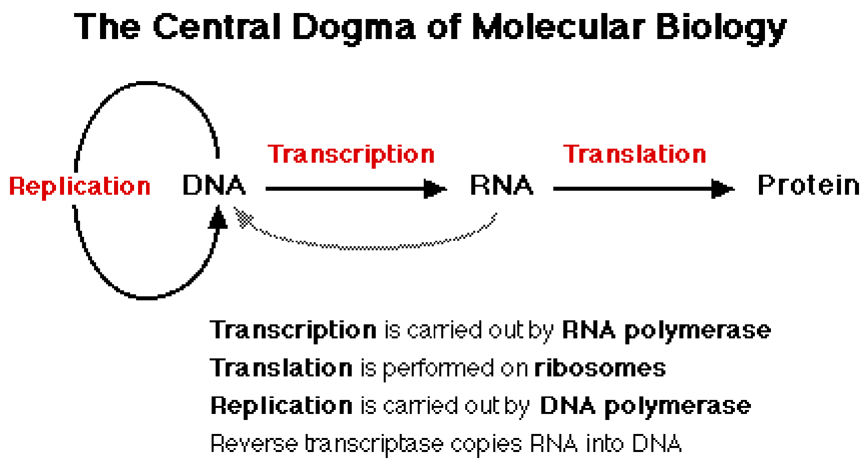
Lecture 15

Transcription and Translation

Central Dogma

The 'Central Dogma' is the process by which the instructions in DNA are converted into a functional product. It was first proposed in 1958 by Francis Crick, discoverer of the structure of DNA. The central dogma of molecular biology explains the flow of genetic information, from DNA to RNA, to make a functional product, protein. The central dogma suggests that DNA contains the information needed to make all of our proteins, and that RNA is a messenger that carries this information to the ribosome. The ribosomes serve as factories in the cell where the information is 'translated' from a code into the functional product.



Reverse Transcription

The synthesis of DNA from an RNA template, via reverse transcription, produces complementary DNA (cDNA). Reverse transcriptases (RTs) use an RNA template and a short primer complementary to the 3' end of the RNA to direct the synthesis of the first strand cDNA, which can be used directly as a template for the Polymerase Chain Reaction (PCR). Some RNA viruses, called "retroviruses" do Reverse Transcription.

RNA

RNA stands for ribonucleic acid. It is an important molecule with long chains of nucleotides. A nucleotide contains a nitrogenous base, a ribose sugar, and a phosphate. Just like DNA, RNA is vital for living beings. RNA can both store information (like DNA) and catalyze chemical reactions (like proteins). RNA/protein hybrid structures are involved in protein synthesis (ribosome). Recently it has been found that very small RNA molecules are involves in gene regulation.

RNA role in Protein Synthesis

1) Messenger RNA (mRNA)

Messenger RNA (mRNA) carries the genetic information copied from DNA in the form of a series of three-base code “words,” each of which specifies a particular amino acid.

2) Transfer RNA (tRNA)

Transfer RNA (tRNA**)** is the key to deciphering the code words in mRNA. Each type of amino acid has its own type of tRNA, which binds it and carries it to the growing end of a polypeptide chain if the next code word on mRNA calls for it. The correct tRNA with its attached amino acid is selected at each step because each specific tRNA molecule contains a three-base sequence that cans base-pair with its complementary code word in the mRNA.

3) Ribosomal RNA (rRNA)

Ribosomal RNA (rRNA) associates with a set of proteins to form ribosomes. These complex structures, which physically move along an mRNA molecule, catalyze the assembly of amino acids into protein chains. They also bind tRNAs and various accessory molecules necessary for protein synthesis. Ribosomes are composed of a large and small subunit, each of which contains its own rRNA molecule

RNA vs DNA

1) RNA contains the sugar ribose.	**1)** DNA contains deoxyribose.
2) RNA contains the base uracil.	**2)** DNA contains thymine instead.
3) RNA is usually single stranded.	**3)** DNA is usually double stranded.
4) RNA is short: one gene long at most	**4)** DNA is long, containing many genes.

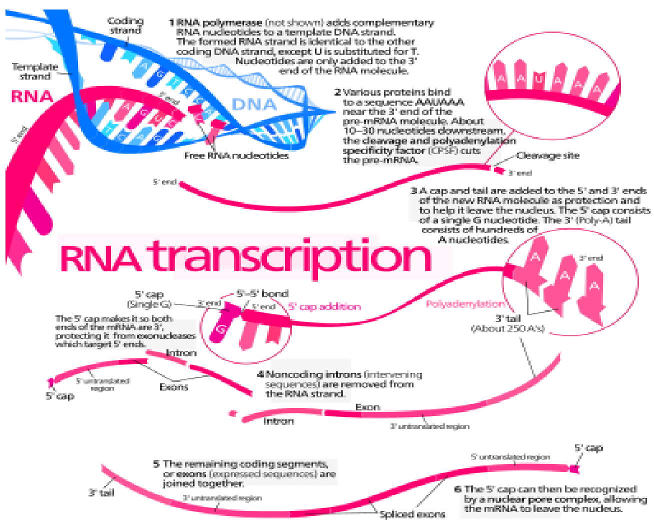
Transcription

Transcription is the first step of gene expression, in which a particular segment of DNA is copied into RNA (mRNA) by the enzyme RNA polymerase.

Both RNA and DNA are nucleic acids, which use base pairs of nucleotides as a complementary language. The two can be converted back and forth from DNA to RNA by the action of the correct enzymes. During transcription, a DNA sequence is read by an RNA polymerase, which produces a complementary, antiparallel RNA strand called a primary transcript.

Transcription proceeds in the following general steps.

1. One or more sigma factor protein binds to the RNA polymerase holoenzyme, allowing it to bind to promoter DNA.
2. RNA polymerase creates a transcription bubble, which separates the two strands of the DNA helix. This is done by breaking the hydrogen bonds between complementary DNA nucleotides.
3. RNA polymerase adds matching RNA nucleotides to the complementary nucleotides of one DNA strand.
4. RNA sugar-phosphate backbone forms with assistance from RNA polymerase to form an RNA strand.
5. Hydrogen bonds of the untwisted RNA–DNA helix break, freeing the newly synthesized RNA strand.
6. If the cell has a nucleus, the RNA may be further processed. This may include polyadenylation, capping, and splicing.
7. The RNA may remain in the nucleus or exit to the cytoplasm through the nuclear pore complex.



After Transcription

• In prokaryotes, translation starts before transcription is finished.

• In eukaryotes, the primary RNA transcript of a gene needs “RNA processing” before it can be translated.

• Also, it needs to be transported out of the nucleus into the cytoplasm.

• Steps in RNA processing:

1. Add a cap to the 5’ end
2. Add a poly-A tail to the 3’ end
3. Splice out introns.

5’ and 3’ Modification

• RNA is inherently unstable, especially at the ends.

• The ends are modified to protect it.

• At the 5’ end, a slightly modified guanine (7-methyl G) is attached.

• At the 3’ end, the primary transcript RNA is cut at a specific site.

• 100-200 adenines are attached to 3'.
• These A's are not coded in the DNA of the gene.

Introns

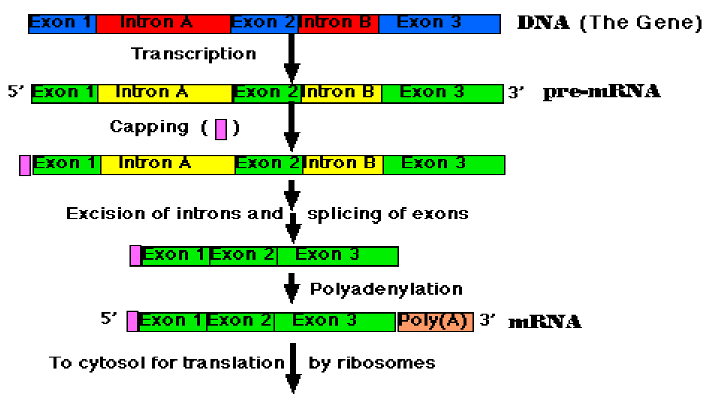
In some genes, not all of the DNA sequence is used to make protein. Introns are noncoding sections of an RNA transcript, or the DNA encoding it, that are spliced out before the RNA molecule is translated into a protein. The sections of DNA (or RNA) that code for proteins are called exons. Following transcription, new, immature strands of messenger RNA, called pre-mRNA, may contain both introns and exons.

Intron Splicing

RNA sequences between exons that are removed by splicing. Introns have been found in eukaryotic mRNA, tRNA and rRNA, as well as chloroplast, mitochondrial and a phage of *E. coli*. Eubacteria are the only species in which introns have not been found. In general, genes that are related by evolution have related organizations with conservation of the position at least some introns.

Summary of RNA processing

• RNA polymerase produces a primary transcript.
• A cap is put on the 5' end.
• The RNA is terminated and poly-A is added to the 3' end.
• All introns are spliced out to messenger RNA.
• It is then transported out of the nucleus
• In the cytoplasm, it is translated.



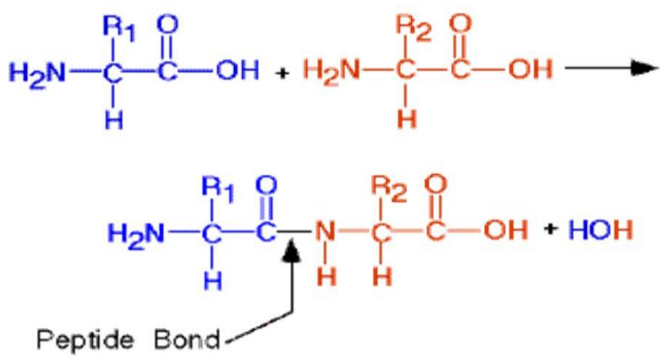
Proteins

Proteins are large biomolecules, or macromolecules, consisting of one or more long chains of amino acid residues. A linear chain of amino acid residues is called a polypeptide. A protein contains at least one long polypeptide. Short polypeptides, containing less than 20-30 residues, are rarely considered to be proteins and are commonly called peptides. The individual amino acid residues are bonded together by peptide bonds and adjacent amino acid residues.

Amino Acids and Peptide Bonds

• There are 20 different amino acids coded in DNA.
• They have an (-NH2) group on one end, and (-COOH) group on the other end.
• They attached to the central carbon is an R group.
• Two amino acids attach to each other by forming a peptide bond.

Peptide Bond Formation



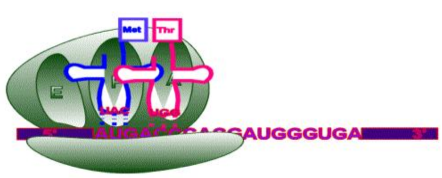
Amino Acids Classification
Translation

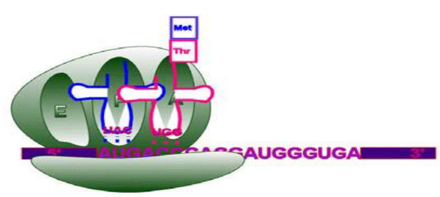
Translation is the process in which cellular ribosomes create proteins. In translation, messenger RNA (mRNA) produced by transcription from DNA is decoded by a ribosome to produce a specific amino acid chain, or polypeptide. The polypeptide later folds into an active protein and performs its functions in the cell. The ribosome facilitates decoding by inducing the binding of complementary tRNA anticodon sequences to mRNA codons. The tRNAs carry specific amino acids that are chained together into a polypeptide as the mRNA passes through and is "read" by the ribosome. The entire process is a part of gene expression.

Translation proceeds in three phases:

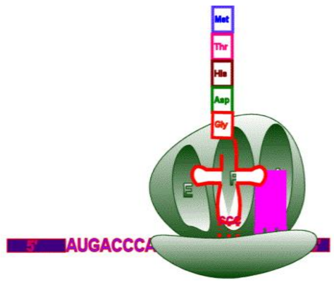
1) Initiation: The ribosome assembles around the target mRNA. The first tRNA is attached at the start codon.

2) Elongation: The tRNA transfers an amino acid to the tRNA corresponding to the next codon. The ribosome then moves (*translocates)* to the next mRNA codon to continue the process, creating an amino acid chain.





3) Termination: When a stop codon is reached, the ribosome releases the polypeptide.



The Genetic Code

• Each group of 3 nucleotides on the mRNA is a codon.

• Since there are 4 bases, there are 43 = 64 possible codons, which must code for 20 different amino acids.

• More than one codon is used for most amino acids: the genetic code is degenerate.

• AUG is used as the start codon.

• All proteins are initially translated with methionine in the first position.

• Proteins end in a stop codon, which codes for no amino acid.

Eg Stop codons are UGA, UAA and UAG

Codon table

Second letter

First letter	U		C		A		G		Third letter
U	UUU	Phenyl-alanine	UCU	Serine	UAU	Tyrosine	UGU	Cysteine	U
	UUC		UCC		UAC		UGC		C
	UUA	Leucine	UCA		UAA	Stop codon	UGA	Stop codon	A
	UUG		UCG		UAG	Stop codon	UGG	Tryptophan	G
C	CUU	Leucine	CCU	Proline	CAU	Histidine	CGU	Arginine	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	Glutamine	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	Isoleucine	ACU	Threonine	AAU	Asparagine	AGU	Serine	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	Lysine	AGA	Arginine	A
	AUG	Methionine; start codon	ACG		AAG		AGG		G
G	GUU	Valine	GCU	Alanine	GAU	Aspartate	GGU	Glycine	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	Glutamate	GGA		A
	GUG		GCG		GAG		GGG		G

Ref:

http://www.yourgenome.org/facts/what-is-the-central-dogma
https://www.neb.com/applications/cloning-and-synthetic-biology/dna-preparation/reversetranscription-cdna-synthesis
http://www.news-medical.net/life-sciences/What-is-RNA.aspx
http://www.ncbi.nlm.nih.gov/books/NBK216/
https://en.wikipedia.org/wiki/Transcription_%28genetics%29
http://www.nature.com/scitable/definition/intron-introns-67
https://www.ndsu.edu/pubweb/~mcclean/plsc731/transcript/transcript4.htm
https://en.wikipedia.org/wiki/Protein
https://en.wikipedia.org/wiki/Translation_%28biology%29