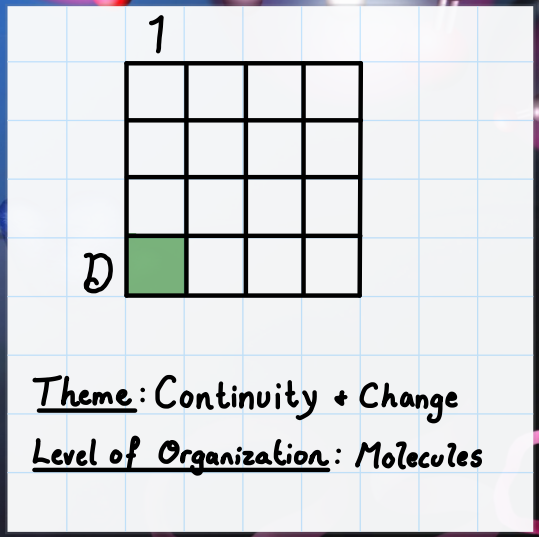
D1.2 Protein Synthesis

Guiding Questions
How does a cell produce a sequence of amino acids from a sequence of DNA bases?
How is the reliability of protein synthesis ensured?

Linking Questions
How does the diversity of proteins produced contribute to the functioning of a cell?
What biological processes depend on hydrogen bonding?



Theme: Continuity + Change

Level of Organization: Molecules

Written and drawn by:

Peter Marier

NOTEWORTHY SCIENCE

MARIER

SL Learning Outcomes

D1.2.1	Transcription as the synthesis of RNA using a DNA template	Students should understand the roles of RNA polymerase in this process.
D1.2.2	Role of hydrogen bonding and complementary base pairing in transcription	Include the pairing of adenine (A) on the DNA template strand with uracil (U) on the RNA strand.
D1.2.3	Stability of DNA templates	Single DNA strands can be used as a template for transcribing a base sequence, without the DNA base sequence changing. In somatic cells that do not divide, such sequences must be conserved throughout the life of a cell.
D1.2.4	Transcription as a process required for the expression of genes	Limit to understanding that not all genes in a cell are expressed at any given time and that transcription, being the first stage of gene expression, is a key stage at which expression of a gene can be switched on and off.
D1.2.5	Translation as the synthesis of polypeptides from mRNA	The base sequence of mRNA is translated into the amino acid sequence of a polypeptide.
D1.2.6	Roles of mRNA, ribosomes and tRNA in translation	Students should know that mRNA binds to the small subunit of the ribosome and that two tRNAs can bind simultaneously to the large subunit.
D1.2.7	Complementary base pairing between tRNA and mRNA	Include the terms “codon” and “anticodon”.
D1.2.8	Features of the genetic code	Students should understand the reasons for a triplet code. Students should use and understand the terms “degeneracy” and “universality”.
D1.2.9	Using the genetic code expressed as a table of mRNA codons	Students should be able to deduce the sequence of amino acids coded by an mRNA strand.
D1.2.10	Stepwise movement of the ribosome along mRNA and linkage of amino acids by peptide bonding to the growing polypeptide chain	Focus on elongation of the polypeptide, rather than on initiation and termination.
D1.2.11	Mutations that change protein structure	Include an example of a point mutation affecting protein structure.

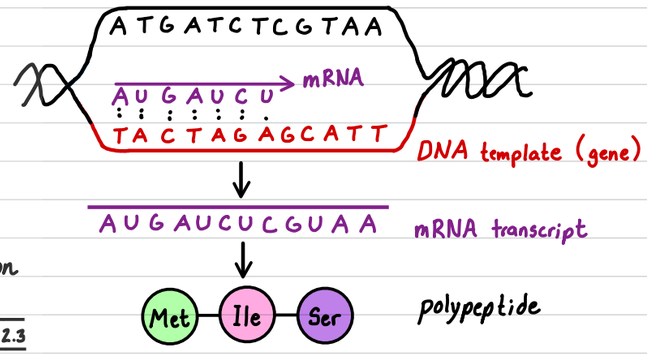
HL Learning Outcomes

D1.2.12	Directionality of transcription and translation	Students should understand what is meant by 5' to 3' transcription and 5' to 3' translation.
D1.2.13	Initiation of transcription at the promoter	Consider transcription factors that bind to the promoter as an example. However, students are not required to name the transcription factors.
D1.2.14	Non-coding sequences in DNA do not code for polypeptides	Limit examples to regulators of gene expression, introns, telomeres and genes for rRNAs and tRNAs in eukaryotes.
D1.2.15	Post-transcriptional modification in eukaryotic cells	Include removal of introns and splicing together of exons to form mature mRNA and also the addition of 5' caps and 3' polyA tails to stabilize mRNA transcripts.
D1.2.16	Alternative splicing of exons to produce variants of a protein from a single gene	Students are only expected to understand that splicing together different combinations of exons allows one gene to code for different polypeptides. Specific examples are not required.
D1.2.17	Initiation of translation	Include attachment of the small ribosome subunit to the 5' terminal of mRNA, movement to the start codon, the initiator tRNA and another tRNA, and attachment of the large subunit. Students should understand the roles of the three binding sites for tRNA on the ribosome (A, P and E) during elongation.
D1.2.18	Modification of polypeptides into their functional state	Students should appreciate that many polypeptides must be modified before they can function. The examples chosen should include the two-stage modification of pre-proinsulin to insulin.
D1.2.19	Recycling of amino acids by proteasomes	Limit to the understanding that sustaining a functional proteome requires constant protein breakdown and synthesis.

D1.2.1—Transcription as the synthesis of RNA using a DNA template.
D1.2.2—Role of hydrogen bonding and complementary base pairing in transcription.
D1.2.4—Transcription as a process required for the expression of genes

Central Dogma of Molecular Biology: genetic information generally flows in one direction: DNA → RNA → proteins

↳ Protein synthesis involves two processes: transcription + translation which together result in gene expression



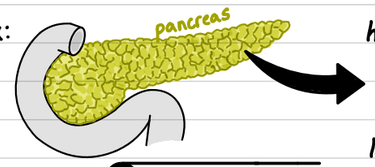
Gene expression: the process by which the information in DNA (genes) are converted into a functional product / observable effect like proteins

↳ all cells in an organism have the same **genome** - set of all genetic material, but they don't transcribe all genes and perform different functions, thus vary in gene expression

↳ gene expression can be controlled by factors both inside and outside cells resulting in genes being **switched ON** or **switched OFF**

cell specialization B2.3

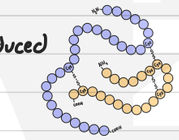
ex:



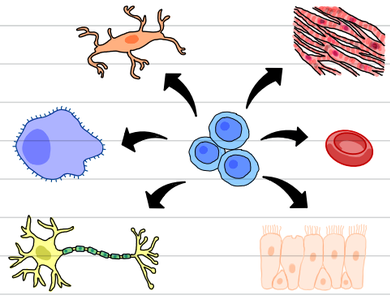
homeostasis D3.3

high blood glucose → β-cell
- insulin gene: ON → transcription → insulin produced
- glucagon gene: OFF → no transcription

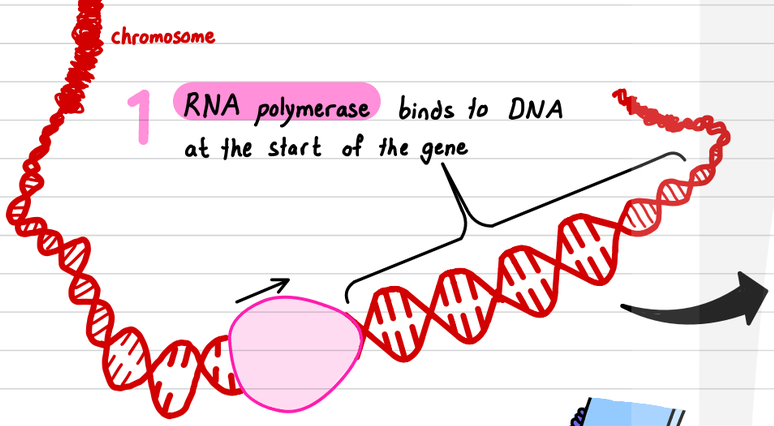
low blood glucose → α-cell
- insulin gene: OFF → no transcription
- glucagon gene: ON → transcription → glucagon produced

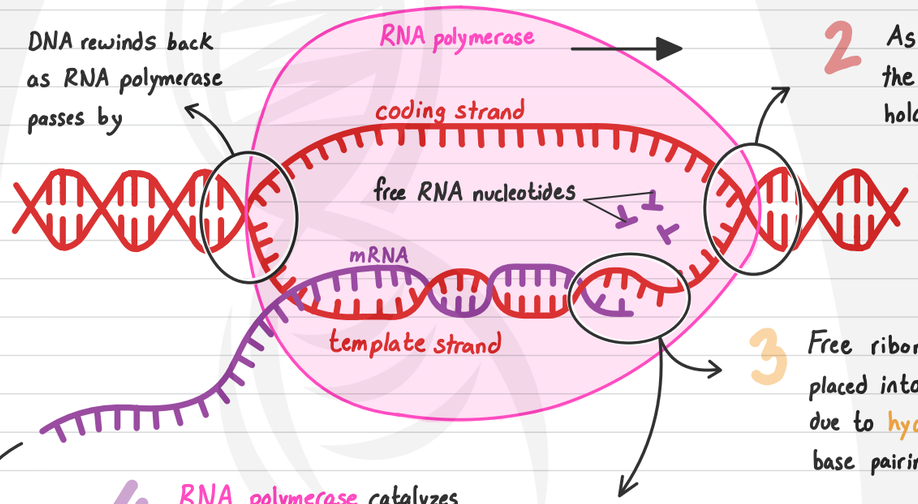


※ **Stem cells** are undifferentiated and have the ability to specialize into any cell type by switching on and off certain genes during development. Specialized cells are those whose role is fixed and only express certain genes to produce specific products related to their function.



Transcription: synthesis of complementary mRNA (messenger RNA) using a DNA template base sequence (gene)



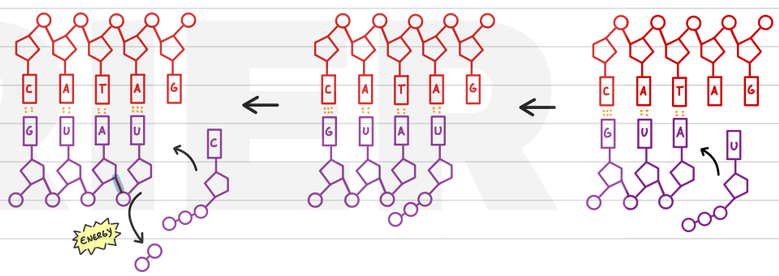


2 As RNA polymerase moves down gene, it separates the DNA double helix by breaking hydrogen bonds holding complementary base pairs together

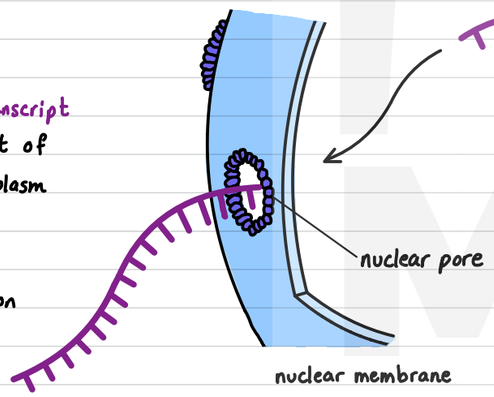
↳ this exposes the template strand which is used for transcription (coding strand is not used)

3 Free ribonucleoside triphosphates are correctly placed into position across the template strand due to hydrogen bonding and complementary base pairing: Adenine (A) ⋯ Uracil (U), Cytosine (C) ⋯ Guanine (G)

4 RNA polymerase catalyzes condensation reactions: catalyzing a covalent phosphodiester bond between the phosphate of a free RNA nucleotide to the ribose of growing mRNA strand. Once it reaches the end of the gene, it unbinds



5 In eukaryotes, mRNA transcript is modified and sent out of the nucleus into the cytoplasm via the nuclear pore to start translation. In prokaryotes, translation occurs right away



coding / sense / non-template strand
strand whose base sequence is identical to the mRNA transcript (with T and U swapped)

AATCGGTTCAGG
AAUCGGUUCAGG

template / antisense / non-coding strand
strand that is used to make an mRNA transcript ∴ is complementary to both the mRNA and coding strand

D1.2.5—Translation as the synthesis of polypeptides from mRNA. D1.2.6—Roles of mRNA, ribosomes and tRNA in translation. D1.2.7—Complementary base pairing between tRNA and mRNA. D1.2.10—Stepwise movement of the ribosome along mRNA and linkage of amino acids by peptide bonding to the growing polypeptide chain

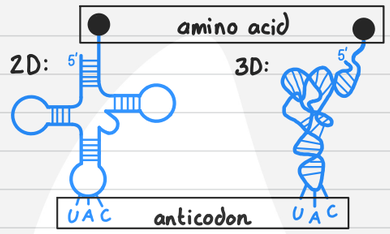
Translation : synthesis of polypeptides (proteins) from mRNA transcript – the base sequence of mRNA (codons) are translated into the amino acid sequence of the polypeptide

↳ **messenger RNA (mRNA)**

provides instructions to synthesize a polypeptide in the form of codons (3 bases)
has a binding site for the ribosome and a start and stop codon to direct translation

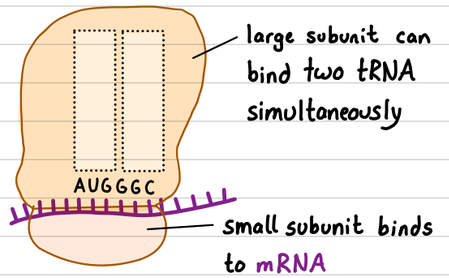
↳ **transfer RNA (tRNA)**

carries amino acid and binds complementarily to mRNA codon using its anticodon. Enzymes ensure each tRNA is bound to a specific amino acid according to its anticodon



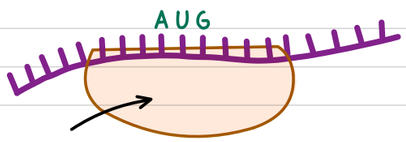
↳ **ribosome**

made of small and large subunit
'reads' mRNA one codon at a time and catalyzes peptide bond between adjacent amino acids.
Made of rRNA and proteins



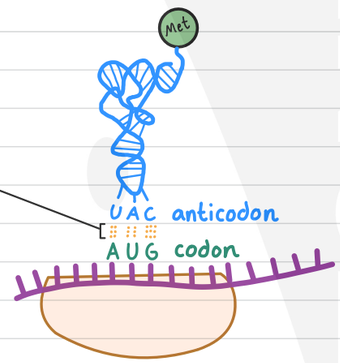
Initiation (1,2,3)

1 small subunit of ribosome binds to mRNA and moves along until START codon is reached

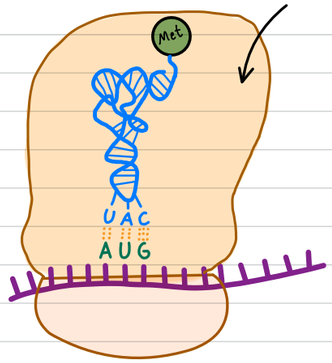


2 tRNA with complementary anticodon to start codon binds to mRNA forming hydrogen bonds
this brings the first amino acid

※ complementary base pairing between tRNA anticodon and mRNA codon ensures the correct amino acid is brought and used



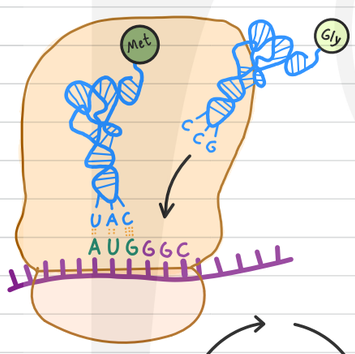
3 Large subunit of ribosome binds, sandwiching the mRNA in between

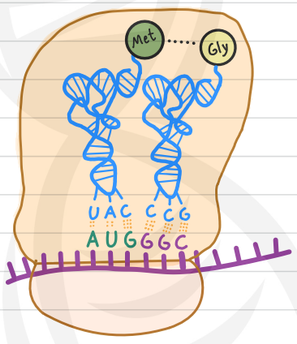


Elongation (4,5,6)

4 A tRNA with a complementary anticodon to the next codon enters the large ribosome subunit and binds

※ Many different tRNAs are present but only those with a complementary anticodon can bind, others cannot

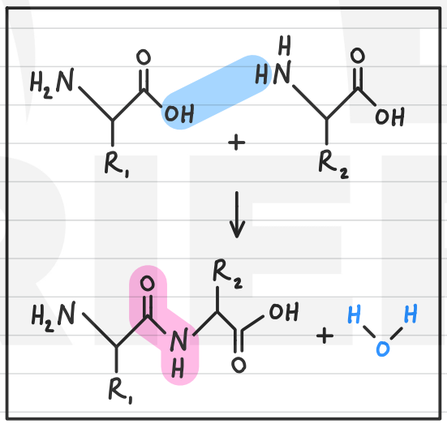




elongation continues until STOP codon

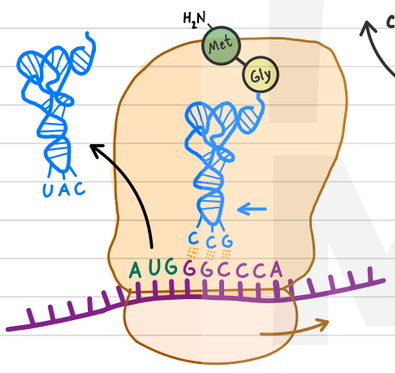
5 rRNA within large subunit catalyzes the formation of a covalent peptide bond (condensation reaction) between the amino acid of previous tRNA and the amino acid on the new tRNA

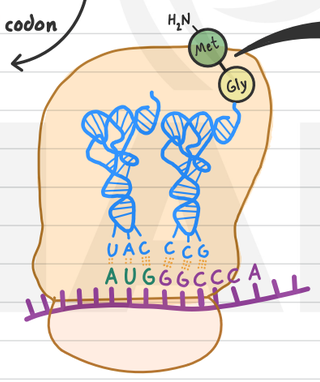
※ this moves growing polypeptide from previous to the new tRNA



6 Ribosome moves down mRNA by one codon, tRNA with polypeptide moves over and empty tRNA is released, ready to be reused and bound to new amino acid

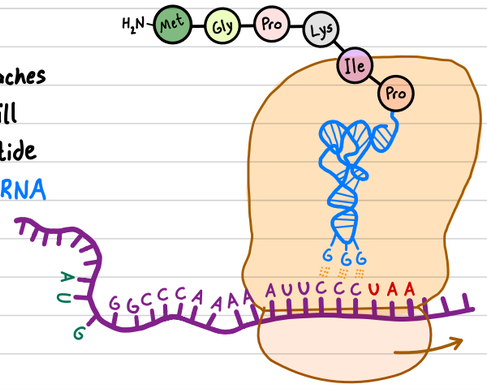
※ each cycle, ribosome moves down mRNA by one codon and polypeptide grows by one amino acid



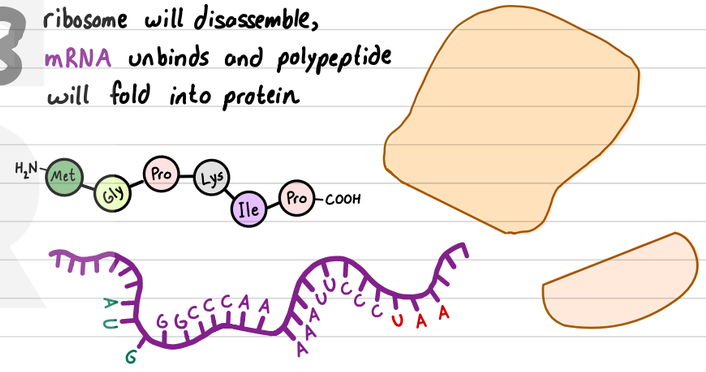


Termination (7,8)

7 Once ribosome reaches STOP codon it will cause the polypeptide to unbind from tRNA



8 ribosome will disassemble, mRNA unbinds and polypeptide will fold into protein



D1.2.3—Stability of DNA templates. D1.2.8—Features of the genetic code. D1.2.9—Using the genetic code expressed as a table of mRNA codons D1.2.11—Mutations that change protein structure

genetic code: the base sequence of nucleotides which provide information to make a protein – the order determines amino acid sequence

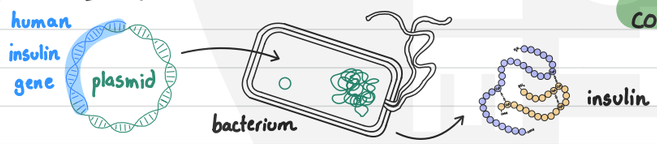
↳ a 3-base code is necessary as a single and two base code could only provide enough combinations for 4 (4^1) and 16 (4^2) amino acids respectively
∴ a triplet code provides more than enough combinations ($4^3 = 64$ codon combinations) for all 20 amino acids and STOP signals as well

↳ the triplet code is **universal**: all organisms (and viruses) use the same code to synthesize proteins, with only very few exceptions

nucleic acids A1.2

- Evidenced by transgenics: when a gene from one species is introduced into another allowing the organism to transcribe and translate protein of interest

ex:



- Shows connectivity between all organisms and that all life shares a last universal common ancestor (LUCA)

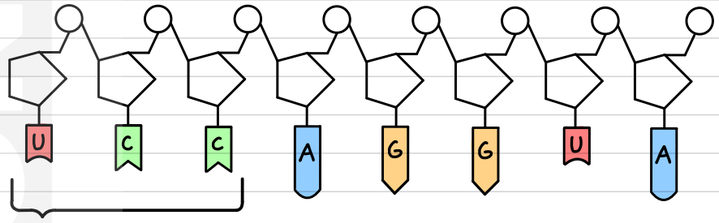
origin of cells A2.1

↳ the triplet code is **degenerate**: most amino acids are coded for by more than one codon ex: glycine is coded by 4 codons: GGA, GGC, GGG, GGU

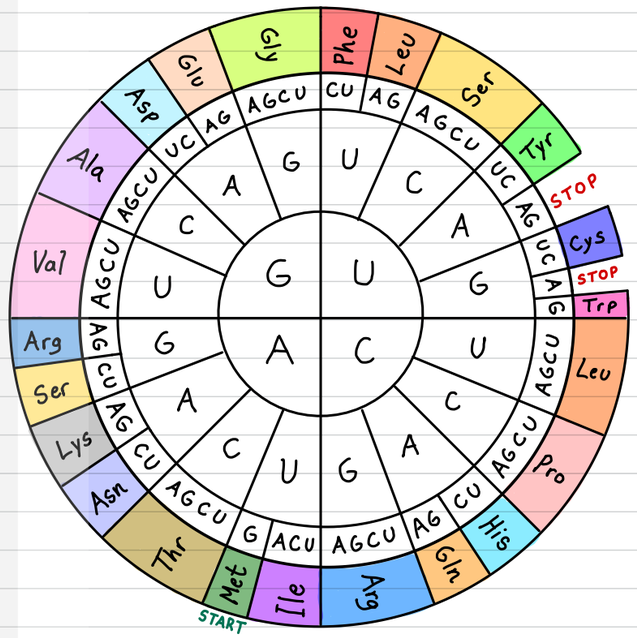
- having multiple codons makes allowances for possibility of mutation as if a change to codon occurs, it may not alter which amino acid is coded for and ∴ the gene's product will be unaltered
  – this is called a silent mutation as even though base sequence is altered, there is no change in phenotype (observable features)

ex:

original template	···ATA··	mutated template ···ATG··
mRNA	···UAU···	mRNA ···UAC···
polypeptide	···–Tyr–···	polypeptide ···–Tyr–···



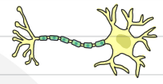
codon: 3 nucleotide base sequence which codes for an amino acid

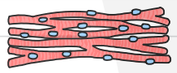


DNA is a very stable molecule due to the covalently-linked phosphate-sugar backbone and many hydrogen bonds holding strands together. When transcription occurs, DNA strands are separated (making them briefly more vulnerable to mutation) but generally DNA template base sequences don't change, allowing the same mRNA to be transcribed

↳ the stability of the DNA template is crucial to ensure the same correct protein is always translated. This is especially important for non-dividing somatic (body) cells as they will transcribe the same template throughout the lifetime of an individual and as more of these cells will not be produced, any mutations would accumulate and the proteins may be altered or even non-functional

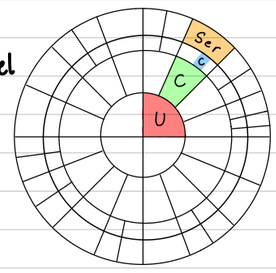
ex: neurons (nerve) and cardiac (heart) muscle cells





do not divide and cannot replace themselves

how to read: inner to outer wheel



codon: first base, second base, third base, amino acid

ex: UCC → Ser

point mutation: a genetic mutation where a single nucleotide base is changed, inserted, or deleted

D1.3 Mutation

↳ base substitution mutations swap one base for another and this alteration to the gene can affect the translated protein structure and functionality

ex: Haemoglobin is a protein composed of 4 polypeptides (two β + two α) and is responsible for carrying O_2 in red blood cells

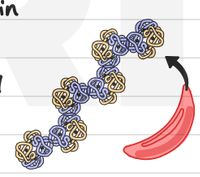
B1.2 Proteins HL

In the normal gene, the β chains produced join with the α chains to form a globular shape

···CTC···
···GAG···
···GAG···
··–Glu–···

A point mutation in the β chain gene causes the mRNA codon transcribed to change, resulting in the hydrophobic amino acid Valine to be translated in the chain rather than hydrophilic Glutamic acid, altering the protein's shape

↳ results in haemoglobin forming filaments and changing the overall shape of the red blood cell into a sickle-shape. This condition (sickle-cell anemia) can lead to blood clots and reduced O_2 transport



···CAC···
···GTG···
···GUG···
··–Val–···

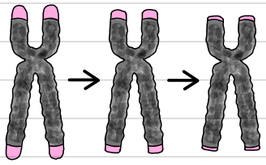
Example translations

template	TTAGCCAAGTCC	coding	CGTGGCATTTGA
mRNA	AAUCGGUUCAGG	mRNA	CGUGGCAUUUGA
polypeptide	Asn–Arg–Phe–Arg	polypeptide	Arg–Gly–Ile

HL D1.2.12—Directionality of transcription and translation. D1.2.13—Initiation of transcription at the promoter. D1.2.14—Non-coding sequences in DNA do not code for polypeptides. D1.2.15—Post-transcriptional modification in eukaryotic cells. D1.2.16—Alternative splicing of exons to produce variants of a protein from a single gene

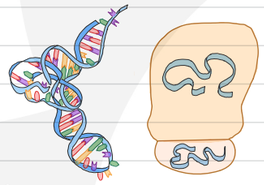
Not all DNA sequences are genes. In eukaryotes the majority of DNA is **non-coding** (does not code for polypeptides). At one point this was termed 'junK DNA' but in actuality they serve many functions:

↳ **telomeres**: region of repetitive nucleotide sequences at the ends of linear chromosomes

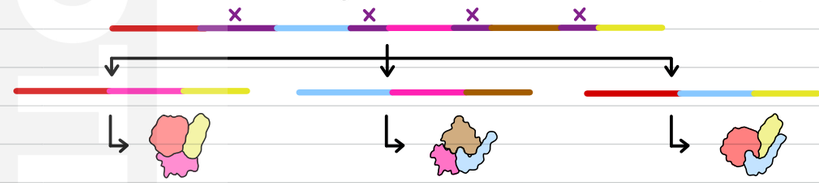


after each round of DNA replication DNA is shortened (due to lagging strand not being fully copied). Telomeres act as a cap and barrier to prevent genetic deterioration and once they are gone, the cell no longer divides and can undergo programmed cell death. This is a deterrent to cancer and the reason for aging

↳ **RNA-coding genes**: sequences of DNA are used to transcribe transfer RNA (tRNA) and ribosomal RNA (rRNA) which are crucial for protein synthesis

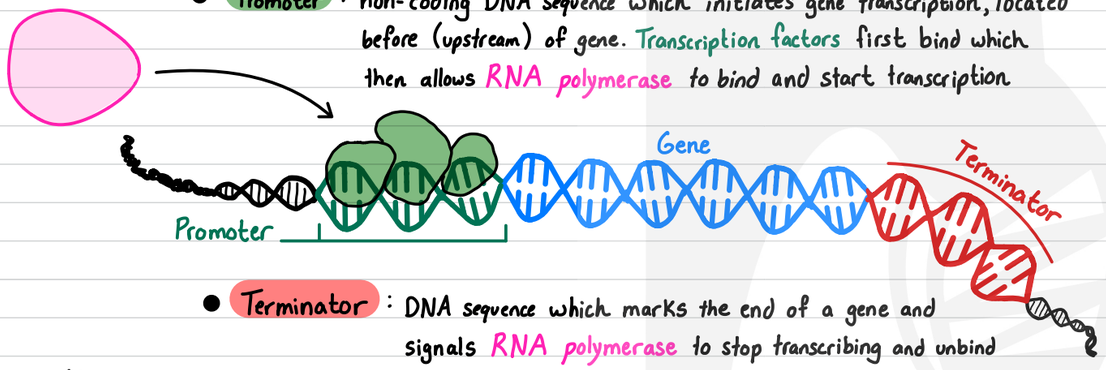


↳ **Introns**: non-coding sequences within genes. Introns are between exons (coding sections) and after transcription (before translation) introns are excised out via alternative splicing to make many different mRNA and proteins



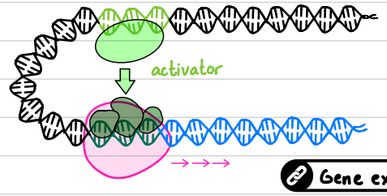
↳ **Regulators of gene expression**: specific non-coding DNA sequences which initiate, promote or suppress gene expression, controlling relative gene activity via the binding of transcription factors (regulatory proteins)

- **Promoter**: non-coding DNA sequence which initiates gene transcription, located before (upstream) of gene. Transcription factors first bind which then allows RNA polymerase to bind and start transcription

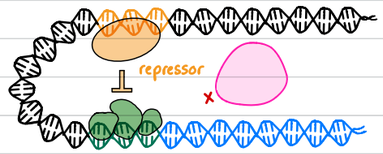


- **Terminator**: DNA sequence which marks the end of a gene and signals RNA polymerase to stop transcribing and unbind

- **Enhancer**: non-coding DNA sequence which promotes gene transcription when bound by an activator protein (transcription factor)
  - act as a ON switch for genes, increasing expression



- **Silencer**: non-coding DNA sequence which block or reduce gene transcription when bound by a repressor protein (transcription factor)
  - act as an OFF switch for genes, decreasing expression

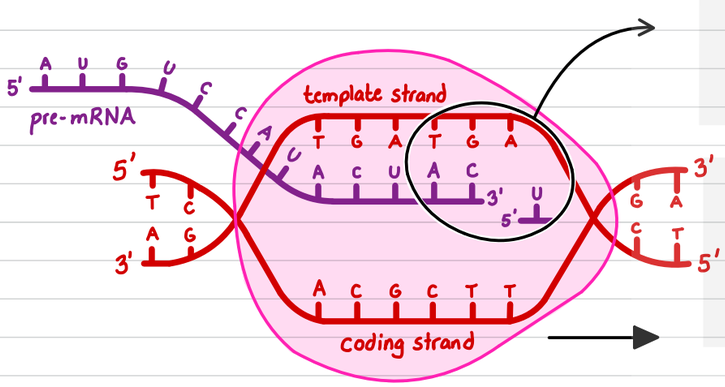


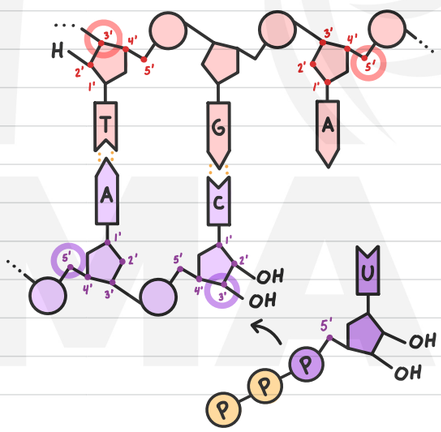
Gene expression D2 2

Transcription

RNA polymerase catalyzes a condensation reaction between the phosphate of the 5' end of a free ribonucleotide triphosphate to the OH on the 3' end of a growing mRNA, forming a phosphodiester bond

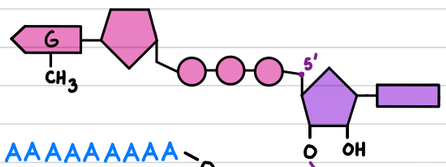
- moves along template/antisense DNA strand in a 3' to 5' direction until terminator is reached
- synthesizes mRNA in a 5' to 3' direction (adding to 3' end of ribose only)



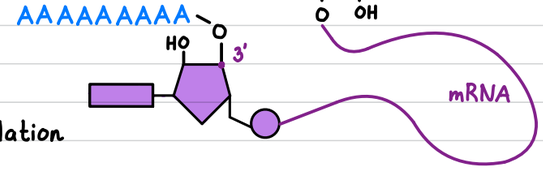


Post-transcriptional modification – In eukaryotes the RNA synthesized (pre-mRNA) is modified before translation

1. **Addition of 5' cap**: a methylated-guanosine 'cap' is added to the 5' end of the RNA

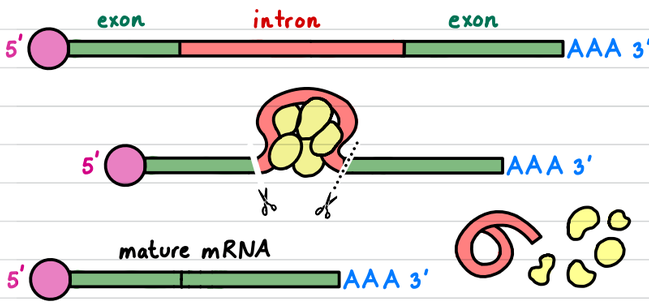


2. **Polyadenation**: a chain of hundreds of adenine nucleotides 'poly A tail' is added to the 3' end of the RNA



※ 5' cap and 3' poly A tail: ✓aid in nuclear export ✓assist in translation ✓improve stability and prevent degradation from nuclease enzymes

3. **Splicing**: Introns (non-coding sequences) within RNA are removed using a complex of small nuclear ribonucleoproteins (snRNPs) called a spliceosome. As the intron loops, exons are brought closer together and are joined (the intron is excised and snRNPs detach)



※ mature mRNA consists of 5' cap, exons, and polyA tail

HL

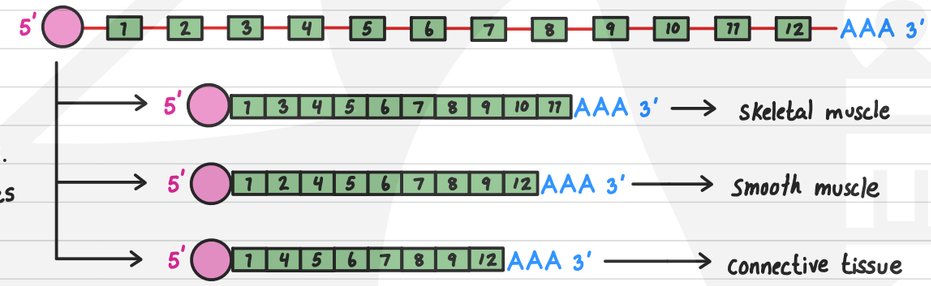
D1.2.16—Alternative splicing of exons to produce variants of a protein from a single gene
D1.2.17—Initiation of translation. D1.2.18—Modification of polypeptides into their functional state
D1.2.19—Recycling of amino acids by proteasomes

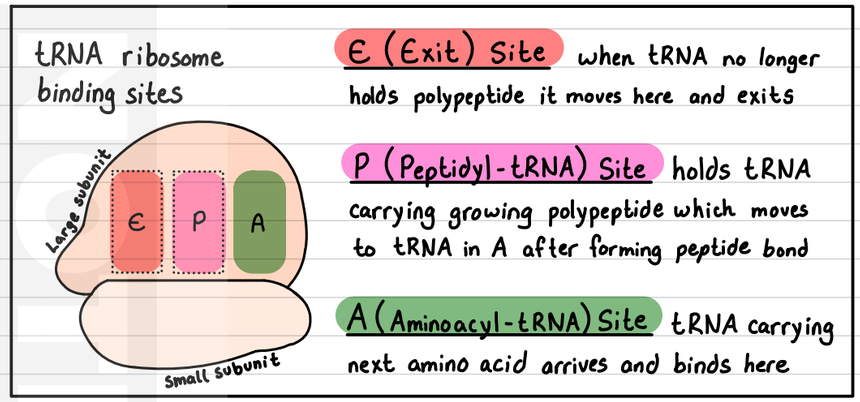
Alternative splicing : splicing exons together in different ways in order to produce different protein variants from a single gene

↳ after introns have been excised, the remaining exons can be rejoined in several different ways either by splicing them in different order or excluding some

↳ allows a single gene to result in many different mature mRNA transcripts and produce a variety of polypeptides.
∴ Number of different proteins made >>> number of genes

ex: α-tropomyosin gene is spliced alternatively to make proteins in many different tissues



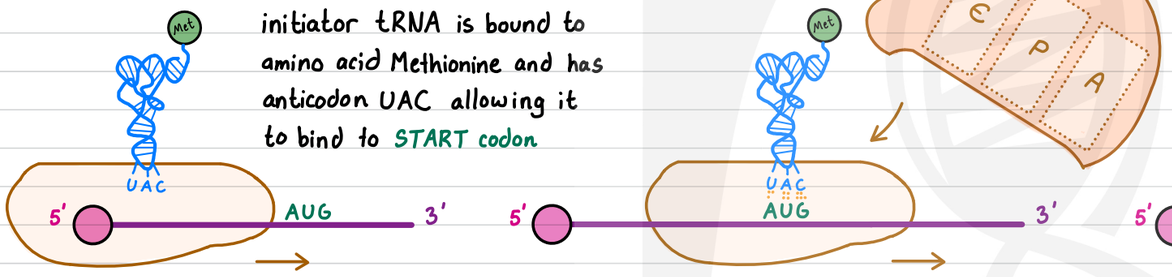


Initiation of Translation

1 An **initiator tRNA** binds to a small ribosome subunit and to the 5' cap of a mature mRNA forming complex. complex moves downstream 5' to 3' scanning for START codon – AUG

2 Once the START codon is reached and the initiator tRNA hydrogen binds, this signals the binding of the large ribosome subunit

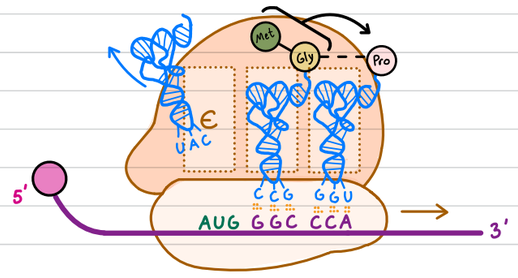
3 The large ribosome subunit binds to the initiator tRNA at its P site, completing the initiation complex. Its A site is open for the next tRNA Elongation begins



initiator tRNA is bound to amino acid Methionine and has anticodon UAC allowing it to bind to START codon

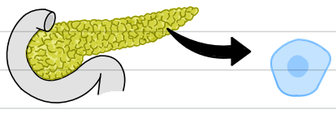
Elongation of Translation

1. ribosome moves 1 codon downstream 5' to 3'
2. tRNA with complementary anticodon binds at A site
3. polypeptide in P forms peptide bond with amino acid in A and chain transferred over to tRNA in A
4. empty tRNA in P moves to E and exits
   tRNA with growing polypeptide in A moves to P

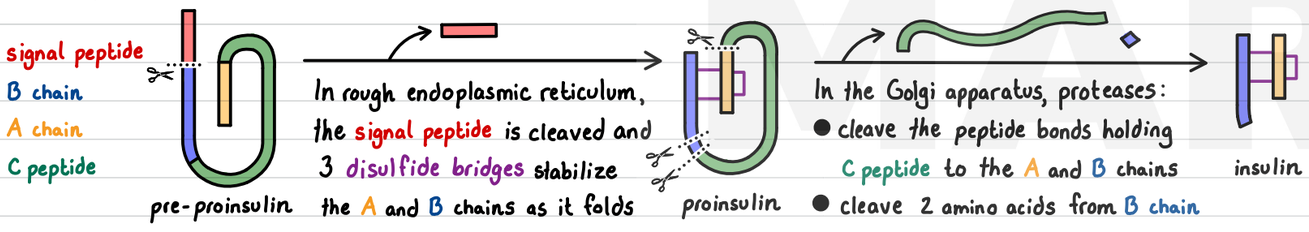


Post-translational modification - many polypeptides are modified after translation to become functional

↳ many possible modifications:
- alter R group composition of amino acids in the chain
- excising sections of the polypeptide by breaking peptide bonds
- forming disulfide bridges between cysteine amino acids R groups
- combine polypeptides together and/or to prosthetic groups



Example: In pancreatic β cells the insulin gene is expressed, resulting in pre-proinsulin which is then modified

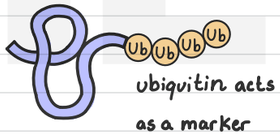


Recycling amino acids

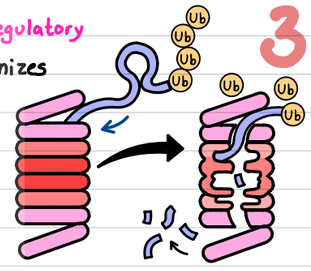
↳ Sustaining a functional **proteome** (entire set of proteins that can be expressed by cell/tissue/organism) requires constant protein breakdown and synthesis as proteins may become damaged (and need to be replaced) or are no longer required due to change in cell's activities

↳ **Proteasomes** : protein complexes which breakdown proteins marked for destruction via proteolysis

1 Proteins which are no longer functional or needed are tagged with **ubiquitin**



2 Proteasome's regulatory subunits recognizes tagged protein It is unfolded and fed into central chamber for breakdown

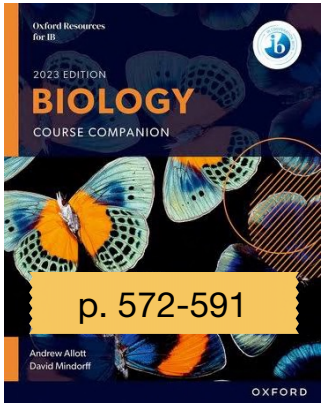


3 Inside central chamber, the protein is hydrolyzed by protease enzymes (whose active sites face inward), breaking down into smaller peptides. These are then ejected and further broken down in the cell, ready to be reused

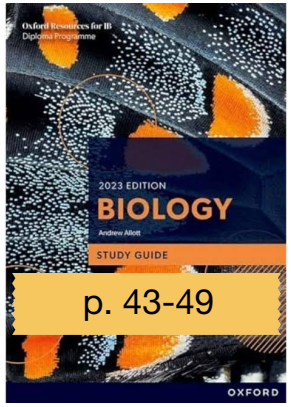
Resource Links

each resource is hyperlinked

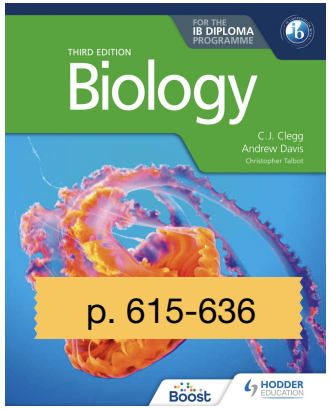
Textbooks



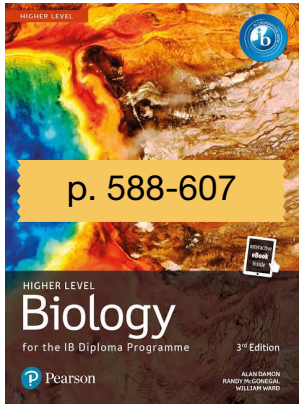
p. 572-591



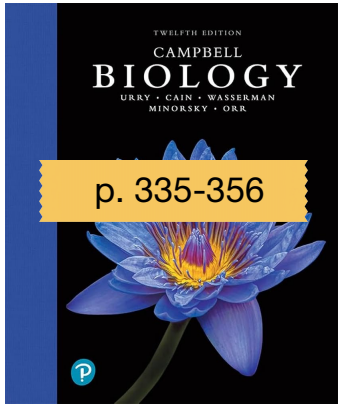
p. 43-49



p. 615-636

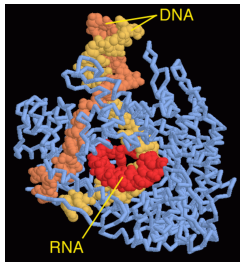


p. 588-607

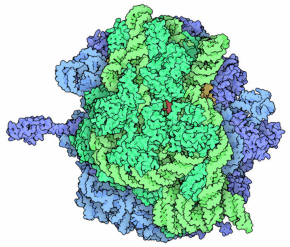


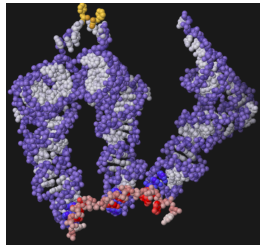
p. 335-356

3D models

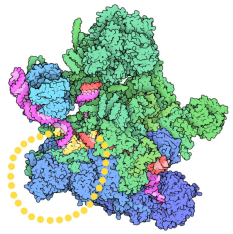


RNA polymerase


Ribosome


tRNA


Proteasome


Spliceosome

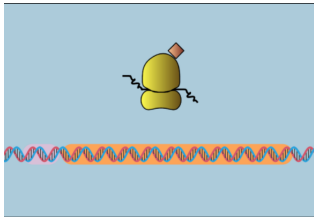
Articles

Fagundes, N. J., Bisso-Machado, R., Figueiredo, P. I., Varal, M., & Zani, A. L. (2022). What we talk about when we talk about “Junk DNA.” Genome Biology and Evolution, 14(5). https://doi.org/10.1093/gbe/evac055

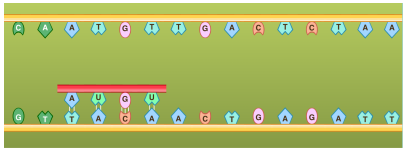
Ramakrishnan V. (2002). Ribosome structure and the mechanism of translation. Cell, 108(4), 557–572. https://doi.org/10.1016/s0092-8674(02)00619-0

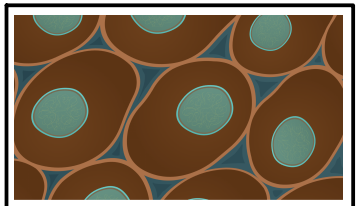
Srinivas, N., Rachakonda, S., & Kumar, R. (2020). Telomeres and Telomere Length: A General Overview. Cancers, 12(3), 558. https://doi.org/10.3390/cancers12030558

Simulators / Interactives


Gene expression

Transcription in prokaryotes


Transcription


Transcription

DNA —> RNA —> protein