Introduction to nucleic acids and their structure

The questions and answers below constitute an introduction to the nucleic acids DNA and RNA. They are all available at **[link]**.

What are nucleic acids?

Nucleic acids are the carriers of genetic information. In all living organisms, the hereditary information is stored in deoxyribonucleic acid (DNA), which is a molecule formed by the repetition of *nucleotides* (making DNA a *polymer*). There are four different nucleotides in DNA, which form a universal code for hereditary information.

Ribonucleic acid (RNA), the other kind of nucleic acid, is a related molecule to DNA. It is also a polymer of four nucleotides, three of which are the same as in DNA while the fourth one is slightly different. It has many functions in cells, notably acting as the intermediate between DNA and proteins. Some viruses even store their genome in the form of an RNA molecule rather than DNA.

What are nucleotides?

Nucleotides are the building blocks of nucleic acids: they are the *monomers* which, repeated many times, form the *polymers* DNA and RNA. Nucleotides are composed of a five-carbon sugar covalently attached to a phosphate group and a base containing nitrogen atoms. Figure 1 shows the structure of the nucleotides making up nucleic acids.

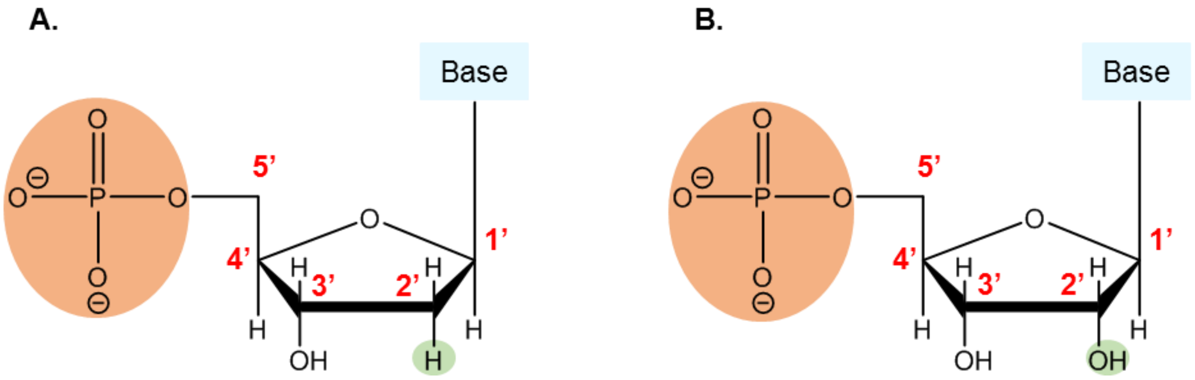


Figure 1 | The chemical structure of a nucleotides. A nucleotide comprises a five-carbon sugar molecule: deoxyribose in DNA (**A**) and ribose in RNA (**B**). The carbon atoms on the sugar molecule are numbered in red. Deoxyribose (**A**) is different from ribose (**B**) in that it lacks an –OH group at carbon 2'. The 5'-carbon atom is attached to a phosphate group (here a monophosphate in orange) and the 1'-carbon is attached to a base (blue).

The main difference between nucleotides from DNA and those from RNA is the nature of the sugar. Nucleotides making up RNA (Figure 1B) contain ribose, making them *ribonucleotides*. In DNA, however, the sugar lacks an –OH group at the 2'-carbon, making it *deoxyribose* and the corresponding nucleotides *deoxyribonuleotides*.

A nucleotide may contain more than one phosphate at its 5'-carbon, for instance the nucleotide adenosine triphosphate has three, as shown in Figure 2. When there is no phosphate group, the molecule is no longer called a *nucleotide*, but a *nucleoside*.

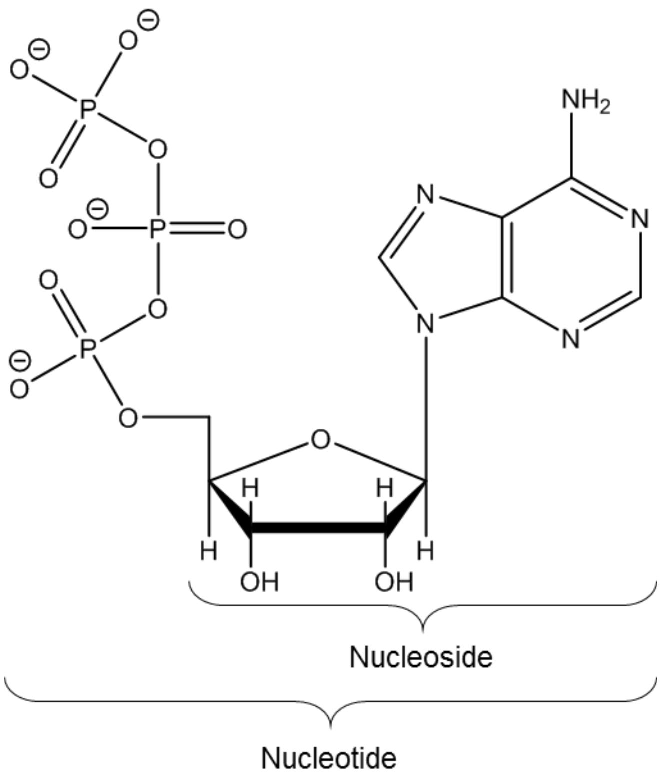
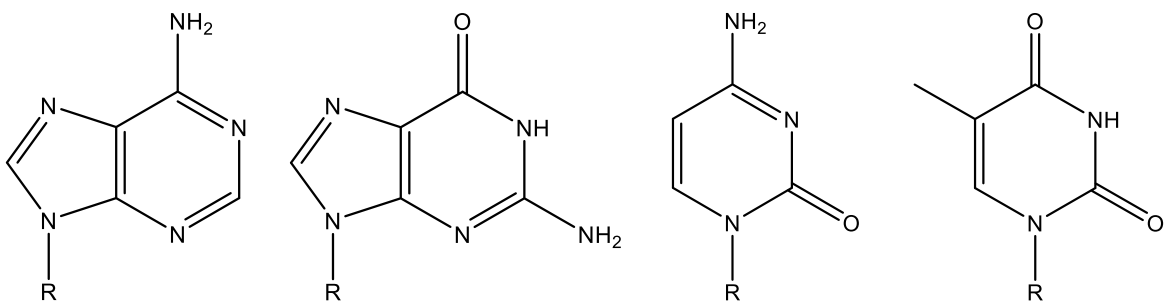


Figure 2 | Adenosine triphosphate, often abbreviated to ATP.

<u>What are the bases in nucleic acids?</u>

The nucleotides making up DNA contain one of four *nitrogenous bases* (i.e. bases that contain nitrogen atoms). From a chemical perspective, two of those bases are *purines*, while the other two are *pyrimidines*. To each base corresponds a name (e.g. adenine), a nucleoside (e.g. adenosine) and a one-letter code (e.g. A). This information is included in Table 1.

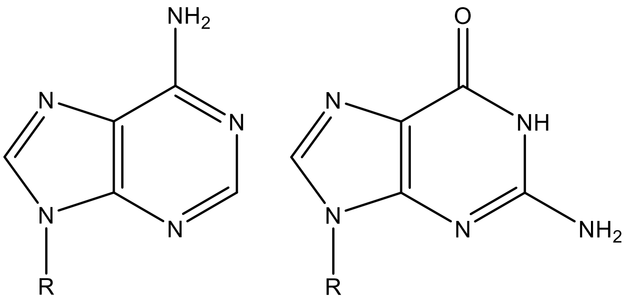
Table 1 | The four bases of DNA. The ‘R’ represents the deoxyribose covalently attached to the base to form the nucleoside named in the third row.



Adenine	Guanine	Cytosine	Thymine
Deoxyadenosine	Deoxyguanosine	Deoxycytidine	Deoxythymidine
A	G	C	T
Purines		Pyrimidines	

As mentioned above, the sugar in RNA is ribose rather than deoxyribose. However, there is another difference between DNA and RNA in the base composition. RNA contains three of the bases found in DNA (adenine, guanine and cytosine) but thymine is replaced by the related base, uracil. The four bases found in RNA, along with the names of their corresponding nucleosides, are in Table 2.

Table 2 | The four bases of RNA. The 'R' represents the ribose covalently attached to the base to form the nucleoside named in the third row.



Adenine	Guanine	Cytosine	Uracil
Adenosine	Guanosine	Cytidine	Uridine
A	G	C	U
Purines		Pyrimidines	

What is the 3D structure of DNA?

DNA is predominantly found as a double helix: two strands of polynucleotides wind about the same axis to form a right-handed helix. Each nucleotide provides a ribose and a phosphate to the backbone. The bases project towards the centre of the helix, away from the surrounding water. The DNA double helix is shown in Figure 3.

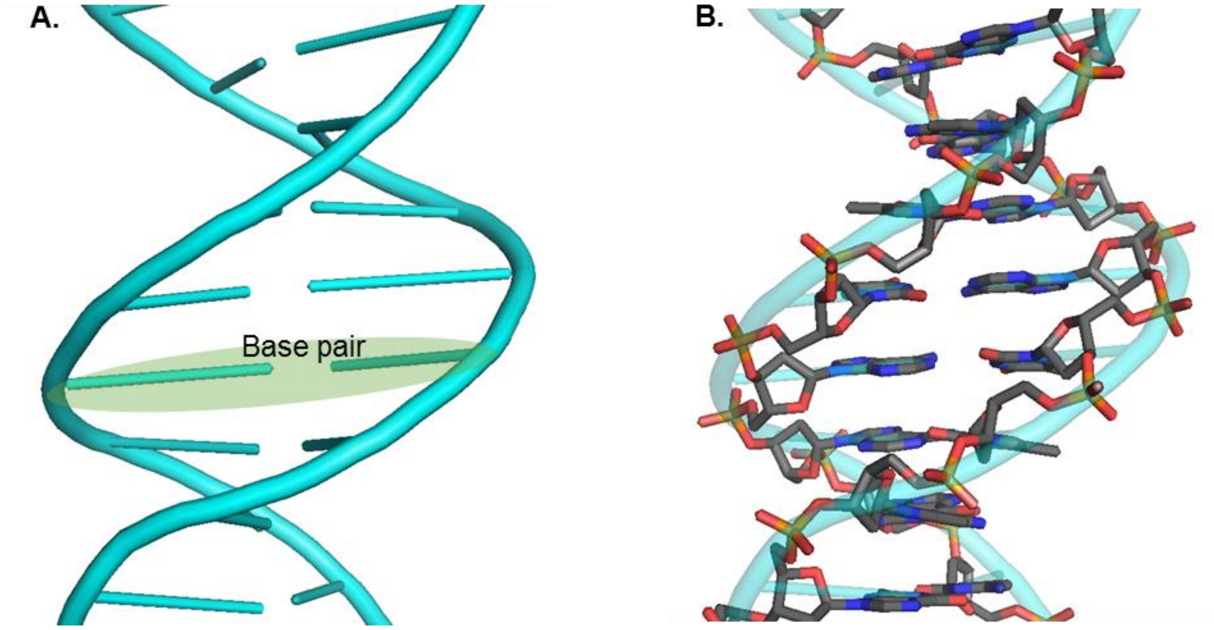


Figure 3 | The double-helical structure of DNA. **A.** DNA shown as a cartoon. **B.** DNA shown as sticks, with a cyan cartoon highlighting the sugar-phosphate backbone. Green: base pair; grey: carbon; red: oxygen; blue: nitrogen; white: hydrogen; orange: sulphur.

Two bases (each from a different strand) come together to form a *base pair*, shown in green in Figure 3A. A base pair is held together by hydrogen bonds between the two bases (cf. Watson-Crick base pairing explained below).

DNA can adopt slightly different kinds of 3D structure, but the majority of the DNA inside a cell at any given point will have the structure shown in Figure 3, called *B-DNA*. It has 10 base pairs per helical turn and a rise of 3.4Å per base pair.

What is Watson-Crick base pairing?

The double helix shown in Figure 3 can only accommodate two kinds of base pairs, due to the geometry of the bases. Adenine and thymine bases always pair with each other while guanine and cytosine bases always pair with each other. This kind of pairing, called *Watson-Crick base pairing*, is mediated by hydrogen bonds between the two bases of a pair, as shown in Figure 4.

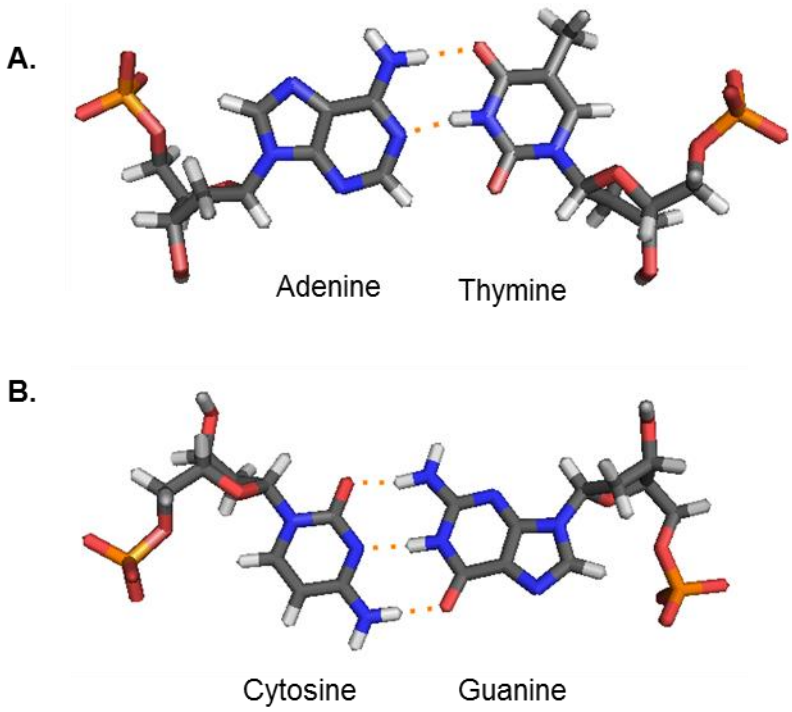


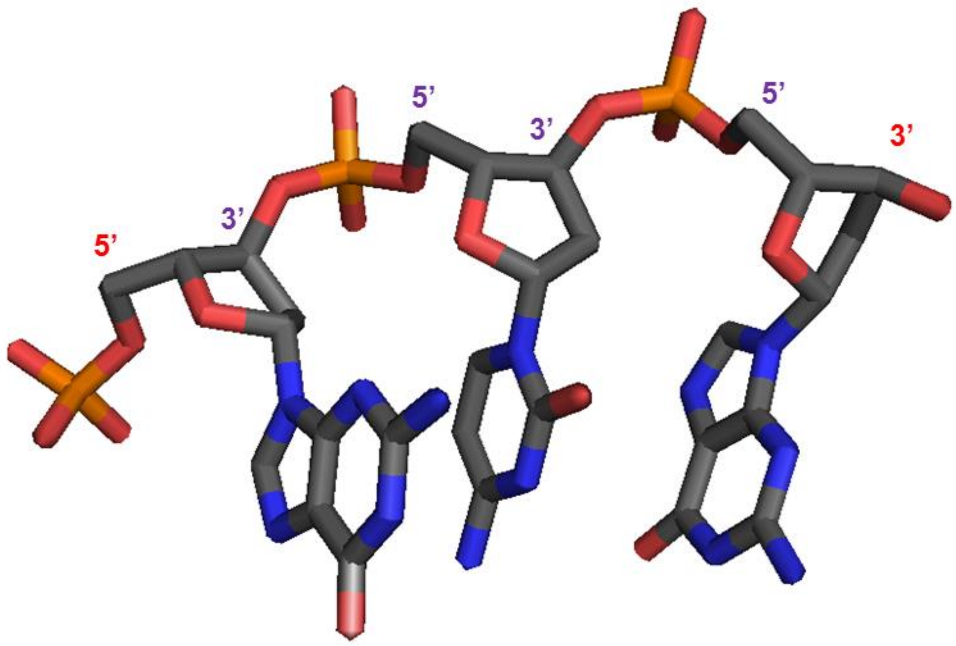
Figure 4 | A. Watson-Crick base pairing between deoxyriboadenosine monophosphate and deoxyribothymidine monophosphate. **B.** Watson-Crick base pairing between deoxyribocytidine monophosphate and deoxyriboguanosine monophosphate. Only the name of the base is given below each nucleotide. The hydrogen bonds are shown by orange dotted lines. Grey: carbon; red: oxygen; blue: nitrogen; white: hydrogen; orange: sulphur.

Note that an AT base pair is only held by two hydrogen bonds whereas a CG base-pair has three, making the latter more stable.

What is the directionality of DNA?

A strand of DNA is the result of the polymerisation of several nucleotides, with the backbone formed by the deoxyribose sugars and the phosphate groups. Each nucleotide *residue* (i.e. a nucleotide within a strand of DNA) contains a phosphate group covalently attached to the 5'-carbon of its deoxyribose, but also has its deoxyribose 3'-carbon covalently attached to the phosphate of the next nucleotide residue in the strand. The only exception is the final nucleotide, which does not have a phosphate at its 3'-carbon (of the deoxyribose), but rather a free –OH group. We define this end of the strand as the 3'-end. The very first nucleotide residue, on the other hand, has a free phosphate group attached to its 5'-carbon. We define that end of the strand as the 5'-end.

DNA is always read from the 5'-end to the 3'-end, as shown in Figure 5.



Sequence: **5’-G-C-G-3’**

Figure 5 | The directionality of DNA. A stretch of 3 nucleotide residues is shown with their 5’- and 3’-carbons numbered. In red are the 5’-end (characterised by a free phosphate group) and the 3’-end (characterised by a free –OH group).

Note that, when studying DNA in the lab, it is common to remove the phosphate at the 5’-end, therefore many experimentally determined structures will actually show an –OH group rather than a phosphate at the 5’-end.

How does the sructure of RNA differ from that of DNA?

As mentioned above, RNA is made of ribonuleotides rather than deoxyribonucleotides: the 2’-carbon of its ribose is covalently attached to an –OH group. Furthermore, RNA contains the base uracil instead of thymine.

The other main difference between RNA and DNA is that RNA is often single-stranded and does not form the regular double-helical structure of DNA. However, it is quite common for a single RNA strand to fold on itself and to form complex 3D structures, with some helical character. When that is the case, the 3D structure is often stabilised by the same Watson-Crick base-pairing as in DNA, although some deviations may be allowed (often disrupting helices).

The directionality of RNA, however, is the same as that of DNA: the sequence is read from the 5’-end to the 3’-end.