UNIT I: Protein Structure and Function

Structure of 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 serum albumin shuttle molecules essential to life, whereas immunoglobulins destroy 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 biological functions.

II. STRUCTURE OF AMINO ACIDS

Although more than 100 different amino acids have been described in nature, only 20 of these species are commonly found as constituents of mammalian proteins. Each amino acid (except for proline) has a carboxyl group, an amino group, and a distinctive side chain ("R-group") bonded to the α-carbon (Figure 1.1). At physiologic pH (approximately 7.4) the carboxyl group is dissociated to form the negatively charged carboxylate ion (—COO^-), and the amino group is protonated (—NH_3^+). In proteins these carboxyl and amino groups are combined in peptide linkage and are not available for chemical reaction (except for hydrogen bonds as described on p. 17). Thus, it is the nature of the side chains that ultimately dictates the role an amino acid will play in a protein. It is therefore useful to classify the amino acids according to the polarity of their side chains, that is, whether they are nonpolar, uncharged polar, acidic, or basic (Figure 1.2).

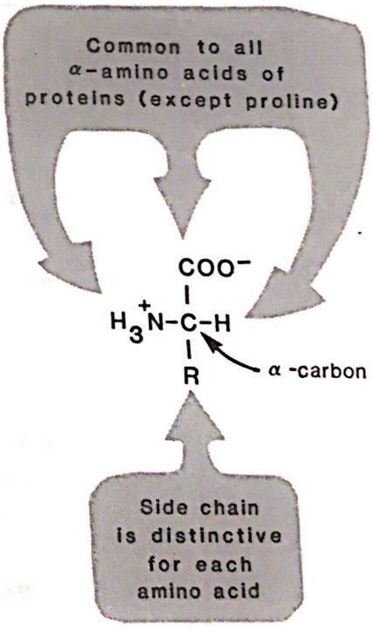
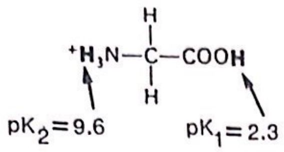


Figure 1.1
Structural features of amino acids.

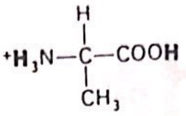
A. Amino acids with nonpolar side chains

Glycine = Gly
Alanine = Ala
Valine = Val
Leucine = Leu

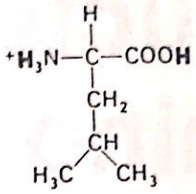
NON-POLAR SIDE CHAINS



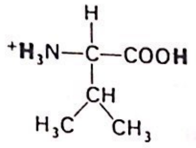
Glycine



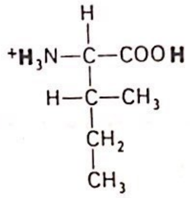
Alanine



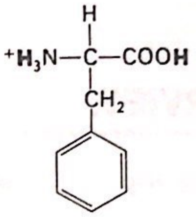
Leucine



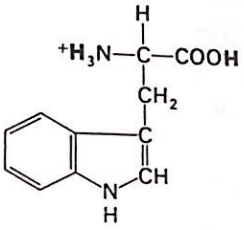
Valine



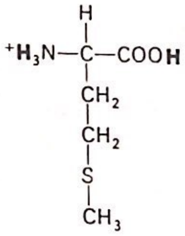
Isoleucine



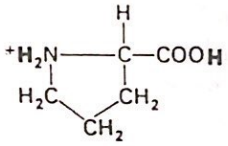
Phenylalanine



Tryptophan

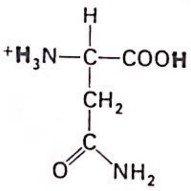


Methionine

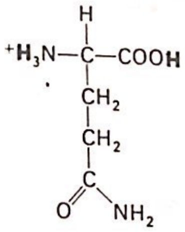


Proline

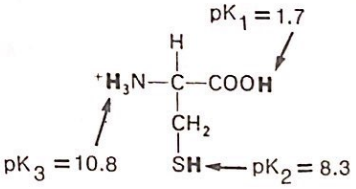
UNCHARGED POLAR SIDE CHAINS



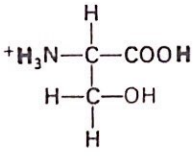
Asparagine



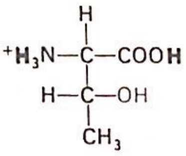
Glutamine



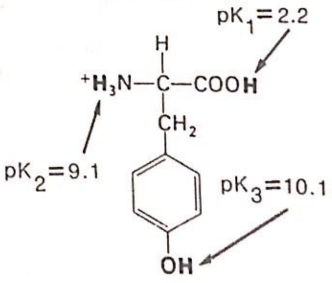
Cysteine



Serine

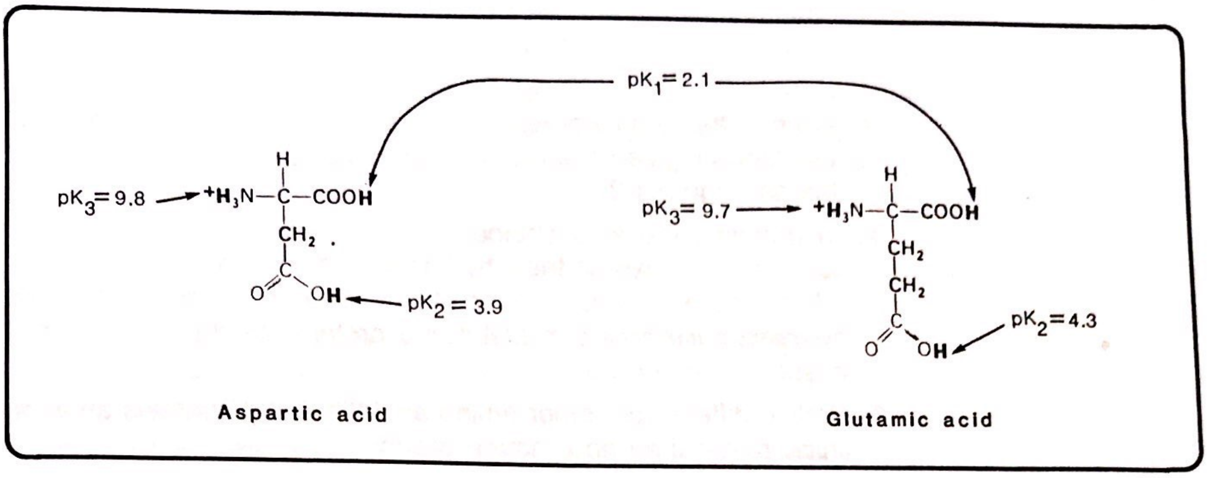


Threonine



Tyrosine

ACIDIC SIDE CHAINS



BASIC SIDE CHAINS

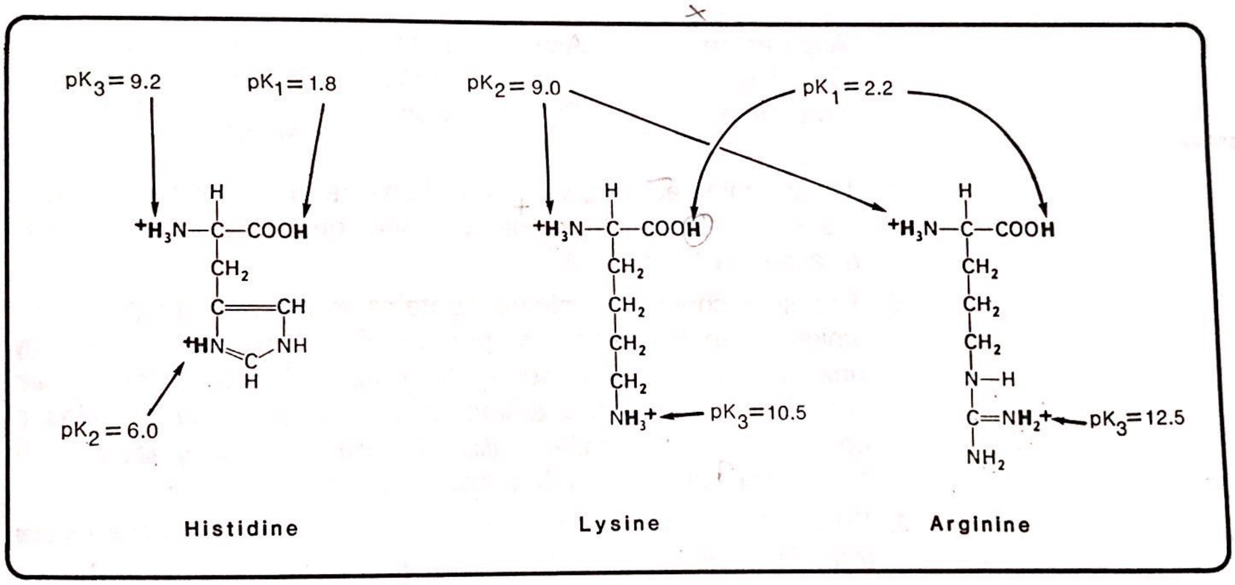
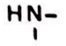


Figure 1.2
Classification of the 20 amino acids found in proteins according to the charge and polarity of their side chains. Each amino acid is shown in its fully protonated form with dissociable hydrogen ions represented in bold print. (For convenience the protonated amino group is shown as —NH_3^+ with the understanding that only one hydrogen of the amino group can dissociate.) The pK values for the α-carboxyl and α-amino groups of the nonpolar amino acids are similar to those shown for glycine.

Isoleucine = Ile Phenylalanine = Phe
Tryptophan = Trp Methionine = Met
Proline = Pro

1. Each of these amino acids has a nonpolar side chain that does not bind or give off protons, or participate in hydrogen or ionic bonds (Figure 1.2).
2. In proteins, the side chains of these amino acids can cluster together because of their hydrophobicity, much like droplets of oil coalesce in aqueous solutions. The importance of these ***hydrophobic interactions*** in stabilizing protein structure will be discussed on p. 17.
3. Proline differs from other amino acids in that it contains an ***imino group,*** rather than an α-amino group.


$HN-$

Imino group

H_2N-

Amino group

B. Amino acids with uncharged polar side chains

Asparagine = Asn Glutamine = Gln
Cysteine = Cys Serine = Ser
Threonine = Thr Tyrosine = Tyr

1. These amino acids have zero net charge at neutral pH, although the side chains of cysteine and tyrosine can lose a proton at alkaline pH (Figure 1.2).
2. 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 condense to form the amino acid cystine, which contains a covalent cross-link called a ***disulfide bond*** (—S—S—). (See p. 18 for a discussion of disulfide bond formation.)
3. Serine, threonine, and tyrosine all contain a polar ***hydroxyl group*** that, in proteins, can participate in hydrogen bond formation (Figure 1.3) or serve as the site of attachment for a phosphate group or a carbohydrate. 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 or serve as a site of attachment for carbohydrates. (See p. 17 for a discussion of hydrogen bonds.)

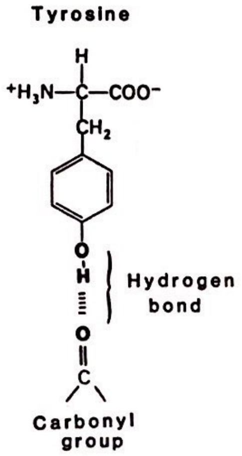


Figure 1.3
Hydrogen bond between the phenolic hydroxyl group of tyrosine and another compound containing a carbonyl oxygen.

C. Amino acids with acidic side chains

Aspartic acid = Asp Glutamic acid = Glu

1. The amino acids aspartic and glutamic acid are proton donors: At neutral pH the side chains of these amino acids are fully ionized and contain a ***negatively charged*** carboxylate group (—COO⁻). They are therefore called aspartate or glutamate to emphasize that at physiologic pH these amino acids are negatively charged (Figure 1.2).

D. Amino acids with basic side chains

Lysine = Lys Arginine = Arg
Histidine = His

1. The side chains of the basic amino acids bind protons (Figure 1.2): At physiologic pH the side chains of lysine and arginine are fully ionized and ***positively charged.***
2. In contrast, histidine is weakly basic and the free amino acid largely uncharged at physiologic pH. In proteins, however, the side chain of histidine can be either positively charged or neutral depending on the ionic environment provided by the polypeptide chains of the protein.

Study Questions

Answer A if 1, 2, and 3 are correct
B if 1 and 3 are correct
C if 2 and 4 are correct
D if only 4 is correct
E if all are correct

1.1 Which of the following statements describe(s) the side chain of the amino acid serine?

1. Contains a hydroxyl group
2. Can form disulfide bonds
3. Can participate in hydrogen bonds
4. Is charged at physiologic pH

1.2 Which of the following amino acids has/have a charged side chain at physiologic pH?

1. Aspartic acid
2. Lysine
3. Glutamic acid
4. Asparagine

1.3 Glutamine

1. contains three titratable groups.
2. contains an amide group.
3. is classified as an acidic amino acid.
4. contains a side chain that can form hydrogen bonds in proteins.

1.4 Which of the following statements about amino acids is(are) true?

1. Glycine contains two dissociable hydrogens.
2. Tyrosine is a site of attachment of phosphate groups in proteins.
3. Cysteine is a sulfur-containing amino acid.
4. Glutamine is classified as a basic amino acid.

III. AMINO ACIDS AS BUFFERS

Amino acids contain weakly acidic α-carboxyl groups and weakly basic α-amino groups. In addition, each of the acid and basic amino acids contains an ionizable group in its side chain. Thus, both free amino acids and amino acids combined in peptide linkage can potentially act as buffers. The quantitative relation between $[H^+]$ and weak acids is described by the Henderson-Hasselbalch equation.

A. The Henderson-Hasselbalch equation

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

$$\underset{\substack{\text{weak}\\\text{acid}}}{HA} \rightleftarrows \underset{\text{proton}}{H^+} + \underset{\substack{\text{salt form}\\\text{or conjugate base}}}{A^-}$$

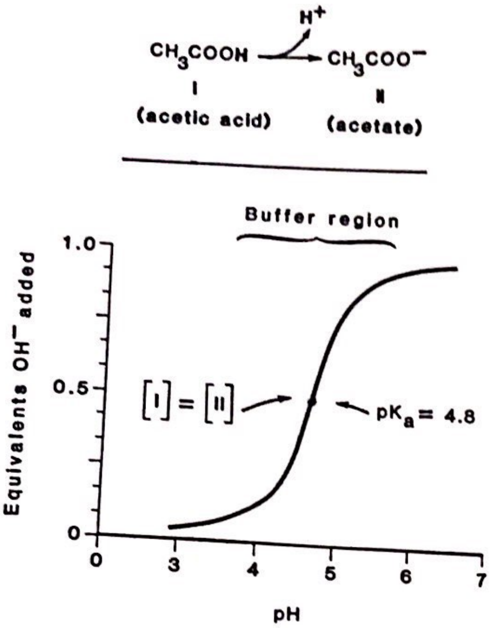


Figure 1.4
Titration curve of acetic acid.

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

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

By solving for the $[H^+]$ in the above equation, taking the logarithm of both sides of the equation, multiplying both sides of the equation by -1, and substituting $pH = -\log[H^+]$ and $pK_a = -\log K_a$, we obtain the Henderson-Hasselbalch equation:

$$pH = pK_a + \log\frac{[A^-]}{[HA]}$$

which can also be written as

$$pH = pK_a + \log\frac{[\text{conjugate base}]}{[\text{acid form}]}$$

2. The Henderson-Hasselbalch equation can be used to calculate the pH of a solution containing a weak acid after the addition of strong acid or base. For example, Figure 1.4 illustrates the change in pH that occurs when NaOH is added to a solution of acetic acid. This titration curve demonstrates several important concepts:
   a. A ***buffer*** is a solution that resists changes in pH following the addition of strong acid or base. A weak acid and its conjugate base, for example, acetic acid (CH_3—COOH) and acetate (CH_3—COO^-), can serve as a buffer when the pH of a solution is within ±1 pH unit of the pK_a of the weak acid. Maximum buffering capacity occurs at the pK_a. Therefore, a solution containing the acetic acid/acetate buffer pair will resist a change in pH from pH 3.8 to 5.8, with maximum buffering at pH 4.8.
   b. When the pH of a solution is equal to the pK_a, the amount of the salt form is equal to the acid form. As noted above, at this pH the solution will have its maximum buffering capacity.
   c. At pH values less than the pK_a, the protonated acid form (e.g., CH_3—COOH) is the predominant species.
   d. At pH values greater than the pK_a, the deprotonated salt (e.g., CH_3—COO^-) is the predominant species in solution.

B. The titration of alanine

1. The titration curve of an amino acid can be analyzed in the same way as that described above for acetic acid. For example, consider the amino acid alanine, which contains both a carboxyl and an amino group. At an acidic pH, both of these groups of alanine are protonated (shown in Figure 1.5). The —COOH group of form I can dissociate by donating a proton to the medium, where it binds to a water molecule, producing H_3O^+. This release of a proton results in the formation of the carboxylate group, —COO^-. The structure is shown as form II, which is the dipolar form of the molecule (Figure 1.5). (Note: This form is also called a ***zwitterion***, and is the ***isoelectric form*** of alanine.

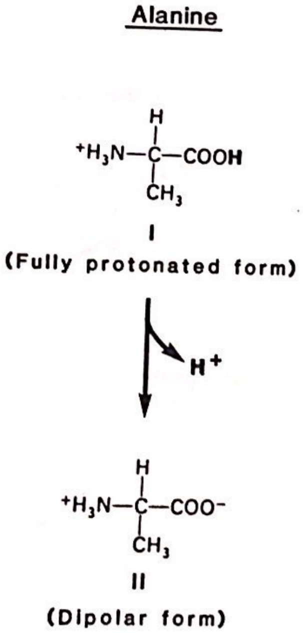


Figure 1.5
Dissociation of the carboxyl group of alanine.

2. The dissociation constant of the carboxyl group is defined as K_1, rather than K_a, since the molecule contains a second titratable group. The Henderson-Hasselbalch equation can be used to analyze the dissociation of the carboxyl group of alanine, as described for acetic acid.

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

This equation can be rearranged to yield

$$pH = pK_a + \log\frac{[II]}{[I]}$$

or

$$pH = pK_a + \log\frac{[\text{conjugate base}]}{[\text{acid form}]}$$

3. The second titratable group of alanine is the amino ($—NH_3^+$) group shown in Figure 1.6. This is a much weaker acid than the —COOH group and therefore has a much smaller dissociation constant, K_2. Release of a proton from the protonated amino group of form II results in the fully deprotonated form of alanine, structure III (Figure 1.6).

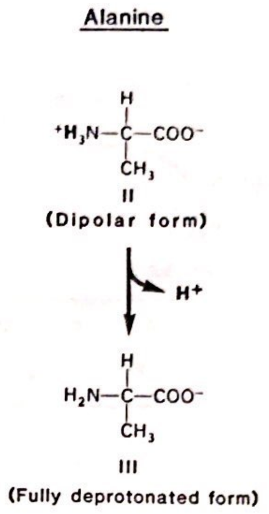


Figure 1.6
Dissociation of the protonated amino group of alanine.

4. The sequential dissociation of protons from the carboxyl and amino groups of alanine is summarized below (Figure 1.7). Each of the titratable groups has a pK that is numerically equal to the pH at which exactly half of the protons have been removed from that group. The pK for the most acidic group is pK_1 (—COOH); the pK for the next most acidic group ($—NH_3^+$) is pK_2.

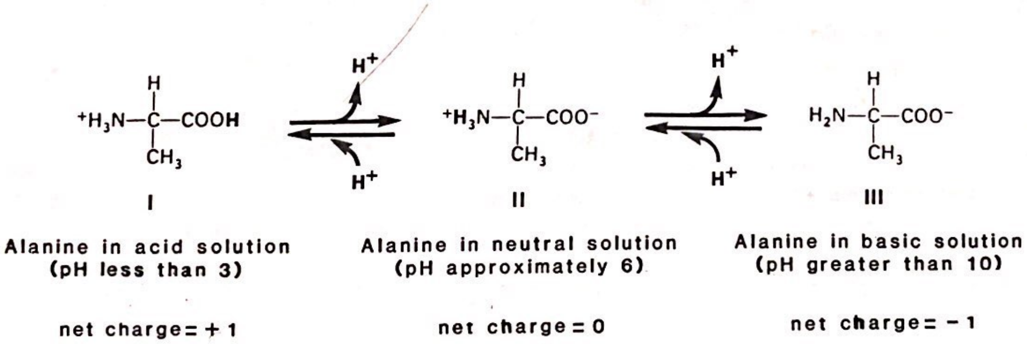


Figure 1.7
Ionic forms of alanine in acidic, neutral, and basic solutions.

5. By applying the Henderson-Hasselbalch equation to each dissociable acidic group, it is possible to calculate the complete titration curve. Figure 1.8 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:

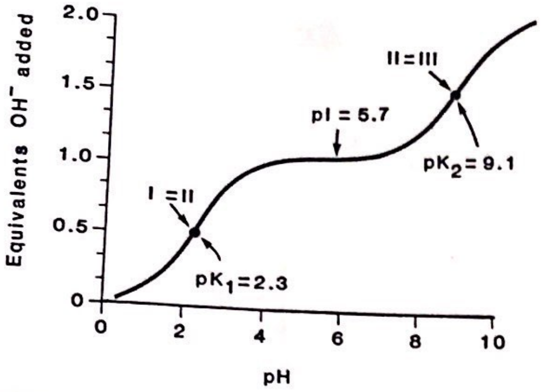


Figure 1.8
The titration curve of alanine.

   a. The —COOH/—COO$^-$ pair can serve as a buffer in the region around pK_1, and the —NH_3^+/—NH_2 pair buffers in the region around pK_2.

   b. 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. 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.

6. The ***isoelectric point*** (pI) is the pH at which a molecule is electrically neutral, that is, where the sum of the positive charges equals the sum of the negative charges.

   a. For an amino acid such as alanine, which 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, pI = (2.3 + 9.1)/2 = 5.7 (Figure 1.8). The pI is thus midway between pK_1 (2.3) and pK_2 (9.1). It corresponds to the pH at which structure II (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).

C. The titration of histidine

1. Histidine is an example of an amino acid that contains three chemical groups, each of which can reversibly gain or lose a proton: the α-carboxyl group, the imidazole group of the side chain, and the α-amino group (Figure 1.9).

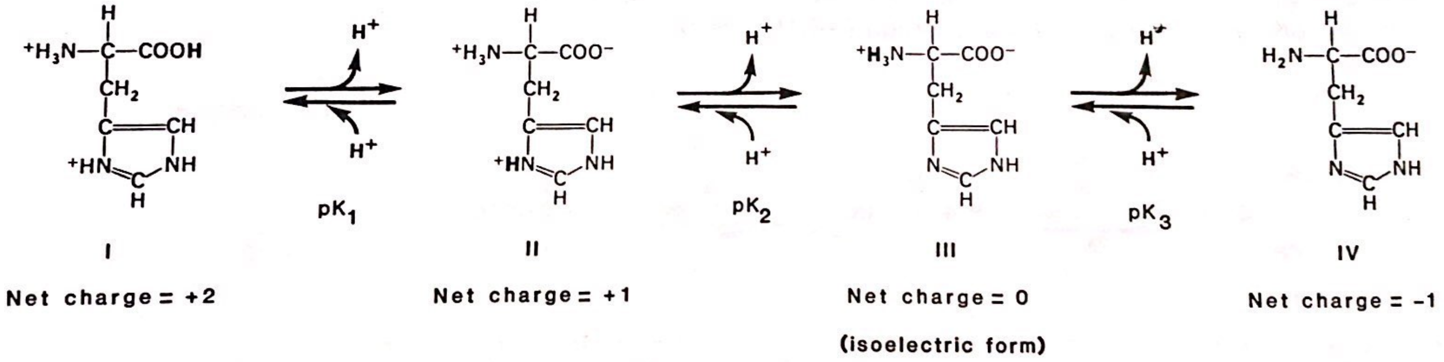


Figure 1.9
Ionic forms of histidine.

2. The R-group of histidine has a pK of 6.0 and can serve as a buffer at physiologic pH.

3. The incremental addition of base to fully protonated histidine results in the sequential removal of protons from the α-carboxyl group (pK_1 = 1.8), the imidazole group (pK_2 = 6.0), and the α-amino group (pK_3 = 9.2). The titration curve is shown in Figure 1.10.

4. The isoelectric point (pI) for histidine is calculated by first identifying the isoelectric form (which has a net charge of zero) of

the amino acid (III in Figure 1.9), then averaging the values of the nearest pK's ($(pK_2 + pK_3)/2 = (6.0 + 9.2)/2 = 7.6$) (Figure 1.10). This calculation requires that the isoelectric species be identified. This in turn depends on knowing the order in which the protons are lost from a particular amino acid, since this determines the charges of the intermediate species. The pK values for some of the 20 commonly encountered amino acids are shown in Figure 1.2.

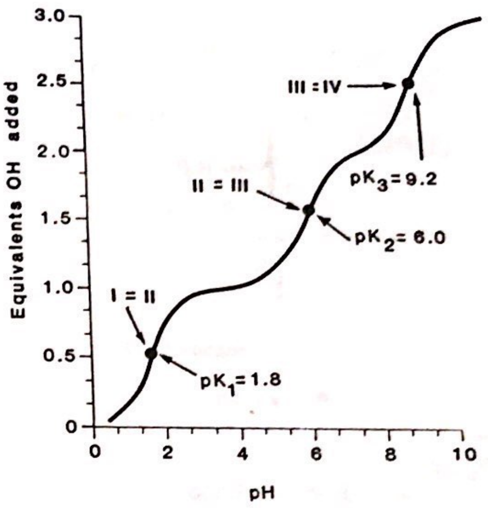


Figure 1.10
Titration of histidine.

Study Questions

Choose the ONE best answer:

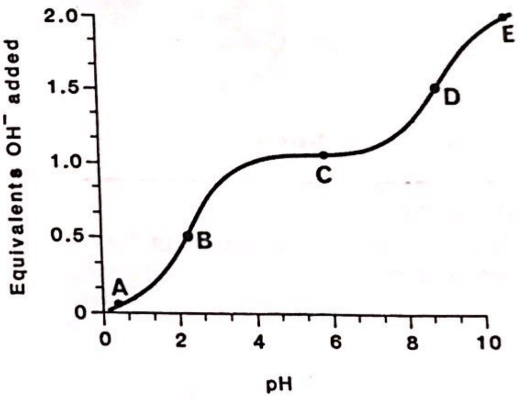
1.5 Lysine ($pK_1 = 2.2$, $pK_2 = 9.0$, $pK_3 = 10.5$)

A. contains an amide group. + ε
B. has a charge of +1 when the carboxyl group is protonated. + +2
C. has a pI of 5.6 ×
D. will migrate toward the cathode ($\overset{-ve}{\text{negative}}$ electrode) during electrophoresis at pH 7.0.
E. has a nonpolar side chain.

1.6 Histidine has a side chain that has a pK_2 of 6.0. Which one of the following amino acids also has a side chain that titrates within about 1.5 pH units of neutrality?

A. Lysine
B. Arginine
C. Threonine
D. Cysteine
E. Hydroxyproline

1.7 The letters A, B, C, D, and E designate regions on the titration curve of glycine ($pK_1 = 2.3$, $pK_2 = 9.6$) shown to the right. Which one of the following descriptions is correct?



A. A is a region of maximal buffering. ×
B. B is a region of minimal buffering. ×
C. C is at the isoelectric point (pI) ✓
D. The concentration of $^+NH_3—CH_2—COOH$ is maximal in region D. ×
E. The concentration of $^+NH_3—CH_2—COO^-$ is maximal in region E. ✓

IV. OPTICAL PROPERTIES OF AMINO ACIDS ✓

A. The α-carbon of each amino acid is attached to four different chemical groups and is therefore a ***chiral*** or ***optically active*** carbon atom. Glycine is the single exception since the α-carbon of glycine has two hydrogen substituents and therefore is optically inactive (see p. 46 for a discussion of optical activity).

B. All the amino acids found in proteins are of the L-configuration. However, D-amino acids are found in some antibiotics and in bacterial cell walls.

V. PEPTIDE BOND

A. In proteins, amino acids are joined covalently by peptide bonds that are amide linkages between the α-carboxyl group of one amino acid and the α-amino group of another. For example, glycine and

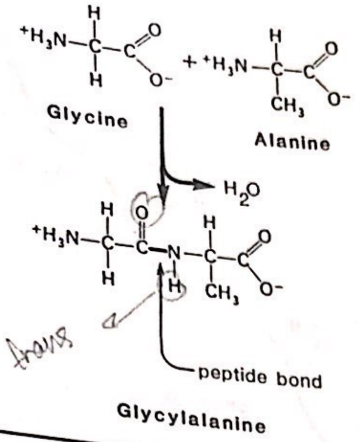


Figure 1.11
Formation of a peptide bond.

alanine can form the dipeptide glycylalanine through the formation of a peptide bond (Figure 1.11). By convention the free amino-end (N-terminal) is written to the left and the free carboxyl-end (C-terminal) to the right. Therefore, all amino sequences are read from the N- to the C-terminal end of the peptide. For example, in Figure 1.11, the order of the amino acids is "glycine, alanine" ***not*** "alanine, glycine".

B. 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***. When the compounds are named,all amino acid residues have the end of their names changed from -ine to -yl, except the C-terminal amino acid. For example, a tripeptide composed of N-terminal valine, a central glycine, and C-terminal leucine is called valylglycylleucine.

C. The peptide bond has a ***partial double-bond character*** and therefore is rigid and planar (Figure 1.12); this prevents free rotation around the bond between the carbonyl carbon and the nitrogen of the peptide bond.

D. Like all amide linkages, the peptide bond neither accepts nor gives off protons over the pH range of 2 to 12. Thus, the charged groups present in polypeptides consist solely of the N-terminal amino group, the C-terminal carboxyl group, and any ionized groups present in the side chains of the constituent amino acids.

E. Peptide bonds are not broken by normal handling or by conditions (such as heating alone or high concentrations of urea) that denature proteins. Strong acid or base at elevated temperatures is required to hydrolyze these bonds.

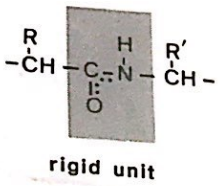


Figure 1.12
Partial double bond character of the peptide bond.

Study Questions

Answer A if 1, 2, and 3 are correct
B if 1 and 3 are correct
C if 2 and 4 are correct
D if only 4 is correct
E if all are correct

1.8 The peptide bond

1. is a special type of amide linkage. —
2. has a partial double bond character.—
3. is not ionized at physiologic pH.—
4. is cleaved by agents that denature proteins, such as organic solvents and high concentrations of urea. ×

1.9 The peptide alanylglycylserine

1. contains alanine with a free α-amino group. —
2. contains three peptide bonds. × tripep → two
3. is optically active.—
4. is a dipeptide. × —