

Part I: Molecular Biology

1. Review

1.1. Structure and Function of Genes

1.1.1. Structure of Nucleic Acids (DNA and RNA)

Structure of DNA

Deoxyribonucleic acid (DNA) is a polymer composed of monomers called nucleotides (a polynucleotide).

A nucleotide consists of:

- **A sugar:** A five-carbon sugar (pentose) called 2'-deoxyribose, because the OH group on the second carbon of ribose is replaced by a hydrogen atom.

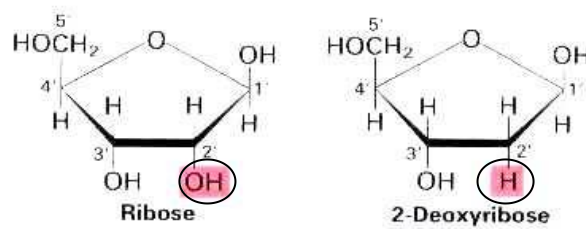


Figure 1 : Ribose and deoxyribose

- **A nitrogenous base:** A purine base (adenine and guanine) or a pyrimidine base (cytosine and thymine).

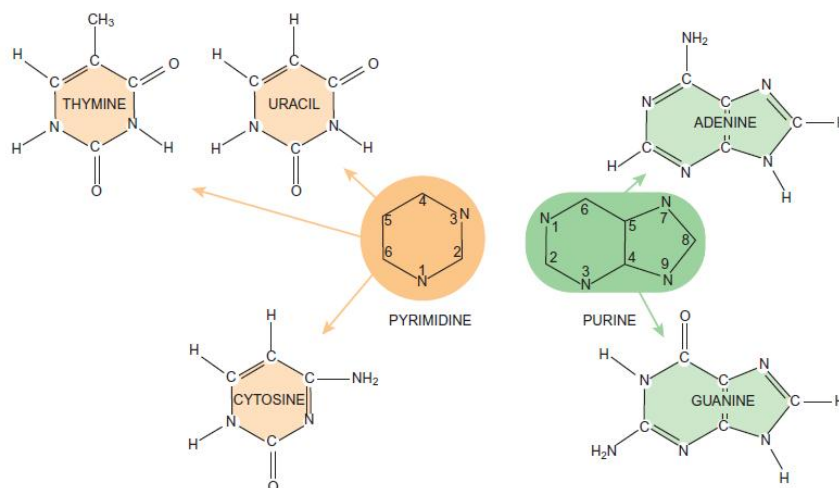


Figure 2: The Bases of the Nucleic Acids: The four bases of DNA are adenine, guanine, cytosine, and thymine. In RNA, uracil replaces thymine. Pyrimidine bases contain one-ring structures, whereas purine bases contain two-ring structures (Clark et al., 2019).

The bases are attached to the sugar by a bond between:

- The 1' carbon of the sugar and the 9 nitrogen atom of purine bases.
- The 1' carbon of the sugar and the 1 nitrogen atom of pyrimidine bases.

The combination of a sugar and a base is called a nucleoside.

- A phosphate group (PO₄)

Phosphate groups are attached to the 5' carbon of the sugar.

Deoxyribonucleic acids are very large molecules, generally composed of two polynucleotide chains coiled around each other to form a double helix with a diameter of 2 nm. In the cell, nucleotides are found in their

free form (where they play an important role as energy carriers for enzymatic reactions) or polymerized in the form of DNA or RNA.

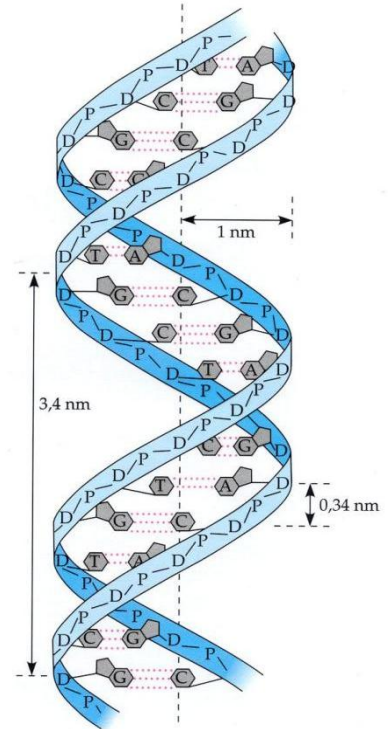
Nucleotides are joined by linking the phosphate on the 5' -carbon of the deoxyribose of one nucleotide to the 3' -hydroxyl of the other; The phosphate group is joined to the sugar on either side by ester linkages, and the overall structure is therefore a phosphodiester linkage. The phosphate group linking the sugars has a negative charge. The polynucleotide chain has a free 5' phosphate group at one end and a free 3' hydroxyl group at the other end. The 5'-end is regarded as the beginning of a DNA or RNA Strand.

The sugar and phosphate part of the DNA molecule forms its backbone and is located on the outside of the helix. The nitrogenous bases face each other in the center of the helix and are stacked on top of one another. The bases on each strand protrude into the center of the double helix where they are paired with the bases in the other strand by means of **hydrogen bonds**.

The double helix makes one complete turn approximately every 10 base pairs. The **pitch** of the helix is 34 Å (3.4 nm).

The double helix has an antiparallel structure: one strand is oriented in the 5'→3' direction, and the other in the 3'→5' direction.

- Adenine (A) is always paired with thymine (T) by two hydrogen bonds.
- Guanine (G) is always paired with cytosine (C) by three hydrogen bonds.
- This base pairing between A–T and G–C means that the two strands in a DNA double helix are complementary.



Chargaff's rules: the concentrations of adenine (A) and thymine (T) are equal, as are those of guanine (G) and cytosine (C). Therefore, $A/T = G/C = 1$ and $(A + G)/(T + C) = 1$.

The GC percentage is characteristic of the species.

There are several forms of DNA, such as A-DNA, B-DNA, and Z-DNA. The form that is predominantly found in cells is called B-DNA.

Structure de l'ARN

There are several important differences between DNA and RNA:

- In RNA, ribose replaces the 2'-deoxyribose found in DNA (Figure 1).
- Thymine is replaced by uracil (U), which also pairs with adenine.
- RNA is single-stranded. An RNA strand can fold back on itself by pairing complementary bases to form a hairpin structure.
- There are three main types of RNA: messenger RNA, ribosomal RNA, and transfer RNA.

1-1-2. Physicochemical Properties of Nucleic Acids

(See tutorial/tutorial exercises.)

1-1-3- DNA Organization in Cells

DNA appears to be much more organized in eukaryotic chromatin and is associated with several proteins, the most important being histones. Histones are small basic proteins rich in the amino acids lysine and/or arginine (which are positively charged). These positive charges allow them to bind strongly to negatively charged DNA (due to the negative charges of the phosphate groups in DNA).

Nearly all eukaryotic cells contain five types of histones: H1, H2A, H2B, H3, and H4. A nucleosome is composed of an octamer of eight histone molecules (two each of H2A, H2B, H3, and H4) wrapped in approximately 200 base pairs (bp) of DNA. The combination of DNA and proteins that makes up chromosomes is called chromatin. The H1 histone molecule is involved in DNA compaction.

Non-histone proteins (which are acidic) play a role in several aspects of chromatin function, particularly the regulation of gene expression.

Although DNA exists as a double helix in both prokaryotes and eukaryotes, its organization differs between the two cell types. DNA is organized into a closed circle in almost all prokaryotes (though the *Borrelia* chromosome consists of a linear DNA molecule). This circular double helix is further wound into a supercoil and is associated with basic proteins rather than the histones complexed with almost all eukaryotic DNA. These histone-like proteins apparently function to organize bacterial DNA into a coiled, chromatin-like structure.

1-1-4- Gene Structure

Biological information is encoded by the sequence of nitrogenous bases in DNA, and it is organized into a large number of genes, each containing the information necessary for the synthesis of a peptide.

A gene is a unit of information, and it corresponds to a DNA segment that encodes the amino acid sequence of a protein.

A gene can also be defined as a polynucleotide sequence encoding a polypeptide, a tRNA, or an rRNA. Genes vary greatly in size, ranging from 100 base pairs to several million base pairs.

Introns and Exons

Prokaryotic genes have a very different structure from eukaryotic genes. In bacterial and viral genes, the coding information is usually continuous (though some bacterial genes do contain introns), whereas in eukaryotes, genes contain coding sequences (exons) that are regularly interrupted by non-coding sequences (introns). An interesting exception to this rule involves eukaryotic histone genes, which do not contain introns.

Gene Promoters

Gene expression is regulated by a fragment of the DNA sequence located upstream of the coding sequence: the promoter.

Conserved DNA sequences within the promoter are recognized by RNA polymerase, which binds to them along with associated proteins called transcription factors, allowing the gene to be transcribed into RNA.

