1-1Amino acids

Of the four major groups of biomolecules—lipids, carbohydrates, proteins, and nucleic acids—proteins have the widest array of functions. Keratin and collagen, for example, form long insoluble fibers, giving strength and support to tissues.

Hair, horns, hooves, and fingernails are all made up of keratin. Collagen is found in bone, connective tissue, tendons, and cartilage.

Membrane proteins transport small organic molecules and ions across cell membranes. Insulin, the hormone that regulates blood glucose levels, and hemoglobin, which transports oxygen from the lungs to tissues, are proteins.

Enzymes are proteins that catalyze and regulate all aspects of cellular function. In Lect. 7 and 8 we will discuss proteins and their primary components, the amino acids.

Proteins are biomolecules that contain many amide bonds, formed by joining amino acids together.

Proteins occur widely in the human body, accounting for approximately 50% of its dry weight (Figure 1).

Fibrous proteins, like keratin in hair, skin, and nails and collagen in connective tissue, give support and structure to tissues and cells.

Protein hormones and enzymes regulate the body's metabolism. Transport proteins carry substances through the blood, and storage proteins store elements and ions in organs.

Contractile proteins control muscle movements, and immunoglobulins are proteins that defend the body against foreign substances.
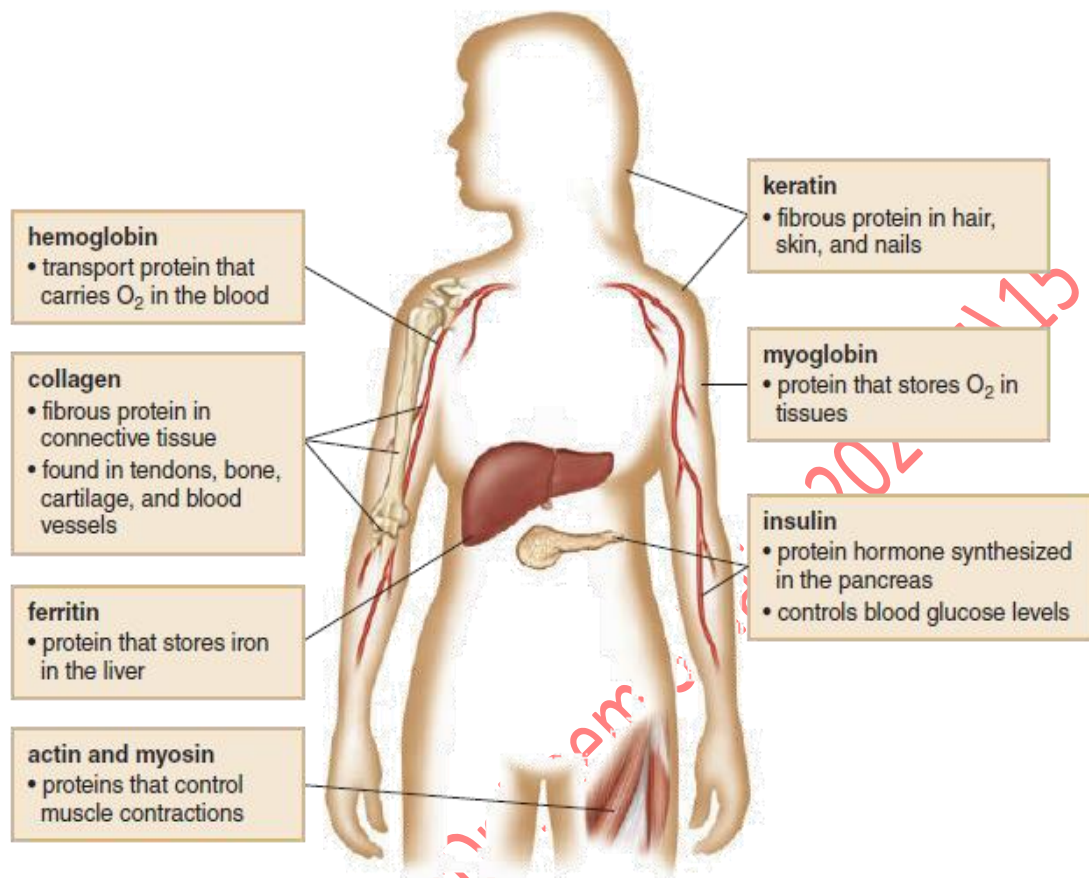
<u>Lect.5</u>

hemoglobin
• transport protein that carries O_2 in the blood

collagen
• fibrous protein in connective tissue
• found in tendons, bone, cartilage, and blood vessels

ferritin
• protein that stores iron in the liver

actin and myosin
• proteins that control muscle contractions

keratin
• fibrous protein in hair, skin, and nails

myoglobin
• protein that stores O_2 in tissues

insulin
• protein hormone synthesized in the pancreas
• controls blood glucose levels

Fig. 1: <u>Proteins in human body</u>

Unlike lipids and carbohydrates, which the body stores for use when needed, protein <u>is not stored</u> so it must be consumed on a daily basis.

Lect.5.

The current recommended daily intake for adults is 0.8 grams of protein per kilogram of body weight. Since children need protein for both growth and maintenance, the recommended daily intake is higher, as shown in Table below.

Table 1: Recommended daily protein intake

| Seq.	Group	Daily protein intake (g protein / Kg body weight)
1	**Children (1-3 years).**	1.10
2	**Children (4-13 years).**	0.95
3	**Children (14-18 years)**	0.85
4	**Adult**	0.80

1.2Amino Acids

To understand protein properties and structure, we must first learn about the amino acids that compose them.

1-3:General Features of Amino Acids.

Amino acids contain two functional groups—an amino group (NH2) and a carboxyl group (COOH).

In most naturally occurring amino acids, the amino group is bonded to the α carbon, the carbon adjacent to the carbonyl group, making them **α- amino acids.**

Lect.5.
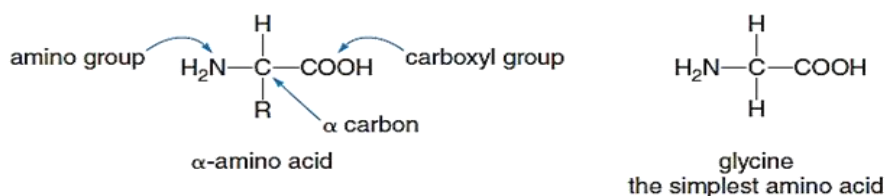
The 20 amino acid acids that occur naturally in proteins differ in the identity of the R group bonded to the α carbon.

The R group is called the side chain of the amino acid. The simplest amino acid, called glycine, has R = H. Other side chains may be simple alkyl groups, or have additional functional groups such as OH, SH, COOH, or NH_2.

Table 2 lists the structures of the 20 common amino acids that occur in proteins:

Amino acids with an additional COOH group in the side chain are called acidic amino acids.

Those with an additional basic N atom in the side chain are called basic amino acids.

All others are neutral amino acids.

All amino acids have common names, which are abbreviated by a three-letter or one-letter designation.

For example, glycine is often written as the three-letter abbreviation Gly, or the one-letter abbreviation G.

These abbreviations are also given in Table 2.

Lect.5.

Amino acids never exist in nature as neutral molecules with all uncharged atoms. Since amino acids contain a base (NH_2 group) and an acid (COOH), proton transfer from the acid to the base forms a salt called a zwitterion, which contains both a positive and a negative charge. These salts have high melting points and are water soluble.
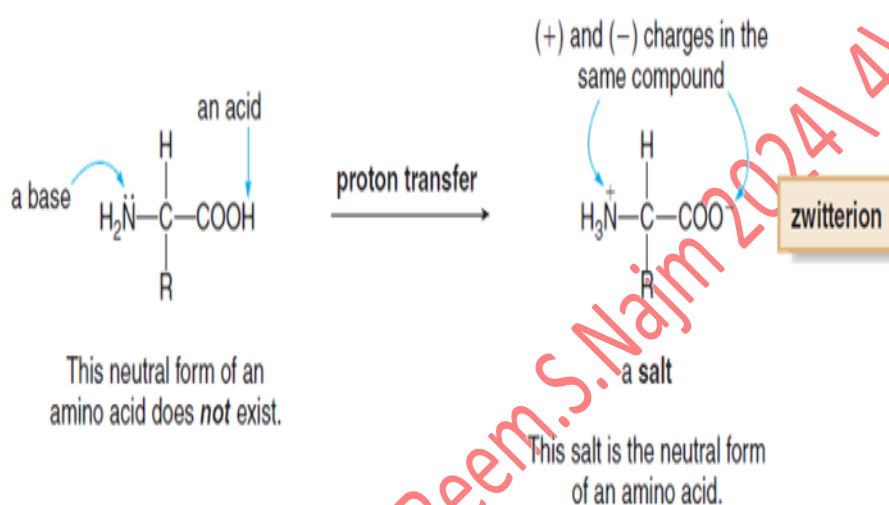
Humans can synthesize only 10 of the 20 amino acids needed for proteins.

The remaining 10, called essential amino acids, must be obtained from the diet and consumed on a regular, almost daily basis.

Diets that include animal products readily supply all of the needed amino acids. Since no one plant source has sufficient amounts of all of the essential amino acids, vegetarian diets must be carefully balanced. Grains—wheat, rice, and corn—are low in lysine, and legumes—beans, peas, and peanuts—are low in methionine, but a combination of these foods provides all the needed amino acids.
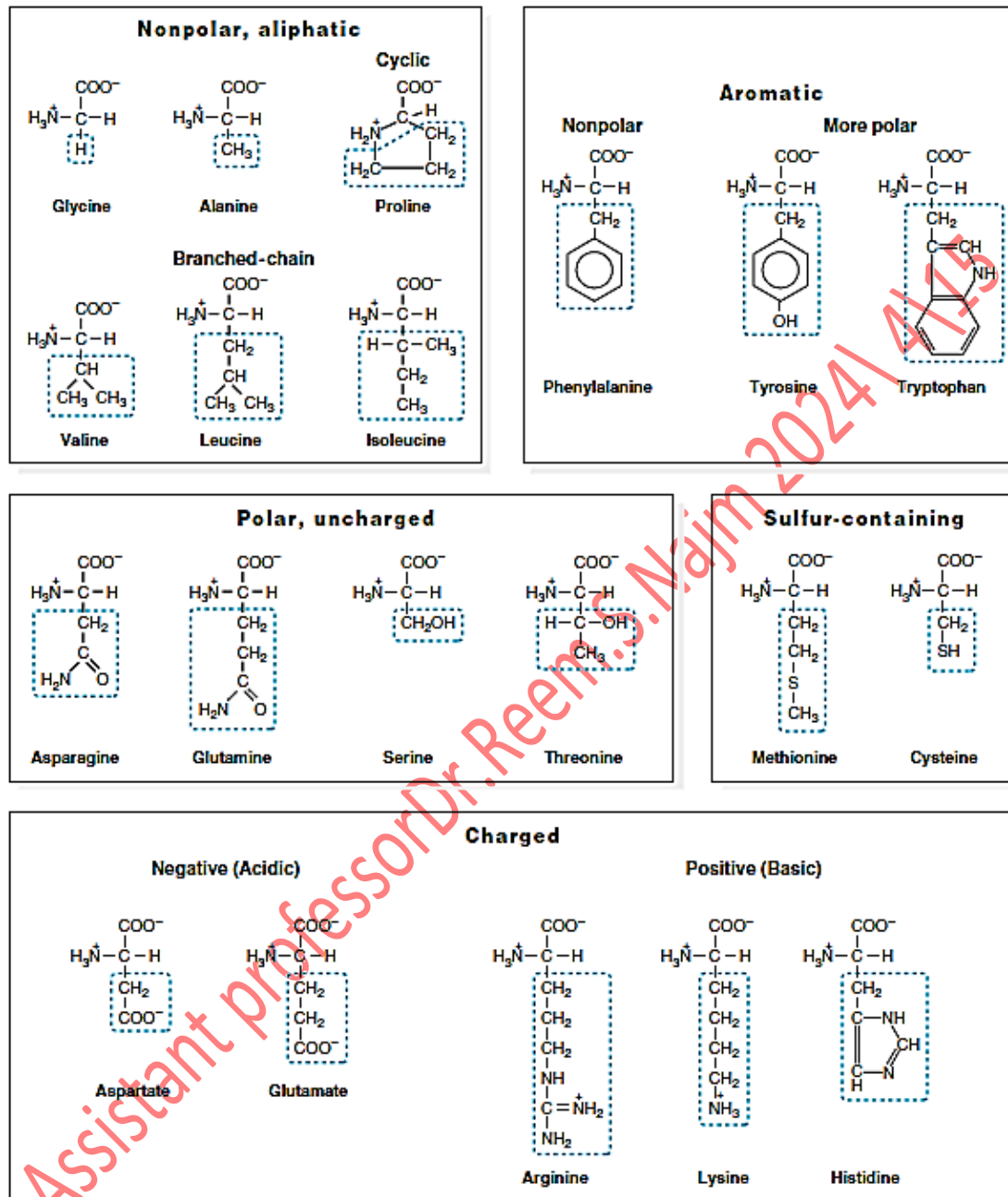
Lect.5.

1-4:Stereochemistry of Amino Acids

Except for the simplest amino acid, glycine, all other amino acids have a chirality center—a carbon bonded to four different groups—on the **α** carbon.

Thus, an amino acid like alanine (R = CH3) has two possible enantiomers, drawn below in both three-dimensional representations with wedges and dashed bonds, and Fischer projections.

Like monosaccharide's, the prefixes D and L are used to designate the arrangement of groups on the chirality center of amino acids. When drawn with a vertical carbon chain having the –COO– group at the top ‹and the R group at the bottom.

❖ L Amino acids have the $-NH_3^+$ group on the *left* side in the Fischer projection. Common naturally occurring amino acids are L isomers.

❖ D Amino acids have the $-NH_3^+$ group on the *right* side in the Fischer projection. D Amino acids occur infrequently in nature.
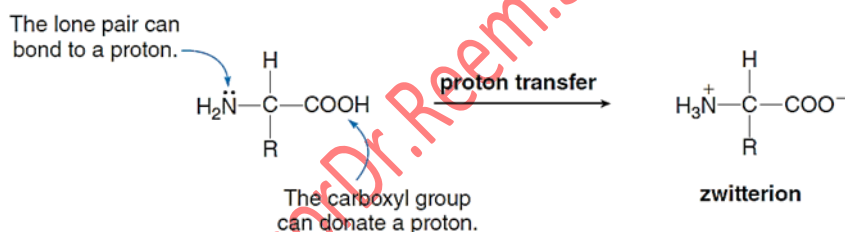
Lect.5.

Nonpolar, aliphatic

Glycine

Alanine

Cyclic

Proline

Branched-chain

Valine

Leucine

Isoleucine

Aromatic

Nonpolar

Phenylalanine

More polar

Tyrosine

Tryptophan

Polar, uncharged

Asparagine

Glutamine

Serine

Threonine

Sulfur-containing

Methionine

Cysteine

Charged

Negative (Acidic)

Aspartate

Glutamate

Positive (Basic)

Arginine

Lysine

Histidine

Lect.5.

Fig 2 : The 20 amino acids that are the building blocks of proteins can be classified as (a) nonpolar (hydrophobic), (b) polar but neutral, (c) acidic, or (d) basic. Also shown here are the one-letter and three-letter codes used to denote amino acids. For each amino acid, the ball-and-stick (left) and space-filling (right) models show only the side chain.

1-5 :Acid–Base Behavior of Amino Acids

As mentioned in Section 7.2.2, an amino acid contains both a basic amino group (NH2) and an acidic carboxyl group (COOH). As a result, proton transfer from the acid to the base forms a zwitterion, a salt that contains both a positive and a negative charge. The zwitterion is neutral; that is, the net charge on the salt is zero.
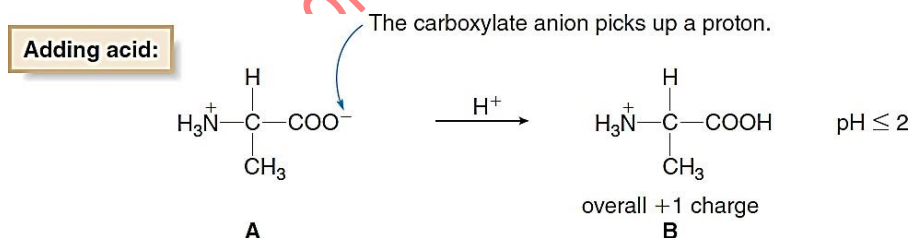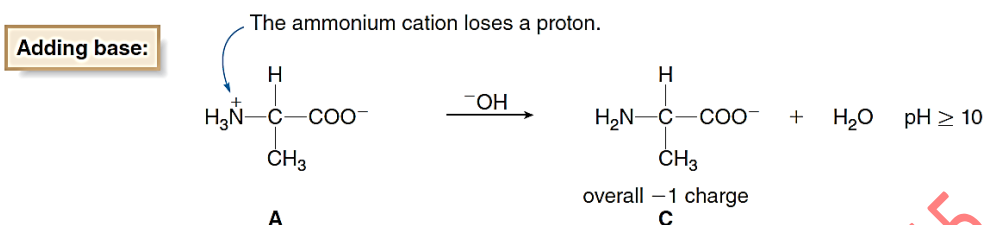
Lect.5.

In actuality, an amino acid can exist in different forms, depending on the pH of the aqueous solution in which it is dissolved. When the pH of a solution is around 6, alanine (R = CH3) and other neutral amino acids exist in their zwitterionic form (A), having no net charge. In this form, the carboxyl group bears a net negative charge—it is a carboxylate anion—(and the amino group bears a net positive charge (an ammonium cation).
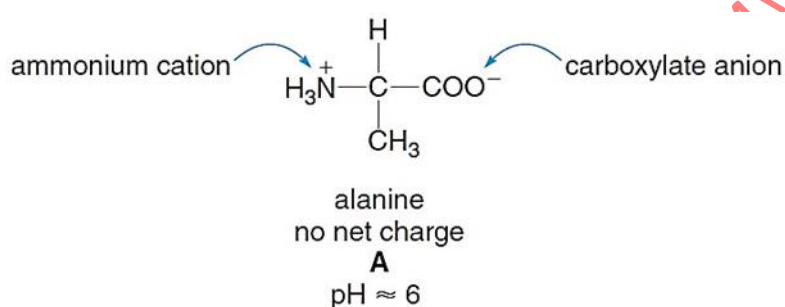
When strong acid is added to lower the pH to 2 or less, the carboxylate anion gains a proton and the amino acid has a net positive charge (form B).

When strong base is added to A to raise the pH to 10 or higher, the ammonium cation loses a proton and the amino acid has a net negative charge (form C).

Lect.5.

Adding base:

The ammonium cation loses a proton.

$H_3\overset{+}{N}$—C—COO$^-$ $\xrightarrow{\ ^-OH\ }$ H_2N—C—COO$^-$ + H_2O pH ≥ 10

overall −1 charge

A   C

Thus, alanine exists in one of three different forms depending on the pH of the solution in which it is dissolved. At the physiological pH of 7.4, neutral amino acids are primarily in their zwitterionic forms.
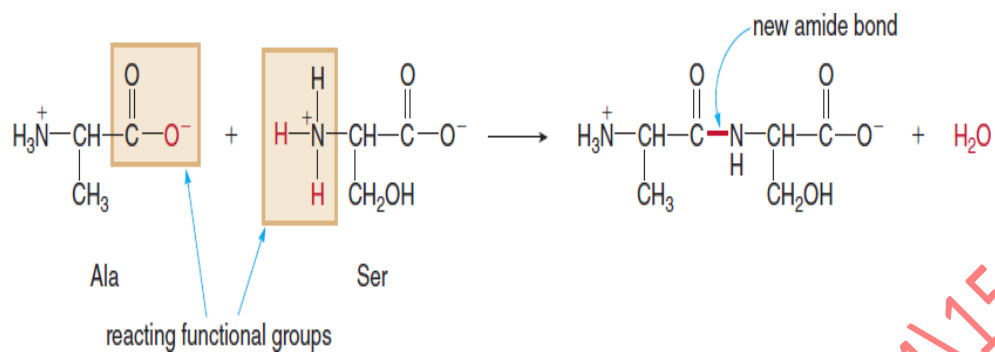
The pH at which the amino acid exists primarily in its neutral form is called its isoelectric point, abbreviated as pI.

The isoelectric points of neutral amino acids are generally around 6. Acidic amino acids (Table 2), which have an additional carboxyl group that can lose a proton, have lower pI values (around 3). The three basic amino acids, which have an additional basic nitrogen atom that can accept a proton, have higher pI values (7.6– 10.8).

The fact that amino acids are zwitterions explains their physical properties. All of them are solids with high melting points (for example, glycine melts at 262°C), just as we would expect for ionic compounds. The 20 common amino acids are also fairly soluble in water, as ionic compounds generally are. If they had no charges, we would expect only the smaller ones to be soluble.
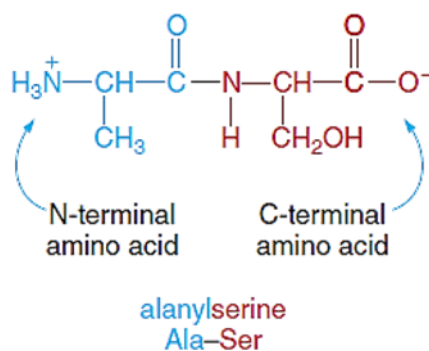
1-6Peptides

When amino acids are joined together by amide bonds, they form larger molecules called peptides and proteins.

Lect.5.

A dipeptide has two amino acids joined together by one amide bond.

A tripeptide has three amino acids joined together by two amide bonds.

[Amide bonds are shown in red.]

Polypeptides and proteins both have many amino acids joined together in long linear chains, but the term protein is usually reserved for polymers of more than 40 amino acids.

The amide bonds in peptides and proteins are called peptide bonds•
The individual amino acids are called amino acid residues

To form a dipeptide, the $-NH^+$ group of one amino acid forms an amide bond with the carboxylate ($-COO-$) of another amino acid, and the elements of H2O are removed. For example, reaction of the $-COO-$ group of alanine with the $-NH3^+$ group of serine forms a dipeptide with one new amide bond, as shown. The dipeptide has an ammonium cation ($-NH^+$) at one end of its chain and a carboxylate anion

($-COO-$) at the other.

Lect.5.

The amino acid with the free $-NH_3^+$ group on the **α** carbon is called the N- terminal amino acid.

the amino acid with the free $-COO-$ group on the **α** carbon is called the C-terminal amino acid.

By convention, the N-terminal amino acid is always written at the left end of the chain and the C-terminal amino acid at the right.

Lect.5.

Peptides are named as derivatives of the C-terminal amino acid. To name a peptide:

Name the C-terminal amino acid using the names in Table 2.

Name all other amino acids from left to right as substituents of the C- terminal amino acid.

Change the -ine or -ic acid ending of the amino acid name to the suffix –yl.

Thus, the dipeptide, which has serine as its C-terminal amino acid, is named as alanylserine. The peptide can be abbreviated by writing the one- or three-letter symbols for the amino acids in the chain from the N-terminal to the C-terminal end. Thus, Ala–Ser has alanine at the N-terminal end and serine at the C-terminal end.

Lect.5.

1-7Proteins

To understand proteins, the large polymers of amino acids that are responsible for so much of the structure and function of all living cells, we must learn about four levels of structure, called the primary, secondary, tertiary, and quaternary structure of proteins.
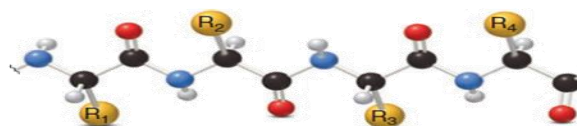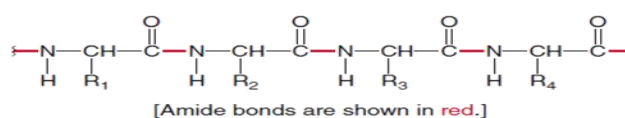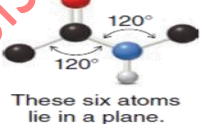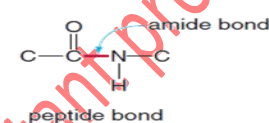
1-Primary Structure

The primary structure of a protein is the particular sequence of amino acids that is joined together by peptide bonds. The most important element of this primary structure is the amide bond that joins the amino acids.

The carbonyl carbon of the amide has trigonal planar geometry.

All six atoms involved in the peptide bond lie in the same plane.

All bond angles are 120° and the C=O and N-H bonds are oriented 180° from each other.

As a result, the backbone of the protein adopts a zigzag arrangement as shown in the three-dimensional structure of a portion of a protein molecule.

Lect.5.

The primary structure of a protein—the exact sequence of amino acids—determines all properties and function of a protein. As we will see in Section 16.7, substitution of a single amino acid by a different amino acid can result in very different properties.

2- Secondary Structure

 The three-dimensional arrangement of localized regions of a protein is called its secondary structure. These regions arise due to hydrogen bonding between the N H proton of one amide and the C=O oxygen of another. Two arrangements that are particularly stable are called the **α**-helix and the **β**-pleated sheet.

The **α**-helix forms when a peptide chain twists into a right-handed or clockwise spiral, as shown in Figure 3a. The C=O group of one amino acid is hydrogen bonded to an N-H group four amino acid residues farther along the chain. The R groups of the amino acids extend outward from the core of the helix. Both the myosin in muscle and the α-keratin in hair are proteins composed almost entirely of α-helices. The **β**- pleated sheet forms when two or more peptide chains, called strands, line up side-by side, as shown in Figure 3b. Hydrogen bonding often occurs between the N-H and C=O groups of nearby amino acid residues. The R groups are oriented above and below the plane of the sheet, and alternate from one side to the other along a given strand.
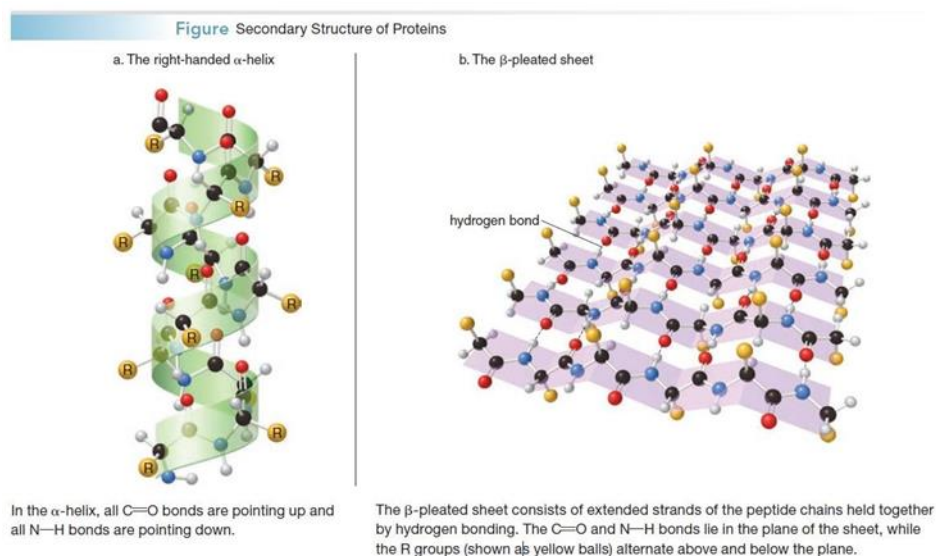
Lect.5.

Figure Secondary Structure of Proteins

a. The right-handed α-helix

b. The β-pleated sheet

hydrogen bond

In the α-helix, all C═O bonds are pointing up and all N─H bonds are pointing down.

The β-pleated sheet consists of extended strands of the peptide chains held together by hydrogen bonding. The C═O and N─H bonds lie in the plane of the sheet, while the R groups (shown as yellow balls) alternate above and below the plane.

**Fig. 3: α-helix and
β- sheet structure**

The β-pleated sheet arrangement is favored by amino acids with small R groups, like alanine and glycine.

With larger R groups, steric interactions prevent the chains from getting close together, so the sheet cannot be stabilized by hydrogen bonding.

Most proteins have regions of α-helix and β-pleated sheet, in addition to other regions that cannot be characterized by either of these. arrangements. Shorthand symbols are often used to indicate these regions of secondary structure. In particular,

 a flat helical ribbon is used for the α-helix, while a flat wide arrow is used for the β- pleated sheet.

These representations are often used in **ribbon diagrams** to illustrate protein structure.

Lect.5.

α-helix β-sheet random region Proteins are drawn in a variety of ways to show different aspects of their structure.

Figure 5 illustrates three different representations of the protein lysozyme, an enzyme found in both plants and animals.

Lysozyme catalyzes the hydrolysis of bonds in bacterial cell walls, weakening them, often causing the bacteria to burst.
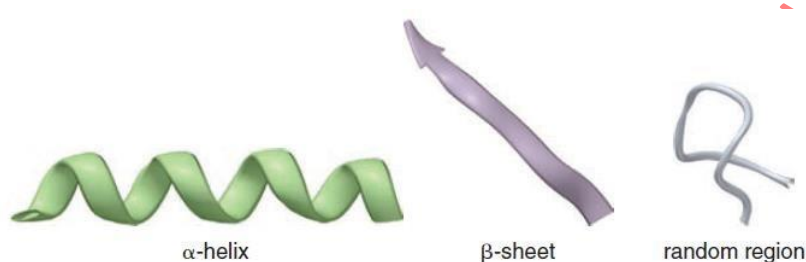

α-helix   β-sheet   random region

Fig. 5: α-helix β-sheet random region Proteins.

3-Tertiary and Quaternary Structure

The three-dimensional shape adopted by the entire peptide chain is called its tertiary structure.

A peptide generally folds into a shape that maximizes its stability. In the aqueous environment of the cell, proteins often fold in such a way as to place a large number of polar and charged groups on their outer surface, to maximize the dipole–dipole and hydrogen bonding interactions with water. This generally places most of the nonpolar side chains in the interior of the protein, where London dispersion forces between these hydrophobic groups help stabilize the molecule, too. In addition, polar functional groups hydrogen bond with each other (not just water), and amino acids with charged side chains like –COO– and –NH + can stabilize tertiary structure by electrostatic interactions. Finally,

Lect.5.

disulfi de bonds are the only covalent bonds that stabilize tertiary structure. These strong bonds form by the oxidation of two cysteine residues on either the same polypeptide chain or another polypeptide chain of the same protein.

Insulin, for example, consists of two separate polypeptide chains (labeled the A and B chains) that are covalently linked by two intermolecular disulfi de bonds.

The A chain, which also has an intramolecular disulfi de bond, has 21 amino acid residues, whereas the B chain has 30.

Figure 6 schematically illustrates the many different kinds of intramolecular forces that stabilize the secondary and tertiary structures of polypeptide chains.
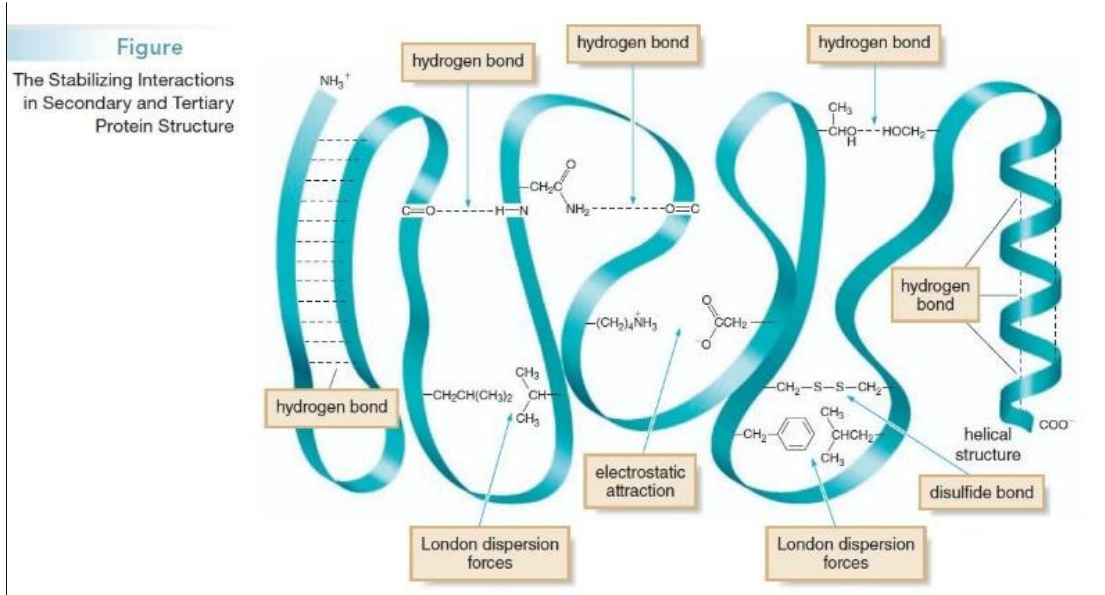
Nearby amino acid residues that have only nonpolar carbon– carbon and carbon–hydrogen bonds are stabilized by London dispersion forces. Amino acids that contain hydroxyl (OH) and amino groups (NH2) in their side chains can intermolecularly hydrogen bond to each other.

The shape adopted when two or more folded polypeptide chains come together into one protein complex is called the quaternary structure of the protein.

Each individual polypeptide chain is called a subunit of the overall protein. Hemoglobin, for example, consists of two α and two β subunits held together by intermolecular forces in a compact three-dimensional shape.

The unique function of hemoglobin is possible only when all four subunits are together.

Lect.5.

Figure

The Stabilizing Interactions in Secondary and Tertiary Protein Structure

hydrogen bond

hydrogen bond

hydrogen bond

hydrogen bond

hydrogen bond

hydrogen bond

London dispersion forces

London dispersion forces

electrostatic attraction

disulfide bond

helical structure