

Biomolecules and biopolymers

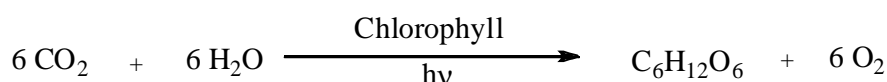
Department of Chemistry
The Open University of Sri Lanka

Biomolecules and biopolymers

Introduction

In the lesson on synthetic polymers and their applications, we discussed the synthesis and some applications of man-made synthetic polymers such as polyethylene, poly(vinyl chloride), polyamides and polyesters. Macromolecules or biopolymers such as *carbohydrates* (e.g. *polysaccharides*), *proteins*, and *nucleic acids* together with water, make up most of the human body. Millions of chemical reactions take place within each cell per every second. We depend on *fats and carbohydrates* to supply energy for our activities. Proteins provide building materials needed to grow and develop. In this lesson, we intend to discuss the chemistry of carbohydrates (monosaccharides, disaccharides, and polysaccharides), peptides, proteins and nucleic acids.

The most important reaction for the existence of life is photosynthesis. During photosynthesis, simple molecules such as carbon dioxide (CO₂) and water (H₂O) are combined to produce glucose (C₆H₁₂O₆).



Glucose is stored in plants as **cellulose** which is a **macromolecule** consisting of thousands of glucose units (in Greek *makro* means large or long). In fact, cellulose is a **condensation polymer of glucose** produced by removing a water molecule between every two glucose residues. Let us look at the chemistry of carbohydrates.

1. Carbohydrates

Carbohydrates are naturally occurring compounds containing C, H and O. They are also called *sugars*. Carbohydrates play an important role in *biochemistry* - they store chemical energy, provide energy through oxidation, and supply carbons for synthesis of cell components. Simple carbohydrates have the *empirical formula* CH₂O, *i.e.* a hydrate of carbon, C(H₂O). A more general molecular formula can be written as C_x(H₂O)_y, for example, molecular formulae of the two most common sugars **glucose** (C₆H₁₂O₆) and **sucrose** (C₁₂H₂₂O₁₁) can be written as C₆(H₂O)₆ and C₁₂(H₂O)₁₁, respectively.

A carbohydrate consisting of one sugar unit is called a **monosaccharide** (e.g. glucose). A carbohydrate having two sugar units is called a **disaccharide** (e.g. sucrose). Most simple sugars are *soluble in water*. Sugars consisting of two different monosaccharides are called **heterosaccharide**. Carbohydrates composed of two to eight units of monosaccharides are referred to as oligosaccharides (in Greek *oligo* means a few). Carbohydrates composed of

Glyceraldehyde is an aldotriose and is a **chiral** (*Greek cheir, the hand*) molecule as carbon-2 (C^2) has four different groups attached to it. These two molecules have the same molecular formula, but the spatial (or 3D) arrangement of groups is different as shown below. **D**-glyceraldehyde rotates the plane of the plane-polarized light to the **right** and **L**-glyceraldehyde to the **left**.

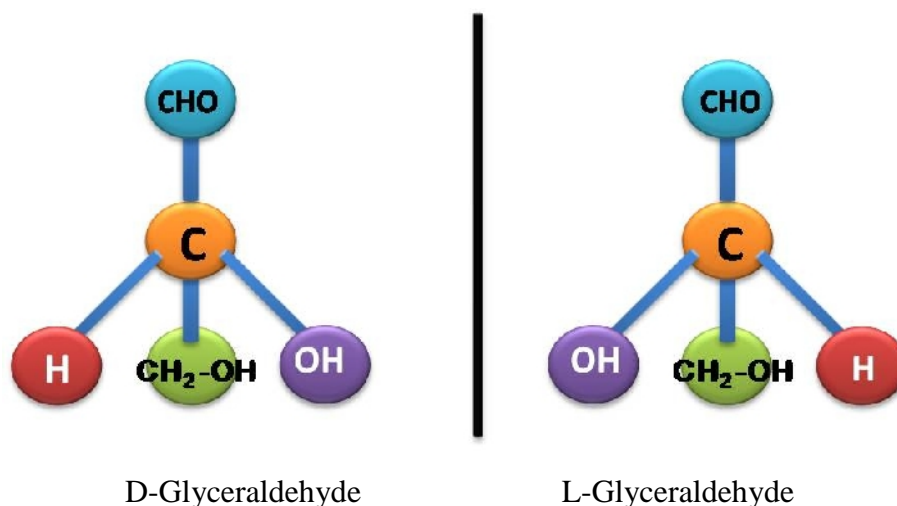


Figure 2 3D view of the D-isomer with the mirror image, i.e. L-isomer. D-isomer is not superimposable with the L-isomer, i.e. they are optically active.

1.2 Chirality and enantiomers

You have learnt previously that an *asymmetric* or *chiral* carbon has four different groups attached to it, the **mirror image** (L-isomer) of the molecule is **not** superimposable with **the other** (D-isomer) (*superimposable means one form can be placed exactly over the mirror image of it*). Chiral molecule with **one** chiral centre produces two *optically active stereoisomers* called **enantiomers**. Hands, ears, shoes and hand gloves are all chiral. If the mirror image of a molecule is superimposable with the molecule itself, it is said to be **symmetrical** or **achiral**. This is because the molecule has a *plane of symmetry*. For example, glycerol, glycine and *n*-propanol are achiral (*i.e.* they are *optically inactive*).



Activity

- Which of the following compounds are chiral?

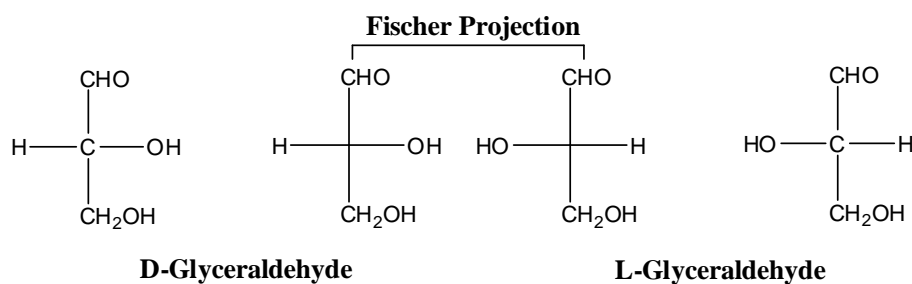
(i) lactic acid	(ii) isopropyl alcohol	(iii) vinyl alcohol
(iv) 2-chloropropanol	(v) 2-hydroxy butane	(vi) $CH_3CH(OH)CH(=O)$

1.3 Fischer projection

It is time-consuming to draw the 3D shape of molecules. German chemist Emil Fischer introduced an easy method to represent more-complex 3D-molecular structures in two dimensions (2D).

In the Fischer projection, the *carbon chain* is written **vertically**, with *substituent groups* to the **left** and **right**. The principal functional group (*aldo* or the *keto* group) is always at the **top** of the carbon chain. The *asymmetric carbon* is in the plain of the paper, carbons attached to the asymmetric carbon are placed *vertically and below the plane*. The groups to the right and left of the asymmetric carbons are *horizontally above the plane*.

If the hydroxyl group is on the right hand side, it is known as the **D-isomer**. If the hydroxyl group is on the left hand side, it is known as the **L-isomer**. Normally, naturally occurring form of sugars is D.



Activity

2. Draw the Fischer projection of the aldotetrose, D-erythrose ($\text{C}_4\text{H}_8\text{O}_4$).

1.4 Furanose and pyranose structures

It is convenient to draw the structures of common sugars (pentoses and hexoses) as open chain compounds, but they mainly form five- and six-membered rings as shown below for D-glucose and D-fructose.

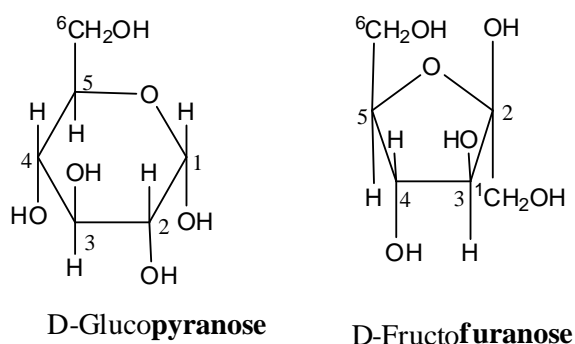


Figure 3 Pyranose and furanose structures

The five-membered and six-membered ring systems are called **furanose** and **pyranose**, respectively. In **Haworth projections**, the attached groups have been drawn above and below the plane of the ring. Thus, it gives the *stereochemistry* of the carbon-1 (C¹) depending on the position of the hydroxyl group; in **α -anomer**, the hydroxyl group is placed *below the ring* whilst in the **β -anomer**, the hydroxyl group is placed *above the ring*. According to Haworth projection, the α - and β -anomers of D-glucopyranose and D-fructofuranose are shown below.

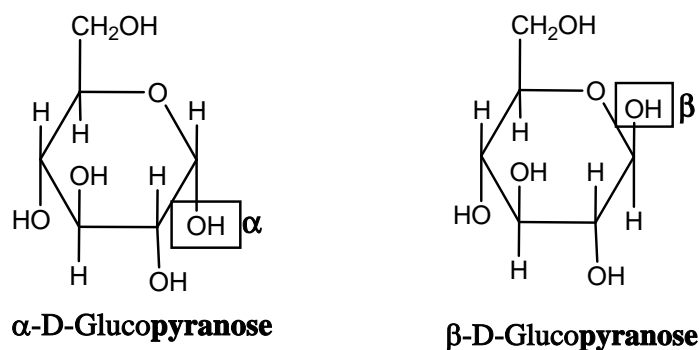


Figure 4 Haworth projections of α - and β -anomers of D-glucopyranose

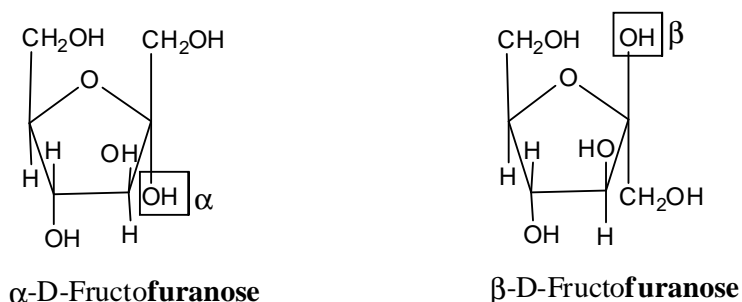


Figure 5 Haworth projections of α - and β -anomers of D-fructofuranose

2. Monosaccharides

Ribose and **deoxyribose** are extremely important **pentose sugars** because they are involved in the synthesis of nucleic acids, which are essential in protein synthesis and transfer of genetic materials.

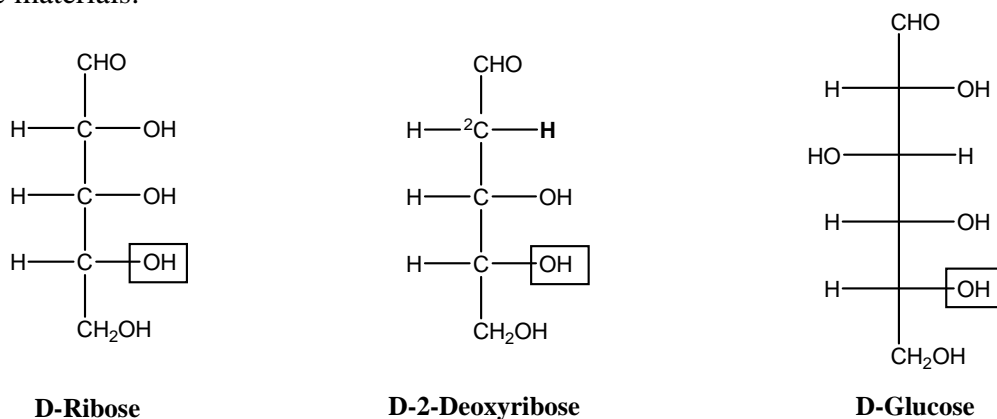


Figure 6 The structures of some monosaccharides: ribose, deoxyribose and glucose

Of **hexose sugars**, glucose is the most important nutritionally and it is the most abundant sugar in nature. Glucose is present in honey and fruits such as grapes and dates; ripe grapes contain 20-30% of glucose. It is also known as **blood sugar** because it is transported in blood to all parts of the body in order to generate energy. Glucose is commonly used as a sweetener.

3. Disaccharides

Disaccharides are sugars composed by **two** monosaccharide units linked by a **glycosidic bond**. A disaccharide can be hydrolysed to yield its monosaccharide building blocks by boiling it with dilute acid or with appropriate enzyme. Nutritionally, the most important members of this group are maltose, lactose and sucrose; their components and sources are given in Table 1.

Table 1 Components of some disaccharides

Disaccharide	Monosaccharide	Type of linkage	Source
Maltose	α -D-glucose, α -D-glucose	(1 4)	Starch
Lactose	-D-galactose, α -D-glucose	(1 4)	Mammalian milk
Sucrose	-D-glucose, -D-fructose	(1 2)	Sugar cane and beet

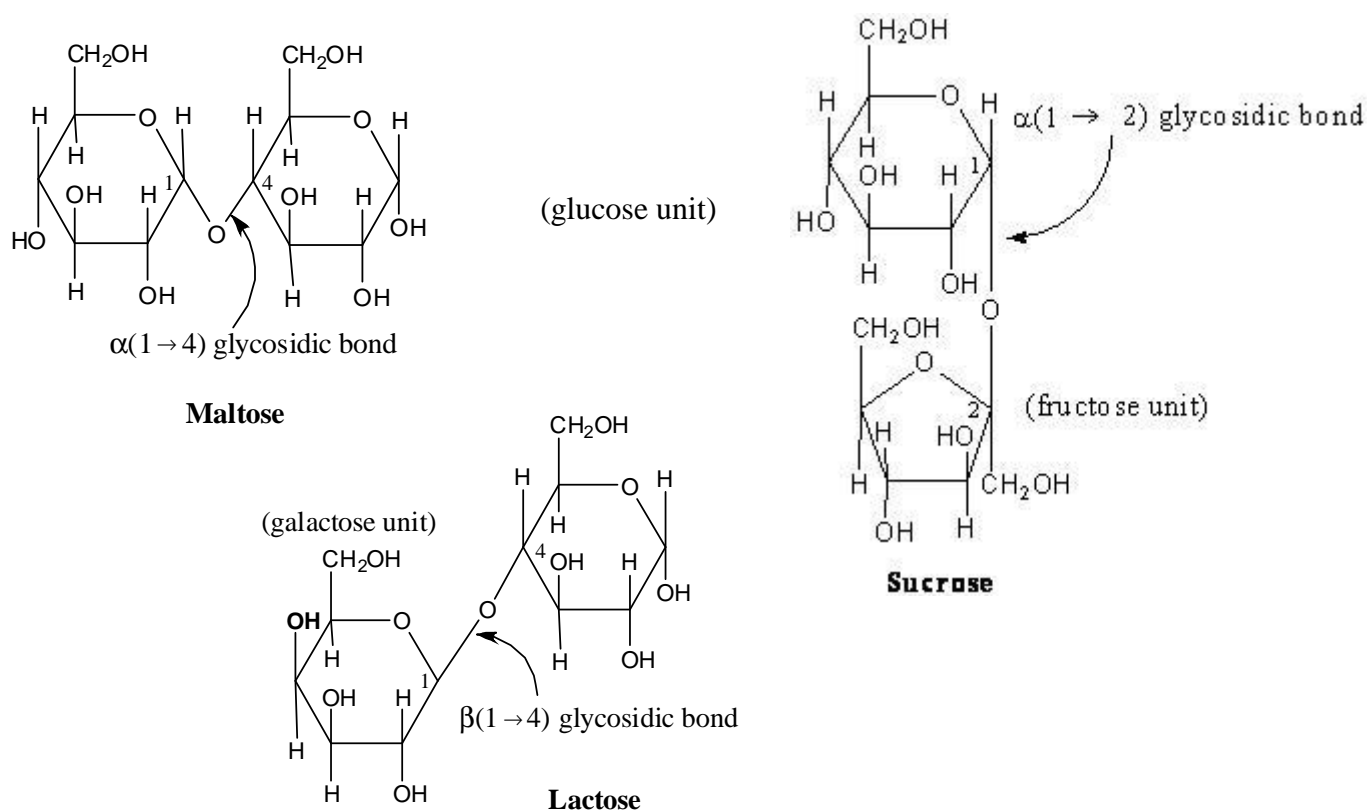


Figure 7 Structures of some disaccharides: maltose, sucrose, and lactose

Sucrose is used as a food additive. It could contribute to obesity and type-II diabetes. It causes dental caries. Sucrose is hydrolysed by bees in making honey.

4. Polysaccharides

Just as two sugar units are linked together to form a disaccharide, additional units can be added to form higher saccharides. Because of their size, polysaccharides are not water soluble, but many hydroxyl groups become hydrated individually when exposed to water and some polysaccharides form thick colloidal dispersions when heated in water. Starch, cellulose and glycogen are the most common examples of polysaccharides.

4.1 Starch

Starch is a polymer consisting entirely of D-glucose units connected by (1 → 4) linkages. It is the major storage form of D-glucose in plants. Two starch fractions, amylose (10-20%) and amylopectin (80-90%), can usually be isolated from plants. Amylose is made up of long *unbranched* chains of glucose units connected by (1 → 4) linkages. The long chain, often containing between 1000 to 2000 glucose units, is flexible enough to allow the molecules *to twist to the shape of a helix*. Iodine reacts with starch (coiled form of amylose) to give a deep blue colour. This blue colour gradually fades and finally disappears as starch is hydrolysed either by acids or enzymes.

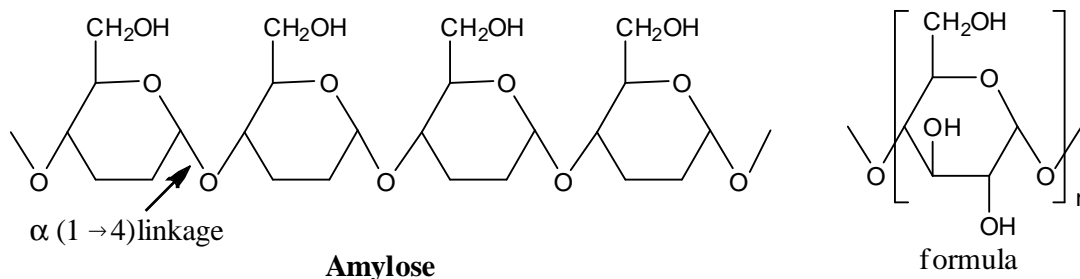


Figure 8 Partial structure of amylose (some H and OH groups are not included for clarity)

Amylopectin, the second component of starch, is not a straight-chain molecule like amylose but contains random branches. The branching point is (1 → 6) glycosidic linkage. There are usually 24 to 30 D-glucose units, all connected by (1 → 4) linkages, between each branch point of amylopectin. Amylopectin contains as many as 100,000 glucose units in one gigantic molecule.

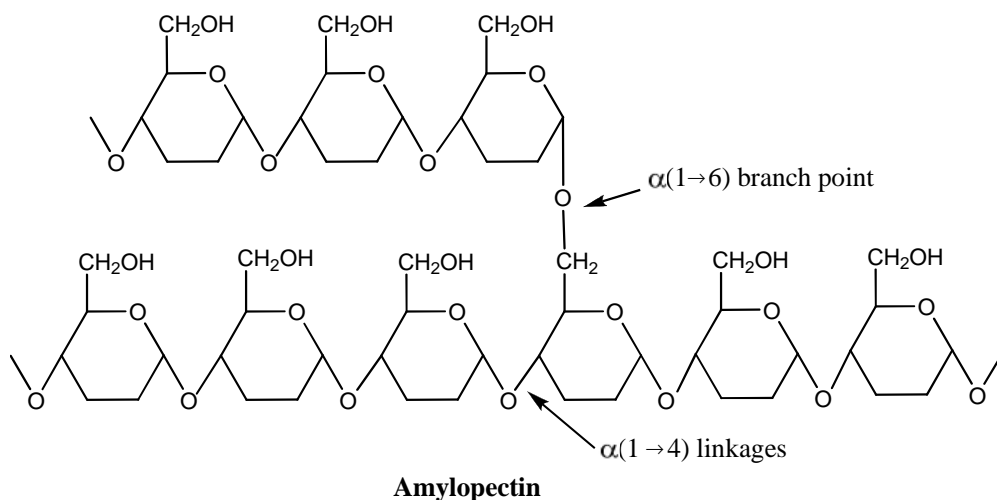


Figure 9 Partial structure of amylopectin (some H and OH groups are not included for clarity)

4.2 Cellulose

Cellulose (fig.10) is the most important structural polysaccharide and is the single most abundant organic compound on earth. It is the material in plant cell wall that provides strength and rigidity. Wood is about 50% cellulose. Like amylose (fig.8), cellulose is a **linear polymer** consisting of D-glucose units joined by (1 → 4) linkages. It may contain 300 to 3000 glucose units in each molecule. The main structural difference between amylose and cellulose is that all the (1 → 4) glycosidic linkages in cellulose are β instead of α . This seemingly small difference causes tremendous differences between amylose and cellulose. The shapes of the molecules are quite different. Amylose with (1 → 4) links tends to form loose spiral structures, whereas cellulose with (1 → 4) links tends to form extended straight chains. These chains become aligned side-by-side to form well-organised, water insoluble fibres in which the hydroxyl groups form numerous **hydrogen bonds** with the neighbouring chains. These parallel chains of cellulose provide rigidity and strength to crystalline molecules.

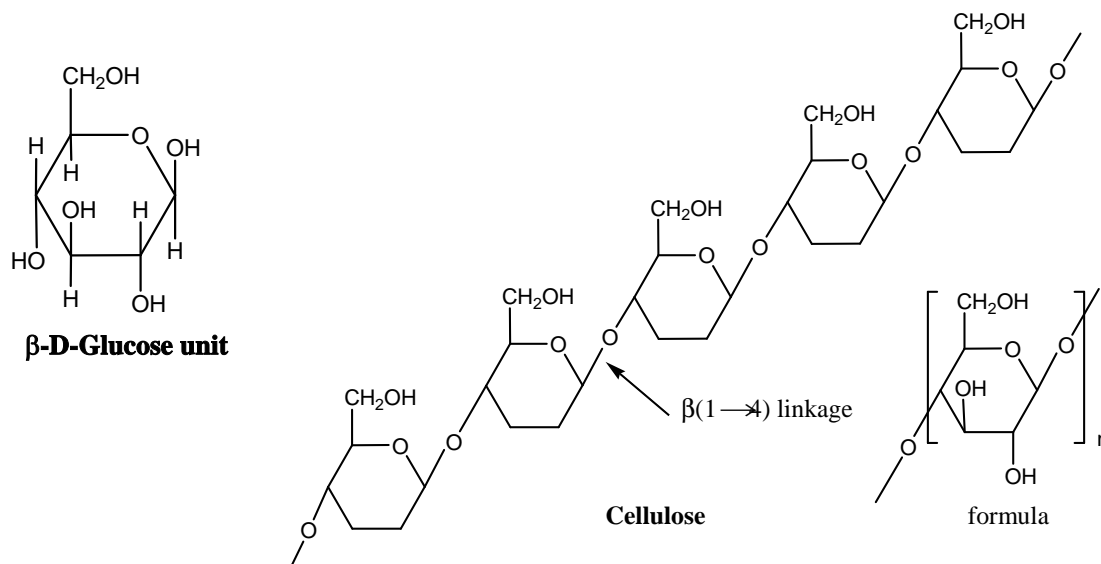


Figure 10 Partial structure of Cellulose (some H and OH groups are not included for clarity)

Most animals lack the necessary enzymes to hydrolyse linkages of cellulose. Animals that use cellulose as a food do so only with the assistance of bacteria that possess the necessary enzymes. Herbivores such as cows, sheep and horses are animals in this category. Each has a colony of such bacteria somewhere in the digestive system and uses the simple carbohydrates resulting from bacterial hydrolysis of cellulose.

4.3 Glycogen

Glycogen is sometimes called animal starch because it is a storage carbohydrate for animals that is analogous to the starch of plants. It is especially abundant in liver and muscles, where much of the excess glucose taken in by an animal is stored for future use. Upon hydrolysis, glycogen forms D-glucose, which helps maintain the normal blood sugar level and provides the muscles with energy. Structurally, glycogen is very similar to amylopectin, containing both (1 → 4) and (1 → 6) linkages between glucose units. The main difference between amylopectin and glycogen is that glycogen is even more highly branched.

5. Nucleic acids

Nucleic acids possess genetic information and are the control centres of the cell. Nucleic acids are classified into two categories: (i) **Ribonucleic acids** (RNA), which are found mainly in cytoplasm in living cells and are involved in protein biosynthesis, and (ii) **Deoxyribonucleic acids** (DNA), which are found primarily in the nucleus of the cell, and also in mitochondria and chloroplast. Its sequence of bases gives the genetic code. It is responsible for transferring of this genetic information from one generation to the other and also in determining the RNA and protein synthesis in cells.

Nucleic acids are macromolecules consisting of long, linear molecules. The repeating structural unit or the **monomer** of nucleic acids is called **nucleotide**. The components of DNA and RNA are given in Table 2.

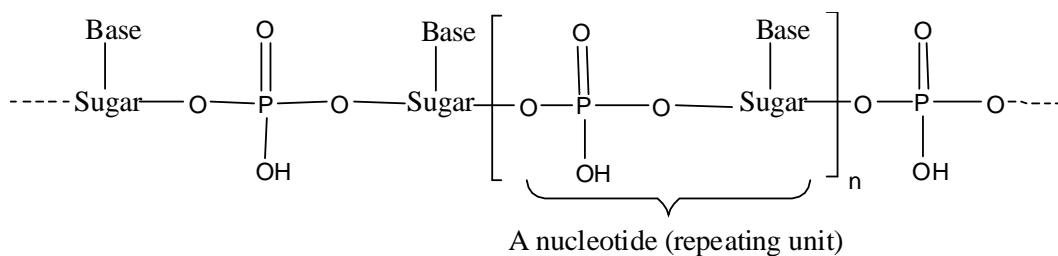


Figure 11 Polymer chain of a nucleic acid

Table 2 Components of DNA and RNA

	DNA	RNA
Purine bases	Adenine	Adenine
	Guanine	Guanine
Pyrimidine bases	Cytosine	Cytosine
	Thymine	Uracil
Pentose sugar	Deoxyribose	Ribose
Inorganic acid	Phosphoric acid	Phosphoric acid

Nucleotides and nucleosides

Nucleotides are composed of **three** simpler components: a phosphate, a sugar, and a heterocyclic base. Nucleoside consists of only **two** components - a base and a sugar. In a polymer chain, nucleosides are placed between two phosphate units.

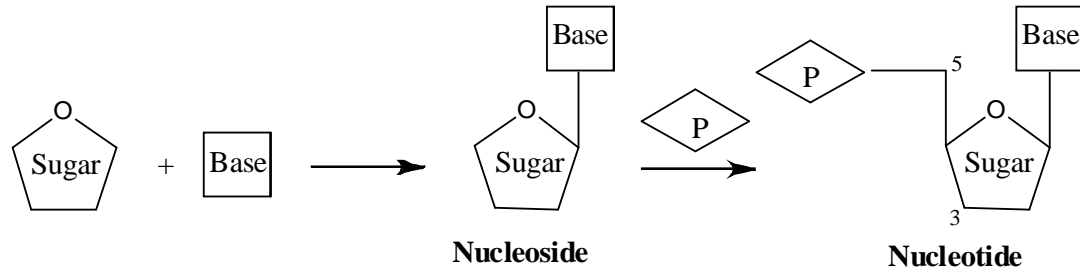


Figure 12 Pictorial presentations of nucleosides and nucleotides

The Sugars

There are two sugars, (i) ribose in RNA and (ii) 2-deoxyribose in DNA

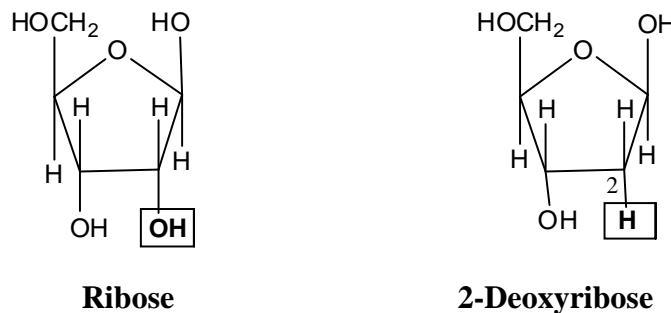


Figure 13 Structures of ribose and deoxyribose

Pyrimidine and purine bases

The three **pyrimidine bases** are cytosine, thymine and uracil usually abbreviated C, T and U.

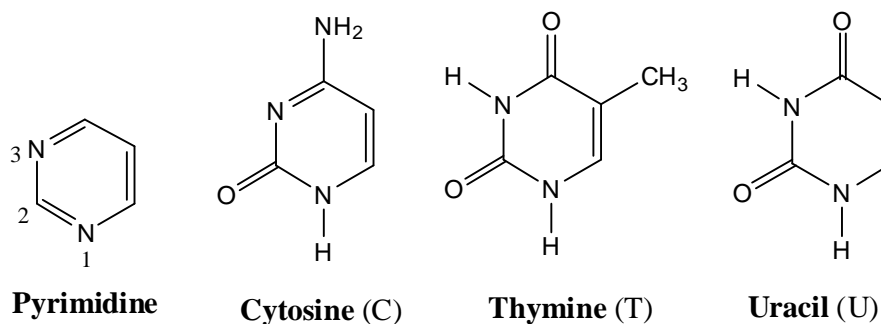


Figure 14 Structures of pyrimidine and pyrimidine bases

Adenine (A) and guanine (G) are the two **purine bases**.

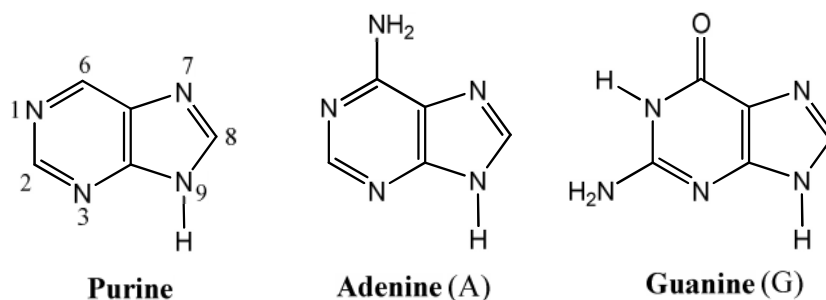


Figure 15 Structures of purine and purine bases

Adenine, guanine and cytosine are found in both DNA and RNA, but uracil is found only in RNA and thymine only in DNA.

Phosphate

The third component of nucleotides is derived from phosphoric acid (H_3PO_4), which under cellular pH conditions exists in the ionic form HPO_4^{2-} , *i.e.* ($^-\text{O}-\text{P}(=\text{O})(\text{OH})-\text{O}^-$).

Nucleosides formed with the ribose sugar and bases A, G, C and U are called adenosine, guanosine, cytosine and uridine. Nucleosides formed with the 2-deoxyribose sugar and bases A, G, C and T are called deoxyadenosine, deoxyguanosine, deoxycytidine and deoxythymidine. The two nucleotides (in the RNA and DNA) containing the base cytosine and one phosphate group are shown below.

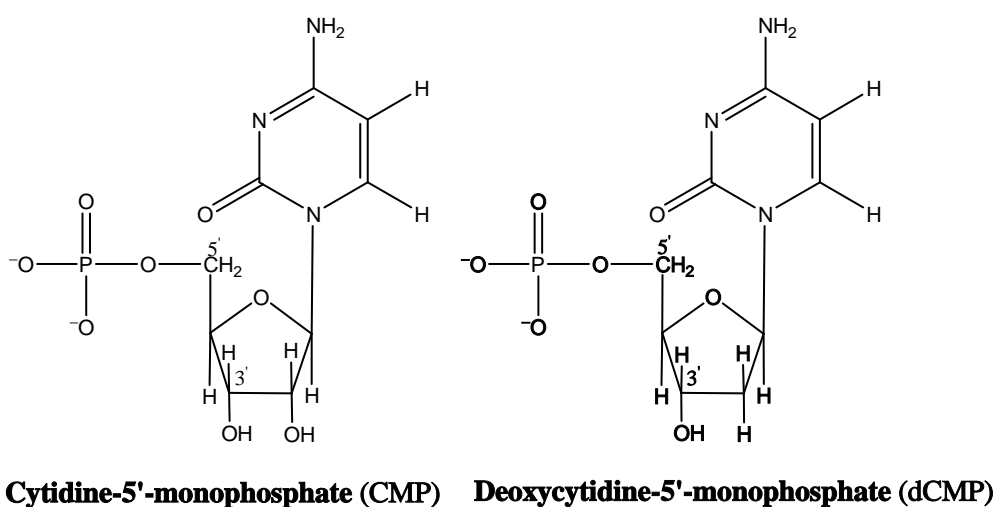


Figure 16 Nucleotides with the base cytosine and one phosphate group

5.1 Structure of DNA

DNA molecules, among the largest molecules known, contain between 1 and 100 million nucleotide units. Nucleosides are joined together by **phosphate** units linking the 5' carbon of one nucleoside to the 3' carbon of the other. These linkages are referred to as **phosphodiester bonds**. The **phosphate-sugar chain** is referred to as the nucleic acid **backbone** and it is constant throughout the entire DNA molecule. Thus one DNA molecule differs from another only in the sequence or order of attachment of the bases along the backbone.

In DNA, molar amount of **adenine** (A) is equal to the molar amount of **thymine** (T). Similarly, molar amount of **guanine** (G) is equal to the molar amount of **cytosine** (C). To account for this balance, the bases must be paired by hydrogen bonding, A to T and G to C. The two scientists, James D. Watson and Francis Crick worked out the **double helix** model of DNA in 1953. They won the Nobel Prize for Chemistry in 1968. (S = sugar; P = phosphate).

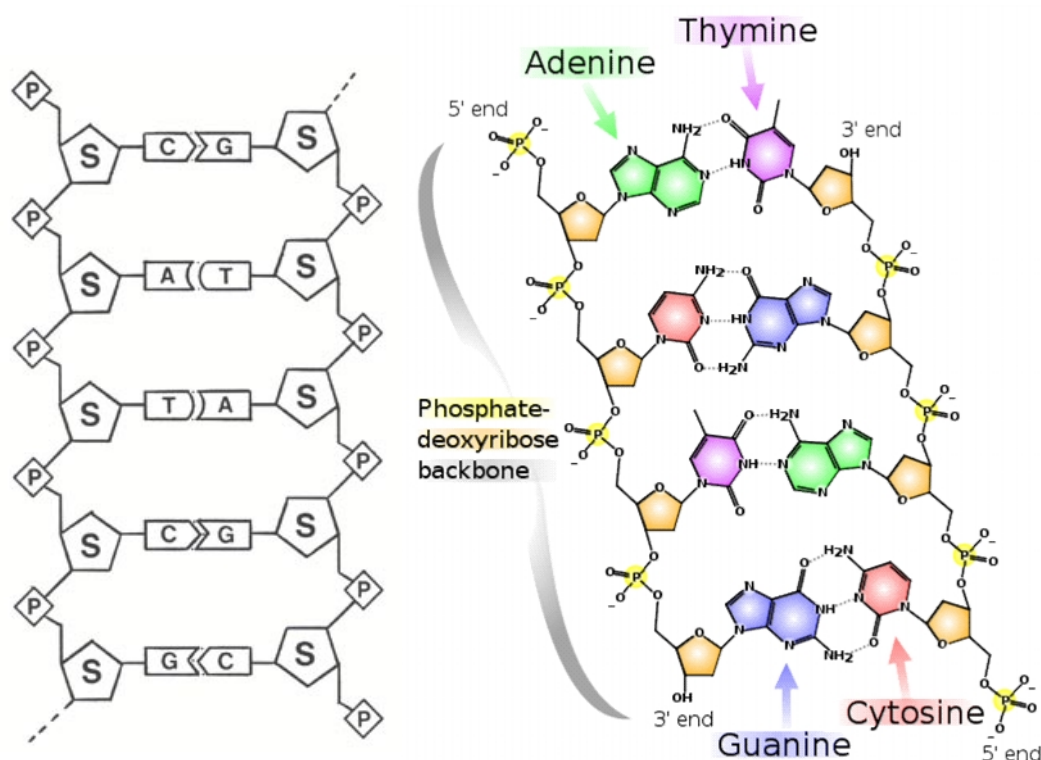


Figure 17 Pictorial presentation of a DNA molecule; there are hydrogen bonds between bases: A and T (two hydrogen bonds); C and G (three hydrogen bonds).

6. Amino acids, peptides and proteins

As indicated by the name itself, amino acids are **bifunctional** molecules. In naturally occurring **α -amino acids**, $\text{H}_2\text{NCH(R)C(=O)OH}$, both functional groups (NH_2 and CO_2H) are attached to the **same carbon**. R = H or an organic group and it is called the *side chain*. The carbon next to the carboxylic carbon is called the **α -carbon**. Since four different groups are attached to the α -carbon (*i.e.* carbon is asymmetric) most amino acids are *optically active*.

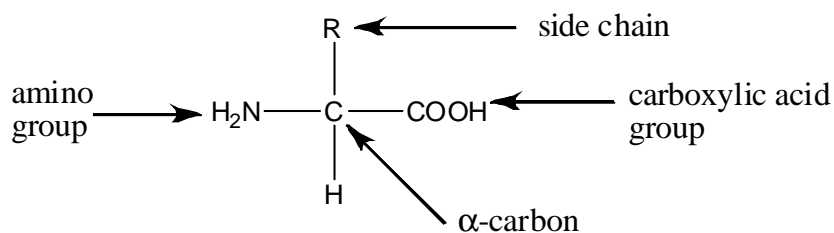


Figure 18 Structure of a α -amino acid

About 20 essential amino acids are found in proteins. Glycine is the **only achiral** amino acid and others could have both D- and L-forms. Only L-amino acids are associated with proteins. Some simple α -amino acids are given below.

Name (Abbreviation)	R group
α -Glycine (Gly)	H
α -Alanine (Ala)	CH ₃
α -Phenyl alanine (Phe)	CH ₂ Ph
α -Valine (Val)	CH(CH ₃) ₂
α -Leucine (Leu)	CH ₂ CH(CH ₃) ₂

A **peptide** (*e.g.* Ala-Gly-Val) is a sequence of α -amino acids linked *via* **amide bonds**, which are also known as **peptide bonds**. In simple di-, tri- and tetra-peptide, there is *one free terminal amino group* and *one free terminal carboxylic group at the end of the molecule*. Peptides with more than 10 amino acid residues are generally known as **polypeptides**. Bovine insulin (a peptide hormone) has a total of 51 amino acids in two polypeptide chains. Insulin deficiency causes diabetes. **Proteins** are biological macromolecules with the molecular weight over 5000 Dalton and contain a number of long polypeptide chains.

Names of peptides are always written from **left to right** starting with the *N-terminal amino acid*. Ala-Gly-Val is the abbreviation used for the **tripeptide**, alanyl-glycyl-valine. The structure of Ala-Gly-Val is shown below.

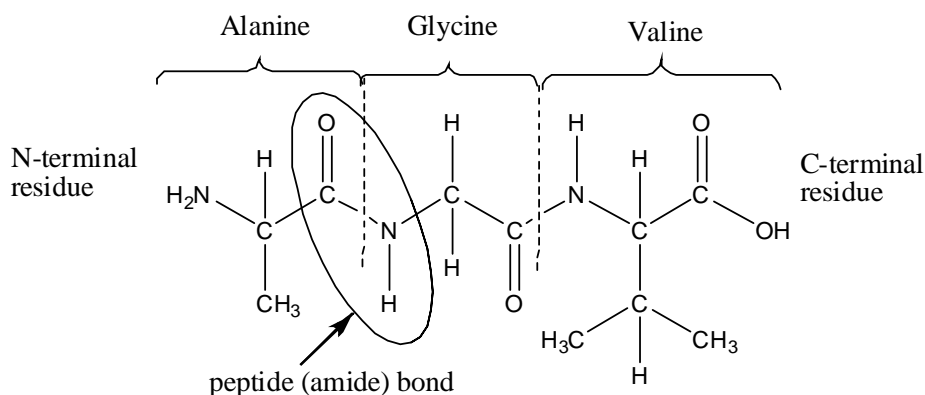


Figure 19 The structure of the tripeptide, alanyl-glycyl-valine (Ala-Gly-Val).

Summary

- Glucose is stored in plants as cellulose, which is a condensation polymer of glucose.
- Functions of carbohydrates are:- store chemical energy, provide energy through oxidation, and supply carbons for synthesis of cell components. Sugars are polyhydroxy aldehydes or ketones, with the general molecular formula $C_x(H_2O)_y$.
- Aldoses containing three, four, five and six carbons are called aldotriose, aldotetrose, aldopentose, aldohexose, respectively. Ketoses are classified similarly.
- Glyceraldehyde is a chiral molecule and exists in D- and L-forms (enantiomers). D-form rotates the plane of the plane-polarized light to the right and L-form to the left. Achiral molecules such as glycerol and glycine are optically inactive as each has a plane of symmetry.
- Five- and six-membered cyclic sugars are called furanoses and pyranoses, respectively. Haworth projection gives the stereochemistry at C-1; in α -form, the OH is placed below the ring whilst in the β -form, the OH is placed above the ring.
- Ribose, deoxyribose and glucose are the most important monosaccharides; in disaccharides two monosaccharide units are linked with a glycosidic bond. Cellulose, starch and glycogen are polysaccharides containing many glycosidic linkages.
- Nucleic acids (DNA and RNA) are macromolecules consisting of long, linear molecules. The repeating structural unit of nucleic acids is called nucleotide, which is made up of a phosphate, a sugar, and a heterocyclic base. Nucleoside consists of only a sugar and a heterocyclic base.
- The three pyrimidine bases are uracil (U), thymine (T) and cytosine (C); Adenine (A) and guanine (G) are the two purine bases.
- A peptide is a sequence of α -amino acids linked *via* amide bonds; peptides with more than 10 amino acid residues are generally known as polypeptides; proteins are biological macromolecules containing long polypeptide chains.



Learning Outcomes

Once you have finished studying this lesson you should be able to

- classify carbohydrates as aldoses, ketoses or mono-, di- and poly-saccharides
- state whether a sugar is in D- or L-form; to identify α - and β -anomers of a given sugar
- draw the Fischer projection of aldotriose and aldotetroses
- comment on structural features of cellulose, starch, amylose, amylopectin, and glycogen
- explain structural features of nucleoside, nucleotide and nucleic acids
- comment on the three components of DNA and RNA
- explain what α -amino acids, peptides, polypeptides and proteins are



Activity

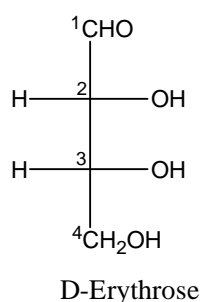
3. Draw the structure of L-threose which is the mirror image of D-threose.
 4. Draw possible structure(s) for a keto-tetrose.
 5. How many chiral carbons are there in glucose? Identify the carbon(s) which is/are not chiral.
 6. Draw Fischer projections of the D- and L-forms of the following compounds.
 - (i) lactic acid, $\text{CH}_3\text{CH}(\text{OH})\text{CO}_2\text{H}$
 - (ii) alanine, $\text{CH}_3\text{CH}(\text{NH}_2)\text{CO}_2\text{H}$
 7. Draw the structure of the disaccharide, maltose. Identify the type of glycosidic linkage in maltose. What is the relationship between maltose and starch?
 8. What are the differences between amylose and amylopectin?
 9. What are the DNA components which are not found in RNA?
 10. Comment on the structure of DNA.
 11. Which purine and pyrimidine bases are common to both DNA and RNA?
 12. Draw the structure of the tripeptide, valinyl-glycyl-phenylalanine (Val-Gly-Phe).
 13. Alanine is an α -amino acid, write the β -amino acid with the same chemical formula.
-
-

Answer guide to activities

1. Chiral molecules have an asymmetric carbon centre and the two mirror images are not superimposable.

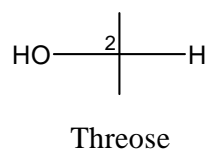
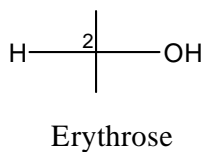
(i) chiral (ii) achiral (iii) achiral (iv) chiral
(v) chiral (vi) chiral

2.

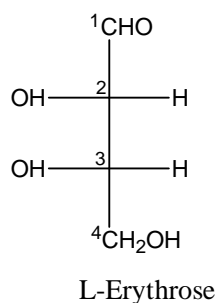


Aldotetroses have four carbon centres, of which the CHO group and CH₂OH units are placed at the top and bottom, respectively. The OH group on the C-3 is placed on the **right** as this is the D-isomer. Depending on the position of the OH group on C-2, the two aldotetroses (erythrose and threose) are named.

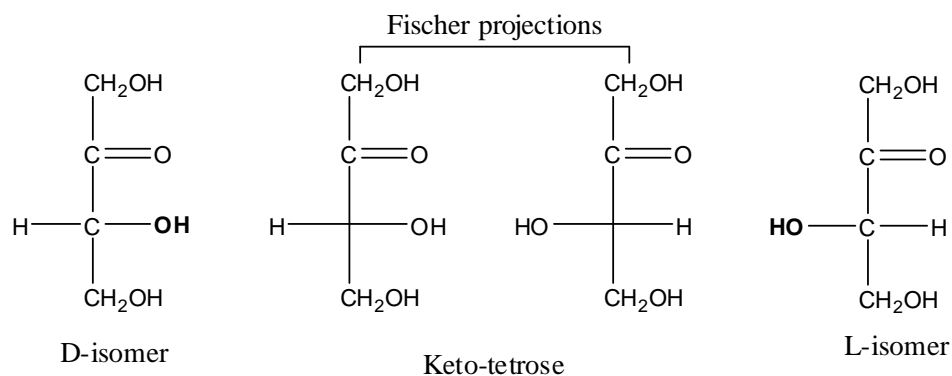
You can see, the OH group on C-2 in erythrose is on the **right** hand side. In threose, the other aldotetrose, the OH group on C-2 is on the **left** hand side.



3.

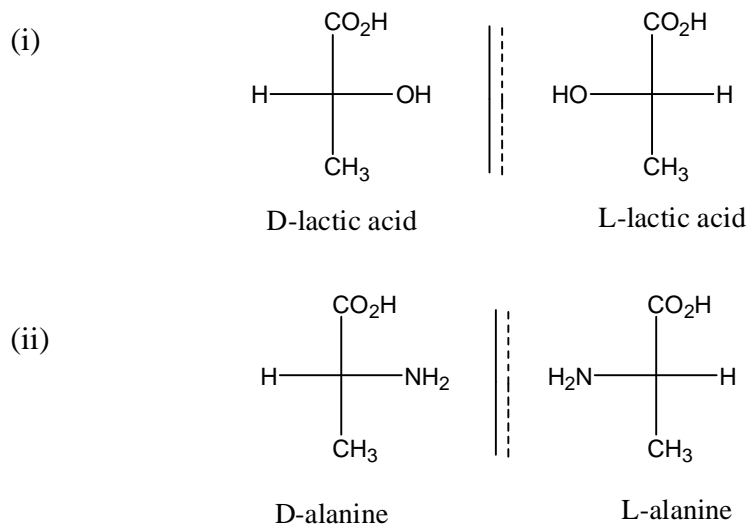


4.



5. Four chiral carbons. They are C-2, C-3, C-4 and C-5. Achiral carbons are C-1 and C-6.

6.



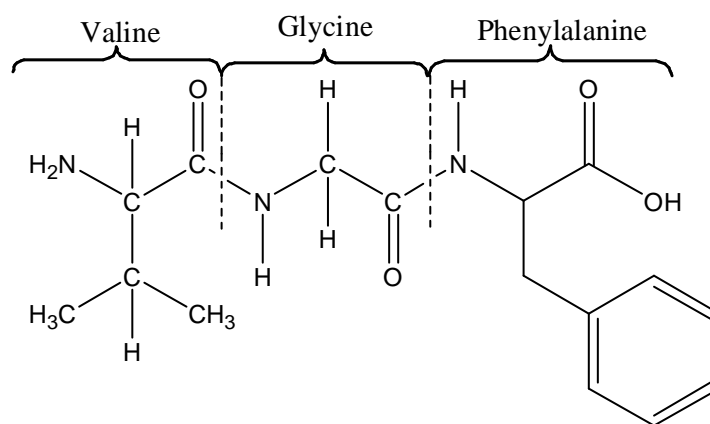
7. Refer section 3

Maltose and starch are made up of D-glucose units with the $\alpha(1 \rightarrow 4)$ glycosidic bonds.

8. Refer section 4.1

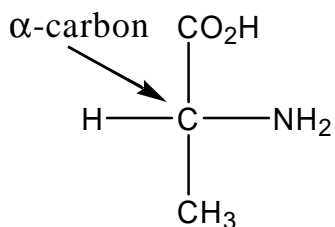
9. Deoxyribose and thymine
10. Refer section 5.1
11. Adenine, guanine and cytosine

12.

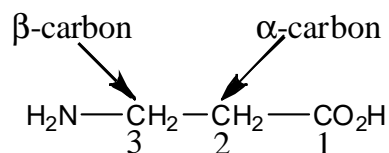


Structure of Val-Gly-Phe

13. In Alanine ($C_3H_7O_2N$) both NH_2 and CO_2H groups are attached to α -carbon.



The structure of β -amino acid with the same composition, $C_3H_7O_2N$ is,





Study Questions

1. How many members are in the aldotetrose family? Write their names.
2. Name the basic components of DNA.
3. Comment on structure of cellulose.
4. How many different aldopentoses (aldo-sugars with five carbons) are there?
5. Draw Fischer projections of the D- and L-forms of the following compounds.
 - (i) ribose
 - (ii) glucose
6. What is the main difference between a nucleoside and a nucleotide?
7. Name the components of RNA.
8. Name the two purine bases and the three pyrimidine bases.
9. What are the differences in cellulose and glycogen?
10. Draw the structure of the tetrapeptide, Val-Gly-Phe-Gly.
11. Write the exact names of aldoses and ketoses given in figure 1.

References

1. Understanding chemistry for advanced level, T. Lister and J. Renshaw, 1991, Stanely Thornes (Publishers) Ltd.
2. Chemistry and our world, C. G. Gebelein, 1997, Wm. C. Brown Publishers.
3. Advanced Chemistry, P. Matthews, 1992, Cambridge University press.
4. General Chemistry Selected Topics, J. W. Hill and R. H. Petrucci, 1996, Prentice Hall.
5. Chemistry in Context, L. P. Eubanks, C. H. Middlecamp, N. J. Pienta, C. E. Heltzel, G. C. Weaver, 5th Edition, 2006, McGraw-Hill.
6. Chemistry for today, S. L. Seager and M. R. Slabaugh, 2nd Edition, 1994, West Publishing Company.

Course Team

Author

Professor K. Sarath D. Perera
Senior Professor in Chemistry

Content Editor

Ms. Chandani Ranasinghe
Lecturer in Chemistry/OUSL

Language Editor

Mrs. Nirmalie Kannangara

Desk Top Publishing

Miss. K. K. H. De Silva
Mr. R. M. Wimal W. Wijenayake

Graphic Artists

Miss. K. K. H. De Silva
Mr. R. M. Wimal W. Wijenayake

Word Processing

Miss. K. K. H. De Silva
Professor K. Sarath D. Perera

Web Content Developers

Miss. Hashika Abeysuriya
Miss. L. Melani Silva

Cover Page Designing

Professor K. Sarath D. Perera
Mr. R. M. Wimal W. Wijenayake

The Open University of Sri Lanka
Nawala, Nugegoda, Sri Lanka

First published 2012

ISBN: 978-955-23-1351-6

