BME 1532-CELL BIOLOGY

Transcription and Translation

by Assist. Prof. Görke Gürel Peközer

Yıldız Technical University
Biomedical Engineering Department
Spring 2020

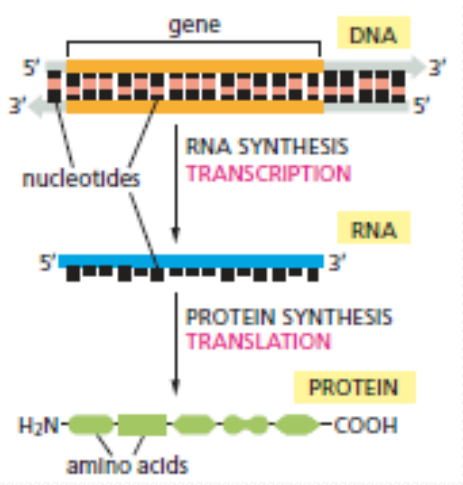
Last Week on BME 1532

- Discovery of the DNA as the Hereditary Material
- DNA Structure
  - Double Helical Conformation of DNA
  - Genome, Gene, Karyotype, Complexity Concepts
- Higher Level DNA Packaging
  - Chromatin
    - Nucleosomes, Histone Proteins
  - Chromosomes
  - Heterochromatin-Euchromatin
- DNA Replication
  - Semiconservative Nature of DNA Replication
  - Replication Forks, Okazaki Fragmens, Primers
  - DNA proofreading
- DNA Repair
  - Nucleotide Excission Repair
  - Mismatch Repair
  - Nonhomologous End Joining

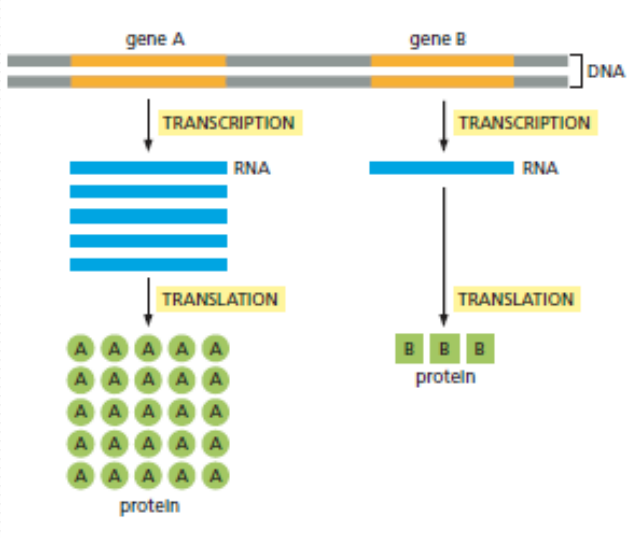
Decoding the Genetic Instructions in DNA

- Once the double-helical structure of DNA had been determined in the early 1950s, it became clear that the hereditary information in cells is encoded in the linear order—or *sequence*—of the four different nucleotide subunits that make up the DNA.
- Genetic instructions written in an alphabet of just four "letters" direct the formation of the simplest organisms to the most complex ones.
- Cells decode and use the information contained in the genes to direct the synthesis of proteins.
- Proteins are the principal constituents of cells and determine not only cell structure but also cell function.
- The properties and function of a protein molecule are determined by the sequence of the 20 different amino acid subunits in its polypeptide chain: each type of protein has its own unique amino acid sequence, which dictates how the chain will fold to form a molecule with a distinctive shape and chemistry.
- The genetic instructions carried by DNA must therefore specify the amino acid sequences of proteins.

- DNA does not synthesize proteins itself, but it acts like a manager, delegating the various tasks to a team of workers.
- When a particular protein is needed by the cell, the nucleotide sequence of the appropriate segment of a DNA molecule is first copied into another type of nucleic acid—*RNA* (*ribonucleic acid*). That segment of DNA is called a gene, and the resulting RNA copies are then used to direct the synthesis of the protein.
- The flow of genetic information in cells is therefore from DNA to RNA to protein.
- All cells, from bacteria to humans, express their genetic information in this way—a principle so fundamental that it has been termed ***the central dogma*** of molecular biology.
- The mechanism by which cells copy DNA into RNA is called ***transcription*** and use the information in RNA to make protein is called ***translation***.



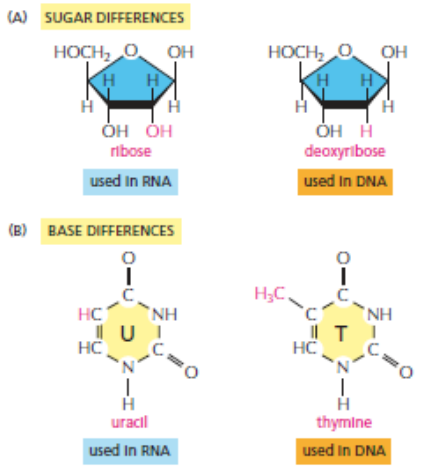
- Transcription and translation are the means by which cells read out, or *express*, the instructions in their *genes*. Many identical RNA copies can be made from the same gene, and each RNA molecule can direct the synthesis of many identical protein molecules.
- This successive amplification enables cells to rapidly synthesize large amounts of protein whenever necessary. At the same time, each gene can be transcribed, and its RNA translated, at different rates, providing the cell with a way to make vast quantities of some proteins and tiny quantities of others.
- A cell can also regulate the expression of each of its genes according to the needs of the moment through regulation of ***gene expression.***

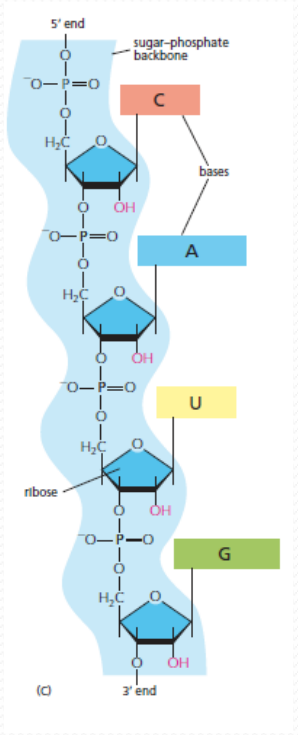


- The first step a cell takes in expressing one of its many thousands of genes is to copy the nucleotide sequence of that gene into RNA.
- This process is called ***transcription*** because the information, though copied into another chemical form, is still written in essentially the same language— the language of nucleotides.
- Like DNA, RNA is a linear polymer made of four different nucleotide subunits, linked together by phosphodiester bonds.
- It differs from DNA chemically in two respects:

1. The nucleotides in RNA are *ribonucleotides*—that is, they contain the sugar ribose (hence the name *ribo*nucleic acid) rather than deoxyribose;
2. Although, like DNA, RNA contains the bases adenine (A), guanine (G), and cytosine (C), it contains uracil (U) instead of the thymine (T) found in DNA.

- Because U, like T, can base-pair by hydrogen-bonding with A, the complementary base-pairing properties of DNA apply also to RNA.

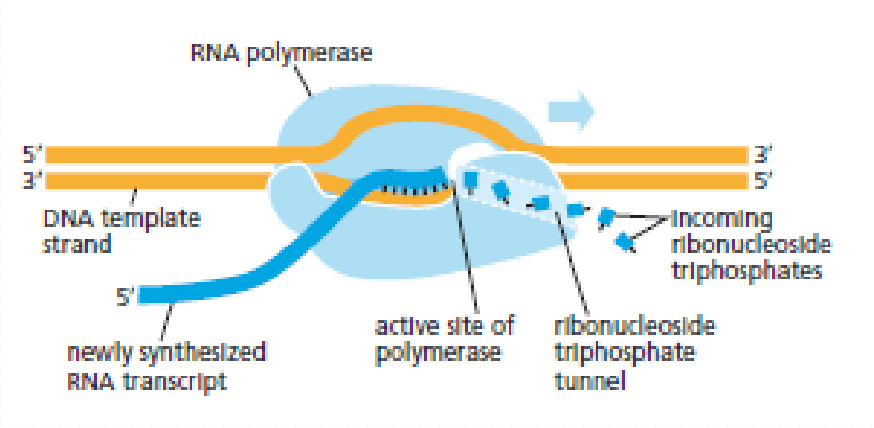




- Although their chemical differences are small, DNA and RNA differ quite dramatically in overall structure.
- DNA always occurs in cells as a double-stranded helix, weheras RNA is single-stranded.
- This difference has important functional consequences:
  - Because an RNA chain is single stranded, it can fold up into a variety of shapes, just as a polypeptide chain folds up to form the final shape of a protein; double stranded DNA cannot fold in this fashion.
  - The ability to fold into a complex three-dimensional shape allows RNA to carry out various functions in cells, in addition to conveying information between DNA and protein. Whereas DNA functions solely as an information store, some RNAs have structural, regulatory, or catalytic roles.

- All the RNA in a cell is made by transcription, a process that has certain similarities to DNA replication .
- Transcription begins with the opening and unwinding of a small portion of the DNA double helix to expose the bases on each DNA strand.
- One of the two strands of the DNA double helix then acts as a template for the synthesis of RNA. Ribonucleotides are added, one by one, to the growing RNA chain; as in DNA replication, the nucleotide sequence of the RNA chain is determined by complementary base-pairing with the DNA template.
- When a good match is made, the incoming ribonucleotide is covalently linked to the growing RNA chain by the enzyme ***RNA polymerase***.
- The RNA chain produced by transcription—the RNA transcript—is therefore elongated one nucleotide at a time and has a nucleotide sequence exactly complementary to the strand of DNA used as the template.
- Transcription differs from DNA replication in several crucial respects:
  - Unlike a newly formed DNA strand, the RNA strand does not remain hydrogen-bonded to the DNA template strand. Instead, just behind the region where the ribonucleotides are being added, the RNA chain is displaced and the DNA helix re-forms.
  - For this reason—and because only one strand of the DNA molecule is transcribed—RNA molecules are single-stranded.
  - Further, because RNAs are copied from only a limited region of DNA, RNA molecules are much shorter than DNA molecules.

- Like the DNA polymerase that carries out DNA replication, RNA polymerases catalyze the formation of the phosphodiester bonds that link the nucleotides together and form the sugar–phosphate backbone of the RNA chain.
- The RNA polymerase moves stepwise along the DNA, unwinding the DNA helix just ahead to expose a new region of the template strand for complementary base-pairing.
- In this way, the growing RNA chain is extended by one nucleotide at a time in the 5′-to-3′ direction.
- The incoming ribonucleoside triphosphates (ATP, CTP, UTP, and GTP) provide the energy needed to drive the reaction forward.



- Although RNA polymerase catalyzes essentially the same chemical reaction as DNA polymerase, there are some important differences between the two enzymes.
  - First, and most obviously, RNA polymerase uses ribonucleoside for phosphates as substrates, so it catalyzes the linkage of ribonucleotides, not deoxyribonucleotides.
  - Second, unlike the DNA polymerase involved in DNA replication, RNA polymerases can start an RNA chain without a primer.
- This difference likely evolved because transcription need not be as accurate as DNA replication; unlike DNA, RNA is not used as the permanent storage form of genetic information in cells, so mistakes in RNA transcripts have relatively minor consequences for a cell.
- RNA polymerases make about one mistake for every 10^4 nucleotides copied into RNA, whereas DNA polymerase makes only one mistake for every 10^7 nucleotides copied.

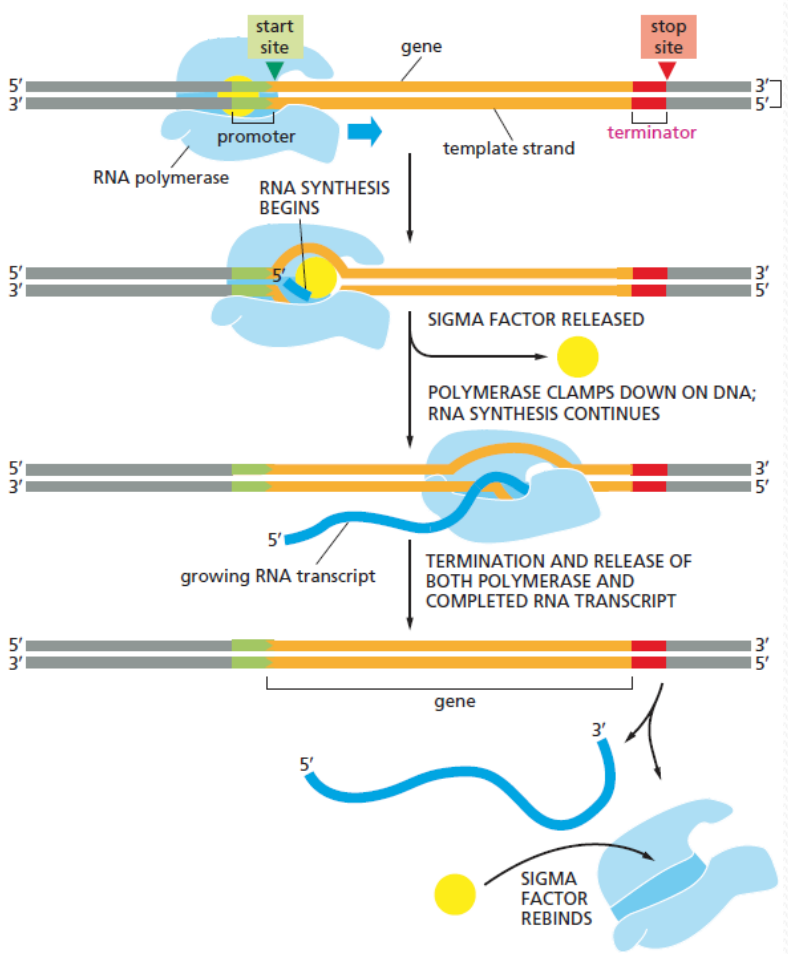
- The vast majority of genes carried in a cell's DNA specify the amino acid sequences of proteins.
- The RNA molecules encoded by these genes—which ultimately direct the synthesis of proteins—are called messenger RNAs (mRNAs).
- The final product of other genes, however, is the RNA itself. These nonmessenger RNAs, like proteins, have various roles, serving as regulatory, structural, and catalytic components of cells.
- They play key parts, for example, in translating the genetic message into protein: *ribosomal RNAs* (*rRNAs*) form the structural and catalytic core of the ribosomes, which translate mRNAs into protein, and *transfer RNAs* (*tRNAs*) act as adaptors that select specific amino acids and hold them in place on a ribosome for their incorporation into protein. Other small RNAs, called *microRNAs* (*miRNAs*), serve as key regulators of eukaryotic gene expression

Type of RNA	Function
messenger RNAs (mRNAs)	code for proteins
ribosomal RNAs (rRNAs)	form the core of the ribosome's structure and catalyze protein synthesis
microRNAs (miRNAs)	regulate gene expression
transfer RNAs (tRNAs)	serve as adaptors between mRNA and amino acids during protein synthesis
other noncoding RNAs	used in RNA splicing, gene regulation, telomere maintenance, and many other processes

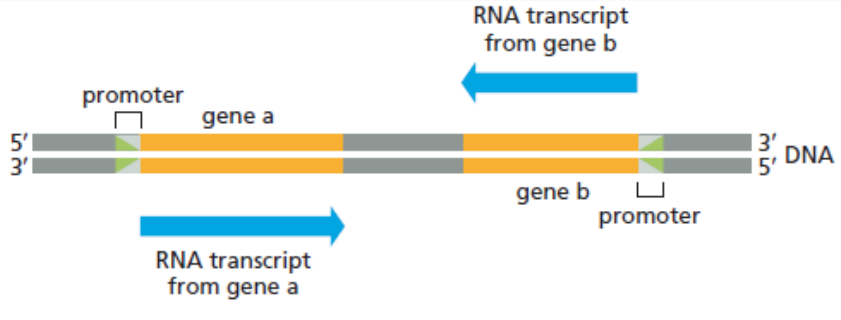
Prokaryotic Transcription

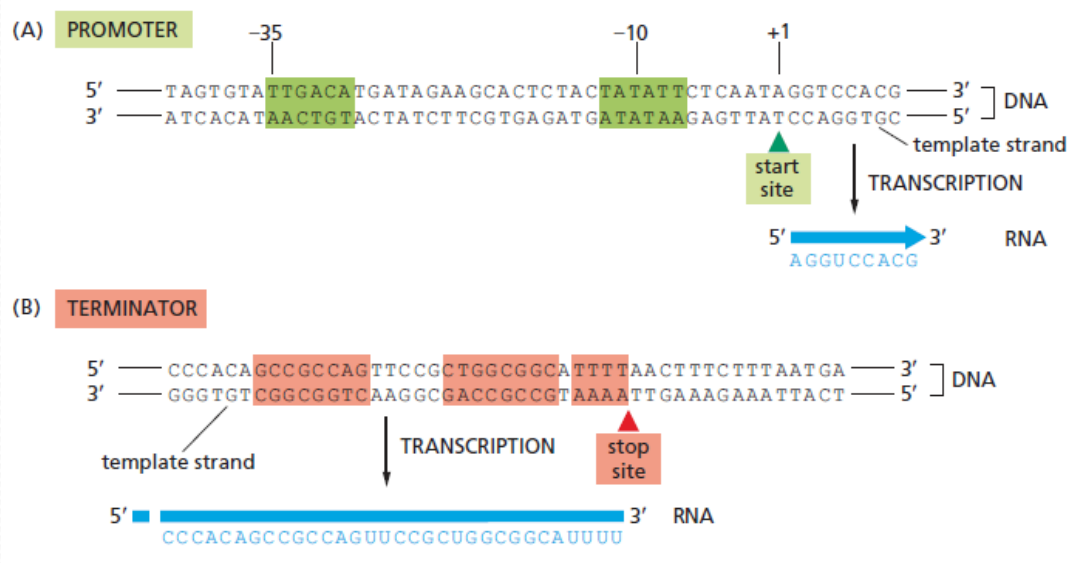
- The initiation of transcription is an especially critical process because it is the main point at which the cell selects which proteins or RNAs are to be produced. To begin transcription, RNA polymerase must be able to recognize the start of a gene and bind firmly to the DNA at this site.
- The way in which RNA polymerases recognize the *transcription start site* of a gene differs somewhat between bacteria and eukaryotes.
- However, both require a specific sequence of nucleotides called ***promoter*** that lies immediately upstream of the starting point for RNA synthesis.
- Bacterial trancription is simpler than eukaryotic transcription.

- When a bacterial RNA polymerase collides randomly with a DNA molecule, the enzyme sticks weakly to the double helix and then slides rapidly along its length. RNA polymerase latches on tightly only after it has encountered the ***promoter***.
- Once bound tightly to this sequence, the RNA polymerase opens up the double helix immediately in front of the promoter to expose the nucleotides on each strand of a short stretch of DNA.
- One of the two exposed DNA strands then acts as a template for complementary base pairing with incoming ribonucleoside triphosphates, two of which are joined together by the polymerase to begin synthesis of the RNA chain.
- Chain elongation then continues until the enzyme encounters a second signal in the DNA, the ***terminator*** (or stop site), where the polymerase halts and releases both the DNA template and the newly made RNA transcript.
- This terminator sequence is contained within the gene and is transcribed into the 3′ end of the newly made RNA.



- In bacteria, it is a subunit of RNA polymerase, the *sigma* (σ) *factor,* that is primarily responsible for recognizing the promoter sequence on the DNA.
- Each base presents unique features to the outside of the double helix, allowing the sigma factor to find the promoter sequence without having to separate the DNA strands.
- The next problem an RNA polymerase faces is determining which of the two DNA strands to use as a template for transcription: each strand has a different nucleotide sequence and would produce a different RNA transcript.
- The secret lies in the structure of the promoter itself. Every promoter has a certain polarity: it contains two different nucleotide sequences upstream of the transcriptional start site that position the RNA polymerase, ensuring that it binds to the promoter in only one orientation.
- Because the polymerase can only synthesize RNA in the 5′-to-3′ direction once the enzyme is bound it must use the DNA strand oriented in the 3′-to-5′ direction as its template.
- This selection of a template strand does not mean that on a given chromosome, transcription always proceeds in the same direction. With respect to the chromosome as a whole, the direction of transcription varies from gene to gene.
- But because each gene typically has only one promoter, the orientation of its promoter determines in which direction that gene is transcribed and therefore which strand is the template strand.





Bacterial promoters and terminators have specific nucleotide sequences that are recognized by RNA polymerase.

(A) The *green*-shaded region represent the nucleotide sequences that specify a promoter. The numbers above the DNA indicate the positions of nucleotides counting from the first nucleotide transcribed, which is designated +1. The polarity of the promoter orients the polymerase and determines which DNA strand is transcribed.
All bacterial promoters contain DNA sequences at –10 and –35 that closely resemble those shown here. (B) The *red*-shaded regions represent sequences in the gene that signal the RNA polymerase to terminate transcription.
Note that the regions transcribed into RNA contain the terminator but not the promoter nucleotide sequences. By convention, the sequence of a gene is that of the non-template strand, as this strand has the same sequence as the transcribed RNA (with T substituting for U).

Eukaryotic Transcription

- Many of the principles we just outlined for bacterial transcription also apply to eukaryotes. However, transcription initiation in eukaryotes differs in several important ways from that in bacteria:
- The first difference lies in the RNA polymerases themselves. While bacteria contain a single type of RNA polymerase, eukaryotic cells have three—*RNA polymerase I*, *RNA polymerase II*, and *RNA polymerase III*. These polymerases are responsible for transcribing different types of genes.
  - RNA polymerases I and III transcribe the genes encoding transfer RNA, ribosomal RNA, and various other RNAs that play structural and catalytic roles in the cell.
  - RNA polymerase II transcribes the vast majority of eukaryotic genes, including all those that encode proteins and miRNAs.

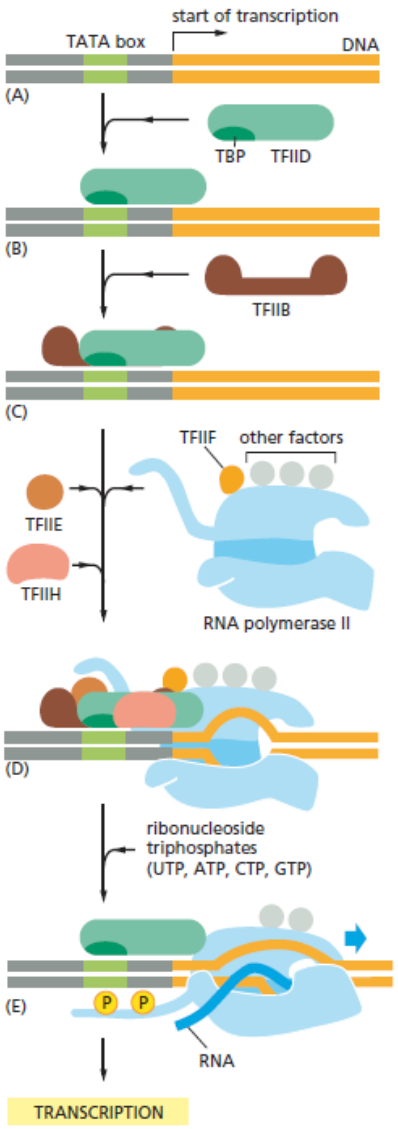
Type of Polymerase	Genes Transcribed
RNA polymerase I	most rRNA genes
RNA polymerase II	all protein-coding genes, miRNA genes, plus genes for other noncoding RNAs (e.g., those in spliceosomes)
RNA polymerase III	tRNA genes 5S rRNA gene genes for many other small RNAs

- A second difference is that, whereas the bacterial RNA polymerase (along with its sigma subunit) is able to initiate transcription on its own, eukaryotic RNA polymerases require the assistance of a large set of accessory proteins. Principal among these are the *general transcription factors*, which must assemble at each promoter, along with the polymerase, before the polymerase can begin transcription.
- A third distinctive feature of transcription in eukaryotes is that the mechanisms that control its initiation are much more elaborate than those in prokaryotes.
- In bacteria, genes tend to lie very close to one another in the DNA, with only very short lengths of nontranscribed DNA between them. But in plants and animals, including humans, individual genes are spread out along the DNA, with stretches of up to 100,000 nucleotide pairs between one gene and the next.
- This architecture allows a single gene to be controlled by a large variety of ***regulatory DNA sequences*** scattered along the DNA, and it enables eukaryotes to engage in more complex forms of transcriptional regulation than do bacteria.
- Last but not least, eukaryotic transcription initiation must take into account the packing of DNA into *nucleosomes* and more compact forms of chromatin structure.

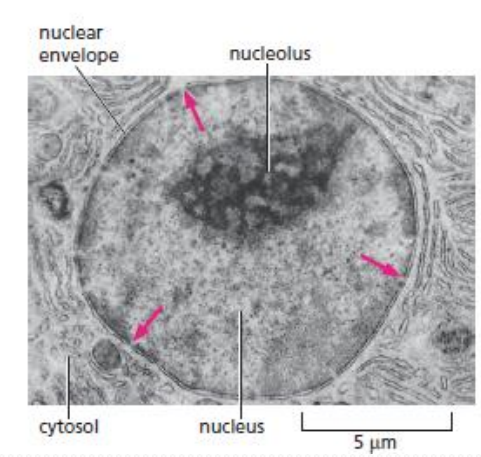
Eukaryotic Transcription

- ***RNA polymerase II*** cannot initiate transcription on its own. It needs ***general transcription factors.***
- These accessory proteins assemble on the promoter, where they position the RNA polymerase and pull apart the DNA double helix to expose the template strand, allowing the polymerase to begin transcription.
- Thus the general transcription factors have a similar role in eukaryotic transcription as sigma factor has in bacterial transcription.

- The assembly of general trancription factors at the promoter process typically begins with the binding of the general transcription factor TFIID to a short segment of DNA double helix composed primarily of T and A nucleotides; because of its composition, this part of the promoter is known as the *TATA box*.
- Upon binding to DNA, TFIID causes a dramatic local distortion in the DNA double helix, which helps to serve as a landmark for the subsequent assembly of other proteins at the promoter.
- The TATA box is a key component of many promoters used by RNA polymerase II, and it is typically located 25 nucleotides upstream from the transcription start site.
- Once TFIID has bound to the TATA box, the other factors assemble, along with RNA polymerase II, to form a complete ***transcription initiation complex.***
- After RNA polymerase II has been positioned on the promoter, it must be released from the complex of general transcription factors to begin its task of making an RNA molecule.
- A key step in liberating the RNA polymerase is the addition of phosphate groups to its "tail". This liberation is initiated by the general transcription factor TFIIH, which contains a protein kinase as one of its subunits.
- Once transcription has begun, most of the general transcription factors dissociate from the DNA and then are available to initiate another round of transcription with a new RNA polymerase molecule.
- When RNA polymerase II finishes transcribing a gene, it too is released from the DNA; the phosphates on its tail are stripped off by protein phosphatases, and the polymerase is then ready to find a new promoter. Only the dephosphorylated form of RNA polymerase II can initiate RNA synthesis.



- Although the templating principle by which DNA is transcribed into RNA is the same in all organisms, the way in which the RNA transcripts are handled before they can be used by the cell to make protein differs greatly between bacteria and eukaryotes.
- Bacterial DNA lies directly exposed to the cytoplasm, which contains the *ribosomes* on which protein synthesis takes place. As an mRNA molecule in a bacterium starts to be synthesized, ribosomes immediately attach to the free 5′ end of the RNA transcript and begin translating it into protein. In eukaryotic cells, by contrast, DNA is enclosed within the *nucleus*.
- Transcription takes place in the nucleus, but protein synthesis takes place on ribosomes in the cytoplasm. So, before a eukaryotic mRNA can be translated into protein, it must be transported out of the nucleus through small pores in the nuclear envelope.



mRNA Processing

- Before it can be exported to the cytosol, however, a eukaryotic RNA must go through several RNA processing steps, which include ***capping, splicing, and polyadenylation***.
- These steps take place as the RNA is being synthesized. The enzymes responsible for RNA processing ride on the phosphorylated tail of eukaryotic RNA polymerase II as it synthesizes an RNA molecule, and they process the transcript as it emerges from the polymerase.

Capping and Polyadenylation

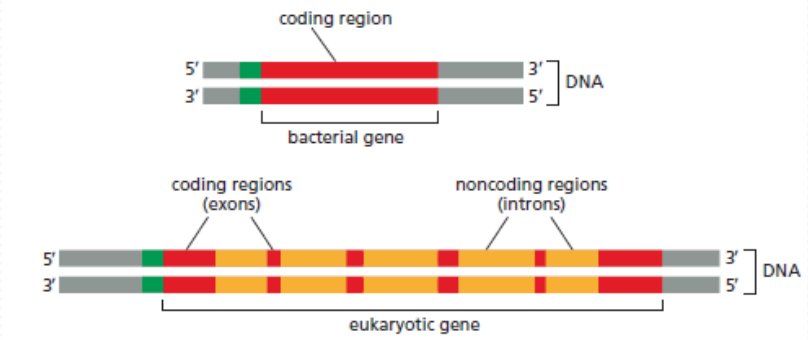
- Different types of RNA are processed in different ways before leaving the nucleus. Two processing steps, capping and polyadenylation, occur only on RNA transcripts destined to become mRNA molecules (called *precursor mRNAs*, or *pre-mRNAs*).

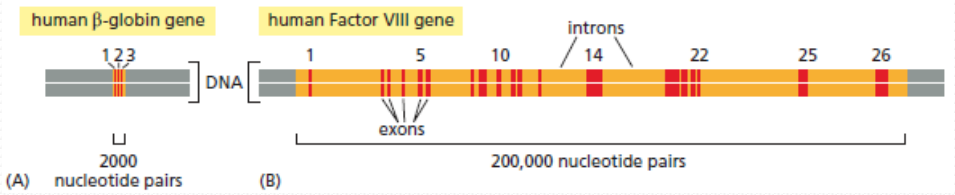
1. RNA capping modifies the 5′ end of the RNA transcript, the end that is synthesized first. The RNA is capped by the addition of an atypical nucleotide—a guanine (G) nucleotide bearing a methyl group, which is attached to the 5′ end of the RNA in an unusual way . This capping occurs after RNA polymerase II has produced about 25 nucleotides of RNA, long before it has completed transcribing the whole gene.
2. Polyadenylation provides a newly transcribed mRNA with a special structure at its 3′ end. In contrast with bacteria, where the 3′ end of an mRNA is simply the end of the chain synthesized by the RNA polymerase, the 3′ end of a forming eukaryotic mRNA is first trimmed by an enzyme that cuts the RNA chain at a particular sequence of nucleotides. The transcript is then finished off by a second enzyme that adds a series of repeated adenine (A) nucleotides to the cut end. This *poly-A tail* is generally a few hundred nucleotides long

- These two modifications—capping and polyadenylation—increase the stability of a eukaryotic mRNA molecule, facilitate its export from the nucleus to the cytoplasm, and generally mark the RNA molecule as an mRNA. They are also used by the protein-synthesis machinery to make sure that both ends of the mRNA are present and that the message is therefore complete before protein synthesis begins.

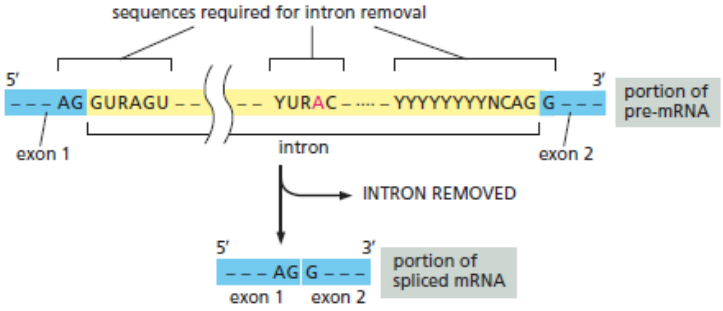
Introns

- Most eukaryotic pre-mRNAs have to undergo an additional processing step before they are functional mRNAs. This step involves a far more radical modification of the pre-mRNA transcript than capping or polyadenylation, and it is the consequence of a surprising feature of most eukaryotic genes.
- In bacteria, most proteins are encoded by an uninterrupted stretch of DNA sequence that is transcribed into an mRNA that, without any further processing, can be translated into protein.
- Most protein-coding eukaryotic genes, in contrast, have their coding sequences interrupted by long, noncoding, *intervening sequences* called ***introns.*** The scattered pieces of coding sequence—called ***exons***—are usually shorter than the introns, and they often represent only a small fraction of the total length of the gene.

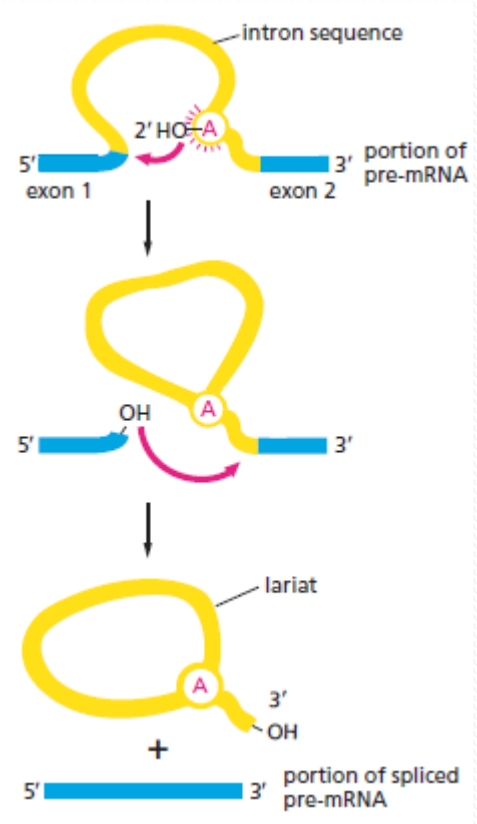




- To produce an mRNA in a eukaryotic cell, the entire length of the gene, introns as well as exons, is transcribed into RNA.
- After capping, and as RNA polymerase II continues to transcribe the gene, the process of RNA splicing begins, in which the introns are removed from the newly synthesized RNA and the exons are stitched together. Each transcript ultimately receives a poly-A tail.
- Once a transcript has been spliced and its 5′ and 3′ ends have been modified, the RNA is now a functional mRNA molecule that can leave the nucleus and be translated into protein.
- Most of the nucleotide sequence of an intron is unimportant. Each intron contains a few short nucleotide sequences that act as cues for its removal from the pre-mRNA. These special sequences are found at or near each end of the intron and are the same or very similar in all introns.
- Guided by these sequences, an elaborate splicing machine cuts out the intron in the form of a “lariat” structure formed by the reaction of the “A” nucleotide.



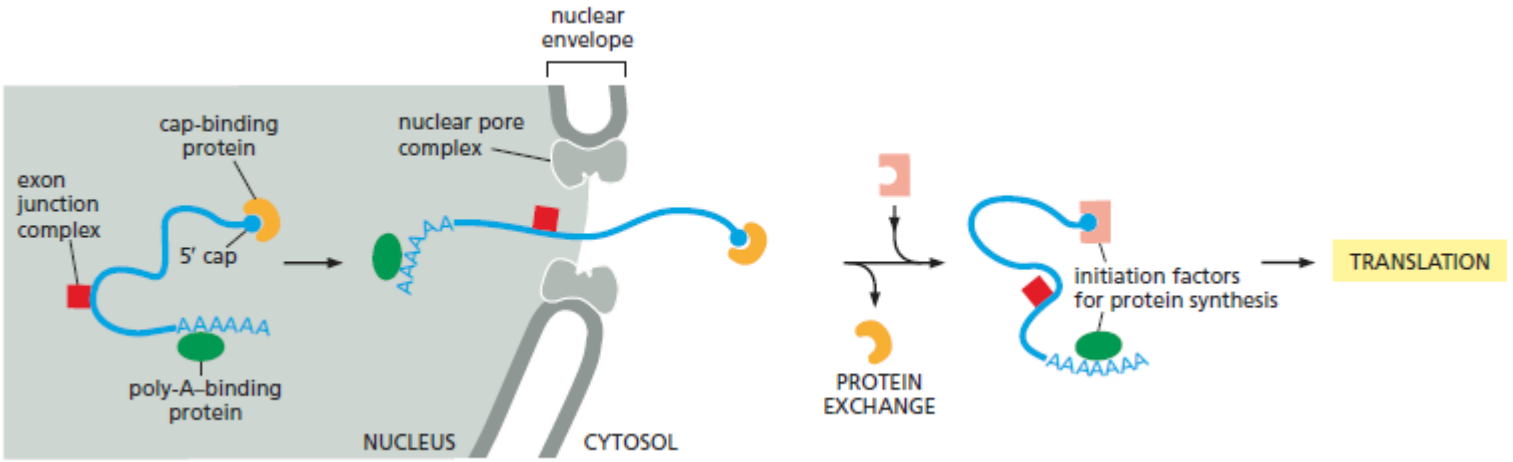
- RNA splicing is carried out largely by RNA molecules rather than proteins.
- These RNA molecules, called small nuclear RNAs (snRNAs), are packaged with additional proteins to form ***small nuclear ribonucleoproteins*** (*snRNPs*, pronounced "snurps").
- The snRNPs recognize splice-site sequences through complementary base-pairing between their RNA components and the sequences in the pre-mRNA, and they also participate intimately in the chemistry of splicing.
- Together, these snRNPs form the core of the ***spliceosome***, the large assembly of RNA and protein molecules that carries out RNA splicing in the nucleus.



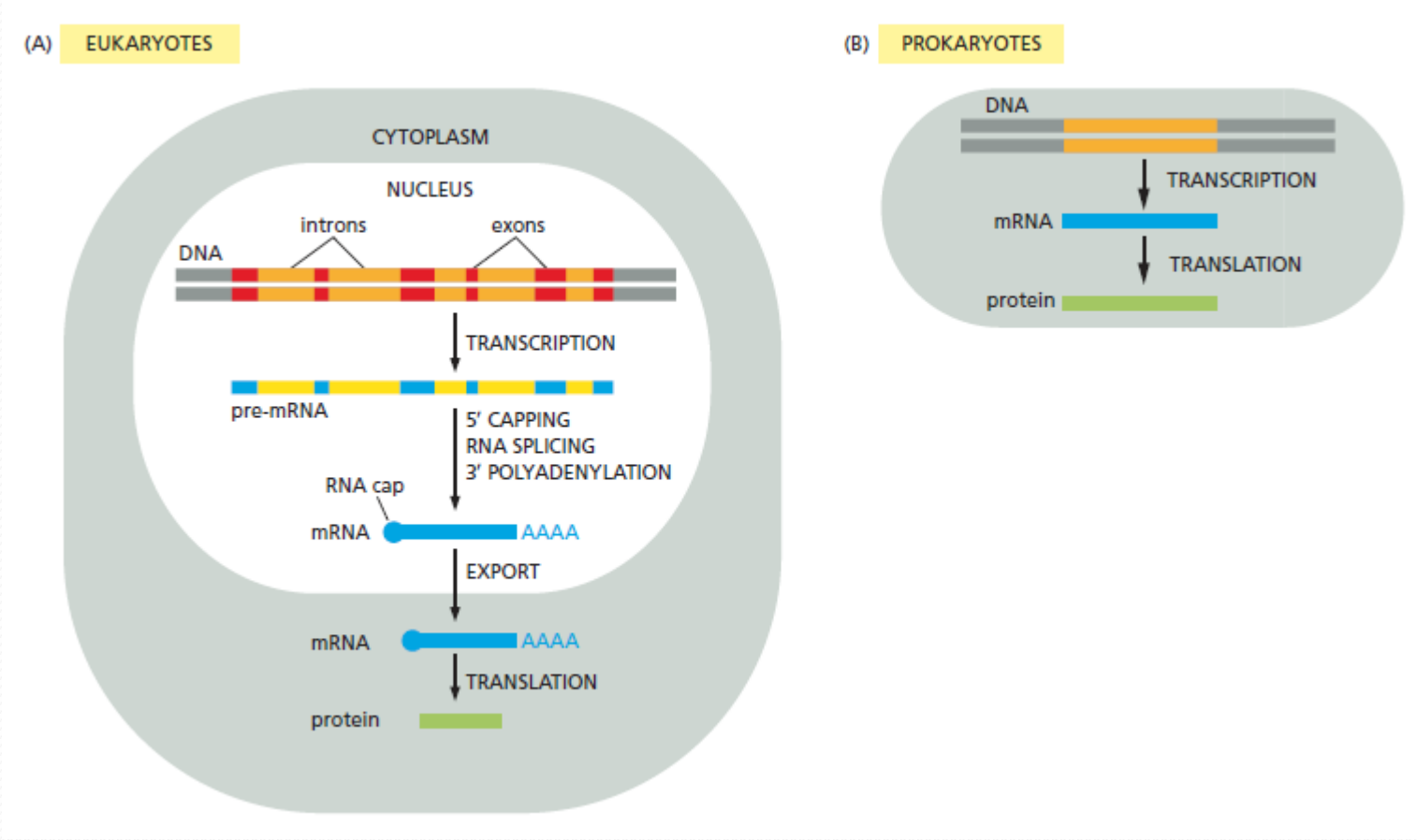
Alternative Splicing

- The intron–exon type of gene arrangement in eukaryotes may, at first, seem wasteful. It does, however, have a number of important benefits.
  - First, the transcripts of many eukaryotic genes can be spliced in different ways, each of which can produce a distinct protein. This process is called ***alternative splicing*** and it allows many different proteins to be produced from the same gene. About 95% of human genes are thought to undergo alternative splicing. Thus RNA splicing enables eukaryotes to increase the already enormous coding potential of their genomes.
  - RNA splicing also provides another advantage to eukaryotes, one that is likely to have been profoundly important in the early evolutionary history of genes: Novel proteins appear to have arisen by the mixing and matching of different exons of preexisting genes.

- Eukaryotic pre-mRNA synthesis and processing take place in an orderly fashion within the cell nucleus.
- Only correctly processed mature mRNA can be transported from the nucleus to the cytosol where it is translated.
- This selective transport is mediated by *nuclear pore complexes*, which connect the nucleoplasm with the cytosol and act as gates that control which macromolecules can enter or leave the nucleus.
- To be "export ready," an mRNA molecule must be bound to an appropriate set of proteins, each of which recognizes different parts of a mature mRNA molecule.
- These proteins include poly-A–binding proteins, a cap-binding complex, and proteins that bind to mRNAs that have been appropriately spliced.
- The entire set of bound proteins, rather than any single protein, ultimately determines whether an mRNA molecule will leave the nucleus.
- The "waste RNAs" that remain behind in the nucleus are degraded there, and their nucleotide building blocks are reused for transcription.



- Each mRNA molecule is eventually degraded into nucleotides by ribonucleases (RNAses) present in the cytosol, but the lifetimes of mRNA molecules differ considerably—depending on the nucleotide sequence of the mRNA and the type of cell.
- Because a single mRNA molecule can be translated into protein many times, the length of time that a mature mRNA molecule persists in the cell affects the amount of protein it produces.
- Different lifetimes are in part controlled by nucleotide sequences in the mRNA itself, most often in the portion of RNA called the *3′ untranslated region (UTR)*, which lies between the 3′ end of the coding sequence and the poly-A tail.
- The different lifetimes of mRNAs help the cell control the amount of each protein that it synthesizes.


(A) EUKARYOTES
CYTOPLASM
NUCLEUS
introns
exons
DNA
TRANSCRIPTION
pre-mRNA
5′ CAPPING
RNA SPLICING
3′ POLYADENYLATION
RNA cap
mRNA
AAAA
EXPORT
mRNA
AAAA
TRANSLATION
protein
(B) PROKARYOTES
DNA
TRANSCRIPTION
mRNA
TRANSLATION
protein

Translation

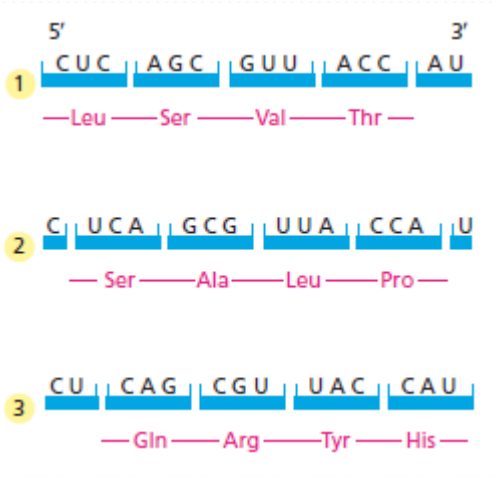
- Transcription as a means of information transfer is simple to understand: DNA and RNA are chemically and structurally similar, and DNA can act as a direct template for the synthesis of RNA through complementary base pairing. The language itself and the form of the message do not change, and the symbols used are closely related.
- However, the conversion of the information in RNA into protein represents a translation of the information into another language that uses different symbols.
- Because there are only 4 different nucleotides in mRNA but 20 different types of amino acids in a protein, this translation cannot be accounted for by a direct one-to-one correspondence between a nucleotide in RNA and an amino acid in protein.
- The rules by which the nucleotide sequence of a gene, through an intermediary mRNA molecule, is translated into the amino acid sequence of a protein are known as the ***genetic code.***

- The sequence of nucleotides in an mRNA molecule is read consecutively in groups of three.
- Each group of three consecutive nucleotides in RNA is called a ***codon***, and each codon specifies one amino acid.
- Because RNA is made of 4 different nucleotides, there are $4 \times 4 \times 4 = 64$ possible combinations of three nucleotides: AAA, AUA, AUG, and so on.
- However, only 20 different amino acids are commonly found in proteins.
- Thus, some amino acids are specified by more than one triplet.
- The same genetic code is used in nearly all present-day organisms. Although a few slight differences have been found, these occur chiefly in the mRNA of mitochondria and of some fungi and protozoa.

codons	GCA GCC GCG GCU	AGA AGG CGA CGC CGG CGU	GAC GAU	AAC AAU	UGC UGU	GAA GAG	CAA CAG	GGA GGC GGG GGU	CAC CAU	AUA AUC AUU	UUA UUG CUA CUC CUG CUU	AAA AAG	AUG	UUC UUU	CCA CCC CCG CCU	AGC AGU UCA UCC UCG UCU	ACA ACC ACG ACU	UGG	UAC UAU	GUA GUC GUG GUU	UAA UAG UGA
amino acids	Ala	Arg	Asp	Asn	Cys	Glu	Gln	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val	stop
	A	R	D	N	C	E	Q	G	H	I	L	K	M	F	P	S	T	W	Y	V	

Reading Frame

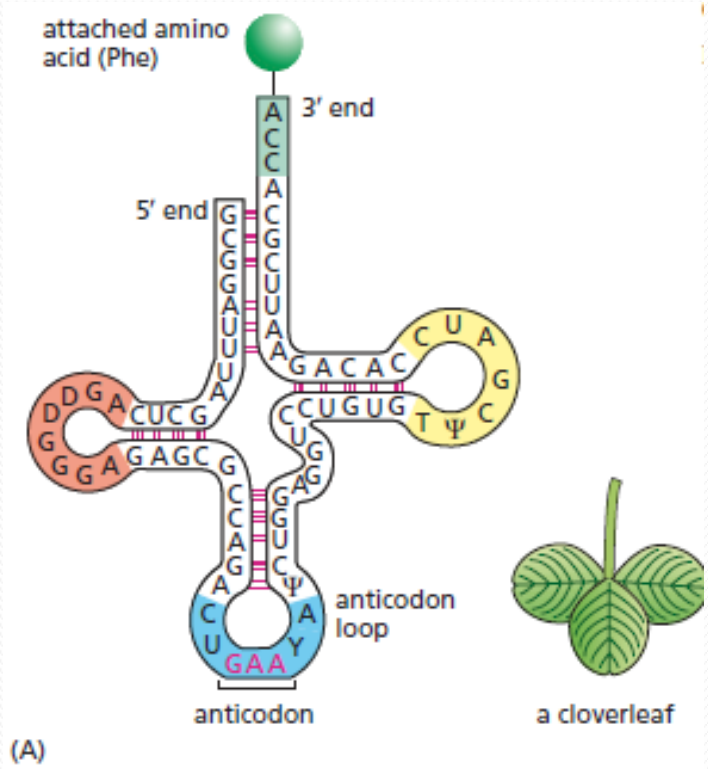
- In the process of translating a nucleotide sequence (*blue*) into an amino acid sequence (*red*), the sequence of nucleotides in an mRNA molecule is read from the 5′ to the 3′ end in sequential sets of three nucleotides.
- In principle, therefore, the same mRNA sequence can specify three completely different amino acid sequences, depending on where translation begins.
- Those three different translation patterns are called the **reading frames.**
- An mRNA sequence can be translated in any one of three different reading frames, depending on where the decoding process begins.
- However, only one of the three possible reading frames in an mRNA specifies the correct protein.



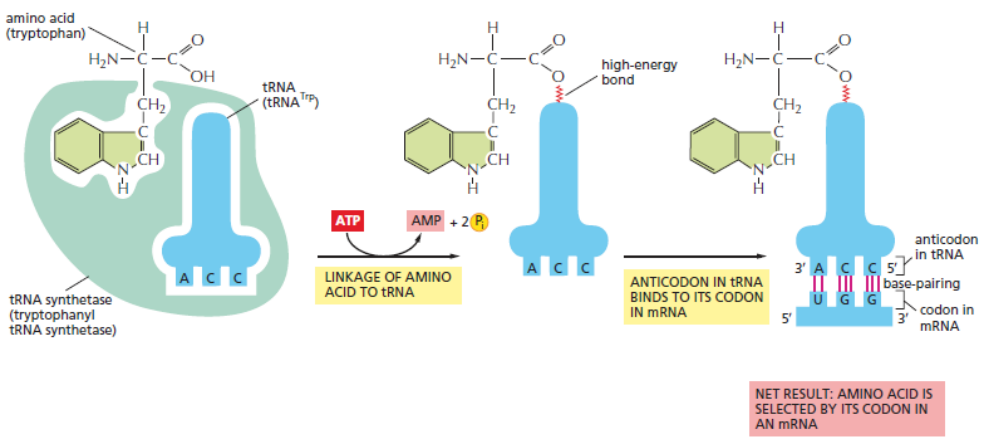
Transfer RNA

- The codons in an mRNA molecule do not directly recognize the amino acids they specify. Rather, the translation of mRNA into protein depends on adaptor molecules that can recognize and bind to a codon at one site on their surface and to an amino acid at another site.
- These adaptors consist of a set of small RNA molecules known as ***transfer RNAs (tRNAs)***, each about 80 nucleotides in length.
- The base paired regions of tRNA fold back on themselves to form secondary structures resembling a cloverleaf.

- Two regions of unpaired nucleotides situated at either end of the L-shaped tRNA molecule are crucial to the function of tRNAs in protein synthesis.
- One of these regions forms the ***anticodon***, a set of three consecutive nucleotides that bind, through base-pairing, to the complementary codon in an mRNA molecule.
- The other is a short single-stranded region at the 3′ end of the molecule; this is the site where the amino acid that matches the codon is covalently attached to the tRNA.

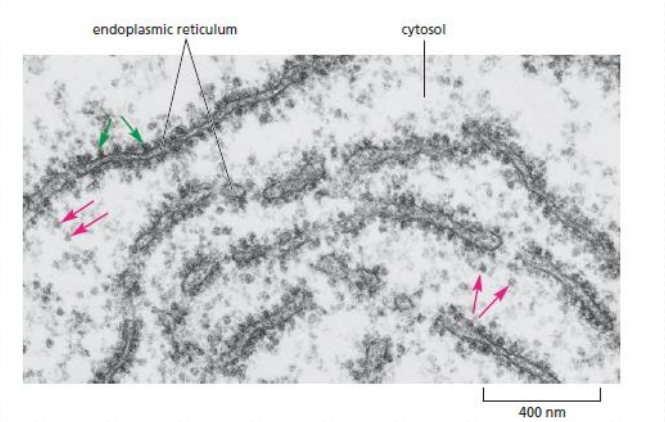


- For a tRNA molecule to carry out its role as an adaptor, it must be linked—or charged—with the correct amino acid.
- Recognition and attachment of the correct amino acid depend on enzymes called aminoacyl-tRNA synthetases, which covalently couple each amino acid to its appropriate set of tRNA molecules.
- There are 20 synthetases for 20 aminoacids: one attaches glycine to all tRNAs that recognize codons for glycine, another attaches phenylalanine to all tRNAs that recognize codons for phenylalanine, and so on.
- Each synthetase enzyme recognizes specific nucleotides in both the anticodon and the aminoacid- accepting arm of the correct tRNA.

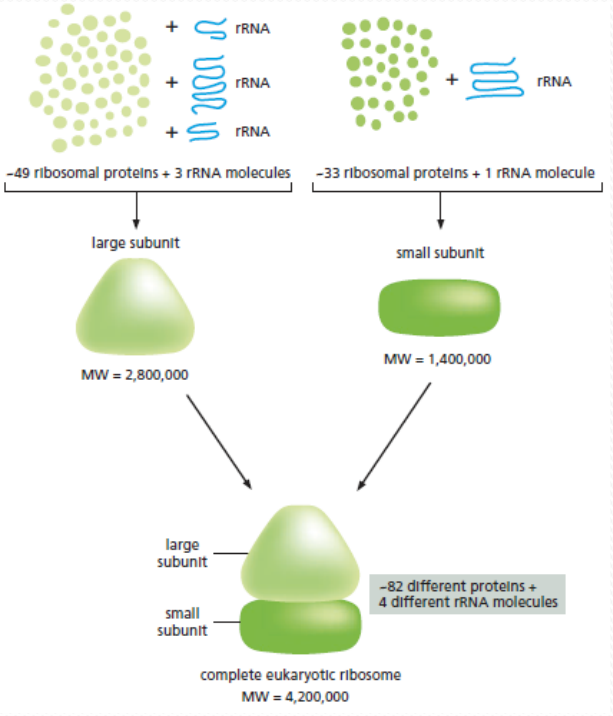


Ribosomes

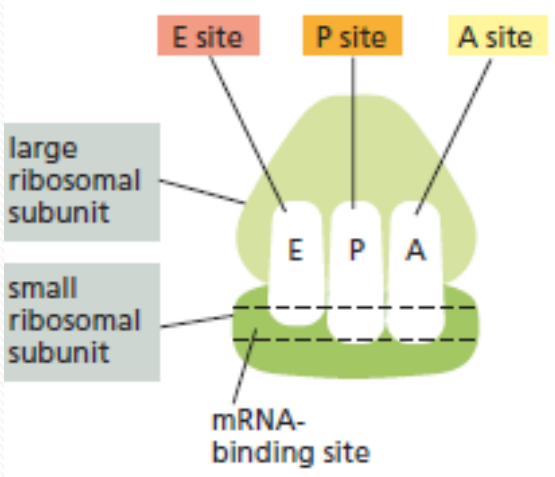
- The recognition of a codon by the anticodon on a tRNA molecule depends on the same type of complementary base-pairing used in DNA replication and transcription.
- However, accurate and rapid translation of mRNA into protein requires a molecular machine that can move along the mRNA, capture complementary tRNA molecules, hold the tRNAs in position, and then covalently link the amino acids that they carry to form a polypeptide chain.
- In both prokaryotes and eukaryotes, the machine that gets the job done is the ribosome—a large complex made from dozens of small proteins (the *ribosomal proteins*) and several crucial RNA molecules called ribosomal RNAs (rRNAs).
- A typical eukaryotic cell contains millions of ribosomes in its cytoplasm.



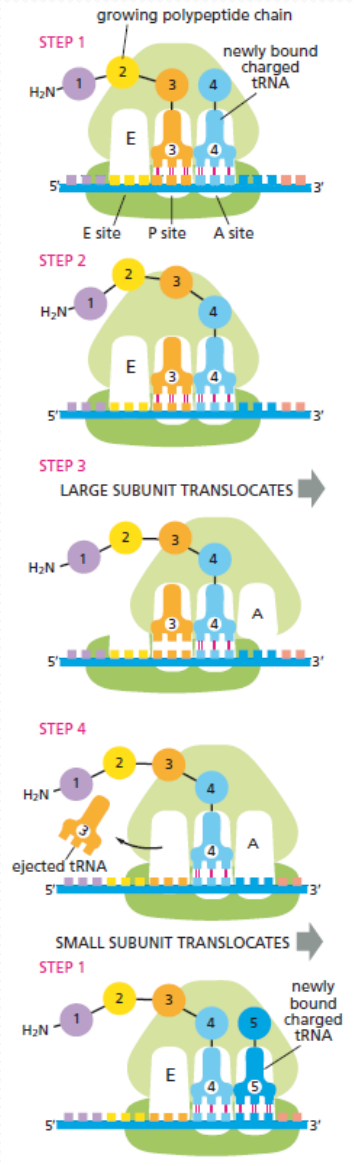
- Eukaryotic and prokaryotic ribosomes are very similar in structure and function.
- Both are composed of one large subunit and one small subunit, which fit together to form a complete ribosome.
- The small ribosomal subunit matches the tRNAs to the codons of the mRNA, while the large subunit catalyzes the formation of the peptide bonds that covalently link the amino acids together into a polypeptide chain.
- These two subunits come together on an mRNA molecule near its 5′ end to start the synthesis of a protein. The mRNA is then pulled through the ribosome. As the mRNA moves forward in a 5′-to-3′ direction, the ribosome translates its nucleotide sequence into an amino acid sequence, one codon at a time, using the tRNAs as adaptors.
- Each amino acid is thereby added in the correct sequence to the end of the growing polypeptide chain. When synthesis of the protein is finished, the two subunits of the ribosome separate.



- In addition to a binding site for an mRNA molecule, each ribosome contains three binding sites for tRNA molecules, called the A site, the P site, and the E site (short for aminoacyltRNA, peptidyl-tRNA, and exit, respectively).
- To add an amino acid to a growing peptide chain, the appropriate charged tRNA enters the A site by base-pairing with the complementary codon on the mRNA molecule.
- Its amino acid is then linked to the peptide chain held by the tRNA in the neighboring P site. Next, the large ribosomal subunit shifts forward, moving the spent tRNA to the E site before ejecting it.
- This cycle of reactions is repeated each time an amino acid is added to the polypeptide chain, with the new protein growing from its amino to its carboxyl end until a stop codon in the mRNA is encountered.



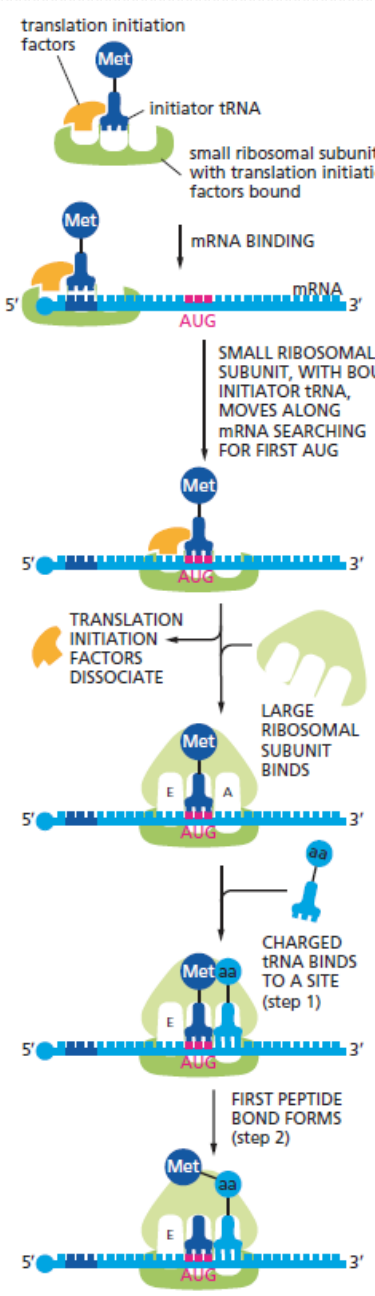
- In *step 1*, a charged tRNA carrying the next amino acid to be added to the polypeptide chain binds to the vacant A site on the ribosome by forming base pairs with the mRNA codon that is exposed there. Because only the appropriate tRNA molecules can base-pair with each codon, this codon determines the specific amino acid added.
- The A and P sites are sufficiently close together that their two tRNA molecules are forced to form base pairs with codons that are contiguous, with no stray bases in between. This positioning of the tRNAs ensures that the correct reading frame will be preserved throughout the synthesis of the protein.
- In *step 2*, the carboxyl end of the polypeptide chain (amino acid 3 in step 1) is uncoupled from the tRNA at the P site and joined by a peptide bond to the free amino group of the amino acid linked to the tRNA at the A site. This reaction is catalyzed by an enzymatic site in the large subunit.
- In *step 3*, a shift of the large subunit relative to the small subunit moves the two tRNAs into the E and P sites of the large subunit.
- In *step 4*, the small subunit moves exactly three nucleotides along the mRNA molecule, bringing it back to its original position relative to the large subunit.
- This movement ejects the spent tRNA and resets the ribosome with an empty A site so that the next charged tRNA molecule can bind.
- As indicated, the mRNA is translated in the 5′-to-3′ direction, and the N-terminal end of a protein is made first, with each cycle adding one amino acid to the C-terminus of the polypeptide chain.



Initiation of Translation

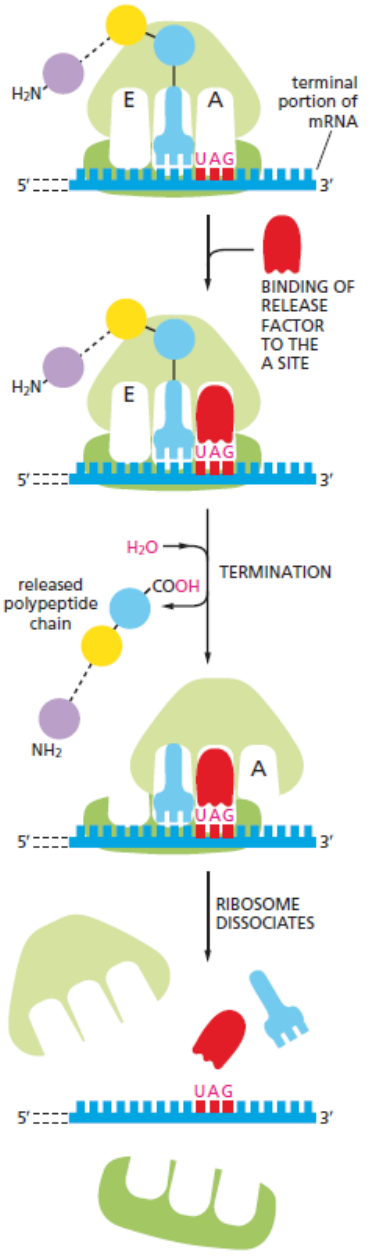
- A specific start signal is required to initiate translation.
- The site at which protein synthesis begins on an mRNA is crucial, because it sets the reading frame for the whole length of the message.
- An error of one nucleotide either way at this stage will cause every subsequent codon in the mRNA to be misread, resulting in a nonfunctional protein with an incorrect sequence of aminoacids.
- The translation of an mRNA begins with the codon AUG, and a special charged tRNA is required to initiate translation.
- This initiator tRNA always carries the amino acid methionine.
- Thus newly made proteins all have methionine as the first amino acid at their N-terminal end, the end of a protein that is synthesized first.
- This methionine is usually removed later by a specific protease.

- In eukaryotes, an initiator tRNA, charged with methionine, is first loaded into the P site of the small ribosomal subunit, along with additional proteins called ***translation initiation factors***.
- The initiator tRNA is distinct from the tRNA that normally carries methionine.
- Of all the tRNAs in the cell, only a charged initiator tRNA molecule is capable of binding tightly to the P site in the absence of the large ribosomal subunit.
- Next, the small ribosomal subunit loaded with the initiator tRNA binds to the 5′ end of an mRNA molecule, which is marked by the 5′ cap that is present on all eukaryotic mRNAs.
- The small ribosomal subunit then moves forward (5′ to 3′) along the mRNA searching for the first AUG.
- When this AUG is encountered and recognized by the initiator tRNA, several initiation factors dissociate from the small ribosomal subunit to make way for the large ribosomal subunit to bind and complete ribosomal assembly.
- Because the initiator tRNA is bound to the P site, protein synthesis is ready to begin with the addition of the next charged tRNA to the A site.



End of Translation

- The end of translation in both prokaryotes and eukaryotes is signaled by the presence of one of several codons, called *stop codons*, in the mRNA. The stop codons—***UAA***, ***UAG***, and ***UGA***—are not recognized by a tRNA and do not specify an amino acid, but instead signal to the ribosome to stop translation.
- Proteins known as ***release factors*** bind to any stop codon that reaches the A site on the ribosome; this binding alters the activity of the peptidyl transferase in the ribosome, causing it to catalyze the addition of a water molecule instead of an amino acid to the peptidyl-tRNA.
- This reaction frees the carboxyl end of the polypeptide chain from its attachment to a tRNA molecule; because this is the only attachment that holds the growing polypeptide to the ribosome, the completed protein chain is immediately released.
- At this point, the ribosome also releases the mRNA and dissociates into its two separate subunits, which can then assemble on another mRNA molecule to begin a new round of protein synthesis.



- Multiple ribosomes usually bind to each mRNA molecule being translated.
- If the mRNA is being translated efficiently, a new ribosome hops onto the 5′ end of the mRNA molecule almost as soon as the preceding ribosome has translated enough of the nucleotide sequence to move out of the way.
- The mRNA molecules being translated are therefore usually found in the form of ***polyribosomes***.
- By this way proteins can be synthesized efficiently in large amounts.