SL IB Biology

Your notes

Protein Synthesis

Contents

* Transcription in Protein Synthesis
* Translation in Protein Synthesis
* The Genetic Code
* Protein Structure & Mutations

Transcription in Protein Synthesis

Your notes

Synthesis of RNA

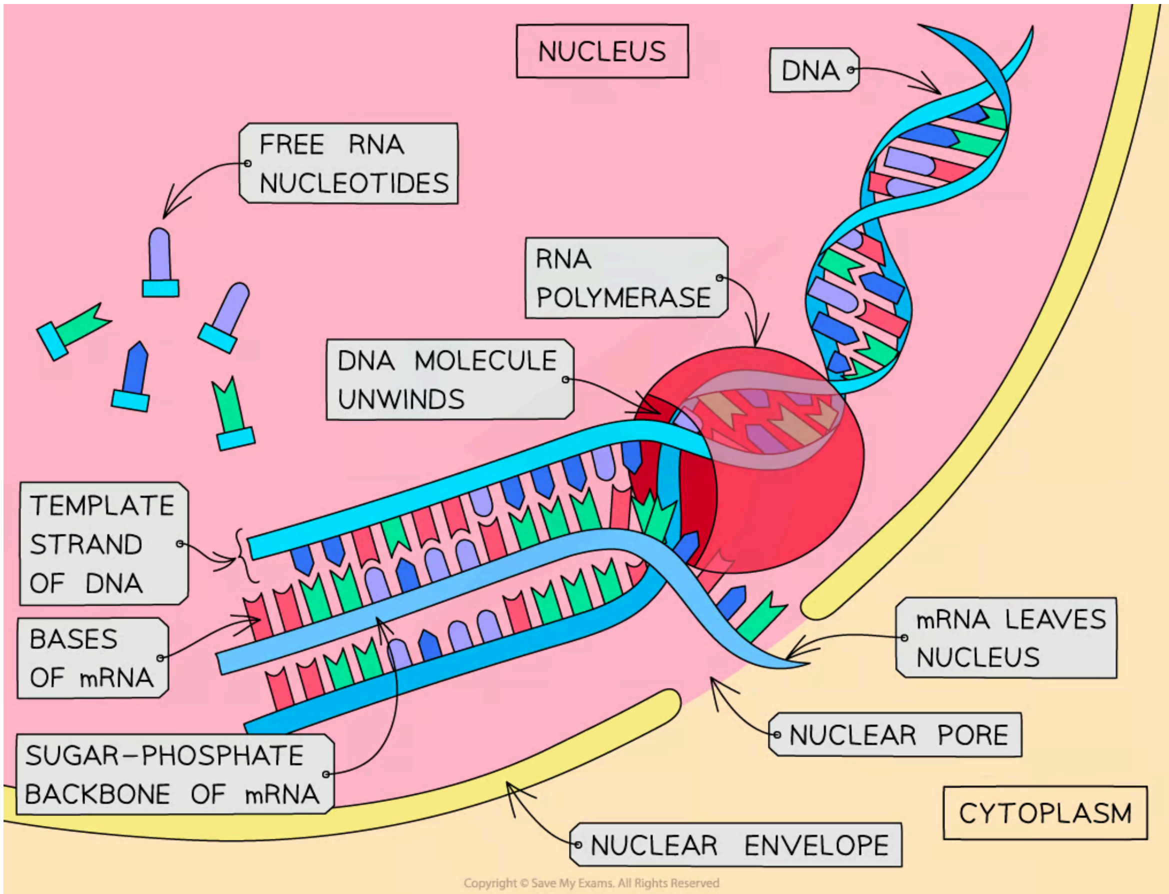
- This process of protein synthesis occurs in **two stages**:
  - **Transcription** – **DNA** is transcribed and an **mRNA** molecule is produced
    - mRNA is a single stranded RNA molecule that transfers the information in DNA from the nucleus into the cytoplasm
    - mRNA production requires the enzyme RNA polymerase
  - **Translation** – **mRNA** (messenger RNA) is translated and an **amino acid sequence** is produced

The process of transcription

- This stage of protein synthesis occurs **in the nucleus** of the cell
- Part of a DNA molecule **unwinds** (the **hydrogen bonds** between the complementary base pairs **break**)
- This exposes the **gene** to be transcribed (the gene from which a particular polypeptide will be produced)
- A complementary copy of the code from the gene is made by building a **single-stranded nucleic acid molecule known as mRNA** (messenger RNA)
- **Free RNA nucleotides** pair up (via hydrogen bonds) with their complementary (now exposed) bases on one strand (the template strand) of the 'unzipped' DNA molecule
- The sugar-phosphate groups of these RNA nucleotides are then **bonded** together by the enzyme **RNA polymerase** to form the sugar-phosphate backbone of the mRNA molecule
- When the gene has been transcribed (when the mRNA molecule is complete), the hydrogen bonds between the mRNA and DNA strands break and the **double-stranded DNA molecule re-forms**
- The mRNA molecule then **leaves the nucleus** via a pore in the nuclear envelope
  - This is where the term *messenger* comes from - the mRNA is despatched, **carrying a message**, to another part of the cell
  - DNA can't make this journey; **it's too big to fit** through the pores in the nuclear envelope

Transcription in the nucleus diagram

Your notes



DNA is transcribed and an mRNA molecule is produced

Examiner Tip

Be careful – DNA polymerase is the enzyme involved in DNA replication; RNA polymerase is the enzyme involved in transcription – don't get these confused.

Your notes

Hydrogen bonding & Complementary Base Pairing

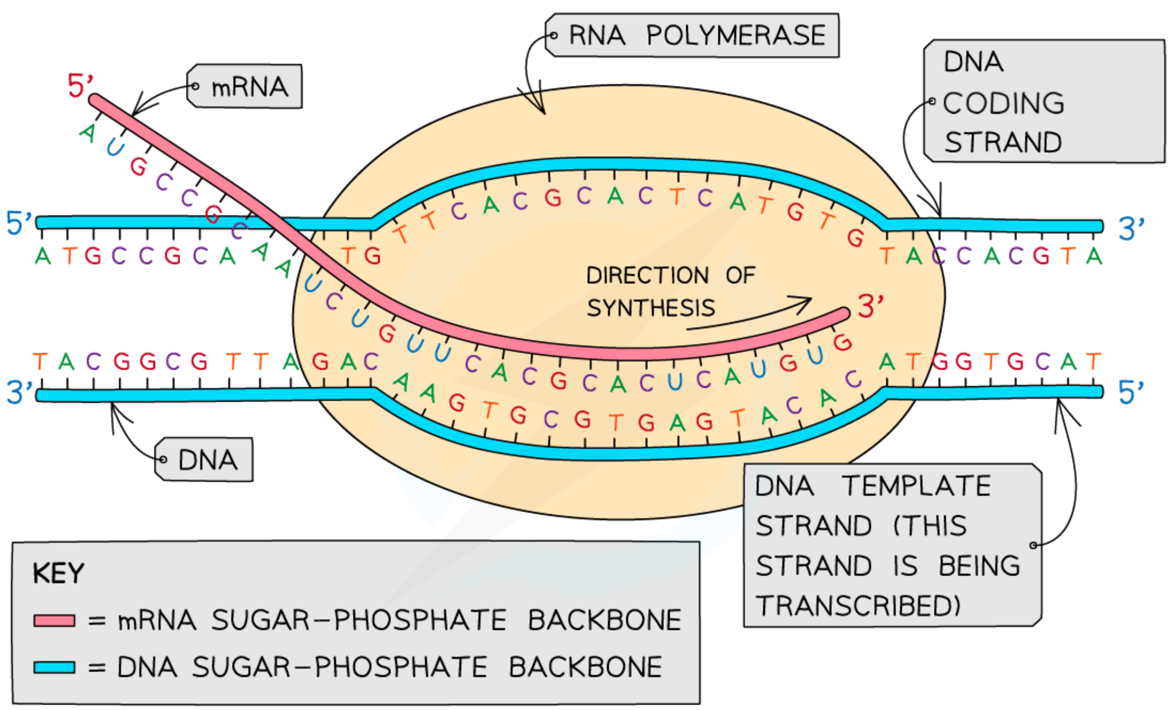
- In the **transcription** stage of protein synthesis, free RNA nucleotides pair up with the exposed bases on the DNA molecule but **only with those bases on one strand of the DNA molecule**
- The RNA will have a complementary base sequence to the DNA strand and will bind to the DNA using **hydrogen bonds**
- The **adenine of the DNA is complementary to uracil on the new RNA strand**, because a thymine RNA nucleotide does not exist

Complementary base pairing between the DNA and the RNA transcript table

DNA template strand code	TAC	GGA	AGA	CTT	GGG
RNA transcript	AUG	CCU	UCU	GAA	CCC

- The strand of the DNA molecule that carries the genetic code is called the **coding strand**
- The opposite DNA strand is called the **template** strand
- To get **an RNA transcript of the coding strand**, the **template strand is the one that is transcribed** to form the mRNA molecule
  - This mRNA molecule will later be translated into an amino acid chain

DNA coding and template strand during transcription diagram



The template strand of the DNA molecule is the one that is transcribed

Your notes

DNA Templates

- DNA is a **very stable** molecule due to the **hydrogen bonding** between the DNA bases of the two strands and the strong phosphodiester bonds between adjacent nucleotides in each strand
  - This means that the genetic code is **not prone to spontaneously breaking or changing**
- This feature allows single DNA strands to act as **reliable templates for transcription** over several generations of cell replication
- In certain types of somatic cells that do not divide during their lifetimes, such as neurones and some types of muscle cells, the genetic sequence is **conserved** due to this stability and **does not degrade over time**

Transcription & Gene Expression

- There are approximately **20,000 protein-coding genes i**n the human genome
- Not every protein is needed in every cell
  - For example, the insulin protein is not needed in cardiac muscles of the heart
- As a result, our specialised cells have a way of **switching certain genes off or on** to match the requirements of the cell. This is called **gene expression**
  - Genes that are expressed are 'switched on' and **undergo the process of transcription and translation**
  - Genes that are not expressed are 'switched off' or silenced, and do not go through the process of transcription and/or translation
- There are **various different mechanisms** in the cell involved in controlling gene expression
- **Transcription is the first stage of gene expression** and so this is a key stage at which gene expression can be switched on or off

Translation in Protein Synthesis

Your notes

Synthesis of Polypeptides

- **Translation** involves taking the genetic code from the mRNA and **synthesising a polypeptide**
  - A polypeptide is a sequence of amino acids covalently bonded together
  - The order of the amino acids is based on the information stored in the genetic code of the mRNA
- This stage of protein synthesis occurs **in the cytoplasm** of the cell
- The **mRNA template** comes from the process of **transcription**, and so translation always takes place following these events
  - After transcription the mRNA moves out of the nuclear pore and diffuses into the cytoplasm towards the ribosome for translation

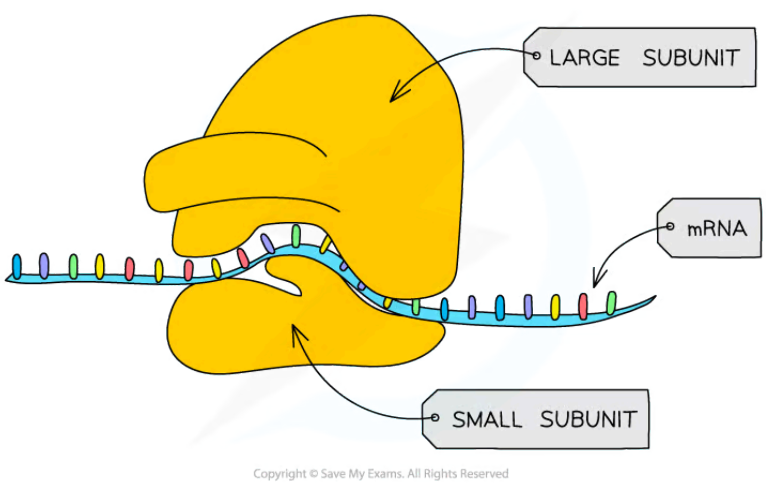
Examiner Tip

Make sure you learn both stages of protein synthesis fully. Don't forget WHERE these reactions take place – transcription occurs in the nucleus but translation occurs in the cytoplasm!

Roles of RNA & Ribosomes in Translation

- After leaving the nucleus, the **mRNA molecule attaches to a ribosome**
- A ribosome is a complex structure that is made of a large and small subunit
  - Ribosomes are themselves made of **proteins** and **RNA** (called ribosomal RNA or **rRNA**)
- There are **binding sites on the subunits** for the various other molecules involved in translation
  - The **mRNA** binds to the **small subunit**
  - **Two tRNA** molecules are able to bind to the **large subunit simultaneously**

mRNA in the ribosome diagram

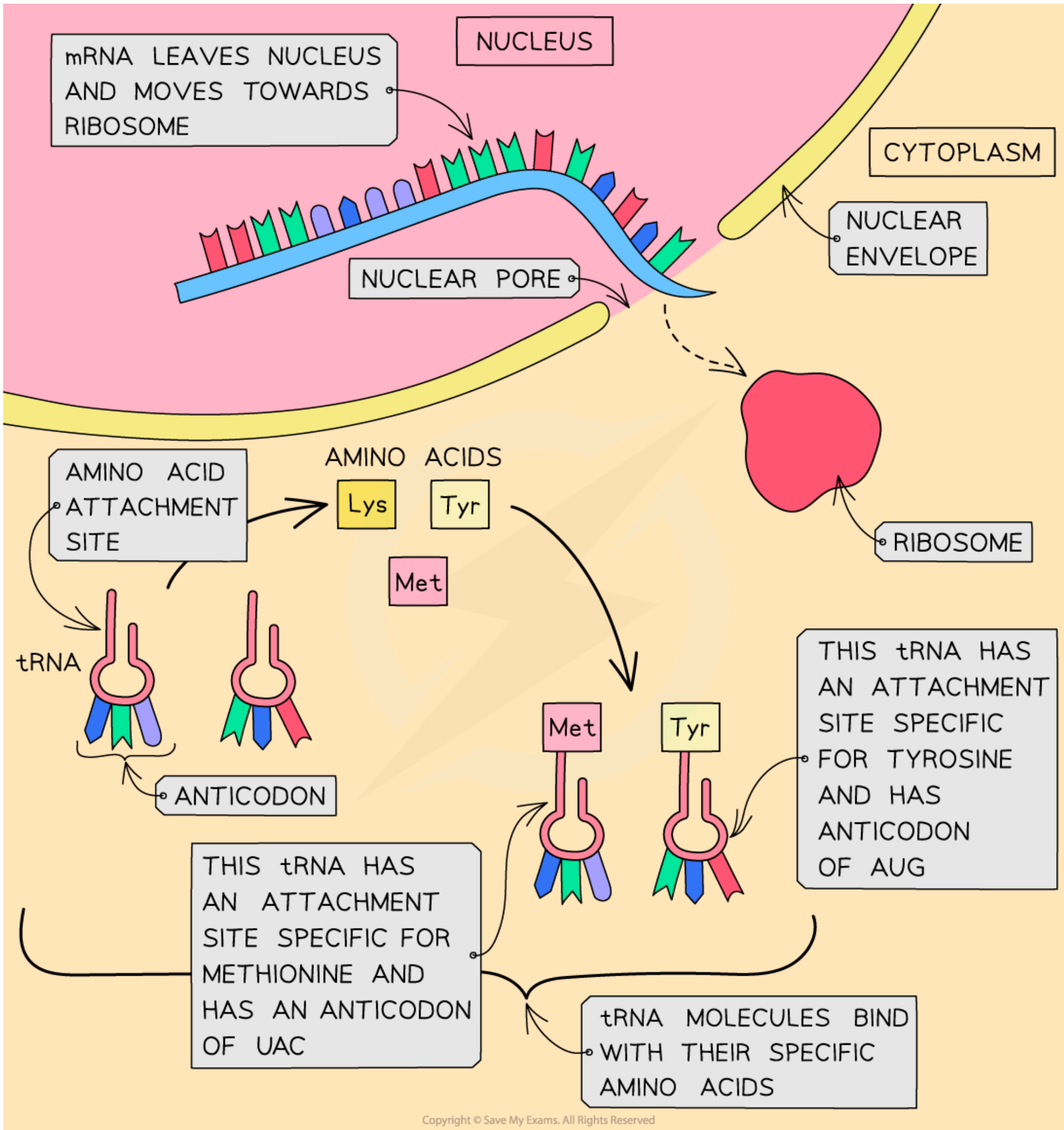


A ribosome is built of large and small subunits, ribosomal RNA and an area on the surface that catalyses the formation of peptide bonds in a newly-synthesised protein

- Translation depends on complementary base pairing between codons on mRNA and anticodons on tRNA
- In the cytoplasm, there are **free molecules of tRNA** (transfer RNA)
- The **tRNA molecules bind with their specific amino acids** (also in the cytoplasm) and bring them to the mRNA molecule on the **ribosome**
- The triplet of bases (anticodon) on each tRNA molecule pairs with a complementary triplet (codon) on the mRNA molecule

tRNA and mRNA before translation diagram

Your notes

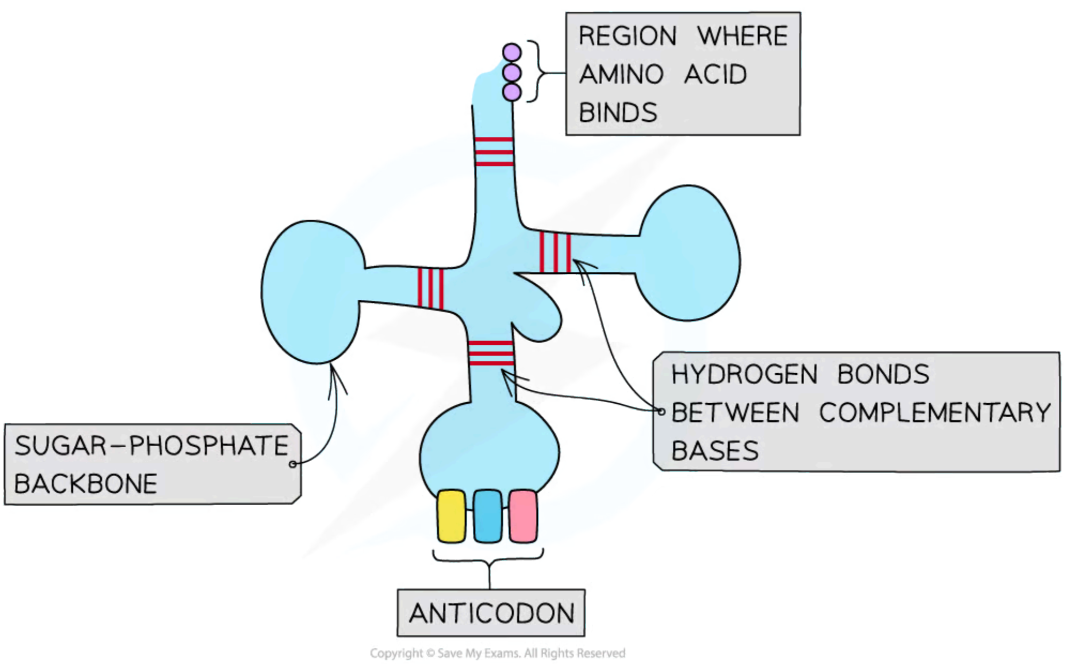


The translation stage of protein synthesis – tRNA molecules bind with their specific amino acids

Codons & Anticodons

- Codons of **three bases** on mRNA correspond to **one amino acid** in a polypeptide
  - A **triplet** is a sequence of three DNA bases that codes for a specific amino acid
  - A **codon** is a sequence of three **mRNA** bases that codes for a specific amino acid
  - A codon is transcribed from the triplet and is complementary to it
- An **anticodon** is a sequence of three **transfer RNA** (tRNA) bases that are complementary to a codon
  - The transfer RNA **carries the appropriate amino acid** to the ribosome
  - The amino acid can then be condensed **onto the growing polypeptide chain**
- Certain codons carry the command to **stop translation** when the polypeptide chain is complete. These are called **stop codons**

Structure of tRNA diagram

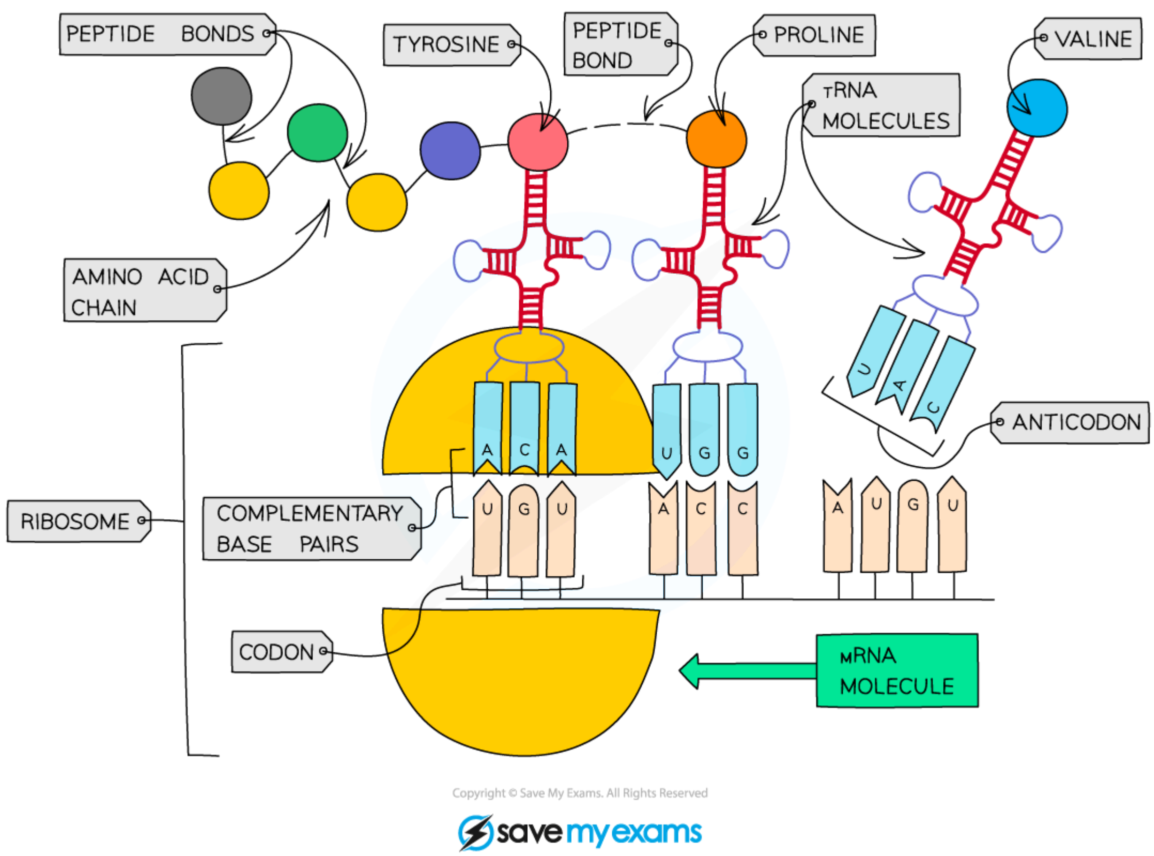


The anticodon is positioned at the bottom of the tRNA molecule and consists of three exposed RNA bases

mRNA and tRNA binding diagram

Your notes

Your notes



Complementary base pairing occurs between the mRNA and the corresponding tRNA molecule, resulting in the correct sequence of amino acids being synthesised into the polypeptide

Analogy: Think of transcription and translation as being like converting between languages

- Each language has its **alphabet**, just as nucleic acids and proteins have their **monomers**
- **Transcription** is like converting text from **English** to **French**
  - The same characters are used, but there are slight differences
  - French uses the same alphabet as English but employs occasionally accented characters like â, é, or ç
  - DNA and RNA employ largely the same monomers but with slight differences in the two pentose sugars and U replacing T.

- **Translation** is like converting text from a Western language to a language that uses a different alphabet, like **Japanese**
  - A completely **different set of characters** is used
  - The sequence of characters is **unrecognisable** from the original
  - If we could see them, a chain of amino acids would look nothing like a chain of nucleotides

Transcription and Translation Can be Likened to the Conversion Between Languages Table

Transcription	DNA → RNA	Similarities	English → French		Similarities
DNA → RNA	TTACAGCTC → AAUGUCGAG	Both use a similar set of monomers (with a slight difference; U replaces T)	"I received biology lessons at my school"	"J'ai reçu des cours de biologie à mon école"	Both use a similar alphabet (with slight differences: ç, à, é, Ô etc)

Translation	RNA → Protein	Differences	French → Japanese		Differences
RNA → protein	AAUGUCGAG → Asn–Val–Glu	Both use different monomers (nucleotides & amino acids)	"J'ai reçu des cours de biologie à mon école"	学校で生物学の授業を受けました	Both use different alphabets

Examiner Tip

Remember that complementary base pairing in RNA means that:

- Adenine (A) will pair up with Uracil (U)
- Cytosine (C) will pair up with Guanine (G)

So if an mRNA codon has a sequence of **CAG**, then its complementary tRNA anticodon will have a sequence of **GUC**.

Your notes

The Genetic Code

Features of the Genetic Code

- The sequence of DNA nucleotide bases found within a gene is determined by a **triplet** (three-letter) **code**
- Each sequence of **three bases** (i.e. each triplet of bases) in a gene codes for **one amino acid**
- These triplets code for different amino acids – there are 20 different amino acids that cells use to make up different proteins
- For example:
  - CAG codes for the amino acid valine
  - TTC codes for the amino acid lysine
  - GAC codes for the amino acid leucine
  - CCG codes for the amino acid glycine

- Some of these triplets of bases code for **start** (TAC – methionine) **and stop signals**
- These start and stop signals **tell the cell where individual genes start and stop**
- As a result, the cell **reads the DNA correctly** and **produces the correct sequences of amino acids** (and therefore the correct protein molecules) that it requires to function properly
- The genetic code is **non-overlapping**
  - Each base is **only read once** in which codon it is part of

- There are **four bases**, so there are **64** different codons (triplets) possible ($4^3 = 64$), yet there are only 20 amino acids that commonly occur in biological proteins
  - This is why the code is said to be **degenerate**: multiple codons can code for the same amino acids
  - The degenerate nature of the genetic code can **limit** the effect of **mutations**

- The genetic code is also **universal**, meaning that almost every organism uses the **same code** (there are a few rare and minor exceptions)
- The **same triplet codes code for the same amino acids in all living things** (meaning that genetic information is transferable between species)
  - The universal nature of the genetic code is why genetic engineering (the transfer of genes from one species to another) is possible

Deducing Amino Acid Sequences

- By observing the **genetic code in the mRNA** it is possible to determine the **sequence of amino acids** that are coded for in the **polypeptide**

mRNA codons and amino acids table

SECOND LETTER

FIRST LETTER	U	C	A	G	THIRD LETTER
U	UUU } Phe UUC } UUA } Leu UUG }	UCU } Ser UCC } UCA } UCG }	UAU } Tyr UAC } UAA Stop UAG Stop	UGU } Cys UGC } UGA Stop UGG Trp	U C A G
C	CUU } Leu CUC } CUA } CUG }	CCU } Pro CCC } CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } Arg CGC } CGA } CGG }	U C A G
A	AUU } Ile AUC } AUA } AUG Met	ACU } Thr ACC } ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G
G	GUU } Val GUC } GUA } GUG }	GCU } Ala GCC } GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } Gly GGC } GGA } GGG }	U C A G

Your notes

Your notes

Worked example

Use the **rules of base-pairing** and the **mRNA Codons and Amino Acids Table** (above) to deduce the amino acid sequence coded for by the following DNA **coding strand** sequence TTC GAG CAT TAC GCC

Answer:

Step 1: Work out the template sequence using A-T and C-G base pairing rules

AAG CTC GTA ATG CGG

Step 2: Work out the mRNA codons, complementary to the template strand

UUC GAG CAU UAC GCC

Step 3: Use the mRNA Codons and Amino Acids Table (above) to work out the first amino acid

First base in codon = U, second base = U, third base = C

So we're looking in the top-left box of the table; this amino acid is **Phe**

Step 4: Repeat for the remaining 4 codons

GAG = Glu

CAU = His

UAC = Tyr

GCC = Ala

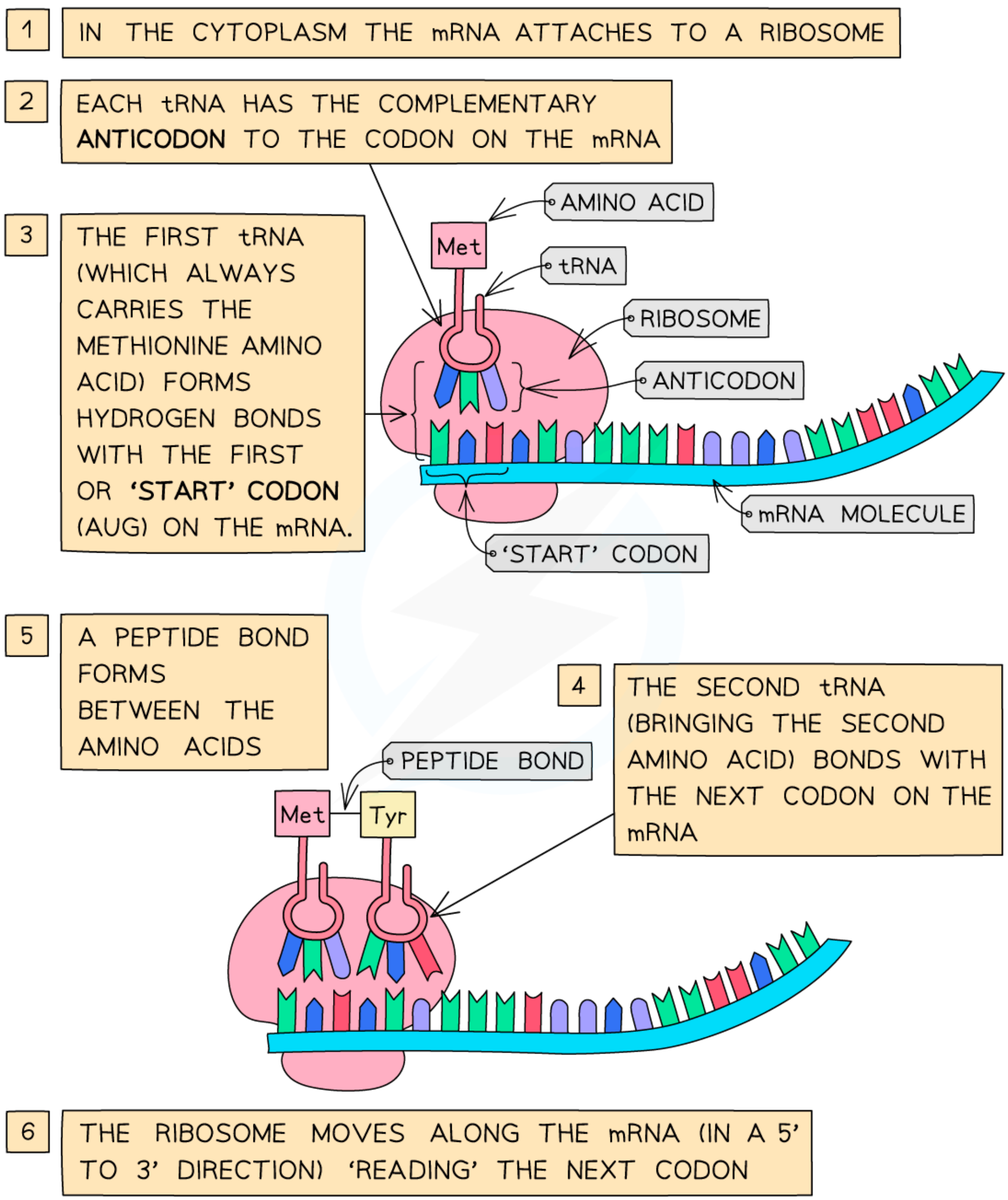
The final sequence of amino acids is Phe-Glu-His-Tyr-Ala

Elongation of the Polypeptide Chain

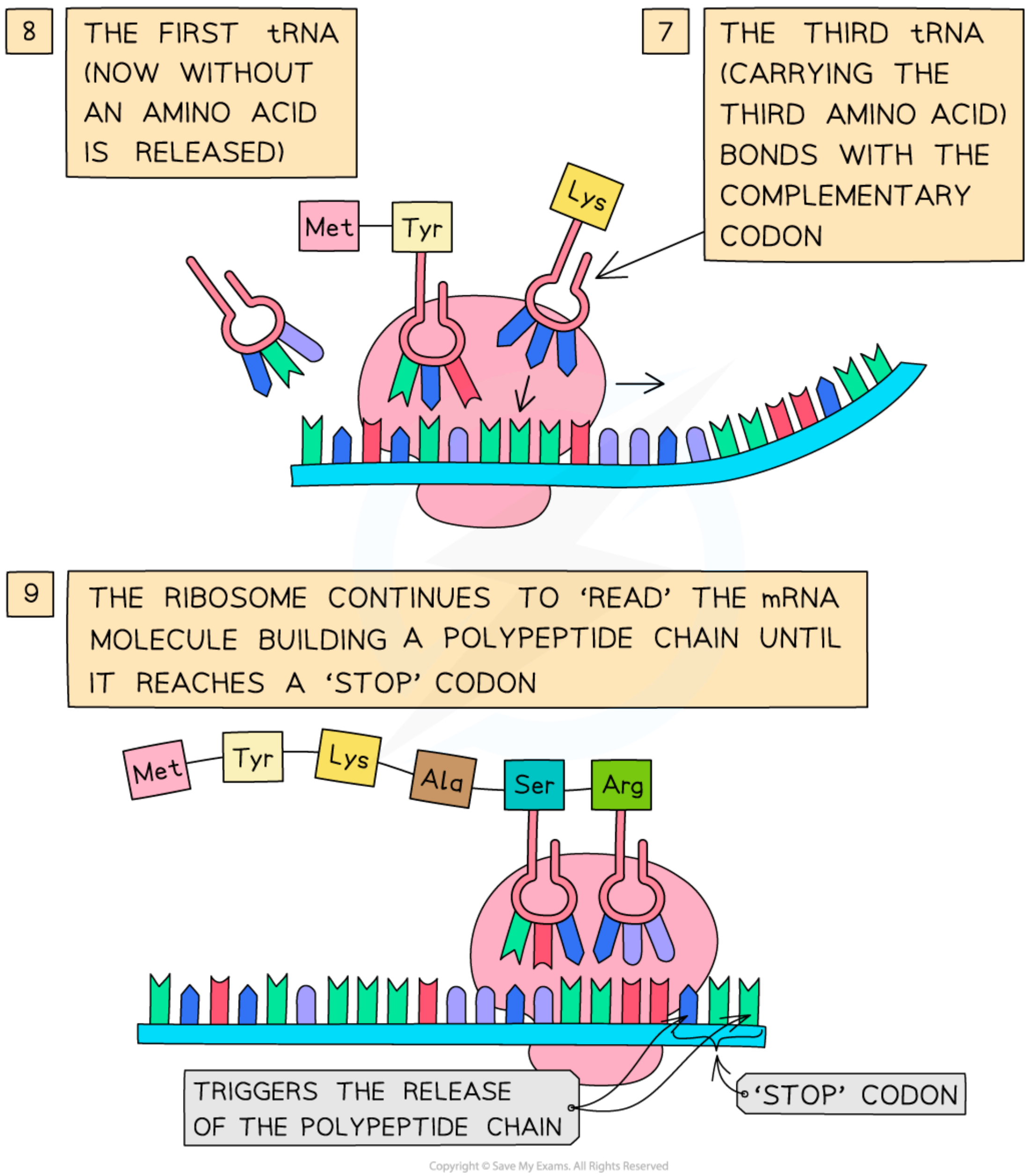
Your notes

- During translation **two tRNA molecules fit onto the ribosome at any one time**, bringing the amino acid they are each carrying side by side
  - The ribosome will move along the mRNA molecule, one codon at a time
- A **peptide bond** is then formed (by condensation) between the two amino acids
  - The formation of a peptide bond between amino acids is an anabolic reaction
  - It **requires energy,** in the form of **ATP**
  - The ATP needed for translation is provided by the **mitochondria** within the cell
- This process continues until a **'stop' codon** on the mRNA molecule is reached – this acts as a signal for translation to stop and at this point the amino acid chain coded for by the mRNA molecule is complete
- This amino acid chain is then **released from the ribosome** and forms the final polypeptide

The process of translation diagram


1
IN THE CYTOPLASM THE mRNA ATTACHES TO A RIBOSOME
2
EACH tRNA HAS THE COMPLEMENTARY **ANTICODON** TO THE CODON ON THE mRNA
3
THE FIRST tRNA (WHICH ALWAYS CARRIES THE METHIONINE AMINO ACID) FORMS HYDROGEN BONDS WITH THE FIRST OR **'START' CODON** (AUG) ON THE mRNA.
AMINO ACID
Met
tRNA
RIBOSOME
ANTICODON
mRNA MOLECULE
'START' CODON
5
A PEPTIDE BOND FORMS BETWEEN THE AMINO ACIDS
4
THE SECOND tRNA (BRINGING THE SECOND AMINO ACID) BONDS WITH THE NEXT CODON ON THE mRNA
PEPTIDE BOND
Met
Tyr
6
THE RIBOSOME MOVES ALONG THE mRNA (IN A 5' TO 3' DIRECTION) 'READING' THE NEXT CODON

Your notes



The translation stage of protein synthesis – an amino acid chain is formed

Protein Structure & Mutations

Protein Structure & Mutations

- A **gene mutation** is a change in the sequence of bases in a DNA molecule; this may result in a new allele
  - Mutations occur **all the time** and occur **randomly**
  - Mutations are **copying errors** that take place when DNA is replicated during **S phase** of interphase
- As the DNA base sequence determines the sequence of amino acids that make up a polypeptide, **mutations in a gene** can sometimes lead to a **change in the polypeptide** for which the gene codes
- Most mutations are **harmful** or **neutral** (have no effect) but some can be **beneficial**
- **Inheritance** of mutations:
  - Mutations present in normal body cells are **not inherited**; they are eliminated once the affected cells die
  - Mutations within gametes are inherited by offspring, so can lead to **heritable** genetic conditions
- **Point mutations** are mutations where **one base** in the DNA sequence is altered; this can result in a **changed amino acid** at this location

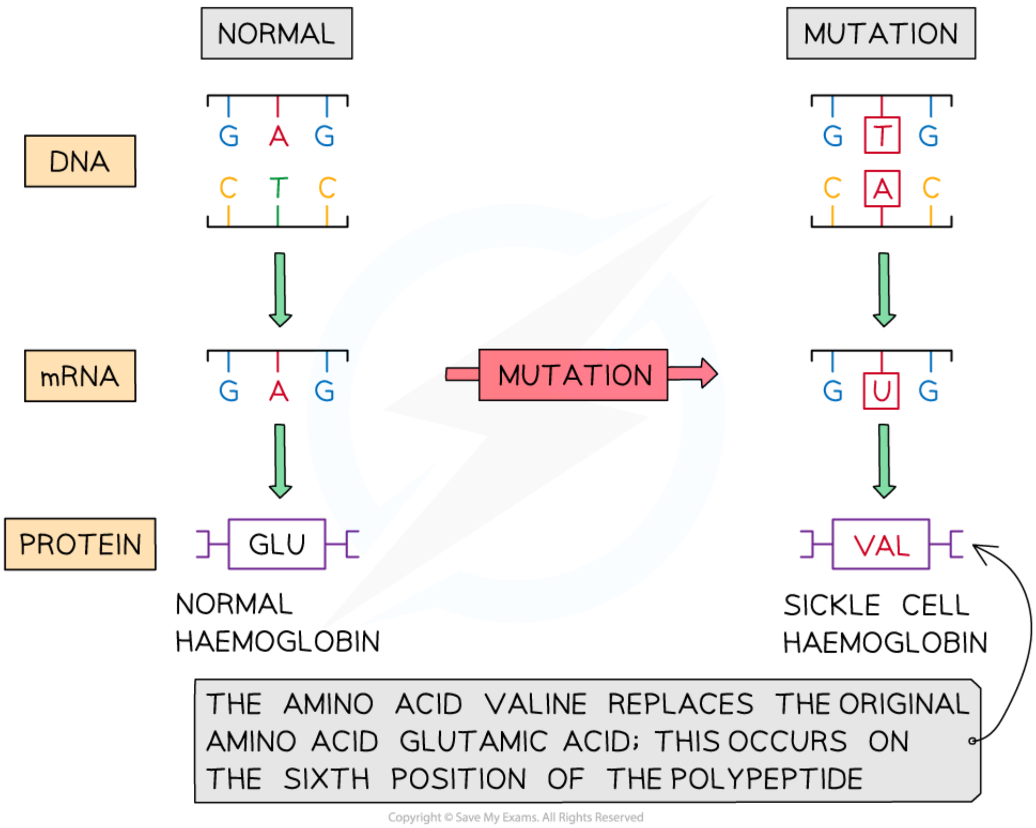
Example of a point mutation: sickle cell disease

- A small change to a gene can have **serious consequences** for an organism
- Sickle cell disease is a genetic disorder caused by **a single point mutation** within the gene that codes for the alpha-globin polypeptide in haemoglobin (Hb)
  - Most humans have the allele **Hb^A**
  - The mutation results in a new allele **Hb^S**

The sickle cell mutation

- Within the haemoglobin gene a point mutation changes the DNA triplet G**A**G to G**T**G on the coding strand
- The resulting DNA triplet (C**A**C) on the template strand is transcribed into the **mRNA codon GUG**, instead of GAG
- During translation the amino acid **valine** (Val) replaces the original amino acid **glutamic acid** (Glu)
  - This occurs at the **sixth position** of the polypeptide

Sickle cell anaemia point mutation diagram

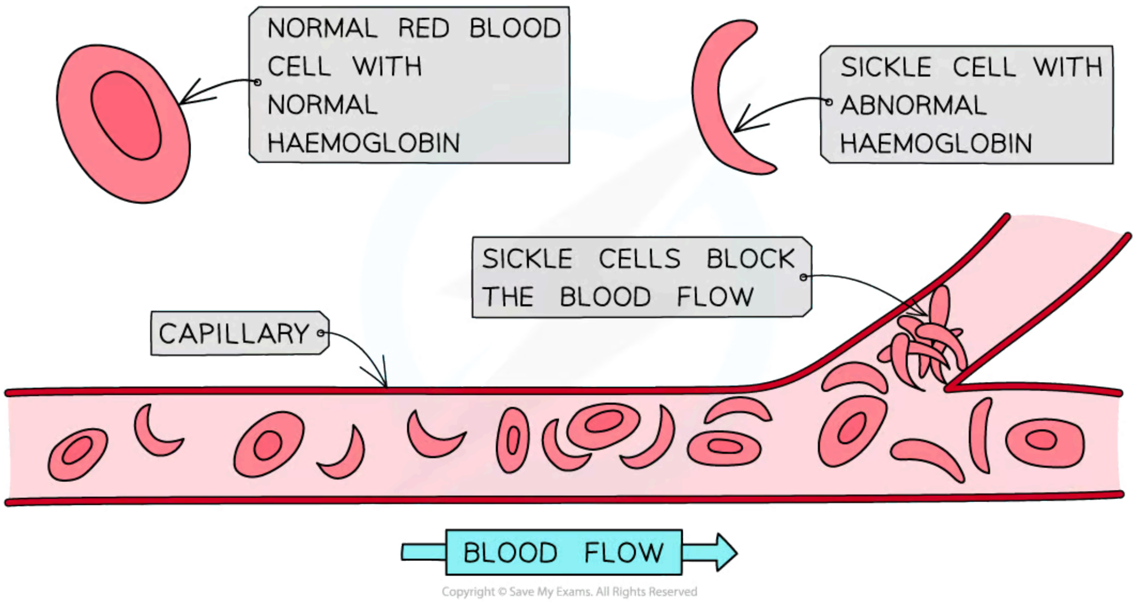


A base substitution on the DNA molecule results in a change in the amino acid at position 6 of the haemoglobin polypeptide, altering the overall structure and function of the protein

The effects

- The protein haemoglobin **S** is produced instead of haemoglobin A; this causes a **distortion in the shape of red blood cells**, resulting in a **sickle shape**
- Sickle-shaped red blood cells:
  - Have a **limited oxygen-carrying capacity**
  - **Block the capillaries** and limit the flow of normal red blood cells
- People with sickle cell anaemia suffer from **acute pain**, **fatigue** and **anaemia**
- There is a **correlation** between the global distribution of sickle cell disease and **malaria**
  - In areas with increased malaria cases there is an increased frequency of sickle cell alleles; this is thought to be due to increased resistance to the malaria parasite in individuals with the Hb^S allele

Sickled cells diagram



Sickled cells can block the flow of blood through the capillaries, restricting oxygen supply to the tissues

- You will cover more on mutations later in the course; see this link

Your notes