**Polynucleotide chains have nitrogenous bases linked to a sugar-phosphate backbone**

The basic building block of nucleic acids is the nucleotide. This has three components:

• A nitrogenous base;

• A sugar;

• A phosphate.

The nitrogenous base is a purine or pyrimidine ring. The base is linked to position 1 on a pentose sugar by a glycosidic bond from N1 of pyrimidines or N9 of purines. To avoid ambiguity between the numbering

systems of the heterocyclic rings and the sugar, positions on the pentose are given a prime (')•

Nucleic acids are named for the type of sugar; DNA has 2'-deoxyribose, whereas RNA has ribose. The difference is that the sugar in RNA has an OH group at the 2' position of the pentose ring. The sugar can be linked by its 5' or 3' position to a phosphate group. A nucleic acid consists of a long chain of nucleotides. **Figure 1.7** shows that the backbone of the polynucleotide chain consists of an alternating

series of pentose (sugar) and phosphate residues. This is constructed by linking the 5' position of one pentose ring to the 3' position of the next pentose ring via a phosphate group. So the sugar-phosphate

backbone is said to consist of 5'-3' phosphodiester linkages. The nitrogenous bases "stick out" from the backbone.

Each nucleic acid contains 4 types of base. The same two purines, adenine and guanine, are present in both DNA and RNA. The two pyrimidines in DNA are cytosine and thymine; in RNA uracil is found instead of thymine. The only difference between uracil and thymine is the presence of a methyl substituent at position C5. The bases are usually referred to by their initial letters. DNA contains A, G, C, T, while RNA contains A, G, C, U.

The terminal nucleotide at one end of the chain has a free 5' group; the terminal nucleotide at the other end has a free 3' group. It is conventional to write nucleic acid sequences in the 5'—*>3'* direction—that is, from the 5' terminus at the left to the 3' terminus at the right.



The observation that the bases are present in different amounts in the DNAs of different species led to the concept that the *sequence of bases is the form in which genetic information is carried.* By the 1950s, the concept of genetic information was common: the twin problems it posed were working out the structure of the nucleic acid, and explaining how a sequence of bases in DNA could represent the sequence of amino acids in a protein. Three notions converged in the construction of the double helix model for DNA by Watson and Crick in 1953: • X-ray diffraction data showed that DNA has the form of a regular helix, making a complete turn every 34 A (3.4 nm), with a diameter of ~20 A (2 nm). Since the distance between adjacent nucleotides is 3.4 A, there must be 10 nucleotides per turn.

• The density of DNA suggests that the helix must contain two polynucleotide chains. The constant diameter of the helix can be explained if the bases in each chain face inward and are restricted so that a purine is always opposite a pyrimidine, avoiding partnerships of purine-purine (too wide) or pyrimidine-pyrimidine (too narrow).

• Irrespective of the absolute amounts of each base, the proportion of G is always the same as the proportion of C in DNA, and the proportion of A is always the same as that of T. So the composition of any DNA can be described by the proportion of its bases that is G + C. This ranges from 26% to 74% for

different species. Watson and Crick proposed that the two polynucleotide chains in the double helix associate by *hydrogen bonding between the nitrogenous bases.* G can hydrogen bond specifically only with C, while A can bond specifically only with T. These reactions are described as base pairing, and the paired bases (G

with C, or A with T) are said to be complementary. The model proposed that the two polynucleotide chains run in opposite directions (antiparallel), as illustrated in **Figure 1.8.** Looking along the helix, one strand runs in the 5'—>3' direction,

while its partner runs 3'—»5'. The sugar-phosphate backbone is on the outside and carries negative charges on the phosphate groups. When DNA is in solution

*in vitro,* the charges are neutralized by the binding of metal ions, typically by Na+. In the cell, positively charged proteins provide some of the neutralizing force. These proteins play an important role in determining the organization of DNA in the cell.



The bases lie on the inside. They are flat structures, lying in pairs They are flat structures, lying in pairs perpendicular to the axis of the helix. Consider the double helix in terms of a spiral staircase: the base pairs form the treads, as illustrated schematically in Figure 1.9. Proceeding along the helix, bases are stacked above one another, in a sense like a pile of plates. Each base pair is rotated ~36° around the axis of the helix relative to the next base pair. So ~10 base pairs make a complete turn of 360°. The twisting of the two strands around one another forms a double helix with a minor groove (~12 A across) and a major groove (~22 A across), as can be seen from the scale model of Figure 1.10. The double helix is right-handed; the turns run clockwise looking along the helical axis.

These features represent the accepted model for what is known as the B-form of DNA. It is important to realize that the B-form represents an average, not a precisely specified structure. DNA structure can change locally. If it has more base pairs per turn it is said to be overwound; if it has fewer base pairs per turn it is underwound. Local winding can be affected by the overall conformation of the DNA double helix in space or by the binding of proteins to specific sites.



