Molecular Biology

MSc Students of Microbiology, Physiology & Pathology

By

Prof. Dr. Inam Jasim Lafta

MSc. Microbiology/ College of Veterinary Medicine/ University of Baghdad, Iraq.

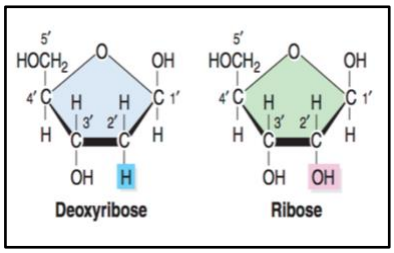
PhD Molecular Biology and Biotechnology/ Faculty of Science & Medical School/ University of Sheffield, UK.

2023/2024

The composition and structure of DNA and RNA

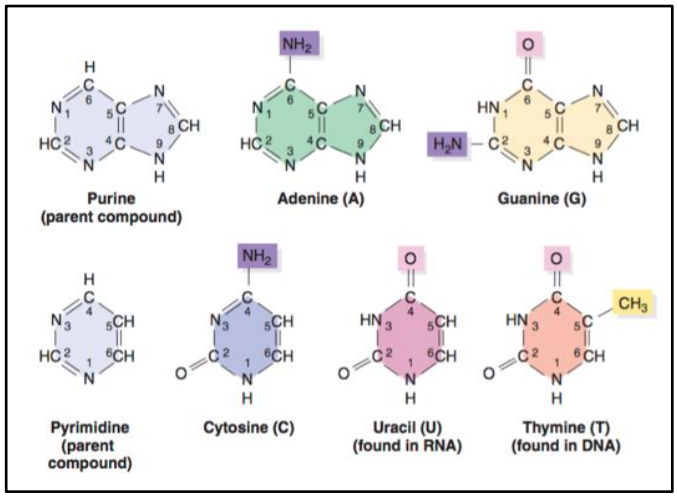
What is the molecular structure of DNA?
DNA and RNA are *polymers*—large molecules that consist of many similar smaller molecules, called *monomers*, linked together. The monomers that make up DNA and RNA are **nucleotides.** Each nucleotide consists of a **pentose** (five-carbon) **sugar,** a **nitrogenous** (nitrogen-containing) **base** (usually just called a **base**), and a **phosphate group.**

In DNA, the pentose sugar is **deoxyribose,** and in RNA it is **ribose**. The two sugars differ by the chemical groups attached to the 2’ carbon: a hydrogen atom (H) in deoxyribose and a hydroxyl group (OH) in ribose. (The carbon atoms in the pentose sugar are numbered 1’ to 5’ to distinguish them from the numbered carbon and nitrogen atoms in the rings of the bases).



There are 2 classes of nitrogenous bases: the **purines,** which are nine-membered, double-ringed structures, and the **pyrimidines,** which are six-membered, single-ringed structures. (The carbons and nitrogens of the purine rings are numbered 1 to 9, and those of the pyrimidines are numbered 1 to 6).

There are two purines—**adenine (A)** and **guanine (G)**—and three different pyrimidines— **thymine (T), cytosine (C),** and **uracil (U)** found in nucleic acid**.** Both DNA and RNA contain adenine, guanine, and cytosine; however, thymine is found only in DNA, and uracil is found only in RNA.



In DNA and RNA, bases are covalently attached to the 1’ carbon of the pentose sugar. The combination of a sugar and a base is called a nucleoside. Addition of a phosphate group (PO_4^{2-}) to a nucleoside yields a **nucleoside phosphate,** which is one kind of nucleotide. The phosphate group is attached to the 5’ carbon of the sugar in both DNA and RNA. Thus, the DNA nucleotide is called (a **deoxyribonucleotide**) and an RNA nucleotide (a **ribonucleotide**).

To form **polynucleotides** of either DNA or RNA, nucleotides are linked together by a covalent bond between the phosphate group of one nucleotide and the 3’ carbon of the sugar of another nucleotide. These 5’-to-3’ phosphate linkages are called **phosphodiester bonds.** The phosphodiester bonds are relatively strong, so the repeated sugar–phosphate–sugar–phosphate backbone of DNA and RNA is a stable structure.

Polynucleotide chains have *polarity*, meaning that the two ends are different: there is a 5’ carbon (with a phosphate group on it) at one end, and a 3’ carbon (with a hydroxyl group on it) at the other end. The ends of a polynucleotide are routinely referred to as the 5’ end and the 3’ end.

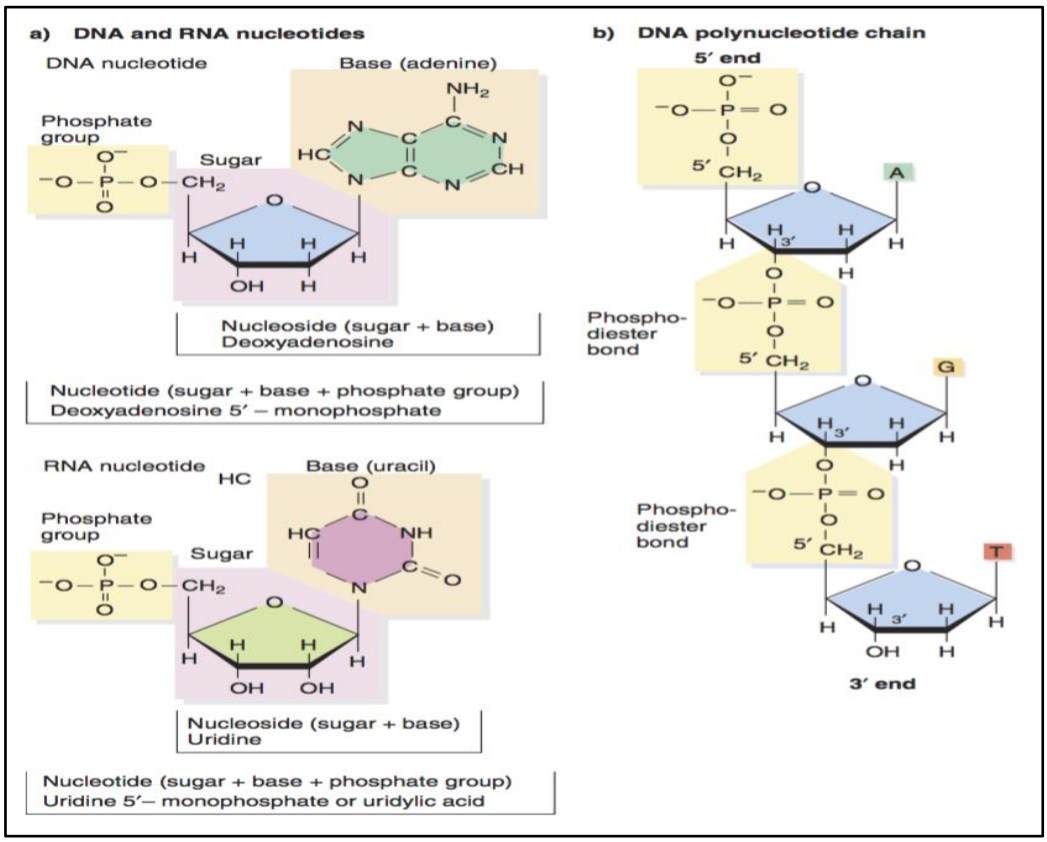


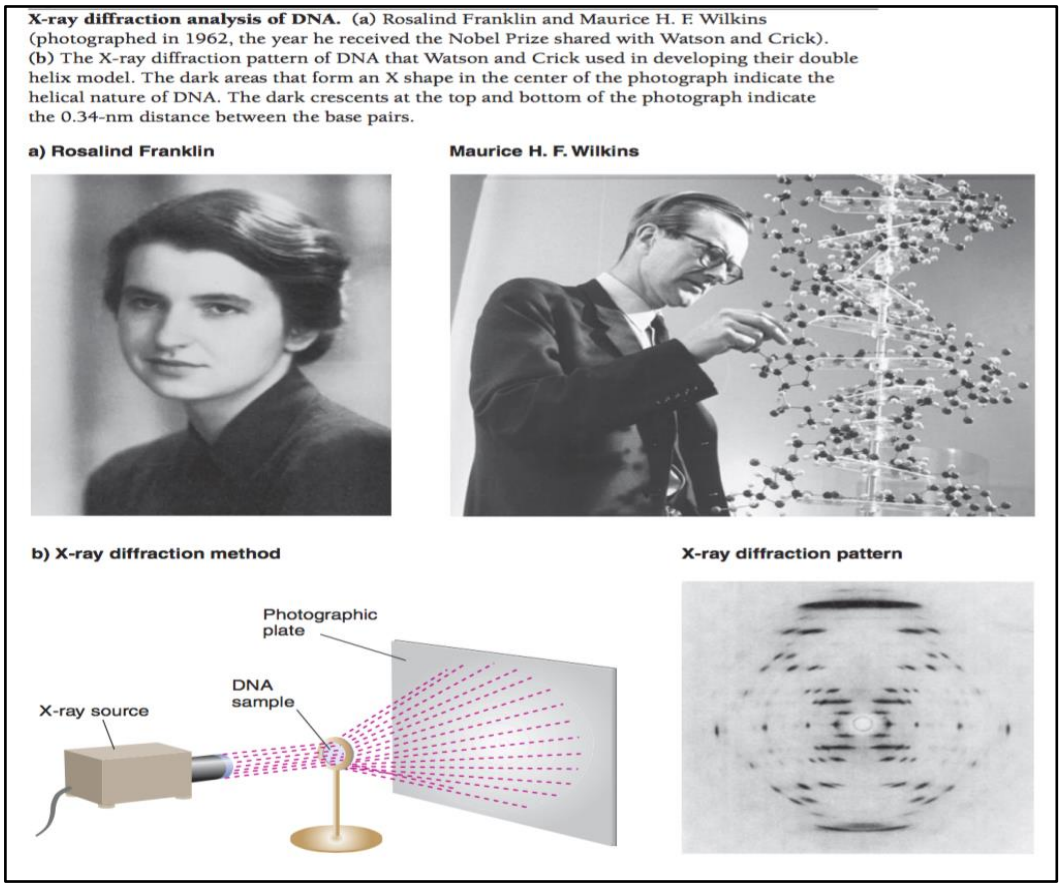
Table 2.1 Nucleic acid nomenclature.

Base	Nucleoside		Nucleotide*	
	DNA	RNA	DNA	RNA
Adenine (A)	Deoxyadenosine	Adenosine	Deoxyadenosine 5′-triphosphate (dATP)	Adenosine 5′-triphosphate (ATP)
Guanine (G)	Deoxyguanosine	Guanosine	Deoxyguanosine 5′-triphosphate (dGTP)	Guanosine 5′-triphosphate (GTP)
Cytosine (C)	Deoxycytidine	Cytidine	Deoxycytidine 5′-triphosphate (dCTP)	Cytidine 5′-triphosphate (CTP)
Thymine (T)	Deoxythymidine	–	Deoxythymidine 5′-triphosphate (dTTP)	–
Uracil (U)	–	Uridine	–	Uridine 5′-triphosphate (UTP)
Generic (N)	Deoxynucleoside	Nucleoside	Deoxynucleoside 5′-triphosphate (dNTP)	Nucleoside 5′-triphosphate (NTP)

* Nucleotides may contain one phosphate unit (monophosphate), two such units (diphosphate), or three (triphosphate). The triphosphate form shown in the table serves as the precursor building block for nucleic acid synthesis.

Base Composition Studies. By chemical treatment, Erwin Chargaff hydrolyzed the DNA of a number of organisms and quantified the purines and pyrimidines released. His studies showed that 50% of the bases were purines and 50% were pyrimidines. More important, the amount of adenine (**A**) was equal to that of thymine (**T**), and the amount of guanine (**G**) was equal to that of cytosine (**C**). These equivalencies have become known as Chargaff's rules. In comparisons of DNAs from different organisms, the **A/T** ratio is 1 and the **G/C** ratio is 1, but the (**A**+**T**)/(**G**+**C**) ratio (typically denoted %**GC**) varies. Because the amount of purines equals the amount of pyrimidines, the (**A**+**G**)/(**C**+**T**) ratio is 1.

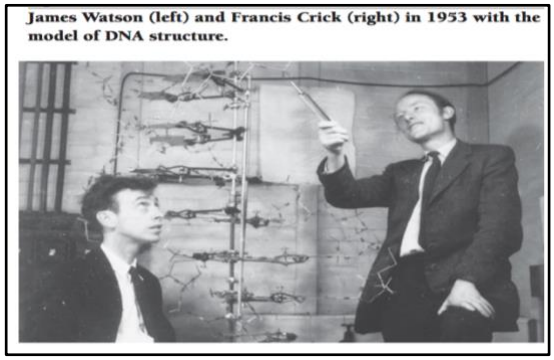
X-Ray Diffraction Studies. Rosalind Franklin, working with Maurice H. F. Wilkins, studied concentrated solutions of DNA pulled out into thin fibers. The analysis technique they used was X-ray diffraction, in which a beam of parallel X-rays is aimed at molecules. The diffracted X- rays were recorded on a photographic plate (Figure b below). By analyzing the photographs, Franklin obtained information about the molecule's atomic structure. In particular, she concluded that DNA is a helical structure with two distinctive regularities of 0.34 nm and 3.4 nm along the axis of the molecule.



X-ray diffraction analysis of DNA. (a) Rosalind Franklin and Maurice H. F. Wilkins (photographed in 1962, the year he received the Nobel Prize shared with Watson and Crick). (b) The X-ray diffraction pattern of DNA that Watson and Crick used in developing their double helix model. The dark areas that form an X shape in the center of the photograph indicate the helical nature of DNA. The dark crescents at the top and bottom of the photograph indicate the 0.34-nm distance between the base pairs.

Watson and Crick's Model (The DNA Double Helix)

James D. Watson and Francis H. C. Crick used some of Franklin's data and some intelligent guesses of their own to build three-dimensional models of the structure of DNA. In 1953, Watson and Crick proposed a model for the physical and chemical structure of the DNA molecule. The model they devised, which fit all the known data on the composition of the DNA molecule, is the now-famous double helix model for DNA.


James Watson (left) and Francis Crick (right) in 1953 with the model of DNA structure.

At the time of Watson and Crick's work, DNA was known to be composed of nucleotides. However, it was not known how the nucleotides formed the structure of DNA. Watson and Crick thought that understanding the structure of DNA would help determine how DNA acts as the genetic basis for living organisms. The data they used to help generate their model came primarily from base composition studies conducted by Erwin Chargaff, and X-ray diffraction studies conducted by Rosalind Franklin and Maurice H. F. Wilkins.

Watson and Crick's double helix model of DNA based on the X-ray crystallography data has the following main features:

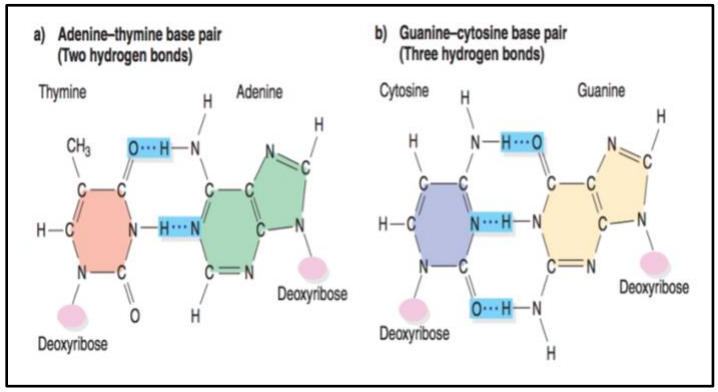
1. The DNA molecule consists of two polynucleotide chains wound around each other in a right-handed double helix; that is, the two strands wind around each other in a clockwise (right-handed) fashion.

2. The two chains are **antiparallel** (show ***opposite polarity***); that is, the two strands are oriented in opposite directions, with one strand oriented in the 5'-to-3' way and the other strand oriented 3' to 5'.

3. The sugar–phosphate backbones are on the outsides of the double helix, with the bases oriented toward the central axis.

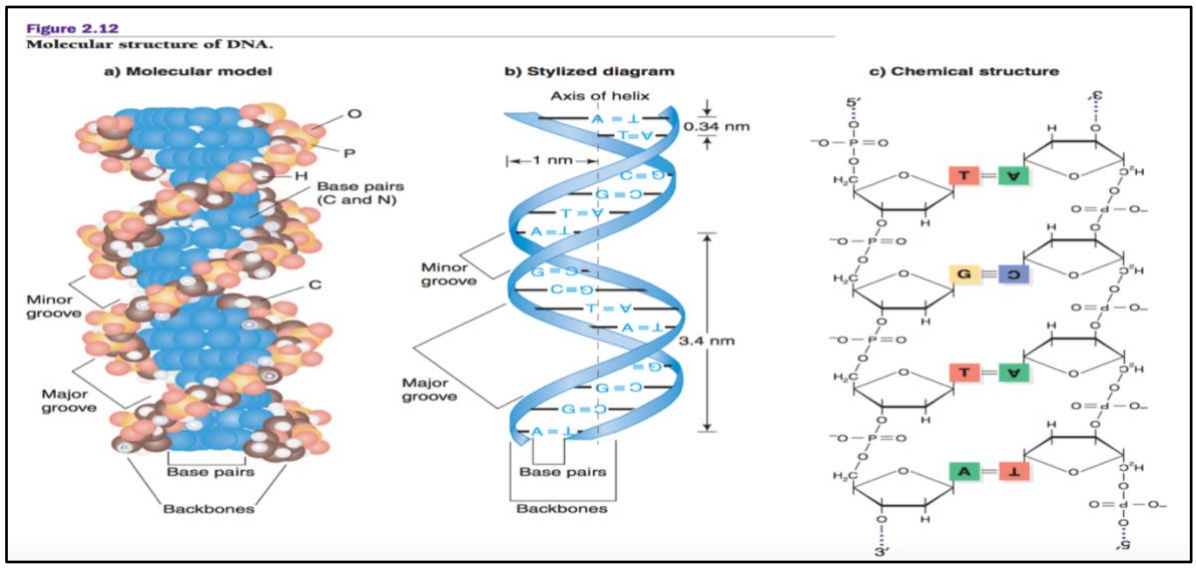
4. The bases in each of the two polynucleotide chains are bound together by hydrogen bonds, which are relatively weak chemical bonds. The specific pairings observed are **A** bound with **T** (two hydrogen bonds) and **G** bound with **C** (three hydrogen bonds). The hydrogen bonds make it relatively easy to separate the two strands of the DNA—for example, by heating. The **A–T** and **G–C** base pairs are the only ones that can fit the physical dimensions of the helical model, and their arrangement is in accord with Chargaff's rules.

The specific **A–T** and **G–C** pairs are called **complementary base pairs,** so the nucleotide sequence in one strand dictates the nucleotide sequence of the other. For instance, if one chain has the sequence **5'-TATTCCGA-3'**, then the opposite, antiparallel chain must bear the sequence **3'-ATAAGGCT-5'**.



5. The base pairs are 0.34 nm apart in the DNA helix. A complete (360°) turn of the helix takes 3.4 nm; therefore, there are 10 base pairs (bp) per turn. The external diameter of the helix is 2 nm.

6. Because of the way the bases bound with each other, the two sugar–phosphate backbones of the double helix are not equally spaced from one another along the helical axis. This unequal spacing results in grooves of unequal size between the backbones; one groove is called the *major* (wider) *groove*, the other the *minor* (narrower) *groove*.



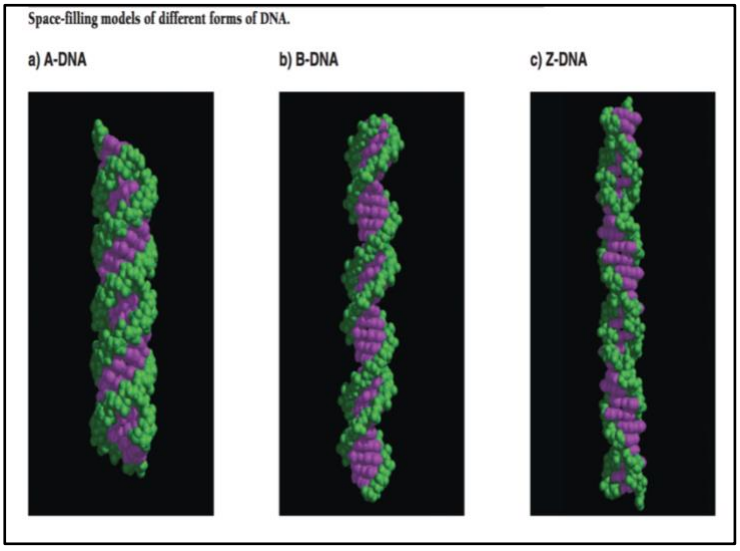
Different DNA Structures

Researchers have now shown that DNA can exist in several different forms—most notably, the A-, B-, and Z- DNA forms.

A-DNA is seen only in conditions of low humidity. B-DNA forms under conditions of high humidity and is the structure that most closely corresponds to that of DNA in the cell. Whether Z-DNA exists in cells has long been a topic of debate among scientists. In

those organisms where there is some evidence for Z-DNA, its physiological significance is unknown.

DNA in the cell is in solution, which is a different state from the DNA used in X-ray crystallography experiments. Experiments have shown that DNA in solution has 10.5 base pairs per turn, which is a little less twisted than B-DNA. However, DNA in the cell most closely resembles B-DNA, and most of the genome is in that form. In certain DNA–protein complexes, though, the DNA assumes the A-DNA structure.



RNA Structure

RNA is molecularly similar to DNA, differing in having ribose as the sugar rather than deoxyribose, and uracil (**U**) as a pyrimidine base instead of thymine. In the cell, the functional forms of RNA such as messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), small nuclear RNA (snRNA), and micro RNA (miRNA) are single-stranded molecules. However, these molecules are not stiff, linear rods. Rather, wherever bases can pair together, they will do so. This means that a single-stranded RNA molecule will fold up on itself to produce regions of antiparallel double-stranded RNA separated by segments of unpaired RNA. This configuration is called the secondary structure of the molecule.

Single-stranded RNA and double-stranded RNA molecules are the genomes of certain viruses. Double-stranded RNA has a structure similar to that of double-stranded DNA, with antiparallel strands, the sugar–phosphate backbones on the outside of the helical molecule, and complementary base pairs formed by hydrogen bonding in the middle of the helix.