# MACROMOLECULES | BCH 101

# BIOMOLECULES

This refers to biological molecules

Hydrogen, oxygen, carbon, and nitrogen constitute more than 99% of the atoms in the human body, with most of the H and O occurring as $H_2O$.

It is the ability of these elements to form covalent bonds by electron-pair sharing. They form the strong covalent bonds.

Two other covalent bond-forming elements, phosphorus and sulfur, also play important roles in biomolecules.

The strength of covalent bonds is inversely proportional to the atomic weights of the atoms involved, H, C, N, and O form the strongest covalent bonds.

# CARBON AS THE BACKBONE

All biomolecules contain carbon. The prevalence of C is due to its unparalleled versatility in forming stable covalent bonds.

Carbon can form as many as four such bonds by sharing each of the four electrons in its outer shell with electrons contributed by other atoms.

Hydrogen can form one such bond by contributing its single electron to formation of an electron pair.

Oxygen, with two unpaired electrons in its outer shell, can participate in two covalent bonds.

Nitrogen, which has three unshared electrons, can form three such covalent bonds.

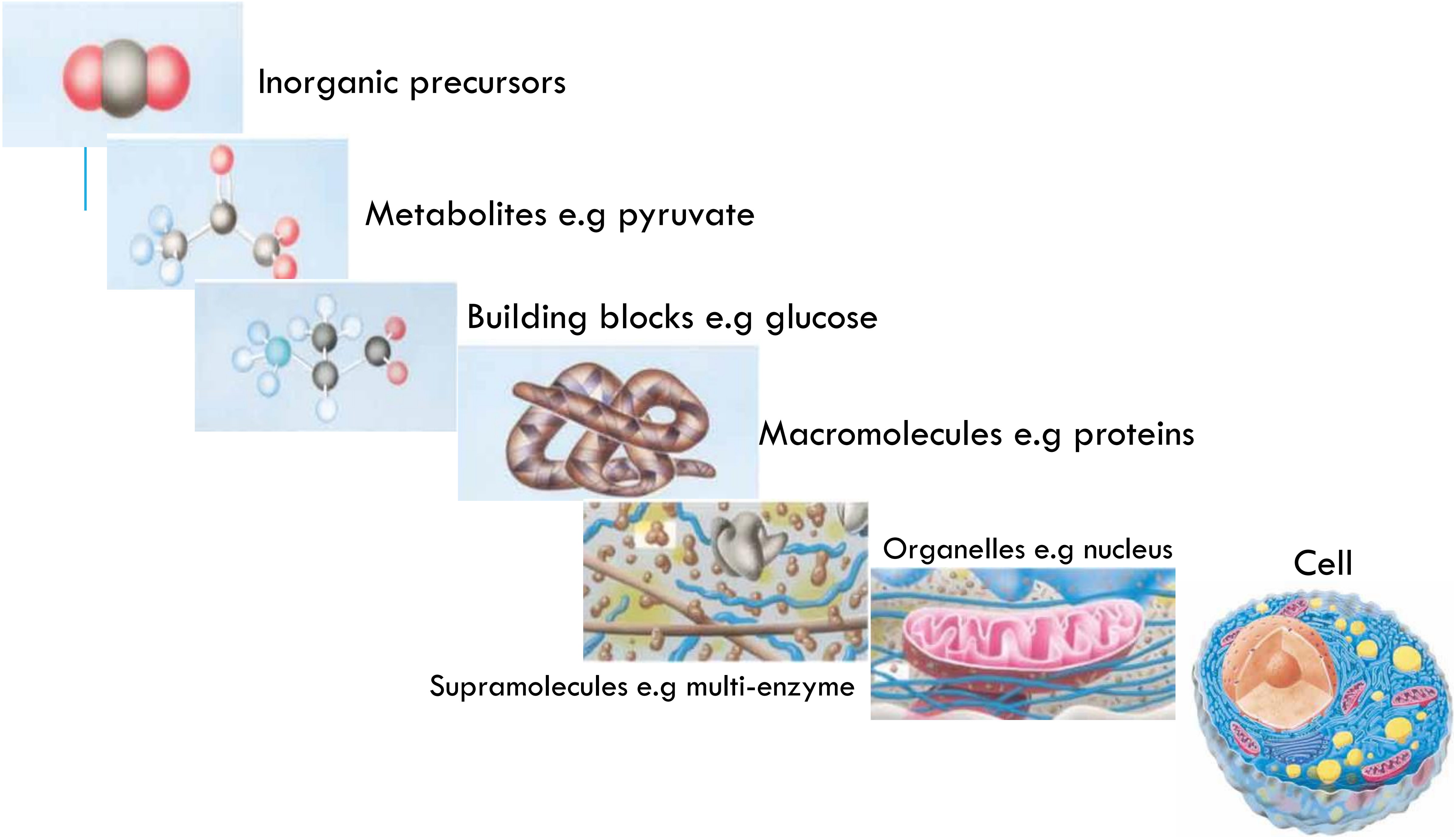
# BIOMOLECULAR HIERARCHY

The major precursors for the formation of biomolecules are water, carbon dioxide, and three inorganic nitrogen compounds—ammonium ($NH_4$ ), nitrate ($NO_3$), and dinitrogen ($N_2$).

These biomolecules are converted to simple organic compounds called **metabolites** that are intermediates in cellular energy transformation and in the biosynthesis of various sets of building blocks: amino acids, sugars, nucleotides, fatty acids.

Covalent linkage of these building blocks form **macromolecules** e.g Polysaccharide, proteins, lipids, polynucleotides (RNA and DNA).

Interactions among macromolecules lead to the next level of structural organization called **supramolecular complexes.**



Inorganic precursors
Metabolites e.g pyruvate
Building blocks e.g glucose
Macromolecules e.g proteins
Supramolecules e.g multi-enzyme
Organelles e.g nucleus
Cell

# CARBOHYDRATES

Carbohydrates are the single most abundant class of organic molecules found in nature.

The name *carbohydrate* arises from the basic molecular formula $(CH_2O)n$, which can be rewritten $(C \,.\, H_2O)n$ to show that these substances are hydrates of carbon, where $n$  3 or more.

Carbohydrates constitute a versatile class of molecules. Energy from the sun captured by green plants, algae, and some bacteria during photosynthesis is stored in the form of carbohydrates.

In turn, carbohydrates are the metabolic precursors of virtually all other biomolecules.

Breakdown of carbohydrates provides the energy that sustains animal life.

Carbohydrates are covalently linked with a variety of other molecules.

Carbohydrates linked to lipid molecules, or **glycolipids**, are common components of biological membranes.

Proteins that have covalently linked carbohydrates are called **glycoproteins.**

These two classes of biomolecules, together called **glycoconjugates**, are important components of cell walls and extracellular structures in plants, animals, and bacteria.

In addition to the structural roles such molecules play, they also serve in a variety of processes involving *recognition* between cell types.

Recognition events are important in normal cell growth, fertilization, transformation of cells, and other processes.

# FEATURES OF CARBOHYDRATES

(1) the existence of at least one and or more asymmetric centers

(2) the ability to exist either in linear or ring structures

(3) the capacity to form polymeric structures via *glycosidic* bonds

(4) the potential to form multiple hydrogen bonds with water or other molecules in their environment.

Carbohydrates are polyhydroxy aldehydes or ketones, or substances that yield such compounds on hydrolysis.

# CLASSES

**Monosaccharides,** or simple sugars, consist of a single polyhydroxy aldehyde or ketone unit. They cannot be hydrolysed into simpler carbohydrates e.g glucose, fructose.

**Oligosaccharides** consist of short chains of monosaccharide units, or residues, joined by characteristic linkages called glycosidic bonds. The most abundant are the **disaccharides,** with two monosaccharide units e.g maltose, sucrose.

The **polysaccharides** are sugar polymers containing more than 20 or so monosaccharide units, and some have hundreds or thousands of units. They are polymers of simple sugars or their derivatives e.g cellulose.

# PROTEIN

Proteins are the most abundant biological macromolecules, occurring in all cells and all parts of cells.

Proteins also occur in great variety; thousands of different kinds, ranging in size from relatively small peptides to huge polymers.

They also exhibit enormous diversity of biological function and are the most important final products of the information pathways.

All proteins, are constructed from the same ubiquitous set of 20 amino acids, covalently linked in characteristic linear sequences.

Cells can produce proteins with very different properties and activities by joining the same 20 amino acids in different combinations and sequences.

# AMINO ACIDS

Proteins are polymers of amino acids, with each **amino acid residue** joined to its neighbour by a specific type of covalent bond.

Proteins can be broken down (hydrolyzed) to their constituent amino acids by a variety of methods.

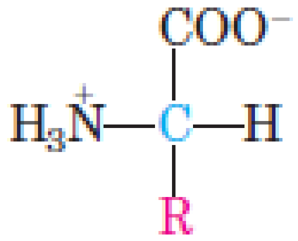
The first to be discovered was asparagine, in 1806. The last of the 20 to be found, threonine, was not identified until 1938.

Asparagine was first found in asparagus, and glutamate in wheat gluten; tyrosine was first isolated from cheese (its name is derived from the Greek *tyros,* "cheese"); and glycine (Greek *glykos,* "sweet") was so named because of its sweet taste.

# STRUCTURES OF AMINO ACIDS

The structure is a tetrahedral alpha carbon (C), which is covalently linked to both the amino group and the carboxyl group.

Also bonded to this carbon is a hydrogen and a variable side chain. It is the side chain, the so-called R group, that gives each amino acid its identity.



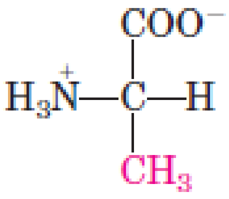

In addition to these 20 amino acids there are many less common ones.

Because the resulting amino acid contains one positive and one negative charge, it is a neutral molecule called a **zwitterion.**

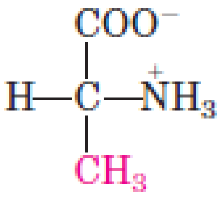
# STRUCTURES (CONTD)

Amino acids are also **chiral** molecules because the α-carbon is bonded to four different molecules (except in glycine).

With this four different molecules attached to it, the α-carbon is said to be **asymmetric.**



L-Alanine



D-Alanine

The two possible configurations for the α-carbon constitute non-identical mirror image isomers or **enantiomers.**

All molecules with a chiral center are also **optically active—**that is, they rotate plane-polarized light i.e laevo or dextro.

# ABBREVIATIONS

| | | |
|---|---|---|
| **Nonpolar, aliphatic R groups** | | |
| Glycine | Gly | G |
| Alanine | Ala | A |
| Proline | Pro | P |
| Valine | Val | V |
| Leucine | Leu | L |
| Isoleucine | Ile | I |
| Methionine | Met | M |
| **Aromatic R groups** | | |
| Phenylalanine | Phe | F |
| Tyrosine | Tyr | Y |
| Tryptophan | Trp | W |

| | | |
|---|---|---|
| **Polar, uncharged R groups** | | |
| Serine | Ser | S |
| Threonine | Thr | T |
| Cysteine | Cys | C |
| Asparagine | Asn | N |
| Glutamine | Gln | Q |
| **Positively charged R groups** | | |
| Lysine | Lys | K |
| Histidine | His | H |
| Arginine | Arg | R |
| **Negatively charged R groups** | | |
| Aspartate | Asp | D |
| Glutamate | Glu | E |

# PROTEIN CLASSIFICATION

**Simple proteins** contain only amino acid residues and no other chemical constituents.

**Conjugated proteins** contain permanently associated chemical components in addition to the amino acid.

The non–amino acid part of a conjugated protein is usually called its **prosthetic group.**

Conjugated proteins are classified on the basis of the chemical nature of their prosthetic groups.

**Lipoproteins** contain lipids e.g lipoproteins

**Glycoproteins** contain carbohydrates e.g Immunoglobulin G

**Metalloproteins** contain specific metals e.g Ferritin

# LIPIDS

The biological functions of the lipids are as diverse as their chemistry.

Fats and oils are the principal stored forms of energy in many organisms.

Phospholipids and sterols are major structural elements of biological membranes.

Other lipids, although present in relatively small quantities, play crucial roles as enzyme cofactors, electron carriers

The fats and oils used almost universally as stored forms of energy in living organisms are derivatives of **fatty acids.**

There are two major classes of lipids: Storage (e.g triacylglycerol) and structural lipids (e.g phospholipids)

# FATTY ACIDS

Fatty acids are carboxylic acids with hydrocarbon chains ranging from 4 to 36 carbons long (C4 to C36).

In some fatty acids, this chain is unbranched and fully saturated (contains no double bonds); in others the chain contains one or more double bonds.

The most commonly occurring fatty acids have even numbers of carbon atoms in an unbranched chain of 12 to 24 carbons.

The physical properties of the fatty acids, and of compounds that contain them, are largely determined by the length and degree of unsaturation of the hydrocarbon chain.

The nonpolar hydrocarbon chain accounts for the poor solubility of fatty acids in water.

# EXAMPLES

| *Carbon skeleton* | *Structure** | *Systematic name*† | *Common name (derivation)* | *Melting point (°C)* |
|---|---|---|---|---|
| 12:0 | $CH_3(CH_2)_{10}COOH$ | *n*-Dodecanoic acid | Lauric acid (Latin *laurus*, "laurel plant") | 44.2 |
| 14:0 | $CH_3(CH_2)_{12}COOH$ | *n*-Tetradecanoic acid | Myristic acid (Latin *Myristica*, nutmeg genus) | 53.9 |
| 16:0 | $CH_3(CH_2)_{14}COOH$ | *n*-Hexadecanoic acid | Palmitic acid (Latin *palma*, "palm tree") | 63.1 |
| 18:0 | $CH_3(CH_2)_{16}COOH$ | *n*-Octadecanoic acid | Stearic acid (Greek *stear*, "hard fat") | 69.6 |
| 20:0 | $CH_3(CH_2)_{18}COOH$ | *n*-Eicosanoic acid | Arachidic acid (Latin *Arachis*, legume genus) | 76.5 |

# TRIACYLGLYCEROL

The simplest lipids constructed from fatty acids are the **triacylglycerols**, also referred to as triglycerides, fats, or neutral fats.

Triacylglycerols are composed of three fatty acids each in ester linkage with a single glycerol.

Those containing the same kind of fatty acid in all three positions are called simple triacylglycerols and are named after the fatty acid they contain. E.g tristearin, tripalmitin, and triolein.

Most naturally occurring triacylglycerols are mixed; they contain two or more different fatty acids.

Triacylglycerol provides stored energy. This is why many foods contain triacylglycerol.

# NUCLEIC ACIDS

**Nucleotides** and **nucleic acids** are biological molecules that possess heterocyclic nitrogenous bases as principal components of their structure.

The biochemical roles of nucleotides are numerous; they participate as essential intermediates in virtually all aspects of cellular metabolism.

Nucleic acids, on the other hands, are the elements of heredity and the agents of genetic information transfer.

The two basic kinds of nucleic acids are **deoxyribonucleic acid (DNA**) and **ribonucleic acid (RNA).**

Complete hydrolysis of nucleic acids liberates nitrogenous bases, a five-carbon sugar, and phosphoric acid in equal amounts.

# NITROGENOUS BASES

The bases of nucleotides and nucleic acids are derivatives of either **pyrimidine** or **purine.**

The common naturally occurring pyrimidines are **cytosine, uracil**, and **thymine** (5-methyluracil).

Cytosine and thymine are the pyrimidines typically found in DNA, whereas cytosine and uracil are common in RNA.

Various pyrimidine derivatives, such as dihydrouracil, are present as minor constituents in certain RNA molecules.

**Adenine** (6-amino purine) and **guanine** (2-amino-6-oxy purine), the two common purines, are found in both DNA and RNA.

Other naturally occurring purine derivatives include **hypoxanthine, xanthine,** and **uric acid**

# PENTOSES

Pentoses are five-carbon sugars.

RNA containns the pentose D-ribose, while 2-deoxy-D-ribose is found in DNA.

**Nucleosides** are compounds formed when a base is linked to a sugar via a **glycosidic bond**

Nucleosides are named by adding the ending *–idine* to the root name of a pyrimidine or *-osine* to the root name of a purine.

The common nucleosides are thus **cytidine, uridine, thymidine, adenosine,** and **guanosine.**

Nucleosides are much more water-soluble than the free bases because of the hydrophilicity of the sugar moiety.

# NUCLEOTIDES

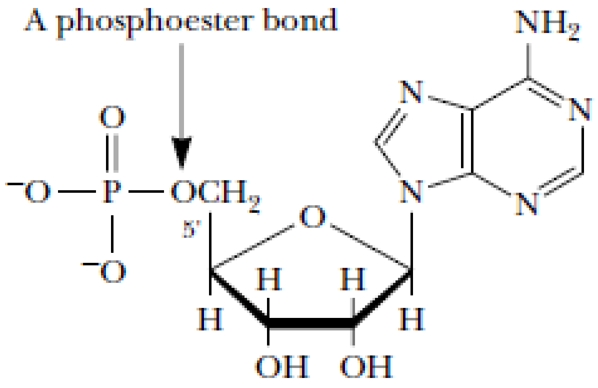
A **nucleotide** results when phosphoric acid is esterified to a sugar -OH group of a nucleoside.

The vast majority of monomeric nucleotides in the cell are **ribonucleotides** having 5-phosphate groups.

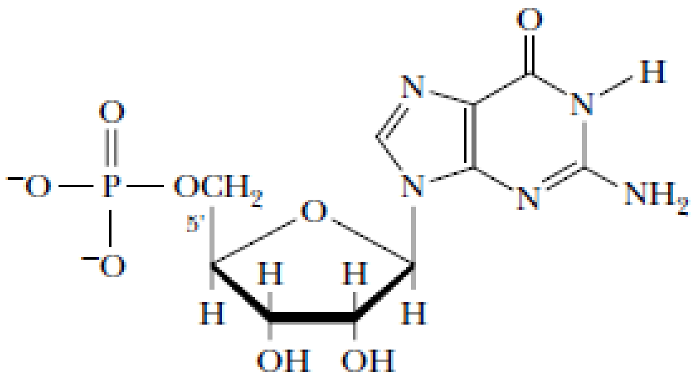
Example are **adenosine 5-monophosphate, guanosine 5-monophosphate, cytidine 5-monophosphate**, and **uridine 5- monophosphate.**

These compounds are more often referred to by their abbreviations: **5-AMP, 5-GMP, 5-CMP,** and **5-UMP,** or even more simply as **AMP, GMP, CMP,** and **UMP.**

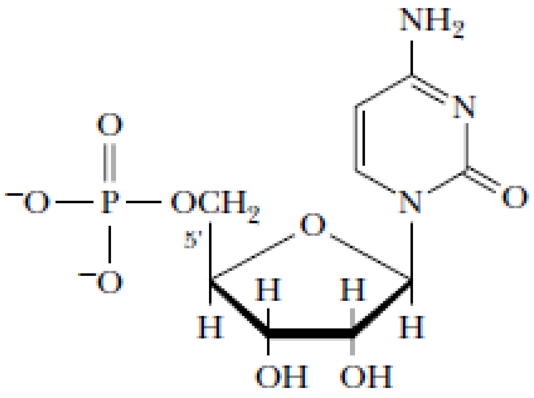
# STRUCTURES



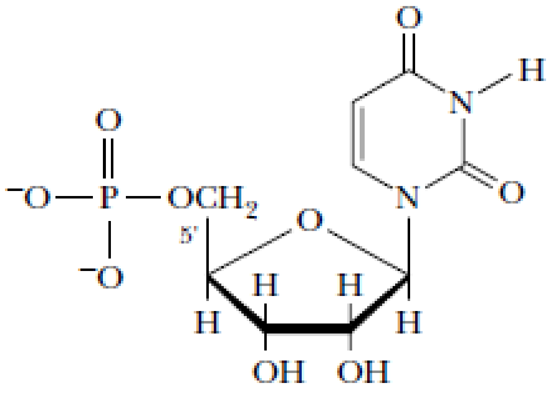

**Adenosine 5'-monophosphate**
**(or AMP or adenylic acid)**



**Guanosine 5'-monophosphate**
**(or GMP or guanylic acid)**



**Cytidine 5'-monophosphate**
**(or CMP or cytidylic acid)**



**Uridine 5'-monophosphate**
**(or UMP or uridylic acid)**