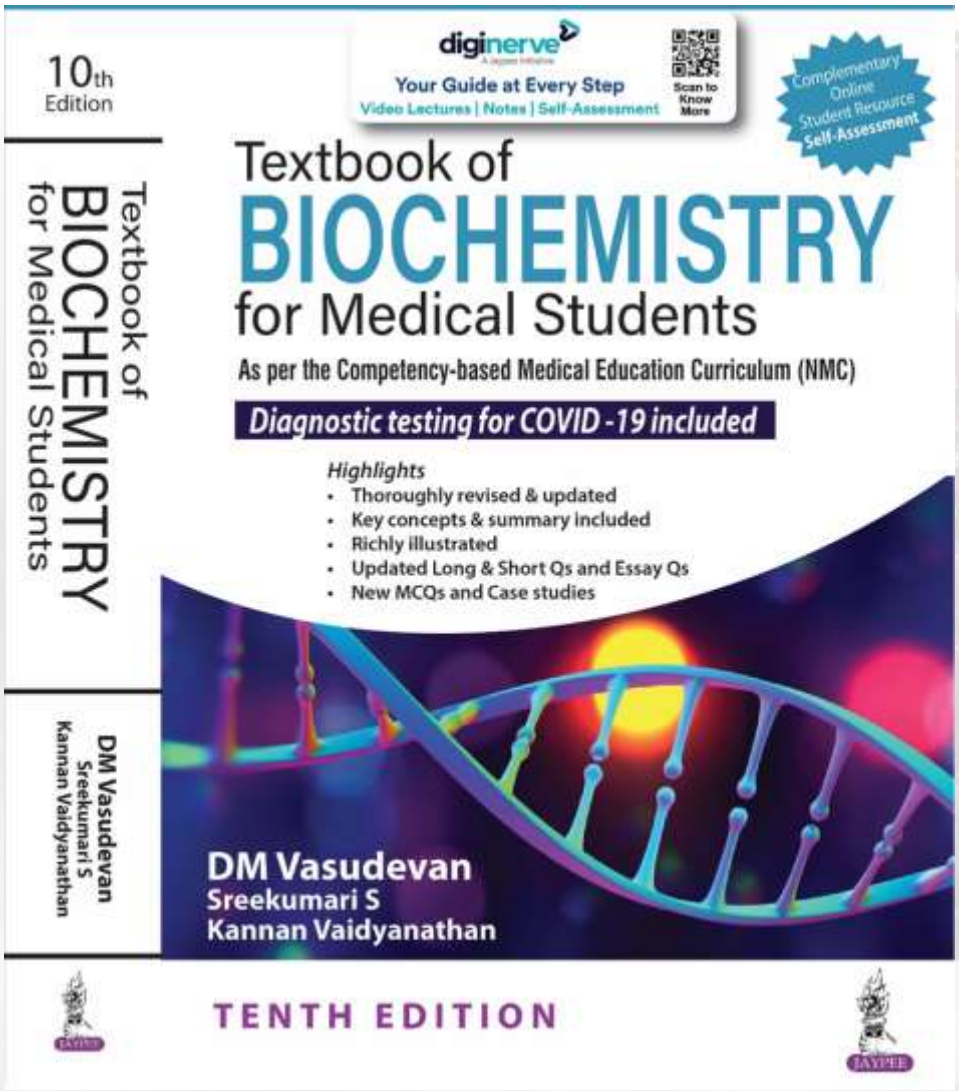


Chapter 35:

Transcription and Translation

Textbook of

BIOCHEMISTRY

for Medical Students

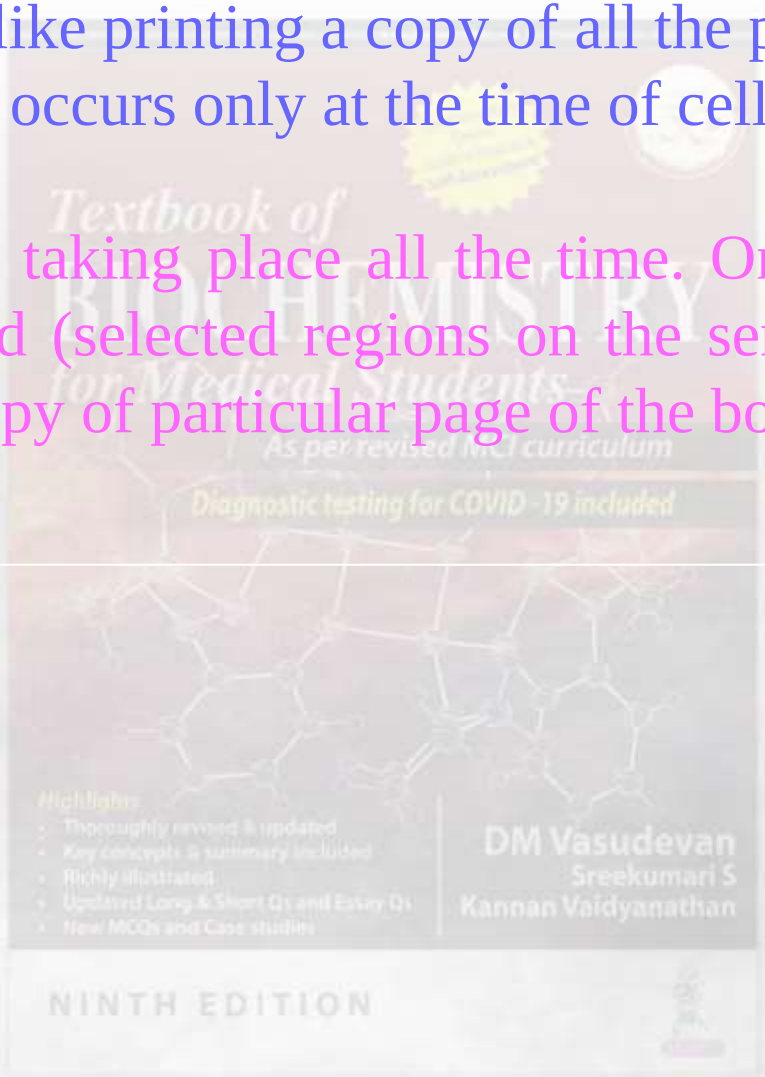
By DM Vasudevan, *et al.*

TENTH EDITION

Transcription

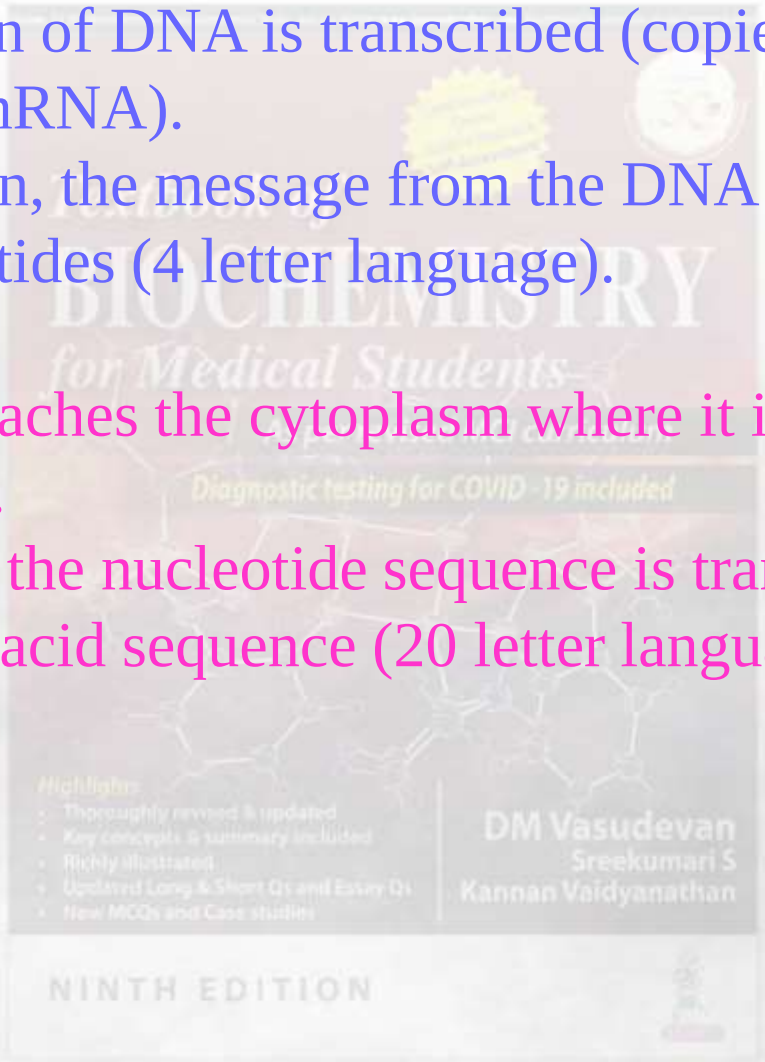
DNA replication is like printing a copy of all the pages of a book. Replication process occurs only at the time of cell division.

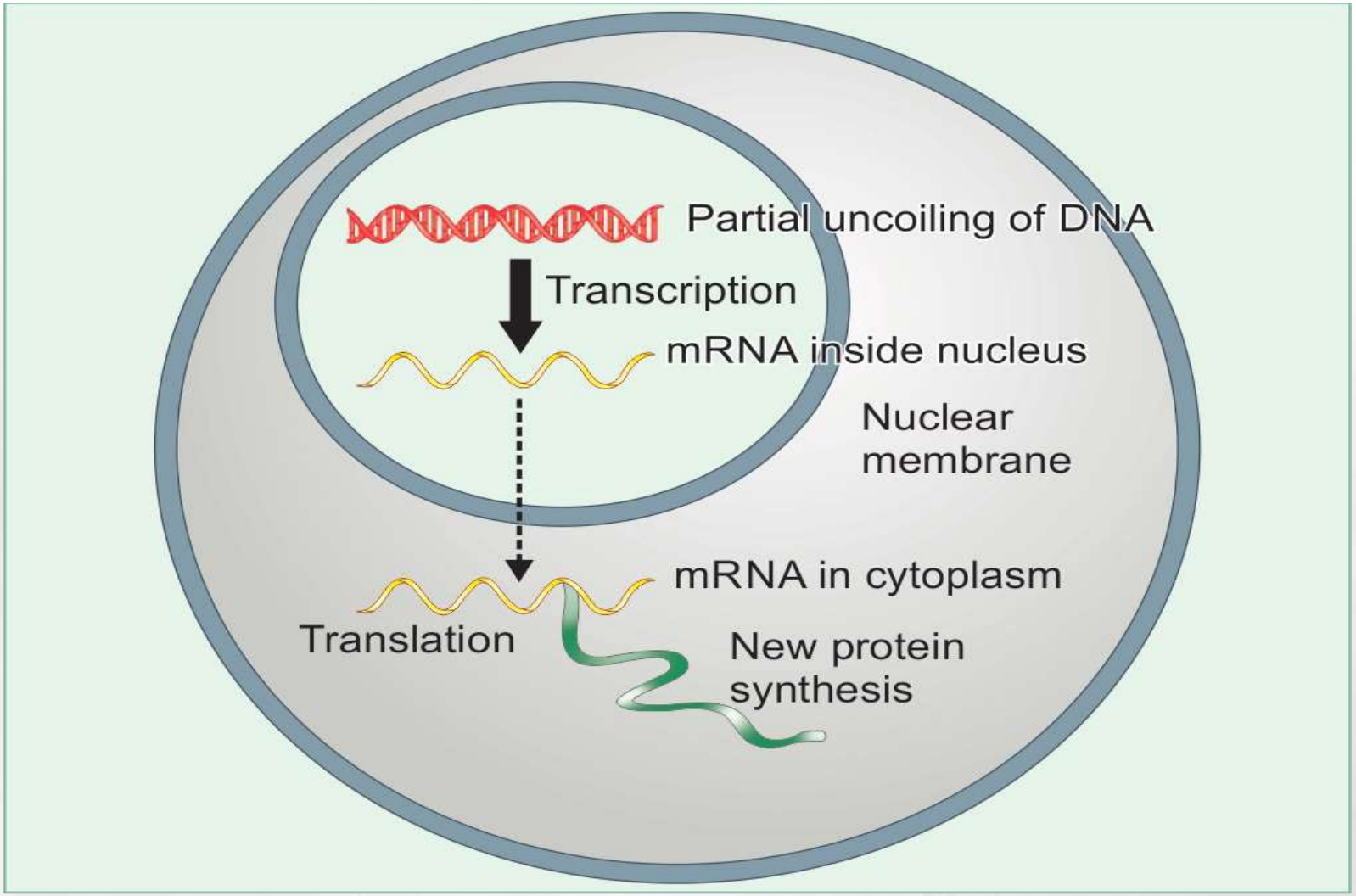
But transcription is taking place all the time. Only certain areas of the DNA are copied (selected regions on the sense strand). This is like taking xerox copy of particular page of the book.



Genetic information of DNA is transcribed (copied) to the messenger RNA (mRNA).
During transcription, the message from the DNA is copied in the language of nucleotides (4 letter language).

The mRNA then reaches the cytoplasm where it is translated into functional proteins.
During translation, the nucleotide sequence is translated to the language of amino acid sequence (20 letter language)



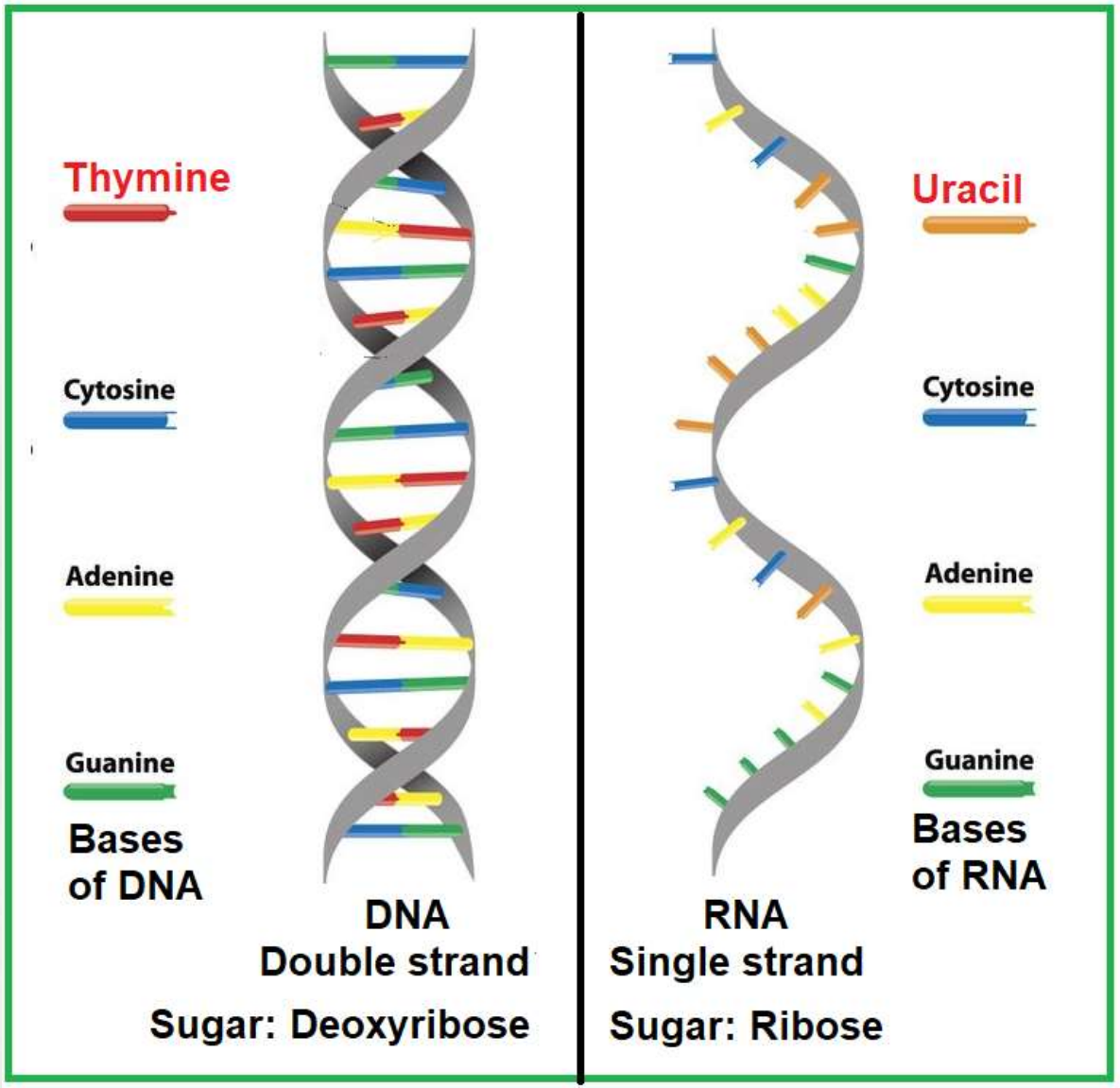


Central dogma of molecular biology.

Differences between RNA and DNA

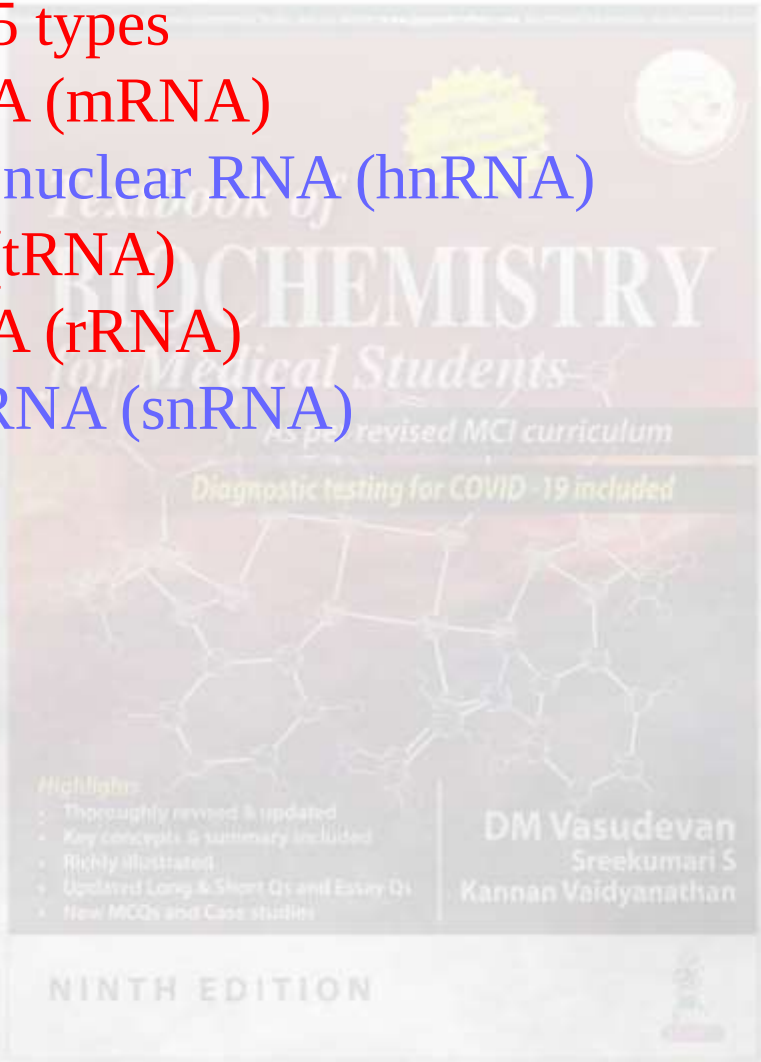
RNA	DNA
1. Mainly seen in cytoplasm	Mostly inside nucleus
2. Usually 100-5000 bases	Millions of base pairs
3. Generally single stranded	Double stranded
4. Sugar is ribose	Sugar is deoxyribose
5. Purines: Adenine, Guanine Pyrimidines: Cytosine, Uracil	Adenine, Guanine Cytosine, Thymine
6. Guanine content is not equal to cytosine and adenine is not equal to uracil	Guanine is equal to cytosine and adenine is equal to thymine
7. Easily destroyed by alkali	Alkali resistant

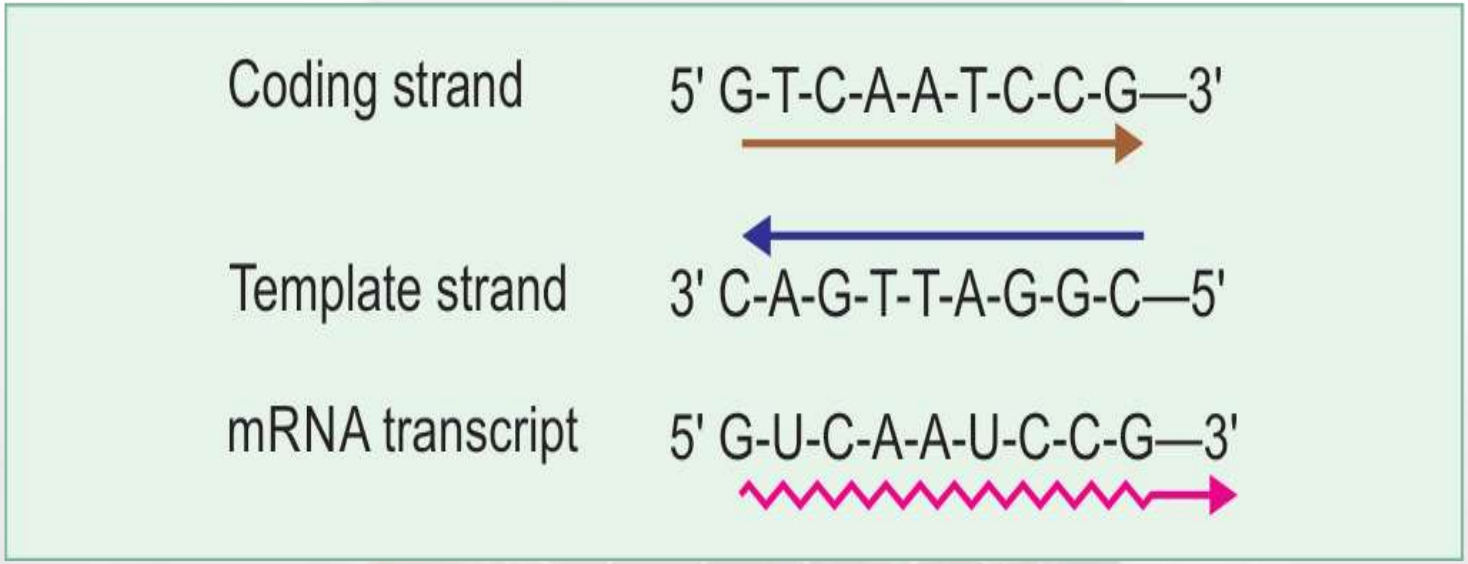
Differences between DNA and RNA



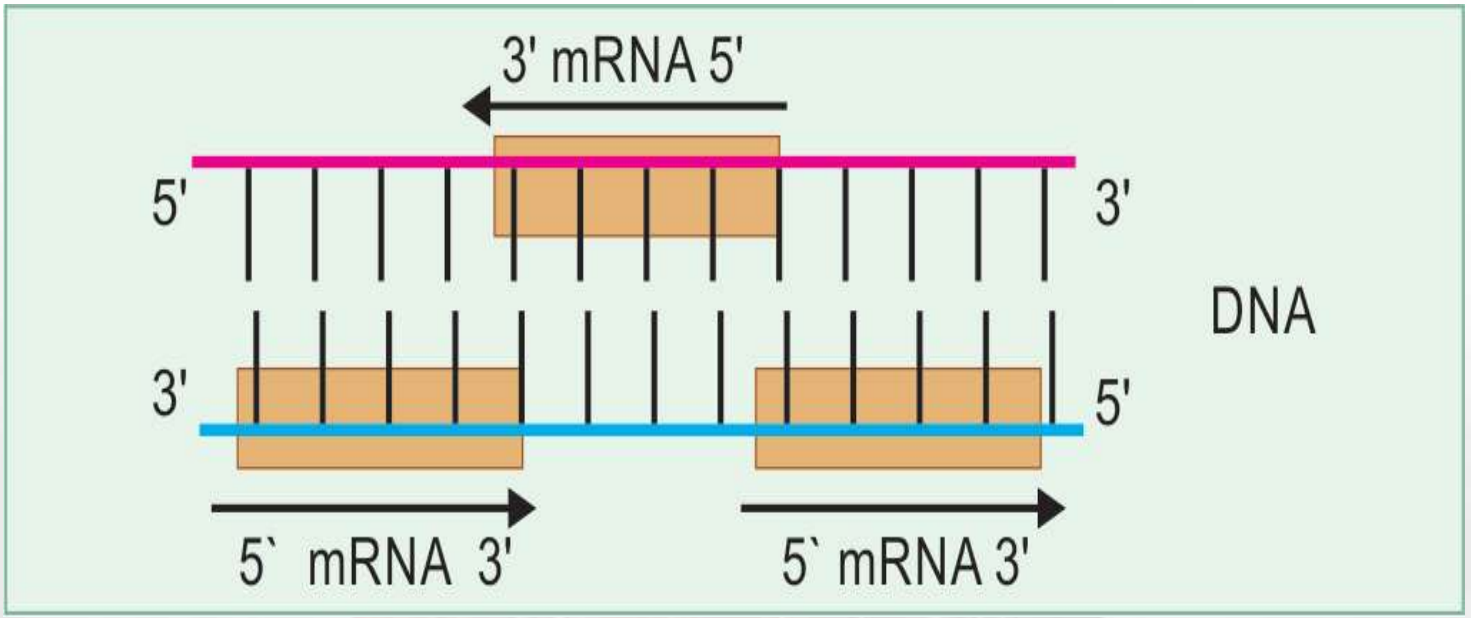
Cellular RNAs are 5 types

a) Messenger RNA (mRNA)
b) Heterogeneous nuclear RNA (hnRNA)
c) Transfer RNA (tRNA)
d) Ribosomal RNA (rRNA)
e) Small nuclear RNA (snRNA)

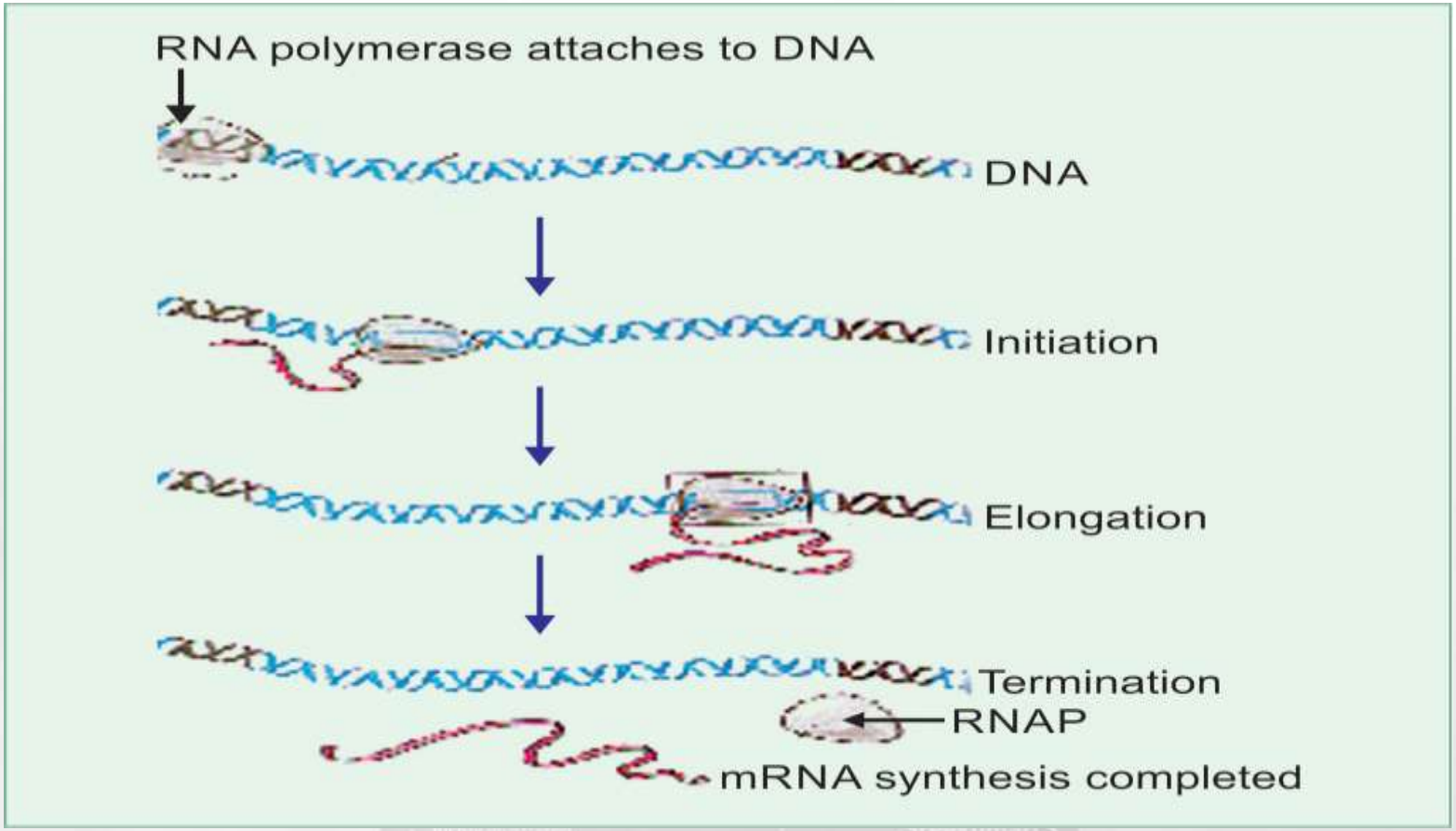




The mRNA base sequence is complementary to that of the template strand and identical to that of the coding strand.



Genes may be on any strand of DNA. Transcription is in 3' to 5' direction of the template.



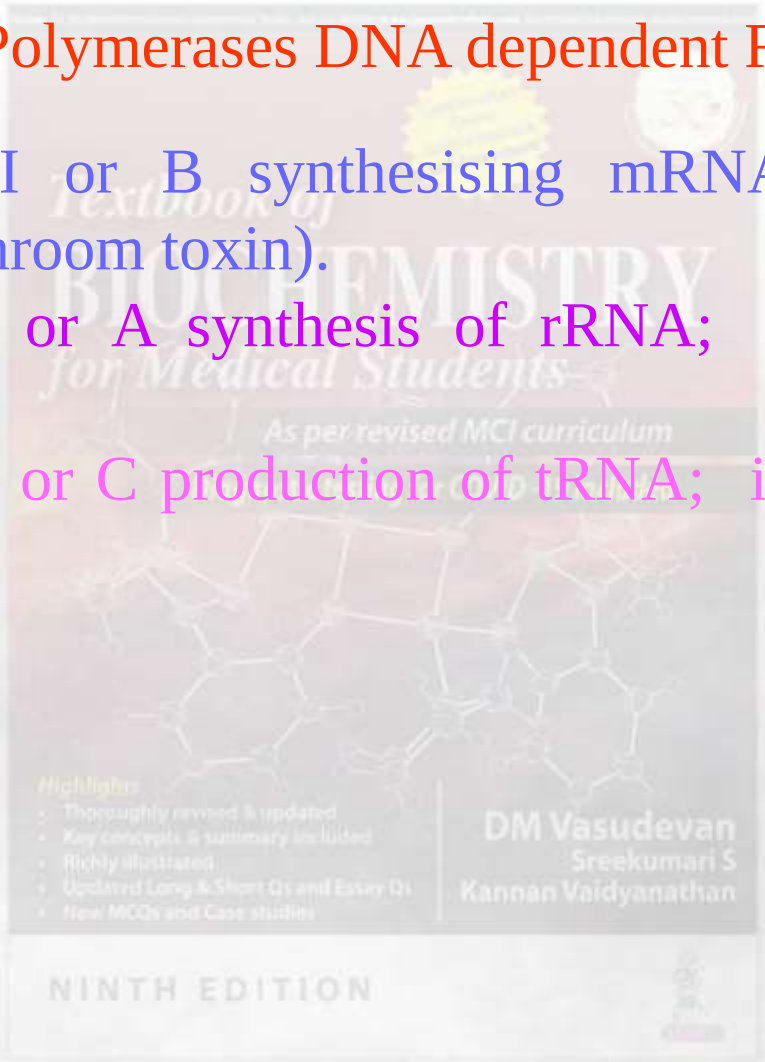
Transcription process.

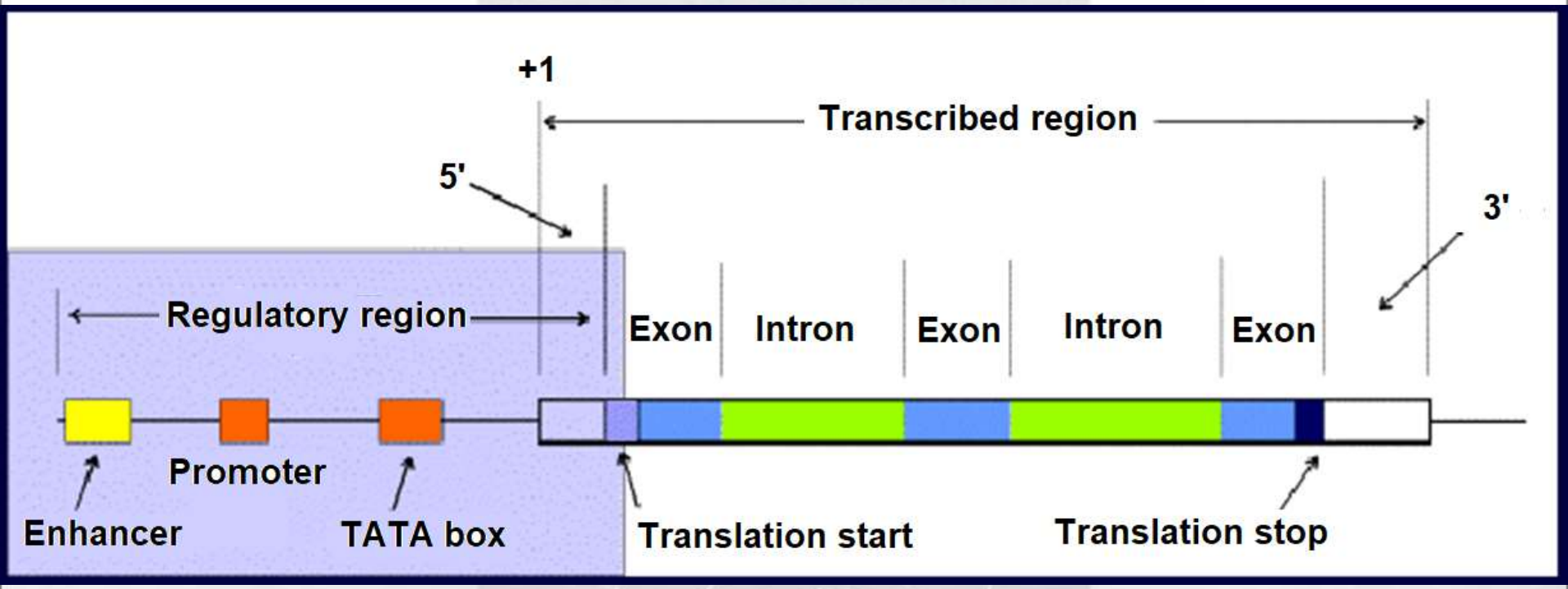
Mammalian RNA Polymerases DNA dependent RNA polymerases

i) RNAP type II or B synthesising mRNAs. Inhibited by amanitin (mushroom toxin).

ii) RNAP type I or A synthesis of rRNA; not inhibited by amanitin.

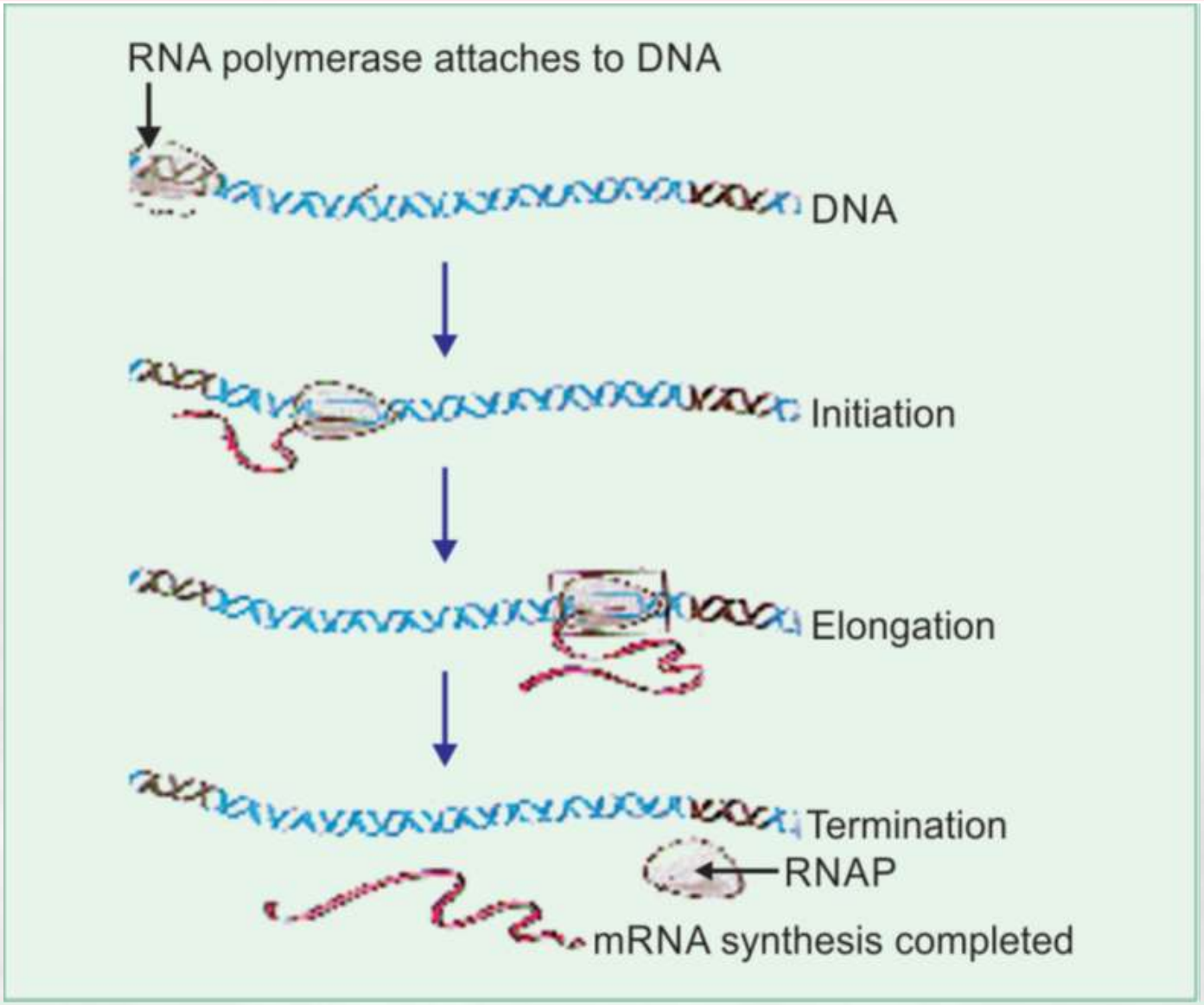
iii) RNAP type III or C production of tRNA; it is insensitive to amanitin.

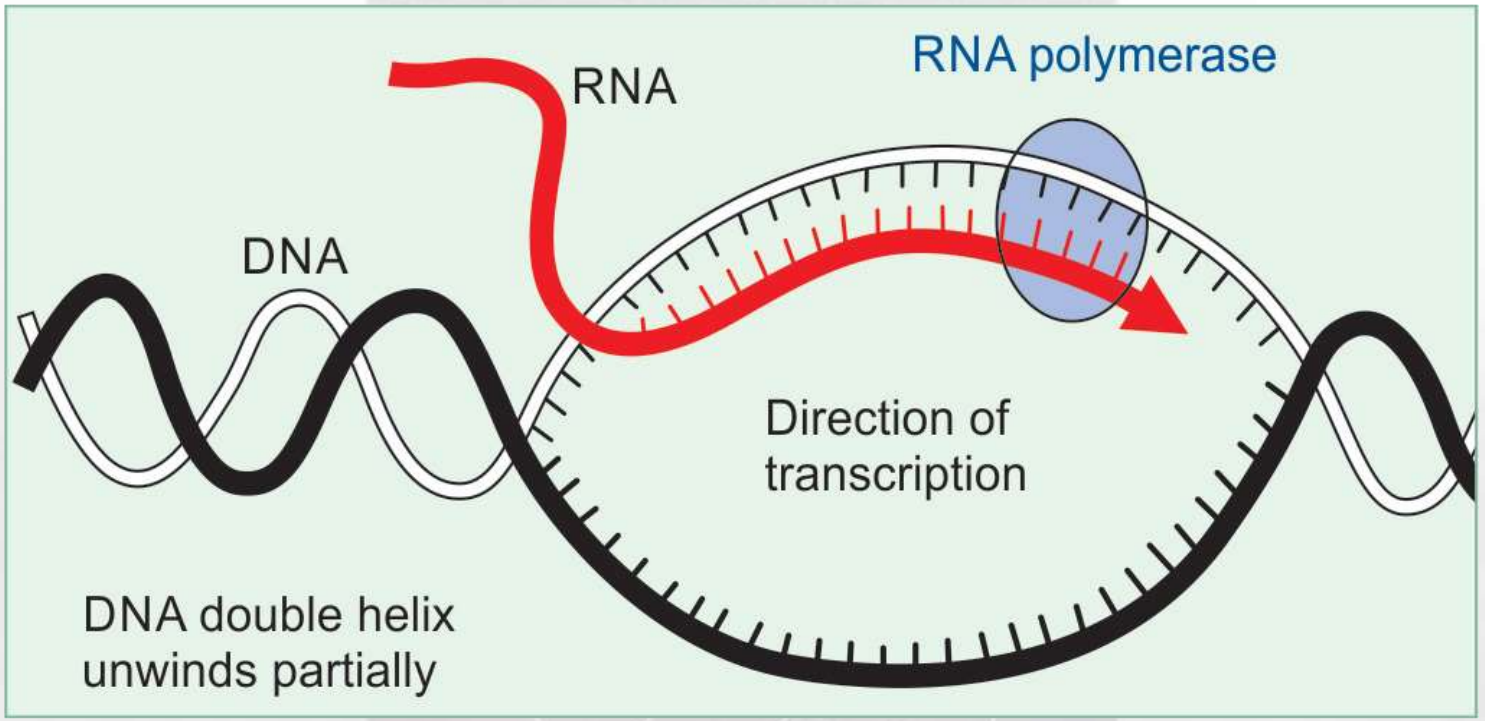




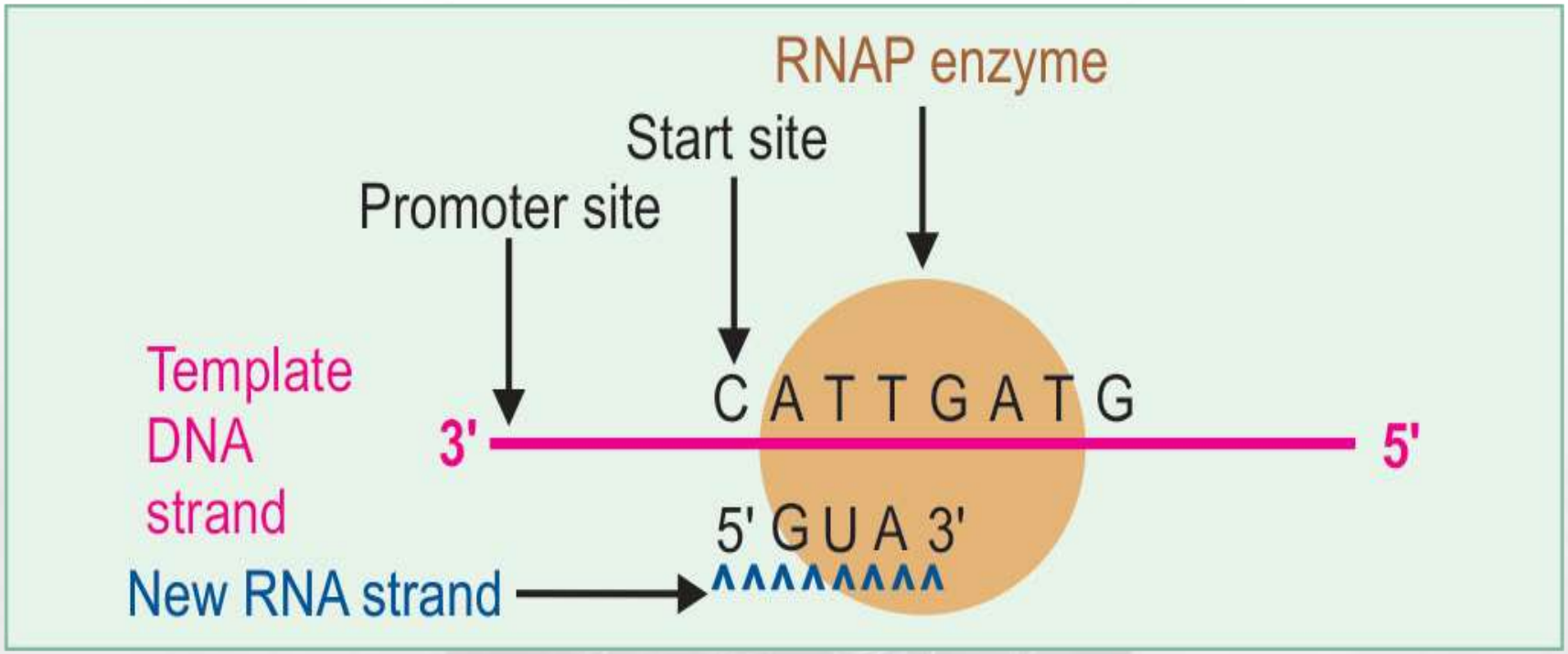
TATA box in relation to the introns.

Transcription Process

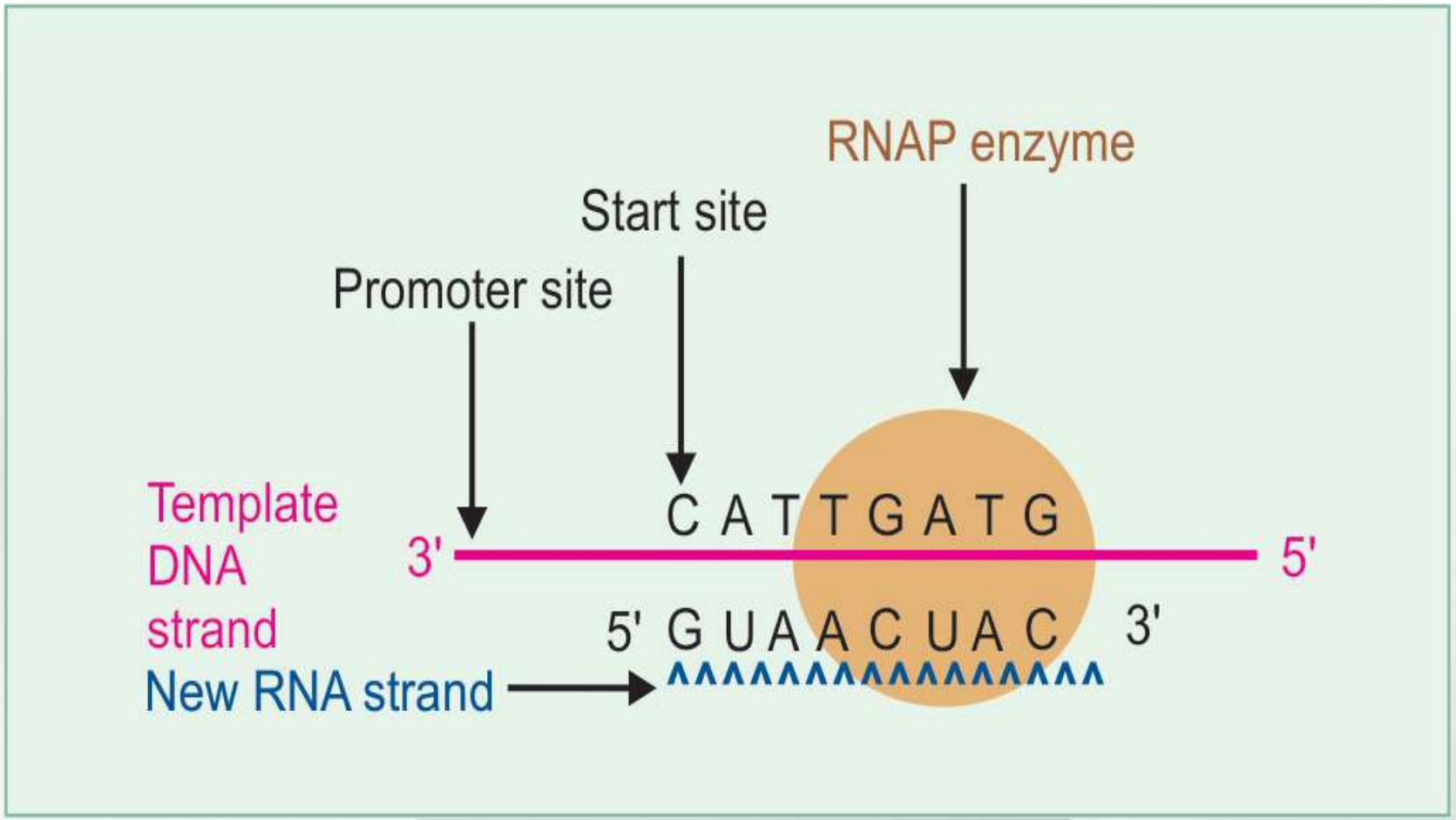




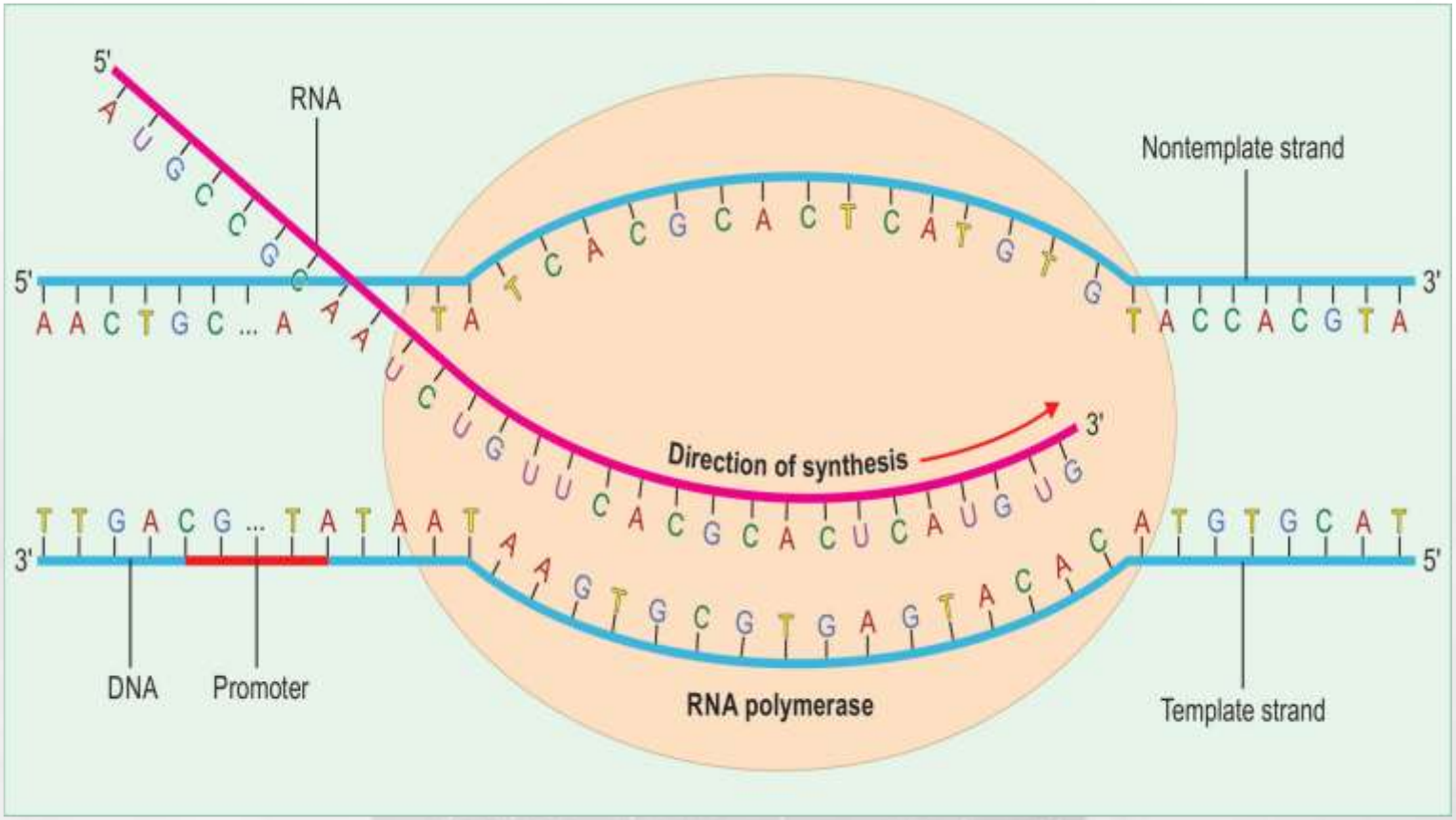
DNA double helix unwinds partially for transcription process



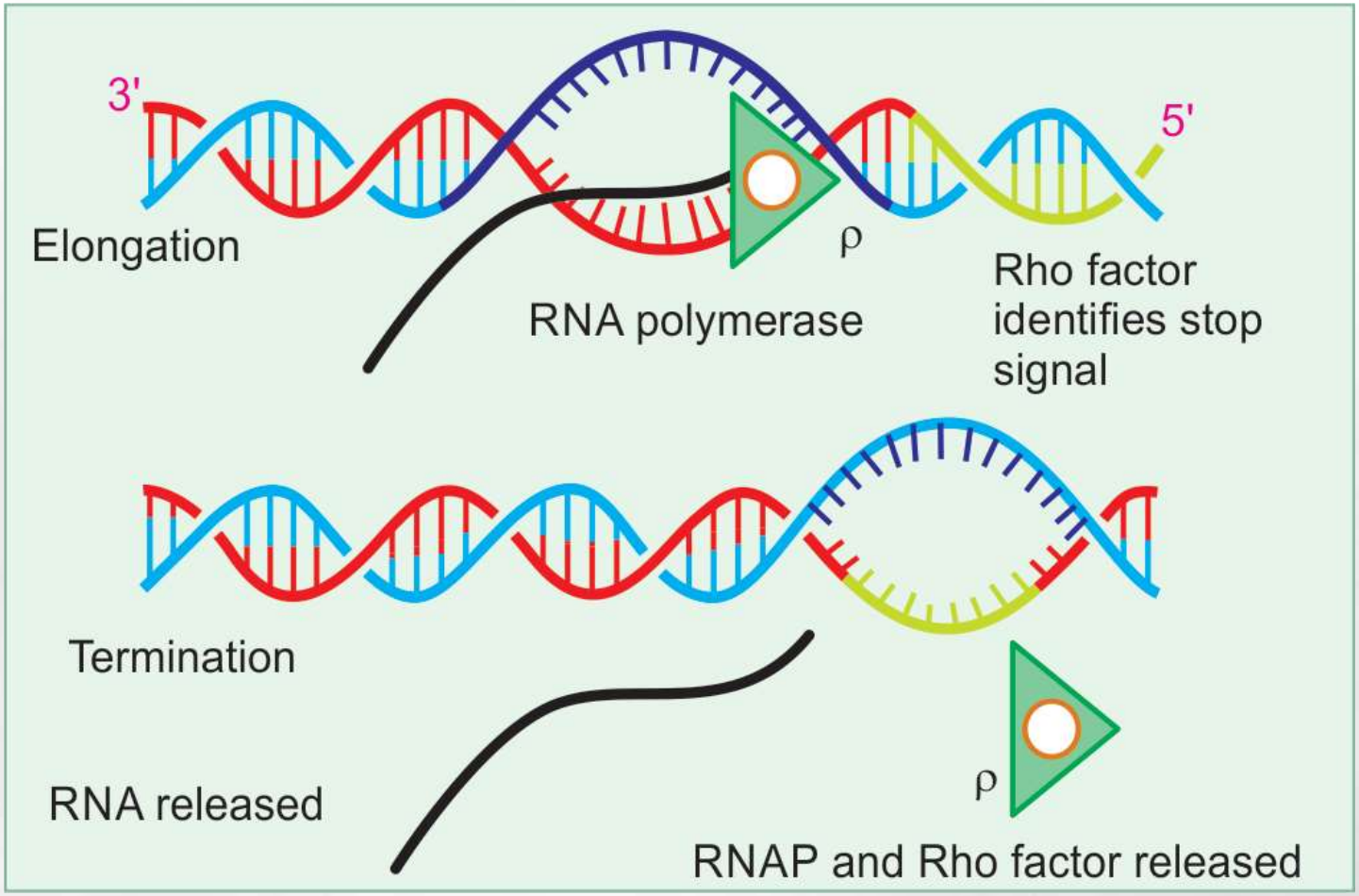
Initiation of transcription.



Elongation process of transcription.



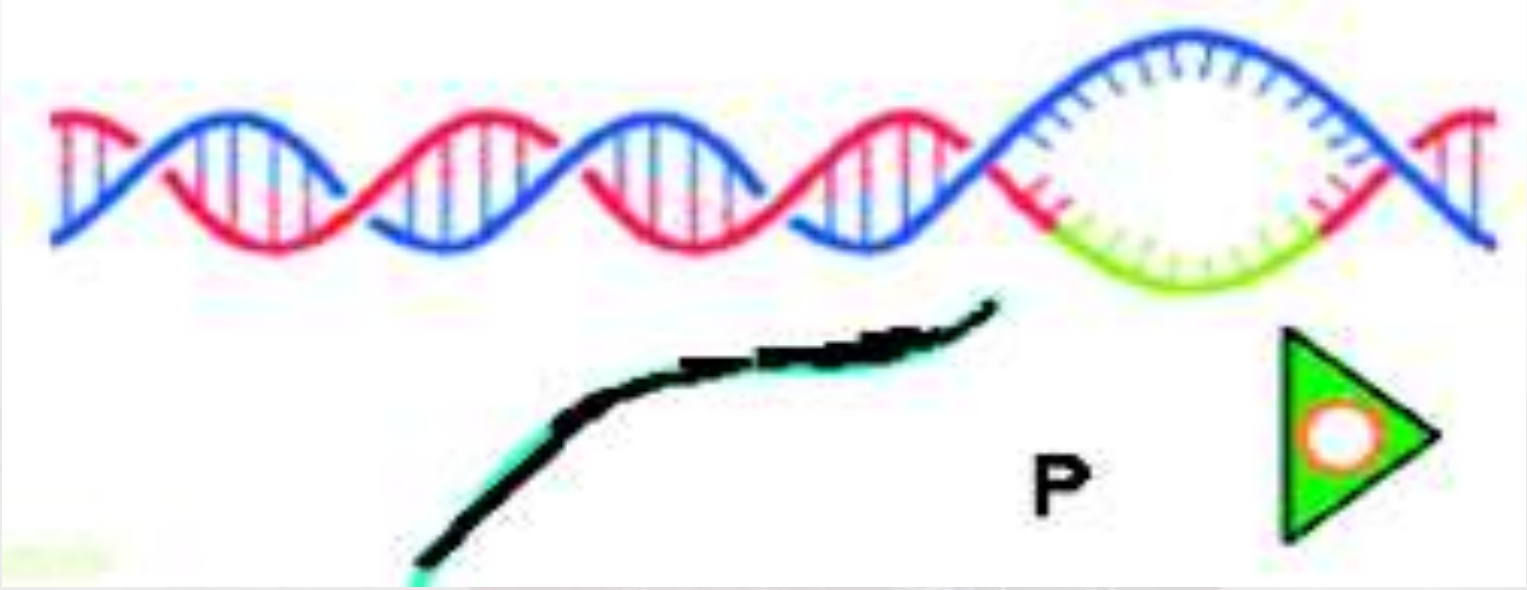
Transcription process.



Termination of transcription.

Transcription Termination

Mammals Rho independent termination



mRNA released

RNA polymerase

Differences between Prokaryotic and Eukaryotic Transcription

Feature	Prokaryotes	Eukaryotes
Site	Nucleus	Nucleus
Enzyme	DNA dependent RNA polymerase	RNAP II for mRNA transcription, RNAP1 for rRNA and RNAP3 for tRNA
Initiation	Sigma subunit binds to promoter site	The recognition of promoter requires transcription factors, TATA binding protein
Start signals	TATA box (Pribnow box)	Goldberg Hogness Box
Termination	Termination may be Rho factor dependent or Rho independent	Termination may be Rho factor dependent or independent. A lariat structure is formed and the RNAP dissociates.

Differences between Prokaryotic and Eukaryotic Transcription

Feature	Prokaryotes	Eukaryotes
Post-transcriptional Modifications	mRNA transcribed is immediately translated without any modification	mRNA undergoes several post-transriptional modifications like 5 methyl capping, poly A tailing, splicing etc. in the nucleoplasm .
Action of antibiotics	Inhibited by actinomycin and rifampicin	Inhibited by actinomycin and alpha amanitin

Differential Processing

Processing of hnRNA at different sites may produce different proteins.

Examples:

IgM molecules: Same lymphocyte produce 2 types of IgM-

- **Secretary type (Ms) with 20 hydrophilic aa in the end. So secreted out.**
- **Membrane bound (Mm) with 38 hydrophobic aa and hence anchored to membranes.**

Apo B-48 (intestine) and Apo B-100(liver)

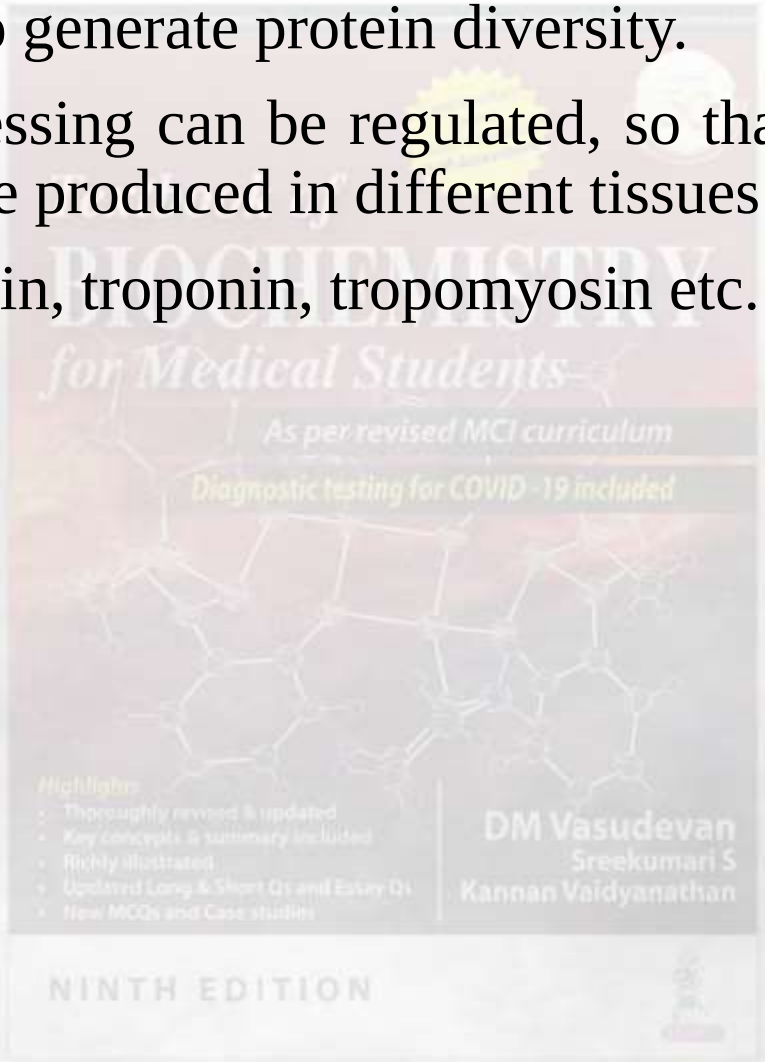
Both proteins are coded on the same gene

- Apo B-48 has 2152 N terminal a.acid
- Apo B-100 -4536 N terminal a.acid residues

Alternate Splicing

- Is a mechanism to generate protein diversity.
- Differential processing can be regulated, so that different forms of a protein can be produced in different tissues from same gene.

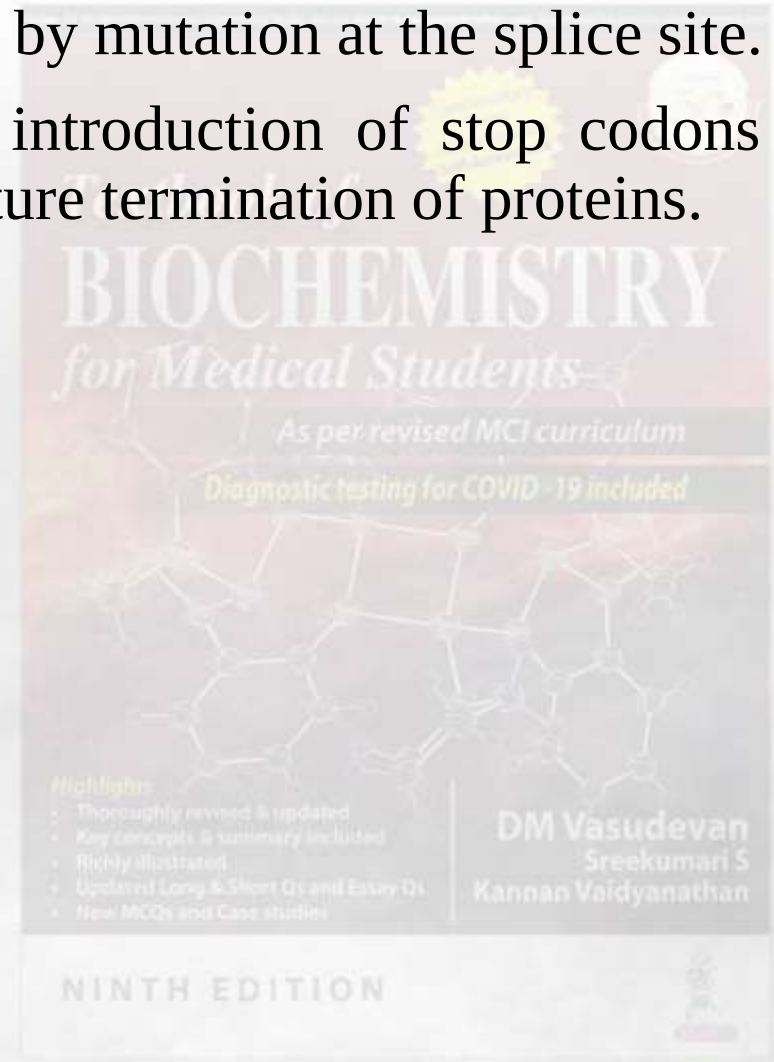
Eg. Actin, fibronectin, troponin, tropomyosin etc.



Aberrant Splicing

- May be produced by mutation at the splice site.
- This may result introduction of stop codons at abnormal sites producing premature termination of proteins.

  Eg. Thalassemias

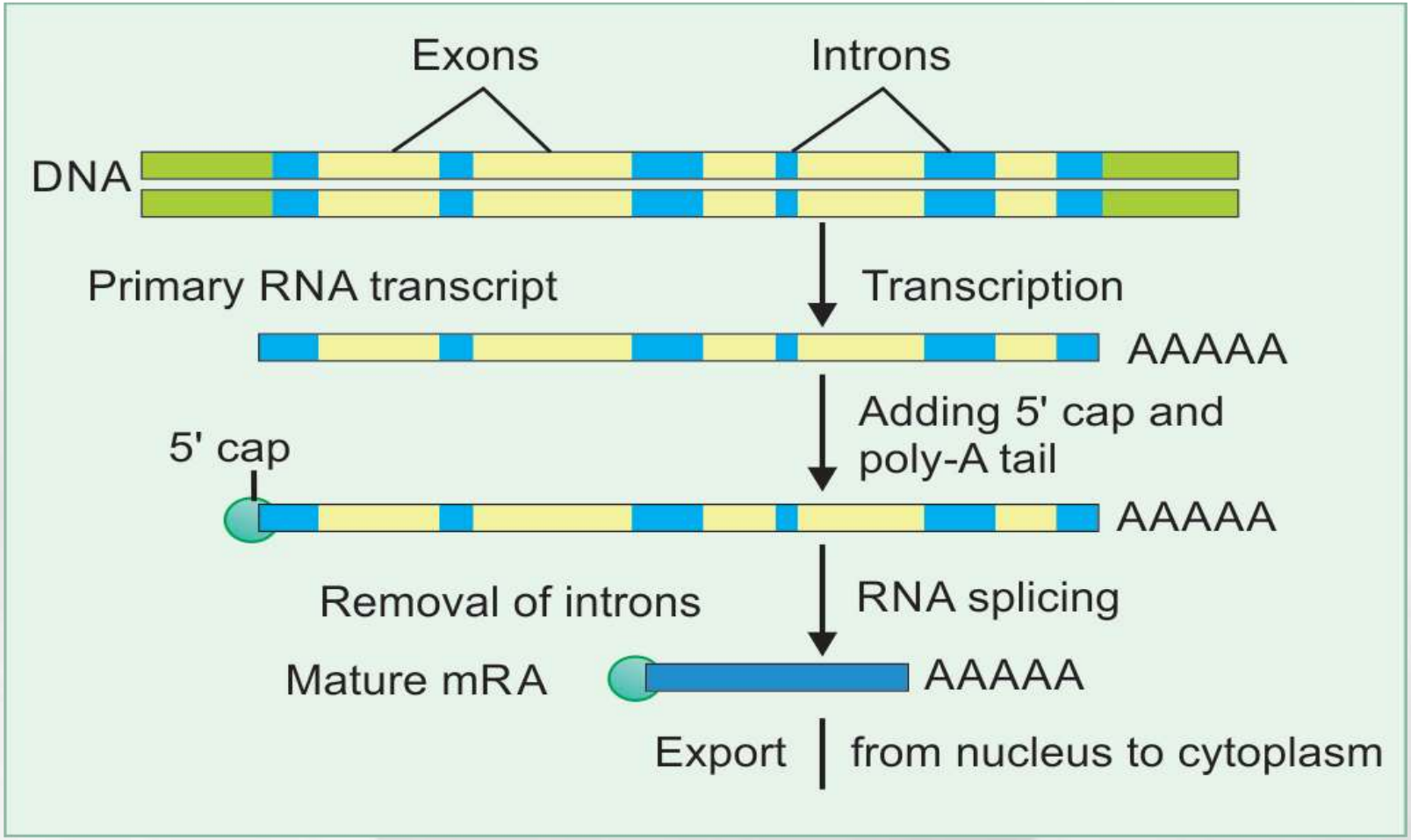


Post-transcriptional Processing

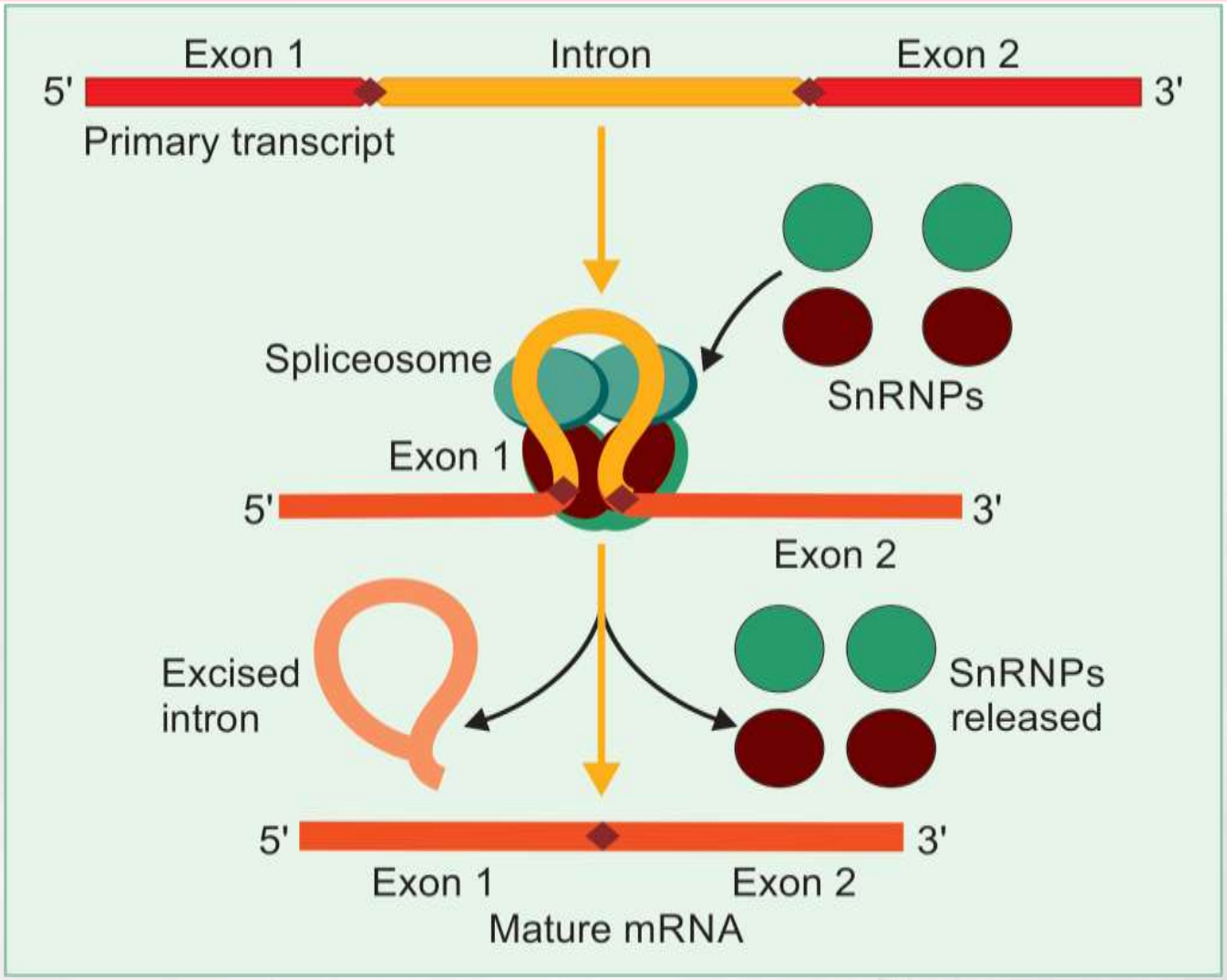
The mRNA formed and released from the DNA template is known as the **primary transcript**. It is also known as **heteronuclear mRNA** or hnRNA. In mammalian system, it undergoes extensive processing to become the **mature mRNA**. These modifications are:

- **Endonuclease cleavage**
- **Poly-A tailing**
- **5' capping**
- **Methylation**
- **Removal of introns**
- **Splicing of exons (connect together).**

The processing occurs mainly in the nucleoplasm. In bacteria, mRNA is not changed; and translation of mRNA starts even before completion of transcription. Posttranscriptional processing is not only for mRNA but for tRNA and rRNA as well.



Splicing process; removal of introns.



Excise of intron by the spliceosome.

Ribozymes

They are enzymes made up of RNA. Ribozymes or RNA enzymes are catalytic RNA molecules with sequence specific cleavage activity.

In the pre-cellular epoch, nucleic acids were biological catalysts; and in course of evolution, proteins took up this activity. In that sense, the ribozymes are vestigial remnants.

Spliceosomes contain ribozymes as well as protein components which serve to stabilise the structure of ribozymes.

RNAse-P is another ribozyme, which generates the ends of tRNAs.

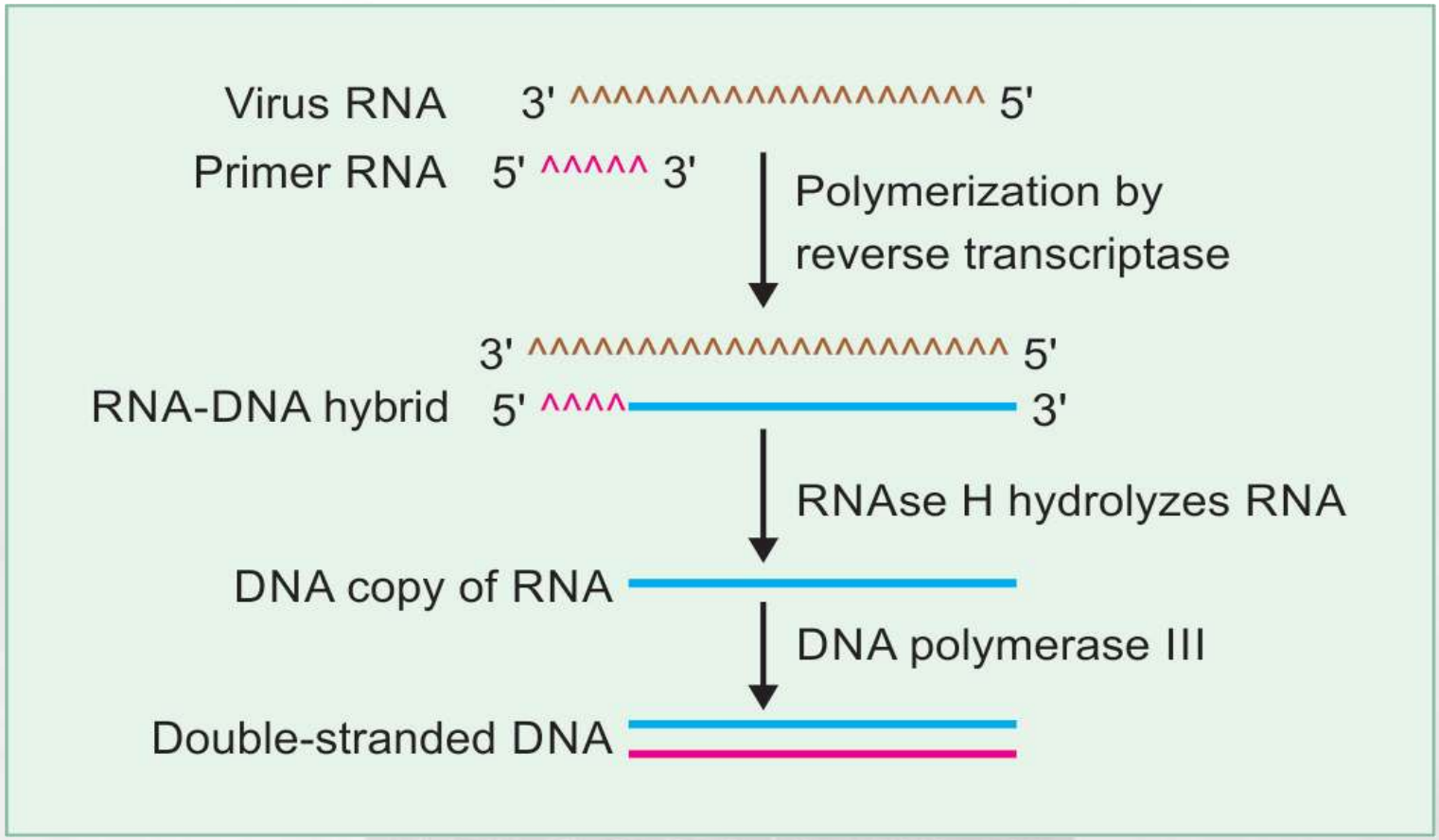
Peptidyl transferase present in ribosomes (used for protein biosynthesis) is another example of ribozyme.

Non-Coding Sequences

Even after editing, mature mRNA is 2 or 3 times bigger than the size required for coding the specific protein.

Long stretches of untranslated areas are present at both 3' and 5' ends. These are non-coding sequences (NCS).

NCS is different from introns. Introns are cleaved out; whereas, NCS are retained in mature mRNA.



Action of reverse transcriptase.

Inhibitors of RNA Synthesis

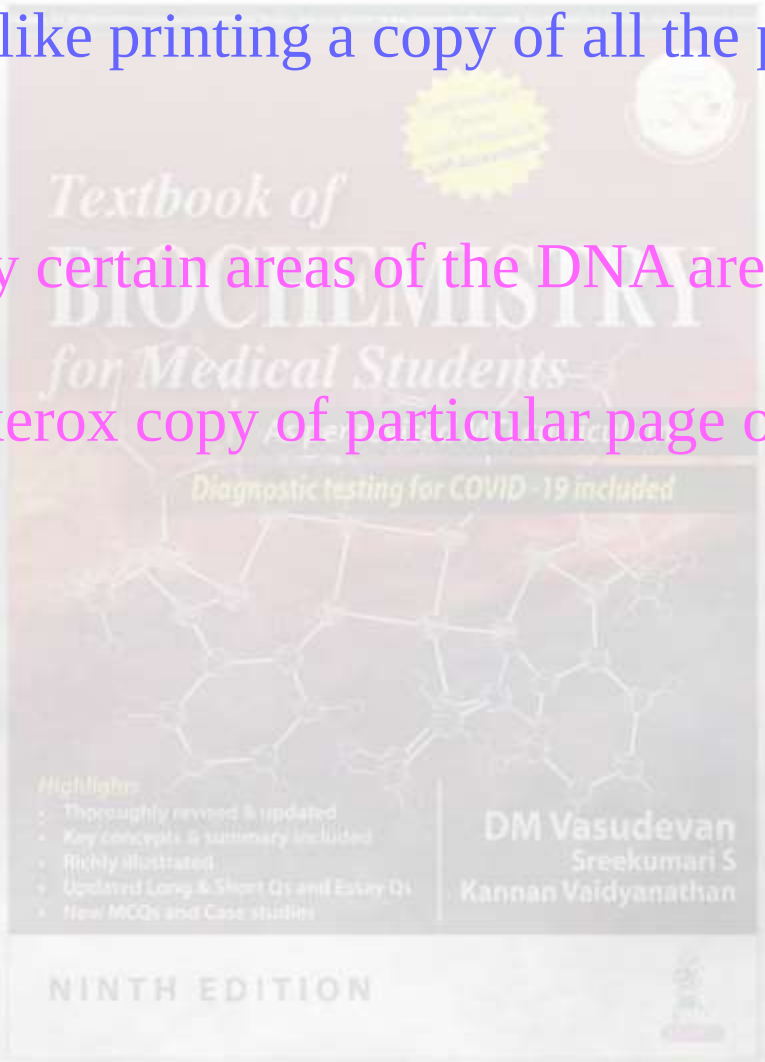
Inhibitor	Source	Mode of action
Actinomycin-D	Antibiotics from streptomyces	Insertion of phenoxazone ring between two G-C bp of DNA
Rifampicin	Synthetic derivative of rifamycin	Binds to beta subunit of RNA polymerase which is inactivated
Alpha amanitin	Toxin from mushroom	Inactivates RNA polymerase II
3'–deoxy adenosine	Synthetic analog	Incorrect entry into chain causing chain termination

Translation

DNA replication is like printing a copy of all the pages of a book.

In transcription only certain areas of the DNA are copied.

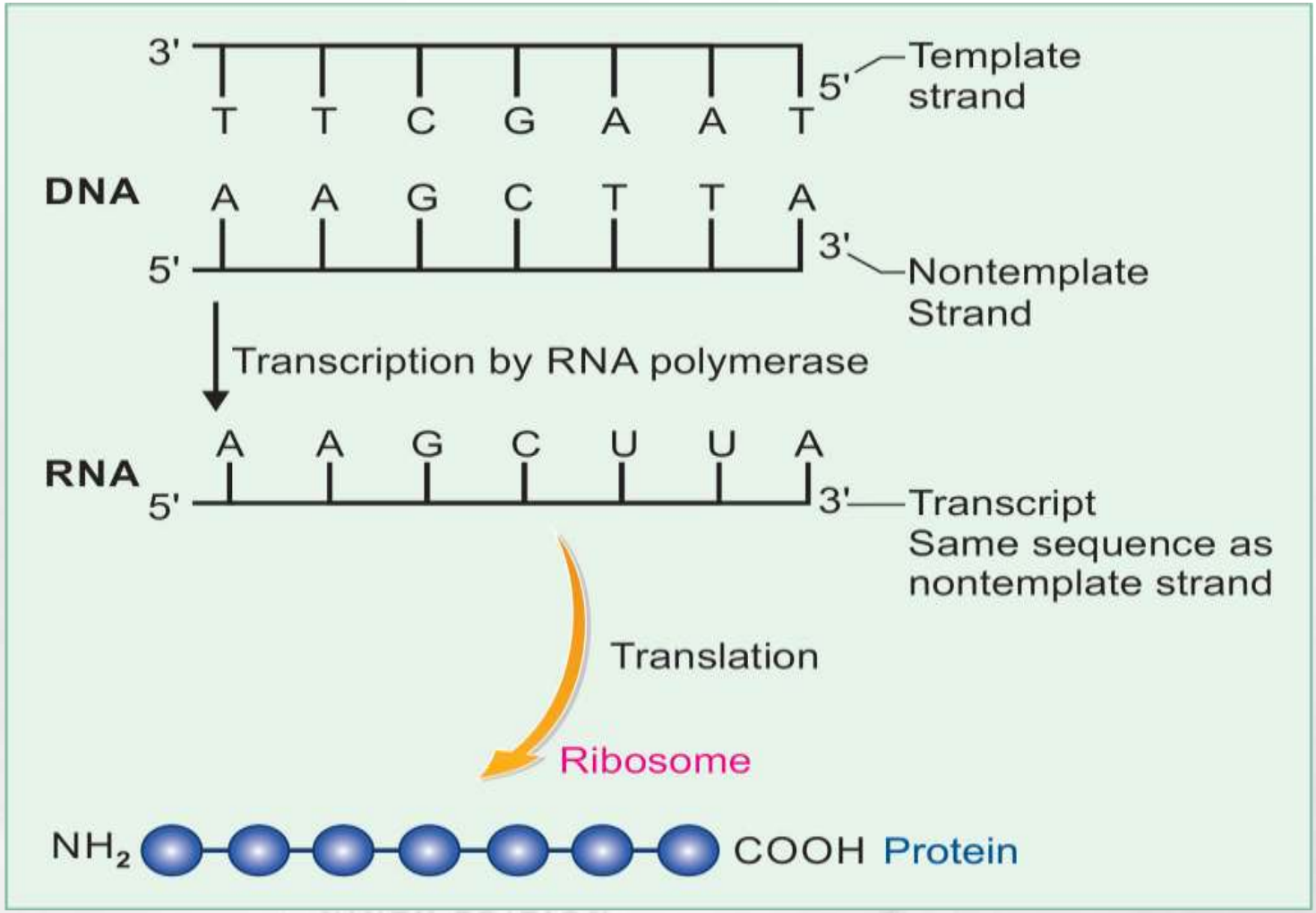
This is like taking xerox copy of particular page of the book.



Genetic information of DNA is transcribed (copied) to the messenger RNA (mRNA).
Then the mRNA is translated into functional proteins.

During transcription, the message from the DNA is copied in the language of nucleotides (4 letter language).

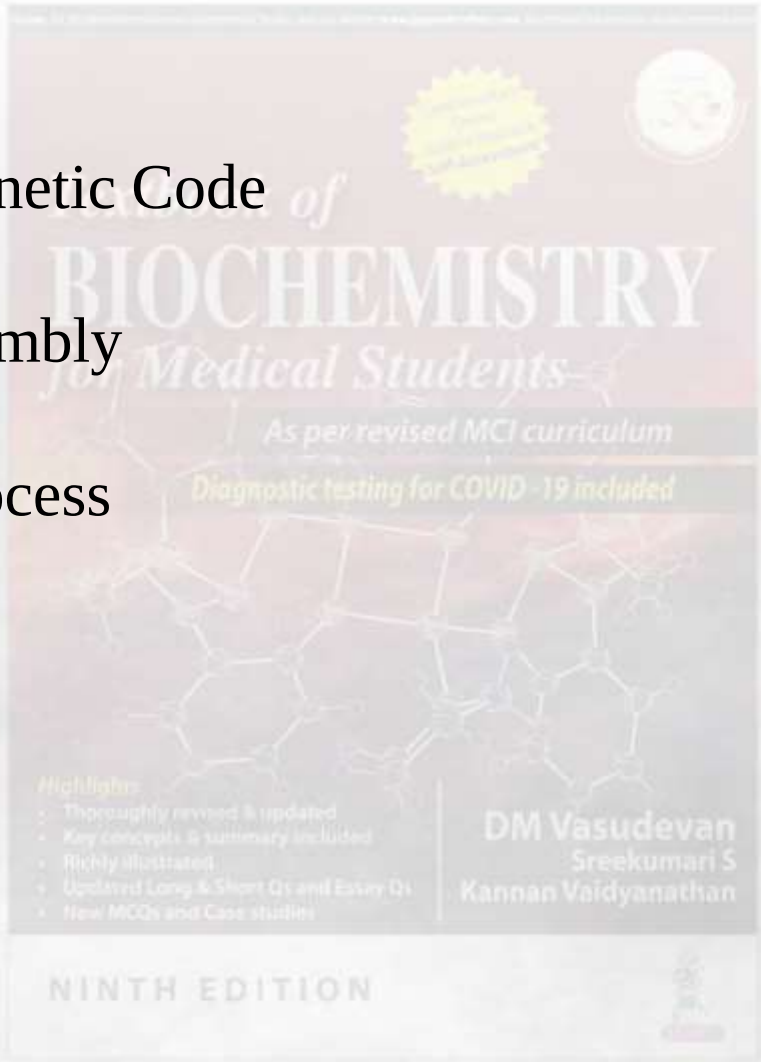
During translation, mRNA (nucleotide, 4 letter language) is translated into protein (amino acids, 20 letter language)

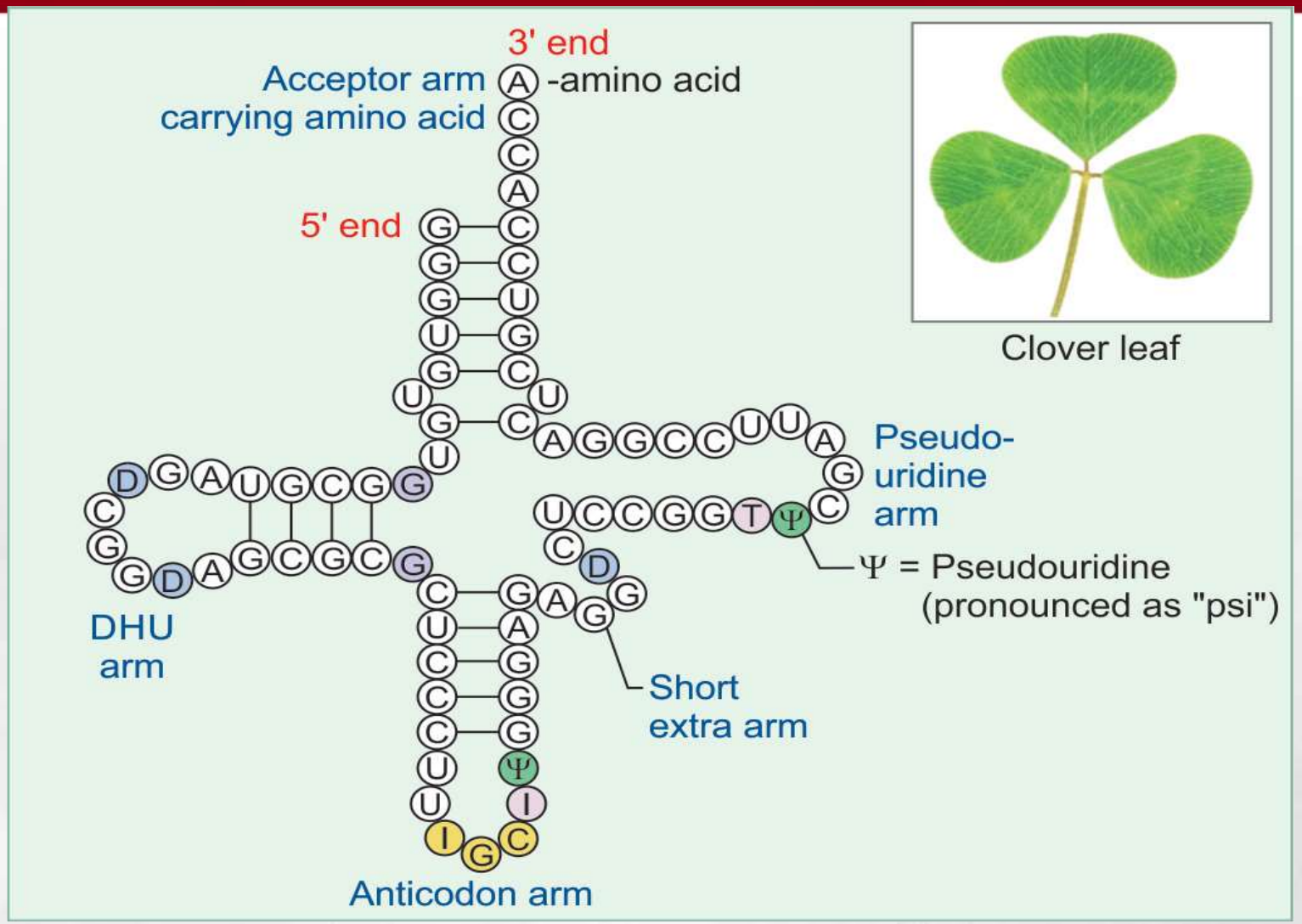


Expression of a gene into a protein.

Translation

1. Transfer RNA
2. mRNA and Genetic Code
3. Ribosome Assembly
4. Translation Process



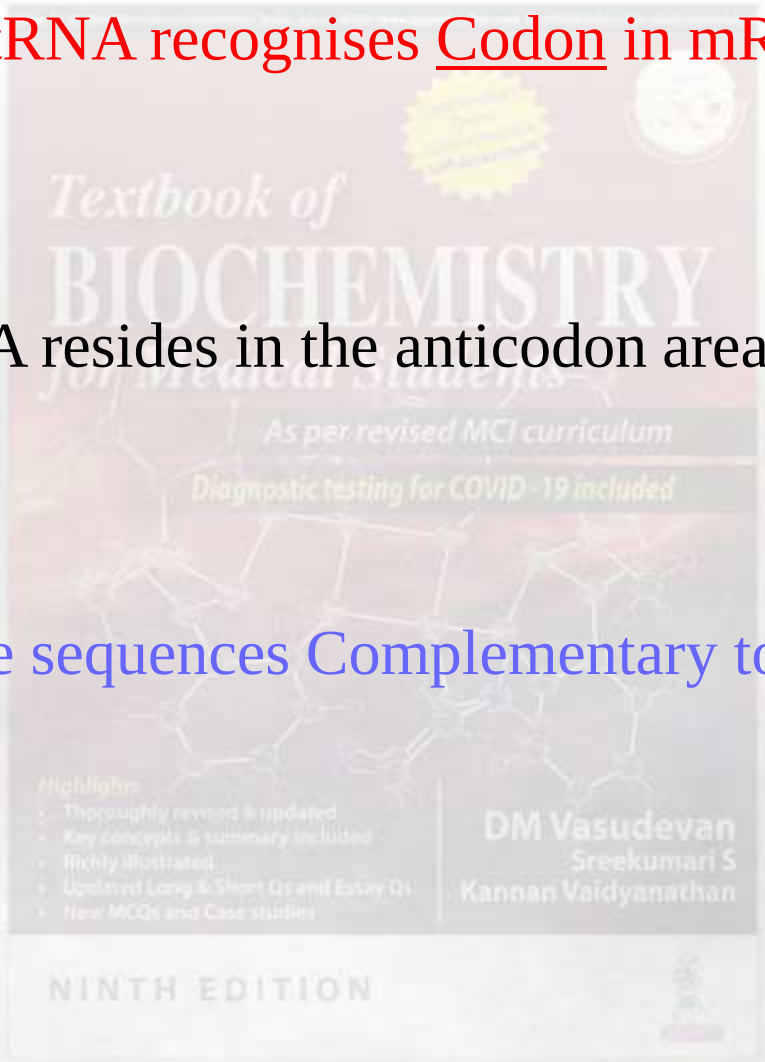


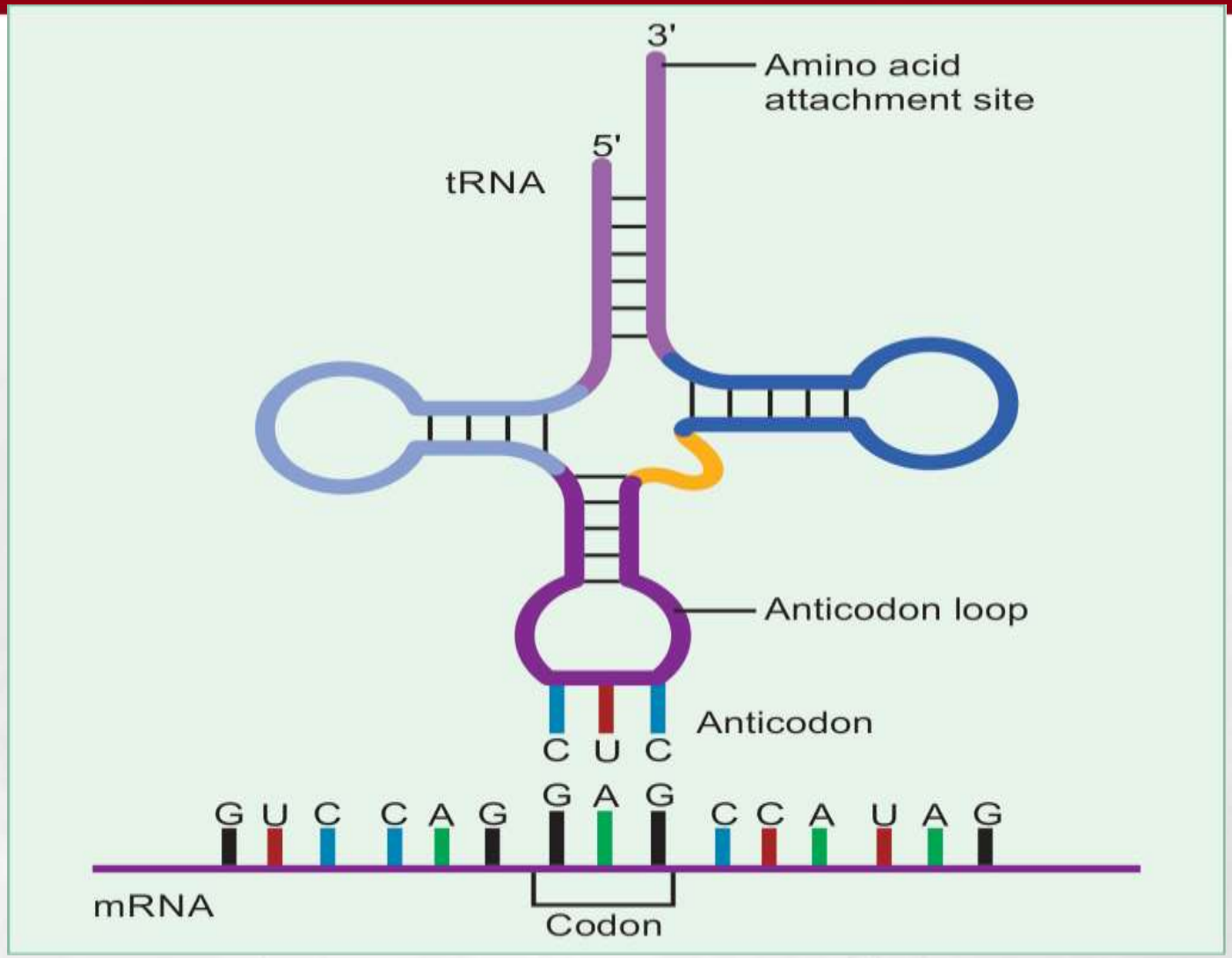
Transfer RNA (tRNA) general structure.

Anticodon arm of tRNA recognises Codon in mRNA

Specificity of tRNA resides in the anticodon area

Anticodon has base sequences Complementary to that of Codon in mRNA

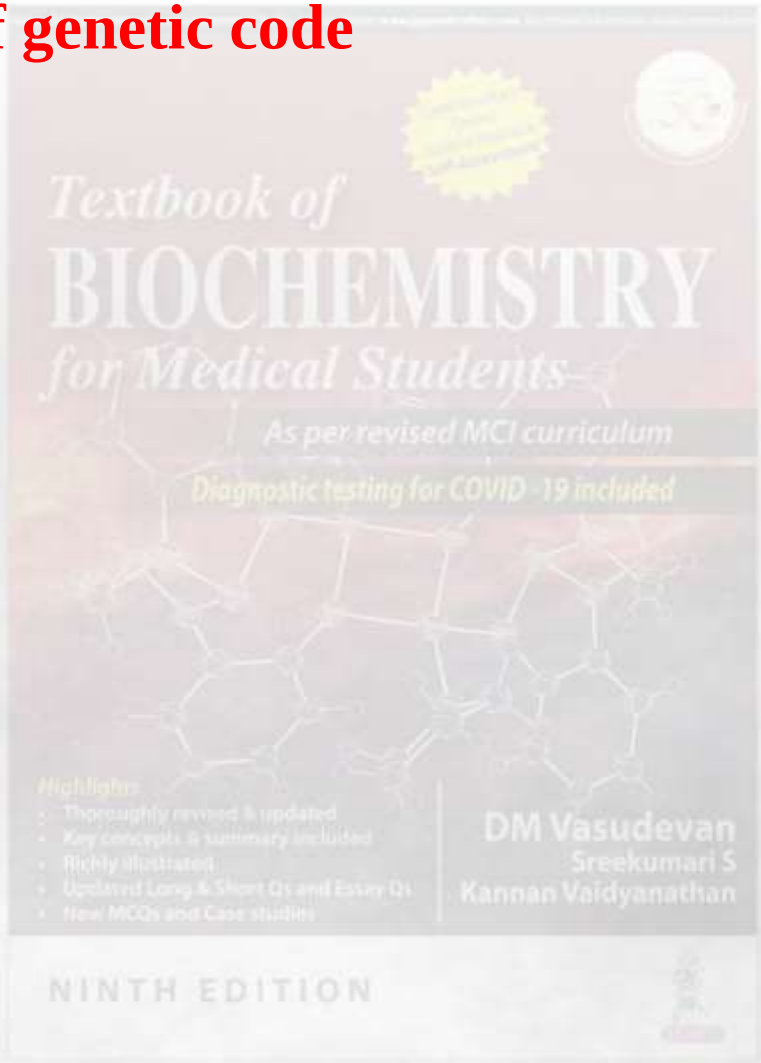




The anticodon on tRNA correctly fits on the codon on mRNA.

Salient features of genetic code

- Triplet codon
- Universal
- Degenerate
- Unambiguous
- Nonoverlapping
- Nonpunctuated



Salient features of Genetic Code

i) Triplet Codons

The codes are on the mRNA. Each codon is a consecutive sequence of three bases on the mRNA

U U U - Phenyl alanine
A U G - Methionine

ii) Non-overlapping

The codes are consecutive. Therefore the starting point is extremely important.
The codes are read one after another in a continuous manner.

AUG, CAU, CAU, GCA, ……

Salient features of Genetic Code

iii) Non-punctuated

There is no punctuation between the codons.

It is consecutive or continuous

AUG,CAU,CAU,GCA, ……

UGC,AUC,AUG,CA……..

.

iv) Degenerate

61 codes; but 20 amino acids.

One amino acid has more than one codon.

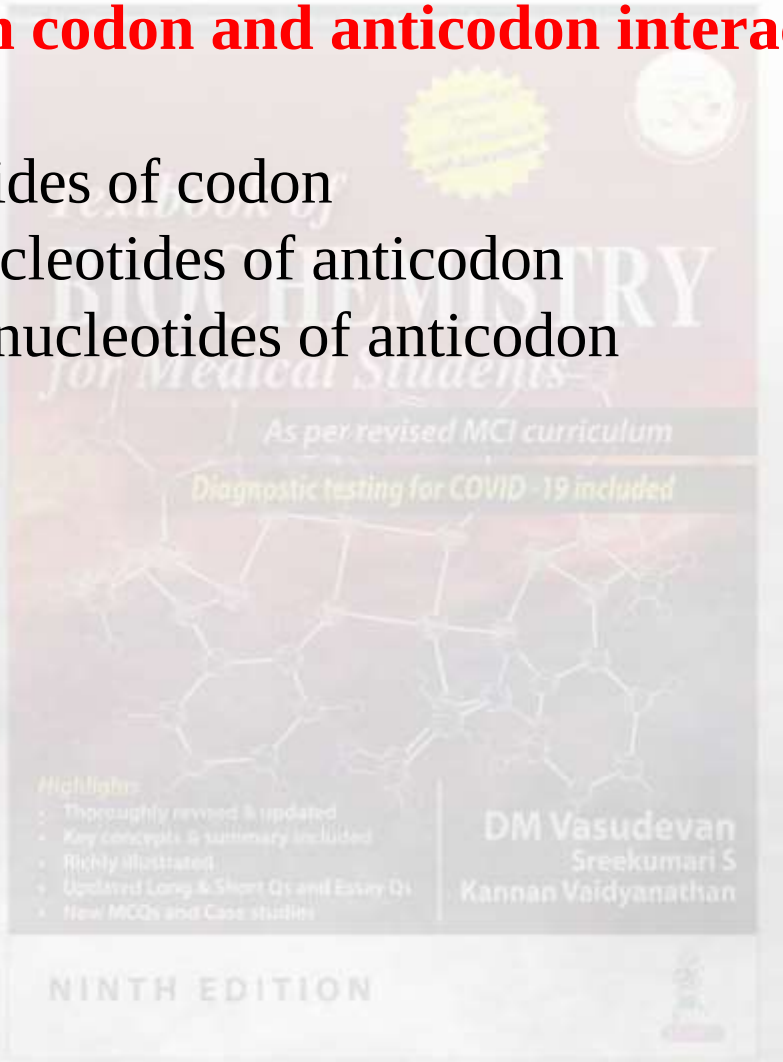
Serine has 6 codons

Glycine has 4 codons .

This is degeneracy of the code.

Wobble positions in codon and anticodon interactions

C A G U - Nucleotides of codon
G U C A - Usual nucleotides of anticodon
I I U G - Wobble nucleotides of anticodon



Salient features of Genetic Code

v) Unambiguous

without any doubtful meaning.

One codon stands only for one amino acid.

61 codes; but 20 amino acids.

vi) Universal

The codons are the same for the same amino acid in all species.

"Elephant and E.coli".

vii) Terminator Codons UAA, UAG, and UGA.

ix) Initiator Codon AUG acts as the initiator codon.

Codes are different in cytoplasm and mitochondria

Codon	Translated in cytoplasm	Translated in mitochondria
AUA	Isoleucine	Methionine
UGA	Termination	Tryptophan
AGA	Arginine	Termination
AGG	Arginine	Termination

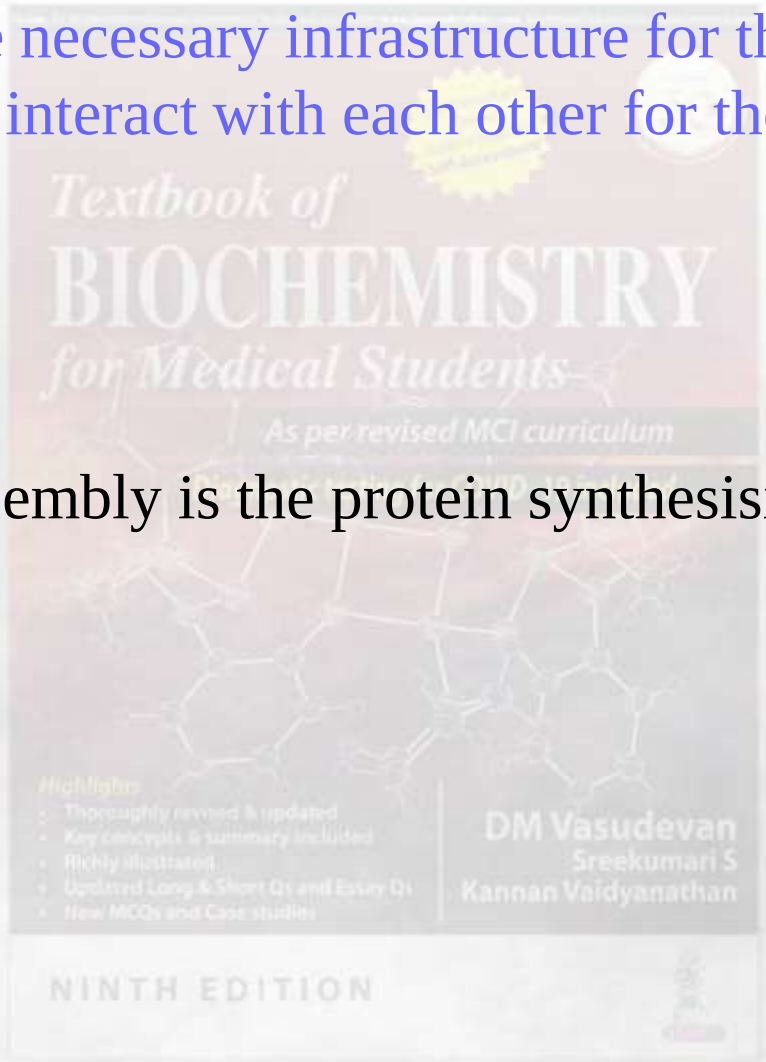
Translation Process

Translation is a cytoplasmic process. The mRNA synthesized in the nucleus is transported to the cytoplasm, where the mRNA is translated from **5' to 3' end**. In the polypeptide chain synthesized, the first amino acid is the amino terminal one. The chain growth is from amino terminal to carboxyl terminal. The process of translation can be conveniently divided into the following five phases:

1. **Activation of amino acid**
2. **Initiation**
3. **Elongation**
4. **Termination and**
5. **Post-translational processing.**

Ribosomal RNA (rRNA)

Ribosomes provide necessary infrastructure for the mRNA, tRNA and amino acids to interact with each other for the translation process.

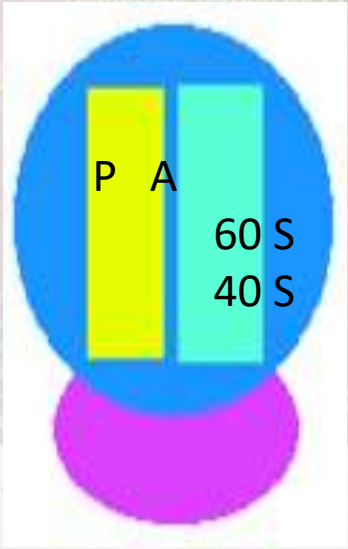


Thus ribosomal assembly is the protein synthesising machinery.

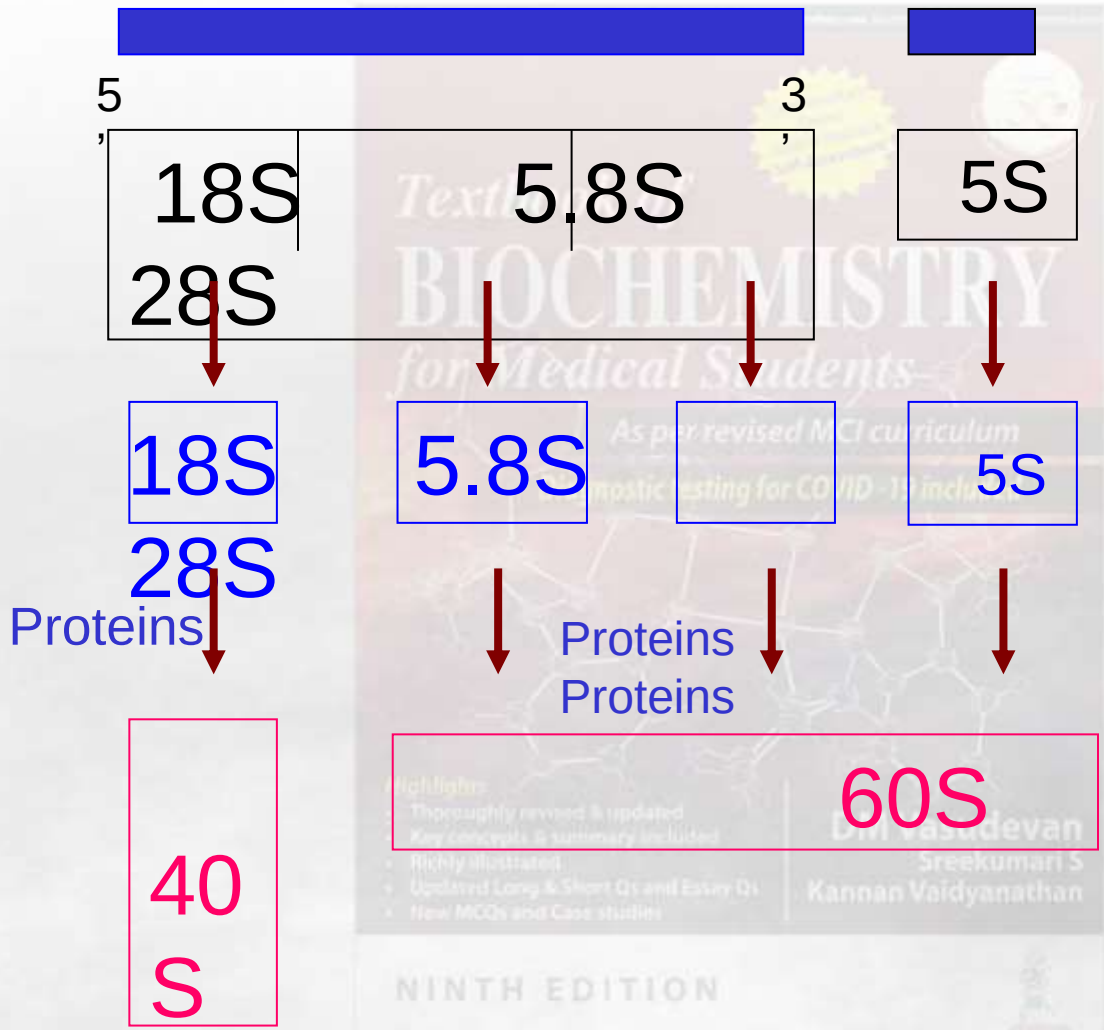
Components of rRNA

The mammalian ribosome has a sedimentation velocity of 80S

Larger 60S subunit and smaller 40S subunit.
They contain different rRNAs and specific proteins.



rRNA Processing-Synthesized in Nucleolus



Differences between Eukaryotic and Prokaryotic Ribosomes

		Eukaryotes	Bacteria
Whole Ribosome	Sedimentation coefficient	80 S	70 S
Large subunit	Sedimentation coefficient	60 S	50 S
	Proteins	47	33
	rRNAs	28 S rRNA 5.8 S rRNA	23S rRNA 5S rRNA
Small subunit	Sedimentation coefficient	40 S	30 S
	No. of Proteins	32	20
	rRNAs	18S rRNA	16S rRNA

4. Translation Process

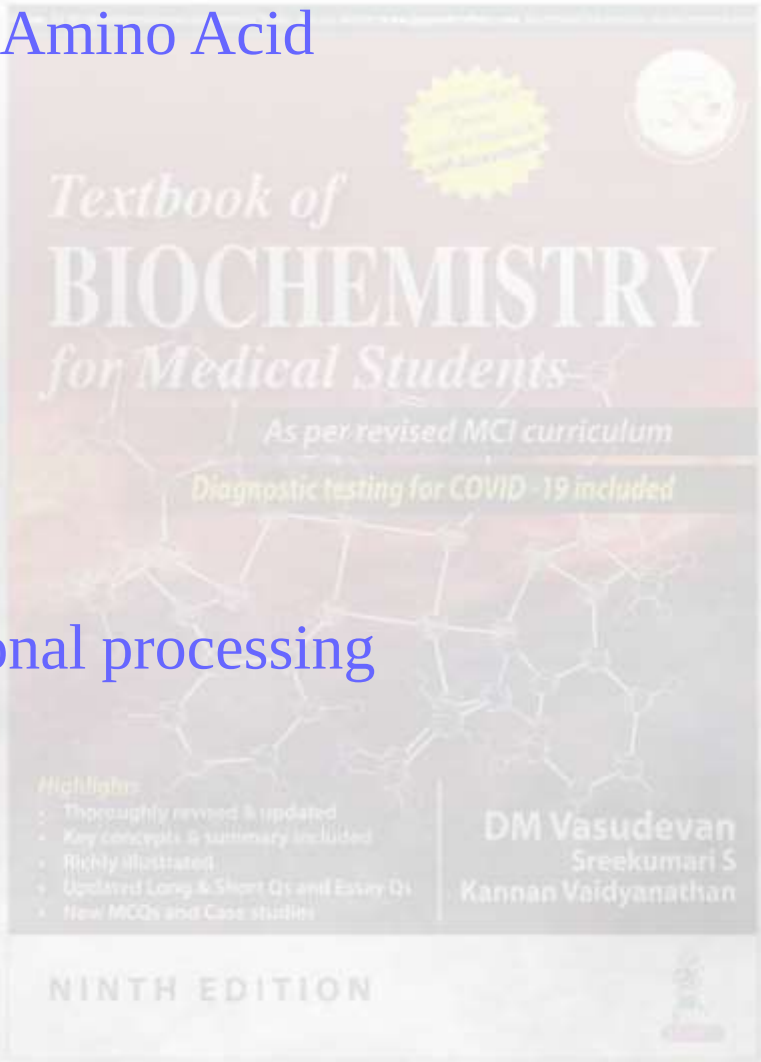
4-A. Activation of Amino Acid

4-B. Initiation

4-C. Elongation

4-D. Termination

4.E. Post-translational processing



Activation of Amino Acid (Charging Reaction)

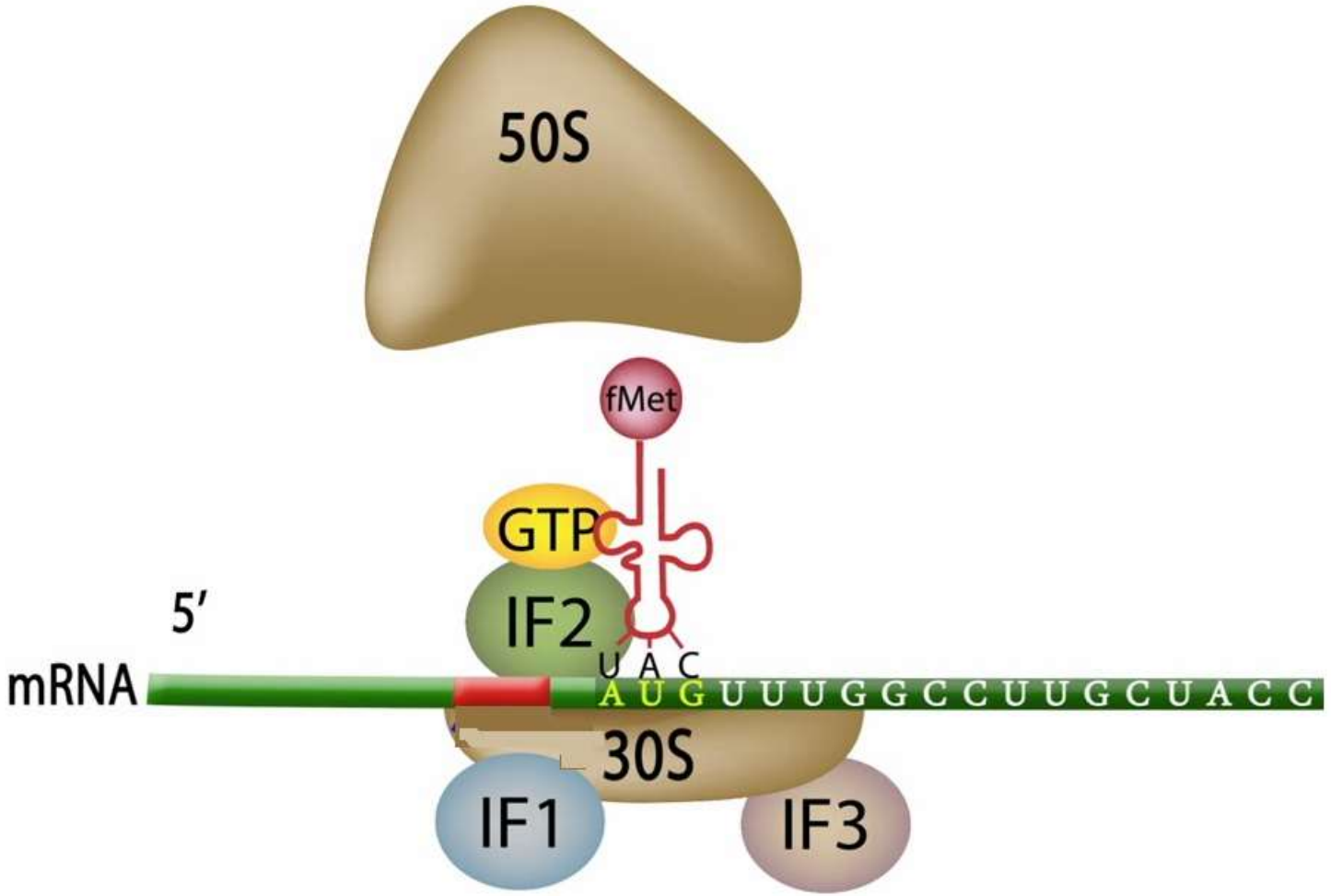
The enzymes **aminoacyl tRNA synthetases** activate the amino acids. There is at least one tRNA for each of the 20 amino acids. The D arm of tRNA is very important for the recognition by the enzyme. The CCA 3' terminus of the acceptor arm carries amino acid.

Amino acid is first activated with the help of ATP. Then the carboxyl group of the amino acid is esterified with 3' hydroxyl group of tRNA.

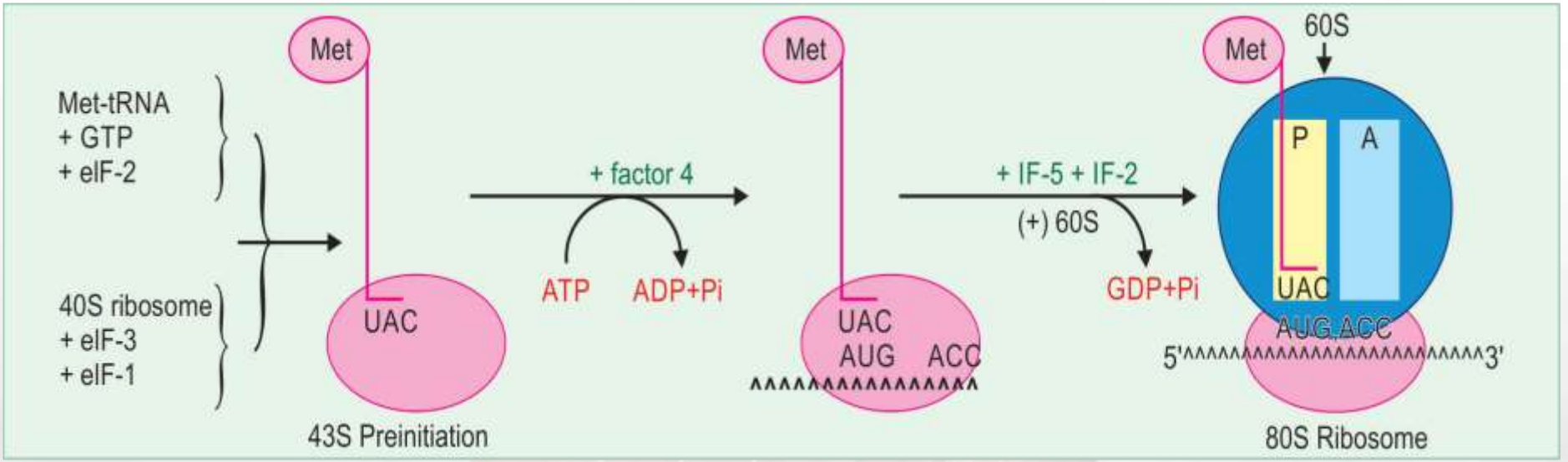
Aminoacyl tRNA synthetase

Amino acid + tRNA + ATP ----------→ Aminoacyl tRNA+ AMP

In this reaction, ATP is hydrolyzed to AMP level, and so two high energy phosphate bonds are consumed.



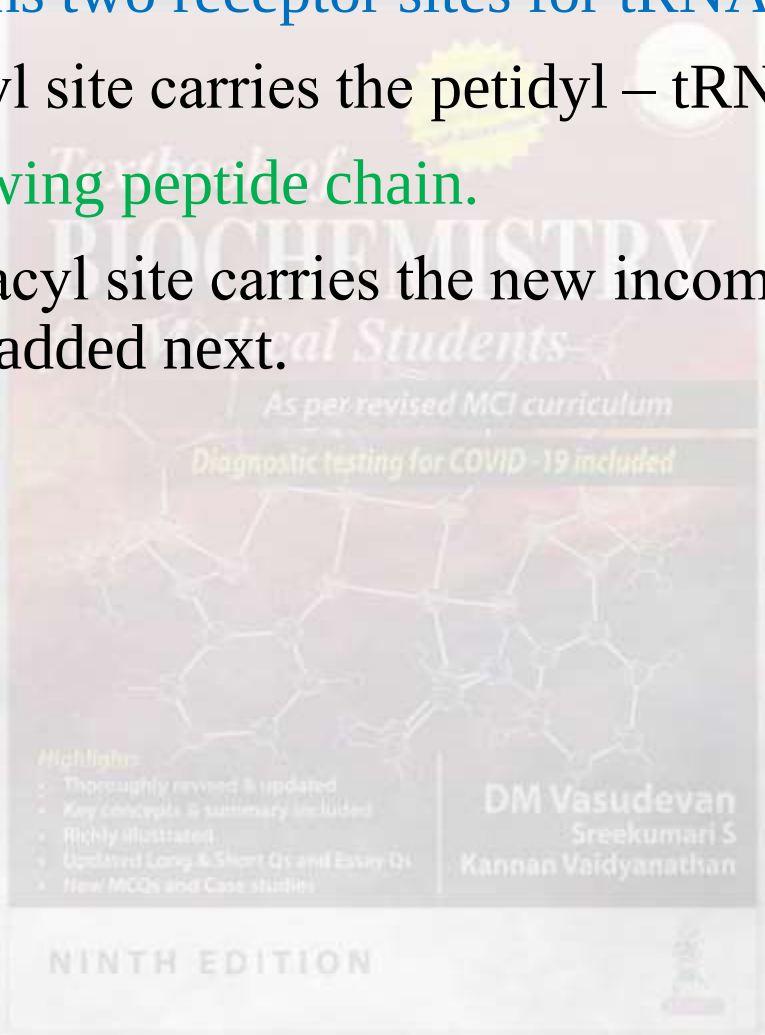
Pre-initiation complex of protein synthesis in prokaryotes.

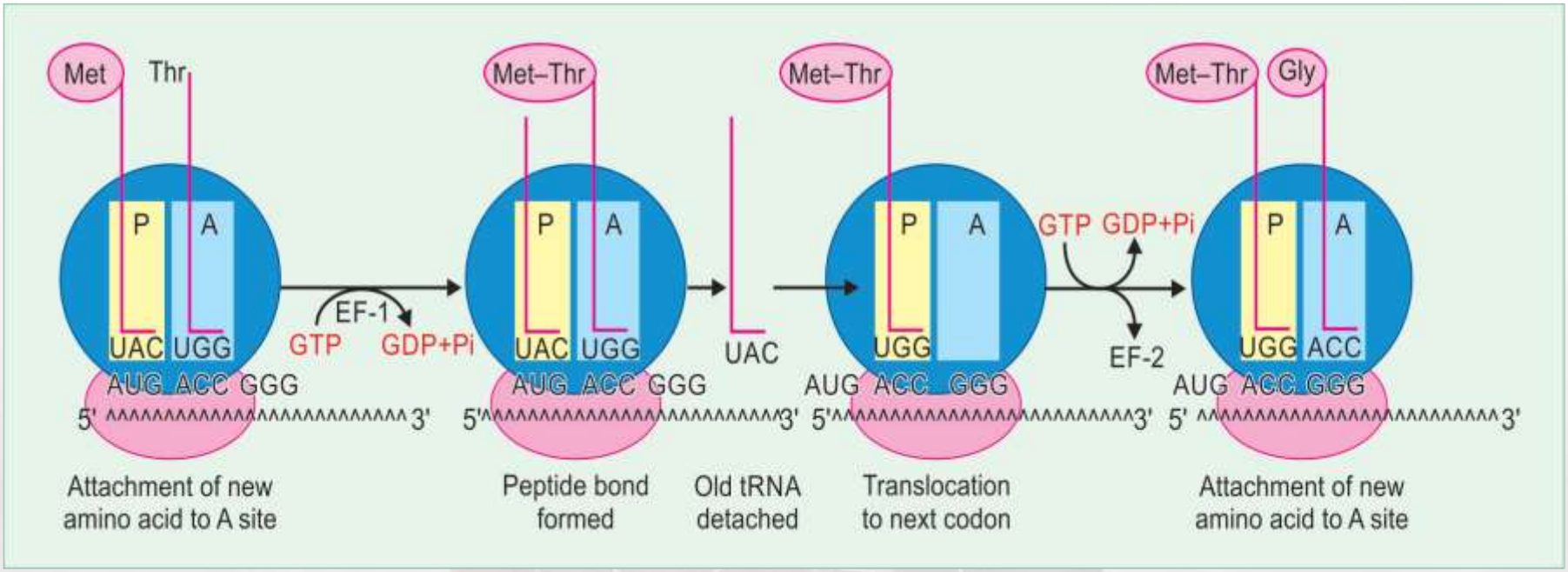


Initiation steps
(UAC: anticodon on Met-tRNA; AUG: start signal;
P: peptidyl site; A: aminoacyl site).

‘P’ and ‘A’ sites of Ribosomal Assembly

- ✓Ribosome contains two receptor sites for tRNA molecules.
- ✓‘P’ site or peptidyl site carries the petidyl – tRNA.
- ✓It carries the growing peptide chain.
- ✓‘A’ site or aminoacyl site carries the new incoming tRNA with the amino acid to be added next.





Elongation phase.
(P: peptidyl site; A: aminoacyl site).

Translocation Process

The tRNA fixed at the "P" site does not carry any amino acid and is therefore released from the ribosome.

Then the whole ribosome moves over the mRNA through the distance of one codon (3 bases). The peptidyl tRNA is translocated to the "P" site; this is done with the help of **elongation factor 2 (EF2)**.

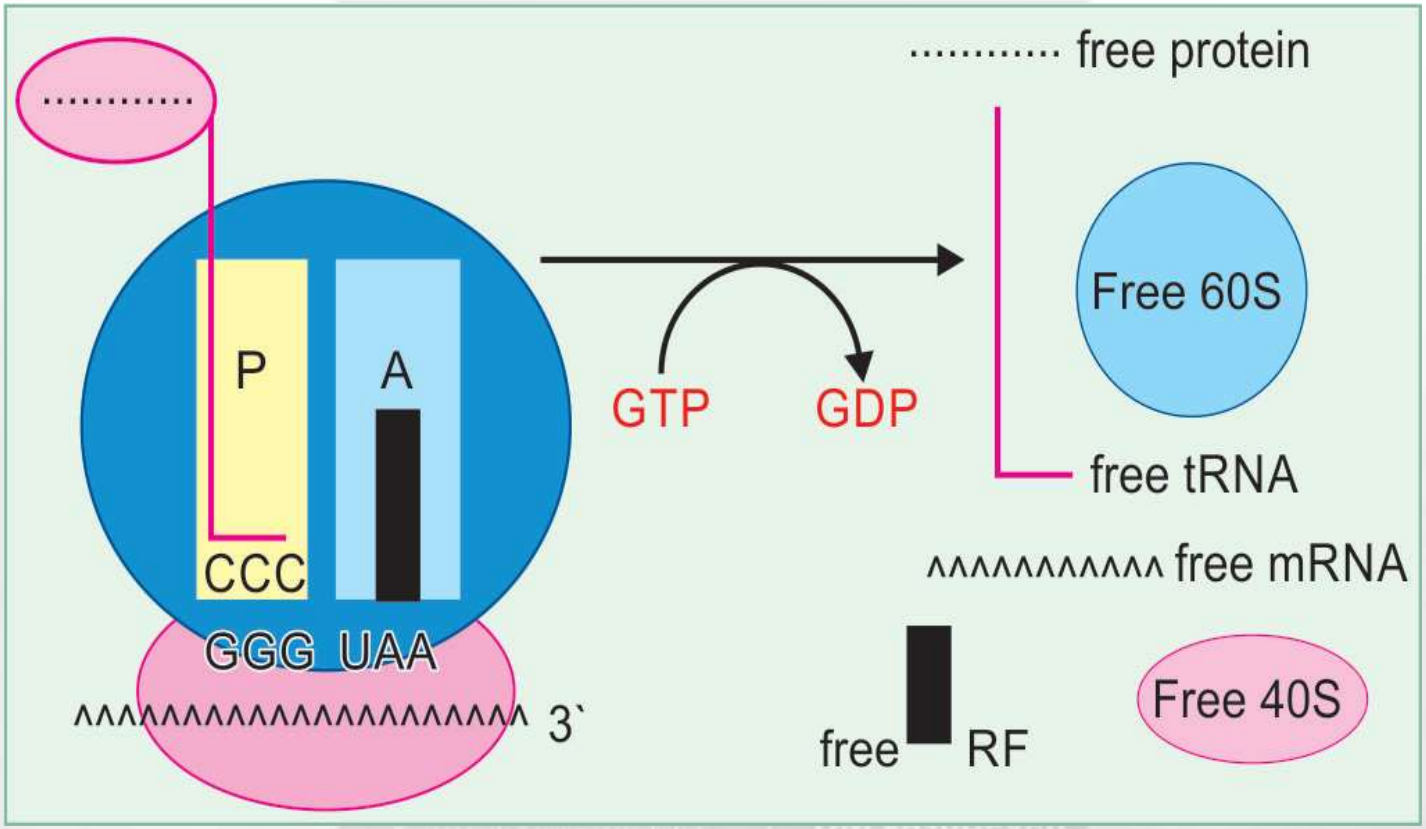
The "A" site is ready to receive another aminoacyl tRNA bearing the appropriate anticodon. The new aminoacyl tRNA is fixed to the "A" site, by base pairing with the mRNA codon. Translocation requires **hydrolysis of GTP** to GDP

Energy Requirements

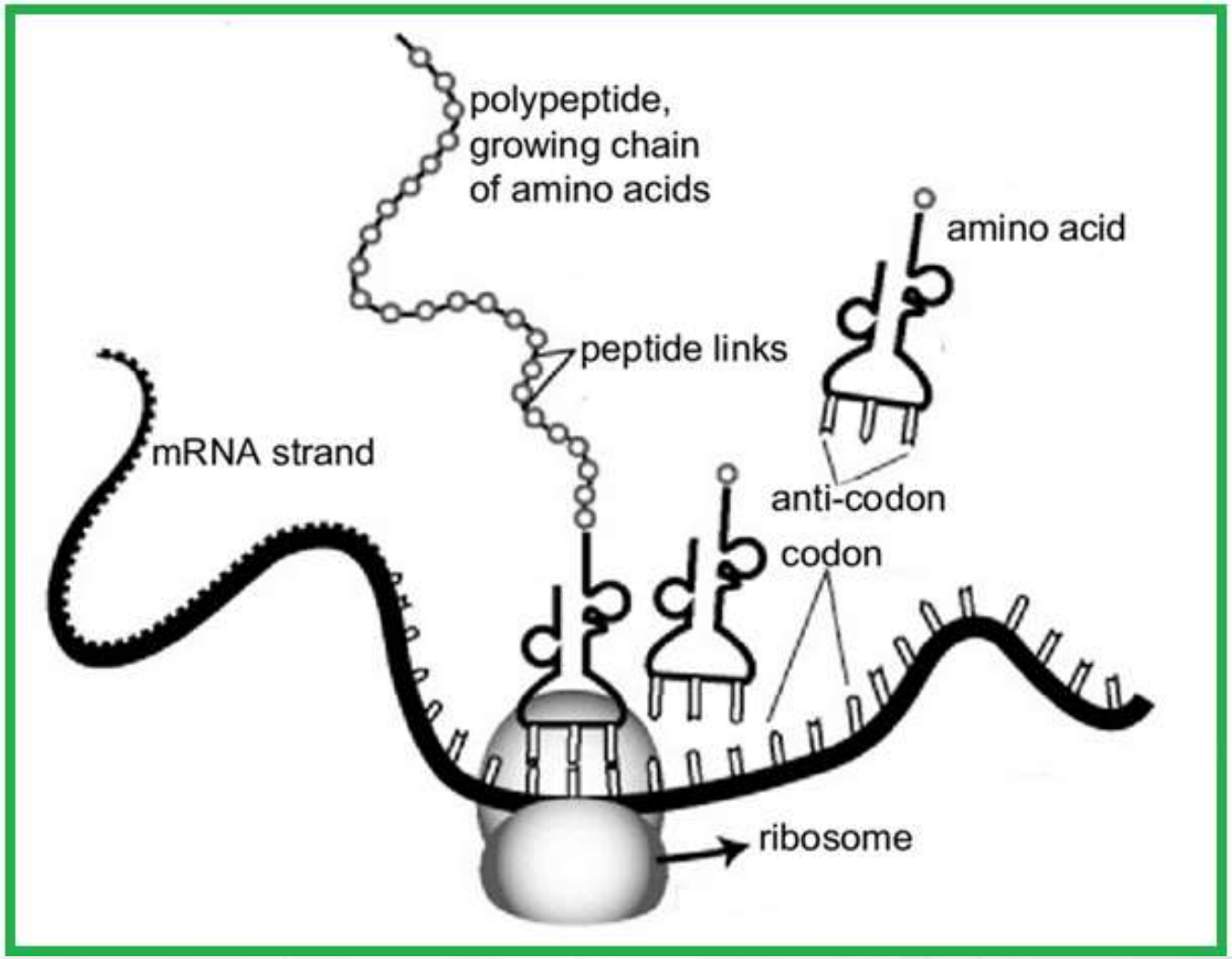
For each peptide bond formation, **four high energy** phosphate bonds are used.
Two for the initial activation and one for EF-1 step (GTP to GDP), and one for EF-2 step (GTP to GDP).

Actual peptide bond formation **(peptidyl transferase step) does not require** any energy, because the amino acids are already activated.

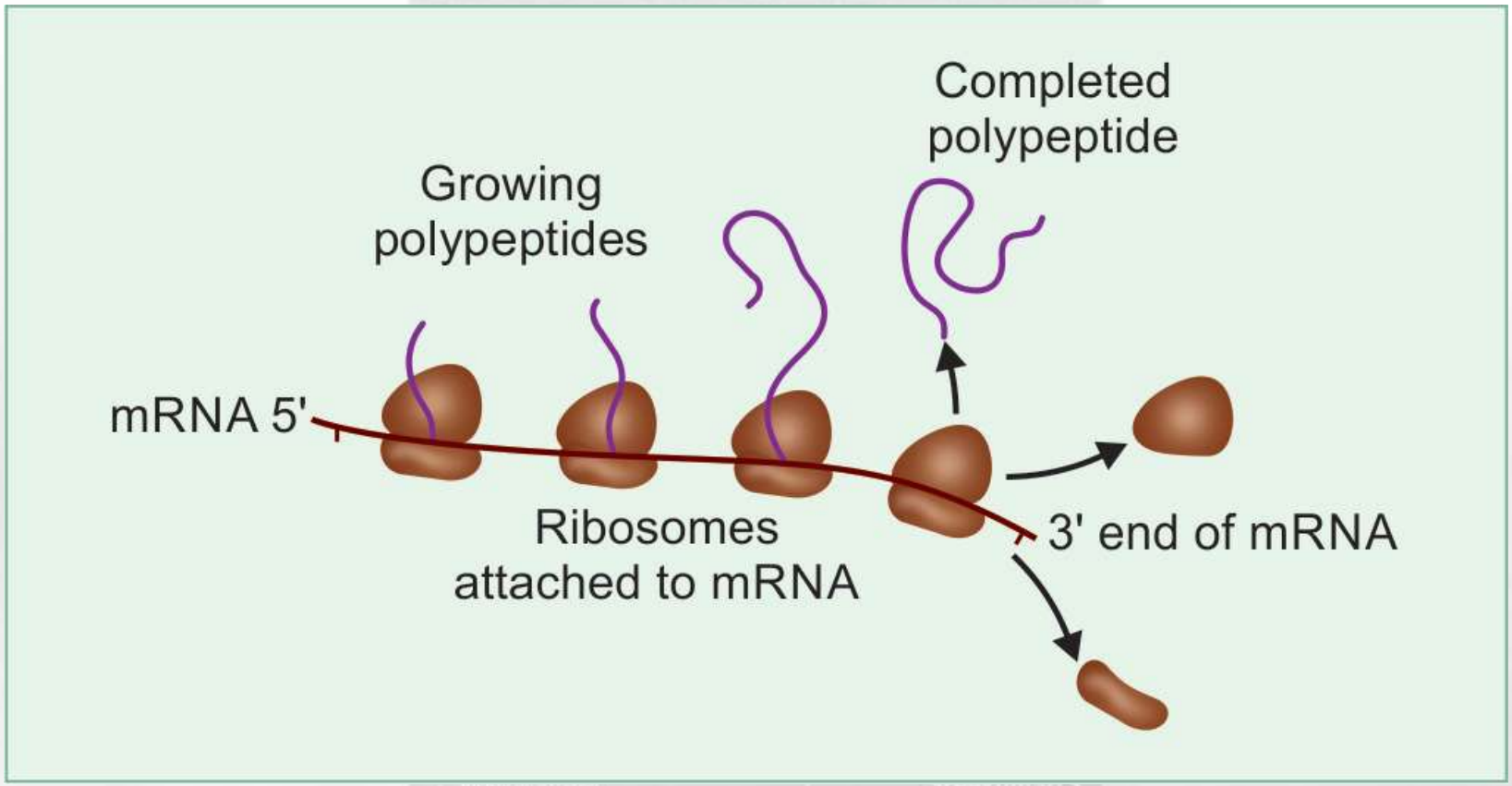
Further, 1 ATP is used for initiation complex formation; 1 GTP for 80S ribosome formation and 1 GTP for termination.



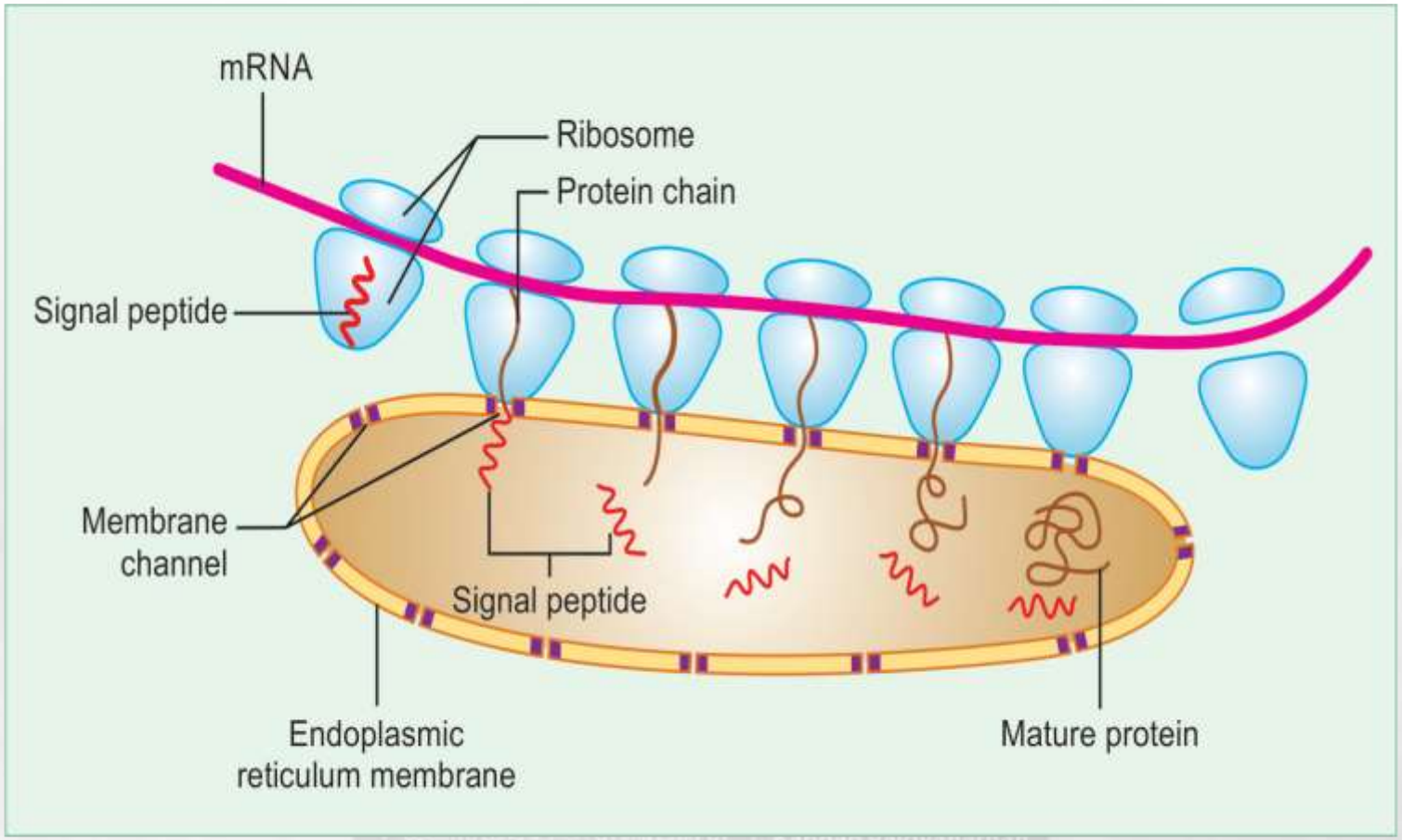
Termination of translation.



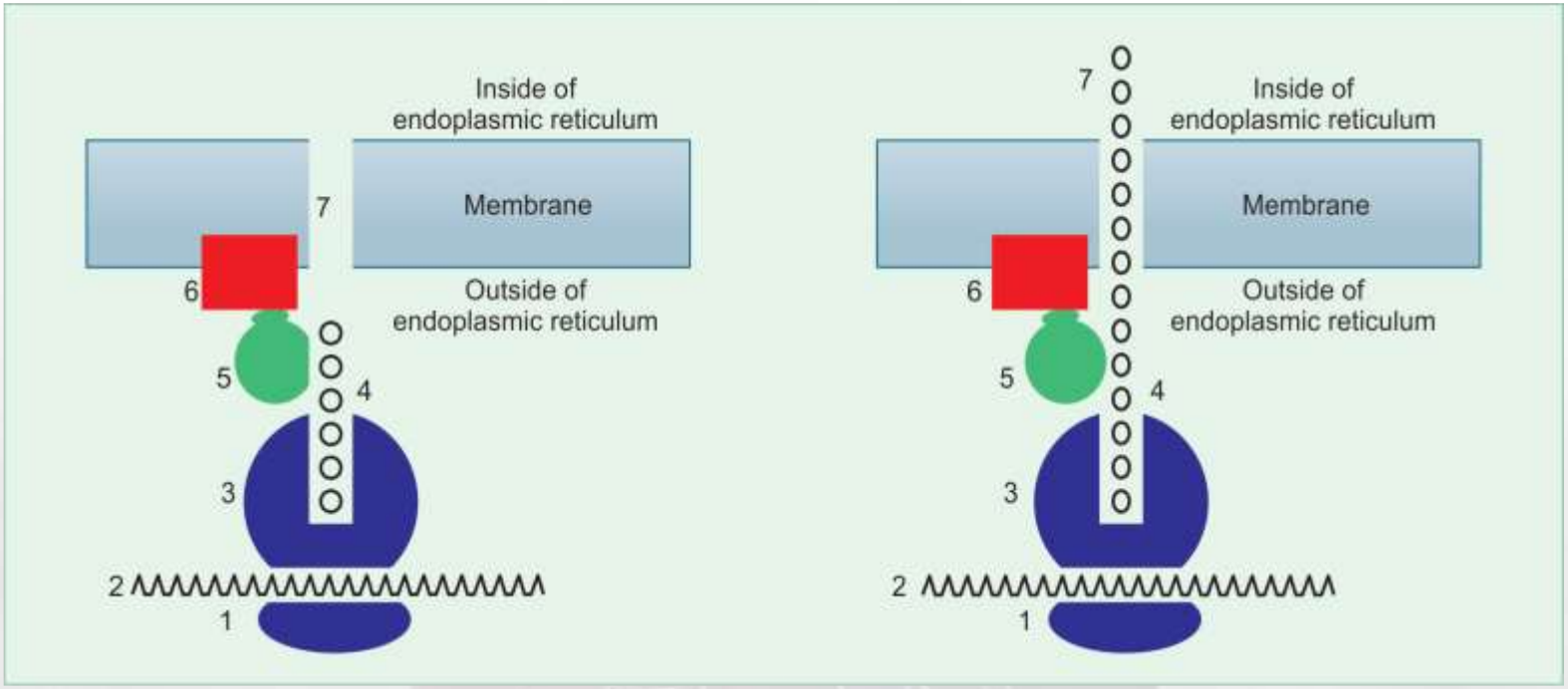
Summary of protein biosynthesis.



Polyribosomes.



Signal hypothesis. Signal peptide leads the protein into endoplasmic reticulum.



Synthesis of secretory proteins. (A) Protein synthesis initiated. (B) Protein is directed into the endoplasmic reticulum. 1 = 40S ribosome; 2 = mRNA attached to the ribosome; 3 = 80S ribosome; 4 = SP (signal peptide) new protein synthesis started; 5 = SRP or signal peptide recognition particle; 6 = SRPR or SRP-receptor; 7 (left side) = The SR protein is correctly aligned through the pore of endoplasmic reticulum. 7 (right side) = Peptide is passing into the lumen of endoplasmic reticulum.

Post-translational Processing

1. Proteolytic cleavage, eg, pro-insulin to insulin

2. Modification of amino acids

2-a Gamma carboxylation of glutamic acid of prothrombin (vitamin K dependent)

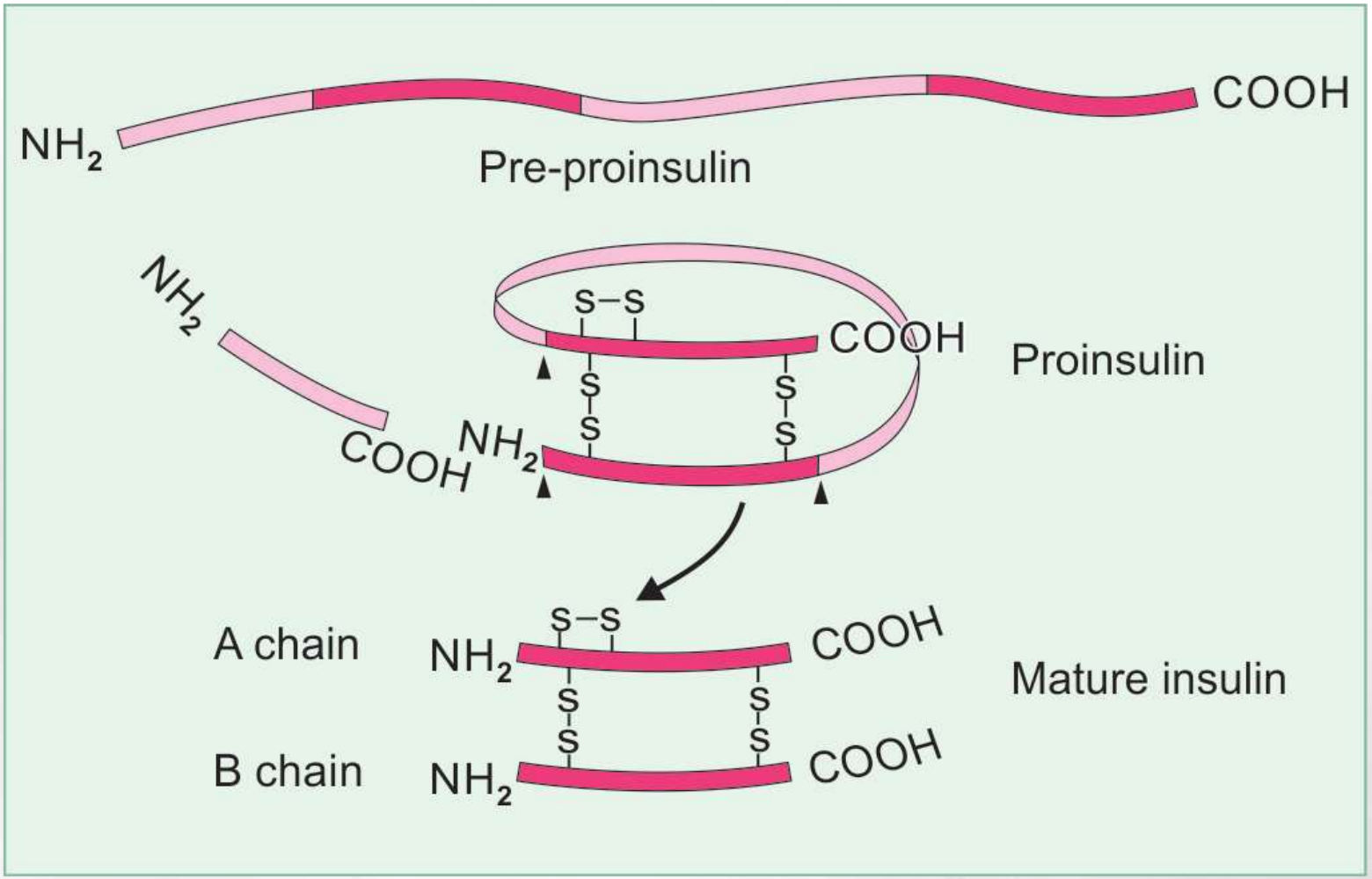
2-b Hydoxylation of proline and lysine in collagen (vitamin C dependent)

2-c Phosphorylation

2-d Glycosylation

2-e Methylation of histones

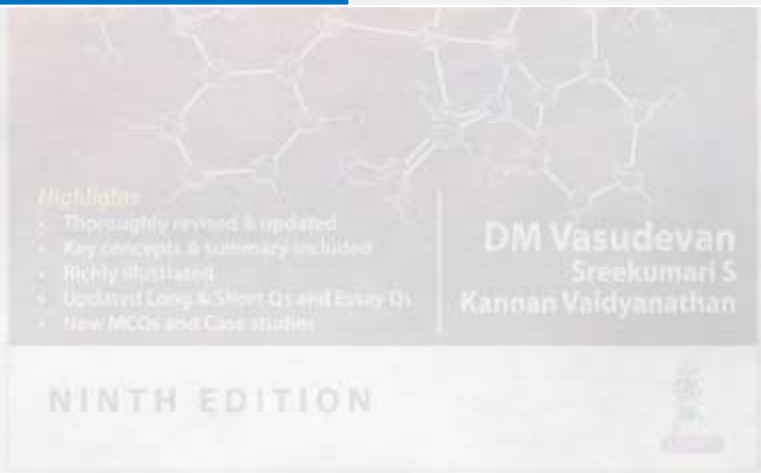
3. Subunit aggregation, eg. Hemoglobin, Immunoglobulin
4. Protein folding by Chaperones



Post-translational processing of insulin by proteolytic cleavage.

Post-translational Modifications

Reversible	Irreversible
Disulfide bridge **Glycosylation** **Phosphorylation Acylation** **N-acetylation** **ADP-ribosylation**	Proteolysis Ubiquitination Lysine hydroxylation Proline hydroxylation Methylation



Defects in Protein Targeting

Zellweger syndrome is due to defective oxidation of very long chain fatty acids (VLCFA). Here the correct "address" is not printed on the protein packet; so that it could not be delivered to the correct locality. Peroxisomal enzymes are produced; but their entry into peroxisome is denied. This leads to insufficient oxidation of VLCFA. Accumulation of VLCFA in CNS causes neurological impairment and death in childhood.

Another example is **primary hyperoxaluria**, which causes kidney stones at an early age. The defect is due to protein targeting defect and the enzyme alanine glyoxalate aminotransferase is seen in mitochondria, instead of its normal peroxisomal location.

Familial hypercholesterolemia is due to deficient transport signals.

Inclusion cell disease is due to non-entry of normal enzymes into lysosomes. Mannose-6-phosphate which is the marker to target enzymes to lysosomes; is absent in this disease.

Protein Folding and Chaperones

When a protein is being synthesized, it may assume different three-dimensional structures, out of which only one will have the biological activity. Abnormal folding of proteins may lead to **protein misfolding diseases**.

Chaperones help to produce the correct spatial arrangement. Chaperones attach to nascent polypeptide chains and prevent wrong foldings; so that the folding is allowed only in the correct direction. They help in the assembly of tertiary and quaternary structure of proteins.

Chaperonopathies are disorders resulting from "sick chaperones". These diseases progress with age.

Differences between Prokaryotic and Eukaryotic Translation

Feature	Prokaryotes	Eukaryotes
Ribosomes	30S small subunit and 50S large subunit associate to form 70S ribosome	40S small subunit and 60S large subunit together form the 80S ribosome
Initiation	Translation starts before transcription is completed. No mRNA processing	Transcription and translation are well separated and the processed mRNA is translated
Consensus sequences	Initiating codon AUG is recognised by the Shine Dalgarno sequence	Initiating codon AUG is recognised by Kozak sequence
Initiator tRNA	The initiating tRNA carries formylated methionine	The initiating tRNA carries methionine

Differences between Prokaryotic and Eukaryotic Translation

Feature	Prokaryotes	Eukaryotes
Formation of initiation complex	Initiation factors and GTP hydrolysis are required for the binding of the small ribosomal subunit to the mRNA	Several IF s and GTP hydrolysis are required for the formation of the initiation complex.
Peptide bond formation	The peptidyl transferase ribozyme is in the 23S r RNA	Peptidyl transferase activity in the 28S rRNA of the large subunit.
Post-translational events	Mainly co-translational modifications	Mainly post-translational modification, sorting, export and localisation
Inhibition	Antibiotics can inhibit different stages of translation	Toxins like diphtheria, ricin, puromycin and cycloheximide

Inhibitors of Protein Synthesis

They generally act only on bacteria and are nontoxic to human beings. This is because mammalian cells have 80S ribosomes, while bacteria have 70S ribosomes.

Reversible Inhibitors in Bacteria

These antibiotics are **bacteriostatic.**

Tetracyclins inhibit attachment of aminoacyl tRNA to the A site of ribosomes.

Chloramphenicol inhibits the peptidyl transferase activity of bacteria.

Erythromycin and **clindamycin** prevent the translocation process.

Irreversible Inhibitors in Bacteria

These antibiotics are **bactericidal.**

Streptomycin binds to 30S subunit of bacterial ribosomes. They cause misreading of mRNA and totally inhibit protein synthesis.

Inhibitors of Protein Synthesis

Inhibitors of Protein Synthesis in Eukaryotes

Puromycin is structurally similar to tyrosine-tRNA and gets attached to the "A" site of the ribosome. So, the incomplete peptide is released.

Cycloheximide inhibits peptidyl transferase in 60S subunit.

Diphtheria toxin, liberated by the bacteria, Corynebacterium diphtheria, causes inactivation of EF-2 by attachment of ADP to EF-2 and consequent inhibition of protein biosynthesis in mammalian systems.

Inhibitors of transcription (described previously in this Chapter) will also, in turn, inhibit translation process.

Mitochondrial DNA

Some of the mitochondrial protein synthesis is under the control of mitochondrial DNA. The mtDNA has information for synthesis of components of electron transport chain. However, most of mitochondrial proteins are encoded by nuclear DNA and synthesised in the cytoplasm. Important proteins of the outer membrane of the mitochondria are synthesised under the influence of nuclear DNA.

Mitochondria are similar to bacteria than mammalian cells. This fact supports the theory that mitochondria are derived from prokaryotes symbiotically adapted to multicellular organisms.

Comparison of translation in eukaryotes and prokaryotes. Mitochondria are similar to prokaryotes.

Feature	Eukaryotes (mammalian cells)	Prokaryotes (bacteria)	Mitochondria
DNA	Open	Circular	Circular
Ribosomes	80S	70S	70S
tRNA (No.)	31	22	22
Initiating amino acid	Methionine	Formyl methionine	Formyl methionine
Effect of tetracyclin	Not affected	Inhibited	Inhibited

Mitochondrial DNA

Maternal inheritance: Since the mitochondria are inherited cytoplasmically, the mtDNA is inherited from the mother.

There are hundreds of copies of mtDNA in each cell (nuclear DNA has only 2 copies).

If mutation occurs in mtDNA, the daughter cells may inherit the mutant or normal mtDNA (Heteroplasty)

High mutation rate.

Accumulation of mutations in mtDNA may be responsible for age related degenerative diseases.

OXPHOS (Oxidative Phosphorylation) Diseases

Syndrome	Features
Leber's hereditary neuropathy (LHON)	Complex I defect; blindness, cardiac conduction defects
Myoclonic epilepsy ragged red fiber disease (MERRF)	Myoclonic epilepsy, myopathy, dementia
Mitochondrial encephalopathy lactic acidosis stroke like episodes (MELAS)	Complex I defect; lactic acidosis, strokes,myopathy, seizures, dementia
Leigh's syndrome	Complex I defect; NDUFS gene defect; movement disorders

Micro -RNA

Micro-RNAs or **miRNAs** are about 21–25 bases in length. They have RNA hairpin structure (showing internal hybridization to make it two strands), and are called **short hairpin RNA (shRNA)**. In cytoplasm, out of the two strands, one is broken by **dicer nuclease**. The selected strand is called the **guide strand**, which is incorporated into the **RNA-induced silencing complex (RISC)** to form functional silencer of mRNA.

The micro-RNAs bind to matching pieces of messenger RNA, turn it into a double strand and keep it from doing its job. The process effectively blocks the production of the corresponding protein, causing translation arrest.

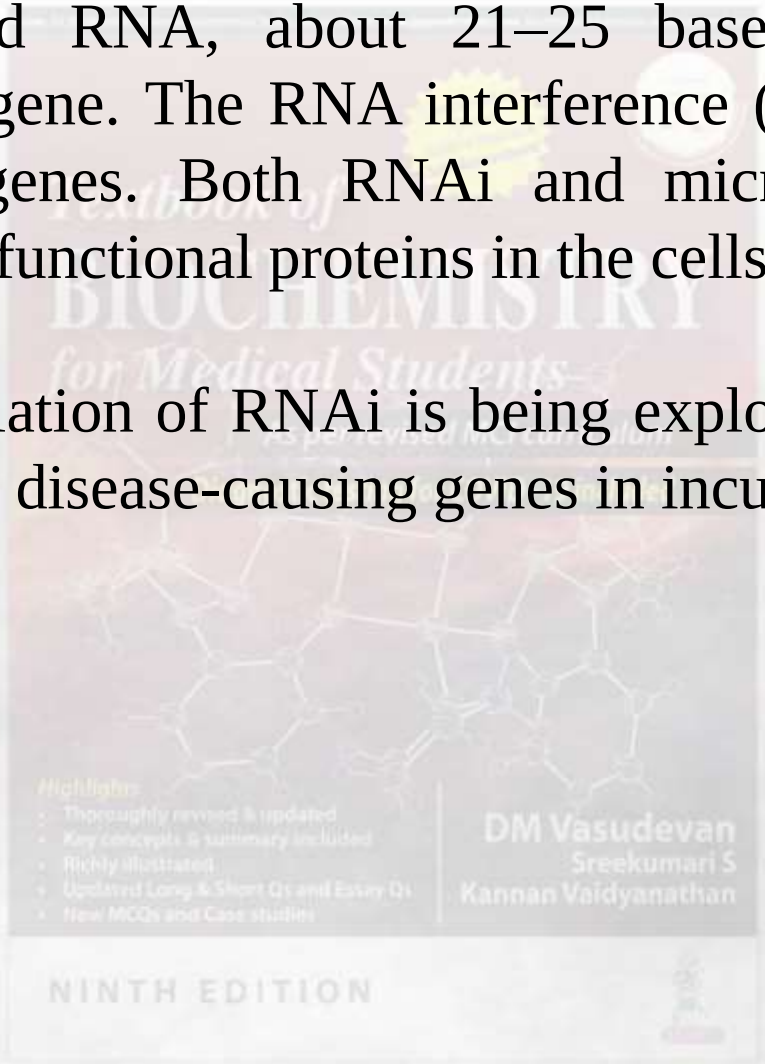
More than 2,500 human miRNAs have been identified.

Circulating miRNA is a rich source for potential disease biomarkers.

Interfering RNA or RNAi or siRNA

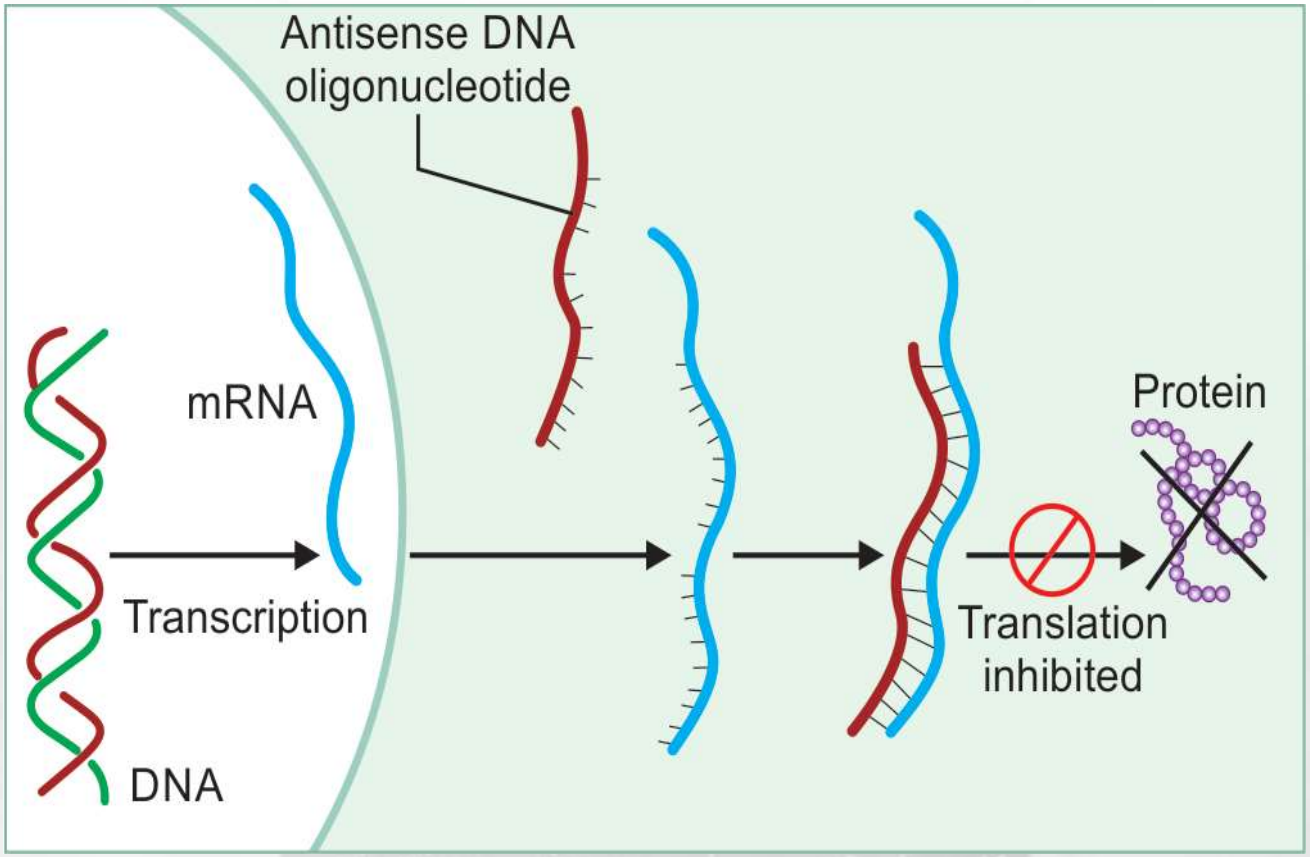
Short double-strand RNA, about 21–25 bases, would silence the corresponding gene. The RNA interference (RNAi) is a faster way to turn off genes. Both RNAi and micro-RNA result in decreased levels of functional proteins in the cells.

Exogenous manipulation of RNAi is being explored as a powerful method of silencing disease-causing genes in incurable neurological disorders.



DNA directed RNA interference (ddRNAi)

The RNAi pathway is initiated by the enzyme **Dicer**, which cleaves long double stranded RNA (dsRNA) into short double stranded fragments of about 21 nucleotide long siRNAs. Each siRNA is unwound into two single-stranded RNAs (ssRNAs), the passenger strand and the guide strand. The passenger strand is degraded and the **guide strand** is incorporated into the RNA induced silencing complex (RISC). The guide strand pairs with a complementary sequence in mRNA molecules and induces cleavage by the catalytic component of the RISC. Thus the mRNA is degraded. The small interfering RNA (siRNA) silences the genes only transiently. But the ddRNAi are continually transcribed, replenishing the cellular of shRNA. Thus the targeted genes are silenced for a long time by introducing small DNA constructs into the cell.



Principle of antisense therapy.