

# Contents

Sections marked with an asterisk may be skipped on a first reading.

## 1 The photosynthetic reaction centre: protein structure in a microcosm 1

The reaction centre from *Rhodospseudomonas viridis* 3

Conclusions 12

Recommended reading and references 12

Exercises and problems 13

## 2 *In vivo, in vitro, in silicio* 15

Why study proteins? 15

Genomics 15

Protein structure and conformation 19

The amino acids 21

The amino acids vary in size, hydrogen-bonding potential, and charge 21

Protein folding 22

Protein folding pathways 23

Protein structures depend on a variety of chemical forces for their stability and for their affinity and specificity for ligands 23

Conformation of the polypeptide chain 31

The Sasisekharan–Ramakrishnan–Ramachandran diagram 32

Sidechain conformation 35

The known protein structures 36

Protein structure historically considered 36

Methods of structure determination 37

X-ray crystallography 37

Model building and refinement 38

How accurate are the structures? 39

Measures of structural quality 40

Nuclear Magnetic Resonance 40

Low-temperature electron microscopy (cryo-EM) 42

The Protein Data Bank 42

The World Wide Web 43

Data banks for molecular biology 44

Archival data banks 45

Specialized data banks 45

Information-retrieval tools 45

Summary 47

Glossary 49

Useful web sites 50

Recommended reading and references 51

Exercises, problems and weblems 53

## 3 Pattern and form in protein structure 59

Helices and sheets 59

Hydrogen-bonding patterns of helices and sheets 59

Hydrogen-bonding patterns in sheets 68

The $\beta$ bulge	70
$\beta$ -barrels	70
The hierarchical nature of protein architecture	70
Assignment of helices and sheets	78
An album of small structures	80
Classification of protein structures	84
Comparisons of protein sequences and structures	85
Classification of protein topologies	89
Protein structure prediction	90
Structural interpretation of genome information	94
Loops	95
Sequence-structure relationships in short $\beta$ hairpins	96
Sequence-structure relationships in two-residue $\beta$ hairpins	97
Structural determinants of medium-sized loops	98
Protein-ligand interactions	103
Water molecules	110
Useful web sites	111
Recommended reading and references	111
Exercises, problems and weblems	112

#### 4 The varieties of protein structure 127

Catalogues of protein structure	127
SCOP	127
CATH	129
FSSP, and the DALI domain dictionary	131
The known structures	132
$\alpha$ -helical proteins	132
Principles of the architecture of $\alpha$ -helical proteins	135*
Complete enumeration within restricted classes of folds	135*
Geometry: the polyhedral model of $\alpha$ -helical globular proteins	140*
Structural chemistry: tertiary-structural interactions	141*
The structure of helix-helix packings	141*
$\beta$ -sheet proteins	143
$\beta$ -sheet 'sandwiches'	145
Other $\beta$ -sheet proteins	147
$\alpha + \beta$ proteins	147
$\alpha / \beta$ proteins	151
The $\beta$ - $\alpha$ - $\beta$ unit	151
Linear or open $\beta$ - $\alpha$ - $\beta$ proteins	151
Closed $\beta$ - $\alpha$ - $\beta$ barrel structures	152
The TIM barrel	152
Other $\beta$ -barrel structures	158
Irregular structures	158
Conclusions	160
Useful web sites	160
Recommended reading and references	160
Exercises, problems and weblems	161

#### 5 Molecular evolution 165

Evolution of DNA and proteins	165
-------------------------------	-----

- Direct access to the genome—nucleotide sequences 166
- Evolutionary changes in protein sequences 166
- Variability in selective constraints in protein molecules 171

## Evolution of protein structures 172

- The evolution of proteins with altered function 173
- Neutral mutations 174
- Domain combination and recombination 175

## Structural relationships among related molecules 176

- A general relationship between divergence of amino acid sequence and protein conformation in families of related proteins 176
- Point mutations 182
- Closely related and distantly related proteins 183
- Application to homology modelling 185

Useful web sites 186

Recommended reading and references 186

Exercises, problems and weblems 186

## Evolution in selected protein families 195

Evolution of the globins 195

Phycocyanin and the globins 200\*

Evolution of serine proteinases of the chymotrypsin family 204

The structure of serine proteinases of the chymotrypsin family 207

Structures of individual domains 207

The domain-domain interface 210

The specificity pocket 211

The  $\beta$ -barrels in serine proteinase domains and the packing of residues in their interiors 213\*

Conclusions 217

NAD-binding domains of dehydrogenases 217

The sequence motif G\*G\*\*G 218

Comparison of NAD-binding domains of dehydrogenases 219

Cofactor binding 224

Binding of NAD vs NADP 224

Conclusion 226

Useful web sites 226

Recommended reading and references 226

Exercises, problems and weblems 227

## Some proteins of the immune system 229

Antibody structure 229

The structures of immunoglobulins 231

The antigen-binding site 236

Somatic mutation and the maturation of the antibody response 237

Canonical structures of antigen-binding loops of antibodies 240

Greater variability in the H3 loop 240

Conclusions: how does the immune system generate and then refine molecules of such a wide range of specificity? 246

Sequence 246

Structure 247

Proteins of the Major Histocompatibility Complex 247

Structures of MHC proteins 249

Specificities of the MHC system	254
Class I and class II MHC proteins function in parallel, selecting different immune responses to extracellular and intracellular pathogens	255
Peptide binding	255
Conclusion	260
T-cell receptors	261
Canonical structures for TCR binding loops	262
The TCR-MHC-peptide complex	263
Useful web sites	266
Recommended reading and references	266
Exercises, problems and weblems	267

## 8 Conformational changes in proteins 271

Structural changes arising from change in state of ligation	271
Sperm whale myoglobin	271
Hinge motions in proteins	274
Hinge motion in lactoferrin	274
Hinge motion in myosin	274
The 'helix interface shear' mechanism of conformational change	277
Insulin	277
How are these movements accommodated: why does the close-packing of interfaces between packed helices not hinder them severely?	281
Citrate synthase	281
The allosteric change in haemoglobin	284
Structural differences between deoxy- and oxyhaemoglobin	286
Tertiary structural change between deoxy- and oxyhaemoglobin	286
Intersubunit interactions in haemoglobin	287
Serpins: SERine Proteinase INhibitors	290
Structural states of serpins	291
Different folding topologies in one protein family	292
Mechanism of the S→R transition	295
Higher-level structural changes	297
The GroEL-GroES chaperonin complex	297
The GroEL-GroES conformational change	299
Operational cycle	299
ATPase	302
Useful web sites	305
Recommended reading and references	305
Exercises and problems	306

## Appendix 1. An album for browsing 309

## Appendix 2. Useful web sites 335

Index of structures illustrated	341
Subject index	345