The Double Helical Structure of DNA

An important clue to the structure of DNA came from the work of Erwin Chargaff and his colleagues in the late 1940s. They found that the four nucleotide bases of DNA occur in different ratios in the DNAs of different organisms and that the amounts of certain bases are closely related. These data, collected from DNAs of a great many different species led Chargaff to the following conclusions (Chargaff's Rules)

1. The basic composition of DNA generally varies from one species to another.
2. DNA specimens isolated from different tissues of the same species have the same base composition.
3. The base composition of DNA in a given species does not change with an organism's age, nutritional state, or changing environment.
4. In all cellular DNAs, regardless of the species, the number of adenosine residues is equal to the number of thymidine residues (that is, A = T), and the number of Guanosine residues is equal to the number of cytidine residues (G = C). From these relationships it follows that the sum of the purine residues equals the sum of the pyrimidine residues; that is, A+G = T + C.

These quantitative relationships, sometimes called **"Chargaff's rules,"** were confirmed by many subsequent researchers. They were a key to establishing the three-dimensional structure of DNA and yield clues to how genetic information is encoded in DNA and passed from one generation to the next.

To shed more light on the structure of DNA, Rosalind Franklin and Maurice Wilkins used the powerful method of x-ray diffraction to analyse DNA fibres. They showed in the early 1950s that DNA produces a characteristic x-ray diffraction pattern. From this pattern it was deduced that DNA molecules are helical with two periodicities along their long axis, a primary one of 3.4 Å and a secondary one of 34 Å. The problem then was to formulate a three-dimensional model of the DNA molecule that could account not only for the x-ray diffraction data but also for the specific A = T and G = C base equivalences discovered by Chargaff and for the other chemical properties of DNA. **James Watson** and **Francis Crick** relied on this accumulated information about DNA to set about deducing its structure. In 1953 they postulated a three-dimensional model of DNA structure that accounted for all the available data.

Based on the contribution of several researchers finally the structure of DNA was established and got popularised as double helical model. The DNA consists of two helical DNA chains wound around the same axis to form a right-handed double helix. The hydrophilic

backbones of alternating deoxyribose and phosphate groups are on the outside of the double helix, facing the surrounding water. The β-furanose ring of each deoxyribose is in the C-2' endo conformation. The purine and pyrimidine bases of both strands are stacked inside the double helix, with their hydrophobic and nearly planar ring structures very close together and perpendicular to the long axis. Each nucleotide base of one strand is paired in the same plane with a base of the other strand. Base pairs are hydrogen bonded, G with C and A with T, and this pairing fit best within the structure, providing a rationale for Chargaff's rule that in any DNA, G = C and A = T. It is important to note that three hydrogen bonds can form between G and C, symbolized G ≡ C, but only two can form between A and T, symbolized A = T. This is one reason for the finding that separation of paired DNA strands is more difficult the higher the ratio of G : C to A : T base pairs. The two antiparallel polynucleotide chains of double-helical DNA are not identical in either base sequence or composition. Instead they are **complementary** to each other. Wherever adenine occurs in one chain, thymine is found in the other; similarly, wherever guanine occurs in one chain, cytosine is found in the other. That is the base at the 5' end of one strand is paired with the base at the 3' end of the other strand. The strands are said to have an **antiparallel orientation.** The strictness of the rules for this "Watson-Crick" pairing derives from the complementarity both of shape and of hydrogen bonding properties between adenine and thymine and between guanine and cytosine. Adenine and thymine match up so that a hydrogen bond can form between the exocyclic amino group at C6 on adenine and the carbonyl at C4 in thymine; and likewise, a hydrogen bond can form between N1 of adenine and N3 of thymine. A corresponding arrangement can be drawn between a guanine and a cytosine, so that there is both hydrogen bonding and shape complementarity in this base pair as well. Likewise, a hydrogen bond can form between N1 of guanine and N3 of cytosine and between the carbonyl at C6 of guanine and the exocyclic NH_2 at C4 of cytosine. Watson-Crick base pairing also requires that the bases be in their preferred tautomeric states. An important feature of the double helix is that the two base pairs have exactly the same geometry; having an A:T base pair or a G:C base pair between the two sugars does not perturb the arrangement of the sugars because the distance between the sugar attachment points are the same for both base pairs. Neither does T:A or C:G. The hydrogen bonds between complementary bases are a fundamental feature of the double helix, contributing to the thermodynamic stability of the helix and the specificity of base pairing. Hydrogen bonding might not, at first glance, appear to contribute importantly to the stability of DNA for the following reason. A second important contribution comes from stacking interactions between the bases. The bases are flat, relatively water-insoluble molecules, and they tend to stack above each other roughly perpendicular to the direction of the helical axis. Electron cloud interactions (π-π) between bases in the helical stacks contribute significantly to the stability of the double helix. Hydrogen bonding is particularly important for the specificity of base pairing.

The overall diameter of a DNA double helix is 2 nm or 20 Angtrom, while one complete turn of DNA gives the DNA a helical rise of 34 Angstrom. As there are 10 base pairs per turn of the DNA (10.5 in realistic conditions), the gap between two base pairs is approximately 3.4 A. There is a minor groove and a major groove. It is a simple consequence of the geometry of the base pair. The angle at which the two sugars protrude from the base pairs (i.e., the angle between the glycosidic bonds) is about 120° (for the narrow angle) or 240° (for the wide angle). As a result, as more and more base pairs stack on top of each other, the narrow angle between the sugars on one edge of the base pairs generates a **minor groove** and the large angle on the other edge generates a **major groove**. (If the sugars pointed away from each other in a straight line, i.e., at an angle of 180°, then the two grooves would be of equal dimensions and there would

be no minor and major grooves.). The edges of each base pair are exposed in the major and minor grooves, creating a pattern of hydrogen-bond donors and acceptors and of van der Waals surfaces that identifies the base pair. The major and minor grooves are structurally significant features of the DNA double helix, and they play essential roles in interactions with proteins and other molecules. The major groove is wider and deeper than the minor groove, and both grooves are accessible sites on the DNA helix where proteins can recognize and bind specific sequences of DNA. Proteins, such as transcription factors, can recognize and bind to specific base sequences in the major or minor groove, leading to the regulation of gene expression. The three-dimensional shape of the major and minor grooves allows for specific hydrogen bonding and van der Waals interactions between the amino acid residues of the protein and the exposed bases of the DNA. This protein-DNA binding specificity is crucial for processes like transcription, replication, and repair, where proteins must accurately recognize and interact with specific DNA sequences to carry out their functions. The ability of proteins to bind to the major and minor grooves of DNA is fundamental to the precise control of gene expression and the overall regulation of cellular activities.

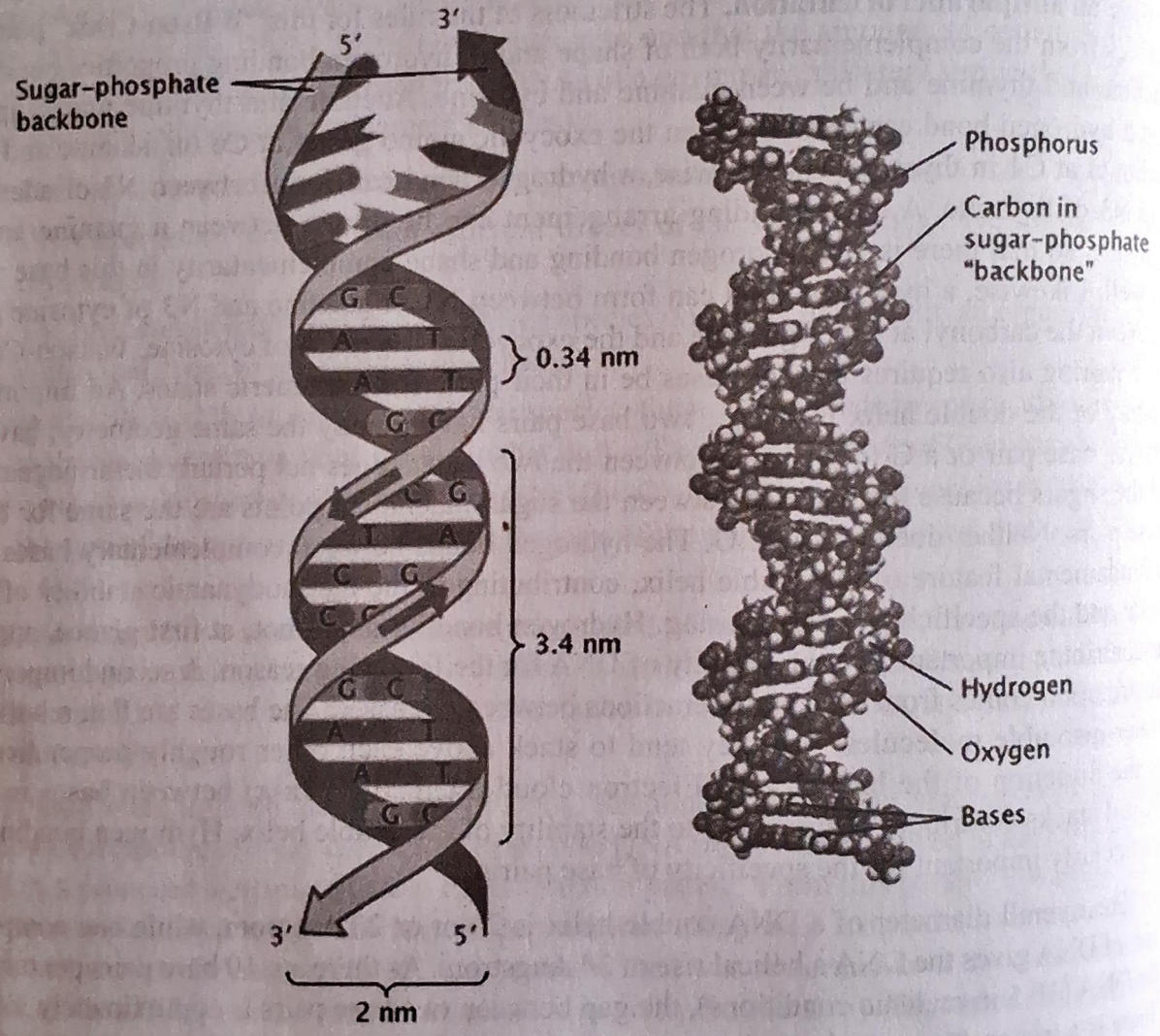

Fig. 1.3. *Structure of B DNA with antiparallel strands linked via hydrogen bonding.*