The type-H isozyme sequence is used as the reference sequence; only the amino acid residues that are nonidentical in the type-L isozyme and rabbit muscle enzyme are indicated. (From *J. Biol. Chem.*, 266 (28), 18453 (1991)).

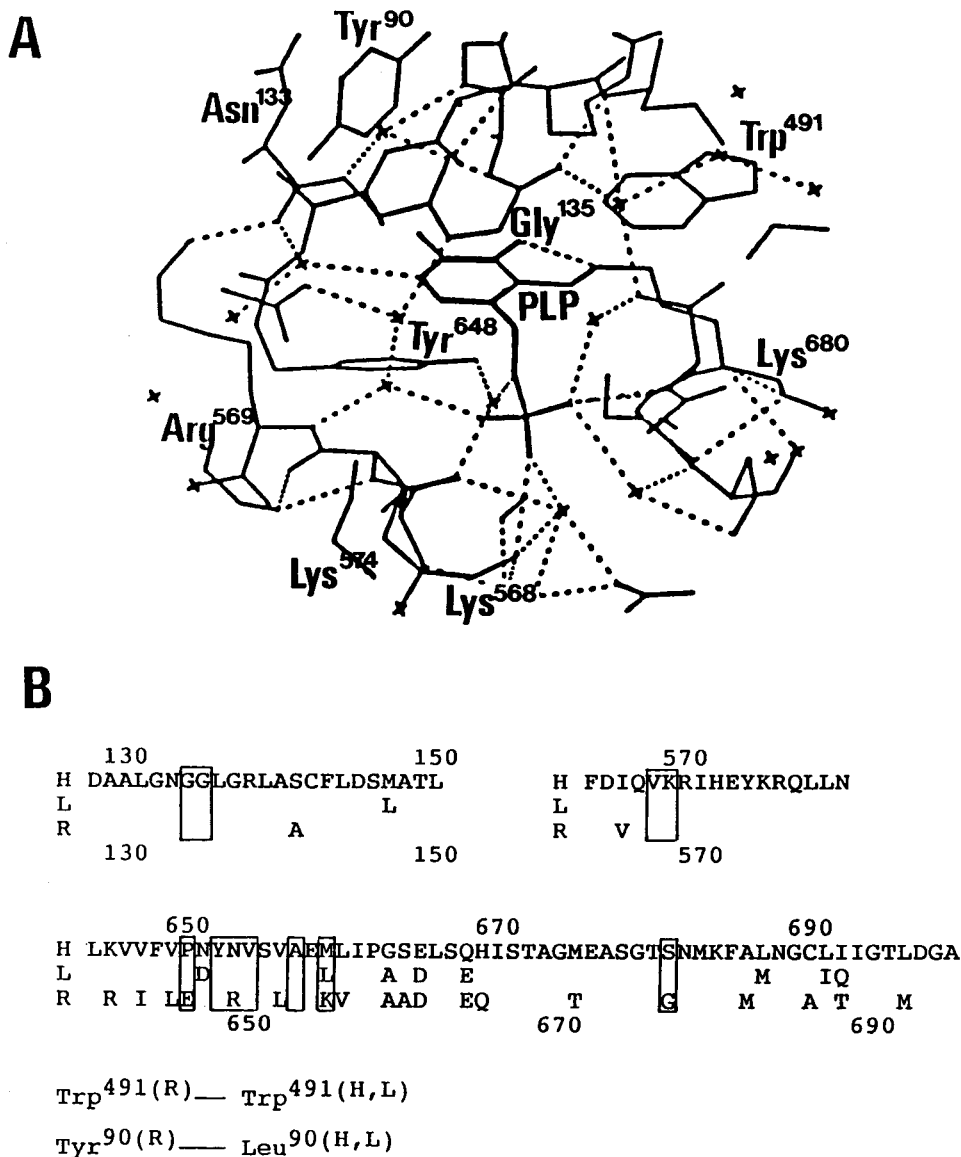


Fig. 6.2 Spatial structure of the active site of rabbit muscle phosphorylase⁸⁵ (A) and sequence comparison of the active site residues (B). A, Hydrogen bonds of less than 3.3 Å (dotted lines) and water molecules (crosses) are also indicated. B, Sequences constituting the active-site region are compared among potato type-H isozyme (H), type-L isozyme (L), and rabbit muscle enzyme (R). Residues making van der Waals contact with the pyridoxal moiety and the 5'-phosphate group are boxed. (From *J. Biol. Chem.*, 261 (18), 8233 (1986)).

conserved (63%) between the two potato isozymes.

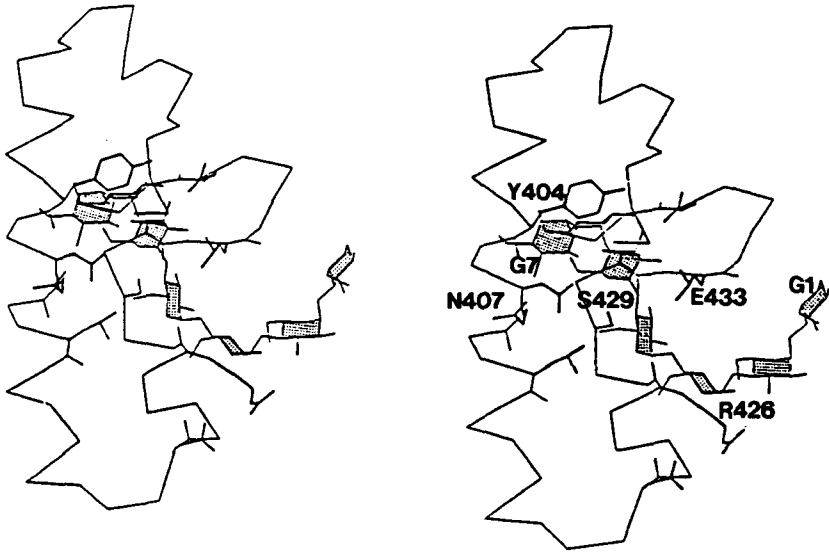
Figure 6.2(A) shows the spatial structure of the active-site region in the rabbit muscle phosphorylase,⁸⁵⁾ in which amino acid residues interacting with the cofactor pyridoxal-P have been indicated. Comparison of partial sequences constituting this active-site region is shown in Fig. 6.2(B). A marked homology of the sequence around the cofactor-binding lysyl residue has already been noted among phosphorylases from rabbit muscle,¹⁶⁾ yeast,⁸⁶⁾ potato tuber (the type-L isozyme⁶¹⁾), *E. coli*,⁸⁷⁾ and human tissues.²⁾ The sequence of residue 644–716 in the potato type-H isozyme is also very similar to this region. Therefore, it seems likely that Lys⁶⁸⁴ of the type-H isozyme is the pyridoxal-P-binding residue.

Affinity labeling studies with P¹,P²-bis-(5'-pyridoxal)diphosphate revealed the presence of a reactive amino group near the phosphate group of the cofactor, and the labeled residue was identified as Lys⁵⁷⁴ in the rabbit muscle enzyme⁸⁸⁾ and Lys⁶⁵² in the potato type-L isozyme (which is matched to Lys⁵⁷⁴ of the rabbit enzyme).⁸⁹⁾ This lysyl residue is conserved in the potato type-H isozyme, as well. In addition, Arg⁵⁶⁹ of the rabbit muscle enzyme has been shown to be essential for the activity,⁹⁰⁾ which is also found at the corresponding positions in both potato isozyme sequences. The high conservation of residues in the sequence including these Lys⁵⁷⁴ and Arg⁵⁶⁹ among phosphorylases from various sources has been previously noted.⁶⁵⁾ The 17-residue sequence (residue 564–580) of the type-H isozyme is identical with that of the type-L isozyme sequence and very similar to those from other phosphorylase sequences. These conservations support the view that the sequence of residue 564–580 constitutes a part of the phosphorylase active site together with the region of residue 644–716 containing the pyridoxal-P-binding Lys⁶⁸⁴.

Most of the residues in contact with pyridoxal-P are also generally conserved in the three enzymes compared (Fig. 6.2(B)). In particular, the residues close to the 5'-phosphate group of pyridoxal-P are completely conserved except for Gly⁶⁷⁷ of the rabbit enzyme. However, several residues of the rabbit enzyme, which are in van der Waals contact with the pyridine ring of pyridoxal-P, are substituted in the plant isozymes: Tyr⁹⁰ to Leu, Glu⁶⁴⁶ to Pro, Arg⁶⁴⁹ to Asn, and Lys⁶⁵⁵ to Met or Leu. The higher conservation of residues around the 5'-phosphate group than those around the pyridine ring is compatible with the catalytic importance of the 5'-phosphate group in phosphorylase (see below).

The glycogen-storage site in the X-ray structure of the rabbit muscle enzyme is a subdomain consisting of 58 residues (position 383–440) and has an unusual $\beta\alpha_3\beta_2$ structure located on the globular surface of the molecule at 30 Å apart from the active site.^{91,92)} The glycogen-storage site shows much higher affinity for glucan than the active site itself, and serves as an anchoring point through which the enzyme dimer binds to a glycogen particle. A distant terminus of the highly branched glucan enters the active-site crevice. This idea could explain why branched glucans are better substrates for the rabbit muscle enzyme than linear glucans.⁸³⁾ The X-ray diffraction analysis of crystals soaked with maltoheptaose, a low molecular weight analogue of an A-chain of glycogen, revealed eight residues that could interact directly with glycogen⁹²⁾ (Fig. 6.3). These residues are either exactly conserved or are of conservative replacement in the sequences of yeast and mammalian enzymes, both of which show high affinity for glycogen, whereas only three of the eight residues and only 42% of the residues that comprise the entire glycogen-storage site are conserved in the sequence of the bacterial enzyme,⁸⁷⁾ which prefers unbranched maltodextrins as substrates. Although the potato type-H isozyme shows very high affinity for both glycogen and amylopectin,⁵²⁾ none of the eight residues and only 32% of the residues in the region

A



B

	390	400	410	420	430	440
H	LEKWSQAVMWKLLPRHMEIIE	EDIKRFVATI	-----	MSERPDLNKMPSM	RIIDENATKPVVH	
L	YEL Q	V	A	EEL HE (insertion)	AAAEK IDK T	VSPEPAMIPDKK- R
E	R DVKLVKG	Q	N	NT KTLV	-----	EKTW GD KVVAKLAVV- HDKQ
Y	PRRLFGH	L	YD NWF LQDV	-----	AKKF KDVDLLSRIS	IEE SPERQIR
HL	R PVDLVE	L	Y	NOKHLDR	-----	VALF KDVDPLRR SLIEEGS R-IN
R	R PVHLLET	LQ	Y NC	LNRV	-----	AAAF GDVDRIIR SLIEEG V R-IN
	390	400	410	420	430	440

Fig. 6.3 Stereoview of the glycogen-storage site of rabbit muscle phosphorylase⁹²⁾ (A) and sequence comparison of the region (B). A, Sugar residues in maltoheptaose bound to the glycogen-storage site are stippled. Side chains in contact with the saccharide are shown in thick lines. B, The sequence constituting the glycogen-storage site of the rabbit enzyme (R) is compared with those of the potato type-H isozyme (H), type-L isozyme (L), *E. coli* maltodextrin phosphorylase (E), yeast enzyme (Y), and human liver enzyme (HL). Eight residues being matched with those of the rabbit enzyme that interact with the saccharide are boxed. (Reproduced with permission from Oikonomakos *et al. Biochemistry*, **26** (25), 8384 (1987)).

corresponding to the glycogen-storage site are conserved (Fig. 6.4(B)). Therefore, the presence in the type-H isozyme of a high affinity site like the glycogen-storage site of the rabbit muscle enzyme is uncertain from the sequence comparison but may exist in the corresponding region on the basis of our recent study⁹³⁾ with a chimeric phosphorylase composed of the type-L and -H isozymes as described below.

The 78-residue region uniquely found in the middle of the potato type-L isozyme sequence is rich in charged amino acid residues and appears to form a helical structure.⁶⁵⁾ The region has been predicted to be located at a site close to the glycogen-storage site of the rabbit muscle enzyme⁹¹⁾. The potato type-L isozyme shows relatively low affinities for glucans. Kinetic analyses and binding studies have suggested that glucans bind directly to the active site of the type-L isozyme.^{50,94)} Thus, it has been concluded that the potato type-L isozyme possesses no glucan-storage site. Furthermore, the 78-residue insertion may block the entry of highly branched glucans into the active site, while allowing access to the smaller, linear oligo- and polysaccharides. In supporting this, removal of the 78-residue insertion from the type-L isozyme affected dramatically its substrate discrimination⁹³⁾ (see below). Since the sequence of the type-H isozyme lacks the 78-residue insertion, it may be reasonable to assume again that the type-H isozyme shows high affinity for both glycogen and amylopectin.

Previously, it has been shown that most of the residues directly interacting with AMP as well as the phosphorylatable Ser¹⁴ and its surroundings in the rabbit muscle enzyme are far less conserved in the potato type-L isozyme sequence.⁶³⁾ Likewise, the amino-terminal region of the potato type-H isozyme is completely different from that of the rabbit muscle enzyme over the first 80 amino acid residues, in which the sites of covalent phosphorylation and of allosteric regulation by AMP are all included. These variances in sequence are compatible with the lack of regulation in the plant phosphorylase isozymes.

Cyclodextrins have been known to inhibit strongly the potato type-L phosphorylase, competing with the substrate amylopectin.⁹⁵⁾ This inhibition has also been shown with phosphorylases from other plants such as spinach leaf (both types) and with the potato type-H isozyme, as well,⁵⁰⁾ but the compounds do not inhibit the rabbit muscle phosphorylase.²⁸⁾ Since cyclodextrins are fairly large molecules made of 6 to 8 glucosyl residues, it is presumed that they cannot enter the narrow active-site crevice. In fact, it was shown that cyclodextrins interact with residues in the area near the crevice at the surface of the molecule.⁹⁴⁾ The superimposition of the potato type-L isozyme on the X-ray structure of the rabbit muscle enzyme predicted that in the potato enzyme cyclodextrins come into contact with the region of residues 713–773, which shows less than 20% similarity to the corresponding region of the rabbit muscle enzyme.⁶⁵⁾ The predicted cyclodextrin-binding site corresponds to the region of residues 718–779 of the type-H isozyme, which shows a high similarity (64%) to the region of the type-L isozyme, suggesting that cyclodextrins bind in the same region of both isozymes.

Willnecker, Jahnke and Buehner⁹⁶⁾ reported recently the preliminary results of X-ray crystallographic analysis of the potato type-L isozyme at 2.6 Å resolution, and showed that the overall folding of the potato phosphorylase is essentially the same as that of the rabbit muscle enzyme (more similar to the *R*-form than to the *T*-form). Only small differences were found in the core of the subunit, although local structures on the surface including the loops deviated significantly, being consistent with the sequence differences with some minor insertions and deletions as compared with the rabbit muscle enzyme (see above). As

predicted from the sequence, most of the 78-residue insertion was invisible probably due to the disordered structure.

6.3.4 Functional Role of the 78-Residue Insertion in Type-L Isozyme

Recently, Mori, Tanizawa and Fukui^{79,93)} constructed bacterial expression systems for cDNAs encoding potato type-H and -L phosphorylase isozymes using the pET vectors carrying the strong T7 RNA polymerase promoter.⁹⁷⁾ Interestingly enough, induction by isopropyl- β -D-thiogalactoside of the active enzymes was entirely dependent on the temperature in both cases; the active enzymes were undetectable in a cell extract of the bacterium induced at the normal growing temperature, 37°C, but were produced considerably in the cells induced at a low temperature, being optimum at 22°C. The reason why the bacterial expression of enzymically active plant phosphorylase is possible only at a low temperature is unclear, but may be ascribable to inefficient folding at a higher temperature into active forms of very large polypeptides (M_r 95,000–104,000). A similar observation was previously reported in the temperature-sensitive bacterial expression of rabbit muscle phosphorylase cDNA.⁹⁸⁾

To examine whether the 78-residue insertion uniquely occurring in the type-L isozyme plays a role in enzymic functions, Mori, Tanizawa and Fukui⁹³⁾ also constructed a chimeric phosphorylase in which the 78-residue insertion and its flanking regions are replaced by the corresponding sequence of the type-H isozyme lacking the insertion. The chimeric enzyme thus prepared genetically and expressed in *E. coli* was about five times less active than the parental type-L isozyme, but the activity was considerably high for an engineered protein, suggesting that the protein has been folded into roughly correct conformations.

A dramatic finding with the chimeric phosphorylase lacking the insertion was that it had a much higher affinity for large, branched glucans (amylopectin and glycogen) than the type-L isozyme, as shown by affinity electrophoresis (Fig. 6.4) and kinetic analysis. The K_m and K_d values of the chimeric enzyme for amylopectin and glycogen were even comparable with the values of the type-H isozyme. This result is fully consistent with the above-mentioned idea that the 78-residue insertion causes steric hindrance for large, branched glucan molecules approaching the active site of the type-L isozyme. Previous studies using affinity electrophoresis⁹⁴⁾ and kinetic analysis of cyclodextrin inhibition⁵⁰⁾ also showed that the type-L isozyme has only a single substrate-binding site located in the catalytic region, and suggested that the 78-residue insertion blocks the direct binding of branched glucans to the active site.

An alternative explanation for the increased affinity of the chimeric enzyme for branched glucans is also possible by assuming that the newly introduced segment of the type-H isozyme without the insertion contains a region with very high affinity for branched glucans. This region may correspond to a high affinity site shown to exist in the type-H isozyme separately from the active site.⁵⁰⁾ Although the amino acid sequence of this region of the type-H isozyme is rather unlike that of the rabbit muscle enzyme (see above), it may function like the glycogen-storage site in the rabbit enzyme, which has high affinity for both amylopectin and glycogen, but an abnormally low affinity for maltooligosaccharides.²⁸⁾ The approximate coincidence of K_m and K_d values for branched glucans of the type-H isozyme, the chimeric enzyme, and the rabbit enzyme⁹¹⁾ suggests the presence of a high affinity site on the molecular surface of these enzymes, which is probably different from the active site.

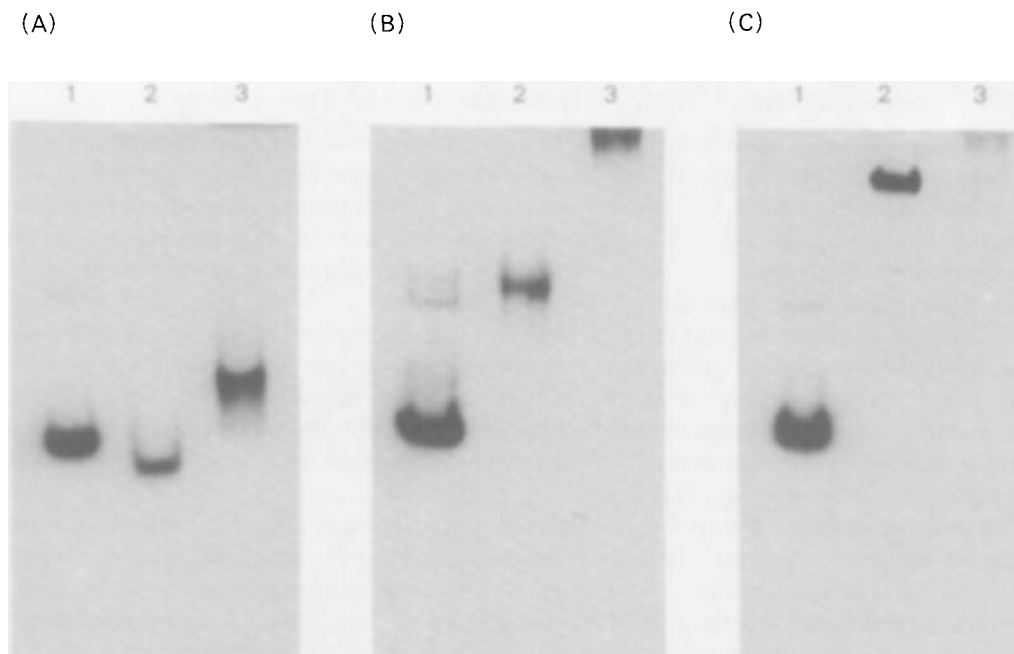


Fig. 6.4 Affinity electrophoresis of the recombinant phosphorylases from potato tuber.⁹³⁾ A crude bacterial cell extract containing the type-L isozyme (*lane 1*), the chimeric enzyme (*lane 2*), or the type-H isozyme (*lane 3*) was electrophoresed in 5% polyacrylamide gels supplemented with 0 (A), 50 (B), and 500 µg/ml (C) glycogen. After electrophoresis, the gels were stained for enzyme activity in KI-I₂ solution. (Reproduced with permission from Goldsmith and Fletterick, *Pure and Appl. Chem.*, **55** (4), 583 (1983)).

6.4 Catalytic Mechanism

Despite the significant differences in regulatory properties and substrate specificities among the enzymes from diverse origins, the catalytic mechanism of phosphorylase reaction is thought to be essentially the same.¹⁾ All the phosphorylases so far isolated contain one mole of pyridoxal-P per monomer, which is indispensable for the enzyme activity.⁹⁹⁾ The cofactor pyridoxal-P in phosphorylase participates in the catalysis in a manner quite different from other pyridoxal-P-dependent enzymes, because the enzyme reduced with sodium borohydride still shows more than 50% of the original activity¹⁰⁰⁾; the borohydride reduction generally leads to the complete inactivation of pyridoxal-P-dependent enzymes involved in amino acid metabolism, in which the 4-aldehyde group of the cofactor is involved in catalysis through the formation of a Schiff base with substrate amino acids.¹⁰¹⁾

Reconstitution experiments of the inactive apophosphorylase with various cofactor analogs provided the first evidence that the 5'-phosphate group but not the 4-aldehyde group of the cofactor is important in the catalytic activity.¹⁰²⁾ The binding site in the enzyme for the phosphate group was also characterized by using several analogues modified at the 5' position; the site is large enough to accommodate a bulky group and possesses cationic groups.¹⁰³⁾ In addition, X-ray crystallographic studies revealed that the 5'-

phosphate group of pyridoxal-P is in proximity to the phosphate group of the substrate glucose 1-phosphate, strongly supporting the view that the cofactor phosphate is directly involved in catalysis.¹⁰⁴⁾ Withers *et al.*¹⁰⁵⁾ and Takagi, Fukui and Shimomura¹⁰⁶⁾ found that pyridoxal(5')diphospho(1)- α -D-glucose, a synthetic conjugate of pyridoxal-P and glucose 1-phosphate through a pyrophosphate linkage, was bound to apophosphorylase at a reasonably high rate, and occupied the same site as the coenzyme itself. The reconstituted enzyme with this compound was catalytically inactive but, upon addition of glycogen, rapidly hydrolyzed to regenerate the catalytically inactive pyridoxal 5'-diphosphate-reconstituted enzyme, and the glucose moiety liberated was transferred to the nonreducing end of glycogen, forming a new α -1,4-glucosidic linkage. These findings provided the first evidence for the direct interaction between the coenzyme phosphate and the substrate phosphate in the phosphorylase catalytic mechanism. On the basis of precise analysis of the reaction of the conjugate-reconstituted enzyme,¹⁰⁷⁾ a possible catalytic mechanism of phosphorylase was proposed, which involves a constrained phosphate of pyridoxal-P acting as an electrophile to the phosphate oxygen of glucose 1-phosphate (Fig. 6.5(A)). Withers, Madsen and Sykes¹⁰⁸⁾ analyzed by ^{31}P -NMR the ternary complex of rabbit muscle phosphorylase *b*, maltopentaose, and glucose 1,2-cyclic phosphate, and suggested that the phosphorus of the constrained dianion of the coenzyme phosphate could be an electrophile interacting with the substrate phosphate.^{109,110)}

On the other hand, based on the ^{31}P -NMR analysis of the ternary complex of potato

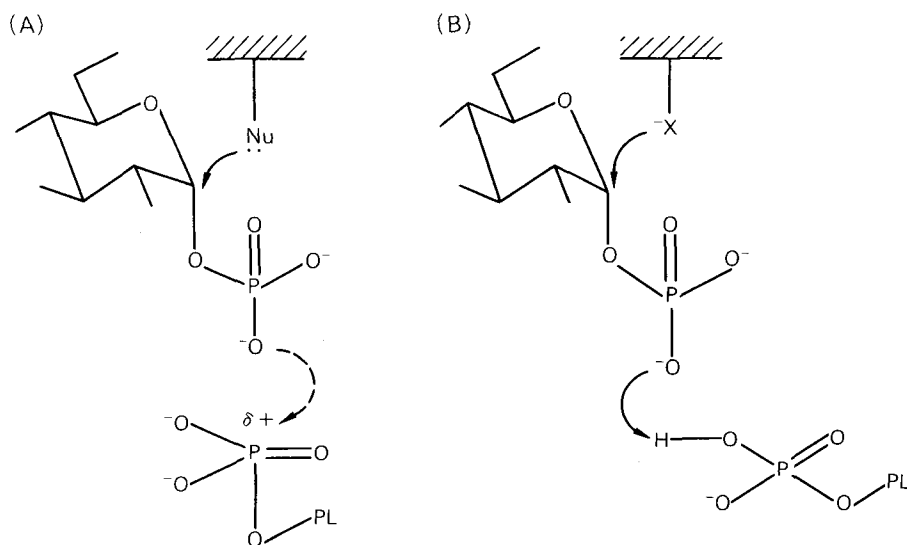


Fig. 6.5 Two proposed models for the role of the 5'-phosphate group of pyridoxal-P bound to phosphorylase. A, the cofactor phosphate acts as an electrophile to the phosphate oxygen of glucose 1-phosphate.^{105,106)} B, the cofactor phosphate acts as an acid-base catalyst donating a proton to the phosphate group of glucose 1-phosphate.¹¹¹⁾ (From *J. Biol. Chem.*, **268**, 5574 (1993)).

phosphorylase, Helmreich and Klein⁽¹¹⁾ proposed a mechanism in which the coenzyme phosphate acts as an acid-base catalyst promoting the protonation of the substrate phosphate (Fig. 6.5(B)). This acid-base mechanism has been supported by studies on phosphorylase catalysis of glycosidic substrates which yield a glucosyl residue on protonation⁽¹¹²⁻¹¹⁵⁾ and by the direct interaction observed between the cofactor 5'-phosphate and the product phosphate of heptulose 2-phosphate in the crystallographic observation.^(116,117) Recently, Palm *et al.*⁽¹¹⁸⁾ discussed the role of pyridoxal-P in the phosphorylase catalysis according to the acid-base mechanism. Thus, in the phosphorolysis direction, the reaction would be started by protonation of the glycosidic oxygen by orthophosphate, followed by stabilization of the incipient oxocarbenium ion by the phosphate anion and subsequent covalent binding of the phosphate to form α -glucose 1-phosphate. In the reverse direction, protonation of the phosphate of glucose 1-phosphate destabilizes the glycosidic bond and promotes formation of a glucosyl oxocarbenium ion-phosphate anion pair, then the phosphate anion promotes the nucleophilic attack of a terminal glucosyl residue on the carbonium ion, leading to the α -1,4-glycosidic bond formation and primer elongation.

Although further studies are needed to elucidate the exact role of pyridoxal-P and clarify the mechanism of phosphorylase catalysis, it is clear that the 5'-phosphate group of the cofactor bears a novel and interesting function that is unique among other pyridoxal-P-dependent enzymes. Studies by site-directed mutagenesis such as those on the recombinant *E. coli* maltodextrin phosphorylase replacing amino acid residues involved in catalysis⁽¹¹⁹⁾ or within hydrogen bond distance to the phosphate group of the cofactor⁽¹²⁰⁾ would be valuable in understanding in molecular detail the basic chemical principles and the structure-function relationships, which make phosphorylases favor phosphorolysis over hydrolysis.

6.5 Evolution of Plant Phosphorylases

The high sequence conservation among phosphorylases from a wide variety of sources as described above strongly suggests that these enzymes have evolved from a common ancestral protein, and sequence variations related to the regulatory properties (allosteric control and covalent modification), substrate specificities (including the glycogen-storage site), and subcellular localizations (the transit peptide in the plant type-L isozyme), and the unique 78-residue insertion in the type-L isozyme are all thought to have derived in a later stage of evolution. Furthermore, considering that the tertiary structure of the rabbit enzyme is composed mainly of two domains of similar size and that the 78-residue insertion in the plant type-L isozyme sequence is positioned near the junction of both domains, Nakano and Fukui⁽⁶⁵⁾ proposed the idea that the ancestral phosphorylase gene was generated by the fusion of two exons of nearly equal molecular size, each encoding a protein of about M_r 50,000. There is an intron in the human muscle phosphorylase gene⁽¹²¹⁾ within the region of the glycogen-storage site and exactly at the position of the plant 78-residue insertion. Elucidation of the complete intron positions in the two plant phosphorylase genes and comparison with those in the human muscle gene would be helpful to further assess this proposal.

The plant type-L isozyme is the only phosphorylase that has a large insertion in the middle of the polypeptide chain; the type-H isozyme has no such insertion, like other enzymes with or without regulatory properties. The sequence containing the insertion and

its flanking regions (residues 403–510 in the potato type-L isozyme) is unusual in that it contains many ionic amino acid residues (Asp, Glu, and Lys) but few arginyl and aromatic amino acid residues.⁶⁵⁾ The sequence of these regions is therefore likely to have derived from other unrelated proteins (or nearby introns through a change in mRNA splicing) and might have been inserted into the type-L isozyme polypeptide, probably sometime later than the period of the gene fusion generating various types of phosphorylases, along with the addition of the 50-residue putative transit peptide to the amino terminus.

6.6 Concluding Remarks

In view of the subcellular localization of the type-L isozyme in the amyloplast of potato tuber where starch is accumulated,¹²²⁾ the 78-residue insertion appears to be disadvantageous for acting on the highly branched glucans coexisting abundantly in the amyloplast, posing a question regarding the physiological significance of the type-L isozyme and its large insertion. It has previously been shown that the mRNA for the potato type-L isozyme is transcribed in the cells from the very early stage of tuber formation.⁷⁹⁾ However, this does not necessarily mean that the expressed type-L isozyme plays a role in the synthesis of starch. Rather, it is conceivable to assume that the type-L isozyme is stored in the amyloplast in an inefficient form until it is needed for degradation of starch during sprouting. Aged potato tubers contain many phosphorylase isoforms of smaller molecular size, which are derived from limited proteolysis.¹²³⁾ The insertion in the type-L isozyme may be nicked or cleaved off from the protein during sprouting to be converted into a high affinity form. The fact that the type-L isozyme is much more easily nicked in the insertion region than the type-H isozyme⁵⁹⁾ supports this speculation, although the change in affinity for substrate glucans of the type-L isozyme by the limited proteolysis remains to be studied. In any event, it is tempting to regard the 78-residue insertion in the plant type-L phosphorylase isozyme as a product of protein engineering by nature, designing the enzyme less efficient for branched substrates, which are otherwise very rapidly degraded.

In conclusion, it should be pointed out that in marked contrast to the very extensive studies on rabbit muscle phosphorylase, little attention has been paid to enzymes from other sources. However, primary structures of plant phosphorylases have now been determined and bacterial expression systems for the plant enzymes have also been made available as reviewed in this article. We hope that future studies on the structure and function of plant phosphorylases without allosteric regulation and comparison with those of the highly regulated animal enzyme will provide valuable information on this interesting group of enzymes, phosphorylases.

REFERENCES

1. D. J. Graves, J. H. Wang, *The Enzymes* (3rd Ed.), **7**, 435 (1972).
2. C. B. Newgard, P. K. Hwang, R. J. Fletterick, *Crit. Rev. Biochem. Mol. Biol.*, **24**, 69 (1989).
3. R. J. Fletterick, N. B. Madsen, *Ann. Rev. Biochem.*, **49**, 31 (1980).
4. E. G. Krebs, J. A. Beavo, *Ann. Rev. Biochem.*, **48**, 923 (1979).
5. E. H. Fischer, D. J. Graves, E. R. S. Crittenden, E. G. Krebs, *J. Biol. Chem.*, **234**, 1698 (1959).
6. E. H. Fischer, E. G. Krebs, *J. Biol. Chem.*, **216**, 121 (1955).

7. C. F. Cori, G. T. Cori, *Proc. Soc. Exp. Biol. Med.*, **34**, 702 (1936).
8. G. T. Cori, S. P. Colowick, C. F. Cori, *J. Biol. Chem.*, **123**, 381 (1938).
9. M. Fosset, L. W. Muir, L. D. Nielsen, E. H. Fischer, *Biochemistry*, **10**, 4105 (1971).
10. M. Schwartz, M. Hofnung, *Eur. J. Biochem.*, **2**, 132 (1967).
11. G. S. Chen, I. H. Segal, *Arch. Biochem. Biophys.*, **127**, 164 (1968).
12. Y.-P. Lee, *Biochim. Biophys. Acta*, **43**, 18 (1960).
13. Y.-P. Lee, *Biochim. Biophys. Acta*, **43**, 25 (1960).
14. A. Kamogawa, T. Fukui, Z. Nikuni, *J. Biochem.*, **63**, 361 (1968).
15. K.-H. Schaechtele, E. Schiltz, D. Palm, *Eur. J. Biochem.*, **92**, 427 (1978).
16. K. Titani, A. Koide, J. Hermann, L. H. Ericsson, S. Kumar, R. D. Wada, K. A. Walsh, H. Neurath, E. H. Fischer, *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 4762 (1977).
17. S. Iwata, T. Fukui, *FEBS Lett.*, **36**, 222 (1973).
18. H. D. Engers, S. Shechosky, N. B. Madsen, *Can. J. Biochem.*, **48**, 746 (1970).
19. A. M. Gold, R. M. Johnson, J. K. Tseng, *J. Biol. Chem.*, **245**, 2564 (1970).
20. H. D. Engers, W. A. Bridger, N. B. Madsen, *Can. J. Biochem.*, **48**, 755 (1970).
21. O. H. Lowry, D. W. Schulz, J. V. Passonneau, *J. Biol. Chem.*, **242**, 755 (1967).
22. H. D. Engers, W. A. Bridger, N. B. Madsen, *J. Biol. Chem.*, **244**, 5936 (1969).
23. V. T. Maddaiah, N. B. Madsen, *J. Biol. Chem.*, **241**, 3873 (1966).
24. A. M. Gold, R. M. Johnson, G. R. Sanchez, *J. Biol. Chem.*, **246**, 3444 (1971).
25. J. Chao, G. F. Johnson, D. J. Graves, *Biochemistry*, **8**, 1459 (1969).
26. D. H. Brown, C. F. Cori, *The Enzymes* (2nd Ed.), **5B**, 207 (1961).
27. S. H. Goldenberg, *Biochim. Biophys. Acta*, **56**, 357 (1962).
28. A. M. Liddle, D. J. Manners, A. Wrights, *Biochem. J.*, **80**, 304 (1961).
29. E. Y. C. Lee, J. H. Brown, *Arch. Biochem. Biophys.*, **156**, 276 (1973).
30. K. Nakano, P. K. Hwang, R. J. Fletterick, *FEBS Lett.*, **204**, 283 (1986).
31. R. J. Fletterick, J. Sygusch, H. Semple, N. B. Madsen, *J. Biol. Chem.*, **51**, 6142 (1976).
32. L. N. Johnson, N. B. Madsen, J. Mosely, K. S. Wilson, *J. Mol. Biol.*, **90**, 703 (1974).
33. K. R. Acharya, D. I. Stuart, K. M. Varvill, L. N. Johnson, *Glycogen Phosphorylase b: Description of the Protein Structure*, World Scientific, Singapore (1991).
34. D. Barford, L. N. Johnson, *J. Mol. Biol.*, **28**, 233 (1991).
35. M. F. Browner, R. J. Fletterick, *Trends Biochem. Sci.*, **17**, 66 (1992).
36. L. D. Gottlieb, *Science*, **216**, 373 (1982).
37. J. Preiss, *Annu. Rev. Plant Physiol.*, **33**, 431 (1982).
38. M. Steup, *Biochem. Plants*, **14**, 255 (1988).
39. E. Beck, P. Ziegler, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **40**, 95 (1989).
40. P. D. Simcox, D. T. Dennis, *Plant Physiol.*, **61**, 871 (1978).
41. M. Nishimura, H. Beevers, *Plant Physiol.*, **64**, 31 (1979).
42. G. Kakefuda, S. H. Duke, M. S. Hostak, *Planta*, **168**, 175 (1986).
43. T.-P. Lin, S. R. Spilatro, J. Preiss, *Plant Physiol.*, **86**, 251 (1988).
44. P. Ziegler, E. Beck, *Plant Physiol.*, **82**, 1119 (1986).
45. G. Kakefuda, S. H. Duke, *Plant Physiol.*, **91**, 136 (1989).
46. T. W. Okita, E. Greengerg, D. N. Kuhn, J. Preiss, *Plant Physiol.*, **64**, 187 (1979).
47. C. Schaechtele, M. Steup, *Planta*, **167**, 444 (1986).
48. C. Mateyka, C. Schnarrenberger, *Plant Sci. Lett.*, **36**, 119 (1984).
49. C. Mateyka, C. Schnarrenberger, *Plant Physiol.*, **86**, 417 (1987).
50. S. Shimomura, M. Nagai, T. Fukui, *J. Biochem.*, **91**, 703 (1982).
51. M. Steup, C. Schaechtele, *Planta*, **168**, 222 (1986).
52. S. J. Gerbrandy, A. Doorgeest, *Phytochemistry*, **11**, 2403 (1972).
53. E. E. Smith, *Arch. Biochem. Biophys.*, **146**, 380 (1971).
54. S. J. Gerbrandy, *Biochim. Biophys. Acta*, **370**, 410 (1974).
55. E. M. Schneider, J. V. Becker, D. Volkmann, *Planta*, **151**, 124 (1981).
56. N. Brisson, H. Giroux, M. Zollinger, A. Camirand, C. Simard, *Plant Cell*, **1**, 559 (1989).
57. J. Conrads, von J. Barkel, C. Schaechtele, M. Steup, *Biochim. Biophys. Acta*, **882**, 452 (1986).
58. M. Suda, T. Watanabe, M. Kobayashi, K. Matsuda, *J. Biochem.*, **102**, 471 (1987).
59. T. Fukui, K. Nakano, M. Tagaya, H. Nakayama, in: *Biochemistry of Vitamin B₆* (T. Korpela, P. Christen, Eds.) p. 267, Birkhauser Verlag, Basel and Boston (1987).
60. J. Preiss, C. Levi, in: *Photosynthesis II, Photosynthetic Carbon Metabolism and Related Processes* (M. Gibbs, E. Latzko, Eds.) p. 282, Springer-Verlag, Berlin (1979).
61. K. Nakano, S. Wakabayashi, T. Hase, H. Matsubara, T. Fukui, *J. Biochem.*, **83**, 1085 (1978).
62. K. Nakano, T. Fukui, H. Matsubara, *J. Biochem.*, **87**, 919 (1980).
63. K. Nakano, T. Fukui, H. Matsubara, *J. Biol. Chem.*, **255**, 9255 (1980).

64. K. Nakano, Y. Tashiro, Y. Kikumoto, M. Tagaya, T. Fukui, *J. Biol. Chem.*, **261**, 8224 (1986).
65. K. Nakano, T. Fukui, *J. Biol. Chem.*, **261**, 8230 (1986).
66. C.-T. Lin, K.-W. Yeh, P.-D. Lee, J.-C. Su, *Plant Physiol.*, **95**, 1250 (1991).
67. T. Fukui, S. Shimomura, K. Nakano, *Mol. Cell. Biochem.*, **42**, 129 (1982).
68. K. Nakano, H. Mori, T. Fukui, *J. Biochem.*, **106**, 691 (1989).
69. G. Heidecker, J. Messing, *Ann. Rev. Plant Physiol.*, **37**, 439 (1986).
70. M. Senser, F. Schotz, E. Beck, *Planta*, **126**, 1 (1975).
71. N. S. Scott, M. J. Tymms, J. V. Possingham, *Planta*, **161**, 12 (1984).
72. D. Macherel, H. Kobayashi, T. Akazawa, *Biochem. Biophys. Res. Commun.*, **133**, 140 (1985).
73. M. A. Catley, C. M. Bowman, M. W. Bayliss, M. D. Gale, *Planta*, **171**, 416 (1987).
74. G. W. Schmidt, M. L. Mishkind, *Ann. Rev. Biochem.*, **55**, 879 (1986).
75. S. V. Tingley, E. L. Walker, G. M. Coruzzi, *EMBO J.*, **6**, 1 (1987).
76. H. Brinkmann, P. Martinez, F. Quigley, W. Maetin, R. Cerff, *J. Mol. Evol.*, **26**, 320 (1987).
77. S. Smeeckens, D. Geerts, C. Baferle, P. Weisbeek, *Mol. Gen. Genet.*, **216**, 178 (1989).
78. M. Steup, C. Schaechtele, E. Latzko, *Planta*, **148**, 168 (1980).
79. H. Mori, K. Tanizawa, T. Fukui, *J. Biol. Chem.*, **266**, 18446 (1991).
80. J.-P. Boissel, T. J. Kasper, S. C. Shah, J. I. Malone, H. F. Bunn, *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 8448 (1985).
81. D. Twell, G. Ooms, *Plant Mol. Biol.*, **9**, 365 (1987).
82. D. Palm, R. Goerl, G. Weidinger, R. Zeier, B. Fischer, R. Schinel, *Z. Naturforsch., Teil C. Biochem. Biophys. Biol. Virol.*, **42**, 394 (1987).
83. R. J. Fletterick, S. R. Sprang, *Accts. Chem. Res.*, **15**, 261 (1982).
84. M. S. P. Sanson, Y. S. Babu, J. Hajdu, D. I. Stuart, E. A. Stura, L. N. Johnson, in: *Chemical and Biological Aspects of Vitamin B₆ Catalysis* (A. E. Evangelopoulos, Ed.) Part A, p. 127 (1984).
85. N. G. Oikonomakos, L. N. Johnson, K. R. Acharya, D. I. Stuart, D. Barford, J. Hajdu, K. M. Varvill, A. E. Melpidou, T. Papageorgiou, D. J. Graves, D. Palm, *Biochemistry*, **26**, 8381 (1987).
86. K. Lerch, E. H. Fischer, *Biochemistry*, **14**, 2009 (1975).
87. D. Palm, R. Goerl, K. J. Burger, *Nature*, **313**, 500 (1985).
88. S. Shimomura, K. Nakano, T. Fukui, *Biochem. Biophys. Res. Commun.*, **82**, 462 (1978).
89. M. Tagaya, K. Nakano, S. Shimomura, T. Fukui, *J. Biochem.*, **91**, 599 (1982).
90. B. Vandenbunder, M. Dreyfus, O. Bertrand, M. J. Dognin, L. Sibilli, H. Buc, *Biochemistry*, **20**, 2354 (1981).
91. P. J. Kasvinsky, N. B. Madsen, R. J. Fletterick, J. Sygusch, *J. Biol. Chem.*, **253**, 1290 (1978).
92. E. Goldsmith, R. J. Fletterick, *Pure and Appl. Chem.*, **55**, 577 (1983).
93. H. Mori, K. Tanizawa, T. Fukui, *J. Biol. Chem.*, **268**, 5574 (1993).
94. S. Shimomura, T. Fukui, *Biochemistry*, **19**, 2287 (1980).
95. J. Staerk, H. Schlenk, *Biochim. Biophys. Acta*, **146**, 120 (1967).
96. J. Willnecker, K. Jahnke, M. Buchner, in: *Enzymes Dependent on Pyridoxal Phosphate and Other Carbonyl Compounds as Cofactors* (T. Fukui, H. Kagamiyama, K. Soda, H. Wada, Eds.) p. 393, Pergamon Press, Oxford (1991).
97. F. W. Studier, A. H. Rosenberg, J. J. Dunn, J. W. Dubendorff, *Methods Enzymol.*, **185**, 60 (1990).
98. M. F. Browner, P. Rasor, S. Tugendreich, R. J. Fletterick, *Protein Eng.*, **4**, 351 (1991).
99. S. Shaltiel, J. L. Hedrick, E. H. Fischer, *Biochemistry*, **5**, 2108 (1966).
100. E. H. Fischer, A. B. Kent, E. R. Snyder, E. G. Krebs, *J. Am. Chem. Soc.*, **80**, 2906 (1958).
101. E. H. Fischer, E. R. Snyder, E. G. Krebs, *J. Amer. Chem. Soc.*, **80**, 2906 (1958).
102. S. Shaltiel, J. L. Hedrick, A. Pocker, E. H. Fischer, *Biochemistry*, **8**, 5189 (1969).
103. S. Shimomura, T. Fukui, *Biochemistry*, **17**, 5359 (1978).
104. J. Sygusch, N. B. Madsen, P. J. Kasvinsky, R. J. Fletterick, *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 4757 (1977).
105. S. G. Withers, N. B. Madsen, B. D. Sykes, M. Takagi, S. Shimomura, T. Fukui, *J. Biol. Chem.*, **256**, 10759 (1981).
106. M. Takagi, T. Fukui, S. Shimomura, *Proc. Natl. Acad. Sci. U.S.A.*, **79**, 3716 (1982).
107. M. Tagaya, T. Fukui, *J. Biol. Chem.*, **259**, 4860 (1984).
108. S. G. Withers, N. B. Madsen, B. D. Sykes, *Biochemistry*, **20**, 1748 (1981).
109. S. G. Withers, N. B. Madsen, B. D. Sykes, *Biochemistry*, **21**, 6716 (1982).
110. N. B. Madsen, *The Enzymes* (3rd Ed), **17**, 366 (1986).
111. E. J. M. Helmreich, H. W. Klein, *Angew. Chem. Int. Ed. Engl.*, **19**, 441 (1980).
112. H. W. Klein, D. Palm, E. J. M. Helmreich, *Biochemistry*, **21**, 6675 (1982).
113. H. W. Klein, M. J. Im, D. Palm, E. J. M. Helmreich, *Biochemistry*, **23**, 5853 (1984).
114. D. Palm, G. Blumenauer, H. W. Klein, M. Blanc-Muesser, *Biochem. Biophys. Res. Commun.*, **111**, 530 (1983).
115. H. W. Klein, M. J. Im, D. Palm, *Eur. J. Biochem.*, **157**, 107 (1986).
116. P. J. McLaughlin, D. I. Stuart, H. W. Klein, N. G. Oikonomakos, L. N. Johnson, *Biochemistry*, **23**, 5862

- (1984).
117. J. Hajdu, K. R. Acharya, D. I. Stuart, P. J. McLaughlin, D. Barford, N. G. Oikonomakos, H. Klein, L. N. Johnson, *EMBO J.*, **6**, 539 (1987).
118. D. Palm, H. W. Klein, R. Schinzel, M. Buehner, J. M. Helmreich, *Biochemistry*, **29**, 1099 (1990).
119. R. Schinzel, *J. Biol. Chem.*, **266**, 9428 (1991).
120. R. Schinzel, D. Palm, *Biochemistry*, **29**, 9956 (1990).
121. J. Burke, P. K. Hwang, L. Abderson, R. Lebo, F. Gorin, R. J. Fletterick, *Proteins: Struc., Func., and Gen.*, **2**, 177 (1987).
122. E. M. Schneider, J.-U. Becker, D. Volkmann, *Hoppe-Seyler's Z. Physiol. Chem.*, **360**, 369 (1979).
123. S. J. Gerbrandy, V. Shankar, K. N. Shivaram, H. Stegemann, *Phytochemistry*, **14**, 2331 (1975).

7

Tryptophan Synthase

Edith Wilson Miles, Syed Ashrafuddin Ahmed, C. Craig Hyde,
Arvind M. Kayastha*, Xiang-Jiao Yang, Sergei B. Ruvinov, and
Zichun Lu

*National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health
Building 8, Room 2A09, Bethesda, MD 20892*

7.1 Introduction

Tryptophan synthase (EC 4.2.1.20) from bacteria is a classic multienzyme complex that channels a metabolic intermediate (for reviews and commentaries see 1–6)). Structure/function analysis of tryptophan synthase was pioneered by genetic and biochemical investigations beginning in the mid-1940's. This chapter emphasizes the relationship between the function of tryptophan synthase and the three-dimensional structure of the tryptophan synthase $\alpha_2\beta_2$ complex from *Salmonella typhimurium*^{7,8)} and focuses on studies carried out since a recent review.⁶⁾

Tryptophan synthase from *S. typhimurium* and from *Escherichia coli* is an $\alpha_2\beta_2$ complex ($M_r=143,400$) that can be dissociated reversibly into two monomeric α subunits ($M_r=28,700$)⁹⁾ and one dimeric β_2 subunit ($M_r=86,000$),¹⁰⁾ termed here the β subunit. Although bacterial tryptophan synthase is composed of α and β subunits, fungal tryptophan synthase is a single polypeptide chain. The amino acid sequences of the multifunctional fungal enzymes show strong homology with the amino acid sequences of the α and β subunits from bacteria^{11–14)} and of the β subunit from a plant.^{15,16)} The fact that enzymes from various sources have strong sequence homology implies that they have similar three-dimensional structures.

The three-dimensional structure of the tryptophan synthase $\alpha_2\beta_2$ complex from *S. typhimurium* reveals that the four polypeptide subunits are arranged in an extended $\alpha\beta\beta\alpha$ order forming a complex 150 Å long.⁷⁾ A schematic view of a single α/β pair based on the crystal structure is shown in the color plate ①. The $\alpha_2\beta_2$ complex catalyzes the synthesis of L-tryptophan from indole-3-glycerol phosphate and L-serine, termed the $\alpha\beta$ reaction (Fig. 7.1). The α and β subunits can be separated and shown to catalyze two distinct reactions, termed the α and β reactions, respectively (Fig. 7.1). The rates of the α and β reactions are greatly increased when catalyzed by the $\alpha_2\beta_2$ complex. Although the $\alpha\beta$ reaction is formally the sum of the α and β reactions, indole does not appear as a free intermediate in solution in this reaction.^{17–21)} This result indicates that indole is a

* Present address: School of Biotechnology, Banaras Hindu University, Varanasi-221005, INDIA

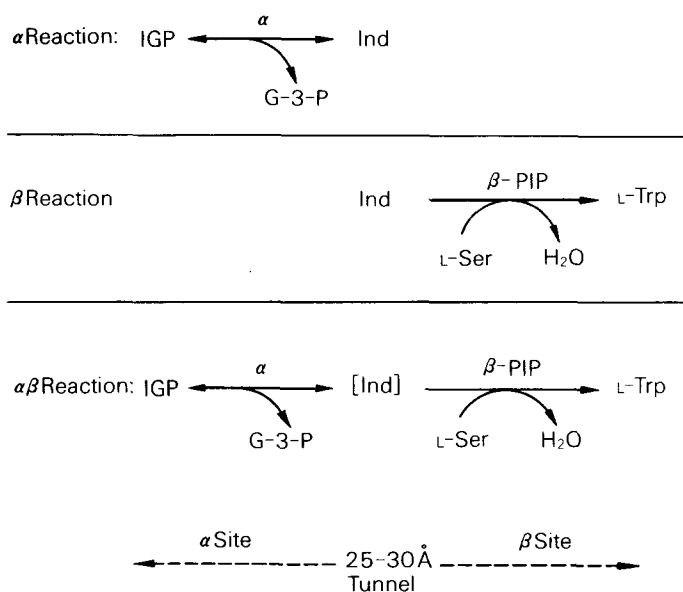


Fig. 7.1 Reactions catalyzed at the active sites of the α subunit (α reaction) and of the β subunit (β reaction) and the coupled, physiological reaction ($\alpha\beta$ reaction). In the $\alpha\beta$ reaction, indole produced by cleavage of indole-3-glycerol phosphate at the α site diffuses through an intramolecular tunnel to the β site 25–30 Å distant where it undergoes a pyridoxal phosphate-dependent β -replacement reaction with L-serine to form L-tryptophan. Abbreviations used: IGP, indole-3-glycerol phosphate; G-3-P, D-glyceraldehyde 3-phosphate; IND, indole; [IND], indole intermediate; PLP, pyridoxal phosphate.

“channeled” intermediate that is transferred intramolecularly from the α site to the β site in the $\alpha\beta$ reaction.

The active sites of the α and β subunits are separated by 25–30 Å and are connected by a tunnel of sufficient size to permit direct diffusion of indole between the heterologous active sites (Fig. 7.1).⁷⁾ Fig. 7.1 shows our current picture of the mechanism of the $\alpha\beta$ reaction based on the crystal structure. The indole intermediate [IND] in Fig. 7.1, which is produced by cleavage of indole-3-glycerol phosphate at the α site, diffuses through an intramolecular tunnel to the β site where it undergoes a pyridoxal phosphate-dependent β -replacement reaction with L-serine to form L-tryptophan. The catalytic mechanism of this reaction is described further in Section 7.7.

7.2 Structure and Function of the α Subunit

7.2.1 Structure

The overall polypeptide fold of the α subunit is that of an 8-fold α/β barrel (Fig. 7.2 and color plate ②).⁷⁾ Similar structures have been observed in more than 18 other

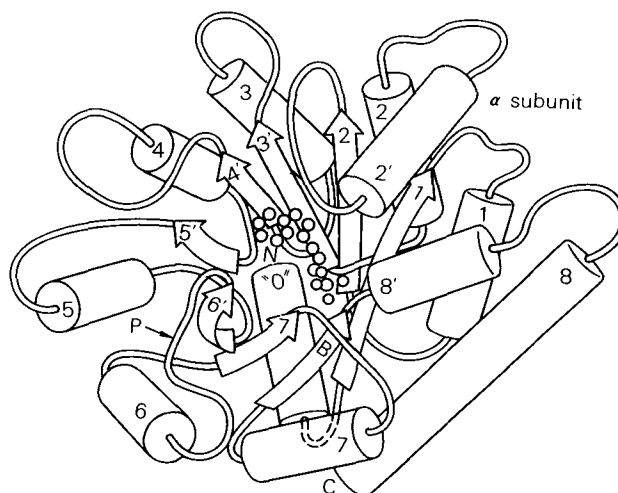


Fig. 7.2 The α subunit has an 8-fold α/β barrel tertiary fold. β -Strands are shown as flattened arrows with arrowheads at their C-termini. α -Helices are represented as cylinders and are labeled on their N-termini. In addition to the eight strands and helices found in a canonical α/β barrel, the α subunit contains three other helices labeled 0, 2' and 8'. N and C mark the N- and C-termini. P marks the site of proteolysis at Arg-188 that occurs in the darkened, flexible loop (residues 179–192) following strand 6. A darkened, disordered loop following strand 2 contains a catalytic residue, Asp-60, and makes important contact with the α active site and with the β subunit. A substrate analog, indole-3-propanol phosphate, binds at the active site as indicated by the ball-and-stick model. (See also color plate). (Reproduced with permission from Hyde *et al.*, *J. Biol. Chem.*, **263**, 17857 (1988)).

enzymes.^{22,23)} The α subunit is most closely related in structure to three of these enzymes: triosephosphate isomerase and both parts of the bifunctional enzyme (phosphoribosyl)anthranilate isomerase: indole-3-glycerol phosphate synthetase.^{22,24–28)}

7.2.2 Active Site Residues

The active center of the α subunit has been located by X-ray crystallography of the $\alpha_2\beta_2$ complex containing a bound substrate analog, indole-3-propanol phosphate.⁷⁾ This analog binds in a hydrophobic “pit” formed just below the surface of the α subunit at the top of the central barrel (Fig. 7.2 and color plate ②). The phosphate binding site is created by residues in the loop between β strand 7 and helix 7 and the N-terminus of an additional helix 8'.²⁷⁾ Interestingly, this same phosphate binding site is shared by the three closely related enzymes: triosephosphate isomerase, (phosphoribosyl)anthranilate isomerase, and indole-3-glycerol phosphate synthetase.²⁷⁾ Nine of the ten residues that are sites of inactivating missense mutations or second-site reversions^{1,29,30)} (residues 22, 49, 60, 175, 177, 211, 213, 234, and 235) are located near the bound inhibitor. Two of these residues, Glu-49 and Asp-60, are thought to be catalytic residues that facilitate the reversible cleavage of the C3'–C3 bond of indole-3-glycerol phosphate as shown in Fig. 7.3.^{31–33)} The tenth site of an inactivating missense mutation or second site-reversion, Thr-183,^{1,29)} may also be an active site residue (see Section 7.2.3).

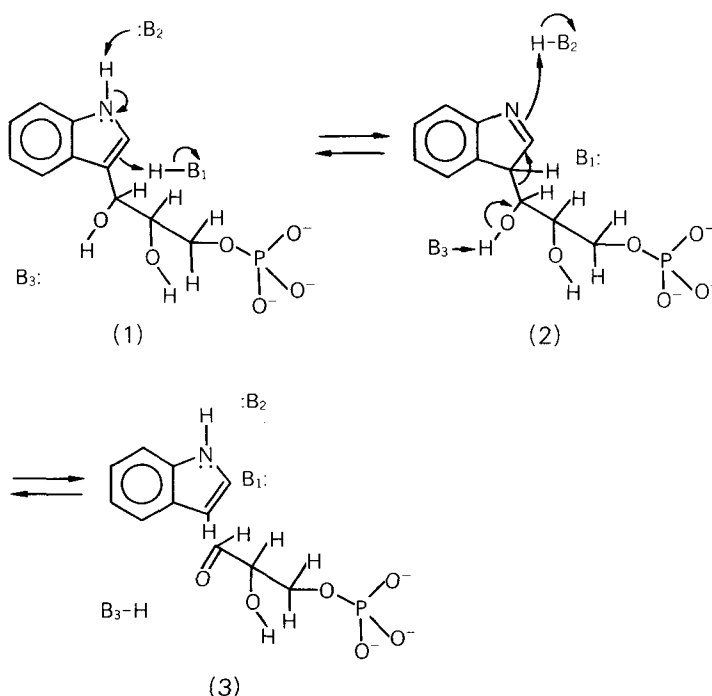


Fig. 7.3 Proposed mechanism of the α reaction. The cleavage of indole-3-glycerol phosphate (1) is activated by tautomerization of the indole ring to yield the indolenine tautomer (2) which undergoes aldol cleavage to yield (3). These steps are catalyzed by three putative residues. Asp-60 is thought to be B_2 which abstracts the proton on N-1 of the indole ring or forms a strong hydrogen bond to polarize the nitrogen atom. Glu-49 is thought to serve as B_3 , which serves as the proton acceptor for removing the hydroxyl proton. Glu-49 may also serve as B_1 that protonates the C3 position of the indole moiety.¹¹²⁾ (Reproduced with permission from Nagata, Hyde, Miles, *J. Biol. Chem.*, **264**, 6288 (1989)).

7.2.3 Flexible Loop

Residues 179–192 (Figs. 7.2 and 7.4 and color plate ①) between strand 6 and helix 6 of the α subunit are disordered and cannot be seen in the crystal structure.⁷⁾ This very long, disordered loop contains a site (Arg-188) that is cleaved by trypsin in the $\alpha_2\beta_2$ complex^{34–37)} (see arrows in Fig. 7.2 and 7.4). Protein folding studies described in Section 7.2.4 provide evidence that the flexible loop connects two folding domains.^{37,38)} It is noteworthy that the loop following β strand 6 is also extraordinarily long and highly flexible in two of the α/β barrel enzymes that are closely related to the α subunit in structure: (phosphoribosyl)anthranilate isomerase²⁷⁾ and triosephosphate isomerase.^{25,26,39)} The loop in triosephosphate isomerase moves toward the active site and becomes less flexible upon the addition of a substrate analog. This conformational change converts triosephosphate isomerase from an “open” to a “closed” form.^{25,26,39,40)}



Fig. 7.4 Sequence of the flexible loop that connects helix 6 and strand 6 in the α subunit (see Fig. 7.2). This loop has very weak electron density and cannot be visualized in the three-dimensional structure of the $\alpha_2\beta_2$ complex.^{7,27)} Conserved residues are underlined. Arg-179,¹¹⁴⁾ and Thr-183⁴²⁾ have been replaced by site-directed mutagenesis; arginine 188 is the only site that is susceptible to proteolysis in the $\alpha_2\beta_2$ complex^{35,36)} (modified from Ref. 42).

To investigate the possibility that substrate binding induces a conformational change in the flexible α subunit loop, we have studied the effect of ligands on the rate of proteolysis of the α subunit at Arg-188 (Fig. 7.4).³⁶⁾ Our finding that the addition of the substrate analog, α -glycerol 3-phosphate, decreases the rate of cleavage of the α subunit in the $\alpha_2\beta_2$ complex about 6-fold suggests that α -glycerol 3-phosphate induces a conformational change that decreases the susceptibility of the loop to tryptic cleavage. This putative conformational change may be stabilized by ligands of the β subunit since addition of L-serine in combination with α -glycerol 3-phosphate completely protects the $\alpha_2\beta_2$ complex from proteolysis.⁴¹⁾ We have replaced another residue in the loop, Thr-183, by alanine. Since ligands of the α subunit and β subunit do not protect this mutant form of the $\alpha_2\beta_2$ complex from proteolysis,⁴²⁾ Thr-183 may be essential for a ligand-induced conformational change in the loop. The hydroxyl of Thr-183 may form a hydrogen bond with the phosphate of the α subunit ligand. Possible allosteric roles of the flexible loop are discussed in Section 7.6.

7.2.4 Protein Folding and Stability

Although the native conformation of the α subunit is a single structural domain,⁷⁾ the α subunit is unfolded by guanidine hydrochloride or urea via a stable intermediate in which the N-terminal domain (residues 1–188) remains folded whereas the C-terminal domain (residues 189–268) becomes disordered.^{37,38,43,44)} Recent kinetic studies of the folding and unfolding of wild type and mutant forms of the α subunit show that the association or “docking” of the amino domain and the carboxyl domain are a late step in the folding process.^{45,46)} Amino acid replacements can selectively affect the energy of interaction between these two autonomous folding units.⁴⁷⁾

The effects of single amino acid substitution on the conformational stability of the α subunit have been extensively investigated.^{43,46–58)} These studies provide important information on factors that influence protein stability. They also help to characterize the effects of mutation on protein conformation.

7.3 Structure and Function of the β Subunit

7.3.1 Structure and Domains

The β subunit contains two structural domains that correspond approximately to the N- and C-terminal halves of the polypeptide chain (Fig. 7.5 and color plates ① and ②).⁷⁾ Since the central core regions of these two domains have very similar folding topologies and are

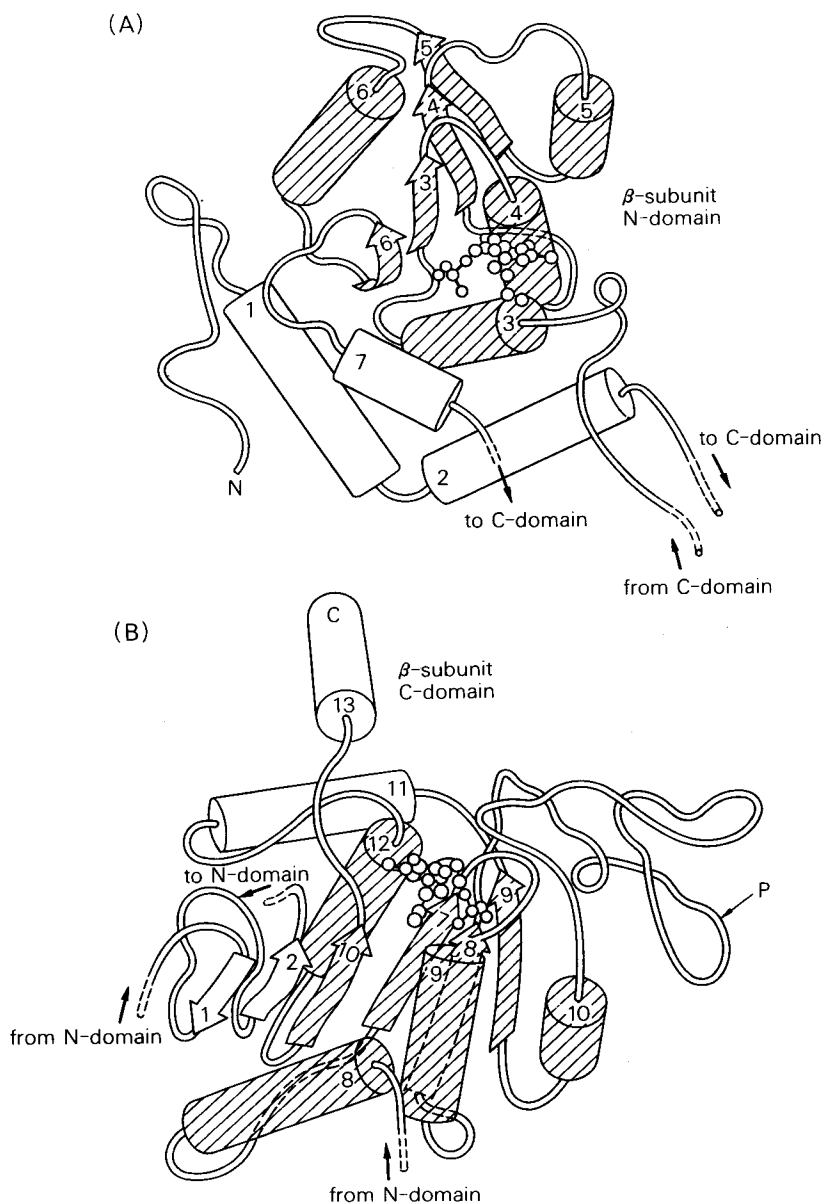


Fig. 7.5 Folding patterns of the two domains of the β subunit. The core regions of these two domains (shown by diagonal stripes) have very similar folding topologies and are nearly superimposable. The pyridoxal phosphate: Lys-87 Schiff base complex, shown in a ball-and-stick model, lies on the surface of each domain. P points to a region that contains several sites that are susceptible to proteolysis in the separate β subunit (see text). The cleavage sites (Lys-272, Arg-275, Lys-283, and Glu-296) are within a stretch of residues (260–310) that contains three long β hairpin loops and apparently lacks any other well-defined secondary structure (modified from Ref. 7). (See also color plates ① and ②.)

nearly superimposable, the β subunit may have evolved via gene duplication and fusion. The pyridoxal phosphate coenzyme is “sandwiched” between the N- and C-terminal domains and is shown on the surface of each domain in Fig. 7.5.

Mild proteolytic treatment of the β subunit results in cleavage of the polypeptide chain in a protease-sensitive region between residues 272 and 296 (see arrow in Fig. 7.5); proteolysis yields an N-terminal fragment (F1) and a C-terminal fragment (F2).^{59–61)} The isolated F1 and F2 fragments refold independently into globular proteins.^{59,62,63)} Recent studies demonstrate that the isolated F2 fragment folds into a stable, organized, nonnative conformation that may represent a precursor to the “molten globule” on the folding pathway of the C-terminal region of the β chain.⁶⁴⁾ The F2 fragment may fail to proceed further on the folding pathway because it lacks the side chains of F1 required for creating the specific contacts leading to the native state.⁶⁴⁾ Investigations of the refolding kinetics of the β subunit include studies using monoclonal antibodies that recognize distinct epitopes on the F1 and F2 fragments. The results provide evidence that refolding proceeds through a series of intermediate steps.^{65–71)} An early folding intermediate is a “molten globule”.⁷²⁾

7.3.2 Pyridoxal Phosphate Intermediates

The chromophoric pyridoxal phosphate coenzyme provides a useful spectrophotometric probe of catalytic events and of conformational changes that occur at the pyridoxal phosphate site of the β subunit and of the $\alpha_2\beta_2$ complex. Tryptophan synthase belongs to a class of pyridoxal phosphate enzymes that catalyze β -replacement and β -elimination reactions.³⁾ The reactions proceed through a series of pyridoxal phosphate-substrate intermediates (Fig. 7.6) that have characteristic spectral properties. Steady-state and rapid kinetic studies of the β subunit and of the $\alpha_2\beta_2$ complex in solution have demonstrated the formation and disappearance of these intermediates.^{73–90)} Fig. 7.7 illustrates the use of rapid-scanning stopped-flow UV-visible spectroscopy to investigate the effects of single amino acid substitutions in the α subunit on the rate of reactions of L-serine at the active site of the β subunit.⁸⁹⁾ Formation of enzyme-substrate intermediates has also been observed with the $\alpha_2\beta_2$ complex in the crystalline state.⁹¹⁾

7.3.3 Active Site Residues

Crystallographic studies have identified the active site of the β subunit by the presence of pyridoxal phosphate, which is bound in Schiff base linkage with the ϵ -amino group of Lys-87.⁷⁾ Although the substrate binding site of the β subunit has not yet been located by X-ray crystallography, this site must be adjacent to the coenzyme binding site since Schiff base intermediates are formed between the substrate and the coenzyme. We are using site-directed mutagenesis to evaluate the roles of residues in the active site of the β subunit. Fig. 7.8 shows the pyridoxal phosphate: Lys-87 Schiff base complex, the tunnel, the putative indole binding site, and several of the amino acid residues that have been selected for site-directed mutagenesis.^{88,92–96)}

The ϵ -amino group of lysine-87 forms external aldimine E in Fig. 7.6 and is released after L-serine binds to form ES I in Fig. 7.6. Lysine-87 is ideally situated to facilitate catalysis by removing the α -proton of L-serine.⁷⁾ We find that a mutant form of the $\alpha_2\beta_2$ complex in which the β subunit lysine-87 is replaced by threonine (K87T) is inactive but binds L-serine and L-tryptophan slowly and extremely tightly.^{92,97)} These results provide evidence that

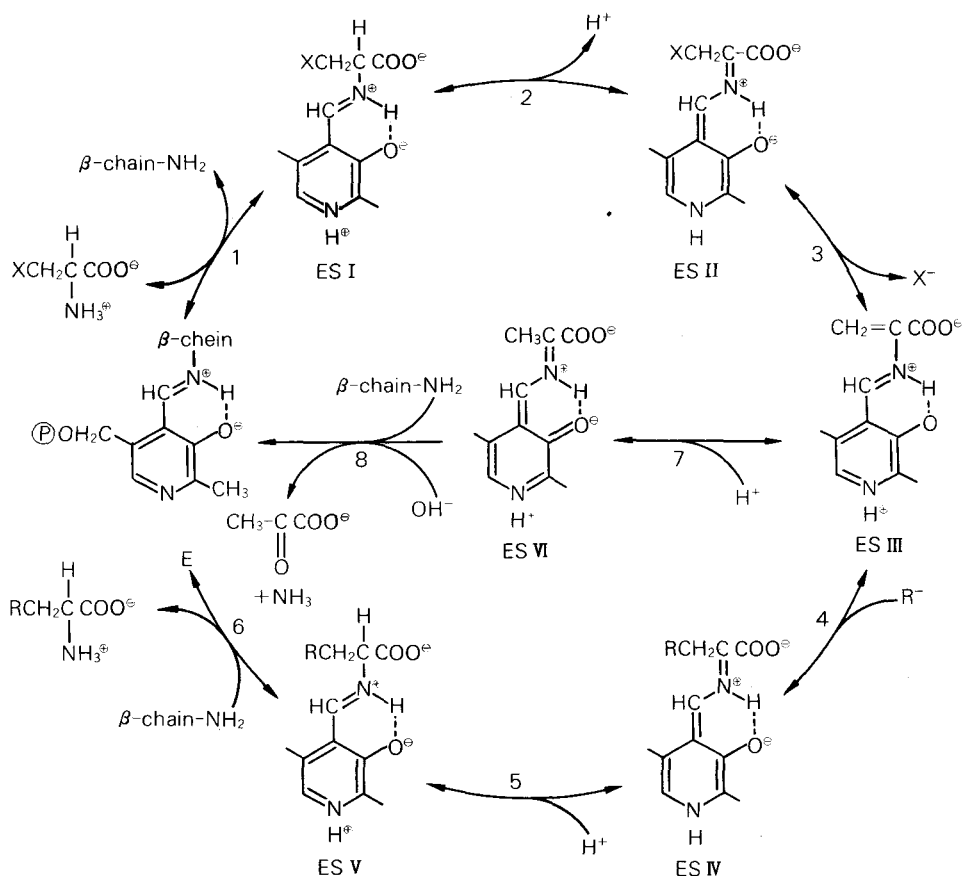


Fig. 7.6 Mechanism of β -replacement and β -elimination reactions with L-serine ($X^- = \text{OH}^-$) or β -chloro-L-alanine ($X^- = \text{Cl}^-$). Formation of the Schiff base intermediate with the amino acid ES I is followed by removal of the α -proton (H^+) and of the leaving group (X^-) to form the Schiff base of amino acrylate ES III, the key intermediate in both types of reaction. ES III can be hydrolyzed to pyruvate and NH_3 (β -elimination) or can add the indole cosubstrate (RH) to form the Schiff base of the quinonoid of L-tryptophan ES IV (β -replacement). Protonation of ES IV leads to release of L-tryptophan. ES IV can also be formed in the reverse direction from L-tryptophan. (Reproduced with permission from Kayastha, *J. Biol. Chem.*, **266**, 7618 (1991)).

lysine-87 in the wild-type β subunit plays important roles in transamination, catalysis, and product release. Lysine-87 appears to catalyze the removal of the α -proton of L-serine and may also protonate the weak hydroxyl leaving group of L-serine. Circular dichroism studies show that bound pyridoxal phosphate has a much weaker ellipticity at 400 nm in the K87T $\alpha_2\beta_2$ complex than in the wild type enzyme.^{96,97} The results demonstrate the importance of the Schiff base linkage for rigid or asymmetric binding of pyridoxal phosphate.

An important intermediate step in the reaction of L-serine and indole to form L-tryptophan is the nucleophilic attack of indole on the Schiff base of amino acrylate (ES III

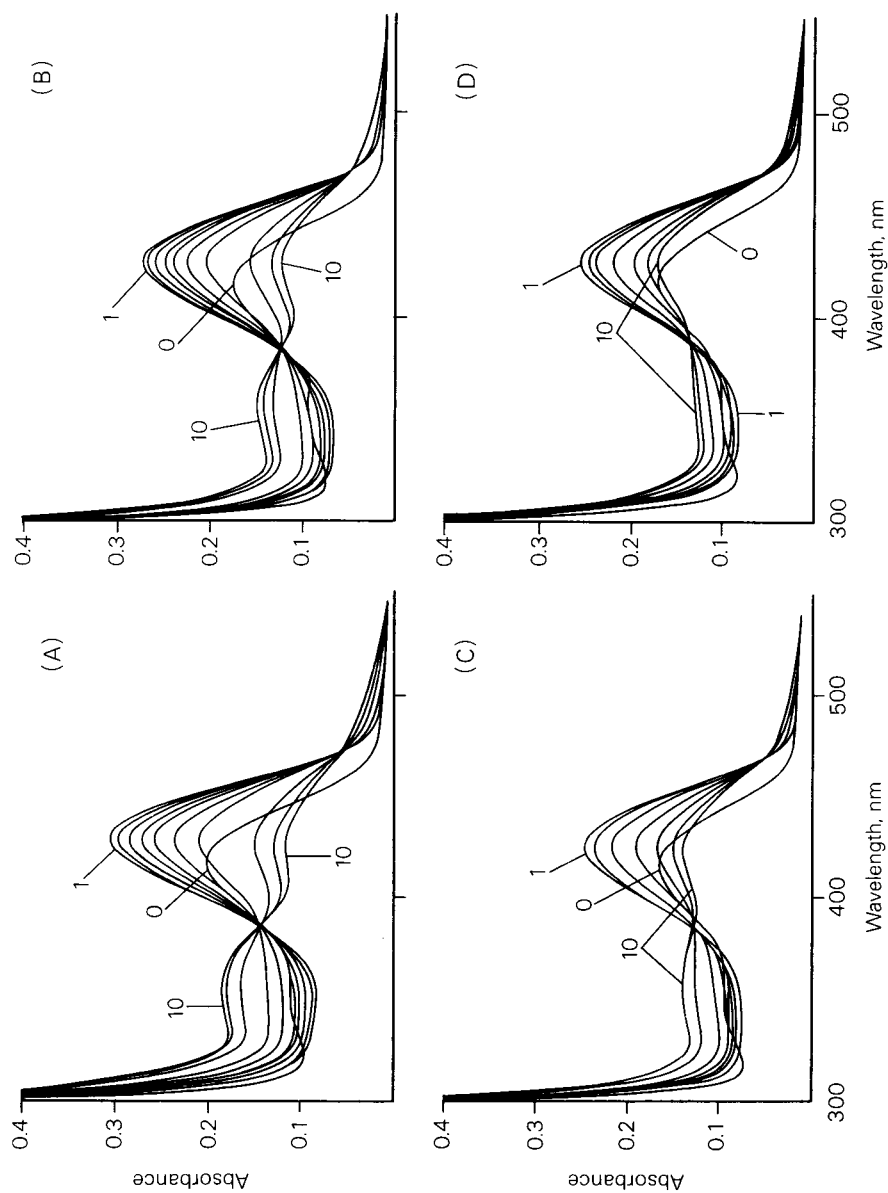


Fig. 7.7 Rapid-scanning stopped-flow spectra for the reaction of 10 μ M wild and mutant $\alpha_2\beta_2$ complexes with 40 mM L-serine at pH 7.8. A) Wild type, B) α E49F, C) α G51L and D) α D60Y. Spectra were collected at various intervals (0–640 msec) after mixing. (Reproduced with permission from Brzoric, *J. Biol. Chem.*, **267**, 13028 (1992)).

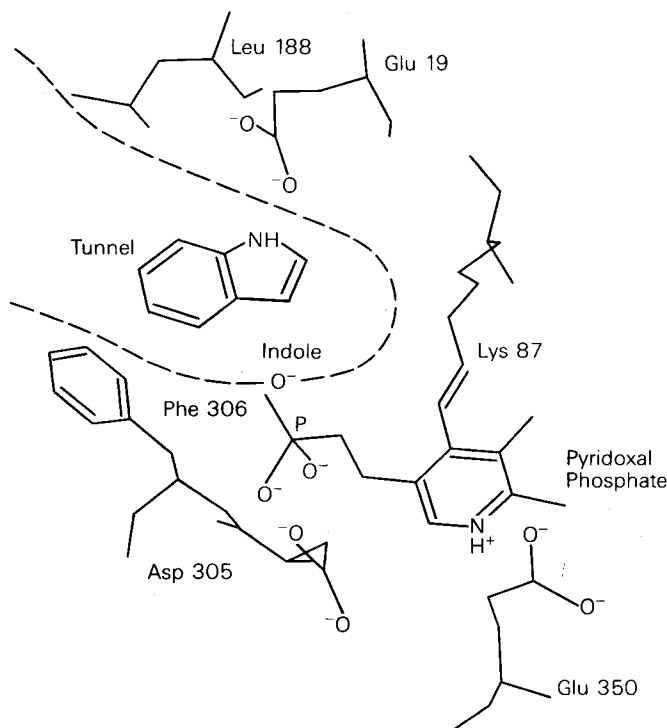


Fig. 7.8 Active site of the β subunit based on X-ray crystallographic results.⁷⁾ The figure shows the pyridoxal phosphate: Lys-87 Schiff base, the putative indole binding site at the end of the tunnel (....) that extends from the α site to the β site, and residues selected for site directed mutagenesis. Sources: Refs. 95 and 102

in Fig. 7.6). Since indole is a very weak nucleophile, the catalysis of this reaction must involve an activated intermediate. The proposal that the pyridoxal phosphate Schiff base of the indolenine tautomer of L-tryptophan (ES IVA in Fig. 7.9) is the activated intermediate⁷⁸⁾ is supported by studies with reaction intermediate analogs.⁹⁸⁻¹⁰⁰⁾ The formation of ES IVA is probably facilitated by a basic group, B₂, which abstracts the proton on N-1 of the heterocyclic ring concomitant with attack of indole on the aminoacrylate. The tautomerization of ES IVA to yield ES IVB is carried out by abstraction of the proton at C₃ of indole by a basic group, B₁, concomitant with protonation of N-1 by the protonated B₂. An examination of the active site of the β subunit in the three-dimensional structure of the $\alpha_2\beta_2$ complex and computer graphics modeling studies suggests that the γ -carboxyl of glutamate-109 may be located near the putative indole binding site (Fig. 7.8) and may play a role in the activation of indole. Several studies have investigated the effects of substitution of β subunit glutamate-109 by aspartate (E109D) or by alanine (E109A) (Fig. 7.8).^{87,88,93,95,101,102)} Rapid-scanning stopped-flow UV-visible spectroscopy⁸⁸⁾ and quench flow kinetic studies under single-turnover conditions⁸⁷⁾ show that the E109D mutation does not seriously affect the pre-steady state reaction of the mutant $\alpha_2\beta_2$ complex

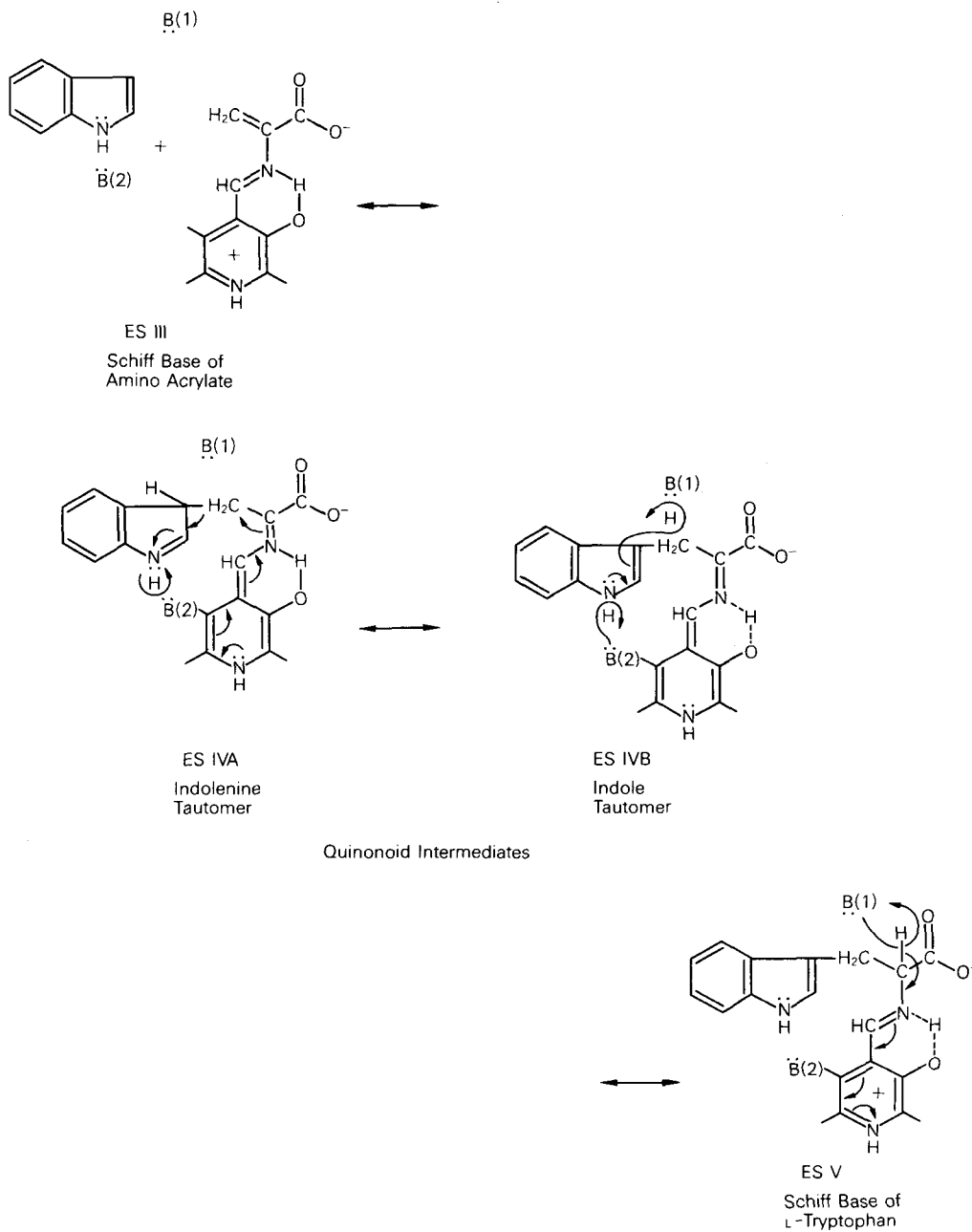


Fig. 7.9 Mechanism of formation of tautomeric quinonoid intermediates in the synthesis of L-tryptophan. The conversion of ES III (see also Fig. 7.6) to ES IVA is probably catalyzed by B₂ which abstracts the proton on N-1 of the heterocyclic ring concomitant with attack of indole on the aminoacrylate. The tautomerization of ES IVA to yield ES IVB is carried out by abstraction of the proton at C₃ of indole by a basic group, B₁, concomitant with protonation of N-1 by B₂H (modified from Ref. 6).

with L-serine to form the amino acrylate intermediate (ES III in Fig. 7.6). However, the E109D mutation drastically reduces the rate of addition of indole to the amino acrylate intermediate to form the quinonoid intermediate (ES IV in Fig. 7.6). A structural change within the β subunit active site may be responsible for a large alteration in the nucleophile specificity of the E109D $\alpha_2\beta_2$ complex.⁸⁸⁾

The carboxylate of glutamic acid 350 is located near the pyridinium nitrogen of pyridoxal phosphate in the three-dimensional structure (Fig. 7.8).⁷⁾ The functional role of glutamic acid 350 has been probed by substitution by glutamine (E350Q) or alanine (E350A).⁹⁴⁾ The two mutant $\alpha_2\beta_2$ complexes are virtually inactive in reactions with L-serine but are active in reactions with β -chloro-L-alanine. Several mutant forms of the β subunit (E109A, L188F, D305N, and F306A) display similar alterations in reaction specificity.^{95,102)} These mutations may alter the conformation of the active site in a way which decreases the ability of the enzymes to catalyze the β -elimination of the weak hydroxide leaving group of L-serine. Our finding that these mutant $\alpha_2\beta_2$ complexes, unlike the wild type $\alpha_2\beta_2$ complex, undergo irreversible substrate-induced inactivation by β -chloro-L-alanine has led us to propose that the mutant enzymes exist in an "open" form whereas the wild type enzyme exists in a "closed" form.^{95,102)} The effects of mutation on substrate specificity and on catalytic activity will be investigated further by rapid kinetic studies and by probes of conformational change, including tritium exchange experiments and X-ray crystallography.

7.4 The Indole Tunnel and Indole Channeling

Early studies provided evidence that indole is not liberated as a free intermediate in the $\alpha\beta$ reaction and suggested that indole was a channeled intermediate.¹⁷⁻²¹⁾ The crystallographic results now provide a structural basis for understanding the mechanism of channeling. The three-dimensional structure of the tryptophan synthase $\alpha_2\beta_2$ complex from *S. typhimurium* shows that a 25–30 Å-long hydrophobic tunnel connects the active sites of the α and β subunits.⁷⁾ About one-third of the tunnel is within the α subunit whereas about two thirds passes between the N-terminal and C-terminal domains of the β subunit. Although the dimensions of the tunnel are sufficient to accommodate four indoles between the α and β sites, there is no direct, structural evidence for the presence of one or more indoles in the tunnel. The cartoon in Fig. 7.10⁸⁹⁾ based on the computer graphics modeling studies⁷⁾ shows the active sites of an α/β pair with one indole bound at the α site and four indoles in the tunnel. The indole near pyridoxal phosphate in the β site is also illustrated in the close up view of the β site in Fig. 7.8. This tunnel probably facilitates the diffusion of indole from the α site, where it is produced by cleavage of indole-3-glycerol phosphate, to the β site, where it condenses with pyridoxal phosphate Schiff base of aminoacrylate (Fig. 7.10).

Recent kinetic studies provide additional support for this mechanism of indole channeling.^{86,87,89,103,104)} There is also evidence that indole can pass from solution through the α site and the tunnel to the β site.^{86,89,105)} This hypothesis is favored by rapid-scanning stopped-flow investigations that show that a ligand that binds to the α site, α -glycerol 3-phosphate, decreases the rate, but not the yield, of the reaction of indole at the β site. The decreased rate may result from a ligand-induced conformational change that restricts the access of indole to the α site or to the tunnel. Single amino acid substitutions in the α

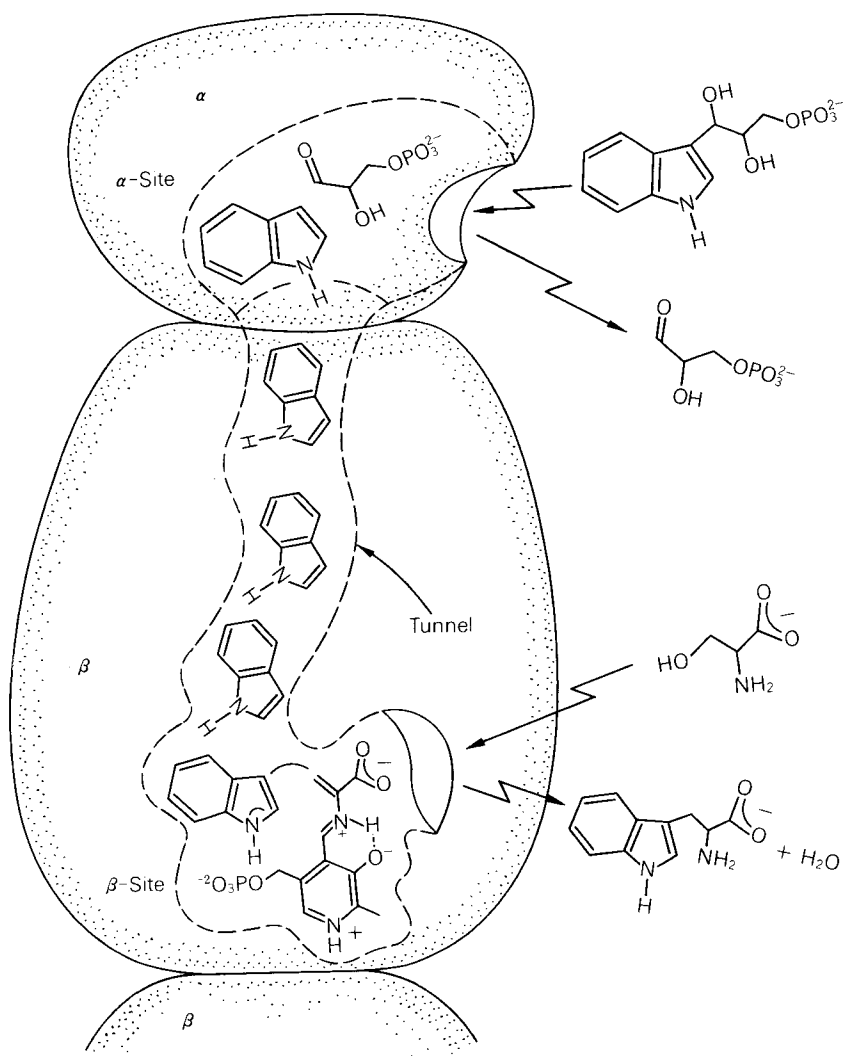


Fig. 7.10 Cartoon depicting the $\alpha\beta$ reaction. L-serine is converted to the Schiff base of amino acrylate at the β -site. Indole, which is produced by cleavage of indole-3-glycerol phosphate at the α -site, diffuses through the tunnel and reacts with the Schiff base of amino acrylate at the β -site to form L-tryptophan. (Reproduced with permission from Brzovic *et al.*, *J. Biol. Chem.*, **267**, 13028 (1992)).

subunit (E49F, G51L, and D60Y) strongly reduce the effect of α -glycerol 3-phosphate on the rate of reaction of indole at the β site.⁸⁹⁾ The results suggest that the mutant $\alpha_2\beta_2$ complexes are unable to undergo the ligand-induced conformational change that serves to obstruct the direct access of indole into the tunnel in the wild type enzyme.⁸⁹⁾

Investigations of the mechanism of the $\alpha\beta$ reaction by steady-state and transient kinetic methods have determined the rate constants for intermediate steps in the reaction.¹⁰⁴⁾ The transient kinetic results show that diffusion of indole and condensation of indole with the aminoacrylate intermediate (ES III in Fig. 7.6) are rapid steps that occur without a lag; reprotonation of the nascent tryptophan carbanion (ES IV in Fig. 7.6) is the rate-limiting step. The data rule out diffusion of free indole through the bulk solvent and support the channeling mechanism.

The mechanism of channeling has also been probed by rapid quench-flow kinetic studies with wild type and mutant forms of the $\alpha_2\beta_2$ complex.⁸⁷⁾ Experiments designed to look for the accumulation of indole in a single turnover during the net conversion of radiolabeled indole-3-glycerol phosphate to L-tryptophan detected an insignificant amount of indole in the reaction with the wild type enzyme. Analysis of the reaction kinetics shows that indole does not accumulate because the rate of diffusion of indole through the tunnel and the reaction of indole with the amino acrylate intermediate (ES III in Fig. 7.6) are much faster than the rate-limiting step that leads to L-tryptophan release. The results suggest that a large reduction in the rate of either of these rapid steps should lead to accumulation of indole. This suggestion is verified by the finding that substantial indole is accumulated by a mutant form of the $\alpha_2\beta_2$ complex (β E109D) which has a greatly reduced rate in the chemical step (ES III $\rightarrow$ ES IV in Fig. 7.6).⁸⁷⁾

Another important aspect of the channeling of indole in the $\alpha\beta$ reaction is the role of D-glyceraldehyde 3-phosphate. This product, which is formed by cleavage of indole-3-glycerol phosphate, is released from the α -site after the reaction of indole with ES III at the β -site.¹⁰⁴⁾ Recent studies of the pre-steady state kinetics of the $\alpha\beta$ reaction have lead to the hypothesis that the tightly bound D-glyceraldehyde 3-phosphate product helps to prevent indole, sequestered within the $\alpha_2\beta_2$ complex, from escaping into solution from the α -site.⁹⁰⁾ The physical picture which emerges is that D-glyceraldehyde 3-phosphate, like the analog α -glycerol 3-phosphate, alters the conformation of the α subunit from an open to a closed structure.⁹⁰⁾ The closed structure prevents access of indole from solution in the β reaction and escape of indole from the tunnel in the $\alpha\beta$ reaction.

7.5 Interactions Between the α and β Subunits

The contact surface between the α and β subunits is very broad ($\sim 1100 \text{ \AA}^2$) and is mostly hydrophobic in character.⁷⁾ The α subunit loops following strands 2,3,4, and 5 (Fig. 7.2) make important contacts with the β subunit. The 50-residue subdomain of the β subunit (residues 260–310) comprised of several long β hairpin loops makes several contacts with the α subunit (Fig. 7.5).⁷⁾ Contacts in this region probably prevent tryptic cleavage of the β subunit in the $\alpha_2\beta_2$ complex.^{34,35)} The fact that the sites of tryptic cleavage of the β subunit (Lys-272, R-275, and K-283) are close to the α/β contact surface may explain why proteolytically cleaved β subunit does not associate with the α subunit.^{59,106)} In contrast, cleavage of the β subunit by proteinase K at Glu-296, a residue at a position in the subdomain (residues 260–310) that does not make direct contact with the α subunit, yields an enzymatically active “nicked” β subunit that interacts weakly with α subunits.⁶¹⁾

Several single amino acid substitutions alter association between the α and β subunits. Two of the six conserved prolines in the α subunit (residues 57 and 132) are located close to or at the α/β contact surface.¹⁰⁷⁾ Substitution of residues 57 and 132 by glycine or alanine

alters interaction with the β subunit whereas substitution of the other four conserved prolines (residues 28, 62, 96, and 207) does not. Substitution of proline 132 by glycine (P132G) greatly decreases the subunit association constant and the association enthalpy.¹⁰⁷⁾

The β subunit of tryptophan synthase from *E. coli* is inactivated by a mutation at Gly-281 (G281R).¹⁰⁸⁾ Gly-281 is located at a sharp turn in the trypsin-sensitive, subdomain (residues 260–310) that makes several contacts with the α subunit.⁷⁾ The G281R mutation alters the catalytic properties of the isolated β subunit and weakens association with the α subunit. The mutation may interfere with hydrophobic interactions between the N-terminal and C-terminal domains of the β subunit and prevent a conformational change that affects catalytic properties and subunit interaction.¹⁰⁸⁾ Insertion of arginine or tryptophan between tyrosine 279 and phenylalanine 280 of the β subunit greatly weakens subunit interaction and decreases catalytic activity (X.-J. Yang and E.W. Miles, unpublished results).

7.6 Reciprocal Communication Between the α and β Subunits

One important property of tryptophan synthase that has not been clarified by the X-ray crystallographic studies is the mechanism of mutual activation of the α and β subunits upon formation of the $\alpha_2\beta_2$ complex.⁷⁾ This activation results from conformational changes that occur during assembly of the $\alpha_2\beta_2$ complex and upon ligand binding.^{85–87,89,90,105,109–113)} Events that occur at the active site of one subunit in the $\alpha_2\beta_2$ complex influence the properties of the other active site, apparently over the distance of 25–30 Å. For example, α subunit ligands (D-glyceraldehyde 3-phosphate, indole-3-propanol phosphate, and α -glycerol 3-phosphate) alter the kinetic and spectroscopic properties of the β subunit in the $\alpha_2\beta_2$ complex.^{76,77,85,86,90,105,114)} In a reciprocal way, the presence of bound L-serine at the β site in the $\alpha_2\beta_2$ complex increases the affinity of the α subunit for ligands.^{78,79,114)} The fact that the rate of cleavage of indole-3-glycerol phosphate is 20 times larger in the presence of L-serine ($\alpha\beta$ reaction) than in the absence of L-serine (α reaction) suggests that an intermediate formed from L-serine at the β site activates the α site over a distance of 25–30 Å. However, it is not possible to measure the α reaction alone in the presence of L-serine, because the nascent indole is immediately converted to tryptophan by the β subunit. This obstacle has been circumvented by the use of 6-nitroindole, which readily condenses with D-glyceraldehyde 3-phosphate at the α site but does not condense with L-serine at the β site.¹¹²⁾ The finding that addition of L-serine results in a 10-fold increase in the turnover number for 6-nitroindole-3-glycerol phosphate (in the forward α reaction) or for 6-nitroindole (in the reverse α reaction) proves that a derivative of L-serine at the β site activates the α reaction.¹¹²⁾

There are several types of evidence that the L-serine derivative that activates the α reaction is the Schiff base formed between aminoacrylate and pyridoxal phosphate (ES III in Fig. 7.6). (1) Amino acids including L- or D-tryptophan and glycine that form tetrahedral, Schiff base, or quinoidal intermediates do not appreciably affect the rate of the α reaction.^{90,113)} (2) L-serine and amino acids such as O-methyl-L-serine that form ES III do stimulate the α reaction.¹¹³⁾ (3) The rate of indole-3-glycerol phosphate turnover is roughly correlated with the rate of formation of ES III for each of the amino acids.^{90,113)} (4) The kinetics of the lag in cleavage of indole-3-glycerol phosphate and synthesis of L-tryptophan under single turnover conditions correspond to the rate of ES III formation

and to the rate of a protein conformational change that occurs upon L-serine binding.⁸⁷⁾ (5) The affinity of D-glyceraldehyde 3-phosphate and α -glycerol 3-phosphate at the α site is dependent on the nature of the chemical intermediate present at the β -site.⁹⁰⁾

The flexible α subunit loop (see Section 7.2.3 and Figs. 7.2 and 7.3) may play an important role in reciprocal communication between the α and β subunits. Ligands of the α subunit (indole-3-propanol phosphate and α -glycerol 3-phosphate) strongly inhibit the activity of the wild type $\alpha_2\beta_2$ complex in the β -reaction. Modification of the α subunit loop by substituting leucine for arginine at position 179 (R179L)¹¹⁴⁾ or by tryptic cleavage at Arg-188³⁶⁾ desensitizes the activity of the β subunit in the $\alpha_2\beta_2$ complex to inhibition by α subunit ligands. In a reciprocal way, L-serine bound at the β site alters the α site and the α subunit loop. For example, addition of L-serine increases the affinity of the wild type $\alpha_2\beta_2$ complex for α subunit ligands but does not have this effect on the R179L $\alpha_2\beta_2$ complex.¹¹⁴⁾ Although the presence of α -glycerol 3-phosphate alone decreases the rate of proteolytic cleavage of the $\alpha_2\beta_2$ complex about 6-fold,³⁶⁾ the presence of L-serine in combination with α -glycerol 3-phosphate completely protects the $\alpha_2\beta_2$ complex from this cleavage.⁴¹⁾ This result suggests that L-serine at the β site stabilizes the conformational change in the α subunit loop that is induced by α -glycerol 3-phosphate at the α site.⁴¹⁾

7.7 Catalytic Mechanism of the Physiological $\alpha\beta$ Reaction

Many aspects of the mechanism of the $\alpha\beta$ reaction have been mentioned previously in this chapter and in a recent review.⁶⁾ The results of steady state and rapid kinetic studies can be correlated with the crystallographic results using the simplified mechanisms shown in Fig. 7.1, 7.6 and 7.10. L-serine is converted to the aminoacrylate intermediate (ES III in Fig. 7.6) at the β site in the $\alpha_2\beta_2$ complex. Cleavage of indole-3-glycerol phosphate at the α site is followed by intramolecular diffusion of indole through the tunnel to the active site of the β subunit where it undergoes a pyridoxal phosphate-dependent β -replacement reaction to form L-tryptophan. The rates of most of the individual steps have been determined by transient kinetics.^{78,79,81,82,84,86,87,90,104,105,112,113)} The rate of cleavage of indole-3-glycerol phosphate is activated by formation of the ES III intermediate (Section 7.6). Indole diffuses rapidly through the channel and condenses rapidly with ES III. The rate-limiting step that leads to L-tryptophan release is protonation of the enzyme-Trp⁻ complex (ES IV in Fig. 7.6).¹⁰⁴⁾ The picture of the $\alpha\beta$ reaction that arises from these studies is that the tunnel and allosteric interactions serve to coordinate the reactions at the α and β sites and to prevent loss of the indole intermediate into solution.^{87,90,104,105,112)}

7.8 Unanswered Questions and Future Directions

There still remain many unanswered questions about channeling, site-site interactions, and reaction mechanisms. We hope that future crystallographic studies of wild type and mutant forms of the $\alpha_2\beta_2$ complex with substrates or substrate analogs bound to the β site will lead to the identification of residues in the substrate binding site of the β subunit. This structural information will allow us to select residues for site-directed mutagenesis and to study the properties of the mutant enzymes by a wide variety of techniques. Mutant enzymes that undergo partial reactions and bind reaction intermediates are especially useful for further crystallographic analysis. Preliminary results of the crystallographic analysis

of the K87T $\alpha_2\beta_2$ complex with bound L-tryptophan show the presence of the external aldimine of L-tryptophan with pyridoxal phosphate (Hyde, C. C., Parris, K. D., Miles, E. W., and Davies, D. R., work in progress). In this structure the indole moiety of L-tryptophan occupies the position of the indole near pyridoxal phosphate in the β site shown in Figs. 7.8 and 7.10. The location of glutamate 109 in the new structure is consistent with the predicted role for this residue in the activation of indole.^{87,88} The K87T $\alpha_2\beta_2$ complex with bound L-tryptophan also exhibits structural changes in the N-domain of the β subunit.

Another important problem is how subunit association controls the reaction rates and reaction specificity of the β subunit. Addition of the α subunit increases the rate of the β subunit in the β -replacement reaction with L-serine and indole ($X=OH$ in Fig. 7.6). Our finding that addition of the α subunit does not increase the rate of the corresponding β -replacement reaction with β -chloro-L-alanine and indole ($X=Cl$ in Fig. 7.6) is evidence that removal of the hydroxyl group of L-serine (see Fig. 7.6) is the step that is activated by the α subunit.^{42,95} Removal of the weak hydroxide leaving group is thought to require protonation by a catalytic residue, whereas removal of the strong chloride leaving group does not.¹¹⁵ Subunit association may stabilize a conformation of the β subunit in which a catalytic group is optimally oriented to protonate the hydroxyl leaving group. Addition of the α subunit inhibits the activity of the β subunit in the β -elimination reaction with L-serine (Fig. 7.6). We have recently suggested that the α subunit may stabilize a "closed" form of the β subunit that disfavors the solvent-catalyzed β -elimination reaction.⁹⁵

The nature of the conformational changes that occur upon subunit association may be revealed by comparing the three-dimensional structures of the $\alpha_2\beta_2$ complex and the separate α subunit or β subunit. We have recently obtained crystals of the β subunit from *S. typhimurium* that may serve this goal (S.A. Ahmed, X.-J. Yang, and E.W. Miles, unpublished results). Another key question is the nature of the conformational changes that occur when ligands bind to the $\alpha_2\beta_2$ complex. Recent kinetic studies mentioned in Section 7.4 suggest that ligands which bind to the α subunit induce a conformational change in the α subunit that serves to obstruct the direct access of indole from solution into the tunnel.^{89,90} Certain mutations in the α subunit (E49F, G51L, and D60Y) appear to block this conformational transition.⁸⁹ Such a conformational change in the wild type $\alpha_2\beta_2$ complex would serve an important function to trap indole that has already gained entrance into the tunnel. Crystallographic analyses of wild type and mutant $\alpha_2\beta_2$ complexes in the presence of ligands of α and β subunits are needed to test this hypothesis. On the basis of crystallographic studies of triose phosphate isomerase^{25,26,39} and of our studies using proteolysis described in Section 7.2.3, one might anticipate that the flexible loop in the α subunit would move toward the active site and close down over the bound α subunit ligand. Ligand binding may also change the orientation of some residues in the tunnel and in the α/β contact region. Phenylalanine 280, a β subunit residue in the wall of the tunnel, appears to enter the tunnel in some crystallographic structures (C.C. Hyde, unpublished observations). In addition, two residues in a disordered loop in the α subunit, residues 58 and 59) appear to enter the tunnel in the structure of the K87T $\alpha_2\beta_2$ complex as previously speculated.⁸⁹ These structural alterations upon ligand binding may represent an elegant mechanism of tunnel gating that is important in controlling the overall $\alpha\beta$ reaction.

REFERENCES

1. C. Yanofsky, I.P. Crawford, in: *The Enzymes* (P.D. Boyer, ed.), Academic Press, New York. (1972) p.1.
2. E.W. Miles, *Adv. Enzymol.*, **49**, 127 (1979).
3. E. W. Miles, in: *Pyridoxal Phosphate and Derivatives* (D. Dolphin, R. Poulson and O. Avramovic, eds.) John Wiley & Sons, New York. (1986) p. 253.
4. C. Yanofsky, *Bioessays*, **6**, 133 (1987).
5. C. Yanofsky, *Biochim. Biophys. Acta*, **1000**, 133 (1989).
6. E.W. Miles, *Adv. Enzymol. Relat. Areas Mol. Biol.*, **64**, 93 (1991).
7. C.C. Hyde, S.A. Ahmed, E.A. Padlan, E.W. Miles, D.R. Davies, *J. Biol. Chem.*, **263**, 17857 (1988).
8. C.C. Hyde, E.W. Miles, *Biotechnology*, **8**, 27 (1990).
9. B.P. Nichols, C. Yanofsky, *Proc. Natl. Acad. Sci. USA*, **76**, 5244 (1979).
10. I.P. Crawford, B.P. Nichols, C. Yanofsky, *J. Mol. Biol.*, **12**, 489 (1980).
11. I.P. Crawford, L. Eberly, *Biochimie*, **71**, 521 (1989).
12. A. Hadero, I.P. Crawford, *Mol. Biol. Evol.*, **3**, 191 (1986).
13. I.P. Crawford, T. Niermann, K. Kirschner, *Proteins*, **2**, 118 (1987).
14. D.M. Burns, C. Yanofsky, *J. Biol. Chem.*, **264**, 3840 (1989).
15. M.B. Berlyn, R.L. Last, G.R. Fink, *Proc. Natl. Acad. Sci. USA*, **86**, 4604 (1989).
16. R.L. Last, P.H. Bissinger, D.J. Mahoney, E.R. Radwanski, G.R. Fink, *Plant Cell*, **3**, 645 (1991).
17. C. Yanofsky, M. Rachmeler, *Biochim. Biophys. Acta*, **28**, 641 (1958).
18. I.P. Crawford, C. Yanofsky, *Proc. Natl. Acad. Sci. USA*, **44**, 1161 (1958).
19. C. Yanofsky, *Biochim. Biophys. Acta*, **31**, 408 (1959).
20. J.A. DeMoss, *Biochim. Biophys. Acta*, **62**, 279 (1962).
21. W.H. Matchett, *J. Biol. Chem.*, **249**, 4041 (1974).
22. G.K. Farber, G.A. Petsko, *Trends Biochem. Sci.*, **15**, 228 (1990).
23. D.K. Wilson, F.B. Rudolph, F.A. Quiocho, *Science*, **252**, 1278 (1991).
24. J.P. Priestle, M.G. Grutter, J.L. White, M.G. Vincent, M. Kania, E. Wilson, T.S. Jardetzky, K. Kirschner, J.N. Jansonius, *Proc. Natl. Acad. Sci. USA*, **84**, 5690 (1987).
25. E. Lolis, G.A. Petsko, *Biochemistry*, **29**, 6619 (1990).
26. E. Lolis, T. Alber, R.C. Davenport, D. Rose, F.C. Hartman, G.A. Petsko, *Biochemistry*, **29**, 6609 (1990).
27. M. Wilmanns, C.C. Hyde, D.R. Davies, K. Kirschner, J.N. Jansonius, *Biochemistry*, **30**, 9161 (1991).
28. M. Wilmanns, J.P. Priestle, T. Niermann, J.J. Jansonius, *J. Mol. Biol.*, **233**, 477 (1992).
29. C. Yanofsky, G.R. Drapeau, J.R. Guest, B.C. Carlton, *Proc. Natl. Acad. Sci. USA*, **57**, 296 (1967).
30. L. Shirvanee, V. Horn, C. Yanofsky, *J. Biol. Chem.*, **265**, 6624 (1990).
31. K. Yutani, K. Ogasahara, T. Tsujita, K. Kanemoto, M. Matsumoto, S. Tanaka, T. Miyashita, A. Matsushiro, Y. Sugino, E.W. Miles, *J. Biol. Chem.*, **262**, 13429 (1987).
32. E.W. Miles, P. McPhie, K. Yutani, *J. Biol. Chem.*, **263**, 8611 (1988).
33. S. Nagata, C.C. Hyde, E.W. Miles, *J. Biol. Chem.*, **264**, 6288 (1989).
34. E.W. Miles, W. Higgins, *J. Biol. Chem.*, **253**, 6266 (1978).
35. W. Higgins, T. Fairwell, E.W. Miles, *Biochemistry*, **18**, 4827 (1979).
36. E.W. Miles, *J. Biol. Chem.*, **266**, 10715 (1991).
37. E.W. Miles, in: *Conformations and Forces in Protein Folding* (B.T. Nall and K.A. Dill, eds.) AAAS, Waldorf, MD (1991) p. 115.
38. E.W. Miles, K. Yutani, K. Ogasahara, *Biochemistry*, **21**, 2586 (1982).
39. R.K. Wierenga, M.E. M. Noble, J.P.M. Postma, H. Groendijk, K.H. Kalk, W.G.J. Hol, F.R. Oppendoes, *Proteins*, **10**, 33 (1991).
40. D.L. Pompliano, A. Peyman, J.R. Knowles, *Biochemistry*, **29**, 3186 (1990).
41. S.B. Ruvinov, E.W. Miles, *FEBS Lett.*, **299**, 197 (1992).
42. X.-J. Yang, E.W. Miles, *J. Biol. Chem.*, **267**, 7520 (1992).
43. A.M. Beasty, M.R. Hurler, J.T. Manz, T. Stackhouse, J.J. Onuffer, C.R. Matthews, *Biochemistry*, **25**, 2965 (1986).
44. T.M. Stackhouse, J.J. Onuffer, C.R. Matthews, S.A. Ahmed, E.W. Miles, *Biochemistry*, **27**, 824 (1988).
45. B.A. Chrnyk, C.R. Matthews, *Biochemistry*, **29**, 2149 (1990).
46. N.B. Tweedy, M.R. Hurler, B.A. Chrnyk, C.R. Matthews, *Biochemistry*, **29**, 1539 (1990).
47. X. Chen, R. Rambo, C.R. Matthews, *Biochemistry*, **31**, 2219 (1992).
48. K. Yutani, K. Ogasahara, M. Suzuki, Y. Sugino, *J. Biochem. (Tokyo)*, **85**, 915 (1979).
49. K. Yutani, N.N. Khechinashvili, E.A. Lapshina, P.L. Privalov, Y. Sugino, *Int. J. Pept. Protein Res.*, **20**, 331 (1982).

50. K. Yutani, K. Ogasahara, K. Aoki, T. Kakuno, Y. Sugino, *J. Biol. Chem.*, **259**, 14076 (1984).
51. K. Yutani, K. Ogasahara, T. Tsujita, Y. Sugino, *Proc. Natl. Acad. Sci. USA*, **84**, 4441 (1987).
52. Y. Sugisaki, K. Ogasahara, E.W. Miles, K. Yutani, *Thermochimica Acta*, **163**, 117 (1990).
53. K. Yutani, S. Hayashi, Y. Sugisaki, K. Ogasahara, *Proteins*, **9**, 90 (1991).
54. C.R. Matthews, M.M. Crisanti, G.L. Gepner, G. Velicelebi, J.M. Sturtevant, *Biochemistry*, **19**, 1290 (1980).
55. C.R. Matthews, M.M. Crisanti, J.T. Manz, G.L. Gepner, *Biochemistry*, **22**, 1445 (1983).
56. M.R. Hurle, N.B. Tweedy, C.R. Matthews, *Biochemistry*, **25**, 6356 (1986).
57. H. Kanzaki, P. McPhie, E.W. Miles, *Arch. Biochem. Biophys.*, **284**, 174 (1991).
58. W.K. Lim, H.J. Shin, D.L. Milton, J.K. Hardman, *J. Bacteriol.*, **173**, 1886 (1991).
59. A. Hogberg-Raibaud, M.E. Goldberg, *Biochemistry*, **16**, 4014 (1977).
60. S.A. Ahmed, T. Fairwell, S. Dunn, K. Kirschner, E.W. Miles, *Biochemistry*, **25**, 3118 (1986).
61. M. Kaufmann, T. Schwarz, R. Jaenicke, K.D. Schnackerz, H.E. Meyer, P. Bartholmes, *Biochemistry*, **30**, 4173 (1991).
62. M.E. Goldberg, A. Hogberg-Raibaud, *J. Biol. Chem.*, **254**, 7752 (1979).
63. B. Friguet, O.L. Djavadi, M.E. Goldberg, *Eur. J. Biochem.*, **160**, 593 (1986).
64. A. Chaffotte, Y. Guillou, M. Delepierre, Hinz-H.-J., M.E. Goldberg, *Biochemistry*, **30**, 8067 (1991).
65. S. Blond, M.E. Goldberg, *J. Mol. Biol.*, **182**, 597 (1985).
66. S. Blond, M.E. Goldberg, *Proc. Natl. Acad. Sci. USA*, **84**, 1147 (1987).
67. B. Friguet, O.L. Djavadi, M.E. Goldberg, *Res. Immunol.*, **140**, 355 (1989).
68. A. Murry-Brelier, M.E. Goldberg, *Biochimie*, **71**, 533 (1989).
69. A. Murry-Brelier, M.E. Goldberg, *Proteins*, **6**, 395 (1989).
70. A. Murry-Brelier, M.E. Goldberg, *J. Biol. Chem.*, **265**, 7987 (1990).
71. S. Blond-Elguindi, M.E. Goldberg, *Biochemistry*, **29**, 4209 (1990).
72. M.E. Goldberg, G.V. Semisotnov, B. Friguet, K. Kuwajima, O.B. Ptitsyn, S. Sugai, *FEBS Lett.*, **263**, 51 (1990).
73. M.E. Goldberg, S. York, L. Stryer, *Biochemistry*, **7**, 3662 (1968).
74. S.S. York, *Biochemistry*, **11**, 2733 (1972).
75. E.W. Miles, M. Hatanaka, I.P. Crawford, *Biochemistry*, **7**, 2742 (1968).
76. E.W. Miles, in *Biochemical and Medical Aspects of Tryptophan Metabolism* (O. Hayaishi, Y. Ishimura and R. Kido, eds.) Elsevier/North-Holland Biomedical Press, Amsterdam (1980) p. 137.
77. A.N. Lane, K. Kirschner, *Eur. J. Biochem.*, **120**, 379 (1981).
78. A.N. Lane, K. Kirschner, *Eur. J. Biochem.*, **129**, 571 (1983).
79. A.N. Lane, K. Kirschner, *Eur. J. Biochem.*, **129**, 561 (1983).
80. M. Roy, E.W. Miles, R.S. Phillips, M.F. Dunn, *Biochemistry*, **27**, 8661 (1988).
81. W.J. Drewe, M.F. Dunn, *Biochemistry*, **24**, 3977 (1985).
82. W.J. Drewe, M.F. Dunn, *Biochemistry*, **25**, 2494 (1986).
83. K.F. Houben, W. Kadima, M. Roy, M.F. Dunn, *Biochemistry*, **28**, 4140 (1989).
84. W.J. Drewe, S.C. Koerber, M.F. Dunn, *Biochimie*, **71**, 509 (1989).
85. K.F. Houben, M.F. Dunn, *Biochemistry*, **29**, 2421 (1990).
86. M.F. Dunn, V. Aguilar, P. Brzovic, W.F.J. Drewe, K.F. Houben, C.A. Leja, M. Roy, *Biochemistry*, **29**, 8598 (1990).
87. K.S. Anderson, E.W. Miles, K.A. Johnson, *J. Biol. Chem.*, **266**, 8020 (1991).
88. P.S. Brzović, A.M. Kayastha, E.W. Miles, M.F. Dunn, *Biochemistry*, **31**, 1180 (1992).
89. P.S. Brzović, Y. Sawa, C.C. Hyde, E.W. Miles, M.F. Dunn, *J. Biol. Chem.*, **267**, 13028 (1992).
90. P. Brzović, K. Ngo, M.F. Dunn, *Biochemistry*, **31**, 3831 (1992).
91. A. Mozzarelli, A. Peracchi, G.L. Rossi, S.A. Ahmed, E.W. Miles, *J. Biol. Chem.*, **264**, 15774 (1989).
92. E.W. Miles, H. Kawasaki, S.A. Ahmed, H. Morita, H. Morita, S. Nagata, *J. Biol. Chem.*, **264**, 6280 (1989).
93. A.M. Kayastha, E.W. Miles, *FASEB J.*, **4**, A2118 (1990).
94. A.M. Kayastha, Y. Sawa, S. Nagata, E.W. Miles, *J. Biol. Chem.*, **266**, 7618 (1991).
95. S.A. Ahmed, S.B. Ruvinov, A.M. Kayastha, E.W. Miles, *J. Biol. Chem.*, **266**, 21540 (1991).
96. A.M. Kayastha, Y. Sawa, S. Nagata, H. Kanzaki, E.W. Miles, *Indian J. Biochem. Biophys.*, **28**, 352 (1991).
97. Z. Lu, S. Nagata, E.W. Miles, *J. Biol. Chem.*, **268**, 8727 (1992).
98. R.S. Phillips, E.W. Miles, L.A. Cohen, *Biochemistry*, **23**, 6228 (1984).
99. R.S. Phillips, E.W. Miles, L.A. Cohen, *J. Biol. Chem.*, **260**, 14665 (1985).
100. E.W. Miles, R.S. Phillips, H.J. Yeh, L.A. Cohen, *Biochemistry*, **25**, 4240 (1986).
- 101.* A.M. Kayastha, E.W. Miles, in: *Enzymes Dependent on Pyridoxal Phosphate and Other Carbonyl Compounds as Cofactors*, Pergamon Press, Elmsford, N.Y., (1991) p. 265.
- 102.* E.W. Miles, S.A. Ahmed, A.M. Kayastha, in: *Enzymes Dependent on Pyridoxal Phosphate and Other Carbonyl Compounds as Cofactors*, Pergamon Press, Elmsford, N.Y., (1991) p. 249.
103. M.F. Dunn, M. Roy, B. Robustell, V. Aguilar, in: *Biochemistry of Vitamin B₆*, (T. Korpela and P. Christen, eds.) Birkhauser Verlag AG, Basel (1987) p. 171.

104. A.N. Lane, K. Kirschner, *Biochemistry*, **30**, 479 (1991).
- 105.* M.F. Dunn, P. Brzovic', C. Leja, K. Houben, M. Roy, A. Aguilar, W.F. Drewe Jr., in: *Enzymes Dependent on Pyridoxal Phosphate and Other Carbonyl Compounds as Cofactors*, Pergamon Press, Elmsford, N.Y., (1991) p. 257.
106. A. Hogberg-Raibaud, M.E. Goldberg, *Proc. Natl. Acad. Sci. USA*, **74**, 442 (1977).
107. K. Ogasahara, K. Hiraga, W. Ito, E.W. Miles, K. Yutani, *J. Biol. Chem.*, **267**, 5222 (1992).
108. G.-P. Zhao, R.L. Somerville, *J. Biol. Chem.*, **267**, 526 (1992).
109. P. Wilhelm, I. Pilz, A.N. Lane, K. Kirschner, *Eur. J. Biochem.*, **129**, 51 (1982).
110. A.N. Lane, K. Kirschner, *Eur. J. Biochem.*, **129**, 675 (1983).
111. A.N. Lane, C.H. Paul, K. Kirschner, *EMBO J.*, **3**, 279 (1984).
112. K. Kirschner, A.N. Lane, A.W. M. Strasser, *Biochemistry*, **30**, 472 (1991).
- 113.* P. Brzovic', E.W. Miles, M.F. Dunn, in: *Enzymes Dependent on Pyridoxal Phosphate and Other Carbonyl Compounds as Cofactors*, Pergamon Press, Elmsford, N.Y., (1991) p. 277.
114. H. Kawasaki, R. Bauerle, G. Zon, S.A. Ahmed, E.W. Miles, *J. Biol. Chem.*, **262**, 10678 (1987).
115. E. Borcsok, R.H. Abeles, *Arch. Biochem. Biophys.*, **213**, 695 (1982).

* The editors for refs. 101, 102, 105 and 113 were omitted. They are "Fukui, T., Kagamiyama, H., Soda, K. and Wada, H. eds."

8

Alanine Racemase:
Structure and Function

Tohru Yoshimura and Kenji Soda

Institute for Chemical Research, Kyoto University Uji-City, Kyoto, 611, Japan

Alanine racemase (EC 5.1.1.1) is a bacterial enzyme that catalyzes racemization of L- and D-alanine, and requires pyridoxal 5'-phosphate (PLP) as a cofactor. The enzyme was first discovered and isolated in *Streptococcus faecalis* by Wood and Gunsalus.¹⁾ It has been purified from *Pseudomonas putida*,²⁾ *Bacillus subtilis*,³⁾ *Salmonella typhimurium*,^{4,5)} *Streptococcus faecalis*,⁶⁾ and *Bacillus stearothermophilus*,⁷⁾ and plays a very important role in bacterial growth by providing D-alanine, a constituent of the peptidoglycan layer of bacterial cell wall. As a result the enzyme is regarded as a target of antibiotics. In this chapter, we describe the physiological, genetic and evolutionary aspects, structure, reaction mechanism, and inhibitors of alanine racemase.

8.1 Physiological Function of Alanine Racemase

The peptidoglycan layer of the bacterial cell wall provides resistance to osmotic lysis in both Gram-negative and -positive bacteria. The basic unit of the peptidoglycan is a muropeptide, a disaccharide of N-acetyl-D-glucosamine and N-acetylmuramic acid in β -(1-4) linkage with tetrapeptide side chain. It contains D-amino acids (D-alanine, D-glutamate, D-glutamine, and meso-diaminopimelic acid), which presumably provide resistance against proteolytic degradation.^{8,9)} The parallel polysaccharide chains of the cell wall are cross-linked through the peptide side chains. The terminal D-alanine residue of the side chain of one polysaccharide chain is covalently bound to the peptide side chain of the next polysaccharide chain. In Gram-negative bacteria such as *Escherichia coli* and *Salmonella typhimurium*, peptidoglycan is assembled by a UDP-N-acetylmuramylpentapeptide, the pentapeptide of which consists of L-Ala-D-Glu-meso-diaminopimelate-D-Ala-D-Ala. This is synthesized from UDP-N-acetylglucosamine, phosphoenolpyruvate, and five amino acids. The last step of the synthesis of UDP-N-acetylmuramylpentapeptide is the attachment of the dipeptide, D-alanyl-D-alanine, to the tetrapeptide side chain of the sugar unit. N-Acetylmuramylpentapeptide is enzymatically transferred from UDP to a membrane-bound lipid intermediate, undecaprenyl phosphate. The other monosaccharide unit, N-acetylglucosamine, is then added to the membrane-bound N-acetylmuramylpentapeptide to yield the recurring disaccharide of the peptidoglycan backbone, N-acetylmuramylpentapeptide- β -(4 $\rightarrow$ 1)-N-acetylglucosamine, attached to the undecaprenyl phosphonate. The disaccharide unit is released from the undecaprenyl moiety

and transferred to the growing end of the peptidoglycan polymer. Then cross-linking between parallel peptidoglycan chains occurs. In Gram-negative bacteria, the ϵ -amino group of a *meso*-diaminopimelate residue is cross-linked directly to the penultimate D-Ala in a second pentapeptide strand, while a pentaglycine bridge connects the strands in Gram-positive bacteria such as *Staphylococcus aureus*. The new interstrand peptide bond (diaminopimelate-D-Ala or Gly₅-D-Ala) is formed concomitant with the break of intra-strand D-Ala-D-Ala bond. D-Alanine is thus a central molecule in the peptidoglycan assembly and cross-linking. Alanine racemase catalyzes the racemization of L- and D-alanine and supplies D-alanine for the peptidoglycan synthesis: it is indispensable for bacterial growth.

8.2 Genetic and Evolutional Aspects of Alanine Racemase

Walsh and his coworkers cloned the alanine racemase genes.¹⁰ Two distinct *Salmonella* DNA clones that complement the D-alanine requirement of *E. coli* strain that is temperature sensitive for growth in the absence of D-alanine have been isolated.¹⁰ Both contain distinct alanine racemase genes: one mapped at minute 37 on the chromosome is termed the *dadB*, and the other mapped at minute 91 is termed the *alr* gene.⁹ The two distinct alanine racemase genes have been also demonstrated in *E. coli*.¹¹ The *dadB* alanine racemase is inducibly formed and functions in the catabolism of L-alanine; D-alanine produced by racemization is deaminated to pyruvate by D-alanine dehydrogenase, whose gene (*dadA*) is located adjacent to *dadB* and repressed together with *dadB* and *dadR*.¹⁰ The other alanine racemase, *Alr* enzyme, is constitutively synthesized and functions in the anabolic assembly of peptidoglycan.⁹ Mutation in either one of the alanine racemase genes does not result in D-alanine auxotrophy in *S. typhimurium*, but mutants lacking both genes require exogenous D-alanine for growth.⁹

We compared the complete amino acid sequences of alanine racemase from four different sources (*B. stearothermophilus*;¹² *B. subtilis*¹¹; *S. typhimurium*, *DadB*; *S. typhimurium*, *Alr*) to discover the homologous regions in their primary structures (Fig. 8.1).¹² Despite the fact that *Bacillus* and *Salmonella* belong to Gram-positive and Gram-negative bacteria, respectively, and that the *dadB* and *alr* genes map at two distinctly different regions of the *S. typhimurium* chromosome, sequences of the four racemases display considerable homology: 74 residues match in the sequences of all four enzymes and 53 residues match in the sequences of three (about 33% homology on average for sequences of the four enzymes). The sequence homologies between two of the four racemases were calculated to be 54% (*B. stearothermophilus*: *B. subtilis*); 43% (*DadB*: *Alr*); 35% (*B. stearothermophilus*: *DadB*); 34% (*B. subtilis*: *DadB*); 31% (*B. stearothermophilus*: *Alr*); and 30% (*B. subtilis*: *Alr*). The high homology between *DadB* and *Alr* enzymes strongly suggests that they evolved from a common ancestor by gene duplication. An octapeptide containing the active-site lysyl residue that binds PLP (Lys^(PLP)-Ala-Gln-Ala-Tyr-Gly-His-Gly) occurs in all four enzymes. Low-specificity amino acid racemase of *P. putida* contains the same sequence, and also probably evolved from the common progenitor.¹²

Fig. 8.1 Linear alignment of the protein sequences of alanine racemases from *B. subtilis*, *B. stearothermophilus*, *S. typhimurium* *dadB*, and *S. typhimurium* *alr*. The sequences of four alanine racemases were aligned by introducing gaps (hyphens) to maximize identities. Common residues among the four (■) and three (●) enzymes are shown below. The active-site lysyl residue is indicated with an asterisk. The vertical arrow shows the position where the limited proteolysis occurs. (Reproduced with permission from Tanizawa *et al.*, *Biochemistry*, 27, 1311 (1988)).

8.3 Structure of Thermostable Alanine Racemase of *Bacillus stearothermophilus*

Among the four alanine racemases whose primary structures have been determined, the one from *B. stearothermophilus* is thermostable. It is stable upon heat treatment at 70°C for 80 min in 10 mM potassium phosphate buffer (pH 7.2). Thermostability facilitates the large-scale purification of the enzyme⁷⁾ and is an advantage for structural study including high-resolution X-ray analysis, which is now in progress.¹³⁾ We have studied the structure of the *B. stearothermophilus* alanine racemase.

8.3.1 Subunit Dissociation and Unfolding

Alanine racemase of *B. stearothermophilus* consists of two identical subunits, whereas both *DadB* and *alr* enzymes of *Salmonella typhimurium* and the *Streptococcus faecalis* enzyme occur in a form of monomer. Toyama *et al.*¹⁴⁾ examined whether the monomeric form of the *B. stearothermophilus* enzyme is catalytically active. They studied the guanidine HCl-induced subunit dissociation and unfolding of the enzyme by fluorescence and absorption spectroscopies, circular dichroism (CD) analysis, and gel filtration.¹⁴⁾ The overall process was found to be reversible: more than 75% of the original activity was recovered by decreasing the denaturant concentration.

The enzyme was unfolded by guanidine HCl treatment through two detectable phases: Phases 1 and 2 were observed by fluorescence spectroscopy derived from tryptophan residue.¹⁴⁾ The fluorescence maximum moved gradually from 336 to 331 nm as the

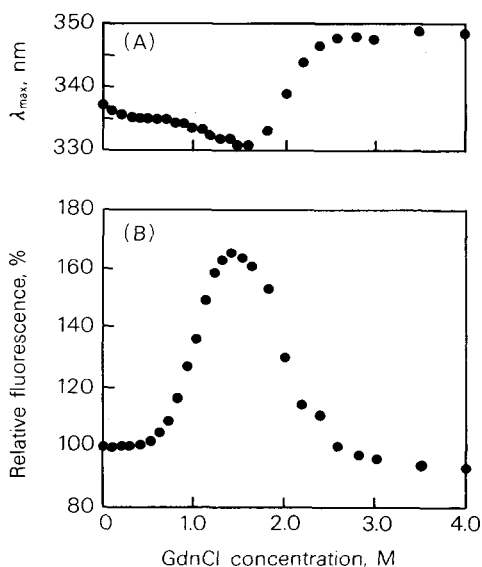


Fig. 8.2 Effect of GdnCl concentrations on fluorescence spectra of *B. stearothermophilus* alanine racemase. (A) The wavelength of the fluorescing maximum was plotted against guanidine HCl concentration. (B) The relative fluorescence intensity at 331 nm was plotted against guanidine HCl concentration.

guanidine HCl concentration was increased from 0.6 to 1.6 M (Fig. 8.2(A)). A transition midpoint was found around 1.0 M. The fluorescence maximum shifted sharply from 331 to 348 nm when guanidine HCl concentration was increased from 1.6 to 2.5 M with a transition midpoint of 2.0 M. The change in relative fluorescence intensity also showed the occurrence of the two transitions, with a midpoint at about 1.0 M for the first phase and at about 2.0 M for the second phase (Fig. 8.2(B)). Increase in the fluorescence intensity of tryptophan residues and the concomitant blue shift of the emission maximum are generally observed when the environment around tryptophan residues in proteins becomes more hydrophobic. The environment around tryptophan residues of alanine racemase thus becomes more hydrophobic in the first transition. The following red shift and the concomitant decrease in the fluorescence intensity in the second phase suggest the exposure of the tryptophan residues to solvent due to protein unfolding.¹⁴⁾

Gel filtration in the presence of guanidine HCl also showed the two-phase unfolding of alanine racemase. The apparent molecular weight obtained from gel filtration suggests that enzyme is dissociated into a monomeric form in the presence of 0.6 to 1.2 M guanidine HCl. This range of guanidine HCl concentration is very similar to that observed in the fluorescence change (the first phase). As the guanidine HCl concentration was increased above 2.4 M, the elution volume began to decrease. This is probably due to an increase in the Stocke's radius of the protein, unfolding at higher concentrations of guanidine HCl.

The denaturation of the enzyme with guanidine HCl was also followed by CD measurement. The mean residue ellipticity at 222 nm decreased with an increase in GdnCl concentrations.¹⁴⁾ The transition midpoints defined by the horizontal base line and the plateau were around 2.0 M. CD spectral change showed only one phase (similar to the second phase appearing in the fluorescence spectral change), suggesting that the structural change of alanine racemase during Phase 1 is small.

Based on the above findings, Toyama *et al.* concluded that the two phases are derived from the following transitions: Phase 1, dissociation of the dimer to monomers; Phase 2, unfolding of the monomer.¹⁴⁾ The monomeric form of the protein was found to be catalytically inactive when directly assayed in the presence of 1.5 M guanidine HCl.¹⁴⁾ The absorption spectrum of the monomeric form indicated that the carbonyl group of PLP is not bound in an internal Schiff base linkage. The reactivity of the carbonyl group of PLP is usually elevated by the formation of an internal Schiff base with the ϵ -amino group of a lysyl residue at the active site of PLP enzymes.¹⁵⁾ The mutant enzymes of aspartate aminotransferase¹⁶⁾ in which the PLP-binding lysyl residue is replaced by arginine do not react with PLP to form a Schiff base, but they show a slight, but distinct activity. Toyama *et al.* suggested that the inactivity of the monomeric form is due to a local and minor structural change which simultaneously occurs with the subunit dissociation.¹⁴⁾

The thermostable alanine racemase undergoes a major structural change when guanidine HCl concentrations are increased to 1.8 M. The unfolded monomer aggregates with each other to form a species of high molecular weight. Isolated monomers of β_2 -subunit of tryptophan synthase¹⁷⁾ and aspartate aminotransferase of *E. coli*¹⁸⁾ also aggregate.

The subunit dissociation is usually accompanied by destruction of the secondary structures when thermolabile proteins are used. The thermostable alanine racemase is very useful for studying the mechanism of subunit dissociation and protein unfolding.

8.3.2 Limited Proteolysis

Galakatos and Walsh showed that the native *DadB* and *Alr* racemases are digested at homologous positions by α -chymotrypsin, trypsin, and subtilisin to produce two non-overlapping polypeptides of *Mr*28,000, 11,000, and a tetrapeptide connecting two fragments¹⁹⁾ (*DadB*, residues 254–257; *Alr*, residues 256–259). Therefore, both enzymes are composed of two domains, which are linked by a tetrapeptide interdomain hinge. The two domains are associated with each other after the cleavage at the interdomain bridge under denaturing conditions. Both clipped enzymes (associated domains) retain about 3% of the original activity; the active-site geometry and secondary structure are not distorted by proteolysis. The hinge region is conserved also in the primary structure of alanine racemases from *Bacillus subtilis* and *B. stearothermophilus*.¹⁹⁾

Toyama *et al.* showed that the *B. stearothermophilus* enzyme also is cleaved into two major fragments with *Mr*14,000 (fragment F-1) and 29,000 (F-2) with subtilisin.²⁰⁾ They isolated F-1 and F-2 fragments by reversed phase HPLC, and showed that the enzyme is cleaved by proteolysis around ²⁶⁴Try to produce F-1 and F-2 based on the amino acid compositions of the isolated fragments. The cleaved position is compatible with those of *DadB* and *Alr* enzymes on the linear alignment of the protein sequences of these enzymes (Fig. 8.1). After incubation with subtilisin for 72 h, approximately 85% of the native protein disappeared and about 50% of the original activity remained.²⁰⁾ This indicates that about 15% of the native enzyme was unchanged, and that the activity corresponding to 35% of the original activity was derived from the peptide fragments formed. The clipped forms of both *DadB* and *Alr* enzymes were only 3% as active as the corresponding native forms.¹⁹⁾ Thus, the peptide fragments of the *B. stearothermophilus* enzyme interact with each other more strongly to form a considerably active structure than that of the thermolabile enzyme. Toyama *et al.* constructed and expressed a mutant gene which tandemly encodes the two peptides corresponding to F-2 and F-1.²¹⁾

8.3.3 Construction and Expression of Fragmentary Enzyme

To examine whether the two polypeptide fragments generated by limited proteolysis correspond to structural domains, Toyama *et al.* attempted to prepare the two polypeptides by means of genetic engineering.²¹⁾ They constructed a mutant alanine racemase gene of *B. stearothermophilus* which contains a new set of termination and initiation codons in the position of the gene corresponding to the putative hinge region of *DadB* and *Alr* racemases (Fig. 8.3).

The gene was found to be expressed to form an active alanine racemase composed of two dissimilar polypeptides.²¹⁾ The mutant enzyme named fragmentary alanine racemase was active in both reactions of the racemization of alanine. In the D- to L-alanine reaction, the K_m value for D-alanine was 1.8 mM, which was 2.2 times larger than that of the wild-type enzyme (0.82 mM), and the maximum velocity (V_{max}) was 930 units/mg, which was about half that of the wild-type enzyme (2,000 units/mg). In the reverse reaction (L- to D-alanine), the K_m value for L-alanine and the V_{max} value were 4.5 mM and 2,000 units/mg for the fragmentary enzyme, respectively, and 1.8 mM and 3,800 units/mg for the wild-type enzyme. Absorption and CD spectra of the fragmentary enzyme were also similar to those of the wild-type enzyme. The thermostability of the fragmentary enzyme was slightly lower than that of the wild-type enzyme. When heated at 80°C, the fragmentary enzyme

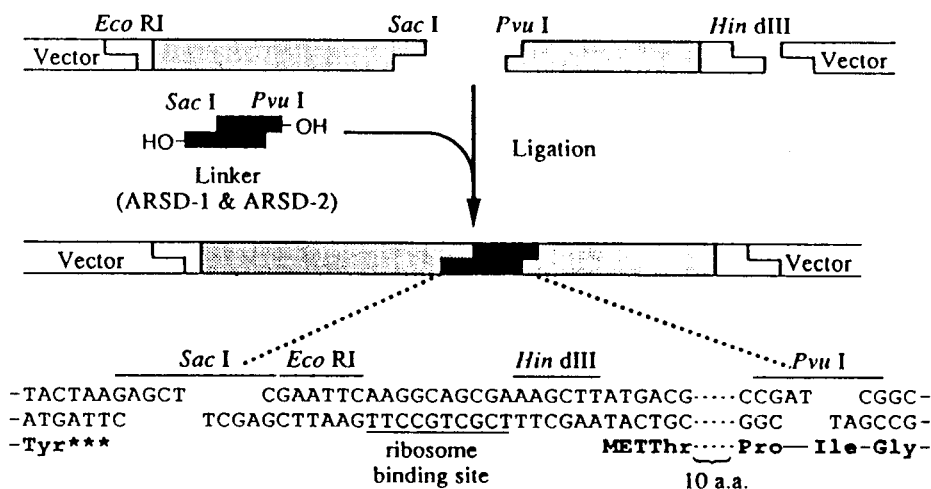


Fig. 8.3 Construction scheme for fragmentary alanine racemase. A termination codon (TAA), a ribosome-binding site (AAGGCAGCGA), and an initiation codon (ATG) were inserted at the sequence corresponding to Thr²⁶⁹–Ala–Gln²⁷¹ of alanine racemase. (Reproduced with permission from Toyama *et al.*, *J. Biol. Chem.*, **266**, 13636 (1991)).

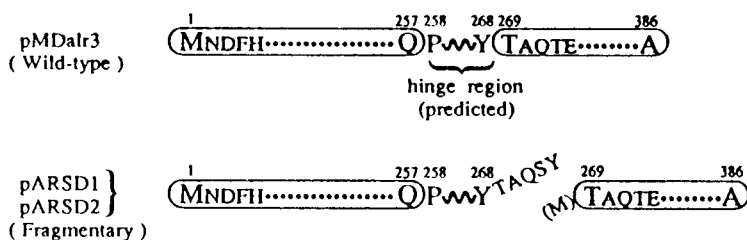


Fig. 8.4 Schematic representation of the wild-type and fragmentary alanine racemase. Only several of the relevant amino acid residues in the terminal portions are shown. The C-terminus of the N-terminal part of the fragmentary enzyme contains five extra amino acids residues. (Reproduced with permission from Toyama *et al.*, *J. Biol. Chem.*, **266**, 13637 (1991)).

was more rapidly inactivated than the wild-type enzyme. The time, $t_{1/2}$, when 50% of the initial activity was lost by incubation at 80°C, was 38 and 18 min for the wild-type and fragmentary enzymes, respectively. Since the fragmentary enzyme was considerably active, Toyama *et al.* suggested that its overall conformation is identical with that of the wild-type enzyme.²¹⁾ Conformational differences, if any, should be confined to a local region at or in the vicinity of the active site, judging from the slight difference in spectral properties of the cofactor bound to the active site. These findings led to the conclusion that the two polypeptide fragments correspond to structural folding units (domains) in the parental polypeptide chain of alanine racemase, as shown in Fig. 8.4.

Although the overall conformation of the fragmentary enzyme appears to be identical with that of the wild-type enzyme, it has only 40% of the activity of the native enzyme. The

decreased activity of the fragmentary enzyme was accounted kinetically to a combined consequence of both the increase in K_m values for substrates and the decrease in the V_{max} values as described above.²¹⁾

Galakatos and Walsh²²⁾ showed by site-directed mutagenesis studies that the hinge region in the *DadB* alanine racemase may play a role as a pivot for movement of the two domains during catalysis and may be located close to the active site, interacting with the bound PLP. Toyama *et al.* suggested that the decreased activity of the fragmentary enzyme relative to that of the native enzyme is due to either the presence of several extra residues in the hinge region or the discontinuity of the hinge sequence itself.²¹⁾ In accordance with this view, one of the *DadB* mutant enzymes in which a fourth Gly was inserted into the intraloop tetrapeptide, showed a 5-fold drop in V_{max}/K_m .²²⁾

The relative activity of the fragmentary enzyme to the wild-type enzyme was much higher than that of the clipped to the native *DadB* enzyme. This was explained by the assumption that the thermostable enzyme has more extensive hydrophobic interdomain interactions than the *DadB* enzyme with less thermostability.¹²⁾ The importance of hydrophobic interdomain interactions for catalysis was pointed out by studies on lactate dehydrogenase.^{23,24)}

Toyama *et al.* attempted to produce a single polypeptide corresponding to each domain separately in *E. coli*.²¹⁾ However, no protein reactive with the antibody against the wild-type alanine racemase was produced. They suggested that the fragments expressed in separate host cells are proteolytically broken down in the cells after translation.²¹⁾ The production of a similar N-terminal fragment from the *DadB* enzyme was also unsuccessful.²²⁾ The two peptide fragments probably fold into an active structure only when they are co-translated. To examine whether the folding process of each fragment into the domain structure is cooperative with each other during translation, Toyama *et al.* attempted to establish the *in vitro* conditions for reversible resolution of the fragmentary enzyme into two polypeptides and their reconstitution into the active enzyme.²⁵⁾

8.3.4 Unfolding and Reconstitution of Fragmentary Enzyme

Toyama *et al.* showed that the fragmentary alanine racemase of *B. stearothermophilus* is unfolded by treatment with guanidine HCl through two detectable phases in the same manner as the wild-type enzyme: Phase 1, dissociation into two $\alpha\beta$ -forms; Phase 2, dissociation into two fragments and unfolding of each fragment.²⁵⁾ However, they found no distinct intermediate processes between the subunit dissociation and the unfolding of the large and the small fragments.

They isolated the two polypeptides from the fragmentary alanine racemase in the presence of 4 M guanidine HCl, and attempted to reconstitute and reactivate the enzyme. The polypeptides were refolded spontaneously into an active form upon removal of guanidine HCl. The spectra obtained by summation of the CD spectra of the isolated subunits were closely similar to that of the native fragmentary enzyme.²⁵⁾ The lysine residue to which PLP is bound in the wild-type enzyme occurs in the large subunit of the fragmentary enzyme. However, neither bound PLP nor activity was found in the large subunit. Both fragments need to be folded together in order to form an active structure comparable with the native fragmentary enzyme.

8.4 Reaction Mechanism of Alanine Racemase

The alanine racemization catalyzed by alanine racemase is considered to be initiated by the transaldimination (Fig. 8.5).²⁶⁾ In this step, PLP bound to the active-site lysine residue forms the external Schiff base with a substrate alanine (Fig. 8.5, 1). The following α -proton abstraction produces the resonance-stabilized carbanion intermediates (Fig. 8.5, 2). If the reprotonation occurs on the opposite face of the substrate-PLP complex on which the proton-abstraction proceeds, the antipodal aldimine is formed (Fig. 8.5, 3). The subsequent hydrolysis of the aldimine complex gives the isomerized alanine and PLP-form racemase. The random return of hydrogen to the carbanion intermediate is the distinguishing feature that differentiates racemization from reactions catalyzed by other pyridoxal enzymes such as transaminases. Transaminases catalyze the transfer of amino group between amino acid and keto acid, and the reaction is initiated by the transaldimination, followed by the α -proton abstraction from the substrate-PLP aldimine to form a resonance-stabilized carbanion. This step is common to racemases and transaminases. However, in the transamination the abstracted proton is then transferred to C4' carbon of PLP in a highly stereospecific manner: The re-protonation occurs on the same face of the PLP-substrate aldimine on which the α -proton is abstracted. With only a few exceptions,^{27,28)} each step of pyridoxal enzymes-catalyzed reaction proceeds on only one side of the planar PLP-substrate complex. However, in the amino acid racemase

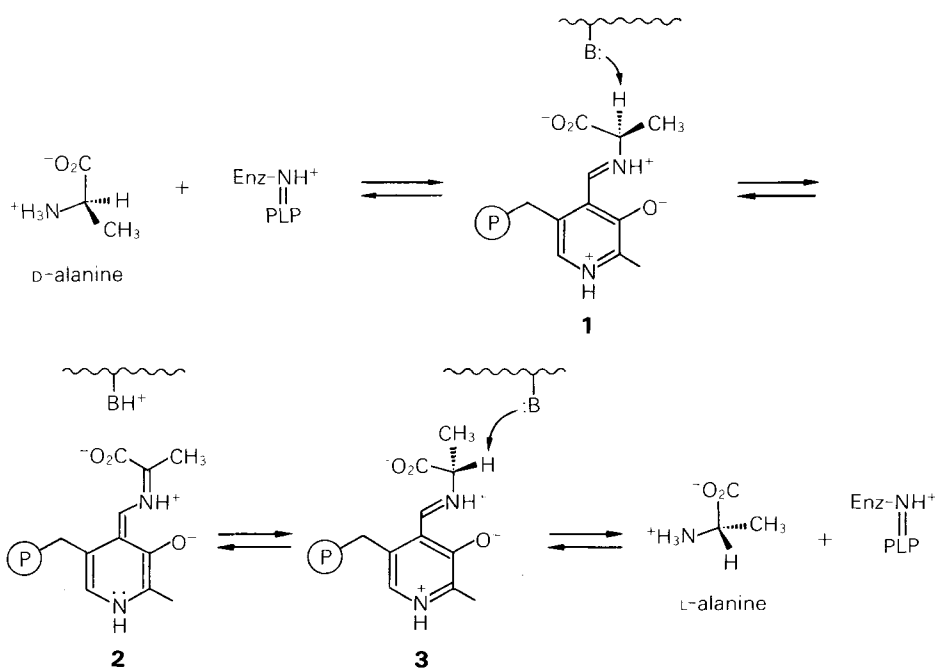


Fig. 8.5 Reaction mechanism of alanine racemase. (Reproduced with permission from Faraci and Walsh, *Biochemistry*, 27, 3268 (1988)).

reactions, each step occurs on both sides of the complex as described above. If the generally accepted hypothesis on pyridoxal enzyme reactions proposed by Dunathan²⁹⁾ is applicable to the alanine racemase, the C α -H σ bond to be broken should be positioned perpendicularly to the plane of the complex to achieve maximal orbital overlap with the π electron system of the complex, resulting in a substantial rate enhancement for the cleavage of that bond.

The racemization is apparently very simple, but the detailed mechanism remains unsolved. The reaction proceeds via either a one-base or a two-base mechanism.^{30,31)} In the one-base mechanism, an amino acid residue of the enzyme abstracts the substrate α -proton of the external Schiff base to form an anionic intermediate. The racemization results from the sterically random return of hydrogen to the α -carbon of the intermediate. In the two-base mechanism, two enantiomer-specific bases juxtaposed on either side of the chiral carbon exist in the active center. One base abstracts the α -proton from the external Schiff base, and the conjugated acid of the second base catalyzes protonation to the anionic intermediate from the other side. These roles are reversed for the racemization of the antipodal substrate.

The one-base mechanism is characterized by the retention of the substrate-derived proton in the product (internal return).³⁰⁾ With this criterion, reactions catalyzed by α -amino- ϵ -caprolactam racemase,³²⁾ amino acid racemase of broad specificity from *Pseudomonas striata*³³⁾ have been considered to proceed through the one-base mechanism. However, such internal returns were not observed in the reactions of alanine racemases from *E. coli* B,³³⁾ *B. stearothermophilus*,²⁶⁾ and *S. typhimurium* (*DadB* and *Alr*).²⁶⁾ The internal return should not be observed in the two-base mechanism, because the base catalyzing the protonation to the intermediate probably obtains the proton from the solvent. But the failure of the observation of the internal return can be also explained by the single-base mechanism in which exchange of the proton abstracted from the substrate α -carbon with the solvent is much faster than its transfer to the α -carbanion. Therefore, lack of the internal return does not directly indicate the two-base mechanism of the alanine racemase reaction.

Adams³⁴⁾ discussed the two-base mechanism based on the asymmetry of hydrogen exchange relative to racemization: more rapid exchange of α -H of D-alanine than that of L-alanine was found in the *B. subtilis* alanine racemase reaction in ²H₂O,³⁵⁾ and the rate of exchange relative to that of the racemization ($K_{\text{exchange}}/K_{\text{racemization}}$) was determined to be about 1 for the L- to D- reaction but about 2.8 for the D- to L- reaction. Such an asymmetry in the kinetics reflects the difference in the active site environment between L- and D-alanyl-PLP aldimines; this was also observed in the reaction by alanine racemase from *E. coli* W³⁶⁾ and in the β -haloalanine-caused inactivation of *DadB* and *Alr* enzymes of *S. typhimurium*.^{4,37)}

Faraci and Walsh²⁶⁾ studied the substrate and solvent deuterium isotope effects of the reactions catalyzed by alanine racemases of *S. typhimurium* (*DadB* and *Alr* enzymes) and *B. stearothermophilus*. Although the kinetic constants for all three alanine racemases obey the Haldane equation, i.e., $K_{\text{eq}} = 1$ (this confirms that the enzymes are "true" racemases), the individual Michaelis-Menten parameters in both directions show marked difference in the binding of each isomer. This suggests a structural asymmetry at the active sites of these enzymes. The asymmetry in the recognition and turnover of substrate enantiomer was also clearly seen in the results of isotope effect experiment with *DadB* enzyme. In the D-

to L-reaction catalyzed by *DadB* enzyme, a strong isotope effect accompanying abstraction of the α -deuteron [$(V/K)_{\alpha\text{-H}}/(V/K)_{\alpha\text{-D}} = 2.9$] was observed; however, no isotope effect was observed in the protonation step [$(V/K)_{\text{H}_2\text{O}}/(V/K)_{\text{D}_2\text{O}}$]. In the opposite direction, abstraction of the α -deuteron led to virtually no isotope effect (1.3), while protonation occurs with a larger isotope effect (2.4). Thus the abstraction of the α -deuteron of the

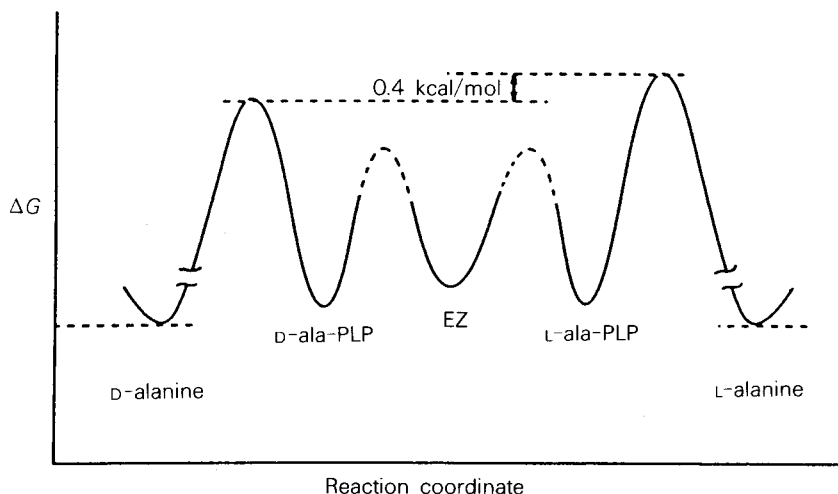


Fig. 8.6 Partial reaction profile of alanine racemization by *B. stearothermophilus*. Dotted lines indicate kinetically insignificant steps. (Reproduced with permission from Faraci and Walsh, *Biochemistry*, 27, 3264 (1988)).

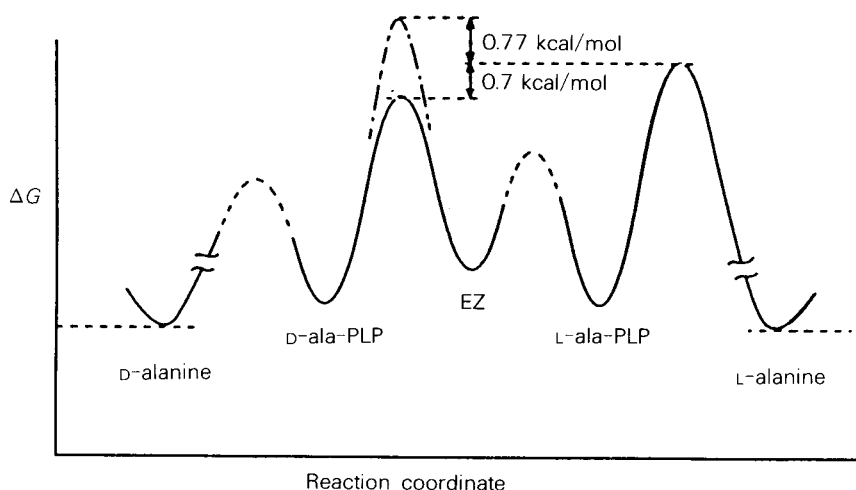


Fig. 8.7 Partial reaction profile of alanine racemization by *DadB* enzyme. Dotted lines indicate kinetically insignificant steps, and the dot-dash line indicates kinetic tritium isotope effect. (Reproduced with permission from Faraci and Walsh, *Biochemistry*, 27, 3275 (1988)).

D-alanyl-PLP aldimine and the protonation of carbanion intermediate to give the D-alanyl-PLP aldimine are more difficult than those with L-isomer. Similar results were obtained with *Alr*. However, little isotope effects were found in the reaction of *B. stearothermophilus* alanine racemase, which suggests that the rate-limiting step of the enzyme is not the proton abstraction but the transaldimination.²⁶⁾ The reaction profiles of alanine racemization catalyzed by *DadB* enzyme and *B. stearothermophilus* enzyme proposed by Faraci and Walsh²⁶⁾ are shown in Fig. 8.6 and 8.7.

An alternative single-base model named the “swinging door” mechanism proposed by Henderson and Johnston³⁸⁾ can also explain the kinetic asymmetries of alanine racemase reactions (Fig. 8.8). In this mechanism, a single base abstracts a proton on one face, then adds it back either on the same face or the opposite face. During the catalysis, either the base moves relative to the substrate or the substrate moves relative to the base. The swinging door mechanism is based on the kinetic asymmetries observed in the inactivation of *B. subtilis* alanine racemase by L- and D- β -chloroalanine³⁸⁾: the D-isomer inhibited the enzyme at a very low concentration ($K_i=0.005$ mM) in a competitive manner, whereas L- β -chloroalanine inhibited at much higher concentrations ($K_i=1.71$ mM) noncompetitively. The position of the “door” relative to the point of deprotonation-reprotonation determines whether a D- or L-substrate interacts with the enzyme. Since the active site of alanine racemase reacts asymmetrically with the enantiomers of the substrates and has a conformation that favors the D-enantiomer, one position (the upper position in Fig. 8.8) is probably more stable than the other. A swinging door mechanism requires the pivot rotation of the substrate PLP Schiff base, and such cofactor motions have been well

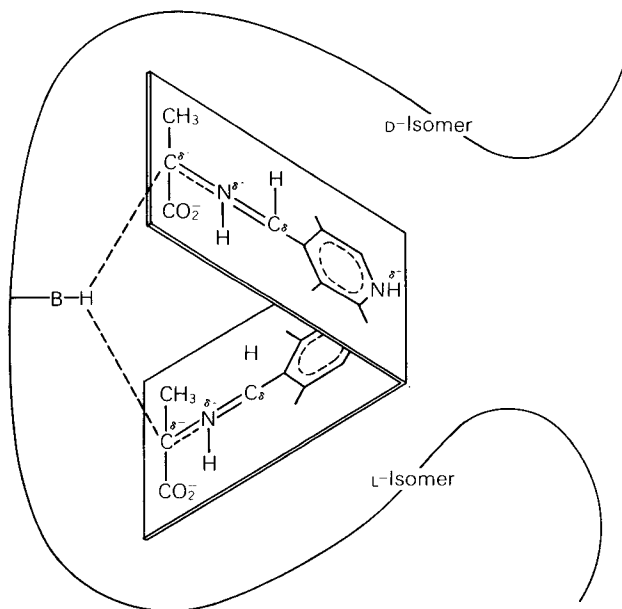


Fig. 8.8 Swinging-door model for mechanism of alanine racemase reaction. (Reproduced with permission from Henderson and Johnston, *Biochem. Biophys. Res. Commun.*, **68**, 797 (1976)).

documented in other PLP enzymes.³⁹⁾ This mechanism can explain the kinetic asymmetry observed in alanine racemase reactions and also provides a simple model for the exchange of the α -hydrogen with the solvent.

Recently, Faraci and Walsh⁴⁰⁾ studied the mechanism of inactivation of *B. stearothermophilus* alanine racemase with β,β,β -trifluoroalanine. When the enzyme was modified with ^3H -labeled trifluoroalanine, incorporation of the label was found only at lysine-38 which binds to PLP. The PLP-binding lysine abstracts the α -proton as a base from the substrate in the catalysis of aspartate aminotransferase.⁴¹⁾ It is thus possible that lysine-38 plays a catalytic base in the racemase reaction. If this is the case, the specific incorporation of the ^3H at lysine-38 sustains the single base mechanism. Lysine-38 is considered to attack the 4'-carbon of PLP and the β -carbon of the trifluoroalanine as well as to abstract the α -proton during the reaction. Therefore, Faraci and Walsh⁴⁰⁾ concluded that the enzyme base and PLP cofactor are flexible during the reaction, which is similar to that depicted in the swinging door mechanism.

There is at present no conclusive evidence for either the two-base mechanism or the single-base mechanism. This will probably be shown by a combination of X-ray crystallography and site-directed mutagenesis study, which showed the mechanism of mandelate racemase reactions.⁴²⁻⁴⁴⁾

8.5 Inhibitors of Alanine Racemase

As described in the first section of this chapter, D-alanine is an essential constituent of the peptidoglycan layer of bacterial cell wall, but it is not a common metabolite in mammalian cells. Therefore, alanine racemase has been recognized as a suitable target for antibiotics, and a variety of natural and synthetic inhibitors (inactivators) of the enzyme have been reported.

β -Substituted alanines are an important group of inhibitors. Manning *et al.*⁴⁵⁾ found that the D-isomer of β -chloro-alanine was bacteriocidal, and alanine racemase activity was lost by treatment of the cells with the compound. Wang and Walsh⁴⁶⁾ have shown that both enantiomers of β -fluoroalanine and β -chloro-alanine, and *O*-carbamoyl- and *O*-acetyl-D-serine inactivate alanine racemase from *E. coli* B. Both *DadB* and *Alr* enzymes of *S. typhimurium*, and the *B. stearothermophilus* alanine racemase also suffered from the irreversible time-dependent inactivation with a few β -substituted alanines.^{4,37,47)} The mechanism of inactivation of alanine racemases is probably similar to that of aspartate transaminase,⁴⁸⁾ glutamate decarboxylase⁴⁹⁾ and the low substrate specificity amino acid racemase from *Pseudomonas striata*⁵⁰⁾ by β -substituted alanines. The racemase recognizes wrongly the β -substituted alanines as a substrate and catalyzes abstraction of the α -proton followed by the elimination of β -substituents such as chloride and fluoride. The resulting α -aminoacrylate-PLP Schiff base intermediate undergoes transaldimination with an ϵ -amino group of the active-site lysine to produce α -aminoacrylate and the PLP form enzyme (Fig. 8.9). There are two possible routes to change the released α -aminoacrylate. This can undergo enamine-imine tautomerization to be hydrolyzed to pyruvate and ammonia (Fig. 8.9, route A). This is harmless to the enzyme, but once per about 800 catalytic turnovers (in the inactivation of *DadB* enzyme), α -aminoacrylate attacks and modifies the C4' carbon of PLP to bring out the inactivation of enzyme (Fig. 8.9, route B). Peptide mapping of the modified *DadB* (with ^{14}C -labeled β -chloro-D, L-alanine) and *Alr*

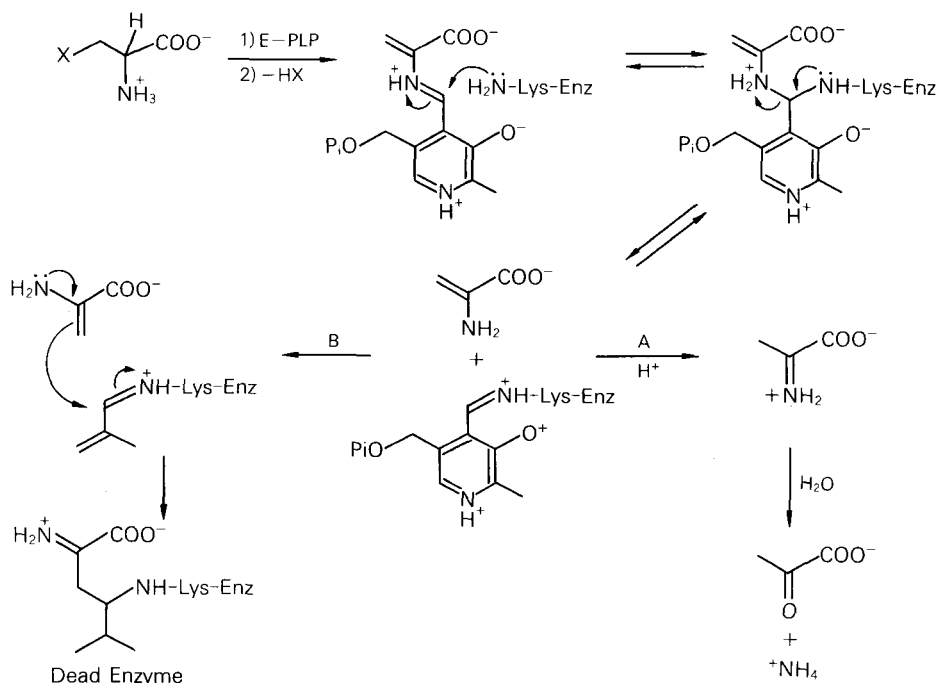


Fig. 8.9 Postulated mechanism of inactivation of alanine racemase by β -chloro-alanine ($\text{X}=\text{Cl}$). (Reproduced with permissions from Roise *et al.*, *Biochemistry*, 23, 200 (1984)).

(with ^{14}C - β -chloro-D-alanine) enzymes showed that the inactivators modified the active-site lysyl residue which binds to PLP in native enzymes.^{4,37)}

The β -substituted alanines seem to serve as a substrate and are converted into the active form by enzymatic catalysis. This type of inactivator is known as mechanism-based inactivators or suicide substrates.⁵¹⁾ The partition ratio is an index of inactivation efficiency. In an attempt to vary the partition ratio, β,β -difluoroalanine and β,β,β -trifluoroalanine isomers were systematically evaluated for substrate processing by alanine racemase.^{40,52)} In contrast to the very low efficiency of D- and L-difluoroalanines in inactivation of the alanine racemase from *E.coli* (partition ratios of 2600–5000), D,L-trifluoroalanine turned over less than 10 times per mole of inactivated enzyme, serving as a very efficient suicide substrate. The difference in inactivation efficiencies between di- and tri-halogenated alanines and the monohalogenated alanine was studied in terms of the stability of intermediates formed in the active site of enzyme and the anticipated oxidation state at the β -carbon of substrate fragment adducted to the enzyme.⁵²⁾ Upon inactivation of *DadB*, *Alr* and *B. stearothermophilus* alanine racemases with β,β,β -trifluoroalanine, two equivalents of fluoride ions are released from one mole of *B. stearothermophilus* racemase, and 5–10 equiv. of fluoride ions are released from each mole of *DadB* and *Alr* enzymes. This suggests that little (*DadB* and *Alr*) or no (*B. stearothermophilus* racemase) turnover occurs prior to the inhibition.⁴⁰⁾ Loss of two equivalents of fluoride ions per mole of the inactivated *B. stearothermophilus* alanine racemase and retention of the label corresponding

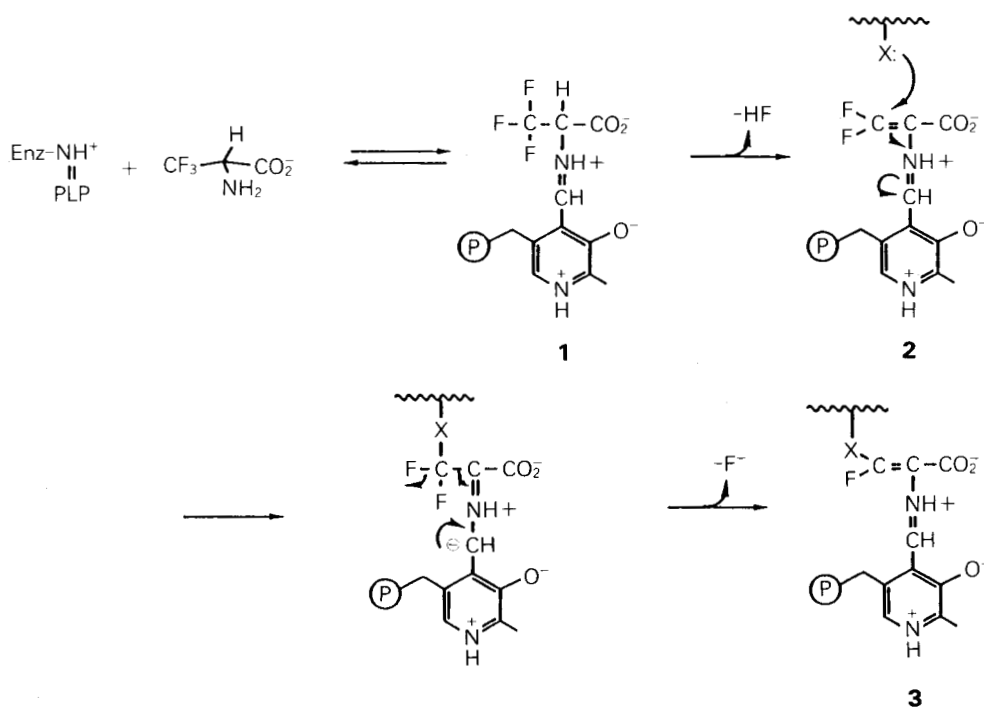
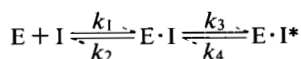


Fig. 8.10 Postulated mechanism of inactivation of alanine racemase by β,β,β -trifluoroalanine. (Reproduced with permissions from Faraci and Walsh, *Biochemistry*, **28**, 434 (1989)).

to one [$1\text{-}^{14}\text{C}$] trifluoroalanine molecule in the protein suggests the presence of a monofluoro enzyme adduct. The proposed reaction mechanism of *B. stearothermophilus* alanine racemase with trifluoroalanine is shown in Fig. 8.10.⁴⁰⁾ Incubation of the inhibitor with the enzyme leads to initial transaldimination from which HF is subsequently eliminated to yield the β -difluoro- α,β -unsaturated imine complex (Fig. 8.10, **2**). This species is a potent electrophile, in which capture of an enzyme nucleophile (PLP-binding lysyl residue, lysine-38) occurs very rapidly.⁴⁰⁾

One of the second type of inhibitor analogs which cause a time-dependent inactivation of alanine racemase is (1-aminoethyl)phosphonic acid, the phosphonate analog of alanine (Ala-P). Ala-P effectively and specifically inactivates alanine racemases from Gram-positive bacteria (*Bacillus*, *Staphylococcus*, *Streptococcus*), and serves as a reversible inhibitor of Gram-negative bacterial (*Escherichia*, *Salmonella*) alanine racemase.^{47,53)} The mechanism of inhibition was studied with *B. stearothermophilus* alanine racemase.⁴⁷⁾ The D- or L-Ala-P leads to an E·I complex with a K_i value of 1 mM, then is slowly isomerized ($K_{\text{inact}} = 6\text{--}9\text{ min}^{-1}$) to a stoichiometric enzyme complex (E·I*). The E·I* dissociates extremely slowly with the



ratio of $k_3/k_4=10^6$, indicating that the inhibited enzyme complex is distributed essentially in the $E \cdot I^*$ form. The estimated recovery time for half the enzyme returned from $E \cdot I^*$ is 25 days. The solid-state ^{15}N -NMR spectra of $[^{15}\text{N}]\text{Ala-P-enzyme}$ complex demonstrated the presence of protonated Ala-P-PLP aldimine,⁵⁴⁾ suggesting that the transaldimination occurs between Ala-P and enzyme.

REFERENCES

1. W. A. Wood, I. C. Gunsalus, *J. Biol. Chem.*, **190**, 403 (1951).
2. G. Rosso, K. Takashima, E. Adams, *Biochem. Biophys. Res. Commun.*, **34**, 134 (1969).
3. K. Yonaha, T. Yorifuji, T. Yamamoto, K. Soda, *J. Ferment. Technol.*, **53**, (1975).
4. N. Esaki, C. T. Walsh, *Biochemistry*, **25**, 3261 (1986).
5. S. A. Wasserman, E. Daub, P. Grisafi, D. Botstein, C. T. Walsh, *Biochemistry*, **25**, 5182 (1984).
6. B. Badet, C. T. Walsh, *Biochemistry*, **24**, 1333 (1985).
7. K. Inagaki, K. Tanizawa, B. Badet, C. T. Walsh, H. Tanaka, K. Soda, *Biochemistry*, **25**, 3268 (1986).
8. J. Strominger, E. Ito, R. H. Threnu, *J. Am. Chem. Soc.*, **82**, 998 (1960).
9. C. T. Walsh, *J. Biol. Chem.*, **264**, 2393, (1989).
10. S. A. Wasserman, C. T. Walsh, D. Botstein, *J. Bacteriol.*, **153**, 1439 (1983).
11. J. Wild, M. Lobočka, W. Walczak, T. Kłopotowski, *Mol. Gen. Genet.*, **198**, 315 (1985).
12. K. Tanizawa, A. Ohshima, A. Scheidegger, K. Inagaki, H. Tanaka, K. Soda, *Biochemistry*, **27**, 1311 (1988).
13. D. J. Neidhart, M. D. Distefano, K. Tanizawa, K. Soda, C. T. Walsh, G. A. Petsko *J. Biol. Chem.*, **262**, 15323 (1987).
14. H. Toyama, N. Esaki, T. Yoshimura, K. Tanizawa, K. Soda, *J. Biochem.*, **110**, 279 (1991).
15. E. E. Snell, S. J. Di Mari, *Enzymes*, 3rd Ed. **2**, 335 (1970).
16. S. Kuramitsu, Y. Inoue, S. Tanase, Y. Morino, H. Kagamiyama, *Biochem. Biophys. Res. Commun.*, **146**, 416 (1987).
17. C. Zetina, M. E. Goldberg, *J. Mol. Biol.*, **137**, 401 (1980).
18. M. Herold, K. Kirschner, *Biochemistry*, **29**, 1907 (1990).
19. N. G. Galakatos, C. T. Walsh, *Biochemistry*, **26**, 8475 (1987).
20. H. Toyama, K. Tanizawa, M. Wakayama, Q. Lee, T. Yoshimura, N. Esaki, K. Soda, *Agric. Biol. Chem.*, **55**, 2881 (1991).
21. H. Toyama, K. Tanizawa, T. Yoshimura, S. Asano, H. -H. Lim, N. Esaki, K. Soda, *J. Biol. Chem.*, **266**, 13634 (1991).
22. N. G. Galakatos, C. T. Walsh, *Biochemistry*, **28**, 8167 (1989).
23. R. Girg, R. Jaenicke, R. Rudolph, *Biochem. Int.*, **7**, 443 (1983).
24. U. Opitz, R. Rudolph, R. Jaenicke, L. Ericsson, H. Neurath, *Biochemistry*, **26**, 1399 (1987).
25. H. Toyama, *Doctoral Thesis, Kyoto University* (1992).
26. W. S. Faraci, C. T. Walsh, *Biochemistry*, **27**, 3267 (1988).
27. Y. Asada, K. Tanizawa, S. Sawada, T. Suzuki, H. Misono, K. Soda, *Biochemistry*, **20**, 6881 (1981).
28. C. C. Chang, A. Laghai, M. H. O'Leary, H. G. Floss, *J. Biol. Chem.*, **257**, 3564 (1982).
29. H. C. Dunathan, *Proc. Natl. Acad. Sci. USA*, **55**, 713 (1966).
30. G. J. Cardinale, R. H. Abeles, *Biochemistry*, **7**, 3970 (1968).
31. K. Soda, H. Tanaka, K. Tanizawa, in: *Vitamin B₆ Pyridoxal Phosphate Part B*, John Wiley & Sons, New York, (1986) p. 223.
32. S. A. Ahmed, N. Esaki, H. Tanaka, K. Soda, *Biochemistry*, **25**, 385 (1986).
33. S. -J. Shen, H. G. Floss, H. Kumagai, H. Yamada, N. Esaki, K. Soda, S. A. Wasserman, C. T. Walsh, *J. Chem. Soc., Chem. Commun.*, **82**, (1983).
34. E. Adams, in: *Advances in Enzymology*, Vol. **44**, John Wiley & Sons, New York, (1976) p. 69.
35. U. N. Babu, R. B. Johnston, L. C. McNeff, *Anal. Biochem.*, **63**, 208 (1975).
36. M. P. Lambert, F. C. Neuhaus, *J. Bacteriol.*, **110**, 978 (1972).
37. B. Badet, D. Roise, C. T. Walsh, *Biochemistry*, **23**, 5188 (1984).
38. L. L. Henderson, R. B. Johnston, *Biochem. Biophys. Res. Commun.*, **68**, 793 (1976).
39. J. C. Vederas, H. G. Floss, *Acc. Chem. Res.*, **3**, 455 (1980).
40. W. S. Faraci, C. T. Walsh, *Biochemistry*, **28**, 431 (1989).
41. Y. M. Torchinsky, in: *Vitamin B₆ Pyridoxal Phosphate, Part B*, John Wiley & Sons, New York, (1986) p. 169.
42. V. M. Powers, C. W. Koo, G. L. Kenyon, J. A. Gerlt, W. K. Kozarich, *Biochemistry*, **30**, 9255 (1991).

43. D. J. Neiderhart, P. L. Howell, G. A. Petsko, V. M. Powers, R. Li, G. L. Kenyon, J. A. Gerlt, *Biochemistry*, **30**, 9264 (1991).
44. J. A. Landro, A. T. Kallarakal, S. C. Ransom, J. A. Gerlt, J. W. Kozarich, D. J. Neidhart, G. L. Kenyon, *Biochemistry*, **30**, 9274 (1991).
45. J. M. Manning, N. E. Merrifield, W. H. Jones, E. C. Gotschlich, *Proc. Natl. Acad. Sci. USA*, **71**, 417 (1974).
46. E. Wang, C. T. Walsh, *Biochemistry*, **17**, 1313 (1978).
47. B. Badet, K. Inagaki, K. Soda, C. T. Walsh, *Biochemistry*, **25**, 3275 (1986).
48. H. Ueno, J. J. Likos, D. E. Metzler, *Biochemistry*, **21**, 4387 (1982).
49. J. J. Likos, H. Ueno, R. W. Feldhaus, D. E. Metzler, *Biochemistry*, **21**, 4377 (1982).
50. D. Roise, K. Soda, T. Yagi, C. T. Walsh, *Biochemistry*, **23**, 5159 (1984).
51. C. T. Walsh, *Trends Biochem. Sci.*, **8**, 254 (1983).
52. E. A. Wang, C. T. Walsh, *Biochemistry*, **20**, 7539 (1981).
53. B. Badet, C. T. Walsh, *Biochemistry*, **24**, 1333 (1985).
54. V. Copie, W. S. Faraci, C. T. Walsh, R. G. Griffin, *Biochemistry*, **27**, 4966 (1988).

9

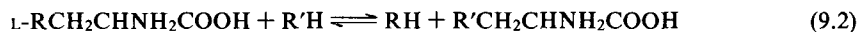
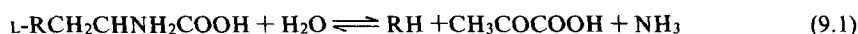
Tryptophanase: Structural, Spectral, and Catalytic Properties

Yuri M. Torchinsky and Yasushi Kawata*

*The Institutes for Applied Research, Ben-Gurion University of the Negev, Beer-Sheva 84110, Israel*** Department of Biotechnology, Faculty of Engineering, Tottori University, Tottori 680, Japan*

9.1 Introduction

Tryptophanase (L-tryptophan indole-lyase (deaminating); EC 4.1.99.1) belongs to the family of the pyridoxal 5'-phosphate (PLP)-dependent enzymes. It serves *in vivo* to degrade L-tryptophan, is induced by L-tryptophan, and found in various bacteria, particularly in enteric species. Tryptophanase catalyzes α,β -elimination¹⁾ and β -replacement reactions on interaction with L-tryptophan and various other β -substituted amino acids²⁾:



where R is an indole, -SH, -SCH₃, -OH, or -OCH₃ group, and R' is an indole group. Monovalent cations K⁺, NH₄⁺, Rb⁺, or Tl⁺ are required for the catalytic activity²⁾. Tryptophanase also catalyzes the synthesis of L-tryptophan from L-serine and indole, like the biosynthetic enzyme, tryptophan synthase. In the early study of tryptophanase, the possibility of contamination by tryptophan synthase was not excluded. However, after finding in 1959 a constitutive tryptophan auxotroph, *Escherichia coli* B/1t7, that lacks the genes encoding for the biosynthetic tryptophan synthase,^{3,4)} studies of the structure and function of tryptophanase have progressed rapidly. A comprehensive review of these studies was written by Snell⁵⁾ in 1975.

Since 1975 the structural genes of tryptophanase from *E. coli* K-12 and *E. coli* B/1t7-A have been isolated and sequenced, the primary structure of the enzyme determined, new catalytic intermediates detected, and single enzyme crystals suitable for X-ray analysis have recently been obtained. These and other studies on the structure and catalytic mechanism of tryptophanase are reviewed in this chapter.

9.2 Structural Properties

9.2.1 Primary and Quaternary Structure

Tryptophanase has been purified from various species, including *E. coli*,^{2,6)} *Bacillus alvei*,⁷⁾ *Aeromonas liquefaciens*,^{8,9)} and *Proteus rettgeri*.^{10,11)} Preparations of tryptophanase from different sources are similar in molecular mass (about 210 kDa), quaternary structure

K-12	-----C-----	
B/1t7-A	ATGGAAACTTTAAACATCTCCCTGAACCGTTCCGCATTCTGTTATTGAGCCAGTAAACGTACTACTCGCGCT	25
	M E N F K H L P E P F R I R V I E P V K R T T R A	
K-12	-----G-----	
B/1t7-A	TATCGTGAAGAAGCAATTATTAATCCGGTATGAACCCGTTCTGCTGGATAGCGAAGATGTGTTATCGATTTA	50
	Y R E E A A I I K S G M N P F L L D S E D V F I D L	
K-12	-----G-G-G-C-----T-----	
B/1t7-A	CTGACCGACAGCGGCACCGGGCGGTAAACCCAAAGTATGCAGGCAGCGATGATGCGCGGCGACGAAGCCTACAGC	75
	L T D S G T G A V T Q S M Q A A M M R G D E A Y S	
K-12	-----T-T-----C-----C-----	
B/1t7-A	GGCAGCCGCACTACTATGCGTTAGCCGAGTCAGTGAAAAATATCTTTGGTTATCAATATACTATTCCGACTCAC	100
	G S R S Y Y A L A E S V K N I F G Y Q Y T I P T H	
K-12	-----C-----	
B/1t7-A	CAGGGCGTGGCGCAGAGCAATCTATATCCGGTACTGATTAACCGAGCAGGAGGAAAGGCTGGATCGC	125
	Q G R G A E Q I Y I P V L I K K R E Q E K G L D R	
K-12	-----G-----C-----	
B/1t7-A	AGCAAAATGGTGGCATTCTCTAACTATTTCTTTGATACACGCAGGGCCATAGCCAGATTAAACGGCTGTACCGTG	150
	S K M V A F S N Y F F D T T Q G H S I N G C T V	
K-12	-----T-G-----	
B/1t7-A	CGTAACGTCTATATCAAAGAAGCCTTCGATACGGCGTGCGTTACGACTTTAAAGGCAACTTTGACCTCGAAGGA	175
	R N V Y I K E A F D T G V R Y D F K G N F D L E G	
K-12	-----T-----	
B/1t7-A	TTAGAACGCGGTATTGAAGAAGTTGGCCCGAATAACGTGCCGTATATCGTTGCAACCATCACCAGTAACCTCTGCA	200
	L E R G I E E V G P N N V P Y I V A T I T S, N S A	
K-12	-----	
B/1t7-A	GGTGGTCAGCCGGTTTCACTGGCAAACTTAAAGCGATGTACAGCATCGCGAAGAAATACGATATCCGGTGGTA	225
	G G Q P V S L A N L K A M Y S I A K K Y D I P V V	
K-12	-----G-----T-----	
B/1t7-A	ATGGACTCCGCACGCTTTGCTGAAACCGCTATTTTCATCAAGCAGCGTGAAGCAGAATACAAGACCTGGACCATC	250
	M D S A R F A E N A Y F I K Q R E A E Y K D W T I	
K-12	-----G-----	
B/1t7-A	GAGCAGATCACCCGCGAAACCTACAATATGCCGATATGCTGGCGATGTCCGCCAAGAAAGATGCCATGGTACCG	275
	E Q I T R E T Y K Y A D M L A M S A K K D A M V P	
K-12	-----C-----	
B/1t7-A	ATGGGCGGCTTGCTGTGCATGAAAGACGACAGCTTCTTTGATGTGTACACCGAGTGCGAACCCTTTGCGTGGTG	300
	M G G L L C M K D D S F F D V Y T E C R T L C V V	
K-12	-----C-----A-----	
B/1t7-A	CAGGAAGGGTTCCCGACATATGGCGGCTTGAAGGCGGCGGATGGAGCGTCTGGCGGTAGGCTCTGTATGACGGC	325
	Q E G F P T Y G G L E G G A M E R L A V G L Y D G	
K-12	-----	
B/1t7-A	ATGAATCTCGACTGGCTGGCTTATCGTATCGCGCAGGTACAGTATCTGGTCGATGGTCTGGAAGAGATTGGCGTT	350
	M N L D W L A Y R I A Q V Q Y L V D G L E E I G V	
K-12	-----T-----	
B/1t7-A	GTCTGCCAGCAGGCGGGCGGTACGCGGCATTCTGTTGATGCCGGTAAACTGCTGCCGATATCCGGCAGACCAG	375
	V C Q Q A G G H A A F V D A G K L L P H I P A D Q	
K-12	-----C-----	
B/1t7-A	TTCCCGGCAACAGGCTGGCGTGGAGCTGTATAAAGTCGCCGGTATCCGTGCGGTAGAAATGGCTCTTTCTCTG	400
	F P A T G L A C E L Y K V A G I R A V E I G S F L	
K-12	-----	
B/1t7-A	TTAGGCGCGATCCGAAACCGGTAACAACCTGCCATGCCCGGCTGAACCTGCTGCGTTTAACCATTCGCGCGCA	425
	L G R D P K T G K Q L P C P A E L R L T I P R A	
K-12	-----	
B/1t7-A	ACATATACTCAACACATATGGACTTCATTATTGAAGCCTTTAAACATGTGAAGAGAAGCGGCGGAATATTAAA	450
	T Y T Q T H M D F I I E A F K H V K E N A A N I K	
K-12	-----G-----G-----T-----	
B/1t7-A	GGATTAACTTTACCTACGAACCAAAAGTATTGCGTCACTTCACGCAAAACTGAAGAAGTTTAA	471
	G L T F T Y E P K V L R H F T A K L K E V *	

Fig. 9.1 Nucleotide sequence and deduced amino acid sequence of tryptophanase from *E. coli* B/1t7-A and K-12. (Reproduced with permission from Tokushige *et al.*, *Biochimie.*, 71, 714 (1989)).

(4 subunits), and PLP content (4 mol/mol-enzyme) but show differences in amino acid composition,^{5,7,12)} some aspects of PLP binding,¹³⁾ and the number of sulfhydryl groups.¹⁴⁾

Structural genes of tryptophanase from *E. coli* K-12¹⁵⁾ and B/1t7-A¹⁶⁾ were cloned and the primary structure deduced as shown in Fig. 9.1. The structure deduced from the nucleotide sequence agrees well with the protein sequence data.¹⁷⁾ Although 29 differences in the nucleotide sequence were found between the *tnaA* genes of *E. coli* K-12 and B/1t7-A, the deduced amino acid sequences are completely identical. The total number of amino acid residues is 471 and the molecular mass of a subunit is 52,242 Da.

The circular dichroism spectrum of *E. coli* tryptophanase¹⁸⁾ showed that the secondary structure of this enzyme seems to be predominantly α -helical. The α -helix content was estimated to be about 50% (Y. Kawata, unpublished results) by the method of Greenfield and Fasman.¹⁹⁾

The three-dimensional structure of tryptophanase has not been determined yet, but crystallization and a preliminary study of the crystals have been reported recently.^{20,21)} Tryptophanase was purified from overproductive cloned cells, *E. coli* K-12 MD55/pMD6B, which produce a large amount of the enzyme corresponding to about 60% of the total soluble protein. Crystallization was performed using polyethylene glycol, potassium phosphate, or ammonium sulfate as precipitants. Fig. 9.2 shows six different types of crystals: holoenzyme (A, B), apoenzyme (C), a complex (D) of holoenzyme and L-alanine all crystallized from polyethylene glycol (4000) solution, a complex (E) of holoenzyme and L-alanine from potassium phosphate buffer, and holoenzyme (F) from ammonium solution. The crystals (Fig. 2(F)) prepared with the highly purified enzyme using ammonium sulfate (13% (w/v)) as a precipitant grew to a maximum size of $0.4 \times 0.4 \times 0.5$ mm in about one week. The buffer used was 25 mM potassium acetate containing 0.25 mM PLP, 2.5 mM dithiothreitol, and 0.04% (w/v) sodium azide at pH 5.25, which is close to the isoelectric point of tryptophanase.²²⁾ The crystals have a rhombic bipyramidal morphology and were obtained reproducibly. These crystals diffracted to a resolution limit of approximately 3 Å in still photographs and the crystals are tetragonal, belonging to space group $P4_12_12$ or its enantiomorph, $P4_32_12$, with unit-cell dimensions of $a=b=113.4$ Å and $c=232.2$ Å. Properties of all crystals are summarized in Table 9.1.

The *E. coli* tryptophanase is a tetrameric enzyme composed of four identical subunits,^{15,16)} each possessing a PLP binding site.¹⁾ PLP affects the stability and conformation of tryptophanase. Morino and Snell^{23,24)} found that apotryptophanase reversibly dissociates into dimers in cold (5°C) at pH 8.0 whereas the holoenzyme does not undergo dissociation at these conditions. The enzyme irreversibly dissociates to a monomeric subunit (protomer) at higher pH values (pH>8.7). Addition of PLP to the tetrameric apotryptophanase significantly increases the sedimentation coefficient and decreases the viscosity, indicating that the holotryptophanase has a more compact form than the apotryptophanase. The free energy change of the reaction of dissociation of the tetrameric apoenzyme to dimers was estimated to be +4 to 6 kcal, and the enthalpy and entropy changes -50 kcal and -198 e.u., respectively, per mole of tetramer, suggesting the predominantly hydrophobic interaction between the two dimers subunits in the tetrameric apoenzyme.²⁴⁾ That the hydrophobic forces are important is also indicated by studies of the effects of anions in the Hofmeister series.²⁵⁾ Studies of the effect of PLP concentration on the activity of the dimeric apotryptophanase indicated that the reassociation of dimers into tetramers is the rate-limiting step in the appearance of enzymatic activity and that the

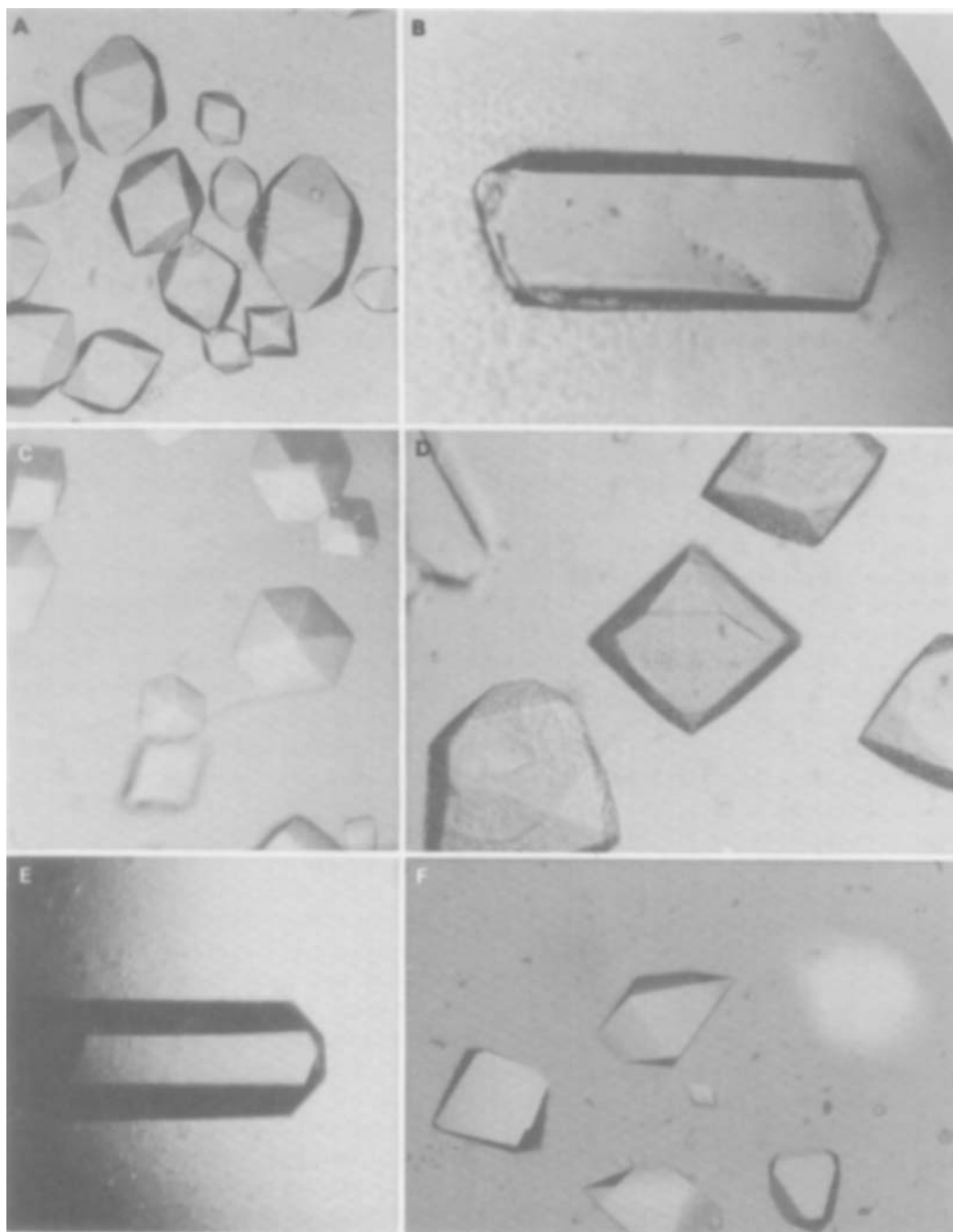


Fig. 9.2 Crystal photographs of *E. coli* tryptophanase. A, B, and F: holoenzyme, C: apoenzyme, D and E: complex of holoenzyme and L-alanine. See Table 9.1 for detailed conditions.

TABLE 9.1 Growth conditions and properties for crystals of *E. coli* tryptophanase

	Precipitant	Buffer	Size (mm)	Crystal data
Holo	13% ammonium sulfate	50 mM potassium acetate, pH 5.25	$0.4 \times 0.4 \times 0.5$ (Fig. 9.2F)	$a = b = 113.4 \text{ \AA}$, $c = 232.2 \text{ \AA}$ P $4_32_12/P 4_12_12$, 3 Å-resolution 2 protomers/asymmetric unit
	18% PEG 4000	10 mM potassium phosphate, pH 7.8	$0.6 \times 0.6 \times 0.8$ (Fig. 9.2A)	$a = b = 153.3 \text{ \AA}$, $c = 141.8 \text{ \AA}$ hexagonal, 7 Å-resolution
	8% PEG 4000	50 mM potassium phosphate, pH 7.0	$0.3 \times 0.3 \times 1.0$ (Fig. 9.2B)	
Apo	27% PEG 4000	10 mM potassium phosphate, pH 7.8	$0.3 \times 0.3 \times 0.3$ (Fig. 9.2C)	
Complex with L-alanine	18% PEG 4000	10 mM potassium phosphate, pH 7.8	$0.6 \times 0.6 \times 0.5$ (Fig. 9.2D)	
	1.22 M potassium phosphate	1.22 M potassium phosphate, pH 7.8	$0.2 \times 0.2 \times 1.0$ (Fig. 9.2E)	$a = 90 \text{ \AA}$, $b = 140 \text{ \AA}$, $c = 190\text{--}200 \text{ \AA}$, 4 Å-resolution

tryptophanase dimer is not functional.²⁵⁾ The existence of tight coupling between protomers within tryptophanase was suggested by the finding that hybrid molecules containing two holo- and two aposubunits remain tetrameric at low temperatures, whereas the apoenzyme dissociates into dimers.²⁶⁾ Also using hybrid molecules, the transmission of a structural change from holoprotomers to apoprotomers was observed; the presence of three holoprotomers containing the reduced PLP in a hybrid molecule imposes a holo conformation on the fourth apoprotomer.²²⁾ The kinetic anticooperativity in the binding of PLP was detected and was interpreted in relation to the coupling between the protomers.²²⁾

Studies of the radii of gyration of the holotryptophanase by using the synchrotron small-angle X-ray scattering confirmed that the holoenzyme does not dissociate into dimers at 20°C or 5°C irrespective of the presence or absence of K⁺ ions.¹⁶⁾ The PLP analogs, which lack the 4'-carbonyl group, *i.e.* pyridoxine 5'-phosphate and 5'-phosphopyridoxyl tryptophan, promote the conversion of dimers to the tetramer at 5°C²⁴⁾ but do not give the catalytically active enzyme. These analogs also promote the dimer to tetramer transition of the *Bacillus alvei* enzyme at 9°C.²⁷⁾

Honda and Tokushige reinvestigated the effect of temperature and monovalent cations on the quaternary structure of tryptophanase using the HPLC gel-filtration analysis.²⁸⁾ In contrast to the above-mentioned data, they found that in the absence of K⁺ or NH₄⁺ ions the tetrameric holoenzyme undergoes dissociation into dimers and inactivation at 5°C. Their results also indicate that formation of the active holoenzyme from the apoenzyme and PLP proceeds in two steps: the inactive tetrameric species are first formed then converted to the active species; only the second step requires the presence of K⁺ or NH₄⁺ ions (Fig. 9.3)

The three-dimensional and primary structures of the closely related enzyme, tyrosine phenol-lyase, have recently been determined.^{28a)} Since 43% of the amino acid sequences of this enzyme and tryptophanase is identical, their three-dimensional structures should be very similar. The structure of the lyase is like a butterfly. The tetrameric molecule can be considered from a crystallographic point of view as an $\alpha_2\alpha'_2$ -structure consisting of two

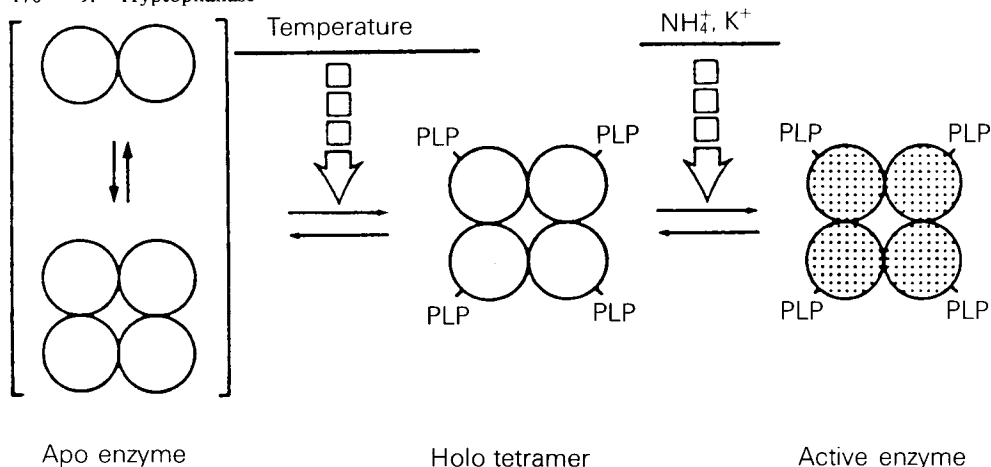


Fig. 9.3 A model for formation of active holotryptophanase. (Reproduced with permission from Honda and Tokushige, *J. Biochem.*, **100**, 684 (1986)).

dimers (α, α and α', α') related by a non-crystallographic symmetry axis. The “dimers” are bound by the highly hydrophobic cluster in the center of the molecule and by the intertwined amino-terminal stretches of the polypeptide chains. Each of the “dimers” contains two active sites located at the subunit interface and formed by residues from both subunits. The domain architecture in the “dimer” of the lyase proved to be similar to that in the dimeric molecule of aspartate aminotransferase.

9.2.2 Unfolding and Refolding

Reconstitution experiments with *E. coli* tryptophanase after unfolding in 8 M urea using dialysis indicate that the enzyme refolds reversibly, but the refolding reaction is highly dependent on the protein concentration.²⁹⁾ This dependence is due to a refolding intermediate which has a strong tendency toward irreversible aggregation at high protein concentrations.

Recently, the unfolding and refolding properties of *E. coli* holotryptophanase in guanidine hydrochloride (Gdn-HCl) have been studied at 25°C by measuring fluorescence and enzymatic activity (T. Mizobata and Y. Kawata, unpublished results). Kinetic experiments have been performed by monitoring the tryptophan fluorescence at 340 nm, the PLP fluorescence at 500 nm, and the Schiff base absorption at 420 nm at various concentrations of Gdn-HCl. It was found that the unfolding kinetics of tryptophanase involves two intermediate species. One is the native apoenzyme and the other is an intermediate which has a hydrophobic region exposed to the solvent and is characterized by a tendency to aggregate irreversibly. Unfolded tryptophanase regains only part of its initial activity upon dilution of the denaturant. This is due to the irreversible aggregation of folding intermediates.^{29–31)} When a molecular chaperone protein, GroE from *E. coli*, is included in the refolding mixture, the refolding yield of tryptophanase is greatly improved.³¹⁾ This effect was observed only with GroE but not with bovine serum albumin or polyethylene glycol, indicating that the GroE facilitates the refolding of tryptophanase specifically (Fig. 9.4).

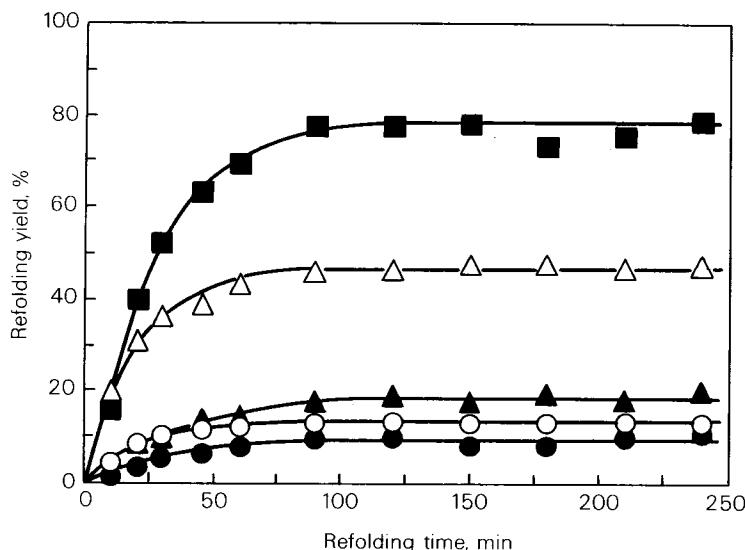


Fig. 9.4 Effects of GroE on the refolding of *E. coli* tryptophanase. The refolding of tryptophanase in the presence of various factors which facilitate folding, as monitored by enzymatic activity at 25°C. The concentration of tryptophanase was 5 µg/ml (9.6×10^{-8} M tryptophanase monomer). Tryptophanase was unfolded in 4 M Gdn-HCl and diluted into refolding mixture containing the indicated amount of GroE complex. (○), refolding in the presence of buffer only; (△), refolding in the presence of a 5-fold molar excess of GroE complex relative to the monomer concentration of tryptophanase; (■), refolding in the presence of a 15-fold molar excess of GroE; (▲), refolding in the presence of 1.0 mg/ml bovine serum albumin; (●), refolding in the presence of 10 mg/ml polyethylene glycol. Refolding experiments were performed in 50 mM Hepes-KOH, pH 7.8, containing 50 mM KCl, 10 mM $\text{Mg}(\text{CH}_3\text{COO})_2$, 2 mM DTT, 0.06 mM PLP, and 2 mM ATP. (Reproduced with permission from Mizobata *et al.*, *J. Biol. Chem.*, **267**, 17775 (1992)).

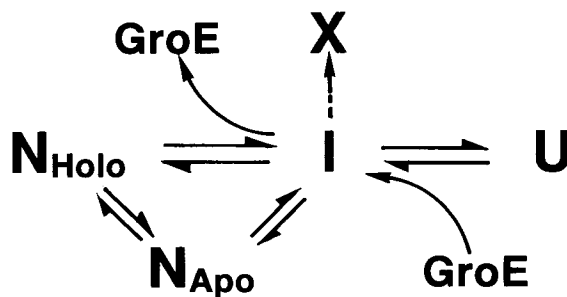


Fig. 9.5 Unfolding and refolding mechanism of *E. coli* tryptophanase. See text for details.

The unfolding and refolding mechanism of *E. coli* tryptophanase is summarized in Fig. 9.5 (T. Mizobata and Y. Kawata, unpublished results), where N is the native state, I the intermediate having a tendency toward aggregation, X the irreversible aggregates, and U the unfolded state. At below 2 M Gdn-HCl, holotryptophanase unfolds via apoenzyme ($N_{\text{holo}} \rightarrow N_{\text{apo}} \rightarrow I \rightarrow U$ (or X)), whereas at high concentrations of Gdn-HCl (> 2 M),

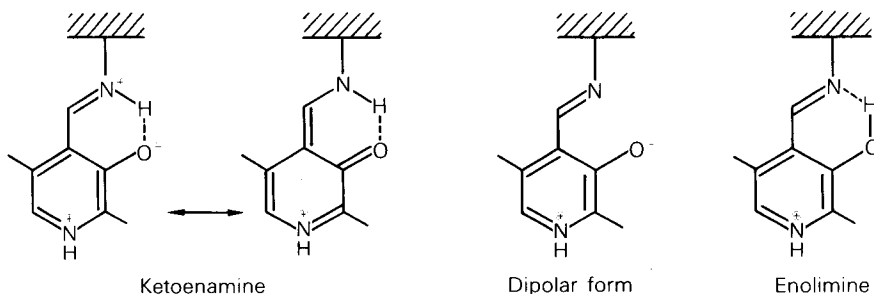
apo form (N_{apo}) and aggregates (X) are not observed ($N_{\text{holo}} \longrightarrow \text{I} \longrightarrow \text{U}$). In the refolding reaction, GroE stabilizes the I state and blocks the irreversible aggregation of this folding intermediate.

9.3 Spectral Properties of the Free Enzyme and Catalytic Intermediates

9.3.1 Absorption Spectra

The characteristic spectral properties of tryptophanase in the visible and near UV region are due to pyridoxal-P-lysine aldimine (Schiff base).³²⁾ In the presence of K^+ or NH_4^+ ions the enzyme displays pH-dependent spectra with maxima at 420 and 337 nm (Fig. 9.6). The 420-nm band predominates at acid pH and gradually diminishes as pH is raised from 6 to 9, whereas another band at 337 nm increases in intensity. From a plot of absorbance *versus* pH at these wavelengths the $\text{p}K_{\text{a}}$ value of the coenzyme chromophore was found to be 7.2³²⁾ or 7.5.³³⁾ The species with λ_{max} 337 nm is the active enzyme form because it prevails at the pH-optimum for tryptophanase activity.

Morino and Snell³²⁾ ascribed the 420-nm species to a protonated internal pyridoxal-P-lysine aldimine and the 337-nm species to either a dipolar ionic form of the aldimine or to a substituted aldamine (an adduct at C-4'). However, the dipolar ionic form absorbs at about 360 nm in aspartate transaminase and in the model systems containing Schiff bases of *N*-methylated PLP.³⁴⁾ It was suggested that the 337-nm species may be the enolimine tautomer.^{35,36)}



A striking feature of tryptophanase spectra is their dependence on monovalent cations. In the absence of cations, e.g. in $(\text{CH}_3)_4\text{N}/\text{Hepps}$ buffer or in the presence of Na^+ ions, the enzyme exhibits an increased pH-independent absorption peak at 420 nm and little or no absorption at 337 nm even at pH 8.0.³²⁾ Addition of activating cations, e.g. K^+ , to the enzyme solution in Hepps buffer at pH 8 results in a decrease of absorbance at 420 nm and an increase at 337 nm, *i.e.* produces an effect similar to that of raising pH.³⁷⁾

Addition of substrates (L-tryptophan, *S*-methyl-L-cysteine and others) or competitive amino acid inhibitors to tryptophanase solutions induces an intense narrow absorption peak near 500 nm with a shoulder at 470–475 nm; the peak rapidly disappears as the substrates are decomposed, but it is relatively stable in the presence of amino acid

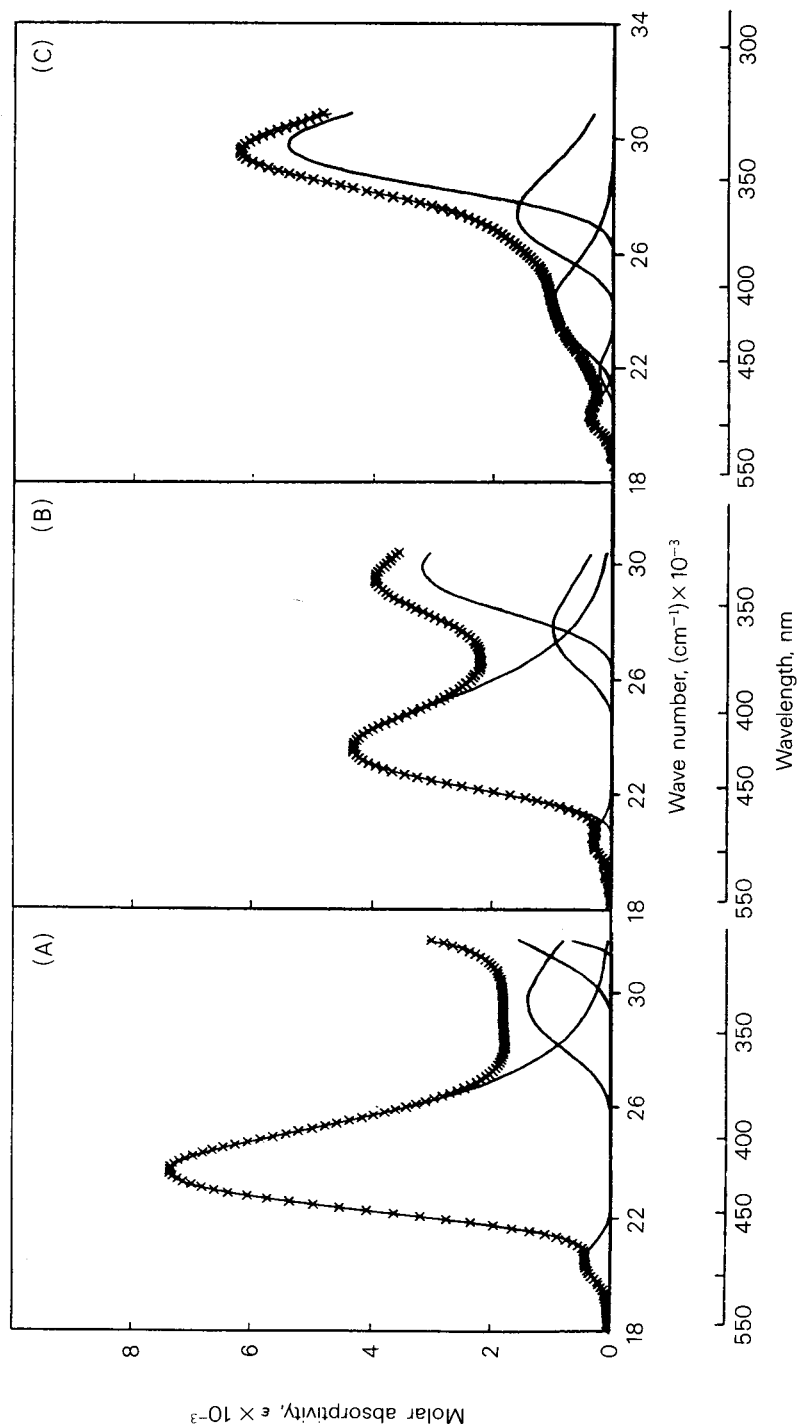


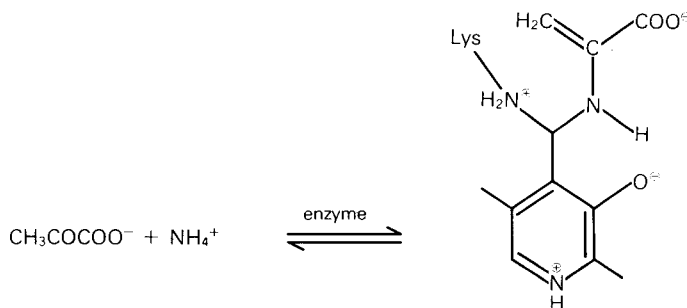
Fig. 9.6 Absorption spectra of three ionic forms of tryptophanase in 0.1 M phosphate. A, the diprotonated, low pH form; B, the monoprotonated form; C, the unprotonated, high pH species. The spectra were resolved with log-normal curves. (Reproduced with permission from Metzler *et al.*, *J. Biol Chem.*, **266**, 9376 (1991)).

inhibitors.*¹ The shape and position of this peak (a strong shift to longer wavelengths) indicate that it belongs to a quinonoid or carbanion form that lacks the α -proton of the amino acid. Consistent with this proposal is the finding that the α -proton of inhibitory amino acids as well as substrates undergoes isotopic exchange in D₂O or T₂O.^{32,39)}

Unusually many amino acids and their derivatives (alanine, oxindolyl-alanine, phenylalanine, dihydroxyphenylalanine, homophenylalanine, α -aminobutyrate, norvaline, norleucine, leucine, allylglycine, β -cyanoalanine, homoserine, *O*-methylserine,*² methionine, ethionine, tyrosine, *S*-benzylcysteine,*² cysteine,*² *S*-methylcysteine,*² asparagine, glutamine, arginine, histidine, ornithine) act as competitive inhibitors of tryptophanase and also evoke the 500-nm band.⁴⁰⁾ Amino acids that have an asymmetric carbon at the β -position (L-threonine*², L-allothreonine*², L-isoleucine, β -phenyl-L-serine, β -methyl-L-tryptophan) as well as L-valine, phenylglycine and glycine do not evoke this band.⁴⁰⁾

A quinonoid intermediate is formed only in the presence of activating cations, which seem to be required for labilization of the α -hydrogen of the enzyme-substrate complex. Suelter and Snell³⁷⁾ found that $V_{\max}$ observed with most cations except NH₄⁺ is directly proportional to the extent to which they elicit formation of the 500-nm absorbing species.

Rapid scanning and single-wavelength stopped-flow spectrophotometry has made it possible to detect new transient reaction intermediates. During the reaction of tryptophanase with L-homophenylalanine, a transient intermediate absorbing at 340 nm was observed which decayed at the same rate as the 508-nm quinonoid peak was formed.⁴¹⁾ It has been suggested that this intermediate is the *gem*-diamine of PLP with the α -amino group of the substrate analog and the ϵ -amino group of Lys-270.⁴¹⁾ Phillips⁴²⁾ found that benzimidazole, an indole analog, can trap an intermediate with $\lambda_{\max}$ 345 nm which is formed during the decomposition of the quinonoid complex with L-tryptophan. Addition of benzimidazole to a solution of enzyme, pyruvate and NH₄Cl results in the formation of a similar absorption peak at 340 nm. Phillips⁴²⁾ has suggested that this peak may belong to an α -aminoacrylate intermediate bound as a *gem*-diamine:



9.3.2 Circular Dichroism

The binding of a symmetrical molecule of PLP to the active site of tryptophanase generates so-called induced optical activity ("extrinsic" Cotton effect), which was detected

*¹ Torchinsky *et al.*³⁸⁾ found that the 502-nm band observed with L-alanine markedly decreases during several hours at 30°C and simultaneously pyruvate and pyridoxamine phosphate are formed. It was concluded that L-alanine undergoes slow transamination in the active site tryptophanase.

*² These amino acids are also substrates of tryptophanase.

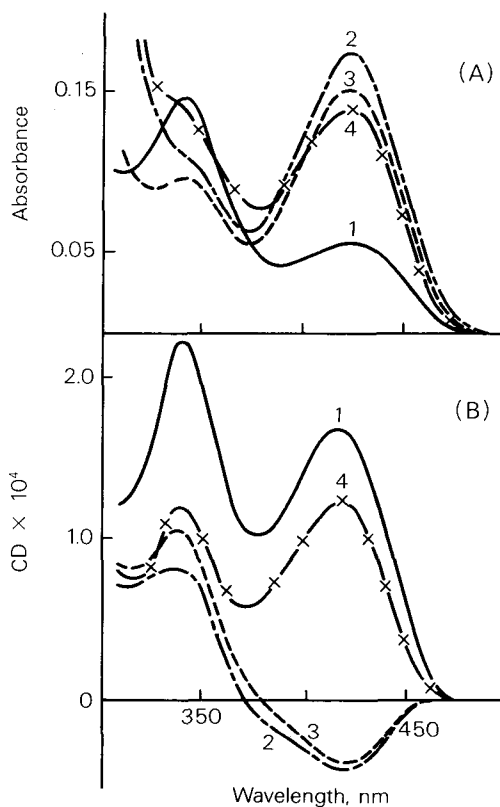


Fig. 9.7 Effects of β -phenyl-DL-serine, L-threonine and α -methyl-DL-serine on absorption (A) and CD (B) spectra of tryptophanase in 100 mM potassium phosphate buffer, pH 7.8, containing 2 mM dithiothreitol and 2 mM EDTA. Protein concentration 1.8 mg/ml. Curve 1, unliganded enzyme. Curve 2, same as 1 + 15 mM β -phenyl-DL-serine (*threo* form). Curve 3, same as 1 + 165 mM L-threonine. Curve 4, same as 1 + 500 mM α -methyl-DL-serine. (Reproduced with permission from Zakomirdina *et al.*, *Biochimie.*, 71, 548:549 (1989)).

by recording circular dichroism (CD).^{38,43,44)} The unliganded tryptophanase displays positive CD in the absorption bands at 337 and 420 nm. Interaction of the enzyme with the quasisubstrates, β -phenyl-DL-serine (*threo* form) and L-threonine, at pH 7.8 leads to a decrease in absorbance at 337 nm and to a strong increase at 425 nm (Fig. 9.7A). This is associated with inversion of the CD sign, *i.e.* with disappearance of the positive CD in the 420-nm band and its replacement by a negative CD (Fig. 9.7B). L-Phenylalanine and D-alanine also cause a strong increase in enzyme absorption at 425–430 nm and a large diminution in the positive CD (Figs. 9.8 and 9.9). The simultaneous addition of D-alanine and indole evokes a negative CD (Fig. 9.9B). The dissymmetry (anisotropy) factor ($\Delta A/A$) in the 425-nm band is reduced manifold in the presence of the above-mentioned inhibitory amino acids (Table 9.2).

Breaking of the internal PLP-lysine aldimine bond by reaction with carbonyl reagents, hydroxylamine and aminooxyacetate was also accompanied by a strong diminution of the positive CD and by a reduction of the dissymmetry factor.⁴⁴⁾ It has been inferred that the

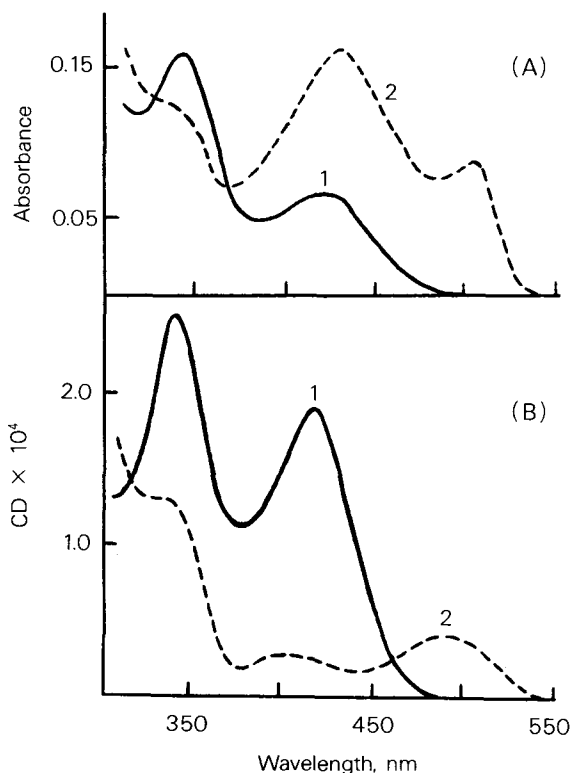


Fig. 9.8 Effect of L-phenylalanine on absorption (A) and CD (B) spectra of tryptophanase. For conditions see legend to Fig. 9.7. Curve 1, unliganded enzyme. Curve 2, same as 1 + 100 mM L-phenylalanine. (Reproduced with permission from Zakomirdina *et al.*, *Biochimie.*, **71**, 548:549 (1989)).

intense 425-nm band observed in the presence of β -phenyl-DL-serine, L-threonine, L-phenylalanine and D-alanine, belongs to the external PLP-quasisubstrate aldimine, which results from the transamination reaction, *i.e.* displacement of the ϵ -NH₂ group of a lysine residue from its bond with PLP by the α -NH₂ group of a substrate analog.^{38,44} Such an interpretation is in agreement with the linear dichroism findings (see Section 9.5.2) and also with the results of stopped-flow kinetic studies of amino acid inhibitor binding, which showed that the enzyme species with $\lambda_{\max}$ 337 nm is converted to the 420-nm species before the α -proton of an amino acid is removed to form quinonoid.⁴⁵ It seems likely that the 425-nm species observed on addition of quasisubstrates mimics a true intermediate of the catalytic reaction.

A positive CD was found in the 500-nm quinonoid band which is formed on reaction of tryptophanase with L-alanine and oxindolyl-L-alanine (Fig. 9.10). The dissymmetry factor in this band is much smaller than in the absorption bands of the unliganded enzyme (Table 9.2). A negative 315-nm peak, which appears in the presence of L-alanine (Fig. 9.10), may be caused by interaction of an aromatic amino acid residue with the quinonoid coenzyme ring.

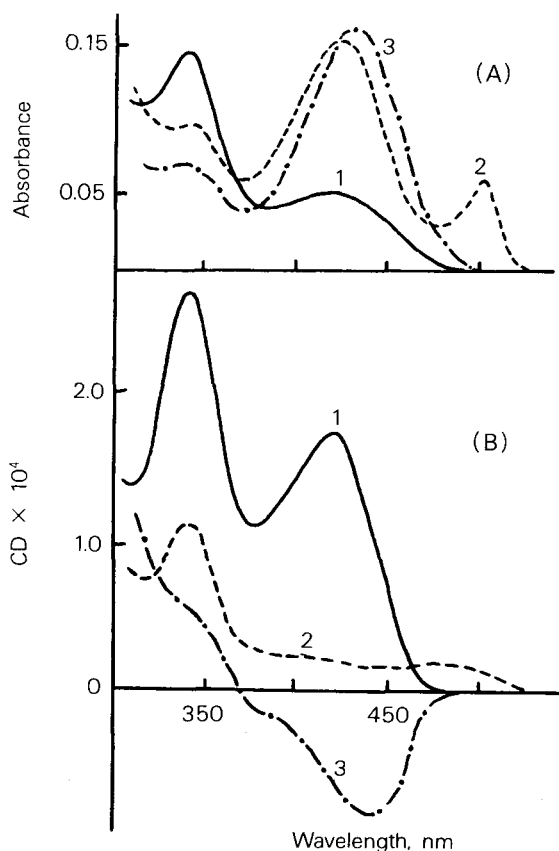


Fig. 9.9 Effects of D-alanine and indole on absorption (A) and CD (B) spectra of tryptophanase. For conditions see legend to Fig. 9.7. Curve 1, unliganded enzyme. Curve 2, same as 1 + 300 mM D-alanine. Curve 3, same as 1 + 300 mM D-alanine and 4 mM indole. (Reproduced with permission from Zakomirdina *et al.*, *Biochimie.*, **71**, 548:549 (1989)).

9.3.3 Fluorescence

On excitation of holotryptophanase in the absorption bands at 420 nm and 337 nm, the fluorescence is maximally emitted at about 500 and 385 nm, respectively⁴⁶⁾ (see Fig. 9.11). The fluorescence lifetime depends on the presence of monovalent cations.¹⁶⁾

When the holoenzyme was excited at 290 nm, two emission peaks were observed, one at 340 nm, arising from the tryptophan residues, and one at about 500 nm.^{46,47)} The intensity of fluorescence of the holoenzyme at 340 nm was about 60% of that of the apoenzyme. Therefore, binding of PLP induces a marked quenching of tryptophan fluorescence.

The emission peak at 500 nm was absent when the apoenzyme was excited at 290 nm⁴⁷⁾ (Fig. 9.11). This peak was also diminished in a mutant holoenzyme in which Trp-248 was replaced by phenylalanine.¹⁸⁾ It was proposed that this peak is due to energy transfer from the tryptophan residues to the coenzyme. Binding of anthranilic acid or alanine quenches the 500-nm peak and, therefore, interferes with the energy transfer.⁴³⁾

TABLE 9.2^{†1} The dissymmetry factor and reduced linear dichroism values in the coenzyme absorption bands of tryptophanase and its complexes with inhibitory amino acids^{†2}

Enzyme form	Ligand concentration mM	Absorption maximum nm	Dissymmetry factor $\times 10^4$	Reduced linear dichroism $\times 10^3$
Unliganded enzyme	—	337	15.5	10
	—	420	27.8	5
Complex with L-threonine	165	425	-2.6	56
Complex with β -phenyl-DL-serine (threo form)	15	425	-2.3	66
Complex with L-phenylalanine	100	430	1.7	—
Complex with D-alanine	300	425	1.5	42
Complex with D-alanine and 4 mM indole	300	435	-5.0	65
Complex with L-alanine	100	502	3.2	51
Complex with oxindolyl-L-alanine	0.22	504	3.0	54

^{†1} The table is compiled from data in Refs. 44 and 69.^{†2} Spectra were measured in 100 mM potassium phosphate buffer, pH 7.8.

9.3.4 NMR Spectra

In the free tryptophanase the ^{31}P chemical shift of bound PLP was found to be pH-independent between pH 5.8 and 8.2, indicating that the phosphate group is in the dianionic form throughout this pH range.⁴⁸⁾ The same pH-independent ^{31}P resonance signal was observed in the presence of β -phenyl-DL-serine. In contrast, the quinonoid complex with L-alanine displayed pH dependence of the ^{31}P chemical shift with a $\text{p}K_{\text{a}}$ of 5.93, which was only 0.4 pH unit lower than the $\text{p}K_{\text{a}}$ of the model Schiff base between PLP and ϵ -aminocaproate.⁴⁸⁾ These data may indicate that formation of the quinonoid intermediate is accompanied by loosening of contacts between the phosphate group of PLP and the protein.

Downfield ^1H -NMR spectra of tryptophanase in H_2O were measured for studying hydrogen-bonded network around PLP in the active site.⁴⁹⁾ The peaks at 17.7 and 16.3 ppm observed in the presence of L-threonine and L-alanine, respectively, were tentatively ascribed to the proton on the ring nitrogen of PLP.

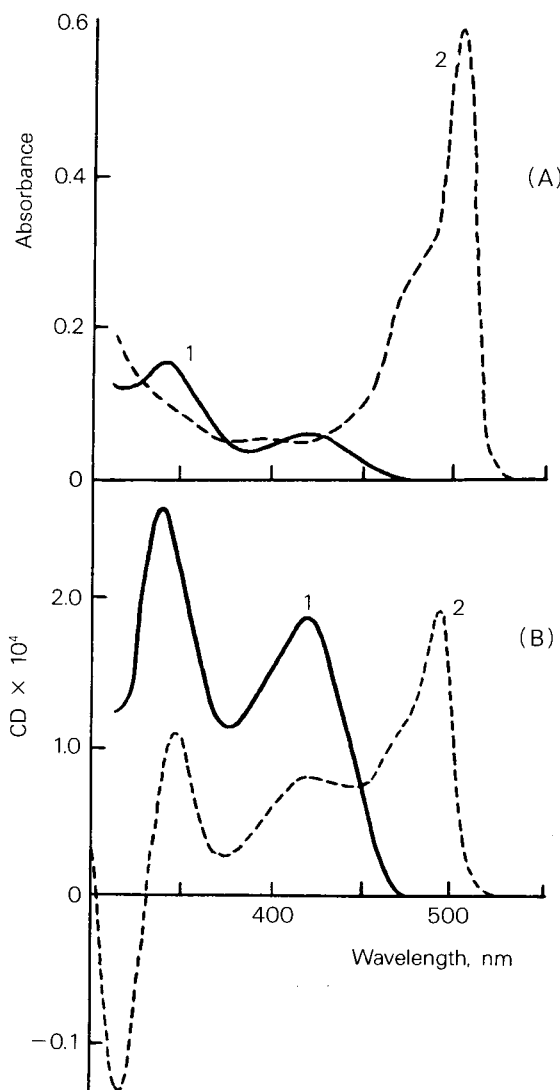


Fig. 9.10 Effect of L-alanine on absorption (A) and CD (B) spectra of tryptophanase. Protein concentration 2 mg/ml. For other conditions see legend to Fig. 9.7. Curve 1, unliganded enzyme. Curve 2, same as 1 + 100 mM L-alanine. (Reproduced with permission from Zakomirdina *et al.*, *Biochimie.*, **71**, 548:549 (1989)).

9.4 Chemical Modification and Site-Directed Mutagenesis

Group-specific chemical modification remains a useful method for studies of structure-function relationships in protein molecules, although unambiguous identification of essential amino acid residues and elucidation of their function are nowadays accomplished mainly by X-ray crystallography and site-directed mutagenesis. Chemical modifications

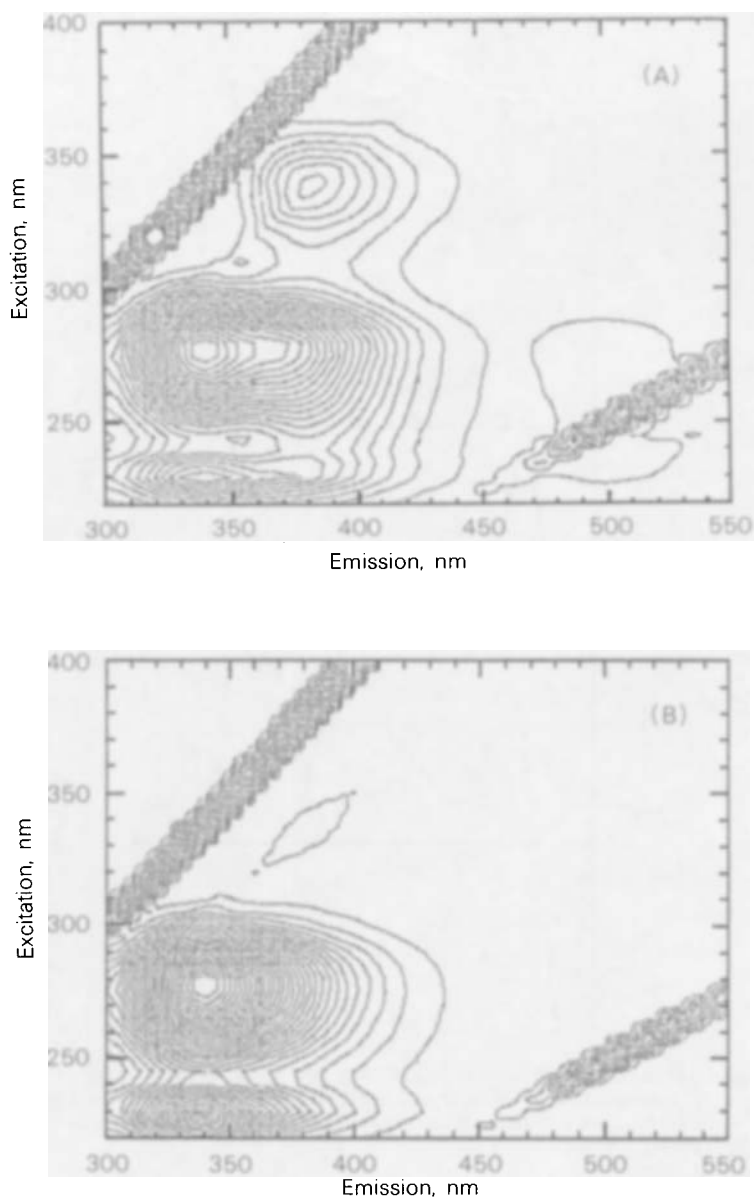


Fig. 9.11 Two-dimensional fluorescence spectra of *E. coli* holo- (A) and apotryptophanase (B) in 100 mM potassium phosphate, pH 7.8. Fluorescence spectra were measured at 5-nm intervals of excitation wavelength and represented in contour. The two cross lines at the upper left-hand and lower right-hand corners in maps were due to primary and secondary scattering, respectively. Source: Ref. 18

of *E. coli* tryptophanase have been carried out extensively by several groups of authors. Their findings as well as the results of recently performed site-directed mutations are briefly described below.

Lysine residues Morino and Snell³²⁾ reduced holotryptophanase with NaBH₄ and isolated ϵ -pyridoxyl-lysine from the hydrolyzed protein. Subsequent isolation of the pyridoxyl-containing peptide⁵⁰⁾ and determination of its position in the total amino acid sequence of the enzyme led to identification of Lys-270, which forms a Schiff base with PLP.

The adjacent residue, Lys-269 was replaced by arginine by site-directed mutagenesis.⁵¹⁾ The mutant enzyme exhibited the k_{cat} value about 90% lower than that of the wild-type enzyme, and the k_{cat}/K_m value about 98% lower. The reaction of the mutant enzyme with substrates and L-alanine was not accompanied by accumulation of the quinonoid intermediate. Phillips *et al.*⁵¹⁾ suggested that Lys-269 plays a significant role in the electrostatic interactions or conformational changes obligatory to the formation and breakdown of the quinonoid intermediate, although it is not a catalytically essential residue.

Arginine residues Kazarinoff and Snell⁵²⁾ showed that apo- and holotryptophanase were inactivated by the arginine-specific reagent phenylglyoxal, and that inactivation was correlated with modification of two arginine residues (out of 24) per subunit. The modified holoenzyme did not form quinonoid complexes with substrates, and its absorption spectra were essentially invariant between pH 6 and 8.5. A curious finding was that it proved impossible to resolve the PLP from the phenylglyoxal-treated holoenzyme by a procedure which removed 90% of PLP from the native enzyme. Since substrate analogs protected the holoenzyme from inactivation and arginines from modification, it was suggested that an arginine residue was located at the substrate binding site.

Histidine and tyrosine residues Nihira, Toraya and Fukui⁵³⁾ found that irradiation of holotryptophanase solutions with visible light led to rapid, pH-dependent loss of enzymatic activity. The apoenzyme was not inactivated under the same conditions; therefore, the inactivation was PLP-sensitized. L-Tryptophan and L-alanine markedly decreased the rate of photoinactivation. From a plot of the inactivation rate *versus* pH, the pK_a was found to be 7.2, which is close to the pK_a of the imidazole group. Photoinactivation was accompanied by destruction of one histidine residue per subunit. No concomitant loss of tryptophan, tyrosine or methionine was detected, and it was suggested that tryptophanase contains an essential histidine residue near the active site.

The presence of an essential tyrosine residue near the active site was suggested on the basis of experiments with tetranitromethane.⁵⁴⁾ Treatment of apotryptophanase with this reagent caused almost complete loss of catalytic activity, a great reduction of affinity for PLP and modification of about one tyrosine residue. The modified enzyme was unable to form the quinonoid intermediate with L-tryptophan or L-alanine.⁵⁵⁾ PLP protected the apoenzyme from inactivation only in the presence of activating cations (K^+ , NH_4^+ , Rb^+). It was shown that inactivation by tetranitromethane was not caused by oxidation of SH-groups, but partial modification of methionine (0.8 residue) was detected and might also be responsible for inactivation. It is worthy of note that modification of tryptophanase with chloramine T indicated that some methionine residues may be important for maintaining the catalytically active conformation of the enzyme.⁵⁶⁾

Tryptophan residues Tryptophanase from *E. coli* has two tryptophans per subunit. Tokushige, Fukada and Watanabe⁵⁷⁾ found that ozonization of apotryptophanase led to oxidation of one tryptophan residue and rapid loss of catalytic

activity. The affinity of the ozonized apoenzyme for PLP was decreased almost 100-fold. The holoenzyme was resistant to ozonization. It was suggested that one tryptophan residue is located near the coenzyme. This suggestion was recently checked by site-directed mutagenesis. Phillips and Gollnick⁵⁸⁾ prepared two tryptophanase mutants in which one of the tryptophan residues, Trp-248 or Trp-330, was replaced by phenylalanine. Both mutants had full catalytic activity, and their affinity for PLP was only slightly decreased: the K_d of PLP was found to be 0.5 μ M for the wild-type enzyme and about 1.3 μ M for the mutants. Kawata *et al.*¹⁸⁾ found however that the affinity of the Trp248Phe mutant for PLP is reduced 10-fold, and its k_{cat} is about 85% of that of the wild-type enzyme. In any case, it is evident that none of the two tryptophan residues is catalytically essential.

Cysteine residues Tryptophanase contains both exposed and buried cysteine residues; their total number is seven per subunit. Two SH-groups per subunit were found to react with 5,5'-dithiobis(2-nitrobenzoate) in the native *E. coli* B enzyme, and enzymatic activity is completely lost coincident with their modification.²³⁾ However, in the native *E. coli* K12 enzyme four SH groups were reported to react with 5,5'-dithiobis(2-nitrobenzoate) and *N*-ethylmaleimide.^{59,60)} The kinetics of inactivation by *N*-ethylmaleimide indicated that loss of 95% activity resulted from the alkylation of a single, highly reactive cysteine residue.⁶⁰⁾ Similarly, inactivation of the enzyme by bromopyruvate^{40,61)} or 2-nitro-5-thiocyanobenzoate⁶²⁾ was shown to be caused by modification of one SH-group per subunit. PLP protected the apoenzyme from inactivation. Tokushige *et al.*¹⁶⁾ found that modification of SH-groups with 5,5'-dithiobis(2-nitrobenzoate) abolished the stabilizing effect of NH_4^+ and K^+ ions on the enzyme and suggested that an SH-group may be involved in the formation of the cation binding site.

Watanabe and Snell⁴⁰⁾ isolated the labeled cysteinyl peptide from ¹⁴C-bromopyruvate-inactivated apotryptophanase. Comparison of this peptide with the total amino acid sequence indicated that Cys-298 was the alkylated residue. Honda and Tokushige⁶³⁾ found that Cys-298, Cys-148 and Cys-352 reacted with *N*-(4-anilino-1-naphthyl)-maleimide and that Cys-298 was the most extensively modified.

Phillips and Gollnick³³⁾ replaced Cys-298 and Cys-294 by serines using site-directed mutagenesis. Both mutants proved to be highly active, indicating that neither of the two cysteine residues is catalytically essential. The Cys294Ser mutant was virtually identical to the wild-type enzyme with respect to most properties. In contrast, the Cys298Ser mutant exhibited a three-fold reduced affinity for PLP, a change in the pH profile for catalysis, a shift of the absorption maximum from 420 to 414 nm, and an altered coenzyme pK_a ; which was raised from 7.5 for the wild-type enzyme to 8.5 for the mutant. An alkaline shift of the coenzyme pK_a , *i.e.* a change in the equilibrium between the 337 and 420-nm species, had been described earlier for tryptophanase with an air-oxidized or cyanylated SH-group.⁶²⁾ Thus, Cys-298 seems to be involved in the interconversion of the two forms of the enzyme, and may be located in or near the active site.

9.5 Catalytic Properties

9.5.1 Activation by Cations

Several monovalent cations (NH_4^+ , K^+ , Tl^+ , Rb^+) have been shown to activate tryptophanase and increase its affinity for PLP.^{2,37,59)} The cations also strongly affect the

spectral properties of the enzyme (Section 9.3.1) and reactivity of a cysteine residue.⁶⁰⁾ The pK_a values for the cations and their effect on the catalytic parameters of tryptophanase are presented in Table 9.3. Högborg-Raibaud, Raibaud and Goldberg⁵⁹⁾ reported that Na^+ ions do not activate the enzyme and even decrease the activating effect of K^+ and NH_4^+ ions. Suelter and Snell³⁷⁾ showed, however, that in the presence of high concentrations of Na^+ ions nearly 30% of enzymatic activity observed with NH_4^+ ions can be achieved (Table 9.3). Earlier findings⁵⁹⁾ indicating very low activity ($\sim 5\%$) in the presence of Na^+ ions can be explained by their low affinity for the enzyme. The stoichiometry of the cation binding was determined in experiments with radioactive thallium; two moles of thallium (I) were found to be bound per mole of protomer.³⁷⁾

Little is known about the mechanism of activation of tryptophanase by cations. Morino and Snell²⁴⁾ found that the sedimentation rate, $S_{20,w}$, of the tetrameric holotryptophanase is different in the absence and presence of K^+ ions (9.5 and 10.5S, respectively), indicating a significant conformational change. On the other hand, synchrotron small-angle X-ray scattering experiments showed that the radius of gyration is constant, irrespective of the presence or absence of K^+ ions.¹⁶⁾ Fluorescence decay measurements of the holoenzyme have revealed that the fluorescence lifetime of the coenzyme is markedly dependent on the presence of the monovalent cations, while the fluorescence lifetime of the tryptophan residues is independent on the presence of the cations.¹⁶⁾ These data suggest that the cations elicit a conformational change near the active site. From studies of the absorption and CD spectra, Suelter and Snell³⁷⁾ inferred that a competitive inhibitor, ethionine, forms external aldimine with the enzyme in the absence of monovalent cations; however, formation of the quinonoid complex occurs only after addition of the cations, which are therefore required for labilization of the α -hydrogen.

TABLE 9.3 The K_a values for monovalent cations and the effect of each cation on the catalytic parameters of tryptophanase³⁷⁾

Cation	K_a (mM)	Concentration of activating ion (mM)	SOPC	
			K_m (mM)	V_{max} (units/mg)
Li^+	54	850	0.068	4.3
Na^+	40	500	0.245	18.8
K^+	1.44	100	0.071	35.6
Tl^+	0.95	50	0.090	39
NH_4^+	0.23	50	0.055	57.9
Rb^+	3.5	100	0.064	33.5
Cs^+	14.6	100	0.155	21

Note: The kinetic constant for activation with each cation and the catalytic parameters for S-*o*-nitrophenyl-L-cysteine (SOPC) (0.6 mM) were determined from double reciprocal plots at 25°C and pH 8.0.

9.5.2 Coenzyme Reorientations in the Active Site

Reorientations of the coenzyme associated with the enzymatic catalytic cycle were first proposed by Ivanov and Karpeisky⁶⁴⁾ for aspartate transaminase on the basis of general topochemical considerations and studies of induced optical activity. Torchinsky and Koreneva^{65,66)} found that induced optical activity of PLP bound to the transaminase is

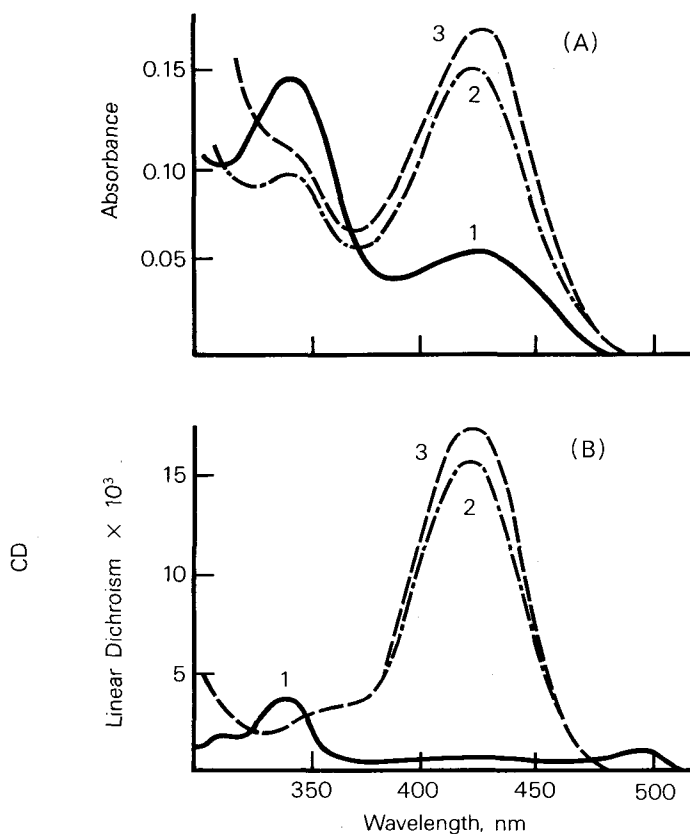


Fig. 9.12 Effects of β -phenyl-DL-serine and L-threonine on absorption (A) and linear dichroism (B) spectra of gel-oriented tryptophanase in 100 mM potassium phosphate buffer, pH 7.8. Curve 1, unliganded enzyme. Curve 2, same as 1 + 165 mM L-threonine. Curve 3, same as 1 + 15 mM β -phenyl-DL-serine (*threo* form). (Reproduced with permission from *Enzymes Dependent on Pyridoxal Phosphate and Other Carbonyl Compounds as Cofactors*, p. 282, Pergamon, N. Y. (1991)).

greatly reduced upon formation of the external aldimine, and its sign is inverted upon formation of the quinonoid intermediate. In the case of tryptophanase the sign of induced optical activity, *i.e.* of the coenzyme's CD, is inverted upon formation of the external aldimine, and the dissymmetry factor is greatly diminished in the quinonoid band (*see* Section 9.3). Torchinsky, Sakharova and Zakomirdina^{38,44} have inferred from these data that reorientations of the coenzyme also occur in the course of the catalytic action of tryptophanase. This inference was supported by measurements of linear dichroism (LD) spectra of tryptophanase oriented in a compressed slab of polyacrylamide gel by the method of Rosenberg, Makarov and Torchinsky.⁶⁷⁾

It has been found⁶⁸⁻⁷⁰⁾ that the reaction of tryptophanase with the inhibitory amino acids, β -phenyl-DL-serine (*threo* form), L-threonine and D-alanine, is accompanied by a manifold increase in the reduced LD, *i.e.* in the ratio of LD to absorbance ($\Delta A/A$) in the 420–425 nm band (Fig. 9.12; Table 9.2). This band belongs to the protonated internal PLP-lysine

aldimine in the free enzyme, and was attributed to the external aldimine in the enzyme-quasisubstrate complexes on the basis of the CD data (Section 9.3.1).

The value of reduced LD depends on the degree of orientation (G) of the enzyme molecules in the gel slab and angle (α) between the direction of the transition dipole moment in the chromophore ring and the axis of orientation of the enzyme molecules in the gel^[67,71]:

$$\Delta A/A = 3/2(3 \cos^2\alpha - 1)G \quad (9.3)$$

Therefore, the difference in reduced LD values in the 420–425-nm band between the internal and external aldimines of tryptophanase can be explained either by reorientation of the PLP ring, or by a change in the direction of the transition dipole moment in the ring. The latter assumption seems unlikely. The protonated internal and external PLP-aldimines have a very similar electronic structure, and, therefore, their dipole moment directions should be very close or identical; that they are indeed so have been confirmed by polarized-light spectroscopy of single crystals of aspartate transaminase.^[72,73] Thus, it was concluded that the large difference in LD values in the 420 nm band between the free tryptophanase and its quasisubstrate complexes is caused by reorientation of the coenzyme ring in the transamination step.^[68–70]

The reaction of tryptophanase with α -methyl-DL-tryptophan, like that with β -phenyl-DL-serine, L-threonine and D-alanine, results in a decrease in absorbance at 337 nm and an increase at 420 nm. However, in contrast to those amino acids, α -methyltryptophan does not cause any change of the reduced LD or the ratio of CD to absorbance in the 420-nm band.^[69,70] Thus, this substrate analog does not form an external aldimine; its binding only shifts the pK value of the internal PLP-lysine aldimine. It follows that the noncovalent binding of substrate analogs does not influence the degree of orientation of tryptophanase molecules in a compressed gel slab. This supports the conclusion that the changes of LD described above are caused by reorientation of the coenzyme ring which occurs upon formation of the external aldimine.

The value of reduced LD in the quinonoid band of the enzyme complexes with L-alanine and oxindolyl-L-alanine proved to be close to that in the absorption band of the external aldimine formed with L-threonine (Table 9.2). It does not follow from this coincidence that the external aldimine and quinonoid have a similar orientation in the active site. The two intermediates, in fact, differ in their orientation; this is because the directions of the transition dipole moment in the pyridine ring of the aldimine and quinonoid are very different.^[72]

The LD and CD data are consistent with the results of stereochemical analysis reported by Floss and Vederas.^[74,75] They reduced the aldimine bond in the active site of tryptophanase with sodium borohydride and found that reduction proceeds predominantly from the *re* face at the C-4' atom of PLP in the free enzyme, and from the opposite side (*si* face) in the enzyme complex with L-alanine. Similar stereochemical findings have been reported for aspartate transaminase and a number of other PLP-dependent enzymes.^[75] The change of face of the imine double bond accessible to NaBH₄ can be explained either by reorientation of the coenzyme ring, or by rotation of the imine bond around the C-4–C-4' bond during substrate binding and transamination. The former explanation proved to be true for aspartate transaminase^[76] and, judging from the LD data, also holds for tryptophanase. The change of the ³¹P-NMR signal of tryptophanase upon

formation of the quinonoid intermediate⁴⁸⁾ may also result from reorientation of the coenzyme ring.

9.5.3 Catalytic Mechanism

General aspects of the stereochemistry and mechanism of PLP-catalyzed reactions have been reviewed by Floss and Vederas⁷⁵⁾ and Miles.⁷⁷⁾ In this section we briefly describe the catalytic cycle of tryptophanase. New transient intermediates have recently been detected in this cycle by Phillips *et al.*,^{41,42,78)} using rapid-scanning stopped-flow spectrophotometry, and they are included in the reaction mechanism depicted in Fig. 9.13.

In the active site of the enzyme PLP forms an internal aldimine (Schiff base) with Lys-270 (Fig. 9.13, I). When the substrate is bound at the active site, its α -amino group attacks the C-4' atom of the coenzyme and replaces the ϵ -amino group of Lys-270 from its bond with PLP. This transaldimination reaction probably proceeds via a tetrahedral intermediate (a *gem*-diamine). Spectral evidence for formation of a *gem*-diamine in this step has recently been obtained in studies of the reaction of tryptophanase with L-homophenylalanine.⁴¹⁾ The *gem*-diamine is subsequently converted to an external, PLP-substrate aldimine, and the ϵ -amino group of Lys-270 is released (Fig. 9.13, II). The equilibrium constant of this step with L-tryptophan is determined to be 11.6 mM.⁷⁸⁾

The α -proton of the external aldimine is then abstracted by a basic residue B₁ and the quinonoid intermediate is formed (Fig. 9.13, III). The rate constants of formation of the quinonoid intermediate from L-tryptophan and *S*-benzyl-L-cysteine are estimated to be 940 and 47.6 sec⁻¹, respectively; this demonstrates a drastic effect of the leaving group on labilization of the α -proton.⁷⁸⁾

The next step is tautomerization of indole to the reactive indolenine (Fig. 9.13, III $\longrightarrow$ IV). In this step the acid catalyst B₁-H presumably comes into proximity of the indole ring and donates its proton, while another basic group, B₂, abstracts the ring nitrogen proton of L-tryptophan. Phillips⁷⁸⁾ suggested that tautomerization of indole occurs via a concerted "push-pull" general acid-base catalysis. This suggestion is consistent with the results of studies on the steady-state pH-dependence of the reaction with L-tryptophan,⁷⁹⁾ and the intramolecular proton transfer demonstrated by Vederas *et al.*⁸⁰⁾ The latter showed that degradation of α -[²H]-L-tryptophan by tryptophanase produces indole enriched in deuterium in the 3-position. The proton transfer and tautomerization of indole may be the rate-limiting steps in the catalytic cycle.⁷⁸⁾

Elimination of indole from the second quinonoid intermediate (Fig. 9.13, IV) results in formation of the aminoacrylate Schiff base intermediate (Fig. 9.13, V). The latter may either decompose to the final products completing the reaction of α,β -elimination or may add a new reactive group (R-H, where R is a -SH, -SCH₃, -OH, or -OCH₃ group) at C-3, resulting in the β -replacement reaction (V $\longrightarrow$ IV $\longrightarrow$ III $\longrightarrow$ II $\longrightarrow$ I). Recent findings⁴²⁾ suggest that conversion of intermediate V into products and restoration of the internal PLP-lysine aldimine occur via a transient *gem*-diamine of α -aminoacrylate and the ϵ -amino group of Lys-270 (Fig. 9.13, VI). Iminopyruvate is then released from the *gem*-diamine and decomposes to pyruvate and ammonia.

Experiments on solvent deuterium isotope effect with L-tryptophan and *S*-methyl-L-cysteine indicated that the proton abstraction (II $\longrightarrow$ III) and breakdown of the aminoacrylate (V $\longrightarrow$ I) steps are not solvent deuterium-sensitive.^{79,81)} The solvent deuterium-sensitive step(s) in the mechanism appears to be formation of the tetrahedral

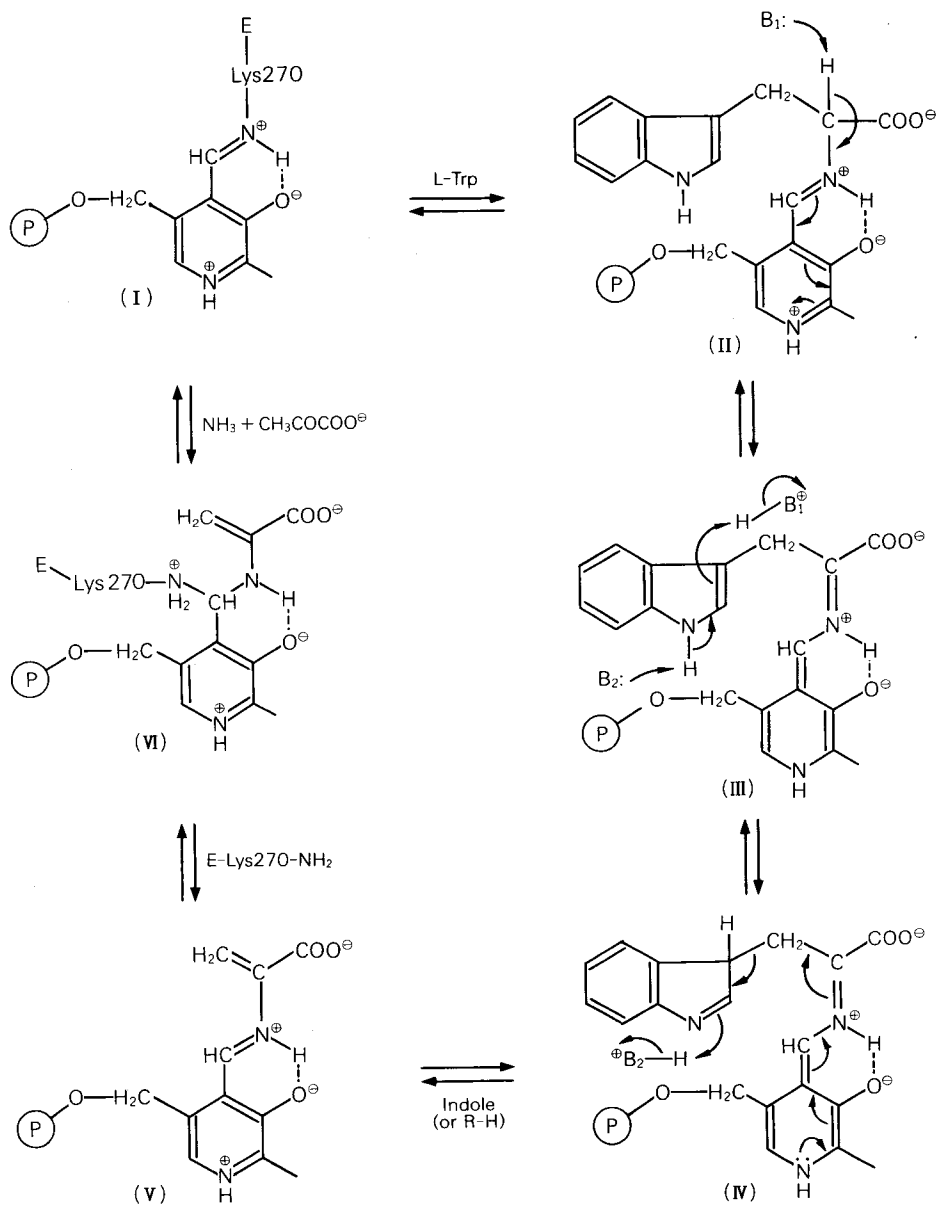


Fig. 9.13 Catalytic mechanism of tryptophanase. See text for details.

intermediate or the elimination of indole.

Analysis of the pH-dependence of the kinetic parameters and competitive inhibitor K_i values indicates that the pK_a values of the B_1 and B_2 groups are 7.6 and 6.0, respectively.⁷⁹⁾ These groups have not been identified, but the observed low degree of deuterium transfer ($\sim 8\%$)^{80,82)} suggests that the basic group B_1 is polyprotic and, therefore, may be the ε -amino group of a lysine residue.

The three-dimensional structure of a closely related enzyme, tyrosine phenol-lyase has recently been determined;^{28a)} its active site proved similar to that of aspartate transaminase from *E. coli*. Computer modeling of tryptophanase based on the X-ray coordinates of tyrosine phenol-lyase and homology of the primary structures of the two lyases suggested that the PLP-binding site is nearly identical in all three PLP-enzymes (R.S. Phillips, personal communication). This similarity makes it even more intriguing to elucidate the molecular mechanisms which direct different PLP-dependent enzymatic reactions along diverse specific pathways. The detailed comparative structural and functional studies of various PLP-enzymes including tryptophanase will shed light on this fundamental problem of enzymology.

REFERENCES

1. W.A. Newton, Y. Morino, E.E. Snell, *J. Biol. Chem.*, **240**, 1211 (1965).
2. W.A. Newton, E.E. Snell, *Proc. Natl. Acad. Sci. USA*, **51**, 382 (1964).
3. C. Yanofsky, E.S. Lennox, *Virology*, **8**, 425 (1959).
4. C. Yanofsky, I.P. Crawford, *Proc. Nat. Acad. Sci. USA*, **45**, 1016 (1959).
5. E.E. Snell, *Adv. Enzymol.*, **42**, 287 (1975).
6. J. London, M.E. Goldberg, *J. Biol. Chem.*, **247**, 1566 (1972).
7. J.A. Hoch, E.J. Sempron, R.D. DeMoss, *Biochemistry*, **5**, 2229 (1966).
8. J.L. Cowell, R.D. DeMoss, *J. Biol. Chem.*, **248**, 6262 (1973).
9. J.L. Cowell, K. Moser, R.D. DeMoss, *Biochim. Biophys. Acta.*, **315**, 449 (1973).
10. H. Yoshida, H. Kumagai, H. Yamada, *Agric. Biol. Chem.*, **38**, 463 (1974).
11. H. Yoshida, H. Kumagai, H. Yamada, H. Mastubara, *Biochim. Biophys. Acta*, **391**, 494 (1975).
12. C. Simard, A. Mardini, L.M. Bordeleau, *Can. J. Microbiol.*, **21**, 828 (1975).
13. C. Simard, A. Mardini, L.M. Bordeleau, *Can. J. Microbiol.*, **21**, 834 (1975).
14. C. Simard, A. Mardini, L.M. Bordeleau, *Can. J. Microbiol.*, **21**, 841 (1975).
15. M.C. Deeley, C. Yanofsky, *J. Bacteriol.*, **147**, 787 (1981).
16. M. Tokushige, N. Tsujimoto, T. Oda, T. Honda, N. Yumoto, S. Ito, M. Yamamoto, E.-H. Kim, Y. Hiragi, *Biochimie*, **71**, 711 (1989).
17. H. Kagamiyama, H. Matsubara, E.E. Snell, *J. Biol. Chem.*, **247**, 1576 (1972).
18. Y. Kawata, N. Tsujimoto, S. Tani, T. Mizobata, M. Tokushige, *Biochem. Biophys. Res. Commun.*, **173**, 756 (1990).
19. N. J. Greenfield, G. D. Fasman, *Biochemistry*, **8**, 4108 (1969).
20. S. Tani, N. Tsujimoto, Y. Kawata, M. Tokushige, *Biotechnol. Appl. Biochem.*, **12**, 28 (1990).
21. Y. Kawata, S. Tani, M. Sato, Y. Katsube, M. Tokushige, *FEBS Lett.*, **284**, 270 (1991).
22. O. Raibaud, M.E. Goldberg, *J. Biol. Chem.*, **251**, 2814 (1976).
23. Y. Morino, E.E. Snell, *J. Biol. Chem.*, **242**, 5602 (1967).
24. Y. Morino, E.E. Snell, *J. Biol. Chem.*, **242**, 5591 (1967).
25. O. Raibaud, M.E. Goldberg, *J. Biol. Chem.*, **251**, 2820 (1976).
26. C. Skrzynia, J. London, M.E. Goldberg, *J. Biol. Chem.*, **249**, 2325 (1974).
27. H.C. Isom, R.D. DeMoss, *Biochemistry*, **14**, 4298 (1975).
28. T. Honda, M. Tokushige, *J. Biochem.*, **100**, 679 (1986).
- 28a. A.A. Antson, T.V. Demidkina, P. Gollnick, Z. Dauter, R. L. Von Tersch, J. Long, S.N. Berezhnoy, R.S. Phillips, E.H. Harutyunyan, K.S. Wilson, *Biochemistry*, **32**, 4195 (1993).
29. J. London, C. Skrzynia, M.E. Goldberg, *Eur. J. Biochem.*, **47**, 409 (1974).
30. T. Mizobata, N. Yumoto, Y. Kawata, *J. Prot. Chem.*, **11**, 413 (1992).

31. T. Mizobata, Y. Akiyama, K. Ito, N. Yumoto, Y. Kawata, *J. Biol. Chem.*, **267**, 17773 (1992).
32. Y. Morino, E.E. Snell, *J. Biol. Chem.*, **242**, 2800 (1967).
33. R.S. Phillips, P. D. Gollnick, *J. Biol. Chem.*, **264**, 10627 (1989).
34. R.G. Kallen, T. Korpela, A.E. Martell, Y. Matsushima, C. M. Metzler, D.E. Metzler, Yu.V. Morozov, I.M. Ralston, F.A. Savin, Yu.M. Torchinsky, H. Ueno, in: *Transaminases*, (P. Christen, D.E. Metzler, eds.) John Wiley & Sons, New York (1985) p. 70.
35. D.S. June, C.H. Suelter, J.L. Dye, *Biochemistry*, **20**, 2707 (1981).
36. C.M. Metzler, R. Viswanath, D.E. Metzler, *J. Biol. Chem.*, **266**, 9374 (1991).
37. C.H. Suelter, E.E. Snell, *J. Biol. Chem.*, **252**, 1852 (1977).
38. Yu.M. Torchinsky, I.S. Sakharova, L.N. Zakomirdina, in: *Biochemistry of Vitamin B₆* (T. Korpela, P. Christen, eds.) Birkhauser Verlag, Basel (1987) p. 233.
39. N.G. Faleev, S.B. Ruvinov, L.N. Zakomirdina, I.S. Sakharova, Yu.M. Torchinsky, V.M. Belikov, *Molek. Biol. (Moscow)*, **25**, 752 (1991).
40. T. Watanabe, E.E. Snell, *J. Biochem. (Tokyo)*, **82**, 733 (1977).
41. R.S. Phillips, S.L. Bender, P. Brzovic, M.F. Dunn, *Biochemistry*, **29**, 8608 (1990).
42. R.S. Phillips, *Biochemistry*, **30**, 5927 (1991).
43. J.D. Fenske, R.D. DeMoss, *J. Biol. Chem.*, **250**, 7554 (1975).
44. L.N. Zakomirdina, I.S. Sakharova, Yu.M. Torchinsky, *Biokhimie*, **71**, 545 (1989).
45. D.S. June, C.H. Suelter, J.L. Dye, *Biochemistry*, **20**, 2714 (1981).
46. H.C. Isom, R.D. DeMoss, *Biochemistry*, **14**, 4291 (1975).
47. M. Tokushige, K. Iimura, M. Yamamoto, Y. Nishijima, *Biochem. Biophys. Res. Commun.*, **96**, 863 (1980).
48. K.D. Schnackerz, E.E. Snell, *J. Biol. Chem.*, **258**, 4839 (1983).
49. C.M. Metzler, D.E. Metzler, A. Kintanar, R.D. Scott, M. Marceau, *Biochem. Biophys. Res. Commun.*, **178**, 385 (1991).
50. H. Kagamiyama, Y. Morino, E.E. Snell, *J. Biol. Chem.*, **245**, 2819 (1970).
51. R.S. Phillips, I. Richter, P. Gollnick, P. Brzovic, M.F. Dunn, *J. Biol. Chem.*, **266**, 18642 (1991).
52. M.N. Kazarinoff, E.E. Snell, *J. Biol. Chem.*, **252**, 7598 (1977).
53. T. Nihira, T. Toraya, S. Fukui, *Eur. J. Biochem.*, **101**, 341 (1979).
54. T. Nihira, T. Toraya, S. Fukui, *Eur. J. Biochem.*, **119**, 273 (1981).
55. T. Kakizono, T. Nihira, H. Taguchi, *Biochem. Biophys. Res. Commun.*, **137**, 964 (1986).
56. T. Oda, M. Tokushige, *J. Biochem. (Tokyo)*, **104**, 178 (1988).
57. M. Tokushige, Y. Fukada, Y. Watanabe, *Biochem. Biophys. Res. Commun.*, **86**, 976 (1979).
58. R.S. Phillips, P. Gollnick, *FEBS Letters*, **268**, 213 (1990).
59. A. Högberg-Raibaud, O. Raibaud, M.E. Goldberg, *J. Biol. Chem.*, **250**, 3352 (1975).
60. O. Raibaud, M.E. Goldberg, *Eur. J. Biochem.*, **73**, 591 (1977).
61. T. Honda, M. Tokushige, *J. Biochem. (Tokyo)*, **97**, 851 (1985).
62. T. Nihira, T. Yasuda, T. Kakizono, H. Taguchi, M. Ichikawa, T. Toraya, S. Fukui, *Eur. J. Biochem.*, **149**, 129 (1985).
63. T. Honda, M. Tokushige, *J. Chromatog.*, **371**, 353 (1986).
64. V.I. Ivanov, M.Ya. Karpeisky, *Adv. Enzymol.*, **32**, 21 (1969).
65. Yu.M. Torchinsky, L.G. Koreneva, *Biokhimiya*, **28**, 1087 (1963).
66. Yu.M. Torchinsky, L.G. Koreneva, *Biokhim. Biophys. Acta*, **79**, 426 (1964).
67. M.V. Rosenberg, V.L. Makarov, Yu.M. Torchinsky, *Molek. Biol. (Moscow)*, **19**, 1669 (1985). English translation by Plenum Publ. Co., New York (1986) p. 1360.
68. Yu.M. Torchinsky, M.V. Rosenberg, L.N. Zakomirdina, I.S. Sakharova, *Annals. New York Acad. Sci.*, **585**, 68 (1990).
69. L.N. Zakomirdina, I.S. Sakharova, Yu.M. Torchinsky, *Eur. J. Biochem.*, **193**, 243 (1990).
70. Yu.M. Torchinsky, L.N. Zakomirdina, I.S. Sakharova, in: *Enzymes Dependent on Pyridoxal Phosphate and Other Carbonyl Compounds as Cofactors* (T. Fukui, H. Kagamiyama, K. Soda, H. Wada, eds.) Pergamon Press, Oxford (1991) p. 281.
71. B. Nordén, *Appl. Spectr. Rev.*, **14**, 157 (1978).
72. M.G. Vincent, D. Picot, G. Eichele, J.N. Jansonius, H. Kirsten, P. Christen, in: *Chem. and Biol. Aspects of Vitamin B₆ Catalysis* (A.E. Evangelopoulos, ed.) Alan R. Liss, New York (1984) part B, p. 233.
73. C.M. Metzler, J. Mitra, D.E. Metzler, M.W. Makinen, C.C. Hyde, P.H. Rogers, A. Arnone, *J. Mol. Biol.*, **203**, 197 (1988).
74. J.C. Vederas, H.G. Floss, *Acc. Chem. Res.*, **13**, 455 (1980).
75. H.G. Floss, J.C. Vederas, in: *Stereochemistry* (C. Tamm, ed.) Elsevier, Amsterdam (1982) p. 161.
76. A. Arnone, P. Christen, J.N. Jansonius, D.E. Metzler, in: *Transaminases* (P. Christen, D.E. Metzler, eds.) John Wiley & Sons, New York (1985) p. 326.
77. E.W. Miles, *Coenzymes Cofactors*, **1**, 253 (1986).
78. R.S. Phillips, *J. Am. Chem. Soc.*, **111**, 727 (1989).

79. D.M. Küick, R.S. Phillips, *Biochemistry*, **27**, 7339 (1988).
80. J.C. Vederas, E. Schleicher, M.-D. Tsai, H.G. Floss, *J. Biol. Chem.*, **253**, 5350 (1978).
81. D.M. Küick, *J. Am. Chem. Soc.*, **113**, 8499 (1991).
82. E. Schleicher, K. Mascara, R. Potts, D.R. Mann, H.G. Floss, *J. Am. Chem. Soc.*, **98**, 1043 (1976).

10

Superoxide Dismutase

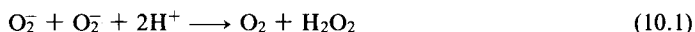
Sumio Kanematsu and Kozi Asada*

*Department of Food Science and Technology, Minami-Kyushu University, Takanabe, Miyazaki 884, Japan*** The Research Institute for Food Science, Kyoto University, Uji, Kyoto 611, Japan*

10.1 Introduction

Utilization of dioxygen in organisms as electron acceptors and oxygenase substrates and of water as electron donors inevitably produces reactive species of oxygen. These include redox intermediates in the reduction and oxidation between dioxygen and water; superoxide anion radical (O_2^-), hydrogen peroxide (H_2O_2) and hydroxyl radical ($HO\cdot$), and the electronically-excited oxygen species, singlet excited oxygen ($^1O_2$). Many lines of evidence have shown that these activated oxygens cause cell damage unless they are immediately scavenged at the site where they are generated. Thus, organisms acquired defense systems against activated oxygens upon the utilizations of dioxygen and water in the course of evolution.¹⁾

Superoxide dismutase (SOD) is a key enzyme which constitutes the first line of defense against oxygen toxicity and catalyzes the disproportionation of the superoxide anions to dioxygen and hydrogen peroxide²⁾:



By lowering the steady state concentration of superoxide, SOD protects cells from harmful effects of O_2^- itself and of other reactive oxygen species derived from O_2^- .³⁻⁵⁾

Three distinct types of SOD based on their prosthetic metals have been found: CuZn-, Mn- and Fe-containing enzymes, which fall into two evolutionary families. Mn-SOD and Fe-SOD show a high degree of homology in their amino acid sequences and similarity in their three-dimensional structures, and are considered to have evolved from the ancestor Fe-SOD. In contrast Fe- and Mn-SODs show little similarity in their amino acid sequences and three-dimensional structures to those of CuZn-SOD. Despite the high degree of homology in amino acid sequence between Fe- and Mn-SODs, there exists specificity of metal binding; Fe-substituted Mn-SOD and Mn-substituted Fe-SOD have no SOD activity with few exceptions.⁶⁻⁹⁾ Thus, the microenvironments of their metal binding sites differ to some extent from each other. Distribution studies of the three types of SOD among wide varieties of organisms at different evolutionary stages showed the hierarchical appearance of enzymes along the phylogenic evolution of organisms.¹⁾

Nevertheless, with respect to the catalysis of the dismutation of O_2^- , little difference has been found among the three types of SOD.¹⁰⁾ The reaction rate constants between O_2^- and the enzyme are about $2 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ at neutral pH. However, although the catalytic rate of the CuZn-SOD is constant between pH 5 and 9.5, those of the Fe- and Mn-SODs become progressively lower as the pH is raised above 8.5.

Structural information on the three types of SOD have revealed structure-function relationships. In this chapter, we describe the phylogenic distribution, gene regulation and general properties of the three types of SOD. The reaction mechanism of SOD is also discussed based on its structure.

10.2 Distribution of SOD among Organisms

The distribution of the three types of SOD among organisms is characteristic of the evolutionary stages of organisms, and of cell organelles. This characteristic distribution reflects the evolutionary history of the acquisition of defense against oxygen toxicity due to increase in atmospheric oxygen in the biosphere by photosynthetic organisms.

10.2.1 Phylogenic Distribution

Figure 10.1 shows the general distribution of the three types of SOD among organisms at different stages of evolution.^{1,11-15)} Fe-SOD has been found in prokaryotes including anaerobic bacteria,^{16,17)} protozoa, eukaryotic algae,¹⁸⁾ and in limited species of plants,¹⁹⁾ but is absent from animals and fungi, indicating the earliest appearance among the three types of SOD. Mn-SOD occurs in aerobic bacteria but not in anaerobic bacteria, and also localizes in mitochondria matrix of eukaryotes²⁰⁾ and chloroplast thylakoids of plants. Since Mn-SOD is a homologous protein with Fe-SOD in their amino acid sequences, it would be derived from Fe-SOD and acquired at the time when nonsulfur photosynthetic bacteria appeared on the earth.¹⁾ To the contrary, CuZn-SOD has not been detected in prokaryotes with several exceptions,²¹⁾ most eukaryotic algae and protozoa, but it has been found in eukaryotes^{1,11)} including non-vertebrates, vertebrates, fungi, most evolved eukaryotic algae, mosses and ferns, and seed plants, suggesting its latest acquisition among SODs.

10.2.2 Subcellular Localization

The major sites of O_2^- production in eukaryotic cells are chloroplasts and mitochondria, where the evolution of dioxygen and the reduction of dioxygen proceed at a high flux rate. Due to the low permeability of O_2^- toward membranes,²²⁾ O_2^- must be removed effectively at the site of its production. The production and scavenging of active species of oxygen in chloroplasts have been well elucidated.⁵⁾

Chloroplast type of CuZn-SOD isozyme is localized in the chloroplast stroma of most plants, whereas Fe-SOD occurs in chloroplasts of *Euglena*, the moss *Marchantia polymorpha* and several species of seed plant. So far Mn-SOD has not been detected in chloroplasts in a soluble form but occurs in a thylakoid membrane-bound form.^{23,24)} The cytosolic CuZn-SOD, which is distinguishable from the chloroplastic CuZn-SOD in terms of amino acid sequence, occurs in cell compartments other than chloroplasts and in nonphotosynthetic tissues.¹⁴⁾

In mitochondria of animals, plants and fungi, Mn-SOD is localized in the matrix. In

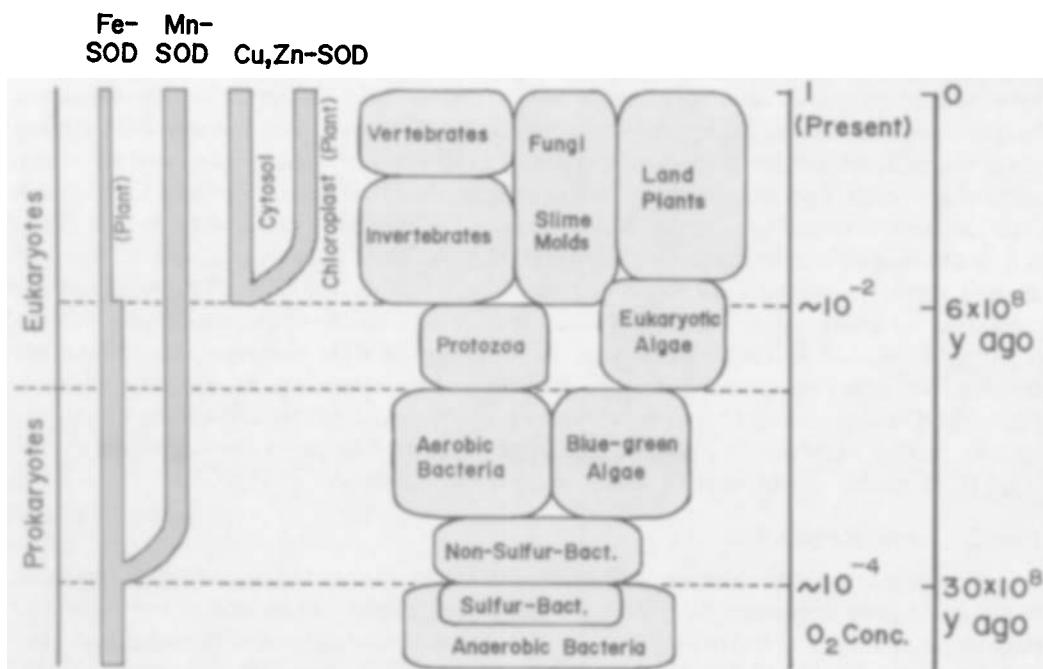


Fig. 10.1 Phylogenetic distribution of CuZn-, Mn- and Fe-SODs in organisms at various stages of evolution.¹⁾ O₂ conc. shows the concentrations of O₂ in the atmosphere, relative to the present concentration, when the indicated organisms appeared on the earth.

addition, CuZn-SOD is present in the intermembrane space, although its molecular properties are not characterized.²⁰⁾ CuZn-SOD and Mn-SOD are localized in peroxisomes (glyoxysomes).²⁵⁾

10.3 Physiological Functions

10.3.1 Defense against Oxygen Toxicity

The sole physiological function of SOD so far established is the defense against highly reactive oxygen species by scavenging O₂⁻. There is a close relationship between the viability of organisms under aerobic conditions and occurrence of SOD. Not only aerobes, but anaerobes also contain SOD, indicating that dioxygen even at ultra low concentrations is toxic to the cells. In exception, the lactate fermenting bacterium *Lactobacillus plantarum* lacks SOD. However, millimolar levels of Mn ions chelated to low molecular weight ligands are contained in this bacterium, and they can catalyze the dismutation of O₂⁻.^{26,27)} Although the disproportionation activity for O₂⁻ of the Mn-

chelates is only 1% that of the enzyme on a molar basis, high contents of Mn ions allow lowering the concentration of O_2^- as does SOD at micromolar concentrations.

The conditions leading to the production of active oxygen within cells, *i.e.* under hyperbolic conditions or methyl viologen-addition, induce the biosynthesis of Mn-SOD in bacteria and give the cells increased tolerance against active oxygen. However, bacteria manipulated in these ways to increase SOD content also induce other anti-oxidative enzymes such as catalase and peroxidase, indicating that it is difficult to correlate directly the acquired tolerance against active oxygen to SOD content only. Using a recombinant technology which appears to avoid these problems, the effect of dosage or lack of SOD in cells has been determined. Scott, Meshnick and Eaton²⁸⁾ have demonstrated that *E. coli* cells harboring a plasmid expressing Fe-SOD is less resistant than that of the wild type cells against methyl viologen, and explained that the resulting H_2O_2 is responsible for less tolerance. The increased Mn-SOD also renders the cells less tolerant against oxygen toxicity.²⁹⁾ Hence, at least in some cases, introduction of SOD gene into cells perturbs the regulation of other proteins that may be involved in the defense against oxygen toxicity.³⁰⁾ CuZn-SOD expressed in *E. coli*, which lacks both Mn- and Fe-SODs, restores tolerance against methyl viologen.³¹⁾ Since CuZn-SOD does not occur in the wide type *E. coli* (Fig. 10.1), the dismutation of O_2^- seems to be the sole function of SOD.

10.3.2 Gene Regulation

Gene regulation has been extensively studied in *E. coli*, which contains Mn-SOD encoded by the *sodA* gene (located at 88 min on chromosome), Fe-SOD encoded by the *sodB* gene (located at 38 min) and the hybrid Fe/Mn-SOD that consists of one subunit each of Mn- and Fe-SOD.³²⁾ Fe-SOD is a constitutive enzyme and found even in anaerobically cultured cells. Mn-SOD, however, is induced by O_2 and redox-cycling compounds such as methyl viologen that increase the flux of O_2^- . In addition, even during anaerobic growth, chelators for ferrous ion,³³⁾ and nitrate³⁴⁾ that acts as an electron sink resulting in an increase of redox state, also induce the biosynthesis of Mn-SOD, indicating that the induction of Mn-SOD reflects the cellular redox state. Moody and Hassan³³⁾ have proposed that the *trans*-acting negative repressor protein containing iron is a sensor for the redox state (Fig. 10.2), *i.e.* the Fe^{2+} -repressor complex is active and represses the expression of Mn-SOD, while the Fe^{3+} -repressor complex which is formed upon an oxidative state is inactive and derepresses the expression. The binding site for a regulatory protein(s) at the -35 region of the *sodA* gene of *E. coli* has been identified by site-directed mutagenesis.³⁵⁾ This region is shown to be the potential binding sites for the regulatory proteins Fur (ferric uptake regulation), Arc (aerobic respiratory control), and Fnr (fumarate nitrate reduction/regulator of anaerobic respiration), which act as anaerobic repressors of *sodA*³⁶⁾ (Fig. 10.2).

The regulation of the *sodA* gene in *E. coli* is independent of the oxidative stress regulon (*oxyR*)³⁴⁾ that controls the biosynthesis of catalase and glutathione peroxidase, the inducible DNA repair system (SOS), and heat shock regulon (*hspR*).³⁷⁾ Recently, it has been reported that Mn-SOD is positively regulated by the superoxide response regulon (*soxRS*)³⁸⁾ (located at 92.2 min), which is not involved in the regulation of *sodA* by Fur, Arc and Fnr.³⁶⁾ Biosynthesis of CuZn-SOD also is induced by the conditions that lead to the production of superoxide in eukaryotes. However, its mechanism of gene regulation remains to be elucidated.

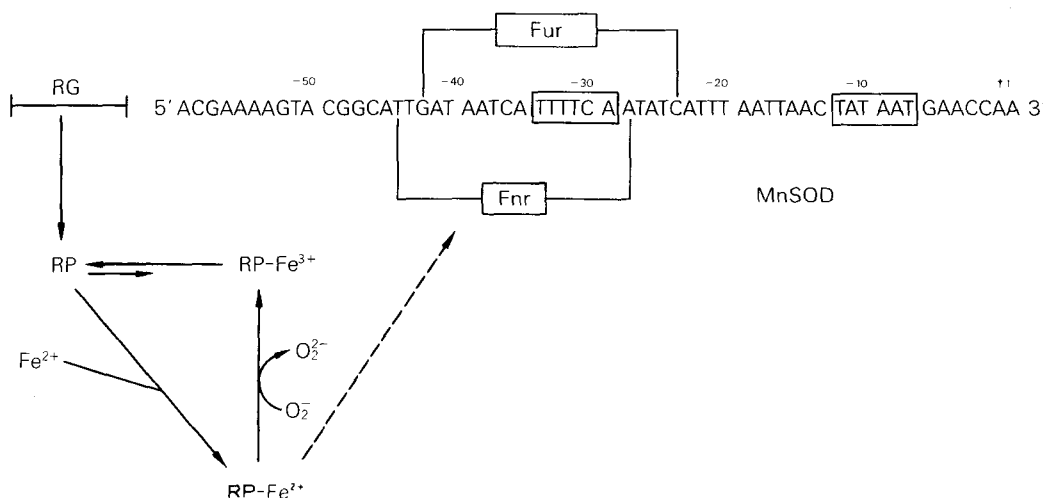


Fig. 10.2 Schematic model for the regulation of Mn-SOD in *E. coli*.^{33,36)} RG, regulatory gene; RP, apo-repressor protein (inactive); RP-Fe³⁺, ferric repressor (inactive); RP-Fe²⁺, ferrous repressor (active). Mn-SOD, nucleotide sequence of the 5' regulatory region of *sodA* (nucleotides -59 to +1 are shown). The +1 nucleotide designates the start point of transcription. The -35 and -10 (Pribnow box) regions for RNA polymerase binding are boxed. Fur-, and Fnr-binding sites are bracketed.

10.4 Molecular Properties

10.4.1 CuZn-SOD

A. Physicochemical Properties³⁹⁾

CuZn-SOD was the first enzyme found possessing catalytic activity of the disproportionation of O_2^- . It was isolated from bovine erythrocyte by McCord and Fridovich²⁾ in 1969. The molecular weight of the enzyme is 32,000, and it is a dimer constituted by two identical subunits (16,000) without covalent linkage. Each subunit contains one atom of Cu and one atom of Zn, and the Cu participates in the catalytic reaction and the Zn in the stabilization of the enzyme structure. The active site copper is ligated by four histidine residues, and Zn is ligated by three histidine and one aspartic acid residues. One histidine residue is a common ligand to the Cu and the Zn. Tryptophan and tyrosine residues lack or are low in CuZn-SOD, thus the absorption spectra of the enzyme give an atypical ultraviolet spectrum showing a triplet at 260 nm due to the phenylalanine residue, unlike most proteins. In the visible range the Cu gives a small absorption at around 680 nm with an absorption coefficient of $300 \text{ M}^{-1} \text{ cm}^{-1}$, which disappears upon the addition of reductants.

The CuZn-SOD activity is completely inhibited by cyanide, and also partly by azide. The Cu^{2+} and Zn^{2+} ions can be substituted by other metal ions such as Co^{2+} , Ni^{2+} etc., but the activity is found only when those substituted enzymes retain Cu^{2+} in the native site (Table 10.1).

In addition to the intracellular CuZn-SOD, the extracellular CuZn-SOD (EC-SOD)

TABLE 10.1 General properties of CuZn-, Mn- and Fe-SODs

	CuZn-SOD	Mn-SOD	Fe-SOD
Molecular weight	32,000	44,000/88,000	42,000/84,000
Subunit structure and molecular weight	homodimer 16,000	homodimer/homotetramer 22,000	homodimer/homotetramer 21,000
Metal content (atom/subunit)	1 Cu and 1 Zn	1 Mn	1 Fe
Ligands of metals	Cu: 4 His Zn: 3 His and 1 Asp	3 His and 1 Asp	3 His and 1 Asp
$\lambda_{\max}$ (nm)	258, 680	280, ~470	280, ~350
Absorbance coefficient ($M^{-1} cm^{-1}$)	300 (680 nm)	900 (470 nm)	3,500 (350 nm)
EPR signal	$g_{\perp} = 2.04$, $g_{\parallel} = 2.24$	—	$g = 4.9, 3.9, 3.5 (9.2, 1.98)$
Amino acid composition	No or low Tyr and Trp	No bias	No bias
Amino acid sequence	No homology with Mn-/Fe-SOD	High homology with Fe-SOD	High homology with Mn-SOD
Specific activity (McCord-Fridovich unit/mg enzyme)	8,000	6,000	6,000
Rate constant with O_2^- ($M^{-1} s^{-1}$)	2×10^9	2×10^9	2×10^9
K_m for O_2^- (M)	3.5×10^{-3} (5.5°C, pH 9.3)	3×10^{-6} (25°C, pH 9.7)	8×10^{-5} (25°C, pH 8.4)
Catalytic center activity, k_{cat} (s^{-1})	1×10^6	1.3×10^3	2.6×10^4
E_m (V , at pH 7)	~0.28	0.26–0.31	0.23–0.27
Inhibition by cyanide	Yes	No	No
Other inhibitors	Azide, Halide anions, Diethyldithiocarbamate	Azide	Azide, Fluoride, Toluene 3,4-dithizone
Inactivation by H_2O_2	Yes	No	Yes

terminus of the enzyme from animals is acetylated, but those from others are not.

The three-dimensional structures have been determined for bovine (2.0 Å resolution),^{41,42} yeast (2.5 Å),⁴³ spinach (2.0 Å)⁴⁴ and human (2.5 Å)⁴⁵ CuZn-SODs. The structure of cobalt-substituted bovine CuCo-SOD (2.0 Å) has also been determined.⁴⁶ The overall structure of the enzyme is a barrel structure consisting of eight antiparallel β strands (46% of total amino acid residues) and three loops (48%). The two large loops are fixed by an internal disulfide bond and contain ligating amino acid residues to the Cu and the Zn. The content of α helix is only 5%, and it is localized in the loop region (Fig. 10.4).

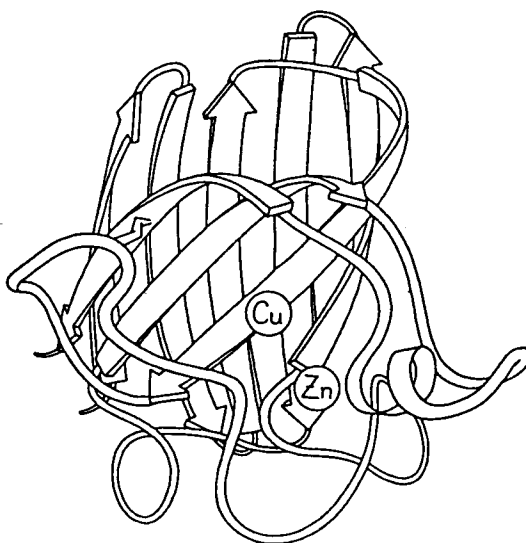


Fig. 10.4 Chain fold of bovine CuZn-SOD.⁴¹⁾ The β strands are shown by arrows and the disulfide bridge by a zig-zag. The Cu and Zn lie at the bottom of the active-site cavity.
(Reproduced with permission from J.A. Tainer *et al.*, *J. Mol. Biol.*, **160**, 189 (1982)).

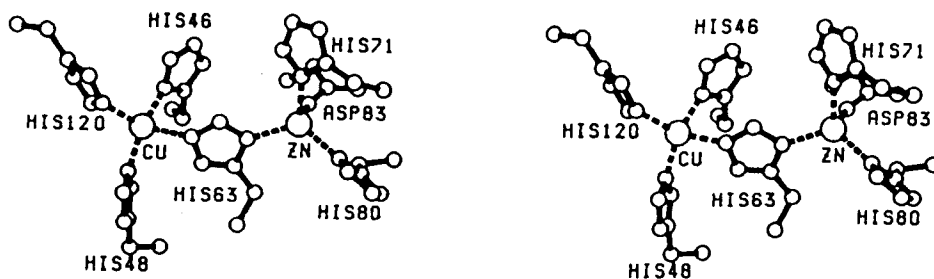


Fig. 10.5 Active site of spinach CuZn-SOD.⁴⁴⁾ The geometry of the Zn ligands is approximately tetrahedral. The Cu ligands form a distorted square plane.
(Reproduced with permission from Y. Kitagawa *et al.*, *J. Biochem.*, **109**, 477 (1991)).

10.4.2. Fe-SOD and Mn-SOD

A. Physicochemical properties

After the discovery of CuZn-SOD, Mn-SOD was isolated by Keele, McCord and Fridovich⁴⁷⁾ and Fe-SOD by Yost and Fridovich⁴⁸⁾ from *E. coli*. Mn- and Fe-SOD are homodimers or homotetramers, and the molecular weight of their subunits is 21,000–22,000. Most Fe-SODs from bacteria are dimers, while Mn-SODs are dimers or tetramers, and Mn-SODs from mitochondria are tetramers. Mn-SOD and Fe-SOD show absorption maxima at 470 nm and 340 nm, respectively, due to the *d-d* transition of the metals, which disappear upon reduction (Table 10.1). Ligand environments of Mn- and Fe-SODs are specific to the respective enzymes, and the SODs reconstituted with a metal different from the original one show no enzymatic activity. However the SODs from *Propionibacterium shermanii*,⁶⁾ *Bacteroides fragilis*,⁷⁾ *Bacteroides thetaiotaomicron*,⁸⁾ and

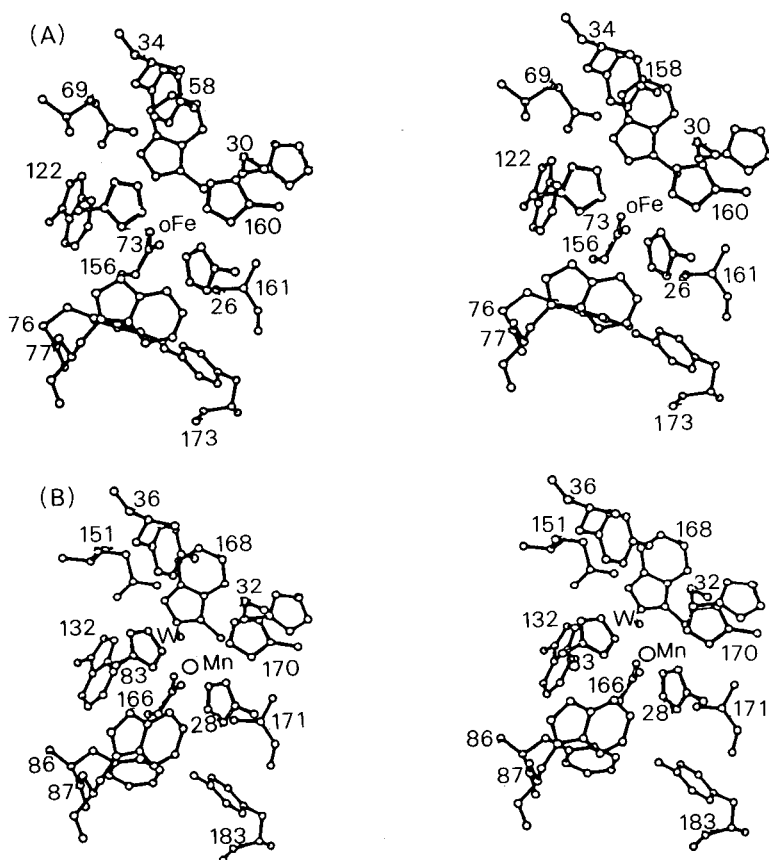


Fig. 10.7 Active centers of *E. coli* Fe-SOD (A) and *T. thermophilus* Mn-SOD (B).⁴⁹⁾ The metal ligands are shown in open bonds and several other residues in closed bonds. The amino acid residues of *T. thermophilus* Mn-SOD are identical with those found in *E. coli* Fe-SOD structure, and the two active sites can almost be superimposed, except for residues 86 and 151.

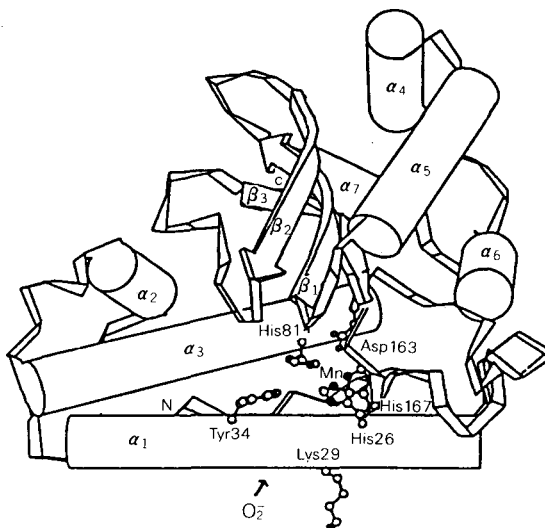


Fig. 10.8 Chain fold of *Bacillus stearothermophilus*.⁵¹⁾ α -Helices are designated by α 1 to 7 and β -strands β 1 to 3. The location of the Mn ion and its ligands are shown.
(Reproduced with permission from M.W. Parlier and C.C.F. Blake, *J. Mol. Biol.*, **199**, 657 (1988)).

*Streptococcus mutants*⁹⁾ are exceptions in that they can be reactivated from the respective apoenzymes by either Mn or Fe ions.

B. Primary and Three-Dimensional Structures

Figure 10.6 shows the alignment of amino acid sequences of Fe- and Mn-SODs so far determined. The amino acid sequences of Fe-SODs show a high degree of homology with those of Mn-SODs. His-26, His-73, Asp-156, and His-160⁴⁹⁾ in *E. coli* Fe-SOD serve as the metal ligands (Fig. 10.7) and are conserved in both Mn- and Fe-SODs from other organisms (Fig. 10.6). The three-dimensional structures have been determined for Mn-SODs from *Thermus thermophilus* HB8 (2.4 Å)⁵⁰⁾ and *Bacillus stearothermophilus* (2.4 Å)^{51,52)} and for Fe-SODs from *E. coli* (3.1 Å)⁵³⁾ and *Pseudomonas ovalis* (2.9 Å).⁵⁴⁾ The polypeptide chain of the monomer of Mn-SOD is composed of two domains; one has an all- α structure and the other an α/β structure, with the Mn ion bound between them (Fig. 10.8).

The Mn ion in *B. stearothermophilus* Mn-SOD is ligated by His-26, His-81, Asp-163, and His-167, with a tetrahedral arrangement strongly distorted toward trigonal pyramidal. Tyr-34, whose hydroxy group is only 5 Å apart from the metal, is involved in the catalytic reaction, and the Mn-site is particularly rich in aromatic amino acid residues (Fig. 10.8).

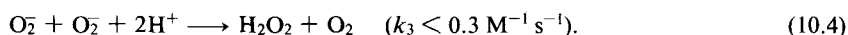
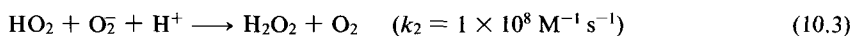
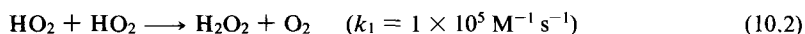
10.5 Properties of Superoxide Radical and Analytical Methods

The substrate of SOD, O_2^- , has unique properties as a reactive and unstable radical, as compared to stable substrates for most enzymes. O_2^- is spontaneously disproportionated at an appreciable rate, and its life time is only 5 s when produced at $1 \mu M s^{-1}$ at pH 7. SOD catalyzes the disproportionation of O_2^- at a diffusion-controlled rate, and to study such a

fast reaction, special methods are necessary.

10.5.1. Superoxide Anion Radical as a Substrate

Since O_2^- is a conjugate base of weak acid, HO_2 ($pK_a=4.9$), and also an anion radical at physiological pH, it can perform both ionic and radical reactions. As an anion O_2^- has reactivities such as nucleophilic substitution and addition reactions, and as a radical, it performs hydrogen-withdrawal reactions, one-electron reduction or oxidation reactions and the disproportionation reaction. In aqueous medium, the radical reactions are significant because of the short life of O_2^- to perform ionic reactions. Spontaneous disproportionation rate is largely pH dependent⁵⁵:



Thus, one O_2^- barely reacts with the other, and the decay of most superoxide proceeds through the interaction between HO_2 and O_2^- . In aprotic environments such as membranes, the life time of O_2^- is prolonged due to the unavailability of protons.

10.5.2 Diffusion-Controlled Reaction⁵⁶

The initial step of chemical reactions is an encounter of reactants by diffusion, and the subsequent reactions proceed to give products from the activated complex. The diffusion energy in solution is 15 kJ/mol, while many chemical reactions need an activation energy of 40 kJ–100 kJ/mol. If the activation energy of the reaction is low enough compared to the diffusion energy, then the diffusion determines the overall reaction, which has been referred to as a diffusion-controlled or -limited reaction. From Debye's equation on the diffusion-limited bimolecule reaction, the maximum value for the second-order reaction rate constant is estimated to be 10^9 – $10^{10} \text{ M}^{-1} \text{ s}^{-1}$ (25°C). The fastest reaction in aqueous solution is that of oxonium and OH^- ions at a rate constant of $1.4 \times 10^{11} \text{ M}^{-1} \text{ s}^{-1}$ (25°C): $H_3O^+ + OH^- \longrightarrow 2H_2O$. The SOD-catalyzed disproportionation of O_2^- is the fastest enzymatic reaction, and proceeds at a near diffusion-limited rate.

10.5.3 Assay of SOD

The SOD activity is directly determined by following the decay of O_2^- or indirectly by competition between a monitoring molecule of O_2^- and SOD for O_2^- . Cytochrome *c*, nitroblue tetrazolium, tetranitromethane and epinephrine have been used as the monitoring molecules of O_2^- , and their reaction products are spectrophotometrically followed. Since the indirect methods do not require special apparatus, they have been widely used for the assay of SOD. For elucidation of the reaction mechanism, however, the direct methods using special apparatus are indispensable.

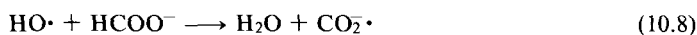
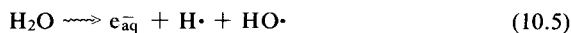
A. Xanthine-Xanthine Oxidase-Cytochrome *c* Method^{2,57}

In this indirect assay system, O_2^- is generated continuously by xanthine and xanthine oxidase, since O_2^- disproportionates spontaneously in the absence of a reactant for O_2^- . In

the presence of cytochrome *c* at a concentration that suppresses the spontaneous disproportionation of O_2^- , the O_2^- thus generated is followed by monitoring the reduction of cytochrome *c* by O_2^- at 550 nm. SOD activity is assayed by the inhibition of cytochrome *c* reduction by O_2^- . One unit of SOD is defined as the amount of enzyme required to inhibit the reduction of cytochrome *c* by 50% under the assay conditions. The enzymatic unit is equal to $(V/v - 1)$, where V and v represent the reduction rates in the absence and presence of SOD, respectively.⁵⁷⁾

B. Pulse Radiolysis Method⁵⁸⁾

Pulse radiolysis is a method that determines the reactivity of the unstable radical species produced by radiation using an electron pulse. When oxygenated water is irradiated with an electron pulse of 10^{-8} – 10^{-6} -s duration in the presence of formate, O_2^- is produced by the following equations.



Formate converts $HO\cdot$ to O_2^- via COO^- ; thus, all of the radiochemically produced radicals are converted to O_2^- . The reaction of O_2^- with reactants is determined by following spectrophotometrically the O_2^- or reaction products. This method gives an O_2^- concentration of 20–30 μM , less than that obtained by the stopped-flow method.

C. Flash Photometric Method⁵⁹⁾

The principle of this method is similar to that of the pulse radiolysis method. O_2^- is pulse-generated by a flash-light through the excitation of FMN in the presence of *N,N,N',N'*-tetramethylethylenediamine (TEMED) and dioxygen. O_2^- is produced within 10 ms by a transfer of electron to dioxygen from the electron donor TEMED via photochemically-excited FMN. The decay of O_2^- is followed by observing the absorbance change at 240 nm due to O_2^- directly, where the absorbance of FMN is minimal and that of TEMED is almost negligible (Fig. 10.9), or by observing the reduction of cytochrome *c* in a semi-direct fashion.⁵⁹⁾

D. Stopped-Flow Method

O_2^- is unstable in aqueous media, but stable in aprotic media. *Potassium* superoxide (KO_2) is only slightly soluble in dimethyl sulfoxide (DMSO), but can be dissolved at 0.15 M in DMSO with the aid of 0.3 M dicyclohexyl-18-crown-6.^{60,61)} KO_2 solution thus prepared can be stored at room temperature under dry argon. KO_2 solution is rapidly mixed with

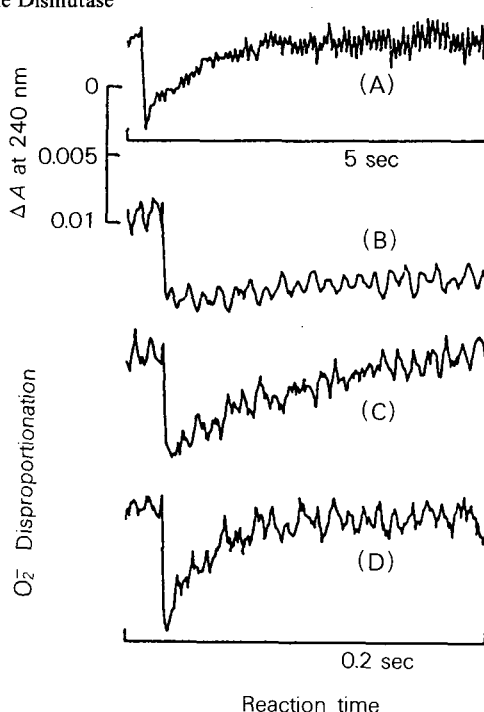


Fig. 10.9 Decay kinetics of O_2^- in the presence or absence of SOD by flash photometric method.⁵⁹⁾ Traces A and B show the decay of O_2^- monitored at 240 nm in the absence of SOD with different time scales. For traces C and D, 11.9 and 23.9 nM CuZn-SOD was added, respectively. (Reproduced from *J. Biochem.*, **91**, 895 (1982)).

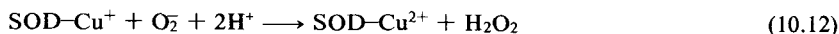
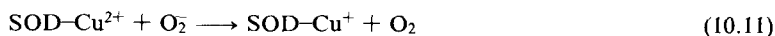
SOD or O_2^- -reactants in aqueous media, and the kinetics of the resulting reaction can be followed in a stopped-flow apparatus.^{62,63)} Bull, McClune and Fee⁶⁴⁾ have developed a specific flow system to mix rapidly water and DMSO and shown the saturation of O_2^- toward SOD.

10.6 Kinetics and Reaction Mechanisms

10.6.1 CuZn-SOD

A. Kinetics

The catalytic mechanism of CuZn-SOD involves alternate reduction and reoxidation of the Cu ion during successive interactions with O_2^- ⁶⁵⁾:



Each reaction is first order with respect to the substrate and enzyme, and proceeds with a second-order rate constant of $2 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ at 25°C over the pH range of 5–9.5, thus, the

overall reaction proceeds at near the diffusion-limited rate. Although the reverse reaction of the reaction (10.12) is unfavorable in such a diffusion-limited reaction, it has been demonstrated by the reduction of tetranitromethane with O_2^- generated by H_2O_2 .⁶⁶⁾

Under the condition of high O_2^- concentration, CuZn-SOD exhibits saturation behavior, indicating the presence of a rate-limiting first-order process in the reaction.⁶⁷⁾ The Michaelis-Menten parameters are estimated to be $3.5 \times 10^{-3} \text{ M}$ for K_m and $1 \times 10^6 \text{ s}^{-1}$ for k_{cat} . The rate constant for O_2^- binding to the active site (k_{cat}/K_m) is $2.6 \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$ (pH 9.3, 5.5°C). When O_2^- is added at saturated concentrations, the rate-limiting step of the reaction will be the first-order process, which has been suggested to be the proton transfer to the Cu-peroxo complex.

B. Facilitated Encounter of O_2^- with the Active Site by Electrostatic Guidance

The dismutation rate of O_2^- catalyzed by CuZn-SOD, $2 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$, is 10% of the rate of diffusion-limited reaction expected for a small molecule (O_2^-) and a macromolecule (SOD), whereas the Cu site comprises only 0.1% of the total surface area of CuZn-SOD. This indicates that the interaction of O_2^- with the Cu site is facilitated 100-fold, for which an electrostatic interaction is responsible.

Using a computer graphic analysis of the three-dimensional structure of CuZn-SOD, Getzoff *et al.*⁶⁸⁾ calculated the electrostatic vector field of the enzyme. The charged amino acid residues form a cluster of positive charges around the Cu site to provide a large target area and facilitates the interaction of O_2^- with the enzyme through the guidance of O_2^- to the Cu site and the complementary of the charge. The active site channel consists of the following charged amino acid residues: Thr-58, Glu-132, Glu-133, Lys-136, Thr-137, and Arg-143 (Fig. 10.10(A)). Among them, Glu-133 and Lys-136 are laid on the rim of the active site cavity and form a salt link with each other. These residues play an important role in long-range guidance, and Arg-143 positioned near the Cu ion orients O_2^- to the Cu (local orientation) (Fig. 10.10(B)). Sharp, Fine and Honig⁶⁹⁾ have indicated that the electrostatic field facilitates the association rate of O_2^- more than 30-fold. The participation of Lys-122 and Lys-136 in the approach of O_2^- to the Cu^{2+} has been shown by Argese *et al.*⁷⁰⁾

Site-directed mutagenesis of Arg-143 to Lys, Ile or Glu lowers the activity in human CuZn-SOD,⁷¹⁾ but it does not affect the metal coordination geometry,⁷²⁾ a fact which further supports the participation of Arg-143 in the electrostatic attraction of O_2^- . This is also shown by molecular dynamic calculation.⁷³⁾ Furthermore, Getzoff *et al.*⁷⁴⁾ have indicated the involvement of a hydrogen bonding network formed by Glu-132, Glu-133 and Lys-136 in the electrostatic guidance. Local increase of positive charge by changes of Glu-132 and Glu-133 to Gln maintaining the network of hydrogen bonding facilitates the guidance of O_2^- to the Cu site, and this mutant SOD shows a higher reaction rate than the wild SOD. However, the displacement of Glu-133 to Lys that increases the positive charge but disturbs the network of hydrogen bonding did not increase the rate any more than that of Glu $\rightarrow$ Gln.

The residues participating in the electrostatic guidance of CuZn-SOD are highly conserved during its molecular evolution.⁷⁵⁾ CuZn-SODs from organisms at different evolutionary levels display large differences in the net protein charge and distribution of charged surface residues, but show identical distribution of electrostatic potential in the proximity of the active site.

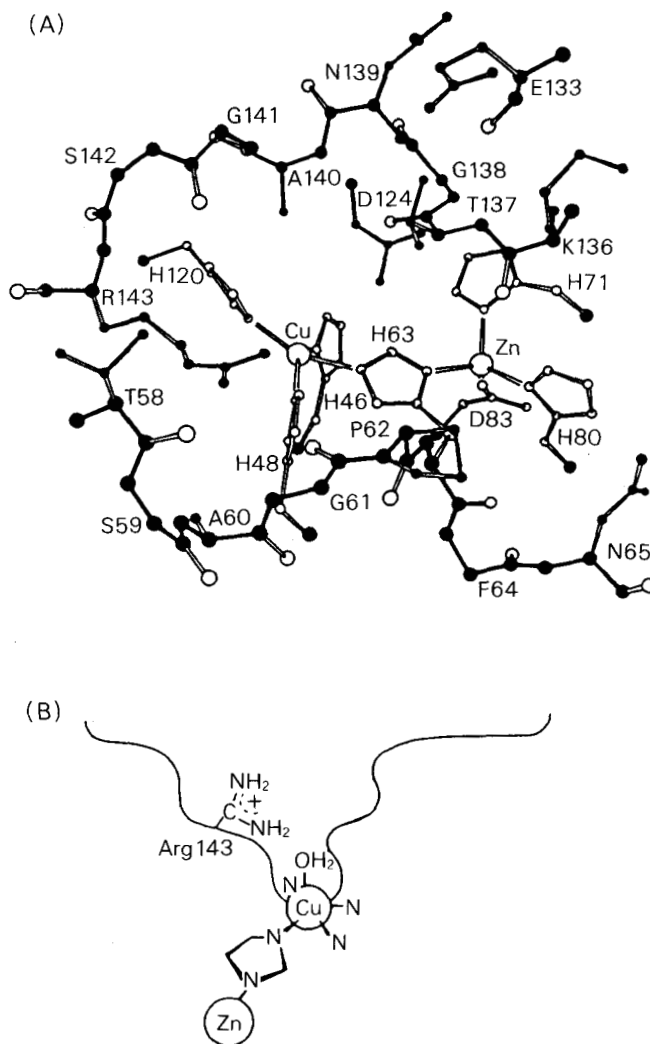


Fig. 10.10 Surroundings of the active site of bovine CuZn-SOD (A)⁴² and schematic drawing of its active cavity (B)³⁰. In (A), main chain shown in black bonds, ligand side chain open bonds, and other side chains with solid atoms and open bonds.

C. Reduction and Oxidation at Active Site Metal

When the O_2^- concentration is below saturation, the rate-limiting step in the catalytic process of CuZn-SOD is the diffusion of O_2^- to the active site channel and its binding to the Cu^{2+} . After entering the active cavity, O_2^- diffuses to the coordination sphere of the Cu. This step involves a displacement of the Cu^{2+} -bound water molecule by O_2^- and the formation of bonds among O_2^- , the Cu^{2+} ion and Arg-143. The reduction of the Cu^{2+} is

accompanied by simultaneous breaking of a bond between the Cu and His-63, and the release of dioxygen (reaction 10.11). The second molecule of O_2^- interacts with the Cu^+ , His-63 and Arg-143 residues and accepts the electron from the Cu^+ , together with the proton from His-63, releasing hydrogen peroxide (reaction 10.12).

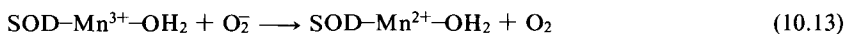
D. Source of Proton⁷⁶⁾

In the dismutation of O_2^- , two protons must be incorporated to form H_2O_2 , but the reaction rate is independent of pH in a range between 5 and 9.5, indicating that the protons do not derive directly from H_3O^+ . The Cu^+ in the reduced SOD is triligated, which is revealed by ESR and absorption spectrum of the metal substituted CuCo-SOD, ^1H NMR, EXAFS and pulse radiolysis and rapid spectroscopy. These results indicate that the proton is donated from the protonated imidazole group of the bridging His-63 between the Cu and the Zn. His-63 shuttles a proton from medium to peroxide at a fast rate, probably not only from H_3O^+ but also from the protonated buffers.

10.6.2 Fe-SOD and Mn-SOD

The reaction mechanisms of Fe-SOD and Mn-SOD are not yet as well characterized as those of CuZn-SOD, but the ligands of the metals have been confirmed to be three histidyl imidazoles and one aspartyl carboxylate. The Fe- and Mn-SODs undergo alternate reduction and reoxidation during the catalytic cycle, similar to CuZn-SOD, at a rate constant of $2 \times 10^9 M^{-1} s^{-1}$. The electron transfer to O_2^- from the metals proceeds through an outer sphere mechanism for Fe- and Mn-SODs, while it proceeds via an inner sphere mechanism for CuZn-SOD. Fe- and Mn-SODs exhibit substrate saturation⁷⁷⁾ as does CuZn-SOD.

A pulse radiolysis study of Mn-SOD from *B. stearothermophilus* has indicated a fast outer sphere mechanism of the interaction of O_2^- with the metals at low O_2^- concentrations ($5.6 \times 10^9 M^{-1} s^{-1}$) and a slow inner sphere mechanism at higher O_2^- concentrations ($4.8 \times 10^7 M^{-1} s^{-1}$). The latter mechanism is becoming more important at high temperatures.⁷⁸⁾ NMR studies support an outer sphere mechanism by which O_2^- interacts with the Mn^{3+} through Mn-bound water.⁷⁹⁾ The proposed reaction for the fast cycle is:



The activities of Fe- and Mn-SODs are decreased above pH 8.5, consistent with the involvement of an ionizable group of the protein with a pK_a value between 9 and 10 in the catalytic cycle.⁷⁷⁾ The most likely residue appears to be Tyr-34, which is only 5 Å apart from the Mn.⁵¹⁾ A possible route of O_2^- to the Mn site in *T. thermophilus* Mn-SOD⁵⁰⁾ runs across the helix 1 between the Lys-29 and Tyr-34, and the O_2^- binds to the Mn^{3+} *trans* to Asp-163 with octahedral coordination (Figs. 10.7 and 10.8). The route to the metal is lined by aromatic and histidine residues including His-26, His-30, Tyr-34, His-81, Phe-84, Trp-165 and His-167.

With increase in ionic strength, the activities of Fe- and Mn-SODs that bear a net negative charge at neutral pH are decreased, suggesting that they also have a cluster of positive charges around the active site.⁸⁰⁾ Acetylation of the lysine residues of Fe- and

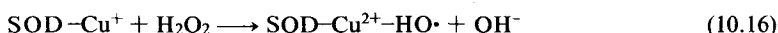
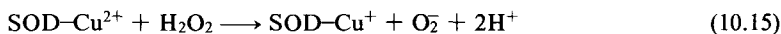
Mn-SODs causes decrease in the activity, indicating that lysine residues are implicated in the electrostatic field around the active site.⁸¹⁾ Thus, Mn-SOD and Fe-SOD exhibit a guiding electrostatic field around its active site in a way similar to that of CuZn-SOD.⁶⁸⁾

10.6.3 Inactivation by H₂O₂

The dismutation product of O₂⁻, H₂O₂, inactivates CuZn-SOD⁸²⁾ and Fe-SOD, but not Mn-SOD⁸³⁾ (Table 10.1).

A. CuZn-SOD

CuZn-SOD is inactivated by H₂O₂ through the following mechanism: H₂O₂ first reduces the enzyme-Cu²⁺ to the Cu⁺ form (reaction 10.15), then the enzyme-Cu⁺ reacts with H₂O₂ to form a strong oxidant at the active site by a Fenton-like reaction (reaction 10.16). The oxidant destroys the Cu-ligated His-63.⁸⁴⁾



The inactivation is a first-order process with respect to both SOD and H₂O₂, and increases with pH exhibiting a pK_a of 10.2, indicating that HO₂⁻ rather than H₂O₂ is the active species.⁸⁵⁾ One histidine residue per subunit is lost during the inactivation; this is assigned to be His-63.⁸⁶⁾ Proper electron acceptors such as histidine and imidazole can prevent the inactivation by competition of the oxidant generated by the interaction of SOD with H₂O₂ between His-63 and the acceptors. Thus, CuZn-SOD shows a weak peroxidase activity for such acceptors.⁸⁷⁾

B. Fe-SOD

While CuZn-SOD is completely inactivated by H₂O₂, Fe-SOD is inactivated to a limit of 90%.⁸⁸⁾ The inactivation is first order with respect to the residual activity if the H₂O₂-resistant portion is subtracted. The apoenzyme of Fe-SOD and the Mn-substituted enzyme are unaffected by H₂O₂. Thus, the reaction of H₂O₂ with the iron at the active site generates a potent oxidant for the inactivation. The inactivation of the *E. coli* Fe-SOD with H₂O₂ is accompanied by loss of the tryptophan residue and some iron, but there is no detectable degradation of histidine or other amino acid residues. *Pseudomonas ovalis* Fe-SOD, however, loses tryptophan, histidine and cysteine residues by H₂O₂.⁸⁹⁾ The oxidant generated in the reactive center oxidizes the ligating histidine residue, then the oxidized histidine residue abstracts the electron from the adjacent tryptophan residue. The difference in the degradation of amino acid residues between the two Fe-SODs is explained by the difference in the conductivity of electron from primarily attacked histidine residue to tryptophan residue.⁸⁸⁾ This reaction is faster in the *E. coli* enzyme than in the *P. ovalis* enzyme; thus, only tryptophan residue is destroyed in the *E. coli* enzyme.

The Fe-SOD activity remaining after the treatment with H₂O₂ is different from that of the native enzyme, with respect to several criteria including heat stability and inhibition by azide. Thus, the H₂O₂-resistant activity of Fe-SOD represents an activity of the H₂O₂-modified Fe-SOD.

REFERENCES

1. K. Asada, S. Kanematsu, S. Okada, T. Hayakawa, in: *Chemical and Biochemical Aspects of Superoxide and Superoxide Dismutase*, Vol. I, p. 136 Elsevier/North-Holland, New York (1980).
2. J.M. McCord, I. Fridovich, *J. Biol. Chem.*, **244**, 6049 (1969).
3. J.V. Bannister, W.H. Bannister, G. Rotilio, *CRC Critical Reviews in Biochemistry*, **22**, 111 (1987).
4. I. Fridovich, *J. Biol. Chem.*, **264**, 7761 (1989).
5. K. Asada, in: *Molecular Biology of Free Radical Scavenging Systems*, p. 173, Cold Spring Harbor Laboratory Press, New York (1992).
6. B. Meier, D. Barra, F. Bossa, L. Calabrese, G. Rotilio, *J. Biol. Chem.*, **257**, 13977 (1982).
7. E.M. Gregory, *Arch. Biochem. Biophys.*, **238**, 83 (1985).
8. C.D. Pennington, E.M. Gregory, *J. Bacteriol.*, **166**, 528 (1986).
9. M.E. Martin, B.R. Byers, M.O.J. Olson, M.L. Salin, J.E.L. Arceneaux, C. Tolbert, *J. Biol. Chem.*, **261**, 9361 (1986).
10. H.M. Steinman, in: *Superoxide Dismutase*, Vol. I, p. 11, CRC Press, Florida (1982).
11. K. Asada, S. Kanematsu, in: *Evolution of Protein Molecules*, p. 361, Japan Scientific Societies Press, Tokyo (1978).
12. S. Kanematsu, K. Asada, *Plant Cell Physiol.*, **30**, 381 (1989).
13. S. Kanematsu, K. Asada, *Plant Cell Physiol.*, **30**, 717 (1989).
14. S. Kanematsu, K. Asada, *Plant Cell Physiol.*, **31**, 99 (1990).
15. S. Kanematsu, K. Asada, *Free Rad. Res. Comms.*, **12**, 383 (1991).
16. S. Kanematsu, K. Asada, *Arch. Biochem. Biophys.*, **185**, 473 (1978).
17. S. Kanematsu, K. Asada, *FEBS Lett.*, **91**, 94 (1978).
18. S. Kanematsu, K. Asada, *Arch. Biochem. Biophys.*, **195**, 535 (1979).
19. S.M. Bridges, M.L. Salin, *Plant Physiol.*, **68**, 275 (1981).
20. R.A. Weisiger, I. Fridovich, *J. Biol. Chem.*, **248**, 4793 (1973).
21. K. Puget, A.M. Michelson, *Biochem. Biophys. Res. Commun.*, **58**, 830 (1974).
22. M. Takahashi, K. Asada, *Arch. Biochem. Biophys.*, **226**, 558 (1983).
23. S. Okada, S. Kanematsu, K. Asada, *FEBS Lett.*, **103**, 106 (1979).
24. T. Hayakawa, S. Kanematsu, K. Asada, *Planta*, **166**, 111 (1985).
25. L.M. Sandalio, L.A. del Rio, *Plant Physiol.*, **88**, 1215 (1988).
26. F.S. Archibald, I. Fridovich, *J. Bacteriol.*, **145**, 442 (1981).
27. F.S. Archibald, I. Fridovich, *Arch. Biochem. Biophys.*, **214**, 452 (1982).
28. M.D. Scott, S.R. Meshnick, J.W. Eaton, *J. Biol. Chem.*, **262**, 3640 (1987).
29. C.A. Bloch, F.M. Ausubel, *J. Bacteriol.*, **168**, 795 (1986).
30. E.D. Harris, *FASEB J.*, **6**, 2675 (1992).
31. D.O. Natvig, K. Imlay, D. Touati, R.A. Halliwell, *J. Biol. Chem.*, **262**, 14697 (1987).
32. H.M. Hassan, I. Fridovich, *J. Bacteriol.*, **129**, 1574 (1977).
33. C.S. Moody, H.M. Hassan, *J. Biol. Chem.*, **259**, 12821 (1984).
34. H.M. Hassan, C.S. Moody, *J. Biol. Chem.*, **262**, 17173 (1987).
35. S.M. Naik, H.M. Hassan, *Proc. Natl. Acad. Sci. USA*, **87**, 2618 (1990).
36. H.M. Hassan, H-C.H. Sun, *Proc. Natl. Acad. Sci. USA*, **89**, 3217 (1992).
37. S.W. Bowen, H.M. Hassan, *J. Biol. Chem.*, **263**, 14808 (1988).
38. J.T. Greenberg, P. Monach, J.H. Chou, P.D. Josephy, B. Demple, *Proc. Natl. Acad. Sci. USA*, **87**, 6181 (1990).
39. K. Asada, *Metalloproteins*, p. 331, Elsevier, Amsterdam (1988).
40. L. Tibell, K. Hjalmarsson, T. Edlund, G. Skogman, Å. Engström, S.L. Marklund, *Proc. Natl. Acad. Sci. USA*, **84**, 6634 (1987).
41. J.A. Tainer, E.D. Getzoff, K.M. Beem, J.S. Richardson, D.C. Richardson, *J. Mol. Biol.*, **160**, 181 (1982).
42. J.A. Tainer, E.D. Getzoff, J.S. Richardson, D.C. Richardson, *Nature*, **306**, 284 (1983).
43. K. Djinoic, G. Gatti, A. Coda, L. Antolini, G. Pelosi, A. Desideri, M. Falconi, F. Marmocchi, G. Rotilio, M. Bolognesi, *J. Mol. Biol.*, **225**, 791 (1992).
44. Y. Kitagawa, N. Tanaka, Y. Hata, M. Kusunoki, G.P. Lee, Y. Katsube, K. Asada, S. Aibara, Y. Morita, *J. Biochem.*, **109**, 477 (1991).
45. H.E. Parge, R.A. Halliwell, J.A. Tainer, *Proc. Natl. Acad. Sci. USA*, **89**, 6109 (1992).
46. K. Djinoic, A. Coda, L. Antolini, G. Pelosi, A. Desideri, M. Falconi, G. Rotilio and M. Bolognesi, *J. Mol. Biol.*, **226**, 227 (1992).
47. B.B. Keele, Jr., J.M. McCord, I. Fridovich, *J. Biol. Chem.*, **245**, 6176 (1970).
48. F.J. Yost, Jr., I. Fridovich, *J. Biol. Chem.*, **248**, 4905 (1973).

49. A. Carlouz, M.L. Ludwig, W.C. Stallings, J.A. Fee, H.M. Steinman, D. Touati, *J. Biol. Chem.*, **263**, 1555 (1988).
50. W.C. Stallings, K.A. Pattridge, R.K. Strong, M.L. Ludwig, *J. Biol. Chem.*, **260**, 16424 (1985).
51. M.W. Parker, C.C.F. Blake, *J. Mol. Biol.*, **199**, 649 (1988).
52. C.J. Brock, J.E. Walker, *Biochemistry*, **19**, 2873 (1980).
53. W.C. Stallings, T.B. Powers, K.A. Pattridge, J.A. Fee, M.L. Ludwig, *Proc. Natl. Acad. Sci. USA*, **80**, 3884 (1983).
54. D. Ringe, G.A. Petsko, F. Yamakura, K. Suzuki, D. Ohmori, *Proc. Natl. Acad. Sci. USA*, **80**, 3879 (1983).
55. B.H.J. Bielski, R.L. Arudi, *Anal. Biochem.*, **133**, 170 (1983).
56. M. Eigen, W. Kruse, G. Maass, L. de Mayer, in: *Progress in Reaction Kinetics*, Vol. 2, p. 287, Pergamon Press, New York (1964).
57. K. Asada, M. Takahashi, M. Nagate, *Agric. Biol. Chem.*, **38**, 471 (1974).
58. E.M. Fielden, P.B. Robers, R.C. Bray, D.J. Lowe, G.N. Mautner, G. Rotilio, L. Calabrese, *Biochem. J.*, **139**, 49 (1974).
59. M. Takahashi, K. Asada, *J. Biochem.*, **91**, 889 (1982).
60. J.S. Valentine, A.B. Curtis, *J. Amer. Chem. Soc.*, **97**, 224 (1975).
61. C.J. Pederson, *J. Amer. Chem. Soc.*, **89**, 7017 (1967).
62. G.J. McClune, J.A. Fee, *Biophys. J.*, **24**, 65 (1978).
63. C. Bull, G.J. McClune, J.A. Fee, *J. Amer. Chem. Soc.*, **105**, 5290 (1983).
64. C. Bull, G.J. McClune, J.A. Fee, *J. Amer. Chem. Soc.*, **107**, 3295 (1983).
65. D. Klug, I. Fridovich, J. Rabani, *J. Amer. Chem. Soc.*, **95**, 2786 (1973).
66. E.K. Hodgson, I. Fridovich, *Biochem. Biophys. Res. Commun.*, **54**, 270 (1973).
67. J.A. Fee, C. Bull, *J. Biol. Chem.*, **261**, 13000 (1986).
68. E.D. Getzoff, J.A. Tainer, P.K. Weiner, P.A. Kollman, J.S. Richardson, D.C. Richardson, *Nature*, **306**, 287 (1983).
69. K. Sharp, R. Fine, B. Honig, *Science*, **236**, 1460 (1987).
70. E. Argese, P. Viglino, G. Rotilio, M. Scarpa, A. Rigo, *Biochemistry*, **26**, 3224 (1987).
71. W.F. Beyer, Jr., I. Fridovich, G.T. Mullenbach, R. Hallewell, *J. Biol. Chem.*, **262**, 11182 (1987).
72. L. Banci, I. Bertini, C. Luchinat, R.A. Hallewell, *J. Amer. Chem. Soc.*, **110**, 3629 (1988).
73. L. Banci, P. Carloni, G.L. Penna, P.L. Orioli, *J. Amer. Chem. Soc.*, **114**, 6994 (1992).
74. E.D. Getzoff, D.E. Cabelli, C.L. Fisher, H.E. Parge, M.S. Viezzoli, L. Banci, R.A. Hallewell, *Nature*, **358**, 347 (1992).
75. A. Desideri, M. Falconi, F. Polticelli, M. Bolognesi, K. Djinovic, G. Rotilio, *J. Mol. Biol.*, **223**, 337 (1992).
76. M.E. McAdam, E.M. Fielden, F. Lavelle, L. Calabrese, D. Cocco, G. Rotilio, *Biochem. J.*, **167**, 271 (1977).
77. J.A. Fee, G.J. McClune, P. O'Neill, E.M. Fielden, *Biochem. Biophys. Res. Commun.*, **100**, 377 (1981).
78. M.E. McAdam, F. Lavelle, R.A. Fox, E.M. Fielden, *Biochem. J.*, **165**, 81 (1977).
79. J.J. Villafranca, F.J. Yost, Jr., I. Fridovich, *J. Biol. Chem.*, **249**, 3532 (1974).
80. A. Cudd, I. Fridovich, *J. Biol. Chem.*, **257**, 11443 (1982).
81. J. Benovic, T. Tillman, A. Cudd, I. Fridovich, *Arch. Biochem. Biophys.*, **221**, 329 (1983).
82. E.M. Fielden, P.B. Roberts, R.C. Bray, G. Rotilio, *Biochem. Soc. Trans.*, **1**, 52 (1973).
83. K. Asada, K. Yoshikawa, M. Takahashi, Y. Maeda, K. Enmanji, *J. Biol. Chem.*, **250**, 2801 (1975).
84. E.K. Hodgson, I. Fridovich, *Biochemistry*, **14**, 5294 (1975).
85. D.M. Blech, C.L. Borders, Jr., *Arch. Biochem. Biophys.*, **224**, 579 (1983).
86. R.C. Bray, S.A. Cockle, E.M. Fielden, P.B. Roberts, G. Rotilio, L. Calabrese, *Biochem. J.*, **139**, 43 (1974).
87. M. Takahashi, Y. Kono, K. Asada, *Plant Cell Physiol.*, **21**, 1431 (1980).
88. W.F. Beyer, Jr., I. Fridovich, *Biochemistry*, **26**, 1251 (1987).
89. F. Yamakura, *Biochem. Biophys. Res. Commun.*, **122**, 635 (1984).
90. I. Bertini, L. Banci, C. Luchinat, B.H.J. Bielski, D.E. Cabelli, G.T. Mullenbach, R.A. Hallewell, *J. Amer. Chem. Soc.*, **111**, 714 (1989).

11

H⁺-ATPase:
Catalysis and H⁺ Translocation

Masamitsu Futai, Atsuko Iwamoto and Masatomo Maeda

*Department of Organic Chemistry and Biochemistry
The Institute of Scientific and Industrial Research
Osaka University, Ibaraki, Osaka 567, Japan*

11.1 Introduction

Three types (P-, F- and V-type) of H⁺-ATPases transport protons across biological membranes coupling with the energy generated by the hydrolysis of ATP, resulting in the formation of electrochemical gradient of H⁺ (H⁺ gradient and membrane potential).¹⁾ Interests has naturally focused on their catalytic sites and mechanisms. Structures of ion pathways have been challenging and unique questions for these osmotic enzymes. Furthermore, the mechanism of coupling between the chemical reaction (ATP hydrolysis/synthesis) and vectorial movements of H⁺ are also important questions to be answered.

The P-type H⁺-ATPases such as fungal H⁺-ATPase and gastric H⁺/K⁺-ATPase are similar to Na⁺/K⁺-ATPase and Ca²⁺-ATPase in structure and mechanism, and were previously known collectively as E₁E₂ ATPases from their two distinct conformations during catalytic cycles.²⁾ These enzymes form phospho-enzyme intermediates during their catalytic cycles. Their subunit structures are rather simple: *Sacharomyces cerevisiae* or *Neurospora crassa* H⁺-ATPase has a single subunit of molecular weight of about 90,000, while gastric H⁺/K⁺-ATPase has the α (molecular weight, 100,000) and the β (molecular weight, 30,000) subunits similar to Na⁺/K⁺-ATPase.

The F-type H⁺-ATPase isolated originally as a coupling factor of oxidative phosphorylation³⁾ can be found widely in mitochondria, chloroplasts or bacterial (eubacterial) plasma membranes.⁴⁻¹⁰⁾ The enzyme fuctions, in most cases, as an ATP synthase coupling with H⁺ transport, but can reversibly transport H⁺ coupling with ATP hydrolysis. An anaerobe such as *Streptococcus faecalis* has this enzyme functioning only as an ATP-dependent H⁺ pump for regulation of the cytoplasmic pH.¹¹⁾ The F-type ATPase has a complicated subunit structure, and is formed from two sectors F₁ (membrane extrinsic sector) and F₀ (membrane intrinsic sector).⁴⁻¹⁰⁾ The subunit structure of *Escherichia coli* F₀F₁ has been defined both genetically and biochemically: F₁ has five subunits α , β , γ , δ and ϵ , while F₀ is composed of the *a*, *b*, and *c* subunits⁴⁻⁸⁾ (Fig. 11.1). The subunit stoichiometry is known as $\alpha_3 \beta_3 \gamma_1 \delta_1 \epsilon_1 a_1 b_2 c_{10}$. The subunit structure of the chloroplast enzyme is similar and formed from the α , β , γ , δ , ϵ , I, II, III, and IV subunits.¹⁰⁾ On the other hand mitochondrial enzymes seem to have other regulatory subunits including ATPase inhibitor.¹²⁾

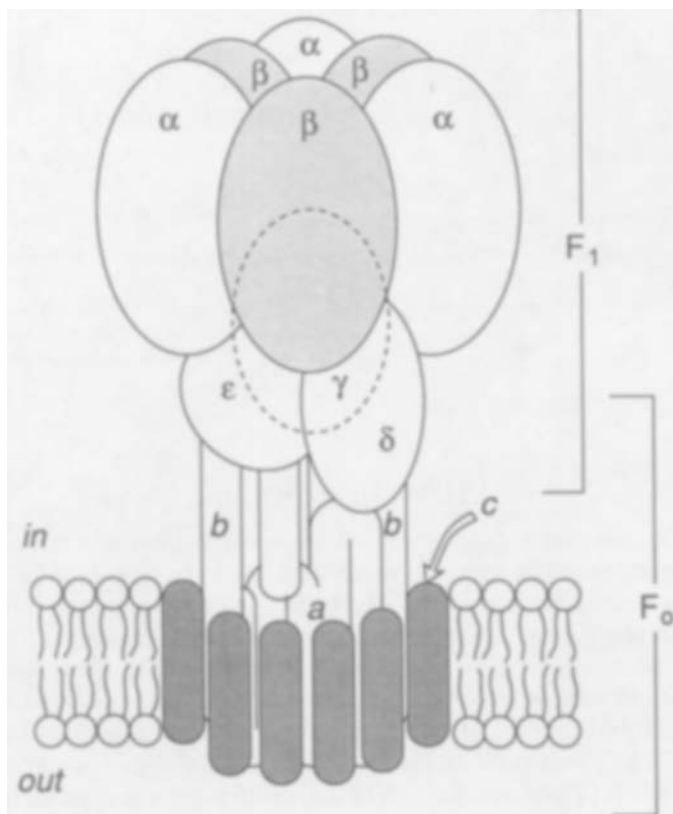


Fig. 11.1 A model of H^+ -ATP (F_0F_1).

This model is based on subunit stoichiometry, predicted secondary structure and subunit-subunit interactions. F_1 sector (α , β , γ , δ , and ϵ subunits) has catalytic sites and F_0 sector (a , b and c subunits) forms a proton pathway.

The V-type (or vacuolar-type) H^+ -ATPase was first found in fungal vacuoles and plant tonoplasts,^{13,14} and is now known to be widely distributed in mammalian endomembrane systems.¹⁵ Like F-type ATPase, the subunit structure of the V-type ATPase is complicated. The membrane extrinsic sector is composed of 73 kDa, 58 kDa, 38 kDa, and 34 kDa subunits, while the membrane intrinsic sector is composed of 40 kDa, 20 kDa and 16 kDa subunits.

The V-type H^+ -ATPases acidify or form electrochemical gradients of H^+ in endomembranes such as coated vesicles, endosomes, lysosomes, Golgi vesicles, synaptic vesicles, etc.¹⁵ V-type ATPases from these eukaryotic organelles have similar molecular weight subunits which cross-react with antibodies against corresponding subunits from chromaffin granules.¹⁶ The 72 kDa and 54 kDa subunits of V-type H^+ -ATPase are homologous to the β and α subunits, respectively, of F-type H^+ -ATPase.¹⁷⁻²⁰ The proteolipid or intrinsic membrane 16 kDa subunit of V-type ATPase and the c subunit of F-type ATPase may have evolved from the same ancestral protein after gene duplication,²¹ because amino and

carboxyl terminal halves of the 16 kDa subunit are homologous to the *c* subunits²²⁻²⁶) and molecular weight of the *c* subunit is about half that of the 16 kDa subunit. Similarities in the two type ATPases were also indicated for the mechanism of ATP hydrolysis.^{27,28}) Thus results from F-type ATPase may be applicable in principle for V-type ATPase, or, at least, results from the two ATPases may be mutually useful.

In this chapter, we discuss our current understanding of the structure and mechanism of the F-type ATPase. However, a comprehensive review is not intended, and the results discussed are mainly from our recent studies on *E. coli* enzyme. Naturally discussion for important results from mitochondria or chloroplasts are included. For other aspects of the enzyme, readers are referred to excellent review articles that have appeared recently.³⁻¹⁰)

11.2 Structure and Function of F-type ATPase

11.2.1 Preparation of F₀F₁

The F-type H⁺-ATPase (F₀F₁) of *E. coli* is coded by *unc* (*atp*) operon. Following a mutation in any one of the eight cistrons, the enzyme becomes defective in ATP synthesis and the bacteria are unable to grow on succinate as the sole carbon source.⁴⁻⁸) The *E. coli* F₀F₁ has been purified in the presence of detergents by combined column chromatographies or a series of density gradient centrifugations.²⁹⁻³¹)

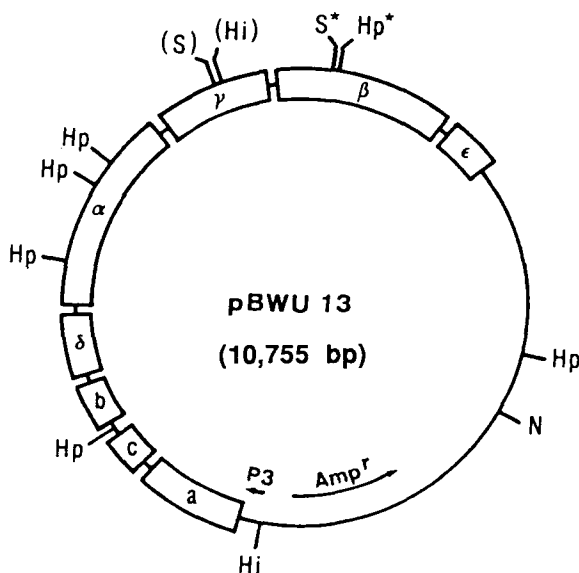


Fig. 11.2 A recombinant plasmid carrying *unc* operon.

The recombinant plasmid pBWU13 was constructed from a *Hind* III-*Nde* I segment (8,660 bp) carrying *E. coli unc* operon.³²) Strain DK8 harboring pBWU13 has at least ten times more F₀F₁ than the wild type. Restriction sites are shown: Hp, *Hpa*I; Hi, *Hind* III; N, *Nde* I; S, *Sac* I. Sites in parentheses are deleted from the original DNA and those with asterisks are newly introduced.

A rapid one-step purification procedure for F_0F_1 was established recently.³²⁾ A recombinant plasmid (pBWU13) carrying the entire *unc (atp)* operon was constructed and introduced into strain DK8 lacking the same operon (Fig. 11.2). The membranes of DK8/pBWU13 had at least tenfold more F_0F_1 than those of wild-type K12 strain; strikingly about 30% inner membrane protein of DK8/pBWU13 was F_0F_1 . The F_0F_1 was solubilized from the inner membrane with octylglucoside and applied to a glycerol gradient centrifugation in the presence of the same detergent. The F_0F_1 obtained was practically homogeneous (Fig. 11.3), and could be reconstituted into proteoliposomes capable of forming proton gradient (Fig. 11.4(A)) and membrane potential (Fig. 11.4(B)). A large amount of F_0F_1 for structural and functional studies became easily available by the rapid purification. Strain DK8 carrying derivatives of pBWU13 with single mutations in F_0F_1 genes could be constructed for analyzing the roles of subunits or amino acid residues. This approach was used extensively for investigating the β subunit mutations as discussed below.

11.2.2 Assembly and Structure of F_0F_1

The complicated subunit complex of F_0F_1 could be dissociated and reconstituted. The F_1 sector can be solubilized from membranes with dilute buffer containing EDTA and has

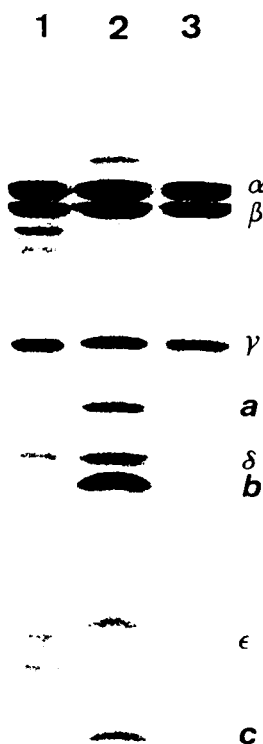


Fig. 11.3 Polyacrylamidegel electrophoresis of F_0F_1 purified from an overproducing strain. F_0F_1 solubilized from membranes of DK8 carrying pBWU13 was applied to glycerol gradient, and fractions near the bottom had practically pure F_0F_1 ; lane 1, EDTA extract of DK8/pBWU13; lane 2, purified F_0F_1 ; lane 3, purified F_1 . Details can be found in Moriyama *et al.*³²⁾

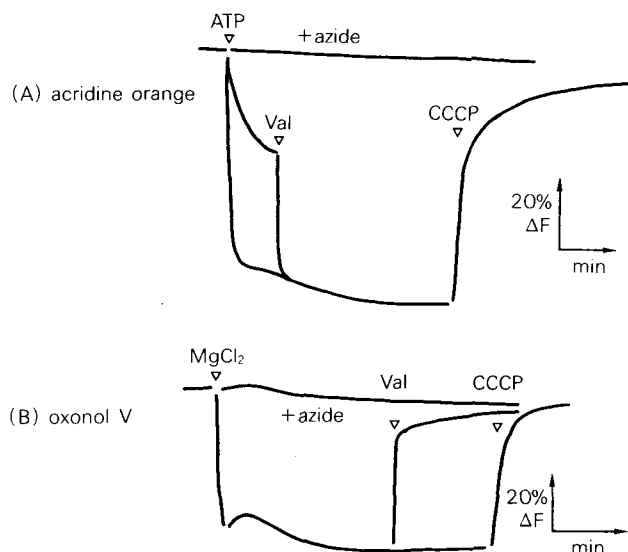


Fig. 11.4 Formation of proton gradient and membrane potential in proteoliposomes reconstituted with purified F_0F_1 .

Reconstituted liposomes (6 μg protein) were suspended in buffer containing 1 μM acridine orange (A) or 1 μM oxonol V (B), and fluorescence was monitored. At indicated times, 1 mM ATP, 0.1 μM valinomycin and 1 μM CCCP were added.³²⁾

ATPase activity but no activity for ATP synthesis or ATP-driven H^+ transport. Higher-ordered structure of rat liver F_1 at 3.6 $\text{\AA}$ resolution was determined by X-ray diffraction³³⁾; the overall dimensions of F_1 are 120 $\text{\AA} \times 120 \text{\AA} \times 74 \text{\AA}$ and the enzyme (α and β subunits) exhibited sixfold symmetry, although the locations of the minor subunits (γ , δ , and ϵ) were not determined. X-ray diffraction may solve the structure of F_1 at the level of amino acid residues in the near future.

Any one of the purified subunits of F_1 does not have ATPase activity,^{34,35)} although the α and β subunits have ATP binding activities.^{35,36)} The complex with ATPase activity could be reconstituted only from combinations of α , β and γ subunits.^{34,35)} The reconstitution of the $\alpha\beta$ complex with ATPase activity was reported for a thermophilic F_1 ,^{37,38)} but such a complex was not shown for *E. coli* enzyme. The *E. coli* $\alpha_3\beta_3\gamma$ complex could bind to F_0 only after F_1 was reconstituted with the δ and ϵ subunits.³⁵⁾ Consistent with the *in vitro* reconstitution experiments, the strain defective in translation initiation (altered Shine-Dalgarno sequence) of the ϵ subunit (138 amino acid residues) or nonsense mutations such as Gln-72 $\rightarrow$ end gave soluble (cytoplasmic) F_1 sector unable to bind to F_0 sector.³⁹⁾ The F_0F_1 with the ϵ subunit lacking 16 residues from the amino terminus was also defective and F_1 was mostly in the cytoplasmic fraction.⁴⁰⁾ A similar observation was made for the δ subunit (177 amino acid residues); missense mutation such as δ Gly-150 $\rightarrow$ Asp or nonsense mutations (δ Val-174 $\rightarrow$ end, δ Ala-172 $\rightarrow$ end, δ Arg-170 $\rightarrow$ end) rendered F_1 soluble in the cytoplasmic fraction.⁴¹⁾

An active F_0 sector having an H^+ pathway and capable of binding F_1 sector could be reconstituted from the *a*, *b*, and *c* subunits into liposomes.^{30,31)} On the other hand,

individual subunits in liposomes did not have activity. Genetic results also indicated that the three subunits are required for the active F_0 .³¹⁾ A large amount of F_0 could be obtained from F_0F_1 purified from DK8/pBWU13 and its structural analysis after two-dimensional crystallization may produce important structural information.

11.2.3 Uni-site and Multi-site Catalysis by F_1

The solubilized F_1 has the active sites of the entire enzyme and catalyzes ATP hydrolysis. Similar to mitochondrial F_1 ,⁴²⁾ *E. coli* F_1 has six nucleotide binding sites (three for catalytic and three for non-catalytic sites).⁴³⁾ The role(s) of the non-catalytic sites is still unknown. The *E. coli* F_1 having at least two vacant catalytic sites, or F_1 having less than one vacant catalytic site could be prepared.⁴⁴⁾

Detailed kinetics of ATP hydrolysis in single-site (uni-site) and steady-state (multi-site) conditions by beef heart F_1 were studied extensively (Fig. 11.5).^{45,46)} At an ATP/ F_1 ratio of less than 1, ATP binds to a catalytic site and is hydrolyzed slowly (uni-site catalysis). The equilibrium constant between bound substrate (ATP) and products (ADP and P_i) bound at the catalytic site of F_1 was close to 1, indicating that the equilibrium can occur without change in free energy. In the presence of excess ATP (multi-site catalysis), ATP binds to all three catalytic sites, and the ATP at the first site is hydrolyzed at a rate that is at least 10^6 times higher than is the case in uni-site catalysis.

Uni-site and multi-site catalyses were essentially confirmed with *E. coli* F_1 , although the ratio of the rates of multi- to uni-site hydrolysis with this enzyme was slightly lower than beef heart F_1 .^{44,47,48)} The mutation with the α or β subunit defective only in multi-site catalysis was isolated. Typically, α Ser-373→Phe or α Arg-376→Cys mutant enzyme with amino acid substitution in the α subunit had low multi-site activity (less than 0.1% that of the wild type), but essentially normal uni-site activity.^{49,50)} These mutant enzymes were defective in catalytic cooperativity, but normal in the catalytic reaction in a single catalytic

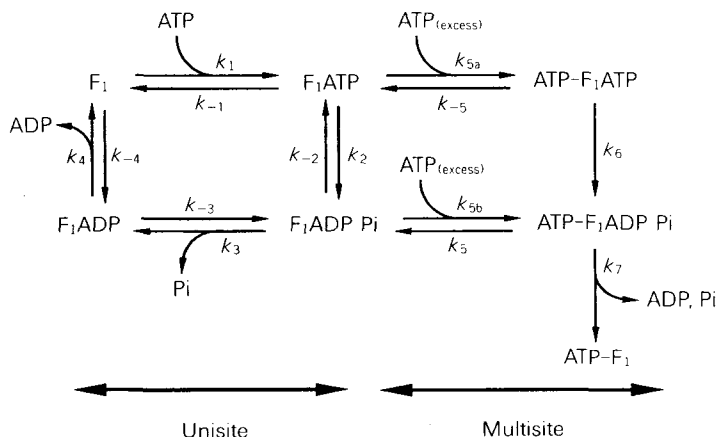


Fig. 11.5 Uni-site (single-site) and multi-site (steady-state) catalysis by F_1 -ATPase. The scheme of ATP hydrolysis (synthesis) by F_1 is cited and modified from Penefsky and coworkers.^{45,46)}

site. Similarly, inhibitors such as sodium azide⁵¹⁾ and DCCD⁵²⁾ inhibited multi-site catalysis almost completely, whereas they had only slight effects on uni-site catalysis.

The three β subunits in F_1 have been proposed to undergo a series of conformational changes during multi-site catalysis.⁵³⁾ Reconstitution experiments of the $\alpha_3\beta_3\gamma$ complex using varied ratios of normal and defective β subunits together with the α and γ subunits suggested that a complex with one defective β subunit was not active and all three β subunits must be normal during multi-site catalysis.⁴⁸⁾ Experiments using defective and normal α subunits reached the similar conclusion that all three α subunits must be normal during multi-site catalysis.⁵⁴⁾

11.3 Catalytic Site of F-type ATPase

11.3.1 Affinity Labeling of Catalytic Site of F_1 ATPase

At least part of the catalytic site on the β subunit has been defined by affinity labeling with ATP analogues and analysis of random and directed mutants including pseudorevertants. A combination of different approaches was useful for proposing a model of the structure of the catalytic site near the γ -phosphate moiety of ATP (Fig. 11.6). It may be noted that the model was obtained before the higher-ordered structure of F_1 with defined amino acid residues became available.

Affinity labeling and chemical modification studies were carried out with purified F_1 . The β Thr-287, β Ile-290 and β Tyr-297 residues (*E. coli* numbering) of beef heart F_1 were

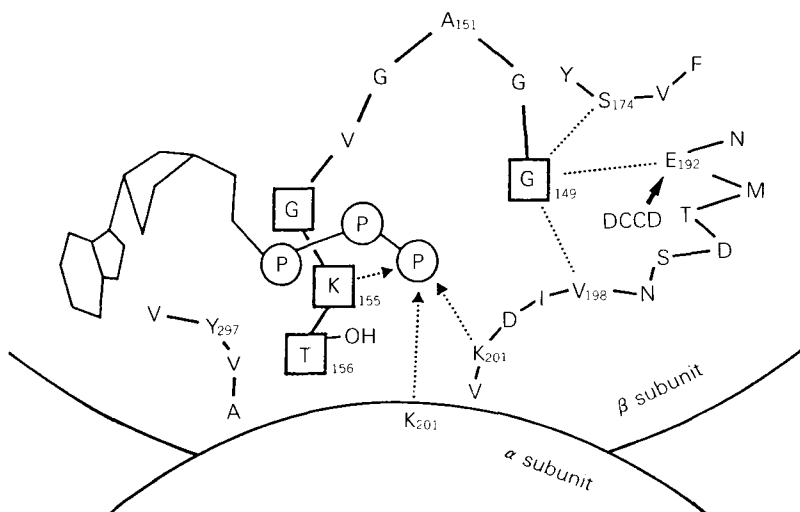


Fig. 11.6 A model of H^+ -ATPase catalytic site near the γ -phosphate moiety of ATP. Results from affinity labeling and mutagenesis (random and directed mutants and pseudorevertants) are shown. Dotted lines indicate the close interactions between two residues suggested from genetic study. Arrows with dotted line show residues near the γ phosphate of ATP suggested by affinity labeling. An arrow shows the DCCD binding residue.

targets of a photoaffinity labeling reagent 8-azido-ATP, indicating that these residues are near the adenine moiety of ATP in the catalytic site.⁵⁵⁾ The β Thr-285 residue is in a sequence (Ile-Thr-Ser-Thr-Lys-Thr-Gly-ser-Ile-Thr, residues 282–291 of the *E. coli* β subunit) which is similar to the sequence (Ile-Cys-Ser-Asp-Lys-Thr-Gly-Thr-Lys-Thr) conserved in the phosphorylation site of P-type ATPases. The β Thr-285 corresponds to the essential aspartate residue for the P-type ATPase (see residues underlined above). The mutation β Thr-285→Asp changed the kinetic of uni-site and multi-site catalysis,⁵⁶⁾ consistent with the affinity labeling experiment with 8-azido-ATP.⁵⁵⁾ However, β Asp-285 enzyme could not form phosphoenzyme. Webb *et al.* showed that the mechanism of ATP hydrolysis by mitochondrial F₁ is a one-step replacement of ADP by a water oxygen, unlike the two-step transfer carried out by P-type ATPases.⁵⁷⁾ It is reasonable to assume that a single amino acid substitution was not enough to change the mechanism of the F-type ATPase to that of P-type ATPase.

The β Tyr-297 residue of the beef heart F₁ reacted with 7-chloro-4-nitrobenzofurazan⁵⁸⁾ and β Ile-290, β Gln-294 and β Tyr-297 reacted with 4-azido-2-nitrophenyl phosphate (*E. coli* numbering). The nitrobenzofurazan moiety was transferred from β Tyr-297 to β Lys-155, suggesting that these two regions may be close together.⁵⁹⁾ 2-Azido-ADP reacted with region between β Leu-328 and β Pro-332.⁶⁰⁾ The β Tyr-331 and β Asp-338 residues are the targets of 3'-O-(4-benzoyl) benzoyl ATP,⁶¹⁾ and β Tyr-331 was modified with 5'-*p*-(fluorosulfonyl) benzoyl inosine.⁶²⁾ These results suggest that positions 328–338 are in or near the catalytic site.

Binding of one mole of adenosine triphosphopyridoxal (AP₃-PL) to F₁ caused complete inhibition of uni- and multi-site catalysis.⁶³⁾ Target residues of AP₃-PL identified were α Lys-201 and β Lys-155 of the α and β subunits, respectively.⁶⁴⁾ In the presence of Mg²⁺, β Lys-155 and β Lys-201 of the β subunit were labeled predominantly.⁶⁵⁾ These results suggest that three residues, β Lys-201, β Lys-155 and α Lys-201, are near the γ phosphate moiety of ATP. As discussed above, β Lys-155 was also suggested to be near β Tyr-297 from studies on beef heart enzyme.⁵⁹⁾ It must be noted that β Lys-155 is in the glycine-rich sequence (Gly-Gly-Ala-Gly-Val-Gly-Lys-Thr-Val, *E. coli* β subunit residues 149–157) conserved in nucleotide binding proteins.⁸⁾ The glycine-rich sequence (Gly-X-X-X-X-



Fig. 11.7 The glycine-rich sequences of adenylate kinase, *ras* protein, and the β subunit of F₁-ATPase. Amino acid sequence of adenylate kinase (Adk), *ras* protein and the β subunit are aligned. The conserved glycine-rich sequence is shown (bottom). The secondary structures of adenylate kinase and *ras* protein are known from crystallography; ~~~~~, β sheet; ~~~~~, α helix. Identical (asterisks) or homologous (vertical lines) residues are shown.

Gly-Lys-Thr/Ser) is found in many proteins capable of binding ATP or GTP including adenylate kinase^{66,67)} and the *ras* p21 protein^{68,69)} (Fig. 11.7). The sequence corresponds to a flexible loop structure between α helix and β sheet: the sequence of adenylate kinase and *ras* protein are residues 15–23 (Gly-Gly-Pro-Gly-Ser-Gly-Lys-Gly-Thr) and 10–17 (Gly-Ala-Gly-Gly-Val-Gly-Lys-Ser), respectively.

11.3.2 Mutations in the Glycine-Rich Sequence

Consistent with the importance in catalysis, mutations in the glycine-rich sequence had profound effects on F_1 . Although we recognized that the sequence in the β subunit is similar to adenylate kinase,⁷⁰⁾ our real interest was initiated when we found that β Ala-151→Val mutant in our stock of random *unc* mutants was isolated by hydroxylamine treatment.⁷¹⁾ The β Val-151 enzyme had about 60% multi-site catalysis of the wild type and much faster uni-site catalysis. A mutation (Gly-12→Val) in the corresponding residue of the glycine-rich sequence of the *ras* protein decreased the GTPase activity to 10% of that of the wild type.⁷²⁾ The mutation of corresponding residue in the adenylate kinase (Pro-17→Gly or Val) also drastically changed the kinetic properties of the enzyme.⁷³⁾ These results suggested that the glycine-rich sequences of the three proteins may have similar role(s) and structure.

Results of affinity labeling (AP₃-PL binding to β Lys-155) and the β Val-151 mutant prompted us to study the glycine-rich sequence in more detail by directed mutagenesis (Fig. 11.6). Directed mutations were introduced in the β subunit gene of the *unc* operon carried by pBWU13³²⁾ (Fig. 11.2). The entire glycine-rich sequence of the β subunit was replaced by the sequence of the *ras* protein and adenylate kinase.⁷⁴⁾ The mutant with the *ras* protein sequence could grow on succinate by oxidative phosphorylation, and had about 40% of the membrane ATPase activity of the wild type. This result clearly suggests that the glycine-rich sequence of *ras* protein and the β subunit have similar higher-ordered structures and catalytic functions, although the two sequences have three different residues (β Gly-150→Ala, β Ala-151→Gly and β Thr-156→Ser). On the other hand, the mutant with the adenylate kinase sequence was defective in oxidative phosphorylation and had essentially no membrane ATPase activity. The adenylate kinase sequence mutant had an extra glycine residue between β Lys-155 and β Thr-156. Simulation of the secondary structure indicated that the projection of the side chain of the threonine residue (originally β Thr-156) next to the Gly residue is changed 120° from the wild-type orientation. Similarly, the Gly insertion between β Lys-155 and β Thr-156 residues gave defective enzyme. These results suggest that the direction of the projection of Thr-156 side chain, or Thr residue itself, is essential for the catalytic activity.

Importance of the threonine residue was established by directed mutagenesis. We changed β Thr-156 to Cys, Asp, or Ala and found that neither mutant had any detectable membrane ATPase activity and was naturally unable to grow by oxidative phosphorylation.⁷⁵⁾ The position of the threonine residue should be at position 156, because β Ala-156/ β Thr-157 mutant was defective.⁷⁶⁾ Furthermore, the purified F_1 with β Cys-156 or β Ala-156 mutation had no uni- or multi-site catalysis, indicating that β Thr-156 residue is an essential residue for catalysis. However, β Thr-156 could be replaced by Ser residue without loss of activity, consistent with the finding that the glycine-rich sequence of the β subunit could be replaced by that of *ras* protein having Ser residue at the corresponding position to β Thr-156.

Directed mutations were also introduced into the β Lys-155 residue.⁷⁶⁾ The β Ser-155, β Thr-155 and β Ala-155 mutants did not grow by oxidative phosphorylation, and had no membrane ATPase activity. The double mutant β Thr-155/ β Lys-156 had no membrane ATPase activity. The purified β Ser-155 or β Ala-155 enzyme had essentially no multi- or uni-site catalysis indicating that the lysine residue at position 155 is essential for catalysis. The β Lys-155 residue was also replaced by other groups: β Glu-155 and β Gln-155 enzymes had no activity⁷⁷⁾ (A. E. Senior, personal communication).

11.3.3 Amino Acid Residues Interacting with Glycine-rich Sequence

Analysis of pseudorevertants (or suppressor) of *unc* mutants is a unique approach to find functional and/or structural interaction(s) between the two amino acid residues in the complex enzyme. Residues near the γ -phosphate moiety of ATP could be mapped relative to the residues in the glycine-rich sequence by identifying pseudorevertants of the mutants in the glycine-rich sequence. Conversely, other residues can also be mapped closely to the glycine-rich sequence, if effects of their mutations are suppressed by mutations in the glycine-rich sequence.

We found that the pseudorevertant (β Ser-174 $\rightarrow$ Phe/ β Gly-149 $\rightarrow$ Ser) of the mutant (β Ser-174 $\rightarrow$ Phe) had active ATPase similar to that of the wild-type, indicating that the effect of the first mutation (β Phe-174) was suppressed by the second mutation (β Ser-149).^{75,78)} However, this suppression was not due to the replacement of the functional role of the β Ser-174 (wild) by β Ser-149 (mutant), since β Phe-174 mutation was also suppressed by β Ala-149 mutation.⁷⁹⁾ Replacement of β Ser-174 by Gly, Ala, Leu and Phe residues indicated that the enzyme activity decreased with increase in the size of the side chain at position 174. Thus Ser residue at position 174 is not essential and increase in the size of the side chain affected the conformation of the catalytic residues, possibly including β Lys-155 and β Thr-156 residues. β Phe-174 mutation could not be suppressed by β Gly-149 $\rightarrow$ Thr⁷⁹⁾ or β Gly-150 $\rightarrow$ Ser⁷⁵⁾ mutation. Thus the size of the side chain at position 149 appears to be important for suppressing β Phe-174 mutant. β Cys-149/ β Phe-174 mutant was active in oxidative phosphorylation, whereas β Cys-149 mutant was defective, indicating that the effects of the two mutations were mutually suppressed. These results strongly suggest that the β Gly-149 and β Ser-174 residues functionally interact with each other, and are possibly located near each other (Fig. 11.6).

Pseudorevertants for β Cys-149 mutant were also isolated; β Ser-174 $\rightarrow$ Phe, β Val-198 $\rightarrow$ Ala and β Glu-192 $\rightarrow$ Val mutations suppressed the effect of β Cys-149 mutation. The β Val-198 is three residues downstream of the β Lys-201 in primary structure. As discussed above β Lys-201 is near the γ phosphate moiety of ATP and β Lys-155, since both residues were labeled with AP₃PL.⁶⁵⁾ These results are consistent with the notion that β Val-198 may be actually structurally near β Gly-149. The F₁ lost ATPase activity by binding of DCCD to β Glu-192,⁸⁰⁾ supporting the idea that this residue is also near the catalytic site. From these results structure for the catalytic site near γ -phosphate moiety of ATP was proposed (Fig. 11.6).

Analysis of mutations and their suppression as discussed above is a tedious process but yields a great deal of information for determining residues in (or near) the catalytic site of F₁ for which the higher ordered structure at the level of amino acid residues is as yet unknown.

11.4 Pathway and Regulation of H⁺ Transport

11.4.1 Assembly of F₀

Genetic analysis and *in vitro* reconstitution studies suggested that all three F₀ subunits (*a*, *b*, *c*) were required for a functional proton pathway.^{30,31)} A single subunit or a combination of two could not form the H⁺ pathway when incorporated into liposomes. However, these results do not necessarily suggest that amino acid residues from the three subunits actually form a proton pathway.

Close interactions of the membrane spanning domains of the three subunits have been suggested by a genetic method: the effect of *b*Gly-9→Asp mutation in the membrane domain of the *b* subunit was suppressed by a second mutation in the *a* subunit (*a*Pro-240→Ala, Leu)⁸¹⁾ or in the *c* subunit (*c*Ala-62→Ser).⁸²⁾ The three nonsense mutants (*a*Trp-111→end; *a*Trp-231→end; *a*Gln-252→end) had 50–70% of the membrane ATPase activity, but could not form active proton pathways.⁸³⁾ Thus at least part of the F₁ binding site could be formed without a region between *a*Trp-111 and *a*His-271 (carboxyl terminus) of the *a* subunit.

Subunit *b* (156 amino acid residues) has two domains that can easily be defined from their amino acid sequences; a hydrophobic domain (26 amino terminal residues), presumably embedded in the membrane; a hydrophilic domain of 130 residues between *b*Pro-27 and *b*Leu-156 (carboxyl terminus), probably exposed to the cytoplasm.^{4,8)} The amino terminal domain may be essential for the H⁺ pathway, and the carboxyl domain for the F₁ binding site. Experiments using proteolytic enzymes^{84,85)} and photoreactive protein chemical reagents⁸⁶⁾ supported a model in which the amino terminal domain is essential for the H⁺ pathway, and the carboxyl terminal domain for the F₁ binding. However, results of mutant analyses indicated that the two roles of the *b* subunit could not be strictly separated; *b*Gly-9→Asp^{87,88)} and *b*Gly-131→Asp^{87,89)} replacements impaired the formation of both the H⁺ pathway and F₁ binding site. In the case *b*Gly-131→Asp, a mutation in the hydrophilic region impaired the role of the hydrophobic region.

Furthermore, the loss of the carboxyl terminal *b*Leu-156 residue resulted in small, but significant reductions of both F₁ binding and H⁺ translocation, and loss of two residues (*b*Glu-155 and *b*Leu-156) abolished both activities almost completely. Thus the carboxyl terminal region of the *b* subunit is required for correct assembly of the amino terminal (membrane-embedded) region of the *b* subunit with the *a* and *c* subunits. In contrast to the *b* subunit, the seven residues from the carboxyl terminus of the *a* subunit was able to be removed without loss of activity for F₁ binding and H⁺ translocation,^{83,90)} and even the *a* subunit lacking more than half the amino acid residues from the carboxyl terminus was able to form F₀ capable for F₁ binding activity.⁸³⁾ Although not discussed in detail here, mutations in the *c* subunit also impaired assembly of the enzyme.^{4,30)}

11.4.2 Proton Pathway in F₀

Models with five or six transmembrane helices for the *a* subunit have been proposed from combined interpretation of hydrophobicity estimation, amino acid conservation, and mutational analysis. As discussed in detail by Fillingame,⁶⁾ the first three helices in the three models are similar, while one to three helices were proposed from different

				* * * *	
At1	-ASLRELDR	IDSVKNTQKI	TEAMKLVAAA	KVRRAQEAVV	NGRPFSETLV
At2	-AGIRELRER	IDSVKNTQKI	TEAMRLVAAA	RVRRAQDAVI	KGRPFETETLV
So	--ANLRELDR	IGSVKNTQKI	TEAMKLVAAA	KVRRAQEAVV	NGRPFSETLV
Cr	--GLKEVRDR	IASVKNTQKI	TDAMKLVAAA	KVRRAQEAVV	NGRPFSENLV
Sco	MANLKAIRDR	IKSVRNTRKI	TEAMRLVAAA	KLNRAEQNAK	SFRPFADRLA
Scy	MPNLKAIRDR	IQSVKNTKKI	TEAMRLVAAA	KVRRAQEQVL	STRPFADALA
Ana	MPNLKSIRDR	IQSVKNTKKI	TEAMRLVAAA	RVRRAQEQVI	ATRPFADRLA
Rr	MASLKDLRSR	ITSVKSTQKI	TSAMKMVAAS	RLRRAQDTAE	AARPYTQRM
Rb	MPSLKDLKNR	IGSVKNTRKI	TKAMQMVAAS	KLRRAQDSAE	AARPYAERM
Ec	-AGAKEIRSK	IASVQNTQKI	TKAMEMVAAS	KMRKSQDRMA	ASRPYAETMR
Va	MAGAKEIRNK	IGSVKSTQKI	TKAMEMVAAS	KMRSSQDAME	ASRPYAETMR
Bm	MASLRDIQTR	ITSTKKTQKI	TKAMEMVSAK	TKAMEMVSAK	SFRPYMEKI
PS3	-ASLRDIKTR	INATKKTQKI	TKAMEMVLTS	KLNR-AEKRE	IVRPMYMEKI
bom	-ATLKDITRR	LKSIKNIQKI	TKSMKMVAAS	KYARAEREK	PARVYGVGSL
				*	
At1	KGLGLEITYVI	SVGKKGNSYF	LRRPYIPVDK	YLE-AGTL-P	TAKEAQAVAD
At2	KQRGIDCVVI	SVGKKGNAYF	SRRDDFDVVK	CIEGGGVF-P	TTKEAQVSAD
So	KKLGVDYTI	SIGKKGNTYF	IRRPEIPVDR	YFD-GTNL-P	TAKEAQAIAD
Cr	TAMGVKVNLV	CVGRKGAQYF	ARRKQYNIVK	SFS-LGAA-P	STKEAQGIAD
Sco	SAQGLDYKFV	IVGRKAGQYF	QRREQ-PIEA	TYSGLEQI-P	TAQEANDIAD
Scy	KNQGIAYKLV	LVGSKAKQYF	GRRDY-DVAA	SYANLEQI-P	NASEAAQIAD
Ana	KAEGLDYTFV	IVGRKAEQYF	RRREQ-PIDA	SYTGLEQI-P	TAEANKIAD
Rr	EAQKTVKLL	CIGKKGRDGL	-KREFPKQII	GGI-ADQ---	SS-KAIGFSD
Rb	VAQKTVKIL	TVGKKGREQL	-KRDWASAFV	GHV--DL---	SDVRRLLGYSN
Ec	TDKGVQCDLA	MIGSKGVSPF	NSVGGNVVAQ	VTGMGDN--P	SLSELIGPVK
Va	REKGAIEILA	VVGSKATAFF	KHGGAKVAAQ	VSGLDN--P	SLEDLIGSVG
Bm	HQSPDEYGI	AIGRVGRDFF	VKRGIPVLE	ITGLADQ--P	AFADIQGIAS
PS3	HASPDYAI	VIGRVGLSPF	RKRNPVILD	ITRLPDQ--P	SFADIKEIAR
bom	AAAGKEVKII	GVGDKIRSIL	HRTSDQFLV	TFKEVGRPP	TFGDASVIAL
				*	
At1	LTTKEGKLT	ERETFRTPA	DFSPILQFEQ	DPVQILDALL	PLYLNSQILR
At2	LTSKDGKLAV	ERTKLEVEKP	EISPLMQFEQ	DPVQILDAMM	PLYLNSQILR
So	LTTKEGKLT	ERDMIKTETP	AFSPILQFEQ	DPAQILDALL	PLYLNSQILR
Cr	LTTKGGEFAV	EREKTTIETE	ALDPSLIFEQ	EPAQILDALL	PLYMSSCLLR
Sco	LTTGGGSFTV	EREKLTSEVA	PLPRDMIFEQ	DPAQILSALL	PLYLNSQLLR
Scy	LITRGGKFQV	EREKVEAPVE	SFPQDMIFEQ	DPVQILEALL	PLYNTNQLLR
Ana	LTTGGGFQV	ERQTVTSQAR	PLPRDSIFEQ	DPVQILDSLL	PLYLNSQLLR
Rr	--PFAVPTTV	AAGNDNDRTA	GPKAIYVEP	SEEEILADLL	PKNVAIQVFR
Rb	-----	---PAKFEAA	ETNALYDYE	SEEEILADLL	PRGVATQIFT
Ec	-----	-----DDDLK	HKSWDYLYEP	DPKALLDTLL	RRYVESQVYQ
Va	-----	-----SEEMQR	EHSWDYIYEP	EPKPLDITLL	VRYVESQVYQ
Bm	-----	-----PSGK	LV--GYEFEP	SQEEILEVLL	PQYAESLIYG
PS3	-----	-----AENKQR	TV---YEFEP	SQEEILDVLL	PQYAESLIYG
bom	-----	-----SSAESM	S-----IYDD	IDADVLRNYQ	EYSLANIYY

Fig. 11.8 Amino acid sequences of γ subunits from various sources.

The sequences of the γ subunits of *Arabidopsis thaliana* (At1, At2),⁽¹⁰⁴⁾ spinach (So),⁽¹⁰⁵⁾ *Cramidomonas reinhardtii* (Cr),⁽¹⁰⁶⁾ *Synechococcus* 6301 (Sco),⁽¹⁰⁷⁾ *Synechocystis* 6803 (Scy),⁽¹⁰⁸⁾ *Anabaena* (Ana),⁽¹⁰⁹⁾ *Rhodospirium rubrum* (Rr),⁽¹¹⁰⁾ *Rhodopseudomonas blastica* (Rb),⁽¹¹¹⁾ *E. coli* (Ec),⁽¹¹²⁾ *Vibrio alginoliticus*

interpretations of the topology of about 100 amino acid residues from the carboxyl terminus. The model with the four helices in the carboxyl terminal region is difficult to accept, because residues shown to be required for the H⁺ translocation and a residue (*a*Pro-240) interacting with the membrane portion of the *b* subunit were mapped outside the phospholipid bilayer. A five-helices model would accommodate the genetic evidence as suggested.⁶⁾ However, more studies are required to reach a conclusion.

When the sequences of the subunit *a* and *6* (mitochondrial counterpart of the subunit *a*) are aligned, conserved hydrophilic residues are found mostly in the carboxyl terminal region. It is reasonable to assume that some of them may be candidates participating in the H⁺ translocation. Directed mutagenesis suggested that the *a*Arg-210 residue of the *a* subunit is essential for proton translocation. The subunit with a Lys, Glu, Gln, Val or Leu residue at position 210 had essentially no activity for proton conduction. The *a*His-245 and *a*Glu-219 residues may also be directly involved in proton translocation.⁹⁰⁻⁹⁷⁾ These residues may be located in the fourth and fifth transmembrane helix of the five-helices model and in the fifth and sixth helices of the six-helices model. Slight structural changes of the *a* subunit lowered the proton conduction through F₀; replacements of proline residues (*a*Pro-230, *a*Pro-190 or *a*Pro-143) by other residues lowered proton conduction,^{90,97)} although these proline residues may not be directly involved in H⁺ transport.

Analyses of the nonsense mutants suggested that the carboxyl terminal region is important for the structure of the subunit^{83,90)}: *a*Glu-269→end and *a*Ser-265→end mutations were essentially normal, whereas *a*Tyr-263→end mutation was defective. However, *a*Tyr-263→Phe mutation had no effect, indicating that *a*Tyr-263 is not a functional residue. Vik, Lee and Marshall showed that *a*Tyr-263→end mutant showed a temperature-sensitive phenotype⁹⁸⁾: membranes of the mutant grown at 25°C showed ATP dependent proton-translocation, while those grown at 37°C were defective. These results suggest that *a*Tyr-263 and *a*Leu-264 are important to maintain functional assembly of the *a* subunit. Interestingly the seven amino acid residues between *a*Ser-265 and *a*His-271 (carboxyl terminus) of the *a* subunit could be truncated by nonsense mutation without loss of activity, whereas nine amino acid residues between *a*Tyr-263 and *a*His-271 could not be removed. The removal of these nine residues caused drastic change in the assembly of the *a* subunit or the entire proton pathway of F₀.

A membrane-embedded hairpin structure for the *c* subunit was established from genetic studies.³¹⁾ The role of *c*Asp-61 in the second membrane helix of the *c* subunit has been studied in detail: the *c*Asp-61→Asn or Gly mutants showed no proton conduction or ATP synthesis.⁹⁹⁾ Binding of DCCD to this residue blocked proton conduction. These results are consistent with the notion that *c*Asp-61 is part of the proton pathway. The carboxyl residue at position 61 was able to be transferred to the corresponding position of the first helix.¹⁰⁰⁾ Vacuolar-type ATPase has a similar proteolipid subunit possibly evolved from the same ancestral protein as the *c* subunit.²⁵⁾ The vacuolar proteolipid has glutamate in the middle of the fourth membrane-spanning helix, corresponding to the position of *c*Asp-61 of the *c* subunit. The glutamate may play an important role in proton conduction similar to the *c* subunit.¹⁰¹⁾

The membrane domain of the *b* subunit may not play any role in proton-translocation. Thus we may conclude that the *a* and *c* subunits form a proton pathway, and residues such as *a*Arg-210, *a*Glu-219, *a*His-245 and *c*Asp-61 function as the pathway. Further studies will give more detailed understanding of proton translocation.

11.4.3 Regulation of Proton Conduction

The most difficult and interesting question about H⁺-ATPase is how chemical reaction (ATP synthesis/hydrolysis) is coupled with vectorial H⁺ conduction. The mechanism for the stoichiometric coupling between chemical reaction and vectorial transport of ions is a universal question for ion-motive ATPases. The F₀ portion is a passive proton pathway but becomes a regulated pathway after the binding of F₁. Mutant analyses suggest that the γ subunit has regulatory role(s) for proton conduction.

When all the sequences of the γ subunits so far sequenced are aligned, 28 residues are mostly in the carboxyl terminal region (Fig. 11.8). Membranes with γ Thr-277→end, γ Gln-269→Leu, γ Glu-275→Lys or a frame-shift (downstream of γ Thr-277) mutation in the γ subunit had about 15% of the wild-type ATPase activity, but showed different extent of ATP-dependent proton translocation.¹⁰²⁾ The electrochemical gradient of H⁺ formed by γ Gln-275→Lys mutant membranes and the gradient formed was at least two times higher than that formed by γ Gln-269→Lue mutant (Fig. 11.9). This difference was not due to the non-specific leakiness of mutant membranes, since mutant membranes were able to form essentially similar degrees of electrochemical gradient dependent on respiration. The γ Met-23→Lys or Arg mutant showed more uncoupling phenotype between H⁺ transport and ATP hydrolysis/synthesis.¹⁰³⁾ The γ Lys-23 and γ Arg-23 mutant had essentially the same membrane ATPase active as the wild-type whereas they could grow only very slowly by oxidative phosphorylation. Membranes of the mutants were able to form much lower proton gradients. However, they had normal F₀ sectors, since their membranes could form similar proton gradients upon binding of wild-type F₁. These suggest that the γ subunit regulates proton conduction. It may also be possible that the γ subunit functions as gate for the proton pathway. We isolated second site mutations (mutations in the

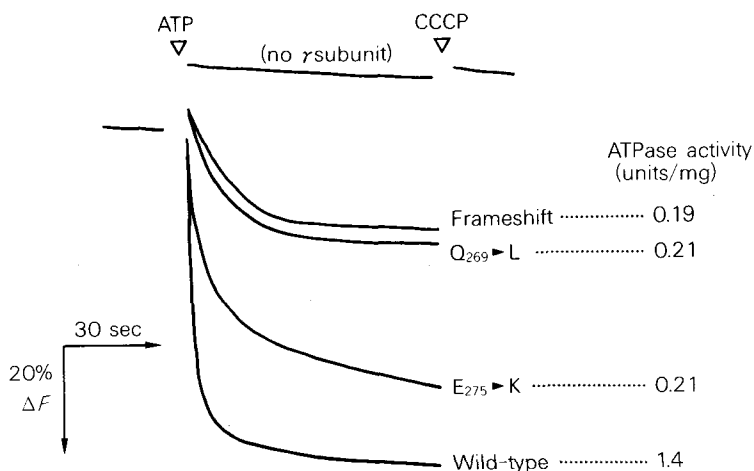


Fig. 11.9 Formation of proton gradient by membrane vesicles with different mutant γ subunits. Membrane vesicles (100 μ g protein) from strain KF10rA (γ Gln-14→end) harboring mutant or wild-type plasmids were suspended in buffer containing 1 μ M acridine orange, and fluorescence was monitored at 25°C. At the indicated times, 1 mM ATP or 1 μ M CCCP was added. Membrane ATPase activities of mutants and wild type are also shown.

carboxyl terminal region) which can suppress γ Met-23-Lys.¹¹⁷⁾ These results suggest that close interaction between the two regions (region near γ -Met-23 and carboxyl terminal) is required for efficient coupling between ATP synthesis/hydrolysis and proton transport.

11.5 Conclusion

This chapter summarized recent studies on proton-transporting ATP synthase (H⁺-ATPase) of *E. coli* carried out mainly in our laboratories. The catalytic site, catalytic cooperativity, H⁺ translocation, mechanism of coupling between catalysis and proton translocation, and subunit-subunit interaction (assembly) have been discussed at the level of amino acid residues. Combined results of mutagenesis and affinity labeling together with structural simulation can provide a picture of at least a part of the catalytic site of the enzyme. This time-consuming approach may still provide further interesting information. The crystal of F₁ is now being analyzed, although higher-ordered structures with visual amino acid residues are as yet unavailable. Another line of structural studies, the crystallization of F₀F₁ and F₀, may become possible because of the progress in rapid and easy preparation of F₀F₁. Location of the amino acid residues discussed in this chapter in higher ordered structures will be of great importance in understanding the mechanism and dynamic structure of the enzyme, once the crystal structure is determined.

Acknowledgments

We are grateful to Dr. T. Noumi for collaboration in the early part of the work discussed in this article. We thank Dr. Toshio Fukui for collaboration in affinity labeling of F₁ with AP₃-PL.

The most recent work from our laboratory was in collaboration with Dr. Robert K. Nakamoto, Dr. Hiroshi Omote, Dr. Yoshinori Moriyama, and students whose names appear in the references. Their contribution is greatly appreciated.

REFERENCES

1. P. L. Pedersen, E. Carafoli, *Trends Biochem. Sci.*, **12**, 146, 187 (1987).
2. M. Nakao, in: *New Era of Bioenergetics* (Y. Mukohata, ed.), Academic Press, Tokyo, New York, p. 1 (1991).
3. E. Racker, *A New Look at Mechanism in Bioenergetics*, Academic Press, New York (1976).
4. M. Futai, H. Kanazawa, *Microbiol. Review*, **47**, 285 (1983).
5. M. Futai, T. Noumi, M. Maeda, *Annu. Rev. Biochem.*, **58**, 111 (1989).
6. R. H. Fillingame, *The Bacteria*, **12**, 345 Academic Press, New York (1990).
7. A. E. Senior, *Annu. Rev. Biophys. Chem.*, **19**, 7 (1990).
8. J. E. Walker, M. Saraste, W. J. Gay, *Biochim. Biophys. Acta*, **768**, 164 (1984).
9. Y. Kagawa, *New Comp. Biochem.*, **9**, 149 (1984).
10. H. Strotmann, S. Bickel-Sandkötter, *Annu. Rev. Plant Physiol.*, **35**, 97 (1984).
11. H. Kobayashi, N. Murakami, T. Unemoto, *J. Biol. Chem.*, **257**, 13246 (1982).
12. Y. Matsubara, T. Hase, T. Hashimoto, K. Tagawa, *J. Biochem.*, **90**, 1159 (1981).
13. Y. Kakinuma, Y. Ohsumi, Y. Anraku, *J. Biol. Chem.*, **256**, 10859 (1981).
14. B. J. Bowman, E. J. Bowman, *J. Memb. Biol.*, **94**, 83 (1986).
15. M. Forgac, *Physiol. Rev.*, **69**, 765 (1988).
16. Y. Moriyama, N. Nelson, *Biochim. Biophys. Acta*, **980**, 241 (1989).
17. L. Zimniak, P. Dittrich, L. P. Gogarten, H. Kiback, L. Taiz, *J. Biol. Chem.*, **263**, 9102 (1988).
18. E. J. Bowman, K. Tenney, B. J. Bowman, *J. Biol. Chem.*, **263**, 13994 (1988).

19. B. J. Bowman, R. Allen, M. A. Wechsler, E. J. Bowman, *J. Biol. Chem.*, **263**, 14002 (1988).
20. M. F. Manolson, B. F. F. Ouellette, M. Filon, R. J. Poole, *J. Biol. Chem.*, **263**, 127987 (1988).
21. N. Nelson, L. Taiz, *TIBS*, **14**, 113 (1989).
22. M. Mandel, Y. Moriyama, J. D. Hulmes, Y.-C. E. Pan, H. Nelson, N. Nelson, *Proc. Natl. Acad. Sci. USA*, **85**, 5521 (1988).
23. S. Birman, F.-M. Meunier, B. Lesbats, J.-P. Le Caer, J. Rossier, M. Israel, *FEBS Lett.*, **261**, 303 (1990).
24. L. Meagher, P. McLean, M. E. Finbow, *Nucleic Acids Res.*, **18**, 6712 (1990).
25. H. Nelson, N. Nelson, *FEBS Lett.*, **247**, 147 (1989).
26. H. Hanada, M. Hasebe, Y. Moriyama, M. Maeda, M. Futai, *Biochem. Biophys. Res. Commun.*, **176**, 1062 (1991).
27. E. Uchida, Y. Ohsumi, Y. Anraku, *J. Biol. Chem.*, **263**, 45 (1988).
28. H. Hanada, Y. Moriyama, M. Maeda, M. Futai, *Biochem. Biophys. Res. Commun.*, **170**, 873 (1990).
29. D. L. Foster, R. H. Fillingame, *J. Biol. Chem.*, **254**, 8230 (1979).
30. E. Schneider, K. Altendorf, *EMBO J.*, **4**, 515 (1985).
31. E. Schneider, K. Altendorf, *Microbiol. Rev.*, **51**, 477 (1987).
32. Y. Moriyama, A. Iwamoto, H. Hanada, M. Maeda, M. Futai, *J. Biol. Chem.*, **266**, 2214 (1991).
33. M. Bianchet, X. Ysern, J. Hüllihen, P. L. Pedersen, L. M. Amzel, *J. Biol. Chem.*, **266**, 21197 (1991).
34. M. Futai, *Biochem. Biophys. Res. Commun.*, **79**, 1231 (1977).
35. S. D. Dunn, M. Futai, *J. Biol. Chem.*, **255**, 113 (1980).
36. M. Hirano, K. Takeda, H. Kanazawa, M. Futai, *Biochemistry*, **23**, 1652 (1982).
37. Y. Kagawa, S. Ohta, Y. Otawara-Hamamoto, *FEBS Lett.*, **249**, 67 (1989).
38. K. Miwa, M. Yoshida, *Proc. Natl. Acad. Sci. USA*, **86**, 6484 (1989).
39. M. Kuki, T. Noumi, M. Maeda, T. Amemura, M. Futai, *J. Biol. Chem.*, **263**, 17437 (1988).
40. M. Jounouchi, M. Takeyama, T. Noumi, Y. Moriyama, M. Maeda, M. Futai, *Arch. Biochem. Biophys.*, **292**, 87 (1992).
41. M. Jounouchi, M. Takeyama, P. Chaiprasert, T. Noumi, Y. Moriyama, M. Maeda, M. Futai, *Arch. Biochem. Biophys.*, **292**, 376 (1992).
42. R. L. Cross, C. M. Nalin, *J. Biol. Chem.*, **257**, 2874 (1982).
43. J. G. Wise, T. M. Duncan, L. R. Latchney, D. N. Cox, A. E. Senior, *Biochem. J.*, **215**, 343 (1983).
44. H. Hanada, T. Noumi, M. Maeda, M. Futai, *FEBS Lett.*, **257**, 465 (1989).
45. C. Grubmeyer, R. L. Cross, H. S. Penefsky, *J. Biol. Chem.*, **257**, 12092 (1982).
46. R. L. Cross, C. Grubmeyer, H. S. Penefsky, *J. Biol. Chem.*, **257**, 12101 (1982).
47. T. M. Duncan, A. E. Senior, *J. Biol. Chem.*, **260**, 4901 (1985).
48. T. Noumi, M. Taniai, H. Kanazawa, M. Futai, *J. Biol. Chem.*, **261**, 9196 (1986).
50. T. Noumi, M. Futai, H. Kanazawa, *J. Biol. Chem.*, **268**, 10076 (1984).
51. S. Soga, T. Noumi, M. Takeyama, M. Maeda, M. Futai, *Arch. Biochem. Biophys.*, **268**, 643 (1989).
52. T. Noumi, M. Maeda, M. Futai, *FEBS Lett.*, **213**, 381 (1987).
53. M. Tommasion, R. A. Capaldi, *Biochemistry*, **24**, 3972 (1985).
54. P. D. Boyer, in: *Of Oxygen, Fuels and Living Matter* (Semenza, G., ed.) p. 229 (1981).
55. R. Rao, Q. E. Senior, *J. Biol. Chem.*, **262**, 17450 (1987).
56. M. Hollemans, M. J. Runswick, I. M. Fearnley, J. E. Walker, *J. Biol. Chem.*, **258**, 9307 (1983).
57. T. Noumi, M. Maeda, M. Futai, *J. Biol. Chem.*, **263**, 8765 (1988).
58. M. R. Webb, C. Grubmeyer, H. S. Penefsky, D. R. Trentham, *J. Biol. Chem.*, **255**, 11637 (1980).
59. W. W. Andrews, F. C. Hill, W. S. Allison, *J. Biol. Chem.*, **259**, 8219 (1984).
60. W. W. Andrews, F. C. Hill, W. S. Allison, *J. Biol. Chem.*, **259**, 14378 (1984).
61. J. Garin, F. Boulay, J. P. Issartel, J. Lunardi, P. V. Vignais, *Biochemistry*, **25**, 4431 (1986).
62. A. Admon, G. G. Hammes, *Biochemistry*, **26**, 3193 (1987).
63. D. A. Bullough, W. S. Allison, *J. Biol. Chem.*, **261**, 14171 (1986).
64. T. Noumi, M. Tagaya, K. Miki-Takeda, M. Maeda, T. Fukui, M. Futai, *J. Biol. Chem.*, **262**, 7686 (1987).
65. M. Tagaya, T. Noumi, K. Nakano, M. Futai, T. Fukui, *FEBS Lett.*, **233**, 347 (1988).
66. K. Ida, T. Noumi, M. Maeda, T. Fukui, M. Futai, *J. Biol. Chem.*, **266**, 5424 (1991).
67. W. Sachesenheimer, G. E. Schulz, *J. Mol. Biol.*, **114**, 23 (1977).
68. P. Dreusike, P. A. Karplus, G. E. Schulz, *J. Mol. Biol.*, **199**, 359 (1988).
69. A. M. de Vos, L. Tong, M. V. Milburn, P. M. Matias, J. Jancarik, J. Noguchi, S. Nishimura, K. Miura, E. Ohtsuka S.-H. Kim, *Science*, **239**, 888 (1988).
70. L. Tong, A. M. de Vos, M. V. Milburn, J. Jancarik, S. Noguchi, S. Nishimura, K. Miura, E. Ohtsuka, S.-H. Kim, *Nature*, **337**, 90 (1989).
71. H. Kanazawa, T. Kayano, T. Kiyasu, M. Futai, *Biochem. Biophys. Res. Commun.*, **105**, 1257 (1982).
72. S.-Y. Hsu, T. Noumi, M. Takeyama, M. Maeda, S. Ishibashi, M. Futai, *FEBS Lett.*, **218**, 222 (1987).
73. R. W. Sweet, S. Yokoyama, T. Kamata, J. R. Feramisco, M. Rosenberg, M. Gross, *Nature*, **311**, 273 (1984).
74. M. Tagaya, T. Yagami, T. Noumi, M. Futai, F. Kishi, A. Nakazawa, T. Fukui, *J. Biol. Chem.*, **264**, 990

- (1989).
75. M. Takeyama, K. Ihara, Y. Moriyama, T. Noumi, K. Ida, N. Tomioka, A. Itai, M. Maeda, M. Futai, *J. Biol. Chem.*, **265**, 21279 (1990).
 76. T. Iwamoto, H. Omote, H. Hanada, N. Tomioka, A. Itai, M. Maeda, M. Futai, *J. Biol. Chem.*, **266**, 16350 (1991).
 77. H. Omote, M. Maeda, M. Futai, *J. Biol. Chem.*, **267**, 20571 (1992).
 78. D. Parsonage, S. Wilke-Mounts, A. E. Senior, *J. Biol. Chem.*, **262**, 8023 (1987).
 79. J. Miki, K. Fujiwara, M. Tsuda, T. Tsuchiya, H. Kanazawa, *J. Biol. Chem.*, **265**, 21567 (1990).
 80. A. Iwamoto, M. Maeda, M. Futai, *J. Biol. Chem.*, **268**, 3156 (1993).
 81. M. Yoshida, W. S. Allison, F. S. Esch, M. Futai, *J. Biol. Chem.*, **257**, 10033 (1982).
 82. C. A. Kumamoto, R. P. Simoni, *J. Biol. Chem.*, **261**, 10037 (1986).
 83. C. A. Kumamoto, R. P. Simoni, *J. Biol. Chem.*, **262**, 3060 (1987).
 84. S. Eya, T. Noumi, M. Maeda, M. Futai, *J. Biol. Chem.*, **263**, 10056 (1988).
 85. J. Hoppe, P. Friedl, H. U. Schairer, W. Sebald, K. von Meyernburg, B. B. Jørgensen, *EMBO J.*, **2**, 105 (1983).
 86. D. S. Perlin, D. N. Cox, A. E. Senior, *J. Biol. Chem.*, **258**, 9793 (1983).
 87. J. Hoppe, C. Montecucco, P. Friedl, *J. Biol. Chem.*, **258**, 2882 (1983).
 88. A. C. G. Porter, C. A. Kumamoto, K. Aldape, R. D. Simoni, *J. Biol. Chem.*, **260**, 8182 (1985).
 89. D. A. Jans, L. Hatch, A. L. Fimmel, F. Gibson, G. B. Cox, *J. Bacteriol.*, **420** (1985).
 90. D. A. Jans, A. L. Fimmel, L. Hatch, F. Gibson, G. B. Cox, *J. Biochem.*, **221**, 43 (1984).
 91. S. Eya, M. Maeda, M. Futai, *Arch. Biochem. Biophys.*, **284**, 71 (1991).
 92. B. D. Cain, R. D. Simoni, *J. Biol. Chem.*, **261**, 10043 (1986).
 93. B. D. Cain, R. D. Simoni, *J. Biol. Chem.*, **263**, 6606 (1988).
 94. B. D. Cain, R. D. Simoni, *J. Biol. Chem.*, **264**, 3292 (1989).
 95. R. N. Lightowlers, S. N. Howitt, L. Hatch, F. Gibson, G. B. Cox, *Biochem. Biophys. Acta*, **894**, 399 (1987).
 96. R. N. Lightowlers, S. N. Howitt, L. Hatch, F. Gibson, G. B. Cox, *Biochem. Biophys. Acta*, **933**, 241 (1988).
 97. S. M. Howitt, F. Gibson, G. B. Cox, *Biochim. Biophys. Acta*, **936**, 74 (1988).
 98. S. B. Vik, B. D. Cain, K. T. Chunk, R. D. Simoni, *J. Biol. Chem.*, **263**, 6599 (1988).
 99. S. B. Vik, D. Lee, P. A. A. Marshall, *J. Bacteriol.*, **173**, 4544 (1991).
 100. W. Sebald, J. Hoppe, *Curr. Top. Bioenerg.*, **12**, 1 (1981).
 101. M. J. Miller, M. Oldenburg, R. H. Fillingame, *Proc. Natl. Acad. Sci. USA*, **87**, 4900 (1990).
 102. T. Noumi, C. Beltran, H. Nelson, N. Nelson, *Proc. Natl. Acad. Sci. USA*, **88**, 1938 (1991).
 103. A. Iwamoto, J. Miki, M. Maeda, M. Futai, *J. Biol. Chem.*, **265**, 5043 (1990).
 104. K. Shin, R. K. Nakamoto, M. Maeda, M. Futai, *J. Biol. Chem.*, **267**, 20835 (1992).
 105. N. Inohara, A. Iwamoto, Y. Moriyama, S. Shimomura, M. Maeda, M. Futai, *J. Biol. Chem.*, **266**, 7333 (1991).
 106. J. Miki, M. Maeda, Y. Mukohata, M. Futai, *FEBS Lett.*, **232**, 221 (1988).
 107. M. L. Yu, R. B. Selman, *J. Biol. Chem.*, **263**, 19342 (1988).
 108. L. A. Cozens, E. J. Walker, *J. Mol. Biol.*, **194**, 359 (1987).
 109. S. Werner, J. Schumann, H. Stromann, *FEBS Lett.*, **261**, 204 (1990).
 110. F. D. McCarn, A. R. Whitaker, J. Alam, M. J. Vrba, E. S. Curtis, *J. Bacteriol.*, **107**, 3448 (1988).
 111. G. Falk, A. Hampe, E. J. Walker, *Biochem. J.*, **228**, 391 (1985).
 112. J. L. V. Tybulewicz, G. Falk, G. Falk, E. J. Walker, *J. Mol. Biol.*, **179**, 185 (1984).
 113. H. Kanazawa, T. Kayano, K. Mabuchi, M. Futai, *Biochem. Biophys. Res. Commun.*, **103**, 604 (1981).
 114. R. L. Krumholz, U. Esser, D. R. Simoni, *Nucleic Acids Res.*, **17**, 7993 (1989).
 115. A. S. W. Brusilow, A. M. Scarpetta, A. C. Haethorne, P. W. Clark, *J. Biol. Chem.*, **264**, 1528 (1989).
 116. S. Ohta, M. Yoshida, M. Ishizuka, H. Hirata, T. Hamamoto, Y. Otawara-Hamamoto, K. Matsuda, Y. Kagawa, *Biochim. Biophys. Acta*, **933**, 141 (1988).
 117. E. J. Walker, M. I. Fearnly, J. N. Gay, W. B. Gibson, D. F. Northrop, J. S. Powell, J. M. Runswick, M. Saraste, J. L. V. Tybulewicz, *J. Mol. Biol.*, **184**, 677 (1985).
 118. R. K. Nakamoto, M. Maeda, M. Futai, *J. Biol. Chem.*, **268**, 867 (1993).

12

Stabilization of Proteases by
an Engineered Disulfide Bond

Hiroshi Matsuzawa and Takahisa Ohta

*Department of Agricultural Chemistry, The University of Tokyo,
Bunkyo-ku, Tokyo 113, Japan*

Oligonucleotide-directed site-specific mutagenesis is a powerful technique for the study of proteins and has brought about great progress in research on the structure and function of proteins. This technique makes it possible in theory to obtain any mutant enzyme desired. One of the goals of such protein engineering is to make the proteins more stable.

The disulfide bond differs from other types of interactions in folded proteins, such as hydrogen bonds and hydrophobic, electrostatic and van der Waals interactions. The disulfide bond is a covalent bond that is able to significantly stabilize folded conformations by 2–5 kcal/mol for each disulfide.¹⁾ The effect is presumed to be due mainly to a decrease in the configurational chain entropy of the unfolded polypeptide.²⁾ On the other hand, another view is that the disulfide bond destabilizes folded structures entropically, but stabilizes them enthalpically to a greater extent.³⁾

Since there are strict stereochemical requirements for the relative positions and orientations of the two participating cysteine residues,¹⁾ addition of new disulfides to existing proteins by site-directed mutagenesis has not always produced the desired increase in stability. Introduction of disulfide bonds has been attempted for phage T4 lysozyme,^{4–7)} phage λ repressor,⁸⁾ dihydrofolate reductase,⁹⁾ and subtilisins.^{10–13)} Among them the most extensive study has been performed on T4 lysozyme, and enhancement of protein stability has been successful.

Natural thermostable enzymes are expected to be a good model for engineering stabilization of proteins. Study of thermostable enzymes from thermophilic microorganisms has revealed that the hydrophobicity of hydrophobic core inside the protein molecule and the electrostatic interactions of amino acid residues within the folded protein seem to be the cause of their stability. The enzymes from thermophilic microorganisms known so far do not contain the disulfide bond, but aqualysin I. Aqualysin I is an extracellular protease while the others are intracellular enzymes.

As for subtilisin BPN', the first attempt at molecular modeling of disulfide mutants was performed with computer graphics using coordinates from the crystal structure of the enzyme,^{10–12)} but increasing enzyme stability was unsuccessful. We have studied a thermostable subtilisin-type protease, aqualysin I, and the introduction site of a disulfide bond was chosen on structural homology between aqualysin I and subtilisin E. Here we describe a successful study to increase the stability of subtilisin E and others done for subtilisin enzymes.

12.1 Aqualysin I is a Thermostable Protease

Aqualysin I is an alkaline serine protease, extracellularly secreted by *Thermus aquaticus* YT-1,¹⁴⁾ an extremely thermophilic, Gram-negative bacterium. The enzyme is thermostable. Optimum temperature for the proteolytic activity of aqualysin I is 80°C in the presence of 1 mM Ca^{2+} and 70°C in its absence (Fig. 12.1).¹⁵⁾ 85% of the original activity remains after treatment at 80°C for 3 h in the presence of 1 mM Ca^{2+} , but only about 20% remains in its absence.¹⁵⁾ These results indicate that calcium ion is essential for the structural stability of the enzyme.

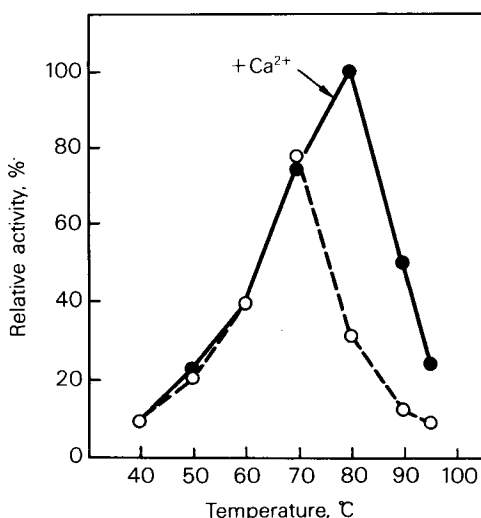


Fig. 12.1 Effect of temperature on the proteolytic activity of aqualysin I.¹⁵⁾ The enzyme activity was assayed at the indicated temperatures in the presence (●) and absence (○) of 1 mM CaCl_2 . (Reproduced with permission from *European Journal of Biochemistry*, 171, 445 (1988)).

12.2 Aqualysin I Belongs to a Family of Subtilisin-type Serine Proteases

The gene encoding aqualysin I has been cloned and sequenced.¹⁶⁾ The amino acid sequence deduced from the nucleotide sequence indicates that aqualysin I is produced as a large precursor and eventually processed to a mature enzyme. The mature enzyme consists of 281 amino acid residues, and the molecular weight is 28,350. The amino acid sequence has high homology with those of the subtilisin-type serine proteases (Fig. 12.2); 43% identity with proteinase K,¹⁷⁾ 37–39% with subtilisins,^{18–21)} and 34% with thermitase.²²⁾ Extremely high sequence identity is observed in the regions containing the catalytic residues of active site, Asp32, His64, and Ser221, the residue numbers corresponding to those of subtilisin BPN' (Fig. 12.2).

The large precursor of aqualysin I comprises four structurally distinguishable domains: an N-terminal signal peptide (14 amino acid residues), an N-terminal pro-sequence (113

	10	20	30	40	50	
AQU:	ATQSPAPWGLDRIDQRLDPLSNSYTYTATGRGVNVYVIDTGIRTTTHREF	G	GRAR			
PRO:	-A-TN-----A--SSTSPGT-TY-YDESA-Q-SC-----EAS-P--	E	---	QM		
THE:	YTPNDPYFS- RQY-PQK-Q	APQAW DI-E-S-AKIAIV--VQSN-PDLA-KVV-GWDF				
BPN:	- -- V-Y-VSQ-K	APALHSQGY--SN-K-A--S--DSS-PDL	KVA-G-SM			
E :	- -- V-Y-ISQ-K	APALHSQGY--SN-K-A--S--DSS-PDL	NVR-G-SF			
	10	20	30	40	50	
	60	70	80	90	100	110
AQU:	VGYDALGGNGQDCNGHGTHVAGTIG	GVTYGVAKAVNLYAVRVLDNCGSGSTSGVIAGV				
PRO:	-KTYYYSS R-G-----C---V-	S R-----KTQ-FG-K---D-----QY-TI---M				
THE:	-DN-STP -NG-----C--IAAAVTNNST-I	A-T-PKAS-L-----NS---TWTA-AN-I				
BPN:	-PSETN PF--N-S-----VAAL NNSI--	L--PSAS---K--GAD---QY-WI-N-I				
E :	-PSETN PY--GSS-----AAL NNSI--	L--SPSAS---K---ST---QY-WI-N-I				
	60	70	80	90	100	110
	120	130	140	150	160	
AQU:	DWV T RNHRPA	VANMSLGGGV	STALDNAVKNSIAAGVVYAVAAGNDNANACNYS	PA		
PRO:	-F-ASDK-N-NCPKGV--SL-----YS-SVNSA-A	RLQSS--MV-----N--D-R--S	--			
THE:	TY AAD QGAK	-ISL-----T-GNSG-QQ--NYAWN-K-S-VVA---AGNT-P--	--			
BPN:	E- AIA -NMD	-I-----PSG-A--KA--DKAV-S---VVA---EGTSGSSSTVGY-G				
E :	E- AIS -NMD	-I-----PTG---KTV-DKAVSS-I-V-A---EGSSGSTSTVGY--				
	120	130	140	150	160	
	170	180	190	200	210	220
AQU:	RVAEALTVGATTSSDARASFSNYGSCVDLFAPGASIPSAWYTS	SDTATQTTLNGTSMATPHVAGVAA				
PRO:	SEPSVC-----SDRY-R-S-----VL-I-G--T--L-T-IGG	S -RSIS-----L--				
THE:	Y-SN-IA-AS-DQN-NKS---T---V--VA---SW-Y-TYP--	- YAS-S-----G				
BPN:	KYPSVIA---VD--NQ-----SV-PEL-VM---V--Q-TLPGN	K YGAY-----S-----A--				
E :	KYPSTIA---VN--NQ-----SA--EL-VM---V--Q-TLPGG	- YGAY-----A--				
	170	180	190	200	210	220
	240	250	260	270	280	
AQU:	LYLEQNPSATPAS	VASAILNGATT	GRLSGI	GSGSPNRLLYSLLSSGSG		
PRO:	YLMTLGKTTA --ACR Y-ADT-NK -D--N-	PF-TV-L-A-NNYQA				
THE:	-LAS-GR- --NIRA--E-T-DKISG	T-TYWAK-RV- A-KAVQY				
BPN:	-I-SKH-NW- NTQ-R-SLE-TT-KL-D	-FYY -K-LI-VQA AAQ				
E :	-I-SKH-TW- NAQ-RDRLEST--YL-N	-FYY -K-LI-VQA AAQ				
	240	250	260	270		

Fig. 12.2 Comparison of the amino acid sequence of aqualysin I with those of other subtilisin-type proteases. The sequence of aqualysin I⁽⁶⁾ (AQU) is shown compared with those of proteinase K⁽⁷⁾ (PRO), thermolysin⁽²²⁾ (THE), subtilisin BPN⁽¹⁸⁾ (BPN), and subtilisin E⁽¹⁹⁾ (E). Identical amino acids with those of aqualysin I are shown by hyphen (-). Open space is the position where a corresponding amino acid is absent. The numbering above the sequences refers to aqualysin I, and that below the sequences to subtilisins. Asterisks indicate the active-site residues, Asp, His, and Ser.

residues), the protease domain and a C-terminal pro-sequence (105 residues).²³⁾ Such a four-domain structure is also found in the precursor of *Achromobacter* protease I,²⁴⁾ a mammalian-type serine protease. Two types of precursors of bacterial extracellular proteases having a three-domain structure are known as well; the precursors of subtilisins,¹⁸⁻²¹⁾ proteinase K,¹⁷⁾ and α -lytic protease²⁵⁾ have an N-terminal signal sequence and an N-terminal pro-sequence, and the precursors of proteases produced by *Neisseria*

*gonorrhoeae*²⁶⁾ and *Serratia marcescens*²⁷⁾ have an N-terminal signal sequence and a C-terminal pro-sequence.

In an *Escherichia coli* expression system for the aqualysin I precursor, the precursor is processed autoproteolytically into the mature 28-kDa enzyme by treatment at 65 °C.²³⁾ In this case, the N-terminal pro-sequence is required for the production of active enzyme and functions to stabilize the precursor structure.²⁸⁾ The C-terminal pro-sequence is not essential for the production of active aqualysin I,²⁹⁾ but seems to be involved in the translocation of the precursor across the cytoplasmic membrane.³⁰⁾ In a *Thermus thermophilus* expression system,³¹⁾ the C-terminal pro-sequence is required for the production and extracellular secretion of active aqualysin I.³²⁾ In an *E. coli* expression system for the subtilisin E gene, the N-terminal pro-sequence is essential for the production of active enzyme,³³⁾ as in the case of aqualysin I. The requirement of the pro-sequence is also shown *in vitro* for the refolding of the inactive mature protein to an active enzyme.^{34,35)} The functions of the N-terminal pro-sequences of aqualysin I and subtilisin E seem to be similar.

12.3 Comparison of Known Three-dimensional Structures of Subtilisin-type Serine Proteases: Probable Structural Similarity between Aqualysin I and Other Subtilisin-type Enzymes

Three-dimensional structures of four subtilisin-type enzymes, subtilisin BPN',^{36,37)} subtilisin Carlsberg,^{37,38)} thermitase,^{39,40)} and proteinase K,^{40,41)} are known, but that of aqualysin I has not yet been determined. The C α atoms of the known structures were superimposed to obtain maximal overlap of the backbone structures, and large parts of all four structures overlap very well (Fig. 12.3).⁴²⁾ On the basis of such analyses, structurally equivalent core residues (194 residues) are identified, and higher sequence identity was found to correspond to a closer overlap of mainchain atoms in the core (Table 12.1).⁴²⁾

As for aqualysin I, its high sequence identity in the core residues with the four proteases (Table 12.1) suggests that the three-dimensional structure of aqualysin I is similar to those of the proteases. The similarity to subtilisins and proteinase K is probably higher than to thermitase.

12.4 Aqualysin I Contains Two Disulfide Bonds

In the family of subtilisin-type serine proteases, primary sequences of about 40 enzymes are known.⁴²⁾ Among them, no subtilisin produced by *Bacillus* species has cysteine residues. Aqualysin I (four cysteine residues per molecule),¹⁶⁾ proteinase K (five residues),¹⁷⁾ and thermitase (one residue)²²⁾ are cysteine-containing enzymes (Fig. 12.2).

Aqualysin I contains two disulfide bonds which are formed between Cys67 and Cys99, and between Cys163 and Cys194 (Fig. 12.2).⁴³⁾ Proteinase K contains one free cysteine residue which is located at the corresponding position to the free cysteine residue of thermitase, and the other four cysteine residues form two disulfide bonds between two N-terminal-side residues and between two other C-terminal-side residues.⁴⁴⁾ Positions of the disulfide bond-forming cysteine residues of the two enzymes are different in the primary and

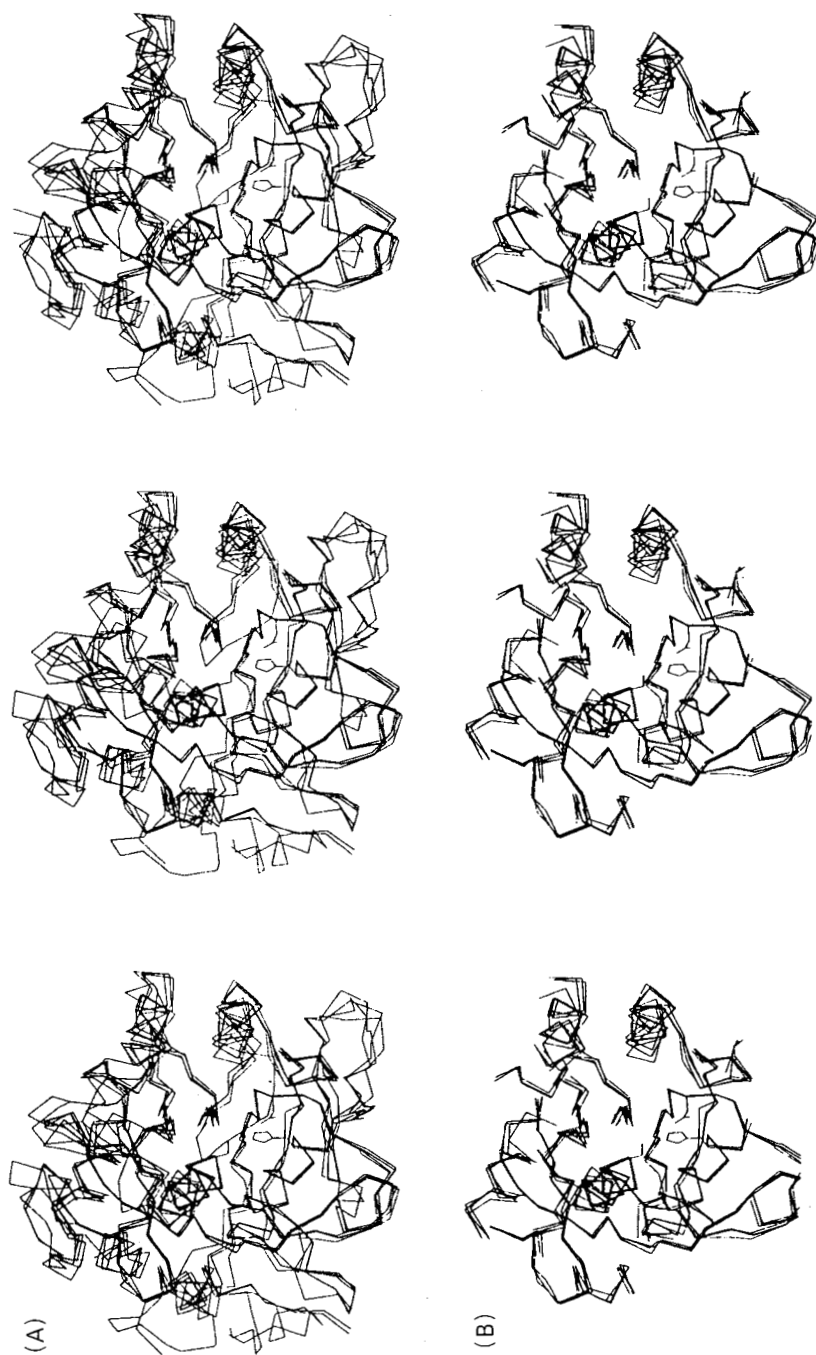


Fig. 12.3 Stereo-diagram of the superposition of the C_{α} -atom backbones of subtilisin BPN', subtilisin Carlsberg, thermolysin, and proteinase K.⁴²⁾ Only the side chains of catalytic triad residues, Asp32, His64, and Ser221 (numbering of subtilisin BPN'), are shown. (A) Complete backbones, and (B) core of 194 residues defined as the structurally conserved regions. (Reproduced with permission from Siezen, *Protein Engineering*, 4, 723 (1991)).

TABLE 12.1 Root-mean-square deviation values (Å) for the superposition of C α atoms (upper right) and amino acid sequence identity (%) (lower left) of the 194 core residues between subtilisin BPN', subtilisin Carlsberg, thermitase, and proteinase K,⁴²⁾ together with the sequence identity (%) of aqualysin I with the proteases

	Subtilisin BPN'	Subtilisin Carlsberg	Thermitease	Proteinase K
Subtilisin BPN'	—	0.405	0.646	0.888
Subtilisin Carlsberg	75	—	0.621	0.940
Thermitease	52	56	—	0.949
Proteinase K	44	42	41	—
Aqualysin I	49	51	42	55

three-dimensional structures.

At present, 16 cysteine-containing subtilisin-type enzymes are known and the position of the cysteine residues is restricted to the nine corresponding sites described above.⁴²⁾ Of the 16 enzymes, six enzymes other than aqualysin I and proteinase K have cysteine residues at positions where the cysteine residues are able to form disulfide bond(s) like the two enzymes. Although these disulfide bonds seem to have been acquired to increase protein stability, only four kinds of disulfide bonds are found in the subtilisin-type enzymes, suggesting that the positions of the disulfide bonds have been selected strictly in the process of molecular evolution of the enzyme.

12.5 Introduction of a Disulfide Bond into Subtilisin by Site-directed Mutagenesis

As for aqualysin I, the disulfide bond seems to be one of the causes of its thermostability, because the enzyme is unstable in the presence of 2-mercaptoethanol.¹⁵⁾ Accordingly, increasing the stability of subtilisin E was attempted by engineering disulfide bonds at the sites corresponding to those of aqualysin I.

Amino acid residues of subtilisin E corresponding to the four cysteine residues of aqualysin I are Gly61, Ser98, Ser161, and Glu195 in the primary structure (Fig. 12.2). These residues were replaced with cysteine by site-directed mutagenesis for the subtilisin E gene and mutant enzymes produced in the *E. coli* expression system were analyzed.¹³⁾

12.5.1 Mutant Subtilisin E Having a Cys61-Cys98 Disulfide Bond Gains Thermostability

Mutant subtilisin E, which has two replacements of Gly61 and Ser98 with cysteine residues, forms a Cys61-Cys98 disulfide bond. Kinetic constants, K_m , k_{cat} and k_{cat}/K_m , of the disulfide mutant is almost identical to those of the wild-type enzyme (Table 12.2), suggesting that the active-site structure is unaffected by the introduction of the disulfide bond. In contrast, k_{cat} s of the single-cysteine mutants are greatly lowered, although their K_m s are rather decreased. The disulfide mutant is much more active than wild-type enzyme at about 60°C, while both enzymes show no difference in optimum temperature for the activity (about 50°C).

Thermostability of subtilisin E is increased by the introduction of the disulfide bond (Table 12.3). The half-life of the disulfide mutant is 2–3 times longer than that of the wild-type enzyme at 45°–60°C. On the other hand, those of the single-cysteine mutants are

TABLE 12.2 Kinetic constants of wild-type, Cys61–Cys98 disulfide, and single-cysteine mutant subtilisin Es¹³⁾

Enzyme	K_m (mM)	k_{cat} (sec ⁻¹)	k_{cat}/K_m (sec ⁻¹ /mM)
Wild type	1.9	21	11
Cys61–Cys98	2.2	23	10
Cys61	1.5	11	7.3
Cys98	1.4	6.0	4.3

Assays were carried out in 50 mM Tris-HCl (pH 8.5) and 1 mM CaCl₂ at 37°C using a synthetic peptide, succinyl-L-Ala-L-Ala-L-Pro-L-Phe-*p*-nitroanilide.

TABLE 12.3 Effect of the disulfide bond introduced in subtilisin E on the half-life of the enzyme activity due to thermal inactivation¹³⁾

Enzyme	Half-life at		
	45°C	55°C	60°C
	(hour)	(min)	(min)
Wild type	4.4	31	13
Cys61–Cys98	10.2	81	55
Cys61		37	
Cys98		40	

After incubation of the enzymes for various times at the indicated temperatures in the presence of 1 mM CaCl₂, remaining activity was measured by assaying as described for Table 12.2. Half-life was determined from semi-log plots of log₁₀[remaining activity] *versus* time.

essentially identical to that of the wild-type enzyme.

The increased thermostability of the disulfide mutant is also shown by denaturation temperature estimated by differential scanning calorimetry (Fig. 12.4). T_m (midpoint temperature in the thermally induced transition from the folded to the unfolded state) of the mutant enzyme is 63.0°C, which is 4.5°C higher than that of the wild-type enzyme (58.5°C). Under reducing conditions, the T_m of the mutant enzyme is decreased to 50.5°C, which is lower than that of the wild-type enzyme, indicating that the disulfide bond formation is required for the thermostability and that the mutant enzyme with free cysteine residues does not have a stable structure like the wild-type enzyme.

Identity of amino acid sequences between subtilisins E and BPN' is 86%, so three-dimensional structures of the two enzymes are considered to be very similar. In the case of subtilisin BPN', residues 61 and 98 are located on the loop and turn structure, respectively, both of which connect β -strand and α -helix (Fig. 12.5). Solvent exposures of the residues are both 9,⁴⁵⁾ indicating their presence on the surface of the enzyme molecule. The distance between the α -carbons of the two residues is 5.8 Å. Accordingly, the positions seem appropriate for cysteine residues to form a disulfide bond without any strain in the enzyme structure. The disulfide bond formed is located close to the active site so as to stabilize the wall of the active-site pocket (Fig. 12.5).

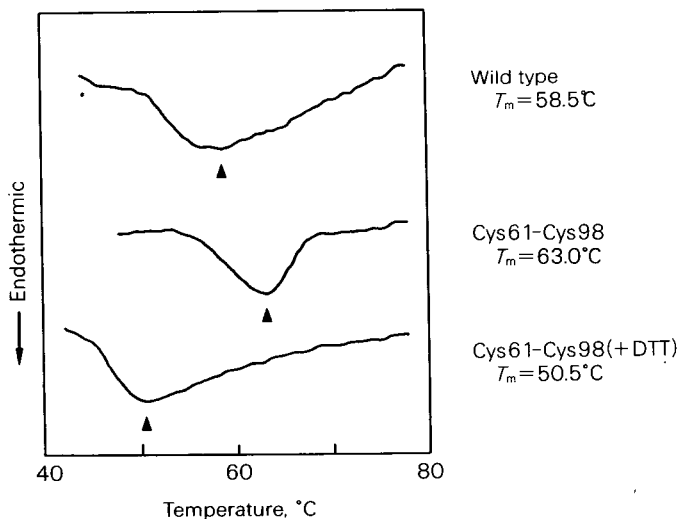


Fig. 12.4 Differential scanning calorimetry data for wild-type enzyme and Cys61-Cys98 mutant subtilisin E.¹³⁾ The results were obtained in the presence of 1 mM CaCl_2 and 1 mM phenylmethanesulfonyl fluoride. The reduced form of the disulfide mutant was prepared by the addition of 100 mM dithiothreitol (DTT). The samples were scanned at a heating rate of $0.6^\circ\text{C}/\text{min}$ from 20 to 90°C . Arrowheads indicate the T_m (midpoint of denaturation temperature) of the enzyme. (Reproduced with permission from Takagi *et al.*, *J. Biol. Chem.*, **265**, 6877 (1990)).

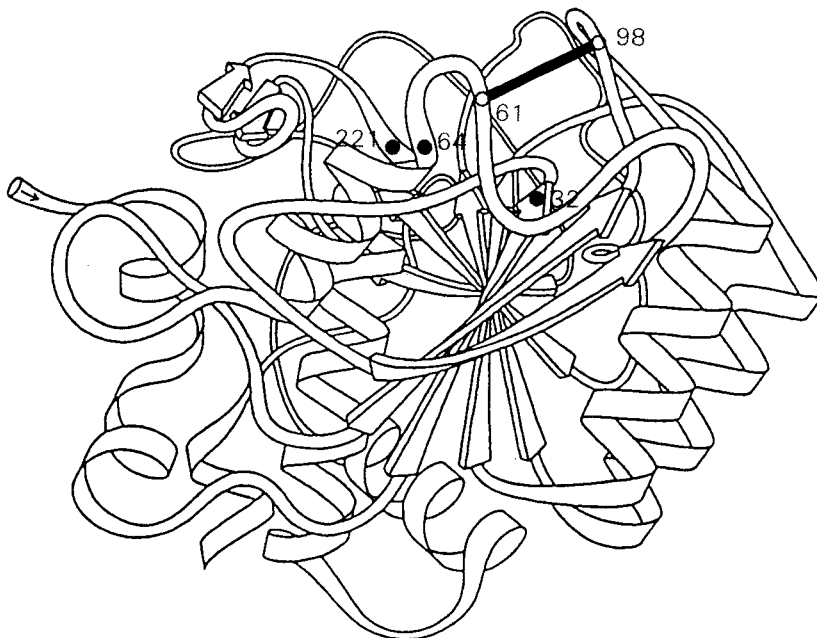


Fig. 12.5 Position of Cys61-Cys98 disulfide bond introduced into subtilisin E. Ribbon diagram of subtilisin⁵¹⁾ is shown with the positions of substituted cysteines and active-site residues, Asp32, His64, and Ser221. The disulfide bond is represented by a dark line. (Reproduced with permission from Richardson, *Adv. Protein Chem.*, **34**, 266 (1981)).

12.5.2 Another Mutant Did Not Form Disulfide Linkage

In the mutant subtilisin E with Cys161 and Cys195, no disulfide bond was detected. By another alignment of the sequences,⁴²⁾ it is Ser163 of subtilisin E, and not Ser161, to which Cys163 of aqualysin I corresponds. It is possible that the original alignment was wrong.

12.6 Other Attempts to Increase the Thermostability of Subtilisin are Unsuccessful

As for subtilisin BPN', besides the molecular modeling of disulfide mutants, Cys22–Cys87, Cys24–Cys87, Cys26–Cys232, Cys36–Cys210, and Cys41–Cys80, through computer graphics using coordinates from the crystal structure of the enzyme,^{10–12)} engineering of a disulfide bond, Cys29–Cys119 or Cys148–Cys243, was carried out on the basis of structural homology to the disulfides of proteinase K.¹²⁾ In the seven cases, disulfide bonds are formed, and all of the mutant enzymes except Cys148–Cys243 mutant have virtually the same kinetic constants as the wild-type enzyme.^{10–12)} However, increase in enzyme stability is not observed in any case.

Calcium is known to stabilize subtilisins,⁴⁶⁾ and two Ca^{2+} -binding sites are observed for subtilisin BPN'.³⁷⁾ In the case of the mutant subtilisin BPN' having a Cys22–Cys87 disulfide bond, the disulfide has an effect on the enzyme stability only in the absence of Ca^{2+} .¹¹⁾ In the presence of CaCl_2 , however, the effect is not observed.^{10,11)} It is likely that the binding of calcium results in conformational changes that introduce strain at the disulfide bond, thereby diminishing its effectiveness in stabilization. In the case of the mutant subtilisin E having Cys61–Cys98, the effect of the disulfide bond on the thermostability is observed in the presence of Ca^{2+} ,¹³⁾ suggesting that the structure stabilized by calcium is further stabilized by the disulfide bond.

The X-ray structures of four disulfide mutants of subtilisin BPN' with Cys22–Cys87, Cys24–Cys87, Cys26–Cys232 or Cys29–Cys119, have been analyzed.⁴⁷⁾ Each of the disulfide bonds exhibits atypical sets of dihedral angles compared with known structures of natural disulfide bridges in proteins and affects its local structural environment to a different extent. The disulfide located in buried regions, Cys26–Cys232 and Cys29–Cys119, induce larger changes than do Cys22–Cys87 and Cys24–Cys87, which are located on the surface of the enzyme molecule. In the latter cases, the disulfide formation is accommodated by relatively minor changes in local main-chain conformation.⁴⁸⁾

12.7 Another Example, an Alkaline Protease from *Aspergillus oryzae*

On the basis of sequence alignment, subtilisin-type proteases can be subdivided into class I and class II.⁴²⁾ Subtilisins, thermolysin and others, none of which has a disulfide bond, belong to class I, and ten proteases including aqualysin I and proteinase K to class II. An alkaline protease from *Aspergillus oryzae*, which has no cysteine residue, belongs to class II. The sequence identity between aqualysin I and the alkaline protease is 44%.⁴⁹⁾

Disulfide mutants of the alkaline protease were constructed on the basis of the positions of two disulfide bonds of aqualysin I. Cys169–Cys200 disulfide bond, which corresponds to Cys163–Cys194 of aqualysin I, increases the thermostability of the enzyme, but the other

N-terminal side disulfide bond, Cys69–Cys101, has no effect.⁴⁹⁾ On the other hand, another mutant enzyme which has the two disulfide bonds together shows higher thermostability than the mutant with Cys169–Cys200.⁵⁰⁾ The effect of the engineered disulfide bond(s) on the thermostability is presumed to be due to the increase in activation enthalpy overcoming a concurrent increase in activation entropy during thermal denaturation of the enzyme.⁵¹⁾

12.8 Conclusions

Because of the strict stereochemical requirements, it is not easy to find optimal sites for the introduction of disulfide bonds into proteins. Introduction of disulfide bonds into T4 lysozyme has been engineered by theoretical calculations and computer modeling.^{4–7)} The results obtained from the mutant lysozymes illustrate several points relevant to the use of disulfide bonds for improving protein stability.⁶⁾ (i) Introduction of the cysteine(s) should minimize the disruption or loss of interactions that stabilize the native structure. (ii) The size of the loop formed by the crosslink should be as large as possible. (iii) The strain energy introduced by the disulfide bond should be kept as low as possible. For this purpose, a location within the flexible part of the molecule is desirable.

We have described a protein engineering study on increasing the stability of proteases by the introduction of disulfide bond(s) as observed in naturally occurring thermostable enzymes. Expectations for protein engineering are certainly in general use. At present, however, computer graphics cannot always present the correct modeling to improve protein stability, as seen in the case of subtilisin BPN'. Therefore, it is useful to imitate naturally occurring enzymes for protein engineering. The results thus obtained will be used for further progress in protein engineering.

REFERENCES

1. T.E. Creighton, *BioEssays*, **8**, 57 (1988).
2. D.C. Poland, H.A. Scheraga, *Biopolymers*, **3**, 379 (1965).
3. A.J. Doig, D.H. Williams, *J. Mol. Biol.*, **217**, 389 (1991).
4. L.J. Perry, R. Wetzel, *Science*, **226**, 555 (1984).
5. R. Wetzel, L.J. Perry, W.A. Baase, W.J. Becktel, *Proc. Natl. Acad. Sci. USA*, **85**, 401 (1988).
6. M. Matsumura, W.J. Becktel, M. Levitt, B.W. Matthews, *Proc. Natl. Acad. Sci. USA*, **86**, 6562 (1989).
7. M. Matsumura, G. Signor, B.W. Matthews, *Nature*, **342**, 291 (1989).
8. R.T. Sauer, K. Hehir, R.S. Stearman, M.A. Weiss, A. Jeitler-Nilsson, E.G. Suchanek, C.O. Pabo, *Biochemistry*, **25**, 5992 (1986).
9. J.E. Villafranca, E.E. Howell, S.J. Oatley, N.-H. Xuong, J. Kraut, *Biochemistry*, **26**, 2182 (1987).
10. J.A. Wells, D.B. Powers, *J. Biol. Chem.*, **261**, 6564 (1986).
11. M.W. Pantoliano, R.C. Ladner, P.N. Bryan, M.L. Rollence, J.F. Wood, T.L. Poulos, *Biochemistry*, **26**, 2077 (1987).
12. C. Mitchinson, J.A. Wells, *Biochemistry*, **28**, 4807 (1989).
13. H. Takagi, T. Takahashi, H. Momose, M. Inouye, Y. Maeda, H. Matsuzawa, T. Ohta, *J. Biol. Chem.*, **265**, 6874 (1990).
14. H. Matsuzawa, M. Hamaoki, T. Ohta, *Agric. Biol. Chem.*, **47**, 25 (1983).
15. H. Matsuzawa, K. Tokugawa, M. Hamaoki, M. Mizoguchi, H. Taguchi, I. Terada, S.-T. Kwon, T. Ohta, *Eur. J. Biochem.*, **171**, 441 (1988).
16. S.-T. Kwon, I. Terada, H. Matsuzawa, T. Ohta, *Eur. J. Biochem.*, **173**, 491 (1988).
17. F.A. Gunkel, H.G. Gassen, *Eur. J. Biochem.*, **179**, 185 (1989).
18. J.A. Wells, E. Ferrari, D.J. Henner, D.A. Estell, E.Y. Chen, *Nucleic Acids Res.*, **11**, 7911 (1983).

19. M.L. Stahl, E. Ferrari, *J. Bacteriol.*, **158**, 411 (1984).
20. M. Jacobs, M. Eliasson, M. Uhlén, J.-I. Flock, *Nucleic Acids Res.*, **13**, 8913 (1985).
21. R. Kaneko, N. Koyama, Y.-C. Tsai, R.-Y. Juang, K. Yoda, M. Yamasaki, *J. Bacteriol.*, **171**, 5232 (1989).
22. B. Meloun, M. Baudyš, V. Kostka, G. Hausdorf, C. Frömmel, W.E. Höhne, *FEBS Lett.*, **183**, 195 (1985).
23. I. Terada, S.-T. Kwon, Y. Miyata, H. Matsuzawa, T. Ohta, *J. Biol. Chem.*, **265**, 6576 (1990).
24. T. Ohara, K. Makino, H. Shinagawa, A. Nakata, S. Norioka, F. Sakiyama, *J. Biol. Chem.*, **264**, 20625 (1989).
25. J.L. Silen, C.N. McGrath, K.R. Smith, D.A. Agard, *Gene*, **69**, 237 (1988).
26. J. Pohlner, R. Halter, K. Beyreuther, T.F. Meyer, *Nature*, **325**, 458 (1987).
27. N. Yanagida, T. Uozumi, T. Beppu, *J. Bacteriol.*, **166**, 937 (1986).
28. Y.-C. Lee, Y. Miyata, I. Terada, T. Ohta, H. Matsuzawa, *Agric. Biol. Chem.*, **55**, 3027 (1991).
29. Y.-C. Lee, T. Ohta, H. Matsuzawa, *FEMS Microbiol. Lett.*, **92**, 73 (1992).
30. Y.-C. Lee, I. Terada, Y. Miyata, T. Ohta, H. Matsuzawa, (Manuscript in preparation).
31. N. Touhara, H. Taguchi, Y. Koyama, T. Ohta, H. Matsuzawa, *Appl. Environ. Microbiol.*, **57**, 3385 (1991).
32. Y.-C. Lee, H. Koike, H. Taguchi, T. Ohta, H. Matsuzawa, (Manuscript in preparation).
33. H. Ikemura, H. Takagi, M. Inouye, *J. Biol. Chem.*, **262**, 7859 (1987).
34. H. Ikemura, M. Inouye, *J. Biol. Chem.*, **263**, 12959 (1988).
35. X. Zhu, Y. Ohta, F. Jordan, M. Inouye, *Nature*, **339**, 483 (1989).
36. C.S. Wright, R.A. Alden, J. Kraut, *Nature*, **221**, 235 (1969).
37. C.A. McPhalen, M.N.G. James, *Biochemistry*, **27**, 6582 (1988).
38. W. Bode, E. Papamokos, D. Musil, *Eur. J. Biochem.*, **166**, 673 (1987).
39. A.V. Teplyakov, I.P. Kuranova, E.H. Harutyunyan, B.K. Vainshtein, C. Frömmel, W.E. Höhne, K.S. Wilson, *J. Mol. Biol.*, **214**, 261 (1990).
40. C. Betzel, A.V. Teplyakov, E.H. Harutyunyan, W. Saenger, K.S. Wilson, *Protein Engng.*, **3**, 161 (1990).
41. C. Betzel, G.P. Pal, W. Saenger, *Eur. J. Biochem.*, **178**, 155 (1988).
42. R.J. Siezen, W.M.de Vos, J.A.M. Leunissen, B.W. Dijkstra, *Protein Engng.*, **4**, 719 (1991).
43. S.-T. Kwon, H. Matsuzawa, T. Ohta, *J. Biochem.*, **104**, 557 (1988).
44. K.-D. Jany, G. Lederer, B. Mayer, *FEBS Lett.*, **199**, 139 (1986).
45. W. Kabsch, C. Sander, *Biopolymers*, **22**, 2577 (1983).
46. G. Voordouw, C. Milo, R.S. Roche, *Biochemistry*, **15**, 3716 (1976).
47. B. Katz, A.A. Kossiakoff, *Proteins Struct. Funct. Genet.*, **7**, 343 (1990).
48. B.A. Katz, A. Kossiakoff, *J. Biol. Chem.*, **261**, 15480 (1986).
49. K. Ikegaya, Y. Ishida, K. Murakami, A. Masaki, N. Sugio, K. Takechi, S. Murakami, H. Tatsumi, Y. Ogawa, E. Nakano, H. Motai, H. Kawabe, *Biosci. Biotech. Biochem.*, **56**, 326 (1992).
50. K. Ikegaya, Y. Ishida, K. Murakami, S. Murakami, H. Tatsumi, E. Nakano, H. Motai, H. Kawabe, (Manuscript in preparation).
51. J.S. Richardson, *Adv. Protein Chem.*, **34**, 167 (1981).

Index

- α/β pair color plate ②
 absorption spectra 172
 abzyme 9
 acetyl group 26
 acid-base catalyst 95, 122
 acid-base mechanism 122
 actin filament 60
 activation by cation 182
 active site 116
 active site residue 94, 129
 acyl amino acid-releasing enzyme 26
 adenosine diphosphopyridoxal 75
 adenosine tetraphosphopyridoxal 75
 adenosine triphosphopyridoxal 75, 218
 adenylate kinase 78, 219
 affinity electrophoresis 119, 120
 affinity labeling 8, 73, 116
 Ala-P (phosphonate analog of alanine) 161, 162
 aldamine 172
 aldimine 95, 99, 172
 aldimine bond 90
 D-alanine 147, 148
 alanine racemase 147, 150, 155
 allosteric regulation 118, 122
 D-amino acid 147
 amino acid analyzer 4
 D-amino acid transaminase 95, 96
 amino acid composition analysis 25
 amino acid hydantoin 6
 amino terminal 5
 α -aminoacrylate intermediate 174
 (1-aminoethyl)phosphonic acid 161
 aminopeptidase M 26
 AMP 107, 118
 amylopectin 108, 119
 amyloplast 108–110, 123
 amylose 108, 119
 anhydrotrypsin 27
 anisotropy 175
 anthranilate synthase 64
 apotryptophanase 167
 aqua-lysin I 230
 Arg-266 94
 Arg-292 94
 Arg-292* 99, 100, 101, 102
 Arg-386 94, 99, 101
 AroAT 103
 Asn-194 94
 Asp-222 94, 98, 99
 aspartase 62
 aspartate proteinase 8
 aspartate transcarbamoylase 55, 59, 63
 AspAT (aspartate aminotransferase) 17, 87, 90, 93, 96
 ATP 217
 automated amino acid analyzer 24
 automated sequencer 22, 28
 8-azido-ATP 218
 bacterial expression 119
 β -barrel structure 198
 blocked amino-terminal 26
 ^{13}C NMR spectra 40
 C-terminal end determination 26
 Ca^{2+} -binding site 237
 Ca^{2+} -transporting ATPase 83
 caged group 8
 cAMP-dependent protein kinase 65
 carboxyl terminal 5
 carboxypeptidase 5
 carboxypeptidase Y 27
 cAspAT (cytosolic AspAT) 17, 90, 92, 93, 101, 103
 catalytic base 96
 catalytic intermediate 172
 catalytic mechanism 142, 185, 204
 catalytic property 182
 cation binding site 182
 CD (circular dichroism) 174
 cDNA 110
 cDNA cloning 30
 CE (capillary electrophoresis) 24
 chaperonin 68
 chemical modification 178
 chimeric phosphorylase 119
 β -chloroalanine 158, 159
 chloroplast 109
 circular dichroism 134, 174
 CNBr 22, 24
 coenzyme reorientation 183
 column chromatography 2
 computer graphics 237
 conformational change 130, color plate ④
 consensus sequence for nucleotide-binding 78
 creatine kinase 61
 cross-linking analysis 57
 cryo-X-ray crystallography 8
 crystallization 167
 CuZn-SOD 191, 192

- CuZn-SOD isozyme, chloroplast type of 192
 cyclodextrin 118
 Cys residues 101, 102
 cysteine proteinase 8
 cytoplasm 108
- Dansyl-Edman method 27
 database 32
d-d transition 200
 Debye's equation 202
 denaturation 235
 diffusion-controlled reaction 202
 diffusion-limited reaction 202, 205
 difluoroalanine 160
 dismutation rate of O₂⁻ 205
 disproportionation of O₂⁻ 195, 203
 dissociation 167
 disulfide bond 229
 domain 47, 152–154
- E. coli* 194
E. coli B/1t7-A 165
E. coli F₁-ATPase 81
 eAroAT (*E. coli* aromatic amino acid transaminase) 102, 103
 eAspAT (*E. coli* AspAT) 17, 24, 28, 29, 90, 92, 93, 96, 101–103, color plate ③
 EC-SOD (extracellular CuZn-SOD) 195
 Edman degradation 22, 25, 27, 28
 eglin c 49
 electrophile 121
 electrostatic guidance 205
 β-elimination reaction 133
 enantiomer analysis 4
 energy transfer 177
 engineered disulfide bond 229
 enolimine tautomer 172
 enzyme application 10
 enzyme kinetics 7
 enzyme
 number of 1
 substrate specificity of 7
 enzyme structure 3
 evolution 122, 192
 exon 122
 expression system 119
 external aldimine 187
- F₁-ATPase 81
 Fe-SOD 191, 192
 Fenton reaction 208
 F₀F₁ 213, 214
 flash photometric method 203
 flexible loop 130
 fluoescamine 25
 fluorescence 177
 fluorescence energy transfer 66
 8-fold α/β barrel 128
 folding domain 130
 fragmentary alanine racemase 152, 154
 fragmentary enzyme 153, 154
 FSBA (fluorosulfonylbenzoyl adenosine) 74
- FSBI (fluosulfonylbenzyl inosine) 82
 FSH-RH (LH-RH), 33
 gas phase sequencer 5, 22, 29
gem-diamine 187
 α-glucan phosphorylase 107, 114
 glucose 1-phosphate 121
 glycine-rich sequence 219
 glycogen 107, 108, 116
 glycogen-storage site 116–119, 122
 glycogen synthase 76
 GroE 170
 GUS (β-glucuronidase) 113
- H⁺-ATPase 56, 211
 H⁺-ATPase, F-type 211, 213, 217
 P-type 211, 218
 V-type 212
 half-life of disulfide mutant 234
 His-143 98, 99
 histidine 208
 holotryptophanase 167
 HPLC (high performance liquid chromatography) 2
 hydrazinolysis 6
- in vitro* mutagenesis 94
 indole channeling 138
 indole tunnel 138
 induced-fit 7
 inhibitor-proteinase complex 39
 inner sphere mechanism 207
 internal aldimine (Schiff base) 187
 intron 122
 irreversible aggregation 170
 isozyme 108
 Type L 109
 Type H 111
- K₂O₂ (potassium superoxide) 203
- LD (linear dichroism) 183
 Leu-18 103
 L-leucine dehydrogenase 3
 ligand-induced conformational change 139
 limited proteolysis 123, 152
 liquid-phase sequencer 28
 lock and key theory 7
 lutropin 59
 Lys-258 90, 94–96
 lysylendopeptidase 22
- maltodextrin 108
 mAspAT (mitochondrial AspAT) 17, 22, 24, 28, 29, 31, 90, 92, 93, 101, 103
 mature protein 31
 α-Me-Asp (methylaspartate) 95
 mechanism-based inactivator 160
 4*N*-methanesulfonic acid 25
 Met residue, three 43
 methyl viologen 194
 Michaelis complex 52
 Michaelis intermediate 46

- Michaelis-Menten parameter 205
 missense mutation 129
 Mn-SOD 191, 192
 modified amino-terminal 26
 molecular chaperone 170
 molecular modeling 237
 molecular weight 3
 M⁷³R SSI 51
 mRNA 112
 MS (mass spectrometer) 32
 MS/MS (tandem mass spectrometer) 32
 multi-site 216
 multienzyme complex 127

 Na⁺,K⁺-ATPase 83
¹⁵N-¹³C-double labeling method 39
 N-end rule 112
 ninhydrin 24
 NMR 37
¹H-NMR 178
 nucleophilic attack 122
 nucleoside polyphosphopyridoxal 74
 nucleotide-binding site 74

 one-base mechanism 156
 one-electron reduction 202
 OPA (*O*-phthalaldehyde) 25
 outer sphere mechanism 207
 overproductive cloned cell 167
 oxidant 208

³¹P chemical shift 177
 patatin 112
 pBWU13, recombinant plasmid 214
 PDI (protein disulfide isomerase) 69
 peptidoglycan 147, 148
 phosphoglycerate mutase 56
 phosphorolysis 122
 phosphorylase isozyme 109
 phosphorylase kinase, 83
 phosphorylation 107, 118
 photoinactivation 181
 phylogenetic distribution 192
 plant phosphorylase 123
 plastid 108
 PLDP-glucose (pyridoxal-diphosphoglucose) 80
 PLP (pyridoxal 5'-phosphate) 87, 107, 147, 148, 151, 155, 159
 PLP-dependent enzyme 87, 165
 poly(A) tail 111
 polyadenylation signal 110, 112
 potato tuber 108
 PPIase (peptidyl prolyl isomerase) 69
 precursor protein 31
 primary structure 166
 of AspAT 17
 of protein 5
 of tryptophanase 166
 proline- β -naphthylamidase 58
 proplastids 110
 protease 230
 protein engineering 229

 protein folding 131
 proteinase K 230
 proteolipid 212
 α -proton abstraction 95, 96, 98
 1,3-prototropic shift 95, 96
 pseudorevertant 220
 PTC (phenylthiocarbamate derivatives)-amino acid 5
 PTH (phenylthiohydantoin)-amino acid 5, 22, 27, 28
 pulse radiolysis 203, 207
 purification, enzyme 1
 push-pull general acid-base catalysis 187
 pyridoxal diphosphate 79
 pyridoxal-P 107, 116, 120, 122, 133
 pyroglutamyl peptidase 26

 quaternary structure 55, 166, 169
 quinonoid intermediate 97, 174, 187

 rabbit muscle phosphorylase 113, 117, 123
 rabbit skeletal muscle 108
 rapid-scanning stopped-flow UV-visible spectroscopy 133
ras protein 219
 rate-limiting step 187
 reactive site 38
 reciprocal communication 141
 recombinant DNA technology 30
 refolding 170
 renaturation process 61
 β -replacement reaction 133
 78-residue insertion 109, 113, 118, 119, 122
 reverse double labeling experiment 41
 ribozyme 9
 ribulose biphosphate carboxylase/oxygenase 63

 Schiff base 90, 96, 97
 scissile bond 39, 46
 secondary structure 6
 sequence homology 17, 90, 230
 sequencer (sequenator) 5
 serine preteinase 8
 site-directed mutagenesis 122, 133, 178, 205, 229
 small-angle X-ray scattering 183
 SOD (superoxide dismutase) 191
sodA gene 194
sodB gene 194
 solid phase sequencer 28
 spectral property 172
 spinning cup sequencer 28
 SSI* (modified SSI) 46
 SSI (*Streptomyces* subtilisin inhibitor) 37
 SSI-subtilisin BPN' complex 42, 44
 starch 108, 123
 stereospecificity of enzyme 7, 10
 stopped-flow method 203
Streptomyces 52
 structural domain 131
 structural mechanism 7
 subcellular localization 123
 substrate recognition 101

- substrate specificity of enzyme 7
- subtilisin 230
- subtilisin-type serine protease 230
- subtilisin variants 49
- subunits (*a*, *b*, *c*) 221, 224
- β subunit 217
- γ subunit 225
- subunit association 143
- subunit dissociation 150, 151
- subunit interface 58
- suicide substrate 8, 160
- superoxide anion radical, O_2^- 191, 202
- swinging door mechanism 158, 159
- synzyme 11

- TEM (transmission electron microscopy) 57
- tetrahedral intermediate 45
- thermitase 230
- thermodynamic mechanism 7
- Thermophilic bacillus* AspAT 19, 30
- thermostability 234
- thermostable alanine racemase 150
- thermostable enzyme 229
- three-dimensional structure 6, 127, 232
 - of protein 6
 - of subtilisin-type serine protease 232
 - of tryptophan synthase 127
- time-resolved X-ray crystallography 8
- T_m (midpoint temperature) 235

- transient intermediate 185
- transit peptide 110–113, 122, 123
- transition-state analogue 9
- trifluoroalanine 159, 160, 161
- Trp-140 94, 99, 103
- trypsin-like enzyme 51
- tryptophan 208
- L-tryptophan indole-lyase 165
- tryptophan synthase 3, 127, color plate ①
- tryptophanase 68, 165, 166
- two-base mechanism 156
- Tyr-225 94, 97
- Tyr-70 94
- Tyr-70* 95–97
- tyrosine phenol-lyase 169, 188
- tyrosyl-tRNA synthetase 66

- UDP-glucose pyrophosphorylase 78
- unc* (*atp*) operon 213
- unfolding 150, 151, 170
- unit-site 216
- 5'-upstream region 112
- uridine diphosphopyridoxal 75

- V8 protease 22

- X-ray crystallographic analysis 19, 120
- X-ray crystallographic structure of AspATs 93
- X-ray crystallography 129