COURSE CODE: BCH 211

COURSE TITLE: GENERAL BIOCHEMISTRY 1

Course Content: Chemistry of amino acids, proteins and their derivatives. Primary, Secondary, tertiary and quaternary structures of proteins. Methods of isolation and identification of isolation and identification of proteins; determination and biochemical application of these structure.

Chemistry of Amino Acids

Amino acids belong to a group of organic compounds containing a **basic amino group (-NH_2), an acidic carboxyl group (-COOH)** and an organic alkyl group as side chain as basic structural components. The amino acids are building monomers of proteins because they are **building blocks** of proteins. Amino acids may be **proteinogenic** (protein creating) or **non-proteinogenic** (non-protein creating). Some amino acids can be synthesized in the body (**non-essential**) while some cannot be synthesized (**Nutritionally essential**).

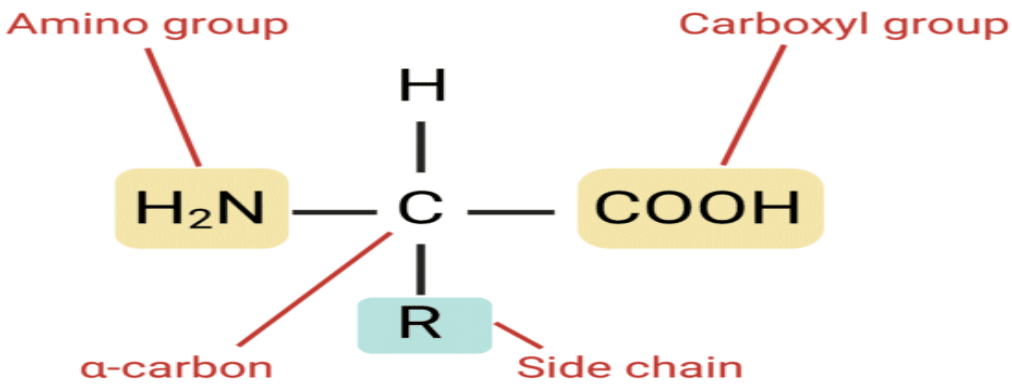


Fig 1: Basic structure of an amino acid

There are about 300 amino acids occurring in nature, each molecules contain a centrally placed carbon atom called **alpha carbon(α-carbon)** to which both **an amino group** and **a carboxylic group** are attached, while the remaining two bonds of the alpha carbon are generally attached to an hydrogen atom(H) and an alkyl group (R) or sides chain. The **R-group** contains additional carbon atoms in a chain which are designated **α, β, γ, δ, ε, ζ……**

General nature of Amino acids

At present, there are over 300 amino acids occurring in nature in various animal, plant and microbial systems. Twenty two(22) of these amino acids are genetically encoded i.e.

proteinogenic amino acids and twenty (20) out of these 22 amino acids are referred to as standard amino or primary or normal amino acids because they are repeatedly found in the structure of proteins. The remaining or additional two amino acids (i.e. **selenocysteine and pyrrolysine**) are also incorporated during special translation mechanism in some species. However some may proteins contained modified or unsual amino acids.

The structure of these standard amino acids are shown below;

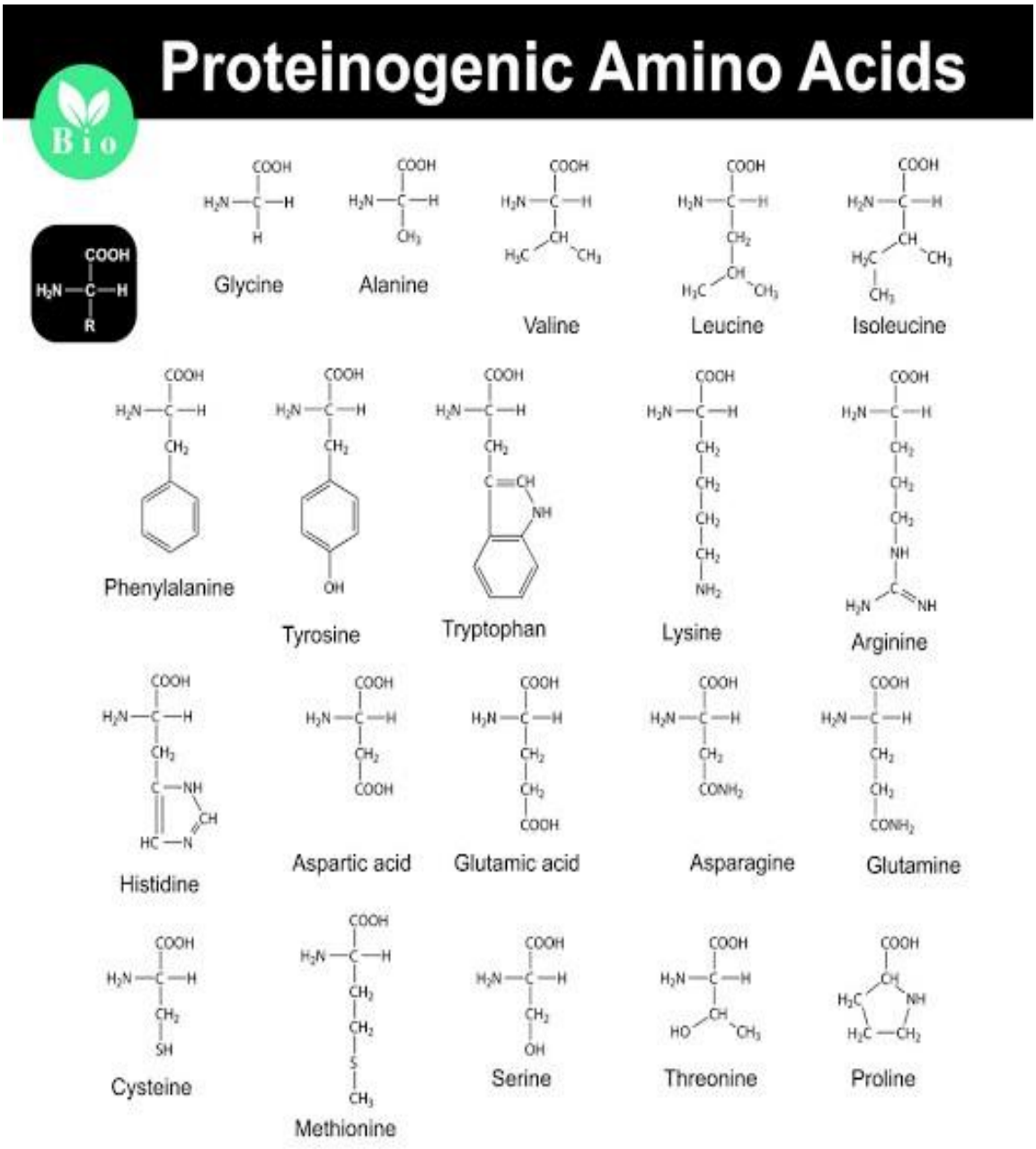


Fig 2: Structure of the standard amino acids

From the above structure, it is clear that all the amino acids differ from each other in their side chain that is attached to the alpha carbon causing variation in structure, size, and charges, thus influencing it solubility in water. Among all the 20 amino acids, **only proline contained imino group** instead of amino group.

Nomenclature and symbol of amino acids

The international Union of Biochemistry and Molecular Biology has officially designated **three letters abbreviation** and **one letter symbol** for easier identification and representation of all the proteinogenic amino acids. For example, amino acid Alanine was designated "**Ala"** as it three letters code representation and letter "**A"** as one letter code. The three letters abbreviation and one letter symbol for all the 20 amino acids are shown in the fig 3.

Amino Acid	Three Letter Code	One Letter Code
Alanine	Ala	A
Arginine	Arg	R
Aspartic Acid	Asp	D
Asparagine	Asn	N
Cysteine	Cys	C
Glutamic Acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Fig 3: Nomenclature and symbol of amino acids

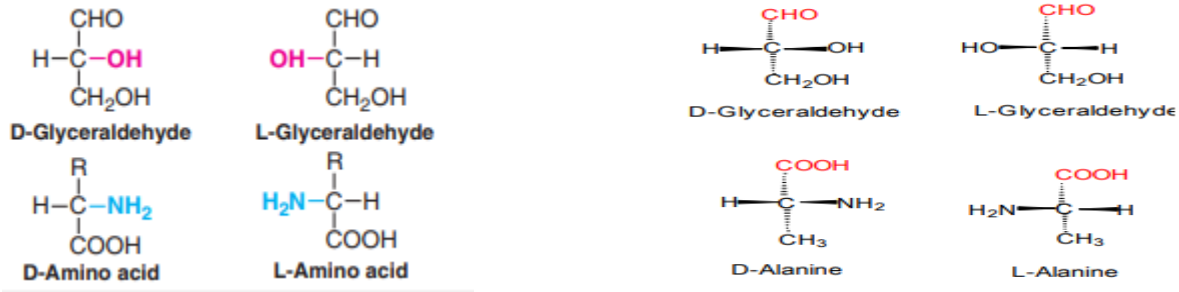
Chirality in amino acids

Amino acids except glycine exhibit chirality as a compound because they have at least one chiral carbon atom (i.e. a carbon atom to which four different functional groups are covalently linked. This is illustrated in fig 1. The carboxyl group (-COOH), amino group ($-NH_2$), hydrogen atom (H) and the variable side chain tagged as R, were the four different functional groups attached to the α-carbon.

Optical isomers of amino acids

Since of α-carbon of all the primary amino acids except glycine exhibit chirality, hence it is asymmetric and therefore, it is exhibit optical isomerism. Amino acids exhibit L and D configurations depending on how they rotates a plane polarized light either to left or right.

The L and D nomenclature system of configuration was designed in 1891 by Emil Fischer. As stated by Emil Fischer, any stereoisomers that have similar configuration to that of D-glyceraldehyde are designated L, while those configuration identical to L-glycceraldehyde are designated L. A typical illustration for this is shown below, done by simply aligning the carboxyl group (COOH) of the amino acid with the aldehyde group (CHO) of glyceraldehydes as illustrated for Alanine



A **B**

Fig 4: Designation of D and L forms of amino acid based on the structure of glyceraldehyde for a basic structure of amino acids and alanine

Note: All the amino acids found in proteins are exclusively of the L-configuration, but some D-amino acids are found in some micro-organism which expresses antibiotic, likewise **D-serine and D-aspartic acid** are also found in the fore brain and peripheral region of the brain.

Classification of amino acids

Generally, amino acids are classified on the basis of the following criteria

1) Chemical nature of the amino acid in the solution
2) Chemical structure of the side chain of the amino acids
3) Nutritional requirement of amino acids

4) Metabolic product of amino acids
5) Polarity of the side chain of the amino acids
6) Functional group of the side chain

A. Classification based on the chemical nature of the amino acid in the solution

On the basis of chemical nature of amino acids in solution, all amino acids are classified into three groups namely:

a) Acidic amino acids
b) Basic amino acids.
c) Neutral amino acids

Neutral amino acids

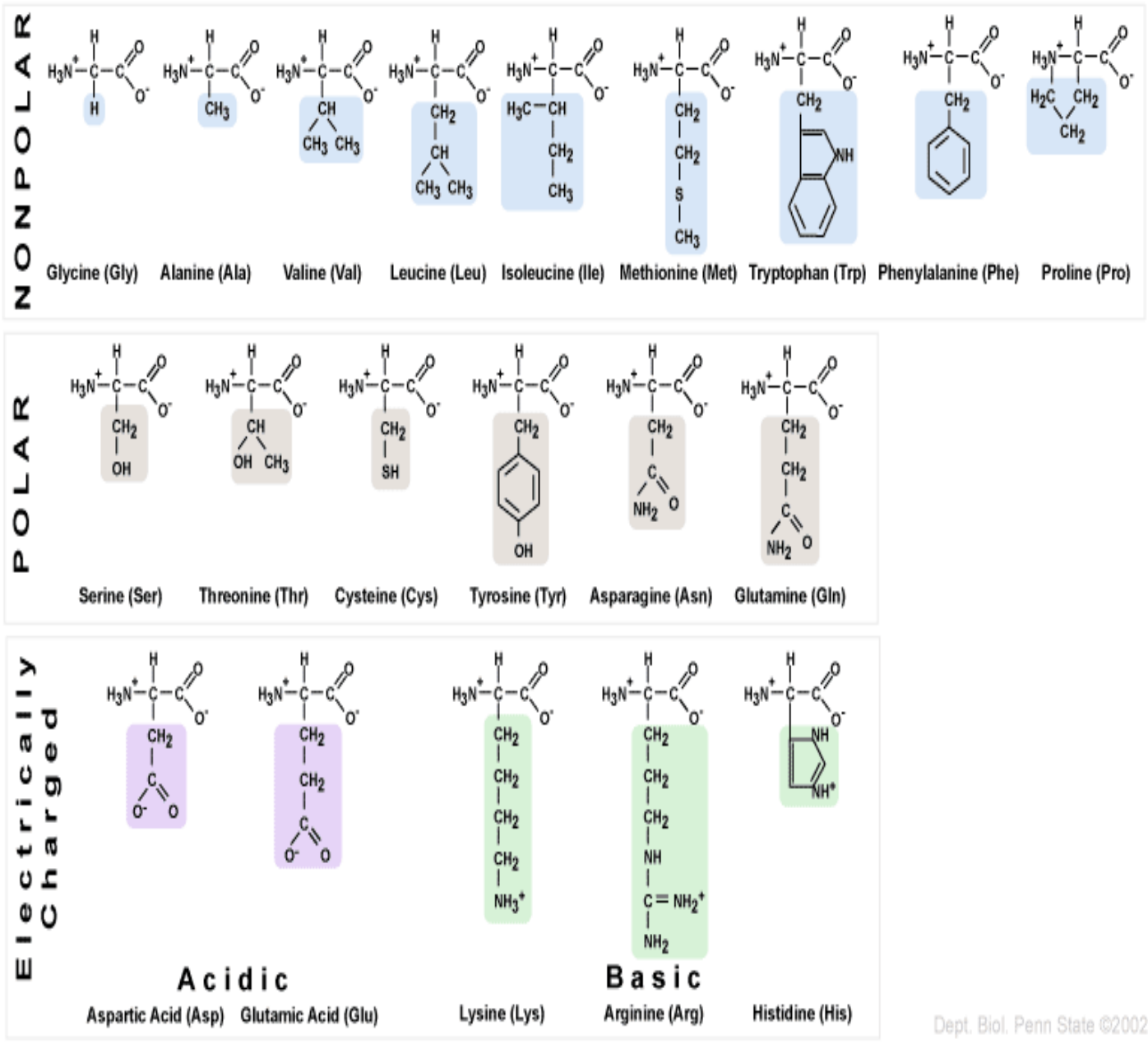
Neutral amino acids are amino acids which are neither acidic nor basic in nature thus they are neutral in solution. This is because they have equal number of amino and carboxyl groups. They are referred to as monoamino-monocarboxylic acids. Examples includes **glycine, alanine, threonine, tyrosine, valine, cysteine, tryptophan, leucine, methionine, aspargine, isoleucine, proline, glutamine, serine and phenylalanine**

Basic amino acids

This group of amino acids are basic in solution because they have extra nitrogen group in their structure that tends to attract a hydrogen atom, hence they are named as diamino-monocarboxylic acids. Examples includes **Lysine, arginine and histidine**

Acidic amino acids

Acidic amino acids are acidic in solution, because they have a side chain containing carboxylic acid moiety which tends to donates it hydrogen atom, hence they are referred to as monoamino dicarboxylic acids. Examples includes **aspartic acid and glutamic acid.**



B. Classification on the basis of chemical structure of the side chain of the amino acids

On the basis of the chemical structure of side chain of the amino acid, amino acids are classified as;

i. **Hydroxy amino acids**: This group of amino acids has a hydroxyl group in their chain. Example includes **serine, threonine and tyrosine.**

ii. **Dicarboxylic acid and their amides**: This group of amino acids has carboxylic group in their side chain. Example includes **aspartic acid and glutamic acids** while asparagine and glutamine are their amides respectively. They are known as dicarboxylic monoamino acids.

iii. **Diamino acids**: This group of amino acids is highly basic in character. The amino acids in this group have amino or amino related group in their side chain. Example includes **lysine**, **arginine** (with guanidino group) and **histidine** (with imidazole ring).

iv. **Aliphatic amino acids**: These are amino acids having aliphatic side chains (either be straight or branched carbon chains). They are monoamino monocarboxylic acids. This group consists of those having simple amino acids structure such as glycine, alanine, valine, leucine and isoleucine and those having branched chain structure such as Leucine, Isoleucine and Valine.

v. **Sulfur containing amino acids**: They are amino acids having sulfur in the side chain such as Cysteine having sulfhydryl group, methionine having thioether group and cysteine, a modified amino acids formed by condensation of two molecules of cysteine.

vi. **Aromatic amino acids**: Phenylalanine, tyrosine and tryptophan (with indole ring) make up this group. These are amino acids structurally have aromatic ring in their sides chain. Amino acids histidine is also placed in this group.

vii. **Imino acids or heterocyclic amino acids**: Amino acid proline is a classic example here as it contains pyrrolidine ring and an imino group (NH) instead of an amino group (-NH2) as compared to other. This pyrrolidine ring makes it to exhibit D-Isomer.

C. Classification on the basis of Nutritional requirement of amino acids

All the standard amino acids are required for the synthesis of many proteins, but not all these amino acids can be obtained through intake of adequate diet as some were synthesized in the body, hence on the basis of nutritional requirement, amino acids are classified into two groups namely:

i. Nutritionally essential or indispensable amino acids
ii. Non-essential amino acids

Nutritionally essential or indispensable amino acids

This group of amino acids cannot be synthesized by the body hence, they are required to be supply through diet as they are required for proper growth and maintenance of individual. Presently there are ten amino acids classified as essential amino acids, these include:

Arginine, **V**aline, **H**istidine, **I**soleucine, **L**eucine, **L**ysine, **M**ethionine, **P**henylalanine, **T**hreonine, **T**ryptophan.

For easier memorization, an acronym can be derived using their first letter e.g. **PVT TIM HALL or PH. VILLMA, TT or H. VITTAL, LMP.** A deficiency of an essential amino acid result in impaired protein synthesis which leads to negative nitrogen balance.

Note: Among these ten amino acids, arginine and histidine are known to be synthesized in adults only and not in growing children, hence they are classified as **semi–essential amino acids,** while the remaining 8 are classified as **absolutely essential amino acids**.

Non-essential amino acids

These are amino acids that can synthesized in the body in order to meet the metabolic demand in a living system since they cannot supply through diet. These about ten in number i.e **glycine, alanine, serine, cysteine, aspartate, asparagine, glutamate, glutamine, tyrosine and proline.**

D. Classification based on their metabolic fate

The structural nature of this carbon skeleton differs and these are used for further classification of amino acids. On the basis of the metabolic fate of the amino acids carbon skeleton, they are classified as:

i. **Glycogenic amino acids**: The carbon skeleton of this group of amino acids can serve as precursors for the synthesis of glucose or glycogen. About 14 out the 20 standard amino acids are classified as glycogenic amino acids. These are Glycine, alanine, serine, cysteine, aspartic acid, aspargine, glutamic acid, glutamine, proline, histidine, arginine, methionine, threonine, and valine.

ii. **Ketogenic amino acids**: The carbon skeleton of this group of amino acids are ketogenic in nature (Ketone), and can serve as precursors for the synthesis of ketone bodies. leucine and lysine are the two ketogenic amino acids

iii. **Glycogenic and ketogenic amino acids**: For this group, their carbon skeleton of has dual functions, as it can be serve as precursor of for synthesis both glucose and ketone bodies. Example includes isoleucine, phenylalanine, tryptophan and tyrosine.

E. Classification on the basis of the polarity of the side chain of the amino acids

The interaction of amino acids with water varies and this may influences the charges on the amino acids. Based on this, amino acids are classified into two groups namely:

i. Hydrophilic or polar amino acids

ii. Hydrophobic or non-polar amino acids

i. Hydrophilic or polar amino acids

Hydrophilic amino acids are group of amino acids having sides chains which are polar in nature and can interact favorably with water. This group is further subdivided into two categories based on whether their polar group contained charge or not.

Charged hydrophilic amino acids

The charge on hydrophilic amino acids can either be positive or negative, hence they are also subdivided into two groups namely:

1) **Negatively charge side chain:** The side chains of the dicarboxylic monoamino acids which are acidic in nature are negatively charged since they have carboxyl group which is negatively charged at physiological pH range. Example of these amino acid are aspartic and glutamic acids.
2) **Positively charged side chain:** The side chain of basic amino acids are negatively charged due to the presence of amino or amino related group which confer negative charge on them at physiological pH range. Examples of such amino acids are lysine, Histidine, and arginine.

Uncharged hydrophilic amino acids

This group of amino acids have no charge on their sides chain (-R)**.** Although they possess functional groups or atom such as -OH, -SH and -NH which enable them to for hydrogen bonds with water, thus enabling their interaction with water. Examples of these include:

- Serine and threonine which possessed –OH group in the side chain
- Glutamine and Aspargine which possed –NH group in the side chain.
- Glycine with –H in the side chain

ii. Hydrophobic or non-polar amino acids

This group of amino acids have side chains which interact poor with water, hence they have no charge on their side chains. Structurally, they usually contained aliphatic or aromatic as sides chains. exampls includes tryptophan, proline alanine, leucine, isoleucine, valine, methionine and phenylalanine.

Recent advances of standard amino acids

1. Recently, amino acids selenocysteine discovered by biochemist Thressa Stadtman in 1974 was recently tagged as the 21st amino acid. It exist as compositional amino acids

existing at the active sites of some enzymes/proteins such as glutathione peroxidase, glycine reductase and thioredoxin reductase. Synthesis from cysteine during translation process. Structurally, the amino resembled cysteine except that the sulfur atom is replaced by selenium. The IUPBMB has officially assigned three later symbol SEC and one letter symbol U for this amino acid. Debates on this development is ongoing as this amino acid is encoded by codon UGA which is one of the stop codon and it is only apparently ubiquitous in human and animal only.

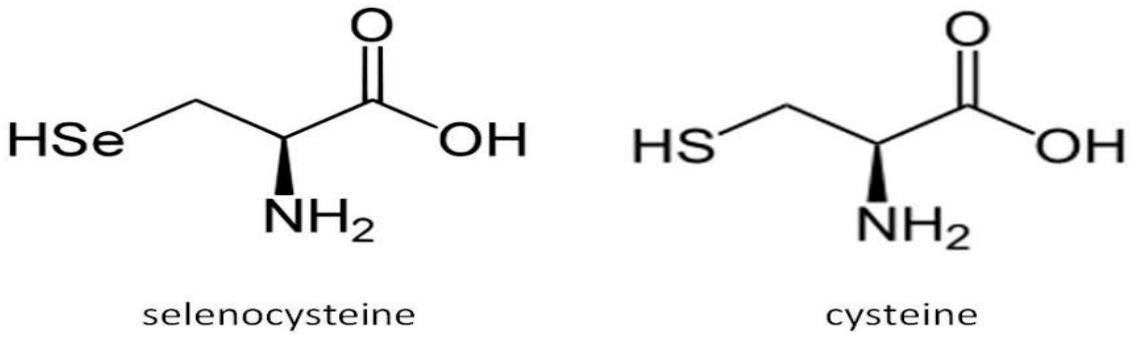


Fig 5: Structural comparison of amino acids selenocyteine and Cysteine

2. In 2002, a modified amino acids of lysine called Pyrrolysine was also tagged as the 22nd amino acid. This amino acids is formed from lysine upon modification of the amide group in lysine with 4-methylpyrroline-5-carboxylate group as shown below. This amino acids was discovered first in methanogenic archea and in some bacteria. The amino acids exist in some protein and has been assigned the stop codon UAG.

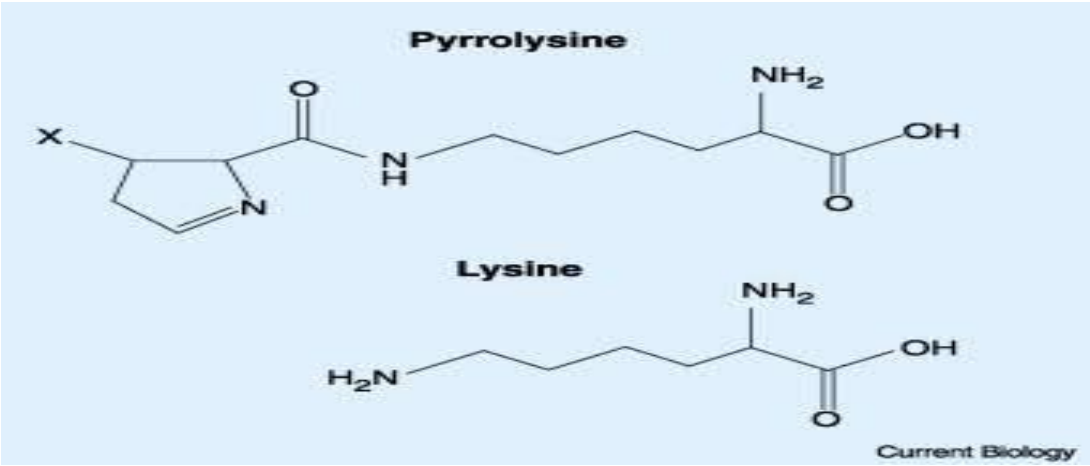


Fig 6: Structural comparison of amino acids pyrrolysine and lysine

Naturally occurring non-proteins amino acids

Aside the 20 standard amino acids, there are other amino acids which are biologically important which may or may not found proteins. These include the amino acid derivatives found in proteins and those are non-protein amino acids. These group of amino acids are called Non-Standard or Modified amino acids. Non-proteins amino acids. Classical examples of the modified amino acids are below;

- **Cystine**

  Cystine is a classic example of modified amino acid formed by covalent interaction of sulfhydryl (-SH) group in two cysteine molecules. This interaction leads to formation of disulfide bond. This amino acid is found in some immune cells, digestive enzymes, skin, in hair, skeletal and connective tissue. This is bond is essential in providing stability to the 3D-structure of proteins.

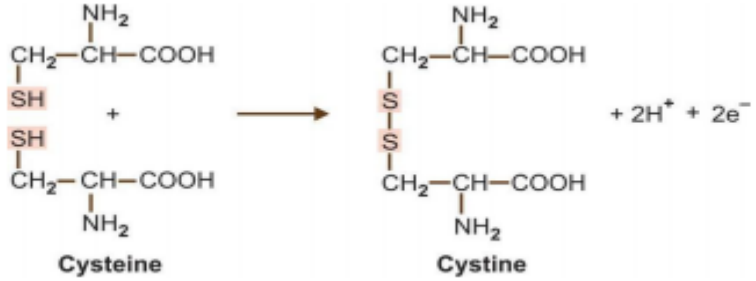


Fig 7: Structural formation of disulfide bond of cysteine

- **Gamma Carboxyglutamate**

γ- Carboxyglutamate is another modified amino acids in clotting proteins e.g. promthrombin and other calcium binding proteins). This is formed by carboxylation of amino acid glutamic acid side chains. Inability of the body to form this modified amino acid results in the bleeding disorder.

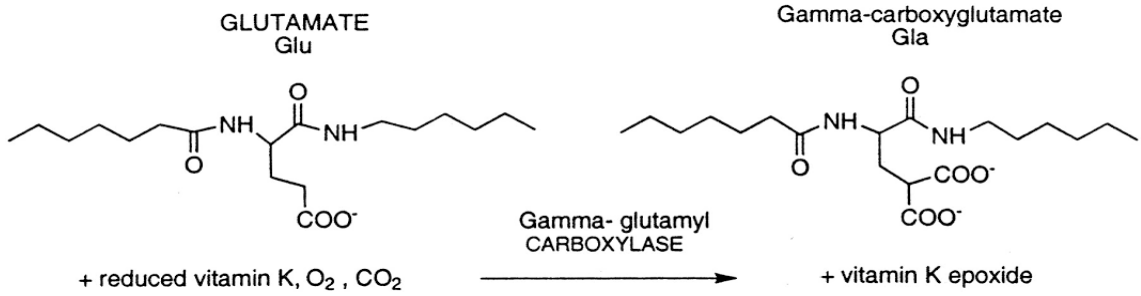


Fig 8: Structural formation of amino acid γ- Carboxyglutamate

- **Hydroxy-proline and hydroxylysine**

  These are another structural amino acids found primarily in collagen of the connective tissue. They are formed by selective hydroxylation of lysine and proline residues in protein collagen.

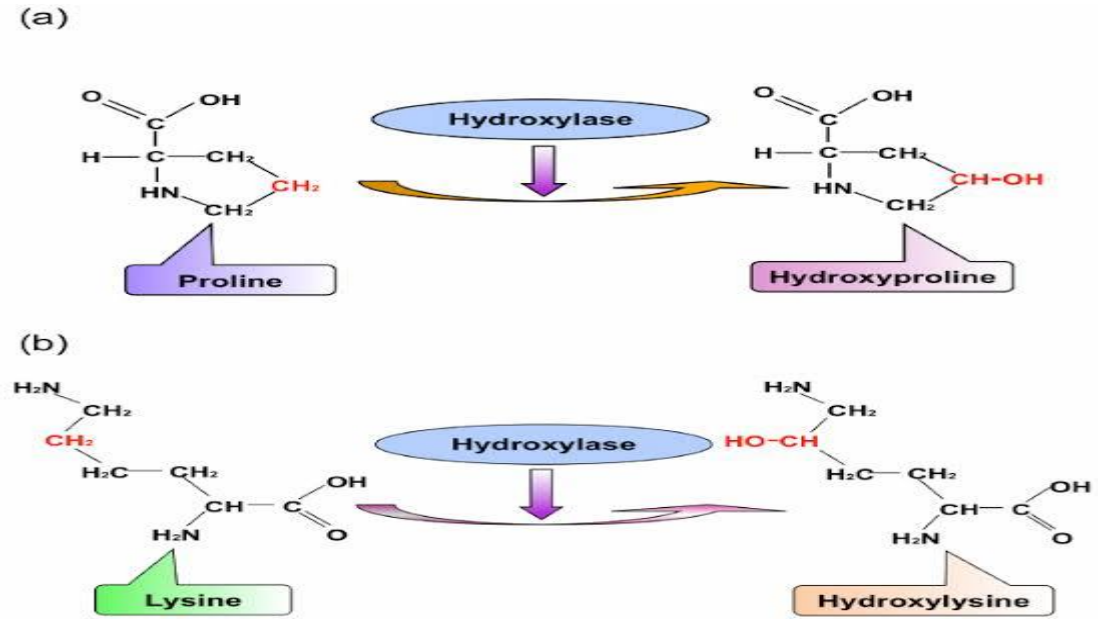


Fig 9: Formation of (b) Hydroxylysine and (a) hydroxyproline

Other modified amino acids include:

i. N-methylysine
ii. Phosphoserine, phosphothreonine, and phosphotyrosine
iii. Desmosine and isodesmosine

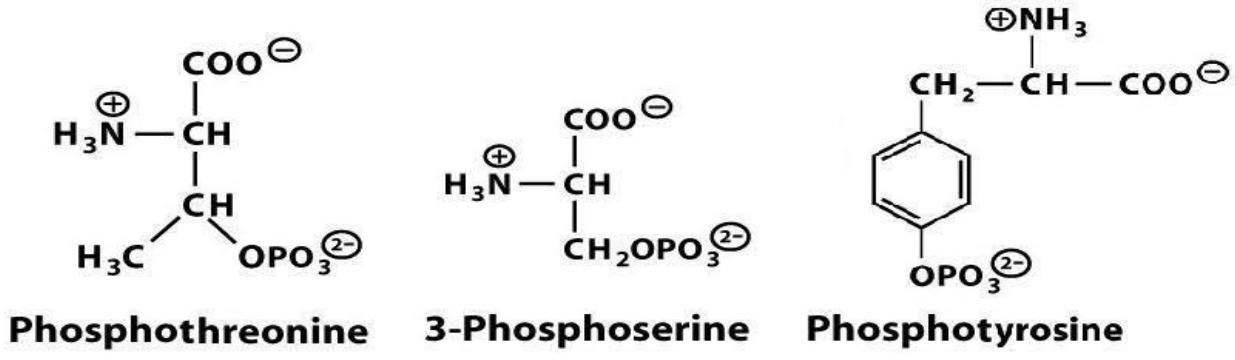


Fig 10: Structure of modified amino acid phosphoserine, phosphothreonine, and phosphotyrosine

General Properties of amino acids

Physical properties of amino acids

1. Amino acids are colorless, crystalline solid.
2. All amino acids have a high melting point greater than 200^{o}
3. Solubility: They are soluble in water, slightly soluble in alcohol, and dissolve with difficulty in methanol, ethanol, and propanol. R-group of amino acids and pH of the solvent play important role in solubility.
4. On heating to high temperatures, they decompose.
5. All amino acids (except glycine) are optically active.

Chemical properties of amino acids

1. Ionization behavior

Amino acids in solution ionizes and the ionizing behavior here depends on the pH of their solution, nature of system and structure of the amino acids. Their ability to ionize enables them to exert:

- Acid base behavior
- Amphoteric properties (zwitter ion formation)
- Buffering activity

Acid-base behavior

The presence of the amino and carboxyl groups attached to the α-carbon of amino acids and on the polarity of the sides attaching to this α-carbon enable them exhibit acid/base behavior. Since the carboxyl (-COOH) group of an amino acid can donate proton (H^+) forming a negatively charged anionand the amino group (-NH_2) of an amino acid can accept the proton (H^+) forming positively charged cation, thus acting as an acid and as a base respectively. The uncharged or unionized structure of amino acids cannot exist. At pH of 7.4, the physiological pH of the blood plasma, the carboxyl group(COOH) exist almost as carboxylate(COO^-) while the amino group (NH_2) exist in protonated form as NH_3^+.

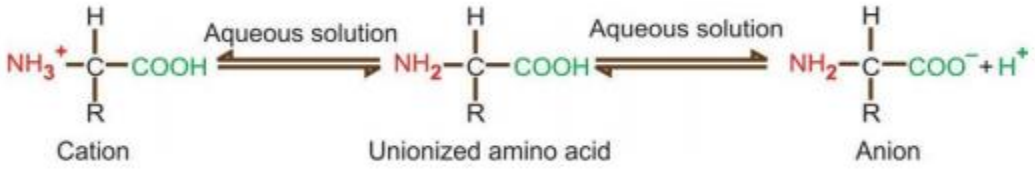


Fig 11: Ionization of amino acid

Amphoteric properties of Amino acids

Since amino acids can behave as an acid and as a base, thus exhibiting acidic and basic properties, thus amino acids is **Amphoteric in nature**. Also at Isoelectric pH tagged as PI (the pH at which amino acids bears no net charge and thus show no movement in an electric field), amino acids is exhibit amphoteric property.

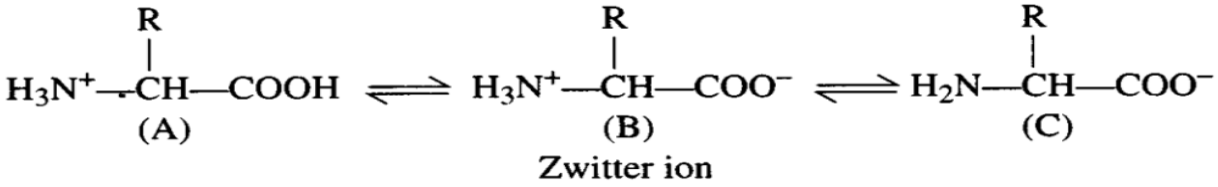


Fig 12: Amphoteric nature of amino acids, (A) Cationic form, (B) Zwitter ion form, (C) Anionic form.

Buffering activity

Buffer solution is a solution that resist slight change in pH when small quantity of acid or base is added. This type of solution do contained weak acid component and weak base component. In aqueous solution, amino acid contains weak acidic component i.e α-carboxylic group (α-COOH) and weak basic α-amino component i.e. (α-NH_2), thus amino acids can act as buffers.

Note: Among all the standard amino acids, histidine has best buffering capacity at physiological pH, due to the fact that the histidine side chain has pKa of 6.0 while other amino acids have pKa value too far from pH 7. Maximum buffering capacity occurred when pH = pKa.

2. Zwitter ion formation

A zwitterion can be defined as a molecule having functional groups, of which at least one has a positive and one has a negative electrical charge, thus conferring a net charge of zero on the entire molecule. Zwitter ion is also known as dipolar molecule. Monoamino monocarboxylic acids exist in aqueous solution as dipolar molecule. They contain an amine group (basic) and a carboxylic group (acidic).

The $–NH_2$ group is the stronger base, and so it picks up H^+ from the -COOH group to leave a zwitterion. Although the (neutral) zwitterion is the usual form of amino acids that exist in the solution.

3. Ninhydrin reaction

On heating of amino acids solution with excess ninhydrin, all amino acids having free α-amino group gives a purple colored solution while proline which has imino group in place of amino group gives yellow colored product. This equation for this reaction is shown below;

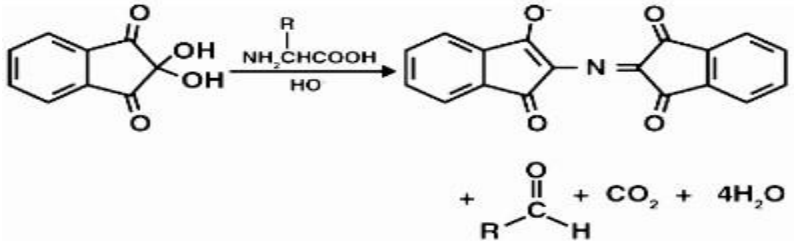


Fig 13: Reaction of Ninhydrin with amino acids

This reaction is used for colorimetric estimation of amino acids and in chromatographic separation.

3. Peptides formation

Peptide formation is an important reaction as this reaction results in formation of proteins, which is an important macromolecules. Two amino acids can reacts together by undergoing condensation reaction leading to formation of peptide bond (covalent interaction) linking them together as dipeptide with elimination of a water molecule as shown below;

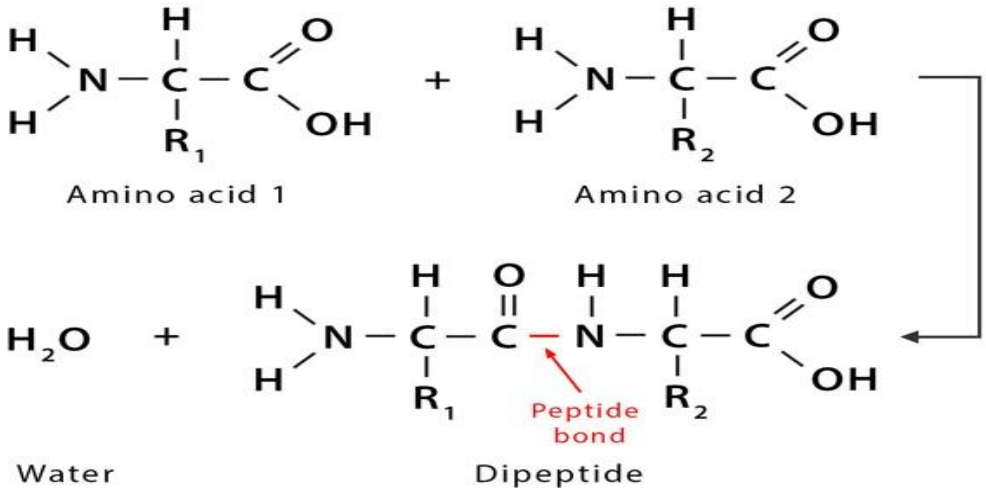


Fig 14: Peptide bond formation between two amino acids

This peptide bond is basically a amide linkage and the link is formed between α-carbonyl group of one amino acid and α-amino group of another. Three amino acids can be joined together by two peptide bonds to form tripeptides while four amino acids can be joined together to form tetrapeptides. Similarly five amino acids can be joined together to form pentapeptides and few amino acids can joined together to form oligopeptide, while many amino acids can joined together to form polypeptides.

Proteins are basically polymer of amino acids (polypeptides) as many amino acids were covalently connected between the N-terminal of one amino acid to the C- terminal of the next amino acids. The amino acid units in this polypeptides is called residue.

Some biologically important peptides are:

1) Insulin (30-amino acid residues)
2) Glucagon (29-amino acid residues)
3) Thyrotropin releasing hormone (3-amino acid residues)
4) Oxytocin (9-amino acids residues)
5) Carnosine and anserine (2- amino acid residues)

Chemical test for amino acids

The following are chemical test used for detecting presence of amino acids in solution.

- **Ninhydrin test**.
- **Xanthoproteic test**

The xanthoproteic test is performed for the detection of aromatic amino acids (tyrosine, tryptophan, and phenylalanine) in a protein solution. The nitration of benzoid radicals present in the amino acid chain occurs due to a reaction with nitric acid, giving the solution yellow coloration.

- **Reaction with Sanger's reagent**

Sanger's reagent (1-fluoro-2, 4-dinitrobenzene) reacts with a free amino group in the peptide chain in a mild alkaline medium under cold conditions.

Functions of amino acids

1) **Building blocks of proteins**: Proteins is polymer of amino acids. Amino acids is the molecular building uints in proteins and only L-amino acids are polymerized to form proteins, although both D-amino acids and non-L-amino acids found in nature
2) **Biological buffers**: Amino acids are amphoteric and therefore they can act as buffers in solutions, resisting slight changes in pH.

3) **Nitrogen storage**: Amino acids contains nitrogen and hence they act as reservoir for nitrogen. Good examples are asparagine and glutamine which are amide derivatives of aspartic acid and glutamic acid respectively. They serve as storage of nitrogen
4) **Formation other compounds**.
   (i) Tyrosine produces the hormones thyroxin and adrenaline and the skin pigment melanin,
   (ii) Glycine forms heme
   (iii) Tryptophan produces vitamin, nicotinamide and plant hormone Indole Acetic acid (IAA). The coenzyme A, a vitamin pantothenic acid, coenzyme glutathione and alkaloids are some other compounds formed by amino acids
5) **Antibiotics**: The non-protein amino acids are useful compounds of antibiotics e.g. Azaserine, Valinomycin etc.
6) **Genetic defects**: Inborn errors in the metabolism of amino acids cause several disorders e.g. Phenylketonuria.
7) **Detoxification reactions**: Glycine, cysteine and methionine are involved in the detoxification of toxic substances
8) **Formation of glucose**: Glucogenic amino acids are converted to glucose in the body.
9) **Enzyme activity**: The thiol (-SH) group of cysteine has an important role in certain enzyme activity.

PROTEINS

As mentioned earlier proteins are polymeric molecules made by polymerization of the amino acids. A protein is a molecule made from many chains of amino acid linked together by covalent peptide bonds. Each protein has its own specific and unique amino acids sequence which defined its 3D- structure and biologic functions. Proteins is one of the major classes of biomolecules and food. They are known to perform various functions. Some proteins contained single polypeptide chains, while two or more polypeptides associated non-covalently together are major constituents of some other proteins.

Proteins are one of most abundant organic molecules of life. They perform diverse static (structural) and dynamic functions in the living cells. The dynamic functions of proteins are highly diversified such as enzymes, hormones, clotting factors, immunoglobulins, storage proteins and membrane receptors. Half of the amino acids (about 10) that occur in proteins have to be consumed by humans in the diet, hence they are essential. A protein is said to be complete (or

first class) protein if all the essential amino acids are present in the required proportion by the human body e.g. egg albumin

Structure of proteins

Each native proteins has its own native structure i.e. it three dimensional structure called **conformation.** They are made or build up from the 20 standard amino acids. The amino acids organization in each protein determine the structural conformation of the proteins. The protein structure are classified into four level of organization namely:

1) Primary structure,
2) Secondary structure
3) Tertiary structure
4) Quaternary structure

There are several bonds or interaction that is responsible for formation, stabilization and maintenance of each of these structural level or organizations. These bonds are grouped into two categories namely

- Covalent bonds which comprises of
    I. Peptides bonds
    II. Disulfide bond
    III. Lysinonorleucine bond
- Non-covalent bonds which comprises of
    I. Hydrogen bond
    II. Hydrophobic interaction
    III. Electrostatic interaction
    IV. Van der waals interaction

Covalent bonds

- **Peptide bonds:** A peptide bond is type of covalent interaction formed by condensation of the amino group of one amino acid with the carboxyl group of another amino acid with the release of a water molecule. This bond is responsible for polymerization of the protein monomer i.e. amino acid.

- **Disulfide bond:** This is a covalent interaction formed by interaction between the sulfhydryl group(-SH) of the side chain of cysteine residue in same or different polypeptide chain
- **Lysinonorleucine bond:** This covalent bond is oxidized lysine residue and an unmodified lysine side chain forming a crosslink between them both with or between the triple helical units. This is very significant as it give tensile strength to the structure.

Non covalent bonds

- **Hydrogen bond:** This bond type is formed as a result of non-covalent interaction between the –NH and –CO groups of peptide bond within the same polypeptide chain or different polypeptide chain. Side chains of these eleven (11) amino acids can do participate in hydrogen bonding (tryptophan, glutamic acid, glutamine, aspartic acid, aspargine, threonine, serine, lysine, tyrosine, histidine and arginine).
- **Hydrophobic bond**: This type of bond exist through the interaction between nonpolar hydrophobic R groups (side chain) of amino acids like methionine, phenylalanine, alanine, valine, leucine, isoleucine, methionine, phenylalanine and tryptophan.
- **Electrostatic interaction**: This exist between the oppositely charged groups when they are in contact or closer, oppositely charge group like amino (NH_3^+) terminal and carboxyl (COO–) terminal groups of the peptide and the oppositely charged R-groups of polar amino acid residues. This is also referred to as ionic bond or salt bond or salt bridge
- **Van der Waals interaction**: Van der Waals forces are known to be an extremely weak forces which act only at extremely short distances. This force constitutes both the attractive and a repulsive forces existing between both polar and nonpolar side chain of amino acid residues.

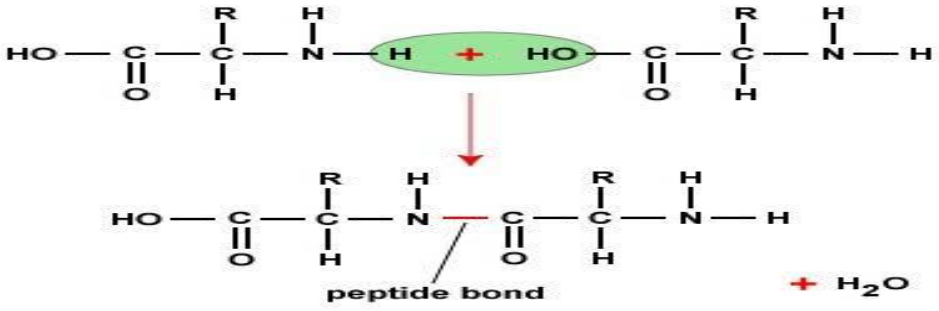


Fig 12: Peptide bond formation

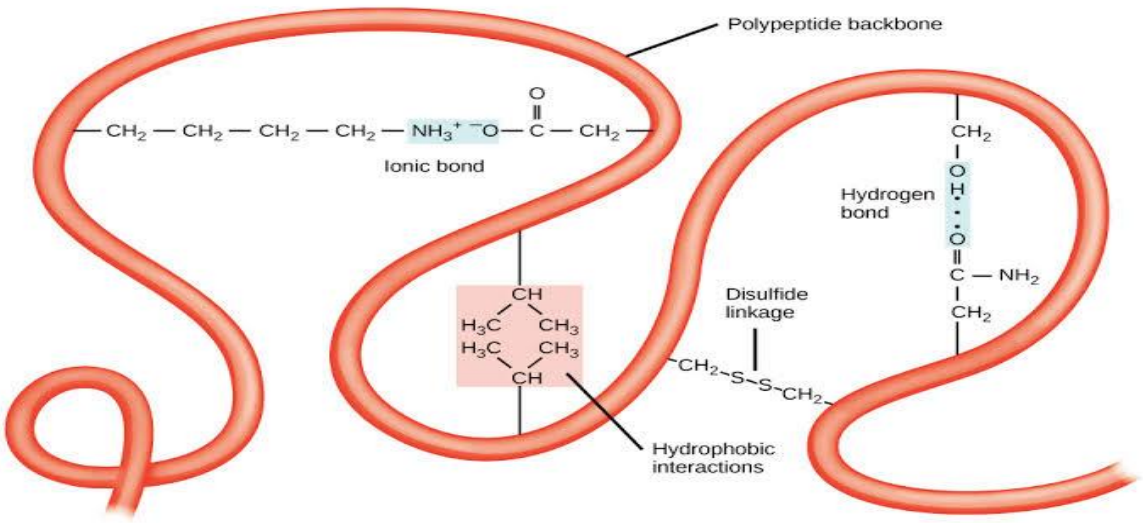


Fig 13: Illustration of bonds responsible for protein structure

STRUCTURAL LEVELS OF ORGANIZATION IN PROTEINS

1) **Primary structure of proteins**

The primary structure of proteins basically deals with the amino acids sequence which formed the backbone of proteins and location of any disulfide bond in a protein. It has established that in proteins the amino acids are joined covalently by peptide bonds resulting in formation of unbranched polypeptide chains. Each proteins have unique amino acid sequences that are specified by genes and this determine the primary structure of proteins.

For better identification, the amino acids are read from the N-terminal to C-terminal ends of protein. Each amino acid in a polypeptide is called a residue or moiety as shown below;

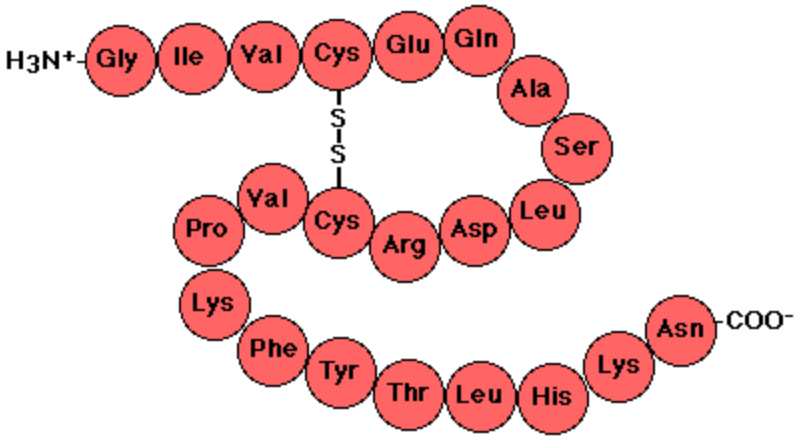


Fig 14: Primary structure of protein, structure showing the amino acids sequence and arrangement; and the location of disulfide bond

For experimental purpose, determination of amino acid composition by acid/alkaline hydrolysis or enzyme treatment; degradation of protein or polypeptide into smaller fragments and determination of the amino acid sequence using Sanger's or Edman's reagent are done in studying the primary structure of proteins.

A pancreatic hormone insulin is a classic example of protein which has been sequenced. The hormone was first sequence by Frederick Sanger in 1953. His study revealed that the hormone has a molecular weight of about 5808 and composed of 51 amino acids arranged into two polypeptides chains tagged as A(21) and B(30) chains as shown below;

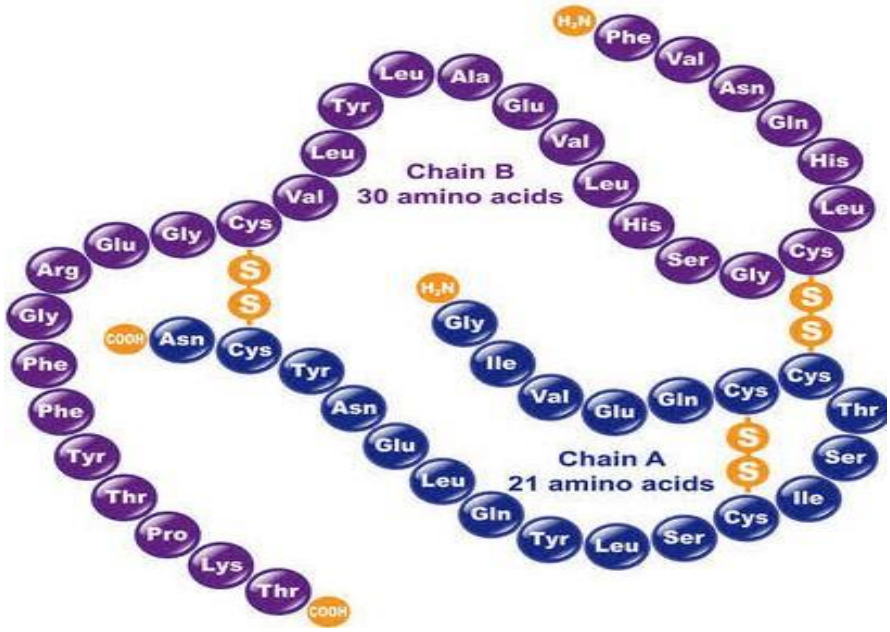


Fig 15: Amino acid sequence of a typical protein (insulin)

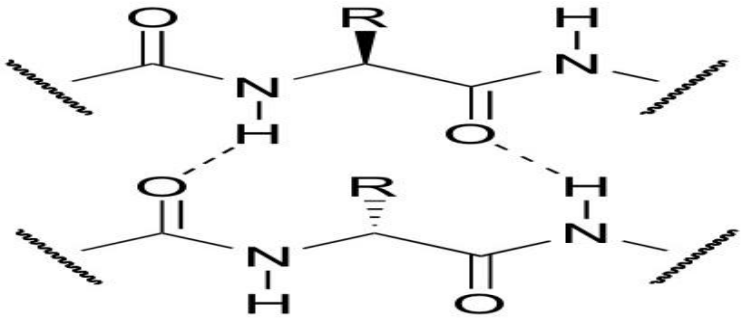
Clinical significance of primary structure of proteins

A clear understanding of the primary structure of a protein is crucial for the molecular study of disease molecular epidemiology, this is because many genetic diseases have been linked incorporation of an abnormal amino acid sequences e.g. sickle cell, therefore, if the primary structure of the normal and mutated proteins are known, this information would help in diagnosing or studying of the disease.

Secondary structure of proteins

The secondary structure constitutes the regular folding and twisting of the polypeptide chain brought about by hydrogen bonding. The stability of the primary structure is maintained by

hydrogen formed between the hydrogen of amide group (-NH) and oxygen of the carbonyl group within polypeptide chain which result in twisting of the primary structure. A typical illustration of this is shown below;

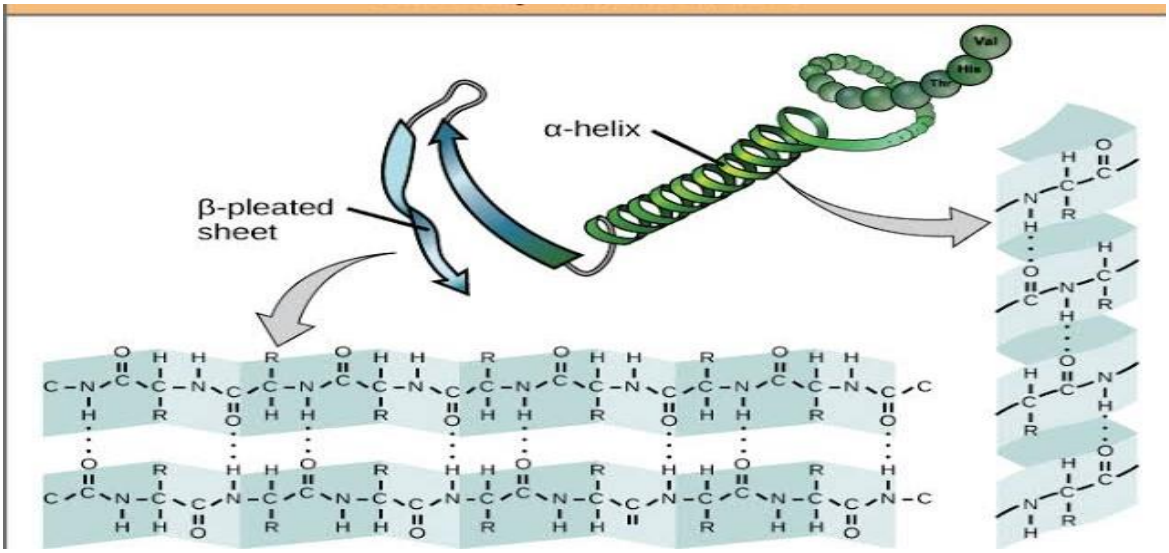


Hydrogen bond formation in polypeptide chain

The twisting by the hydrogen bonding gives rise to different kinds of secondary structure namely:

1) α-Helix or Helicoidal state
2) β-pleated sheet (strectched state)
3) loop regions
4) β-bend
5) disordered region
6) Triple helix

The **α-Helix** is also known as helicoidial state, the **alpha** in the name alpha helicoidal arises based on the fact that it was the first secondary structure of protein elucidated and this was done by **Pauling and Correy**. This structure is established if the backbone of polypeptide chain is twisted by an equal amounts about each α-carbon such that it forms a coil or helix or rod-like structure and this structure is maintained by hydrogen bonds between the NH and CO groups of amino acids with the same chain. In forming this, the CO group of each amino acid is hydrogen bonded to the –NH group of another amino acid that is situated approximately four residues ahead in the linear sequence, resulting in axial distance of about 1.5 Å between adjacent amino acids. Thus giving 3.6 amino acid residues per turn of helix shown below;



Secondary Structure of proteins

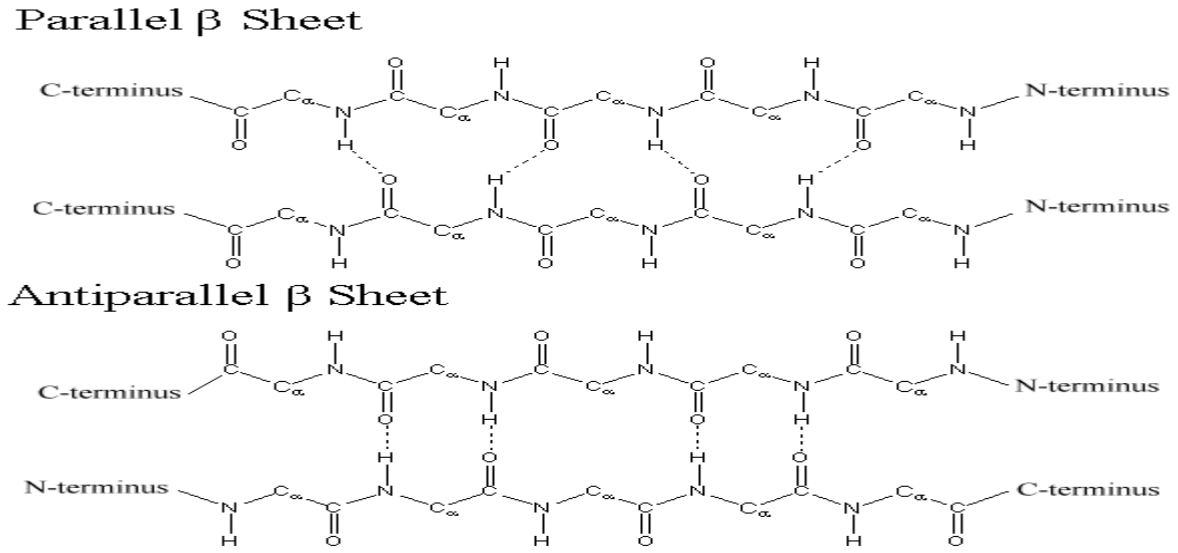
Amino acids **glycine and proline** are known to be the **Helix destabilizing amino acids**. Proteins like **α-keratin** of the hair, **myosin** and **tropomyosin** of the muscle and **hemoglobin** of the muscle exhibit this helical structure. The α-helix structure is significant in that it can coil round one another forming a structure having strong stiff bundles of fiber which provides mechanical support.

β-pleated structure is the second structure discovered by Paul and Corey as the another type of secondary structure of protein. In this the hydrogen bonding interaction can twist the polypeptides chain to be like **β-pleat sheet** fully extended rather than being tightly coiled as in the α-helix. β-pleated sheet is created between two or polypeptides chain rather one in α-helix, and the structure is stabilized by hydrogen bonds between NH and C=O groups in a **different or the same polypeptide** chains creating an **axis distance of 3.5 Å** between the adjacent amino acids. The hydrogen bonds in **β-pleated structure** are perpendicular to the polypeptide backbone rather than parallel in α-helix.

Depending on the arrangement of polypeptide chains in β-pleated sheet conformation, the conformation can exist in two different ways as shown below;

a) **Parallel pleated**
b) **Anti-parallel pleated sheets**

Illustrations about the arrangement is shown below;



Parallel pleated and anti-parallel pleated sheets as secondary structure of proteins

As shown above in **Parallel pleated structure,** the polypeptide chains are arranged such that it lies side-by-side and in the same direction (with respect to N- and C-terminal), so that their N-terminal residue of polypeptides chain faced oppositely to the N-terminal residues of another polypeptides chain i.e N-terminal faces to N-terminal and stabilized by hydrogen bonding while in antiparallel pleated structure, N-terminal residue of polypeptides chain faced oppositely to the C-terminal residues of another polypeptides chain and stabilized by hydrogen bonding.

Classical example of β-pleated sheet structure are found in silk fibroin and in amyloid protein expressed in Alzheimer's diseases, and in both fibrous and globular proteins.

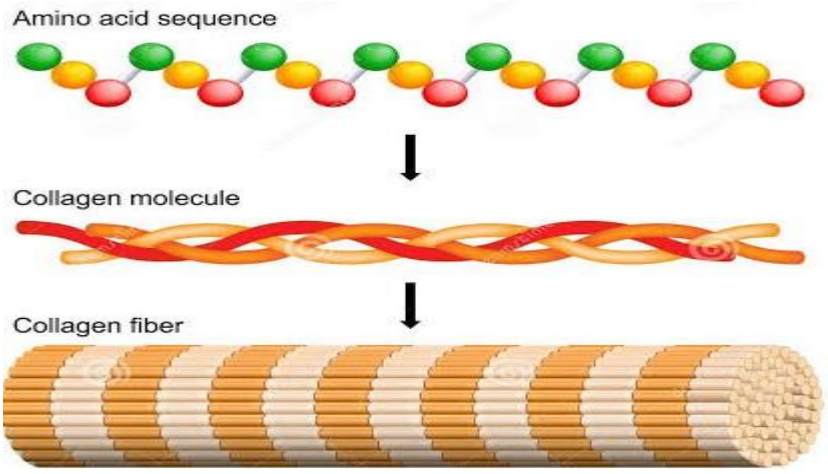
The loop or coil conformation is an irregularly ordered region which confers the remaining region left by α- and β- region in globular protein, although the structure lacks regular secondary structure but it is ordered. This region forms the antigen binding sites of antibodies.

The β-turn or bend or hair pin turn is another secondary structure of proteins, this region of β-turn shows abruptly reverse direction and often connecting the ends of the adjacent antiparallel β- strands. This region usually consist of four residues of amino acids (proline or glycine included) stabilized by hydrogen bond between the first and

fourth Amino acid residues. The structure is responsible for the formation of **a compact globular structure**

Disordered Regions consist of other several conformations in solution other than the above listed secondary structure is called disordered region. This structure becomes ordered upon binding to ligands.

Triple helix is found in all collagen protein. This structure helps to confer tensile strength of steel in protein and host defense protein. Triple helical conformation exist where there are a lots of helix destabilizing amino acids such proline, hydroxyproline and glycine, prevention the formation of α-helical structure. In collagen, we have three polypeptide chains, each twisted into the left handed of the helix having three amino acids per turn as shown below;

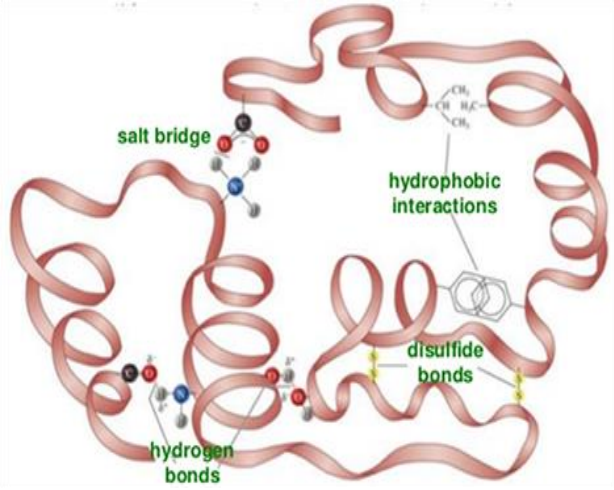


Structure of collagen

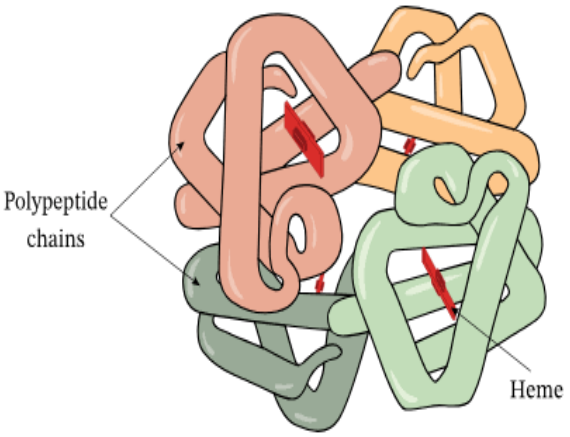
The **super secondary structure** of proteins expressed the combination of all the secondary structural features in protein. A protein which contained all arrangement in different part such that a part form an **α-helix** followed by a region of **pleated structure** which may include parallel or anti-parallel intervening by **β-turns, loop regions and disordered regions** is said to have super-secondary structure

Super secondary structure of proteins

The three dimensional folded compact and biologically active conformation of a protein is referred to as it **tertiary structure,** the structure is stabilized by hydrogen bonds, hydrophobic interactions, Van der waals force, disulfide bond and ionic interaction. After establishment of secondary structure, the protein may fold and twist further to for three dimensional arrangement of the polypeptide chains. This folding involved interaction of amino acids at distant with each other after being brought together and this is responsible for the formation of active or catalytic sites of enzymes. All the bonds responsible for the formation of the tertiary structure are shown in the diagram



Tertiary structure of proteins



Quaternary structure of proteins

The arrangement of polypeptides subunits in three dimensional complexes is called the quaternary structure of the proteins. Only proteins with more than one polypeptides chains (polymeric) can have quaternary conformation because not all proteins are polymeric in nature.

Each sub-units may function independently or may work in co-operation with other such as in case of heme protein. Protein subunits in quaternary structure are held together by non-covalent interaction. Examples of such is found in hemoglobin.

CLASSIFICATION AND FUNCTIONS OF PROTEINS

Proteins have been classified in several ways, but the most commonly used is the classification based:

- Function
- Physical and chemical properties
- Nutritional importance

A. Classification of Proteins Based on Function

On the Basis of functions that proteins perform, they are classified into the following groups (with examples)

1) **Catalytic Proteins or Enzymes**: These proteins act as enzymes, e.g. Hydrolytic enzymes e.g. pepsin andtrypsin, glucokinase, dehydrogenases and transaminases, etc.
2) **Transport Proteins**: These proteins are involved in the process of transportation, examples includes Hemoglobin for transports oxygen, transferrin for transports iron and Albumin that act as career of fatty acids and bilirubin.
3) **Storage Proteins**: Many proteins serve as storage form, examples include apoferritin which stores iron in the form of ferritin and Myoglobin that stores oxygen in muscles.
4) **Contractile Proteins**: These proteins have the ability to contract, thus functioning as component of contractile system of skeletal muscle. Examples includes actin and Myosin of the muscle.
5) **Structural Proteins**: This class of proteins provides supporting framework for cells to give biological structure, strength or protection, e.g. Collagen in bone, cartilage, elastin of ligaments and keratin of hair, nail.
6) **Defence Proteins**: These groups of proteins are involved in defence mechanism against invasion of foreign substances such as viruses, bacteria and cells. Examples includes Immunoglobulins or antibodies that fight against pathogen, fibrinogen and thrombin which are blood clotting proteins that prevent loss of blood when the vascular system is injured.

7) **Regulatory Proteins**: This class of proteins participates in regulating cellular or physiological activity. Examples includes hormones such as insulin and glucagon that helps in regulating sugar metabolism and growth hormone of pituitary gland that regulates growth of the cells.

B. Classification of Proteins Based on Physical and Chemical Properties of Protein

Here proteins are classified into three main sub-groups, namely

1) Simple proteins
2) Conjugated proteins

1) Derived proteins.**Simple proteins** are those proteins that upon hydrolysis, yield only amino acids residue or their derivatives. They are further according to their solubility and heat coagulability as:

- **Albumins**: These proteins are soluble in water and coagulated by heat. Examples includes egg albumin, serum albumin and Lactalbumin of milk
- **Globulins**: These proteins are insoluble in water, but they are soluble in dilute neutral salt solution and are heat coagulable. Examples includes ovaglobulin of egg yolk, Serum globulin and myosin of muscle.
- **Glutelins**: These proteins are soluble in dilute acids and alkalis, but are insoluble in neutral solvents. They exist mostly as plant proteins e.g. glutelin of wheat, oryzenin of rice
- **Prolamins or alcohol soluble proteins**: Prolamins are soluble in 70 to 80% alcohol, but they are insoluble in water, neutral solvent or absolute alcohol. These proteins are rich in amino acids proline but are deficient in lysine. e.g. Zein of corn and Gliadin of wheat
- **Histones**: These proteins are soluble in water, but are not coagulated by heat. They are rich in basic amino acids. Histones usually occur in tissues in salt combinations with acidic substances, such as nucleic acids (RNA and DNA). Examples includes nucleoprotein.
- **Protamine**
- **Scleroproteins (fibrous proteins)**

2. **Conjugated Proteins**: These proteins are composed of simple protein combined with some non-protein substance. The non-protein group is referred to as the prosthetic (additional) group. Examples of conjugated proteins includes:

- **Nucleoproteins**; These composed of simple basic proteins (e.g. histones or protamines) with nucleic acids (RNA and DNA) as the prosthetic groups. Example includes nucleohistone and nucleoprotamine.
- Glycoproteins and proteoglycans or mucoproteins: These conjugated proteins consist of simple protein and carbohydrate as a prosthetic group. If the carbohydrate content is less than 4% of protein it is called glycoprotein such as e.g. Mucin of saliva while if the carbohydrate content is more than 4%, it is called mucoprotein or proteoglycans, e.g. glycosaminoglycans.
- Chromoproteins: This conjugated proteins composed of simple proteins with a colored prosthetic group, e.g. Hemoglobin, Cytochromes, and Catalase which all have heme groups as prosthetic group
- Phosphoproteins: These conjugated proteins are formed by a combination of protein with phosphoric acid as prosthetic group e.g. casein of milk and vitellin of egg yolk.
- Lipoproteins
- Metaloproteins

3. **Derived Proteins**: These proteins are obtained simple and conjugated proteins. They are further sub-divided into two group:
   (1) Primary derived proteins (denatured proteins)
   (2) Secondary derived proteins

The primary derived are the denatured or coagulated or first hydrolysed products of proteins e.g Coagulated protein, metaloproteins and proteans while the secondary derived are the degraded (due to breakdown of peptide bonds) products of proteins e.g. proteoses, peptones, polypeptides and peptides.