Biomolecules And Biological Processes

Content

14. Biomolecules And Biological Processes

14.1 Introduction

The branch of science that deals with the study of the chemical composition and structure of living organisms and also various changes taking place within them is known as **Biochemistry,** while the complex organic molecules which form the basis of life, i.e., which build up living organisms and are also required for their growth and maintenance are called **Biomolecules** . For example, carbohydrates, proteins, lipids, vitamins, etc..

These biomolecules interact with each other and constitute the living systems. Thus, Biomolecules may be defined as complex lifeless chemical substances which form the basis of life, i.e., they not only build up living systems (creatures) but are also responsible for their growth, maintenance and their ability to reproduce.

The various biomolecules are carbohydrates, proteins, enzymes, nucleic acids, lipids, hormones and compounds for storage and exchange of energy such as adenosine triphosphate (ATP). In addition, some simple molecules like vitamins and mineral salts also play an important role in the functions of organisms.

14.2 Carbohydrates

One of the most important classes of biomolecules is carbohydrates. These include compounds like sugars, starches, glycogen, cellulose, dextrins and gums. Although they are widely distributed both in animal and plant kingdom yet they are obtained mainly from plants. They are formed in plants by a process known as photosynthesis and make up about 70% of the solid plant material.

$$xCO_2 + yH_2O \xrightarrow[\text{Chlorophyll}]{\text{Sunlight}} \underset{\text{Carbohydrate}}{C_x(H_2O)_y} + x\,O_2$$

Carbohydrates play a vital role in our daily life. They provide us with three basic necessities of life, i.e., food (in the form of starch), clothing (cotton, linen and rayon are essentially cellulose) and shelter (wood in making houses and furniture is almost cellulose).

The name carbohydrate (meaning hydrate of carbon) was originally given to a class of compounds containing only carbon, hydrogen and oxygen; the hydrogen and oxygen being present in the same ratio as in water. These compounds would therefore be represented by the general formula, $C_2H_2O_Y$

Modern definition – Carbohydrates are defined as optically active polyhydroxy aldehydes or polyhydroxy ketones or substances which give these on hydrolysis.

It may be noted here that aldehydic and ketonic groups in carbohydrates are not present as such but usually exist in combination with one of the hydroxyl groups of the molecule in the form of hemiacetals and hemiketals respectively.

14.2.1 Classification of Carbohydrates

Carbohydrates are also known as saccharides. These are classified into the following three classes depending upon their behavior towards hydrolysis.

1. **Monosaccharides** These are the simplest carbohydrates which cannot be hydrolysed to smaller molecules. Their general formula is $(CH_2O)_n$ where n = 3 – 7.
2. **Oligosaccharides** (Greek, oligo means a few). These are carbohydrates which on hydrolysis give 2-10 molecules of monosaccharides. Depending upon the number of monosaccharide molecules actually obtained upon hydrolysis, they are further classified as di, tri, tetrasaccharides, etc. For example.

(i) Disaccharides Carbohydrates which upon hydrolysis give two molecules of the same or different monosaccharides are called disaccharides. For example, sucrose, maltose, lactose, etc. Their general formula is $C_{12}H_{22}O_{11}$.

(ii) Trisaccharides Carbohydrates which on hydrolysis give three molecules of the same or different monosaccharides are called trisaccharides. For example, raffinose upon hydrolysis

gives one molecule each of glucose, fructose and galactose. Their general formula is $C_{18}H_{32}O_{16}$.

(iii) Tetrasaccharides Carbohydrates which upon hydrolysis give four molecules of the same or different monosaccharides are called tetrasaccharides. For example, stachyrose upon hydrolysis gives one molecule each of glucose and fructose and two molecules of galactose. Their general formula is $C_{24}H_{42}O_{21}$.

3. Polysaccharides Carbohydrates which upon hydrolysis give a large number of monosaccharide molecules are called polysaccharides. The most commonly occurring polysaccharides are starch, cellulose and glycogen. Their general formula is $(C_6H_{10}O_5)_n$ where n = 100 – 3000.

Sugars and non-sugars Carbohydrates are further classified as sugars and non-sugars. All the monosaccharides and oligosaccharides are crystalline solids, soluble in water and sweet in taste. These are collectively called sugars. The polysaccharides, on the other hand, are amorphous solids, insoluble in water and tasteless and thus are called non-sugars.

Reducing and non-reducing sugars or carbohydrates. Carbohydrates are also classified as reducing and non-reducing sugars. All those carbohydrates which contain aldehydic or ketonic group in the hemiacetal or hemiketal form and reduce Tollens reagent or Fehling's solution are called reducing carbohydrates or sugars while others which do not reduce these reagents are called non-reducing carbohydrates or sugars. All monosaccharides whether aldoses or ketoses are reducing sugars. Disaccharides such as sucrose in which the two monosaccharide units are linked through their reducing centres, i.e., aldehydic or ketonic groups are non-reducing while others in which these groups are in the hemiacetal or hemiketal form are reducing sugars. However, all polysaccharides (starch, cellulose, glycogen, dextrins, etc.) are non-reducing carbohydrates.

Let us now briefly discuss each one of these classes of carbohydrates.

14.2.2 Monosaccharides – Structures of pentoses and hexoses.

Monosaccharides are the simplest carbohydrates which cannot be hydrolysed to smaller molecules. They contain 3 to 7 carbon atoms. There are about twenty monosaccharides which occur in nature. These are of two types :

(i) Aldoses : Monosaccharides containing an aldehyde (- CHO) group are called aldoses. Since the aldehyde group is monovalent, therefore, it is always present at one end of the carbon chain, i.e., at C_1.

(ii) Ketoses: Monosaccharides containing a keto (> C = O) group are called ketoses. Since the keto group is divalent, it can be present anywhere along the carbon chain. However, in all the naturally occurring ketoses, keto group is always present at a carbon atom next to the terminal carbon, i.e., at C_2.

Aldoses and ketoses are further classified as trioses, tetroses, pentoses, hexoses, heptoses, etc. according as they contain three, four, five, six, seven, etc. carbon atoms respectively. This classification of monosaccharides is summarized in Table.

TABLE : Classification of monosaccharides

Class	Molecular formula	Structural formula	Examples
		Aldoses	
Aldotrioses	$C_3H_6O_3$	$CH_2OH.CHOH.CHO$	Glyceraldehyde
Aldotetroses	$C_4H_8O_4$	$CH_2OH.(CHOH)_2.CHO$	Erythrose, threose
Aldopentoses	$C_5H_{10}O_5$	$CH_2OH.(CHOH)_3.CHO$	Arabinose, ribose, xylose, lyxose
Aldohexoses	$C_6H_{12}O_6$	$CH_2OH.(CHOH)_4.CHO$	Glucose, mannose, galactose, gulose, talose, idose, allose, altrose
Aldoheptoses	$C_7H_{14}O_7$	$CH_2OH.(CHOH)_5.CHO$	-
		Ketoses	
Ketotrioses	$C_3H_6O_3$	$CH_2OH.CO.CH_2OH$	Dihydroxyacetone
Ketotetroses	$C_4H_8O_4$	$CH_2OH.CO.CHOH,CH_2OH$	Erythrulose
Ketopentoses	$C_5H_{10}O_5$	$CH_2OH.CO.(CHOH)_2.CH_2OH$	Ribulose, xylulose
Ketohexoses	$C_6H_{12}O_6$	$CH_2OH.CO.(CHOH)_3.CH_2OH$	Fructose, sorbose, tagatose, etc.
Ketoheptoses	$C_7H_{14}O_7$	$CH_2OH.CO.(CHOH)_4.CH_2OH$	-

14.2.3 Glucose (Dextrose, Grape sugar) $C_6H_{12}O_6$

Glucose occurs in nature in free as well as in the combined state. It occurs in large quantities (about 20%) in ripe grapes and that is why it is called grape sugar. It also occurs in honey and most of the other sweet fruits (such as mangoes, apples, peaches, pears, etc.) In the combined state, it occurs in glycosides, dissacharides and polysaccharides.

Preparation of glucose

(i) From sucrose (Cane-sugar) When sucrose is hydrolysed by boiling with dil. HCl or H_2SO_4 in alcoholic solution, an equimolar mixture of glucose and fructose is obtained.

$$\underset{\text{Sucrose}}{C_{12}H_{22}O_{11}} + H_2O \xrightarrow{H^+} \underset{\text{Glucose}}{C_6H_{12}O_6} + \underset{\text{Fructose}}{C_6H_{12}O_6}$$

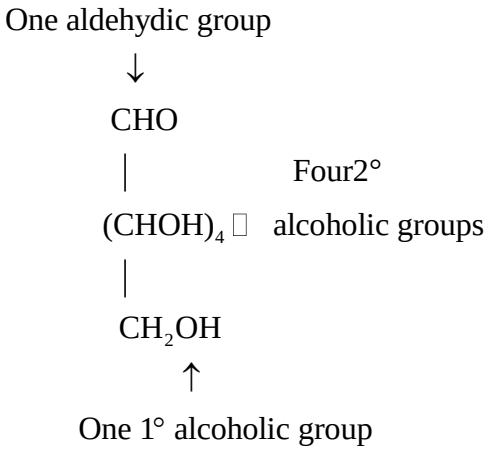
(ii) From starch Commercially glucose is obtained by hydrolysis of starch by boiling it with dil. H_2SO_4 at 393 K under pressure.

$$\underset{\text{Starch}}{(C_6H_{10}O_5)_n} + n\,H_2O \xrightarrow[393\text{ K, }2\text{ - }3\text{ atm}]{H^+} n\,\underset{\text{Glucose}}{C_6H_{12}O_6}$$

14.2.4 Open Chain Structure of Glucose

Glucose is an aldohexose. It is also called dextrose. It is the monomer of many polysaccharides such as starch, glycogen, cellulose, etc. It is probably the most abundant compound on the earth.

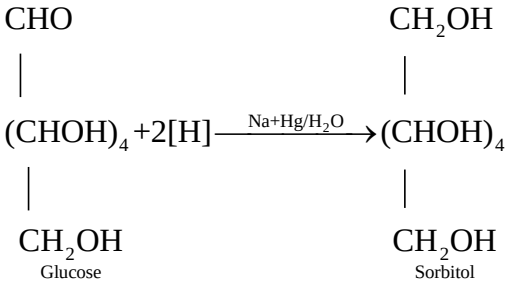
The open chain structure of glucose was proposed by Baeyer. It contains one aldehyde (- CHO) group, one primary alcoholic (- CH_2OH) group and four secondary alcoholic (- CHOH) groups.



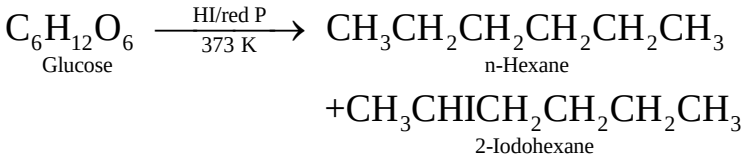
This structure was assigned on the basis of following evidence.

1. **Molecular formula –** The molecular formula of glucose is $C_6H_{12}O_6$.
2. **Straight chain**

   (i) On reduction with sodium amalgam, glucose yields sorbitol, a hexahydric alcohol.

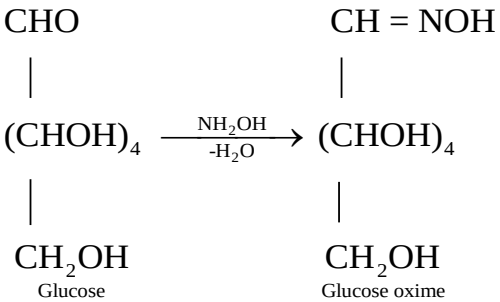


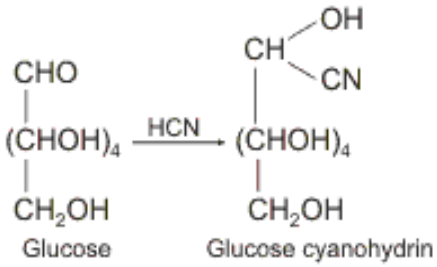
   (ii) Glucose on reduction with HI and red P at 373 K gives a mixture of n-hexane and 2-iodohexane.



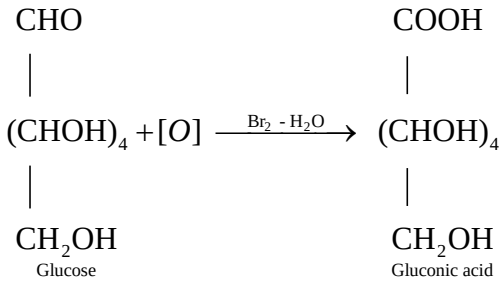
   This reaction suggests that all the six carbon atoms in glucose are arranged in a straight chain.

3. **Presence of an aldehyde group. (a)** Glucose reacts with hydroxylamine (NH_2OH) to form an oxime and adds a molecule of hydrogen cyanide to form cyanohydrins.



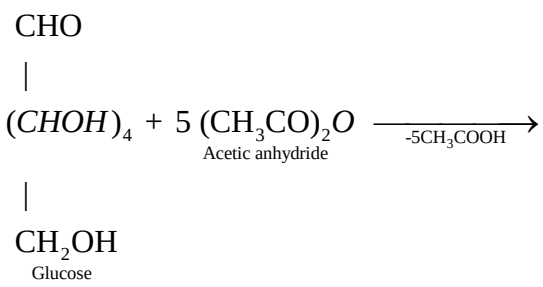


These reactions suggest that glucose contains a carbonyl group (C = O) group.
(b) Glucose on oxidation with a mild oxidizing agent like bromine water, gives gluconic acid containing the same six carbon atoms as present is glucose. This suggests that the carbonyl group present in glucose in an aldehydic group.

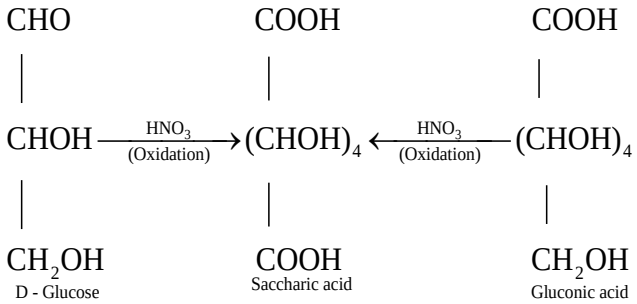


Since aldehyde (- CHO) is monovalent, it is always present at the end of the carbon chain.

4. **Presence of five hydroxyl groups** On acetylation with acetic anhydride, glucose gives a pentaacetate. This confirms that glucose contains five – OH groups. We know that the presence of two or more OH groups on the same carbon atom makes the molecule unstable. Glucose is a stable compound, therefore, the five – OH groups must be present on different carbon atoms.



5. **Presence of one primary alcoholic group** On oxidation with nitric acid, both glucose and gluconic acid give the same dicarboxylic acid, saccharic acid (also called glycaric acid). This indicates that glucose contains one primary alcoholic group.

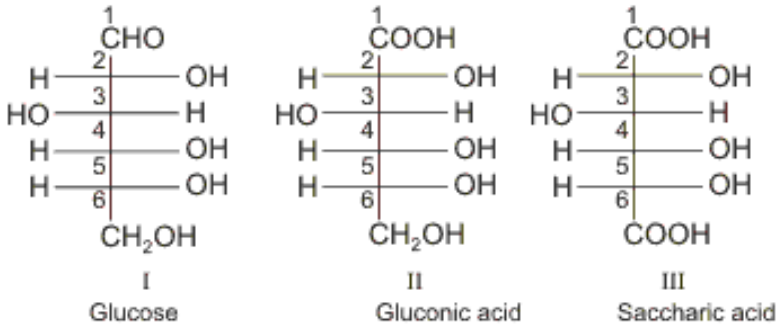


The primary alcohol group (- CH_2OH) is always present at the end of the carbon chain.

Open chain structure of glucose From the above discussion, it is evident that glucose has one aldehydic (- CHO) group and one primary alcoholic (- CH_2OH) group. Since, in all, glucose has five – OH groups, one of which is - CH_2OH group, the other four – OH groups must be secondary (- CHOH -) . In other words, in between the - CHO and – CH_2OH groups, there are four CHOH groups. Thus, the open chain structure of glucose is

$$OHC - \overset{*}{C}HOH\overset{*}{C}HOH\overset{*}{C}HOH\overset{*}{C}HOH - CH_2OH$$

The exact spatial arrangement of the different – OH groups was determined by **Emil Fischer** a German Scientist in 1891 through a remarkable piece of research work for which he got Nobel Prize in 1902. The exact configuration of glucose is represented as I. So, gluconic acid must have configuration II and saccharic acid configuration III.



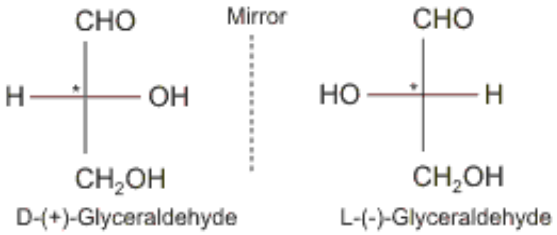
It may be noted here that in glucose, the - OH group at C – 3 chiral carbon is oriented toward left while the remaining three OH groups at C – 2, C – 4 and C – 5 chiral carbon atoms are oriented towards right.

14.2.5 Configuration of Monosaccharides

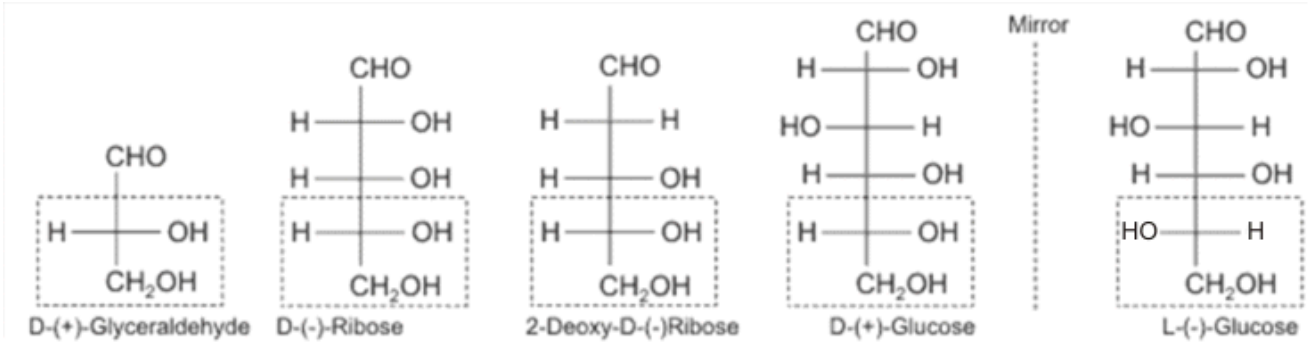
The correct name of glucose is D-(+) – Glucose. Since glucose is optically active, therefore, its enantiomer is called L-(-) – glucose. The letter 'D' and 'L' used before the name of glucose represent the configuration of the OH group at the penultimate (last but one) carbon atom of the carbon chain whereas the algebraic signs '+' and '-' within brackets after the letters D and L refer to the sign of optical rotation, i.e., dextrorotatory and laevorotatory respectively. It may, however, be noted that 'D' and 'L' have no relation with the sign of optical rotation, i.e., a carbohydrate having 'D' configuration may be either dextrorotatory or laevorotatory and vice versa.

For example, D-glucose is dextrorotatory while D-fructose in laevorotatory. The meaning of D- and L-notations is explained below:

D and L-Notations To assign D and L-configurations to monosaccharides, the simplest carbohydrate, glyceraldehydes, is chosen as the standard. It contains one chiral carbon atom and hence exists in two enantiomeric forms as shown below:



All the compounds which can be chemically correlated to (+) – isomer of glyceraldehydes are assigned D-configuration whereas all those compounds which can be correlated to (-) isomer of glyceraldehydes are assigned L-configuration.
Thus, a monosaccharide is assigned D-configuration if the – OH group at the last chiral carbon like that in D-(+) – glyceraldehydes lies towards right hand side. On the other hand, it is assigned L-configuration if the – OH group on the last chiral carbon atom like that in L- (-) – glyceraldehydes lies on the left hand side. For making this comparison, the structures of the monosaccharides are written is such as a way that the most oxidized carbon (i.e., - CHO group) is at the top. In the light of the above discussion (-) – ribose, 2-deoxy- (-) – ribose and (+) – glucose are assigned D configuration while (-) glucose is assigned L configurations respectively as illustrated below:



14.2.6 Cyclic structure of D-Glucose

Limitation of the open chain structure Although the open chain structure of D-(+) –glucose explains most of its reactions yet it fails to explain the following facts :

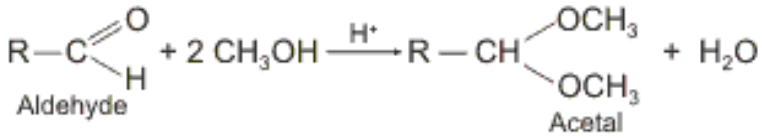
1. **D(+)-Glucose does not undergo certain characteristic reactions of aldehydes** For example, glucose does not form $NaHSO_3$ addition product, aldehyde-ammonia adduct, 2, 4-DNP derivative and does not respond to Schiff's reagent test.
2. **Glucose reacts with NH_2OH to form an oxime but glucose pentaacetate does not** This implies that the aldehyde group is absent in glucose pentaacetate.
3. **D(+)-Glucose exists in two stereoisomeric forms, i.e., α-glucose and β-glucose** α-D(+)-Glucose is obtained when a concentrated aqueous or alcoholic solution is crystallized at 303 K. It has a m.p. of 419 K and has a specific rotation of + 111°in a freshly prepared aqueous solution. However, when glucose is crystallized from water above 371 K, β-D(+)-glucose is obtained. It has a m.p. of 423 K and has a specific rotation of +19.2° in a freshly prepared aqueous solution.
4. **Both α-D-glucose and β-D-glucose undergo mutarotation in aqueous solution** Although the crystalline forms of α-and β-D(+)-glucose are quite stable in aqueous solution but each form slowly changes into the same equilibrium mixture of both containing 36% of α-D-glucose and 64% of β-D-glucose. This is evident from the fact that the specific rotation of a freshly prepared aqueous solution of α-D(+)-glucose falls gradually from + 111° to + 52.5° with time and that of β-D(+)-glucose increases from + 19.2° to 52.5°. Thus,

α-D(+) - Glucose ⇌ Equilibrium mixture ⇌ β-D(+) - Glucose

$[\alpha]_D^{298K} = +111^\circ$ $[\alpha]_D^{298K} = +52.5^\circ$ $[\alpha]_D^{298K} = +19.2^\circ$

This spontaneous change in specific rotation of an optically active compound with time, to an equilibrium value, is called **mutarotation.**

5. **D(+)-Glucose forms two isomeric methyl glucosides** Aldehydes normally react with two moles of methanol per mole of the aldehyde to form an acetal but D(+)-glucose when treated with methanol in presence of dry HCl gas, reacts with only one mole of methanol per mole of glucose to form a mixture of two methyl D-glucosides i.e. methyl α-D-glucoside (m.p.438 K; specific rotation + 158°) and methyl β-D-glucoside (m.p. 380 K; specific rotation - 33°)

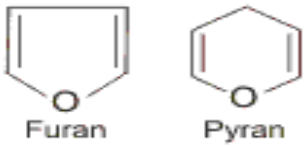


$$C_6H_{12}O_6 + CH_3OH \xrightarrow{H^+} (C_6H_{11}O_5)\,OCH_3 + H_2O$$

D - Glucose — Methyl α - and β - D - glucosides

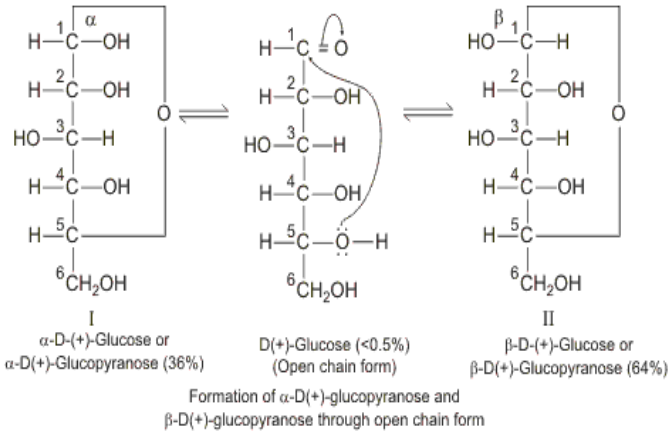
The two methyl D-glucosides behave as full acetals. Like acetals, these methyl glucosides are formed under anhydrous conditions but are hydrolysed back to glucose and methanol under aqueous acidic conditions. Further, these glucosides do not reduce Tollens' reagent and Fehling's solution and also do not react with HCN or NH_2OH indicating the absence of - CHO group.

Types of Cyclic structures of monosaccharides Monosaccharides do not have a free aldehyde or a ketonic group but instead have cyclic hemiacetal or hemiketal structures. Since monosaccharides contain a number of hydroxyl groups and an aldehyde or a keto group, therefore, any one of the – OH groups (usually C_4 or C_5 in aldohexoses and C_5 or C_6 in ketohexoses) may combine with the aldehyde or the keto group to form intramolecular hemiacetal or hemiketal. As a result, monosaccharides have five or six-membered cyclic structures containing an oxygen atom. The five membered ring containing one oxygen atom because of its similarity with furan is called the furanose form and the six membered ring containing one oxygen atom because of its resemblance with pyran is called the pyranose form.



All the monosaccharides (pentoses and hexoses) in the free state always exist in the pyranose form. However, in the combined state some monosaccharides such as ribose, 2-deoxyribose, fructose etc. usually exist in the furanose form. In contrast, trioses and tetroses have open chain structures.

Cyclic structure of Glucose-Anomers During hemiacetal formation, C_5 – OH of glucose combines with the C_1 becomes chiral or asymmetric and thus has two possible arrangements of H and OH groups around it. In other words, D-glucose exists in two stereoisomeric forms, i.e., α-D-glucose and β-D-glucose. Their Fischer projection formulae are shown in Fig.



Formation of α-D(+)-glucopyranose and β-D(+)-glucopyranose through open chain form

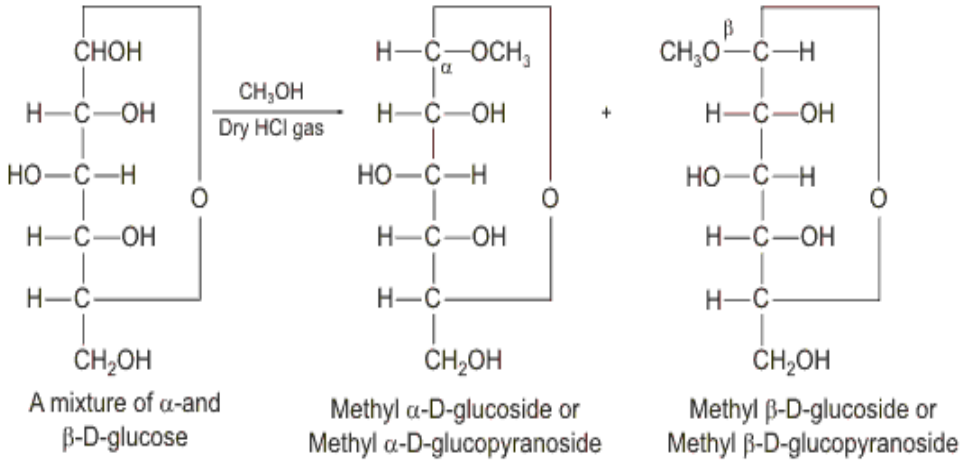
In α-D-glucose, the OH group at C_1 is towards right while in β-D-glucose, the OH group at C_1 is towards left. Such a pair of stereoisomer's which differ in configuration only around C_1 carbon is called the **anomeric carbon (or the glycosidic carbon).**

Explanation of the defects of the open chain structure The above cyclic hemiacetal structure of glucose explains all the objections raised against the open chain form as discussed below :

(i) The phenomenon of mutarotation exhibited by glucose anomers can be easily explained in terms of the cyclic structure. Each of these anomers is a typical hemiacetal and can be easily hydrolysed in the aqueous solution to give the open chain form of glucose. Therefore, in the aqueous solution either anomer is converted, through the open chain form, into an equilibrium mixture consisting of both the anomers with a very small amount of the open chain form.
Thus, in aqueous solution, glucose exists as an equilibrium mixture of two anomers, with a small amount (< 0.5%) of the open chain form. The α-anomer being less stable constitutes 36% while β-form being more stable constitutes 64% of the equilibrium mixture.
Thus, during mutarotation, the ring opens up to give the open chain form which then recloses either in the inverted position or in the original position giving an equilibrium mixture of two anomers with a small amount of the open chain form.
All reducing carbohydrates, i.e., monosaccharides and reducing disaccharides (maltose, lactose, etc. but not sucrose) undergo mutarotation in aqueous solutions.

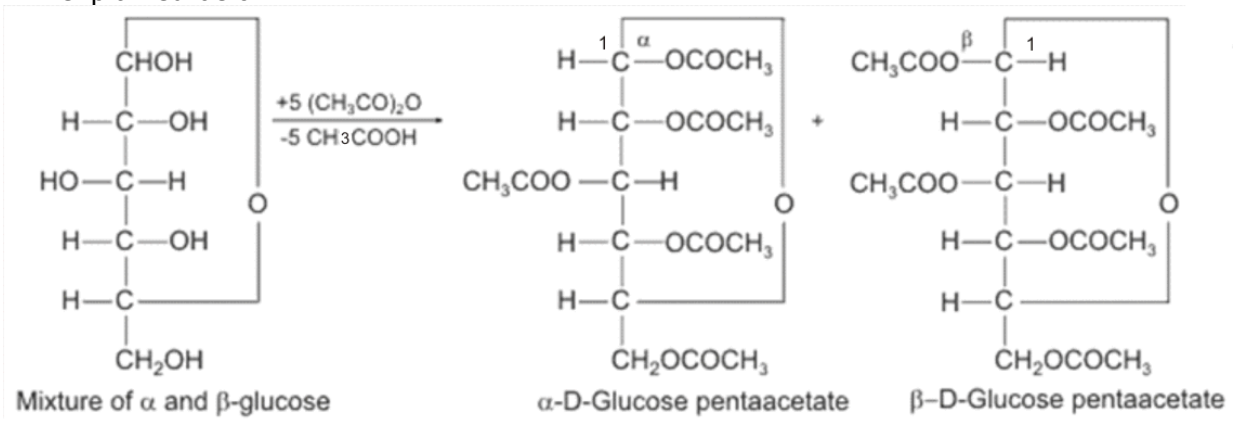
(ii) The typical reactions of D-(+)-glucose due to the presence of an aldehyde group such as cyanohydrins, oxime and osazone formation, reduction of the Fehling's reagent, etc. can be easily explained in terms of the small amount of the open chain form present in the equilibrium mixture of the two anomers.

(iii) The formation of two anomeric methyl glucosides can be explained in terms of the cyclic structure. The equilibrium mixture of α-and β-glucose react separately with methanol in presence of dry HCl gas to form the corresponding methyl D-glucosides.



These glucosides are full acetals since they are formed by the further reaction of hemiacetals with methanol. Being acetals, they are hdrolysed back to methanol and glucose only in the aqueous acidic medium. In other words, they are not hydrolysed in aqueous solution and hence do not undergo mutarotation. Further like ethers, acetals are stable towards alkalies. As such they are not hydrolysed by the alkali present in Fehling's solution or Tollens' reagent and hence do not reduce these reagents.

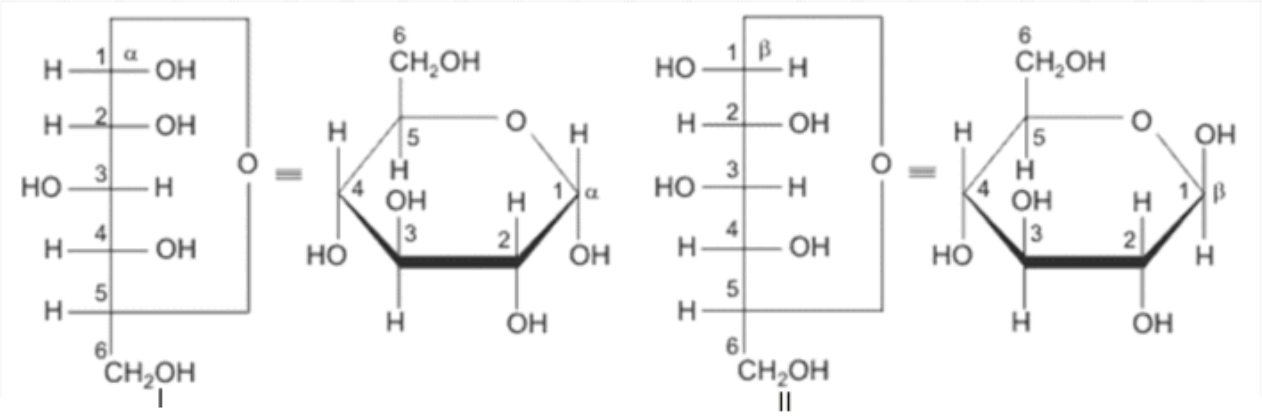
(iv) Like methyl glucosides, glucose pentaacetate also exists in two anomeric form as explained below :



These pentaacetates do not have a free – OH group at C_1 and hence are not hydrolysed in aqueous solution to produce the open chain aldehydic form and hence do not react with NH_2OH to form glucose oxime.

Haworth Projection Formulae of Monosaccharides

The cyclic structures of monosaccharides can be better represented by Haworth Projection Formulae. To get such a formula for any monosaccharide (say α-and β-D-glucose), draw a hexagon with its oxygen atom at the upper right hand corner. The lower edge is thickened to indicate that this edge is towards the observer. Place all the groups (on C_1, C_2, C_3, and C_4) which are present on left hand side in structures I and II, above the plane of the ring and all those groups on the right hand side below the plane of the ring. The terminal – CH_2OH group is always placed above the plane of the hexagon ring (in D-series). Following the above procedure, Haworth Projection Formulae for α-D-glucose (I) and β-D-glucose (II) are obtained as shown in Fig.



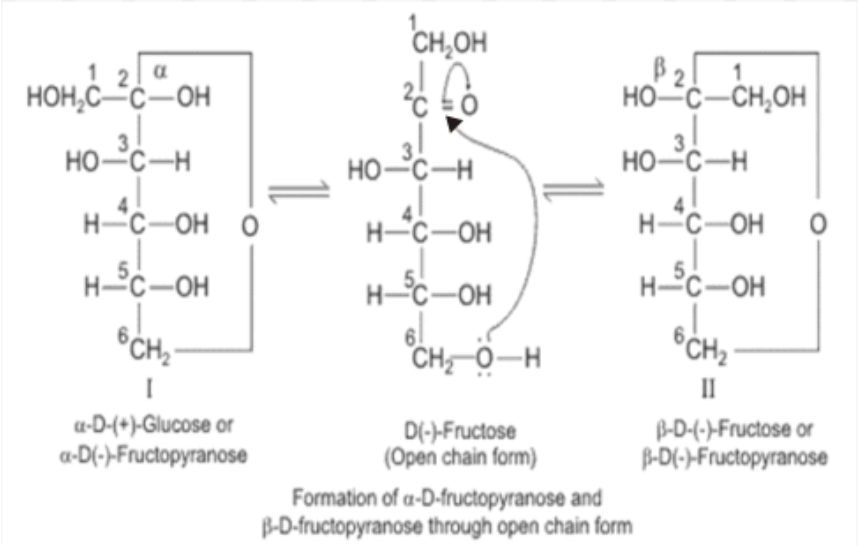
Fischer and Haworth projection formulae of α-D and β-D-glucopyranose

14.2.7 Fructose (Laevulose; Fruit sugar), $C_6H_{12}O_6$ Fructose is the most important of all the ketoses. It occurs free along with glucose (in the form of invert sugar) in honey and sweet fruits and hence the name fruit sugar. In the combined state, it is present in dissacharide (sucrose) and polysaccharide (inulin from dahlia tubers, etc.)

14.2.8 Structure of Fructose

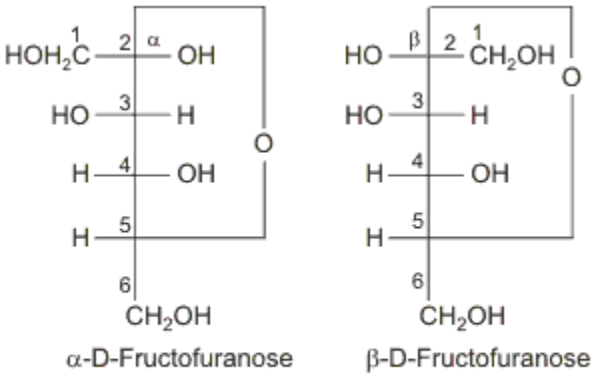
Fructose is obtained along with glucose by the hydrolysis of disaccharide, sucrose. Its molecular formula is $C_6H_{12}O_6$. On the basis of its reactions, it has been established that fructose contains a keto group at C-2 and the six carbon atoms are arranged in a straight chain as in case of glucose. It belongs to D-series and is laevorotatory.

Cyclic structure of Fructose Like glucose, fructose also has a cyclic structure. Since fructose contains a keto group, it forms an intramolecular hemiketal. In the hemiketal formation, C_6 – OH of the fructose combines with C_2-keto group. As a result, C_2 becomes chiral and thus has two possible arrangements of CH_2OH and OH groups around it. Thus, D-fructose exists in two stereoisomeric forms. i.e., α-D-fructopyranose and β-D-fructopyranose as shown in Fig.



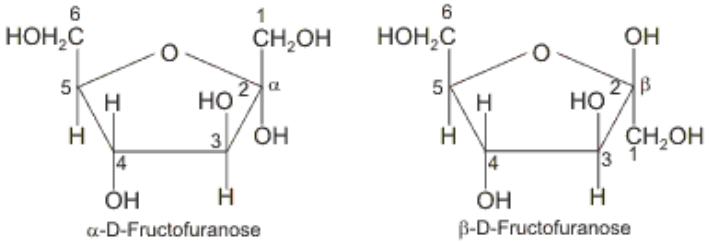
Formation of α-D-fructopyranose and β-D-fructopyranose through open chain form

However, in the combined state (such as in sucrose), fructose exists in the furanose form (Fig.)



Fischer projection formulae of α-D-fructofuranose and β-D-fructofuranose.

Their Haworth projection formulae are given in Fig.



Howorth projection formulae of α- and β-fructofuranose.

14.2.9 Disaccharides

Carbohydrates which upon hydrolysis give two molecules of the same or different monosaccharides are called ***disaccharides***. Their general formula is $C_{12}H_{22}O_{11}$. The three most important disaccharides are sucrose, maltose and lactose. Each one of these on hydrolysis with either an acid or an enzyme gives two molecules of the same or different monosaccharides as shown below :

$$\underset{\text{Sucrose}}{C_{12}H_{22}O_{11}} + H_2O \xrightarrow[\text{or Invertase}]{H^+} \underset{\text{Glucose}}{C_6H_{12}O_6} + \underset{\text{Fructose}}{C_6H_{12}O_6}$$

$$\underset{\text{Sucrose}}{C_{12}H_{22}O_{11}} + H_2O \xrightarrow[\text{or Maltase}]{H^+} \underset{\text{Glucose}}{C_6H_{12}O_6} + \underset{\text{Glucose}}{C_6H_{12}O_6}$$

$$\underset{\text{Sucrose}}{C_{12}H_{22}O_{11}} + H_2O \xrightarrow[\text{or Lactase}]{H^+} \underset{\text{Glucose}}{C_6H_{12}O_6} + \underset{\text{Galactose}}{C_6H_{12}O_6}$$

The two monosaccharide units are joined together through an ethereal or oxide linkage formed by loss of a molecule of H_2O. *Such a linkage between two monosaccharide units through oxygen atom is called* ***glycosidic linkage****.*

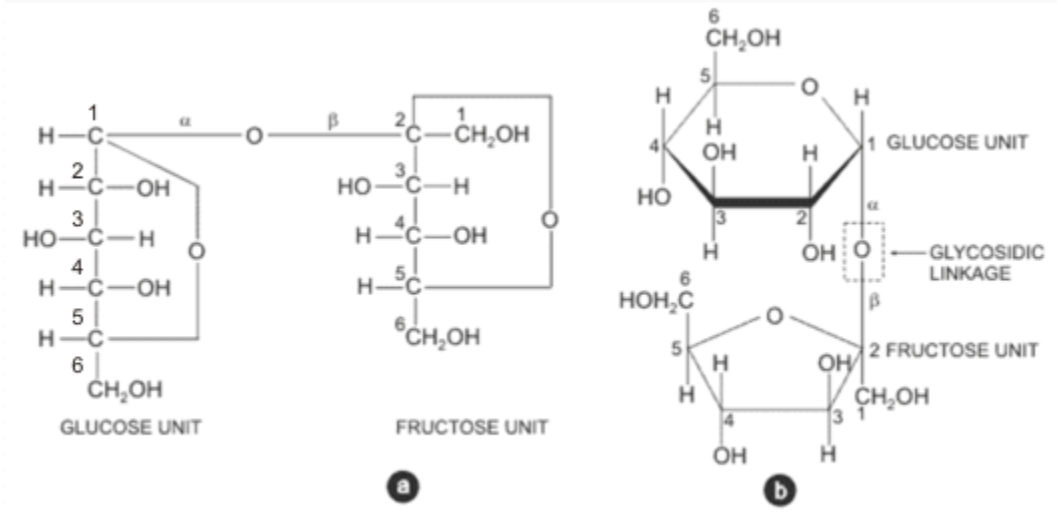
If the two monosaccharide units are linked through their respective carbonyl groups (i.e. reducing centres), the disaccharide is said to be non-reducing, e.g., sucrose. On the other hand, if the carbonyl group of any one of the two monosaccharide units is free, the disaccharide is said to be reducing, e.g., maltose, lactose, etc.

1. **Sucrose, Cane-sugar (table sugar), $C_{12}H_{22}O_{11}$** It is manufactured either from sugarcane or beet root. It is a colourless, crystalline, water soluble sweet substance. On hydrolysis with dilute acids or enzyme invertase or sucrose, it gives an equimolar mixture of D(+)- glucose and D(-)-fructose

$$\underset{\substack{\text{Sucrose}\\ [\alpha]_D = +66.5^\circ}}{C_{12}H_{22}O_{11}} + H_2O \xrightarrow{HCl} \underset{\substack{\text{D-Glucose}\\ [\alpha]_D = +52.5^\circ}}{C_6H_{12}O_6} + \underset{\substack{\text{D-Fructose}\\ [\alpha]_D = -92.4^\circ}}{C_6H_{12}O_6}$$

Inversion of sugar Sucrose is dextrorotatory with a specific rotation of $[\alpha]_D = +66.5^\circ$ but on hydrolysis, it gives dextrorotatory glucose and laevorotatory fructose. Since the laevorotation of fructose (-92.4°) is more than dextrorotation of glucose (52.5°), therefore, the resulting solution becomes laevorotatory. Further since the hydrolysis of sucrose to an equimolar mixture of D (+)-glucose and D(-)-fructose is accompanied by a change in the sign of optical rotation from dextrorotatory to laevorotatory, the overall process is called inversion of sugar and the equimolar mixture of D(+)-glucose and D(-)-fructose obtained as a result of hydrolysis is called the invert sugar. Invert sugar (130) is sweeter than sucrose (100), glucose (70) galactose (32) and lactose (10) but less sweet than fructose (170).

Structure Enzymatic studies have revealed that C_1 - α of glucose is connected to C_2 - β of fructose. Further, determination of ring size has revealed that in sucrose, glucose is present in the pyranose from while fructose is present in the furanose form. On the basis of these evidences, Haworth in 1927, suggested the following structure for sucrose.



(a) Fischer (b) Haworth projection formulae of surcrose.

1. **Maltose (Malt sugar), $C_{12}H_{22}O_{11}$** It is obtained by partial hydrolysis of starch by the enzyme diastase present in malt, i.e., sprouted barley seeds.

$$\underset{\text{Starch}}{2(C_6H_{10}O_5)n} + n\,H_2O \xrightarrow{\text{Diastase}} n\,\underset{\text{Maltose}}{C_{12}H_{22}O_{11}}$$

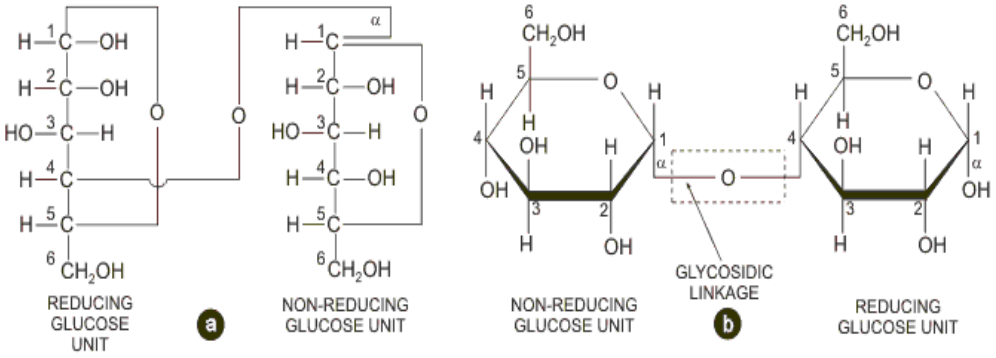
Maltose is a reducing sugar since it forms an osazone*, undergoes mutarotation and also reduces Tollens' and Fehling's solutions, Methylation studies have revealed that

(i) Both glucose units are present in the pyranose form.

(ii) C_1 of one glucose unit is linked to C_4 of the other.

Further since maltose is hydrolysed by the enzyme maltase which specifically hydrolyses α-glycosidic linkages, therefore, the non-reducing glucose unit in maltose must be present in the

α-form. In other words, C_1 - α of non-reducing glucose unit is attached to C_4 of the reducing glucose unit as shown in Fig.



(a) Fischer (b) Haworth projection formulae of α-D-maltose

2. Lactose (Milk sugar), $C_{12}H_{22}O_{11}$

Lactose occurs in milk (6-8% in human milk and 4-5% in cow's milk) that is why it is also called milk sugar.

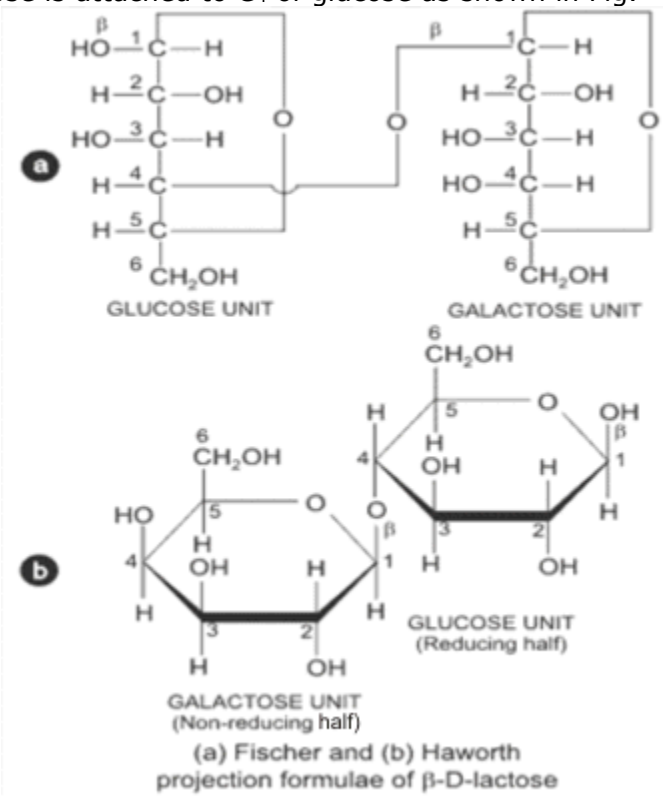
Lactose, on hydrolysis with dilute acids, yields an equimolar mixture of D-glucose and D-galactose. It is a reducing sugar since it forms an osazone, undergoes mutarotation and also reduces Tollens' or Fehling's solution. Methylation studies have revealed that

(i) Both glucose and galactose are present in the **pyranose form**

(ii) Glucose is the reducing half while galactose is the non-reducing half.

(iii) C_1 of galactose unit is connected to C_4 of glucose unit.

Further since emulsin, an enzyme which specifically hydrolyses β-glycosidic linkage also hydrolyses lactose, therefore, galactose must be present in the β-form. In other words, in lactose, C_1 - β of galactose is attached to C_4 of glucose as shown in Fig.



(a) Fischer and (b) Haworth projection formulae of β-D-lactose

14.2.10 Polysaccharides:

Polysaccharides are formed when a large number (hundreds to even thousands) of monosaccharide molecules join together with the elimination of water molecules. Thus, polysaccharides may be regarded as condensation polymers in which the monosaccharides (monomers) are joined together by glycosidic linkage. Some important polysaccharides are :

1. Starch
2. Cellulose
3. Glycogen
4. Dextrins

Starch, Amylum, $(C_6H_{10}O_5)n$*

Occurrence The value of n (200-1000) varies from source to source. It is the chief food reserve material or storage polysaccharide of plants and is found mainly in seeds, roots, tubers, etc. Wheat, maize, rice, potatoes, barley, bananas and sorghum are the main sources of starch.

Properties (i) Starch is a white amorphous powder sparingly soluble in water.

(ii) On hydrolysis with dilute mineral acids or enzymes, starch breaks down first to smaller molecules (n > n'), then to maltose and finally to D-glucose.

$$\underset{\text{Starch}}{(C_6H_{10}O_5)_n} \xrightarrow{H^+/H_2O} (C_6H_{10}O_5)_{n'}$$

$$\downarrow H^+/H_2O$$

$$\underset{\text{D-Glucose}}{C_6H_{12}O_6} \xleftarrow[\text{or Maltase}]{H^+/H_2O} \underset{\text{Maltose}}{C_{12}H_{22}O_{11}}$$

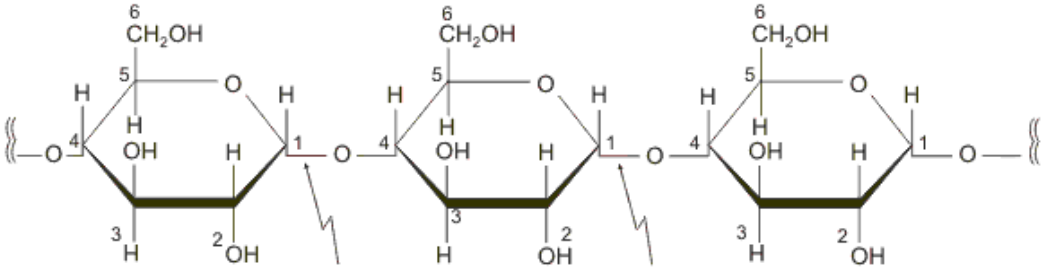
(iii) Starch is a non-reducing saccharide. It neither reduces Tollens' reagent (or Fehling's solution) nor forms an osazone. This suggests that all hemiacetal OH groups of glucose units at C_1 are not free but are involved in glycosidic linkage.

Composition Starch is not a single compound but is a mixture of two components – a water soluble component called amylase (15-20%) and a water insoluble component called amylopectin (80-85%). The aqueous solution of amylase gives a blue colour with iodine solution due to the formation of an inclusion complex. The blue colour, however, disappears on heating and reappears on cooling.

Amylopectin, on the other hand, does not give blue colour with iodine solution.

Both amylase and amylopectin are polymers of α-D(+)-glucose. A molecule of amylase may contain 200-1000 units whereas a molecule of amylopectin may contain 2000-3000 glucose units.

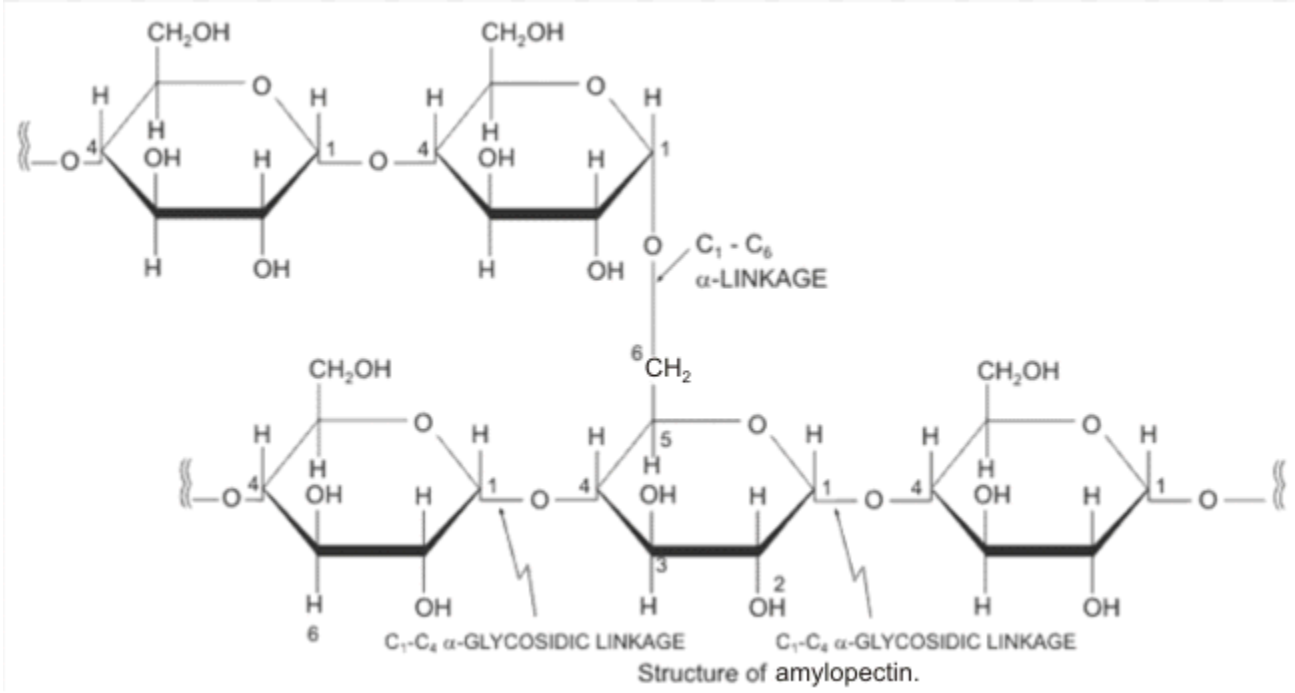
Structure of amylase It is a linear polymer of α-D-glucose in which C_1 of one glucose unit is attached to C_4 of the other through α-glycosidic linkage as shown in Fig.



Structure of amylose.

Structure of amylopectin Amylopectin, on the other hand, is a highly branched polymer. It consists of a large number (several hundred) of short chains each containing 20-25 glucose units which are joined together through α-glycosidic linkages involving C_1 of one glucose unit

with C_4 of the other. The C_1 of terminal glucose unit in each chain is further linked to C_6 of some other glucose unit in the next chain through $C_1 - C_6$ α-glycosidic linkage. This gives amylopectin a highly branched structure.



Structure of amylopectin.

Starch is the major food material for us. It is hydrolysed by the enzymes amylase present in saliva. The end product of this hydrolysis is glucose which is an essential nutrient.

14.2.12 Cellulose, $(C_6H_{10}O_5)_n$

The molecular mass of cellulose varies from 50,000 to 500,000 suggesting thereby that cellulose may contain 300-3000 glucose units.

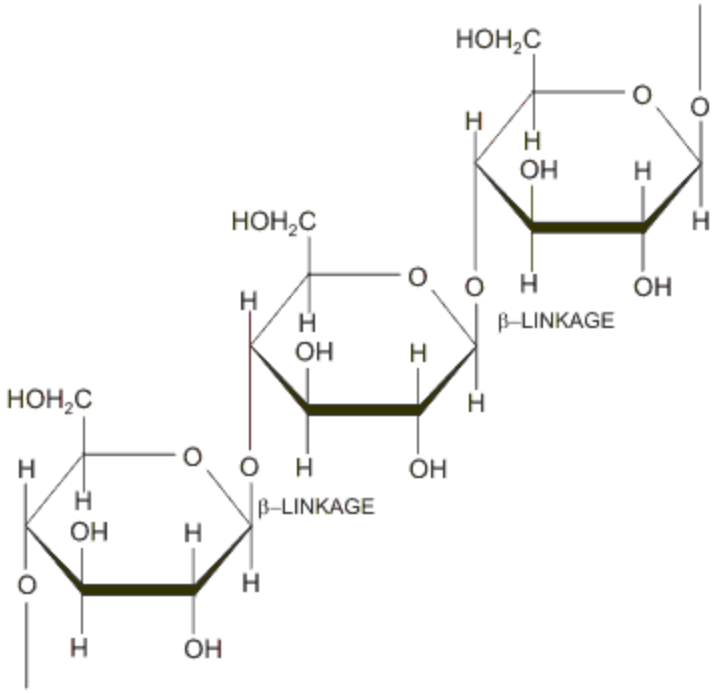
It is insoluble in water but dissolves is Schweitzer's reagent, i.e., ammoniacal solution of cupric hydroxide from which it can be precipitated by addition of alcohols, acids or salts. This property is used in the manufacture of artificial silk (rayon) from cellulose.

Cellulose is widely distributed in plants. In fact, more than 50% of all the living matter is cellulose. It is the chief structural material of cell walls of all plants. It is also the chief component of cotton, wood, jute, etc.

Properties Like starch, cellulose is also non-reducing since it does not reduce Tollens' reagent or Fehling's solution. It also does not form an osazone and is not fermented by yeast. Unlike starch, it does not undergo hydrolysis easily. However, on heating with dil. H_2SO_4 under pressure, it does undergo hydrolysis to give only D-glucose.

Concentrated NaOH acts upon cellulose to form gelatinous translucent mass which imparts luster to cotton. This process is called mercerization and the cotton thus treated is called mercerized cotton.

Cellulose is a linear polymer of β-D-glucose in which C_1 of one glucose unit is connected to C_4 of the other through β-D-glucosidic linkage as shown in Fig.



Structure of cellulose

X-ray analysis has shown that cellulose is made up of several linear long chains (each containing about 3000 or more glucose units). These chains lie side by side to form bundles held together by H-bonding between $^{6}CH_2$ – OH of one chain with C_2 – OH group of the other chain.

14.2.13 Glycogen $(C_6H_{10}O_5)_n$ Like starch, glycogen is also a condensation polymer of α-D-glucose since it is hydrolysed by the enzyme β-amylase to form maltose. Just as glucose is stored in plants as starch, it is stored as glycogen in human beings. It is present in liver, muscles and brain.
Glycogen is also called animal starch because its structure is similar to that of amylopectin. Glycogen is more highly branched than amylopectin.
14.2.14 Importance of Carbohydrates The main functions and uses of carbohydrates are described below :
1. Living systems
(a) Biofuels Carbohydrates are essential for life in both plants and animals. They form a major portion of our food. Honey has been used for a long time as an instant source of energy by Vaids in Ayurvedic system of medicine. Thus, carbohydrates provide energy for the functioning of living systems and thus act as biofuels.

$$\underset{\text{Glucose}}{C_6H_{12}O_6} + 6\,O_2 \rightarrow 6\,CO_2 + 6\,H_2O + 2880\ \text{kJ}$$

In the living systems, the polysaccharides such as starch and glycogen are first hydrolysed by enzymes to glucose. This is then transported from one cell to the other by the blood in animals and cell sap in plants. Oxidation of glucose to carbon dioxide and water by a series of enzyme-catalysed reactions provides energy for the functioning of the cells.
(b) Reserve food materials The polysaccharide starch is the major reserve food material in plants. It is stored in seeds and acts as the reserve food material for the tiny plant till it is capable of making its own food by photosynthesis. Similarly, glycogen stored in liver and muscles acts as the reserve food material in animals.
(c) Constituents of biomolecules Two aldopentoses, i.e., D-ribose and 2-deoxy-D-ribose are the essential components of RNA and DNA respectively. These nucleic acids essentially control

the transmission of hereditary effects from one generation to the other and also the biosynthesis of proteins.

14.3 Proteins

Proteins Another class of biomolecules, essential for living beings, are proteins. The chief sources of proteins are milk, cheese, pulses, peanuts, fish, meat, etc. These are high molecular mass complex biopolymers of amino acids. They are so named because proteins are vital chemical substances essential for the growth and maintenance of life.

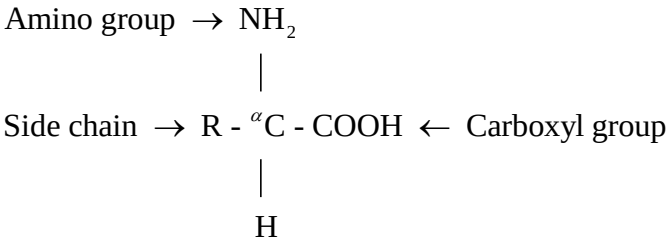
Chemically proteins are condensation polymers (actually polyamides) in which the monomeric units are the α-amino acids.

All proteins contain the elements carbon, hydrogen, oxygen, nitrogen and sulphur. Some of these may also contain phosphorus, iodine and traces of metals such as Fe, Cu, Zn and Mn. All proteins on partial hydrolysis give peptides of varying molecular masses which upon complete hydrolysis give α-amino acids.

$$\text{Proteins} \xrightarrow{\text{Hydrolysis}} \text{Peptides} \xrightarrow{\text{Hydrolysis}} \alpha\text{-Amino acids}$$

14.3.1 α-Amino Acids

Amino acids contain amino (-NH_2) and carboxyl (- COOH) as functional groups. Depending upon the relative position of the – NH_2 group w.r.t. – COOH group, the amino acids are classified as α-, β-, γ-,δ-, etc. amino acids. Only α-amino acids are obtained by hydrolysis of proteins. Thus, α-amino acids are the building blocks of proteins.

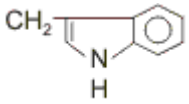
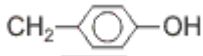
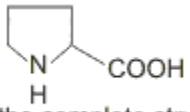
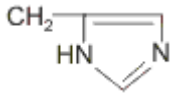


14.3.2 Nomenclature of α-Amino acids

Although amino acids can be named by the IUPAC system, they are generally known by their common names. For example, H_2NCH_2COOH is better known as glycine rather than α-aminoacetic acid or 2-aminoethanoic acid. These trivial names are derived either from the name of the source or the name of some characteristic property of that amino acid. For example, glycine is so named since it has sweet taste and tyrosine was first obtained from cheese. For sake of simplicity, each amino acid has been given a standard abbreviation or a code which usually consists of the first three letters of the common name. for example, glycine is abbreviated as Gly and alanine is abbreviated as Ala.

The structures, names, abbreviations and one letter symbols of the 20 most commonly occurring amino acids are given in the table.

Table: Structure of natural amino acids

Amino acid	Three letter Symbol	One letter Code	Side chain (R)	Isoelectric Point***
	Neutral aminoacids			
1. Glycine	Gly	G	H	6·1
2. Alanine	Ala	A	CH_3	6·1
3. Valine	Val	V	$CH(CH_3)_2$	6·0
4. Leucine	Leu	L	$CH_2CH(CH_3)_2$	6·0
5. Isoleucine	Ileu	I	$CHCH_2CH_3$ | CH_3	6·0
6. Phenylalanine	Phe	F	$CH_2C_6H_5$ or CH_2Ph**	5·9
7. Methionine	Met	M	$CH_2CH_2 - S - CH_3$	5·7
8. Tryptophan	Trp	W		5·9
9. Serine	Ser	S	CH_2OH	5·7
10.Crysteine	Cys	C	CH_2SH	5·1
11. Glutamine	Gln	O	$CH_2CH_2CONH_3$	5·7
12. Threonine*	Thr	T	$CHCH_3$ | OH	5·7
13. Tyrosine	Tyr	Y		5·6
14.Proline	Pro.	P	 (It is the complete structure)	6·3
	Acidic amino acids			
	Asp	D	CH_2COOH	3·0
	Asn	N	CH_2CONH_2	5·4
15. Aspartic acid	Glu	E	CH_2CH_2COOH	3·2
	Basic amino aicds			
16. Asparagine	Lys	K	$CH_2(CH_2)_3NH_2$	10·8
17. Glutamic acid	Arg	R	$CH_2(CH_2)_2NH - C - NH_2$ || NH	
18. Lysine*	His	H		7·6
19. Arginine*				
20. Histidine*				

14.3.3 Classification of α-Amino acids

Neutral, acidic and basic amino acids Amino acids can be broadly classified into three classes, i.e., neutral, acidic or basic amino acids. Amino acids such as glycine, alanine, valine, etc. which contain one – NH_2 and one – COOH group, are called neutral amino acids. Those amino acids such as aspartic acid, asparagines and glutamic acid which contain two – COOH groups and one – NH_2 group, are called acidic amino acids, and amino acids like lysine, arginine

and histidine which contain two – NH_2 groups and one – COOH group, are called basic amino acids.

Essential and non-essential α-amino acids Amino acids have also been classified as essential and non-essential amino acids. Out of the 20 amino acids required for protein synthesis, human body can synthesize only 10. These ten amino acids which the body can synthesize are called non-essential or dispensable amino acids while the remaining ten which the human body cannot synthesize are called essential or indispensable amino acids.)

Complete and incomplete proteins Proteins that provide all the essential amino acids in about the right proportion for human nutrition are called complete proteins. For example, meat, fish, milk, and eggs contain complete proteins. Proteins which are deficient in one or more of the essential amino acids are called incomplete proteins.

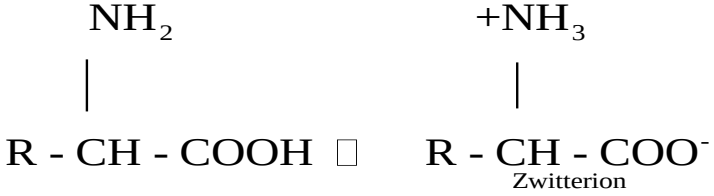
14.3.4 Physical Properties of α-Amino Acids

(i) Amino acids are usually colourless, non-volatile crystalline solids.

(ii) They are insoluble in non-polar solvents like petroleum ether, benzene but are appreciably soluble in water.

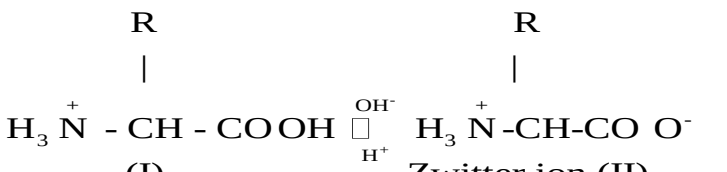
(iii) Their aqueous solutions behave like solution of substances of high dipole moment.

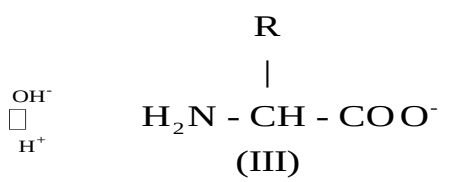
(iv) As is evident from their general formula, amino acids contain both an acidic carboxyl group and a basic amino group. These two groups interact and thus affect the properties of each other. They neutralize each other involving transfer of a proton from the carboxyl group to the amino group within the molecule. As a result, amino acids largely exist as dipolar ions.



The dipolar structure is commonly known as internal salt or zwitterions.

(v) Isoelectric points of α-amino acids Because of the zwitterions structure, α-amino acids are high melting crystalline solids. They are moderately soluble in water. In acidic medium, α-amino acids exist as cations (I) and thus migrate towards cathode under the influence of an electric field. On the other hand, in alkaline medium, α-amino acids exist as anions (III) and thus migrate towards anode under the influence of an electric field. However, at some intermediate value of pH, the concentration of the cationic form (I) and anionic form (III) will become equal and consequently the α-amino acid will exist primarily as the neutral dipolarion (II). At this pH, there would be no net migration of the amino acid in an electric field. This pH at which there is no net migration of the amino acid under the influence of an applied electric field is called isoelectric point.

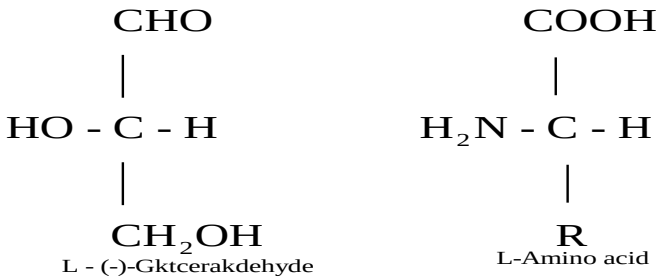




Each amino acid has a characteristic isoelectric point.
For neutral amino acids, isoelectric point is slightly less than pH7 (glycine = 6.1), for acidic amino acids, it lies between pH 3.0 – 5.4 (aspartic acid = 3.0) and for basic amino acid, it lies between pH 7.6 – 10.8 (lysine = 9.7)

(vi) The acidity and basicity constants of amino acids are very low. For example, glycine has $K_a = 1.6 \times 10^{-10}$ and $K_b = 2.5 \times 10^{-12}$ whereas most carboxylic acids have K_a values of about 10^{-5} and most aliphatic amines have K_b values of about 10^{-4}. The reason being that in α-amino acids, it is the NH_3 group which acts as an acid and $-COO^-$ group acts as the base.

(vii) **Configuration** The α-carbon of all the amino acids (except glycine) is chiral (asymmetric) and hence amino acids can exist in two stereoisomeric forms, i.e., D and L*. However, all the naturally occurring amino acids belong to the L-series.

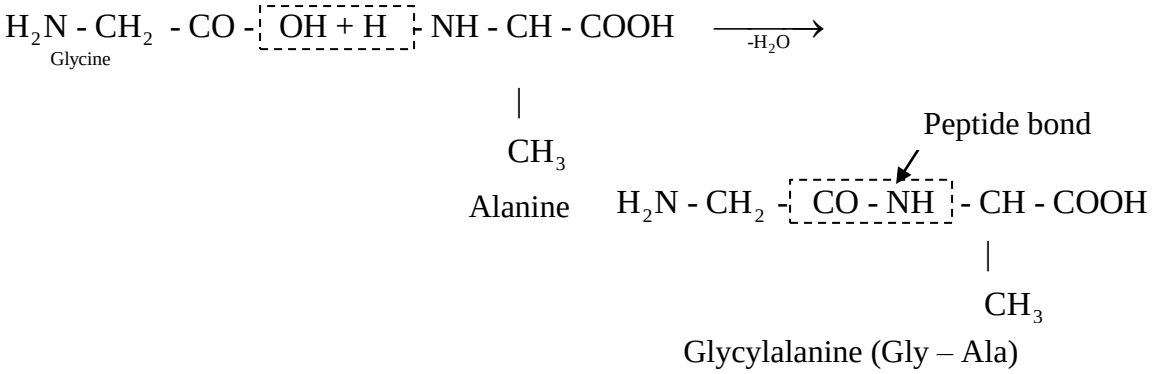


In other words, proteins exclusively consist of L-amino acids. In contrast, all the sugars belong to D-series.

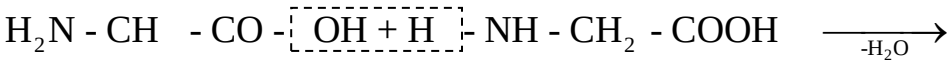
14.3.5 Peptides and their Classification

Peptides are amides formed by the condensation of amino group of one α-amino acid with the elimination of a molecule of water. They are classified as di-, tri-, tetra-, pentapeptide, etc.

For example, when – COOH group of glycine combines with the NH_2 group of alanine, we get the dipeptide, glycylalanine.



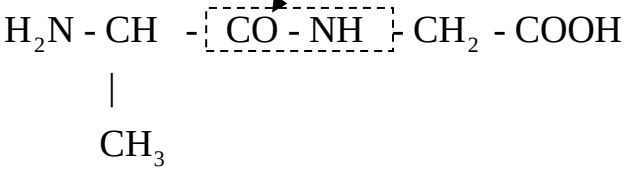
Alternatively, the – NH_2 group of glycine may react with - COOH group of alanine resulting in the formation of a different dipeptide, alanylglycine (Ala – Gly).



Glycine

Alanine

Peptide bond

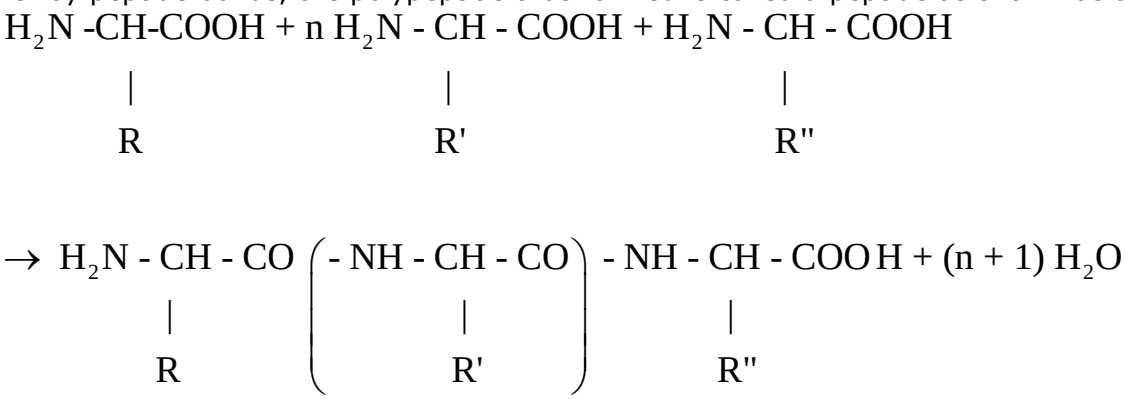


Alanylglycine (Ala – Gly)

The $-C-NH-$ bond is called the peptide bond or the peptide linkage.

In both the dipeptides, i.e., glycylalanine and alanylglycine, there is a free amino group at one end and a free carboxyl group at the other end. These groups can further react with appropriate functional groups of other α-amino acids to form tri-, tetra-, pentapeptides, etc.

14.3.6 Polypeptides If a large number (hundreds to thousands)of α-amino acids are joined together by peptide bonds, the polypeptide thus formed is called a peptide as shown below:

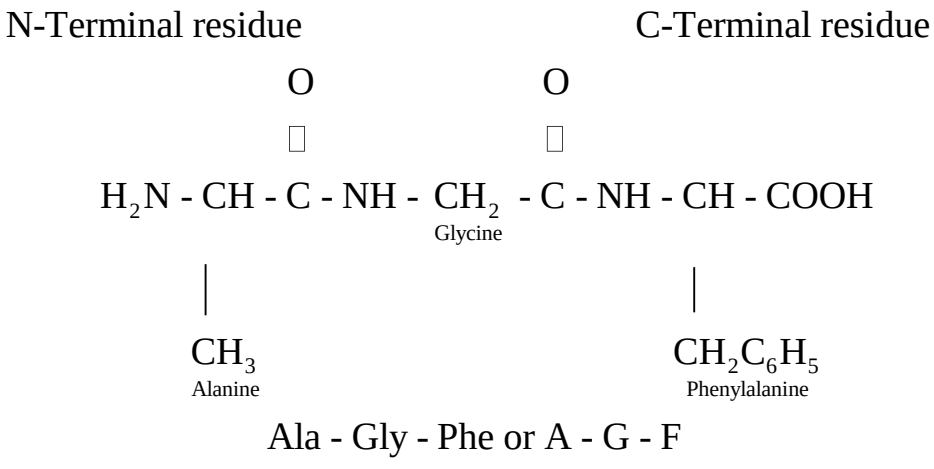


A polypeptide

Formation of a polypeptide

The amino acid unit having the free – NH_2 group is called the **N-terminal amino acid** whereas the amino acid unit having the free – COOH group is called the **C-terminal amino acid**.

Writing and naming of polypeptides As a matter of convention, the structures of polypeptides are written in such a way that amino acid with the free amino (- NH_2) group is written on the left hand side of the polypeptide chain while the amino acid with the free carboxyl (-COOH) group is written on the right hand side of the chain. Thus a tripeptide, alanyl glycylphenylalamine (Al-Gly-Phe) is represented as follows :



Ala - Gly - Phe or A - G - F

The name of any polypeptide is written starting from the N-terminal residue. While writing the name, the suffix 'ine' in the name of the amino acid is replaced by 'yl' (e.g., glycyl for glycine and alanyl for alanine, etc.) for all the constituent α-amino acids except the C-terminal residue. This nomenclature is usually not used. Instead the three letter abbreviations or one letter codes for the various α-amino acids except the C-terminal residue. This nomenclature is usually not used. Instead the three letter abbreviations or one letter codes for the various α-amino acids present in the chain is used. For example, the above tripeptide is named as Ala-Gly-Phe or A – G – F.

14.3.7 Oligopeptides, Polypeptides and Proteins

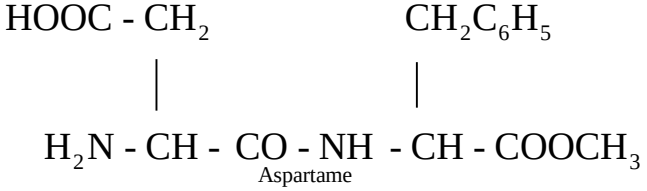
Peptides having relatively shorter chains (having 2-10 amino acid residues) are called oligopeptides while those with longer chains (greater than 10 amino acid residues) are called polypeptides. A polypeptide with more than hundred or so α-amino acid residues having molecular mass greater than 10,000 u is called a protein.

If we represent the different amino acids present in a protein by the symbols, aa_1, aa_2, aa_3, aa_4 …….etc. and a peptide bond by a dash (-), then a protein molecule may be represented as $aa_1 - aa_2 - aa_3 - aa_4$……..

Properties (i) Polypeptides are amphoteric in character because of the presence of terminal ammonium and carboxylate ions as well as the ionized side chains of amino acid residues. Therefore, like α-amino acids, they neutralize both acids as well as bases and possess isoelectric points. At isoelectric points, polypeptides have least solubility and hence can be separated.

(ii)The total content of smaller peptides in tissues is small as compared to proteins. Nevertheless, like proteins, they also carry out many important functions in bio-systems. Most of the toxins (poisonous substances) in animal venoms and in plant sources are polypeptides. Some oligopeptides such as oxytoxin, vasopressin and angiotensin-II are effective hormones while others act as sweetening agents.

For example, the dipeptide, aspartylphenylalanine methyl ester, called aspartame, is 100 times sweeter than sucrose and is used as a sugar substitute for diabetic patients.



14.3.8 Classification of proteins

Proteins have been classified in a number of different ways. The two most widely used systems of classification are discussed below :

1. On the basis of molecular structure, proteins have been classified as :

(i) Fibrous proteins and

(ii) Globular proteins

(i) Fibrous proteins These proteins consist of linear thread-like molecules which tend to lie side by side to form fibres. The polypeptide chains in them are held together usually at many points by hydrogen bonds and some disulphide bonds. As a result, intermolecular forces of attraction are very strong and hence fibrous proteins are insoluble in water. Further, these proteins are stable to moderate changes in temperature and pH.

(ii) Globular proteins The polypeptide chain in these proteins is folded around itself in such a way so as to give the entire protein molecule an almost spheroidal shape. The folding takes place in such a manner that lipophilic (non-polar, fat soluble) parts are pushed inwards and hydrophilic (polar, water soluble) parts are pushed outwards. As a result, water molecules interact strongly with the polar groups and hence globular proteins are water soluble. As compared to fibrous proteins, these are very sensitive to small changes in temperature and pH.

This class of proteins includes all enzymes, many hormones such as insulin from pancreas, thyroglobulin from thyroid gland (which produces the hormone thyroxine), antibodies (responsible for allergies and defence against foreign organisms), haemoglobin which transports oxygen from the blood to the muscles, fibrinogen which is converted into the insoluble protein fibrin and thus causes clotting of the blood, albumin in eggs etc.

Each globular protein has its own characteristic geometry which results through the interaction between different sites in the same polypeptide chain. These interactions are usually of four types :

(i) Disulphide bridging
(ii) Intramolecular hydrogen bonding
(iii) Van der Waals interactions
(iv) Dipolar interactions

2. On the basis of chemical composition or the nature of products they give upon hydrolysis , proteins can also be classified as follows :

(i) Simple proteins: These proteins on hydrolysis give only α-amino acids. For example, albumin in the white of eggs, glutenin in wheat, oxyzenin in rice, keratin in hair, nails, horns, etc.

(ii) Conjugated proteins These proteins on hydrolysis give a non-protein portion in addition to the α-amino acids. This non-protein is called the prosthetic group.

The main function of the prosthetic group is to control the biological functions of the protein.

Types of conjugated proteins

Name of the protein	**Prosthetic group**
Nucleoproteins	Nucleic acids
Glycoproteins or Mucoproteins	Sugars
Lipoproteins	Lipids such as lecithin
Phosphoproteins	Phosphoric acid residues
Chromoproteins (Haemoglobin and Myoglobin)	Some colouring matter such as red coloured protoporhyrin (Heme)

(iii) Derived proteins These are the degradation products obtained by partial hydrolysis of simple or conjugated proteins with acids, alkalies or enzymes. For example, proteoses, peptones and polypeptides.

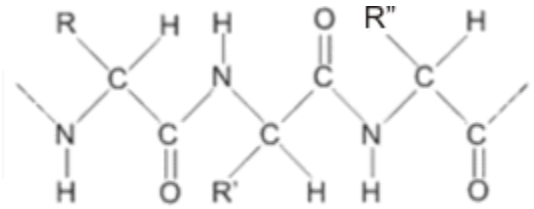
$$\text{Proteins} \rightarrow \text{Proteoses} \rightarrow \text{Peptones} \rightarrow \text{Polypeptides}$$

14.3.9 Structures of Proteins

Proteins are biopolymers obtained by condensation of a large number of α-amino acids joined together through peptide bonds having three-dimensional (3 D) structures.

The structures of proteins are quite complex. Their complete structures are usually discussed at four different levels, i.e., primary, secondary, tertiary and quaternary structures, each level being more complex than the previous one.

(i) Primary structure Proteins may have one or more polypeptide chains. Each polypeptide chain has a large number of α-amino acids linked to one another in a specific sequence. The specific sequence in which the various α-amino acids present in a protein are linked to one one another is called its primary structure. Any change in this primary structure, i.e., sequence of amino acids creates a different protein.



Primary structure of a protein

In other words, primary structure of a protein refers to the covalent structure including the disulphide bridges of each polypeptide chain.

The primary structure of a protein is usually determined by its successive hydrolysis with either enzymes or mineral acids into various products having decreasing molecular mass as shown below :

$$\text{Proteins} \rightarrow \text{Proteoses} \rightarrow \text{Peptones} \rightarrow \text{Polypeptides} \rightarrow \text{Simple peptides} \rightarrow \alpha\text{-Amino acids}$$

The first ever primary structure of a protein i.e. insulin* was determined by the British chemist, Frederic Sanger and for this work he was awarded the Nobel Prize in 1958. Since then, the primary structures of hundreds of different proteins have been determined.

The importance of the primary structure of a protein in determining its biological activity is shown by the fact that replacement of just one amino acid in the sequence of a protein destroys its biological activity. For example, the replacement of one specific amino acid (glutamic acid by valine) in haemoglobin, the protein of the blood, results in defective haemoglobin which causes a disease called sickle cell anaemia.

-Val-His-Leu-Thr-Pro- Glu -Glu-Lys-

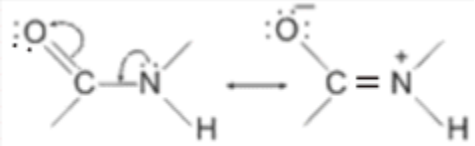
Normal haemoglobin

-Val-His-Leu-Thr-Pro- Val -Glu-Lys-

Sickle cell haemoglobin

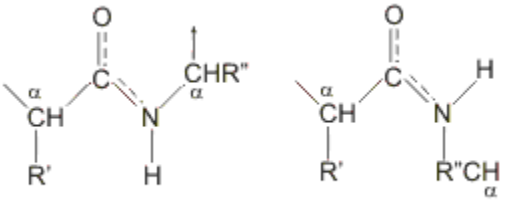
Configuration and conformation of the peptide bond in polypeptides Before discussing the secondary structure of proteins, let us first examine the nature of the peptide

bond. The lone pair of electrons on the N-atom in the peptide bond is delocalized over the C = O group. As a result, carbon-nitrogen bond acquires some double bond character.



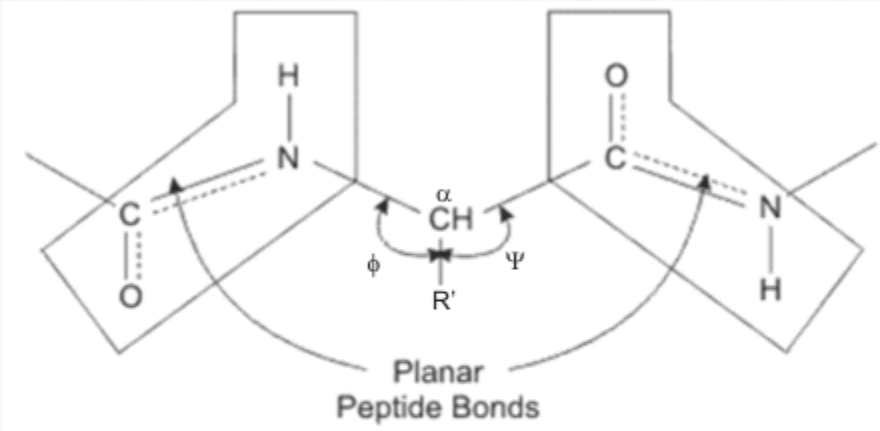
Resonance structure of peptide bond

In other words, the rotation about the C – N bond is hindered and as a result of this hindered rotation, the peptide bond can show geometrical isomerism. Further because of much larger steric repulsions between R' and R'' groups, in the cisisomer, the trans isomer is more stable. Thus, the atoms forming the peptide bond, i.e., CONH group lie in a plane with the O and H atoms in trans-orientation.



trans (more stable) cis (less stable)

Free rotation of a peptide chain can occur only around the bonds joining the nearly planar amide groups to the α-carbons. To describe the conformation of the protein molecule or the polypeptide chain in terms of the angle Φ between R' – CH – NH bonds and the angle Ψ between R' – CH – CO bonds. These angles are called Ramachandran angles after the name of the Indian Biophyscist, G.N.A. Ramachandran.

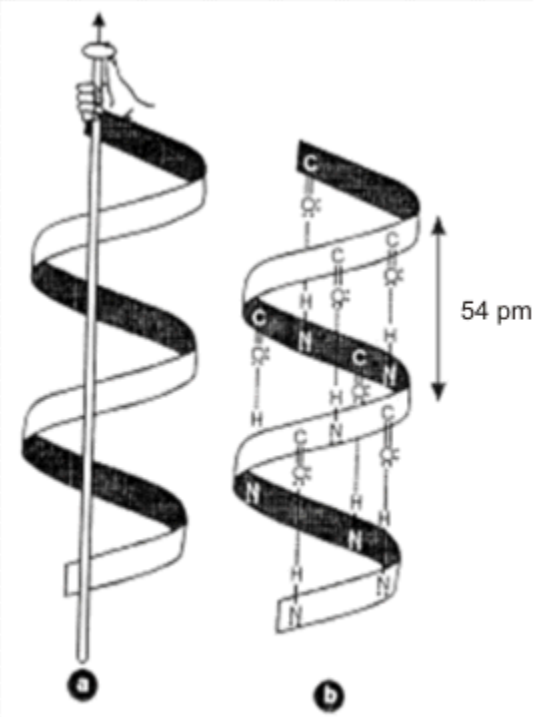


Planar Peptide Bonds

Model of a segment of a polypeptide showing planar peptide bonds in boxes and Ramachandran angles of rotation (Φ and Ψ) around α-carbon.

(ii) Secondary structure The conformation which the polypeptide chains assume as a result of hydrogen-bonding is called the secondary structure of the protein. Depending upon the size of the R groups, the following two different secondary structures are possible.

(a) α-Helix structure The α-helix model was postulated by Linus Pauling in 1951 purely on the oretical grounds and was later on verified experimentally. If the size of the R groups is quite large, the H-bonds (intramolecular) are formed between the C = O of one amino acid residue and the N – H of the fourth amino acid residue in the chain. This causes the polypeptide chain to coil up into a spiral structure called right handed α-helix* structure.(Figs a & b)



(a) A right handed α-Helix
(b) Hydrogen bonding in α-Helix

The α-helix is also known as 3.6_{13} helix since each turn of the helix has approximately 3.6 amino acids and a 13-membered ring is formed by H-bonding. The pitch or the distance between two successive turns = 54 pm.

(b) β-Pleated sheet structure or simply β-structure This structure was also proposed by Linus Pauling in 1951. In this conformation, the polypeptide chains lie side by side in a zig-zag manner with alternate R groups on the same side situated at fixed distance apart. The two such neighbouring polypeptide chains are held together by intermolecular H-bonds. A number of such chains can be interbounded to form a sheet. These sheets are then stacked one above the other like the pages of this book to form a three-dimensional structure. This structure resembles pleated folds of drapery and hence is also called β-pleated sheet structure.

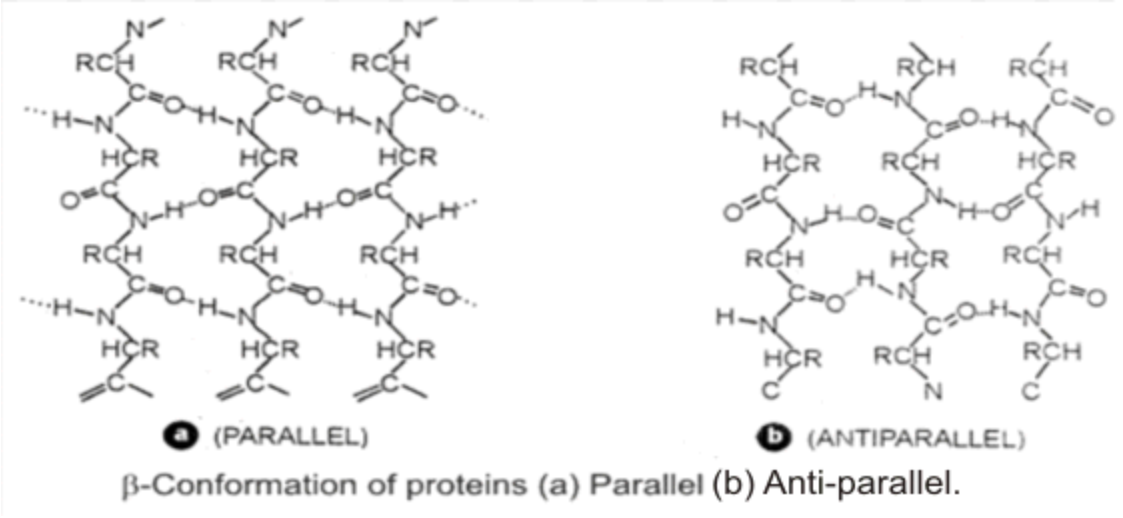
Two types of pleated sheets are possible, i.e., parallel and antiparallel. In the parallel conformation, all the polypeptide chains run in the same direction. In other words, N-termini are aligned head to head, i.e., on the same side. In the anti-parallel conformation, the polypeptide chains run alternately in the opposite directions, i.e., N-termini are aligned head to tail or N-terminus of one chain and C-terminus of another chain are on the same side. Keratin the protein present in hair has parallel β-pleated sheet structure while the silk protein fibroin has antiparallel β-pleated sheet structure.

It is important to remember that a protein may or may not have the same secondary structure throughout its length (i.e., secondary structure is a local structure). Some parts may have α-helix structure, while others may have β-pleated sheet structure. Some parts of the chain may even have no secondary structure at all. Such a structureless part is called a **random coil.**

(iii) Teritary structure The tertiary structure of proteins represents overall folding of the polypeptide chains, i.e., further folding of the secondary structure. In other words, tertiary structure refers to the manner in which the entire protein molecule folds up in the three-dimensional space to produce a specific shape (or compact form). At normal pH and temperature, each protein will take up a shape (tertiary structure) that is energetically most stable. This shape is specific to a given amino acid sequence and is called the native shape of the protein. In other words, primary structure of a protein dictates its tertiary structure.

Two major molecular shapes, i.e., tertiary structures which are found to occur naturally are fibrous and globular.

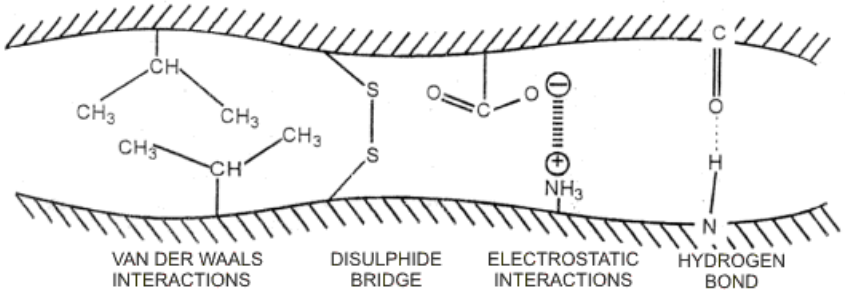
(a) Teritary structure of fibrous proteins Fibrous proteins have almost the same secondary (α-helix or β-pleated) structure throughout the length of the protein. For example, the α-keratins, the major proteins of hair and wool have the α-helix structure. In these proteins, several α-helices are coiled about each other to form molecular ropes or rods. Thus, the tertiary structure of fibrous proteins is rope-like or rod-like. The structure of collagen triple helix is shown in Fig.



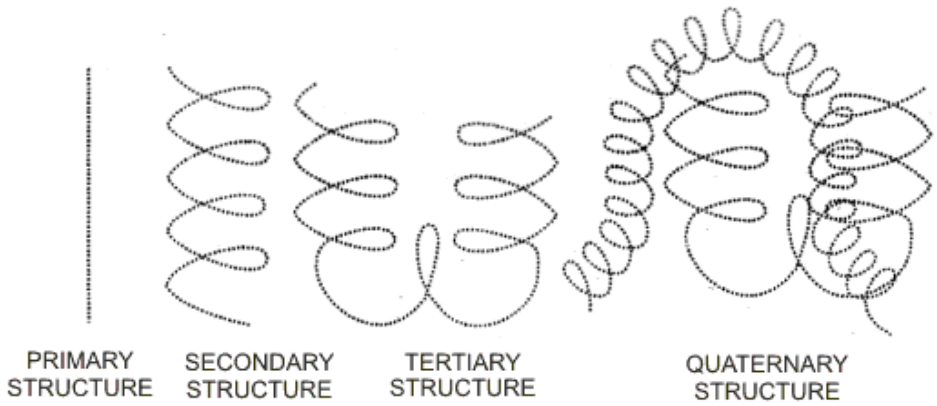
β-Conformation of proteins (a) Parallel (b) Anti-parallel.

Structure of collagen triple helix.

(b) Tertiary structures of globular proteins Unlike fibrous proteins, globular proteins do not have the same secondary structure throughout the length of the molecule. Parts of the molecule may have the α-helical structure, while the other parts may have the β-pleated sheet structure and still other parts may be random coils. These different segments of the protein then fold up to give the entire molecule a spherical shape. This folding results through the interaction of the various side chain groups of the constituent amino acids. These interactions involve several types of forces such as van der Waals interaction, disulphide bridges, ionic or salt bridges and hydrogen bonding as shown in Fig.



Forces which stabilize the teritary structure.



Diagrammatic representation of four levels of protein structure
(two subunits of two types in quaternary structure)

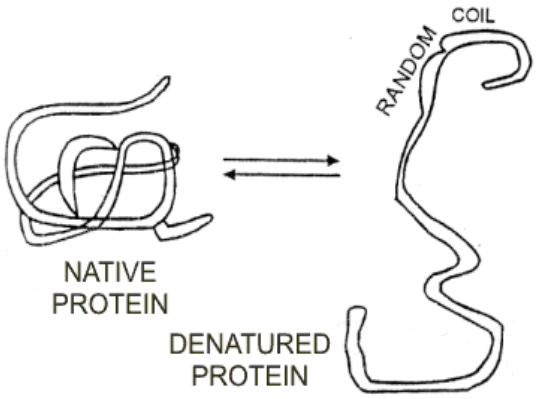
Perutz and Kendrew shared the 1962 Nobel Prize in chemistry for determining the tertiary structure of haemoglobin and myoglobin through X-ray diffraction studies.

(iv) Quaternary structure Although many proteins exist as a single polypeptide chain, there are certain proteins which exist as assemblies of two or more polypeptide chains called subunits or protomers. These subunits may be identical or different and are held together by non-covalent forces such as hydrogen bonds, electrostatic interactions and van der Waals interactions. The quatemary structure refers to the determination of the number of subunits and their spatial arrangement w.r.t. each other in an aggregate protein molecule. The best known example of a protein possessing quaternary structure is haemoglobin which transports oxygen from the lungs to the cells and carbon dioxide from the cells to the lungs through the blood stream. It is an aggregate of four polypeptide chains or subunits, two identical alpha chains (each containing 141 amino acid residue and two identical beta chains (each containing 146 amino acid residues). These four subunits lie more or less at the vertices of a regular tetrahedron. Each polypeptide chain carries a heme (iron-protoporphyrin complex) group at its end.

A diagrammatic representation of four levels of protein structures is given in Figure where each dot represents an amino acid.

14.3.10 Denaturation of Proteins

A protein found in a biological system having a unique three-dimensional structure and specific biological activity is called a **native protein**. When a protein in its native form, is subjected to a physical change like change in temperature, or a chemical change, like change in PH, the native conformation of the molecule is disrupted. As a result, globules unfold and helices get uncoiled. In other words, globular proteins (which are soluble in water) undergo coagulation or precipitation to give fibrous proteins (which are insoluble in water). Due to coagulation, the native shape of the protein is destroyed and biological activity is lost. That is why coagulated proteins so formed are called **denaturated proteins.** Chemically, during denaturation, the secondary and tertiary structures are destroyed but primary structure remains intact.



Denaturation of Globular proteins

The most common example of denaturation is the coagulation of white of an egg on boiling. When the egg is boiled hard, the soluble globular protein albumin present in the egg white is denaturated resulting in the formation of insoluble fibrous protein. Another example is the curdling of milk which is caused due to formation of lactic acid (from lactose present in milk) by the bacteria present in milk.

Yet another example of denaturation is the coagulation that occurs when milk is heated with an acid (lemon juice or tartaric acid) leading to the formation of cheese. During this denaturation, the globular milk protein, lactalbumin, becomes fibrous.

The denaturation may be reversible or irreversible. The coagulation of egg white on boiling of egg protein is an example of irreversible protein denaturation.

However, it has been shown now that in some cases, the process is actually reversible. The reverse process is called renaturation. In such cases, when the temperature and pH of a denatured protein are brought back to conditions under which the native protein is stable, secondary and tertiary structures of the protein are restored.

14.3.11 Biological functions of Proteins

Proteins constitute an indispensable part of our food. These are required for building new cells and repair of worn out ones. Obviously children and expectant mothers require more proteins. Some of the important biological functions of proteins are listed below :

1. **Structural materials** One of the most important functions of proteins is that they serve as the chief structural material of animal tissues. Structural proteins are approximately one-half of the body's total protein content. Some common examples of structural proteins are : keratin in skin, hair, nails, wool, horn, feathers etc., collagen in tendons, myosin in muscles, fibroin in silk, etc.
2. **Transport agents** Certain proteins serve as transport agents or delivery vans in the body. They transport the cellular necessities such as oxygen, metals, fatty acids, hormones, etc. for example, haemoglobin present in blood, transports oxygen from lungs to all the parts of the body.
3. **Enzymes** Some proteins function as enzymes which catalyse all biological reactions.
4. **Metabolic regulators** Some globular proteins regulate metabolic processes. For example, the protein hormone insulin from pancreas maintains glucose level in the blood (60 – 100*mg/100 cm^3 of blood); the glycoprotein thyroglobulin from thyroid gland helps in the synthesis of amine hormone thyroxine; fibrinogen which is converted into the insoluble protein fibrin and thus causes clotting of blood; nucleoproteins (constituents of genes) carry genetic information from parents to the off springs.
5. **Antibodies** These are proteins which defend the body against the invasion of foreign organisms. When an infectious species (virus, bacteria or a foreign protein) enters the blood, it releases foreign substances called antigens. In response to these antigens, the body proteins produce specific proteins called antibodies. The antibodies thus produced then specifically combine with the antigens (which induced their synthesis) thereby protecting us from disease. Gamma globulins present in blood plasma are examples of antibodies.

14.4 Enzymes

An enzyme is a specialized protein produced within an organism which is capable of catalyzing a specific chemical reaction. Almost all enzymes are proteins. There are some nucleic acids that behave like enzymes. These are called riboenzymes. Since the enzyme acts as a catalyst is sometimes referred to as a **biolcatalyst**. An enzymes facilitates a biochemical reaction by providing alternate lower activation energy pathways thereby increasing the rate of the reaction. For example, activation energy for acid hydrolysis of sucrose is 6.22 kJ mol^{-1} while the activation energy is only 2.15 Kj mol^{-1} when hydrolysed by the **enzyme sucrase.** One can depictan enzyme by a line diagram. An enzyme like any protein has the secondary and the tertiary structure.

Nomenclature A common method of naming enzymes is to add the suffix ase to the name of the compound or the class of the compounds on which they work. For example, the enzyme which catalyses the hydrolysis of maltose into glucose is named as maltase.

$$\underset{\text{Maltose}}{C_{12}H_{22}O_{11}} + H_2O \xrightarrow{\text{Maltase}} 2\,\underset{\text{Glucose}}{C_6H_{12}O_6}$$

Similarly, esterase acts on esters, amylase acts on starch (amylum), protease on proteins, urease on urea, etc. Some enzymes have, however, retained their trivial names, e.g., emulsin, pepsin, trypsin, etc.

According to IUPAC system of nomenclature, enzymes are named according to the nature of the reaction they catalyse. For example, enzymes which catalyse the oxidation of one substrate with the simultaneous reduction of another substrate are called oxidoreductase enzymes. Similarly, hydrolases are used for hydrolytic reactions, transferases for transfer of functional groups, isomerases for isomerization reactions, etc.

Composition Chemically all enzymes are globular proteins. However, some enzymes are also associated with some non-protein component called the cofactor for their activity. These cofactors are of two types :

(a) Inorganic ions Such as Zn^{2+}, Mg^{2+}, Mn^{2+}, Fe^{2+}, Cu^{2+}, Co^{2+}, Mo^{3+}, K^{+}, Na^{+}, etc.

(b) Organic molecules These are also of two types :

(i) Coenzymes are small organic molecules which are only loosely held to the protein and can be easily separated by dialysis. These are usually derived from vitamins such as thiamine, riboflavin, niacin, etc.

(ii) Prosthetic groups are also organic molecules which are tightly held to the protein by covalent bonds but can be separated only by careful hydrolysis. Most of these are also derived from vitamins such as biotin.

Regardless of the nature of the cofactor, the protein – cofactor complex is called the holoenzyme while the inactive protein part left after the removal of the cofactor is called the apoenzyme. Thus,

$$\text{Holoenzyme} \rightarrow \underset{\text{(Inactive)}}{\text{Apoenzyme}} + \text{Cofactor}$$

Every biological reaction requires a different kind of enzyme. Since there is a large number of such biological reactions, therefore, there is a large number of enzymes functioning in a living system. A typical cell, on the average, contains about 3000 different kinds of enzymes, each catalyzing a different reaction.

14.4.1 How do enzymes bring about such High Rates of Chemical Conversions?

Enzymes are needed only in small quantities for the progress of a reaction. Similar to the action of chemical catalysts, enzymes are said to reduce the magnitude of activation energy. For

example, activation energy for acid hydrolysis of sucrose is 6.22 kJ mol-1, while the activation energy is only 2.15 kJ mol-1 when hydrolysed by the enzyme, sucrose.

The chemical which is converted into the product is called a "substrate".Hence enzymes i.e., proteins with three dimentional structures including an "active site", convert a substate(S) into a product(P).symbolically, this can be depicted as:

$$S \rightarrow P$$

It is now understood that the substrate"S' has to bind the enzyme at its active site' within a given cleft or pocket. The substrate has to diffuse towards the active site .there is thus, an obligatory formation of an 'ES' complex.E stands for enzyme.This complex formation is a transient phenomenon. During the state where substrate is bound to the enzyme active site, a new structure of the substrate called transition state structure is formed.Very soon, after the expected bond breaking/making is completed, the product is released from the active site. In other words, the structure of substratye gets transformed into the struicture of product(s).The pathway of this transformation must go through the so called transition state structure. There could be many more 'altered structural states' between the stable substrate and the product. Implicit in this statement is the fact that all other intermediate structural states are unstable. Stability is something related to energy status of the molecule or the structure. Hence, when we look at this pictorially through a graph it looks like something as in fig

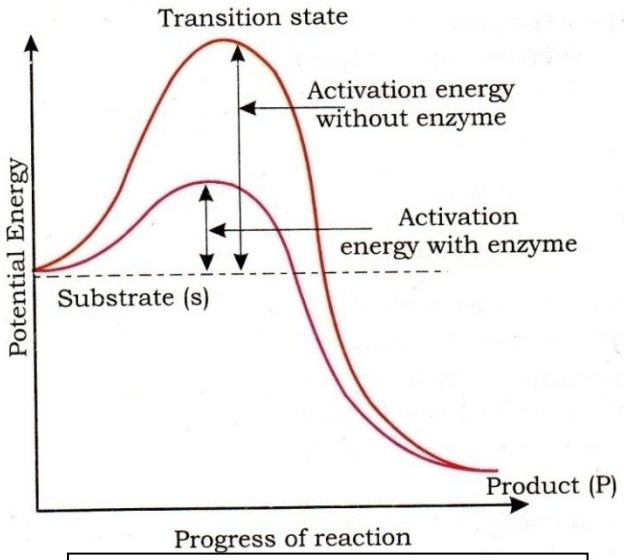


Fig: Concept of activation energy

The y-axis represents the potential energy content. The x-axis represents the progression of the structural transformation or states through the 'transition state'. You would notice two things. The energy level difference between S and P. If 'P' is at a lower level than 'S', the reaction is an exothermic reaction. One need not supply energy (by heating)in order to form the product. However, whether it is an exothermic or spontaneous reaction or an endothermic or energy requiring reaction, the 'S' has to go through a much higher energy state or transition state. The difference in average energy content of 'S' from that of this transition state is called 'activation energy'.

Enzymes eventually bring down this energy barrier making the transition of 'S' to 'P' more easy.

14.4.2 Applications of enzymes:

Enzymes are widely used :

(i) In the manufacture of beer and wine by the fermentation of carbohydrates.

(ii) In the manufacture of sweet syrup from corn starch.

(iii) In the production of cheese by coagulation of milk.

14.4.3 Disease Caused by Enzyme Deficiency There are certain diseases which are caused by enzyme deficiency. For example, the congenital disease phenylketone urea (PKU) is caused by the deficiency of the enzyme called phenylalanine hydroxylase required for converting phenylalanine to tyrosine. Due to deficiency of this enzyme, other enzymes in the cell convert phenylalanine to phenyl pyruvate. The accumulation of phenylalanine and phenylpyruvate in the blood can lead to severe brain damage and mental retardation. If PKU is detected in new born babies, a diet is prescribed which is low in phenylalanine content. As a result, build up of phenylpyruvate is prevented which ensures normal growth and development.

Another disease which is caused by the deficiency of the enzyme tyrosinase is called albinism. Due to deficiency of tyrosinase enough melanin (the pigment which gives colour to the skin or hair) is not produced and hence the persons and animals suffering from this disease have white skin or hair.

Enzymes have also been used to treat heart disease. Many heart attacks occur due to blood clot formation in a coronary artery. Recently, it has been possible to check heart attacks by using the enzyme streptokinase which dissolves the blood clot.

14.5 Vitamins

Definition Besides oxygen, water, fats, carbohydrates, proteins and certain inorganic salts, a number of organic compounds are needed for the normal metabolic processes and for the life, growth and health of human beings and animals. These compounds, though required in very small amounts, are vital to life processes. Therefore, they are called vitamins* (Latin, vita = life). Hence,

vitamins may be defined as a group of biomolecules other than fats, carbohydrates and proteins) which are required in small amounts for normal metabolic processes and for the life, growth and health of human beings and animal organisms.

Importance Vitamins neither supply energy nor help in building tissues of the cells. Nevertheless, they play an important role in keeping good health of human beings and animals. Their deficiency causes serious disturbances diseases in the body.

Most of the vitamins cannot be synthesized by our body. Therefore, these must be supplied in the food. Plants can synthesize all vitamins but only a few are synthesized in animals. Vitamin D, however, may either be supplied in the food or may be produced in the skin by the irradiation of ergosterol (a sterol present in our body) with ultraviolet light.

Human body can synthesize vitamin A from carotenes, some members of vitamin B complex and vitamin K are synthesized by microorganisms present in the intestinal tract.

Sources The main sources of vitamins is our food which consists of milk, butter, cheese, fruits, green vegetables, meat, fish, eggs, etc. These have also been synthesized in the laboratory and are available in the market either as tablets and capsules which can be taken orally or as injections which may be given in case of acute deficiency.

14.5.1 Classification of Vitamins

Vitamins are complex organic molecules. From the view point of chemical structures, there is very little common to the various vitamins. Therefore, for simplicity, these are usually designated by alphabet letters such as A, B, C, D, E and K. Some of these are named as sub groups, e.g., B_1, B_2, B_6, B_{12}, etc. About 25 vitamins are known to-date. These have been broadly classified into the following two categories :

(i) Water soluble vitamins These include vitamin B-complex (B_1, B_2, B_5 i.e. nicotinic acid, B_6, B_{12}, pantothenic acid and folic acid) and vitamin C.

Water soluble vitamins must be supplied regularly in diet because they are regularly excreted in urine and cannot be stored (except vitamin B_{12}) in our body.

(ii) Fat soluble vitamins These are oily substances not readily soluble in water. However, they are soluble in fat. These include vitamins A, D, E and K. They are stored in liver and adipose (fat storing) tissues. Excess intake of these vitamins (A and D) is harmful and may cause hypervitaminoses.

Biotin (vitamin H) is, however, an exception since it is neither soluble in water nor in fat.

Lack of a particular vitamin causes a specific deficiency disease. Multiple deficiencies caused by lack of more than one vitamin is called avitaminoses. It is quite common in human beings.

Some important vitamins, their characteristics, sources and deficiency disease are summarised in the following table.

	Vitamin	**Characteristics**	**Sources**	**Deficiency disease**
1.	**Vitamin A** (Retinol or bright eye vitamin). Carotenoids (present in carrots) are precursors of vitamin A.	Soluble in oils and fats, insoluble in water, stable to heat. Promotes growth and vision in animals and increases resistance to diseases.	Milk, butter, eggs, fish liver oil, rice, polishing, liver, kidney, green vegetables (cabbage), carrots, tomatoes, mangoes, sweet potatoes, etc.	Xerophthalmia (hardening of cornea of eye), night blindness and xerosis (drying of skin).
2.	**Vitamin B_1** (Thiamine, aneurin or antineuritic vitamin)	Soluble in water, insoluble in oils and fats, destroyed by heat above 313 K.	Pulses, nuts, whole cereals (rice, wheat, etc.), rice polishings, yeast, egg yolk, milk, green vegetables and fruits.	Beriberi (paralysis of legs and general weakness) and loss of appetite.
3.	**Vitamin B_2** (Riboflavin or Lactoflavin)	Soluble in water, insoluble in oils and fats, sensitive to light but stable to heat. Essential for growth and health of animals.	Milk, yeast, green vegetables, meat, liver, kidney, egg white, etc. Daily dosage is 2-3 mg.	Retards growth, causes general inflammation of tongue i.e. dark red tongue (glossitis) dermatitis and cheilosis (cracking or fissuring at corners of mouth and lips).
4.	**Vitamin B_6** (Adermin or pyridoxine). In	Soluble in water, insoluble in oils and fats.	Rice bran, yeast, molasses, meat, fish, egg	Causes specific dermatitis in rats (acrodynia),

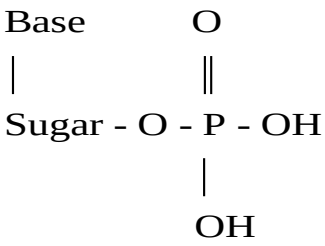
	fact, vitamin B_6 is a mixture of three substances called pyridoxine, pyridoxal and pyridoxamine.		yolk, whole cereals (wheat, gram), etc.	causes pellagra (shrivelled skin) and anaemia (bloodlessness) in man, affects central nervous system, causes general weakness, convulsions, nervousness, insomnia and irritability.
	Vitamin H (Biotin)	Neither soluble in water not in fat.	Yeast, liver, kidney and milk	Dermatitis, (skin disease) loss of hair and paralysis
5.	**Vitamin B_{12}** (Cyanocobalamin) It contains cobalt.	Water soluble, stable to heat.	Milk, eggs and liver of ox, sheep, pig, fish, etc.	Pernicious anaemia (RBC deficient in haemoglobin), inflammation of tongue and mouth.
6.	**Vitamin B-Complex** It is a group of vitamins which contains vitamins B_1, B_2, B_6, B_{12}, biotin, folic acid, pantothenic acid and nicotinic acid.	Required to release energy from food and to promote healthy skin and muscles.	Occur together in foods such as milk, liver and cereals.	Beriberi (Vitamin B_1 and pernicious anaemia (Vitamin B_{12}).
7.	**Vitamin C** (Ascorbic acid or Antiscorbutic vitamin)	Soluble in water, destroyed by cooking and prolonged exposure to air. To avoid the loss, vegetables rich in vitamin C must be cooked in closed pans and pressure cookers. Vitamin C increases	Citrous fruits, e.g., oranges, lemons, amla, tomatoes, green leafy vegetables (cabbage), chilies, sprouted pulses and germinated grains. Daily dosage is about 75 mg.	Causes scurvy (bleeding of gums), pyorrhea (loosening and bleeding of teeth).

		resistance of the body towards diseases. Maintains healthy skin and helps cuts and abrasions to heal properly.		
8.	**Vitamin D** (Ergocalciferol or antirachitic vitamin or sunshine vitamin)	Soluble in oils and fats, stable to heat and resistant to oxidation. It controls calcium and phosphorus metabolism.	Fish liver oils, butter, milk, eggs, liver and meat. Daily dosage of vitamin D is about 0.025 mg.	Rickets (bending of bones) in children, and osteomalacia (soft bones and joint pain) in adults.
9.	**Vitamin E** (Antisterility vitamin). It is a mixture of four vitamins called α, β, γ and δ-tocopherols. Amongst these α-tocopherols is the most active.	Soluble in oils and fats, stable to heat and oxidation.	Vegetables oils like wheat germ oil, cotton seed oil, soyabean oil, peanut oil, sunflower oil, etc., eggs, milk and fish, etc. Daily dosage of vitamin E is about 5 mg.	Sterility, i.e., loss of sexual power of reproduction. Increased fragility of RBCs and muscular weakness.
10.	**Vitamin K** (Phylloquinone or antihaemorrhagic vitamin) is a mixture of two vitamins called K_1 and K_2.	Soluble in oils and fats, can withstand cooking, sensitive to light and alkali.	Vitamin K_1; alfalfa, leafy vegetables like cabbage and spinach, and carrot tops. Vitamins K_2 and occurs mainly in bacteria.	Haemorrhage, lengthens the time of blood clotting.
11.	**Coenzyme** Q_{10}	–	Chloroplasts of green plants and mitochondria of animals.	Low order of immunity of body against many diseases.

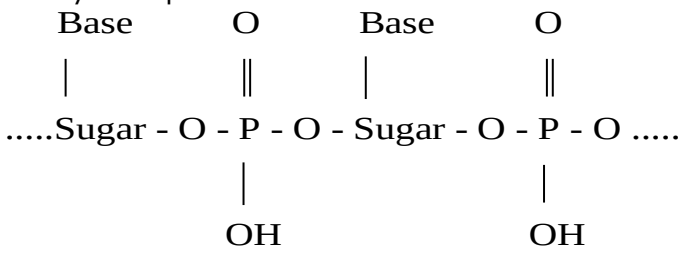
14.6 Nucleic acids

Nucieic Acids: The acid insoluble fraction of any living tissue is the nucleic acid. Nucleic acids are biopolymers (i.e. polymers present in the living system). They are also called polynucleotides since the repeating structural unit (monomeric unit) of nucleic acids is a nucleotide.Together with polysaccharides any polypeptides these comprise the true

macromolecular fraction of any living tissue or cell. Each nucleotide is made up of three parts, i.e., a sugar molecule, a heterocyclic nitrogenous base and phosphoric acid.
Thus, in general, a nucleotide may be represented as :



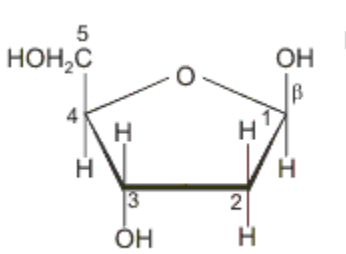
And hence a nucleic acid may be represented as :

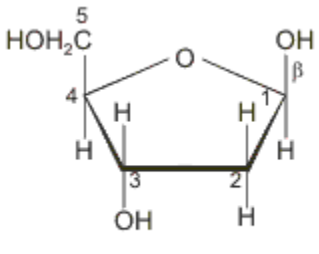


14.6.1 Chemical composition of Nucleic Acids

Complete hydrolysis of DNA or RNA gives a mixture of three different compounds:
(i) A pentose sugar (ii) nitrogen containing heterocyclic compounds also called nitrogenous base and (iii) phosphoric acid.

Sugars Two pentose sugars have been isolated. Whereas DNA contains, β-D-2 deoxyribose, RNA contains β-D-ribose. Both these sugars are found in the furanose form as shown below :

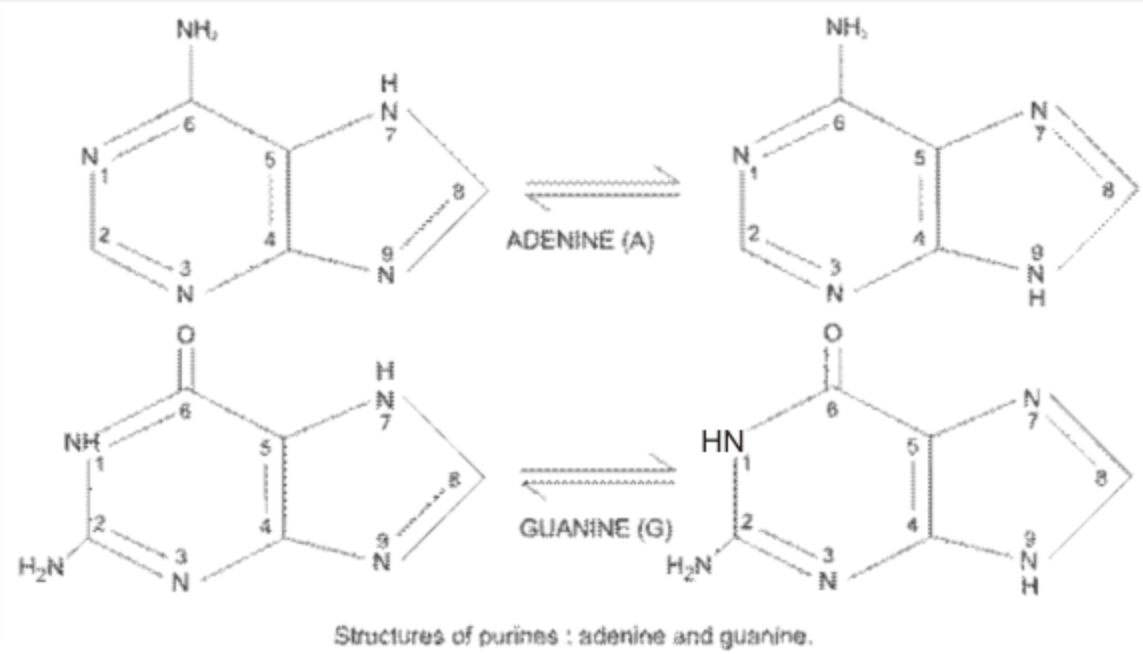




Bases There are two different classes of heterocyclic nitrogenous bases which have been isolated by the hydrolysis of nucleic acids. These are :

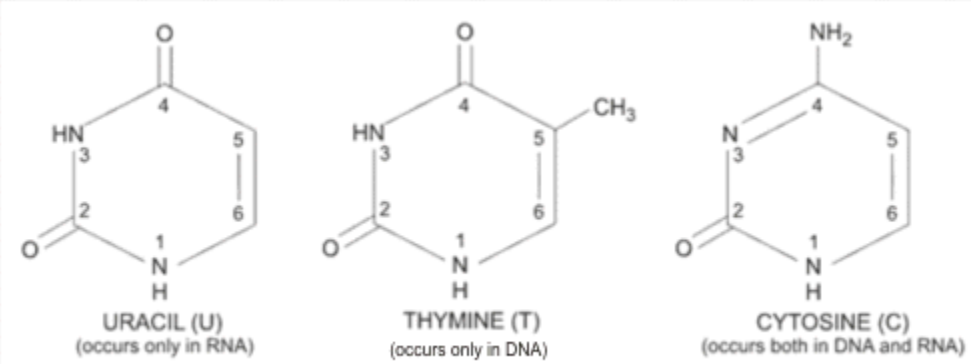
(i) Purines
(i)Purines (ii)Pyrimidines

(i) Purines Two purines which are most commonly found in nucleic acids are adenine (A) and guanine (G).



Structures of purines : adenine and guanine.

(ii) Pyrimidines Three most commonly occurring pyrimidines in nucleic acids are : Uracil,(U), thymine (T) and cytosine(C).



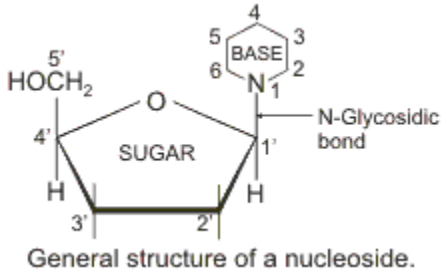
Structures of pyrimidines : uracil, thymine and cytosine.

14.6.2 Nucleosides and Nucleotides

Nucleosides A nucleoside contains only two basic components of nucleic acids, i.e., a pentose sugar and a nitrogenous base. During their formation, 1-position of the pyrimidine or 9-position of the purine moiety is linked to C_1' of the sugar (ribose or deoxyribose) by a β-linkage. Thus, in general, nucleosides may be represented as :

Sugar – Base

In terms of structural formula, a nucleoside may be represented as shown in Fig.



General structure of a nucleoside.

The carbon atoms of the sugar are numbered as 1′, 2′, 3′, 4′, 5′, etc. in order to distinguish them from the numbering of the atoms of the bases.
Depending upon the type of sugar present, nucleosides are of two types :

(i) Ribonucleosides and **(ii)** Deoxyribonucleosides
The names of the ribonucleosides of the five bases are :

Base	Abbreviation	Nucleoside
Adenine	A	Adenosine
Guanine	G	Guanosine
Cytosine	C	Cytidine
Thymine	T	Thymidine
Uracil	U	Uridine

The names of the deoxyribonucleosides of the five bases are obtained by adding the prefix deoxy to the name of the corresponding ribonucleoside.For example, deoxyadenosine, deoxyguanosine, deoxycytidine, deoxythymidine and deoxyuridine.

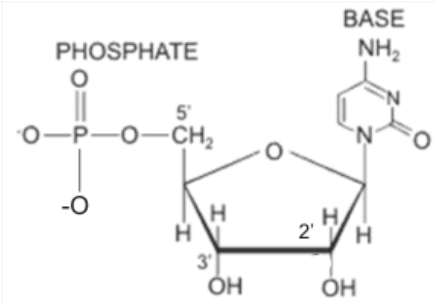
Nucleotides: A nucleotide contains all the three basic components of nucleic acids, i.e., a phosphoric acid group, a pentose sugar and a nitrogenous base. These are formed by esterification of C_5′ – OH of the sugar of the nucleoside with phosphoric acid. In other words, nucleotides are nucleoside monophosphates.
Thus, in general, nucleotides may be represented as :

Base

|

Sugar - Phosphate

In terms of structural formula, a nucleotide may be represented as shown in Fig.



Depending upon the type of the sugar present, nucleotides like nucleosides are of two types :
(i) Ribonucleotides and
(ii) Deoxyribonucleotides
The nucleotides are abbreviated by three capital letters, preceded by d-in case of deoxy series. For example,

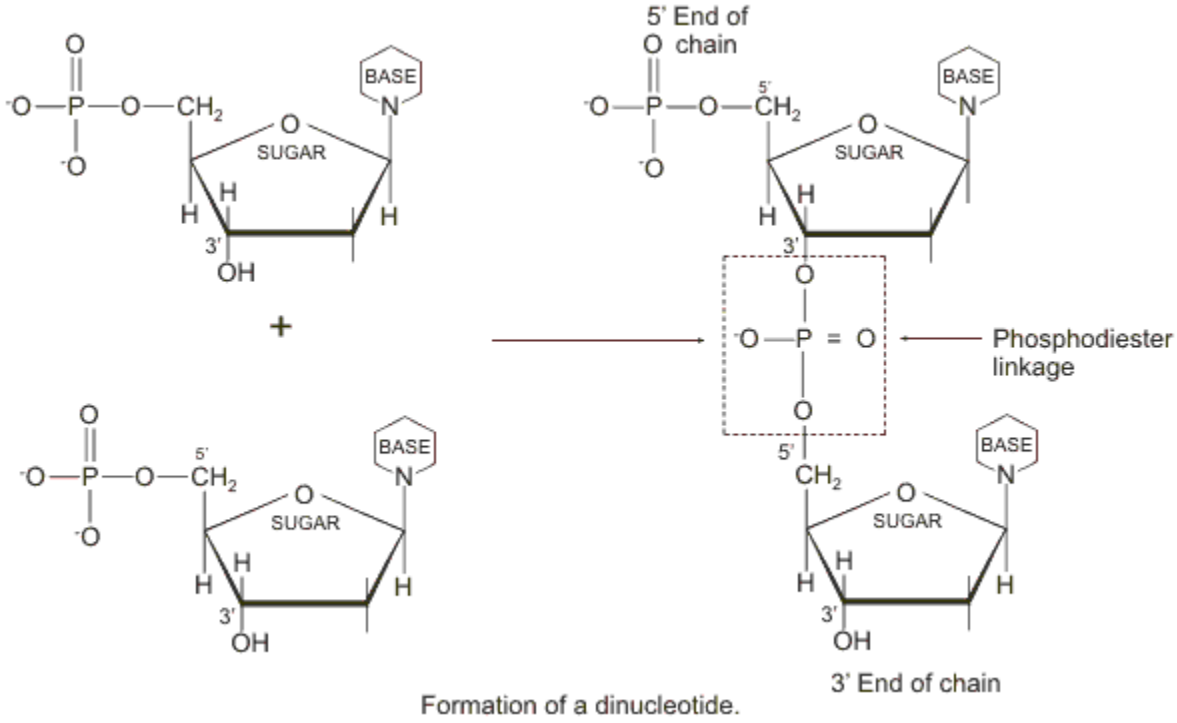
AMP = Adenosine monophosphate
dAMP = Deoxyadenosine monophosphate
UDP = Uridine diphosphate
ATP = Adenosine triphosphate

14.6.3 Structure of Nucleic Acids

The structure of nucleic acids is discussed at the following two levels:

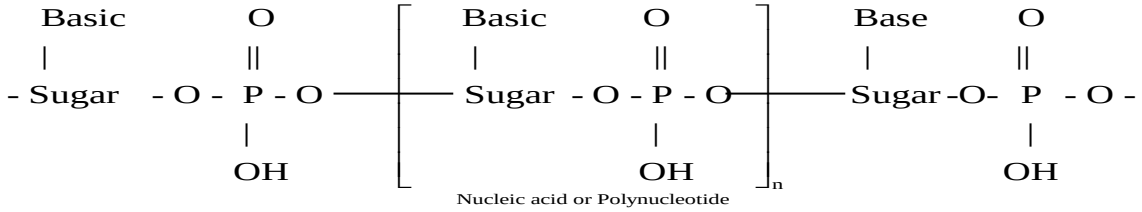
(i) Primary structure
(ii) Secondary structure

Primary structure Nucleic acids are polynucleotides, i.e., they are formed by the combination of thousands of molecules of nucleotides Di-, tri-, oligo- and polynucleotides are obatained by joining together two, three, few(upto 10) and thousands of nucleotides respectively through phosphodiester linkages between 5' and 3' carbon atoms of the pentose sugar.For example, formation of a typical dinucleotide is illustrated in fig



Formation of a dinucleotide.

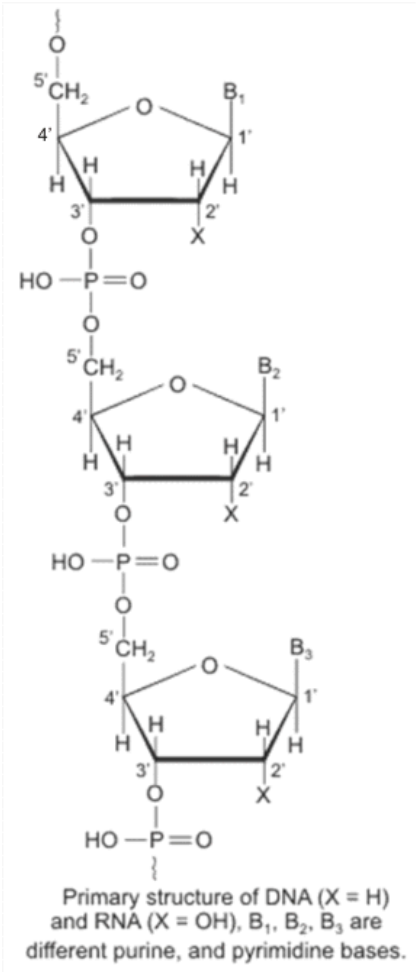
If a large number of nucleotides are connected in the manner as discussed above, we get a nucleic acid. A simplified version of nucleic acid chain may be represented as follows :



Nucleic acid or Polynucleotide

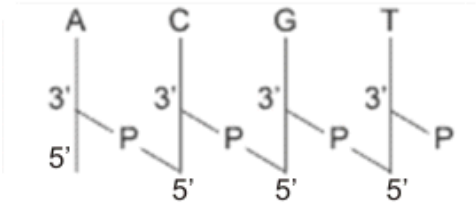
with the elimination of water molecules. During this polymerization, the $^{5}CH_2$ – OH hydroxy group of the sugar residue of one nucleotide combines with one of the - OH groups of the

phosphoric acid group present at C_3' of the other nucleotide to form a long polynucleotide chain. Thus, the backbone of nucleic acids consists of alternating sugar-phosphate residues; each sugar on this backbone is further connected to one of the four nitrogenous bases. Thus, the sequence in which the four nitrogen bases are attached to the sugar-phosphate backbone of a nucleotide chain is called the primary structure of nucleic acids. This is shown in Fig.



Primary structure of DNA (X = H) and RNA (X = OH), B_1, B_2, B_3 are different purine, and pyrimidine bases.

From the representation shown in Fig. it is evident that it is extremely time consuming to write the complete primary structures of nucleic acid chains or even oligonucleotides. In order to simplify, the bases are represented by their one letter symbols (A, C, G, T, U) the phosphate bond is represented by the symbol 'P' and the sugar is drawn by simple Fischer projection. The complete oligonucleotide chain or nucleic acid chain is then written with the 5' end at the bottom. For example, a tetranucleotide having adenine, cytosine, guanine and thymine bases from 5' end to 3' end is represented as ACGT as shown in Fig.



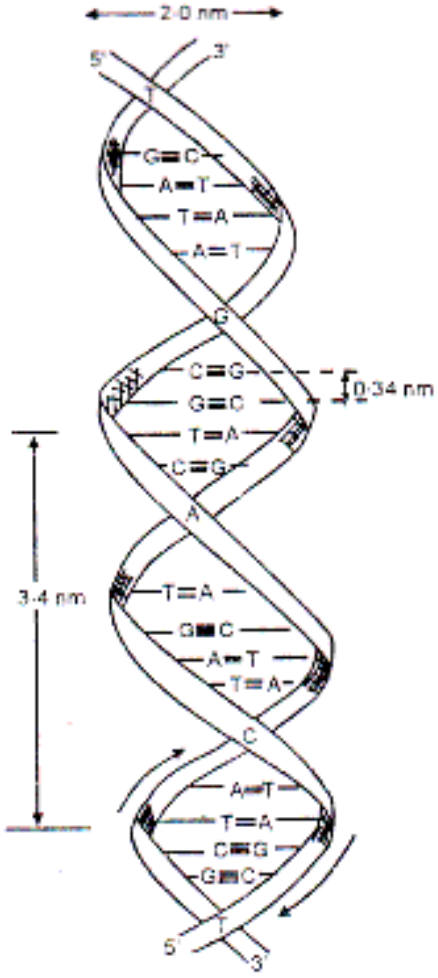
Representation of a tetranucleotide ACGT.

(i) Secondary structure Before discussing the secondary structure, we must know Chargaff rule.

Chargaff rule It was discovered by **E. Chargaff** that the base composition in DNA varied from one species to the other but in all cases, the amount (in moles) of adenine was equal to that of thymine (A = T) and that of cytosine was equal to that of guanine (C = G). In other words, the total amount of pyrimidines (A + G = C + T). However, the AT/CG ratio differed considerably between different species. For example, the ratio is 1.52 in man while in E. Coli, it is 0.93.

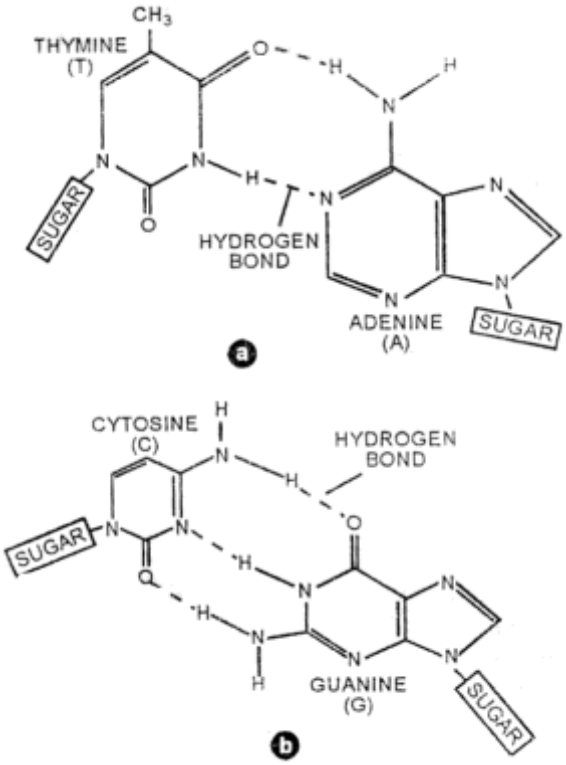
In 1953, **J.D. Watson, F.H.C. Crick** and **M.Wilkins** from their X-ray studies proposed a double helical structure for DNA. For this discovery, they shared the 1962 Nobel Prize in Physiology and Medicine. This model not only explained the base equivalence (A = T; C = G) but other important properties such as replication of DNA.

According to this model, DNA consists of two right handed polynucleotide strands (or chains). These strands run in opposite directions (i.e., the free phosphate residues at 3' (or 5') positions of the two strands lie on the opposite sides of the α-helix) giving a double helix structure to DNA. The backbone of each strand consists of sugar-phosphate units and the base units of each strand are pointed into the interior of the helix. The base pairs of the two strands are



The double α-helix structure of DNA.The two nucleotide strands are held together by specific hydrogen bonding between bases.The numerals 5′,3′indicate that the free hydroxyl groups of the terminal deoxyribose units are present at 5′ and 3′ positions

linked together through hydrogen bonds. A critical point of Watson – Crick model is that hydrogen bonding can occur between specific bases. A purine base of one strand is always paired with a pyrimidine base of the other strand. Because of sizes and geometries of the bases, the only possible pairings in DNA are between G (guanine) and C (cytosine) through three H-bonds (i.e. C ≡ G) and between A (adenine) and T (thymine) through two H-bonds (i.e. A = T) as shown in Fig.



Pairing of complementary bases,(a)Thymine(T) with Adenine(A) through two H-bonds and (b) Cytosine(C) and Guanine(G)through three H-bonds

The two strands of the double helix are complementary and not identical since the base sequence of one strand automatically fixes that of the other due to the above base-pairing principle. The paired bases are stacked together one above or below the other and the distance between any two successive turns of the helix is 3.4 nm suggesting that there are ten base pairs in each turn. The diameter of the helix is about 2.0 nm.

On heating, the two strands of DNA separate from each other. This is called **melting**. However, on cooling these again hybridize. This is called **annealing**. *The temperature at which the two strands separate completely is called its melting temperature (Tm) which is specific of each specific sequence.*

Unlike DNA, RNA has a single strand Besides primary and secondary structures. DNA also have higher level structures. These higher level structures deal with the way these molecules are bound to proteins, folded and super coiled to make chromatin and chromosomes. Such structures explain as how four meters of DNA can be accommodated inside a single cell.

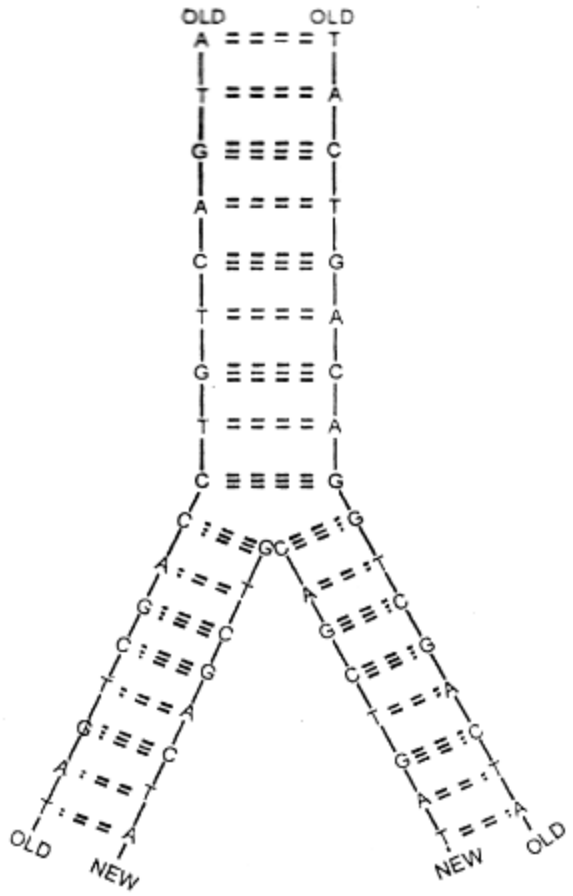
14.6.4 Biological Functions of Nucleic Acids

The two important functions of nucleic acids are (i) Replication and (ii) Protein synthesis.

(i) Replication The process by which a single DNA molecule produces two identical copies of itself is called cell division (mitosis) or replication.

As stated above, DNA is the chemical basis of heredity and hence may be regarded as reserve of genetic information. It is exclusively responsible for maintaining the identity of different species of organism over millions of years. A DNA molecule is capable of replication (self duplication) during cell division and the two identical DNA strands are transferred to the daughter cells.

Replication of DNA is an enzyme-catalysed process. It begins by partial unwinding of the double helix through breaking of the H-bonds between pairs of bases leading to the formation of two isolated strands.



Schematic representation of replication of DNA; for simplicity only base sequences are shown

Each strand then acts as the template (or pattern) for the synthesis of two new strands. Because of the *base-pairing principle*, each new strand is complementary to its old template strand. As a result, two identical copies of DNA from the parent DNA are produced. Each copy is then passed on to the two new cells resulting from cell division. In this way, hereditary effects are transmitted from one cell to the other.

It may be noted that DNA replication is **semiconservative.**

(ii) Synthesis of proteins Another important function of DNA is the synthesis of proteins.+

14.6.5 Points of difference in DNA and RNA

The main points of difference between DNA and RNA are given in the following table.

Points of difference between DNA and RNA

Deoxyribonucleic acid (DNA)	Ribonucleic acid (RNA)
1. The sugar present in DNA is 2-deoxy D-(-)- ribose.	1. The sugar present in RNA is D-(-)-ribose.
2. DNA contains cytosine and thymine as pyrimidine bases and, guanine and adenine as purine bases.	2. RNA contains cytosine and uracil as pyrimidine bases and, guanine and adenine as purine bases.
3. DNA has double stranded α-helix structure.	3. RNA has single stranded α-helix structure.
4. DNA chiefly occurs in the nucleus of the cell.	4. RNA mainly occurs in the cytoplasm of the cell.
5. DNA molecules are very large; their molecular mass may vary from $6 \times 10^6 - 16 \times 10^6$ u.	5. RNA molecules are much smaller with molecular mass ranging from 20,000 to40,000 u.
6. DNA has the unique property of replication.	6. RNA usually does not replicate.
7. DNA controls the transmission of hereditary effects.	7. RNA controls the synthesis of proteins.

Some points requiring

Special attention

1. Monosaccharides which differ in configuration at C_1 in aldoses and C_2 in ketoses are called **anomers**, Thus, α-D-glucose and β-D-glucose are anomers and so are α-D-fructose and β-D-fructose.
2. The C_1 carbon atom in aldoses and C_2 carbon atom in ketoses around which the configuration of anomers differs is called the **anomeric** or the **glycosidic carbon**.
3. Monosaccharides which differ in configuration at a carbon atom other than the anomeric carbon are called **epimers.** Thus, glucose and mannose which differ in configuration at C_2 are called **C_2-epimers** while glucose and galactose which differ in configuration at C_4 are called **C_4-epimers.**
4. All monosaccharides (aldoses and ketoses) and disaccharides except sucrose reduce Fehling's solution, Benedict's solution and Tollen's reagent and hence are called **reducing sugars.** Others (sucrose, starch, cellulose) which do not reduce these reagents are called **non-reducing sugars.**
5. Although starch and cellulose both contain an aldehydic group at the end of the chain but still they do not reduce Tollen's reagent. The reason being that due to high molecular weights of these polysaccharides their reducing properties are masked and hence no detectable reducing properties are observed.
6. The spontaneous change of specific rotation with time to an equilibrium value is called **mutarotation**. All reducing carbohydrates i.e. monosaccharides (glucose, fructose, mannose etc.) and disaccharides (maltose, lactose etc.) undergo mutarotation in aqueous solutions.
7. Since glucose (grape sugar) is dextrorotatory, it is also called dextrose. Similarly, fructose being laevorotatory is also called laevulose.
8. α-Amino acids are the building blocks of proteins or proteins are the condensation polymers of α-amino acids.
9. All the α-amino acids forming proteins have L- configuration while all the naturally occurring carbohydrates have D-configuration.
10. Keratin in skin, hair, nails and wool, collagen in tendons, fibroin in silk and myosin in muscles are all **fibrous proteins** and have linear structures.
11. Enzymes, hormones (insulin, thyroglobulin), antibodies, haemoglobin, fibrinogen, alubimin etc. are all globular proteins. These have folded structures. The folding of globular proteins occurs due to (i) disulphide bridges, (ii) intramolecular H-bonding, (iii) van der Waals' interactions and (iv) dipolar interactions.
12. Insulin is a protein or peptide hormone. It consists of 51 amino acids arranged in two polypeptide chains containing 21 and 30 α-amino acid residues respectively. The two peptide chains are held together by two cystine disulphide cross-links.
13. The disease sickle cell anaemia is caused by defective haemoglobin which is obtained by replacement of just one amino acid (i.e., glutamic acid by valine) in the sequence of the protein haemoglobin.
14. **Enzymes** are biological catalysts. Chemically all enzymes are globular proteins.

15. Certain enzymes are associated with **coenzymes** mostly derived from vitamins for their biological activity.

Competition window

1. Both glucose and fructose reduce Tollen's reagent, Fehling's solution, Benedict's solution etc.

2. All monosaccharides and reducing disaccharides (maltose, lactose etc.) react with three molecules of $C_6H_5NHNH_2$ to form crystalline osazones which are used for their identification and characterization. Further all monosaccharides which differ in configuration at C_1 and C_2 give the same osazone, i.e., glucose and fructose.

3. Both starch and cellulose are condensation polymers of glucose. Whereas glucose is a polymer of α-glucose, cellulose is a polymer of β-glucose.

4. Starch is a mixture of two components, i.e. amylase and amylopectin. Whereas amylase is a linear polymer, amylopectin has branched chain structure.

5. Amylose gives a blue colour with iodine solution due to the formation of an inclusion complex.

6. Sucrose is the only non-reducing disaccharide.

7. In lactose, glucose unit is in the reducing form.

8. Proteins are the condensation polymers of α-amino acids.

9. –CO-NH- is called the peptide bond.

10. At isoelectric point, the amino acids primarily exist as the neutral dipolar or zwitterions and hence have the minimum solubility.

11. In amino acids, $-COO^-$ group acts as the **base** while $-\overset{+}{N}H_3$ acts as the acid.

12. The biological activity of proteins is due to their secondary and tertiary structures.

13. During denaturation, the primary structure of proteins remains intact while secondary and tertiary structures are destroyed.

14. DNA and RNA are polynucleotides.

15. A nucleotide consists of three components, i.e., base, sugar, phosphate and has the structure, base-sugar-phosphate. A nucleoside has only two components and has the structure, base-sugar.

16. Both DNA and RNA contain the same tow purine bases, i.e., adenine (A) and guanine (G), Their pyrimidine bases are, however, different whereas DNA contains cytosine (C) and thymine (T), RNA contains cytosine (C) and uracil (U).

17. DNA has double standard α-helix structure in which the two complementary stands are held together by H-bonds between the bases A and T and C and G. RNA, on the other hand, has only one strand.

18. Oils and fats consist of triglycerides which on hydrolysis give glycerol and saturated and unsaturated fatty acids. The most commonly occurring saturated fatty acids are:
Myristic acid ($C_{13}H_{27}COOH$), palmitic acid ($C_{15}H_{31}COOH$) and stearic acid ($C_{17}H_{35}COOH$) while the common unsaturated fatty acids are : oleic acid ($C_{17}H_{33}COOH$), linoleic acid ($C_{17}H_{31}COOH$) and linolenic acid ($C_{17}H_{29}COOH$).

19. Waxes are esters of long chain fatty acids with long chain alcohols. For example, Bees wax is myricyl palmitate, carnauba wax is myricyl cerotate while spermaceti is cetyl palmitate.

20. Molish reagent is 1% alcoholic solution of α-naphthol. It is used to detect carbohydrates.

21. Biuret test and ninhydrin test are used for detection of amino acids and proteins.

22. Adenosine (ribose + adenine) is a **nucleoside** while adenosine monophosphate (AMP_, adenosine diphosphate and adenosine triphosphate (ATP) are all **nucleotides.**

23. ATP is called energy currency of the cells. It is a nucleotide containing adenine as the purine base, ribose as the sugar and three inter-linked phosphate groups.

24. Energy is stored in living cells in form of ATP.

25. During photosynthesis, synthesis of each molecule of glucose consumes 18 molecules of ATP.

26. Formaldehyde is formed during photosynthesis.
27. Amylases or Ptylins present in saliva partially hydrolyse starch into maltose while amylases secreted by pancreas and small intestinal walls hydrolyse it completely to glucose.
28. Human body can digest starch but no cellulose.
29. Enzymes present in liver convert galactose into glucose while fructose is either converted into glucose or used as such.
30. The PH of the gastric juice (containing strong HCl) is 2 which causes denaturation of proteins in the stomach.
31. The proteolytic enzymes, i.e, pepsin, trypsin and chymotripsin all hydrolyse proteins to α-amino acids.
32. Bile, acids present in bile juice cause emulsification of the fats while the enzymes lipases hydrolyse them to glycerol and fatty acids.
33. The end products of citric acid cycle or Kreb's cycle or thricarboxylic acid cycle (TCA), are two molecules of CO_2 and 12 molecules of ATP per molecule of acetyl coenzyme-A are generated in this process.
34. During complete oxidation of one molecule of glucose into CO_2 and H_2O, 38 molecules of ATP are generated.
35. During complete oxidation of one molecule of palmitic acid to CO_2 and H_2O, 130 molecules of ATP are produced.
36. **Haemoglobin** is a globular protein and the red colour of haemoglobin is due to the iron-protopor-phyrin complex called the **heme**.
37. The bicarbonate/carbonic acid system, i.e., HCO^-_3/H_2CO_3 acts as the **buffer** and maintains the pH of blood between 7.36 – 7.42.
38. **White Blood Cells (WBC)** produce antibodies or immunoglobulins which are Y-shaped protein molecules which protect the body against infections.
39. Water soluble vitamins are B_1, B_2, B_6 and C.
40. Water insoluble vitamins are A, D, E and K.
41. Vitamins which act as antioxidants are C, D and E.
42. Calorific value of various food stuffs follows the sequence ; Fats > Carbohydrates > proteins.
43. Allergy is caused by production of histamines.
44. While writing the structure of peptides, the **N-terminal amino acid** is written on the left while the C-terminal amino acid is written on the right hand side. Similarly, the name of any peptide is written starting from the **N-terminal residue.**