ENZYMOLOGY VOLUME 3

Enzyme Structure and Function

S. Blackburn

Enzyme Structure and Function

S. Blackburn

Leeds
England

MARCEL DEKKER, INC. New York and Basel

PREFACE

The search for an understanding of the chemical basis of enzymic catalysis forms a most important part of current biochemical research and in Great Britain in 1971 enzyme chemistry was selected as an area of science worthy of special encouragement. The study is being conducted by protein chemists, x-ray crystallographers and nuclear magnetic resonance experts using a variety of techniques. Protein crystallographers in particular have developed a whole range of weapons which can be employed to attack problems of increasing biochemical complexity.

A wealth of information on many enzymes is becoming available. General structural and catalytic features are emerging as the primary sequences and three-dimensional structures of a considerable number of enzymes are being unravelled. The x-ray analyses and other studies are providing answers to questions which have nuzzled enzymologists for a long time and from a preoccupation with the tactics of structure determination investigators are beginning to devote their attention to learning about the control processes of living systems. Structure analysis is becoming a tool and emphasis is now on function.

It is thus possible to consider the relationship between structure and function of enzymes at the molecular level. The present book discusses the relatively few enzymes where structural and other investigations have advanced to the point where this relationship is reasonably clear. These enzymes, however, cover a fairly wide spectrum both with respect to their action mechanisms and the substrates whose attack they catalyze. They are probably representative of a large number, possibly the majority, of enzymes. Remarkable similarities in structural features of enzymes of widely different origins, but of similar function, have been discovered. The evidence on which our knowledge of the structure and function of these enzymes is based is discussed in the book, which will interest biochemists, protein chemists, and scientists generally who wish to understand how enzyme structure is related to the ability to catalyze reactions.

My thanks are due to several workers who very kindly helped me with reprints, photographs, and advice, often giving up much of their valuable time. Among these should be mentioned: Dr. C. B. Anfinsen, Prof. M. L. Bender, Dr. D. M. Blow, Dr. I. M. Chaiken, Prof. F. A. Cotton, Dr. P. Cuatrecasas, Prof. J. Drenth, Prof. E. M. Finn, Dr. J. C. Fletcher, Prof. A. N. Glazer, Dr. D. Gratecos, Prof. G. P. Hess, Prof. K. Hofmann, Prof. O. Jardetzky, Prof. E. T. Kaiser, Dr. J. F. Kirsch, Prof. J. Kraut, Dr. M. C. Lin, Prof. W. N. Lipscomb, Prof. G. Lowe, Prof. A. P. Mathias, Dr. S. Moore, Prof. H. Neurath, Prof. A. C. T. North, Prof. D. C. Phillips, Dr. M. A. Raftery, Dr. G. W. Robinson, Dr. J. A. Rupley, Prof. E. L. Smith, Dr. R. M. Stroud, Dr. A. Thomson, Prof. B. L. Vallee, Dr. A. Williams, Dr. H. Witzel, Dr. H. T. Wright, and Prof. H. W. Wyckoff.

S. BLACKBURN

CONTENTS

Chapter 1

THE IMPORTANCE OF ENZYME STUDIES: TECHNIQUES OF INVESTIGATION

I. INTRODUCTION

The importance of enzymes and their study nowadays hardly needs to be emphasized. A measure of this importance is the fact that in 1970 Britain's Science Research Council "selected enzyme chemistry and technology as an area of science worthy of special encouragement, in the belief that its development could lead to a new era in pure and applied chemistry" [1]. The fact that this step was considered necessary in a country which has been responsible for many of the most important advances in enzyme chemistry and structure only serves to underline the importance of the field.

All chemical reactions in living cells are catalyzed by enzymes. Knowledge of the structure of enzymes is therefore essential for the understanding of biologic function. The mechanism of enzyme action, which forms the theme of this book, is one of the central problems of biochemistry. Enzymes are proteins and one can therefore ask the ultimate question associated with enzyme mechanism: How does a protein catalyze a chemical reaction? To answer this question many laboratories have put forward a prodigious effort using a variety of approaches. Although no final answer has been obtained, a great deal of progress has been made. It is the purpose of this book critically and selectively to examine some of this progress. Little more than a decade ago our knowledge of the basis of enzymic catalysis was fragmentary and meager in the extreme. The increase in the number of proteins, including a large proportion of enzymes, whose primary structure has been determined is evident from the increase in the number recorded every year in the *Atlas of Protein Sequence and Structure* [2]. The

increase in the extent of our knowledge of enzyme structure and mechanism is also made evident by comparing the number of pages devoted to a particular enzyme in the successive editions of a comprehensive treatise, *The Enzymes* [3]. This treatise gives full and excellent descriptions of a large number of enzymes. The past few years have also seen the first synthesis in two different laboratories of an enzyme, ribonuclease, which was also the first enzyme to be sequenced (see Chap. 8).

II. EXPERIMENTAL TECHNIQUES

The profound change in the quality of our information about enzymes that has occurred recently would have been impossible without improved investigational procedures. The description of experimental techniques falls outside the scope of the present monograph, but the brief survey of methods which follows should assist the reader to realize the debt recent advances in our knowledge owe to the newer procedures.

A. X-Ray Diffraction

One of the most important techniques which enables us to obtain information about the detailed conformations of enzyme molecules, how they interact with their substrates, and how they are related to each other, is x-ray diffraction. The striking successes of the method have depended on obtaining suitable crystalline forms of the enzyme under study. Following the production of large crystals of pepsin by Philpott in Uppsala in 1934 (see, e.g., Ref. 4), Bernal and Crowfoot [5] obtained x-ray diffraction photographs of them. These photographs, obtained by keeping the crystals in contact with their mother liquor throughout the experiment, showed that atoms in the crystal adopted a regular arrangement, which in principle could be determined. It was not before 1954, however, that Perutz [6] showed how these structures might be solved by attaching heavy atoms to the molecule at specific sites, and 1960 before Kendrew and coworkers [7] showed how the atoms are arranged in myoglobin.

Very few, if any, enzyme crystals give images in which all the atoms appear as separately resolved peaks. Nevertheless, a proportion of the amino acid residues can be confidently identified in a satisfactory x-ray image, even though all may not be. The side chains of

several residues are readily identified; tryptophan is uniquely large, and tyrosine is distinguishable from phenylalanine or histidine by its hydroxyl group. Proline, methionine, cystine, arginine, lysine, glycine, alanine, and serine can generally be recognized. Side chains that are easy to distinguish on the inside of the molecule may be difficult when surrounded by intermolecular solution.

It is often possible to determine enzyme structures in considerable detail from partially resolved images with the help of other information to aid in interpretation. A knowledge of the type of conformation we may expect is often of value in this respect. We have this information largely because of the work of Pauling and Corey, who have analyzed the structure of relatively simple molecules related to proteins, such as glycine [8] and L-threonine [9]. They were then able to predict with remarkable accuracy the polypeptide conformations likely to be encountered in protein molecules. Particularly important findings were that the peptide group is generally planar and that hydrogen bonding plays an important part in determining peptide conformations. Examples of both types of polypeptide conformation which they described, the α helix and β structure [10, 11] have been found in enzymes.

Lipscomb [12] has discussed the extent of the information on structure and sequence that can be obtained from x-ray and chemical studies. In theory the x-ray structure of a protein should yield a sequence, but in practice serious limitations are seen due to intrinsic disorder in the side chains and sometimes even in the polypeptide backbone. Several pairs of side chains are very similar in shape; x-ray techniques alone are thus unlikely to establish a sequence. There are few chemical sequences that have not required later revision, and x-ray studies have often helped in this revision. This was the case, for example, with myoglobin, lysozyme, papain, and carboxypeptidase A.

The x-ray consecutive numbering of amino acid residues is unambiguous provided disorder does not cause uncertainty in the trace of the polypeptide backbone. The x-ray sequence can thus be used to order chemical fragments, groups such as tryptophan being used as markers. Ambiguities in the x-ray determination generally occur at individual positions rather than in clusters. These positions are determined by factors different from those which cause uncertainty in chemical

sequence determination. The combination of both methods should thus lead to the early unraveling of the complete, or nearly complete, sequence. Neurath et al. [18] consider that a polemic of the relative merits of x-ray and chemical methods for the determination of the structure of an enzyme is rather futile, an opinion with which the present author agrees.

Table 1-1 shows the interpretation of protein structures at various resolutions. The accurate assignment of 60% of the amino acid residues was possible with the excellent 2 Å carboxypeptidase map. This probably approaches the best that can be obtained at the resolution. At 2.8 Å resolution, but without knowledge of the amino acid sequence, the main chain of carboxypeptidase could not be followed with certainty. At the same resolution, and with the availability of an amino acid sequence which contained some errors, the main chain of papain could be traced. With information on the sequence, electron-density maps of ribonuclease S and elastase could be interpreted even at 3.5 Å resolution. In these cases, however, a homologous structure was already known. At resolutions lower than 4 Å complete interpretation is probably only possible for structures containing substantial amounts of α helix.

X-ray diffraction investigations are also able to yield information on the specificity of enzymes. Due to the permeability of the crystals it is possible to investigate directly the interactions between enzymes and molecules related to their substrates, such as competitive inhibitors, if they form stable complexes. The experiment is basically simple. The inhibitor molecules are cocrystallized with the enzyme and the diffraction pattern of the complex is measured. An alternative approach is to diffuse the inhibitor into existing enzyme crystals. If the conformation of the enzyme molecules do not alter too greatly on formation of the complex, and if their relative positions and orientations are the same as in the native enzyme, a good approximation of the electron density of the inhibitor can be calculated when the original enzyme structure is known. An electron-density map of this nature can reveal small conformational changes in the enzyme molecule and changes in the arrangement of bound water molecules produced by the binding of the inhibitor.

TABLE 1-1. Interpretation of protein structures at various resolutions[a]

Protein	Resolution of multiple isomorphous replacement (Å)	Amino acid sequence available?	Could main chain be followed?	Amino acid sequence--residues correctly assigned
Myoglobin	2.0	No	99%	48%[b]
Lysozyme	2.0	Yes	All	Some discrepancies resolved
Ribonuclease A	2.0	Yes	All	
Chymotrypsin	2.0 (incomplete data from 2.5)	Yes	98%	
Carboxypeptidase A	2.0	No	All	60%[b]
	2.8	No	Not with certainty	
Subtilisin BPN'	2.5	Yes	All	
Papain	2.8	Yes	All	Several corrections to chemically determined sequence
Ribonuclease S	3.5	Yes	All[c]	
Elastase	3.5	Partial	All[d]	
Streptococcal nuclease	4.0	Yes	Probably 60%	

[a]Reprinted from Ref. 14 by courtesy of Annual Reviews, Inc.

[b]Based on those parts of the structure interpreted by x-ray diffraction before the amino acid sequence was known.

[c]Subsequent to a verbal presentation of the ribonuclease A structure.

[d]Subsequent to publication of the chymotrypsin structure.

B. Chemical Procedures

Amino acid sequence studies have helped to increase our understanding of the activation of zymogens, and a satisfactory explanation of enzyme activity and specificity requires knowledge of complete primary and secondary structure. Determination of the amino acid sequence of an enzyme is therefore one of the first requirements in attempting to arrive at an understanding of its action mechanism.

Although the size of an enzyme molecule is considerable, the catalytic action is limited to one or more restricted areas of the molecule, the active site(s) or active center(s). During the past few years the use of irreversible inhibitors, particularly site-specific reagents, has given direct evidence for the presence of particular amino acid residues in the active sites of enzymes. Many reagents used in protein-modification studies are group reagents with a preference for particular amino acids. In order to limit chemical modification to the active site of the enzyme and obtain information about the functional residues there, an approach is adopted in which reagents for a particular enzyme combine substrate-like features and a chemically reactive group. The substrate part of the reagent structure localizes the reagent at the active site initially in the dissociable enzyme-inhibitor complex. The chemically reactive group then combines covalently, with the irreversible formation of an inhibitor-enzyme compound. The reagents are often specific to a particular enzyme and their reaction with the active-site amino acid is enzymically catalyzed, no reaction occurring with a nonfunctioning enzyme, e.g., in urea solution. With the exception of x-ray crystallographic procedures, this is the only method available for locating active-site amino acids in the amino acid sequence of enzymes. The application of these and similar reagents is described in more detail under the appropriate enzymes.

C. Kinetic Studies

Kinetic studies with synthetic substrates have been very rewarding in the investigation of enzyme function but interpretation of the results is not always a simple matter. The specificity of proteolytic enzymes has generally been expressed in terms of the amino acid residue

whose peptide linkages are cleaved. This criterion alone, however, is not satisfactory since exceptions are found to the rules of specificity that are formulated in this simple way. Konigsberg and coworkers [15, 16] have shown that chymotrypsin and pepsin can hydrolyze a particular peptide linkage in a polypeptide or protein at a much faster rate than other potentially sensitive linkages, and suggest that other factors, in addition to the nature of sensitive amino acid residues, are involved in determining the particular bond that is cleaved by the enzyme. These factors may be a longer sequence of the peptide, or the conformation of the peptide itself.

It is hence becoming increasingly necessary to describe the specificity of an enzyme in broader terms. For this purpose the hydrolysis of model peptides which have up to five or six amino acid residues on either side of the bond that is cleaved, is studied. This type of approach has been used with chymotrypsin, subtilisin, papain, and carboxypeptidase among others. As a corollary the active site of the enzyme is then described in terms of a number of subsites, rather than of a unique site. The concept of subsites has also been successfully applied to nonproteolytic enzymes, such as lysozyme and staphylococcal nuclease.

Fastrez and Fersht [17] (cf. Chap. 2) have discussed the difficulties that may arise in attempting to decide between different theories of enzyme action solely on the basis of kinetic measurements. A single series of measurements of k_{cat} and K_m on a series of substrates cannot, in general, distinguish between the possibilities. The situation is particularly complex in enzymes with an extended active site. Complications may arise in the interpretation of kinetic studies due to the nonproductive binding of small substrates. These questions are discussed in greater detail in the appropriate chapters of the book.

D. Nuclear Magnetic Resonance

One of the problems that for a long time has exercised enzyme chemists is how to assess the effect of a unique environment within a protein on the behavior of an ionizing group. One of the most successful approaches to the question arises from the discovery that the C-2 ring protons of histidine side chains can be resolved in proton

magnetic resonance spectra, so that in favorable cases the titration curves of the individual histidine of a protein can be determined. The technique has been applied, for example, to ribonuclease (see Chap. 8).

III. SCOPE OF THIS STUDY

The concluding paragraphs of this chapter briefly describe the arrangement and scope of the book. Consideration is limited to those enzymes whose investigation has proceeded to the extent that we have some understanding of their action mechanism and structure. The enzymes that have been studied in detail so far are not representative. They are generally extracellular, monomeric, highly stable, and relatively small. They catalyze the hydrolysis of polymeric substrates. A glance at the contents list of this book, however, shows that they include enzymes concerned with the degradation of each of the principal classes of biologic polymers. Together these structures give a great deal of information about the conformation of proteins and their interactions with other molecules, which gives a firm basis for further deductions about the nature and genesis of biologic structures in general.

The first chapters of the book deal with proteolytic enzymes. Hartley [18] in 1960 proposed a classification of these enzymes based on their different mechanisms of action: (a) serine proteinases, e.g., chymotrypsin, trypsin, elastase, thrombin, plasmin, and subtilisin, the active centers of which contain a serine residue that reacts specifically with organophosphorus compounds; (b) thiol proteinases, e.g., papain, ficin, bromelain; (c) acid proteinases, e.g., pepsin and rennin, which are distinguished by their low optimum pH; (d) metal proteinases, e.g., carboxypeptidases, aminopeptidases, and dipeptidases, which are generally exopeptidases. These classes represent different functional solutions to the chemical problem of enzymic peptide bond cleavage, each requiring a unique active site. No further class of proteinases has since been discovered.

The following two chapters describe enzymes which catalyze hydrolysis of phosphate esters. These are ribonuclease and staphyloccal nuclease. The final descriptive chapter is concerned with lysozyme, which catalyzes the hydrolysis of bacterial cell walls. The

concluding chapter attempts to summarize the present state of our knowledge on enzyme mechanism and structure and draw conclusions.

REFERENCES

1. *Support of Enzyme Chemistry and Technology*, 1969-1971, Science Research Council, London, 1971.
2. M. O. Dayhoff ed., *Atlas of Protein Sequence and Structure*, Vol. 5, National Biomedical Research Foundation, Silver Spring, Maryland, 1972.
3. P. D. Boyer, ed., *The Enzymes*, 3rd ed., Academic, New York.
4. D. C. Hodgkin and D. P. Riley, *Structural Chemistry and Molecular Biology*, Freeman, San Francisco, Calif., 1968, p. 15.
5. J. D. Bernal and D. Crowfoot, *Nature*, *133*, 794 (1934).
6. D. W. Green, V. M. Ingram, and M. F. Perutz, *Proc. Roy. Soc., Ser. A*, *225*, 287 (1954).
7. J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips, and V. C. Shore, *Nature*, *185*, 422 (1960).
8. G. Albrecht and R. B. Corey, *J. Amer. Chem. Soc.*, *61*, 1087 (1939).
9. D. P. Shoemaker, J. Donohue, V. Schomaker, and R. B. Corey, *J. Amer. Chem. Soc.*, *72*, 2328 (1950).
10. L. Pauling and R. B. Corey, *Proc. Natl. Acad. Sci. U.S.*, *37*, 729 (1951).
11. L. Pauling, R. B. Corey, and H. R. Branson, *Proc. Natl. Acad. Sci. U.S.*, *37*, 205 (1951).
12. W. N. Lipscomb, *Accts. Chem. Res.*, *3*, 81 (1970).
13. H. Neurath, K. A. Walsh, and R. A. Bradshaw, *Nature*, *225*, 881 (1970).
14. D. M. Blow and T. A. Steitz, *Annu. Rev. Biochem.*, *39*, 63 (1970).
15. W. Konigsberg and R. J. Hill, *J. Biol. Chem.*, *237*, 2547 (1962).
16. J. Goldstein, W. Konigsberg, and R. J. Hill, *J. Biol. Chem.*, *238*, 2016 (1963).
17. J. Fastrez and A. R. Fersht, *Biochemistry*, *12*, 1067 (1973).
18. B. S. Hartley, *Annu. Rev. Biochem.*, *29*, 45 (1960).