

An International Series of Monographs and Textbooks

Protein Crystallography

T.L. Blundell and L.N. Johnson

PROTEIN CRYSTALLOGRAPHY

(内部交流)

T.L. BLUNDELL

*School of Biological Sciences, University of Sussex
Falmer, Brighton, Sussex, England*

and

L.N. JOHNSON

*Laboratory of Molecular Biophysics, University of Oxford
South Parks Road, Oxford, England*



ACADEMIC PRESS

NEW YORK • LONDON • SAN FRANCISCO

A Subsidiary of Harcourt Brace Jovanovich, Publishers

ACADEMIC PRESS INC. (LONDON) LTD.
24/28 Oval Road,
London NW1

United States Edition published by
ACADEMIC PRESS INC.
111 Fifth Avenue
New York, New York 10003

Copyright © 1976 by
ACADEMIC PRESS INC. (LONDON) LTD.

All Rights Reserved

No part of this book may be reproduced in any form by photostat, microfilm, or
any other means, without written permission from the publishers

Library of Congress Catalog Card Number: 76-1068
ISBN: 0 12 108350-0

Typeset in IBM Press Roman by PREFACE LTD., SALISBURY
and printed in Great Britain by
WHITSTABLE LITHO LTD.,
WHITSTABLE, KENT

PROTEIN CRYSTALLOGRAPHY

MOLECULAR BIOLOGY

An International Series of Monographs and Textbooks

Editors: BERNARD F. HUTCHINSON, NATHAN O. KAMMAN, JUDITH A. KAMMAN

and HAROLD A. SCHERAGA

A complete list of titles in this series appears at the end of this volume.

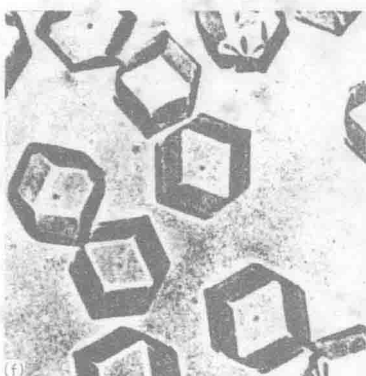
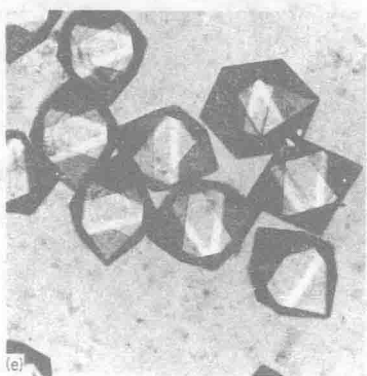
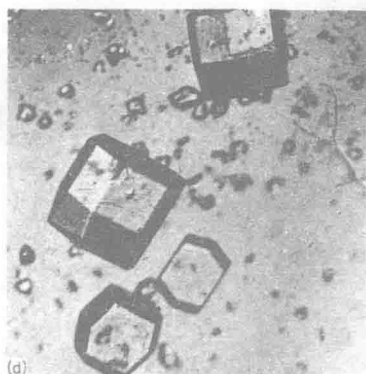
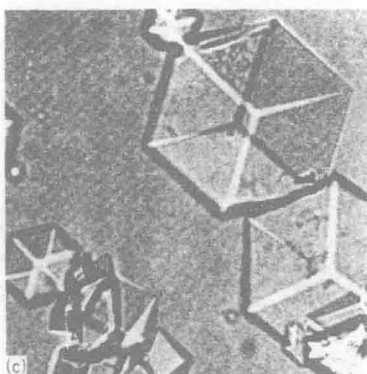
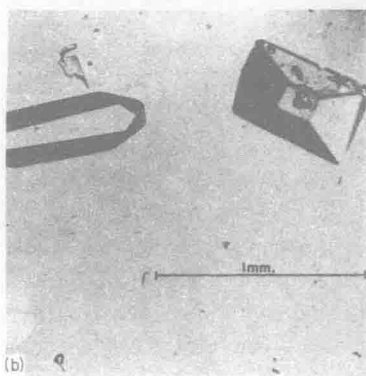
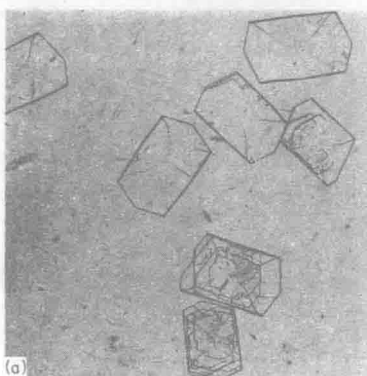
MOLECULAR BIOLOGY

An International Series of Monographs and Textbooks

**Editors: BERNARD HORECKER, NATHAN O. KAPLAN, JULIUS MARMUR,
AND HAROLD A. SCHERAGA**

A complete list of titles in this series appears at the end of this volume.

Protein Crystals. a) monoclinic aldolase. b) orthorhombic triosephosphate isomerase. c) trigonal insulin. d) tetragonal lysozyme. e) hexagonal aldolase. f) cubic glucagon.



PREFACE

The scope of this book is simply defined: we have tried to cover all aspects of the X-ray diffraction analysis of protein crystals, with special emphasis on the differences between the analysis of protein crystals and the analysis of crystals of small molecules. In several cases we have had a difficult task in selection of material. In general we have limited our detailed descriptions to those procedures which have contributed to successful structure determinations. We also include chapters on neutron diffraction, γ -ray resonance and electron microscopy. The information obtained from the application of these techniques to protein crystals is complementary to that obtained by X-ray analysis and in recent years there have been considerable advances in these fields.

T. L. BLUNDELL

L. N. JOHNSON

April, 1976

ACKNOWLEDGEMENTS

We wish to acknowledge the debt we owe to our teachers, Professor D. C. Hodgkin and Professor D. C. Phillips. Their continued help and encouragement has been a great inspiration to us in our work.

We wish to thank the many colleagues who have generously helped us in the preparation of this book, in particular: Dr. R. Ashaffenberg, Dr. S. H. Banyard, Dr. J. Beaumont, Dr. C. W. Bunn, Dr. R. Diamond, Dr. G. G. Dodson, Dr. P. A. M. Eagles, Dr. P. R. Evans, Mr. J. A. Jenkins, Dr. K. R. Leonard, Dr. S. Mason, Dr. J. Mosley, Dr. S. Neidle, Professor, A. C. T. North, Dr. K. W. Snape, Dr. I. J. Tickle, Miss I. T. Weber, Dr. K. S. Wilson, Dr. S. Wood, Dr. D. G. R. Yeates and Professor M. Zeppezauer.

We are most grateful to those authors, publishers and learned societies who allowed us to reproduce published figures and photographs and due acknowledgement is made in each case in the text.

Finally we wish to thank Gail Gibbs and Reiko Sakuma for typing the manuscript.

CONTENTS

| | |
|---|-----------|
| Preface | vii |
| Acknowledgements | xv |
| 1. Introduction | 1 |
| 1.1. Introduction | 1 |
| 1.2. The Nature and Function of Globular Proteins | 2 |
| 1.3. The Development of Protein Crystallography and its Limitations | 10 |
| 1.4. Steps in the X-ray Analysis of Crystalline Proteins | 13 |
| 2. The Principles of Protein Structure | 18 |
| 2.1. Introduction | 18 |
| 2.2. Amino Acids — the Building Blocks | 18 |
| 2.3. The Primary Structure | 25 |
| 2.4. The Nature of the Peptide Link | 25 |
| 2.5. Peptide Structure | 28 |
| 2.6. Secondary Structure of Polypeptides | 31 |
| 2.7. Protein Tertiary Structure | 45 |
| 2.8. Quaternary Structure | 51 |
| 2.9. Protein Interactions with Small Molecules and Ions | 57 |
| 3. Crystallization of Proteins | 59 |
| 3.1. Introduction | 59 |
| 3.2. Factors affecting the Solubility of Proteins | 60 |
| 3.3. Factors affecting Crystal Nucleation and Growth | 66 |
| 3.4. Techniques for Crystallization | 68 |
| 3.5. Protein Crystals: the Solvent of Crystallization | 79 |
| 3.6. Crystal Mounting | 81 |
| 4. Symmetry, Space Groups and Optical Properties of Crystals | 83 |
| 4.1. Introduction | 83 |
| 4.2. The Thirty-two Classes of Symmetry | 86 |
| 4.3. The Seven Crystal Systems | 90 |
| 4.4. The Fourteen Bravais Lattices | 90 |

| | | |
|-------|--|------------|
| 4.5. | Internal Symmetry Elements | 93 |
| 4.6. | The two-hundred-and-thirty Space Groups | 94 |
| 4.7. | Optical Properties of Crystals | 97 |
| 4.8. | The Use of the Polarizing Microscope by the Protein Crystallographer | 103 |
| 4.9. | The Determination of Molecular Weight from Protein Crystals | 104 |
| 5. | The Principles of X-ray Diffraction | 107 |
| 5.1. | Introduction — the Wave Nature of X-rays | 107 |
| 5.2. | Bragg's Law | 109 |
| 5.3. | The Analogy with Light Scattering | 110 |
| 5.4. | Resolution | 117 |
| 5.5. | The Scattering of X-rays by an Atom | 119 |
| 5.6. | The Scattering of X-rays by a Molecule | 123 |
| 5.7. | Scattering of X-rays by a Crystal | 124 |
| 5.8. | The Relationship between the Laue Conditions for Diffraction and Bragg's Law. The Reciprocal Lattice | 130 |
| 5.9. | The Ewald Construction | 132 |
| 5.10. | Friedel's Law | 133 |
| 5.11. | The Electron Density Equation | 134 |
| 5.12. | The Phase Problem | 136 |
| 5.13. | The Patterson Function | 137 |
| 5.14. | Direct Methods | 140 |
| 5.15. | Refinement | 144 |
| 5.16. | Summary | 149 |
| 6. | Isomorphous Replacement | 151 |
| 6.1. | Introduction | 151 |
| 6.2. | Single Isomorphous Replacement | 153 |
| 6.3. | Centric Zones | 159 |
| 6.4. | Multiple Isomorphous Replacement | 160 |
| 6.5. | The Size of the Heavy Atom Contribution | 161 |
| 6.6. | Tests for Isomorphism | 162 |
| 7. | Anomalous Scattering | 165 |
| 7.1. | Introduction | 165 |
| 7.2. | The Dispersion and Absorption Effects | 165 |
| 7.3. | The Use of Anomalous Dispersion to Simulate Isomorphism | 168 |

| | | |
|-------|---|------------|
| 7.4. | The Breakdown of Friedel's Law | 170 |
| 7.5. | Fourier Syntheses for Structures with an Anomalous Scatterer | 175 |
| 7.6. | The Use of Anomalous Scattering Differences in the Determination of Phases | 177 |
| 7.7. | Choosing the Absolute Configuration using Anomalous Scattering | 180 |
| 8. | Preparation of Heavy Atom Derivatives | 183 |
| 8.1. | Introduction | 183 |
| 8.2. | Metal Ion Replacement in Metalloproteins | 184 |
| 8.3. | Heavy Atom Labelled Specific Inhibitors of Enzymes and Carrier Proteins | 188 |
| 8.4. | Replacement of an Amino Acid with a Heavy Atom Labelled Analogue | 191 |
| 8.5. | Direct Binding of the Heavy Atom Salts | 192 |
| 8.6. | Result of Direct Binding of Heavy Atom Reagents | 199 |
| 8.7. | The Introduction of New Functional Groups as Heavy Atom Binding Sites | 235 |
| 8.8. | Techniques for Carrying out Modification of Protein Crystals | 236 |
| 8.9. | Warning | 238 |
| 9. | Data Collection | 240 |
| 9.1. | Introduction | 240 |
| 9.2. | X-Rays | 242 |
| 9.3. | Factors which affect the Measurement of Integrated Intensities | 246 |
| 9.4. | Cylindrical Polar Co-ordinates | 257 |
| 9.5. | Photographic Methods | 260 |
| 9.6. | Diffractionmeters | 284 |
| 10. | Data Processing | 310 |
| 10.1. | Introduction | 310 |
| 10.2. | Primary Data Processing: Photographic Data | 310 |
| 10.3. | Primary Data Processing: Diffractionmeter Data | 321 |
| 10.4. | Comparison of Equivalent Reflections | 331 |
| 10.5. | Scaling Different sets of Data | 331 |
| 10.6. | Absolute Scale and the Scaling of Derivative to Native Data | 333 |

| | |
|--|------------|
| 11. The Determination of Heavy Atom Positions | 337 |
| 11.1. Introduction | 337 |
| 11.2. Estimation of F_H | 338 |
| 11.3. Heavy Atom Vector Maps | 343 |
| 11.4. Difference Fourier Techniques for finding Heavy Atom Positions | 349 |
| 11.5. The Use of Direct Methods in finding Heavy Atom Positions | 351 |
| 11.6. Refinement of Atomic Parameters for Heavy Atoms | 352 |
| 11.7. Correlation of origins | 360 |
| 11.8. General Approach to finding Heavy Atom Positions | 361 |
| 12. The Calculation of Phases | 363 |
| 12.1. Introduction | 363 |
| 12.2. The Treatment of Errors in the Method of Isomorphous Replacement | 364 |
| 12.3. Inclusion of Anomalous Scattering | 371 |
| 12.4. The Determination of the Absolute Configuration of the Molecule | 373 |
| 12.5. A Simplified Representation of the Phase Probability Distributions, The Hendrickson-Lattman Method | 375 |
| 13. The Interpretation of the Electron Density Map | 381 |
| 13.1. Introduction | 381 |
| 13.2. Resolution | 381 |
| 13.3. Low Resolution Studies, $dm > 4.5\text{\AA}$ | 387 |
| 13.4. Medium Resolution Electron Density Maps $3.5\text{\AA} > dm > 2.5\text{\AA}$ | 389 |
| 13.5. High Resolution Electron Density Maps $dm < 2\text{\AA}$ | 389 |
| 13.6. The Display of Electron Density Maps and Model Building for Medium and High Resolution Studies | 390 |
| 13.7. Interpretation of Electron Density Maps at Medium and High Resolution | 399 |
| 13.8. Determination of the Amino Acid Sequence | 400 |
| 14. Difference Fouriers | 404 |
| 14.1. Introduction | 404 |
| 14.2. Peak Heights in Difference Fouriers | 407 |

| | | |
|--------|--|-----|
| 14.3. | Errors in Difference Fouriers | 409 |
| 14.4. | Conformational Changes | 411 |
| 14.5. | Other Types of Difference Fouriers | 416 |
| 15. | Refinement and Improvement of Resolution | 420 |
| 15.1. | Introduction | 420 |
| 15.2. | Techniques Available for the Refinement of Protein Models | 421 |
| 15.3. | Refinement of Atomic Positions | 422 |
| 15.4. | Examples of Refinement of Protein Structures | 432 |
| 15.5. | Direct Methods of Refinement of Protein Phases | 436 |
| 16. | Molecular Replacement | 443 |
| 16.1. | Introduction | 443 |
| 16.2. | Orientation of Molecules | 444 |
| 16.3. | Positioning the Oriented Molecule: the Translation Problem | 456 |
| 16.4. | Molecular Replacement and the Calculation of Phases | 462 |
| 17. | The Use of Neutrons and Gamma-rays in Protein Crystallography | 465 |
| 17.1. | Introduction | 465 |
| 17.2. | Generation of Neutron Beams | 465 |
| 17.3. | The Scattering of Neutrons by Atoms | 467 |
| 17.4. | Crystal Damage | 471 |
| 17.5. | Data Collection | 472 |
| 17.6. | Neutron Fouriers | 477 |
| 17.7. | Comparison of X-ray and Neutron Amplitudes | 478 |
| 17.8. | Isomorphous Replacement with Neutron Diffraction | 478 |
| 17.9. | Anomalous Scattering of Neutrons | 479 |
| 17.10. | Gamma-ray Resonance | 482 |
| 18. | Electron Microscopy | 483 |
| 18.1. | Introduction | 485 |
| 18.2. | The Principle and Design of the Electron Microscope | 487 |
| 18.3. | Preparation of Biological Specimens | 496 |
| 18.4. | Radiation Damage | 504 |
| 18.5. | High Voltage Electron Microscopy | 507 |

| | | |
|---------------|---|-----|
| 18.6. | Image Analysis | 509 |
| 18.7. | The High Resolution Transmission Scanning Electron Microscope (STEM) | 513 |
| 19. | Achievements of Protein Crystallography | 518 |
| REFERENCES | | 541 |
| SUBJECT INDEX | | 559 |
| 15. | Molecular Replacement | 443 |
| 15.1 | Introduction | 443 |
| 15.2 | Refinement of Molecular Models | 444 |
| 15.3 | Refining the Combined Molecular Model | 445 |
| 15.4 | Model Building and the Calculation of Phases | 447 |
| 16. | The Use of Neutrons and Synchrotrons in Protein Crystallography | 462 |
| 16.1 | Introduction | 462 |
| 16.2 | Generation of Neutron Beams | 463 |
| 16.3 | The Scattering of Neutrons by Atoms | 464 |
| 16.4 | X-ray Diffraction | 471 |
| 16.5 | Data Collection | 472 |
| 16.6 | Neutron Powder | 473 |
| 16.7 | Comparison of X-ray and Neutron Techniques | 478 |
| 16.8 | Isomorphous Replacement with Neutrons | 478 |
| 16.9 | Diffraction | 479 |
| 16.10 | Anomalous Scattering of Neutrons | 480 |
| 16.11 | Comparative Advantages | 482 |
| 17. | Electron Microscopy | 483 |
| 17.1 | Introduction | 483 |
| 17.2 | The Principle and Design of the Electron Microscope | 484 |
| 17.3 | Properties of Biological Specimens | 485 |
| 17.4 | Electron Damage | 486 |
| 17.5 | High Voltage Electron Microscopy | 487 |

INTRODUCTION

We ought to be neither like spiders which spin things out of their own insides nor like ants which merely collect but like bees which both collect and arrange.

Sir Francis Bacon

1.1 INTRODUCTION

Protein Crystallography is the application of the techniques of X-ray diffraction (and to a lesser extent, those of neutron, γ ray and electron diffraction) to crystals of one of the most important classes of biological molecules, the proteins. Proteins are ubiquitous molecules. They include the enzymes, the biological catalysts which control all chemical processes in living systems; the storage and carrier molecules such as haemoglobin which is responsible for the transport of oxygen in the blood; the hormones, which act as messengers between different parts of the organism; and the antibodies which provide immunity against infection. It is known that the diverse biological functions of these complex molecules are determined by and are dependent upon their three-dimensional structure and upon the ability of these structures to respond to other molecules by changes in shape. At the present time X-ray analysis of protein crystals forms the only method by which detailed structural information (in terms of the spatial co-ordinates of the atoms) may be obtained. The results of these analyses have provided firm structural evidence which, together with biochemical and chemical studies, immediately suggests proposals concerning the molecular basis of biological activity.

The success of these physical techniques in describing biological structure and function does not mean that biology has become a branch of physics. As Bernal (1969) has written: "The use of physical or chemical knowledge to explain mechanical, electrical or chemical aspects of living organisms has brought into greater relief their biological aspects. These phenomena however well they could be described in physical terms do not occur in mechanisms made by some divine craftsman from ideal models laid down from all eternity, but in self-regulating

and self-producing entities whose present form is the result of an evolution stretching back over billions of years".

Protein Crystallography has advanced extremely rapidly in the last decade and in certain cases protein structures may now be determined relatively easily using standard methods. (Previously such analyses had proved difficult and time consuming.) Many of the techniques which were being tentatively tried out a few years ago are now well established; other approaches have still to prove fruitful and in certain areas there is a need for new ideas. This recent knowledge is scattered in many different journals and doctoral theses or is simply passed by word of mouth from one generation of research students to the next. It seemed to us that the subject has now reached a stage of maturity when a detailed review in the form of a book is needed. Our aim has been to draw together in a systematic and comprehensible manner those methods that are now well established and to indicate those areas where further expansion is needed.

The subject attracts workers from many different disciplines, such as chemistry, biochemistry, biophysics, zoology, physics and mathematics, and the different insights and skills contributed by these workers are vitally important in this multi-disciplinary subject. The book is primarily intended for those who wish to do research in the field. However, we have attempted to write it in a style which will be comprehensible to workers from these various disciplines.

The principles of protein structure and the implications of the X-ray analyses of biological molecules form important parts of many undergraduate courses in both biological and physical sciences. We hope that this book will also prove useful to undergraduates especially where a deeper insight into the scope and limitations of the technique than is usually provided in review articles is required.

In this introductory chapter we first survey the biochemistry of the proteins which have been and are being studied by protein crystallographers and which recur frequently in our discussions of techniques. We then review the history of our subject and provide a synopsis of the special problems encountered in protein crystallography and the means by which they have been overcome. Thus this chapter to some extent forms a resumé of the book.

1.2 THE NATURE AND FUNCTION OF GLOBULAR PROTEINS

We have already seen that globular proteins may have many roles in organisms including biological catalysis and regulation of cellular processes. It was in fact their wide range of biological functions and their consequent importance in living systems which led to the name "proteins" which is derived from a Greek word meaning first. We shall see in Chapter 2 that proteins are characterized by their hydrolysis to amino acids in dilute solutions of mineral acids at boiling point. In fact all proteins are built up from linear polymers of amino acids called

polypeptide chains. They may be divided into two classes, fibrous proteins and globular proteins, on the basis of their structure and function.

Fibrous proteins are relatively insoluble in aqueous solutions. These include the keratins of wool, hair and horn, the myosin of muscle, the fibroin of silk and the collagen of tendon. These fibrous proteins have a structural role. They are either protective materials or they are nature's ropes and girders. The fibrous nature and the insolubility of these materials make crystallization with good three dimensional order almost impossible. Consequently the information we can gain about their structure is much less precise and the techniques of their structure analysis are quite different from that of globular proteins. The techniques are reviewed by Holmes and Blow (1966); we do not intend to discuss them in this volume.

In globular proteins the polypeptide chain has a complex folded structure. Most globular proteins are soluble in aqueous solutions, but others which are normally located in or on membrane structures will be more soluble in nonpolar solvents. The difficulty of purifying and stabilizing the membrane proteins has meant that most proteins studied by X-ray analysis at high resolution are water soluble proteins and so our present knowledge of structure is really rather unrepresentative of globular proteins as a whole. Further, the crystallographer has tended to study the more stable proteins and these have often been extracellular proteins.

Let us consider in more detail the globular proteins which have been the subject of X-ray analysis. For the reader who has no biochemical training we recommend one of the standard texts such as "Biological Chemistry" by Mahler and Cordes (1968) or "Biochemistry" by Lehninger (1975). Here we will try to outline the biological roles of some of the most intensely studied proteins. We hope that this will give the reader some impression of the purpose of the structural study when we are discussing the techniques of protein crystallography.

Almost all of the proteins studied by X-ray analysis can be classified in one of the five following groups:

1. Enzymes
2. Redox proteins
3. Carrier and storage proteins
4. Hormones
5. Antibodies

Let us consider enzymes first.

Enzymes are catalysts; they alter the rate of biochemical reactions without altering the position of equilibrium. They are important in all biological processes. They are classified according to the kind of reactions catalysed; the classification is illustrated in Table 1.1 with the enzymes which are the subject of discussion in this book.