BIOLOGICAL PHYSICS

CHAPTER 4 –PROTEIN

Poopalasingam Sivakumar
Department of Physics, SIUC

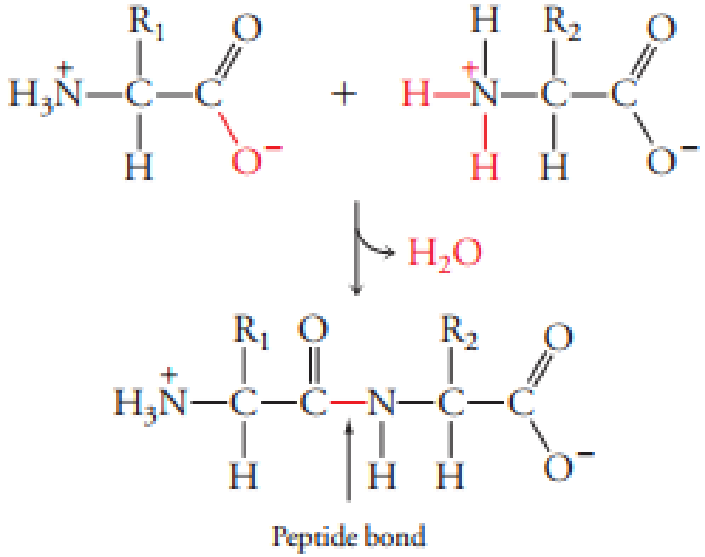
Peptide Bond

Peptide Bond Formation:

Type of Reaction: Condensation (dehydration) reaction.

Process:

- Occurs between the α-carboxyl group of one amino acid and the α-amino group of another.
- Releases a water molecule (H_2O) during bond formation.



Peptide Types:

- Dipeptide: Two amino acids linked.
- Oligopeptide: Short chains (2–20 amino acids).
- Polypeptide: Long chains (>20 amino acids), forming the basis of proteins.

Peptide Bond

Peptide Bond Hydrolysis:

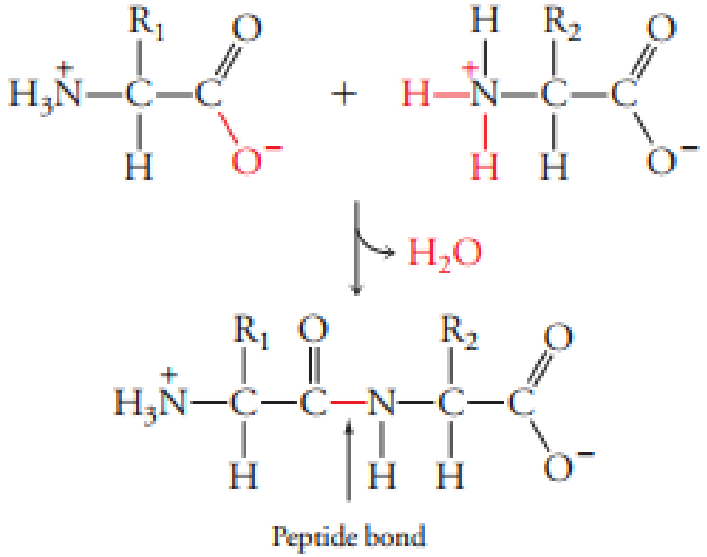
Reverse of condensation: Breaks peptide bonds by adding water.

Catalyzed by enzymes:

Peptidases/Proteases

Exopeptidases: Act on terminal peptide bonds.

Endopeptidases: Target bonds within the peptide chain



Peptide Bond

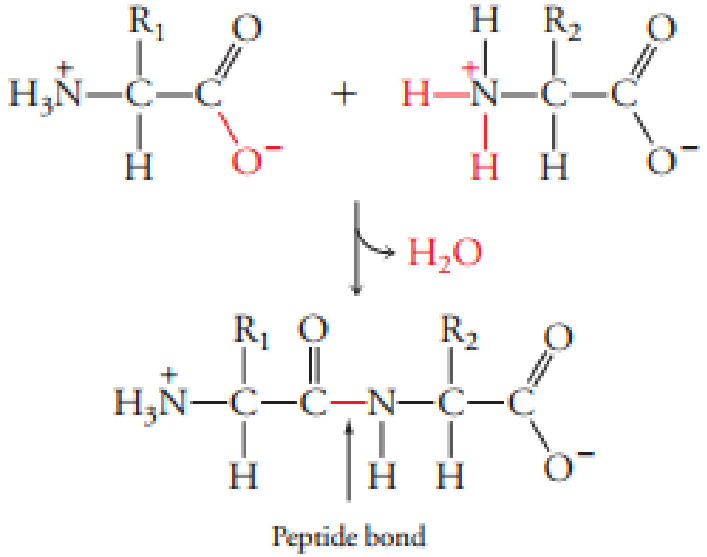
Peptide Bond Stability:

Remarkably stable in proteins under physiological conditions.

Half-life ($t_{1/2}$): Approximately 7 years without enzymatic catalysis.

Hydrolysis characteristics:

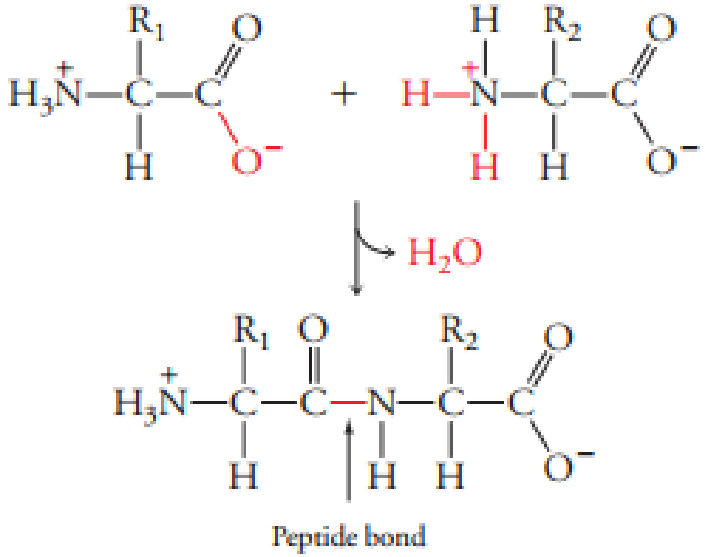
- Exergonic (energy-releasing) but slow due to high activation energy.
- Requires enzymes or harsh conditions to proceed efficiently.



Peptide Bond

Hydrolysis:

- **Water Addition**: A water molecule attacks the carbonyl carbon of the peptide bond.
- **Bond Cleavage**: The bond between the Carbon (C) of the carboxyl group and the Nitrogen (N) of the amino group is broken.



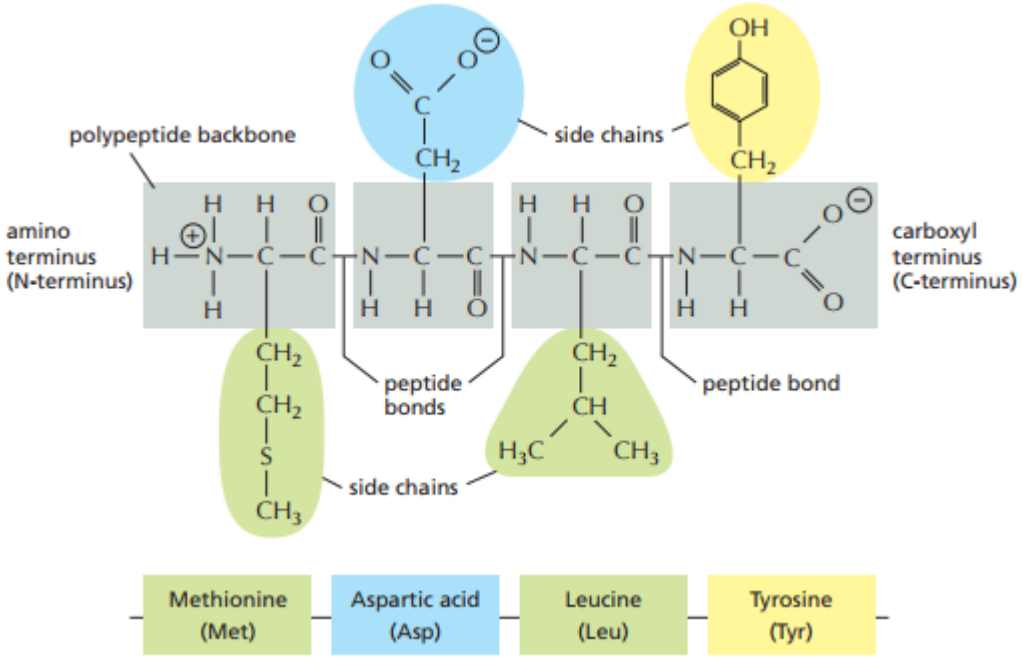
Restoration:

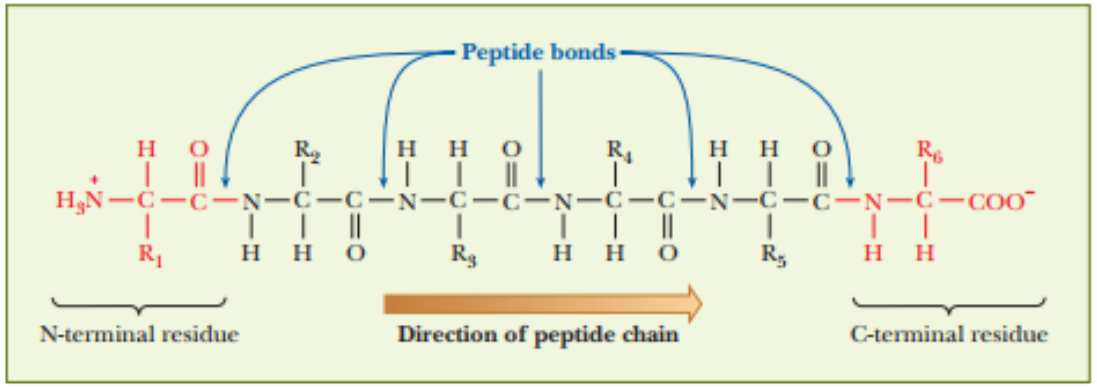
The Hydroxyl group (-OH) from the water attaches to the carboxyl group, recreating the carboxylic acid.

The Hydrogen atom (H^+) from the water attaches to the amino group, recreating the amine.

Polypeptide

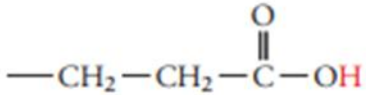
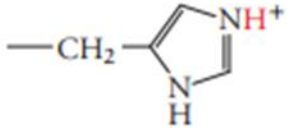
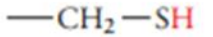
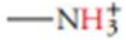
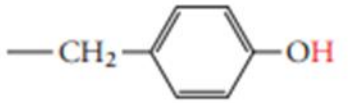
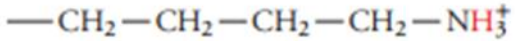
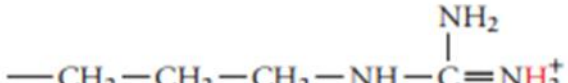
- The N-terminus (left end) has a free amino group, while the C-terminus (right end) has a free carboxylate group.
- The electrostatic properties of a polypeptide are primarily determined by the side chains (R groups) extending from the backbone.
- Peptides can be distinguished by their ionization behavior, which influences their charge and solubility.
- Like free amino acids, peptides exhibit characteristic titration curves and have a specific isoelectric point (pI), where their net charge is zero.
- The pK_a values of ionizable R groups can shift when an amino acid becomes a residue within a peptide, due to interactions with neighboring groups and the peptide environment.





Peptides and Proteins

- The provided table lists pK_a values for all charged and ionizable groups in amino acids, reflecting their tendency to ionize.
- In polymerized amino acids, side chain behavior differs from free amino acids due to:
  - Electronic effects from the peptide bond.
  - Influence of nearby functional groups during polypeptide folding.
  - Microenvironment effects, where neighboring side chains alter polarity and proton exchange capacity.
- The chemical and physical properties of proteins depend on their amino acid composition, leading to distinct behaviors under specific laboratory conditions.

Asp		3.9
Glu		4.1
His		6.0
Cys		8.4
N-terminus		9.0
Tyr		10.5
Lys		10.5
Arg		12.5

Polypeptide Interaction and Protein Folding

- Protein folding is influenced by multiple weak, noncovalent bonds that form between different regions of the polypeptide chain.
- These interactions involve atoms in both the polypeptide backbone and amino acid side chains.
- The three primary types of weak noncovalent bonds that stabilize protein structure are:
  - Hydrogen bonds – Form between partially positive hydrogen and electronegative atoms (O, N).
  - Electrostatic (ionic) attractions – Occur between charged side chains.
  - Van der Waals forces – Weak, short-range attractions between nonpolar atoms.

AMINO ACID			SIDE CHAIN
Aspartic acid	Asp	D	negative
Glutamic acid	Glu	E	negative
Arginine	Arg	R	positive
Lysine	Lys	K	positive
Histidine	His	H	positive
Asparagine	Asn	N	uncharged polar
Glutamine	Gln	Q	uncharged polar
Serine	Ser	S	uncharged polar
Threonine	Thr	T	uncharged polar
Tyrosine	Tyr	Y	uncharged polar

POLAR AMINO ACIDS

AMINO ACID			SIDE CHAIN
Alanine	Ala	A	nonpolar
Glycine	Gly	G	nonpolar
Valine	Val	V	nonpolar
Leucine	Leu	L	nonpolar
Isoleucine	Ile	I	nonpolar
Proline	Pro	P	nonpolar
Phenylalanine	Phe	F	nonpolar
Methionine	Met	M	nonpolar
Tryptophan	Trp	W	nonpolar
Cysteine	Cys	C	nonpolar

NONPOLAR AMINO ACIDS

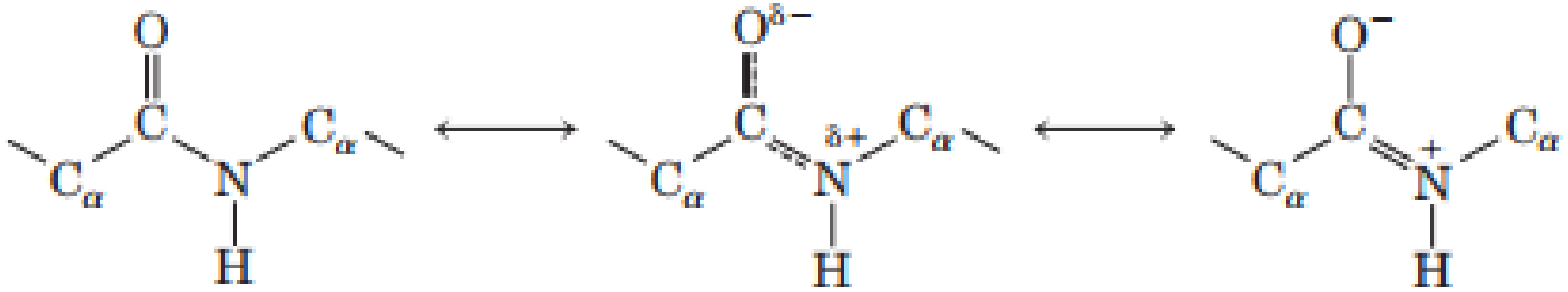
Peptide Bond Struture and Backbone Flexibility

Resonance and Partial Double-Bond Character

- The peptide bond exhibits resonance due to electron delocalization between the carbonyl group and the amide nitrogen:

$$\boldsymbol{O = C - NH \leftrightarrow O - C^{-} = N^{+}}$$

- Because of this resonance:
- The C–N bond has **partial double-bond character (~40%)**
- The peptide bond is **shorter and stronger** than a typical C–N single bond
- The atoms involved in the peptide bond are **planar**



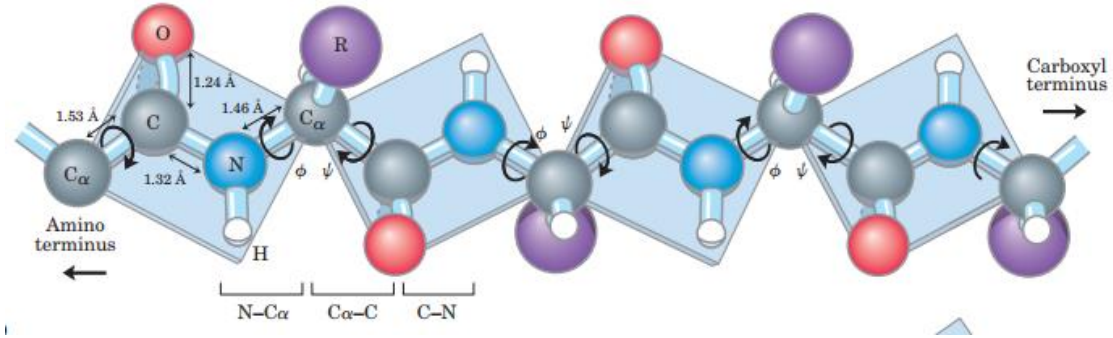
Peptide Bond Struture and Backbone Flexibility

Due to its partial double-bond character:

❌ **No free rotation occurs around the peptide C–N bond**

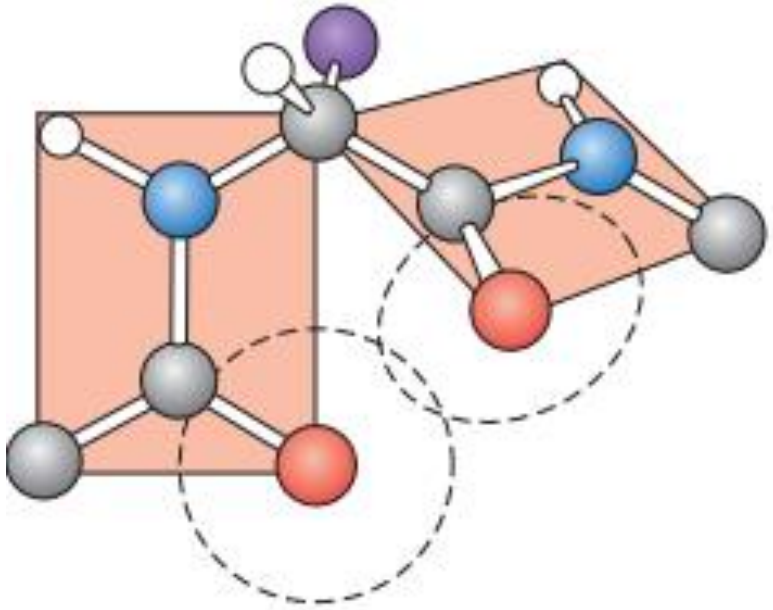
The peptide group is therefore **rigid and planar**.

- The Polypeptide backbone consists of **repeating N-C_α-C units,** where each unit is a **planar peptide group** with all atoms in the same plane.
- Rotation is only possible:
  - ϕ angle- Rotation around the $N - C_\alpha$ bond
  - ψ angle -Rotation around the $C_\alpha - C$ bond,



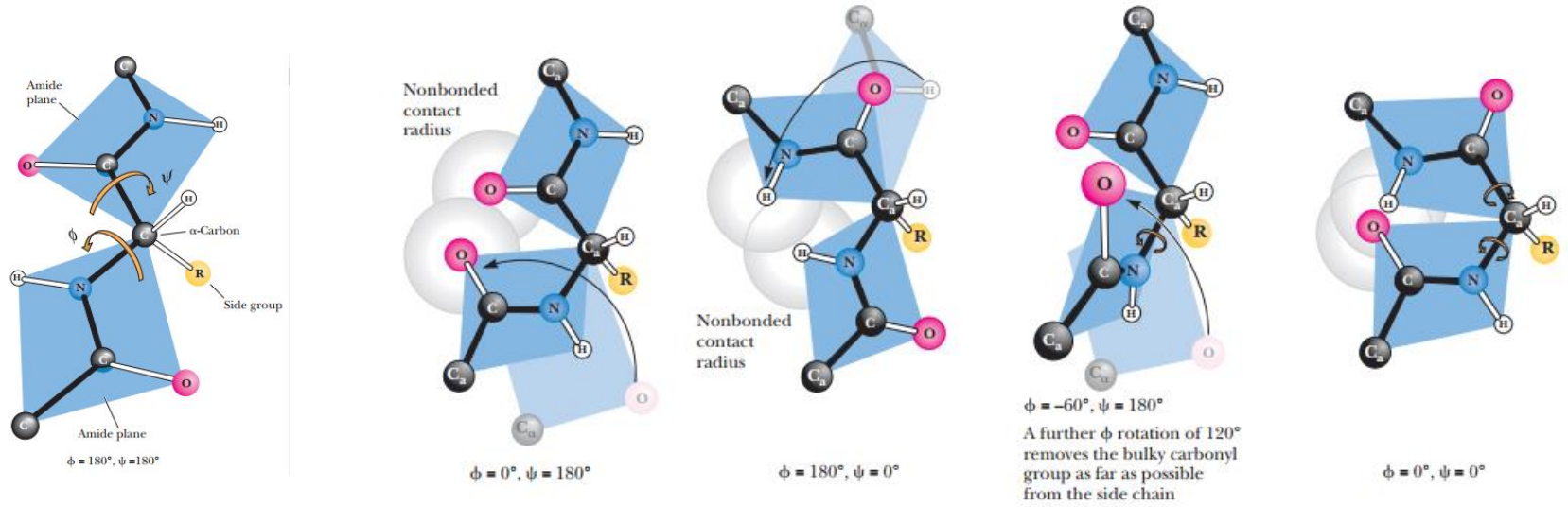
Peptide Bond Backbone Rotation & Steric Hindrance

Steric hindrance (physical crowding of atoms) restricts backbone rotation, preventing conformations where carbonyl oxygens, side chains, or hydrogen atoms clash.



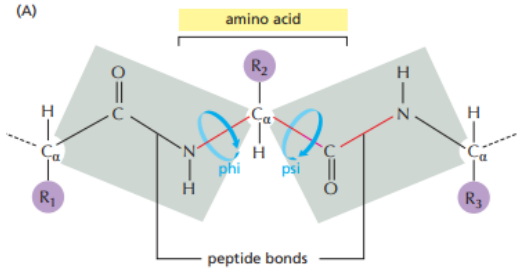
Peptide Bond Backbone Rotation & Steric Hindrance

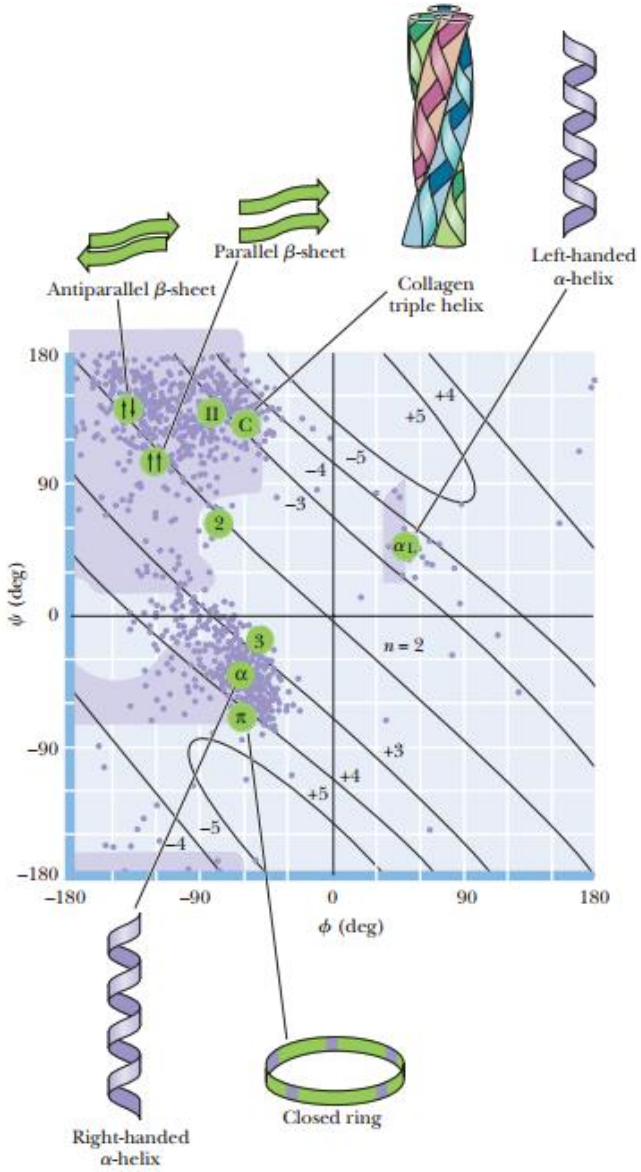
- .In proteins, certain ϕ and ψ angle combinations are sterically disallowed due to nonbonded atom crowding.
  - Example: $\phi = 180°$ & $\psi = 0°$ (and vice versa) are unfavorable due to steric clashes.
  - These steric restrictions help define secondary structure preferences and play a crucial role in maintaining proper protein folding.
- Rotation & Handedness:
  - Right-handed rotation (clockwise) corresponds to positive ϕ and ψ values
  - .Biological structures exhibit specific handedness, such as right-handed α-helices (common in proteins) and left-handed polyproline helices (e.g. collagen triple helices).



Peptide Bond Backbone Rotation & Steric Hindrance

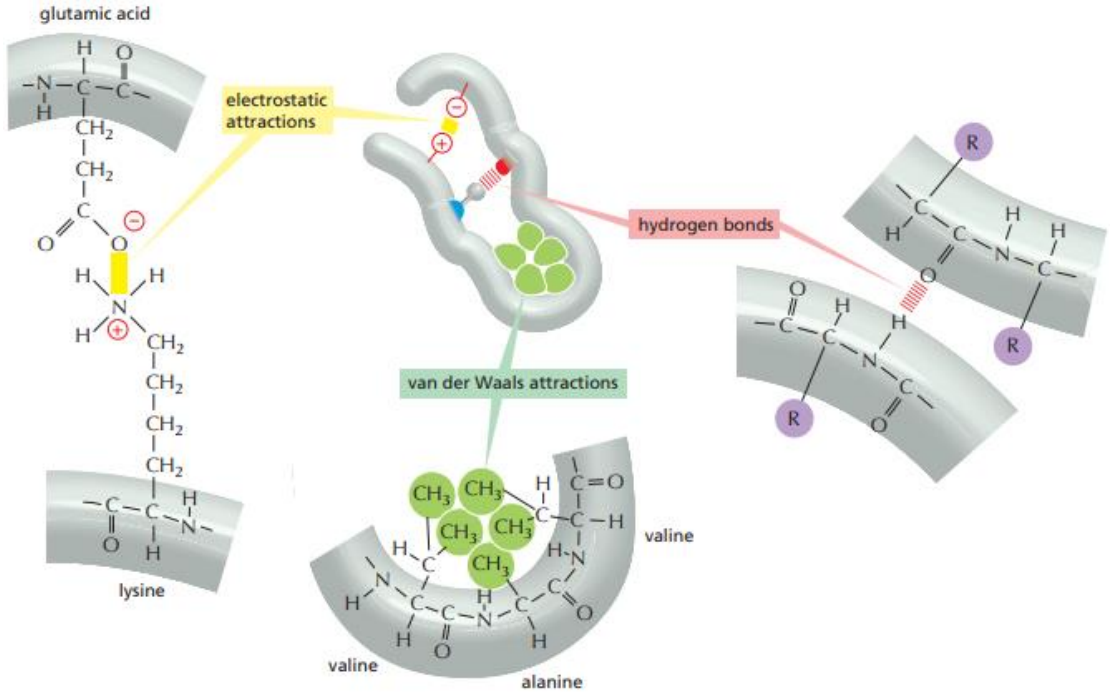
- G. N. Ramachandran and his colleagues in Madras, India, introduced the Ramachandran plot to visualize the allowed and disallowed combinations of ϕ and ψ angles in proteins (*Adv. Prot. Chem. 34:174–175, 1981*).
- Shaded regions indicate energetically favorable conformations, corresponding to common secondary structures: α-helices (right- and left-handed), β-sheets, and collagen helices
- Purple dots on the plot represent experimentally observed ϕ and ψ angles for 1000 amino acid residues (excluding glycine, which has a broader range due to its small side chain).
- This plot is a valuable tool for analyzing protein folding, validating protein structures, and predicting secondary structure formation.

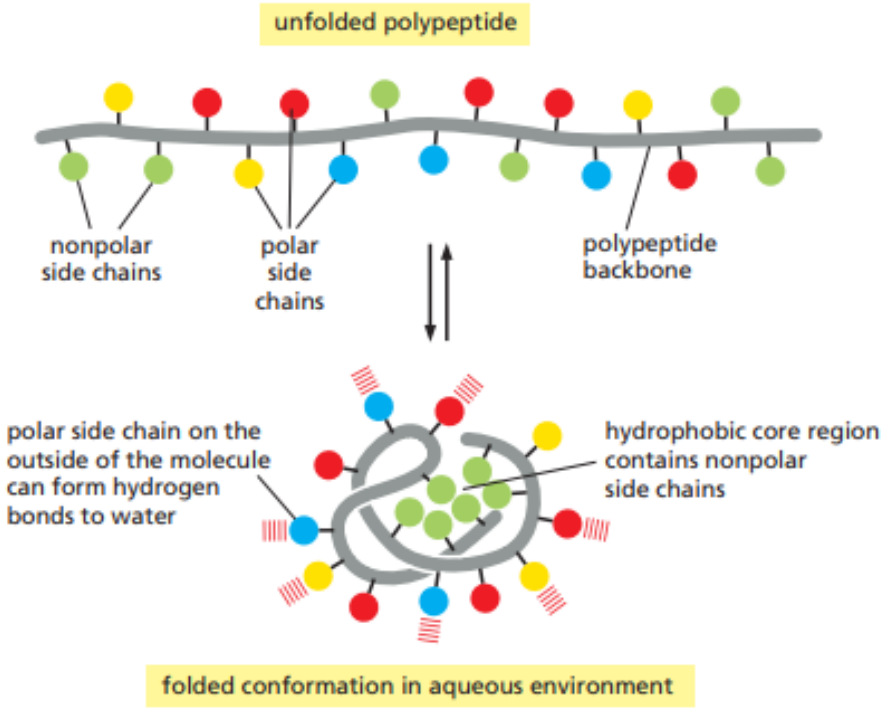




Non-Covalent Bonds Help Protein Fold

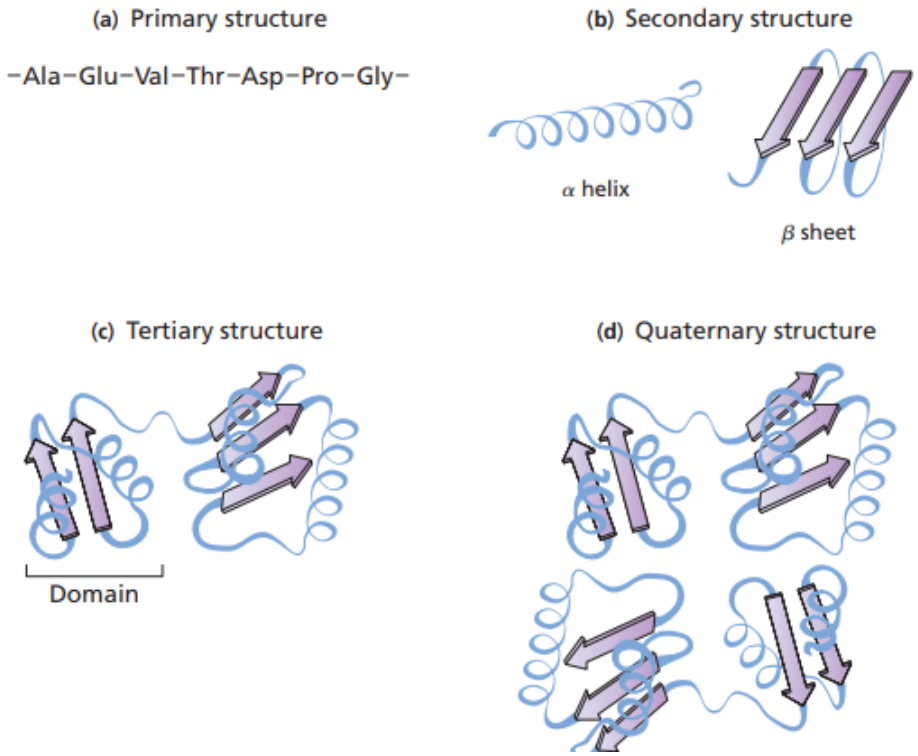
While a single noncovalent bond is relatively weak, multiple bonds working together create a strong and stable bonding network within the protein structure. These interactions collectively stabilize the folded conformation of proteins.





Levels of Protein Structure

- Protein structure is determined by both its amino acid sequence and noncovalent interactions (e.g., hydrogen bonds, van der Waals forces, electrostatic interactions).
- Although there are numerous possible folding patterns, only a limited number are biologically favorable due to energetic constraints and steric hindrance
- Proteins are organized into four hierarchical levels of structure:
  - Primary structure – The linear sequence of amino acids.
  - Secondary structure – Localized folding patterns (e.g., α-helices, β-sheets).
  - Tertiary structure – The three-dimensional shape of a single polypeptide.
  - Quaternary structure – The assembly of multiple polypeptide subunits into a functional protein



Levels of Protein Structure

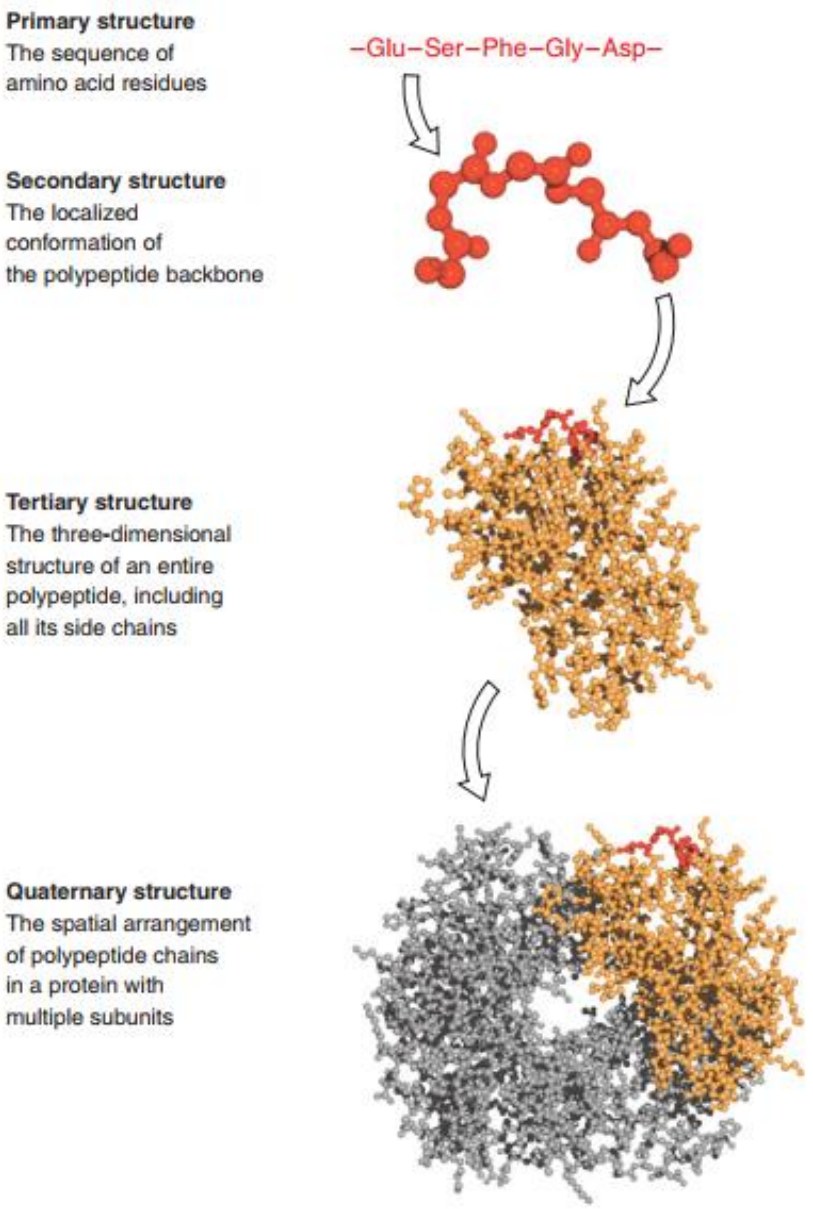
Primary structure (1^o structure): The Amino Acid Sequence. Linear chain of amino acids with peptide bonds highlighted. Basic blueprint of the protein. It is held together by covalent peptide bonds. Disulfide bond s (covalent) can also occur.

1° structure contains all information for protein to achieve its intricate architecture.

Secondary structure (2^o structure) : the local folding of the polypeptide backbone into regular patterns.. Noncovalent forces (hydrogen bonds, ionic, van der Waals, hydrophobic interactions) determine shape and stability. Noncovalent bonds are formed whenever possible within a given protein structure. Peptide backbone tends to form noncovalent bonds with one another.

Tertiary structure (3^o structure) : the overall 3D shape of a polypeptide, determined by interactions between the side chains.

Quaternary structure (3^o structure): the spatial arrangement of two or more polypeptide chains in a protein, held together by non-covalent interactions.

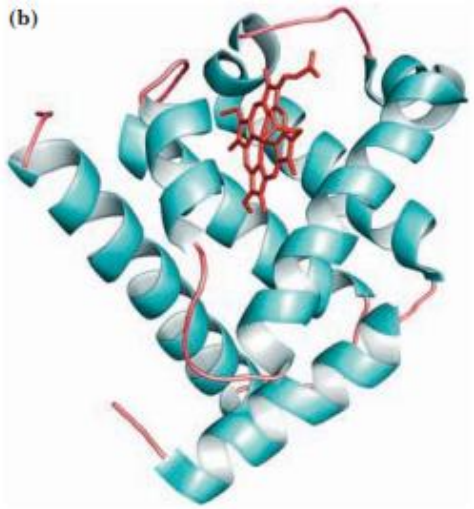


Classes According to Shape and Solubility

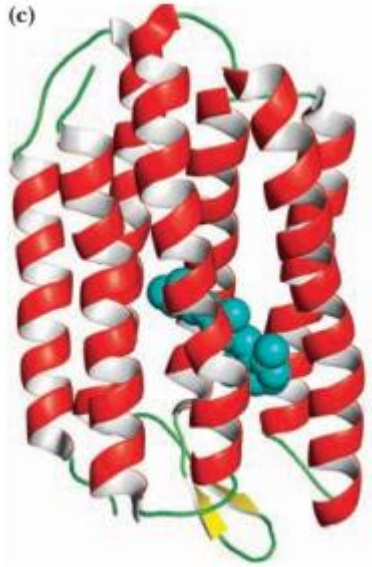
Proteins are categorized into three main types based on their shape and solubility: fibrous, globular, and membrane proteins.

- **Fibrous Proteins**:
  - Have elongated, linear, or helical structures.
  - Primarily serve structural roles (e.g., collagen, keratin).
  - Typically, insoluble in water or dilute salt solutions due to extensive intermolecular interactions.
- **Globular Proteins**
  - Compact, roughly spherical structures formed by folding polypeptide chains.
  - Hydrophobic side chains are buried inside, while hydrophilic side chains are exposed, enhancing solubility.
  - Generally soluble in aqueous solutions (e.g., cytosolic enzymes, hemoglobin).

Collagen,
a fibrous protein



Myoglobin, a globular protein

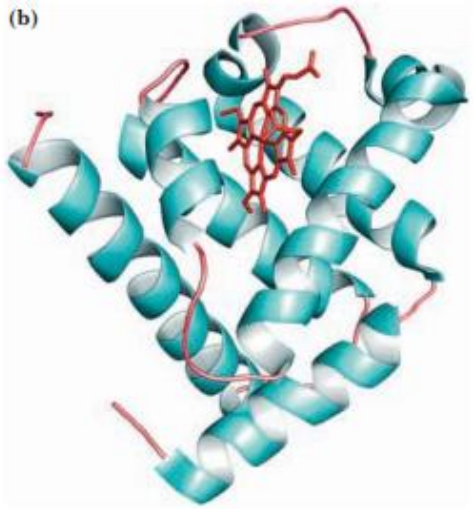


Bacteriorhodopsin, a membrane protein

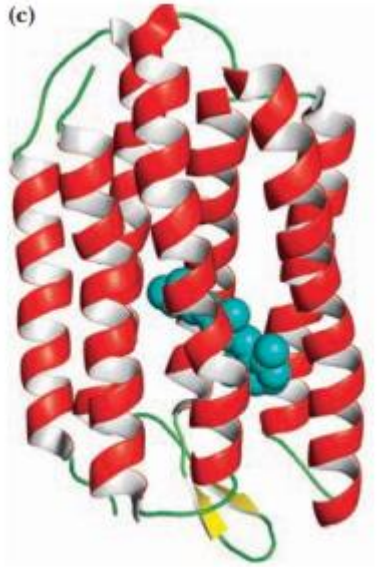
Classes According to Shape and Solubility

- **Membrane Proteins**
  - Embedded in or associated with cell membranes.
  - Contain hydrophobic side chains on their exterior to interact with the lipid bilayer.
  - Insoluble in water but can be solubilized in detergents.
  - Have fewer hydrophilic amino acids compared to cytosolic proteins.

Collagen,
a fibrous protein



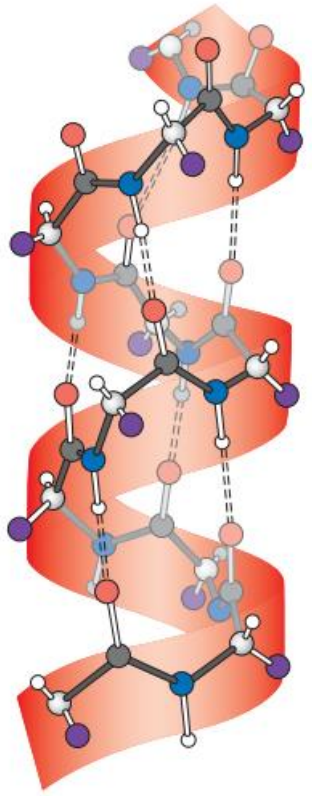
Myoglobin, a globular protein



Bacteriorhodopsin, a membrane protein

Structure of Protein

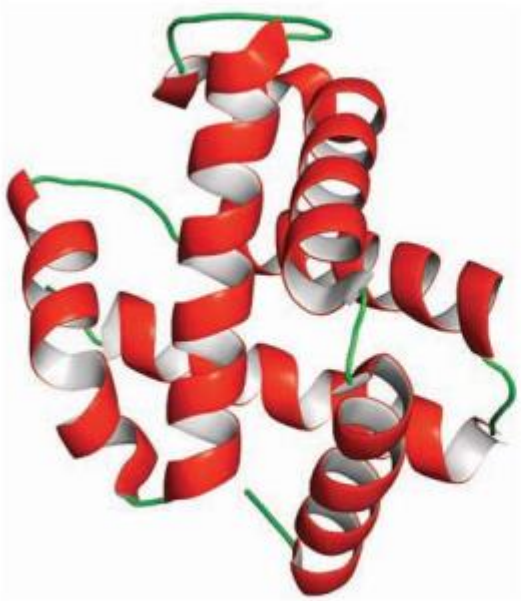
- William Astbury, while at the University of Leeds, pioneered research on protein secondary structure using X-ray diffraction, particularly in wool fibers.
- He observed structural differences between stretched and unstretched fibers:
  - Unstretched fibers had a helical structure, which he termed the alpha form.
  - Stretching caused the fibers to adopt an extended conformation, later called the beta form.
  - Astbury was the first to propose that hydrogen bonds play a crucial role in stabilizing protein structures.
- His work influenced Linus Pauling and Robert Corey, who further explored the role of hydrogen bonding in stabilizing α-helices and β-sheets, aligning C=O and N-H groups in peptide bonds.



α helix. the polypeptide backbone twists in a right-handed fashion so that hbonds form between C=O and N-H groups four residues farther along. Atoms are color-coded: Cα light gray, carbonyl C dark gray, O red, N blue, side chain purple, H white.

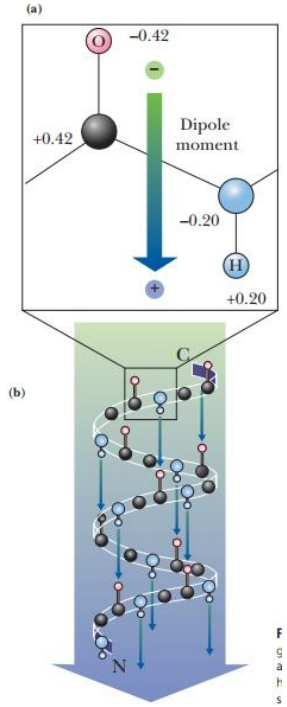
Structure of Protein

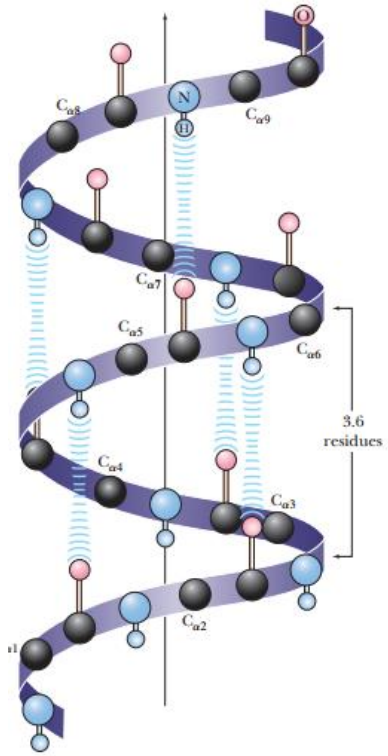
Experiments in model-building have demonstrated that both L- and D-amino acids can contribute to the formation of an α helix in polypeptides. However, it's crucial that all residues belong to the same stereoisomeric series; the presence of a D-amino acid within a sequence of L-amino acids, or vice versa, disrupts the regular structure. While naturally occurring L-amino acids can assemble into α helices with either right- or left-handed twists, extended left-handed helices have not been observed in proteins



β-Hemoglobin subunit

Myohemerythrin





Structure of Protein

The way different amino acids affect the stability of α helices varies between typical proteins and polyamino acids. A summary of how common amino acids occur in helices is provided in the table. Proline (and hydroxyproline) are particularly noteworthy as they disrupt helices due to their unique structure, which locks the value of the $C_\alpha - \mathrm{N} - \mathrm{C}$ bond angle. Helices can consist of either D- or L-amino acids, but each helix must be made entirely of amino acids of the same configuration. **It's not possible to form an α-helix from a mix of D- and L-amino acids**. If an α-helix is composed of D-amino acids, it is left-handed.

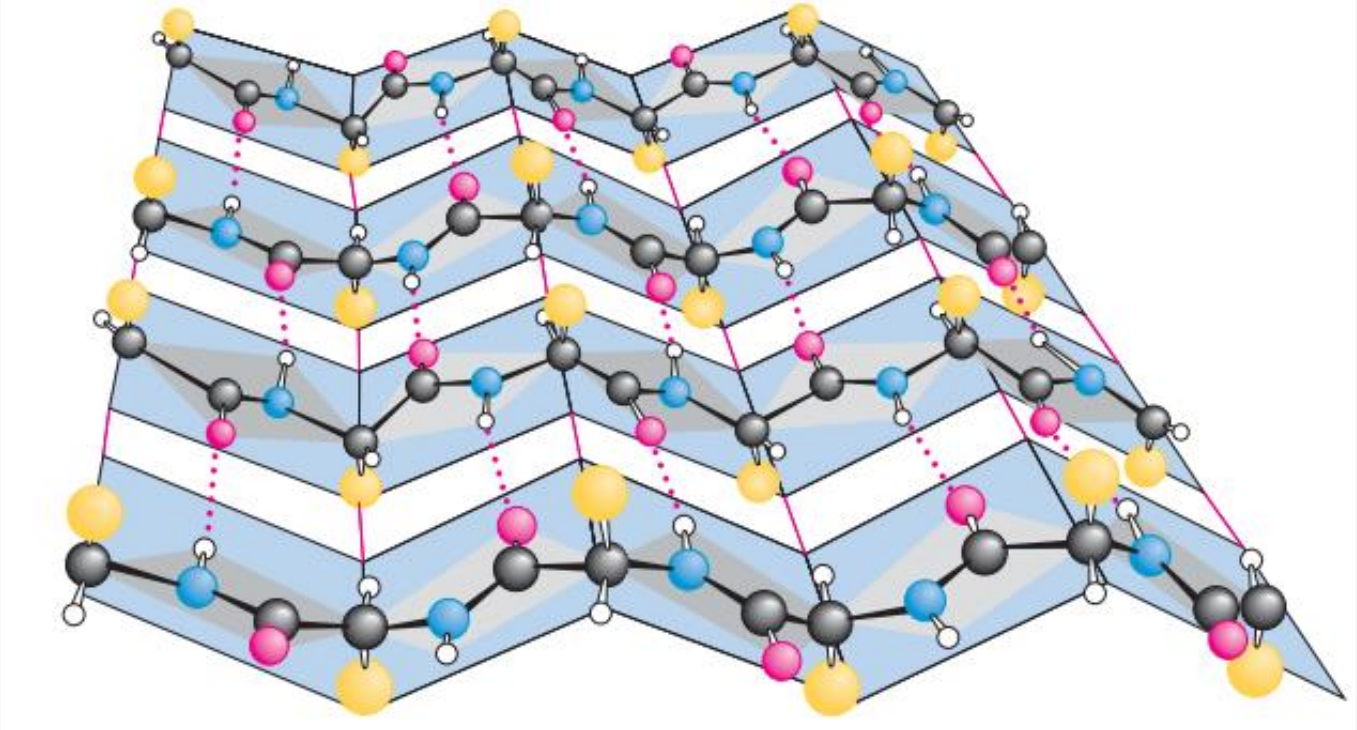
TABLE 6.1 Helix-Forming and Helix-Breaking Behavior of the Amino Acids

Amino Acid		Helix Behavior*	
A	Ala	H	(I)
C	Cys	Variable	
D	Asp	Variable	
E	Glu	H	
F	Phe	H	
G	Gly	I	(B)
H	His	H	(I)
I	Ile	H	(C)
K	Lys	Variable	
L	Leu	H	
M	Met	H	
N	Asn	C	(I)
P	Pro	B	
Q	Gln	H	(I)
R	Arg	H	(I)
S	Ser	C	(B)
T	Thr	Variable	
V	Val	Variable	
W	Trp	H	(C)
Y	Tyr	H	(C)

*H = helix former; I = indifferent; B = helix breaker; C = random coil; ( ) = secondary tendency.

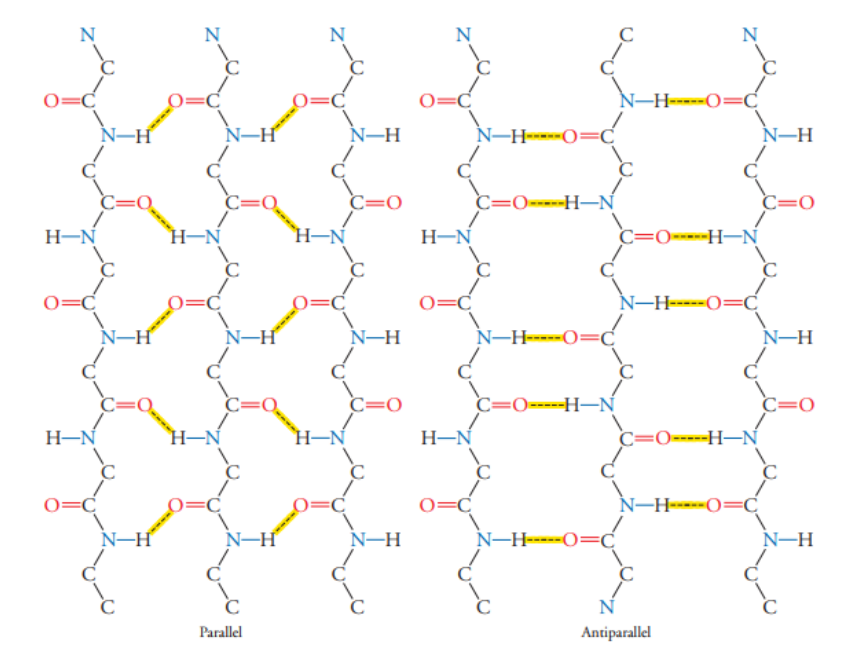
β Sheet

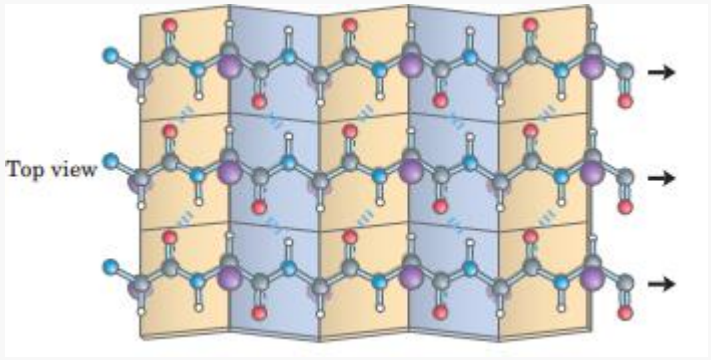
β sheet, a core structure in proteins, is a secondary structure with several polypeptide strands. Pauling and Corey postulated this structure in 1951, which has polypeptide strands that line up and form hydrogen bonds with each other. A β-pleated sheet can be imagined as thin strips of paper laid side by side in a pleated manner, forming a sheet with a characteristic fold.

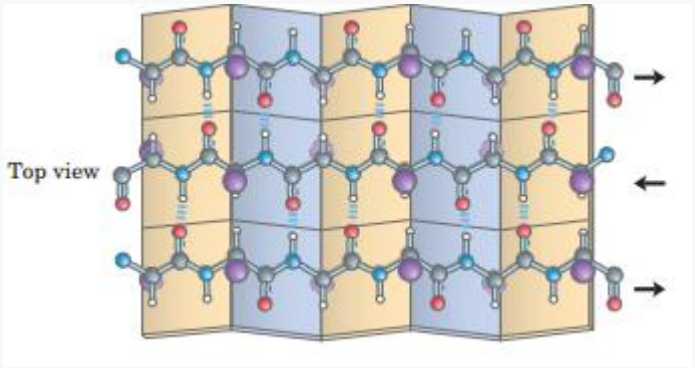


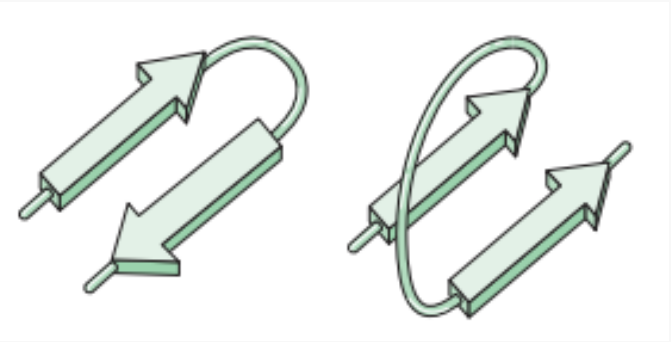
β Sheet

The strands in a β sheet can be **parallel or antiparallel**, depending on whether they go the same way or opposite ways. β turns are common in proteins. Parallel β-sheet tend to be more regular than antiparallel. The range of ϕ and ψ angles for the peptide bonds in parallel sheets is much smaller (typical $\phi = -120^o$ and $\psi = 105^o$) than that for antiparallel sheets (typical $\phi = -135^o$ and $\psi = 140^o$) . Parallel sheets are typically large structures, compose of less than 5 strands are rare, and Antiparallel sheets may consist of as few as two strands.



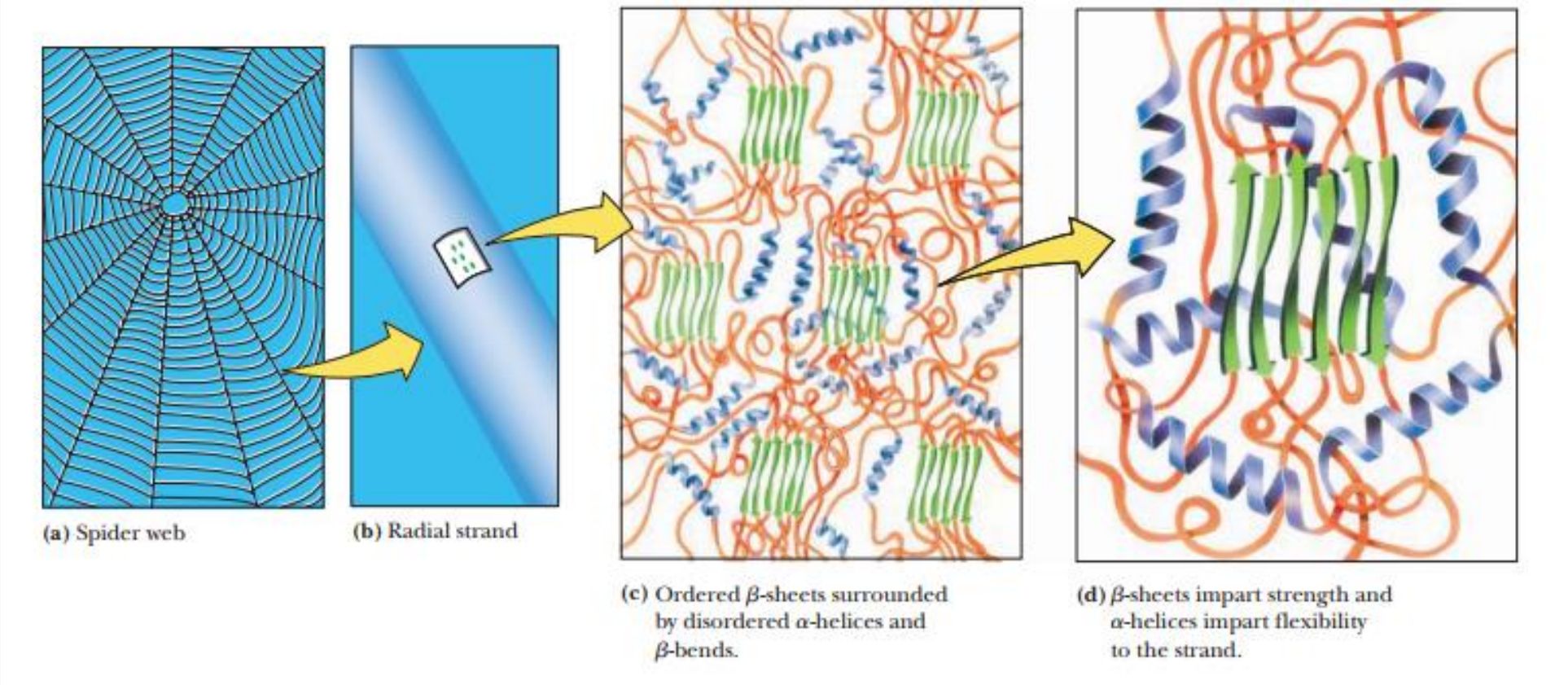






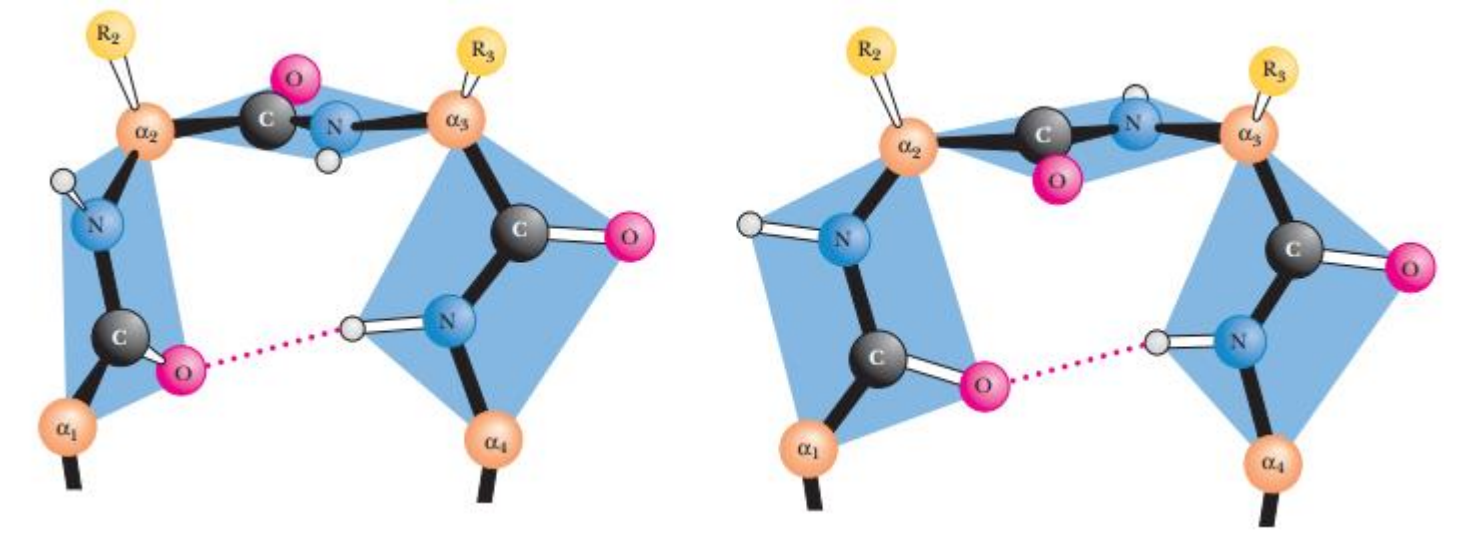
β Helix-Sheet Composites in Spider Silk

Spider web silks are composites of α-helices and β-sheets. The radial strands of webs must be strong and rigid . silks in radial strands contain a higher percentage of β-sheets



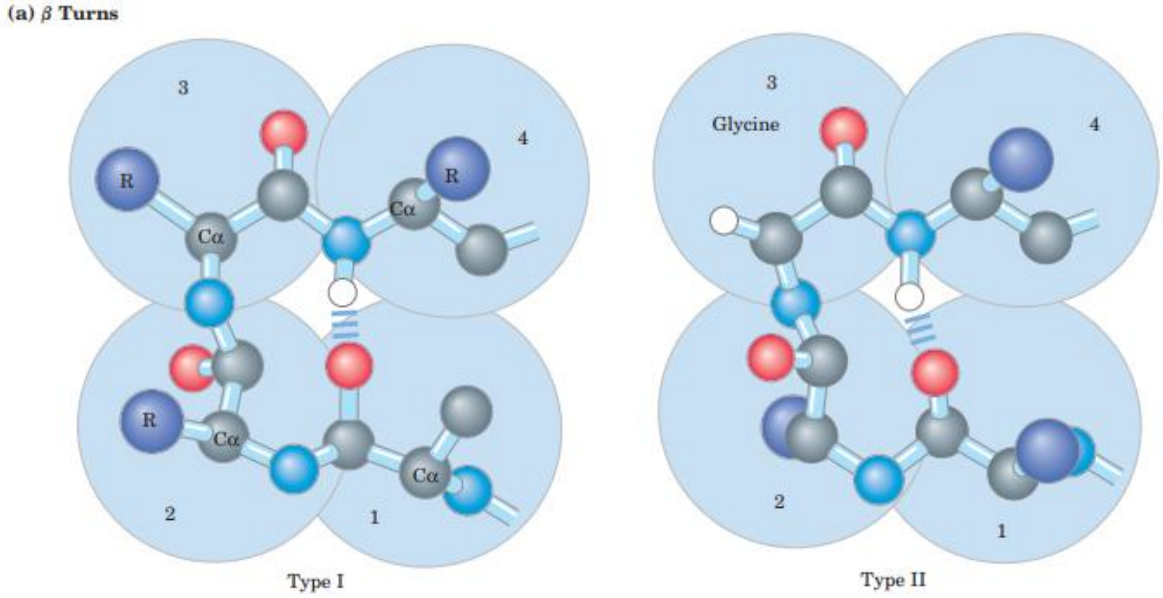
β Sheet

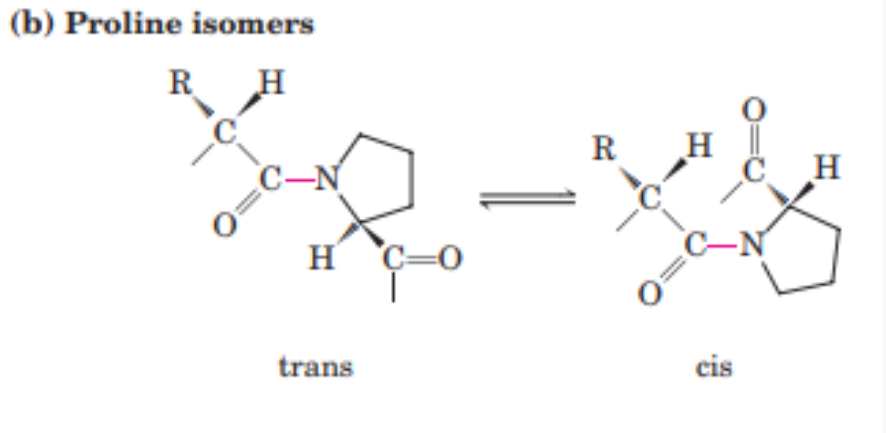
The majority of proteins adopt globular shapes. Consequently, the polypeptide chain needs to be flexible enough to bend, turn, and adjust its orientation to achieve the compact, globular structures required. A common structural feature found in many proteins is the β-turn, also called the tight turn or β-bend. In this arrangement, the peptide chain forms a tight loop where the carbonyl oxygen of one residue is hydrogen bonded with the amide proton of the residue three positions down the chain. Tow major type of β turn (type I: common and Type II-less common)



β Sheet

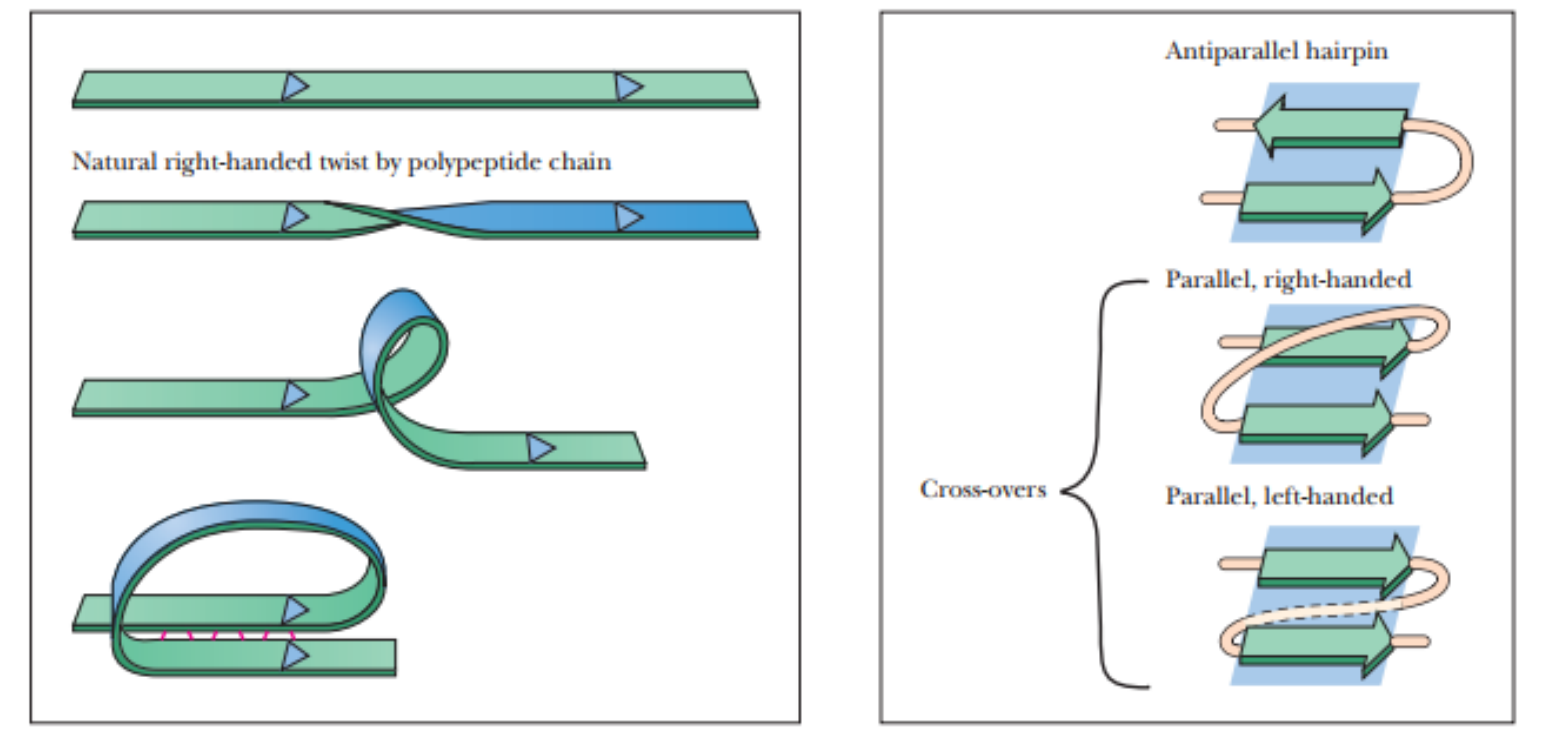
The most common β turns are type I and type II, with type I being more than twice as common as type II. Gly is always the third residue in type II β turns. The first and fourth residues of the turns form a hydrogen bond between their peptide groups (b) A peptide bond with proline's imino nitrogen can be trans or cis. Most peptide bonds (more than 99.95%) are trans, except for about 6% of those with proline's imino nitrogen. These are often cis and found in β turns





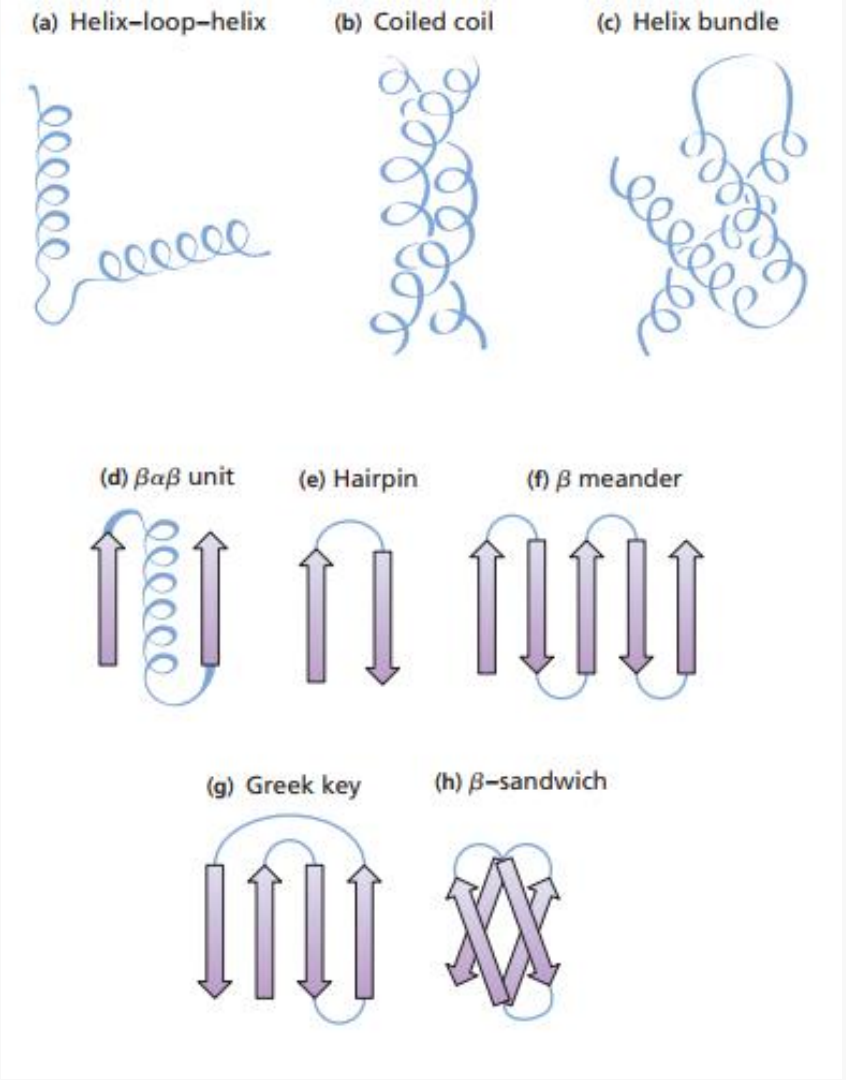
β Sheet

The natural right-handed twist exhibited by polypeptide chains, and the types of connections between β-strands



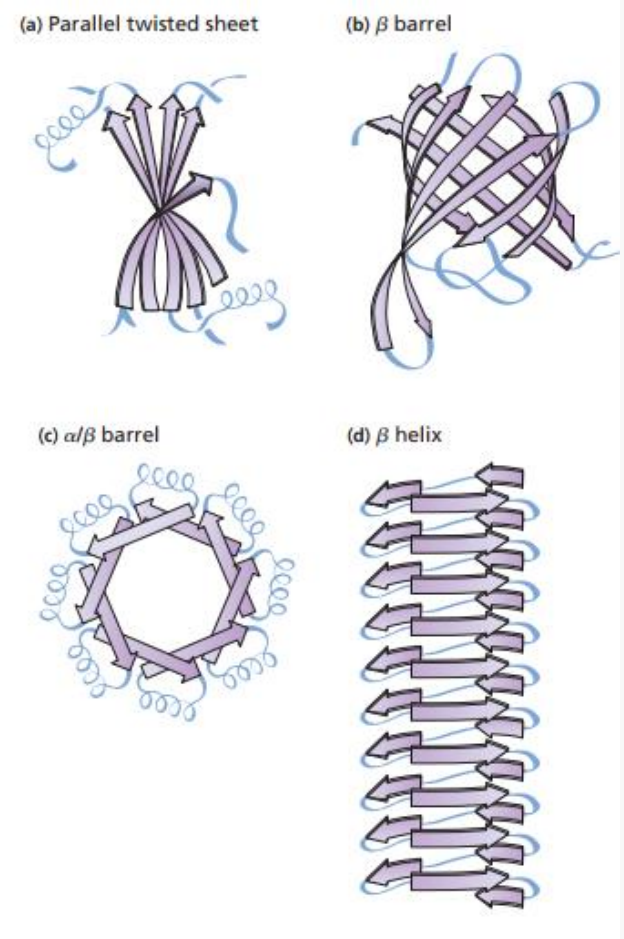
Super-secondary Structures (motifs)

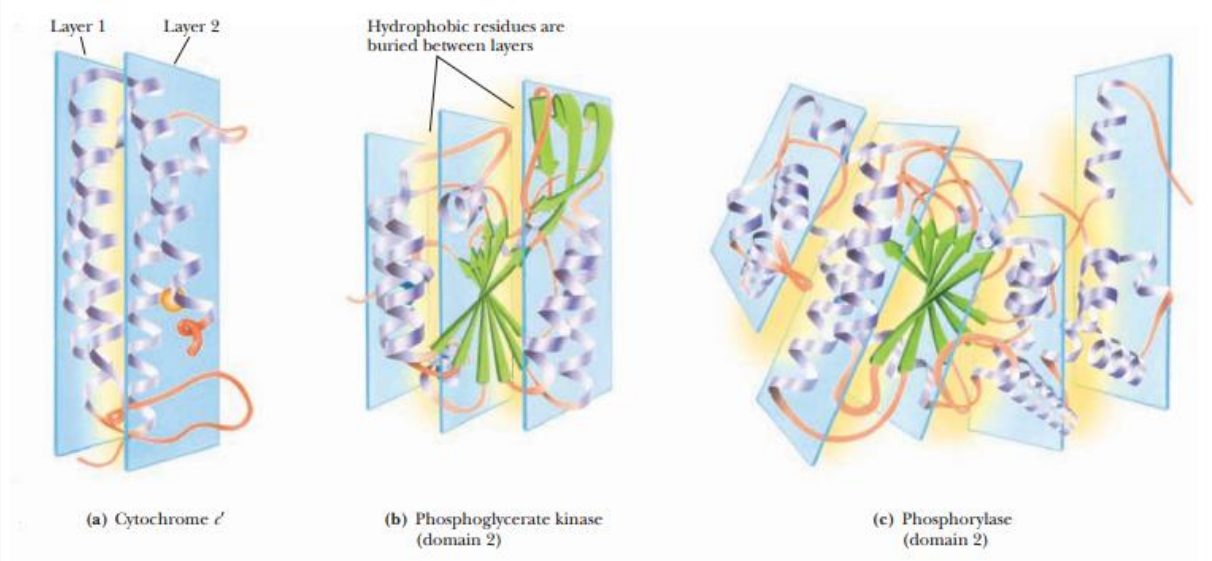
Supersecondary structures, also known as motifs, are identifiable arrangements of α helices, β strands, and loops found in various proteins. These motifs can be linked to specific functions, although similar structures may serve different functions in different proteins. Below are some examples of commonly observed motifs.

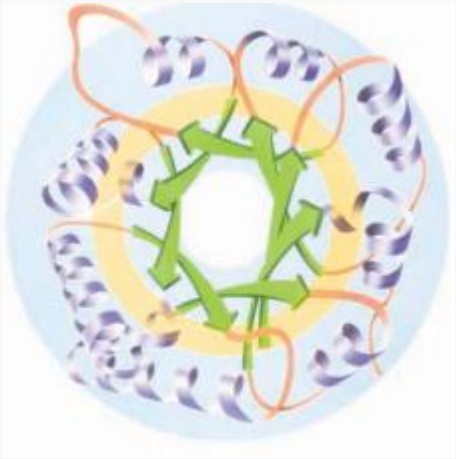


Domain Folds

Common domain Folds

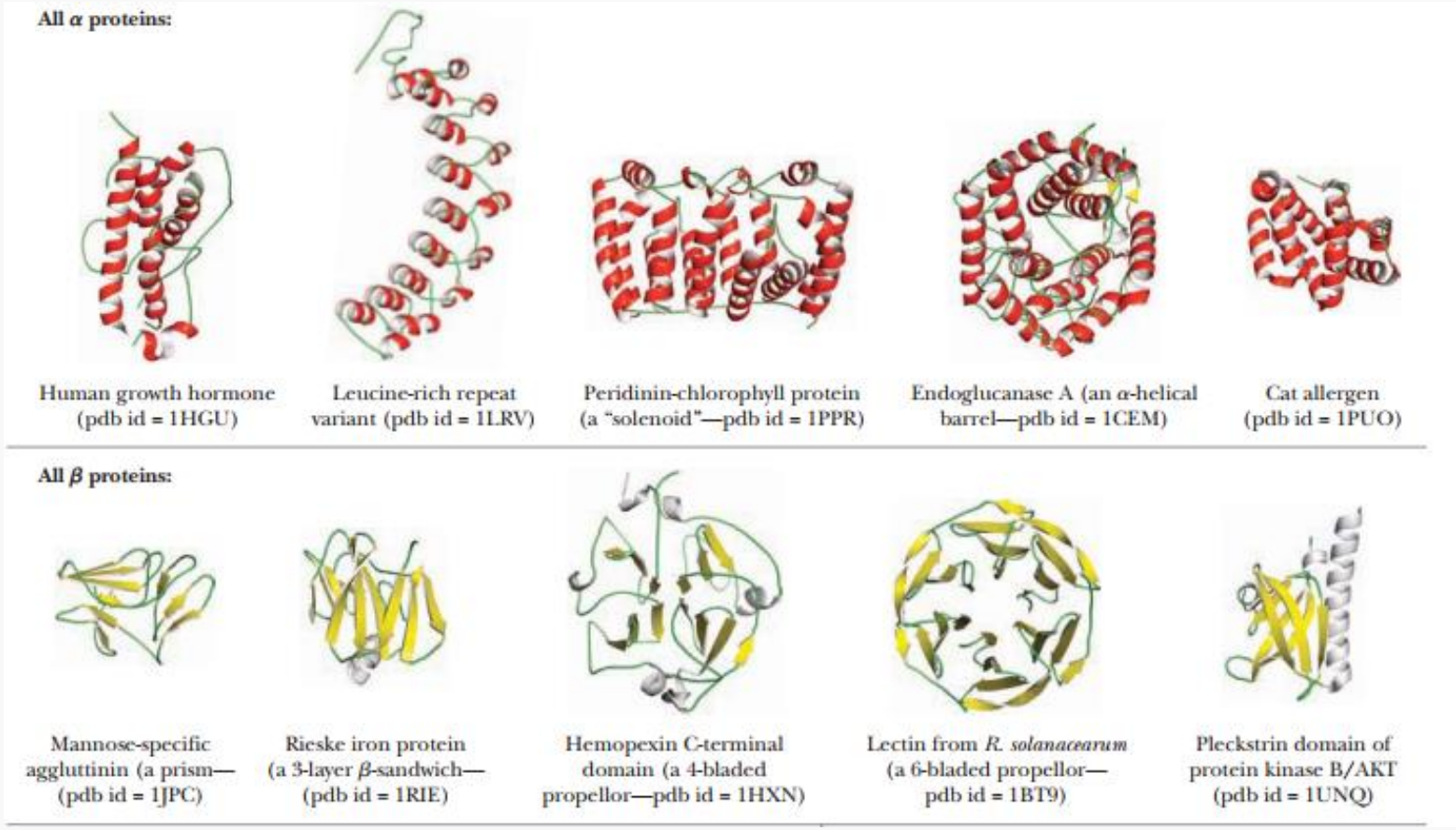




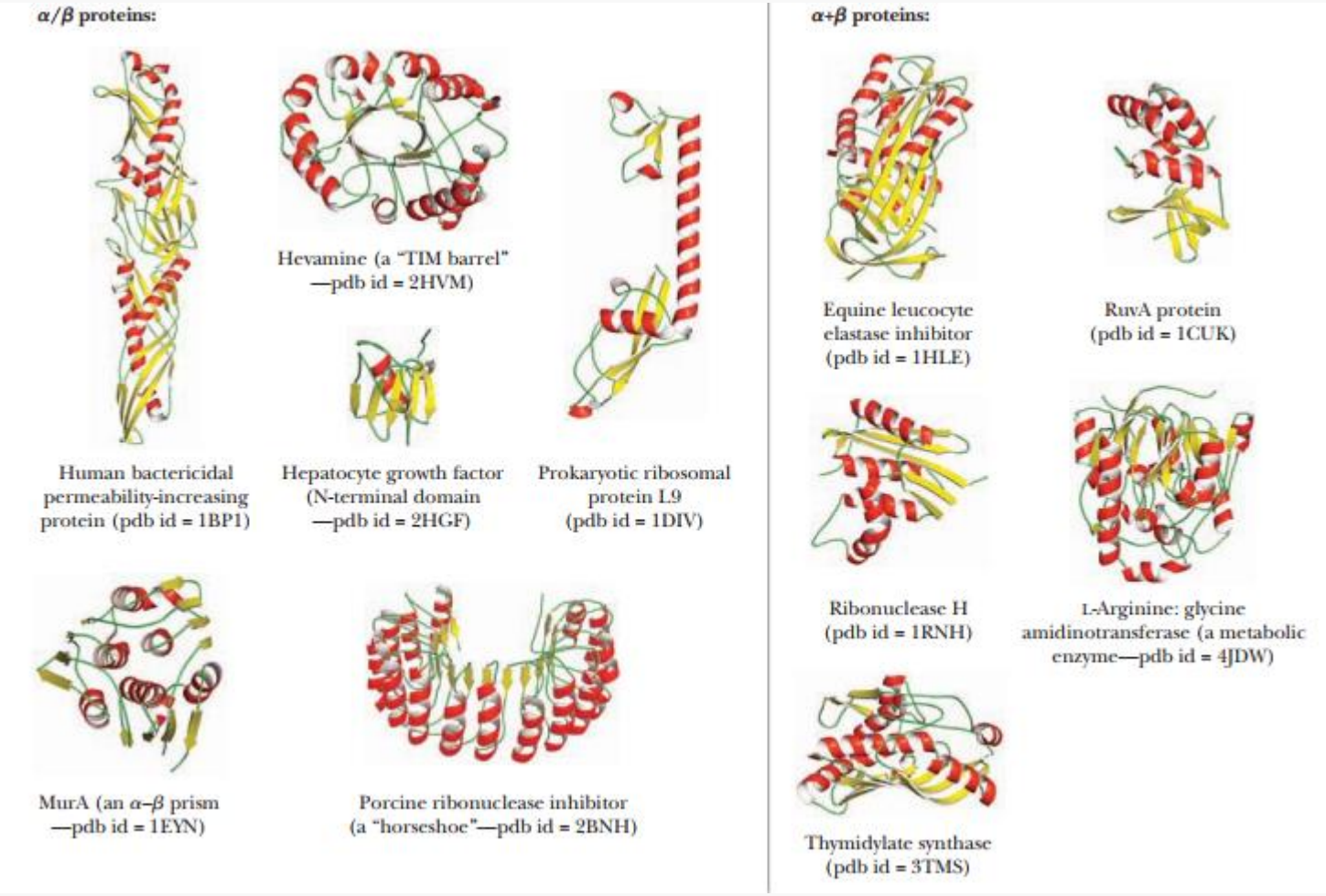


Four major Class of Protein

all proteins and all proteins (in which the structures are dominated by α −helices and β −sheets, respectively), α/β proteins (in which helices and sheets are intermingled), and $\alpha + \beta$ proteins (in which α −helical and β −sheet domains are separated for the most part)

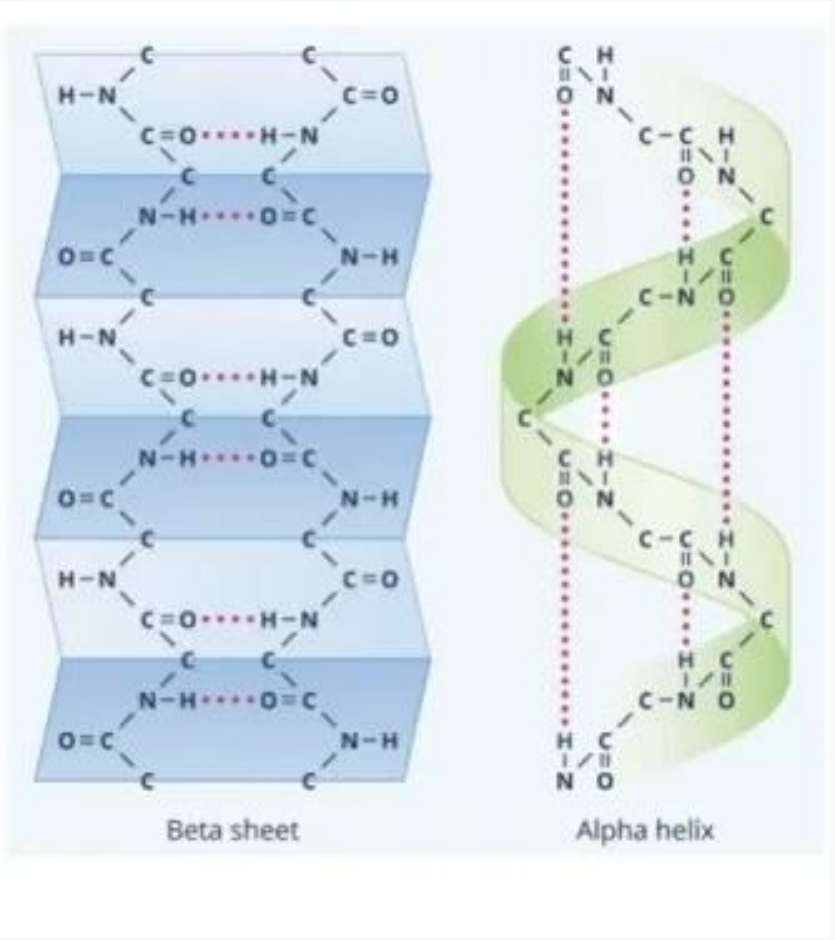


Four major Class of Protein



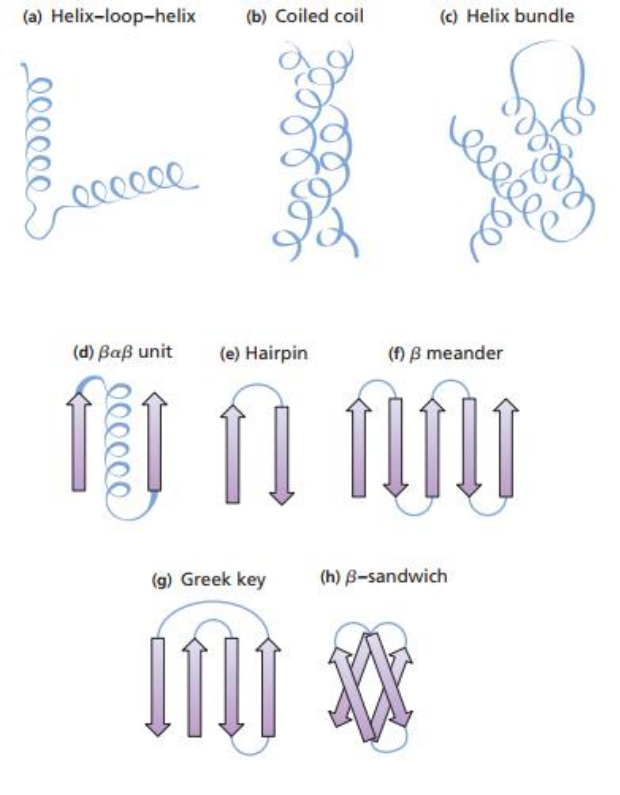
Recapitulation of Last Class: Secondary Structure

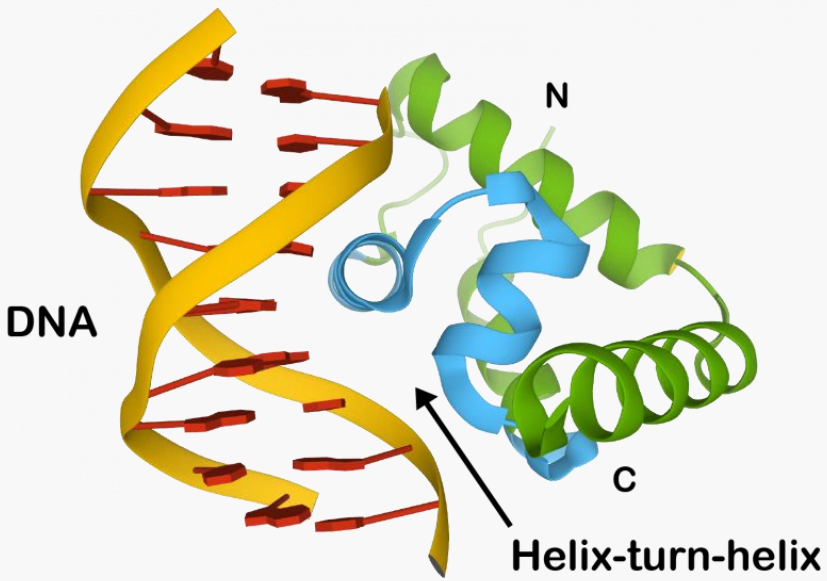
- It (polypeptides chain) consists of several repeating patterns, mostly observer types are α-helix and β-sheet.
- **Stability**: Both (α-helix and β-sheet) are stabilized by hydrogen bonds between backbone atoms.
- **Flexibility**: Peptide bonds are rigid, limiting rotation. Rotation around the α-carbon provides chain flexibility. The rotation around the alpha carbon is defined by the phi (ϕ) and psi (ψ) angles.
- **Amino acid influence**: Side chain (R group) properties (size, charge) influence folding angles (φ and ψ). These angles determine which secondary structures form. Certain amino acids have preferences for forming specific secondary structures

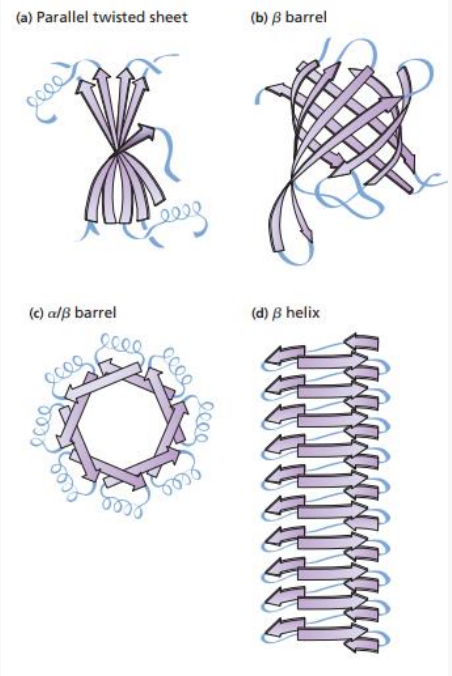


Super-secondary Structures (Motifs)

Super-secondary structures (motifs) are specific combinations of secondary structure elements (α-helices and β-sheets). Motifs often play a role in protein function and folding. They can serve as building blocks for larger protein domains. These are recognizable 3D structures that appear in a variety of proteins.



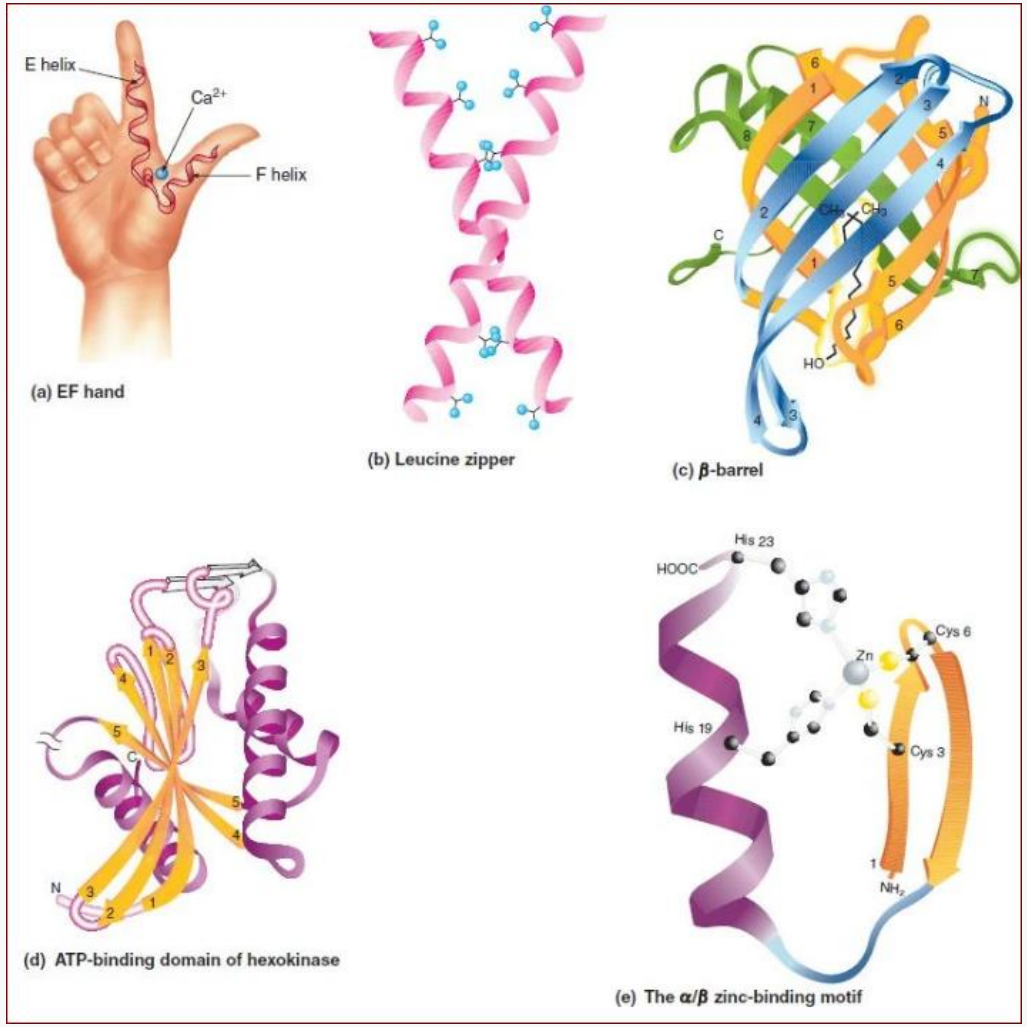




Common domain Folds

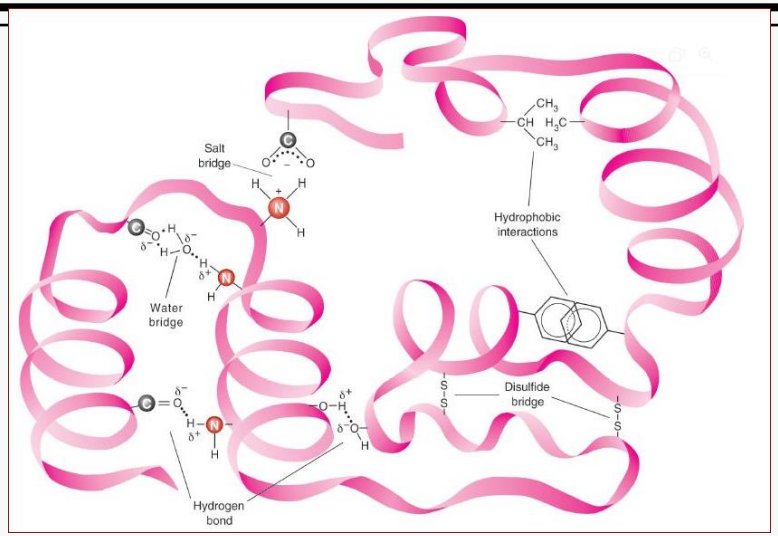
Tertiary Structure

- Tertiary structure is the unique 3D shape of a single polypeptide chain.
- This 3D shape is essential for protein biological function.
- Tertiary Structure is also known as protein folding.
- The 3D shape is determined by interactions between amino acid side chains (R groups). Key interactions include:
  - Hydrogen bonds
  - Hydrophobic interactions
  - Disulfide bonds (covalent).
  - Electrostatic (ionic) interactions



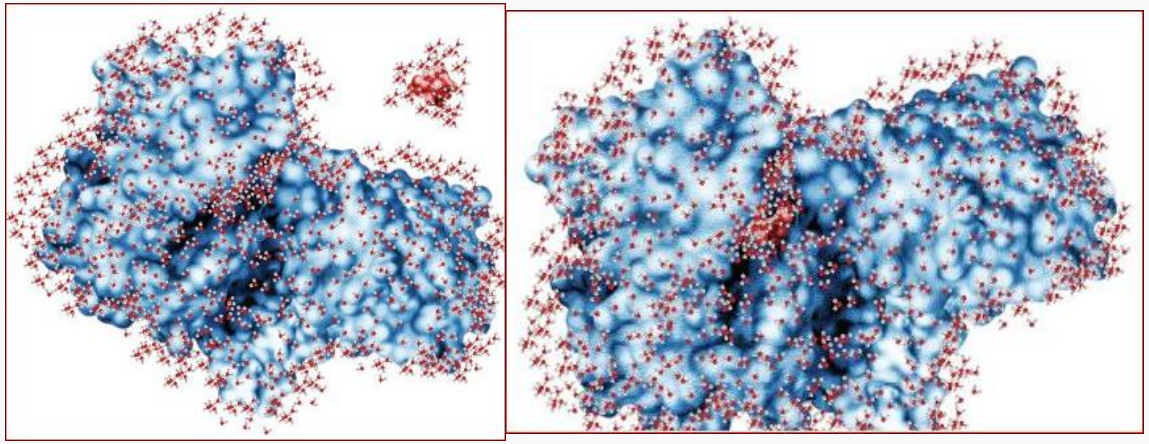
Tertiary Structure

Top Figure and Hydration (bottom) stabilize tertiary structure. The specific forces that drive protein folding are not fully understood, but it is known to be a thermodynamically favorable process with an overall decrease in energy.



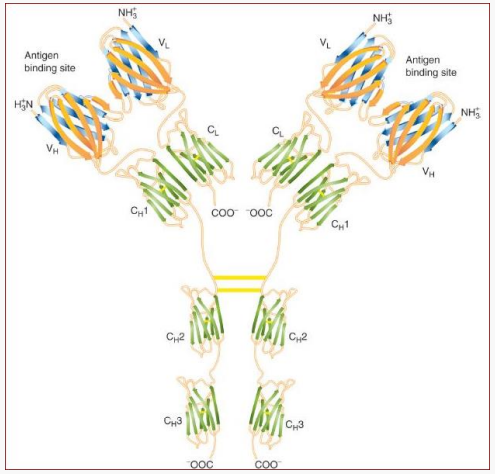
$$\Delta G^{o} = \Delta H^{o} - T\Delta S^{o}$$

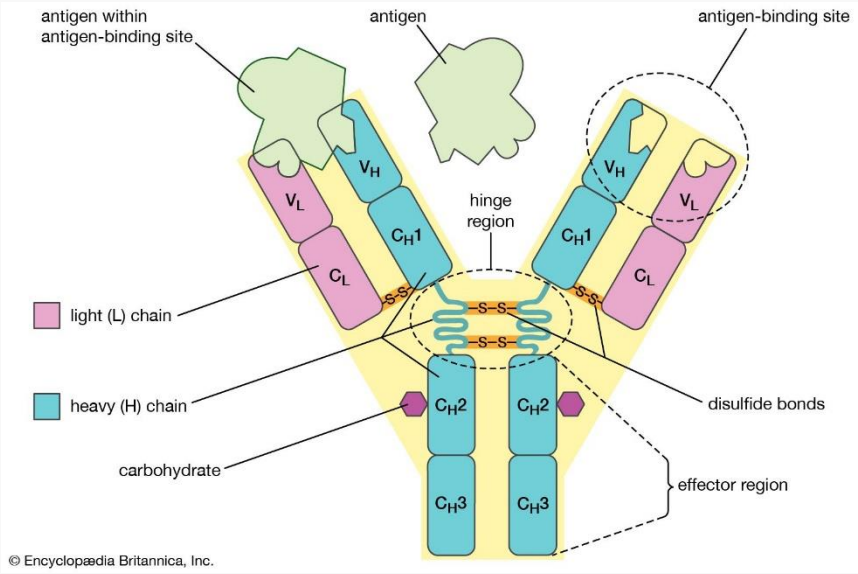
ΔG is change in Gibbs free energy, ΔH is the change in enthalpy, and ΔS is change in entropy.



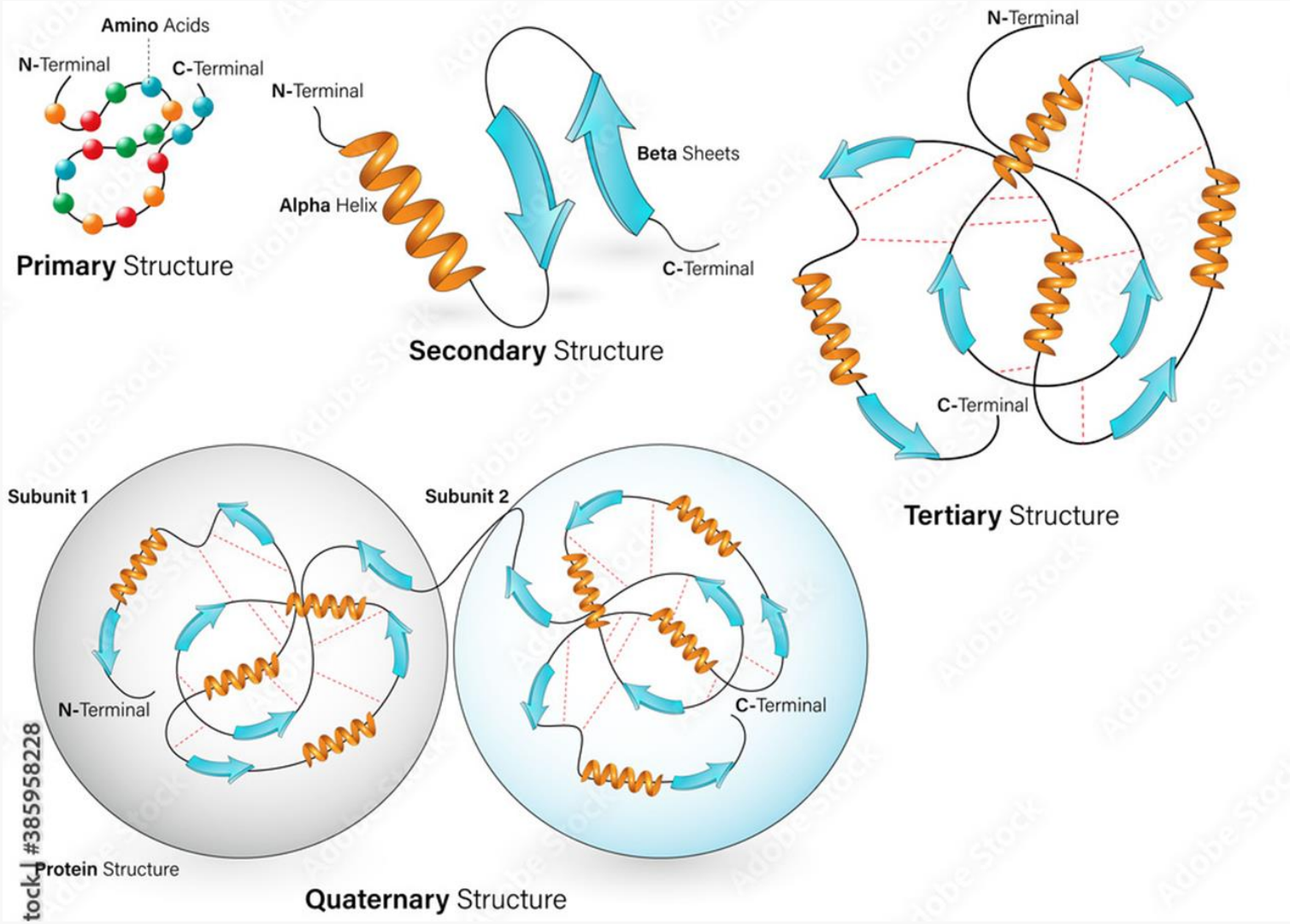
Oligomeric Proteins: Structure & Classification

- Many proteins, especially those with high molecular weights, consist of multiple **polypeptide chains** (subunits).
- Subunits can be:
  - **Homomeric** – Identical subunits (e.g. Actin)
  - **Heteromeric** – Distinct subunits (e.g Antibody, hemoglobin)
- These multi-subunit proteins are collectively called **oligomers**, which are composed of a small number of repeating **monomer** units.
- **Oligomers vs. Polymers**: Oligomers are generally smaller than polymers.
- The smallest functional unit within an oligomer is the **protomer**, which may consist of one or more subunits.





Protein Structure



Intrinsically Disordered Proteins (IDPs) & Regions (IDRs)

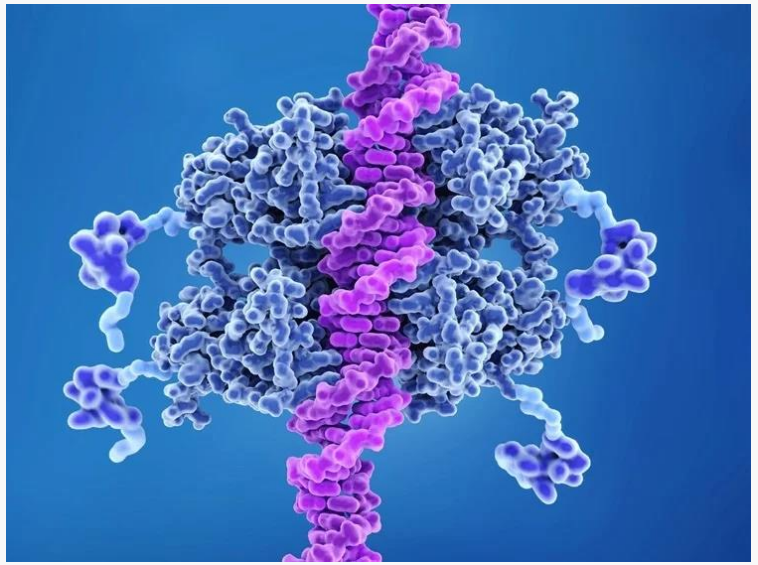
- Not all proteins have a fixed 3D structure—some are partially or completely unstructured.
- **Intrinsically Disordered Proteins (IDPs**) lack a rigid structure due to their unique amino acid composition:
  - **Low in hydrophobic amino acids** (Leu, Val, Phe, Trp) that usually stabilize folded structures.
  - **Rich in charged & polar amino acids** (Ser, Gln, Lys, Glu) that favor interaction with water, preventing rigid folding.

- Intrinsically Disordered Regions (IDRs): Unstructured segments within otherwise structured proteins

Intrinsically Disordered Proteins (IDPs) & Regions (IDRs)

p53, a crucial tumor suppressor protein, is a remarkable example of the utility of intrinsically disordered protein (IDP) domains in biological functions.

- Mutations in p53 (especially in its DNA-binding domain) are found in over 50% of human cancers, highlighting its role as a critical tumor suppressor.
- Targeting the disordered regions of p53 is an emerging strategy in cancer therapy.



Why is p53's Unstructured Domain Important?

1. Intrinsically Disordered Regions (IDRs) Enable Versatile Interactions.
   - The **N-terminal transactivation domain (TAD) and C-terminal regulatory domain** of p53 are intrinsically disordered.
   - This allows **dynamic binding** to multiple proteins, including transcription factors, regulatory proteins, and ubiquitin ligases.

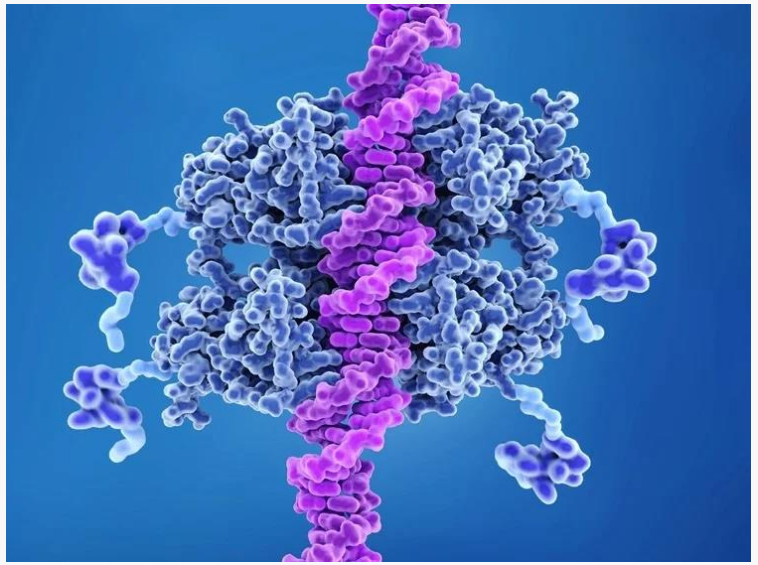
Intrinsically Disordered Proteins (IDPs) & Regions (IDRs)

2. Adaptive Functionality in Cellular Signaling

- Unlike rigidly structured proteins, p53 can adopt multiple conformations, enabling it to interact with over 100 different partners.
- This flexibility is crucial for responding to cellular stress, DNA damage, and oncogenic signals.

3. Facilitates Rapid Degradation and Recycling

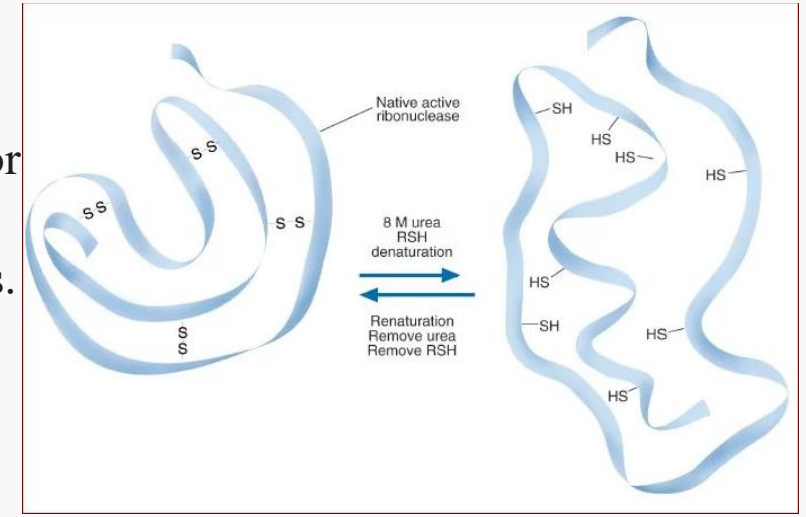
- The flexible C-terminal domain helps p53 be quickly degraded when not needed, preventing unnecessary cell cycle arrest or apoptosis



Protein Structure & Denaturation

- Proteins are highly sensitive to environmental factors because the free energy difference between folded and unfolded states is small.

- What is Denaturation?
  - Loss of native conformation without necessarily unfolding completely.
  - No peptide bond breaking, but the protein may lose (partial or complete) its biological function.
  - Can be reversible or irreversible, depending on the conditions.



- Effects of Denaturation
  - Loss of biological activity (e.g., enzymes losing catalytic function).
  - Changes in physical properties, such as solubility and transparency.
  - Example: Cooking eggs
    - Heat causes egg albumin to become insoluble & opaque—an irreversible process.

Protein Stability and Denaturation

- **Strong Acids/Bases**:
  - Alter protonation states of amino acid side chains.
  - Disrupt hydrogen bonding & salt bridges, destabilizing the protein.
  - Near the isoelectric point (pI), proteins become less soluble and may precipitate.
- **Organic Solvents (e.g., Ethanol, Acetone)**:
  - Interfere with hydrophobic interactions, causing protein unfolding.
  - Nonpolar solvents can disrupt protein-lipid interactions in membrane proteins.
- **Detergents (e.g., SDS, Triton X-100)**
  - Amphipathic molecules that integrate into proteins, disrupting hydrophobic interactions
  - .Commonly used to denature proteins in electrophoresis (e.g., SDS-PAGE).
- **Reducing Agents (e.g., DTT, β-mercaptoethanol)**
  - Break disulfide bridges, converting them into free sulfhydryl (-SH) groups.
  - Urea & Guanidine Hydrochloride further disrupt hydrogen bonds & hydrophobic interactions.

Protein Stability and Denaturation

- **High Salt Concentration ("Salting Out")**
  - Competes for water molecules, reducing protein solubility and causing aggregation/precipitation.
  - Reversible process, commonly used in protein purification (e.g., ammonium sulfate precipitation).

- **Heavy Metal Ions** (such as Hg^{2+} Pb^{2+})**:** \
  - Disrupt salt bridges and bind to sulfhydryl (-SH) groups, leading to protein misfolding.
  - Toxic effects: Can interfere with essential enzymes (e.g., lead inhibits heme synthesis, contributing to anemia).

- **Temperature Changes**:
  - Higher temperatures increase molecular vibrations, disrupting weak interactions (hydrogen bonds, van der Waals forces).
  - Leads to protein unfolding (denaturation), with some proteins more heat-resistant than others (useful in protein purification).
  - Example: Heat denaturation in cooking (e.g., egg proteins solidifying).

Denaturing Conditions

- **Mechanical Stress**:
  - Physical forces like stirring, grinding, or shearing can break non-covalent interactions.
  - Example: Whipped egg whites—proteins unfold and aggregate, stabilizing the foam.

Protein Folding and Problems

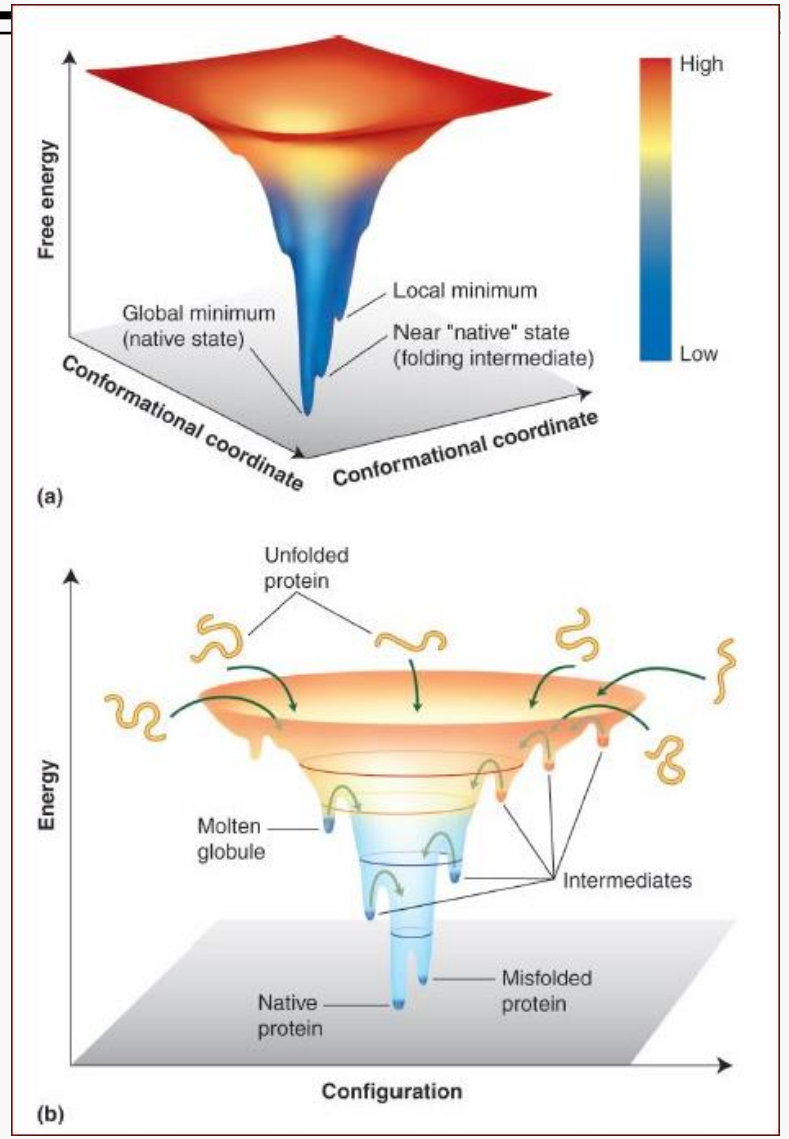
Christian Anfinsen's experiment: Suggested proteins fold based solely on their amino acid sequence. Direct relationship between protein's primary sequence and its final 3D conformation.

Folding process: Stepwise with early secondary structure formation.

Hydrophobic interactions: Major driving force for folding.

Surface vs. core amino acids: Surface changes often have minimal impact on structure. Core changes can lead to significant structural alterations.

Problem is not completely resolved despite advances using X-ray, NMR, mutagenesis, and modeling. hallenge remains: Predicting structure from sequence is not fully solved.



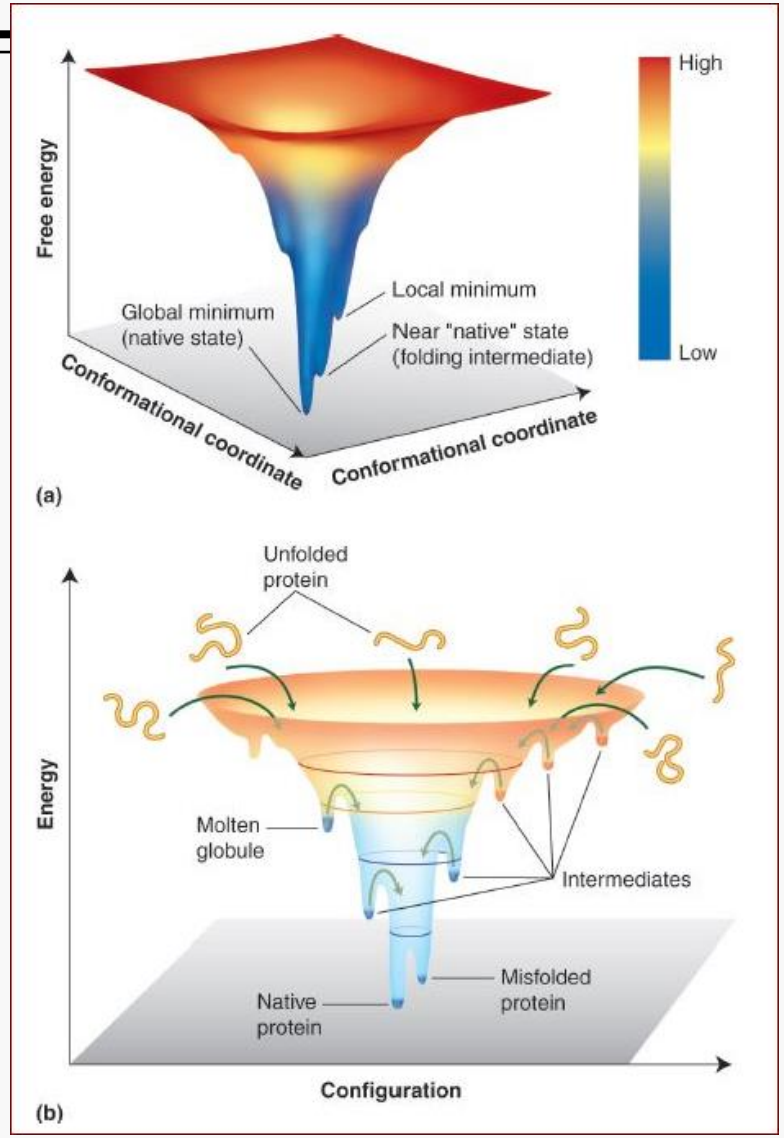
Protein Folding and Problems

Folding Routes: Numerous paths exist for each polypeptide to reach its native state.

Energy Landscape: Folding depicted as a funnel, with polypeptide navigating to lowest energy (native) state.

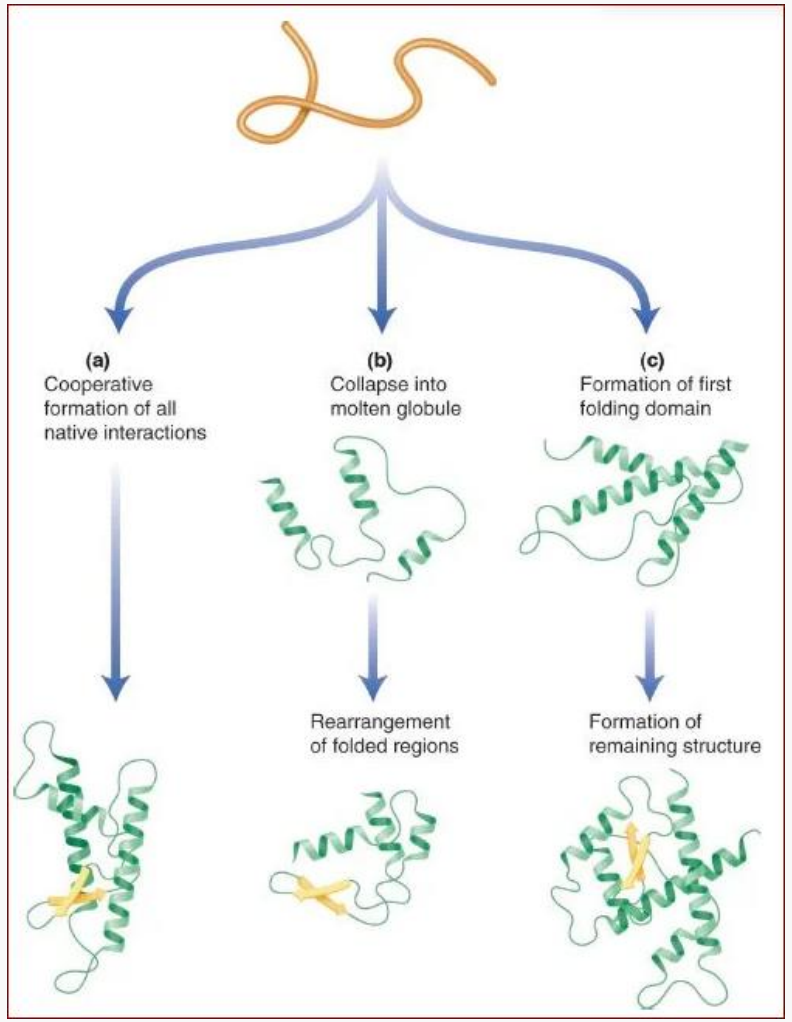
Intermediate States: Small proteins often fold directly without intermediates. Larger proteins frequently form intermediates: (i) **Local Energy Wells**: Briefly trap the polypeptide during folding. (ii) **Molten Globule**: Partially folded intermediate resembling the native state, with fluctuating internal interactions.

Determinants of Folding: (i) **Size**: Smaller proteins fold faster and with fewer intermediates. (ii) **Constraints**: Amino acid sequence, posttranslational modifications, and cellular environment (temperature, pH, crowding) influence folding



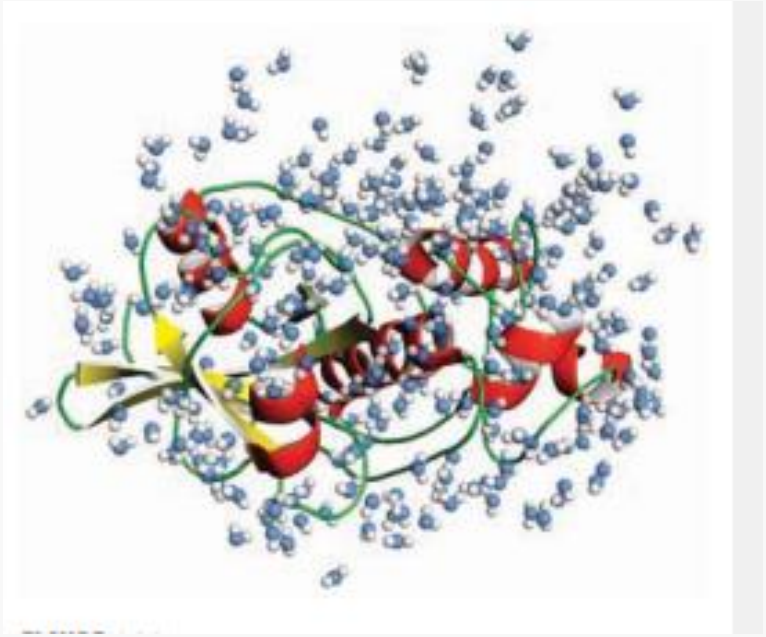
Protein Folding and Problems

- **Small proteins**, folding is cooperative with no intermediates formed.
- **Some larger proteins**, folding involves the initial formation of a molten globule followed by rearrangement into the native conformation.
- **Large multi-domain proteins:** follow a more complex pathway, with each domain folding separately before its native conformation.



Structure of Protein

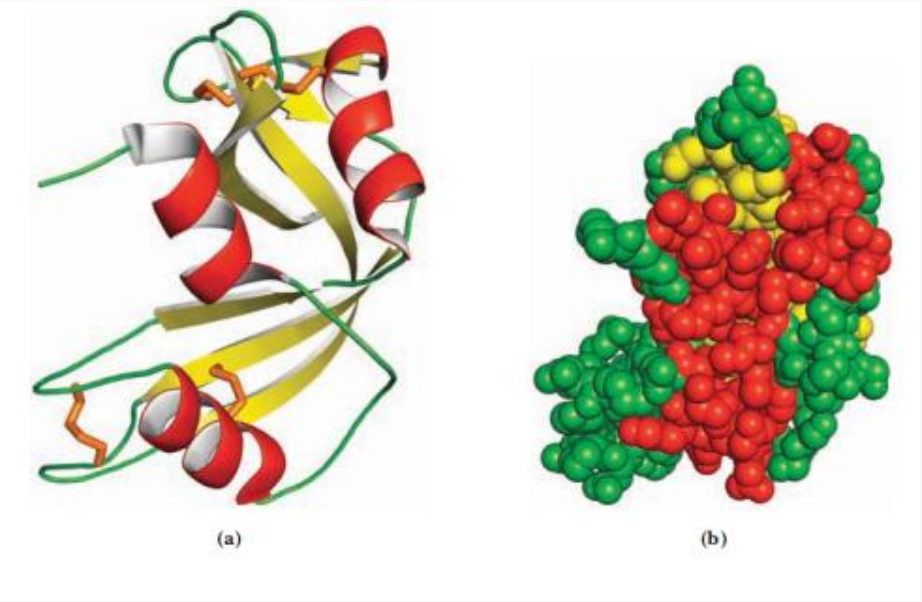
Waters on the protein surface stabilize the structure. The surface structure of a globular protein incorporates water molecules, with numerous polar groups from both the backbone and side chains interacting via hydrogen bonds with solvent water molecules.



Structure of Protein

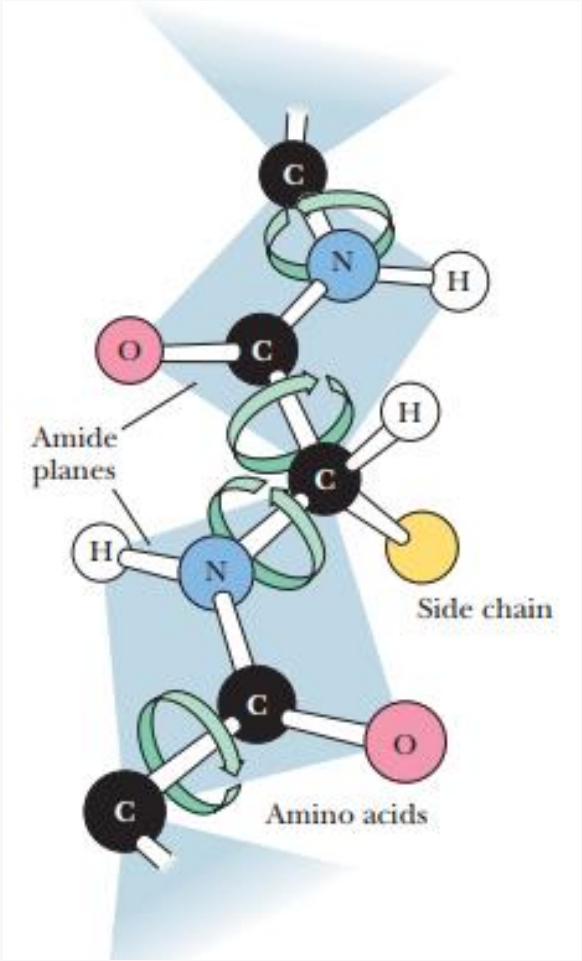
Helices and Sheets make up the core of most Globular proteins. 3D structure of bovine ribonuclease A (pdb=1FS3).

Packing: The tertiary structures of proteins rely heavily on packing. Secondary structures are closely packed together and also fit between extended polypeptide chains. When you calculate the sum of the van der Waals volumes of the amino acids in a protein and divide it by the total volume occupied by the protein, you typically get packing densities ranging from 0.72 to 0.77. These densities are similar to those found in collections of solid spheres. This indicates that even though proteins are tightly packed, about 25% of their total volume isn't taken up by protein atoms. Most of this unused space consists of very tiny cavities.



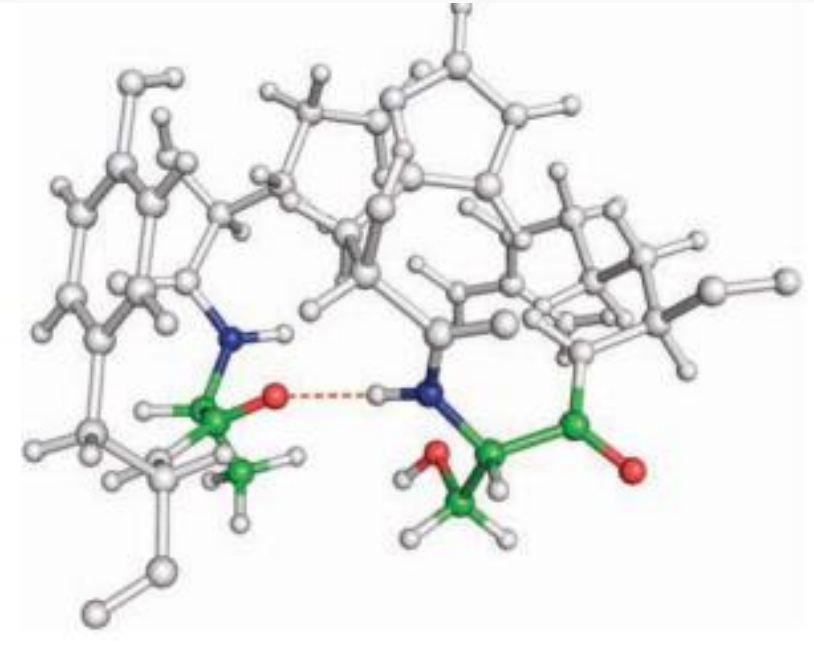
Higher-order structure or Proteins

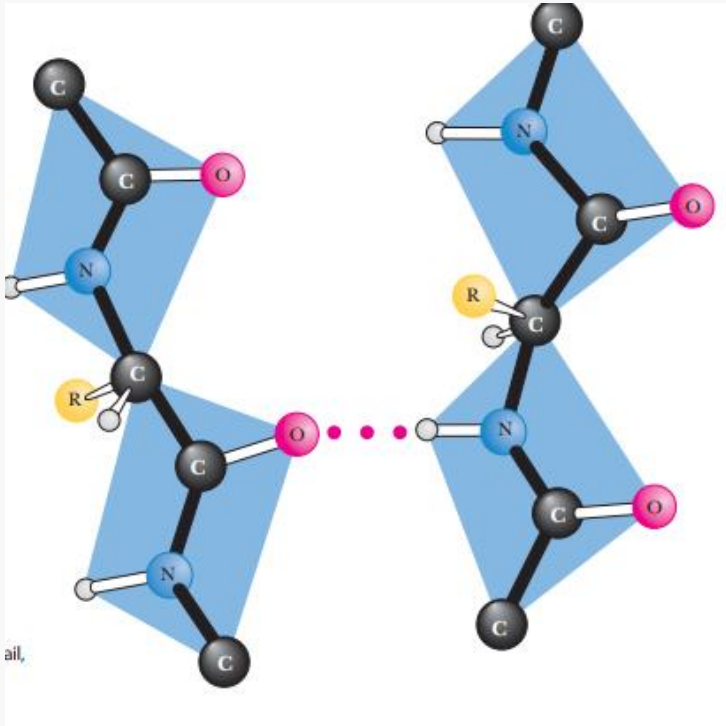
- **Protein conformation** refers the overall three-dimensional shape of a protein. **Protein configuration** refers. The geometric arrangement of atoms in a molecule.
- **Conformation vs configuration**:
  - Configuration changes require breaking and rearranging covalent bonds.
  - Conformation changes do not require breaking covalent bonds.
- **Conformational possibilities**:
  - Rotations about single bonds along peptide backbone can alter polypeptide chain shape and can create many possible orientation for the protein chain, called conformational possibilities.
  - Only a few conformations are energetically favorable under physiological conditions.



Structure of Protein

In the acetylcholine-binding protein of the snail, Lymnaea stagnalis (PDB ID: 1I9B), a hydrogen bond forms between the backbone C=O of Ala191 and the backbone N-H of Ser147. This interaction contributes to the protein's stability and function. The crystal structure of this protein provides valuable insights into ligand binding and receptor interactions.





Conjugated Proteins

- Typically, proteins are made of only amino acids, e.g. ribonuclease A and chymotrypsinogen.
- **Conjugated proteins** have amino acids and other chemical components, e.g. hemoglobin and nucleoproteins.
- The non-amino acid part of a conjugated protein is usually called its **prosthetic group**, e.g. heme and nucleic acids. They are classified on the basis of their prosthetic groups.

Class	*Prosthetic group*	*Example*
Lipoproteins	Lipids	β_1-Lipoprotein of blood
Glycoproteins	Carbohydrates	Immunoglobulin G
Phosphoproteins	Phosphate groups	Casein of milk
Hemoproteins	Heme (iron porphyrin)	Hemoglobin
Flavoproteins	Flavin nucleotides	Succinate dehydrogenase
Metalloproteins	Iron	Ferritin
	Zinc	Alcohol dehydrogenase
	Calcium	Calmodulin
	Molybdenum	Dinitrogenase
	Copper	Plastocyanin

Separation and Purification of Proteins

- Proteins are separated based on varying properties like size, charge, and binding properties.
- The initial step in protein purification involves breaking open cells to release proteins into a crude extract.
- The crude extract undergoes treatments for fractionation, separating proteins based on specific properties.
- Early fractionation utilizes protein solubility differences influenced by pH, temperature, salt concentration, etc.
- Solutions containing the target protein may need alterations for further purification steps.
- Dialysis is used to separate proteins from solvents due to their larger size.
- Large Proteins must be sequenced in smaller segments.

Carbohydrate

Understand the complex process in living organisms by reviewing the structure and function of the most common carbohydrates and glycoconjugates.

Origin and Abundance

- Most abundant biomolecules, linking solar energy to living organisms.
- Over half of organic carbon comes from carbohydrates.
- Formed through photosynthesis, using light energy to create energy-rich molecules.

Composition and Functions:

- Primarily composed of carbon, hydrogen, and oxygen $(CH_2O)_n$.
- Diverse functions including:
  - Energy source: e.g., glucose
  - Structure: e.g., cellulose, chitin
  - Cellular communication and identity
  - Precursor for other biomolecules

Carbohydrate

Classification:

- Monosaccharides: Single sugar units.
- Disaccharides: Two sugar units joined by glycosidic bond.
- Oligosaccharides: Few sugar units (3-10).
- Polysaccharides: Many sugar units linked together.

Beyond Simple Sugars:

- Carbohydrates often linked to other molecules:
- Glycoproteins: Proteins with carbohydrate groups.
- Glycolipids: Lipids with carbohydrate groups.
- Nucleotides and nucleic acids: Use sugar molecules (ribose/deoxyribose) for structure.

Importance and Distribution:

- Glycoconjugates found in all living things, especially eukaryotes.
- Play crucial roles in various biological processes.

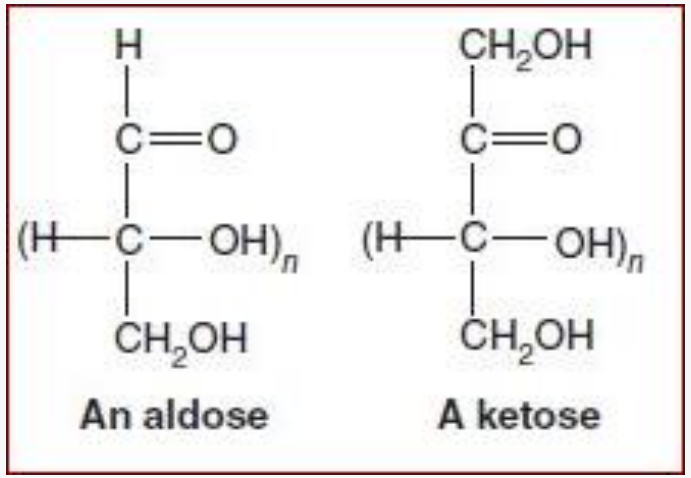
Monosaccharides

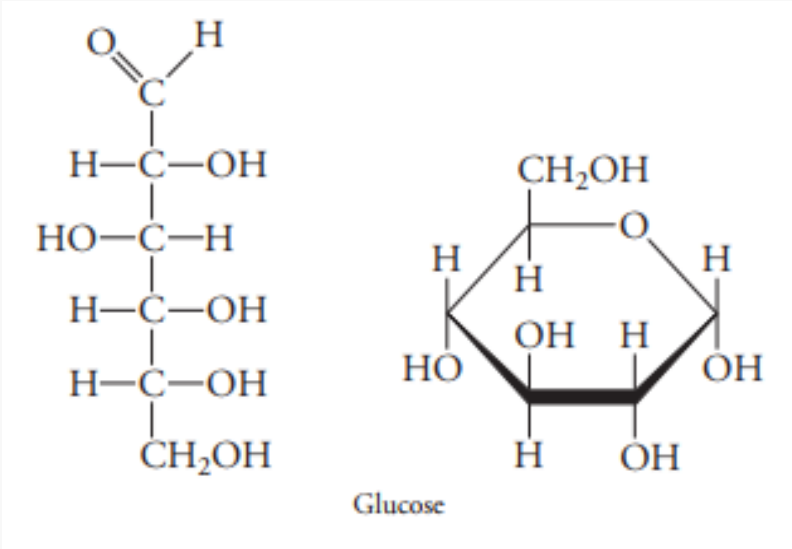
They are the simplest form of carbohydrates, also known as simple sugars. They are the building blocks of more complex carbohydrates like disaccharides and polysaccharides.

They can be classification based on the functional group:

- **Aldoses**: Aldehyde functional group (e.g., glucose).
- **Ketoses**: Ketone functional group (e.g., fructose).
- **Based on carbon atoms**: Trioses (3 C), Tetroses (4 C) , Pentoses (5 C) - important in nucleic acids (RNA, DNA), Hexoses (6 C) - most common (e.g., glucose, fructose), and so on.

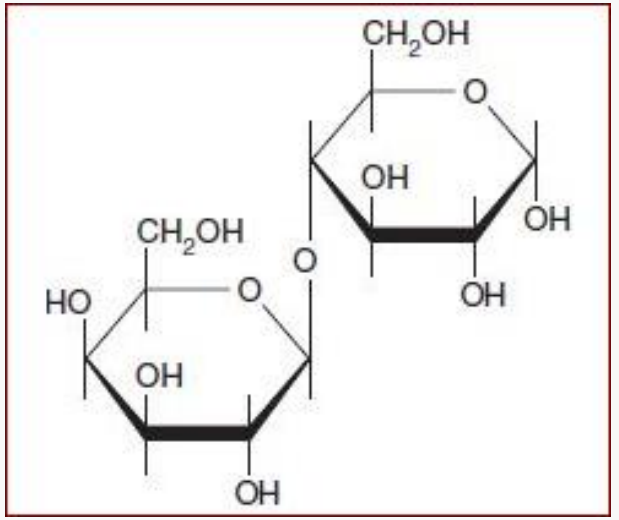
It has the formula $(CH_2O)_n$ where $n \geq 3$. Example, glucose, a monosaccharide with six carbon atoms, has the formula $C_6H_{12}O_6$





Disaccharides

Composed of two monosaccharides linked by a glycosidic bond. Often referred to as "double sugars".



Examples:

- Sucrose (table sugar - glucose + fructose)
- Lactose (milk sugar - galactose + glucose)
- Maltose (sugar from starch - glucose + glucose)

More complex than monosaccharides: Not immediately usable for energy by cells. Require digestion to break down into monosaccharides.

Functions:

- Energy source after digestion (e.g., lactose in infants)
- Transport and storage of sugars (e.g., sucrose in plants)
- Structural component in some organisms (e.g., trehalose in insects)

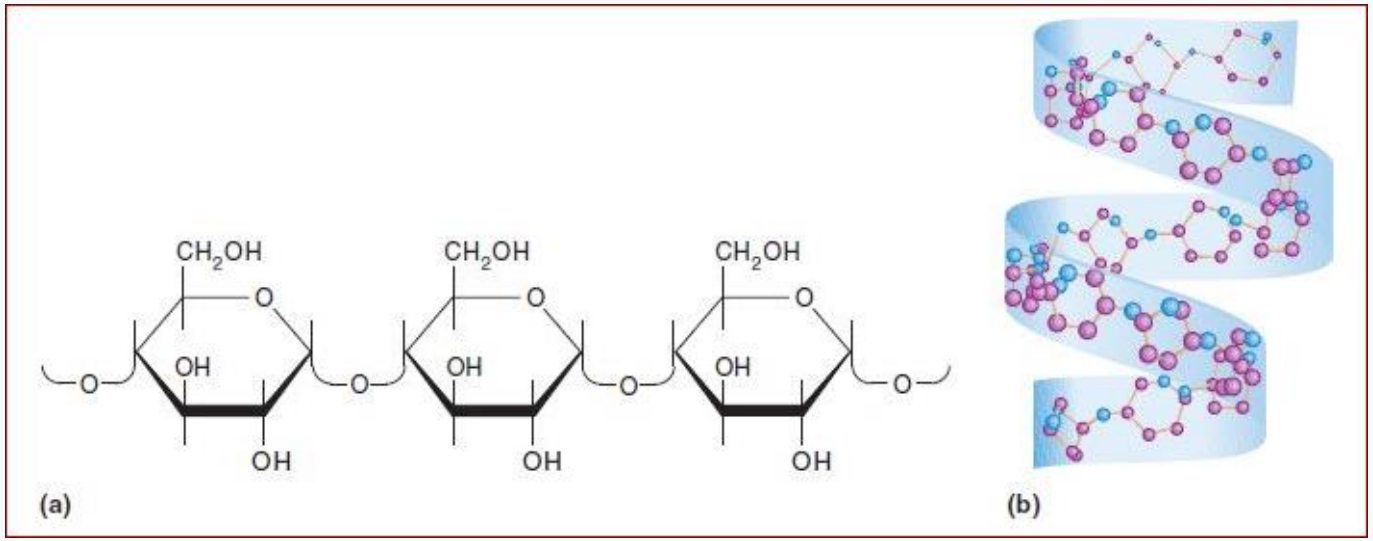
Polysaccharide (Glycans)

Structure:

- Large molecules composed of many monosaccharide units linked by glycosidic bonds.
- Smaller versions (oligosaccharides) have 3-10 units, often attached to proteins/lipids.
- Larger polysaccharides have hundreds/thousands of units.
- Can be linear or branched.

Classification:

- **Homoglycans**: One type of monosaccharide (e.g., starch, cellulose).
- **Heteroglycans**: Two or more types of monosaccharides (e.g., hemicellulose, pectin).



Polysaccharide (Glycans)

Examples:

- Homoglycans: Starch (energy storage), cellulose (plant cell walls), chitin (insect exoskeletons).
- Heteroglycans: Agarose (gelling agent), hyaluronic acid (tissue lubrication).

Functions:

- Energy storage: Starch (broken down into glucose for energy).
- Structure: Cellulose (plant cell walls), chitin (insect exoskeletons).
- Lubrication: Hyaluronic acid (joint lubrication).
- Cell communication and recognition: Glycoproteins on cell surfaces.

Additional points:

- Glycans have diverse and crucial roles in various biological processes.
- Their structure and function are highly dependent on the specific monosaccharide types and linkages.

Polysaccharide (Glycans)

Examples:

- Homoglycans: Starch (energy storage), cellulose (plant cell walls), chitin (insect exoskeletons).
- Heteroglycans: Agarose (gelling agent), hyaluronic acid (tissue lubrication).

Functions:

- Energy storage: Starch (broken down into glucose for energy).
- Structure: Cellulose (plant cell walls), chitin (insect exoskeletons).
- Lubrication: Hyaluronic acid (joint lubrication).
- Cell communication and recognition: Glycoproteins on cell surfaces.

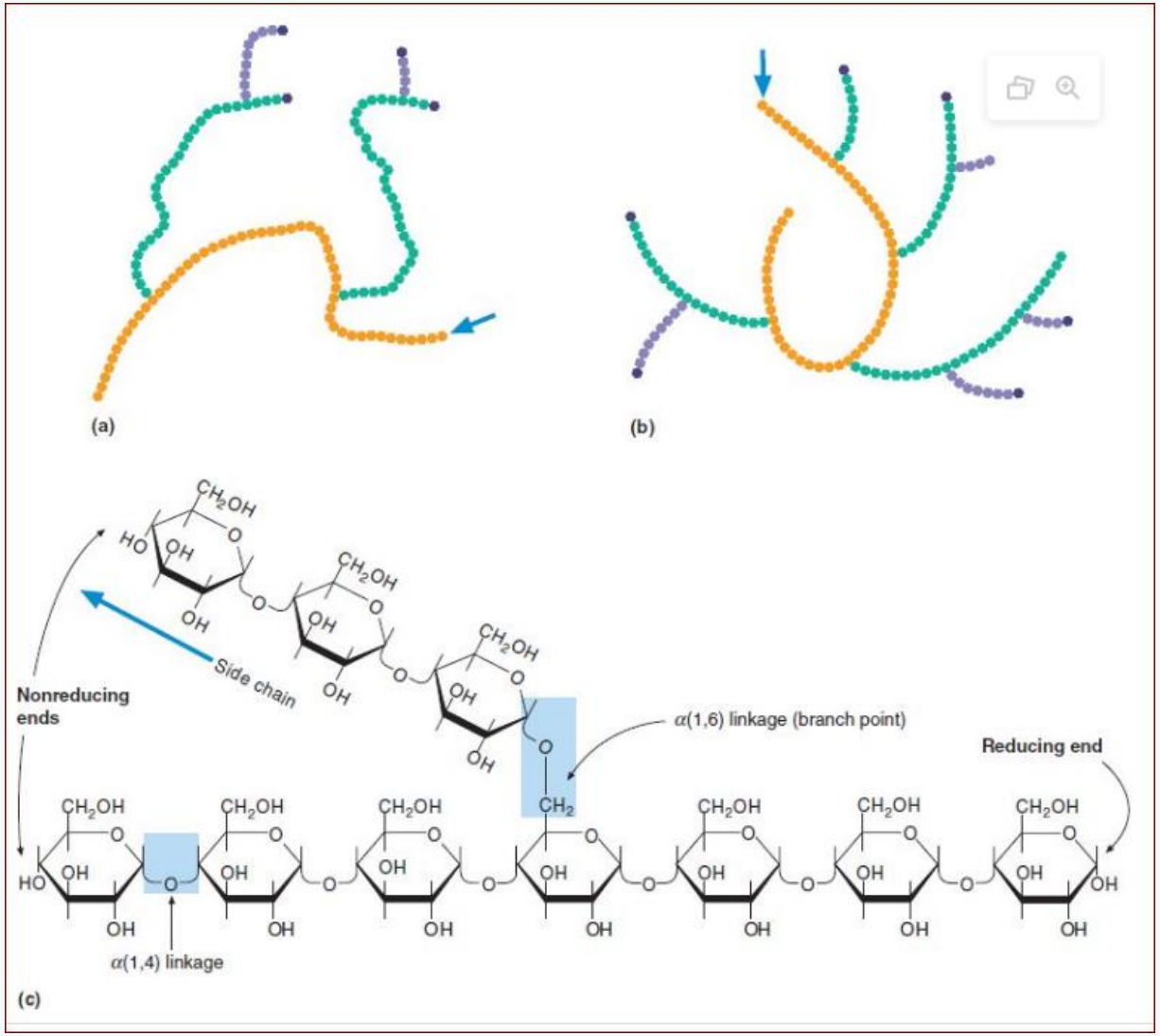
Additional points:

- Glycans have diverse and crucial roles in various biological processes.
- Their structure and function are highly dependent on the specific monosaccharide types and linkages.

Polysaccharide (Glycans)

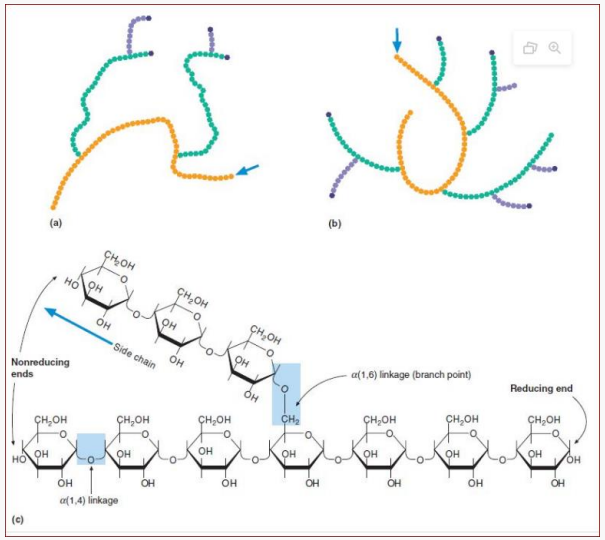
(a) Amylopectin and (b) Glycogen

Each hexagon represents a glucose molecule. (c) Detail from (a) and (b)



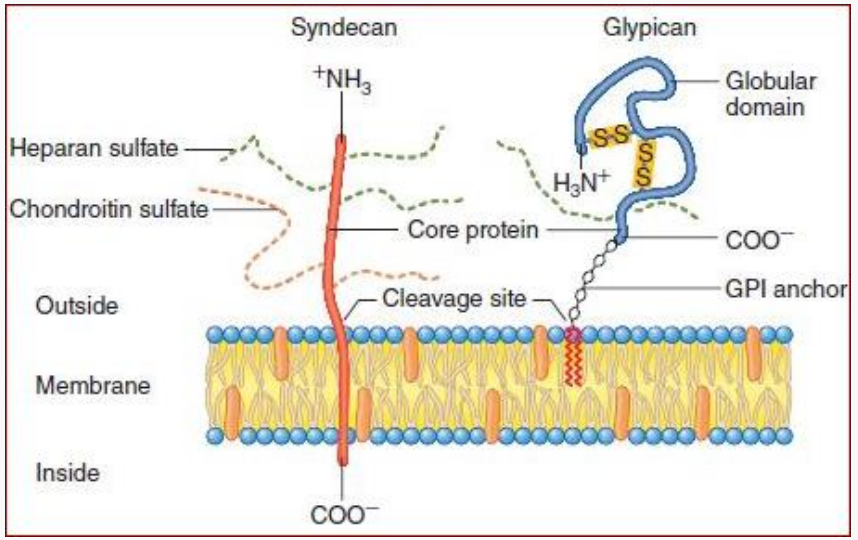
Glycoconjugates

The molecules that result from the covalent linkages of carbohydrate molecules to both proteins and lipids are collectively known as the glycoconjugates. There are two classes of carbohydrate-protein conjugate: proteoglycans and glycoproteins



(1) Amylopectin and (b) Glycogen. Each hexagon represents a glucose molecule. (c) Detail from (a) and (b)

(2) Structure of the Syndecans and Glypicans

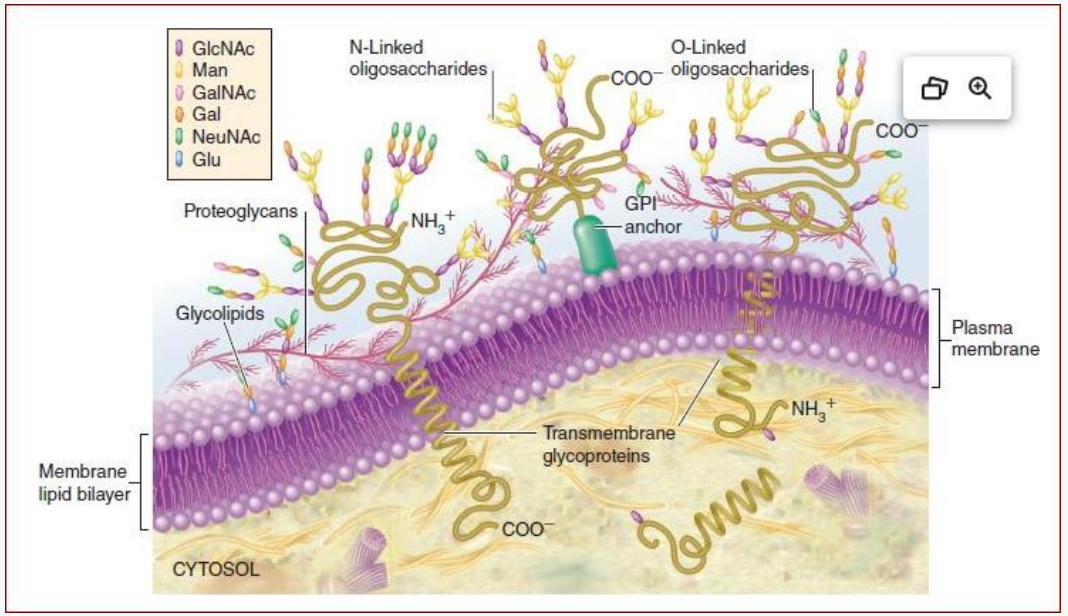


Glycoproteins

Proteins with covalently bound carbohydrates (1-85% by weight). Glycoproteins are a diverse group of molecules that are ubiquitous constituents of most living organisms.

Key Features:

- Lack uronic acids, sulfate groups, and repeating units found in Proteoglycans.
- Diverse functions: cell adhesion, recognition, immune response, etc.
- Carbohydrate composition varies greatly across different glycoproteins.
- Plays crucial roles in various biological processes.

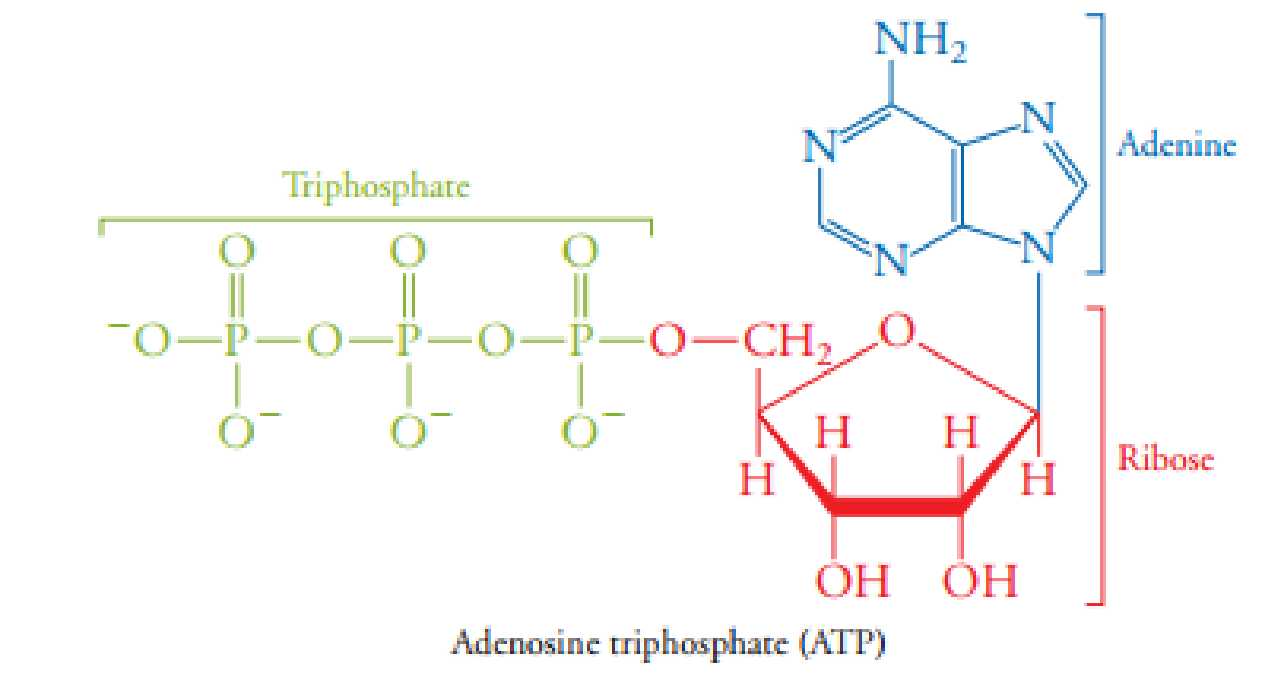


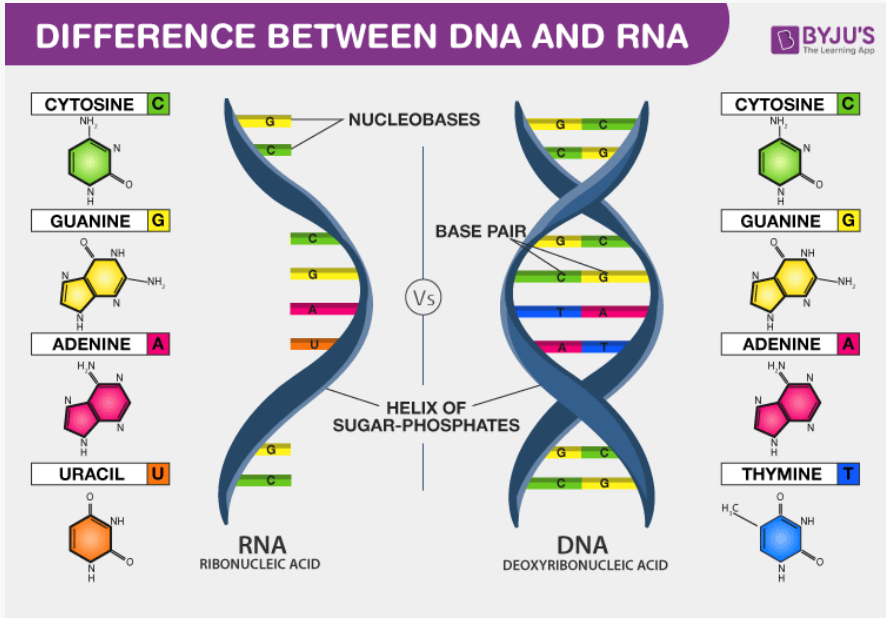
Glycoproteins

Type	Example	Source	Molecular Mass (Da)
Enzyme	Ribonuclease B	Bovine	14,700
Immunoglobulin	Immunoglobulin A Immunoglobulin M	Human Human	160,000 950,000
Hormone	Chorionic gonadotropin Follicle-stimulating hormone	Human placenta Human	38,000 34,000
Membrane protein	Glycophorin	Human red blood cells	31,000
Lectin (carbohydrate-building proteins)	Potato lectin Soybean agglutinin Ricinus lectin	Potato Soybean Castor bean	50,000 120,000 120,000
Viral envelope protein	gp120	HIV	120,000

Nucleic Acids (later)

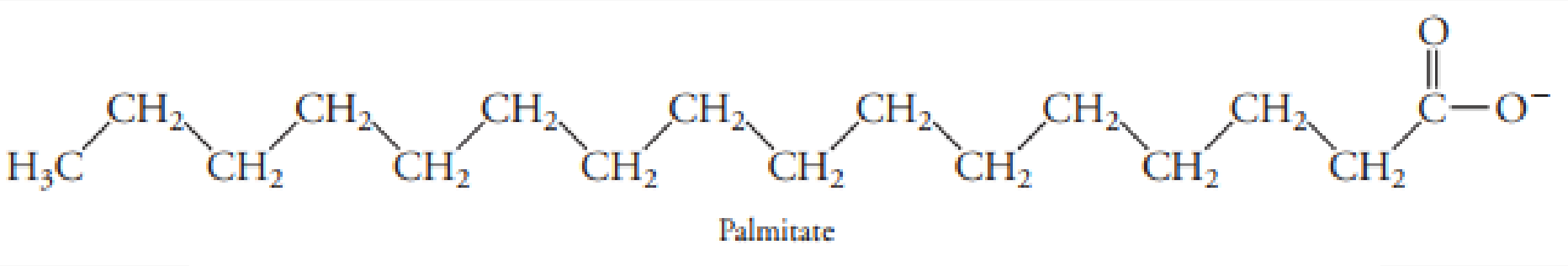
A nucleotide is **the basic building block of nucleic acids**. ... A nucleotide consists of a sugar molecule (either ribose in RNA or deoxyribose in DNA) attached to a phosphate group and a nitrogen-containing base. The bases used in DNA are adenine (A), cytosine (C), guanine (G), and thymine (T).



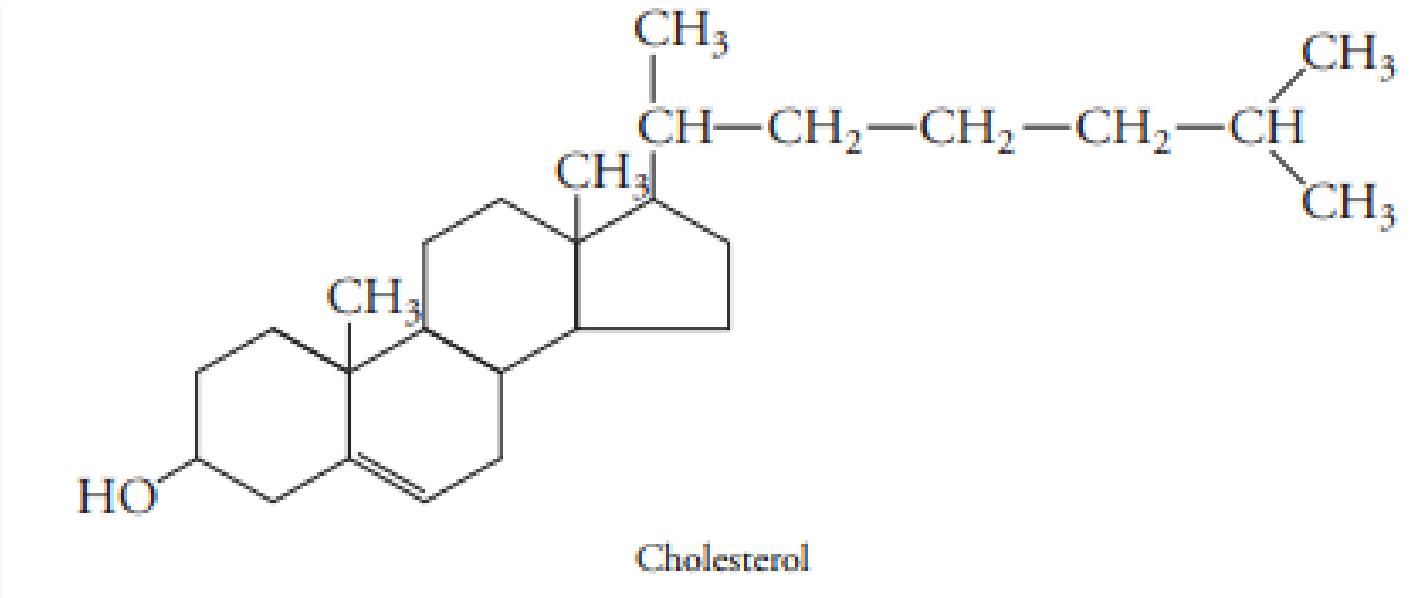


Lipids

Cannot be described by a single structural formula since they are a diverse collection of molecules. However, they all have in common a tendency to be poorly soluble in water because the bulk of their structure is hydrocarbon-like.



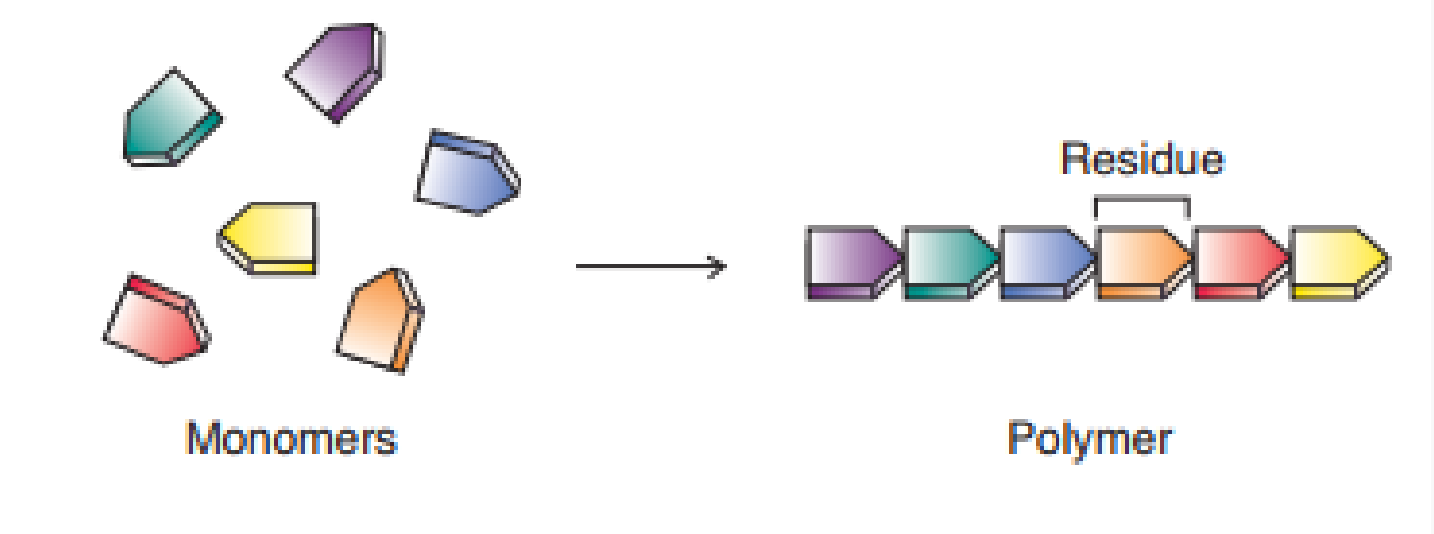
Palmitate



Cholesterol

Major Biological Polymers

Addition to small molecules, organisms contain macromolecules that may consist of thousands of atoms where these huge molecules are not synthesized in one piece but are built from smaller units.



Biomolecules

Biopolymer	Encode Information	Carry Out Metabolic Reactions	Store Energy	Support Cellular Structures
Proteins	—	✔	✓	✔
Nucleic acids	✔	✓	—	✓
Polysaccharides	✓	—	✔	✔

✔ major function
✓ minor function

- All life form fallows the basic rule of Central Dogma of Molecular Biology (Francis Crick, discoverer of DNA double-helix): $\boldsymbol{DNA \rightarrow RNA \rightarrow Proteins}$.
- DNA: Genetic information- very stable, it almost lasts entire lifetime. It makes a copy by RNA.
- RNA: "Copy" of genetic information. Rapidly react when outside environment needs, but much more than this.
- Proteins: Workhorse of Cells. It does every work such structural components, do catalysis (enzymes), make reactions go which normally don't happened.

Thank you very much for your attention