

# Biomolecules: Proteins and Nucleic acids

BIO 101

Adu, Oluwatosin B.

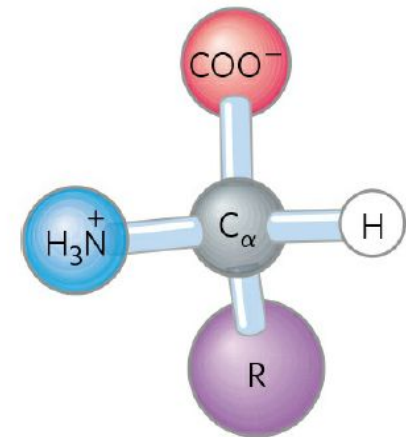
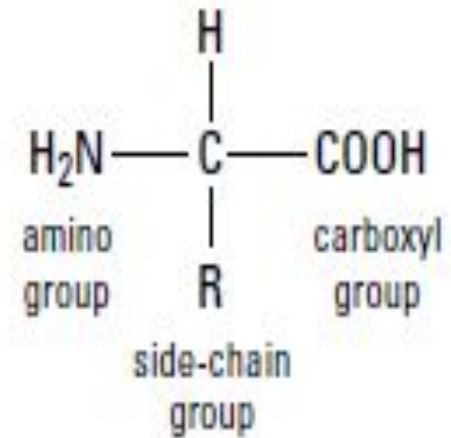
C19-

# PROTEIN

- A protein is a large organic molecule constructed of a chain of amino acids. The word protein, derived from the Greek *proteios*, loosely means “holding first place”
- They orchestrate the business of life by controlling multiple bioprocesses including metabolism, cell growth, and neurotransmission.
- Proteins also provide structure and can act as energy source, the main reason they are so important is their role as enzymes—enabling chemical reactions that are critical to life.
- Proteins are **polymers of amino acids**, with each amino acid residue joined to its neighbor by a specific type of covalent bond called **peptide bond**.
- Twenty different amino acids are commonly found in proteins.

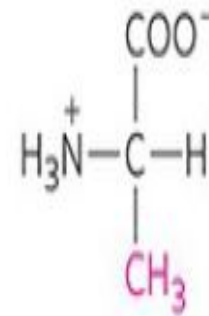
# Amino Acids

- All 20 of the common amino acids are  **$\alpha$ -amino acids**.
- They have a carboxyl group and an amino group bonded to the same carbon atom (the  $\alpha$  carbon).
- They differ from each other in their side chains, or **R groups**, which vary in structure, size, and electric charge, and which influence the solubility of the amino acids in water.
- In addition to these 20 amino acids there are many less common ones.
- The common amino acids of proteins have been assigned three letter abbreviations and one-letter symbols, which are used as shorthand to indicate the composition and sequence of amino acids polymerized in proteins.

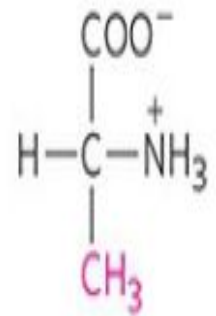


General structure of an amino acid

- For all the common amino acids except glycine, the  $\alpha$  carbon is bonded to four different groups: a carboxyl group, an amino group, an R group, and a hydrogen atom (in glycine, the R group is another hydrogen atom).
- The  $\alpha$ -carbon atom is thus a **chiral center**. With two possible stereoisomers- D and L.
- The amino acid residues in protein molecules are exclusively L stereoisomers.
- D-Amino acid residues have been found in only a few, generally small peptides, including some peptides of bacterial cell walls and certain peptide antibiotics.



L-Alanine



D-Alanine

**The two stereoisomers of alanine, L- and D-alanine,**

- Amino acids in solution at neutral pH exist mainly as **zwitterions** rather than as unionized molecules.
- Zwitterions have both a positive and a negative charge.
- The amino group carries a positive charge ( $\text{NH}_3^+$ ) and the carboxyl group carries a negative charge ( $\text{CO}_2^-$ ).
- The presence of both charges means that the molecule can act as both an acid and a base, that is, **amphoteric**.

# Classification of amino acids

- Amino acids can be classified on the following bases:

1. **Nature of the R- group:** Amino acids can be grouped into five main classes based on the properties of their R groups, particularly their polarity, or tendency to interact with water at biological pH. These are
  - Nonpolar, aliphatic R groups e.g., Gly, Ala, Val
  - Aromatic R groups e.g., Phe, Tyr, Trp
  - Polar, uncharged R groups e.g., Ser, Thr, Cys
  - Positively charged R groups (Acidic) e.g., Lys, His , Arg,
  - Negatively charged R groups (Basic) eg., Asp and Glu

**2. Fate of the carbon skeleton:** Amino acids can be classified into two groups based on the metabolic fate of their carbon skeleton. These are

- **Glucogenic** e.g., Phe, Tyr, Ile
- **Ketogenic** e.g., Leu, Lys

Some amino acids are both glucogenic and ketogenic e.g., Trp, Phe and Tyr

**3. Nutritional need:** Based on the ability of mammals to synthesise them, amino can be classified as

- **Essential** amino acids cannot be synthesized in the body and so must be obtained from food e.g., His, Ile
- **Nonessential** amino acids can be synthesized by the body e.g., Pro, Ser

10 amino acids are essential for children but only 8 are essential for adults

# Assignment

- List all the 20 amino acids with their three letter and 1 letter notations and classify them based on all the 3 criteria stated above.

# Proteins

- Two amino acid molecules can be covalently joined through a substituted amide linkage, termed a **peptide bond**, to yield a dipeptide.
- Three amino acids can be joined by two peptide bonds to form a tripeptide; similarly, four amino acids can be linked to form a tetrapeptide, five to form a pentapeptide, and so forth.
- When a few amino acids are joined in this fashion, the structure is called an **oligopeptide**.
- When many amino acids are joined, the product is called a **polypeptide**. Proteins may have thousands of amino acid residues.
- Molecules referred to as **polypeptides** generally have molecular weights below 10,000, and those called **proteins** have higher molecular weights.
- In a peptide, the amino acid residue at the end with a free  $\alpha$ -amino group is the **amino-terminal** (or N-terminal) residue; the residue at the other end, which has a free carboxyl group, is the **carboxyl-terminal** (C-terminal) residue.

# Conjugated proteins

- Many proteins, e.g., chymotrypsin, contain only amino acid residues and no other chemical constituents; these are **simple proteins**.
- Some proteins contain associated chemical components in addition to amino acids; these are **conjugated proteins**.
- The non–amino acid part of a conjugated protein is its **prosthetic group**.
- Conjugated proteins are classified on the basis of the chemical nature of their prosthetic groups e.g.  
**glycoproteins** = protein + sugar
- prosthetic group plays an important role in the protein's biological function.

## Conjugated Proteins

Class	Prosthetic group	Example
Lipoproteins	Lipids	$\beta_1$ -Lipoprotein of blood
Glycoproteins	Carbohydrates	Immunoglobulin G
Phosphoproteins	Phosphate groups	Casein of milk
Hemoproteins	Heme (iron porphyrin)	Hemoglobin
Flavoproteins	Flavin nucleotides	Succinate dehydrogenase
Metalloproteins	Iron Zinc Calcium Molybdenum Copper	Ferritin Alcohol dehydrogenase Calmodulin Dinitrogenase Plastocyanin

# Structural organization of proteins

- The spatial arrangement of atoms in a protein or any part of a protein is called its **conformation or shape**. This conformation is essential for the function of proteins.
- The shape of a protein develops in stages. Protein structures generally are described at four levels: **primary, secondary, tertiary, and quaternary**.
- **Primary structure** is simply the two-dimensional **linear sequence** of amino acids in the peptide chain. The primary structure is a long string of amino acids forming polypeptide chains stabilised by **covalent bond**
- **Secondary structure** is formed as the protein twists in accordance with chemical forces within the primary chain. The hydrogen in the amine groups and the oxygen in the carboxyl groups of the backbone form hydrogen bonds that stabilize the secondary structures.

- Secondary structure commonly takes one of two forms.
  1. **left-handed helix** that is developed as the molecule twines around itself. This form is called an  **$\alpha$ -helix**. **Globular proteins** typically contain the  $\alpha$ -helix form. Hormones are globular proteins.
  2. The other secondary form is a crimped shape called a  **$\beta$ -pleated sheet**. A  $\beta$ -pleated sheet looks a little like a piece of tin.  $\beta$ -Sheets are usually found in **structural proteins**, such as silk and the  $\beta$ -keratin proteins found in claws, beaks, and shells of turtles.

- **Tertiary structure** is the final 3-Dimensional shape of one folded amino acid chain.
  - The final shape of any polypeptide chain is unique and will have specific areas that are necessary for the function of the protein.
  - The tertiary structure is stabilized by **hydrogen bond, ionic bond, hydrophobic interactions** and **disulfide bonds**
- A **quaternary structure** may be formed from the association of two or more peptide strands.
  - Some proteins are large and complex, consisting of many polypeptide chains. These proteins have quaternary structure.
  - Hemoglobin (with four polypeptide chains) and (with three chains) are proteins with quaternary structure.
  - The formation of **disulfide bonds** is important in the quaternary structure of many proteins.

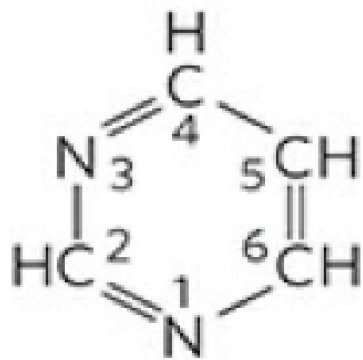
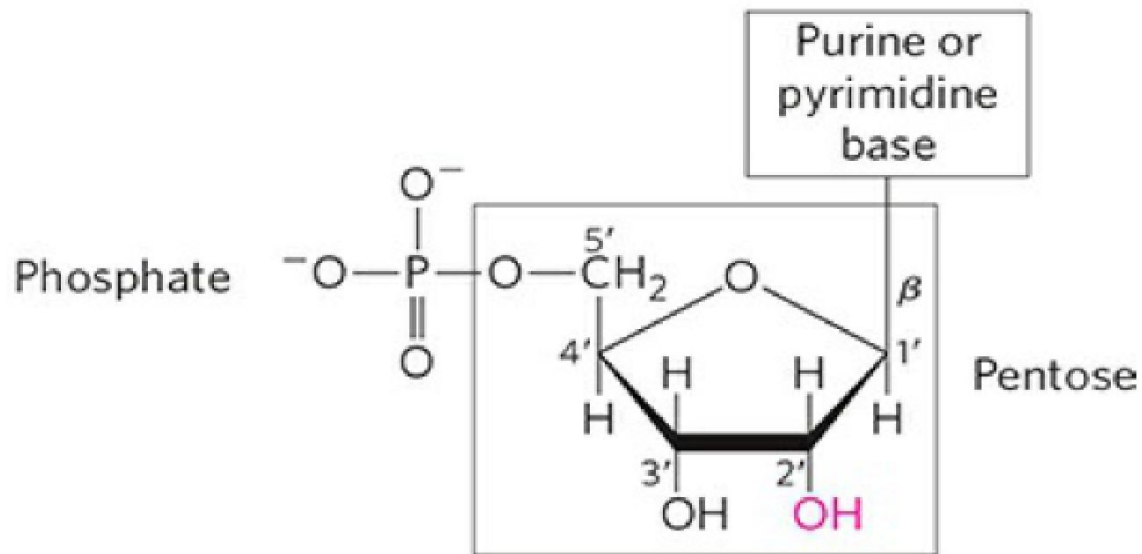
# Functions of Proteins

- Some of the important functions of proteins include:
  - ✓ **Proteins are enzymes.** Enzymes make chemical reactions happen faster.
  - ✓ **Proteins reinforce structures.** Proteins are part of the structure of plasma membranes, and cytoskeletal proteins. Proteins in the extracellular matrix support cells e.g. collagen.
  - ✓ **Proteins transport materials into and out of the cell.** Proteins in the plasma membrane help molecules enter and exit the cell.
  - ✓ **Proteins are involved in cellular identity.** Glycoproteins on cell surfaces act as a marker that identifies the cell type.

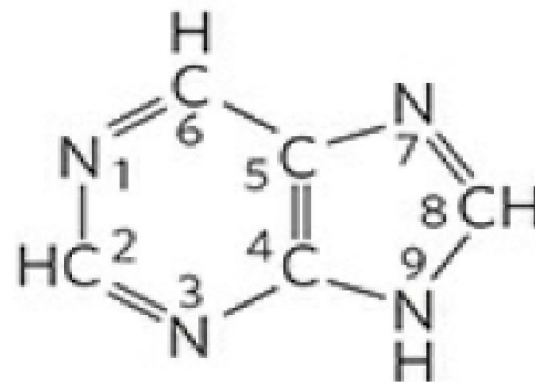
- ✓ **Proteins help cells move.** Cytoskeletal proteins power the movement of flagella and allow cells to crawl.
- ✓ **Proteins help cells communicate.** Proteins are receptors for signals. Also, some signaling molecules, like insulin, are proteins.
- ✓ **Proteins organize molecules within the cell.** Chaperone proteins assist folding of new proteins and guide proteins to their proper locations within the cell.

# NUCLEIC ACIDS

- **Nucleotides** are one of the most important classes of molecules in the body.
  - They act as the basis for sources of energy (adenosine triphosphate [ATP], guanosine triphosphate [GTP]) that drive biochemical reactions.
  - Important constituent in several coenzymes,
  - The most famous role of the nucleotides is as a component of the nucleic acids- RNA and DNA.
- A nucleotide has three characteristic components:
  - (1) a nitrogenous base,
  - (2) a pentose, and
  - (3) one or more phosphates
- The molecule without a phosphate group is called a **nucleoside**.
- The nitrogenous bases are derivatives of two parent compounds, **pyrimidine** and **purine**. The bases and pentoses of the common nucleotides are heterocyclic compounds.

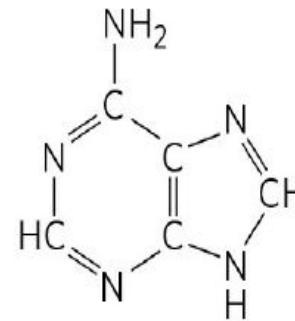


Pyrimidine

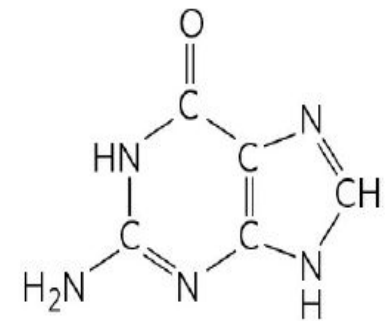


Purine

- Both DNA and RNA contain two major purine bases, **adenine (A)** and **guanine (G)**, and two major pyrimidines.



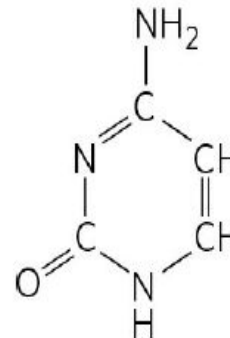
Adenine



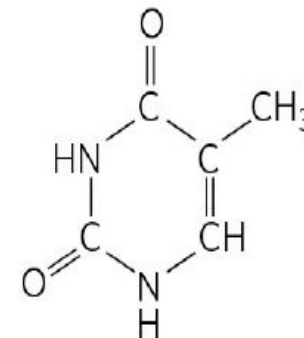
Guanine

### Purines

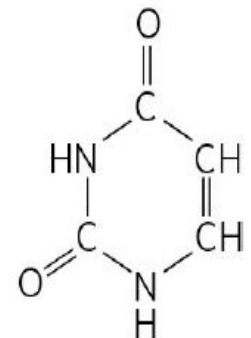
- In both DNA and RNA one of the pyrimidines is **cytosine (C)**, but the second common pyrimidine is not the same in both: it is **thymine (T)** in DNA and **uracil (U)** in RNA.



Cytosine



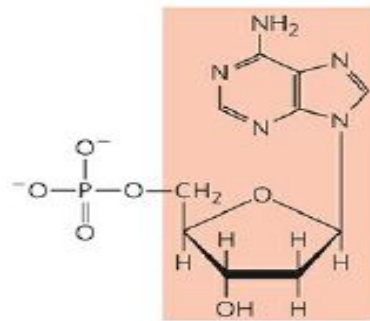
Thymine  
(DNA)



Uracil  
(RNA)

### Pyrimidines

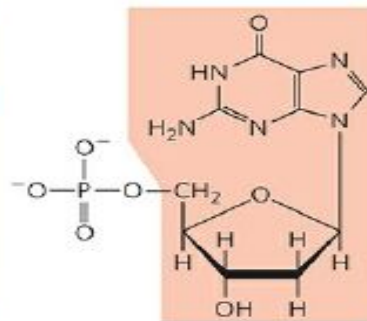
**Major purine and pyrimidine bases of nucleic acids.**



**Nucleotide:** Deoxyadenylate  
(deoxyadenosine  
5'-monophosphate)

**Symbols:** A, dA, dAMP

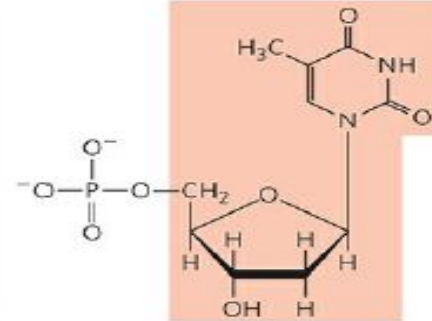
**Nucleoside:** Deoxyadenosine



**Nucleotide:** Deoxyguanylate  
(deoxyguanosine  
5'-monophosphate)

**Symbols:** G, dG, dGMP

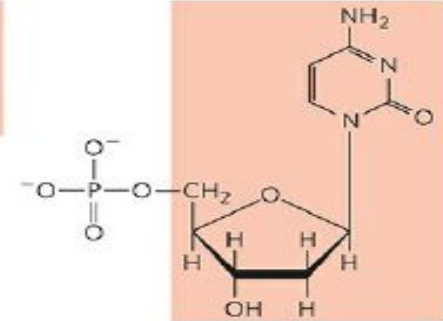
**Nucleoside:** Deoxyguanosine



**Nucleotide:** Deoxythymidylate  
(deoxythymidine  
5'-monophosphate)

**Symbols:** T, dT, dTMP

**Nucleoside:** Deoxythymidine

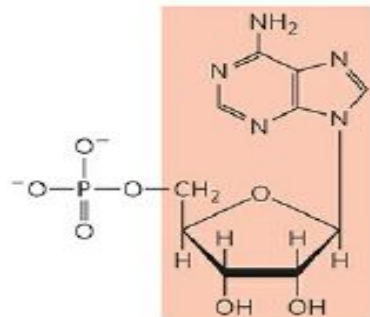


**Nucleotide:** Deoxycytidylate  
(deoxycytidine  
5'-monophosphate)

**Symbols:** C, dC, dCMP

**Nucleoside:** Deoxycytidine

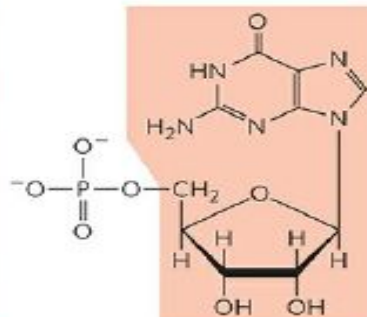
#### (a) Deoxyribonucleotides



**Nucleotide:** Adenylate (adenosine  
5'-monophosphate)

**Symbols:** A, AMP

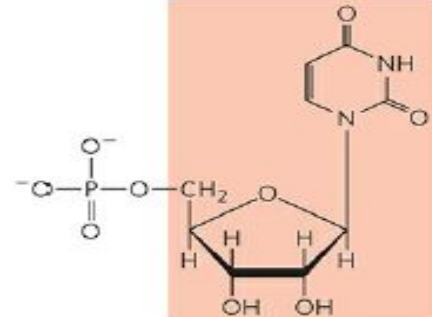
**Nucleoside:** Adenosine



**Nucleotide:** Guanylate (guanosine  
5'-monophosphate)

**Symbols:** G, GMP

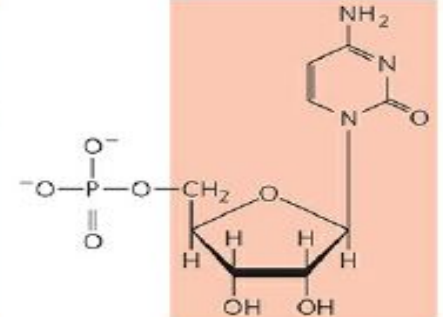
**Nucleoside:** Guanosine



**Nucleotide:** Uridylate (uridine  
5'-monophosphate)

**Symbols:** U, UMP

**Nucleoside:** Uridine



**Nucleotide:** Cytidylate (cytidine  
5'-monophosphate)

**Symbols:** C, CMP

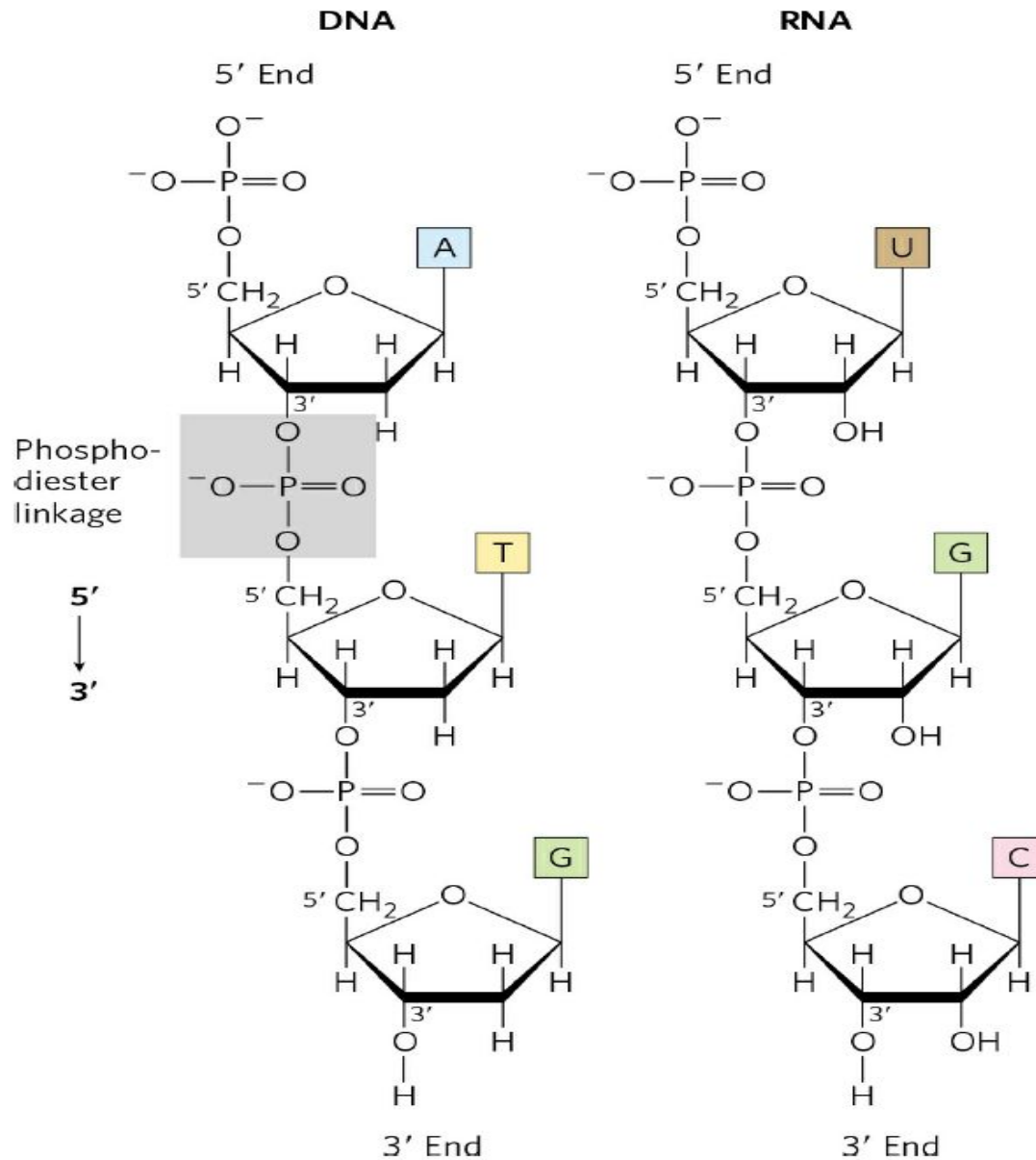
**Nucleoside:** Cytidine

#### (b) Ribonucleotides

- Nucleic acids have two kinds of pentoses: deoxyribonucleotide units of DNA contain 2'-deoxy-D-ribose, and the ribonucleotide units of RNA contain D-ribose. In nucleotides, both pentoses are in their  $\beta$ -furanose form.

# Nucleic Acids

- Successive nucleotides of DNA and RNA are covalently linked through **phosphodiester bond** in which the 5'-phosphate group of one nucleotide unit is joined to the 3'-hydroxyl group of the next.
- The covalent backbones of nucleic acids consist of alternating phosphate and pentose residues, and the nitrogenous bases may be regarded as side groups joined to the backbone at regular intervals.
- Polymers containing 50 or fewer nucleotides are generally called **oligonucleotides**.
- A longer nucleic acid is called a **polynucleotide**.

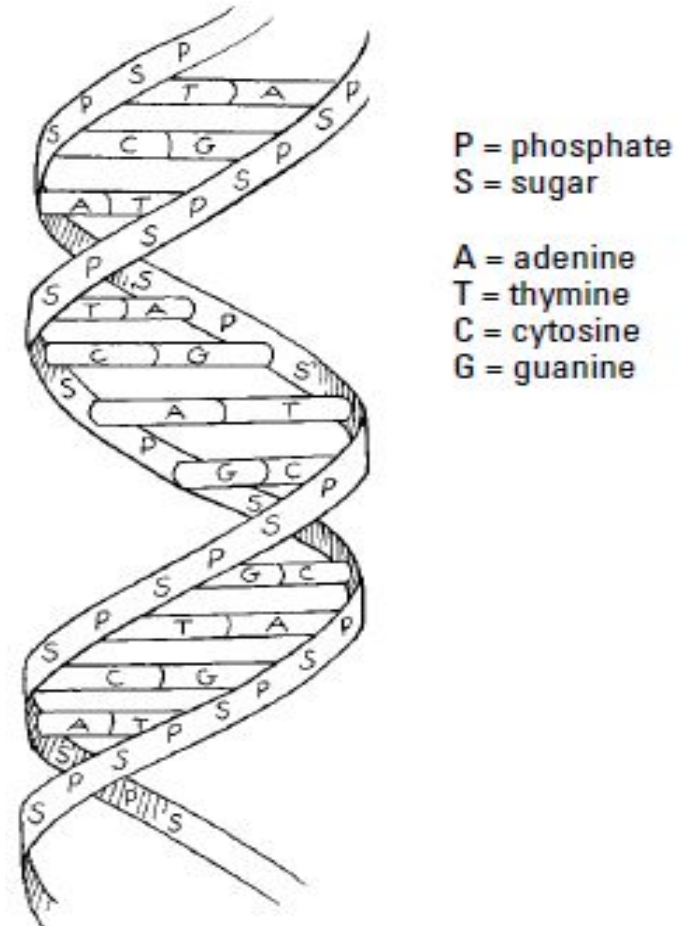


**Phosphodiester linkages in the covalent backbone of DNA and RNA.**

- The order and type of nucleotides in a polynucleotide strand make up its **primary structure**.
- The order of nitrogenous bases within a polynucleotide chain is called its **sequence**.
- Both DNA and RNA are made of polynucleotide chains, but the two types of molecules are unique.
- Three major differences exist between DNA and RNA:
  - ✓ DNA contains the nitrogenous base thymine, while RNA contains uracil. (Both contain adenine, guanine, and cytosine.)
  - ✓ DNA nucleotides have the sugar deoxyribose, while RNA nucleotides have ribose.
  - ✓ DNA molecules are double-stranded, while RNA molecules are single stranded.

# The double helix of DNA

- Polynucleotide chains interact with each other and with themselves to give nucleic acids three-dimensional shape, called their **secondary structure**.
- The secondary structure of DNA is a **double helix**.
- Two polynucleotide chains join together, forming a molecule with the shape of a twisted ladder. The sides of the ladder are made of the sugar phosphate backbones of the two strands.
- The nitrogenous bases project off of the sugar-phosphate backbone and join together by hydrogen bonds, forming the “rungs” of the ladder.
- 



The double helix structure of DNA.

- The two chains of the double helix are **antiparallel** to each other, i.e. they are opposite in polarity.
- The two antiparallel strands of DNA are held together by **hydrogen bonds** between their nitrogenous bases.
- In order for the hydrogen bonds to form, Adenine (A) forms hydrogen bonds with thymine (T), and cytosine (C) forms hydrogen bonds with guanine (G).
- In DNA, **A** is **complementary** to **T**, and **C** is **complementary** to **G**.
- When these bases match up, they form complementary base pairs. The base pairing rules for DNA are A=T and C=G.

# The Function of DNA and RNA

- DNA and RNA molecules are involved in information storage and retrieval in cells.
- DNA stores the information that ultimately determines the characteristics of cells and organisms. The information is written in a chemical **code** determined by the **order of nitrogenous bases** within the DNA.
- When cells decode the DNA message, it provides instructions for important cell structures and functions, including the following:
  - ✓ Protein structure is determined by the sequence of nucleotides in DNA. Through the processes of **transcription** and **translation**, the code in DNA is copied and used to specify the sequence of amino acids (primary structure) in polypeptide chains.
  - ✓ RNA structure is also determined by the sequence of nucleotides in DNA. Through the process of **transcription**, the sequence of nucleotides in DNA is used to build RNA molecules.

# Types of RNA

- Several types of RNA molecules are built from the information in DNA. Each type of RNA has a different function in cells:
  - ✓ **Messenger RNA (mRNA)** carries the code for protein structure from the DNA to ribosomes where it can be used to produce proteins.
  - ✓ **Transfer RNA (tRNA)** decodes the message in mRNA by matching amino acids to the mRNA code.
  - ✓ **Ribosomal RNA (rRNA)** is part of the structure of the ribosome.