background sheet

How is gene expression regulated?

Gene expression is a controlled process. Regulation enables organisms to respond, immediately and effectively in an energy efficient way, to extracellular and intracellular environmental stimuli.

A gene is switched 'on', or expressed, only when a cell is signalled to produce a particular molecular product.

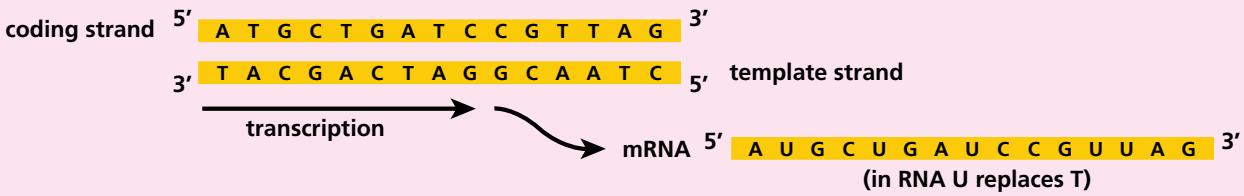
Regulated gene expression enables cells to produce products the body needs, at the appropriate time. Regulation occurs at all phases of gene expression: transcription, transcript processing and translation.

Transcription regulation

Transcription involves the molecule RNA polymerase decoding DNA, to produce a strand of messenger RNA (mRNA). Most gene regulation occurs during this stage of gene expression.

This document focuses on eukaryotic transcription of mRNA via RNA polymerase II.

There are three phases of transcription: initiation, elongation and termination.



For a single gene, only one DNA strand is the coding strand however, genes are located on both strands of DNA.

DNA's code is always read in a 5′ → 3′ direction from the coding strand, however the RNA message is transcribed from the template strand reading from 3′ → 5′. This results in the transcript (mRNA) replicating the coding strand.

Transcription initiation

Transcription is initiated when a protein is needed by a cell. Before transcription can begin a complex of regulatory molecules is formed. This is known as the transcription initiation complex (TIC).

The first molecules to begin forming the TIC are **transcription factors**: proteins that are key regulators of gene expression. Human cells contain thousands of transcription factors. Transcription factors bind to a specific DNA sequence or motif.

Transcription factors bind to an area of the gene called the promoter, recognising a specific nucleotide sequence, the TATA box. This region is found upstream of the transcription start site. Once transcription factors are in place, RNA polymerase locates and binds to the promoter, clamping onto the DNA.

Once RNA polymerase is in place the **enhancer**, a DNA sequence which may be located some distance away from the TIC, moves into position, often looping above the TIC. Regulatory proteins facilitate positioning of the enhancer.

The promoter's role is regulatory; it's not transcribed. It has three main functions: indicating the direction of transcription; determining which DNA strand acts as the template strand; and determining the starting point of transcription.

Finally the **mediator**, a large complex of proteins, binds to activator proteins on the enhancer, building a bridge between RNA polymerase, transcription factors and the enhancer sequence.

Transcription can only begin when all regulatory genes and molecules are active and in the right position.

Transcription elongation

As the elongation phase of transcription begins regulatory molecules that make up the transcription initiation complex are released, and RNA polymerase moves from the promoter.

Transcription elongation occurs by repetitive addition of complementary nucleotides to produce a strand of mRNA. Addition of nucleotides is controlled by **transcription elongation factors**: proteins that stabilise the connection between RNA polymerase and DNA, and determine speed of transcription. Average speed of elongation is 60 bases per second.

The newly emerging RNA transcript is open to attack by cellular enzymes. To prevent this the 5′ end of the transcript is protected with a methylated cap.

Transcription termination

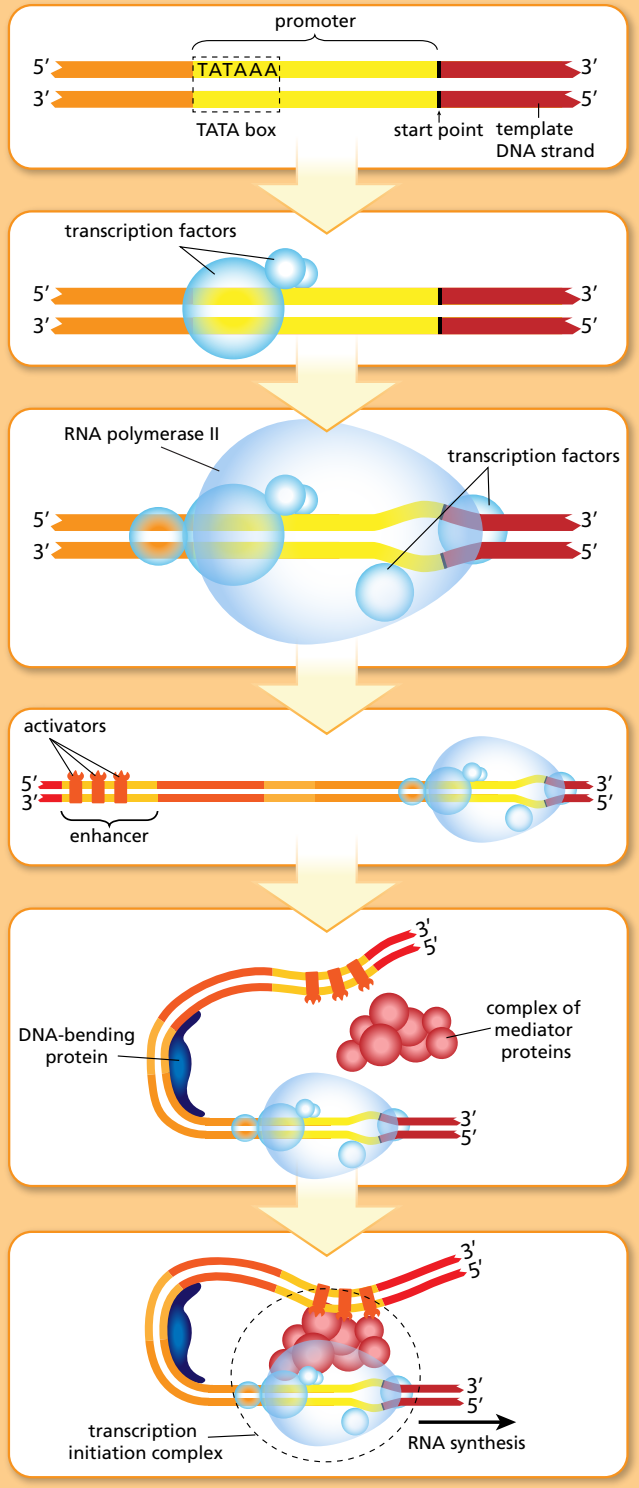
Elongation continues until regulatory proteins recognise a specific nucleotide sequence. Transcription termination occurs and RNA polymerase and the newly formed mRNA are released from the DNA template. As with other transcription events this process is complex and involves multiple molecules.

The 3′ end of the newly terminated transcript needs protection. This is done by attaching a series of adenines to the transcript (polyadenylation). This poly-A tail provides stability to mRNA.

Splicing

Before mRNA moves into the cytoplasm it undergoes further processing. Splicing removes introns, with the remaining exons joined together to form an mRNA molecule. Splicing is carried out by complexes of RNA and proteins known as **spliceosomes** and is controlled by motif-driven recognition. Splicing machinery recognises where cuts in the mRNA should be made and anneals the ends of the remaining exons back together.

Once processed mRNA leaves the nucleus and travels into the cytoplasm for the next stage of protein synthesis: translation.

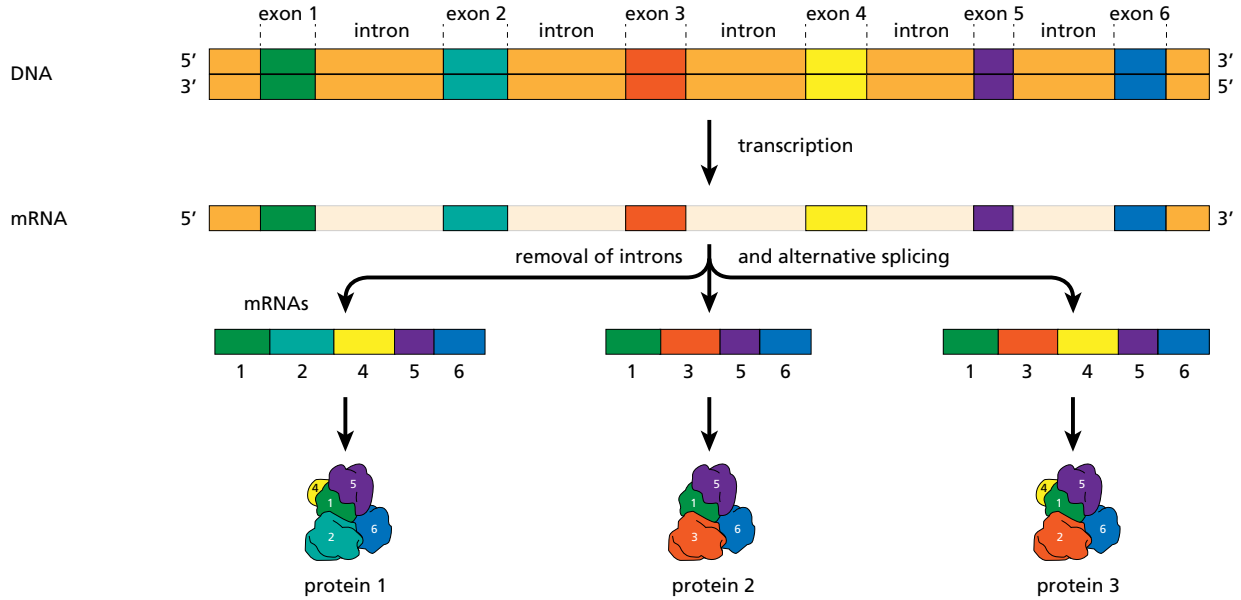


Regulation of transcription initiation.

If introns aren't essential for making proteins why have them?

Early researchers did not know the function of introns and they were initially called 'junk' DNA. It is now known that some introns play a role in regulation of gene expression. The number of introns within a genome also increases with organism complexity. For instance, there are around 25 000 genes in the human genome, less than the number found in fruit flies, but humans have significantly more introns.

Up to 90% of human genes undergo **alternative splicing**, where different exons are included or excluded in the mature mRNA product, resulting in different isoforms and protein products. Alternative splicing occurs in the cell nucleus and is important evolutionarily as it enables increased variation.



Alternative splicing: from a single gene different combinations of exons are joined together to produce multiple protein products.

Translation regulation

During translation an mRNA nucleotide sequence is translated into a protein's amino acid sequence. Translation requires complex molecular machinery: a ribosome, transfer RNA molecules and many supporting regulatory molecules.

Similar to transcription, translation involves three controlled phases: initiation, elongation and termination.

Translation initiation

Initiation involves assembling a ribosome complex at the start of mRNA, specified by a start codon: AUG. A ribosome includes a small and large subunit.

Initiation factors and enzymes regulate translation initiation. These regulatory molecules form an initiation complex that includes proteins, ribosome, mRNA and tRNA.

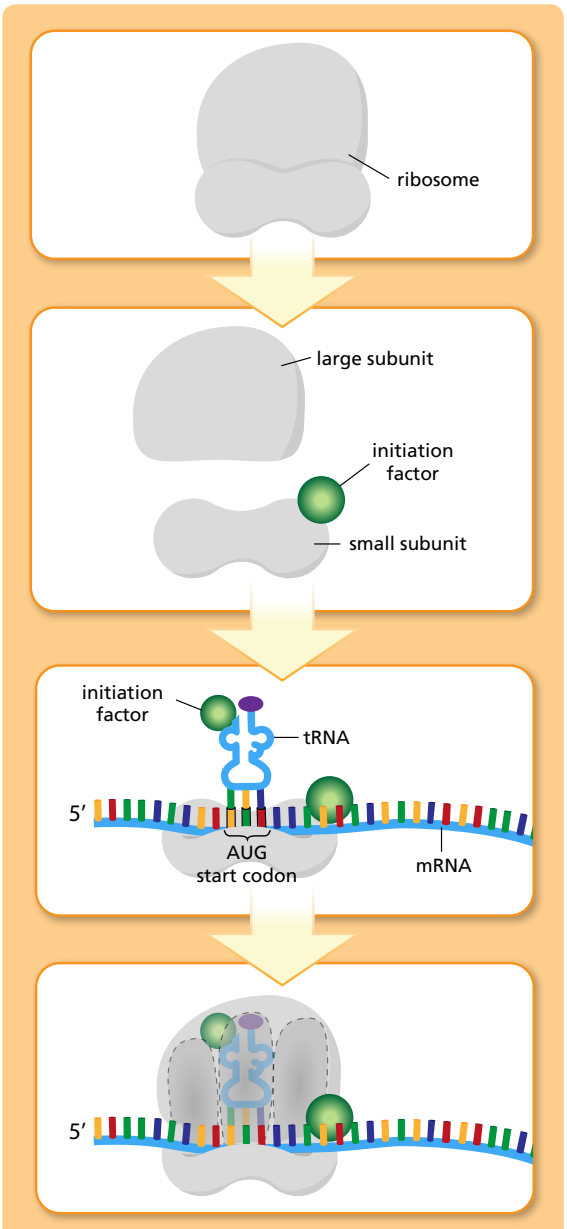
Translation elongation

Elongation involves adding amino acids, in a specific order determined by mRNA, to a growing polypeptide chain. This process is controlled by proteins known as **elongation factors** that help deliver appropriate tRNA to the ribosome.

Translation termination

Translation stops when the ribosome encounters a stop codon, which determines the length of a protein. Proteins called **release factors** are involved in this process. At termination the peptide chain is released and the ribosomal unit dissociates from mRNA.

Multiple interactions, between many different molecules, regulate gene expression in eukaryotic organisms. Gene regulation allows for increasing organism complexity. It's also efficient: only those genes required by the cell are expressed. Gene regulation controls how much of a protein is produced and how often it's produced, reducing waste and conserving energy.



Regulation of translation initiation.