Chem Factsheet

www.curriculum-press.co.uk Number 232

The Structure of DNA

Johannes Miescher
(1844 – 1895)

DNA (**d**eoxyribo**n**ucleic **a**cid) is arguably, along with RNA (**r**ibo**n**ucleic **a**cid) and proteins, amongst the most important molecules in nature. It is a polymeric molecule which is essential for all forms of life for storing genetic information in a chemically encoded format.

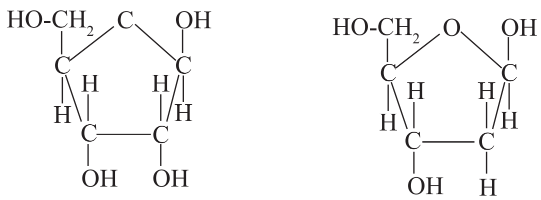
DNA was first isolated and identified in 1869 by the Swiss doctor and biologist Johannes Miescher.

The monomers of which this polymer is composed are generally called **nucleotides** each of which is made of three components - a 5-carbon sugar molecule (2-deoxyribose), a phosphate group ($H_2PO_4^-$) and a nitrogenous base (a purine or a pyrimidine). The sugar and phosphate group are common to all nucleotides with 4 different bases possible. This gives only four different nucleotides from which all DNA is composed.

2-Deoxyribose

Sugars usually have a general formula $C_m(H_2O)_n$. In the case of ribose, m = n = 5 with the structure shown in fig. 1.

Fig. 1



Ribose 2-Deoxyribose

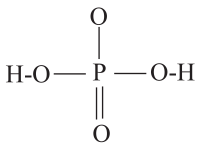
However, **2-deoxy**ribose (also shown in fig.1) differs in that there is an oxygen atom missing (hence "deoxy") at the C2 position.

Note: C1 is the C atom to the "right" of the molecules as they are shown. C2, 3, 4 and 5 are then counted clockwise round the ring structure.

Phosphate Group

Essentially this derived from phosphoric acid (H_3PO_4) by loss of a proton. Its structure is shown in fig. 2.

Fig. 2

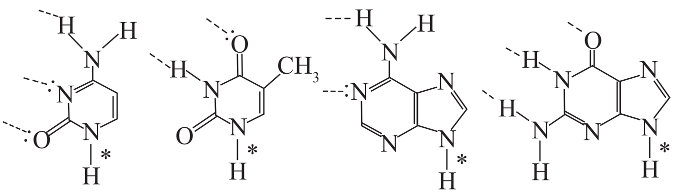


The OH groups are used to bond to 2-deoxyribose to form the main backbone of DNA – see later.

Nitrogenous Bases

As already mentioned, there are 4 different nitrogenous bases involved in DNA – **cytosine (C), thymine (T), adenine (A) and guanine (G)**. The structures are shown in fig. 3.

Fig 3



Cytosine Thymine Adenine Guanine
Pyrimidines Purines

Note: C and T both belong to the group of nitrogenous bases called **pyrimidines** because they have the same basic ring structure. Similarly, A and G belong to the **purines**.

Note: These are all bases because the can accept protons via the lone pair on the N atom of the N-H groups.

Note: A-level chemistry syllabuses will not require you to know these structures. However, how they link to 2-deoxyribose during DNA formation is important.

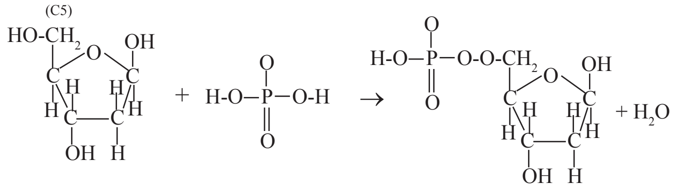
As also show in fig. 3, these bases have the ability to form hydrogen bonds (----). Adenine and thymine have the appropriate atoms with lone pairs, $H^{\delta+}$ atoms and geometry to form two hydrogen bonds per molecule whereas guanine and cytosine have the ability to form 3 such bonds per molecule.

Moreover, since A and T both have one lone pair and one $H^{\delta+}$ atom each with similar geometry they can hydrogen bond to each other via two hydrogen bonds. Similarly, C has two lone pairs and one $H^{\delta+}$ atom while G has one lone pair and two $H^{\delta+}$ atoms with similar geometry, so they can hydrogen bond to each other via three hydrogen bonds. Hence, A and T are said to be a **base pair**. Similarly G and C are another base pair. These pairings are absolutely vital in determining the detailed structure of DNA and also the mechanism by which genetic information is replicated or passed to the next generation.

Deoxyribose Combining with Phosphate

One of the OH groups in a phosphate group (see fig. 2) reacts with the OH group on C5 of 2-deoxyribose to create a covalent bond between the two particles by eliminating water. Hence, this is a **condensation reaction**. The product (see fig. 4) can be seen later as the basic unit for forming the "backbone" of DNA.

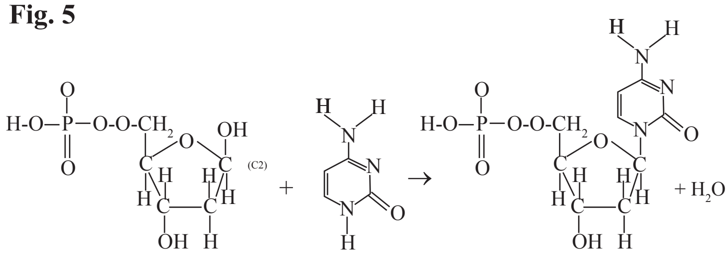
Fig. 4



Nucleotide Formation

The 2-deoxyribose-phosphate unit shown in fig. 4 combines with one of the nitrogenous bases by a second condensation reaction to form a nucleotide. This occurs by eliminating the OH at C1 in the 2-deoxyribose along with the H marked * in fig. 3 to from a covalent link and water. This is represented in fig. 5 for the formation of a cytosine nucleotide.

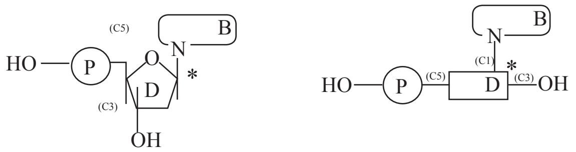
Fig. 5



Note: Four different nucleotides are possible using A, T, G or C.

Note: To simplify matters, a nucleotide will now be represented by the block diagram shown in fig. 6 where P is phosphate, D is 2-deoxyribose and B is the nitrogenous base

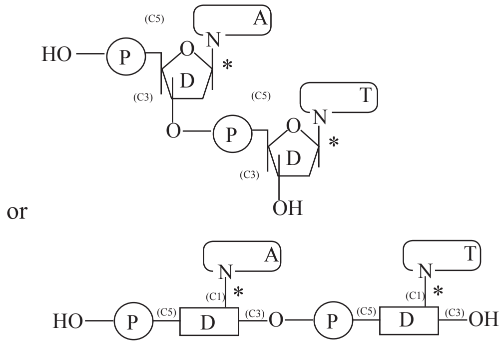
Fig. 6



Polymerisation of Nucleotides

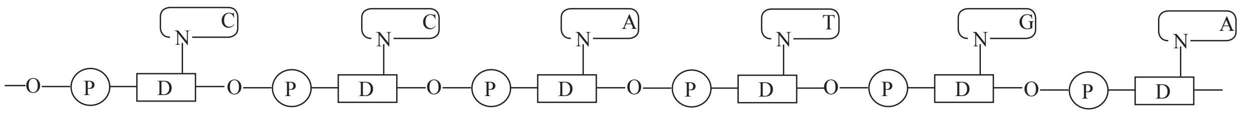
The OH of the phosphate group in one nucleotide and the OH at C3 of the deoxyribose in a second nucleotide undergo yet another condensation reaction to form a dimer as shown in fig. 7 for adenine and thymine nucleotides.

Fig. 7



The OH of the phosphate group in one dimer and the OH at C3 of the deoxyribose in another dimer can then condense over and over again producing a polymer. As shown in fig. 8 the sequence of nucleotides (determined by A, C, T and G) can vary depending on the genetic function of this section of the molecule.

Fig. 8



Thousands and thousands of nucleotides link in this manner to form a single strand of DNA!

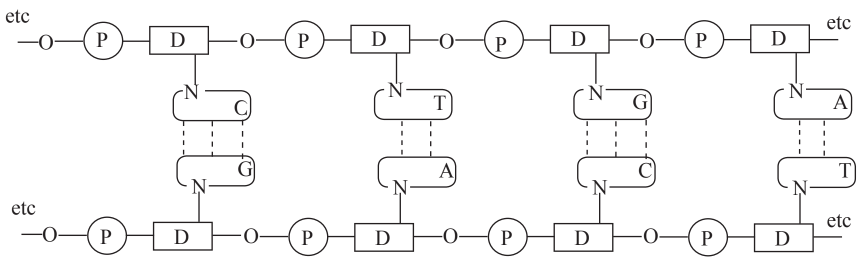
Formation of the Double Strand of DNA

DNA exists as a **double strand** with one linked to the other by hydrogen bonds between the base side-chains. One strand is said to be the **complementary strand** of the other because the composition of one pre-determines the composition of the other. This is because of the hydrogen bonding characteristics of the base pairs discussed earlier.

As illustrated in fig. 9, since guanine (G) and cytosine (C) both form three hydrogen bonds per molecule, any G in one strand is always bonded to a C in the complementary strand. Consequently, C is always bonded to G.

Similarly, since adenine (A) and thymine (T) both form two hydrogen bonds per molecule, any A in one strand is always bonded to a T in the complementary strand. Consequently, T is always bonded to A.

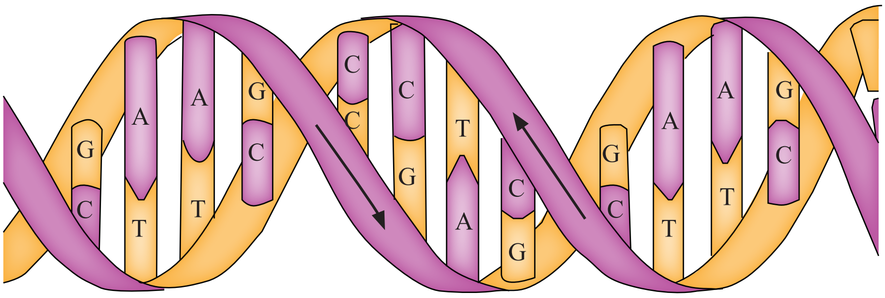
Fig. 9



Double Helix Structure of DNA

So far the DNA double strand is represented as a ladder-like structure with the deoxyribose-phosphate chains as the sides of the ladder and the A-T and G-C base pairs as the rungs of the ladder. However, as discovered by Watson and Crick in 1953, the "ladder" is twisted to form a double helix as shown in fig. 10.

Fig 10



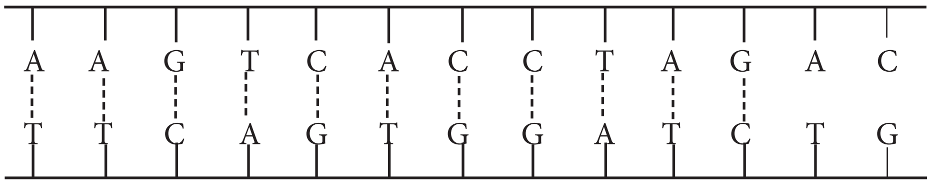
As indicated by the arrows, the two helices coil in opposite directions.

DNA Replication

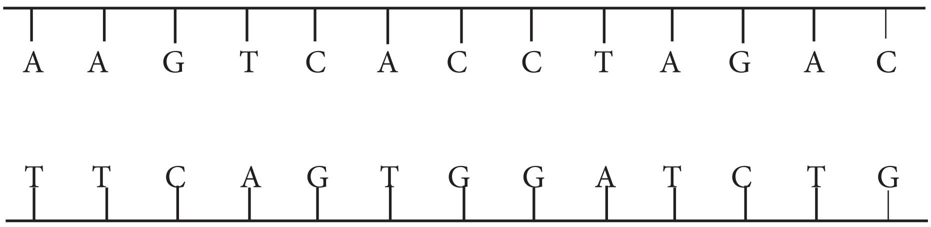
In order to maintain and transmit genetic information, it is essential that a DNA molecule can accurately reproduce itself.
In simple terms, as represented in fig. 11, the double strand shown in 11(a) "unzips" by breaking the hydrogen bonds to form two separate complementary strands as shown in 11(b).

Fig. 11

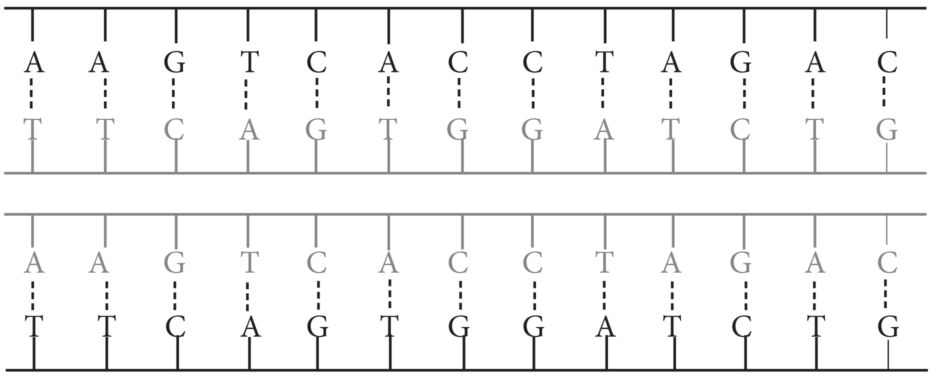
(a)



(b)



(c)



Under the control of enzymes such as DNA polymerase, the complementary strands of both unzipped strands are then built up using "fresh" nucleotides. As always, A is hydrogen bonded to T and G to C resulting in a "copy" of the original double strand.

Acknowledgements: *This Factsheet was researched and written by Mike Hughes.* ***Curriculum Press, Bank House, 105 King Street, Wellington, Shropshire, TF1 1NU.***
 ISSN 1351-5136