E-content

M.Sc. Zoology (Semester II)
CC7- Biochemistry

Unit: 3.1

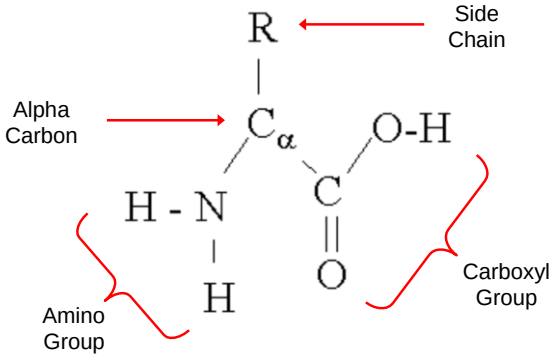
Primary, secondary, tertiary, quaternary and domain structure of proteins

Dr Gajendra Kumar Azad
Assistant Professor
Post Graduate Department of Zoology
Patna University, Patna
Email: gajendraazad@outlook.com

Protein are built from a Repertoire of 20 amino acids

Amino-acids are natural compounds composed of amine (-NH2) and carboxylic acids (-COOH) functional groups, linked to the same carbon atom.

The key elements of an amino acid are carbon, hydrogen, oxygen and nitrogen.



Amino acids are the building blocks of proteins.

Amino acids in the solution at neutral pH exist predominantly as dipolar ions (also called zwitterions). In the dipolar form, amino group is protonated (-NH3+) and the caroxyl group is deprotonated (-coo-).

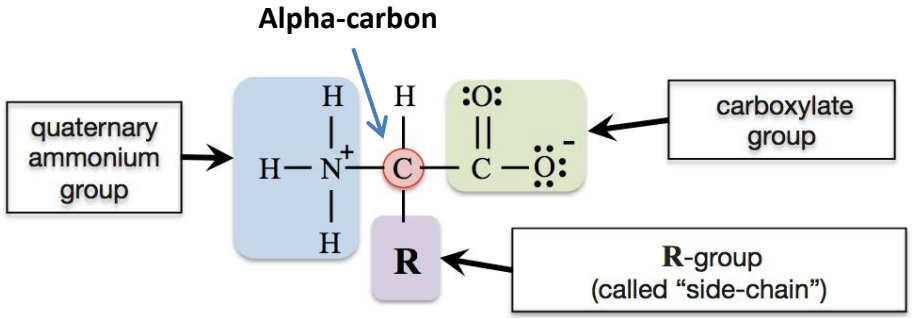


Figure 1: Basic structure of amino acids

There are 20 different, naturally occurring amino acids.

The properties of each amino acid are determined by its specific side chain

Amino acids are classified into various groups as mentioned below

Amino Acid Class	Side Chain Polarity	Side-Chain Charge at Physiological pH
Nonpolar	nonpolar (hydrophobic side-chain)	zero
Polar neutral	polar (hydrophilic side-chain)	zero
Polar acidic	polar (hydrophilic side-chain)	negative
Polar basic	polar (hydrophilic side-chain)	positive

Nonpolar Amino Acids

Nonpolar amino acids have *nonpolar (hydrophobic) side-chains* and their predominant forms have *uncharged* side-chains at physiological pH.

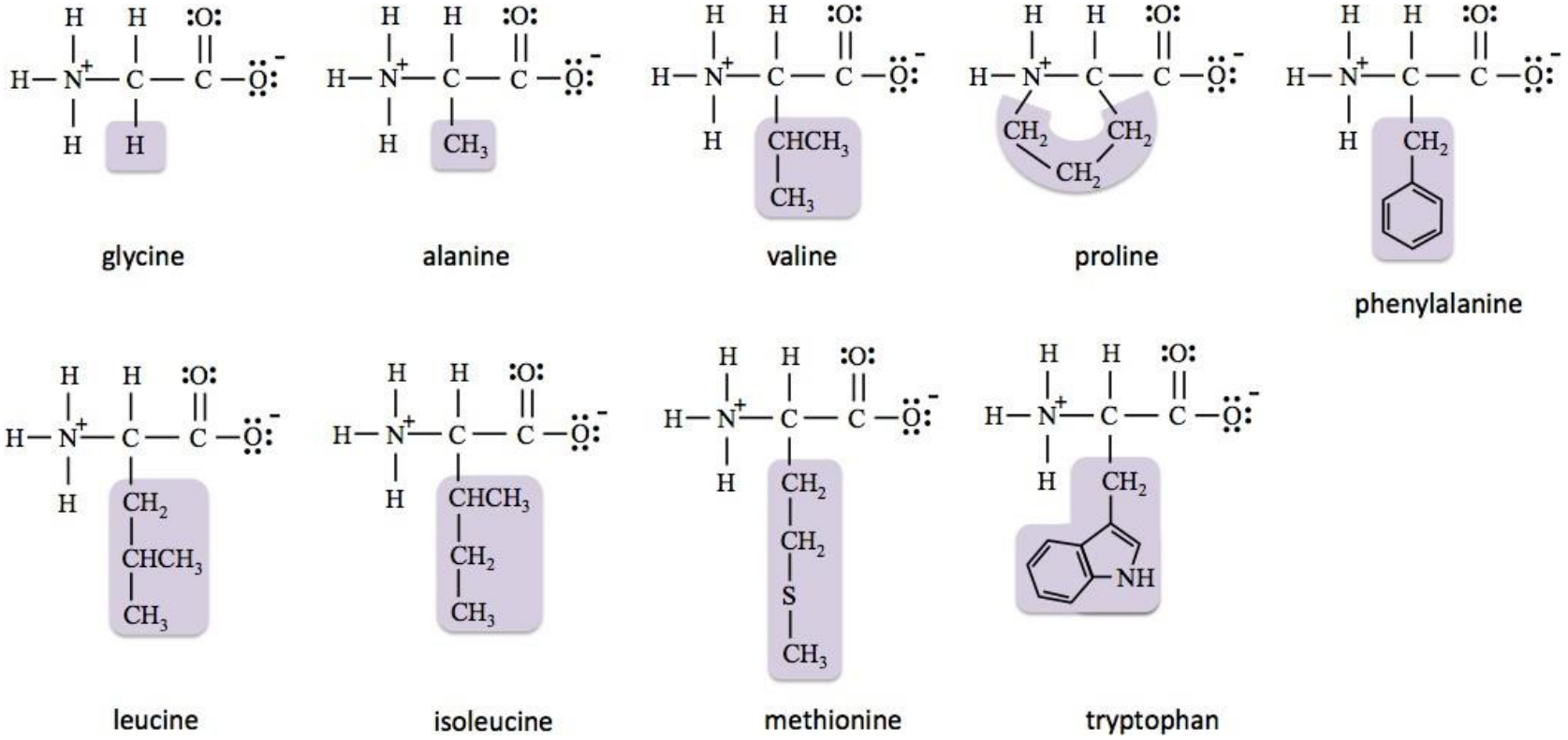
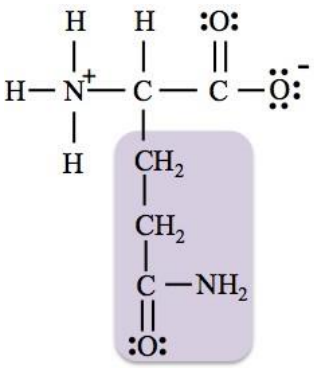


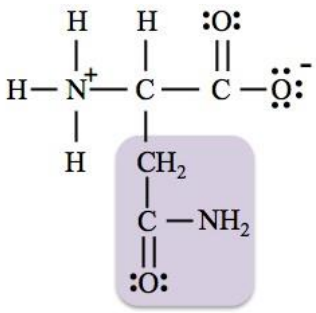
Figure 2: non-polar amino acids

Polar Neutral Amino Acids

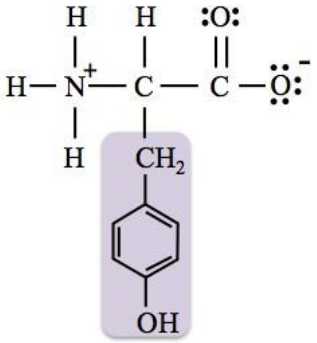
Polar neutral amino acids have *polar (hydrophilic) side-chains* and their predominant forms have *uncharged* side-chains at physiological pH.



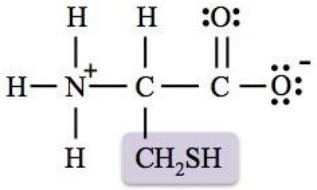
glutamine



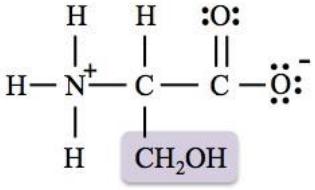
asparagine



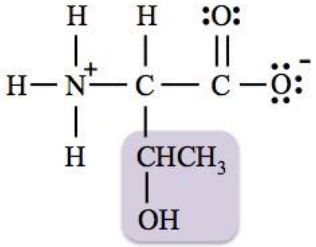
tyrosine



cysteine



serine

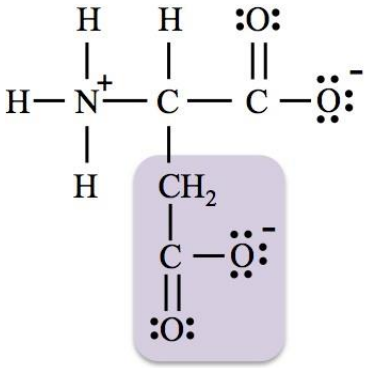


threonine

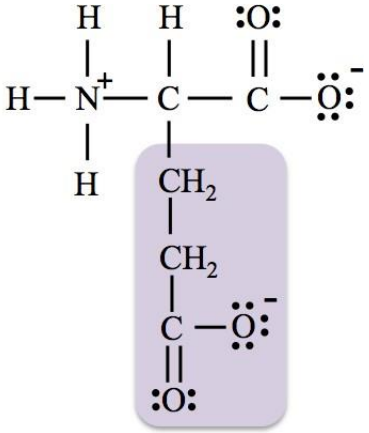
Figure 3: Polar neutral amino acids

Polar Acidic Amino Acids

Polar acidic amino acids have *polar (hydrophilic) side-chains* and, their predominant forms have side-chains with negative *formal charge* at physiological pH.



aspartic acid



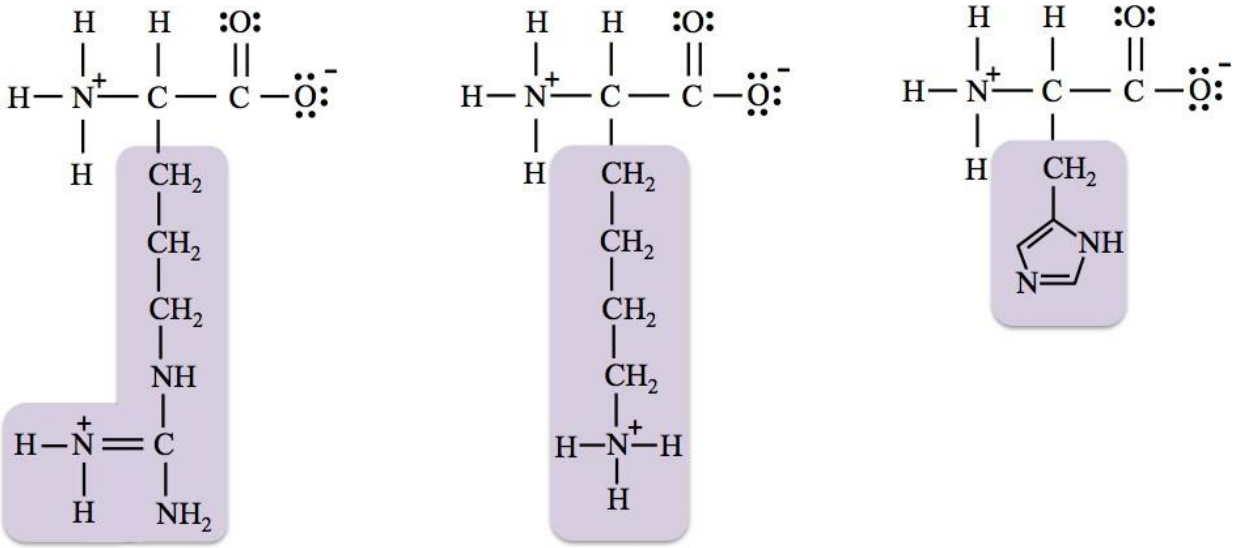
glutamic acid

Figure 4: Polar acidic amino acids

Polar Basic Amino Acids

Polar basic amino acids have *polar (hydrophilic) side-chains* and, except for *histidine*, their predominant forms have side-chains with positive *formal charge* at physiological pH.

This formal charge is from a *quaternary ammonium group*.



arginine lysine histidine

Figure 5: Polar basic amino acids

Primary structure of polypeptide

Amino acids are linked together by peptide bonds to form polypeptide chains
Proteins are linear polymers formed by linking the α- carboxyl of one amino acid to the α- amino group of another amino acid. This type of linkage is called peptide bond (or an amide bond).

Formation of a Peptide Bond

Step 1: The *oxygen of first amino acid (*from the *carboxylate group) and two* hydrogen atoms (from the *ammonium group) of second* amino acid combines to form a water molecule.

Step 2: A *new bond* is made between the carbonyl carbon and the nitrogen.

The *new bond* between the two amino acid residues is called a peptide bond.

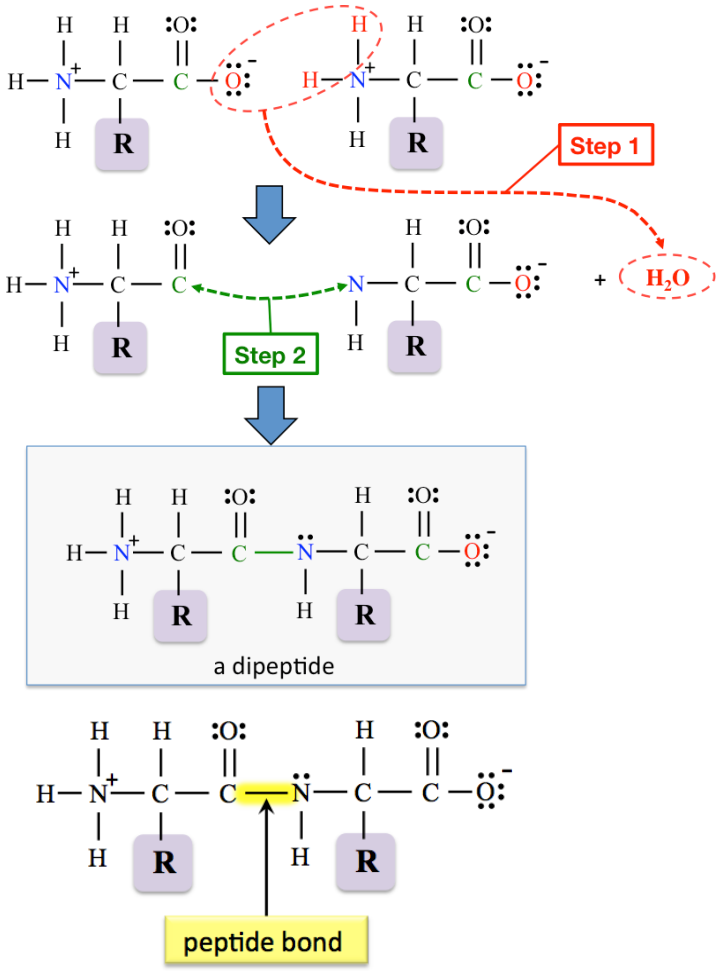


Figure 6: amino acids are joined by peptide bonds

Primary structure of polypeptide

The peptide bonds are quite stable kinetically because the rate of hydrolysis is extremely slow; the lifetime of a peptide bond in aqueous solution in the absence of a catalyst is around 1000 years.

A series of amino acids joined by peptide bonds form a polypeptide chain, and each amino acid unit in polypeptide is called residue.

A polypeptide chain has polarity because its ends are different: an α-amino group is present at one end and a carboxyl group at the other.

By convention, the amino end is taken to be the beginning of a polypeptide chain and called N-terminus. Similarly, the carboxyl end is called C-terminus of polypeptide chain

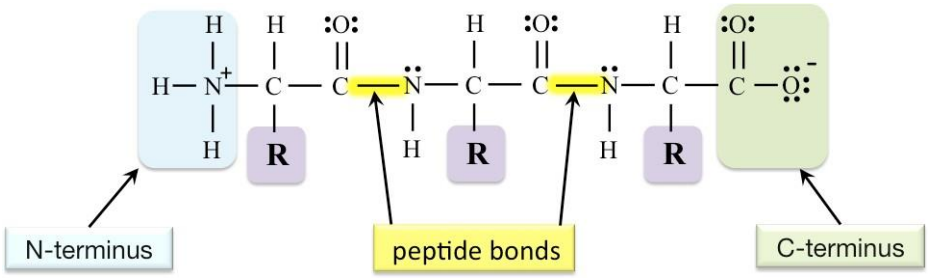


Figure 7: Peptide bonds in polypeptide chain

Primary structure of polypeptide

Proteins are composed of a long polypeptide chains.

Chains that are less than 40-50 residues are often referred to as polypeptide chains since they are too small to form a functional domain.

Larger than this size, they are called proteins.

The structure, function and general properties of a protein are all determined by the sequence of amino acids that make up its primary sequence.

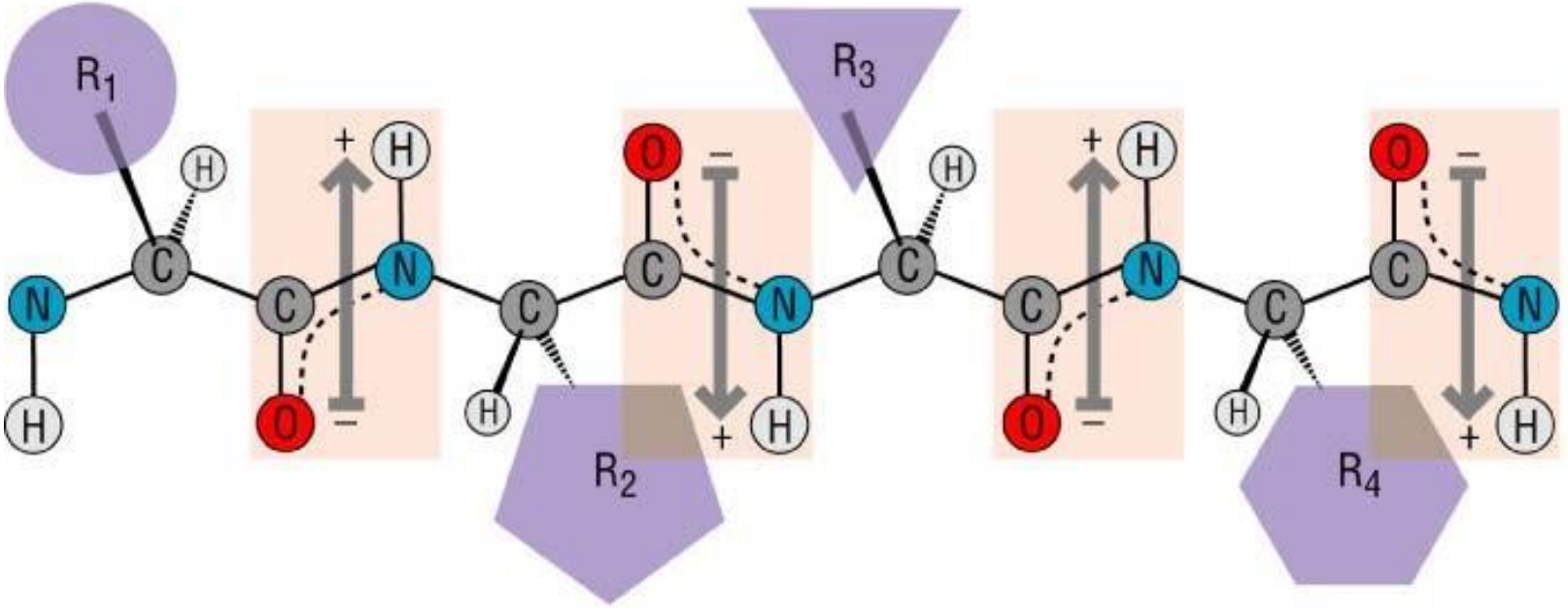


Figure 8: Primary structure of polypeptide chain

Polypeptide chains are flexible yet conformationally restricted

The geometry of protein backbone reveals several important features.

1. The peptide bond is essentially planar: for a pair of amino acids linked by peptide bond, six atoms lie in the same plane; the α-carbon and CO group of the fist amino acid and the NH group and α-carbon atom of the second amino acid.

2. The peptide bond have considerable double bond character., which prevents rotation about this bond. The C-N distance in a peptide bond is typically 1.32Å, which is between the values expected for a C-N single bond (1.49Å) and a C=N double bond (1.27Å).

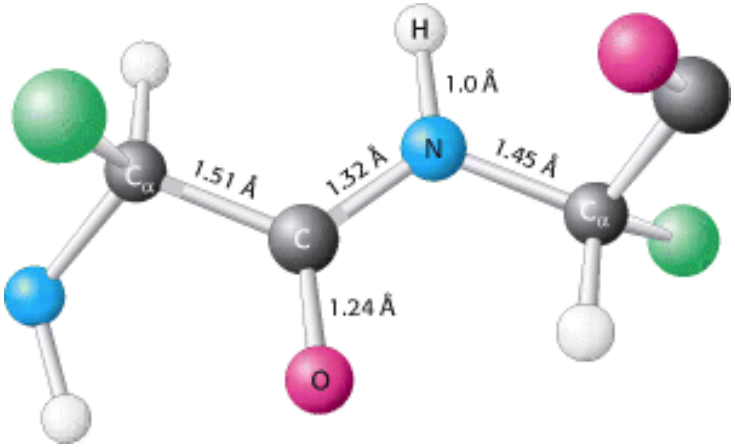


Figure 9: Typical bond length within a peptide unit

3. The peptide bond is uncharged, allowing polymers of amino acids linked by peptide bonds to form tightly packed globular structure.

Trans and cis peptide bonds

Two configuration are possible for a planar peptide bonds; trans and cis.

In the *trans* configuration, the two α-carbon atoms are on the opposite sides of the peptide bond.

In *cis* configuration, the two α-carbon atoms are on the opposite sides of the peptide bond.

Almost all peptides are in trans.

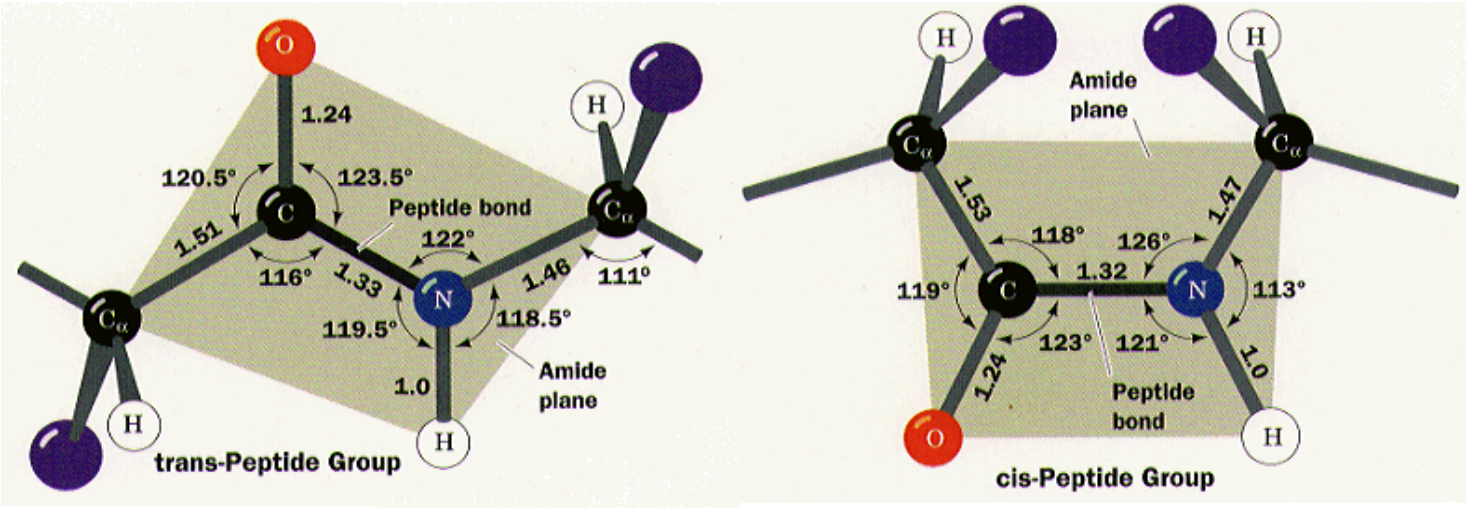


Figure 10: trans is strongly favored over cis form

Rotation of bonds in a polypeptide

Peptide conformations is defined by three dihedral angles (also known as torsion angles) called ϕ (phi), ψ (psi), and ω(omega), reflecting rotation about each of the three repeating bonds in the polypeptide backbone.

Within the peptide bond, the bond between the amino group and the α-carbon atom and between the α-carbon atom and the carbonyl group are pure single bonds.

The two adjacent rigid peptide unit s may rotate about these bonds, taking various orientations.

The angle of rotation between the nitrogen and the α-carbon atom is called ϕ.

The angle of rotation between the α-carbon atom and the carbonyl group is called ψ.

Ω is not often considered. It involves the carbon and nitrogen atom of peptide bond, where rotation is constrained.

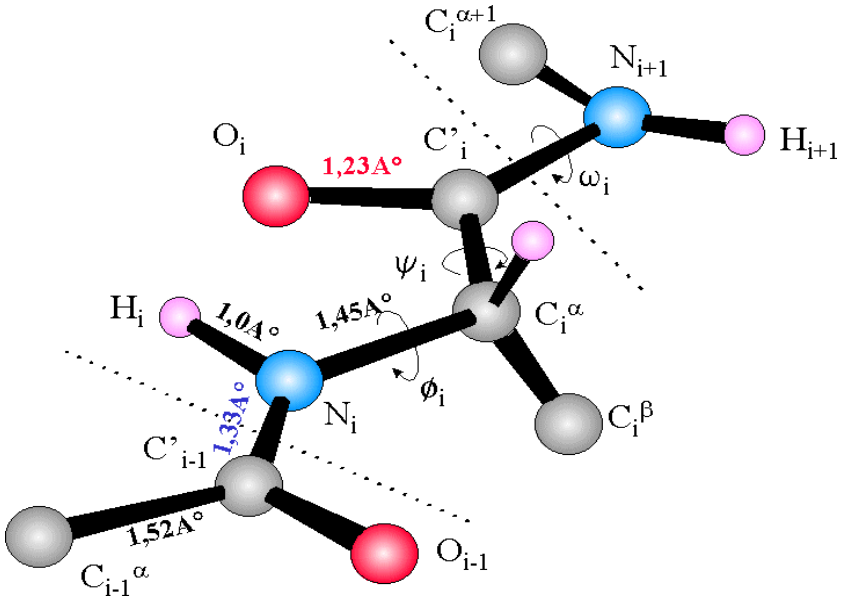


Figure 11: dihedral angles in peptide backbone

Ramachandran Plot

In principle, the ϕ and ψ can have any value between -180° to +180°, but many values are prohibited by steric interference between atoms in the polypeptide backbone and amino acid side chains (glycine is an exception).

Three- quarters of the possible (ϕ and ψ) combinations are excluded simply by local steric clashes based on calculations using known van der Waals radii and dihedral angles.

The area shaded dark blue represents conformations that involves no steric overlap. Medium blue represents conformations allowed at the extreme limits for unfavourable atomic contacts. Lightest blue represents conformations that are permissible if a little flexibility is allowed in the dihedral angle. Yellow region are conformation that are not allowed.

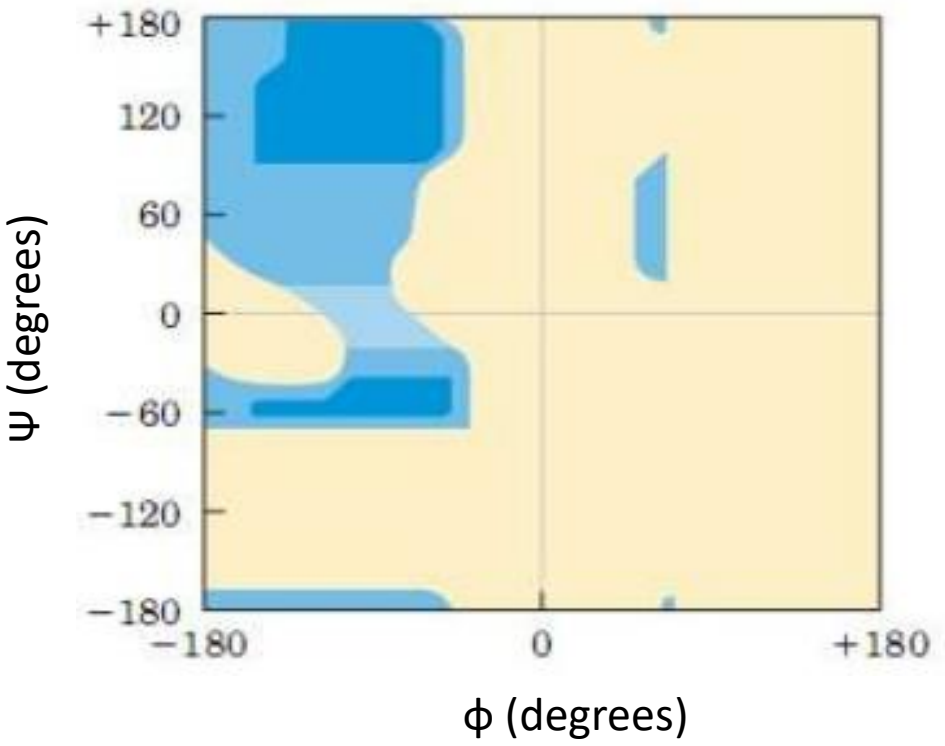


Figure 12: Ramachandran plot for l-Alanine residues

Protein Secondary Structure

The term secondary structure refers to any chosen segment of a polypeptide chain and describes the local spatial arrangement of its main-chain atoms, without regard to the conformation of its side chains or its relationship to other segments.

The two most common regular secondary structure elements are α helices and β sheets, formed by repeating amino acids with the same (ϕ and ψ) angles.

There are other secondary structure elements such as β turns, Ω loops.

α helices , β sheets, and turns are formed by regular pattern of hydrogen bond between the peptide N-H and C=O groups of amino acids that are near one another in linear sequence.

The path of polypeptide backbone in almost any protein is not random; rather, it is typically unchanging and highly specific to the structure and function of that particular protein.

Secondary structure: α helices

The α helix is the most common secondary structure.

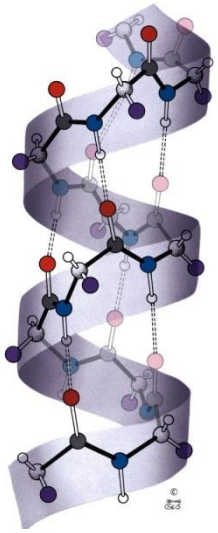
They are regular structures that repeats every 5.4 Å.

The poly peptide backbone is tightly wound around an imaginary axis down longitudinally through the middle of the helix, and the R-group of amino acid residues protrude outward from the helical backbone.

The amino acid residues in the α helix have conformations with ϕ= -57° and ψ= -47°, and each helical turn includes 3.6 amino acid residues.) angles.

Linus Pauling and Corey *were pioneer in proposing* α helix structure and build model in *1951.*

Helical twist of the α helix found in all protein is **right handed**



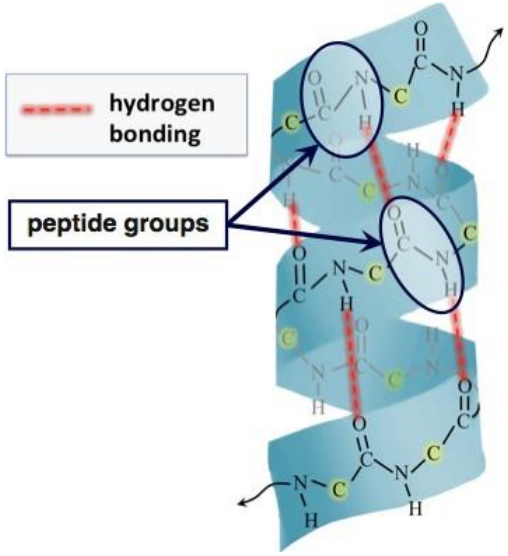


Figure 13: alpha helix structure

Secondary structure: α helices

The α helix is stabilised by hydrogen bonds.

The hydrogen bonds are formed between hydrogen attached to electronegative nitrogen atom of the peptide linkage and the electronegative carbonyl oxygen atom of the fourth amino acid on the amino terminal side of the peptide bond.

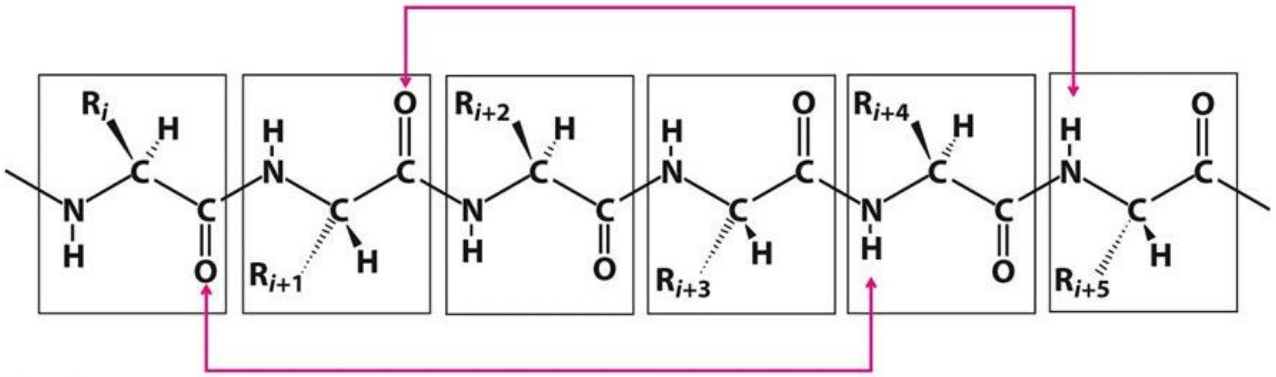


Figure 14: Hydrogen bonding pattern in alpha helix

Within the α helix every peptide bond participates in hydrogen bonding

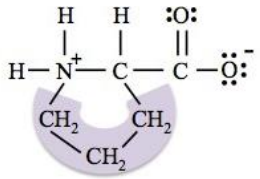
All hydrogen bonds together provide the stability to the α helix

Secondary structure: α helices

Not all polypeptide can form a stable helix. Interactions between amino acids can stabilize or destabilize it.

Such as, if a polypeptide chain has a long block of Glutamic acid residues, this segment will not form an α helix. Negative charged carboxyl group of the adjacent glutamic residues repel each other and strongly prevents formation of the α helix.

A polypeptide rich in proline residue also not form α helix. In proline, the nitrogen atom is part of rigid ring and rotation about the N- Cα bond is not possible. Therefore, proline introduces destabilising kink in the polypeptide.



proline

The bulk shape of Asn, Ser, Thr, and Cys residues can also destabilise an α helix if they are close together in the chain.

Both right and left handed helices lie in regions of allowed conformations in the Ramachandran diagram. However, essentially all helices in protein are right handed

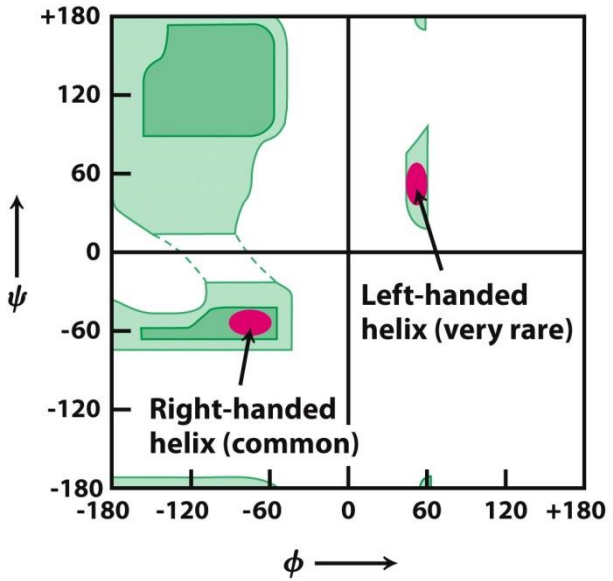


Figure 15: Ramachandran diagram for helices

Secondary structure: β sheets

The β conformations organises polypeptide chains into sheets. The β conformations is an extended form of polypeptide. The backbone of the zigzag rather than helical structures.

The zigzag polypeptide chains can be arranged side by side to form a structure resembling a series of pleats called β sheets.

β sheets are composed of two or more polypeptide chains called β strands.

The structure is stabilized by hydrogen bonds. The H-bond is formed between the adjacent segments of the chains.

The R-groups of adjacent amino acids protrude from the zigzag structure in opposite directions creating the alternating patterns.

The adjacent polypeptide chain in a β sheet can be either parallel or antiparallel (having the same or opposite amino-to-carboxyl orientations, respectively).

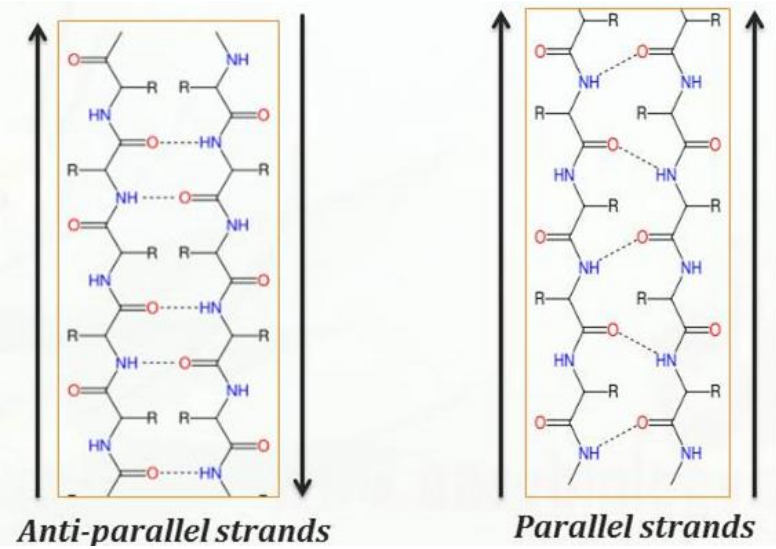
The idealized structures corresponds to ϕ= -119° and ψ= +113 (parallel) and ϕ=-139, ψ=+135 (antiparallel); these values vary somewhat in real proteins, resulting into some structural variations.

Secondary structure: β sheets

Two types of β conformations organises polypeptide chains into sheets.

a) anti-parallel β sheets

b) Parallel β sheets



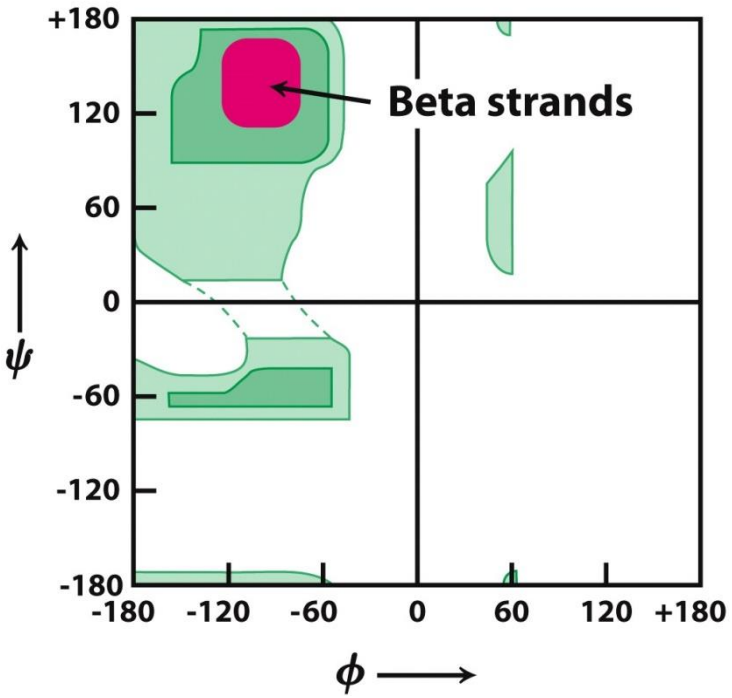


Figure 16: Ramachandran diagram for β strands.

The red areas show the sterically allowed conformations of extended, β strand-like structures

Secondary structure: β turns and Ω loops

Most proteins have compact folded structures and nearly one-third of the amino acid residues are in turns or loops where polypeptide chain reverse directions.

These are connecting elements that links successive runs of α helix and β stands.

The most common structural element is called β turn or hairpin turn or reverse turn. In many reverse turn the CO group of first residue of a polypeptide is hydrogen bonded to NH group of fourth. This interaction stabilises abrupt changes in the direction of polypeptide chain.

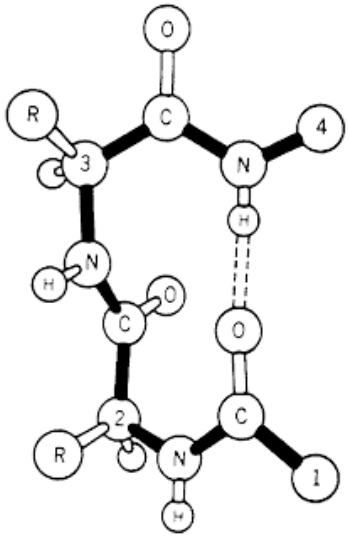


Figure 17: Turn structure

The more elaborate structures comprises the Ω loops. The loops do not have regular, periodic structures. However, loop structures are often rigid and well defines.

Turns and loops are mostly present on the surface of the proteins and thus often participates in interactions between proteins and other molecules.

Tertiary and quaternary structure of proteins

The overall three dimensional arrangement of all atoms in a protein is referred as ***tertiary structure*** of the protein. It also includes the spatial arrangement of amino acids that are situated distant in the polypeptide chain.

The amino acids that are far apart in polypeptide sequence and are in different types of secondary structures may interact within the completely folded structure of the protein.

Tertiary structure is stabilized by covalent bond other than several kinds of weak interactions. One of the covalent bond found in tertiary structures are disulfide bonds (between two Cys residues).

Proteins with a single subunit can have up to tertiary structure.

Some proteins contain two or more separate polypeptide chains, or subunits, which may be identical or different. The arrangement of these subunits in three-dimensional complexes constitute ***quaternary structures***.

Tertiary and quaternary structure of proteins

In considering higher levels of structure in proteins, it is useful to classify proteins into two major groups.

1. Fibrous protein
2. Globular proteins

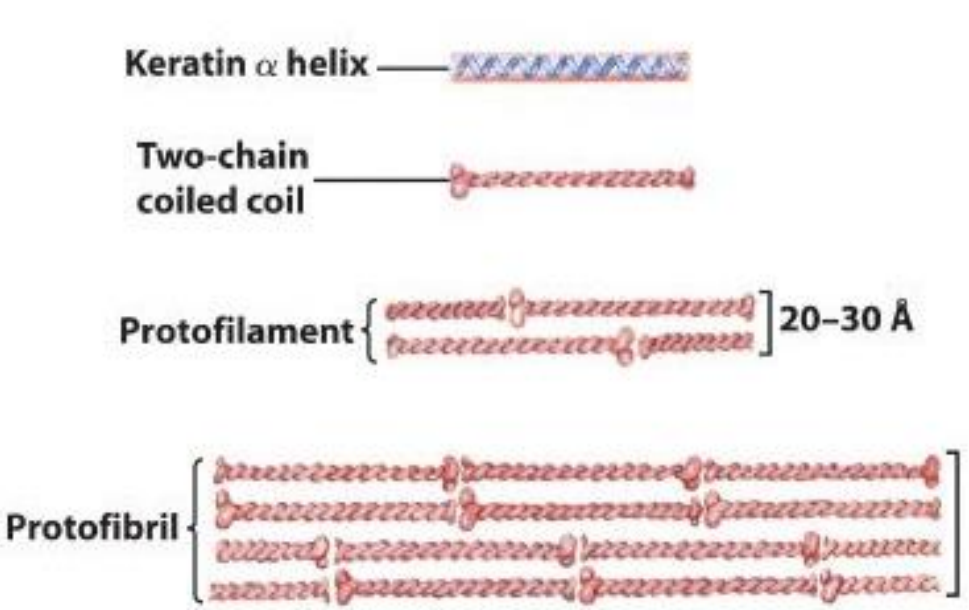
1. Fibrous protein

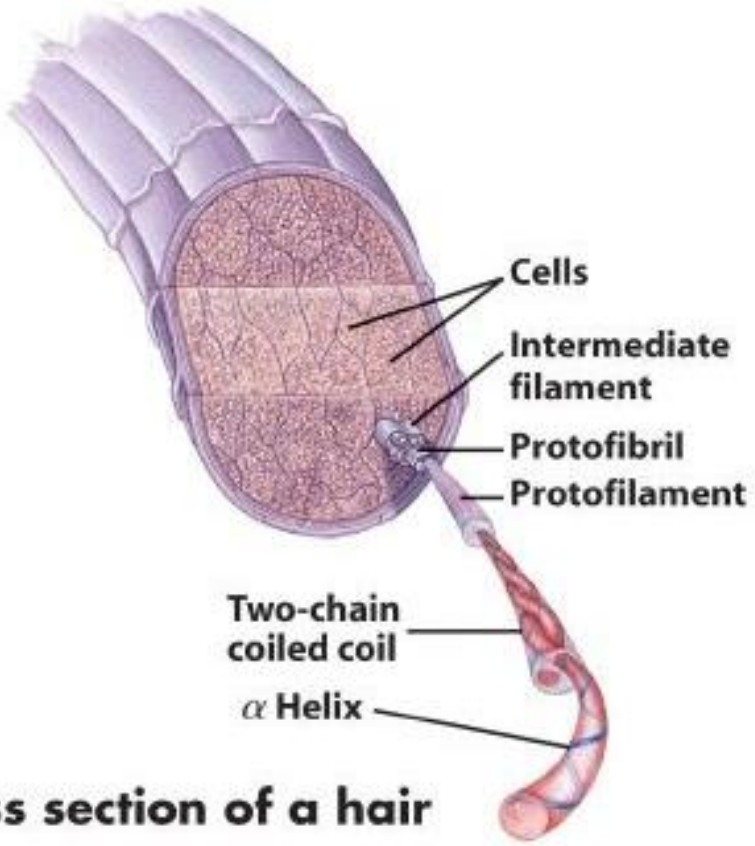
In fibrous proteins the polypeptide chains are arranged in long strands or sheets. The fibrous proteins usually consists largely of a single type of secondary structures, and their tertiary structures are relatively simple.

Few example of fibrous proteins are

A) α-Keratin
B) Collagen
C) Silk fibroin

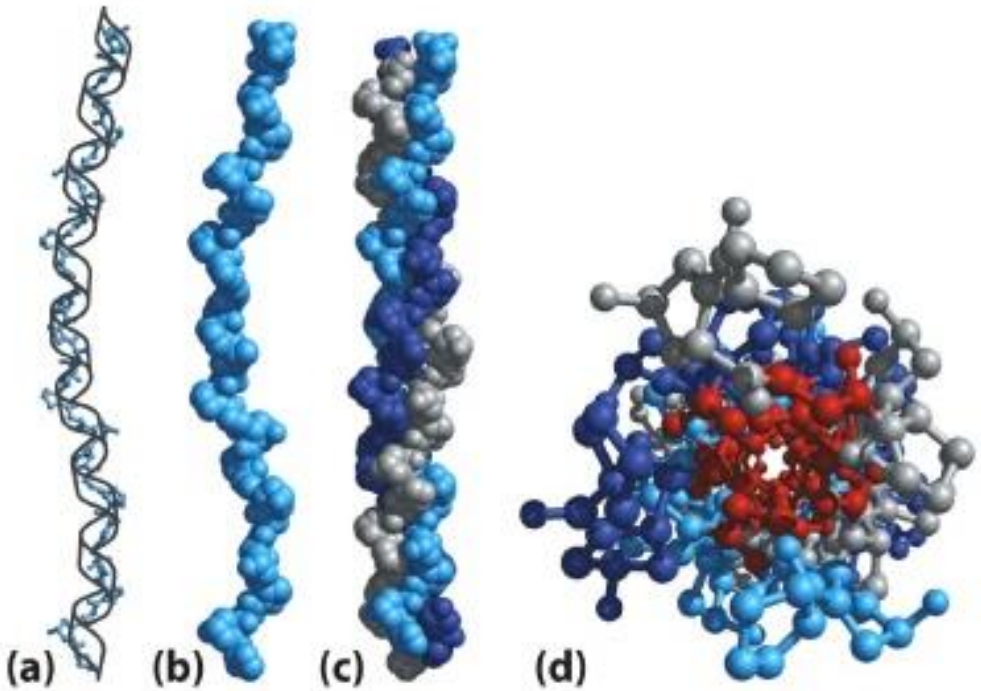
A) α-Keratin



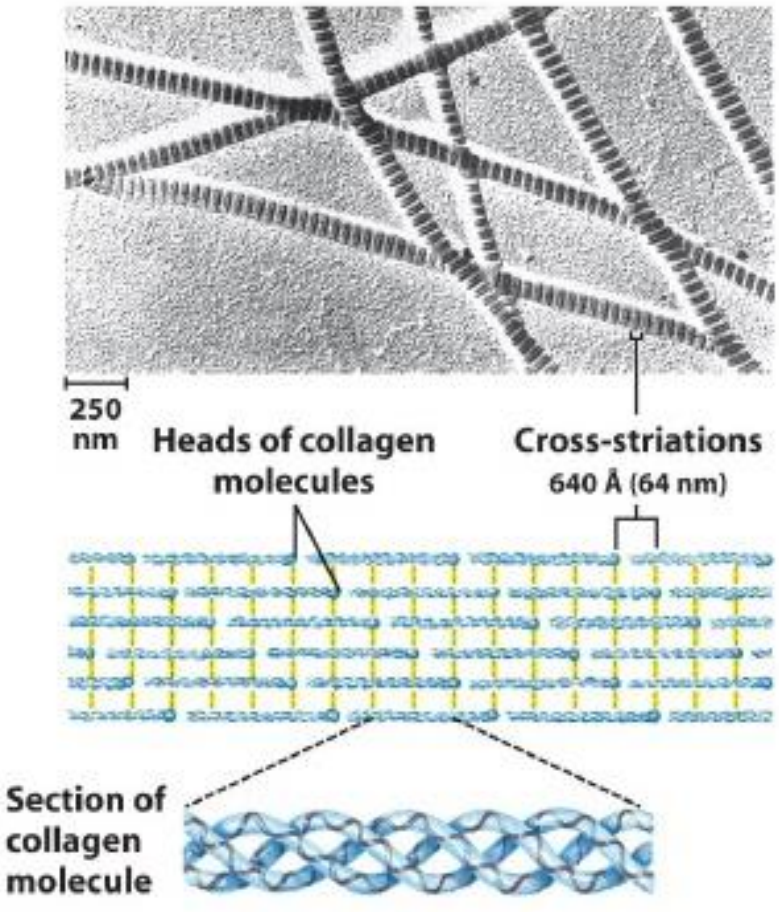


-Alpha keratins belong to the intermediate filament (IF) protein family.
-An all α-helix protein.
-Rich in hydrophobic amino acids: Ala, Val, Leu, Ile, Met, Phe

B) Collagen

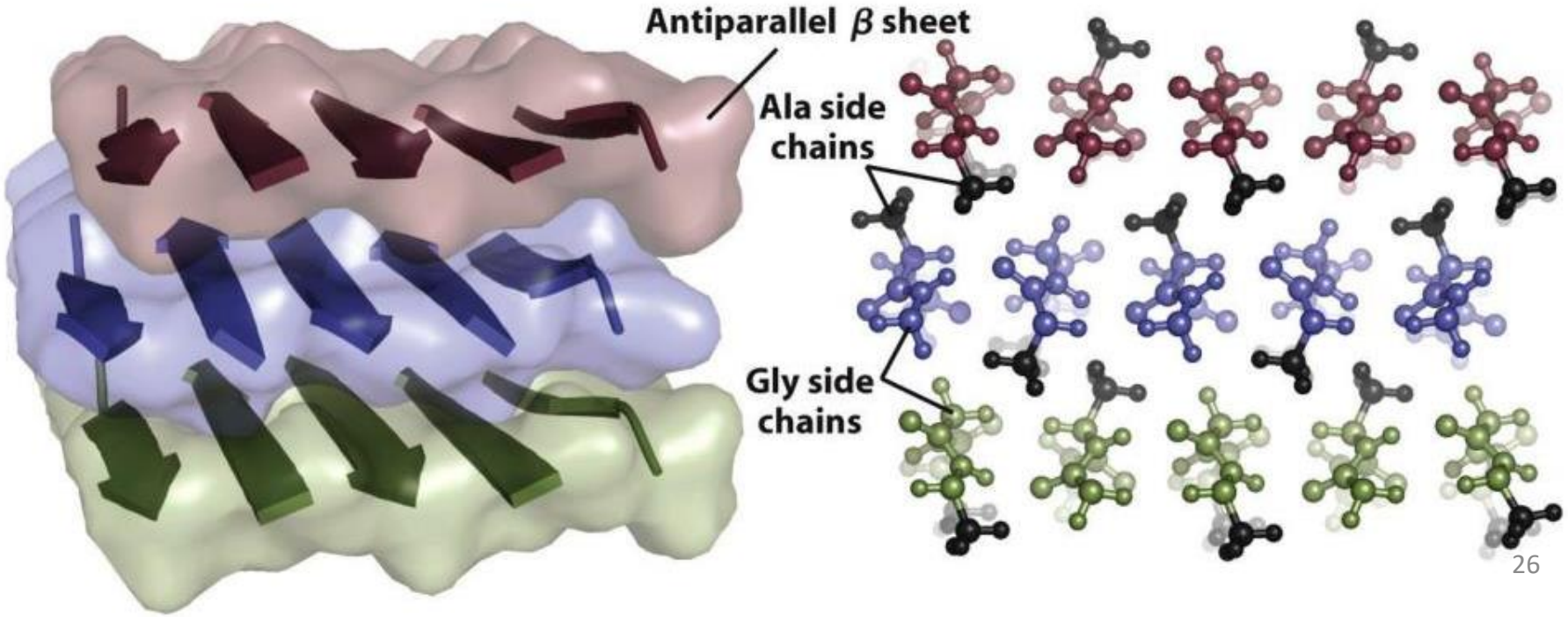


-A repeating tripeptide: Gly-X-Pro or Gly-X-Hyp
-Left-handed helical structure: 3 residues per turn
-3 helices wrap around each other in a right-handed twist
-Great tensile strength!



C) Silk Fibroin

- Fibroin is the main protein in silk from moths and spiders
- Antiparallel β sheet structure
- Small side chains (Ala and Gly) allow the close packing of sheets
- Structure is stabilized by
  - hydrogen bonding within sheets



2. Globular proteins

In a globular proteins, different segments of the polypeptide chain (or multiple polypeptide chains) fold back to each other, generating a more compact shape then is seen in fibrous proteins.

The 3-D structure of typical globular protein can be considered an assemblage of polypeptide segments in the α-helical and β-sheet conformations, linked by connecting segments.

The folding also provides the structural diversity necessary for proteins to carryout wide array of biological functions.

Globular proteins includes enzymes, transport proteins, motor proteins, regulatory proteins, immunoglobulin's and many other proteins with other functions.

One of the most common example of globular protein is

Myoglobin

Myoglobin

Myoglobin is relatively small oxygen binding protein of muscle cells.

Myoglobin contains a single polypeptide chain of 153 amino acid residues of known sequence and a single polypeptide chain of 153 amino acid residues and a single iron protoporphyrin, or heme, group.

The backbone of the myoglobin molecule consists of 8 straight segments of α helix interrupted by bends, some of which are β turns.

The flat heme group rests in a pocket in the myoglobin molecule.

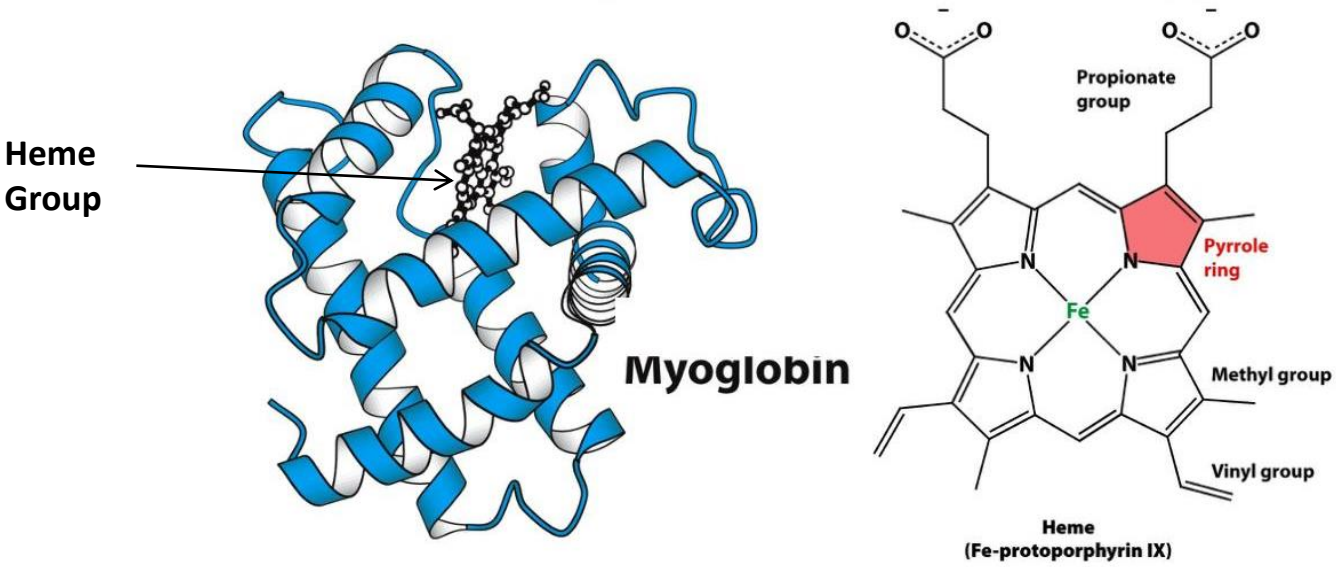


Figure 18: Myoglobin structure with heme group

Structural patterns of tertiary structure

The tertiary structure of proteins have distinct localized structures or folding patterns which is termed as A) **motif** and B) **domains**

A) Motif

A motif is similar 3-D structure conserved among different proteins that serves a similar function. They are recognizable regions of protein structure that may (or may not) be defined by a unique chemical or biological function.

The motifs are also known as supersecondary structure or fold.

Some common examples are:

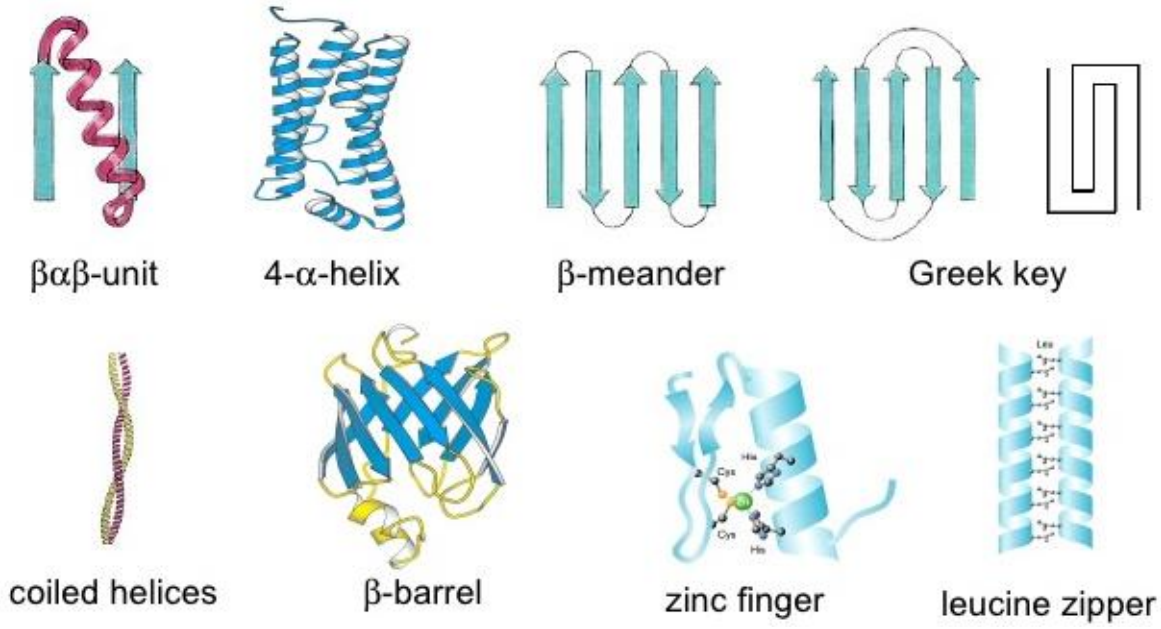


Figure 19: various kind of motifs in protein structure

Structural patterns of tertiary structure

B) Domains

Domains, on the other hand, are regions of a protein that has a specific function and can (usually) function independently of the rest of the protein.

For example a single protein can have a DNA binding domain located towards the N terminus of the protein, and a catalytic domain that is located closer to the C-terminus.

Theoretically the domains can be separated from each other and they can still function, such as the DNA binding domain will still bind DNA and the catalytic domain will still perform catalysis.

A large globular proteins may consist of several domains linked by stretches of polypeptide. Separate domain may have distinct functions .

One of the example is Glyceraldehyde-3-phosphate dehydrogenase

Tertiary structure: Glyceraldehyde-3-phosphate dehydrogenase

G3P dehydrogenase is comprised of two distinct globular domains.

The dinucleotide binds in the N-terminal domain of that contains Rossmann fold

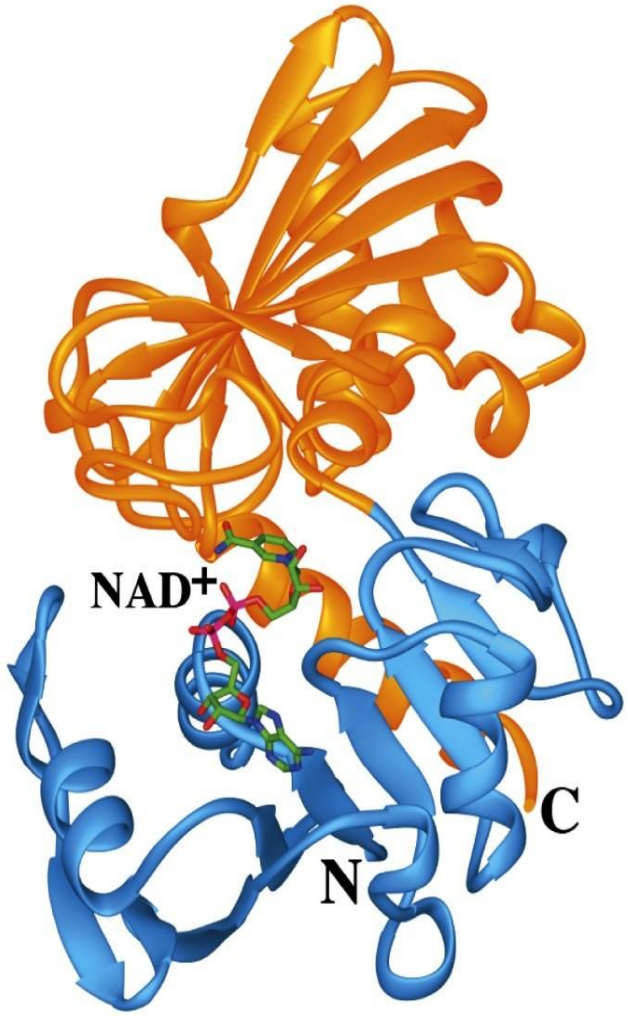


Figure 20: tertiary structure of Glyceraldehyde-3-phosphate dehydrogenase

Quaternary structure

Quaternary structure exists in proteins consisting of two or more identical or different polypeptide chains (**subunits**).

These proteins are called **oligomers** because they have two or more subunits.

The quaternary structure describes the manner in which subunits are arranged in the native protein.

Subunits are held together by **non-covalent forces**; as a result, oligomeric proteins can undergo rapid conformational changes that affect biological activity.

If the final protein is made of two subunits, the protein is said to be a dimer. If three subunits must come together, the protein is said to be a trimer, four subunits make up a tetramer, etc.

If the subunits are identical, the prefix “homo” is used, as in “homodimer.” If the subunits are different, we use “hetero,” as in “heterodimer.”

One of the most common example is Hemoglobin

Quaternary structure: Hemoglobin

It is made up of four polypeptides: two α- and two β-subunits.

One α-subunit and one β-subunit will come together to form a heterodimer, and two of these heterodimers will interact to form one hemoglobin molecule.

Hemoglobin can therefore be thought of as a dimer of dimers, which come together to give the final protein its quaternary structure.

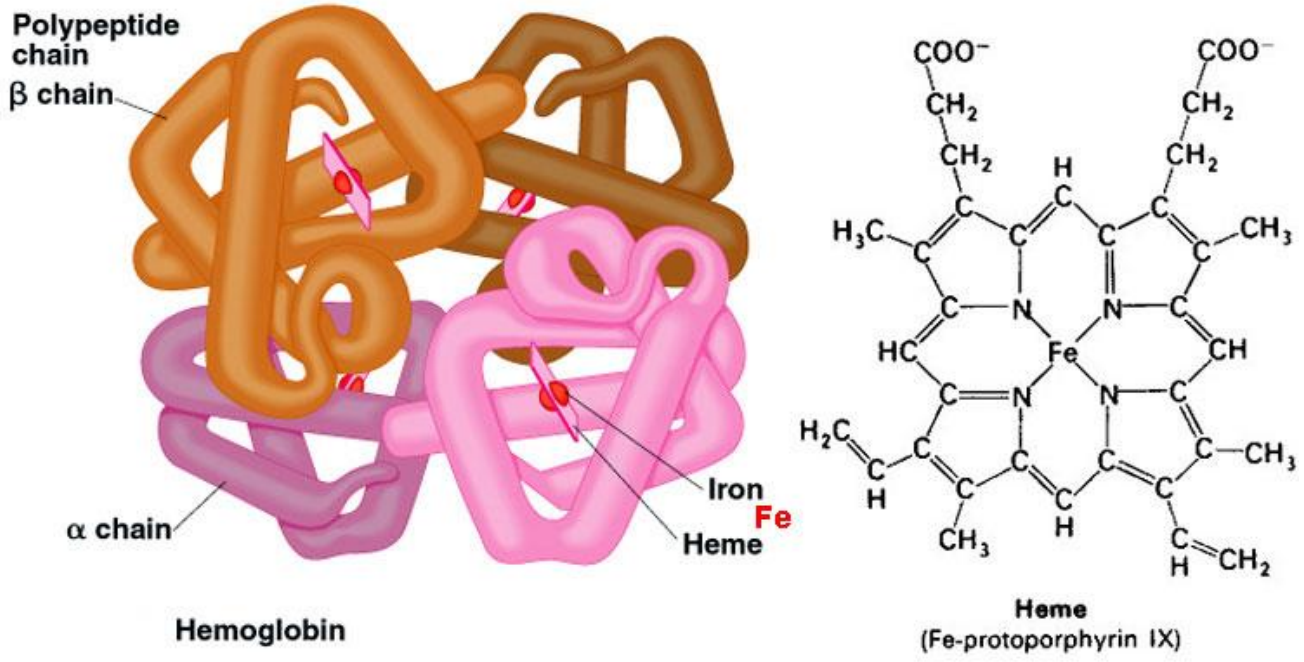
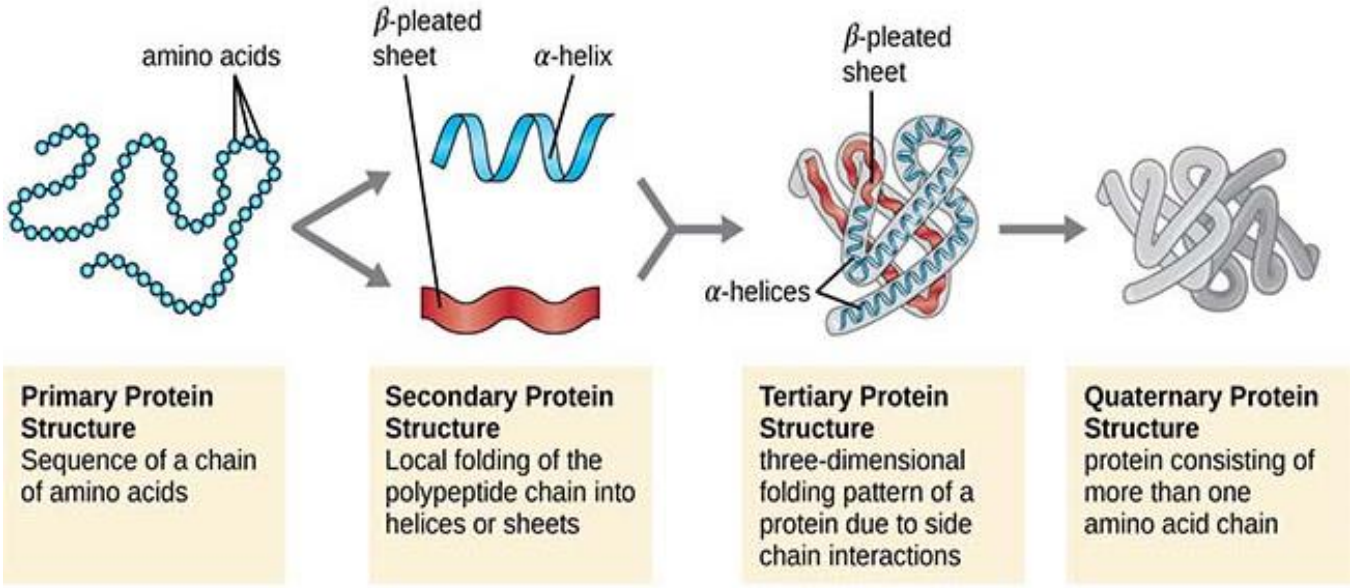


Figure 21: Hemoglobin structure with heme group

Summary



References

[1]Lehninger Principles of biochemistry

[2]Lubert Stryer Biochemistry

Voet and Voet Biochemistry

End