CHAPTER 3

Protein Structure and Function

Proteins are macromolecules that play central roles in all the processes of life. Chapter 3 begins with a discussion of key properties of proteins and continues with a description of the chemical properties of amino acids—the building blocks of proteins. It is essential that you learn the names, symbols, and properties of the 20 common amino acids at this point, as they will recur throughout the text in connection with protein structures, enzymatic mechanisms, metabolism, protein synthesis, and the regulation of gene expression. It is also important to review the behavior of weak acids and bases, either in the appendix to Chapter 3 or in an introductory chemistry text. Following the discussion of amino acids, the chapter turns to peptides and to the linear sequences of amino acid residues in proteins. Next, it describes the folding of these linear polymers into the specific three-dimensional structures of proteins. The primary structure (or sequence of amino acids) dictates the higher orders of structure including secondary (α, β, etc.), tertiary (often globular), and quaternary (with multiple chains). You should note that the majority of functional proteins exist in water and that their structures are stabilized by the forces and interactions you learned about in Chapter 1. This chapter concludes with a discussion of the theory of how proteins fold, including attempts to predict protein folding from amino acid sequences.

LEARNING OBJECTIVES

When you have mastered this chapter, you should be able to complete the following objectives.

Introduction

1. List the key properties of proteins.
2. Explain how proteins relate one-dimensional gene structure to three-dimensional structure in the cell, and their complex interactions with each other and various substrates.

Proteins Are Built from a Repertoire of 20 Amino Acids (Text Section 3.1)

3. Draw the structure of an *amino acid* and indicate the following features, which are common to all amino acids: *functional groups, side chains, ionic forms,* and *isomeric forms.*
4. Classify each of the 20 amino acids according to the side chain on the α *carbon* as *aliphatic, aromatic, sulfur-containing, aliphatic hydroxyl, basic, acidic,* or *amide derivative.*
5. Give the name and one-letter and three-letter *symbol* of each amino acid. Describe each amino acid in terms of *size, charge, hydrogen-bonding capacity, chemical reactivity,* and *hydrophilic* or *hydrophobic* nature.
6. Define *pH* and *pKa.* Use these concepts to predict the *ionization state* of any given amino acid or its side chain in a protein.
7. State *Beer's Law*. Understand how it can be used to estimate protein concentration.

Primary Structure: Amino Acids Are Linked by Peptide Bonds to Form Polypeptide Chains (Text Section 3.2)

8. Draw a *peptide bond* and describe its *conformation* and its role in *polypeptide* sequences. Indicate the *N-* and *C-terminal residues* in *peptides.*
9. Define *main chain, side chains,* and *disulfide bonds* in polypeptides. Give the range of *molecular weights* of proteins.
10. Explain the origin and significance of the unique *amino acid sequences* of proteins.
11. Understand why nearly all peptide bonds are *trans*.
12. Define the ϕ and ψ angles used to describe a peptide bond, and be able to read a *Ramachandran plot.*

Secondary Structure: Polypeptide Chains Can Fold into Regular Structures Such as the Alpha Helix, the Beta Sheet, and Turns and Loops (Text Section 3.3)

13. Differentiate between two major *periodic structures* of proteins: the α *helix* and the β *pleated sheet.* Describe the patterns of hydrogen bonding, the shapes, and the dimensions of these structures.
14. List the types of interactions among amino acid side chains that stabilize the *three-dimensional structures* of proteins. Give examples of *hydrogen bond donors* and *acceptors.*
15. Describe α*-helical coiled coils* in specialized proteins and the role of β *turns* or *hairpin turns* in the structure of common proteins.

Tertiary Structure: Water-Soluble Proteins Fold into Compact Structures with Nonpolar Cores (Text Section 3.4)

16. Using *myoglobin* and *porin* as examples, describe the main characteristics of native folded protein structures.

Quaternary Structure: Polypeptide Chains Can Assemble into Multisubunit Structures (Text Section 3.5)

17. Describe the *primary, secondary, tertiary,* and *quaternary structures* of proteins. Describe *domains.*

The Amino Acid Sequence of a Protein Determines Its Three-Dimensional Structure (Text Section 3.6)

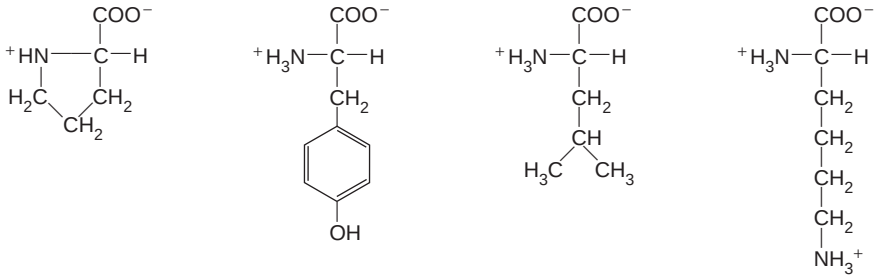
18. Using *ribonuclease* as an example, describe the evidence for the hypothesis that all of the information needed to specify the three-dimensional structure of a protein is contained in its amino acid sequence.
19. Rationalize the conformational preferences of different amino acids in proteins and polypeptides.
20. Give evidence that protein folding appears to be a cooperative transition, and explain why that means it is an "all or none" process.
21. Explain how *protein folding* proceeds through stabilization of *intermediate states* rather than through a sampling of all possible conformations.
22. Discuss the methods and advances in the prediction of three-dimensional structures of proteins.
23. List examples of the *modification* and *cleavage* of proteins that expand their functional roles.

SELF-TEST

Introduction

Proteins Are Built from a Repertoire of 20 Amino Acids

1. (a) Examine the four amino acids given below:



A B C D

Indicate which of these amino acids are associated with the following properties:

(a) aliphatic side chain
(b) basic side chain
(c) three ionizable groups
(d) charge of +1 at pH 7.0
(e) pK ~10 in proteins
(f) secondary amino group
(g) designated by the symbol K
(h) in the same class as phenylalanine
(i) most hydrophobic of the four
(j) side chain capable of forming hydrogen bonds

(b) Name the four amino acids.
(c) Name the other amino acids of the same class as D.

2. Draw the structure of cysteine at pH 1.
3. Match the amino acids in the left column with the appropriate side chain types in the right column.

(a) Lys	(1) nonpolar aliphatic
(b) Glu	(2) nonpolar aromatic
(c) Leu	(3) basic
(d) Cys	(4) acidic
(e) Trp	(5) sulfur-containing
(f) Ser	(6) hydroxyl-containing

4. Which of the following amino acids have side chains that are negatively charged under physiologic conditions (i.e., near pH 7)?

(a) Asp
(b) His
(c) Trp
(d) Glu
(e) Cys

5. Why does histidine act as a buffer at pH 6.0? What can you say about the buffering capacity of histidine at pH 7.6?

Primary Structure: Amino Acids Are Linked by Peptide Bonds to Form Polypeptide Chains

6. How many different dipeptides can be made from the 20 L amino acids? What are the minimum and the maximum number of pK values for any dipeptide?
7. For the pentapeptide Glu-Met-Arg-Thr-Gly,
   (a) name the carboxyl-terminal residue.
   (b) give the number of charged groups at pH 7.
   (c) give the net charge at pH 1.
   (d) write the sequence using one-letter symbols.
   (e) draw the peptide bond between the Thr and Gly residues, including both side chains.
8. If a polypeptide has 400 amino acid residues, what is its approximate mass?

(a) 11,000 daltons
(b) 22,000 daltons
(c) 44,000 daltons
(d) 88,000 daltons

9. Which amino acid can stabilize protein structures by forming covalent cross-links between polypeptide chains?
   (a) Met
   (b) Ser
   (c) Gln
   (d) Gly
   (e) Cys

10. Discuss the significance of *Ramachandran plots*. Contrast the conformational states of Gly and Pro in proteins compared with other amino acid residues.

Secondary Structure: Polypeptide Chains Can Fold into Regular Structures Such as the Alpha Helix, the Beta Sheet, and Turns and Loops

11. Which of the following statements about the peptide bond are true?
   (a) The peptide bond is planar because of the partial double-bond character of the bond between the carbonyl carbon and the nitrogen.
   (b) There is relative freedom of rotation of the bond between the carbonyl carbon and the nitrogen.
   (c) The hydrogen that is bonded to the nitrogen atom is trans to the oxygen of the carbonyl group.
   (d) There is no freedom of rotation around the bond between the α carbon and the carbonyl carbon.

12. Which of the following statements about the α helix structure of proteins is correct?
   (a) It is maintained by hydrogen bonding between amino acid side chains.
   (b) It makes up about the same percentage of all proteins.
   (c) It can serve a mechanical role by forming stiff bundles of fibers in some proteins.
   (d) It is stabilized by hydrogen bonds between amide hydrogens and amide oxygens in polypeptide chains.
   (e) It includes all 20 amino acids at equal frequencies.

13. Which of the following properties are common to α-helical and β pleated sheet structures in proteins?
   (a) rod shape
   (b) hydrogen bonds between main-chain CO and NH groups
   (c) axial distance between adjacent amino acids of 3.5 Å
   (d) variable numbers of participating amino acid residues

14. Explain why α helix and β pleated sheet structures are often found in the interior of water-soluble proteins.

Tertiary Structure: Water-Soluble Proteins Fold into Compact Structures with Nonpolar Cores

15. Which of the following amino acid residues are likely to be found on the inside of a water-soluble protein?
   (a) Val
   (b) His
   (c) Ile
   (d) Arg
   (e) Asp

16. Which of the following statements about the structures of water-soluble proteins, exemplified by myoglobin, are **not** true?
    (a) They contain tightly packed amino acids in their interior.
    (b) Most of their nonpolar residues face the aqueous solvent.
    (c) The main-chain NH and CO groups are often involved in H-bonded secondary structures in the interior of these proteins.
    (d) Polar residues such as His may be found in the interior of these proteins if the residues have specific functional roles.
    (e) All of these proteins contain β sheet structural motifs.

Quaternary Structure: Polypeptide Chains Can Assemble into Multisubunit Structures

17. Match the levels of protein structures in the left column with the appropriate descriptions in the right column.

(a) primary	(1) association of protein subunits
(b) secondary	(2) overall folding of a single chain, can include α-helical and β sheet structures
(c) tertiary	(3) linear amino acid sequence
(d) quaternary	(4) repetitive arrangement of amino acids that are near each other in the linear sequence

The Amino Acid Sequence of a Protein Determines Its Three-Dimensional Structure

18. Which of the following statements are true?
    (a) Ribonuclease (RNase) can be treated with urea and reducing agents to produce a random coil.
    (b) If one oxidizes random-coil RNase in urea, it quickly regains its enzymatic activity.
    (c) If one removes the urea and oxidizes RNase slowly, it will renature and regain its enzymatic activity.
    (d) Although renatured RNase has enzymatic activity, it can be readily distinguished from native RNase.

19. When most proteins are exposed to acidic pH (e.g., pH 2), they lose biological activity. Explain why.

20. Which one of the following amino acids may alter the direction of polypeptide chains and interrupt α helices?
    (a) Phe
    (b) Cys
    (c) Trp
    (d) His
    (e) Pro

21. If we know that a solution of protein is half-folded, what will we find in solution?
    (a) 100% half-folded protein
    (b) 50% fully folded, 50% unfolded
    (c) 33% fully folded, 34% half-folded, and 33% unfolded

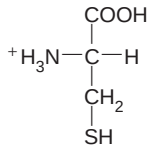
22. Several amino acids can be modified after the synthesis of a polypeptide chain to enhance the functional capabilities of the protein. Match the type of modifying group in the left column with the appropriate amino acid residues in the right column.

(a) phosphate	(1) Glu
(b) hydroxyl	(2) Thr
(c) γ-carboxyl	(3) Pro
(d) acetyl	(4) Ser
	(5) N-terminal
	(6) Tyr

23. How can a protein be modified to make it more hydrophobic?

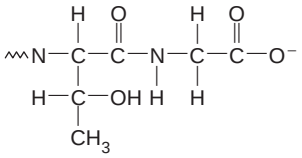
ANSWERS TO SELF-TEST

1. (a) (a) C (b) D (c) B, D (d) D (e) B, D (f) A (g) D (h) B (i) C (j) B, D (k) D.
   (b) A is proline, B is tyrosine, C is leucine, and D is lysine.
   (c) Histidine and arginine (basic amino acids).
2. See the structure of cysteine. At pH 1, all the ionizable groups are protonated.



Cysteine

3. (a) 3 (b) 4 (c) 1 (d) 5 (e) 2 (f) 6
4. a, d
5. Histidine acts as a buffer at pH 6.0 because this is the pK of the imidazole group. At pH 7.6, histidine is a poor buffer because no one ionizing group is partially protonated and therefore capable of donating or accepting protons without markedly changing the pH.
6. The 20 L amino acids can form $20 \times 20 = 400$ dipeptides. The minimum number of pK values for any dipeptide is two; the maximum is four.
7. (a) glycine
   (b) 4, namely the 2 carboxyl groups of glutamate, the R group of arginine, and the alpha amino group of glycine.
   (c) +2, contributed by the N-terminal amino group and the arginine residue
   (d) E-M-R-T-G
   (e) See the structure of the peptide bond below.



Peptide bond

8. c
9. e
10. a. A Ramachandran plot gives the possible ϕ and ψ angles for the main polypeptide chain containing different amino acid residues. The fact that glycine lacks an R group means that it is much less constrained than other residues. In Figure 3.31, the left-handed helix region, which occurs rarely, generally includes several Gly residues. In contrast to glycine, proline is more highly constrained than most residues because the R group is tied to the amino group. This fixes ϕ at about -65°. In Figure 3.26, the rare cis form of the peptide bond is shown as occurring about half of the time in X-Pro peptide bonds.
11. a, c
12. c, d
13. b, d
14. In both α-helical and β sheet structures, the polar peptide bonds of the main chain are involved in internal hydrogen bonding, thereby eliminating potential hydrogen bond formation with water. Overall the secondary structures are less polar than the corresponding linear amino acid sequences.
15. a, c. Specific charged and polar amino acid residues may be found inside some proteins, in active sites, but most polar and charged residues are located on the surface of proteins.
16. b, e. Statement (b) is incorrect because globular, water-soluble proteins have most of their nonpolar residues buried in the interior of the protein. Statement (e) is incorrect because not all water-soluble proteins contain β sheet secondary structures. For example, myoglobin is mostly α-helical and lacks β sheet structures.
17. (a) 3 (b) 4 (c) 2 (d) 1
18. a, c
19. A low pH (pH 2) will cause the protonation of all ionizable side chains and will change the charge distribution on the protein; furthermore, it will impart a large net positive charge to the protein. The resulting repulsion of adjacent positive charges and the disruption of salt bridges often cause unfolding of the protein and loss of biological activity.
20. e
21. b
22. (a) 2, 4, 6 (b) 3 (c) 1 (d) 5
23. The attachment of a fatty acid chain to a protein can increase its hydrophobicity and promote binding to lipid membranes.

PROBLEMS

1. The net charge of a polypeptide at a particular pH can be determined by considering the pK value for each ionizable group in the protein. For a linear polypeptide composed of 10 amino acids, how many α-carboxyl and α-amino groups must be considered?
2. For the formation of a polypeptide composed of 20 amino acids, how many water molecules must be removed when the peptide bonds are formed? Although the hydrolysis of a peptide bond is energetically favored, the bond is very stable in solution. Why?

3. Where stereoisomers of biomolecules are possible, only one is usually found in most organisms; for example, only the L amino acids occur in proteins. What problems would occur if, for example, the amino acids in the body proteins of herbivores were in the L isomer form, whereas the amino acids in a large number of the plants they fed upon were in the D isomer form?
4. Many types of proteins can be isolated only in quantities that are too small for the direct determination of a primary amino acid sequence. Recent advances in gene cloning and amplification allow for relatively easy analysis of the gene coding for a particular protein. Why would an analysis of the gene provide information about the protein's primary sequences? Suppose that two research groups, one in New York and the other in Los Angeles, are both analyzing the same protein from the same type of human cell. Why would you not be surprised if they publish exactly the same primary amino acid sequence for the protein?
5. Each amino acid in a run of several amino acid residues in a polypeptide chain has ϕ values of approximately -140° and ψ values of approximately $+147$. What kind of structure is it likely to be?
6. A survey of the location of reverse turns in soluble proteins shows that most reverse turns are located at the surface of the protein, rather than within the hydrophobic core of the folded protein. Can you suggest a reason for this observation?
7. Wool and hair are elastic; both are α-keratins, which contain long polypeptide chains composed of α helices twisted about each other to form cablelike assemblies with crosslinks involving Cys residues. Silk, on the other hand, is rigid and resists stretching; it is composed primarily of antiparallel β pleated sheets, which are often stacked and interlocked. Briefly explain these observations in terms of the characteristics of the secondary structures of these proteins.
8. In a particular enzyme, an alanine residue is located in a cleft where the substrate binds. A mutation that changes this residue to a glycine has no effect on activity; however, another mutation, which changes the alanine to a glutamate residue, leads to a complete loss of activity. Provide a brief explanation for these observations.
9. Glycophorin A is a glycoprotein that extends across the red blood cell membrane. The portion of the polypeptide that extends across the membrane bilayer contains 19 amino acid residues and is folded into an α helix. What is the width of the bilayer that could be spanned by this helix? The interior of the bilayer includes long acyl chains that are nonpolar. Which of the 20 L amino acids would you expect to find among those in the portion of the polypeptide that traverses the bilayer?
10. Before Anfinsen carried out his work on refolding in ribonuclease, some scientists argued that directions for folding are given to the protein during its biosynthesis. How did Anfinsen's experiments contradict that argument?
11. Early experiments on the problem of protein folding suggested that the native three-dimensional structure of a protein was an automatic consequence of its primary structure.
    (a) Cite experimental evidence that shows that this is the case.

    Later, the discovery that proteins are synthesized directionally on ribosomes, from the amino to the carboxy terminus, complicated the earlier view of protein folding.

    (b) Explain what the complicating circumstance might be.

    The discovery of chaperone proteins allows both earlier views to be reconciled.

    (c) Explain how that might be the case.

12. Suppose you are studying the conformation of a monomeric protein that has an unusually high proportion of aromatic amino acid residues throughout the length of the polypeptide chain. Compared with a monomeric protein containing many aliphatic residues, what might you observe for the relative α-helical content for each of the two types of proteins? Would you expect to find aromatic residues on the outside or the inside of a globular protein? What about aliphatic residues?
13. As more and more protein sequences and three-dimensional structures become known, there is a proliferation of computer algorithms for the prediction of folding based on sequence. How might it be possible to winnow through the possibilities and find the best computer programs? Bear in mind that if the sequence and the structure are available, it is too easy to "reverse engineer" a routine that will produce the correct answer.
14. In its discussion of protein modification and cleavage, the text refers to the synthesis and cleavage of a large polyprotein precursor of virus proteins, as well as to the synthesis of multiple polypeptide hormones from a single polypeptide chain. Is there an advantage to synthesizing a large precursor chain and then cleaving it to create a number of products?
15. What is the molarity of pure water? Show that a change in the concentration of water by ionization does not appreciably affect the molarity of the solution.
16. When sufficient H^+ is added to lower the pH by one unit, what is the corresponding increase in hydrogen ion concentration?
17. You have a solution of HCl that has a pH of 2.1. What is the concentration of HCl needed to make this solution?
18. The charged form of the imidazole ring of histidine is believed to participate in a reaction catalyzed by an enzyme. At pH 7.0, what is the probability that the imidazole ring will be charged?
19. Calculate the pH at which a solution of cysteine would have no net charge.

ANSWERS TO PROBLEMS

1. Only the N-terminal α-amino group and the C-terminal α-carboxyl group will undergo ionization. The internal groups will be joined by peptide bonds and are not ionizable.
2. For a peptide of n residues, $n - 1$ water molecules must be removed. A significant activation energy barrier makes peptide bonds kinetically stable.
3. All metabolic reactions in an organism are catalyzed by enzymes that are generally specific for either the D or the L isomeric form of a substrate. If an animal (an herbivore in this case) is to be able to digest the protein from a plant and build its own protein from the resulting amino acids, both the animal and the plant must make their proteins from amino acids having the same configuration.
4. Because the sequence of DNA specifies, through a complementary sequence of RNA, the amino acid sequence of a protein, knowledge about any one of the three types of sequences yields information about the other two. One would also expect the coding sequence for a particular protein to be the same among members of the same species, allowing for an occasional rare mutation. For that reason, the published primary amino acid sequences are likely to be the same.
5. From the Ramachandran plot in Figure 3.35 of the text, we see that β conformation is accommodated by ϕ values of approximately -140° and ψ values of approximately $+147^\circ$. The structure is most likely a β sheet. In fact, the "low" numbers here imply that it is an antiparallel beta sheet. The parallel β sheet would have higher numbers, more like $\phi = -160^\circ$ and $\psi = +160^\circ$.

6. Figure 3.42 in the text shows that in a reverse turn the CO group of residue 1 is hydrogen-bonded to the NH group of residue 4. However, there are no adjacent amino acid residues available to form intrachain hydrogen bonds with the CO and NH groups of residues 2 and 3. These groups cannot form hydrogen bonds in the hydrophobic environment found in the interior portion of a folded protein. They are more likely to hydrogen bond with water on the surface of the protein.
7. When the α helices in wool are stretched, intrahelix hydrogen bonds are broken as are some of the interhelix disulfide bridges; maximum stretching yields an extended β sheet structure. The Cys cross-links provide some resistance to stretch and help pull the α helices back to their original positions. In silk, the β sheets are already maximally stretched to form hydrogen bonds. Each β pleated sheet resists stretching, but since the contacts between the sheets primarily involve van der Waals forces, the sheets are somewhat flexible.
8. Both alanine and glycine are neutral nonpolar residues with small side chains, whereas the side chain of glutamate is acidic and bulkier than that of alanine. Either feature of the glutamate R group could lead to the loss of activity by altering the protein conformation or by interfering with the binding of the substrate.
9. Since each residue in the α helix is 1.5 Å from its neighbor, the length of the chain that spans the membrane bilayer is 19×1.5 Å = 28.5 Å, which is also the width of the membrane. One would expect to find nonpolar amino acid residues in the polypeptide portion associated with the membrane bilayer. These would include Ala, Val, Leu, Ile, Met, and Phe (FILMV + A). The actual sequence of the buried chain is

   I–T–L–I–I–F–G–V–M–A–G–V–I–G–T–I–L–L–I.

10. The fact that ribonuclease folded in vitro to yield full activity indicated that the biosynthetic machinery is not required to direct the folding process for this protein.
11. (a) The experiment by Anfinsen on ribonuclease, described in Section 3.6 of the text, is the classic observation. When native ribonuclease is treated with mercaptoethanol to disrupt disulfide bonds and with urea as a denaturant, it unfolds, as indicated by the fact that it becomes enzymically inactive. When urea is removed by dialysis and disulfide bonds reform by oxidation, it regains enzymic activity, suggesting that its native structure has been restored.
    (b) The discovery that proteins are synthesized directionally on ribosomes beginning at the amino terminus complicates matters somewhat because folding of the amino end of the polypeptide chain could begin before the carboxyl end had been synthesized. Such folding could represent the most stable conformation over a short range, but there would be no guarantee that it would be part of the energy minimum for the entire molecule.
    (c) Chaperone proteins could bind to an initially synthesized polypeptide and prevent it from undergoing final folding until the entire molecule was synthesized.
12. The higher the proportion of aromatic side chains (such as those of phenylalanine) in the protein, the more likely that steric hindrance among closely located residues could interfere with the establishment of the regular repeating structure of the α helix. Smaller aliphatic side chains like those of leucine, isoleucine, and valine would be less likely to interfere. Structural studies on many proteins reveal that the number of aromatic residues in α-helical segments is relatively low, while the content of aliphatic side chains in such segments is unremarkable, compared to that of other nonhelical regions of a folded protein. Both aliphatic and aromatic side chains (especially that of phenylalanine) are hydrophobic, so that many of them are buried inside a globular protein, away from water molecules.

13. Protein scientists have devised a competition called CASP, or Critical Assessment of Techniques for Protein Structure Prediction, which is held every other year. Laboratories that are working on determination of three-dimensional structure by x-ray crystallography (or nmr) announce that they expect to release the structure in a few months. They give a description of the sequence of the protein and its use in the cell, and withhold the actual structural coordinates until a certain date. In the meantime, laboratories with predictive algorithms publicly post the structure they think the protein will have. The success or failure of the prediction takes place in a public arena, and the better predictors have bragging rights. CASP-4 in 2000 showed that there are several effective programs available, notably ROSETTA, used by David Baker of the University of Washington. Results of the competition are published in the journal *Protein* and online (in technical language) at the website http://predictioncenter.llnl.gov/.

14. The primary advantage of precursor chain synthesis is that the production of related proteins can be coordinated. This could be important in viral infection, and it may also be important for coordinated synthesis of hormones with related activities. It is worth noting that there are other reasons for the synthesis of polyprotein precursors. For example, the genome of the poliovirus consists of a single RNA molecule that acts as a messenger on entering the cytoplasm of the host. In eukaryotic cells a messenger RNA molecule can be translated into only one polypeptide chain. Therefore the poliovirus can reproduce only by synthesizing its proteins by sequential cleavages.

15. The molarity of water equals the number of moles of water per liter. A liter of water weighs 1000 grams, and its molecular weight is 18, so the molarity of water is

$$\mathrm{M} = \frac{1000}{18} = 55.6$$

At 25°C, K_w is 1.0×10^{-14}; at neutrality, the concentration of both hydrogen and hydroxyl ions is each equal to 10^{-7} M. Thus, the actual concentration of H_2O is (55.6 − 0.0000001) M; the difference is so small that it can be disregarded.

16. Because pH values are based on a logarithmic scale, every unit change in pH means a tenfold change in hydrogen ion concentration. When pH = 2.0, $[H^+] = 10^{-2}$ M; when pH = 3.0, $[H^+] = 10^{-3}$ M.

17. Assume that HCl in solution is completely ionized to H^+ and Cl^-. Then find the concentration of H^+, which equals the concentration of Cl^-.

$$\mathrm{pH} = \log\left[\mathrm{H}^+\right] = 2.1$$

$$\left[\mathrm{H}^+\right] = 10^{2.1}$$

$$= 10^{0.9} \quad 10^{3}$$

$$= 7.94 \quad 10^{3}\,\mathrm{M}$$

$$\text{Thus, } \left[\mathrm{H}^+\right] = \left[\mathrm{Cl}^-\right] = \left[\mathrm{HCl}\right] = 7.94 \quad 10^{3}\,\mathrm{M}$$

18. Use the Henderson-Hasselbalch equation to calculate the concentration of histidine, whose imidazole ring is ionized at neutral pH. The value of pK for the ring is 6.0 for a histidine residue in a protein (see Table 3.1).

$$\text{pH} = \text{p}K + \log \frac{[\text{His}]}{[\text{His}^+]}$$

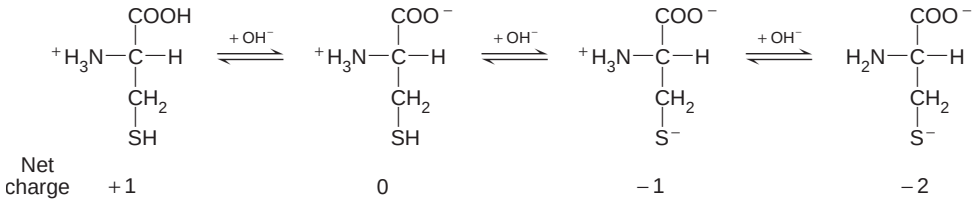
$$7.0 = 6.0 + \log \frac{[\text{His}]}{[\text{His}^+]}$$

$$\log \frac{[\text{His}]}{[\text{His}^+]} = 1.0$$

$$\frac{[\text{His}]}{[\text{His}^+]} = 10$$

At pH 7.0, the ratio of uncharged histidine to charged histidine is 10:1, making the probability that the side chain is charged only 9%.

19. To see which form of cysteine has no net charge, examine all the possible forms, beginning with the one that is most protonated:



The pH of the cysteine solution at which the amino acid has no net charge will be that point at which there are equal amounts of the compound with a single positive charge and a single negative charge. This is, in effect, the average of the two corresponding pK values (see the Appendix to Chapter 3), one for the α-carboxyl group and the other for the side chain sulfhydryl group. Thus, (1.8 + 8.3)/2 = 5.05. This value is also known as the *isoelectric point.*

EXPANDED SOLUTIONS TO TEXT PROBLEMS

1. (a) Since tropomyosin is double-stranded, each strand will have a mass of 35 kd. If the average residue has a mass of 110 d, there are 318 residues per strand (35,000/110), and the length is 477Å (1.5 Å/residue × 318).
   (b) Since 2 of the 40 residues formed the hairpin turn, 38 residues formed the antiparallel β pleated sheet which is 19 residues long (38/2). In β pleated sheets, the axial distance between adjacent amino acids is 3.5 Å. Hence, the length of this segment is 66.5 Å (3.5 × 19).
2. Branching at the β carbon of the side chain (isoleucine), in contrast to branching at the γ carbon (leucine), sterically hinders the formation of a helix. This fact can be shown with molecular models.

3. Changing alanine to valine results in a bulkier side chain, which prevents the correct interior packing of the protein. Changing a nearby, bulky, isoleucine side chain to glycine apparently alleviates the space problem and allows the correct conformation to take place.
4. The amino acid sequence of insulin does not determine its three-dimensional structure. By catalyzing a disulfide-sulfhydryl exchange, this enzyme speeds up the activation of scrambled ribonuclease because the native form is the most thermodynamically stable. In contrast, the structure of active insulin is not the most thermodynamically stable form. The three-dimensional structure of insulin is determined by the folding of preproinsulin, which is later processed to mature insulin.
5. Appropriate hydrogen-bonding sites on the protease might induce formation of an intermolecular β pleated sheet with a portion of the target protein. This process would effectively fully extend α helices and other folded portions of the target molecule.
6. Being the smallest amino acid, glycine can fit into spaces too small to accommodate other amino acids. Thus, if sharp turns or limited spaces for amino acids occur in a functionally active conformation of a protein, glycine is required; no substitute will suffice. In view of this, it is not surprising that glycine is highly conserved.
7. To answer this question one needs to know some of the characteristics of the guanidinium group of the side chain of arginine and of the other functional groups in proteins. Most of the needed information is presented in Figure 3-42 in your textbook; note that the guanidinium group has a positive charge at pH 7 and contains several hydrogen bond donor groups. The positive charge can form salt bridges with the negatively charged groups of proteins (glutamate and asparatate side chains and the terminal carboxylate). As a hydrogen bond donor, the guanidinium group can react with the various hydrogen bond acceptors shown in Figure 3-42 (glutamine, asparagine, aspartate, and the main chain carbonyl). It can also hydrogen bond with the hydroxyl group of serine and threonine (not shown in Figure 3-42). Hydroxyl groups accept hydrogen bonds much like water does.
8. The keratin of hair is essentially a bundle of long protein strands joined together by disulfide bonds. If these bonds are broken (reduced) by the addition of a thiol and the hair curled, the keratin chains slip past each other into a new configuration. When an oxidizing agent is added, new disulfide bonds are formed, thus stabilizing the new "curled" state.
9. There is a considerable energy cost for burying charged groups of non-hydrogen-bonded polar groups inside a hydrophobic membrane. Therefore, an α-helix with hydrophobic side chains is particularly suited to span a membrane. The backbone hydrogen-bonding requirements are all satisfied by intramolecular interactions within the α-helix. Good candidate amino acids with hydrophobic side chains would include Ala, Ile, Leu, Met, Phe, and Val. (Pro is also hydrophobic but will cause a bend in the helix.) Additionally, the aromatic (and amphipathic) amino acids Trp and Tyr are often found toward the ends of membrane-spanning helices, near the phospholipid head groups in the membrane/water interface region.
10. The protein is not at equilibrium, but is in a state where the peptide bond is "kinetically stable" against hydrolysis. This situation is due to the large activation energy for hydrolyzing a peptide bond.

11. One can effectively apply the Henderson-Hasselbalch equation successively to the amino group and to the carboxyl group and multiply the results to arrive at a ratio of 10^{-5}.

    So, considering first the amino group, pH = pK + log $[NH_2]/[NH_3^+]$. With pH = 7 and pK = 8, one has 7 = 8 + log $[NH_2]/[NH_3^+]$ or $[NH_2]/[NH_3^+] = 10^{-1}$. Considering now the carboxyl group with pK of 3, one has 7 = 3 + log $[COO^-]/[COOH]$, or $[COO^-]/[COOH] = 10^{+4}$ or $[COOH]/[COO^-] = 10^{-4}$. Then to consider the two simultaneous ionizations that relate the zwitterionic form to the neutral form of an amino acid such as alanine, one needs to multiply the ratio of $[NH_2]/[NH_3^+]$ by the ratio of $[COOH]/[COO^-]$, i.e., $(10^{-1}) \times (10^{-4}) = (10^{-5})$.

12. The presence of the larger sulfur atom (next to the beta carbon of Cys) alters the relative priorities of the groups attached to the α carbon. The stereochemical arrangement of the β carbon with respect to the α hydrogen does not change, but the convention for assigning the R configuration changes when the Cβ-sulfur is present. (With methionine, the sulfur is too far removed for Cβ to influence the group priority.)

13. LEARNING SCIENCE IS GREAT!

14. No. Unlike the Pro nitrogen in X-Pro, the nitrogen of X in the peptide bond of Pro-X is *not* bonded between two tetrahedral carbon atoms. Therefore, the steric preference for the *trans* conformation will be similar to that of other (non-proline) peptide bonds.

15. Model A shows the reference structure for extended polypeptide chain with $\phi = 180^\circ$ and $\psi = 180^\circ$, so the answer is c. Models C and E have one torsion angle identical to model A and the other angle changed to 0°. In model C ϕ is changed to 0° (answer d), and in model E ψ is changed to 0° (answer b). Comparing model B with a reference for which $\phi = 0^\circ$ (model C), we see a 60° *counter*clockwise rotation of ϕ, when viewed from Cα, so answer e is correct for model B. Finally, comparing model D with the $\phi = 0^\circ$ reference in model C, we see a 120° clockwise rotation of ϕ, when viewed from Cα (answer a).

16. One should use Beer's Law and remember that each mole of protein contains 3 moles of tryptophan. Then for the protein, $A = 3\varepsilon cl$, where ε is the molar extinction coefficient for tryptophan at 280 nm. With A = 0.1, $\varepsilon = 3400\ M^{-1}\ cm^{-1}$ and l = 1.0 cm, one has $c = A / (3\varepsilon l)$. Therefore $c = (0.1)/((3)(3400\ M^{-1})) = 9.8 \times 10^{-6}$ M. For the concentration in grams per liter, one multiplies 9.8×10^{-6} moles/liter by 100,000 grams per mole to arrive at 0.98 g/liter, or 0.98 mg/mL.