D1.2 Protein synthesis

Continuity and change—Molecules

Standard level and higher level: 3 hours

Additional higher level: 3 hours

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?

SL and HL

D1.2.1—Transcription as the synthesis of RNA using a DNA template
Students should understand the roles of RNA polymerase in this process.
Transcription is the process by which **messenger RNA (mRNA)** is synthesized from a DNA template (**antisense** strand) in the nucleoid or nucleus, producing a sequence identical to the **sense** strand (complementary to the template). **RNA polymerase** catalyzes transcription by: (1) Binding to the DNA template (antisense) at the start of the gene (2) Unwinding the double helix by breaking hydrogen bonds between base pairs (3) **Elongation**: adding ribonucleotides in a 5' to 3' direction through complementary base pairing and connecting their backbones via phosphodiester bonds (4) Detaching from the DNA template when a **terminator sequence** is reached and rewinding of helix

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.
The **messenger RNA (mRNA)** produced by transcription is made up of adenine, uracil (pairs with adenine), cytosine, and guanine. As the RNA polymerase carries out elongation, A forms 2 hydrogen bonds with U and C forms 3 hydrogen bonds with G, ensuring accuracy.

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.
Since DNA contains all information for cell activities, its sequences must remain fixed and safely stored. In somatic cells that do not divide, such sequences must be conserved throughout the life of a cell, which are achieved via a wide range of mechanisms: • Transcription occurs without any changes to the template sequence and disseminates information within DNA quickly and efficiently • Structural protection (i.e. nucleus) • Intrinsic chemical stability from hydrogen bonds • Complex proofreading and repair mechanisms

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.

Gene expression is the process by which genetic information is used to direct changes in the organism's phenotype, which mostly occurs by regulating protein synthesis.

Not all genes in a cell are expressed at any given time; genes can be switched on and off by regulating their transcription rate (which is the first stage of gene expression). Increasing transcriptional activity of a gene increases the extent to which the gene's function is expressed within a cell.

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.

Translation is the synthesis of polypeptides from mRNA in the cytoplasm in eukaryotes. In prokaryotes, transcription and translation occur simultaneously as there is no nucleus.

D1.2.7—Complementary base pairing between tRNA and mRNA

Include the terms "codon" and "anticodon".

A **codon** is a sequence of three consecutive nucleotides in a DNA or RNA molecule that 'code' for a specific amino acid. An **anticodon** is a sequence of three consecutive nucleotides in a tRNA molecule that is complementary to the sequence of a codon in an RNA molecule.

During translation, the anticodon binds to its respective codon in mRNA in an antiparallel manner through complementary base pairing, ensuring accurate and reliable protein synthesis.

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.

- mRNA contains the codons required for initiating and carrying out translation of nucleotides into amino acids
- tRNA is an RNA molecule that binds to an amino acid at its 3' end and contains the anticodon for that amino acid
- Small ribosomal subunit contains a binding site for mRNA
- Large ribosomal subunit contains the majority of the A (aminoacyl), P (peptidyl), and E (exit) sites, allowing for 2 tRNAs to simultaneously bind

D1.2.8—Features of the genetic code

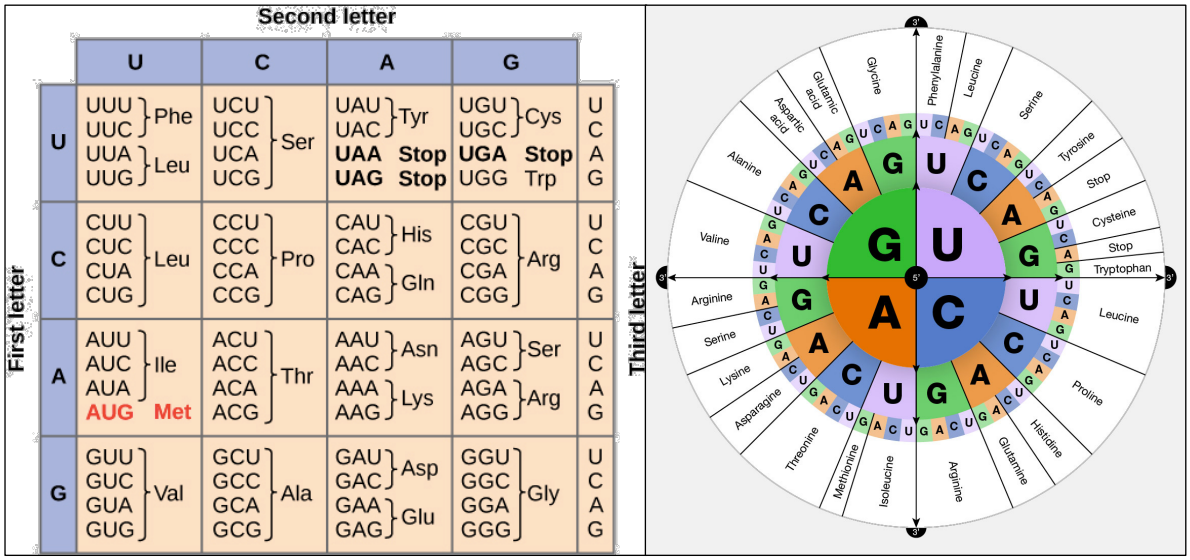
Students should understand the reasons for a triplet code. Students should use and understand the terms "degeneracy" and "universality".

The **genetic code** is the set of rules by which a polypeptide chain is synthesized from the nucleotide sequence of a gene via an intermediary mRNA molecule. The genetic code has 2 main features:

- **Degeneracy (redundancy)**: multiple anticodons code for the same amino acid, which helps make point (base) substitutions less harmful.
- **Universality**: all organisms use the same codons and anticodons for the same amino acids, which is crucial in biotechnology.

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.



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

Figure 1: tables of the genetic code (Khan Academy, NHGRI)

Tables of the genetic code specify the **mRNA sequence** of each codon and its respective amino acid read in a 5' to 3' direction.

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.

The process of translation is as follows:

1. tRNA is activated (attached to an amino acid) by an enzyme via ATP hydrolysis
2. Activated Met-tRNA binds to the small ribosomal subunit
3. mRNA binds to small subunit and once Met-tRNA scans and finds the start codon (AUG), its anticodon forms hydrogen bonds with the start codon
4. Large ribosomal subunit binds at the P-site
5. Another activated tRNA binds to the A-site and forms hydrogen bonds with its respective anticodon
6. A peptide bond forms between the 2 tRNAs' amino acids, catalyzed by the ribosome
7. Translation complex is **translocated** (moves) by 1 codon (5' to 3'); P-site tRNA moves to E-site where it exists and A-site tRNA moves to P-site
8. Another activated tRNA binds to the A-site and forms hydrogen bonds with its respective anticodon, then a peptide bond forms between the 2 tRNAs' amino acids; **elongation cycle** continues until a **stop codon** is reached and the process is terminated

D1.2.11—Mutations that change protein structure

Include an example of a point mutation affecting protein structure.

Base (point) substitutions can change a codon within mRNA, which changes the identity of the translated amino acid and thus affects the overall polypeptide structure.

Sickle cell anemia is an autosomal recessive disease caused by a point mutation in the **Hemoglobin (Hb) Beta Subunit**, changing the amino acid from Glutamic acid (**Glu**) to Valine (**Val**). This distorts the overall structure of Hb, leading to sickle-shaped RBCs as a result of abnormal Hb clustering under certain conditions (like high Hb concentrations or low oxygen levels). This causes RBCs to become trapped in small blood vessels, which causes blockage and pain. Heterozygous individuals are immune to malaria.

Additional higher level

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.

Most of human DNA is composed of non-coding sequences like (**GRRIT**):

- **Gene expression regulators** (promoters, enhancers, etc.)
- **Repetitive DNA sequences** (satellite DNA, transposons, etc.)
- **RNA coding sequences** (genes that produce RNA like rRNA and tRNA as the final product)
- **Introns** (non-coding sequences within coding genes)
- **Telomeres** (repetitive sequences towards the ends of chromosomes to protect DNA from damage)

D1.2.12—Directionality of transcription and translation

Students should understand what is meant by 5' to 3' transcription and 5' to 3' translation.

In transcription, RNA polymerase binds to the 3’ end of the DNA template and begins synthesizing the mRNA in a 5’ to 3’ direction *with respect to the mRNA*. In translation, the ribosome moves in a 5' to 3’ direction along mRNA.

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.

The **promoter** consists of a few sequences located **upstream** (5’ end) of the gene that RNA polymerase binds to in order for transcription to be initiated. In prokaryotes, one transcription factor (and its types) binds to the promoter to initiate transcription. In eukaryotes, **general transcription factors** are a large group of proteins that bind to the promoter and to RNA polymerase in order to ensure accurate initiation. Once the gene is transcribed, a **terminator sequence** is reached, causing RNA polymerase to detach. Thus, the promoter is <u>not</u> part of the nascent mRNA but the terminator sequence is.

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.

The **reading frame** encompasses the different ways in which a DNA or RNA nucleotide sequence can be divided into codons. Initiation of translation needs to be done accurately to ensure that the correct (a) reading frame (b) start codon are identified. In **eukaryotes**:

1. A specific enzyme attaches a methionine to its tRNA via ATP hydrolysis, 'activating' the initiator tRNA
2. Initiator tRNA binds to small ribosomal subunit (**ternary complex**)
3. Ternary complex attaches to 5' end of mRNA and slides along it until reaching the start codon (AUG)
4. Hydrogen bonding between tRNA anticodon and mRNA codon occurs
5. Large ribosomal subunit binds to the initiator tRNA on the P-site
6. tRNA carrying the anticodon for the next codon attaches to A site
7. rRNA catalyzes the formation of the peptide bond between the 2 amino acids, creating a dipeptide that is attached to the A-site tRNA

Prokaryotic initiation of translation is different but elongation and termination are identical to 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.

Post-transcriptional modification occurs only in eukaryotes in order to help protect the mRNA during nuclear export (prokaryotes have no nucleus).

1. **Capping**: a methylated guanine base is added to the 5' end of the mRNA, which helps to prevent degradation by exonucleases (enzymes).
2. **Polyadenylation**: addition of ~200 adenine bases (**polyA tail**) to the 3' end of the mRNA, which:
   (a) prevents degradation by exonucleases
   (b) assists in export from nucleus
3. **Splicing**: genes are composed of **exons** (sequences that are expressed and translated into protein) and **introns** (intervening sequences that do not code for amino acids, thought to be evolutionary remnants), which are alternating in sequence. Splicing involves the removal of introns and joining together of exons with the aid of proteins.

Both capping and splicing occur as the mRNA is still being synthesized, but polyadenylation occurs after mRNA synthesis is complete.

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.

Alternative splicing involves the splicing (gluing) together of exons in different combinations, leading to the synthesis of different polypeptide chains and thus allowing for one gene to code for multiple polypeptides.

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.

Polypeptides usually undergo several **post-translational changes** to reach their functional state, like (**CIS**):

- **Combining polypeptide chains** and synthetic groups to form quaternary structures
- **Intramolecular interactions** (i.e. forming disulfide bridges)
- **Side chain alterations** (i.e. adding a carbohydrate to an amino acid side chain)

Insulin requires several post-translational modifications before becoming functional:

1. Insulin gene is transcribed and its mRNA is translated inside the lumen of the rough endoplasmic reticulum (RER) into **pre-proinsulin**
2. A protease within the RER removes the **signal sequence** from pre-proinsulin's N-terminal, forming **proinsulin**
3. **2 disulfide bridges** form between the A and B chains and **1** within the A chain
4. Proinsulin is transported to the Golgi apparatus
5. **C-peptide** (in between the A and B chains) is cleaved

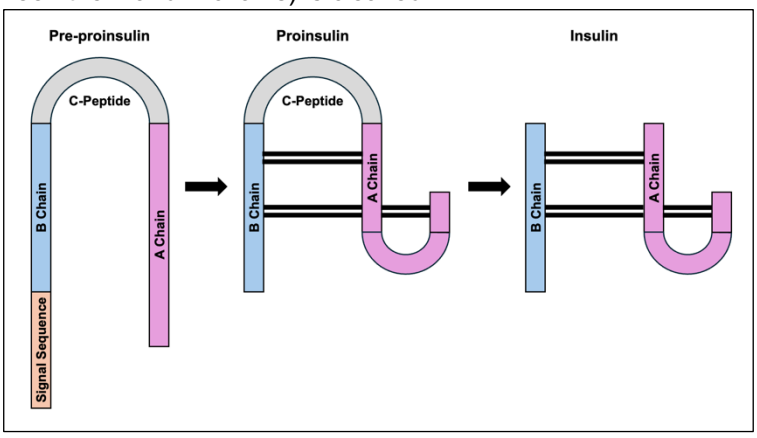


Figure 2: post-translational modification of pre-proinsulin to insulin.

D1.2.19—Recycling of amino acids by proteasomes

Limit to the understanding that sustaining a functional proteome requires constant protein breakdown and synthesis.

Misfolded or damaged proteins need to be removed by the cell, in addition to those that only serve short-term functions, in order to maintain the functionality of the proteome.

1. Specific amino acid sequences are exposed when a protein is misfolded or accumulates sufficient damage
2. Enzymes are recruited to catalyze (using ATP) the addition of several **ubiquitin** (a protein) units to the damaged protein, forming a **polyubiquitin chain**
3. Polyubiquitinated proteins are recognized by and bind to the **stoppers** of a proteosome
4. The stoppers unfold the protein and thread it into the **central cylinder**
5. Threaded protein comes in contact with the active sites of **proteases** (protein-digesting enzymes) lining the inner surface of the central cylinder, degrading it via hydrolysis to amino acids that are later recycled

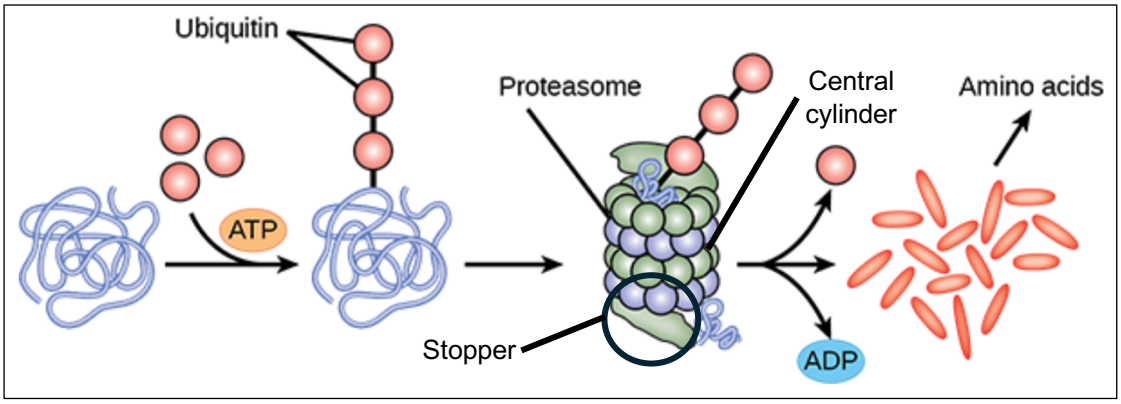


Figure 3: recycling of proteins by proteosomes (Ann Clark).

Linking questions

• How does the diversity of proteins produced contribute to the functioning of a cell?

• What biological processes depend on hydrogen bonding?

Review questions

SL and HL

- Define gene expression. [1]
- Outline the features of the genetic code and their benefits. [2]
- Explain why different types of human cells express different genes. [2]
- Explain the role of hydrogen bonding in transcription. [4]
- Compare and contrast transcription and translation. [4]
- Explain the role of RNA polymerase in transcription. [4]
- Describe transcription. [4]
- Explain how DNA molecules retain their stability. [4]
- Using an example, explain how base substitutions can lead to disease. [4]
- Explain the roles of different types of RNA molecules in protein synthesis. [6]
- Describe translation. [7]
- Describe how a cell produces a sequence of amino acids from a sequence of DNA bases. [8]

Additional Higher Level

- Outline the differences you would observe between eukaryotes and prokaryotes with regards to protein synthesis under a light microscope. [2]
- Outline how protein diversity is achieved during protein synthesis. [2]
- Outline how the modifications that occur after transcription and translation increase diversity. [2]
- Outline how transcription initiation is regulated. [3]
- Explain how an error in polyadenylation can lead to disease. [3]
- Describe how eukaryotic cells modify their mRNA post-transcription. [5]
- Describe the post-translational modifications that produce insulin. [5]
- Describe how proteins are recycled. [6]
- Describe how translation is initiated. [6]
- Compare and contrast post-transcriptional and post-translational modification. [8]
- Describe how insulin is synthesized, modified, and secreted by pancreatic cells. [8]
- Describe the journey of an intracellular protein molecule, starting from translation. [8]

References

“Definition of Codon - NCI Dictionary of Genetics Terms.” *Cancer.gov*, www.cancer.gov/publications/dictionaries/genetics-dictionary/def/codon.

Ann Clark, Mary, et al. *Biology 2e*. E-book, OpenStax, 2018, https://openstax.org/books/biology-2e/pages/1-introduction. OpenStax.

Gopalan, Chaya, and Erik Kirk. “Endocrine Pancreas.” *Elsevier eBooks*, 2022, pp. 209–22. https://doi.org/10.1016/b978-0-12-823421-1.00008-1.

Gordon Betts, J., et al. *Anatomy and Physiology 2e*. E-book, OpenStax, 2022, https://openstax.org/books/anatomy-and-physiology-2e/pages/1-introduction. OpenStax.

National Center for Biotechnology Information (US). Genes and Disease [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 1998-. Anemia, sickle cell. Available from: https://www.ncbi.nlm.nih.gov/books/NBK22238/