DP IB Biology: HL

Your notes

7.3 Translation

Contents

7.3.1 Translation

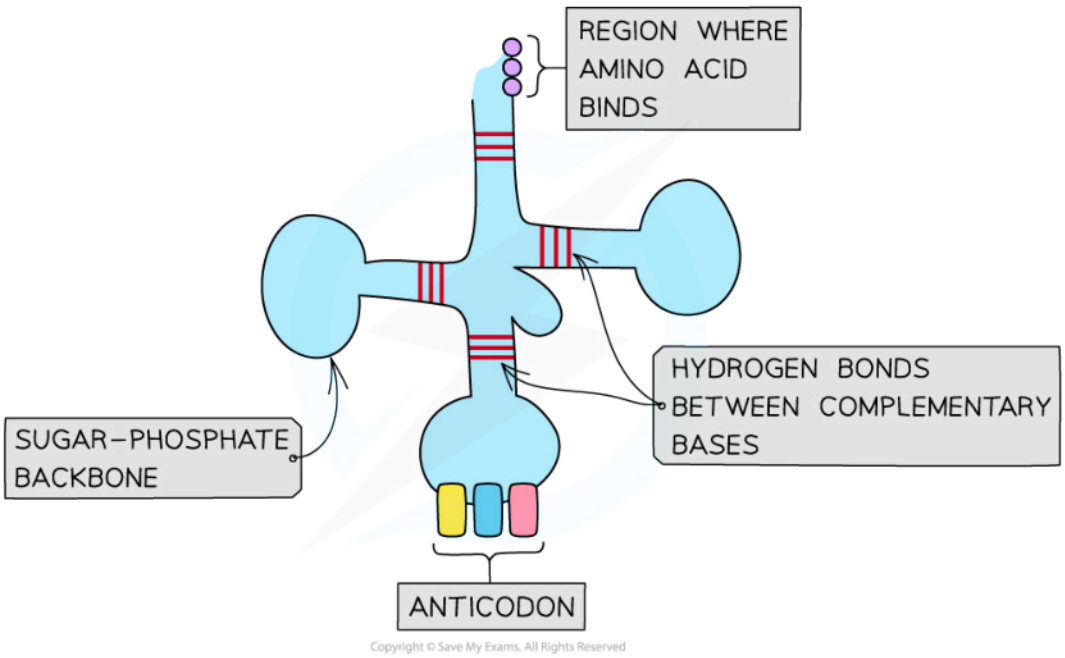

Your notes

Initiation of Translation

Initiation of translation involves assembly of the components that carry out the process.

- During translation, the specific sequence of messenger RNA (**mRNA**) is translated to produce a polypeptide chain consisting of amino acids
  - mRNA is a single stranded, linear, RNA molecule that transfers the information in DNA from the nucleus into the cytoplasm
- Translation is categorised into three stages: **initiation**, **elongation** and **termination**
- Translation occurs in the cytoplasm at complex molecules made of protein and RNA called **ribosomes**
  - Ribosomes have a **two-subunit** (large and small) structure that helps bind mRNA
  - Ribosomes have **three tRNA binding sites** termed "E" (**exit**), "P" (**peptidyl**) and "A" (**aminoacyl**)
    - At the **A site** the mRNA codon joins with the tRNA **a**nticodon
    - At the **P site** the amino acids attached to the tRNA are joined by **p**eptide bonds
    - At the **E site** the tRNA **e**xits the ribosome
- Another key molecule in translation is **transfer RNA** (tRNA) that decodes mRNA
  - tRNA molecules are single stranded RNA molecules that **fold** to form a clover-shaped structure
    - The folded structure is held together by **hydrogen bonds** between bases at different points on the strand
    - tRNA molecules are the shortest of the RNA molecules, being only around 80 nucleotides in length
    - There are 20 different types of tRNA molecule, one for each of the amino acids involved in protein synthesis
  - tRNA molecules have a region that binds to a **specific amino acid** as well as a three-nucleotide region called an **anticodon** that is **complementary to the codon on mRNA**
  - The role of tRNA molecule is to carry a specific amino acid to the ribosome

Your notes



Structure of tRNA

- In eukaryotic cells, the **mRNA molecule leaves the nucleus** through the nuclear pores
- Translation is initiated by the following process
  - A **small ribosomal subunit** attaches to the 5' end of mRNA
  - An **initiator tRNA** molecule carrying the amino acid **methionine** binds to the small ribosomal subunit
    - The initiator tRNA occupies the "P" site on the ribosome
  - The ribosome moves along the mRNA until it locates a start codon (AUG)
  - The **large ribosomal subunit binds to the small subunit**
    - **Elongation** of the polypeptide can begin

Elongation of the Polypeptide

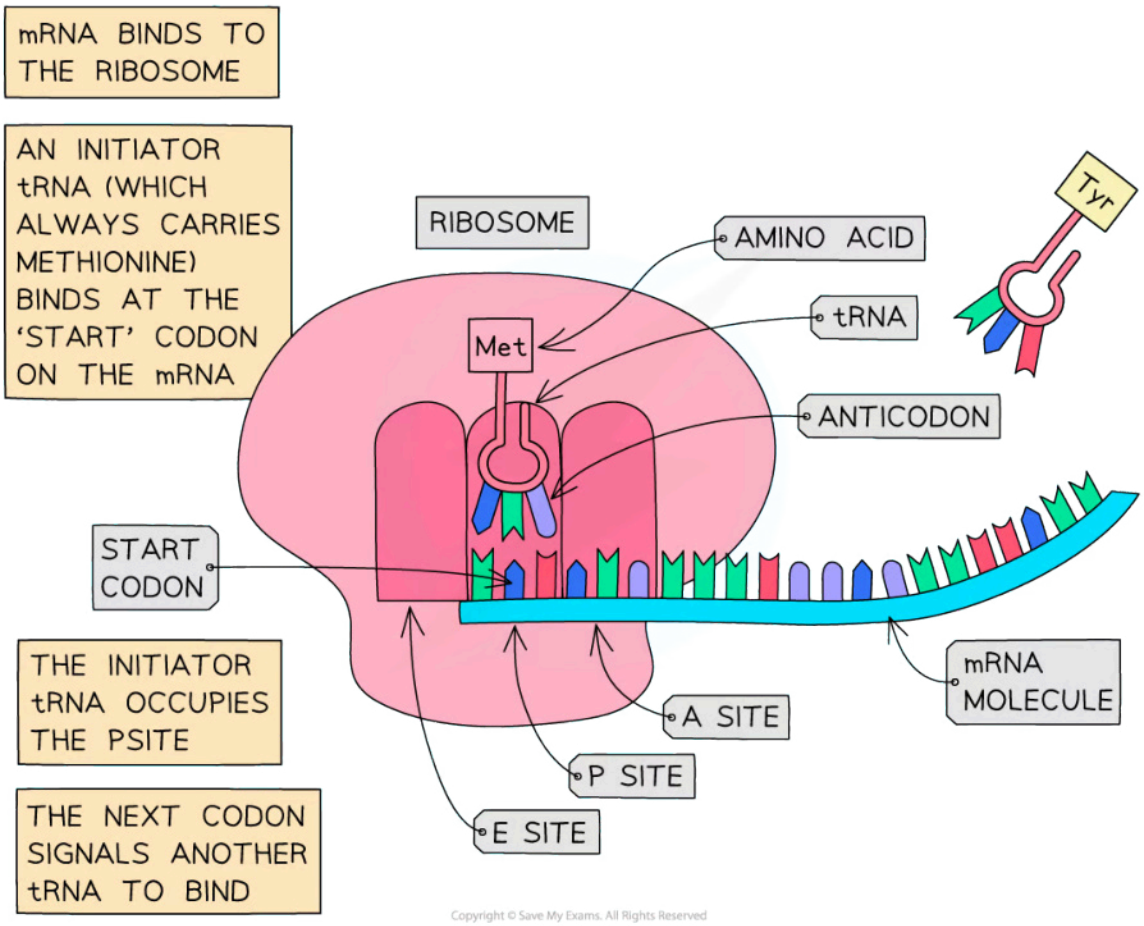

Your notes

- The initiator tRNA currently occupies the "**P**" site, the next codon on the mRNA signals for the corresponding tRNA to bind at the "**A**" site
  - The two amino acids (attached to the tRNAs) are **linked with a peptide** bond, forming a dipeptide
- Synthesis of the peptide chain now involves a **repeated cycle of events**
  - In the cytoplasm, free tRNA molecules bind to their corresponding amino acids and transport them to the ribosome
  - The ribosome shifts along the mRNA one codon (three bases) at a time
    - The initiator tRNA in the "P" site moves to the "E" site which **releases** it
    - The tRNA carrying the peptide chain moves from the "A" site to the "P" site
    - The **next mRNA codon** is exposed and a tRNA with the complementary anticodon binds to the unoccupied "A" site whilst its amino acid is linked to the polypeptide chain
- The cyclical process is repeated as **new amino acids** are **added to the growing chain**

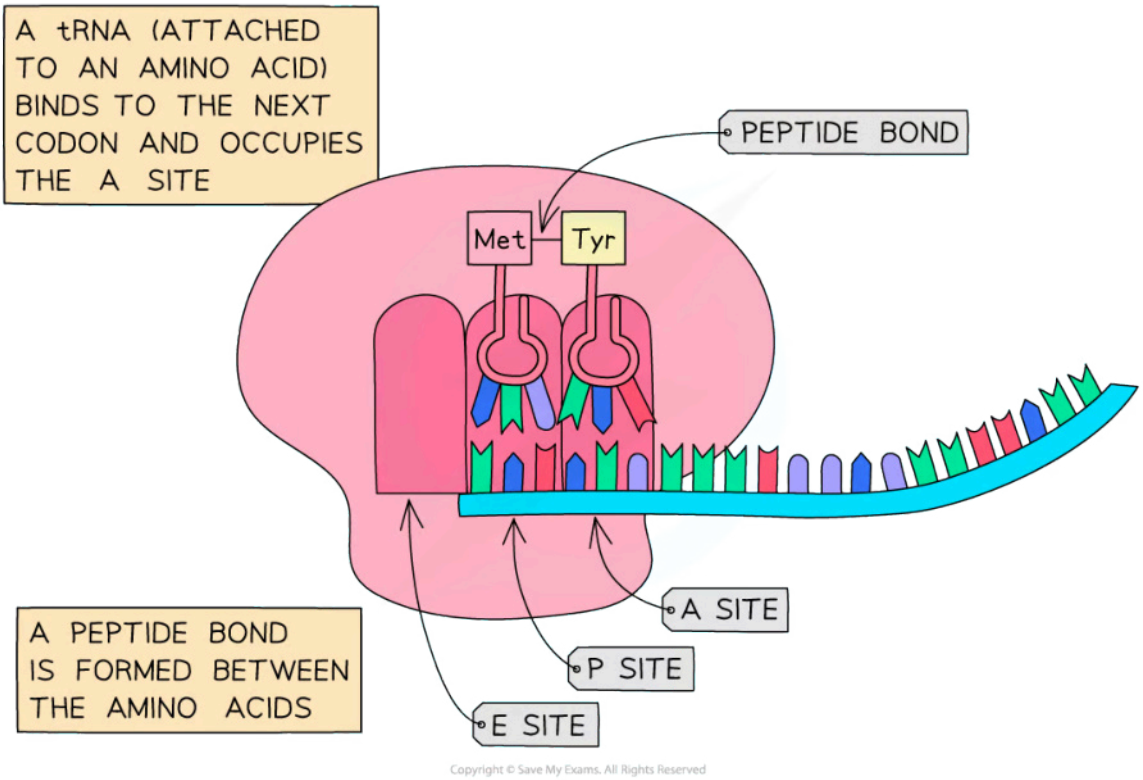
Termination of Translation

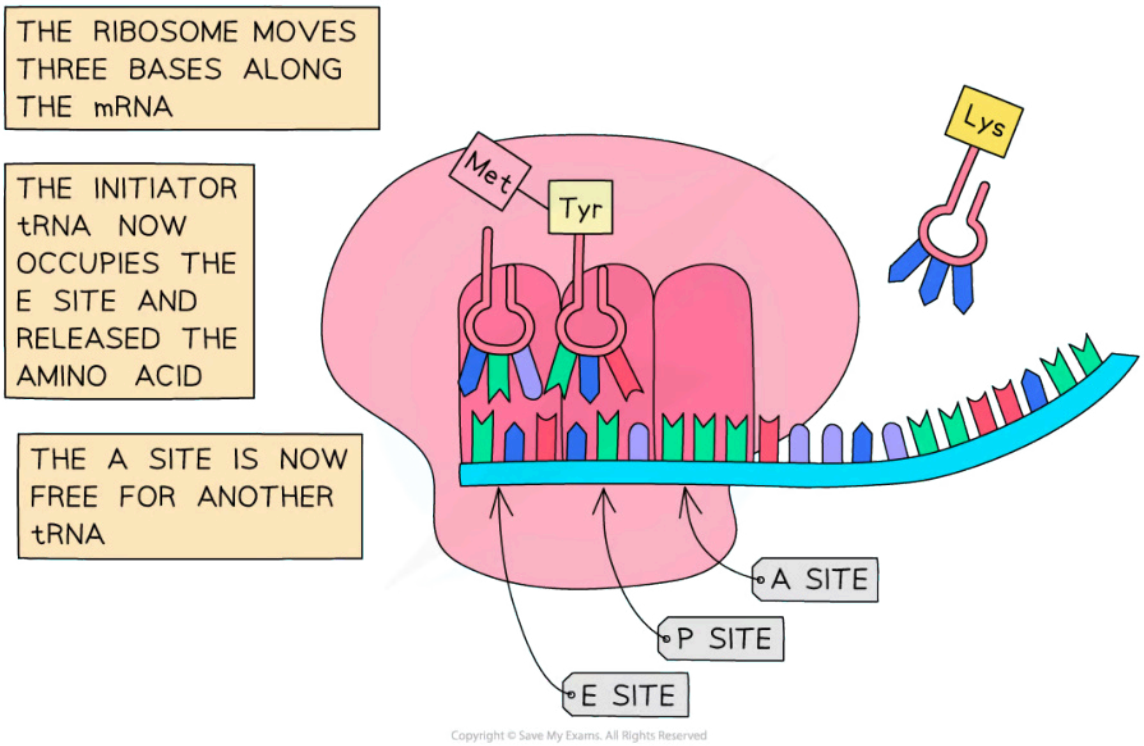
Your notes

- The process of elongation continues until one of three **'stop' codons** (UAA, UAG and UGA) on the mRNA molecule is reached
  - Stop codons do not code for a tRNA molecule but act as a **signal for translation to stop**
- The **polypeptide chain** and **mRNA** are **released** from the ribosome
- The **ribosome disassembles** back into two separate subunits
  - And can await the arrival of the next mRNA molecule



Your notes



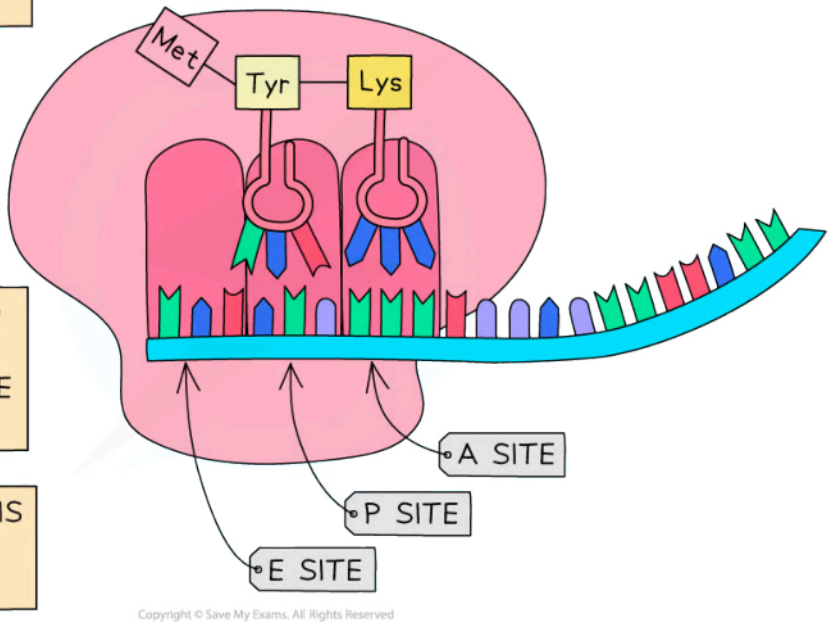


Your notes

THE INITIATOR tRNA IS RELEASED FROM THE E SITE

A tRNA BINDS TO THE NEXT CODON AND OCCUPIES THE A SITE

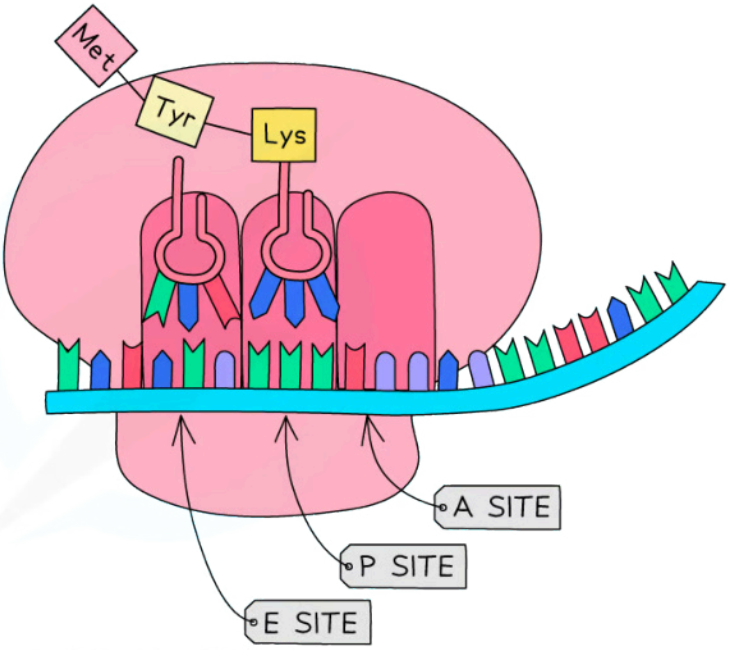
A PEPTIDE BOND IS FORMED BETWEEN THE AMINO ACIDS

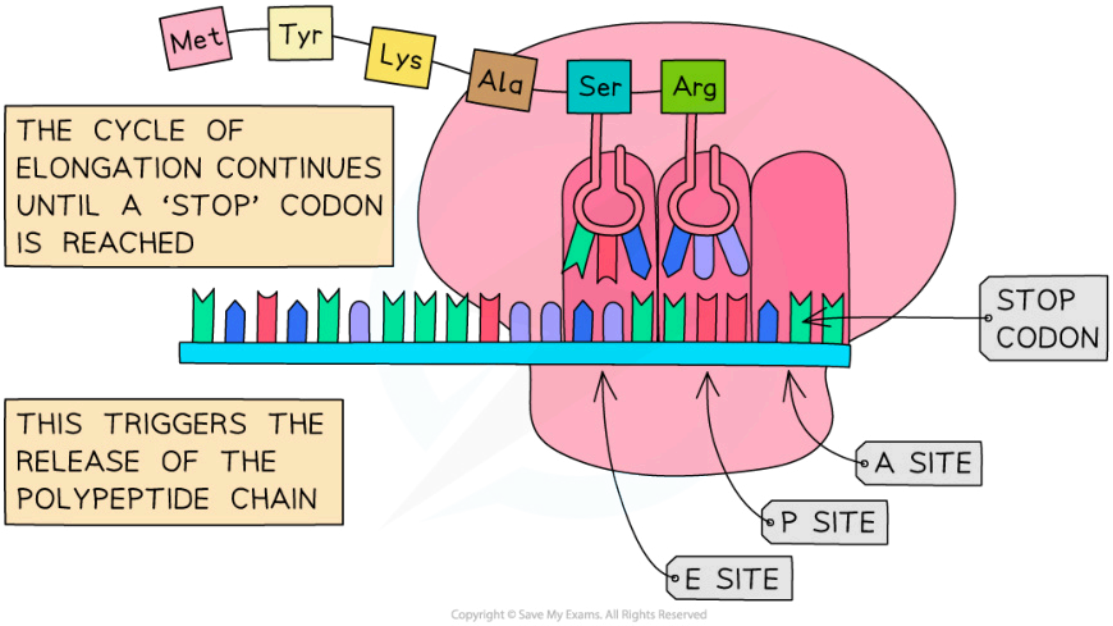


THE RIBOSOME MOVES THREE BASES ALONG THE mRNA

THE A SITE IS NOW FREE FOR ANOTHER tRNA

THE CYCLE OF ELONGATION CONTINUES TO BUILD THE POLYPEPTIDE CHAIN





Following the initiation of protein synthesis, translation involves a repeated cycle of events to build the polypeptide chain, tRNA molecules move into the A, P and E sites as the ribosome reads the mRNA

Examiner Tip

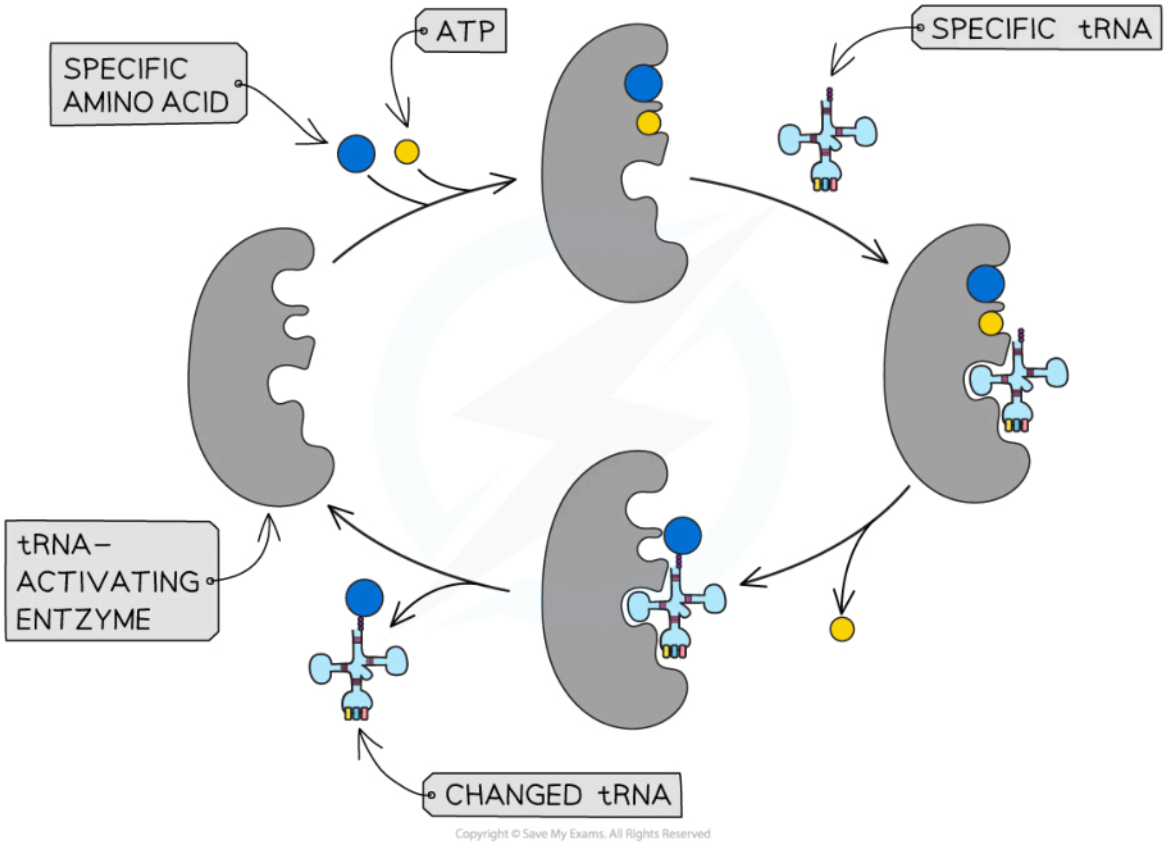
You don't need to remember the precise base sequences of start and stop codons for your examination.

Your notes

tRNA-activating Enzymes

Your notes

- Amino acids are paired to specific tRNA molecules through the action of **tRNA-activating enzymes**
  - Each tRNA activating enzyme recognises a **specific** tRNA molecule
- tRNA-activating enzymes, in common with most enzymes, are **substrate-specific** and recognise the correct tRNA molecules by their shape
  - Nucleotide sequence variability between tRNA molecules results in **variation in their three-dimensional structure**
  - Active sites of tRNA-activating enzymes are optimised **to bind a specific tRNA**
- Initially, a tRNA-activating enzyme binds to **ATP** and a **specific amino acid**
- The active site of the enzyme attracts a **conformationally-specific tRNA molecule**
- The tRNA molecule is bound to the amino acid using ATP (phosphorylation) to **create a high energy bond**
  - The stored energy in this bond will be used later in **peptide bond formation** to link the amino acid to the growing polypeptide chain
  - This is an example of how an **anabolic reaction** like protein synthesis utilises the energy stored in ATP
- A tRNA molecule with an amino acid attached is called a **charged tRNA**



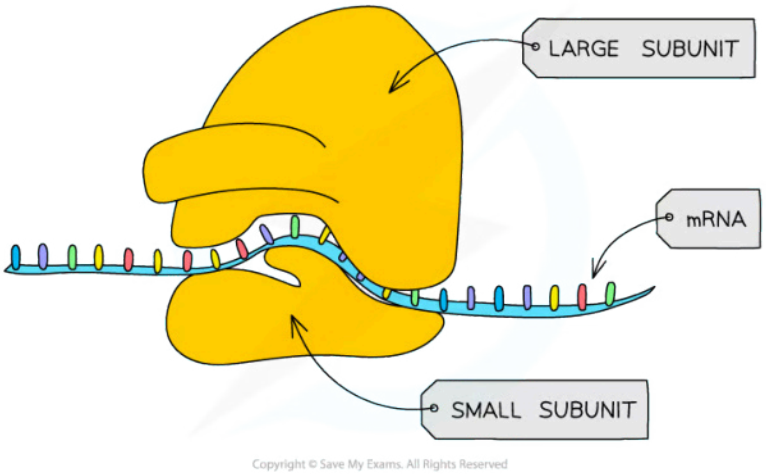
Specific tRNA-activating enzymes are involved in charging an amino acid to a specific tRNA molecule

7.3.2 Ribosomes

Your notes

Structure of Ribosomes

- **Ribosomes** are found in cells
  - Either **freely** in the cytoplasm (of all cells)
  - Or bound to the **endoplasmic reticulum** (ER) to form **rough ER** (only in eukaryotic cells)
- Ribosomes are the **site of protein synthesis**
- They consist of a **large** and a **small subunit** composed of **protein** and **ribosomal RNA** (rRNA)
  - Protein provides **structure** to the ribosome
  - rRNA **facilitates the binding of mRNA and tRNA** and catalyses the formation of peptide bonds between amino acids
- Ribosomes have **three tRNA binding sites** and **one mRNA binding site**
- mRNA sits in a **groove** between the two subunits and the **ribosome moves along**, forming a polypeptide as it travels



A diagram of a ribosome, showing the small and large subunits

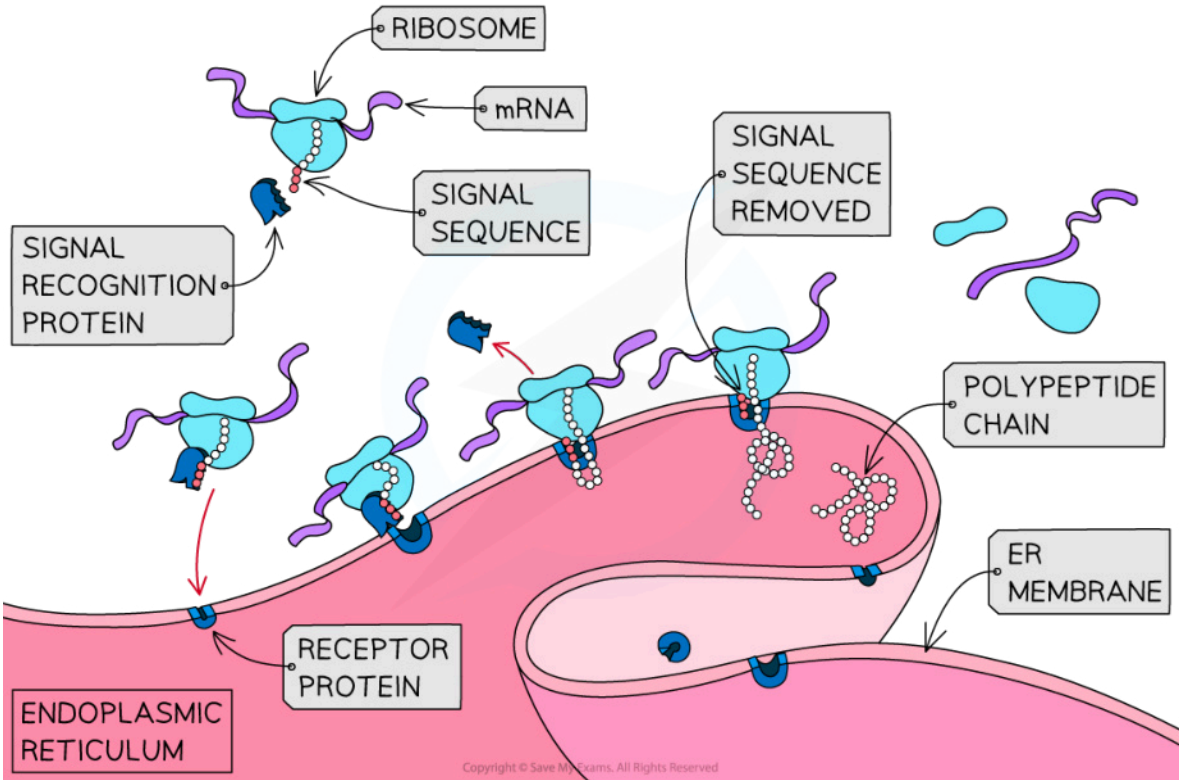
Free Ribosomes

- In eukaryotic cells, protein synthesis commonly occurs at **free ribosomes** in the cytoplasm
- Free ribosomes can **move** within the cytoplasm and synthesise proteins for use primarily **within the cell**
  - As opposed to proteins destined to be secreted extracellularly
  - Proteins synthesised on free ribosomes are destined for use **within the cytosol** (the fluid part of the cytoplasm)
    - And within **large organelles** such as mitochondria and chloroplasts

Bound Ribosomes

Your notes

- Eukaryotic cells make thousands of proteins that need to be **delivered to the correct location**, sometimes in different tissues/organs altogether
- When free ribosomes make proteins destined for lysosomes, or secretion from the cell, the ribosome becomes **bound to the endoplasmic reticulum (ER)**
- Signal sequences in the growing polypeptide chain dictate whether the free ribosome needs **to move to the ER**
  - The signal sequence occurs at the beginning polypeptide
  - Signal recognition proteins bind to the polypeptide, pausing translation
  - The free ribosome binds to a receptor on the ER, **forming rough ER**
  - **Translation is re-initiated** and the polypeptide chain **moves inside the ER**
- The synthesised protein can be **carried via a vesicle** to the **Golgi** apparatus before being **secreted out** of the cell



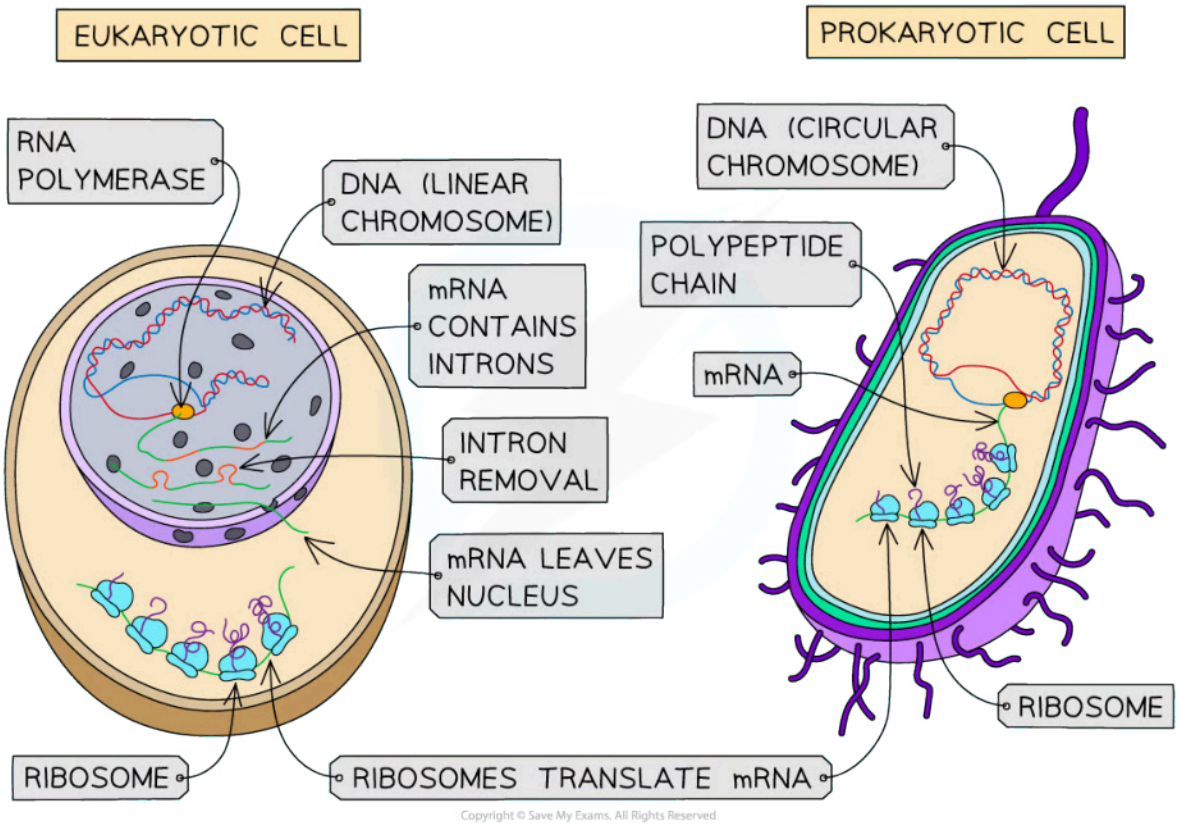
Proteins destined for lysosomes or secretion out of the cell are synthesised by ribosomes bound to the endoplasmic reticulum

7.3.3 Translation in Prokaryotes

Your notes

Translation in Prokaryotes

- Prokaryotic cells have a **less complex ultrastructure** than eukaryotic cells
  - Eukaryote cells are divided up into membrane-bound compartments called organelles
  - Transcription and translation happen in different compartments because ribosomes are separated from the nucleus
- The lack of a nucleus is a **defining cellular feature of prokaryotes**, allowing transcription and translation to take place in the **same compartment**
- Translation can occur **immediately after transcription** due to the absence of a nuclear membrane
  - Both processes **proceed simultaneously** and likely in a coupled fashion
  - Translation starts even **before the mRNA has finished being transcribed** from the DNA



Transcription and translation occur simultaneously in prokaryotes due to the absence of a nuclear membrane. Ribosomes start translating the mRNA whilst it is still being synthesised.

7.3.4 Bioinformatics

Your notes

Bioinformatics

- The 21st Century has seen a tremendous **increase** in the amount of **biological data**
  - This has been due to rapid advances in **DNA sequencing** and other technologies
- Developments in scientific research have been accompanied by **improvements in computing**, enabling scientists to interpret complex biological data using bioinformatics applications
- **Bioinformatics** is an **interdisciplinary field** that develops methods and software to help further our understanding of life by making sense of this data
  - Although many new bioinformatics applications are at the forefront of applied computing, most scientific research uses **standard tools and databases**
- Data related to gene sequence, protein structure, gene expression or metabolites is curated, annotated and stored in databases such as **GenBank, NCBI, EBI, PDB**
- A range of **open source software tools** is available to query this data

Sequence similarity

- If a scientist has an unknown DNA sequence, they can determine if it codes for a gene
- **BLAST** (**B**asic **L**ocal **A**lignment **S**earch **T**ool) search can compare the unknown DNA sequence to **all known gene sequences** in a particular database
  - BLAST **finds regions of similarity** between sequences
  - The search returns 'hits' which are the sequences most related to the search sequence (depending on the parameters set)
- There are **many variations of BLAST** that can be used for different analyses such as **protein sequences** or comparing multiple input sequences at once

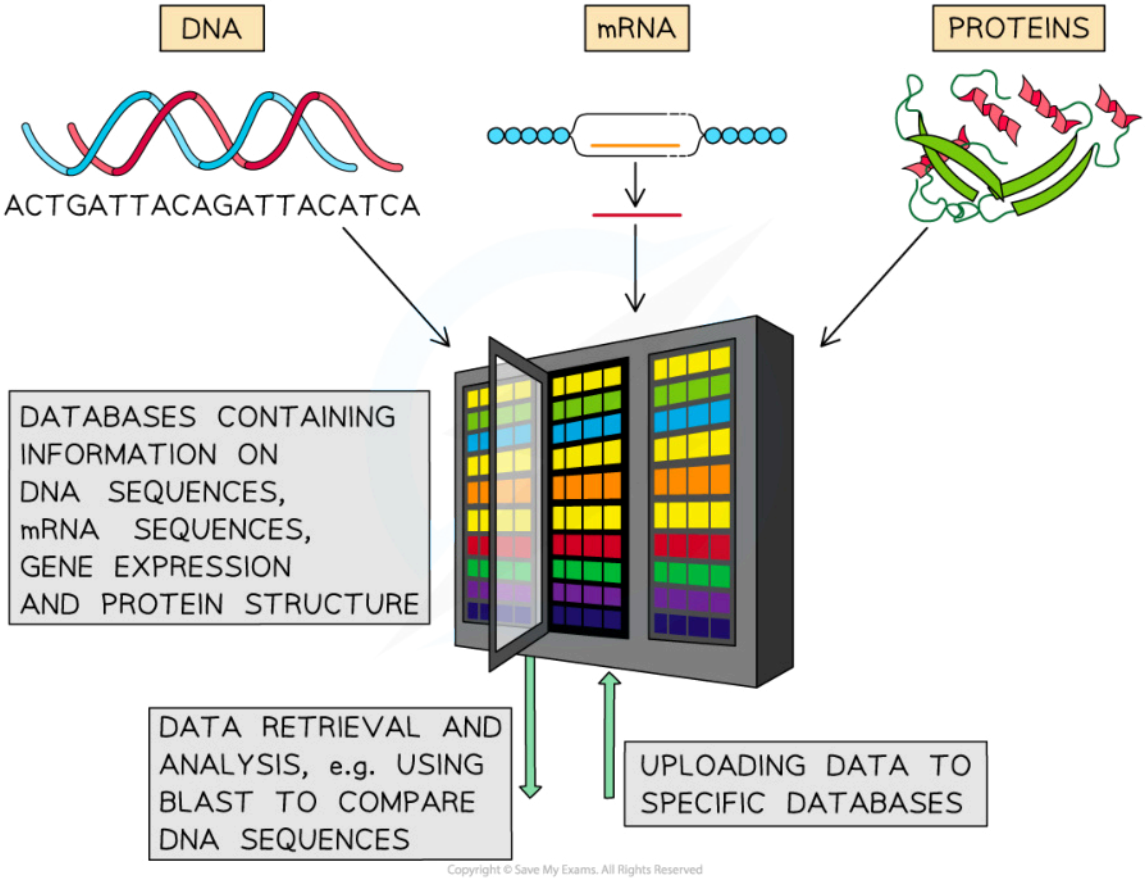
Genetic variation and evolutionary relationships

- Scientists can **compare homologous gene sequences** between many organisms
  - Sequences are compared using an alignment tool such as **Clustal W** (there are many alternatives)
  - This aligns (stacks) the sequences based on similar regions so that **variable regions can be identified**
- This determines the **degree of similarity** between organisms which **gives an indication of how closely related** the organisms are
  - There may be a **common ancestral origin** but in some organisms, the gene might have accumulated differences over times from random mutations
- Tree-like evolutionary diagrams (**phylogenetic trees**) can be constructed with software such as **PhyloWin** to show the **degree of relatedness** to a recent common ancestor
- **Phylogenetic analysis** is useful for biological classification, conservation studies, forensics or molecular epidemiology which can help dictate public health policy
  - **Variants** of highly infectious pathogens such as SARS-CoV-2 (a well-known coronavirus) can be identified using these techniques

Sequencing DNA to determine protein sequences

Your notes

- The **genetic code** can be used to determine the **amino acid sequence** within a protein
  - This primary structure information can be used to predict **how proteins will fold** into their tertiary structure
  - This gives a greater level of understanding of **how a protein functions** or **interacts** with other proteins or molecules
- Such information can be used for a range of applications, such as **drug design** or **novel protein engineering** in synthetic biology



Bioinformatics allows for large amounts of biological data to be available instantly to researchers across the globe

7.3.5 Levels of Protein Structure


Your notes

Primary Structure

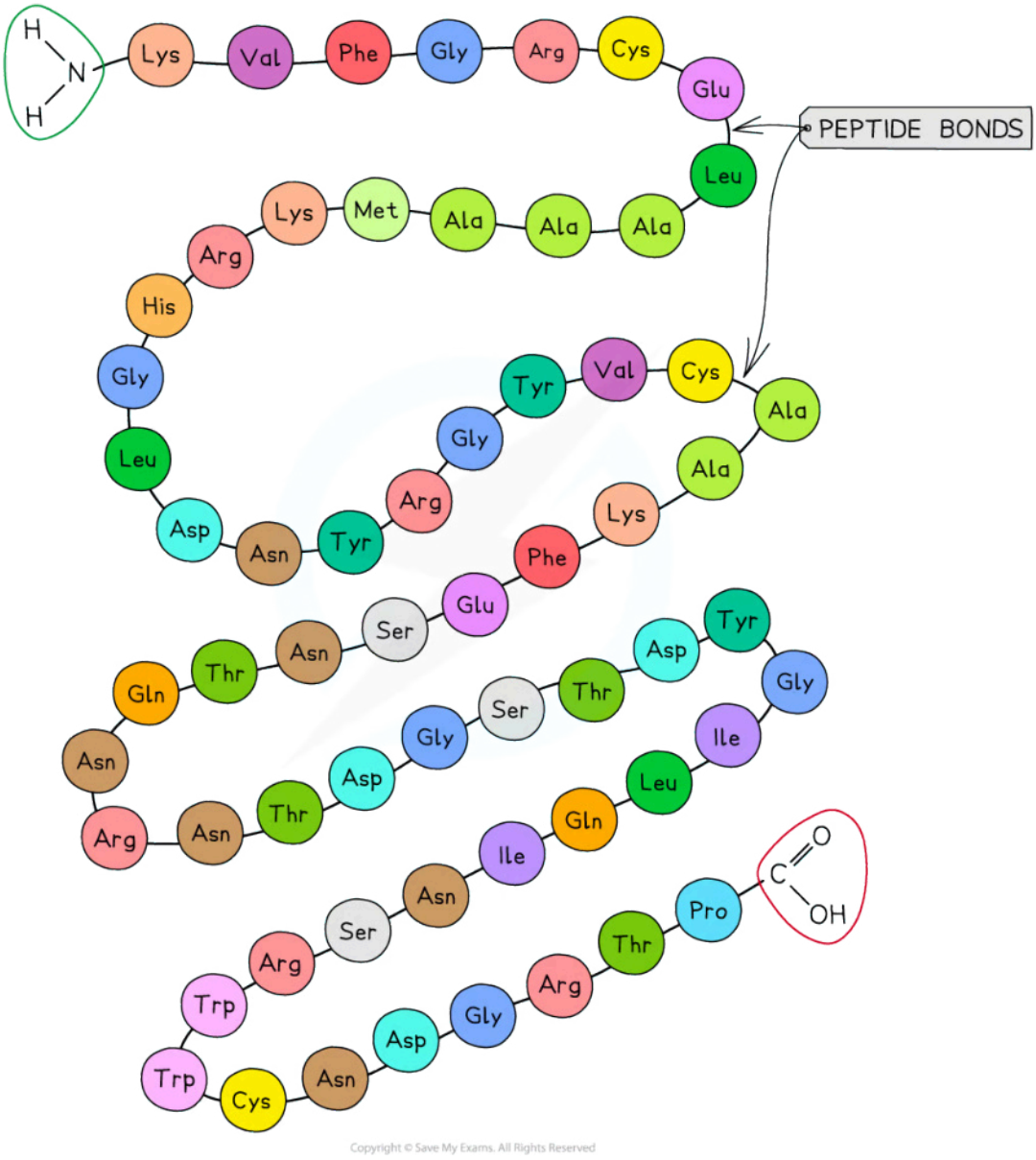
Levels of Protein Structure

- Proteins are relatively **large**, **complex molecules** that contain **one or more chains** of amino acids known as **polypeptides**
- The **three-dimensional arrangement** of polypeptide chains dictates a protein's structure and function
- There are **four levels of structure in proteins**
  - Three levels are structural aspects of a single polypeptide chain
  - The fourth level relates to a protein that has more than one polypeptide chain

Primary structure

- The sequence of amino acids bonded by covalent peptide bonds is the **primary structure** of a protein
- **The DNA** of a cell **determines** the primary structure of a protein by instructing the cell to add certain amino acids in specific quantities in a specific, ordered sequence
- This affects the **shape,** and therefore the **function,** of the protein
- The primary structure is **specific** for each protein
- Some mutations can lead to the incorrect amino acid being incorporated into the polypeptide chain which can affect the function of the protein

Your notes

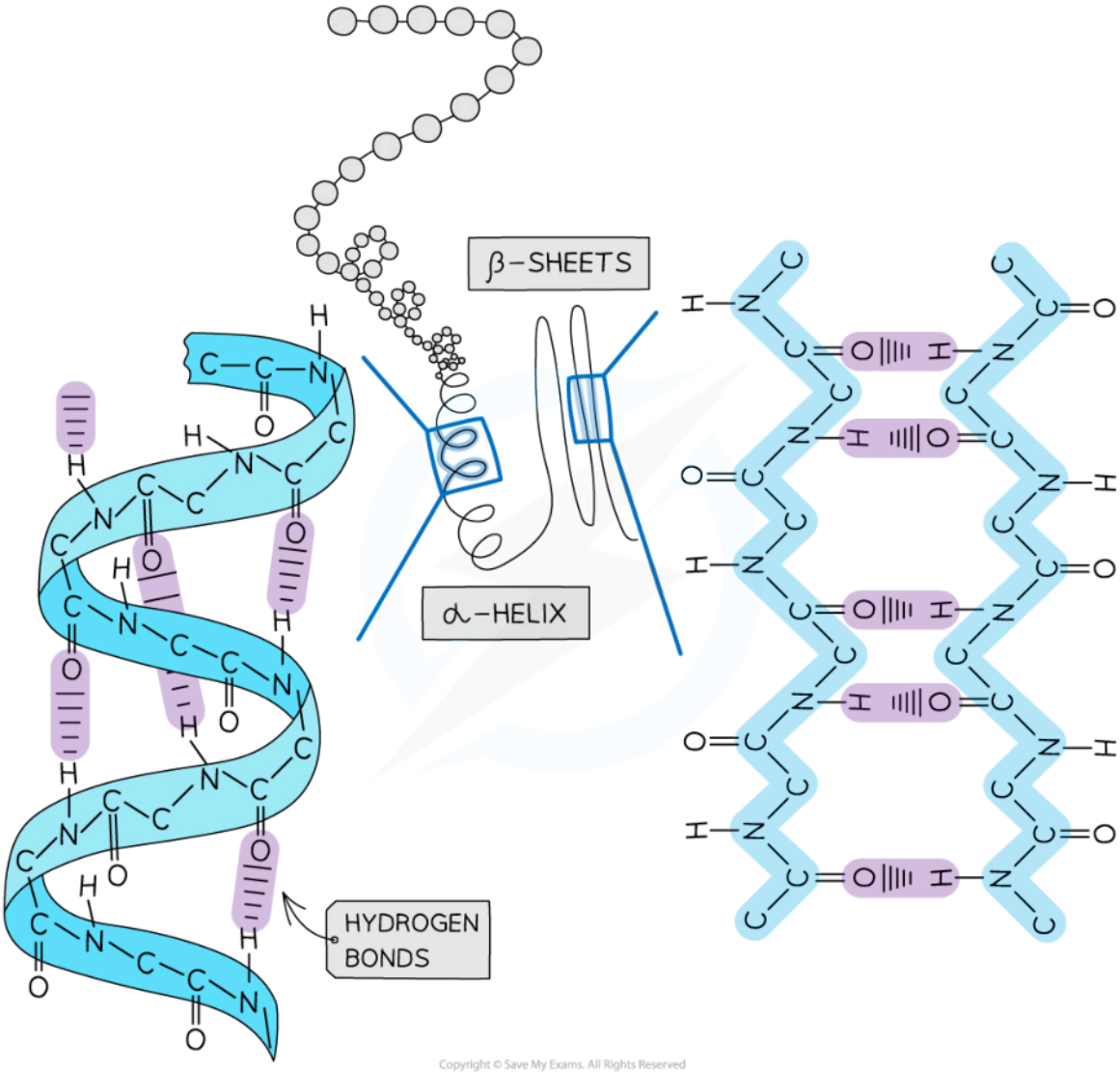


The primary structure of a protein. The three-letter abbreviations indicate the specific amino acid (there are 20 commonly found in cells of living organisms).

Secondary Structure

Your notes

- Secondary structure is the **formation of complex shapes** within the polypeptide chain
- Secondary structure of a protein occurs due to **weak hydrogen bonds**
  - Hydrogen bonds form between **carboxyl** (C=O) groups and **amino** (H-N-H) groups
  - The bonds usually form **between non-adjacent amino acids** resulting in a change in shape of the linear polypeptide chain
- There are **two shapes** that can form within proteins due to the hydrogen bonds:
  - Alpha-helix (or **α-helix**)
  - Beta-pleated sheet (or **β-pleated sheet**)



The secondary structure of a protein with the α-helix and β-pleated sheets.
The magnified regions illustrate how the hydrogen bonds form between peptide bonds.

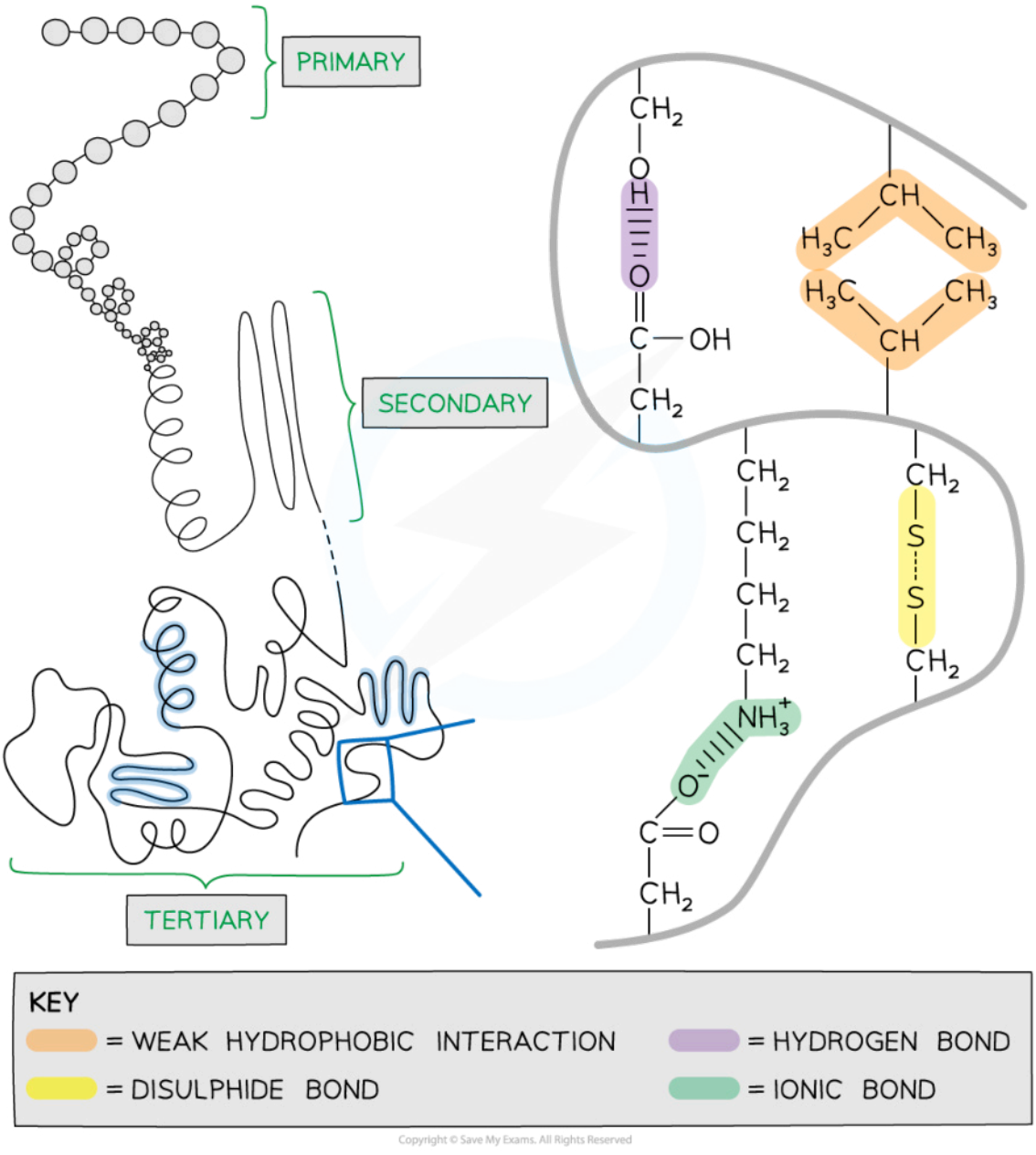
Tertiary Structure

Polar and non-polar amino acids are relevant to the bonds formed between R groups

- Tertiary structure refers to how the polypeptide chain folds to form a complex, three-dimensional shape
- Tertiary structure gives proteins a **very specific shape** that is important for function
  - Such as **receptor sites** on cell membranes and **active sites** in enzymes
- Folding results from **interactions between R groups** (side chains) of the amino acids and the surrounding environment
- A number of different interactions between R-groups contribute to the tertiary structure
  - **Hydrogen bonds** form between polar R-groups
  - **Hydrophobic interactions** form between the R-groups of non-polar amino acids **within the interior of proteins** to avoid contact with water
  - **Covalent bonds** form between the R-groups of cysteine amino acids to form disulphide bridges
  - **Ionic bonds** form between positively and negatively charged R-groups


Your notes

Your notes



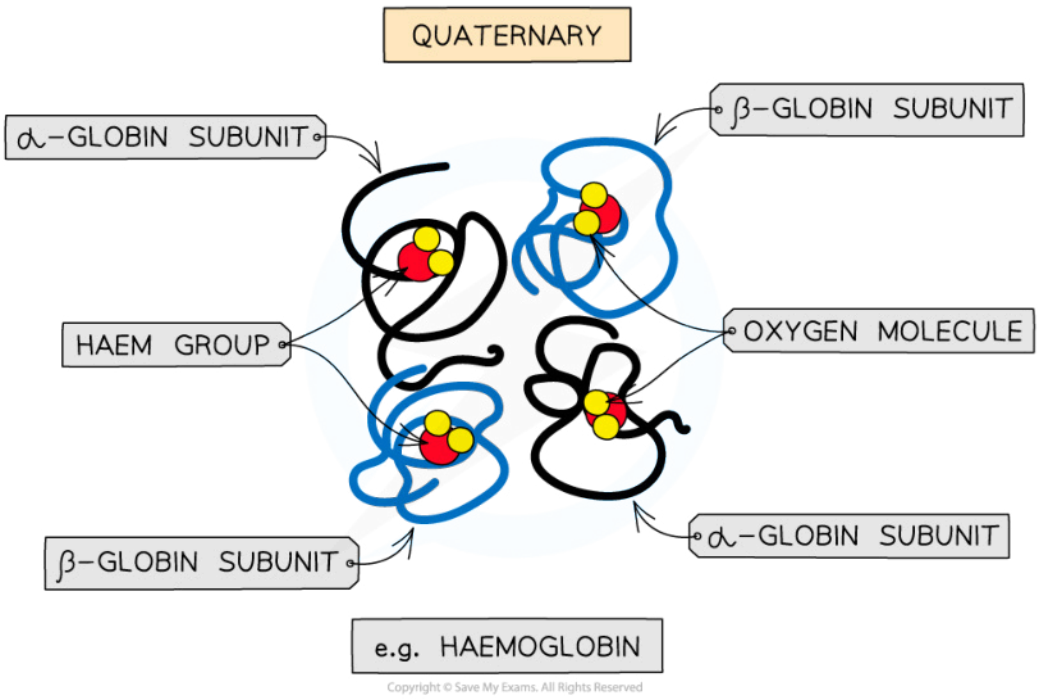
The interactions that occur between the R groups of amino acids determine the tertiary structure and function of a protein

Quaternary Structure

Your notes

Quaternary structure

- Large proteins often consist of **multiple polypeptide chains** functioning together as a larger biologically active macromolecule
  - Each polypeptide chain is referred to as a **subunit** of the protein
- Many proteins also contain non-polypeptide components (**prosthetic groups**) and are classed as **conjugated proteins**
- Quaternary structure refers to how polypeptides and other components are arranged
  - This relates closely to function
  - Proteins with only one polypeptide chain do not have a quaternary structure
- **Haemoglobin** is a conjugated protein, having quaternary structure, as it consists of **multiple polypeptide chains** (making **four subunits**) each with a prosthetic group
  - There are two pairs of identical polypeptide chains (**α-globins** and **β-globins**)
  - Each subunit has a prosthetic **haem** group which contains an **iron** atom (Fe)



The quaternary structure of haemoglobin
Four subunits (polypeptide chains) and prosthetic haem groups work together to carry oxygen

Summary of Bonds in Proteins Table

Your notes

	Level		
Bonds	Primary	Secondary	Tertiary
Peptide	✓	✓	✓
Hydrogen		✓ (only between the amino and carboxyl groups)	✓ (R groups + amino and carboxyl groups)
Disulphide			✓
Ionic			✓
Hydrophobic interactions			✓

Examiner Tip

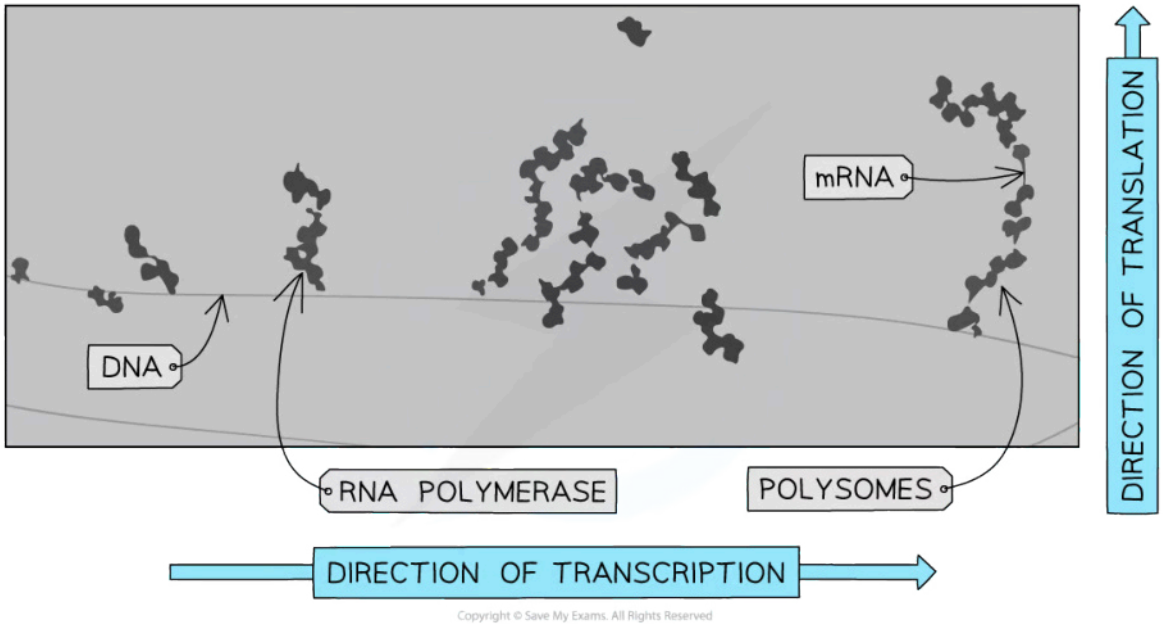
Familiarise yourself with the difference between the four structural levels found in proteins, noting which bonds are found at which level. Remember that the hydrogen bonds in tertiary structures are between the R groups whereas in secondary structures the hydrogen bonds form between the amino and carboxyl groups.

7.3.6 Skills: Polysomes & Ribosomes

Your notes

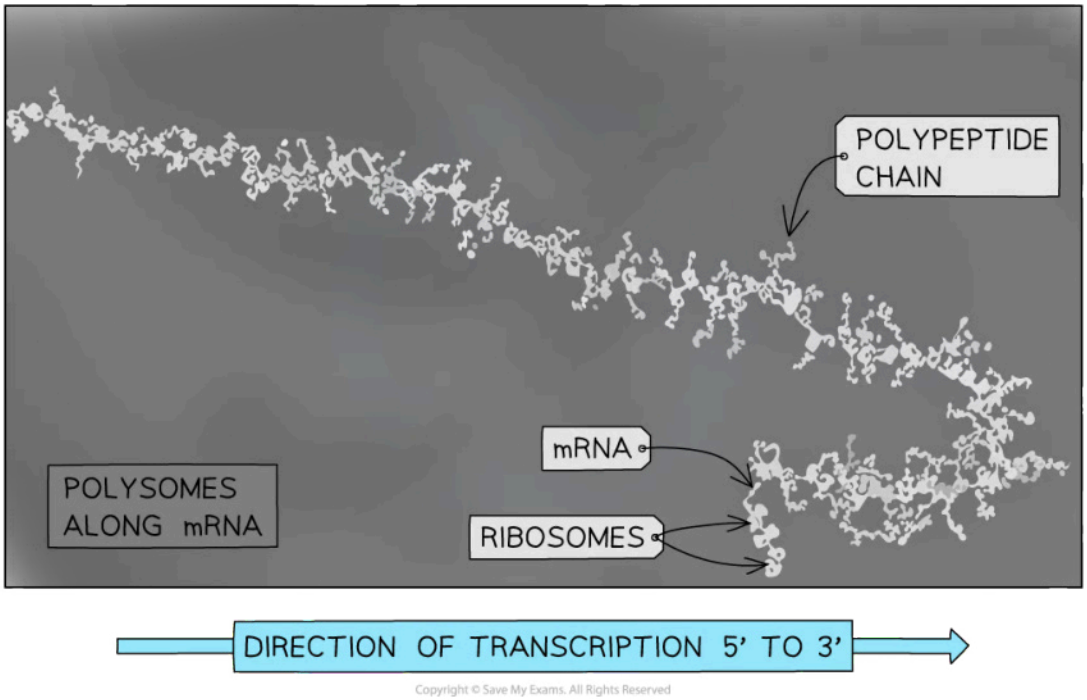
Identifying Polysomes

- Translation can occur **simultaneously** at **multiple positions** along mRNA
- **Polysomes** (or polyribosomes) are groups of two or more ribosomes translating **the same mRNA** transcript
- Multiple copies of the same polypeptide chain can be made simultaneously from a single mRNA transcript
  - Polysomes effectively **increase the amount of polypeptide produced**
- In electron micrographs, polysomes look like 'beads on a string' with each bead representing a ribosome
  - There are **visible differences** between eukaryotes and prokaryotes
- In prokaryotes, the lack of a nucleus means transcription and translation are coupled
  - Translation starts **before the mRNA has finished being transcribed** from the DNA
  - On an electron micrograph, **multiple polysomes** can appear **on growing mRNA strands** along the DNA molecule
- In eukaryotes, mRNA is transported **out of the nucleus** prior to translation
  - On an electron micrograph, **polysomes are seen on the mRNA** with no involvement of DNA
- As ribosomes move in the same 5' to 3' direction along the mRNA, ribosomes towards the 3' end have longer polypeptide chains being synthesised



Electron micrograph of prokaryotic polysomes, the image shows simultaneous transcription and translation of a bacterial gene

Your notes



Electron micrograph of eukaryotic polysomes, polypeptide chains can be seen emerging from the ribosomes

Visualising the Structure of Ribosomes


Your notes

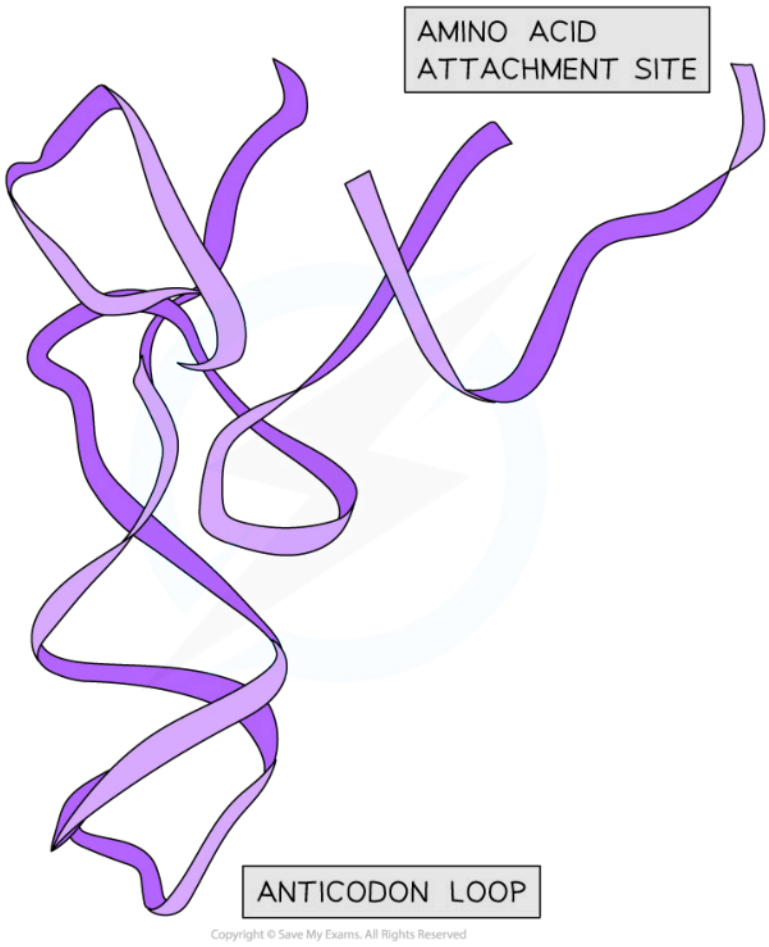
- There are many 'open source' **databases** that contain data relating to the **three-dimensional structure of proteins**
  - The most commonly used ones are **PDB** (Protein Data Bank) and **Swiss-Prot**
- Such databases allow researchers to **analyse biological molecules** and **study interactions** between them
  - This helps relate **structure** to biological **function**
- Data relating to the three-dimensional structure of biological molecules can be visualised using molecular visualisation software such as **Mol*** or **Jmol**
- Molecules can be represented in many different ways including **ball and stick** atom models or **simplified ribbon** representations that show the protein backbone
- Most molecular visualisation software is **freely available** on the Internet or can be accessed through many bioinformatics repositories

Analysing the structure of the eukaryotic ribosome

- Visit the **PDB** and search for: **Yeast 80S ribosome 4V7R** (do not put the search term in quotes)
- Select the "**3D view**" to view the protein structure in **mol***
- Eukaryotic ribosome are complex molecules consisting predominantly of ribosomal RNAs and multiple proteins - these are represented in **different colours**
- The molecule appears to be made of two distinct halves (**subunits**)
- **Rotate** or **zoom** into the image to visualise the different components
- Try and identify the **ribosomal RNA** in the two subunits (hover the cursor over)
- An opening (**groove**) between the subunits should be visible - mRNA passes through here

Analysing the structure of a tRNA molecule

- Search for: **1YFG** in the **PDB** (1YFG is the database identifier for yeast initiator tRNA)
- Select the "**3D view**" to view the protein structure in **mol***
- You should be able to recognise the **loop structures**
- Consider how this structure is **specific** to the **tRNA-activating enzyme**
- Try and identify the **anticodon region** and amino acid binding site (**acceptor arm**)
- Try selecting a different viewer such as **JSmol**
- Investigate changing settings in the viewer



Structure of yeast tRNA

Your notes