Chapter 4

Amino Acids

To hold, as 'twere, the mirror up to nature.

WILLIAM SHAKESPEARE, *Hamlet*

All objects have mirror images. Like many biomolecules, amino acids exist in mirror-image forms (stereoisomers) that are not superimposable. Only the L*-isomers of amino acids commonly occur in nature.* (The Mirror of Venus *(1898), Sir Edward Burne-Jones/Museu Calouste Gulbenkian Lisbon/The Bridgeman Art Library)*

OUTLINE

Proteins are the indispensable agents of biological function, and **amino acids** are the building blocks of proteins. The stunning diversity of the thousands of proteins found in nature arises from the intrinsic properties of only 20 commonly occurring amino acids. These features include (1) the capacity to polymerize, (2) novel acid–base properties, (3) varied structure and chemical functionality in the amino acid side chains, and (4) chirality. This chapter describes each of these properties, laying a foundation for discussions of protein structure (Chapters 5 and 6), enzyme function (Chapters 14–16), and many other subjects in later chapters.

4.1 • Amino Acids: Building Blocks of Proteins

Structure of a Typical Amino Acid

The structure of a single typical amino acid is shown in Figure 4.1. Central to this structure is the tetrahedral alpha (α) carbon (C_α), which is covalently linked to both the amino group and the carboxyl group. Also bonded to this α-carbon is a hydrogen and a variable side chain. It is the side chain, the so-called R group, that gives each amino acid its identity. The detailed acid–base properties of amino acids are discussed in the following sections. It is sufficient for now to realize that, in neutral solution (pH 7), the carboxyl group exists as $—COO^-$ and the amino group as $—NH_3^+$. Because the resulting amino acid contains one positive and one negative charge, it is a neutral molecule called a **zwitterion.** Amino acids are also *chiral* molecules. With four different groups attached to it, the α-carbon is said to be *asymmetric.* The two possible configurations for the α-carbon constitute nonidentical mirror image isomers or *enantiomers.* Details of amino acid stereochemistry are discussed in Section 4.4.

Amino Acids Can Join via Peptide Bonds

The crucial feature of amino acids that allows them to polymerize to form peptides and proteins is the existence of their two identifying chemical groups: the amino ($—NH_3^+$) and carboxyl ($—COO^-$) groups, as shown in Figure 4.2. The amino and carboxyl groups of amino acids can react in a head-to-tail fashion, eliminating a water molecule and forming a covalent amide linkage, which, in the case of peptides and proteins, is typically referred to as a **peptide bond.** The equilibrium for this reaction in aqueous solution favors peptide bond hydrolysis. For this reason, biological systems as well as peptide chemists in the laboratory must carry out peptide bond formation in an indirect manner or with energy input.

Iteration of the reaction shown in Figure 4.2 produces **polypeptides** and **proteins.** The remarkable properties of proteins, which we shall discover and come to appreciate in later chapters, all depend in one way or another on the unique properties and chemical diversity of the 20 common amino acids found in proteins.

Common Amino Acids

The structures and abbreviations for the 20 amino acids commonly found in proteins are shown in Figure 4.3. All the amino acids except proline have both free α-amino and free α-carboxyl groups (Figure 4.1). There are several ways to classify the common amino acids. The most useful of these classifications is based on the polarity of the side chains. Thus, the structures shown in Figure 4.3 are grouped into the following categories: (1) nonpolar or hydrophobic

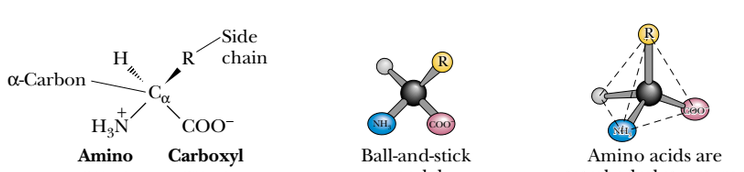


FIGURE 4.1 • Anatomy of an amino acid. Except for proline and its derivatives, all of the amino acids commonly found in proteins possess this type of structure.

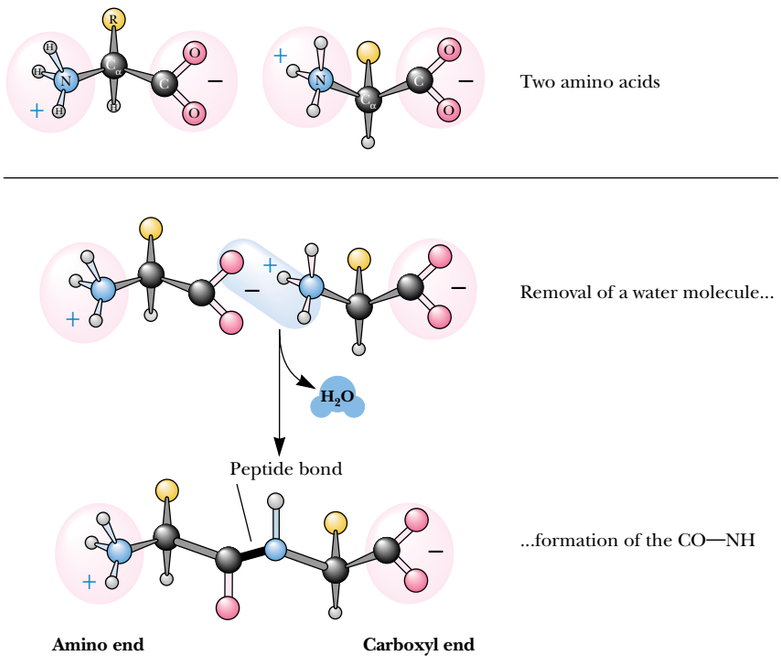


FIGURE 4.2 • The α-COOH and α-NH_3^+ groups of two amino acids can react with the resulting loss of a water molecule to form a covalent amide bond. *(Irving Geis.)*

amino acids, (2) neutral (uncharged) but polar amino acids, (3) acidic amino acids (which have a net negative charge at pH 7.0), and (4) basic amino acids (which have a net positive charge at neutral pH). In later chapters, the importance of this classification system for predicting protein properties becomes clear. Also shown in Figure 4.3 are the three-letter and one-letter codes used to represent the amino acids. These codes are useful when displaying and comparing the sequences of proteins in shorthand form. (Note that several of the one-letter abbreviations are phonetic in origin: arginine = "Rginine" = R, phenylalanine = "Fenylalanine" = F, aspartic acid = "asparDic" = D.)

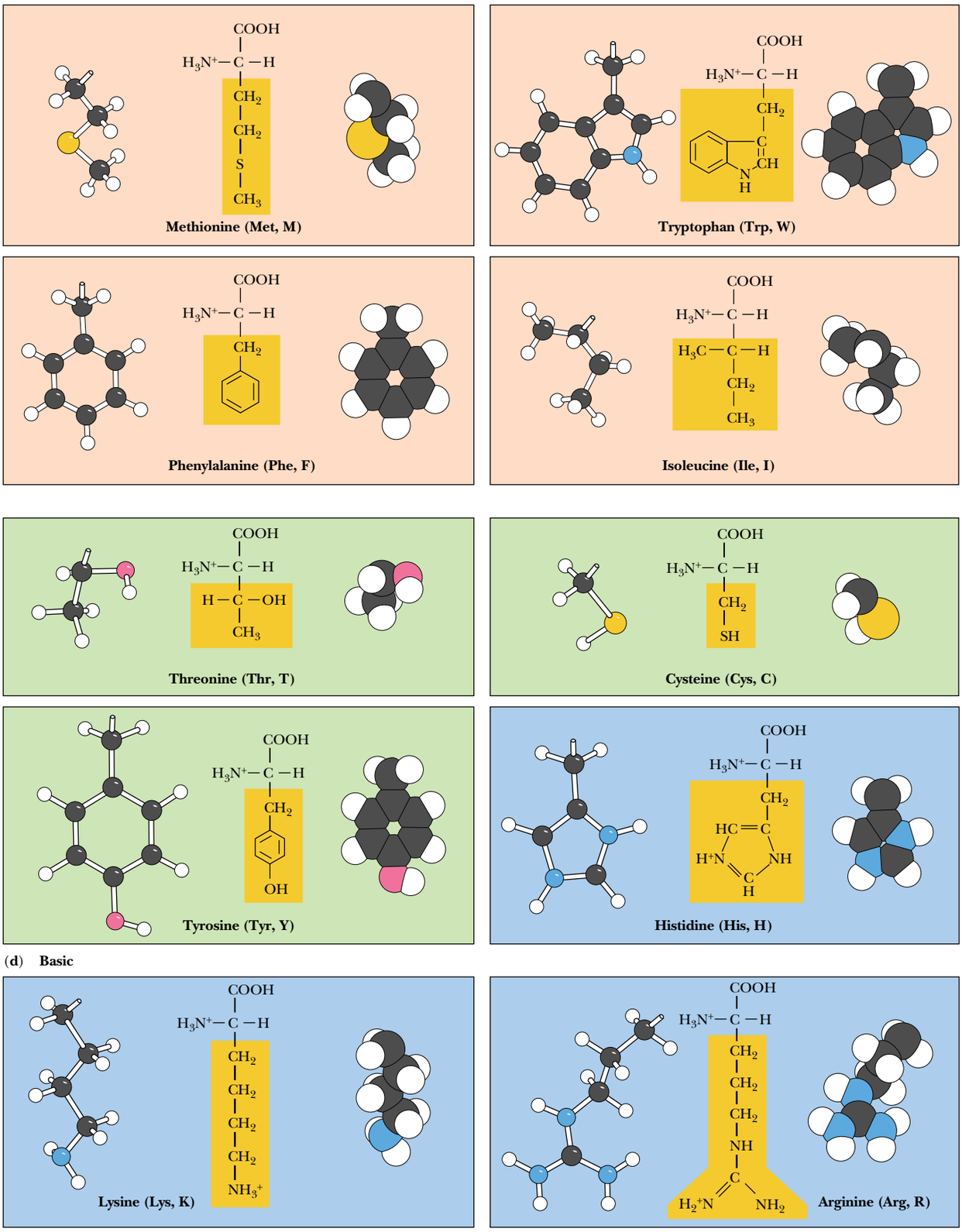
Nonpolar Amino Acids

The nonpolar amino acids (Figure 4.3a) include all those with alkyl chain R groups (alanine, valine, leucine, and isoleucine), as well as proline (with its unusual cyclic structure), methionine (one of the two sulfur-containing amino acids), and two aromatic amino acids, phenylalanine and tryptophan. Tryptophan is sometimes considered a borderline member of this group because it can interact favorably with water via the N–H moiety of the indole ring. Proline, strictly speaking, is not an amino acid but rather an α-imino acid.

Polar, Uncharged Amino Acids

The polar, uncharged amino acids (Figure 4.3b) except for glycine contain R groups that can form hydrogen bonds with water. Thus, these amino acids are usually more soluble in water than the nonpolar amino acids. Several exceptions should be noted. Tyrosine displays the lowest solubility in water of the 20 common amino acids (0.453 g/L at 25°C). Also, proline is very soluble in water, and alanine and valine are about as soluble as arginine and serine. The amide groups of asparagine and glutamine; the hydroxyl groups of tyrosine, threonine, and serine; and the sulfhydryl group of cysteine are all good hydrogen

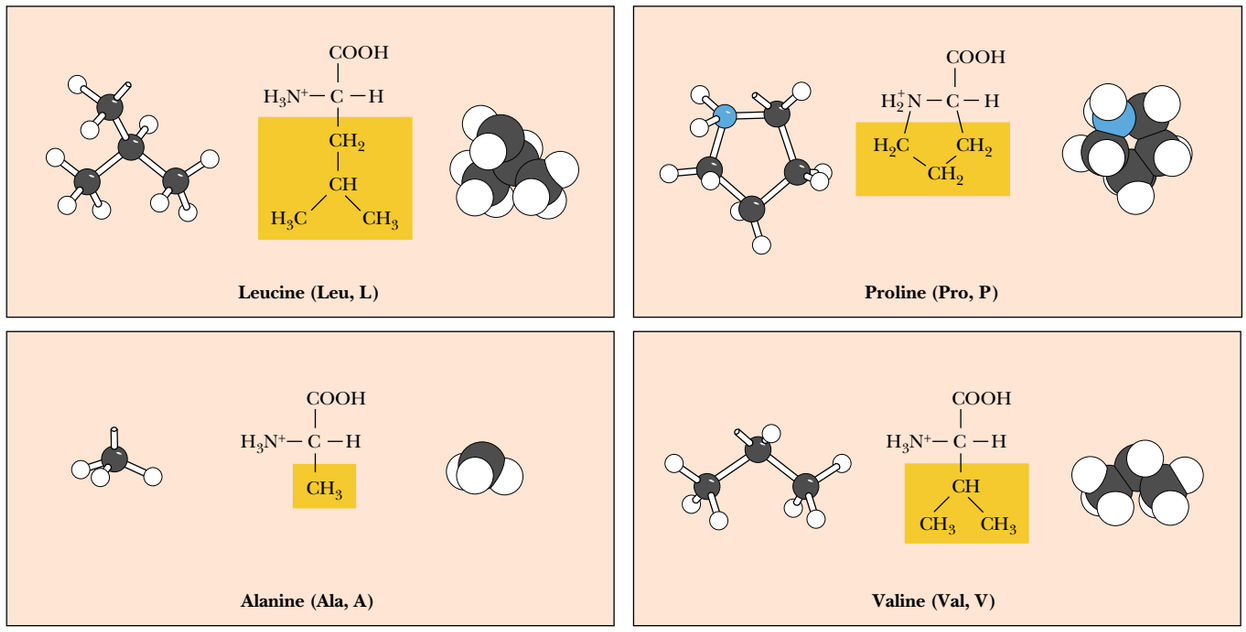
(Text continues on page 86.)



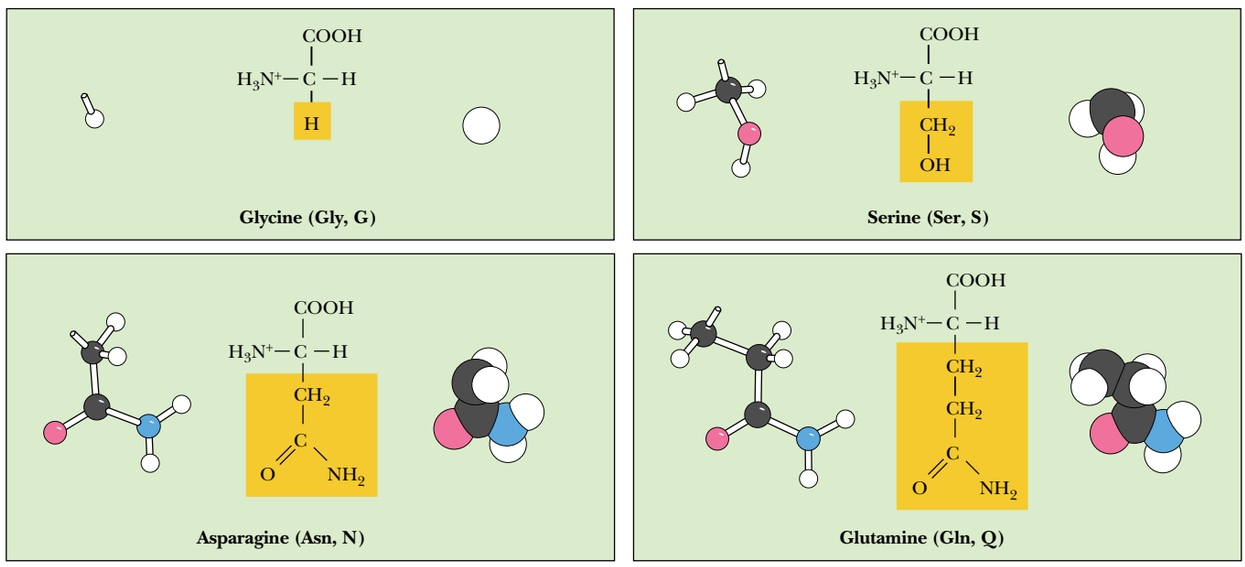
(**d**) **Basic**

FIGURE 4.3 • The 20 amino acids that are the building blocks of most proteins can be classified as (a) nonpolar (hydrophobic), (b) polar, neutral, (c) acidic, or (d) basic.

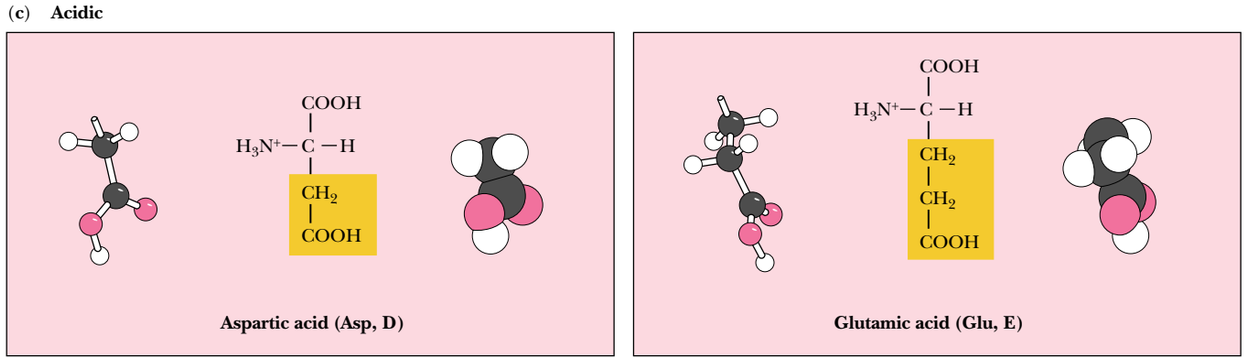
(a) **Nonpolar (hydrophobic)**



(b) **Polar, uncharged**



(c) **Acidic**



Also shown are the one-letter and three-letter codes used to denote amino acids. For each amino acid, the ball-and-stick *(left)* and space-filling *(right)* models show only the side chain. *(Irving Geis)*

bond–forming moieties. Glycine, the simplest amino acid, has only a single hydrogen for an R group, and this hydrogen is not a good hydrogen bond former. Glycine's solubility properties are mainly influenced by its polar amino and carboxyl groups, and thus glycine is best considered a member of the polar, uncharged group. It should be noted that tyrosine has significant nonpolar characteristics due to its aromatic ring and could arguably be placed in the nonpolar group (Figure 4.3a). However, with a pK_a of 10.1, tyrosine's phenolic hydroxyl is a charged, polar entity at high pH.

Acidic Amino Acids

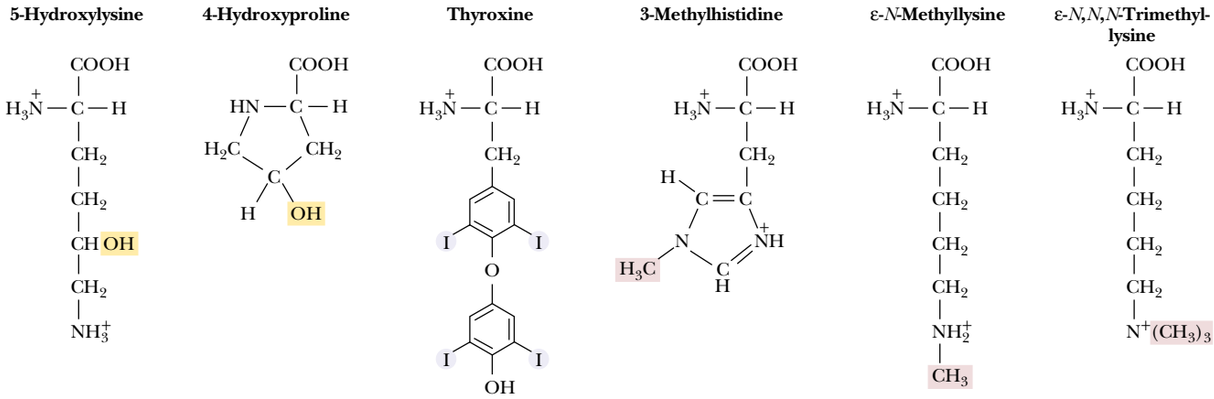
There are two acidic amino acids—aspartic acid and glutamic acid—whose R groups contain a carboxyl group (Figure 4.3c). These side chain carboxyl groups are weaker acids than the α-COOH group, but are sufficiently acidic to exist as $-COO^-$ at neutral pH. Aspartic acid and glutamic acid thus have a net negative charge at pH 7. These negatively charged amino acids play several important roles in proteins. Many proteins that bind metal ions for structural or functional purposes possess metal binding sites containing one or more aspartate and glutamate side chains. Carboxyl groups may also act as nucleophiles in certain enzyme reactions and may participate in a variety of electrostatic bonding interactions. The acid–base chemistry of such groups is considered in detail in Section 4.2.

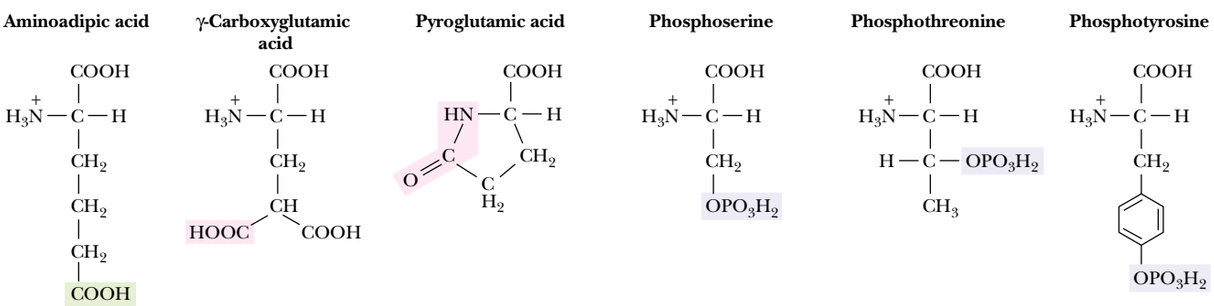
Basic Amino Acids

Three of the common amino acids have side chains with net positive charges at neutral pH: histidine, arginine, and lysine (Figure 4.3d). The ionized group of histidine is an imidazolium, that of arginine is a guanidinium, and lysine contains a protonated alkyl amino group. The side chains of the latter two amino acids are fully protonated at pH 7, but histidine, with a side chain pK_a of 6.0, is only 10% protonated at pH 7. With a pK_a near neutrality, histidine side chains play important roles as proton donors and acceptors in many enzyme reactions. Histidine-containing peptides are important biological buffers, as discussed in Chapter 2. Arginine and lysine side chains, which are protonated under physiological conditions, participate in electrostatic interactions in proteins.

Uncommon Amino Acids

Several amino acids occur only rarely in proteins (Figure 4.4). These include **hydroxylysine** and **hydroxyproline,** which are found mainly in the collagen and gelatin proteins, and **thyroxine** and **3,3′,5-triiodothyronine,** iodinated amino acids that are found only in thyroglobulin, a protein produced by the thyroid gland. (Thyroxine and 3,3′,5-triiodothyronine are produced by iodination of tyrosine residues in thyroglobulin in the thyroid gland. Degradation of thyroglobulin releases these two iodinated amino acids, which act as hormones to regulate growth and development.) Certain muscle proteins contain methylated amino acids, including **methylhistidine,** **ϵ-*N*-methyllysine,** and **ϵ-*N*,*N*,*N*-trimethyllysine** (Figure 4.4). **γ-Carboxyglutamic acid** is found in several proteins involved in blood clotting, and **pyroglutamic acid** is found in a unique light-driven proton-pumping protein called bacteriorhodopsin, which is discussed elsewhere in this book. Certain proteins involved in cell growth and regulation are reversibly phosphorylated on the —OH groups of serine, threonine, and tyrosine residues. **Aminoadipic acid** is found in proteins isolated from corn. Finally, ***N*-methylarginine** and ***N*-acetyllysine** are found in histone proteins associated with chromosomes.





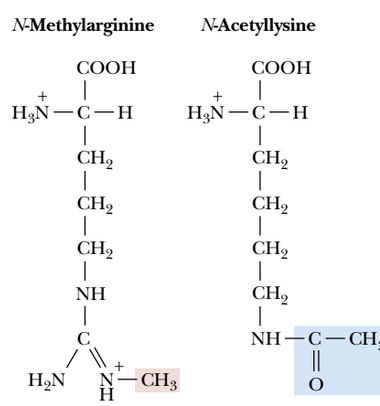


FIGURE 4.4 • The structures of several amino acids that are less common but nevertheless found in certain proteins. Hydroxylysine and hydroxyproline are found in connective-tissue proteins, pyroglutamic acid is found in bacteriorhodopsin (a protein in *Halobacterium halobium*), and aminoadipic acid is found in proteins isolated from corn.

Amino Acids Not Found in Proteins

Certain amino acids and their derivatives, although not found in proteins, nonetheless are biochemically important. A few of the more notable examples are shown in Figure 4.5. **γ-Aminobutyric acid,** or **GABA,** is produced by the decarboxylation of glutamic acid and is a potent neurotransmitter. **Histamine,** which is synthesized by decarboxylation of histidine, and **serotonin,** which is derived from tryptophan, similarly function as neurotransmitters and regulators. **β-Alanine** is found in nature in the peptides carnosine and anserine and is a component of pantothenic acid (a vitamin), which is a part of coenzyme A. **Epinephrine** (also known as **adrenaline**), derived from tyrosine, is an important hormone. **Penicillamine** is a constituent of the penicillin antibiotics. **Ornithine, betaine, homocysteine,** and **homoserine** are important metabolic intermediates. **Citrulline** is the immediate precursor of arginine.

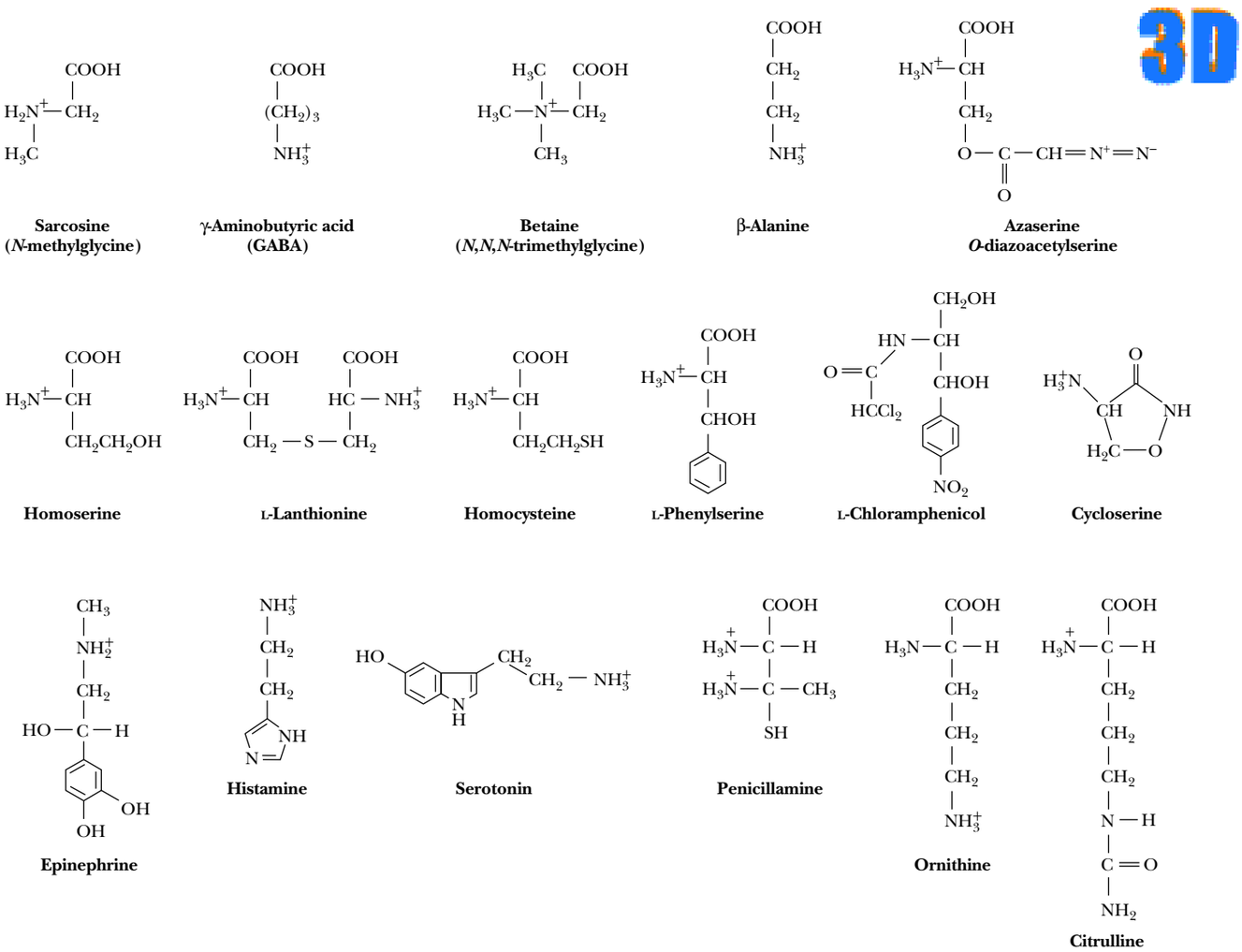


FIGURE 4.5 • The structures of some amino acids that are not normally found in proteins but that perform other important biological functions. Epinephrine, histamine, and serotonin, although not amino acids, are derived from and closely related to amino acids.

4.2 • Acid–Base Chemistry of Amino Acids

Amino Acids Are Weak Polyprotic Acids

From a chemical point of view, the common amino acids are all weak polyprotic acids. The ionizable groups are not strongly dissociating ones, and the degree of dissociation thus depends on the pH of the medium. All the amino acids contain at least two dissociable hydrogens.

Consider the acid–base behavior of glycine, the simplest amino acid. At low pH, both the amino and carboxyl groups are protonated and the molecule has a net positive charge. If the counterion in solution is a chloride ion, this form is referred to as **glycine hydrochloride.** If the pH is increased, the carboxyl group is the first to dissociate, yielding the neutral zwitterionic species Gly^0 (Figure 4.6). Further increase in pH eventually results in dissociation of the amino group to yield the negatively charged **glycinate.** If we denote these

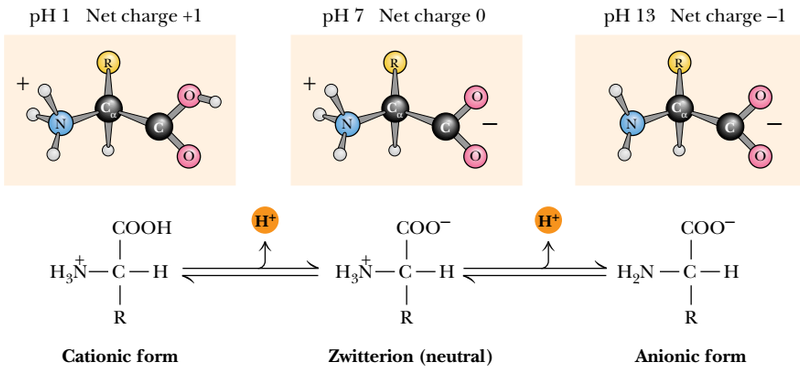


FIGURE 4.6 • The ionic forms of the amino acids, shown without consideration of any ionizations on the side chain. The cationic form is the low pH form, and the titration of the cationic species with base yields the zwitterion and finally the anionic form. (*Irving Geis*)

three forms as Gly^+, Gly^0, and Gly^-, we can write the first dissociation of Gly^+ as

$$Gly^+ + H_2O \rightleftharpoons Gly^0 + H_3O^+$$

and the dissociation constant K_1 as

$$K_1 = \frac{[Gly^0][H_3O^+]}{[Gly^+]}$$

Values for K_1 for the common amino acids are typically 0.4 to 1.0×10^{-2} *M*, so that typical values of pK_1 center on values of 2.0 to 2.4 (see Table 4.1). In a similar manner, we can write the second dissociation reaction as

$$Gly^0 + H_2O \rightleftharpoons Gly^- + H_3O^+$$

Table 4.1

pK_a Values of Common Amino Acids

Amino Acid	α-COOH pK_a	α-NH_3^+ pK_a	R group pK_a
Alanine	2.4	9.7	
Arginine	2.2	9.0	12.5
Asparagine	2.0	8.8	
Aspartic acid	2.1	9.8	3.9
Cysteine	1.7	10.8	8.3
Glutamic acid	2.2	9.7	4.3
Glutamine	2.2	9.1	
Glycine	2.3	9.6	
Histidine	1.8	9.2	6.0
Isoleucine	2.4	9.7	
Leucine	2.4	9.6	
Lysine	2.2	9.0	10.5
Methionine	2.3	9.2	
Phenylalanine	1.8	9.1	
Proline	2.1	10.6	
Serine	2.2	9.2	~13
Threonine	2.6	10.4	~13
Tryptophan	2.4	9.4	
Tyrosine	2.2	9.1	10.1
Valine	2.3	9.6	

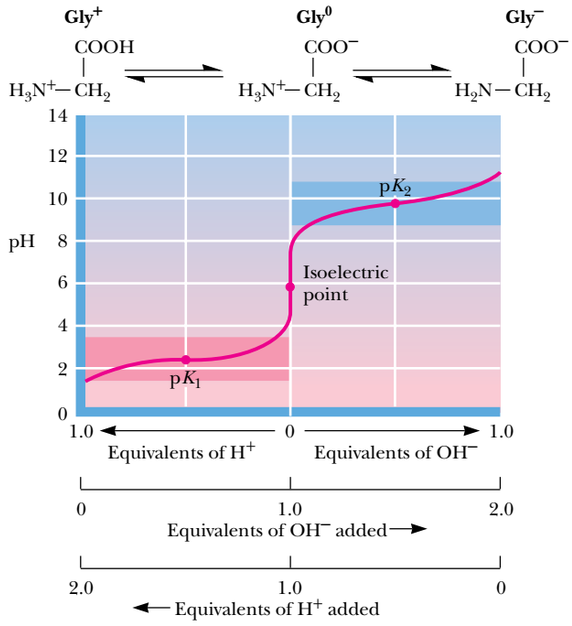


FIGURE 4.7 • Titration of glycine, a simple amino acid. The isoelectric point, pI, the pH where the molecule has a net charge of 0, is defined as $(pK_1 + pK_2)/2$.

and the dissociation constant K_2 as

$$K_2 = \frac{[\text{Gly}^-][\text{H}_3\text{O}^+]}{[\text{Gly}^0]}$$

Typical values for pK_2 are in the range of 9.0 to 9.8. At physiological pH, the α-carboxyl group of a simple amino acid (with no ionizable side chains) is completely dissociated, whereas the α-amino group has not really begun its dissociation. The titration curve for such an amino acid is shown in Figure 4.7.

EXAMPLE

What is the pH of a glycine solution in which the α-NH_3^+ group is one-third dissociated?

SOLUTION

The appropriate Henderson–Hasselbalch equation is

$$\text{pH} = pK_a + \log_{10}\frac{[\text{Gly}^-]}{[\text{Gly}^0]}$$

If the α-amino group is one-third dissociated, there is one part Gly^- for every two parts Gly^0. The important pK_a is the pK_a for the amino group. The glycine α-amino group has a pK_a of 9.6. The result is

$$\text{pH} = 9.6 + \log_{10}(1/2)$$
$$\text{pH} = 9.3$$

Note that the dissociation constants of both the α-carboxyl and α-amino groups are affected by the presence of the other group. The adjacent α-amino group makes the α-COOH group more acidic (that is, it lowers the pK_a) so

that it gives up a proton more readily than simple alkyl carboxylic acids. Thus, the pK_1 of 2.0 to 2.1 for α-carboxyl groups of amino acids is substantially lower than that of acetic acid ($pK_a = 4.76$), for example. What is the chemical basis for the low pK_a of the α-COOH group of amino acids? The α-NH_3^+ (ammonium) group is strongly electron-withdrawing, and the positive charge of the amino group exerts a strong field effect and stabilizes the carboxylate anion. (The effect of the α-COO^- group on the pK_a of the α-NH_3^+ group is the basis for Problem 4 at the end of this chapter.)

Ionization of Side Chains

As we have seen, the side chains of several of the amino acids also contain dissociable groups. Thus, aspartic and glutamic acids contain an additional carboxyl function, and lysine possesses an aliphatic amino function. Histidine contains an ionizable imidazolium proton, and arginine carries a guanidinium function. Typical pK_a values of these groups are shown in Table 4.1. The β-carboxyl group of aspartic acid and the γ-carboxyl side chain of glutamic acid exhibit pK_a values intermediate to the α-COOH on the one hand and typical aliphatic carboxyl groups on the other hand. In a similar fashion, the ϵ-amino group of lysine exhibits a pK_a that is higher than the α-amino group but similar to that for a typical aliphatic amino group. These intermediate values for side-chain pK_a values reflect the slightly diminished effect of the α-carbon dissociable groups that lie several carbons removed from the side-chain functional groups. Figure 4.8 shows typical titration curves for glutamic acid and lysine, along with the ionic species that predominate at various points in the

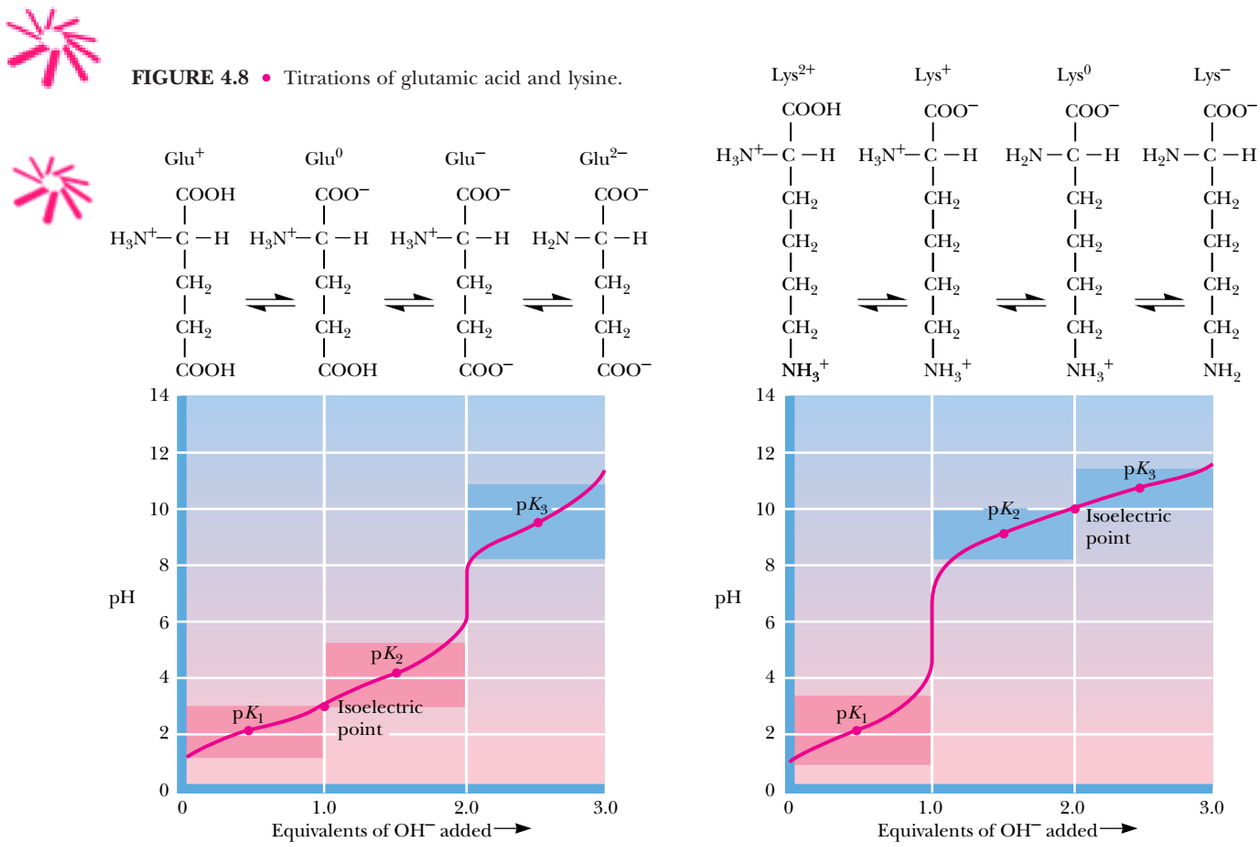


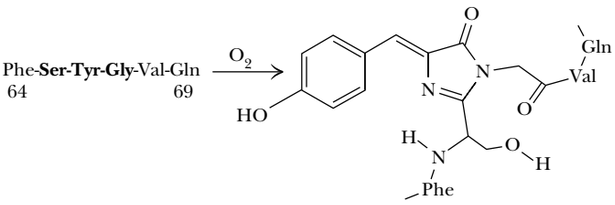
FIGURE 4.8 • Titrations of glutamic acid and lysine.

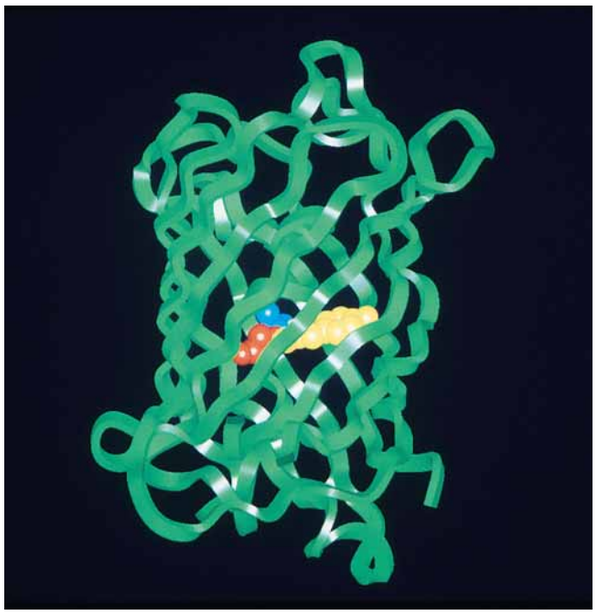
CRITICAL DEVELOPMENTS IN BIOCHEMISTRY

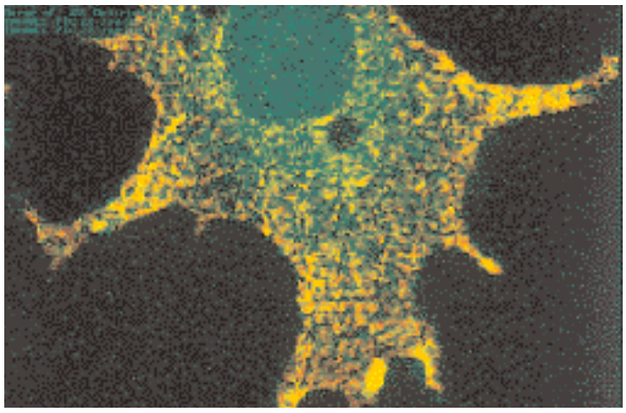
Green Fluorescent Protein—The "Light Fantastic" from Jellyfish to Gene Expression

Aquorea victoria, a species of jellyfish found in the northwest Pacific Ocean, contains a **green fluorescent protein (GFP)** that works together with another protein, **aequorin,** to provide a defense mechanism for the jellyfish. When the jellyfish is attacked or shaken, aequorin produces a blue light. This light energy is captured by GFP, which then emits a bright green flash that presumably blinds or startles the attacker. Remarkably, the fluorescence of GFP occurs without the assistance of a **prosthetic group**—a "helper molecule" that would mediate GFP's fluorescence. Instead, the light-transducing capability of GFP is the result of a reaction between three amino acids in the protein itself. As shown below, adjacent **serine, tyrosine,** and **glycine** in the sequence of the protein react to form the pigment complex—termed a **chromophore.** No enzymes are required; the reaction is autocatalytic.

Because the light-transducing talents of GFP depend only on the protein itself (upper photo, chromophore highlighted), GFP has quickly become a darling of genetic engineering laboratories. The promoter of any gene whose cellular expression is of interest can be fused to the DNA sequence coding for GFP. Telltale green fluorescence tells the researcher when this fused gene has been expressed (see lower photo and also Chapter 13).







Autocatalytic oxidation of GFP amino acids leads to the chromophore shown on the left. The green fluorescence requires further interactions of the chromophore with other parts of the protein.

Boxer, S.G., 1997. Another green revolution. *Nature* **383:**484–485.

titration. The only other side-chain groups that exhibit any significant degree of dissociation are the *para*-OH group of tyrosine and the —SH group of cysteine. The pK_a of the cysteine sulfhydryl is 8.32, so that it is about 12% dissociated at pH 7. The tyrosine *para*-OH group is a very weakly acidic group, with a pK_a of about 10.1. This group is essentially fully protonated and uncharged at pH 7.

4.3 • Reactions of Amino Acids

Carboxyl and Amino Group Reactions

The α-carboxyl and α-amino groups of all amino acids exhibit similar chemical reactivity. The side chains, however, exhibit specific chemical reactivities, depending on the nature of the functional groups. Whereas all of these reactivities are important in the study and analysis of isolated amino acids, it is the characteristic behavior of the side chain that governs the reactivity of amino acids incorporated into proteins. There are three reasons to consider these reactivities. Proteins can be chemically modified in very specific ways by taking advantage of the chemical reactivity of certain amino acid side chains. The detection and quantification of amino acids and proteins often depend on reactions that are specific to one or more amino acids and that result in color, radioactivity, or some other quantity that can be easily measured. Finally and most importantly, the biological functions of proteins depend on the behavior and reactivity of specific R groups.

The carboxyl groups of amino acids undergo all the simple reactions common to this functional group. Reaction with ammonia and primary amines yields unsubstituted and substituted amides, respectively (Figure 4.9a,b). Esters

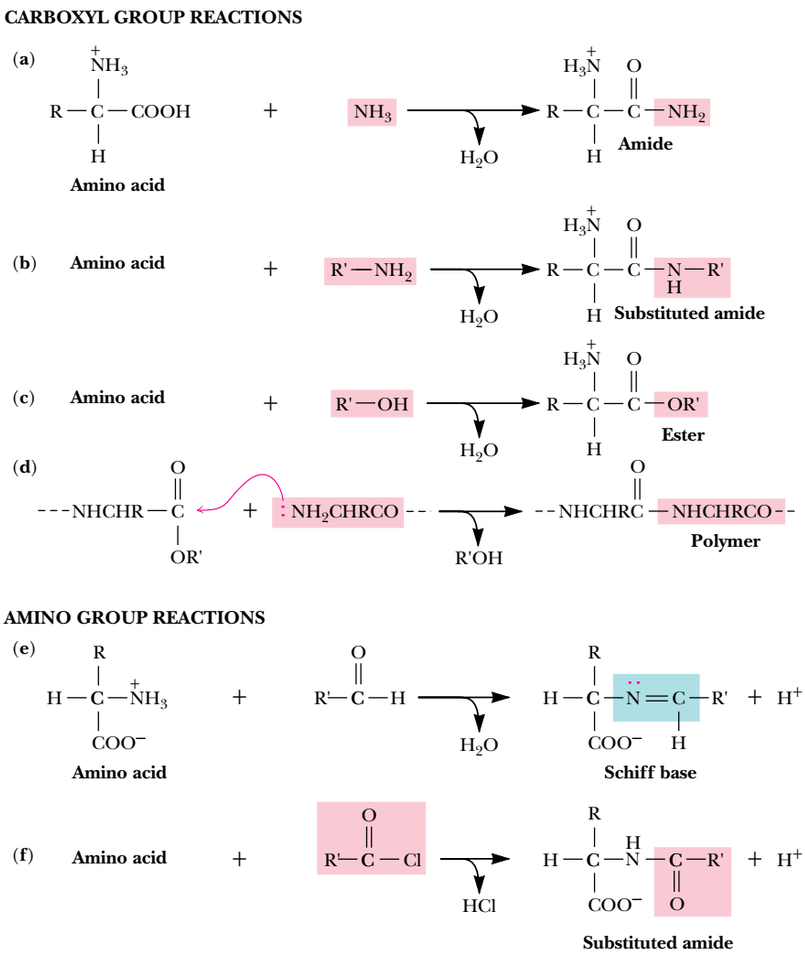


FIGURE 4.9 • Typical reactions of the common amino acids (see text for details).

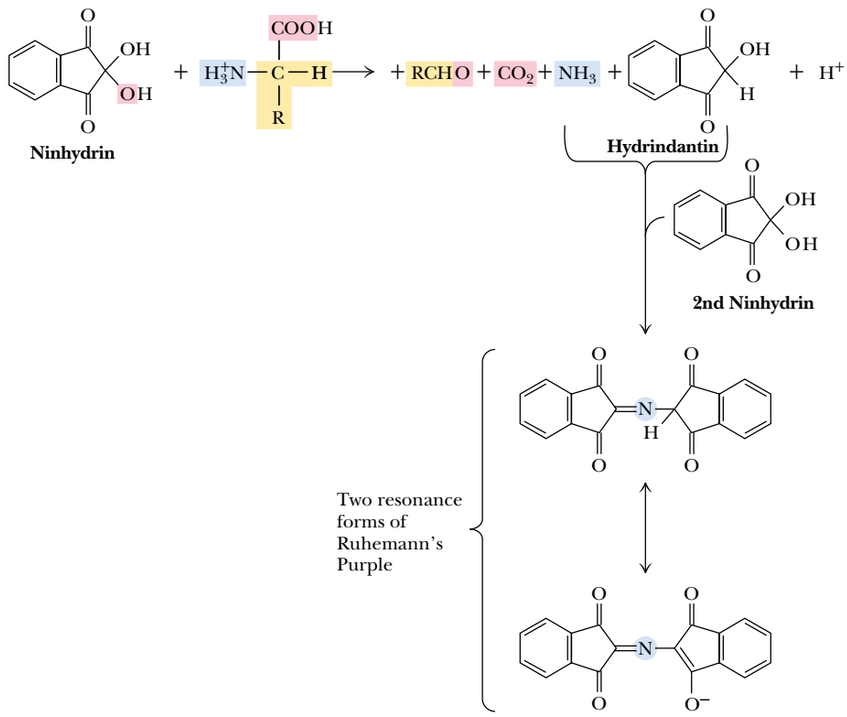


FIGURE 4.10 • The pathway of the ninhydrin reaction, which produces a colored product called "Ruhemann's Purple" that absorbs light at 570 nm. Note that the reaction involves and consumes two molecules of ninhydrin.

and acid chlorides are also readily formed. Esterification proceeds in the presence of the appropriate alcohol and a strong acid (Figure 4.9c). Polymerization can occur by repetition of the reaction shown in Figure 4.9d. Free amino groups may react with aldehydes to form Schiff bases (Figure 4.9e) and can be acylated with acid anhydrides and acid halides (Figure 4.9f).

The Ninhydrin Reaction

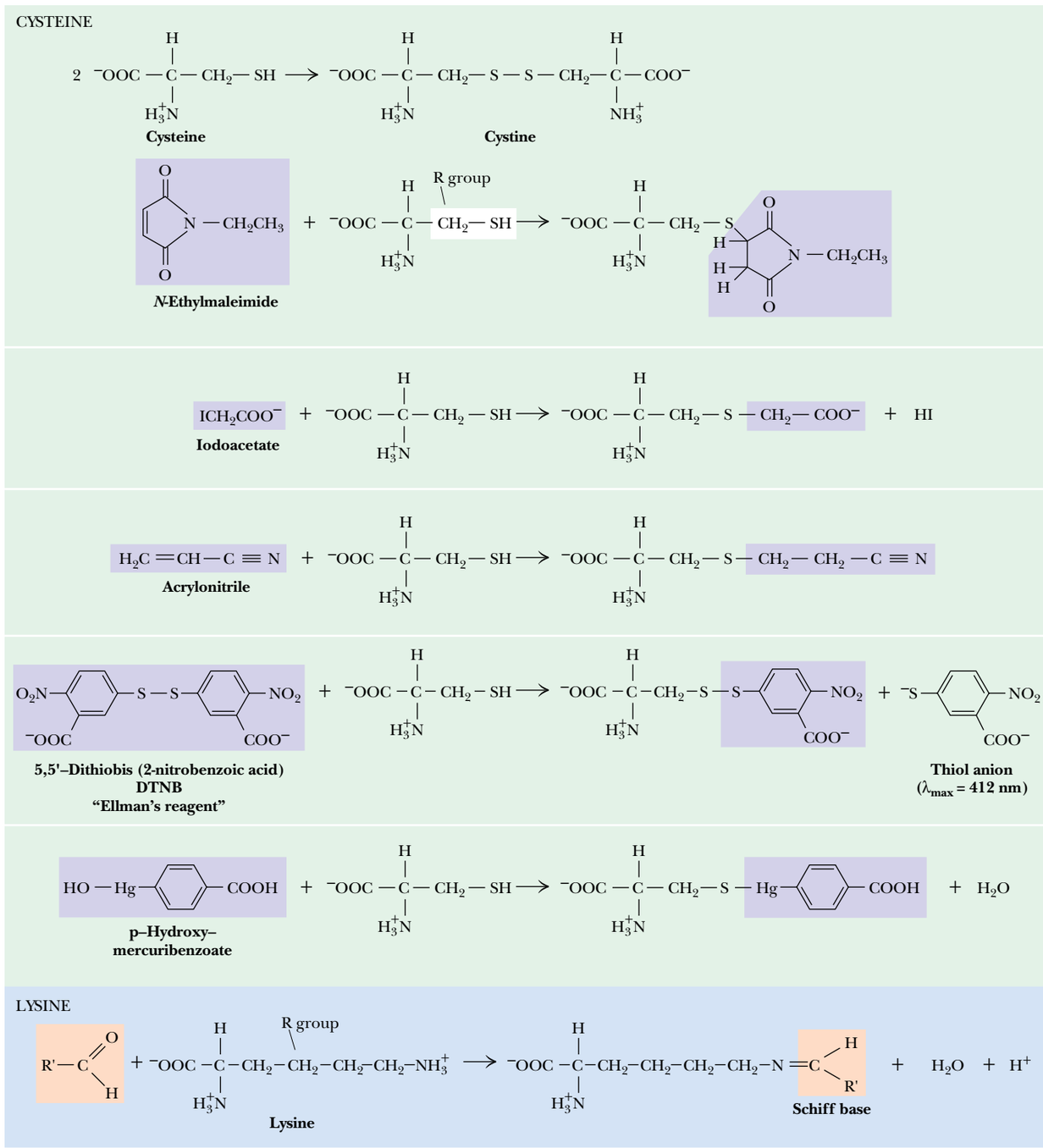
Amino acids can be readily detected and quantified by reaction with ninhydrin. As shown in Figure 4.10, *ninhydrin,* or triketohydrindene hydrate, is a strong oxidizing agent and causes the oxidative deamination of the α-amino function. The products of the reaction are the resulting aldehyde, ammonia, carbon dioxide, and hydrindantin, a reduced derivative of ninhydrin. The ammonia produced in this way can react with the hydrindantin and another molecule of ninhydrin to yield a purple product (Ruhemann's Purple) that can be quantified spectrophotometrically at 570 nm. The appearance of CO_2 can also be monitored. Indeed, CO_2 evolution is diagnostic of the presence of an α-amino acid. α-Imino acids, such as proline and hydroxyproline, give bright yellow ninhydrin products with absorption maxima at 440 nm, allowing these to be distinguished from the α-amino acids. Because amino acids are one of the components of human skin secretions, the ninhydrin reaction was once used extensively by law enforcement and forensic personnel for fingerprint detection. (Fingerprints as old as 15 years can be successfully identified using the ninhydrin reaction.) More sensitive fluorescent reagents are now used routinely for this purpose.

Specific Reactions of Amino Acid Side Chains

A number of reactions of amino acids have become important in recent years because they are essential to the degradation, sequencing, and chemical synthesis of peptides and proteins. These reactions are discussed in Chapter 5.

In recent years, biochemists have developed an arsenal of reactions that are relatively specific to the side chains of particular amino acids. These reactions can be used to identify functional amino acids at the active sites of enzymes or to label proteins with appropriate reagents for further study. Cysteine residues in proteins, for example, react with one another to form disulfide species and also react with a number of reagents, including maleimides (typically *N*-ethylmaleimide), as shown in Figure 4.11. Cysteines also react effectively

FIGURE 4.11 • Reactions of amino acid side-chain functional groups.



with iodoacetic acid to yield *S*-carboxymethyl cysteine derivatives. There are numerous other reactions involving specialized reagents specific for particular side chain functional groups. Figure 4.11 presents a representative list of these reagents and the products that result. It is important to realize that few if any of these reactions are truly specific for one functional group; consequently, care must be exercised in their use.

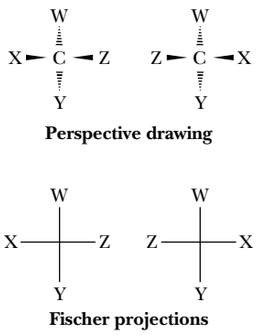


FIGURE 4.12 • Enantiomeric molecules based on a chiral carbon atom. Enantiomers are nonsuperimposable mirror images of each other.

4.4 • Optical Activity and Stereochemistry of Amino Acids

Amino Acids Are Chiral Molecules

Except for glycine, all of the amino acids isolated from proteins have four different groups attached to the α-carbon atom. In such a case, the α-carbon is said to be **asymmetric** or **chiral** (from the Greek *cheir*, meaning "hand"), and the two possible configurations for the α-carbon constitute nonsuperimposable mirror image isomers, or **enantiomers** (Figure 4.12). Enantiomeric molecules display a special property called **optical activity**—the ability to rotate the plane of polarization of plane-polarized light. Clockwise rotation of incident light is referred to as **dextrorotatory** behavior, and counterclockwise rotation is called **levorotatory** behavior. The magnitude and direction of the optical rotation depend on the nature of the amino acid side chain. The temperature, the wavelength of the light used in the measurement, the ionization state of the amino acid, and therefore the pH of the solution, can also affect optical rotation behavior. As shown in Table 4.2, some protein-derived amino acids at a given pH are dextrorotatory and others are levorotatory, even though all of them are of the L configuration. The direction of optical rotation can be specified in the name by using a (+) for dextrorotatory compounds and a (−) for levorotatory compounds, as in L(+)-leucine.

Nomenclature for Chiral Molecules

The discoveries of optical activity and enantiomeric structures (see the box, page 97) made it important to develop suitable nomenclature for chiral molecules. Two systems are in common use today: the so-called D,L system and the (*R*,*S*) system.

In the **D,L system** of nomenclature, the (+) and (−) isomers of glyceraldehyde are denoted as **D-glyceraldehyde** and **L-glyceraldehyde,** respectively (Figure 4.13). Absolute configurations of all other carbon-based molecules are referenced to D- and L-glyceraldehyde. When sufficient care is taken to avoid racemization of the amino acids during hydrolysis of proteins, it is found that all of the amino acids derived from natural proteins are of the L configuration. Amino acids of the D configuration are nonetheless found in nature, especially as components of certain peptide antibiotics, such as valinomycin, gramicidin, and actinomycin D, and in the cell walls of certain microorganisms.

In spite of its widespread acceptance, problems exist with the D,L system of nomenclature. For example, this system can be ambiguous for molecules with two or more chiral centers. To address such problems, the (***R*,*S***) **system** of nomenclature for chiral molecules was proposed in 1956 by Robert Cahn, Sir Christopher Ingold, and Vladimir Prelog. In this more versatile system, priorities are assigned to each of the groups attached to a chiral center on the basis of atomic number, atoms with higher atomic numbers having higher priorities (see the box, page 100).

The newer (*R*,*S*) system of nomenclature is superior to the older D,L system in one important way. The configuration of molecules with more than one

Table 4.2

Specific Rotations for Some Amino Acids

Amino Acid	Specific Rotation $[\alpha]_D^{25}$, Degrees
L-Alanine	+1.8
L-Arginine	+12.5
L-Aspartic acid	+5.0
L-Glutamic acid	+12.0
L-Histidine	−38.5
L-Isoleucine	+12.4
L-Leucine	−11.0
L-Lysine	+13.5
L-Methionine	−10.0
L-Phenylalanine	−34.5
L-Proline	−86.2
L-Serine	−7.5
L-Threonine	−28.5
L-Tryptophan	−33.7
L-Valine	+5.6

CRITICAL DEVELOPMENTS IN BIOCHEMISTRY

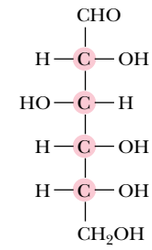
Discovery of Optically Active Molecules and Determination of Absolute Configuration

The optical activity of quartz and certain other materials was first discovered by Jean-Baptiste Biot in 1815 in France, and in 1848 a young chemist in Paris named Louis Pasteur made a related and remarkable discovery. Pasteur noticed that preparations of optically inactive sodium ammonium tartrate contained two visibly different kinds of crystals that were mirror images of each other. Pasteur carefully separated the two types of crystals, dissolved them each in water, and found that each solution was *optically active*. Even more intriguing, the specific rotations of these two solutions were equal in magnitude and of opposite sign. Because these differences in optical rotation were apparent properties of the dissolved molecules, Pasteur eventually proposed that the molecules themselves were mirror images of each other, just like their respective crystals. Based on this and other related evidence, in 1847 van't Hoff and LeBel proposed the tetrahedral arrangement of valence bonds to carbon.

In 1888, Emil Fischer decided that it should be possible to determine the *relative* configuration of (+)-glucose, a six-carbon sugar with four asymmetric centers (see figure). Because each of the four C could be either of two configurations, glucose conceivably could exist in any one of 16 possible isomeric structures. It took three years to complete the solution of an elaborate chemical and logical puzzle. By 1891, Fischer had reduced his puzzle to a choice between two enantiomeric structures. (Methods for determining *absolute* configuration were not yet available, so Fischer made a simple guess, selecting the structure shown in the figure.) For this remarkable feat, Fischer received the Nobel Prize in chemistry in 1902. Sadly, Fischer, a brilliant but troubled chemist, later committed suicide.

The absolute choice between Fischer's two enantiomeric possibilities would not be made for a long time. In 1951, J.M. Bijvoet in Utrecht, the Netherlands, used a new X-ray diffraction technique to determine the absolute configuration of (among other things) the sodium rubidium salt of (+)-tartaric acid. Because the tartaric acid configuration could be related to that of glyceraldehyde and because sugar and amino acid configurations could all be related to glyceraldehyde, it became possible to determine the absolute configuration of sugars and the common amino acids. The absolute configuration of tartaric acid determined by Bijvoet turned out to be the configuration that, up to then, had only been assumed. This meant that Emil Fischer's arbitrary guess 60 years earlier had been correct.

It was M.A. Rosanoff, a chemist and instructor at New York University, who first proposed (in 1906) that the isomers of glyceraldehyde be the standards for denoting the stereochemistry of sugars and other molecules. Later, when experiments showed that the configuration of (+)-glyceraldehyde was related to (+)-glucose, (+)-glyceraldehyde was given the designation D. Emil Fischer rejected the **Rosanoff convention,** but it was universally accepted. Ironically, this nomenclature system is often mistakenly referred to as the **Fischer convention.**



The absolute configuration of (+)-glucose.

chiral center can be more easily, completely, and unambiguously described with (*R*,*S*) notation. Several amino acids, including isoleucine, threonine, hydroxyproline, and hydroxylysine, have two chiral centers. In the (*R*,*S*) system, L-threonine is (2*S*,3*R*)-threonine. A chemical compound with n chiral centers can exist in 2^n-isomeric structures, and the four amino acids just listed can thus each take on four different isomeric configurations. This amounts to two pairs of enantiomers. Isomers that differ in configuration at only one of the asymmetric centers are non–mirror image isomers or **diastereomers.** The four stereo-

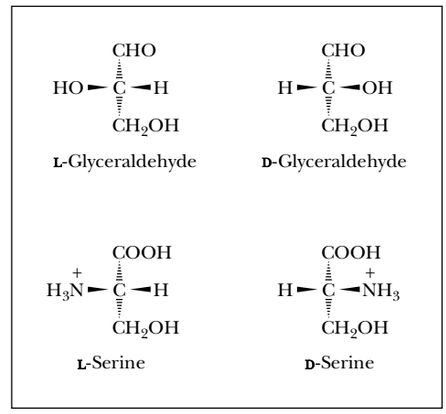


FIGURE 4.13 • The configuration of the common L-amino acids can be related to the configuration of L(−)-glyceraldehyde as shown. These drawings are known as Fischer projections. The horizontal lines of the Fischer projections are meant to indicate bonds coming out of the page from the central carbon, and vertical lines represent bonds extending behind the page from the central carbon atom.

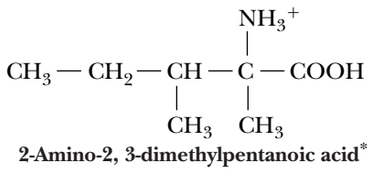
The Murchison Meteorite—Discovery of Extraterrestrial Handedness

The predominance of L-amino acids in biological systems is one of life's most intriguing features. Prebiotic syntheses of amino acids would be expected to produce equal amounts of L- and D-enantiomers. Some kind of enantiomeric selection process must have intervened to select L-amino acids over their D-counterparts as the constituents of proteins. Was it random chance that chose L- over D-isomers?

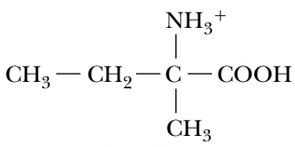
Analysis of carbon compounds—even amino acids—from extraterrestrial sources might provide deeper insights into this mystery. John Cronin and Sandra Pizzarello have examined the enantiomeric distribution of unusual amino acids obtained from the Murchison meteorite, which struck the earth on September 28, 1969, near Murchison, Australia. (By selecting unusual amino acids for their studies, Cronin and Pizzarello ensured that they were examining materials that were native to the meteorite and not earth-derived contaminants.) Four α-dialkyl amino acids—α-methylisoleucine, α-methylalloisoleucine, α-methylnorvaline, and isovaline—were found to have an L-enantiomeric excess of 2 to 9%.

This may be the first demonstration that a natural L-enantiomer enrichment occurs in certain cosmological environments. Could these observations be relevant to the emergence of L-enantiomers as the dominant amino acids on the earth? And, if so, could there be life elsewhere in the universe that is based upon the same amino acid handedness?

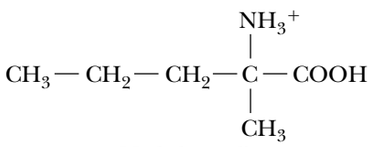
Amino acids found in the Murchison meteorite



2-Amino-2, 3-dimethylpentanoic acid*



Isovaline



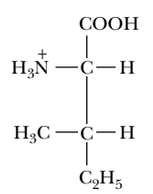
α-Methylnorvaline

**The four stereoisomers of this amino acid include the D- and L-forms of α-methylisoleucine and α-methylalloisoleucine.*

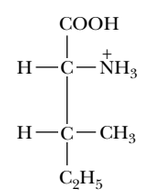
Cronin, J.R., and Pizzarello, S., 1997. Enantiomeric excesses in meteoritic amino acids. Science ***275:****951–955.*

isomers of isoleucine are shown in Figure 4.14. The isomer obtained from digests of natural proteins is arbitrarily designated L-isoleucine. In the (*R*,*S*) system, L-isoleucine is (2*S*,3*S*)-isoleucine. Its diastereomer is referred to as L-alloisoleucine. The D-enantiomeric pair of isomers is named in a similar manner.

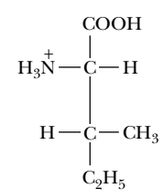
FIGURE 4.14 • The stereoisomers of isoleucine and threonine. The structures at the far left are the naturally occurring isomers.



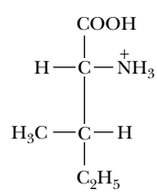
L-Isoleucine
(2*S*,3*S*)-Isoleucine



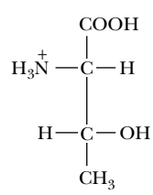
D-Isoleucine
(2*R*,3*R*)-Isoleucine



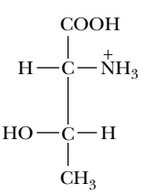
L-Alloisoleucine
(2*S*,3*R*)-Isoleucine



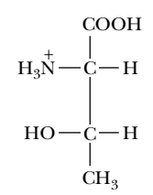
D-Alloisoleucine
(2*R*,3*S*)-Isoleucine



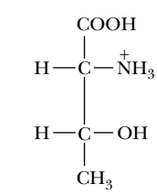
L-Threonine



D-Threonine



L-Allothreonine



D-Allothreonine

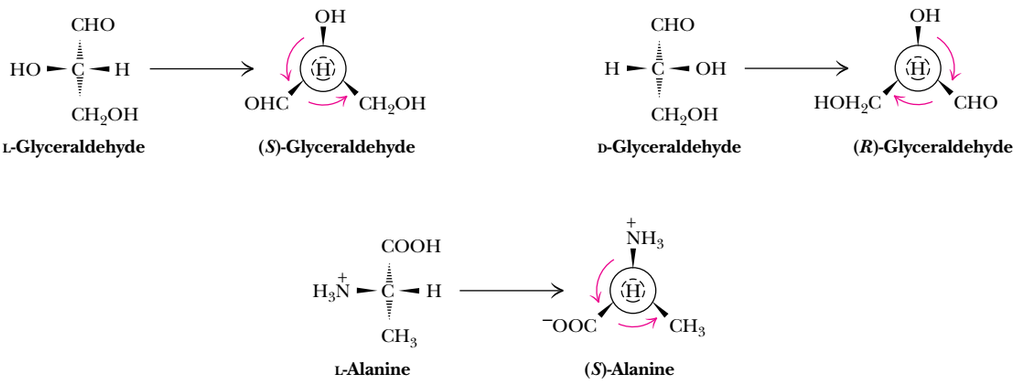
CRITICAL DEVELOPMENTS IN BIOCHEMISTRY

Rules for Description of Chiral Centers in the (R,S) *System*

Naming a chiral center in the (*R*,*S*) system is accomplished by viewing the molecule from the chiral center to the atom with the lowest priority. If the other three atoms facing the viewer then decrease in priority in a clockwise direction, the center is said to have the (*R*) configuration (where *R* is from the Latin *rectus* meaning "right"). If the three atoms in question decrease in priority in a counterclockwise fashion, the chiral center is of the (*S*) configuration (where *S* is from the Latin *sinistrus* meaning "left"). If two of the atoms coordinated to a chiral center are identical, the atoms bound to these two are considered for priorities. For such purposes, the priorities of certain functional groups found in amino acids and related molecules are in the following order:

$$SH > OH > NH_2 > COOH > CHO > CH_2OH > CH_3$$

From this, it is clear that D-glyceraldehyde is (*R*)-glyceraldehyde, and L-alanine is (*S*)-alanine (see figure). Interestingly, the α-carbon configuration of all the L-amino acids *except for cysteine* is (*S*). Cysteine, by virtue of its thiol group, is in fact (*R*)-cysteine.



The assignment of (*R*) and (*S*) notation for glyceraldehyde and L-alanine.

4.5 • Spectroscopic Properties of Amino Acids

One of the most important and exciting advances in modern biochemistry has been the application of **spectroscopic methods,** which measure the absorption and emission of energy of different frequencies by molecules and atoms. Spectroscopic studies of proteins, nucleic acids, and other biomolecules are providing many new insights into the structure and dynamic processes in these molecules.

Ultraviolet Spectra

Many details of the structure and chemistry of the amino acids have been elucidated or at least confirmed by spectroscopic measurements. None of the amino acids absorbs light in the visible region of the electromagnetic spectrum. Several of the amino acids, however, do absorb **ultraviolet** radiation, and all absorb in the **infrared** region. The absorption of energy by electrons as they rise to higher energy states occurs in the ultraviolet/visible region of the energy spectrum. Only the aromatic amino acids phenylalanine, tyrosine, and tryptophan exhibit significant ultraviolet absorption above 250 nm, as shown in Figure 4.15. These strong absorptions can be used for spectroscopic determinations of protein concentration. The aromatic amino acids also exhibit relatively weak fluorescence, and it has recently been shown that tryptophan can exhibit *phos-*