



Protein synthesis

Revision notes · IB Biology D1.2: transcription, translation, the genetic code, and mutations that change proteins.

For IB Diploma Biology SL undefined. **HL only** is additional higher level content, for HL students only.

Protein synthesis Spec D1.2

Transcription and gene expression Spec D1.2.1, D1.2.2, D1.2.3, D1.2.4

Transcription is the synthesis of RNA using a DNA template. Only one of the two DNA strands, the **template strand**, is transcribed. The other strand is the coding strand: the RNA has the same base sequence as the coding strand, except that uracil replaces thymine.

Roles of RNA polymerase

1. RNA polymerase binds to DNA at the start of a gene.
2. It unwinds the double helix and separates the two strands by breaking hydrogen bonds between bases, over a short stretch of DNA.
3. Free RNA nucleotides pair with exposed bases on the template strand by complementary base pairing.
4. RNA polymerase links the RNA nucleotides with covalent bonds to form a single strand of RNA.
5. Behind the polymerase, the RNA separates from the template and the DNA strands re-form their hydrogen bonds and rewind.

Hydrogen bonding and complementary base pairing

Each RNA nucleotide is held in place by **hydrogen bonds** to the complementary base on the template strand. The pairs are: **adenine (A) on the DNA template with uracil (U) on the RNA**, thymine (T) on the template with adenine (A), cytosine (C) with guanine (G) and guanine (G) with cytosine (C). The hydrogen bonds are weak, so the RNA can separate from the template once its nucleotides are linked.

Stability of DNA templates

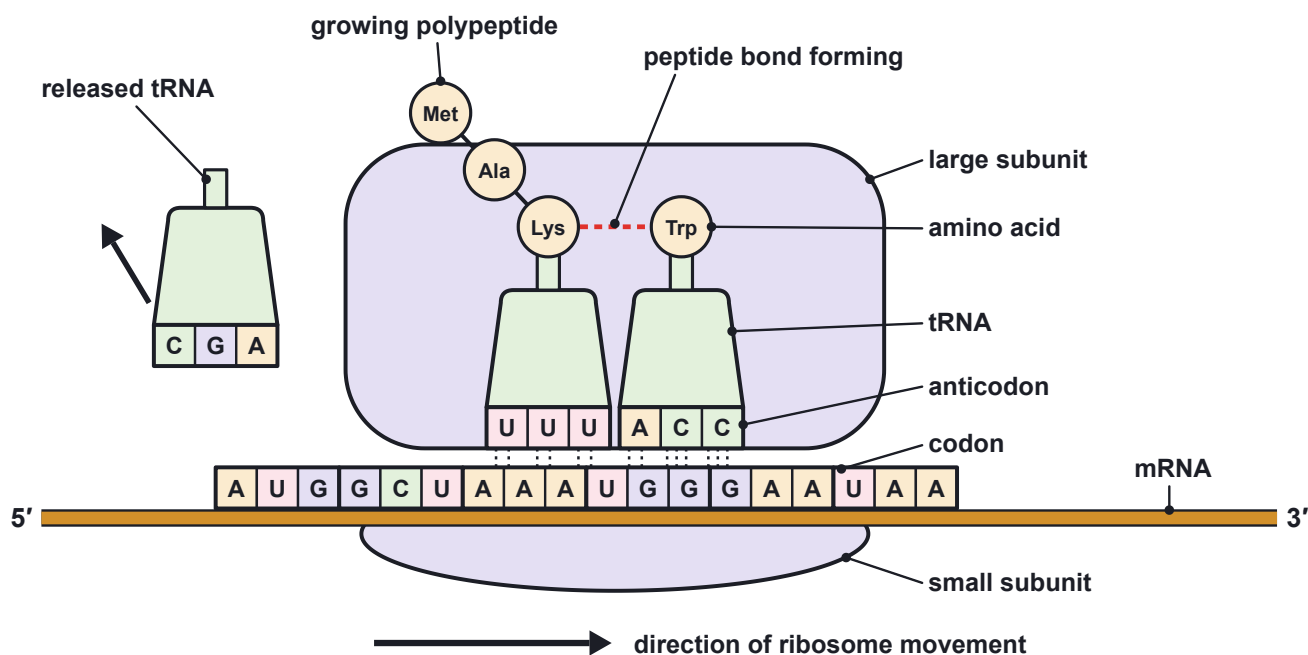
A single DNA strand can be used as a template for transcribing a base sequence without the DNA base sequence changing: only hydrogen bonds are broken and then re-formed. In **somatic cells that do not divide**, such as many neurones, the same DNA must be transcribed again and again, so its sequences must be conserved throughout the life of the cell.

Transcription and gene expression

Not all genes in a cell are expressed at any given time. Transcription is the **first stage of gene expression**, so it is a key stage at which the expression of a gene can be switched on or off. A gene that is not transcribed produces no mRNA and so no polypeptide.

2. A second tRNA, whose anticodon is complementary to the next codon, binds alongside it, bringing its amino acid.
3. A **peptide bond** forms between the amino acid on the second tRNA and the polypeptide, so the polypeptide is transferred to the second tRNA and becomes one amino acid longer.
4. The ribosome moves along the mRNA by one codon (three bases). The first tRNA, now without an amino acid, is released and can collect another amino acid.
5. The next codon is exposed and the cycle repeats, adding one amino acid at a time until a stop codon is reached.

This stepwise movement of the ribosome along the mRNA, one codon at a time, means the amino acids are linked in exactly the order set by the codons.



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

Elongation: **two tRNAs bind to the large subunit at once**, so a peptide bond can join the polypeptide to the next amino acid.

Common mistake: Codons are on mRNA and anticodons are on tRNA. Do not describe the anticodon as part of the mRNA or the codon as part of the tRNA.

The genetic code Spec D1.2.8, D1.2.9

Why a triplet code

There are four bases in mRNA and 20 amino acids in proteins. A code of single bases would give 4 combinations, and pairs of bases would give $4^2 = 16$, which is not enough. Triplets give $4^3 = 64$ codons, which is more than enough for 20 amino acids plus start and stop signals.

Degeneracy	More than one codon can code for the same amino acid. For example, GAA and GAG both code for glutamic acid. Of the 64 codons, 61 code for amino acids and 3 are stop codons.
Universality	The same codons code for the same amino acids in almost all living organisms, with only minor variations. This is evidence for common ancestry and allows genes to be transferred between species.
Start codon	AUG, which codes for methionine and marks where translation begins.
Stop codons	UAA, UAG and UGA, which do not code for an amino acid and signal the end of translation.

Using a table of mRNA codons

Codon tables give the amino acid for each mRNA codon. Read the mRNA from the start codon in groups of three, without overlapping, and look up each codon. If you are given DNA, first work out the mRNA: it is complementary to the template strand, and the same as the coding strand with U in place of T.

Part of the mRNA codon table (practice extract)			
Codon	Amino acid	Codon	Amino acid
AUG	Met (start)	GAA	Glu
GCU	Ala	GAG	Glu
AAA	Lys	GUG	Val
UGG	Trp	UUU	Phe
CAU	His	UAA	stop

Maths skill: deducing an amino acid sequence

The template strand of a short DNA sequence reads TAC CGA TTT ACC CTT ATT. Deduce the amino acid sequence.

1. Write the complementary mRNA (A→U, T→A, C→G, G→C): AUG GCU AAA UGG GAA UAA.
2. Look up each codon: AUG = Met, GCU = Ala, AAA = Lys, UGG = Trp, GAA = Glu, UAA = stop.

Answer: Met - Ala - Lys - Trp - Glu (then stop).

Exam tip: Check whether a sequence is DNA (contains T) or mRNA (contains U) before using a codon table. Codon tables are always written for mRNA.

Mutations that change protein structure Spec D1.2.11

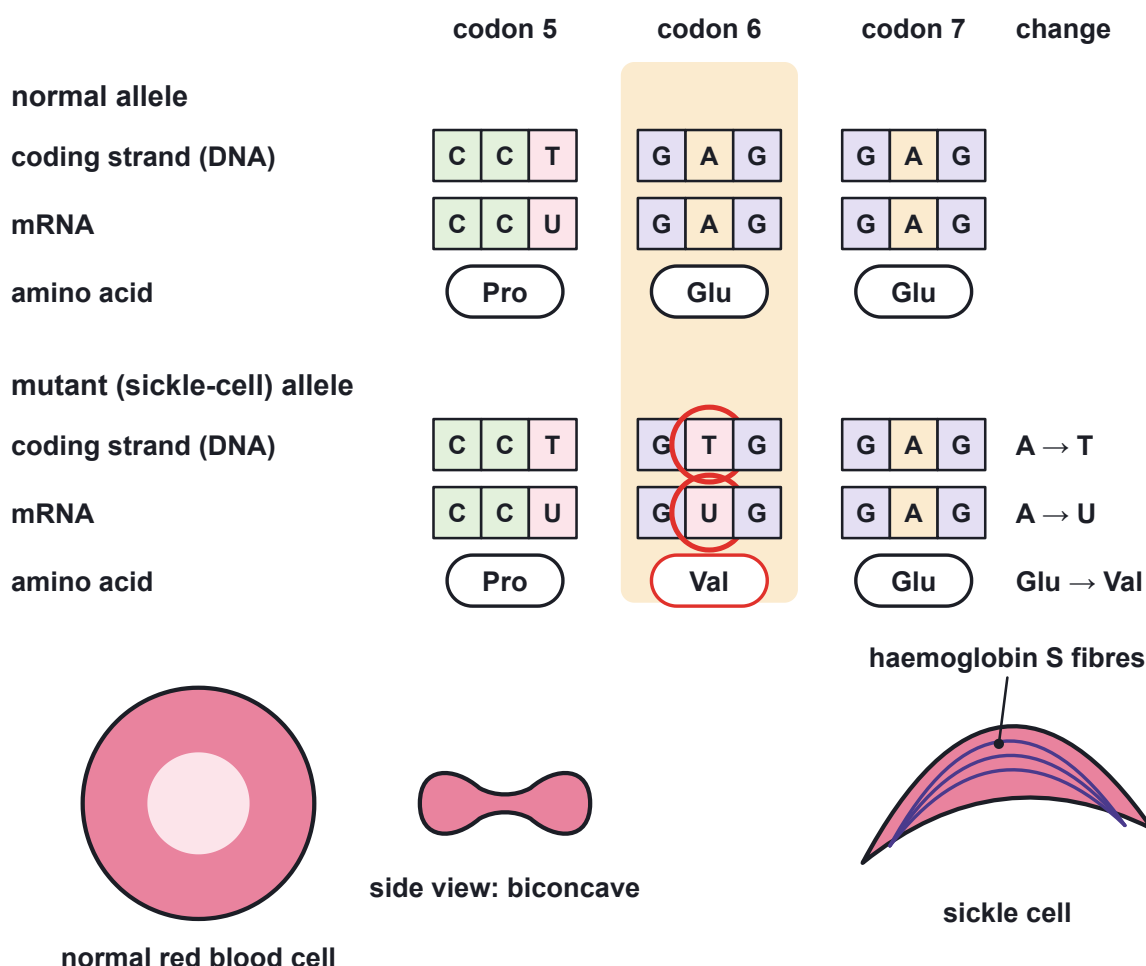
The amino acid sequence of a polypeptide decides how it folds and therefore its three-dimensional structure. A change in the base sequence of a gene can change the codons in the mRNA, and so the amino acid sequence, the folding and the function of the protein.

Example: sickle-cell anaemia

1. The gene for the beta chain of haemoglobin has a **base substitution**: on the coding strand, A is replaced by T (GAG becomes GTG).
2. The mRNA codon becomes **GUG instead of GAG**.

- At the sixth amino acid of the beta chain, **valine replaces glutamic acid**. The resulting form is haemoglobin S.
- Glutamic acid is charged and hydrophilic; valine is non-polar and hydrophobic. At low oxygen concentrations, haemoglobin S molecules stick together into long fibres.
- The fibres distort red blood cells into a rigid sickle shape. Sick cells can block capillaries and carry less oxygen, and they are broken down faster, causing anaemia.

A change to just one base of the gene is enough to alter the protein. This shows how closely protein structure depends on base sequence.



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

Sickle-cell anaemia: one base substitution (GAG → GTG) changes **glutamic acid to valine** in haemoglobin.

Exam tip: In a sickle-cell answer, give the base change, the codon change (GAG → GUG), the amino acid change (Glu → Val) and the effect on haemoglobin and red blood cells. Each is a separate marking point.

Quick check

- Which base on RNA pairs with adenine on the DNA template strand?
- Why must DNA in non-dividing somatic cells stay unchanged by transcription?
- What is meant by the degeneracy of the genetic code?
- How many tRNAs can bind to the large ribosomal subunit at once?

Quick check answers

Protein synthesis

1. Uracil.
2. The same sequences must be transcribed throughout the life of the cell, so they must be conserved.
3. More than one codon can code for the same amino acid.
4. Two.