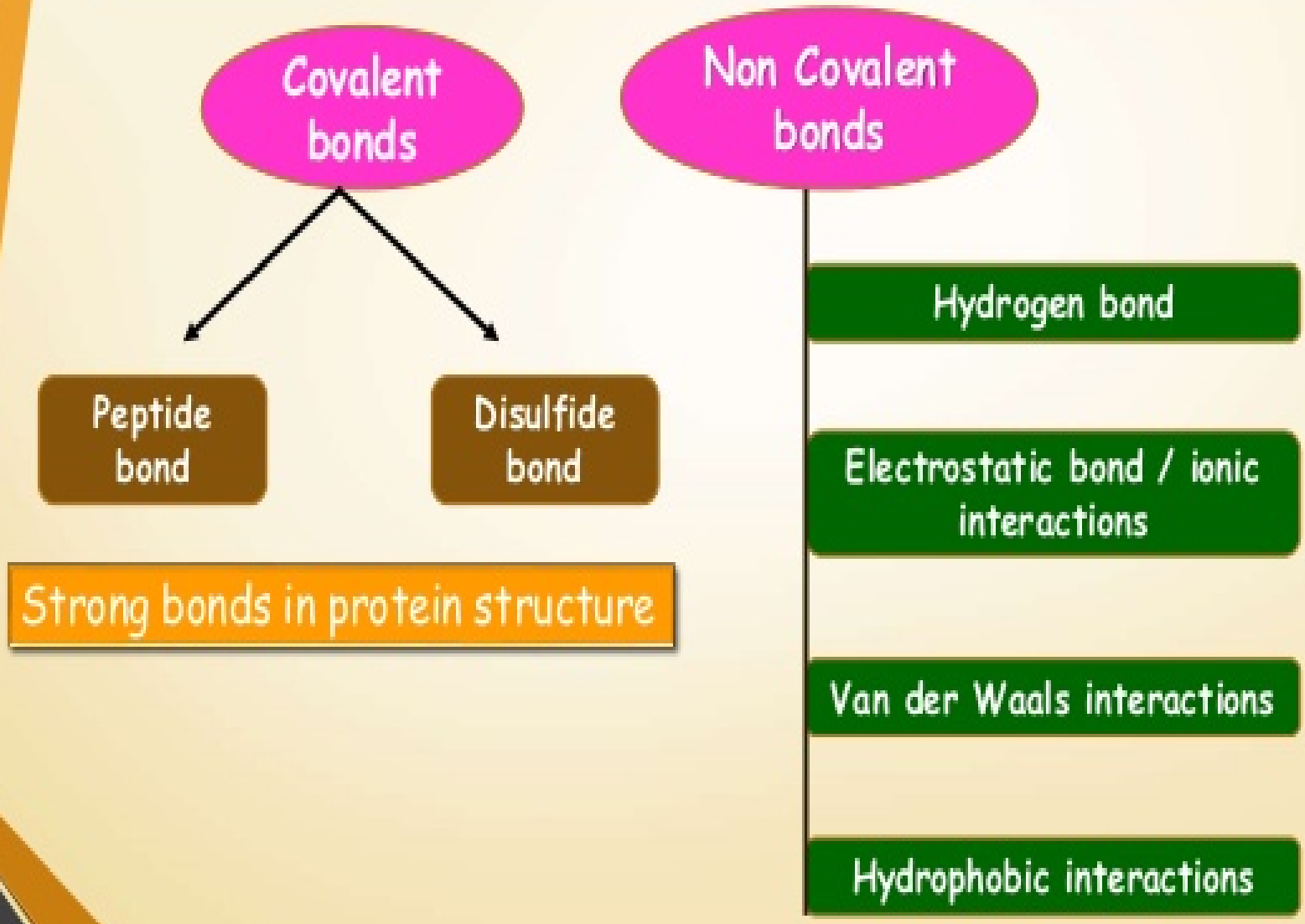
Structural Organisation of Protein

DR. ROSHNI SADARIA

Associate Professor

Biochemistry Department

Proteins are formed of A.A. linked together by the following types of bonds



Strong bonds in protein structure

Bonds responsible for protein structure

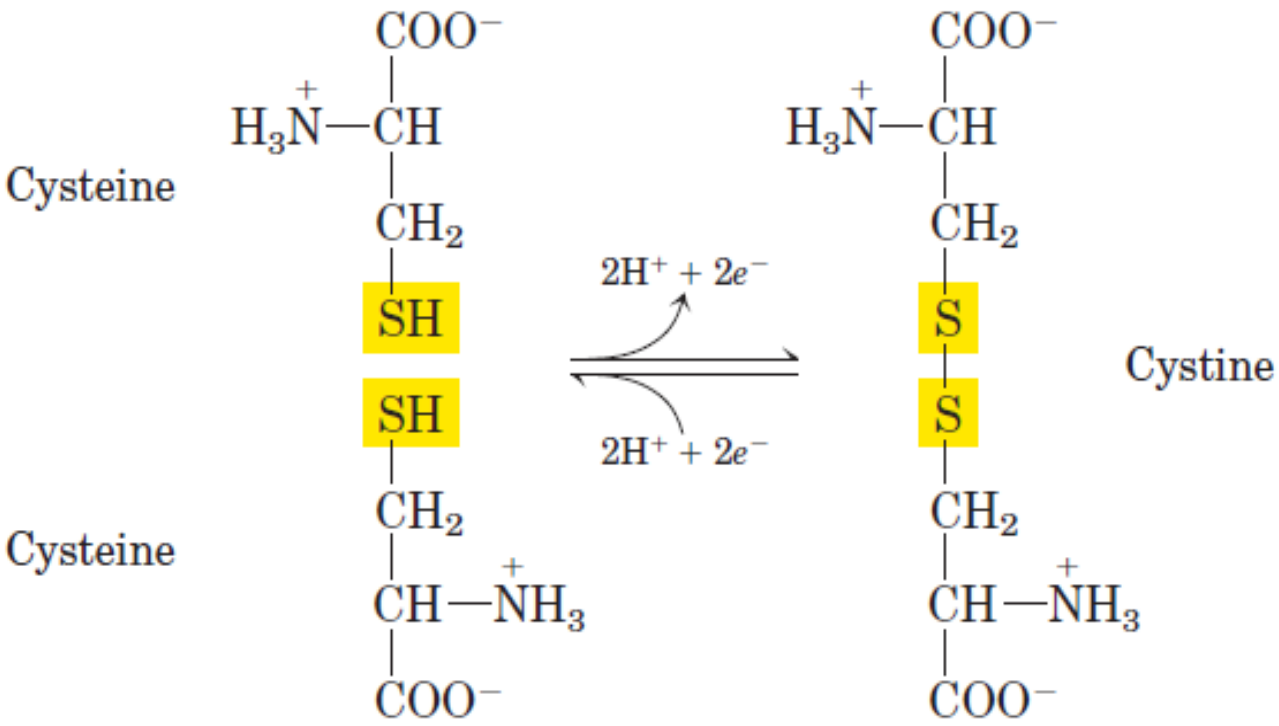
- Two types of bonds

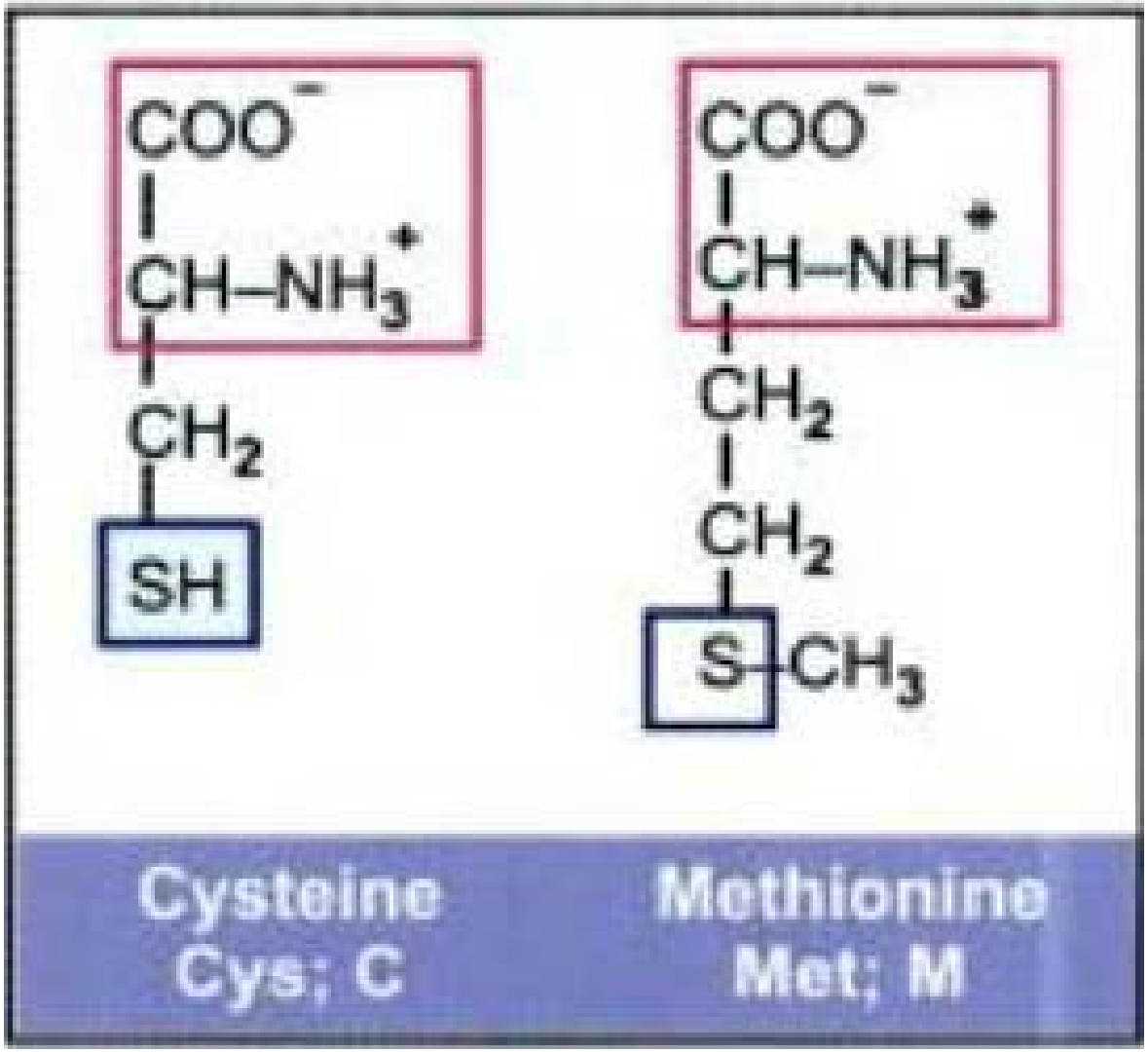
  Covalent and non-covalent bonds

- Covalent bonds:
- The peptide and disulfide bonds are the strong bonds in protein structure.

- **Disulfide bond:**

- A disulfide bond (-S-S-) is formed by oxidation of sulfhydryl groups (-SH) of two cysteine residues to produce cystine.
- Formed in a single polypeptide chain or between different polypeptides.
- Contribute to the structural conformation and stability of protein.


Cysteine
Cysteine
Cystine
COO⁻
H₃N⁺—CH
CH₂
SH
SH
CH₂
CH—NH₃⁺
COO⁻
2H⁺ + 2e⁻
2H⁺ + 2e⁻
COO⁻
H₃N⁺—CH
CH₂
S
S
CH₂
CH—NH₃⁺
COO⁻


COO⁻
CH–NH₃⁺
CH₂
SH
COO⁻
CH–NH₃⁺
CH₂
CH₂
S–CH₃
Cysteine
Cys; C
Methionine
Met; M

Non-covalent bonds

- ***Hydrogen bonds (H-bonds):***
- The H-bonds are formed by sharing of H-atoms between the Nitrogen and carbonyl oxygen of different peptide bonds.
- Each H-bond is weak but collectively they are strong.
- Significantly contribute to the protein structure.

- ***<u>Hydrophobic bonds</u>***
- Formed by interactions between non-polar hydrophobic side chains of neutral amino acids by eliminating water molecule.
- This serves to hold the lipophilic side chains together.

- ***Electrostatic bonds:***

- These bonds are formed by interactions between negatively charged groups of acidic amino acids with positively charged groups of basic amino acids.

- ***Van der Waals forces:***

- weak, short-range electrostatic attractive forces between uncharged molecules.
- Collectively contribute maximum towards the stability of protein structure.

STRUCTURAL ORGANIZATION

- Proteins are made up of one or more polypeptide chains.

❑ STRUCTURAL ORGANISATION OF PROTEINS

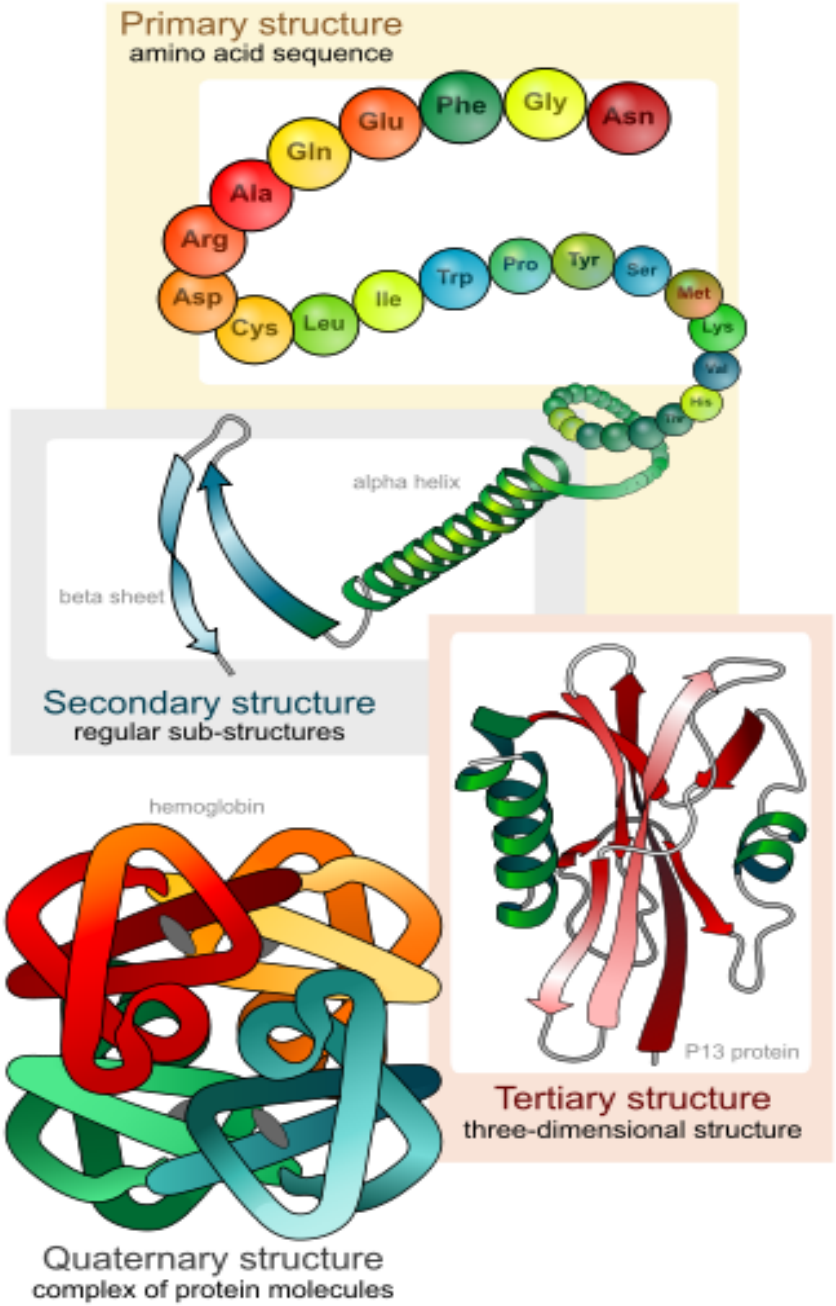
Proteins exhibit **four different levels** of structural organization

1)Primary structure

2)Secondary structure

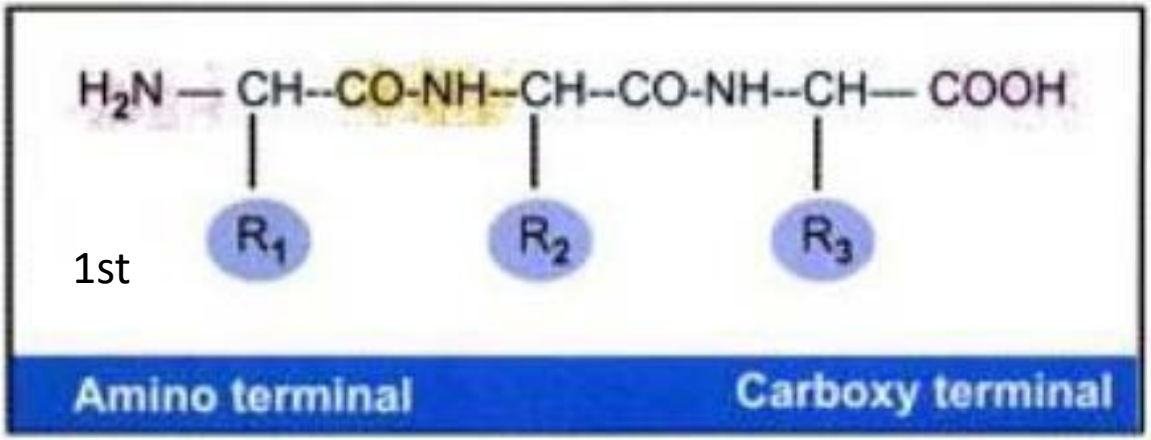
3)Tertiary structure

4)Quaternary structure


Primary structure
amino acid sequence
Glu
Phe
Gly
Asn
Gln
Ala
Arg
Asp
Cys
Leu
Ile
Trp
Pro
Tyr
Ser
Met
Lys
His
alpha helix
beta sheet
Secondary structure
regular sub-structures
P13 protein
Tertiary structure
three-dimensional structure
hemoglobin
Quaternary structure
complex of protein molecules

PRIMARY STRUCTURE

- **Number and sequence of amino acid** residues from N-terminal to the C-terminal forms Primary structure.



- NH2-Gly-Ala-Val-COOH or
- Gly-Ala-Val or GAV
- Gly-Val-Ala : Different sequence

➢ Naming the peptide:

- Amino end & carboxyl end**.**
- Therefore, all amino sequences are read from the N- to the C-terminal end of the peptide.
- When a polypeptide is named, all amino acid residues have their suffixes (-ine, -an, -ic, or -ate) changed to **–yl,** with the exception of the C-terminal amino acid.
- For example, a tripeptide composed of an N-terminal valine, a glycine, and a C-terminal leucine is called valylglycylleucine.

- **Features:**

- **Peptide bonds** maintain and form the **backbone** of primary structures.
- Amino acids are arranged in **linear chain** linked by peptide bonds.
- Primary structure of each protein has a **unique amino acid sequence and numbers**, decided by genes.
- Primary structure is **not affected during denaturation.**

➢ Peptide Bond:

- Rigid, Planar
- Partial double bond- 1.32A^{o} – no freedom of rotation
- Trans configuration
- Numbered from the N terminal to the C terminal.
- Has an electric dipole
- Side chains are free to rotate on either side of peptide bond. The angles of rotation known as Ramachandran angles which determines the spatial orientation of the peptide chain.

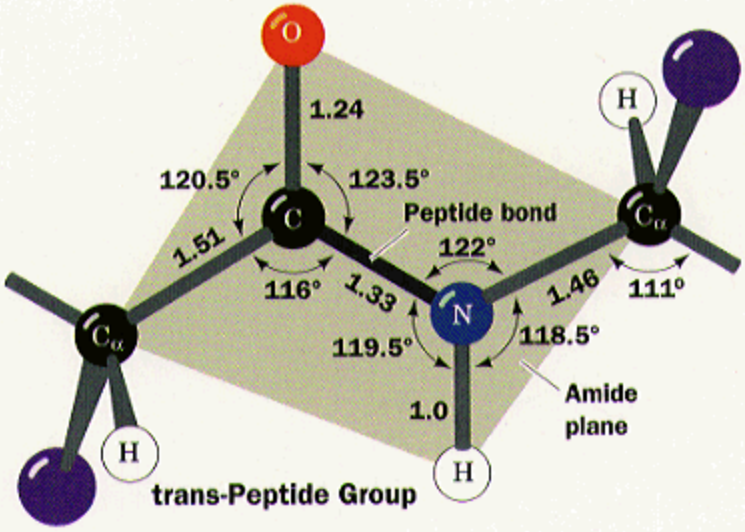


Figure 7-1. The trans-peptide group.

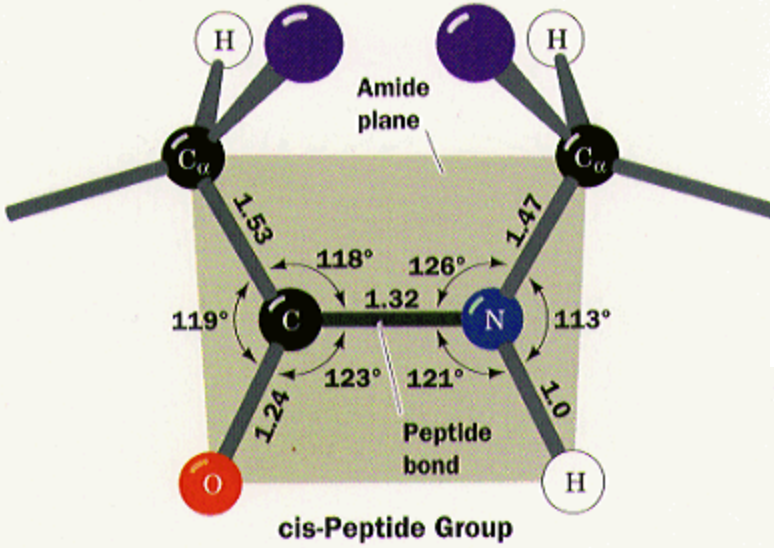
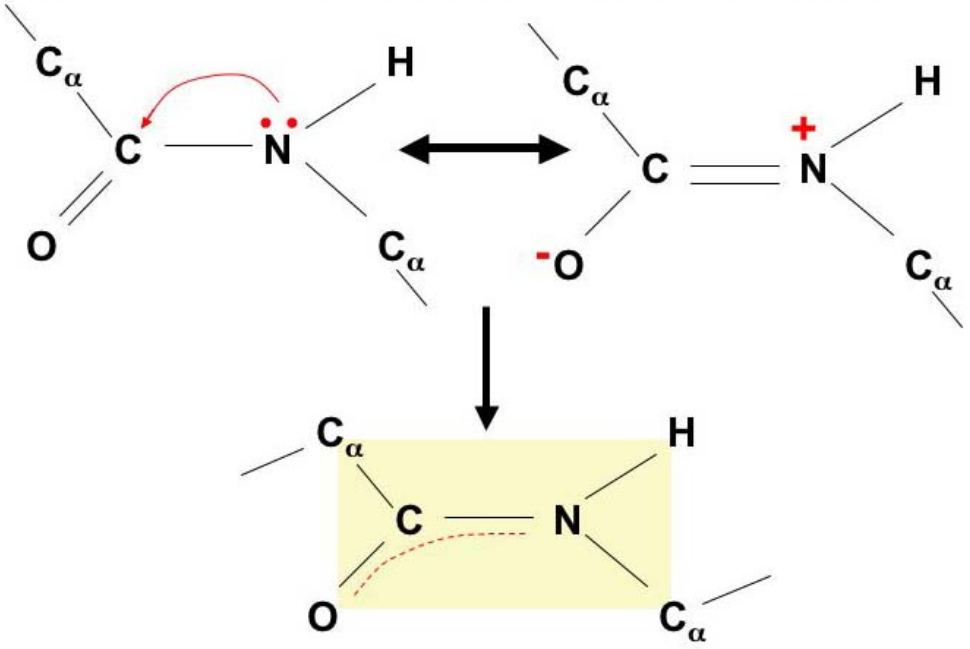


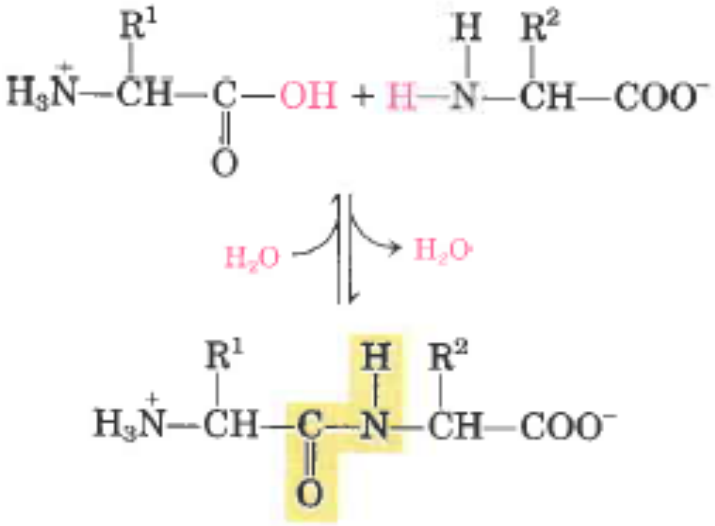
Figure 7-2. The cis-peptide group.

Peptide Bond has partial double bond characteristics

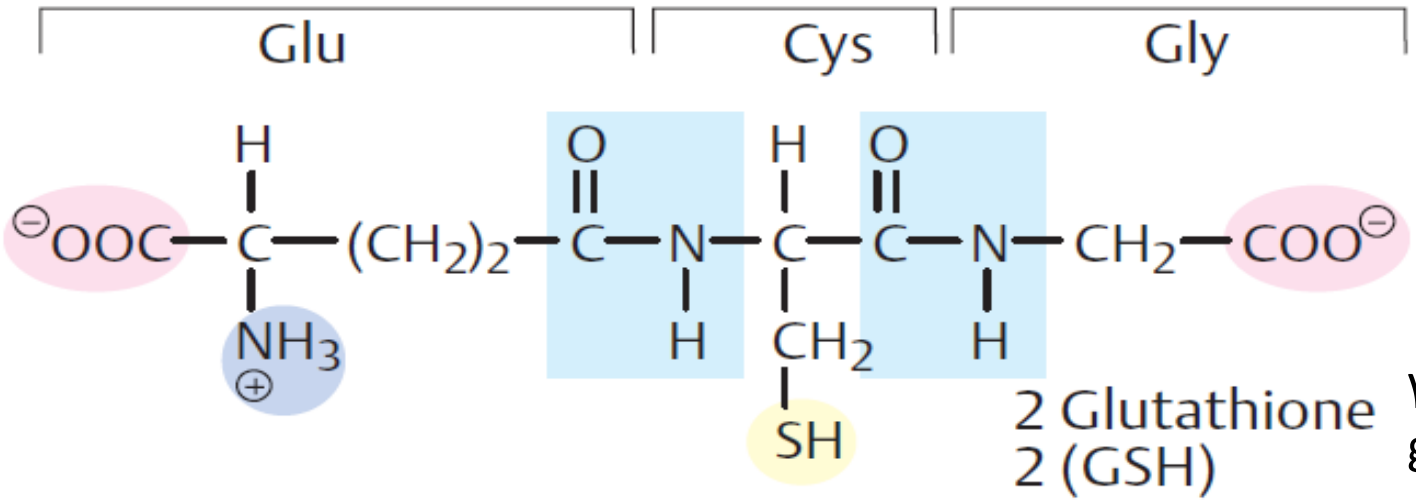


Pseudopeptides

- Peptide bond is read in N→C end. In some proteins **peptide bond is formed by COOH group other than α position.** Such proteins are called Pseudopeptides.



3–13 Formation of a peptide bond by condensation.



➢ Polarity of the peptide bond:

- Like all amide linkages, the -C= 0 and -N H groups of the **peptide bond are uncharged**, and neither accept nor release protons over the pH range of 2 to 12.
- Thus, the charged groups present in polypeptides consist solely of the N-terminal α-amino group, the C-terminal α-carboxyl group, and any ionized groups present in the side chains of the constituent amino acids.

➢ Branched and Circular proteins:

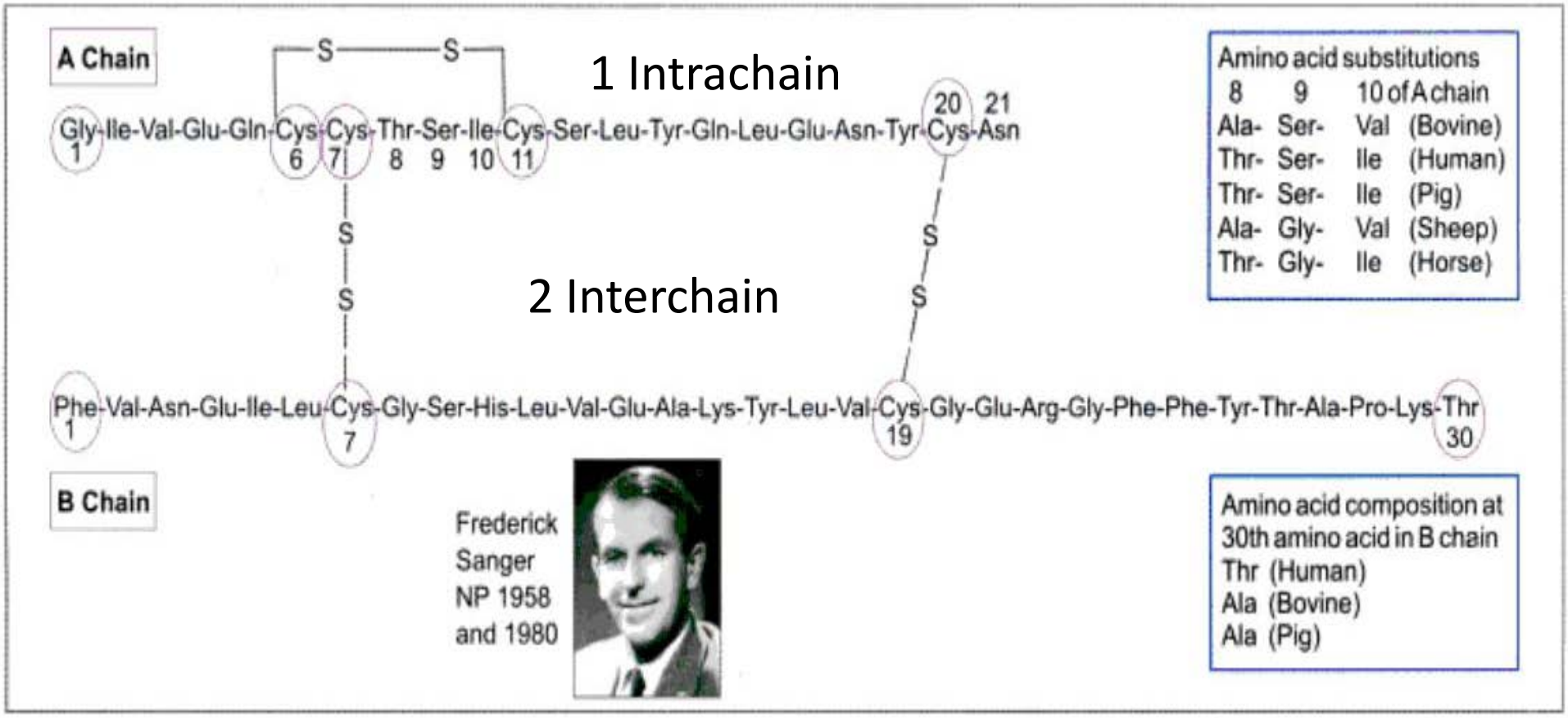
- Generally polypeptide chains are linear.
- Branching points in the chains may be produced by interchain disulphide bridges.
- Rarely proteins may be in circular form: Gramicidine

Primary structure of Insulin – Branched Protein

- Insulin has *2 polypeptide chains*. The A(Glycine chain) has 21 amino acids and B(Phenylalanine chain) has 30 amino acids.

- They are held together by *2 interchain disulphide bonds.* A chain 7^{th} cysteine and B chain 7^{th} cysteine are connected. Similarly A chain 20^{th} cysteine and B chain 19^{th} cysteine are connected.

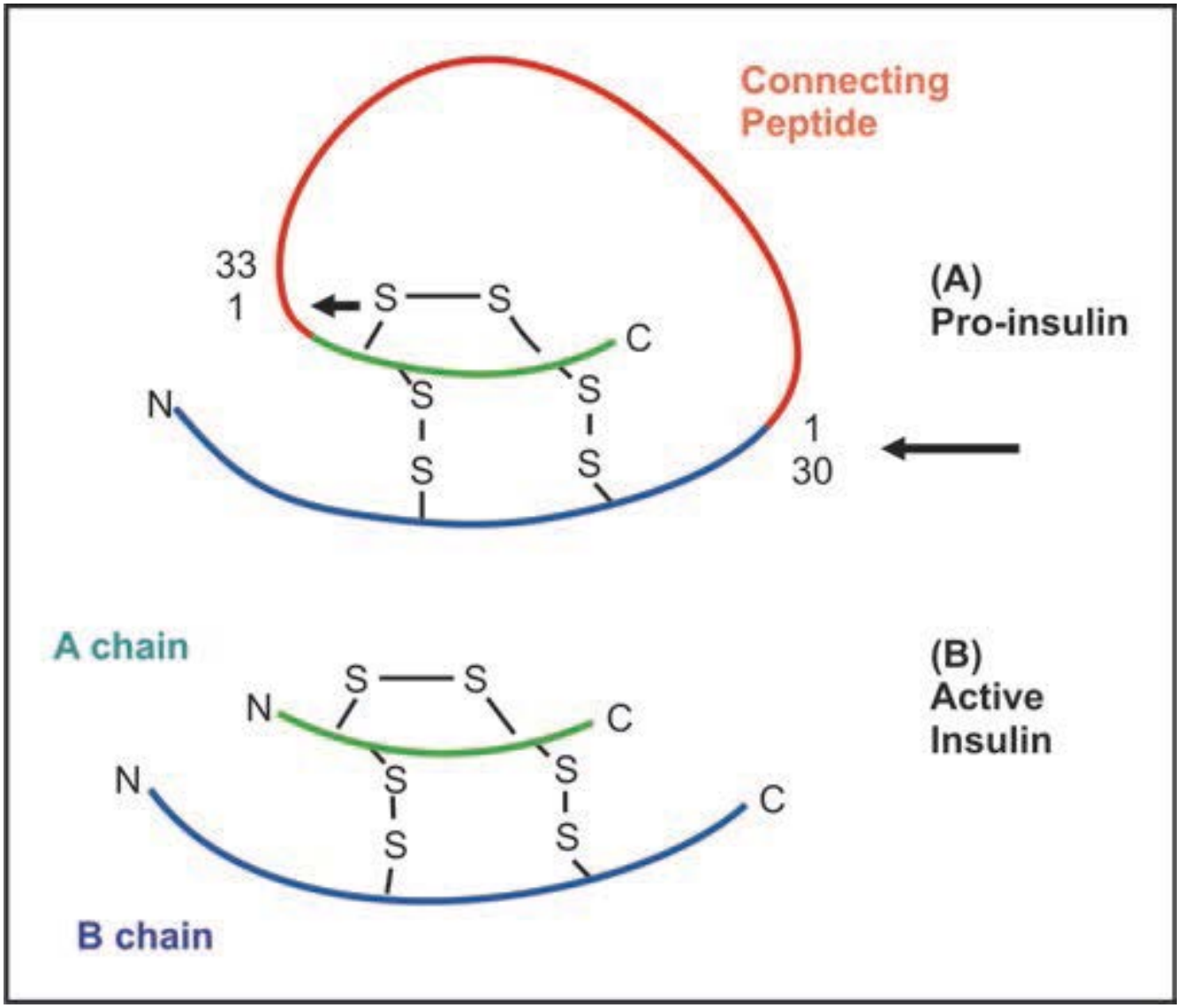
- There is another *intrachain disulphide bond* between 6^{th} and 11^{th} cysteine residues of A chain.

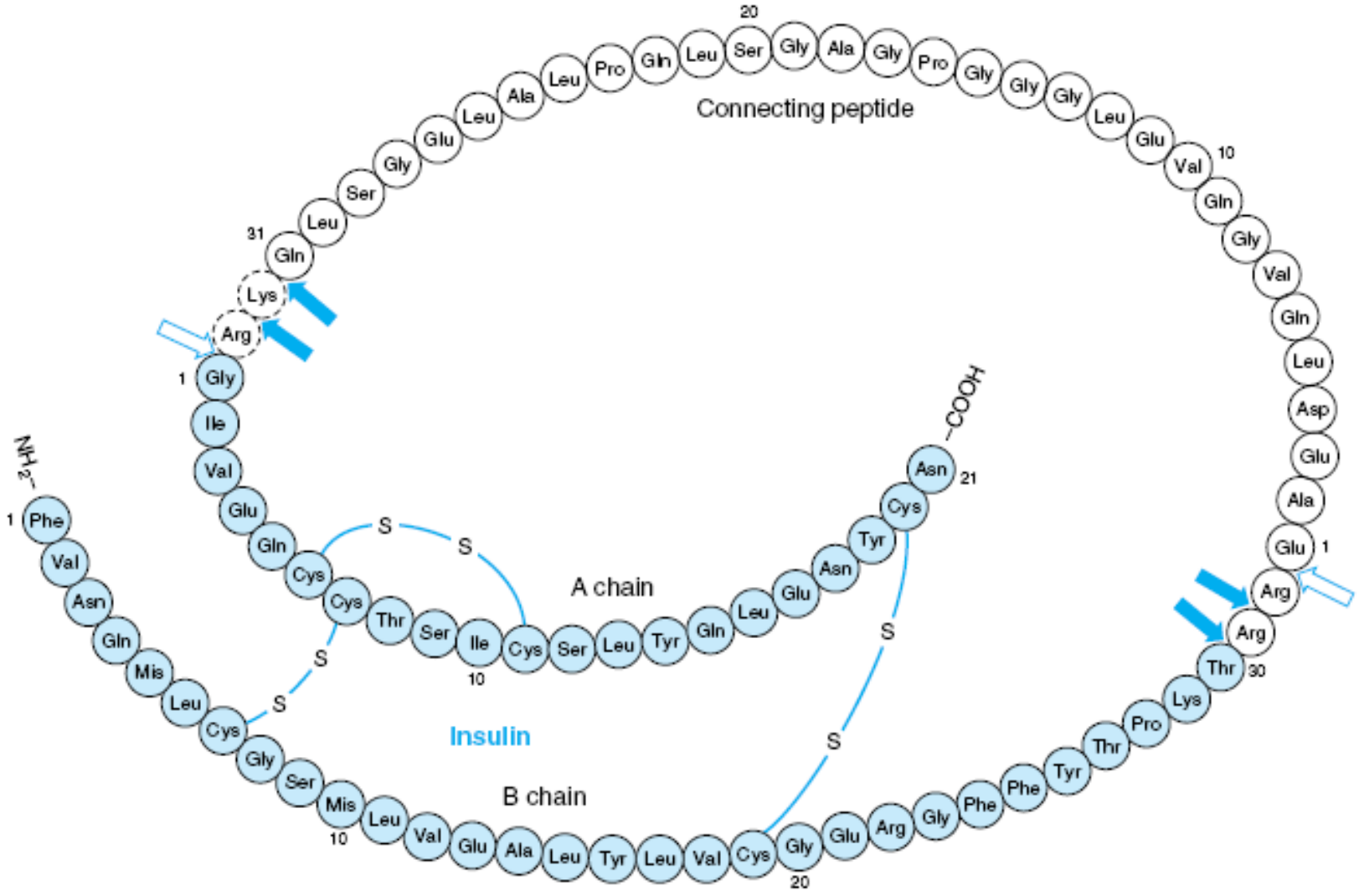

A Chain
S
S
1 Intrachain
Gly-Ile-Val-Glu-Gln-Cys-Cys-Thr-Ser-Ile-Cys-Ser-Leu-Tyr-Gln-Leu-Glu-Asn-Tyr-Cys-Asn
1
6
7
8
9
10
11
20
21
S
S
2 Interchain
S
S
Phe-Val-Asn-Glu-Ile-Leu-Cys-Gly-Ser-His-Leu-Val-Glu-Ala-Lys-Tyr-Leu-Val-Cys-Gly-Glu-Arg-Gly-Phe-Phe-Tyr-Thr-Ala-Pro-Lys-Thr
1
7
19
30
B Chain
Amino acid substitutions
8 9 10 of A chain
Ala- Ser- Val (Bovine)
Thr- Ser- Ile (Human)
Thr- Ser- Ile (Pig)
Ala- Gly- Val (Sheep)
Thr- Gly- Ile (Horse)
Frederick
Sanger
NP 1958
and 1980
Amino acid composition at
30th amino acid in B chain
Thr (Human)
Ala (Bovine)
Ala (Pig)

- The bovine insulin and human insulin are structurally similar, except the terminal amino acid in B chain (Thr → Ala).
- Bovine Insulin may produce antibodies in humans by repeated injections. But de-alaninated porcine Insulin, bearing no antigenic difference from human Insulin will not produce antibodies in diabetic patients even after a long-term use.
- Nowadays human Insulin is being produced by recombinant DNA technology.

➢ Pro-insulin

- Beta cells of pancreas synthesize insulin as a prohormone.
- Proinsulin is a **single polypeptide chain** with 86 amino acids.
- Biologically active insulin (2 chains) is formed by removal of the central portion of the pro-insulin before release.
- The **C-peptide** (connecting peptide) is also released into the circulation.


Connecting
Peptide
33
1
S
S
C
S
S
N
S
S
1
30
(A)
Pro-insulin
A chain
S
S
N
C
S
S
N
C
S
S
B chain
(B)
Active
Insulin


Connecting peptide
A chain
B chain
Insulin
NH2-
-COOH
S
S
S
S
S
S
Phe
Val
Asn
Gln
His
Leu
Cys
Gly
Ser
His
Leu
Val
Glu
Ala
Leu
Tyr
Leu
Val
Cys
Gly
Glu
Arg
Gly
Phe
Phe
Tyr
Thr
Pro
Lys
Thr
Arg
Arg
Glu
Ala
Glu
Asp
Leu
Gln
Val
Gly
Gln
Val
Glu
Leu
Gly
Gly
Pro
Gly
Ala
Gly
Ser
Leu
Gln
Pro
Leu
Ala
Leu
Glu
Gly
Ser
Leu
Gln
Lys
Arg
Gly
Ile
Val
Glu
Gln
Cys
Cys
Thr
Ser
Ile
Cys
Ser
Leu
Tyr
Gln
Leu
Glu
Asn
Tyr
Cys
Asn
1
10
20
30
31
1
10
20
21

➢ IMPORTANCE OF PRIMARY STRUCTURE

- A protein with a specific primary structure will automatically form its natural 3D shape. So the primary structure of proteins **determine the higher levels of protein structure**. So the biological activities of proteins are dependent on the primary structure.
- Understanding the primary structure of proteins is important because many **genetic diseases** result in proteins with abnormal amino acid sequences, which cause improper folding and loss or impairment of normal function. E.g. Sickle cell anemia.
- The primary structures of the normal and the mutated proteins are known, this information may be used to diagnose or study the disease.

SECONDARY STRUCTURES

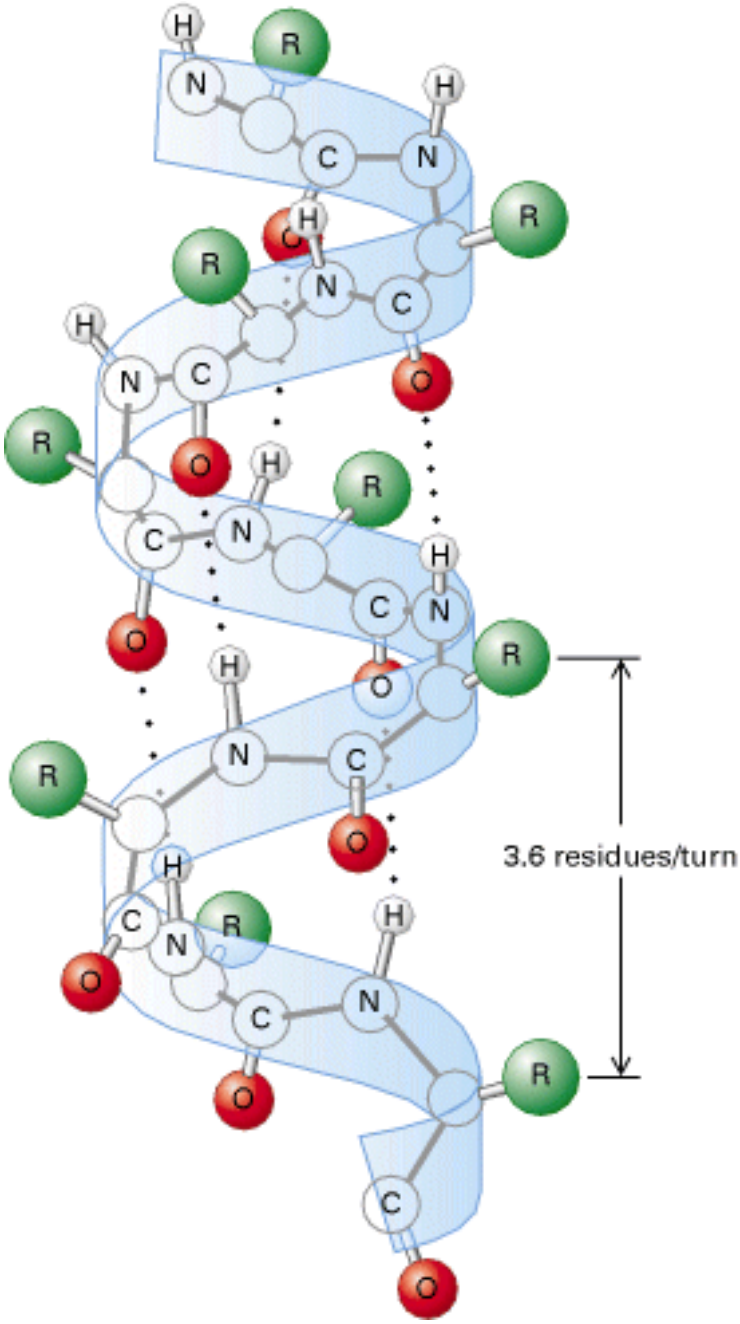
- It is the steric relationship of residues which are about 3-4 AA s apart in the linear sequence.
- Formed by folding of primary structure into geometrically ordered units.
- Stabilized by Noncovalent bonds.

- Types –

1. α Helix
2. β Pleated sheet
3. β Bends
4. Loop regions
5. Triple helix

α-Helix

- Most common & stable conformation
- Spiral structure, right handed.
- Tightly packed, coiled polypeptide backbone core.
- Side chain extend outwards
- Stabilized by H bonding b/w NH & C=O groups of the main chain.
- 3.6 residues of Amino acids per turn
- Distance between each AA residue is 1.5 A.
- Pitch is 5.4 A



- Hydrogen bonds are parallel to central axis and R-groups are roughly perpendicular to central axis.

- Left handed alpha helix is rare, because amino acids found in proteins are of L-variety, which exclude left-handedness.

- **Proline** form kinks in helix due to absence of α-amino group for hydrogen bonding. Hence these are called helix breakers. Cannot form α-Helix.
- **Charged amino acids** (for example, glutamate, aspartate, histidine, lysine, or arginine) also disrupt the helix by forming ionic bonds, or by electrostatically repelling each other.
- Amino acids with bulky side chains, such as **tryptophan**, or aminoacids, such as **valine or isoleucine**, that branch at the β-carbon can interfere with formation of the α-helix if they are present in large numbers.

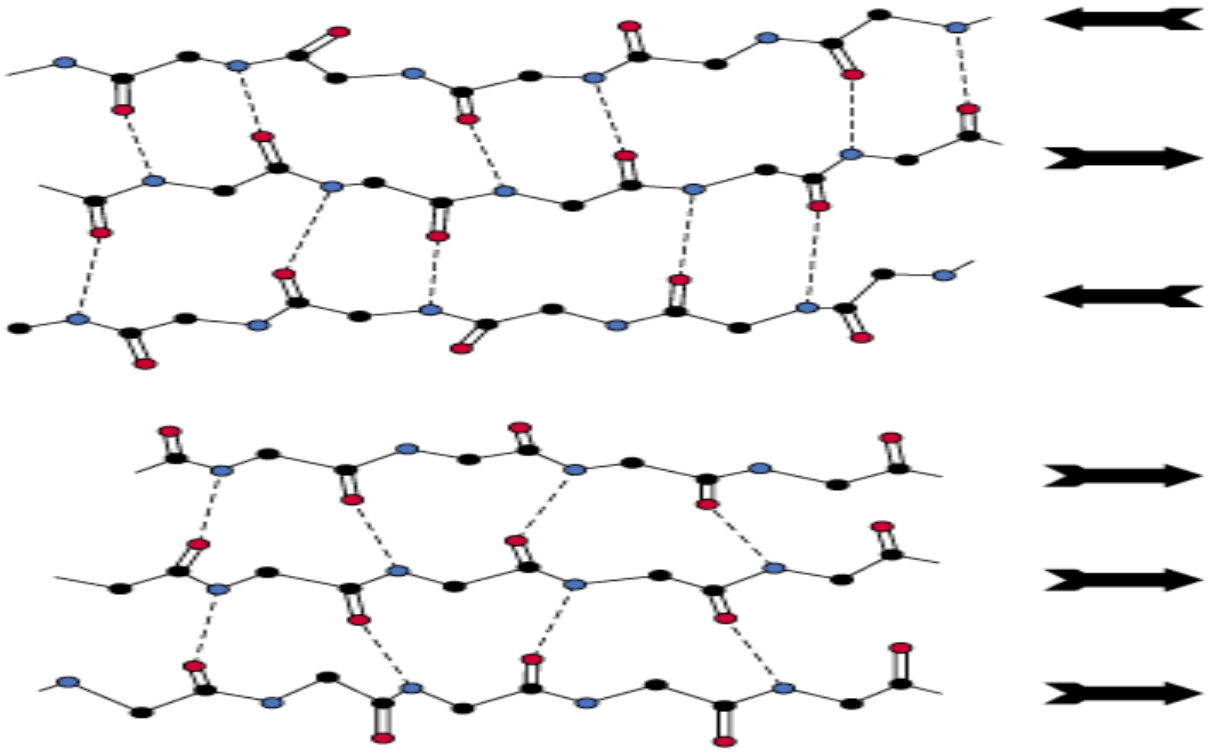
- Clinical importance – gives mechanical support by forming strong stiff bundles of fibres.

- Example – fibrous proteins – keratin, myosin, tropomyosin, fibrin

  globular proteins – hemoglobin, myoglobin

β Pleated sheet

- Almost fully extended structure composed of more than 2 peptide chains.
- Distance between amino acid residues is 3.4A



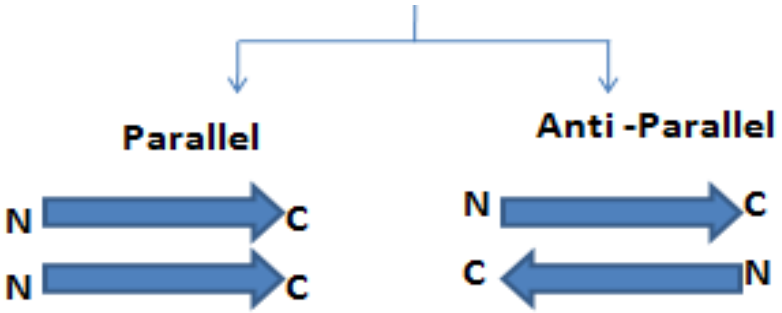
- **Features:**

- These are non-helical, extended, zig-zag linear strands held closely by H2 bonds between NH & C=0.

- The hydrogen bonds between adjacent strands are perpendicular to long axis of strands

- All peptide bonds are involved in ‘H’ bonding

- **2 types**



- Eg for parallel – Flavodoxin, carbonic anhydrase

- Eg for anti parallel – silk fibroin

Reverse Turns or β-bends

- β-Bends reverse the direction of a polypeptide chain, helping it form a compact, globular shape.
- β bends play a role in connecting the successive strands of antiparallel β sheets
- stabilized by hydrogen and ionic bonds.
- Proline and glycine are seen mostly in beta bends.

Loops/ Coils

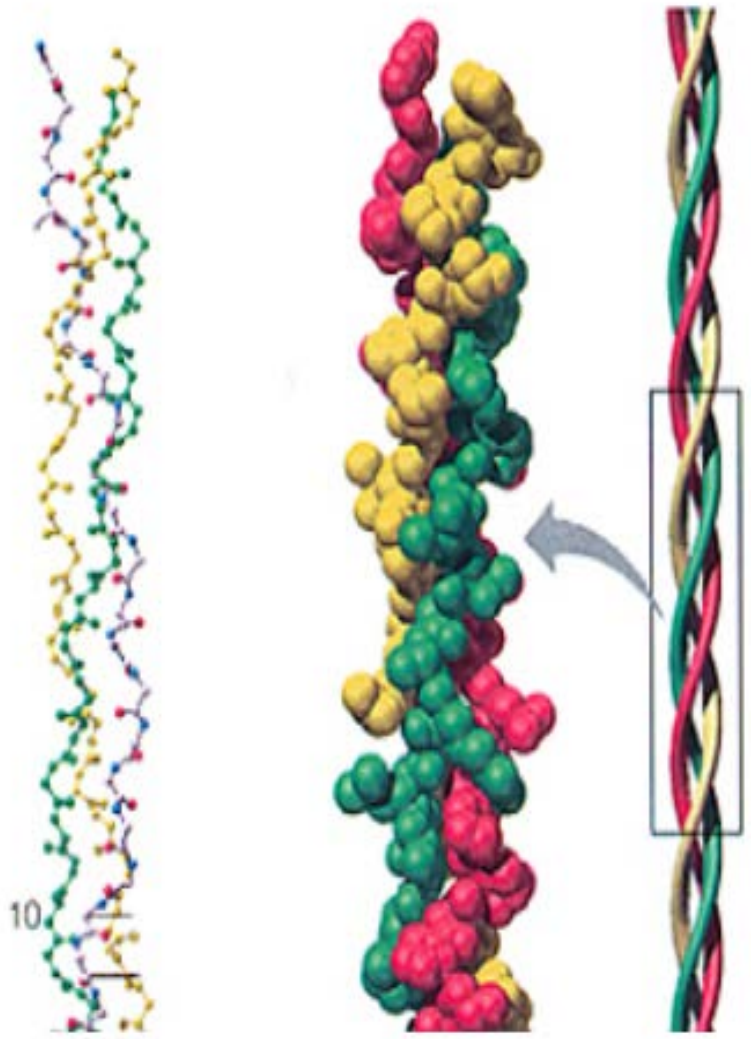
- Approximately one half of globular protein is organized into repetitive structures, such as the α-helix and/or β-sheet.

- The remainder of the polypeptide chain is having a **loop or coil conformation**. These nonrepetitive secondary structures have *a less regular structure* than those α-helix and/or β-sheet.

TRIPLE HELIX

- **Collagen** forms a triple helix. Collagen is rich in proline and hydroxy proline and cannot form a α-helix or β-Pleated sheet.

- Stabilized by both non covalent & covalent bonds.



- Has 3 polypeptide chains, each called the α chain.
- Rich in proline, OHproline, OHlysine and glycine.
- Each chain twist as a left handed helix and inturn the 3 chains wrap around each other with a right handed twist.
- 3 residues per turn

Super secondary structures (Motifs)

- Various combinations of secondary structure, called super secondary structure, are commonly found in *globular proteins*.
- Intermediate between secondary and tertiary structures. E.g.
- β- α- β unit
- Greek key
- β-meander

- **β- α- β unit :**

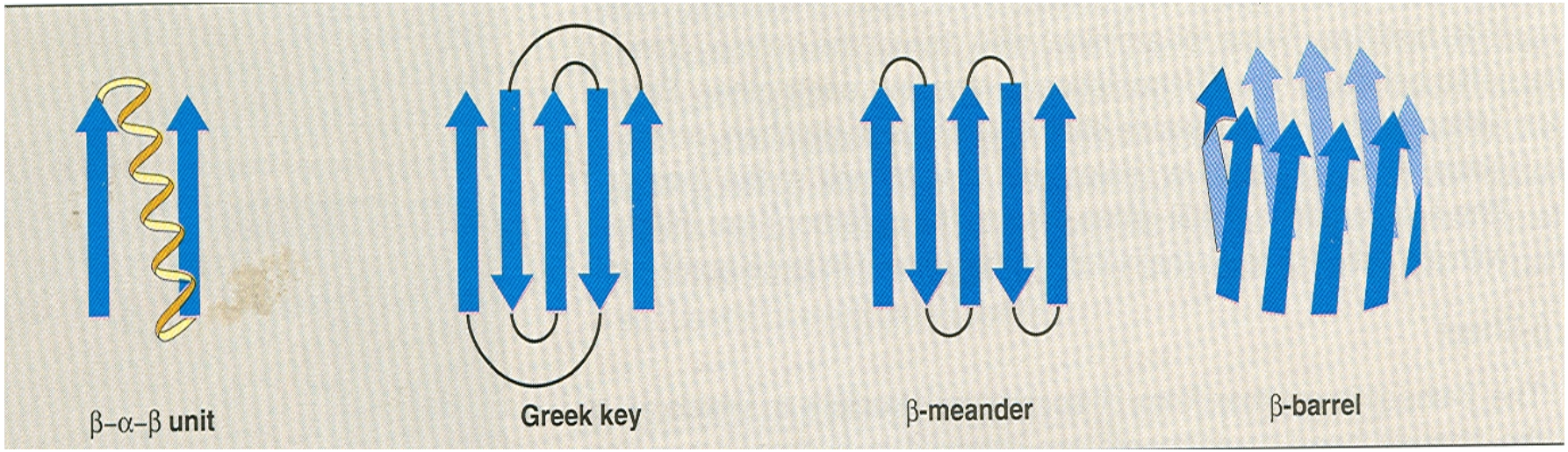
- Two parallel β-pleated sheets connected by an intervening strand of α –helix.

- **Greek key :**

- A conformation design often found on classical greek pottery.

➢ β-meander :

- Five β-Pleated sheets connected by reverse turns.
- The β-meander contains nearly as many hydrogen bonds as an α-helix and it is very stable.

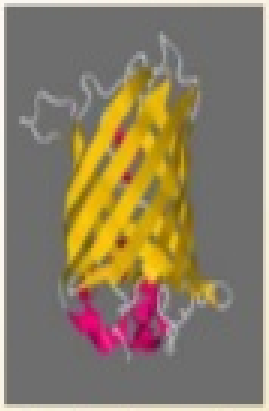


Super secondary structures

- Some of the common structural motifs combining α - helices and β - sheets
- found in globular proteins

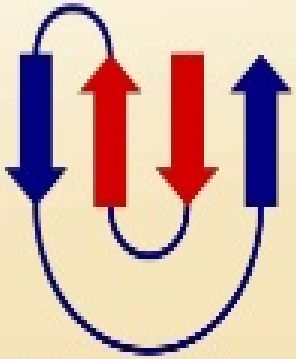
These are :-

- β- α- β unit
- Greek key
- β-meander



Beta barrel

beta-alpha-beta motif



Greek key motif

β-meander motif

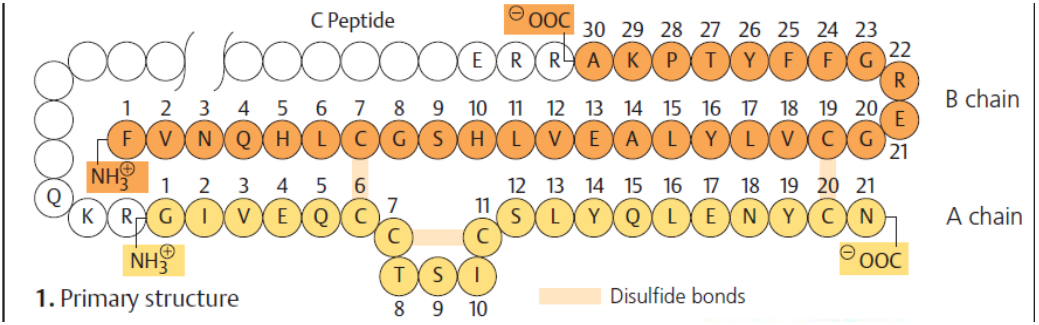
TERTIARY STRUCTURE

- It refers to further folding of polypeptide chain of secondary structure giving the **compact three dimensional conformation** leading to formation of **domains.**
- It refers to the spatial relationship between the amino acid residues located far apart from each other in their linear structure, but brought closer as a result of folding.
- "Tertiary" refers both to the folding of domains and the final arrangement of domains in the polypeptide.

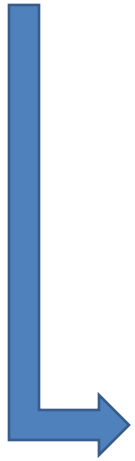
- Stabilized primarily by noncovalent interactions.
- This type of arrangement ensures thermodynamical stability of the molecule.
- A polypeptide with 200 amino acids normally consists of two or more domains.
- Phenyl alanine hydroxylase contains 3 domains.

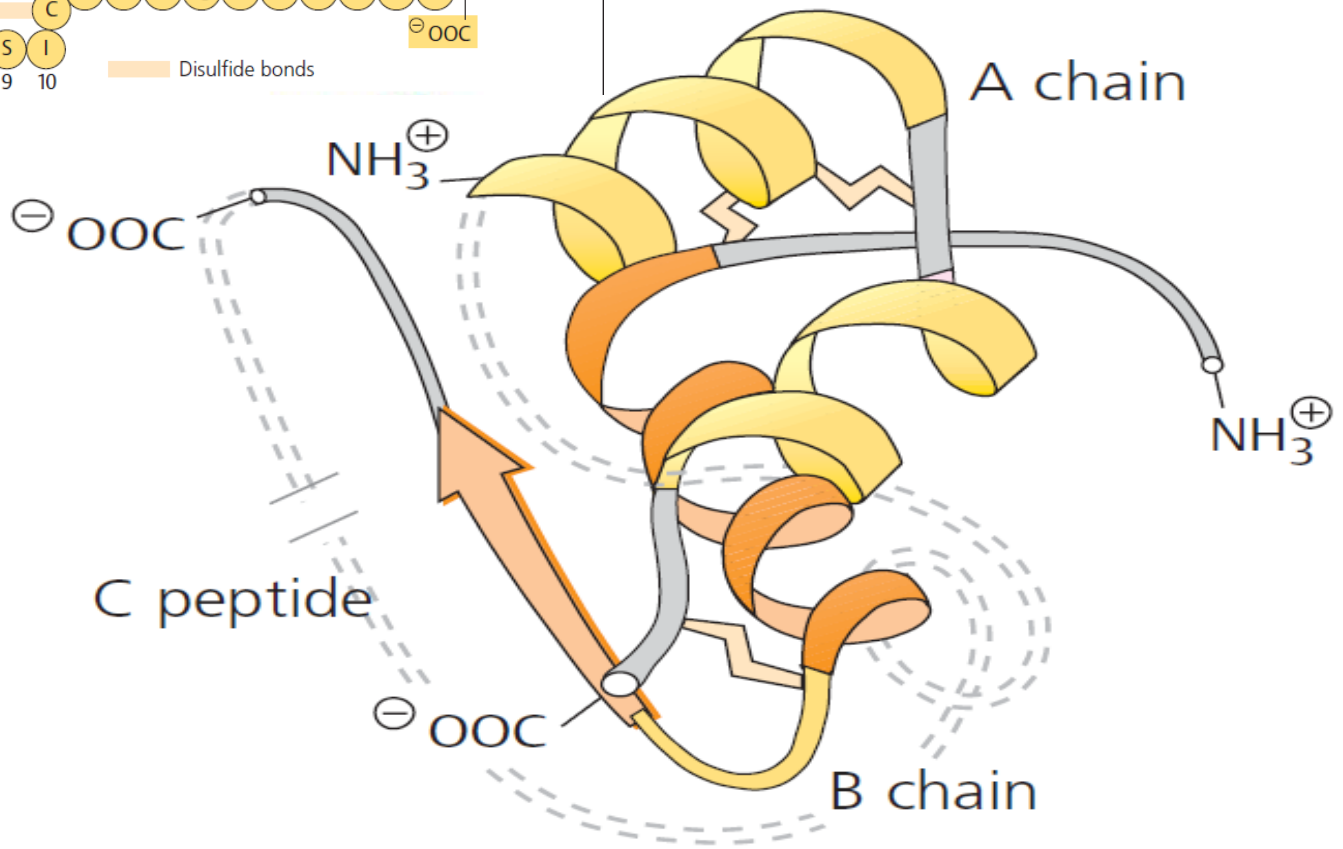
- The core of a domain is built from combinations of super-secondary structural elements (**motifs**).
- Folding of the peptide chain within a domain usually occurs independently of folding in other domains. Therefore, each domain has the characteristics of a small, **compact globular protein** that is structurally independent of the other domains in the polypeptide chain.

3D structure of insulin



1. Primary structure





Specific structural motifs in common proteins

Protein	Structural Motifs present
Myoglobin	Alpha helix and beta pleated sheet
Collagen	Triple helix
Keratin	Coiled coil
Elastin	No specific motif
Superoxide Dismutase	Antiparallel beta pleated sheet

QUATERNARY STRUCTURE

- Proteins with > 1 polypeptid chains(polymers) – form functional protein.
- Proteins will lose its function when subunits are dissociated.
- Stabilized by noncovalent bonds.

- Such proteins are termed as oligomers.

  1 chain – monomer/subunit

  2 chains – Dimer

  4 chains – tetramer

- Homodimer – contains 2 copies of same polypeptide chain
- Heterodimer - contains 2 different polypeptide chain

HAEMOGLOBIN

- Has 4 polypeptde chains – heteromeric.
- α, β, γ, δ, ε

- HbA major form in adults→ 2α, 2β chains
  -HbA2 minor form in adults→2 α, 2δ
  -HbF major form in Fetus→ 2 α, 2γ
- Abnormal Hb –

  thalassemia – a or b

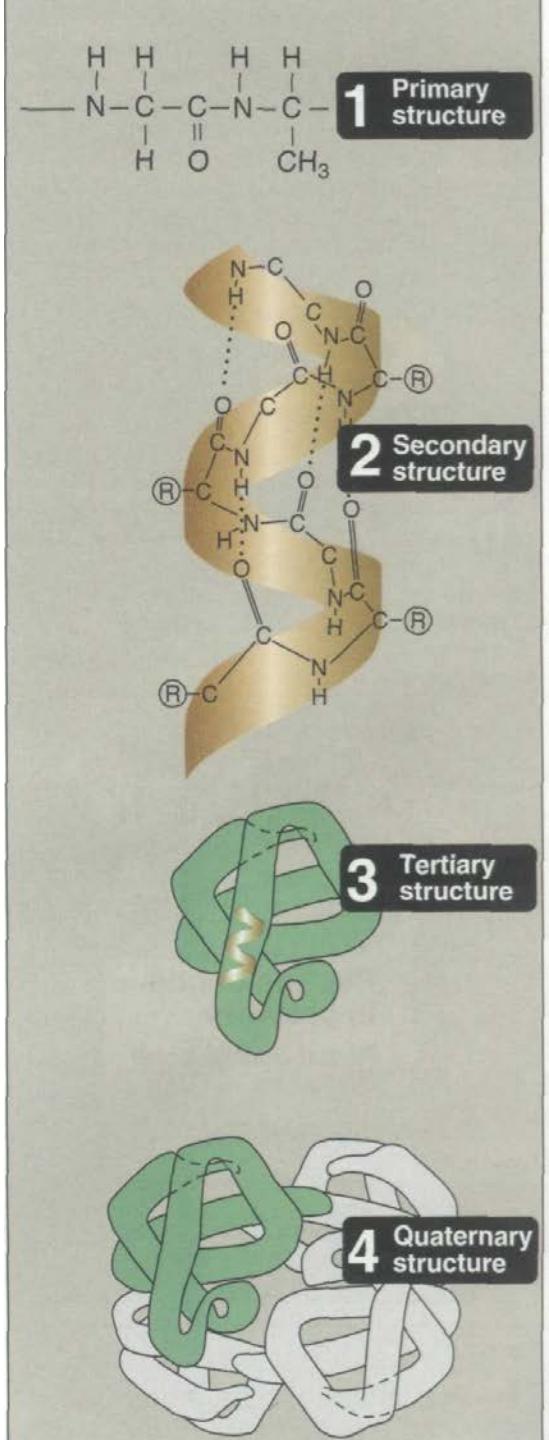
  Hb S – glutamic acid of b6 by valine

  Hb C – glutamic acid of b6 by lysine

Importance of higher structures of proteins

- Proteins are **biologically active** only in their three dimensional conformations. So the biological activities of proteins are attributed to their higher structures.
- If a protein loses its three dimensional form(as in denaturation), it loses its biological activity.

1. **Primary structure** of protein means the order of amino acids in the polypeptide chain and the location of disulfide bonds, if any.
2. **Secondary structure** is the steric relationship of amino acids, close to each other.
3. **Tertiary structure** denotes the overall arrangement and interrelationship of the various regions, or domains of a single polypeptide chain.
4. **Quaternary structure** results when the proteins consist of two or more polypeptide chains held together by non-covalent forces.


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

MCQs

1. Stable Peptide bond occurs in the following isomeric form

a) cis form
b) Trans form
c) Both
d) None

- 2. Covalent bond is

a) Hydrogen bond
b) Electrostatic bond
c) Disulphide Bond
d) Hydrophobic Bond

- 3. β Pleated sheets are component of

a) Primary structure
b) Secondary structure
c) Tertiary structure
d) Quaternary structure

- 4. Triple Helix is found in

a) Collagen
b) Elastin
c) Hemoglobin
d) Myoglobin

- 5. Domains are component of

a) Primary structure
b) Secondary structure
c) Tertiary structure
d) Quaternary structure