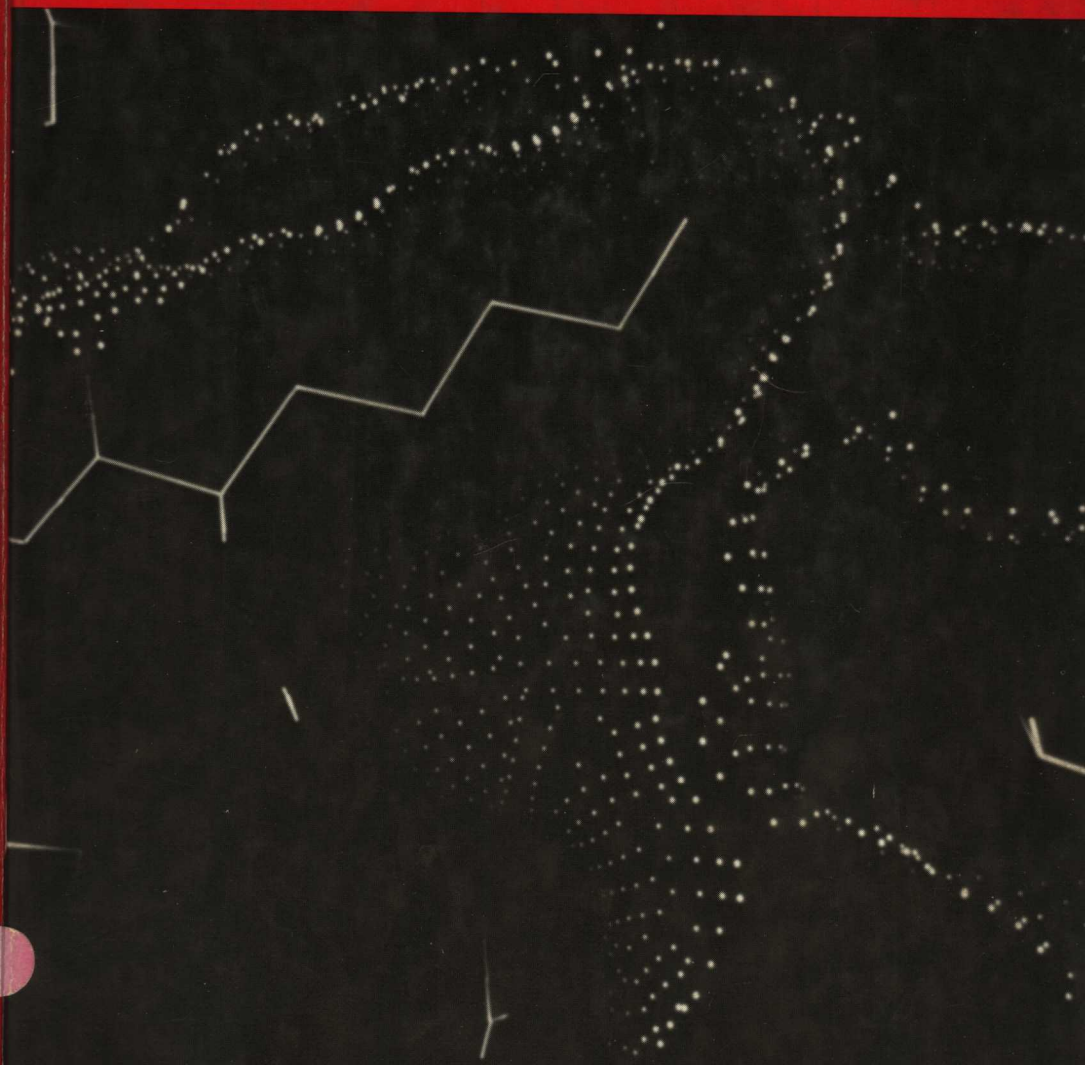
Special Collections Program

Protein and Nucleic Acid Structure and Dynamics

Jonathan King, Editor



Protein and Nucleic Acid Structure and Dynamics

Jonathan King, *Editor*

Department of Biology
Massachusetts Institute of Technology

A volume in the Annual Reviews Special Collections Program

The Benjamin/Cummings Publishing Company, Inc.
Menlo Park, California • Reading, Massachusetts • Wokingham, U.K.
Don Mills, Ontario • Amsterdam • Sydney • Singapore
Tokyo • Mexico City • Bogota • Santiago • San Juan

Front Cover: Computer graphic representation of a bovine trypsin inhibitor-trypsin complex. Surfaces and skeletons shown. (Courtesy of Robert Langridge, Computer Graphics Laboratory, University of California, San Francisco. Copyright © Regents of the University of California.)

Copyright ©1985 by The Benjamin/Cummings Publishing Company, Inc.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. Printed in the United States of America. Published simultaneously in Canada.
ISBN 0-8053-5403-4

ABCDEFGHIJ-MA-8987654

The Benjamin/Cummings Publishing Company, Inc.
2727 Sand Hill Road
Menlo Park, California 94025

Annual Reviews Sources

Rossmann, M.G. and P. Argos, from *Ann. Rev. Biochem.* 1981. 50: 497-532. E.E. Snell, P.D. Boyer, A. Meister, and C.C. Richardson (Eds.).
Copyright ©1981 by Annual Reviews Inc. All rights reserved.

Chothia, C., from *Ann. Rev. Biochem.* 1984. 53: 537-572. C.C. Richardson, P.D. Boyer and A. Meister (Eds.). Copyright ©1984 by Annual Reviews Inc. All rights reserved.

Blundell, T. and S. Wood, from *Ann. Rev. Biochem.* 1982. 51: 123-154. E.E. Snell, P.D. Boyer, A. Meister and C.C. Richardson (Eds.). Copyright ©1982 by Annual Reviews Inc. All rights reserved.

Markley, J.L. and E.L. Ulrich, from *Ann. Rev. Biophys. Bioeng.* 1984. 13: 493-521. D.M. Engelman, C.R. Cantor and T.D. Pollard (Eds.).
Copyright ©1984 by Annual Reviews Inc. All rights reserved.

Kim, P.S. and R.L. Baldwin, from *Ann. Rev. Biochem.* 1982. 51: 459-489. E.E. Snell, P.D. Boyer, A. Meister and C.C. Richardson (Eds.).
Copyright ©1982 by Annual Reviews Inc. All rights reserved.

Karplus, M. and J.A. McCammon, from *Ann. Rev. Biochem.* 1983. 53: 263-300. E.E. Snell, P.D. Boyer, A. Meister, and C.C. Richardson (Eds.).
Copyright ©1983 by Annual Reviews Inc. All rights reserved.

Eisenberg, D., from *Ann. Rev. Biochem.* 1984. 53: 595-623. C.C. Richardson, P.D. Boyer and A. Meister (Eds.). Copyright ©1984 by Annual Reviews Inc. All rights reserved.

Harrington, W.F. and M.E. Rodgers, from *Ann. Rev. Biochem.* 1984. 53: 35-73. C.C. Richardson, P.D. Boyer and A. Meister (Eds.). Copyright ©1984 by Annual Reviews Inc. All rights reserved.

Zimmerman, S.B., from *Ann. Rev. Biochem.* 1982. 51: 395-427. C.C. Richardson, P.D. Boyer, and A. Meister (Eds.).

Record, M.T., S.J. Mazur, P. Melancon, J.H. Roe, S.L. Shaner, and L. Unger from *Ann. Rev. Biochem.* 1981. 50: 997-1024. E.E. Snell, P.D. Boyer, A. Meister, and C.C. Richardson (Eds.). Copyright ©1981 by Annual Reviews Inc. All rights reserved.

Rich, A., A. Nordheim, and A.H.-J. Wang from *Ann. Rev. Biochem.* 1984. 53: 791-846. C.C. Richardson, P.D. Boyer, and A. Meister (Eds.). Copyright ©1984 by Annual Reviews Inc. All rights reserved.

Noller, H.F. from *Ann. Rev. Biochem.* 1984. 53: 119-162. C.C. Richardson, P.D. Boyer, and A. Meister (Eds.). Copyright ©1984 by Annual Reviews Inc. All rights reserved.

Reid, B.R. from *Ann. Rev. Biochem.* 1981. 50: 969-996. E.E. Snell, P.D. Boyer, A. Meister, and C.C. Richardson (Eds.). Copyright ©1981 by Annual Reviews Inc. All rights reserved.

Pabo, C.O. and R.T. Sauer from *Ann. Rev. Biochem.* 1984. 53: 293-321. P.D. Boyer, A. Meister, and C.C. Richardson (Eds.). Copyright ©1984 by Annual Reviews Inc. All rights reserved.

Busch, H., R. Reddy, L. Rothblum and Y.C. Choi from *Ann. Rev. Biochem.* 1982. 51: 617-654. E.E. Snell, P.D. Boyer, A. Meister, and C.C. Richardson (Eds.). Copyright ©1982 by Annual Reviews Inc. All rights reserved.

McGhee, J.D. and G. Felsenfeld from *Ann. Rev. Biochem.* 1980. 49: 1115-1156. E.E. Snell, P.D. Boyer, A. Meister and C.C. Richardson (Eds.). Copyright ©1980 by Annual Reviews Inc. All rights reserved.

此为试读,需要完整PDF请访问: www.ertongbook.com

Publisher's Foreword

The proliferation of scientific information in recent years has been so rapid that carefully written and well referenced reviews are of greater importance to scientists than ever before. Such reviews are critical resources for students just entering a research field and for researchers whose interests are broadening, as well as essential references for many specialists. Recognizing the broad utility of review coverage at this level, Benjamin/Cummings has joined with Annual Reviews Inc. to provide access to its outstanding scientific reviews in new formats: The Benjamin/Cummings—Annual Reviews Special Collections Program. Each volume in this program is dedicated to a single topic of current scientific interest and consists of articles taken from one or more of the Annual Review series. Compiled and introduced by an eminent scientist, the articles in each volume provide review coverage and exhaustive referencing of the original literature in the area discussed. By bringing together the rigorous scholarly standards of Annual Reviews articles and Benjamin/Cummings' worldwide resources and commitment to educational publishing in science, we believe this unique program will be of real utility to those active in science today as well as those who will be active tomorrow.

James W. Behnke
Editor-in-Chief
The Benjamin/Cummings Publishing Company, Inc.
Menlo Park, California
November 1984

Preface

Molecular biology is progressing from the static description of protein and nucleic acid conformation at the atomic level to an understanding of the forces that determine these structures and their dynamics. These form and function relationships underlie catalysis, the regulation of gene expression, the formation of mature structures from their precursors, and, in short, govern much of the normal activity of the cell.

In recent years *Annual Reviews* has published a number of excellent and authoritative reviews covering this emerging area. This volume collects them in one place both for ease of access, and for the ways in which the articles illuminate each other when considered together. Advances in protein structure, folding, and dynamics are covered first, followed by nucleic acid structure and dynamics, and then by a final section examining aspects of protein-nucleic acid interactions.

Protein Structure

The structures of a large number of water soluble proteins have by now been determined by X-ray crystallography. From this body of information some general features of the organization of units of secondary structure into structures of higher order have emerged. Rossmann and Argos catalogue classes of helices, sheets, and turns, and then consider their organization into super-secondary structures. Through comparison of specific domains among families of proteins these authors consider the convergence and divergence of structural features during evolution. The detailed patterns of packing of alpha helices and beta sheets are analyzed by Chothia, with considerable attention to the interacting surfaces between secondary structure units. Despite the apparent variety of interactions, a number of rules governing the packings of these units and the topology of the links between them can be derived.

The conformational flexibility of polypeptide chains is particularly clear from the study of polypeptide hormones. Despite their variation between carrier and receptor environments, a number have been crystallized and these are reviewed by Blundell and Wood. These systems offer the opportunity to identify which features of a peptide interact with the external environment to determine chain conformation. Nuclear magnetic resonance has emerged as a complementary technique to X-ray diffraction, providing details of atomic interactions in polypeptides in the solution state. New insights through NMR concerning the conformation and motions of the atoms in polypeptide chains in solution are covered by Markley and Ulrich.

Though proteins have been traditionally represented as static structures, their various atoms, helices, and domains are in motion. Karplus and McCammon review the motions and modes of subunits at various levels including atoms, side chains, loop displacement, hinge motions, as well as collective motions of the entire molecules. The biological significance of these movements is made particularly clear by the authors' discussion of the entry of oxygen into the myoglobin and hemoglobin heme cavities and of the effects of ligands on this accessibility.

Despite detailed knowledge of both the structures and amino acid sequences of numerous proteins, the mechanisms through which amino acid sequence determines protein structure remain obscure. Solving this problem requires knowledge of the actual folding pathways for polypetide chains. Kim and Baldwin summarize our knowledge of defined refolding pathways for polypeptides and the character of those intermediate stages which have been identified.

Membrane proteins have been much less accessible to X-ray analysis than water soluble enzymes due to the lack of suitable crystals. Eisenberg's review describes the structure of bacterio-rhodopsin and of a group of small peptide toxins which bind to or insert into membranes. Though few crystal structures of membrane proteins have been determined, a considerable amount of sequence data has accumulated and these sequences are analyzed in terms of the location and orientation of hydrophobic residues.

Myosin is representative of a very different kind of molecule from either the membrane proteins or the carrier proteins. Harrington and Rodgers succinctly summarize an extensive body of non-crystallographic data on the structure and conformation of the myosin molecule and relate that to the assembly of myosin molecules into thick filaments, and to the mechanism of the force-generating event.

Nucleic Acid Structure

The last few years have seen a change of views concerning DNA. The realization that DNA structure is varied under physiological conditions and that this variation is a feature of its biological activity have come as recent surprises. Associated with this has been a more detailed understanding of the interactions both along the chain and between chains in B-form DNA, as summarized in the Zimmerman article. Record et al provide a critical analysis of the manner in which solution conditions and interaction with ligands cause the collapse of DNA into a folded or condensed state.

Considerable excitement has come from the identification and characterization of left-handed Z-DNA. Rich, Nordheim and Wang summarize the characteristics of Z-DNA with special emphasis on its physiological role within cells. Recognition that the flexibility and conformation of the DNA double helix are properties of its sequence and its context opens up new views on modes of gene expression and regulation.

The role of nucleic acids in providing structural, rather than coding, information is manifest in the two large RNA molecules that serve as the backbones of all

ribosomes. Noller summarizes the evidence for a very complex secondary and tertiary structure of the ribosomal RNA. These structures probably not only define sites of interaction with ribosomal proteins, but are involved in the motions of different parts of the ribosome during the steps in the protein synthetic cycle.

The simplest of the nucleic acid molecules under consideration is transfer RNA. Studies of its molecular biology reveal that very subtle modifications of bonding and structure alter its biological activity both as a receptor for the RNA synthetases, and in the translation process itself. Reid provides a clear account of the use of NMR spectroscopy to identify the base-pairing interactions within tRNA molecules and the use of these as reporters for tRNA conformation.

Protein-Nucleic Acid Interactions

Although most of our images of protein and nucleic acid structure come from analysis of molecules not undergoing interactions with other macromolecules, such interactions are the crux of biological activity. Using the X-ray structure of isolated phage lambda repressors, Sauer and Pabo examine models of repressor-DNA interactions. These models suggest a wrapping of the protein arms around the DNA as part of the binding interaction, and identify regions of the protein likely to be involved in sequence recognition, confirmed by studies of mutant repressors. The models assume little alteration in DNA conformation, but data from crystals of protein-DNA complexes should soon be able to assess this.

The small RNA-protein complexes involved in RNA processing appear to play a pivotal role in gene expression in higher organisms. Though the structures of such complexes have not been solved at the atomic level, their sequences have been determined. Busch et al summarize these results and discuss the role of these species in RNA processing and other transcriptional and post-transcriptional processes.

The best studied and most general of DNA-protein complexes is the nucleosome, the basic unit of chromosome structure in higher organisms. McGhee and Felsenfeld synthesize the considerable body of work on histone-histone and histone-DNA interactions, and explore how there interactions influence the higher order structure of chromatin and the processes of gene expression and replication.

Jonathan King
Cambridge, Massachusetts
November 1984

CONTENTS

此为试读,需要完整PDF请访问: www.ertongbook.com

I
Protein Structure

Ann. Rev. Biochem. 1981. 50:497–532
Copyright © 1981 by Annual Reviews Inc. All rights reserved

1

PROTEIN FOLDING

Michael G. Rossmann and Patrick Argos

Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907

CONTENTS

PERSPECTIVE

Linderstrøm-Lang and his co-workers (1) were the first to recognize structural levels of organization within a protein. They introduced the terms primary, secondary, and tertiary structure. Although a variety of helical secondary structures had been proposed (cf 2, 3), it was Pauling (4, 5) who recognized the α-helix and β-pleated sheet, which provide an acceptable interpretation of Astbury's α- and β-diffraction patterns for fibrous proteins. Nevertheless, details of the α-helix were not seen at high resolution until the advent of the myoglobin structure (6), while the first atomic resolution observation of a β-sheet as a small antiparallel segment in lysozyme was not published until 1965 (7). Since that time well over 100 distinct structures have been determined. This wealth of information has led to a detailed examination of structural hierarchy as displayed by folded polypeptide chains.

The notion was entertained, even in the 1930s, that a protein would spontaneously refold after in vitro denaturation (8, 9). While three-dimensional structures demonstrate the uniqueness of a general fold with respect to a given protein, they do not directly discern the folding pathways. It was not until the 1960s, when the properties of proteins were better understood,

that the concept of spontaneous renaturation enjoyed wide acceptance. The pivotal work was that of Anfinsen and his co-workers, who "scrambled" ribonuclease, with its eight sulfhydryl groups, by allowing the reduced protein to reoxidize under denaturing conditions of 8 M urea (10). Removal of the denaturant and addition of mercaptoethanol resulted in a stable, functionally active conformation, though "unscrambling" frequently took hours to complete, an obvious discrepancy with in vivo rates. Denaturation-renaturation investigations have since been performed on a variety of other proteins including myoglobin (11), staphylococcal nuclease (12), lysozyme (13), and pancreatic trypsin inhibitor (14–16). The most detailed work has centered on disulfide proteins, where the covalent formation of S–S bonds can be used to characterize intermediates. In this way Creighton (17–19) was able to draw a folding pathway for bovine pancreatic trypsin inhibitor (Figure 1). It is noteworthy that essential intermediates exhibit some incorrect S–S pairing (17).

With the prompting of crystallographic results that revealed the form of folded proteins and renaturation experiments that demonstrated the spontaneity of refolding, a significant understanding emerged of the physical principles underlying the folding operation. Elementary principles of thermodynamics state that folding in a constant physiological environment

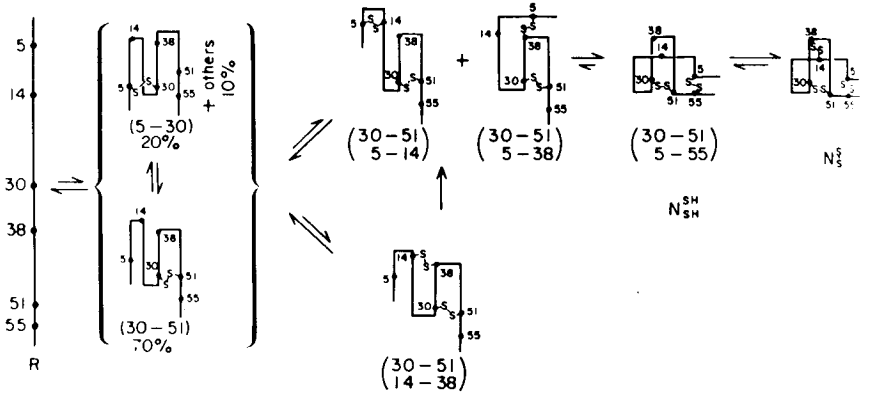


Figure 1 Schematic diagram of the pathway of folding and unfolding of normal bovine pancreatic trypsin inhibitor. The solid line represents the polypeptide backbone, with the positions of the cysteine residues indicated. The configurations of species N^{SH}_{SH} and N^{S}_{S} approximate the conformation of the native inhibitor; those of the others are relatively arbitrary except for the relative positions of the cysteine residues involved in disulfide bonds.

The brackets around the single-disulfide intermediates indicate that they are in rapid equilibrium; only the two most predominant species are depicted. The + between intermediates (30–51,5–14) and (30–51,5–38) signifies that both are formed directly from the single-disulfide intermediates, that both are converted directly to N^{SH}_{SH}, and that either or both are intermediates in the arrangement of (30–51,14–38) to N^{SH}_{SH}. [Reprinted with permission from Creighton (17). Copyright by Academic Press Inc. (London) Ltd.]

must be synonymous with the reduction of Gibbs free energy, though the folded protein may not have attained a "global" minimum (20, 21). Physically this implies burying of hydrophobic groups within the folded molecular core, creation of ion pairs and hydrogen bonds, and reduction of molecular surface. However, such considerations do not provide any information on pathway, since they are concerned with energy. Accordingly, the complex interactions of the polypeptide chain with itself and the environment must be considered. This requires a great simplification of the appearance of the polypeptide (20, 22–24). Alternative simplifications, suggested by light scattering and hydrodynamic measurements, assume nucleation centers, such as helices, around which the polypeptide can condense.

Methods to predict secondary structure from the primary amino acid sequence have been developed to avoid the difficult thermodynamic and statistical calculations. The predictive algorithms are statistical and rely on known protein structures. Perhaps the best known and easiest to apply is the technique of Chou & Fasman (25–27) who rank the amino acids as helix, sheet, and turn formers. They then elaborate on the number and kind of residues required to nucleate and terminate a given structural element. Two international competitions have been held to determine the accuracy of various techniques in the prediction of the adenylate kinase (28) and phage lysozyme (29) secondary structures. It is clear that these methods work better on some proteins than others, but they generally predict with a moderate degree of accuracy (21). Nevertheless, secondary structural predictions have found wide applicability in the analysis of amino acid sequences where the structure is unknown, and are particularly valuable when other functional properties of the structures are known (30). Recently, attempts have been made to extend these methods to predict tertiary structure by analyzing such variables as the packing of α-helices (31, 32) or the frequency of topological arrangements within β-sheets (cf 33).

The relationship between sequences and fold is not rigorous. The code that relates sequence to structure is highly degenerate, and yet remains responsive to the protein solvent. Furthermore there appear to be only a limited number of amino acid sequences that can provide a unique structure in a given environment; all others are nonsense. This may in part be the basis for the apparently small number of folds or architectural classes that have so far been observed. Albeit, with only a few exceptions (34), these observations have been confined to aqueous-soluble proteins.

Enzymes often utilize to their advantage a change of environment to alter conformation, as is implied by the term "induced fit" (35, 36). This is typified by the movement of the loop in lactate dehydrogenase (LDH), which is controlled by NAD binding (37, 38). Huber (39) has drawn attention to the order-disorder phenomenon that can occur in the formation of the trypsin specificity pocket or in the Fc fragment of the immunoglobulins.

Similarly, certain sections of viral coat proteins may fold as α-helices only in the presence of RNA (40, 41). The disordered segments of a polypeptide chain frequently start and end with glycines and contain few, if any, aromatic residues.

The degeneracy of the relationship between sequence and structure is essential in the process of evolution, as it permits an alteration of specific amino acids without destruction of the fold and, hence, without loss of function [see Lesk & Chothia (42) who analyze this degeneracy for the globin structure]. Indeed, the great conservation of residues in the active center of enzymes and the associated conservation of fold clearly demonstrate that function is a controlling aspect in protein evolution [see Doolittle (43) for a recent discussion of protein evolution]. Quaternary interactions can have a regulating effect on the function of each subunit. The presence of functional globin monomers such as lamprey hemoglobin or sperm whale myoglobin is one of many examples that show that the evolution of monomers usually precedes the subsequent evolution of allosteric oligomeric proteins as hemoglobin.

Excellent reviews on protein structure and fold include those of Jane Richardson (44), Schulz & Schirmer (21) and Cantor & Schimmel (45). The valuable book by Dickerson & Geis (46), an updated version of which is soon to be published by Benjamin, has been a standard text for most students and scholars in this area. In addition, there are a long line of reviews on protein structure in the *Annual Review of Biochemistry* (e.g. 47–52) and in other journals (e.g. 53, 54). Mention should also be made of reviews relevant to the dynamics of folding (e.g. 55–64). The present discourse confines itself to the analysis of folds of known protein structures.

SUMMARY

After some general remarks on protein structure, there follows a discussion on primary, secondary, and tertiary organization. The account of primary structure includes a discussion of the conformation of disulfide bonds. Types of helices, sheets, and turns are described in the section on secondary structure, followed by a discussion of super-secondary structure and the effects of metals and prosthetic groups on protein fold.

The crux of the review lies in an examination of tertiary structure, or specifically of domains that are defined, in part, as functional units within a polypeptide chain. An assembly of domains can in turn result in a protein whose function is quite sophisticated. Some consideration of domain recognition is given in the section on taxonomy and in the appendix. The key part of the tertiary structure section concentrates on a taxonomic protein classification dependent not only on structure but also on function. A discussion

此为试读,需要完整PDF请访问: www.ertongbook.com