FOURTH CHAPTER: Structure and physicochemical properties of Proteins

I. Definition and importance in biology

Proteins are large biomolecules and macromolecules that comprise one or more long chains of amino acid residues. Proteins differ from one another primarily in their sequence of amino acids, which is dictated by the nucleotide sequence of their genes, and which usually results in protein folding into a specific 3D structure that determines its activity.

Like other biological macromolecules such as polysaccharides and nucleic acids, proteins are essential parts of organisms and participate in virtually every process within cells. They provide many of the structural elements of a cell, and they help to bind cells together into tissues. Many proteins are enzymes that catalyse biochemical reactions and are vital to metabolism. Some proteins act as contractile elements to make movement possible. Others are responsible for the transport of vital materials from the outside of the cell ("extracellular") to its inside ("intracellular"). Proteins, in the form of antibodies, protect animals from disease and, in the form of interferon, mount an intracellular attack against viruses that have eluded destruction by the antibodies and other immune system defenses. Many hormones are proteins. Last but certainly not least, proteins control the activity of genes ("gene expression").

In animals, proteins are needed in the diet to provide the essential amino acids that cannot be synthesized.

II. Amino acids

All amino acids possess common structural features, including an α-carbon to which an amino group, a carboxyl group, and a variable side chain (R) are bonded, which determines the chemical properties of each amino acids.

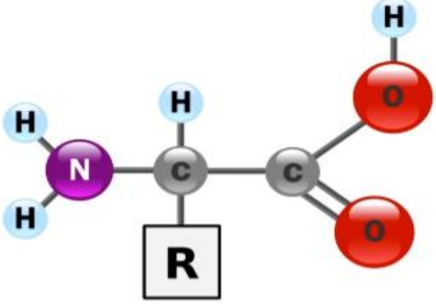


Fig 1. Amino acid schematic.

Although over 500 amino acids exist in nature, we distinguish:

- The 20 AA constituents of natural proteins or standard amino acids. They are encoded in DNA and incorporated into the peptide chain during mRNA translation.
- And the others, which are found either in the free state or in peptides synthesized by microorganisms or plants.

Based on side chains, we can group the 20 amino acids found in proteins as follows:

- **Aromatic** (phenylalanine, tyrosine, tryptophan);
- **Aliphatic** (leucine, isoleucine, alanine, methionine, valine);
- **Hydroxyl/Sulfhydryl** (threonine, serine, tyrosine, cysteine);
- **Carboxyamide** (glutamine, asparagine);
- **R-Acids** (glutamic acid, aspartic acid);
- **R-Amines** (lysine, histidine, arginine);
- **Odd** (glycine, proline).

Note that tyrosine has a hydroxyl group and fits into two categories.

Proteins are often called polypeptides in reference to the fact that they are composed of amino acids held together by peptide bonds. Amino acids can be linked by a condensation reaction in which an —OH is lost from the carboxyl group of one amino acid along with a hydrogen from the amino group of a second, forming a molecule of water and leaving the two amino acids linked via an amide.

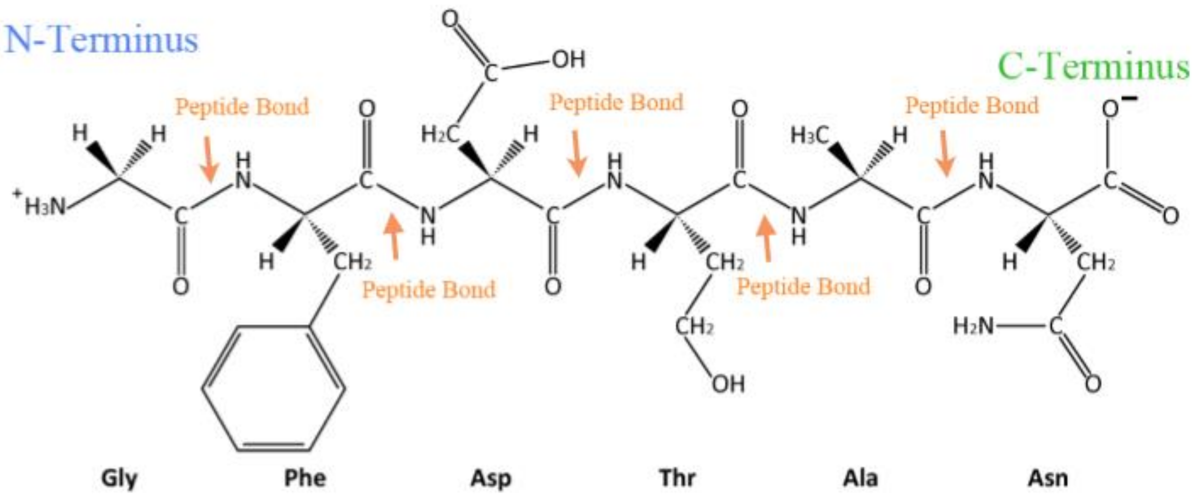
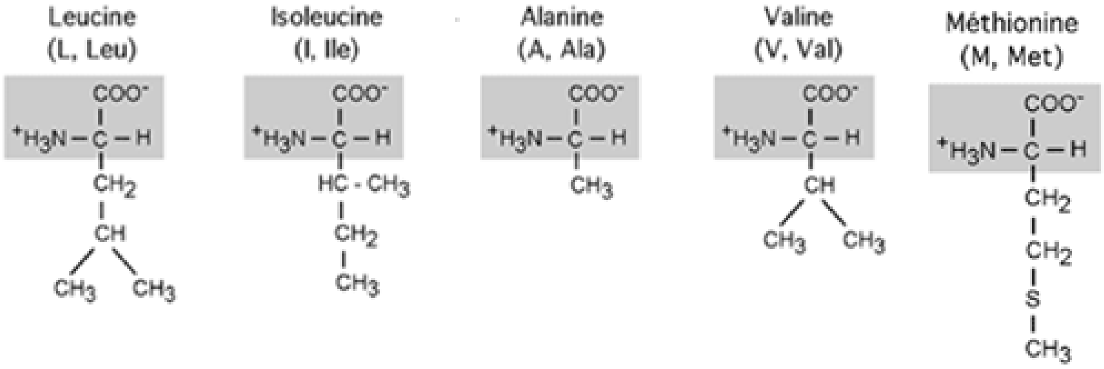


Fig 2. Peptide bonds.

Amino acids joined by a series of peptide bonds are said to constitute a peptide. After they are incorporated into a peptide, the individual amino acids are referred to as amino acid residues. Small polymers of amino acids (fewer than 50) are called oligopeptides, while larger ones (more than 50) are referred to as polypeptides. Hence, a protein molecule is a polypeptide chain composed of many amino acid residues, with each residue joined to the next by a peptide bond. The lengths for different proteins range from a few dozen to thousands of amino acids, and each protein contains different relative proportions of the 20 standard amino acids.

II.1. Aliphatic amino acids (leucine, isoleucine, alanine, methionine, valine);

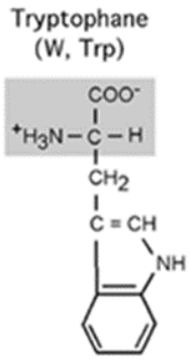
The R groups of these amino acids is made up only of carbon and hydrogen atoms. This makes them hydrophobic (apolar AA).



- Valine, leucine and isoleucine cannot be biosynthesized by the body. They are therefore part of the essential AAs (AAs which cannot be synthesized by the body; they must be provided by food).
- Alanine is an AA very common in proteins, its radical R is a methyl group.
- Methionine is one of the two amino acids that possess a sulfur atom. Methionine plays a central role in protein biosynthesis (translation) as it is almost always the initiating amino acid. It is one of the essential AAs. Its radical R is nonpolar (hydrophobic).

II.2. Aromatic amino acids

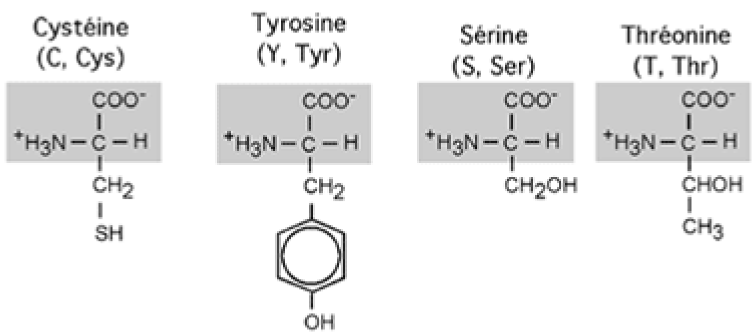
The R groups of these amino acids have aromatic groups. This makes them hydrophobic. The chemical structures of this amino acids group are:



- Tryptophan contains an indole ring attached to the alanyl side chain. It is one of the essential AAs.
- Phenylalanine, as the name implies, consists of a phenyl group attached to alanine. It is one of the essential AAs.
- Tyrosine obtained by hydroxylation of phenylalanine, its side chain becomes hydrophilic. These two amino acids are important because they serve as precursors for the biosynthesis of cathecolamines (adrenaline and norepinephrine). Tyrosine participates in the formation of thyroid hormones.

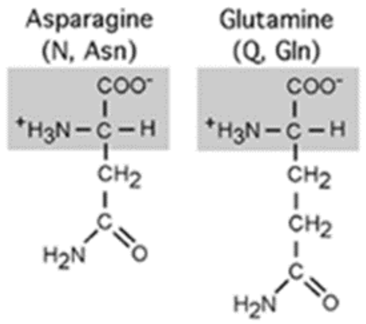
II.3. Hydroxyl/Sulfhydryl (threonine, serine, tyrosine, cysteine);

- Two amino acids, serine and threonine, contain aliphatic hydroxyl groups (that is, an oxygen atom bonded to a hydrogen atom, represented as —OH).
- Tyrosine possesses a hydroxyl group in the aromatic ring, making it a phenol derivative. The alcohol function makes their side chain polar (hydrophilic).
- Threonine is also one of the essential AAs.
- Cysteine has a thiol function. It makes the side chain polar (hydrophilic). It is an important AA because it contributes to the stabilization of the tertiary structure of proteins through the formation of disulfide bonds.



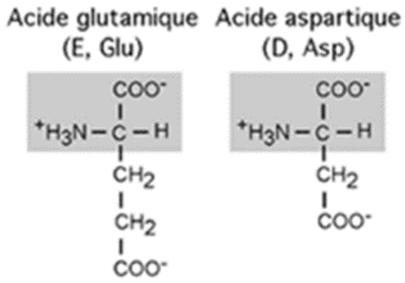
II.4. Carboxyamide (glutamine, asparagine);

Amide amino acids have an amide function at the radical level. Non-ionizable but polar; the side chain has a hydrophilic behavior.



II.5. (glutamic acid, aspartic acid);

The two amino acids in this group are aspartic acid and glutamic acid. Each has a carboxylic acid on its side chain that gives it acidic (proton-donating) properties. In an aqueous solution at physiological pH, all three functional groups on these amino acids will ionize, thus giving an overall charge of −1. The chemical structures of R-Acids amino acids group are:

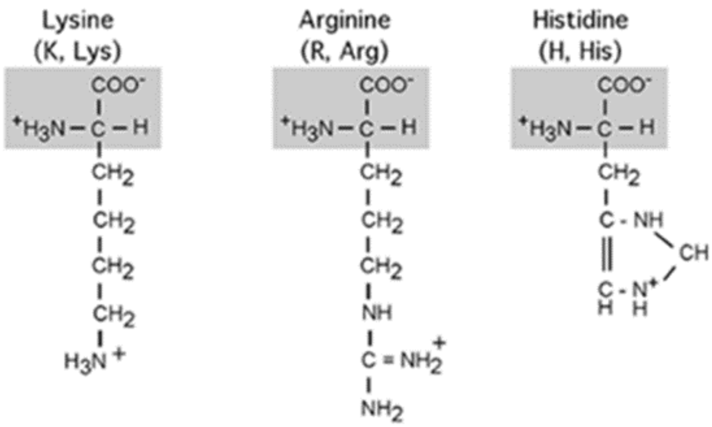


Glutamate is the most abundant excitatory neurotransmitter in the central nervous system.

II.6. R-Amines (lysine, histidine, arginine);

The three amino acids in this group are arginine, histidine, and lysine. Each side chain is basic. they exist with an overall charge of +1 at physiological pH.

- Histidine is an amino acid that most often makes up the active sites of protein enzymes, it is considered an essential AA in children.
- Arginine and histidine are present in significant quantities in histones (basic proteins allowing the condensation of DNA in the nucleus).
- Lysine is one of the essential AAs and Arginine is essential in infants.



II.7. Odd (glycine, proline).

In addition to being part of the composition of proteins, glycine participates in detoxification processes in the liver or even in the formation of bile salts.

Proline is unique among the standard amino acids in that it does not have both free α-amino and free α-carboxyl groups. Instead, its side chain forms a cyclic structure as the nitrogen atom of proline is linked to two carbon atoms.