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.

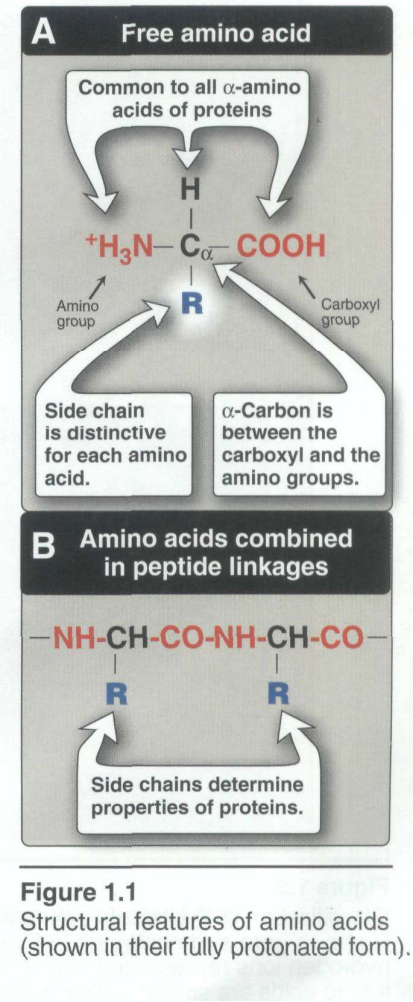

Figure 1.1
Structural features of amino acids (shown in their fully protonated form).

II. STRUCTURE OF THE AMINO ACIDS

Although more than 300 different amino acids have been described in nature, only twenty 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. 393).] Each amino acid (except for proline, which is described on p. 4) has a **carboxyl group,** an **amino group,** and a distinctive side chain **("R-group")** bonded to the α-carbon atom (Figure 1.1 A). At physiologic pH (approximately pH = 7.4), the carboxyl group is dissociated, forming the negatively charged carboxylate ion ($-COO^-$), and the amino group is protonated ($-NH_3^+$). In proteins, almost all of these carboxyl and amino groups are combined in 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 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 (that is, have an even distribution of electrons) or polar (that is, 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 bind or give off protons or participate in hydrogen or ionic bonds (see 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 2.9, p. 18).

1. **Location of nonpolar amino acids in proteins:** In proteins found in aqueous solutions, the side chains of the nonpolar amino acids tend to cluster together in the interior of the protein (Figure 1.4). This phenomenon is the result of the hydrophobicity of the nonpolar

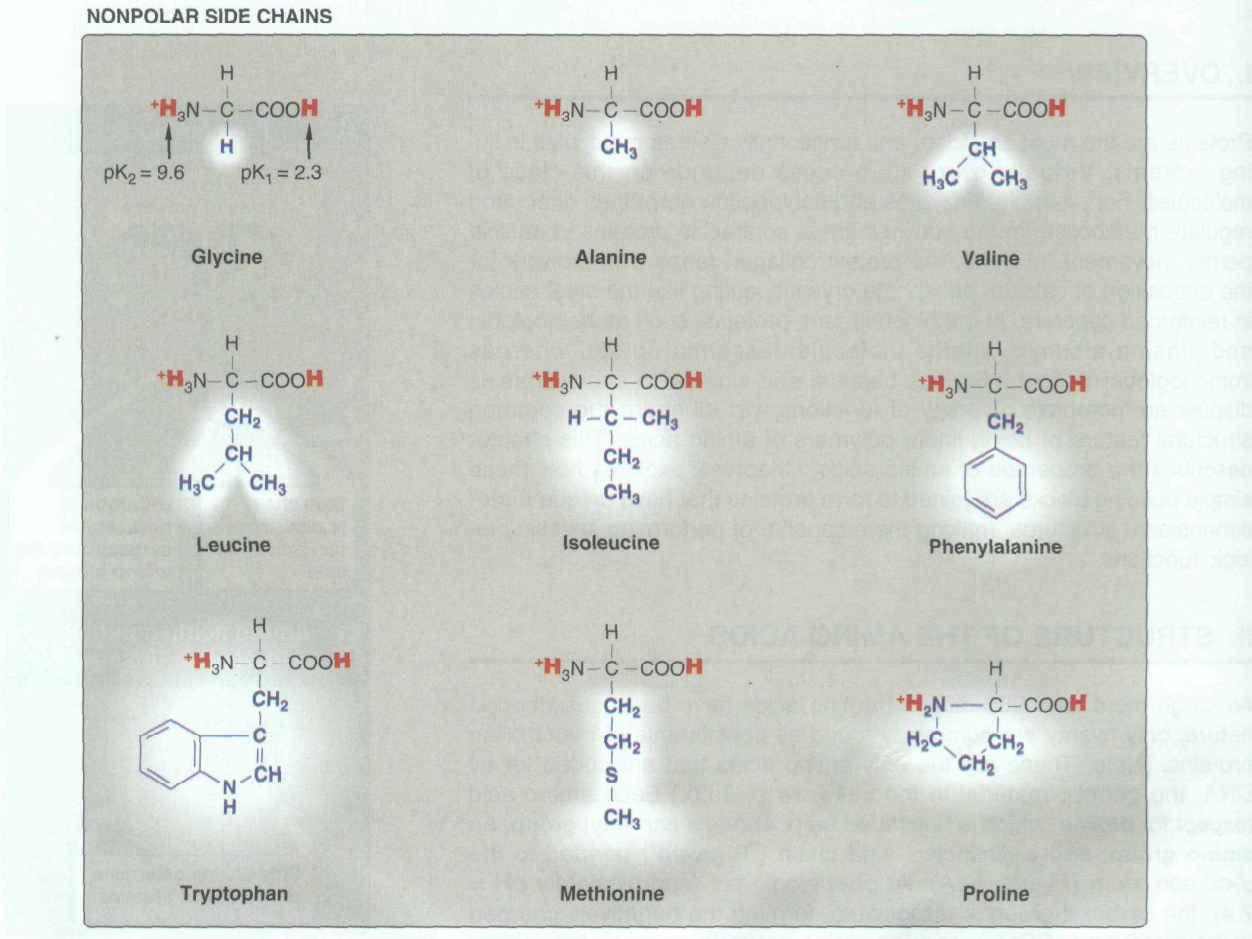

Figure 1.2
Classification of the twenty amino acids found in proteins, according to the charge and polarity of their side chains, is shown here and continues in Figure 1.3. Each amino acid is shown in its fully protonated form, with dissociable hydrogen ions represented in red print. The pK values for the α-carboxyl and α-amino groups of the nonpolar amino acids are similar to those shown for glycine. (Continued on Figure 1.3.)

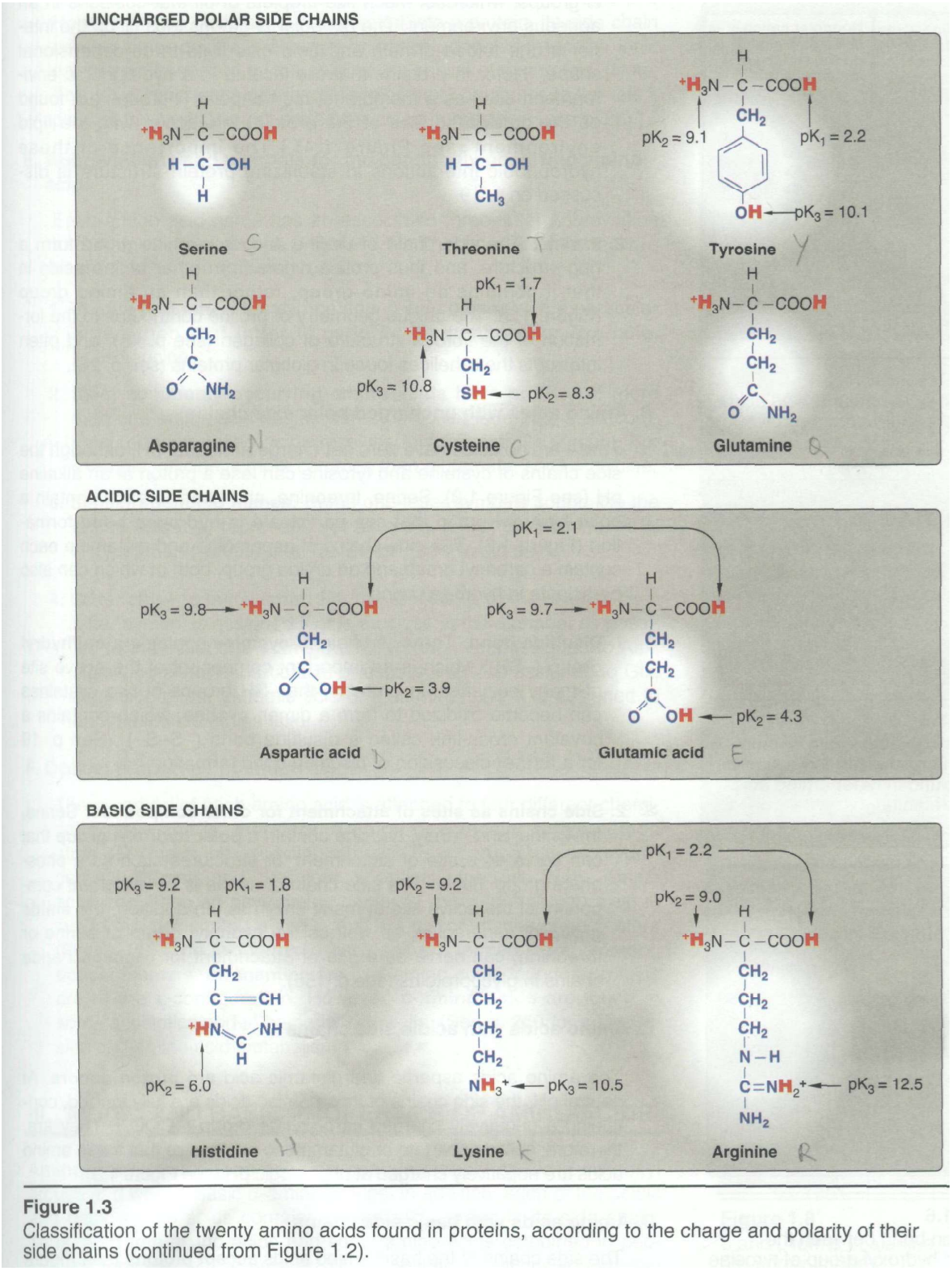

Figure 1.3
Classification of the twenty amino acids found in proteins, according to the charge and polarity of their side chains (continued from Figure 1.2).

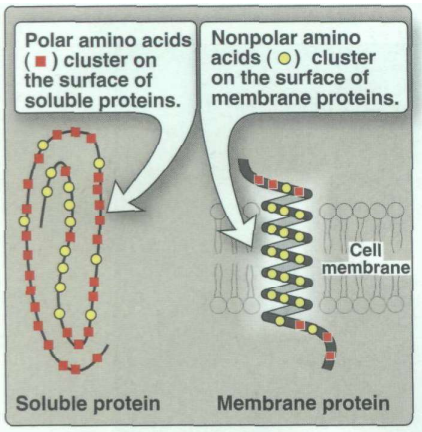

Figure 1.4
Location of nonpolar amino acids in soluble and membrane proteins.

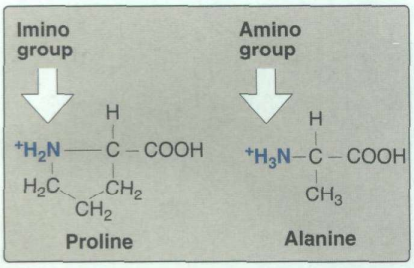

Figure 1.5
Comparison of the imino group found in proline with the α-amino group found in other amino acids, such as alanine.

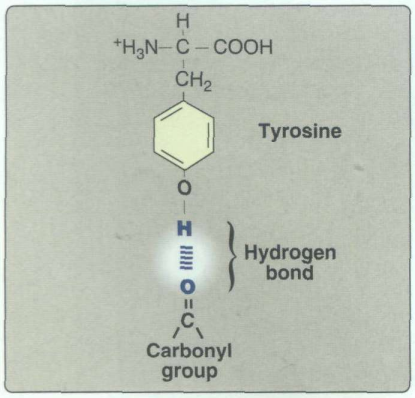

Figure 1.6
Hydrogen bond between the phenolic hydroxyl group of tyrosine and another molecule containing a carbonyl group.

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. [Note: In 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.

2. **Proline:** The side chain of proline and its α-amino group form a ring structure, and thus proline differs from other amino acids in that it contains an **imino group,** rather than an amino group (Figure 1.5). 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.)

2. **Side chains as sites of attachment for other compounds:** Serine, threonine, and, rarely, tyrosine contain a **polar hydroxyl group** that can serve as a site of attachment for structures such as a phosphate group. [Note: The side chain of serine is an important component of the active site of many enzymes.] 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. 156).

C. Amino acids with acidic side chains

The amino acids aspartic and glutamic acid are **proton donors.** At neutral pH, the side chains of these amino acids are fully ionized, containing a negatively charged **carboxylate group** (-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. [Note: This is an important property of histidine that contributes to the role it plays in the functioning of proteins such as hemoglobin (see p. 26).]

E. Abbreviations and symbols for the 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 of the amino acid. Further, 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.

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	("glutEmate")
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.

F. Optical properties of amino acids

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 exception because its α-carbon has two hydrogen substituents and, therefore, is optically inactive. [Note: 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 bacterial cell walls. (See p. 250 for a discussion of D-amino acid metabolism.)

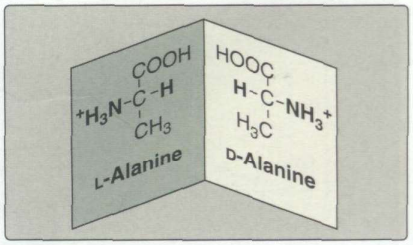

Figure 1.8
D and L forms of alanine are mirror images.

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.** The quantitative relationship between the concentration of a weak acid (HA) and its conjugate base (A^-) is described by the **Henderson-Hasselbalch equation.**

A. Derivation of the equation

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

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

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

$$K_a \qquad [HA]$$

[Note: The larger the K_a, the stronger the acid, because most of the HA has been converted 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 $[H^+]$ 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:

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

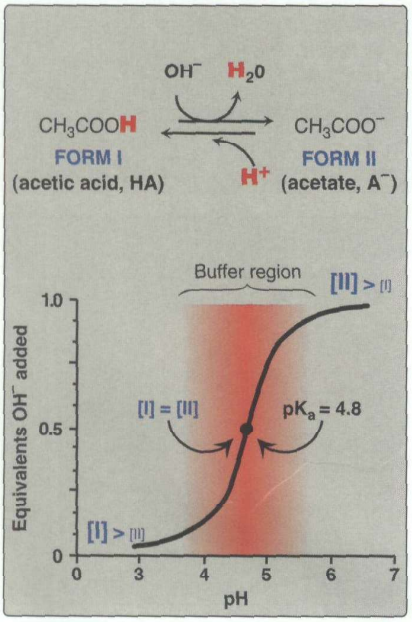

Figure 1.9
Titration curve of acetic acid.

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 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 pK_a, but a conjugate acid/base pair can still serve as an effec tive buffer when the pH of a solution is within approximately + 1 pH unit of the pK_a. [Note: If the 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 = CH_3–COOH) and acetate (A^- = CH_3–COO^-) with a pK_a of 4.8 resists a change in pH from pH 3.8 to 5.8, with maxi mum buffering at pH = 4.8. [Note: At pH values less than the pK_a, the protonated acid form (CH_3–COOH) is the predominant species. At pH values greater than the pK_a, the deprotonated base form (CH_3–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