Question 1: "1. (a) Justify the following statements (any five):"

- **(i) Glycine and Proline are often present at turns of polypeptide chain.**
    - **Glycine:** Is the smallest amino acid, lacking a side chain (its R-group is just a hydrogen atom). This minimal steric hindrance provides exceptional flexibility to the polypeptide backbone, allowing it to easily adopt sharp turns and loops that are often critical for protein folding and function.
    - **Proline:** Is unique because its side chain forms a covalent bond with its own amino group, creating a rigid ring structure. This rigidity introduces a "kink" or "bend" in the polypeptide chain and restricts the rotation around the N-Cα bond. This structural constraint makes proline ideally suited for initiating or breaking helices and is frequently found at the turns and bends of protein structures.
- **(ii) Negatively charged and positively charged amino acids are often present near the amino and Carboxy terminal end of the helical segment respectively.**
    - **Helix Dipole:** An α-helix has a net dipole moment due to the alignment of the individual dipoles of the peptide bonds along the helix axis. The N-terminus of the helix has a partial positive charge, and the C-terminus has a partial negative charge.
    - **Charge Stabilization:** Negatively charged amino acids (like Aspartate or Glutamate) are attracted to and can stabilize the partial positive charge at the amino (N-) terminal end of the helix. Conversely, positively charged amino acids (like Lysine, Arginine, or Histidine) are attracted to and can stabilize the partial negative charge at the carboxy (C-) terminal end of the helix. This electrostatic interaction helps to stabilize the overall helical structure.

- **(iii) Iron should be present in ferrous form in heme part of the hemoglobin.**
    - **Oxygen Binding:** The ferrous (Fe^{2+}) state of iron in the heme group of hemoglobin is crucial because it is the only oxidation state that can reversibly bind oxygen.
    - **Oxidation State:** When Fe^{2+} binds oxygen, it undergoes a change in its electronic configuration but does not become oxidized to Fe^{3+} (ferric state). If the iron were to be oxidized to the Fe^{3+} (ferric) state, as in methemoglobin, it would lose its ability to bind oxygen, rendering the hemoglobin molecule non-functional for oxygen transport. This is why the ferrous state is essential for physiological oxygen transport.
- **(iv) Rotation around peptide bond is restricted.**
    - **Partial Double Bond Character:** The peptide bond, which links amino acids in a polypeptide chain, exhibits partial double-bond character. This is due to the resonance between the carbonyl carbon and the nitrogen atom, where electrons are delocalized across the C-N-C-O atoms.
    - **Planar Structure:** This partial double-bond character makes the peptide bond rigid and planar. All four atoms involved in the peptide bond (the carbonyl carbon, the carbonyl oxygen, the amide nitrogen, and the amide hydrogen) lie in the same plane.
    - **Restricted Rotation:** The presence of this partial double bond significantly restricts rotation around the C-N bond within the peptide backbone. This rigidity is a critical factor in limiting the conformational possibilities of a polypeptide chain, allowing proteins to fold into specific, stable three-dimensional structures. Rotation is primarily permitted around the Cα-N bond (ϕ) and the Cα-C bond (ψ).
- **(v) Fetal Hb has high affinity for oxygen binding as compared to adult.**

- **Oxygen Transfer:** Fetal hemoglobin (HbF) has a higher affinity for oxygen compared to adult hemoglobin (HbA). This difference is biologically essential for the efficient transfer of oxygen from the mother's blood (via the placenta) to the fetal blood.
- **2,3-Bisphosphoglycerate (BPG) Binding:** The primary reason for this higher affinity lies in the weaker binding of 2,3-bisphosphoglycerate (BPG) to HbF. BPG is an allosteric effector that binds to a central cavity in the deoxygenated (T-state) form of hemoglobin, stabilizing this state and reducing oxygen affinity.
- **Structural Difference:** HbF is composed of two α and two γ subunits ($\alpha_2\gamma_2$), whereas HbA is $\alpha_2\beta_2$. The γ subunits in HbF have a serine residue at position 143 instead of a histidine (which is present in the β subunits of HbA). This single amino acid difference reduces the number of positive charges in the BPG binding pocket, leading to weaker electrostatic interactions with the negatively charged BPG molecule. Consequently, BPG binds less effectively to HbF, allowing it to maintain a higher affinity for oxygen even in the presence of BPG, thus facilitating oxygen uptake from maternal blood.

- **(vi) Collagen forms left handed alpha-helix secondary structure.**
  - **Collagen Helix, not α-helix:** This statement is partially inaccurate in its terminology. Collagen does *not* form a typical α-helix secondary structure. Instead, individual collagen polypeptide chains form a unique type of helix called a **left-handed helix** (sometimes referred to as a polyproline II-like helix).
  - **Triple Helix:** Three of these individual left-handed collagen helices then intertwine to form a characteristic **right-handed triple helix**, which is the primary structural unit of collagen fibers.

- **Amino Acid Composition:** This specific helical structure is dictated by collagen's unique amino acid composition, particularly its high content of Glycine, Proline, and Hydroxyproline, which facilitate the formation of this loose, left-handed helix in individual strands, and the subsequent tight winding into the right-handed triple helix. So, while an individual chain forms a left-handed helix, it's not the "alpha-helix" found in other proteins like keratin or hemoglobin.

(b) Define the following proteins with example.

- **(i) Conjugated proteins:**
  - **Definition:** Proteins that are covalently or non-covalently associated with a non-protein component (called a prosthetic group) that is essential for their function. Without the prosthetic group, the protein (called the apoprotein) is typically inactive.
  - **Example:**
    - **Hemoglobin:** Contains a heme group (a porphyrin ring with an iron atom) as its prosthetic group, essential for oxygen binding.
    - **Glycoproteins:** Proteins with covalently attached carbohydrate chains (e.g., antibodies, cell surface receptors).
    - **Lipoproteins:** Proteins associated with lipids (e.g., LDL, HDL for lipid transport in blood).
    - **Metalloproteins:** Proteins containing metal ions as prosthetic groups (e.g., ferroproteins like ferritin, flavoproteins like succinate dehydrogenase containing FAD).
- **(ii) Multimeric proteins:**

- **Definition:** Proteins composed of two or more polypeptide chains (subunits) that associate non-covalently to form a single functional protein complex. Each individual polypeptide chain is called a monomer or subunit.
- **Example:**
  - **Hemoglobin:** A heterotetramer composed of four subunits: two α chains and two β chains.
  - **Immunoglobulins (Antibodies):** Typically composed of four polypeptide chains (two heavy chains and two light chains).
  - **Lactate Dehydrogenase:** A tetrameric enzyme composed of four identical or non-identical subunits.

- **(iii) Fibrous proteins:**
  - **Definition:** Proteins characterized by elongated, insoluble structures, often composed of repeating amino acid sequences that form long, rod-like or filamentous shapes. They primarily serve structural, protective, or contractile roles in cells and tissues.
  - **Example:**
    - **Collagen:** The most abundant protein in mammals, forming strong, insoluble fibers in connective tissues like tendons, ligaments, bone, and skin. It has a characteristic triple helical structure.
    - **Keratin (e.g., α-Keratin):** Found in hair, nails, skin, and wool. It forms tough, insoluble fibers and provides structural integrity.
    - **Elastin:** Provides elasticity to tissues like skin, blood vessels, and lungs.

    - **Fibrin:** Involved in blood clotting, forming an insoluble fibrous network.

- **(iv) Membrane proteins:**
    - **Definition:** Proteins that are associated with biological membranes (e.g., cell membranes, organelle membranes). They play crucial roles in transport, signaling, cell adhesion, and enzymatic activities across membranes.
    - **Example:**
        - **Integral Membrane Proteins (Transmembrane Proteins):** Span the entire lipid bilayer, having hydrophobic regions embedded within the membrane and hydrophilic regions exposed to the aqueous environments on both sides (e.g., **Ion Channels**, **G-protein coupled receptors**, **Glucose Transporters**).
        - **Peripheral Membrane Proteins:** Associate loosely with the membrane surface, often by interacting with integral proteins or the polar heads of lipids (e.g., **Spectrin** on the inner surface of red blood cell membranes).
        - **Lipid-Anchored Proteins:** Covalently attached to a lipid molecule that is inserted into the membrane.
- **(v) Globular proteins:**
    - **Definition:** Proteins that adopt a compact, roughly spherical or globular shape when folded. They are typically soluble in aqueous solutions and perform a wide variety of dynamic functions such as catalysis, transport, regulation, and defense.
    - **Example:**
        - **Enzymes:** Almost all enzymes are globular proteins (e.g., **Lysozyme**, **Hexokinase**, **DNA Polymerase**).

- **Hemoglobin:** Though multimeric, its individual subunits and the overall complex adopt a globular shape.
- **Myoglobin:** An oxygen-binding protein in muscle, it is a single polypeptide chain folded into a compact globular structure.
- **Antibodies (Immunoglobulins):** Involved in immune defense.
- **Hormones:** Many protein hormones (e.g., **Insulin**).

Question 2: "2. (a) Differentiate between the following :"

- **(i) Parallel and Antiparallel β-pleated sheets.**
  - **Parallel β-pleated sheets:**
    - **Orientation:** Adjacent polypeptide strands run in the **same direction** (i.e., N-terminus to C-terminus).
    - **Hydrogen Bonding:** The hydrogen bonds between the strands are **diagonal**, connecting each amino acid in one strand to two different amino acids in the adjacent strand. This results in slightly weaker and less stable hydrogen bonds compared to antiparallel sheets.
    - **Structure:** Often found buried within the core of globular proteins and usually require more than two strands to form a stable structure.
    - **Diagrammatic Representation:** N -> C // N -> C (strands running parallel)
  - **Antiparallel β-pleated sheets:**
    - **Orientation:** Adjacent polypeptide strands run in **opposite directions** (i.e., N-terminus to C-terminus, then C-terminus to N-terminus).

- **Hydrogen Bonding:** The hydrogen bonds between the strands are **straight and perpendicular** to the strands. Each amino acid in one strand forms hydrogen bonds with a single amino acid in the adjacent strand. This results in stronger and more stable hydrogen bonds.
- **Structure:** Can be formed by as few as two strands and are generally more stable and commonly found than parallel sheets.
- **Diagrammatic Representation:** N -> C // C <- N (strands running antiparallel)

- **(ii) Protein sequence and structure databases.**
  - **Protein Sequence Databases:**
    - **Content:** Store the primary amino acid sequences of proteins, usually derived directly from DNA or mRNA sequences (translated from nucleotide sequences). They typically contain information about gene names, species origin, function predictions, and sometimes post-translational modifications.
    - **Purpose:** Primarily used for sequence similarity searches (e.g., BLAST) to infer function or evolutionary relationships, identifying motifs, and predicting domains. They are fundamental for genomics and proteomics.
    - **Examples:** UniProt (Universal Protein Resource), NCBI Protein Database, TrEMBL.
  - **Protein Structure Databases:**
    - **Content:** Store experimentally determined three-dimensional (3D) atomic coordinates of proteins (and nucleic acids). The structures are typically solved using techniques like X-ray crystallography, NMR spectroscopy, or cryo-electron microscopy. They provide detailed

information about secondary structures, domains, ligand binding sites, and overall protein architecture.

- **Purpose:** Used to understand protein function based on shape, active site analysis, drug design, protein engineering, and to study protein evolution and folding.
- **Examples:** Protein Data Bank (PDB), Electron Microscopy Data Bank (EMDB), Biological Magnetic Resonance Bank (BMRB).

- **(iii) Concerted and sequential model.**
  - **Concerted Model (Monod-Wyman-Changeux or MWC Model):**
    - **Applicability:** Describes allosteric regulation in multimeric proteins (proteins with multiple identical subunits).
    - **Conformational States:** Proposes that all subunits of a multimeric protein simultaneously exist in one of two distinct conformational states: a low-affinity T (Tense) state or a high-affinity R (Relaxed) state.
    - **Binding:** Ligand binding (e.g., substrate or activator) shifts the equilibrium towards the R-state, increasing the affinity of all other subunits for the ligand. Inhibitors shift the equilibrium towards the T-state.
    - **Symmetry:** Assumes that the protein maintains symmetry in its transitions; all subunits are in the same state (either all T or all R) at any given time.
    - **Cooperativity:** Explains positive cooperativity (where binding of one ligand increases affinity for subsequent ligands) by the complete shift of all subunits to the R-state.

- **Sequential Model (Koshland-Nemethy-Filmer or KNF Model):**
  - **Applicability:** Also describes allosteric regulation in multimeric proteins.
  - **Conformational States:** Proposes that ligand binding to one subunit induces a conformational change in that specific subunit, which then sequentially influences the conformation of adjacent subunits.
  - **Binding:** Ligand binding to one subunit causes it to transition from a low-affinity to a high-affinity state. This change then facilitates the conformational change and subsequent ligand binding in neighboring subunits.
  - **Asymmetry:** Allows for intermediate states where some subunits are in the T-state and others are in the R-state, leading to asymmetry in the protein structure during the transition.
  - **Cooperativity:** Explains both positive and negative cooperativity (where binding of one ligand decreases affinity for subsequent ligands) through induced conformational changes that sequentially affect neighboring subunits.

- **(iv) R and T-state of Hb.**
  - **T-state (Tense or Taut State):**
    - **Oxygen Affinity:** Low affinity for oxygen.
    - **Structure:** Characterized by a more constrained or "tense" conformation. The central cavity between the subunits is larger, allowing for the binding of allosteric regulators like 2,3-Bisphosphoglycerate (BPG).

        - **Dominance:** Predominant in the tissues, where oxygen partial pressure is low, facilitating the release of oxygen from hemoglobin.
        - **Stability:** Stabilized by a network of ion pairs and hydrogen bonds between the α and β subunits.
    - **R-state (Relaxed State):**
        - **Oxygen Affinity:** High affinity for oxygen.
        - **Structure:** Characterized by a more relaxed or "open" conformation. The central cavity between the subunits becomes smaller, making it difficult for BPG to bind.
        - **Dominance:** Predominant in the lungs, where oxygen partial pressure is high, facilitating the binding of oxygen to hemoglobin.
        - **Stability:** Formed upon oxygen binding, which causes a conformational change that breaks some of the ion pairs stabilizing the T-state.

- **(v) Motif and Domain.**
    - **Motif (or Supersecondary Structure):**
        - **Definition:** A recurring combination of two or more secondary structural elements (e.g., α-helices and β-sheets) that are often found in various proteins. They represent common folding patterns.
        - **Size:** Relatively small, typically consisting of only a few secondary structures.
        - **Function:** While they can sometimes be associated with a specific function (e.g., DNA binding motifs), they are more often primarily structural elements. They may not be independently stable or functional outside the context of a larger domain or protein.

        - **Examples:** Helix-loop-helix, β-hairpin, $\beta - \alpha - \beta$ motif, Greek key motif.

    - **Domain:**

        - **Definition:** A distinct, compact, stable, and independently folding globular unit within a larger polypeptide chain. A protein may consist of one or more domains.
        - **Size:** Larger than motifs, typically 50-200 amino acids, but can be larger.
        - **Function:** Often associated with a specific function (e.g., enzymatic activity, binding site, regulatory function) or structural role, and can sometimes function independently when separated from the rest of the protein. Proteins with multiple domains may combine different functions.
        - **Examples:** The DNA-binding domain of a transcription factor, the catalytic domain of an enzyme, the immunoglobulin domain found in antibodies.

(b) What are the important features of peptide bond?

- **1. Planar Structure:** The peptide bond (composed of the carbonyl carbon, carbonyl oxygen, amide nitrogen, and amide hydrogen) is rigid and planar. All these four atoms lie in the same plane. This planarity is a critical determinant of polypeptide conformation.
- **2. Partial Double Bond Character:** Due to resonance, the C-N bond in the peptide linkage has about 40% double-bond character. This restricts free rotation around this bond, contributing to the rigidity of the peptide backbone.
- **3. Trans Configuration (Preferred):** The Cα carbons of adjacent amino acids are almost always in a *trans* configuration relative to the peptide bond. This arrangement minimizes steric hindrance between the R-groups of adjacent amino acids, making it more energetically

favorable than the *cis* configuration (which is rare, primarily seen only when proline is involved).

- **4. Uncharged but Polar:** The peptide bond itself is uncharged, preventing electrostatic repulsion that would restrict folding. However, the carbonyl oxygen has a partial negative charge (δ^-) and the amide nitrogen has a partial positive charge (δ^+), making the peptide bond highly polar.
- **5. Hydrogen Bonding Potential:** The polar nature of the peptide bond allows the carbonyl oxygen to act as a hydrogen bond acceptor and the amide hydrogen to act as a hydrogen bond donor. These hydrogen bonds are crucial for stabilizing secondary structures like α-helices and β-pleated sheets.
- **6. Fixed Length:** The C-N bond length (1.33 Å) is shorter than a typical single bond (1.49 Å) but longer than a typical double bond (1.27 Å), reflecting its partial double-bond character.
- **7. Limited Rotational Freedom:** While rotation is restricted around the peptide bond (C-N bond), there is free rotation around the N-Cα (phi, ϕ) and Cα-C (psi, ψ) bonds. These dihedral angles determine the conformation of the polypeptide backbone and are crucial for protein folding.

(c) Mention the effects of following chemicals on protein structure/functions :

- **(i) Hydrazine:**
  - **Effect:** Hydrazine ($NH_2 - NH_2$) is a strong nucleophile and reducing agent. Its primary effect on proteins is **cleavage of peptide bonds**, particularly at the C-terminal end, through a reaction called hydrazinolysis.
  - **Mechanism:** Hydrazine attacks the carbonyl carbon of the peptide bond, leading to the formation of amino acid hydrazides, and releasing the free C-terminal amino acid. If the

protein is completely hydrazinolyzed, all amino acids (except the C-terminal one) are converted to hydrazides.

- **Impact on Structure/Function:** Complete hydrazinolysis leads to the **complete denaturation and fragmentation** of the protein, abolishing all levels of protein structure (secondary, tertiary, quaternary) and completely **destroying its biological function**. It is used in protein sequencing to identify the C-terminal amino acid.

- **(ii) Performic acid:**
  - **Effect:** Performic acid (HCO_3H) is a powerful **oxidizing agent**. Its primary effect on proteins is the **irreversible oxidation of disulfide bonds** and methionine residues.
  - **Mechanism:**
    - **Disulfide Bonds (-S-S-):** Performic acid oxidizes disulfide bridges (covalent bonds between two cysteine residues) to two cysteic acid residues. Cysteic acid is a highly polar, charged residue.
    - **Methionine (Met):** Oxidizes methionine residues to methionine sulfone or methionine sulfoxide.
  - **Impact on Structure/Function:**
    - **Loss of Tertiary/Quaternary Structure:** By irreversibly breaking disulfide bonds, performic acid **disrupts the tertiary structure** (and sometimes quaternary structure) that is stabilized by these covalent cross-links. This often leads to unfolding or denaturation.
    - **Altered Folding Pathways:** The introduction of charged cysteic acid residues and oxidized methionine residues can significantly alter the physical and chemical properties of the protein, preventing correct refolding.

- **Complete Loss of Function:** This structural disruption almost always leads to a **complete loss of biological function** because the correct three-dimensional active site or binding site is no longer maintained. Performic acid treatment is used in protein sequencing and purification to permanently break disulfide bonds prior to further analysis (e.g., Edman degradation after reduction).

Question 3: "3. (a) Describe the steps involved in the synthesis of a dipeptide Gly-Ala using Solid-Phase Peptide synthesis."

Solid-Phase Peptide Synthesis (SPPS) is a widely used method for synthesizing peptides. It involves attaching the growing peptide chain to an insoluble polymeric support (resin), allowing for easy washing and filtration between steps. Here's how to synthesize Gly-Ala using SPPS:

- **Objective:** Synthesize the dipeptide Glycylalanine (Gly-Ala), where Glycine (Gly) is the N-terminal amino acid and Alanine (Ala) is the C-terminal amino acid.
- **Key Principles:**
  - **C-terminal to N-terminal:** Synthesis proceeds from the C-terminus to the N-terminus.
  - **Solid Support:** The C-terminal amino acid is anchored to an insoluble resin.
  - **Amino Group Protection:** The α-amino group of each incoming amino acid is temporarily protected (e.g., with Fmoc or Boc group) to prevent unwanted reactions.
  - **Side Chain Protection:** Any reactive side chains of amino acids are permanently protected throughout the synthesis until the very end.
  - **Coupling Agents:** Used to activate the carboxyl group of the incoming protected amino acid, facilitating peptide bond formation.

- **Deprotection:** Removal of the α-amino protecting group after each coupling step.
- **Cleavage:** Final removal of the peptide from the resin and deprotection of side chains.

- **Steps for Synthesizing Gly-Ala using Fmoc-SPPS:**
  - **Step 1: Preparation of the Resin and Loading of the C-Terminal Amino Acid (Alanine)**
    - **Resin Selection:** Choose a suitable resin, commonly Wang resin or Rink Amide resin, which has a linker that can be cleaved under mild acidic conditions (e.g., with trifluoroacetic acid, TFA) at the end of synthesis. The resin needs to have a reactive group (e.g., hydroxyl or amino group) for attachment.
    - **Loading Alanine:** The first amino acid, Fmoc-Alanine (Fmoc-Ala-OH), with its α-amino group protected by Fmoc and its carboxyl group free, is coupled to the resin's linker. This is done using a coupling agent (e.g., DIC/HOBt or HATU/DIPEA) to form an ester bond (with Wang resin) or an amide bond (with Rink Amide resin) between the carboxyl group of Alanine and the functional group on the resin.
    - *Result:* Resin-O-CO-Ala-Fmoc (assuming Wang resin, and the peptide chain grows from Ala's C-terminus).
  - **Step 2: Fmoc Deprotection (of Alanine)**
    - **Removal of Fmoc:** The Fmoc protecting group from the α-amino group of the resin-bound Alanine is removed. This is typically achieved by treating the resin with a dilute solution of a base, most commonly **piperidine** (e.g., 20% piperidine in DMF).

- **Washing:** After deprotection, the resin is thoroughly washed with DMF (dimethylformamide) to remove excess piperidine and the cleaved Fmoc byproduct.
- *Result:* Resin-O-CO-Ala-NH_2 (free amino group of Alanine exposed).

- **Step 3: Coupling of the Next Amino Acid (Glycine)**
  - **Activation and Coupling:** The next amino acid, Fmoc-Glycine (Fmoc-Gly-OH), with its α-amino group protected, is activated using a coupling agent (e.g., DIC/HOBt or HATU/DIPEA). This activated Glycine is then added to the resin-bound Alanine.
  - **Peptide Bond Formation:** The free α-amino group of Alanine (on the resin) nucleophilically attacks the activated carbonyl carbon of Fmoc-Glycine, forming a peptide bond between Glycine and Alanine.
  - **Washing:** The resin is washed extensively with DMF to remove unreacted amino acid, coupling reagents, and byproducts.
  - *Result:* Resin-O-CO-Ala-CO-Gly-Fmoc (the dipeptide Gly-Ala is formed, with its N-terminus protected).
- **Step 4: Final Fmoc Deprotection (of Glycine)**
  - **Removal of Fmoc:** The Fmoc protecting group from the α-amino group of the newly coupled Glycine is removed using piperidine, as in Step 2.
  - **Washing:** Thorough washing with DMF.
  - *Result:* Resin-O-CO-Ala-CO-Gly-NH_2 (the complete dipeptide Gly-Ala is attached to the resin, with its N-terminus free).

- **Step 5: Cleavage from the Resin and Side Chain Deprotection**
    - **Cleavage Cocktail:** The resin-bound dipeptide is treated with a strong acid mixture, typically **Trifluoroacetic Acid (TFA)**, often with scavengers (e.g., triisopropylsilane, water) to prevent side reactions. This cocktail simultaneously cleaves the peptide from the resin and removes any permanent protecting groups from amino acid side chains (though for Gly and Ala, no side chain protection is typically needed).
    - **Precipitation and Purification:** The cleaved peptide is then precipitated (e.g., with cold diethyl ether), filtered, and purified (e.g., by HPLC) to obtain the pure dipeptide Gly-Ala.
    - *Result:* Free dipeptide H_2N-Gly-Ala-COOH.

Question 3: "(b) Discuss the various accessory proteins/enzymes involved in protein folding and prevent their misfolding or aggregation."

The process of protein folding *in vivo* is complex and often requires the assistance of various cellular accessory proteins and enzymes, collectively known as **chaperones** and **folding catalysts**. These proteins help nascent polypeptide chains achieve their correct three-dimensional conformation and prevent misfolding or aggregation, which can lead to dysfunctional proteins and diseases.

- **1. Chaperones (Heat Shock Proteins - HSPs):**
    - **General Function:** Chaperones are a class of proteins that assist in the proper folding of other proteins, particularly nascent (newly synthesized) polypeptides and proteins undergoing stress (e.g., heat shock). They do not become part of the final folded protein structure.

- **Mechanism:** They typically bind to exposed hydrophobic regions of unfolded or partially folded proteins, preventing premature aggregation with other hydrophobic regions (either on the same protein or other proteins). They often use ATP hydrolysis to facilitate cycles of binding and release, promoting proper folding.
- **Types and Examples:**
  - **Hsp70 family (e.g., DnaK in bacteria, BiP in ER, Hsc70 in cytosol):** These chaperones bind to short hydrophobic segments of nascent polypeptide chains as they emerge from the ribosome, preventing aggregation and allowing time for proper folding. They work with co-chaperones (like Hsp40/DnaJ) and nucleotide exchange factors.
  - **Chaperonins (e.g., GroEL/GroES in bacteria, Hsp60/Hsp10 in eukaryotes):** These form large, barrel-shaped complexes. They encapsulate partially folded or misfolded proteins within their central cavity, providing an isolated environment where the protein can fold correctly without interference from other proteins. ATP hydrolysis drives cycles of conformational change within the chaperonin, assisting in the folding process.
  - **Hsp90 family:** Involved in the folding and activation of specific client proteins, often signal transduction proteins like steroid hormone receptors and protein kinases. They help refine the folding of proteins that are already partially folded.
  - **Small Heat Shock Proteins (sHSPs):** Act as holdases, binding to partially denatured proteins under stress conditions (like heat) to prevent their irreversible aggregation. They hold proteins in a folding-competent state until larger chaperones (like Hsp70 or chaperonins) become available.

- **2. Folding Catalysts (Enzymes that Catalyze Rate-Limiting Steps):**
    - **General Function:** These are enzymes that specifically catalyze chemical reactions that can be rate-limiting steps in protein folding, accelerating the formation of specific conformations or covalent bonds. They facilitate conformational changes rather than simply preventing aggregation.
    - **Types and Examples:**
        - **Protein Disulfide Isomerase (PDI):**
            - **Function:** Catalyzes the formation, reduction, and rearrangement (isomerization) of disulfide bonds. Disulfide bonds are crucial for stabilizing the tertiary and quaternary structures of many secreted and membrane proteins.
            - **Location:** Primarily found in the endoplasmic reticulum (ER), where proteins that will be secreted or integrated into membranes acquire their disulfide bonds.
            - **Mechanism:** PDI contains active site cysteines that can form transient disulfide bonds with the substrate protein, correcting incorrectly formed disulfide linkages.
        - **Peptidyl Prolyl Isomerase (PPI):**
            - **Function:** Catalyzes the *cis-trans* isomerization of peptide bonds involving proline residues. The *cis* configuration of X-Pro peptide bonds (where X is any amino acid) can be a slow step in protein folding.
            - **Location:** Found in both the cytoplasm and the ER.

- **Mechanism:** PPIs facilitate the rapid interconversion between the *cis* and *trans* isomers of prolyl peptide bonds, which can be crucial for the protein to adopt its correct conformation.

- **3. Degron-Mediated Degradation Systems:**
    - **Function:** While not directly involved in *folding*, these systems are crucial for preventing the accumulation of misfolded or aggregated proteins, which can be toxic to the cell.
    - **Mechanism:** They recognize terminally misfolded proteins (which chaperones could not rescue) and target them for degradation, primarily through the ubiquitin-proteasome system or autophagy. This serves as a quality control mechanism.
    - **Examples:** Ubiquitin ligases (e.g., components of the ER-associated degradation, ERAD pathway) that mark misfolded proteins for proteasomal degradation.

- **Summary:** The intricate network of chaperones and folding catalysts ensures that proteins fold efficiently and correctly, maintaining cellular proteostasis (protein homeostasis). When these systems fail or are overwhelmed (e.g., due to mutations or stress), misfolded proteins can accumulate, leading to aggregation and various protein misfolding diseases (e.g., Alzheimer's, Parkinson's, Huntington's, cystic fibrosis).

Question 3: "(c) Discuss the contributions of following Scientists :"

- **(i) Linus Pauling:**
    - **Contributions:** A towering figure in 20th-century chemistry and biochemistry, Pauling made groundbreaking contributions to understanding the nature of chemical bonds and molecular structure, particularly proteins.
    - **Chemical Bonding:** Developed the concept of electronegativity and applied quantum mechanics to explain the nature of the

chemical bond, culminating in his seminal book "The Nature of the Chemical Bond" (1939).

- **Protein Secondary Structure:** Crucially, he **proposed the α-helix and β-pleated sheet** as the fundamental secondary structures of proteins in 1951, based on precise measurements of bond lengths and angles in amino acids and peptides. These theoretical predictions were later confirmed by X-ray crystallography and remain central to protein structure analysis.
- **Sickle Cell Anemia:** Identified sickle cell anemia as a "molecular disease" in 1949, demonstrating that a specific genetic defect (a single amino acid change in hemoglobin) could lead to a particular disease phenotype. This pioneered the field of molecular medicine.
- **Awards:** Awarded the Nobel Prize in Chemistry in 1954 for his research into the nature of the chemical bond and its application to the elucidation of the structure of complex substances. He also received the Nobel Peace Prize in 1962 for his efforts against nuclear weapons.

- **(ii) Frederick Sanger:**
  - **Contributions:** A highly influential biochemist, Sanger is known for pioneering methods for sequencing both proteins and nucleic acids, revolutionizing molecular biology.
  - **Insulin Sequencing:** Developed methods for determining the **amino acid sequence of proteins**. In 1953, he successfully determined the complete primary amino acid sequence of insulin, the first protein to be fully sequenced. This work demonstrated that proteins have a precise, defined amino acid sequence, a crucial step in understanding their structure and function.
  - **DNA Sequencing:** Later developed the "Sanger sequencing" or dideoxy chain-termination method for **DNA sequencing** in

1977, which became the predominant method for DNA sequencing for decades and was instrumental in the Human Genome Project.

- **Awards:** Awarded two Nobel Prizes in Chemistry: the first in 1958 for his work on the structure of proteins, especially that of insulin; and the second in 1980 (shared) for his contributions concerning the determination of base sequences in nucleic acids.

- **(iii) Robert Bruce Merrifield:**
  - **Contributions:** Developed the revolutionary **Solid-Phase Peptide Synthesis (SPPS)** method, which significantly simplified and accelerated the synthesis of peptides.
  - **SPPS Development:** In the early 1960s, Merrifield conceived the idea of attaching the growing peptide chain to an insoluble polymeric support (resin) to enable easy removal of reagents and byproducts by washing, thus eliminating laborious purification steps between each coupling reaction.
  - **Automation:** His method paved the way for the automation of peptide synthesis, making it possible to synthesize long and complex peptides with high efficiency.
  - **Impact:** SPPS revolutionized biochemistry, pharmacology, and medicine by making synthetic peptides widely available for research, drug development, and therapeutic applications (e.g., synthetic hormones, vaccines).
  - **Awards:** Awarded the Nobel Prize in Chemistry in 1984 for his development of the methodology for chemical synthesis on a solid matrix.

- **(iv) John Kendrew:**

- **Contributions:** A British biochemist who, along with Max Perutz, made pioneering contributions to the understanding of protein structure through X-ray crystallography.
- **Myoglobin Structure:** He was primarily responsible for determining the **first atomic-resolution three-dimensional structure of a protein, myoglobin, in 1958**. This landmark achievement demonstrated how a polypeptide chain folds into a complex, specific 3D structure and confirmed the existence of α-helices and the compact globular nature of proteins.
- **Methodology:** This work involved extensive X-ray diffraction analysis and the use of heavy atom derivatives to solve the phase problem, a major technical challenge.
- **Awards:** Shared the Nobel Prize in Chemistry in 1962 with Max Perutz for their studies of the structures of globular proteins (Perutz for hemoglobin, Kendrew for myoglobin).

- **(v) Christian Anfinsen:**
  - **Contributions:** A crucial figure in protein folding research, most notably for his work demonstrating that the information for a protein's three-dimensional structure is encoded solely in its primary amino acid sequence.
  - **Ribonuclease Refolding Experiment:** In the early 1960s, Anfinsen conducted a classic experiment using the enzyme ribonuclease A. He showed that when denatured (unfolded) ribonuclease was allowed to spontaneously refold under appropriate conditions, it regained its full enzymatic activity and native three-dimensional structure.
  - **Anfinsen's Dogma (Thermodynamic Hypothesis of Protein Folding):** This experiment led to his seminal hypothesis that "the native three-dimensional structure of a protein is determined solely by its amino acid sequence." This implied

that protein folding is a spontaneous process driven by the thermodynamic principles of minimizing free energy.

- **Impact:** This fundamental principle revolutionized the understanding of protein folding, laying the groundwork for much of modern structural biology and bioinformatics.
- **Awards:** Awarded the Nobel Prize in Chemistry in 1972 (shared with Stanford Moore and William Stein) for his work on ribonuclease, especially concerning the connection between the amino acid sequence and the biologically active conformation.

Question 4: "4. (a) Define the cooperativity effect. Discuss the effect of following on binding of oxygen to hemoglobin: CO_2, BPG, H^+."

- **Cooperativity Effect (Allostery/Cooperative Binding):**
  - **Definition:** Cooperativity refers to the phenomenon where the binding of one ligand molecule to a multisubunit protein influences the binding affinity of subsequent ligand molecules to other subunits of the same protein.
  - **Positive Cooperativity:** If the binding of the first ligand molecule *increases* the affinity of other subunits for the same ligand, it is called positive cooperativity. This leads to a sigmoidal (S-shaped) binding curve, rather than a hyperbolic one. Hemoglobin's oxygen binding exhibits positive cooperativity.
  - **Negative Cooperativity:** If the binding of the first ligand molecule *decreases* the affinity of other subunits, it is called negative cooperativity (less common in physiological systems for primary ligands).
  - **Mechanism:** Involves conformational changes induced by ligand binding to one subunit, which are transmitted to other

subunits, altering their ligand-binding sites. This is a characteristic of allosteric proteins.

- **Effect of Following on Binding of Oxygen to Hemoglobin:** Hemoglobin's oxygen binding affinity is exquisitely sensitive to several physiological factors, acting as allosteric regulators. These factors bind at sites other than the oxygen-binding heme group and stabilize either the T-state (low affinity) or R-state (high affinity) of hemoglobin, thus modulating oxygen release.
    - **i)** CO_2 **(Carbon Dioxide):**
        - **Effect:** CO_2 **decreases** hemoglobin's affinity for oxygen. This is part of the **Bohr Effect**.
        - **Mechanism:**
            - **Carbamate Formation:** CO_2 binds directly to the uncharged α-amino groups at the N-termini of each of the four polypeptide chains of hemoglobin, forming carbaminohemoglobin. This reaction releases protons (H^+).
            - **Stabilization of T-state:** The formation of carbamate groups contributes to the stabilization of the T-state (deoxygenated form) of hemoglobin, leading to the release of bound oxygen.
        - **Physiological Relevance:** In metabolically active tissues, where CO_2 levels are high, CO_2 binding to hemoglobin facilitates oxygen unloading, providing more oxygen to the tissues. In the lungs, where CO_2 levels are low, CO_2 dissociates from hemoglobin, promoting oxygen binding.
    - **ii) BPG (2,3-Bisphosphoglycerate):**
        - **Effect:** BPG **decreases** hemoglobin's affinity for oxygen.

- **Mechanism:** BPG is a highly negatively charged molecule that binds specifically to the central cavity formed by the four subunits in the **T-state (deoxygenated form)** of hemoglobin. It forms electrostatic interactions with positively charged amino acid residues in the pocket.
    - **Stabilization of T-state:** By binding to and stabilizing the T-state, BPG shifts the oxygen-binding equilibrium towards the deoxygenated form, making it easier for oxygen to be released.
    - **Exclusion from R-state:** When hemoglobin transitions to the R-state (oxygenated form), the central cavity narrows, expelling BPG.
- **Physiological Relevance:** BPG levels are higher in individuals living at high altitudes (to compensate for lower atmospheric O_2), and in conditions like chronic hypoxia or anemia. This allows hemoglobin to release more oxygen to the tissues under these challenging conditions. Fetal hemoglobin has a lower affinity for BPG, ensuring efficient oxygen transfer from mother to fetus.

- **iii) H^+ (Protons/pH):**
    - **Effect:** Increased H^+ concentration (lower pH) **decreases** hemoglobin's affinity for oxygen. This is also part of the **Bohr Effect**.
    - **Mechanism:**
        - **Protonation of Key Residues:** Protons bind to specific amino acid residues (like Histidine residues) in hemoglobin, particularly those that are deprotonated in the R-state and protonated in the T-state.

- **Stabilization of T-state:** The protonation of these residues forms salt bridges or strengthens existing ones, stabilizing the T-state (deoxygenated form) and promoting oxygen release.

- **Physiological Relevance:** In active tissues, metabolism produces CO_2 (which forms carbonic acid, H_2CO_3, and then H^+ and HCO_3^-) and lactic acid, leading to a drop in pH. This lower pH causes hemoglobin to release oxygen more readily to the tissues where it is most needed. In the lungs, where CO_2 is expelled, the pH rises, favoring oxygen binding.

4. (b) Describe the structure and function of α-Keratin."

- **Structure of α-Keratin:**
  - **Fibrous Protein:** α-Keratin is a prominent example of a fibrous protein, meaning it has an elongated, filamentous structure, unlike the compact globular proteins. It is highly insoluble in water.
  - **Primary Structure:** Characterized by a repeating seven-residue pseudo-repeat known as a **heptad repeat (a-b-c-d-e-f-g)**, where positions 'a' and 'd' are typically occupied by hydrophobic residues (e.g., Leucine, Isoleucine, Valine). This hydrophobic periodicity is crucial for its higher-order structure.
  - **Secondary Structure:** The basic structural unit of α-keratin is the **right-handed α-helix**. Each polypeptide chain coils into an α-helix.
  - **Quaternary Structure (Coiled-Coil):**
    - **Dimer Formation:** Two α-helices, running parallel to each other and with their hydrophobic faces aligned, intertwine to form a **left-handed supercoil** called a **coiled-coil dimer**. The hydrophobic residues at positions

'a' and 'd' in the heptad repeat of each helix interact to form a hydrophobic stripe along one side of the helix, facilitating this stable coiled-coil interaction.

- **Protofilament Formation:** These coiled-coil dimers then associate in a staggered, antiparallel fashion to form **protofilaments**.
- **Protofibril Formation:** Two protofilaments intertwine to form a **protofibril**.
- **Intermediate Filaments:** Multiple protofibrils then assemble to form larger, highly stable structures called **intermediate filaments**. These filaments have a diameter of about 10 nm, which is "intermediate" between actin microfilaments (7 nm) and myosin thick filaments (15 nm) in muscle cells.

- **Disulfide Bonds:** α-Keratin is rich in cysteine residues. Extensive **disulfide bonds** form between the individual coiled-coil dimers and between different intermediate filaments, providing immense strength, rigidity, and insolubility. The more disulfide bonds, the harder and more rigid the keratin (e.g., in nails and horns) compared to softer keratin (e.g., in skin and hair).

- **Function of α-Keratin:**
  - **Structural Support and Protection:**
    - α-Keratin is the primary structural protein in various protective biological structures, providing mechanical strength, rigidity, and resilience.
    - **Skin:** Forms the protective outer layer of the epidermis, acting as a barrier against pathogens, water loss, and physical trauma.

        - **Hair and Wool:** Provides the characteristic strength, flexibility, and elasticity of hair fibers. The degree of curl in hair is influenced by the pattern of disulfide bonds.
        - **Nails, Claws, Hooves, Horns:** Forms very hard, durable structures due to a high density of disulfide cross-links.
    - **Cytoskeletal Component:** As intermediate filaments, α-keratin forms a crucial part of the cytoskeleton in epithelial cells. It provides internal structural support, maintains cell shape, and helps cells resist mechanical stress. It connects cells to each other and to the extracellular matrix, distributing applied forces across tissues.
    - **Waterproofing:** The hydrophobic nature of α-keratin, along with its extensive cross-linking, contributes to the waterproofing properties of skin, preventing excessive water loss and entry.
    - **Disease Association:** Mutations in keratin genes can lead to genetic disorders like epidermolysis bullosa simplex, a blistering skin disease, demonstrating its critical role in tissue integrity.

4. (c) Proteins are diversified in their Biological functions. Explain."

Proteins are arguably the most versatile macromolecules in living systems, performing an astonishing array of functions essential for life. Their functional diversity stems from their ability to fold into incredibly varied and specific three-dimensional structures, dictated by their amino acid sequence.

- **1. Catalysis (Enzymes):**
    - **Explanation:** The vast majority of biological reactions are catalyzed by enzymes, which are almost exclusively proteins. Proteins fold into specific active sites that bind to substrates, lower activation energy, and dramatically increase reaction rates.

  - **Example: Hexokinase** phosphorylates glucose, initiating glycolysis. **DNA Polymerase** synthesizes new DNA strands. **Amylase** breaks down starch.

- **2. Structural Support:**
  - **Explanation:** Many proteins serve as building blocks and provide mechanical strength, shape, and support to cells, tissues, and organisms. They form the extracellular matrix and cytoskeletal elements.
  - **Example: Collagen** provides tensile strength to skin, tendons, and bones. **Keratin** forms hair, nails, and the outer layer of skin. **Actin** and **Tubulin** form the cytoskeleton, maintaining cell shape and enabling cell movement.
- **3. Transport and Storage:**
  - **Explanation:** Proteins facilitate the movement of molecules (ions, gases, nutrients, waste products) across membranes or within the body, and they can store certain substances.
  - **Example: Hemoglobin** transports oxygen in the blood. **Myoglobin** stores oxygen in muscles. **Transferrin** transports iron in the blood. **Ion channels** and **transporters** (e.g., glucose transporter GLUT1) move specific molecules across cell membranes.
- **4. Immune Defense:**
  - **Explanation:** Proteins are central to the immune system, recognizing and neutralizing foreign invaders like bacteria, viruses, and toxins.
  - **Example: Antibodies (Immunoglobulins)** specifically bind to antigens to neutralize pathogens. **Complement proteins** help destroy pathogens.
- **5. Signaling and Communication:**

- **Explanation:** Proteins act as messengers, receptors, and transducers, allowing cells to communicate with each other and respond to their environment.
- **Example: Hormones** (e.g., insulin) carry signals between cells. **Receptors** (e.g., G-protein coupled receptors) bind signaling molecules and transmit signals into the cell. **Kinases** phosphorylate other proteins, amplifying and relaying signals.

- **6. Movement and Contraction:**
  - **Explanation:** Proteins are essential components of motor systems, enabling cellular and organismal movement.
  - **Example: Actin** and **Myosin** are key proteins in muscle contraction. **Dynein** and **Kinesin** are motor proteins involved in intracellular transport and cell division.
- **7. Regulation and Gene Expression:**
  - **Explanation:** Proteins play critical roles in regulating gene expression, controlling cell growth, differentiation, and metabolism.
  - **Example: Transcription factors** bind to DNA to regulate gene transcription. **Repressors** and **activators** control gene expression. **Cyclins** and **Cyclin-Dependent Kinases** regulate the cell cycle.
- **8. Storage of Amino Acids (Nutrient Storage):**
  - **Explanation:** Some proteins serve as a reserve source of amino acids for growth and development, particularly in seeds and eggs.
  - **Example: Ovalbumin** in egg white. **Casein** in milk. **Zein** in corn seeds.

- **Conclusion:** The remarkable functional diversity of proteins arises from their vast chemical diversity (20 common amino acids) and their ability to fold into intricate, highly specific three-dimensional structures. Each unique structure dictates a unique function, enabling proteins to carry out virtually every biological process necessary for life.

Question 5: "5. (a) Discuss the experiment which explained that primary structure of proteins contains all the necessary information of protein folding."

The seminal experiment that demonstrated that the primary amino acid sequence of a protein contains all the necessary information for its three-dimensional structure and function was performed by **Christian Anfinsen** in the early 1960s, using the enzyme **ribonuclease A (RNase A)**.

- **Background:**
  - Ribonuclease A is a small protein (124 amino acids) that catalyzes the hydrolysis of RNA.
  - It contains four disulfide bonds, which are crucial for stabilizing its native, active conformation.
  - At the time, it was unknown whether protein folding was an inherent property of the amino acid sequence or if it required additional cellular machinery (e.g., a "template" or "chaperones," though the latter term wasn't yet established in its modern sense).
- **The Experiment (Denaturation and Renaturation of Ribonuclease A):**
  - **Step 1: Denaturation (Unfolding and Inactivation)**
    - **Disulfide Bond Reduction:** Anfinsen treated active, native RNase A with a high concentration of a **reducing agent**, β-mercaptoethanol (or dithiothreitol, DTT). This

broke the four disulfide bonds by reducing them to free sulfhydryl (-SH) groups.

- **Disruption of Non-covalent Interactions:** Simultaneously, he added a strong **denaturing agent**, such as a high concentration of urea (e.g., 8 M urea) or guanidine hydrochloride. This disrupted all the non-covalent interactions (hydrogen bonds, hydrophobic interactions, ionic bonds) that stabilize the protein's secondary and tertiary structures, causing the polypeptide chain to completely unfold into a random coil.
- **Result:** The protein became fully denatured and lost all its enzymatic activity. The disulfide bonds were cleaved, and the polypeptide chain was in an unfolded, random-coil conformation.

- **Step 2: Renaturation (Refolding and Regaining Activity)**
  - **Removal of Denaturants:** Anfinsen then carefully removed both the reducing agent (β-mercaptoethanol) and the denaturing agent (urea) by dialysis.
  - **Spontaneous Refolding:** Upon removal of these denaturing conditions, the denatured, inactive RNase A spontaneously refolded into its native, catalytically active three-dimensional conformation.
  - **Disulfide Bond Re-formation:** The correct disulfide bonds (the original four) re-formed spontaneously during this refolding process, even though there were many incorrect disulfide pairings possible (105 different ways to form four disulfide bonds from eight cysteine residues).
  - **Result:** The refolded RNase A regained virtually 100% of its original enzymatic activity, indicating that it had returned to its correct, native three-dimensional structure.

- **Conclusion (Anfinsen's Dogma / Thermodynamic Hypothesis):**
    - The striking observation that denatured RNase A could spontaneously refold and regain its full activity led Anfinsen to conclude that:
        - **The primary amino acid sequence of a protein contains all the necessary information to specify its unique three-dimensional structure.**
        - Protein folding is a **spontaneous process** driven by the protein seeking its lowest free energy (most thermodynamically stable) conformation under physiological conditions.
- **Significance and Impact:**
    - This experiment was a landmark achievement, fundamentally changing the understanding of protein folding. It established the central dogma of protein folding and laid the foundation for much of modern structural biology, bioinformatics, and protein engineering.
    - It implied that the complex process of protein folding is not guided by an external template but is an intrinsic property determined by the sequence of amino acids.
    - While we now know that *in vivo* folding is often assisted by chaperones, Anfinsen's principle holds true: chaperones help *prevent misfolding and aggregation* and *accelerate* folding, but they do not provide the folding information itself. The final native state is still dictated by the primary sequence.

5. (b) Mention the important forces/bonds responsible for stabilizing the protein structure."

Protein structure is stabilized by a combination of strong covalent bonds and weaker non-covalent interactions. These forces act hierarchically to maintain the protein's precise three-dimensional conformation.

- **1. Covalent Bonds:**
    - **Peptide Bonds:** These are the primary covalent bonds that link amino acids together to form the polypeptide backbone. While they define the primary structure, their rigid planar nature also imposes crucial constraints on higher-order structures.
    - **Disulfide Bonds (S-S bonds):** Formed by the oxidation of two cysteine sulfhydryl (-SH) groups to create a covalent bond between them. These bonds are relatively strong and play a critical role in stabilizing the tertiary and quaternary structures of many extracellular and secreted proteins. They also help to lock a protein into a particular conformation.
- **2. Non-Covalent Interactions (Weak Interactions):** These interactions are individually weak but collectively provide significant stability to the folded protein structure. They are highly sensitive to changes in temperature, pH, and solvent conditions.
    - **i) Hydrophobic Interactions:**
        - **Mechanism:** The most significant driving force for protein folding. Nonpolar (hydrophobic) amino acid side chains tend to cluster together in the interior of the protein, away from the aqueous solvent. This reduces the unfavorable interactions between nonpolar groups and water, minimizing the disruption of water's hydrogen bonding network and increasing the entropy of the surrounding water molecules.
        - **Location:** Primarily in the protein's core, away from the surface.
        - **Impact:** Crucial for the formation of the compact, globular shape of proteins.
    - **ii) Hydrogen Bonds:**

- **Mechanism:** Formed between a hydrogen atom (covalently bonded to an electronegative atom like N or O) and another electronegative atom (like O or N) with a lone pair of electrons.
- **Location:**
  - **Backbone-Backbone H-bonds:** Form the basis of secondary structures (α-helices and β-pleated sheets) by forming between the carbonyl oxygen and amide hydrogen atoms of the polypeptide backbone.
  - **Side Chain-Side Chain H-bonds:** Occur between polar amino acid side chains.
  - **Side Chain-Backbone H-bonds:** Occur between polar side chains and backbone atoms.
- **Impact:** Provide significant stability to secondary and tertiary structures and are critical for defining specific protein folds.

- **iii) Ionic Interactions (Salt Bridges):**
  - **Mechanism:** Electrostatic interactions between oppositely charged amino acid side chains (e.g., between a positively charged Lysine or Arginine and a negatively charged Aspartate or Glutamate).
  - **Location:** Can occur on the protein surface or in the interior, depending on the surrounding environment.
  - **Impact:** Contribute to tertiary and quaternary structure stability. They are highly sensitive to pH changes.
- **iv) Van der Waals Forces (London Dispersion Forces):**
  - **Mechanism:** Weak, transient attractive forces that arise from temporary fluctuations in electron distribution around

atoms, creating transient dipoles that induce complementary dipoles in nearby atoms. These forces are very weak and occur between any atoms that are very close to each other.

- **Location:** Present throughout the protein's interior, wherever atoms are in close proximity.
- **Impact:** While individually weak, their cumulative effect can be substantial, especially in the tightly packed interior of a folded protein, contributing to the overall stability and precise packing of the protein core.

- **Summary:** The overall stability of a protein's native structure is a delicate balance of these various forces. Covalent bonds dictate the primary sequence and provide strong linkages, while the cumulative effect of non-covalent interactions drives the folding process and maintains the precise 3D architecture essential for biological function.

5. (c) Determine the sequence of small peptide based on the following observations :"

- **Given Information:**
  - Heptapeptide (7 amino acids).
  - Complete hydrolysis reveals amino acids: Val, Asp, Lys, Met, Gly, Leu (missing one, must be 2 of one kind or one is missing from list - re-check. "Heptapeptide" means 7 total. The list has 6. Let's assume one is missing from the list or there's a typo in "heptapeptide" vs count. Let's assume there are 7 total. If a full list were given and only 6 unique AAs, then one must be repeated. Let's proceed based on fragments and infer the list is comprehensive for *unique* AAs and the actual list from fragments will sum to 7).
  - Let's re-evaluate after processing fragments.
- **Observations:**

- **(i) Complete hydrolysis revealed that heptapeptide peptide contain following amino acids: Val, Asp, Lys, Met, Gly, Leu.**
    - Count: Val (1), Asp (1), Lys (1), Met (1), Gly (1), Leu (1). This is 6 amino acids. For a heptapeptide, one amino acid must be repeated. Let's track this as we deduce the sequence.
- **(ii) Reaction with Edman reagent gives PTC-Val adduct formation.**
    - **Conclusion:** Edman degradation removes the N-terminal amino acid. PTC-Val adduct formation means **Val is the N-terminal amino acid.**
    - Sequence starts with: **Val - ...**
- **(iii) CNBr treatment released:**
    - **CNBr cleavage:** Cleaves specifically at the C-terminal side of **Methionine (Met)** residues. If Met is at the C-terminus, it is released as a free amino acid.
    - **Fragments:**
        - **(i) Tetrapeptide having Val, Met, Lys and Asp.**
            - This fragment contains Met. Since it's a CNBr product, Met must be at its C-terminus (unless it was the last Met in the original peptide and this is the C-terminal fragment).
            - So, the sequence of this tetrapeptide is **... - Met**.
            - The amino acids are Val, Met, Lys, Asp.
        - **(ii) Dipeptide having Gly and Met.**
            - This fragment also contains Met. So, its sequence is **... - Met**.

    - The amino acids are Gly, Met. So, it's **Gly - Met**.

- **(iii) Free amino acid was released as Leu.**
    - This means Leu was the C-terminal amino acid of the *original* heptapeptide, or it was immediately C-terminal to a Met, and it's the *only* C-terminal fragment that is a single amino acid. However, CNBr cleaves *after* Met. If Leu is a free amino acid, it implies it was the C-terminal amino acid *of the peptide that resulted from a CNBr cleavage*. This means the original peptide must have ended with **... - Met - Leu**. (If Met was the C-terminal amino acid of the full peptide, it would be a tetrapeptide, etc., not free Leu).
    - Let's reconsider: If CNBr released a free Leu, this must be the very last amino acid of the original peptide, and the last Met was followed by it. So, the end of the original peptide is **- Met - Leu**. This makes Leu the C-terminal.

- **Consolidating CNBr data:**
    - Fragment 1: (Val, Lys, Asp) - Met
    - Fragment 2: Gly - Met
    - Fragment 3: Leu (free C-terminal amino acid)
- This implies the original peptide had two Met residues.
    - The first Met (from fragment 1) is followed by Gly-Met.
    - The second Met (from fragment 2) is followed by Leu.

- So the sequence structure is: **(Val, Lys, Asp) - Met - Gly - Met - Leu**
- Now we have: Val (1), Asp (1), Lys (1), Met (2), Gly (1), Leu (1). This is a total of 7 amino acids. The initial list of unique amino acids was incomplete, implying one Met was repeated.

- **(d) Trypsin treatment released:**
  - **Trypsin cleavage:** Cleaves specifically at the C-terminal side of **Lysine (Lys)** and **Arginine (Arg)** residues.
  - **Fragments:**
    - **(i) Tripeptide containing Val, Lys and Asp**
      - This fragment contains Lys. So, its sequence must be **... - Lys**.
      - The amino acids are Val, Lys, Asp.
      - Given our N-terminal is Val, this fragment must be **Val - Asp - Lys** (or Val-Lys-Asp, but we'll see if it fits).
    - **(ii) Tetrapeptide containing Gly, Leu and Met** (Note: this list has 3, but says tetrapeptide. This is a common typo in problems. Assuming it implies a 4-amino acid peptide, let's re-evaluate after fitting).
      - Wait, the original list from hydrolysis was Val, Asp, Lys, Met, Gly, Leu. That is 6 unique amino acids. If this is a heptapeptide (7 total), then one must be repeated. We deduced two Met from CNBr fragments. So the total amino acids are: Val, Asp, Lys, Met, Met, Gly, Leu. This is 7 amino acids. So the CNBr deduction is correct.

- Let's check Trypsin fragment (ii) again: "Tetrapeptide containing Gly, Leu and Met". This implies one Met is repeated here, or there's a third Met, which contradicts the overall count. Let's assume the listed AAs (Gly, Leu, Met) are what are present in the fragment. If it is a tetrapeptide, it must contain 4 amino acids. Since we only have Gly, Leu, and Met listed, it implies *one of them is repeated*. From our total AA count (Val, Asp, Lys, Met, Met, Gly, Leu), we have two Met. So, this tetrapeptide must contain **Gly, Leu, Met, Met**.

- **Reconstructing the Sequence:**

  a. **N-terminus:** From Edman, the peptide starts with **Val**.

     - Sequence: **Val - ? - ? - ? - ? - ? - ?**

  b. **C-terminus and Met positions:** From CNBr, the peptide ends with **- Met - Leu**, and there's another Met followed by Gly.

     - So, the two Met are at positions 4 and 6, or some other combination.
     - Let's try to fit the CNBr fragments:
       - Fragment 1: (Val, Lys, Asp) - Met
       - Fragment 2: Gly - Met
       - Fragment 3: Leu (free C-terminal)
     - This means we have: **(Val, Lys, Asp) - Met - Gly - Met - Leu**. This confirms the 7 amino acids: Val, Asp, Lys, Met, Gly, Met, Leu.

  c. **Using Trypsin fragments to refine:**

- Trypsin cleaves after Lys.
- Fragment (i): Tripeptide containing Val, Lys and Asp.
    - Since Val is N-terminal, this fragment must be **Val - Asp - Lys**.
- Fragment (ii): Tetrapeptide containing Gly, Leu and Met.
    - From our CNBr deductions, the fragment **Gly - Met - Leu** is a piece after a Met. If this is a tetrapeptide from trypsin, and it starts after Lys, then it must be **Gly - Met - Met - Leu** (from the list of 4 amino acids). This makes sense, as the first Met must be part of the first trypsin fragment, and the second Met (from Gly-Met-Leu) is part of the second trypsin fragment.

d. **Putting it all together:**

- Start: **Val - ...**
- Trypsin fragment (i) tells us: **Val - Asp - Lys** (since Val is N-terminal and it contains Lys).
- So, the peptide starts with: **Val - Asp - Lys - ...**
- The next fragment from Trypsin is a tetrapeptide: **Gly - Met - Met - Leu**. This must follow the Lys.
- Therefore, the full sequence is: **Val - Asp - Lys - Gly - Met - Met - Leu**

- **Verification:**
    - **Heptapeptide:** Yes, 7 amino acids.
    - **Amino Acids:** Val, Asp, Lys, Gly, Met, Met, Leu. (Matches initial set if one Met is repeated).
    - **Edman:** Starts with Val. Correct.

- **CNBr cleavage (after Met):**
  - **Val - Asp - Lys - Gly - Met** / **Met** / **Leu**
  - Cleaves after the first Met: Gives **Val - Asp - Lys - Gly - Met** (a pentapeptide with Val, Asp, Lys, Gly, Met).
  - Cleaves after the second Met: Gives **Met** (a free amino acid).
  - Gives **Leu** (a free amino acid at the C-terminus, as per CNBr's definition - it means Leu was the C-terminal amino acid of the last fragment).
  - **This does not match CNBr results precisely.** Let's re-read the CNBr results carefully.
    - "(i) Tetrapeptide having Val, Met, Lys and Asp." -> **Val-Asp-Lys-Met** (This means Met is the C-terminal of this fragment.)
    - "(ii) Dipeptide having Gly and Met." -> **Gly-Met** (This means Met is the C-terminal of this fragment.)
    - "(iii) Free amino acid was released as Leu."
  - If Leu is released as a free amino acid by CNBr, it implies the sequence was X-Met-Leu at the end.
  - If we have **Val-Asp-Lys-Met** and **Gly-Met**, and then **Leu** (free), the original peptide must have been assembled from these fragments.
  - Possibility 1: **Val-Asp-Lys-Met - Gly-Met - Leu**
    - This gives: Val, Asp, Lys, Met, Gly, Met, Leu (7 amino acids).
    - CNBr cleavage after the first Met gives: Val-Asp-Lys-Met (Tetrapeptide).

- The next Met is followed by Leu, so cleavage after the second Met gives: Gly-Met and Leu (free).
- This **matches the CNBr results perfectly.**

- **Trypsin cleavage (after Lys):**
  - **Val - Asp - Lys / Gly - Met - Met - Leu**
  - Fragment 1: **Val - Asp - Lys** (Tripeptide, contains Val, Lys, Asp). This matches.
  - Fragment 2: **Gly - Met - Met - Leu** (Tetrapeptide, contains Gly, Leu, Met, Met). This matches.

- **Final Sequence:** Therefore, the sequence of the heptapeptide is: **Val - Asp - Lys - Gly - Met - Met - Leu**

Question 6: "6. Write short note on :"

- **(i) Ramachandran plot**
  - **Definition:** The Ramachandran plot (or Ramachandran diagram) is a graphical representation used in biochemistry to visualize the energetically allowed regions for the backbone dihedral angles (ϕ and ψ) of amino acid residues in protein structures.
  - **Dihedral Angles:**
    - ϕ **(phi):** The torsion angle around the N-Cα bond (between the nitrogen atom of the peptide bond and the alpha carbon).
    - ψ **(psi):** The torsion angle around the Cα-C bond (between the alpha carbon and the carbonyl carbon of the peptide bond).
  - **Purpose:** Since the peptide bond itself is rigid and planar, the flexibility of the polypeptide backbone is primarily determined by

the rotation around these two bonds. However, not all combinations of ϕ and ψ angles are sterically allowed due to collisions between atoms.

- **Plot Structure:** The plot maps ϕ (x-axis) against ψ (y-axis), typically ranging from -180° to +180°.
- **Allowed Regions:**
  - **Favored Regions:** Represent conformations where there are no steric clashes and the atoms are in energetically favorable positions. These regions correspond to the common secondary structures:
    - **α-helices:** Located in a specific region (typically around $\phi = -57°$, $\psi = -47°$).
    - **β-sheets (parallel and antiparallel):** Located in another specific region (typically around $\phi = -119°$, $\psi = +113°$ for antiparallel).
  - **Allowed Regions:** Represent conformations that are sterically permissible but less common.
  - **Disallowed Regions:** Represent conformations where atoms would clash sterically, making these arrangements highly unfavorable and generally not found in protein structures (except for glycine due to its lack of a side chain).
- **Glycine and Proline:** Glycine residues have a much larger allowed region on the Ramachandran plot due to its small side chain (H atom), allowing it to adopt conformations inaccessible to other amino acids. Proline, with its rigid ring structure, has a highly restricted and unique allowed region.
- **Significance:**

- **Protein Structure Validation:** Used as a powerful tool to assess the quality of experimentally determined protein structures (e.g., from X-ray crystallography). A high-quality structure should have most of its residues falling within the allowed or favored regions.
- **Predicting Secondary Structure:** Helps in understanding how amino acid sequences contribute to specific secondary structures.
- **Understanding Folding:** Provides insights into the conformational space accessible to polypeptide chains during folding.

- **(ii) Alzheimer's Disease**
  - **Definition:** Alzheimer's Disease (AD) is a progressive neurodegenerative disorder and the most common cause of dementia. It is characterized by a gradual decline in cognitive function, including memory loss, impaired judgment, language difficulties, and changes in personality and behavior, ultimately leading to severe brain atrophy and death.
  - **Pathology (Protein Misfolding Disease):** AD is classified as a protein misfolding disease (or proteopathy), specifically linked to the accumulation of two distinct types of abnormal protein aggregates in the brain:
    - **Amyloid-beta (Aβ) Plaques:** Extracellular deposits formed from the aggregation of amyloid-beta peptides. Aβ peptides are fragments of a larger protein called amyloid precursor protein (APP). In AD, APP is abnormally processed by enzymes (beta-secretase and gamma-secretase), leading to the generation of sticky Aβ monomers, which then aggregate into oligomers, protofibrils, and eventually insoluble amyloid plaques.

These plaques are toxic to neurons and disrupt synaptic function.

- **Neurofibrillary Tangles (NFTs):** Intracellular aggregates formed by the hyperphosphorylation and aggregation of the tau protein. Tau is a microtubule-associated protein that normally stabilizes microtubules (part of the neuronal cytoskeleton). In AD, tau becomes abnormally phosphorylated, detaches from microtubules, and self-assembles into insoluble paired helical filaments that form neurofibrillary tangles within neurons. These tangles disrupt neuronal transport and eventually lead to neuronal death.

- **Disease Progression:** The accumulation of Aβ plaques and tau tangles is thought to precede symptoms by many years. They disrupt neuronal communication, lead to inflammation, synaptic loss, and widespread neurodegeneration, particularly in areas of the brain critical for memory and cognition (e.g., hippocampus, cerebral cortex).

- **Etiology:** The exact cause of AD is unknown. Most cases are sporadic (late-onset), likely involving a complex interplay of genetic (e.g., APOE4 gene), environmental, and lifestyle factors. A small percentage of cases are familial (early-onset), caused by mutations in specific genes (APP, PSEN1, PSEN2) that lead to increased Aβ production.

- **Symptoms:** Early symptoms include mild memory loss, difficulty with problem-solving, and mood changes. As the disease progresses, symptoms worsen, including disorientation, confusion, difficulty speaking and swallowing, and complete dependence on caregivers.

- **Treatment:** Currently, there is no cure for AD. Treatments focus on managing symptoms (e.g., cholinesterase inhibitors, memantine) and providing supportive care. Research is

ongoing to develop therapies that target the underlying protein pathologies (e.g., anti-amyloid or anti-tau therapies).

- **Significance:** Alzheimer's disease highlights the critical importance of proper protein folding and degradation in maintaining cellular health, particularly in the brain. Misfolding and aggregation of key proteins can have devastating consequences, leading to debilitating neurodegenerative disorders.