Master Success Series

TEXT BOOK OF

GENERAL BIOCHEMISTRY

For B.S. & M.Sc. Zoology

Dr. Riaz-ul-Haq Ramay

According to syllabus approved
by the Higher Education Commission (HEC) Islamabad
Applicable to all the Universities of Pakistan

Master Success **Text Book**

GENERAL BIOCHEMISTRY

For B.S. (4 & 5th semester)
M.Sc. and M.Phil.

Dr. M. Riaz-ul-Haq. Ramay, PhD
Assistant Professor
Govt. Postgraduate College of Science, Multan

Consultants:

Honorable Prof. Dr. Muhammad Ali
Vice Chancellor,
Quad Azam University, (QAU) Islamabad

Prof. Dr. Khawaja Raees,
University of Sargodha (UOS), Sargodha

Prof. Mrs. Shehnaz Akhtar,
Govt. Degree College Rawalpindi

Prof. Mrs. Noshaba Rana,
Govt. PG College for women, Mirpur (AJK)

Prof. Liaqat Ali Khan
Govt. Postgraduate College, Peshawar (KPK)

Prof. Mrs. Farahat Zahra,
Queen Mary College, Lahore

Prof. Syed Shabir Hussain Shirazi
Govt. Mary College, Sialkot

Prof. Afsari Bano,
Ibne-Rushd Girls College, Mirpur Khas (Sindh)

Prof. Zain-ul-abideen
Govt. College of Science, Quetta

M.I. PUBLISHERS
Ph: 0300-6366320

Future Zoologist Group of PGCG (By Muhammad ARSALAN)

Our New Vision!

The introduction of B.S. at different universities and colleges marked a new era in the educational history of Pakistan. The HEC has thoroughly revised the syllabi of B.S. programs. New subjects have been introduced for inculcating new vision of thoughts in the minds on our students. B.S. Zoology is one of these disciplines in which B.S. have been started. Out General Books of Zoology with the name of *Master Success Series Zoology* are already catering to the educational needs of students of first four semesters of B.S Zoology. We are approached many a time by academia to introduce books for last four semesters. We also felt the need of writing these books as early as possible. Hence our thoughts and tools again took the stand to complete this herculean task. In this connection, our first book *Master success General Biochemistry* is your hand. It has following important features.

1. This book written in line with syllabus provided by the HEC.
2. Book is prepared from the books recommended by the HEC. Principles of Biochemistry by Albert Leningher have been used as Principle reference book.
3. Written material is elaborated with diagrammatic approach.
4. At the end of each chapter long questions, short questions and multiple choice questions (MCQS) has been given.

I am sure this book will prove to be an invaluable asset for the student and teachers. If you found any problem please contact me through phone or E-mail. I my response will be quick and productive.

Prof. Riaz ul Haq. Ramay
E-mail: ramay11@hotmail.com
Phone: 03006366320

Contents

No.	Chapter	Page
1	Amino Acids, Peptides and Proteins	9
2	Proteins II	43
3	Enzymes	58
4	Carbohydrates	86
5	Lipids	124
6	Vitamins and Cofactors	149
7	Nucleic Acids	155
8	Bioenergetics	193
9	Glycolysis	204
10	Citric Acid Cycle	250
11	ETC and Its Components	270
12	Lipid Metabolism	291
13	Nitrogen Metabolism	333

Future Zoologist
Group of PGCG
(By Muhammad
ARSALAN)

Chapter 1

Amino Acids, Peptides and Proteins

Syllabus 2012-13
Chapter 1: Amino acids, peptides and proteins:
Standard amino acids, their structure and classification; acid/base properties of amino acids and their titration curves; peptides, their ionic behavior and amino acid composition, cytochrome c; Proteins: level of structural organization, example of structural and functional proteins.
Syllabus 2018-19
2. Amino acids, peptides & proteins: a. Types of amino acids & their classification; b. Uncommon amino acids; Acid/base behavior of amino acids. c. Titration curves in amino acids and their importance: d. Peptides & proteins; e. Biologically active peptides & polypeptides; f. Amino acid sequence in proteins & their importance; **Conjugated proteins;**

INTRODUCTION

Proteins are the most abundant biological macro- molecules. These occur in all cells and all parts of cells.

Diversity in structure: Proteins also occur in great variety. These occur in thousands of different kinds. Proteins range in size from relatively small peptides to huge polymers. The molecular weights of these polymers are in the millions. But these may be found in a single cell.

Diversity in function: Moreover, proteins show enormous diversity of biological function. These are the most important final products of the information path-ways. Proteins are the molecular instruments for expressing genetic information.

Role of amino acids in formation of proteins: Amino acids are relatively simple monomeric subunits. But they are key for the structure of the thousands of different proteins. Proteins may be from the most ancient lines of bacteria. Or these may be from the most **complex** forms of life. All such proteins are constructed from the same ubiquitous set of 20 amino acids. These amino acids covalently linked in characteristic linear sequences.

Each amino acid has a side chain. This side chain has distinctive chemical properties. Therefore, these groups of 20 pre-cursor molecules may be regarded as alphabet. Language of protein structure is written in these alphabets.

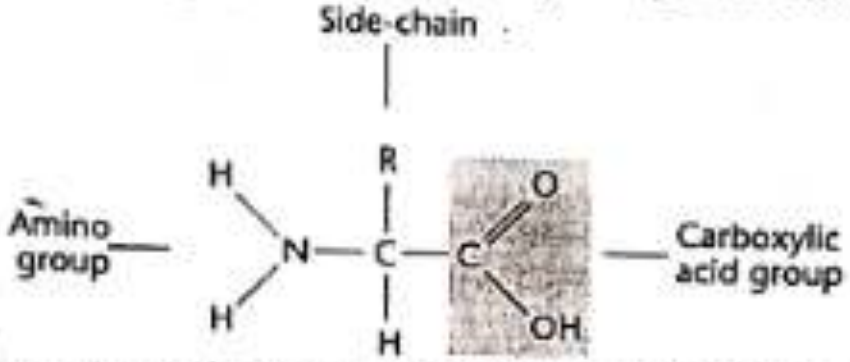
Role of proteins in cells

There is most remarkable feature of cells. It can produce proteins with different properties and activities. These are produced by joining the same 20 amino acids. These amino acids are jointed in many different combinations and sequences. Different organisms can make such widely diverse products from these building blocks. These product include enzymes, hormones, antibodies, transporters, muscle fibers, lens protein of the eye, feathers, spider webs, rhinoceros horn, milk proteins, antibiotics, mushroom poisons, and myriad other substances. These substances have distinct biological activities. Enzymes are one of these protein products. Enzymes are the most varied and specialized. Almost all cellular reactions are catalyzed by enzymes.

AMINO ACIDS

Standard Amino Acids

All amino acids have an amino group (NH_2) and a carboxyl group (COOH), attached to the same carbon atom, also known as alpha carbon. 170 types of amino acids are present in cells and tissues. Only 25 amino acids form proteins. However, most of the protein contains 20 types of amino acids. They have general formula as:



R may be a hydrogen atom as in **glycine** amino acid. It may be CH_3 as in **alanine** amino acid. Or it may be any other group. Thus amino acids mainly differ due to the type or nature of R group.

Discovery of Amino Acids

Asparagine was the first amino acid discovered was, in 1806. The last of the 20 amino acids was threonine. It was not identified in 1938.

Naming of Amino Acids

All the amino acids have trivial or common names. In some cases, their names derived from the source. They were first isolated from this source.

- Asparagine was first found in asparagus
- Glutamate was found in wheat gluten
- Tyrosine was first isolated from cheese. Its name is derived from the Greek word tyros, means cheese.
- Glycine has sweet taste. Greek word glykos means sweet.

Structure of Amino Acids

Amino Acids Share Common Structural Features

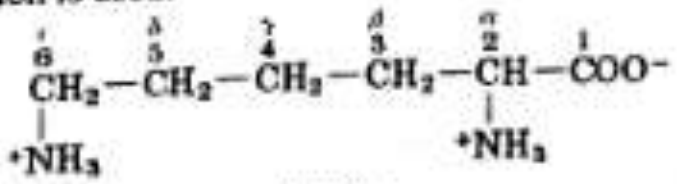
All 20 of the common amino acids are α-amino acids. They have a carboxyl group and an amino group bonded to the same carbon atom (α carbon). They differ from each other in their side chains, or R groups. R groups vary in structure, size, and electric charge. It influences the solubility of the amino acids in water.

There less are many less common amino acids in addition to these 20 amino acids. Some are residues. These are modified after synthesis of a protein. Others are amino acids are present in living organisms. But they do not form proteins. The common amino acids of proteins are assigned three-letter abbreviations and one-letter symbols. These symbols are used as shorthand. These indicate the composition and sequence of amino acids in proteins.

Conventions to Indentify Amino Acids

Two conventions are used to identify carbons in an amino acid.

1. The additional carbons in an R group are commonly designated as β, γ, δ, ε and so forth. This numbering starts from α carbon.
2. For most other organic molecules, carbon atoms are simply numbered from one end. It gives highest priority (C-1) to the carbon with the substituent. This carbon contains the atom of highest atomic number. Within this latter convention, the carboxyl carbon of an amino acid will be C-1 and α carbon will be C-2. In some cases, amino have heterocyclic R groups. In this case, Greek lettering system is ambiguous. Therefore, numbering convention is used.



Lysine

Optical Activity in Amino Acids

For all amino acids, except glycine, the carbon is bonded to four different groups in common amino acids. These include a carboxyl group, an amino group, an R group, and

a hydrogen atom. In glycine, R group is another hydrogen atom. The α-carbon atom is thus a **chiral** center. There is tetrahedral arrangement of the bonding orbitals around α - carbon atom. Here four different groups can occupy two unique spatial arrangements. Therefore, amino acids have two possible **stereoisomers**. They are non-superimposable mirror images of each other. Hence two forms represent a class of stereoisomers called **enantiomers**. All molecules with a chiral center are also optically active. Therefore, they rotate plane-polarized light.

(a) L-Alanine D-Alanine

(b) L-Alanine D-Alanine

(c) L-Alanine D-Alanine

D, L System

Special nomenclature is used to specify absolute configuration of four substituents of asymmetric carbon atoms. The absolute configurations of simple sugars and amino acids are specified by D, L system. It is based on the absolute configuration of three-carbon sugar **glyceraldehyde**. This convention was proposed by **Emil Fischer** in 1891. It was later confirmed by x-ray diffraction analysis. For all chiral compounds:

a) Stereoisomers with configuration related L-glyceraldehyde are designated L.

b) Stereoisomers related to D-glyceraldehyde are designated D.

The functional groups of L-alanine are matched with L-glyceraldehyde. It is matched by aligning the functional groups which can be interconverted by simple one-step chemical reactions. An aldehyde is readily converted to a carboxyl group via a one-step oxidation.

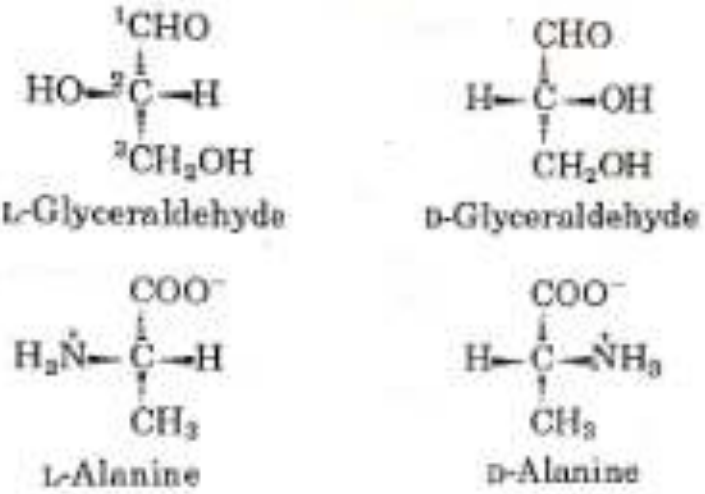
Therefore, carboxyl group of L alanine occupies the same position about chiral carbon like aldehyde group of L-glyceraldehyde.

The *l* and *d* system

Historically, similar *l* and d designations were used:

a) The *l* is used for levorotatory (rotating light to the left)

b) The *d* is used for dextrorotatory (rotating light to the right).



By Fischer's convention, L and D refer only to absolute configuration of four substituents around chiral. They do not refer the carbon not showing optical properties of the molecule.

RS system

RS system is another system of specifying configuration around a chiral center. It is used in the systematic nomenclature of organic chemistry. It describes more precisely configuration of molecules with more than one chiral center.

Amino Acid Residues in Proteins L Stereoisomers

Nearly all biological compounds with a chiral center occur naturally in only one stereoisomeric form. They may be D or L.

- The amino acid residues in protein molecules are exclusively L stereoisomers.
- D-Amino acid residues are found only in a few, small peptides. These include some peptides of bacterial cell walls and certain peptide antibiotics.

All amino acid residues in proteins are L stereoisomers. Some chiral compounds are formed by ordinary chemical reactions. They form racemic mixture of D and L isomers. Such mixtures are difficult to seperate. But in a living system, D and L isomers are different like right hand and the left. The constituent amino acids must be of one stereochemical series for formation of stable proteins. Enzymes in cells are asymmetric. Hence they specifically synthesize L isomers of amino acids.

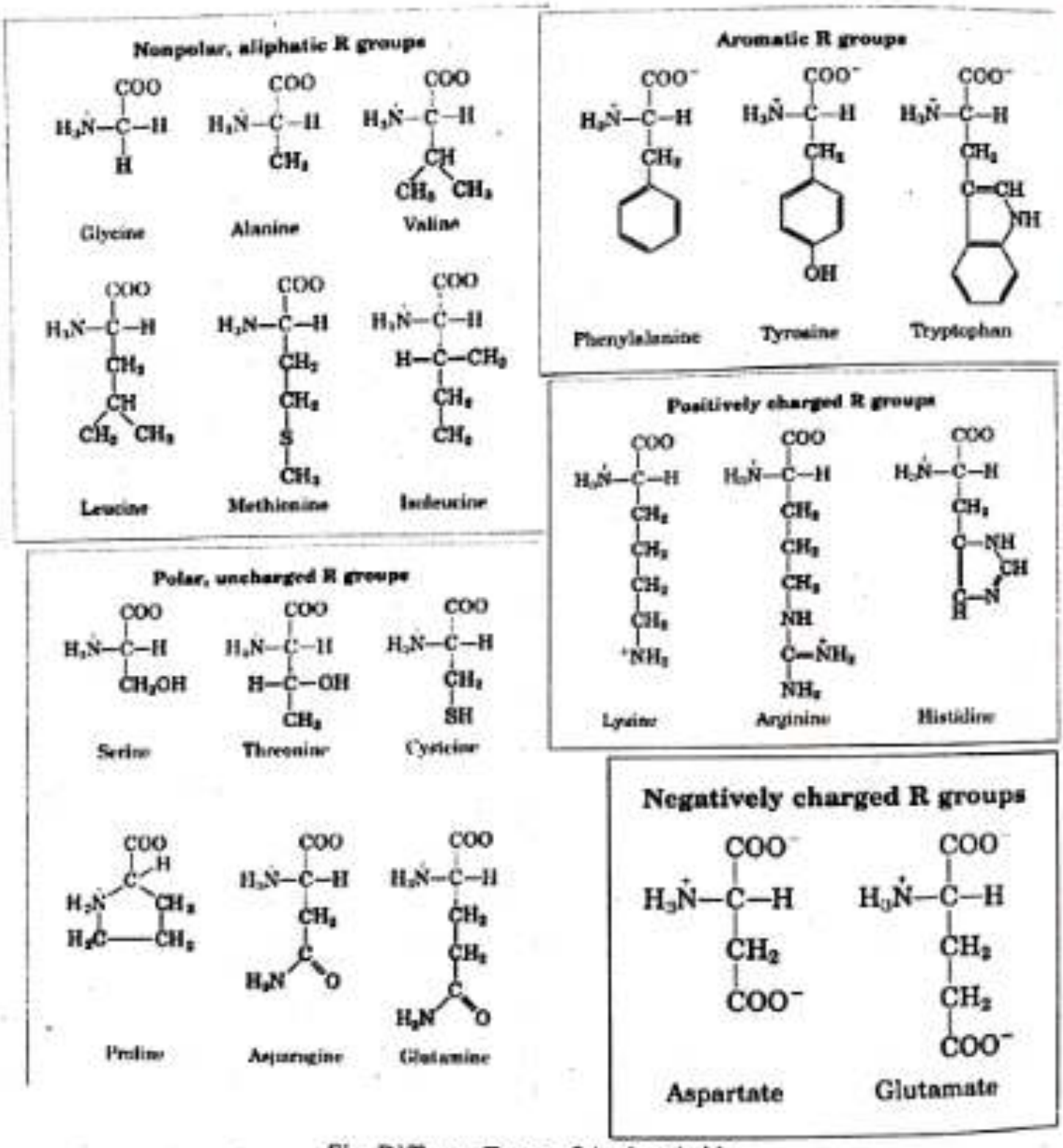


Fig: Different Types of Amino Acids

Classification of Amino Acids

Basis of Classification

Amino acids can be classified by R group. Amino acids are classified into five main classes. This classification is based on the properties of their R groups (See Table). Particularly these properties are:

a) Polarity or R

b) Or their tendency to interact with water at biological pH (near pH 7.0).

The polarity of the R groups varies widely. These vary from non-polar and hydrophobic (water-insoluble) to highly polar and hydrophilic (water-soluble). There are gradations of polarity, size, and shape of R groups within each class.

(A) Non-polar, Aliphatic R Groups

The R groups in this class of amino acids are non-polar and hydrophobic. The side chains of alanine, valine, leucine, and isoleucine tend to cluster together within proteins. It stabilizes protein structure by hydrophobic interactions.

a) **Glycine:** It has the simplest structure. It is formally non-polar. But it has very small side chain. It has no real role in hydrophobic interactions.

b) **Methionine:** It is sulfur-containing amino acids. It has a non- polar thioether group in its side chain.

c) **Proline:** It has an aliphatic side chain. This chain has a distinctive cyclic structure. Proline has secondary amino (imino) group of proline residues. This group is held in a rigid conformation. It reduces structural flexibility of polypeptide regions containing proline.

(B) Aromatic R Groups

This group includes **phenylalanine, tyrosine**, and **tryptophan**. These amino acids have their aromatic side chains. These amino acids are relatively non-polar (hydrophobic). All can participate in hydrophobic interactions.

- The hydroxyl group of tyrosine can form hydrogen bonds. This it is an important functional group in some enzymes.
- Tyrosine has hydroxyl group and tryptophan has nitrogen in indole ring. Therefore, tyrosine and tryptophan are significantly more polar than phenylalanine.
- Tryptophan and tyrosine and phenylalanine, absorb ultraviolet light. Hence most proteins absorb characteristic strong light at a wavelength of 280 nm. Researchers use this property to characterize proteins.

(C) Polar Uncharged Group

This group includes amino acids, **serine**, **threonine**, **cysteine**, **asparagine** and **glutamine**. R groups of these amino acids are more soluble in water. Their R groups are more hydrophobic than non-polar amino acids. This is because of functional groups of these amino acids. Their functional groups form hydrogen bonds with water. Following groups contribute in polarity of these amino acids:

- Hydroxyl group in serine and threonine.
- Sulfhydryl group in cysteine.
- Amide groups in asparagine and glutamine.

Asparagine and glutamine are amides of two other amino acids asparatate and glutamate.

Disulphide bond or bridge

Cysteine is readily oxidized to form **cystine**. It is covalently linked diameric amino acid. Two amino acids are joined by disulphide in cysteine. Disulfide linked residues are

strongly hydrophobic (non-polar). Disulphide bond play a special role in structures of many proteins.

(D) Positively Charged R group

The most hydrophilic R groups may be positively or negatively charged are pH 7. The examples of **positively charged R groups** are lysine, arginine and histidine.

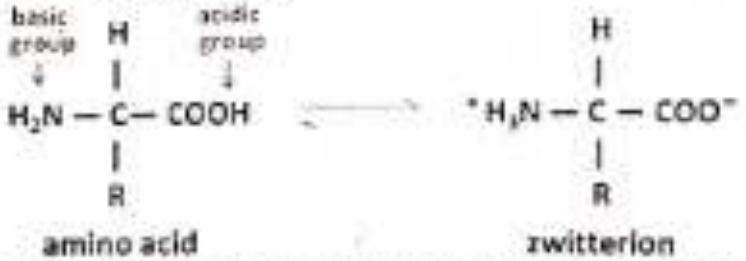
- **Lysine** has a second primary amino group on this aliphatic chain.
- **Arginine** has positively charged guanidino group.
- **Histidine** has imidazole group.

(E) Negatively charged R groups:

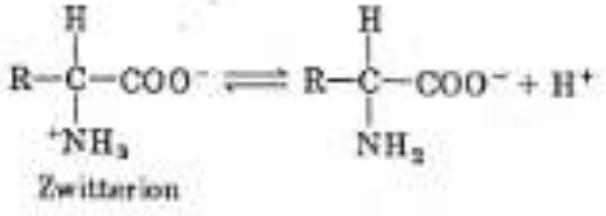
This group includes two amino acids, **asparate** and **glutamate**. They have net negative charge in their R groups at pH 7. Both these amino acids have secondary carboxyl groups.

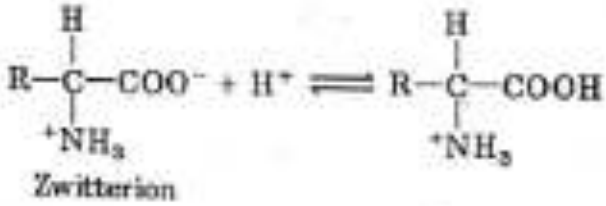
Acid Base Properties of Amino Acids

Amino acid is dissolved in water. Amino acid is dipolar ion in this solution. It is called **Zwitterions** (German for "hybrid ion").

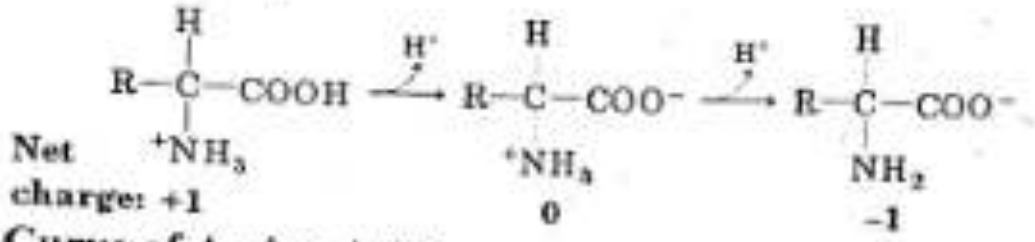


A zwitterion can act as either an acid (proton donor): or a base (proton acceptor):





Substances having this dual nature are amphoteric. They are often called ampholytes (from amphoteric electrolytes). A simple monoamino monocarboxylic amino acid is like alanine. It is a diprotic acid when fully protonated. It has two groups, -COOH group and $-NH_3^+$ groups. They can yield protons:



Titration Curve of Amino Acids

Amino acids have characteristic titration curves. Acid-base titration involves the gradual addition or removal of protons. Figure shows the titration curve of the diprotic form of glycine. The plot has two distinct stages. These stages correspond to deprotonation of two different groups on glycine. Each of the two stages resembles in shape the titration curve of a monoprotic acid like acetic acid. Hence it can be analyzed in the same way.

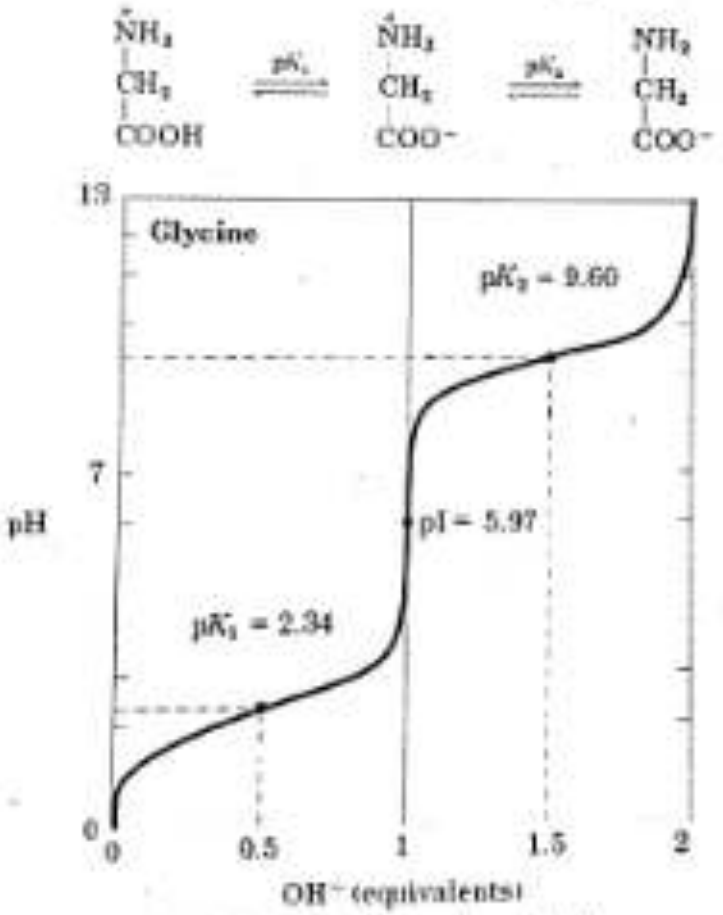


Fig: Titration of Amino Acid

1. At very low pH, predominant ionic species of glycine is fully protonated form, $^{+}H_3N$-CH_2O- COOH.
2. At the midpoint in the first stage of the titration, -COOH group of glycine loses its proton. Now equimolar concentrations of proton-donor ($^{+}H_3N$-CH_2-COOH) and proton-acceptor (H_3N-CH_2COO^{-}) species are present.

   At the midpoint of any titration, a point of inflection is reached. Al this point, pH is equal to pKa of the protonated group being titrated. For glycine, pH at the midpoint is

2.34. Thus its -COOH group has a pKa (labeled pK_1 in Fig) of 2.34. We know that pH is notations for proton concentration. The pKa is notation for the equilibrium constant for ionization. The pKa is a measure of the tendency of a group to give up a proton. This tendency decreases tenfold as the pKa increases by one unit.

3. As the titration proceeds, another important point is reached at pH 5.97. Here there is another point of inflection. Here removal of the first proton is essentially complete. But removal of the second has just begun. At this pH glycine is present largely as the dipolar ion (H_3N-CH_2COO^-).
4. Second stage of titration corresponds to removal of a proton from $-NH_3^+$ group of glycine. The pH at the midpoint of this stage is 9.60. It is equal to Pka (labeled pK2 in Fig) for $-NH_3^+$ group.
5. The titration is complete at a pH of about 12. At this point, predominant form of glycine is H_3N-CH_2COO^-.

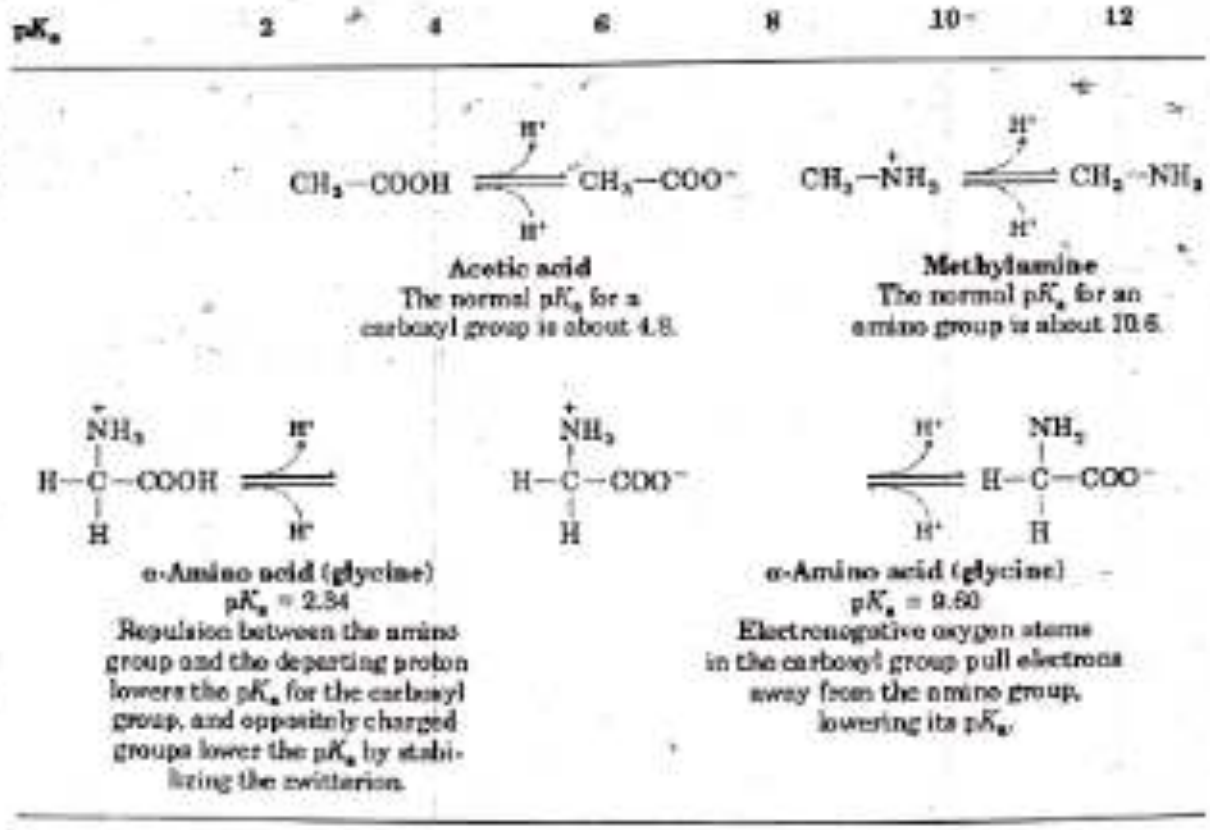


Fig: Effect of chemical environment in Kpa

Important Information Given by Titration Curve (Applications of TC)

(a) Quantitative Measurement

Titration curve gives a quantitative measure of pKa of each of two ionizing groups. This is 2.34 for -COOH group and 9.60 for $-NH_3^+$ group.

1. We note that carboxyl group of glycine is over 100 times more acidic (more easily ionized) than carboxyl group of acetic acid. It has a pKa of 4.76. This value is about average for a carboxyl group attached to an un-substituted aliphatic hydrocarbon.

2. The change in pKa of glycine is caused by repulsion. This repulsion occurs between departing proton and nearby positively charged amino group on -carbon atom. The opposite charges on resulting zwitterion are stabilizing. It pushes the equilibrium farther to the right. Similarly, pKa of amino group in glycine is moved downward relative to average pKa of an amino group.
3. This effect is partly due to electronegative oxygen atoms in the carboxyl groups. This oxygen tends to pull electrons toward them. It increases the tendency of the amino group to give up a proton. Hence amino group has a pKa is lower than of an aliphatic amine like methylamine (Fig).

(b) Buffering power of glycine

Titration curve of glycine provides second information. This curve shows that this amino acid has two regions of buffering power.

1. One of these is relatively flat portion of the curve. It extends for approximately 1 pH unit on either side of first pKa of 2.34. It shows that glycine is a good buffer near this pH.
2. The other buffering zone is centered around pH 9.60. We use Henderson-Hasselbalch equation within buffering ranges of glycine. It is used to calculate proportions of proton-donor and proton-acceptor species of glycine. It is required to make a buffer at a given pH. Glycine is not a good buffer at pH of intracellular fluid or blood, about 7.4.

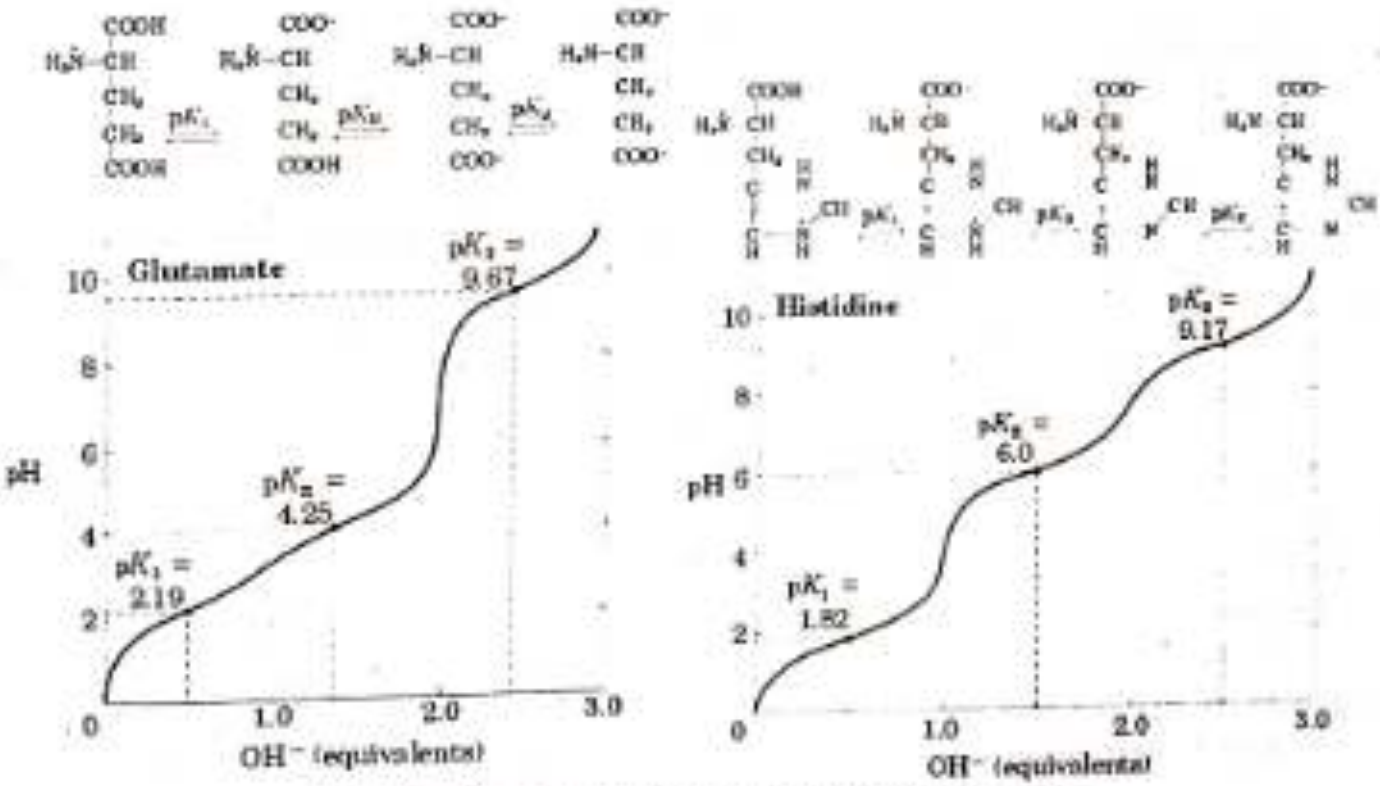


Fig: Titration curve of (a) glutamate (b) Histidine

(c) Titration Curves Predict the Electric Charge of Amino Acids

There is another important piece of information derived from the titration curve of an amino acid. It is the relationship between its net electric charge and pH of the solution. The pH 5.97 is point of inflection between two stages in its titration curve. At this point glycine is present pre-dominantly as its dipolar form. It is fully ionized but with no net electric charge. The characteristic pH at which the net electric charge is zero is called **isoelectric point or isoelectric pH**. It is designated **pI**. Glycine has no ionizable group in its side chain. Its isoelectric point is simply arithmetic mean of the two pKa values:

$$pI = \frac{1}{2}(pK_1 + pK_2) = \frac{1}{2}(2.34 + 9.60) = 5.97$$

Above figure shows that glycine has a net negative charge at any pH above its pI. Thus it will move toward positive electrode (the anode) when placed in an electric field. At any pH below its pI, glycine has a net positive charge. It will move toward the negative electrode (the cathode). The farther the pH of a glycine solution is from its isoelectric point, the greater the net electric charge of the population of glycine molecules. For example:

- At pH 1.0, glycine exists almost entirely as $^{+}H_3N\text{-}CH_2\text{-}COOH$ form. It has a net positive charge of 1.0.
- At pH 2.34, it is an equal mixture of $^{+}H_3N\text{-}CH_2\text{-}COOH$ and ($H_3N\text{-}CH_2COO^{-}$). Here average or net positive charge is 0.5.

The sign and the magnitude of the net charge of any amino acid at any pH can be predicted in the same way

Amino Acids Differ in Their Acid-Base Properties

Amino acids share may properties. But they show different acid base behaviors.

1. First, some amino acids with a single -amino group have a single-carboxyl group, and R group. They do not ionize. They have titration curves resembling that of glycine. These amino acids have very similar. But they do not have identical, pKa values. The pKa of the -COOH group in the range of 1.8 to 2.4, and pKa of the $-NH_3$ group in the range of 8.8 to 11.0.
2. Second, some amino acids have an ionizable R group. They have more complex titration curves. It has three stages corresponding to the three possible ionization steps. Thus they have three pKa values. The additional stage for the titration of the ionizable R group merges to some extent with the other two. The titration curves for two amino acids of this type, glutamate and histidine, are shown in Figure. The isoelectric points reflect the nature of ionizing R group. For example:
   - Glutamate has a pI of 3.22. It is lower than that of glycine. This is due to the presence of two carboxyl groups. Their average pKa values are (3.22). These two groups contribute a net charge of −1. This charge balances the +1 contributed by the amino group.

- Similarly, histidine has two positively charged groups when protonated. Its pI is 7.59 (the average of the pKa values of the amino and imidazole groups). It is much higher than that of glycine.

3. Finally, only histidine has an R group (pKa 6.0) in free and open aqueous environment. This R group provides significant buffering power near the neutral pH. It is found in the intracellular and extracellular fluids of most animals and bacteria.

PEPTIDES

Peptides are chains of amino acids. Two amino acid molecules are joined by **peptide bond**. Peptide bond is a substituted **amide linkage**.

1. Dipeptide

Two amino acids join to form a dipeptide. Such a linkage is formed by removal of water (dehydration). OH is removed from carboxyl group of one amino acid. Hydrogen is removed from amino group of another. Peptide bond formation is condensation reaction. Under standard biochemical conditions, equilibrium for the reaction favors the amino acids over dipeptide. Hence carboxyl group must be chemically modified or activated. It makes the reaction thermodynamically more favorable. Thus hydroxyl group can be more readily eliminated.

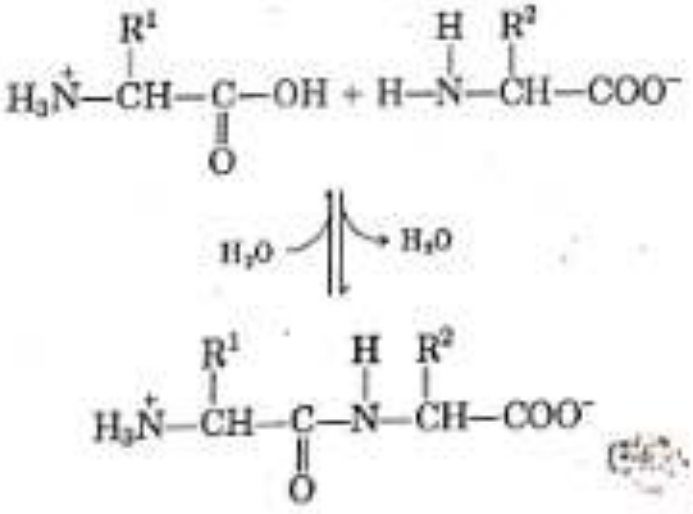


Fig: Formation of peptide bond

2. Tripeptide, tetrapeptides and pentapeptides

Three amino acids can be joined by two peptide bonds. They form a tripeptide. Similarly, amino acids can be linked to form tetrapeptides, pentapeptides, and so forth.

3. Oligopeptide

The structure formed by linking of few amino acids is called **oligopeptide**.

4. Polypeptide

The structure formed by joining of many amino acids is called **polypeptide**.

Proteins

Proteins may have thousands of amino acid residues. Mostly terms protein and polypeptide are used interchangeably. Generally, molecules having molecular weights

below 10,000is are called polypeptides. The molecules having higher molecular weights are called proteins.

Residue

An amino acid unit in a peptide is called a residue. It is part left over after losing a hydrogen atom from its amino group and hydroxyl moiety from its carboxyl group).

a) **Amino-terminal (or N- terminal):** In a peptide, amino acid residue at the end with a free -amino group is called amino-terminal residue;

b) **Carboxyl terminal (C-terminal):** The residue at the other end with a free carboxyl group is called carboxyl terminal (C-terminal) residue.

Stability of amino acids in peptides

The hydrolysis of a peptide bond is an exergonic reaction. But it occurs slowly because of its high activation energy. As a result, peptide bonds in proteins are quite stable. Their average half-life ($t_{1/2}$) is about 7 years under most intracellular condition.

Ionization behavior of Amino Acids

Two factors play role in ionization of amino acids:

1. **Free amine and carboxylic groups:** Peptides contain only one free amino group and one free carboxyl group. Both these are present at opposite ends of the chain. These groups ionize like in free amino acids. But an oppositely charged group is not linked to the carbon in peptides. Therefore, ionization constants of amino acids are different from that of peptides.
2. **Role of R group:** Amino and carboxyl groups of all non-terminal amino acids are covalently joined in peptide bonds. Thus they do not ionize. Hence they do not contribute to total acid-base behavior of peptides. But R groups of some amino acids can ionize. Thus in a peptide these contribute to the overall acid-base properties of the molecule.

Therefore, acid-base behavior of a peptide can be predicted from:

a) Its free amino and carboxyl groups
b) Nature and number of its ionizable R groups.

Peptides have characteristic titration curves like free amino acids. They also have characteristic isoelectric pH (pI). They do not move in an electric field at this pH. These properties are used in techniques used to separate peptides and proteins. An amino acid becomes a residue in a peptide.

Amino acid	Number of residues per molecule of protein*	
	Bovine cytochrome c	Bovine chymotypshogen
Ala	6	22
Arg	2	4
Asn	5	15
Asp	3	8
Cys	2	10
Gln	3	10
Glu	9	5
Gly	14	23
His	3	2
Ile	6	10
Leu	6	19
Lys	18	14
Met	2	2
Phe	4	6
Pro	4	9
Ser	1	28
Thr	8	23
Trp	1	8
Tyr	4	4
Val	3	23
Total	104	245

Then its pKa value for an ionizable R group can change.

Factors affecting pKa values

Following factors affect the pKa values:

a) Loss of charge in the -carboxyl and amino groups.
b) Interactions with other peptide.
c) R groups and other environmental factors.

Amino Acid Composition

Polypeptides have characteristic amino acid compositions. Hydrolysis of peptides or proteins with acid yields a mixture of free amino acids. When completely hydrolyzed, each protein yields a characteristic mixture of different amino acids. There are 20 common amino acids. They almost never occur in equal amounts in a protein. Some amino acids may occur only once or not at all in a given type of protein. Other may occur in large numbers. Table shows composition of amino acid mixtures. These are obtained on complete hydrolysis of bovine cytochrome c and chymotrypsinogen. Chymotrypsinogen is inactive precursor of digestive enzyme chymotrypsin.

Procedure of hydrolysis

These two proteins have very different functions. But they also differ significantly in the relative numbers of each kind of amino acid. Some side reactions occur during the hydrolysis. Therefore, complete hydrolysis is not sufficient for an exact analysis of amino acid composition.

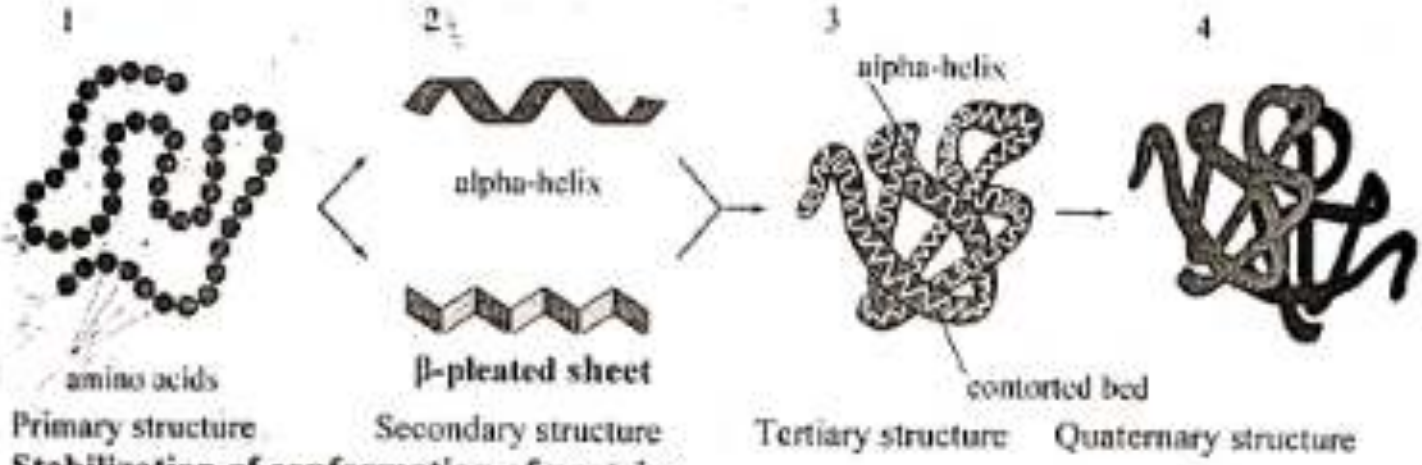
For example, acid treatment breaks amide bonds in the side chains of asparagine and glutamine. It yields aspartate and glutamate. The side chain of tryptophan is almost completely degraded by acid hydrolysis. Similarly, small amounts of serine, threonine and tyrosine are also lost. Biochemists use additional procedures for determining precise amino acid composition. It removes confusion in amino acid composition.

PROTEINS

Level of Structural Organization

Conformation

The spatial arrangement of atoms in a protein is called its conformation. The possible conformations of a protein include any structural state. This state must be achieved without breaking covalent bonds. A change in conformation can occur by rotation about single bonds. The changes must occur in most proteins as they bind to other molecules or catalyze reactions. The conformations existing under a given set of conditions must be thermodynamically most stable. It must have lowest Gibbs free energy (G). **Proteins in any of their functional, folded conformations are called native proteins.**



Stabilization of conformation of proteins

Weak Interactions are involved in conformation of proteins. The term **stability** can be defined as tendency to maintain a native conformation. Native proteins are only marginally stable. There are different chemical interactions that stabilize the native conformation. These include disulfide bonds and the weak (non-covalent) interactions, hydrogen bonds, and hydrophobic and ionic interactions.

About 200 to 460 kJ/mol are required to break a single covalent bond. But weak interactions can be disrupted by a mere 4 to 30 kJ/mol. Some individual covalent bonds contribute to native conformations of proteins. These are like disulfide bonds. They link separate parts of a single polypeptide chain. These bonds must be much stronger than individual weak interactions. But these are so numerous. Hence weak interactions predominate as a stabilizing force in protein structure. Generally, protein conformation with the lowest free energy is the one with the maximum number of weak interactions.

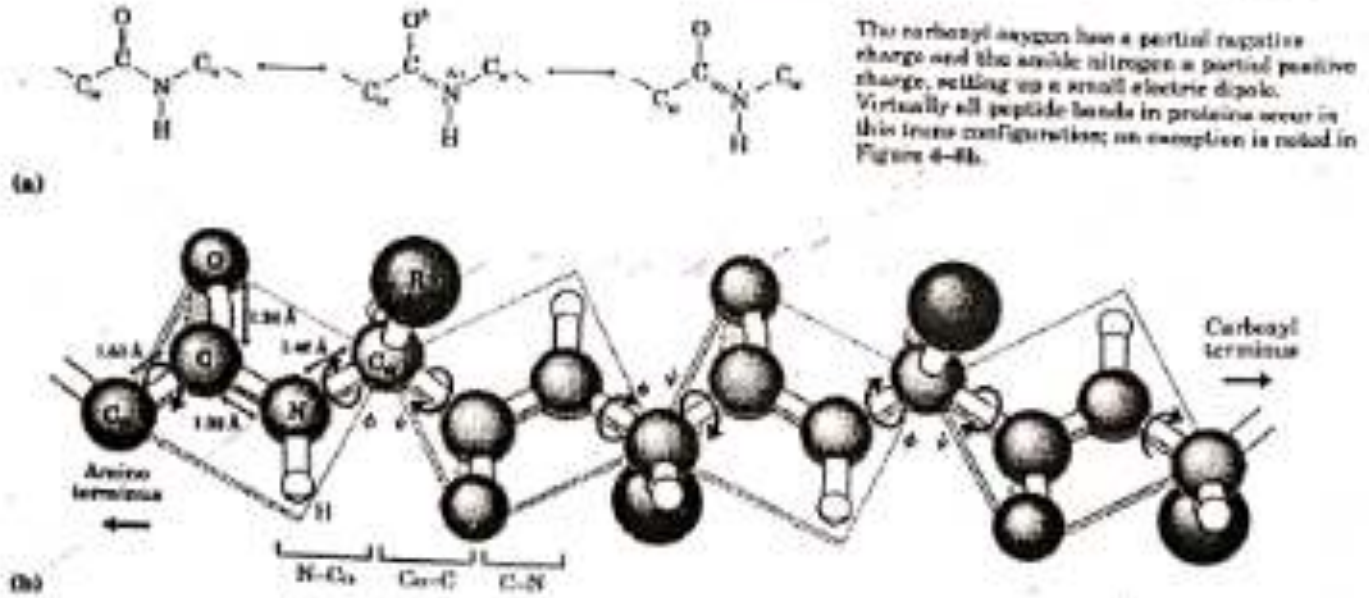


Fig: Primary Structure; Rotation in Peptide bond

Four levels of Protein Structure

Four levels of protein structure are:

(a) Primary Structure

Primary structure is a description of all covalent bonds (mainly peptide bonds and disulfide bonds) linking amino acid residues in a polypeptide chain. The most important element of primary structure is sequence of amino acid residues. Primary structure of a protein determines how it folds up into a unique three-dimensional structure. This structure in turn determines the function of the protein.

Peptide Bond is Rigid and Planar

Covalent bonds in amino acids restrict the conformation of a polypeptide. In the late 1930s, Linus Pauling and Robert Corey started different studies. These studies become basis of our present understanding of protein structure. They began with a careful analysis of the peptide bond. The α carbons of adjacent amino acid residues are separated by three covalent bonds. They are arranged as C_α–C–N–C_α. They made X-ray diffraction studies of crystals of amino acids and of simple dipeptides and tripeptides. They demonstrated that peptide C–N bond is shorter than C–N bond in a simple amine. The atoms associated with the peptide bond are co-planar. This indicated a resonance or partial sharing of two pairs of electrons between carbonyl oxygen and amide nitrogen. The oxygen has a partial negative charge and nitrogen a partial positive charge. It sets up a small electric dipole. The six atoms of the peptide group lie in a single plane. Here oxygen atom of the carbonyl group and hydrogen atom of amide nitrogen are trans to each other. From these findings Pauling and Corey concluded that peptide C–N bonds have partial double-bond character. Therefore, it is unable to rotate freely. Rotation is permitted about N–C_α and C_α-C bonds. Thus backbone of a polypeptide chain is series of rigid planes with consecutive planes. They share a common point of rotation at C_α. The rigid peptide bonds limit the range of conformations.

a) Determining sequence of amino acids

Amino acid sequences are determined by fragmenting polypeptides into smaller peptides.

- Different reagents are used to cleave specific peptide bonds.
- Then automated Edman degradation procedure is used to amino acid sequence of each fragment.
- Then ordering the peptide fragments by finding sequence overlaps between fragments.
- A protein sequence can also be determined from the nucleotide sequence of its corresponding gene in DNA.

Sequence of amino acids in different proteins

- **Sequence of Insulin molecule: F. Sanger** determines the sequence of first protein molecule. Sanger worked on insulin for ten years. He found that insulin is composed of 51 amino acids. It has two chains of amino acids. One chain has 21 amino acids. The other chain has 30 amino acids. Both these chains are held together by disulphide bridges (bonds).
- **Sequence of Haemoglobin molecule:** Haemoglobin is composed of 574 amino acids. These amino acids form four chains of amino acids. There are two alpha and two beta

chains. Each alpha chain has 141 amino acids and each beta chain has 146 amino acids.

b) Size of protein molecule

The size of protein depends on the types and numbers of amino acids present in a protein.

c) Specific arrangement of amino acids

There are 10,000 proteins in the human body. Each protein is composed of unique and specific arrangement of 20 types of amino acids. A sequence of nucleotides is present in the DNA. This sequence of nucleotides determines the sequence of amino acids in a protein molecule. The arrangement of amino acids is highly specific for the proper functioning of protein. If a single amino acid is not in its normal place, the protein fails to perform its normal functions. The best example is **sickle cell Haemoglobin** in human beings. In this case, one molecule of amino acid out of 574 amino acids is not in its normal place. This particular amino acid is replaced by some other amino acid. This Haemoglobin fails to carry sufficient oxygen. It may cause death of the patient.

2. Secondary structure of Proteins

Secondary structure is particularly stable arrangements of amino acid residues giving rise to recurring structural patterns. Secondary structure refers to the spatial arrangement of amino acid residues adjacent in the primary structure.

The term secondary structure is used for local conformation of some part of a polypeptide. Secondary structure focuses on common regular folding patterns of the polypeptide backbone. A few types of secondary structure are particularly stable. They occur widely in proteins. The most prominent are α-helix and β conformations. Pauling and Corey predicted the existence of these secondary structures in 1951.

(a) α-helix

It involves a spiral formation of basic polypeptide chain. Helical structure is the simplest arrangement of polypeptide with its rigid peptide bonds Pauling and Corey called it α-helix. In this structure:

1. Polypeptide backbone is tightly wound around an imaginary axis. This axis is drawn longitudinally through the middle of the helix.
2. The R groups of the amino acid residues protrude outward from helical back- bone.
3. The repeating unit is a single turn of the helix. This helix extends about 5.4 Å along the long axis.
4. The amino acid residues in an α-helix have conformations with $\psi = -45$ to -50 and $\varphi = -60$. Each helical turn includes 3.6 amino acid residues.
5. The helical twist of the helix found in all proteins is right-handed.

The α–helix is the predominant structure in α-keratins. Generally, about one-fourth of all amino acid residues in polypeptides are found in α helices. But their exact fraction varies greatly from one protein to the next.

Role of hydrogen bonding in α helix

The α-helix makes optimal use of internal hydrogen bonds. The structure is stabilized by a hydrogen bond. This bond is formed between hydrogen atom attached to

electronegative nitrogen atom of a peptide linkage and electronegative carbonyl oxygen atom of the fourth amino acid on the amino-terminal side of that peptide bond. Within α helix, every peptide bond (except those close to each end of the helix) participates in such hydrogen bonding. Each successive turn of α helix is held to adjacent turns by three to four hydrogen bonds. All the hydrogen bonds combined. These bonds give the entire helical structure considerable stability.

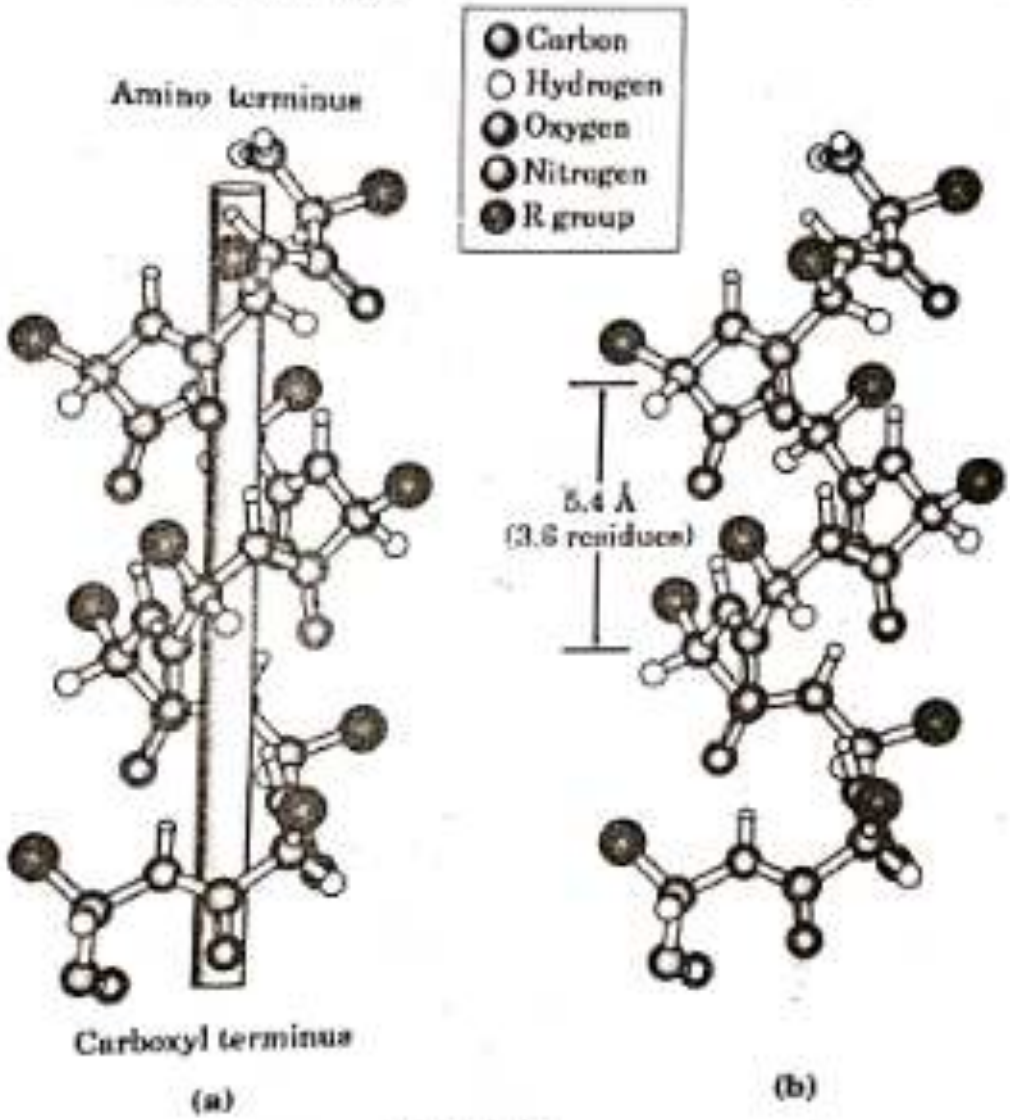


Fig: α helix

(b) β-pleated sheet

β-pleated sheet is formed by folding back of the polypeptide chain. Pauling and Corey predicted this second type of repetitive structure. This is a more extended conformation of polypeptide chains. Its structure is confirmed by x-ray analysis.

1. In the β conformation, backbone of the polypeptide chain is extended into a zigzag manner. The zigzag polypeptide chains can be arranged side by side to form a structure. This structure resembles a series of pleats.
2. In β-sheet, hydrogen bonds are formed between adjacent segments of polypeptide chain.

3. The individual segments that form β-sheet are nearby on the polypeptide chain. But these can also be quite distant from each other in linear sequence of the polypeptide. They may even be segments in different polypeptide chains.

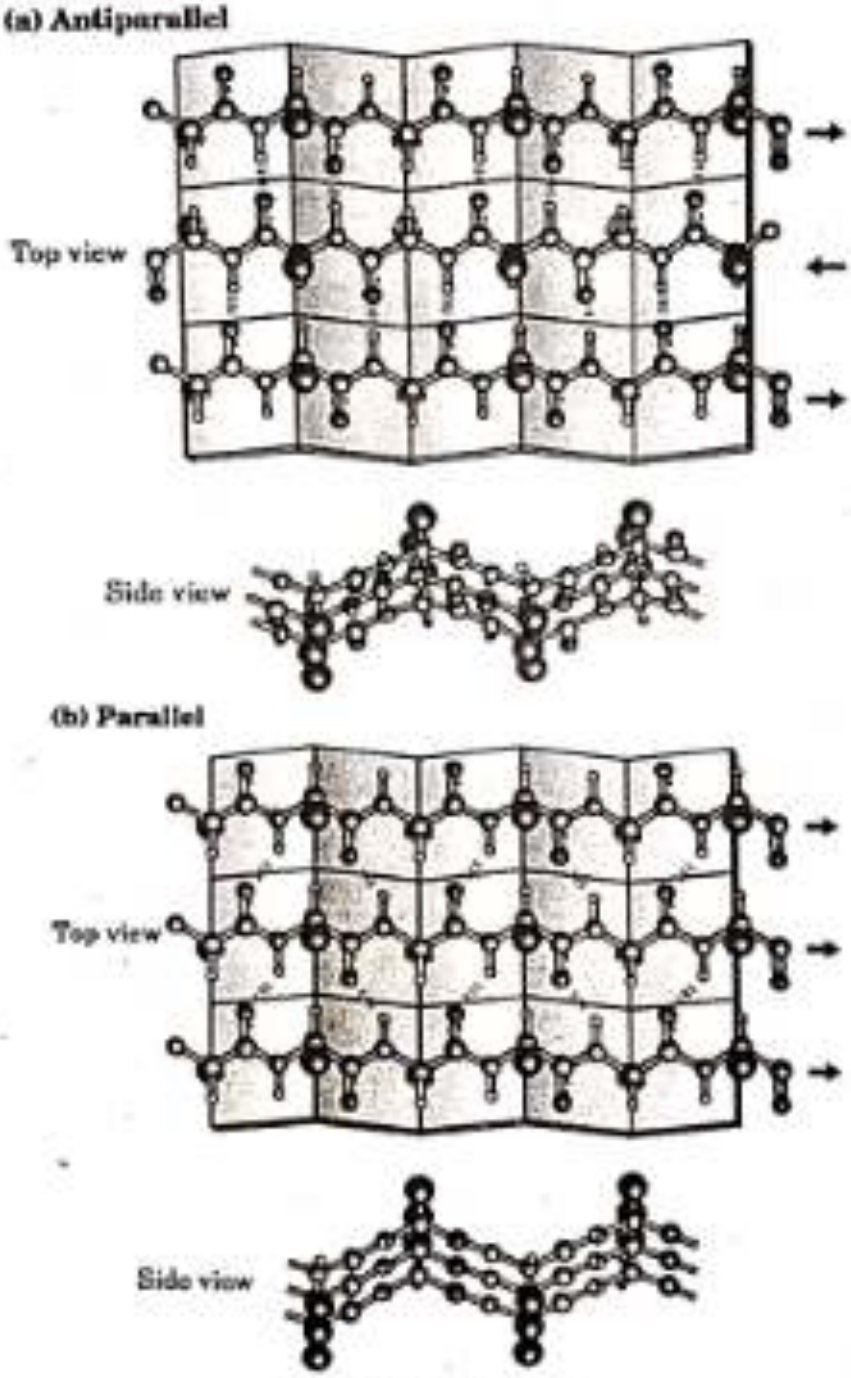


Fig: β-pleated sheet

4. The R groups of adjacent amino acids protrude from zigzag structure in opposite directions. It creates alternate pattern seen in the side views. The adjacent polypeptide chains in β sheet can be parallel or antiparallel. They may have same or opposite amino-to-carboxyl orientations. The structures are similar for the parallel conformation. But hydrogen- bonding patterns are different.
5. Some protein structures limit the kinds of amino acids that can occur in β sheet. In

some proteins, two or more β sheets are layered close together within a protein. In this case, R groups of the amino acid residues on the touching surfaces must be relatively small.

Its example is keratins like silk fibroin and fibroin of spider webs. They have a very high content of Gly and Ala residues. These two amino acids have smallest R groups. Hence in silk fibroin Gly and Ala alternate over large parts of the sequence.

(c) β-Turn

Globular proteins have a compact folded structure. Here nearly one-third of the amino acid residues are in turns or loops. The polypeptide chain reverses direction in these turns. These are the connecting elements. They link successive runs of α-helix or conformation.

β- turns

β- turns are particularly common. They connect the ends of two adjacent segments of an antiparallel β-sheet. The structure is a 180° turn. It involves four amino acid residues. The carbonyl oxygen of first residue forms a hydrogen bond with the amino-group hydrogen of the fourth. The peptide groups of the central two residues do not participate in any interresidue hydrogen bonding. Gly and Pro residues often occur in turns due to.

- Gly is small and flexible.
- Pro form peptide bonds involving imino nitrogen of proline. It readily assumes the cis configuration. It is a form that is particularly used for a tight turn.

There are several types of β- turns. But two are the most common. Beta turns are often found near the surface of a protein. Here peptide groups of the central two amino acid residues in the turn. Thus they can hydrogen-bond with water.

γ-turn

The γ-turn is much less common. It is three- residue turn. Here hydrogen bond is between the first and third residues.

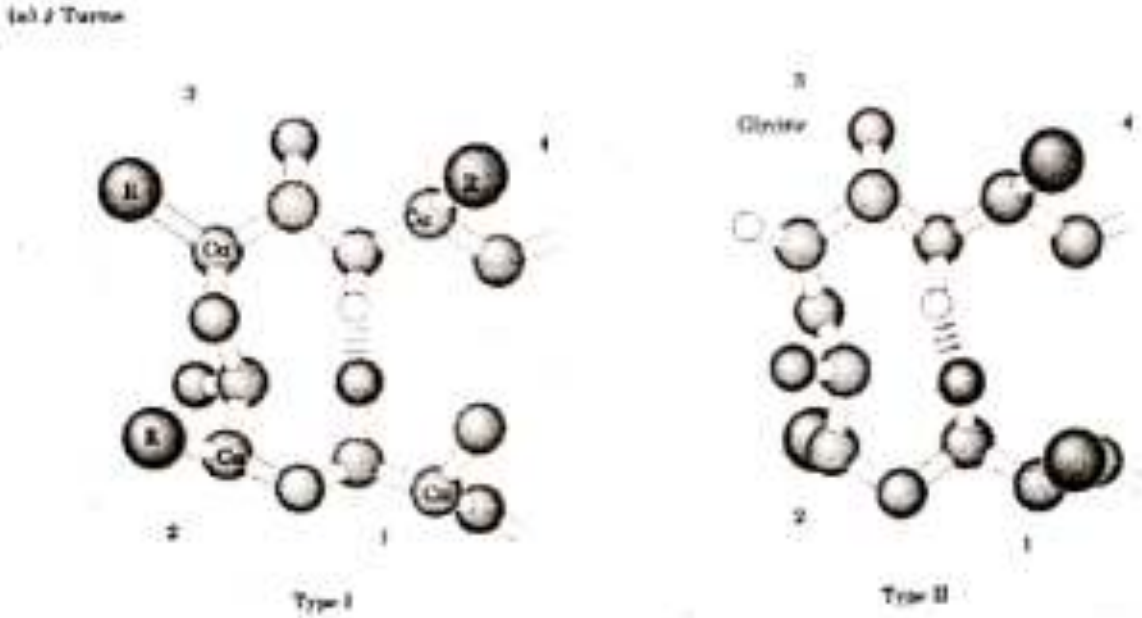
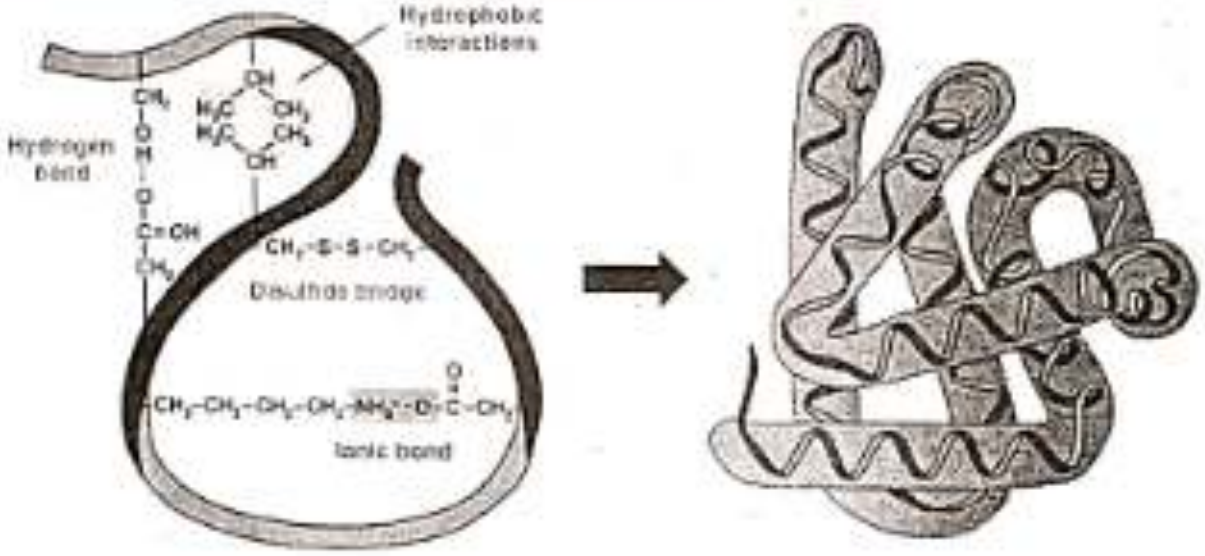


Fig: β-Turn

3. Tertiary structure

The overall three-dimensional arrangement of all atoms in a protein is called protein tertiary structure. The polypeptide chain bends and folds upon itself forming tertiary structure. Tertiary structure includes longer-range aspects of amino acid sequence. Amino acids are far apart in the polypeptide sequence in tertiary structure. These amino acids reside in different types of secondary structure. They may interact within the completely folded structure of a protein. These chains have different number and location of specific bend-producing residues like Pro, Thr, Ser, and Gly. These amino acids determine the location of bends (including β-turns) in the polypeptide chain. They also determine the direction and angle of these bends.



Types of side chain interactions Overall 3D shape (3° Structure)

Fig: Tertiary Structure

There are different kinds of weak bonding interactions between the segments in interacting polypeptide chains. These hold in interacting polypeptide chains in their characteristic tertiary positions. Sometimes, they are held by covalent bonds like disulfide cross-links. Tertiary confirmation of protein is maintained by three types of bonds: ionic, hydrogen and disulphide (S-S). For example, in aqueous environment, most stable tertiary confirmation is formed. The **hydrophobic** amino acids are buried inside. But the **hydrophilic** (water loving) amino acids are on the surface of the protein molecule.

Some proteins contain two or more separate polypeptide chains, or subunits. They may be identical or different. The arrangement of these protein subunits in three-dimensional complexes forms **quaternary structure.**

We classify proteins into two major groups for higher levels of structure:

a) **Fibrous proteins:** They have polypeptide chains arranged in long strands or sheets,
b) **Globular proteins:** They have polypeptide chains folded into a spherical or globular shape.

The two groups are structurally distinct.

a) Fibrous proteins consist largely of a single type of secondary structure. Globular

proteins often contain several types of secondary structure.

b) The two groups differ functionally. The structures that provide support, shape, and external protection to vertebrates are made of fibrous proteins. The most enzymes and regulatory proteins are globular proteins.

Certain fibrous proteins played a key role in our modern understanding of protein structure. These provide particularly clear examples of the relationship between structure and function.

4. Quaternary structure

In quaternary structure, polypeptide tertiary chains are aggregated and held together by hydrophobic interactions, hydrogen and ionic bonds. Haemoglobin is the example of the quaternary structure of protein. Haemoglobin carries oxygen of red blood cells.

Many proteins have multiple polypeptide subunits. The association of polypeptide chains can perform a variety of functions. Many multi-subunit proteins have regulatory roles. The binding of small molecules may affect the interaction between subunits. It causes large changes in the protein's activity in response to small changes in concentration of substrate or regulatory molecules. A multi-subunit protein is also called a **multimer**. Multiservice proteins can have from two to hundreds of subunits. A multimer with just a few subunits is often called an **oligomer**. In some cases, multimer is composed of a number of non-identical subunits. In this case, the overall structure of the protein can be asymmetric and quite complicated. However, most multimers have identical sub- units or repeating groups of non-identical subunits. These are symmetric arrangements.

The repeating structural unit in a multimeric protein is a single subunit or a group of sub-units is called a **protomer**.

Discovery of haemoglobin structure

Hemoglobin (Mr 64,500) was the first oligomeric protein for which three- dimensional structure was determined. It contains four polypeptide chains and four heme prosthetic groups. Here iron atoms are in the ferrous (Fe^{2+}) state. The protein portion is called globin. It consists of two α-chains (141 residues each) and two β chains (146 residues each). Max Perutz, John Kendrew in 1959 discovered three-dimensional structure of protein of haemoglobin by x-ray analysis. The subunits of hemoglobin are arranged in symmetric pairs. Each pair has one α and one β subunit. Therefore, hemoglobin can be described as a tetramer or as a dimer of αβ protomers.

Symmetric patterns in quaternary structures

Generally identical subunits of multimeric proteins are arranged in one or more symmetric patterns. Oligomers can have rotational symmetry or helical symmetry.

(i) Rotational symmetry

In this case, individual subunits can be superimposed on others by rotation about one or more rotational axes. In this symmetry, the subunits pack about the rotational axes. It forms closed structures.

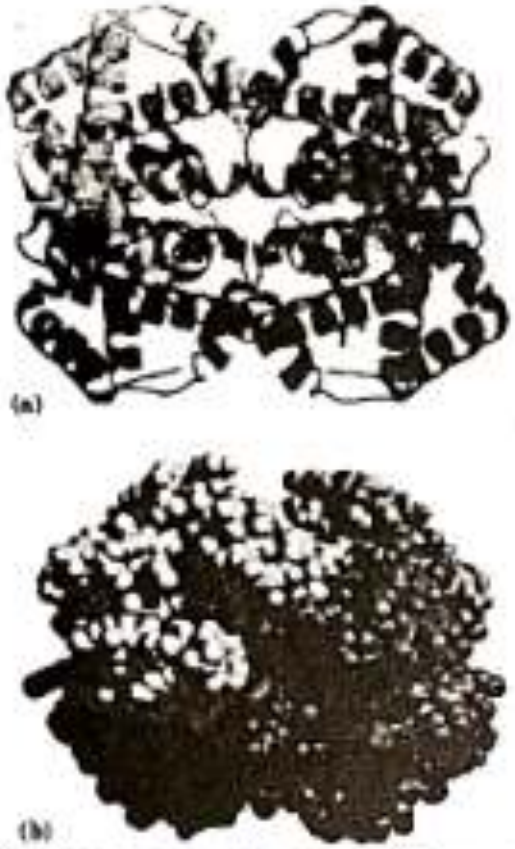
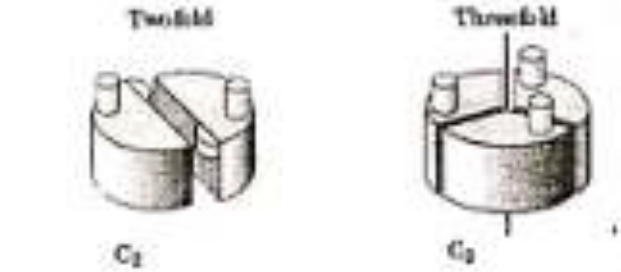
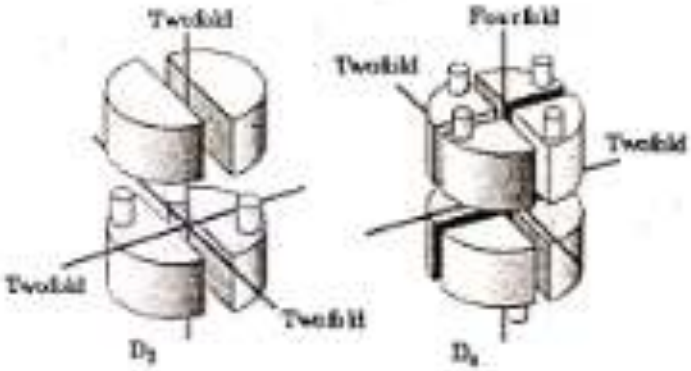


Fig: Quaternary structure of haemoglobin

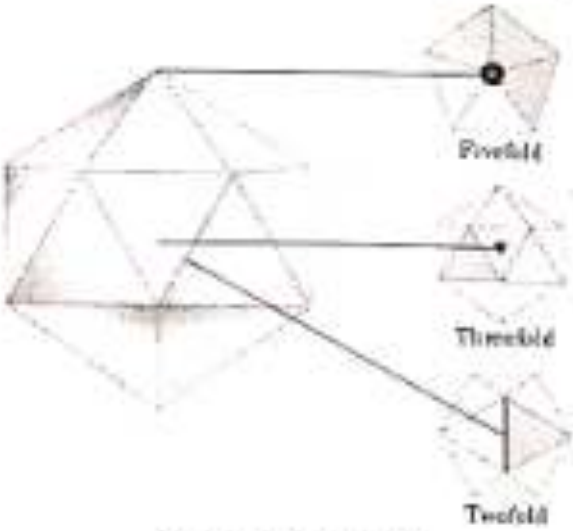


Two types of cyclic symmetry
(a)



Two types of dihedral symmetry
(b)

Fig: Two types of dihedral symmetry



Icosahedral symmetry
(c)

Fig: Icosahedral symmetry

(ii) Helical symmetry

They may be superimposed by a helical rotation. Proteins with helical symmetry tend to form more open-ended form structures. Here subunits added in a spiraling way.

Types of Rotational symmetry

There are several forms of rotational symmetry.

a) **Cyclic symmetry:** The simplest is cyclic symmetry. It involves rotation about a single axis. The protein with cyclic symmetry is defined by convention as C_n. Here C is for cyclic, n for the number of subunits related by the axis. The axis itself is described as an n-fold rotational axis. The αβ protomers of hemoglobin are related by C_2 symmetry.

b) **Dihedral symmetry:** It is a more complicated rotational symmetry. In this case, twofold rotational axis intersects an n-fold axis at right angles. The symmetry is defined as D_n. A protein with dihedral symmetry has 2n protomers.

c) **Icosahedral symmetry:** Proteins with cyclic or dihedral symmetry are particularly common. More complex rotational symmetries are possible. But only a few are regularly found. One example is **icosahedral symmetry**. An icosahedron is a regular 12-cornered polyhedron. It has 20 equilateral triangular faces. Each face can coincidence with another by rotation about one or more of three rotational axes. This is a common structure in virus coats, or capsids. The human poliovirus has an icosahedral capsid. Each triangular face is made up of three protomers. Each protomer containing single copies of four different polypeptide chains. Three of which are accessible at the outer surface. Sixty protomers form 20 faces of icosahedral shell. It encloses the genetic material (RNA).

Examples of structural and functional proteins

Structural Proteins

Proteins are classified into two groups on the basis of structure and shape proteins. These groups are fibrous and globular proteins.

Fibrous proteins

a) These proteins have fibre or filament like shape.
b) They exist in secondary structure during function.
c) These proteins are insoluble in aqueous medium
d) These are elastic in nature
e) These cannot be crystallized.
f) **Examples:** collagen (found in bone and cartilages), fibrinogen (blood clotting protein), actin, myosin (found in muscles and keratin (found in nails).

Globular proteins

a) These proteins have spherical or globules like shape.
b) They exist in tertiary or quaternary structure during function.
c) These proteins are soluble in aqueous medium
d) These are inelastic in nature
e) These can be crystallized.
f) **Examples:** enzymes, hormones, antibodies, channel proteins.

Functional proteins

Proteins are very important molecules in our cells. They are involved in all cell functions. Each protein within the body has a specific function.

1. **Structural role:** Some proteins are involved in support or composition of body parts. Hence they have structural roles
2. **Functional role:** Some proteins are involved in different physiological activities. These activities are like bodily movement or defense against germs. These are functional roles.

Table: List of structural proteins

Types	Role of Proteins
Collagen	It forms matrix (ground matter) of bone and cartilages.
Elastin	Elastin provides support for connective tissues. These are like tendons and ligaments.
Keratin	It strengthens protective coverings like hair, nails, quills, feathers, horns, and beaks.
Histone	It arranges DNA into chromosome.

Table: List of functional proteins

Types	Role of Proteins
Enzymes	Most of enzymes are protein. They control metabolism. They speed up biochemical reactions.
Hormones	Some hormones are protein in nature. These hormones regulate physiological activities like: • regulation of glucose level • regulation of calcium level • regulation digestion • regulation of blood pressure
Antibodies	These proteins are produced by WBCs. These are produced in response to antigen (a foreign particle). They provide immunity.
Haemoglobin	It is found in RBCs. It mainly transports oxygen. It transports small amount of carbon dioxide.
Fibrinogen	It is found in blood plasma. It is involved in blood clotting process.
Ovalbumin and Casein	Ovalbumin is found in egg whites. Casein is a milk-based protein. Both store amino acids.

CONJUGATED PROTEINS

Some proteins contain chemical groups other than amino acids. Many proteins, for example enzymes ribonuclease A and chymotrypsin, contain only amino acid residues and no other chemical constituents. These are called **simple proteins**. However, some proteins contain permanently associated chemical components in addition to amino acids. These are called conjugated proteins. The non–amino acid part of a conjugated protein is usually called its **prosthetic group**.

Classification

Conjugated proteins are classified on basis of chemical nature of their prosthetic groups. For example:

1. Lipoproteins contain lipids, glycoproteins contain sugar groups.
2. The metalloproteins contain a specific metal.
3. Some proteins contain more than one prosthetic group.

Function

The prosthetic group plays an important role in the protein's biological function.

Table: Some Conjugated Proteins

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

EXERCISE

LONG QUESTIONS

1. Describe structure of amino acid.
2. Discuss optical activity in amino acid and proteins.
3. Write note on: DL system and RS systems.
4. Describe different basis of amino acid classification.
5. How will you justify that amino acids are amphoteric?
6. What is meant by titration curve? Give its description for amino acids.
7. What are applications or information imparted by titration curve of amino acids.
8. How do amino acids differ in the acid base properties?
9. What are peptides? Elucidate their classification and formation.
10. Primary structure is the basis of Protein configuration. Discuss different aspects of primary structure of proteins.
11. What is secondary structure of protein?
12. What is alpha helix and beta pleated sheaths? Discuss their role in protein configuration?
13. What is meant by beta turn? How is it formed?

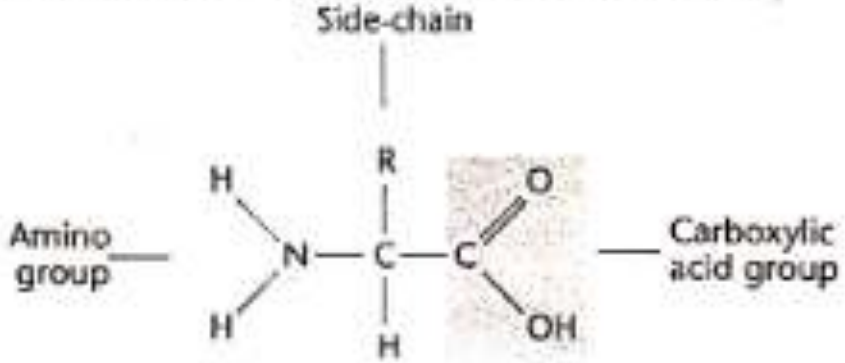
14. What is tertiary structure of protein? How proteins are classified on the basis of tertiary structure.
15. Discuss quaternary structure of protein with an example of haemoglobin.
16. Demonstrate presence of different symmetric patterns in quaternary structure of protein.
17. Discuss structural and functional protein with examples.

SHORT QUESTIONS

1. Define standard amino acid. Draw its general formula.

Ans:

All amino acids have an amino group (NH_2) and a carboxyl group (COOH), attached to the same carbon atom, also known as alpha carbon. 170 types of amino acids are present in cells and tissues. They have general formula as:



2. Give the method of naming of amino acids.

Ans: All the amino acids have trivial or common names. In some cases, their names derived from the source. They were first isolated from this source.

- Asparagine was first found in asparagus
- Glutamate was found in wheat gluten
- Tyrosine was first isolated from cheese. Its name is derived from the Greek word tyros, means cheese.

Glycine has sweet taste. Greek word glykos means sweet.

3. How is carbon in amino acid chiral?

Ans: For all amino acids, except glycine, the carbon is bonded to four different groups in common amino acids. These include a carboxyl group, an amino group, an R group, and a hydrogen atom. In glycine, R group is another hydrogen atom. The α-carbon atom is thus a **chiral** center. There is tetrahedral arrangement of the bonding orbitals around α - carbon atom. Here four different groups can occupy two unique spatial arrangements.

4. How do amino acids have two possible stereoisomers?

Ans: The α-carbon atom is thus a **chiral** center. There is tetrahedral arrangement of the bonding orbitals around α -carbon atom. Here four different groups can occupy two unique spatial arrangements. Therefore, amino acids have two possible **stereoisomers.** They are non-superimposable mirror images of each other.

5. What is meant by DL system of optical activity? Who did discover it?

Ans: Special nomenclature is used to specify absolute configuration of four substituents of asymmetric carbon atoms. The absolute configurations of simple sugars and amino acids are specified by D, L system. It is based on the absolute configuration of three-carbon sugar **glyceraldehyde**. This convention was proposed by **Emil Fischer** in 1891. It was later confirmed by x-ray diffraction analysis.

6. How amino acids are classified? Give its basis.

Ans:

Amino acids can be classified by R group. Amino acids are classified into five main classes. This classification is based on the properties of their R groups (See Table). Particularly these properties are:

a) Polarity or R

b) Or their tendency to interact with water at biological pH (near pH 7.0).

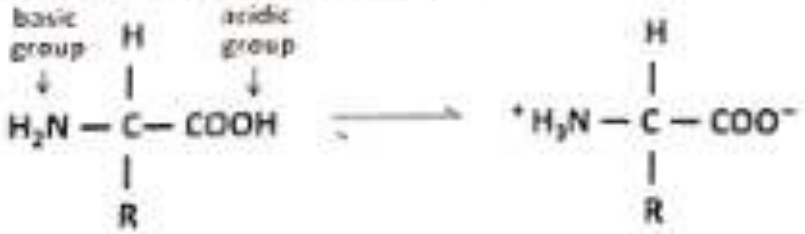
The polarity of the R groups varies widely. These vary from non-polar and hydrophobic (water-insoluble) to highly polar and hydrophilic (water-soluble). There are gradations of polarity, size, and shape of R groups within each class.

7. What is disulphide bond or bridge? In which amino acid is it formed?

Ans: Cysteine is readily oxidized to form **cystine**. It is covalently linked diameric amino acid. Two amino acids are joined by disulphide in cysteine. Disulfide linked residues are strongly hydrophobic (non-polar). Disulphide bond play a special role in structures of many proteins.

8. How do amino acids act as zwitterions? Give example.

Ans: Amino acid is dissolved in water. Amino acid is dipolar ion in this solution. It is called **Zwitterions** (German for "hybrid ion").



amino acid zwitterion

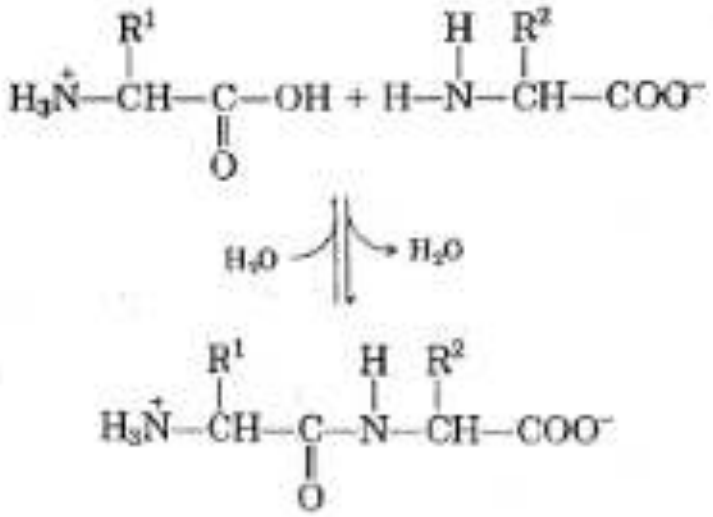
A zwitterion can act as either an acid (proton donor): or a base (proton acceptor)

9. Draw titration curve of glycine.

Ans: See in the text of chapter.

10. What are dipeptides? How are they formed?

Ans: Two amino acids join to form a dipeptide. Such a linkage is formed by removal of water (dehydration). OH is removed from carboxyl group of one amino acid. Hydrogen is removed from amino group of another. Peptide bond formation is condensation reaction. Under standard biochemical conditions, equilibrium for the reaction favors the amino acids over dipeptide. Hence carboxyl group must be chemically modified or activated. It makes the reaction thermodynamically more favorable. Thus hydroxyl group can be more readily eliminated.



11. Why do bonds in amino acids are stable?

Ans: The hydrolysis of a peptide bond is an exergonic reaction. But it occurs slowly because of its high activation energy. As a result, peptide bonds in proteins are quite stable. Their average half-life ($t_{1/2}$) is about 7 years under most intracellular condition.

12. What factors do affect the pka values of amino acids?

Ans:

Factors affecting pKa values

a) Following factors affect the pKa values:
b) Loss of charge in the -carboxyl and amino groups.
c) Interactions with other peptide.
d) R groups and other environmental factors.

13. What are meant by conformation and native proteins?

Ans: The spatial arrangement of atoms in a protein is called its conformation. The possible conformations of a protein include any structural state. This state must be achieved without breaking covalent bonds. **Proteins in any of their functional, folded conformations are called native proteins.**

14. What is primary structure of proteins?

Ans: Primary structure is a description of all covalent bonds (mainly peptide bonds and disulfide bonds) linking amino acid residues in a polypeptide chain. The most important element of primary structure is sequence of amino acid residues. Primary structure of a protein determines how it folds up into a unique three-dimensional structure. This structure in turn determines the function of the protein.

15. Why do we use term secondary structure for proteins?

Ans: Secondary structure is particularly stable arrangements of amino acid residues giving rise to recurring structural patterns. Secondary structure refers to the spatial arrangement of amino acid residues adjacent in the primary structure.

The term secondary structure is used for local conformation of some part of a polypeptide. Secondary structure focuses on common regular folding patterns of the polypeptide backbone. A few types of secondary structure are particularly stable. They

occur widely in proteins. The most prominent are α-helix and β conformations. Pauling and Corey predicted the existence of these secondary structures in 1951.

16. What is alpha helix?

Ans: Alpha helix involves a spiral formation of basic polypeptide chain. Helical structure is the simplest arrangement of polypeptide with its rigid peptide bonds Pauling and Corey called it α-helix.

17. Elucidate the role of hydrogen bonding alpha helix formation.

Ans: The α-helix makes optimal use of internal hydrogen bonds. The structure is stabilized by a hydrogen bond. This bond is formed between hydrogen atom attached to electronegative nitrogen atom of a peptide linkage and electronegative carbonyl oxygen atom of the fourth amino acid on the amino-terminal side of that peptide bond. Within α helix, every peptide bond (except those close to each end of the helix) participates in such hydrogen bonding. Each successive turn of α helix is held to adjacent turns by three to four hydrogen bonds. All the hydrogen bonds combined. These bond give the entire helical structure considerable stability.

18. What is beta sheet? Name some proteins in which it is formed.

Ans: β-pleated sheet is formed by folding back of the polypeptide chain. Pauling and Corey predicted this second type of repetitive structure. This is a more extended conformation of polypeptide chains. Its example is keratins like silk fibroin and fibroin of spider webs.

19. What is the function of β- turns

Ans: β- turns are particularly common. They connect the ends of two adjacent segments of an antiparallel β-sheet. The structure is a 180° turn. It involves four amino acid residues. The carbonyl oxygen of first residue forms a hydrogen bond with the amino-group hydrogen of the fourth.

20. What is the tertiary structure? How is it formed?

Ans: The overall three-dimensional arrangement of all atoms in a protein is called protein tertiary structure. The polypeptide chain bends and folds upon itself forming tertiary structure. Tertiary structure includes longer-range aspects of amino acid sequence. Amino acids are far apart in the polypeptide sequence in tertiary structure. These amino acids reside in different types of secondary structure. They may interact within the completely folded structure of a protein. These chains have different number and location of specific bend-producing residues like Pro, Thr, Ser, and Gly. These amino acids determine the location of bends (including β-turns) in the polypeptide chain. They also determine the direction and angle of these bends.

21. What are fibrous and globular proteins?

Ans:

a) **Fibrous proteins:** They have polypeptide chains arranged in long strands or sheets,
b) **Globular proteins:** They have polypeptide chains folded into a spherical or globular shape.

22. What is quaternary structure? Give its example.

Ans: In quaternary structure, polypeptide tertiary chains are aggregated and held together by hydrophobic interactions, hydrogen and ionic bonds. Haemoglobin is the example of the quaternary structure of protein. Haemoglobin carries oxygen of red blood cells.

23. What are rotational and helical symmetries of proteins molecules?

Ans:

a) **Rotational symmetry:** In this case, individual subunits can be superimposed on others by rotation about one or more rotational axes. In this symmetry, the subunits pack about the rotational axes. It forms closed structures.

b) **Helical symmetry:** They may be superimposed by a helical rotation. Proteins with helical symmetry tend to form more open-ended form structures. Here subunits added in a spiraling way.

24. Write different characteristics and example of fibrous proteins.

Ans:

Fibrous proteins

a) These proteins have fibre or filament like shape.
b) They exist in secondary structure during function.
c) These proteins are insoluble in aqueous medium
d) These are elastic in nature
e) These cannot be crystallized.
f) **Examples:** collagen (found in bone and cartilages), fibrinogen (blood clotting protein), actin, myosin (found in muscles and keratin (found in nails).

25. Write different characteristics and example of globular proteins.

Ans:

Globular proteins

a) These proteins have spherical or globules like shape.
b) They exist in tertiary or quaternary structure during function.
c) These proteins are soluble in aqueous medium
d) These are inelastic in nature
e) These can be crystallized.

Examples: enzymes, hormones, antibodies, channel proteins.

MULTIPLE CHOICE QUESTIONS (MCQs)

1. Which was the first amino acid discovered?

(a) Asparagine (b) threonine (c) Methionine (d) Lycine

2. The amino acid first found from cheese was:

(a) Glutamate (b) threonine (c) Methionine (d) Tyrosine

3. Which amino acid has sweet taste?

(a) Asparagine (b) threonine (c) Methionine (d) Lycine

4. The chiral centre in amino acid is:

(a) α-carbon (b) β-carbon (c) γ-carbon (d) ε-carbon

5. Amino acids have number possible stereoisomers?

(a) 1 (b) 2 (c) 3 (d) 4

6. DL configuration is based on:

(a) glucose (b) glyceraldehyde (c) starch (d) lycine

7. The system of specifying configuration around a chiral center is;

(a) DL system (b) ld system (c) RS system (d) none

8. Which of the amino acid has Non-polar, Aliphatic R Groups?

(a) serine (b) Methionine (c) tryptophan (d) Lysine

9. Which of the amino acid has Aromatic Groups?

(a) serine (b) Methionine (c) tryptophan (d) Lysine

10. Which of the amino acid has Positively charged Groups?

(a) serine (b) Methionine (c) tryptophan (d) Lysine

11. Which of the amino acid has Polar uncharged Groups?

(a) serine (b) Methionine (c) tryptophan (d) Lysine

12. The pKa value for COOH is:

(a) 1.34 (b) 2.34 (c) 3.34 (d) 4.34

13. Peptide bond is a substituted:

(a) ester linkage (b) disulphide linkage (c) glycosidic linkage (d) amide linkage

14. molecules having molecular weights below _____is are called polypeptides

(a) 4000 (b) 6000 (c) 10,000 (d) 20000

15. The number of amino acids in insulin are:

(a) 41 (b) 31 (c) 51 (d) 61

16. The alpha chain of haemoglobin have amino acids:

(a) 174 (b) 171 (c) 141 (d) 146

17. Each turn of alpha helix has amino acids:

(a) 2 (b) 3 (c) 3.5 (d) 3.6

18. Which of the followings is not fibrous protein?

(a) Silk (b) Fibrin (c) Keratin (d) Haemoglobin

19. Which of the followings is Globular protein?

(a) Silk (b) Fibrin (c) Keratin (d) enzyme

20. Most part of the body of the animals is composed of:

(a) Protein (b) Lipids (c) Carbohydrates (d) Nucleic acid

21. The primary structure of protein is composed of:

(a) Number of amino acids (b) Sequence of amino acids
(c) (a) and (b) (d) Coiling and folding of amino acids

22. Which of the followings is not a primary structure of protein?

(a) Number of amino acids (b) Sequence of amino acids
(c) arrangement of amino acids (d) coiling folding of amino acids

23. When we say haemoglobin is composed of 574 molecules. It is a:

(a) Primary structure (b) Secondary structure
(c) Tertiary structure (d) Quaternary structure

24. Haemoglobin has:

(a) Primary and secondary
(b) Tertiary structure
(c) Quaternary structure
(d) All structures

25. Enzyme are chemically

(a) Primary proteins
(b) Fibrous proteins
(c) Globular proteins
(d) None of the above

Ans:

1. (a) 2. (d) 3. (d) 4. (a) 5. (b) 6. (b) 7. (d) 8. (b) 9. (c) 10. (d) 11. (c) 12. (b) 13. (d) 14. (c) 15. (c) 16. (d) 17. (c) 18. (d) 19. (d) 20. (a) 21. (c) 22. (d) 23. (a) 24. (d) 25. (c)

Future Zoologist Group of PGCG (By Muhammad ARSALAN)

Chapter 2

Proteins II

Syllabus 2012-13
Not included in this syllabus
Syllabus 2018-19
2.1 Purification Techniques for Proteins a. An outline of purification techniques for proteins; column chromatography, electrophoresis; Isoelectric focusing;
2.2 Organization of proteins: a. Structural levels of proteins; Covalent structure of proteins; b. function of some structural & functional proteins; Hemoglobin, Cytochrome-c: Chymotrypsin, alpha Keratin and Collagen; c. Proproteins, their examples and role;

PURIFICATION TECHNIQUES FOR PROTEINS

Biochemists have used many techniques to understand protein structure and function. The researcher must separate a protein from other proteins in pure form. Then he can study the structure and function of proteins. There are different techniques to determine structure and properties of proteins.

Proteins Can Be Separated and Purified

A pure preparation is essential for determination of properties and activities of proteins. The cells contain thousands of different kinds of proteins. There are different methods to purify these proteins. Classical methods for separating proteins are based on properties of proteins. These properties vary from one protein to the next. These properties include size, charge, and binding properties. Some additional modern methods involves DNA cloning and genome sequencing. These can simplify the process of protein purification.

Preparation of Extract

The source of a protein is generally tissue or microbial cells. The first step in any protein purification procedure is to break open these cells. Their proteins are released into a solution called a **crude extract**. In some cases, differential centrifugation can be used to

prepare sub-cellular fractions or to isolate specific organelles. Now extract of organelle is ready. Now various methods are available for purifying one or more of the proteins it contains.

Fractionation

This extract is subjected to **fractionation**. Fractionation separates the proteins into different fractions. This treatment is based on a property like size or charge. Early fractionation steps in purification utilize differences in protein **solubility**. Solubility is a complex function of pH, temperature, salt concentration, and other factors. The solubility of proteins is lowered at high salt concentrations. This effect is called **salting out**. The addition of certain salts in the right amount can selectively precipitate some proteins. But others remain in solution. Ammonium sulfate ($(NH_4)_2SO_4$) is particularly effective. It is often used to salt out proteins. The precipitated proteins are removed from those remaining in solution by **low-speed centrifugation**. A solution containing the protein of interest must be further altered. Then subsequent purification steps are possible.

Dialysis

Dialysis is used to separates proteins from small solutes from large proteins molecules. The partially purified extract is placed in a bag or tube. This tube is made of a semi-permeable membrane. This extract is suspended in a much larger volume of buffered solution of appropriate ionic strength. The membrane allows the exchange of salt and buffer but not proteins. Thus dialysis retains large proteins within the membranous bag or tube. But it changes the concentration of other solutes in the protein preparation. Finally, they come into equilibrium with the solution outside the membrane. For example dialysis may be used to remove ammonium sulfate from the protein preparation.

COLUMN CHROMATOGRAPHY

Column chromatography is the most powerful methods for fractionating proteins. Chromatography is based on differences in protein charge, size, binding affinity, and other properties.

A porous solid material with appropriate chemical properties is selected as stationary phase. This material is held in a column. A buffered solution percolates through it. This solution acts as a mobile phase. The protein-containing solution is layered on the top of the column. This solution percolates through the solid matrix as an ever-expanding band within the larger mobile phase. Individual proteins migrate faster or more slowly through the column depending on their properties.

Procedure

1. The standard elements of a chromatographic column include a solid, porous material (matrix). It is supported inside a column. This column is generally made of plastic or glass.
2. A solution is used as a mobile phase. It flows through the matrix. This matrix acts as stationary phase.
3. The solution that passes out of the column at the bottom is called **effluent**. It is constantly replaced by solution supplied from a reservoir at the top.

Future Zoologist Group of PGCG (By Muhammad ARSALAN)

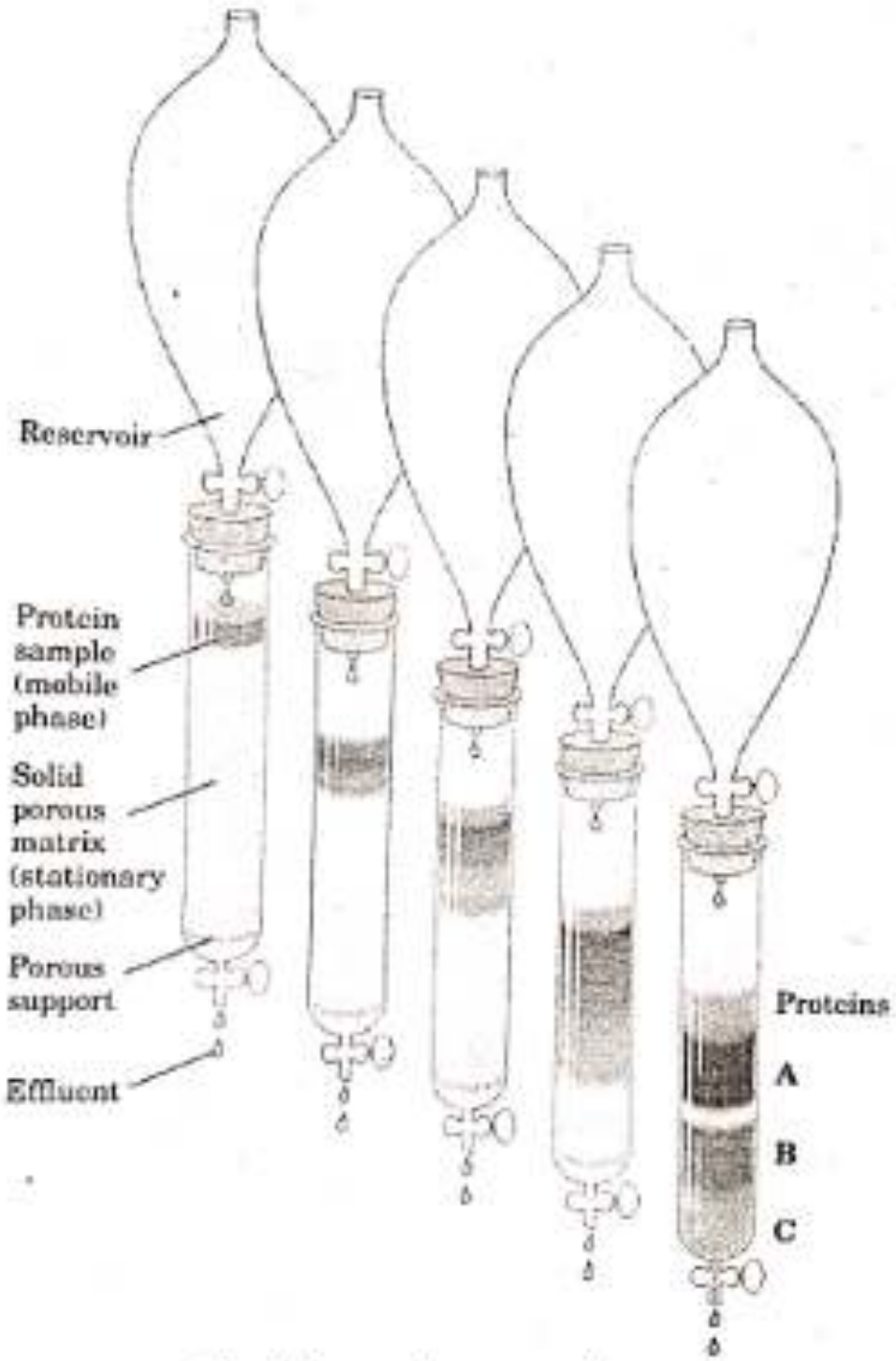


Fig: Column chromatography.

4. The protein solution is layered on top of the column. This layer is allowed to percolate into the solid matrix.
5. Additional solution is added on top. The protein solution forms a band within the mobile phase. The proteins migrate through the column. These proteins are retarded to different degrees by their different interactions with the matrix material. The overall protein band thus widens as it moves through the column.
6. Individual types of proteins (like A, B, and C,) gradually separate from each other. They form bands within the broader protein band. Separation improves as the length

of the column increases. However, each individual protein band also broadens with time. It is broaden due to diffusional spreading. It is a process that decreases resolution.

7. In this example, protein A is well separated from B and C. but diffusional spreading prevents complete separation of B and C under these conditions.

ELECTROPHORESIS

Proteins can be separated and characterized by electrophoresis. It is another important technique for the separation of proteins. It is based on the migration of charged proteins in an electric field. This process is called electrophoresis.

Disadvantages

Electrophoresis is not generally used to purify proteins in large amounts due to two reasons:

a) Simpler cheap alternatives methods are available
b) Electrophoretic methods often adversely affect structure. It in turn affects the function of proteins.

Advantage

a) Electrophoresis is especially useful as an analytical method.
b) Its advantage is that proteins can be visualized during separation. It permits a researcher to estimate number of different proteins in a mixture or degree of purity of a particular protein preparation.
c) Electrophoresis allows determination of crucial properties of a protein. These properties include isoelectric point and approximate molecular weight.

Procedure

1. Use of gel

Electrophoresis of proteins is carried out in gels. These gels are made up of the cross-linked polymer polyacrylamide. The polyacrylamide gel acts as a molecular sieve. It slows down the migration of proteins in proportion to their charge-to-mass ratio. Migration may also be affected by protein shape.

2. Electrophoretic mobility (μ)

In electrophoresis, electrical potential (E) is the force moving the macromolecule. The **electrophoretic mobility (μ)** of a molecule is ratio of its velocity (V) to the electrical potential. Electrophoretic mobility is also equal to the net charge (Z) of the molecule divided by frictional coefficient (f). It reflects shape of protein. Thus:

$$\mu = \frac{V}{E} = \frac{Z}{f}$$

Therefore, migration of a protein in a gel during electrophoresis is a function of its size and its shape.

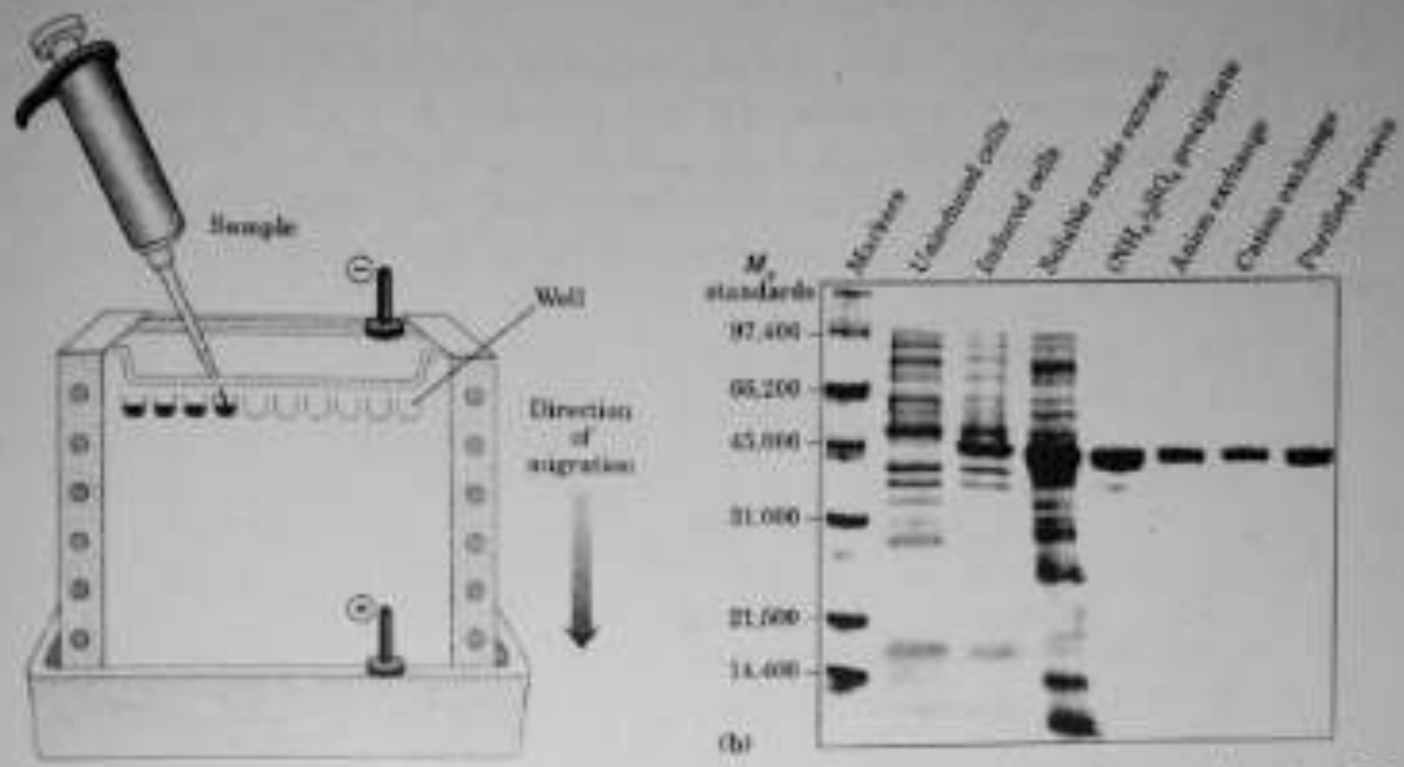
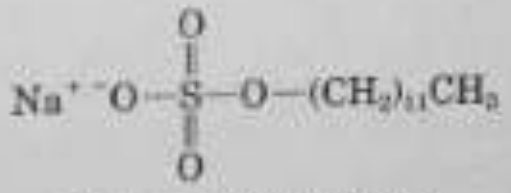


Fig: Electrophoresis. (a) Different samples are loaded in wells or depressions at the top of the polyacrylamide gel. (b) Proteins can be visualized after electrophoresis by treating the gel with a stain such as Coomassie blue

3. Electrophoretic method

An electrophoretic method is commonly used for estimation of purity and molecular weight. It uses detergent sodium dodecyl sulfate (SDS) (dodecyl denoting a 12-carbon chain).



Sodium dodecyl sulfate
(SDS)

SDS binds to most proteins in proportional to molecular weight of the protein. About one molecule of SDS binds for every two amino acid residues. The bound SDS contributes a large net negative charge. It makes the intrinsic charge of the protein insignificant. Thus it gives each protein a similar charge-to-mass ratio. In addition, SDS binding partially unfolds proteins. Therefore, most SDS-bound proteins assume a similar shape. Therefore, electrophoresis in the presence of SDS separates proteins almost exclusively on the basis of mass (molecular weight). Here smaller polypeptides migrating more rapidly.

4. Visualization of Proteins

After electrophoresis, the proteins are visualized by adding a dye like **Coomassie blue**. It binds to proteins but not to the gel itself. Thus, a researcher can monitor the progress of a protein purification procedure. He can visualize number of protein bands visible on the gel decreases after each new fractionation step.

5. Determination of Position of Unidentified Protein and Molecular Weight

The number of bands is compared with the positions to which proteins of known molecular weight migrate in the gel. It gives the position of an unidentified protein. It can provide a good approximation of its molecular weight. Some protein has two or more different subunits. These subunits are separated by the SDS treatment. Thus a separate band appears for each.

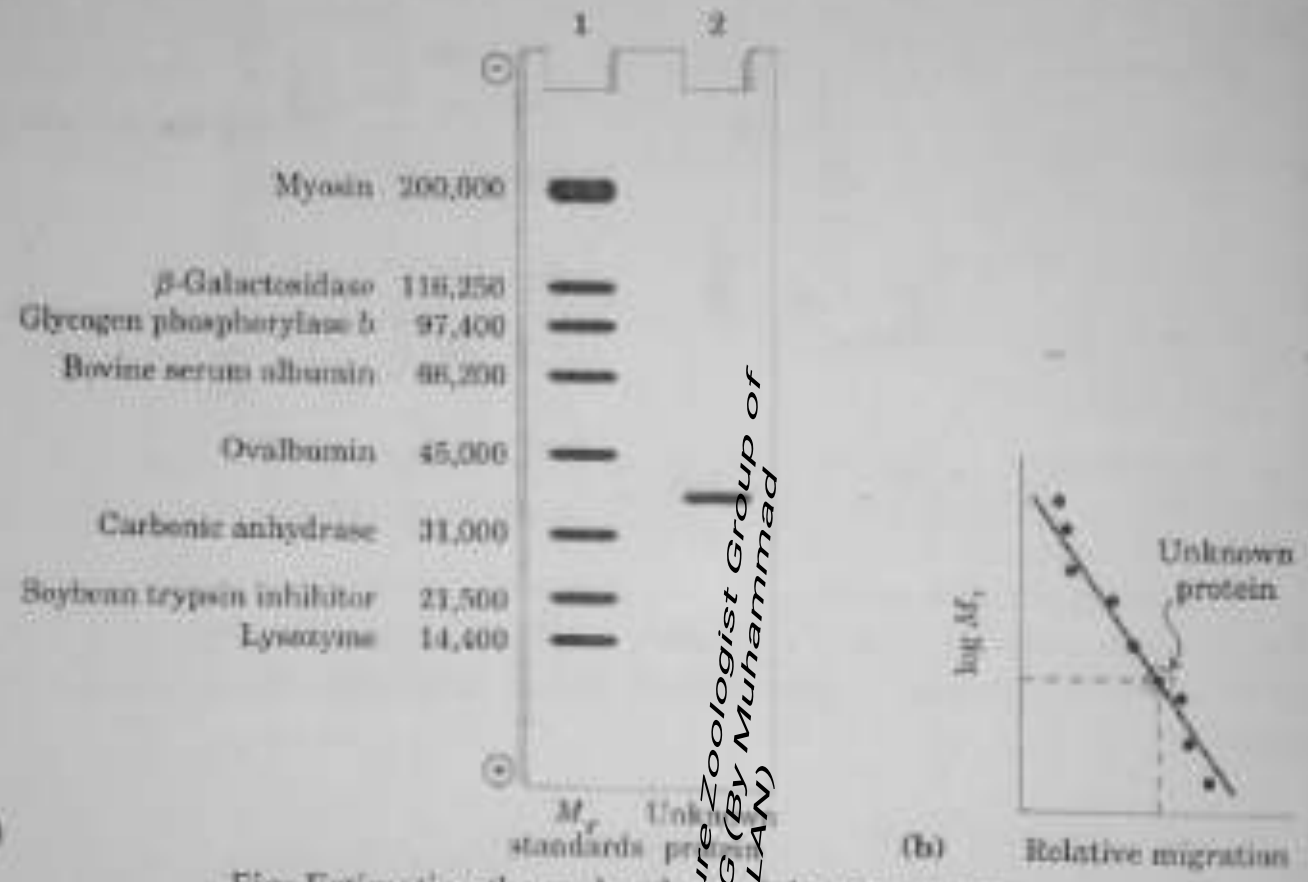


Fig: Estimating the molecular weight of a protein

Future Zoologist Group of PGCG (By Muhammad ARSALAN)

ISOELECTRIC FOCUSING

Isoelectric focusing is used to determine the isoelectric point (pI) of a protein.

1. In this procedure, a pH gradient is established. It established by allowing a mixture of low molecular weight organic acids and bases (ampholytes; p. 79) to distribute themselves in an electric field. This electric field is generated across the gel.
2. Then a protein mixture is applied. Each protein migrates until it reaches the pH that matches its pI. Proteins with different isoelectric points are distributed differently throughout the gel.
3. In two dimensional electrophoresis, isoelectric focusing and SDS electrophoresis are sequentially combined. It causes resolution of complex mixtures of proteins. This is a more sensitive analytical method than either electrophoretic method alone. Two-dimensional electrophoresis separates proteins of identical molecular weight. These proteins differ in pI, or proteins with similar pI values but different molecular weights.

Isoelectric points of Some Proteins	
Protein	**pI**
Pepsin	<1.0
Egg albumin	4.6
Serum albumin	4.9
Urease	5.0
β-Lactoglobulin	5.2
Hemoglobin	6.8
Myoglobin	7.0
Chymotrypsinogen	9.5
Cytochrome c	10.7
Lysozyme	11.0

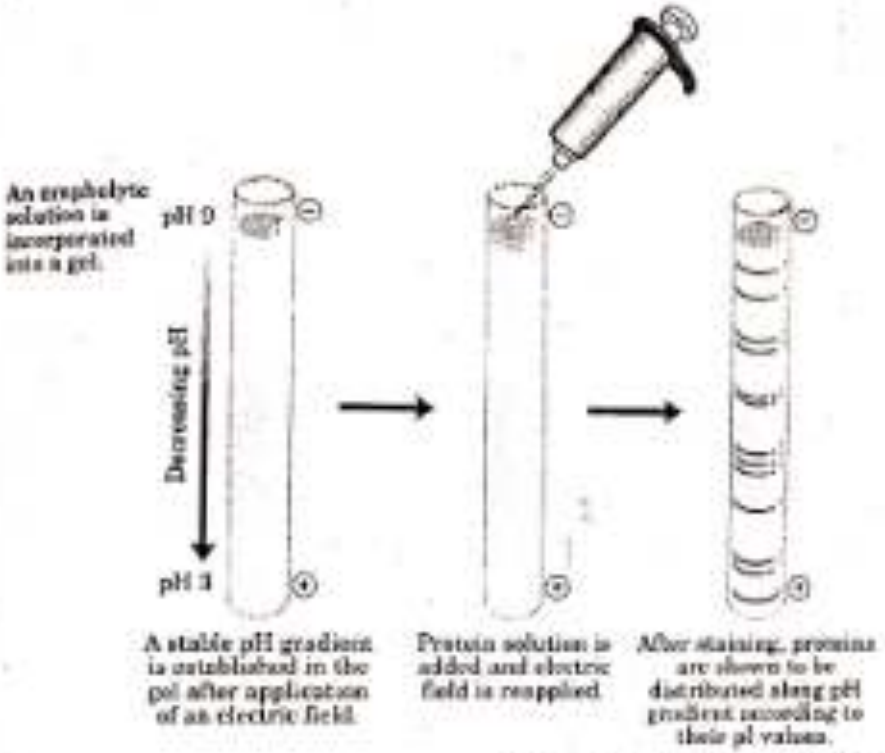


Fig: Isoelectric focusing

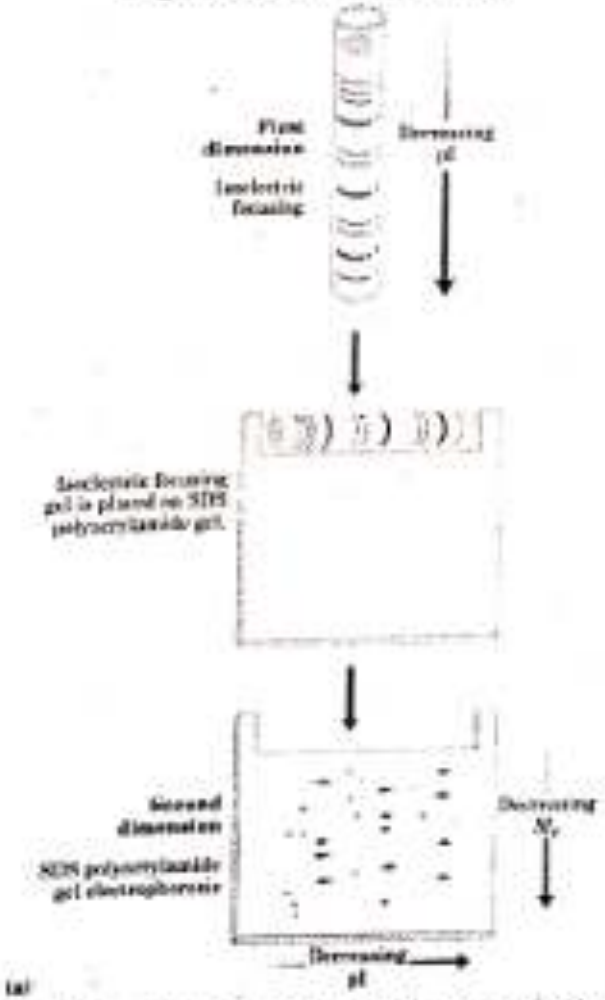


Fig: Two-dimensional electrophoresis

ALPHA-KERATIN

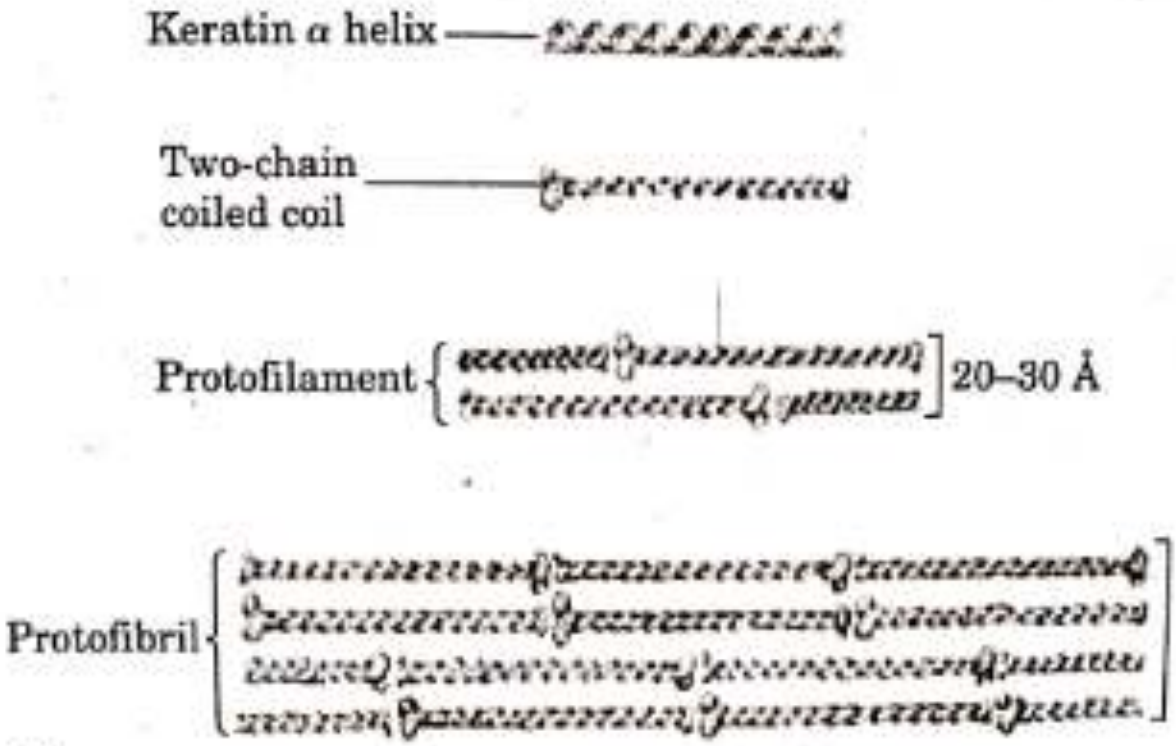
Function

The α-keratins are used to provide strength. It is found only in mammals. These proteins constitute almost the entire dry weight of hair, wool, nails, claws, quills, horns, hooves, and much of the outer layer of skin.

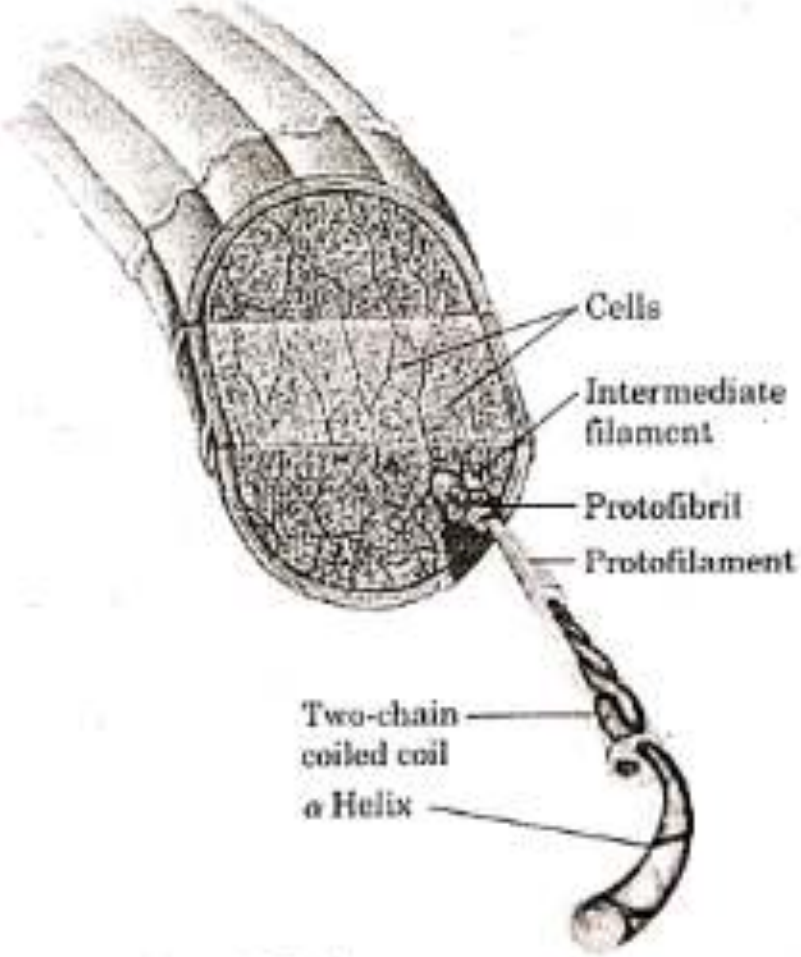
Structure

The α-keratins are part of a broader family of proteins called **intermediate filament (IF) proteins**. Other IF proteins are found in the cytoskeletons of animal cells. All IF proteins share the structural features of α-keratins.

1. **Supertwisting coiled coil:** The α-keratin helix is a right-handed helix. Francis Crick and Linus Pauling in the early 1950s suggested that helices of keratin were arranged as a coiled coil. Two strands of α-keratin are oriented in parallel (with their amino terminate the same end). They are wrapped about each other to form a supertwisted coiled coil. The supertwisting amplifies the strength of overall structure like rope.
2. **Distance per turn:** Pauling and Corey predicted that twisting of axis of α helix to form a coiled coil had 5.4Å per turn. But actually 5.15 to 5.2 Å repeating structure was observed in x-ray diffraction of hair.
3. **Formation of regular interlocking pattern:** The helical path of supertwists is left-handed. The surfaces where the two helices touch are made up of hydrophobic amino acid residues. The R groups of these amino acids are meshed together in a regular interlocking pattern. This permits a close packing of polypeptide chains within left-handed supertwist. The α -keratin is rich in hydrophobic residues Ala, Val, Leu, Ile, Met, and Phe.
4. **Tertiary structure:** An individual polypeptide in α-keratin coiled coil has a relatively simple tertiary structure. It is dominated by an α-helical secondary structure. Its helical axis twisted in a left-handed superhelix.
5. **Quaternary structures:** The intertwining of the two α-helical polypeptides is an example of quaternary structure. Coiled coils of this type are common structural elements. These are found in filamentous proteins and in the muscle protein myosin. The quaternary structure of α-keratin can be quite complex. Many coiled coils can be assembled into large supramolecular complexes. Such arrangement of α-keratin forms the intermediate filament of hair.
6. **Formation of covalent cross links:** The strength of fibrous proteins is enhanced by covalent cross-links. These cross-links are present between polypeptide chains in multihelical ropes and between adjacent chains in a supramolecular assembly. In α-keratins, cross-links stabilizing quaternary structure are disulfide bonds. Hardest and toughest α-keratins are found in rhinoceros horn. Its up to 18% of the residues are cysteines involved in disulfide bonds.



(a)



(b) Cross section of a hair

Fig: Structure of hair

COLLAGEN

Function

Collagen provides strength. It is found in connective tissue like tendons, cartilage, organic matrix of bone, and cornea of the eye.

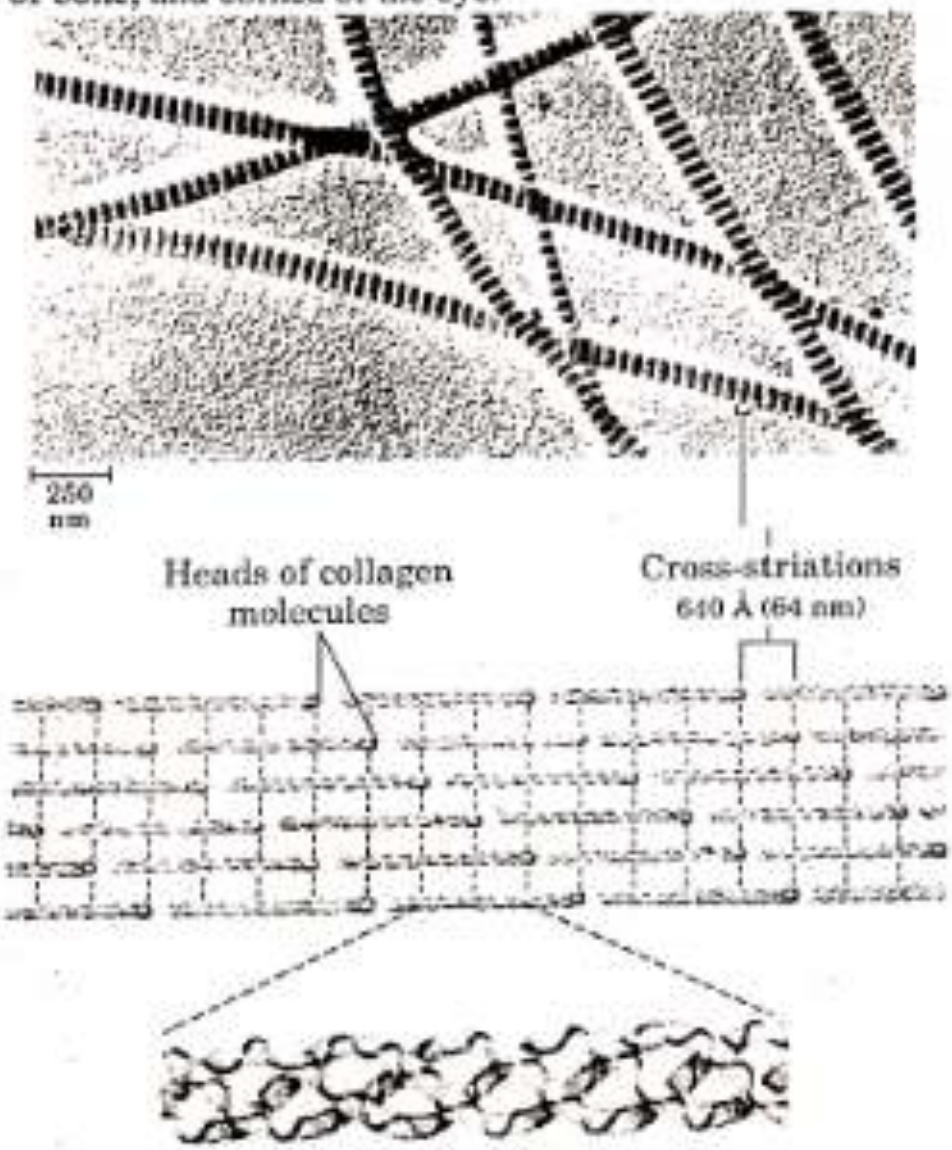


Fig: Structure of collagen fibrils

Structure

1. The collagen helix is a unique secondary structure. It has coiled coil. But it has distinct tertiary and quaternary structures. It has three separate polypeptides chains (not helixes). These chains are supertwisted about each other. The superhelical twisting is right-handed in collagen. It is opposite in sense to lefthanded helix of the chains.
2. There are many types of vertebrate collagen. Typically they contain about 35% Gly, 11% Ala, and 21% Pro and 4-Hyp (4-hydroxyproline, an uncommon amino acid).

3. The food product gelatin is derived from collagen. It has little nutritional value as a protein. The collagen is extremely low in many amino acids that are essential in the human diet. The unusual amino acid content of collagen is related to structural constraints unique to collagen helix.
4. The amino acid sequence in collagen is a repeating tripeptide unit, Gly–X–Y. Here X is often Pro, and Y is often 4-Hyp. Only Gly residues can be accommodated at the very tight junctions between the individual chains. The Pro and 4-Hyp residues permit sharp twisting of collagen helix. The amino acid sequence and supertwisted quaternary structure of collagen allow a very close packing of its three polypeptides. 4-Hydroxyproline has a special role in the structure of collagen.
5. There is tight wrapping of the chains in the collagen triple helix. It provides tensile strength greater than that of a steel wire of equal cross section. Collagen fibrils are supramolecular assemblies consisting of triple-helical collagen molecules. These are sometimes called **tropocollagen molecules**. These molecules are associated in a variety of ways to provide different degrees of tensile strength.

HEMOGLOBIN

Hemoglobin is iron-containing oxygen-transport metalloprotein. It is found in the red blood cells (erythrocytes) of almost all vertebrates and tissues of some invertebrates. Haemoglobin in blood carries oxygen from the lungs or gills to the rest of the body (i.e. the tissues). There it releases the oxygen to permit aerobic respiration. It provides energy to power the functions of the organism in metabolism. A healthy individual has 12 to 20 grams of haemoglobin in every 100 ml of blood.

In mammals, the protein makes up about 96% of the red blood cells' dry content (by weight), and around 35% of the total content (including water). Haemoglobin has an oxygen-binding capacity of 1.34 mL O_2 per gram. It increases the total blood oxygen capacity seventy-fold compared to dissolved oxygen in blood. The mammalian hemoglobin molecule can bind (carry) up to four oxygen molecules.

Hemoglobin is involved in the transport of other gases: It carries some of the body's respiratory carbon dioxide (about 20–25% of the total) as carbaminohemoglobin. Here CO_2 is bound to the heme protein. The molecule also carries the important regulatory molecule nitric oxide bound to a globin protein thiol group, releasing it at the same time as oxygen.

Structure

Hemoglobin was first oligomeric protein whose three dimensional structure was determined. It contains four polypeptide chains and four heme prosthetic groups. These chains have iron atoms are in the ferrous (Fe^{2+}) state. The protein portion is **globin**. It consists of two α chains (141 residues each) and two β chains (146 residues each). Hemoglobin is four times as large as myoglobin. Therefore, much more time and effort were required to solve its three-dimensional structure by x-ray analysis. It was finally achieved by Max Perutz, John Kendrew, and their colleagues in 1959. The subunits of hemoglobin are arranged in symmetric pairs. Each pair has one α and one β subunit.

Therefore hemoglobin can be described either as a tetramer or as a dimer of αβ protomers.

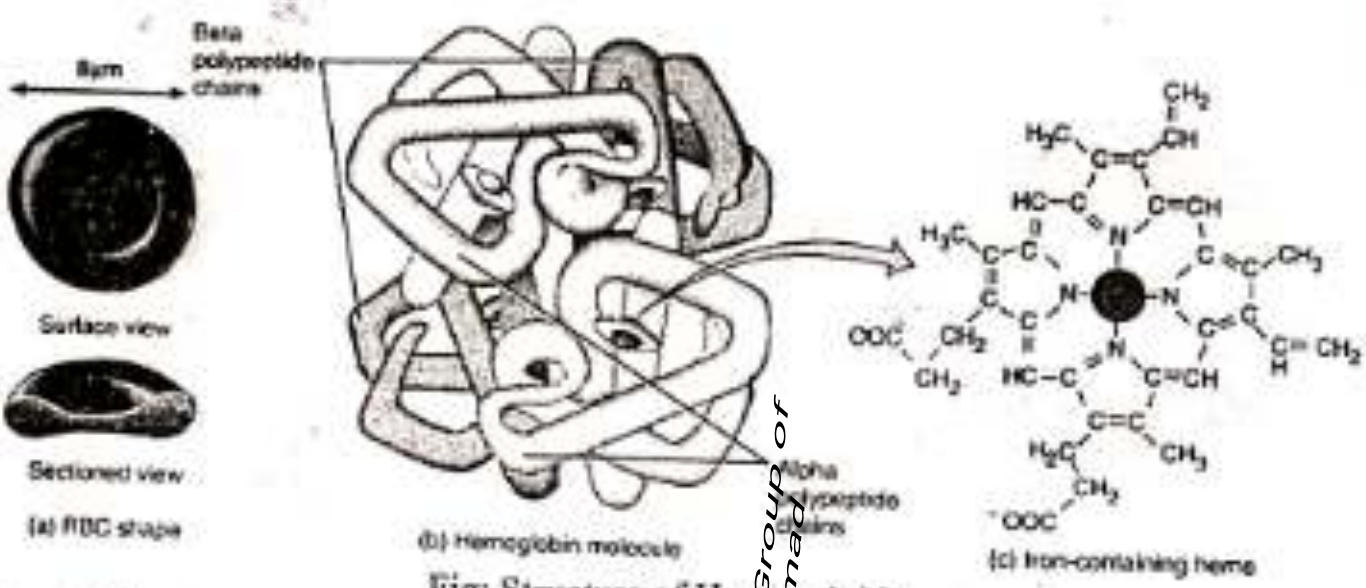


Fig: Structure of Haemoglobin

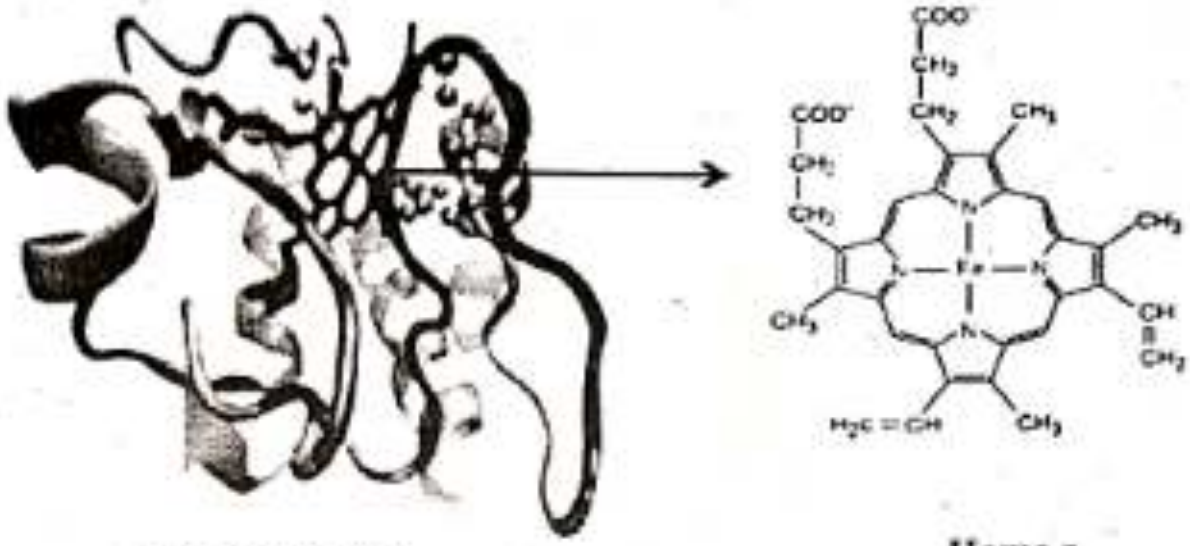
Hemoglobin has a quaternary structure characteristic of many multi-subunit globular proteins. Most of the amino acids in hemoglobin form alpha helices. These helices are connected by short non-helical segments. Hydrogen bonds stabilize the helical sections inside this protein, causing attractions within the molecule. Hydrogen bonding helps polypeptide chain to fold into a specific shape. Hemoglobin's quaternary structure comes from its four subunits in roughly a tetrahedral arrangement.

A heme group consists of an iron (Fe) ion (charged atom). This group is held in a heterocyclic ring called a **porphyrin**. The porphyrin ring consists of four pyrrole molecules. They are cyclically linked together (by methine bridges) with the iron ion bound in the center. The iron ion is the site of oxygen binding. It coordinates with the four nitrogen atoms in the center of the ring. These all rings lie in one plane. The iron is bound strongly (covalently) to the globular protein via N atoms of imidazole ring of F8 histidine residue (also known as the proximal histidine). A sixth position can reversibly bind oxygen by a coordinate covalent bond. It completes the octahedral group of six ligands. Oxygen binds in an end-on bent geometry where one oxygen atom binds to Fe and the other protrudes at an angle. When oxygen is not bound, a very weakly bonded water molecule fills the site. It forms a distorted octahedron. The iron ion may be either in the Fe^{2+} or in the Fe^{3+} state. In binding, oxygen temporarily and reversibly oxidizes (Fe^{2+}) to (Fe^{3+}). As a result, oxygen temporarily turns into the superoxide ion.

CYTOCHROME-C

The cytochrome complex or **cyt *c*** is a small hemeprotein. It is found loosely associated with the inner membrane of the mitochondrion. It belongs to the cytochrome c family of proteins. Cytochrome c is highly water-soluble. It is an essential component of the electron transport chain. It carries one electron. It is capable of undergoing oxidation and reduction. But it does not bind oxygen. It transfers electrons between Complexes III (Coenzyme Q – Cyt C reductase) and IV (Cyt C oxidase).

Future Zoologist Group of PGCG (By Muhammad ARSALAN)



Cytochrome *c* Heme *c*

Structure

Cytochrome c belongs to class I of c-type cytochrome family. It contains a characteristic CXXCH (cysteine-any-any-cysteine-histidine) amino acid motif that binds heme. This motif is located towards N-terminus of the peptide chain. It contains a histidine as the fifth ligand of the heme iron. The sixth ligand is provided by a methionine residue found towards the C-terminus. The protein backbone is folded into five α-helices. These are numbered as α1-α5 from N-terminus to C-terminus.

Heme

The heme group of cytochrome c makes thioether bonds with two cysteine side chains of the protein. One of the main properties of heme c is its ability to have different reduction potentials in nature. This property determines the kinetics and thermodynamics of an electron transfer reaction.

Function

1. Cytochrome c is a component of the electron transport chain in mitochondria. The heme group of cytochrome c accepts electrons from bc_1 complex. It then transfers electrons to the complex IV. Cytochrome c is also involved in initiation of apoptosis. Upon release of cytochrome c to cytoplasm, the protein binds apoptotic protease activating factor-1.
2. Cytochrome c can also catalyze several redox reactions like hydroxylation and aromatic oxidation.
3. It shows peroxidase activity by oxidation of various electron donors.
4. A bacterial cytochrome *c* functions as a nitrite reductase.

CHYMOTRYPSIN

Chymotrypsin is a digestive enzyme component of pancreatic juice. It acts in the duodenum. Here it performs proteolysis. The breakdown of proteins and polypeptides is called proteolysis. Chymotrypsin cleaves peptide amide bonds. The side chain of the amino acid C-terminal to the scissile amide bond (P_1 position) is a large hydrophobic amino acid (tyrosine, tryptophan, and phenylalanine). These amino

acids contain an aromatic ring in their side chain. This chain fits into a hydrophobic pocket of the enzyme. It is activated in the presence of trypsin. Chymotrypsin also hydrolyzes other amide bonds in peptides at slower rates. Structurally, it is the archetypal structure for its superfamily PA clan of proteases.

Fig: Structure of Chymotrypsin

PROPROTEINS

Pro-protein or pro-peptide is a protein precursor. It is an inactive protein (or peptide) that can be turned into an active form by post-translational modification like breaking off a piece of the molecule or adding on another molecule. The name of the precursor for a protein is often prefixed by *pro-*.

Examples

1. **Zymogens or proenzymes:** Protein precursors are often used by an organism when subsequent protein is potentially harmful. But it needs to be available on short notice in large quantities. Enzyme precursors are called **zymogens or proenzymes**. Examples are enzymes of the digestive tract in humans like pepsinogen, trypsinogen etc.
2. **Prohormones:** These include proinsulin and proopiomelanocortin. Both these are prohormones.
3. **Pre-pro-peptides** Some protein precursors are secreted from the cell. Many of these are synthesized with an N-terminal signal peptide that targets them for secretion. Like other proteins that contain a signal peptide, their name is prefixed by *pre*. They are thus called pre-pro-proteins or pre-pro-peptides. The signal peptide is cleaved off in the endoplasmic reticulum. Its example is preproinsulin.
4. **Pro-sequences:** These are areas in the protein that are essential for its correct folding. Pro-sequences may also be involved in pro-protein
5. **Procollagen:** The connective tissue protein collagen is initially synthesized as the soluble precursor procollagen.

6. **Blood clotting proteins, Fibrinogen and Prothrombin:** Blood clotting is mediated by a complicated cascade of proteolytic activations. The fibrin is protein of blood clots. It is produced by proteolysis of fibrinogen. Fibrinogen is inactive proprotein. The protease responsible for this activation is thrombin itself. It is produced by proteolysis of a proprotein called prothrombin.

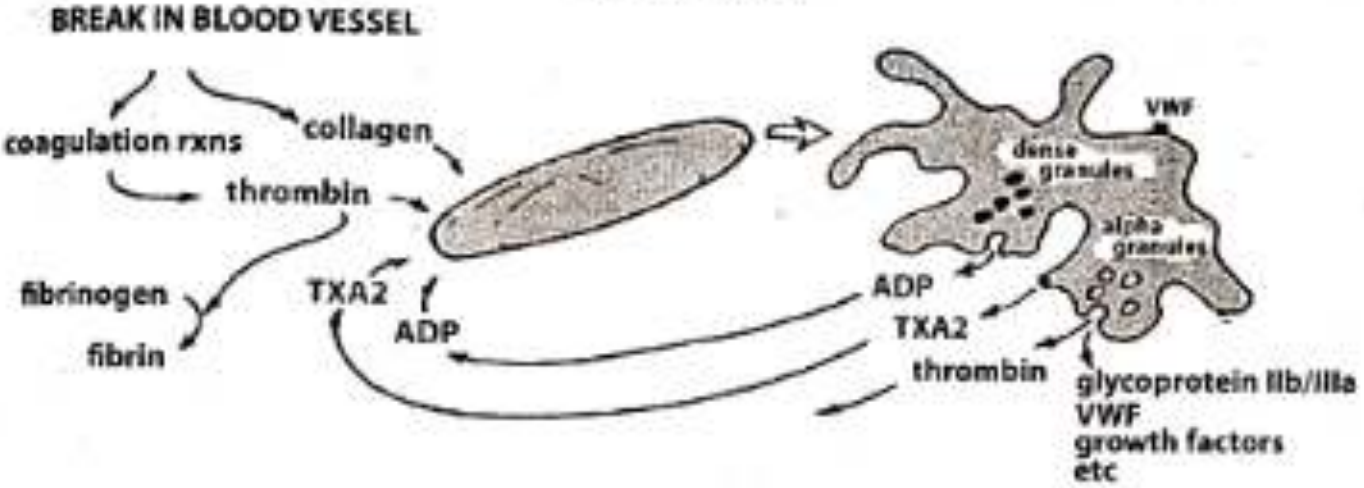


Fig: Mechanism of clotting of blood

Chapter 3

ENZYMES

Syllabus 2012-13
Chapter 2: Enzymes
Introduction; important characteristics of enzymes; immobilized enzymes; how enzymes work; example of enzymatic reaction; enzyme kinetics, enzyme rate of reaction and substrate concentration, how pH and temperature effect on enzyme activity.
Syllabus 2018-19
3. Enzymes a. Enzymes, their importance, classification & nomenclature, Function & **inhibition**.

INTROUDUCTION

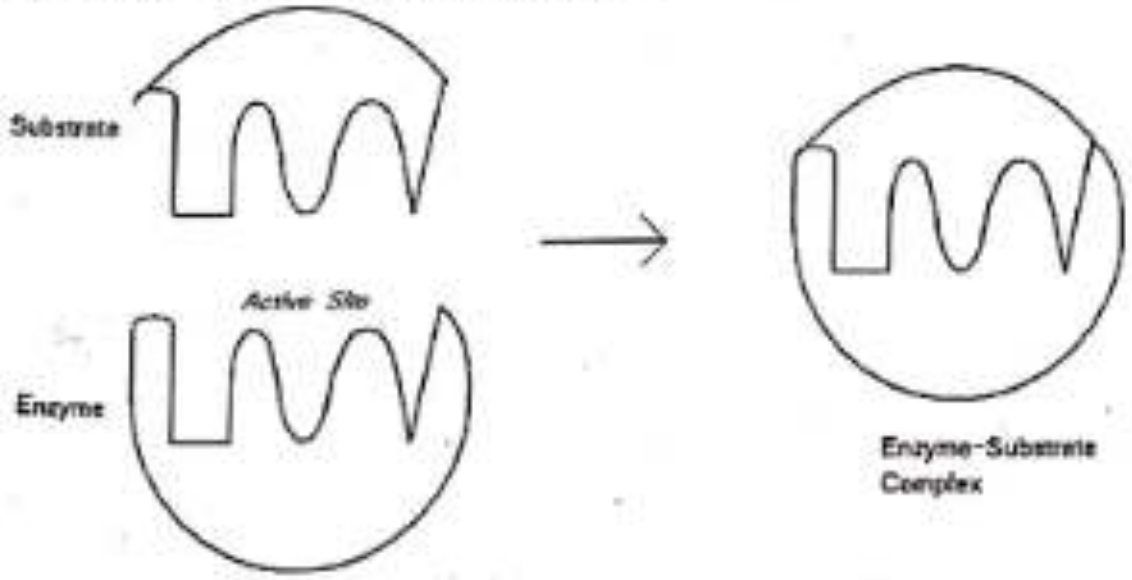
Enzymes are the most important biologically active group of proteins or sometimes RNA which increase the efficiency of biochemical reactions and are specific for each type of reaction. Enzymes increase the efficiency of the biological reactions. They are specific for each type of reaction. The reactions proceed at a very slow speed without enzymes. Thus without enzymes life is impossible.

History of Discovery of Enzyme

Much of the history of biochemistry is the history of enzyme research.

1. Enzyme was first recognized and described in the late 1700s. It was recognized in studies on the digestion of meat. This meat was digested by secretions of the stomach.
2. Its research continued in the 1800s. Biochemists studied conversion of starch to sugar by saliva and various plant extracts.

3. In the 1850s, Louis Pasteur concluded that fermentation of sugar into alcohol by yeast. It is catalyzed by **ferments**. He postulated that these ferments were inseparable from living yeast cells. This view was called **vitalism**.
4. Then in 1897, Eduard Buchner discovered that yeast extracts can ferment sugar to alcohol. It proved that fermentation was promoted by molecules. These molecules continued to function when removed from cells.
5. Frederick W. Kuhne called these molecules enzymes.
6. The vitalistic views of life were disproved. Then new enzymes were isolated. Their advanced properties were discovered. James Sumner in 1926 isolated and crystallized enzyme **urease**. It was a great breakthrough in early enzyme studies. Sumner found that urease crystals consisted entirely of protein. He postulated that all enzymes are proteins.
7. After that John Northrop and Moses Kunitz crystallized pepsin, trypsin, and other digestive enzymes. They found them also to be proteins.
8. During this period, J. B. S. Haldane wrote a paper entitled **Enzymes**. Haldane suggested that weak bonding interactions occur between an enzyme and its substrate. This interaction is used to catalyze a reaction. This concept is basis of current enzymatic catalysis.
9. In the latter part of the twentieth century, one or more metal ions were studied for activity. A coenzyme or metal ion that is very tightly or even covalently bound to the enzyme protein is called a **prosthetic group**.



Characteristic of Enzymes

Nature

1. **Protein:** All enzymes except ribozymes are globular proteins. They are made up of one or more polypeptides.
2. **Ribozymes:** Ribozymes are enzymes. They consist of RNA. They are found in ribosomes. For example, peptidyl transferase is a ribozyme. It controls polypeptide elongation during translation process.

Globular Structure and Active Site

Enzymes are composed of hundreds of amino acids. They join together and form a **globular structure.**

Active site

Catalytic activity of an enzyme is located in its active site. **Active site is a specific charge bearing three dimensional cavity.**

i) **Attachment with substrate: The reactant which will be converted into product is called substrate.** Substrate molecule is attached to active site. It is attached by non-covalent interactions. These interactions are like hydrogen bonding and hydrophobic interactions.

ii) **Arrangement of amino acids in active site:** Active site consists of 3-12 amino acids. These amino acids may be scattered in the polypeptide. But secondary and tertiary folding of the protein molecule occurs. These folding bring amino acids together. Thus they are arranged in a particular fashion. For example, active site for **aldolase.** It consists of glycine, histidine, and alanine amino acids.

iii) **Regions of active site:** An active site consists of two functional regions. These are binding site and catalytic site.

a) **Binding site:** Some amino acids have active site. It makes bonds with substrate. It is called binding site

b) **Catalytic site:** Other amino acids convert substrate into product (catalysis). They form catalytic site.

a) **Shape of active site:** Shape of active site is designed according to substrate. Therefore, only a particular substrate can attach to the active site. But sometime related substrate can also bind to the active site.

Co-factors

Some enzymes also require a non-protein part. It is called cofactor. Cofactors have two functions:

a) They attach substrate to active site

b) They also participate in catalytic process.

c) The attachment of cofactor gives final shape to active site.

There are two types of enzyme on the basis of presence or absence of cofactors:

a) **Holoenzyme:** Some enzyme requires a cofactor. They become active only on combing with cofactor. Such an active enzyme is called **holoenzyme.**

b) **Apoenzyme:** Sometimes, cofactor is not available. Now the remaining protein part of enzyme becomes catalytically inactive. This inactive enzyme is called **apoenzyme.**

4. Enzymes without cofactors

Some enzymes do not require cofactor. They can also show active and inactive states. Its example is pepsin. It has two forms:

a) **Pepsinogen (inactive form):** It is secreted by gastric gland from stomach wall. Here it in an inactive state called **pepsinogen.** In this state, it has an additional polypeptide

fragment. This fragment is attached to its active site. It does not allow binding of substrate. Hence binding site remains inactive.

b) **Pepsin: (active form):** The pepsinogen is exposed in HCl in stomach cavity. Now the additional polypeptide fragment is removed. Thus inactive (apoenzyme) pepsinogen is changed into its active (holoenzyme) pepsin.

Types of Cofactors

Cofactor may be inorganic or organic molecules. Inorganic cofactors are activators. organic cofactors are co-enzymes or prosthetic groups.

1. Activator

Activators are inorganic cofactors. These include different metallic ions like Fe^{2+}, Mg^{2+}, Cu^{2+}, etc. These attach to enzymes on binding of substrate. Thus they are detachable cofactors.

$$\text{Glucose} + \text{ATP} \xrightarrow{\text{Hexokinase}+Mg^{2+}} \text{Glucose-6-phosphate} + \text{ADP}$$

2. Coenzyme

Coenzymes are derivatives of vitamins. For example, ATP, NAD, FAD are common coenzymes. They also attached to enzymes on binding of substrate. Thus they are also detachable cofactors.

$$\text{Ethyl alcohol} \xrightarrow{\text{Alcohol dehydrogenase}+NAD} \text{Acetaldehyde} + NADH_2$$

3. Prosthetic group

Prosthetic group is covalently bonded part of an enzyme. It is permanently attached to enzyme. It does not detach after the completion of reaction. Its example is an iron containing porphyrin ring. It is attached to some enzymes like cytochromes.

Modification of enzyme activities

Some enzyme proteins are modified covalently by phosphorylation, glycosylation, and other processes. Many of these alterations are involved in the regulation of enzyme activity.

Properties of enzymes

Enzymes are organic catalysts. But they show many properties of inorganic catalyst. These properties are:

1. **Speed up reactions:** Enzymes increases speed of chemical reaction. Now reaction occurs 10^6-10^{14} times faster than uncatalyzed reaction.
2. **Need in very small amount:** A reaction requires a very small amount of enzymes. Very few enzymes can convert huge amount of substrate into products. They do it in very short time.
3. **Sensitivity:** Enzymes are highly sensitive to pH and temperature changes of the system.
4. **Specificity:**
   a) Some enzymes may be highly specific. For example sucrase can only digest sucrose.

b) Some enzyme may be slightly less specific. Such enzyme can catalyze the related reaction. For example carbonic anhydrase can add O_2 in haemoglobin. It can control the formation of carbonic acid and bicarbonates in blood.

5. **Working place:** Enzymes can work in vivo (living cells) and in vitro (glassware).
6. **Cofactor:** Some enzymes require co-factor for proper activity.
7. **Lowering activating energy:** Enzymes lower activation energy.
8. **No effect on equilibrium:** Enzymes only speed up a reaction. They do not affect the equilibrium of the reaction.

IMMOBLIZED ENZYMES

An **immobilized enzyme** is an enzyme attached to an inert, insoluble material like calcium alginate. It is produced by reacting a mixture of sodium alginate solution and enzyme solution with calcium chloride. This can provide increased resistance to changes in conditions like pH or temperature. It also holds enzymes in place throughout the reaction. After reaction, they are easily separated from the products. Thus these may be used again. Their reaction is more efficient process. Hence immobilized enzymes are used in industry for enzyme catalysed reactions.

Commercial Use

Immobilized enzymes are very important for commercial. They have benefits of having less expensive. They processes of the reaction of which include:

1. **Convenience:** Very small amounts of protein dissolve in the reaction. Hence work can be much easier. On completion, reaction mixtures typically contain only solvent and reaction products.
2. **Economy:** The immobilized enzyme is easily removed from the reaction. It makes easy to recycle the biocatalyst. This is particularly useful in processes like production of lactose free milk. In this reaction, milk can be drained from a container. It leaves enzyme (Lactase) inside ready for the next batch.
3. **Stability:** Immobilized enzymes typically have greater thermal and operational stability than the soluble form of the enzyme.

Immobilization of an Enzyme

There are different methods to immobilize an enzyme:

1. **Affinity-tag binding:** Enzymes may be immobilized to a surface like porous material. In this process, we use non-covalent or covalent protein tags. This technology is used for protein purification purposes. This technique can be used without prior enzyme purification with a pure preparation. Therefore, porous glass and derivatives are used. Here porous surface can be adapted for hydrophobicity to suit the enzyme.
2. **Adsorption on glass, alginate beads or matrix:** Enzyme is attached to the outside of an inert material. This method is the slowest in these processes. The adsorption is not a chemical reaction. Therefore, active site of the immobilized enzyme may be blocked by the matrix or bead. It greatly reduces the activity of the enzyme.

3. **Entrapment:** The enzyme is trapped in insoluble beads or microspheres like calcium alginate beads. But these insoluble substances hinder the arrival of the substrate, and the exit of products.
4. **Cross-linkage:** Enzyme molecules are covalently bonded to each other. It creates a matrix consisting of only enzyme. The reaction ensures that the binding site does not cover the enzyme's active site.
5. **Covalent bond:** The enzyme is bound covalentely to an insoluble support. These supports are like silica gel or macroporous polymer beads with epoxide groups. It provides the strongest enzyme/support interaction. Hence lowest protein leakage occurs during catalysis.

HOW ENZYME WORKS

MECHANISM OF ENZYME ACTION

Need of Enzyme

Enzymatic catalysis of reactions is essential for living systems.

1. Uncatalyzed reactions are slow under biological conditions. Most biological molecules are quite stable in the neutral-pH, mild- temperature, aqueous environment inside cells.
2. Many common reactions in biochemistry are unfavorable in the cellular environment. They form transition of unstable charged intermediates. These reactions need collision of two or more molecules in precise orientation.
3. Reactions are required to digest food, send nerve signals, or contract a muscle. These reactions do not occur at a useful rate without catalysis.

Role of Enzyme

An enzyme solves these problems.

1. Enzyme provides a specific environment. A given reaction can occur more rapidly this environment.
2. An enzyme-catalyzed reaction has important feature. This reaction takes place within a pocket on the enzyme. This pocket is called the **active site.** The molecule bound in the active site and on which enzyme acts is called the **substrate**. The surface of the active site is lined with amino acid residues. These residues have substituent groups. These groups bind the substrate and catalyze its chemical transformation.
3. Often, active site encloses a substrate. Here it is completely separated from solution.
4. The **enzyme-substrate complex** is central to the action of enzymes. The existence of this complex was first proposed by **Charles-Adolphe Wurtz** in 1880. It is also the starting point for mathematical treatments. It defines kinetic behavior of enzyme-catalyzed reactions and for theoretical descriptions of enzyme mechanisms.

Enzymes Affect Reaction Rates, Not Equilibria

A simple enzymatic reaction is written as:

$$E + S \rightleftharpoons ES \rightleftharpoons EP \rightleftharpoons E + P$$

Here E, S, and P represent the enzyme, substrate, and product. The ES and EP are transient complexes of the enzyme with substrate and with product.

There is an important difference between reaction equilibria and reaction rates. The function of a catalyst is to increase the rate of a reaction. Catalysts do not affect reaction equilibria.

Reaction coordinate diagram

Any reaction, such as S P, can be described by a **reaction coordinate diagram**. It is a picture of the energy changes during the reaction. Energy in biological systems is described in terms of free energy G. In the coordinate diagram, free energy of the system is plotted against the progress of the reaction (reaction coordinate).

Ground state

The starting point for forward or the reverse reaction is called **ground state. Ground state is the contribution of an average molecule (S or P) to the free energy of the system under a given set of conditions.**

Standard set of conditions

Chemists define a standard set of conditions to describe the free-energy changes for reactions. These are temperature 298 K; partial pressure of each gas 1 atm, or 101.3 kPa; concentration of each solute 1 M.

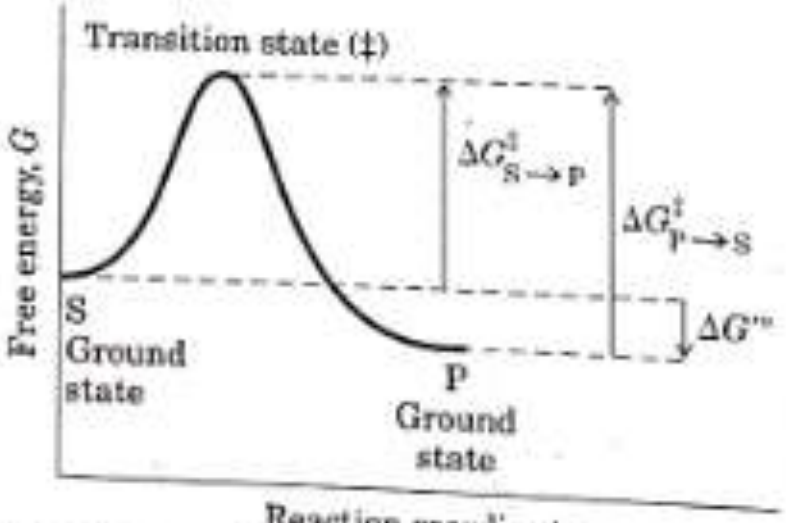


Fig: Reaction coordinate diagram for a chemical reaction

Standard free energy change

Chemists also express the free-energy change for this reacting system as ΔG^{o}. The ΔG^{o} is **standard free- energy change**. The biochemical systems commonly involve H^{+} concentrations far below 1 M. Therefore, biochemists define a **biochemical standard free-energy change**, $\Delta G'^{o}$. It is the standard free-energy change at pH 7.0.

Formation of equilibrium

The equilibrium between S and P shows difference in the free energies of their ground states. In the example in Figure, the free energy of the ground state of P is lower than that of S. Hence $\Delta G'^{o}$ for the reaction is negative. Thus equilibrium favors P. Any catalyst does not affect the position and direction of equilibrium.

Overcoming energy barrier

A favorable equilibrium does not mean that S→P conversion will occur at a detectable rate. The rate of a reaction is dependent on an entirely different parameter. There is an energy barrier between S and P. These are:

- Energy required for alignment of reacting groups,
- Formation of transient unstable charges
- Bond re-arrangements
- Other transformations required for the reaction to proceed in either direction.

This is shown by **energy "hill"** in Figures. The molecules must overcome this barrier to undergo reaction. Therefore, it must be raised to a higher energy level.

Transition state

At the top of the energy hill there is a point. The decay to S or P state is equally probable at this point. This is called **transition state**. Transition state is not a chemical species with any significant stability. It should not be confused with a reaction intermediate (such as ES or EP). It is simply a fleeting molecular moment. In this state, bond breakage, bond formation, and charge development have proceeded to the precise point. At this point decay to either substrate or product is equally likely.

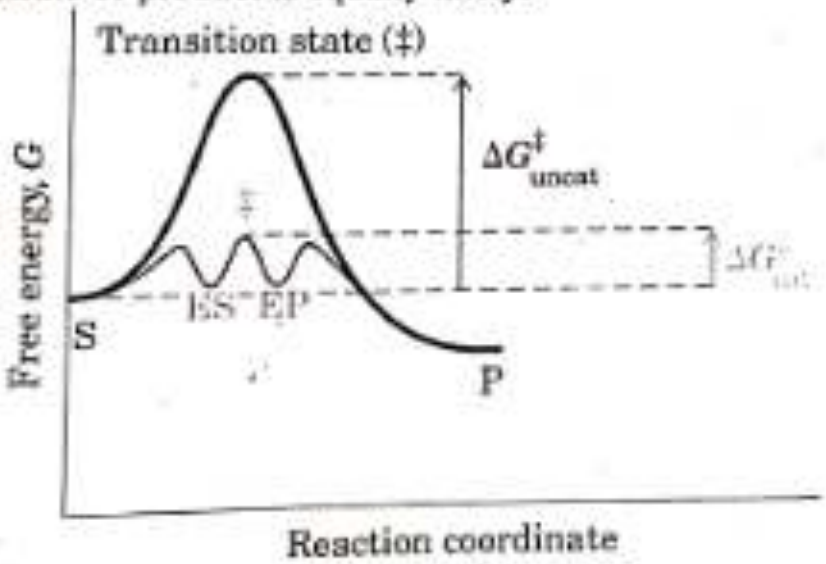


Fig: Reaction coordinate diagram comparing enzyme- catalyzed and uncatalyzed reactions

Activation energy

The difference between the energy levels of the ground state and transition state is the **activation energy** ($\Delta G^{\ddagger}$). The rate of a reaction shows this activation energy. A higher activation energy corresponds to a slower reaction. Reaction rates can be increased by raising the temperature. Hence it increases the number of molecules with sufficient energy. They overcome the energy barrier.

Role of catalyst

The activation energy can be lowered by adding a **catalyst**. Catalysts increases reaction rates by lowering activation energies. Enzymes follow the rule that catalysts do not affect reaction equilibria. Any enzyme that catalyzes the reaction S→P also catalyzes the reaction P→S. The role of enzymes is to accelerate the inter-conversion of S and P. The

enzyme is not used up in the process. Here equilibrium point is unaffected. However, the reaction reaches equilibrium much faster in the presence of enzyme because the rate of the reaction is increased.

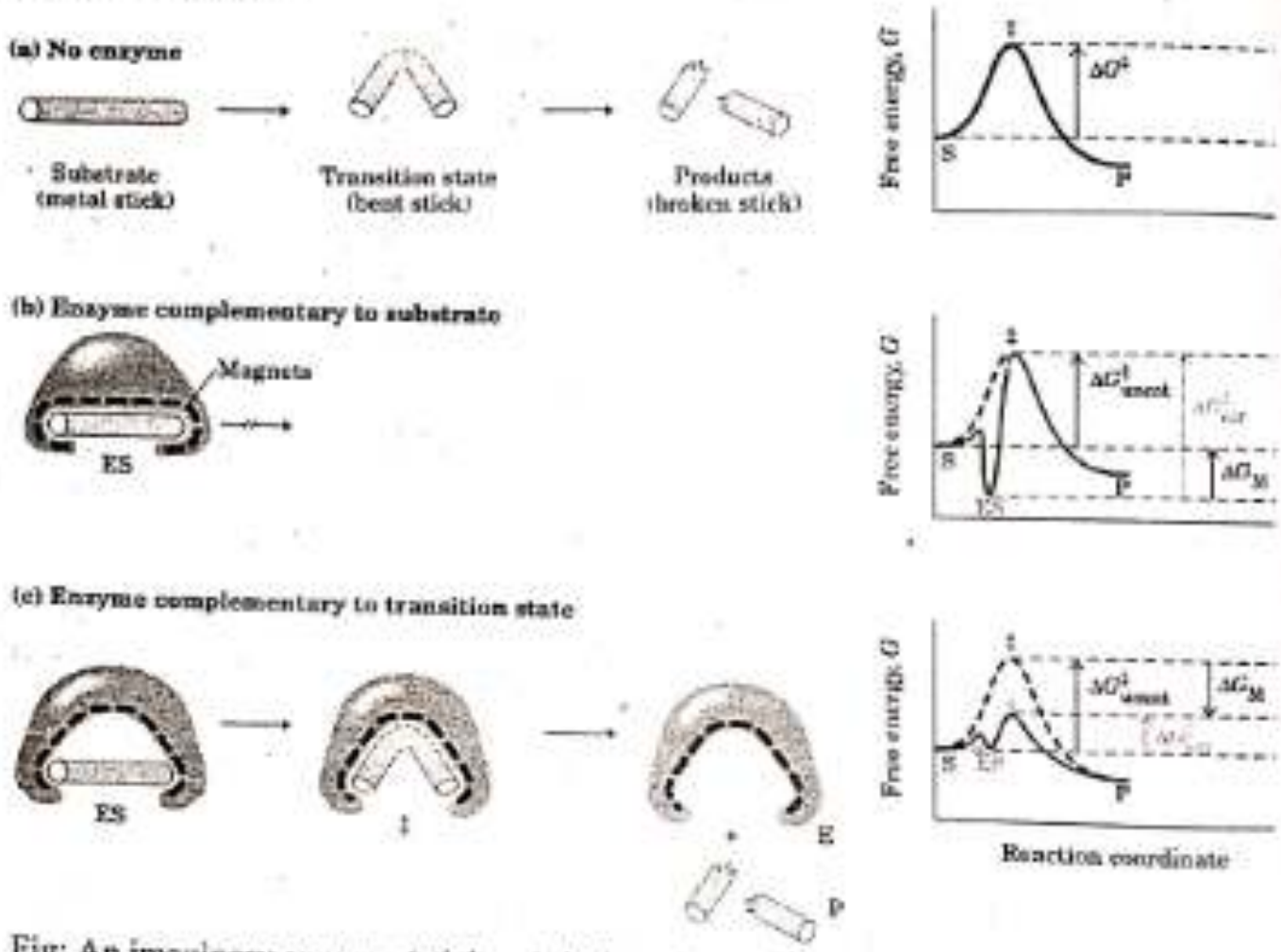


Fig: An imaginary enzyme (stickase) designed to catalyze breakage of a metal stick

Reaction intermediates

The steps of reaction involving formation and decay of transient chemical species are called **reaction intermediates**. A reaction intermediate is any species having a finite chemical lifetime (longer than a molecular vibration, ~10 13 seconds). An enzyme catalyzes a S↔P reaction. Here ES and EP complexes are intermediates. The ES and EP complexes occupy valleys in the reaction coordinate diagram. Some less stable chemical intermediates often exist in an enzyme-catalyzed reaction. Thus inter-conversion of two sequential reactions inter-mediates form a reaction step.

Rate determining step

In some reactions, several steps occur in a reaction. Here overall rate is determined by the step (or steps) with the highest activation energy. Such step is called **rate-limiting step**. Thus the rate-limiting step is the highest-energy point in the diagram for interconversion of S and P.

Activation energy barrier

Activation energies are energy barriers to chemical reactions. These barriers are crucial to life. The activation barrier for a reaction increases with decreases of reaction rate. Complex macro-molecules will revert spontaneously to much simpler molecular forms without such energy barriers. Thus complex and highly ordered structures and metabolic processes of cells will not exist.

Models of Enzyme Actions

There two different models to explain mechanism of enzyme action.

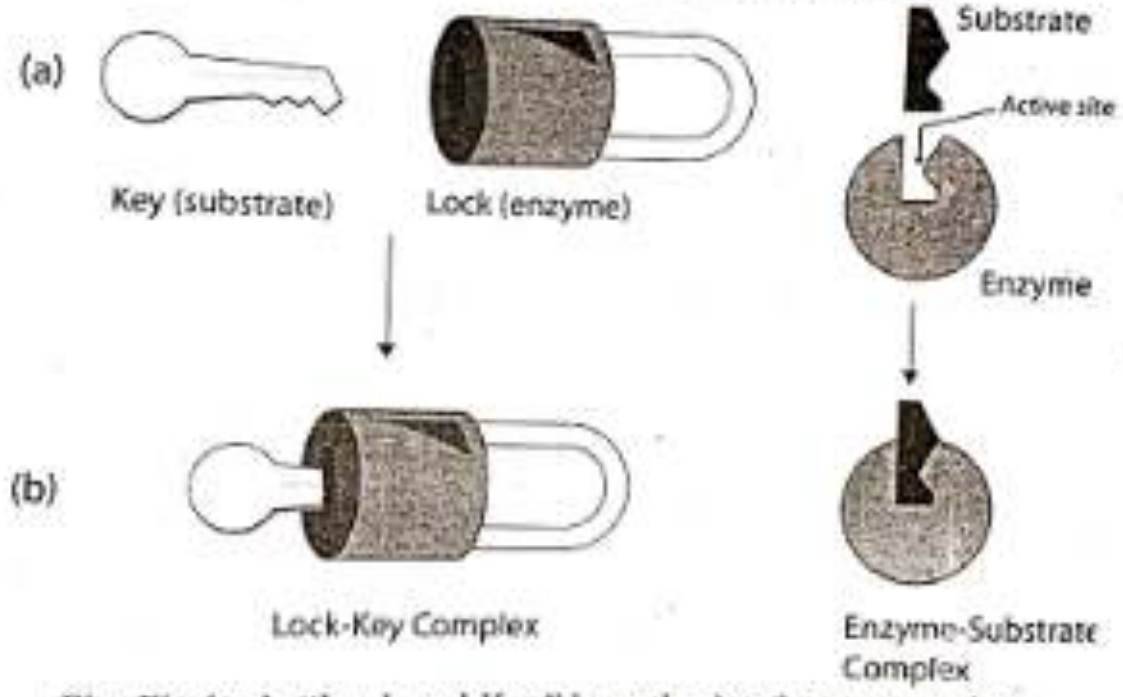


Fig: Fischer's "Lock and Key" hypothesis of enzyme action

1. Lock and Key model

Emil Fischer proposed Lock and key model in 1894. According to this model:

i) Active site of the enzyme has definite shape. It is rigid structure. Shape of active site is complementary to the shape of substrate.

ii) Therefore, a particular substrate can only bind to the active site.

iii) Active site remains unchanged during or after the reaction.

(a) Comparison of key and lock with enzyme and substrate

According to lock and key model, a particular key opens a particular lock. Similarly, specific enzyme (key) acts on a particular substrate (lock). Actually, key has notched portion. This portion is equivalent to active site of enzyme.

(b) Non-regulatory enzymes

The enzymes are highly specific in their action. Hence each enzyme can carry out only one particular reaction. **The enzymes working according to lock and key model are called non-regulatory enzymes.**

(c) Absolute specificity of enzyme

A few enzymes follow this model. Its examples are sucrase, maltase and etc. Such enzymes can catalyze one specific reaction. It is its most significant property of enzyme. But many enzymes show a broad range of specificity for substrate. **The specificity of enzyme in which an enzyme can catalyze only one substrate not others is called absolute specificity.** Its example is urease.

$$\text{Urea + Water} \xrightarrow{\text{Urease}} \text{Ammonia + Carbon dioxide}$$

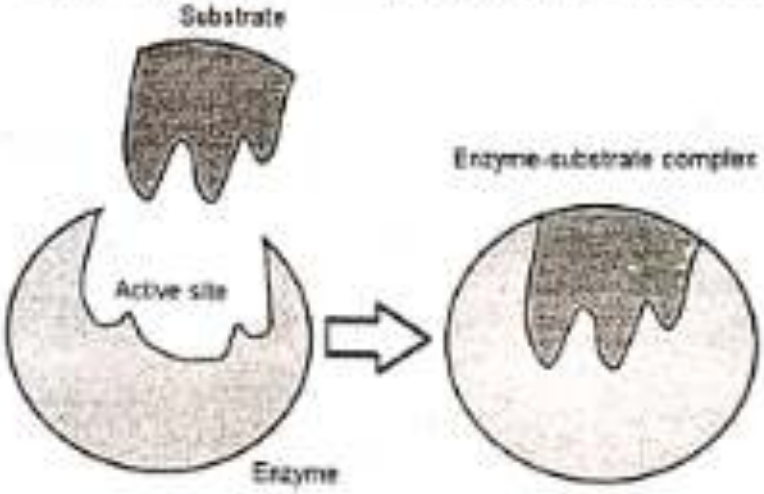


Fig: Koshland "Induced Fit" model of enzyme action

2. Induced fit model

Koshland proposed Induced fit model in 1959. **According to this model, active site is flexible. Therefore, it is modified as the substrate interacts with enzyme.**

a) Some amino acids form active site. These amino acids are molded into a precise shape. It enables the enzyme to catalyze reaction more effectively.
b) Change is induced in the shape of active site. This change converts substrate into product.
c) Finally the reaction is completed. Now the active site regains its original shape.
d) Hence there is the flexibility of active site. It allows more than one type of related substrates to attach on active site. Therefore, an enzyme can carry out more than one type of related reactions.

Example: Its example is carbonic anhydrase. It can add O_2 in haemoglobin. It can also control the formation of carbonic acid and bicarbonates in blood.

Regulatory or allosteric enzymes

The enzymes which follow induced fit mechanism are called regulatory or allosteric enzymes. Its example is hexokinase.

$$Hb + 4O_2 \xrightarrow{\text{Carbonic anhydrase}} Hb_4O_2 \text{ (Oxyhaemoglobin)}$$

$$CO_2 + H_2O \xrightarrow{\text{Carbonic anhydrase}} H_2CO_3 \xrightarrow{\text{Carbonic anhydrase}} H^+ + HCO^-$$

ENZYME KINETICS

Determination of rate of the reaction and how it changes in response to changes in experimental parameters is called enzyme kinetics. This is the oldest approach to understanding enzyme mechanisms. But it is remains the most important.

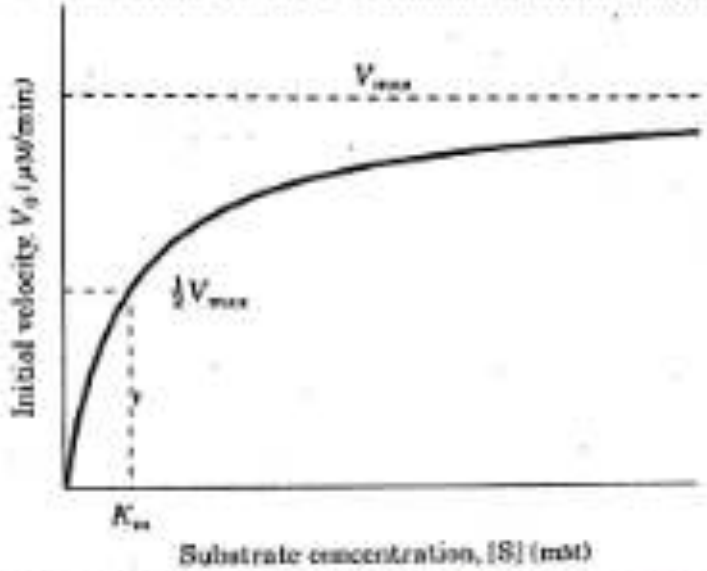


Fig: Effect of substrate concentration on the initial velocity of an enzyme-catalyzed reaction

Substrate Concentration Affects the Rate of Enzyme-Catalyzed Reactions

Concentration of substrate [S] is a key factor affecting the rate of a reaction. There is one simple approach in kinetics experiments. We measure initial rate (or initial velocity V_o) when [S] is much greater than concentration of enzyme, [E]. In a typical reaction, the enzyme may be present in nanomolar quantities. But [S] may be five or six orders of magnitude higher. We only monitor the beginning of the reaction (often the first 60 seconds or less). Here changes in [S] can be limited to a few percent. Thus [S] can be regarded as constant. The V_o can then be taken as a function of [S]. It is adjusted by the investigator.

a) At relatively low concentrations of substrate, V_o increases almost linearly with an increase in [S].
b) At higher substrate concentrations, V_o increases by smaller and smaller amounts in response to increases in [S].
c) Finally, a point is reached beyond which increases in V_o are much small as [S] increases. This plateau-like V_o region is close to **maximum velocity, V_{max}**.

The ES complex is the key to understanding this kinetic behavior. Victor Henri and Wurtz in 1903 proposed an enzyme combine with its substrate molecule to form an **ES complex.** Formation of this complex is a necessary step in enzymatic catalysis. Leonor Michaelis and Maud Menten in 1913 expanded this idea into a **general theory of enzyme action.** They postulated that the enzyme first combines reversibly with its substrate. They form an enzyme-substrate complex in a fast reversible step:

$$E + S \underset{K_{-1}}{\overset{K_1}{\rightleftharpoons}} ES \quad \text{.....} \quad (i)$$

This ES complex then breaks down in a slower second step. It yield the free enzyme and the reaction product

$$E + S \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E + P \ldots\ldots(ii)$$

The second reaction is slower. Therefore, it must limit the rate of the overall reaction. The overall rate must be proportional to concentration of species that reacts in the second step. This species is ES.

Saturation kinetics

At any given instant in an enzyme-catalyzed reaction, the enzyme exists in two forms. It may be in free or uncombined form E or in combined form ES.

1. At low [S], most of the enzyme is in the uncombined form E. Here, equilibrium of Equation (i) is pushed toward formation of more ES as [S] increases. Therefore, rate is proportional to [S].
2. At one stage, virtually all the enzyme is present as ES complex. Thus [E] is much small. At this stage, maximum initial rate of the catalyzed reaction (V_{max}) is observed. Under these conditions, the enzyme is **saturated** with its substrate. Therefore, further increases in [S] have no effect on rate. This condition exists when [S] is sufficiently high. Hence all the free enzyme has been converted to the ES form.
3. Then the ES complex breaks down to yield the product P. Now the enzyme is free to catalyze reaction of another molecule of substrate. The saturation effect is a distinguishing characteristic of enzymatic catalysts. It is responsible for the plateau observed in Figure. The pattern seen in Figure is called **saturation kinetics**.

Steady-state kinetics

1. **Pre–steady state:** There is an initial period when the enzyme is first mixed with a large excess of substrate.. This period is called **pre–steady state**. During this state, concentration of ES builds up. This period is usually too short to be easily observed. It last for just microseconds.
2. **Steady state:** The reaction quickly achieves a steady state. In this state, [ES] (and the concentrations of any other intermediates) remains constant over time. The concept of a steady state was introduced by **G. E. Briggs and Haldane** in 1925. Generally, V_o shows steady state. The analysis of these initial rates is called **steady-state kinetics**.

Michaelis Menten Equation

Quantitative Relationship between Substrate Concentration and Reaction Rate

The curve expressing the relationship between [S] and V_0 has the same general shape for most enzymes. It becomes a rectangular hyperbola). Thus it can be expressed algebraically by **Michaelis Menten equation**. Michaelis and Menten derived this equation starting from their basic hypothesis that rate-limiting step in enzymatic reactions is the breakdown of ES complex to product and free enzyme. The equation is:

$$V_0 = \frac{V_{max}[S]}{K_m + [S]} \quad \text{..... (1)}$$

The important terms are [S], V_0, V_{max} and a constant called Michaelis constant, K_m. All these terms are readily measured experimentally.

Derivation of Equation

The derivation starts with the two basic steps of formation and breakdown of ES. Early in the reaction, concentration of the product, [P] is negligible. Thus we make the simplifying assumption that reverse reaction, P → S (described by k_{-2}), can be ignored. The overall reaction then reduces to:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P \quad \text{..... (2)}$$

V_0 is determined by the breakdown of ES to form product. It is determined by [ES]:

$$V_0 = k_2[ES] \quad \text{..... (3)}$$

The [ES] in Equation (3) is not easily measured experimentally. Thus we must begin by finding an alternative expression for this term. First, we introduce the term $[E_t]$. It represents the total enzyme concentration (the sum of free and substrate-bound enzyme). Free or unbound enzyme can then be represented by:

$$[E_t] - [E_S]$$

The [S] is far greater than $[E_t]$. Therefore, the amount of substrate bound by the enzyme at any given time is negligible compared with total [S]. From these conditions, following steps lead from an expression for V_0 in terms of easily measurable parameters.

Step 1

The rates of formation and breakdown of ES are determined by steps governed by the rate constants k_1 (formation) and $k_1 - k_2$ (breakdown to reactants and products, respectively). According to the expressions:

$$\text{Rate of ES formation} = k_1([E_t] - [ES])[S] \quad \text{..... (4)}$$

$$\text{Rate of ES breakdown} = k_{-1}[ES] + k_2[ES] \quad \text{..... (5)}$$

Step 2

We now make an important assumption: that the initial rate of reaction reflects a steady state. Here [ES] is constant. Hence the rate of formation of ES is equal to the rate of its breakdown. This is called **steady-state assumption**. The expressions in Equations (4) and (5) can be equated for the steady state. It gives:

$$k_1([E_t] - [ES])[S] = k_{-1}[ES] + k_2[ES] \quad \text{..... (6)}$$

Step 3

In a series of algebraic steps, we now solve Equation (6) for [ES]. First, the left side is multiplied out and right side simplified to give:

$$k_1[E_t][S] - k_1[ES][S] = (k_{-1} + k_2)[ES \quad \text{..... (7)}$$

Adding the term $k_1[ES][S]$ to both sides of the equation and simplifying gives:

$$k_1[E_t][S] = (k_1[S] + k_{-1} + k_2)[ES] \quad \text{..... (8)}$$

We then solve this equation for [ES]:

$$[ES] = \frac{k_1[E_t][S]}{k_1[S] + k_{-1} + k_2} \quad \text{..... (9)}$$

This can now be simplified further. We combine the rate constants into one expression:

$$[ES] = \frac{[E_t][S]}{[S] + (k_{-1} + k_2)/k_1} \quad \text{..... (10)}$$

The term $(k_{-1} + k_2)/k_1$ is defined as **Michaelis constant (K_m)**. Substituting this into Equation (10) and simplifies the expression to:

$$[ES] = \frac{[E_t][S]}{K_m + [S]} \quad \text{..... (11)}$$

Step 4

We can now express V_0 in terms of [ES]. Substituting the right side of Equation (11) for [ES] in Equation (3). It gives

$$V_0 = \frac{k_2[E_t][S]}{K_m + [S]} \quad \text{..... (12)}$$

This equation can be further simplified. The maximum velocity occurs when the enzyme is saturated (that is, with $[ES] = [E_t]$) V_{max} can be defined as $k_2[E_t]$. Substituting this in Equation (12). It gives Equation (1)

$$V_0 = \frac{V_{max}[S]}{K_m + [S]}$$

This is the **Michaelis-Menten equation**. It is the rate equation for a one-substrate enzyme-catalyzed reaction. It is a statement of the quantitative relationship between initial velocity V_0, maximum velocity V_{max}, and initial substrate concentration [S]. These all are related through **Michaelis constant K_m**. K_m has units of concentration.

Confirmation and Application

We can confirm this equation by considering limiting situations where [S] is very high or very low. An important numerical relationship comes from Michaelis-Menten equation in the special case when V_0 is exactly one-half V_{max} Then:

$$\frac{V_{max}}{2} = \frac{V_{max}[S]}{K_m + [S]} \quad \text{..... (13)}$$

On dividing by V_{max}, we obtain

$$\frac{1}{2} = \frac{[S]}{K_m + [S]} \qquad \text{..... (14)}$$

Solving for K_m, we get Km + [S] = 2[S], or

$$K_m = [S], \text{ when } V_0 = \frac{1}{2} V_{max} \qquad \text{..... (15)}$$

This is a very useful, practical definition of K_m: K_m is equivalent to substrate concentration at which V_0 is one-half V_{max}.

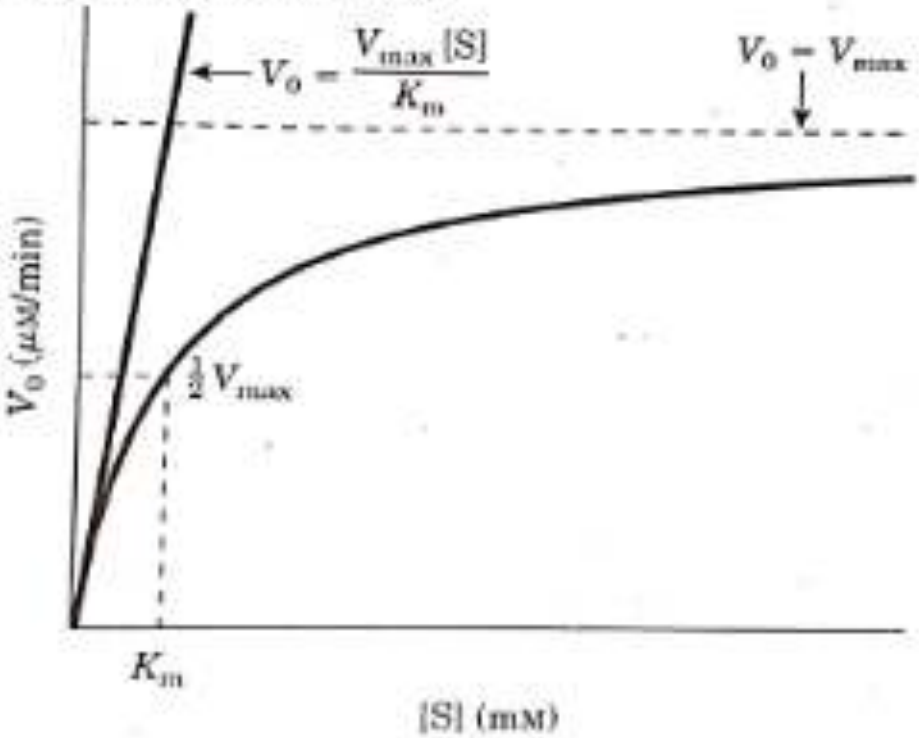


Fig: Dependence of initial velocity on substrate concentration

EXAMPLE OF ENZYMATIC REACTIONS

Hexokinase Undergoes Induced Fit on Substrate Binding

Yeast hexokinase is a bisubstrate enzyme. It catalyzes the reversible reaction ATP and ADP. It always binds to enzymes as a complex with the metal ion Mg^{2+}. There is a hydroxyl (OH) at C-6 of glucose. Phosphoryl of ATP is transferred to this carbon in the hexokinase reaction. This OH is similar in chemical reactivity to water. Hence water freely enters the enzyme active site. But hexokinase favors the reaction with glucose by a factor of 10^6. Conformational change occurs in the enzyme when the correct substrate binds. From this change enzyme can differentiate between glucose and water. Thus hexokinase is a good example of induced fit. When glucose is not present, the enzyme is in an inactive conformation with the active-site amino acid side chains out of position for reaction. When glucose (but not water) and Mg^{2+} ATP bind, the binding energy derived from this interaction induces a conformational change in hexokinase to the catalytically active form. This model has been reinforced by kinetic studies. The five-carbon sugar xylose, stereochemically similar to glucose but one carbon shorter, binds to hexokinase

but in a position where it cannot be phosphorylated. Nevertheless, addition of xylose to the reaction mixture increases the rate of ATP hydrolysis. Evidently, the binding of xylose is sufficient to induce a change in hexokinase to its active conformation, and the enzyme is thereby "tricked" into phosphorylating water. The hexokinase reaction also illustrates that enzyme specificity is not always a simple matter of binding one compound but not another. In the case of hexokinase, specificity is observed not in the formation of the ES complex but in the relative rates of subsequent catalytic steps. Water is not excluded from the active site, but reaction rates increase greatly in the presence of the functional phosphoryl group acceptor (glucose). Induced fit is only one aspect of the catalytic mechanism of hexokinase—like chymotrypsin, hexokinase uses several catalytic strategies. For example, the active site amino acid residues (those brought into position by the conformational change that follows substrate binding) participate in general acid-base catalysis and transition-state stabilization.

ENZYME INHIBITION

Enzyme inhibitors are molecules that interfere with catalysis. They slow or halt enzymatic reactions. Enzymes are subject to reversible or irreversible inhibition. Enzymes catalyze virtually all cellular processes. Therefore, enzyme inhibitors are among the most important pharmaceutical agents known. For example, aspirin (acetylsalicylate) inhibits the enzyme that catalyzes the first step in the synthesis of prostaglandins. These compounds are involved in many processes like producing of pain. The study of enzyme inhibitors also has provided valuable information about enzyme mechanisms. It has helped to define some metabolic pathways. There are two broad classes of enzyme inhibitors: reversible and irreversible.

(a) Reversible Inhibition

They form weak linkages with the enzymes. Their effect is completely or partly neutralized by increasing the concentration of the substrate. They are further divided into three major types:

(i) Competitive Inhibition

A competitive inhibitor competes with the substrate for the active site of an enzyme. The inhibitor (I) occupies the active site. It prevents binding of the substrate to the enzyme. Many competitive inhibitors are structurally similar to the substrate. They combine with the enzyme to form an EI complex. But it does not cause catalysis. The combinations of this type will reduce the efficiency of the enzyme. We take into account the molecular geometry of inhibitors. It tells about which parts of the normal substrate bind to the enzyme.

Competitive inhibition can be analyzed quantitatively by steady-state kinetics. In the presence of a competitive inhibitor, **Michaelis Menten equation** becomes

$$V_0 = \frac{V_{max}[S]}{\alpha K_m + [S]} \qquad \text{..... (i)}$$

Where

$$\alpha = 1 + \frac{[I]}{K_I} \quad \text{and} \quad K_I = \frac{[E][I]}{[EI]}$$

Equation (i) describes the important features of competitive inhibition. The experimentally determined variable is αK_m. The K_m is observed in the presence of the inhibitor. It is often called **apparent K_m**.

The inhibitor binds reversibly to the enzyme. Therefore, competition favor the substrate simply by adding more substrate.

1. When [S] far exceeds [I], the probability that an inhibitor molecule will bind to the enzyme is minimized. Therefore, reaction exhibits a normal V_{max}.
2. However, [S] at which $V_0 = \frac{1}{2} V_{max}$, the apparent K_m increases in the presence of inhibitor by the factor α. This effect on apparent K_m, combined with the absence of an effect on V_{max}, is diagnostic of competitive inhibition. It is readily revealed in a double-reciprocal plot. The equilibrium constant for inhibitor binding, K_I, can be obtained from the same plot.

(ii) Uncompetitive Inhibition

Two other type of reversible inhibition is uncompetitive. It is often defined in terms of one substrate enzymes. In practice, these are observed only with enzymes having two or more substrates. **An uncompetitive inhibitor binds at a site distinct from the substrate active site.** It binds only to the ES complex. In the presence of an uncompetitive inhibitor, the Michaelis-Menten equation is altered to:

$$V_0 = \frac{V_{max}[S]}{K_m + \alpha'[S]}$$

Where

$$\alpha' = 1 + \frac{[I]}{K_I'} \quad \text{and} \quad K_I' = \frac{[ES][I]}{[ESI]}$$

At high concentrations of substrate, V_0 approaches $V_{max/\alpha'}$. Thus, an uncompetitive inhibitor lowers the measured V_{max}. Apparent K_m also decreases, because [S] required to reach one-half V_{max} decreases by the factor α'.

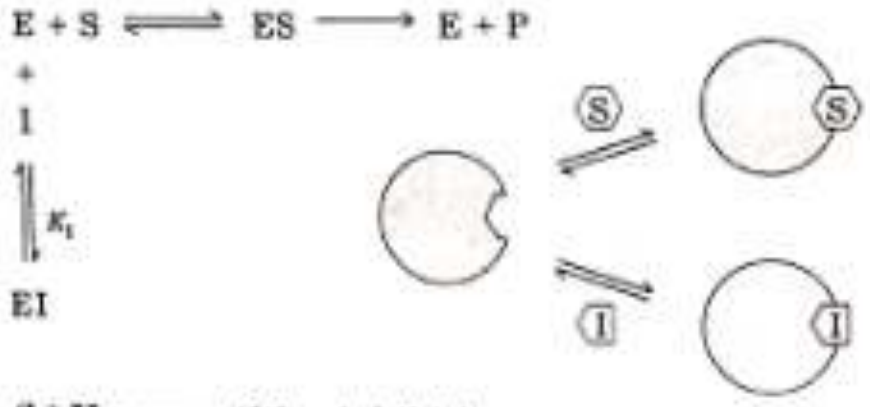
(iii) Mixed Reversible Inhibition

A mixed inhibitor also binds at a site distinct from the substrate active site. But it binds to either E or ES. The rate equation describing mixed inhibition is:

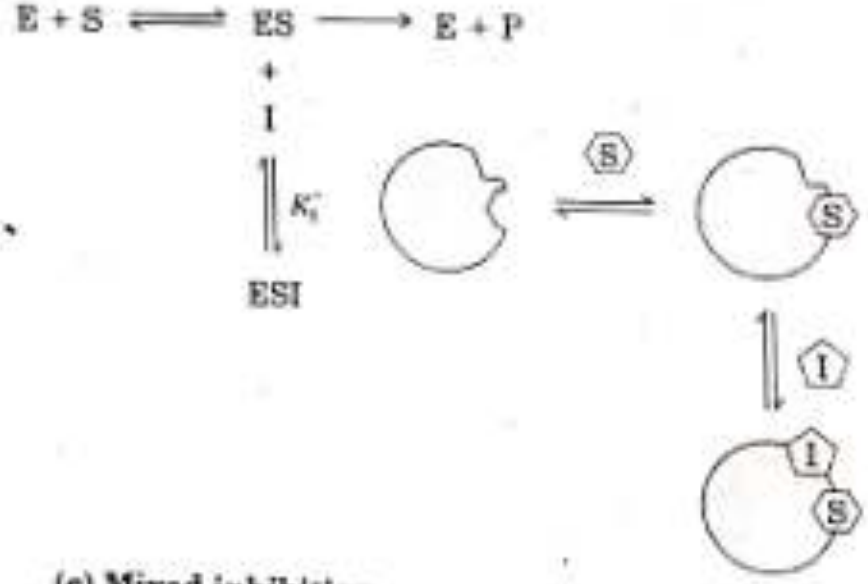
$$V_0 = \frac{V_{max}[S]}{\alpha K_m + \alpha'[S]} \qquad \text{..... (ii)}$$

A mixed inhibitor affects both K_m and V_{max}.

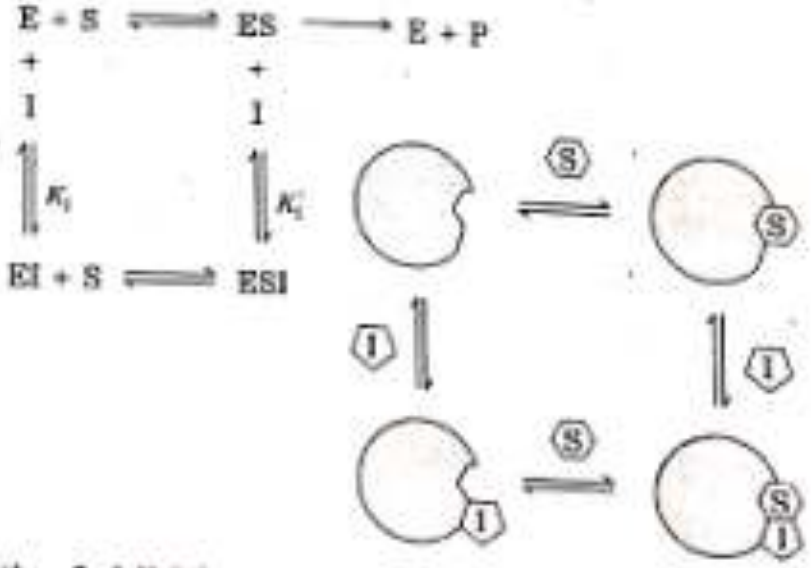
(a) Competitive inhibition



(b) Uncompetitive inhibition



(c) Mixed inhibition



(iv) Non-competitive Inhibition

The special case of α =α', rarely occur in experiments. It is classically defined as non-competitive inhibition.

Expression for competitive and uncompetitive inhibition

Equation (ii) is a general expression for the effects of reversible inhibitors. We simplify the expressions for competitive and uncompetitive inhibition when $\alpha' = 1.0$ or $\alpha = 1.0$, respectively. From this expression we can summarize the effects of inhibitors on individual kinetic parameters.

1. **For all reversible inhibitors**, apparent $V_{max} = V_{max}/\alpha'$. The right side of Equation (ii) always simplifies to $V_{max} = V_{max}/\alpha'$ at sufficiently high substrate concentrations.
2. **For competitive inhibitors**, $\alpha' = 1.0$. It can thus be ignored. Taking this expression for apparent V_{max}, we can also derive a general expression for apparent K_m to show how this parameter changes in the presence of reversible inhibitors. Apparent K_m is always equals [S] at which V_0 is one-half apparent V_{max} or, more generally, when $V_0 = V_{max}/2\alpha'$. This condition is met when $[S] = \alpha K_m/a'$. Thus, apparent $K_m = \alpha K_m/a'$. This expression is simpler when either α or α' is 1.0 (for uncompetitive or competitive inhibitors).
3. **For uncompetitive and mixed inhibition:** In practice, uncompetitive and mixed inhibition are observed only for enzymes with two or more substrates—say, S_1 and S_2. These are very important in the experimental analysis of such enzymes.
   a) An inhibitor may bind to the site normally occupied by S_1. Then it may act as a competitive inhibitor in experiments in which [S1] is varied.
   b) An inhibitor may bind to the site normally occupied by S2. Then it may act as a mixed or uncompetitive inhibitor of S_1.

The actual inhibition patterns observed depend on whether S_1- and S_2-binding events are ordered or random. The order in which substrates bind and products leave the active site can be determined. Use of one of the reaction products as an inhibitor is often particularly informative. If only one of two reaction products is present, no reverse reaction can take place. However, a product generally binds to some part of the active site. Thus it acts as an inhibitor.

Table: Effects of Reversible Inhibitors on Apparent V_{max} and Apparent K_m

Inhibitor type	Apparent V_{max}	Apparent K_m
None	V_{max}	K_m
Competitive	V_{max}	αK_m
Uncompetitive	V_{max}/α'	K_m/α'
Mixed	V_{max}/α'	$\alpha K_m/\alpha'$

(b) Irreversible Inhibition

The irreversible inhibitors bind covalently with or destroy a functional group on an enzyme that is essential for enzyme's activity, or form a particularly stable non-covalent association. Formation of a covalent link between an irreversible inhibitor and an enzyme is common.

Application

Irreversible inhibitors are another useful tool for studying reaction mechanisms. Some amino acids have key catalytic functions in the active site. They can be identified by determining which residue is covalently linked to an inhibitor after the enzyme is inactivated.

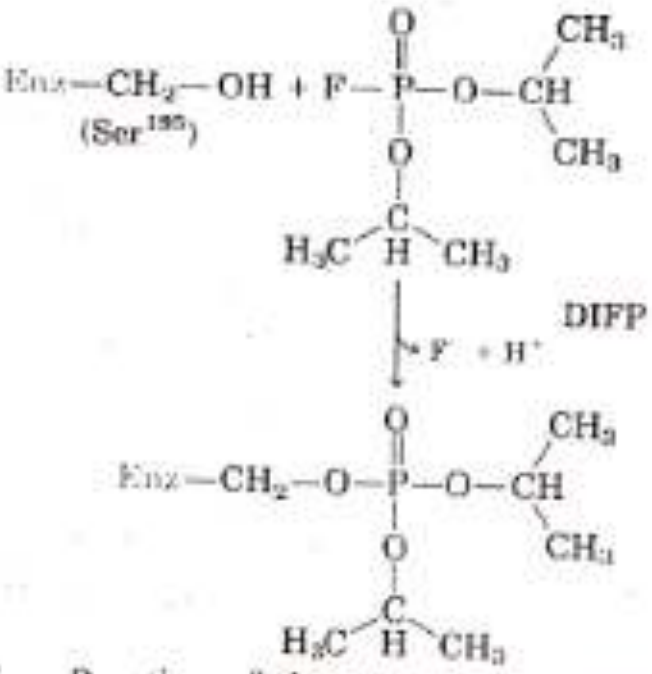


Fig: Irreversible inhibition. Reaction of chymotrypsin with diisopropylfluorophosphate (DIFP). It modifies Ser^{195}, irreversibly inhibits the enzyme. This has led to the conclusion that Ser^{195} is the key active-site Ser residue in chymotrypsin.

Suicide inactivators

A special class of irreversible inhibitors is **suicide inactivators**.

1. These compounds are relatively unreactive until they bind to the active site of a specific enzyme.
2. A suicide inactivator undergoes the first few chemical steps of normal enzymatic reaction. But it is not transformed into the normal product. Instead, the inactivator is converted to a very reactive compound.
3. The reactive compound combines irreversibly with the enzyme. These compounds are also called **mechanism-based inactivators**. They hijack the normal enzyme reaction mechanism to inactivate the enzyme.

Application

Suicide inactivators play a significant role in rational drug design. A well-designed suicide inactivator is specific for a single enzyme. It is unreactive until it is within that enzyme's active site. Therefore, drugs based on this approach can offer the important advantage of few side effects.

EXERCISE

LONG QUESTIONS

1. **Write down the characteristics of enzymes.**
2. **What are cofactors? Give their role in functioning of enzymes.**
3. **Write down the properties of enzymes.**
4. **What is meant by immobilized enzyme? Write its commercial applications.**
5. **Describe in detail mechanism of Enzyme action.**
6. **Discuss different models of enzyme action.**
7. **Illustrate enzyme kinetics along with its types and functioning.**

SHORT QUESTIONS

1. What is an enzyme or define an enzyme? Give its role or importance in cell.

Ans: Enzymes are the most important biologically active group of proteins or sometimes RNA which increase the efficiency of biochemical reactions and are specific for each type of reaction. Enzymes increase the efficiency of the biological reactions. They are specific for each type of reaction. The reactions proceed at a very slow speed without enzymes. Thus without enzymes life is impossible.

2. What are Ribozymes?

Ans: Ribozymes are enzymes. They consist of RNA. They are found in ribosomes. For example, peptidyl transferase is a ribozyme. It controls polypeptide elongation during translation process.

3. What is active site of an enzyme? What are its two regions?

Ans: Active site: The catalytic activity is restricted to small portion of the enzyme called active site. Substrates are attached to the active site. Active site is composed of a few amino acids. Rest of the bulk of amino acids maintains globular structure of the enzymes.

Regions of active site: The active site of the enzymes is made up of two regions

b) **Binding site:** It helps the enzyme in recognition and binding of proper substrate and form ES **(enzyme substrate)** complex. This reaction activates the catalytic site.

c) **Catalytic site:** The activated catalytic site changes the substrate into products.

4. What are cofactors? Give their functions.

Ans: Cofactors: Some enzymes contain a non-protein part, called co-factor. The types of cofactors are activator, prosthetic group and coenzymes.

Functions of cofactors

- They are essential for proper functioning of enzymes.
- The co-factors act as a bridge between the enzyme and substrate.
- They often directly take part in the chemical reactions during catalysis.
- Sometimes, the co-factors provide source of chemical energy. This energy is necessary to drive a reaction.

5. Why are cofactors essential for the functioning of an enzyme?

Ans: The co-factors act as a bridge between the enzyme and substrate. The co-factors often directly take part in the chemical reactions. Sometimes, the co-factors provide source of chemical energy.

6. What is meant by immobilized enzyme? Give its applications.

Ans: An **immobilized enzyme** is an enzyme attached to an inert, insoluble material like calcium alginate. It is produced by reacting a mixture of sodium alginate solution and enzyme solution with calcium chloride. This can provide increased resistance to changes in conditions like pH or temperature. It also holds enzymes in place throughout the reaction. After reaction, they are easily separated from the products. Thus these may be used again. Their reaction is more efficient process. Hence immobilized enzymes are used in industry for enzyme catalysed reactions.

7. What is the difference between a Prosthetic group and Coenzyme?

Ans:

Prosthetic group	Coenzyme
1. The covalently bonded non-protein part cofactor is known prosthetic group.	1. The loosely attached non-protein part to the protein part is known as coenzyme.
2. Mostly these are attached permanently with enzyme.	2. These are attached loosely and are removed after reaction.
3. They composed are non-protein parts.	3. Coenzymes are related to vitamins.

4. What is the difference between Apoenzyme and Holoenzyme?

Ans:

Apoenzyme	Holoenzyme
1. An enzyme, with coenzyme or prosthetic group removed is called apoenzyme.	1. An activated enzyme, consisting of polypeptide chain and cofactor is called holoenzyme.
2. Apoenzyme remains inactive.	2. It is active form of enzyme.
3. Addition of concentrated coenzyme to the apoenzyme restores its activity.	3. Removal of cofactor block its activity.

5. Why are the enzymes so specific in action?

Ans: An enzyme is a three dimensional globular protein. Each enzyme has a specific chemical composition. It has specific amino acids and has specific shape. Due to this specificity, each enzyme recognizes and reacts with a specific substrate. It then changes this substrate into product. In the end, the enzyme releases unaltered.

6. What is the function of catalytic site of the enzyme?

Ans: The activated catalytic site changes the substrate into products.

7. Write three characteristics of enzymes.

Ans: Characteristics of enzymes

1. All enzymes are globular proteins.
2. They increase the rate of reactions. But they are not used in the reactions.
3. Their presence does not affect the nature or properties of end products.

8. What is Lock and key model? Who did propose it?

Ans: Lock and key model: According to this model, one specific key can open only a specific lock. Similarly, a specific enzyme can transform only specific substrate into products. According to this model, the active site is a rigid structure. There is no

modification or flexibility in the active site during and after enzyme action. It is used only as a template. Later studies did not support this model in all the reactions.

Propose: Emil Fischer proposed the lock and key model in 1890.

9. What is induced fit model? Who did propose it?

Ans: This model was proposed Koshland in 1959. He describes that when a substrate combines with an enzyme, it induces changes in the enzyme structure. These changes enable the enzyme to perform its catalytic activity.

9. Differentiate between Lock and key model and Induced fit models. Or compare Lock and key model and Induced fit models.

Ans:

Lock and key model	Induced fit models
1. According to this model, one specific key can open only a specific lock. Similarly, a specific enzyme can transform only specific substrate into products.	1. According this model a substrate combines with an enzyme, it induces changes in the enzyme structure. These changes enable the enzyme to perform its catalytic activity.
2. According to this model, the active site is a rigid structure.	2. According to this model, active site is not rigid structure.
3. There is no modification or flexibility in the active site during and after enzyme action.	3. Active site is flexible and substrate induces changes in it.
4. Active site is used only as a template.	4. It is not used as template

11. What is optimum temperature for enzymes? Give optimum temperature of human body.

Ans: The specific temperature at which an enzyme can work at maximum rate is called optimum temperature. For example, 37^0 C is the optimum temperature for the enzymes of the human body.

12. What is ground state?

13. The starting point for forward or the reverse reaction is called **ground state. Ground state is the contribution of an average molecule (S or P) to the free energy of the system under a given set of conditions.**

14. Write effect of enzyme concentration on enzyme action.

Ans: Enzyme Concentration

a) The rate of reaction depends directly on the enzyme concentration. At this specific time, the substrate concentration should be unlimited. The rate of reaction becomes double by doubling the amount of the enzyme.

b) The increase of number of enzyme molecules increases the number of active sites. Therefore, more active sites convert more substrates into products.

c) After sometime, the amount of substrate becomes limited. Now the reaction rate will no longer depend upon this increase.

15. What happens with the enzyme at high temperature?

Ans: Heat energy increases the vibrations of atoms of the enzymes molecules. Further increase in vibration can destroy the globular structure of the enzymes.

16. What is optimum pH? How pH effect enzyme action?

Ans: Optimum pH: Every enzyme has an optimum pH value. An optimum pH is a narrow range of pH in which an enzyme functions effectively.

Effect of pH on enzyme action: A slight change in pH can change the ionization of the active site of enzyme. It may affect the ionization of the substrate. In these conditions, activity of enzyme is retard blocked completely. Extreme changes in pH break the bonds of the enzyme. It causes enzyme denaturation

17. What are enzyme inhibitors? Give examples.

Ans: Inhibitors: Inhibitor is a chemical substance which can react (in place of substrate) with enzyme but is not transferred into products and thus blocks the active site of the enzyme temporarily or permanently.

Examples: For example, poisons like, cyanide, antibodies, anti-metabolites and some drugs

16. What is energy of activation? What is its relation with enzyme?

Ans: The minimum amount of energy required to start a reaction is called activation energy. Enzyme decreases the activation enzyme. So reaction takes place in less time.

17. Do you agree with the Lock and key model of Emil Fischer?

Ans: To some extant I agree with the model of Emil Fisher, but it is too limited to applied to all the reactions. Some substance called allostaric substance can change the enzyme (key).

18. Can you develop a relation between Lock and Key Model of Emil Fisher and Induced fit model of Koshland?

Ans: Actually, Induce fit model is a modified Lock and key model. It has decreased the intensity of rigidity of Emil Fisher's model.

19. Can reaction rate continue to increase with the increase of substrate concentration?

Ans: Yes, reaction rate go on increasing with the increase of substrate up to a limit. Further increase in the amount of substrate does not increase the rate of the reaction. The substrate molecules occupy all the active sites of the enzymes at the high concentration of substrate.

20. What is optimum pH? Why can pepsin not work in intestine?

Ans: An optimum pH is a narrow range of pH in which an enzyme functions effectively. Pepsin works in pH of 2. This pH is present only in stomach, while the pH of intestine is 7. So pepsin can not work in intestine.

21. What is Standard free energy change?

Ans: Chemists also express the free-energy change for this reacting system as ΔG^{o}. The ΔG^{o} is **standard free- energy change**. The biochemical systems commonly involve H^{+} concentrations far below 1 M. Therefore, biochemists define a **biochemical standard free-energy change**, $\Delta G'^{o}$. It is the standard free-energy change at pH 7.0.

MULTIPLE CHOICE QUESTIONS (MCQs)

1. Reaction can take place without enzyme but

(a) reaction will be more fast
(c) at the same speed as that with enzyme
(c) speed the reaction is much slow.
(d) reaction cannot take place without enzyme

2. Active site:

(a) is the part of globular part of enzyme.
(b) is separate part of the enzyme.
(c) is within the globular part of enzyme.
(d) none of the above

3. Cofactor is a non- protein part of enzyme:

(a) It is present in all enzyme
(b) It is present in some enzymes
(c) it is not necessary for the functioning of most of the enzyme
(d) None of the above

4. Which of the followings is a activator?

(a) It can be detached from enzyme.
(b) It is covalently bonded with the enzyme.
(c) It is loosely attached with the enzyme.
(d) none of above.

5. Which of the followings is a prosthetic group?

(a) detach after the completion of reaction
(b) covalently bonded with the enzyme.
(c) loosely attached with the enzyme.
(c) none of above.

6. Which of the followings is a coenzyme?

(a) detach after the completion of reaction
(b) covalently bonded with the enzyme.
(c) loosely attached with the enzyme.
(c) none of above.

7. Vitamins are essential for the body because:

(a) They are part of some enzymes
(b) Prosthetic groups of enzymes are synthesized from them.
(c) Coenzyme of enzymes is synthesized from them.
(d) Some activators have direct link with them.

8. In which form enzyme cannot function?

(a) In holoenzyme form (b) In apoenzyme form
(c) In activated form (d) non of the above

9. Activation energy is the amount of energy required to:

(a) activate the enzyme (b) start a reaction
(c) speed up a reaction (d) none of the above.

10. Enzyme always react with a specific substrate because of:
(a) specific shape and specific amino acids
(b specific prosthetic group
(c) capability to form specific bonds.
(d) None of the above

11. The binding site of the enzyme is that site which:
(a) recognize the substrate (b) bind with the cofactors
(c) change substrate into product. (d) none of the above

12. The catalytic site of the enzyme is that site which:
(a) recognize the substrate (b) bind with the cofactors
(c) change substrate into product. (d) none of the above

13. Lock and key model says that:
(a) The active site of an enzyme can be changed by key
(b) Only a specific key can open a specific lock
(c) Any key can open any lock
(d) None of the above.

14. Induce fit model says that:
(a) The active site of an enzyme can be changed by key
(b) Only a specific key can open a specific lock
(c) Any key can open any lock
(d) None of the above.

15. The reaction rate is decreased at high concentration of substrate because:
(a) Substrate make the active site as inactive.
(b) Substrate destroys the enzyme
(c) All enzymes are used in the reaction.
(d) Substrate occupies all the active site.

16. Heat accelerate the chemical reaction because:
(a) It increases the activation energy.
(b) It activates the active sites of enzyme.
(c) It increase the size of enzyme
(d) None of above.

17. High amount of heat checked the reaction rate because:
(a) It increases the movement of molecules of the enzymes
(b) It increase the vibratory movement of the molecules the enzymes
(c) It makes the enzyme inactive and destroying their active sites.
(d) None of the above.

18. A change in pH affect the enzyme reaction rate by:
(a) Ionizing the active site of the enzyme.
(b) Decreasing the movement of the molecules of the enzymes

(c) Destroying the globular structure of the enzyme
(d) None of the above

19. An irreversible inhibitor inhibit the activity of enzyme by:
(a) Occupying the active site
(b) Destroying the substrate
(c) Changing the structure by attaching.
(d) None of the above.

20. An reversible inhibitor inhibit the activity of the enzyme by:
(a) Occupying the active site
(b) Destroying the substrate
(c) Changing the structure by attaching.
(d) None of the above.

21. Which one of the following scientists believes that an enzyme structure cannot be changed?
(a) Palad (b) Fisher (c) Koshland (d) Sanger

Ans:
1. (c) 2. (b) 3. (b) 4. (a) 5. (b) 6. (c) 7. (c) 8. (b) 9. (b) 10. (a) 11. (a) 12. (c) 13. (b) 14. (a) 15. (d) 16. (b) 17. (b) 18. (a) 19. (a) 20 (c) 21. (b)

Chapter 4

CARBOHYDRATES

Syllabus 2012-13
Carbohydrates:
Classification, types, important characteristics and structure of carbohydrates; cyclic structure of monosaccharides; cyanohydrins formation; disaccharides their types structure and function; polysaccharides, storage and structural types; structure and major functions of polysaccharides.
Syllabus 2018-19
1. Introduction to Macromolecules a. Structure, types and role of various building blocks their respective macromolecules. b. Carbohydrates: Introduction; Classification Stereoisomerism in carbohydrate, Structure, types and role of monosaccharides, oligosaccharides and c. polysaccharides; **Glycosaminoglycans and glycoconjugates; d. Carbohydrates as an information carrier molecule.**

CARBOHYDRATES

Old definition of carbohydrate
Carbohydrates are compounds of carbon, hydrogen and oxygen. Literally word carbohydrate means **hydrates of carbon**. It means a carbon associated with water. **Any compound composed of carbon, hydrogen and oxygen and in which hydrogen and oxygen have same ratio as in water (2:1) is called carbohydrate.** This definition is applied to many carbohydrates. These are like glucose ($C_6H_{12}O_6$), sucrose ($C_{12}H_{22}O_{11}$) and starches and cellulose $(C_6H_{10}O_5)_n$.

Problem with old definition

a) Some compounds contain hydrogen and oxygen. It is in the same ratio as in water (2:1). These are not actually carbohydrates. Its example is acetic acid ($C_2H_4O_2$) and lactic acid ($C_3H_6O_3$).

b) We have some carbohydrates like rhamnose $(C_6H_{12}O_5)_n$. They do not contain hydrogen and oxygen in the ratio as in water.

Therefore, classical concept of carbohydrates as hydrates of carbon is incorrect.

More valid definition of carbohydrates

Now we have more valid definition of carbohydrates:

Carbohydrates are polyhydroxy aldehydes or ketones, or substances that yield such compounds on hydrolysis. Many, but not all, carbohydrates have empirical formula $(CH_2O)_n$. Some carbohydrates also contain nitrogen, phosphorus, or sulfur.

Importance or Functions of Carbohydrates

Carbohydrates occur abundantly in living organisms. They have following importance:

1. They are found in all organisms. They are present in all parts of the cells.
2. They form different structures, like cellulose of wood, cotton, and papers.
3. They act as storage compounds like **starch** and **glycogen**. Starch is present in cereals and root tubers (potatoes). Carbohydrates are also present in cane sugar and milk sugar.
4. Carbohydrates play both structural and functional role.
5. Simple carbohydrates are the main source of energy in cell.

Some carbohydrates form cell wall in plants and microorganisms.

CLASSIFICATION OF CARBOHYDRATES

More familiar carbohydrates have sweet taste. Therefore, carbohydrates are commonly called **sugars**. In Greek, **sakcharon** word is used for sugars. Hence term **saccharide** is derived from this word. Therefore carbohydrates are also called **saccharides**.

- A polyhydroxy aldehyde or polyhydroxy ketone is a single saccharide unit.
- But its derivative carbohydrates are composed of more than one saccharide units.

Classification of carbohydrates is based on number of saccharide units. Carbohydrates are classified into three groups. These are monosaccharides, oligosaccharides and polysaccharides.

1. **Monosaccharides or simple sugars:** They consist of a single polyhydroxy aldehyde or ketone unit. The most abundant monosaccharide in nature is the six-carbon sugar D-glucose. Sometimes it is called dextrose. Monosaccharides of four or more carbons tend to have cyclic structures.
2. **Oligosaccharides:** They consist of short chains of monosaccharide units, or residues, joined by characteristic linkages called glycosidic bonds. The most abundant are the disaccharides. They have two monosaccharide units. Typical is sucrose (cane sugar) consists of six-carbon sugars D-glucose and D-fructose. All common monosaccharides and disaccharides have names ending with suffix "-ose." In cells, most oligosaccharides consisting of three or more units do not occur as free entities. They are joined to non-sugar molecules (lipids or proteins) in glycoconjugates.
3. **Polysaccharides:** They are sugar polymers containing more than 20 or so monosaccharide units. Some have hundreds or thousands of units. Some polysaccharides like cellulose are linear chains. But others like glycogen are branched. Both glycogen and cellulose consist of repeating units of D-glucose. But they differ in the type of glycosidic linkage. Hence they have strikingly different properties and biological roles.

Table: Comparison of characteristics of carbohydrates

Monosaccharides	Oligosaccharides	Polysaccharides
They have single saccharide unit.	They have 2 to 10 saccharide units.	They have more than 10 saccharide units.
They are simplest carbohydrates Hence they cannot be further hydrolyzed.	They have less complex structure. On hydrolysis, they yield at least 2 and maximum 10 monosaccharides.	They have highly complex structure. On hydrolysis they yield at least 11 monosaccharides.
They are highly soluble in water.	They are less soluble in water.	They are insoluble in water
They are sweetest carbohydrates	They are less sweet in taste.	They are tasteless

MONOSACCHARIDES

i) Monosaccharides are true carbohydrates. They may be polyhydroxy aldehydes or polyhydroxy ketones.

ii) They contain 3 to 7carbons in monosaccharide.

iii) Except one, all carbon atoms in a monosaccharide have a hydroxyl group (–OH). The remaining carbon atom has aldehyde or ketone group.

iv) General formula of monosaccharides is $C_nH_{2n}O_n$. Here *n* is the number of carbon atoms in monosaccharides.

v) Many of the carbon atoms to which hydroxyl groups are attached are chiral centers. These centers give rise to the many sugar stereoisomers found in nature.

Classification of monosaccharides

They are classified on the basis of functional group and number of carbon atoms.

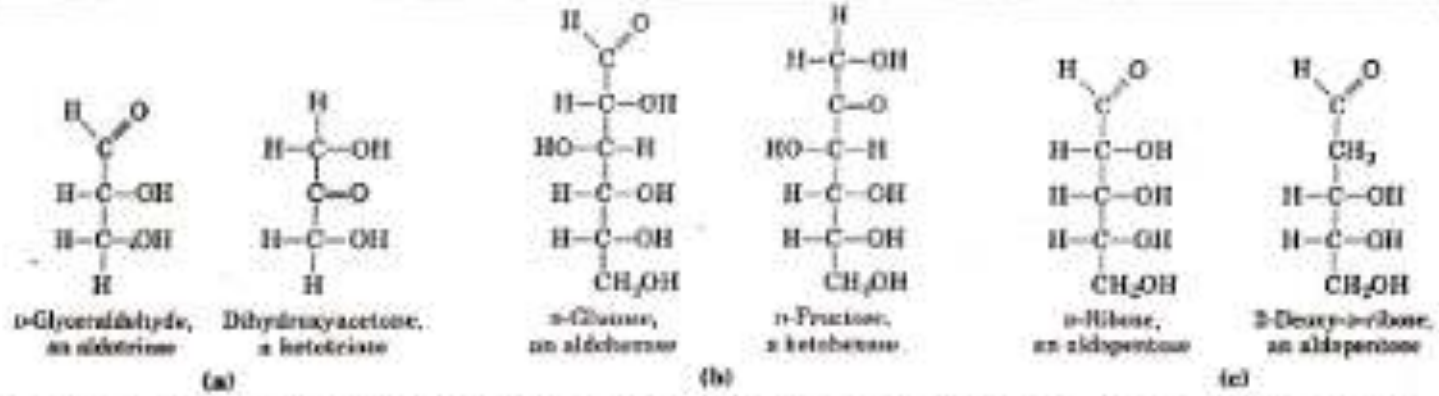
a) **Functional group basis:**
- Monosaccharides containing aldehyde are called aldoses.
- Monosaccharides containing ketone are called ketoses.

b) **Number of carbon basis:** On this basis, they are classified into five groups. These are trioses (3C), tetroses (4C), pentoses (5C), hexoses (6C) and heptoses (7C).

Table: Examples and functions of monosaccharide

Class	Formula	Aldoses	Ketoses	Functions
Trioses (3C)	$C_3H_6O_3$	Glyceraldehyde	Dihydroxy acetone	Intermediates in photosynthesis and cellular respiration..
Tetroses (4C)	$C_4H_8O_4$	Erythrose	Erythrulose	Intermediate in bacterial photosynthesis.
Pentoses (5C)	$C_5H_{10}O_5$	Ribose	Ribulose	• Ribose and deoxyribose are components of RNA and DNA.

				• Ribulose is an intermediate in photosynthesis.
Hexoses (6C)	$C_6H_{12}O_6$	Glucose Glactose	Fructose	• Glucose is respiratory fuel (initial substrate) • Fructose is an intermediate in respiration. • Galactose is the component of milk sugar.
Heptoses (7C)	$C_7H_{14}O_7$	Glucoheptose	Sedoheptulose	Intermediates in photosynthesis.



Representative monosaccharides. (a)Two trioses an aldose and a ketose. The carbonyl group in each is shaded. (b) Two common hexoses. (c) The pentose components of nucleic acids. D-Ribose is a component of ribonucleic acid (RNA), and 2-deoxy-D-ribose is a component of deoxyribonucleic acid (DNA).

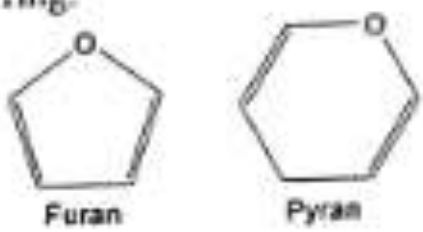
Chemical Structure of Monosaccharides

a) Monosaccharides have open chain structure in crystalline form.

b) Most of pentoses and hexoses are dissolved in water. Now they are converted into ring chain structure. Two types of ring structures are found in monosaccharides. These are furanose and pyranose.

   i) **Furanose:** Furanose is a five membered ring. This ring has one oxygen atom and four carbon atoms. Here oxygen atom is linked with C_1 and C_4. All pentoses and ketohexoses are converted into furanose ring.

   ii) **Pyranose:** Pyranose is a six membered ring. This ring has one oxygen atom and five carbon atoms. Oxygen atom is linked with C_1 and C_5. Only aldohexoses are converted into pyranose ring.



Two Families of Monosaccharides

Two families of monosaccharides are aldoses and ketoses. Monosaccharides are colorless, crystalline solids. They are freely soluble in water but insoluble in nonpolar solvents. Most have a sweet taste.

1. The backbones of common monosaccharides are unbranched carbon chains. All the carbon atoms are linked by single bonds to this chain. In open-chain form, one of the carbon atoms is double-bonded to an oxygen atom to form a carbonyl group. Each of the other carbon atoms has a hydroxyl group.
2. If the carbonyl group is at an end of the carbon chain (aldehyde group) the monosaccharide is called **aldose**. If the carbonyl group is at any other position (in a ketone group) monosaccharide is a **ketose**.
3. The simplest monosaccharides are two three-carbon trioses. These are glyceraldehyde, an aldotriose, and dihydroxyacetone, a ketotriose.
4. Monosaccharides with four, five, six, and seven carbon atoms in their backbones are called, respectively, tetroses, pentoses, hexoses, and heptoses. There are aldoses and ketoses of each of these chain lengths: aldotetroses and ketotetroses, aldopentoses and ketopentoses, and so on.
5. The hexoses include aldohexose D-glucose and ketohexose D-fructose. These are the most common monosaccharides in nature.
6. The aldopentoses Dribose and 2-deoxy-D-ribose are components of nucleotides and nucleic acids.

Monosaccharides Have Asymmetric Centers

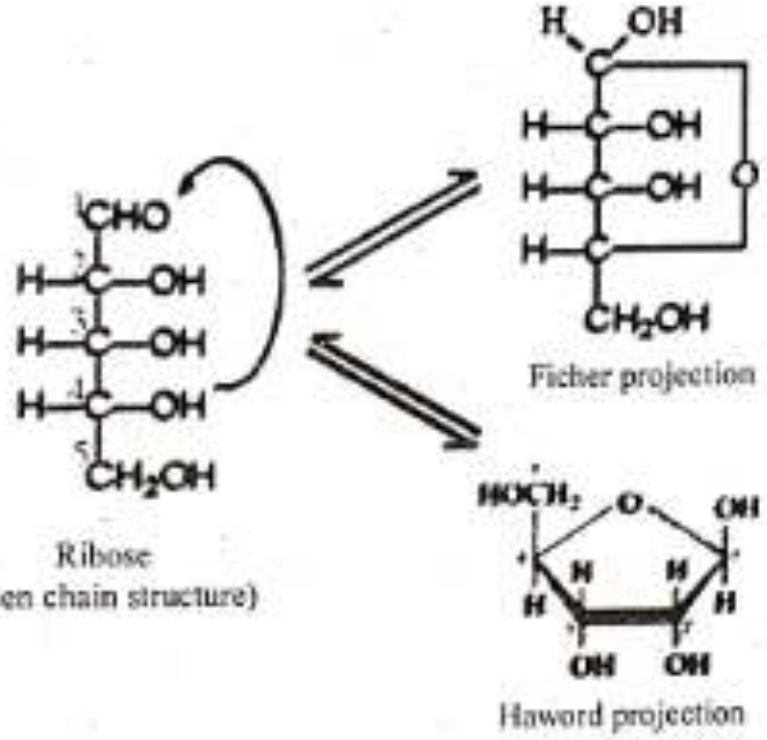
All monosaccharides except dihydroxyacetone contain one or more asymmetric (chiral) carbon atoms. Thus these occur in optically active isomeric forms. The simplest aldose, glyceraldehyde contains one chiral center (middle carbon atom). Therefore, it has two different optical isomers, or enantiomers.

Ribose is an aldopentose. It has molecular formula $C_5H_{10}O_5$. It exists in open chain structure in dried form. But it exists in furanose ring in aqueous medium.

Ribose is dissolved in water. As a result, oxygen atom from aldehyde group reacts with penultimate carbon (second last carbon C_4). Hence oxygen atom forms a link between C_1 and C_4. But OH group of C_4 is shifted to C_1. These changes form ring structure of ribose. It has two forms:

a) **Fischer projection:** This ring structure was demonstrated by **Emil Fischer**. Hence it is called **Fischer projection.** It is a two dimensional representation of ring structure. As for other biomolecules with chiral centers, absolute configurations of sugars are known from x-ray crystallography. We use Fischer projection formulas to represent three-dimensional sugar structures on paper. In Fischer projection formulas, horizontal bonds project out of plane of the paper are toward the reader. The vertical bonds project behind the plane of the paper is away from the reader.

b) **Haworth projection:** English chemist **Norman Haworth** also represented a ring structure. It is called as **Haworth projection**. It is a three dimensional representation of ring structure.

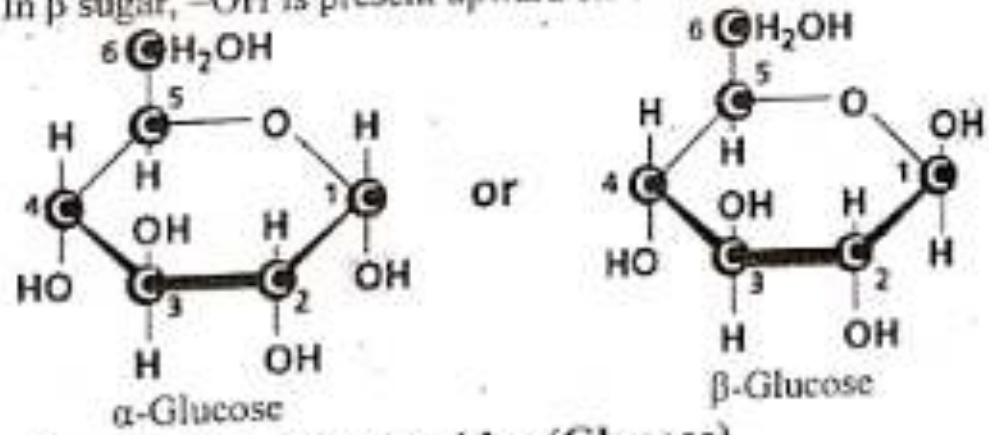


The α or β form

Each pentose or hexose molecule in ring structure may be α or β form. It depends on the position of –H and –OH group on C-1.

a) **α sugar:** In α sugar, –OH group is present downward on C-1.

b) **β- sugar:** In β sugar, –OH is present upward on C-1.



Stereoisomerism in Monosaccharides (Glucose)

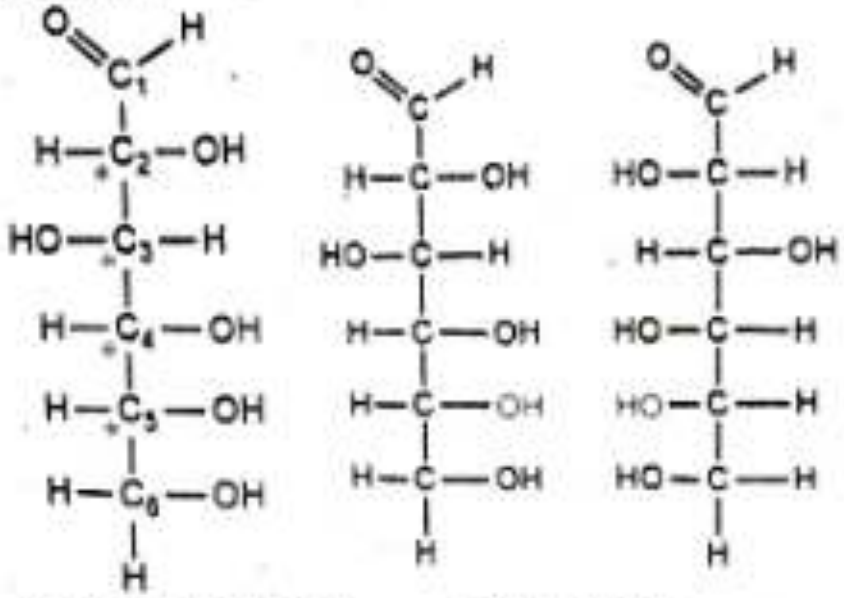
The isomers in which –H and –OH groups are arranged in different pattern to asymmetric carbon atoms are called stereoisomers. An asymmetric carbon atom form bonds with four different atoms around it.

Example

In glucose, C-2, C-3, C-4 and C-5 are asymmetric carbon atoms.

Calculation of stereoisomers

Number of stereoisomers depends on number of asymmetric carbons in monosaccharide. It can be calculated by formula 2^n. Here n is number of asymmetric carbon atoms. Hence glucose has 16 stereoisomers.



Asymmetric carbon Enantiomers

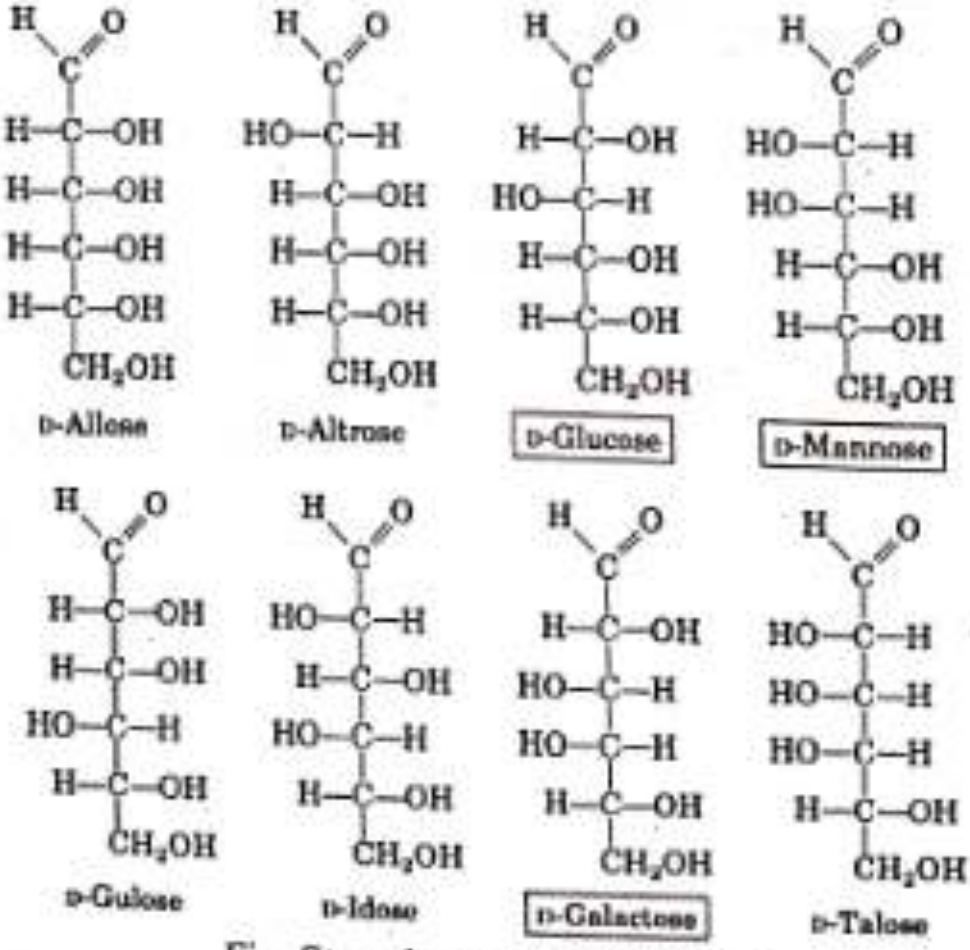


Fig: Stereoisomers of D-glucose

Classification of stereoisomers

Stereoisomers can be classified in thr

Asymmetric carbon Enantiomers

(i) **Enantiomers**

The stereoisomers which are non-superimposable mirror images of one another are called enantiomers. For example, enantiomers of glucose are D and L isomers.

a) **D isomers (right handed form):** In D isomer, asymmetric carbon atom is farthest from aldehyde group. It is second last carbon or C-5 in glucose. It is also called **penultimate carbon.** This carbon has –OH group on right side.

b) **L isomers (left handed form):** In L isomers, –OH group is projected on left side at penultimate carbon atom.

Glucose has 16 stereoisomers. Its 8 stereoisomers are enantiomers of other 8. (see above)

(ii) **Diastereoisomers**

The stereoisomers which have different arrangement of –H and –OH groups at more than one asymmetrical carbon atoms are called diastereoisomers. Diastereoisomers are not mirror images. For example, there are two carbohydrates D-Glucose and D-Altrose. They are diastereoisomers.

(iii) **Epimers**

The stereoisomers which have different arrangement of –H and –OH groups at only one asymmetrical carbon atom are called epimers. Example of epimer is D-Glucose and D-Mannose.

Cyclic Structure of Monosaccharides

For simplicity, the structures of aldoses and ketosesare represented as straight-chain molecules. Actually, in aqueous solution, aldotetroses and all monosaccharides with five or more carbon atoms occur predominantly as cyclic (ring) structures. In this structure, carbonyl group forms a covalent bond with oxygen of a hydroxyl group along the chain.

Formation of ring structure

These ring structures are formed by general reaction between alcohols and aldehydes or ketones. They form derivatives called **hemiacetals** or **hemiketals.** Hemiacetals contain an additional asymmetric carbon atom. Hence they can exist in two stereoisomeric forms. For example, D-glucose exists in solution as an intramolecular hemiacetals. Here free hydroxyl group at C-5 reacts with aldehydic C-1. It makes the latter carbon asymmetric. Thus it produces two stereoisomers, (See fig).

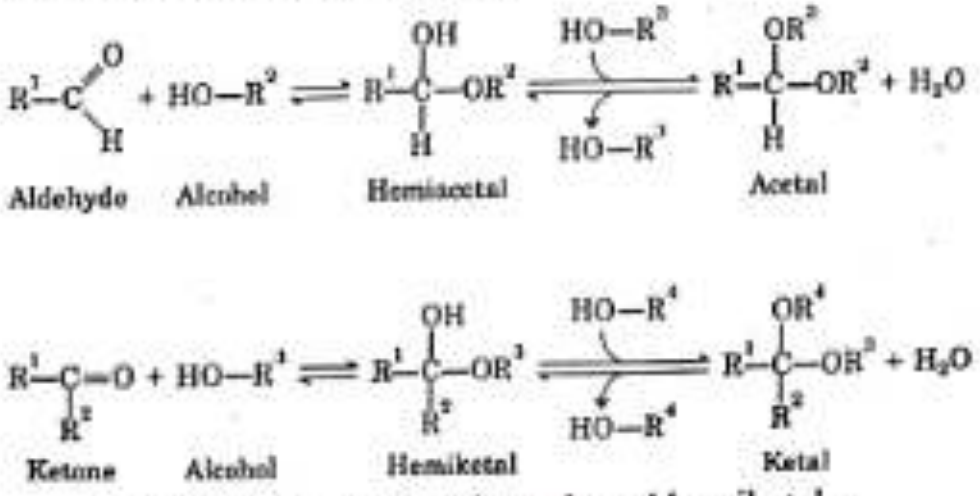


Fig: Formation of hemiacetals and hemiketals.

Types of compounds on the basis of rings

There are two types of compounds on the basis of ring structures:

a) **Pyranoses:** These six-membered ring compounds are called **pyranoses**. They resemble the sixmembered ring compound pyran. The systematic names for the two ring forms of D-glucose are α-D-glucopyranose and β-D-glucopyranose.

b) **Furanoses:** Aldohexoses also exist in cyclic forms. They have five membered rings. They ring resemble the fivemembered ring compound furan. Hence they are called **furanoses**.

The six-membered aldopyranose ring is much more stable than aldofuranose ring. Therefore, aldopyranose predominates in aldohexose solutions. Only aldoses having five or more carbon atoms can form pyranose rings.

Anomers

Isomeric forms of monosaccharides differing only in their configuration about the hemiacetal or hemiketal carbon atom are called **anomers**. The hemiacetal (or carbonyl) carbon atom is called the **anomeric carbon**.

Mutarotation

The α and β anomers of D-glucose interconvert in aqueous solution by a process called **mutarotation**. Thus, a solution of α-D-glucose and a solution of β-D-glucose form identical equilibrium mixtures. These mixtures have identical optical properties. This mixture consists of about one-third α-D-glucose, two-thirds β-D-glucose. It also has very small amounts of the linear and five-membered ring (glucofuranose) forms.

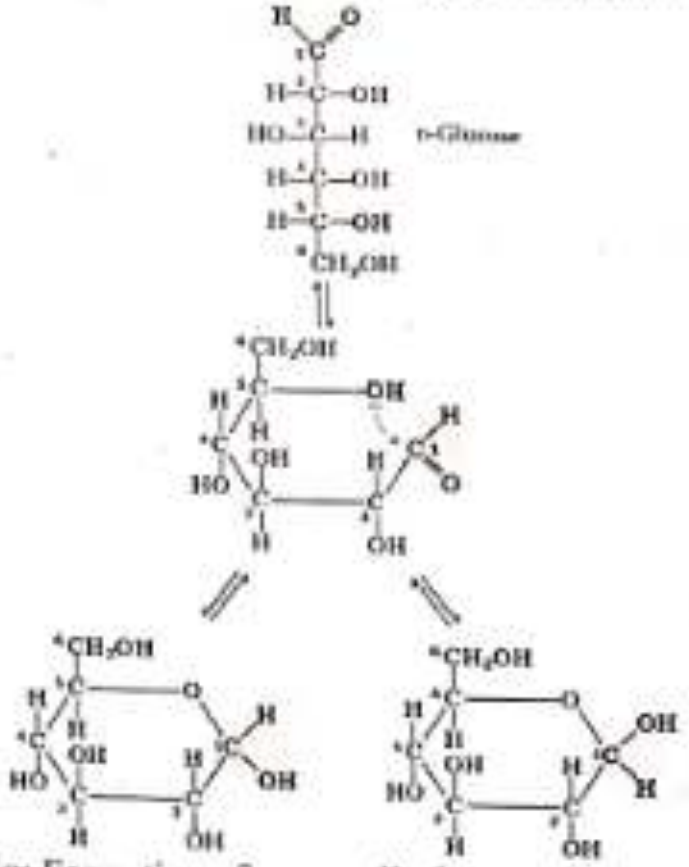
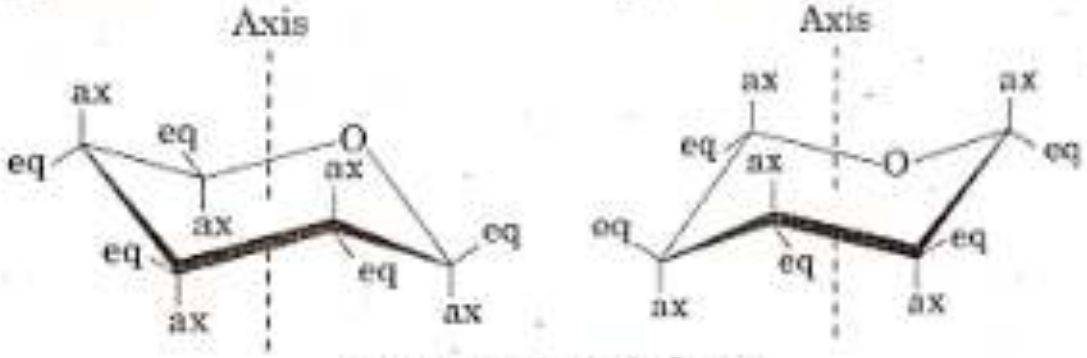


Fig: Formation of two cyclic forms of D-glucose

Ketohexoses also occur in α and β anomeric forms. In these compounds, hydroxyl group at C-5 (or C-6) reacts with the keto group at C-2. They a furanose (or pyranose) ring. This ring contains a hemiketal linkage. D-Fructose readily forms furanose ring. The more common anomer of this sugar in combined forms or in derivatives is -D-fructofuranose.

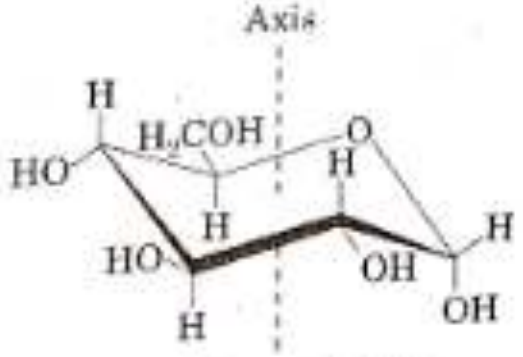
Haworth perspective formulas

Haworth perspective formulas are commonly used to show stereochemistry of ring forms of monosaccharides. However, the six-membered pyranose ring is not planar, as Haworth perspectives suggest. But it tends to assume either of two chair conformations. The two conformations of a molecule are interconvertible without the breakage of covalent bonds. But two configurations can be interconverted only by breaking a covalent bond.



Two possible chair forms

(a)



α-D-Glucopyranose

(b)

Fig: Conformational formulas of pyranoses. (a) Two chair forms of the pyranose ring. (b) The preferred chair conformation of -D-glucopyranose.

Organisms Contain a Variety of Hexose Derivatives

In addition to simple hexoses like glucose, galactose, and mannose, organisms also contains number of sugar derivatives. In these sugar derivatives, a hydroxyl group in the parent compound is replaced with another substituent or a carbon atom is oxidized to a carboxyl group.

1. In glucosamine, galactosamine, and mannosamine, hydroxyl at C-2 of parent compound is replaced with an amino group.

2. The amino group is nearly always condensed with acetic acid, as in N-acetylglucosamine. This glucosamine derivative is part of many structural polymers like bacterial cell wall.

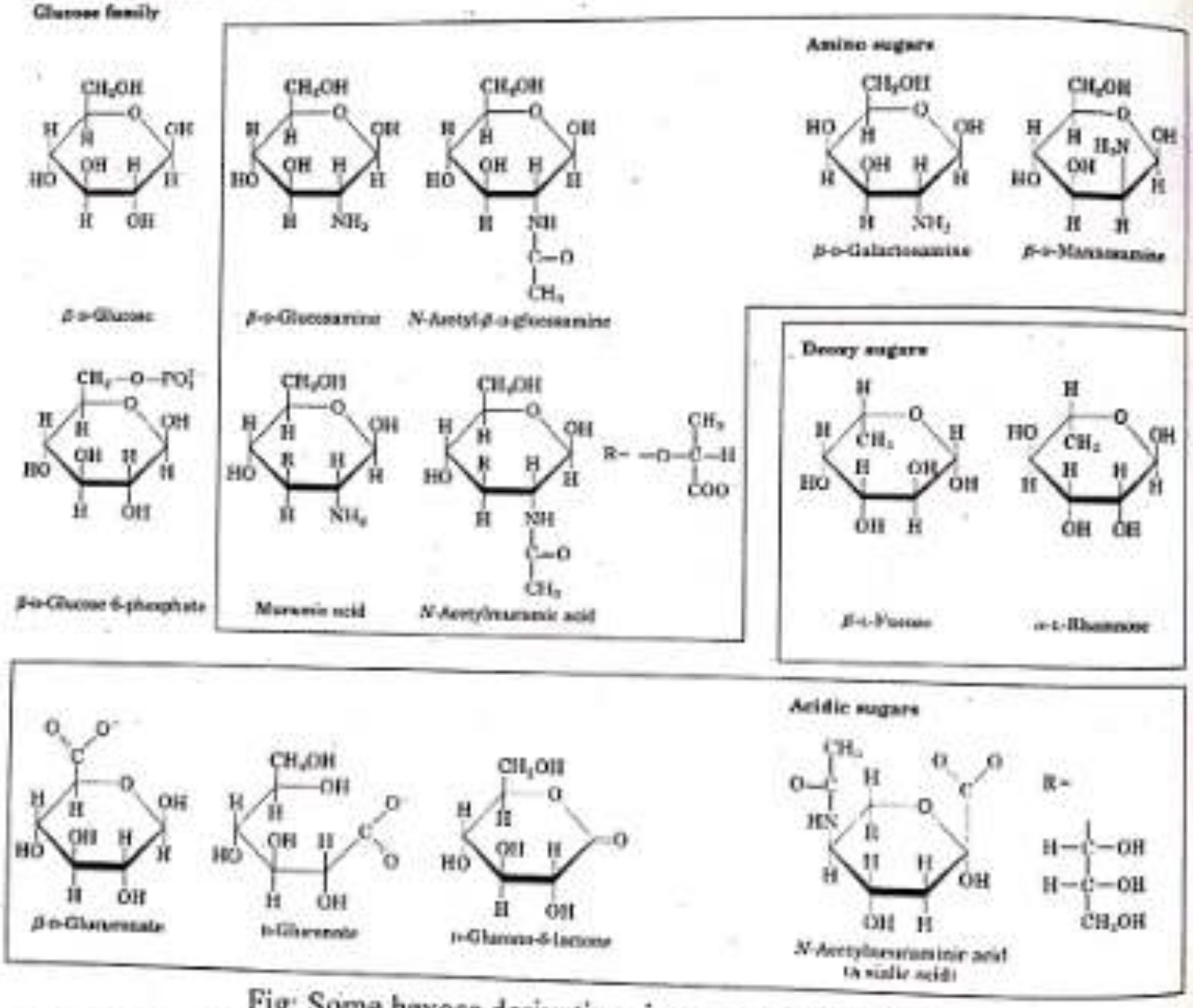


Fig: Some hexose derivatives important in biology

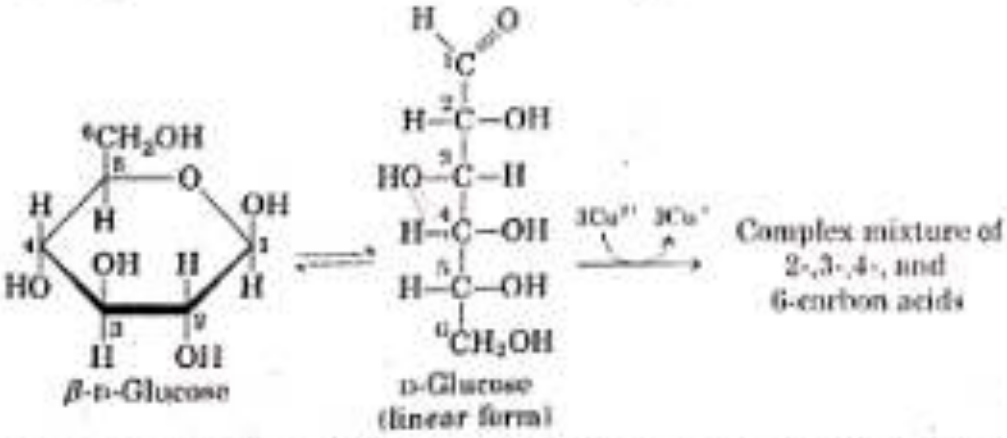
3. Bacterial cell walls also contain a derivative of glucosamine, N-acetylmuramic acid. In this derivative, lactic acid (a three carbon carboxylic acid) is ether-linked to the oxygen at C-3 of N-acetylglucosamine.
4. The substitution of hydrogen for the hydroxyl group at C-6 of L-galactose or L-mannose produces L-fucose or L-rhamnose, respectively. L-Fucose is found in the complex oligosaccharide components of glycoproteins and glycolipids; L-rhamnose is found in plant polysaccharides.
5. Oxidation of the carbonyl (aldehyde) carbon of glucose to the carboxyl level produces gluconic acid. Other aldoses yield other aldonic acids.
6. Oxidation of carbon at the other end of carbon chain C-6 of glucose, galactose, or mannose forms corresponding uronic acid. These acids are glucuronic, galacturonic,

or mannuronic acid. Both aldonic and uronic acids form stable intramolecular esters called lactones.

7. N-acetylneuraminic acid (a sialic acid) is a a derivative of N-acetylmannosamine. It is a component of many glycoproteins and glycolipids in animals.
8. The carboxylic acid groups of acidic sugar derivatives are ionized at pH 7. Hence these compounds are called as carboxylates—glucuronate, galacturonate, and so forth.
9. In the synthesis and metabolism of carbohydrates, the intermediates are very often not the sugars themselves. But they are their phosphorylated derivatives. Condensation of phosphoric acid with one of the hydroxyl groups of a sugar forms a phosphate ester, as in glucose 6-phosphate. Sugar phosphates are relatively stable at neutral pH and bear a negative charge. One effect of sugar phosphorylation within cells is to trap the sugar inside the cell. Most cells do not have plasma membrane transporters for phosphorylated sugars. Phosphorylation also activates sugars for subsequent chemical transformation. Several important phosphorylated derivatives of sugars are components of nucleotides.

Reducing non-reducing Sugars: Monosaccharides Are Reducing Agents

a) **Reducing sugars:** Monosaccharides can be oxidized by relatively mild oxidizing agents like cupric (Cu^{2+}) ion. The carbonyl carbon is oxidized to a carboxyl group. Glucose and other sugars which can reduce cupric ion are called reducing sugars. They form enediols. The enediols are converted to aldonic acids and then to a complex mixture of 2-, 3-, 4-, and 6-carbon acids. This is the basis of **Fehling's reaction**. This reaction a qualitative test for the presence of reducing sugar. We can measure the amount of oxidizing agent reduced by a solution of a sugar from this we can estimate the concentration of that sugar. For many years this test was used to detect and measure elevated glucose levels in blood and urine in the diagnosis of diabetes mellitus. All monosaccharides like glucose are reducing sugars. A disaccharide can be a reducing sugar or a non-reducing sugar. Maltose and lactose are reducing sugars.



b) **Non-reducing sugars:** Some other sugars cannot reduce cupric ion. These are called non-reducing sugars. Non-reducing sugars do not have an OH group attached to the

anomeric carbon. Hence they cannot reduce other compounds. All monosaccharides such as glucose are reducing sugars. Disaccharide sucrose is a non-reducing sugar.

DISA CCHARIDES

i) Disaccharides are most common oligosaccharide. Two monosaccharides combine to form a disaccharide.
ii) Disaccharides are less sweet in taste.
iii) They are less soluble in water.
iv) These can be hydrolyzed to give monosaccharides.
v) General formula of disaccharide is $C_{12}H_{22}O_{11}$
vi) Examples: maltose, lactose, sucrose. On. Some common disaccharides are as follows:

Glycosidic bond formation in Disaccharide

Disaccharides consist of two monosaccharides. They are joined covalently by an **O-glycosidic bond**. This bond is formed when a hydroxyl group of one sugar reacts with the anomeric carbon of the other. This reaction represents the formation of an acetal from a hemiacetal (such as glucopyranose) and an alcohol. Glycosidic bonds are readily hydrolyzed by acid. But resist cleavage by base. Thus disaccharides are boiled with dilute acid. They are hydrolyzed to form their free monosaccharide components.

Reducing properties

The oxidation of a sugar's anomeric carbon by cupric or ferric ion occurs only with the linear form. Their linear form exists in equilibrium with the cyclic form(s). In some disaccharides, anomeric carbon is involved in a glycosidic bond, such sugar residue cannot take the linear form. Therefore, it becomes a nonreducing sugar. In describing disaccharides or polysaccharides, the end of a chain with a free anomeric carbon (one not involved in a glycosidic bond) is commonly called the **reducing end**.

Common disaccharides

Common disaccharides are sucrose, lactose and maltose.

Naming of Oligosaccharides

By convention, the name describes the compound written with its non-reducing end to the left. Then we can build up the name in the following order.

1. Give the configuration (α or β) at the anomeric carbon joining the first monosaccharide unit (on the left) to the second.
2. Name the non-reducing residue to distinguish five- and six-membered ring structures. Insert **furano** or **pyrano** into the name.
3. Indicate in parentheses the two carbon atoms joined by glycosidic bond with an arrow. This arrow connects the two numbers; for example, (1→4) shows that C-1 of the first-named sugar residue is joined to C-4 of second.

4. Name the second residue. If there is a third residue, describe the second glycosidic bond by same conventions.

Shortening the Description in name

We can shorten the description of complex polysaccharides. For this purpose, three-letter abbreviations or colored symbols for monosaccharides are used. Following this convention for naming oligosaccharides, maltose is α-Dglucopyranosyl-(1→4)-D-glucopyranose. Most sugars in living organisms are D enantiomers. In these compounds, pyranose form of hexoses predominates. Therefore, generally use a shortened version of the formal name of such compounds. We give the configuration of the anomeric carbon and naming the carbons joined by the glycosidic bond. In this abbreviated nomenclature, maltose is Glc (α1→4) Glc.

Table: Symbols and Abbreviations of Common Oligosaccharides

Abequose	Abe	Glucuronic acid	◆ GlcA
Arabinose	Ara	Galactosamine	□ GalN
Fructose	Fru	Glucosamine	◥ GlcN
Fucose	▲ Fuc	*N*-Acetylgalactosamine	□ GalNAc
Galactose	○ Gal	*N*-Acetylglucosamine	■ GlcNAc
Glucose	● Glc	Iduronic acid	◇ IdoA
Mannose	◎ Man	Muramic acid	Mur
Rhamnose	Rha	*N*-Acetylmuramic acid	Mur2Ac
Ribose	Rib	*N*-Acetylneuraminic acid (a sialic acid)	◆ Neu5Ac
Xylose	☆ Xyl		

Maltose

Occurrence

i) It is commonly called malt sugar. It is an intermediate disaccharide. It is produced during breakdown of starch and glycogen.

ii) Maltose is found in germinating seeds.

iii) Disaccharide retains a free anomeric carbon. Therefore, maltose is a reducing sugar.

Use

It is produced in brewing industry. It is produced from breakdown of barley starch. This reaction occurs with help of amylase enzyme. This process is known as **malting**.

Formation

It yields two molecules of α-D-glucoses on hydrolysis. Therefore, maltose is formed by condensation of α-D-glucoses. In this reaction,–OH group at C-1 of one glucose reacts with –OH group at C-4 of other glucose. It releases a water molecule. As a result, linkage is formed between C-1 of one glucose and C-4 of other glucose. This linkage is called **α-1,4-glycosidic linkage.**

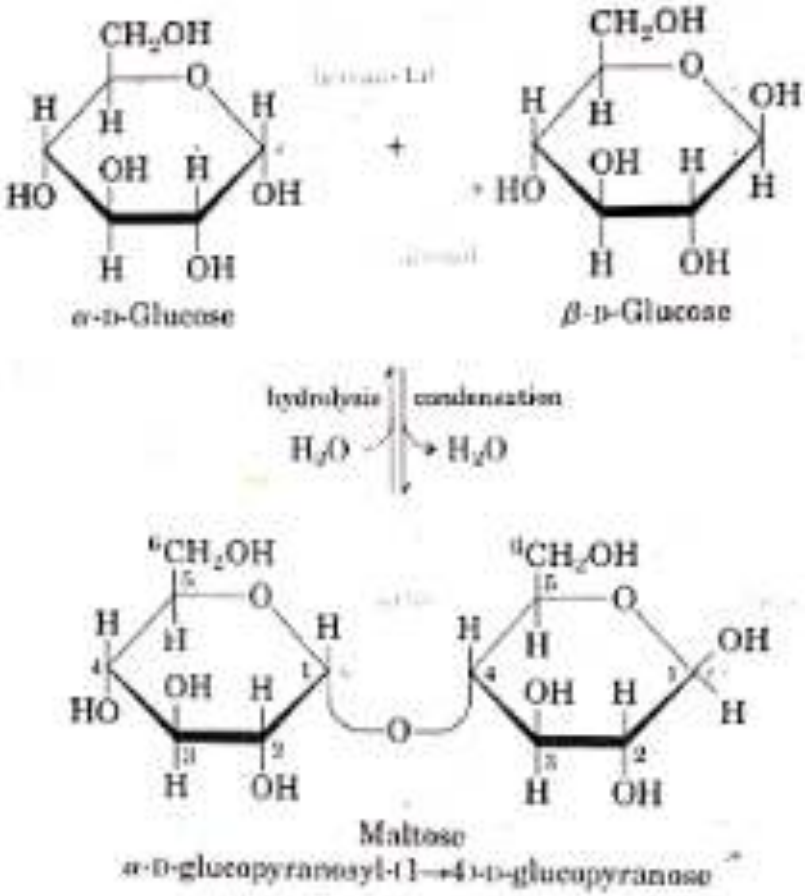
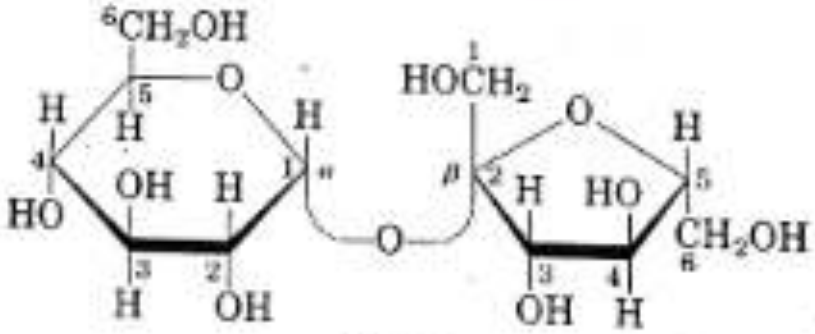


Fig: Formation of maltose

Sucrose

Sucrose contains no free anomeric carbon atom. The anomeric carbons of both monosaccharide units are involved in the glycosidic bond. Therefore, sucrose is a non-reducing sugar. In the abbreviated nomenclature, a double-headed arrow connects the symbols specifying anomeric carbons and their configurations. For example, abbreviated name of sucrose is Glc(α1↔2β) Fru or Fru (β2↔1α)Glc.



Sucrose
β-D-fructofuranosyl α-D-glucopyranoside
Fru(2β↔α1)Glc = Glc(α1↔2β)Fru

Occurrence

- It occurs abundantly in sugarcane. Therefore, it is commonly called cane sugar.
- Many other plants also have sucrose in large amount. These are like sugar beet, pine apple, sorghum and all sweet fruits.

Uses

- It is widely used as sweetener at homes. It is used for making sweet dishes.
- Prepared food in plants is transported in the form of sucrose. Hence it is also called transport disaccharide in plants

Formation

Sucrose yields two molecules on hydrolysis. One is α-glucose and other is β-fructose. Therefore, sucrose is formed by condensation of glucose and fructose. In this reaction, –OH group at C-1 of glucose reacts with–OH group at C-2 of fructose. It releases a water molecule.

As a result, a linkage is formed between C-1 of glucose and C-2 of fructose. This linkage is called **α-1,2-glycosidic linkage**.

Lactose

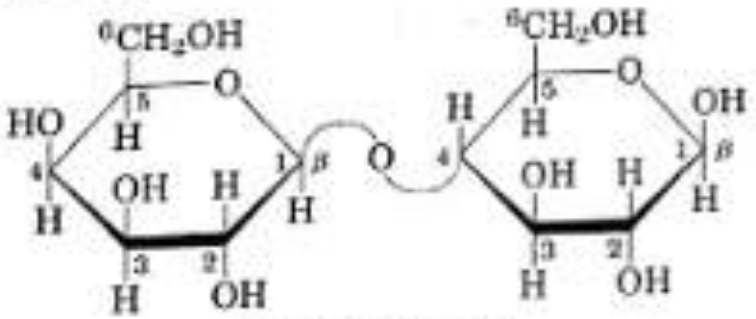
The disaccharide lactose yields D-galactose and D-glucose on hydrolysis. The anomeric carbon of the glucose residue is available for oxidation. Thus lactose is a reducing disaccharide. Its abbreviated name is Gal(β1→4)Glc

Occurrence:

- It is commonly called **milk sugar**. It is found in milk of mammals. It is 4-6% in cow's milk and 5-8% in human milk.
- It is also a byproduct (waste) in manufacture of cheese.

Formation

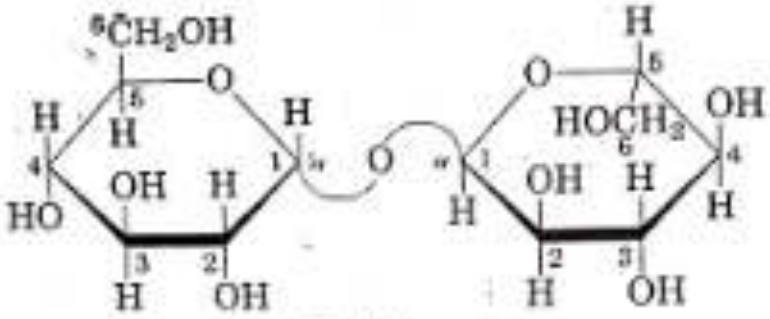
It yields a molecule of β-galactose and molecule of β-glucose on hydrolysis. Therefore, lactose is formed by the condensation of β-galactose and β-glucose. In this reaction, –OH group at C-1 of galactose reacts with –OH group at C-4 of glucose. It releases a water molecule. As a result linkage is formed between C-1 of galactose and C-4 of glucose. This linkage is called β-1, 4-glycosidic linkage.



Lactose (β form)
β-D-galactopyranosyl-(1→4)-β-D-glucopyranose
Gal(β1→4)Glc

Trehalose

Its abbreviated formula is Glc(α1↔1α) Glc. It is a disaccharide of D-glucose. It is a non-reducing sugar. It is a major constituent of the circulating fluid (hemolymph) of insects. It acts as as an energy-storage compound. Fungi also contain trehalose. These are used as a commercial source of this sugar.



Trehalose
α-D-glucopyranosyl α-D-glucopyranoside
Glc(α1↔1α)Glc

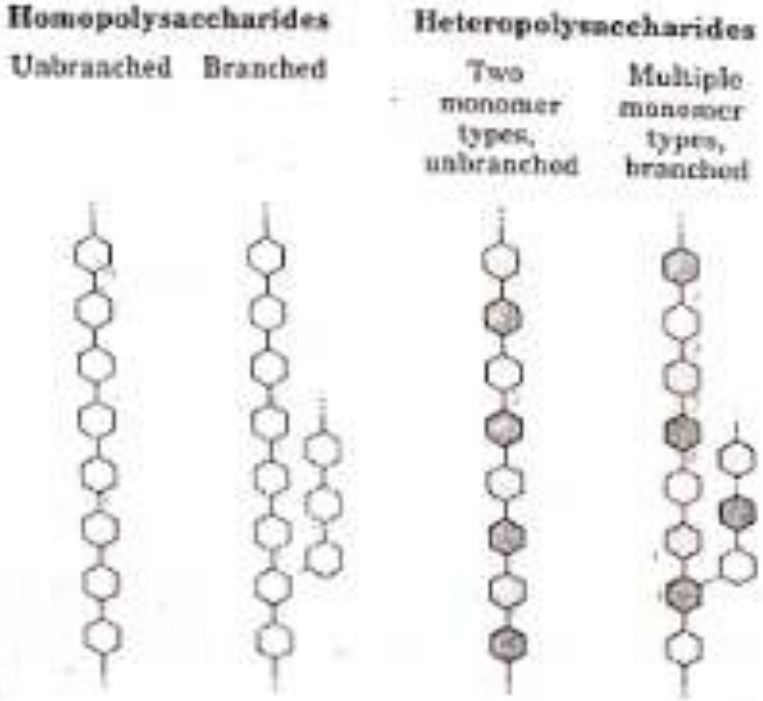
POLYSACCHARIDES

The carbohydrates on hydrolysis yielding more than ten monosaccharide units are called polysaccharides. This is largest group of carbohydrates. Polysaccharides, also called glycans. They differ from each other in:

- The identity of their recurring monosaccharide units
- In length of their chains in the types of bonds linking the units
- It the degree of branching

Polysaccharides are of two types:

a) **Homopolysaccharides: The polysaccharides formed by condensation of only one kind of monosaccharides are called homopolysaccharides.** Its examples are starch, glycogen, cellulose and chitin. Some homopolysaccharides serve as storage forms of monosaccharides. These are used as fuels. Starch and glycogen are homopolysaccharides of this type. Other homopolysaccharides (cellulose and chitin,) serve as structural elements in plant cell walls and animal exoskeletons.



b) **Heteropolysaccharides: The polysaccharides formed by condensation of different kind of monosaccharides are called heteropolysaccharides.** Its examples are agar, pectin and peptidoglycan. Heteropolysaccharides provide extracellular support for organisms of all kingdoms. For example, rigid layer of the bacterial cell envelope (peptidoglycan) is composed in part of a heteropolysaccharide. It is built from two alternating monosaccharide units. In animal tissues, extracellular space is occupied by several types of heteropolysaccharides. They form a matrix that holds individual cells together. It also provides protection, shape, and support to cells, tissues, and organs

Importance or functions of polysaccharides

i) **Storage molecules:** Polysaccharides act as food and energy stores. These storage molecules are like starch, glycogen. They are convenient storage molecule. It has several reasons.

- They have large size. It makes them more or less insoluble in water. Hence they exert no osmotic or chemical influence in the cell
- They fold into compact shapes.
- They are easily converted to sugars by hydrolysis.

ii) **Structural material**: Its examples are cellulose and chitin.

Examples

Some common polysaccharides are starch, cellulose and chitin.

Starch

i) **Formation:** Starch is a homopolysaccharides. It is formed by condensation of hundreds of **α-glucoses**.

ii) **Occurrence:** It is storage carbohydrate of plants. It is mainly stored in root, stem and seeds. Cereal grains and potato tubers are rich sources of starch in human diet.

iii) **Digestion of starch:** Starch is digested in oral cavity and in small intestine. It is digested by enzyme amylase. On hydrolysis it yields maltose first. Then maltose is further digested by maltase enzyme. It yields glucoses.

iv) **Test:** We use iodine test to test presence of starch in a given sample. It gives blue colour with iodine solution.

v) **Types:** There are two types of starches. These are amylose and amylopectin.

(a) **Amylose**

- Amylose is un-branched. It is linear chain of glucoses.
- Here glucoses are attached together by α-1,4-glycosidic linkages.
- It is soluble in hot water only.

(b) **Amylopectin**

- Amylopectin has branched structure. It is a linear chain of glucoses. But it has more chains of glucoses in the form of branches.
- These branches are also attached by α-1,6-glycosidic linkages.
- It is completely insoluble in water.

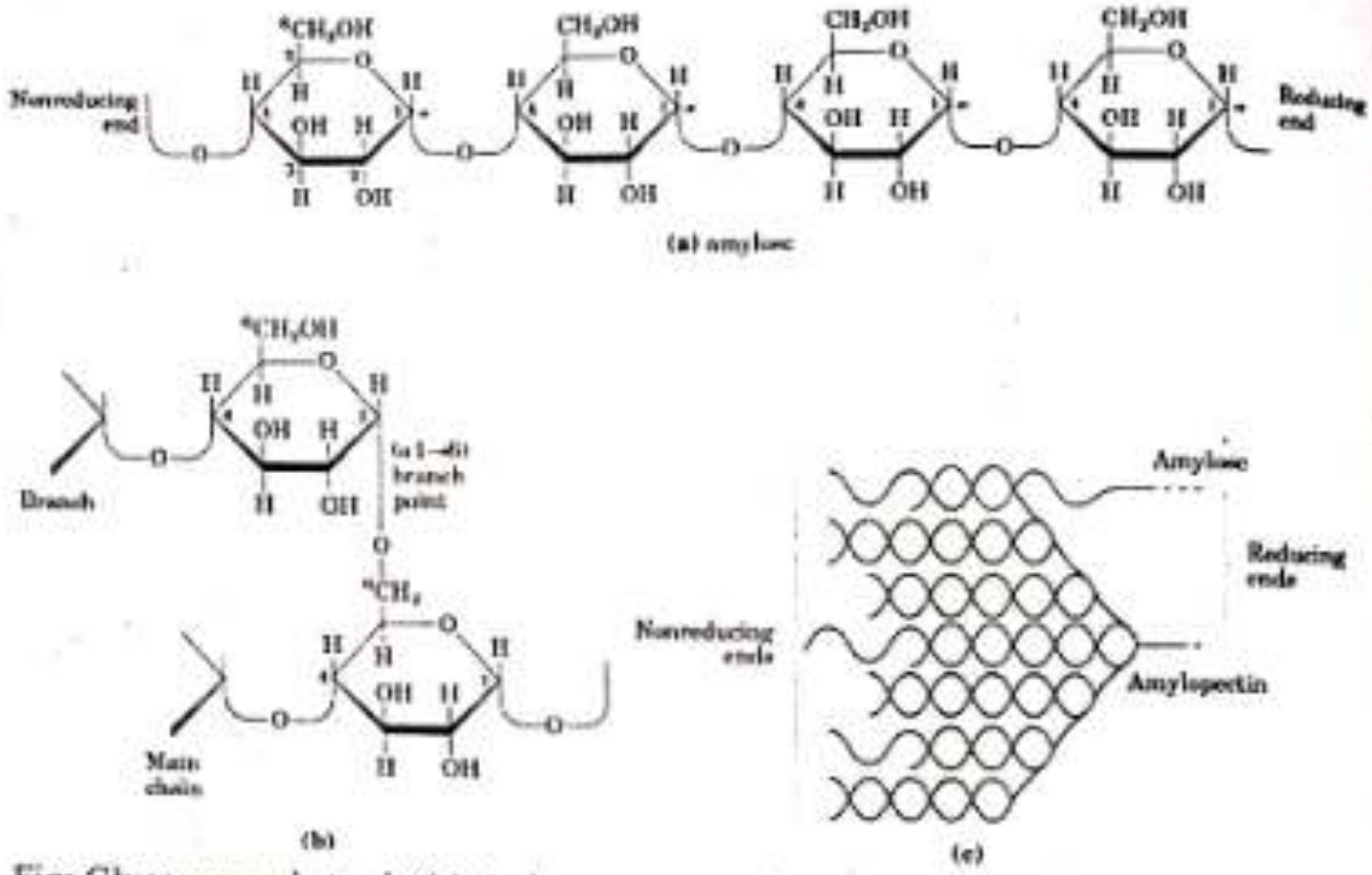


Fig: Glycogen and starch. (a) A short segment of amylose, a linear polymer of D-glucose residues in (1→4) linkage. (b) An (1→6) branch point of glycogen or amylopectin. (c) A cluster of amylose and amylopectin like that believed to occur in starch granules.

Glycogen

i) **Formation:** Glycogen is also a homopolysaccharides. It is composed of α-glucoses.

ii) **Occurrence:** It is storage carbohydrate of animals. It is mainly stored in liver and muscles. Therefore it is also called **animal's starch.**

iii) **Digestion:** Its digestion is similar to digestion of starch.

iv) **Test:** We use iodine test to determine presence of glycogen in a given sample. This iodine test gives red colour with iodine solution.

v) **Structure:** Its structure resembles with amylopectin starch. But glycogen has much more branching than amylopectin.

Cellulose

Cellulose is most abundant carbohydrate on earth.

i) **Formation:** It is a homopolysaccharides. It is formed by the condensation of hundreds of β-glucoses.

ii) **Role:** It is structural carbohydrate of plants. It is a major constituent of plant cell wall. Cotton and paper are pure forms of cellulose.

iii) **Test:** Cellulose shows no colour with iodine solution.

iv) **Structure:** Its structure resembles with amylose starch. It has un-branched structure. But it has -1, 4-glycosiciic linkages between glucose residues. Therefore, cellulose

chain has β-glucoses. These are alternatively arranged in upright and inverted manner.

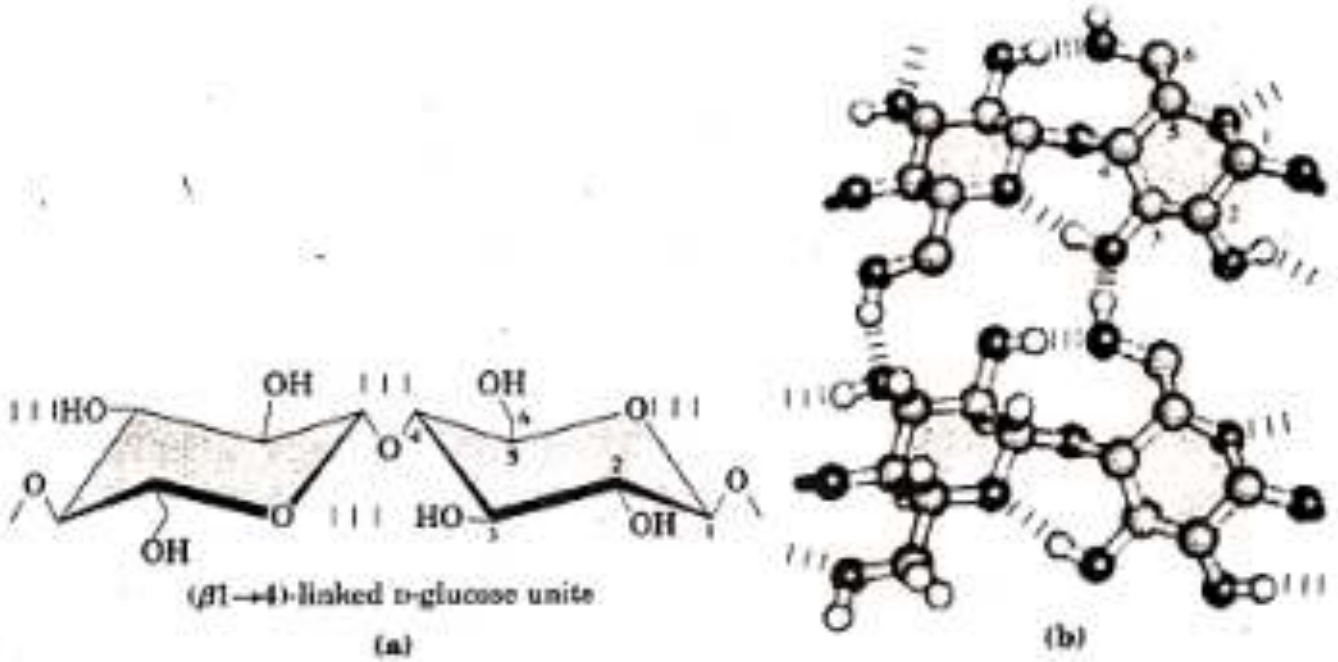


Fig: Cellulose. (a) Two units of a cellulose chain; (b) Scale drawing of segments of two parallel cellulose chains.

Digestion of Starch

Glycogen and starch ingested in the diet. These are hydrolyzed by α-amylases and glycosidases. These enzymes are found in saliva and the intestine that break (α1→4) glycosidic bonds between glucose units. Most animals cannot use cellulose as a fuel source. They lack an enzyme to hydrolyze the (β1→4) linkages. Termites readily digest cellulose (and therefore wood). Their intestinal tracts contain a symbiotic microorganism. These microorganisms digest cellulose.

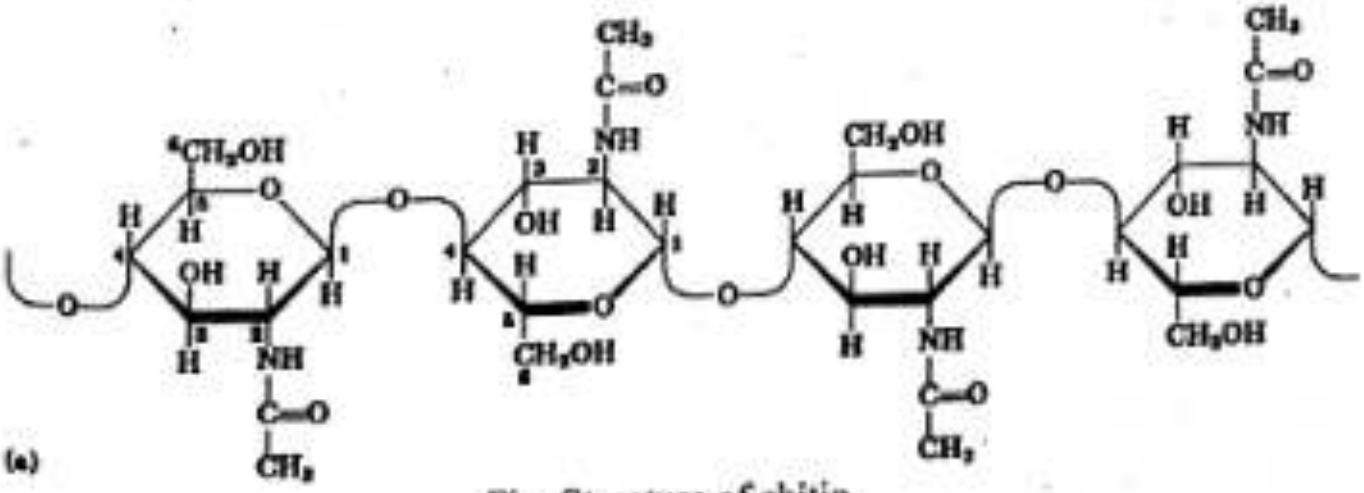


Fig: Structure of chitin

Chitin

i) **Nature:** Chitin is second most abundant organic molecule on earth. It is also a homopolysaccharides. It is a structural carbohydrate.

ii) **Occurrence:** It is found in the cell walls of fungi and exoskeleton of arthropods. It occurs in fungal cell wall. Therefore it is also called **fungal cellulose.**

iii) **Structure:** Chitin is derivative of N-acetyl glucosamine. It is a modified form of glucose. It has an un-branched structure like cellulose. It has alternative upright and inverted N-acetyl glucosamine residues. These are linked together by -1, 4-glycosidic linkages.

Table: Some polysaccharides with their functions

Polymer	Type*	Repeating unit†	Size (number of monosaccharide units)	Roles/significance
Starch				Energy storage: in plants
Amylose	Homo-	(α1→4)Glc, linear	50-5,000	
Amylopectin	Homo-	(α1→4)Glc, with (α1→6)Glc branches every 24-30 residues	Up to 10^6	
Glycogen	Homo-	(α1→4)Glc, with (α1→6)Glc branches every 8-12 residues	Up to 50,000	Energy storage: in bacteria and animal cells
Cellulose	Homo-	(β1→4)Glc	Up to 15,000	Structural: in plants, gives rigidity and strength to cell walls
Chitin	Homo-	(β1→4)GlcNAc	Very large	Structural: in insects, spiders, crustaceans, gives rigidity and strength to exoskeletons
Dextran	Homo-	(α1→6)Glc, with (α1→3) branches	Wide range	Structural: in bacteria, extracellular adhesive
Peptidoglycan	Hetero-; peptides attached	4)Mur2Ac(β1→4) GlcNAc(β1	Very large	Structural: in bacteria, gives rigidity and strength to cell envelope
Agarose	Hetero-	3)D-Gal(β1→4)3,6-anhydro-L-Gal(α1	1,000	Structural: in algae, cell wall material
Hyaluronate (a glycosaminoglycan)	Hetero-; acidic	4)GlcA(β1→3) GlcNAc(β1	Up to 100,000	Structural: in vertebrates, extracellular matrix of skin and connective tissue; viscosity and lubrication in joints

Dextrans

Dextrans are bacterial and yeast polysaccharides. They are made up of (1-6)-linked poly-D-glucose. They all have (1-3) branches, and some also have (1-2) or (1-4) branches. Dental plaque is formed by bacteria growing on the surface of teeth. It is rich in dextrans. Synthetic dextrans are used in several commercial products. The dextrans in these products are chemically cross-linked. They form insoluble materials of various porosities.

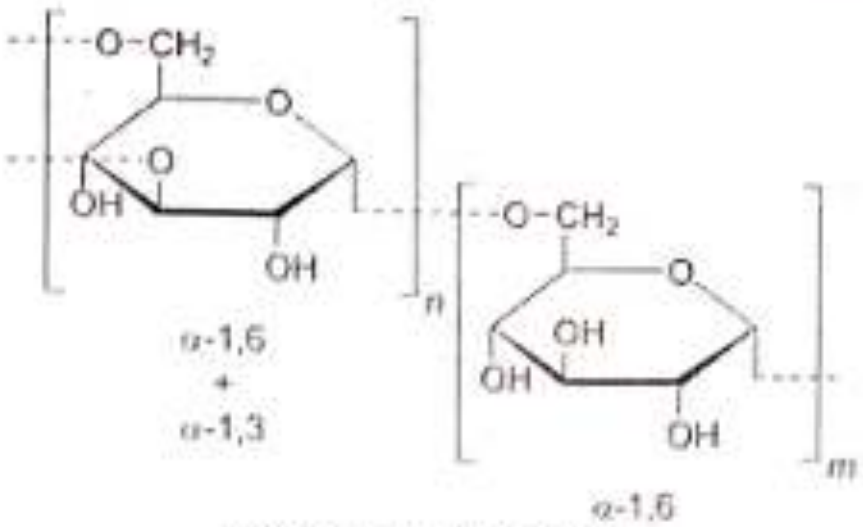


Fig: Structure of Dextrans

GLYCOSAMINOGLYCANS

Glycosaminoglycans are heteropolysaccharides of extracellular matrix. The extracellular space in the tissues of multicellular animals is filled with a gel-like material called **extracellular matrix (ECM)**. It is also called ground substance. It holds the cells together and provides a porous pathway for diffusion of nutrients and oxygen to individual cells. **Reticular ECM** surrounds fibroblasts and other connective tissue cells. It is composed of an interlocking meshwork of heteropolysaccharides and fibrous proteins like fibrillar collagens, elastin, and fibronectin. Basement membrane is a specialized ECM. It underlies epithelial cells. It consists of specialized collagens, laminin, and heteropolysaccharides. These heteropolysaccharides are glycosaminoglycans.

Composition and Structure

These are a family of linear polymers composed of repeating disaccharide units. They are unique to animals and bacteria and are not found in plants. One of the two monosaccharides is always either N-acetylglucosamine or N-acetylgalactosamine. The other is in most cases an uronic acid. It is usually D-glucuronic or L-iduronic acid. These have some modifications:

1. Some glycosaminoglycans contain esterified sulfate groups. The combination of sulfate groups and carboxylate groups of uronic acid residues gives glycosaminoglycans a very high density of negative charge.
2. There are some repulsive forces among neighboring charged groups. To minimize these repulsive forces, these molecules assume an extended conformation in solution. They form a rod-like helix in which the negatively charged carboxylate groups occur on alternate sides of the helix. The extended rod form also provides maximum separation between negatively charged sulfate groups.
3. There are specific patterns of sulfated and non-sulfated sugar residues in glycosaminoglycans. They provide for specific recognition by a variety of protein ligands. These ligands bind electrostatically to these molecules. The sulfated glycosaminoglycans are attached to extracellular proteins to form proteoglycans.

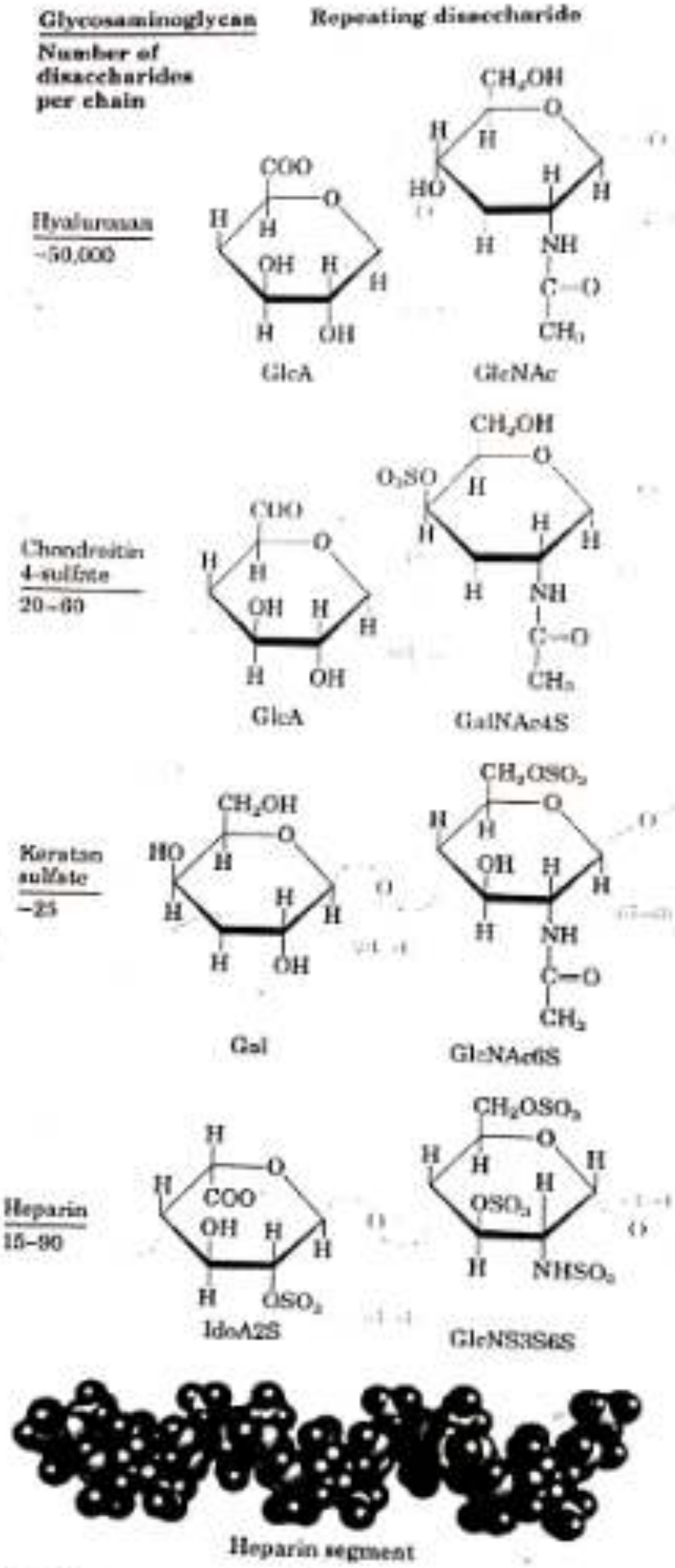


Fig: Repeating units of some common glycosaminoglycans of extracellular matrix

Forms of Glycosaminoglycans

Glycosaminoglycan Hyaluronan

The **glycosaminoglycan hyaluronan** (hyaluronic acid) contains alternating residues of D-glucuronic acid and N-acetylglucosamine. It has up to 50,000 repeats of the basic disaccharide unit. Thus hyaluronan has a molecular weight of several million. It forms clear, highly viscous solutions.

a) These solutions serve as lubricants in the **synovial fluid** of joints.
b) They also give the vitreous humor of the vertebrate eye its jellylike consistency (Greek hyalos means glass; hyaluronan can have a glassy or translucent appearance).
c) Hyaluronan is also a component of the extracellular matrix of cartilage and tendons. They provide tensile strength and elasticity to them as a result of its strong interactions with other components of the matrix.

Hyaluronidase is an enzyme secreted by some pathogenic bacteria. It can hydrolyze glycosidic linkages of hyaluronan. Thus it makes the tissues more susceptible to bacterial invasion. In many species, a similar enzyme in sperm hydrolyzes an outer glycosaminoglycan coat around the ovum. It allows sperm penetration.

Other glycosaminoglycans

Other glycosaminoglycans differ from hyaluronan in three respects.

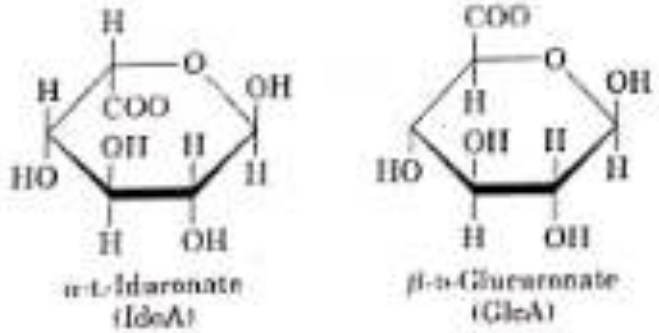
a) They are generally much shorter polymers.
b) They are covalently linked to specific proteins (proteoglycans)
c) One or both monomeric units differ from those of hyaluronan.

1. **Chondroitin sulfate** (Greek chondros, cartilage)

It contributes to the tensile strength of cartilage, tendons, ligaments, and the walls of the aorta.

2. **Dermatan sulfate** (Greek derma, "skin").

It contributes to the pliability of skin. It is also present in blood vessels and heart valves. In this polymer, many of glucuronate residues present in chondroitin sulfate are replaced by their 5-epimer, L-iduronate.



3. **Keratan sulfates** (Greek keras, horn)

They have no uronic acid and their sulfate content is variable. They are present in cornea, cartilage, bone, and a variety of horny structures. They form of dead cells like horn, hair, hoofs, nails, and claws.

4. **Heparan sulfate** (Greek h par, liver)

It is produced by all animal cells. It contains variable arrangements of sulfated and non-sulfated sugars. The sulfated segments of chain allow it to interact with a large number of proteins, including growth factors and ECM components and various enzymes and factors present in plasma.

Heparin is a fractionated form of heparan sulfate. It is derived mostly from mast cells (a type of leukocyte). Heparin is a therapeutic agent used to inhibit coagulation through its capacity to bind the protease inhibitor antithrombin. Heparin binding causes antithrombin to bind to and inhibit thrombin. Thrombin is a protease essential to blood clotting. The interaction is strongly electrostatic. Heparin has the highest negative charge density of any known biological macromolecule. Purified heparin prevents clotting. Thus it is routinely added to blood samples obtained for clinical analysis. It is also added to blood donated for transfusion.

GLYCOCONJUGATES:

PROTEOGLYCANS, GLYCOPROTEINS AND GLYCOLIPIDS

Glycoconjugate are biologically active molecule in which informational carbohydrate is covalently joined to a protein or a lipid.

Glycoconjugate performs different functions:

1. Some glycoconjugate provides communication between cells and their extracellular surroundings.
2. Some label proteins for transport in specific organelles. Some are used for destruction when the protein is malformed or superfluous.
3. Some are used for recognition sites for extracellular signal molecules (growth factors, for example) or extracellular parasites (bacteria or viruses).
4. On almost every eukaryotic cell, specific oligosaccharide chains attached to components of plasma membrane form a carbohydrate layer (glycocalyx). It is several nanometers thick. It serves as an information-rich surface that a cell shows to its surroundings. These oligosaccharides are central players in cell-cell recognition and adhesion, cell migration during development, blood clotting, immune response, wound healing, and other cellular processes.

Form of Carbohydrates

These are of different types:

1. **Proteoglycans**

These are macromolecules of the cell surface or extracellular matrix. In these molecules, one or more sulfated **glycosaminoglycan** chains are joined covalently to a membrane protein or a secreted protein. The glycosaminoglycan chain can bind to extracellular proteins through electrostatic interactions with the negatively charged groups on the polysaccharide. Proteoglycans are major components of all extracellular matrices.

2. Glycoproteins

Glycoproteins have one or several oligosaccharides of varying complexity. They are joined covalently to a protein. They are found on the outer face of the plasma membrane (as part of the glycocalyx), in the extracellular matrix, and in blood. Inside cells they are found in specific organelles like Golgi complexes, secretory granules, and lysosomes. The oligosaccharide portions of glycoproteins are very heterogeneous. Like glycosaminoglycans, they are rich in information. They form highly specific sites for recognition and high affinity binding by carbohydrate-binding proteins called **lectins**. Some cytosolic and nuclear proteins can be glycosylated as well.

3. Glycolipids

Glycolipids are membrane sphingolipids in which hydrophilic head groups are oligosaccharides. As in glycoproteins, the oligosaccharides act as specific sites for recognition by lectins. The brain and neurons are rich in glycolipids. They help in nerve conduction and myelin formation. Glycolipids also play a role in signal transduction in cells.

Proteoglycans

Proteoglycans are glycosaminoglycan-containing macromolecules of cell surface and extracellular matrix. Mammalian cells can produce 40 types of proteoglycans. These molecules act as tissue organizers. They influence various cellular activities like growth factor activation and adhesion.

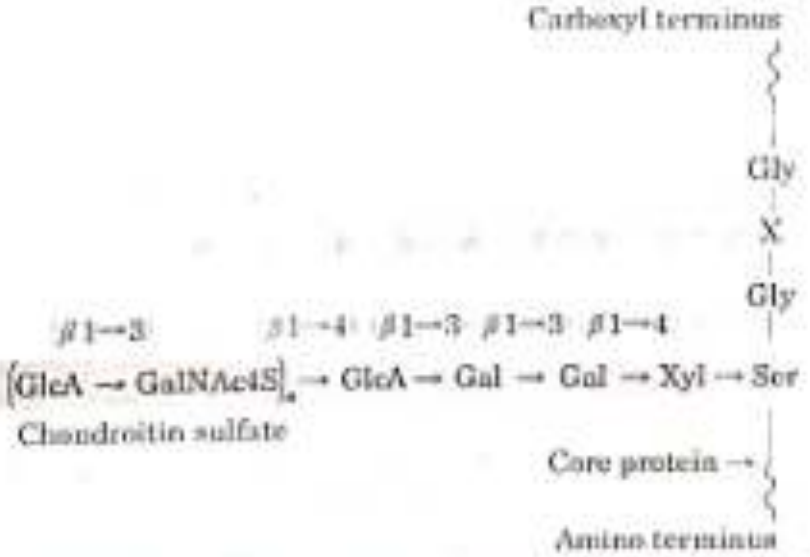


Fig: Proteoglycan structure, showing the tetrasaccharide bridge

Structure of basic proteoglycan unit

The basic proteoglycan unit consists of a core protein with covalently attached glycosaminoglycan(s). The point of attachment is a Ser residue. The glycosaminoglycan is joined to this Ser residue through a tetrasaccharide bridge. The Ser residue is generally in the sequence –Ser–Gly–X–Gly– (where X is any amino acid residue). But every protein is not with this sequence has an attached glycosaminoglycan.

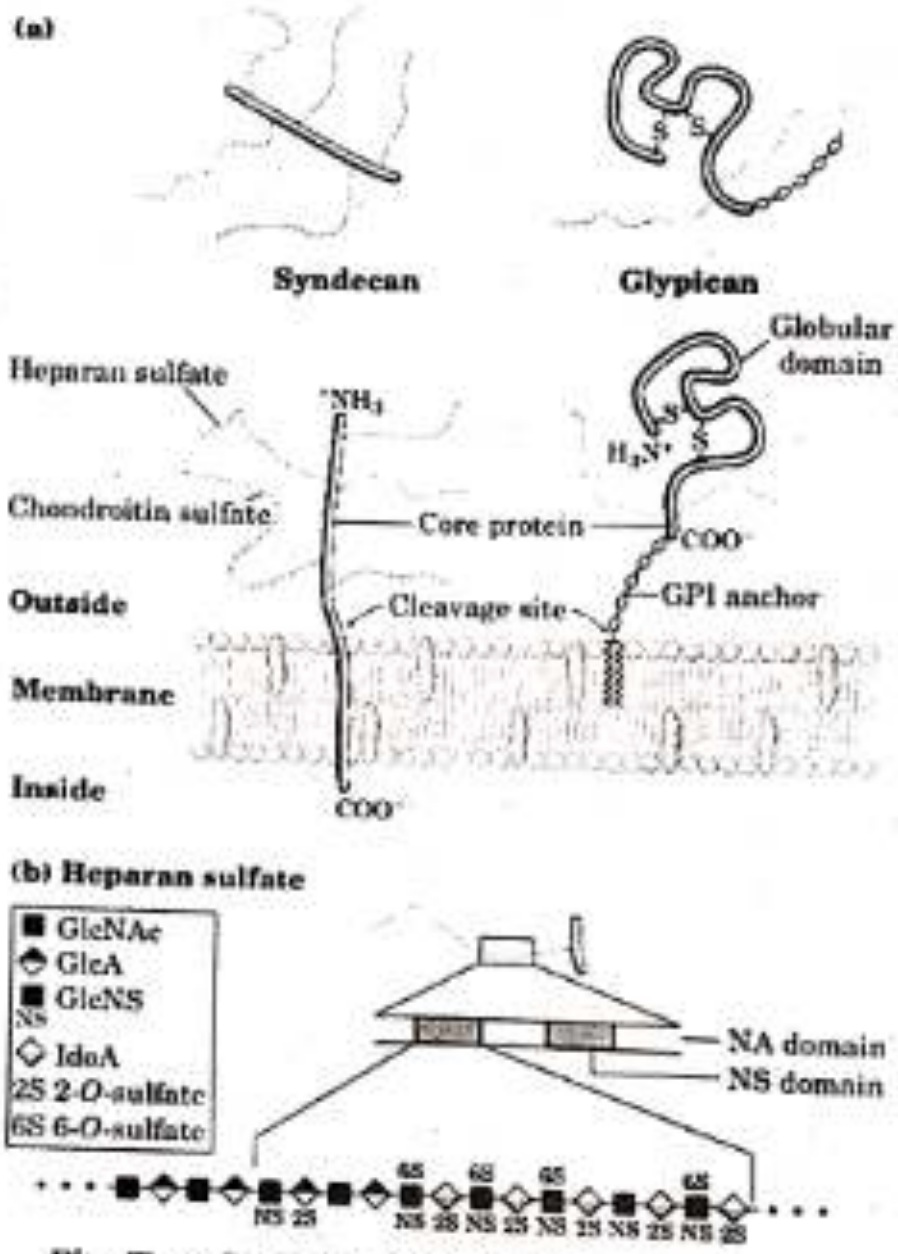


Fig: Two families of membrane proteoglycans

1. Protein core

Many proteoglycans are secreted into the extracellular matrix. But some are integral membrane proteins. For example, sheet-like extracellular matrix (**basal lamina**) separates organized groups of cells from other groups. This matrix contains a family of core proteins (Mr 20,000 to 40,000). Each protein is with several covalently attached heparan sulfate chains. There are two major families of membrane **heparan sulfate proteoglycans**.

a) **Syndecans:** These have a single transmembrane domain and an extracellular domain. This domain bears three to five chains of heparan sulfate and in some cases chondroitin sulfate.

b) **Glypicans:** These are attached to membrane by a lipid anchor. It is a derivative of the membrane lipid phosphatidylinositol.

2. Glycosaminoglycan chains

The glycosaminoglycan chains can bind to a variety of extracellular ligands. Hence they modulate the ligands' interaction with specific receptors of the cell surface. Detailed studies of heparan sulfate demonstrate a domain structure. It is not random. Some domains (typically 3 to 8 disaccharide units long) differ from neighboring domains in sequence and in ability to bind to specific proteins. Highly sulfated domains (called 'NS' domains) alternate with domains having unmodified GlcNAc and GlcA residues (N-acetylated, or NA, domains). The exact pattern of sulfation in NS domain depends on the particular proteoglycan. It gives the number of possible modifications of the GlcNAc–IdoA dimer. Thus at least 32 different disaccharide units are possible.

Glycoproteins

Glycoproteins have covalently attached oligosaccharides. Glycoproteins are carbohydrate-protein conjugates. In glycoproteins, glycans are smaller, branched, and more structurally diverse than glycosaminoglycans of proteoglycans. The carbohydrate is attached at its anomeric carbon. It is attached through:

a) **O-linked:** In this case, glycosidic link to –OH of a Ser or Thr residue.

b) **N-linked:** In this case, an N-glycosyl link to the amide nitrogen of an Asn residue.

Some glycoproteins have a single oligosaccharide chain. But many have more than one chain. The carbohydrate may constitute from 1% to 70% or more of the glycoprotein by mass.

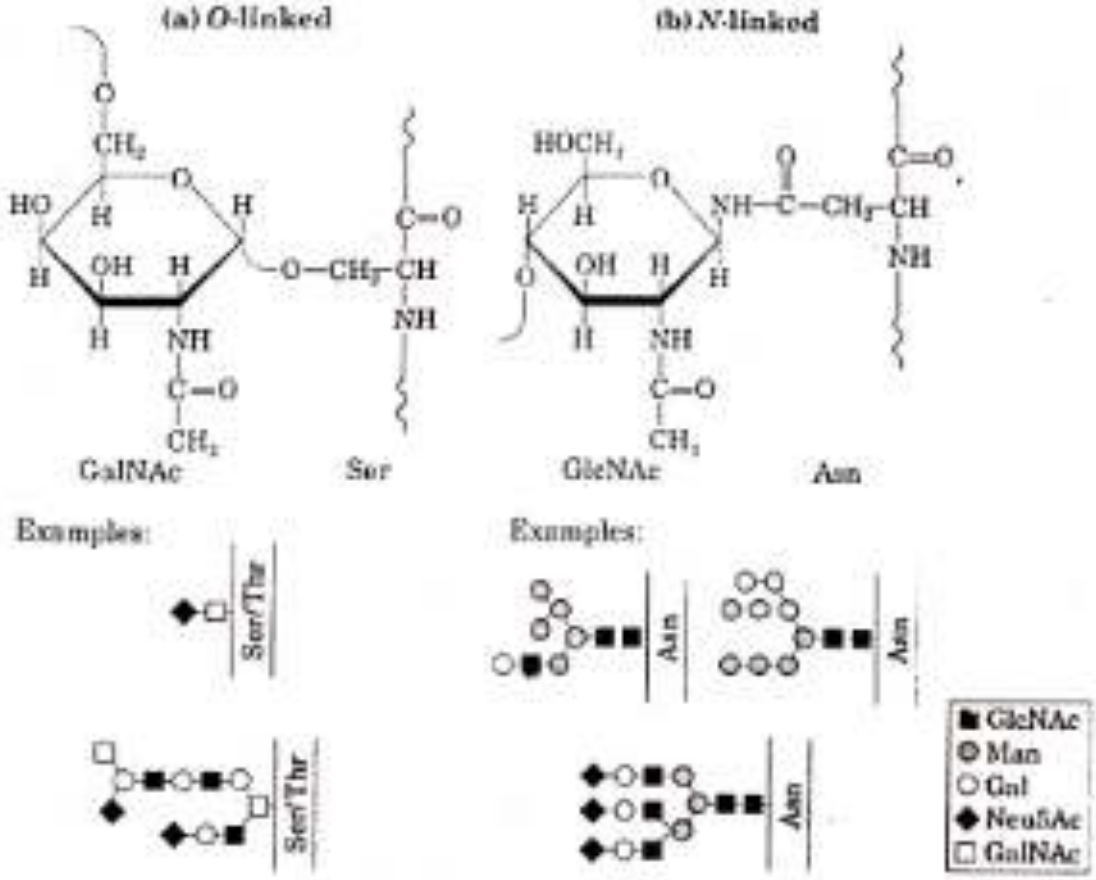


Fig: Oligosaccharide linkages in glycoproteins. (a) O-linked oligosaccharides (b) N-linked oligosaccharides

Different types of glycoproteins are found in cell:

1. **Mucins:** Mucins are secreted or membrane glycoproteins. It can contain large numbers of O-linked oligosaccharide chains. Mucins are present in most secretions. Mucins give mucus its characteristic slipperiness.
2. **Regulation of protein activity:** One class of glycoproteins found in the cytoplasm and nucleus. In this glycoproteins, glycosylated positions in the protein carry only single residues of N-acetylglucosamine in O-glycosidic linkage to hydroxyl group of Ser side chains. This modification is reversible. It often occurs on the same Ser residues. These residues are phosphorylated at some stage in the protein's activity. This type of glycosylation is important in the regulation of protein activity.
3. **Surface membrane glycoproteins:** The external surface of the plasma membrane has many membrane glycoproteins. These glycoproteins have covalently attached oligosaccharides of varying complexity. The first well-characterized membrane glycoprotein was **glycophorin A** of erythrocyte membrane. It contains 60% carbohydrate by mass, in the form of 16 oligosaccharide chains (totaling 60 to 70 monosaccharide residues). These oligosaccharides are covalently attached to amino acid residues near amino terminus of polypeptide chain. Fifteen of the oligosaccharide chains are O-linked to Ser or Thr residues. One of these is N-linked to an Asn residue.
4. **Secretions:** Many of the proteins secreted by eukaryotic cells are glycoproteins. These include most of the proteins of blood. For example, immunoglobulins (antibodies) and certain hormones, like follicle-stimulating hormone, luteinizing hormone, and thyroid-stimulating hormone are glycoproteins.
5. **Milk proteins:** Many milk proteins, including lactalbumin are glycoproteins
6. **Glycoproteins secreted by pancreases:** Some of the proteins secreted by the pancreas (like ribonuclease) are glycosylated. Such proteins are found in lysosomes.
7. **Role in genetic disorders:** The importance of normal protein glycosylation is clear from the finding of at least 18 different genetic disorders of glycosylation in humans. All of thee severely cause defective physical or mental development. Some of these disorders are fatal.

Glycolipids and Lipopolysaccharides

Both of these are membrane components.

1. **Gangliosides** are membrane lipids of eukaryotic cells. In gangliosides, polar head group is a complex oligosaccharide. This head group contains a sialic acid and other monosaccharide residues. Polar head group is a part of lipid that forms the outer surface of the membrane.
2. Some of the oligosaccharide moieties of gangliosides determine human blood groups. These are identical with those found in certain glycoproteins. Therefore, these also contribute to blood group type.
3. These membrane lipids are generally found on the outer face of the plasma membrane.

4. Lipopolysaccharides are the dominant surface feature of outer membrane of gram-negative bacteria.

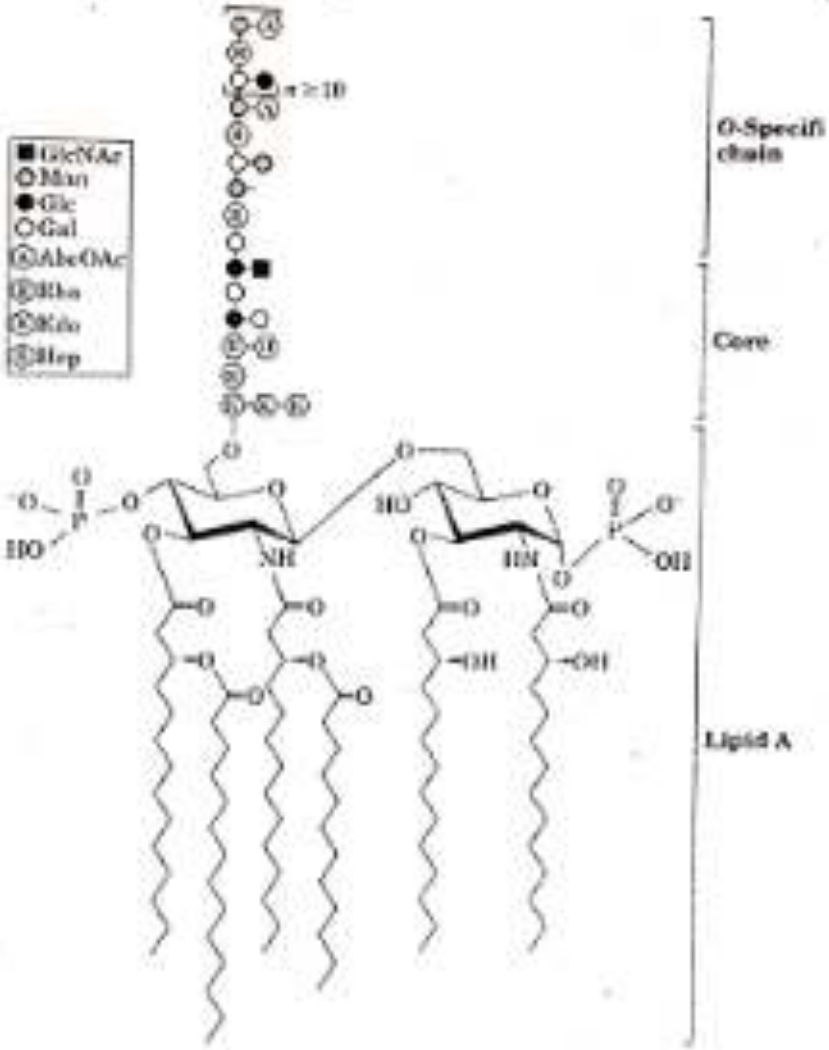


Fig: Bacterial lipopolysaccharides

CARBOHYDRATES AS AN INFORMATION CARRIER MOLECULE

Glycobiology is the study of the structure and function of glycoconjugates. It is one of the most important areas of biochemistry and cell biology. The cells use specific oligosaccharides to encode important information. This information is about intracellular targeting of proteins, cell-cell interactions, cell differentiation and tissue development and xtracellular signals.

'versity found in oligosaccharides of glycoproteins and glycolipids

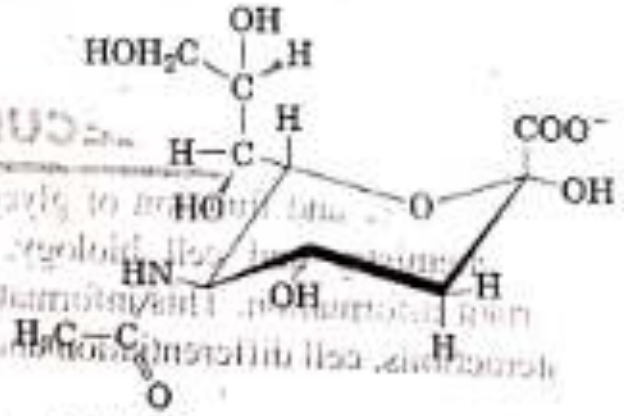
narkable complexity and diversity is found in oligosaccharides of glycoproteins and olipids. Some oligosaccharide chains are found in many glycoproteins. The most lex of those contains 14 monosaccharide residues of four different kinds. They are ly linked as (1→2), (1→3), (1→4), (1→6), (2→3), and (2→6). Some are linked and some with β configuration. Branched structures are common in harides. Suppose we have 20 different monosaccharide subunits are available

for construction of oligosaccharides. From this we can calculate that many billions of different **hexameric** oligosaccharides are possible. Thus 6.4×10^7 (20^6) different hexapeptides are possible with 20 common amino acids. Similarly, 4,096 (4^6) different hexanucleotides are possible with the four nucleotide subunits. Variations in oligosaccharides may occur from sulfation of one or more residues. Hence the number of possible oligosaccharides increases by two orders of magnitude. In reality, only a subset of possible combinations is found. Each of the oligosaccharides represents a unique, three-dimensional face. It is a word in the sugar code. It is readable by the proteins that interact with it.

Lectins

Lectins are proteins that read sugar code and mediate many biological processes. Lectins are found in all organisms. Lectins are proteins that bind carbohydrates with high specificity and with moderate to high affinity. Lectins perform different functions of cell-cell recognition, signaling, and adhesion processes and in intracellular targeting of newly synthesized proteins.

1. Plant lectins are abundant in seeds. It acts as deterrents to insects and other predators.
2. In the laboratory, purified plant lectins are useful reagents for detecting and separating glycans and glycoproteins with different oligosaccharide moieties.
3. Some peptides hormones circulate in the blood have oligosaccharide moieties. This oligosaccharide strongly influences their circulatory half-life. Luteinizing hormone and thyrotropin are polypeptide hormones. These are produced in the adrenal cortex. They have N-linked oligosaccharides that end with the disaccharide GalNAc4S (β1→4)GlcNAc. It is recognized by a lectin (receptor) of hepatocytes. Receptor-hormone interaction mediates the uptake and destruction of luteinizing hormone and thyrotropin. It reduces their concentration in the blood. Thus the blood levels of these hormones undergo a periodic rise and fall.



N-Acetylneuraminic acid (Neu5Ac) (a sialic acid)

4. The residues of Neu5Ac (a sialic acid) situated at the ends of the oligosaccharide chains of many plasma glycoproteins. It protects those proteins from uptake and degradation in the liver. For example, ceruloplasmin is a copper-containing serum glycoprotein. It has several oligosaccharide chains ending in Neu5Ac. The plasma

membrane of hepatocytes has lectin molecules. It specifically binds oligosaccharide chains with galactose residues no longer protected by a terminal Neu5Ac residue.

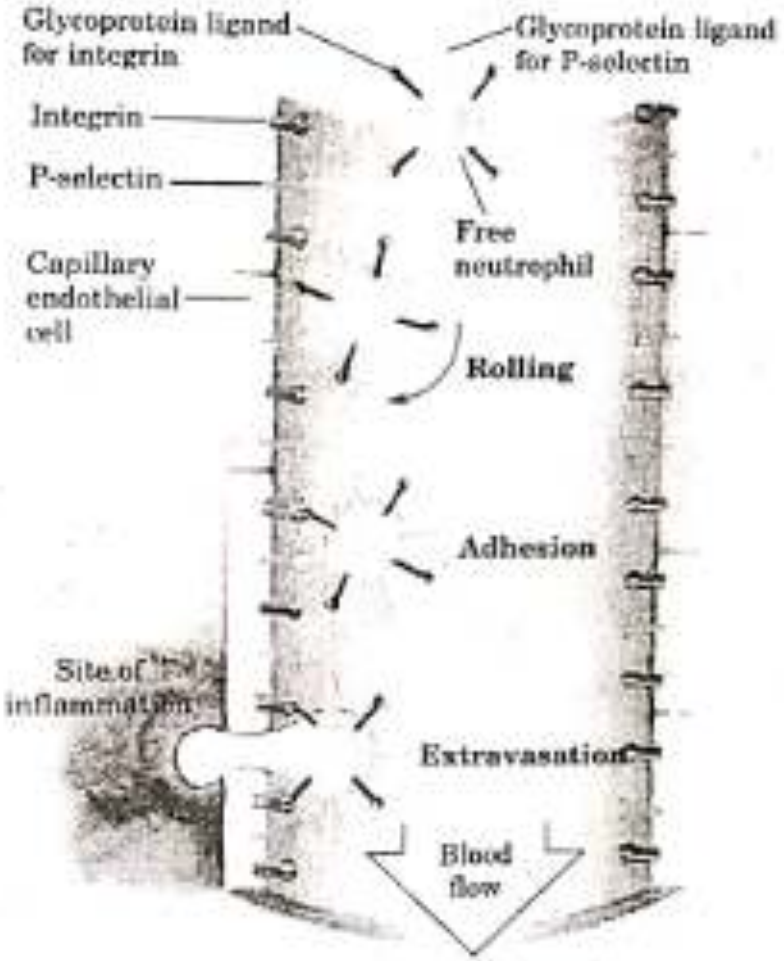


Fig: Role of lectin-ligand interactions in lymphocyte movement

5. A neutrophil circulate through a capillary. It is slowed by transient interactions between P-selectin molecules in plasma membrane of capillary endothelial cells and glycoprotein ligands for P-selectin on the neutrophil surface. As a result, the neutrophil rolls along capillary surface. Near a site of inflammation, stronger interactions between integrin in capillary surface and its ligand in the neutrophil surface lead to tight adhesion. The neutrophil stops rolling and, under the influence of signals sent out from the site of inflammation. It escapes through the capillary wall as it moves toward the site of inflammation.
6. Receptor-ceruloplasmin interaction triggers endocytosis and destruction of the ceruloplasmin.
7. A similar mechanism is responsible for removing old erythrocytes from the mammalian bloodstream. Newly synthesized erythrocytes have several membrane glycoproteins with oligosaccharide chains. These chains end in Neu5Ac. An enzyme sialidase was treated with erythrocytes. These erythrocytes disappear from the bloodstream within a few hours. Erythrocytes with intact oligosaccharides continue to circulate for days.

8. Other lectins act in other kinds of protein sorting. Any newly synthesized protein in the endoplasmic reticulum already has a complex oligosaccharide attached. It can be bound by either of two ER lectins that are also chaperones: **calnexin** (membrane-bound) or **calreticulin** (soluble). These lectins link the new protein with an enzyme. This enzyme brings about rapid disulfide exchange as the protein tries various ways to fold. It forms native conformation. At this point, enzymes in the ER trim the oligosaccharide moiety to a form recognized by another lectin, ERGIC53. This lectin draws the folded protein (glycoprotein) into the Golgi complex for further maturation. If a protein has not folded effectively, the oligosaccharide is trimmed to another form. This form is recognized by a lectin, EDEM. The EDEM initiates movement of the defectively folded protein into the cytosol. Here it will be degraded.

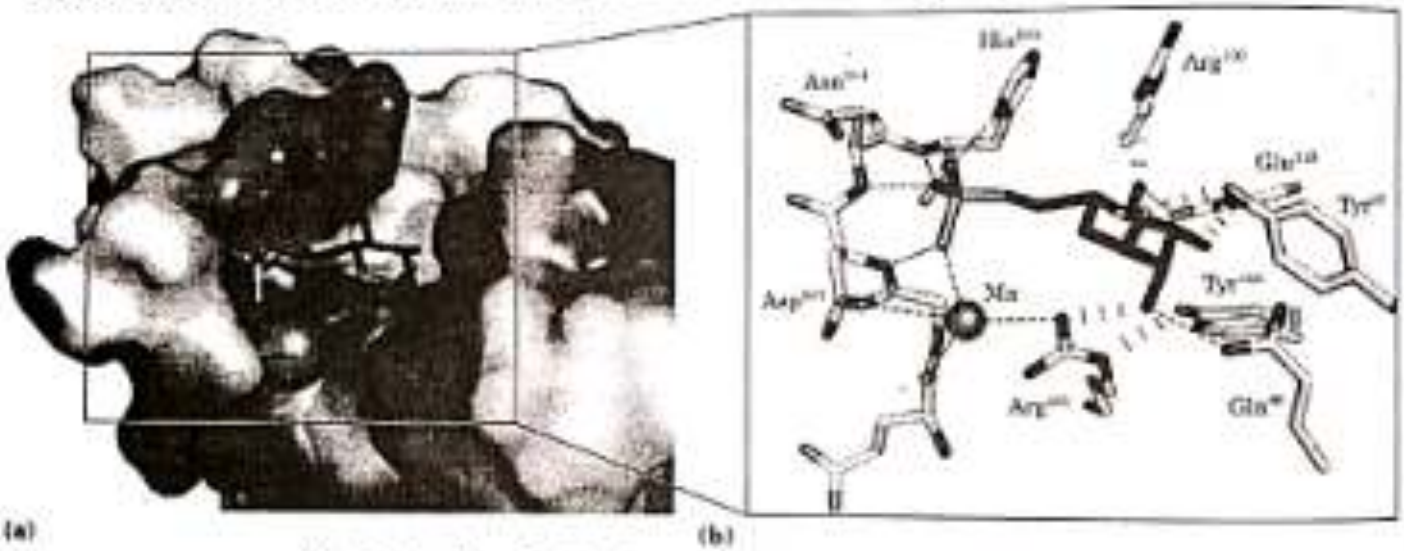


Fig: Details of a lectin-carbohydrate interaction

Lectin-Carbohydrate Interactions

Lectin-carbohydrate interactions are highly specific and often polyvalent.

Formation of interactions

The oligosaccharide should have unique structure in lectins for all the functions of lectins. Therefore, recognition by lectin is highly specific.

1. The high density of information in oligosaccharides provides a sugar code. It has an unlimited number of unique words. These words are small enough to be read by a single protein.
2. In their carbohydrate-binding sites, lectins have a unique molecular complementarity. It allows interaction only with their correct carbohydrate cognates. The result is an extraordinarily high specificity in these interactions.
3. The affinity between an oligosaccharide and each carbohydrate binding domain (CBD) of a lectin is modest (micromolar to millimolar Kd values). But effective affinity is in many cases greatly increased by lectin multivalency. In multivalency, a single lectin molecule has multiple CBDs.

4. Some cluster of oligosaccharides is commonly found on a membrane surface. In these clusters, each oligosaccharide can engage one of the lectin's CBDs. It strengthens the interaction.
5. When cells express multiple receptors, the chance of the interaction can be very high. It enables highly cooperative events like cell attachment and rolling.

Lectin-sugar interaction

X-ray crystallographic studies were done for structures of several lectin-carbohydrate complexes. These studies gave the details of lectin-sugar interaction. In humans, a family of 11 lectins binds to oligosaccharide chains ending in sialic acid residues. This family plays some important biological roles. All of these lectins bind sialic acids at β sandwich domains like found in immunoglobulins. Therefore, proteins are called **siglecs 1 to 11**(sialic acid-recognizing Ig-superfamily lectins), or sometimes **sialoadhesins**. The interaction of a siglec with sialic acid (Neu5Ac) involves each of ring substituents unique to Neu5Ac:

a) The acetyl group at C-5 undergoes both hydrogenbond and van der Waals interactions with the protein.
b) The carboxyl group makes a salt bridge with a conserved Arg residue;
c) The hydroxyls of glycerol moiety hydrogen-bond with the protein.

Siglecs regulate activities in the immune and nervous systems and in blood cell development. For example Siglec-7 binds to a specific ganglioside (GD3) containing two sialic acid residues. It suppresses the activity of NK (natural killer) cells in the immune system. Thus it saves the cells targeted for immune destruction from NK killing activity. There is elevation of GD3 levels in tumors like malignant melanoma and neuroblastoma. This elevation may be a mechanism for evading protective action of immune system.

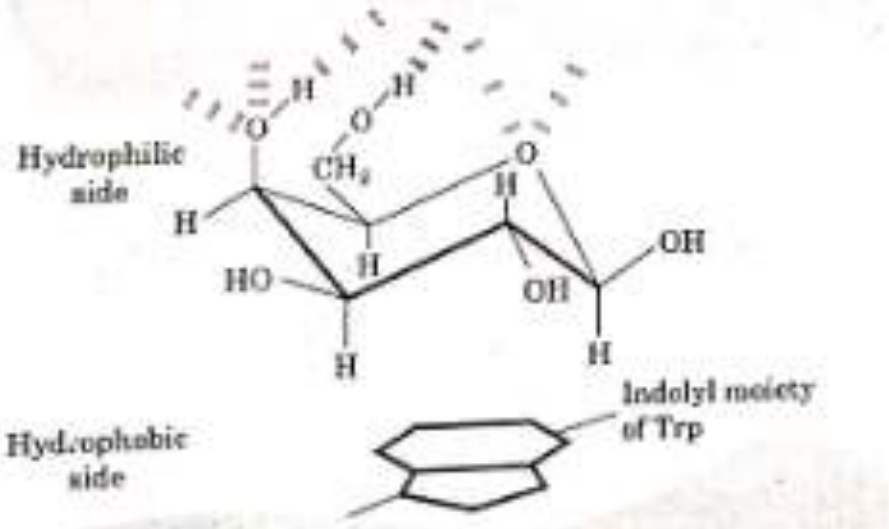


Fig: Hydrophobic interactions of sugar residues

Specific interactions

The structure of mannose 6-phosphate receptor/lectin reveals details of its interaction with mannose 6-phosphate. It explains the specificity of the binding and role for a

divalent cation in the lectin-sugar interaction. His^{105} is hydrogen-bonded to one of oxygen atoms of the phosphate. The protein tagged with mannose 6-phosphate reaches the lysosome (which has a lower internal pH than Golgi complex). Then the receptor loses its affinity for mannose 6-phosphate. Protonation of His105 may be responsible for this change in binding.

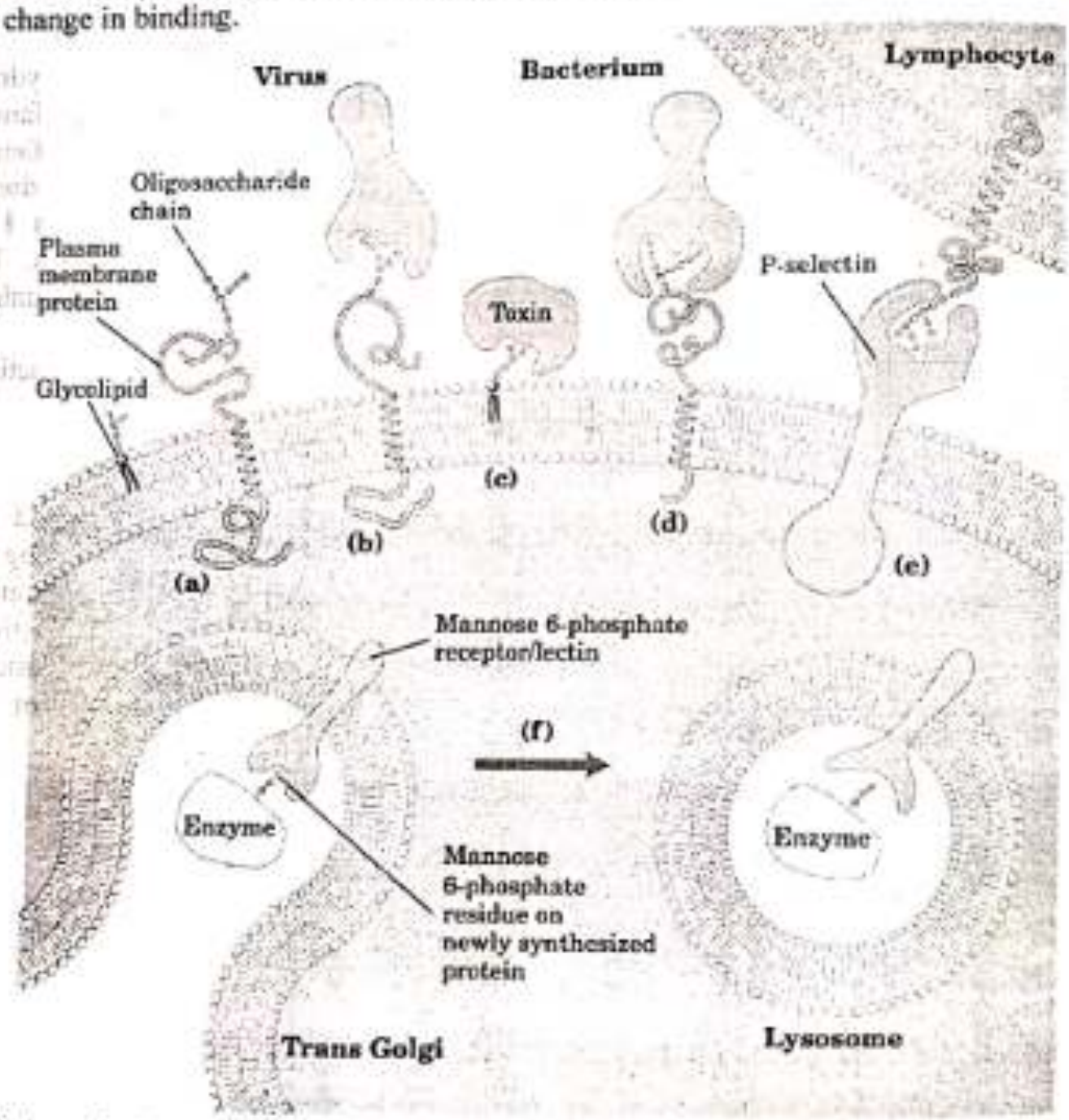


Fig: Roles of oligosaccharides in recognition and adhesion at the cell surface. (a) Oligosaccharides with unique structures (b) Viruses that infect animal cells (c) Bacterial toxins bind to a surface glycolipid before entering a cell. (d) Some bacteria, such as H. pylori, adhere to and then colonize or infect animal cells. (e) Selectins (lectins) in the plasma membrane of certain cells mediate cell-cell interactions, (f) The mannose 6-phosphate receptor/lectin of the trans Golgi complex binds to oligosaccharide of lysosomal enzymes, targeting them for transfer into the lysosome.

General interactions

There are more general interactions in addition to these very specific interactions. These general interactions contribute to binding of many carbohydrates to their lectins. For example, many sugars have a more polar and a less polar side. The more polar side form hydrogen-bonds with the lectin. The less polar undergoes hydrophobic interactions with nonpolar amino acid residues. The sum of all these interactions produces high-affinity binding and high specificity of lectins for their carbohydrates. This represents a kind of information transfer. This transfer is clearly central in many processes within and between cells. (Figure summarizes some of the biological interactions mediated by the sugar code.)

EXERCISE

LONG QUESTIONS

1. **What are carbohydrates? Write its definitions. Illustrate their importance.**
2. **Write note on (i) ribose (ii) Stereoisomerism**
3. **Discuss chemical structure of monosaccharides.**
4. **Demonstrate Stereoisomerism in monosaccharides.**
5. **Explain cyclic structure of monosaccharides.**
6. **Write note on: (a) Mutarotation (b) Haworth perspective formulas**
7. **What are disaccharides? Explain their different types with examples.**
8. **What is lactose? Demonstrate its formation.**
9. **Write note on polysaccharides.**
10. **What are polysaccharides? Write note on: (a) starch (b) cellulose (c) chitin**
11. **What is starch? Describe its structure and types.**

SHORT QUESTIONS

1. **What are carbohydrates? Discuss its old and new definitions.**
2. **Why are carbohydrates called as hydrates carbon? Is there any problem with this old concept?**
3. **Give comparison of characteristics of monosaccharides, oligosaccharides and polysaccharides.**
4. **What are monosaccharides? On which basis are they classified?**
5. **Write formulae of heptoses and Glucoheptose. Give their functions.**
6. **Differentiate between furanose and pyranose.**

Furanose	Pyranose
1. Furanose is a five membered ring.	1. Pyranose is a six membered ring.
2. This ring has one oxygen atom and four carbon atoms.	2. This ring has one oxygen atom and five carbon atoms.
3. Here oxygen atom is linked with C_1 and C_4.	3. Oxygen atom is linked with C_1 and C_5.
4. All pentoses and ketohexoses are converted into furanose ring.	4. Only aldohexoses are converted into pyranose ring.

7. What are Fischer projection and Haworth projection? Differentiate between α sugar and β- sugars.

8. Define stereoisomerism. How we can calculate stereoisomers?

9. Explain: (a) enantiomers (b) diastereoisomers (c) epimers

10. What are laboratory manufactured and naturally occurring sugars? How are artificial sweeter not consumed in our body?

11. What are disaccharides? Write their two examples.

12. Write note on Lactose and maltose.

13. What are polysaccharides? Differentiate between homopolysaccharides and heteropolysaccharides. Give examples.

14. What is starch? Differentiate between amylose and amylopectin starch.

Amylose starch	Amylopectin starch
1. Amylose starch have unbranched chain of glucose.	1. Amylopectin starch has branched chains.
2. It has 1- 4 glycosidic bonding.	2. It has 1- 4 and 1- 6 glycosidic bonding
3. It is soluble in hot water.	3. It is insoluble in hot or cold water.
4. It is most abundant starch.	4. It is less abundant

14. Write note on chitin.

MULTIPLE CHOICE QUESTIONS (MCQs)

1. Which of the following atoms do not occur in carbohydrates?
(a) Carbon (b) Hydrogen (c) Oxygen (d) Nitrogen
2. Monosaccharides which are rare in nature and occur in some bacteria are:
(a) Trios (b) Tetrose (c) Pentose (d) Hexose
3. Most of the monosaccharide form ring structure in:
(a) water (b) solution (c) solvent (d) stomach
4. Covalent bond between monosaccharide is:
(a) glycosidic bond (b) peptide bond (c) ester bond (d) hydrogen bond
5. Which one of following is not a polysaccharide?
(a) glycogen (b) lactose (c) starch (d) chitin
6. The most abundant carbohydrate in nature is:
(a) starch (b) cellulose (c) glycogen (d) agar
7. Starch with iodine gives colour:

(a) blue (b) red (c) green (d) yellow

8. Which of the followings is a polysaccharide?
(a) chitin (b) cutin (c) lactose (d) fructose

9. Oxygen in furanose are linked by bond:
(a) C1-C2 (b) C1-C3 (c) C2-C4 (d) C1-C4

10. Which of the following carbon is not asymmetric in glucose:
(a) C1 (b) C2 (c) C3 (d) C4

11. Number of enantiomers of glucoses are:
(a) 6 (b) 8 (c) 16 (d) 12

12. Number of stereoismers of glucoses are:
(a) 6 (b) 8 (c) 16 (d) 12

13. The stereoisomers which are non-superimposable mirror images of one another are called
(a) Enantiomers (b) Diastereoisomers (c) Epimers (d) none

14. The stereolsomers which have different arrangement of –H and –OH groups at more than one asymmetrical carbon atoms are called:
(a) Enantiomers (b) Diastereoisomers (c) Epimers (d) none

15. Artificial sweeteners are:
(a) L- form (b) D form (c) both D and L form (d) not isomers

16. Sucrose gives on hydrolysis two molecules:
(a) α-glucose and α-fructose (b) α-glucose and β-glucose
(c) α-glucose and β-fructose (d) β-glucose and β-fructose

17. The sugar found in germinating seeds is:
(a) sucrose (b) maltose (c) lactose (d) fructose

18. Which of the followings is milk sugar?
(a) sucrose (b) maltose (c) lactose (d) fructose

19. Which of the followings is not Homopolysaccharides?
(a) starch (b) chitin (c) pectin (d) glycogen

20. Which of the followings is animal starch?
(a) starch (b) chitin (c) pectin (d) glycogen

21. The fungal cellulose is:
(a) starch (b) chitin (c) pectin (d) glycogen

Ans:

1. (d) 2. (b) 3. (a) 4. (a) 5. (b) 6. (b) 7. (a) 8. (a) 9. (d) 10. (a) 11. (b) 12. (d) 13. (a) 14. (b) 15. (a) 16. (c) 17. (b) 18. (c) 19. (c) 20. (d) 21. (b)

Chapter 5

LIPIDS

Syllabus 2012-13
Syllabus Chapter 4: Lipids:
Fatty acids, their types and major characteristics; storage lipids, acylglycerols; waxes; structural lipids in membranes; major functions of lipids; lipoproteins, their types and major functions.
Syllabus 2018-19
4. Lipids: a. Introduction & classification of lipids; Fatty acids, their types; Storage lipids; Classification and important characteristics;
a. Triacylglycerols; waxes Structural/membrane lipids; Glycerophospholipids with Ether and Ester linkages Galactolipids & Sulfolipds: **Sphingolipids their types & importance: Sterols, their structure, types & functions. Examples of Functional diversity of Lipids as Signaling molecules, Cofactors, Electron carrier, antioxidants, pigments etc.**

LIPIDS

Lipid is defined as a heterogeneous (different nature) group of organic compounds insoluble in water (hydrophobic) but soluble in organic solvent like acetone, alcohol, and ether etc.

- Lipid is collective name for different organic compounds found in body. These compounds are like fats, oils, waxes and fat-like molecules (steroids).
- Lipids are composed of carbon, hydrogen and oxygen like carbohydrates. But they have less proportion of oxygen to carbon and hydrogen than in carbohydrates.

- For example, **tristearin** is a simple lipid. Its molecular formula is $C_{57}H_{110}O_6$. It has high contents of carbon and hydrogen. Therefore, they contain double amount of energy than carbohydrates.

Functions or importance of Lipids

1. Lipids are components of cell membranes (phospholipids and cholesterol).
2. They act as energy stores (triglycerides)
3. They act as chemical messengers (steroid)
4. They are involved in protection, waterproofing, insulation and buoyancy.

Classification of lipids

Lipids are classified into simple, complex and derived lipids. They are further subdivided into different groups.

a) **Simple lipids:** These are esters of fatty acids with various alcohols. Its examples are acylglycerols (fats and oils) and waxes.

b) **Compound/complex lipids:** These contain other groups in addition to an alcohol and fatty acid. These groups are like phospholipids, glycolipids and lipoproteins.

c) **Derived lipids:** Derived lipids are derivatives of simple and complex lipids. Its examples are terpenes, steroids, prostaglandins and cholesterols.

FATTY ACIDS

A fatty acid is an organic acid containing one carboxylic acid group at attached to a hydrocarbon.

Structure of fatty acids

1. Fatty acids contain even number of carbons from C4 to C36. Each fatty acid is represented as R-COOH. Here R is a hydrocarbon tail.
2. The most commonly occurring fatty acids have even numbers of carbon atoms. These atoms are present in an unbranched chain of 12 to 24 carbons. The even number of carbons results from the mode of synthesis of these compounds. This mode involves condensation of two-carbon (acetate) units.
3. There is also a common pattern in the location of double bonds. In most monounsaturated fatty acids the double bond is between C-9 and C-10 (9). The double bonds of polyunsaturated fatty acids are generally 12 and 15. (except in arachidonic acid) The double bonds of polyunsaturated fatty acids are almost never conjugated (alternating single and double bonds, as in –CH=CH–CH=CH–). But these are separated by a methylene group: –CH=CH–CH_2–CH=CH–.
4. In all naturally occurring unsaturated fatty acids, double bonds are in the cis configuration. Trans fatty acids are produced by fermentation in the rumen of dairy animals. These are obtained from dairy products and meat. They are also produced during hydrogenation of fish or vegetable oils.

Diets high in trans fatty acids correlate with increased blood levels of LDL (bad cholesterol) and decreased HDL (good cholesterol). Therefore, it is recommended that we

should avoid large amounts of these fatty acids. French fries, dough- nuts, and cookies are high in trans fatty acids.

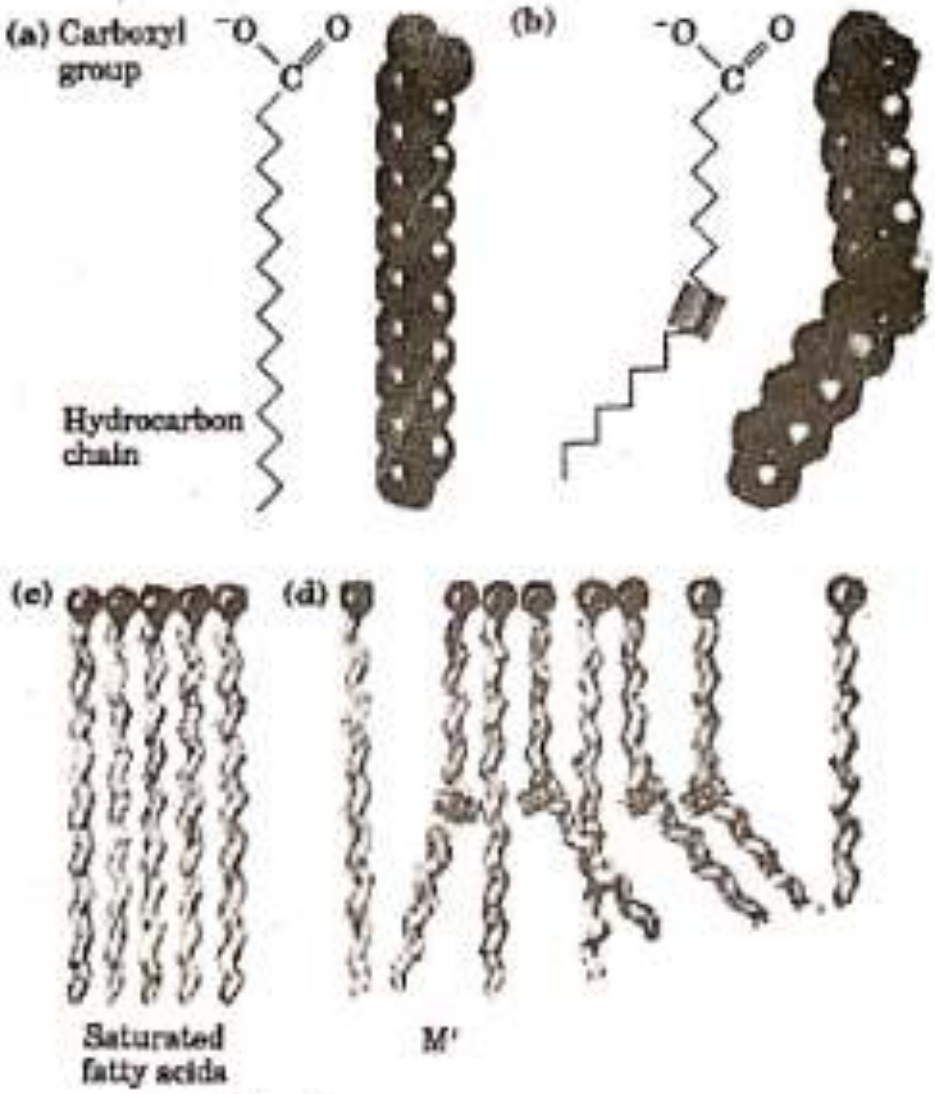


Fig: c

Nomenclature

a) A simplified nomencl
number of double b
- 16-carbon sat
- 18-carbon
- with

b) The posit
(delta). For
10 (C-1 is
designated as 20

Properties of fatty
Length and degree of u
of fatty acids.

1. There are about 30 differe

2. Fatty acids vary in length.
   - Acetic acid (2C) and butyric acid (4C) are simplest fatty acid.
   - Palmitic acid (16C) and stearic acid (18C) are most common fatty acids.
3. Some properties of fatty acid are increased with an increase in number of carbon atoms. These properties are like melting point, solubility in organic solvent and hydrophobic nature.
4. The non-polar hydrocarbon chain cause poor solubility of fatty acids in water. For example, lauric acid (12:0), has a solubility in water of 0.063 mg/g. It is much less than that of glucose (Mr 180). Glucose has solubility 1,100 mg/g. The longer the fatty acyl chain and the fewer is the double bonds. Thus lower is the solubility in water. The carboxylic acid group is polar. It is responsible for slight solubility of short-chain fatty acids in water.
5. Melting points are also strongly influenced by the length and degree of unsaturation of hydrocarbon chain. At room temperature (25° C), saturated fatty acids from 12:0 to 24:0 have a waxy consistency. But unsaturated fatty acids of these lengths are oily liquids. This difference in melting points is due to different degrees of packing of the fatty acid molecules.
6. In the fully saturated compounds, there is free rotation around each carbon-carbon bond. It gives the hydrocarbon chain great flexibility. The most stable conformation is the fully extended form. In this form, steric hindrance of neighboring atoms is minimized. Thus these molecules can pack together tightly in crystalline form. Here atoms all along their lengths in van der Waals contact with the atoms of neighboring molecules. In unsaturated fatty acids, a cis double bond forces a kink in the hydrocarbon chain. Fatty acids with one or several such kinks cannot pack together. Therefore, their interactions with each other are weaker. It takes less thermal energy to disorder these poorly ordered arrays of unsaturated fatty acids. Therefore, they have markedly lower melting points than saturated fatty acids of the same chain length.

Types of fatty acids

In some fatty acids, this chain is unbranched and fully saturated (contains no double bonds).In others the chain contains one or more double bonds. A few contain three-carbon rings, hydroxyl groups, or methyl- group branches. Thus fatty acids may be saturated or unsaturated.

a) **Saturated fatty acids: Fatty acids in which all of the internal carbon atoms possess hydrogen side groups are called saturated fatty acids.** They contain maximum possible number of hydrogen atoms. Its example is palmitic acid.
   - Saturated fatty acids are solid at room temperature. Thus they have higher melting point.
   - They are more common in animal lipids (fats).

b) **Unsaturated fatty acids: They have one or more pairs of carbon atoms joined by a double bond.** Therefore, they are not fully saturated with hydrogen. Its example is oleic acid.

- Unsaturated fatty acids are liquid at room temperature. Thus they have lower melting point.
- They are more common in plant lip (oils).

Triglycerides contain hydrocarbon chains. It melts at a low temperature. This is useful for living things.

Table: Common types of fatty acids

Name	Sources	No. of carbon	Condensed Formula	Melting point (C^{o})
Saturated fatty acids				
1. Lauric	Coconut oil	12	$OH_3(CH_2)_{10}COOH$	44
2. Myristic	Butter fat	14	$OH_3(CH_2)_{12}COOH$	58
3. Palmitic	Most fats and oils	16	$OH_3(CH_2)_{14}COOH$	63
4.Stearic	Most fats and oils	18	$OH_3(CH_2)_{16}COOH$	70
Unsaturated fatty acids				
5.Oleic	Olive oil	18	$OH_3(OH_2)_7CH = CH(CH_2)_7COOH$	4
6.Linoleic	Vegetable oils	15	$CH_3(CH_2)_4CH=CHCH_2CH=CH(CH_2)_7COOH$	–5
7.Linolenic	Soybeans and canola oils	18	$CH_3CH_2CH=CHCH_2CH=CH(CH_2)_7CHOH$	41
8.Arachidonic	Lard (present in chicken etc.)	20	$CH_3(CH_2)_4CH=CHCH_2CH=CH-CH_2CH=CHCHCH_2=CH(CH_2)_3COOH$	–50

ACYLEGLYCEROL

Acylglycerol is most abundant lipids in living things. **Acylglycerols is defined as esters of glycerol and fatty acids.**

Ester and esterification

An ester is produced as result of a chemical reaction of an alcohol with acid. A water molecule is released in this reaction. This reaction is called esterification.

a) **Glycerol: Glycerol is a trihydroxy alcohol.** It contains three carbons, each bears an OH group.

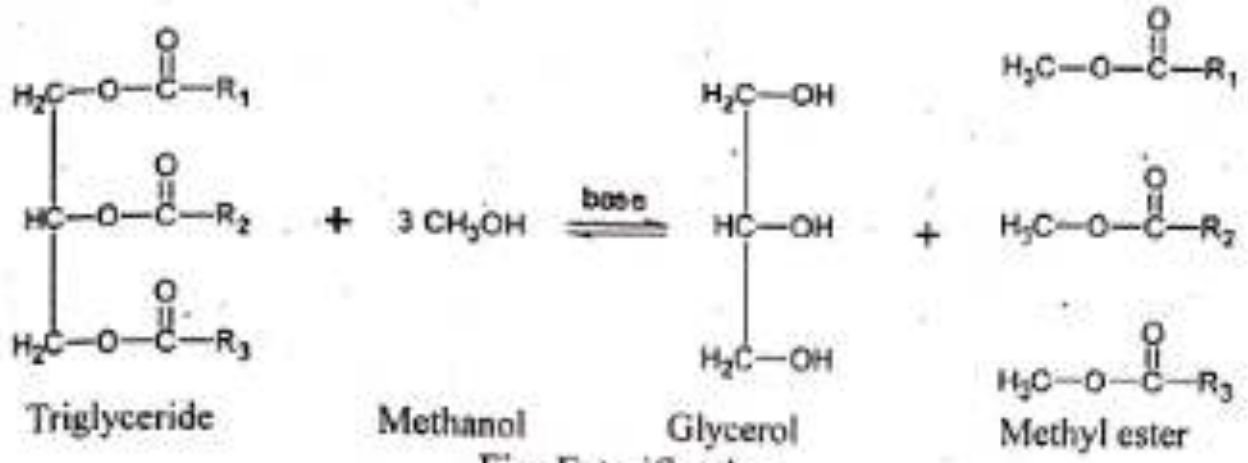


Fig: Esterification

b) **Fatty acid: A fatty acid is an organic acid containing one carboxylic acid group attached to a hydrocarbon.** Fatty acids contain even number of carbons from 2 to 30. Each fatty acid is represented as R-COOH. Here R is a hydrocarbon tail.

Formation of acylglycerol

a) **Monoacylglycerol:** A glycerol molecule combines with one fatty acid. As a result a Monoacylglycerol (monoglyceride) is formed.

b) **Diacylglycerol:** Now two fatty acids combine with a glycerol. As a result, diacylglycerol (diglyceride) is formed.

c) **Triacylglycerols:** Three fatty acids combine with one glycerol molecule. As a result, triacylglycerol (triglyceride) is formed. Triacylglycerols are also called **neutral lipid.** Their all three OH groups of glycerol are occupied by fatty acids. Here no charge bearing OH group is left.

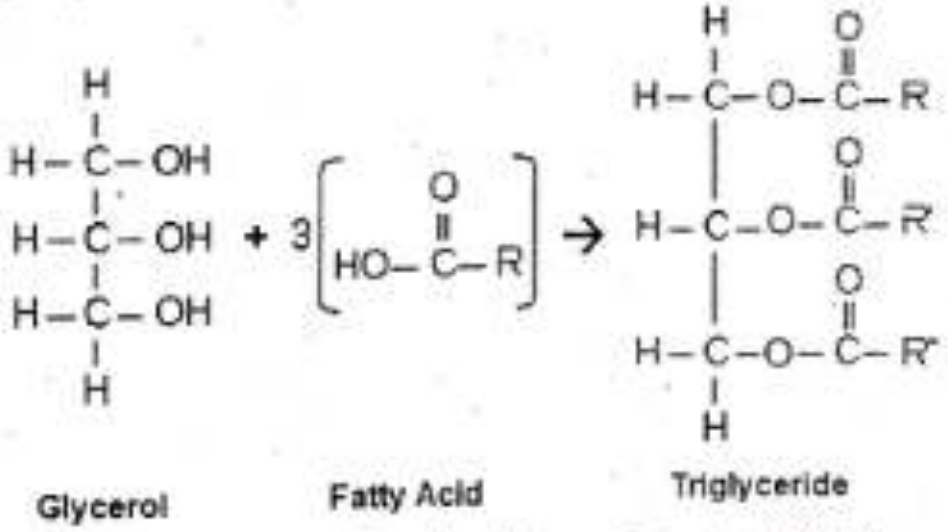


Fig: Formation of a triacylglycerol (neutral lipid)

Storage Lipids

Triacylglycerols Provide Stored Energy and Insulation

In most eukaryotic cells, triacylglycerols form a separate phase of microscopic, oily droplets in the aqueous cytosol. They are used as depots of metabolic fuel. Vertebrates have specialized cells called **adipocytes**, or fat cells. They store large amounts of triacylglycerols as fat droplets. These droplets nearly fill the cell. Triacylglycerols are

also stored as oils in the seeds of many types of plants. They provide energy and biosynthetic precursors during seed germination. Adipocytes and germinating seeds contain lipases enzymes. These enzymes catalyze the hydrolysis of stored triacylglycerols. They release fatty acids for export.

Advantage of fats as storage compounds

There are two advantages for using triacylglycerols as stored fuels than polysaccharides like glycogen and starch.

a) The carbon atoms of fatty acids are more reduced than those of sugars. Therefore, oxidation of triacylglycerols yields more than twice energy than the oxidation of carbohydrates.

b) Triacylglycerols are hydrophobic. Therefore, unhydrated organism carries fat as fuel. They do not have to carry the extra weight of water of hydration associated with stored polysaccharides (2 g per gram of polysaccharide) **in human.**

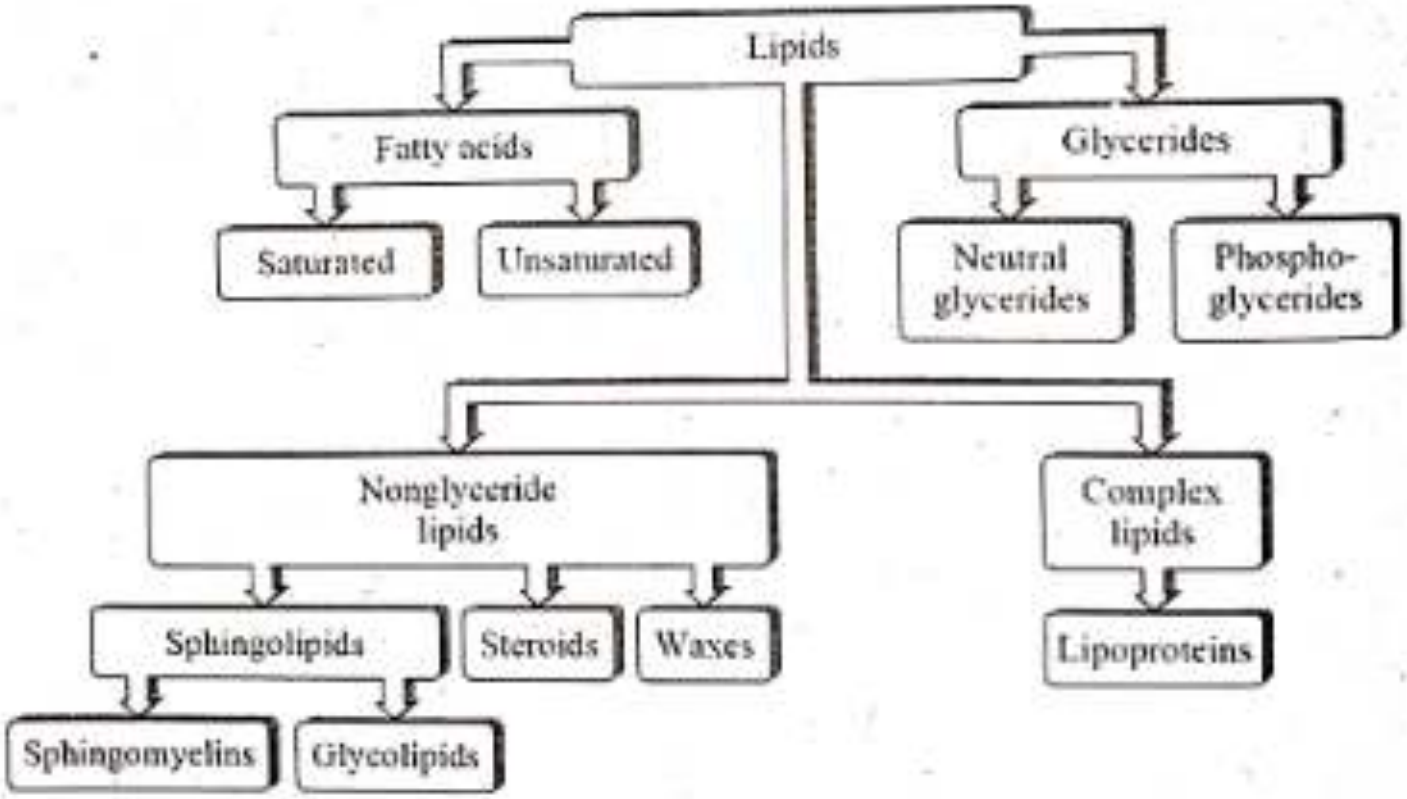


Fig: Some simple types of lipids

Uses of stored fats

1. Stored fats used as energy reserves

Humans have fat tissue composed primarily of adipocytes. They are present under the skin, in the abdominal cavity, and in the mammary glands. Moderately obese people have 15 to 20 kg of triacylglycerols deposited in their adipocytes. It can meet their energy needs for months. But human body can store less than a day's energy supply in the form of glycogen. Carbohydrates like glucose and glycogen have certain advantages. They are quick sources of metabolic energy. They are readily solubility in water.

2. Use as energy reserve and insulating substance

Triacylglycerols are stored under the skin in some animals. They are used as energy stores. They are also used for insulation against low temperatures. Its examples seals, walruses, penguins, and other warm-blooded polar animals. In hibernating animals like bears, the huge fat reserves are stored before hibernation. It gives dual purposes of insulation and energy storage.

3. Role in buoyancy

The low density of triacylglycerols has another important function. Sperm whales have a store of triacylglycerols and waxes. These compounds match the buoyancy of sperms to their surroundings during deep dives in cold water.

WAXES

Biological waxes are esters of long-chain (C_{14} to C_{36}) saturated and unsaturated fatty acids with long-chain (C_{16} to C_{30}) alcohols. Their melting points are 60^0 to 100^0 C. Their melting points are higher than triacylglycerols. Waxes are highly are highly hydrophobic compounds. There are two types of waxes:

(a) Natural waxes

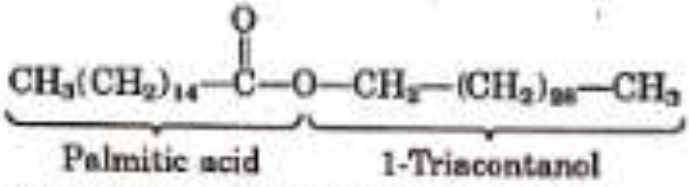
These are simple waxes. They are ester of long chain fatty acids and long chain alcohols. These are like:

- **Bee's wax:** It is found in honeycomb
- **Lanolin:** It is obtained from sheep wool.
- **Cutin:** It is found on leaf surfaces of plants
- **Suberin:** It is found in cell wall of **endodermis** of plant roots.

Functions

These are chemically inert. Thus these are resistant to atmospheric oxidation. Waxes have protective functions in plants and animals.

1. Some plankton (free-floating microorganisms) lives at the bottom of the food chain for marine animals. Waxes are the chief storage compounds for these planktons.
2. Waxes have water-repellent properties. They have firm consistency. Certain skin glands of vertebrates secrete waxes to protect hair and skin. Waxes keep hair pliable, lubricated, and waterproof.
3. Birds like waterfowl, secrete waxes from their preen glands. It keeps their feathers water-repellent.
4. The shiny leaves of holly, rhododendrons, poison ivy, and many tropical plants are coated with a thick layer of waxes. These waxes prevent excessive evaporation of water and protects against parasites.
5. Biological waxes have many applications in the pharmaceutical, cosmetic, and other industries. Lanolin (from lamb's wool), beeswax, carnauba wax (from a Brazilian palm tree), and wax extracted from spermaceti oil (from whales) are widely used in the manufacture of lotions, ointments, and polishes.



Biological wax: Triacontanoylpalmitate

(b) Synthetic waxes

Synthetic waxes are derived from petroleum or polyethylene. They consist of mixtures of long-chain hydrocarbons (alkanes), alcohols, aldehydes, ketones and fatty acid.

Example: Its example is paraffin wax. It is used to make candles, wax paper, lubricants, and sealing materials.

STRUCTURAL LIPIDS IN MEMBRANES

Biological membrane is a double layer of lipids. It acts as a barrier for the passage of polar molecules and ions. Membrane lipids are **amphipathic**. One end of the molecule is **hydrophobic**. The other end is **hydrophilic**. Their hydrophobic interactions with each other and their hydrophilic interactions with water pack it into sheets called **membrane bilayers.**

There are five general types of membrane lipids. These are:

1. **Glycerophospholipids:** In these compounds, hydrophobic regions are composed of two fatty acids joined to glycerol
2. **Galactolipids and sulfolipids:** These also contain two fatty acids esterified to glycerol. But they lack the characteristic phosphate of phospholipids.
3. **Archaebacterial tetraether lipids:** These have two very long alkyl chains. These chains are ether-linked to glycerol at both ends
4. **Sphingolipids:** These have single fatty acid joined to a fatty amine
5. **Sphingosine and sterols:** These have a rigid system of four fused hydrocarbon rings.

Hydrophilic group in these amphipathic compounds may be simple a single –OH group. This group is present at one end of the sterol ring system. Or they may be much more complex. There are different types of combinations of fatty acid tails and polar heads in these groups of membrane lipids. It forms enormous diversity.

Phospholipids

Glycerophospholipids and some sphingolipids are phospholipids. They have a polar head group is joined to the hydrophobic moiety. This joining occurs by a phosphodiester linkage.

Glycolipids

Other sphingolipids lack phosphate. But they have a simple sugar or complex oligosaccharide at their polar ends. These are glycolipids.

A. Glycerophospholipids or Phospholipids

Phospholipid is a type of compound/complex lipids. Glycerophospholipids are also called phosphoglycerides. They are membrane lipids. Phospholipids are derived from

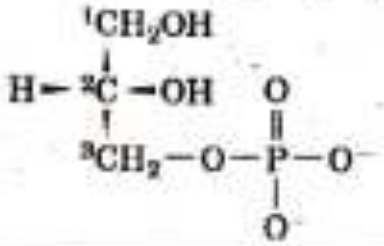
phosphatidic acid. In phospholipid, phosphatidic acid combines with one of the four organic compounds. These compounds are like:

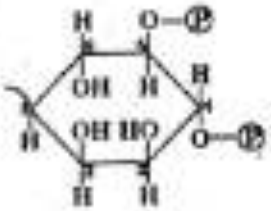
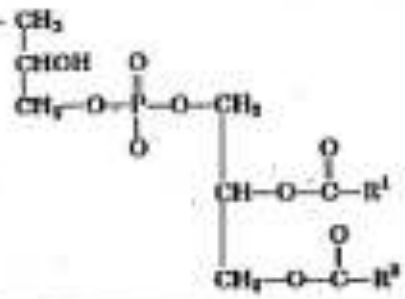
- Choline: It is a nitrogenous base
- Ethanolamine: It is an amino alcohol
- Inositol. It is an amino alcohol
- Serine: It is an amino acid.

Structure

Phosphatidic acid molecule is much similar to diglyceride. It contains a glycerol, two fatty acids.

a) Two fatty acids are esterified with first and second OH groups of glycerol.
b) A phosphate group esterified with third OH group of glycerol.



Name of glycerophospholipid	Name of X	Formula of X	Net charge (at pH 7)
Phosphatidic acid	—	— H	−1
Phosphatidylethanolamine	Ethanolamine	$-CH_2-CH_2-\overset{+}{N}H_3$	0
Phosphatidylcholine	Choline	$-CH_2-CH_2-\overset{+}{N}(CH_3)_3$	0
Phosphatidylserine	Serine	$-CH_2-\underset{COO^-}{CH}-\overset{+}{N}H_3$	−1
Phosphatidylglycerol	Glycerol	$-CH_2-\underset{OH}{CH}-CH_2-OH$	−1
Phosphatidylinositol 4,5-bisphosphate	myo-Inositol 4,5-bisphosphate		−4
Cardiolipin	Phosphatidyl-glycerol		−2

They have two fatty acids attached in ester linkage to the first and second carbons of glycerol. Thus a highly polar or charged group is attached through a phosphodiester linkage to the third carbon. Glycerol is **prochiral**. It has no asymmetric carbons. But attachment of phosphate at one end converts it into a chiral compound. Hence these can be correctly named either L-glycerol 3-phosphate, D-glycerol 1-phosphate, or sn-glycerol 3-phosphate. Glycerophospholipids are named as derivatives of the parent compound, phosphatidic acid according to the polar alcohol in the head group. For example, phosphatidylcholine and phosphatidylethanolamine have choline and ethanolamine in their polar head groups. In all these compounds, the head group is joined to glycerol through a phosphodiester bond. Here phosphate group bears a negative charge at neutral pH. The polar alcohol may be:

- Negatively charged: As phosphatidylinositol 4,5-bisphosphate
- Neutral: As phosphatidylserine
- Positively charged: As phosphatidylcholine, phosphatidylethan-olamine.

These charges contribute greatly to the surface properties of membranes. The fatty acids in glycerophospholipids can be any of a wide variety. Hence a given phospholipid (like phospha-tidylcholine) may consist of a number of molecular species. Each species has its unique complement of fatty acids. The distribution of molecular species is specific for different organisms, different tissues of same organism, and different glycerophospholipids in the same cell or tissue. Generally, glycerophospholipids contain a C16 or C18 saturated fatty acid at C-1 and a C18 to C20 unsaturated fatty acid at C-2.

Example: Phosphatidylcholine

Most common type of phospholipid is phosphatidylcholine. It is also called **lecithin.** Here choline is attached to phosphate group of phosphatidic acid. It has two ends:

a) **Hydrophilic end:** One end of the phospholipid molecule contain phosphate group. This additional compound is hydrophilic. It is polar and readily soluble in water.
b) **Hydrophobic end:** The other end contains fatty acid side chains. It is hydrophobic. It is non-polar and insoluble in water.

Function

These phosholipids are major constituents of lipid bilayer of cell membrane.

B. Galactolipids and Sulfolipids

Chloroplasts Contain galactolipids and sulfolipids. This is second group of membrane lipids. These predominate in plant cells:

(a) Galactolipids

In these, one or two galactose residues are connected by a glycosidic linkage to C-3 of a 1,2-diacylglycerol. Galactolipids are localized in the thylakoid membranes (internal membranes) of chloroplasts. They make up 70% to 80% of the total membrane lipids of a vascular plant. They are probably the most abundant membrane lipids in the biosphere.

(b) Sulfolipids

Plant membranes also contain sulfolipids. In this compound, sulfonated glucose residue is joined to a di-acylglycerol in glycosidic linkage. In sulfolipids, sulfonate on the head group bears a fixed negative charge like phosphate group in phospholipids.

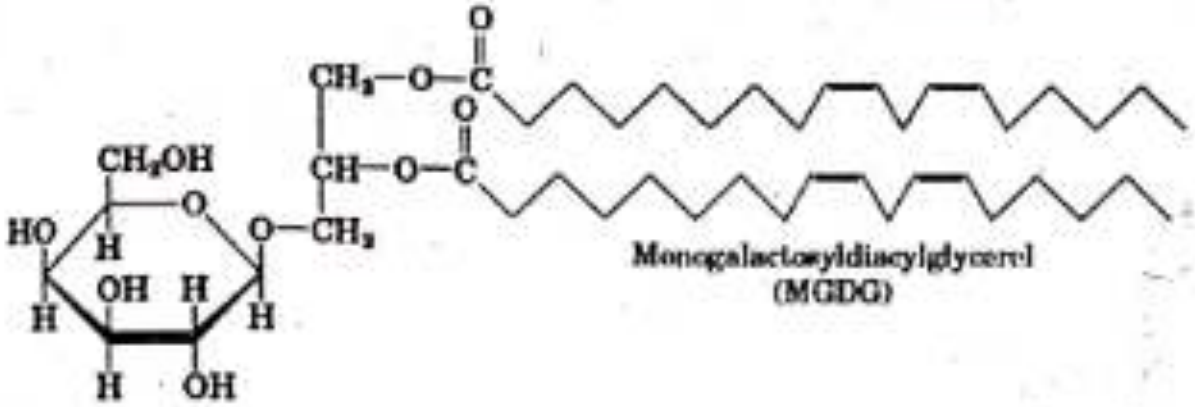


Fig: A Galactolipids

(c) Archaebacterial tetraether lipids

Archaebacteria contain unique membrane lipids. Most of archaebacteria live in habitat with high temperatures (boiling water), low pH and high ionic strength. They have membrane lipids containing long-chain (32 carbons) branched hydrocarbons linked at each end to glycerol. These linkages occur through ether bonds. Thus it is much more stable to hydrolysis at low pH and high temperature. In their fully extended form, these archaebacterial lipids are twice the length of phospholipids and sphingolipids. At each end of the extended molecule is a polar head of glycerol. This head is linked to either phosphate or sugar residues.

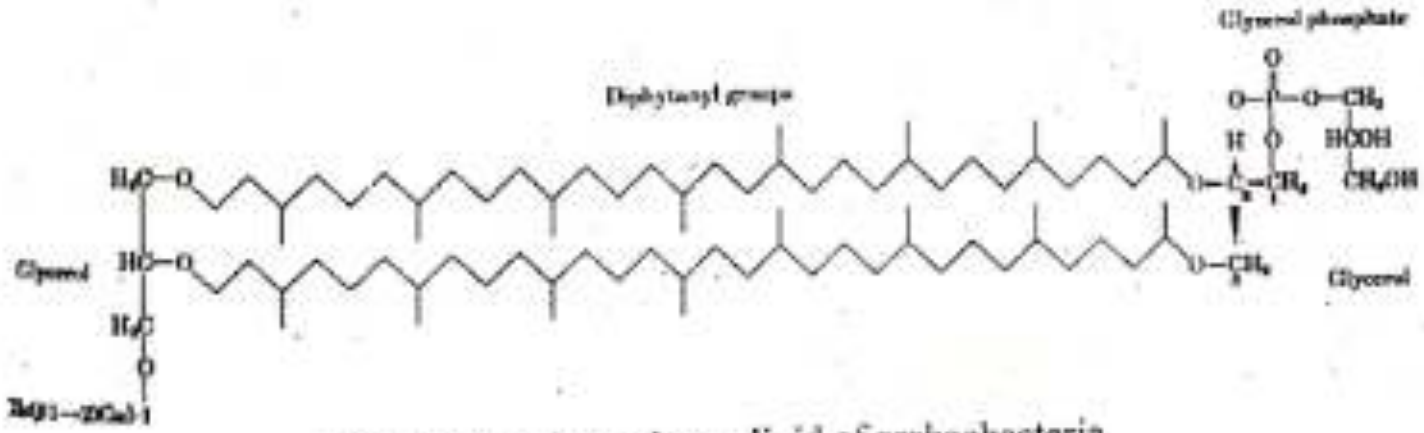


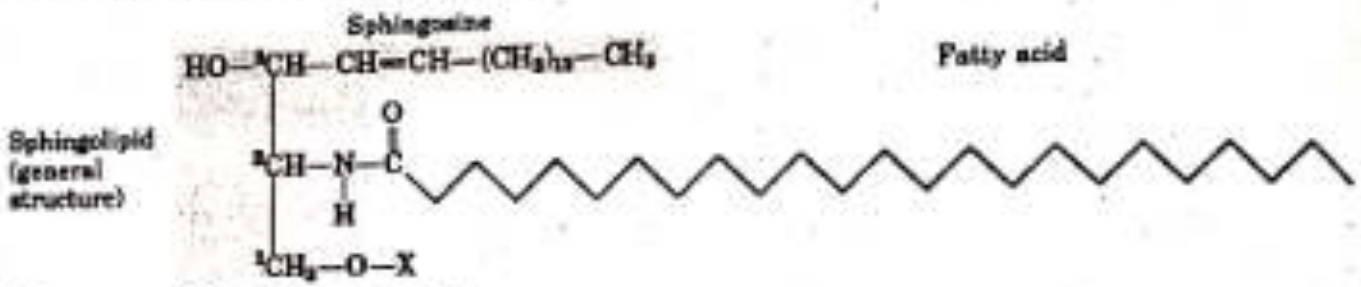
Fig: A typical membrane lipid of archaebacteria

(d) SPHINGOLIPIDS

Sphingolipids are derivatives of sphingosine. Sphingolipids are fourth large class of membrane lipids. They also have a polar head group and two nonpolar tails. But they contain no glycerol. Sphingolipids are composed of:

a) One molecule of the long-chain amino alcohol sphingo- sine (also called 4-sphingenine) or one of its derivatives
b) One molecule of a long-chain fatty acid
c) A polar head group that is joined by a glycosidic linkage

Carbons C-1, C-2, and C-3 of the sphingosine molecule are structurally similar to the three carbons of glycerol in glycerophospholipids. A fatty acid is attached in amide linkage to the –NH2 on C-2. Now the resulting compound is a **ceramide.** Ceramide is the structural parent of all sphingolipids.

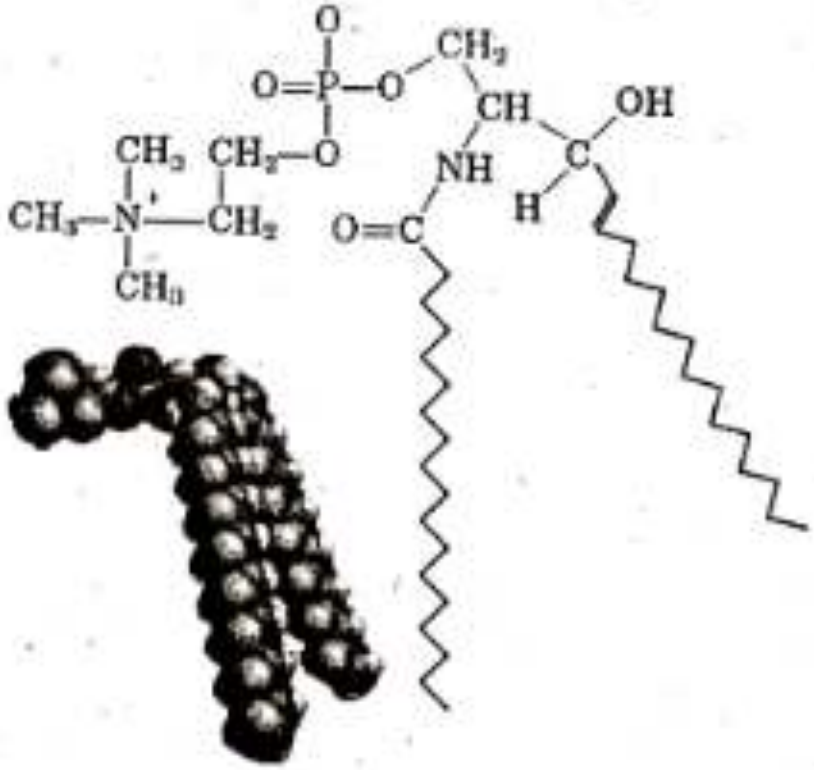


Subclasses of Sphingolipids

There are three subclasses of sphingolipids. All are derivatives of ceramide. But they differ in their head groups. These three groups are sphingomyelins, neutral (uncharged) glycolipids and gangliosides.

(a) Sphingomyelins

Sphingomyelins contain phosphocholine or phosphoethanolamine as their polar head group. Therefore, these are classified along with glycerophospholipids as phospholipids. The sphingomyelins resemble phosphatidylcholines in their general properties and three-dimensional structure. Both have no net charge on their head groups. Sphingomyelins are present in the plasma membranes of animal cells. These are especially prominent in myelin. Myelin is a membranous sheath that surrounds and insulates the axons of some neurons. Hence these are the name sphingomyelins due to presence in myelin sheath.



Sphingomyelin

(b) Glycosphingolipids

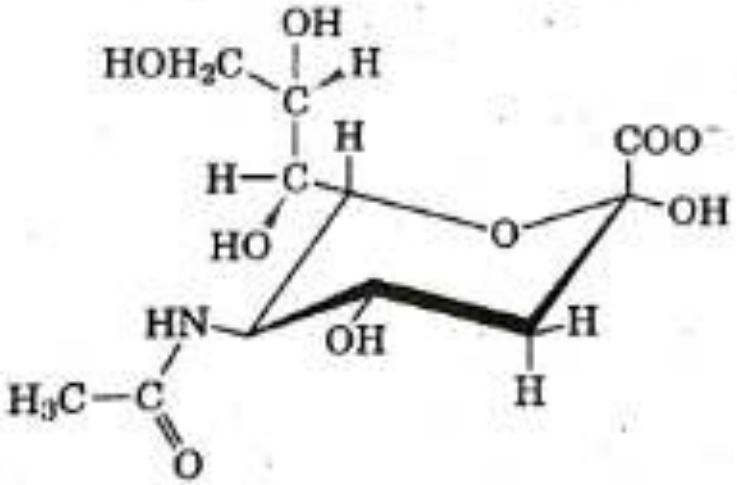
Glycosphingolipids occur largely in the outer face of plasma membranes. They have head groups with one or more sugars. These sugars are connected directly to OH at C-1 of the ceramide moiety. They do not contain phosphate.

Cerebrosides are glycosphingolipids. They have a single sugar linked to ceramide. Those cerebrosides with galactose are found in the plasma membranes of cells in neural tissue. Those with glucose in the plasma membranes of cells are found in non-neural tissues.

Globosides are glycosphingolipids with two or more sugars. These sugars are D-glucose, D-galactose, or N-acetyl-D-galactosamine. Cerebrosides and globosides are sometimes called **neutral glycolipids**. They have no charge at pH 7.

(c) Gangliosides

Gangliosides are the most complex sphingolipids. They have oligosaccharides as their polar head groups and one or more residues of N-acetylneuraminic acid (Neu5Ac), a sialic acid, at the termini. Sialic acid gives gangliosides the negative charge at pH 7. It distinguishes them from globosides. Gangliosides with one sialic acid residue are in the GM (M for mono-) series. Those with two are in the GD (D for di-) series, and so on (GT, three sialic acid residues; GQ, four).



N-Acetylneuraminic acid (a sialic acid)
(Neu5Ac)

Importance of Sphingolipids

In humans, at least 60 different sphingolipids have been identified in cellular membranes. They perform different functions:

1. Many of these are especially prominent in the plasma membranes of neurons. Some have prominent sites on the cell surface. But a specific function for only a few sphingolipids has been discovered thus far.
2. The carbohydrate moieties of certain sphingolipids define the human blood groups. Therefore, it determines the type of blood that individuals can safely receive in blood transfusions.
3. Sphingomyelins are present in the plasma membranes of animal cells. These are

especially prominent in myelin.

4. Gangliosides are concentrated in the outer surface of cells. Here they present points of recognition for extracellular molecules or surfaces of neighboring cells. The kinds and amounts of gangliosides in the plasma membrane change dramatically during embryonic development.
5. Tumor formation induces the synthesis of a new complement of gangliosides.
6. Very low concentrations of a specific ganglioside have been found to induce differentiation of cultured neuronal tumor cells.

LIPOPROTEINS

A lipoprotein is a conjugate protein. Their primary purpose is to transport hydrophobic lipid molecules in water, as in blood or extracellular fluid.

Structure

Lipoproteins have single-layer phospholipids and cholesterol outer shell. Their hydrophilic portions are present outward toward surrounding water. Their lipophilic portions oriented inwards toward the lipids molecules within the particles. Apolipoproteins are embedded in the membrane.

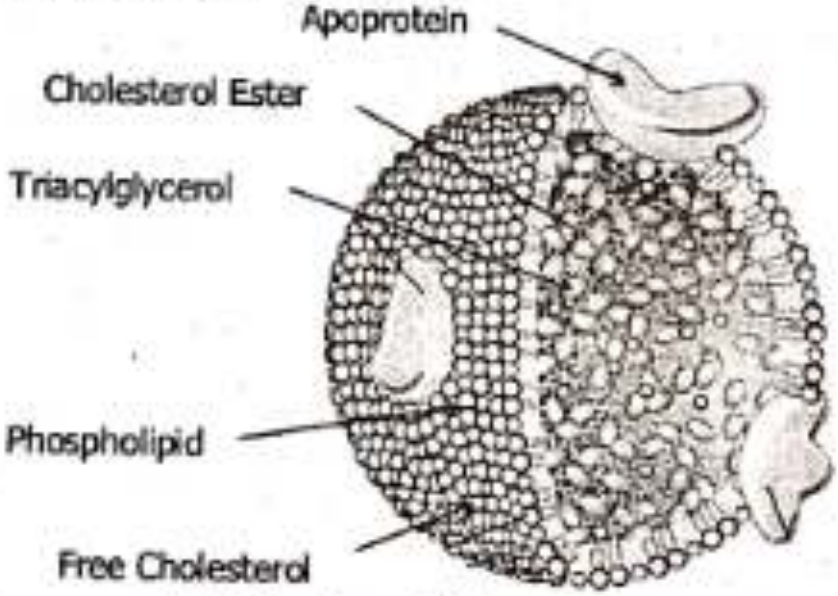


Fig: Structure of lipoproteins

Classification

Lipoproteins may be classified as five major groups follows:

1. **Chylomicrons:** They carry triglycerides (fat) from the intestines to the liver, to skeletal muscle, and to adipose tissue.
2. **Very-low-density lipoproteins (VLDL):** They carry triglycerides from the liver to adipose tissue.
3. **Intermediate-density lipoproteins (IDL):** They are intermediate between VLDL and LDL
4. **Low-density lipoproteins (LDL):** They carry 3,000 to 6,000 fat molecules (phospholipids, cholesterol, triglycerides, etc.) around the body. LDL particles are

sometimes called as **bad lipoprotein**. Their concentrations, dose related, correlate with atherosclerosis formation.

5. **High-density lipoproteins (HDL):** They collect fat molecules (phospholipids, cholesterol, triglycerides, etc.) from cells/tissues, and take it back to the liver. HDLs are also called as **good lipoprotein**. Their higher concentrations cause low rates of atherosclerosis progression.

Functions

Many enzymes, transporters, structural proteins, antigens, adhesions, and toxins are lipoproteins. Its examples are plasma lipoprotein particles. These particles are classified as HDL, LDL, IDL, VLDL and ULDL. These complex protein capsules enable fats to be carried in all extracellular water, including blood stream. Proteolipids are a different kind of protein-lipid combination that are insoluble in water. Proteolipids are abundant in brain tissue, and are also present in many other animal and plant tissues.

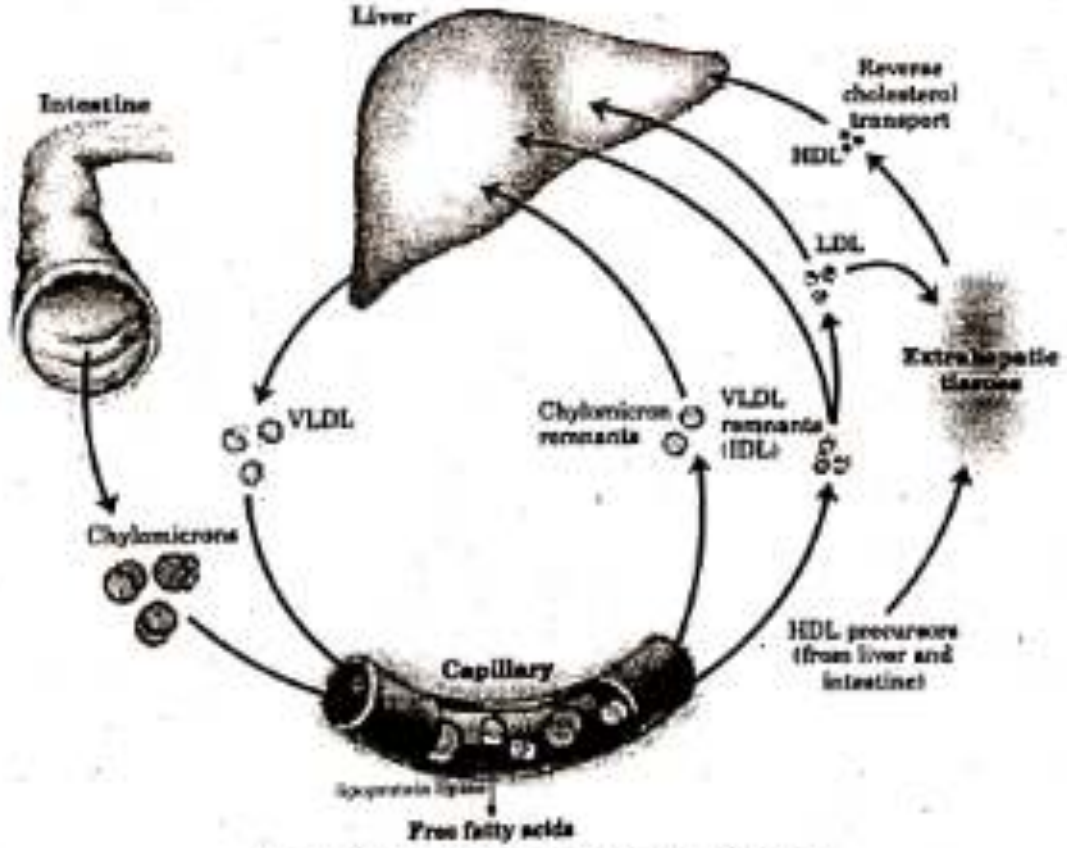


Fig: Lipoproteins and lipid transport

Plasma lipoprotein particles

Fats are insoluble in water. Therefore, fats cannot be transported in extracellular water. Thus fats are surrounded by hydrophilic proteins. Now these function as transport vehicles. The role of lipoprotein particles is to transport all hydrophobic fat molecules. These molecules are like triacylglycerols, phospholipids, cholesterol.

The lipoprotein particles have hydrophilic groups of phospholipids, cholesterol, and apoproteins directed outward. Such characteristics make them soluble in the salt water-

based blood pool. Phospholipid monolayer and apoproteins protect the triglyceride-fats and cholesteryl esters from the water.

STEROLS

Sterols are structural lipids present in the membranes of most eukaryotic cells.

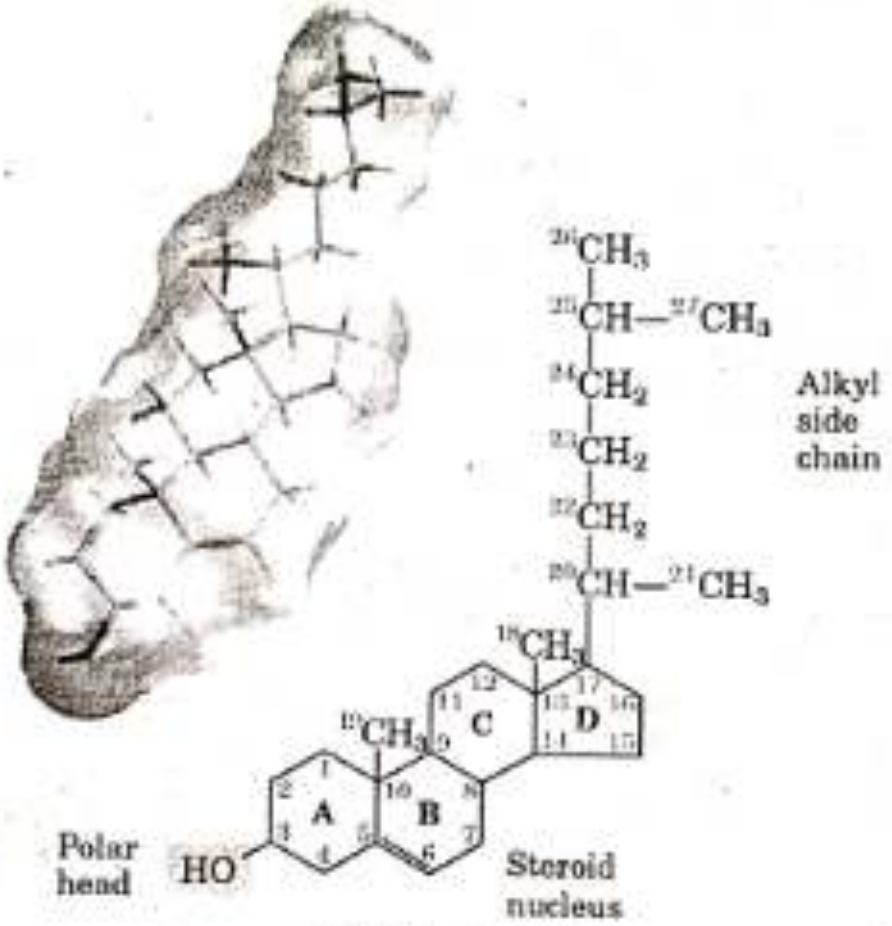


Fig: Cholesterol

Steroid Nucleus

The characteristic structure of sterol is the steroid nucleus. This nucleus consists of four fused rings. Three are fused with six carbons and one fused with five carbon ring. The steroid nucleus is almost planar. It is relatively rigid. The fused rings do not allow rotation about COC bonds.

Different forms

Different forms of sterols are found in organisms:

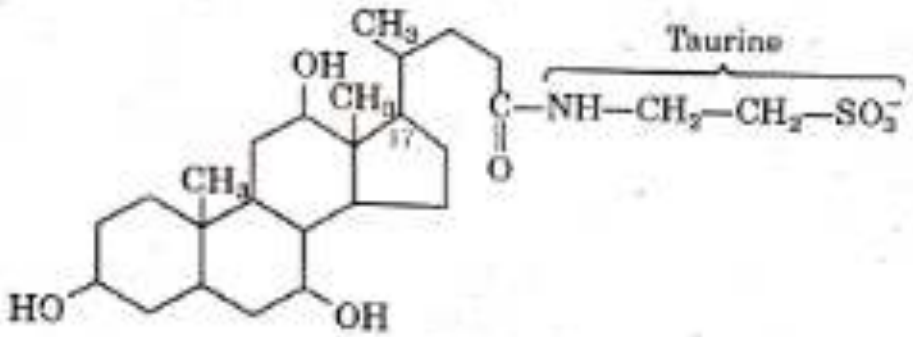
1. **Cholesterol:** Cholesterol is the major sterol in animal tissues. It is amphipathic. It has a polar head group (hydroxyl group at C-3) and a non-polar hydrocarbon body (steroid nucleus and hydrocarbon side chain at C-17). It is about as long as a 16-carbon fatty acid in its extended form.
2. **Stigma sterol and ergosterol:** Similar sterols are found in other eukaryotes. For example, stigma sterol is found in plants and ergosterol is found in fungi.
3. Bacteria cannot synthesize sterols. A few bacterial species can incorporate exogenous sterols into their membranes.

Biosynthesis

The sterols of all eukaryotes are synthesized from simple five-carbon isoprene subunits.

Roles

1. Sterols have important role in membrane.
2. The sterols act as precursors for a variety of products. These products have specific biological activities.
3. Steroid hormones are potent biological signals. They regulate gene expression.
4. Bile acids are polar derivatives of cholesterol. They act as detergents in the intestine. They emulsify dietary fats to make them more readily accessible to digestive lipases.



Taurocholic acid
(a bile acid)

LIPIDS AS SIGNALS, COFACTORS AND PIGMENTS

Another group of lipids is present in much smaller amounts. They have active roles in the metabolic traffic as metabolites and messengers.

1. **Messengers:**
   - Some serve as potent signals as hormones. They are carried in the blood from one tissue to another.
   - Some acts as intracellular messengers. They are generated in response to an extracellular signal (hormone or growth factor).
2. **Enzyme cofactors:** Others function as enzyme cofactors in electron-transfer reactions in chloroplasts and mitochondria. Some are involved in transfer of sugar moieties in a variety of glycosylation reactions.
3. **Pigments:** A third group consists of lipids with a system of conjugated double bonds. These include pigment molecules. They absorb visible light.
   - Some of these act as light-capturing pigments in vision and photosynthesis
   - Some of these produce natural colorations like orange of pumpkins and carrots and yellow of canary feathers.
4. **As volatile lipids:** A very large group of volatile lipids produced in plants. They serve as signals that pass through the air. They allow the plants to communicate with each other and to invite animal friends and deter enemies.

Lipids as Chemical Messenger

A. Phosphatidylinositols and Sphingosine Derivatives

Signals Phosphatidylinositol and its phosphorylated derivatives act at several levels. They regulate cell structure and metabolism.

1. Phosphatidylinositol 4,5-bisphosphate

Phosphatidylinositol 4,5-bisphosphate is found in the cytoplasmic (inner) face of plasma membranes. It serves as a reservoir of messenger molecules. These molecules are released inside the cell in response to extracellular signals interacting with specific surface receptors.

2. Vasopressin

Extracellular signals like hormone vasopressin activate a specific enzyme phospholipase C in the membrane. This enzyme hydrolyzes **phosphatidylinositol 4, 5- bisphosphate** to release two products. These products act as intracellular messengers. These products are inositol 1,4,5-trisphosphate (IP_3) and diacylglycerol. The inositol 1,4,5-trisphosphate (IP_3) is water-soluble. The diacylglycerol remains associated with plasma membrane.

a) IP_3 triggers release of Ca^{2+} from endoplasmic reticulum. The combination of diacylglycerol and elevated cytosolic Ca^{2+} activates the enzyme protein kinase C.

b) This enzyme phosphorylates the specific proteins. It starts the cell's response to the extracellular signal.

3. As points of nucleation

Inositol phospholipids also act as points of nucleation for supramolecular complexes. These complexes are involved in signaling or in exocytosis. Certain signaling proteins bind specifically to phosphatidylinositol 3,4,5-trisphosphate in the plasma membrane. They initiate the formation of multienzyme complexes at the membrane's cytosolic surface. Therefore, formation of phosphatidylinositol 3,4, 5-trisphosphate in response to extracellular signals brings the proteins together in signaling complexes at the surface of plasma membrane.

Roles of membrane Sphingolipids

Membrane sphingolipids also can serve as sources of intracellular messengers. Both ceramide and sphingomyelin are potent regulators of protein kinases. The ceramide or its derivatives are involved in the regulation of cell division, differentiation, migration, and programmed cell death.

B. Eicosanoids

Eicosanoids are paracrine hormones. The paracrine are substances that act only on cells near the point of hormone synthesis. These fatty acid derivatives have a variety of effects on vertebrate tissues.

Functions

1. They are involved in reproductive function
2. They are involved in the inflammation, fever and pain associated with injury or disease.
3. They are involved in the formation of blood clots and regulation of blood pressure.

4. They are involved in gastric acid secretion.
5. They are involved in various other processes important in human health or disease.

Nature

All eicosanoids are derived from **arachidonic acid**. It is a 20-carbon polyunsaturated fatty acid. eicosanoids take their name from arachidonic acid (Greek eikosi twenty).

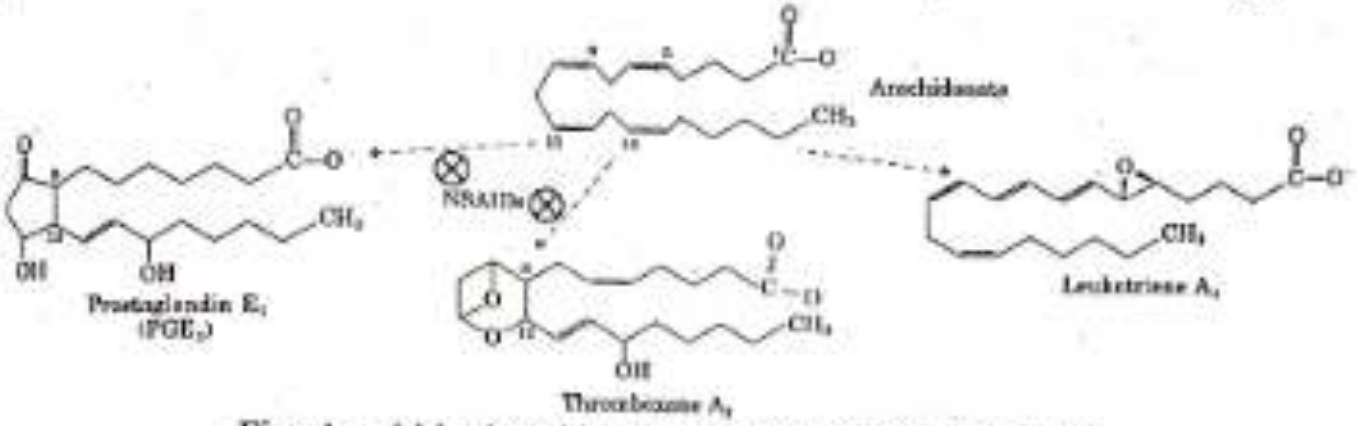


Fig: Arachidonic acid and some eicosanoid derivatives

Classes

There are three classes of eicosanoids: prostaglandins, thromboxanes, and leukotrienes.

(i) Prostaglandins (PG)

Prostaglandins contain a five-carbon ring. This ring originates from the chain of arachidonic acid. Their name derives from **prostate gland**. It was tissue from which they were first isolated by Bengt Samuelsson and Sune Bergstrom. There are two groups of prostaglandins: PGE (ether-soluble) and PGF (buffer-soluble). Each group contains numerous subtypes, named PGE1, PGE2, PGF1, and so forth.

Functions

Prostaglandins have different functions.

1. Some stimulate contraction of the smooth muscle of uterus during menstruation and labor.
2. Some others affect blood flow to specific organs.
3. Some are involve in wake-sleep cycle.
4. Some cause responsiveness of certain tissues to hormones like epinephrine and glucagon.
5. Prostaglandins in a third group elevate body temperature (producing fever). They also cause inflammation and pain.

(ii) Thromboxanes

Thromboxanes have a six-membered ring containing an ether. They are produced by platelets (also called thrombocytes). They act in the formation of blood clots and reduction of blood flow to site of a clot.

The non-steroidal anti-inflammatory drugs (NSAIDs) are like aspirin, ibuprofen, and meclofenamate. These drugs inhibit the enzyme **prostaglandin H_2 synthase** enzyme. This enzyme catalyzes an early step in the pathway from arachidonate to prostaglandins and thromboxanes.

(ii) Leukotrienes

Leukotrienes first found in leukocytes. They contain three conjugated double bonds. They are powerful biological signals. For example, leukotriene D4 is derived from leukotriene A4. It induces contraction of smooth muscle lining airways to the lung. Overproduction of leukotrienes causes asthmatic attacks. Leukotriene synthesis is one target of antiasthmatic drugs such as prednisone.

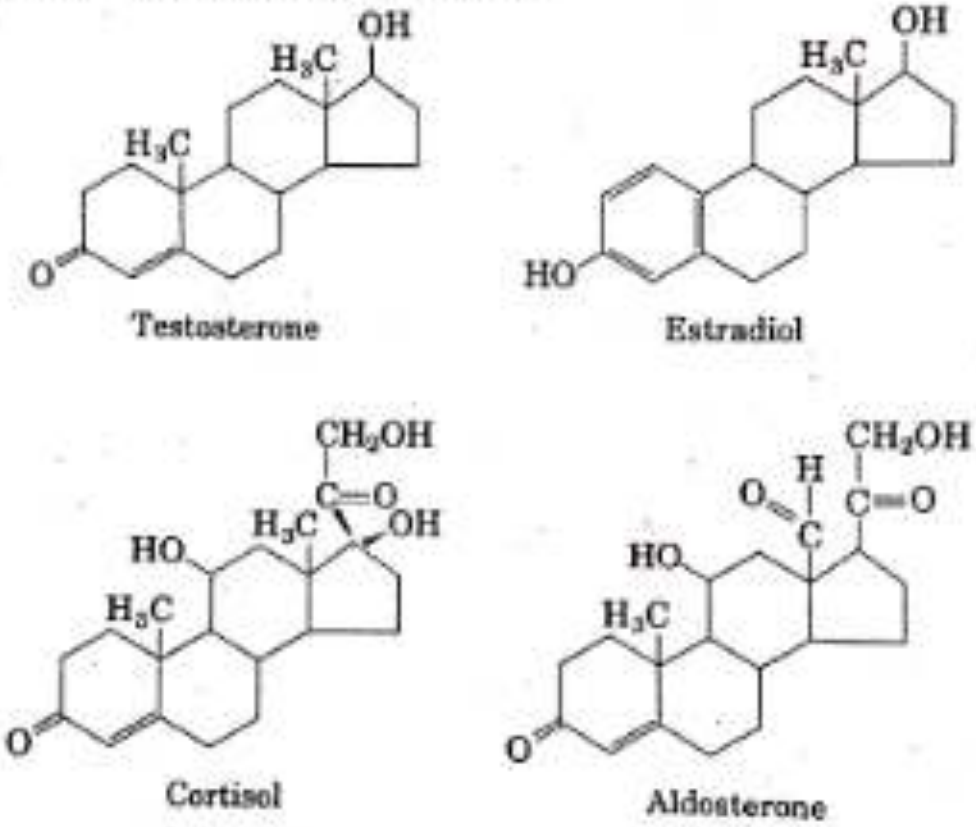


Fig: Different Types of Steroid Hormones

C. Steroid Hormones

Nature

Steroids are oxidized derivatives of sterols. They have the sterol nucleus. But they lack the alkyl chain attached to ring D of cholesterol. They are more polar than cholesterol.

Mechanism of Action

Steroid hormones move through bloodstream (on protein carriers) from their site of production to target tissues. They enter the target cells. They bind to highly specific receptor proteins in the nucleus. Thus cause trigger changes in gene expression and metabolism.

Types

The major groups of steroid hormones are male and female sex hormones and hormones produced by the adrenal cortex, cortisol and aldosterone.

Lipids as Cofactors and Electron Carrier

Vitamins E and K and Lipid Quinones are Oxidation-Reduction Cofactors

Vitamin A

Vitamin E is the collective name for a group of closely related lipids called **tocopherols**. All of these contain a substituted aromatic ring and a long isoprenoid side chain.

Functions

1. They are hydrophobic. Therefore, tocopherols associate with cell membranes, lipid deposits, and lipoproteins in the blood.
2. Tocopherols are biological antioxidants. The aromatic ring reacts with and destroys most reactive forms of oxygen radicals and other free radicals. Hence they protect unsaturated fatty acids from oxidation. Thus they prevent oxidative damage to membrane lipids.

Source

Tocopherols are found in eggs and vegetable oils. They are especially abundant in wheat germ. Vitamin E deficiency in humans is very rare. Its principal symptom is fragile erythrocytes.

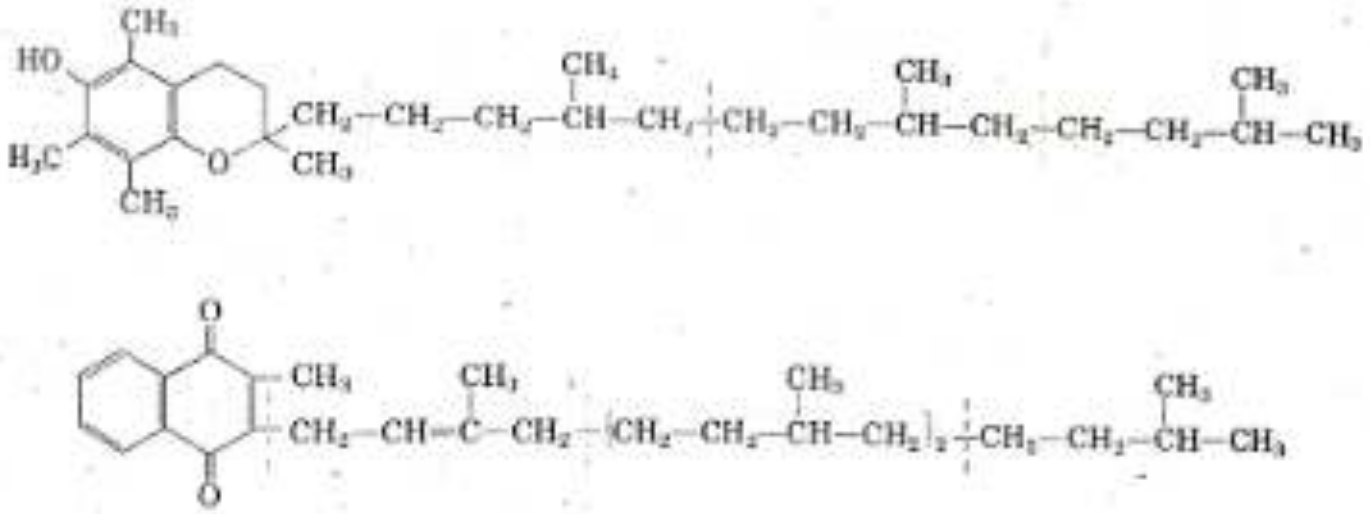


Fig: Vitamin E and K

Vitamin K

The aromatic ring of vitamin K undergoes a cycle of oxidation and reduction during formation of active prothrombin. Prothrombin is a blood plasma protein essential in blood clotting. Prothrombin is a proteolytic enzyme. It splits peptide bonds in the blood protein fibrinogen to convert it to fibrin. Fibrin is insoluble fibrous protein. It holds blood clots together. The vitamin K deficiency causes slows blood clotting, which can be fatal.

Sources

Vitamin K_1 (phylloquinone) is found in green plant leaves. The vitamin K_2 (menaquinone), is formed by bacteria living in the vertebrate intestine.

Ubiquinone (Coenzyme Q) and Plastoquinone

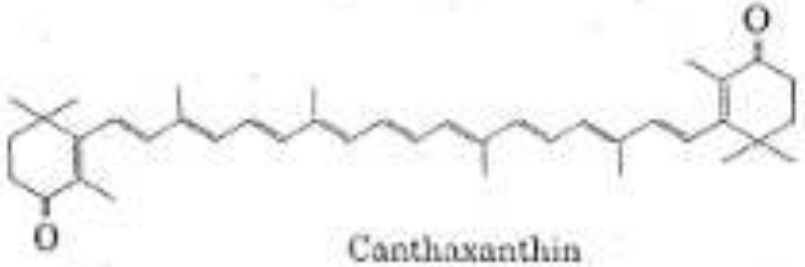
Ubiquinone and plastoquinone are isoprenoids. They function as lipophilic electron carriers in the oxidation-reduction reactions. This reaction drives ATP synthesis in

mitochondria and chloroplasts. Both ubiquinone and plastoquinone can accept either one or two electrons or either one or two protons.

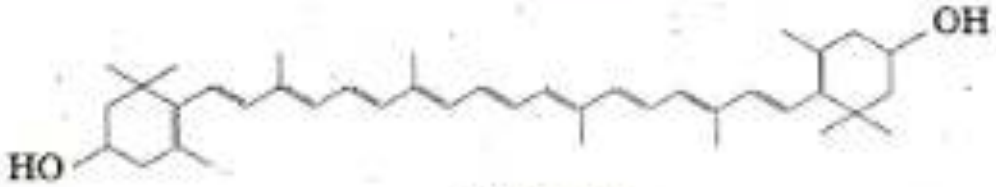
Lipids as Pigments

Many Natural Pigments Are Lipidic Conjugated Dienes

Conjugated dienes have carbon chains with alternating single and double bonds. This structural arrangement allows delocalization of electrons. Therefore, these compounds can be excited by low-energy electromagnetic radiation (visible light). It gives them colors visible to humans and other animals. **Carotene** is yellow orange; similar compounds give bird feathers their striking reds, oranges, and yellows. These pigments are synthesized from five-carbon isoprene derivatives.



Canthaxanthin
(bright red)



Zeaxanthin
(bright yellow)

Fig: Lipids as pigments in plants and bird feathers

EXERCISE

LONG QUESTIONS

1. **What are lipids? Describe their function and classification.**
2. **What are fatty acids? Describe their structure and types.**
3. **What are acylglycerol? Give it components and structure.**
4. **What are storage lipids? Give their types and functions.**
5. **What are waxes? Give their nature, types and importance.**
6. **Discuss structural lipids and membranes. Give their types and structure.**
7. **What are Glycerophospholipids or Phospholipids? Discuss their structure and functions.**
8. **Describe structure of spingolipids.**
9. **Write note on: (a) waxes (b) phospholipids (c) terpenes (d) steroids**

10. What are lipoproteins? Illustrate their structure and function.
11. Write note on spingolipids, their types and roles.
12. Discuss role of lipids as messengers.

SHORT QUESTIONS

1. What are lipids? Write their functions.
2. What are three classes of lipids?
3. Define acylglycerol. How they are formed?
4. Differentiate between saturated and unsaturated fatty acids.

Saturated fatty acids	Unsaturated fatty acids
1. Fatty acids in which all of the internal carbon atoms possess hydrogen side groups are called saturated fatty acids.	1. They have one or more pairs of carbon atoms joined by a double bond.
2. They contain maximum possible number of hydrogen atoms.	2. They are not fully saturated with hydrogen.
3. Saturated fatty acids are solid at room temperature.	3. Unsaturated fatty acids are liquid at room temperature.
4. They have higher melting point.	4. They have lower melting point.
5. They are more common in animal lipids (fats).	5. They are more common in plant lip (oils).
6. Its example is palmitic acid.	6. Its example is oleic acid.

5. Write nomenclature of fatty acids.
6. What are adipocytes?

Ans: In most eukaryotic cells, triacylglycerols form a separate phase of microscopic, oily droplets in the aqueous cytosol. They are used as depots of metabolic fuel. Vertebrates have specialized cells called **adipocytes**, or fat cells. They store large amounts of triacylglycerols as fat droplets. These droplets nearly fill the cell.

7. What are advantages of fats as storage compounds?

And: There are two advantages for using triacylglycerols as stored fuels than polysaccharides like glycogen and starch.

a) The carbon atoms of fatty acids are more reduced than those of sugars. Therefore, oxidation of triacylglycerols yields more than twice energy than the oxidation of carbohydrates.
b) Triacylglycerols are hydrophobic. Therefore, unhydrated organism carries fat as fuel. They not have to carry the extra weight of water of hydration associated with stored polysaccharides (2 g per gram of polysaccharide) in human.

8. Write two uses of stored fats.
9. What are waxes? Give their types.
10. What are synthetic waxes?
11. Name the fiver general types of membranous lipids.
12. What are phospholipids? Discuss Phosphatidylcholine choline

12. What are glactolipids?
13. What are steriods? What types of rings are present in them?
14. Name different classes of lipoproteins.
15. What do you know about plasma lipoprotein particles?

MULTIPLE CHOICE QUESTIONS (MCQs)

1. Heterogeneous group of compounds are:
(a) proteins (b) lipids (c) nucleic acid (d) carbohydrates

2. Formula of tristearin:
(a) $C_5H_{10}O_6$ (b) $C_{50}H_{115}O_6$ (c) $C_{52}H_{110}O_4$ (d) $C_{57}H_{110}O_6$

3. Acylglycerol is:
(a) simple lipid (b) complex lipid (c) derived lipid (d) complicated lipid

4. Cholesterol is:
(a) simple lipid (b) complex lipid (c) derived lipid (d) complicated lipid

5. Which is ester?
(a) fatty acid (b) cholesterol (c) acylglycerol (d) terpenes

6. Number of types of fatty acids are:
(a) 10 (b) 20 (c) 30 (d) 40

7. Which of the followings is an unsaturated fatty acid?
(a) Acetic acid (b) Butyric acid(c) Oleic acid (d) Palmitic acid

8. Wax found in sheep wool is:
(a) cutin (b) lanolin (c) suberin (d) bee wax

9. Which of the following is present on endodermis of plants:
(a) cutin (b) lanolin (c) suberin (d) bee wax

10. Which of the followings is nitrogenous base:
(a) inositol (b) serine (c) ethanolamine (d) choline

11. Lipid found in plasma membrane is:
(a) inositol (b) serine (c) ethanolamine (d) lecithin

12. Isoprene unit forms:
(a) waxes (b) prostaglandin (c) terpenes (d) steriods

13. Number of carbon in steriods are:
(a) 4 (b) 14 (c) 17 (d) 25

14. Vitamin D is:
(a) waxes (b) prostaglandin (c) terpenes (d) steriods

15. Good lipoprotein is:
(a) HDL (b) LDL (c) VLDL (d) IDL

16. Bad lipoprotein is:
(a) HDL (b) LDL (c) VLDL (d) IDL F

Ans:
1. (b) 2. (d) 3. (a) 4. (c) 5. (c) 6. (c) 7. (c) 8. (b) 9. (c) 10. (c) 11. (d) 12. (c) 13. (c) 14. (d) 15. (a) 16. (b)

Chapter 6

VITAMINS AND COFACTORS

Syllabus Chapter 5: Vitamins and cofactors:
Occurrence, structure and biochemical function of vitamins B complex group.

VITAMINS

The essential organic compounds which are needed in very small amount in the body are called vitamins. Normal metabolic activity of animals depends on vitamins. Vitamins are chemically unrelated, organic substances. They occur in many foods in small amounts. They are necessary for normal metabolic functioning. There are two types of vitamins.

(a) **Water soluble vitamins:** Most water-soluble vitamins act as coenzymes. They are needed in metabolism. Examples: vitamins B and vitamin C.

(b) **Fat soluble vitamins:** The far-soluble vitamins perform different functions.

Requirement of Vitamins in vertebrates:
The vitamin C and fat soluble vitamins (A, D, E and K) are need in limited amount in the vertebrates. Different animals have different requirements of vitamins. For example, humans and guinea pigs require vitamin C. But rabbits do not require it. Some birds require vitamin A. But other birds do not require it.

Water Soluble Vitamins

Vitamin	Characteristics	Functions	Occurrence
Thiamin (vitamin B_1)	Destroyed by heat and oxygen, especially in alkaline environment	Part of coenzyme needed for oxidation of carbohydrates; Coenzyme needed an	Lean meats, liver, eggs, whole gain cereal, Leafy green vegetables, legumes

		synthesis of ribose	
Riboflavin (vitamin B_2)	1. Stable to heat acids, and oxidation 2. destroyed by alkalis and light	Part of enzymes and coenzymes needed for oxidation of glucose and fatty acids and for cellular growth	Meats, dairy products, leafy green vegetables, whole-grain cereals
Niacin (nicotinic acid)	1.Stable to heat, acids, and alkalis 2.converted to niacin amide by cells 3. synthesized from tryptophan	Part of coenzymes needed for oxidation of glucose and synthesis of proteins, fats, and nucleic acids	Liver, lean meats, poultry, peanuts, legumes
Vitamin B_6	1.Group of three compounds 2. stable to heat and acids 3. destroyed by oxidation, alkalis, and ultraviolet light	1.Coenzyme needed for synthesis of proteins and different amino acids 2.For conversion of tryptophan to niacin, for production of antibodies 3.For synthesis of nucleic acids	Liver, meats, fish, poultry, bananas, avocados, beans, peanuts, whole-grain cereals, egg yolk
Pantothenic acid	Destroyed by heat, acids, and alkalis	Part of coenzyme needed for oxidation of carbohydrates and fats	Meats, fish, whole-grain cereals, legumes, milk, fruits, vegetables
Cyanocobalamin (vitamin B_{12})	1.Complex, cobalt-containing compound; 2. Stable to heat; 3.Inactivated by light, strong acids, and strong alkalis; 4.Absorption regulated by stimulation of gastric Glands 5. Stored In liver	1.Part of coenzyme needed for synthesis of nucleic acids and for metabolism of carbohydrates 2.plays role in synthesis of myelin	Liver, meats, poultry, fish, milk, cheese, eggs
Folate (folic acid)	1.Occurs an several forms 2.destroyed by oxidation in acid environment or by heat in alkaline environment; 3. Stored in live; where it is converted into folinic acid	1. Coenzyme needed for metabolism of certain amino acids and for synthesis of DNA 2.promotes production of normal red blood cells	Liver, leafy green vegetables, whole-grain cereals, legumes
Biotin	1. Stable to heat, acids, and light 2. destroyed by oxidation and alkalis	Coenzyme needed for metabolism of amino acids and fatty acids and for synthesis of nucleic acids	Liver, egg yolk, nuts, Legumes, mushrooms
Ascorbic acid	1.Closely related to monosaccharide	1. Needed for production of collagen,	Citrus fruits, citrus juices, tomatoes,

	2. stable in acids, but destroyed by oxidation, heat, light, and alkalis	2. conversion of folate to folinic acid, and metabolism of certain amino acids 3.Promotes absorption of iron and synthesis of hormones from cholesterol	cabbage, potatoes, leafy green vegetables, fresh fruits

Fat Soluble Vitamins

Vitamin	Characteristics	Functions	Occurrence
Vitamin A	1.Occurs in several forms synthesized from carotenes 2. stored in liver 3. Stable in hear, acids, and alkalis 4. Unstable in light	1.Necessary for synthesis of visual pigments, mucoproteins, and mucopolysaccharides 2.For normal development of bones and teeth 3.Maintenance of epithelial cells	Liver, fish, whole milk, butter, eggs, leafy , green vegetables, and yellow and orange vegetables and fruits
Vitamin D	1.A group of sterols 2.Resistant to heat, oxidation, acids, and alkalis 3.Stored in liver, skin, brain, spleen, and bones	1.Promotes absorption of calcium and phosphorus 2. Promotes development of teeth and bones	Produced in skin exposed to ultraviolet light; in milk, egg yolk, fish-liver oils, fortified foods (creamy food)
Vitamin E	1. A group of compounds; resistant to heat and visible light 2. Unstable in presence of oxygen and ultraviolet light 3.stored in muscles and adipose tissue	1.An antioxidant 2. prevents oxidation of vitamin A and polyunsaturated fatty acids 3. Maintain stability of cell membranes	Oils from cereal seeds, salad oils, margarine, shortenings, fruits, nuts, and vegetables
Vitamin K	1.Occurs in several forms 2.Resistant to heat, but destroyed by acids, alkalis, and light; stored in liver	1.Needed for synthesis of prothrombin 2.needed for blood clotting	Leafy green vegetables, egg yolk, pork liver, soy oil, tomatoes, cauliflower

Vitamin B Complex

Vitamin B includes following types:

Thiamine

It is water soluble.

Deficiency may cause beriberi and Wernicke-Korsakoff syndrome.

Good sources include: yeast, pork, cereal grains, sunflower seeds, brown rice, whole-grain rye, asparagus, kale, cauliflower, potatoes, oranges, liver, and eggs.

Vitamin B2

Chemical name: Riboflavin

It is water soluble

Deficiency may cause ariboflavinosis

Good sources include: asparagus, bananas, persimmons, okra, chard, cottage cheese, milk, yogurt, meat, eggs, fish, and green beans

Vitamin B3

Chemical names: Niacin, niacinamide

It is water soluble.

Deficiency may cause pellagra, with symptoms of diarrhea, dermatitis, and mental disturbance.

Good sources include: liver, heart, kidney, chicken, beef, fish (tuna, salmon), milk, eggs, avocados, dates, tomatoes, leafy vegetables, broccoli, carrots, sweet potatoes, asparagus, nuts, whole-grains, legumes, mushrooms, and brewer's yeast.

Vitamin B5

Chemical name: Pantothenic acid

It is water soluble.

Deficiency may cause paresthesia, or "pins and needles."

Good sources include: meats, whole-grains (milling may remove it), broccoli, avocados, royal jelly, and fish ovaries.

Vitamin B6

Chemical names: Pyridoxine, pyridoxamine, pyridoxal

It is water soluble.

Deficiency may cause anemia, peripheral neuropathy, or damage to parts of the nervous system other than the brain and spinal cord.

Good sources include: meats, bananas, whole-grains, vegetables, and nuts. When milk is dried, it loses about half of its B6. Freezing and canning can also reduce content.

Vitamin B7

Chemical name: Biotin

It is water soluble.

Deficiency may cause dermatitis or enteritis, or inflammation of the intestine.

Good sources include: egg yolk, liver, some vegetables.

Vitamin B9

Chemical names: Folic acid, folinic acid

It is water soluble.

Deficiency during pregnancy is linked to birth defects. Pregnant women are encouraged to supplement folic acid for the entire year before becoming pregnant.

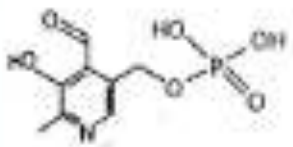
Good sources include: leafy vegetables, legumes, liver, baker's yeast, some fortified grain products, and sunflower seeds. Several fruits have moderate amounts, as does beer.

Vitamin B_{12}

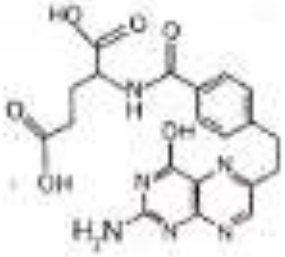
Chemical names: Cyanocobalamin, hydroxocobalamin, methylcobalamin

It is water soluble.

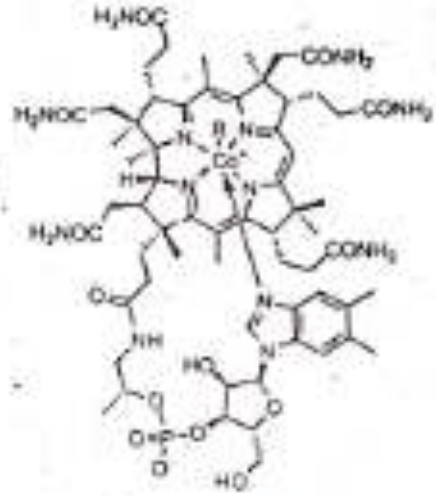
Deficiency may cause megaloblastic anemia, a condition where bone marrow produces unusually large, abnormal, immature red blood cells.
Good sources include: fish, shellfish, meat, poultry, eggs, milk and dairy products, some fortified cereals and soy products, as well as fortified nutritional yeast.
Vegans are advised to take B_{12} supplements.



Vitamin B6 (Peridoxine)



B9 (Folate)



B12 (Cobalamine)

EXERCISE

LONG QUESTIONS

1. Discuss nature, physiological role and sources of vitamin B complex.

2. Write note on characteristic, function and occurrence of fat soluble vitamins.

SHORT QUESTIONS

1. What are vitamins? Why are they necessary for our body?

Ans: The essential organic compounds which are needed in very small amount in the body are called vitamins. Normal metabolic activity of animals depends on vitamins. Vitamins are chemically unrelated, organic substances. They occur in many foods in small amounts. They are necessary for normal metabolic functioning. There are two types of vitamins.

2. What are water soluble and fat soluble vitamins?

Ans:

a) **Water soluble vitamins:** Most water-soluble vitamins act as coenzymes. They are needed in metabolism. Examples: vitamins B and vitamin C.

b) **Fat soluble vitamins:** The far-soluble vitamins perform different functions.

3. What are requirement of Vitamins in vertebrates?

Ans: The vitamin C and fat soluble vitamins (A, D, E and K) are need in limited amount in the vertebrates. Different animals have different requirements of vitamins. For example, humans and guinea pigs require vitamin C. But rabbits do not require it. Some birds require vitamin A. But other birds do not require it.

4. What is Thiamine? Write its role.

Ans: See from text of this chapter.

5. Write role, sources and deficiency of Vitamin B12.

Ans: See from text of this chapter.

6. What are function water soluble vitamins? Give their examples.

Ans: Most water-soluble vitamins act as coenzymes. They are needed in metabolism. Examples: vitamins B and vitamin C.

7. What are the requirements of Vitamins in vertebrates?

Ans: The vitamin C and fat soluble vitamins (A, D, E and K) are need in limited amount in the vertebrates. Different animals have different requirements of vitamins. For example, humans and guinea pigs require vitamin C. But rabbits do not require it. Some birds require vitamin A. But other birds do not require it.

8. Name water soluble vitamins.

Ans: Thiamin (vitamin B_1), Riboflavin (vitamin B_2), Niacin (nicotinic acid) , Vitamin B_6, Pantothenic acid, Cyanocobalamin (vitamin B_{12}), Folate (folic acid), Biotin and Ascorbic acid.

9. What are the functions of vitamin B_1?

Ans: They are part of coenzyme needed for oxidation of carbohydrates.

10. Name different fat soluble vitamins.

Ans: Vitamin A, D, E and K.

11. What are the sources of vitamin D?

Ans: Produced in skin exposed to ultraviolet light; in milk, egg yolk, fish-liver oils, and fortified foods.

MULTIPLE CHOICE QUESTIONS (MCQs)

1. The vitamin produced in skin is:

(a) A (b) K (c) C (d) D

2. Which of the following vitamins is necessary for synthesis of visual pigment?

(a) A (b) K (c) C (d) D

3. Which of the following is water soluble vitamin?

(a) A (b) B (c) K (d)

4. Thiamin is:

(a) Vitamins A (b) B (c) B1 (d) 12

5. Riboflavin is:

(a) Vitamins B2 (b) B (c) B1 (d) 12

6. Folic acid is type of Vitamin:

(a) A (b) K (c) C (d) B

7. Beriberi is caused by deficiency of:

(a) A (b) B (c) C (d) D

Ans:

1. (b) 2. (a) 3. (b) 4. (c) 5. (a) 6. (d) 7. (b)

Chapter 7

NUCLEIC ACIDS

Syllabus 2012-13
Not included in this syllabus
Syllabus 2018-19
Nucleic acids a. Nucleic acids and their types; Structure and role of various Bases in nucleic acids, b. Nucleoside & Nucleotides; c. Structure of DNA and RNA molecules; d. Organization and Chemistry of Double helical structure of DNA with their details.

NUCLEIC ACIDS AND THEIR TYPES

Nucleic acid is polymer of nucleotides. Nucleic acids are very important compounds in the cells. They act as hereditary material. They also regulate and control all the activities of the cell. There are two types of nucleic acids found in living cells. These are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Most of the DNA is present in nucleus. Small amount of DNA is also present in chloroplast and mitochondria. RNA is found in both nucleus and cytoplasm.

Discovery

A German Chemist, Friedrich Miescher, discovered Nucleic in 1869. He discovered DNA only four years after the publication of Mendel's work. Miescher extracted a white substance from the nuclei of human cells and fish sperm. He called this substance **nuclein** as it was associated with the nucleus. The nuclein was acidic in character. So the nuclein was called **nucleic acid.**

Research revealed its significance in the last half of the 20th century. RNA was first made by laboratory synthesis in 1955. In 1965 the nucleotide sequence of tRNA was determined. In 1967 the synthesis of biologically active DNA was achieved. The amount of RNA varies from cell to cell. But the amount of DNA is normally constant for all

typical cells of a given species of plant or animal.

Significance

Nucleic acids are found in the chromosomes of living cells and viruses. They play a central role in the storage and replication of hereditary information. They have central role in expression of this information through protein synthesis. In most organisms, nucleic acids occur in combination with proteins. The combined substances are called **nucleoproteins**. Nucleic acid molecules are complex chains of varying length. The two chief types of nucleic acids are DNA (deoxyribonucleic acid). DNA carries the hereditary information from generation to generation. RNA (ribonucleic acid) provides instructions to ribosome for protein synthesis.

Chemical composition

P.A. Levene determined the basic structure of nucleus acid in 1920's. Nucleic acid molecules are made of repeating units called nucleotide.

Nucleotides

Each nucleotide is made up of three subunits pentose sugar, nitrogen base and phosphoric acid.

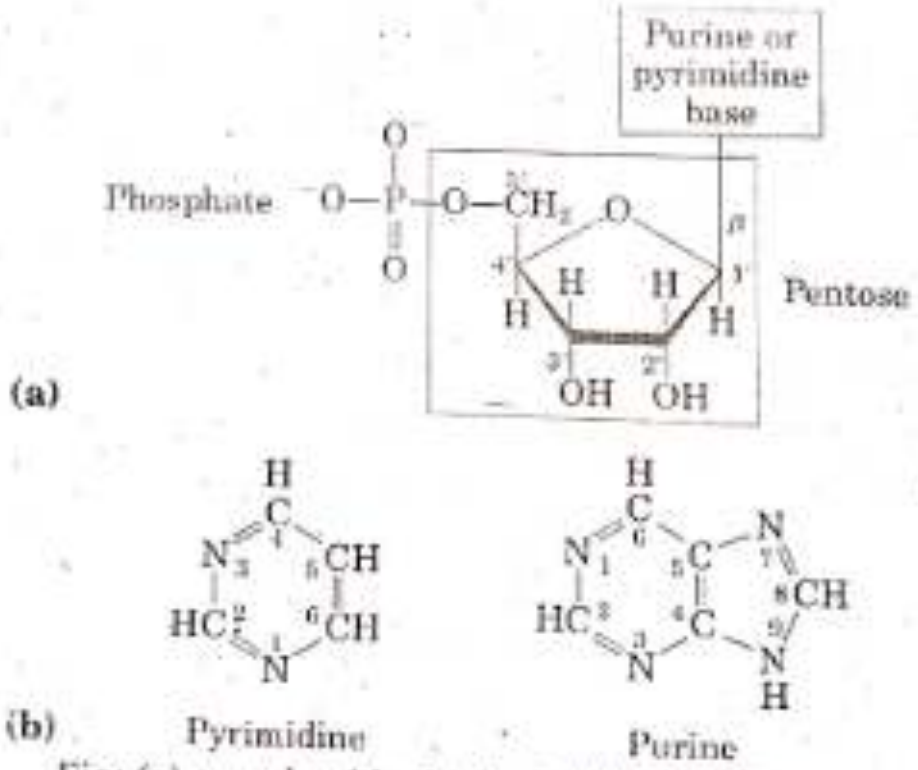


Fig: (a) a nucleotide (b) Pyrimidine and Purine

(a) Pentose sugar

Nucleic acids have two kinds of pentoses. The continuous deoxyribonucleotide units of DNA contain 2- deoxy-D-ribose. The ribonucleotide units of RNA contain D-ribose. In nucleotides, both types of pentoses are in their β -furanose (closed five-membered ring) form. The pentose ring is not planar but occurs in one of a variety of conformations generally described as puckered.

The DNA and RNA have two distinctions. They have different pentoses and presence of uracil in RNA and thymine in DNA. But actually, it is the pentoses that define the

identity of a nucleic acid. If the nucleic acid contains 2-deoxy-D-ribose, it is DNA by definition even though it may contain some uracil. Similarly, if the nucleic acid contains D-ribose, it is RNA regardless of its base composition.

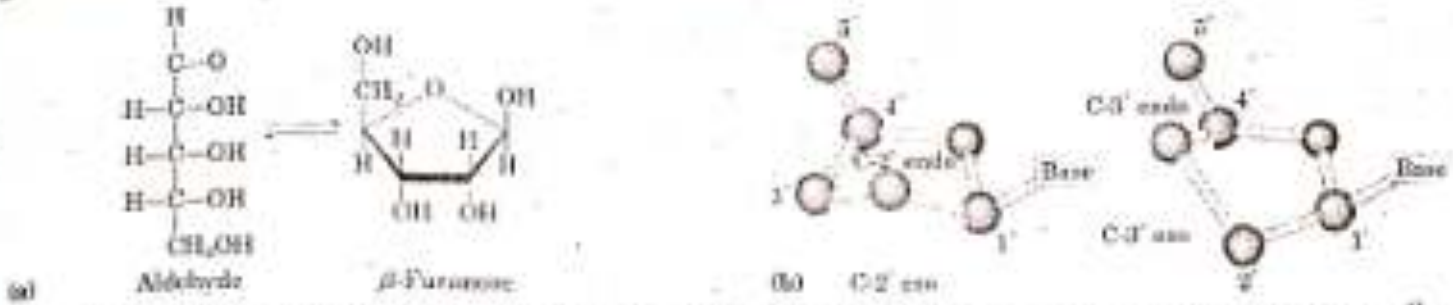


Fig: Conformations of ribose. (a) In solution, the straight chain (aldehyde) and ring (β furanose) forms of free ribose are in equilibrium. RNA contains only the ring form, β-D-ribofuranose. Deoxyribose undergoes a similar interconversion in solution. But in DNA it exists solely as β-2-deoxy-D-ribofuranose. (b) Ribofuranose rings in nucleotides can exist in four different puckered conformations. In all cases, four of the five atoms are in a single plane. The fifth atom (C-2 or C-3) is on either the same (endo) or the opposite (exo) side of the plane relative to the C-5 atom.

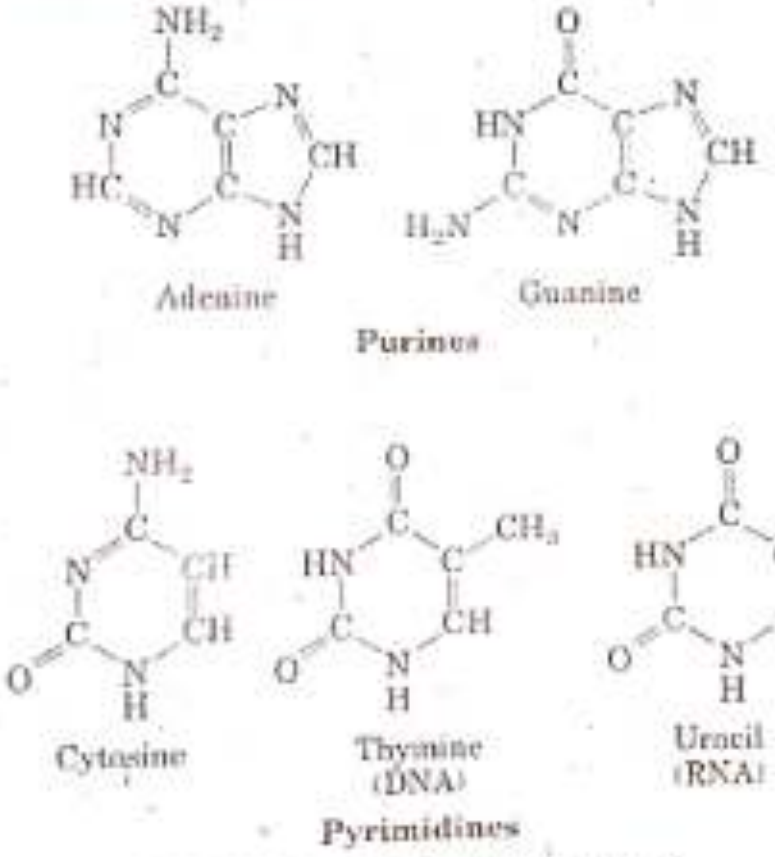


Fig: Different Nitrogenous Bases

(b) Nitrogen bases

Nitrogen bases are of two types.

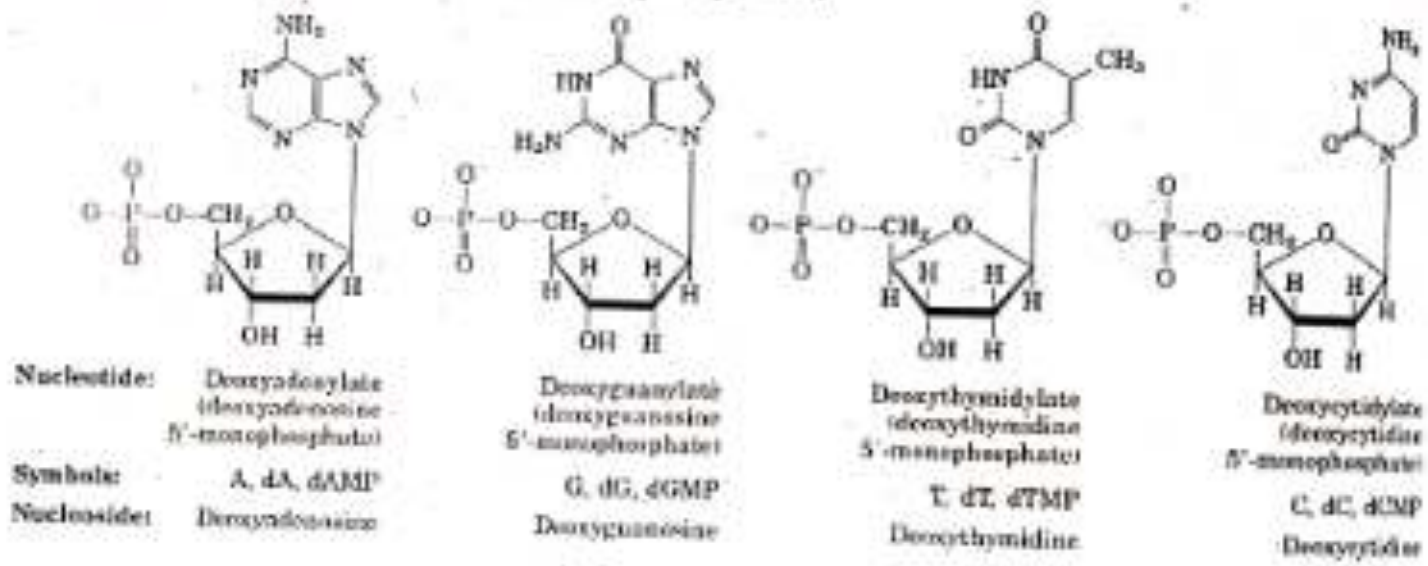
a) **Pyrimidines:** These are composed of single ring. These are **Cytosine (C), Thymine (T) and Uracil (U).**

b) **Purines:** These are composed of double rings. These are **Adenine (A) and Guanine (G).**

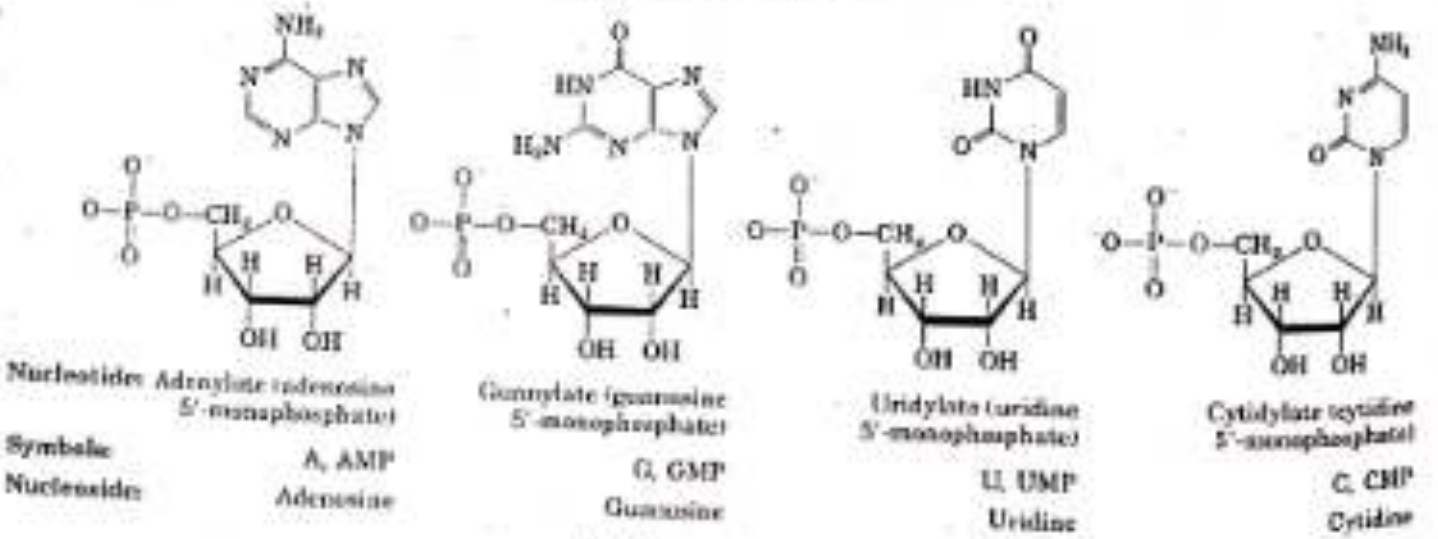
(c) Phosphoric acid (H_3PO_4):

It has the ability to form ester linkage with OH group of Pentose sugar. The nitrogen base is attached to the **carbon 1** of the pentose sugar in a typical nucleotide. The phosphoric acid is attached with the **carbon 3** of the pentose sugar.

The structural units of DNAs have four major deoxyribonucleotides (deoxyribonucleoside 5-monophosphates). The structural units of RNAs have four major ribonucleotides (ribonucleoside 5-monophosphates).



(a) Deoxyribonucleotides



(b) Ribonucleotides

Minor Basis

The nucleotides bearing the major purines and pyrimidines are most common. But both DNA and RNA also contain some minor bases. In DNA the most common of these are methylated forms of major bases. In some viral DNAs, certain bases may be hydroxymethylated or glucosylated. Altered or unusual bases in DNA molecules have roles in regulating or protecting genetic information. Minor bases of many types are also found in RNAs, especially in tRNAs.

Cells also contain nucleotides with phosphate groups in positions other than 5 carbon:

- Ribonucleoside 2,3-cyclic monophosphates are isolatable intermediates

ribonucleoside

- 3 - monophosphates are end products of hydrolysis of RNA by certain ribonucleases.
- Other variations are adenosine 3,5-cyclic monophosphate (cAMP) and guanosine 3,5-cyclic monophosphate (cGMP).

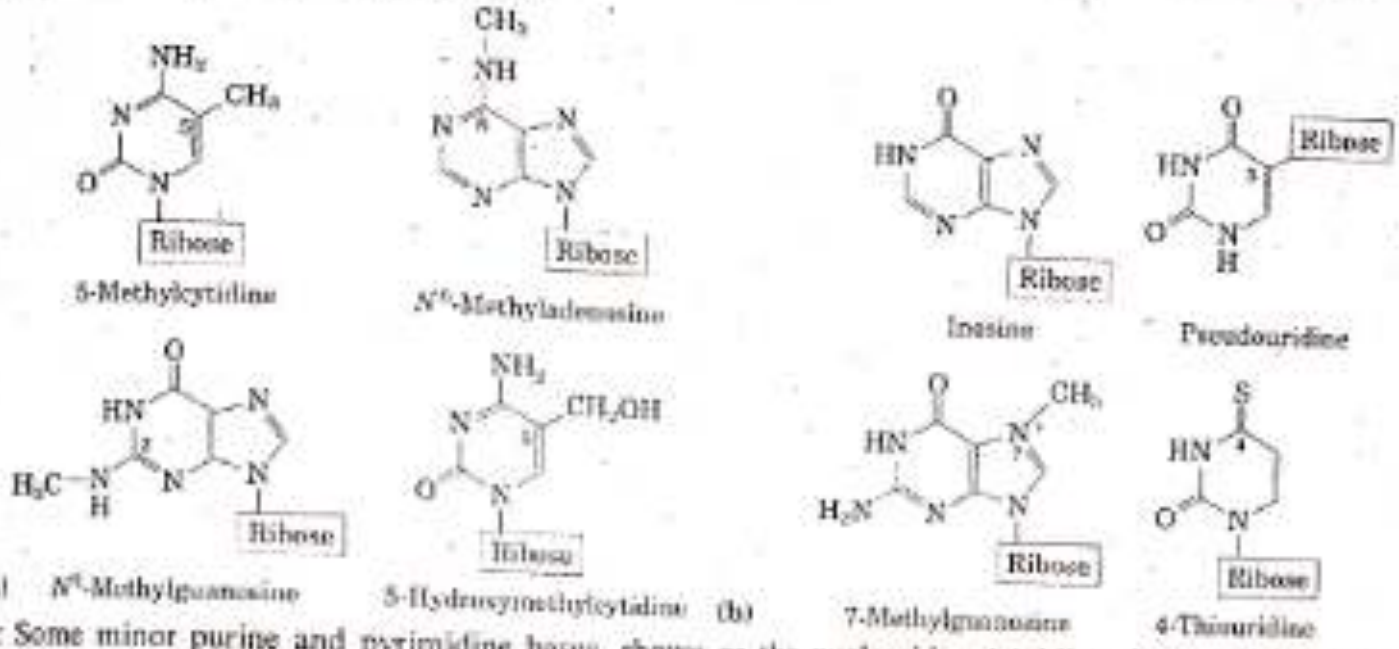


Fig: Some minor purine and pyrimidine bases, shown as the nucleosides. (a) Minor bases of DNA. (b) Some minor bases of tRNAs

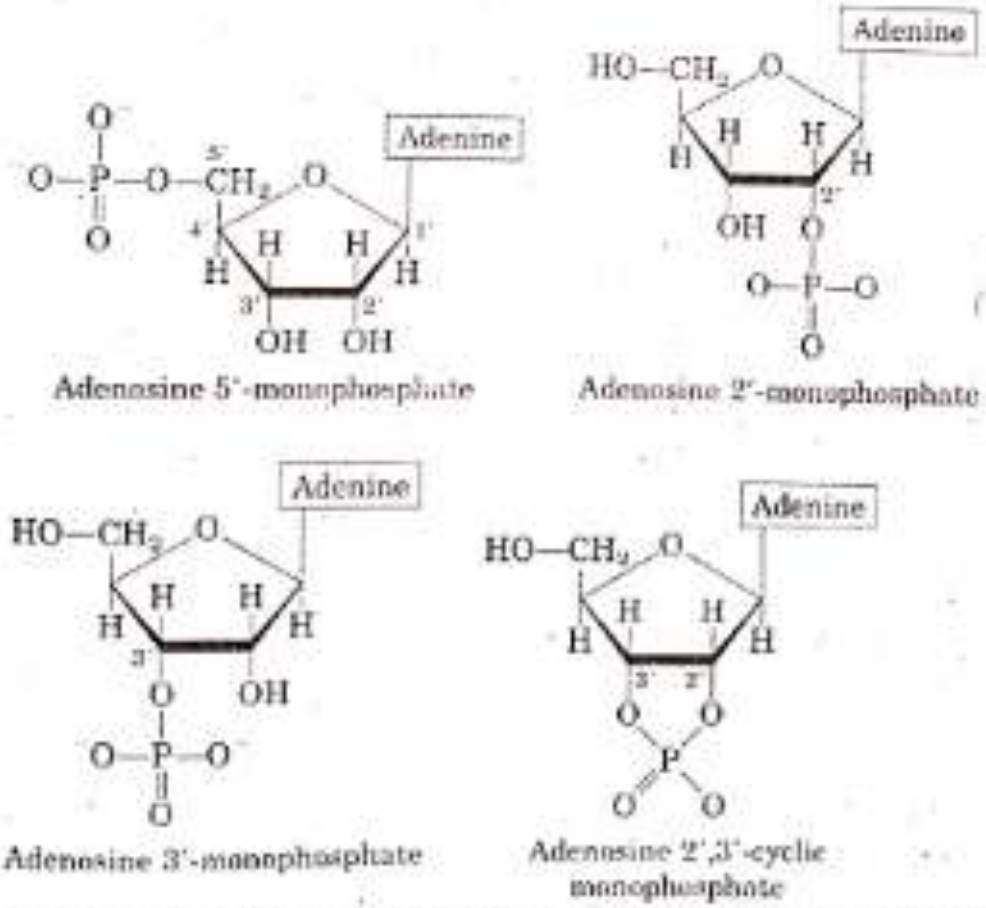
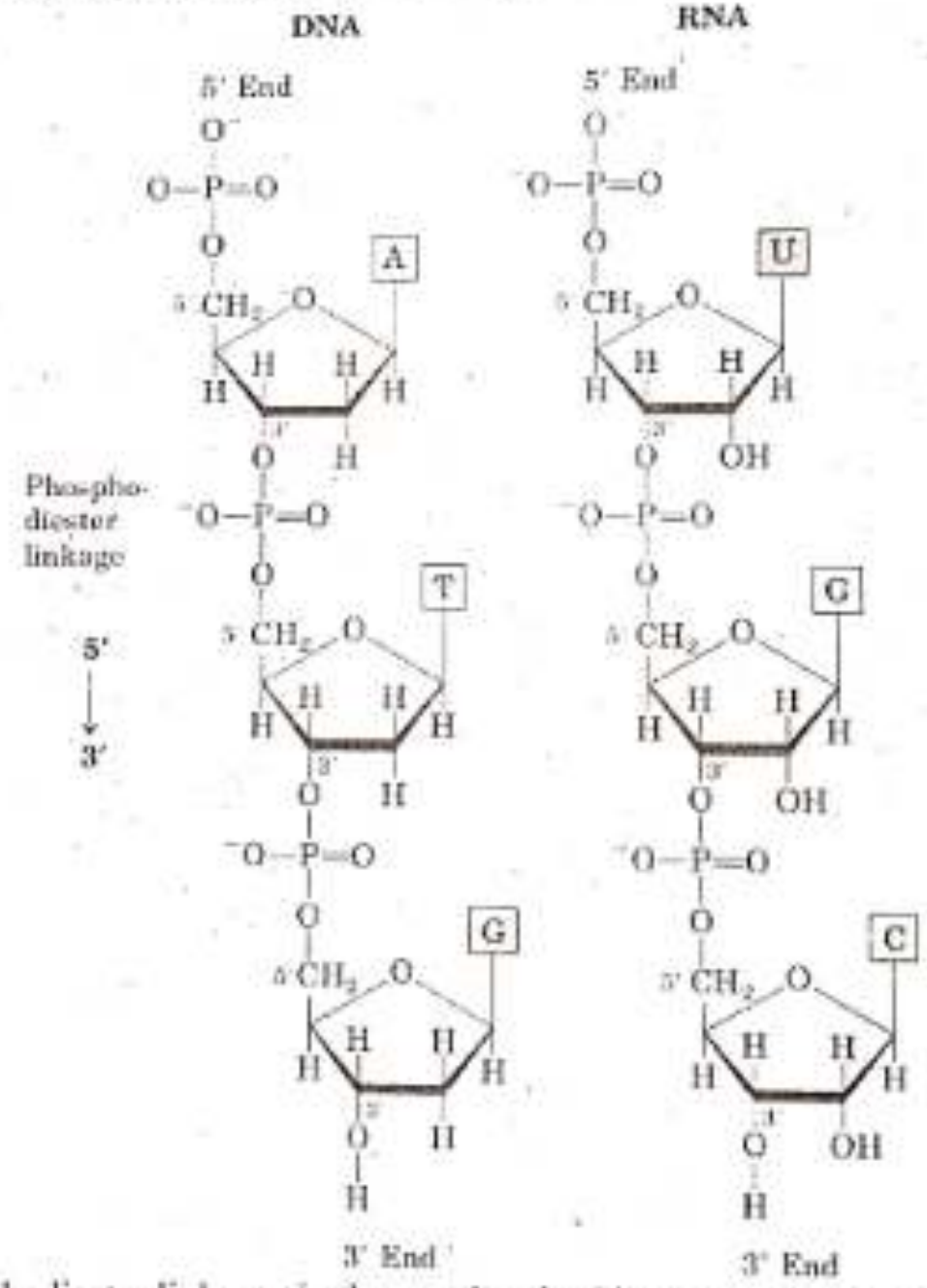


Fig: Some adenosine monophosphates. Adenosine 2-monophosphate, 3-monophosphate, and 2,3-cyclic monophosphate are formed by enzymatic and alkaline hydrolysis of RNA.

Nucleoside

The compound formed by the nitrogenous base and pentose sugar is called nucleoside. A nucleoside and phosphoric acid combine to form a nucleotide. [2,8] ATP (adenosine triphosphate) is also an important nucleotide. The cells use ATP as energy currency.

Erwin Chargaff proved that the amount of adenine in DNA always equals the amount of thymine. Similarly amount of guanine always equals the amount of cytosine. Thus there is always equal proportion of purine (A+G) and pyrimidine (C+T) in DNA.



Phosphodiester linkages in the covalent backbone of DNA and RNA

Phosphodiester Bond Formation

1. Nitrogen base is attached to carbon number 1 of a pentose sugar in nucleotide.
2. Phosphate group is attached to carbon number 5 of the sugar.

3. A free hydroxyl (OH) group is attached to the 3' carbon atom. The 5' phosphate and 3 hydroxyl groups react chemically with each other. A covalent bond is formed between them. It links two nucleotides. This reaction releases a water molecule. So it is called **dehydration reaction**. It allows DNA and RNA to form long chains of nucleotides.
4. The phosphate group is linked to the two sugars by a pair of ester (P-O-C) bonds. Thus the linkage between two nucleotides is called a **phosphodiester bond.**
5. The two unit polymer of this reaction still has a free 5' phosphate group at one end and a free 3' hydroxyl group at the other. Thus it can link to other nucleotides. In this way, many thousands of nucleotides join together in long chains.
6. Linear strands of DNA or RNA always have a free 5' phosphate group at one end and a free 3' hydroxyl group at the other.
7. A short nucleic acid is called as **oligonucleotide**. The polymers containing 50 or fewer nucleotides are generally called oligonucleotides. A longer nucleic acid is called a **polynucleotide**.

Representation of Nucleotide Sequence

The nucleotide sequences of nucleic acids can be represented schematically. The phosphate groups are symbolized by (P).Each deoxyribose is symbolized by a vertical line, from C-1 at the top to C-5 at the bottom. The connecting lines between nucleotides (which pass through P) are drawn diagonally from middle (C-3) of the deoxyribose of one nucleotide to the bottom (C-5) of the next.

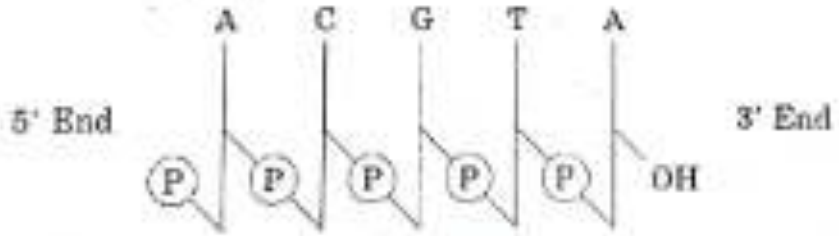


Fig: Some simpler representations of this pentadeoxyribonucleotide are pA-C-G-T-AOH, pApCpGpTpA, and pACGTA

Properties of Nucleotide Bases Affect the Three- Dimensional Structure of Nucleic Acids

Free pyrimidines and purines are weakly basic compounds. Hence they are called bases. The purines and pyrimidines common in DNA and RNA are aromatic molecules. This property affects the structure, electron distribution, and light absorption of nucleic acids.

1. **Delocalization of electrons:** There is delocalization of electrons among atoms in the ring. It gives most of the bonds partial double-bond character. Its one result is that pyrimidines are planar molecules and purines are very nearly planar, with a slight pucker.
2. **Tautomeric forms:** Free pyrimidine and purine bases may exist in two or more tautomeric forms. Such existence depends on pH. For example uracil, occurs in lactam, lactim, and double lactim forms. The structures shown in Figure are the tautomers that predominate at pH 7.0. All nucleotide bases absorb UV light. The

nucleic acids are characterized by a strong absorption at wavelengths near 260 nm.

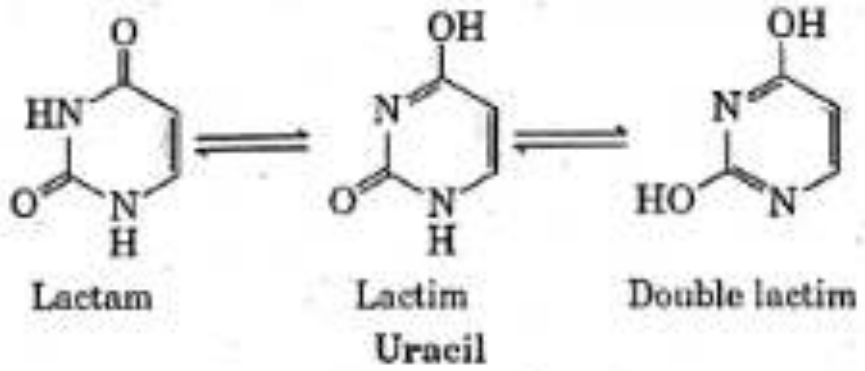


Fig: Tautomeric forms of uracil

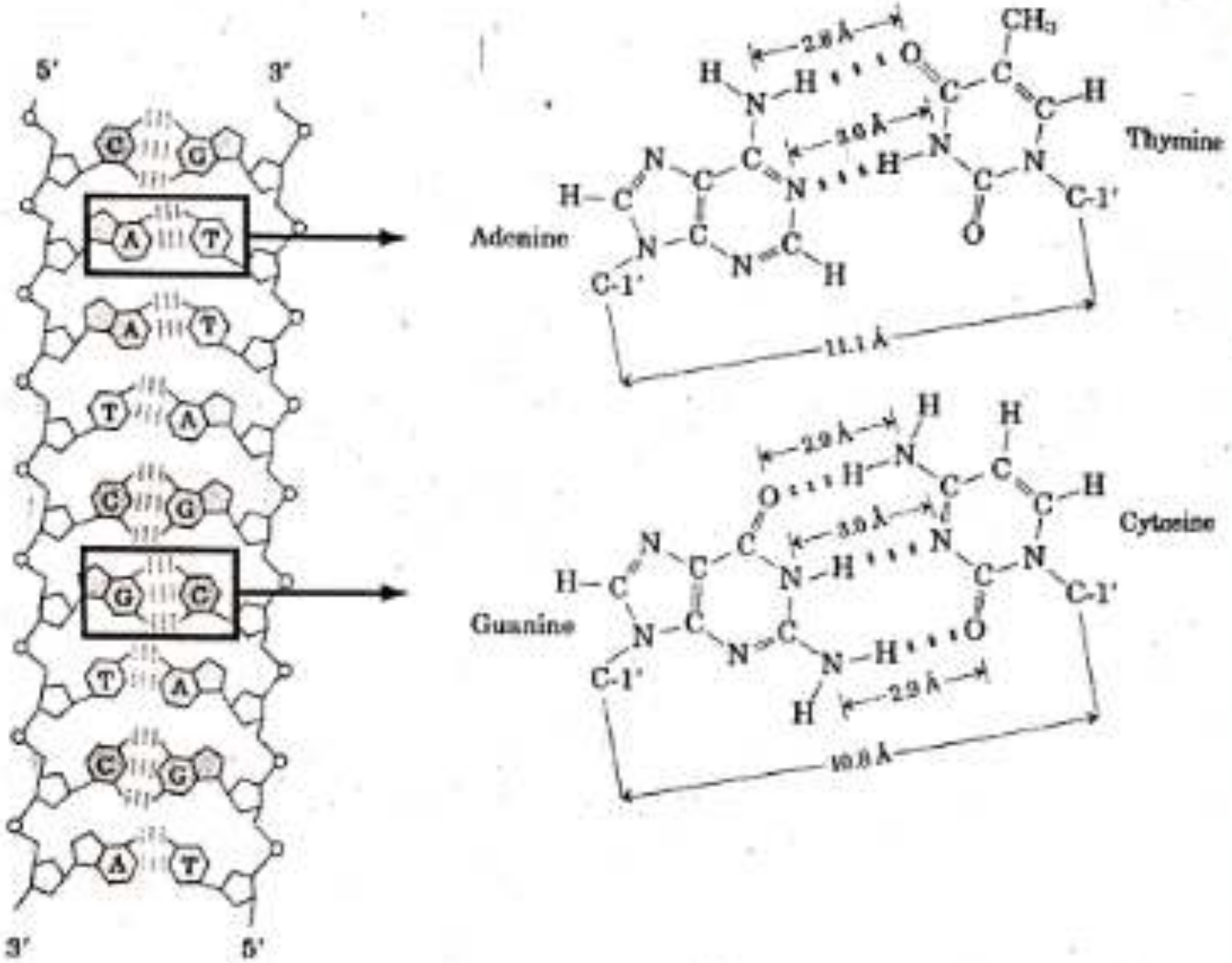


Fig: Hydrogen-bonding patterns in the base pairs defined by Watson and Crick.

3. **Hydrophobic nature:** The purine and pyrimidine bases are hydrophobic. They are relatively insoluble in water at the near neutral pH of the cell. At acidic or alkaline pH the bases become charged and their solubility in water increases.
4. **Hydrophobic stacking interactions:** Hydrophobic stacking interactions occur in nucleic acid. In these interactions, two or more bases are positioned with the planes of

their rings parallel (like a stack of coins). These are one of two important modes of interaction between bases in nucleic acids. The stacking also involves a combination of van der Waals and dipole-dipole interactions between the bases. Base stacking helps to minimize contact of the bases with water. The base-stacking interactions are very important in stabilizing the three dimensional structure of nucleic acids.

5. **Hydrogen bonding:** The functional groups of pyrimidines and purines are ring nitrogens, carbonyl groups, and exocyclic amino groups. Hydrogen bonds involve the amino and carbonyl groups. These bonds are the most important mode of interaction between two (and occasionally three or four) complementary strands of nucleic acid. The most common hydrogen-bonding patterns are defined by James D. Watson and Francis Crick in 1953. In this bonding, A bonds specifically to T (or U) and G bonds to C. These two types of base pairs predominate in double-stranded DNA and RNA. Tautomers are responsible for these patterns. It is this specific pairing of bases that permits the duplication of genetic information.

NUCLEIC ACID STRUCTURE

The discovery of the structure of DNA by Watson and Crick in 1953 was great event in science. It gave rise to entirely new disciplines and influenced the course of many established ones. The DNA structure is described at different levels. These are primary, secondary and tertiary structures.

a) **Primary structure:** The primary structure of a nucleic acid is its covalent structure and nucleotide sequence.

b) **Secondary structure:** Any regular, stable structure taken up by nucleotides in a nucleic acid is called secondary structure. All structures other than primary structures come under the heading of secondary structure.

c) **Tertiary structure:** The complex folding of large chromosomes within eukaryotic chromatin and bacterial nucleoids is tertiary structure.

STRUCTURE OF DNA

DNA Stores Genetic Information

1. **Friedrich Miescher:** DNA was first isolated and characterized by Friedrich Miescher in 1868. He called the phosphorus-containing substance **nuclein.**
2. **Avery and coworkers:** In 1940s, Oswald T. Avery, Colin MacLeod, and Maclyn McCarty proved that DNA was the genetic material. Avery and his colleagues took DNA extracted from a virulent (disease-causing) strain of the bacterium *Streptococcus pneumonia*. They injected it into a non-virulent strain of the same bacterium. It transformed the non-virulent strain into a virulent strain. They concluded that DNA from the virulent strain carried the genetic information for virulence.
3. **Alfred D. Hershey and Martha Chase:** They in 1952 studied the infection of bacterial cells by a virus (bacteriophage). They radioactively labeled DNA and protein. They found that DNA was transferred not protein to next generation. Thus DNA carries the genetic information.

Work of Erwin Chargaff

Erwin Chargaff and his colleagues worked on DNA in the late 1940s. They provide important clue about structure of DNA. They found that the four nucleotide bases of DNA occur in different ratios in the DNAs of different organisms. The amounts of certain bases are closely related. These data, collected from DNAs of a great many different species. It helped Chargaff to form following conclusions:

1. The base composition of DNA generally varies from one species to another.
2. DNA specimens isolated from different tissues of the same species have the same base composition.
3. The base composition of DNA in a given species does not change with an organism's age, nutritional state, or changing environment
4. In all cellular DNAs, number of adenosine residues is equal to number of thymidine residues (that is, A = T). Similarly, number of guanosine residues is equal to the number of cytidine residues (G = C). Thus sum of the purine residues equals the sum of the pyrimidine residues. Hence A + G = T + C.

Significance

These quantitative relationships are called **Chargaff's rules**. These were later confirmed by many subsequent researchers. They were a key to establishing the three- dimensional structure of DNA. It also provided clues to how genetic information is encoded in DNA and passed from one generation to the next.

Work of Rosalind Franklin and Maurice Wilkins

Rosalind Franklin and Maurice Wilkins used powerful method of x-ray diffraction to analyze DNA fibers. They showed in the early 1950s that DNA produces a characteristic x-ray **diffraction pattern**. From this pattern it was deduced that DNA molecules are helical. It has two periodicities along their long axis. These include a primary one of 3.4 Å and a secondary one of 34 Å. The problem then was to formulate a three-dimensional model of the DNA molecule. This model should describe x-ray diffraction data, specific A =T and G = C base equivalences and other chemical properties of DNA.

Watson and Crick Model of DNA

James Watson and Francis Crick used above information about DNA to formulated its structure. In 1953 they postulated a three-dimensional model of DNA structure. It explained all the available data. It has following points:

1. **Helical shape:** The DNA is helical in shape. It is made up of two strands. These strands form a right-handed double helix. The helix has a uniform width of 2 nm.
2. **Arrangement of nitrogenous bases:** Its nitrogenous bases are 0.34 nm apart. The hydrophilic backbones of alternating deoxyribose and phosphate groups are on the outside of the double helix. They face the surrounding water. The furanose ring of each deoxyribose is in the C-2 -endo conformation. The purine and pyrimidine bases of both strands are stacked inside the double helix. Their hydrophobic and nearly planar ring structures are very close together and perpendicular to the long axis. The

offset pairing of the two strands creates a **major groove** and **minor groove** on the surface of the duplex. Both These grooves alternate with each other.

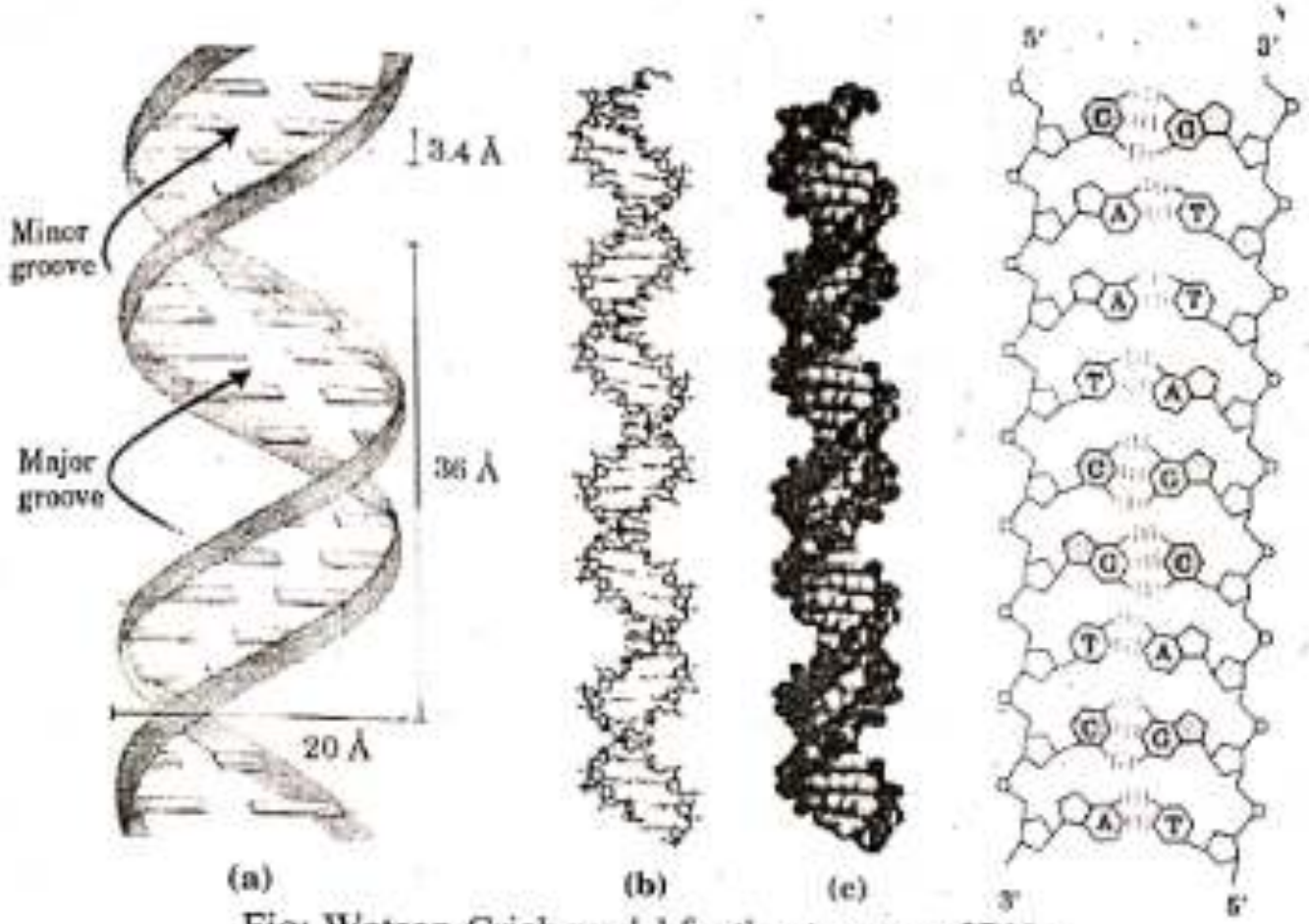


Fig: Watson-Crick model for the structure of DNA

3. **Ladder like appearance:** The double helix is a ladder like. It has rigid rungs. Its ladder twists in a spiral fashion. The side ropes are the equivalent sugar—phosphate backbones. The rungs are pairs of nitrogenous bases.
4. **Complementary base paring:** The pairing of the nitrogenous bases is complement. Adenine pair with thymine (T) and guanine (C) with cytosine. Adenine and Guanine are larger bases. They have two ringed structures. They are called **purine** base. On the other hand, the cytosine and thymine are **pyrimidine** bases. They have single ring. In this way one purine and one pyrimidine combination keep the diameter of the DNA uniform.
5. **Antiparallel strands:** Both strands of the DNA are anti parallel. One strand is in 5 – 3 direction. The other strand is in 3 – 5 strands. Later work with DNA polymerases provided experimental evidence that the strands are indeed antiparallel. This was finding ultimately confirmed by x-ray analysis.
6. **Explanation of periodicities observed in the x-ray diffraction patterns:** Watson and Crick also explained the periodicities observed in the x-ray diffraction patterns of DNA fibers. For this purpose, they manipulated molecular models to describe this structure. According to it, vertically stacked bases inside the double helix are 3.4 Å apart. The secondary repeat distance of about 34 Å occurs by the presence of 10 base

pairs in each complete turn of the double helix. In aqueous solution the structure differs slightly from that in fibers. In this forms, it has10.5 base pairs per helical turn.

7. Non-identical antiparallel polynucleotide chains: The two antiparallel polynucleotide chains of double-helical DNA are not identical in either base sequence or composition. Instead they are complementary to each other. Wherever adenine occurs in one chain, thymine is found in the other. Similarly, wherever guanine occurs in one chain, cytosine is found in the other.
8. **Forces holding helixes:** The DNA double helix, or duplex. These helixes are held together by two forces. These are hydrogen bonding between complementary base pairs and base-stacking interactions.
   - The complementarity between DNA strands is due to hydrogen bonding between base pairs.
   - The base-stacking interactions are largely non-specific with respect to the identity of the stacked bases. But these interactions have major contribution to the stability of the double helix.

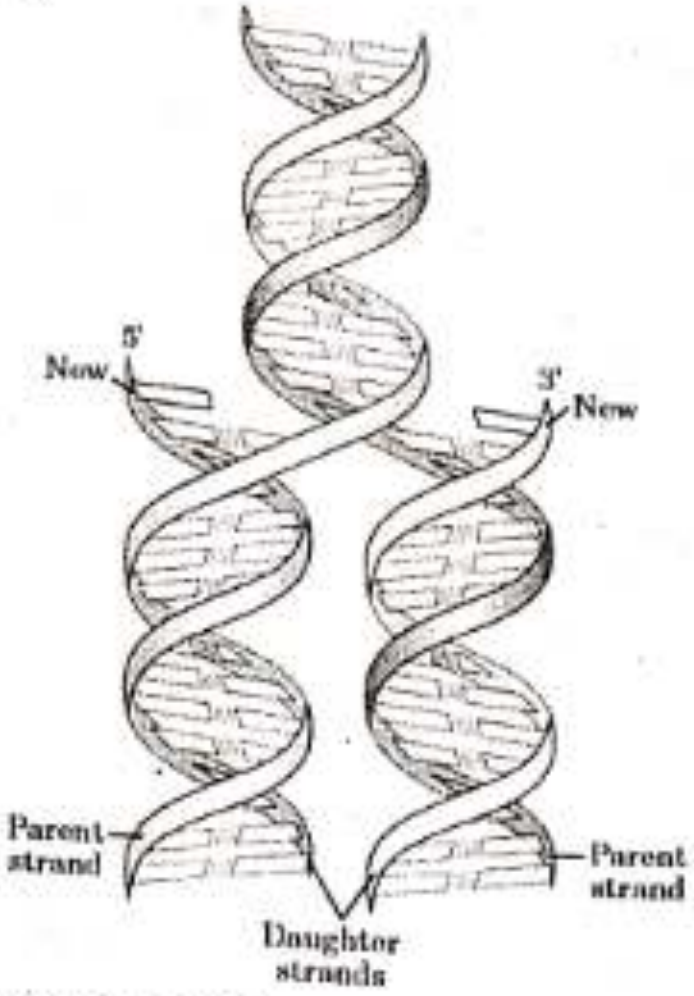


Fig: Replication of DNA as suggested by Watson and Crick

Significance of Model

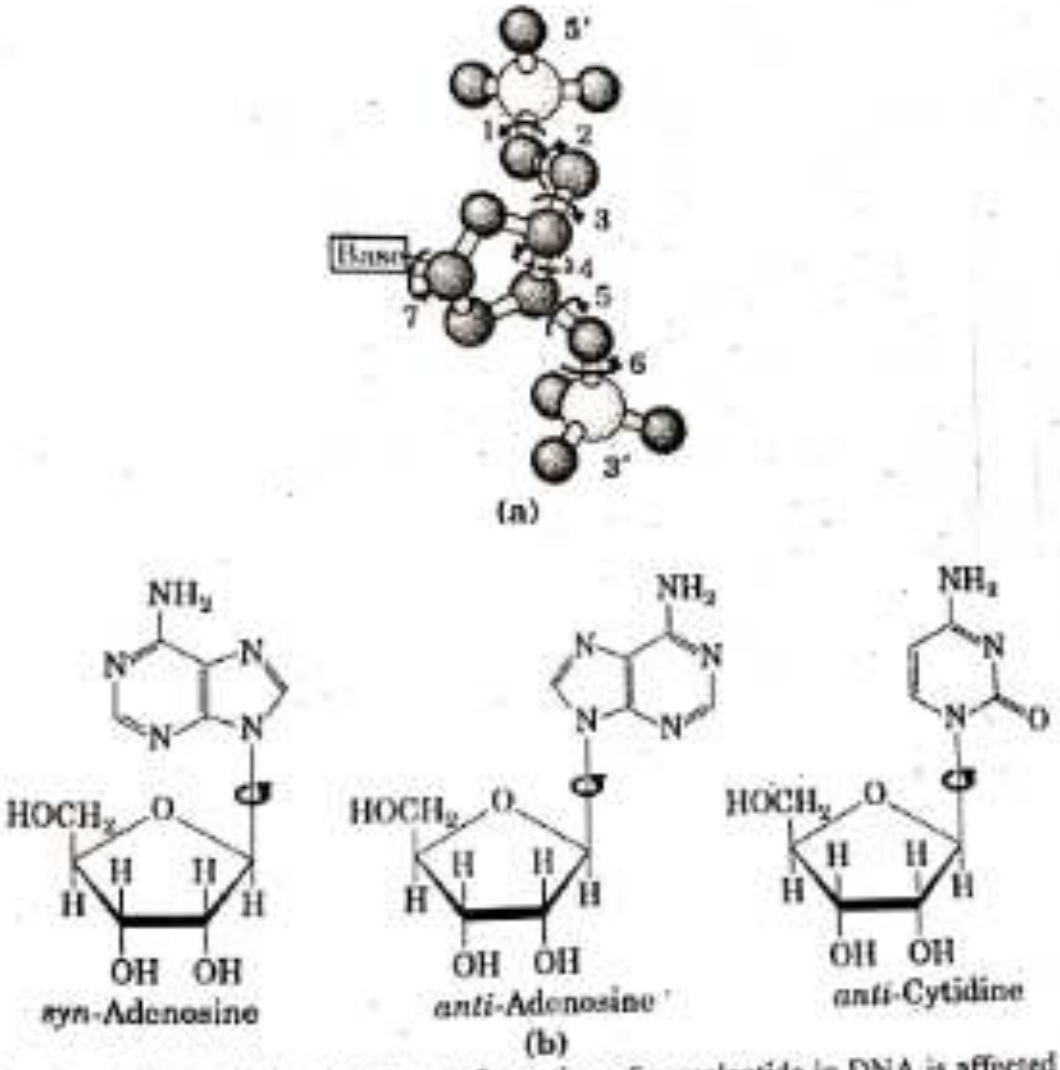
This model was a revolution in our understanding of biological inheritance.

a) The important features of the double-helical model of DNA structure are supported by much chemical and biological evidence.

b) Moreover, model immediately suggested a mechanism for the transmission of genetic information.

c) The essential feature of the model is complementarity of the two DNA strands. It explains the separating the two strands and synthesizing a complementary strand for each.

d) The nucleotides in each new strand are joined in a sequence specified by the base-pairing rules stated above. Therefore, each preexisting strand functions as a template to guide the synthesis of one complementary strand. These expectations were experimentally confirmed.



e)

f) Fig: Structural variation in DNA. (a) The conformation of a nucleotide in DNA is affected by rotation about seven different bonds. (b) For purine bases in nucleotides, only two conformations with respect to the attached ribose units are sterically permitted, anti or syn. Pyrimidines generally occur in the anti conformation.

STRUCTURAL VARIATIONS IN DNA

DNA Can Occur in Different Three-Dimensional Forms

DNA is a remarkably flexible molecule. Considerable rotation is possible around several

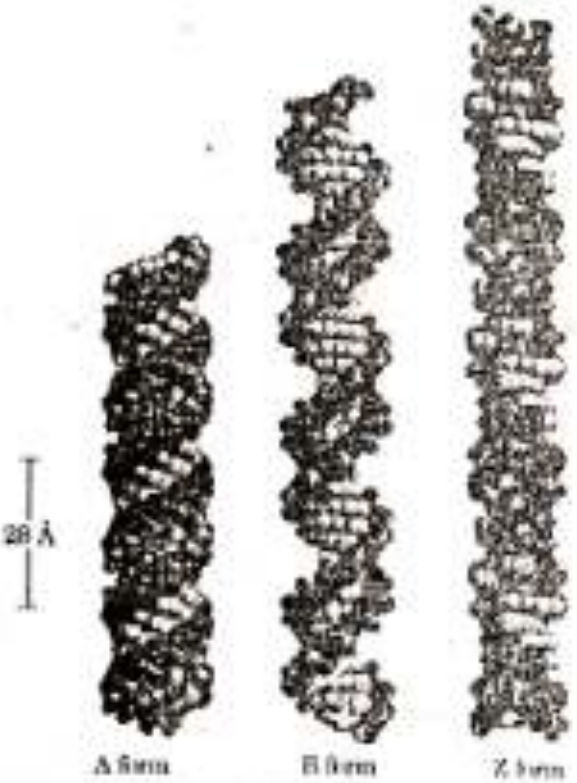
types of bonds in the sugar–phosphate (phosphodeoxyribose) backbone. The thermal fluctuation can produce bending, stretching, and unpairing (melting) of the strands. Many significant deviations from the Watson-Crick DNA structure are found in cellular DNA. Some or all of which may be important in DNA metabolism. These structural variations generally do not affect the properties of DNA defined by Watson and Crick. These properties are strand complementarity, antiparallel strands, and requirement for A= T and G≡C base pairs.

Structural variation in DNA reflects three things:

1. Different possible conformations of the deoxyribose.
2. Rotation about the contiguous bonds that make up the phosphodeoxyribose backbone
3. Free rotation about the C-1–N-glycosyl bond.

There are steric constraints in bases. Therefore, purines in purine nucleotides are restricted to two stable conformations with respect to deoxyribose. This is called syn and anti. Pyrimidines are generally restricted to anti conformation. This occurs due to steric interference between sugar and carbonyl oxygen at C-2 of pyrimidine.

Thus there are three forms of DNA.



	A form	B form	Z form
Helical sense	Right handed	Right handed	Left handed
Diameter	~25 Å	~20 Å	~18 Å
Base pairs per helical turn	11	10.5	12
Helix rise per base pair	2.6 Å	3.4 Å	3.7 Å
Base tilt normal to the helix axis	20°	6°	7°
Sugar pucker conformation	C-3' endo	C-2' endo	C-2' endo for pyrimidines; C-3' endo for purines
Glycosyl bond conformation	Anti	Anti	Anti for pyrimidines; syn for purines

(a) B-form DNA

The Watson-Crick structure is also called **B-form DNA**, or **B-DNA**. The B form is the most stable structure for a random-sequence DNA molecule under physiological conditions. Therefore, it is the standard point of reference in any study of the properties of DNA. There are two structural variants of DNA, **A and Z forms**. These are characterized in crystal structures. These three DNA conformations are shown in Figure with a summary of their properties.

(b) A-form DNA

The A form is favored in many solutions devoid of water. The DNA is still arranged in a right-handed double helix. But the the helix is wider. The number of base pairs per helical turn is 11, rather than 10.5 as in B- DNA. The plane of the base pairs in A-DNA is tilted about 20° with respect to the helix axis. These structural changes deepen the major groove. But it makes the minor groove shallower. The reagents used to promote crystallization of DNA tend to dehydrate it. Thus most short DNA molecules tend to crystallize in the A form.

(c) Z-form DNA

Z-form DNA is a more radical departure from B structure. Here most obvious distinction is the left- handed helical rotation. There are 12 base pairs per helical turn. Thus structure appears more slender and elongated. The DNA backbone takes on a zigzag appearance. Certain nucleotide sequences fold into left-handed Z helices much more readily than others. Prominent examples are sequences in which pyrimidines alternate with purines. It especially alternates C and G or 5-methyl-C and G residues. To form the left- handed helix in Z-DNA, purine residues flip to **syn conformation**. It alternates with pyrimidines in anti conformation. The major groove is barely apparent in Z-DNA. The minor groove is narrow and deep.

Occurrence of A-DNA in cells is uncertain. But there is evidence for some short stretches (tracts) of Z-DNA in both bacteria and eukaryotes. These Z- DNA tracts may play a role in regulating the expression of some genes or in genetic recombination.

Future Zoologist Group of PGCG (By Muhammad ARSALAN)

STRUCTURE OF RNA

Many RNAs have more complex three-dimensional structures. Messenger RNA is only one of several classes of cellular RNA. Transfer RNAs are adapter molecules in protein synthesis. It is covalently linked to an amino acid at one end. They pair with the mRNA in such a way that amino acids are joined to a growing polypeptide in the correct sequence. Ribosomal RNAs are components of ribosomes. There is also a wide variety of special-function RNAs, including some (called ribozymes). They have enzymatic activity. There are diverse and often complex functions of these RNAs. Therefore, they have diversity of structure much richer than that observed in DNA molecules.

1. **Primary Structure: Single stranded but assumes double stranded structure:** The product of transcription of DNA is always single-stranded RNA. The single strand tends to assume a right-handed helical conformation. It is dominated by base-stacking interactions. These interactions are stronger between two purines than between a purine and pyrimidine or between two pyrimidines. The purine-purine interaction is so strong that a pyrimidine may separate two purines. It is often displaced from the stacking pattern so that the purines can interact.
2. **Secondary structure: Formation of self-complementary sequences:** Any self-complementary sequences in the molecule produce more complex structures. RNA can base-pair with complementary regions of either RNA or DNA. Base pairing matches the pattern for DNA: G pairs with C and A pairs with U (or with occasional

T residue in some RNAs). One difference is that base pairing between G and U residues is unusual in DNA. But it is fairly common in RNA.

Fig: Primary structure

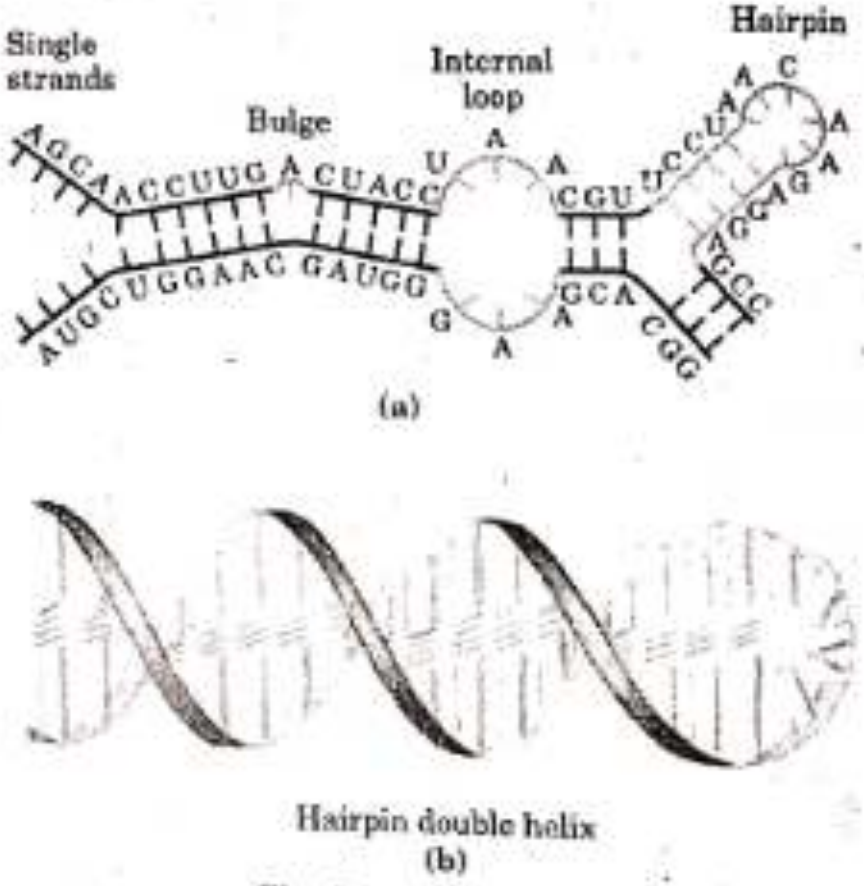


Fig: Secondary structure

3. **Antiparallel paring in strands:** The paired strands in RNA or RNA-DNA duplexes are antiparallel, as in DNA. RNA has no simple, regular secondary structure that serves as a reference point.
4. **Three-dimensional structures of RNAs:** The three-dimensional structures of many RNAs are complex and unique. Weak interactions occur especially in base-stacking interactions. These interactions help stabilize RNA structures just like DNA.
5. **Form of structures:** Where complementary sequences are present, the predominant double-stranded structure is an A- form right-handed double helix. Z-form helices have been made in the laboratory (under very high-salt or high-temperature conditions). The B form of RNA has not been observed.
6. **Breaks and formation of hairpin loops:** Breaks occur in the regular A-form helix. These are caused by mismatched or unmatched bases in one or both strands. Such breaks are common and result in bulges or internal loops. **Hairpin loops** form between nearby self-complementary sequences. The potential for base-paired helical structures in many RNAs is extensive. It results in hairpins. These are the most common type of secondary structure in RNA.
7. **Existence of specific short base sequences:** Specific short base sequences (such as UUCG) are often found at the ends of RNA hairpins. They form particularly tight and stable loops. Such sequences may act as starting points for the folding of an RNA

molecule into its precise three-dimensional structure.

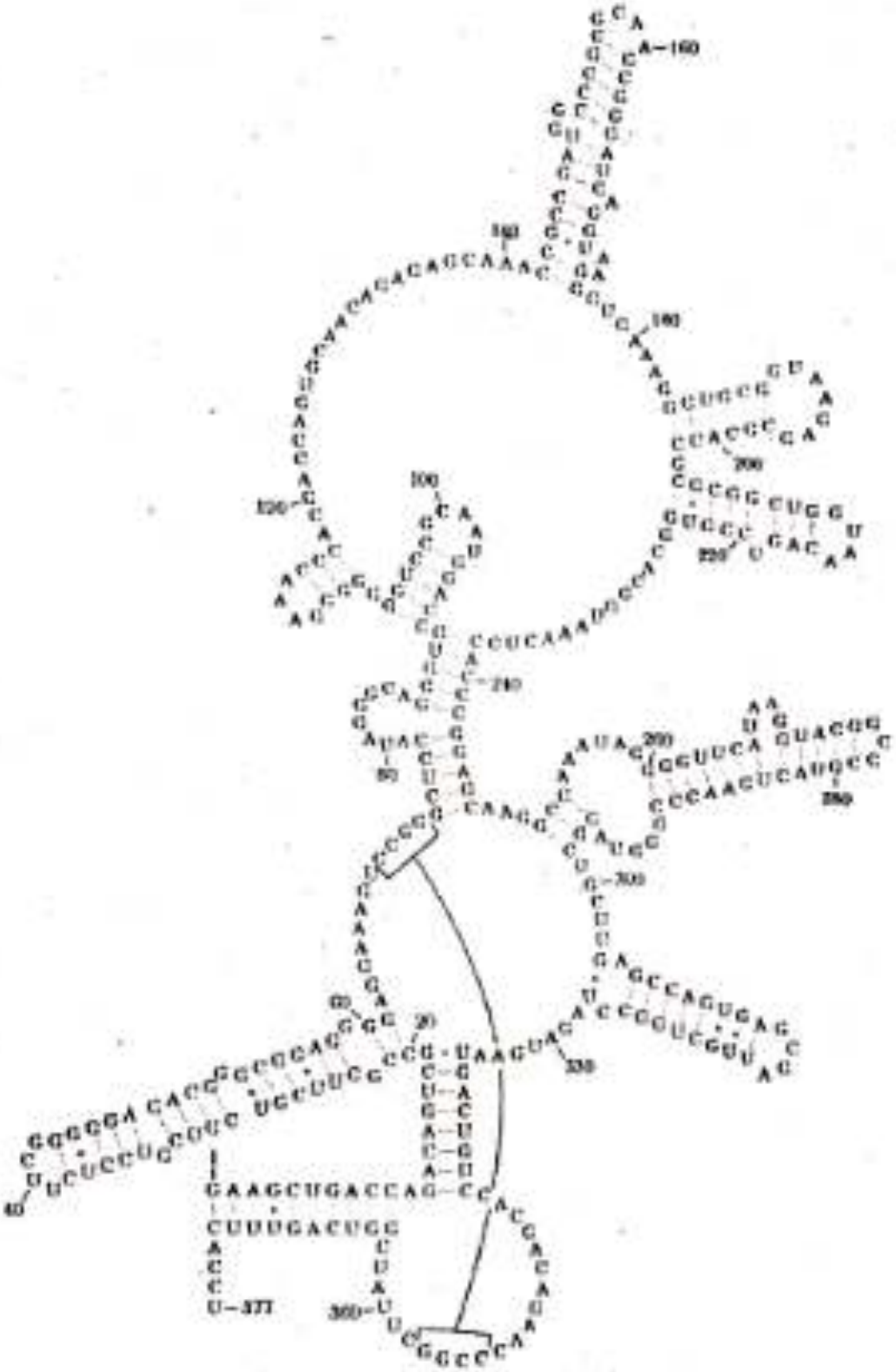


Fig: Base-paired helical structures in an RNA

8. **Role of hydrogen bonding:** Other contributions are made by hydrogen bonds. Such hydrogen bonds are not part of standard Watson-Crick base pairs. For example, 2'-hydroxyl group of ribose can hydrogen-bond with other groups. Some of these properties are evident in the structure of the phenylalanine transfer RNA of yeast. It is the tRNA responsible for inserting Phe residues into polypeptides. These also found

in two RNA enzymes, or ribozymes. The functions of ribozymes depend on their three-dimensional structures.

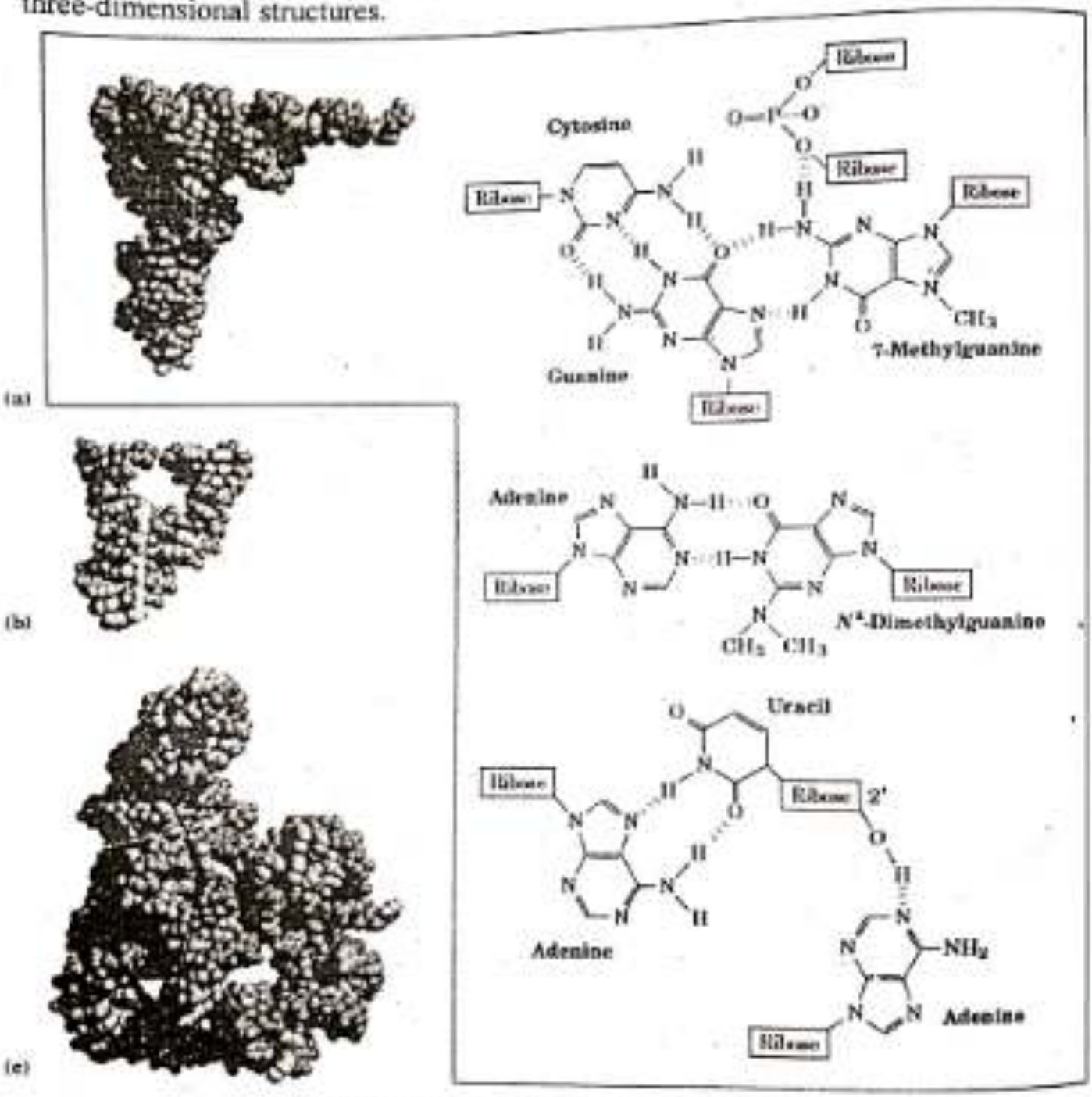


Fig: Three-dimensional structure in RNA

Types of RNA

There are three main types of RNA. These are messenger RNA (mRNA), transfer RNA (tRNA) and ribosomal RNA (rRNA). All three types of RNA are synthesized from DNA in nucleus. After their synthesis they are transferred to cytoplasm. All three RNA (mRNA, tRNA and rRNA) interact with each other. They synthesize proteins from the genetic information (gene).

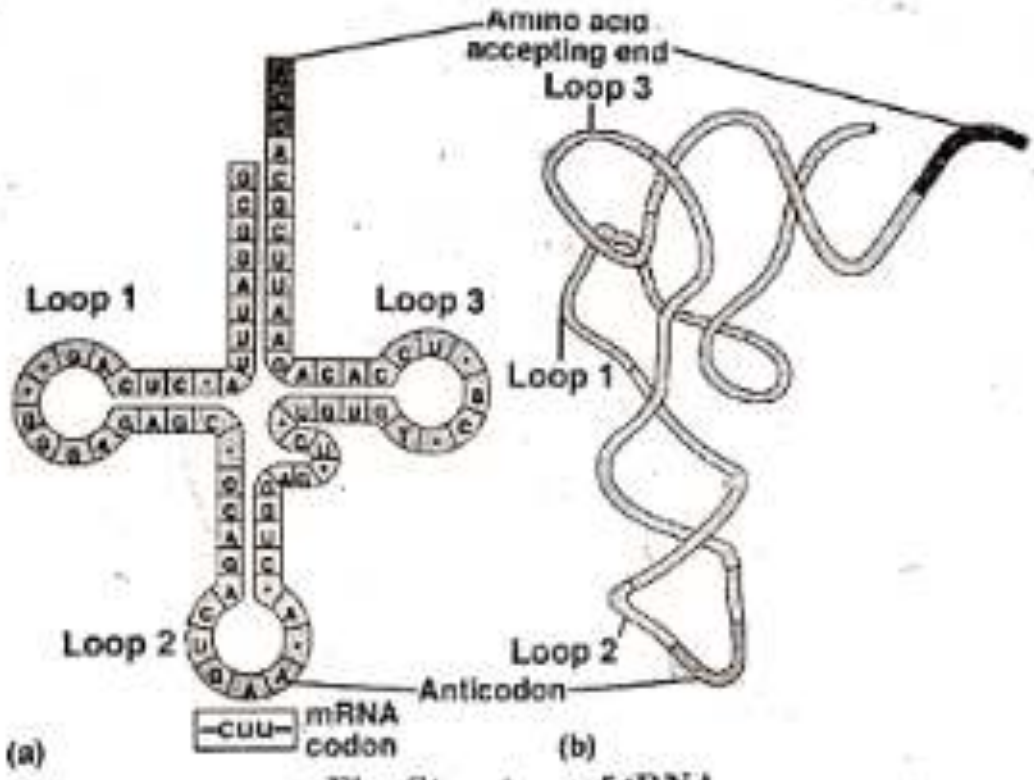


Fig: Structure of tRNA

(a) **Messenger RNA (mRNA)**

The mRNA brings genetic message from nucleus to the ribosome. The mRNA is about **3 to 4%** of the total RNA in the cell. The ribosomes are present in the cytoplasm for the synthesis of the particular protein. DNA transfers its genetic information to mRNA. Now, this mRNA has genetic information for the synthesis of specific protein. This mRNA, attaches with the ribosome for synthesis of protein. The mRNA consists of single strand of variable length. Its size depends on size of gene (on DNA) for the specific protein. For example, the mRNA has 3,000 nucleotides for a protein of 1,000 amino acids.

(b) **Transfer RNA (tRNA)**

The tRNA reads message (code) on mRNA and transfer specific amino acid to the ribosome. It has clover leaf like structure. It has anticodes. These amino acids are linked to form polypeptide chain of the protein. There is one specific tRNA for each amino acid. So the cell contains 20 types of tRNA. It forms about **10 to 20%** of the total cellular RNA. The tRNA has small size. Its chain is composed of 75 to 90 nucleotides.

(c) **Ribosomal RNA (rRNA)**

The rRNA combines with ribosomal proteins and forms ribosome. Ribosome is made up of rRNA and protein. The rRNA forms 50 % of the ribosome. It forms a large part, about **80 %**, of total RNA. It acts as machine for the synthesis of protein.

NUCLEIC ACID CHEMISTRY

DNA has role as a repository of genetic information. This role depends on its inherent stability. The chemical transformations are generally very slow in the absence of an enzyme catalyst. The long-term storage of information without alteration is so important

to a cell. However, processes like carcinogenesis and aging may be linked to slowly accumulating, irreversible alterations of DNA. Other, nondestructive alterations also occur. These are essential to function. These alterations include strand separation that must precede DNA replication or transcription.

Different powerful technologies help us to understand chemistry of DNA. These technologies have applications in molecular biology, medicine, and forensic science. We now examine the chemical properties of DNA and some of these technologies.

Denaturation: Double-Helical DNA and RNA Can Be Denatured

Solutions of carefully isolated native DNA are highly viscous at pH 7.0 and room temperature (25° C). Such a solution is subjected to extremes of pH or to temperatures above 80° C. As a result, its viscosity decreases sharply. It indicates that the DNA has undergone a physical change.

Denaturation

Heat and extremes of pH also cause denaturation, or melting, of double-helical DNA. Disruption of hydrogen bonds between paired bases and of base stacking causes unwinding of double helix. They form two single strands. They completely separate from each other along the entire length or part of the length (partial denaturation) of the molecule. No covalent bonds in the DNA are broken.

Renaturation

Renaturation of a DNA molecule is a rapid one-step process. A double-helical segment of a dozen or more residues still unites the two strands. When the temperature or pH is returned to the range in which most organisms live, the unwound segments of the two strands spontaneously rewind, or anneal. They yield the intact duplex. However, if the two strands are completely separated, renaturation occurs in two steps.

a) The first step is slow step. In this step, two strands find each other by random collisions. They form a short segment of complementary double helix.
b) The second step is much faster. The remaining unpaired bases successively come into register as base pairs. The two strands zipper themselves together to form the double helix.

Applications of Denaturation

1. There is close interaction between stacked bases in a nucleic acid. It has the effect of decreasing its absorption of UV light relative to that of a solution with the same concentration of free nucleotides. The absorption is decreased further when two complementary nucleic acid strands are paired. This is called **hypochromic effect**. Denaturation of a double-stranded nucleic acid produces the opposite result. There is an increase in absorption. It is called **hyperchromic effect**. Thus transition from double-stranded DNA to single-stranded, denatured form can be detected by monitoring UV absorption at 260 nm.
2. Viral or bacterial DNA molecules in solution denature when they are heated slowly. Each species of DNA has a characteristic denaturation temperature, or melting point. The higher its content of G≡C base pairs, the higher is the melting point of the DNA.

This is because G≡C base pairs have three hydrogen bonds. It requires more heat energy to dissociate than A=T base pairs. Thus the melting point of a DNA molecule, determined under fixed conditions of pH and ionic strength. It can yield an estimate of its base composition. If denaturation conditions are carefully controlled, regions that are rich in A=T base pairs will specifically denature. But most of the DNA remains double-stranded. Such denatured regions (called **bubbles**) can be visualized with **electron microscopy**.

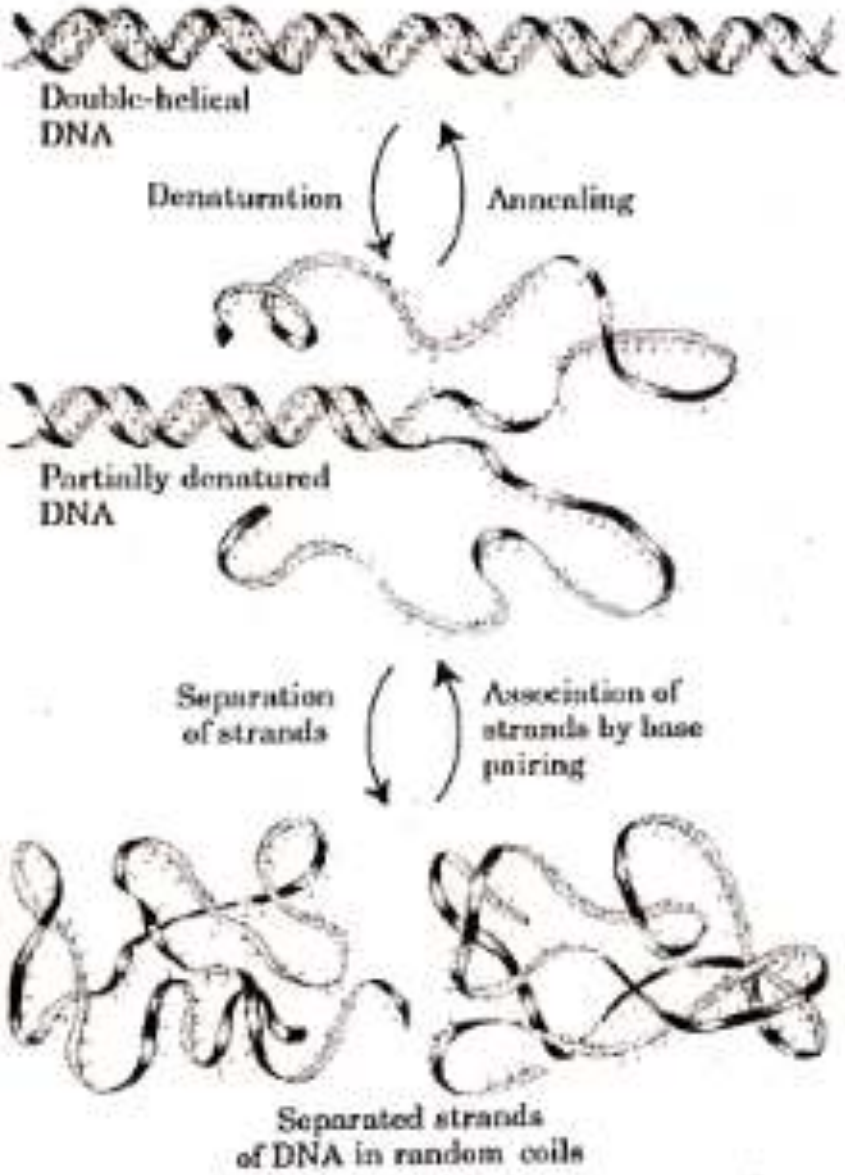


Fig: Reversible denaturation and annealing (renaturation) of DNA

Duplexes of two RNA strands or one RNA strand and one DNA strand (RNA-DNA hybrids) can also be denatured. RNA duplexes are more stable than DNA duplexes. At neutral pH, denaturation of a double- helical RNA often requires temperatures 20° C or higher than those required for denaturation of a DNA molecule with a comparable sequence. The stability of an RNA-DNA hybrid is generally intermediate between that of RNA and that of DNA.

Hybridization: Nucleic Acids from Different Species Can Form Hybrids

Two complementary DNA strands have ability of to pair with one another. It can be used

to detect similar DNA sequences in two different species or within the genome of a single species.

Technique

Suppose duplex DNAs isolated from human cells and from mouse cells. These are completely denatured by heating. Then these are mixed and kept at about 25° C below their tm for many hours. Much of the DNA will anneal. The rate of DNA annealing is affected by temperature, the length and concentration of DNA fragments, concentration of salts in the reaction mixture, and properties of sequence itself (e.g., complexity and G=C content).

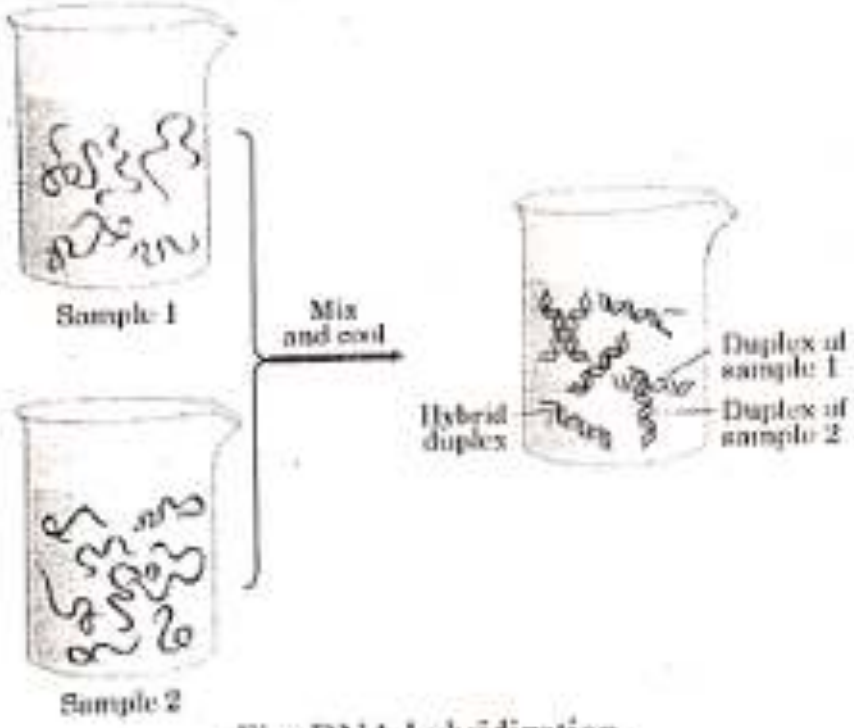


Fig: DNA hybridization

Factors affecting the Hybridization

Temperature is especially important for hybridization.

a) At low temperature, short sequences with coincidental similarity from distant, heterologous parts of the DNA molecules will anneal unproductively. It will interfere with the more general alignment of complementary DNA strands.

b) Temperatures that are too high will favor denaturation.

Process of Reannealing

a) Most of the reannealing occurs between complementary mouse DNA strands to form mouse duplex DNA.

b) Similarly, most human DNA strands anneal with complementary human DNA strands.

c) However, some strands of the mouse DNA will associate with human DNA strands to yield **hybrid duplexes.** In these hybrid duplexes, segments of a mouse DNA strand form base-paired regions with segments of a human DNA strand. This shows a common evolutionary heritage.

Applications

1. **Determination of evolutionary relationship:** Different organisms generally have many proteins and RNAs with similar functions and, often, similar structures. In many cases, the DNAs encoding these proteins and RNAs have similar sequences. The closer the evolutionary relationship between two species, more extensively their DNAs will hybridize. For example, human DNA hybridizes much more extensively with mouse DNA than with DNA from yeast.
2. **Development of different techniques:** The hybridization of DNA strands from different sources forms the basis for a powerful set of techniques. These techniques are essential to the practice of modern molecular genetics. It can detect specific DNA sequence or gene in the presence of many other sequences. The complementary DNA can be from a different species or from the same species.
3. **Identification of DNA or RNA:** Hybridization techniques can be varied to detect a specific RNA rather than DNA. The isolation and identification of specific genes and RNAs rely on these hybridization techniques.
4. **Modern application:** Applications of this technology make possible the identification of an individual on the basis of a single hair left at the scene of a crime or the prediction of the onset of a disease decades before symptoms appear.

Non-enzymatic Transformations: Nucleotides and Nucleic Acids Undergo Non-enzymatic Transformations

Purines and pyrimidines undergo spontaneous alterations in their covalent structure. The rate of these reactions is generally very slow. But they are physiologically significant. The cell has very low tolerance for alterations in its genetic information. Alterations in **DNA structure that produce permanent changes in the genetic information encoded are called mutations**. Much evidence suggests that there is an intimate link between the accumulation of mutations in an individual organism and processes of aging and carcinogenesis.

1. Deamination Reactions

Several nucleotide bases undergo spontaneous loss of their exocyclic amino groups (deamination). For example, under typical cellular conditions, deamination of cytosine (in DNA) to uracil occurs in about one of every 107 cytidine residues in 24 hours. This corresponds to about 100 spontaneous events per day on average in a mammalian cell. Deamination of adenine and guanine occurs at about 1/100th this rate. The slow cytosine deamination reaction is reason why DNA contains thymine rather than uracil. The product of cytosine deamination (uracil) is readily recognized as foreign in DNA. It is removed by a repair system. Cytosine deamination would gradually lead to a decrease in G≡C base pairs and an increase in AUU base pairs in the DNA of all cells. Over the millennia, cytosine deamination could eliminate G≡C base pairs and the genetic code that depends on them. Establishing thymine as one of the four bases in DNA may well have been one of the crucial turning points in evolution. It made the long-term storage of genetic information possible.

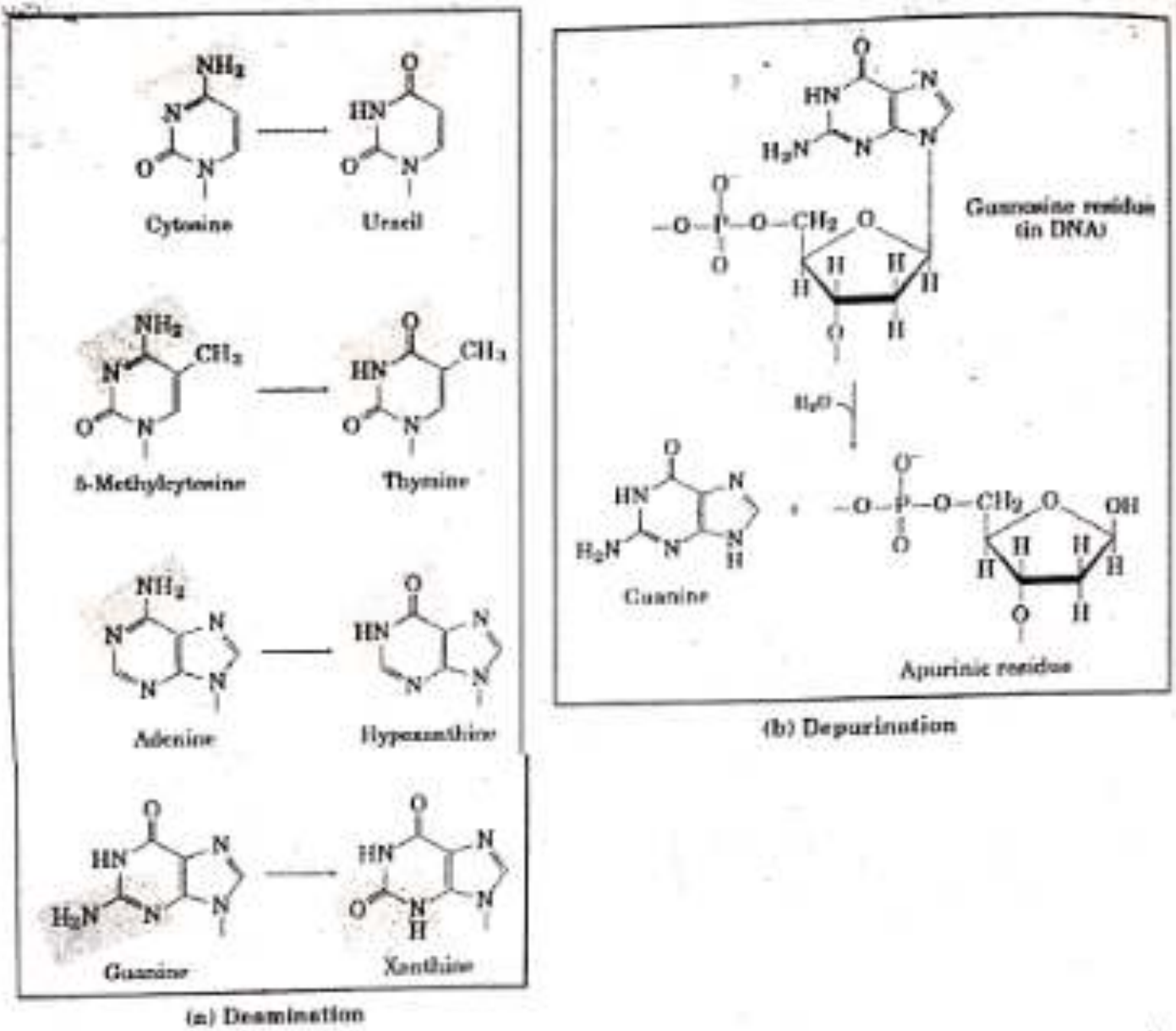


Fig: Nonenzymatic reactions of nucleotides. (a) Deamination reactions. (b) Depurination

2. Depurination: Hydrolysis of N-β-glycosyl bond between base and pentose

Another important reaction in deoxyribonucleotides is the hydrolysis of N-β-glycosyl bond between base and pentose. It creates a DNA lesion called an **AP (apurinic, apyrimidinic) site or abasic site.** This occurs at a higher rate for purines than for pyrimidines. As many as one in 105 purines (10,000 per mammalian cell) are lost from DNA every 24 hours under typical cellular conditions. Depurination of ribonucleotides and RNA is much slower. Thus generally it is not considered physiologically significant.

3. Formation of pyrimidine dimers

Such reactions are promoted by radiation. UV light induces the condensation of two ethylene groups to form a **cyclobutane ring**. In the cell, the same reaction between adjacent pyrimidine bases in nucleic acids forms **cyclobutane pyrimidine dimers**. This happens most frequently between adjacent thymidine residues on the same DNA strand.

A second type of pyrimidine dimer is called a 6-4 photoproduct. It is also formed during UV irradiation. Ionizing radiation (x rays and gamma rays) can cause ring opening and fragmentation of bases as well as breaks in the covalent backbone of nucleic acids.

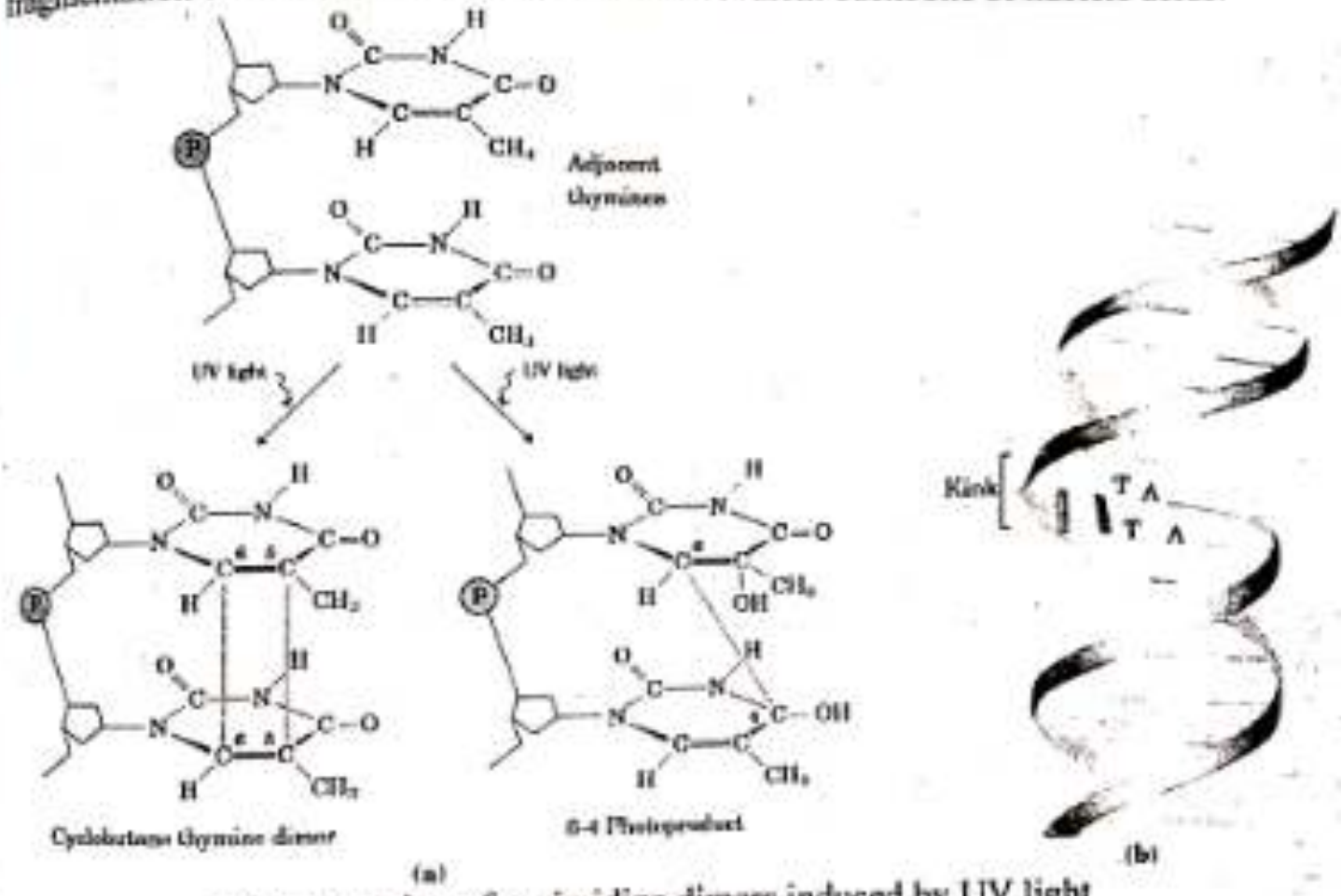


Fig: Formation of pyrimidine dimers induced by UV light

4. **Chemical agents that cause DNA damage**

DNA also may be damaged by reactive chemicals. These chemicals are introduced into the environment as products of industrial activity. Such products may not be injurious per se, but they may be metabolized by cells into active forms. There are two prominent classes of such agents:

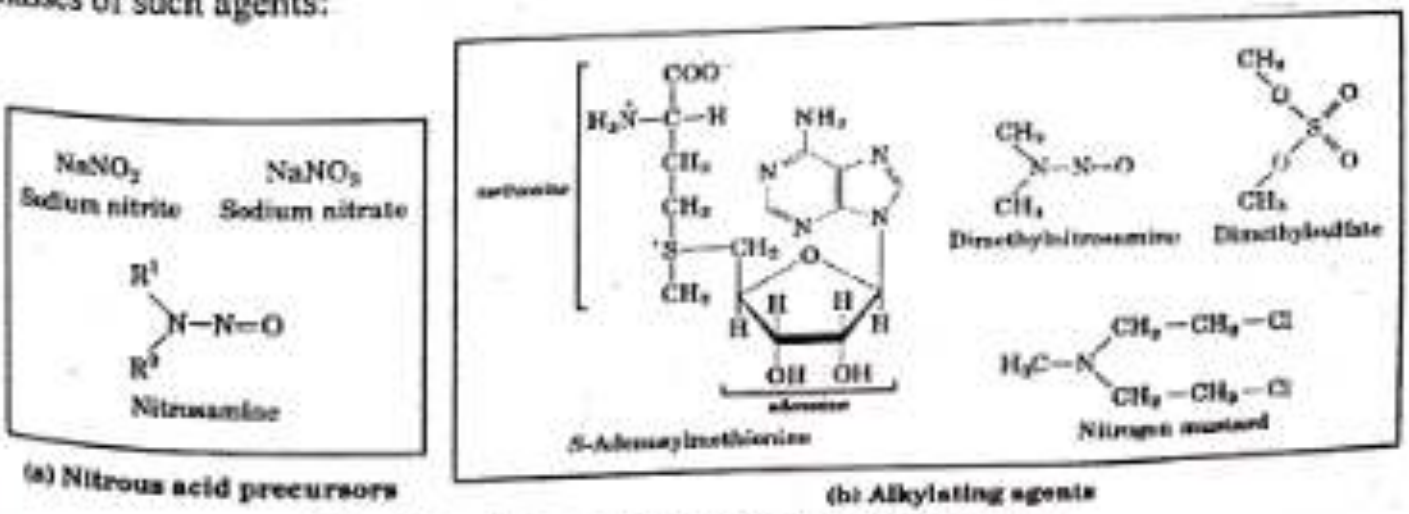


Fig: Chemical agents that cause DNA damage

a) **Deaminating agents:** These are like nitrous acid (HNO_2) or compounds that can be metabolized to nitrous acid or nitrites. Nitrous acids are formed from organic precursors like nitrosamines and from nitrite and nitrate salts. Bisulfite has similar effects. Both agents are used as preservatives in processed foods.

b) **Alkylating agents:** Alkylating agents can alter certain bases of DNA. For example, highly reactive chemical dimethylsulfate can methylate a guanine. It yields O^6-methylguanine. It cannot base-pair with cytosine.

5. Oxidative damage

Oxidative damage is the most important source of mutagenic alterations in DNA. There are some excited-oxygen species like hydrogen peroxide, hydroxyl radicals, and superoxide radicals. These arise during irradiation or as a byproduct of aerobic metabolism. Of these species, hydroxyl radicals are responsible for most oxidative DNA damage. Cells have an elaborate defense system to destroy reactive oxygen species. These include enzymes such as catalase and superoxide dismutase. These enzymes convert reactive oxygen species to harmless products. A fraction of these oxidants escape cellular defenses, however. They damage to DNA by oxidation of deoxyribose and base moieties to strand breaks. Every day the DNA of each human cell is subjected to thousands of damaging oxidative reactions.

Methylation of Some Bases of DNA

Certain nucleotide bases in DNA molecules are enzymatically methylated. Adenine and cytosine are methylated more often than guanine and thymine. Methylation is confined to certain sequences or regions of a DNA molecule. In some cases the function of methylation is well understood. In others the function remains unclear. All known DNA methylases use S-adenosylmethionine as a methyl group donor.

E. coli has two prominent methylation systems.

a) One serves as part of a defense mechanism. It helps the cell to distinguish its DNA from foreign DNA. It marks its own DNA with methyl groups and destroys (foreign) DNA without methyl groups. It is called **restriction-modification system.**

b) The other system methylates adenosine residues within sequence (5') GATC (3') to N^6-methyladenosine. This is mediated by the Dam (DNAadenine methylation) methylase.

In eukaryotic cells, about 5% of cytidine residues in DNA are methylated to 5-methylcytidine. Methylation is most common at C=G sequences. It produces methyl-C=G symmetrically on both strands of the DNA. The extent of methylation of C=G sequences varies by molecular region in large eukaryotic DNA molecules.

Determination of Sequences of Long DNA Strands

A DNA molecule's most important property is its nucleotide sequence. Two new techniques were developed in 1977.

a) One was developed by Alan Maxam and Walter Gilbert.

b) The other was developed by Frederick Sanger.

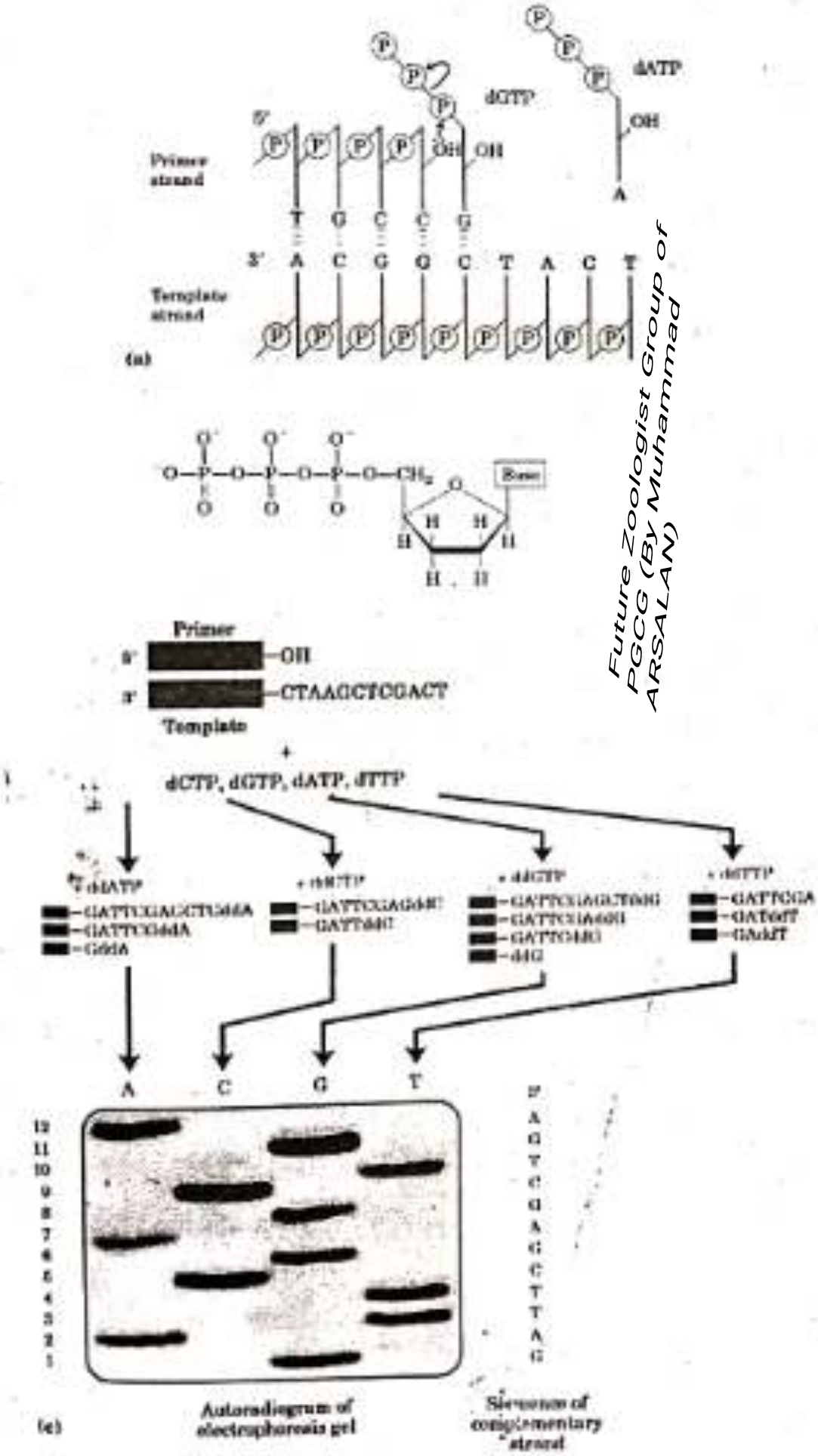


Fig: DNA sequencing by the Sanger method

Future Zoologist Group of PGCG (By Muhammad ARSALAN)

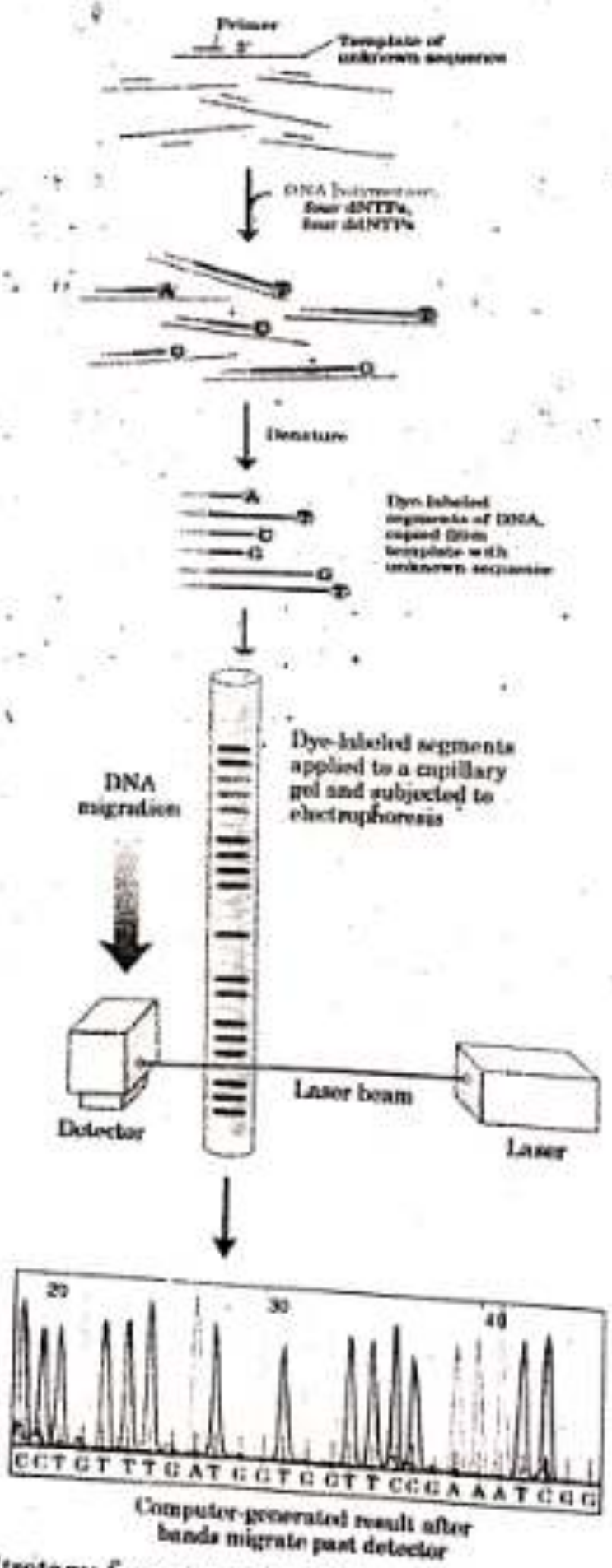


Fig: Strategy for automating DNA sequencing reactions

These techniques made the DNA sequencing easier. The techniques depend on an improved understanding of nucleotide chemistry and DNA metabolism. These techniques also depend on electrophoretic methods for separating DNA strands differing in size by only one nucleotide.

Electrophoresis

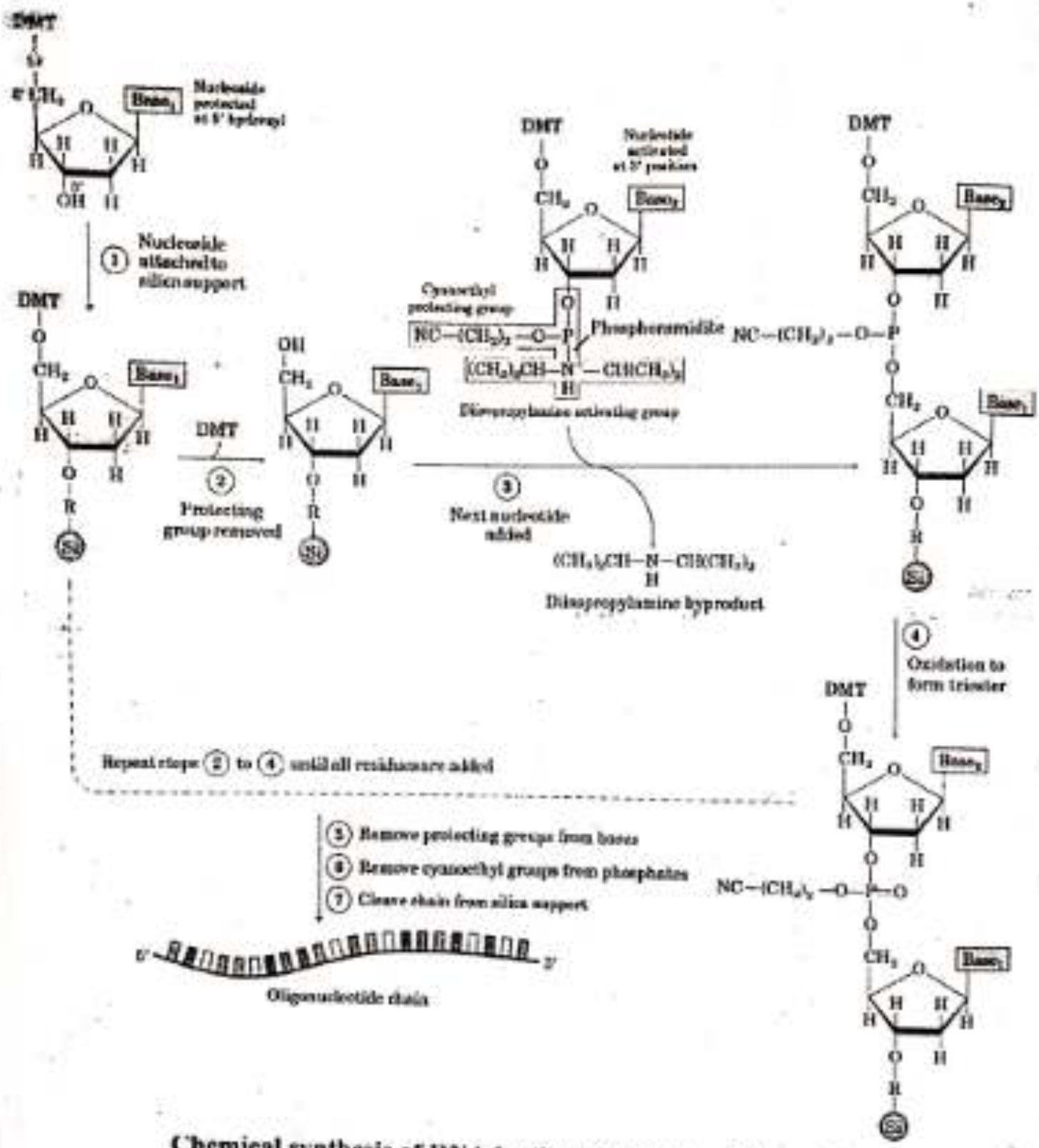
Electrophoresis of DNA is similar to that of proteins. Polyacrylamide is used as a gel matrix in work with short DNA molecules. Agarose is used for longer pieces of DNA.

Method of sequencing

1. In both Sanger and Maxam-Gilbert sequencing, general principle is to reduce the DNA to four sets of labeled fragments. The reaction producing each set is base-specific. Hence the lengths of the fragments correspond to positions in the DNA sequence where a certain base occurs. For example, for an oligonucleotide with the sequence pAATCGACT, labeled at the 5' end (the left end). A reaction that breaks the DNA after each C residue will generate two labeled fragments. It forms a four-nucleotide and a seven-nucleotide fragment. The reaction that breaks the DNA after each G will produce only one labeled, five-nucleotide fragment.
2. The fragments are radioactively labeled at their 5'ends. Therefore, only the fragment to the 5' side of the break is visualized. The fragment sizes correspond to relative positions of C and G residues in the sequence. The sets of fragments corresponding to each of four bases are electrophoretically separated side by side. They then produce a ladder of bands from which sequence can be read directly.
3. The Sanger method is technically easier. Thus it is more commonly used. It requires enzymatic synthesis of a DNA strand. This strand is complementary to the strand under analysis. This is done by using a radioactively labeled **primer** and **dideoxynucleotides.**
4. DNA sequencing is,now automated by a variation of Sanger's sequencing method. Now dideoxynucleotides are used for each reaction are labeled with a differently colored fluorescent tag. With this technology, researchers can sequence DNA molecules containing thousands of nucleotides in a few hours.

Automated Chemical Synthesis of DNA

The rapid and accurate synthesis of short oligonucleotides of known sequence is important in nucleic acid chemistry. The methods were pioneered by H. Gobind Khorana and his colleagues in the 1970s. It was refined by Robert Letsinger and Marvin Caruthers. They formed the method called **phosphoramidite method**. The synthesis is carried out with the growing strand attached to a solid support. This synthesis is readily automated. The efficiency of each addition step is very high. It allows the routine synthesis of polymers containing 70 or 80 nucleotides. Inexpensive DNA polymers are available. They make it easier.



Chemical synthesis of DNA by the phosphoramidite method

Procedure of phosphoramidite method

Automated DNA synthesis is similar to the synthesis of polypeptides on a solid support. The oligonucleotide is built up on the solid support (silica), one nucleotide at a time, in a repeated series of chemical reactions.

1. The first nucleoside (which will be the 3' end) is attached to silica support at the 3' hydroxyl. It is protected at the 5' hydroxyl with an acid-labile dimethoxytrityl group (DMT). The reactive groups on all bases are also chemically protected.

2. The protecting DMT group is removed by washing the column with acid.
3. The next nucleotide has a reactive phosphoramidite at its 3' position. It is a trivalent phosphite. It has one linked oxygen replaced by an amino group or substituted amine. One of phosphoramidite oxygen is bonded to the deoxyribose. The other is protected by a cyanoethyl group. The third position is occupied by a readily displaced diisopropyl- amino group. Reaction with immobilized nucleotide forms a 5', 3' linkage. The diisopropylamino group is eliminated.
4. In this step, phosphite linkage is oxidized with iodine. It produces a phosphotriester linkage. Reactions 2 through 4 are repeated until all nucleotides are added. At each step, excess nucleotide is removed before addition of the next nucleotide.
5. In steps 5 and 6, remaining protecting groups on the bases and the phosphates are removed.
6. In step 7, oligonucleotide is separated from the solid support and purified.

EXERCISE

Long Questions

1. **What is nucleotide? Describe its composition.**
2. **What are major and minor nitrogenous bases? Write their formulae.**
3. **What are nucleosides? How do they differ from nucleotides?**
4. **How is phosphodiester bond formed between two nucleotides?**
5. **How do properties of bases affect the three dimensional structure of nucleic acid?**
6. **Describe structure of DNA.**
7. **Discuss the significance of wok different biochemists in the formulation of DNA structure.**
8. **Describe in detail Watson and Crick model of DNA.**
9. **Illustrate different structural variations in DNA.**
10. **What are A-form, D-form and Z-form of DNAs?**
11. **Illustrate the different aspects of RNA structure.**
12. **What are three types of RNA?**
13. **How can DNA be denatured and renatured? Write its applications.**
14. **How can nucleic acid of two different species form hybrids?.**
15. **What is hybridization of DNA? Write its application.**
16. **Illustrate the method of determination of sequence of DNA.**
17. **Describe phosphoramidite method of synthesis of DNA.**

Short Questions

1. Why the nucleic acids were named so?

Ans: They were named nucleic acid, due to their isolation from nucleus and their acidic nature.

2. Differentiate between nucleotides and nucleosides?

Ans:

Nucleotide	Nucleoside
1. The compounds containing nitrogenous basis, pentose sugar and phosphate group are called nucleotides.	1. The compound formed by the nitrogenous base and pentose sugar is called nucleoside.
2. Nitrogen base + pentose sugar + Phosphate	2. Nitrogen base + pentose sugar
3. A nucleoside and phosphoric acid combine to form a nucleotide.	3. A nucleotide without its phosphate group is nucleoside.

3. Differentiate between RNA and DNA.

Ans:

DNA	RNA
1. It is double stranded	1. It is single stranded
2. Its pentose sugar is deoxyribose	2. Its pentose sugar is ribose
3. The nitrogen bases are Thymine, Adenine guanine and cytosine	3. The nitrogen bases are Uracil , Adenine guanine and cytosine
4. It is present in nucleus, chloroplast, mitochondria	4. It is present in nucleus and cytoplasm.
5. DNA is a hereditary material	5. RNA is not hereditary material. It is involved in the synthesis of protein.
6. The synthesis of DNA is called replication.	6. The synthesis of DNA is called transcription.

4. How hydrogen bonding occur between two strands of DNA.

Ans: The two strands are attached with each other by weak hydrogen bonds. Two hydrogen bonds are formed between A and T and three hydrogen bonds are formed between G and C pairs. Thus adenine (A) is always opposite to Thymine (T) and Guanine (G) is always opposite to Cytosine.

5. What is Watson & Crick model of DNA?

Ans: James D. Watson and **Francis Crick** proposed the model of DNA. It suggests:

1. The DNA is made up of two polynucleotide chains or strands.
2. These two strands coil around each other and form a double helix.
3. The coiling of the two strands is opposite (Antiparallel to each other i.e., one strand is 3-5 the other is 5-3).
4. The two strands are attached with each other by weak hydrogen bonds. Two hydrogen bonds are formed between A and T and three hydrogen bonds are formed between G

and C pairs.

5. Adenine (A) is always opposite to Thymine (T) and Guanine (G) is always opposite to Cytosine.
6. The two strands coil around each other. Each turn of the strands is about 34 Angstrom. This turn contains 10 base pairs.

4. What are gene and genome?

Ans: A gene is a small part of DNA that controls a specific character. The total genes present in the cell of an organism are called genome.

6. What is RNA? Name its three types.

Ans: RNA: RNA is polymers of ribonucleotides. RNA molecule has single strand. Sometimes, this strand may fold back to give double helical characteristics. Nitrogenous bases form complementary pairings. RNA have nitrogenous base Uracil in place of Thymine. Cytosine (C) forms pair with Guanine (G) and Uracil (U) forms pair with Adenine (A). DNA synthesizes RNA. The process of synthesis of RNA from DNA is called transcription.

Types of RNA: There are three main types of RNA. These are messenger RNA (mRNA), transfer RNA (tRNA) and ribosomal RNA (rRNA).

7. What is tRNA? Give its role. (Gujranwala Board 2012)

Ans: Transfer RNA (tRNA): The transfer RNA picks up amino acids and transfers them to ribosome. Amino acids are linked in ribosome to form proteins. It is about 10 to 20% of the total cellular RNA. The tRNA has small size. Its chain is composed of 75 to 90 nucleotides.

Role: The tRNA transfers amino acid to the site of polypeptide chain synthesis (ribosome). It transfers amino acids. There is one specific tRNA for each amino acid. So the cells have 20 types of tRNA.

8. What is the role of different types of RNAs in proteins synthesis?

Ans: rRNA forms ribosome. The mRNA brings message from the DNA and rRNA read this message on mRNA and brings specific amino acid for the synthesis of protein.

9. What are hydrophobic stacking interactions? What is their role in stability of nucleotide chains?

Ans: Hydrophobic stacking interactions occur in nucleic acid. In these interactions, two or more bases are positioned with the planes of their rings parallel (like a stack of coins). These are one of two important modes of interaction between bases in nucleic acids. The stacking also involves a combination of van der Waals and dipole-dipole interactions between the bases. Base stacking helps to minimize contact of the bases with water. The base-stacking interactions are very important in stabilizing the three dimensional structure of nucleic acids.

10. What are primary, secondary and tertiary structures of DNA?

Ans:

The DNA structure is described at different levels. These are primary, secondary and tertiary structures.

a) **Primary structure:** The primary structure of a nucleic acid is its covalent structure

and nucleotide sequence.

b) **Secondary structure:** Any regular, stable structure taken up by nucleotides in a nucleic acid is called secondary structure. All structures other than primary structures come under the heading of secondary structure.

c) **Tertiary structure:** The complex folding of large chromosomes within eukaryotic chromatin and bacterial nucleoids is tertiary structure.

11. What was the work of Rosalind Franklin and Maurice Wilkins? What is its significance?

Ans:

Rosalind Franklin and Maurice Wilkins used powerful method of x-ray diffraction to analyze DNA fibers. They showed in the early 1950s that DNA produces a characteristic x-ray **diffraction pattern**. From this pattern it was deduced that DNA molecules are helical. It has two periodicities along their long axis. These include a primary one of 3.4 Å and a secondary one of 34 Å. The problem then was to formulate a three-dimensional model of the DNA molecule. This model should describe x-ray diffraction data, specific A =T and G = C base equivalences and other chemical properties of DNA.

12. What is the significance of Watson and Crick Model of DNA?

Ans:

Significance of Model

This model was a revolution in our understanding of biological inheritance.

a) The important features of the double-helical model of DNA structure are supported by much chemical and biological evidence.

b) Moreover, model immediately suggested a mechanism for the transmission of genetic information.

c) The essential feature of the model is complementarity of the two DNA strands. It explains the separating the two strands and synthesizing a complementary strand for each.

d) The nucleotides in each new strand are joined in a sequence specified by the base-pairing rules stated above. Therefore, each preexisting strand functions as a template to guide the synthesis of one complementary strand. These expectations were experimentally confirmed.

13. What is meant by B-form DNA?

Ans:

The Watson-Crick structure is also called **B-form DNA**, or **B-DNA**. The B form is the most stable structure for a random-sequence DNA molecule under physiological conditions. Therefore, it is the standard point of reference in any study of the properties of DNA. There are two structural variants of DNA, **A and Z forms**. These are characterized in crystal structures. These three DNA conformations are shown in Figure with a summary of their properties.

14. What is Z-form DNA?

Ans: Z-form DNA is a more radical departure from B structure. Here most obvious distinction is the left- handed helical rotation. There are 12 base pairs per helical turn.

Thus structure appears more slender and elongated. The DNA backbone takes on a zigzag appearance. Certain nucleotide sequences fold into left-handed Z helices much more readily than others. Prominent examples are sequences in which pyrimidines alternate with purines. It especially alternates C and G or 5-methyl-C and G residues. To form the left- handed helix in Z-DNA, purine residues flip to **syn conformation**. It alternates with pyrimidines in anti conformation. The major groove is barely apparent in Z-DNA. The minor groove is narrow and deep.

15. How are hairpin loops formed in DNA?

Ans: Breaks occur in the regular A-form helix. These are caused by mismatched or unmatched bases in one or both strands. Such breaks are common and result in bulges or internal loops. **Hairpin loops** form between nearby self-complementary sequences. The potential for base-paired helical structures in many RNAs is extensive. It results in hairpins. These are the most common type of secondary structure in RNA.

16. What is meant by denaturation of DNA?

Ans: Denaturation

Heat and extremes of pH also cause denaturation, or melting, of double-helical DNA. Disruption of hydrogen bonds between paired bases and of base stacking causes unwinding of double helix. They form two single strands. They completely separate from each other along the entire length or part of the length (partial denaturation) of the molecule. No covalent bonds in the DNA are broken.

17. What are two chemical agents responsible for breaking of DNA?

Ans:

a) **Deaminating agents:** These are like nitrous acid (HNO_2) or compounds that can be metabolized to nitrous acid or nitrites. Nitrous acids are formed from organic precursors like nitrosamines and from nitrite and nitrate salts. Bisulfite has similar effects. Both agents are used as preservatives in processed foods.

b) **Alkylating agents**: Alkylating agents can alter certain bases of DNA. For example, highly reactive chemical dimethylsulfate can methylate a guanine. It yields O^6-methylguanine. It cannot base-pair with cytosine.

18. What is DMT?

Ans:

The first nucleoside (which will be the 3' end) is attached to silica support at the 3' hydroxyl. It is protected at the 5' hydroxyl with an acid-labile dimethoxytrityl group (DMT).

Multiple Choice Questions (MCQs)

1. Nucleic acids were discovered by:

(a) Sanger (b) Miescher (c) Watson (d) None

2. Which of the following bases is absent in DNA?

(a) Purine (b) Uracil (c) Guanine (d) Thymine

3. Each turn of DNA strand is:

(a) 24 Angstrom (b) 34 Angstrom (c) 44 Angstrom (d) None

4. The amount of mRNA in the cells is:
(a) 3 to 4% (b) 10 to 20% (c) 50% (d) 80%

5. The amount of tRNA in the cells is:
(a) 3 to 4% (b) 10 to 20% (c) 50% (d) 80%

6. The amount of rRNA in the cells is:
(a) 3 to 4% (b) 10 to 20% (c) 50% (d) 80%

7. Which of the followings is not a nucleotide?
(a) ATP (b) GTP (c) NADH (d) DNA

8. Which one of the followings is not true sequence of nucleotides in nucleic acids?
(a) Adenine and Thymine (b) Guanine and Cytosine
(c) Adenine and Cytosine (d) Cytosine and Guanine

9. Which RNA is the direct copy of DNA?
(a) mRNA (b) tRNA
(c) rRNA (d) None of the above

10. Hydrogen bonds in nucleic acid involve the amino and:
(a) Alcoholic group (b) carbonyl groups
(c) carboxylic acid group (d) oxygen group

11. Primary structure of DNA involves:
(a) hydrogen boding (b) covalent bonding
(c) ionic bonding (d) all of these

12. Secondary structure of DNA involves:
(a) hydrogen boding (b) covalent bonding
(c) ionic bonding (d) all of these

13. Watson-Crick structure is also called:
(a) A-form DNA (b) B-form DNA
(c) Z-form DNA (d) all of these

14. The _____ is favored in many solutions devoid of water.
(a) A-form DNA (b) B-form DNA
(c) Z-form DNA (d) all of these

15. The structure of DNA with the left- handed helical rotation is:
(a) A-form DNA (b) B-form DNA
(c) Z-form DNA (d) all of these

16. The sequences often found at the ends of RNA hairpins are:
(a) UUCG (b) AUUUC
(c) UUGC (d) UACU

17. Heat and extremes of pH also cause in double-helical DNA:
(a) denaturation (b) joining
(c) bonding (d) hybridization

18. Method of sequencing was discovered by:
(a) Sanger (b) Maxam-Gilbert
(c) Smith (d) a and b

19. Method of automated DNA synthesis was developed by first time:

(a) Sanger (b) Maxam-Gilbert
(c) Smith (d) Khorana

20. Phosphoramidite method is used for the synthesis of:

(a) RNA (b) DNA
(c) Protein (d) Cholesterol

Ans:

1. (b) 2. (b) 3. (b) 4. (a) 5. (b)b 6. (d) 7. (d) 8. (c) 9. (a) 10. (b) 11. (b) 12. (a) 13. (b) 14. (a) 15. (c) 16. (a) 17. (a)18. (d) 19. (d) 20. (b)

Future Zoologist Group of PGCG (By Muhammad ARSALAN)