

UNIT I: Protein Structure and Function

Amino Acids

1

I. OVERVIEW

Proteins are the most abundant and functionally diverse molecules in living systems. Virtually every life process depends on this class of molecules. For example, enzymes and polypeptide hormones direct and regulate metabolism in the body, whereas contractile proteins in muscle permit movement. In bone, the protein collagen forms a framework for the deposition of calcium phosphate crystals, acting like the steel cables in reinforced concrete. In the bloodstream, proteins, such as hemoglobin and plasma albumin, shuttle molecules essential to life, whereas immunoglobulins fight infectious bacteria and viruses. In short, proteins display an incredible diversity of functions, yet all share the common structural feature of being linear polymers of amino acids. This chapter describes the properties of amino acids. Chapter 2 explores how these simple building blocks are joined to form proteins that have unique three-dimensional structures, making them capable of performing specific biologic functions.

II. STRUCTURE OF THE AMINO ACIDS

Although more than 300 different amino acids have been described in nature, only 20 are commonly found as constituents of mammalian proteins. [Note: These are the only amino acids that are coded for by DNA, the genetic material in the cell (see p. 395).] Each amino acid (except for proline, which has a secondary amino group) has a carboxyl group, a primary amino group, and a distinctive side chain ("R-group") bonded to the α -carbon atom (Figure 1.1A). At physiologic pH (approximately pH 7.4), the carboxyl group is dissociated, forming the negatively charged carboxylate ion ($-\text{COO}^-$), and the amino group is protonated ($-\text{NH}_3^+$). In proteins, almost all of these carboxyl and amino groups are combined through peptide linkage and, in general, are not available for chemical reaction except for hydrogen bond formation (Figure 1.1B). Thus, it is the nature of the side chains that ultimately dictates the role

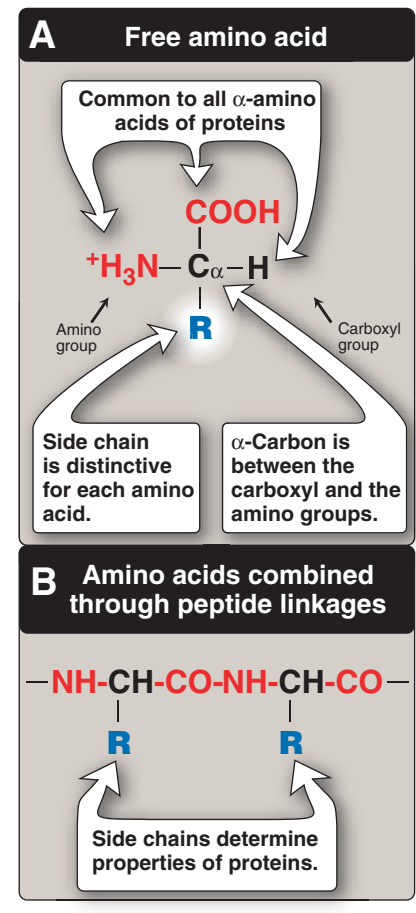


Figure 1.1
Structural features of amino acids
(shown in their fully protonated form).

an amino acid plays in a protein. It is, therefore, useful to classify the amino acids according to the properties of their side chains, that is, whether they are nonpolar (have an even distribution of electrons) or polar (have an uneven distribution of electrons, such as acids and bases; Figures 1.2 and 1.3).

A. Amino acids with nonpolar side chains

Each of these amino acids has a nonpolar side chain that does not gain or lose protons or participate in hydrogen or ionic bonds (Figure 1.2). The side chains of these amino acids can be thought of as “oily” or lipid-like, a property that promotes hydrophobic interactions (see Figure 2.10, p. 19).

1. Location of nonpolar amino acids in proteins: In proteins found in aqueous solutions—a polar environment—the side chains of the nonpolar amino acids tend to cluster together in the interior of the protein (Figure 1.4). This phenomenon, known as the hydrophobic

NONPOLAR SIDE CHAINS

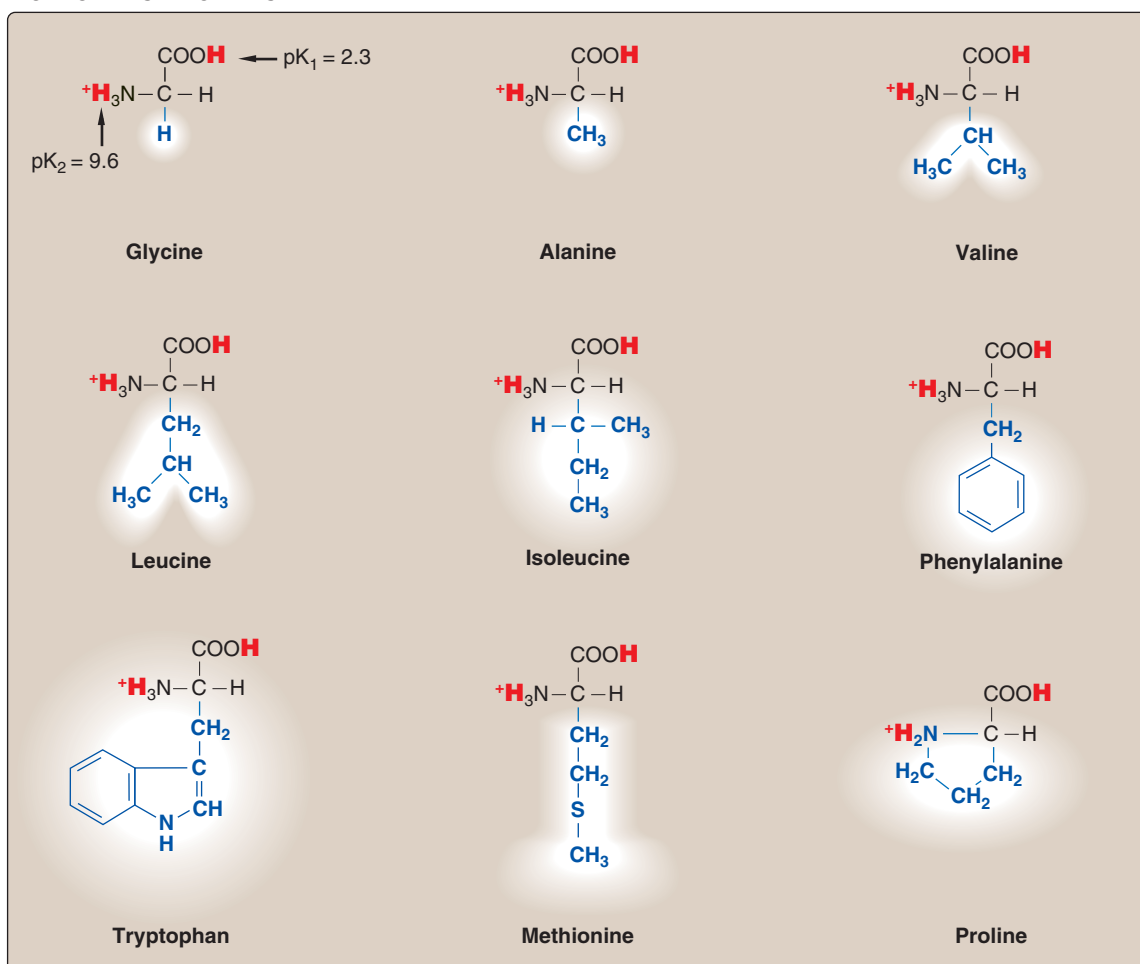
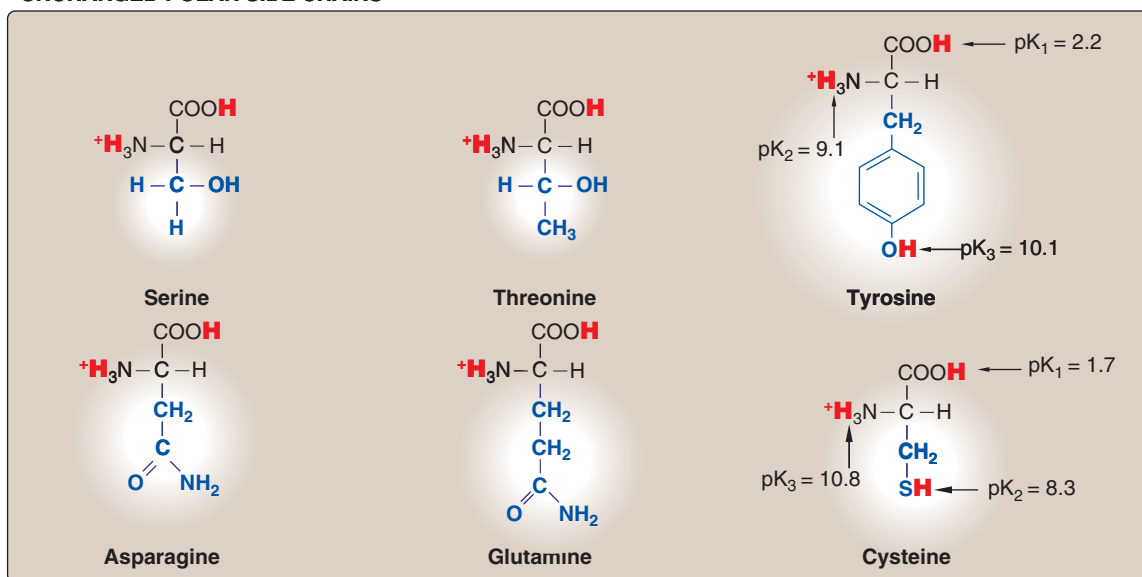


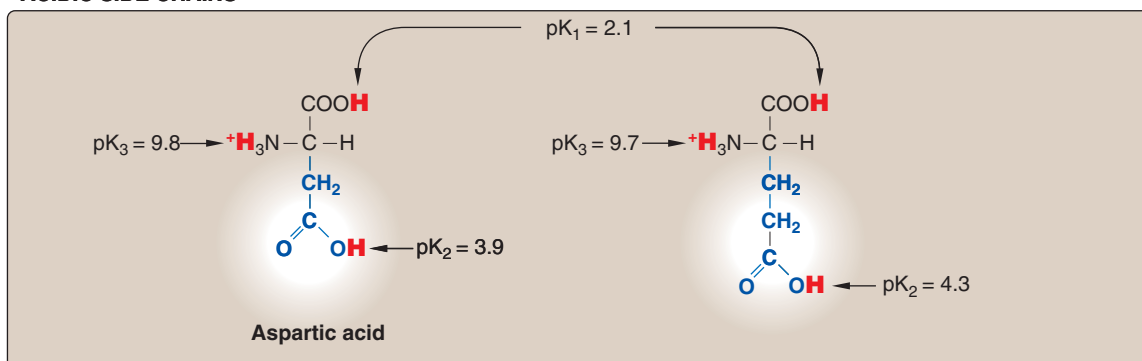
Figure 1.2

Classification of the 20 amino acids commonly found in proteins, according to the charge and polarity of their side chains at acidic pH is shown here and continues in Figure 1.3. Each amino acid is shown in its fully protonated form, with dissociable hydrogen ions represented in red print. The pK values for the α -carboxyl and α -amino groups of the nonpolar amino acids are similar to those shown for glycine. (Continued in Figure 1.3.)

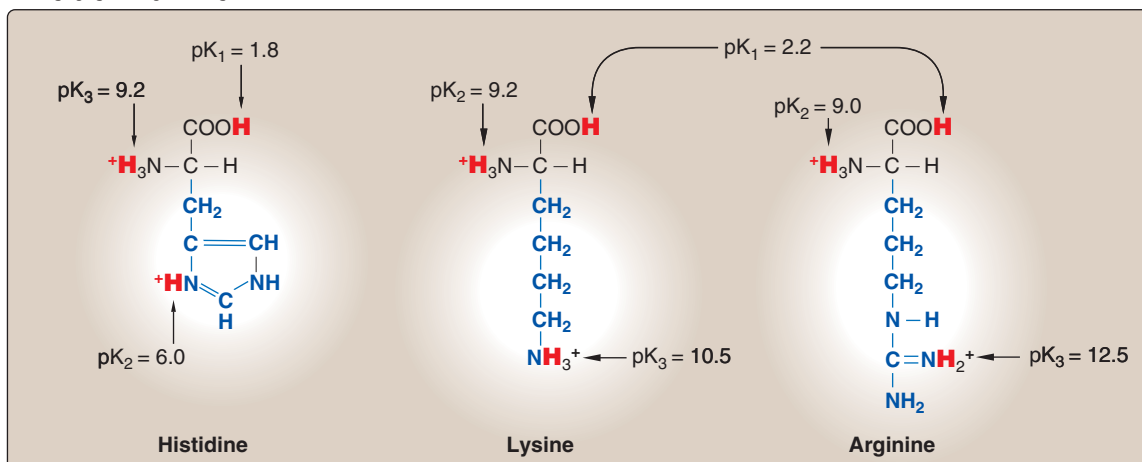
UNCHARGED POLAR SIDE CHAINS



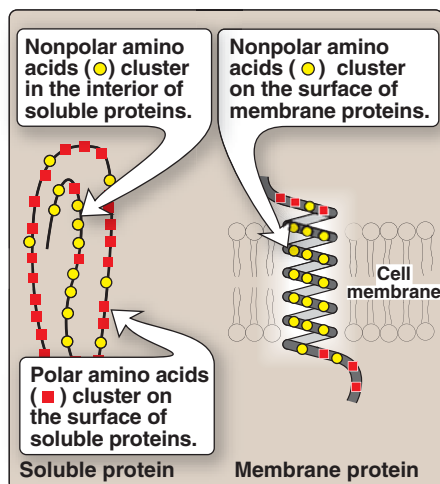
ACIDIC SIDE CHAINS



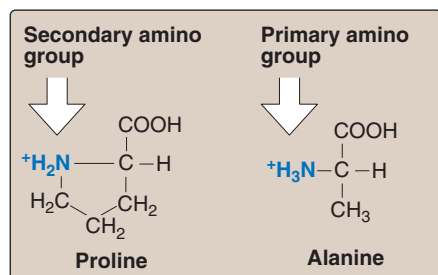
BASIC SIDE CHAINS

**Figure 1.3**

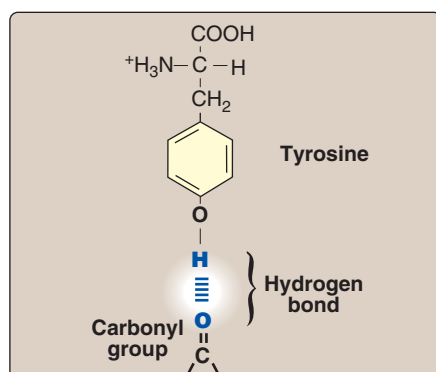
Classification of the 20 amino acids commonly found in proteins, according to the charge and polarity of their side chains at acidic pH (continued from Figure 1.2).

**Figure 1.4**

Location of nonpolar amino acids in soluble and membrane proteins.

**Figure 1.5**

Comparison of the secondary amino group found in proline with the primary amino group found in other amino acids, such as alanine.

**Figure 1.6**

Hydrogen bond between the phenolic hydroxyl group of tyrosine and another molecule containing a carbonyl group.

effect, is the result of the hydrophobicity of the nonpolar R-groups, which act much like droplets of oil that coalesce in an aqueous environment. The nonpolar R-groups thus fill up the interior of the folded protein and help give it its three-dimensional shape. However, for proteins that are located in a hydrophobic environment, such as a membrane, the nonpolar R-groups are found on the outside surface of the protein, interacting with the lipid environment (see Figure 1.4). The importance of these hydrophobic interactions in stabilizing protein structure is discussed on p. 19.

|| Sickle cell anemia, a sickling disease of red blood cells, results from the substitution of polar glutamate by nonpolar valine at the sixth position in the β subunit of hemoglobin (see p. 36).

- 2. Proline:** Proline differs from other amino acids in that proline's side chain and α -amino N form a rigid, five-membered ring structure (Figure 1.5). Proline, then, has a secondary (rather than a primary) amino group. It is frequently referred to as an imino acid. The unique geometry of proline contributes to the formation of the fibrous structure of collagen (see p. 45), and often interrupts the α -helices found in globular proteins (see p. 26).

B. Amino acids with uncharged polar side chains

These amino acids have zero net charge at neutral pH, although the side chains of cysteine and tyrosine can lose a proton at an alkaline pH (see Figure 1.3). Serine, threonine, and tyrosine each contain a polar hydroxyl group that can participate in hydrogen bond formation (Figure 1.6). The side chains of asparagine and glutamine each contain a carbonyl group and an amide group, both of which can also participate in hydrogen bonds.

- 1. Disulfide bond:** The side chain of cysteine contains a sulfhydryl group ($-SH$), which is an important component of the active site of many enzymes. In proteins, the $-SH$ groups of two cysteines can become oxidized to form a dimer, cystine, which contains a covalent cross-link called a disulfide bond ($-S-S-$). (See p. 19 for a further discussion of disulfide bond formation.)

|| Many extracellular proteins are stabilized by disulfide bonds. Albumin, a blood protein that functions as a transporter for a variety of molecules, is an example.

- 2. Side chains as sites of attachment for other compounds:** The polar hydroxyl group of serine, threonine, and, rarely, tyrosine, can serve as a site of attachment for structures such as a phosphate group. In addition, the amide group of asparagine, as well as the hydroxyl group of serine or threonine, can serve as a site of attachment for oligosaccharide chains in glycoproteins (see p. 165).

C. Amino acids with acidic side chains

The amino acids aspartic and glutamic acid are proton donors. At physiologic pH, the side chains of these amino acids are fully ionized, containing a negatively charged carboxylate group ($-\text{COO}^-$). They are, therefore, called aspartate or glutamate to emphasize that these amino acids are negatively charged at physiologic pH (see Figure 1.3).

D. Amino acids with basic side chains

The side chains of the basic amino acids accept protons (see Figure 1.3). At physiologic pH the side chains of lysine and arginine are fully ionized and positively charged. In contrast, histidine is weakly basic, and the free amino acid is largely uncharged at physiologic pH. However, when histidine is incorporated into a protein, its side chain can be either positively charged or neutral, depending on the ionic environment provided by the polypeptide chains of the protein. This is an important property of histidine that contributes to the role it plays in the functioning of proteins such as hemoglobin (see p. 31).

E. Abbreviations and symbols for commonly occurring amino acids

Each amino acid name has an associated three-letter abbreviation and a one-letter symbol (Figure 1.7). The one-letter codes are determined by the following rules:

- 1. Unique first letter:** If only one amino acid begins with a particular letter, then that letter is used as its symbol. For example, I = isoleucine.
- 2. Most commonly occurring amino acids have priority:** If more than one amino acid begins with a particular letter, the most common of these amino acids receives this letter as its symbol. For example, glycine is more common than glutamate, so G = glycine.
- 3. Similar sounding names:** Some one-letter symbols sound like the amino acid they represent. For example, F = phenylalanine, or W = tryptophan ("twyptophan" as Elmer Fudd would say).
- 4. Letter close to initial letter:** For the remaining amino acids, a one-letter symbol is assigned that is as close in the alphabet as possible to the initial letter of the amino acid, for example, K = lysine. Furthermore, B is assigned to Asx, signifying either aspartic acid or asparagine, Z is assigned to Glx, signifying either glutamic acid or glutamine, and X is assigned to an unidentified amino acid.

F. Optical properties of amino acids

The α -carbon of an amino acid is attached to four different chemical groups and is, therefore, a chiral or optically active carbon atom. Glycine is the exception because its α -carbon has two hydrogen substituents and, therefore, is optically inactive. Amino acids that have an asymmetric center at the α -carbon can exist in two forms, designated D and L, that are mirror images of each other (Figure 1.8). The two forms in each pair are termed stereoisomers, optical isomers, or enantiomers. All amino acids found in proteins are of the L-configuration. However, D-amino acids are found in some antibiotics and in plant and bacterial cell walls. (See p. 253 for a discussion of D-amino acid metabolism.)

1 Unique first letter:

Cysteine	=	Cys	=	C
Histidine	=	His	=	H
Isoleucine	=	Ile	=	I
Methionine	=	Met	=	M
Serine	=	Ser	=	S
Valine	=	Val	=	V

2 Most commonly occurring amino acids have priority:

Alanine	=	Ala	=	A
Glycine	=	Gly	=	G
Leucine	=	Leu	=	L
Proline	=	Pro	=	P
Threonine	=	Thr	=	T

3 Similar sounding names:

Arginine	=	Arg	=	R	("aRginine")
Asparagine	=	Asn	=	N	(contains N)
Aspartate	=	Asp	=	D	("asparDic")
Glutamate	=	Glu	=	E	("glutEamate")
Glutamine	=	Gln	=	Q	("Q-tamine")
Phenylalanine	=	Phe	=	F	("Fenylalanine")
Tyrosine	=	Tyr	=	Y	("tYrosine")
Tryptophan	=	Trp	=	W	(double ring in the molecule)

4 Letter close to initial letter:

Aspartate or asparagine	=	Asx	=	B	(near A)
Glutamate or glutamine	=	Glx	=	Z	
Lysine	=	Lys	=	K	(near L)
Undetermined amino acid	=		=	X	

Figure 1.7

Abbreviations and symbols for the commonly occurring amino acids.

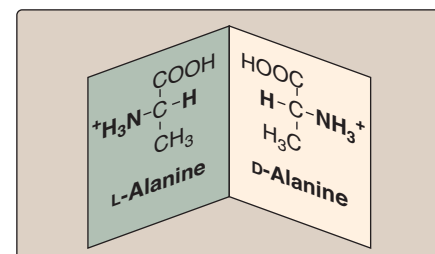


Figure 1.8

D and L forms of alanine are mirror images.

III. ACIDIC AND BASIC PROPERTIES OF AMINO ACIDS

Amino acids in aqueous solution contain weakly acidic α -carboxyl groups and weakly basic α -amino groups. In addition, each of the acidic and basic amino acids contains an ionizable group in its side chain. Thus, both free amino acids and some amino acids combined in peptide linkages can act as buffers. Recall that acids may be defined as proton donors and bases as proton acceptors. Acids (or bases) described as “weak” ionize to only a limited extent. The concentration of protons in aqueous solution is expressed as pH, where $\text{pH} = \log 1/[\text{H}^+]$ or $-\log [\text{H}^+]$. The quantitative relationship between the pH of the solution and concentration of a weak acid (HA) and its conjugate base (A^-) is described by the Henderson-Hasselbalch equation.

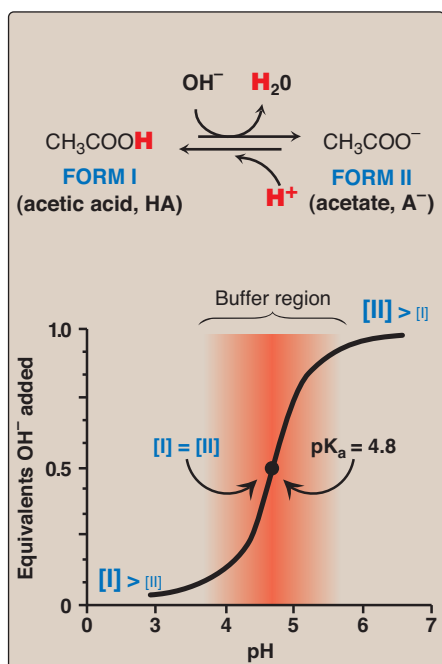
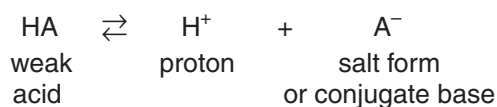


Figure 1.9
Titration curve of acetic acid.

A. Derivation of the equation

Consider the release of a proton by a weak acid represented by HA:



The “salt” or conjugate base, A^- , is the ionized form of a weak acid. By definition, the dissociation constant of the acid, K_a , is

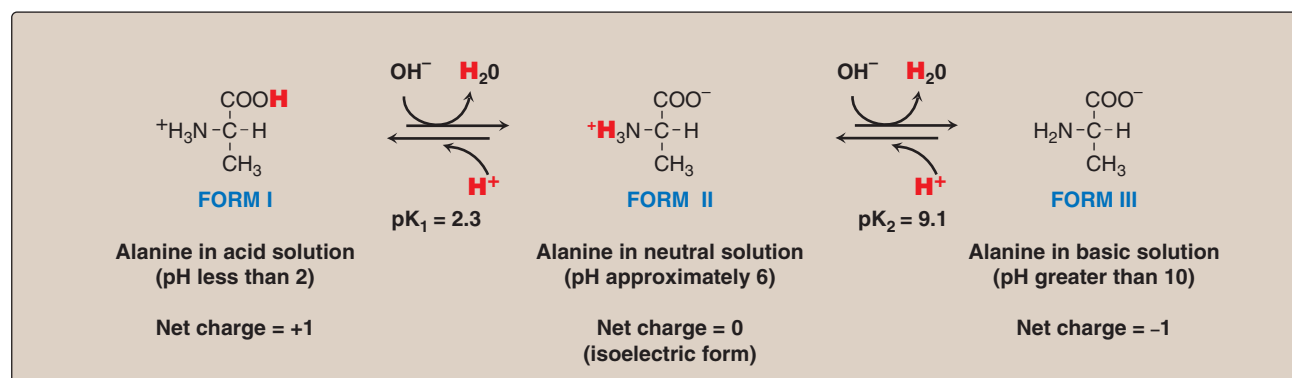
$$K_a = \frac{[\text{H}^+][\text{A}^-]}{[\text{HA}]}$$

[Note: The larger the K_a , the stronger the acid, because most of the HA has dissociated into H^+ and A^- . Conversely, the smaller the K_a , the less acid has dissociated and, therefore, the weaker the acid.] By solving for the $[\text{H}^+]$ in the above equation, taking the logarithm of both sides of the equation, multiplying both sides of the equation by -1 , and substituting $\text{pH} = -\log [\text{H}^+]$ and $\text{p}K_a = -\log K_a$, we obtain the Henderson-Hasselbalch equation:

$$\text{pH} = \text{p}K_a + \log \frac{[\text{A}^-]}{[\text{HA}]}$$

B. Buffers

A buffer is a solution that resists change in pH following the addition of an acid or base. A buffer can be created by mixing a weak acid (HA) with its conjugate base (A^-). If an acid such as HCl is then added to such a solution, A^- can neutralize it, in the process being converted to HA. If a base is added, HA can neutralize it, in the process being converted to A^- . Maximum buffering capacity occurs at a pH equal to the $\text{p}K_a$, but a conjugate acid/base pair can still serve as an effective buffer when the pH of a solution is within approximately ± 1 pH unit of the $\text{p}K_a$. If the

**Figure 1.10**

Ionic forms of alanine in acidic, neutral, and basic solutions.

amounts of HA and A⁻ are equal, the pH is equal to the pK_a. As shown in Figure 1.9, a solution containing acetic acid (HA = CH₃-COOH) and acetate (A⁻ = CH₃-COO⁻) with a pK_a of 4.8 resists a change in pH from pH 3.8 to 5.8, with maximum buffering at pH 4.8. At pH values less than the pK_a, the protonated acid form (CH₃-COOH) is the predominant species. At pH values greater than the pK_a, the deprotonated base form (CH₃-COO⁻) is the predominant species in solution.

C. Titration of an amino acid

1. Dissociation of the carboxyl group: The titration curve of an amino acid can be analyzed in the same way as described for acetic acid. Consider alanine, for example, which contains both an α-carboxyl and an α-amino group. At a low (acidic) pH, both of these groups are protonated (shown in Figure 1.10). As the pH of the solution is raised, the -COOH group of Form I can dissociate by donating a proton to the medium. The release of a proton results in the formation of the carboxylate group, -COO⁻. This structure is shown as Form II, which is the dipolar form of the molecule (see Figure 1.10). This form, also called a zwitterion, is the isoelectric form of alanine, that is, it has an overall (net) charge of zero.

2. Application of the Henderson-Hasselbalch equation: The dissociation constant of the carboxyl group of an amino acid is called K₁, rather than K_a, because the molecule contains a second titratable group. The Henderson-Hasselbalch equation can be used to analyze the dissociation of the carboxyl group of alanine in the same way as described for acetic acid:

$$K_1 = \frac{[\text{H}^+][\text{II}]}{[\text{I}]}$$

where I is the fully protonated form of alanine, and II is the isoelectric form of alanine (see Figure 1.10). This equation can be rearranged and converted to its logarithmic form to yield:

$$\text{pH} = \text{pK}_1 + \log \frac{[\text{II}]}{[\text{I}]}$$

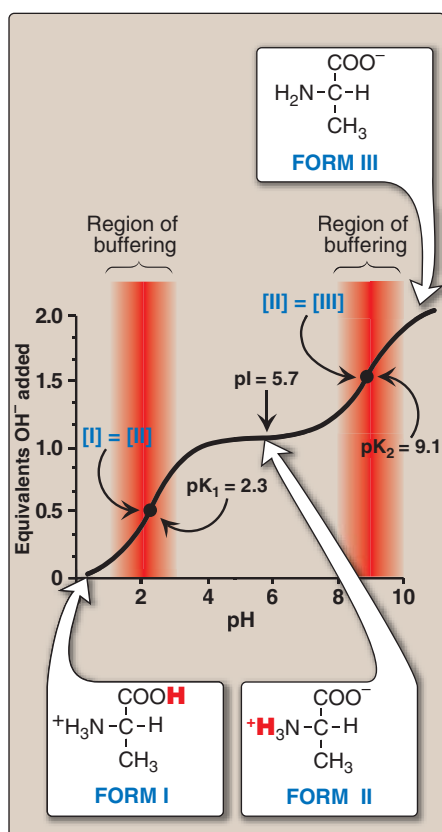


Figure 1.11
The titration curve of alanine.

3. Dissociation of the amino group: The second titratable group of alanine is the amino ($-\text{NH}_3^+$) group shown in Figure 1.10. This is a much weaker acid than the $-\text{COOH}$ group and, therefore, has a much smaller dissociation constant, K_2 . [Note: Its pK_a is therefore larger.] Release of a proton from the protonated amino group of Form II results in the fully deprotonated form of alanine, Form III (see Figure 1.10).

4. pKs of alanine: The sequential dissociation of protons from the carboxyl and amino groups of alanine is summarized in Figure 1.10. Each titratable group has a pK_a that is numerically equal to the pH at which exactly one half of the protons have been removed from that group. The pK_a for the most acidic group ($-\text{COOH}$) is pK_1 , whereas the pK_a for the next most acidic group ($-\text{NH}_3^+$) is pK_2 .

5. Titration curve of alanine: By applying the Henderson-Hasselbalch equation to each dissociable acidic group, it is possible to calculate the complete titration curve of a weak acid. Figure 1.11 shows the change in pH that occurs during the addition of base to the fully protonated form of alanine (I) to produce the completely deprotonated form (III). Note the following:

a. Buffer pairs: The $-\text{COOH}/-\text{COO}^-$ pair can serve as a buffer in the pH region around pK_1 , and the $-\text{NH}_3^+/-\text{NH}_2$ pair can buffer in the region around pK_2 .

b. When pH = pK: When the pH is equal to pK_1 (2.3), equal amounts of Forms I and II of alanine exist in solution. When the pH is equal to pK_2 (9.1), equal amounts of Forms II and III are present in solution.

c. Isoelectric point: At neutral pH, alanine exists predominantly as the dipolar Form II in which the amino and carboxyl groups are ionized, but the net charge is zero. The isoelectric point (pI) is the pH at which an amino acid is electrically neutral, that is, in which the sum of the positive charges equals the sum of the negative charges. For an amino acid, such as alanine, that has only two dissociable hydrogens (one from the α -carboxyl and one from the α -amino group), the pI is the average of pK_1 and pK_2 ($\text{pI} = [\text{pK}_1 + \text{pK}_2]/2 = 5.7$, see Figure 1.11). The pI is thus midway between pK_1 (2.3) and pK_2 (9.1). pI corresponds to the pH at which the Form II (with a net charge of zero) predominates, and at which there are also equal amounts of Forms I (net charge of +1) and III (net charge of -1).

Separation of plasma proteins by charge typically is done at a pH above the pI of the major proteins, thus, the charge on the proteins is negative. In an electric field, the proteins will move toward the positive electrode at a rate determined by their net negative charge. Variations in the mobility pattern are suggestive of certain diseases.

6. Net charge of amino acids at neutral pH: At physiologic pH, amino acids have a negatively charged group ($-\text{COO}^-$) and a positively charged group ($-\text{NH}_3^+$), both attached to the α -carbon. [Note: Glutamate, aspartate, histidine, arginine, and lysine have additional potentially charged groups in their side chains.] Substances, such as amino acids, that can act either as an acid or a base are defined as amphoteric, and are referred to as ampholytes (amphoteric electrolytes).

D. Other applications of the Henderson-Hasselbalch equation

The Henderson-Hasselbalch equation can be used to calculate how the pH of a physiologic solution responds to changes in the concentration of a weak acid and/or its corresponding "salt" form. For example, in the bicarbonate buffer system, the Henderson-Hasselbalch equation predicts how shifts in the bicarbonate ion concentration, $[\text{HCO}_3^-]$, and CO_2 influence pH (Figure 1.12A). The equation is also useful for calculating the abundance of ionic forms of acidic and basic drugs. For example, most drugs are either weak acids or weak bases (Figure 1.12B). Acidic drugs (HA) release a proton (H^+), causing a charged anion (A^-) to form.



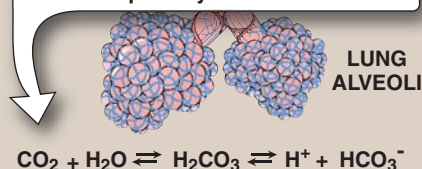
Weak bases (BH^+) can also release a H^+ . However, the protonated form of basic drugs is usually charged, and the loss of a proton produces the uncharged base (B).



A drug passes through membranes more readily if it is uncharged. Thus, for a weak acid such as aspirin, the uncharged HA can permeate through membranes and A^- cannot. For a weak base, such as morphine, the uncharged form, B, penetrates through the cell membrane and BH^+ does not. Therefore, the effective concentration of the permeable form of each drug at its absorption site is determined by the relative concentrations of the charged and uncharged forms. The ratio between the two forms is determined by the pH at the site of absorption, and by the strength of the weak acid or base, which is represented by the pK_a of the ionizable group. The Henderson-Hasselbalch equation is useful in determining how much drug is found on either side of a membrane that separates two compartments that differ in pH, for example, the stomach (pH 1.0–1.5) and blood plasma (pH 7.4).

A BICARBONATE AS A BUFFER

- $\text{pH} = \text{pK} + \log \frac{[\text{HCO}_3^-]}{[\text{CO}_2]}$
- An increase in HCO_3^- causes the pH to rise.
- Pulmonary obstruction causes an increase in carbon dioxide and causes the pH to fall, resulting in respiratory acidosis.



B DRUG ABSORPTION

- $\text{pH} = \text{pK} + \log \frac{[\text{Drug}^-]}{[\text{Drug-H}]}$
- At the pH of the stomach (1.5), a drug like aspirin (weak acid, $\text{pK} = 3.5$) will be largely protonated (COOH) and, thus, uncharged.
- Uncharged drugs generally cross membranes more rapidly than charged molecules.

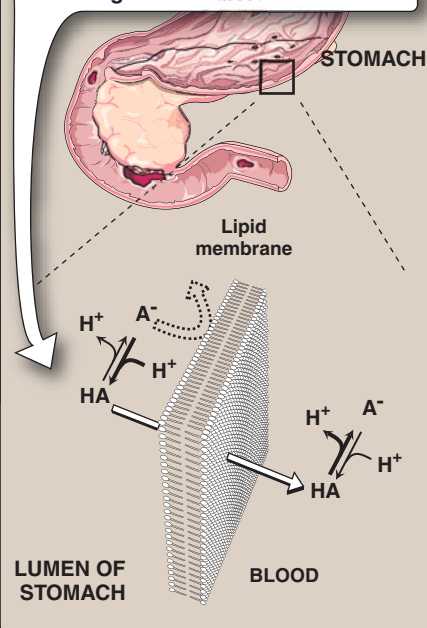


Figure 1.12

The Henderson-Hasselbalch equation is used to predict: A, changes in pH as the concentrations of HCO_3^- or CO_2 are altered; or B, the ionic forms of drugs.

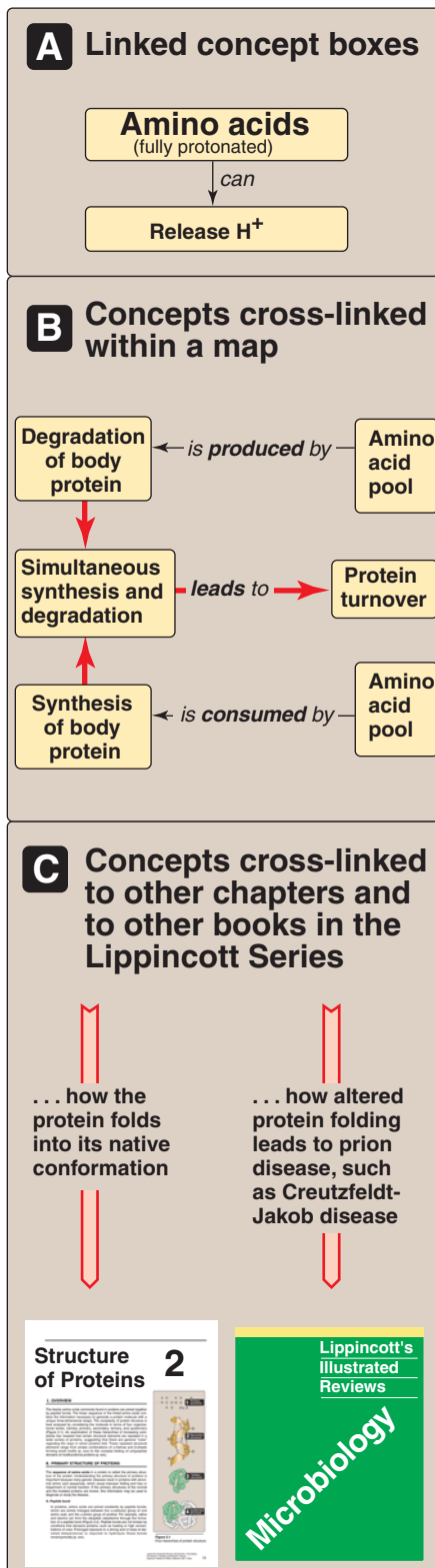


Figure 1.13
Symbols used in concept maps.

IV. CONCEPT MAPS

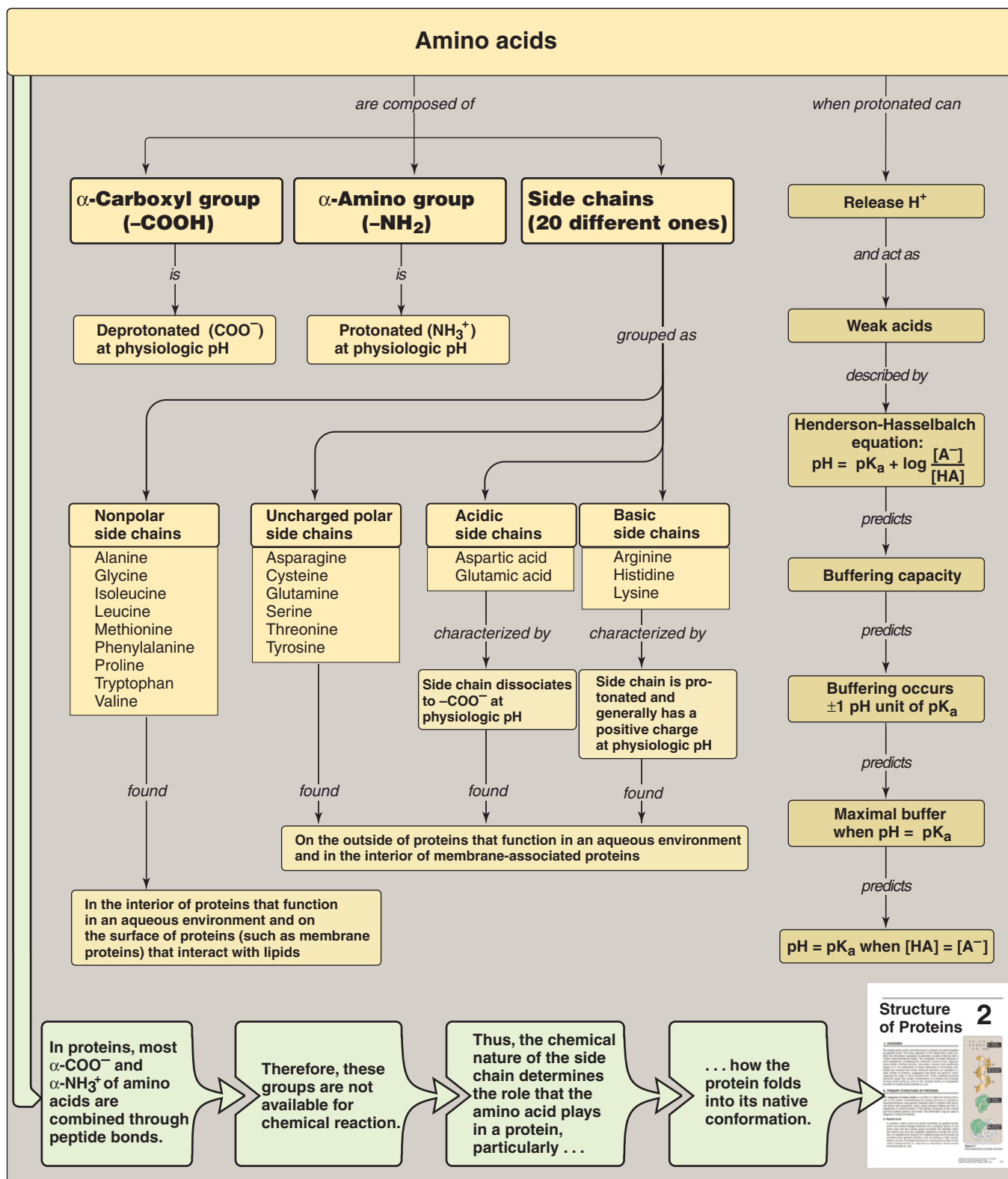
Students sometimes view biochemistry as a blur of facts or equations to be memorized, rather than a body of concepts to be understood. Details provided to enrich understanding of these concepts inadvertently turn into distractions. What seems to be missing is a road map—a guide that provides the student with an intuitive understanding of how various topics fit together to make sense. The authors have, therefore, created a series of biochemical concept maps to graphically illustrate relationships between ideas presented in a chapter, and to show how the information can be grouped or organized. A concept map is, thus, a tool for visualizing the connections between concepts. Material is represented in a hierarchic fashion, with the most inclusive, most general concepts at the top of the map, and the more specific, less general concepts arranged beneath. The concept maps ideally function as templates or guides for organizing information, so the student can readily find the best ways to integrate new information into knowledge they already possess.

A. How is a concept map constructed?

- 1. Concept boxes and links:** Educators define concepts as “perceived regularities in events or objects.” In our biochemical maps, concepts include abstractions (for example, free energy), processes (for example, oxidative phosphorylation), and compounds (for example, glucose 6-phosphate). These broadly defined concepts are prioritized with the central idea positioned at the top of the page. The concepts that follow from this central idea are then drawn in boxes (Figure 1.13A). The size of the type indicates the relative importance of each idea. Lines are drawn between concept boxes to show which are related. The label on the line defines the relationship between two concepts, so that it reads as a valid statement, that is, the connection creates meaning. The lines with arrowheads indicate in which direction the connection should be read (Figure 1.14).
- 2. Cross-links:** Unlike linear flow charts or outlines, concept maps may contain cross-links that allow the reader to visualize complex relationships between ideas represented in different parts of the map (Figure 1.13B), or between the map and other chapters in this book or companion books in the series (Figure 1.13C). Cross-links can thus identify concepts that are central to more than one discipline, empowering students to be effective in clinical situations, and on the United States Medical Licensure Examination (USMLE) or other examinations, that bridge disciplinary boundaries. Students learn to visually perceive nonlinear relationships between facts, in contrast to cross-referencing within linear text.

V. CHAPTER SUMMARY

Each amino acid has an **α -carboxyl group** and a primary **α -amino group** (except for proline, which has a **secondary amino group**). At physiologic pH, the α -carboxyl group is dissociated, forming the negatively charged carboxylate ion ($-\text{COO}^-$), and the α -amino group is protonated ($-\text{NH}_3^+$). Each amino acid also contains one of 20 distinctive

**Figure 1.14**

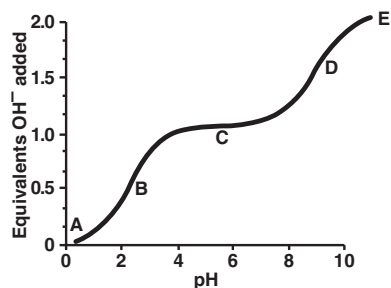
Key concept map for amino acids.

side chains attached to the α -carbon atom. The chemical nature of this side chain determines the function of an amino acid in a protein, and provides the basis for classification of the amino acids as **nonpolar**, **uncharged polar**, **acidic**, or **basic**. All free amino acids, plus charged amino acids in peptide chains, can serve as **buffers**. The quantitative relationship between the pH of a solution and the concentration of a weak acid (HA) and its conjugate base (A^-) is described by the **Henderson-Hasselbalch equation**. Buffering occurs within ± 1 pH unit of the pK_a , and is maximal when $pH = pK_a$, at which $[A^-] = [HA]$. The α -carbon of each amino acid (except glycine) is attached to four different chemical groups and is, therefore, a **chiral** or **optically active** carbon atom. Only the L-form of amino acids is found in proteins synthesized by the human body.

Study Questions

Choose the ONE correct answer.

- 1.1 The letters A through E designate certain regions on the titration curve for glycine (shown below). Which one of the following statements concerning this curve is correct?



Correct answer = C. C represents the isoelectric point or pI, and as such is midway between pK_1 and pK_2 for this monoamino monocarboxylic acid. Glycine is fully protonated at Point A. Point B represents a region of maximum buffering, as does Point D. Point E represents the region where glycine is fully deprotonated.

- A. Point A represents the region where glycine is deprotonated.
 B. Point B represents a region of minimal buffering.
 C. Point C represents the region where the net charge on glycine is zero.
 D. Point D represents the pK of glycine's carboxyl group.
 E. Point E represents the pI for glycine.
- 1.2 Which one of the following statements concerning the peptide shown below is correct?
 Gly-Cys-Glu-Ser-Asp-Arg-Cys
- A. The peptide contains glutamine.
 B. The peptide contains a side chain with a secondary amino group.
 C. The peptide contains a majority of amino acids with side chains that would be positively charged at pH 7.
 D. The peptide is able to form an internal disulfide bond.
- 1.3 Given that the pI for glycine is 6.1, to which electrode, positive or negative, will glycine move in an electric field at pH 2? Explain.

Correct answer = D. The two cysteine residues can, under oxidizing conditions, form a disulfide bond. Glutamine's 3-letter abbreviation is Gln. Proline (Pro) contains a secondary amino group. Only one (Arg) of the seven would have a positively charged side chain at pH 7.

Correct answer = negative electrode. When the pH is less than the pI, the charge on glycine is positive because the α -amino group is fully protonated. (Recall that glycine has H as its R group).

Structure of Proteins

2

I. OVERVIEW

The 20 amino acids commonly found in proteins are joined together by peptide bonds. The linear sequence of the linked amino acids contains the information necessary to generate a protein molecule with a unique three-dimensional shape. The complexity of protein structure is best analyzed by considering the molecule in terms of four organizational levels, namely, primary, secondary, tertiary, and quaternary (Figure 2.1). An examination of these hierarchies of increasing complexity has revealed that certain structural elements are repeated in a wide variety of proteins, suggesting that there are general “rules” regarding the ways in which proteins achieve their native, functional form. These repeated structural elements range from simple combinations of α -helices and β -sheets forming small motifs, to the complex folding of polypeptide domains of multifunctional proteins (see p. 18).

II. PRIMARY STRUCTURE OF PROTEINS

The sequence of amino acids in a protein is called the primary structure of the protein. Understanding the primary structure of proteins is important because many genetic diseases result in proteins with abnormal amino acid sequences, which cause improper folding and loss or impairment of normal function. If the primary structures of the normal and the mutated proteins are known, this information may be used to diagnose or study the disease.

A. Peptide bond

In proteins, amino acids are joined covalently by peptide bonds, which are amide linkages between the α -carboxyl group of one amino acid and the α -amino group of another. For example, valine and alanine can form the dipeptide valylalanine through the formation of a peptide bond (Figure 2.2). Peptide bonds are not broken by conditions that denature proteins, such as heating or high concentrations of urea (see p. 20). Prolonged exposure to a strong acid or base at elevated temperatures is required to hydrolyze these bonds nonenzymically.

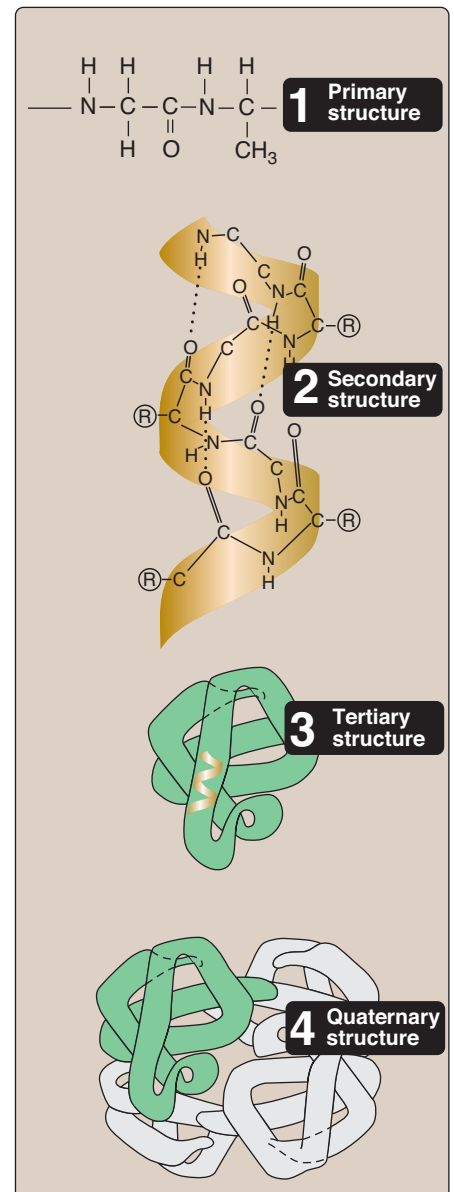
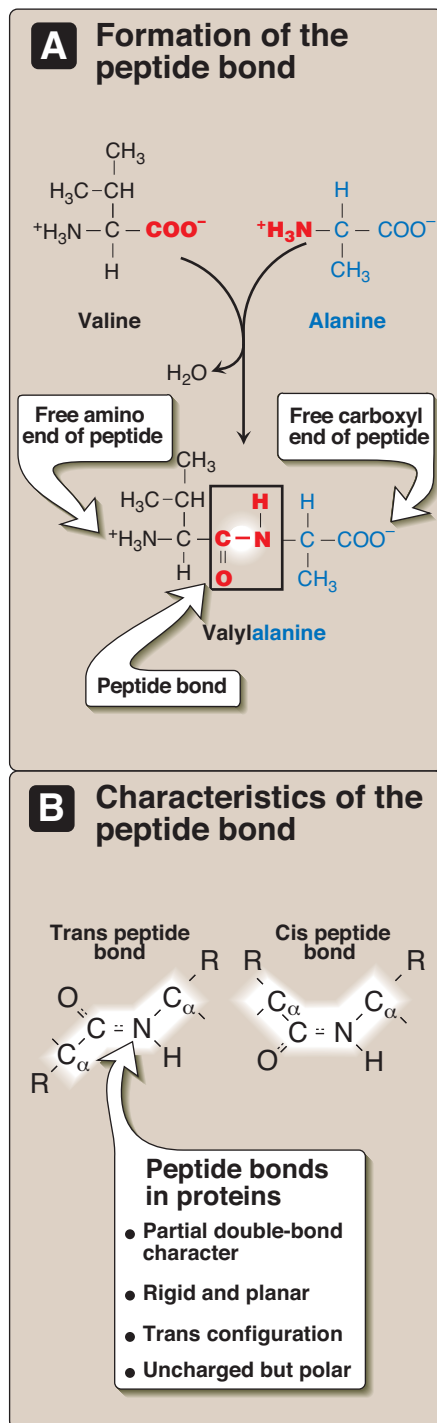


Figure 2.1
Four hierarchies of protein structure.

**Figure 2.2**

A. Formation of a peptide bond, showing the structure of the dipeptide valylalanine.
 B. Characteristics of the peptide bond.

1. Naming the peptide: By convention, the free amino end (N-terminal) of the peptide chain is written to the left and the free carboxyl end (C-terminal) to the right. Therefore, all amino acid sequences are read from the N- to the C-terminal end of the peptide. For example, in Figure 2.2A, the order of the amino acids is “valine, alanine.” Linkage of many amino acids through peptide bonds results in an unbranched chain called a polypeptide. Each component amino acid in a polypeptide is called a “residue” because it is the portion of the amino acid remaining after the atoms of water are lost in the formation of the peptide bond. When a polypeptide is named, all amino acid residues have their suffixes (-ine, -an, -ic, or -ate) changed to -yl, with the exception of the C-terminal amino acid. For example, a tripeptide composed of an N-terminal valine, a glycine, and a C-terminal leucine is called valylglycylleucine.

2. Characteristics of the peptide bond: The peptide bond has a partial double-bond character, that is, it is shorter than a single bond, and is rigid and planar (Figure 2.2B). This prevents free rotation around the bond between the carbonyl carbon and the nitrogen of the peptide bond. However, the bonds between the α -carbons and the α -amino or α -carboxyl groups can be freely rotated (although they are limited by the size and character of the R-groups). This allows the polypeptide chain to assume a variety of possible configurations. The peptide bond is generally a trans bond (instead of cis, see Figure 2.2B), in large part because of steric interference of the R-groups when in the cis position.

3. Polarity of the peptide bond: Like all amide linkages, the —C=O and —NH groups of the peptide bond are uncharged, and neither accept nor release protons over the pH range of 2–12. Thus, the charged groups present in polypeptides consist solely of the N-terminal (α -amino) group, the C-terminal (α -carboxyl) group, and any ionized groups present in the side chains of the constituent amino acids. The —C=O and —NH groups of the peptide bond are polar, and are involved in hydrogen bonds, for example, in α -helices and β -sheet structures, described on pp. 16–17.

B. Determination of the amino acid composition of a polypeptide

The first step in determining the primary structure of a polypeptide is to identify and quantitate its constituent amino acids. A purified sample of the polypeptide to be analyzed is first hydrolyzed by strong acid at 110°C for 24 hours. This treatment cleaves the peptide bonds and releases the individual amino acids, which can be separated by cation-exchange chromatography. In this technique, a mixture of amino acids is applied to a column that contains a resin to which a negatively charged group is tightly attached. [Note: If the attached group is positively charged, the column becomes an anion-exchange column.] The amino acids bind to the column with different affinities, depending on their charges, hydrophobicity, and other characteristics. Each amino acid is sequentially released from the chromatography column by eluting with solutions of increasing ionic strength and pH (Figure 2.3). The separated amino acids contained in the eluate from the column are quantitated by heating them with ninhydrin—a reagent that forms a purple compound with most

amino acids, ammonia, and amines. The amount of each amino acid is determined spectrophotometrically by measuring the amount of light absorbed by the ninhydrin derivative. The analysis described above is performed using an amino acid analyzer—an automated machine whose components are depicted in Figure 2.3.

C. Sequencing of the peptide from its N-terminal end

Sequencing is a stepwise process of identifying the specific amino acid at each position in the peptide chain, beginning at the N-terminal end. Phenylisothiocyanate, known as Edman reagent, is used to label the amino-terminal residue under mildly alkaline conditions (Figure 2.4). The resulting phenylthiohydantoin (PTH) derivative introduces an instability in the N-terminal peptide bond that can be selectively hydrolyzed without cleaving the other peptide bonds. The identity of the amino acid derivative can then be determined. Edman reagent can be applied repeatedly to the shortened peptide obtained in each previous cycle.

D. Cleavage of the polypeptide into smaller fragments

Many polypeptides have a primary structure composed of more than 100 amino acids. Such molecules cannot be sequenced directly from end to end. However, these large molecules can be cleaved at specific sites, and the resulting fragments sequenced. By using more than one cleaving agent (enzymes and/or chemicals) on separate samples of the purified polypeptide, overlapping fragments can be generated that permit the proper ordering of the sequenced fragments, thus providing a complete amino acid sequence of the large polypeptide (Figure 2.5). Enzymes that hydrolyze peptide bonds are termed *peptidases* (*proteases*). [Note: *Exopeptidases* cut at the ends of proteins, and are divided into *aminopeptidases* and *carboxypeptidases*. *Carboxypeptidases* are used in determining the C-terminal amino acid. *Endopeptidases* cleave within a protein.]

E. Determination of a protein's primary structure by DNA sequencing

The sequence of nucleotides in a protein-coding region of the DNA specifies the amino acid sequence of a polypeptide. Therefore, if the nucleotide sequence can be determined, it is possible, from knowledge of the genetic code (see p. 431), to translate the sequence of nucleotides into the corresponding amino acid sequence of that

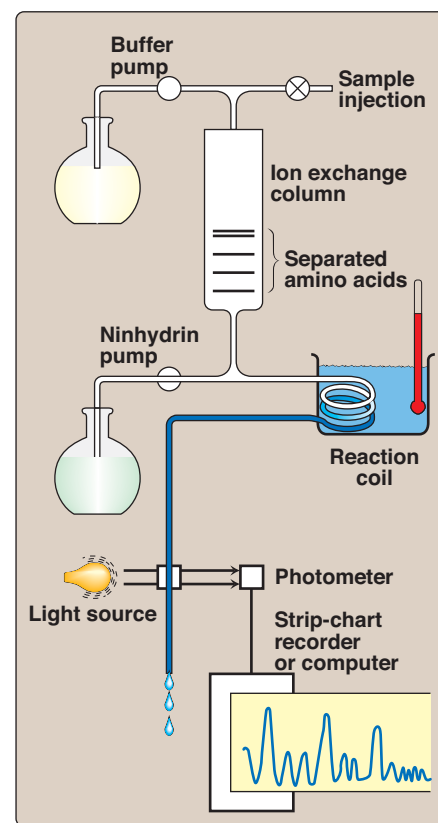


Figure 2.3

Determination of the amino acid composition of a polypeptide using an amino acid analyzer.

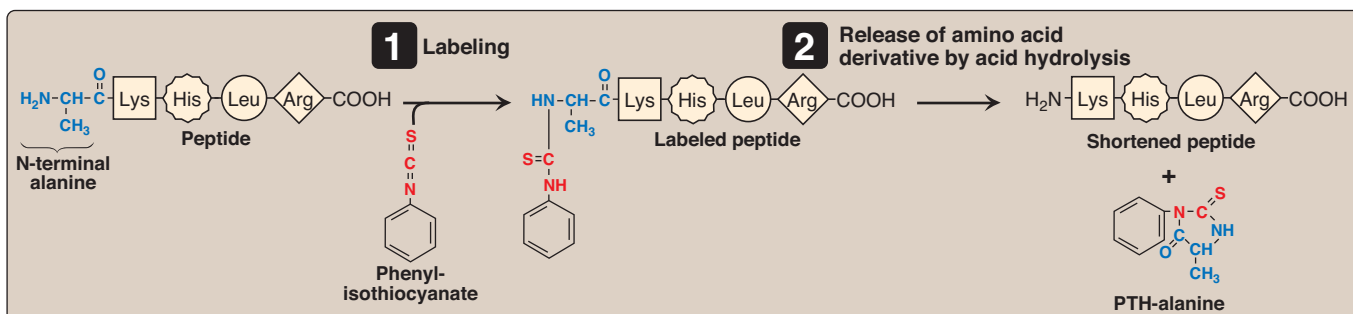
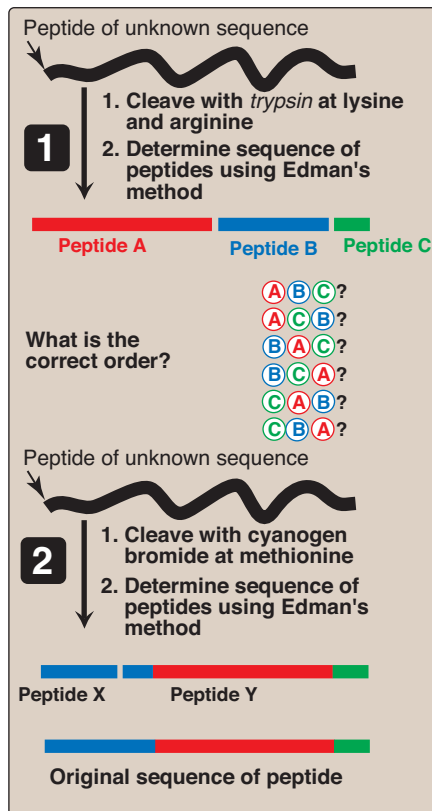
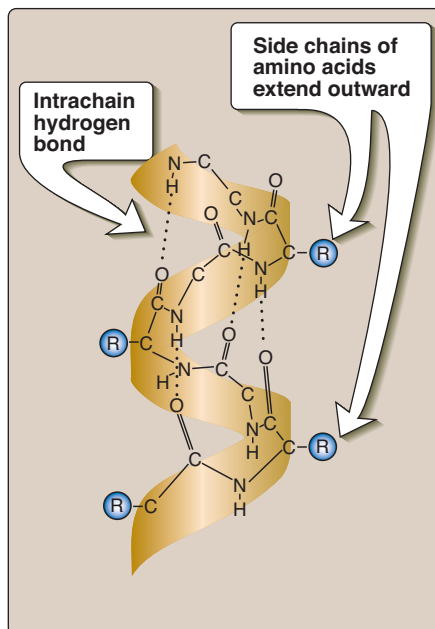


Figure 2.4

Determination of the amino-terminal residue of a polypeptide by Edman degradation.

**Figure 2.5**

Overlapping of peptides produced by the action of *trypsin* and cyanogen bromide.

**Figure 2.6**

α -Helix showing peptide backbone.

polypeptide. This indirect process, although routinely used to obtain the amino acid sequences of proteins, has the limitations of not being able to predict the positions of disulfide bonds in the folded chain, and of not identifying any amino acids that are modified after their incorporation into the polypeptide (posttranslational modification, see p. 443). Therefore, direct protein sequencing is an extremely important tool for determining the true character of the primary sequence of many polypeptides.

III. SECONDARY STRUCTURE OF PROTEINS

The polypeptide backbone does not assume a random three-dimensional structure, but instead generally forms regular arrangements of amino acids that are located near to each other in the linear sequence. These arrangements are termed the secondary structure of the polypeptide. The α -helix, β -sheet, and β -bend (β -turn) are examples of secondary structures frequently encountered in proteins. [Note: The collagen α -chain helix, another example of secondary structure, is discussed on p. 45.]

A. α -Helix

Several different polypeptide helices are found in nature, but the α -helix is the most common. It is a spiral structure, consisting of a tightly packed, coiled polypeptide backbone core, with the side chains of the component amino acids extending outward from the central axis to avoid interfering sterically with each other (Figure 2.6). A very diverse group of proteins contains α -helices. For example, the keratins are a family of closely related, fibrous proteins whose structure is nearly entirely α -helical. They are a major component of tissues such as hair and skin, and their rigidity is determined by the number of disulfide bonds between the constituent polypeptide chains. In contrast to keratin, myoglobin, whose structure is also highly α -helical, is a globular, flexible molecule (see p. 26).

1. **Hydrogen bonds:** An α -helix is stabilized by extensive hydrogen bonding between the peptide-bond carbonyl oxygens and amide hydrogens that are part of the polypeptide backbone (see Figure 2.6). The hydrogen bonds extend up and are parallel to the spiral from the carbonyl oxygen of one peptide bond to the $-\text{NH}-$ group of a peptide linkage four residues ahead in the polypeptide. This ensures that all but the first and last peptide bond components are linked to each other through intrachain hydrogen bonds. Hydrogen bonds are individually weak, but they collectively serve to stabilize the helix.
2. **Amino acids per turn:** Each turn of an α -helix contains 3.6 amino acids. Thus, amino acid residues spaced three or four residues apart in the primary sequence are spatially close together when folded in the α -helix.
3. **Amino acids that disrupt an α -helix:** Proline disrupts an α -helix because its secondary amino group is not geometrically compatible with the right-handed spiral of the α -helix. Instead, it inserts a kink in the chain, which interferes with the smooth, helical structure. Large numbers of charged amino acids (for example, gluta-

mate, aspartate, histidine, lysine, or arginine) also disrupt the helix by forming ionic bonds, or by electrostatically repelling each other. Finally, amino acids with bulky side chains, such as tryptophan, or amino acids, such as valine or isoleucine, that branch at the β -carbon (the first carbon in the R-group, next to the α -carbon) can interfere with formation of the α -helix if they are present in large numbers.

B. β -Sheet

The β -sheet is another form of secondary structure in which all of the peptide bond components are involved in hydrogen bonding (Figure 2.7A). The surfaces of β -sheets appear “pleated,” and these structures are, therefore, often called “ β -pleated sheets.” When illustrations are made of protein structure, β -strands are often visualized as broad arrows (Figure 2.7B).

- 1. Comparison of a β -sheet and an α -helix:** Unlike the α -helix, β -sheets are composed of two or more peptide chains (β -strands), or segments of polypeptide chains, which are almost fully extended. Note also that in β -sheets the hydrogen bonds are perpendicular to the polypeptide backbone (see Figure 2.7A).
- 2. Parallel and antiparallel sheets:** A β -sheet can be formed from two or more separate polypeptide chains or segments of polypeptide chains that are arranged either antiparallel to each other (with the N-terminal and C-terminal ends of the β -strands alternating as shown in Figure 2.7B), or parallel to each other (with all the N-termini of the β -strands together as shown in Figure 2.7C). When the hydrogen bonds are formed between the polypeptide backbones of separate polypeptide chains, they are termed interchain bonds. A β -sheet can also be formed by a single polypeptide chain folding back on itself (see Figure 2.7C). In this case, the hydrogen bonds are intrachain bonds. In globular proteins, β -sheets always have a right-handed curl, or twist, when viewed along the polypeptide backbone. [Note: Twisted β -sheets often form the core of globular proteins.]

|| The α -helix and β -sheet structures provide maximal hydrogen bonding for peptide bond components within the interior of polypeptides.

C. β -Bends (reverse turns, β -turns)

β -Bends reverse the direction of a polypeptide chain, helping it form a compact, globular shape. They are usually found on the surface of protein molecules, and often include charged residues. [Note: β -Bends were given this name because they often connect successive strands of antiparallel β -sheets.] β -Bends are generally composed of four amino acids, one of which may be proline—the amino acid that causes a “kink” in the polypeptide chain. Glycine, the amino acid with the smallest R-group, is also frequently found in β -bends. β -Bends are stabilized by the formation of hydrogen and ionic bonds.

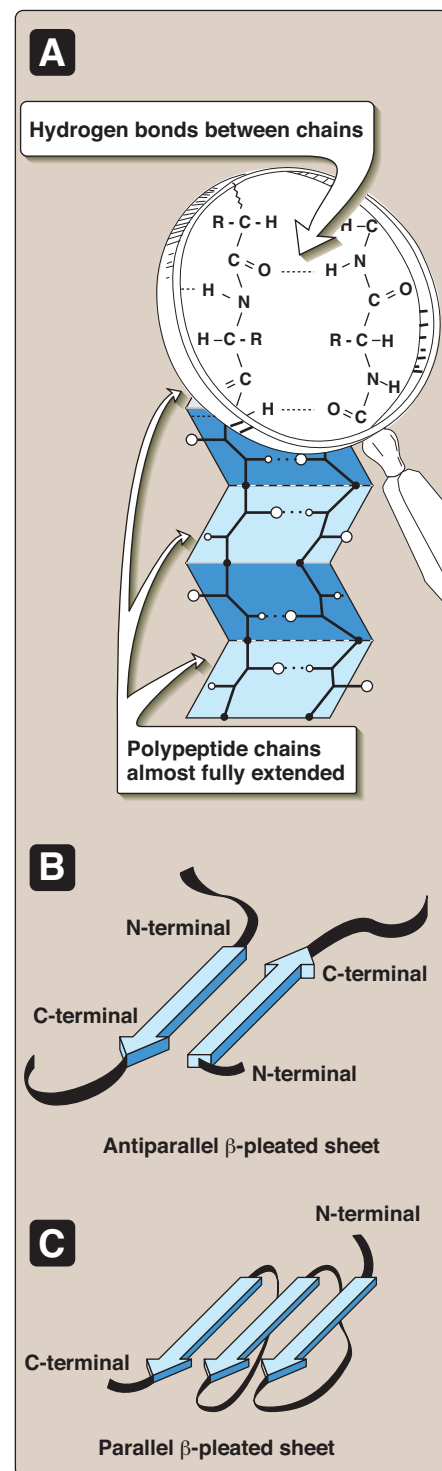
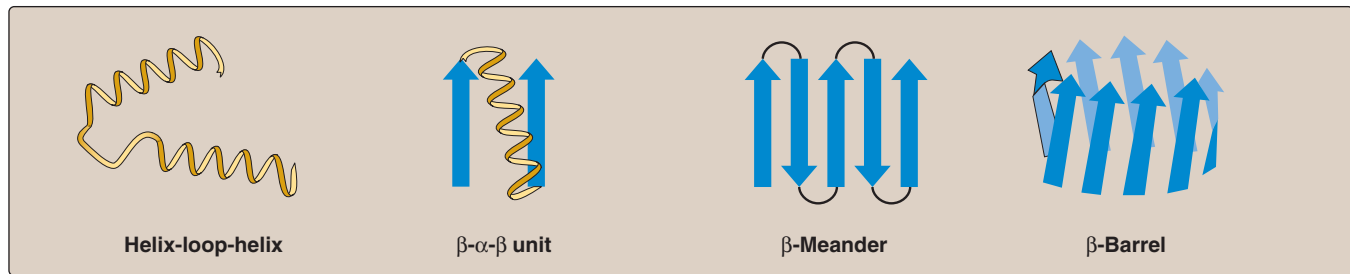


Figure 2.7

A. Structure of a β -sheet. B. An antiparallel β -sheet with the β -strands represented as broad arrows. C. A parallel β -sheet formed from a single polypeptide chain folding back on itself.

**Figure 2.8**

Some common structural motifs combining α -helices and/or β -sheets. The names describe their schematic appearance.

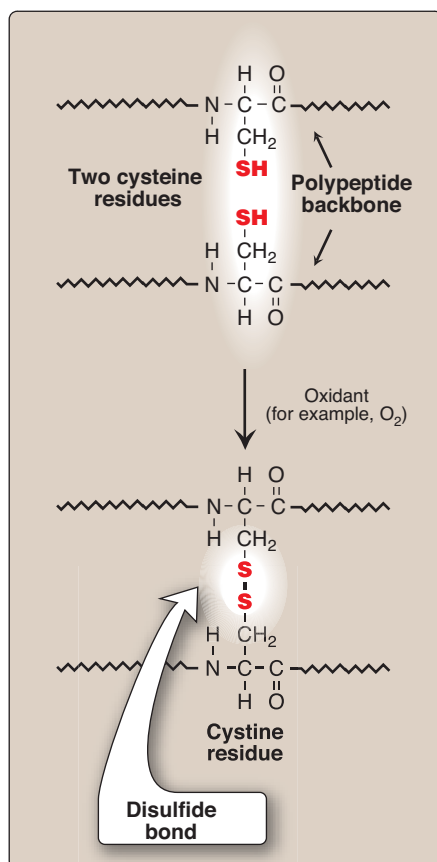
D. Nonrepetitive secondary structure

Approximately one half of an average globular protein is organized into repetitive structures, such as the α -helix and/or β -sheet. The remainder of the polypeptide chain is described as having a loop or coil conformation. These nonrepetitive secondary structures are not “random,” but rather simply have a less regular structure than those described above. [Note: The term “random coil” refers to the disordered structure obtained when proteins are denatured (see p. 20).]

E. Supersecondary structures (motifs)

Globular proteins are constructed by combining secondary structural elements (α -helices, β -sheets, nonrepetitive sequences). These form primarily the core region—that is, the interior of the molecule. They are connected by loop regions (for example, β -bends) at the surface of the protein. Supersecondary structures are usually produced by packing side chains from adjacent secondary structural elements close to each other. Thus, for example, α -helices and β -sheets that are adjacent in the amino acid sequence are also usually (but not always) adjacent in the final, folded protein. Some of the more common motifs are illustrated in Figure 2.8.

Proteins that bind to DNA contain a limited number of motifs. The helix-loop-helix motif is an example found in a number of proteins that function as transcription factors (see p. 455).

**Figure 2.9**

Formation of a disulfide bond by the oxidation of two cysteine residues, producing one cystine residue.

IV. TERTIARY STRUCTURE OF GLOBULAR PROTEINS

The primary structure of a polypeptide chain determines its tertiary structure. “Tertiary” refers both to the folding of domains (the basic units of structure and function, see discussion below), and to the final arrangement of domains in the polypeptide. The structure of globular proteins in aqueous solution is compact, with a high-density (close packing) of the atoms in the core of the molecule. Hydrophobic side chains are buried in the interior, whereas hydrophilic groups are generally found on the surface of the molecule.

A. Domains

Domains are the fundamental functional and three-dimensional structural units of polypeptides. Polypeptide chains that are greater than 200 amino acids in length generally consist of two or more domains. The core of a domain is built from combinations of super-secondary structural elements (motifs). Folding of the peptide chain within a domain usually occurs independently of folding in other domains. Therefore, each domain has the characteristics of a small, compact globular protein that is structurally independent of the other domains in the polypeptide chain.

B. Interactions stabilizing tertiary structure

The unique three-dimensional structure of each polypeptide is determined by its amino acid sequence. Interactions between the amino acid side chains guide the folding of the polypeptide to form a compact structure. The following four types of interactions cooperate in stabilizing the tertiary structures of globular proteins.

- 1. Disulfide bonds:** A disulfide bond is a covalent linkage formed from the sulfhydryl group ($-SH$) of each of two cysteine residues, to produce a cystine residue (Figure 2.9). The two cysteines may be separated from each other by many amino acids in the primary sequence of a polypeptide, or may even be located on two different polypeptide chains; the folding of the polypeptide chain(s) brings the cysteine residues into proximity, and permits covalent bonding of their side chains. A disulfide bond contributes to the stability of the three-dimensional shape of the protein molecule, and prevents it from becoming denatured in the extracellular environment. For example, many disulfide bonds are found in proteins such as immunoglobulins that are secreted by cells.
- 2. Hydrophobic interactions:** Amino acids with nonpolar side chains tend to be located in the interior of the polypeptide molecule, where they associate with other hydrophobic amino acids (Figure 2.10). In contrast, amino acids with polar or charged side chains tend to be located on the surface of the molecule in contact with the polar solvent. [Note: Recall that proteins located in nonpolar (lipid) environments, such as a membrane, exhibit the reverse arrangement (see Figure 1.4, p. 4).] In each case, a segregation of R-groups occurs that is energetically most favorable.
- 3. Hydrogen bonds:** Amino acid side chains containing oxygen- or nitrogen-bound hydrogen, such as in the alcohol groups of serine and threonine, can form hydrogen bonds with electron-rich atoms, such as the oxygen of a carboxyl group or carbonyl group of a peptide bond (Figure 2.11; see also Figure 1.6, p. 4). Formation of hydrogen bonds between polar groups on the surface of proteins and the aqueous solvent enhances the solubility of the protein.
- 4. Ionic interactions:** Negatively charged groups, such as the carboxylate group ($-COO^-$) in the side chain of aspartate or glutamate, can interact with positively charged groups, such as the amino group ($-NH_3^+$) in the side chain of lysine (see Figure 2.11).

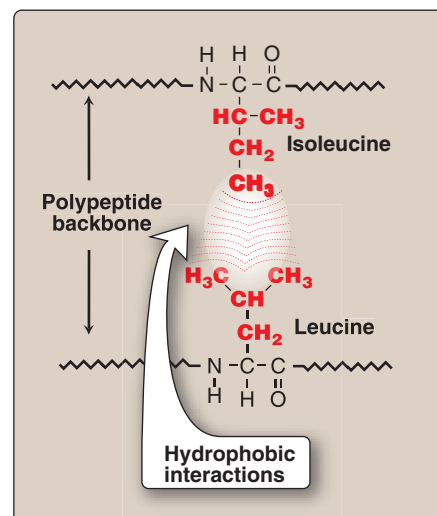


Figure 2.10

Hydrophobic interactions between amino acids with nonpolar side chains.

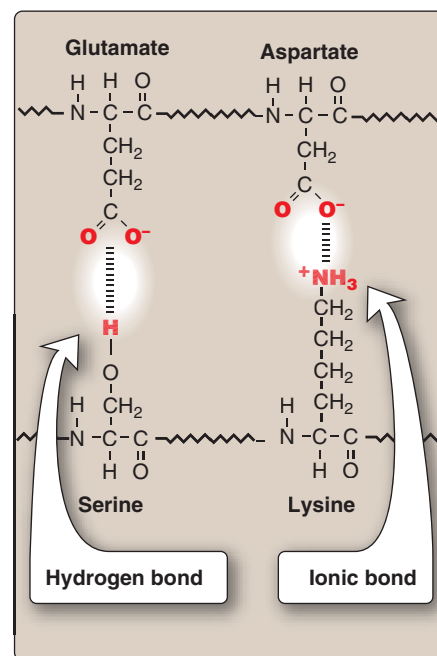


Figure 2.11

Interactions of side chains of amino acids through hydrogen bonds and ionic bonds (salt bridges).

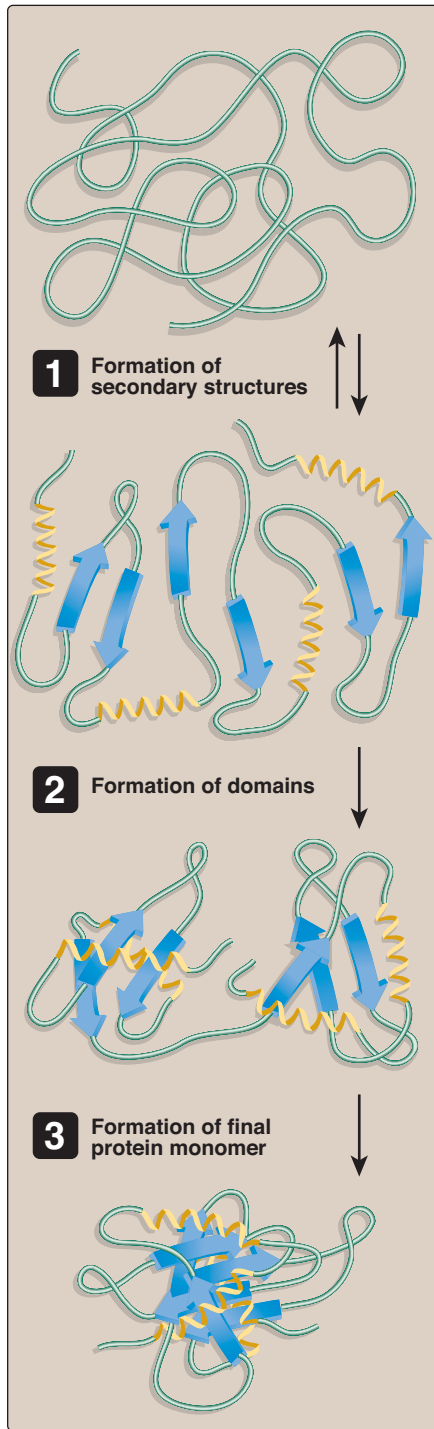


Figure 2.12
Steps in protein folding.

C. Protein folding

Interactions between the side chains of amino acids determine how a long polypeptide chain folds into the intricate three-dimensional shape of the functional protein. Protein folding, which occurs within the cell in seconds to minutes, employs a shortcut through the maze of all folding possibilities. As a peptide folds, its amino acid side chains are attracted and repulsed according to their chemical properties. For example, positively and negatively charged side chains attract each other. Conversely, similarly charged side chains repel each other. In addition, interactions involving hydrogen bonds, hydrophobic interactions, and disulfide bonds all exert an influence on the folding process. This process of trial and error tests many, but not all, possible configurations, seeking a compromise in which attractions outweigh repulsions. This results in a correctly folded protein with a low-energy state (Figure 2.12).

D. Denaturation of proteins

Protein denaturation results in the unfolding and disorganization of the protein's secondary and tertiary structures, which are not accompanied by hydrolysis of peptide bonds. Denaturing agents include heat, organic solvents, mechanical mixing, strong acids or bases, detergents, and ions of heavy metals such as lead and mercury. Denaturation may, under ideal conditions, be reversible, in which case the protein refolds into its original native structure when the denaturing agent is removed. However, most proteins, once denatured, remain permanently disordered. Denatured proteins are often insoluble and, therefore, precipitate from solution.

E. Role of chaperones in protein folding

It is generally accepted that the information needed for correct protein folding is contained in the primary structure of the polypeptide. Given that premise, it is difficult to explain why most proteins when denatured do not resume their native conformations under favorable environmental conditions. One answer to this problem is that a protein begins to fold in stages during its synthesis, rather than waiting for synthesis of the entire chain to be totally completed. This limits competing folding configurations made available by longer stretches of nascent peptide. In addition, a specialized group of proteins, named "chaperones," are required for the proper folding of many species of proteins. The chaperones—also known as "heat shock" proteins—interact with the polypeptide at various stages during the folding process. Some chaperones are important in keeping the protein unfolded until its synthesis is finished, or act as catalysts by increasing the rates of the final stages in the folding process. Others protect proteins as they fold so that their vulnerable, exposed regions do not become tangled in unproductive interactions.

V. QUATERNARY STRUCTURE OF PROTEINS

Many proteins consist of a single polypeptide chain, and are defined as monomeric proteins. However, others may consist of two or more polypeptide chains that may be structurally identical or totally unrelated. The arrangement of these polypeptide subunits is called the quaternary structure of the protein. Subunits are held together by noncovalent interactions (for example, hydrogen bonds, ionic bonds, and hydrophobic

interactions). Subunits may either function independently of each other, or may work cooperatively, as in hemoglobin, in which the binding of oxygen to one subunit of the tetramer increases the affinity of the other subunits for oxygen (see p. 29).

Isoforms are proteins that perform the same function but have different primary structures. They can arise from different genes or from tissue-specific processing of the product of a single gene. If the proteins function as enzymes, they are referred to as isozymes (see p. 65).

VI. PROTEIN MISFOLDING

Protein folding is a complex, trial-and-error process that can sometimes result in improperly folded molecules. These misfolded proteins are usually tagged and degraded within the cell (see p. 444). However, this quality control system is not perfect, and intracellular or extracellular aggregates of misfolded proteins can accumulate, particularly as individuals age. Deposits of these misfolded proteins are associated with a number of diseases.

A. Amyloid disease

Misfolding of proteins may occur spontaneously, or be caused by a mutation in a particular gene, which then produces an altered protein. In addition, some apparently normal proteins can, after abnormal proteolytic cleavage, take on a unique conformational state that leads to the formation of long, fibrillar protein assemblies consisting of β -pleated sheets. Accumulation of these insoluble, spontaneously aggregating proteins, called amyloids, has been implicated in many degenerative diseases—particularly in the age-related neurodegenerative disorder, Alzheimer disease. The dominant component of the amyloid plaque that accumulates in Alzheimer disease is amyloid β ($A\beta$), a peptide containing 40–42 amino acid residues. X-ray crystallography and infrared spectroscopy demonstrate a characteristic β -pleated sheet conformation in nonbranching fibrils. This peptide, when aggregated in a β -pleated sheet configuration, is neurotoxic, and is the central pathogenic event leading to the cognitive impairment characteristic of the disease. The $A\beta$ that is deposited in the brain in Alzheimer disease is derived by proteolytic cleavages from the larger amyloid precursor protein—a single transmembrane protein expressed on the cell surface in the brain and other tissues (Figure 2.13). The $A\beta$ peptides aggregate, generating the amyloid that is found in the brain parenchyma and around blood vessels. Most cases of Alzheimer disease are not genetically based, although at least 5–10% of cases are familial. A second biologic factor involved in the development of Alzheimer disease is the accumulation of neurofibrillary tangles inside neurons. A key component of these tangled fibers is an abnormal form of the tau (τ) protein, which in its healthy version helps in the assembly of the microtubular structure. The defective τ , however, appears to block the actions of its normal counterpart.

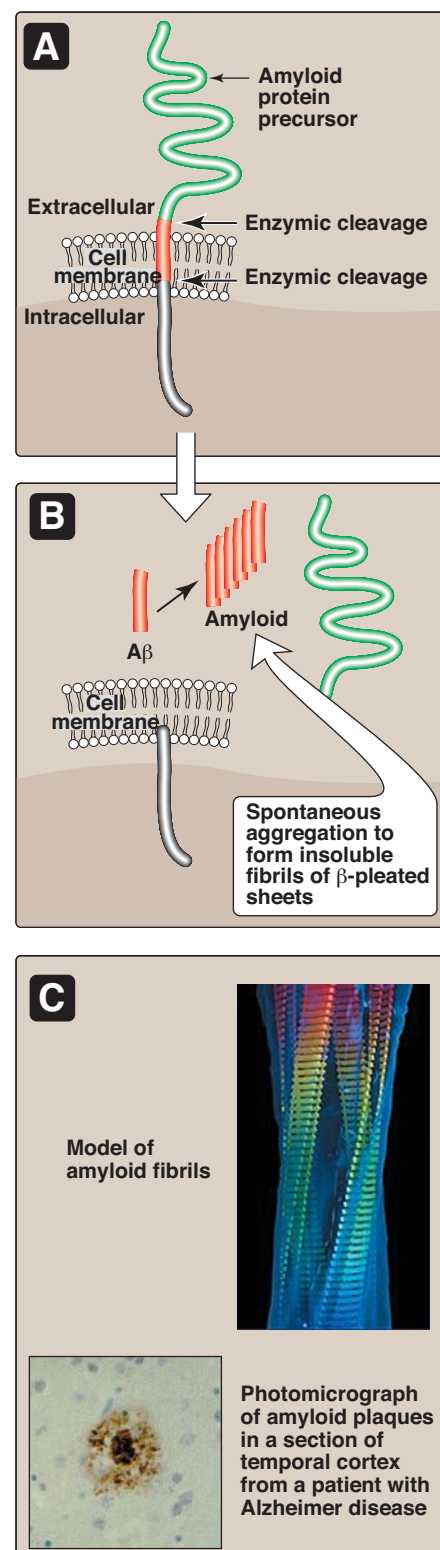


Figure 2.13
Formation of amyloid plaques found in Alzheimer disease.

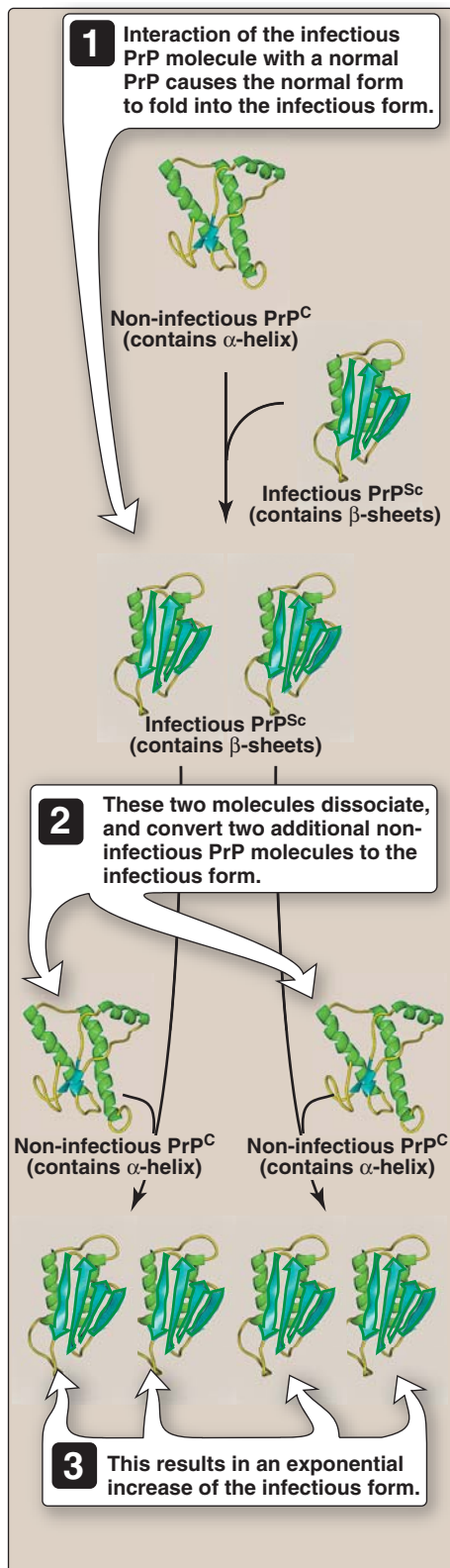


Figure 2.14

One proposed mechanism for multiplication of infectious prion agents.

B. Prion disease

The prion protein (PrP) has been strongly implicated as the causative agent of transmissible spongiform encephalopathies (TSEs), including Creutzfeldt-Jakob disease in humans, scrapie in sheep, and bovine spongiform encephalopathy in cattle (popularly called “mad cow disease”).¹ After an extensive series of purification procedures, scientists were astonished to find that the infectivity of the agent causing scrapie in sheep was associated with a single protein species that was not complexed with detectable nucleic acid. This infectious protein is designated PrP^{Sc} (Sc = scrapie). It is highly resistant to proteolytic degradation, and tends to form insoluble aggregates of fibrils, similar to the amyloid found in some other diseases of the brain. A noninfectious form of PrP^C (C = cellular), encoded by the same gene as the infectious agent, is present in normal mammalian brains on the surface of neurons and glial cells. Thus, PrP^C is a host protein. No primary structure differences or alternate posttranslational modifications have been found between the normal and the infectious forms of the protein. The key to becoming infectious apparently lies in changes in the three-dimensional conformation of PrP^C. It has been observed that a number of α -helices present in noninfectious PrP^C are replaced by β -sheets in the infectious form (Figure 2.14). It is presumably this conformational difference that confers relative resistance to proteolytic degradation of infectious prions, and permits them to be distinguished from the normal PrP^C in infected tissue. The infective agent is thus an altered version of a normal protein, which acts as a “template” for converting the normal protein to the pathogenic conformation. The TSEs are invariably fatal, and no treatment is currently available that can alter this outcome.

VII. CHAPTER SUMMARY

Central to understanding protein structure is the concept of the **native conformation** (Figure 2.15), which is the functional, fully-folded protein structure (for example, an active enzyme or structural protein). The unique three-dimensional structure of the native conformation is determined by its **primary structure**, that is, its amino acid sequence. Interactions between the amino acid side chains guide the folding of the polypeptide chain to form **secondary**, **tertiary**, and (sometimes) **quaternary** structures, which cooperate in stabilizing the native conformation of the protein. In addition, a specialized group of proteins named “**chaperones**” is required for the proper folding of many species of proteins. **Protein denaturation** results in the unfolding and disorganization of the protein’s structure, which are not accompanied by hydrolysis of peptide bonds. Denaturation may be reversible or, more commonly, irreversible. Disease can occur when an apparently normal protein assumes a conformation that is cytotoxic, as in the case of Alzheimer disease and the **transmissible spongiform encephalopathies (TSEs)**, including **Creutzfeldt-Jakob disease**. In **Alzheimer disease**, normal proteins, after abnormal chemical processing, take on a unique conformational state that leads to the formation of neurotoxic **amyloid protein** assemblies consisting of β -pleated sheets. In TSEs, the infective agent is an altered version of a normal **prion protein** that acts as a “template” for converting normal protein to the pathogenic conformation.



¹See Chapter 31 in *Lippincott’s Illustrated Reviews: Microbiology* for a more detailed discussion of prions.

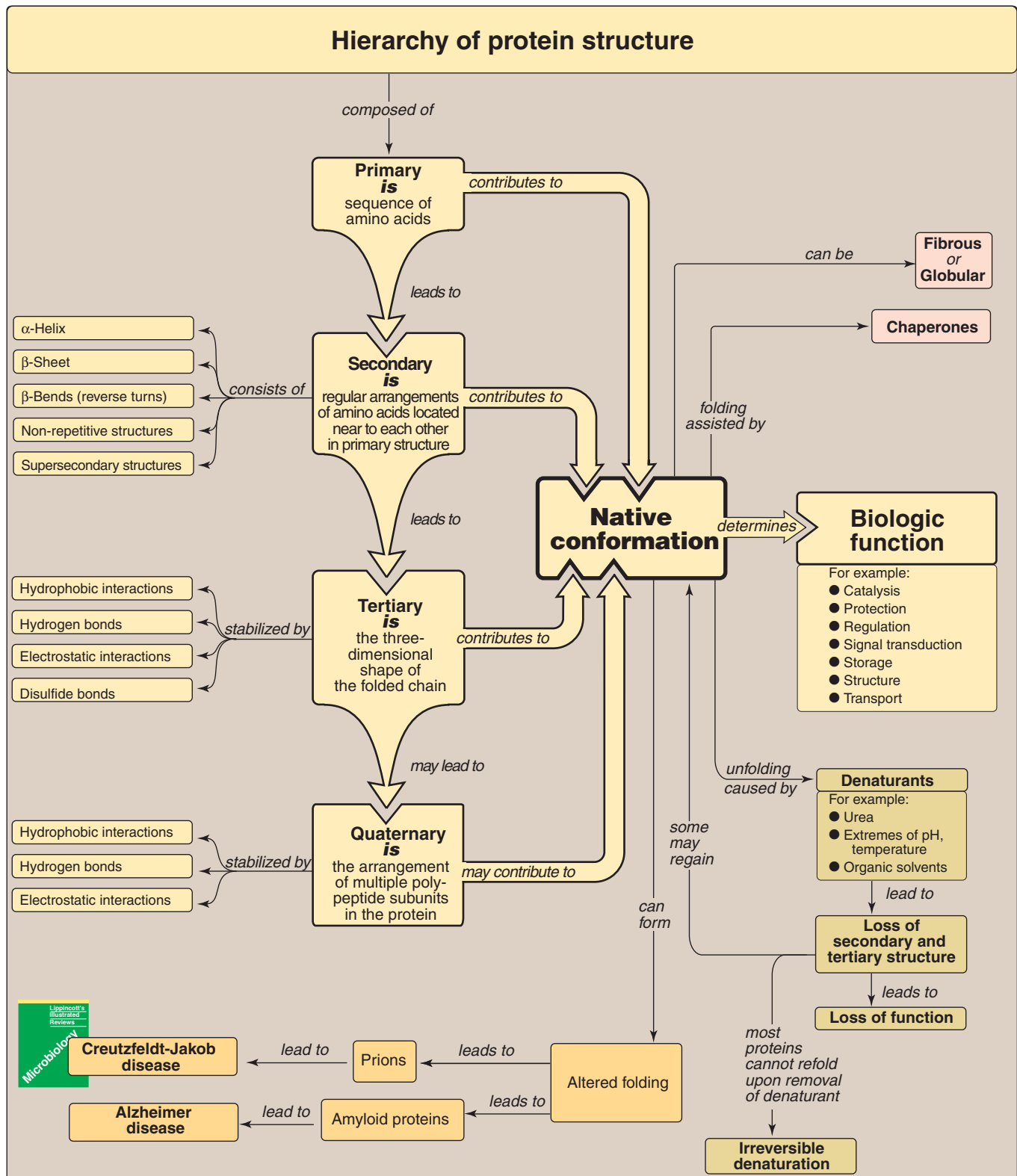


Figure 2.15
Key concept map for protein structure.

Study Questions

Choose the ONE correct answer.

2.1 A peptide bond:

- A. has a partial double-bond character.
- B. is ionized at physiologic pH.
- C. is cleaved by agents that denature proteins, such as organic solvents and high concentrations of urea.
- D. is stable to heating in strong acids.
- E. occurs most commonly in the cis configuration.

Correct answer = A. The peptide bond has a partial double-bond character. Unlike its components—the α -amino and α -carboxyl groups—the $-\text{NH}$ and $-\text{C}=\text{O}$ of the peptide bond do not accept or give off protons. The peptide bond is not cleaved by organic solvents or urea, but is labile to strong acids. It is usually in the trans configuration.

2.2 Which one of the following statements is correct?

- A. The α -helix can be composed of more than one polypeptide chain.
- B. β -Sheets exist only in the antiparallel form.
- C. β -Bends often contain proline.
- D. Domains are a type of secondary structure.
- E. The α -helix is stabilized primarily by ionic interactions between the side chains of amino acids.

Correct answer = C. β -Bends often contain proline, which provides a kink. The α -helix differs from the β -sheet in that it always involves the coiling of a single polypeptide chain. The β -sheet occurs in both parallel and antiparallel forms. Domains are elements of tertiary structure. The α -helix is stabilized primarily by hydrogen bonds between the $-\text{C}=\text{O}$ and $-\text{NH}-$ groups of peptide bonds.

2.3 Which one of the following statements about protein structure is correct?

- A. Proteins consisting of one polypeptide can have quaternary structure.
- B. The formation of a disulfide bond in a protein requires that the two participating cysteine residues be adjacent to each other in the primary sequence of the protein.
- C. The stability of quaternary structure in proteins is mainly a result of covalent bonds among the subunits.
- D. The denaturation of proteins always leads to irreversible loss of secondary and tertiary structure.
- E. The information required for the correct folding of a protein is contained in the specific sequence of amino acids along the polypeptide chain.

Correct answer = E. The correct folding of a protein is guided by specific interactions between the side chains of the amino acid residues of a polypeptide chain. The two cysteine residues that react to form the disulfide bond may be a great distance apart in the primary structure (or on separate polypeptides), but are brought into close proximity by the three-dimensional folding of the polypeptide chain. Denaturation may either be reversible or irreversible. Quaternary structure requires more than one polypeptide chain. These chains associate through noncovalent interactions.

2.4 An 80-year-old man presented with impairment of higher intellectual function and alterations in mood and behavior. His family reported progressive disorientation and memory loss over the last 6 months. There is no family history of dementia. The patient was tentatively diagnosed with Alzheimer disease. Which one of the following best describes the disease?

- A. It is associated with β -amyloid—an abnormal protein with an altered amino acid sequence.
- B. It results from accumulation of denatured proteins that have random conformations.
- C. It is associated with the accumulation of amyloid precursor protein.
- D. It is associated with the deposition of neurotoxic amyloid peptide aggregates.
- E. It is an environmentally produced disease not influenced by the genetics of the individual.
- F. It is caused by the infectious form of a host-cell protein.

Correct answer = D. Alzheimer disease is associated with long, fibrillar protein assemblies consisting of β -pleated sheets found in the brain and elsewhere. The disease is associated with abnormal processing of a normal protein. The accumulated altered protein occurs in a β -pleated sheet configuration that is neurotoxic. The $\text{A}\beta$ amyloid that is deposited in the brain in Alzheimer disease is derived by proteolytic cleavages from the larger amyloid precursor protein—a single transmembrane protein expressed on the cell surface in the brain and other tissues. Most cases of Alzheimer disease are sporadic, although at least 5–10% of cases are familial. Prion diseases, such as Creutzfeldt-Jakob, are caused by the infectious form (PrP^{Sc}) of a host-cell protein (PrP^{C}).