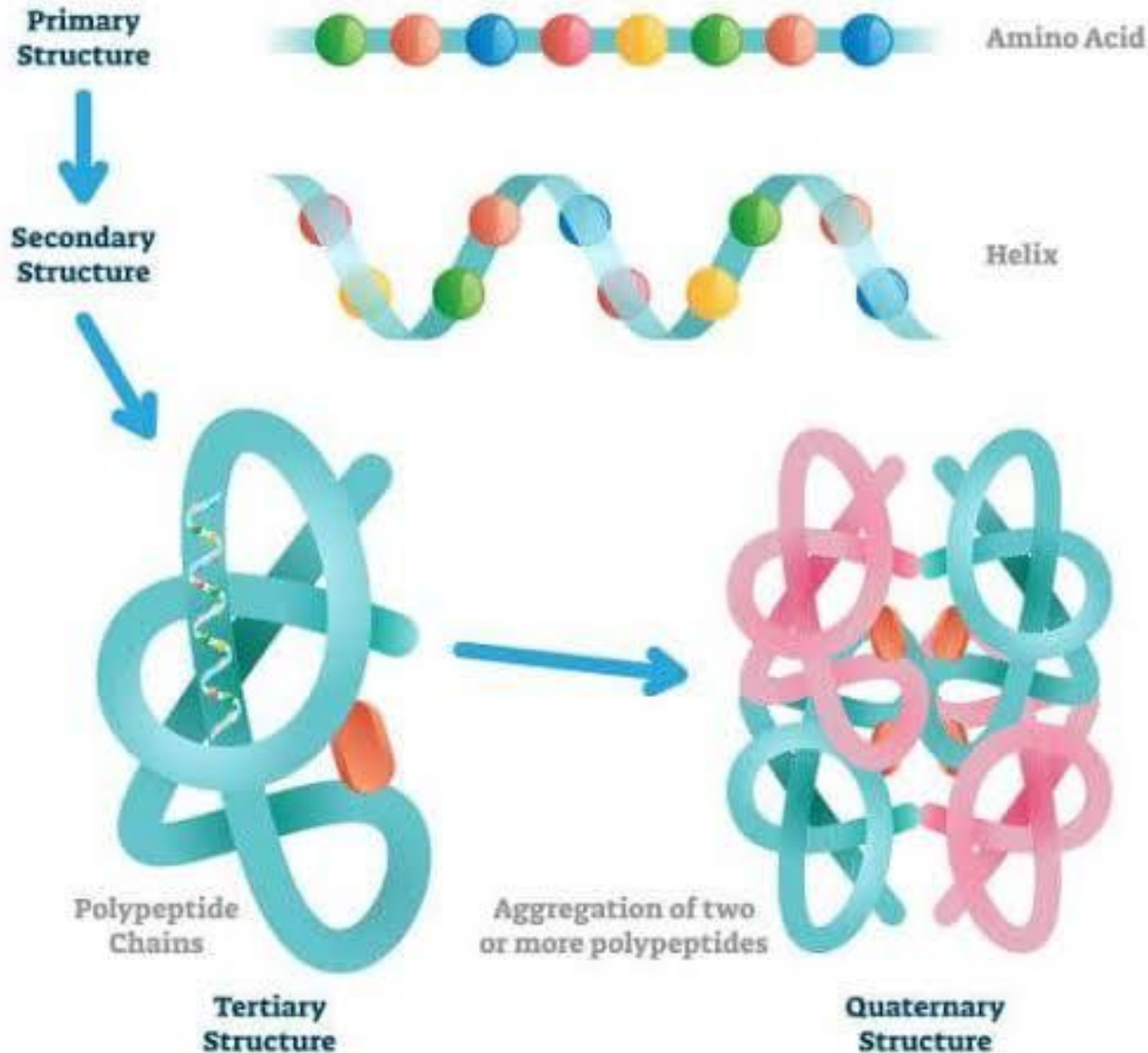


PROTEIN STRUCTURE



Protein Structure

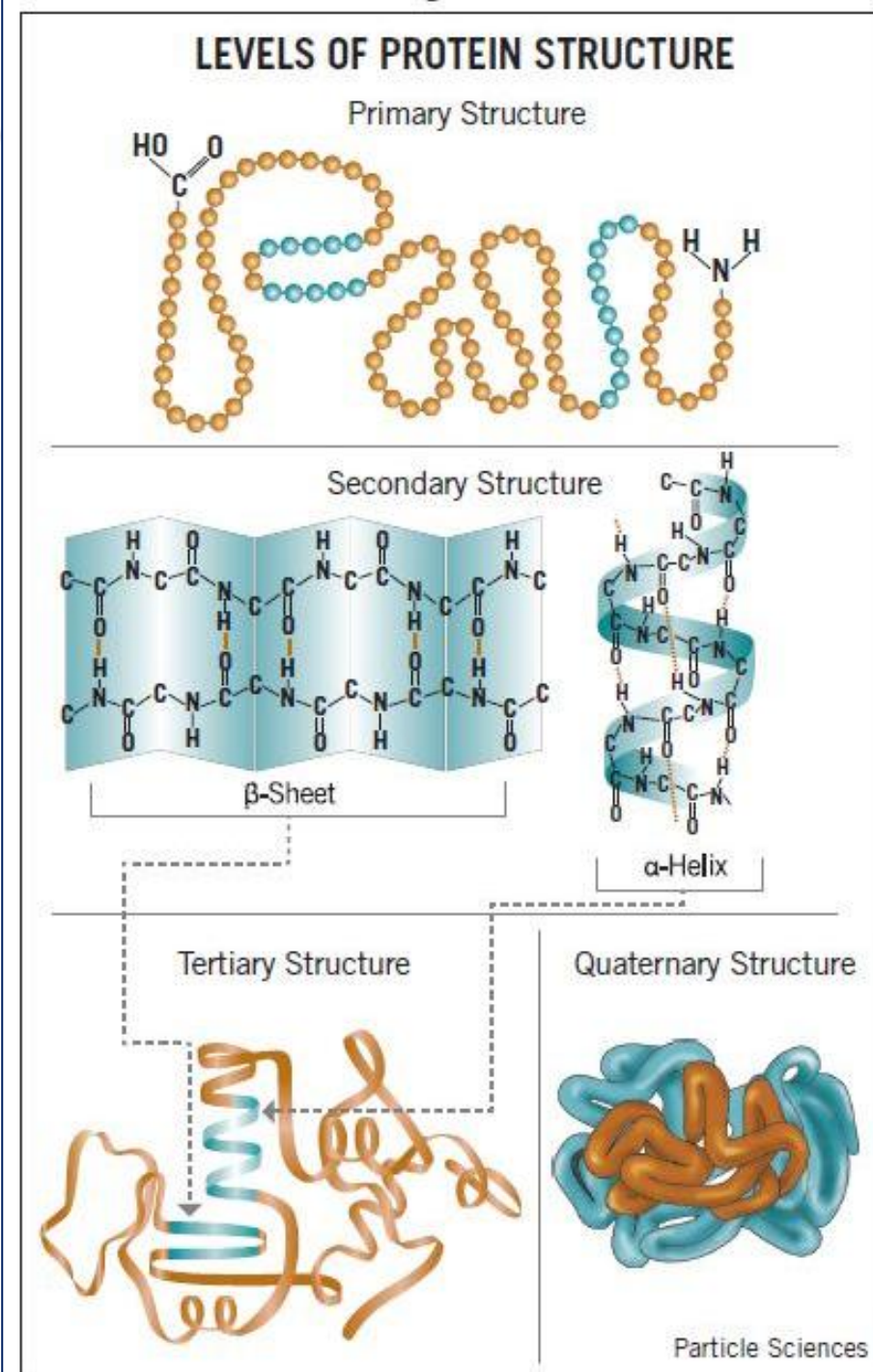
Alkmini Anastasiadi, PhD

From amino acids to proteins

- Each unit of a peptide is called an amino acid residue—that is, what remains after the release of H_2O when an amino acid forms a peptide bond in the peptide chain
 - Dipeptides, tripeptides... oligopeptides ($\sim 12\text{--}20$ amino acid residues), polypeptides.
- Proteins > 50 amino acids (insulin: 51 amino acids)
- A protein usually consists of >1 polypeptide (homopolymer or heteropolymer).

Protein Structure

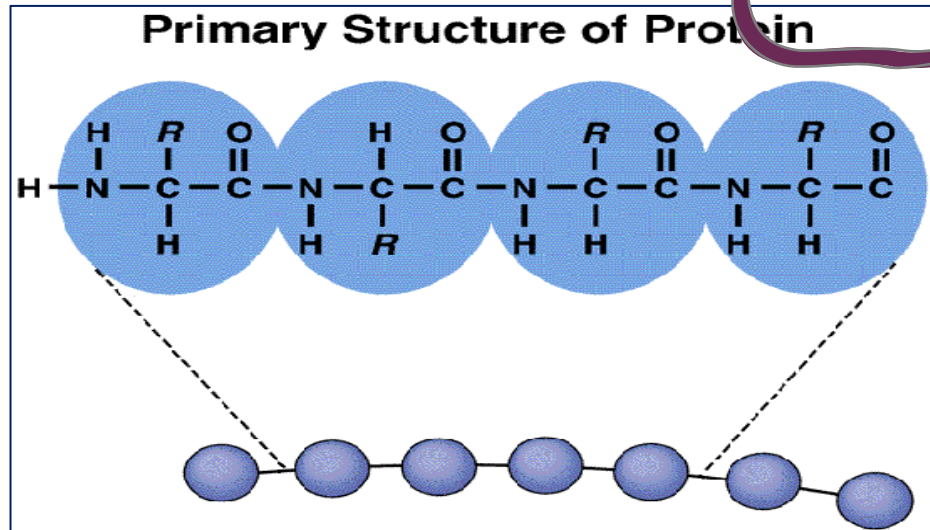
- **Primary Structure:** the amino acid sequence in the peptide chain
- **Secondary Structure:** local spatial conformation of the polypeptide backbone
- **Tertiary Structure:** the three-dimensional shape of a polypeptide
- **Quaternary Structure:** structure of proteins which are composed of two or more smaller polypeptide chains (also referred to as subunits)



Protein Structure

Primary

- Definition: The unique linear sequence of α-L-amino acids linked sequentially by covalent peptide bonds
 - **Constant part**: «backbone» of C and N atoms from consecutive amino acids linked by covalent (peptide) bonds
 - **Variable part**: side chains

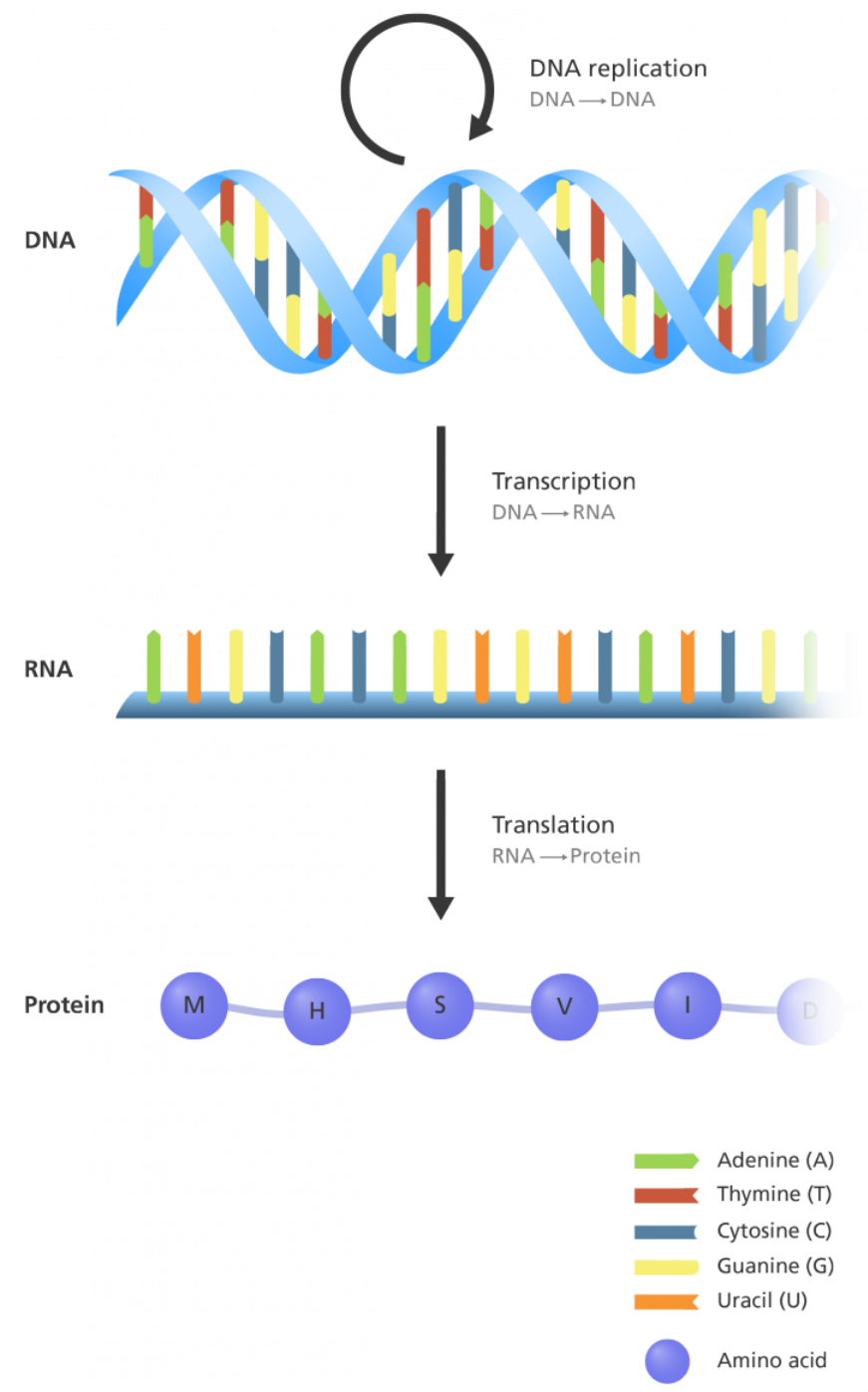
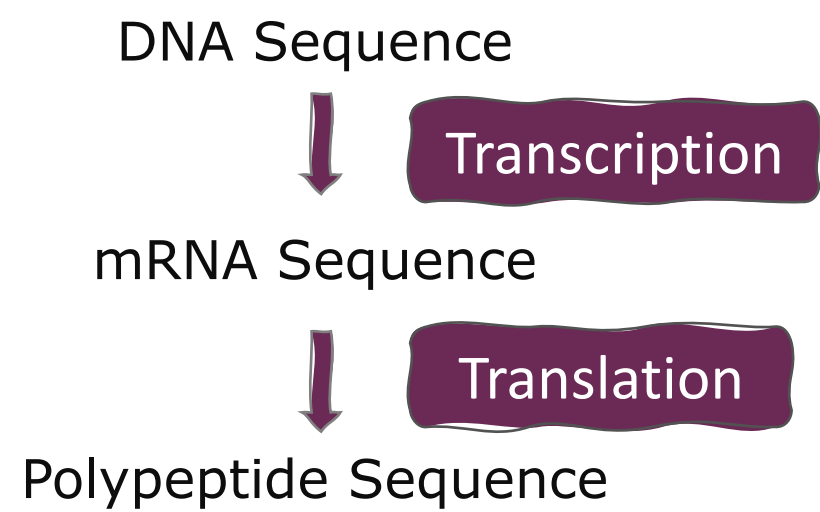


They give each protein its unique characteristics

Protein Structure

Primary

- Genetic Encoding: Directly specified by mRNA codon sequences derived from genomic DNA



Protein Structure

Primary

- Clinical Correlations & Medical Impact:
 - Primary Structure = the Master Blueprint: **Dictates all higher-order folding** (secondary, tertiary, quaternary conformations).
 - Medical Rule: A **single** amino acid **substitution** in a protein of hundreds of residues can **alter** molecular surface **topology**, **collapse** higher-order **structure**, or **eliminate** biological **function**

Protein Structure

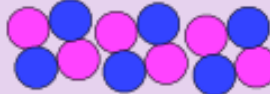
Primary

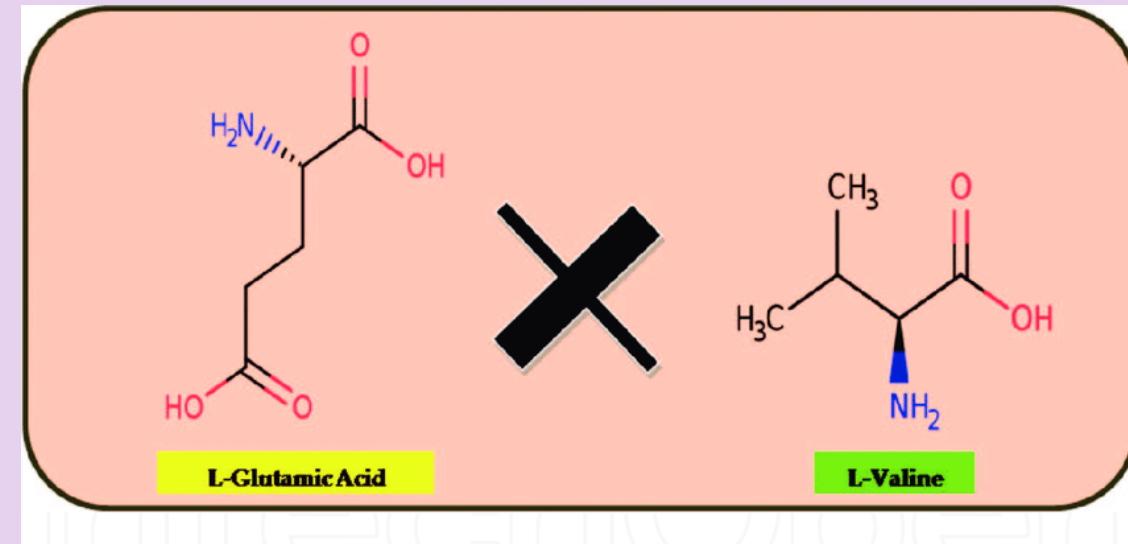
- A change in DNA sequence (mutation) → can affect protein structure and function
 - Neutral mutation → Protein function is not affected
 - Non-functional mutated protein
 - Intermediate cases → the effect on function **depends on the nature of the substituted amino acids and their role in the protein**

Protein Structure

Primary

The example of sickle cell disease

	Normal	Missense Mutation
Partial DNA Sequence of Beta Globin Gene:	CCT GAG GAG GGA CTC CTC	CCT GTG GAG GGA CAC CTC
Partial RNA Sequence:	CCU GAG GAG	CCU GUG GAG
Partial Amino Acid Sequence for Beta Globin:	Pro — Glu — Glu	Pro — Val — Glu
Hemoglobin Molecule:		
Red Blood Cell:		

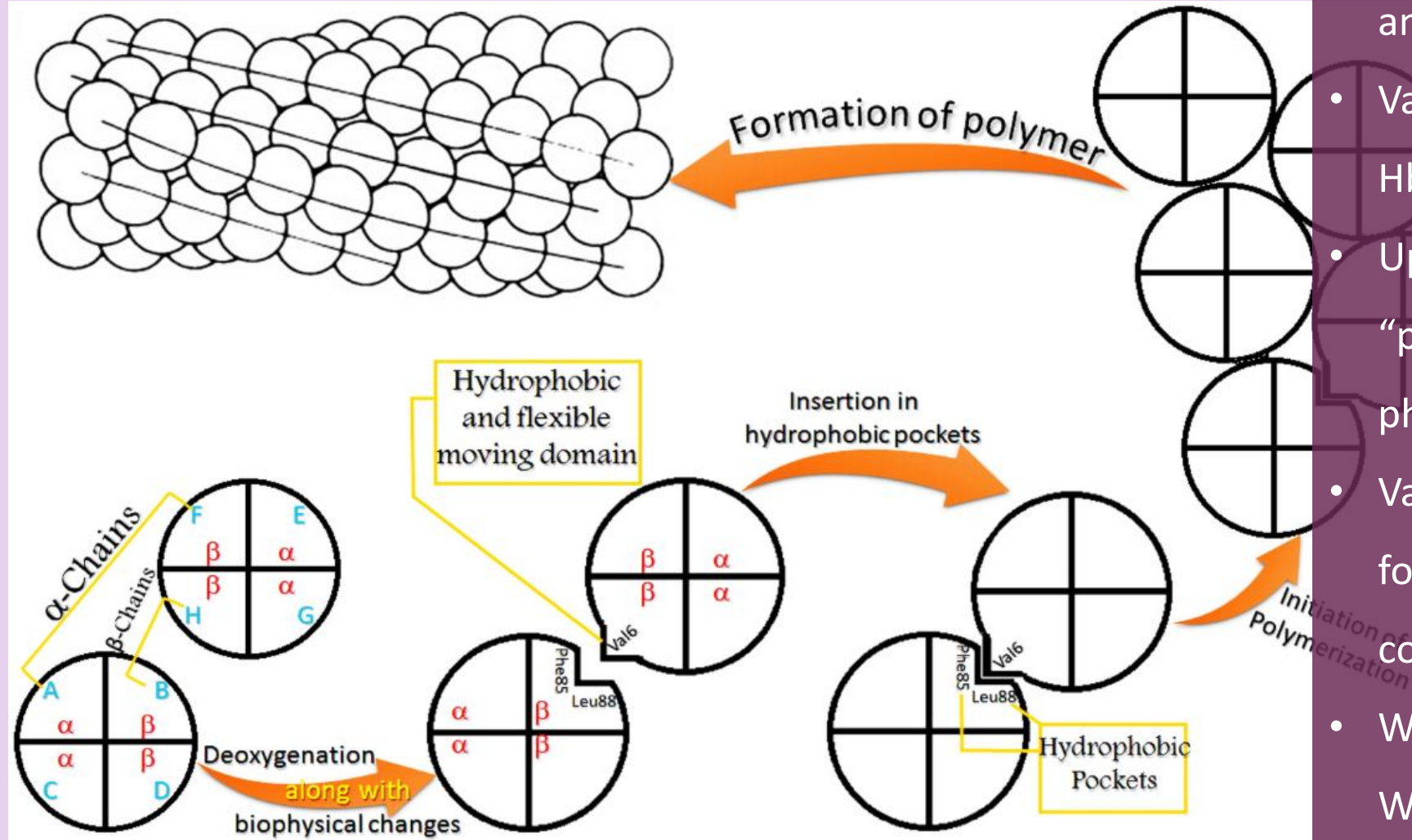


- Why is glutamic acid to Valine such an impactful change?
- Why when valine is exposed it leads to hemoglobin polymerization?

Protein Structure

Primary

The example of sickle cell disease

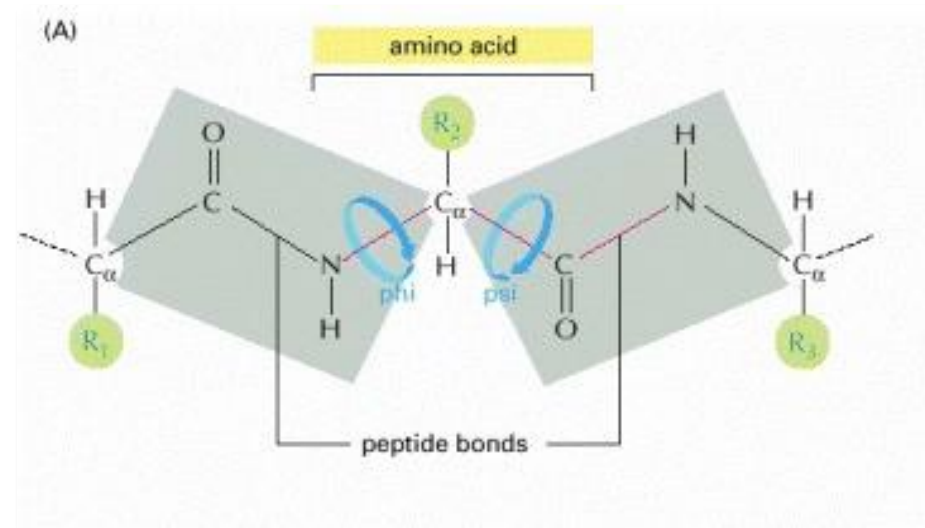


- Acidic amino acid $\rightarrow$ hydrophobic amino acid
- Valine is exposed on the surface of HbS
- Upon deoxygenation a hydrophobic "pocket" containing leucine and phenylalanine is also formed on Hb
- Valine and the hydrophobic pocket are forced by the hydrophobic effect to come close
- When in close proximity, van der Waals forces develop between them

Protein Structure

Secondary

- It is the way in which consecutive segments of the polypeptide chain fold (stereoconformation).
 - How does it fold since the peptide bond does not allow rotation?
 - ➔ The bonds formed by carbon (C_α) with the nitrogen (N) of the amino group and with the carbon (C) of the carboxyl group can rotate around their axes by **angles ϕ and ψ** , respectively, which range from -180° to 180° (depending on the side group).



Protein Structure

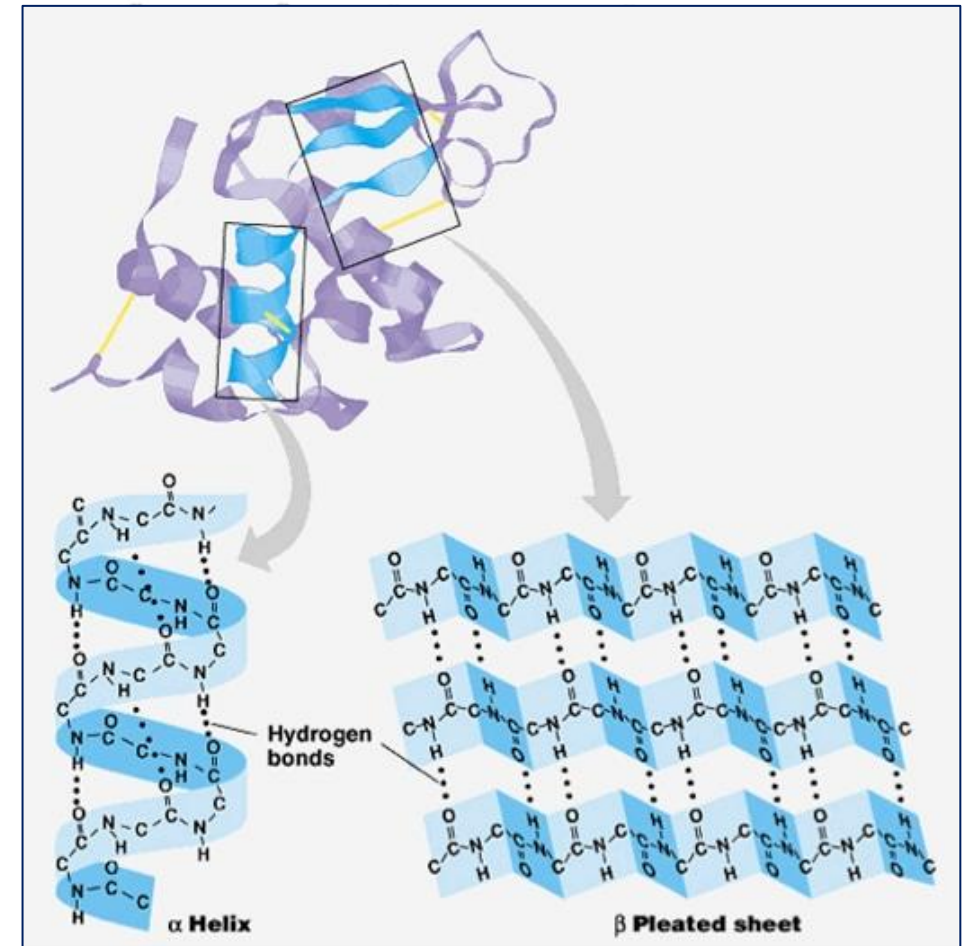
Secondary

- The rotation of these bonds → natural folding of peptide chains
 - Two common folding patterns: **α -helix** and **β -strand**
 - The models of these two structures were described with remarkable precision by Pauling, Corey, and Bransons in 1951.

The folds are stabilized by hydrogen bonds between the N-H and C=O groups of the polypeptide backbone

*THE STRUCTURE OF PROTEINS: TWO HYDROGEN-BONDED
HELICAL CONFIGURATIONS OF THE POLYPEPTIDE CHAIN*

BY LINUS PAULING, ROBERT B. COREY, AND H. R. BRANSON*

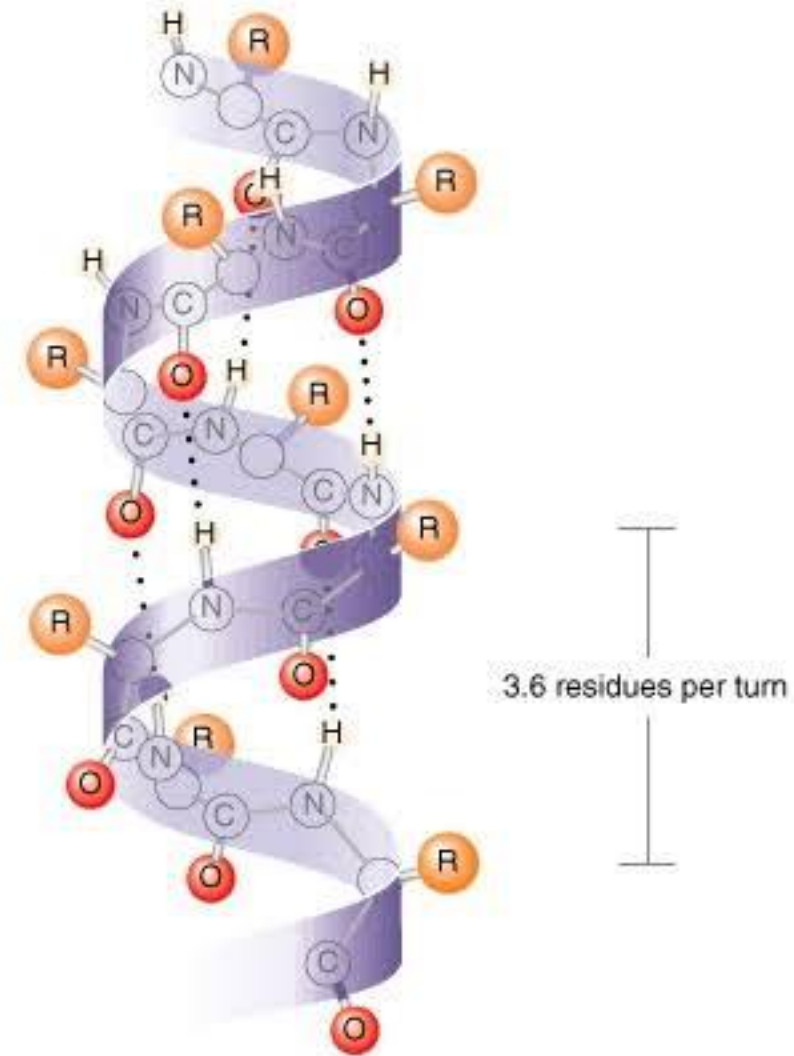


Protein Structure

Secondary

α -helix

- It is most often **right-handed**
- The **side chains** point **outward** from the helix
- **Hydrogen bonds parallel** to the **helix axis** (between the oxygen of the carboxyl group of amino acid n and the hydrogen of the amino group of amino acid $n+4$)

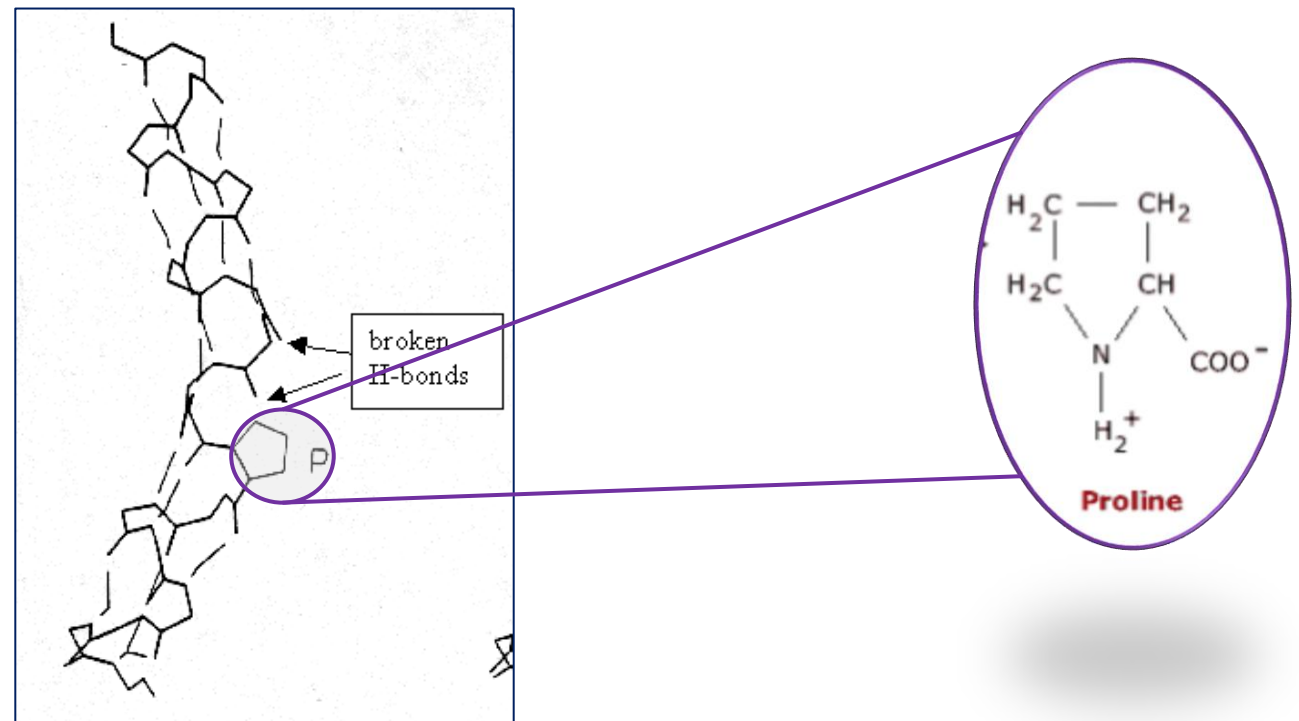


Protein Structure

Secondary

α -helix

- The number of amino acids involved in an α -helix varies from helix to helix and from protein to protein
- Common: polyleucine, polyalanine /
Not common: polyaspartic acid, polyglutamic acid (negative charge)
- Proline = **helix breaker** due to its structure!

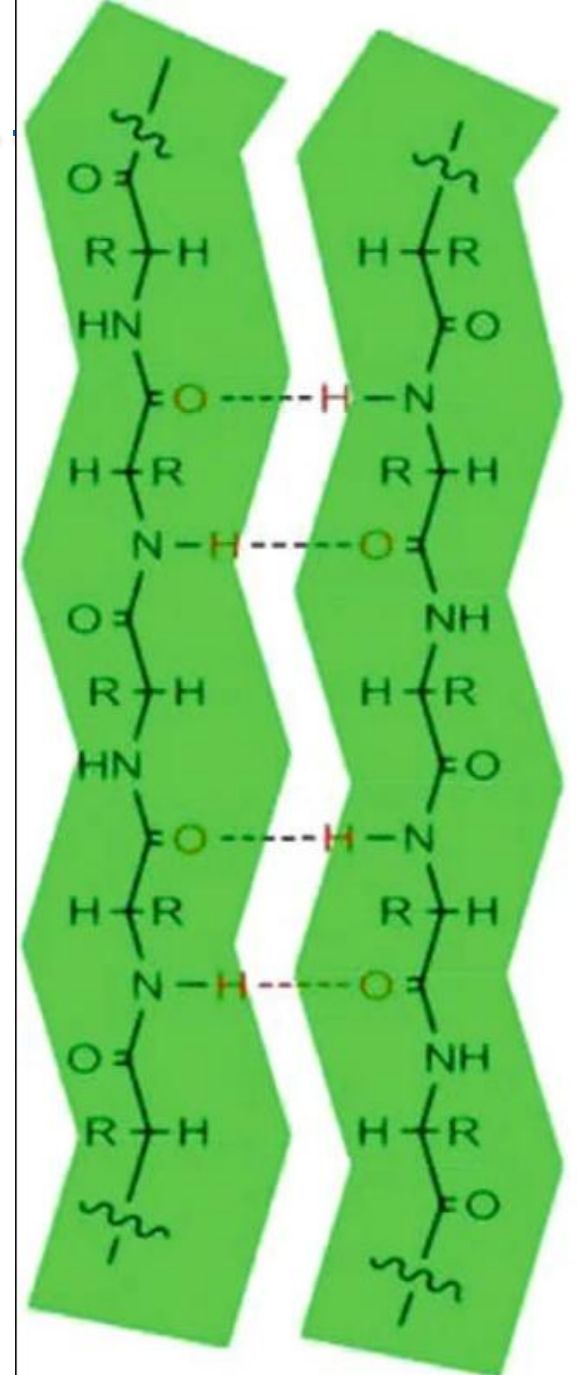


Protein Structure

Secondary

β -strand / β -pleated sheet

- β -pleated sheets are extended β -strands linked by interchain hydrogen bonds
 - β -strand: consists of a stretch of a polypeptide chain (typically 3 to 10 amino acids long) in an almost fully extended, zigzag conformation
- The **hydrogen bonds** of the main chain are formed **between adjacent β -strands**
- The **side chains** of consecutive amino acids → **alternate above and below** the plane of the fold



Beta pleated sheets

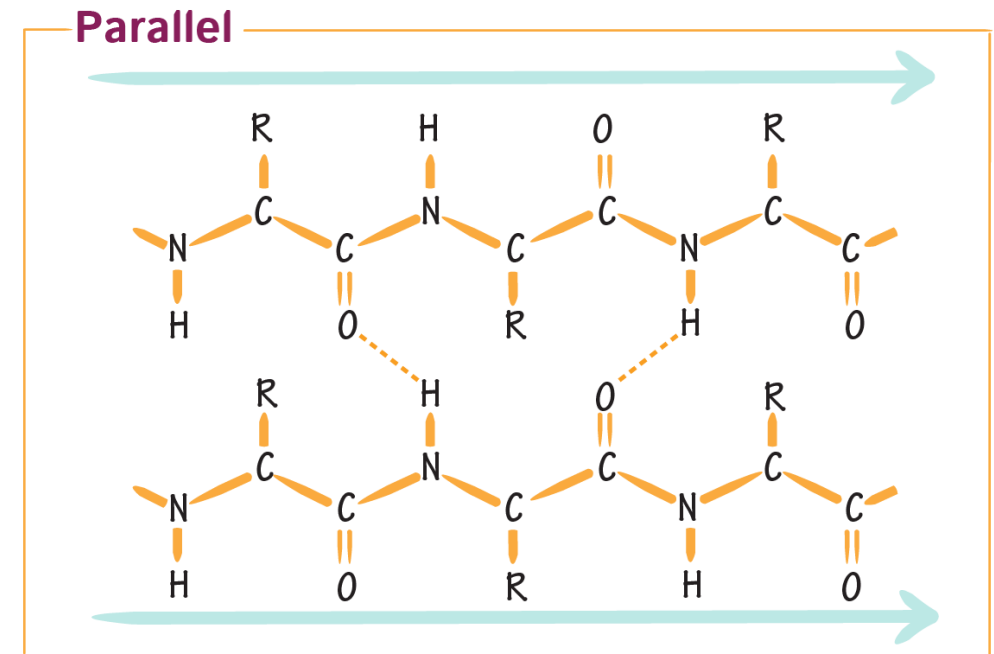
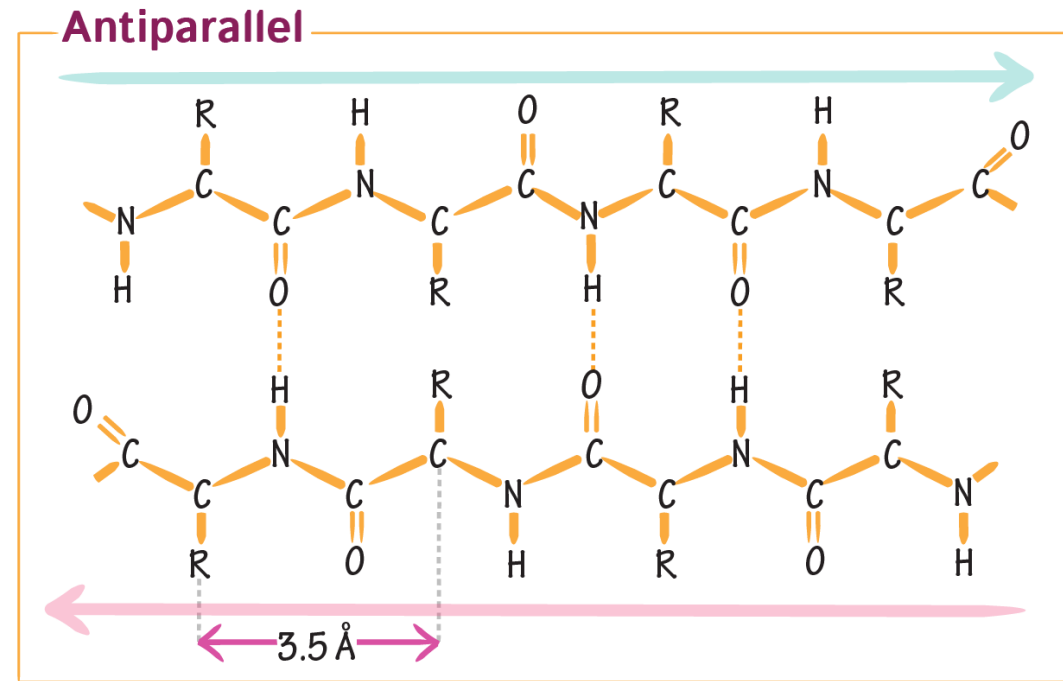
Protein Structure

Secondary

β -strand / β -pleated sheet

- Like folded paper strips placed side by side → each strip = a region of the peptide chain with a zigzag peptide backbone
- Parallel:** β -strands with the same orientation
- Antiparallel:** β -strands with opposite orientations

Directionality of Beta Strand

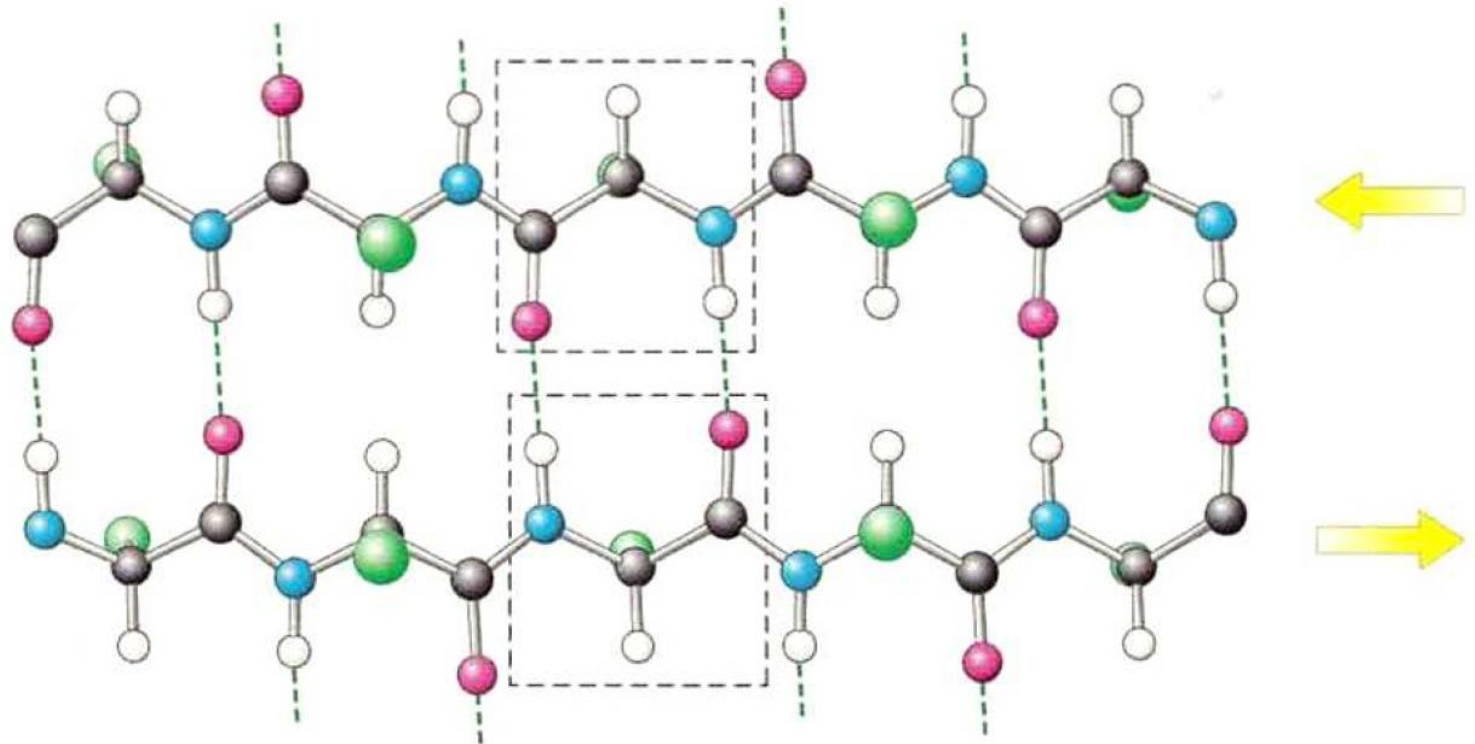


Protein Structure

Secondary

β -strand / β -pleated sheet

- Antiparallel β -sheets: hydrogen bonds between an amino acid in one strand and one in the adjacent strand (C=O–NH)

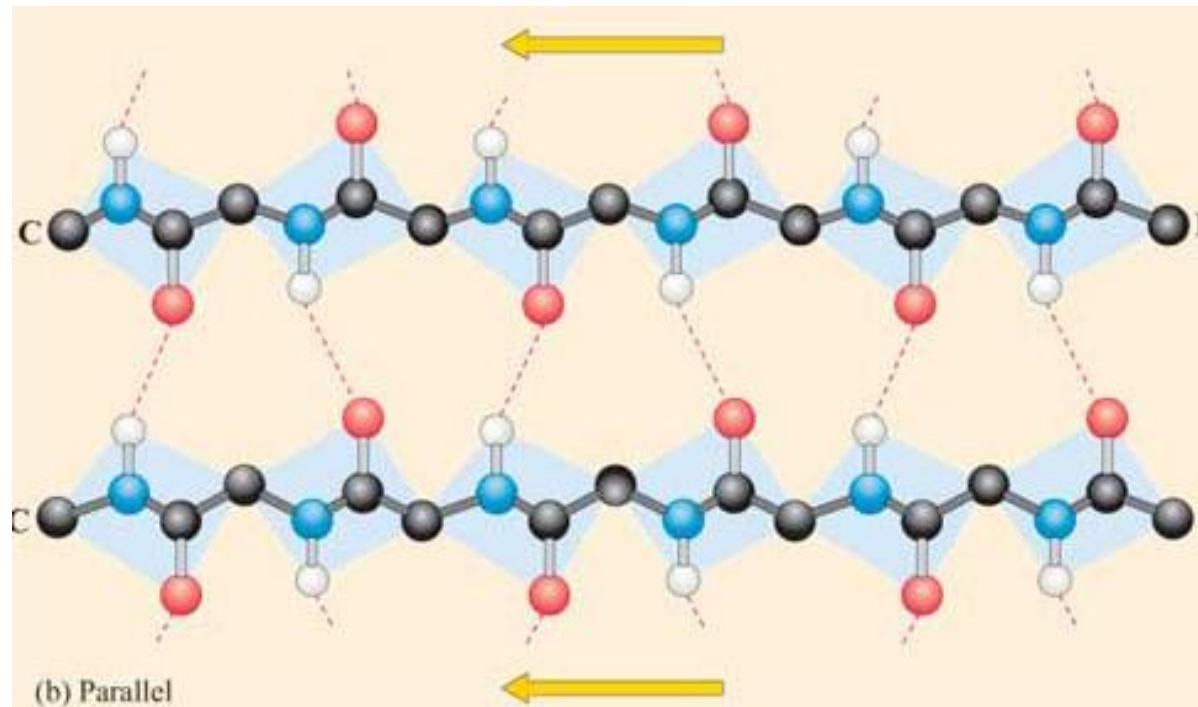


Protein Structure

Secondary

β -strand / β -pleated sheet

- Parallel β -sheets: hydrogen bonds between an amino acid in one β -strand and two amino acids (not necessarily consecutive) in the adjacent β -strand (C=O to NH).

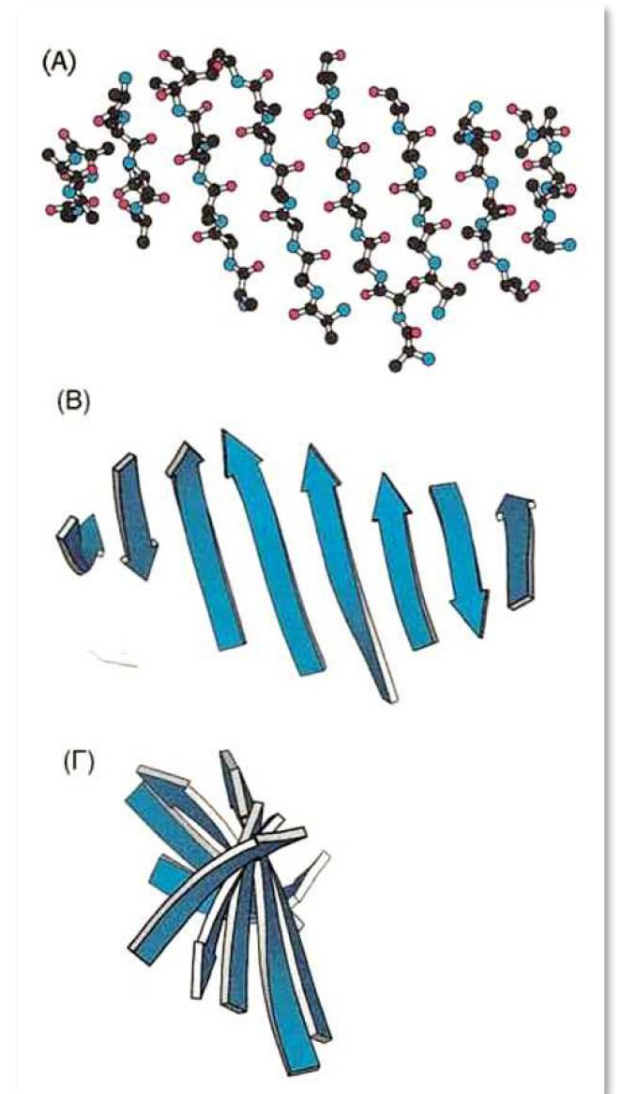
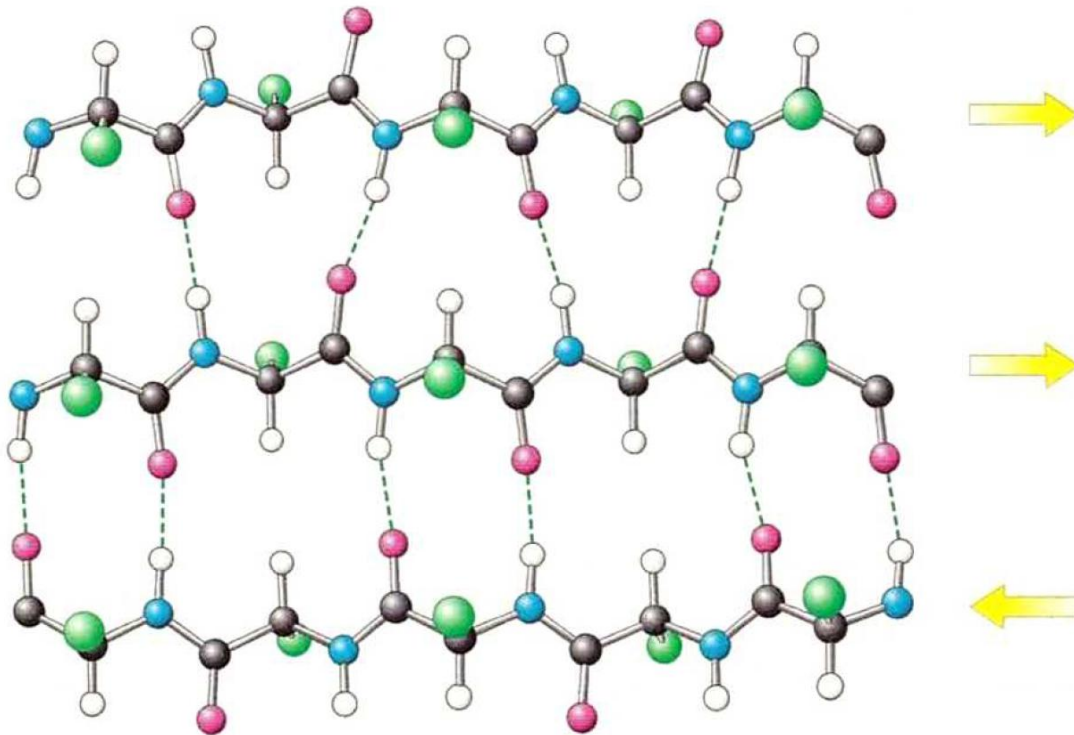


Protein Structure

Secondary

β -strand / β -pleated sheet

- There are also cases of **mixed structures** (parallel and antiparallel)



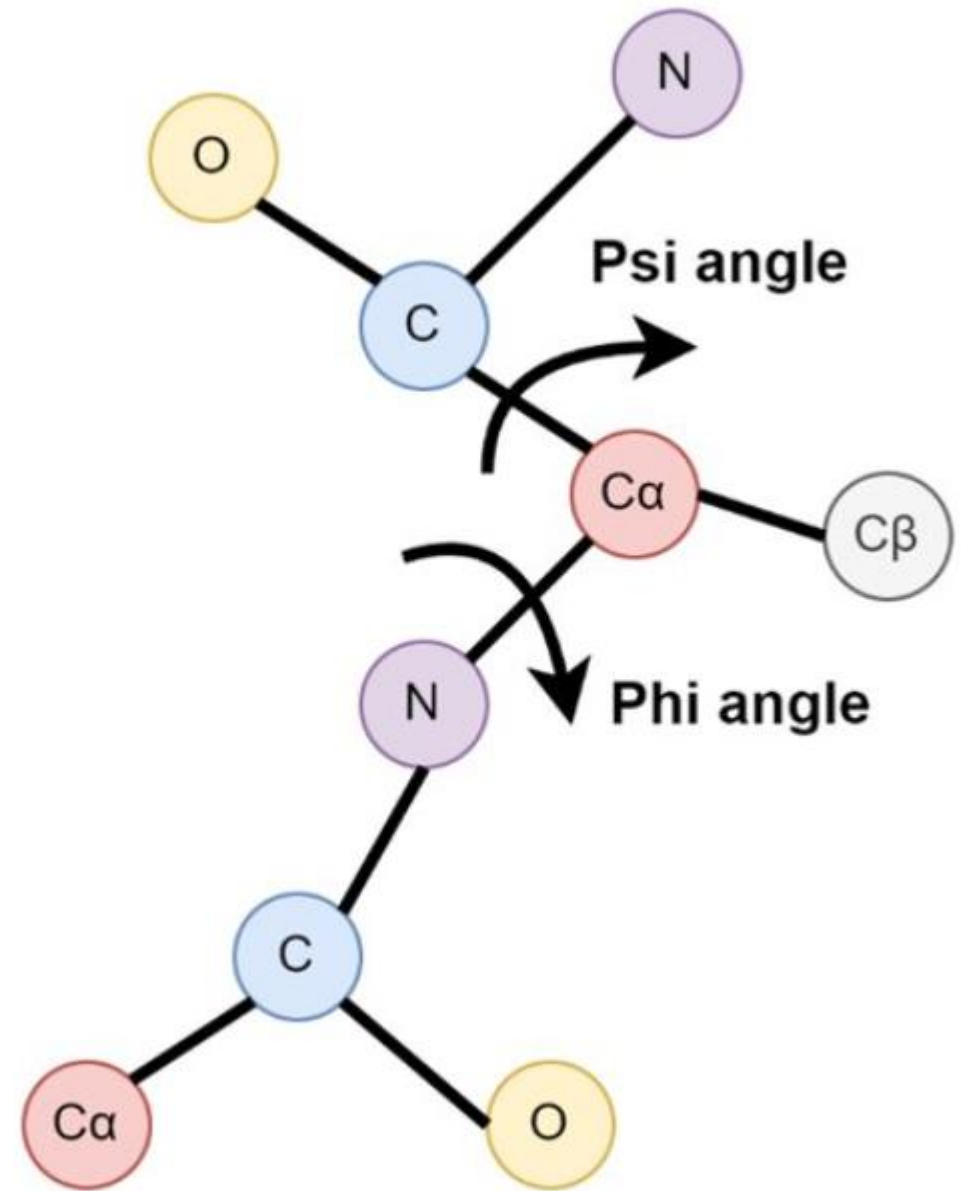
Protein Structure

Secondary

What determines if we will get an α -helix or a β -sheet?

→ The ϕ and ψ angles, which in turn are based on the side chains

- There are combinations of ϕ and ψ angles that are not permitted due to stereochemical constraints.



Secondary

Secondary Protein Structure

The diagram illustrates the two main types of secondary protein structure: the α -helix and the β -pleated sheet. The α -helix is shown as a tight coil, stabilized by hydrogen bonds between the carbonyl oxygen of one residue and the amide hydrogen of a residue four positions ahead. The β -pleated sheet is shown as a zigzag structure, stabilized by hydrogen bonds between adjacent strands. A specific amino acid sequence (Val, His, Thr) is shown at the end of the α -helix. The diagram also includes a detailed view of the hydrogen bonding pattern in the β -pleated sheet, showing the repeating sequence of amide and carbonyl groups.

Lactase subunit: yellow β -strands – fuchsia α -helices

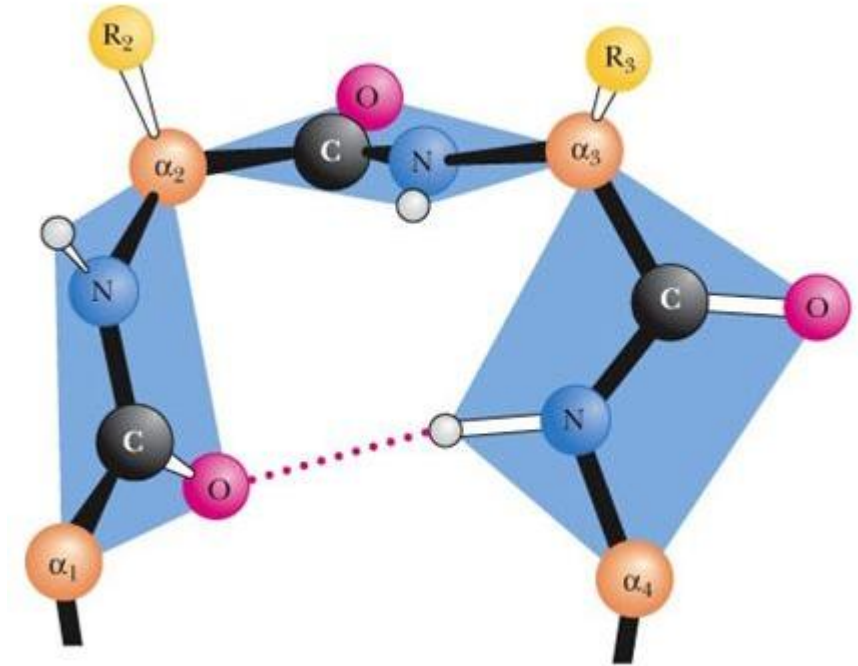
Protein Structure

Secondary

There are also secondary structure elements that are not characterized by a periodic pattern, such as α -helices and β -sheets:

▪ β -turns

- They join the ends of two adjacent segments of the antiparallel β -sheet surfaces.
- a **180° turn** comprising **4 amino acid residues**, with the carbonyl oxygen of the first amino acid forming a **hydrogen bond** with the hydrogen of the fourth amino acid



Protein Structure

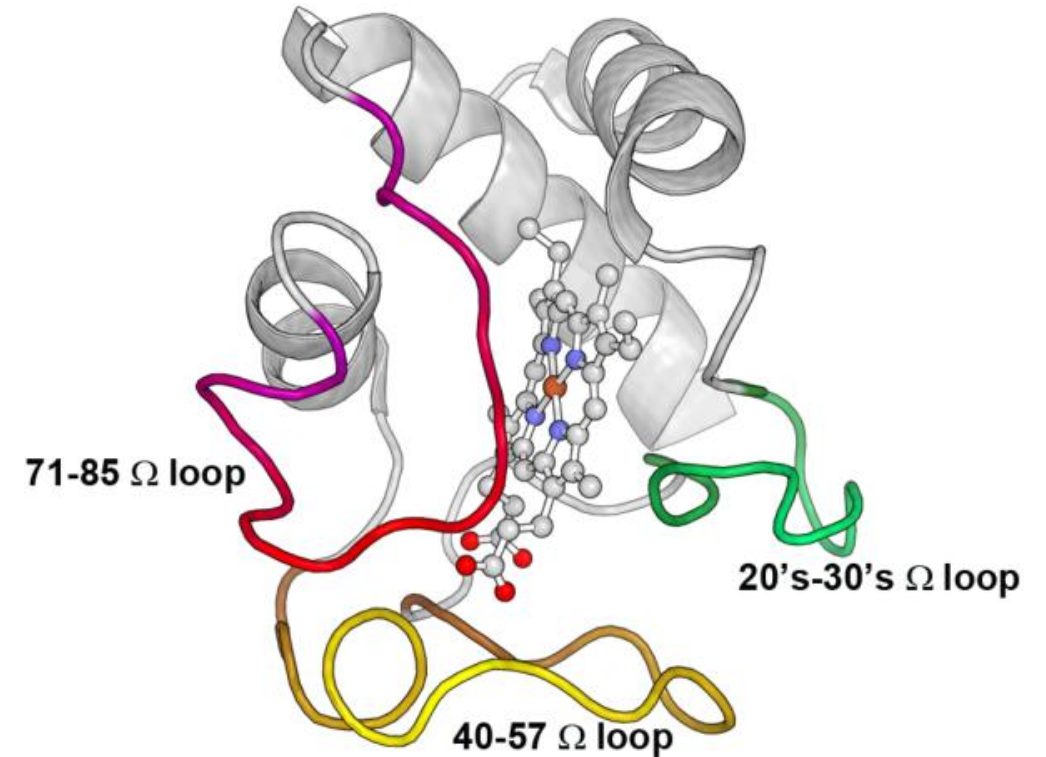
Secondary

There are also secondary structure elements that are not characterized by a periodic pattern, such as α -helices and β -sheets:

- **loop**

- **More complex and more flexible** than turns. They often have a fixed and strictly defined structure

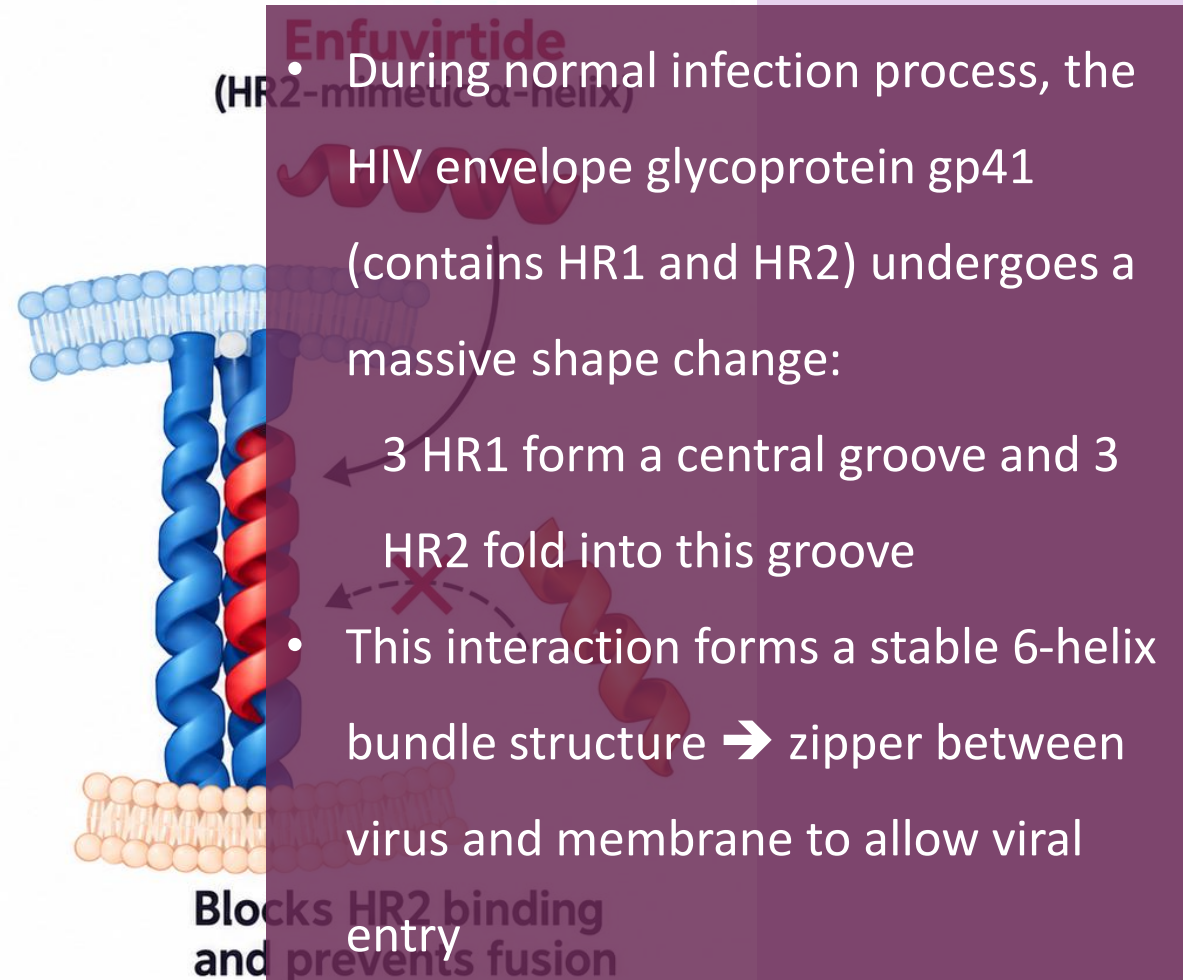
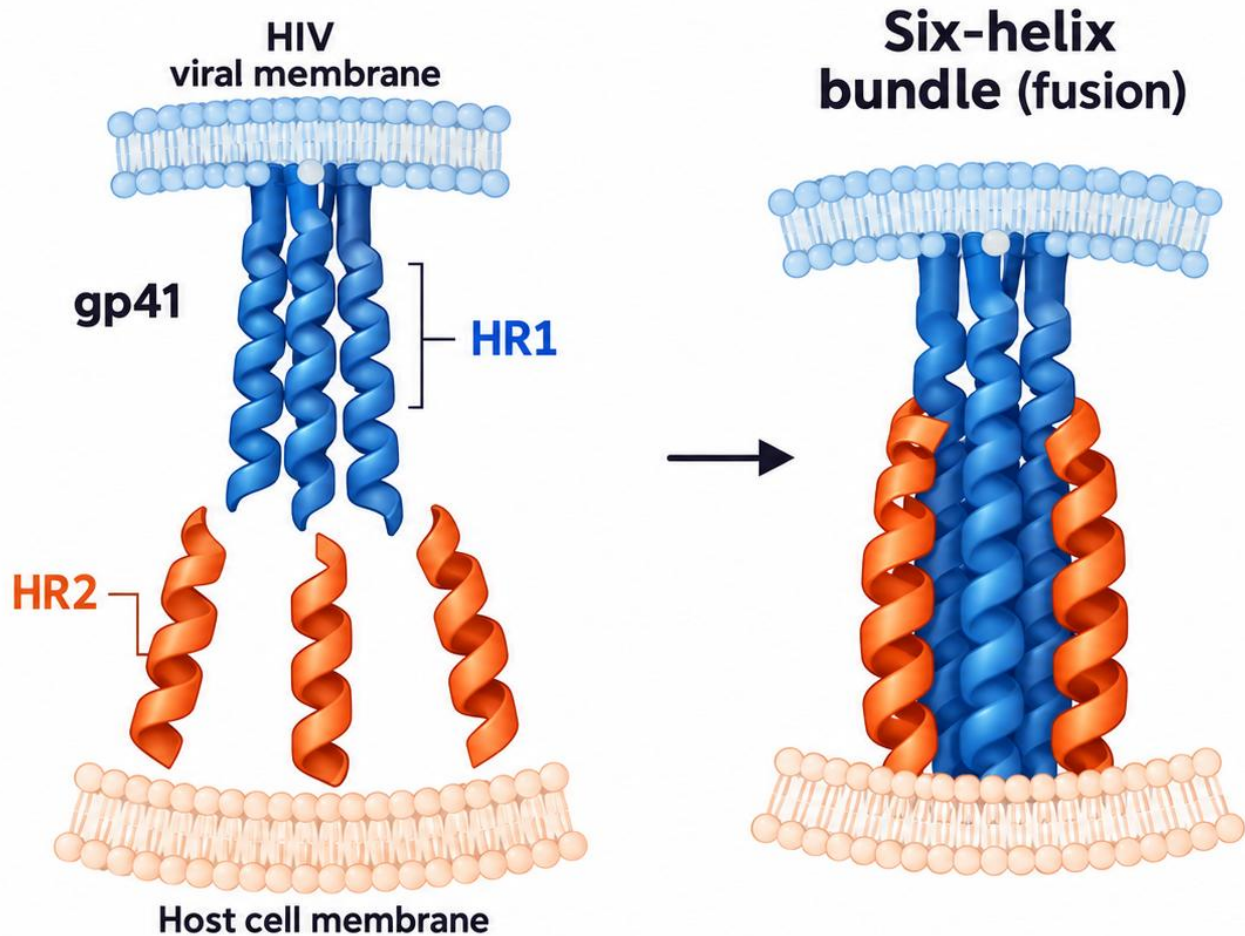
Turns and loops are usually found on the surface of proteins and interact with other molecules



Protein Structure

Secondary

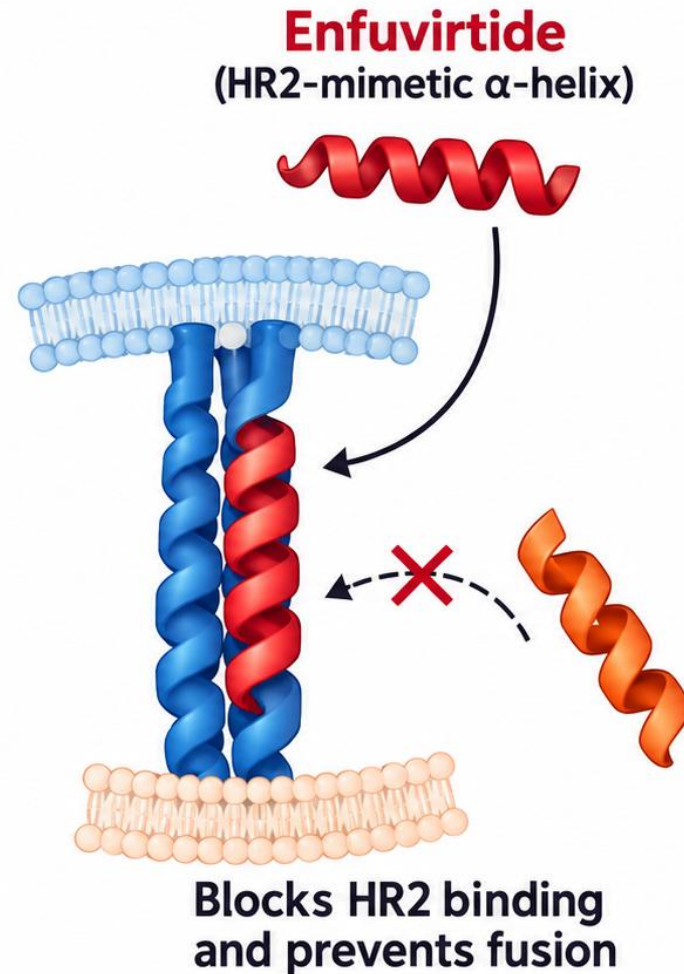
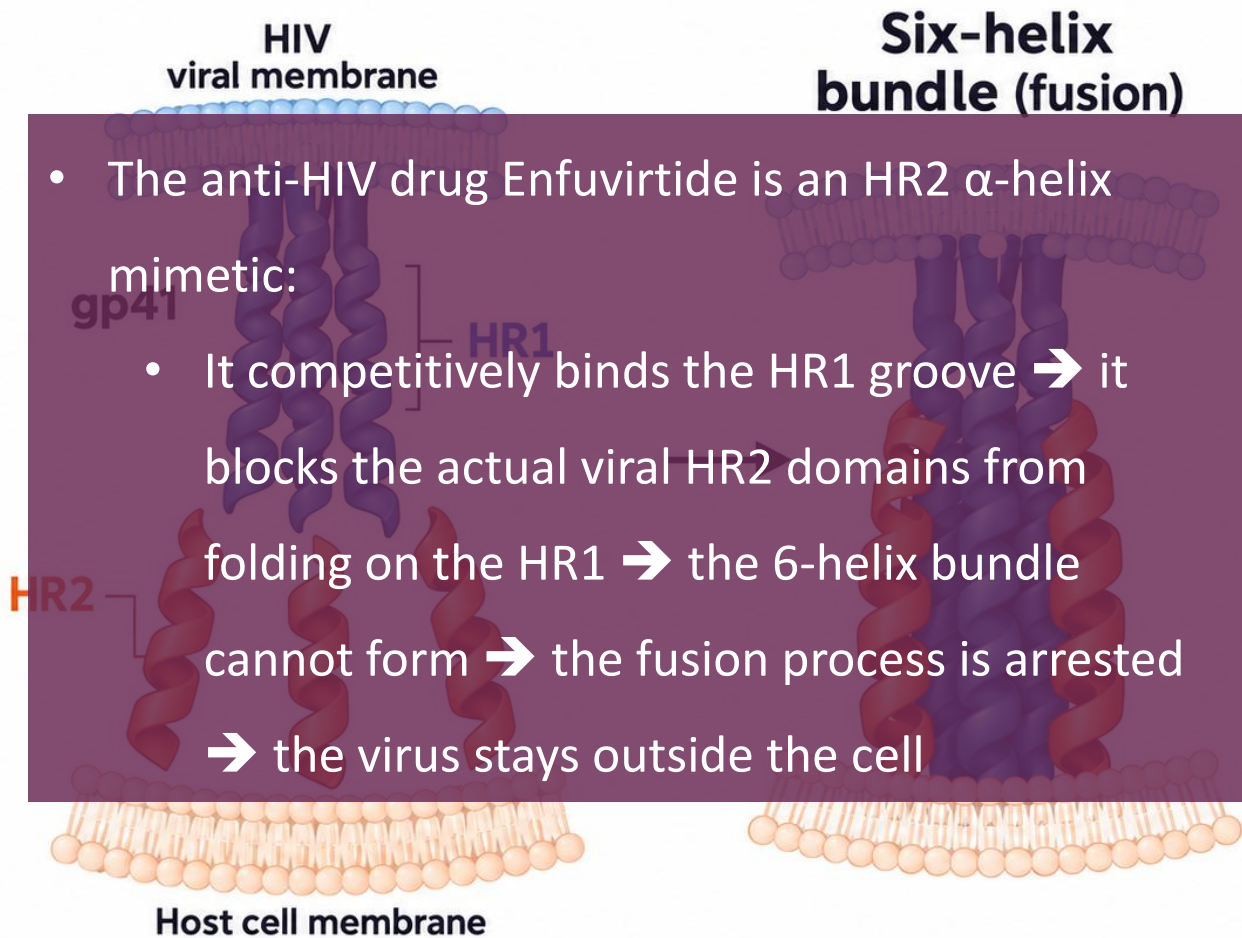
The example of anti-HIV drug Enfuvirtide



Protein Structure

Secondary

The example of anti-HIV drug Enfuvirtide



Protein Structure

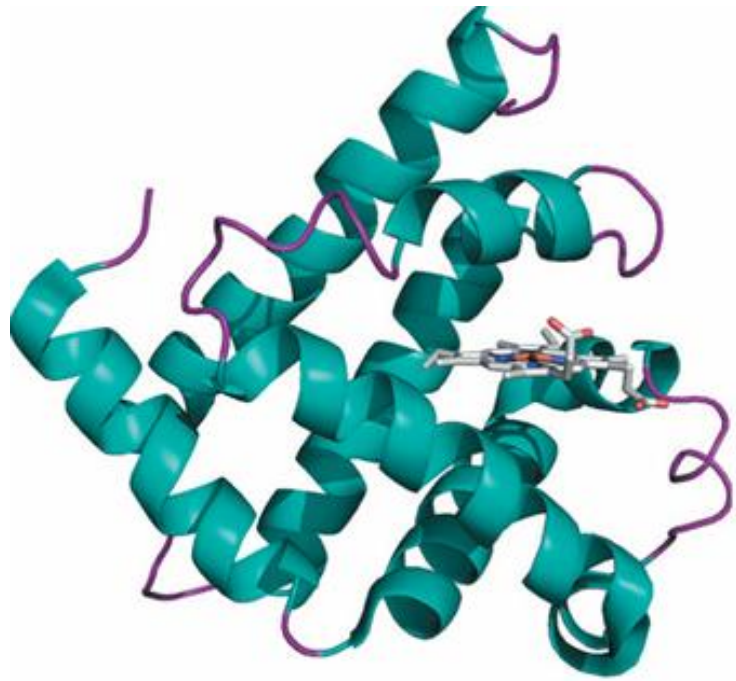
Tertiary

- The complete **3D spatial conformation** of a single polypeptide chain, bringing amino acid residues far apart in linear sequence into close proximity
- If a protein is comprised only by one polypeptide chain → tertiary structure = the final form of the protein
- If the protein is comprised by more than one polypeptide chains → tertiary structure = **subunit** of the protein

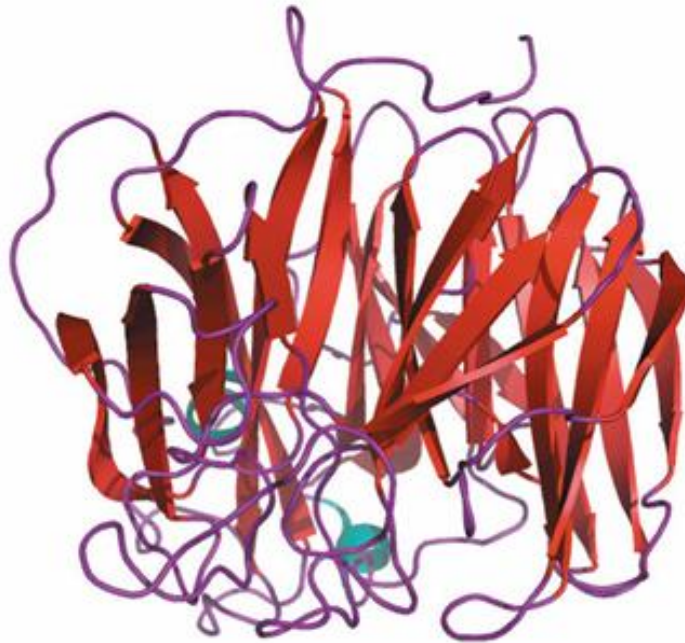
Protein Structure

Tertiary

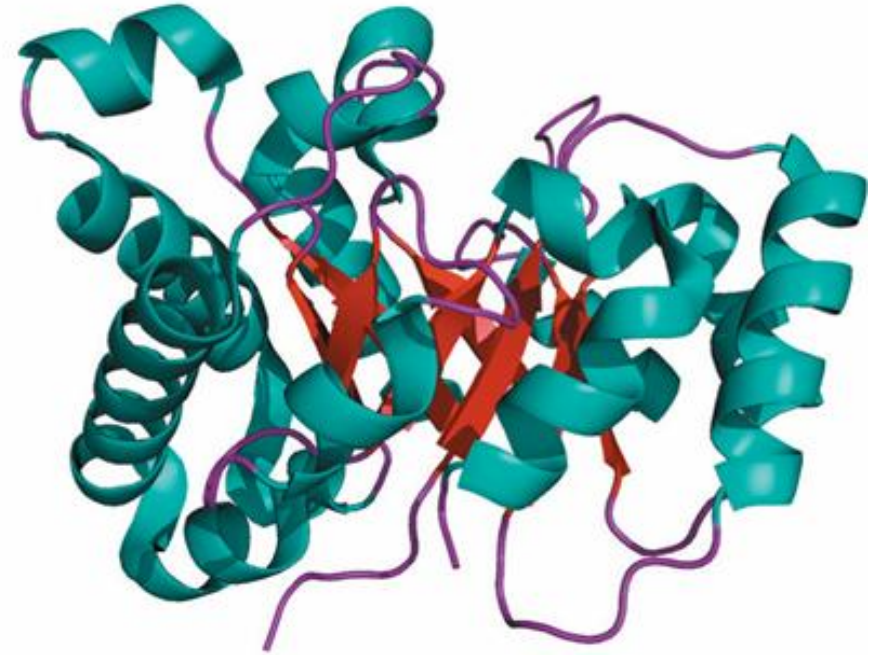
- Examples of diversity in the tertiary structures of proteins (α -helix, β -sheets, turns/loops)



Myoglobin



Neuraminidase



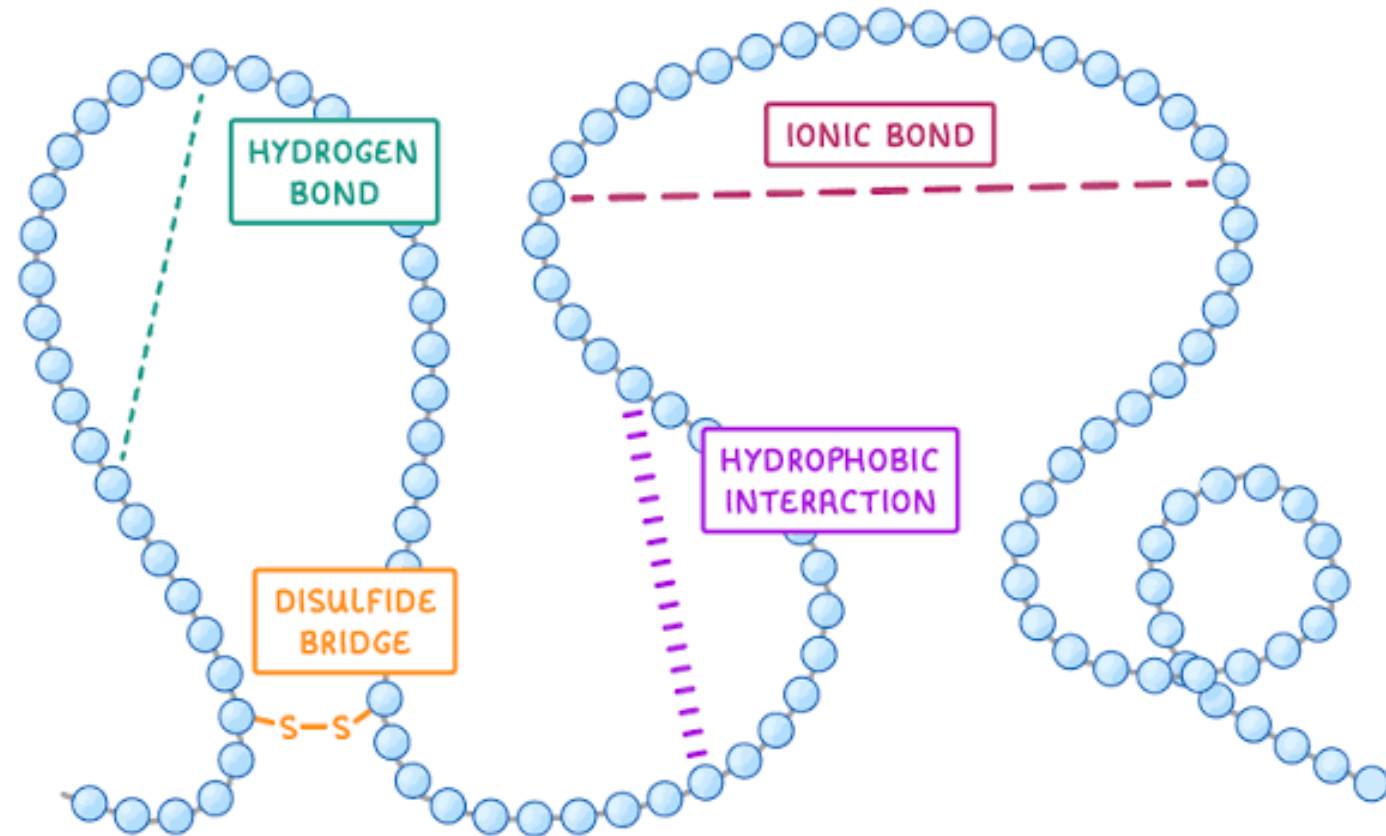
Triosephosphate isomerase

Protein Structure

Tertiary

It's based on **the interaction between side chains** of amino acids, which may be located far apart in the sequence.

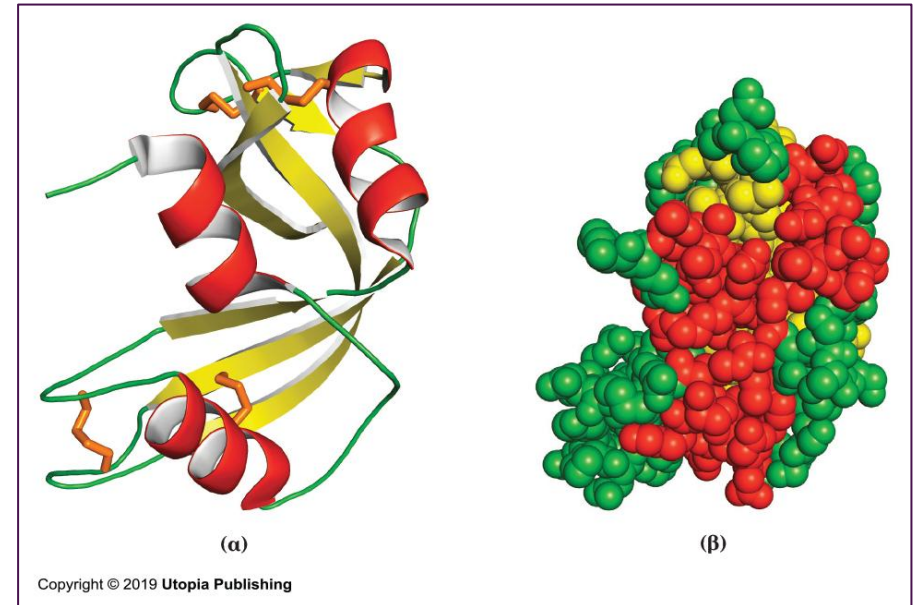
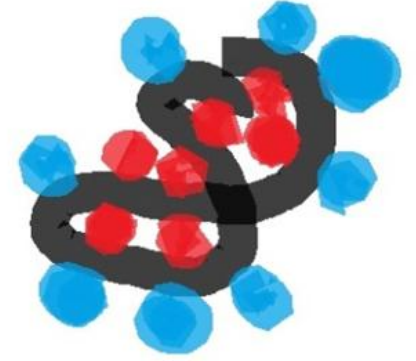
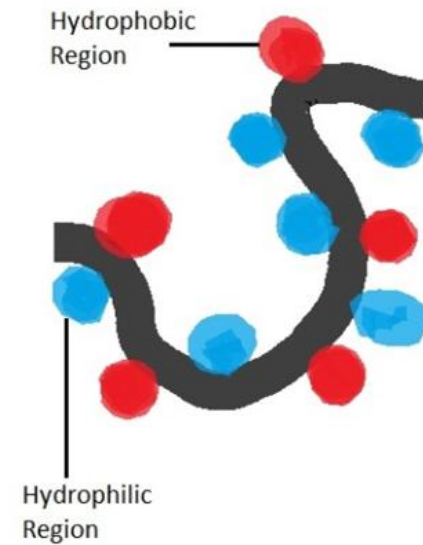
- Disulfide bonds
- van der Waals forces
- Hydrogen bonds
- Ionic bonds / Salt bridges



Protein Structure

Tertiary

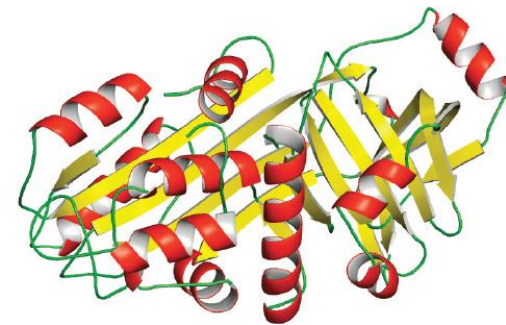
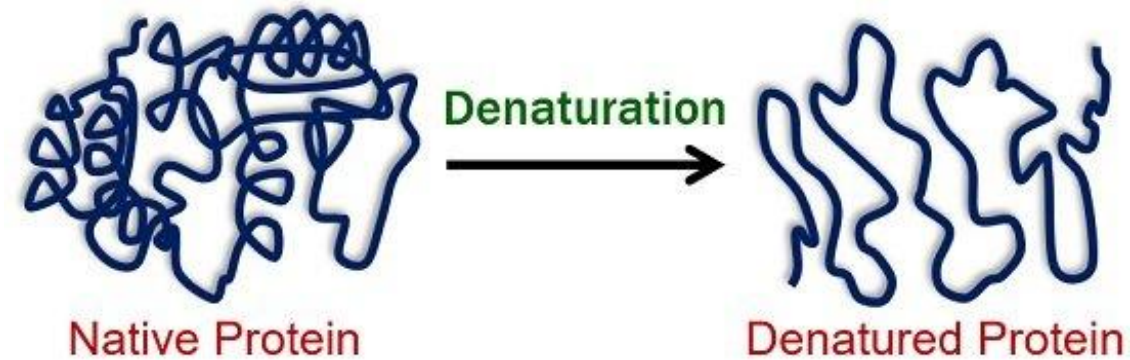
- Secondary structures are “packed” through weak interactions → **minimum free energy**
- e.g., globular proteins: α -helices and β -sheets in the core, and bends and turns (which connect these structures to one another) on the surface



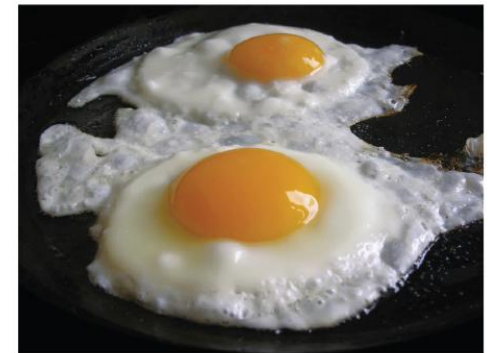
Protein Structure

Tertiary

- The environment affects the tertiary structure, as changes in **temperature**, **pH**, or the **presence of organic molecules** such as alcohols and urea cause changes in the structure and conformation of the protein.
- This process is called **denaturation**



Ovalbumin subunit

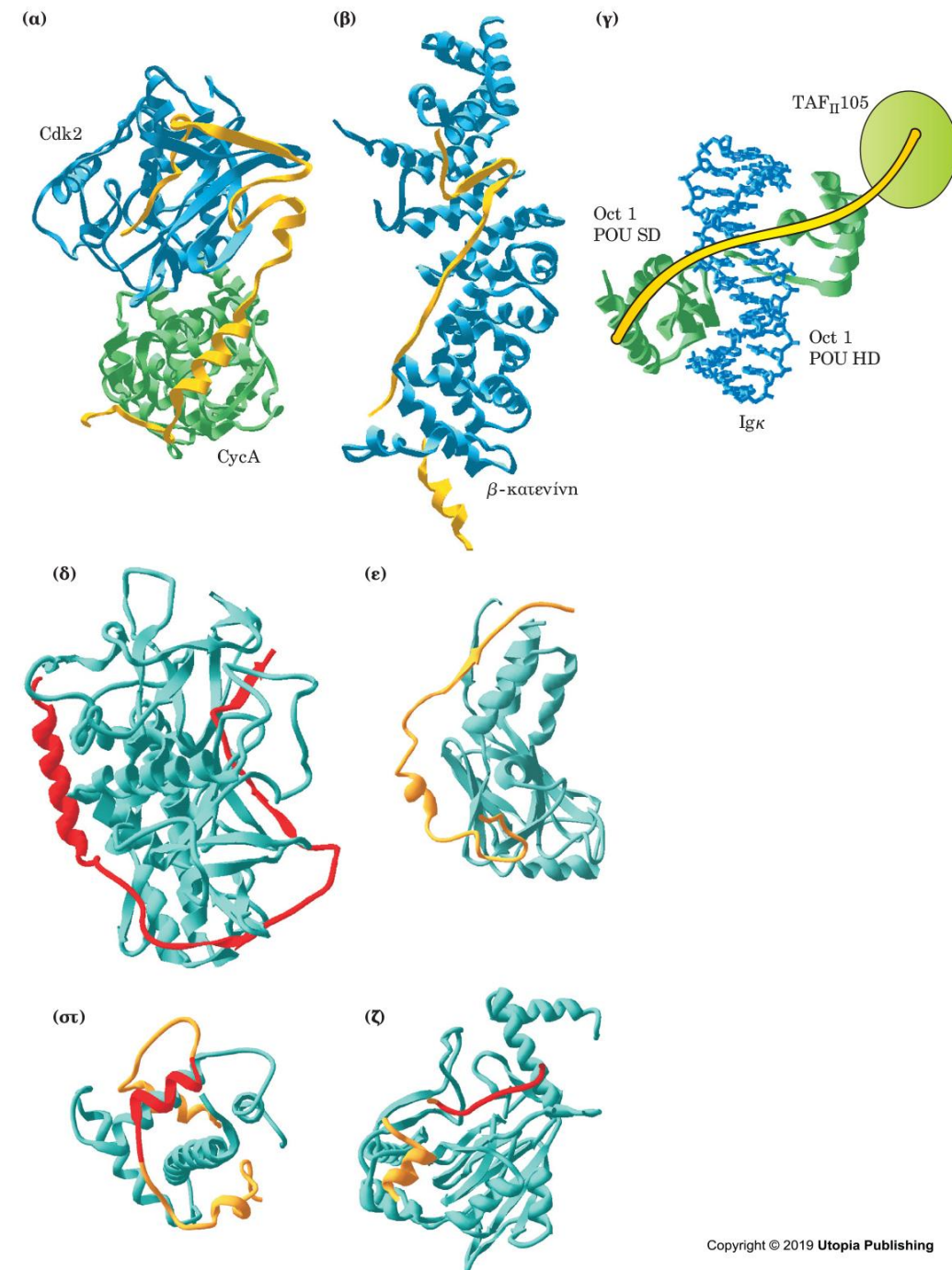


Protein Structure

Tertiary

Certain proteins are functional in a partially unfolded state = **intrinsically disordered proteins**

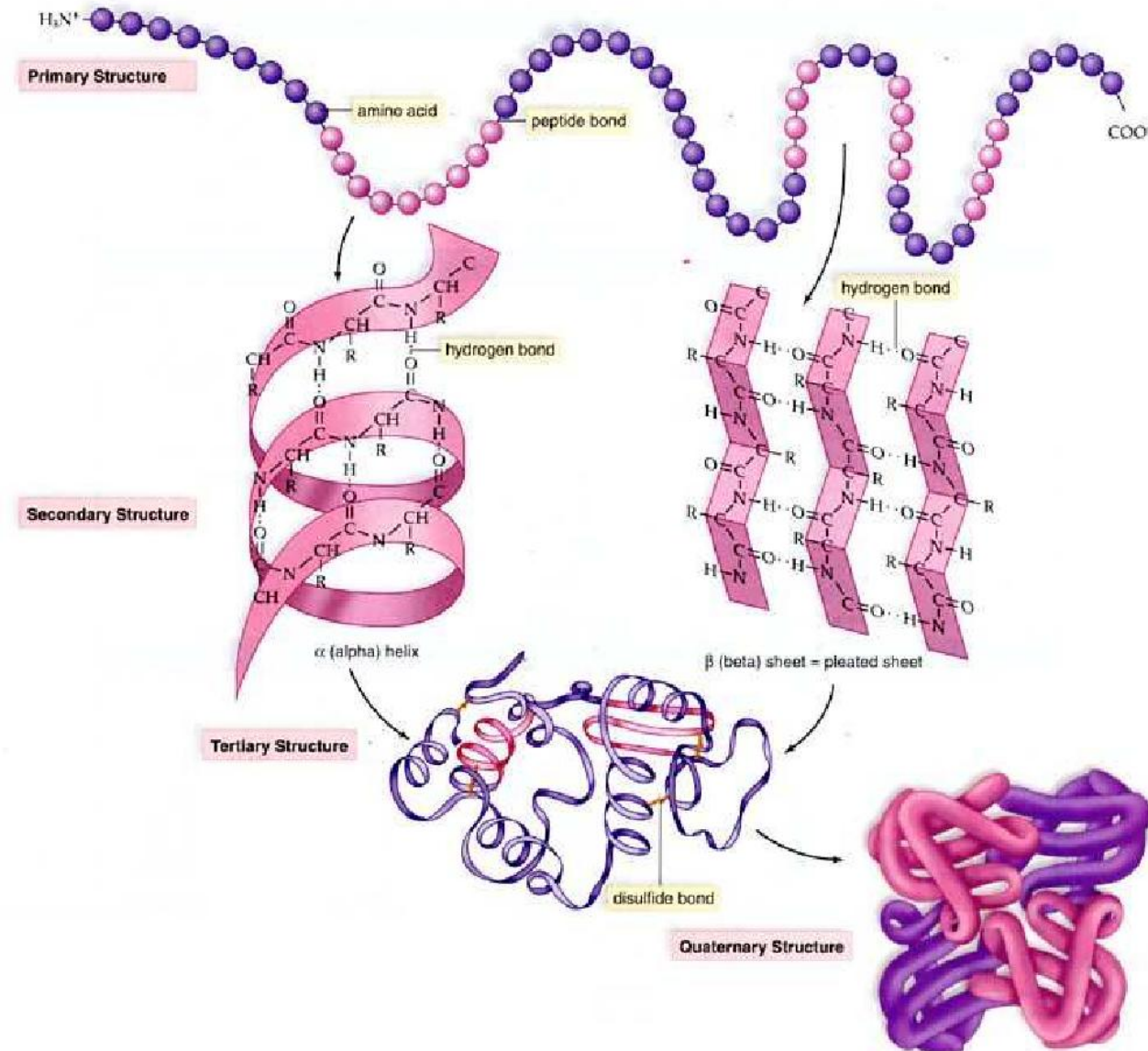
- structures of indeterminate conformation → rich in polar and charged amino acids with few hydrophobic residues
- high intramolecular elasticity
- They can interact with a large portion of the surface of their target proteins



Protein Structure

Quaternary

- The association of polypeptide chains to form structures with **multiple subunits**
- The **forces** that maintain the quaternary structure are the **same** ones that contribute to the folding of the **tertiary** structure (hydrogen bonds, hydrophobic interactions, and electrostatic bonds)

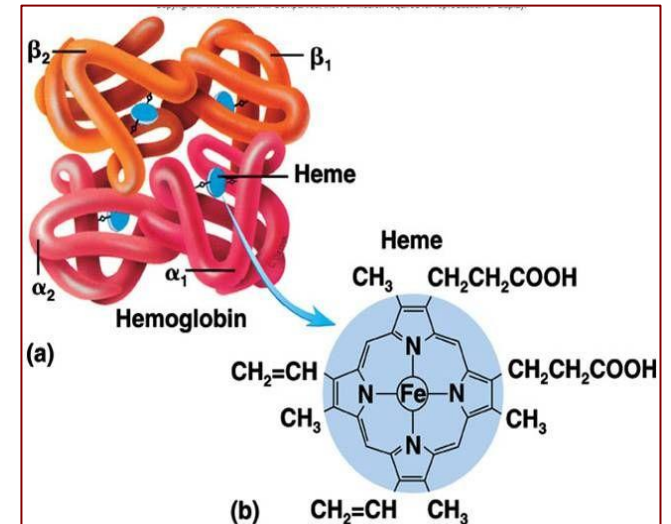


Protein Structure

Quaternary

There are two types of organization of the quaternary structure of proteins:

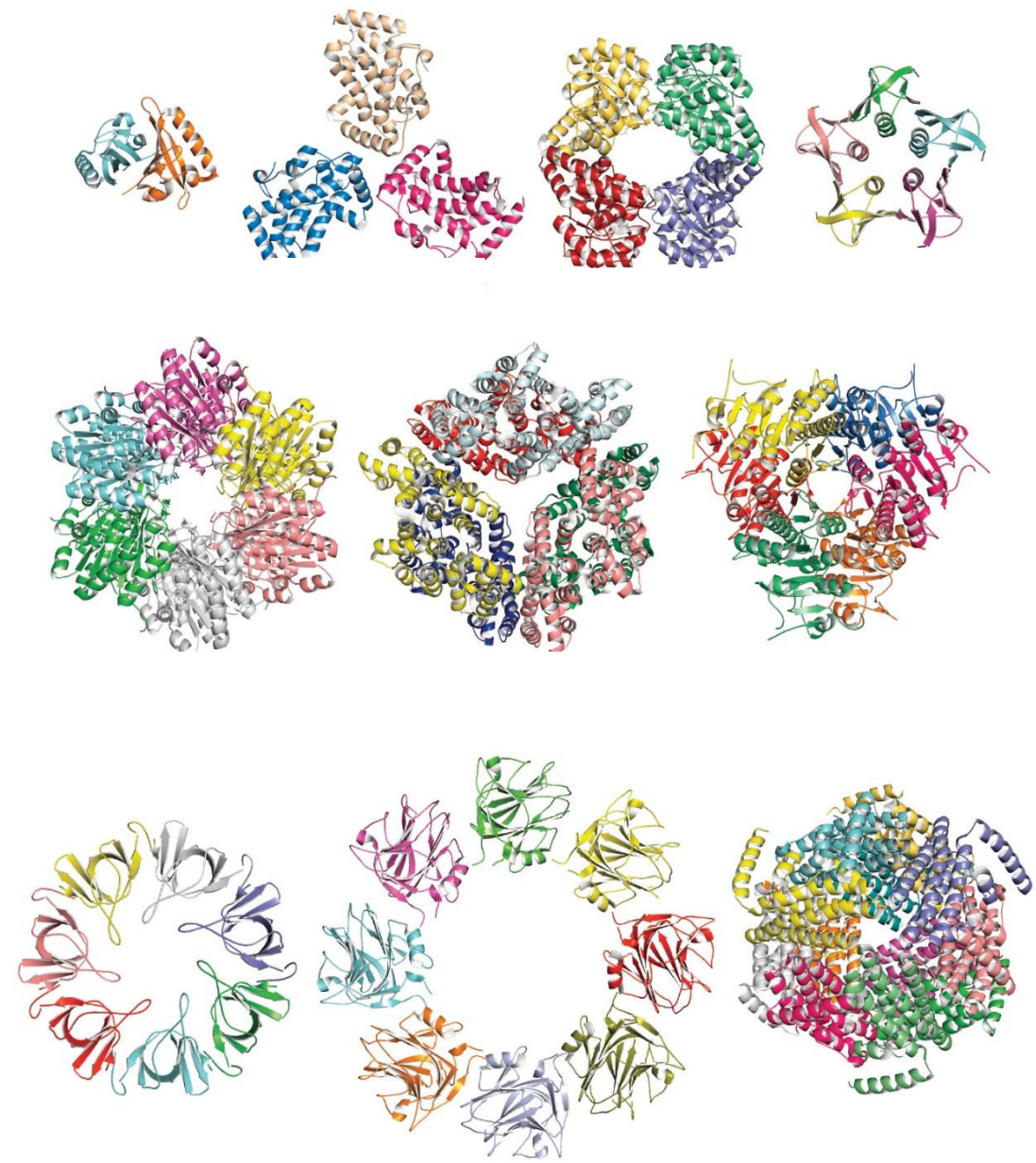
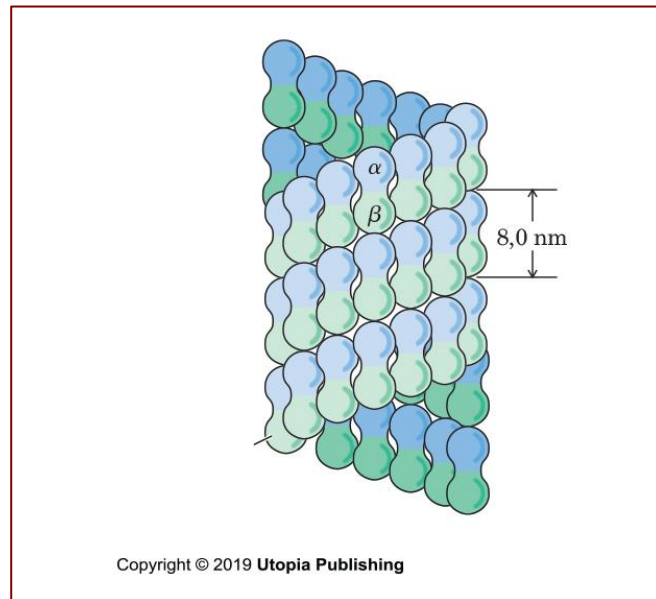
- **Homotypic:** the association between identical or nearly identical subunits
- **Heterotypic:** the association between subunits with different structures



Protein Structure

Quaternary

- Many proteins have a symmetrical, closed structure
- Others have open structures → “continuous” polymerization



Protein Structure

Quaternary

Why choose a quaternary structure?

1. Synergy: regulatory subunits → regulation of other parts of the protein

- e.g., hemoglobin (4 subunits) vs myoglobin (1 subunit):

Hemoglobin loads O_2 almost fully in the lungs yet releases a large fraction in the tissues, across a fairly small drop in pO_2 . Myoglobin cannot do this; it holds on to O_2 until pO_2 is very low. That is why **myoglobin is a store** and **hemoglobin is a transporter**.

2. Regulation: Interfaces between subunits create new binding sites. Regulators bind there and shift the whole complex between active and inactive states.

- e.g., 2,3-BPG and hemoglobin

Protein Structure

Quaternary

Why choose a quaternary structure?

- 3. Mixing and Matching:** Different subunits can combine in different proportions, producing tissue-specific versions of the same protein from a small set of genes
 - e.g., lactate dehydrogenase is a tetramer of H (heart) and M (muscle) subunits in five combinations.
- 4. Genetic economy:** less DNA is required to encode a monomer that forms a homopolymer than to encode a large polypeptide of the same molecular mass

Protein Structure

Quaternary

Why choose a quaternary structure?

5. Signaling across the membrane:

Many receptors are activated by bringing two subunits together

- e.g., HER2 in breast cancer signals by dimerizing with other members of its family

6. Efficiency: In multi-enzyme complexes, intermediates are handed from one active site to the next without diffusing away

