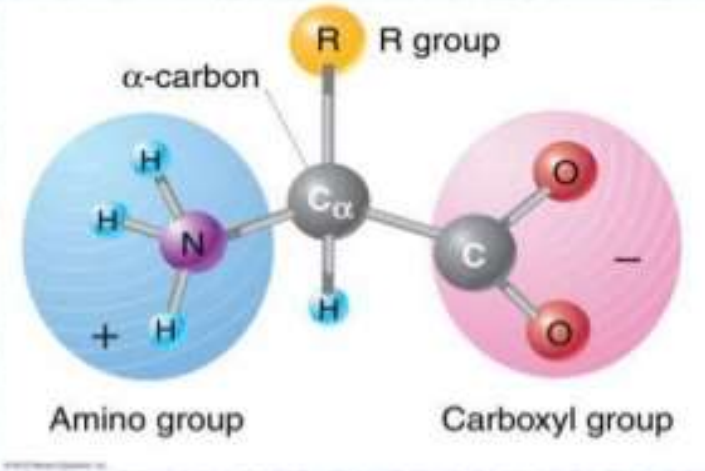
Protein Structure

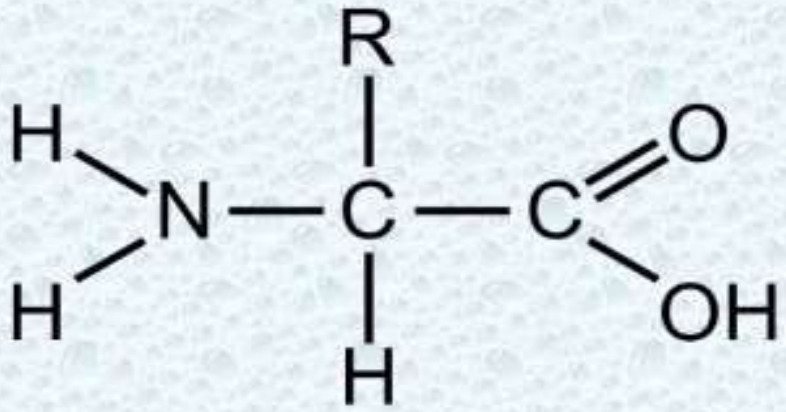
AMINO ACIDS

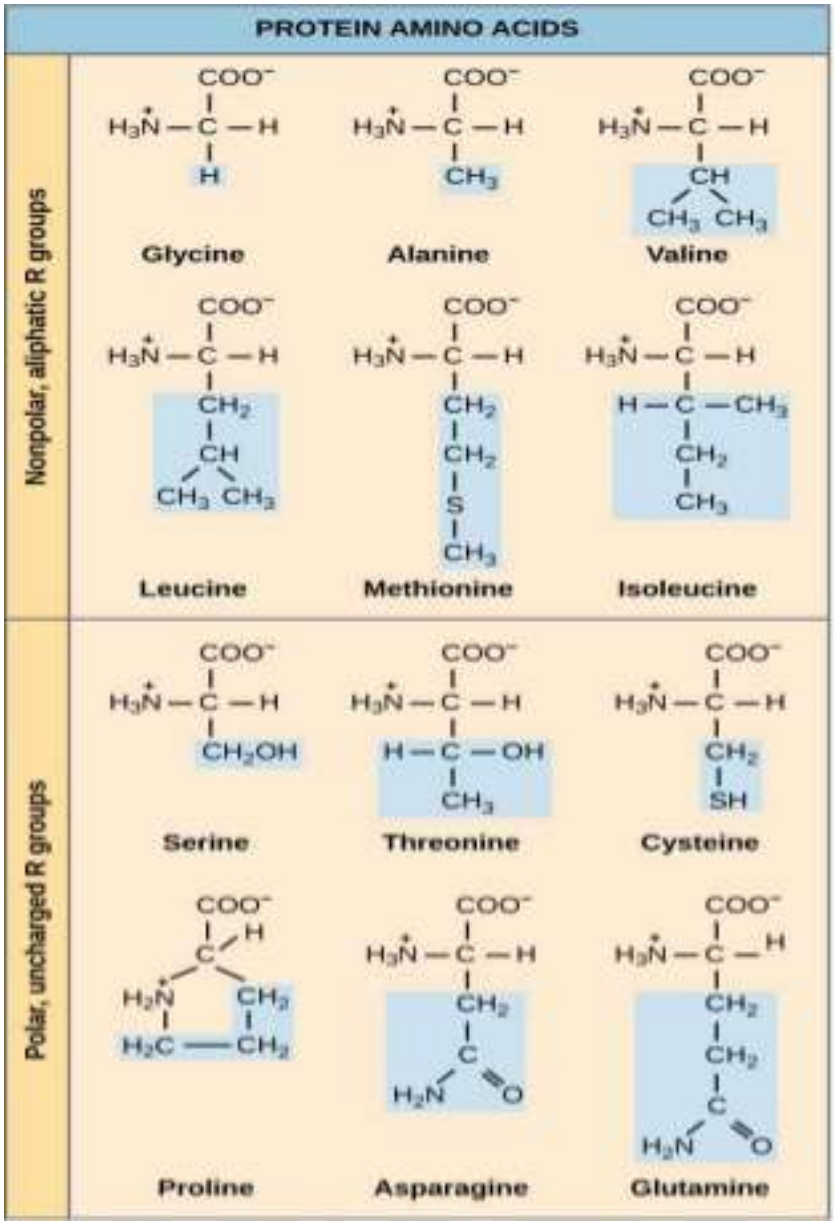
- Amino acids are a group of organic compounds containing two functional groups – amino and carboxyl.
- The amino group[$-NH_2$] is basic while the carboxyl group [-COOH] is acidic in nature.
- There are about 300 aminocids occur in nature. Only 20 of them occur in proteins.

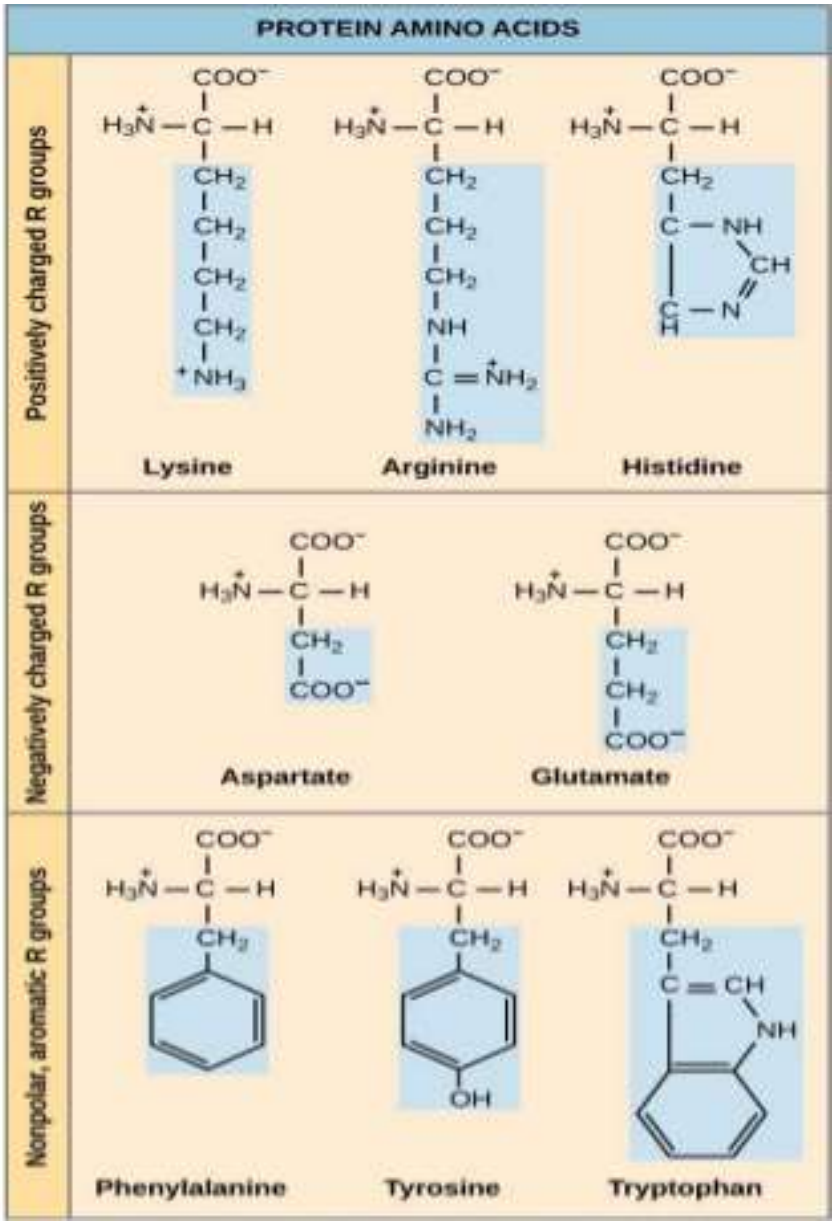


Structure of amino acids

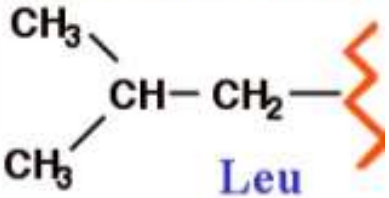
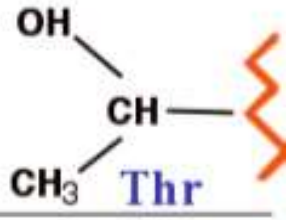
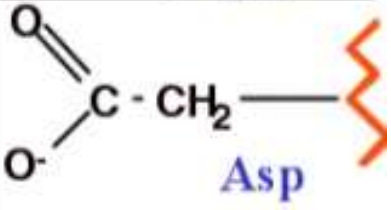
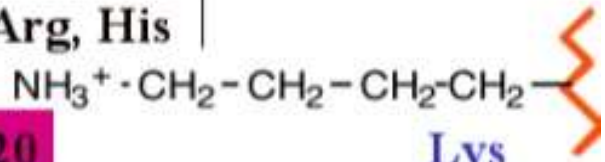
- Each amino acid has 4 different groups attatched to α-carbon (which is C atom next to COOH).These 4 groups are : amino group, COOH group , Hydrogen atom and side chain(R).




PROTEIN AMINO ACIDS
Nonpolar, aliphatic R groups
Glycine
Alanine
Valine
Leucine
Methionine
Isoleucine
Polar, uncharged R groups
Serine
Threonine
Cysteine
Proline
Asparagine
Glutamine


PROTEIN AMINO ACIDS
Positively charged R groups
Lysine
Arginine
Histidine
Negatively charged R groups
Aspartate
Glutamate
Nonpolar, aromatic R groups
Phenylalanine
Tyrosine
Tryptophan

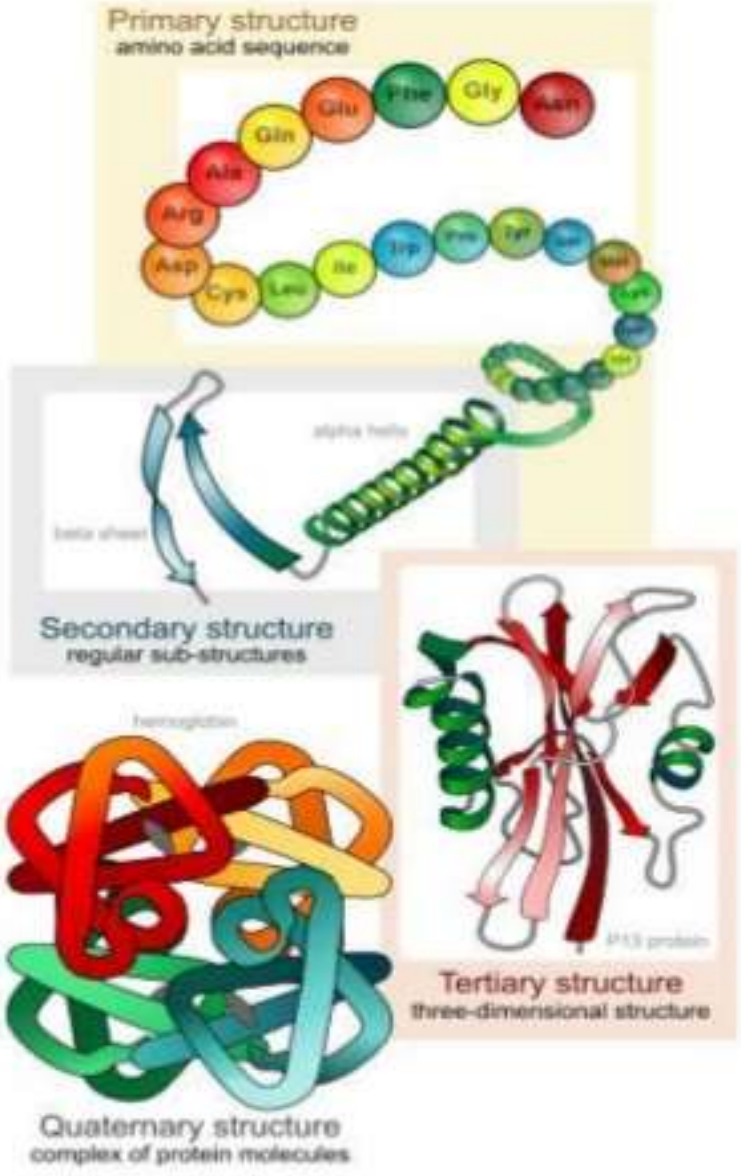
Amino acids groups

Group	Characteristics	Names	Example (-Rx)
non-polar	hydrophobic	Ala, Val, Leu, Ile, Pro, Phe Trp, Met	 Leu
polar	hydrophilic (non-charged)	Gly , Ser, Thr, Cys, Tyr, Asn Gln	 Thr
acidic	negatively charged	Asp, Glu	 Asp
basic	positively charged	Lys, Arg, His	 Lys

Total = 20

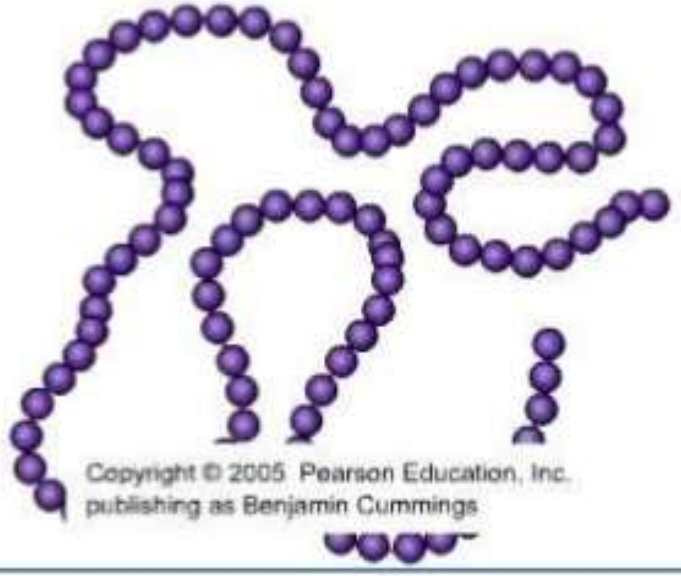
INTRODUCTION

- Proteins are an important class of biological macromolecules which are the polymers of amino acids.
- Biochemists have distinguished several levels of structural organization of proteins. They are:
  - Primary structure
  - Secondary structure
  - Tertiary structure
  - Quaternary structure



PRIMARY STRUCTURE

- The primary structure of protein refers to the sequence of amino acids present in the polypeptide chain.
- Amino acids are covalently linked by peptide bonds.
- Each component amino acid in a polypeptide is called a "residue" or "moiety"
- By convention, the 1^0 structure of a protein starts from the amino-terminal (N) end and ends in the carboxyl-terminal (C) end.



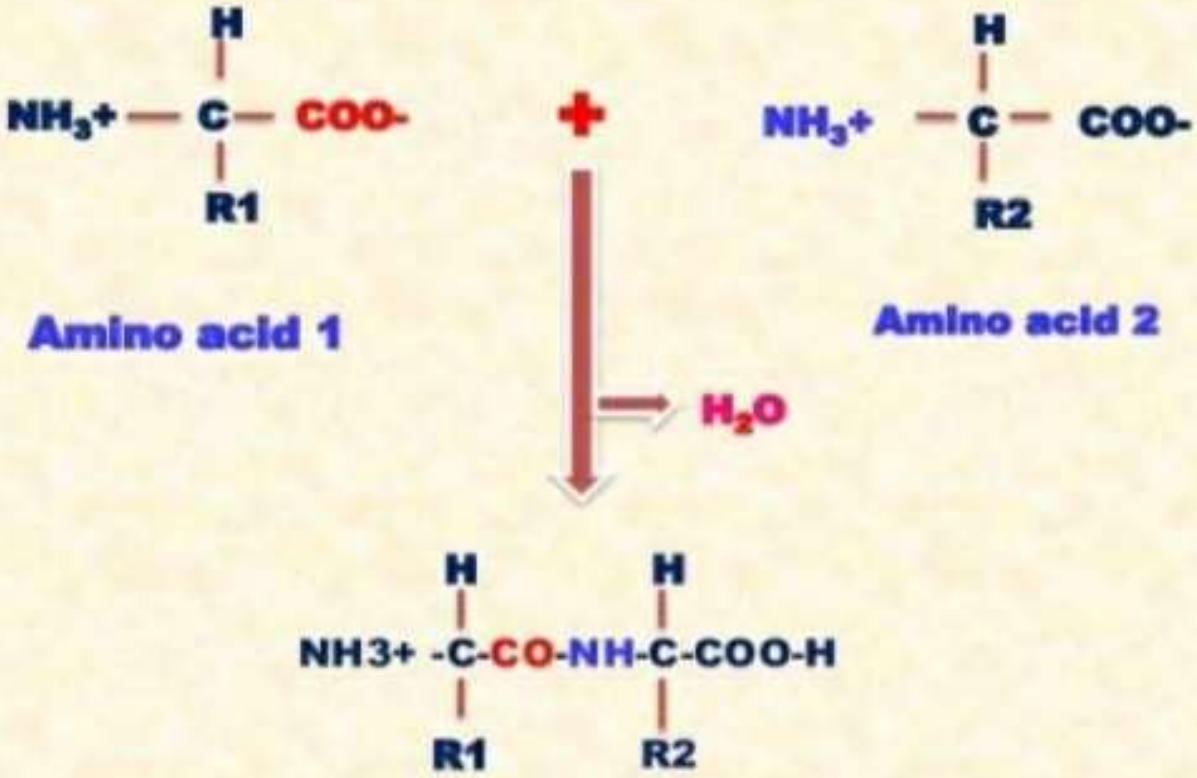
IMPORTANCE OF PRIMARY STRUCTURE

- To predict 2^0 and 3^0 structures from sequence homologies with related proteins. (Structure prediction)
- Many genetic diseases result from abnormal amino acid sequences.
- To understand the molecular mechanism of action of proteins.
- To trace evolutionary paths.

METHODS OF AMINO ACID SEQUENCE DETERMINATION

- End group analysis – Edman degradation.
- Gene sequencing method.

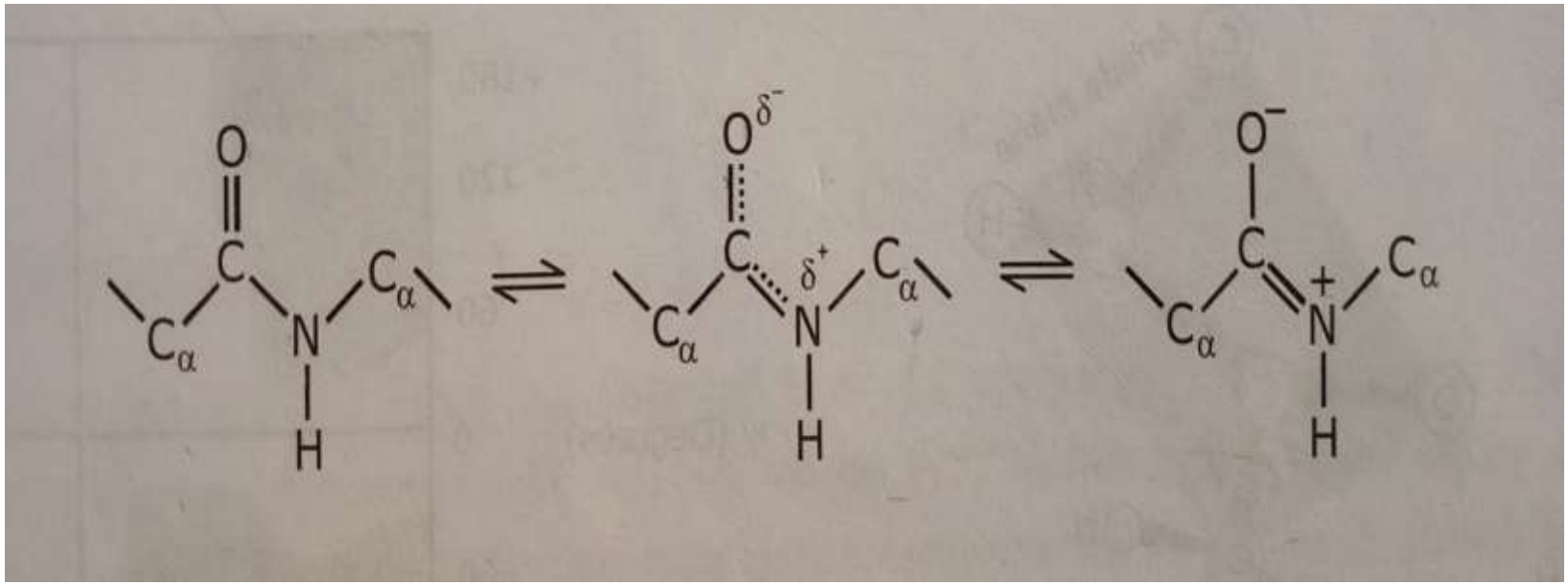
Formation of peptide bond



Peptide Bond

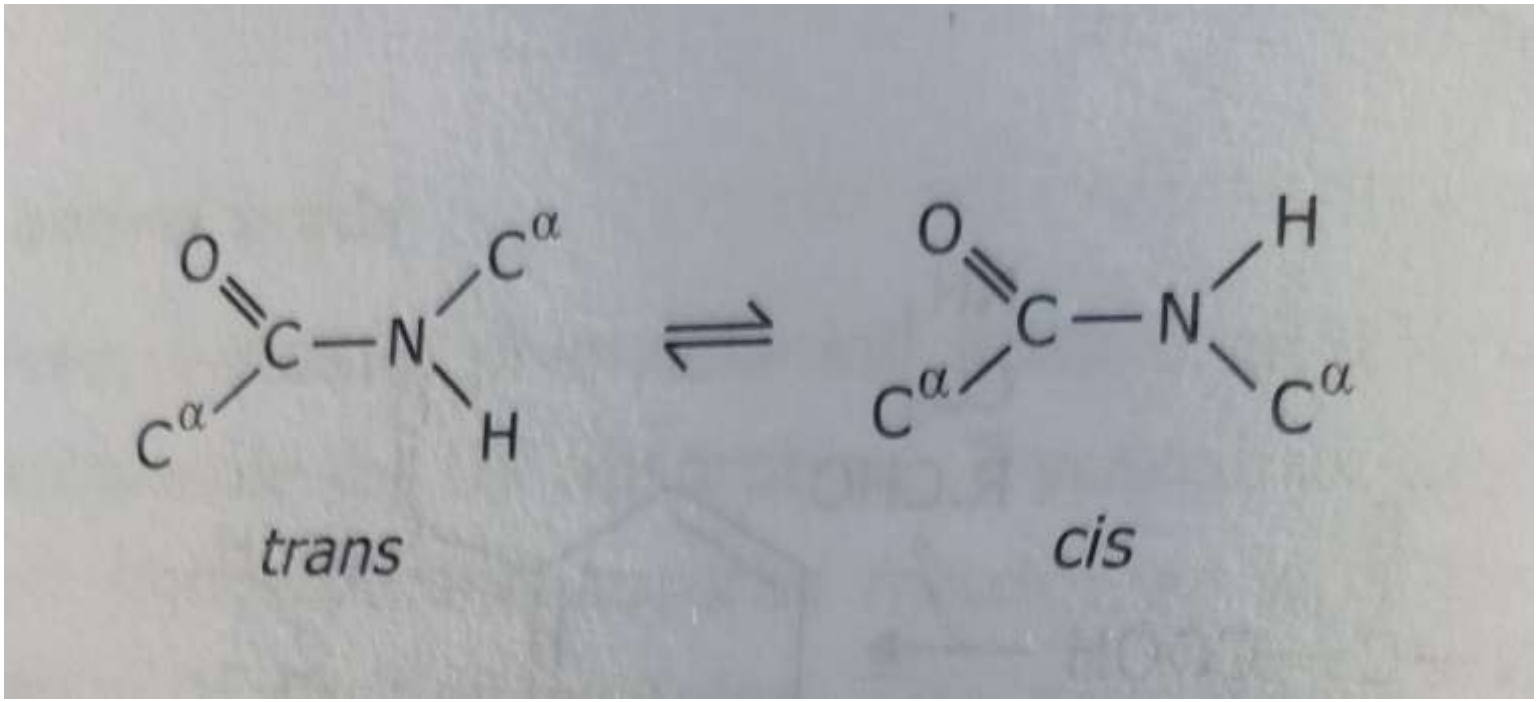
- It has a partial double bond character (40%) that keeps the entire 6-atom peptide group in a rigid planar configuration
- The peptide bond length is only 1.33Å, shorter than the usual C-N bond length of 1.45Å

Resonance of Peptide Bond



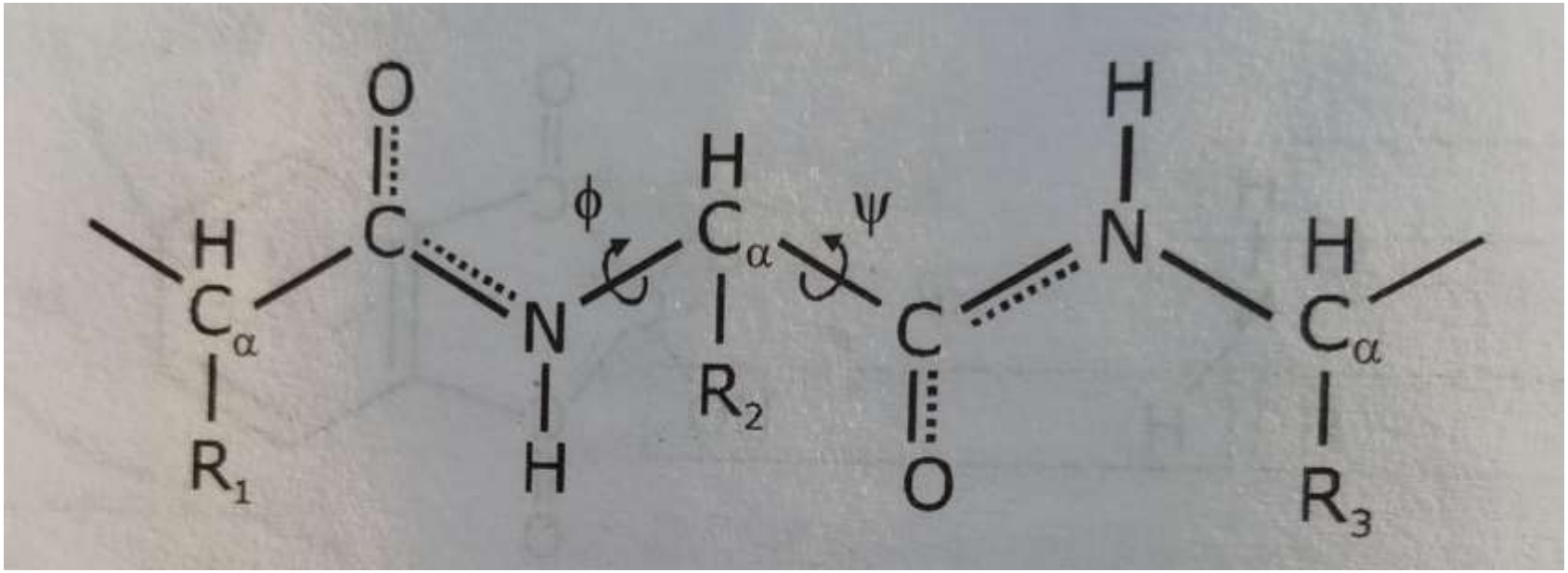
- The angle of rotation around the peptide bond (ω) has the value 180° (trans) and occasionally 0° (cis)
- There are two possible conformations of the planar peptide bond – 1) Trans- the C_α atoms are on opposite sides of the peptide bond & 2) Cis- the C_α atoms are on the same side of the peptide bond

Trans- & Cis- form of Peptide Bond



- The trans- form is favoured over the cis- form because of steric hindrance. In the cis- form the C_α atom and the side chains of neighbouring residues are in too close proximity
- Rotation is permitted about the N - C_α and the C_α – C bonds. Rotation about bonds are described as Torsion / Dihedral / Conformational angle
- By convention, Torsion angle for the N - C_α bond is called Φ (phi) and for the bond C_α – C is called Ψ (psi)

Torsion / Dihedral / Conformational angle

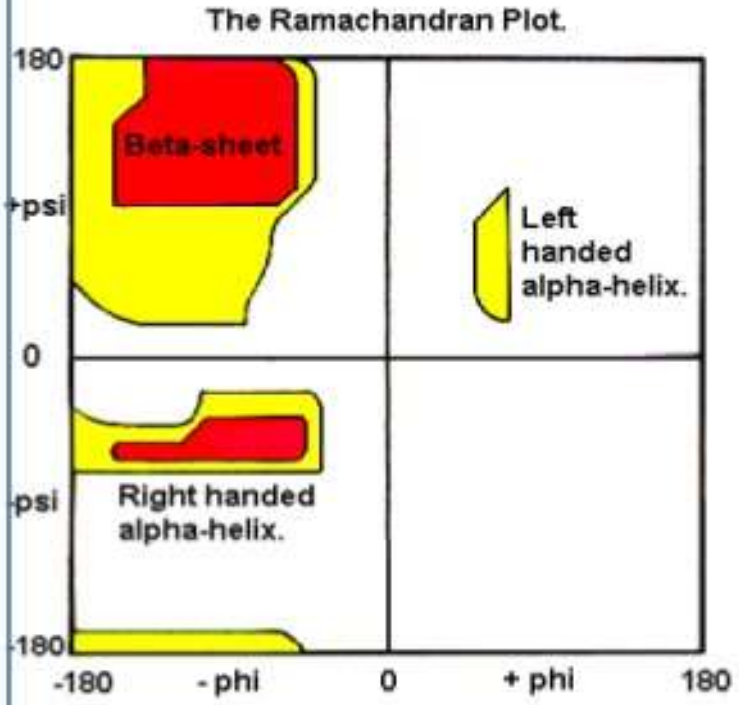


Ramachandran Plot

- Φ and Ψ can have any value between -180° and +180° but many values are prohibited by steric interference between atoms in the polypeptide backbone and amino acid side chains
- The permitted values for Φ and Ψ were first determined by G.N.Ramachandran
- The permitted values are indicated on a 2-D map of the Φ – Ψ plane, known as Ramachandran Plot
- It allows us to identify those conformations (for a particular value of Φ and Ψ) that are sterically favourable or unfavourable

RAMACHANDRAN PLOT

A **Ramachandran plot** (also known as a **Ramachandran diagram** or a **[φ,ψ] plot**), originally developed in 1963 by G. N. Ramachandran.



White regions : Sterically disallowed for all amino acids except glycine.

Red regions : allowed regions namely the a-helical and b-sheet conformations.

Yellow areas : outer limit

SECONDARY STRUCTURE

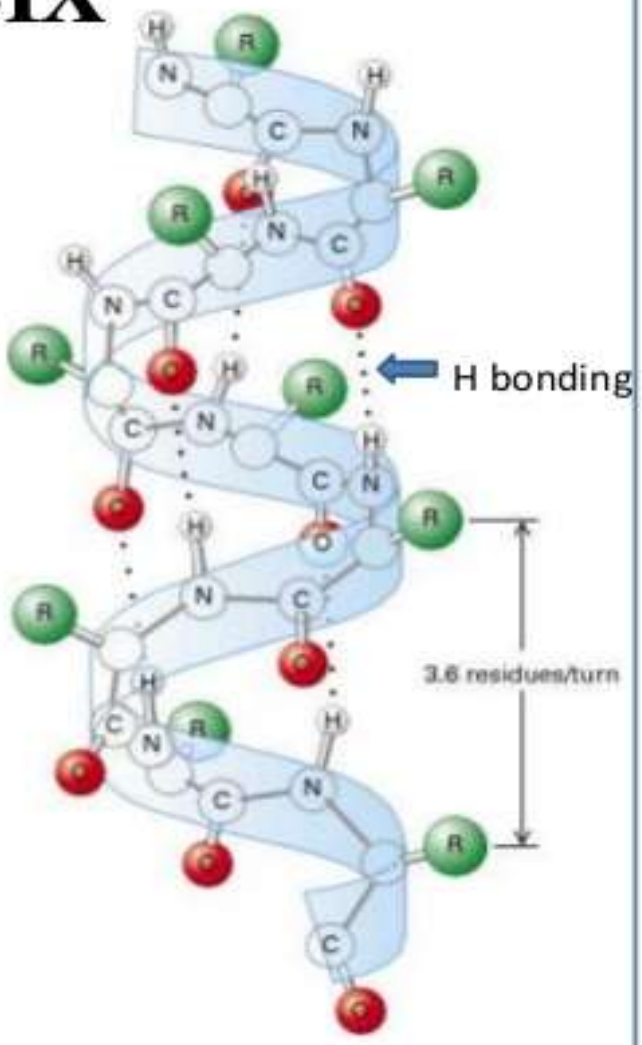
- Localized arrangement of adjacent amino acids formed as the polypeptide chain folds.
- It consists of
  - **α-helix**
  - **β-pleated sheet**
  - **β-bends**
  - **Non repetitive structures**
  - **Super secondary structures**
- Linus Pauling proposed some essential features of peptide units and polypeptide backbone. They are:
  - The amide group is rigid and planar as a result of resonance. So rotation about C-N bond is not feasible.
  - Rotation can take place only about N- C_α and C_α – C bonds.
  - Trans configuration is more stable than cis for R grps at C_α
- From these conclusions Pauling postulated 2 ordered structures **α** helix and **β** sheet

- Secondary structures spontaneously form as an intermediate before the protein folds into its three dimensional tertiary structure
- Secondary structures are stabilized by hydrogen bonds between the carbonyl and N-H groups in the polypeptides backbone
- Amino acids vary in their ability to form various secondary structures
- Methionine, Alanine, Leucine, Glutamate and Lysine prefer to adopt helical conformations in proteins

- Tryptophan, Tyrosine, Phenylalanine, Isoleucine, Valine and Threonine prefer to adopt beta strand conformations
- Proline and Glycine are known as “helix breakers” as they disrupt the regularity of the α-helical backbone conformation. These amino acids are commonly found in turns

ALPHA HELIX

- Spiral structure
- Tightly packed, coiled polypeptide backbone core.
- Side chain extend outwards
- Stabilized by H bonding b/w carbonyl oxygen and amide hydrogen.
- Amino acids per turn – 3.6
- Pitch is 5.4 A
- Alpha helical segments are found in many globular proteins like myoglobins, troponin- C etc.



- α-helix is a rigid, rod like structure that forms when a polypeptide chain twists (clockwise/right-handed or counterclockwise/ left-handed) into a helical conformation
- However, right-handed helices are energetically more favourable
- There are 3.6 amino acid residues per turn of the helix and the pitch (the distance between corresponding points per turn) is 0.54 nm
- Each residue is related to the next one by a rise of 1.5Å (0.15 nm) along the helix axis

Right handed α-Helix

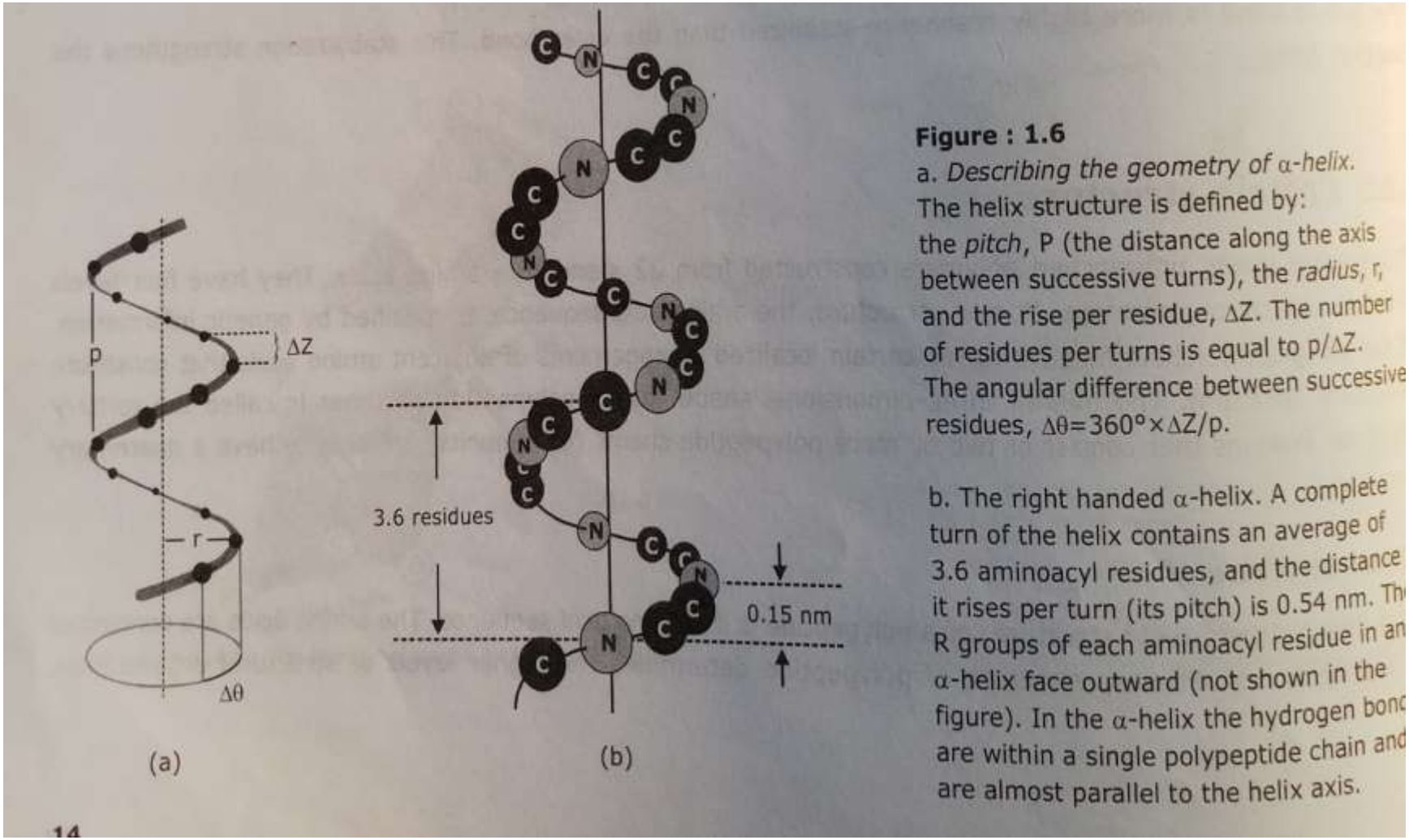


Figure : 1.6

a. *Describing the geometry of α-helix.* The helix structure is defined by: the *pitch*, P (the distance along the axis between successive turns), the *radius*, r, and the rise per residue, ΔZ. The number of residues per turns is equal to p/ΔZ. The angular difference between successive residues, $\Delta\theta = 360° \times \Delta Z/p$.

b. The right handed α-helix. A complete turn of the helix contains an average of 3.6 aminoacyl residues, and the distance it rises per turn (its pitch) is 0.54 nm. Th R groups of each aminoacyl residue in an α-helix face outward (not shown in the figure). In the α-helix the hydrogen bonc are within a single polypeptide chain and are almost parallel to the helix axis.

- A single turn of α-helix involves 13 atoms from O to the H of the hydrogen bond. For this reason, the α-helix is referred to as the 3.6_{13} helix
- Length of α-helix is usually 10-15 amino acid residues
- Intrachain hydrogen bonds form between the N-H group of each amino acids and the carbonyl group of the amino acid **4 residues away**
- Except for amino acids near the ends of an α-helix , all the main-chain CO and NH groups are hydrogen bonded
- The side chains of amino acids extend outward from the helix
- All the H-bonds lie parallel to the helix axis & point in the same direction

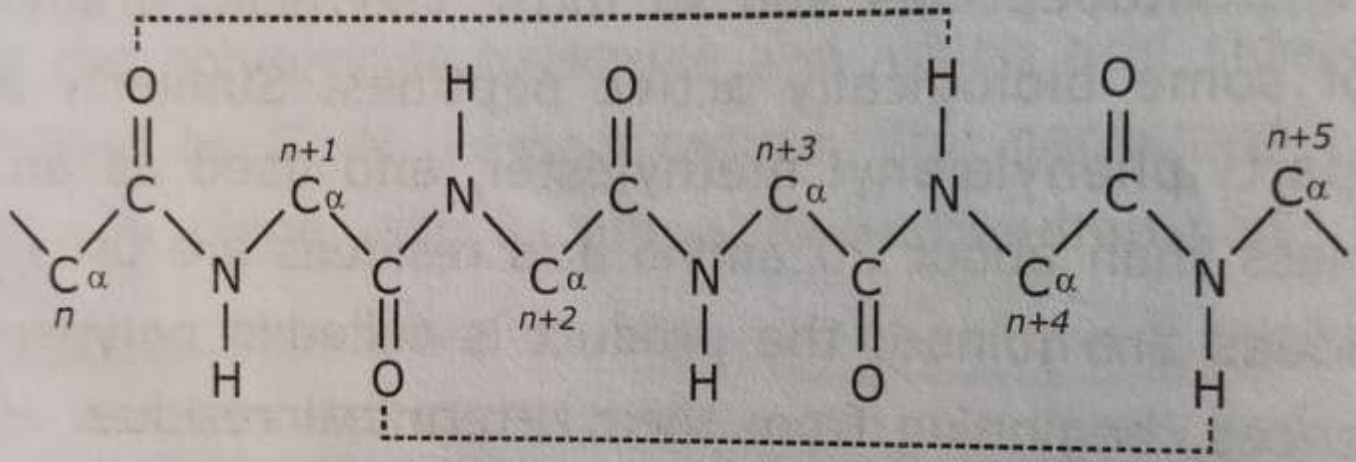


Figure 1.5 : The hydrogen bonding arrangement in the α-helix.

- Many fibrous proteins like hair, skin, nails consist almost entirely of α-helices
- Stability of α-helices is due to –
- The involvement of all the >NH and >C=O groups in the chain in H-bonding
- The way in which the side chains project outward from an α-helix. Bulky side chains therefore do not interfere with the H-bonding, enabling a fairly rigid cylinder to be formed

Types of helix

α-helix

Also called the 3.6_{12}

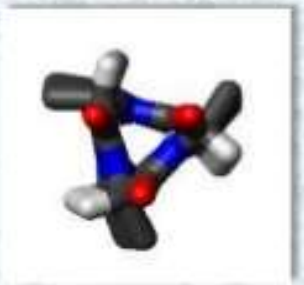
π-helix

Very loosely coiled H-bonding pattern n + 5

Rarely found in nature.

3_{10}-helix

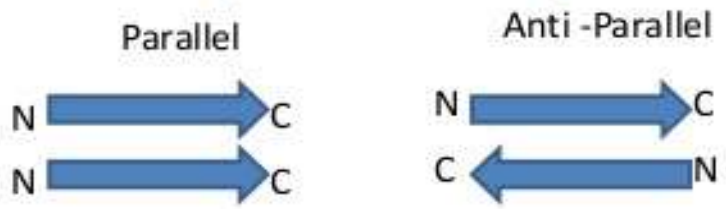
Very tightly coiled H-bonding pattern n+3 rarely found in nature

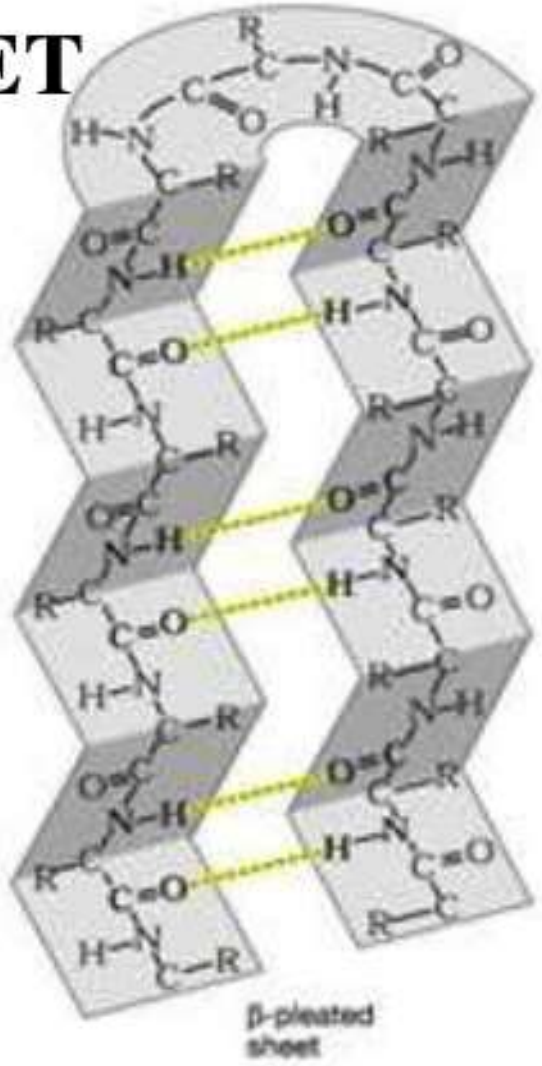


BETA PLEATED SHEET

- Formed when 2 or more polypeptides line up side by side.
- Individual polypeptide - β strand
- Each β strand is fully extended.
- They are stabilized by H bond b/w N-H and carbonyl grps of adjacent chains.

2 types





- β-pleated sheets are stabilized by interchain H-bonds that form between the polypeptide backbone N-H and carbonyl groups of adjacent strands
- Adjacent strands can be either parallel or antiparallel
- In parallel β-pleated sheet structures, the polypeptide chains are arranged in the same direction
- However, in antiparallel β-pleated sheet, chains run in opposite directions
- Antiparallel β-pleated sheets are more stable than parallel β-pleated sheets because fully colinear H-bonds form

Parallel & Antiparallel β-strands

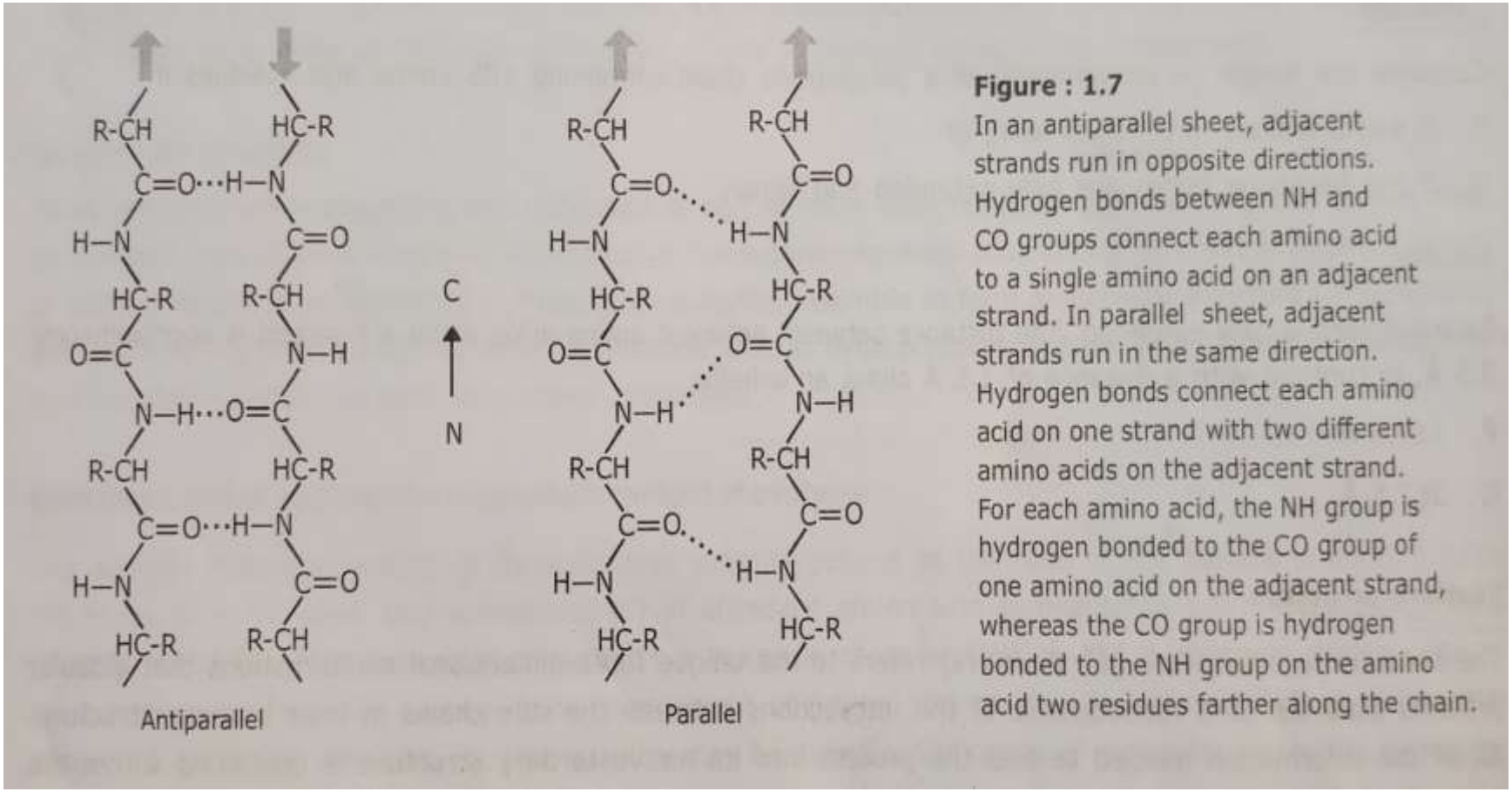
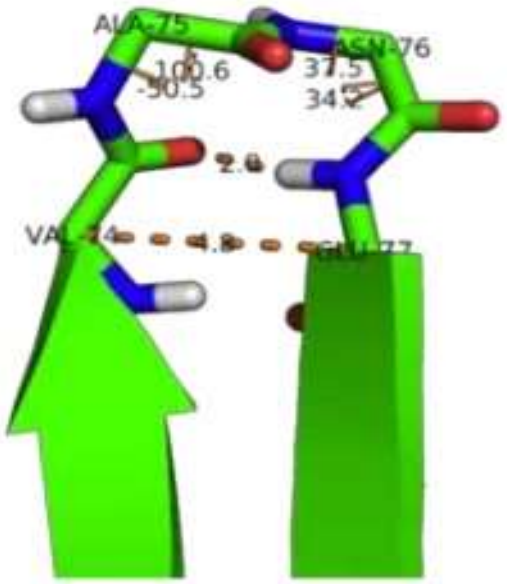


Figure : 1.7
In an antiparallel sheet, adjacent strands run in opposite directions. Hydrogen bonds between NH and CO groups connect each amino acid to a single amino acid on an adjacent strand. In parallel sheet, adjacent strands run in the same direction. Hydrogen bonds connect each amino acid on one strand with two different amino acids on the adjacent strand. For each amino acid, the NH group is hydrogen bonded to the CO group of one amino acid on the adjacent strand, whereas the CO group is hydrogen bonded to the NH group on the amino acid two residues farther along the chain.

- The secondary structure of silk is the β-pleated sheet
- The primary structure of silk contains the amino acids glycine, alanine & serine in specific repeating pattern. These 3 amino acids make up 90% of the protein in silk
- The 10% is comprised of amino acids glutamic acid, valine & aspartic acid. These are used as side chains and affect elasticity and strength
- The β-pleated sheet of silk is connected by H-bonds

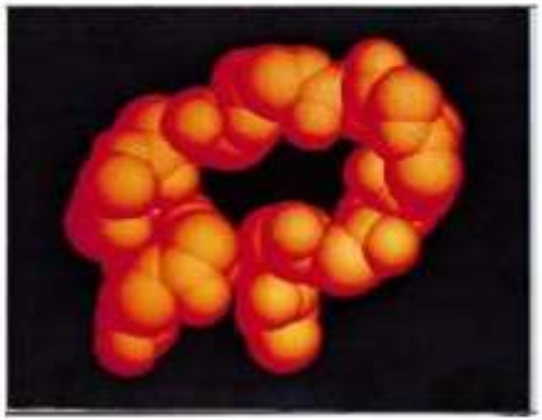
BETA BENDS

- Permits the change of direction of the peptide chain to get a folded structure.
- It gives a protein globularity rather than linearity.
- H bond stabilizes the beta bend structure.
- Proline and Glycine are frequently found in beta turns.
- Beta turns often promote the formation of antiparallel beta sheets.
- Occur at protein surfaces.
- Involve four successive aminoacid residues



NON REPETITIVE STRUCTURES

- A significant portion of globular protein's structure may be irregular or unique.
- They include coils and loops.
- Segments of polypeptide chains whose successive residues do not have similar ϕ and Ψ values are called **coils**.
- Almost all proteins with more than 60 residues contain one or more loops of 6 to 16 residues, called **Ω loops.**



Space-filling model of an **Ω** loop

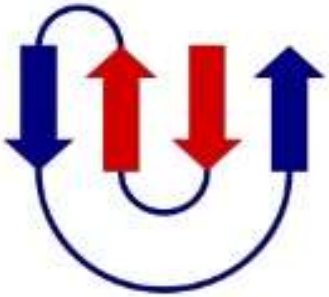
Supersecondary Structures

- Many globular proteins contain combinations of α-helix and β-pleated sheet secondary structures
- These patterns are called supersecondary structures or motifs

SUPER SECONDARY STRUCTURES (MOTIFS)

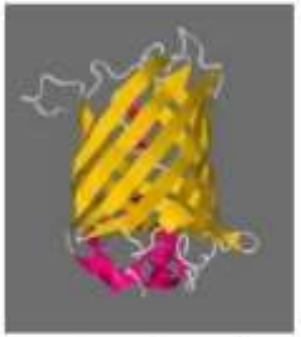
Certain groupings of secondary structural elements are called **motifs.**

beta-alpha-beta motif



Greek key motif

β-meander motif



Beta barrel

Experimental Determination

- Secondary structure content of a biopolymer can be estimated **spectroscopically**
- A common method is **far-ultraviolet (170-250 nm) circular dichroism**. A pronounced double minimum at 208 and 222 nm indicate α-helical structure, whereas a single minimum at 204 nm or 217 nm reflects random coil or beta sheet structure, respectively
- A less commnon method is **infrared spectroscopy** which detects differences in the bond oscillations of amide groups due to hydrogen bonding
- Secondary structure contents can be estimated accurately using the chemical shifts of an initially unassigned **NMR spectrum**

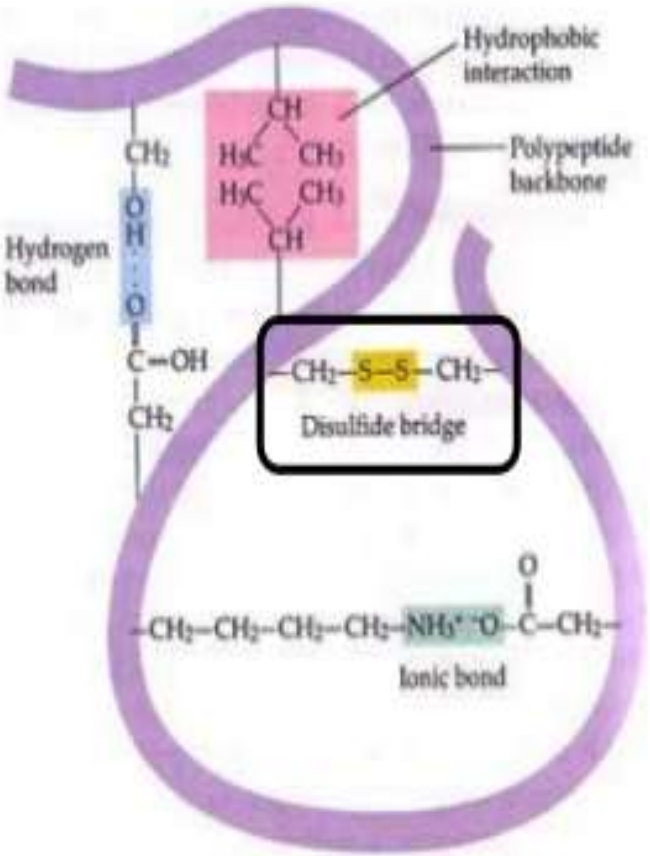
Tertiary Structure

- It is the unique 3D conformation of globular proteins due to the **interactions between the side chains** in their primary structure
- Interactions stabilizing tertiary structure are –
- Hydrophobic interactions (Disulfide bridges)
- Electrostatic interactions (Ionic bonds, e.g., salt bridge)
- Hydrogen bonds (weak ionic interaction)
- van der Waals force (Hydrophobic interaction)

- The fundamental unit of tertiary structure is the Domain
- It is a polypeptide chain or a part of a polypeptide chain that can fold independently into a stable 3D tertiary structure
- It is also a functional unit of protein
- Proteins may comprise a single domain or more than 1 domain
- A small protein often consists of only one domain

Covalent bond: Disulphide links

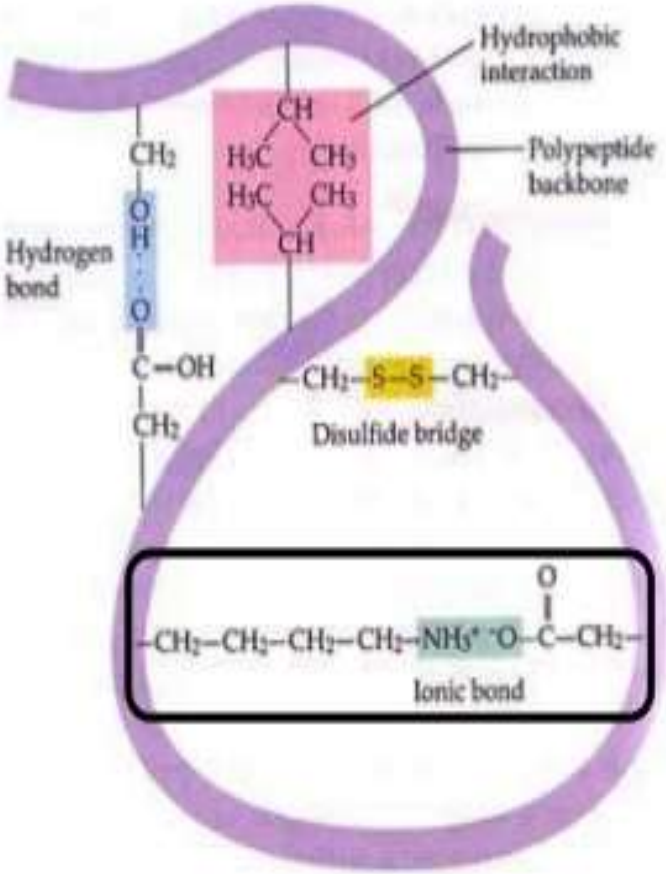
- Cysteine has a residue containing thiol group capable of forming a covalent bond in protein tertiary structure.
- When two such residues are close together, covalent Disulphide bond can be formed as a result of oxidation.
- The two cysteine residues involved in the bond formation may be far apart in the primary structure but are brought close together as a result of protein folding.



Ionic or electrostatic bond

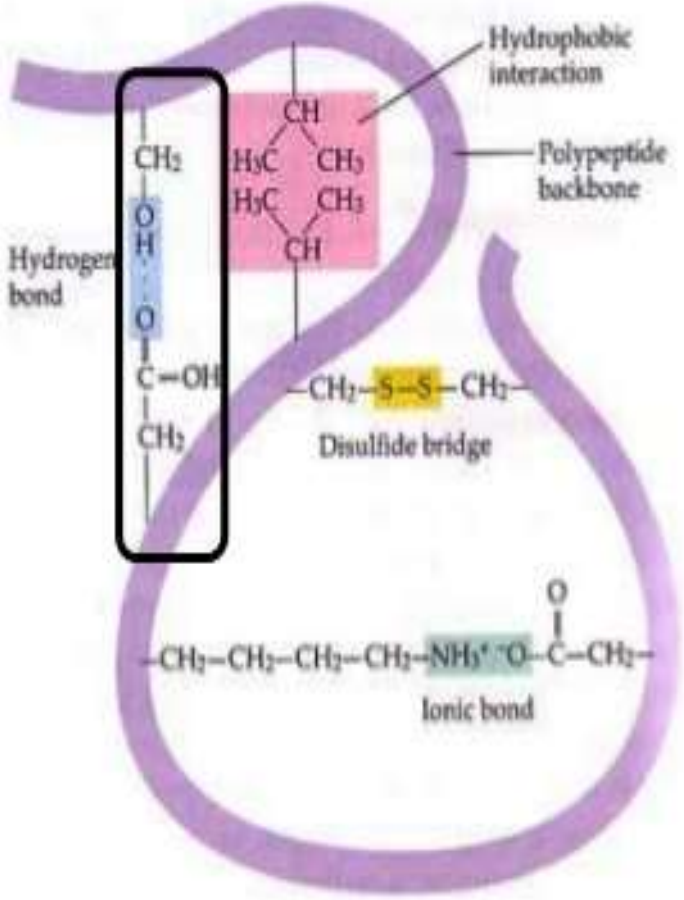
- An ionic bond or salt bridge can be formed between the carboxylate ion of an acidic residue such as aspartic acid or glutamic acid and an ammonium ion of a basic residue such as lysine, arginine or histidine.

- This is the strongest of the intramolecular bonds.



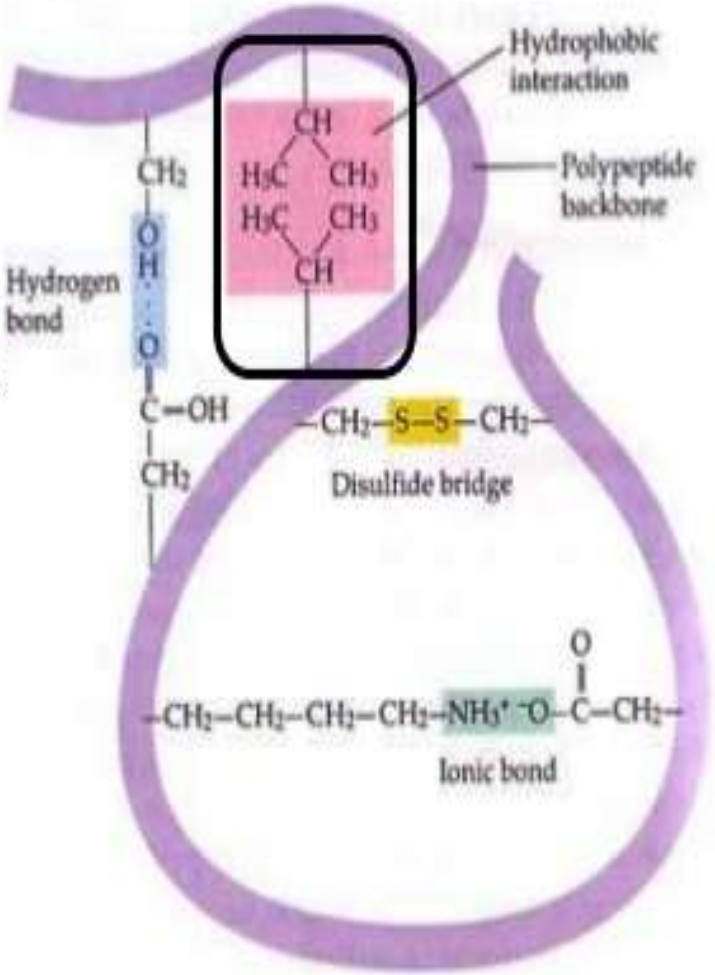
Hydrogen bond

- Weak form of ionic interaction: interaction between the atoms having partial charges.
- Can be formed between large number of amino acid residues: serine, threonine, aspartic acid, glutamic acid, glutamine, lysine, arginine, histidine, tryptophan, tyrosine and asparagine.



Van der waals or hydrophobic interactions

- Weaker than hydrogen bonds: take place between two hydrophobic regions of protein.
- Amino acids alanine, valine, leucine, isoleucine, phenylalanine and proline have hydrophobic residues capable of interacting with hydrogen bond interactions.
- Amino acids methionine, tryptophan, threonine and tyrosine contain polar functional groups, but also have significant hydrophobic character, so van der waal interactions are possible.



Relative importance of bonding interactions

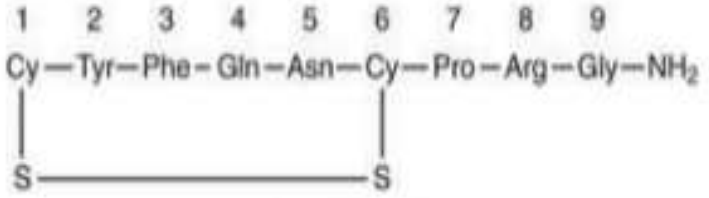
- Most important interactions in tertiary structure: van der waals and hydrogen bonding interactions

  Least important interactions: covalent and ionic bonding

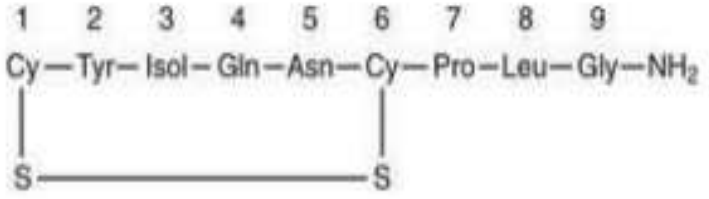
- Two reasons for this:
  1. More opportunity for van der waals and hydrogen bonding interactions
  2. Presence of water surrounding the protein structure

- Only amino acid which can form a covalent Disulphide bond is cysteine, whereas several amino acids can interact with each other through hydrophobic and covalent bonding.

- But Disulphide links are more significant in some small polypeptides such as peptide hormones vasopressin and oxytocin.

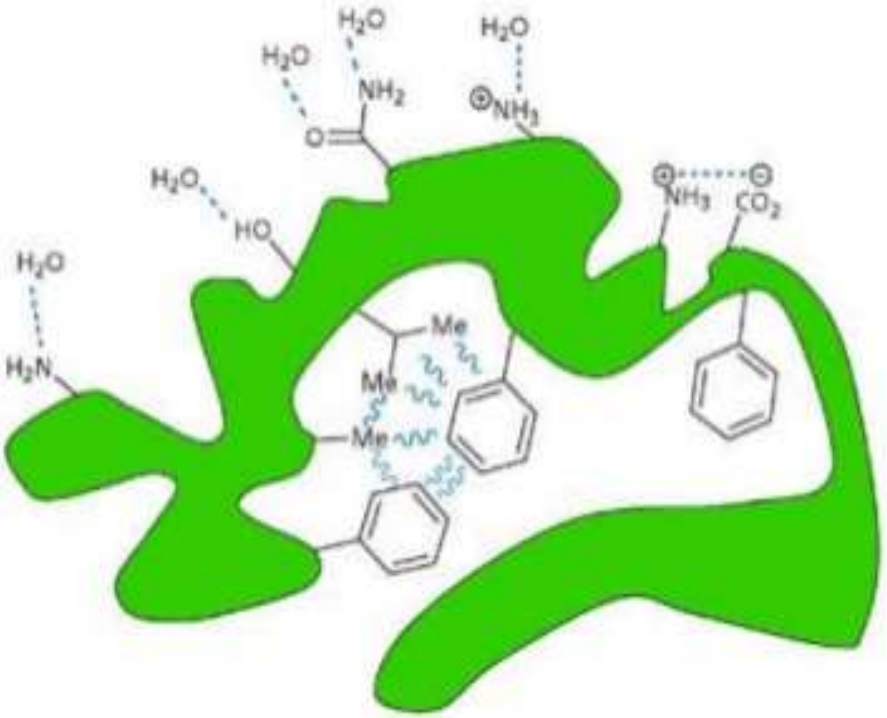


Vasopressin



Oxytocin

- Proteins: surrounded by water
- Water, being highly polar compound interacts readily with polar, hydrophilic amino acid residues capable of forming hydrogen bonds.



- Most stable tertiary structure has most of the hydrophilic groups on the surface so that they interact with water and most of the hydrophobic groups in center so they avoid water and interact with each other.

- Hydrophilic amino acids form a hydrogen bond with water and thus number of ionic and hydrogen bonds contributing to tertiary structure is reduced, this leaves van der waal and hydrophobic interactions to largely determine the three – dimensional shape of the protein.

DETERMINATION OF TERTIARY STRUCTURE

- The known protein structures have come to light through:
- **X-ray crystallographic studies**
- **Nuclear Magnetic Resonance studies**
- The atomic coordinates of most of these structures are deposited in a database known as the Protein Data Bank (PDB).
- It allows the tertiary structures of a variety of proteins to be analyzed and compared.

Quaternary Structure

- Oligomeric proteins are composed of 2 or more polypeptide chains and each polypeptide chain is called a subunit
- Subunits in a multisubunit oligomeric protein may be identical (homomultimer) or different (heteromultimer)
- Polypeptide subunits assemle to form quaternary structure and are held together by non-covalent interactions such as Hydrophobic interactions, Electrostatic interactions, Hydrogen bonds as well as covalent cross links

- For a protein to have Quaternary structure –
- The protein should be formed by more than one polypeptide chain
- The chains cannot be attached by covalent bond among them
- Therefore, all proteins cannot show a quaternary level of organization
- Hemoglobin is formed by 4 peptide chains associated through non-covalent bonds
- Insulin is formed by 2 peptide chains but the chains are linked by disulfide bonds. Therefore, insulin does not have a quaternary structure

Quaternary (4°) Structure

- aggregation of two or more polypeptide subunits
- forms 2 types of proteins: globular and fibrous
- not found in all proteins

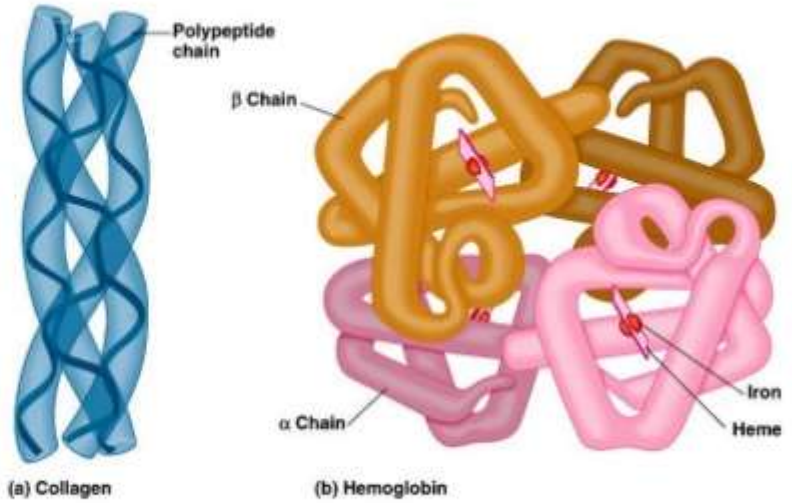
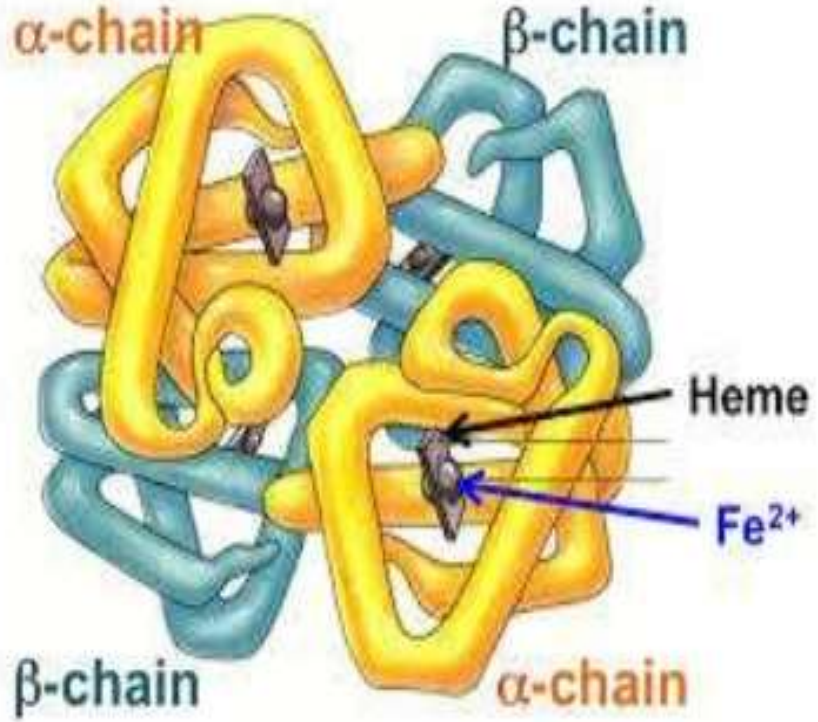


Fig. 5.23

- Hemoglobin is made up of four protein molecules: two identical α subunits and two identical β subunits.



FOLLOWING YOUTUBE LINKS MAY BE FOLLOWED

- https://www.youtube.com/watch?v=PPJ7C3hcnPw
- https://www.youtube.com/watch?v=nTxvoKspIS8