

# Gene expression and regulation

## Learning goal

By the end of this learning material you would have learnt about:

- The process by which the genetic code directs protein synthesis to produce the structures of a cell
- The cellular processes that control the rate and manner of gene expression.

## Gene expression

Gene expression is the process by which the genetic code - the nucleotide sequence - of a gene is used to direct protein synthesis and produce the structures of the cell. Genes that code for amino acid sequences are known as 'structural genes'.

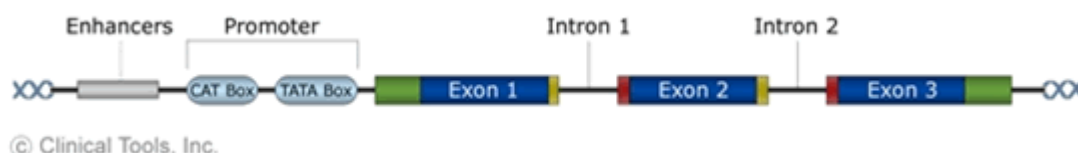
The process of gene expression involves two main stages:

**Transcription:** the production of messenger RNA (mRNA) by the enzyme RNA polymerase, and the processing of the resulting mRNA molecule

**Translation:** the use of mRNA to direct protein synthesis, and the subsequent post-translational processing of the protein molecule.

Some genes are responsible for the production of other forms of RNA that play a role in translation, including transfer RNA (tRNA) and ribosomal RNA (rRNA).

A structural gene involves a number of different components:



- **Exons.** Exons code for amino acids and collectively determine the amino acid sequence of the protein product. It is these portions of the gene that are represented in final mature mRNA molecule.
- **Introns.** Introns are portions of the gene that do not code for amino acids, and are removed (spliced) from the mRNA molecule before translation.

## Control regions

- **Start site.** A start site for transcription.
- **A promoter.** A region a few hundred nucleotides 'upstream' of the gene (toward the 5' end). It is not transcribed into mRNA, but plays a role in controlling the transcription of the gene. Transcription factors bind to specific nucleotide sequences in the promoter region and assist in the binding of RNA polymerases.
- **Enhancers.** Some transcription factors (called activators) bind to regions called 'enhancers' that increase the rate of transcription. These sites may be thousands of

nucleotides from the coding sequences or within an intron. Some enhancers are conditional and only work in the presence of other factors as well as transcription factors.

- **Silencers.** Some transcription factors (called repressors) bind to regions called 'silencers' that depress the rate of transcription.

Note: The term 'gene expression' is sometimes used to refer to the transcription phase alone.

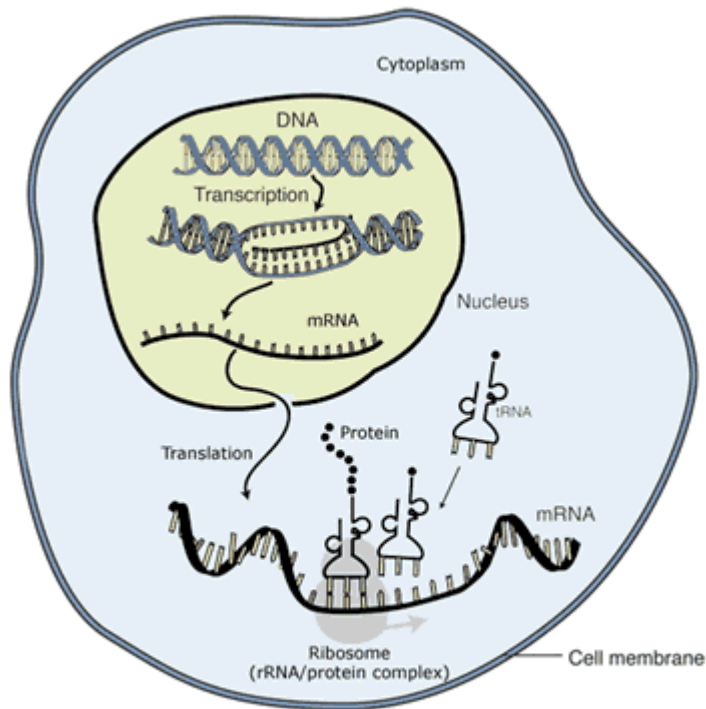


Image adapted from: National Human Genome Research Institute.

## Transcription

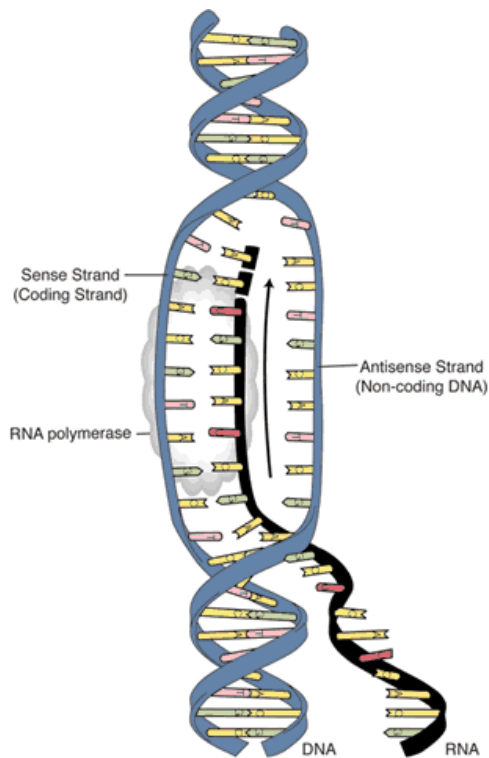


Image adapted from: National Human Genome Research Institute. Talking Glossary of Genetic Terms. Available at: [www.genome.gov/Pages/Information/DIR/VIP/Glossary/Illustration/antisense.shtml](http://www.genome.gov/Pages/Information/DIR/VIP/Glossary/Illustration/antisense.shtml).

Transcription is the process of RNA synthesis, controlled by the interaction of promoters and enhancers. Several different types of RNA are produced, including **messenger RNA (mRNA)**, which specifies the sequence of amino acids in the protein product, plus **transfer RNA (tRNA)** and **ribosomal RNA (rRNA)**, which play a role in the translation process.

Transcription involves four steps:

- **Initiation.** The DNA molecule unwinds and separates to form a small **open complex**. RNA polymerase binds to the promoter of the **template strand** (also known as the 'sense strand' or 'coding strand'). The synthesis of RNA proceeds in a 5' to 3' direction, so the template strand must be 3' to 5'.
- **Elongation.** RNA polymerase moves along the template strand, synthesising an mRNA molecule. In prokaryotes RNA polymerase is a holoenzyme consisting of a number of subunits, including a **sigma factor** (transcription factor) that recognises the promoter. In eukaryotes there are three RNA polymerases: I, II and III. The process includes a proofreading mechanism.
- **Termination.** In prokaryotes there are two ways in which transcription is terminated. In **?-dependent termination**, a protein ? is responsible for disrupting the complex involving the template strand, RNA polymerase and RNA molecule. In **?-independent termination**, a loop forms at the end of the RNA molecule, causing it to detach itself. Termination in eukaryotes is more complicated, involving the addition of additional adenine nucleotides at the 3' of the RNA transcript (a process referred to as **polyadenylation**).
- **Processing.** After transcription the RNA molecule is processed in a number of ways: introns are removed and the exons are spliced together to form a mature mRNA

molecule consisting of a single protein-coding sequence. RNA synthesis involves the normal base pairing rules, but the base thymine is replaced with the base **uracil**.

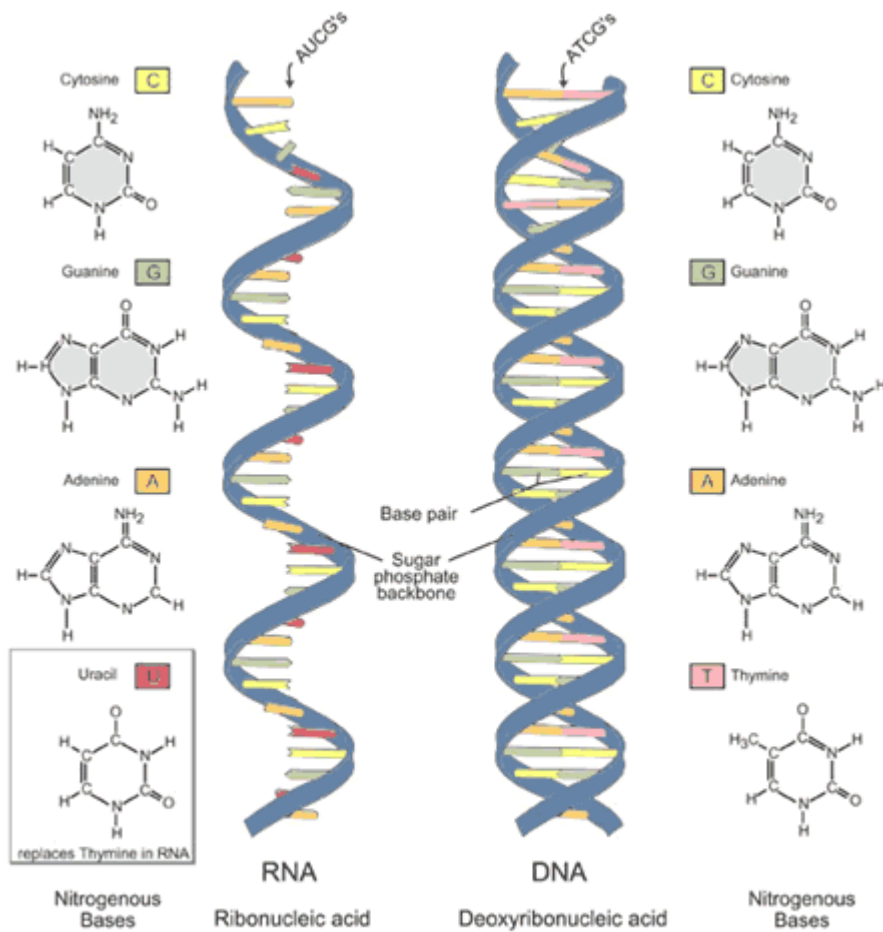


Image adapted from: National Human Genome Research Institute. Talking Glossary of Genetic Terms. Available at: [www.genome.gov/Pages/Information/DIR/VIP/Glossary/Illustration/rna.shtml](http://www.genome.gov/Pages/Information/DIR/VIP/Glossary/Illustration/rna.shtml).

## Translation

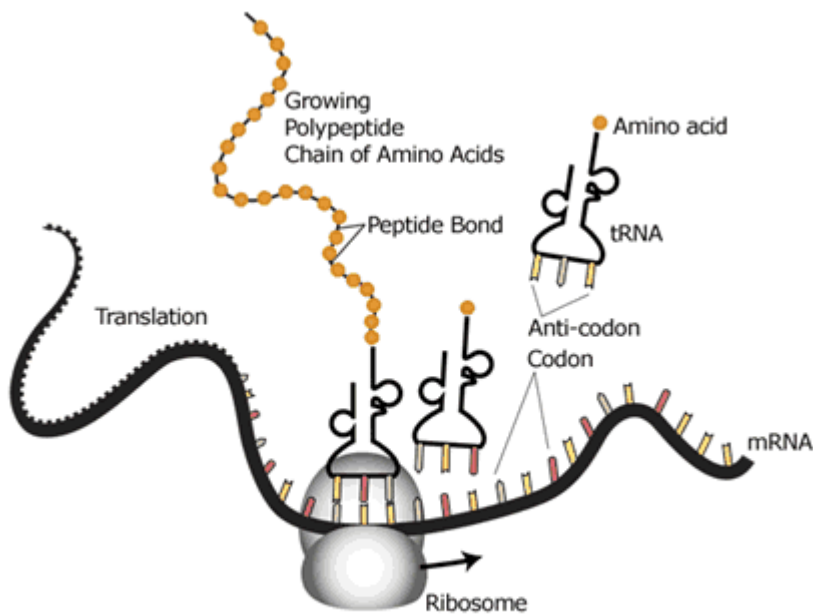


Image adapted from: National Human Genome Research Institute.

In translation the mature mRNA molecule is used as a template to assemble a series of amino acids to produce a polypeptide with a specific amino acid sequence. The complex in the cytoplasm at which this occurs is called a **ribosome**. Ribosomes are a mixture of ribosomal proteins and ribosomal RNA (rRNA), and consist of a large subunit and a small subunit.

Transcription involves four steps:

- **Initiation.** The small subunit of the ribosome binds at the 5' end of the mRNA molecule and moves in a 3' direction until it meets a start codon (AUG). It then forms a complex with the large unit of the ribosome complex and an initiation tRNA molecule.
- **Elongation.** Subsequent codons on the mRNA molecule determine which tRNA molecule linked to an amino acid binds to the mRNA. An enzyme peptidyl transferase links the amino acids together using peptide bonds. The process continues, producing a chain of amino acids as the ribosome moves along the mRNA molecule.
- **Termination.** Translation is terminated when the ribosomal complex reached one or more stop codons (UAA, UAG, UGA).

The ribosomal complex in eukaryotes is larger and more complicated than in prokaryotes. In addition, the processes of transcription and translation are divided in eukaryotes between the nucleus (transcription) and the cytoplasm (translation), which provides more opportunities for the regulation of gene expression.

- **Post-translation processing of the protein**

## Gene regulation

Gene regulation is a label for the cellular processes that control the rate and manner of gene expression. A complex set of interactions between genes, RNA molecules, proteins (including transcription factors) and other components of the expression system determine

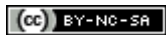
when and where specific genes are activated and the amount of protein or RNA product produced.

Some genes are expressed continuously, as they produce proteins involved in basic metabolic functions; some genes are expressed as part of the process of cell differentiation; and some genes are expressed as a result of cell differentiation.

Mechanisms of gene regulation include:

- Regulating the rate of transcription. This is the most economical method of regulation.
- Regulating the processing of RNA molecules, including alternative splicing to produce more than one protein product from a single gene.
- Regulating the stability of mRNA molecules.
- Regulating the rate of translation.

**Transcription factors** are proteins that play a role in regulating the transcription of genes by binding to specific regulatory nucleotide sequences.



This work is licensed under a [Creative Commons Licence](https://creativecommons.org/licenses/by-nc-sa/4.0/)