

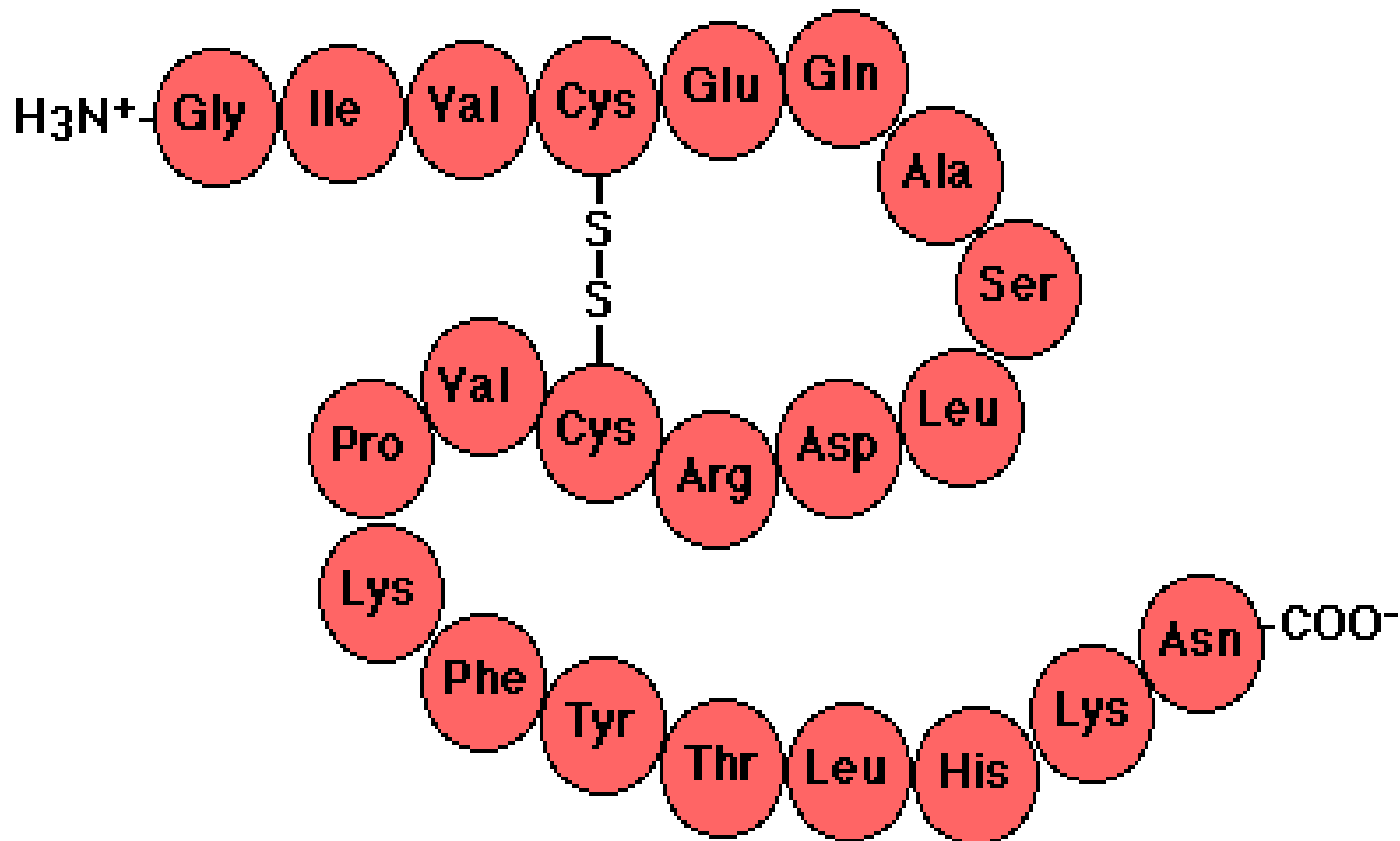
PROTEIN STRUCTURE

BIOCHEMISTRY

Primary Structure of Proteins

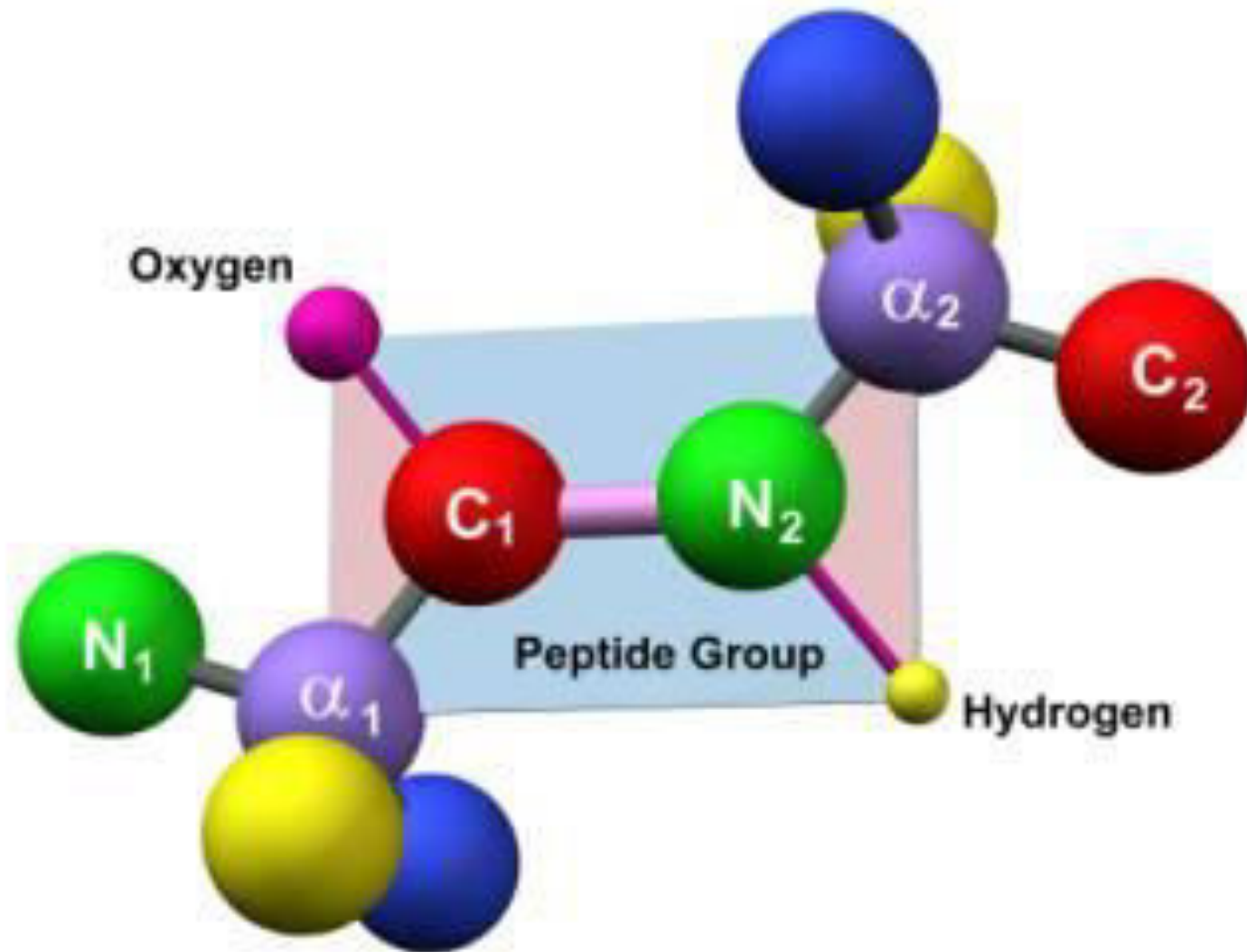
- The primary structure of peptides and proteins refers to the linear number and order of the amino acids present.

- The convention for the designation of the order of amino acids is that:
- The N-terminal end (i.e. the end bearing the residue with the free α -amino group) is to the left (and the number 1 amino acid) and the C-terminal end (i.e. the end with the residue containing a free α -carboxyl group) is to the right.

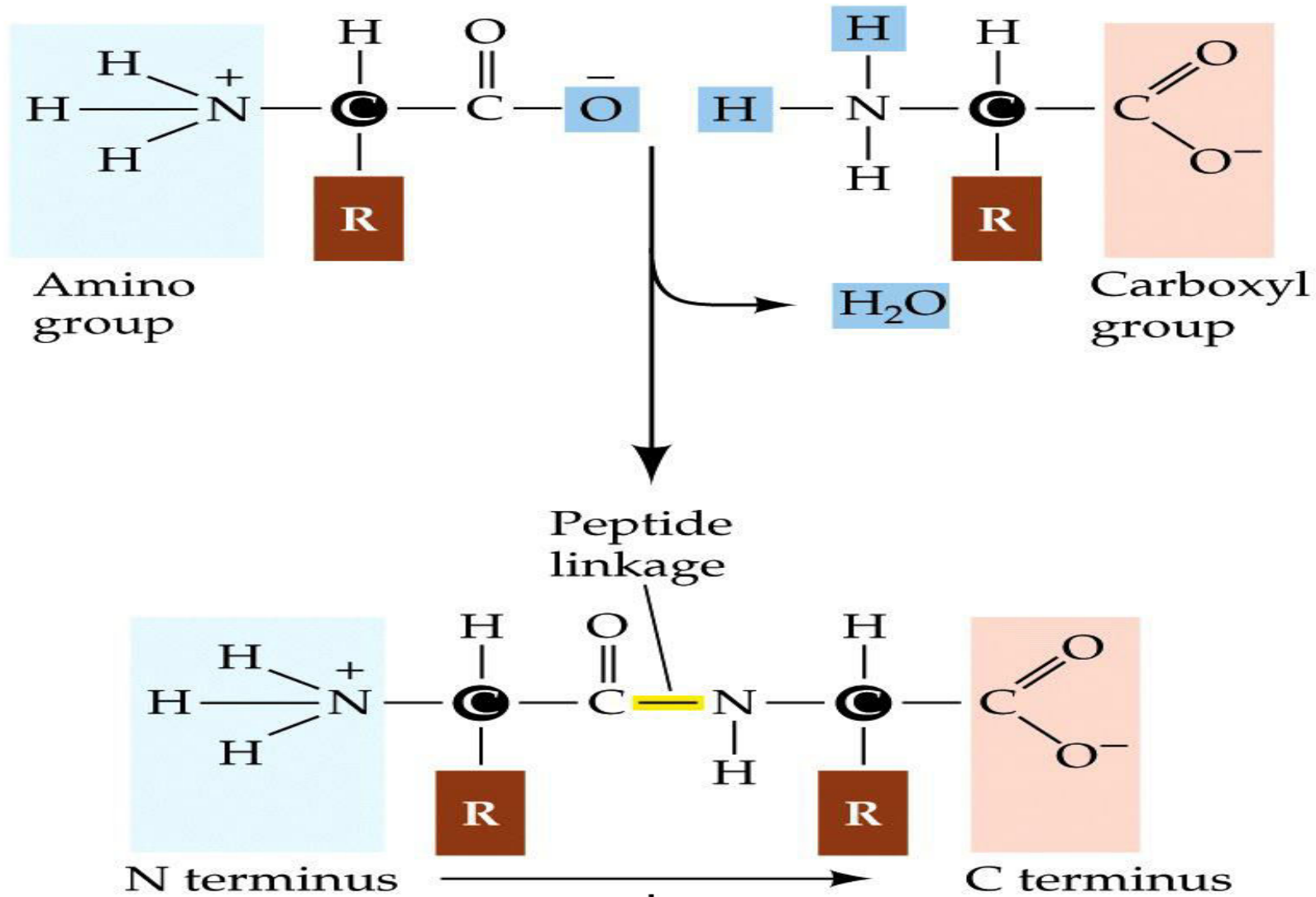


Peptide bond

- The peptide bond combines two amino acids and it is formed by a group of atoms from each side of the bond called the peptide group, illustrated here by a rectangle.



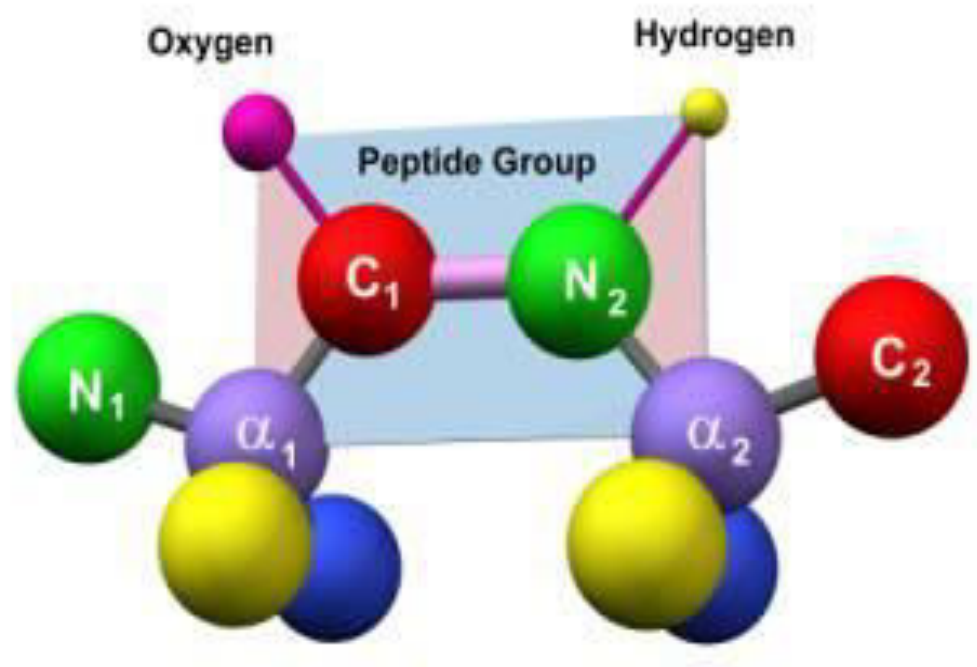
- The first amino acid contributes to the peptide group its **carbonyl carbon (C1), α -carbon (alpha 1) and the carbonyl oxygen (Oxygen)**.
- The second amino acid contributes its **amide nitrogen (N), the α -carbon (alpha 2), and the amide hydrogen (Hydrogen)**.



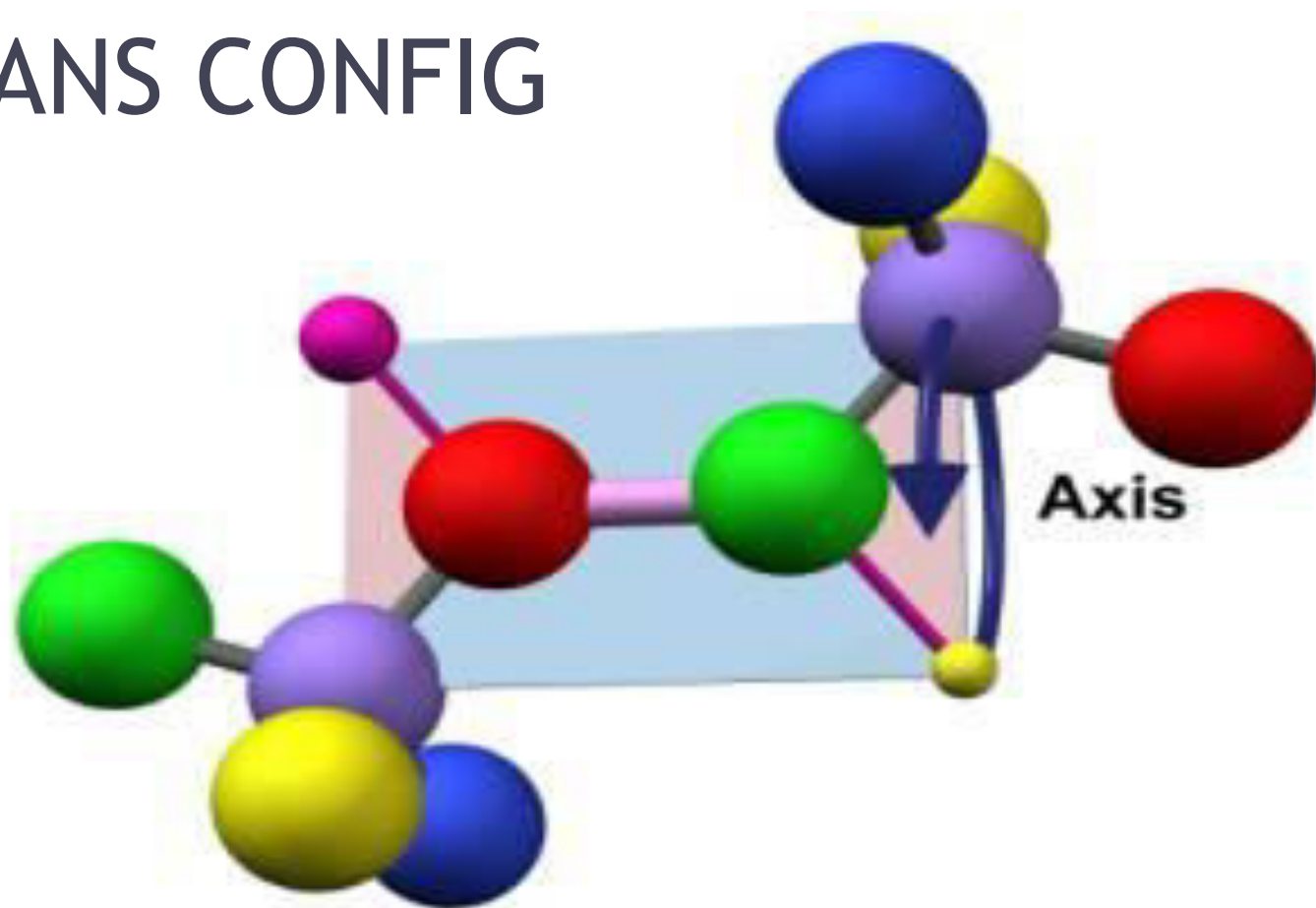
- It is the **partial double-bond** character of the peptide bond that defines the conformations a polypeptide chain may assume.
- It is shorter than a single bond
- Rigid & planar

- The bond is planar, and the group can take one of two major configurations:
- **Cis or Trans**

CIS CONFIG



TRANS CONFIG



- The cis configuration of the peptide group has a much **greater steric hindrance** between the two amino acids. This means that the atoms get in the way of each other.

- For this reason nearly all of the peptide bonds in **naturally occurring proteins are in the trans configuration**, at the very least 95% of them are trans-peptide bonds.
- However, the cis configuration can and does naturally occur.

Amino-Terminal Sequence Determination

- Prior to sequencing peptides it is necessary to **eliminate disulfide bonds** within peptides and between peptides.

- Several different chemical reactions can be used in order to permit separation of peptide strands and prevent protein conformations that are dependent upon disulfide bonds.

- Purified sample of polypeptide is first hydrolyzed by a strong acid at 110°C for 24 hours.
- This cleaves the peptide bonds and releases the individual amino acids.
- Amino acids are separated by chromatography.

- There are three major chemical techniques for sequencing peptides and proteins from the N-terminus. These are:
 - Sanger
 - Dansyl chloride and
 - Edman techniques.

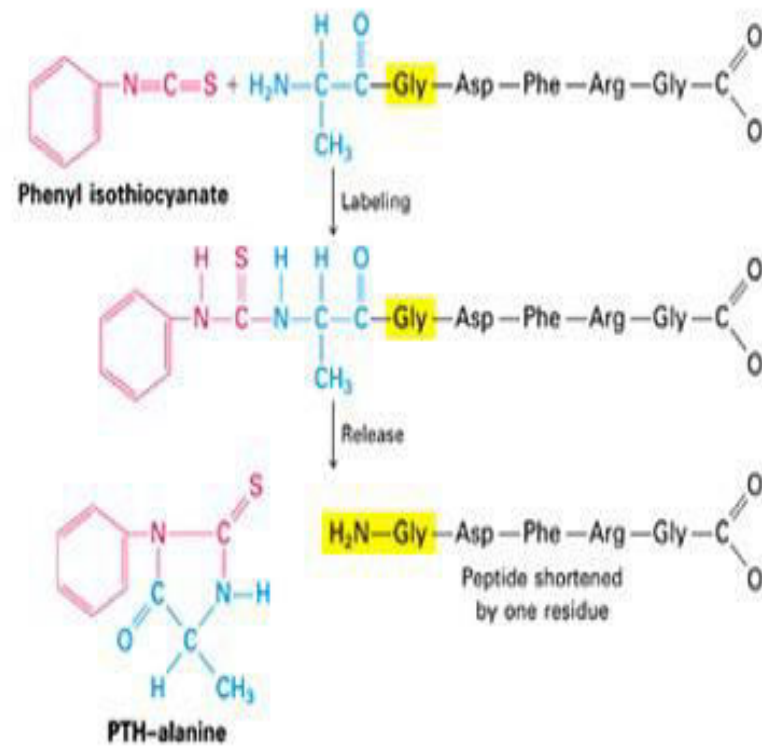
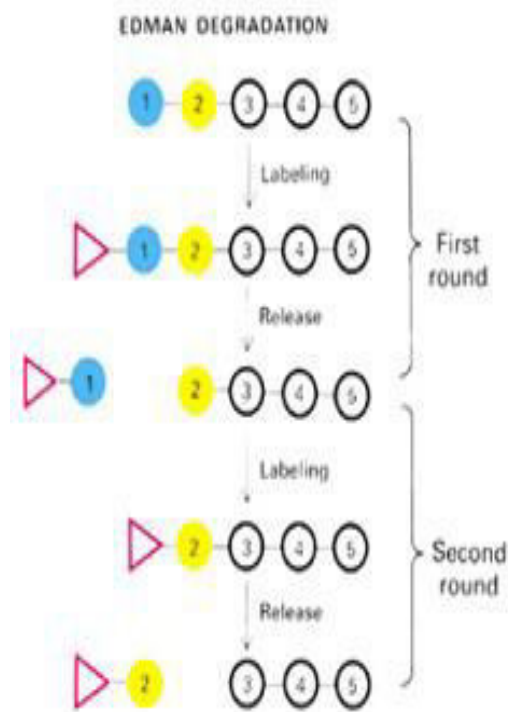
Edman degradation

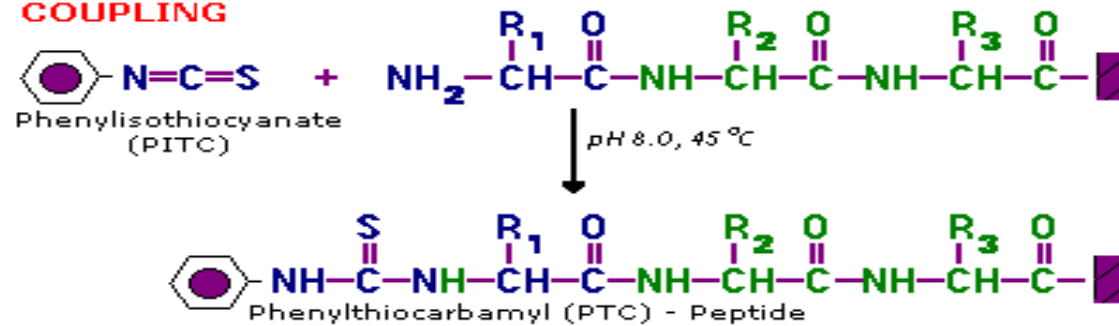
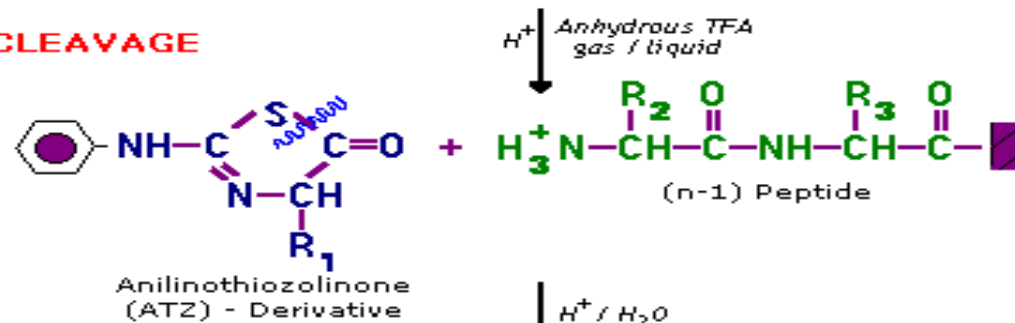
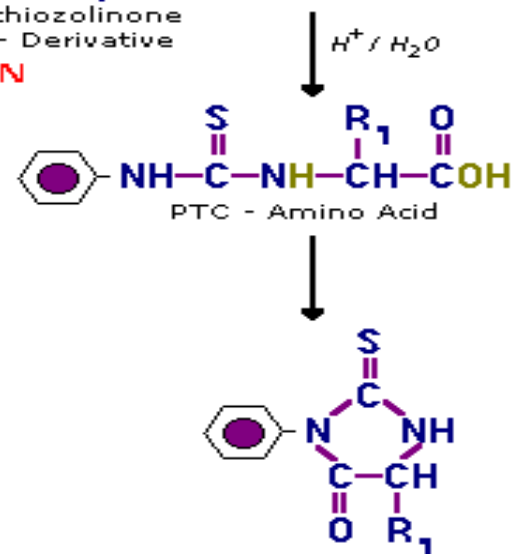
- The utility of the Edman degradation technique is that it allows for amino acid sequence to be obtained from the N-terminus inward.
- Using this method it is possible to obtain the entire sequence of peptides.

- This method utilizes phenylisothiocyanate to react with the N-terminal residue under alkaline conditions.

- The reaction results in a rearrangement of the released N-terminal residue to a **phenylthiohydantoin** derivative.
- The added advantage of the Edman process is that the remainder of the peptide is intact.

- The entire sequence of reactions can be repeated over and over to obtain the sequences of the peptide.
- This process has been automated to allow rapid and efficient sequencing of even extremely small quantities of peptide.



COUPLING**CLEAVAGE****CONVERSION**

Phenylthiohydantoin (PTH)
Amino Acid

Cleavage into smaller fragments

- A single series of Edman degradation reactions is not able to determine the entire sequence of a protein.
- What is needed are smaller fragments, with new amino termini, which can be individually purified and sequenced.

- This is accomplished by cleaving the protein with a **proteolytic enzyme**, such as **trypsin**, or a chemical reagent such as **cyanogen bromide**, which generates a set of peptides, fragments of the original protein, that can be separated and sequenced.

- This method has now been automated, and the machine used is called sequenator.

Determination by DNA sequencing

- Sequence of nucleotides in a DNA specifies the sequence of amino acids in a protein.
- If the nucleotide sequence is known then it can be translated for the aa sequence.
- Cannot identify disulfide bonds and posttranslationally modified a.a.

Secondary Structure in Proteins

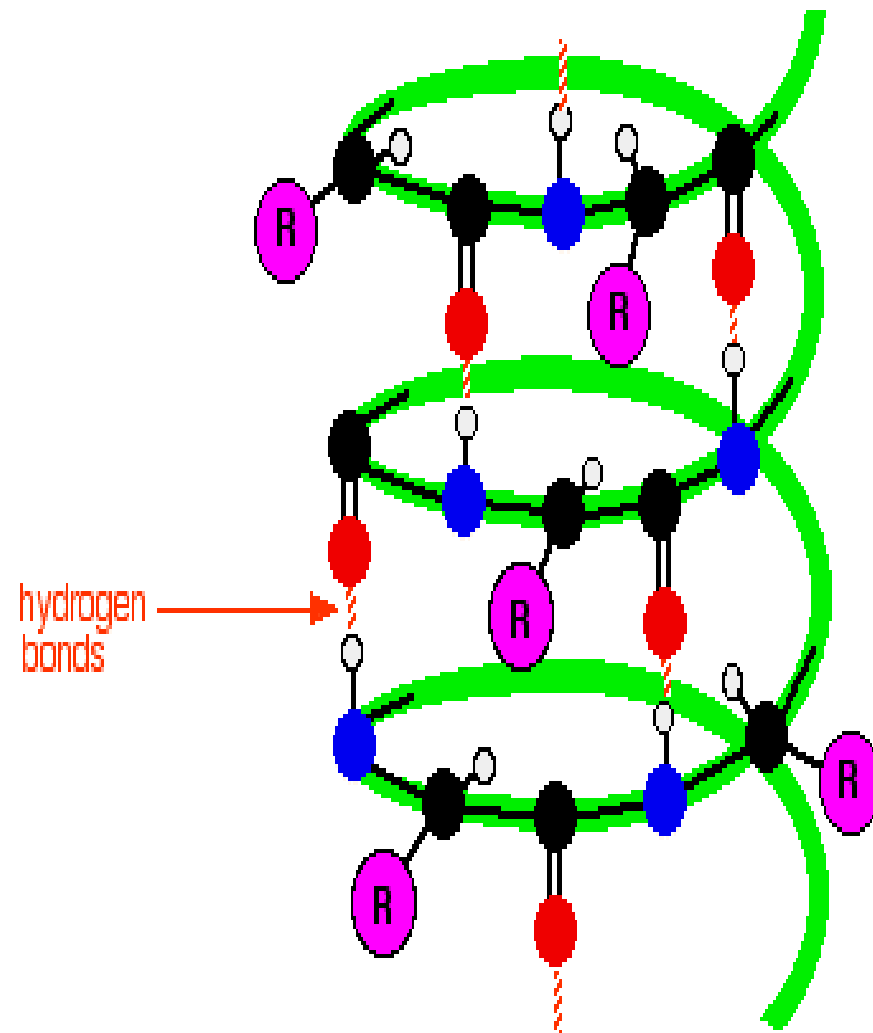
- In general proteins fold into two broad classes of structure termed, globular proteins or fibrous proteins. Globular proteins are compactly folded and coiled, whereas, fibrous proteins are more filamentous or elongated.

The α -Helix

- It is the most common confirmation.
- It is a spiral structure.
- Tightly packed coiled polypeptide backbone, with extending side chains.

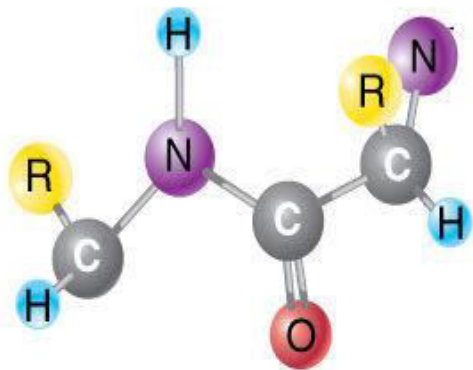
- The formation of the α -helix is spontaneous .
- It is stabilized by H-bonding between amide hydrogens and carbonyl oxygens of peptide bonds.
- It is spaced four residues apart.

- This orientation of H-bonding produces a helical coiling of the peptide backbone such that the R-groups lie on the exterior of the helix and perpendicular to its axis.

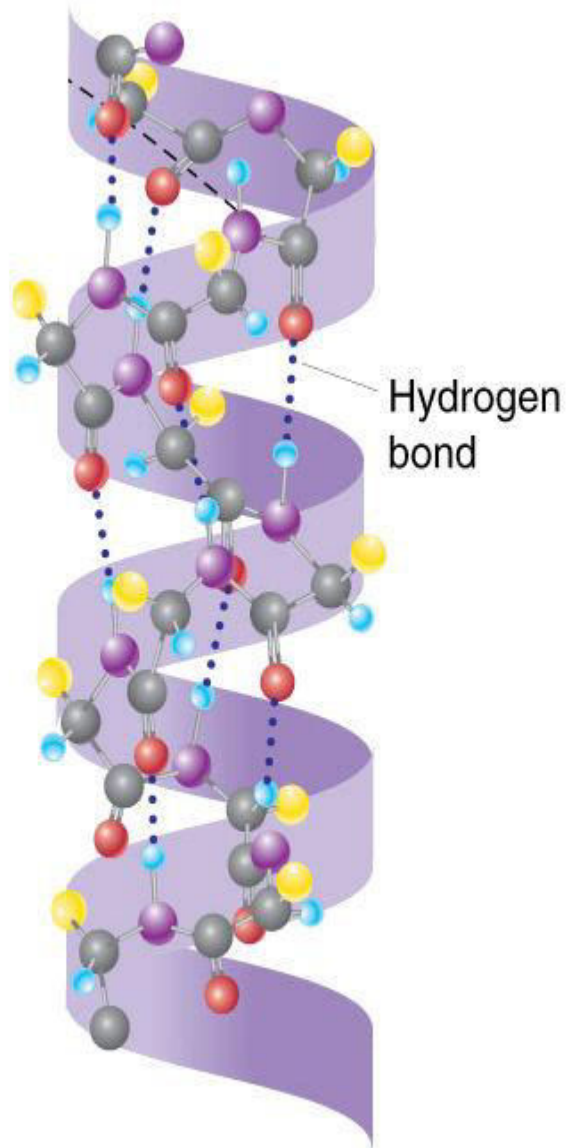


- A complete turn of the helix contains an average of **3.6 aminoacyl residues**, and the distance it rises per turn is 0.54 nm
- The **R groups** of each aminoacyl residue in an α helix face **outward**

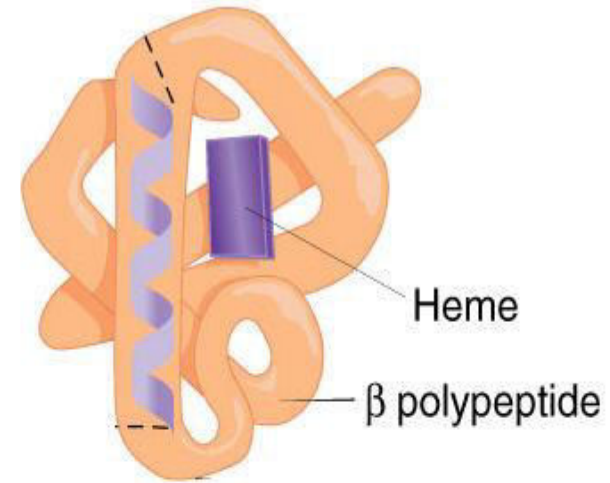
- e.g. the keratins- entirely α -helical
- Myoglobin- 80% helical.



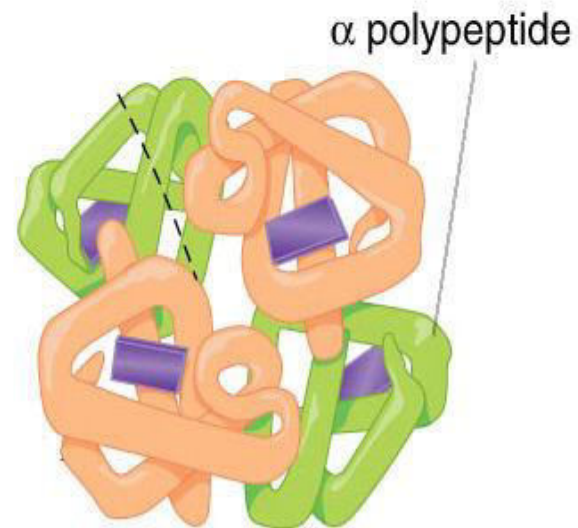
(a) Primary structure



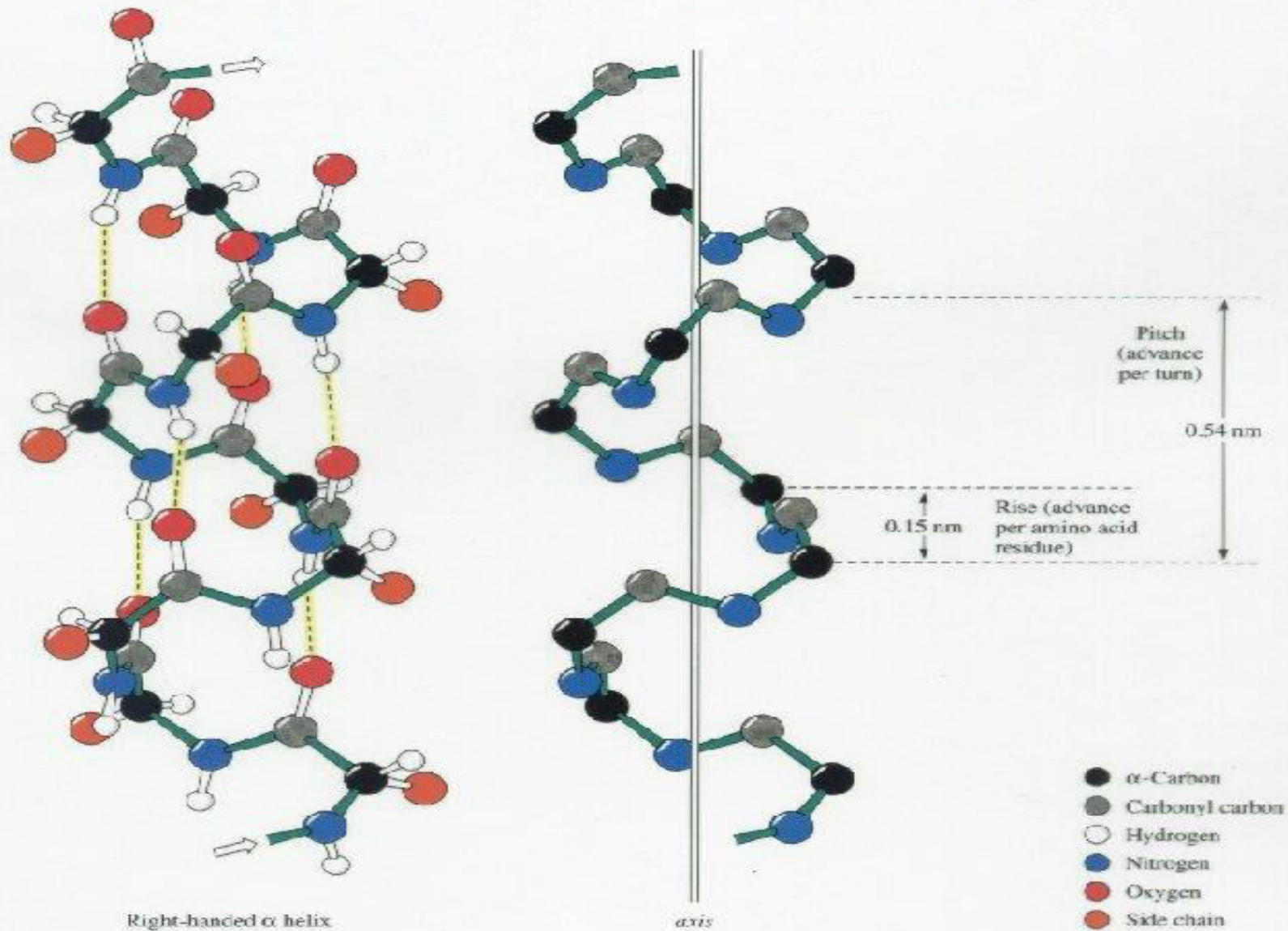
(b) Secondary structure



c) Tertiary structure



(d) Quaternary structure

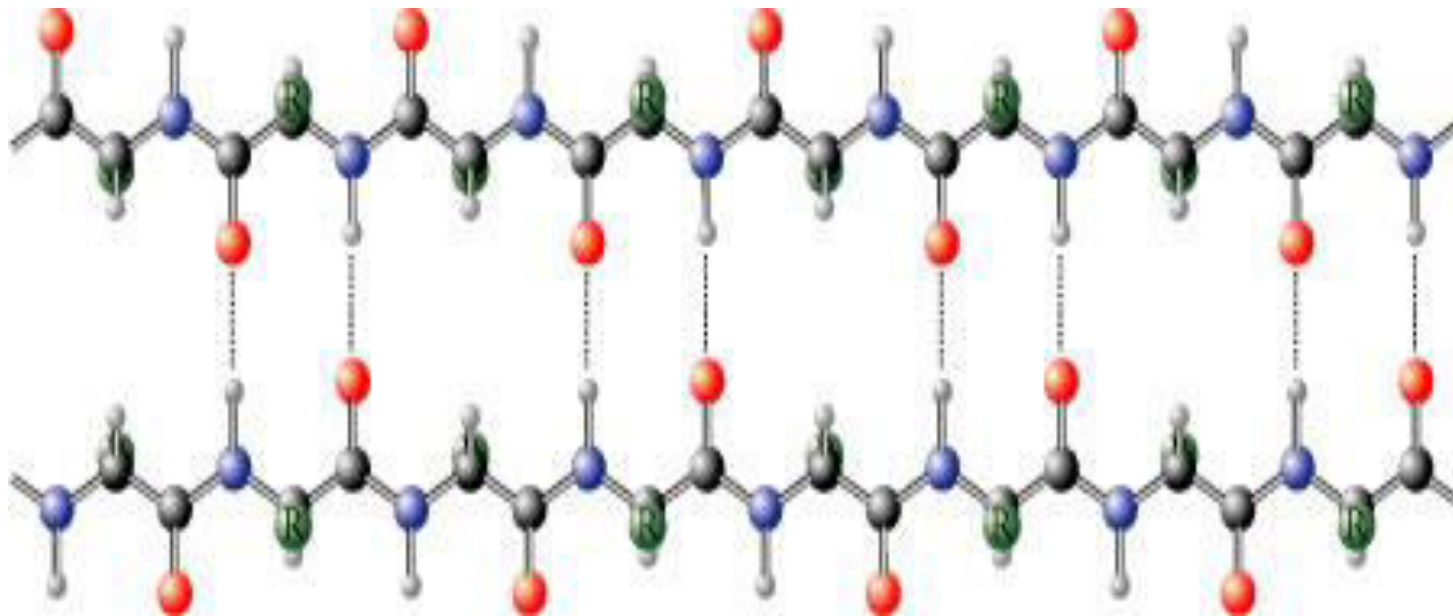


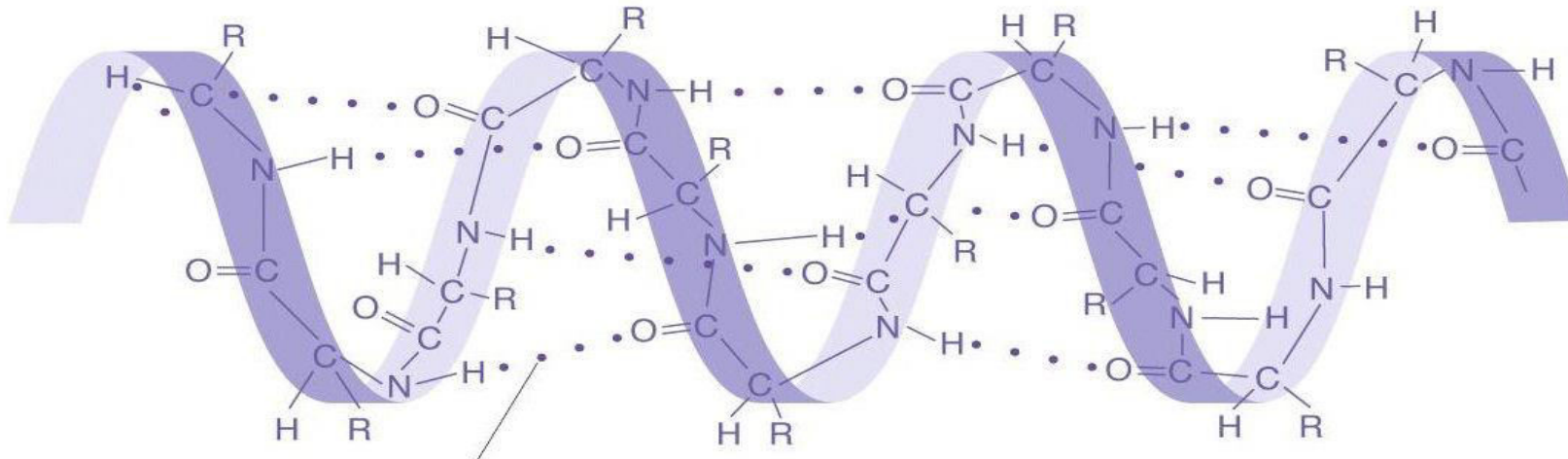
- Glycine and Proline , bulky amino acids, charged amino acids favor disruption of the helix.
- The disruption of the helix is important as it introduces additional folding of the polypeptide backbone to allow the formation of globular proteins.

β-Sheets

- β-sheets are composed of 2 or more different regions of stretches of at least 5-10 amino acids.
- The folding and alignment of stretches of the polypeptide backbone aside one another to form β-sheets is stabilized by **H-bonding** between amide hydrogens and carbonyl oxygens.

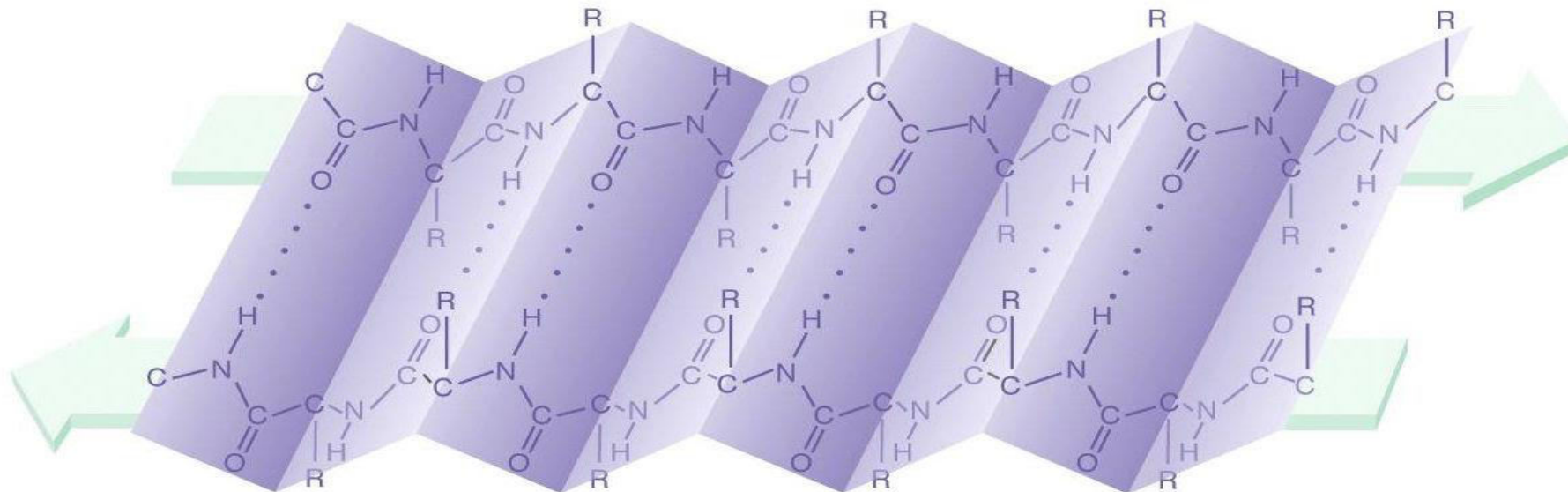
- Unlike the compact backbone of the α helix, the peptide backbone of the β sheet is highly extended.
- β -sheets are said to be pleated in which the R groups of adjacent residues point in opposite directions.
- β -sheets are either parallel or antiparallel.





Hydrogen bonds between amino acids
at different locations in polypeptide chain

α helix



Pleated sheet

- In **parallel sheets** adjacent peptide chains proceed in the **same direction** (i.e. the direction of N-terminal to C-terminal ends is the same).
- In **antiparallel sheets** adjacent chains are aligned in **opposite directions**.

β -Bends & Loops

- Roughly half of the residues in a “typical” globular protein reside in α helices and β sheets
- and half in loops, turns, bends, and other extended conformational features.

- Turns and bends refer to short segments of amino acids that join two units of secondary structure, such as two adjacent strands of an antiparallel β sheet.

Super-Secondary Structure

- Globular proteins are constructed by combining secondary structural elements.
 - They form the core region .
 - Connected by loops.
 - Produced by packing side chains from adjacent secondary structural elements.
 - Examples include:
-

- helix-turn-helix
- helix-loop-helix
- zinc finger
- Greek key
- β -meander
- β -Barrel

Tertiary Structure of Proteins

- Tertiary structure refers to the complete **three-dimensional structure of the polypeptide units of a given protein.**

- Secondary structures of proteins often constitute distinct **domains**.
- Domain is the basic unit of structure and function

- Tertiary structure describes the **relationship** of different domains to one another within a protein.
- OR the final arrangement of domains in a polypeptide.

- The interactions of different domains is governed by several forces:
- These include hydrogen bonding, hydrophobic interactions, electrostatic interactions and van der Waals forces.

- A- helix and B- sheet --- hydrogen bonding and this eliminates the possibility of water to disrupt the structure of proteins.

DOMAINS:

- Polypeptide chains more than 200 amino acids in length generally consists of two or more domains.
- Core of the domain is built from motifs.

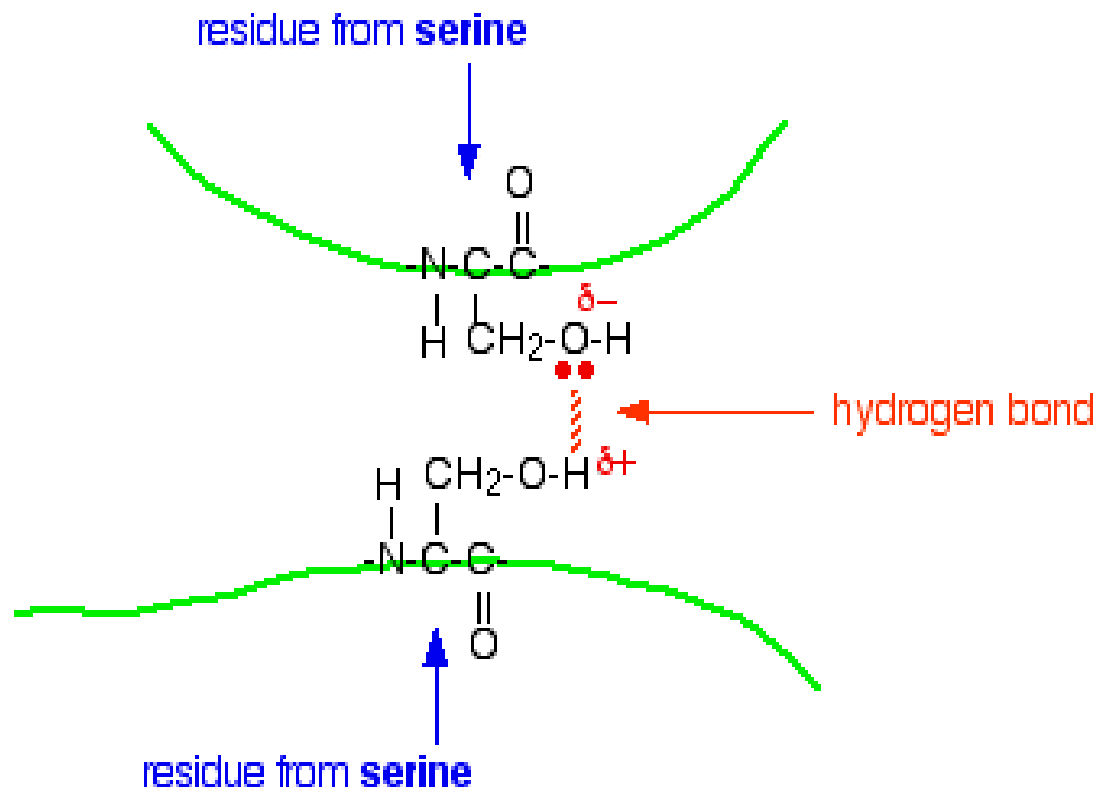
- Folding of the peptide chain within the domain usually occurs independently of folding in other domains.
- So each domain has the characteristics of a small compact globular protein. It is independent of other domains.

Forces Controlling tertiary Structure

Hydrogen Bonding:

- Polypeptides contain numerous proton donors and acceptors both in their backbone and in the R-groups of the amino acids.
- The environment in which proteins are found also contains the H-bond donors and acceptors of the water molecule.

- H-bonding, occurs not only within and between polypeptide chains but with the surrounding aqueous medium.



Hydrophobic Forces:

- Proteins are composed of amino acids that contain either hydrophilic or hydrophobic R-groups.
- It is the nature of the interaction of the different R-groups with the aqueous environment that plays the major role in shaping protein structure.

- The spontaneous folded state of globular proteins is a balance between the opposing energetics of H-bonding between hydrophilic R-groups and the aqueous environment and the repulsion from the aqueous environment by the hydrophobic R-groups.

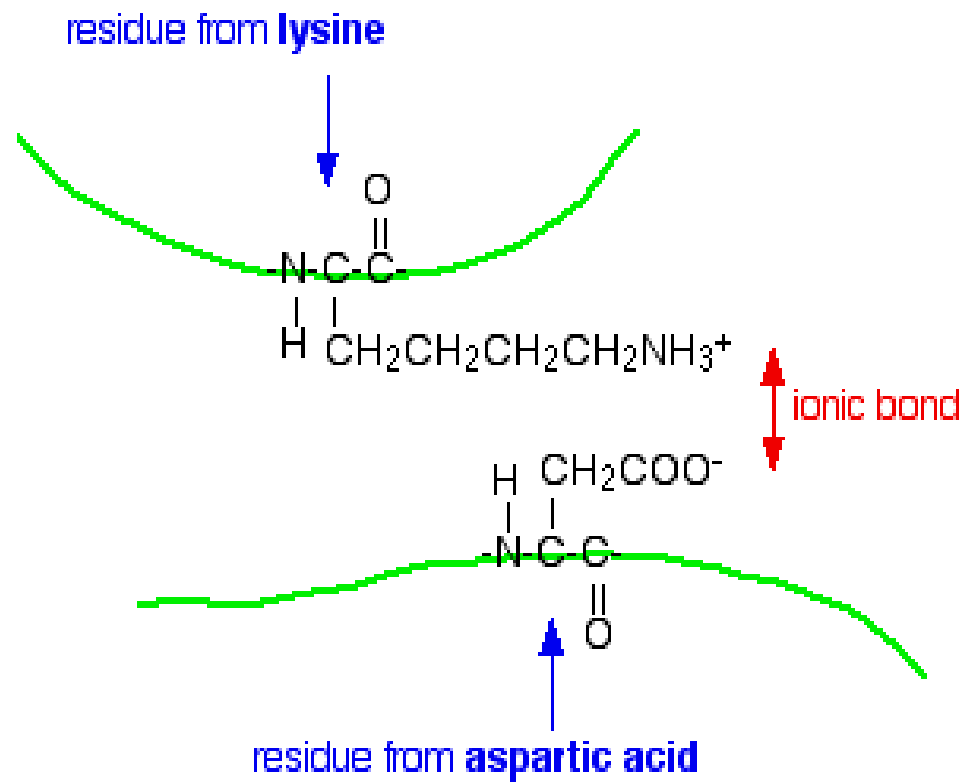
- The hydrophobicity of certain amino acid R-groups tends to drive them away from the exterior of proteins and into the interior. This driving force restricts the available conformations into which a protein may fold.

Electrostatic Forces:

- Electrostatic forces are mainly of three types; charge-charge, charge-dipole and dipole-dipole. Typical charge-charge interactions that favor protein folding are those between oppositely charged R-groups such as K or R and D or E.

Charge-dipole interactions:

- This refers to the interaction of ionized R-groups of amino acids with the dipole of the water molecule.

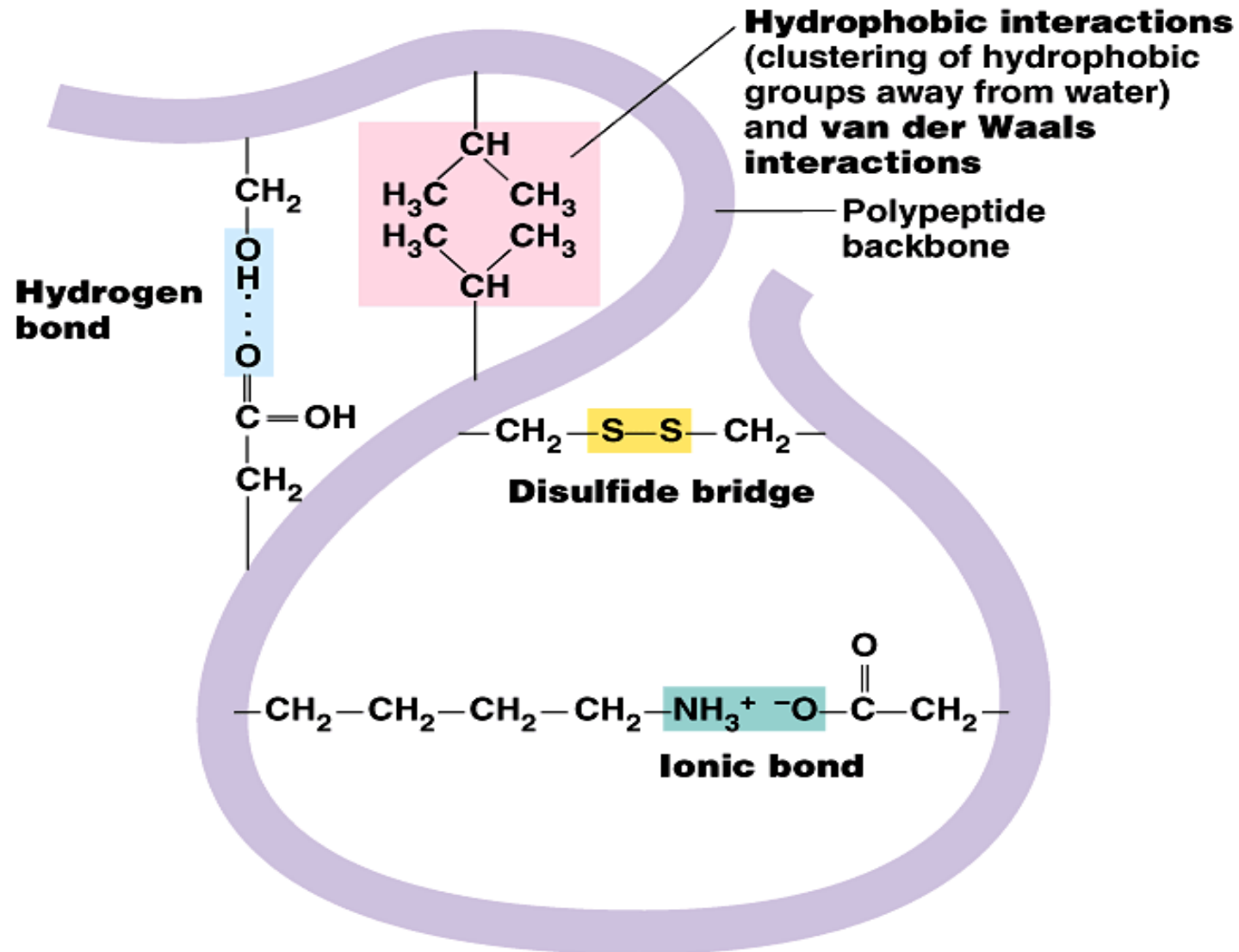


Van der Waals Forces:

- There are both attractive and repulsive van der Waals forces that control protein folding. Attractive van der Waals forces involve the interactions among induced dipoles that arise from fluctuations in the charge densities that occur between adjacent uncharged non-bonded atoms..

- Repulsive van der Waals forces involve the interactions that occur when uncharged non-bonded atoms come very close together. The repulsion is the result of the electron-electron repulsion that occurs as two clouds of electrons begin to overlap.

- Van der Waals forces are extremely weak, relative to other forces, it is the huge number of such interactions that occur in large protein molecules that make them significant to the folding of proteins.



Quaternary Structure

- Many proteins contain 2 or more different polypeptide chains that are held in association by the same non-covalent forces that stabilize the tertiary structures of proteins.

- The arrangement of these **polypeptide subunits** is called the **quaternary structure** of proteins.

- Two subunits- **dimeric**
- Three subunits- **trimeric**
- Proteins with **multiple polypeptide subunits** are **oligomeric proteins**.

- Oligomeric proteins can be composed of multiple identical polypeptide chains or multiple distinct polypeptide chains.
- Proteins with identical subunits are termed homo-oligomers. Proteins containing several distinct polypeptide chains are termed hetero-oligomers.

- Hemoglobin, the oxygen carrying protein of the blood, contains two α and two β subunits arranged with a quaternary structure in the form, $\alpha_2\beta_2$.
- Hemoglobin is, therefore, a hetero-oligomeric protein.

