

Contents in full

The Authors	v
A Note from the Publisher on Primers in Biology	vii
Preface	ix
A Note on the Protein Data Bank	x
Acknowledgements	xii

CHAPTER 1 From Sequence to Structure

1-0 Overview: Protein Function and Architecture	2
Proteins are the most versatile macromolecules of the cell	
There are four levels of protein structure	
1-1 Amino Acids	4
The chemical characters of the amino-acid side chains have important consequences for the way they participate in the folding and functions of proteins	
1-2 Genes and Proteins	6
There is a linear relationship between the DNA base sequence of a gene and the amino-acid sequence of the protein it encodes	
The organization of the genetic code reflects the chemical grouping of the amino acids	
1-3 The Peptide Bond	8
Proteins are linear polymers of amino acids connected by amide bonds	
The properties of the peptide bond have important effects on the stability and flexibility of polypeptide chains in water	
1-4 Bonds that Stabilize Folded Proteins	10
Folded proteins are stabilized mainly by weak noncovalent interactions	
The hydrogen-bonding properties of water have important effects on protein stability	
1-5 Importance and Determinants of Secondary Structure	12
Folded proteins have segments of regular conformation	
The arrangement of secondary structure elements provides a convenient way of classifying types of folds	
Steric constraints dictate the possible types of secondary structure	
The simplest secondary structure element is the beta turn	
1-6 Properties of the Alpha Helix	14
Alpha helices are versatile cylindrical structures stabilized by a network of backbone hydrogen bonds	
Alpha helices can be amphipathic, with one polar and one nonpolar face	
Collagen and polyproline helices have special properties	
1-7 Properties of the Beta Sheet	16
Beta sheets are extended structures that sometimes form barrels	
Amphipathic beta sheets are found on the surfaces of proteins	
1-8 Prediction of Secondary Structure	18
Certain amino acids are more usually found in alpha helices, others in beta sheets	
1-9 Folding	20
The folded structure of a protein is directly determined by its primary structure	
Competition between self-interactions and interactions with water drives protein folding	
Computational prediction of folding is not yet reliable	
Helical membrane proteins fold by condensation of preformed secondary structure elements in the bilayer	
1-10 Tertiary Structure	22
The condensing of multiple secondary structural elements leads to tertiary structure	
Bound water molecules on the surfaces of a folded protein are an important part of the structure	

	Tertiary structure is stabilized by efficient packing of atoms in the protein interior	
1-11	Membrane Protein Structure The principles governing the structures of integral membrane proteins are the same as those for water-soluble proteins and lead to formation of the same secondary structure elements	24
1-12	Protein Stability: Weak Interactions and Flexibility The folded protein is a thermodynamic compromise Protein structure can be disrupted by a variety of agents The marginal stability of protein tertiary structure allows proteins to be flexible	26
1-13	Protein Stability: Post-Translational Modifications Covalent bonds can add stability to tertiary structure Post-translational modification can alter both the tertiary structure and the stability of a protein	28
1-14	The Protein Domain Globular proteins are composed of structural domains Domains have hydrophobic cores Multidomain proteins probably evolved by the fusion of genes that once coded for separate proteins	30
1-15	The Universe of Protein Structures The number of domain folds is large but limited Protein structures are modular and proteins can be grouped into families based on the basis of the domains they contain The modular nature of protein structure allows for sequence insertions and deletions	32
1-16	Protein Motifs Protein motifs may be defined by their primary sequence or by the arrangement of secondary structure elements Identifying motifs from sequence is not straightforward	34
1-17	Alpha Domains and Beta Domains Protein domains can be classified according to their secondary structural elements The two common motifs for alpha domains are the four-helix bundle and the globin fold Beta domains contain strands connected in two distinct ways Antiparallel beta sheets can form barrels and sandwiches	36
1-18	Alpha/Beta, Alpha+Beta and Cross-Linked Domains In alpha/beta domains each strand of parallel beta sheet is usually connected to the next by an alpha helix There are two major families of alpha/beta domains: barrels and twists Alpha+beta domains have independent helical motifs packed against a beta sheet Metal ions and disulfide bridges form cross-links in irregular domains	38
1-19	Quaternary Structure: General Principles Many proteins are composed of more than one polypeptide chain All specific intermolecular interactions depend on complementarity	40
1-20	Quaternary Structure: Intermolecular Interfaces All types of protein-stabilizing interactions contribute to the formation of intermolecular interfaces Inappropriate quaternary interactions can have dramatic functional consequences	42
1-21	Quaternary Structure: Geometry Protein assemblies built of identical subunits are usually symmetric	44
1-22	Protein Flexibility Proteins are flexible molecules Conformational fluctuations in domain structure tend to be local Protein motions involve groups of non-bonded as well as covalently bonded atoms Triggered conformational changes can cause large movements of side chains, loops, or domains	46

CHAPTER 2 From Structure to Function

2-0 Overview: The Structural Basis of Protein Function	50
There are many levels of protein function	
There are four fundamental biochemical functions of proteins	
2-1 Recognition, Complementarity and Active Sites	52
Protein functions such as molecular recognition and catalysis depend on complementarity	
Molecular recognition depends on specialized microenvironments that result from protein tertiary structure	
Specialized microenvironments at binding sites contribute to catalysis	
2-2 Flexibility and Protein Function	54
The flexibility of tertiary structure allows proteins to adapt to their ligands	
Protein flexibility is essential for biochemical function	
The degree of protein flexibility varies in proteins with different functions	
2-3 Location of Binding Sites	56
Binding sites for macromolecules on a protein's surface can be concave, convex, or flat	
Binding sites for small ligands are clefts, pockets cavities	
Catalytic sites often occur at domain and subunit interfaces	
2-4 Nature of Binding Sites	58
Binding sites generally have a higher than average amount of exposed hydrophobic surface	
Binding sites for small molecules are usually concave and partly hydrophobic	
Weak interactions can lead to an easy exchange of partners	
Displacement of water also drives binding events	
Contributions to binding affinity can sometimes be distinguished from contributions to binding specificity	
2-5 Functional Properties of Structural Proteins	60
Proteins as frameworks, connectors and scaffolds	
Some structural proteins only form stable assemblies	
Some catalytic proteins can also have a structural role	
Some structural proteins serve as scaffolds	
2-6 Catalysis: Overview	62
Catalysts accelerate the rate of a chemical reaction without changing its overall equilibrium	
Catalysis usually requires more than one factor	
Catalysis is reducing the activation-energy barrier to a reaction	
2-7 Active-Site Geometry	64
Reactive groups in enzyme active sites are optimally positioned to interact with the substrate	
2-8 Proximity and Ground-State Destabilization	66
Some active sites chiefly promote proximity	
Some active sites destabilize ground states	
2-9 Stabilization of Transition States and Exclusion of Water	68
Some active sites primarily stabilize transition states	
Many active sites must protect their substrates from water, but must be accessible at the same time	
2-10 Redox Reactions	70
A relatively small number of chemical reactions account for most biological transformations	
Oxidation/reduction reactions involve the transfer of electrons and often require specific cofactors	
2-11 Addition/Elimination, Hydrolysis and Decarboxylation	72
Addition reactions add atoms or chemical groups to double bonds, while elimination reactions remove them to form double bonds	
Esters, amides and acetals are cleaved by reaction with water; their formation requires removal of water	
Loss of carbon dioxide is a common strategy for removing a single carbon atom from a molecule	

2-12	Active-Site Chemistry	74
	Active sites promote acid-base catalysis	
2-13	Cofactors	76
	Many active sites use cofactors to assist catalysis	
2-14	Multi-Step Reactions	78
	Some active sites employ multi-step mechanisms	
2-15	Multifunctional Enzymes	80
	Some enzymes can catalyze more than one reaction	
	Some bifunctional enzymes can have only one active site	
	Some bifunctional enzymes contain two active sites	
2-16	Multifunctional Enzymes with Tunnels	82
	Some bifunctional enzymes shuttle unstable intermediates through a tunnel connecting the active sites	
	Trifunctional enzymes can shuttle intermediates over huge distances	
	Some enzymes also have non-enzymatic functions	

CHAPTER 3 Control of Protein Function

3-0	Overview: Mechanisms of Regulation	86
	Protein function in living cells is precisely regulated	
	Proteins can be targeted to specific compartments and complexes	
	Protein activity can be regulated by binding of an effector and by covalent modification	
	Protein activity may be regulated by protein quantity and lifetime	
	A single protein may be subject to many regulatory influences	
3-1	Protein Interaction Domains	88
	The flow of information within the cell is regulated and integrated by the combinatorial use of small protein domains that recognize specific ligands	
3-2	Regulation by Location	90
	Protein function in the cell is context-dependent	
	There are several ways of targeting proteins in cells	
3-3	Control by pH and Redox Environment	92
	Protein function is modulated by the environment in which the protein operates	
	Changes in redox environment can greatly affect protein structure and function	
	Changes in pH can drastically alter protein structure and function	
3-4	Effector Ligands: Competitive Binding and Cooperativity	94
	Protein function can be controlled by effector ligands that bind competitively to ligand-binding or active sites	
	Cooperative binding by effector ligands amplifies their effects	
3-5	Effector Ligands: Conformational Change and Allostery	96
	Effector molecules can cause conformational changes at distant sites	
	ATCase is an allosteric enzyme with regulatory and active sites on different subunits	
	Disruption of function does not necessarily mean that the active site or ligand-binding site has been disrupted	
	Binding of gene regulatory proteins to DNA is often controlled by ligand-induced conformational changes	
3-6	Protein Switches Based on Nucleotide Hydrolysis	98
	Conformational changes driven by nucleotide binding and hydrolysis are the basis for switching and motor properties of proteins	
	All nucleotide switch proteins have some common structural and functional features	
3-7	GTPase Switches: Small Signaling G proteins	100
	The switching cycle of nucleotide hydrolysis and exchange in G proteins is modulated by the binding of other proteins	
3-8	GTPase Switches: Signal Relay by Heterotrimeric GTPases	102
	Heterotrimeric G proteins relay and amplify extracellular signals from a receptor to an intracellular signaling pathway	
3-9	GTPase Switches: Protein Synthesis	104
	EF-Tu is activated by binding to the ribosome, which thereby signals it to release its bound tRNA	

3-10	Motor Protein Switches	106
	Myosin and kinesin are ATP-dependent nucleotide switches that move along actin filaments and microtubules respectively	
3-11	Regulation by Degradation	108
	Protein function can be controlled by protein lifetime	
	Proteins are targeted to proteasomes for degradation	
3-12	Control of Protein Function by Phosphorylation	110
	Protein function can be controlled by covalent modification	
	Phosphorylation is the most important covalent switch mechanism for the control of protein function	
3-13	Regulation of Signaling Protein Kinases: Activation Mechanism	112
	Protein kinases are themselves controlled by phosphorylation	
	Src kinases both activate and inhibit themselves	
3-14	Regulation of Signaling Protein Kinases: Cdk Activation	114
	Cyclin acts as an effector ligand for cyclin-dependent kinases	
3-15	Two-Component Signaling Systems in Bacteria	116
	Two-component signal carriers employ a small conformational change that is driven by covalent attachment of a phosphate group	
3-16	Control by Proteolysis: Activation of Precursors	118
	Limited proteolysis can activate enzymes	
	Polypeptide hormones are produced by limited proteolysis	
3-17	Protein Splicing: Autoproteolysis by Inteins	120
	Some proteins contain self-excising inteins	
	The mechanism of autocatalysis is similar for inteins from unicellular organisms and metazoan Hedgehog protein	
3-18	Glycosylation	122
	Glycosylation can change the properties of a protein and provide recognition sites	
3-19	Protein Targeting by Lipid Modifications	124
	Covalent attachment of lipids targets proteins to membranes and other proteins	
	The GTPases that direct intracellular membrane traffic are reversibly associated with internal membranes of the cell	
3-20	Methylation, N-acetylation, Sumoylation and Nitrosylation	126
	Fundamental biological processes are regulated by other post-translational modifications of proteins	

CHAPTER 4 From Sequence to Function: Case Studies in Structural and Functional Genomics

4-0	Overview: From Sequence to Function in the Age of Genomics	130
	Genomics is making an increasing contribution to the study of protein structure and function	
4-1	Sequence Alignment and Comparison	132
	Sequence comparison provides a measure of the relationship between genes	
	Alignment is the first step in determining whether two sequences are similar to each other	
	Multiple alignments and phylogenetic trees	
4-2	Protein Profiling	134
	Structural data can help sequence comparison find related proteins	
	Sequence and structural motifs and patterns can identify proteins with similar biochemical functions	
	Protein-family profiles can be generated from multiple alignments of protein families for which representative structures are known	
4-3	Deriving Function from Sequence	136
	Sequence information is increasing exponentially	
	In some cases function can be inferred from sequence	
4-4	Experimental Tools for Probing Protein Function	138
	Gene function can sometimes be established experimentally without information from protein structure or sequence homology	
4-5	Divergent and Convergent Evolution	140

	Evolution has produced a relatively limited number of protein folds and catalytic mechanisms	
	Proteins that differ in sequence and structure may have converged to similar active sites, catalytic mechanisms and biochemical function	
	Proteins with low sequence similarity but very similar overall structure and active sites are likely to be homologous	
	Convergent and divergent evolution are sometimes difficult to distinguish	
	Divergent evolution can produce proteins with sequence and structural similarity but different functions	
4-6	Structure from Sequence: Homology Modeling	142
	Structure can be derived from sequence by reference to known protein folds and protein structures	
	Homology modeling is used to deduce the structure of a sequence with reference to the structure of a close homolog	
4-7	Structure From Sequence: Profile-Based Threading and "Rosetta"	144
	Profile-based threading tries to predict the structure of a sequence even if no sequence homologs are known	
	The Rosetta method attempts to predict protein structure from sequence without the aid of a homologous sequence or structure	
4-8	Deducing Function from Structure: Protein Superfamilies	146
	Members of a structural superfamily often have related biochemical functions	
	The four superfamilies of serine proteases are examples of convergent evolution	
	Very closely related protein families can have completely different biochemical and biological functions	
4-9	Strategies for Identifying Binding Sites	148
	Binding sites can sometimes be located in three-dimensional structures by purely computational means	
	Experimental means of locating binding sites are at present more accurate than computational methods	
4-10	Strategies for Identifying Catalytic Residues	150
	Site-directed mutagenesis can identify residues involved in binding or catalysis	
	Active-site residues in a structure can be recognized computationally by their geometry	
	Docking programs model the binding of ligands	
4-11	TIM Barrels: One Structure with Diverse Functions	152
	Knowledge of a protein's structure does not necessarily make it possible to predict its biochemical or cellular functions	
4-12	PLP Enzymes: Diverse Structures with One Function	154
	A protein's biochemical function and catalytic mechanism do not necessarily predict its three-dimensional structure	
4-13	Moonlighting: Proteins With More Than One Function	156
	In multicellular organisms, multifunctional proteins help expand the number of protein functions that can be derived from relatively small genomes	
4-14	Chameleon Sequences: One Sequence with More than One Fold	158
	Some amino-acid sequences can assume different secondary structures in different structural contexts	
4-15	Prions, Amyloids and Serpins: Metastable Protein Folds	160
	A single sequence can adopt more than one stable structure	
4-16	Functions for Uncharacterized Genes: Galactonate Dehydratase	162
	Determining biochemical function from sequence and structure becomes more accurate as more family members are identified	
	Alignments based on conservation of residues that carry out the same active-site chemistry can identify more family members than sequence comparisons alone	
	In well studied model organisms, information from genetics and cell biology can help identify the substrate of an "unknown" enzyme and the actual reaction catalyzed	
4-17	Starting From Scratch: A Gene Product of Unknown Function	164
	Function cannot always be determined from sequence, even with the aid of structural information and chemical intuition	

CHAPTER 5 Structure Determination

5-1 The Interpretation of Structural Information	168
Experimentally determined protein structures are the result of the interpretation of different types of data	
Both the accuracy and the precision of a structure can vary	
The information content of a structure is determined by its resolution	
5-2 Structure Determination by X-Ray Crystallography and NMR	170
Protein crystallography involves summing the scattered X-ray waves from a macromolecular crystal	
NMR spectroscopy involves determining internuclear distances by measuring perturbations between assigned resonances from atoms in the protein in solution	
5-3 Quality and Representation of Crystal and NMR Structures	172
The quality of a finished structure depends largely on the amount of data collected	
Different conventions for representing proteins are useful for different purposes	
Glossary	175
References	181
Index	189