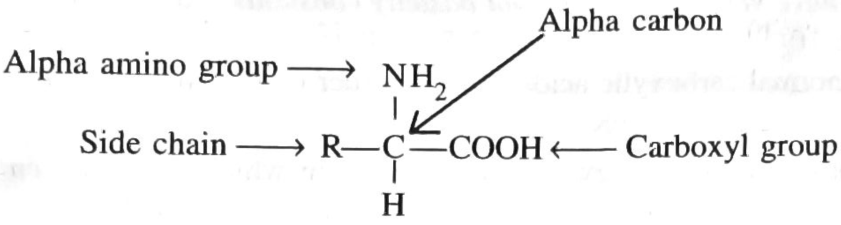
AMINO ACIDS, PEPTIDES AND PROTEINS

3.1. INTRODUCTION

Amino acids, as the name suggests are compounds which contain both an amino group (–NH_2) and a carboxyl group (–COOH) in their molecules. In principle, the amino group can be linked to any carbon atom of the chain. Depending upon the position of amino group, *i.e.*, α-, β-, γ-, etc. on the carbon chain *w.r.t.* the carboxyl group they are classified as α-, β-, γ-, etc. amino acids respectively. Out of these, α-amino acids (*i.e.*, NH_2 group in linked to the carbon atom next to be –COOH group, *i.e.*, α-carbon) are the most important because they are constituents of proteins which are very important biomolecules. The general formula of α-amino acids is



$$\mathrm{R{-}\overset{\displaystyle NH_2}{\underset{\displaystyle H}{C}}{-}COOH}$$

✓ ①

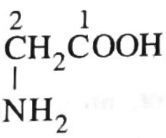
Twenty six different α-amino acids have been isolated by the hydrolysis of various proteins.

Out of these 20 amino acids occur in almost all proteins while the remaining 6 are found in special tissues. These amino acids differ from one another in the nature of the side chain groups R. Therefore, properties of side chain groups R determine the properties of amino acids and the properties of proteins they constitute.

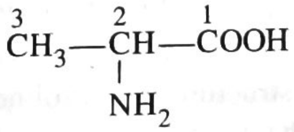
In this unit, we will first describe the chemistry of α-amino acids, then we will discuss as to how these α-amino acids combine to form peptides and proteins. This will be followed by structure elucidation of peptides and proteins at different levels.

3.2. NOMENCLATURE OF α-AMINO ACIDS

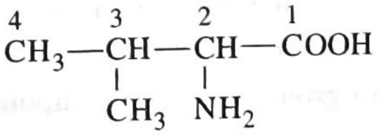
Although amino acids can be named by the IUPAC system, they are generally known by their common names. Their common names generally end in the suffix *'ine'*. Further, for sake of simplcity, each amino acid has been given a standard abbreviation or a code which usually consists of the first three letters of the common name. For example,



$$\mathrm{\underset{\displaystyle NH_2}{\overset{2}{C}H_2}\overset{1}{C}OOH}$$



$$\mathrm{\overset{3}{C}H_3{-}\underset{\displaystyle NH_2}{\overset{2}{C}H}{-}\overset{1}{C}OOH}$$



$$\mathrm{\overset{4}{C}H_3{-}\underset{\displaystyle CH_3}{\overset{3}{C}H}{-}\underset{\displaystyle NH_2}{\overset{2}{C}H}{-}\overset{1}{C}OOH}$$

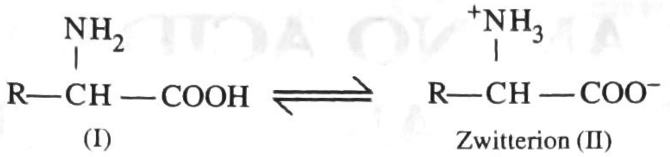
IUPAC name :	2-Aminoethanoic acid	2-Aminopropanoic acid	2-Amino-3-methylbutanoic acid
Common name :	Glycine	Alanine	Valine
Abbreviation	Gly	Ala	Val
One letter code	G	A	V

Pg-①

3.3. STRUCTURE OF α-AMINO ACIDS

(05) Zwitterion Structure

α-Amino acids are generally represented as compounds containing a basic amino group and an acidic carboxyl group as shown in the general formula (I). However, in actual practice, they neutralize each other involving the transfer of a proton from the carboxyl group to the amino group within the molecule. As a result, amino acids largely exist as dipolar ions (II).



$$\underset{\text{(I)}}{R-\underset{}{\overset{NH_2}{\overset{|}{C}H}}-COOH} \rightleftharpoons \underset{\text{Zwitterion (II)}}{R-\overset{^+NH_3}{\overset{|}{C}H}-COO^-}$$

The dipolar structure is commonly known as **internal salt** or **zwitterion.**

3.3.1. Evidence in favour of the dipolar ionic structure

The absence of free amino and carboxyl groups and the existence of dipolar ionic structure is supported by the following factors :

(*i*) *They are soluble in water, but insoluble in common organic solvents like benzene, alcohol, etc.*

(*ii*) *They have got very high dipole moments, 10–15 D which indicate considerable ionic or dipolar character.* Besides this, they raise the dielectric constant of water when dissolved in it–a property characteristic only of polar molecules.

(*iii*) *Neutral amino acids have very low acidity and basicity constants.* For example, the value of K_a and K_b for glycine are $1{\cdot}6 \times 10^{-10}$ ($pK_a = 9{\cdot}8$) and $2{\cdot}5 \times 10^{-12}$ ($pK_b = 11{\cdot}6$) respectively. It may be mentioned here that K_a for normal carboxylic acids is of the order of 10^{-5} (pK_a around 5) while K_b for most aliphatic amines is of the order of 10^{-4} (pK_b around 4).

These values are in accordance with the zwitterion structure in which the *acid centre in α-amino acids is an ammonium ion* ($-\overset{+}{N}H_3$) *instead of a free carboxyl group and the basic centre is the carboxylate* ion ($-COO^-$) instead of a free amino group.

(*iv*) *Infrared and Raman spectra of the neutral solutions indicate the presence of* $-\overset{+}{N}H_3$ *and* $-COO^-$ *and not* $-NH_2$ *and* $-COOH$ *groups.*

3.4. CLASSIFICATION OF AMINO ACIDS

As already stated, depending upon the position of the NH_2 group *w.r.t.* the carboxyl group, amino acids have been broadly classified as α-, β-, γ- amino acids, etc. For example,

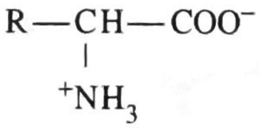
$$H_2N\underset{\alpha}{\overset{2}{C}}H_2\overset{1}{C}OOH \qquad H_2N\underset{\beta}{\overset{3}{C}}H_2\underset{\alpha}{\overset{2}{C}}H_2\overset{1}{C}OOH \qquad H_2N\underset{\gamma}{\overset{4}{C}}H_2\underset{\beta}{\overset{3}{C}}H_2\underset{\alpha}{\overset{2}{C}}H_2\overset{1}{C}OOH$$

2-Aminoethanoic acid (*α-Aminoacetic acid*)

3-Aminopropanoic acid (*β-Aminopropionic acid*)

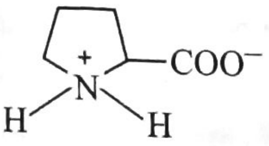
4-Aminobutanoic acid (*γ-Aminobutyric acid*)

However, all the naturally occurring amino acids are α-amino acids containing at least one primary amino group and having dipolar ionic structure (II). Proline is, however, an exception which contains a secondary amino group ; the N-atom of the 2° amino group and α-carbon being a part of the pyrrolidine ring.



$$R-\underset{\underset{^+NH_3}{|}}{CH}-COO^-$$

Dipolar ionic structure (II)



Proline

②

3.4.1. Classification of α-Amino Acids

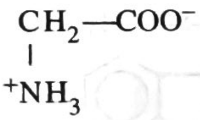
The α-amino acids listed in table 3.1 can be further classified as discussed below :

1. Essential and non-essential α-amino acids. Amino acids have also been classified as essential and non-essential amino acids. Out of the 20 amino acids required for protein synthesis, human body can synthesize only 10. These ten amino acids which the body can synthesize are called **non-essential or dispensable amino acids** *while the remaining ten which the human body cannot synthesize are called* ***essential or indispensable amino acids.*** These essential amino acids are required for the growth of the body and their deficiency causes diseases such as *Kwashiorkor*. Therefore, these must be supplied in the human diet.*

The ten essential amino acids are : valine, leucine, isoleucine, phenylalanine, methionine, tryptophan, threonine, lysine, arginine and histidine. These have been marked by asterisks in the Table 3.1

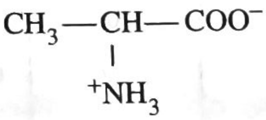
2. Classification according to neutral, acidic and basic character. α-Amino acids can also be classified according to the number of amino and carboxyl groups. For example,

(*i*) **Neutral amino acids.** These are amino acids which contain one $-NH_2$ group and one $-COOH$ groups, *i.e.*,



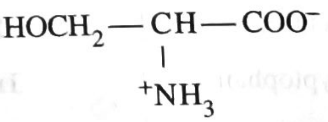
$\underset{\underset{{}^{+}NH_3}{|}}{CH_2}-COO^-$

Glycine



$CH_3-\underset{\underset{{}^{+}NH_3}{|}}{CH}-COO^-$

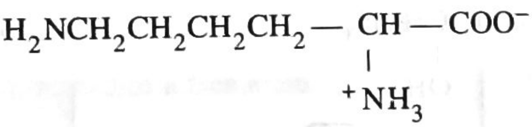
Alanine



$HOCH_2-\underset{\underset{{}^{+}NH_3}{|}}{CH}-COO^-$

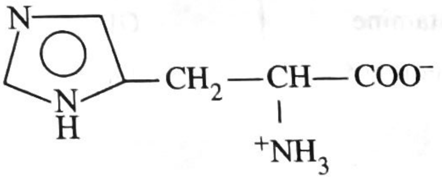
Serine

(*ii*) **Basic amino acids.** These are amino acids which contain two or more basic groups and one $-COOH$ group in their molecules. For example,



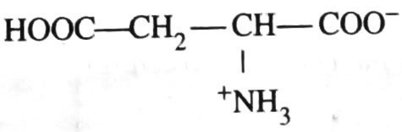
$H_2NCH_2CH_2CH_2CH_2-\underset{\underset{{}^{+}NH_3}{|}}{CH}-COO^-$

Lysine



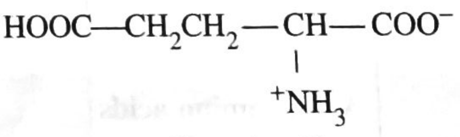
Histidine

(*iii*) **Acidic amino acids.** These are amino acids which contain one $-NH_2$ group and two $-COOH$ groups in their molecules. For example,



$HOOC-CH_2-\underset{\underset{{}^{+}NH_3}{|}}{CH}-COO^-$

Aspartic acid

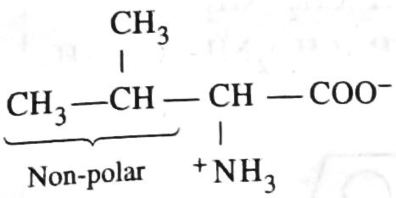


$HOOC-CH_2CH_2-\underset{\underset{{}^{+}NH_3}{|}}{CH}-COO^-$

Glutamic acid

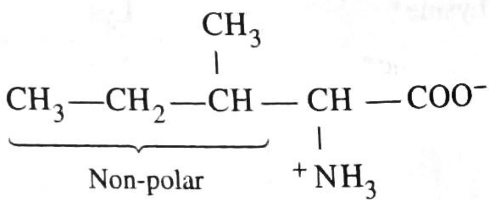
3. Classification according to nature of the side chain. According to this classification, amino acids have been divided into four categories :

(*i*) **Amino acids having non-polar side chain.** For example,



$\underbrace{CH_3-\overset{\overset{CH_3}{|}}{CH}}_{\text{Non-polar}}-\underset{\underset{{}^{+}NH_3}{|}}{CH}-COO^-$

Valine



$\underbrace{CH_3-CH_2-\overset{\overset{CH_3}{|}}{CH}}_{\text{Non-polar}}-\underset{\underset{{}^{+}NH_3}{|}}{CH}-COO^-$

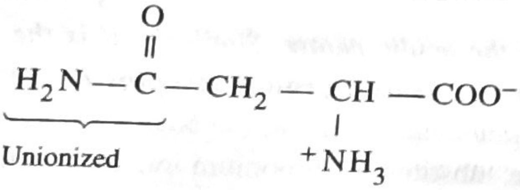
Isoleucine

---

*During this disease, the water balance in the body is distributed. As a result, some organs of body become watery and bloated.

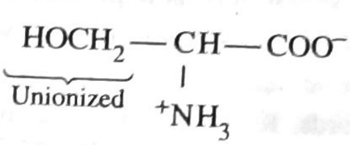
③

(*ii*) **Amino acids having polar but unionised side chains.** For example,



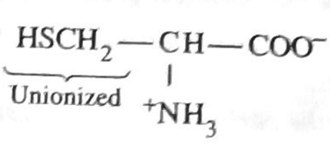
$$\underbrace{H_2N-\overset{\overset{\displaystyle O}{\|}}{C}-CH_2}_{\text{Unionized}}-\underset{\underset{{}^{+}NH_3}{|}}{CH}-COO^-$$

Asparagine



$$\underbrace{HOCH_2}_{\text{Unionized}}-\underset{\underset{{}^{+}NH_3}{|}}{CH}-COO^-$$

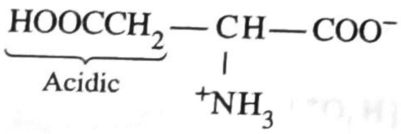
Serine



$$\underbrace{HSCH_2}_{\text{Unionized}}-\underset{\underset{{}^{+}NH_3}{|}}{CH}-COO^-$$

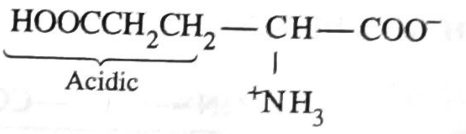
Cysteine

(*iii*) **Amino acids having acidic side chains.** For example,



$$\underbrace{HOOCCH_2}_{\text{Acidic}}-\underset{\underset{{}^{+}NH_3}{|}}{CH}-COO^-$$

Aspartic acid

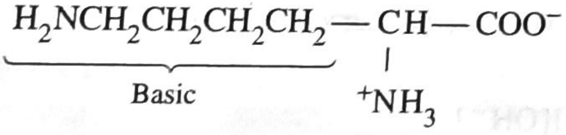


$$\underbrace{HOOCCH_2CH_2}_{\text{Acidic}}-\underset{\underset{{}^{+}NH_3}{|}}{CH}-COO^-$$

Glutamic acid

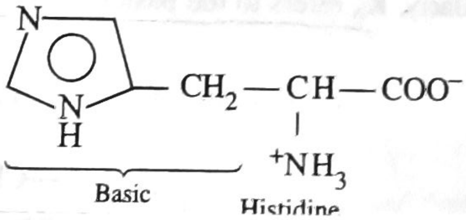
4

(*iv*) **Amino acids having basic side chains.** For example,



$$\underbrace{H_2NCH_2CH_2CH_2CH_2}_{\text{Basic}}-\underset{\underset{{}^{+}NH_3}{|}}{CH}-COO^-$$

Lysine

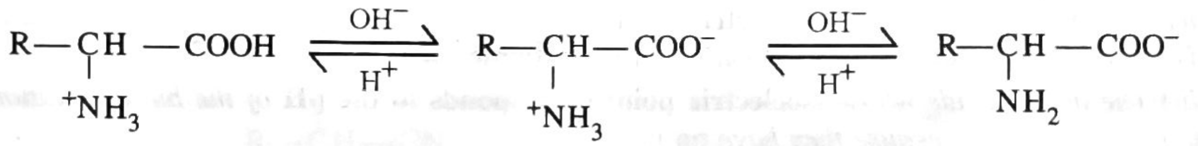


Histidine

3.7. ISOELECTRIC POINT OF α-AMINO ACIDS : ELECTROPHORESIS

When an electric current is passed through the aqueous solution of an amino acid, its behaviour depends upon the *p*H (whether acidic or basic) of the solution.

We have discussed above that in aqueous solution of an α-amino acid, there exists an equilibrium between the dipolar ion (I), cationic (II) and anionic (III) forms of the amino acid.



$$R-\underset{{}^{+}NH_3}{CH}-COOH \underset{H^+}{\overset{OH^-}{\rightleftharpoons}} R-\underset{{}^{+}NH_3}{CH}-COO^- \underset{H^+}{\overset{OH^-}{\rightleftharpoons}} R-\underset{NH_2}{CH}-COO^-$$

Cationic form (II) predominates in strongly acidic solution

Dipolar ion (I)

Anionic form (III) predominates in strongly basic solution

The exact position of the equilibrium depends upon the *p*H of the solution. If the solution is strongly acidic (*i.e.*, low *p*H), the amino acid exists predominantly in the cationic form (II) and thus migrates towards the cathode in the electric field. On the other hand, in strongly basic solution, it predominantly exists in the anionic form (III) and thus migrates towards anode in the electric field. Therefore, at some intermediate value of *p*H, the concentration of the cationic form (II) and the anionic form (III) are minimum and equal to each other. In other words, at this *p*H, the concentration of dipolar ionic form is the maximum *i.e.*, at this *p*H, the amino acid primarily exists as the neutral dipolar ion (I). In other words, at this *p*H, there will no net migration of the amino acid in the electric field. *This pH at which there is no net migration of the amino acid under the influence of the applied electric field is called the* **isoelectric point (pI)***.

---

*At isoelectronic point, sum of the charges of all the molecules of amino acids is zero.

(Contd.)

From the above discussion we conclude, that at a **pH lower than pI**, there will be a net migration of the amino acid towards **cathode** because the amino acid exists predominantly in the cationic form (II). Conversely, **at a pH higher than pI**, there will be a net migration of the amino acid towards **anode** because the amino acid mainly exists in the anionic form (III).

The isoelectric point* of an amino acid depends upon its structure and hence each amino acid has a characteristic isoelectric point (Table 3.1). For example, the isoelectric point of glycine is 6·0.

In general, **neutral amino acids,** *i.e.*, **monoamino monocarboxylic acids** such as glycine, alanine, etc. are slightly more acidic than basic**. When the crystal of such as amino acid is dissolved in water, the resulting solution contains more of the anion (III) than the cation (II). To suppress the excess ionization of (I) into the anion (III), we must add some acid to reach the isoelectric point. That is why, *isoelectric point of all such acids lies in slightly acidic range, i.e., around pH 6*. For example, isoelectric points of alanine, valine, leucine are 6·1, 6·0, 6·0 respectively.

For **basic amino acids,** *i.e.*, **diaminomonocarboxylic acids,** isoelectric point lies in the basic range (*p*H 7·6 – 10·8). For example, isoelectric points of lysine and arginine are 9·7 and 10·8 respectively.

For acidic amino acids, *i.e.*, **monoamino dicarboxylic acids**, the isoelectric point lies in the strongly acidic range (*i.e.* *p*H 3·2 – 3·5). For example, isoelectric points of aspartic acid and glutamic acids are 3·0 and 3·2 respectively.