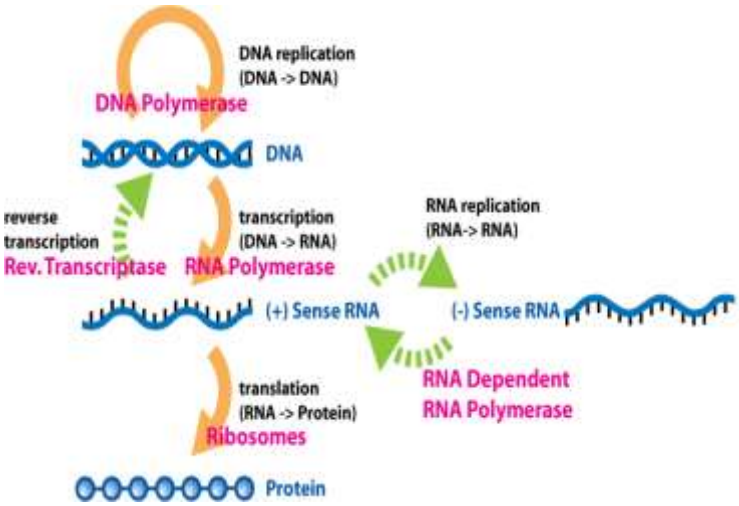
GENE EXPRESSION

- The process by which a gene's DNA sequence is converted into the structure and functions of a cell
- Genetic information chemically determined by DNA structure is transferred to daughter cells by **DNA replication**
- It is then expressed by **Transcription and Translation**
- This series of events is called Central Dogma
- It is found in all the cells (exceptions such as retroviruses where reverse transcription occurs and prion)



TRANSCRIPTION

- The process through which a DNA sequence is enzymatically copied by an RNA Polymerase to produce a complementary DNA
- It is the transfer of genetic information from DNA to RNA
- RNA polymerase, together with one or more transcription factors, binds to promoter DNA.
- RNA polymerase generates a transcription bubble, which separates the two strands of the DNA helix. This is done by breaking the hydrogen bonds between complementary DNA nucleotides.
- RNA polymerase adds RNA nucleotides (which are complementary to the nucleotides of one DNA strand).
- RNA sugar-phosphate backbone forms with assistance from RNA polymerase to form an RNA strand.
- Hydrogen bonds of the RNA–DNA helix break, freeing the newly synthesized RNA strand.

- If the cell has a nucleus, the RNA may be further processed. This may include polyadenylation, capping, and splicing.
- The RNA may remain in the nucleus or exit to the cytoplasm through the nuclear pore complex.

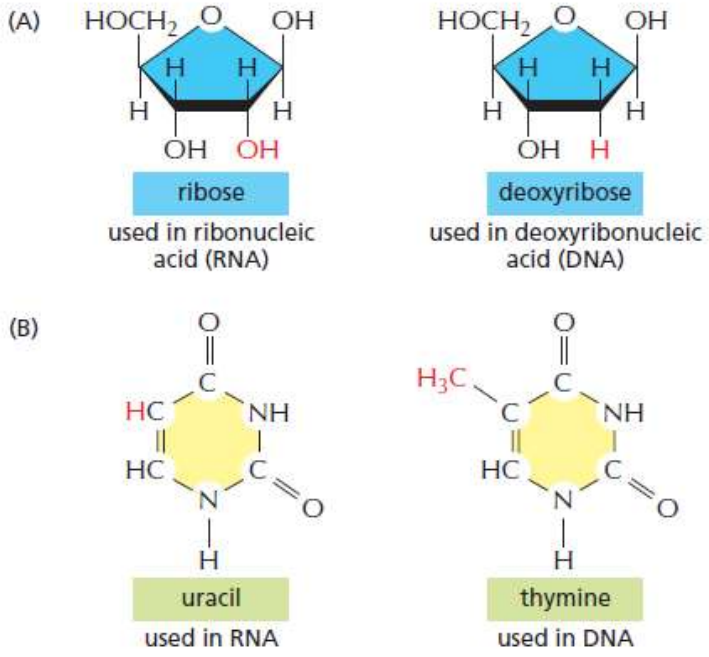


TABLE 6–1 Principal Types of RNAs Produced in Cells

Type of RNA	Function
mRNAs	Messenger RNAs, code for proteins
rRNAs	Ribosomal RNAs, form the basic structure of the ribosome and catalyze protein synthesis
tRNAs	Transfer RNAs, central to protein synthesis as adaptors between mRNA and amino acids
snRNAs	Small nuclear RNAs, function in a variety of nuclear processes, including the splicing of pre-mRNA
snoRNAs	Small nucleolar RNAs, help to process and chemically modify rRNAs
miRNAs	MicroRNAs, regulate gene expression by blocking translation of specific mRNAs and cause their degradation
siRNAs	Small interfering RNAs, turn off gene expression by directing the degradation of selective mRNAs and the establishment of compact chromatin structures
piRNAs	Piwi-interacting RNAs, bind to piwi proteins and protect the germ line from transposable elements
lncRNAs	Long noncoding RNAs, many of which serve as scaffolds; they regulate diverse cell processes, including X-chromosome inactivation

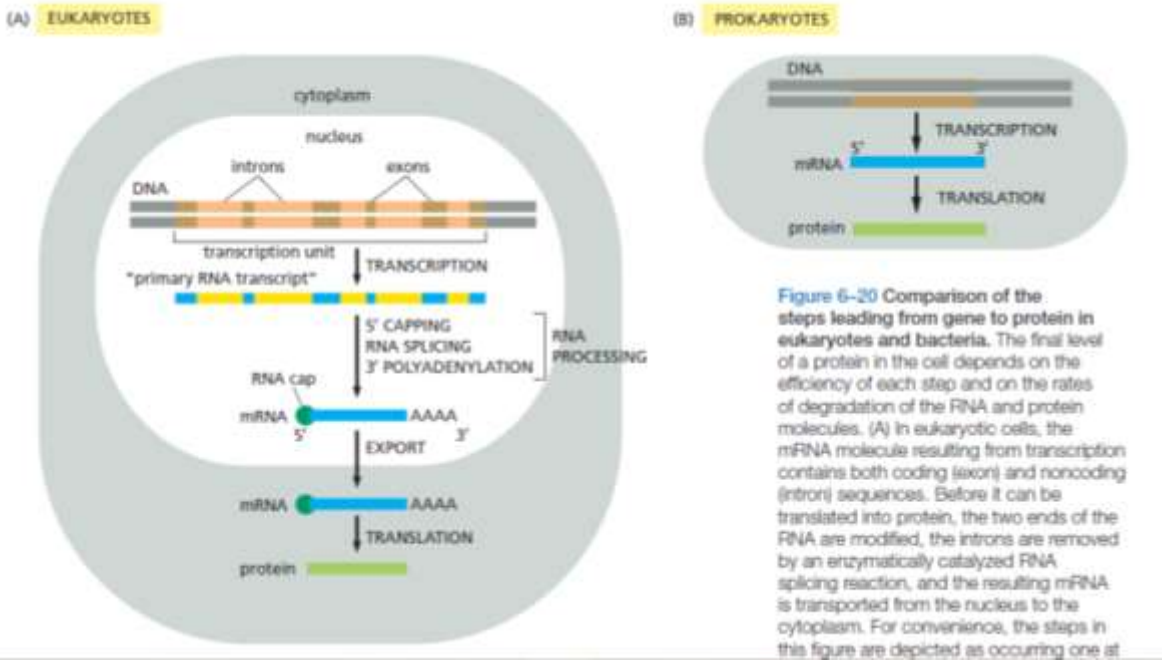
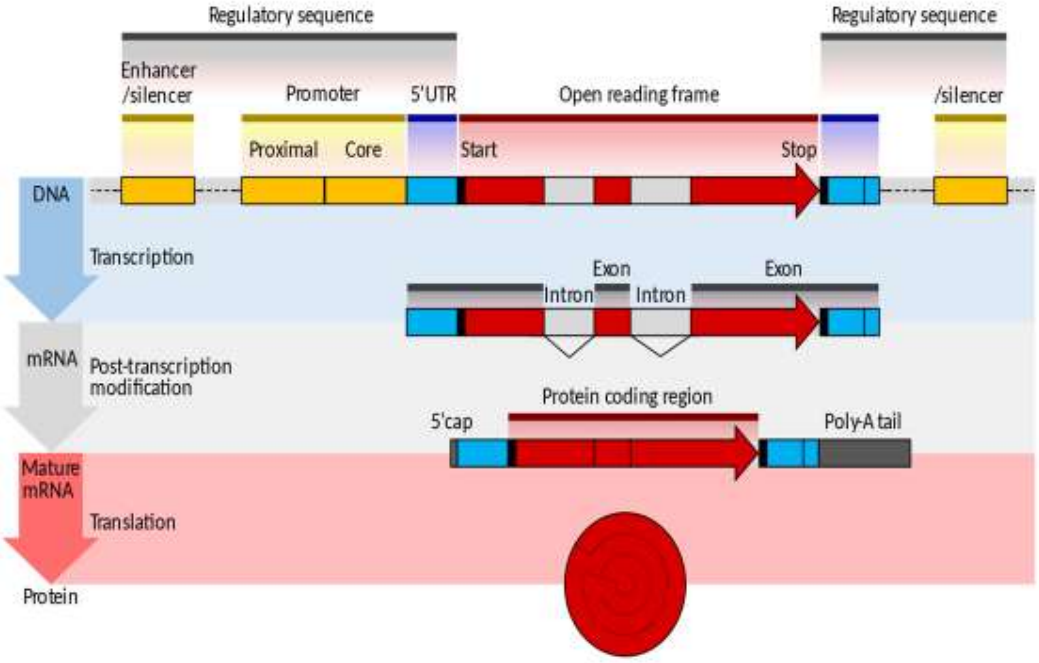


Figure 6-20 Comparison of the steps leading from gene to protein in eukaryotes and bacteria. The final level of a protein in the cell depends on the efficiency of each step and on the rates of degradation of the RNA and protein molecules. (A) In eukaryotic cells, the mRNA molecule resulting from transcription contains both coding (exon) and noncoding (intron) sequences. Before it can be translated into protein, the two ends of the RNA are modified, the introns are removed by an enzymatically catalyzed RNA splicing reaction, and the resulting mRNA is transported from the nucleus to the cytoplasm. For convenience, the steps in this figure are depicted as occurring one at

STEPS IN TRANSCRIPTION

- Initiation
- Elongation
- Termination



INITIATION

- Initiation of transcription requires **promoter regions**, which are specific nucleotide consensus sequence (most frequent nucleotides) that tell the **σ-factor** (sigma) on RNA polymerase where to bind to the DNA. They are most commonly found upstream of the genes they control
- The promoter region is a prime regulator of transcription. It regulate transcription of all genes within bacteria.
- There are two base sequences on the coding strand which RNA Polymerase can recognize for initiation

❑ The **Pribnow Box** (TATA Box) consisting of 6 nucleotide bases (TATAAT). It is located on the left side about 10 bases upstream from the starting point of transcription (Prokaryotes)

❑ The **- 35 sequence**, which is the second recognition site in the promoter region of the DNA and contains a base sequence TTGACA which is located about 35 bases upstream of the transcription starting point (Prokaryotes)

- transcription begins at the start site (+1)
- **Eukaryotic genes** have a conserved promoter sequence called the TATA box, located 25 to 35 base pairs upstream (– 25 to – 35) of the transcription start site
- B Recognition Element (BRE)
- Eukaryotic promoter regulatory sequences typically bind proteins called **transcription factors (TF)** that are involved in the formation of the transcriptional complex
- Single RNA Polymerase in Prokaryotes and Archaebacteria, while 3 types of RNA Polymerases in Eukaryotes
- RNA Polymerase contains a core Mg+ ion that assists the enzyme with its catalytic properties.
- It works by catalysing the nucleophilic attack of 3’ OH of RNA to the alpha phosphate of a complementary NTP molecule to create a growing strand of RNA from the template strand of DNA.
- It also displays exonuclease activities, meaning that if improper base pairing is detected, it can cut out the incorrect bases and replace them with the proper, correct one.
- It transcribes RNA in 5’-3’ direction antiparallel to DNA
- The σ-factor binds to the -35 promoter region where RNA Polymerase binds to the double stranded DNA (ds DNA). This structure is called **closed complex**
- The high concentration of adenine-thymine bonds at the -10 region facilitates the unwinding of the DNA

- After binding of the polymerase, the DNA double helix is partially unwound and becomes single stranded (ss) in the area near the initiation site. This is called **open complex (transcription bubble)**

TABLE 6–2 The Three RNA Polymerases in Eukaryotic Cells

Type of polymerase	Genes transcribed
RNA polymerase I	5.8S, 18S, and 28S rRNA genes
RNA polymerase II	All protein-coding genes, plus snoRNA genes, miRNA genes, siRNA genes, lncRNA genes, and most snRNA genes
RNA polymerase III	tRNA genes, 5S rRNA genes, some snRNA genes, and genes for other small RNAs

The rRNAs were named according to their "S" values, which refer to their rate of sedimentation in an ultracentrifuge. The larger the S value, the larger the rRNA.

ELONGATION

- RNA synthesis proceeds with addition of ribonucleotides ATP, GTP, CTP and UTP
- One DNA strand serves as a template strand for RNA synthesis
- The addition of ribonucleotides is in such a way that G in DNA pairs with C in RNA, C with G, T with A and A with U (T is replaced by U in RNA)
- The newly formed RNA strand is identical to the DNA coding strand (sense strand or non-template strand), except it has uracil substituting thymine, and a ribose sugar backbone instead of a deoxyribose sugar backbone
- Multiple RNA polymerases can be active at once so that many strands of mRNA can be produced very quickly

TERMINATION

- Two termination mechanisms are known –

❑ **Rho-independent** (Intrinsic): Termination sequence within the RNA signals RNA polymerase to STOP. This sequence is usually a palindromic sequence that forms a stem-loop hairpin structure. RNA polymerase dissociates from the template

❑ **Rho-dependent**: It uses a termination factor called **ρ** factor (rho factor) to stop RNA synthesis at specific sites. Rho factor is a protein which binds and runs along mRNA towards RNA polymerase. On reaching the polymerase, it causes RNA polymerase to dissociate from DNA and terminate transcription. A series of G nucleotides at this region causes it to stop

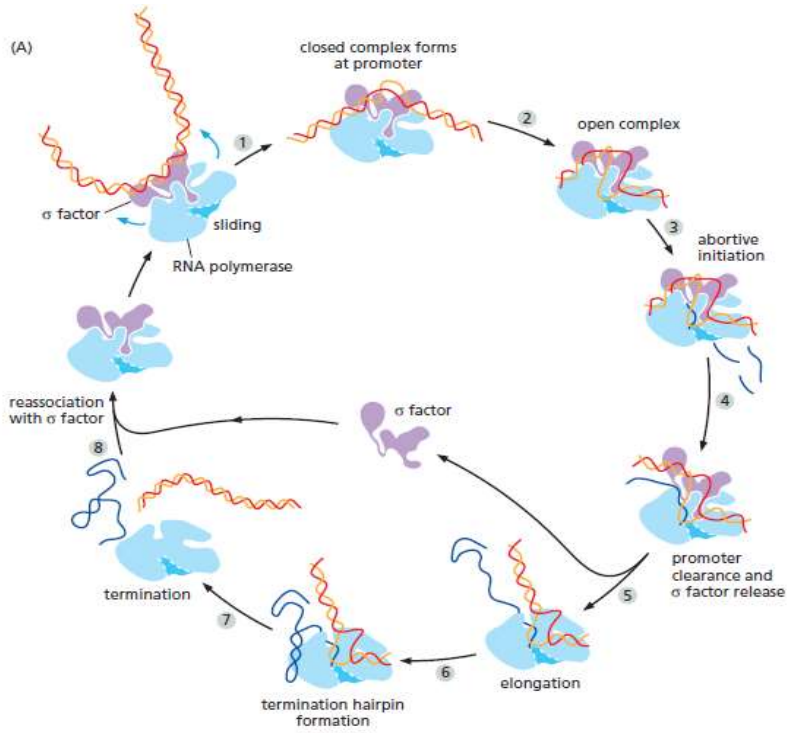


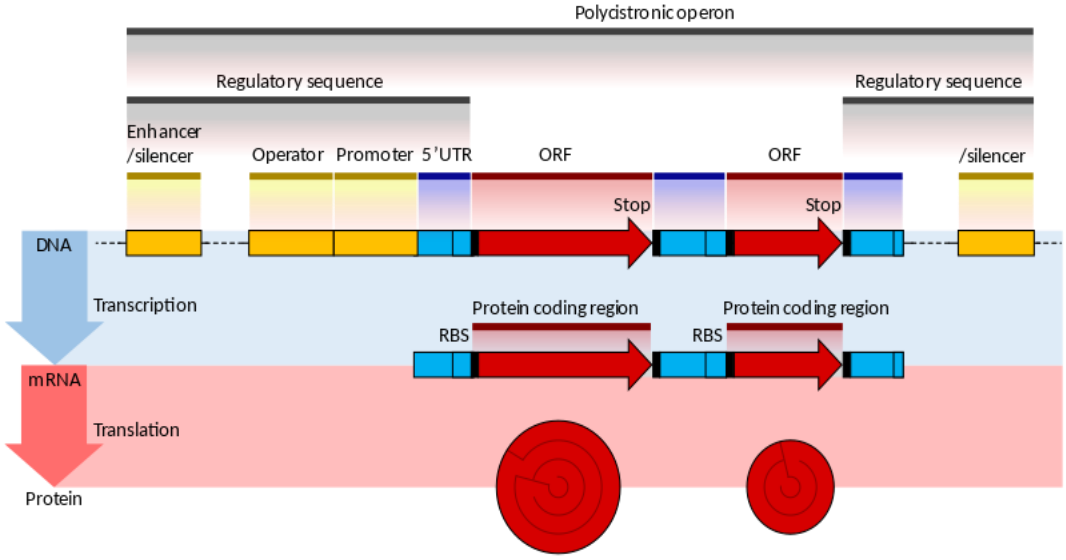
Figure 6–11 The transcription cycle of bacterial RNA polymerase. (A) In step 1, the RNA polymerase holoenzyme (polymerase core enzyme plus σ factor) assembles and then locates a promoter DNA sequence (see Figure 6–12). The polymerase opens (unwinds) the DNA at the position at which transcription is to begin (step 2) and begins transcribing (step 3). This initial RNA synthesis (abortive initiation) is relatively inefficient as short, unproductive transcripts are often released. However, once RNA polymerase has managed to synthesize about 10 nucleotides of RNA, it breaks its interactions with the promoter DNA (step 4) and eventually releases σ factor—as the polymerase tightens around the DNA

POST-TRANSCRIPTIONAL MODIFICATION OR CO-TRANSCRIPTIONAL MODIFICATION

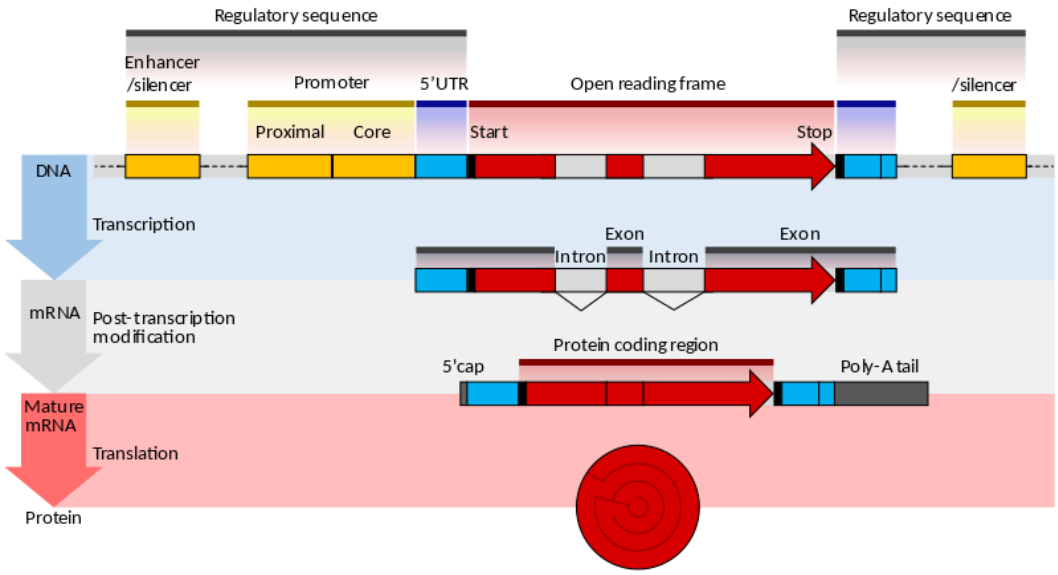
- It is a set of biological processes common to most eukaryotic cells by which an RNA primary transcript is chemically altered following transcription from a gene to produce a mature, functional RNA molecule that can then leave the nucleus and perform any of a variety of different functions in the cell
- In both prokaryotes and eukaryotes, primary RNA transcripts undergo various alterations or processing events to become mature RNAs. The three commonest types are:
  - Nucleotide removal by nucleases,
  - Nucleotide addition to the 5- or 3-end, and
  - Nucleotide modification on the base or the sugar
- The major difference in RNA processing between prokaryotes and eukaryotes is in the processing of messenger RNAs. The noncoding regions must be removed before the mRNA is sent out of the nucleus into the cytoplasm for protein synthesis.
- Bacterial mRNAs are synthesized by the RNA polymerase starting and stopping at specific spots on the genome. The situation in eukaryotes is different.
- Transcription is only the first of several steps needed to produce a mature mRNA molecule.
- Other critical steps are the covalent modification of the ends of the RNA and the removal of *intron sequence* that are discarded from the middle of the RNA transcript by the process of *RNA splicing*

- Both ends of eukaryotic mRNAs are modified: by *capping* on the 5′ end and by *polyadenylation* of the 3′ end
- The precursor RNA is converted to mature RNA (modified)
- There are three main modifications –
- **5' capping** : addition of 7-methlyguanosine cap to 5' end
- **3' polyadenylation** : cleavage of 3' end of the primary transcript followed by addition of Poly A tail
- **RNA splicing** : This is the process by which introns are removed from mRNA and remaining exons are connected to form a single continuous molecule.
- Splicing reaction is catalysed by a large protein complex called **spliceosome**

IN PROKARYOTES

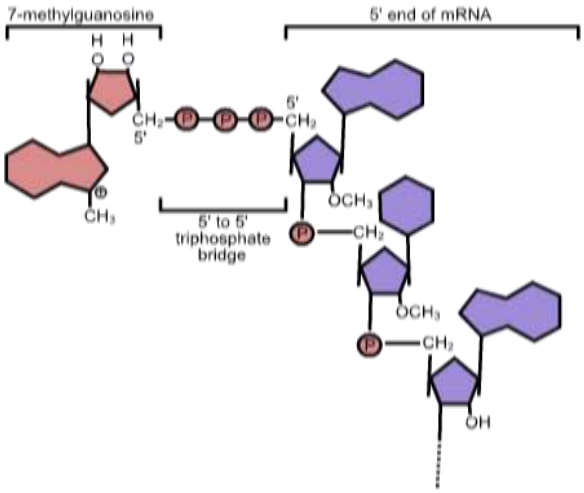


IN EUKARYOTES



CAPPING

- Capping of the pre-mRNA involves the addition of 7-methylguanosine (m^7G) to the 5' end.
- To achieve this, the terminal 5' phosphate requires removal, which is done with the help of a phosphatase enzyme.
- The enzyme guanosyl transferase then catalyses the reaction, which produces the diphophate 5' end.
- The diphosphate 5' end then attacks the alpha phosphorus atom of a GTP molecule in order to add the guanine residue in a 5'5' triphosphate link.
- The enzyme guanine-N^7-methyltransferase ("cap MTase") transfers a methyl group from S-adenosyl methionine to the guanine ring.
- This type of cap, with just the (m^7G) in position is called a cap 0 structure.
- The ribose of the adjacent nucleotide may also be methylated to give a cap 1.
- Methylation of nucleotides downstream of the RNA molecule produce cap 2, cap 3 structures and so on.
- In these cases the methyl groups are added to the 2' OH groups of the ribose sugar.
- The cap protects the 5' end of the primary RNA transcript from attack by ribonucleases that have specificity to the 3'5' phosphodiester bonds



TAILING

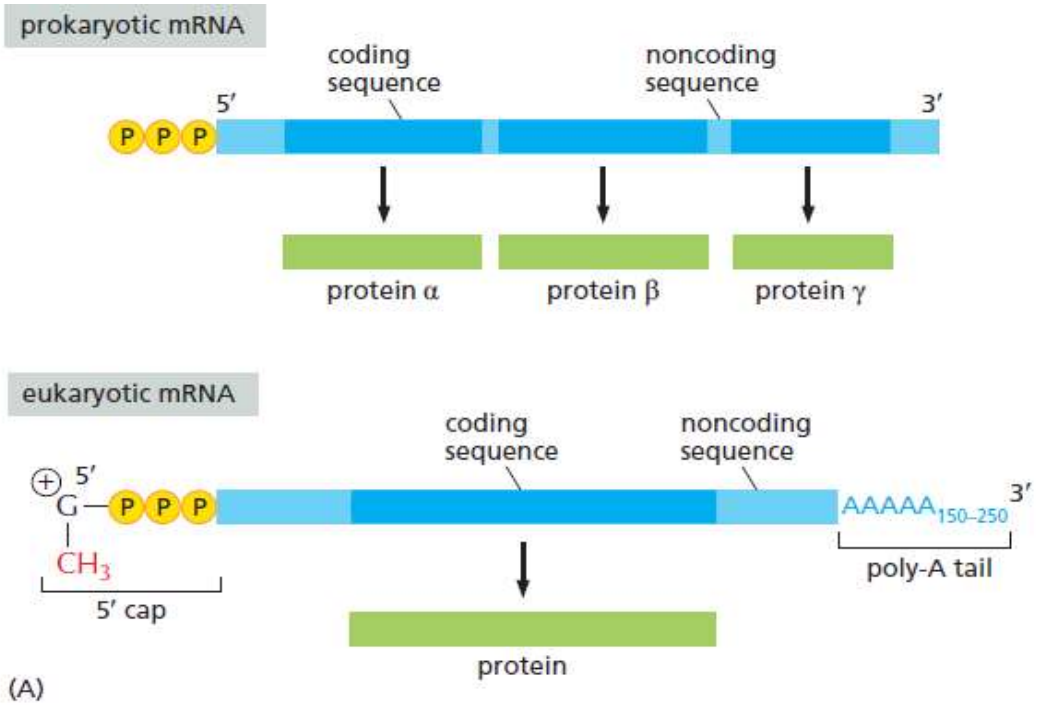
- **It** is the addition of a **poly(A) tail** to a messenger RNA
- The poly(A) tail consists of multiple adenosine monophosphates
- it is a stretch of RNA that has only adenine bases.
- In eukaryotes, polyadenylation is part of the process that produces mature messenger RNA (mRNA) for translation.
- In many bacteria, the poly(A) tail promotes degradation of the mRNA.
- The process of polyadenylation begins as the transcription of a gene terminates.

- ❖ The 3’ segment of the newly made pre-mRNA is first cleaved off by a set of proteins, these proteins then synthesize the poly(A) tail at the RNA's 3′ end
- ❖ The poly(A) tail is important for the nuclear export, translation, and stability of mRNA.
- ❖ The tail is shortened over time, and, when it is short enough, the mRNA is enzymatically degraded

NOTE: mRNA molecules in both prokaryotes and eukaryotes have polyadenylated 3′-ends, with the prokaryotic poly(A) tails generally shorter and less mRNA molecules polyadenylated

SPLICING

- Eukaryotes splice many protein-coding messenger RNAs and some non-coding RNAs.
- Prokaryotes splice rarely and mostly non-coding RNAs.
- Another important difference between these two groups of organisms is that prokaryotes completely lack the spliceosome pathway.
- During splicing, introns (non-coding regions) are removed and exons (coding regions) are joined together.
- For NUCLEAR ENCODED GENES, splicing takes place within the nucleus either during or immediately after transcription.
- For those eukaryotic genes that contain introns, splicing is usually required in order to create an mRNA molecule that can be translated into protein.
- For many eukaryotic introns, splicing is carried out in a series of reactions which are catalysed by the SPLICEOSOME, which is a complex of large RNA-protein complex composed of five small nuclear ribonucleoproteins (snRNPs)
- The **major spliceosome** splices introns containing GU at the 5' splice site and AG at the 3' splice site. It is composed of the U1, U2, U4, U5, and U6 snRNPs and is active in the nucleus.
- The m**inor spliceosome** is very similar to the major spliceosome, but the difference is it splices out rare introns with different splice site sequences.
- **Self-splicing** occurs for rare introns that form a ribozyme, performing the functions of the spliceosome by RNA alone. These are capable of catalysing their own excision from their parent RNA molecule
- There are three kinds of self-splicing introns, *Group I*, *Group II* and *Group III*. Group I and II introns perform splicing similar to the spliceosome without requiring any protein.



Transportation of processed mRNA from the nucleus to the cytoplasm

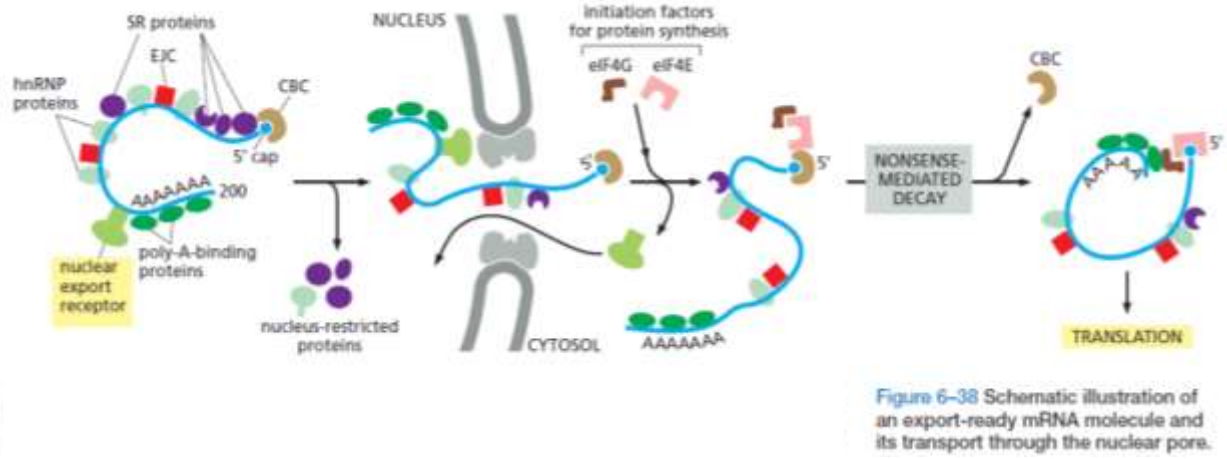


Figure 6–38 Schematic illustration of an export-ready mRNA molecule and its transport through the nuclear pore.

TRANSLATION

- Translation is synthesis of proteins on mRNA template.
- Proteins are made up of a set of 20 amino acids called standard amino acids
- Each of the 20 amino acids is specified by specific codon(s)
- Polypeptide synthesis proceeds from N-terminus to C-terminus
- The ribosome read mRNA in the 5’ to 3’ direction

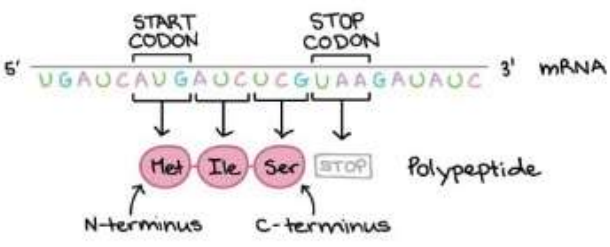
Three types of molecules are involved

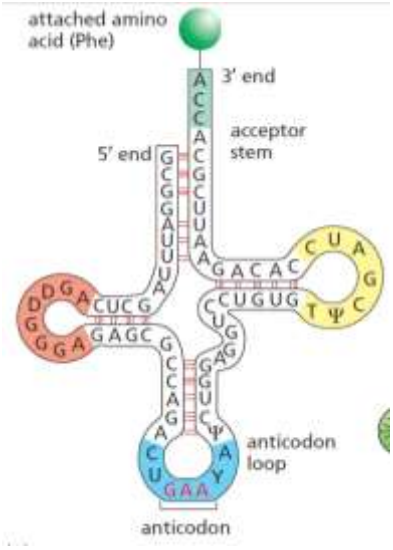
- Messenger RNA (mRNA)

- Transfer RNA (tRNA)
- Ribosomal RNA (rRNA)

mRNA

- Carries the genetic information in the form of codons, each specifying a particular amino acid
- mRNAs are **monocistronic (encodes one protein) / polycistronic (encodes several proteins)**
- All mRNAs consist of two types of regions – coding and untranslated region
- The coding region starts with AUG and ends with a terminating sequence
- Extra regions present in it are called untranslated regions (at **5' end it is called leader** and at **3' end it is called trailer in prokaryotes**; in eukaryotes, processing is involved)
- An mRNA can be translated into **3 reading frames** (one of the three possible ways of reading a nucleotide sequence)
- An **open reading frame (ORF)** is a sequence of DNA consisting of triplets that can be translated into amino acids starting with an initiation codon and ending with a termination codon
- **Initiation codon is AUG** and **STOP codon is one of the three – UAG/ UAA/ UGA**. These are also known as **nonsense codons**
- A reading frame which has termination sequences occurring frequently is a **blocked reading frame**
- A reading frame for which no protein product has been identified is called an **unidentified reading frame (URF)**





tRNA

- Each type of amino acid has its own type of tRNA

- It carries the anticodons
- The tRNA binds to the amino acid and carries it to the growing end of a polypeptide chain

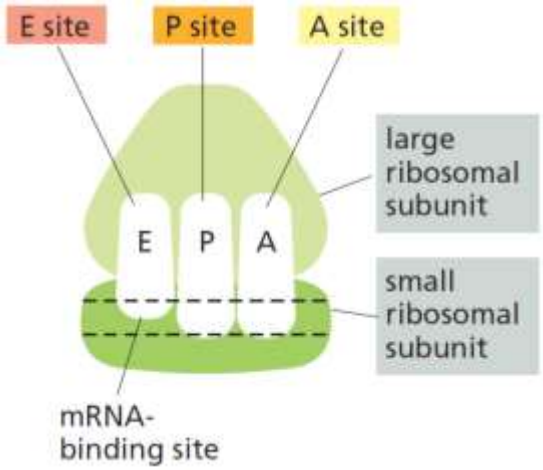
rRNA

- rRNA associates with a set of proteins to form ribosomes
- This complex structure physically moves along the mRNA molecule in 5' to 3' direction and catalyse the assembly of amino acids into protein chains
- They also bind the tRNA and other molecules necessary for protein synthesis

RIBOSOMES

- Ribosomes act like small migratory factory
- Ribosomes are ribonucleoproteins that contain rRNA and r-proteins
- Each ribosome is made up of two subunits – small and large
- The ribosome has 3 tRNA binding sites

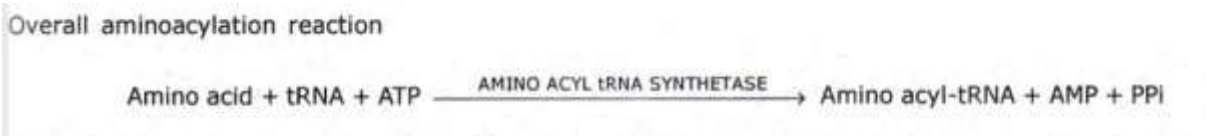
1. An aminoacyl-tRNA enters the **A** site
2. Peptidyl-tRNA (carrying polypeptide chain) is bound in the **P** site
3. Deacylated tRNA (which lacks any amino acid) exits from the **E** site



ACTIVATION OF AMINO ACID & ATTACHMENT TO tRNA

- Attachment of amino acids to tRNA is facilitated by a group of enzymes called **aminoacyl-tRNA synthetases**
- This enzyme activates amino acids by covalently linking them to tRNAs.
- When a tRNA is charged with the amino acid corresponding to its anticodon, it is called **aminoacyl-tRNA**
- The enzyme catalyses the reaction in 2 steps

1. An activated amino acid intermediate is formed by reaction between the amino acid and ATP
2. Amino acid is transferred to the 3' end of the tRNA and the link is formed between –COOH group of amino acid and the –OH group attached to either 2' or 3' carbon on the sugar of the last nucleotide
3. In aminoacylation, amino acid is linked to tRNA by high energy bond and is thus activated



- Organisms have 20 aminoacyl-tRNA synthetases (few exceptions) one for each amino acid
- Synthetases recognize the anticodon loop and acceptor stems of transfer RNA molecules
- The 20 aminoacyl-tRNA synthetases can be divided into 2 categories- Class I and Class II
- Class I – attach to 2'-OH of the terminal nucleotide of the TRNA
- Class II – attach to 3'-OH of the terminal nucleotide of the TRNA
- The translation of the nucleotide sequence of an mRNA molecule into protein takes place in the **cytosol (cytoplasm).**
- It takes place on a large ribonucleoprotein assembly called a **ribosome**.
- Each amino acid used for protein synthesis is first attached to a tRNA molecule which possess an **anticodon.**
- The tRNA recognizes, by complementary base-pair interactions, a particular set of **three nucleotides (codons)** in the mRNA.
- As an mRNA is passed through a ribosome, its sequence of nucleotides is read from one end to the other **in sets of three (triplet)** according to the genetic code.
- **Different sets of accessory protein factors** assist the ribosome at each stage.
- **Energy** is provided at various stages **by the hydrolysis of guanine triphosphate (GTP).**

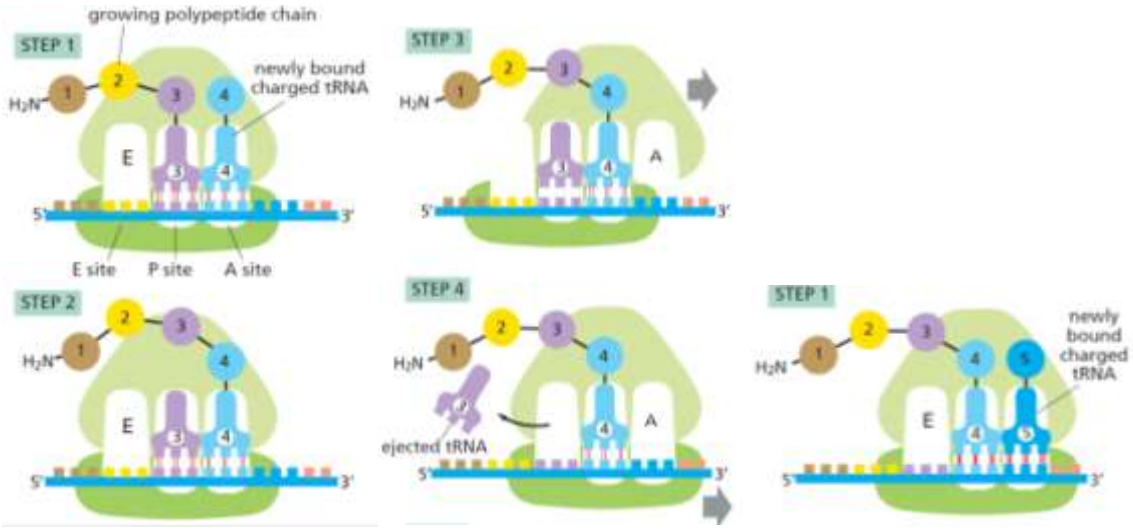
STEPS INVOLVED IN PROTEIN SYNTHESIS

- The translation of the nucleotide sequence of an mRNA molecule into protein takes place in the **cytosol (cytoplasm).**
- It takes place on a large ribonucleoprotein assembly called a **ribosome**.
- Each amino acid used for protein synthesis is first attached to a tRNA molecule which possess an **anticodon.**

- The tRNA recognizes, by complementary base-pair interactions, a particular set of **three nucleotides (codons)** in the mRNA.
- As an mRNA is passed through a ribosome, its sequence of nucleotides is read from one end to the other **in sets of three (triplet)** according to the genetic code.
- **Different sets of accessory protein factors** assist the ribosome at each stage.
- **Energy** is provided at various stages **by the hydrolysis of guanine triphosphate (GTP).**

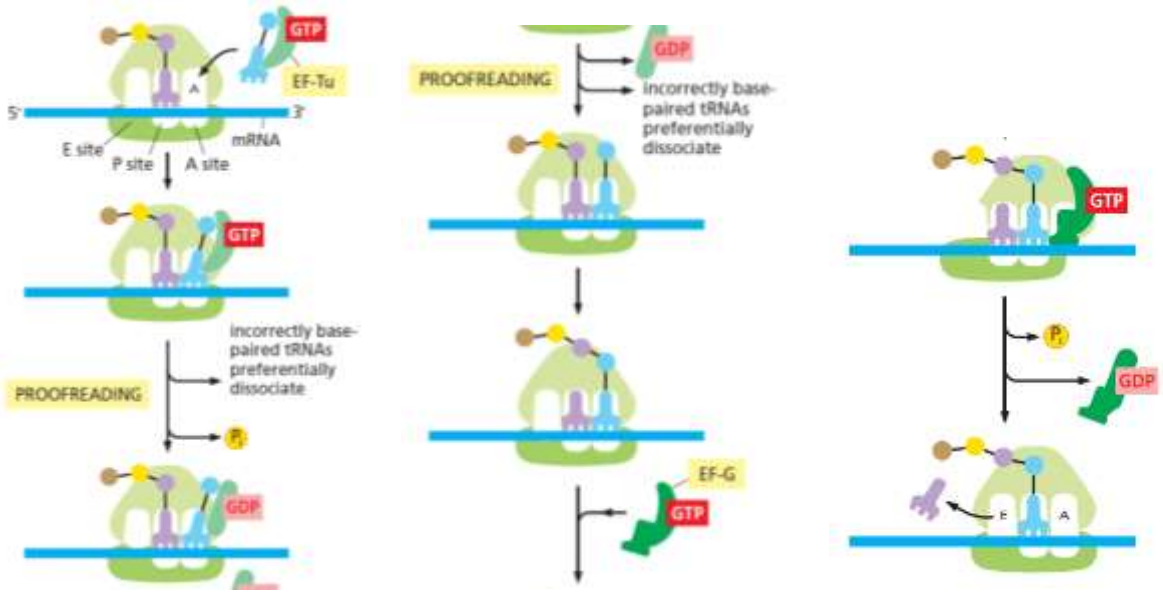
INITIATION

- Involves the reactions that result in the formation of the peptide bond between the first two amino acids of the polypeptide.
- To initiate translation, a small ribosomal subunit binds to the mRNA molecule at a start codon (AUG) that is recognized by a unique initiator tRNA molecule.
- A large ribosomal subunit then binds to complete the ribosome and begin protein synthesis.
- Aminoacyl-tRNAs—each bearing a specific amino acid—bind sequentially to the appropriate codons in mRNA through complementary base-pairing between tRNA anticodons and mRNA codons.
- Each amino acid is added to the C-terminal end of the growing polypeptide in four sequential steps: aminoacyl-tRNA binding, followed by peptide bond formation, followed by two ribosome translocation steps.
- Initiation is <u>a relatively slow step in translation</u> and usually determines the rate at which an mRNA is translated



ELONGATION

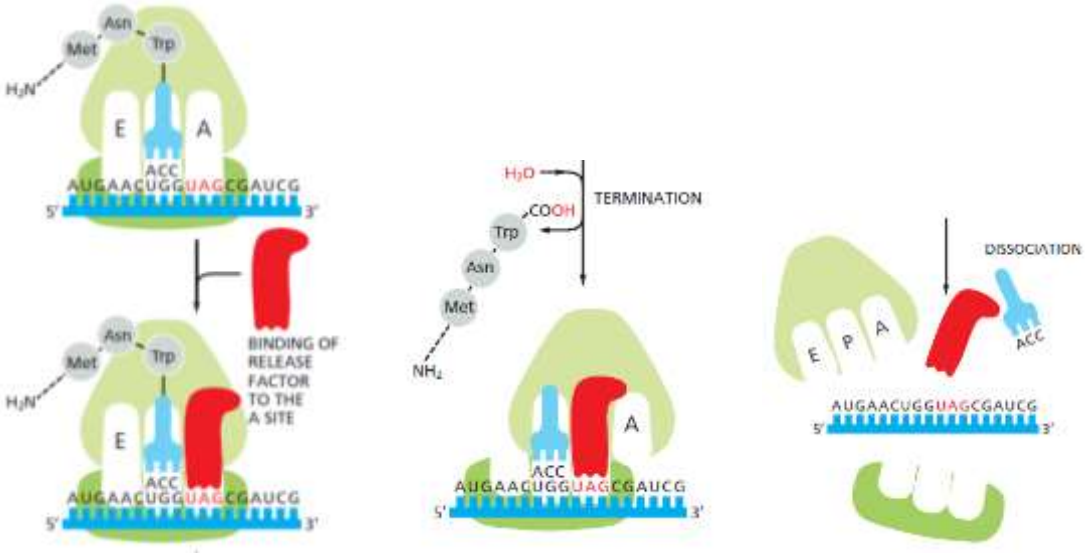
- Elongation includes all the reactions from the formation of the first peptide bond to the addition of the last amino acid.
- Amino acids are added to the chain one at a time.
- The addition of an amino acid is the most rapid step in translation.
- Elongation factors use GTP hydrolysis both to drive these reactions forward and to improve the accuracy of amino acid selection.
- The mRNA molecule progresses codon by codon through the ribosome in the 5′-to-3′ direction until it reaches one of three stop codons.
- EF-Tu is a monomeric G protein whose active form (bound to GTP) binds to aminoacyl-tRNA. The EF-Tu–GTP–aminoacyl-tRNA complex binds to the ribosome's A site.
- Two types of tRNA can carry the amino acid methionine. One is used for initiation, the other for recognizing AUG codons during elongation



TERMINATION

- A release factor binds to the ribosome, terminating translation and releasing the completed polypeptide.
- The ribosome dissociates from the mRNA.
- The codons UAA (ochre), UAG (amber), and UGA (opal) terminate translation.
- In bacteria, they are used most often with relative frequencies of UAA > UGA > UAG.
- Termination codons are recognized by protein release factors, not by aminoacyl-tRNAs.

- The mechanism of termination in bacteria (which have two types of class 1 release factors, RF1 & RF2) is similar to that of eukaryotes (which have only one class 1 release factor).



Post-Translational Modification

- In the final steps of protein synthesis, two distinct types of molecular chaperones, known as **hsp60 and hsp70** guide the folding of polypeptide chains.

An error can be made at several steps in gene expression:

- The enzymes that synthesize RNA may insert a base that is not complementary to the base on the template strand.
- Synthetases may attach the wrong tRNA to an amino acid or the wrong amino acid to a tRNA.
- A ribosome may allow binding of a tRNA that does not correspond to the codon in the A site.

DIFFERENCE IN TRANSLATION BETWEEN PROKARYOTES & EUKARYOTES

- **Initiation of translation in prokaryotes** requires separate 30S and 50S ribosomal subunits.
- Prokaryotes use three initiation factors, numbered **IF-1**, **IF-2**, and **IF-3**. They are needed for both mRNA and tRNA to enter the initiation complex, which bind to 30S subunits.
- A 30S subunit carrying initiation factors binds to an initiation site on the mRNA to form an initiation complex.
- IF-2 binds the initiator fMet-tRNA and allows it to enter the partial P site on the 30S subunit.

- IF-3 must be released to allow the 50S subunit to join the 30S-mRNA complex.
- The Shine-Dalgarno sequence is a ribosomal binding site in bacterial and Archaeal messenger RNA. It is located generally at 8 bases upstream of AUG. An initiation site on bacterial mRNA consists of the AUG or less often, GUG or UUG.
- In bacteria, mitochondria, and chloroplasts, the initiator tRNA carries a methionine residue that has been formylated on its amino group, forming a molecule of **N-formyl-methionyl-tRNA** (*fMet-tRNA)*.

Difference Between Prokaryotic And Eukaryotic Translation

Prokaryotic Translation	Eukaryotic Translation
Definition	
The translation & transcription process is synchronous	The translation and transcription process is discontinuous
mRNA	
Cytoplasm	Nucleus
Cap initiation	
Cap-independent	Cap-dependent and Cap-independent
Performed by	
70S ribosomes	80S ribosomes
Stability of mRNA	
Unstable	Stable
Ribosomes	
30S & 50S = 70S	40S & 60S = 80S
Lifespan of mRNA	
A few seconds to 2 minutes	A few hours to days
Occurrence	
No definite phase	G1 and G2 phase of the cell cycle
Process	
Fast	Slow
Release factor	
RF1, RF2	eRF
Initiation factors	
3	9

SOURCE OF FIGURES:

Molecular Biology of The Cell. Bruce Alberts et al. 6th Edition. Garland Science

Lewin;s Gene XII. Krebs et al. 2018 Jones & Bartlett Learning

Wikipedia

FURTHER READING:

Cell and Molecular Biology. P.K. Gupta. Rastogi Publications

Cell Biology, Genetics, Mol.Biology, Evolution and Ecology. P.S. Verma. S.Chand Publication

Molecular Biology of The Cell. Bruce Alberts et al. 6th Edition. Garland Science

Cell and Molecular Biology- Concepts and Experiments- Gerald Karp. 7th Edition Wiley

Lewin;s Gene XII. Krebs et al. 2018 Jones & Bartlett Learning