

Protein Structure and Function

Dr. Urmi Roy

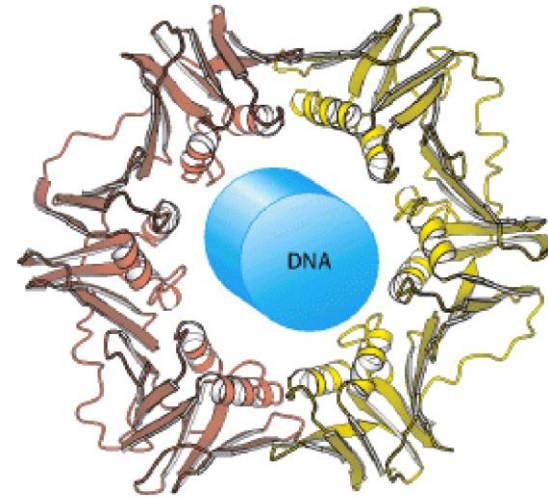
Definition

- Proteins are
 - The most versatile macromolecules in living systems
 - Serve crucial functions in essentially all biological processes.
 - They function as
 - catalysts
 - they transport and
 - store other molecules such as oxygen
 - they provide mechanical support and
 - immune protection
 - they generate movement,
 - they transmit nerve impulses
 - they control growth and differentiation.

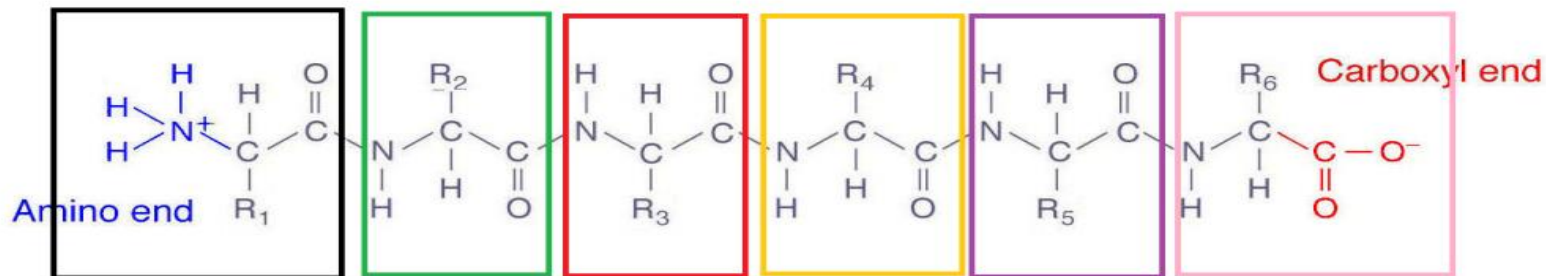


Properties

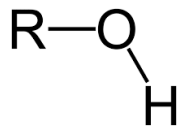
- Proteins are linear polymers
- They are built of monomer units called *amino acid*
 - The function is directly dependent on its 3D structure.
 - Proteins spontaneously fold up into 3D structure: determined by the sequence of amino acids in the protein polymer.
 - Proteins are capable of diverse activities.



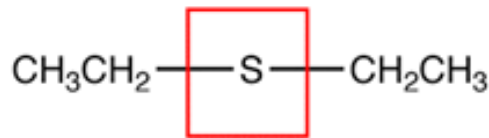
A protein component of the DNA replication machinery surrounds a section of DNA double helix. The structure of the protein allows large segments of DNA to be copied without the replication machinery dissociating from the DNA.



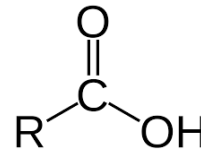
- Proteins contain a wide range of *functional groups*.
 - alcohols,
 - thioethers,
 - carboxylic acids,
 - carboxamides and
 - a variety of basic groups.
 - These functional groups accounts for the broad spectrum of protein function.
 - The chemical reactivity associated with these groups is essential to the function of enzymes (proteins), that catalyze specific chemical reactions.



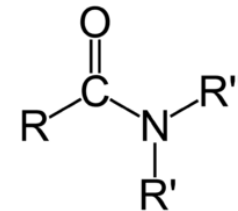
alcohols



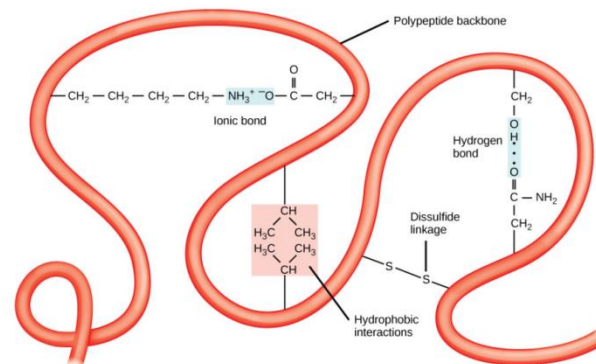
thioethers



carboxylic acids



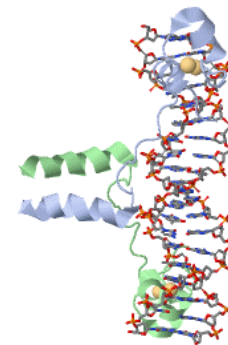
carboxamides



- Proteins can *interact*
 - With one **another** and with **other** biological macromolecules to form complex assemblies.
 - The proteins within these assemblies can act **synergistically** to generate capabilities not afforded by the individual component proteins.
 - These assemblies include macro-molecular machines that carry out the accurate **replication of DNA**, the transmission of signals within cells and many other essential processes.

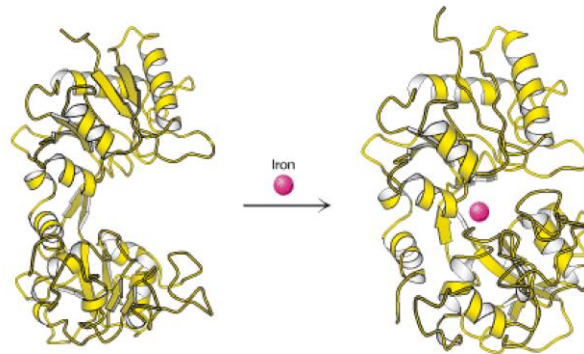


Protein with protein



Protein with DNA

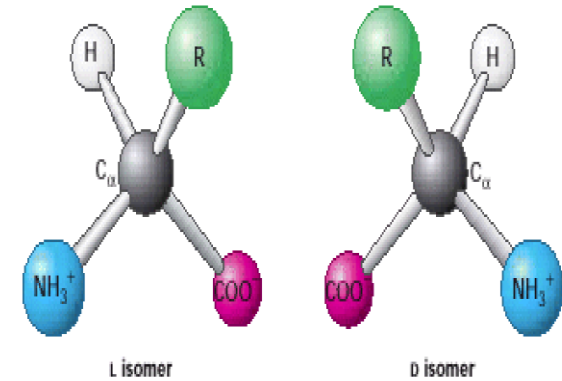
- Some proteins are quite **rigid**, whereas others display **limited flexibility**.
 - **Rigid** units: can function as **structural** elements
 - in the cytoskeleton or
 - in connective tissue.
 - Parts of proteins with **limited flexibility** may act as
 - hinges,
 - springs and
 - levers
 - These are crucial to
 - protein function
 - the assembly of proteins with one another
 - with other molecules into complex units
 - transmission of information within and between cells.



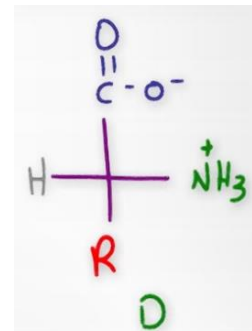
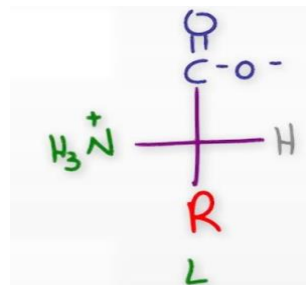
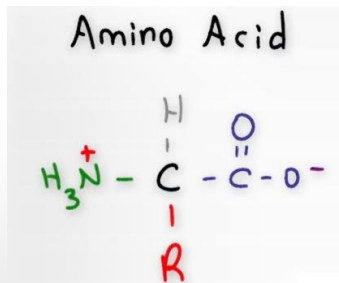
Flexibility and Function. Upon binding iron, the protein lactoferrin undergoes conformational changes that allow other molecules to distinguish between the iron-free and the iron-bound forms.

Proteins Are Built from a Repertoire of 20 Amino Acids

- Amino acids are the building blocks of proteins.
- An α -amino acid consists of
 - a central carbon atom, called the α carbon,
 - linked to an amino group,
 - a carboxylic acid group,
 - a hydrogen atom, and
 - a distinctive R group (referred to as the side chain)
- With four different groups connected to the tetrahedral α -carbon atom, α -amino acids are **chiral**.
- The two mirror-image forms are called the *l isomer* and the *d isomer*.



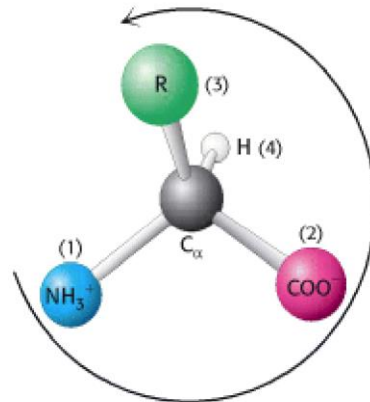
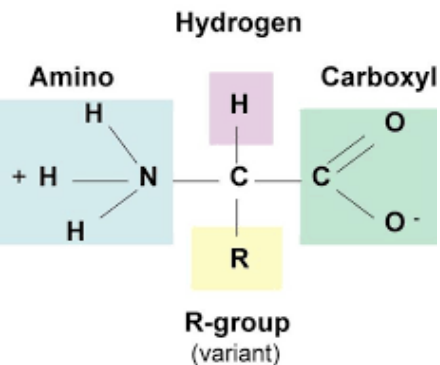
The *l* and *d* Isomers of Amino Acids. R refers to the side chain. The *l* and *d* isomers are mirror images of each other.



Notation for distinguishing stereoisomers

- The **four different substituents** of an **asymmetric carbon** atom are assigned a priority according to atomic number.
- The **lowest-priority** substituent, often **hydrogen**, is pointed away from the viewer.
- Only 1 amino acids are constituents of proteins.
- Considerable effort has gone into understanding why amino acids in proteins have this absolute configuration, no satisfactory explanation has been arrived at.
- It seems plausible that the selection of l over d was arbitrary but, once made, was fixed early in evolutionary history.

Amino Acid Structure

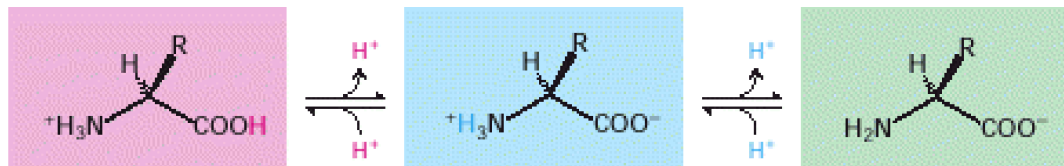


Only l Amino Acids Are Found in Proteins.

Almost all l amino acids have an *S* absolute configuration (from the Latin *sinister* meaning "left"). The counterclockwise direction of the arrow from highest- to lowest-priority substituents indicates that the chiral center is of the *S* configuration.

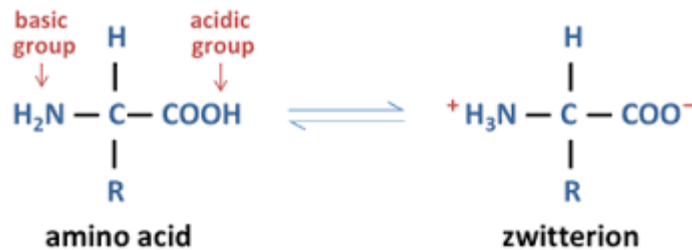
Zwitterions

- Amino acids in solution at neutral pH exist predominantly as dipolar ions (also called *zwitterions*).
- In the dipolar form, the amino group is protonated ($-\text{NH}_3^+$) and the carboxyl group is deprotonated ($-\text{COO}^-$).
- The ionization state of an amino acid varies with pH.
- In acid solution (e.g., pH 1), the amino group is protonated ($-\text{NH}_3^+$) and the carboxyl group is not dissociated ($-\text{COOH}$).
- As the pH is raised, the carboxylic acid is the first group to give up a proton.
- The dipolar form persists until the pH approaches 9, when the protonated amino group loses a proton.



Amphoteric property

- Amino acids are amphoteric in nature
- They act as both acids and base since due to the two amine and carboxylic group present.

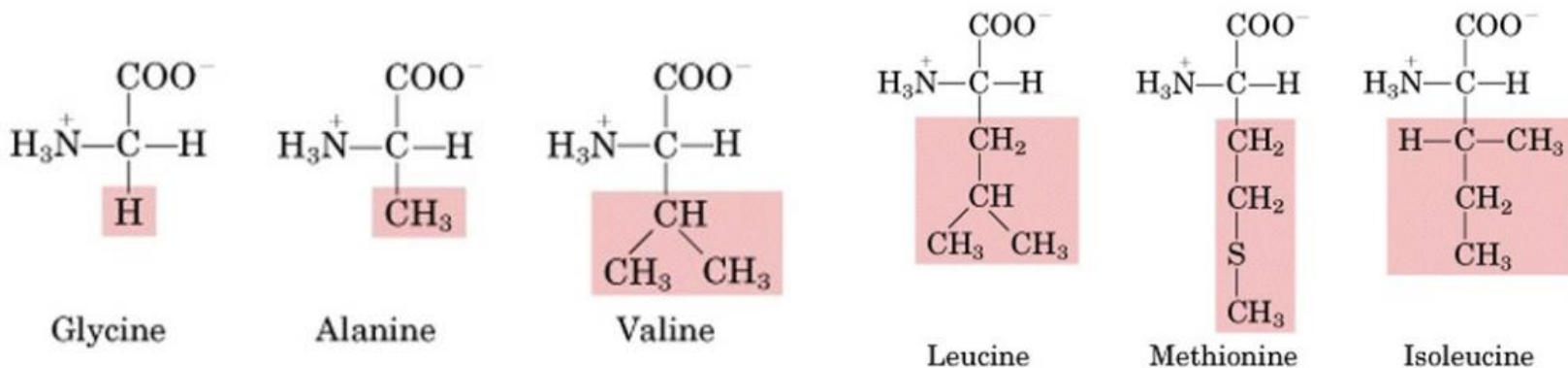


Amino acids

- More than 300 amino acids are found in nature but only 20 amino acids are standard and present in protein because they are coded by genes.
- Other amino acids are modified amino acids and called non-protein amino acids.
- All proteins are constructed from the same set of 20 amino acids.
- The range of functions mediated by proteins results from the **diversity** and **versatility** of these 20 building blocks.

Amino Acids with Non-Polar Aliphatic Side Chains

- The **simplest** one is **glycine**,
 - It has just a hydrogen atom as its side chain.
 - With two hydrogen atoms bonded to the α -carbon atom,
 - Glycine is unique in being **achiral**.
- **Alanine**,
 - The next simplest amino acid,
 - It has a methyl group (-CH₃) as its side chain.
- The larger aliphatic side chains are hydrophobic that is, they tend to cluster together rather than contact water.
- The three-dimensional structures of water-soluble proteins are stabilized by this tendency of hydrophobic groups to come together, called the hydrophobic effect.
- The different sizes and shapes of these hydrocarbon side chains enable them to pack together to form compact structures with few holes.
- Next are the **Valine, Leucine**
- **Isoleucine**:
 - The side chain includes an additional chiral center. Only one isomer is present in proteins.
- **Methionine**: It contains a largely aliphatic side chain that includes a thioether (-S-) group.



Amino Acids with Aromatic side chains

- Three amino acids with relatively simple aromatic side chains are:

- **Phenylalanine**,

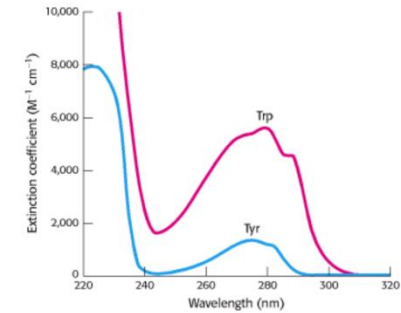
- It is purely hydrophobic
- It contains a **phenyl ring** attached in place of one of the hydrogen of alanine.

- **Tyrosine**

- The aromatic ring contains a **hydroxyl group**.
- They are **less** hydrophobic because of their reactive hydroxyl groups

- **Tryptophan**

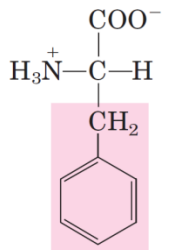
- Has an **indole ring** joined to a methylene (-CH₂-) group
- They are also less hydrophobic because of their -NH groups
- The indole group comprises two fused rings and -NH group.



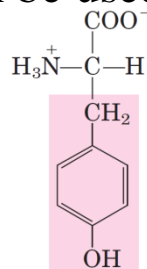
- The aromatic rings of **tryptophan** and **tyrosine** contain delocalized electrons: strongly absorb ultraviolet light

- For tryptophan, absorption is maximum at 280 nm and tyrosine, absorption is maximum at 276 nm
- Phenylalanine absorbs light less strongly and at shorter wavelengths.

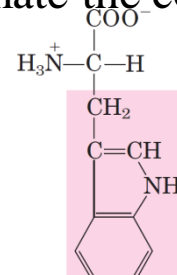
- The absorption of light at 280 nm can be used to estimate the concentration of a protein in solution.



Phenylalanine



Tyrosine

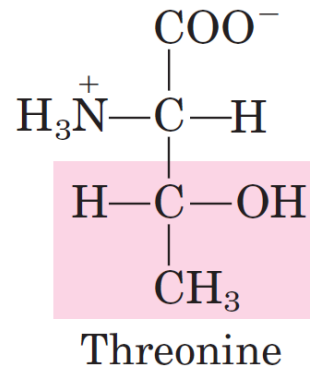
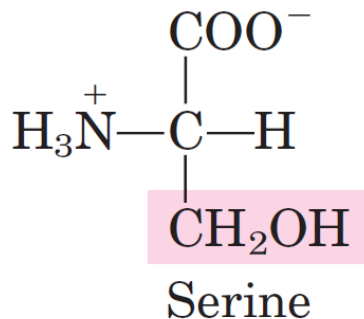


Tryptophan

Amino Acids With Unchanged Side chains

➤ *Amino Acids Containing Aliphatic Hydroxyl Groups*

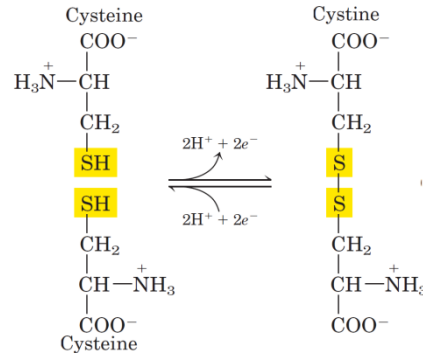
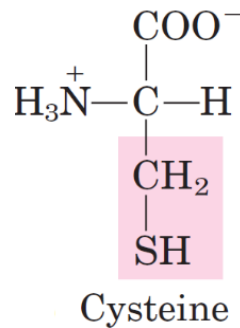
- Serine and threonine,
 - Contain aliphatic hydroxyl.
 - The hydroxyl groups make them much more hydrophilic (water loving) and reactive than alanine and valine.
- **Serine**
 - It can be thought of as a hydroxylated version of alanine
- **Threonine**
 - It resembles valine with a hydroxyl group in place of one of the valine methyl groups.
 - Threonine, contains an additional asymmetric center; only one isomer is present in proteins.



➤ Amino Acid Containing Aliphatic Sulfhydryl Groups

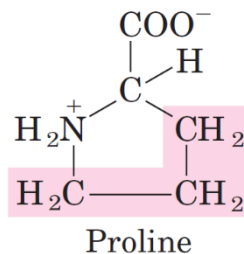
• **Cysteine**

- It is structurally similar to serine but contains a sulfhydryl or thiol (-SH), group in place of the hydroxyl (-OH) group.
- The sulfhydryl group is much more reactive.
- Pairs of sulfhydryl groups may come together to form disulfide bonds, which are particularly important in stabilizing some proteins.



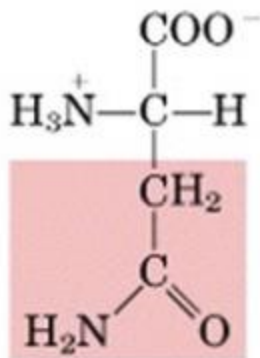
• **Proline**

- It also has an aliphatic side chain, but its side chain is bonded to both the nitrogen and the α -carbon atoms.
- Proline markedly influences protein architecture because its ring structure makes it more conformationally restricted than the other amino acids.

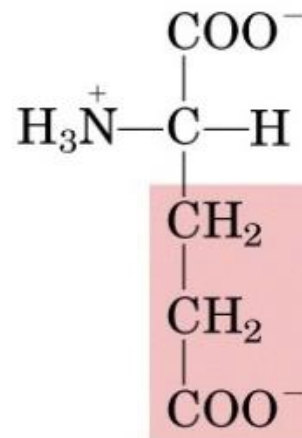


Amino Acids With Unchanged Side chains....continued

- Side chains do accept protons
- The set includes uncharged derivatives of aspartate, the **asparagine** and glutamate, the **glutamine**.
- Each of which contains a terminal carboxamide in place of a carboxylic acid



Asparagine

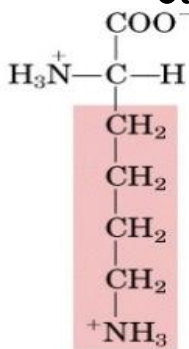


Glutamate

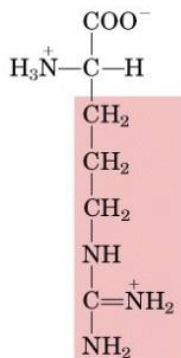
Amino Acids with Positively Charged side Chains

The Basic Amino Acids

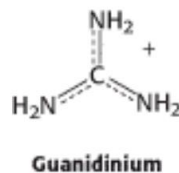
- The amino acids with very polar side chains that render them highly hydrophilic are **lysine**, **arginine** and **histidine**.
- They have relatively long side chains that terminate with groups that are positively charged at neutral pH.
- Lysine: It is capped by a *primary amino group*
- Arginine: It has a *guanidinium* group.
- Histidine:
 - It contains an *imidazole* group, an aromatic ring that also can be positively charged.
 - It is often found in the active sites of enzymes, where the imidazole ring can bind and release protons in the course of enzymatic reactions.



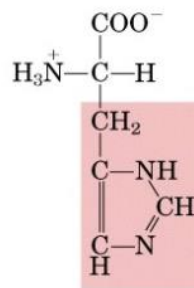
Lysine



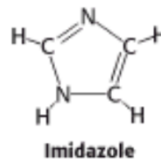
Arginine



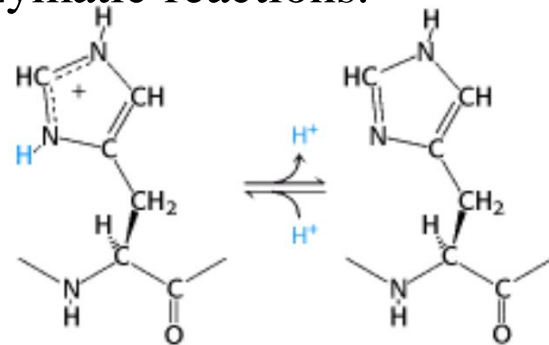
Guanidinium



Histidine



Imidazole

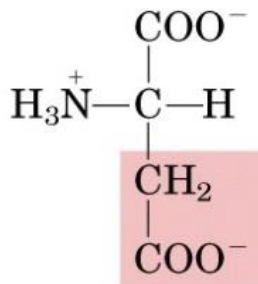


Histidine Ionization

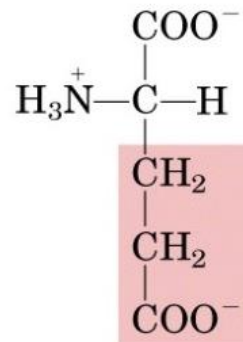
Amino acids containing Negatively Charged Side Chains

The Acidic Amino Acid

- The set of amino acids also contains two with acidic side chains:
 - aspartic acid and
 - glutamic acid.
- These amino acids are often called
 - **aspartate** and
 - **glutamate** to emphasize that their side chains are usually negatively charged at physiological pH.
- In some proteins, these side chains do accept protons.
- This ability is often functionally important.



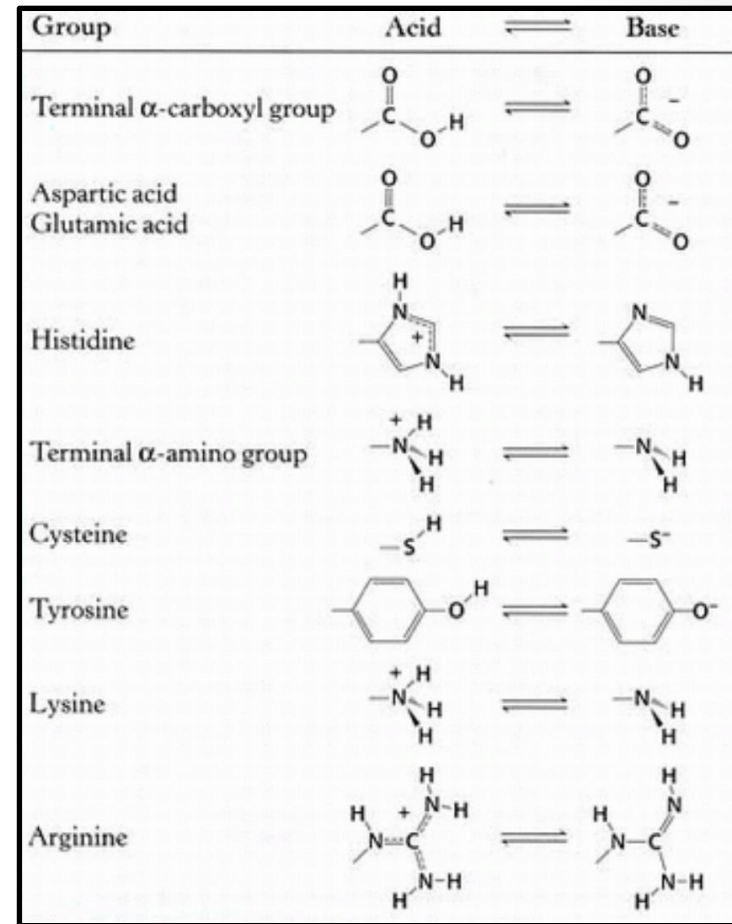
Aspartate



Glutamate

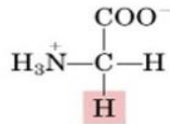
Amino acids with readily ionizable side chains

- There are seven of the 20 amino acids have readily ionizable side chains.
- They are able to *donate or accept* protons to facilitate reactions as well as to form ionic bonds.
 1. Tyrosine,
 2. Cysteine,
 3. Arginine,
 4. Lysine,
 5. Histidine,
 6. Aspartic and
 7. Glutamic acids in proteins.
- Two other groups in proteins can be ionized:
 1. The terminal α -amino group and
 2. The terminal α - carboxyl group

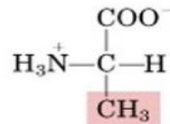


Classification of amino acids on the basis of R-group

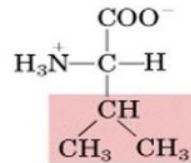
Nonpolar, aliphatic R groups



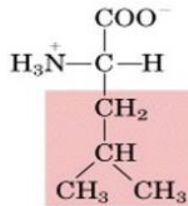
Glycine



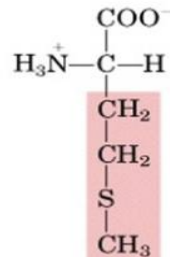
Alanine



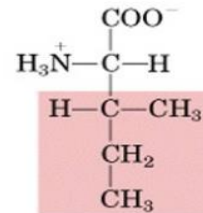
Valine



Leucine

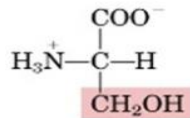


Methionine

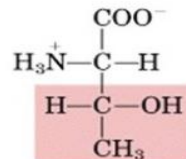


Isoleucine

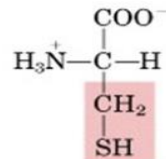
Polar, uncharged R groups



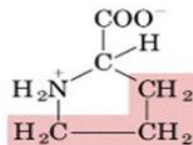
Serine



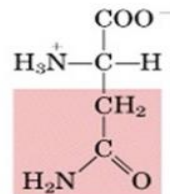
Threonine



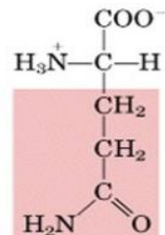
Cysteine



Proline

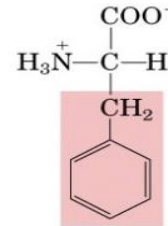


Asparagine

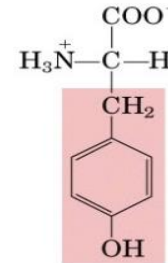


Glutamine

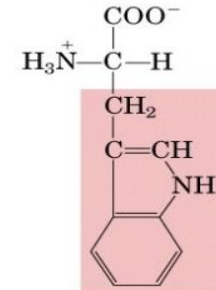
Aromatic R groups



Phenylalanine

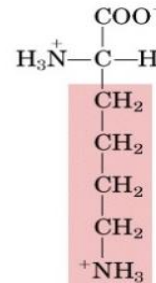


Tyrosine

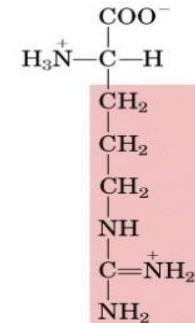


Tryptophan

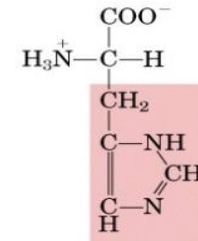
Positively charged R groups



Lysine

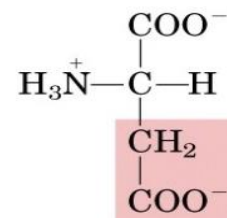


Arginine

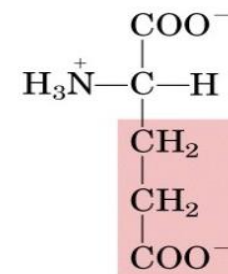


Histidine

Negatively charged R groups



Aspartate



Glutamate

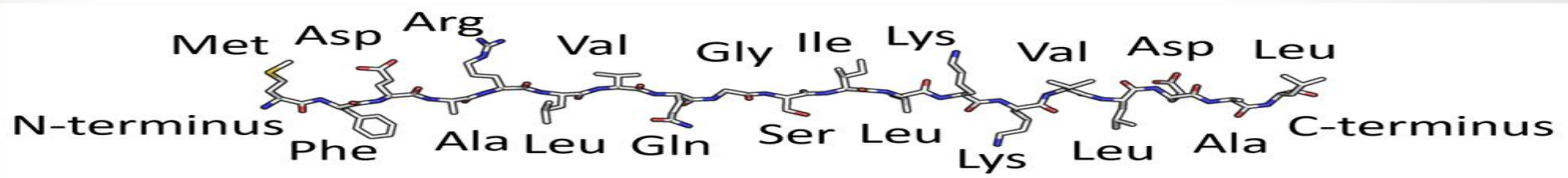
Amino acids: Three-Letter abbreviations

- Amino acids are often designated by
 - A three-letter abbreviation
 - Those are the first three letters of their names,
 - Except for asparagine (Asn), glutamine (Gln), isoleucine (Ile), and tryptophan (Trp).
 - A one-letter symbol
 - The symbols for many amino acids are the first letters of their names
 - e.g., G for glycine and L for leucine;
- The abbreviations for amino acids the other symbols have been agreed on by convention.
- These abbreviations and symbols are an integral part of the vocabulary of biochemists.

Abbreviations for amino acids

Amino acid	Three-letter abbreviation	One-letter abbreviation
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic Acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic Acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V
Asparagine or aspartic acid	Asx	B
Glutamine or glutamic acid	Glx	Z

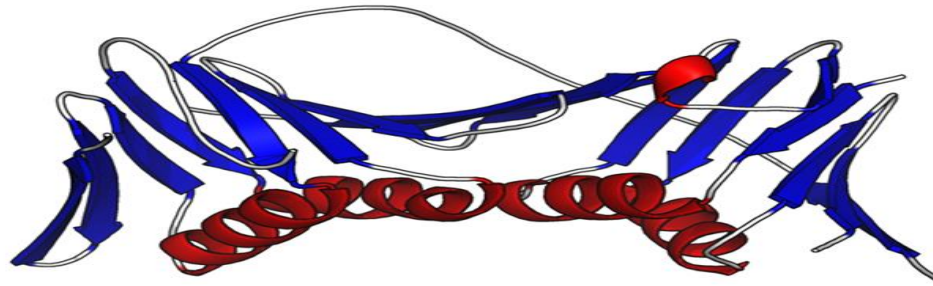
Primary



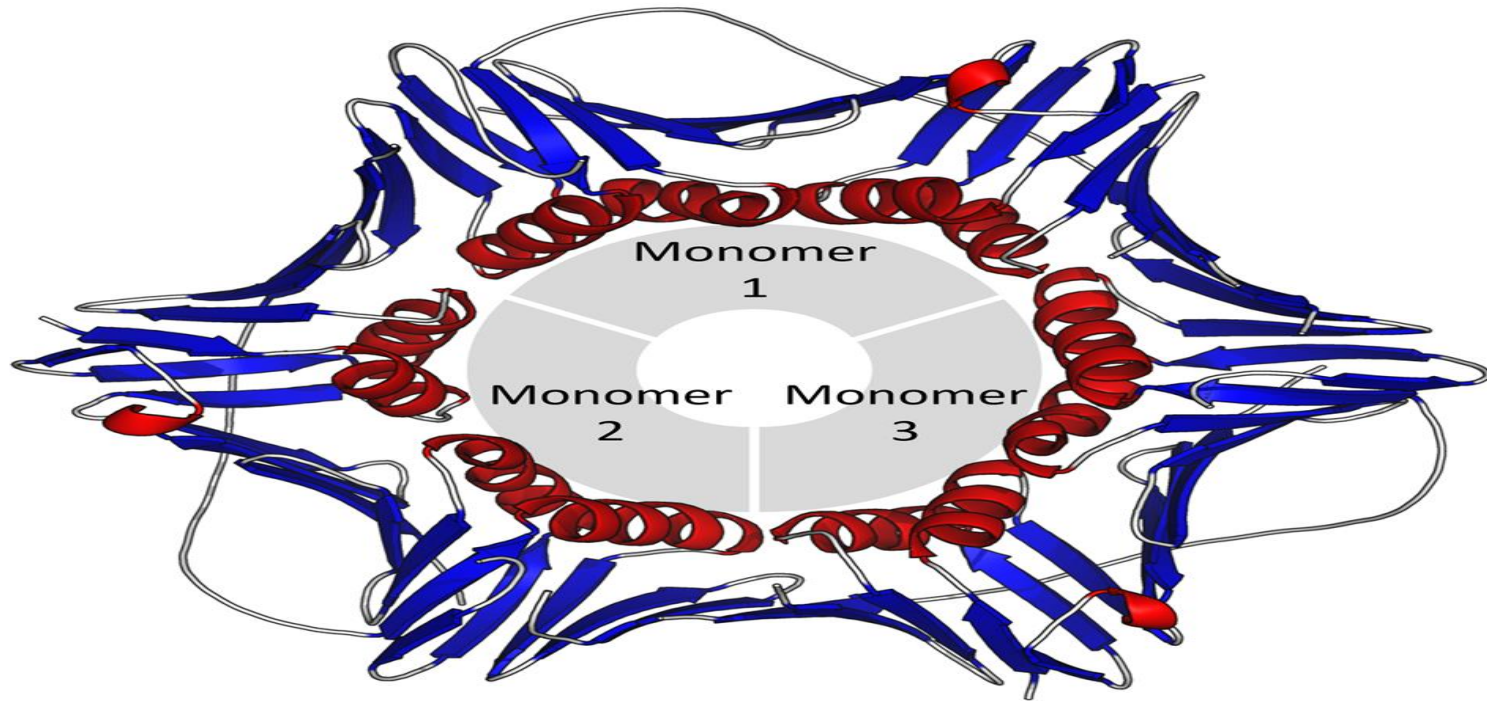
Secondary



Tertiary

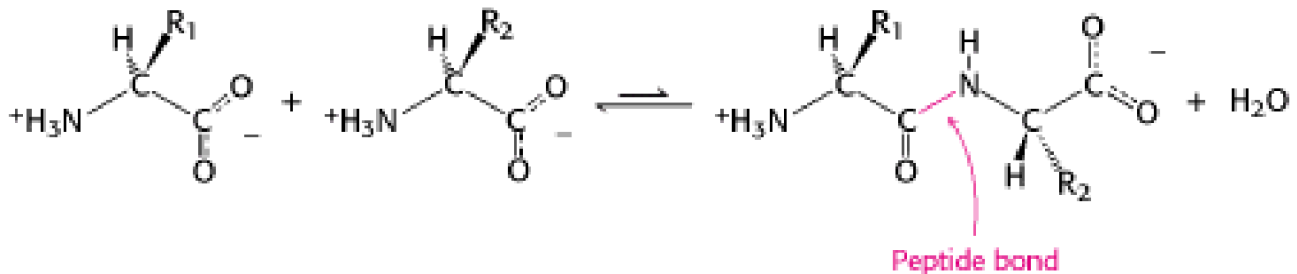
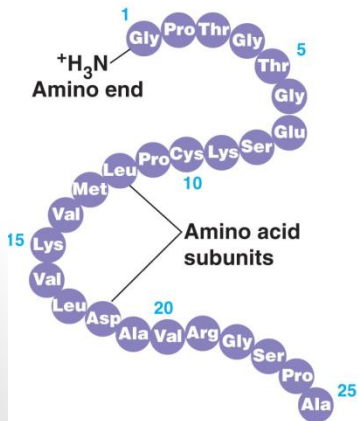


Quaternary

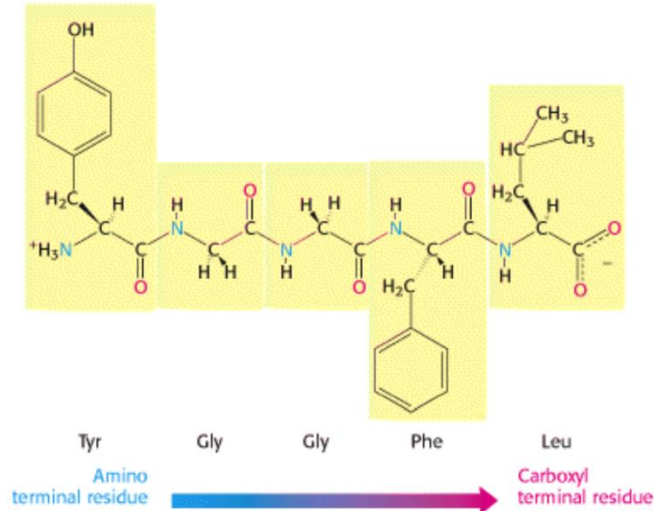


Primary Structure

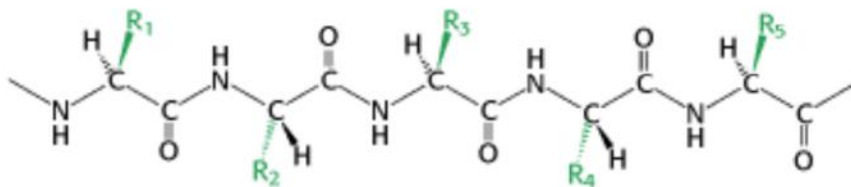
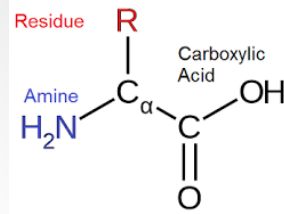
- Primary structure is *the amino acid sequence* of a protein.
- Proteins are linear polymers formed by linking the
 - α -carboxyl group of one amino acid to the
 - α -amino group of another amino acid
 - with a *peptide bond* (also called an *amide bond*).
- The formation of a dipeptide from two amino acids is accompanied by the loss of a *water molecule*.
- The equilibrium of this reaction lies on the side of hydrolysis rather than synthesis.
- The biosynthesis of peptide bonds requires an input of free energy.
- Peptide bonds are quite stable kinetically;
- The lifetime of a peptide bond in aqueous solution in the absence of a catalyst approaches 1000 years.



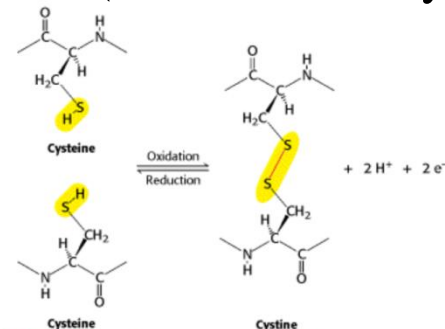
- A series of amino acids joined by peptide bonds form a polypeptide chain
- Each *amino acid unit* in a polypeptide is called a *residue*.
- A polypeptide chain has polarity
 - Its ends are different
 - An α -amino group at one end and
 - An α -carboxyl group at the other.
- By convention,
 - The *amino end* is taken to be the *beginning* of a polypeptide chain
 - The sequence of amino acids: starting with the amino-terminal residue.
- A polypeptide chain consists of
 - a regularly repeating part: *main chain* or *backbone* and
 - a variable part: the *distinctive side chains*.
- The polypeptide backbone is rich in hydrogen-bonding potential.



- Each *residue* contains
 - A **carbonyl group**: a good hydrogen-bond acceptor (except proline)
 - An -NH group: a good hydrogen-bond donor.
- These groups interact with each other and with functional groups from side chains to stabilize particular structures.
- Polypeptide chains with 50 to 2000 amino acid residues: **proteins**.
- Peptides with small numbers of amino acids: **oligopeptides** or **peptides**.
- The mean molecular weight of an amino acid residue: 110
- The mass of a protein is expressed in units of daltons;
 - one dalton=one atomic mass unit.
 - A protein with a molecular weight of 50,000 has a mass of 50,000 daltons, or 50 kd (kilodaltons).
- The linear polypeptide chain may be cross-linked with **disulfide bonds**, formed by the oxidation of a pair of cysteine residues (unit is called cystine).



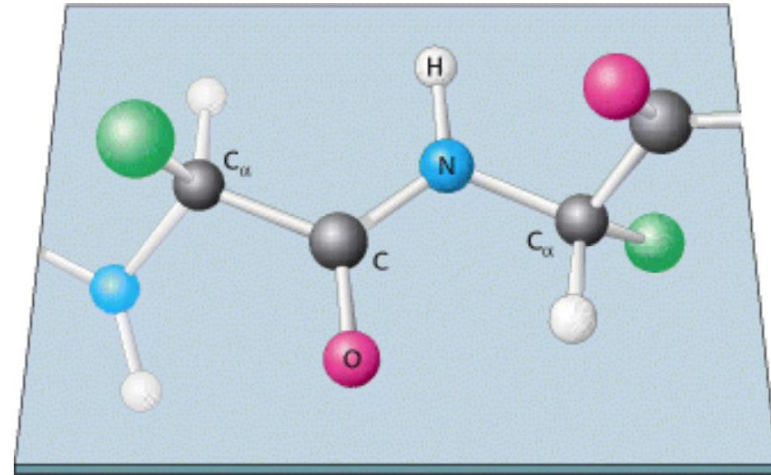
Components of a Polypeptide Chain. A polypeptide chain consists of a constant backbone and variable side chains (shown in green).



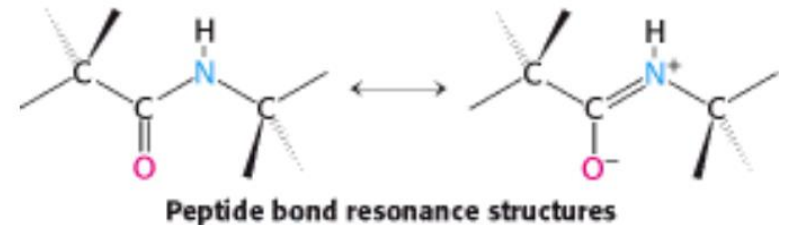
Cross-Links. The formation of a disulfide bond from two cysteine residues is an oxidation reaction.

Polypeptide Chains Are Flexible Yet Conformationally Restricted

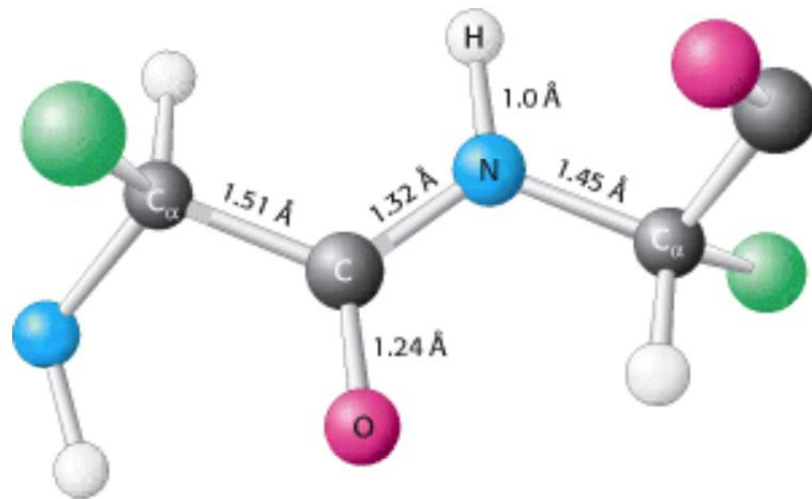
- The peptide bond is essentially planar.
- For a pair of amino acids linked by a peptide bond, six atoms lie in the same plane:
 - the α -carbon atom and
 - CO group from the first amino acid and
 - the NH group and
 - α -carbon atom from the second amino acid.
- The peptide bond has considerable **double-bond character** (between the CO and NH groups)
 - Prevents rotation about this bond.
 - Accounts for the bond's planarity.



Peptide Bonds Are Planar. In a pair of linked amino acids, six atoms (C α , C, O, N, H, and C α') lie in a plane.



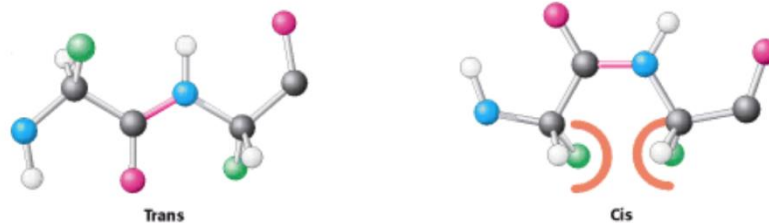
- The **CN distance in a peptide bond** is typically **1.32 Å**, which is between the values expected for
 - a C-N single bond (1.49 Å) and
 - a C=N double bond (1.27 Å).
- The peptide bond is uncharged, allowing polymers of amino acids linked by peptide bonds to form tightly packed globular structures.



Typical Bond Lengths Within a Peptide Unit.

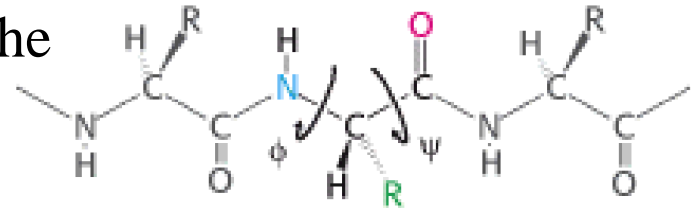
The peptide unit is shown in the trans configuration.

- Two configurations are possible for a planar peptide bond.
 - In the trans configuration: the two α -carbon atoms are on opposite sides of the peptide bond.
 - In the cis configuration, these groups are on the same side of the peptide bond.
- Almost all peptide bonds in proteins are *trans*.
- This preference can be explained by the steric clashes between groups attached to the α -carbon atoms hinder formation of the cis form but do not occur in the trans configuration.

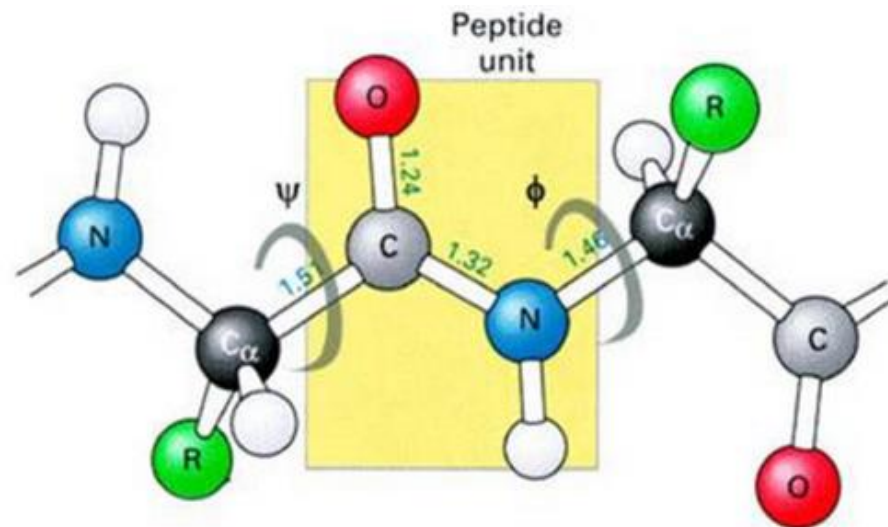


Trans and Cis Peptide Bonds. The trans form is strongly favored because of steric clashes that occur in the cis form.

- **Pure single bonds** are present between the
 - amino group and the α -carbon atom and
 - α -carbon atom and the carbonyl group.

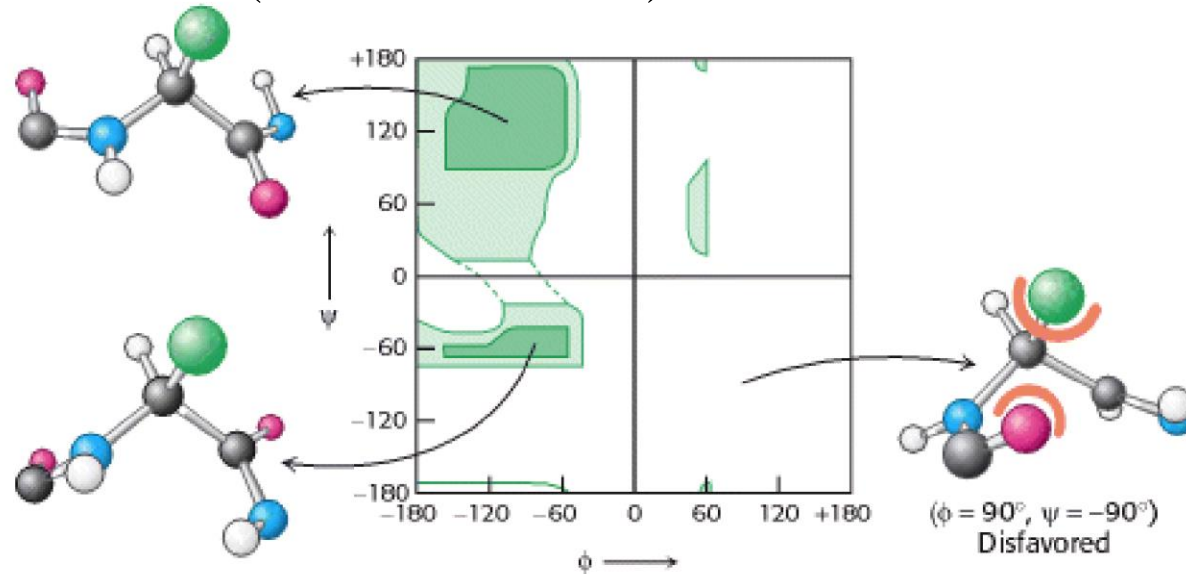


- This freedom of rotation about two bonds of each amino acid allows proteins to fold in many different ways.
- The angle of rotation about the bond between
 - The **nitrogen** and the **α -carbon** atoms is called **phi** (ϕ).
 - **α -carbon** and the **carbonyl carbon** atoms is called **psi** (ψ).
- The ϕ and ψ angles determine the path of the polypeptide chain.



Ramachandran diagram

- All combinations of ϕ and ψ are not possible.
- **G. N. Ramachandran** recognized that many combinations are forbidden because of **steric collisions** between atoms.
- The allowed values can be visualized on a two-dimensional plot called a *Ramachandran diagram*.
- Three-quarters of the possible (ϕ , ψ) combinations are excluded simply by local steric clashes (**steric exclusion**)



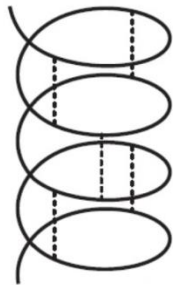
A Ramachandran Diagram Showing the Values of ϕ and ψ .

Not all ϕ and ψ values are possible without collisions between atoms. The most favorable regions are shown in dark green; borderline regions are shown in light green.

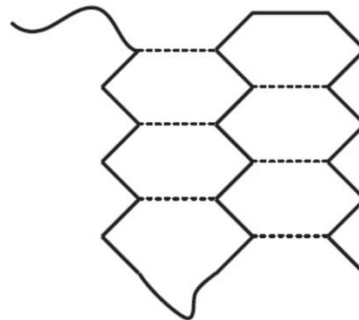
The structure on the right is disfavored because of steric clashes.

Secondary Structure

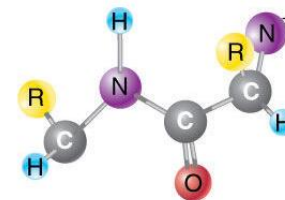
- Secondary structure refers to the **spatial arrangement of amino acid**.
- These are of a regular kind and give rise to a periodic structure.
- Formed by the residues that are nearby in the sequence.
- Linus Pauling and Robert Corey (In 1951) proposed two periodic structures called
 - the α helix (alpha helix) and the
 - β pleated sheet (beta pleated sheet).
- Subsequently, other structures such as the β turn and omega (ω) loop were identified.



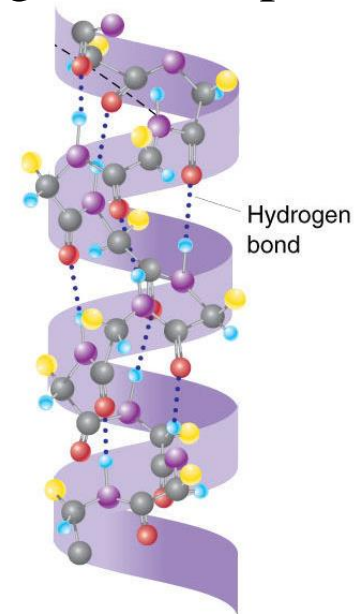
α -Helix
(intramolecular
hydrogen bonds)



β -Pleated sheet
(intramolecular
hydrogen bonds)



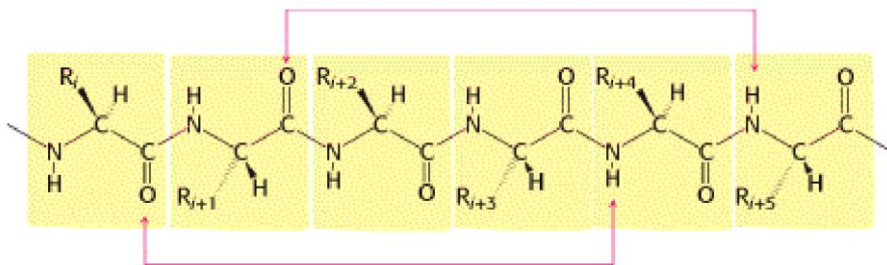
(a) Primary structure



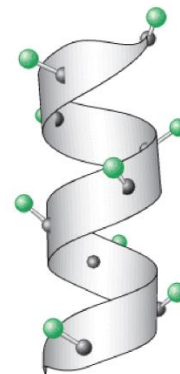
(b) Secondary structure

Alpha Helix

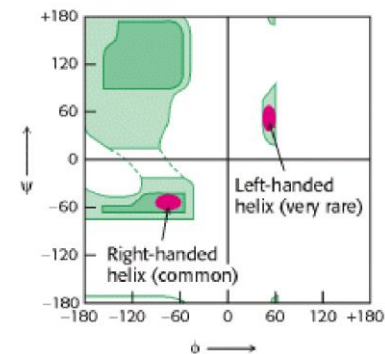
- Coiled structure stabilized by **intra-chain hydrogen** bonds.
- Pauling and Corey proposed the α helix structures.
 - It is a rod-like structure.
 - The α helix is stabilized by **hydrogen** bonds between the **NH** and **CO** groups of the main chain.
 - The CO group of each amino acid forms a hydrogen bond with the NH group of the amino acid that is situated **four residues ahead** in the sequence.
 - All the main-chain CO and NH groups are hydrogen bonded (except near the ends).
 - Each residue is related to the next one by
 - a **rise of 1.5 Å** along the helix axis and
 - a rotation of 100 degrees,
 - This gives 3.6 amino acid residues per turn of helix.
 - Amino acids spaced three and four apart in the sequence are spatially quite close to one another in an α helix.
 - The screw sense: can be right-handed (clockwise) or left-handed (counterclockwise).
 - Essentially all a helices found in proteins are right handed.



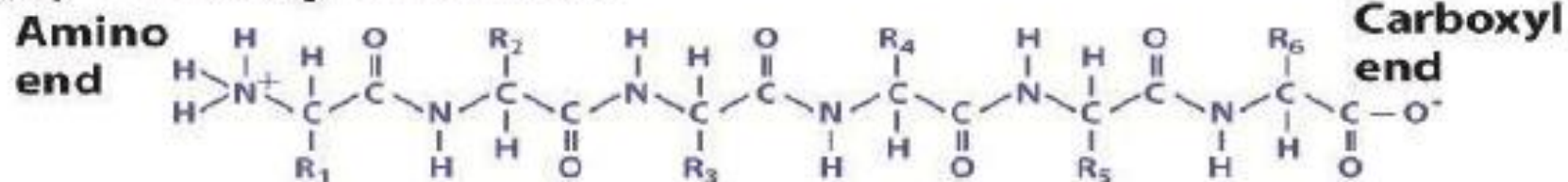
In the α helix, the CO group of residue n forms a hydrogen bond with the NH group of residue $n+4$.



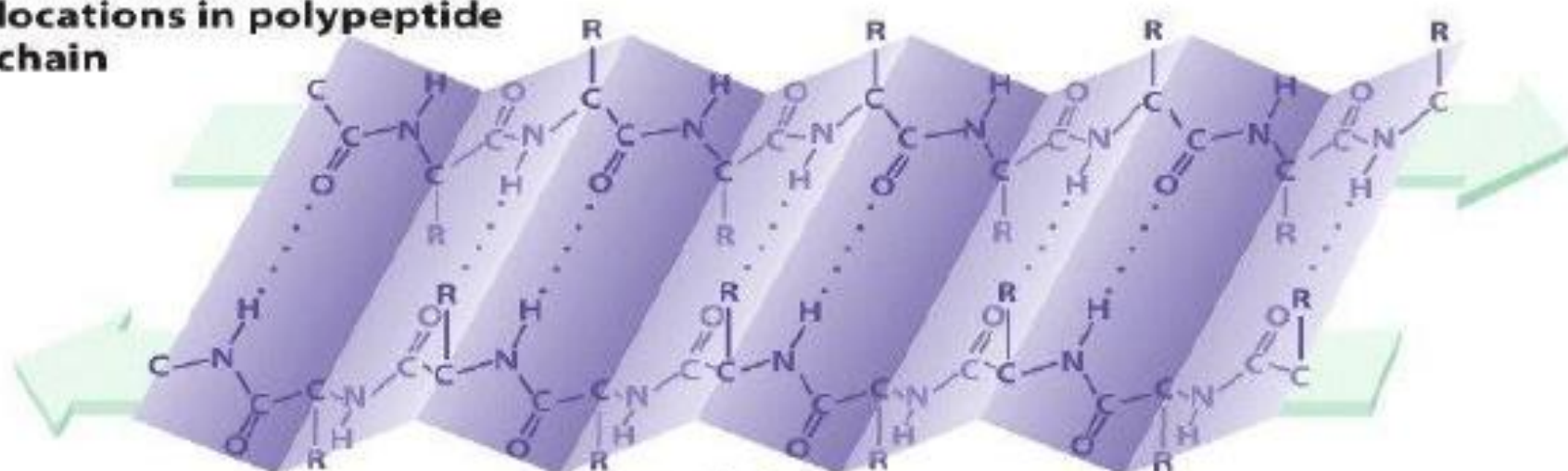
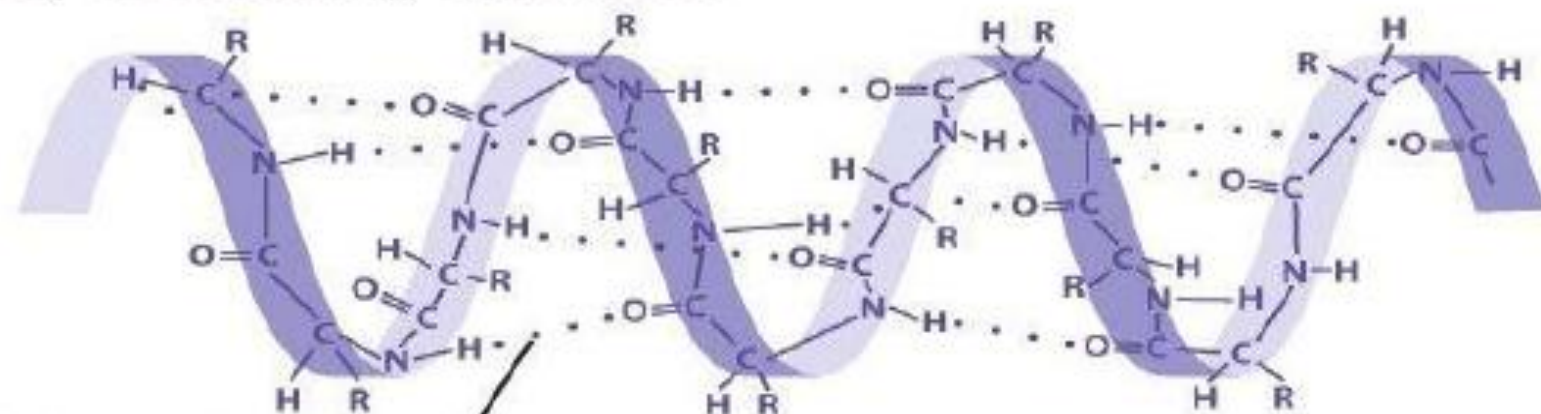
A ribbon depiction with the α -carbon atoms and side chains (green) shown.



(a) Primary structure

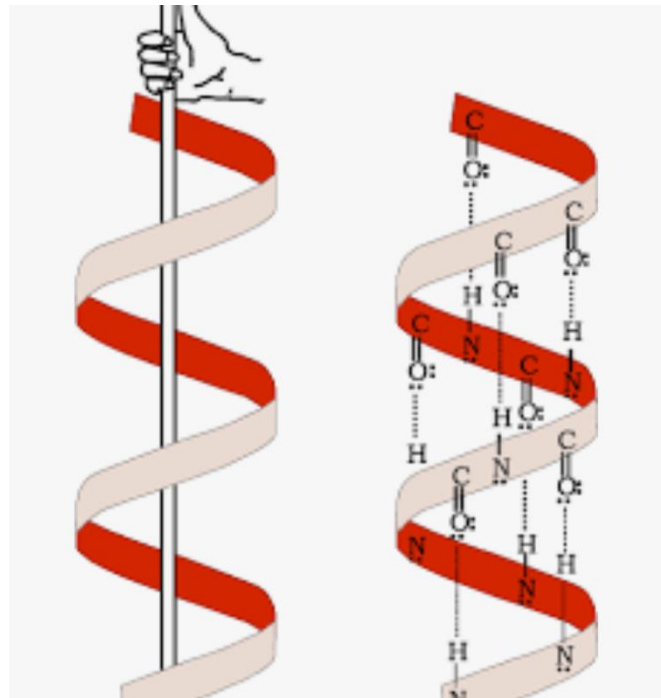


(b) Secondary structure



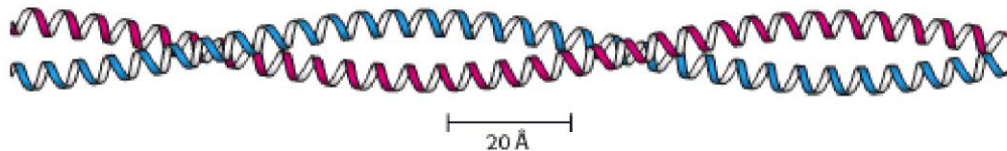
Screw sense

- Describes the direction in which a **helical structure rotates with respect to its axis**.
- If, viewed down the axis of a helix, the chain turns in a clockwise direction, it has a right-handed screw sense.

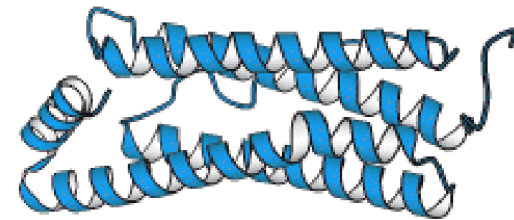


The x-ray reconstruction of the structure of Myoglobin

- It is iron- and oxygen-binding protein found in the skeletal muscle tissue
- Single α helices are usually less than 45 Å long.
- Two or more α helices can entwine to form a very stable structure, which can have a length of 1000 Å (100 nm, or 0.1 μ m).
- Such α -helical coiled coils are found in myosin and tropomyosin in muscle, in fibrin in blood clots, and in keratin in hair.
- The helical cables in these proteins serve a mechanical role in forming stiff bundles of fibers (as in porcupine quills).
- Many proteins that span biological membranes also contain α helices.



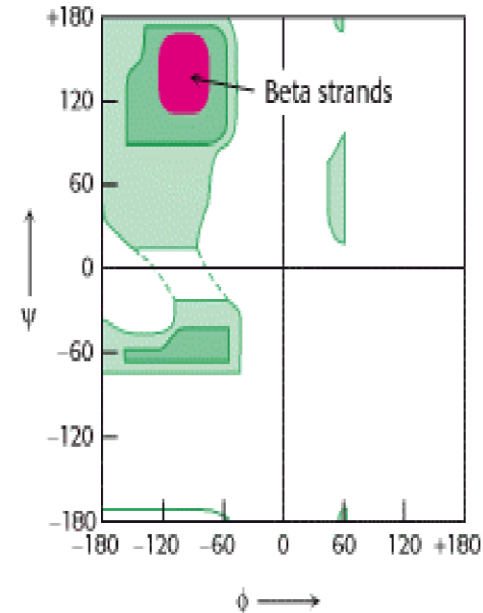
An α -Helical Coiled Coil. The two helices wind around one another to form a superhelix. Such structures are found in many proteins including keratin in hair, quills, claws, and horns.



Ferritin, an iron-storage protein, is built from a bundle of α helices.

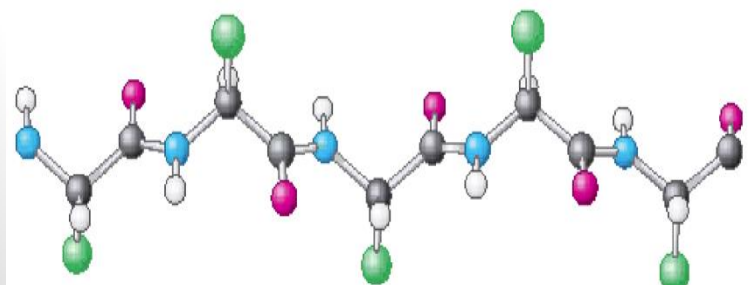
Beta Sheets Are Stabilized by Hydrogen Bonding Between Polypeptide Strands

- Pauling and Corey discovered another periodic structural motif
- They named the β pleated sheet (β because it was the second structure that they elucidated, the α helix having been the first).
- The β pleated sheet (or the β sheet) differs markedly from the rod-like α helix.
- A polypeptide chain, called a β strand, in a β sheet is almost fully extended rather than being tightly coiled as in the α helix.
- A range of extended structures are sterically allowed.
- The distance between adjacent amino acids along a β strand is approximately 3.5 Å (1.5 Å along an α helix).

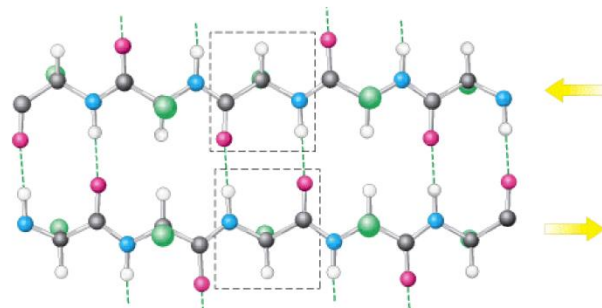


The red area shows the sterically allowed conformations of extended, β -strand-like structures.

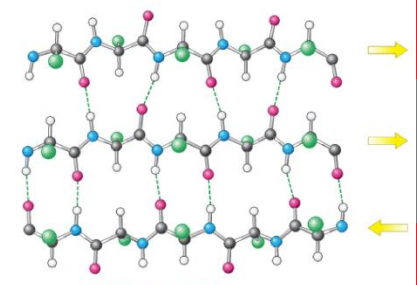
- A β sheet is formed by linking two or more β strands by hydrogen bonds.
- Adjacent chains in a β sheet can run in opposite directions (antiparallel β sheet) or in the same direction (parallel β sheet).
- In the *antiparallel* arrangement,
 - the NH group and the CO group of each amino acid are respectively hydrogen bonded to the CO group and the NH group of a partner on the adjacent chain.
- In the *parallel* arrangement,
 - the hydrogen-bonding scheme is slightly more complicated.
- For each amino acid,
 - the NH group is hydrogen bonded to the CO group of one amino acid on the adjacent strand,
 - whereas the CO group is hydrogen bonded to the NH group on the amino acid two residues farther along the chain.
- Many strands (4 or 5 but as many as 10 or more), can come together in β sheets.



The side chains (green) are alternately above and below the plane of the strand.

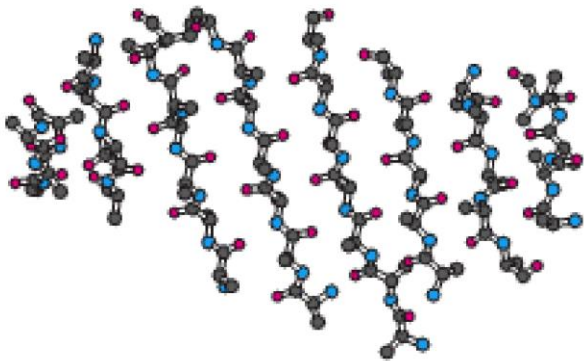


An Antiparallel β Sheet. Adjacent β strands run in opposite directions. Hydrogen bonds between NH and CO groups connect each amino acid to a single amino acid on an adjacent strand, stabilizing the structure.



Structure of a Mixed β Sheet.

- In schematic diagrams, β strands are usually depicted by broad arrows pointing in the direction of the carboxyl-terminal end to indicate the type of β sheet formed parallel or antiparallel.
- More structurally diverse than α helices, β sheets can be relatively flat but most adopt a somewhat twisted shape.
- The β sheet is an important structural element in many proteins.
 - fatty acid-binding proteins,
 - important for lipid metabolism



A ball-and-stick model.
Twisted β Sheet.



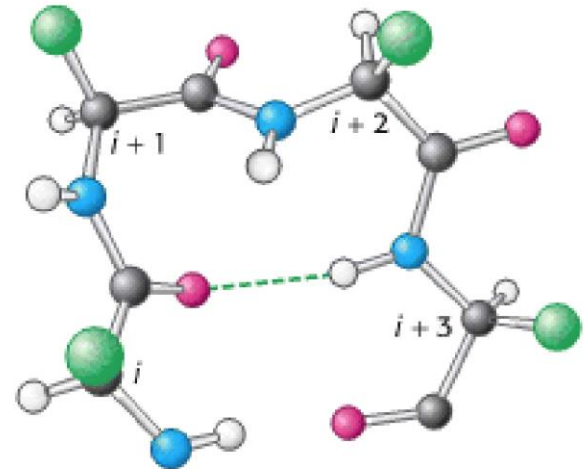
A schematic model.



The schematic view rotated
by 90 degrees

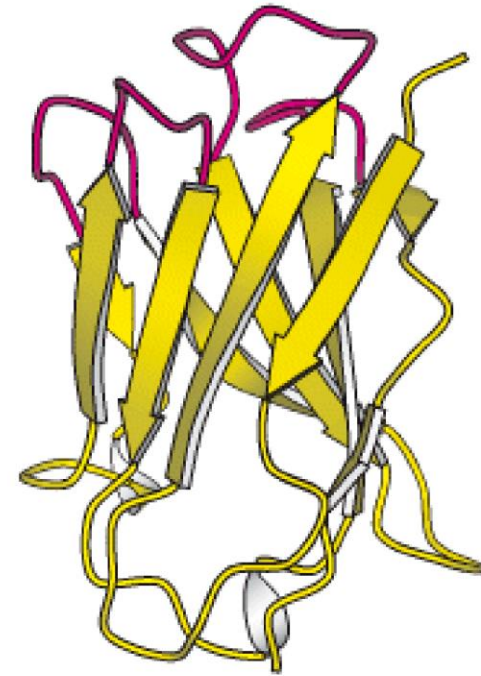
Reverse Turns and Loops

- Most proteins have compact, globular shapes, requiring reversals in the direction of their polypeptide chains.
- Many of these reversals are accomplished by a common structural element called the reverse turn (the **b turn** or **hairpin bend**).
- In many reverse turns, the CO group of residue i of a polypeptide is hydrogen bonded to the NH group of residue $i + 3$.
- This interaction stabilizes abrupt changes in direction of the polypeptide chain.



The CO group of residue i of the polypeptide chain is hydrogen bonded to the NH group of residue $i + 3$ to stabilize the turn.

- More elaborate structures are also responsible for chain reversals.
- These structures are called **loops** or sometimes ω **loops** (omega loops).
- ω loops do not have regular, periodic structures.
- Loop structures are often rigid and well defined.
- Turns and loops invariably lie on the surfaces of proteins and often participate in interactions between proteins and other molecules.
- The distribution of α helices, β strands, and turns along a protein chain is often referred to as its **secondary structure**.

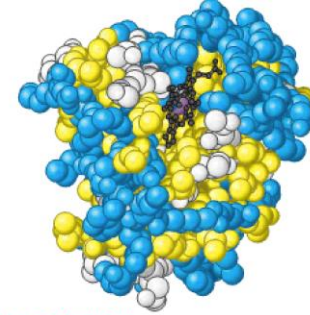


A part of an antibody molecule has surface loops (shown in red) that mediate interactions with other molecules.

Tertiary Structure: Water-Soluble Proteins Fold Into Compact Structures with Nonpolar Cores

- Tertiary structure refers to the *spatial arrangement of amino acid residues that are far apart in the sequence* and to the pattern of disulfide bonds.
- Amino acids are grouped together in a **complete protein**.
- Myoglobin is the first protein to be seen in atomic detail.

Myoglobin



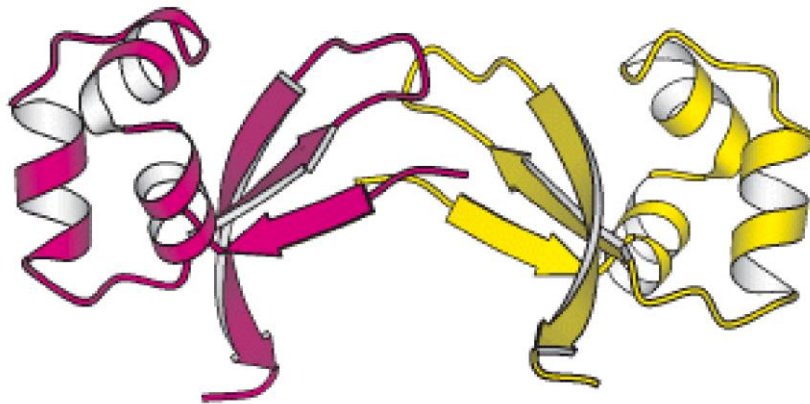
Amino Acids in Myoglobin. (A) A space-filling model of myoglobin with hydrophobic amino acids shown in yellow, charged amino acids shown in blue, and others shown in white.

- It is a **oxygen carrier** in muscle,
- It is a **single** polypeptide chain of 153 amino acids.
- The capacity to bind oxygen depends on the
 - presence of **heme** (a non-polypeptide **prosthetic**) group consisting of protoporphyrin IX
 - a central **iron** atom.
- About 70% of the main chain is folded into eight α helices and the rest of the chain forms turns and loops between helices.
- The overall course of the polypeptide chain of a protein is referred to as its ***tertiary structure***.
- The interior consists almost entirely of nonpolar residues (leucine, valine, methionine and phenylalanine).
- **Charged** residues (aspartate, glutamate, lysine, and arginine) are **absent** from the inside of myoglobin.
- The only **polar** residues **inside** are ***two histidine residues***, which play critical roles in binding iron and oxygen.
- The **outside** of **myoglobin**, on the other hand, consists of both **polar** and **nonpolar** residues.

- In an **aqueous** environment, protein folding is driven by the strong tendency of **hydrophobic residues** to be **excluded** from water.
- A system is more **thermodynamically stable** when hydrophobic groups are **clustered** rather than extended into the aqueous surroundings.
- The polypeptide chain therefore folds so that its hydrophobic side chains are buried and its polar, charged chains are on the surface.
- Many α helices and β strands are **amphipathic**; hydrophobic face points into the protein interior and a more polar face points into solution.
- An unpaired peptide NH or CO group prefers water.
- This pairing is neatly accomplished in an α helix or β sheet.
- Van der Waals interactions between tightly packed hydrocarbon side chains also contribute to the stability of proteins.

Quaternary Structure: Polypeptide Chains Can Assemble Into Multi-subunit Structures

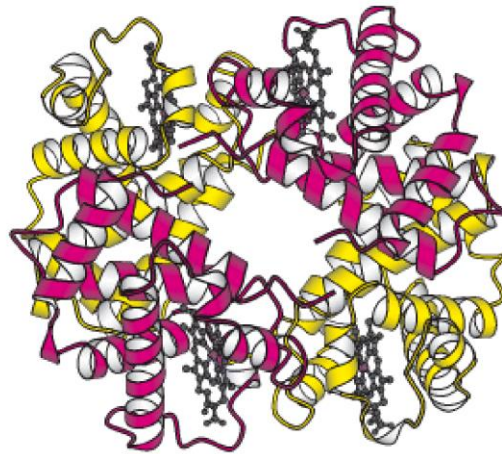
- Proteins also contain more than one polypeptide chain.
- Each polypeptide chain in it, is called a *subunit*.
- Quaternary structure refers to *the spatial arrangement of subunits and the nature of their interactions*.
- The simplest sort of quaternary structure is a dimer, consisting of *two identical subunits*.
- This organization is present in the DNA-binding protein *Cro* found in a bacterial virus called λ .



The Cro protein of bacteriophage λ is a dimer of identical subunits.

Human hemoglobin

- It is the oxygen-carrying protein in blood
- It consists of two subunits of
 - one type (designated α) and
 - another type (designated β).
- The hemoglobin molecule exists as an $\alpha_2 \beta_2$ tetramer.
- Subtle changes in the arrangement of subunits within the hemoglobin molecule allow it to carry oxygen from the lungs to tissues with great efficiency.



The $\alpha_2 \beta_2$ Tetramer of Human Hemoglobin. The structure of the two identical α subunits (red) is similar to but not identical with that of the two identical β subunits (yellow). The molecule contains four heme groups (black with the iron atom shown in purple).

- **Exceptions are:**

- Glycine

- Does not have a side chain.
- Its α -carbon contains two hydrogens.

- Proline

- The nitrogen is part of a ring.

- Each amino acid has an *amine group* at one end and an *acid group* at the other and a distinctive side chain.

- The backbone is the same for all amino acids while the *side chain differs* from one amino acid to the next.

- *All* of the 20 amino acids *except glycine* are of the *L-configuration*, as for all but one amino acid the α -carbon is an asymmetric carbon.

- Because *glycine* does not contain an asymmetric carbon atom, it is *not optically active* and, thus, it is neither D nor L.

Thank You