

Nucleic acids and protein synthesis: summary notes

These notes cover the whole of Topic 6, Nucleic acids and protein synthesis, of the Cambridge International AS & A Level Biology syllabus (9700). They follow the syllabus order: first the structure of nucleic acids and DNA replication (6.1), then protein synthesis (6.2). They cover the structure of nucleotides and ATP, the DNA double helix and base pairing, semi-conservative replication and the enzymes involved, a comparison of DNA and RNA, the genetic code, transcription and translation, and gene mutations and their effects. Worked examples show every step, and there are quick-check questions with answers at the end.

1. Nucleotides: the building blocks

Nucleic acids (DNA and RNA) are **polymers**. Their monomers are called **nucleotides**. Every nucleotide has three parts joined together:

- a **pentose sugar** (a sugar with five carbon atoms): **deoxyribose** in DNA, **ribose** in RNA
- a **phosphate group**
- a **nitrogen-containing (nitrogenous) base**

To picture a nucleotide, think of a pentagon (the sugar) with the phosphate attached to one side (carbon 5) and the base attached to another side (carbon 1).

The bases: purines and pyrimidines

Group	Ring structure	Bases	Found in
Purines	Two rings (double-ring), larger	Adenine (A), guanine (G)	DNA and RNA
Pyrimidines	One ring (single-ring), smaller	Cytosine (C), thymine (T), uracil (U)	C in both; T in DNA only; U in RNA only

Memory aid: the pyrimidines **CUT** the "py" (**C**ytosine, **U**racil, **T**hymine). The purines are the other two, adenine and guanine ("pure As Gold"). Note that pyrimidine is the longer word but the smaller, single-ring molecule.

Joining nucleotides: phosphodiester bonds

Nucleotides are joined by **condensation reactions**, which release a molecule of water each time. The phosphate group on carbon 5 of one nucleotide bonds to the -OH group on carbon 3 of the sugar of the next nucleotide. The link formed (sugar–phosphate–sugar) is called a **phosphodiester bond**. A chain of nucleotides is a **polynucleotide**. The alternating sugars and phosphates make up the **sugar-phosphate backbone**, and the bases stick out to the side.

Because of the way the bonds form, each strand has a direction. One end has a free phosphate on carbon 5 (the **5' end**, "five prime") and the other end has a free -OH on carbon 3 (the **3' end**).

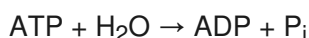
ATP: a phosphorylated nucleotide

ATP (adenosine triphosphate) is built like an RNA nucleotide but carries **three** phosphate groups

instead of one:

- the base **adenine**
- the pentose sugar **ribose**
- a chain of **three phosphate groups**

Adenine joined to ribose is called **adenosine**. Hydrolysis of ATP removes the end phosphate, giving ADP (adenosine diphosphate) and an inorganic phosphate (P_i), and releases energy that the cell can use:



ATP is not part of a polynucleotide chain; it is a free molecule and the cell's universal energy currency. However, ATP also acts as a raw material: when RNA is made, ATP is one of the activated nucleotides that are added to the growing chain (losing two phosphates as it does so).

2. The structure of DNA

Key idea: the DNA double helix

A DNA molecule consists of **two polynucleotide strands** that run in **opposite directions (antiparallel)** and are held together by **hydrogen bonds** between **complementary base pairs**. The two strands are twisted around each other to form a **double helix**.

Features you must be able to describe

- **Sugar-phosphate backbones** are on the outside of the helix. Phosphodiester bonds are covalent and strong, so the backbone is stable and protects the bases.
- **Bases** point inwards and pair across the middle, like the rungs of a twisted ladder.
- **Complementary base pairing**: adenine always pairs with thymine (A–T) and guanine always pairs with cytosine (G–C).
- A–T pairs are held by **two** hydrogen bonds; G–C pairs are held by **three** hydrogen bonds.
- Each pair is always one **purine** with one **pyrimidine**. A two-ring base plus a one-ring base always spans the same distance, so the helix has a **constant width** all along its length.
- **Antiparallel**: one strand runs $5' \rightarrow 3'$ and its partner runs $3' \rightarrow 5'$.
- Individual hydrogen bonds are weak, so the strands can be separated (for replication and transcription), but there are so many of them that the molecule as a whole is stable.

A consequence of base pairing is that in any double-stranded DNA, the amount of A equals the amount of T, and the amount of G equals the amount of C. So $\%A + \%G = 50\%$, and $\%T + \%C = 50\%$.

Worked example 1: base percentages

A sample of double-stranded DNA contains 18% adenine. Calculate the percentage of each of the other bases.

Step 1. A pairs with T, so the amount of T equals the amount of A: $T = 18\%$.

Step 2. A and T together: $18\% + 18\% = 36\%$.

Step 3. The rest must be G and C: $100\% - 36\% = 64\%$.

Step 4. G pairs with C, so they are equal: $64\% \div 2 = 32\%$ each.

Answer: T = 18%, G = 32%, C = 32%.

Check: $18 + 18 + 32 + 32 = 100$ ✓

This method only works for double-stranded DNA. In single-stranded RNA the bases are not paired, so you cannot work out one base from another.

3. Semi-conservative replication of DNA

Before a cell divides, its DNA must be copied exactly, so that each daughter cell receives identical genetic information. This happens during the S phase of interphase.

Key idea: semi-conservative replication

Each new DNA molecule contains **one strand from the original (parent) molecule** and **one newly made strand**. Each original strand acts as a **template** for building a new complementary strand.

The stages of replication

1. **Unwinding and separating.** The enzyme **DNA helicase** unwinds the double helix and breaks the **hydrogen bonds** between complementary bases, separating the two strands. This exposes the bases on both strands. The Y-shaped region where this happens is called a replication fork.
2. **Base pairing.** Free **activated nucleotides** in the nucleus (deoxyribonucleotides carrying three phosphate groups) move in and pair with the exposed bases by **complementary base pairing**: A with T and C with G. Hydrogen bonds form between the paired bases.
3. **Joining the nucleotides.** The enzyme **DNA polymerase** joins adjacent nucleotides of the new strand by forming **phosphodiester bonds** (condensation). The two extra phosphates on each activated nucleotide are released, and this provides energy for the reaction.
4. **Result.** Two DNA molecules are formed, each identical to the original and each containing one old and one new strand. Each rewinds into a double helix.

Leading and lagging strands

DNA polymerase can only add nucleotides to the **3' end** of a growing strand, so a new strand can only be built in the **5' → 3' direction**. Because the two template strands are antiparallel, this causes a problem at each replication fork:

- On one template the new strand grows in the same direction as the fork opens. It can be made **continuously**. This is the **leading strand**.
- On the other template the new strand must grow **away** from the fork. It is made **discontinuously**, in short sections (often called Okazaki fragments) as each new stretch of template is exposed. This is the **lagging strand**.
- The enzyme **DNA ligase** then joins the fragments of the lagging strand together by forming phosphodiester bonds between them, making one continuous strand.

Enzyme	Role in replication
--------	---------------------

DNA helicase	Unwinds the helix and breaks hydrogen bonds between bases to separate the strands
DNA polymerase	Joins free nucleotides to the 3' end of the growing strand, forming phosphodiester bonds; works in the 5' → 3' direction only
DNA ligase	Joins the fragments on the lagging strand by forming phosphodiester bonds

Beyond the syllabus: each fragment is started by a short RNA primer made by an enzyme called primase. You do not need to know this for 9700, but it explains why DNA polymerase always needs an existing 3' end to add to.

Evidence for the semi-conservative model

The classic evidence comes from growing bacteria on a medium containing only the heavy nitrogen isotope ^{15}N , so that all their DNA becomes "heavy", then moving them to a medium containing only the normal light isotope ^{14}N . DNA extracted after each generation is spun in a centrifuge in a density gradient: heavy DNA settles lowest, light DNA highest, and hybrid DNA (one heavy strand and one light strand) in between.

Worked example 2: predicting DNA densities

Bacteria with only ^{15}N DNA are moved to a ^{14}N medium and allowed to divide three times. Assume replication is semi-conservative. Calculate the percentage of DNA molecules that would be heavy, hybrid and light after the third generation.

Step 1: count the molecules. Start with 1 DNA molecule. Each generation doubles the number: $1 \rightarrow 2 \rightarrow 4 \rightarrow 8$. After three generations there are 8 molecules, which contain $8 \times 2 = 16$ strands.

Step 2: count the heavy strands. The original molecule had 2 heavy (^{15}N) strands. Strands are never broken down or remade, so there are still exactly 2 heavy strands. All 14 other strands are light.

Step 3: place the heavy strands. In semi-conservative replication each old strand is always paired with a new one. So the 2 heavy strands are in 2 different molecules, each paired with a light strand. These 2 molecules are hybrid.

Step 4: the rest. $8 - 2 = 6$ molecules contain two light strands.

Step 5: convert to percentages. Hybrid = $2 \div 8 \times 100 = 25\%$. Light = $6 \div 8 \times 100 = 75\%$. Heavy = 0%.

Answer: 0% heavy, 25% hybrid, 75% light.

Compare: if replication were *conservative* (the original molecule stays intact), after three generations there would be 1 heavy molecule out of 8 (12.5%), 7 light (87.5%) and no hybrid. The fact that only hybrid DNA appears after the first generation rules out the conservative model. The *dispersive* model (every strand a patchwork of old and new DNA) would give a single intermediate band that becomes gradually lighter each generation. In the second generation, a light band appears as well as the hybrid band, and this rules out the dispersive model.

4. DNA and RNA compared

RNA (ribonucleic acid) is also a polynucleotide, made of nucleotides joined by phosphodiester bonds, but it differs from DNA in several ways.

Feature	DNA	RNA
---------	-----	-----

Pentose sugar	Deoxyribose	Ribose
Bases	A, T, C, G	A, U, C, G (uracil replaces thymine)
Number of strands	Two (double-stranded), forming a double helix	Usually one (single-stranded)
Base pairing	Complementary pairs along the whole molecule (A–T, G–C)	None in mRNA; tRNA and rRNA fold back on themselves so that some regions pair (A–U, G–C)
Size	Very long; one molecule per chromosome	Much shorter
Where found (eukaryote)	Mainly the nucleus (also mitochondria and chloroplasts)	Made in the nucleus; works in the nucleus and cytoplasm, including ribosomes
Stability	Very stable; lasts the life of the cell	Less stable; mRNA is broken down after use
Types and function	One type; stores genetic information	Several types: messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA); involved in protein synthesis

The three types of RNA

- **mRNA (messenger RNA):** a single, straight strand that carries a copy of the base sequence of a gene from the nucleus to a ribosome.
- **tRNA (transfer RNA):** a short strand that folds into a clover-leaf shape held by hydrogen bonds between some of its own bases. It has an **anticodon** (three unpaired bases) at one end and a site where a **specific amino acid** attaches at the other end. It brings amino acids to the ribosome.
- **rRNA (ribosomal RNA):** together with proteins, makes up the two subunits (large and small) of a ribosome.

5. Genes and the genetic code

Definition: gene

A gene is a sequence of nucleotides that forms part of a DNA molecule and codes for a polypeptide.

The sequence of bases in a gene determines the sequence of amino acids in a polypeptide (the **primary structure**). Since the primary structure decides how the polypeptide folds, the gene ultimately controls the shape and function of the protein. This is the basis of the "one gene, one polypeptide" idea.

Features of the genetic code

- **Triplet code:** each amino acid is coded for by a sequence of three bases. The code is carried in DNA as base triplets; the mRNA copy of each triplet is a **codon**. There are $4^3 = 64$ possible codons, which is more than enough for the 20 amino acids. (A doublet code would give only $4^2 = 16$, too few.)

- **Degenerate:** most amino acids are coded for by more than one codon. For example, lysine is coded for by both AAA and AAG.
- **Non-overlapping:** each base is read as part of only one codon.
- **Read from a fixed starting point** in groups of three, with no gaps or "commas" between codons. This is why adding or removing a single base causes a frameshift (section 8).
- **Start codon:** AUG starts translation and also codes for the amino acid methionine.
- **Stop codons:** UAA, UAG and UGA do not code for any amino acid; they end translation.
- **Universal:** the same codon codes for the same amino acid in (almost) all organisms. This is why a human gene can be expressed in a bacterium to produce the human protein, as long as it is inserted without its introns (for example, a copy made from the mRNA), because bacteria cannot splice (section 6).

The codon table below shows only the codons used in these notes. In the exam you are always given any codon information you need.

mRNA codon	Amino acid	mRNA codon	Amino acid
AUG	Methionine (Met), start	GAA, GAG	Glutamic acid (Glu)
UUU	Phenylalanine (Phe)	GUG	Valine (Val)
UUG	Leucine (Leu)	GCA	Alanine (Ala)
GGC, GGU	Glycine (Gly)	AAU	Asparagine (Asn)
AAA, AAG	Lysine (Lys)	CCU	Proline (Pro)
UGG	Tryptophan (Trp)	UAA, UAG, UGA	Stop

6. Transcription

DNA stays in the nucleus, but polypeptides are made on ribosomes in the cytoplasm. A copy of the gene, in the form of mRNA, carries the information out. Making this mRNA copy is called **transcription**.

Definition: transcription

Transcription is the production of an RNA molecule with a base sequence complementary to the template strand of a gene, catalysed by RNA polymerase.

Steps of transcription

1. **RNA polymerase** binds to the DNA at the start of the gene (at a region called the promoter).
2. The DNA double helix unwinds and the hydrogen bonds between bases break over a short length, exposing the bases of the gene. RNA polymerase does this itself; no separate helicase is needed.
3. Only **one** of the two DNA strands is used as a template. This is the **template strand** (also called the antisense or transcribed strand). The other strand is the **coding strand** (sense strand).

4. Free activated **RNA nucleotides** pair with the exposed bases of the template strand by complementary base pairing (RNA nucleotide : DNA base): C with G, G with C, U with A, and A with T. Uracil on the RNA pairs with adenine on the DNA; there is no U–T pairing.
5. **RNA polymerase** joins the RNA nucleotides together by forming **phosphodiester bonds**, moving along the template strand. The RNA strand grows in the 5' → 3' direction.
6. Behind the enzyme, the new RNA separates from the DNA and the DNA rewinds.
7. When RNA polymerase reaches a terminator sequence at the end of the gene, it detaches and the RNA is released.

Useful shortcut: because the mRNA is complementary to the template strand, it has the same base sequence as the coding strand, except that U replaces T.

Processing the RNA in eukaryotes

In eukaryotes, the RNA made by transcription is called the **primary transcript** (or pre-mRNA). It contains both:

- **exons**: sequences that code for parts of the polypeptide, and
- **introns**: non-coding sequences in between.

Before the RNA leaves the nucleus, the **introns are removed** and the **exons are joined together**. This process is called **splicing**. The result is **mature mRNA**, which leaves the nucleus through a **nuclear pore** and travels to a ribosome. Prokaryotic genes do not contain introns, so their mRNA does not need splicing and can be translated while it is still being made.

7. Translation

Definition: translation

Translation is the process in which the sequence of codons on mRNA is used to assemble amino acids, in the correct order, into a polypeptide at a ribosome.

Before translation: loading tRNA

In the cytoplasm, each tRNA molecule is joined to its **specific amino acid** by a specific enzyme. This step uses energy from **ATP**. Each type of tRNA has a particular **anticodon**, so it always carries the same amino acid.

Steps of translation

1. The mRNA binds to the **small subunit of a ribosome**. The ribosome moves along the mRNA until it reaches the start codon **AUG**.
2. A tRNA with the **complementary anticodon** (UAC) and carrying methionine binds to the start codon by complementary base pairing (hydrogen bonds between codon and anticodon).
3. The ribosome can hold **two tRNA molecules** at a time, which means two codons (six bases) of the mRNA are exposed inside it. A second tRNA, with an anticodon complementary to the next codon, brings its amino acid into the second position.
4. A **peptide bond** forms between the two amino acids by a **condensation reaction**. This is catalysed by the ribosome itself (by its rRNA).

5. The ribosome **moves along the mRNA by one codon**. The first tRNA, now without its amino acid, leaves and can collect another amino acid. A third tRNA moves in, and another peptide bond forms.
6. This repeats, making the polypeptide chain longer by one amino acid at a time.
7. When the ribosome reaches a **stop codon** (UAA, UAG or UGA), there is no tRNA with a matching anticodon. The polypeptide is released and the ribosome subunits separate.

Several ribosomes can translate the same mRNA at the same time, one behind the other, so many copies of the polypeptide are made quickly. After release, the polypeptide may fold and, in eukaryotes, may be modified in the rough endoplasmic reticulum and Golgi body.

Transcription and translation compared

Feature	Transcription	Translation
Where (eukaryote)	Nucleus	Ribosomes in the cytoplasm (free or on rough ER)
Template	Template strand of DNA	mRNA
Product	RNA (primary transcript, then mRNA)	Polypeptide
Monomers used	RNA nucleotides	Amino acids (carried by tRNA)
Bond formed	Phosphodiester bonds	Peptide bonds
Main enzyme / structure	RNA polymerase	Ribosome
Base pairing	DNA bases with RNA nucleotides	mRNA codon with tRNA anticodon

Worked example 3: from DNA to polypeptide

The template strand of a short gene has the base sequence below (written 3' → 5', and split into triplets for clarity):

3' TAC AAA CCG TTT ACC ATT 5'

Use the codon table in section 5 to state (a) the mRNA sequence, (b) the tRNA anticodons used and (c) the amino acid sequence of the polypeptide.

Step 1: transcribe. Pair each template base with its complementary RNA base (A → U, T → A, C → G, G → C):

Template triplet	TAC	AAA	CCG	TTT
mRNA codon	AUG	UUU	GGC	AAA
tRNA anticodon	UAC	AAA	CCG	UUU
Amino acid	Met	Phe	Gly	Lys

Template triplet	ACC	ATT
mRNA codon	UGG	UAA

tRNA anticodon	ACC	none
Amino acid	Trp	stop

(a) mRNA: 5' AUG UUU GGC AAA UGG UAA 3'

Step 2: find the anticodons. Each anticodon is complementary to its codon. Notice that the anticodon has the same sequence as the DNA template triplet, but with U in place of T.

(b) Anticodons: UAC, AAA, CCG, UUU, ACC. No tRNA binds to the stop codon UAA.

Step 3: translate. Read the mRNA codons from the start codon, in order, using the table.

(c) Polypeptide: Met – Phe – Gly – Lys – Trp (5 amino acids). The sixth codon, UAA, is a stop codon and adds no amino acid.

Worked example 4: codons and amino acids

The coding region of a mature mRNA, from the start codon up to and including the stop codon, contains 903 nucleotides. Calculate the number of amino acids in the polypeptide.

Step 1. Number of codons = $903 \div 3 = 301$ codons.

Step 2. The last codon is a stop codon, which codes for no amino acid: $301 - 1 = 300$.

Answer: 300 amino acids.

Do not subtract anything for the start codon: AUG codes for methionine, so it does add an amino acid. (In many proteins this first methionine is later removed, but questions will tell you if that matters.)

8. Gene mutations

Definition: gene mutation

A gene mutation is a change in the sequence of base pairs in a DNA molecule that may result in an altered polypeptide.

Gene mutations happen randomly, for example as copying errors during DNA replication, and their frequency is increased by mutagens such as ionising radiation and some chemicals. A mutation matters because a changed base sequence in DNA leads to a changed mRNA, which may change the amino acid sequence (primary structure), and therefore the folding, shape and function of the protein.

Types of gene mutation

Type	What happens	Possible effects on the polypeptide
Substitution	One base pair is replaced by a different base pair	Only one codon is affected. It may code for: the same amino acid (no change, because the code is degenerate); a different amino acid (one amino acid changed; effect can be small or large); or a stop codon (polypeptide cut short, usually non-functional)
Insertion	One or more extra base pairs are added	If the number added is not a multiple of three, a frameshift occurs

Deletion	One or more base pairs are lost	If the number lost is not a multiple of three, a frameshift occurs
-----------------	---------------------------------	---

Frameshift mutations

Because the code is read in threes from a fixed starting point with no gaps, inserting or deleting one base changes the grouping of **every codon after the mutation**. This is called a **frameshift**. Every amino acid after the mutation is likely to be different, and a stop codon may appear early (giving a short polypeptide) or the normal stop codon may be lost. The protein almost always fails to fold correctly and is usually non-functional. A frameshift near the **start** of a gene is usually more damaging than one near the end, because more of the polypeptide is affected.

If **three** bases (or any multiple of three) are inserted or deleted, the reading frame is restored. If the three bases make up exactly one codon, the polypeptide simply gains or loses one amino acid (or one per three bases). If the three bases span two codons, one amino acid is gained or lost and the neighbouring one may also change (or a stop codon may form). Either way, the reading frame after that point is normal, so the effect may be much smaller than a frameshift.

Worked example 5: the effects of different mutations

Use the mRNA from worked example 3: 5' AUG UUU GGC AAA UGG UAA 3', which codes for Met – Phe – Gly – Lys – Trp. Describe the effect on the polypeptide of each change to the mRNA below (each change results from a mutation in the DNA).

(a) Substitution: the fourth codon changes from AAA to AAG.

Step 1: find the new codon's amino acid. AAG codes for lysine.

Step 2: compare. AAA also codes for lysine, so the amino acid is unchanged.

Effect: the polypeptide is Met – Phe – Gly – Lys – Trp, identical to the original. This is a silent mutation, possible because the code is degenerate.

(b) Substitution: the fourth codon changes from AAA to UAA.

Step 1: UAA is a stop codon.

Step 2: translation stops at codon 4.

Effect: the polypeptide is Met – Phe – Gly only (3 amino acids instead of 5). It is shortened and probably non-functional.

(c) Deletion: the first U of the second codon is lost.

Step 1: write out the mRNA with the base removed: AUG UUGGCAAUGGUAA

Step 2: regroup into triplets from the start codon: AUG UUG GCA AAU GGU AA...

Step 3: translate: Met – Leu – Ala – Asn – Gly – ...

Effect: this is a frameshift. Every amino acid after the first is different, and the original stop codon UAA is no longer in the reading frame, so translation continues into the next bases until a stop codon happens to turn up. The polypeptide is very different and almost certainly non-functional.

An example: sickle cell anaemia

The gene for the β -globin polypeptide of haemoglobin has a well-known substitution mutation. On the

coding strand the triplet **GAG** becomes **GTG** (on the template strand, CTC becomes CAC). The mRNA codon changes from GAG (glutamic acid) to GUG (valine), so the sixth amino acid of β -globin is valine instead of glutamic acid. Valine is non-polar, whereas glutamic acid is charged. At low oxygen concentrations, this makes the haemoglobin molecules stick together into long fibres, distorting red blood cells into a sickle shape. A single base change can therefore have a very large effect on the organism.

This example is also used at A Level (Topic 16). For Topic 6 you only need the principle that one substitution can change one amino acid.

Key terms

Term	Definition
Nucleotide	Monomer of a nucleic acid, made of a pentose sugar, a phosphate group and a nitrogenous base
Purine	Double-ring base: adenine or guanine
Pyrimidine	Single-ring base: cytosine, thymine or uracil
Phosphodiester bond	Covalent link between the phosphate of one nucleotide and the sugar of the next, formed by condensation
ATP	Phosphorylated nucleotide made of adenine, ribose and three phosphate groups; the cell's energy currency
Antiparallel	Describes the two strands of DNA running in opposite directions ($5' \rightarrow 3'$ and $3' \rightarrow 5'$)
Complementary base pairing	Hydrogen bonding of A with T (or U) and G with C
Semi-conservative replication	DNA copying in which each new molecule contains one original strand and one new strand
DNA helicase	Enzyme that unwinds DNA and breaks hydrogen bonds between base pairs
DNA polymerase	Enzyme that joins nucleotides into a new DNA strand, adding them to the $3'$ end
DNA ligase	Enzyme that joins DNA fragments (e.g. on the lagging strand) with phosphodiester bonds
Leading strand	New DNA strand made continuously during replication
Lagging strand	New DNA strand made in short fragments that are later joined by DNA ligase
Gene	A sequence of nucleotides that forms part of a DNA molecule and codes for a polypeptide
Codon	Sequence of three bases on mRNA that codes for one amino acid or a stop signal
Anticodon	Sequence of three bases on tRNA, complementary to a codon on mRNA
Degenerate code	Most amino acids are coded for by more than one codon
Template strand	The DNA strand that is used as a template during transcription

Transcription	Making RNA complementary to the template strand of a gene, using RNA polymerase
Primary transcript	RNA straight after transcription in a eukaryote, before introns are removed
Intron / exon	Non-coding sequence removed from the primary transcript / coding sequence kept in mature mRNA
Translation	Assembling a polypeptide at a ribosome, using the codon sequence of mRNA
Gene mutation	A change in the sequence of base pairs in a DNA molecule that may result in an altered polypeptide
Frameshift	Change in the reading frame caused by inserting or deleting a number of bases that is not a multiple of three

Quick check

1. State the three components of a DNA nucleotide. [2]
2. Name the bond that joins nucleotides in a polynucleotide and the type of reaction that forms it. [2]
3. State **two** differences between the structure of ATP and the structure of an adenine nucleotide in DNA. [2]
4. A sample of double-stranded DNA contains 31% guanine. Calculate the percentage of thymine. Show your working. [2]
5. Explain why the two strands of a DNA molecule are always the same distance apart. [2]
6. Explain what is meant by *semi-conservative* replication. [2]
7. Explain why one new strand is made continuously but the other is made in fragments during DNA replication. [3]
8. The template strand of part of a gene is 3' TAC GGA CTT 5'. Using the codon table in section 5, state the mRNA sequence, the three tRNA anticodons and the amino acid sequence coded for. [3]
9. Explain what is meant by the statement that the genetic code is *degenerate*, and suggest why this is an advantage. [2]
10. Describe the roles of mRNA and tRNA in protein synthesis. [4]
11. Explain why a deletion of one base near the start of a gene usually has a greater effect on the protein than a deletion of three adjacent bases. [4]
12. The coding region of a mature mRNA, including the stop codon, contains 1,236 nucleotides. (a) Calculate the number of amino acids in the polypeptide. (b) The gene that codes for this mRNA in a human cell is much longer than 1,236 base pairs. Suggest why. [3]
13. State **three** structural differences between a molecule of DNA and a molecule of mRNA. [3]
14. Describe how a molecule of mRNA is produced from a gene in a eukaryotic cell. [5]

Answers

1. Deoxyribose (pentose sugar); phosphate group; nitrogenous base (A, T, C or G). All three for [2]; two for [1].

2. Phosphodiester bond [1]; condensation reaction [1].
3. Any two, [1] each: ATP has three phosphate groups, the DNA nucleotide has one; ATP contains ribose, the DNA nucleotide contains deoxyribose. (Not credited: "ATP is a free molecule, not part of a polynucleotide chain". This is not a difference in molecular structure.)
4. $C = G = 31\%$, so $G + C = 62\%$ [1]. $A + T = 100 - 62 = 38\%$, so $T = 38 \div 2 = 19\%$ [1].
5. Each base pair always contains one purine (two rings) and one pyrimidine (one ring) [1]; so every pair has the same total width / A always pairs with T and G with C [1].
6. Each new DNA molecule contains one strand from the original molecule [1] and one newly synthesised strand, made using the original strand as a template [1].
7. The two strands are antiparallel [1]. DNA polymerase can only add nucleotides to the 3' end / build a strand in the $5' \rightarrow 3'$ direction [1]. On one template this is the same direction as the replication fork moves, so synthesis is continuous (leading strand); on the other it is away from the fork, so short fragments are made as more template is exposed, then joined by DNA ligase (lagging strand) [1].
8. mRNA: AUG CCU GAA [1]; anticodons: UAC, GGA, CUU [1]; amino acids: Met – Pro – Glu [1].
9. Most amino acids are coded for by more than one codon [1]. Advantage: some substitution mutations (often in the third base of a codon) do not change the amino acid, so the protein is unaffected [1].
10. mRNA: a copy of the gene's base sequence, made in the nucleus by transcription [1]; carries it to the ribosome, where its codons determine the order of amino acids [1]. tRNA: carries a specific amino acid to the ribosome [1]; its anticodon pairs with the complementary codon on mRNA by hydrogen bonding, so each amino acid is placed in the correct position [1].
11. The code is read in non-overlapping triplets from a fixed start point [1]. Deleting one base causes a frameshift, so every codon after the mutation is changed [1], changing many amino acids and possibly creating an early stop codon or losing the stop codon, so the protein is non-functional [1]. Deleting three adjacent bases removes one codon (or changes at most two neighbouring codons into one), so the reading frame is kept and only one or two amino acids are affected; the rest of the primary structure is normal and the protein may still work [1].
12. (a) $1,236 \div 3 = 412$ codons [1]; $412 - 1$ stop codon = 411 amino acids [1]. (b) The gene contains introns that are transcribed but removed from the primary transcript by splicing, so they are not in the mature mRNA [1]. (Also accept: the gene includes non-coding regions such as the promoter.)
13. Any three, [1] each: DNA contains deoxyribose, mRNA contains ribose; DNA contains thymine, mRNA contains uracil; DNA is double-stranded (a double helix), mRNA is single-stranded; DNA bases are paired by hydrogen bonds, mRNA bases are unpaired; a DNA molecule is much longer than an mRNA molecule.
14. Any five, [1] each: RNA polymerase binds to the DNA at the start of the gene (promoter); the DNA unwinds and hydrogen bonds between bases break; only one strand (the template strand) is used; free RNA nucleotides pair with exposed template bases by complementary base pairing, with U pairing with A; RNA polymerase joins the nucleotides by

phosphodiester bonds; the primary transcript (pre-mRNA) is spliced, with introns removed and exons joined; the mature mRNA leaves the nucleus through a nuclear pore.