

DNA transcription

To transcribe means “to paraphrase or summarize in writing”. The information in DNA is transcribed - or summarized - into a smaller version - RNA - that can be used by the cell. This process is called **transcription**.

The process in which cells make proteins is called **protein synthesis**. It actually consists of two processes: **transcription and translation**.

Transcription takes place in the **nucleus**. It uses DNA as a template to make an RNA molecule. RNA then leaves the nucleus and goes to a ribosome in the cytoplasm, where translation occurs. Translation reads the genetic code in mRNA and makes a protein.

Transcription is the first part of the **central dogma of molecular biology: DNA → RNA**. It is the transfer of genetic instructions in DNA to **messenger RNA (mRNA)**. During transcription, a strand of mRNA is made that is complementary to a strand of DNA.

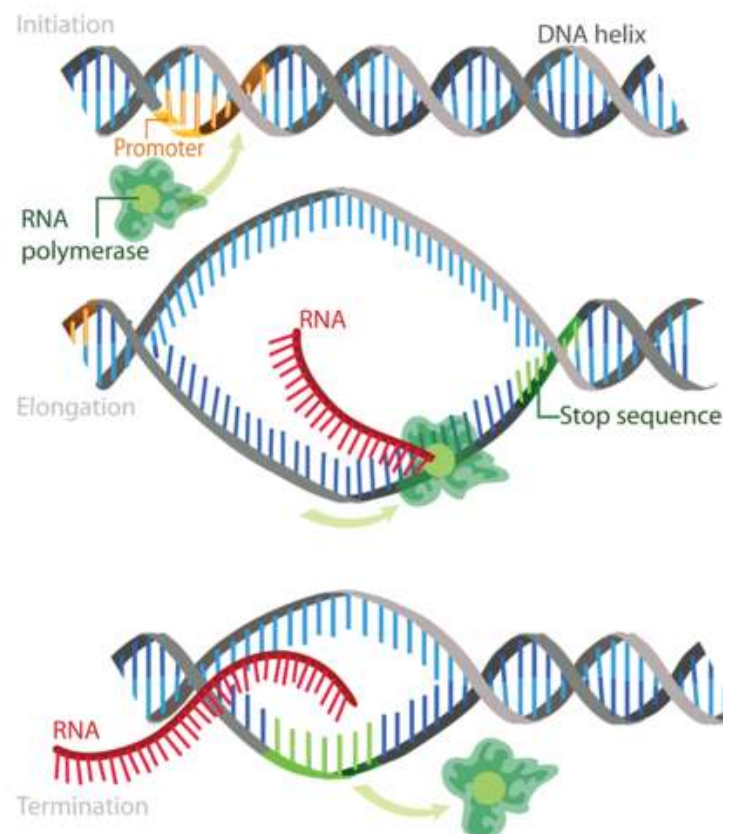
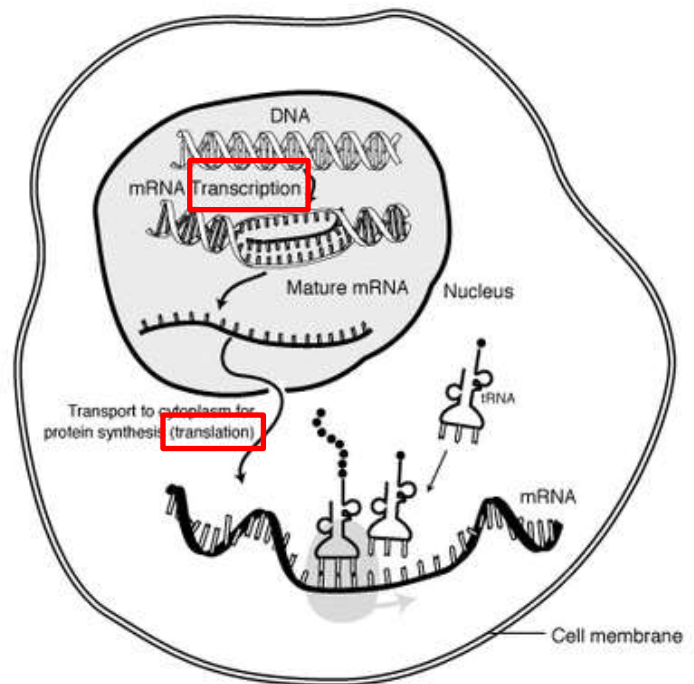
Steps of transcription

Transcription takes place in three steps: initiation, elongation, and termination.

Initiation is the beginning of transcription. It occurs when the enzyme **RNA polymerase** binds to a region of a gene called the **promoter**. This signals the DNA to unwind so the enzyme can “read” the bases in one of the DNA strands. The enzyme is now ready to make a strand of mRNA with a complementary sequence of bases.

Elongation is the addition of nucleotides to the mRNA strand. RNA polymerase reads the unwound DNA strand and builds the mRNA molecule, using complementary base pairs. There is a brief time during this process when the newly formed RNA is bound to the unwound DNA. During this process, an **adenine (A)** in the DNA binds to an **uracil (U)** in the RNA.

Termination is the ending of transcription, and occurs when RNA polymerase crosses a **stop (termination) sequence** in the gene. The mRNA strand is complete, and it detaches from DNA.



Genetics

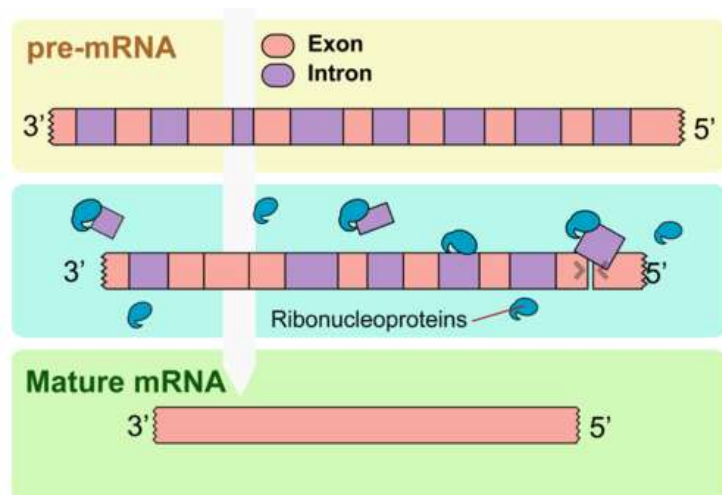
Processing mRNA

In eukaryotes, the new mRNA is not yet ready for translation. It must go through additional processing before it leaves the nucleus. This may include splicing, editing, and polyadenylation. These processes modify the mRNA in various ways. Such modifications allow a single gene to be used to make more than one protein.

Splicing removes introns from mRNA. **Introns** are regions that do not code for proteins. The remaining mRNA consists only of regions that do code for proteins, which are called **exons**. Ribonucleoproteins are nucleoproteins that contain RNA. Small nuclear ribonuclear proteins are involved in pre-mRNA splicing.

Editing changes some of the nucleotides in mRNA. For example, the human protein called APOB, which helps transport lipids in the blood, has two different forms because of editing. One form is smaller than the other because editing adds a premature stop signal in the mRNA.

Polyadenylation adds a “tail” to the mRNA. The tail consists of a **string of As** (adenine bases). It signals the end of mRNA. It is also involved in exporting mRNA from the nucleus. In addition, the tail protects mRNA from enzymes that might break it down.



Translation of mRNA to proteins

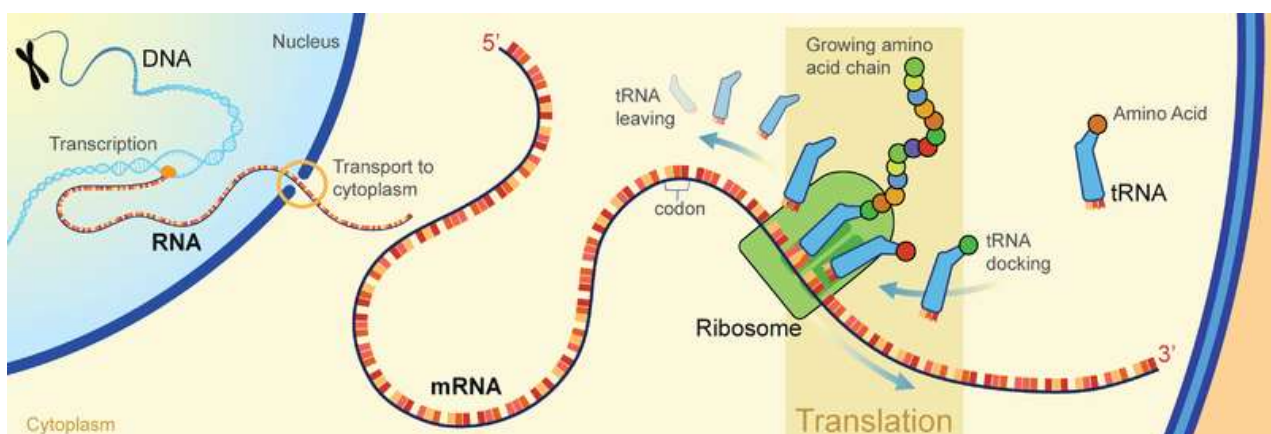
To go from one language to another f. e. Slovak to English, French to German, or nucleotides to amino acids, you must translate. Obviously, the type of translating discussed here translates from the language of nucleotides to the language of amino acids. But you need a code. And the code that changes the information embedded in DNA and RNA into ordered amino acids and proteins is **the genetic code**. And every living organism uses the same genetic code.

The genetic code consists of the sequence of nitrogen bases—A, C, G, U—in an mRNA chain. The four bases make up the “letters” of the genetic code. The letters are combined in **groups of three** to form code “words,” called **codons**. Each codon stands for (encodes) one amino acid, unless it codes for a start or stop signal. There are 20 common amino acids in proteins. There are 64 possible codons, more than enough to code for the 20 amino acids.

Translation is the second part of the **central dogma** of molecular biology: **RNA → Protein**. It is the process in which the genetic code in mRNA is read, one codon at a time, to make a protein. After mRNA leaves the nucleus, it moves to a **ribosome**, which consists of rRNA and proteins. The ribosome reads the sequence of **codons in mRNA**. The codon AUG codes for the amino acid methionine. This codon is also the start codon that begins translation. The **start codon** establishes the **reading frame** of mRNA. The reading frame is the way the letters are divided into codons. After the AUG start codon, the next three letters are read as the second codon. The next three letters after that are read as the third codon, and so on. The mRNA molecule is read, codon by codon, until a **stop codon** is reached. UAG, UGA, and UAA are all stop codons. They do not code for any amino acids. Stop codons are also known as termination codons.

Molecules of **tRNA** bring amino acids to the ribosome in the correct sequence of codons. To understand the role of tRNA, you need to know more about its structure. Each **tRNA molecule** has an **anticodon** for the amino acid it carries. An anticodon is a sequence of 3 bases, and is complementary to the codon for an amino acid. For example, the amino acid lysine has the codon AAG, so the anticodon is UUC. Therefore, lysine would be carried by a tRNA molecule with the anticodon UUC. Wherever the codon AAG appears in mRNA, a UUC anticodon on a tRNA temporarily binds to the codon. While bound to the mRNA, the tRNA gives up its amino acid. Bonds form between adjacent amino acids as they are brought one by one to the ribosome, forming a **polypeptide chain**. The chain of amino acids keeps growing until a **stop codon** is reached. For visualisation you can watch video at: <https://www.youtube.com/watch?v=gG7uCskUOrA>.

After a polypeptide chain is synthesized, it may undergo additional processes. For example, it may assume a folded shape due to interactions among its amino acids. It may also bind with other polypeptides or with different types of molecules, such as lipids or carbohydrates. Many proteins travel to the Golgi apparatus to be modified for the specific job they will do.



Genetics

Characteristics of the Genetic Code

The genetic code has a number of important characteristics.

The genetic code is **universal**. All known living organisms use the same genetic code. This shows that all organisms share a common evolutionary history.

The genetic code is **unambiguous**. Each codon codes for just one amino acid (or start or stop).

The genetic code is **redundant**. Most amino acids are encoded by more than one codon.

		2 nd base			
		U	C	A	G
1 st base	U	UUU (Phe/F) Phenylalanine	UCU (Ser/S) Serine	UAU (Tyr/Y) Tyrosine	UGU (Cys/C) Cysteine
		UUC (Phe/F) Phenylalanine	UCC (Ser/S) Serine	UAC (Tyr/Y) Tyrosine	UGC (Cys/C) Cysteine
		UUA (Leu/L) Leucine	UCA (Ser/S) Serine	UAA Ochre (Stop)	UGA Opal (Stop)
		UUG (Leu/L) Leucine	UCG (Ser/S) Serine	UAG Amber (Stop)	UGG (Trp/W) Tryptophan
	C	CUU (Leu/L) Leucine	CCU (Pro/P) Proline	CAU (His/H) Histidine	CGU (Arg/R) Arginine
		CUC (Leu/L) Leucine	CCC (Pro/P) Proline	CAC (His/H) Histidine	CGC (Arg/R) Arginine
		CUA (Leu/L) Leucine	CCA (Pro/P) Proline	CAA (Gln/Q) Glutamine	CGA (Arg/R) Arginine
		CUG (Leu/L) Leucine	CCG (Pro/P) Proline	CAG (Gln/Q) Glutamine	CGG (Arg/R) Arginine
	A	AUU (Ile/I) Isoleucine	ACU (Thr/T) Threonine	AAU (Asn/N) Asparagine	AGU (Ser/S) Serine
		AUC (Ile/I) Isoleucine	ACC (Thr/T) Threonine	AAC (Asn/N) Asparagine	AGC (Ser/S) Serine
		AUA (Ile/I) Isoleucine	ACA (Thr/T) Threonine	AAA (Lys/K) Lysine	AGA (Arg/R) Arginine
		AUG ^M (Met/M) Methionine	ACG (Thr/T) Threonine	AAG (Lys/K) Lysine	AGG (Arg/R) Arginine
	G	GUU (Val/V) Valine	GCU (Ala/A) Alanine	GAU (Asp/D) Aspartic acid	GGU (Gly/G) Glycine
		GUC (Val/V) Valine	GCC (Ala/A) Alanine	GAC (Asp/D) Aspartic acid	GGC (Gly/G) Glycine
		GUA (Val/V) Valine	GCA (Ala/A) Alanine	GAA (Glu/E) Glutamic acid	GGA (Gly/G) Glycine
		GUG (Val/V) Valine	GCG (Ala/A) Alanine	GAG (Glu/E) Glutamic acid	GGG (Gly/G) Glycine

nonpolar	polar	basic	acidic	(stop codon)
----------	-------	-------	--------	--------------