Protein synthesis

D1.2 · SL and HL

The same gene can produce different amounts and forms of protein

Follow information, then follow the product

Full shared D1.2: transcription, stable DNA templates, gene expression, RNA/ribosome roles, codons, elongation and mutation effects.

AHL: direction, promoters, non-coding DNA, RNA processing/splicing, initiation, protein maturation and recycling.

Sequences and numerical examples are invented unless explicitly identified as named biology.

Name the template, product and purpose

Process	**Information and product**
Replication	DNA $\rightarrow$ DNA; copies for continuity
Transcription	DNA $\rightarrow$ RNA; supplies a transcript
Translation	mRNA $\rightarrow$ polypeptide sequence at a ribosome
Protein synthesis	Amino acids are assembled; DNA and RNA do not turn into amino acids

A polypeptide is not automatically a finished functional protein

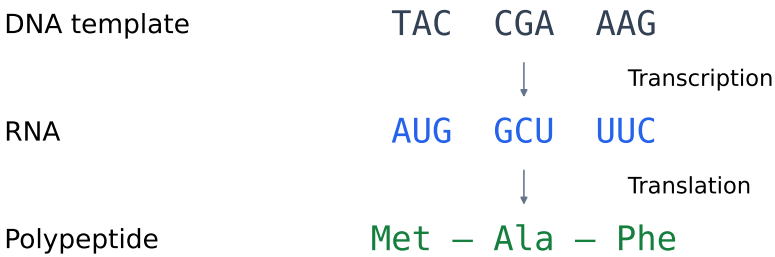
Keep cell type and gene product explicit

Some genes make functional RNAs, including rRNA and tRNA, rather than polypeptides.

Eukaryotic nuclear transcription and cytoplasmic translation are separated by a nuclear envelope; prokaryotes have no nucleus.

A new amino-acid chain may need folding, processing or assembly before functioning.

Follow complementarity before reading codon meanings



Invented aligned segment; complementary RNA, then codon meanings

DNA specifies RNA; RNA codons specify amino-acid order.

Construct the RNA segment from the template bases

1. **Invented template.** T–A–C–C–G–A–A–A–G.
2. **Pair DNA with RNA.** T→A; A→U; C→G; G→C.
3. **RNA product.** A–U–G–G–C–U–U–U–C. In the stated frame: AUG GCU UUC. Do not put thymine in RNA.

Explain polymerase and the two bonding relationships

Feature	Mechanism
Template exposure	A short DNA region opens; one strand is used for this transcript
Base pairing	Hydrogen bonding supports complementary nucleotide selection
RNA polymerase	Catalyses covalent joining into the new RNA backbone
After copying	RNA separates; DNA can pair again without losing its normal base sequence

A long-lived non-dividing cell still uses its stable templates

Transcription does not consume the gene

RNA nucleotides come from the available supply, not pieces removed from the DNA backbone.

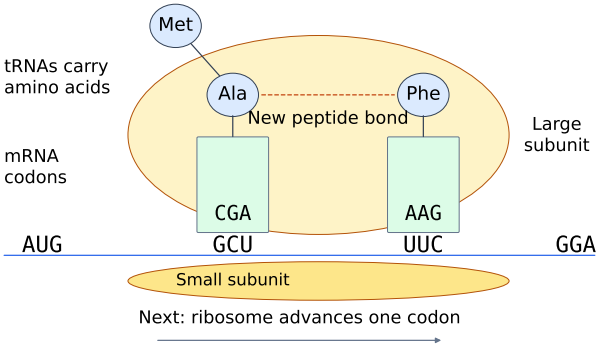
The same gene can be transcribed repeatedly. A cell can make proteins without copying its whole genome for cell division.

Maintained sequence does not mean DNA damage or mutation is impossible.

Do not infer gene loss from a missing product signal

1. **Invented observation.** Both cells contain an intact gene; only A has detectable transcript and protein at this sampling time.
2. **Explanation.** Different **transcription** activity can change mRNA availability and product abundance without different genes.
3. **Evaluate.** Check DNA, RNA and protein separately. Later regulation, detection sensitivity and timing also affect the result.

Two tRNAs connect neighbouring codons with chain transfer



Matching tRNAs position a growing chain and an amino acid.

Distinguish the message, adaptor and assembly machinery

Component	**Role**
mRNA	Ordered codons specify the sequence
tRNA	Anticodon pairs with a **codon**; a separate site carries an amino acid
Small subunit	Binds and positions mRNA
Large subunit	Supports peptide bonding; two tRNAs can participate together

A codon and its aligned anticodon are complementary

1. **Codon GCU.** Aligned anticodon CGA; the appropriate charged tRNA carries alanine.
2. **Codon UUC.** Aligned anticodon AAG; the appropriate charged tRNA carries phenylalanine.
3. **Accuracy.** Correct amino-acid charging matters. The anticodon is not the site to which the amino acid attaches.

Compare the numbers of possible base signals

1. **One base.** $4^1 = 4$ possible signals.
2. **Two bases.** $4^2 = 16$: too few for twenty standard amino acids.
3. **Three bases.** $4^3 = 64$: sufficient for amino acids and stop signals. Three bases form one **codon**.

Keep degeneracy, universality and stopping distinct

Property	**Meaning**
Degeneracy	Several codons can specify one amino acid; GCU and GCC both Ala
Standard meaning	A **codon** does not arbitrarily choose several amino acids
Near universality	Broad shared code with known exceptions, including some mitochondrial codes
Start/stop	AUG commonly starts in appropriate context; UAA/UAG/UGA add no amino acid

Read this excerpt as mRNA codons

Codon(s)	**Amino acid or signal**
AUG; GCU/GCC	Met; Ala
UUU/UUC; UAU/UAC	Phe; Tyr
GAA/GAG; GUG	Glu; Val
GGA; UAA/UAG/UGA	Gly; Stop

Read the specified frame from start to stop

1. **Invented coding region.** AUG GCU UUC GGA UAA.
2. **Use the mRNA table.** Met–Ala–Phe–Gly, then stop.
3. **Count.** Five codons, four residues. A stop is not an amino acid; an anticodon is not the message to look up.

Check what the nucleotide count includes

1. **Model.** 63 coding nucleotides, one terminal stop, no untranslated sequence or introns.
2. **Codons and residues.** Codons $= 63/3 = 21$; residues $= 21 - 1 = 20$.
3. **Peptide bonds.** One linear 20-residue chain has $20 - 1 = 19$ peptide bonds before processing.

Whole-gene length is not a ready-made peptide calculation

Use only the stated translated coding-region model

The relation $L = B/3 - 1$ assumes B includes one terminal stop and no other included sequence.

Genes may contain regulatory, untranslated and intron regions; full mRNA also contains untranslated sequence.

Mature proteins can differ from their initial chains after processing. A 90-nucleotide coding region under the stated model gives 29 residues.

Explain the repeating cycle in its causal order

1. **Match.** Incoming charged tRNA pairs with the next **codon**; another tRNA holds the chain.
2. **Transfer.** Peptide bonding attaches the chain to the incoming amino acid and its tRNA.
3. **Move.** The ribosome advances by one codon, repositions the tRNAs and permits uncharged tRNA exit. mRNA is not consumed as chain material.

One added amino acid creates one more peptide bond

1. **Existing chain.** Met–Ala–Phe–Gly: four residues joined by three peptide bonds.
2. **Add one residue.** Five residues joined by four peptide bonds.
3. **Interpretation.** Count the links within the stated linear chain; processing and multiple-chain proteins require their actual structures.

A base substitution can have different coding effects

Change	Effect under the standard code
GCU $\rightarrow$ GCC	Ala remains Ala; synonymous
UUC $\rightarrow$ UAC	Phe changes to Tyr
UAU $\rightarrow$ UAA	Creates a stop signal
Interpretation	Position and sequence determine the effect; one base change does not always destroy a protein

Connect the sickle-cell substitution to protein behaviour

1. **Coding DNA.** GAG → GTG; corresponding mRNA GAG → GUG.
2. **Residue.** Glu → Val, commonly numbered beta 6 in the mature chain.
3. **Consequence.** Changed side-chain properties favour deoxygenated haemoglobin S polymerization and can distort red cells. This is a named biological example, not invented data.

Regroup the sequence after deleting coding base 4

1. **Original.** AUG GCU UUC GGA UAA.
2. **Delete the G.** AUG CUU CGG AUA A: Met–Leu–Ser–Asp, plus an incomplete final group.
3. **Limit.** No complete stop is displayed in the regrouped segment. More downstream sequence is needed to infer the full product.

A shifted frame does not prove every later amino acid must differ

Predict the grouping, then inspect the actual code and sequence

A one-base insertion/deletion changes downstream triplet grouping. Degeneracy or chance agreement can preserve some amino acids.

A stop may arise later, but need not appear immediately. Do not invent the unseen downstream sequence.

A three-base deletion preserves the downstream frame; a boundary-crossing deletion can also change the joined codon. Frame preservation does not guarantee function.

Name template and coding strands before reporting the RNA

1. **Template.** $3'$-TAC CGA AAG-$5'$.
2. **RNA and coding DNA.** RNA $5'$-AUG GCU UUC-$3'$; coding DNA $5'$-ATG GCT TTC-$3'$.
3. **Direction.** Polymerase follows template 3′ to 5′ and extends RNA 5′ to 3′. The ribosome reads mRNA 5′ to 3′; polypeptides extend amino-to-carboxyl, not from sugar 5′ ends.

Promoter recognition regulates initiation

A regulatory sequence need not encode a polypeptide to function

Transcription factors recognize specific DNA sequences and help recruit or position transcription machinery at an appropriate promoter.

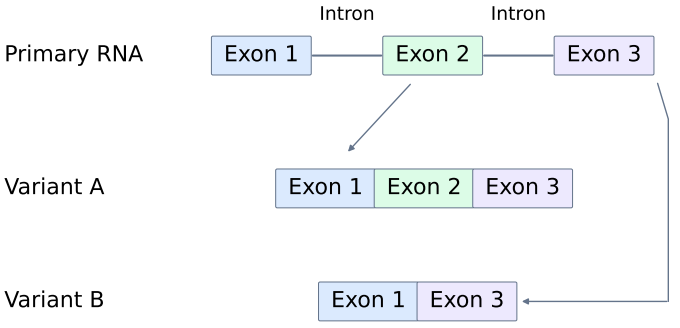
Binding can affect whether transcription begins; the promoter does not first have to be translated.

The template or “non-coding strand” name is different from saying an entire region has no coding information.

Non-coding DNA has several distinct kinds of role

Example	Function or fate
Regulatory sequences	Control gene expression
Introns	Transcribed into primary RNA, then removed
Telomeres	Help protect chromosome ends
rRNA/tRNA genes	Make functional RNAs rather than polypeptides; some exon sequence can also be untranslated

Different retained exon combinations yield different messages



Same gene; different retained exon combinations

Splicing changes retained RNA while the gene's DNA stays intact.

Distinguish the RNA-processing steps

Step	**Result**
Splicing	Introns removed and retained exons joined
5′ cap	Modified end contributes to stability and RNA handling
3′ polyA tail	Added during processing; contributes to stability and handling
Coding region	Cap, tail and untranslated regions are not all translated amino-acid instructions

Calculate a fictional exon-skipping example with stated conditions

1. **Primary segment.** Exons 120/90/180; introns 210/150. Total $= 750$ nucleotides.
2. **Retained coding sequence.** All exons $= 390$; skip middle exon $= 300$ nucleotides.
3. **Products.** With frame compatibility and one terminal stop: $390/3 - 1 = 129$ and $300/3 - 1 = 99$ residues. No UTRs/cap/tail included.

Alternative splicing is not random exon shuffling

Infer the actual permitted variants and their reading frames

Different combinations of retained exons can generate different polypeptides from one gene.

Retained exons keep the appropriate order. Exons may include untranslated sequence.

Do not calculate 2 to the power of the exon count as a guaranteed number of functional proteins. Sequence, reading frame and regulation constrain the products.

Start with the initiator tRNA in the correct position

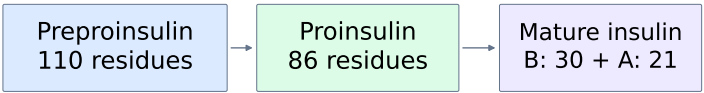
1. **Initiate.** Small subunit associates near the mRNA 5′ end and with initiator tRNA; scans to a suitable start **codon** in this simplified cap-dependent model.
2. **Assemble.** Initiator tRNA pairs at the start in P; large subunit completes the ribosome. Incoming tRNA can then match the next codon in A.
3. **Limit.** The initiator is an exception to "every tRNA first enters A". A stop is recognized by termination machinery, not a stop-carrying tRNA.

Follow one elongation cycle through A, P and E

Stage	Site relationship
Before transfer	P carries chain; A holds incoming charged tRNA
Peptide bonding	Chain transfers onto A-site amino acid/tRNA
One-codon movement	Chain-carrying tRNA shifts to P; uncharged tRNA moves toward E
Exit	Uncharged tRNA leaves from E and can be recharged

Insulin requires two stages of precursor processing

Human insulin: two stages of precursor processing



A and B remain linked by disulfide bonds; retained total = 51 residues

Process diagram; boxes are not molecular structures or lengths to scale

Two processing stages produce mature insulin from its precursor.

Keep C-peptide separate from the additional junction residues

1. **Human precursor.** Preproinsulin 110; signal 24; proinsulin $110 - 24 = 86$ residues.
2. **Mature product.** B 30 plus A 21 gives 51 retained residues; subsequent removal $86 - 51 = 35$.
3. **Removed region.** C-peptide 31 plus $35 - 31 = 4$ basic junction residues. Folding/disulfide bonding and proteolytic processing differ from RNA splicing.

A functional protein can require more than translation

Use the actual product's processing pathway

Folding, cleavage, carbohydrate addition, phosphorylation or chain assembly may be needed in suitable proteins; not all proteins undergo all modifications.

Insulin signal removal yields proinsulin, which folds and forms disulfide bonds; later processing releases C-peptide and removes adjacent residues.

The mature A and B chains remain linked by disulfide bonds. Neither RNA splicing nor precursor cleavage shortens the DNA gene each time.

Protein abundance reflects synthesis and breakdown together

Recover building blocks, then rebuild in a new specified sequence

Proteasomes break appropriate proteins into peptides; further breakdown can release reusable amino acids. Other degradation pathways also exist.

Recycling does not preserve the original protein's function. New synthesis still requires message, ribosomes and energy.

Proteasomes do not remove RNA introns or proofread DNA. Constant abundance can conceal rapid turnover.

Use a balance before judging whether a protein pool is stable

1. **Invented rates.** Synthesis 120 and degradation 80 molecules per hour; no other gains/losses.
2. **Net change.** $120 - 80 = 40$ molecules per hour gained.
3. **Stable abundance.** Synthesis and degradation both 200 per hour give zero net change despite continual replacement.

Attempt the complete handout practice before checking

Name the strand, frame, count and processing step

Construct an RNA complement; explain template stability, expression, ribosome/tRNA roles, code properties and elongation.

Translate the given coding sequence, count residues/bonds, and evaluate synonymous, amino-acid-changing, stop and frameshift changes.

AHL: reverse a supplied template; explain promoter/non-coding roles, RNA processing, initiation, insulin maturation and protein turnover.

Check the shared sequence and chain calculations

1. **RNA complement.** TAC CCG AAA $\rightarrow$ AUG GGC UUU. Practice **translation** AUG GCC UAC GAG UGA gives Met–Ala–Tyr–Glu, then stop.
2. **Coding length.** $156/3 = 52$ codons; $52 - 1 = 51$ residues; $51 - 1 = 50$ peptide bonds in the initial single chain.
3. **Elongation.** Six residues/five bonds become seven/six after one amino acid is added. A stop adds no residue.

AHL practice: orient and process the RNA

1. **Template direction.** Template $5'$-AGC CAT-$3'$ $\rightarrow$ RNA $5'$-AUG GCU-$3'$; coding DNA $5'$-ATG GCT-$3'$.
2. **Primary transcript.** $150 + 120 + 210 + 180 + 240 = 900$ nucleotides.
3. **Variants.** Coding 480/360 nucleotides; products $480/3 - 1 = 159$ and $360/3 - 1 = 119$, forty residues apart. Frame compatibility and one terminal stop assumed.

AHL practice: process and recycle the protein

1. **Insulin.** $110 - 24 = 86$; $30 + 21 = 51$; removal $86 - 51 = 35$, including $35 - 31 = 4$ junction residues.
2. **Pool.** $150 - 110 = 40$ molecules per hour increase, with no other gains/losses.
3. **Interpretation.** RNA splicing acts on a transcript; precursor cleavage acts on polypeptide. Proteasomal breakdown helps recover amino acids but does not replace sequence-directed synthesis.

Check yourself

Move from the identified molecule to a justified inference

Read mRNA codons in the stated frame; distinguish template, coding DNA, message and anticodon.

State which length includes a stop and which includes untranslated sequence. Do not equate a changed frame with every downstream amino acid changing.

AHL: distinguish direction, promoter regulation, RNA processing, initiation sites, protein cleavage and recycling.