

S ABEWHMI Science | Technology | Medicine

LECTURE NOTE

BCM 220 STRUCTURE AND CHEMISTRY OF MOLECULES

PREPARED BY: L.B. BELLO

bello.lukman@tau.edu.ng

COURSE OUTLINE

Structure and chemistry of

Carbohydrates

Amino acids & Proteins

Lipids

Nucleic acids

Water

Enzymes and cofactors

FOCUS...

Structure and chemistry of

- Carbohydrates
- Nucleic acids

DID YOU KNOW?

Carbohydrates are the most abundant organic molecules on Earth.

Each year, photosynthesis converts more than 100 billion metric tons of CO_2 and H_2O into cellulose and other plant products.

Nucleic acids are the molecules that code the genetic information of organisms.

It is said that every human being shares 99.9% of their DNA with every other human.

Water makes up about 60% of the body weight.

A person can live about a month without food, but only about a week without water.

4

CARBOHYDRATES

DESINITIONS

Carbohydrates may be defined as optically active polyhydroxy aldehydes or ketones or compounds, which produce them on hydrolysis.

They are primarily composed of the elements carbon, hydrogen and oxygen.

The name carbohydrate literally means **'hydrates of carbon'**.

Many (but not all) of the carbohydrates possess the empirical formula $(C.H₂O)_n$.

Functions of carbohydrates

 They are the most abundant dietary source of energy (4 Cal/g) for all organisms.

- They are precursors for many organic compounds (fats, amino acids).
- Carbohydrates (as glycoproteins and glycolipids) participate in the structure of cell membrane and cellular functions such as cell growth, adhesion and fertilization.
- \checkmark They are structural components of many organisms. These include the fiber (cellulose) of plants, exoskeleton of some insects and the cell wall of microorganisms.
- \checkmark They also serve as the storage form of energy (glycogen) to meet the immediate energy demands of the body.

Classification of Carbohydrates

Carbohydrates are often referred to as saccharides (Greek: *sakcharon*–sugar), and they are broadly classified into 3 major groups— monosaccharides, oligosaccharides and polysaccharides.

•• This classification is based on the number of sugar units.

Mono- and *oligosaccharides* are sweet to taste, crystalline in character and soluble in water, hence they are commonly known as *sugars*.

Monosaccharides

Monosaccharides (Greek : *mono*-one), also known as simple sugars, are colourless, crystalline solids, which consist of a single polyhydroxy aldehyde or ketone unit, and are freely soluble in water but insoluble in non-polar solvents.

- \cdot They have the general formula $C_n(H_2O)_n$, and they cannot be further hydrolysed.
- The most abundant monosaccharide in nature is the six-carbon sugar D-glucose, sometimes referred to as **dextrose**.
- Monosaccharides are divided into different categories, based on the functional group and the number of carbon atoms.
- Monosaccharides of more than four carbons tend to have cyclic structures.

The backbones of common monosaccharide molecules are unbranched carbon chains in which all the carbon atoms are linked by single bonds.

In the open-chain form, one of the carbon atoms is double-bonded to an oxygen atom to form a **carbonyl group**; each of the other carbon atoms has a hydroxyl group.

 \cdot If the carbonyl group is at an end of the carbon chain (that is, in an **aldehyde group**) the monosaccharide is an **aldose**.

 \cdot If the carbonyl group is at any other position (in a **ketone group**) the monosaccharide is a **ketose**.

The simplest monosaccharides are the two three-carbon trioses: glyceraldehyde (aldotriose), and dihydroxyacetone (ketotriose).

Monosaccharides with four, five, six, and seven carbon atoms in their backbones are called, respectively, tetroses, pentoses, hexoses, and heptoses.

• These terms along with functional groups are used while naming monosaccharides.

For instance, glucose is an aldohexose, while fructose is a ketohexose.

an aldotriose

Glyceraldehyde, Dihydroxyacetone, a ketotriose

D-Glucose. an aldohexose

D-Fructose, a ketohexose

(b)

 $\left(e\right)$

STRUCTURAL FEATURES OF MONOSACCHARIDES

Stereoisomerism is an important characteristic of monosaccharides.

Stereoisomers are the compounds that have the same structural formulae but differ in their spatial configuration.

A carbon is said to be asymmetric when it is attached to four different atoms or groups.

•• The number of asymmetric carbon atoms (n) determines the possible isomers of a given compound which is equal to 2^n .

Glucose contains 4 asymmetric carbons, and thus has 16 isomers.

Glyceraldehyde (triose) is the simplest monosaccharide, with one asymmetric carbon atom.

It exists as two stereoisomers and has been chosen as the reference carbohydrate to represent the structure of all other carbohydrates.

D-and-L- forms of glucose compared with D- and L- glyceraldehydes (the reference carbohydrate) 14

The D and L isomers are mirror images of each other.

•• The spatial orientation of H and OH groups on the carbon atom $(C_5$ for glucose) that is adjacent to the terminal primary alcohol carbon determines whether the sugar is D- or Lisomer.

If the OH group is on the right side, the sugar is of D-series, and if on the left side, it belongs to L-series.

• The naturally occurring monosaccharides in the mammalian tissues are mostly of D-configuration.

•• The enzyme machinery of cells is specific to metabolize Dseries of monosaccharides.

Optical Activity of Carbohydrates

Optical activity is a characteristic feature of compounds with asymmetric carbon atom.

- When a beam of polarized light is passed through a solution of an optical isomer, it will be rotated either to the right or left.
- The term **dextrorotatory (d+)** and **levorotatory (l–)** are used for compounds that rotate the plane of polarized light to the right or to the left respectively.
- It may be noted that the D- and L-configurations of sugars are primarily based on the structure of glyceraldehyde, the optical activities however, may be different.

RACEMIC METHERE

If D- and L-isomers are present in equal concentration, it is known as racemic mixture or *dl mixture*.

Racemic mixture does not exhibit any optical activity, since the dextro- and levorotatory activities cancel each other.

In the medical practice, the term **dextrose** is used for **glucose in solution**. This is because of the dextrorotatory nature of glucose.

ERPIMERS

If two monosaccharides differ from each other in their configuration around a single specific carbon (other than anomeric) atom, they are referred to as epimers to each other.

Glucose and galactose are epimers with regard to carbon 4 $(C_4$ -epimers).

Glucose and mannose are epimers with regard to carbon 2 $(C_2$ -epimers).

• The interconversion of epimers (e.g. glucose to galactose and vice versa) is known as epimerization, and a group of enzymes called **epimerases** catalyze this reaction.

Structures of Epimers

(Glucose and galactose are C⁴ -epimers while glucose and mannose are C² -epimers) 19

ENANTIONERS

Enantiomers are a special type of stereoisomers that are mirror images of each other.

•• The two members are designated as D- and L-sugars.

Majority of the sugars in the higher animals (including man) are of D-type.

The term *diastereomers* is used to represent the stereoisomers that are **not mirror images** of one another.

Enantiomers (mirror images) of glucose

HEMIKROW FALLONDE HEMIKETALE

•The hydroxyl group of monosaccharides can react with its own aldehyde or keto functional group to form hemiacetal and hemiketal respectively, as seen below.

STRUCTURE OF GLUCOSE

 \cdot The aldehyde group of glucose at C_1 reacts with alcohol group at C_5 to form two types of cyclic hemiacetals, namely *α* and *β* as seen below.

•• The configuration of glucose is conveniently represented either by **Fischer formulae** or by **Haworth projection formulae**.

Mutarotation of glucose representing α and β anomers (A) Fischer projections (B) Haworth projections ²⁴

Pyranose and furanose

Haworth projection formulae are depicted by a sixmembered ring pyranose (based on pyran) or a fivemembered ring furanose (based on furan).

The cyclic forms of glucose are known as *α*-Dglucopyranose and *α*-D-glucofuranose.

 \cdot The α and β cyclic forms of D-glucose are known as **anomers**.

• They differ from each other in the configuration only around C_1 known as anomeric carbon (hemiacetal carbon).

In the case of α anomer, the OH group held by anomeric carbon is on the opposite side of the $CH₂OH$ group of the sugar ring.

The reverse is true for β anomer.

The anomers differ in certain physical and chemical properties.

Mutarotation is the change in the specific optical rotation representing the interconversion of α and β forms of D-glucose to an equilibrium mixture.

 \cdot The α and β anomers of glucose have different optical rotations.

The specific optical rotation of a freshly prepared glucose (α) anomer) solution in water is +112.2° , which gradually changes and attains an equilibrium with a constant value of +52.7°.

In the presence of alkali, the decrease in optical rotation is rapid.

•• The optical rotation of β -glucose is +18.7°.

Reactions of Monosaccharides

Tautomerization or enolization: The process of shifting a hydrogen atom from one carbon atom to another to produce *enediols.*

- **Reducing properties** (owing to the free aldehyde or keto group of anomeric carbon): Sugars are classified as reducing or nonreducing. Many laboratory tests such as Benedict's test, Fehling's test, Barfoed's test etc. are employed to identify the reducing action of sugars.
- **Oxidation**: Depending on the oxidizing agent used, the terminal aldehyde (or keto) or the terminal alcohol or both the groups may be oxidized.
- **Reduction**: When treated with reducing agents such as sodium amalgam, the aldehyde or keto group of monosaccharide is reduced to its corresponding alcohol.

Dehydration: When treated with concentrated sulfuric acid, monosaccharides undergo dehydration with an elimination of 3 water molecules.

- **Osazone formation**: When boiled with phenylhydrazine in acetic acid, reducing sugars form **osazones**. Reducing disaccharides also give osazones.
- **Formation of esters**: The alcoholic groups of monosaccharides may be esterified by non-enzymatic or enzymatic reactions.
- **Glycosides**: They are formed when the hemiacetal or hemiketal hydroxyl group of a carbohydrate reacts with a hydroxyl group of another carbohydrate or a non-carbohydrate (e.g. methyl alcohol, phenol, glycerol). The bond so formed is known as **glycosidic bond** and the non-carbohydrate moiety (when present) is referred to as **aglycone**.

Naming of glycosidic bond

The nomenclature of glycosidic bonds is based on the linkages between the carbon atoms and the status of the anomeric carbon (α or β).

- For instance, lactose—which is formed by a bond between C_1 of β-galactose and C₄ of glucose— is named as $\beta(1 \rightarrow 4)$ glycosidic bond.
- Monosaccharides are held together by glycosidic bonds to result in di-, oligo- or polysaccharides.

Oligosaccharides

Oligosaccharides (Greek: *oligo*-few) contain 2-10 monosaccharide units joined by glycosidic bonds.

- In cells, most oligosaccharides consisting of three or more units do not occur as free entities but are joined to non-sugar molecules (lipids or proteins) in glycoconjugates.
- Based on the number of monosaccharide units present, oligosaccharides are further subdivided to disaccharides, trisaccharides etc.

• The most abundant oligosaccharides are the disaccharides.

A disaccharide consists of two monosaccharide units (similar or dissimilar) held together by a glycosidic bond.

•• They are crystalline, water-soluble and sweet to taste.

◆ Disaccharides are of two types:

- **1. Reducing disaccharides** with free aldehyde or keto group e.g. maltose, lactose.
- **2. Non-reducing disaccharides** with no free aldehyde or keto group e.g. sucrose, trehalose.

Examples of Disaccharides

- **Maltose (malt sugar)**: It is composed of two α-D-glucose units held together by $\alpha(1 \rightarrow 4)$ glycosidic bond.
- **Sucrose (cane sugar)**: It is made up of α-D-glucose and β-Dfructose held together by $(\alpha_1 \rightarrow \beta_2)$ glycosidic bond.
- **Lactose (milk sugar)**: It is composed of β-D-galactose and β-Dglucose held together by $\beta(1 \rightarrow 4)$ glycosidic bond.
- **Lactulose**: It is a synthetic disaccharide containing galactose and fructose. It is neither digested nor absorbed in the intestine, but useful for the treatment of **hepatic encephalopathy**.

POLYSACCHARIDES

Polysaccharides (Greek: *poly*-many), also called **glycans**, are polymers of monosaccharide units (or their derivatives) with high molecular weight (up to a million).

- They are usually tasteless (non-sugars) and form colloids with water.
- In the body, they play structural roles and serve as a store of energy.
- They are linear as well as branched.
- **★Polysaccharides are of two types:**
	- **1. Homopolysaccharides**: They yield only a single type of monosaccharide on hydrolysis. E.g. starch, dextrans, inulin, glycogen, cellulose, chitin.
	- **2. Heteropolysaccharides**: They yield a mixture of a few monosaccharides or their derivatives on hydrolysis. E.g. mucopolysaccharides.

Starch: It is a homopolymer composed of D-glucose units held by α-glycosidic bonds. It is known as **glucosan** or **glucan**.

Starch consists of two polysaccharide components: water soluble **amylose** (15-20%) and a water insoluble **amylopectin** (80-85%).

Amylose is a long unbranched chain with 200–1,000 D-glucose units held by α (\rightarrow 4) glycosidic linkages.

Amylopectin is a branched chain with $\alpha(1 \rightarrow 6)$ glycosidic bonds at the branching points and α (1 \rightarrow 4) linkages everywhere else.

Dextrins: They are intermediate breakdown products of starch by the enzyme amylase or dilute acids.
Dextrans: They are polymers of glucose, produced by microorganisms. They are used as plasma volume expanders in transfusion, and chromatography (e.g. gel filtration).

- **Inulin**: It is a low molecular weight (around 5,000) polymer of fructose (i.e. fructosan) that occurs in dahlia bulbs, garlic, onion etc. It is not utilized by the body, but used for assessing kidney function through the measurement of glomerular filtration rate (GFR).
- **Glycogen**: It is the carbohydrate reserve in animals, hence often referred to as animal starch. It is present in high concentration in liver, followed by muscle, brain etc.
- The structure of glycogen is similar to that of amylopectin with more number of branches. Glucose is the repeating unit in glycogen joined together by $\alpha(1 \rightarrow 4)$ glycosidic bonds, and $\alpha(\rightarrow 6)$ glycosidic bonds at branching points.

Cellulose: It is the most abundant organic substance in plant kingdom, and a predominant constituent of plant cell wall. It is totally absent in animal body.

•• Cellulose is composed of β -D-glucose units linked by $\beta(1\rightarrow 4)$ glycosidic bonds.

- Certain ruminants and herbivorous animals contain microorganisms in the gut which produce enzymes that can cleave β-glycosidic bonds.
- Hydrolysis of cellulose yields a disaccharide **cellobiose**, followed by β-D-glucose.
- Though not digested, cellulose has great importance in human nutrition, as it is a major constituent of fiber. The functions of dietary fiber include decreasing the absorption of glucose and cholesterol from the intestine, besides increasing the bulk of feces.

Chitin: It is composed of N-acetyl-D-glucosamine units held together by $\beta(1\rightarrow 4)$ glycosidic bonds. It is a structural polysaccharide found in the exoskeleton of some invertebrates e.g. insects, crustaceans.

- **Mucopolysaccharides**: They are heteroglycans made up of repeating units of sugar derivatives, namely amino sugars and uronic acids. They are more commonly known as **glycosaminoglycans (GAG)**.
- Acetylated amino groups, besides sulfate and carboxyl groups are generally present in GAG structure.
- Some of the mucopolysaccharides are found in combination with proteins to form mucoproteins or mucoids or proteoglycans.

Mucoproteins may contain up to 95% carbohydrate and 5% protein.

- Mucopolysaccharides are essential components of tissue structure.
- The biologically important glycosaminoglycans along with their composition, distribution and functions are as seen below.

NUCLEIC ACIDS

Nucleic acids are the polymers of **nucleotides** held together by phosphate bridges.

Nucleic acids primarily store and transmit genetic information.

• There are two types of nucleic acids: ribonucleic acid (RNA) and deoxyribonucleic acid (DNA).

• The nucleotides in RNA contain ribose, while those of DNA contain deoxyribose.

NUCLEONIDES

Nucleotides are made up of a **nitrogenous base**, a **pentose sugar** and a **phosphate**.

Aside being the monomeric units of nucleic acids, nucleotides serve as structural components of some biologically important molecules such as coenzymes.

A "nucleotide" without phosphate is called **nucleoside**. Therefore, nucleotide is nucleoside + phosphate.

• The nitrogenous bases found in nucleotides are of two types: purines and pyrimidines.

General structure of nitrogen bases: (A) Purine (B) Pyrimidine

Nitrogenous bases

The purines in DNA and RNA are the same: **adenine** (A) and **guanine** (G).

- As for pyrimidines, only **cytosine** (C) is found in both DNA and RNA. The second pyrimidine base in DNA is **thymine** (T), while that of RNA is **uracil** (U).
- Aside the bases mentioned above, other minor and unusual bases are often found in DNA and RNA ,such as 5 methylcytosine, N⁴-acetylcytosine, N⁶-methyladenine, N⁶dimethyladenine, pseudouracil etc.
- The unusual bases in nucleic acids are said to help in the recognition of specific enzymes.

Structures of major purines and pyrimidines found in nucleic acids 46

Hypoxanthine, xanthine and uric acid are also bases of biological importance present in the free state in cells.

While hypoxanthine and xanthine are intermediates in purine synthesis, uric acid is the end product of purine degradation.

Plants contain pharmacologically important menthylated purines, such as caffeine (in coffee), theophylline (in tea) and theobromine (in cocoa).

Pentoses, which are five carbon monosaccharides, are found in the structure of nucleic acids.

While RNA contains D-ribose, DNA contains Ddeoxyribose.

NAMING OF NUCLEOTIDES

A pentose sugar and a base form a **nucleoside**.

If the sugar is ribose, the nucleoside is a **ribonucleosides**.

Adenosine, guanosine, cytidine and uridine are the ribonucleosides of adenine, guanine, cytosine and uracil respectively.

If the sugar is a **deoxyribose**, the resulting nucleoside is a **deoxyribonucleoside**.

A **mononucleotide** is formed when a single phosphate moiety is added to a nucleoside.

Adenosine monophosphate (AMP) contains adenine, ribose and phosphate.

• The prefix 'd' is used to show that the sugar is deoxyribose (e.g. dAMP).

The atoms in purine rings are numbered 1 to 9 and those of pyrimidine are 1 to 6.

• The carbons of sugars are numbered with an associated prime (΄). So the pentose carbons are numbered 1΄ to 5΄.

- \bullet Pentoses are bound to nitrogenous bases by β -N-glycosidic bonds.
- In purines, the N⁹ of the ring binds with $C_{1(1')}$ of a pentose sugar to form a covalent bond in the purine nucleoside.
- As for pyrimidines, the glycosidic linkage is between $N¹$ of the ring and C_1 of a pentose.
- The hydroxyl groups of adenosine are esterified with phosphates to give $5^{\text{-}}$ or $3^{\text{-}}$ monophosphates.

*5'-hydroxyl is the most commonly esterified, so 5 is usually omitted while writing nucleotide names.

• AMP represents adenosine 5-monophosphate, but for adenosine 3'-monophosphate, it is abbreviated as 3'-AMP.

When a nucleoside has only one phosphate moiety, it is a monophosphate (e.g. AMP, TMP).

When a second or third phosphate is added to the nucleoside, it is called a nucleoside diphosphate (e.g. ADP) or triphosphate (e.g. ATP), respectively.

The structures of adenosine 5 ΄ -monophosphate (AMP) and thymidine 5^{² - monophosphate (TMP) 52}

ANALOGS OF NUCLEONIDES

• The heterocyclic ring or sugar moiety can be altered to produce synthetic analogs of purines, pyrimidines, nucleosides and nucleotides, such as allopurinol, azathioprine, arabinosyladenine, 5-Fluorouracil etc.

Some of the synthetic analogs are of medical importance.

For example, allopurinol is used in the treatment of hyperuricemia and gout.

Azathioprine is used during transplantation for the suppression of immunological rejection.

DNA is a polymer of deoxyribonucleotides (deoxynucleotides).

- The monomeric units of DNA are deoxyadenylate (dAMP), deoxyguanylate (dGMP), deoxycytidylate (dCMP) and deoxythymidylate (dTMP).
- In DNA, deoxynucleotides are held together by 3΄ ,5΄ phosphodiester bridges.
- DNA (or RNA) structure is often represented in a short-hand form, in which the horizontal line indicates the carbon chain of sugar with base attached to C_1 .
- \bullet Near the middle of the horizontal line is C_3 phosphate linkage, while at the other end of the line is $C_{5'}$ phosphate linkage.

Structure of DNA and shorthand representation of oligonucleotides

CHARGAFF'S RULE

- In late 1940s, Erwin Chargaff observed that in the DNA hydrolysates of the different species he studied, there was an equal number of adenine and thymine residues $(A = T)$ and equal number of guanine and cytosine residues $(G = C)$.
- This is known as *Chargaff's rule of molar equivalence between the purines and pyrimidines in DNA structure*.
- The double helical structure of DNA derives its strength from Chargaff's rule
- In single-stranded DNA and RNA, Chargaff's rule is not obeyed.
- However, double-stranded RNA, which is the genetic material in certain viruses, satisfies Chargaff's rule.

DNA DOUBLAN HELIXAN

- James Watson and Francis Crick proposed the double helical structure of DNA in 1953, and won a Nobel Prize for it in 1962.
- The salient features of Watson-Crick model of DNA are as follows:
	- 1. The DNA is a right handed double helix, which consists of two polydeoxyribonucleotide chains (strands) twisted around each other on a common axis.
	- 2. The two strands are anti-parallel, i.e., one strand runs in the 5΄ to 3΄ direction while the other in 3΄ to 5΄ direction.
	- 3. The width (or diameter) of a double helix is $20 A^{\circ}$ (2 nm).

Each turn (pitch) of the helix is 34 A° (3.4 nm) with 10 pairs of nucleotides, each pair placed at a distance of about 3.4 A°.

- 5. Each strand of DNA has a hydrophilic deoxyribose phosphate backbone (3΄-5΄ phosphodiester bonds) on the outside (periphery) of the molecule while the hydrophobic bases are stacked inside (core).
- 6. The two polynucleotide chains are not identical but complementary to each other due to base pairing.
- 7. The two strands are held together by hydrogen bonds formed by complementary base pairs. The A-T pair has 2 hydrogen bonds while G-C pair has 3 hydrogen bonds. The G \equiv C is stronger by about 50% than $A=T$.
- The hydrogen bonds are formed between a purine and a pyrimidine only. If two purines face each other, they would not fit into the allowable space. And two pyrimidines would be too far to form hydrogen bonds. The only base arrangement possible in DNA structure, from spatial considerations is A-T, T-A, G-C and C-G.
- 9. The complementary base pairing in DNA helix proves Chargaff's rule. The content of adenine equals to that of thymine $(A = T)$ and guanine equals to that of cytosine $(G = C)$.
- 10. The genetic information resides on one of the two strands known as **template strand** or **sense strand**. The opposite strand is **antisense strand**. The double helix has (wide) major grooves and (narrow) minor grooves along the phosphodiester backbone. Proteins interact with DNA at these grooves, without disrupting the base pairs and double helix. 60

Watson-Crick model of DNA helix

Complementary base pairing in DNA helix

STRUCTURAL CONFORMATIONS OF DNA

DNA double helical structure exists in at least 6 different forms: A to E and Z.

• The B, A and Z forms are important.

The most predominant form under physiological conditions is the B-form of DNA double helix, described by Watson and Crick.

Each turn of the B-form has 10 base pairs spanning a distance of 3.4 nm, with a width of 2 nm.

The A-form is also a right-handed helix, which contains 11 base pairs per turn, and a tilting of the base pairs by 20° away from the central axis.

The Z-form (Z-DNA) is a left-handed helix and contains 12 base pairs per turn.

It is believed that transition between different helical forms of DNA plays a significant role in regulating gene expression.

Besides double helical structure, DNA also exists in certain unusual structures.

It is believed that such structures are important for molecular recognition of DNA by proteins and enzymes.

Comparison of structural features of different conformations of DNA double helix

Source: Satyanarayana and Chakrapani, 2013 **Source: Satyanarayana and Chakrapani**, 2013

Generally, adenine base-containing DNA tracts are rigid and straight.

- However, bent conformation of DNA occurs when Atracts are replaced by other bases or a collapse of the helix into the minor groove of A-tract.
- Bending in DNA structure has also been reported due to photochemical damage or mispairing of bases.
- Certain antitumor drugs (e.g. cisplatin) produce bent structure in DNA.
- Such changed structure can take up proteins that damage the DNA. 67

Triple-stranded DNA

Triple-stranded DNA formation may occur due to additional hydrogen bonds between the bases.

- Thus, a thymine can selectively form two Hoogsteen hydrogen bonds to the adenine of A-T pair to form T-A-T.
- Likewise, a protonated cytosine can also form two hydrogen bonds with guanine of G–C pairs that results in $C–G–C$.
- Triple-helical structure is less stable than double helix due to the fact that the three negatively charged backbone strands in triple helix results in an increased electrostatic repulsion.

Four-stranded DNA

Polynucleotides with very high contents of guanine can form a novel tetrameric structure called *G-quartets*.

- These structures are planar and are connected by Hoogsteen hydrogen bonds. Antiparallel four-stranded DNA structures, referred to as *G-tetraplexes* have also been reported.
- The ends of eukaryotic chromosomes namely **telomeres** are rich in guanine, and therefore form G-tetraplexes.
- In recent years, telomeres have become the targets for anticancer chemotherapies.
- G-tetraplexes have been implicated in the recombination of immunoglobulin genes, and in dimerization of double-stranded genomic RNA of the human immunodeficiency virus (HIV).

SIZE OF DNA MOLECULE

DNA molecules are huge in size.

On an average, a pair of B-DNA with a thickness of 0.34 nm has a molecular weight of 660 daltons.

- For the measurement of lengths, DNA double-stranded structure is considered, and expressed in the form of base pairs (bp).
- A kilobase pair (kb) is 10^3 bp, and a megabase pair (Mb) is 10^6 bp and a gigabase pair (Gb) is 10^9 bp.
- The length of DNA varies from species to species, and is usually expressed in terms of base pair composition and contour length.
- Contour length represents the total length of the genomic DNA in a cell.

The genomic DNA size is usually much larger than the size of the cell or nucleus containing it.

For instance, in humans, a 2-meter long DNA is packed compactly in a nucleus of about 10µm diameter.

• The genomic DNA may exist in linear or circular forms.

• Most DNAs in bacteria exist as closed circles.

Chromosomal DNAs in higher organisms are mostly linear.

Individual human chromosomes contain a single DNA molecule with variable sizes compactly packed.

Thus, the smallest chromosome contains 34 Mb while the largest one has 263 Mb.

DENATURATION OF DNA STRANDS

- The disruption of hydrogen bonds (by change in pH or increase in temperature) results in the separation of polynucleotide strands.
- This phenomenon of loss of helical structure of DNA is known as denaturation.
- However, denaturation does not break the phosphodiester bonds of the nucleotides.
- The loss of helical structure can be measured spectrophotometrically by increase in absorbance at 260 nm.
- The phenomenon of increase in the absorbance of purines and pyrimidines, following denaturation is referred to as **hyperchromicity**.
MENTING TEMPLERENTURE

Melting temperature (Tm) is the temperature at which half of the helical structure of DNA is lost.

- Since G-C base pairs are more stable (due to 3 hydrogen bonds) than A-T base pairs (2 hydrogen bonds), the Tm is greater for DNAs with higher G-C content.
- Therefore, the Tm is 65°C for 35% G-C content while it is 70°C for 50% G-C content.
- Formamide lowers Tm by destabilizing hydrogen bonds of base pairs, so it is used in recombinant DNA experiments.
- **Renaturation** (or reannealing) is the process in which the separated complementary DNA strands can form a double helix.

OCCURRENCE OF DNA IN THE CELL

•The double-stranded DNA helix in each chromosome has a length that is thousand times the diameter of the nucleus.

A 2-meter long DNA in humans is packed in a nucleus of about 10 µm diameter.

This is possible due to a compact and wonderful packaging, and organization of DNA inside the cell.

In prokaryotic cells, the DNA is organized as a single chromosome in the form of a double-stranded circle.

These bacterial chromosomes are packed in the form of nucleoids, by interaction with proteins and certain cations (polyamines).

In eukaryotic cells, the DNA double helix is wrapped around basic proteins called **histones**.

This association forms **nucleosome**, which is the basic unit of **chromatin.**

•• The chromatin fibres make up the chromosomes.

•The number of chromosomes in the nucleus of a cell is a characteristic feature of each specie.

Humans have a total of 46 chromosomes (i.e. 23 pairs).

STRUCTURE OF RNA

RNA is a polymer of ribonucleotides held together by 3´ ,5´ phosphodiester bridges.

- \cdot It is synthesized from DNA, and is primarily involved in the process of protein biosynthesis.
- RNA is usually a single-stranded polynucleotide.
- However, if complementary base pairs are in close proximity, the strand may fold at certain places to give a double-stranded structure.
- Due to the single-stranded nature of RNA, there is no specific relation between purine and pyrimidine contents.

So the Chargaff's rule of having equal guanine and cytosine content (as is the case in DNA) is not obeyed.

- Due to the presence of a hydroxyl group at 2' position, RNA can be hydrolyzed by alkali to form 2´ ,3´-cyclic diesters.
- On the other hand, DNA cannot be subjected to alkali hydrolysis due to the absence of the hydroxyl group at that position.
- Also, RNAs can be histologically identified by **orcinol colour reaction** due to the presence of ribose.

TYPES OF RNA

There are three major types of RNAs, with different cellular composition:

- i. Messenger RNA (mRNA) : 5–10%
- ii. Transfer RNA (tRNA) : 10–20%
- iii. Ribosomal RNA (rRNA) : 50–80%
- Aside the above, there are other types of RNAs present in the cells.
- They include heterogeneous nuclear RNA (hnRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA) and small cytoplasmic RNA (scRNA).

RNAs vary in their structure and function.

Messenger RNA (mRNA)

mRNA in eukaryotic cells is synthesized in the nucleus as heterogeneous nuclear RNA (hnRNA).

•The hnRNA is processed to liberate the functional mRNA, which enters the cytoplasm to participate in protein synthesis.

mRNA has high molecular weight with a short half-life.

Eukaryotic mRNA is capped at the 5´-terminal end by 7 methylguanosine triphosphate.

This cap may be responsible for the prevention of the hydrolysis of mRNA by 5´-exonucleases.

The cap may also be involved in the recognition of mRNA for protein synthesis.

The 3´-terminal end of mRNA contains a polymer of adenylate residues (20-250 nucleotides), which is known as **poly (A) tail**.

This tail may provide stability to mRNA, besides preventing it from the attack of 3´-exonucleases.

mRNA molecules often contain certain modified bases such as 6-methyladenylates in the internal structure.

Transfer RNA (tRNA)

Transfer RNA (soluble RNA) molecule contains 71-80 nucleotides (mostly 75) with a molecular weight of about 25,000.

At least 20 species of tRNAs exist, corresponding to 20 amino acids present in protein structure.

• The structure of tRNA (for alanine) was first elucidated by Holley.

•The structure of tRNA resembles that of a clover leaf. tRNA contains mainly four arms, each arm with a base paired stem.

Arms of tRNA

Acceptor arm: This arm is capped with a sequence CCA (5' to 3'). The amino acid is attached to the acceptor arm.

Anticodon arm: This arm, with the three specific nucleotide bases (anticodon), is responsible for the recognition of triplet codon of mRNA. The codon and anticodon are complementary to each other.

D arm: It is so named due to the presence of dihydrouridine.

- *** TwC arm**: This arm contains a sequence of T, pseudouridine (represented by psi , Ψ) and C.
- **Variable arm**: This arm is the most variable in tRNA. Based on this variability, tRNAs are classified into 2 categories:
	- **a) Class I tRNAs**: The most predominant (about 75%) form with 3-5 base pairs length.
	- **b) Class II tRNAs**: They contain 13-20 base pair long arm.

Structure of transfer RNA

Source: Satyanarayana and Chakrapani, 2013 **Source: Satyanarayana and Chakrapani**, 2013

Base pairs in tRNA

The structure of tRNA is maintained due to the complementary base pairing in the arms. The four arms with their respective base pairs are as follows:

 \cdot The acceptor arm -7 bp

 \cdot The T \cdot C arm $-$ 5 bp

 \cdot The anticodon arm – 5 bp

 \cdot The D arm -4 bp

Ribosomal RNA (rRNA)

•• The ribosomes are the factories of protein synthesis.

- The eukaryotic ribosomes are composed of two major nucleoprotein complexes–60S subunit and 40S subunit.
- The 60S subunit contains 28S rRNA, 5S rRNA and 5.8S rRNA, while the 40S subunit contains 18S rRNA.
- •The function of rRNAs in ribosomes is not clearly known.
- It is believed that they play a significant role in the binding of mRNA to ribosomes and protein synthesis.

Cellular RNAs and their function(s)

Source: Satyanarayana and Chakrapani, 2013 **Source: Satyanarayana and Chakrapani**, 2013

Catalytic RNAs (Ribozymes)

- In certain instances, the RNA component of a ribonucleoprotein (RNA in association with protein) is catalytically active.
- **Such RNAs are termed as ribozymes.**
- ◆ Ribonuclease P (RNase P) is a ribozyme containing protein and RNA component.
- It cleaves tRNA precursors to generate mature tRNA molecules.
- RNA molecules are known to adapt tertiary structure just like proteins (i.e. enzymes).
- The specific conformation of RNA may be responsible for its function as biocatalyst.
- It is believed that ribozymes (RNAs) were functioning as catalysts before the occurrence of protein enzymes, during the course of evolution.

Recombinant Ribozymes (rribozymes) It is now possible to design recombinant ribozymes that will cleave any RNA.

• These ribozymes are now being considered as therapeutic agents to cure diseases.

•• Theoretically it is possible to selectively degrade faulty RNAs by rribozymes.

• This way specific RNAs can be eliminated from the cell that will help to inhibit the disease process.

Biochemical reaction(s)

Source: Satyanarayana and Chakrapani, 2013 **Source: Satyanarayana and Chakrapani**, 2013

Water is undoubtedly the most important and most abundant substance in living systems.

It is the major body constituent, accounting for $55-70\%$ and 45-60% of the body weight of men and women respectively.

A normal man weighing 70 kg contains about 42 litres of water, which is distributed in intracellular (inside the cells 28L) and extracellular (outside the cells 14L) compartments, respectively known as intracellular fluid (ICF) and extracellular fluid (ECF).

The ECF is further divided into interstitial fluid (10.5L) and plasma (3.5L).

Distribution of water in an adult man, weighing 70 kg

Source: Satyanarayana and Chakrapani, 2013 ⁹²

Functions of Water

It provides the aqueous medium, which is essential for the various biochemical reactions in an organism.

It directly participates as a reactant in several metabolic reactions.

It serves as a vehicle for the transport of solutes.

It is closely associated with the regulation of body temperature.

WATER BALANCE

Like other substances and processes in the body, the water content of the body is remarkably regulated.

 \cdot In a healthy individual, this is achieved by balancing the daily water intake and water output.

Water is supplied to the body by exogenous and endogenous sources.

Exogenous water include ingested water and beverages, and water content of solid foods.

•• This largely depends on the social habits and climate.

Ingestion of water is mainly controlled by a *thirst centre* located in the hypothalamus.

Endogenous water is the metabolic water (300-350 ml/day) produced within the body.

It is derived from the oxidation of food.

It is estimated that 1g each of carbohydrate, protein and fat yield 0.6 ml, 0.4 ml and 1.1 ml of water respectively.

•• On an average, about 125 ml of water is generated for 1,000 Cal consumed by the body.

As for **water output**, there are four distinct routes through which the body loses/eliminates water: urine, skin, lungs and feces.

Urine: It is the major route for water loss from the body.

• The urine output in a healthy individual is about 1-2 l/day.

About 180 litres of water is filtered by the glomeruli into the renal tubules everyday, however, most of this is reabsorbed and only 1-2 litres is excreted as urine.

Water excretion by the kidney is tightly controlled by **vasopressin** also known as **antidiuretic hormone (ADH)** of the posterior pituitary gland.

Diabetes insipidus is a disorder characterized by the deficiency of ADH which results in an increased loss of water from the body.

Skin: It is another route for the loss of water (450 ml/day) as a result of perspiration.

It is an unregulated process by the body, and mostly depends on the atmospheric temperature and humidity.

It is estimated that for every 1° C rise in body temperature, about 15% increase is observed in the loss of water (through skin).

Lungs: Some amount of water (about 400 ml/day) is lost through the expired air during respiration.

Expired air usually saturated with water and expelled from the body.

Water loss through the lungs is increased in hot climates and/or when a person is suffering from fever.

The loss of water by perspiration (via skin) and respiration (via lungs) is collectively referred to as **insensible water loss**.

Feces: Most of the water entering the gastrointestinal tract is reabsorbed by the intestine, however, about 150 ml/day is lost through feces in a healthy individual.

Fecal loss of water is tremendously increased in diarrhea.

•• The water balance of the body is regulated predominantly by controlling the urine output.

This happens after an obligatory water loss via skin, lungs and feces.

The abnormalities associated with water balance include **dehydration** and **overhydration.**

Water balance in the body (values are variable)

Source: Satyanarayana and Chakrapani, 2013 **1999** 99

DEHYDRATION

Dehydration is a condition characterized by water depletion in the body due to insufficient intake or excessive water loss or both.

Dehydration is generally classified into two types: due to loss of water alone, and due to deprivation of water and electrolytes.

Dehydration may occur as a result of diarrhea, vomiting, excessive sweating, fluid loss in burns, adrenocortical dysfunction, kidney diseases (e.g. renal insufficiency), deficiency of ADH (diabetes insipidus) etc.

There are three degrees of dehydration: mild, moderate and severe.

CHARACTERISTIC FEATURES OF DEHYDRATION

- 1. The volume of the extracellular fluid (e.g. plasma) is decreased with a concomitant rise in electrolyte concentration and osmotic pressure.
- 2. Water is drawn from the intracellular fluid that results in shrunken cells and disturbed metabolism e.g. increased protein breakdown.
- 3. ADH secretion is increased, causing an increase in water retention in the body and consequently urine volume is very low.
- 4. Plasma protein and blood urea concentrations are increased.
- 5. Water depletion is often accompanied by a loss of electrolytes from the body $(Na^+, K^+$ etc.).
- 6. The principal clinical symptoms of severe dehydration include increased pulse rate, low blood pressure, sunken eyeballs, decreased skin turgor, lethargy, confusion and coma.

The treatment of choice for dehydration is intake of plenty of water.

In patients who cannot take orally, water can be administered intravenously in an isotonic solution (usually 5% glucose).

If the dehydration is accompanied by loss of electrolytes, the same should be administered by oral or intravenous routes.

This is usually done by carefully monitoring the water and electrolyte status of the body.

Overhydration

- Overhydration or water intoxication is caused by excessive retention of water in the body.
- It may occur due to excessive intake of large volumes of salt-free fluids, renal failure, overproduction of ADH etc.
- Overhydration is observed after major trauma or operation, lung infections etc.
- Water intoxication is associated with dilution of ECF and ICF with a decrease in osmolality.
- The clinical symptoms include headache, lethargy and convulsions.
- \triangle The treatment advocated is stoppage of water intake and administration of hypertonic saline.

Hydrophilic interactions

- ◆ Hydrophilic interactions are interactions between polar molecules (hydrophiles) and water.
- ◆ Hydrophiles can easily form a hydrogen bond with water, thereby getting dissolved in water.
- In contrast, hydrophobes are not attracted to water, and therefore do not dissolve in it.
- Amphiphilic molecules on the other hand, have both hydrophilic and hydrophobic properties, e.g. the lipids that comprise the cell membrane, and soap.
- ◆ Hydrophilic interactions are possible with polar chemical groups.
- ◆ Hydrophilic molecules are typically charge-polarized and capable of hydrogen bonding.

Each hydrogen atom in water $(H₂O)$ bears a partial positive ϵ harge (δ^+) and the oxygen atom bears a partial negative charge equal to the sum of the two partial positives (2δ) .

• Consequently, there is an electrostatic attraction between the oxygen atom of one water molecule and the hydrogen of another, called a **hydrogen bond**.

Hydrogen bonds are weak, with a bond dissociation energy (the energy required to break a bond) of about 23 kJ/mol in liquid water compared with 470 kJ/mol for the covalent $O-H$ bond in water or 348 kJ/mol for a covalent C-C bond.

At any given time, most of the molecules in liquid water are engaged in hydrogen bonding, but the lifetime of each hydrogen bond is just 1 to 20 picoseconds (1 ps = 10^{-12} s).

Upon breakage of one hydrogen bond, another hydrogen bond forms, with the same partner or a new one, within 0.1 ps.

Hydrogen bonding confers water with its unusual properties of higher melting point, boiling point, and heat of vaporization when compared with most other common solvents.

Ionization and ion product of water

…refer to class note.

Bibliography and suggested FURTHER READING

- 1. Satyanarayana U., Chakrapani U. (2006): Biochemistry, 3rd ed. Kolkata: Arunabha Sen Books and Allied Ltd.
- 2. David Nelson and Michael Cox (2016). Lehninger Principles of Biochemistry, 7th ed. McGrawHill education.
- 3. Victor Rodwell, David Bender, Kathleen Botham, Peter Kennelly, and Anthony Weil (2018). Harper's Illustrated Biochemistry. McGrawHill education lange.
- 4. Reginald Garrett and Charles Grisham (2010). Biochemistry. Brooks/Cole, Cengage Learning.