Organic Chemistry III

Mohammad Jafarzadeh

Faculty of Chemistry, Razi University

Organic Chemistry, *Structure and Function* **(7th edition)**

By *P. Vollhardt* and *N. Schore*, Elsevier, 2014

26. Amino Acids, Peptides, Proteins, and Nucleic Acids

Organic molecules constitute the chemical bricks of life. A functional definition of life refers to it as a condition of matter manifested by growth, metabolism, reproduction, and evolution.

Amino acids and their polymers (polypeptides), the large natural polypeptides called **proteins** and their biological origin, **DNA.**

Proteins have diversity of functions in living systems. As **enzymes,** they catalyze transformations ranging from the simple hydration of carbon dioxide to the replication of entire chromosomes—great coiled strands of DNA, the genetic material in living cells.

The protein of rhodopsin, the photoreceptor that generates and transmits nerve impulses in retinal cells. Hemoglobin carries oxygen; iron is transported in the blood by transferrin and stored in the liver by ferritin.

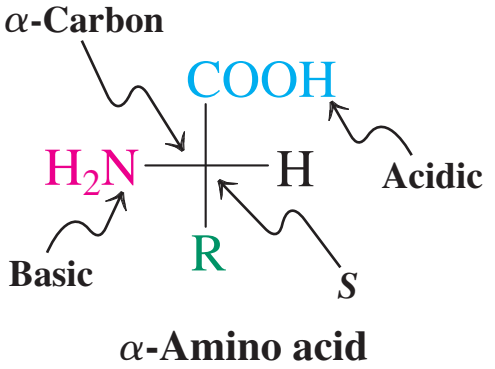
Proteins play a crucial role in coordinated motion (e.g. muscle contraction), and being mechanical support to skin and bone. They are the antibodies responsible for our immune protection; and they control growth and differentiation.

26-1 STRUCTURE AND PROPERTIES OF AMINO ACIDS

Amino acids are carboxylic acids that bear an amine group. The most common of these in nature are the **2-amino acids,** or ***α*-amino acids,** which have the general formula $RCH(NH_2)COOH$; that is, the amino function is located at C2, the α-carbon.

The R group can be alkyl or aryl, and it can contain hydroxy, amino, mercapto, sulfide, carboxy, guanidino, or imidazolyl groups. Because of the presence of both amino and carboxy functions, amino acids are both acidic and basic.

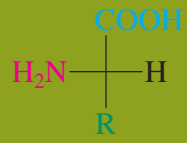
The stereocenter of 2-amino acids has the *S* configuration

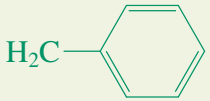
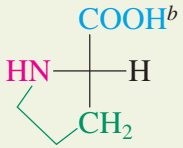
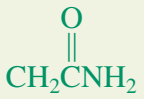


α-Amino acid

More than 500 amino acids exist in nature, but the proteins in all species, from bacteria to humans, consist mainly of only 20. Adult humans can synthesize all but eight, and two only in insufficient quantities. This group is often called the **essential amino acids** because they must be included in our diet.

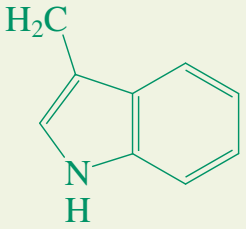
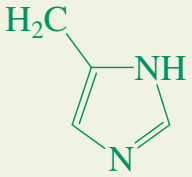
Table 26-1 Natural (2*S*)-Amino Acids



R	Name	Three-letter code	One-letter code	pK_a of α-COOH	pK_a of α-$^+NH_3$	pK_a of acidic function in R	Isoelectric point, p*I*
H	Glycine	Gly	G	2.3	9.6	—	6.0
Alkyl group							
CH_3	Alanine	Ala	A	2.3	9.7	—	6.0
$CH(CH_3)_2$	Valine[a]	Val	V	2.3	9.6	—	6.0
$CH_2CH(CH_3)_2$	Leucine[a]	Leu	L	2.4	9.6	—	6.0
$CHCH_2CH_3$ (*S*) $\lvert$ CH_3	Isoleucine[a]	Ile	I	2.4	9.6	—	6.0
	Phenylalanine[a]	Phe	F	1.8	9.1	—	5.5
	Proline	Pro	P	2.0	10.6	—	6.3
Hydroxy containing							
CH_2OH	Serine	Ser	S	2.2	9.2	—	5.7
$CHOH$ (*R*) $\lvert$ CH_3	Threonine[a]	Thr	T	2.1	9.1	—	5.6
	Tyrosine	Tyr	Y	2.2	9.1	10.1	5.7
Amino containing							
	Asparagine	Asn	N	2.0	8.8	—	5.4

Continued

Table 26-1 Natural (2*S*)-Amino Acids (continued)

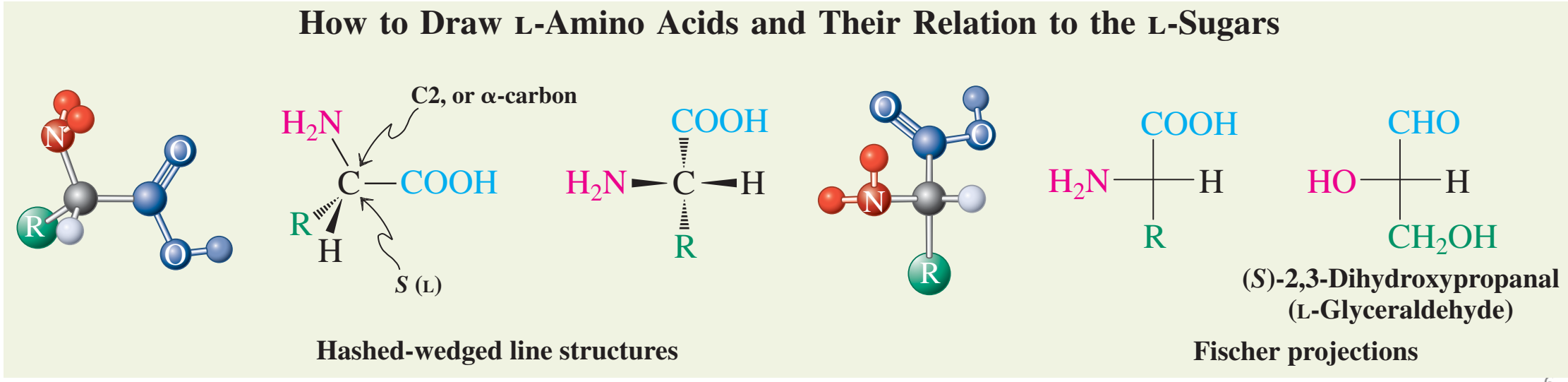
R	Name	Three-letter code	One-letter code	pK_a of α-COOH	pK_a of α-$^+NH_3$	pK_a of acidic function in R	Isoelectric point, p*I*
Amino containing *(continued)*							
$CH_2CH_2\overset{\overset{O}{\|}}{C}NH_2$	Glutamine	Gln	Q	2.2	9.1	—	5.7
$(CH_2)_4NH_2$	Lysine[a]	Lys	K	2.2	9.0	10.5[c]	9.7
$(CH_2)_3NH\overset{\overset{NH}{\|}}{C}NH_2$	Arginine[a]	Arg	R	2.2	9.0	12.5[c]	10.8
	Tryptophan[a]	Trp	W	2.8	9.4	—	5.9
	Histidine[a]	His	H	1.8	9.2	6.1[c]	7.6
Mercapto or sulfide containing							
CH_2SH	Cysteine[d]	Cys	C	2.0	10.3	8.2	5.1
$CH_2CH_2SCH_3$	Methionine[a]	Met	M	2.3	9.2	—	5.7
Carboxy containing							
CH_2COOH	Aspartic acid	Asp	D	1.9	9.6	3.7	2.8
CH_2CH_2COOH	Glutamic acid	Glu	E	2.2	9.7	4.3	3.2

[a]Essential amino acids. [b]Entire structure. [c]pK_a of conjugate acid. [d]The stereocenter is *R* because the CH_2SH substituent has higher priority than the COOH group.

In all but glycine, the simplest of the amino acids, C2 is a stereocenter and usually adopts the *S* configuration. Other stereocenters located in the substituent R may have either *R* (as in threonine) or *S* configuration (as in isoleucine).

As in the names of the sugars, an older amino acid nomenclature uses the prefixes D and L, which relate all the L-amino acids to (*S*)-2,3-dihydroxypropanal (L- glyceraldehyde).

A molecule belonging to the L family is not necessarily levorotatory, for example, both valine ($[a]_D^{25°C}$ = +13.9) and isoleucine ($[a]_D^{25°C}$ = +11.9) are dextrorotatory.



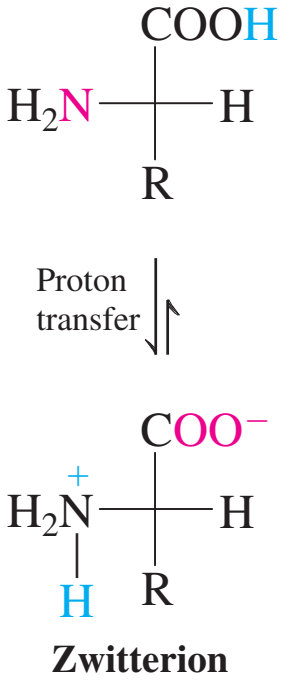
Amino acids are acidic and basic: zwitterions

Because of their two functional groups, the amino acids are both acidic and basic; that is, they are **amphoteric**. The carboxylic acid group protonates the amine function, thus forming a **zwitterion.**

This ammonium carboxylate form is favored because an ammonium ion is much less acidic ($pK_a \approx 10–11$) than a carboxylic acid ($pK_a \approx 2–5$).

The highly polar zwitterionic structure allows amino acids to form particularly strong crystal lattices.

Most of them therefore are fairly insoluble in organic solvents, and they decompose rather than melt when heated.



Zwitterion

The structure of an amino acid in aqueous solution depends on the pH. Consider, for example, the simplest member of the series, glycine.

The major form in neutral solution is the zwitterion. However, in strong acid (pH < 1), glycine exists predominantly as the cat- ionic ammonium carboxylic acid, whereas strongly basic solutions (pH > 13) contain mainly the deprotonated 2-aminocarboxylate ion.

These forms interconvert by acid-base equilibria.

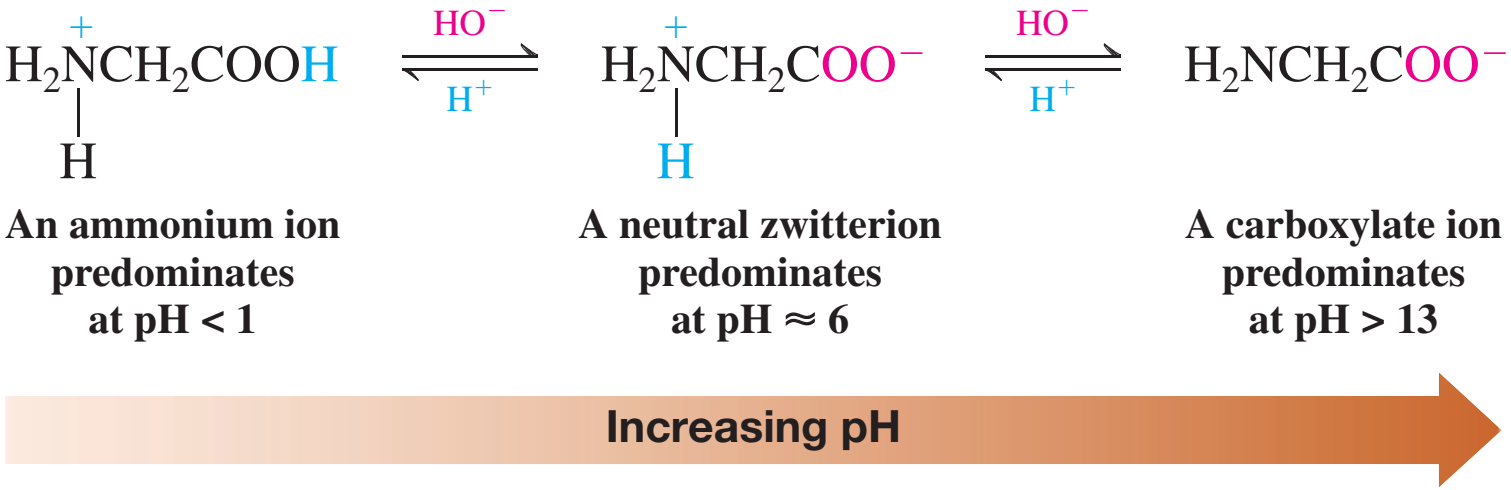
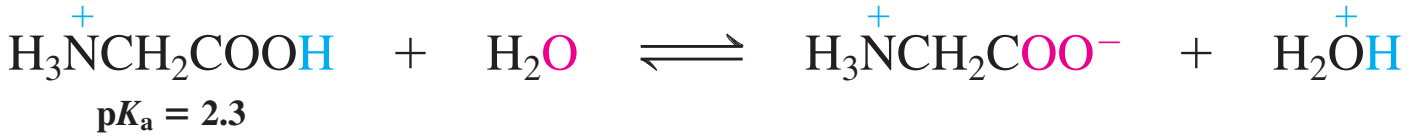
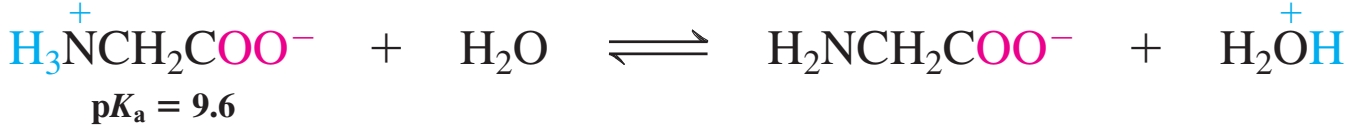


Table 26-1 records pK_a values for each functional group of the amino acids. For glycine, the first value (2.3) refers to the equilibrium:



$$K_1 = \frac{[H_3\overset{+}{N}CH_2COO^-][H_2O\overset{+}{H}]}{[H_3\overset{+}{N}CH_2COOH]} = 10^{-2.3}$$

Note that this pK_a is more than two units less than that of an ordinary carboxylic acid (pK_a CH_3COOH = 4.74), an observation that is true for all the other α-aminocarboxy groups in Table 26-1. This difference is a consequence of the electron-withdrawing effect of the protonated amino group. The second pK_a value (9.6) describes the second deprotonation step:



$$K_2 = \frac{[H_2NCH_2COO^-][H_2O\overset{+}{H}]}{[H_3\overset{+}{N}CH_2COO^-]} = 10^{-9.6}$$

At the isoelectric point, the net charge is zero

The pH at which the extent of protonation equals that of deprotonation is called the **isoelectric pH** or the **isoelectric point (p*I*;** Table 26-1).

At this pH, the amount of positive charge balances that of negative charge and the concentration of the charge-neutralized zwitterionic form is at its greatest.

For an amino acid without any additional acidic or basic groups, such as glycine, the value of its p*I* is the average of its two pK_a values.

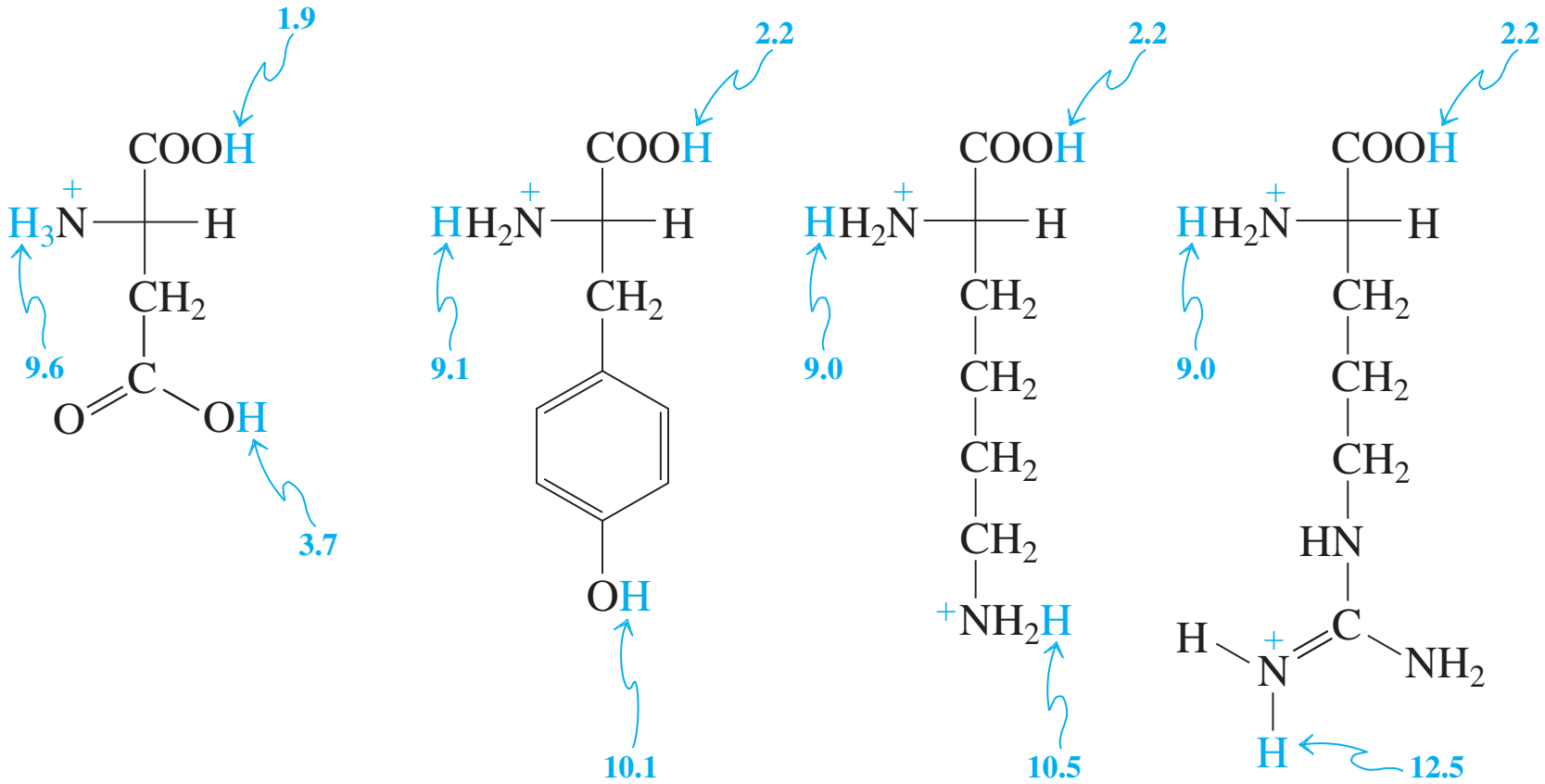
$$\mathrm{p}I = \frac{\mathrm{p}K_{\text{a-COOH}} + \mathrm{p}K_{\text{a-}\overset{+}{\mathrm{N}}\mathrm{H_2H}}}{2} = \text{(for glycine)}\ \frac{2.3 + 9.6}{2} = 6.0$$

pH at which concentration of zwitterion is at its maximum

When the side chain of the acid bears an additional acidic or basic function, the p*I* is either decreased or increased, respectively.

For the four amino acids with an acidic side chain, the p*I* is the average of its two lowest pK_a values. Conversely, for the three amino acids with a basic side chain, the p*I* is the average of its two highest pK_a values.

Assignment of pK_a Values in Selected Amino Acids



Aspartic acid **Tyrosine** **Lysine** **Arginine**

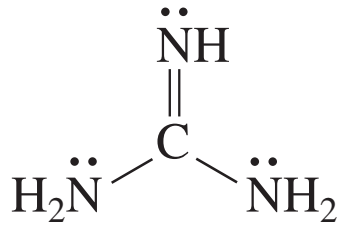
The amino dicarboxylic acid, aspartic acid, will be positively charged at low pH due to the presence of the ammonium substituent. To reach a charge-neutral form, on average one of the carboxy functions has to be deprotonated. This will happen at a pH midpoint between the two respective pK_a values (1.9 and 3.7), at pI = 2.8.

The similar glutamic acid has the corresponding value at 3.2. At physiological pH, both of the carboxy functions are deprotonated, and the molecules exist as the zwitterionic anions aspartate and glutamate. (Monosodium glutamate, MSG, is used as a flavor enhancer in various foods.)

Tyrosine, which bears the relatively nonacidic neutral phenol substituent (at low pH), has a pI = 5.7, which is midway between the pK_a's of the other two more acidic groups.

In contrast, lysine bears an additional basic amino group that is protonated in a strongly acidic medium to furnish a dication. When the pH of the solution is raised, deprotonation of the carboxy group occurs first, followed by respective proton loss from the nitrogen at C2 and the remote ammonium function. The isoelectric point is located halfway between the last two pK_a values, at pI = 9.7.

Arginine carries a substituent derived from the molecule guanidine. The pK_a of its conjugate acid is 12.5, almost three units greater than that of the ammonium ion. The p*I* of arginine lies midway between that of the guanidinium and ammonium groups, at 10.8.

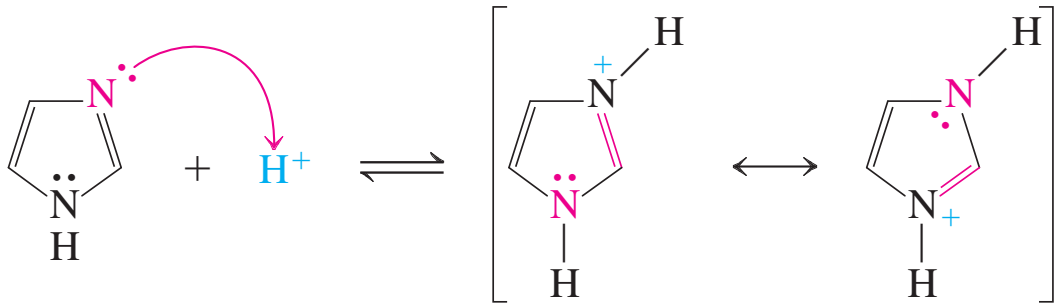


Guanidine

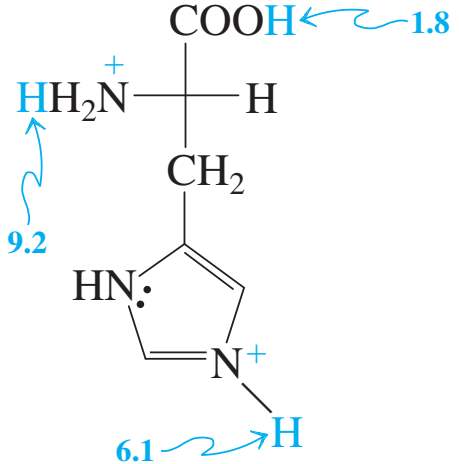
Histidine contains the basic **imidazole** ring. In this aromatic heterocycle, one of the nitrogen atoms is hybridized as in pyridine and the other is hybridized as in pyrrole.

The imidazole ring is relatively basic because the protonated species is stabilized by resonance.

Imidazole is significantly protonated at physiological pH (p*I* = 7.6). It can therefore function as a proton acceptor and donor at the active site of a variety of enzymes (e.g., chymotrypsin).



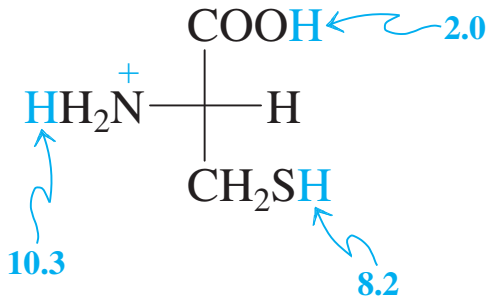
pK_a = 7.0



Histidine

The amino acid cysteine bears a relatively nucleophilic and acidic mercapto substituent (p*K* = 8.2, p*I* = 5.1). In addition, thiols can be oxidized to disulfides under mild conditions.

In nature, various enzymes are capable of oxidatively coupling and reductively decoupling the mercapto groups in the cysteines of proteins and peptides, thereby reversibly linking peptide strands.



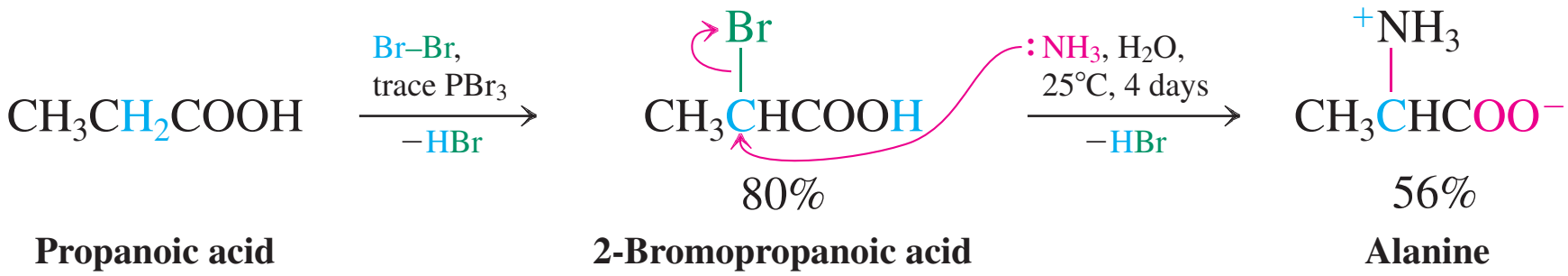
Cysteine

26-2 SYNTHESIS OF AMINO ACIDS: A COMBINATION OF AMINE AND CARBOXYLIC ACID CHEMISTRY

The simple 2-functionalization of an acid is possible by the **Hell-Volhard-Zelinsky** bromination. Furthermore, the bromine in the product can be displaced by nucleophiles, such as ammonia. In these two steps, propanoic acid can be converted into racemic alanine.

Unfortunately, this approach suffers frequently from relatively low yields.

Hell-Volhard-Zelinsky bromination followed by amination converts carboxylic acids to 2-amino acids



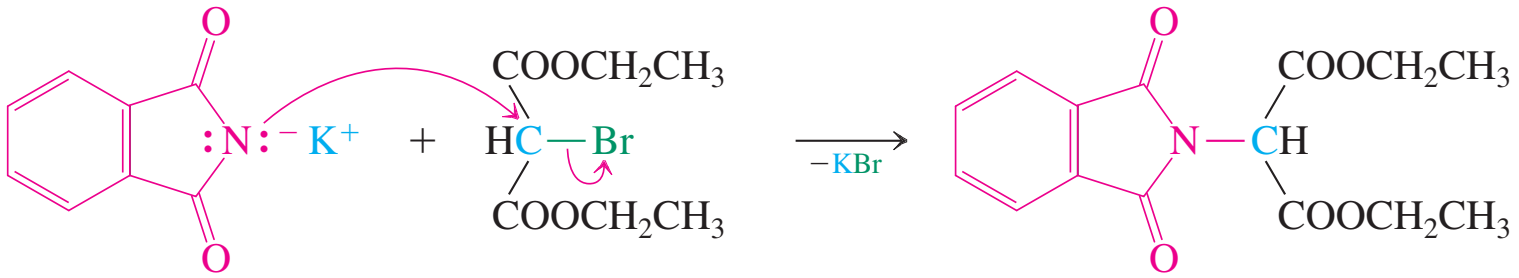
A better synthesis utilizes **Gabriel**'s procedure for the preparation of primary amines. Recall that *N*-alkylation of 1,2-benzenedicarboxylic imide (phthalimide) anion followed by acid hydrolysis furnishes amines.

The Gabriel synthesis can be adapted to produce amino acids

To prepare an amino acid, diethyl 2-bromopropanedioate (diethyl 2-bromomalonate) can be used in the first step of the reaction sequence. This alkylating agent is readily available from the bromination of diethyl propanedioate (malonate). The alkylation product can be hydrolyzed and decarboxylated. Hydrolysis of the imide group then furnishes an amino acid.

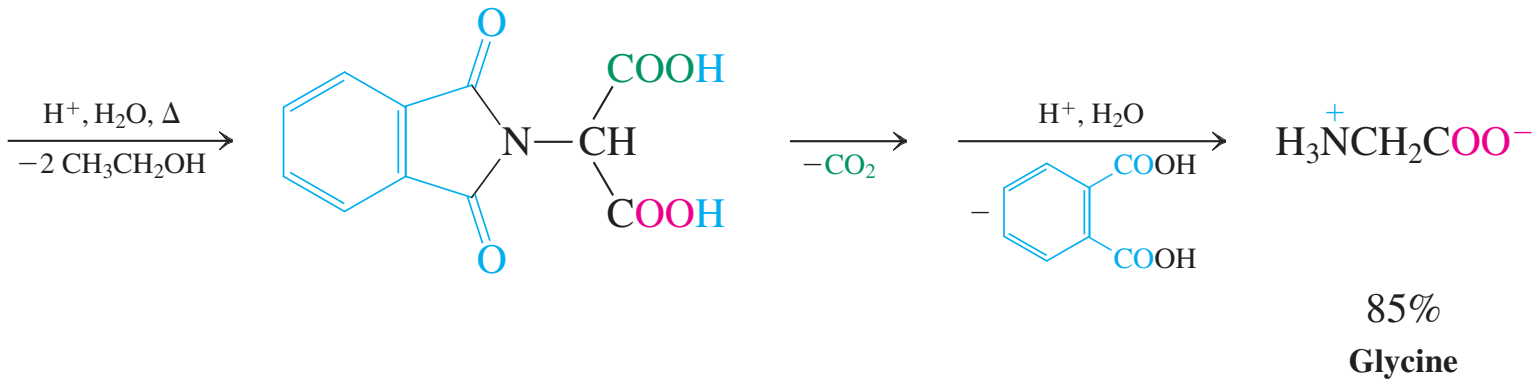
One of the advantages of this approach is the versatility of the initially formed 2-substituted propanedioate. This product can itself be alkylated, thus allowing for the preparation of a variety of substituted amino acids.

Gabriel Synthesis of Glycine

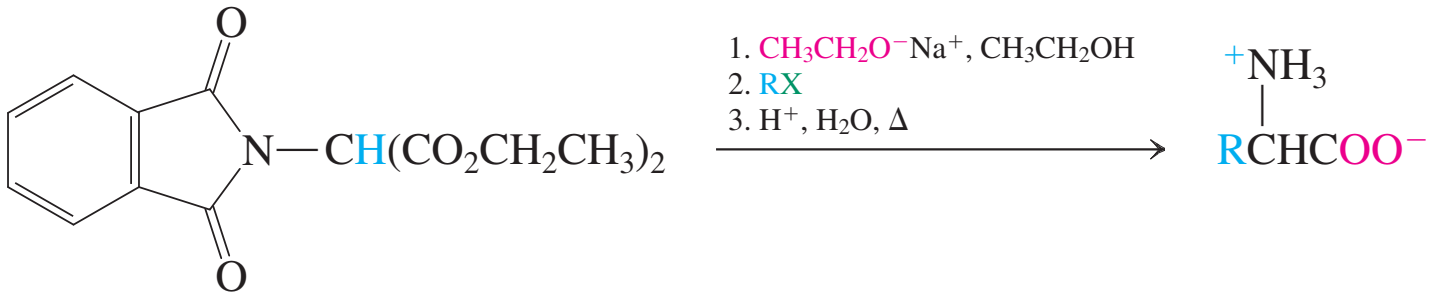


Potassium 1,2-benzene-dicarboxylic imide (Potassium phthalimide)

Diethyl 2-bromo-propanedioate (Diethyl 2-bromo-malonate)

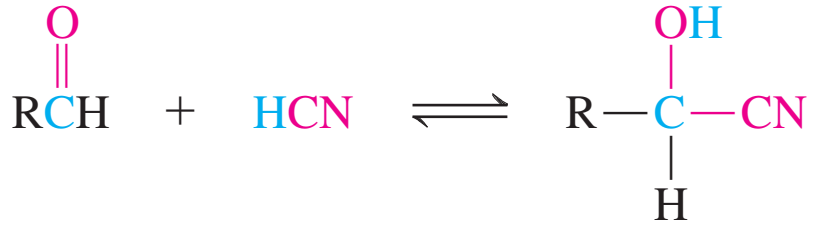


Glycine



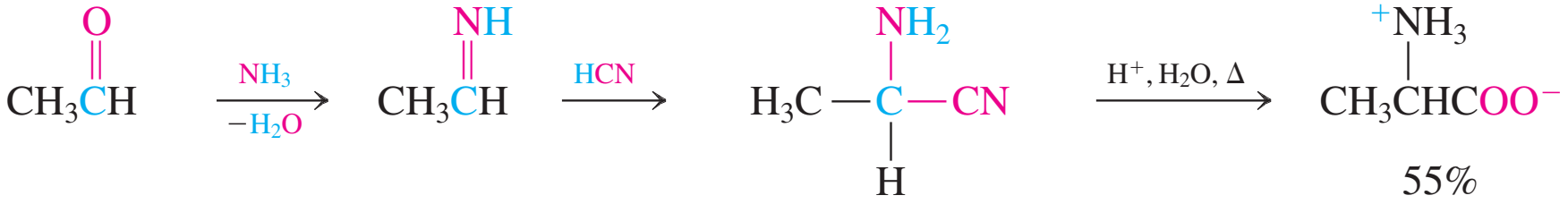
Amino acids are prepared from aldehydes by the Strecker synthesis

The crucial step in the **Strecker synthesis** is a variation of the cyanohydrin formation from aldehydes and hydrogen cyanide. When the same reaction is carried out in the presence of ammonia or ammonium cyanide, $H_4N^+CN^-$, it is the intermediate imine that undergoes addition of hydrogen cyanide, to furnish the corresponding 2-amino nitriles. Subsequent acidic or basic hydrolysis results in the desired amino acids.



Cyanohydrin

Strecker Synthesis of Alanine



Acetaldehyde **Imine intermediate** **2-Aminopropanenitrile** **Alanine**

26-3 SYNTHESIS OF ENANTIOMERICALLY PURE AMINO ACIDS

All the methods of the preceding section produce amino acids in racemic form, while most of the amino acids in natural polypeptides have the *S* configuration. Thus, many synthetic procedures—in particular, peptide and protein syntheses—require enantiomerically pure compounds.

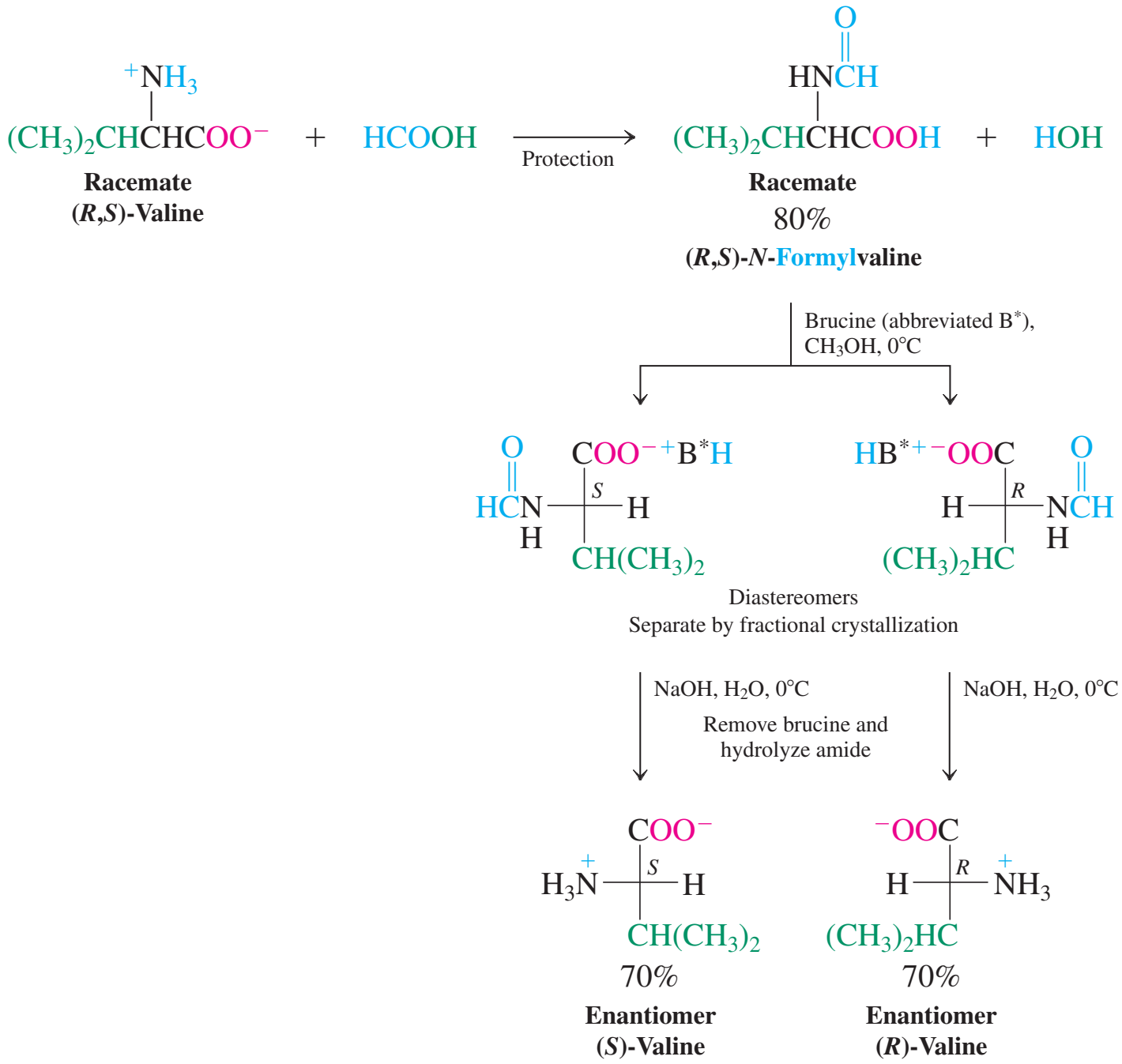
To meet this requirement, either the racemic amino acids must be resolved or a single enantiomer must be prepared by enantioselective reactions.

A conceptually straightforward approach to the preparation of pure enantiomers of amino acids would be resolution of their diastereomeric salts.

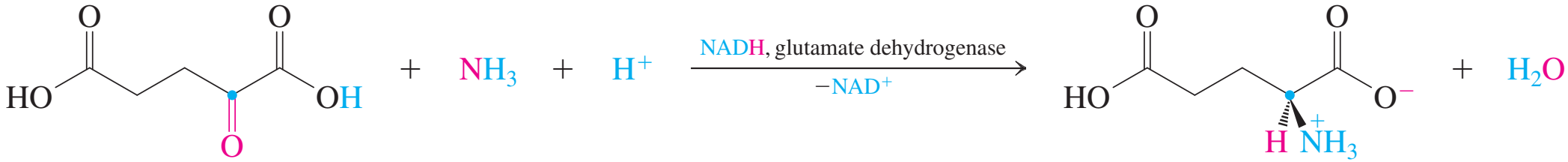
Typically, the amine group is first protected as an amide and the resulting product is then treated with an optically active amine, such as the inexpensive alkaloid brucine. The two diastereomers formed can be separated by fractional crystallization.

Unfortunately, in practice, this method can be tedious and can suffer from poor yields.

Resolution of Racemic Valine



In an alternative approach, the stereocenter at C2 is formed enantioselectively, such as in enantioselective hydrogenations of α,β-unsaturated amino acids. Nature makes use of this strategy in the biosynthesis of amino acids. Thus, the enzyme glutamate dehydrogenase converts the carbonyl group in 2-oxopentanedioic acid into the amine substituent in (*S*)-glutamic acid by a biological reductive amination (for chemical reductive aminations). The reducing agent is NADH.



2-Oxopentanedioic acid **(*S*)-Glutamic acid**

(*S*)-Glutamic acid is the biosynthetic precursor of glutamine, proline, and arginine. Moreover, it functions to aminate other 2-oxo acids, with the help of another type of enzyme, a transaminase, making additional amino acids available.