Proteins- Introduction

- The word protein comes from the Greek ("prota"), meaning "of primary importance" and these molecules were first described and named by the Swedish chemist Jöns Jakob Berzelius in 1838. However, proteins' central role in living organisms was not fully appreciated until 1926, when James B. Sumner showed that the enzyme urease was a protein.
- The first protein to be sequenced was insulin, by Frederick Sanger, who won the Nobel Prize for this achievement in 1958. The first protein structures to be solved included hemoglobin and myoglobin, by Max Perutz and Sir John Cowdery Kendrew, respectively, in 1958. Both proteins’ three-dimensional structures were first determined by x-ray diffraction analysis; the structures of myoglobin and hemoglobin won the 1962 Nobel Prize in Chemistry for their discoverers.

Roles played by proteins

Proteins serve a wide range of functions including:
- Enzymatic catalysis - Enzymes are protein biological catalysts
- Hormonal regulation - Some hormones are proteins (e.g. Insulin, Somatotropin)
- Transport & Storage – e.g. Hemoglobin, Myoglobin, Albumin, Transferrin, Ferritin
- Mechanical support – e.g. collagen
- Immunoprotection – e.g. antibodies, interferon
- Systematic movement – e.g. Actin & Myosin
- Generation and transmission of nerve impulses – e.g. Receptor proteins in nerve cells, Rhodopsin (photoreceptor in retinal cells)
- Toxic proteins – in some plants (like mushroom) and animals (like snake venom)
Proteins

- Proteins are large complex molecules composed of long chain(s) of amino acids called polypeptide(s) folded into a specific 3-dimensional structure called native conformation.

- Proteins are polymers and amino acids are their monomers (building blocks).

- C,H,O, N & S are the elements found in the proteins.

THE AMINO ACID

The general formula of an amino acid is

\[
\text{H}_2\text{N-} \overset{\alpha}{\text{C}} -\text{COOH}
\]

\( \text{amino group} \)
\( \text{carboxyl group} \)
\( \text{side-chain group} \)

\( R \) is commonly one of 20 different side chains. At pH 7 both the amino and carboxyl groups are ionized.

\[
\text{H}_3\text{N-} \overset{+}{\text{C}} -\text{COO}^- \]

\( R \)
Amino Acids

- There are 20 amino acids that are common to all life forms.
- The number and arrangement of these 20 amino acids yields infinite variety of proteins.

Each of 20 different amino acids has different "R-Group," side chain attached to Ca.

Amino Acids

Amino Acids are the building units of proteins. Proteins are polymers of amino acids linked together by what is called “Peptide bond” (see latter).

- There are about 300 amino acids occur in nature. Only 20 (standard amino acids) of them occur in proteins.

Structure of amino acids:

Each amino acid has 4 different groups attached to α- carbon (which is C-atom next to COOH). These 4 groups are: amino group, COOH group, hydrogen atom and side chain (R). Cα is chiral (in all standard amino acids except in Glycine).
- At physiological PH (7.4), -COOH group is ionized (unprotonated) forming a -vely charged carboxylate (COO⁻) and amino gp is protonated forming +vely charged ion (NH₃⁺) forming Zwitterion (isoelectric) form.

- Proline is an imino acid (2° amine) not amino acid (1° amine)

### Classification of Amino Acids

#### I- Chemical classification:
According to the chemical structure of R-.

#### A- Neutral or uncharged:

1- **Glycine**  \( R = H \) (The smallest a. acid which is optically inactive since \( C_α \) is achiral)

2- **Alanine**  \( R = CH_3 \)

3- **Branched chain amino acids:** \( R \) is branched such as in:
   a- **Valine**  \( R = \) isopropyl
   b- **Leucine**  \( R = \) isobutyl → branching occurs on \( \gamma \)- carbon
   c- **Isoleucine**  \( R = \) isobutyl→ branching occurs on \( \beta \)- carbon (chiral)
   *(Isoleucine has two chiral carbons - \( C_α \) & \( C_β \) )

* **Leucine** & **Isoleucine** are structural isomers

---

#### 4- Neutral Sulfur containing amino acids:
**Cysteine** (has thiol group) and **Methionine** (has methyl thioether group)

*Cys* residues can form disulfide bond (important in stabilizing the protein structure)
5- **Neutral, hydroxy amino acids:**  
**Serine** and **Threonine** (The –OH is a site for phosphorylation)  
(Threonine has two chiral carbons - $C_\alpha$ & $C_\beta$ )

6- **Neutral aromatic amino acids:**

a- **Phenylalanine:** It is alanine in which one hydrogen of CH$_3$ is substituted with benzyl group.

b- **Tyrosine:** it is a hydroxylated phenylalanine (phenolic group which is a site for phosphorylation).

c- **Tryptophan:** it contains indole group.

- Only **Tyr** and **Trp** absorb ultraviolet (UV) light maximally at 280nm
- This property is employed for quantitative determination of proteins ($A_{280nm}$ [Protein])

7- **Neutral amino acid with cyclic R-:**

**Proline:** In proline, amino group enters in the ring formation being $\alpha$-imino group so proline is an $\alpha$-imino acid rather than $\alpha$-amino acid (secondary amine). Pro is usually found at bend points of protein sequences.

8- **Neutral amino acids with amide R-:**

**Aspargine and Glutamine:** They are amide forms of aspartate and glutamate in which side chain COOH groups are amidated.

**B- Basic amino acids:** Contain two or more NH$_2$ groups or nitrogen atoms that act as base i.e. can bind proton.

- At physiological pH, basic amino acids will be *vely charged.

a- **Lysine:** contains $\varepsilon$-amnio group

b- **Arginine:** contains guanidino group

c- **Histidine:** contains an imidazole group

**C- Acidic Amino acids:**

at physiological pH will be *vely charged.

a- **Aspartic acid (Aspartate)**

b- **Glutamic acid (Glutamate)**  
see structures in handout.
II- Classification according to polarity of side chain (R):

A- Polar amino acids: in which R contains polar hydrophilic group so can forms hydrogen bond with H₂O. In those amino acids, R may contain:
1- OH group : as in serine, threonine and tyrosine
2- SH group : as in cysteine
3- amide group: as in glutamine and aspargine
4- NH₂ group or nitrogen acting as a base (basic amino acids ): as lysine, arginine and histidine
5- COOH group ( acidic amino acids): as aspartic and glutamic acids.

B- Non polar amino acids:
R is alkyl hydrophobic group which can’t enter in hydrogen bond formation. 9 amino acids are non polar (glycine, alanine, tryptophan, proline, valine, leucine, isoleucine, phenylalanine, methionine).
* The underlined 5 a. acids have hydrophobic R-

III- Nutritional classification:
1- Essential amino acids: These amino acids can’t be formed in the body and so, it is essential to be taken in diet. Their deficiency affects growth, health and protein synthesis.
2- Semiessential amino acids: These are formed in the body but not in sufficient amount for body requirements especially in children.

Summary of essential and semiessential amino acids:
Valine Isoleucine Lysine Leucine Arginine* Histidine*
Methionine Tryptophan Threonine Phenylalanine
*= Arginine and Histidine are semiessential

3- Non essential amino acids: These are the rest of amino acids that are formed in the body in amount enough for adults and children. They are the remaining 10 amino acids.
Amphoteric properties of amino acids: that is they have both basic and acidic groups and so can act as base or acid. Neutral amino acids (monobasic, monocarboxylic) exist in aqueous solution as “Zwitterion” i.e. contain both positive and negative charge. Zwitter ion is electrically neutral (isoelectric) and can’t migrate in electric field.

**Isoelectric point (pI)** = is the pH at which the a.acid is isoelectric (has no net charge). e.g. pI of alanine is 6.0

* The net charge of an amino acid, a peptide or protein depends on its isoelectric point (pI) and the pH of the solution containing it.
  - if pH < pI the net charge is +ve.
  - if pH = pI the net charge is zero.
  - If pH > pI the net charge is -ve

---

### Ionization of amino acids

- All the 20 standard amino acid are diprotic or triprotic weak acids.
- The pKa values of the ionizable groups are shown in the table.
- The ionization of amino acid is started from the fully protonated form (cationic form) ending with the fully unprotonated form (anionic form)
- The isoelectric (zwitterion) form is the form having Zero net charge (equal no. of +ve and –ve charges)
- Migration of the different possible ionic forms in electric field depends on their net charge (cationic form towards cathode, anionic form towards anode and there will be no migration of isoelectric form neither towards cathode nor towards anode).

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>pKa value</th>
<th>Carboxyl group</th>
<th>Amino group</th>
<th>Side chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycine</td>
<td>2.4</td>
<td>9.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alanine</td>
<td>2.4</td>
<td>9.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valine</td>
<td>2.3</td>
<td>9.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leucine</td>
<td>2.3</td>
<td>9.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isoleucine</td>
<td>2.3</td>
<td>9.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methionine</td>
<td>2.1</td>
<td>9.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proline</td>
<td>2.0</td>
<td>10.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>2.2</td>
<td>9.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tryptophan</td>
<td>2.5</td>
<td>9.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serine</td>
<td>2.2</td>
<td>9.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threonine</td>
<td>2.1</td>
<td>9.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cysteine</td>
<td>1.9</td>
<td>10.7</td>
<td>8.4</td>
<td></td>
</tr>
<tr>
<td>Tyrosine</td>
<td>2.2</td>
<td>9.2</td>
<td>10.5</td>
<td></td>
</tr>
<tr>
<td>Asparagine</td>
<td>2.1</td>
<td>8.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glutamine</td>
<td>2.2</td>
<td>9.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>2.0</td>
<td>9.9</td>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>2.1</td>
<td>9.5</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>Lysine</td>
<td>2.2</td>
<td>9.1</td>
<td>10.5</td>
<td></td>
</tr>
<tr>
<td>Arginine</td>
<td>1.8</td>
<td>9.0</td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td>Histidine</td>
<td>1.8</td>
<td>9.3</td>
<td>6.0</td>
<td></td>
</tr>
</tbody>
</table>
Biochemistry 24311
Dr. Awni A. Abu-Hijleh

\[ \text{pH} = pK_a(\text{H}_2\text{COO}^-) + \log \frac{\text{H}_2\text{COO}^-}{\text{H}_2\text{CO}_2\text{H}} \]

\[ \text{pH} = pK_a(\text{H}_2\text{COO}^-) + \log \frac{\text{H}_2\text{COO}^-}{\text{H}_2\text{CO}_2\text{H}} \]

\[ \text{pI} (\text{Ala}) = \frac{pK_a(\text{H}_2\text{COO}^-) + pK_a(\text{H}_2\text{CO}_2\text{H})}{2} \]
Q. What is the pH of a glycine solution in which the α-NH$_3^+$ group is one-third dissociated?

**Answer**

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 1 part Gly$^-$ for every 2 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:

pH $= 9.6 + \log (1/2)$

pH $= 9.3$
Q. What is the pH of a glutamic acid solution if the alpha carboxyl is 1/4 dissociated?

• pH = 2 + log \( \frac{[1]}{[3]} \)
• pH = 2 + (-0.477)
• pH = 1.523

Q. What is the pH of a lysine solution if the side chain amino group is 3/4 dissociated?

• pH = 10.5 + log \( \frac{[3]}{[1]} \)
• pH = 10.5 + (0.477)
• pH = 10.977 = 11.0