

Nucleic acids

Revision notes • IB Biology A1.2 Nucleic acids: nucleotides, DNA and RNA structure, base pairing, the genetic code, and HL directionality, nucleosomes and key evidence.

For IB Diploma Biology, Standard Level and Higher Level (first assessment 2025). **HL only** is additional higher level content, for HL students only.

Nucleic acids Spec A1.2

DNA as the genetic material and nucleotides Spec A1.2.1, A1.2.2, A1.2.3, A1.2.4

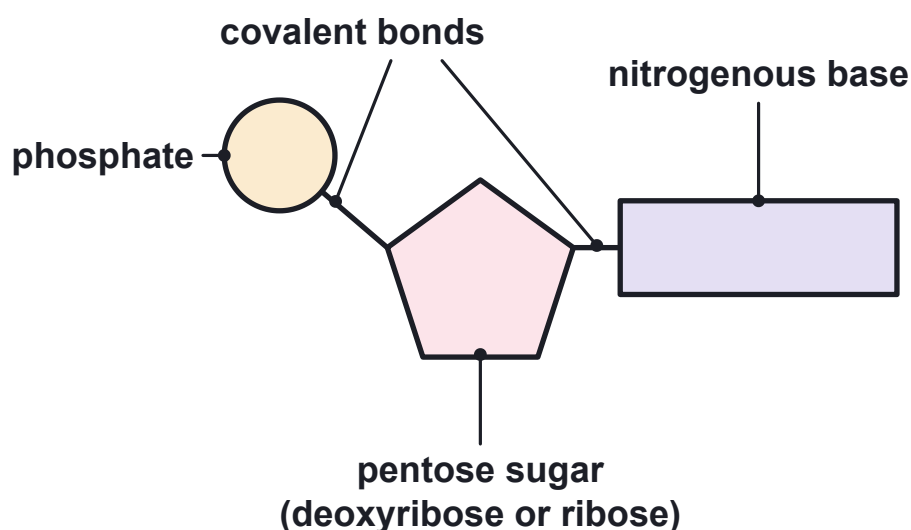
DNA as the genetic material

DNA is the genetic material of all living organisms, from bacteria to plants and animals. Some viruses use RNA as their genetic material, for example influenza virus and HIV, but viruses are not considered to be living.

Components of a nucleotide

Nucleic acids are polymers of **nucleotides**. Every nucleotide has three parts, joined by covalent bonds:

- a **phosphate** group (drawn as a circle)
- a **pentose sugar**, a five-carbon sugar: deoxyribose in DNA, ribose in RNA (drawn as a pentagon)
- a **nitrogenous base** (drawn as a rectangle), attached to the sugar on the opposite side from the phosphate



[\(opens full size in a new tab\)](#)

A nucleotide: circle = **phosphate**, pentagon = **pentose sugar**, rectangle = **base**.

The sugar-phosphate backbone

In a strand of DNA or RNA, the phosphate of one nucleotide is covalently bonded to the sugar of the next. This **sugar-phosphate bonding** makes a continuous chain of covalently bonded atoms along each strand, which forms a strong **backbone**. The bases stick out from the backbone.

Bases form the basis of a code

Nitrogenous bases in DNA and RNA

Nucleic acid	Bases
DNA	adenine (A), thymine (T), guanine (G), cytosine (C)
RNA	adenine (A), uracil (U), guanine (G), cytosine (C)

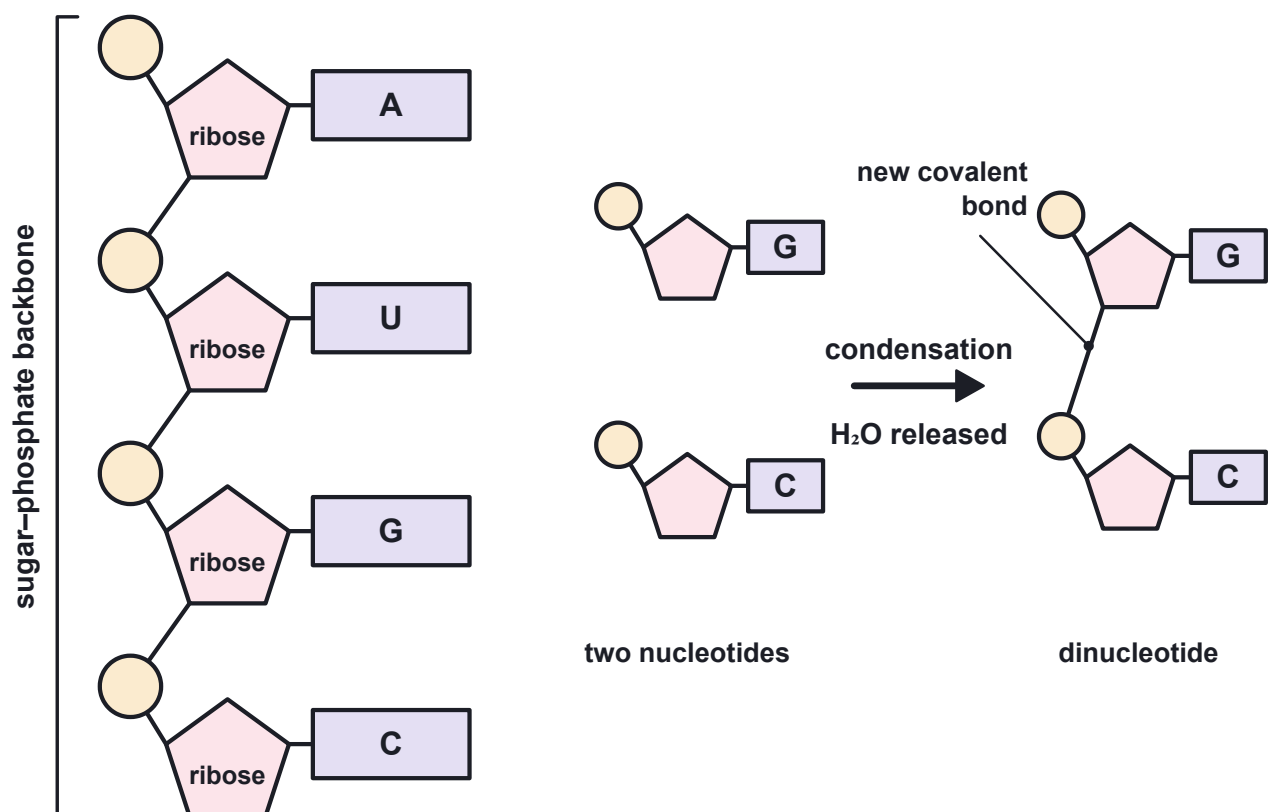
The backbone is the same in every molecule. What varies is the **sequence of bases**, and it is this sequence that stores genetic information, in the way a sequence of letters stores the meaning of a sentence.

Exam tip: You need the full names of the five bases: adenine, thymine, guanine, cytosine and uracil. Spelling matters in written answers.

RNA polymers and the DNA double helix Spec A1.2.5, A1.2.6

RNA as a polymer

RNA is a **polymer formed by condensation of nucleotide monomers**. When two nucleotides are joined, the phosphate of one bonds covalently to the sugar (ribose) of the other and a molecule of water is released. Repeating this builds a single strand of any length, with a sugar-phosphate backbone and one base on each nucleotide.



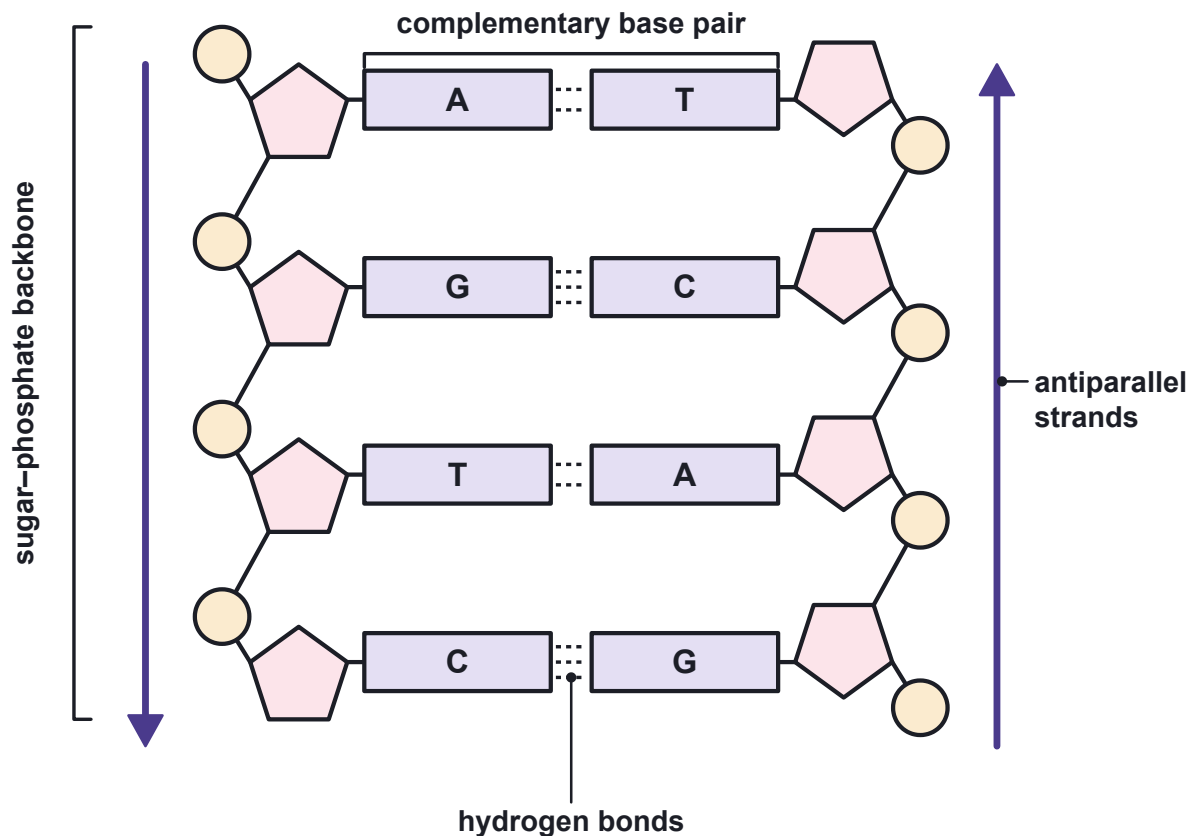
[\(opens full size in a new tab\)](#)

RNA is a polymer of nucleotides. Each new sugar-phosphate bond forms by **condensation**, releasing a molecule of water.

DNA as a double helix

- DNA has **two strands** of nucleotides, wound into a **double helix**.
- The two strands are **antiparallel**: they run in opposite directions.
- The bases face inwards and the two strands are linked by **hydrogen bonds** between bases.

- Pairing is complementary: adenine (A) pairs with thymine (T) and guanine (G) pairs with cytosine (C).



[\(opens full size in a new tab\)](#)

DNA: two antiparallel strands (one drawn upside down), linked by hydrogen bonds between complementary base pairs, A with T and G with C.

Exam tip: In DNA diagrams draw the two strands antiparallel (one strand upside down) and pair A with T and G with C. You do not need to draw the helix, the relative sizes of the bases or the number of hydrogen bonds.

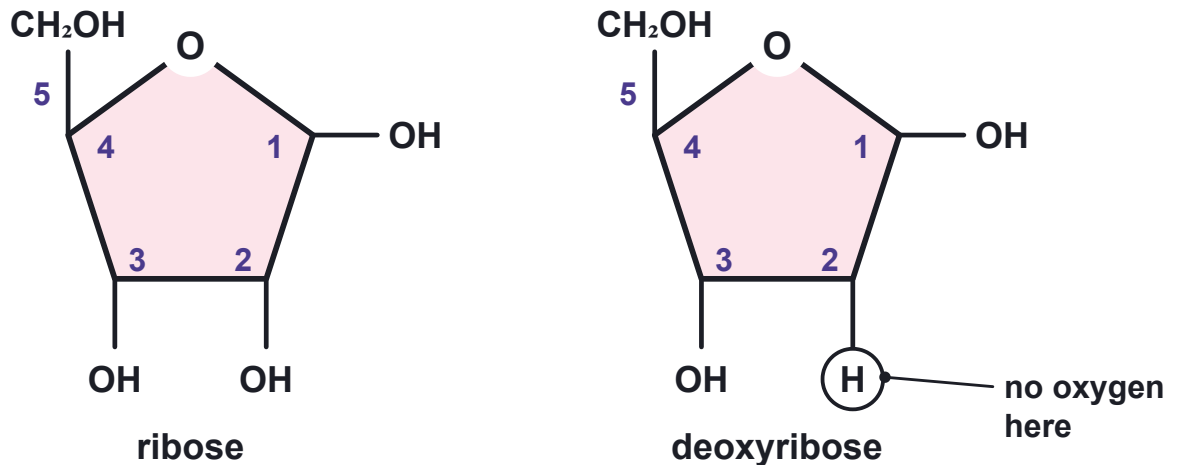
Exam tip: Linking question: how can polymerization result in emergent properties? A single nucleotide stores no information, but a polymer of thousands of them stores a whole gene.

Differences between DNA and RNA Spec A1.2.7

Differences between DNA and RNA		
Feature	DNA	RNA
Number of strands	Two (double helix)	One
Nitrogenous bases	A, T, G, C	A, U, G, C (uracil instead of thymine)
Pentose sugar	Deoxyribose	Ribose

Ribose and deoxyribose

Both sugars are pentoses with a five-membered ring containing an oxygen atom. The difference is on carbon 2: ribose has an OH group there, while deoxyribose has only H, so it has one oxygen atom fewer ('deoxy' means without oxygen).



[\(opens full size in a new tab\)](#)

Ribose has an **OH** on carbon 2; deoxyribose has only **H** there, so it has one oxygen fewer.

Examples of nucleic acids

- **DNA:** the genetic material, found in chromosomes in the nucleus, and in mitochondria, chloroplasts and the cytoplasm of prokaryotes.
- **Messenger RNA (mRNA):** a copy of a gene, carried to ribosomes.
- **Transfer RNA (tRNA):** carries amino acids to the ribosome during translation.
- **Ribosomal RNA (rRNA):** part of the structure of ribosomes.

Common mistake: In a compare question, write each difference as a matched pair, such as 'DNA contains deoxyribose but RNA contains ribose'. Stating only one side rarely scores.

Base pairing, information storage and the genetic code Spec A1.2.8, A1.2.9, A1.2.10

Complementary base pairing

Complementarity is based on hydrogen bonding: A and T (or A and U in RNA) form hydrogen bonds with each other, and so do G and C. Other combinations do not fit and do not bond properly. So the sequence on one strand fixes the sequence on the other.

- **Replication:** the two strands separate and each acts as a template. Free nucleotides pair with complementary bases, so two identical DNA molecules are produced.
- **Transcription** (expression): an mRNA molecule is built with bases complementary to one strand of a gene, with U pairing with A.
- **Translation** (expression): each tRNA anticodon pairs with a complementary mRNA codon, so amino acids are added in the order the code specifies.

Limitless capacity for storing information

There is no restriction on the order of bases along a strand, and a DNA molecule can be any length. So **any base sequence is possible** and the diversity of possible sequences is effectively limitless. With 4 bases, the number of possible sequences of length n is 4^n . DNA also stores data with great economy: the human genome of about 3 billion base pairs fits inside a nucleus only a few micrometres across.

Number of possible base sequences

How many different base sequences are possible for a DNA strand 10 nucleotides long?

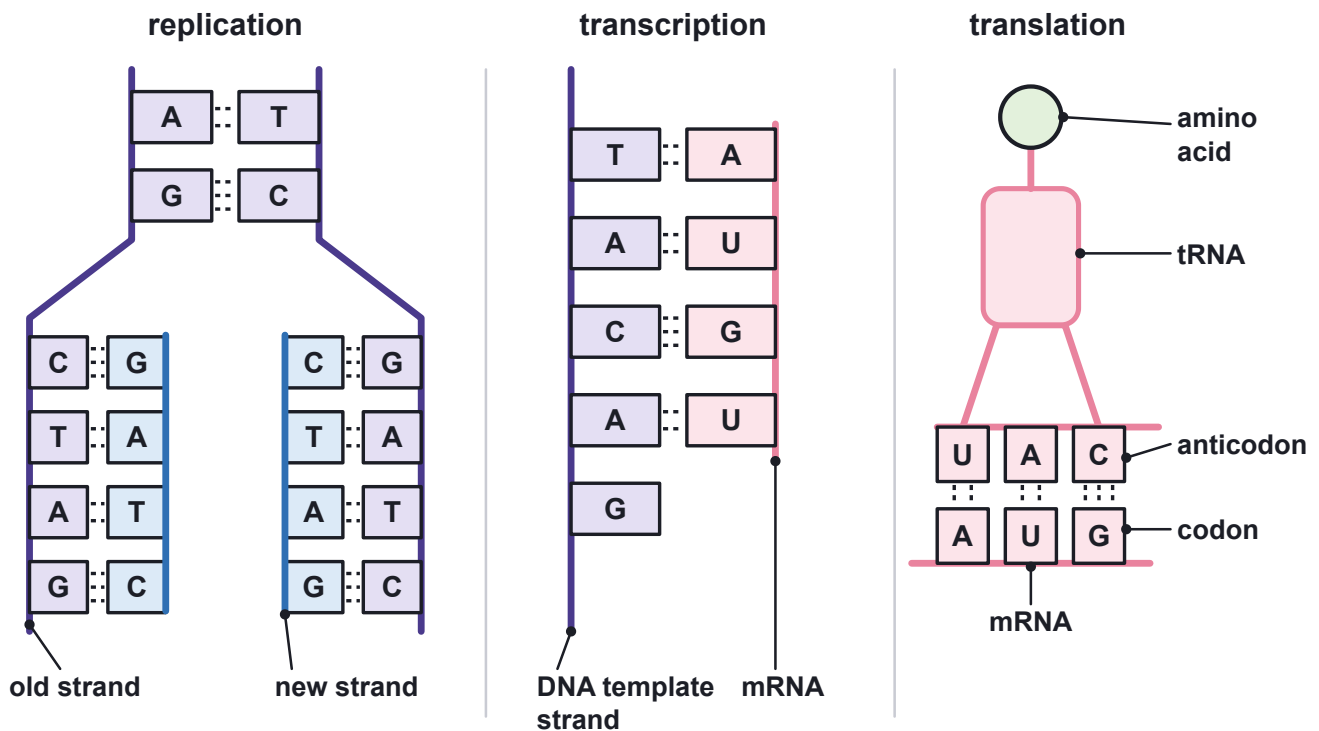
1. Each position can hold any one of 4 bases.
2. Number of sequences = $4 \times 4 \times \dots$ (10 times) = 4^{10} .
3. $4^{10} = 1\,048\,576$.

Answer: About one million (1 048 576) sequences, from just 10 nucleotides.

Maths skill: Possible sequences = 4^n for DNA of length n nucleotides. Each extra nucleotide multiplies the number of possibilities by 4.

A universal genetic code

The genetic code, the way base triplets (codons) correspond to amino acids, is **conserved across all life forms**, with only a few minor variations. It is very unlikely that so many organisms would arrive at the same arbitrary code independently. The best explanation is that all living organisms inherited it from a **universal common ancestor**. This is also why a gene from one species can be expressed in another.



[\(opens full size in a new tab\)](#)

Complementary base pairing by **hydrogen bonds** in replication, transcription and translation.

Directionality and purine-pyrimidine pairing Spec A1.2.11, A1.2.12 HL only

Directionality of RNA and DNA

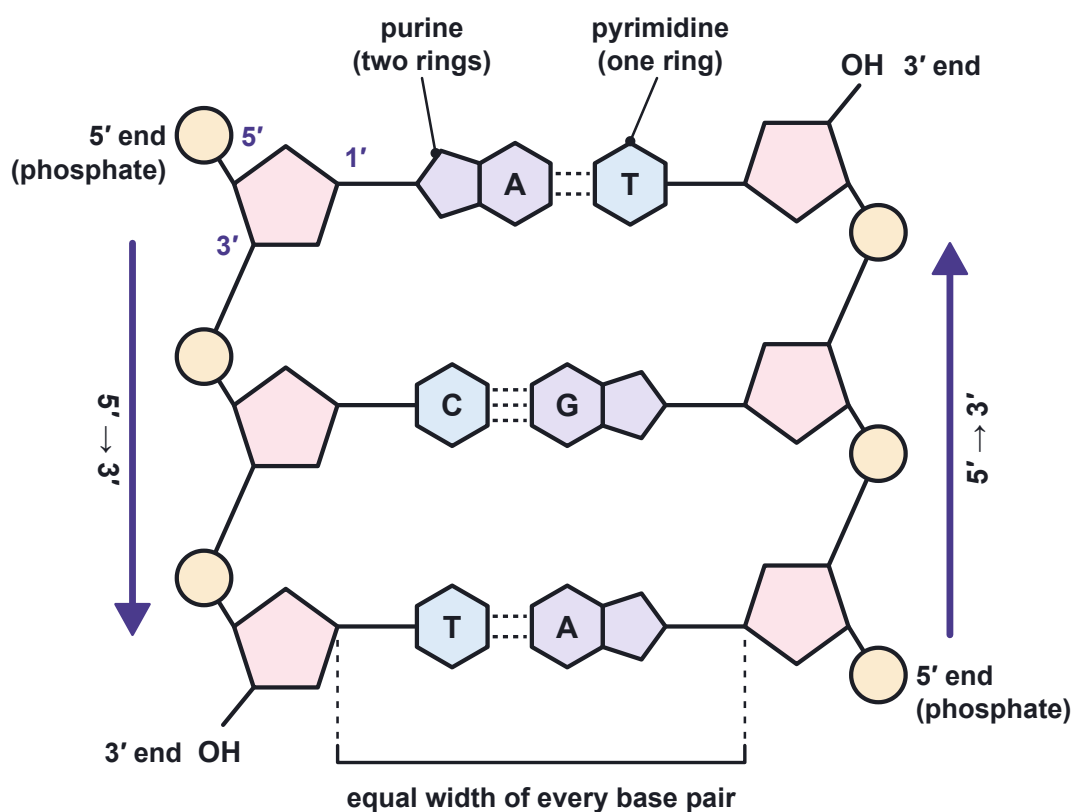
The carbons of the pentose sugar are numbered 1' to 5'. The base is attached to carbon 1', the phosphate of the same nucleotide to carbon 5', and the phosphate of the next nucleotide to carbon 3'. So every link in the backbone runs **from carbon 3' of one sugar to carbon 5' of the next**. One end of a strand has a free phosphate on carbon 5' (the **5' end**) and the other a free OH on carbon 3' (the **3' end**). In DNA the two strands run in opposite directions: this is what antiparallel means.

Why directionality matters

Process	Significance of 5' to 3'
Replication	DNA polymerase can only add nucleotides to the 3' end of a growing strand, so new DNA is built 5' to 3'
Transcription	RNA polymerase builds mRNA 5' to 3', so it moves along the template strand reading it 3' to 5'
Translation	Ribosomes read mRNA from the 5' end towards the 3' end, codon by codon

Purine-to-pyrimidine bonding

Adenine and guanine are **purines**, with two rings. Cytosine, thymine and uracil are **pyrimidines**, with one ring. Each base pair in DNA is one purine bonded to one pyrimidine (A-T and C-G), so all base pairs have equal length. The two backbones are therefore always the same distance apart and the DNA helix has the same three-dimensional structure, regardless of the base sequence. This uniformity contributes to the stability of the helix.



[\(opens full size in a new tab\)](#)

Each strand runs **5' to 3'**, in opposite directions. Every base pair is a purine (two rings) with a pyrimidine (one ring), so all base pairs have **equal width**.

Common mistake: A purine never pairs with a purine. Two purines would be too wide to fit between the backbones and two pyrimidines too narrow.

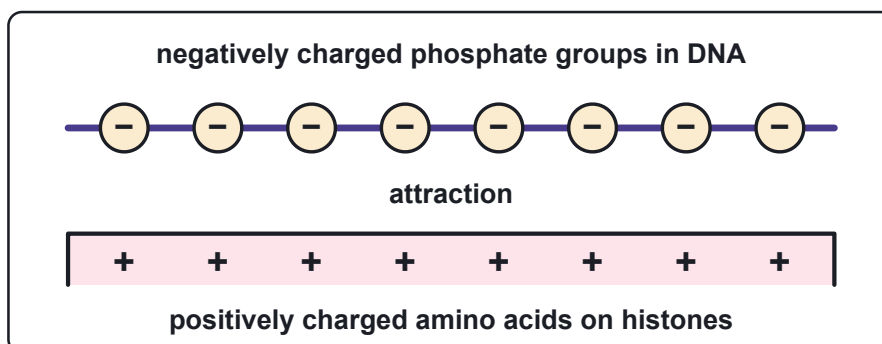
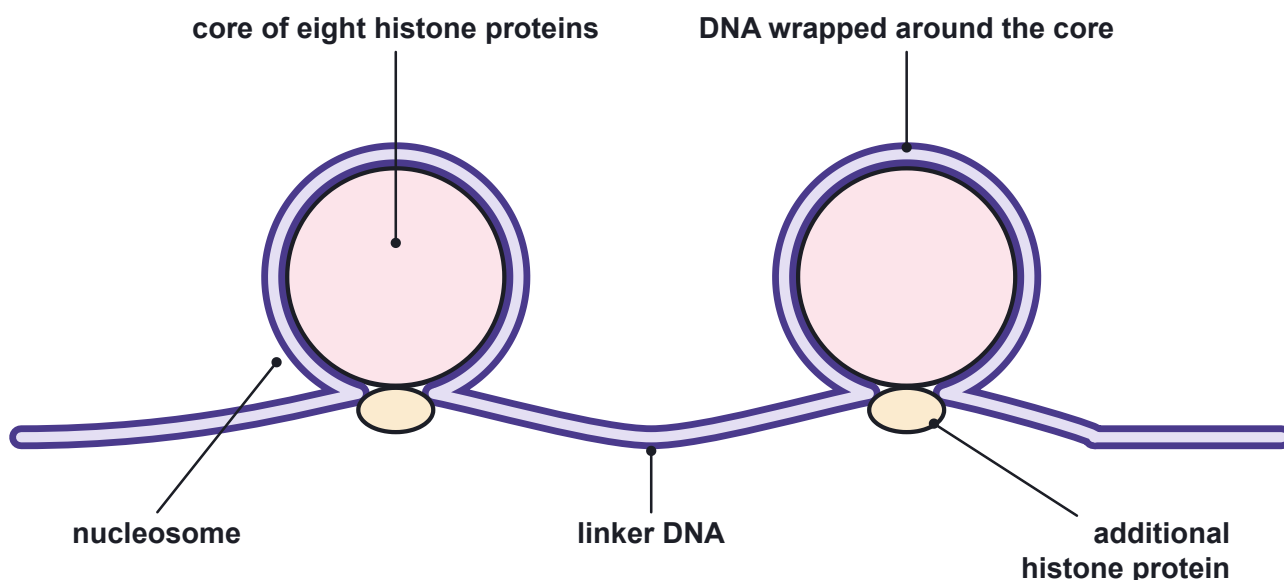
Structure of a nucleosome Spec A1.2.13 HL only

In eukaryotes, DNA in the nucleus is associated with proteins called **histones**. The basic unit of packaging is the **nucleosome**.

- A core of **eight histone proteins**.
- A length of DNA **wrapped around** the core.

- An **additional histone protein** attached to the **linker DNA** where it enters and leaves, holding the DNA in place.
- Linker DNA continues to the next nucleosome, like beads on a string.

The histones have many positively charged amino acids on their surface. These are attracted to the **negatively charged phosphate groups** in the DNA backbone, which is why DNA binds so tightly around the histone core.



[\(opens full size in a new tab\)](#)

Nucleosomes: DNA wrapped around a core of eight histones, held by an additional histone and joined by **linker DNA**. Positive histones attract negative DNA phosphates.

Practical skill: Use molecular visualization software (for example JSmol with the nucleosome core particle structure, PDB entry 1A0I) to study how the protein and DNA associate. Colour the histones and DNA separately and look for positive amino acids next to the DNA phosphates.

Hershey-Chase experiment and Chargaff's data Spec A1.2.14, A1.2.15 HL only

The Hershey-Chase experiment

By the early 1950s it was still debated whether genes were made of protein or DNA. Hershey and Chase used **T2 bacteriophage**, a virus made only of a protein coat and DNA, which infects the bacterium *Escherichia coli*. Proteins contain sulfur but almost no phosphorus; DNA contains phosphorus but no sulfur.

1. One batch of viruses was grown with radioactive sulfur (^{35}S), which labelled their **protein coats**. Another batch was grown with radioactive phosphorus (^{32}P), which labelled their **DNA**.

2. Each batch was allowed to infect bacteria.
3. The mixture was agitated in a blender to detach the virus coats from the outside of the bacteria.
4. It was centrifuged: the heavier bacteria formed a pellet and the lighter virus coats stayed in the supernatant liquid.
5. The radioactivity of the pellet and the supernatant was measured.

Results and conclusion		
Label	Where most radioactivity was found	What it shows
³⁵ S (protein)	Supernatant	Protein stayed outside the bacteria
³² P (DNA)	Pellet (inside bacteria)	DNA entered the bacteria

Because the infected bacteria went on to produce new viruses, the material that entered them must carry the instructions. Only DNA entered, so **DNA is the genetic material**.

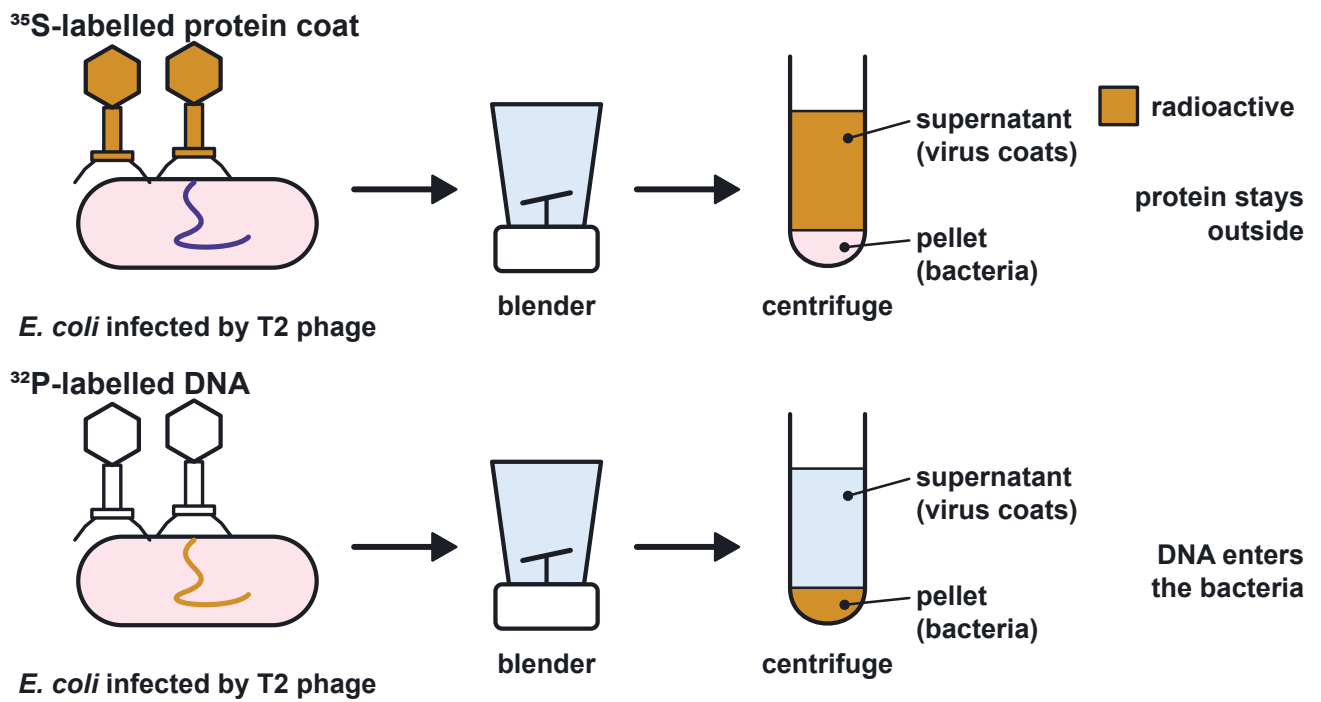
Exam tip: Nature of science: technological developments open up new experiments. The Hershey–Chase experiment only became possible once radioisotopes were available to scientists as research tools.

Chargaff's data

Chargaff measured the relative amounts of the four bases in DNA from diverse life forms. He found that in every species the amount of **adenine was about equal to thymine** and guanine about equal to cytosine, so purines equalled pyrimidines. But the proportion of A + T compared with G + C **varied between species**.

The earlier **tetranucleotide hypothesis** proposed that DNA was a repeating sequence of the four bases, which would mean all four bases are present in equal amounts in every species. Chargaff's data did not fit, so they falsified it, and the idea that DNA was too simple to carry genetic information was abandoned. The A = T and G = C pattern later supported complementary base pairing in the double helix.

Exam tip: Nature of science: the problem of induction is that no number of agreeing observations can prove a general theory true. Falsification is more certain: one reliable set of contrary data, such as Chargaff's, is enough to show a hypothesis is false.



DNA enters the bacteria and is the genetic material

[\(opens full size in a new tab\)](#)

Hershey-Chase: ^{35}S (protein) stayed in the **supernatant**, ^{32}P (DNA) went into the **pellet**, so DNA enters the bacteria and is the genetic material.

Quick check

1. What are the three components of a nucleotide?
2. State three differences between DNA and RNA.
3. What does antiparallel mean in DNA?
4. **HL only** Why does an A-T pair have the same width as a C-G pair?
5. **HL only** Which isotope labelled DNA in the Hershey-Chase experiment?

Quick check answers

Nucleic acids

1. A phosphate group, a pentose sugar and a nitrogenous base.
2. DNA has two strands, RNA one; DNA has thymine, RNA uracil; DNA has deoxyribose, RNA ribose.
3. The two strands run in opposite directions.
4. Each pair is one purine (two rings) bonded to one pyrimidine (one ring).
5. Radioactive phosphorus, ^{32}P .