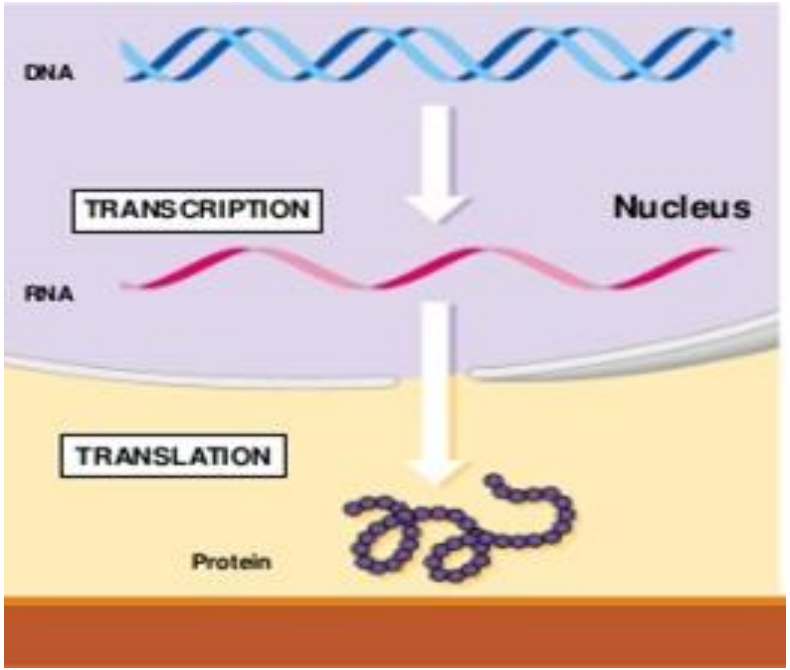
TRANSCRIPTION

Mrs. Samriti Gumber
Assistant Professor
Government College for Girls

- **DEFINITION**---Synthesis of RNA from ssDNA as a template by DNA dependent RNA polymerase enzyme.
- Similar to replication in terms of chemical mechanism, polarity and use of template but differs in :

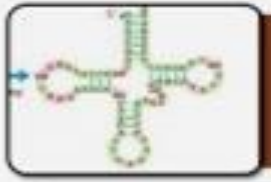
  ---------DOES NOT REQUIRE PRIMER

  ----------ONLY A SHORT SEGMENT OF DNA IS TRANSCRIBED.

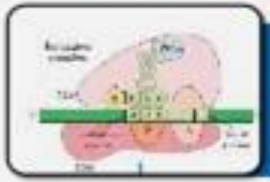


All 3 types of cellular RNA's are copied during transcription

Messenger RNAs (mRNAs) encode the amino acid sequence of one or more polypeptides specified by a gene.



Transfer RNAs (tRNAs) read the information encoded in the mRNA and transfer the appropriate amino acid to a growing polypeptide chain during protein synthesis.



Ribosomal RNAs (rRNAs) are constituents of ribosomes, the intricate cellular machines that synthesize proteins.

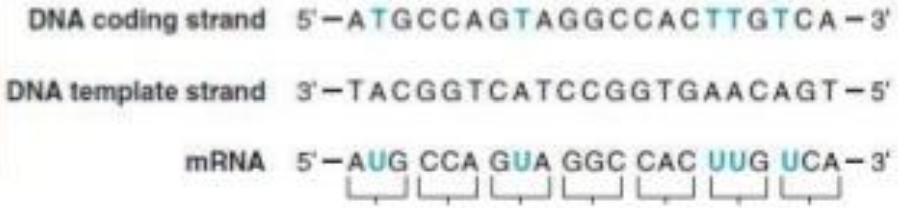
SALIENT FEATURES OF TRANSCRIPTION

- Synthesis of all types of RNA in NUCLEUS
- Only ONE STRAND of DNA participates.
- RIBONUCLEOTIDES are used in RNA synthesis
- RNA synthesis occur in 5′-3′ direction, DNA template is read from 3′-5′ direction.
- Synthesis follows WATSON CRICK MODEL ie. A with U and G with C
- DNA dependent RNA polymerase is used.

Basic Requirements of Transcription

Template

- ss DNA



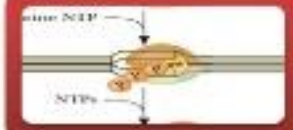
Enzyme

- DNA dependent RNA polymerase

Ribonucleoside triphosphates

- ATP, GTP,CTP,UTP

Steps involved in Transcription



INITIATION

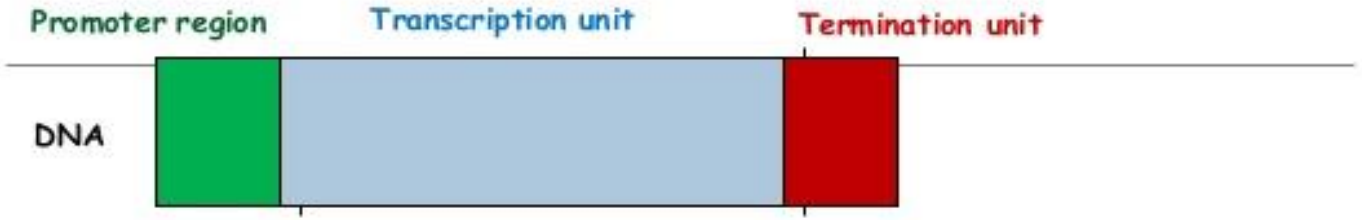
ELONGATION

TERMINATION & Release

TRANSCRIPTION------

(PROKARYOTES)

Template DNA



1.**Promoter region:** It is the specific region in DNA ,where transcription is initiated.

2.**Transcription unit**: It is the region where DNA template is transcribed. Present in b/w promoter and terminating units.

3.**Termination unit**: It is the region where transcription terminates.

Initiation

- Starts with the recognition of promoter sequence on the DNA coding (anti-template) strand by RNA polymerase

- **Promoter sequence**

  Are specific areas on the DNA that act as starting signals for initiation process recognized by RNAP.

- Two common sequences are present on the upstream side of the start site of transcription .

- Start site is denoted by +1

Promoters of prokaryotes

✓Pribnow box or TATAAT box:

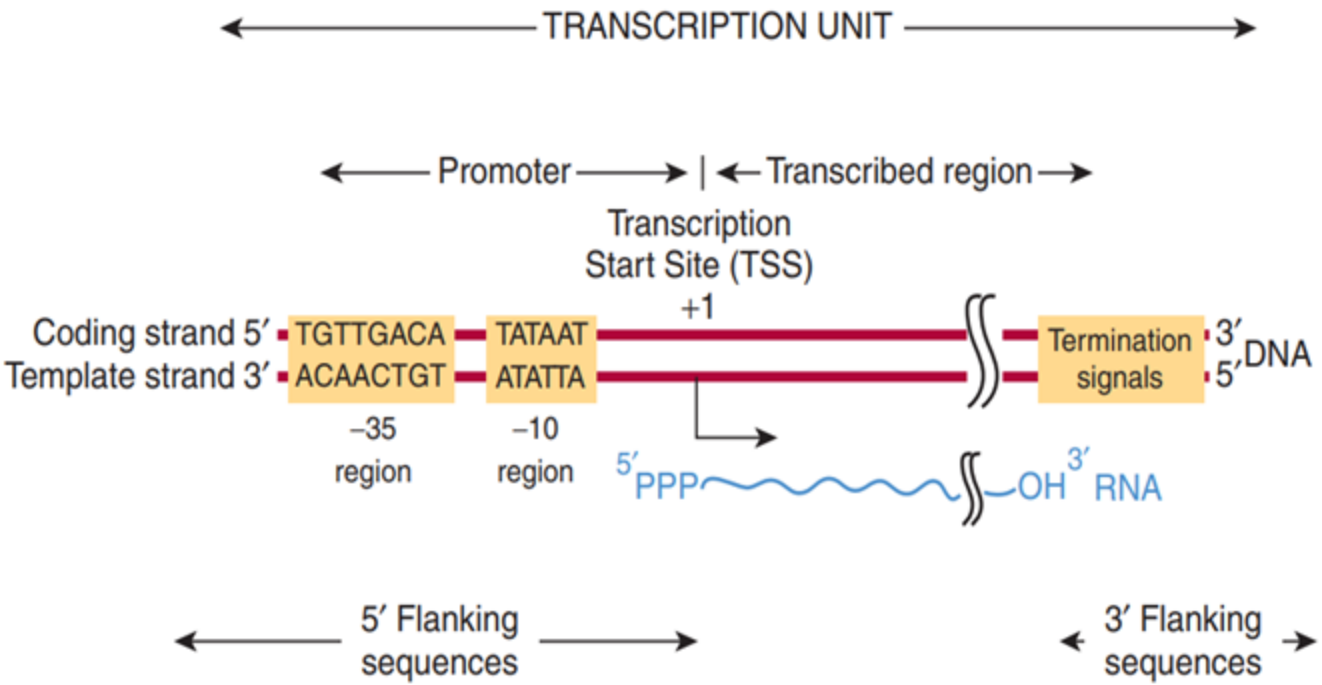
It contains 6 nucleotide bases TATAAT located -10 bases away on the left of origin of transcription.

✓ -35 sequence / 5`-TTGACA-3` seqence:

it is 35 bp upstream of the transcription unit

5`	TTGACA	3`

PROKARYOTIC PROMOTERS



RNA POLYMERASE IN PROKARYOTES

- It has 5 subunits ie 2 alpha units, beta, beta' and sigma subunit.

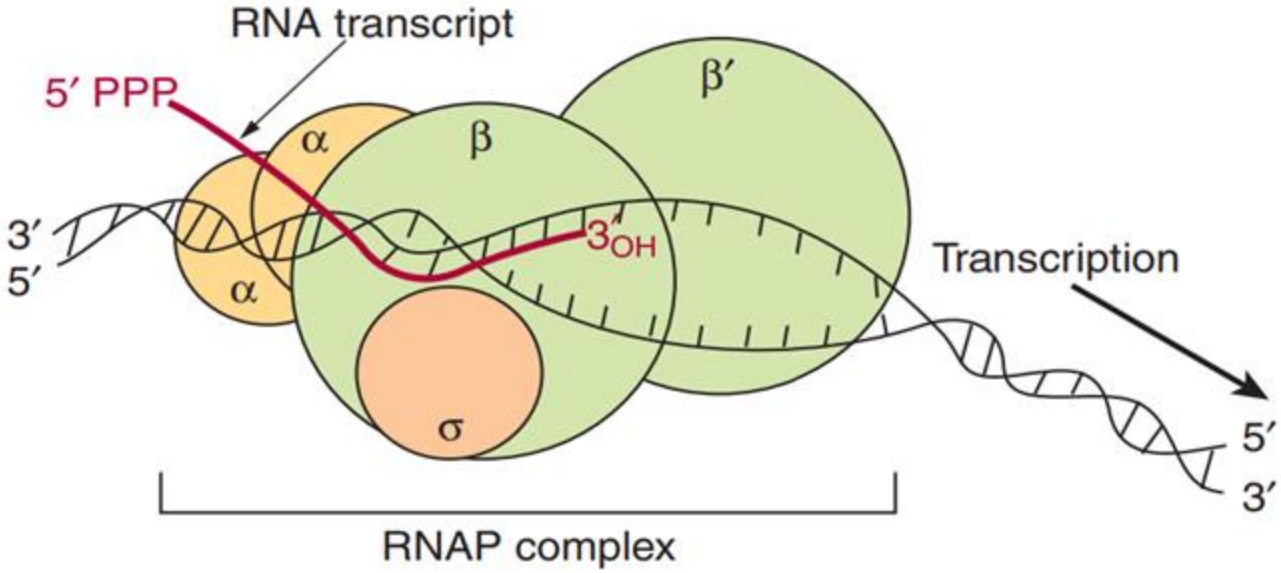
CORE ENZYME

2 alpha subunit (alpha1 and alpha2)
Beta subunit
Beta' subunit
Omega subunit

HOLO ENZYME

2 alpha subunit (alpha1 and alpha2)
Beta subunit
Beta' subunit
Omega subunit
SIGMA SUBUNIT

- Beta' subunit- **LARGEST SUBUNIT**
- Beta subunit- **SECOND LARGEST SUBUNIT**
- Alpha subunit- **THIRD LARGEST SUBUNIT**


RNA transcript
5′ PPP
α
β
β′
3′
5′
α
3′OH
Transcription
σ
5′
3′
RNAP complex

Functions of RNAP

Subunit	Role
α	Binds regulatory proteins
β	Forms phosphodiester bond
β′	Fixes RNAP to DNA template
σ	Recognizes and binds to promoter region of DNA, Initiates transcription.

In bacteria, one species of RNAP can synthesize all the RNA molecules (mRNA, tRNA, rRNA)

RNA polymerase differs from DNA Polymerase in two aspects

1. No primer is required for RNAP.
2. RNAP lacks a separate proofreading 3' to 5' exonuclease activity.

The error rate for transcription is higher.

Error rate is 10^4-10^5 times more than replication

Errors cause little damage – as they are not transmitted to daughte cell/ next generation.

1.INITIATION OF TRANSCRIPTION (PROKARYOTES)

Identification of promoter region:

Sigma unit of RNA Polymerase identifies the promoter region on template DNA .

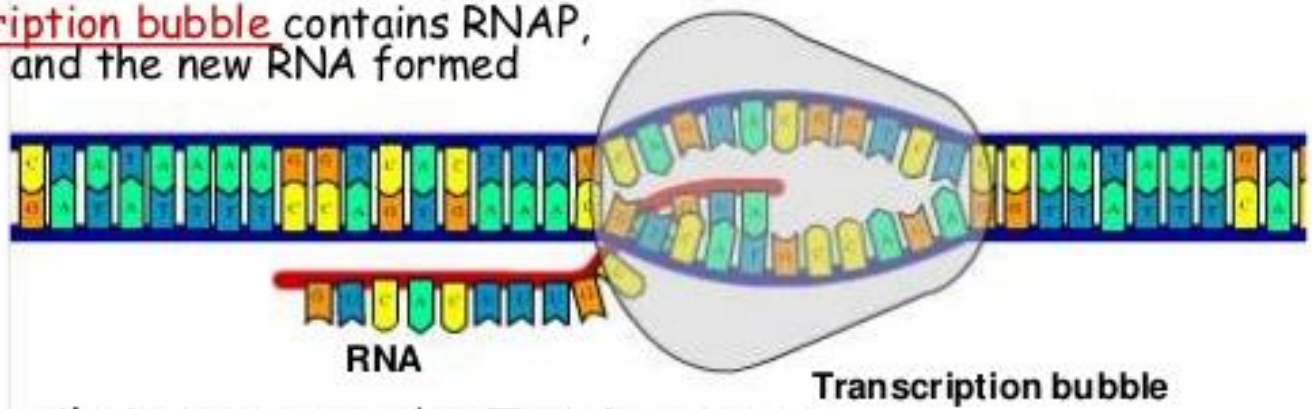
RNA polymerase binds to the promoter region of the template DNA

Beta unit fixes to the promoter region of the template strand and initiates transcription.

Sigma factor is released and RNAP in promoter region unwinds DNA helix.

A **Transcription bubble** is formed as the DNA unwinds down stream

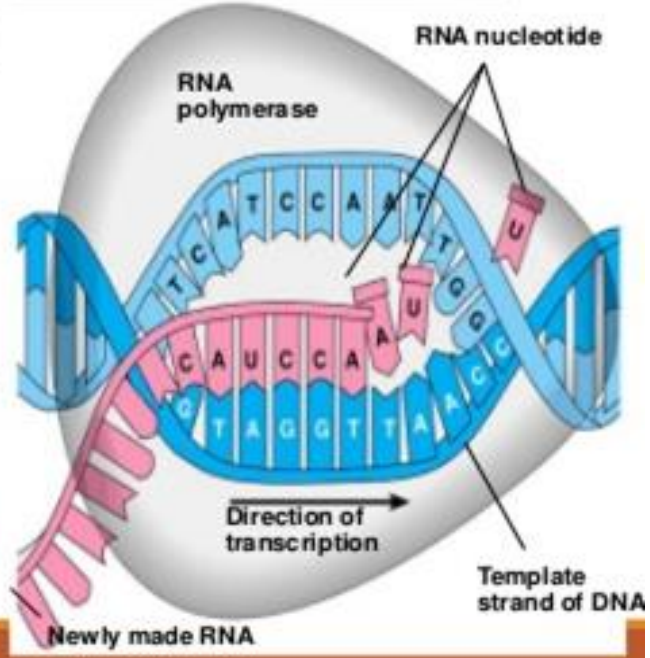
Transcription bubble contains RNAP, DNA and the new RNA formed



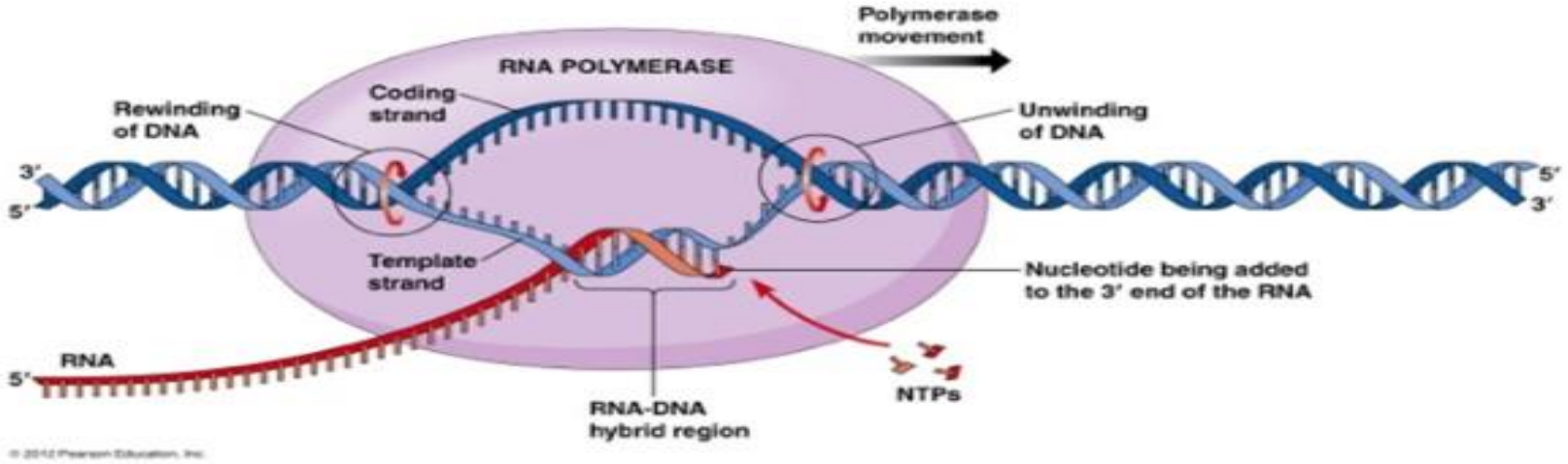
Supercoils is overcome by Topoisomerase

2. ELONGATION

- RNA polymerase elongates an RNA strand by adding ribonucleotide units to the 3'-hydroxyl end.
- RNAP uses ribonucleoside triphosphates, forms 3'5' PDE bond b/w adjacent ribonucleotide and releases ppi each time a nucleotide is added to the growing chain
- Each nucleotide in the newly formed RNA is selected by Watson-Crick base-pairing interaction. (G to C, A to U)



Transcription elongation

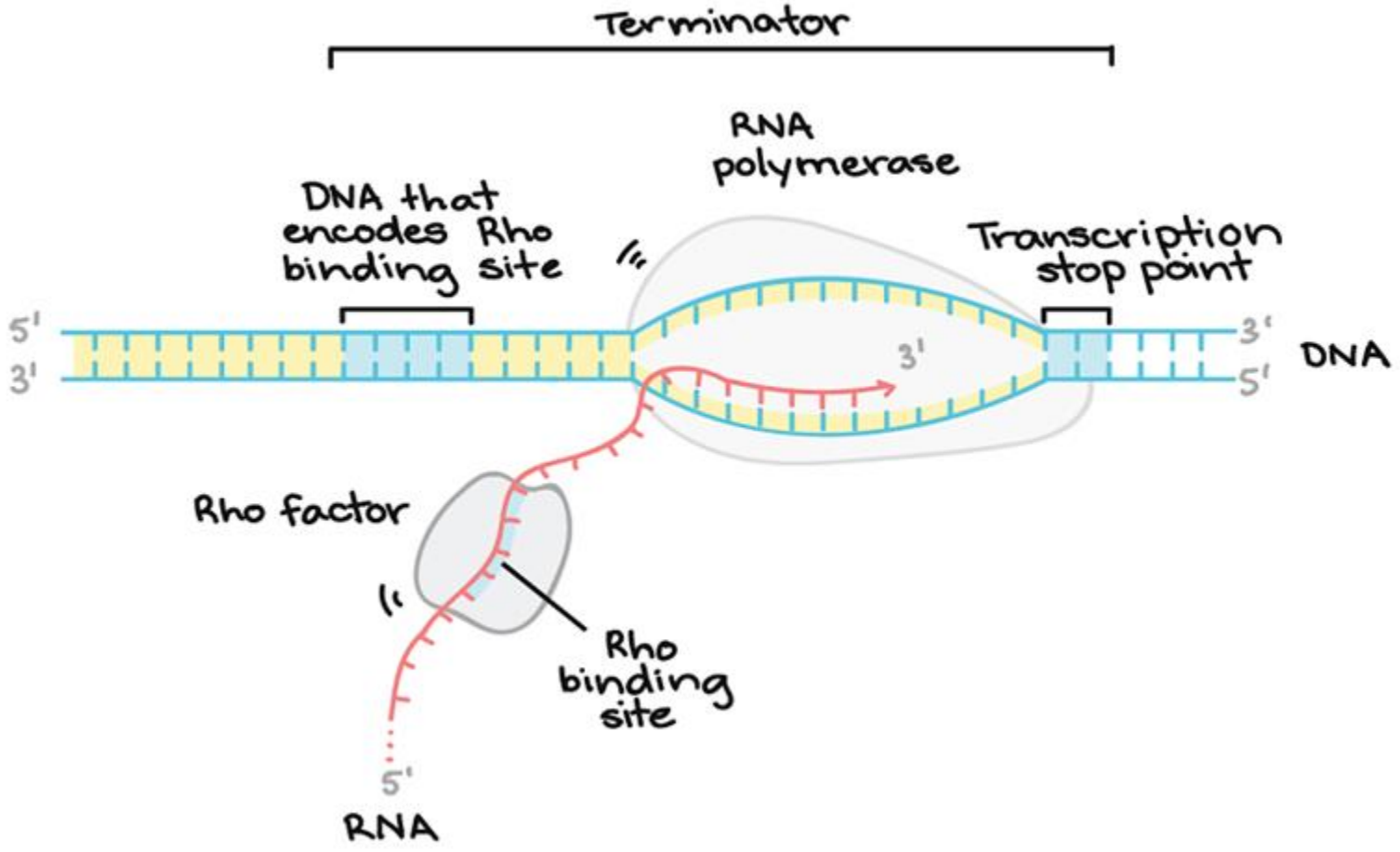


3. TERMINATION

- This is the process of ending transcription, which happens when signaled by a stop sequence known as a terminator sequence.
- This happens when the RNA polymerase transcribes the terminator sequence.
- The RNA polymerase then releases the DNA temple which unwinds back to a double-helical structure.
    - This can be of two types: (1) RHO- DEPENDENT
      (2) RHO- INDEPENDENT

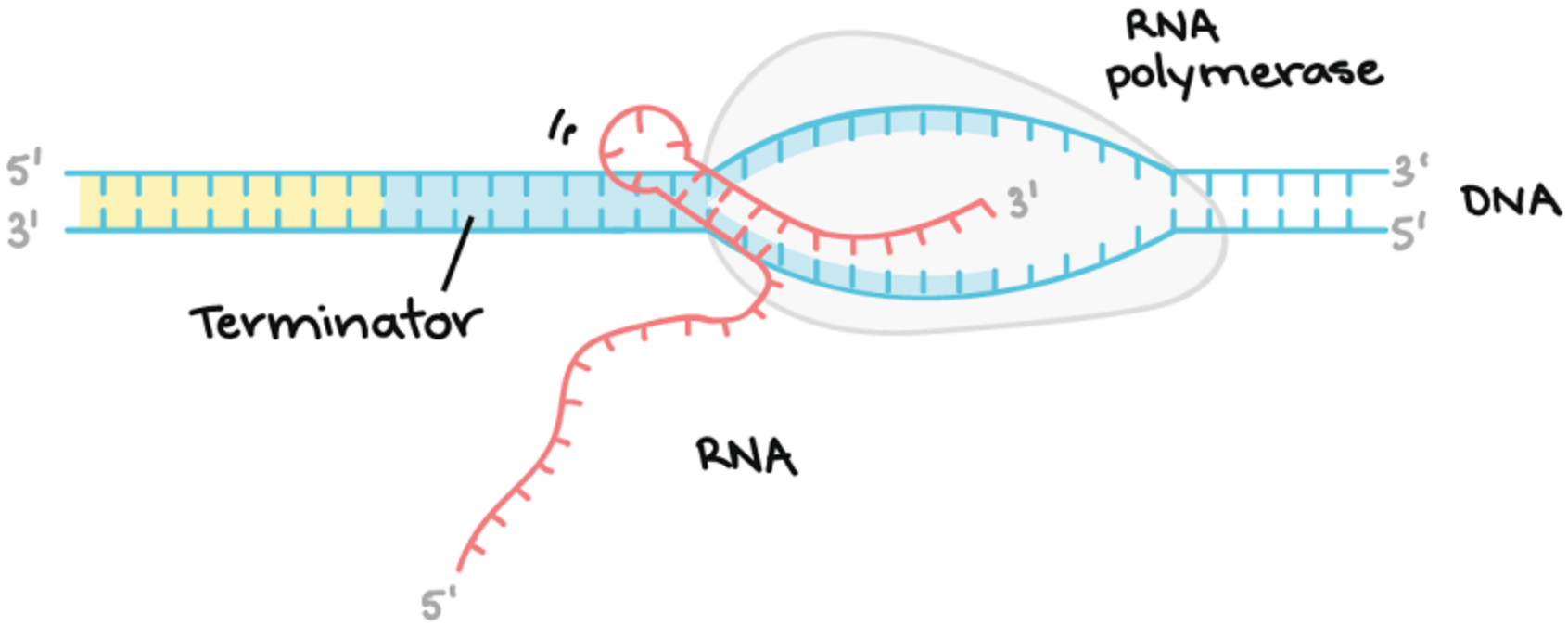
RHO-DEPENDENT TERMINATION

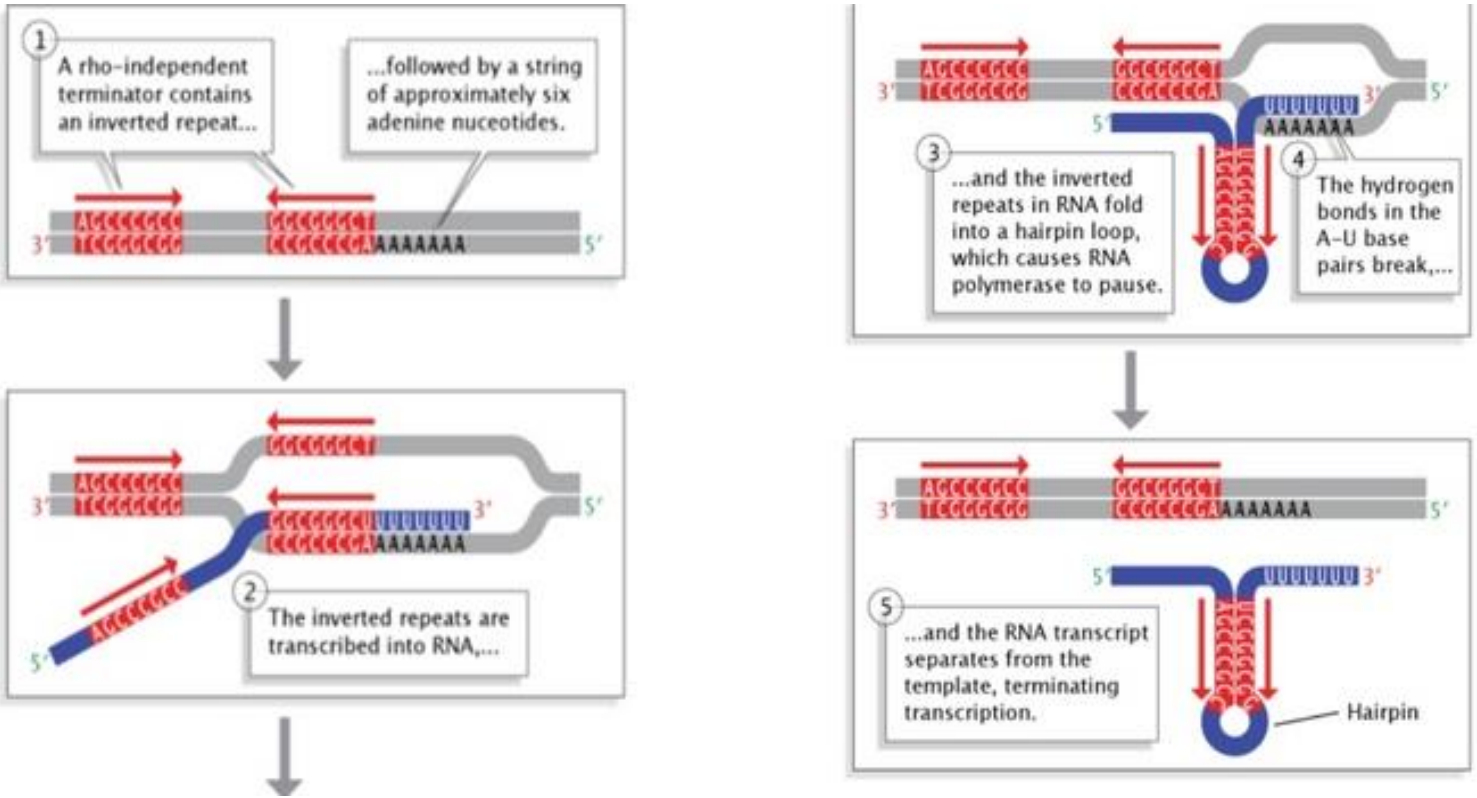
- ★ This is the termination process where the RNA molecule contains a binding site for a protein known as the Rho factor, which binds to the DNA sequence.
- ★ It starts to climb up the transcript towards the RNA polymerase and reaches the transcription bubble.
- ★ At the bubble, the Rho factor pulls the RNA transcript and the DNA template strand apart, releasing the RNA molecule and terminating the transcription process.
- ★ A transcription stop point sequence that is found later in the DNA causes the RNA polymerase to stop and allow the Rho factor to catch up and terminate the process.

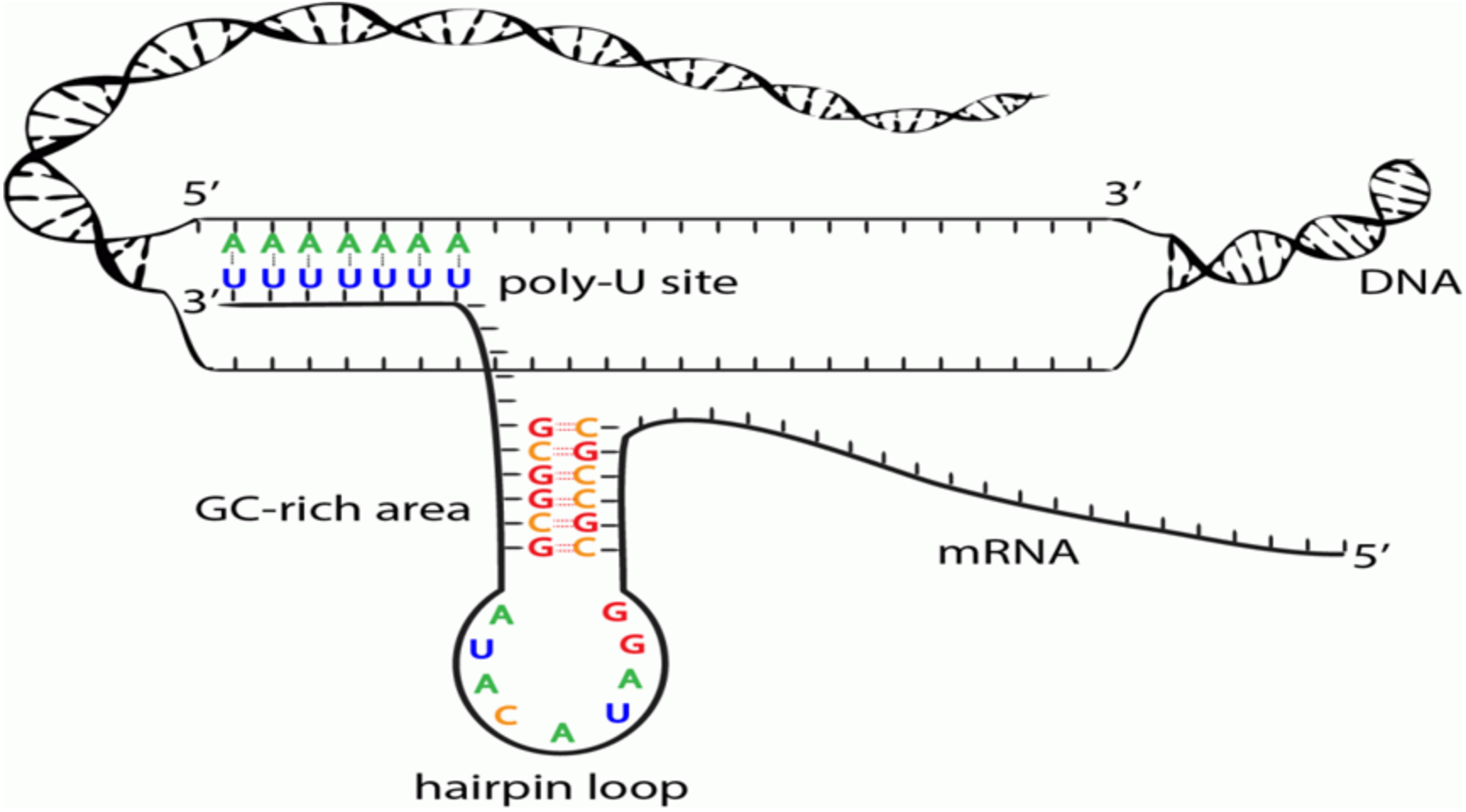

Terminator
RNA polymerase
DNA that encodes Rho binding site
Transcription stop point
5'
3'
3'
5'
DNA
3'
Rho factor
Rho binding site
5'
RNA

RHO INDEPENDENT TERMINATION

- This process depends on a specific sequence found on the DNA template strand.
- During transcription, as the RNA polymerase approaches the endpoint of the gen that is transcribed, it reaches a region that is rich in Cytosine (C) and Guanine(G).
- The RNA that is transcribed from this region folds back on itself, and the complementary C and G bind together forming a stable hairpin that makes the RNA polymerase to stall.
- The hairpin is followed by a Uracil (U) in the RNA terminator which complementary to the DNA template Adenine (A).
- The U-A region forms a weak interaction with the DNA template and with the stalled RNA polymerase causes an instability allowing the enzyme to fall off and end from the new RNA transcript.


RNA
polymerase
5'
3'
3'
3'
5'
DNA
Terminator
RNA
5'


1
A rho-independent terminator contains an inverted repeat...
...followed by a string of approximately six adenine nuceotides.
3′
AAAAAAA
5′
2
The inverted repeats are transcribed into RNA,...
3
...and the inverted repeats in RNA fold into a hairpin loop, which causes RNA polymerase to pause.
4
The hydrogen bonds in the A–U base pairs break,...
5
...and the RNA transcript separates from the template, terminating transcription.
Hairpin


5′
3′
A A A A A A A
U U U U U U U
poly-U site
3′
DNA
G C
C G
G C
G C
C G
G C
GC-rich area
mRNA
5′
A
U
A
C
A
G
G
A
U
hairpin loop

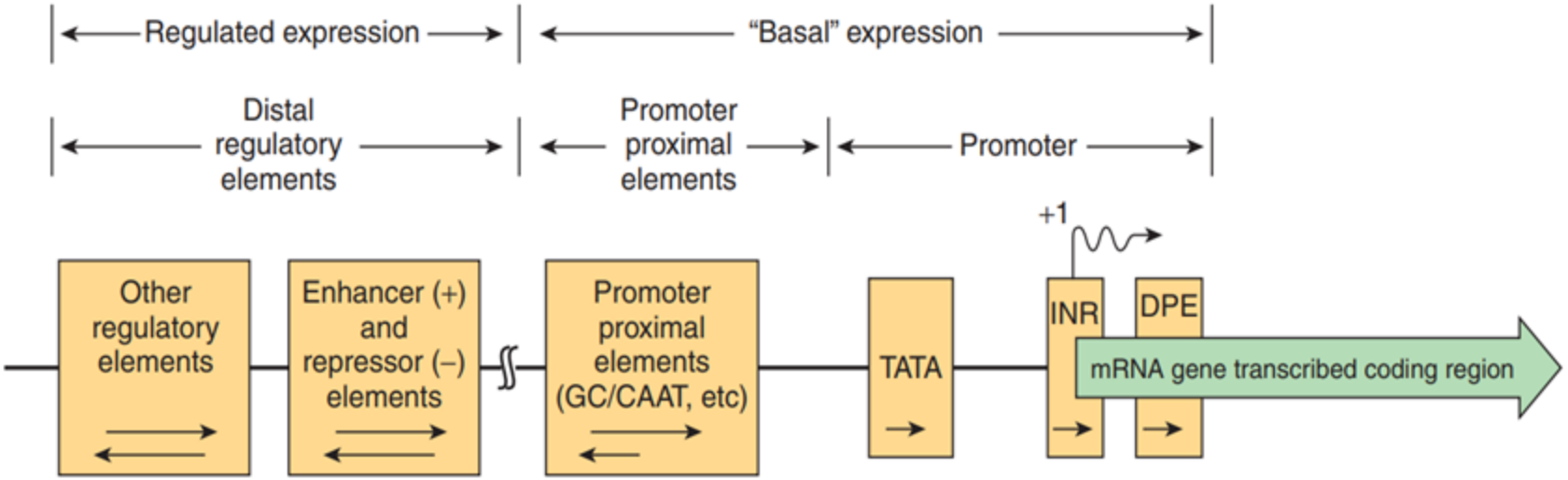
Prokaryotes	Eukaroytes
Simple	More complex
One RNAP	3 distinct RNAP
Promoter site – Pribnow box 35 sequence	Promoter site – TATA box – Hogness box , CAAT box
Initiation – Only requires sigma factor	Initiation – 6 Transcription factors interact with eukaryotic promoter region.
	POST TRANSCRIPTIONAL MODIFICATION

TRANSCRIPTION------
(EUKARYOTES)

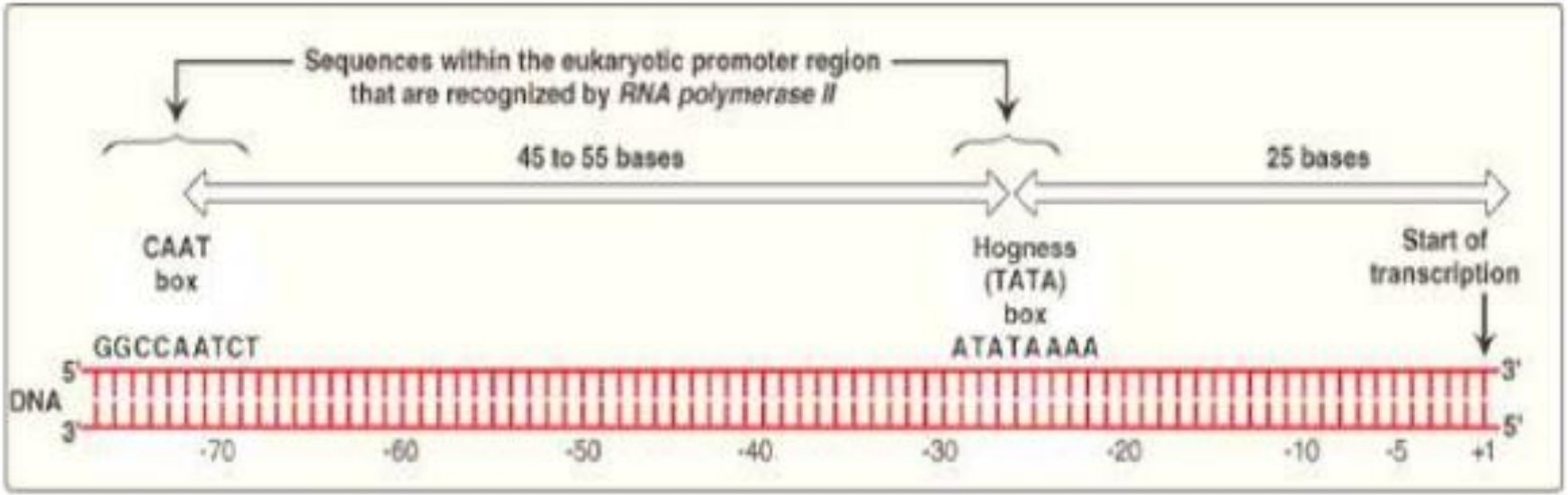
EUKARYOTIC PROMOTERS

- **TATA BOX OR GOLDBERG HOGNESS BOX**- Sequence is TATAAAAG. Located at -25bp. To this box mainly the TATA box binding protein (TBP) of 34 kDa attached in turn to bind TAF's
- **CAAT BOX**- located at 70 to 80 bp. Sequence is CAAT or 5' CCAAT 3'. It is cis acting regulatory element. Mutation of this region lowes the rate of transcription.
- **GC BOX**- it contains the sequence as 5' GGGCGC 3' at the core and is bound by the transcriptional factor Sp1. it is about 110 bp upstream from transcription initiation site (TSS)

PROMOTERS FOR EUKARYOTES



PROMOTER FOR EUKARYOTES



GENE CONTROL REGIONS

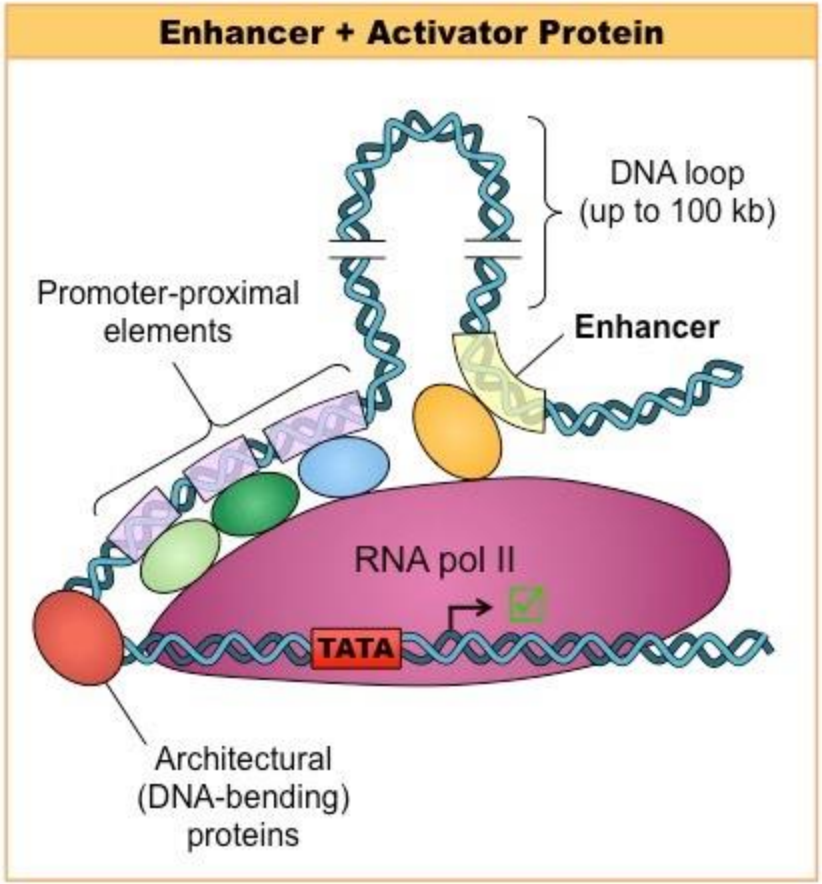
- Transcription is the process of RNA synthesis, controlled by the interaction of promoters and enhancers.
- **Start site**. A start site for transcription.
- **A promoter**. A region a few hundred nucleotides 'upstream' of the gene (toward the 5' end). It is not transcribed into mRNA, but plays a role in controlling the transcription of the gene. Transcription factors bind to specific nucleotide sequences in the promoter region and assist in the binding of RNA polymerases.
- **Enhancers**. Some transcription factors (called activators) bind to regions called 'enhancers' that increase the rate of transcription. These sites may be thousands of nucleotides from the coding sequences or within an intron. Some enhancers are conditional and only work in the presence of other factors as well as transcription factors.
- **Silencers**. Some transcription factors (called repressors) bind to regions called 'silencers' that depress the rate of transcription.

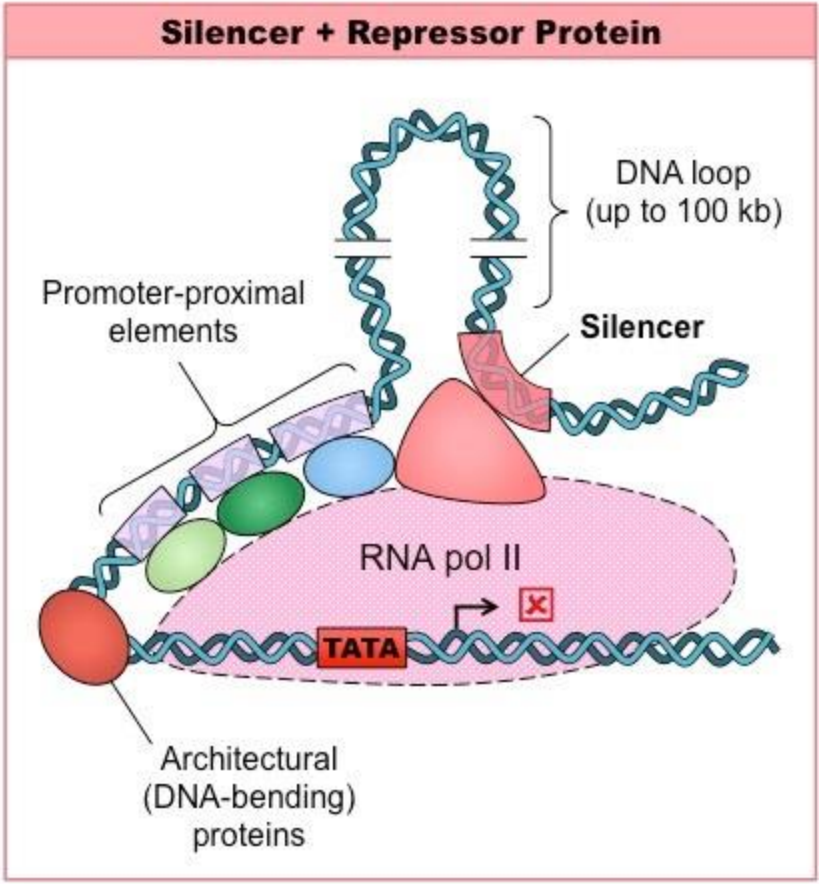
RNA POLYMERASE IN EUKARYOTES

IT IS OF THREE TYPES AS MENTIONED BELOW:

Table 29.2 Eukaryotic RNA polymerases

Type	Location	Cellular transcripts	Effects of α-amanitin
I	Nucleolus	18S, 5.8S, and 28S rRNA	Insensitive
II	Nucleoplasm	mRNA precursors and snRNA	Strongly inhibited
III	Nucleoplasm	tRNA and 5S rRNA	Inhibited by high concentrations


Enhancer + Activator Protein
DNA loop
(up to 100 kb)
Promoter-proximal
elements
Enhancer
RNA pol II
TATA
Architectural
(DNA-bending)
proteins


Silencer + Repressor Protein
DNA loop
(up to 100 kb)
Promoter-proximal
elements
Silencer
RNA pol II
TATA
Architectural
(DNA-bending)
proteins

STAGES OF TRANSCRIPTION

- The transcription is divided into various stages that are:
- **PRE- INITIATION COMPLEX FORMATION**
- **INITIATION**
- **ELONGATION**
- **TERMINATION**
- **POST TRANSCRIPTIONAL MODIFICATIONS**

1. PRE INITIATION COMPLEX (PIC)

- ★ For the process of transcription in eukaryotes the first step is formation of a pre initiation complex
- ★ It includes the binding of TBP (TATA BOX BINDING PROTEIN) that binds to the TATA box
- ★ Some general transcriptional factors (GTF) are also involved, these are: TFIID, TFIIA, TFIIB, TFIIF, TFIIH, TFIIE
- ★ TBP mainly contain TFIId as its subunit
- ★ TBP + TAF(TBP associated factors) = TFIID
- ★ So

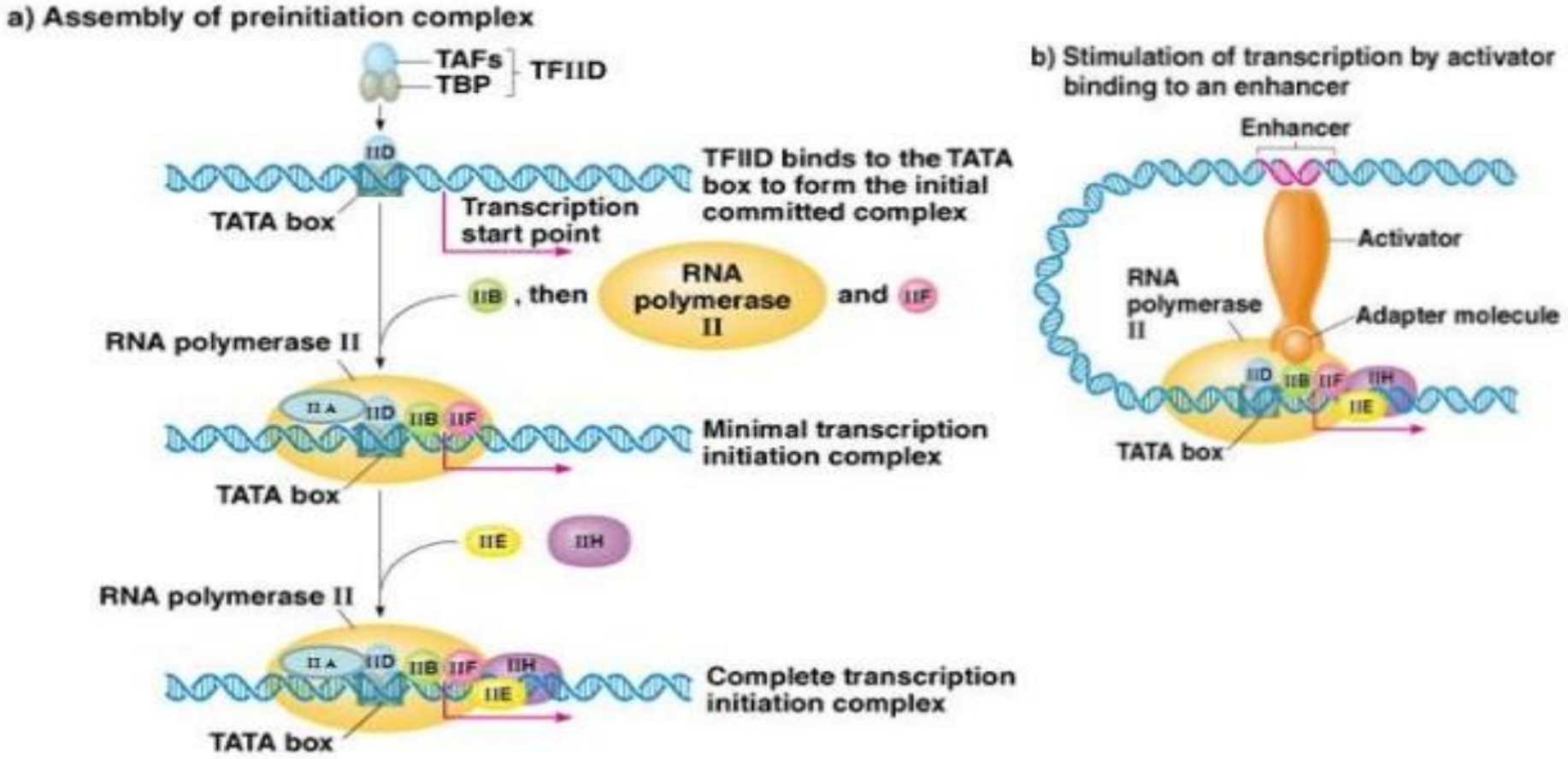
GTF + RNA POLYMERASE II = PIC

GTF	Function
TFIID (TBP component)	Recognition of the TATA box and possibly Inr sequence; forms a platform for TFIIB binding
TFIID (TAFs)	Recognition of the core promoter; regulation of TBP binding
TFIIA	Stabilizes TBP and TAF binding
TFIIB	Intermediate in recruitment of RNA polymerase II; influences selection of the start point for transcription
TFIIF	Recruitment of RNA polymerase II
TFIIE	Intermediate in recruitment of TFIIH; modulates the various activities of TFIIH
TFIIH	**Helicase** activity responsible for the transition from the closed to open promoter complex; possibly influences promoter clearance by phosphorylation of the C-terminal domain (**cinase** activity) of the largest subunit of RNA polymerase II

GTF- dictate the starting point and direction of transcription (low level of transcription)

Initiation in Eukaryote

Order of binding is: IID + IIA + IIB + RNA poly. II + IIF +IIE +IIH



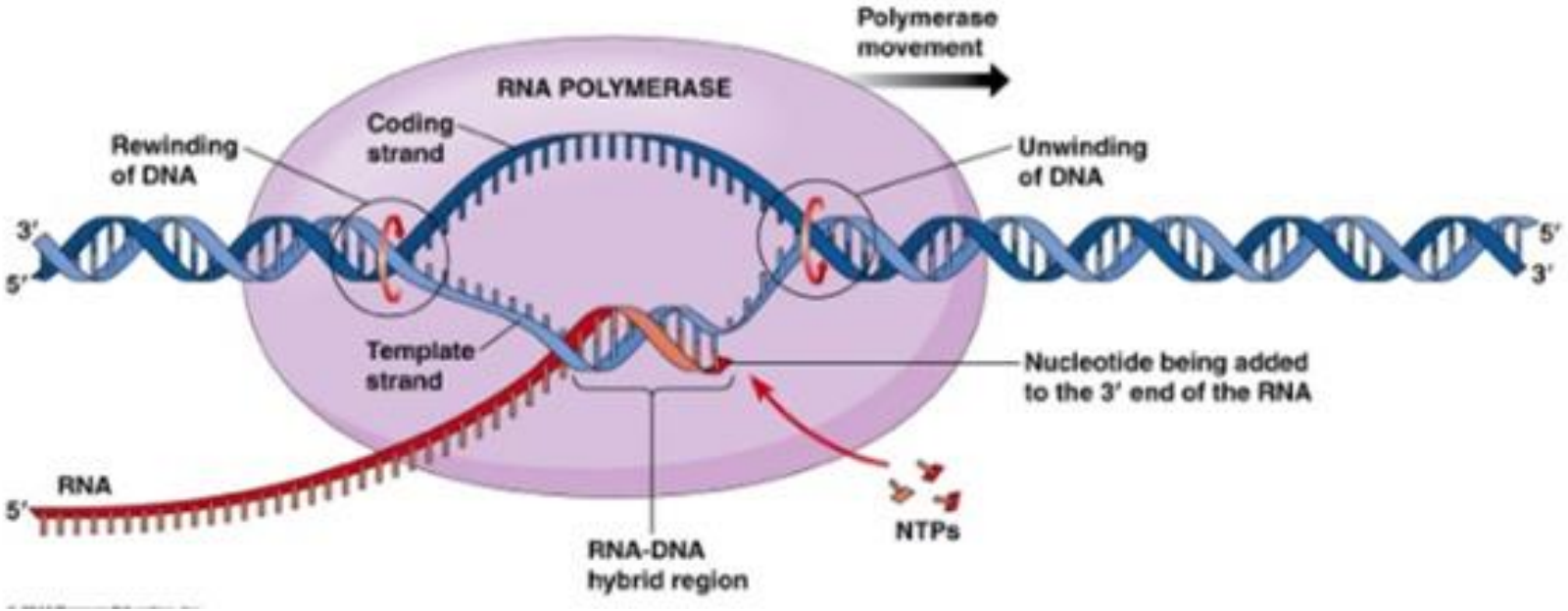
2. ELONGATION

TFIIH has two functions:

1. It is a helicase, which means that it can use ATP to unwind the DNA helix, allowing transcription to begin.
2. In addition, it phosphorylates RNA polymerase II which causes this enzyme to change its conformation and dissociate from other proteins in the initiation complex.

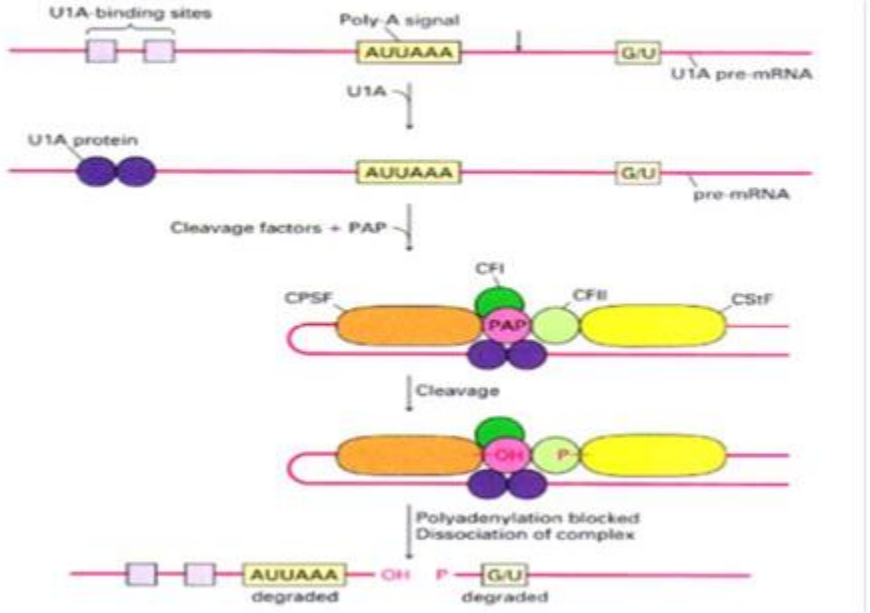
- The key phosphorylation occurs on a long C-terminal tail called the C-terminal domain (CTD) of the RNA polymerase II molecule.
- Interestingly, only RNA polymerase II that has a non-phosphorylated CTD can initiate transcription but only an RNA polymerase II with a phosphorylated CTD can elongate RNA.
- RNA polymerase II now starts moving along the DNA template, synthesizing RNA, that is, the process enters the elongation phase.
- RNA synthesis occurs in the 5′ → 3′ direction with the RNA polymerase catalyzing a nucleophilic attack by the 3-OH of the growing RNA chain on the alpha-phosphorus atom on an incoming ribonucleoside 5-triphosphate.
- The RNA molecule made from a protein-coding gene by RNA polymerase II is called a primary transcript.

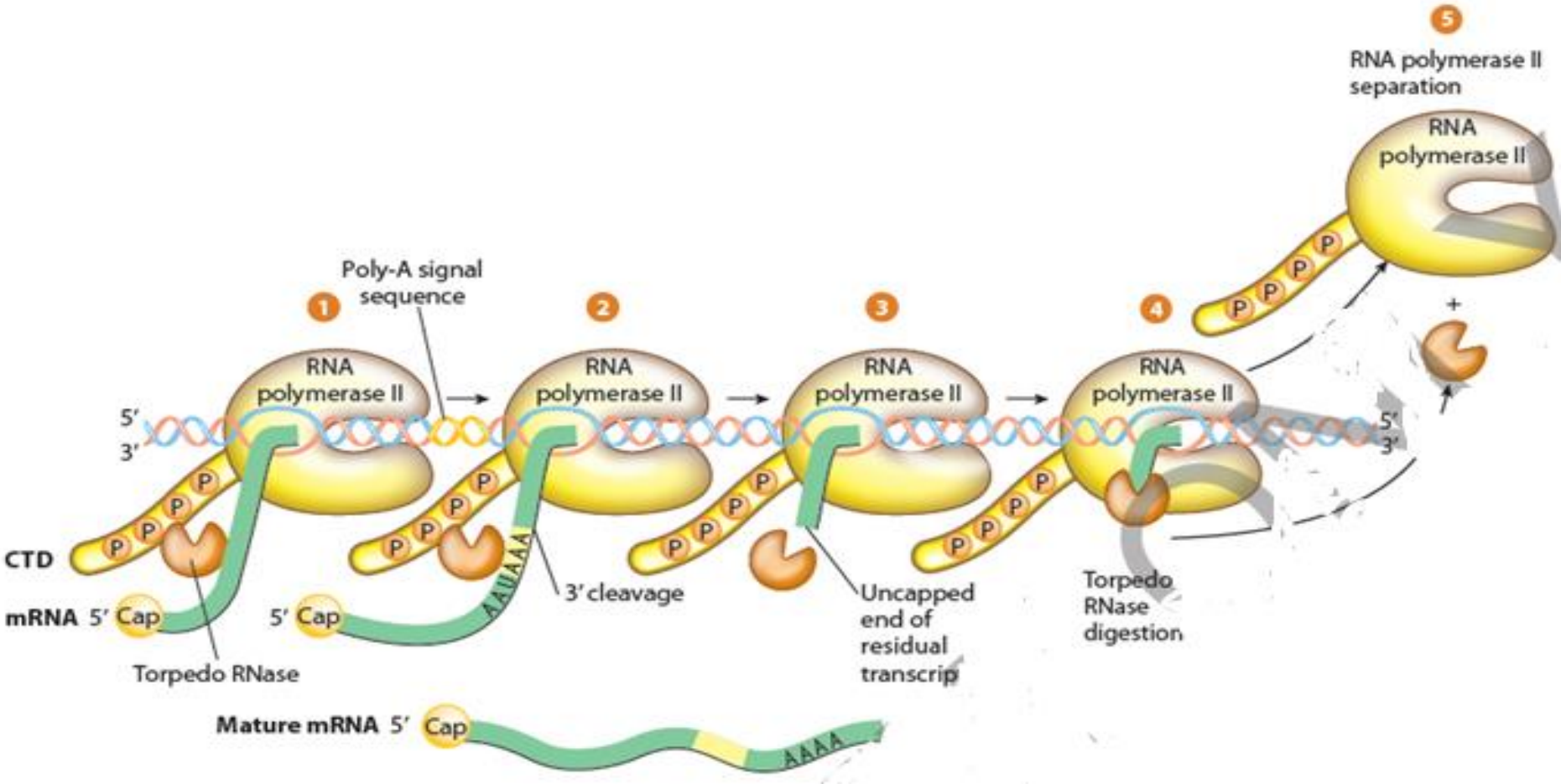
ELONGATION


Polymerase movement
RNA POLYMERASE
Coding strand
Rewinding of DNA
Unwinding of DNA
3′
5′
5′
3′
Template strand
Nucleotide being added to the 3′ end of the RNA
RNA
5′
NTPs
RNA-DNA hybrid region

Termination in eukaryotic transcription

- Eukaryotic protein genes contain a poly-A signal located downstream of the last exon.
- This signal is used to add a series of adenylate residues during RNA processing. Transcription often terminates at 0.5 - 2 kb downstream of the poly-A signal,
- but the mechanism is unclear




5
RNA polymerase II separation
RNA polymerase II
Poly-A signal sequence
1
2
3
4
RNA polymerase II
RNA polymerase II
RNA polymerase II
RNA polymerase II
5′
3′
5′
3′
P
CTD
mRNA 5′ Cap
Torpedo RNase
5′ Cap
AAUAAA
3′ cleavage
Uncapped end of residual transcrip
Torpedo RNase digestion
+
Mature mRNA 5′ Cap
AAAA

POST TRANSCRIPTIONAL MODIFICATIONS

- They are of THREE types:
- CAPPING
- TAILING
- SPLICING

1. CAPPING

- 7-methylguanylate attached by a unusual 5'-5' triphosphate linkage to the ribose at the 5'-end.

- Addition of GTP part of the cap is catalyzed by nuclear enzyme guanylyltransferase.

- Methylation of terminal guanine occurs in the cytosol- **SAM is the source of the methyl group**

- Catalysed by guanine-7-methyl transferase.

Synthesis of 5'-Cap on Eukaryotic mRNAs

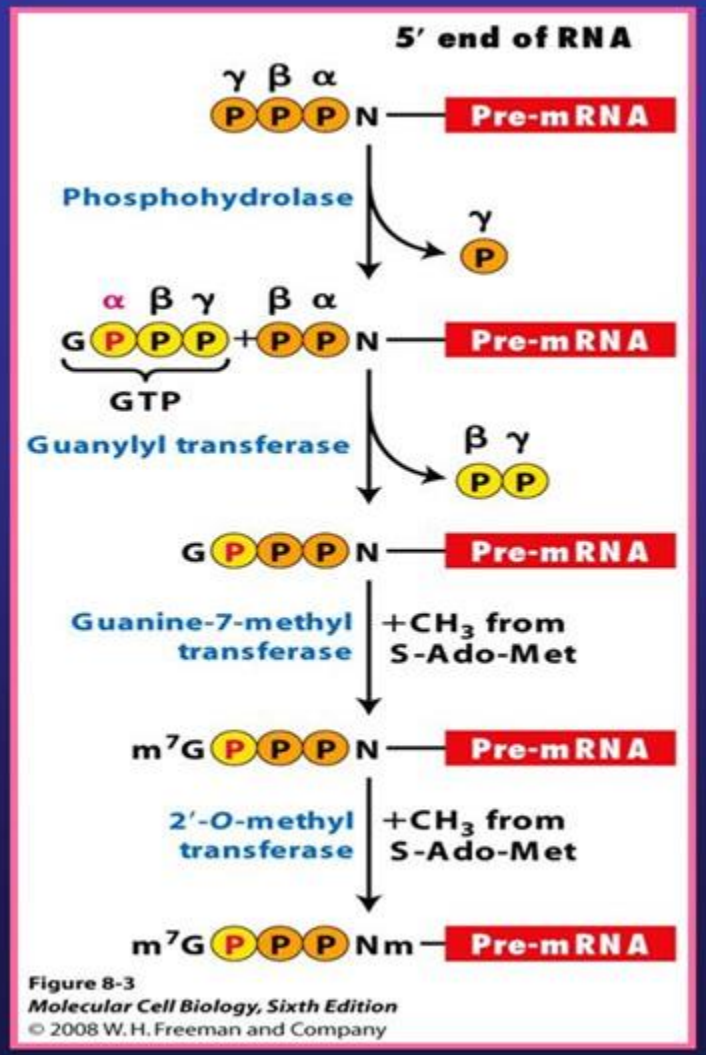
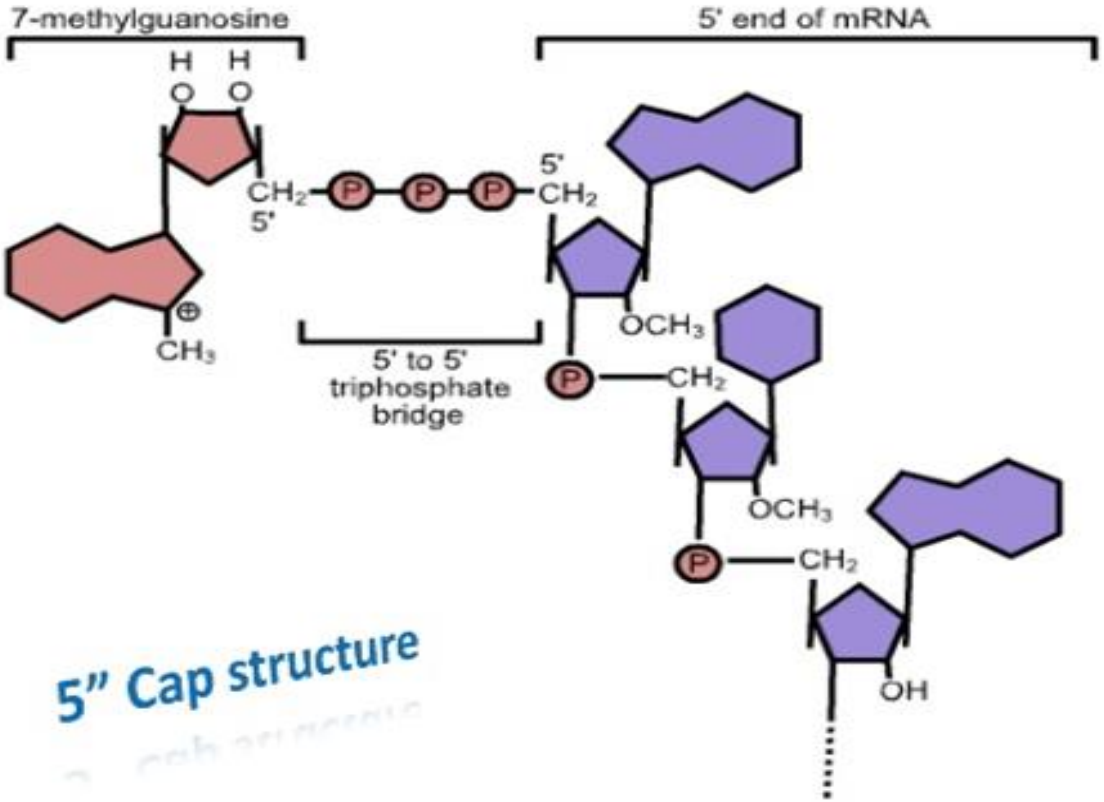


Figure 8-3
Molecular Cell Biology, Sixth Edition
© 2008 W. H. Freeman and Company

- Capping occurs shortly after initiation of transcription
- 7-methyl-G is added in the 5' end of the nascent RNA shortly after transcription initiates, about 25-30 nucleotides in length.
- The enzyme involved in this process is a dimeric capping enzyme associated with the phosphorylated carboxyl-terminal domain (CTD) of Pol II. Capping is specific for transcripts produced by Pol II
- The γ-phosphate is removed from the nascent RNA, replaced with a GMP (5'-5' triphosphate structure), and a methyl group from S-adenosyl-methionine is added to the N7 position of the G and the 2' oxygen of the 5' ribose at the nascent RNA
- Capping of the nascent transcript is coupled to elongation so that all of the transcripts will be capped
- Capping of mRNA will protect it from degradation by 5'-exonuclease



5" Cap structure

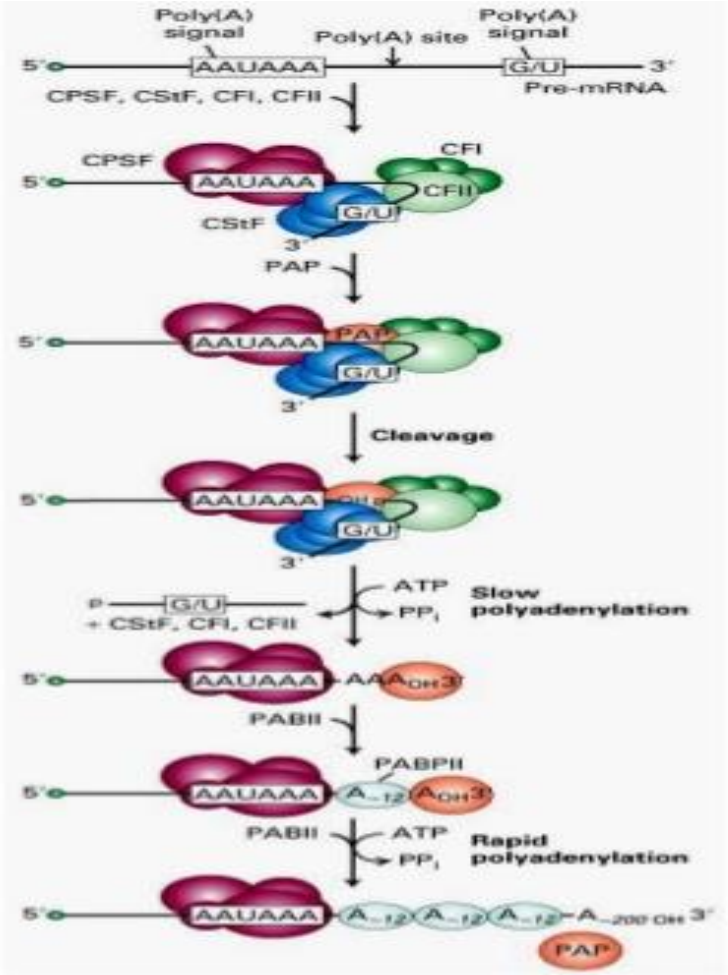
Importance

- The cap binds mature mRNA to the ribosome during protein biosynthesis.
- Cap Stabilizes mRNAs against digestion by ribonucleases.
- Eukaryotic mRNAs lacking the cap are not translated efficiently.

2. POLYADENYLATION OR TAILING

The Poly A Tail

- Post-transcriptional RNA processing at the opposite end of the transcript comes in the form of a **string of adenine bases** attached to the end of the synthesized RNA chain.
- The addition of the adenines is catalyzed by the enzyme poly (A) polymerase.
- The mRNA is first cleaved about 20 nucleotides downstream from an AAUAA recognition sequence
- Another enzyme, poly(A) polymerase, adds a poly(A) tail which is subsequently extended to as many as 200 A residues.
- The **poly(A) tail** appears to protect the 3' end of mRNA from 3' 5' exonuclease attack.
- Histone and interferon's mRNAs lack poly A tail.
- After the m-RNA enters the cytosol, the poly A tail is gradually shortened.



Polyadenylation of mRNA at the 3' end

CPSF: ***cleavage and polyadenylation specificity factor*** **binds upstream AAUAAA poly(A) Signal 5' end.**

CStF: ***cleavage stimulatory factor F*** **interacts with a downstream GU- sequence & bound with CPSF forming a loop in RNA**

CFI & CFII: ***cleavage factor I & II.***

PAP: ***poly(A) polymerase stimulates cleavage at poly A site***
Bound PAP adds ≈12 A residues at a slow rate to 3'-OH group

PABPII: ***poly(A)-binding protein II.***

PABPII (short poly A tail) accelerates rate of addition of A by PAP

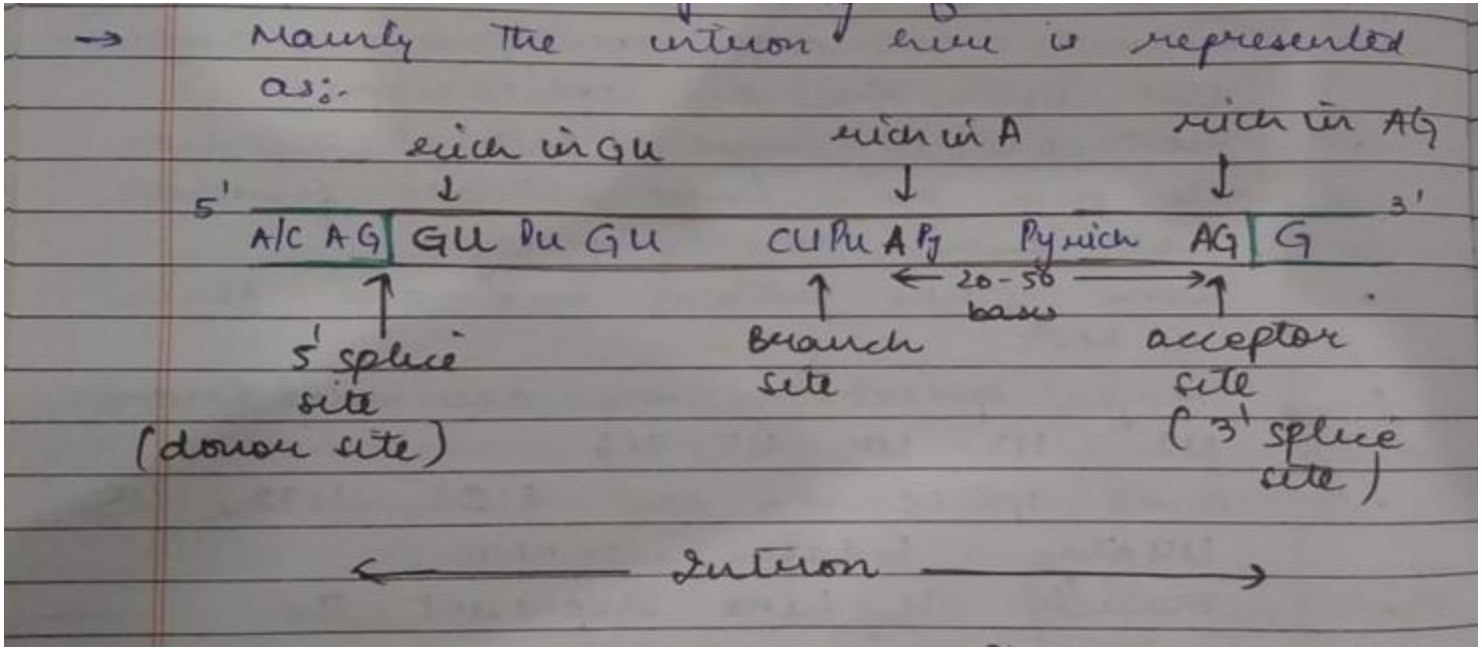
After 200–250 A residues have been added, PABPII signals PAP to stop polymerization

Poly (A) tail controls mRNA stability & influences translation

3. SPLICING

- The eukaryotic primary mRNAs are made up of two types of segments; non-coding introns and the coding exons.
- The introns are removed by a process called RNA splicing where ATP is used to cut the RNA, releasing the introns and joining two adjacent exons to produce mature mRNA.
- The splicing can be of various mechanisms that is:
- **(A) GROUP 1 INTRON**
- **(B) GROUP 2 INTRON**
- **© GROUP 3 INTRON (SPLICEOSOME COMPLEX)**
- **(D) GROUP 4 INTRON**
- **(E) ALTERNATIVE SPLICING**

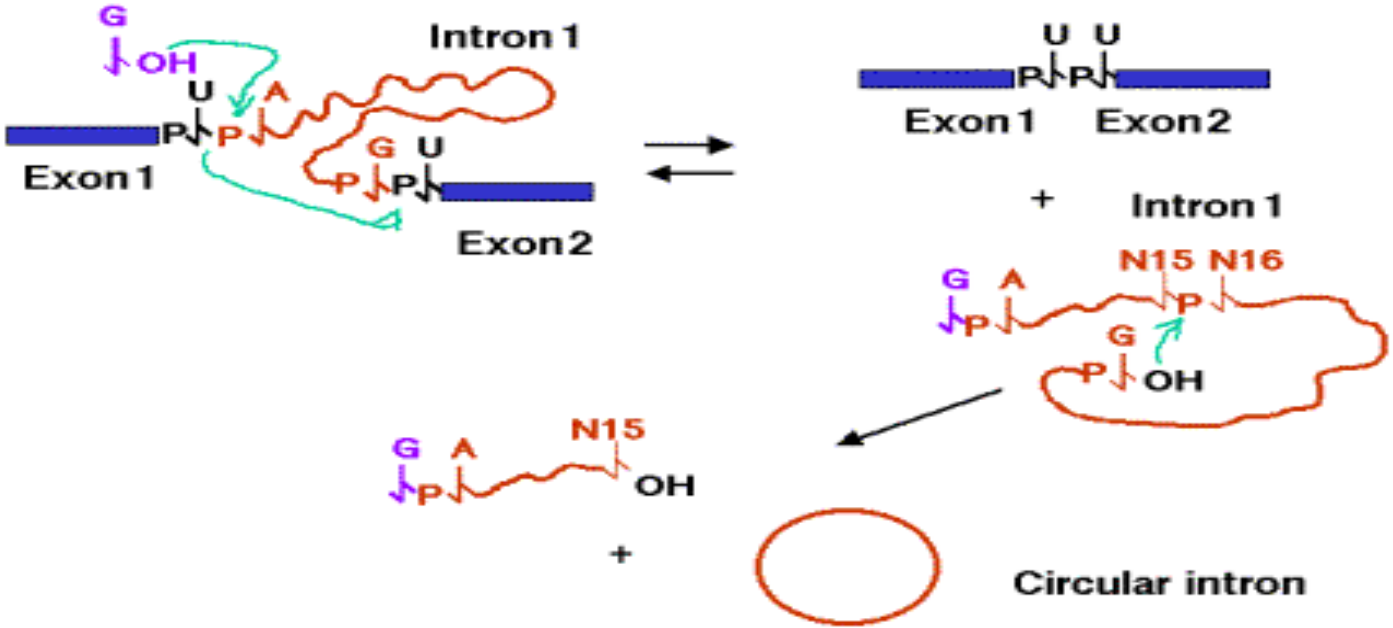
BASIC STRUCTURE OF INTRON



GROUP 1 INTRON - self splicing

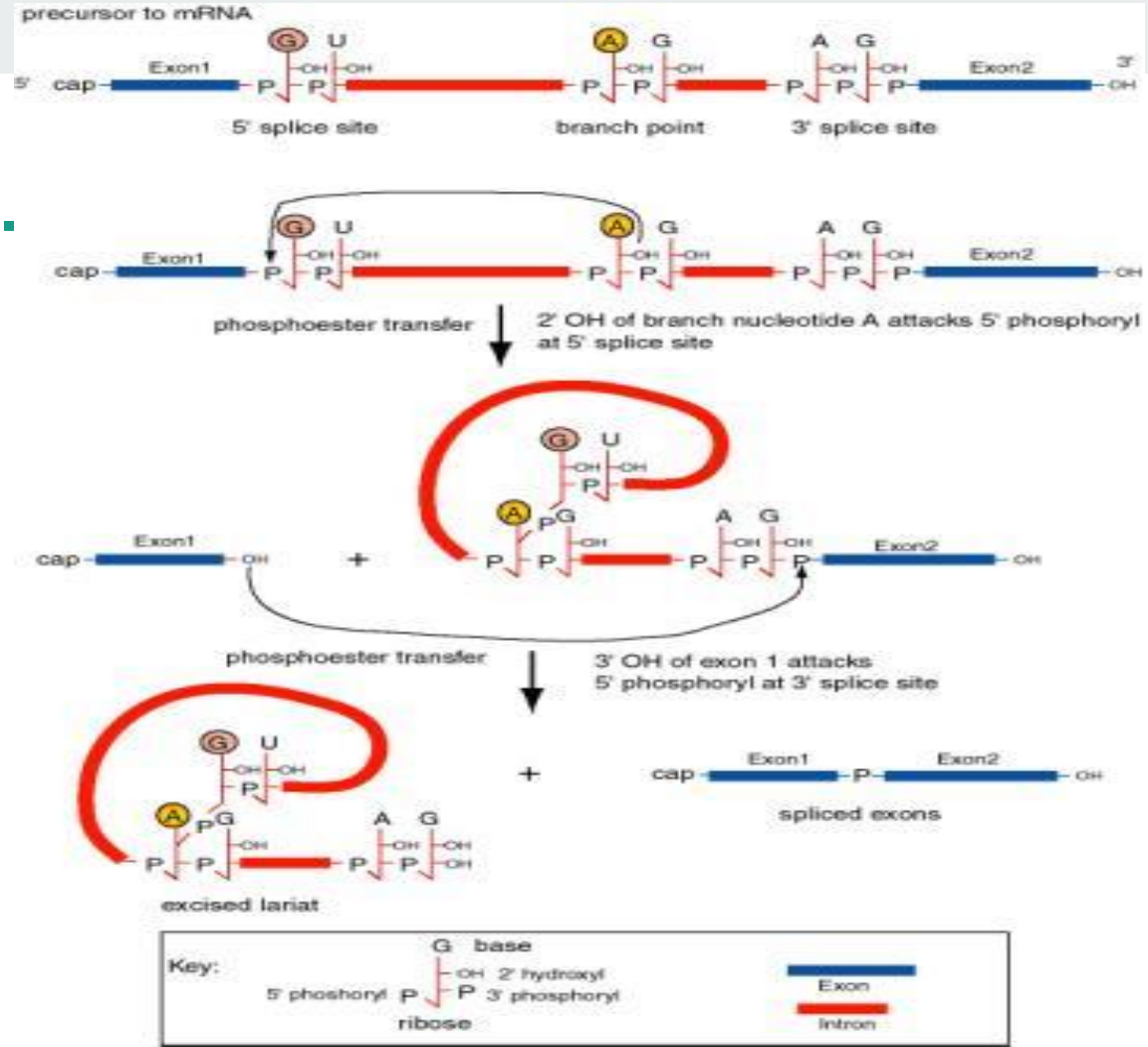
- *The 3'-OH of the guanine nucleotide is the nucleophile that attacks and joins to the 5' phosphate of the first nucleotide of the intron. This leaves the 3'-OH of the last nucleotide of the upstream exon available to attack and join the 5' phosphate of the first nucleotide of the downstream exon. These two phosphoester transfers result in a joining of the two exons and excision of the intron (with the initiating G nucleotide attached to the 5' end.) The excised intron is then circularized by attack of the 3'-OH of the last nucleotide of the intron on the phosphate between the 15th and 16th nucleotides of the introns. Further degradation effectively removes the intron from the reaction and helps prevent the reverse reaction from occurring. Note that the phosphoester transfers are readily reversible unless the products (excised intron) are removed. There is no increase or decrease in the number of phosphoester bonds during this splicing.*

Self-splicing by a phosphoester transfer mechanism



GROUP 2 INTRON- self splicing

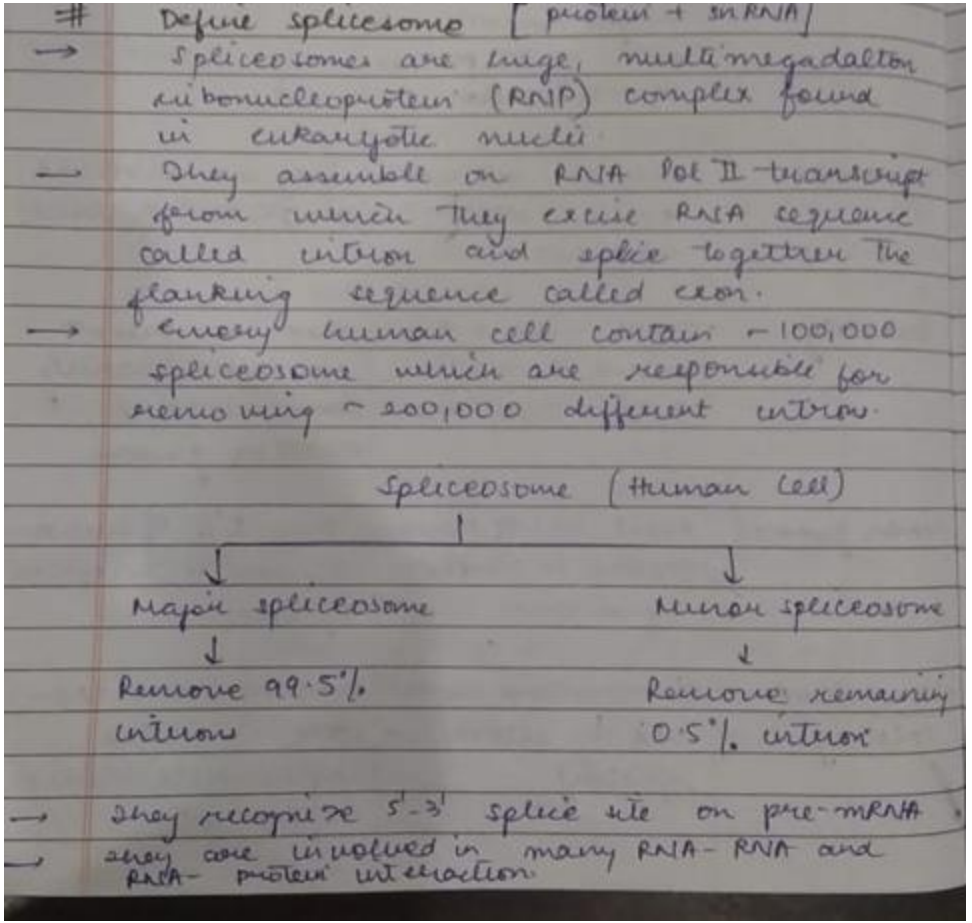
- *The 2'-OH of a highly conserved bulged A nucleotide located within domain 6 of the intron attacks the 5'-splice site. This results in release of the 5'-exon and formation of a lariat structure whose 5'-end of the intron is covalently attached to the 2'-OH of the bulged A. Alternatively, the 5'-exon can be released by hydrolysis in which water is the nucleophile. The second step of splicing involves attack by the 5'-exon on the 3'-splice site. This results in ligation of the flanking exons and release of the lariat or linear intron*



GROUP2 INTRON SPLICING

GROUP 3 INTRON

THIS IS A NOTE ON
SPLICEOSOME AND snurps



⇒ [COMPOSITION]

- They contain both (60) protein and snRNA. (100-150 nucleotides in length)
- Yeast have ~100 spliceosomal protein whereas 300 diff. proteins are associated with human spliceosome.
- Most protein have RNA recognition site and some have NTPase function that ensure the fidelity.
- Some other protein bind and form snRNPs
- Major spliceosome are assembled from U1, U2, U4, U5, U6
- Minio spliceosome are U11, U12, U5, U4atac, U6atac snRNPs
- Mainly U here represent the uridine rich snRNA.
- Later U7 snRNA is required for histone mRNA 3'-end processing

snRNPs [small nuclear Ribo nucleo protein]

→ They are RNA-protein complexes that combine with unmodified pre-mRNA and various other protein to form spliceosome, a large protein complex upon which splicing of pre-mRNA occur

→ They mainly remove the intron and exise the exons.

→ The two components are protein and RNA.

→ Sn RNA in snRNP has a ribosomal RNA like activity as it directly incorporate both enzymatic and structural role

SPLICEOSOME MEDIATED RNA PROCESSING

- The transesterification reactions are mediated by a huge molecular "machine" called the spliceosome.
- The spliceosome is the large complex made up of the snRNPs U1, U2, U4, U5, and U6.
- The snRNPs have three roles in splicing:

i. They recognize the 5' splice site and the branch site.
ii. They bring those sites together as required.
iii. They catalyze (or help to catalyze) the RNA cleavage and joining reactions

- Non-snRNPs are involved in splicing:

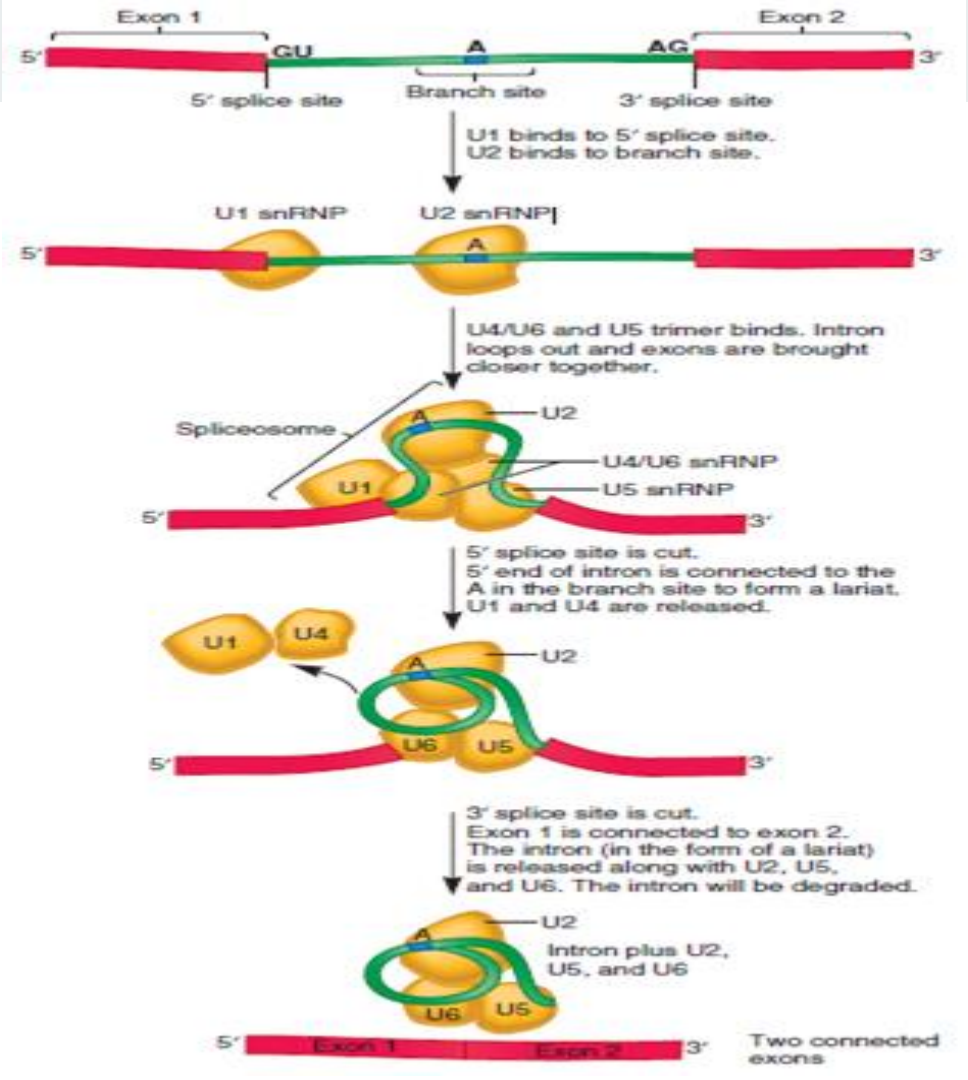
i. U2AF (U2 auxiliary factor)
ii. Branchpoint-binding protein (BBP)

- U11 and U12 components of the alternative spliceosome have the same roles in the splicing reaction as U1 and U2 of the major form, but they recognize distinct sequences. U4 and U6 have equivalent counterparts in both spliceosome forms.

TABLE 28.3 Small nuclear ribonucleoprotein particles (snRNPs) in the splicing of mRNA precursors

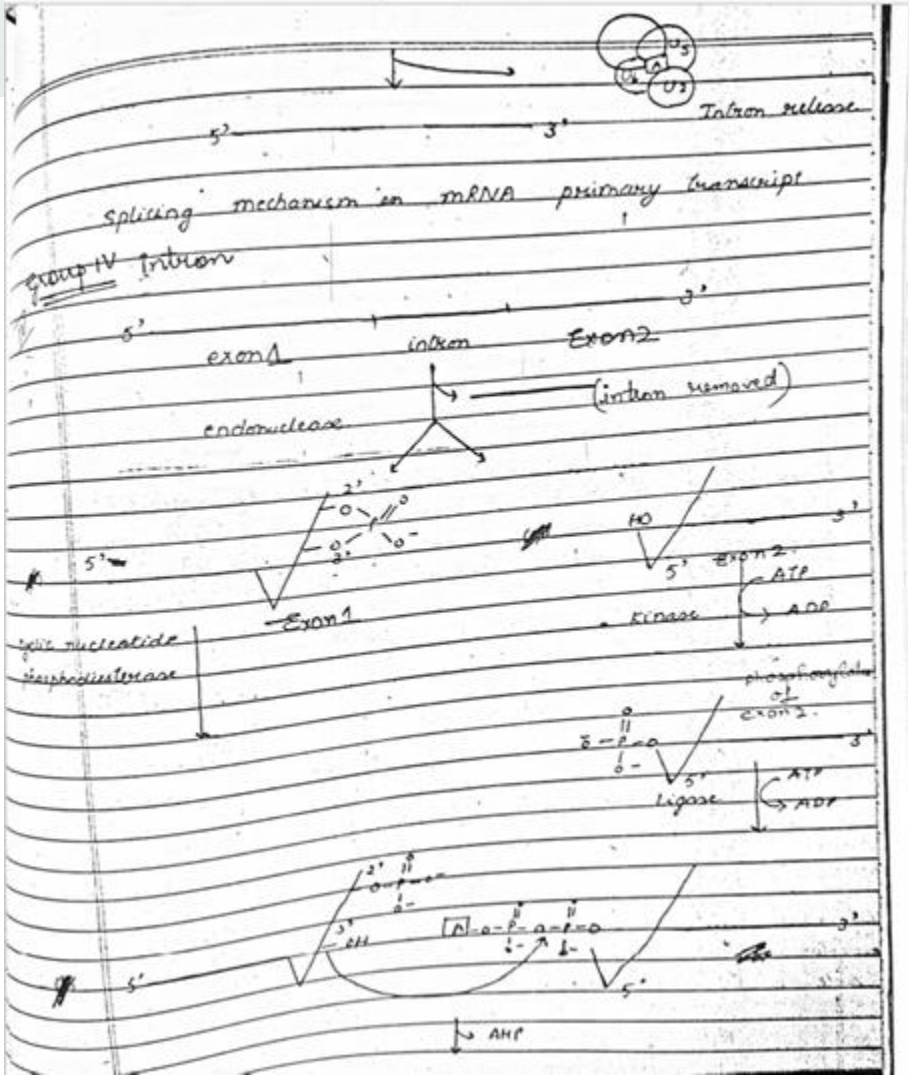
snRNP	Size of snRNA (nucleotides)	Role
U1	165	Binds the 5′ splice site and then the 3′ splice site
U2	185	Binds the branch site and forms part of the catalytic center
U5	116	Binds the 5′ splice site
U4	145	Masks the catalytic activity of U6
U6	106	Catalyzes splicing

ROLE OF SNRNPS IN SPLICING

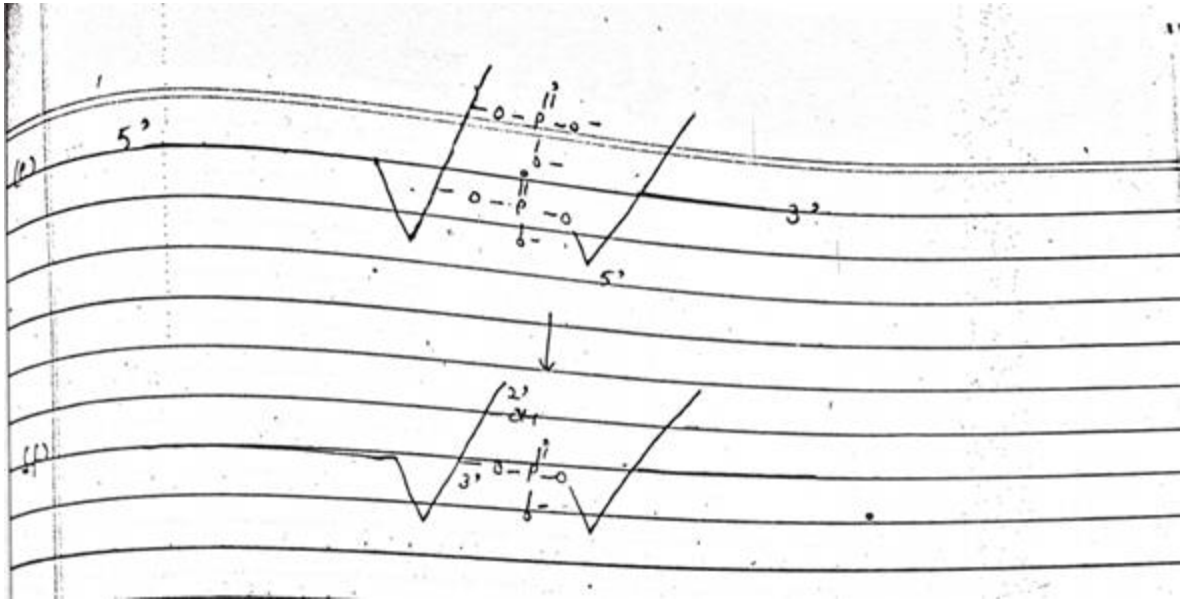


MECHANISM OF SPLICEOSOME

4. GROUP 4 INTRONS



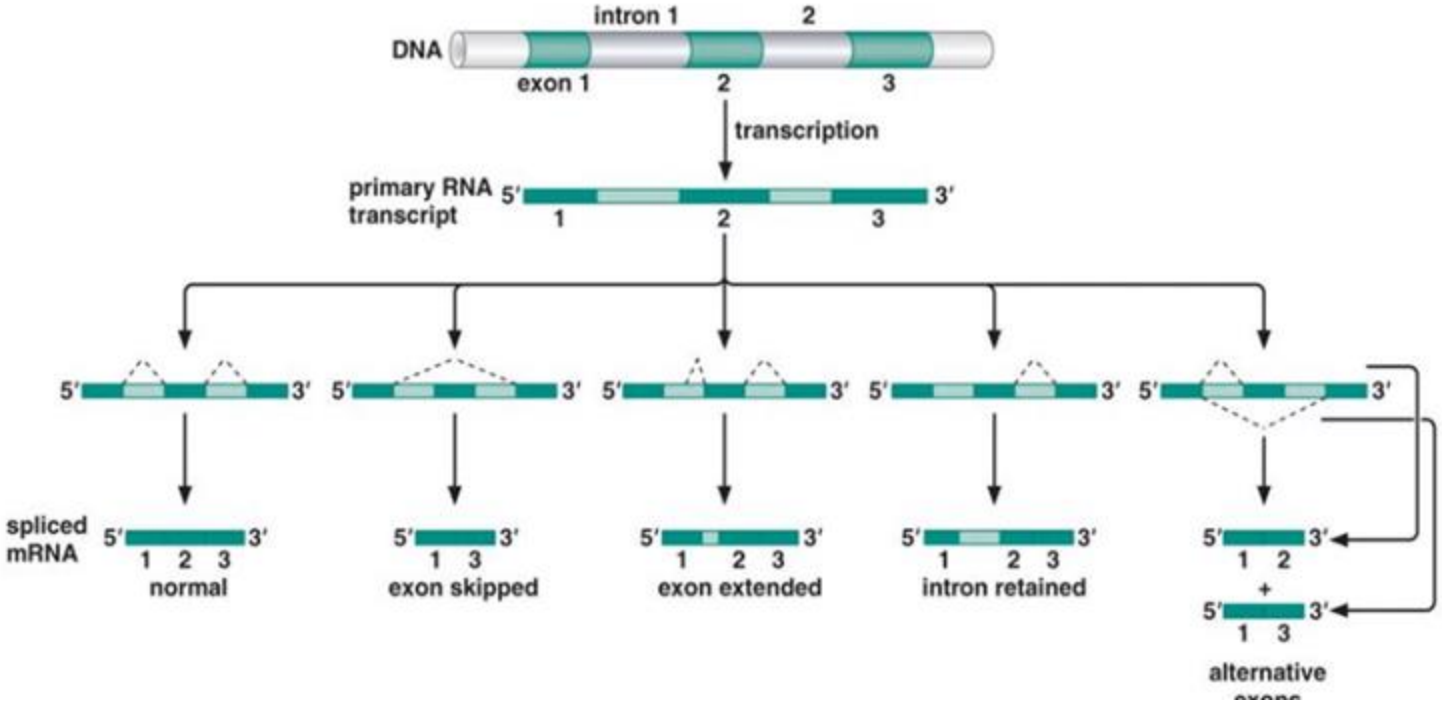
CONTD...

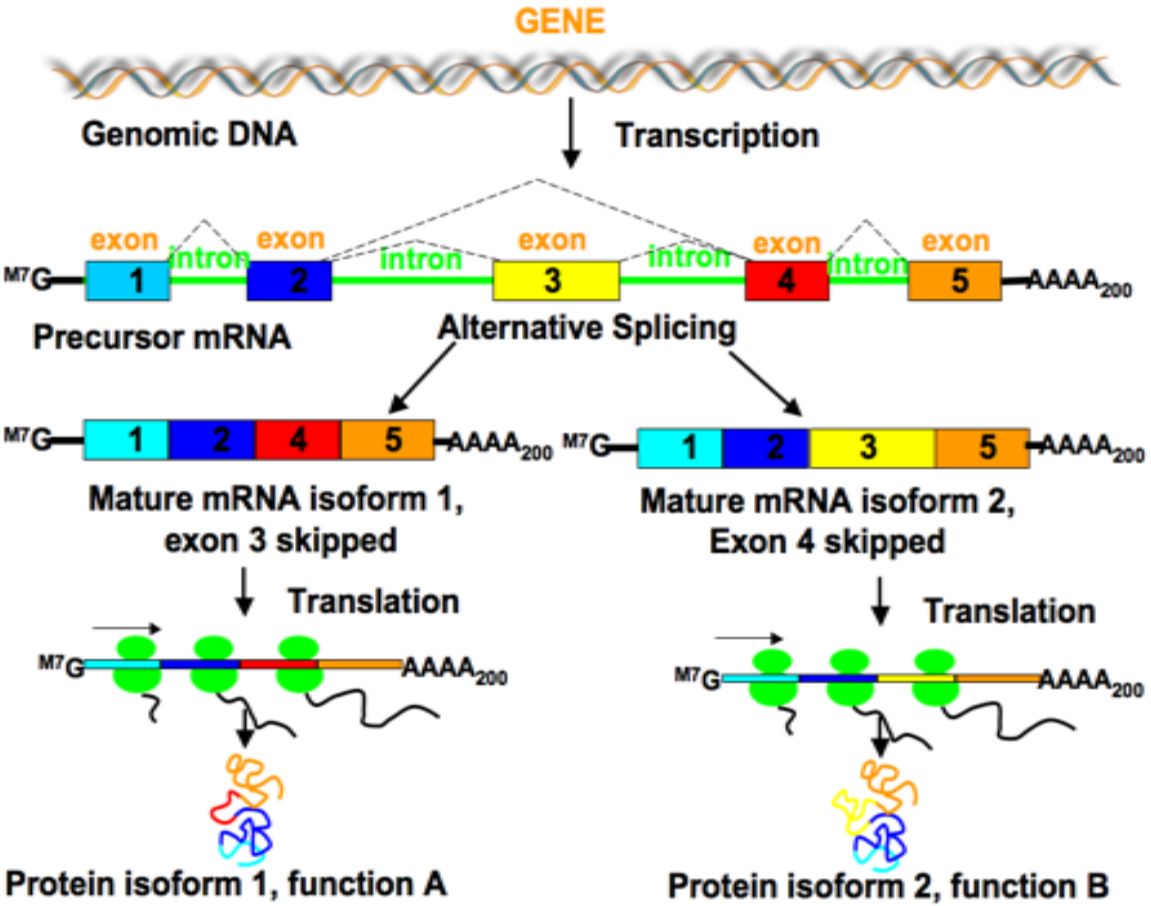


5. ALTERNATIVE SPLICING

- Alternative RNA splicing is a mechanism that allows different protein products to be produced from one gene when different combinations of introns (and sometimes exons) are removed from the transcript.
- This alternative splicing can be haphazard, but more often it is controlled and acts as a mechanism of gene regulation, with the frequency of different splicing alternatives controlled by the cell as a way to control the production of different protein products in different cells, or at different stages of development.
- Alternative splicing is now understood to be a common mechanism of gene regulation in eukaryotes; according to one estimate, 70% of genes in humans are expressed as multiple proteins through alternative splicing.

SlidePlayer





MECHANISM OF ALTERNATIVE SPLICING

THE END