

## Protein

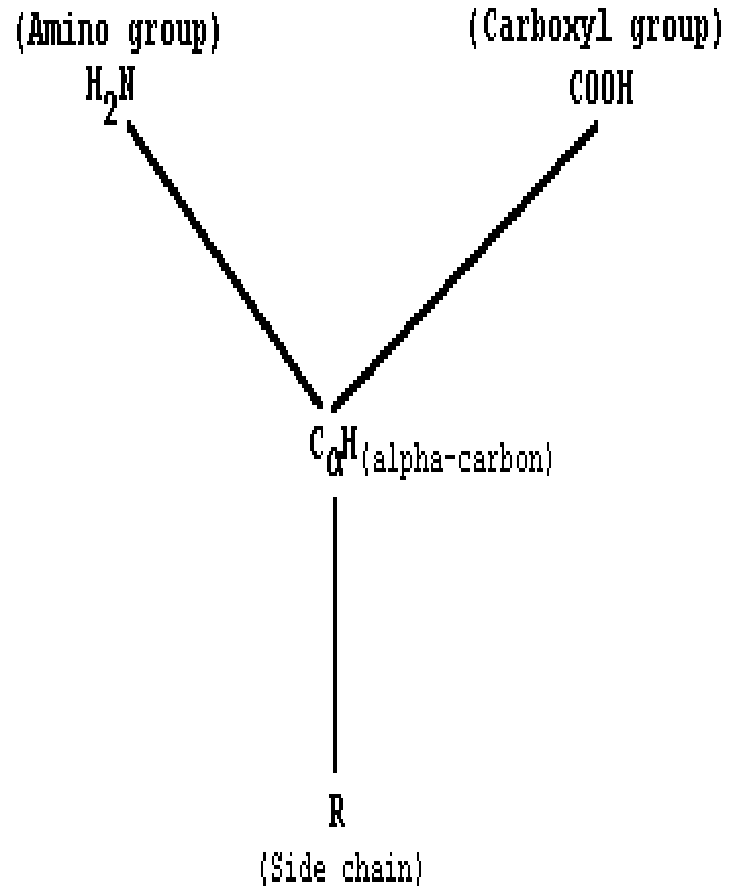
1

- Proteins are an important class of biological macromolecules present in all biological organisms, made up of such elements as carbon, hydrogen, nitrogen, phosphorus, oxygen, and sulfur.
- All proteins are polymers of amino acids.
- The polymers, also known as polypeptides consist of a sequence of 20 different L- $\alpha$ -amino acids, also referred to as residues.
- For chains under 40 residues the term peptide is frequently used instead of protein.
- To be able to perform their biological function, proteins fold into one, or more, specific spatial conformations, driven by a number of noncovalent interactions such as hydrogen bonding, ionic interactions, Van der Waals' forces and hydrophobic packing

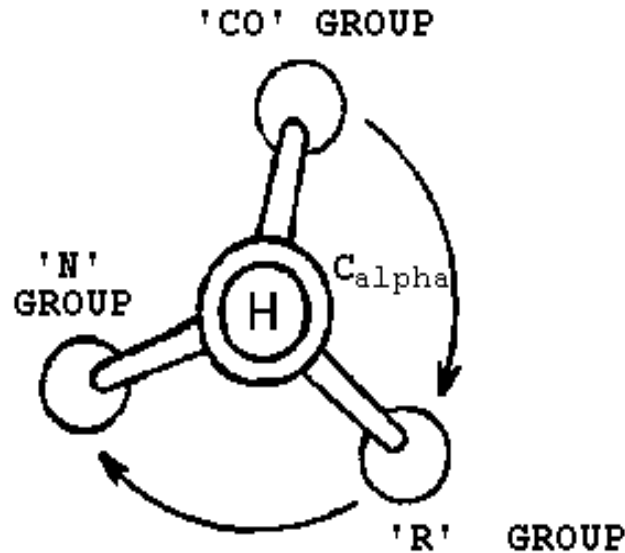
- The commonly occurring amino acids are of 20 different kinds which contain the same dipolar ion group  $\text{H}_3\text{N}^+\text{CHCOO}^-$ .
- They all have in common a central carbon atom to which are attached a hydrogen atom, an amino group ( $\text{NH}_2$ ) and a carboxyl group ( $\text{COOH}$ ).
- The central carbon atom is called the C alpha-atom and is a chiral centre.
- All amino acids found in proteins encoded by the genome have the L-configuration at this chiral centre.

# Protein

3



- This configuration can be remembered as the **CORN** law. Imagine looking along the H-Calpha bond with the H atom closest to you

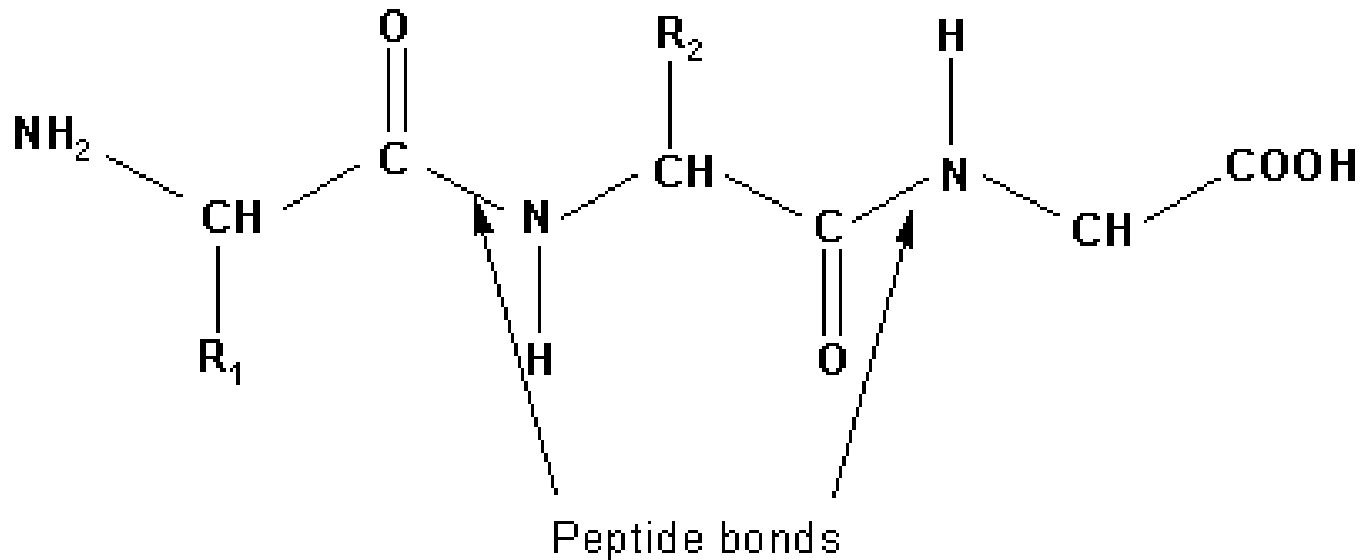


When read clockwise, the groups attached to the Calpha spell the word CORN (Richardson, 1981)

- There are 20 side chains found in proteins encoded by the genetic machinery of the cell.
- The side chains confer important properties on a protein such as the ability to bind ligands and catalyse biochemical reactions.
- They also direct the folding of the nascent polypeptide and stabilise its final conformation. In molecular graphics, atoms can be represented in different ways.
- Amino acids combine to make proteins what is meant by the primary, secondary and tertiary structures of proteins.

- Biochemistry refers to four distinct aspects of a protein's structure:
- **Primary structure** - the amino acid sequence of the peptide chains.
- **Secondary structure** - highly regular sub-structures (*alpha helix* and *strands of beta sheet*) which are locally defined, meaning that there can be many different secondary motifs present in one single protein molecule.
- **Tertiary structure** - Three-dimensional structure of a single protein molecule; a spatial arrangement of the secondary structures.
- **Quaternary structure** - complex of several protein molecules or polypeptide chains, usually called protein subunits in this context, which function as part of the larger assembly or protein complex.

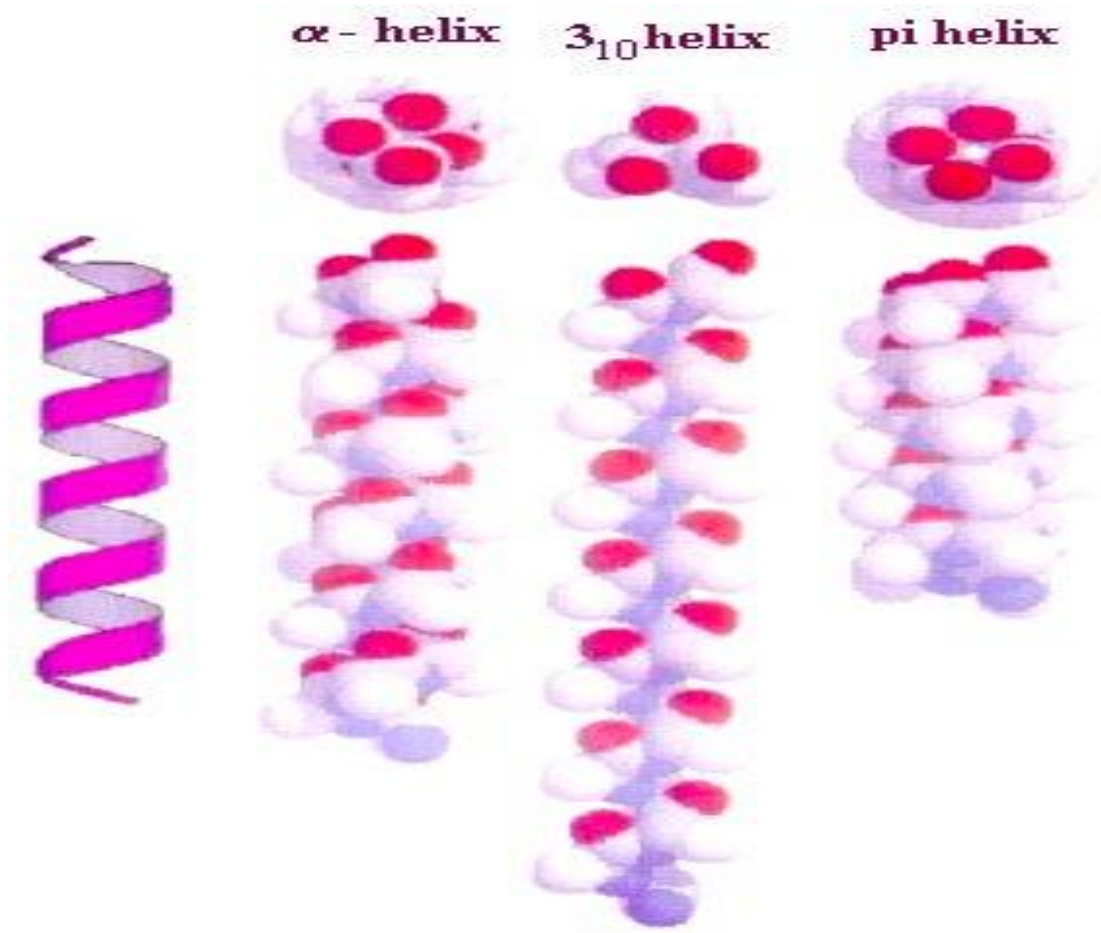
- Amino acids in proteins(or polypeptides) are joined together by peptide bonds. The sequence of R-groups along the chain is called the **primary structure**.



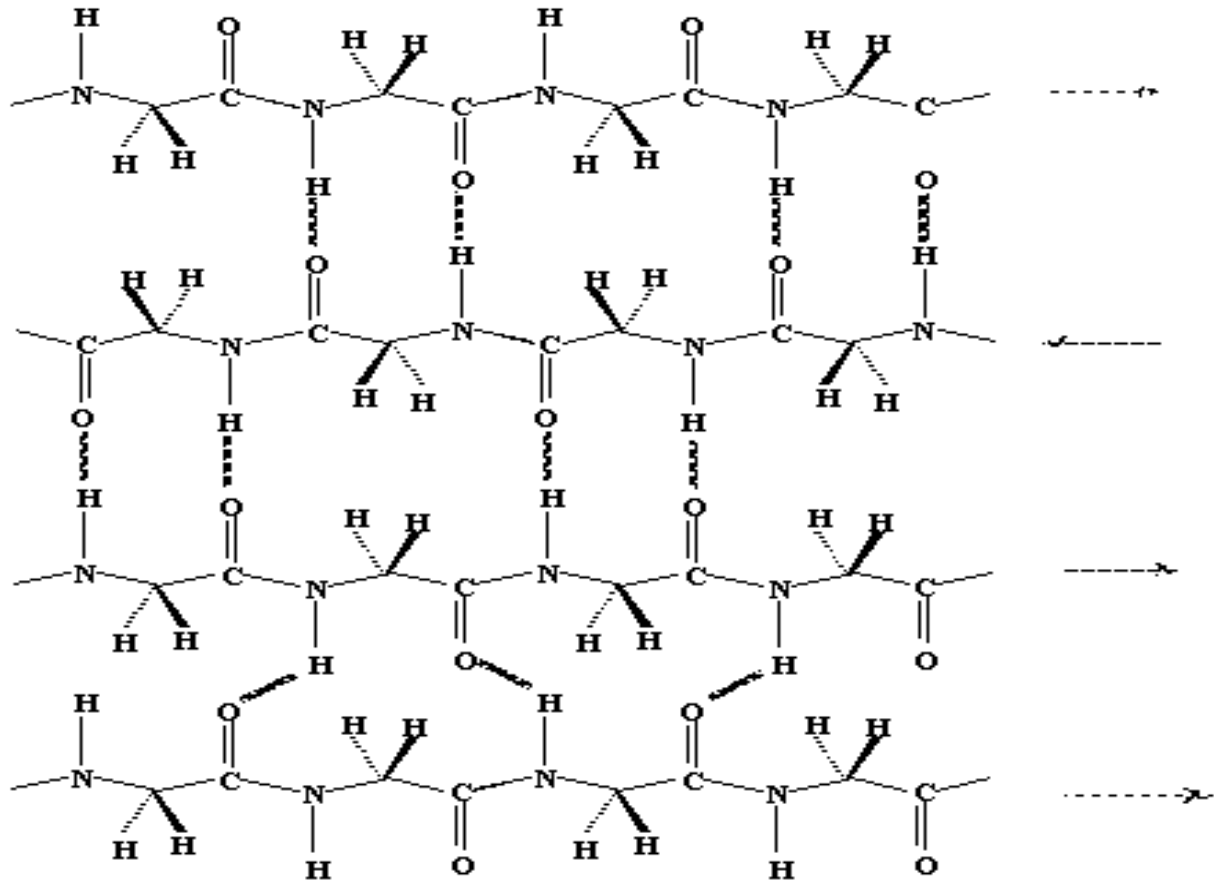
- The sequence of the different amino acids is called the primary structure of the peptide or protein.
- Counting of residues always starts at the N-terminal end (NH<sub>2</sub>-group), which is the end where the amino group is not involved in a peptide bond.
- The primary structure of a protein is determined by the gene corresponding to the protein. A specific sequence of nucleotides in DNA is transcribed into mRNA, which is read by the ribosome in a process called translation.
- The sequence of a protein is unique to that protein, and defines the structure and function of the protein. The sequence of a protein can be determined by methods such as Edman degradation or tandem mass spectrometry.
- Often however, it is read directly from the sequence of the gene using the genetic code.



- The secondary structure of a segment of polypeptide chain is the local spatial arrangement of its main-chain atoms without regard to the conformation of its side chains or to its relationship with other segments.
- There are three common secondary structures in proteins, namely **alpha helices**, **beta sheets** and **turns**.
- The alpha-helix and beta-structure conformations for polypeptide chains are generally the most thermodynamically stable of the regular secondary structures
- CO groups form hydrogen bonds with NH groups of 3 residues along the chain forming a  $3_{10}$  helix. A substantial amount of all  $3_{10}$  helices occur at the ends of alpha-helices.
- They are called  $3_{10}$  because there are 3 residues per turn and 10 atoms enclosed in a ring formed by each hydrogen bond. Dipoles are not aligned as in a normal right-handed alpha-helix.



- $\beta$  Pleats
- The hydrogen bond network in a 2-stranded, antiparallel  $\beta$ -sheet.
- The side chains are sticking out above or below the plane of the picture.
- It less clear cut than in the case of the helix, in which direction to initially trace a beta sheet strand.
- The beta sheet can be infinitely extended due to the repeatable H-bonding pattern to either side of a strand.
- Turns, loops and a few other secondary structure elements such as a 3-10 helix complete the picture
- The beta-sheet is sometimes called the beta pleated sheet since sequential neighbouring C alpha atoms are alternately above and below the plane of the sheet giving a pleated appearance.
- Beta-sheets are found in two forms designated as "Antiparallel" or "Parallel" based on the relative directions of two interacting beta strand



- The alpha-helix and beta-structure conformations for polypeptide chains are generally the most thermodynamically stable of the regular secondary structures.
- However, particular amino acid sequences of a primary structure in a protein may support regular conformations of the polypeptide chain other than alpha-helical or beta-structure.
- Thus, whereas alpha-helical or beta-structure are found most commonly, the actual conformation is dependent on the particular physical properties generated by the sequence present in the polypeptide chain and the solution conditions in which the protein is dissolved

- The elements of secondary structure are usually folded into a compact shape using a variety of loops and turns.
- The formation of tertiary structure is usually driven by the burial of hydrophobic residues, but other interactions such as hydrogen bonding, ionic interactions and disulfide bonds can also stabilize the tertiary structure.
- The tertiary structure encompasses all the noncovalent interactions that are not considered secondary structure, and is what defines the overall fold of the protein, and is usually indispensable for the function of the protein.

- The quaternary structure is the interaction between several chains of peptide bonds.
- The individual chains are called subunits.
- The individual subunits are not necessarily covalently connected, but might be connected by a disulfide bond.
- Not all proteins have quaternary structure, since they might be functional as monomers.
- The quaternary structure is stabilized by the same range of interactions as the tertiary structure.
- Complexes of two or more polypeptides (i.e. multiple subunits) are called multimers.
- Specifically it would be called a dimer if it contains two subunits, a trimer if it contains three subunits, and a tetramer if it contains four subunits.

- Multimers made up of identical subunits may be referred to with a prefix of "homo-" (e.g. a homotetramer) and those made up of different subunits may be referred to with a prefix of "hetero-" (e.g. a heterodimer).
- Tertiary structures vary greatly from one protein to another. They are held together by glycosydic and covalent bonds