MEGA LECTURE

IB DP BIOLOGY

Theme A: Unity and Diversity

A1.2 Nucleic Acids

Revision Notes · Standard and Higher Level

Fahad H. Ahmad

+92 323 509 4443 | Megalecture.com

Original notes prepared for the IB Diploma Programme Biology course (first assessment 2025)

What the syllabus requires

A1.2 is about how living things **store and transmit information**. The central idea is that a simple, repeating chemical structure – a chain of nucleotides – can encode vast amounts of information in the **order** of its bases, and that the shape of DNA makes that information both stable and copyable. Use this list as a final checklist before the exam.

Understanding	You should be able to...
Nucleic acids as information molecules	Explain that DNA and RNA store genetic information in the sequence of their bases.
Nucleotides as monomers	Draw/label a nucleotide as pentose sugar + phosphate + nitrogenous base, and show how they polymerise.
The sugar-phosphate backbone	Describe phosphodiester (condensation) linkage between nucleotides forming a directional strand.
The nitrogenous bases	Name the purines and pyrimidines and state which occur in DNA and in RNA.
DNA double helix	Describe two antiparallel strands, complementary base pairing (A–T, G–C) and hydrogen bonding.
RNA structure	Describe RNA as single-stranded, with ribose and uracil, and outline mRNA, tRNA and rRNA.
Structure and function	Relate base sequence, complementary pairing and stability to storage, replication and information transfer.
Chargaff's data **(HL)**	Use %A = %T and %G = %C in calculations and interpret base composition between species.
Evidence and directionality **(HL)**	Outline evidence that DNA is the genetic material and the significance of antiparallel strands.

> **Exam note:** The core of A1.2 is common to SL and HL. Points flagged **(HL)** – Chargaff calculations, base-composition evidence and the antiparallel/directionality detail – are additional depth that only HL papers assess, though SL students benefit from understanding them.

1. Nucleic acids as information molecules

Nucleic acids are the molecules that carry genetic information in every living organism. There are two types: **deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**. Both are **polymers** – long chains built from many repeating subunits called **nucleotides** (the monomers).

The information is not stored in the sugar or phosphate parts, which repeat unchanged along the whole chain, but in the **sequence (order) of the nitrogenous bases**. Just as the 26 letters of an alphabet can be arranged to spell any word, an alphabet of only four bases, arranged in different orders, can encode the instructions to build every protein an organism needs. A gene is simply a length of DNA whose base sequence carries the code for one product.

Why a polymer is ideal for information storage

Where nucleic acids are found, and the scale of the information

In **eukaryotic** cells the main store of DNA is in the **nucleus**, wound around proteins and packaged into chromosomes; smaller amounts of DNA are also found in **mitochondria** and (in plants) **chloroplasts**. In **prokaryotes** the DNA is a single circular molecule in the cytoplasm. RNA is made in the nucleus but functions mainly in the cytoplasm, at the ribosomes. The amount of information stored is remarkable: a single human cell contains roughly 3 billion base pairs of DNA, which if unravelled would stretch about two metres, yet it is packed into a nucleus only a few micrometres across. It is the same universal four-base system — A, T (or U), G and C — in every organism on Earth, from bacteria to whales, which is powerful evidence that all life shares a common ancestry.

> **Key term:** A **monomer** is a single repeating unit; a **polymer** is a large molecule built by joining many monomers. Nucleotides are the monomers of the nucleic-acid polymers DNA and RNA.

2. The structure of a nucleotide

Every nucleotide is made of **three components** joined by covalent bonds:

Component	Description	Role
Pentose sugar	A 5-carbon sugar: **deoxyribose** in DNA, **ribose** in RNA	Forms the backbone; its carbons are numbered 1′ to 5′
Phosphate group	A negatively charged group derived from phosphoric acid	Links sugars together; gives the backbone its acidity
Nitrogenous base	One of several ring-shaped, nitrogen-containing molecules	The variable part — its identity carries the information

The sugar sits at the centre. The phosphate group is attached to the **5**′ carbon of the sugar, and the nitrogenous base is attached to the **1**′ carbon. The carbon numbering (1′ to 5′, read “one-prime” to “five-prime”) matters because it defines the direction of the whole strand, as we shall see.

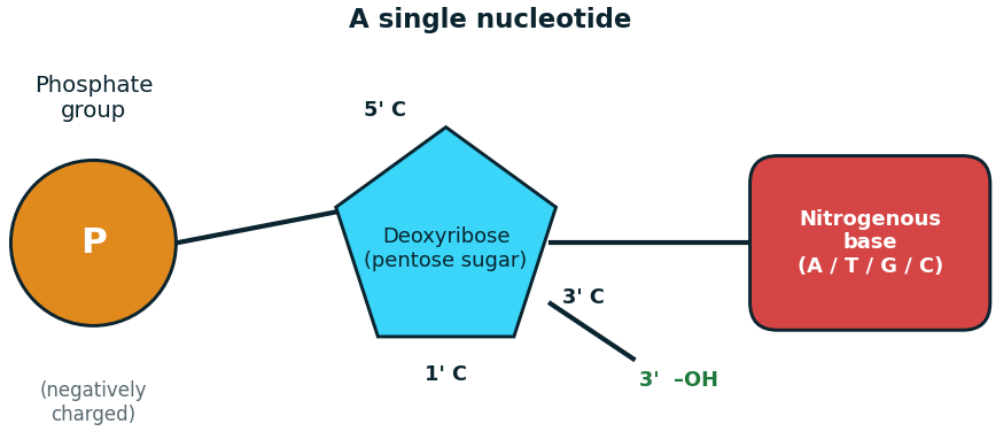

Figure 1. A single nucleotide: a phosphate group, a pentose (deoxyribose) sugar and a nitrogenous base. The phosphate joins the 5′ carbon; the base joins the 1′ carbon; a free –OH sits on the 3′ carbon.

Joining nucleotides: the sugar-phosphate backbone

Nucleotides are joined by a **condensation reaction** (a water molecule is removed). The phosphate group on the **5′** carbon of one nucleotide bonds to the **3′** carbon (specifically its –OH group) of the next nucleotide's sugar. This covalent linkage is called a **phosphodiester bond**.

As nucleotide after nucleotide is added, the alternating sugar–phosphate–sugar–phosphate units form a continuous strand: the **sugar-phosphate backbone**. The bases project sideways from this backbone. Because the backbone is held together by strong covalent phosphodiester bonds, the sequence of bases is chemically very stable.

Directionality: the 5′ and 3′ ends

A single strand is **directional**. At one end there is a free phosphate group on a 5′ carbon — this is the **5′ end**. At the other end there is a free –OH group on a 3′ carbon — the **3′ end**. By convention a base sequence is written and read in the 5′ → 3′ direction, and new nucleotides can only be added to the 3′ end during replication.

> **Condensation vs hydrolysis:** Building the chain (joining nucleotides) is **condensation** and **releases** water; breaking the chain is **hydrolysis** and **uses** water. The same logic applies to all biological polymers (proteins, polysaccharides, nucleic acids).

Labelling a nucleotide in the exam

A common question shows a nucleotide or a short strand and asks you to label the parts. Look for three things: the **phosphate** (often drawn as a circle marked P), the **pentose sugar** (a five-sided ring), and the **nitrogenous base** (a ring or double ring projecting from the sugar). In a strand, identify the alternating sugar–phosphate backbone and note which end carries the free phosphate (5′) and which the free –OH (3′). Marks are lost by muddling the sugar with the base, or by labelling the whole unit “base” when it is a nucleotide. Practise drawing the three parts as three linked shapes until it is automatic.

3. The nitrogenous bases

There are **five** nitrogenous bases in total. They fall into two structural groups distinguished by the number of rings in the molecule.

Group	Ring structure	Bases	Found in
Purines	Double ring (larger)	Adenine (A), Guanine (G)	DNA and RNA
Pyrimidines	Single ring (smaller)	Cytosine (C), Thymine (T), Uracil (U)	C in both; T in DNA only; U in RNA only

So **four** bases appear in any one nucleic acid:

Adenine, guanine and cytosine are shared by both DNA and RNA. The single difference in the base set is that **thymine (DNA) is replaced by uracil (RNA)**. Thymine and uracil are chemically very similar (thymine is essentially uracil with an extra methyl group), and both pair with adenine.

> **Memory hooks:** Pur**A**-**G** (purines are Adenine and Guanine, the “pure silver, Ag” pair). **CUT** the **py** – **C**ytosine, **U**racil, **T**hymine are the pyrimidines. A base pair is always one purine + one pyrimidine, keeping the helix a constant width.

4. The structure of DNA

DNA is a **double helix**: two polynucleotide strands wound around a common axis into a spiral, resembling a twisted ladder. Its structure has several defining features, each with a functional purpose.

Two antiparallel strands

The two strands run in **opposite directions** – they are **antiparallel**. One strand runs 5′ → 3′ while its partner runs 3′ → 5′. Where one strand has its 5′ end, the other has its 3′ end. This opposite orientation is essential for the bases to line up and pair correctly.

Backbones outside, bases inside

The two sugar-phosphate backbones lie on the **outside** of the helix, like the two uprights of a ladder, shielding the information. The nitrogenous bases point **inwards** towards the centre, forming the **rungs** of the ladder, where they pair with bases on the opposite strand.

Complementary base pairing

The bases pair according to strict rules: **adenine always pairs with thymine (A–T)** and **guanine always pairs with cytosine (G–C)**. This is **complementary base pairing**. Each pair is one purine joined to one pyrimidine, so every rung is the same width and the helix stays uniform.

The two bases in a pair are held together by **hydrogen bonds**:

Base pair	Type of pairing	Hydrogen bonds	Note
A–T	Purine + pyrimidine	**2** hydrogen bonds	In RNA, A pairs with U (also 2 bonds)
G–C	Purine + pyrimidine	**3** hydrogen bonds	Stronger; more G–C means a more stable helix

Each hydrogen bond is weak on its own, but along a whole molecule there are millions of them, so together they hold the two strands securely. Crucially, they are still weak enough to be “unzipped” by enzymes when the DNA must be copied or read – a perfect compromise between stability and access.

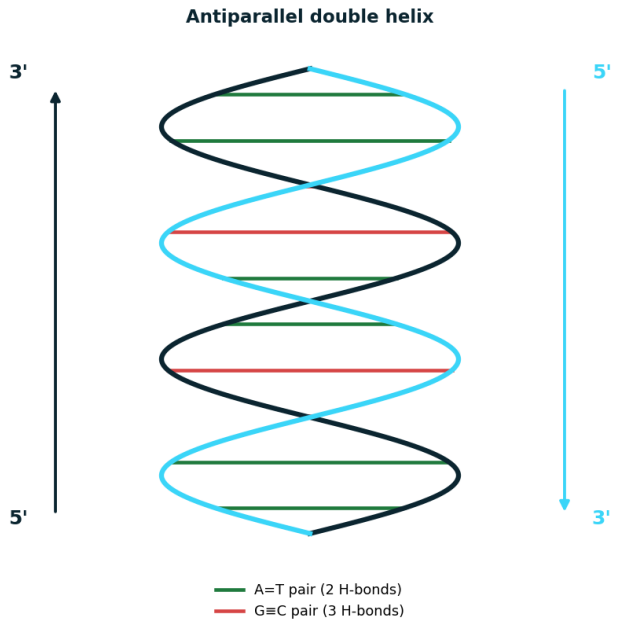

Figure 2. The DNA double helix: two antiparallel sugar-phosphate backbones (one running 5′ → 3′, the other 3′ → 5′) linked by complementary base pairs — A=T (2 H-bonds) and G≡C (3 H-bonds).

The shape of the helix

Because every rung is one purine paired with one pyrimidine, the two backbones stay a **constant distance apart**, giving the helix a uniform diameter along its whole length. If two purines paired together the rung would be too wide, and two pyrimidines would be too narrow — so the pairing rules are also a geometric requirement, not just a chemical preference. The two strands twist around each other into a right-handed spiral, completing one full turn about every ten base pairs. This coiling makes the molecule compact and helps protect the bases held inside. The regular, repeating shape is the same in the DNA of every organism, which is why a mechanism that copies or reads DNA in one species works on the DNA of another.

> **Directionality recap:** Because the strands are antiparallel, the base at the 5′ end of one strand pairs with the base at the 3′ end of the other. Always keep track of which end is which when writing a complementary strand.

5. The structure of RNA

RNA is also a polymer of nucleotides, but it differs from DNA in three key ways:

The three main types of RNA

RNA molecules do the practical work of expressing the information stored in DNA. Three types are central to protein synthesis (covered in detail in D1):

Type of RNA	Full name	Role
mRNA	Messenger RNA	Carries a copy of a gene's base sequence from the DNA in the nucleus to the ribosome; its sequence is read to build a protein

Type of RNA	Full name	Role
tRNA	Transfer RNA	Brings specific amino acids to the ribosome and matches them to the mRNA code through its anticodon
rRNA	Ribosomal RNA	A structural and catalytic component of ribosomes, the organelles where proteins are assembled

The single-stranded, less-stable nature of RNA suits its role as a **short-lived working copy** of genetic information, whereas the stable double helix of DNA suits **long-term storage** of the master copy.

> **Sugar spotting:** The name says it all: **deoxy**ribose in DNA is ribose "minus an oxygen". Ribose (RNA) has the full set of –OH groups; deoxyribose (DNA) lacks the oxygen on the 2′ carbon, which is one reason DNA is the more stable, long-term molecule.

6. DNA versus RNA – summary comparison

This comparison is a very common exam question. Learn it as a table.

Feature	DNA	RNA
Pentose sugar	Deoxyribose	Ribose
Bases used	A, G, C, **T** (thymine)	A, G, C, **U** (uracil)
Number of strands	Two (double helix)	One (single-stranded)
Pairing partner of A	Thymine (T)	Uracil (U)
Relative stability	More stable (long-term)	Less stable (short-lived)
Typical length	Very long (whole chromosomes)	Much shorter (single genes/products)
Main function	Stores genetic information as the master copy	Expresses information: makes proteins (mRNA, tRNA, rRNA)

> **Shared features:** Do not forget what DNA and RNA have in **common**: both are polymers of nucleotides; both use a pentose sugar, a phosphate group and nitrogenous bases; both have a sugar-phosphate backbone; and both share the bases adenine, guanine and cytosine.

7. How structure enables function

The whole point of A1.2 is to link the structure of nucleic acids to what they do. Three structural features explain three functions.

(a) Base sequence stores information

Because any base can follow any other along a strand, the **order of bases** can vary without limit. This variable sequence is a code: a specific order of bases specifies a specific order of amino acids in a protein, and hence the organism's characteristics. The uniform backbone carries this variable message without altering it.

(b) Complementary pairing enables accurate copying and transfer

Because A only pairs with T and G only with C, **each strand carries all the information needed to rebuild the other**. During **replication** the two strands separate and each acts as a **template**; free nucleotides pair with the exposed bases, producing two identical DNA molecules. The same complementary principle allows information to be **transferred** to mRNA during transcription. This is why DNA can be copied with such high fidelity from generation to generation.

(c) Stability suits long-term storage

Several features make DNA a robust archive. The **covalent** sugar-phosphate backbone is strong; the reactive bases are tucked **inside** the helix, protected by the backbones outside; the **double-stranded** structure means a damaged base on one strand can be repaired using the intact complementary strand; and DNA lacks the extra 2′ oxygen of RNA, making it chemically more stable. Together these allow DNA to store genetic information reliably for the whole life of a cell and to be passed on unchanged.

> **The big link: Sequence = information; pairing = copying and transfer; double helix = stability.** If you can state each structural feature and the function it serves, you have the core of this topic.

8. Nucleic acids, unity and diversity

Theme A asks you to see both the **unity** and the **diversity** of life. Nucleic acids illustrate both at once, which is why this topic sits at the start of the course.

Unity: one shared molecular language

Every known organism stores its genetic information in nucleic acids built from the same handful of nucleotides, joined by the same phosphodiester backbone, and read in the same 5′ → 3′ direction. In almost all organisms the same three-base codons specify the same amino acids – the genetic code is essentially **universal**. This deep similarity is strong evidence that all life descended from a common ancestor, and it is what makes genetic engineering possible: a human gene inserted into a bacterium can still be read and translated, because both use the same molecular language.

Diversity: endless variation in one sequence

At the same time, nucleic acids are the source of the **diversity** of life. Because the four bases can be arranged in any order along a chain of millions of units, the number of possible sequences is astronomically large. Differences in base sequence between individuals and species account for the differences in their proteins and therefore their characteristics. Changes to the sequence (**mutations**) create new variation, which is the raw material for evolution. So the same class of molecule provides both the sameness that unites all living things and the variation that makes each species – and each individual – unique.

> **Theme link:** One molecular system, endless sequences: the **structure** of nucleic acids is shared by all life (unity), while the **sequence** of bases differs between all life (diversity). This is the heart of Theme A.

9. Chargaff's rules (HL)

(HL) Before the double helix was known, Erwin Chargaff measured the amounts of each base in DNA from many species and found a consistent pattern, now called **Chargaff's rules**:

These equalities are a direct consequence of complementary base pairing: every A on one strand is matched by a T on the other, and every G by a C. It follows that the total amount of **purines equals the total amount of pyrimidines**:

A + G = T + C (purines = pyrimidines, ratio 1 : 1)

Using Chargaff's rules in calculations

Since the four base percentages must add up to 100%, knowing just one value lets you find all four. The method is:

1. If you know %A, then %T = %A (they are equal). 2. The remaining percentage, 100% − (%A + %T), is shared equally by G and C. 3. So %G = %C = [100% − (%A + %T)] ÷ 2.

Base composition as evidence

While %A = %T and %G = %C hold for every organism, the **ratio of (A+T) to (G+C) differs between species**. Human DNA, a bacterium and a plant each have a characteristic base composition. This variation was early evidence that DNA carries species-specific information, and that the differences between organisms lie in the **sequence and proportion of bases** rather than in the backbone. (Note that two unrelated species can share a similar overall %GC yet still have completely different base **sequences** − composition and sequence are not the same thing.)

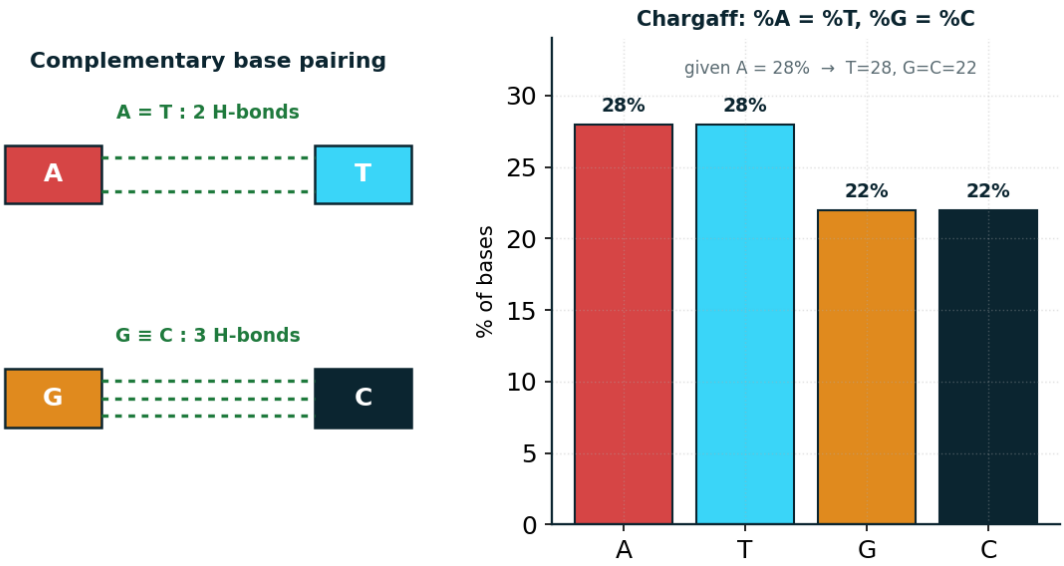

Figure 3. Left: A pairs with T by 2 hydrogen bonds, G pairs with C by 3 hydrogen bonds. Right: applying Chargaff to the worked example − given %A = 28%, then %T = 28% and %G = %C = 22%.

> **HL exam tip:** Chargaff's equalities apply to **double-stranded** DNA. They do **not** hold for a single strand on its own, nor for single-stranded RNA − a favourite trap. If a question gives you one strand only, %A need not equal %T on that strand.

10. DNA as the genetic material and the meaning of antiparallel (HL)

(HL) That DNA — rather than protein — is the hereditary molecule was established by a series of classic experiments. You are not expected to recall every detail, but you should understand the **logic** of the evidence.

Transformation evidence

Early work with bacteria showed that a harmless strain could be permanently changed ("transformed") into a disease-causing strain by something released from dead cells of the harmful strain. When the different chemical components were separated and tested, the **transforming factor** was shown to be **DNA**: destroying the DNA abolished the effect, whereas destroying proteins did not. This pointed to DNA as the carrier of inherited information.

Viral (bacteriophage) evidence

A virus that infects bacteria (a **bacteriophage**) is made of only DNA and a protein coat. By labelling the DNA and the protein separately, experiments showed that it is the phage's **DNA that enters the bacterium** and directs the production of new viruses, while the protein coat stays outside. Since only the DNA went in and carried the instructions, DNA must be the genetic material. Together these lines of evidence convinced biologists that genes are made of DNA.

The significance of antiparallel strands

The antiparallel arrangement is not a trivial detail — it is required for the molecule to work:

> **Directionality in one line:** Antiparallel means "parallel but pointing opposite ways": 5′ → 3′ on one strand sits alongside 3′ ← 5′ on the other. It is what makes complementary pairing and accurate copying possible.

11. Worked examples

Worked example 1 — Chargaff calculation

A sample of double-stranded DNA is found to contain 28% adenine. Calculate the percentage of each of the other three bases. (HL)

Step 1: %A = %T, so %T = **28%**.

Step 2: %A + %T = 28 + 28 = 56%. The rest, 100 − 56 = 44%, is shared by G and C.

Step 3: %G = %C = 44 ÷ 2 = **22%** each.

Answer: A = 28%, T = 28%, G = 22%, C = 22%. Check: 28 + 28 + 22 + 22 = 100%. ✓

Worked example 2 – writing a complementary strand

One strand of DNA reads 5′–A T G C G T A C–3′. Write the sequence of the complementary strand, showing its polarity.

Replace each base with its partner (A–T, G–C) and remember the new strand is antiparallel:

Template: 5′–A T G C G T A C–3′

Partner: 3′–T A C G C A T G–5′

So the complementary strand, written conventionally 5′ → 3′, is **5′–G T A C G C A T–3′**.

Worked example 3 – counting hydrogen bonds

A short double-stranded DNA segment is 20 base pairs long and contains 12 A–T pairs. How many hydrogen bonds hold the two strands together?

A–T pairs: 12, each with 2 hydrogen bonds → 12 × 2 = 24 bonds.

G–C pairs: 20 – 12 = 8, each with 3 hydrogen bonds → 8 × 3 = 24 bonds.

Total = 24 + 24 = 48 hydrogen bonds. Note that a G–C-rich region needs more energy to separate because it has more hydrogen bonds.

Worked example 4 – identifying 5′ and 3′ ends

How can you tell the 5′ end of a strand from the 3′ end, and why does it matter?

The **5′ end** has a free **phosphate group** attached to the 5′ carbon of the terminal sugar.

The **3′ end** has a free **–OH (hydroxyl) group** on the 3′ carbon of the terminal sugar.

It matters because sequences are read 5′ → 3′, the two strands are antiparallel, and enzymes add new nucleotides only to the 3′ end.

Worked example 5 – purine : pyrimidine ratio

In a sample of double-stranded DNA, adenine makes up 31% of the bases. Show that the ratio of purines to pyrimidines is 1 : 1. (HL)

From Chargaff: %T = %A = 31%. Remaining 100 – 62 = 38% is G + C, so %G = %C = 19%.

Purines = %A + %G = 31 + 19 = 50%. Pyrimidines = %T + %C = 31 + 19 = 50%.

Ratio = 50 : 50 = 1 : 1. This is always true in double-stranded DNA, because every purine on one strand is paired with a pyrimidine on the other.

Worked example 6 – interpreting base-composition data

A nucleic acid sample is analysed and found to contain A = 30%, G = 20%, C = 20%, U = 30%, and no thymine. Identify the type of nucleic acid and state whether it is single- or double-stranded, giving reasons.

It contains **uracil and no thymine**, so it must be **RNA** (thymine occurs only in DNA).

Here %A = %U and %G = %C only by coincidence of the numbers; RNA is normally **single-stranded**, so Chargaff-style equalities are not required and cannot be assumed. Because it is RNA, we conclude it is single-stranded.

Answer: single-stranded RNA. Note the trap: equal %A/%U here does not prove double-stranded – that reasoning applies only to double-stranded DNA.

12. Common exam pitfalls

13. Quick reference

Item	Key fact
Monomer	Nucleotide = pentose sugar + phosphate group + nitrogenous base
Backbone bond	Phosphodiester bond (formed by condensation, 5′ phosphate to 3′ –OH)
Purines	Adenine (A), Guanine (G) – double ring
Pyrimidines	Cytosine (C), Thymine (T), Uracil (U) – single ring
DNA bases	A, G, C, T
RNA bases	A, G, C, U (uracil replaces thymine)
Base pairing	A–T (2 H-bonds), G–C (3 H-bonds); A–U in RNA
DNA shape	Antiparallel double helix; backbones outside, bases inside
RNA shape	Single-stranded; ribose sugar
Chargaff (HL)	%A = %T and %G = %C in double-stranded DNA; purines = pyrimidines
Directionality	Read 5′ → 3′; nucleotides added to the 3′ end

14. Test yourself

Attempt these without notes; full answers follow.

1. List the three components of a nucleotide and name the bond that joins one nucleotide to the next. (3)

2. State the base-pairing rules in DNA and give the number of hydrogen bonds in each pair. (3)

3. Explain what is meant by “antiparallel” and why it is important. (3)

4. Give three differences between DNA and RNA. (3)

5. Explain how the structure of DNA makes it suitable for the long-term storage of genetic information. (3)

6. **(HL)** A sample of double-stranded DNA contains 20% cytosine. Calculate the percentage of adenine, thymine and guanine. (3)

7. **(HL)** A double-stranded DNA molecule has 15 A−T base pairs and 10 G−C base pairs. How many hydrogen bonds hold it together? (2)

8. One strand of DNA reads 5′−T A C G G A T−3′. Write the complementary strand in the correct orientation, and write the mRNA sequence that would be transcribed from this strand. (3)

Answers

1. A pentose sugar, a phosphate group and a nitrogenous base (1+1 for all three); nucleotides are joined by a **phosphodiester bond**, formed by condensation (1).

2. Adenine pairs with thymine (A−T) with **2** hydrogen bonds (1+1); guanine pairs with cytosine (G−C) with **3** hydrogen bonds (1).

3. The two strands run in opposite directions − one 5′ → 3′ and the other 3′ → 5′ (1). This is required so that the complementary bases can align and form hydrogen bonds correctly (1), and it determines how the strands are copied during replication (1).

4. Any three: DNA has deoxyribose, RNA has ribose; DNA uses thymine, RNA uses uracil; DNA is double-stranded, RNA is single-stranded; DNA is longer/more stable, RNA is shorter/less stable (1 each, max 3).

5. The sugar-phosphate backbone is held by strong covalent bonds (1); the reactive bases are protected on the inside of the helix, and the double-stranded structure allows damage to one strand to be repaired using the other (1); the base sequence stores information stably and can be copied accurately by complementary base pairing (1).

6. %C = 20%, so %G = 20% (%G = %C) (1). A + T = 100 − (20 + 20) = 60%, shared equally, so %A = %T = 30% each (1+1). **A = 30%, T = 30%, G = 20%.**

7. A−T: 15 × 2 = 30 bonds; G−C: 10 × 3 = 30 bonds (1); total = **60 hydrogen bonds** (1).

8. Complementary DNA strand (antiparallel): 3′−A T G C C T A−5′, i.e. written 5′ → 3′ it is 5′−A T C C G T A−3′ (1+1). The mRNA transcribed from the given strand uses U in place of T and A−U pairing: 3′−A U G C C U A−5′ (1).