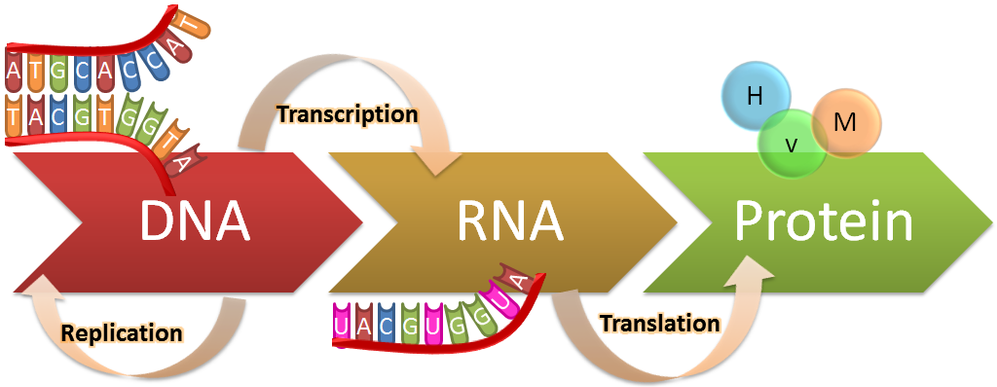
Transcription of Genes

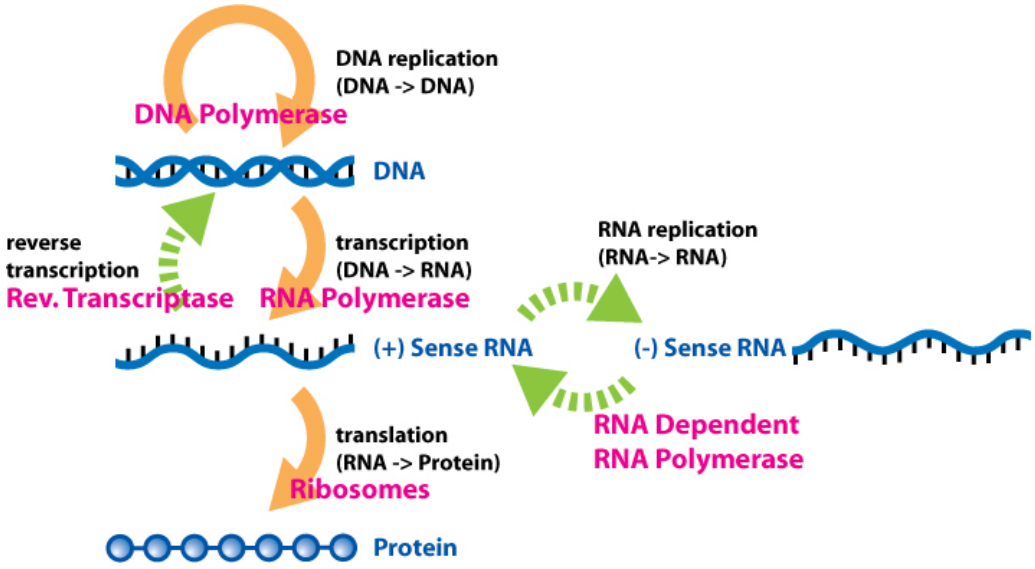


Dr. Nesrin Mwafi

Biochemistry & Molecular Biology Department

Faculty of Medicine, Mutah University

The Central Dogma of Molecular Biology

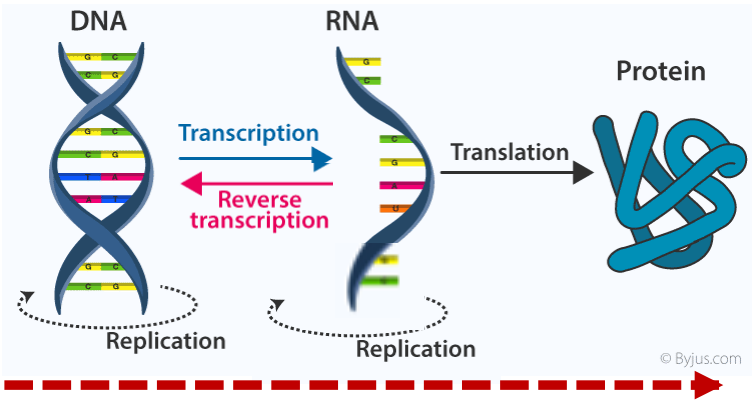


- Central dogma describes the flow of genetic information in living organisms
- Gene expression: is the process by which information from a gene is used in the synthesis of functional gene products (proteins or functional RNAs)
- Transcription is the first step in gene expression

Classes of information transfer

General	Special	Unknown
DNA → DNA	RNA → DNA	protein → DNA
DNA → RNA	RNA → RNA	protein → RNA
RNA → protein	DNA → protein	protein → protein

- General transfer
- Specific transfer
- Unknown transfer

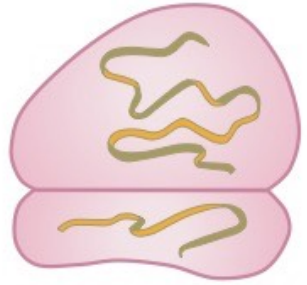


Function of different RNA Molecules

- There are different types of RNA molecules:
  1. Messenger RNA **(mRNA):** it is produced from the transcription of protein-coding genes. In eukaryotic cells, pre-mRNA (primary transcript) is modified to mature mRNA. In prokaryotes, single mRNA (bicistronic / polycistronic transcript) codes for different proteins

Messenger RNA (mRNA)

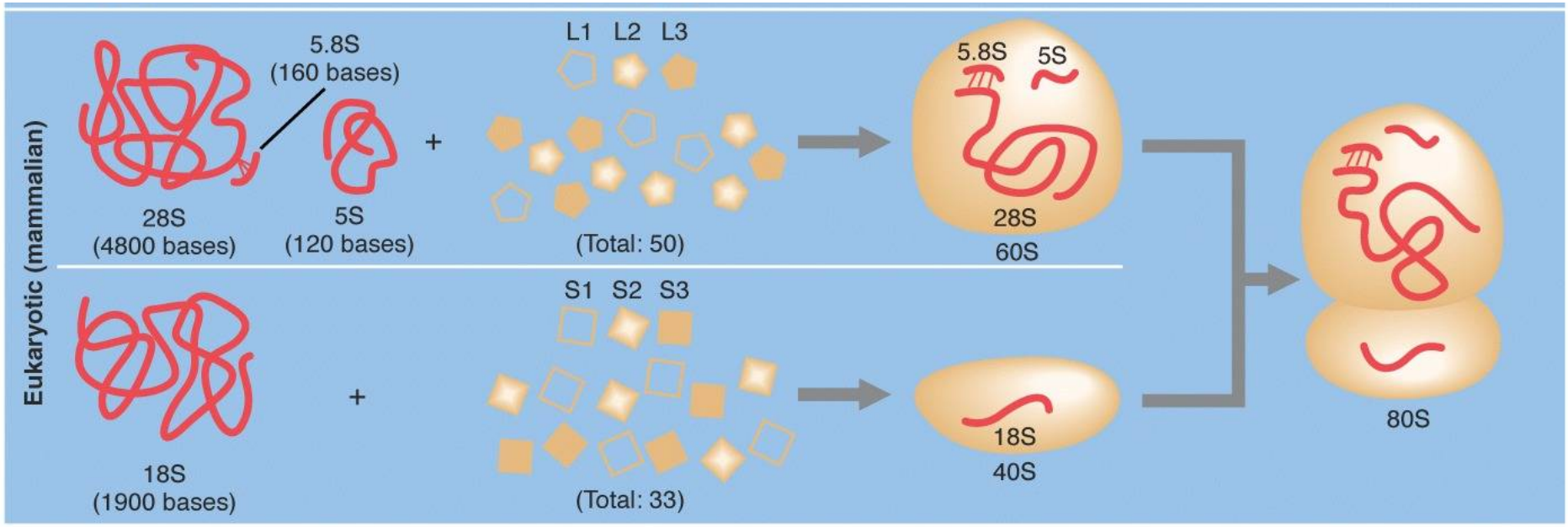
  2. Ribosomal RNA **(rRNA):** are specialized RNA molecules synthesized in the nucleolus. n the cytoplasm, they bind proteins form ribosomes (the machinery synthesize proteins)



Ribosomal RNA (rRNA)

Function of different RNA Molecules

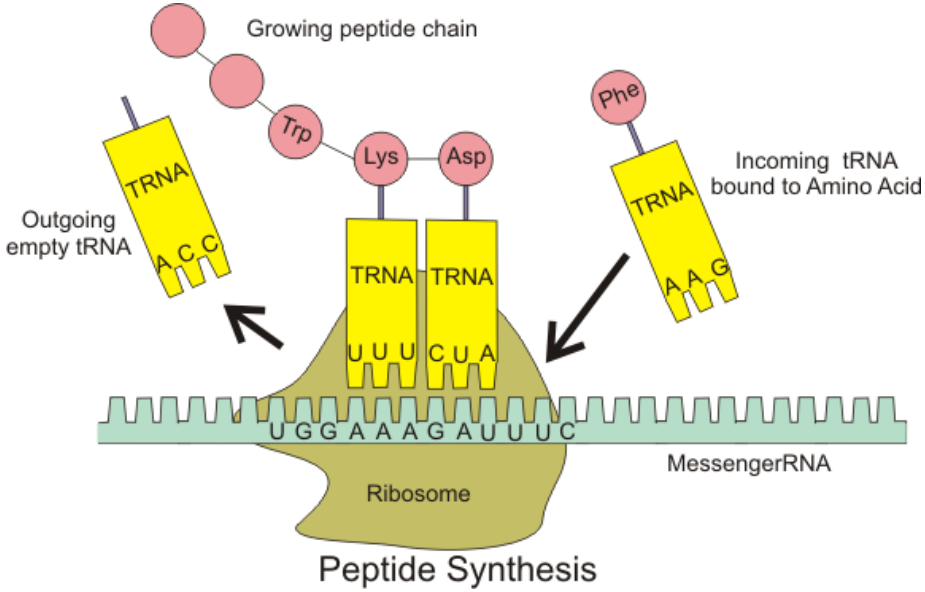
- Ribosome consists of small and large subunits. In eukaryotes (80S) the small subunit (40S) consists of 18S rRNA and 33 proteins while the large subunit (60S) contains: 5S rRNA, 5.8S rRNA, 28S rRNA and 50 proteins



Function of different RNA Molecules

3. Transfer RNA **(tRNA):** are specialized molecules (adaptors) that collect the proper amino acid, bring it to the ribosome and attach it to the growing polypeptide chain

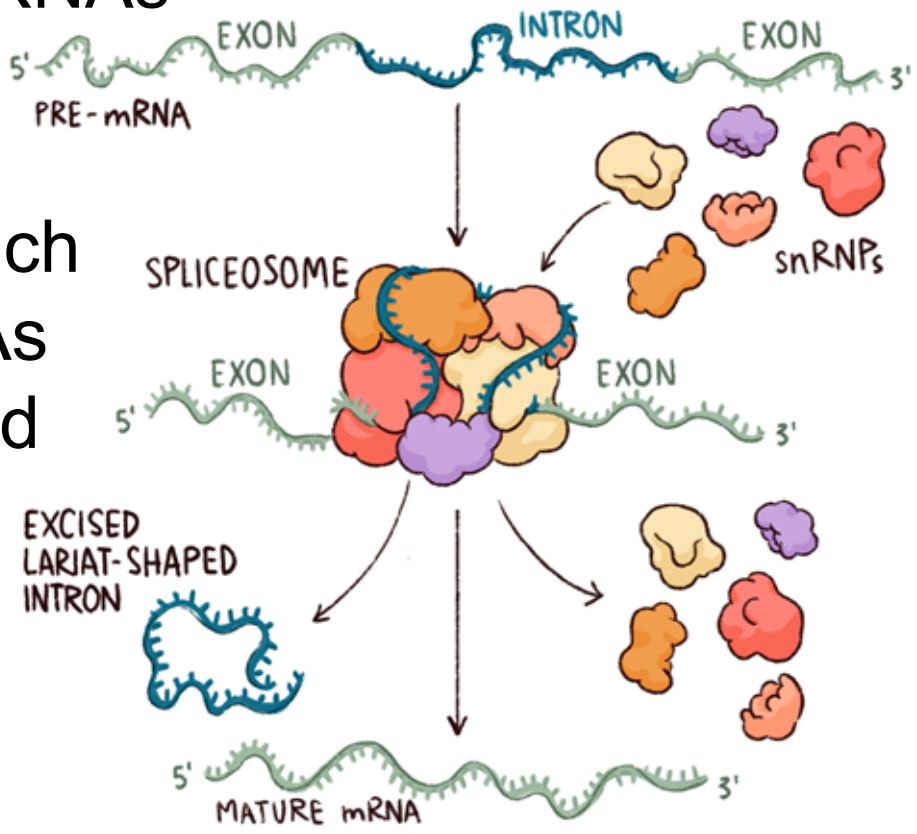
Transfer RNA (tRNA)



Peptide Synthesis

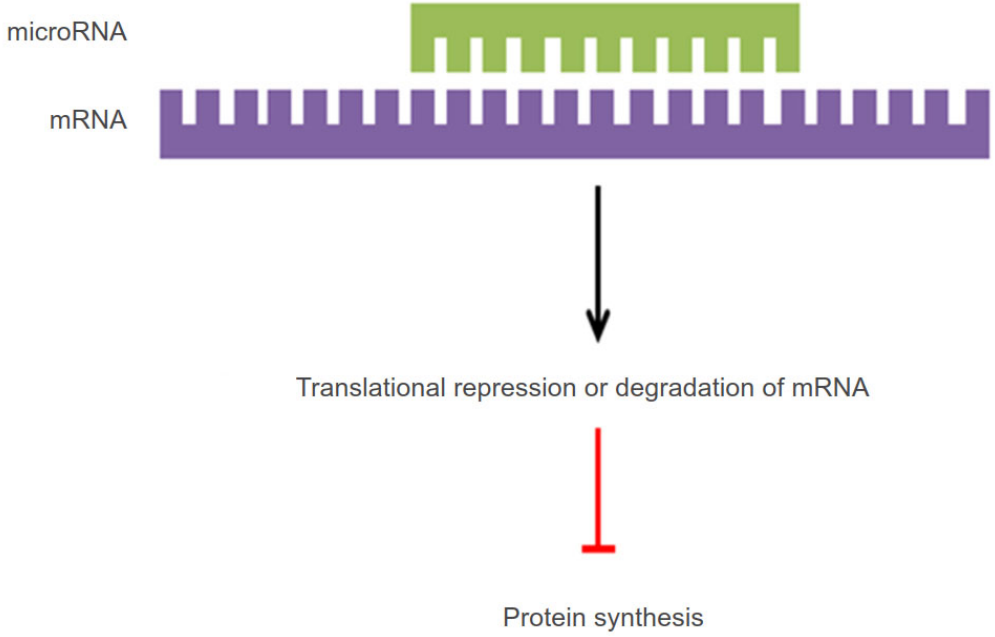
Function of different RNA Molecules

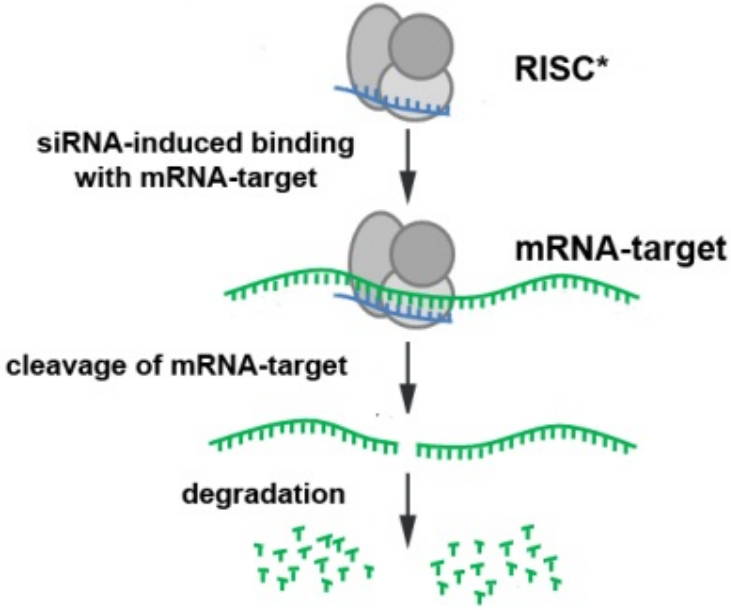
4. Small nuclear RNA **(snRNA):** is a class of small RNA molecules (150 nucleotides) which play an important role in pre-mRNA processing by removing introns (RNA splicing). snRNAs are main components of spliceosome (a large protein-RNA complex) which consists of 5 types snRNAs (U1,U2,U4,U5 and U6) and over 150 proteins.
The RNA-protein complex is known as snRNPs or "snurps".



Function of different RNA Molecules

5. Small regulatory RNAs: small interfering RNA **(siRNA)** and microRNA **(miRNA)** have a role in gene silencing and regulation of gene expression by base-pairing with complementary sequence of target mRNA molecules.





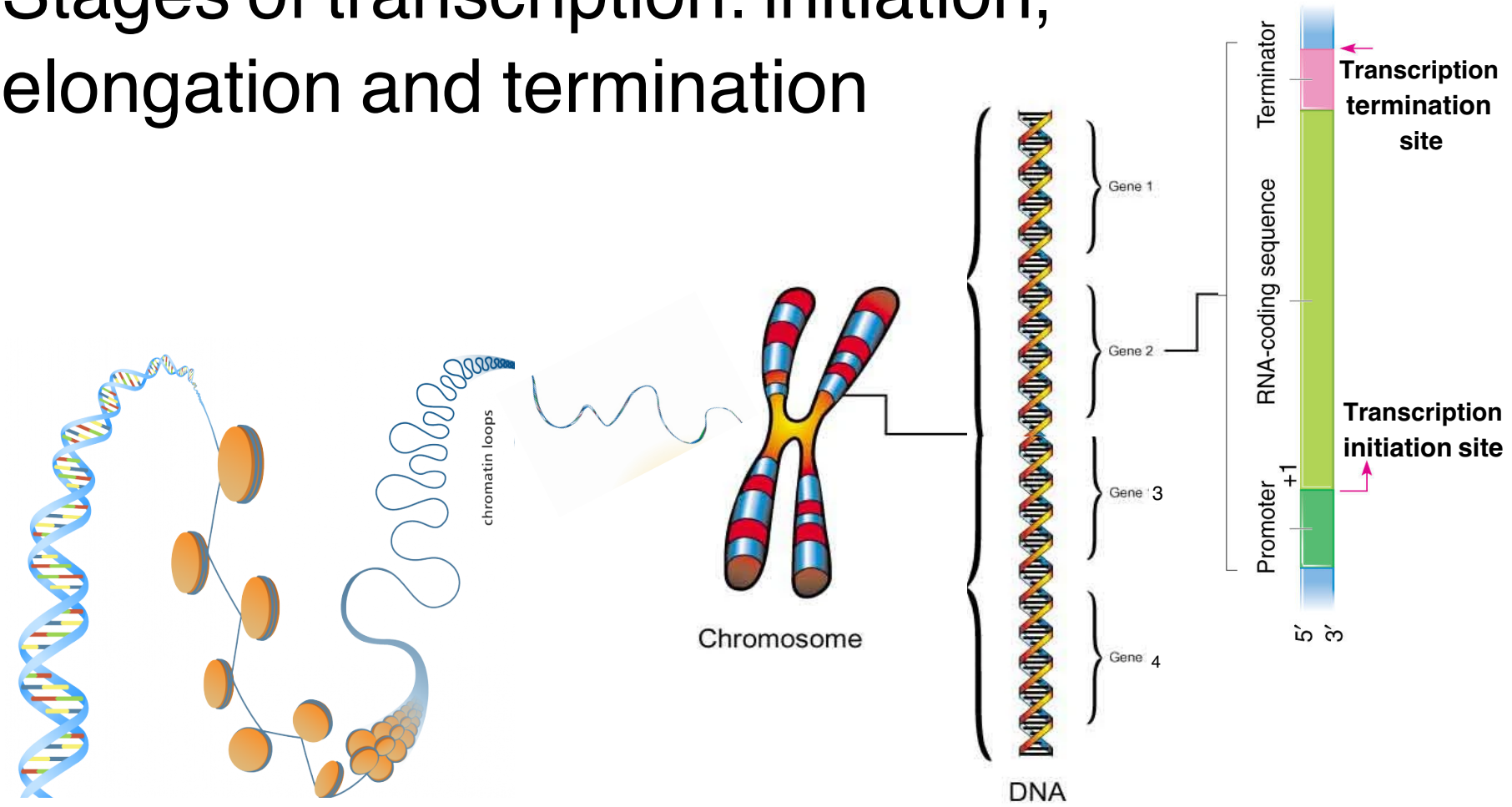
RNA Polymerase Enzyme

- **RNA polymerase:** is an enzyme which synthesizes RNA from DNA template
- Three classes of RNA polymerases in Eukaryotic cells:

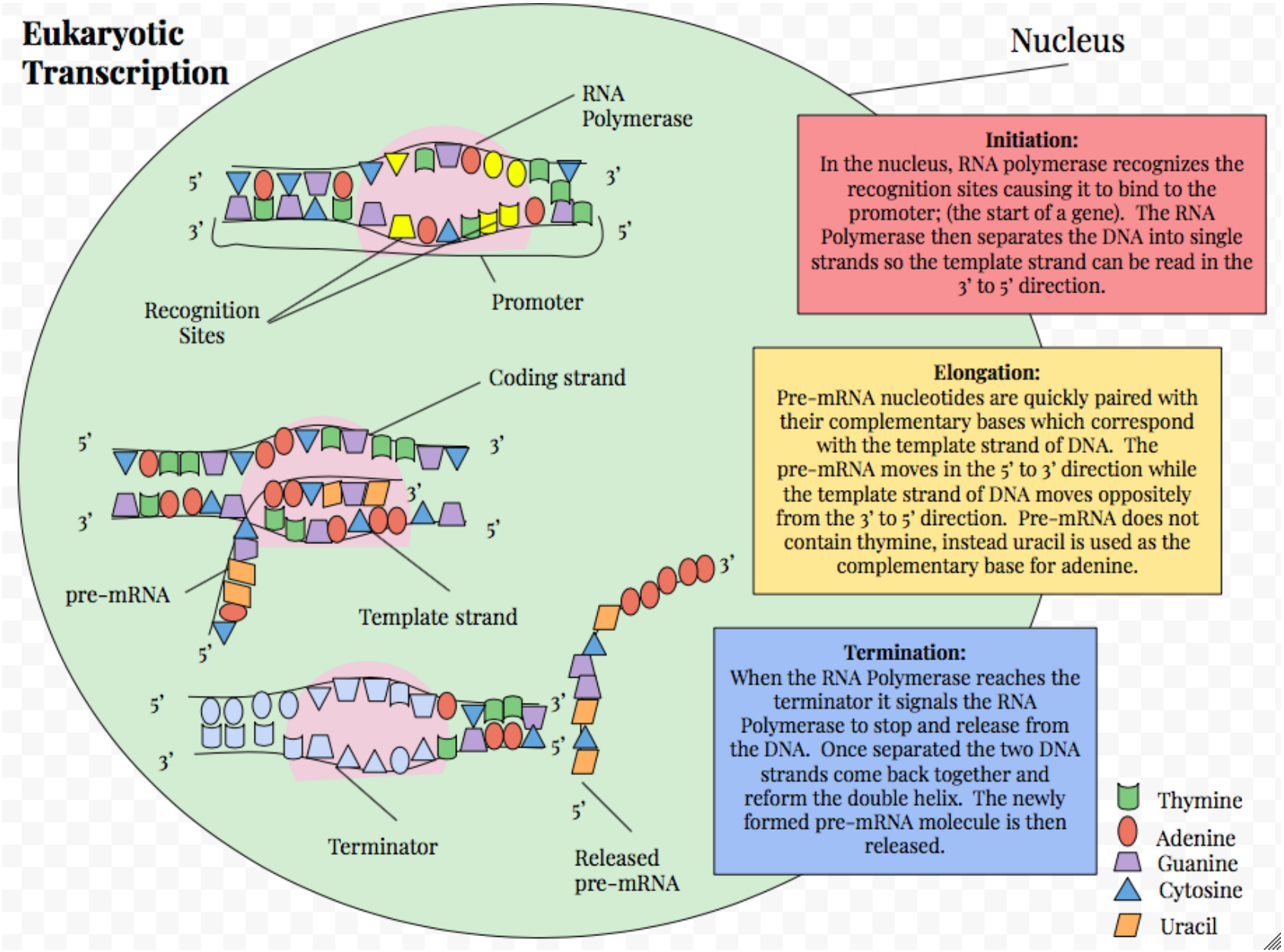
Form	Product
RNA Polymerase I	All rRNAs except 5S rRNA
RNA Polymerase II	All mRNAs, miRNAs, some snRNAs
RNA Polymerase III	All tRNAs, 5S rRNA, other small RNAs

Stages of Transcription

- Gene transcription in eukaryotes is more complex compared to prokaryotes
- Stages of transcription: initiation, elongation and termination

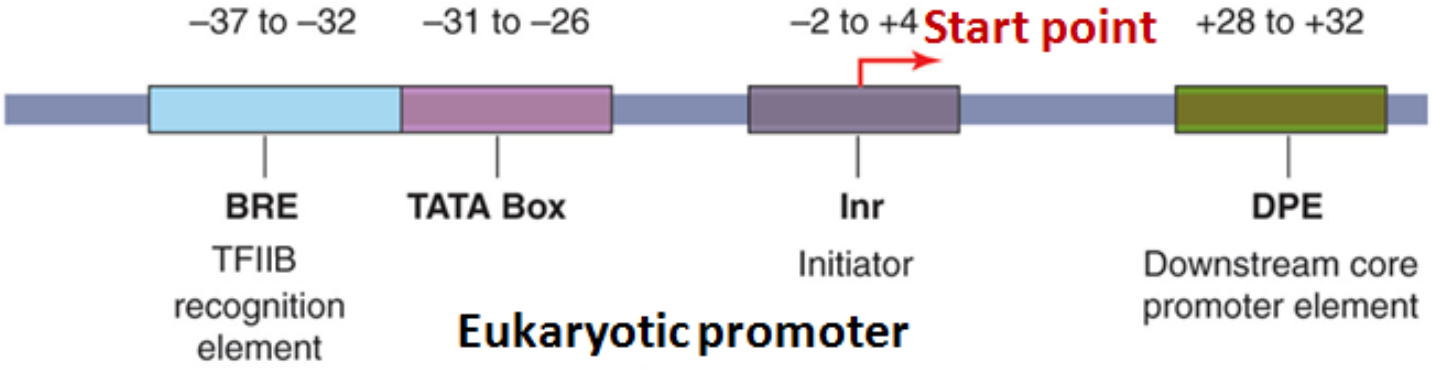


Stages of Transcription



Transcription Initiation

- **Promoter:** is a stretch of DNA sequence (non-coding DNA) where RNA polymerase can bind to start the transcription. Promoters are located upstream the genes that they regulate
- Promotors are found in Eukaryotes and Prokaryotes
- Eukaryotic promoter is a short DNA sequence (~100 bp) which consists of consensus sequences such as TATA box, BRE, INR and DPE



Transcription Initiation

- General transcription factors **(GTFs):** are DNA-binding proteins which recognise specific regions in promoter and correctly position the RNA polymerase II at the transcription start site (TSS)
- There are five types of GTFs: TFIIB, TFIID (TBP and TAFs subunits), TFIIE, TFIIF and TFIIH

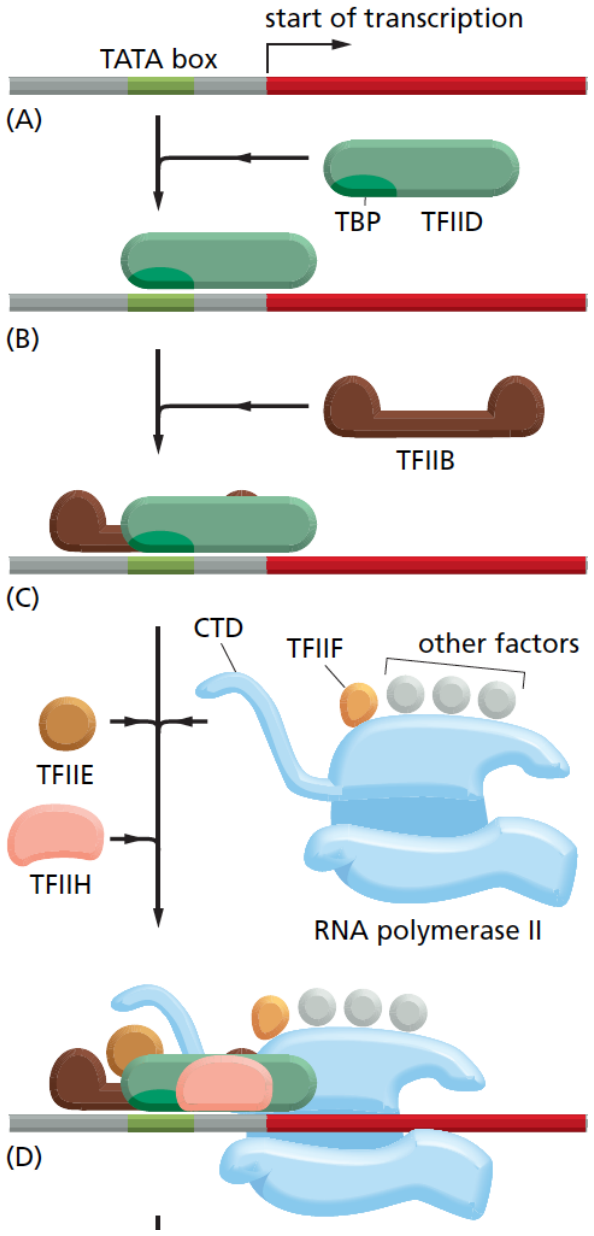
Table 6–3 The General Transcription Factors Needed for Transcription Initiation by Eucaryotic RNA Polymerase II

NAME	NUMBER OF SUBUNITS	ROLES IN TRANSITION INITIATION
TFIID		
TBP subunit	1	recognizes TATA box
TAF subunits	~11	recognizes other DNA sequences near the transcription start point; regulates DNA-binding by TBP
TFIIB	1	recognizes BRE element in promoters; accurately positions RNA polymerase at the start site of transcription
TFIIF	3	stabilizes RNA polymerase interaction with TBP and TFIIB; helps attract TFIIE and TFIIH
TFIIE	2	attracts and regulates TFIIH
TFIIH	9	unwinds DNA at the transcription start point, phosphorylates Ser5 of the RNA polymerase CTD; releases RNA polymerase from the promoter

TFIID is composed of TBP and ~11 additional subunits called TAFs (TBP-associated factors); CTD, C-terminal domain.

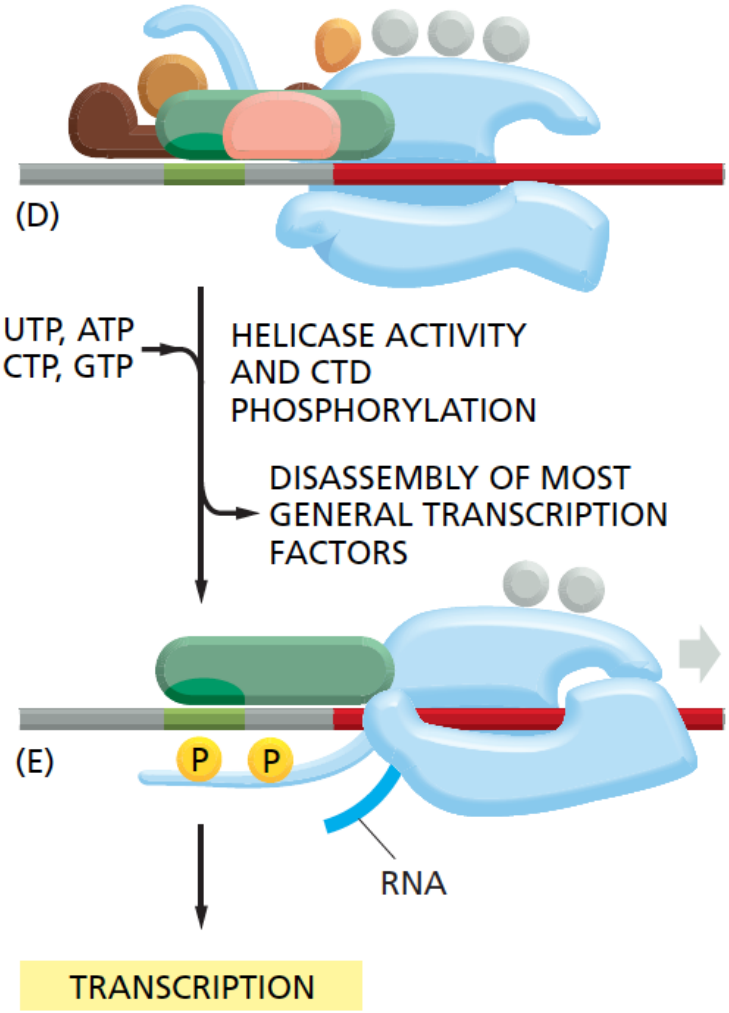
Transcription Initiation

- Pre-initiation complex **(PIC)** formation: is the assembly of GTFs along with RNA polymerase II at the promoter region. It is an important step in the transcription initiation process
- TFIID initially binds TATA box region via its TBP subunit. TFIIB subsequently binds BRE region of the promoter



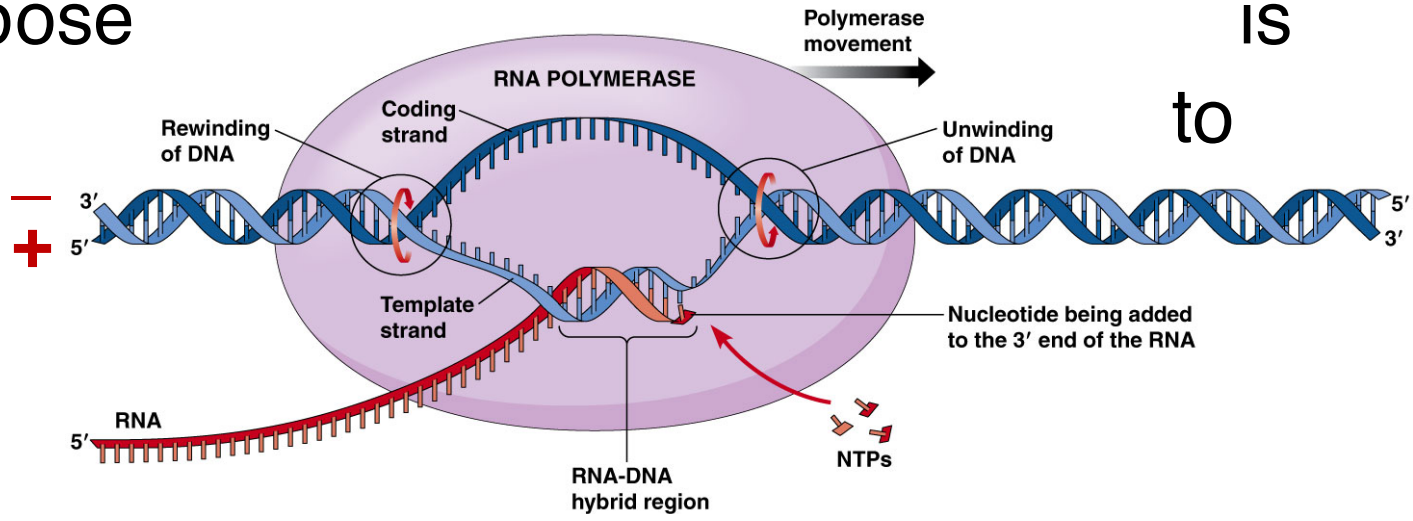
Transcription Initiation

- TFIIH can perform several enzymatic steps required for the transcription elongation.
  1. It unwinds (unzips) the double helix at the transcription start site (TSS) via its helicase activity
  2. It phosphorylates Ser5 located at the tail of RNA polymerase II called CTD (C-terminal domain) via its kinase activity. The CTD tail consists of 52 tandem repeats of 7 -amino acids sequence (Ser2, Ser5). This event is important for RNA polymerase II to escape the promoter and start elongation



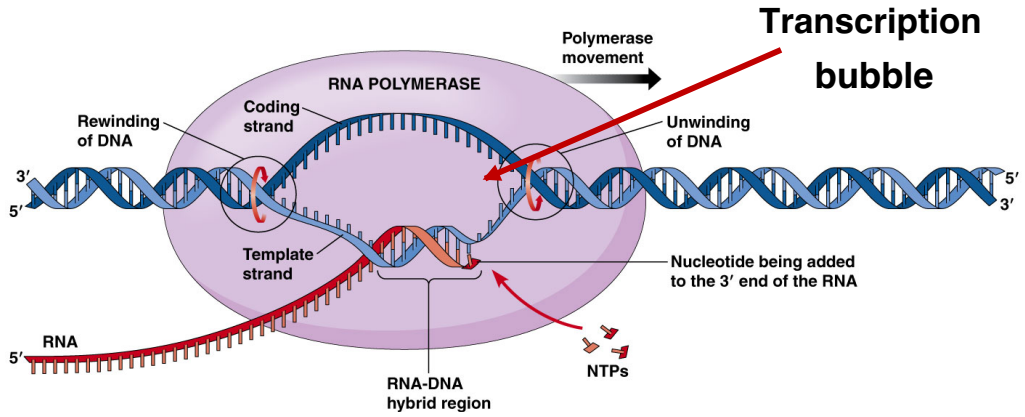
Transcription Elongation

- Elongation: the extension of the newly synthesized transcript
- The template strand (antisense strand, –): 3’- 5’ strand which is used to synthesise the RNA molecule. It is complementary to RNA or transcript strand
- The coding strand (sense strand, +):5’- 3’ strand which resembles the RNA strand except T is replaced with U and deoxyribose is changed to ribose



Transcription Elongation

- RNA polymerase II can only read 3’- 5’ DNA strand and elongate the growing RNA strand in the 5’- 3’ direction (add only to free 3’end)
- A polymerase II unwinds/ unzips the helical DNA and creates a bubble that exposes the single strands of DNA during the elongation process
- Transcription bubble: is a region (10- 20 nucleotides) where the two single strands of DNA are separated and exposed



Transcription Termination

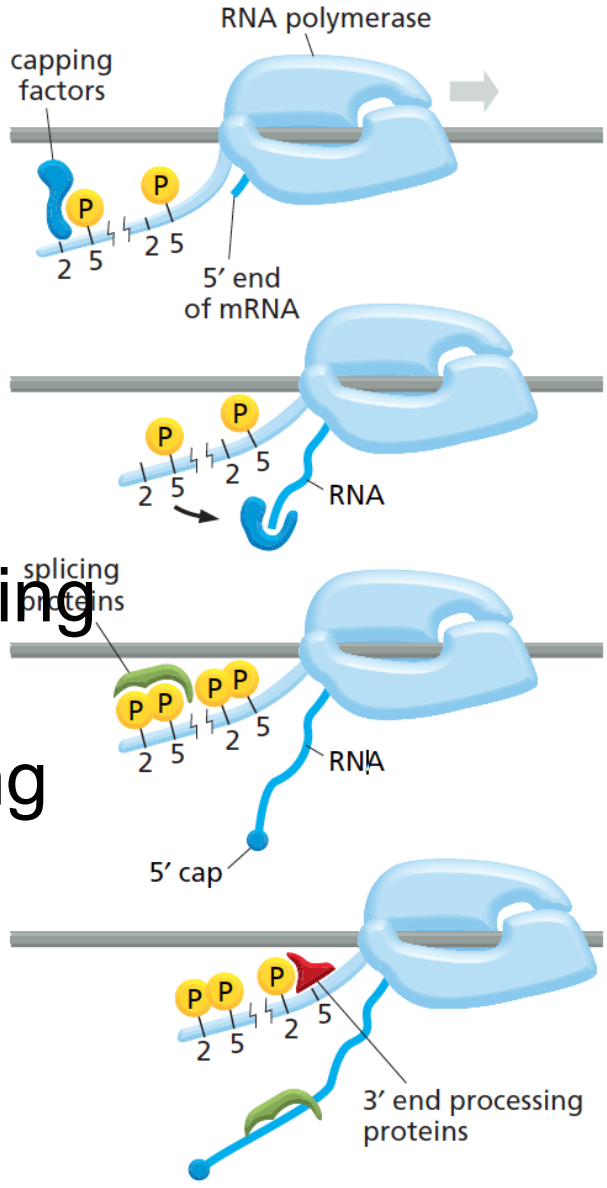
- Transcription termination: in eukaryotes is not well understood. RNA polymerase II meets a termination signal and knows that it is the end of transcription so will stop at that point.
- RNA polymerase II detaches from the template strand to initiate another round of transcription.
- To reinitiate transcription, soluble phosphatases remove the phosphates on CTD tail of RNA polymerase II
- The mature mRNA will leave to cytoplasm to start the second step in gene expression: **Translation**

Post Transcriptional Modification

- Post Transcriptional Modification **(PTM):** is a set of biological processes which modify the chemical structure of RNA and alter the pre- mRNA (primary transcript) to mature mRNA (functional RNA).
- **PTM** occurs only in eukaryotes
- **PTM** occurs in the nucleus. After modification, the functional RNA will leave to cytoplasm
- **PTM** consists of three steps:
  1. 5’ Capping
  2. RNA splicing
  3. 3’ Polyadenylation

Post Transcriptional Modification

- Transcription elongation in eukaryotes is tightly coupled to RNA processing steps
- Phosphorylation of CTD tail (Ser2,Ser5) of RNA polymerase II proceeds gradually during the elongation
- This attracts different RNA processing proteins to the tail (capping, factors, splicing proteins and 3'end processing proteins)
- These factors jump from the tail onto nascent RNA molecule and start processing at the appropriate time

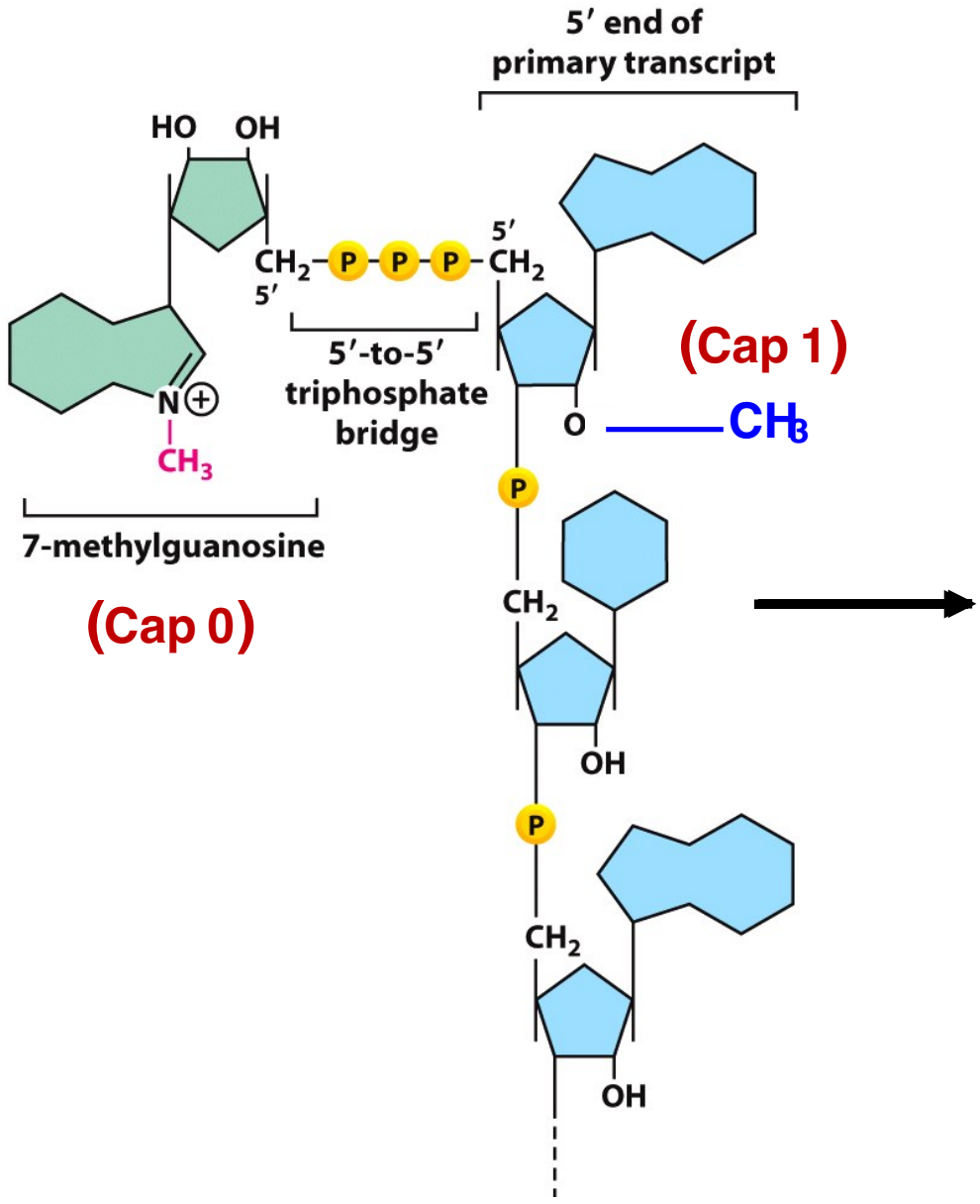


5’Capping

- RNA capping is the first modification of eukaryotic pre-mRNA
- Once the nascent RNA is 25 nucleotides long, a cap is added to the 5’ end
- The cap consists of methylated guanine nucleotide called 7-methylguanosine (m^7G)
- Capping factors present on the CTD tail:

1. RNA triphosphatase,
2. RNA guanylyltransferase
3. RNA methyltransferase

5’ Capping



γ β α

pppNpRNA

γ Pi — RNA Triphosphatase

ppNpRNA

GTP / PPi — RNA Guanylyltransferase

α

GpppNpRNA

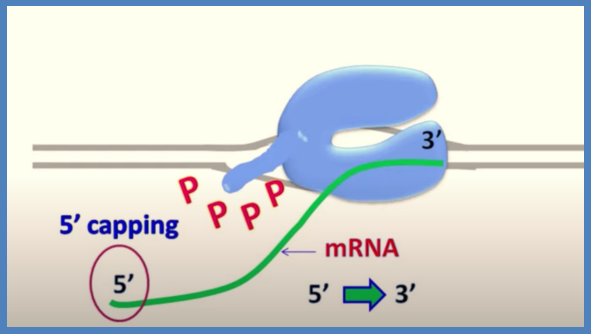
SAM / SAH — Methyltransferase (N7MTase)

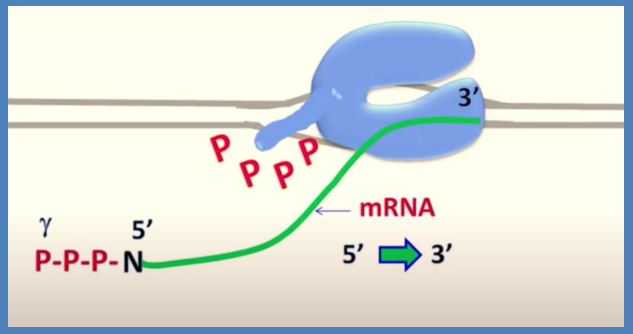
^{m7}GpppNpRNA (Cap-0 RNA)

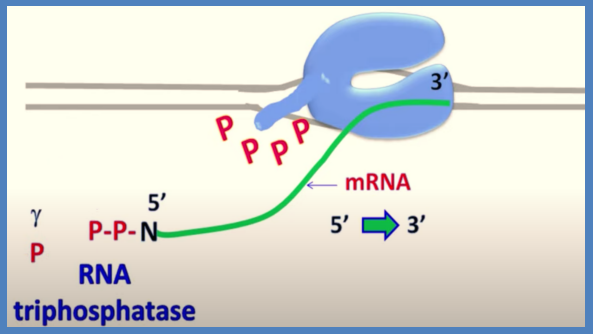
SAM / SAH — RNA (nuceloside-2’-O-)-Methyltransferase (2’OMTase)

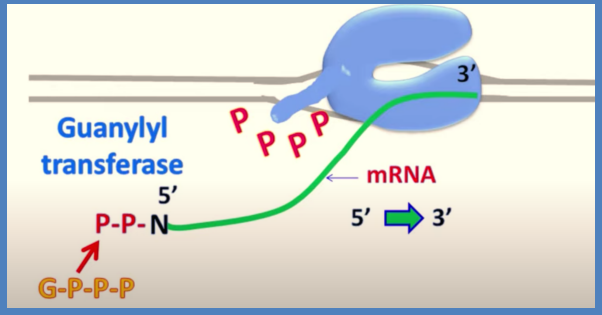
^{m7}GpppN$_m$pRNA (Cap-1 RNA)

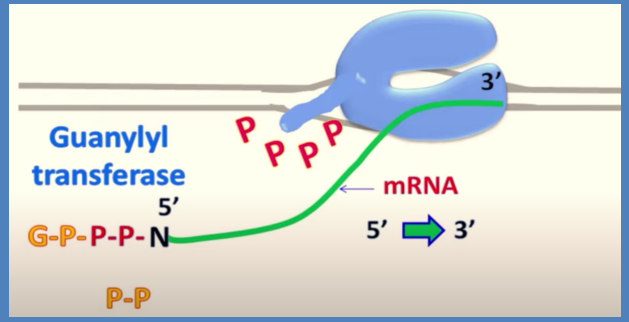
5’ Capping

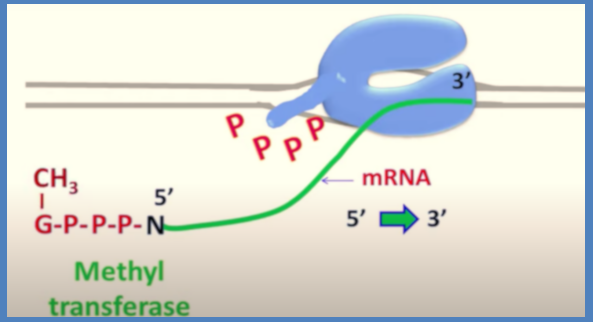

5’ capping
P P P P
3’
mRNA
5’
5’ 3’


P P P P
3’
mRNA
γ
5’
P-P-P-N
5’ 3’


P P P P
3’
mRNA
γ
P
5’
P-P-N
5’ 3’
RNA
triphosphatase


Guanylyl
transferase
P P P P
3’
mRNA
5’
P-P-N
5’ 3’
G-P-P-P


Guanylyl
transferase
P P P P
3’
mRNA
5’
G-P-P-P-N
5’ 3’
P-P


P P P P
3’
mRNA
CH3
5’
G-P-P-P-N
5’ 3’
Methyl
transferase

5’Capping

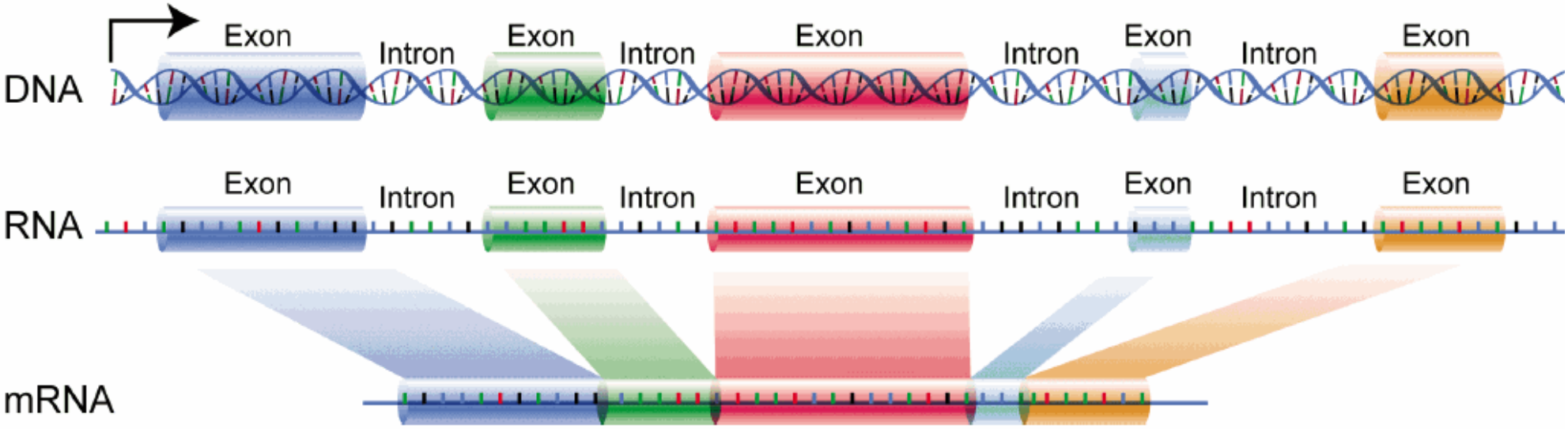
- The 5’-5’ linkage between the 7-methylguanosine and the rest of RNA molecule is unusual

- 5’ Cap has three main functions:

1. Protection from degradation by exonucleases
2. Important for efficient translation
3. Regulation of nuclear transport

RNA Splicing

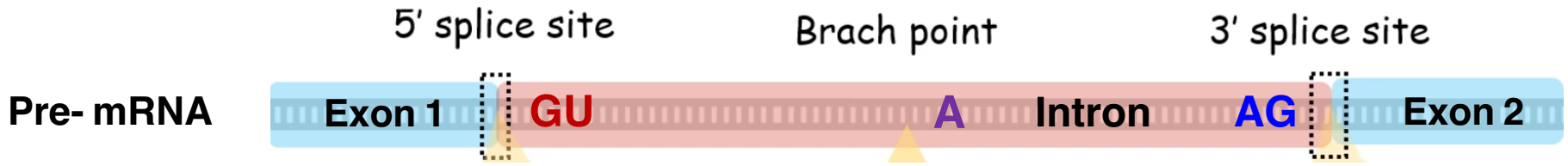
- Genes consists of protein coding sequences (Exons) and non-coding sequences (Introns)
- Both exons and introns are transcribed into RNA (precursor mRNA or pre-mRNA or primary transcript)
- Splicing: is the excision and removal of introns and the re-joining of exons to form the mature mRNA



RNA Splicing

- RNA splicing step occurs during the transcription elongation process and it is catalysed by a large RNA-protein molecular complex called spliceosome
- Three types of consensus and short nucleotide sequences (conserved in eukaryotes) which specify the introns and direct the splicing machinery:

1. 5'splice site **(5'GU)**
2. Branch point **(A)**
3. 3'splice site **(AG3')**



Spliceosome

Spliceosome

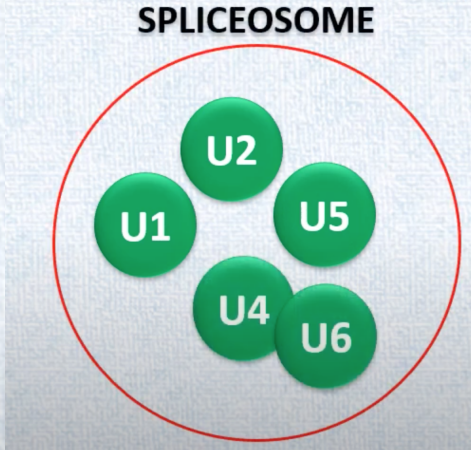
- A set of RNA-protein complexes

5
snRNA
(small nuclear RNA)

150
Protein

snRNP (snurps)

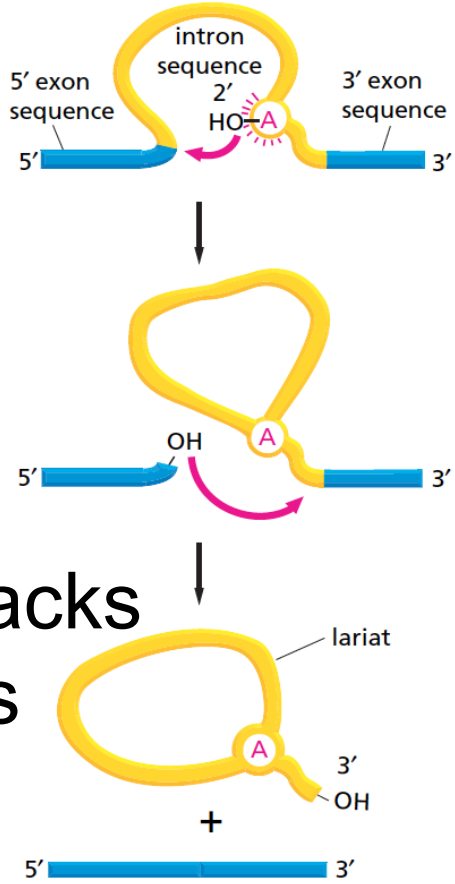
(small nuclear ribonucleoproteins)



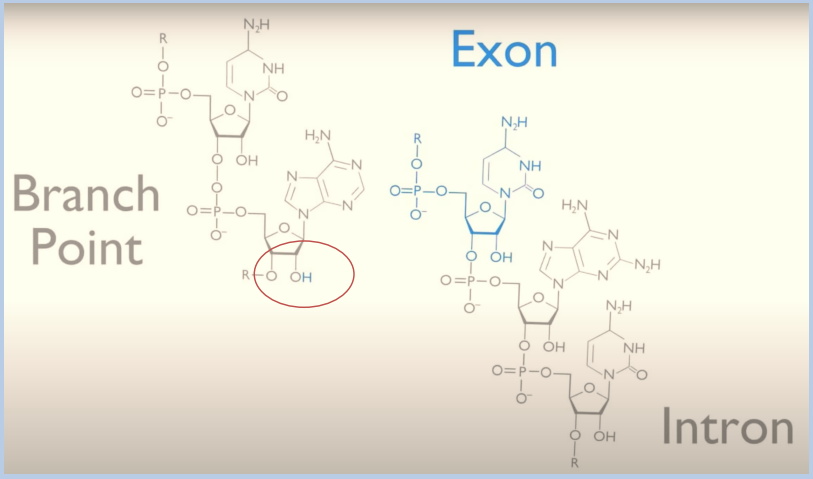
Mechanism of RNA Splicing

- Each splicing event involves two sequential transesterification reactions that join two adjacent exons while removing one intron as a lariat (a rope with a loop)

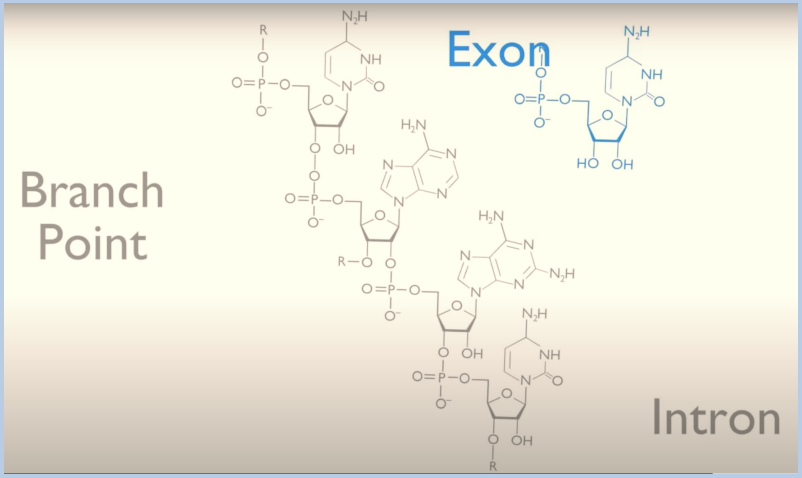
1. The 2'OH of **A** attacks the 5' splice site and break the phosphodiester bond in the backbone of RNA
2. The cut 5'end of the intron becomes covalently linked to **A** forming a loop
3. The OH at the free 3' end of exon 1 attacks the 3'splice site and joins the two exons while the intron is released as lariat

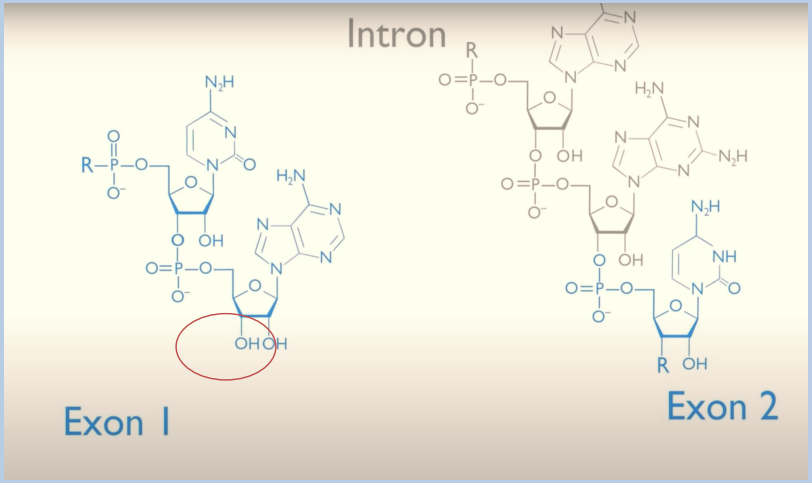


Transesterification

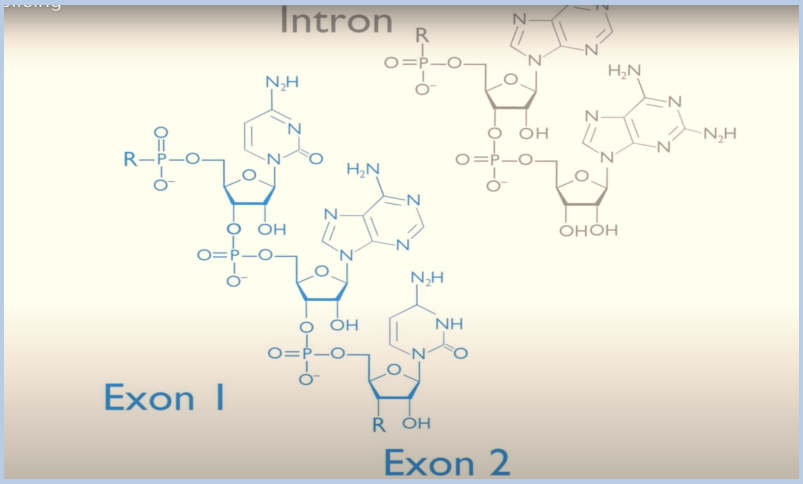


First reaction



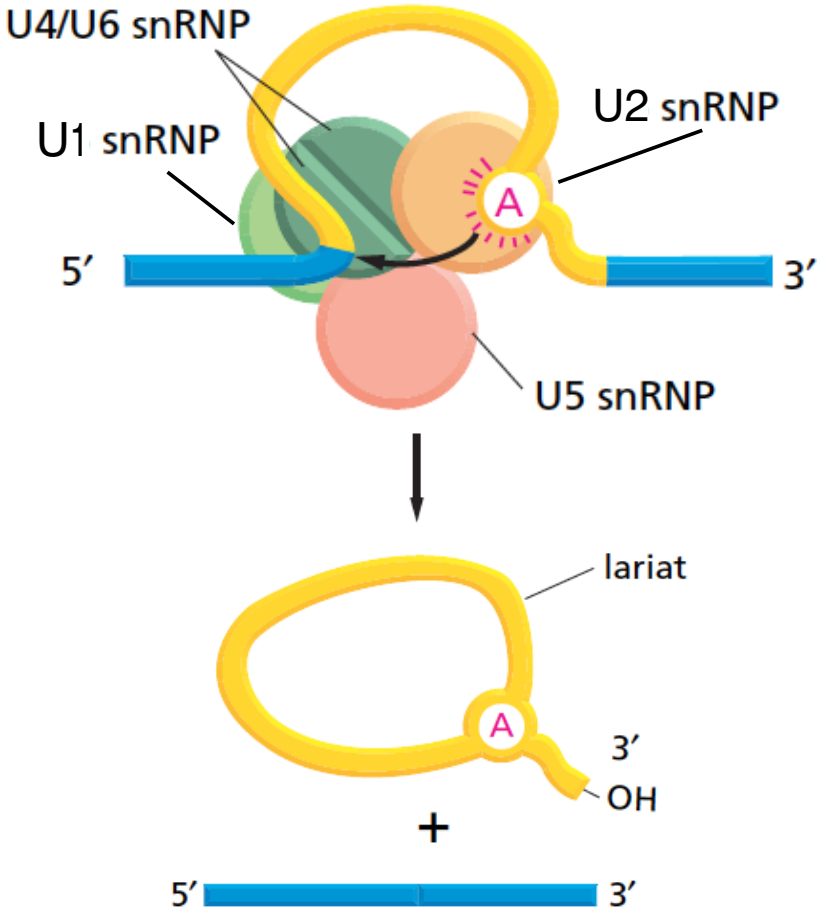


Second reaction



Spliceosome

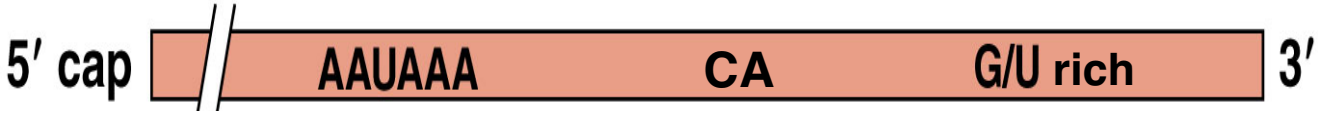
- There are five types of snRNAs (U1,U2,U4,U5 & U6) and over 150 proteins which assemble together at the pre-mRNA to form the functional spliceosome that catalyses the splicing event
- The complex that forms between each type of snRNA and proteins is called snRNPs pronounced as “snurps”



3’ Polyadenylation

- The final step in pre-mRNA is the addition of poly-A tail to the newly synthesized mRNA
- 3’end processing signal which specifies the 3’end of each mRNA is encoded in the genome and transcribed into pre-mRNA
- The signal consists of 3 sequence elements:

1. Polyadenylation sequence **(5’AAUAAA3’)**
2. The site of cleavage **(5’CA3’)**
3. GU rich sequence



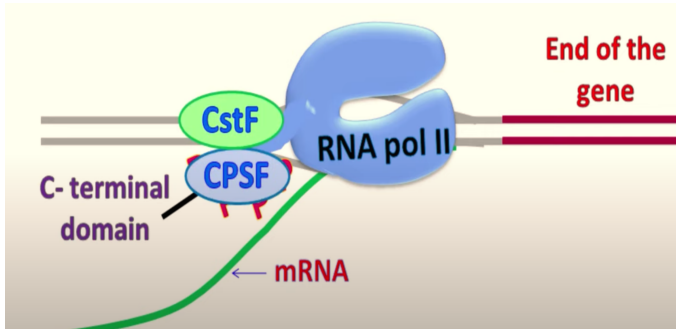
3’ Polyadenylation

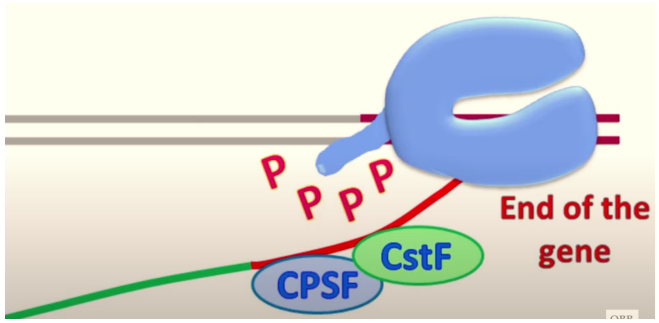
- There are 3 types of 3’end processing proteins or factors involved in the 3’end processing step:

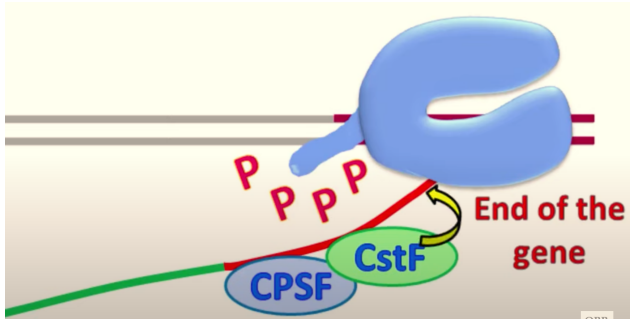
1. Cleavage and polyadenylation specificity factor **(CPSF)**
2. Cleavage stimulation factor **(CStF)**
3. Poly-A polymerase **(PAP)**

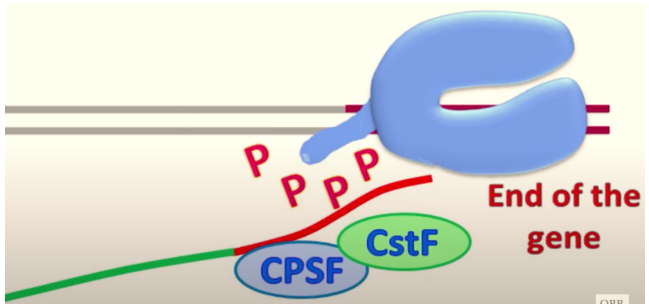
- CPSF and CStF are multi-subunit RNA binding proteins which can jump from the CTD phosphorylated tail of RNA polymerase II and bind the specific sequences on pre-mRNA as it emerges from RNA polymerase II

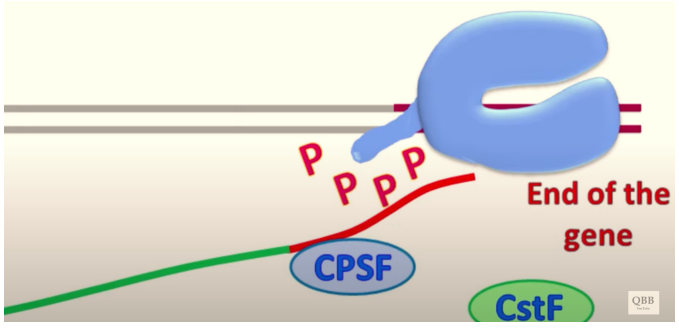
3’ Polyadenylation

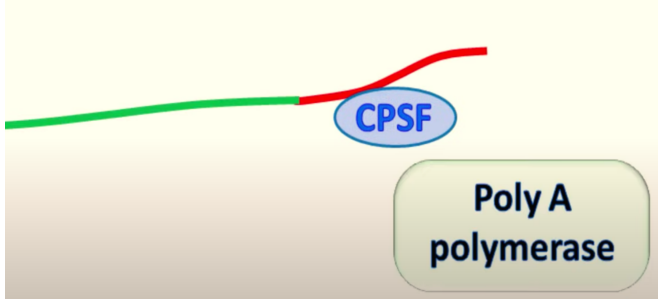


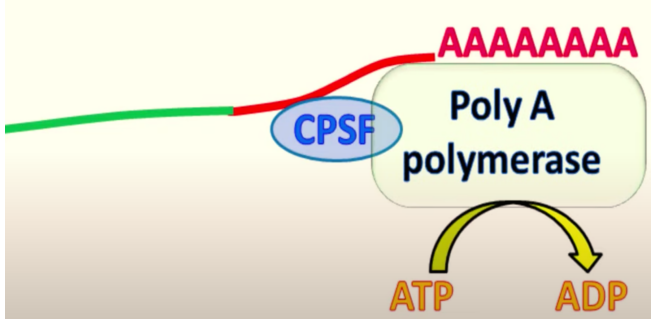


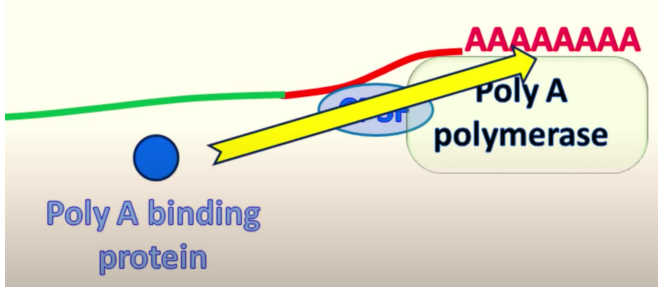


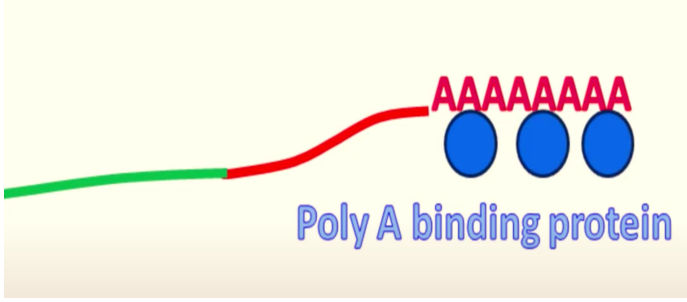










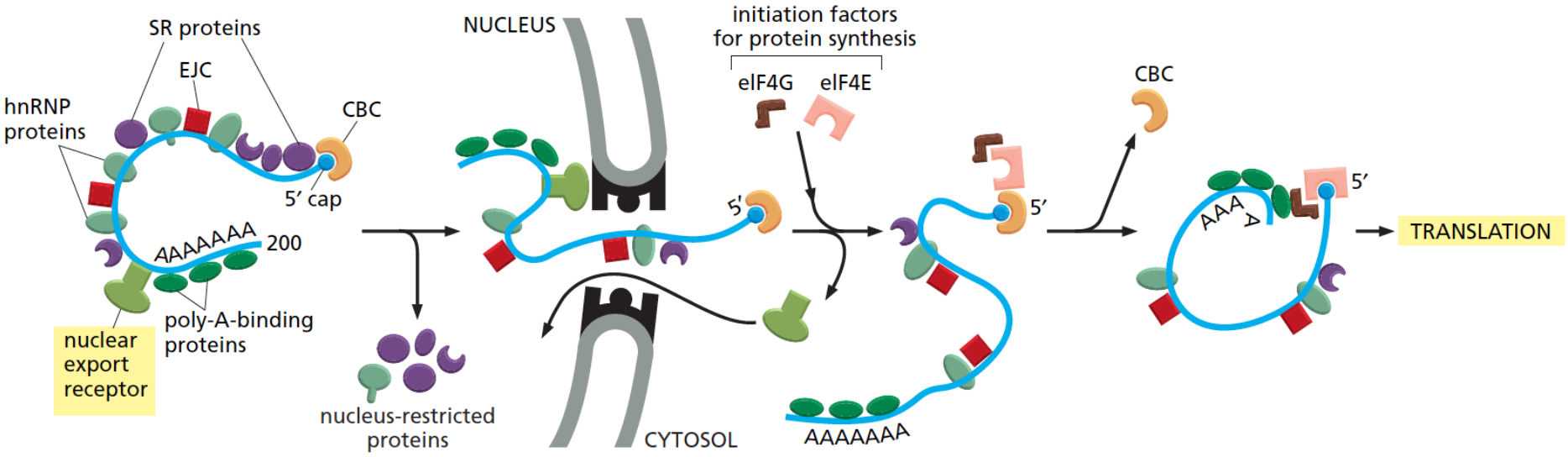


3’ Polyadenylation

- Pre- mRNA is cleaved by CStF at the cleavage site
- PAP is recruited and starts addition of about 200 A nucleotides to 3’ end using ATP as precursor
- Poly- A binding proteins are recruited to **determine the final length of the tail**
- CPSF and PAP dissociate from RNA and some poly - A binding proteins remain bound to protect 3’end from degradation by exonucleases until it travels from the nucleus and pass the check test before the beginning of translation

Transportation of mature mRNA

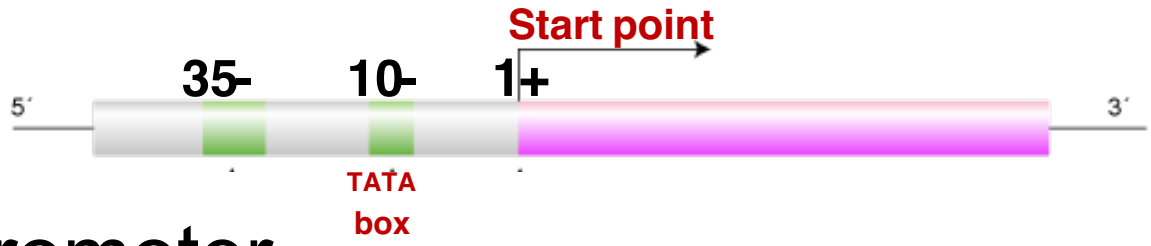
- Only fully processed and mature mRNA can pass through the nuclear pore and travel to the cytosol for translation



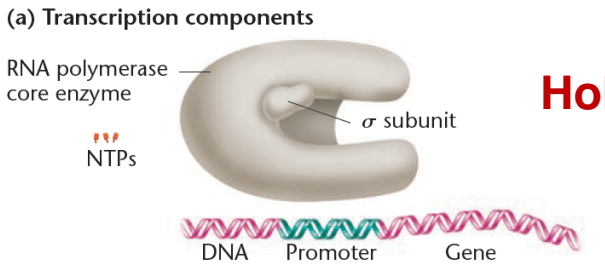
- The final check is performed in the cytosol by initiation factors eIF4E and eIF4G to ensure efficient translation

Transcription in Prokaryotes

- In prokaryotes, two consensus sequences at
  -10 (TATA) and
  -35 located in the promoter



- **σ factor** recognizes the -35 region in the promoter and binds to it
- Once the RNA polymerase starts the transcription, σ factor then dissociates to guide another enzyme to the initiation site.



Holoenzyme