TRANSCRIPTION

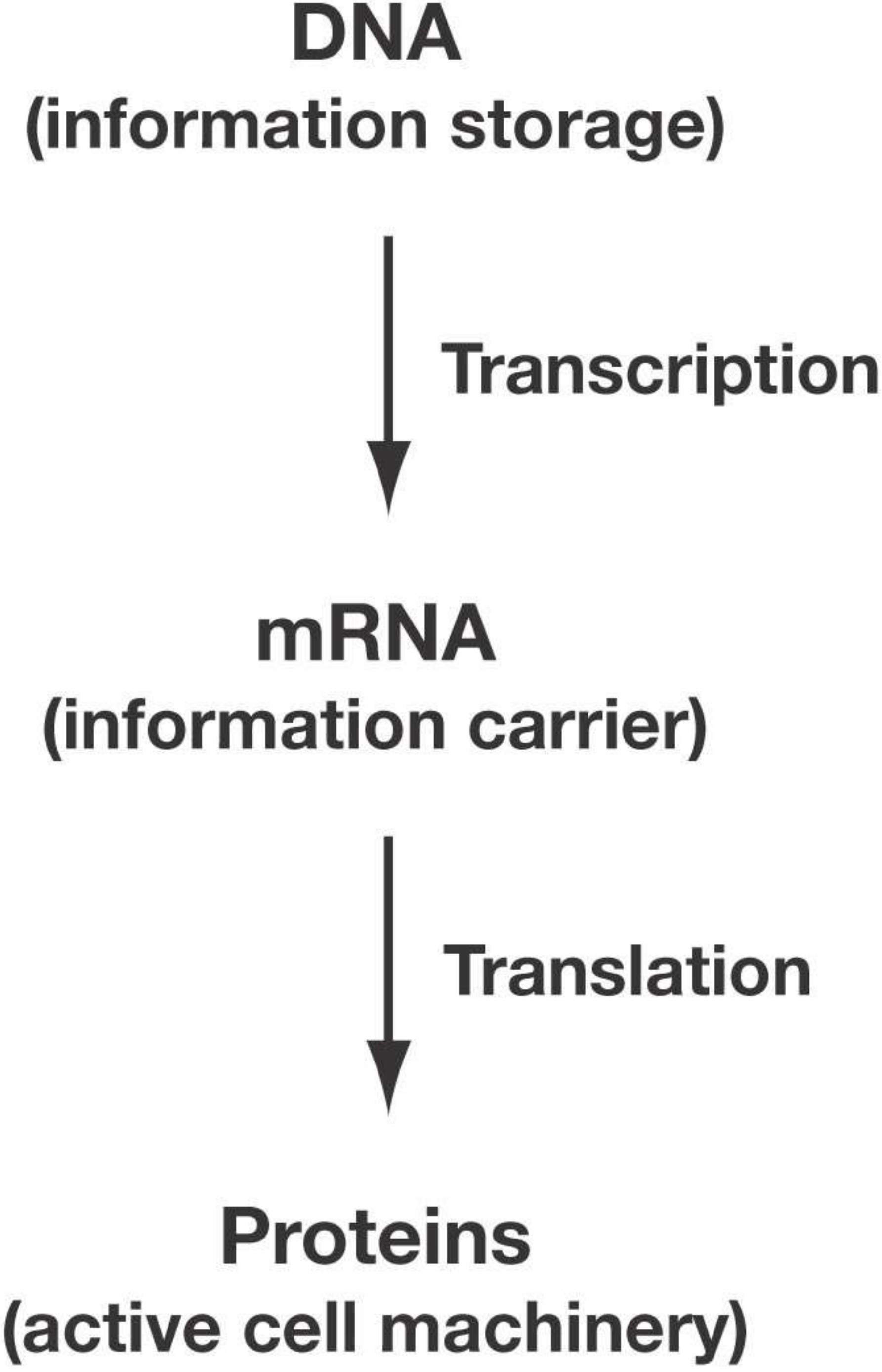
Transcription

- **Prokaryotic transcription**
- The RNA polymerase
- The origin & prokaryotic promoters
- The initiation, elongation,& termination.
- Prokaryotic termination signals
- Prokaryotic transcription product
- **Eukaryotic transcription**
- Eukaryotic RNA polymerases
- Eukaryotic promoters

- Enhancers & transcriptional elements
- The processes of initiation, elongation &
- termination
- Eukaryotic termination signals
- RNA processing & modifications
    - The spliceosomes
    - Thalassemias & globin mRNA splicing
    - Modification to mRNA, tRNA

The Central Dogma

- DNA codes for RNA
- RNA codes for protein



DNA
(information storage)

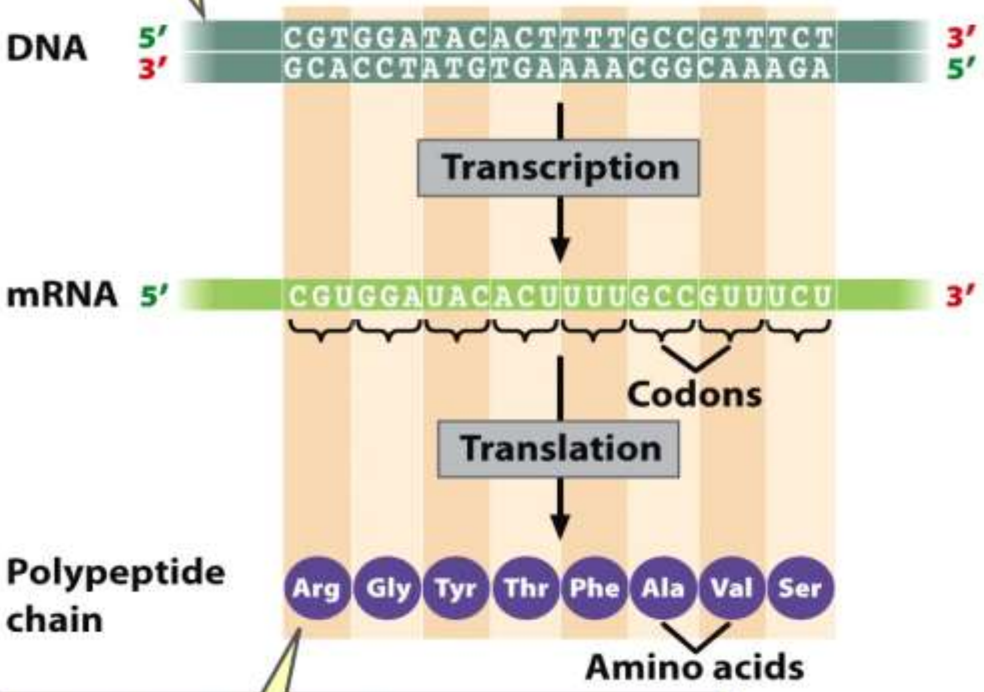
Transcription

mRNA
(information carrier)

Translation

Proteins
(active cell machinery)

1 A continuous sequence of nucleotides in the DNA…

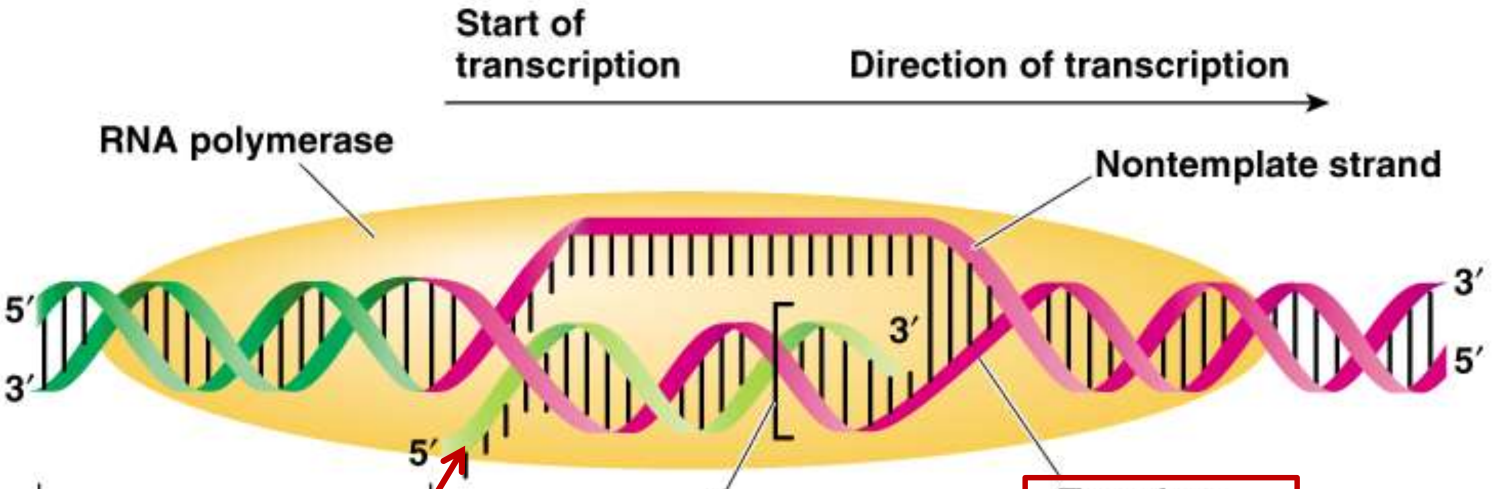


2 …codes for a continuous sequence of amino acids in the protein.

Conclusion: With colinearity, the number of nucleotides in the gene is proportional to the number of amino acids in the protein.

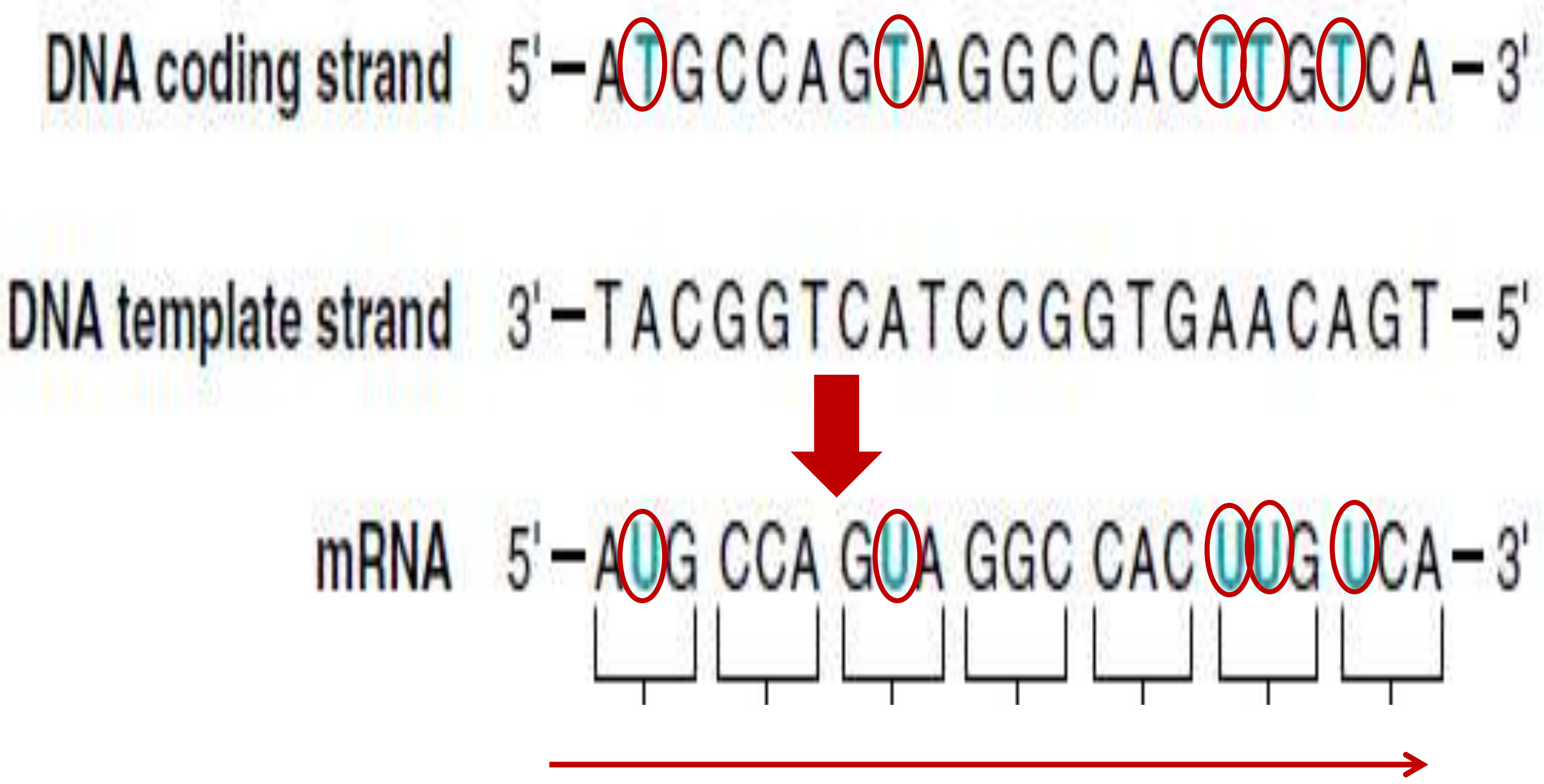
Figure 14.1
Genetics: A Conceptual Approach, Fourth Edition

Transcription: synthesis of one RNA molecule using one of the two DNA strands as a template by the enzyme RNA Polymerase.



Template DNA strand

Newly Synthesized RNA

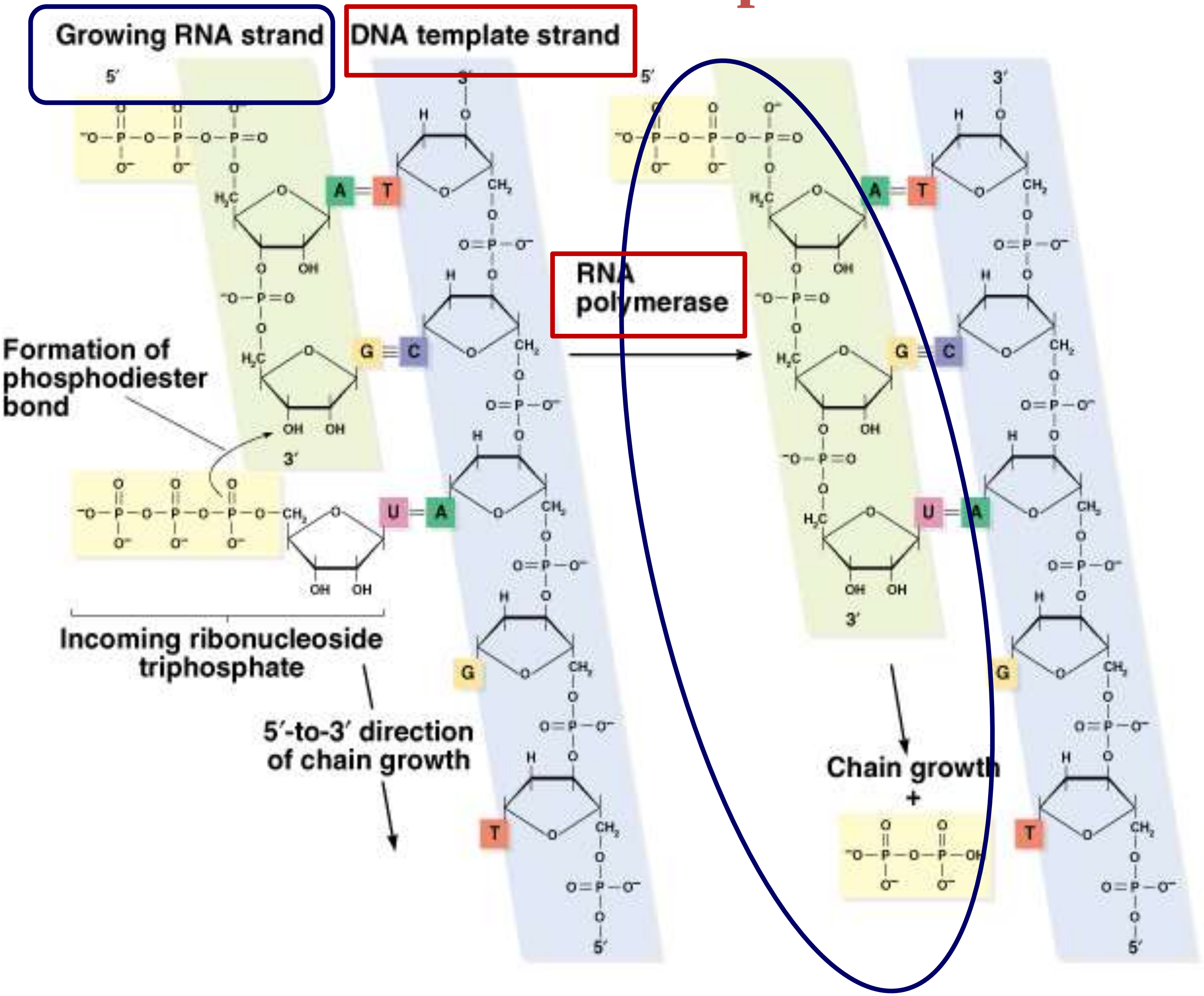


DNA coding strand 5'–ATGCCAGTAGGCCACTTGTCA–3'

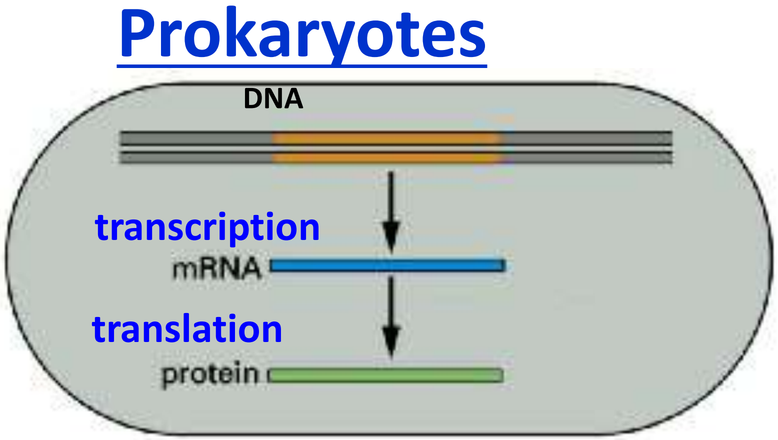
DNA template strand 3'–TACGGTCATCCGGTGAACAGT–5'

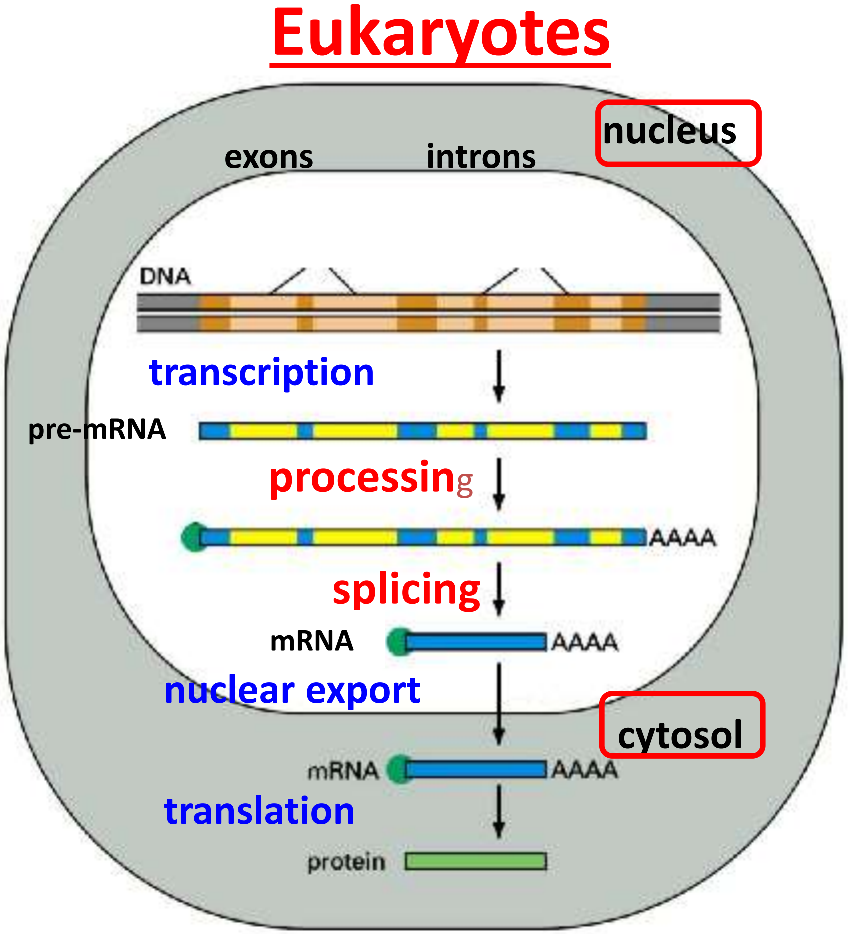
mRNA 5'–AUG CCA GUA GGC CAC UUG UCA–3'

The RNA polymerase-catalyzed synthesis of RNA on a DNA template strand



Eukaryotic Transcription and Tanslation are separated by space and time





RNA Polymerase

- The enzyme responsible for the RNA synthesis is **DNA-dependent RNA polymerase.**
    - The prokaryotic RNA polymerase is a multiple-subunit protein of ~480kD.

RNA Polymerase

$$(NMP)n + NTP \rightarrow (NMP)n+1 + PPi$$

1. Requires no primer for polymerization.
2. Requires DNA for activity and is most active with a double-stranded DNA as template.
3. 5’ → 3’ synthesis.
4. Require Mg^{2+} for RNA synthesis activity.
5. lacks 3’ → 5’ exonuclease activity, and the error rate of nucleotides incorporation is 10^{-4} to 10^{-5}.
6. Usually are multisubunit enzyme.

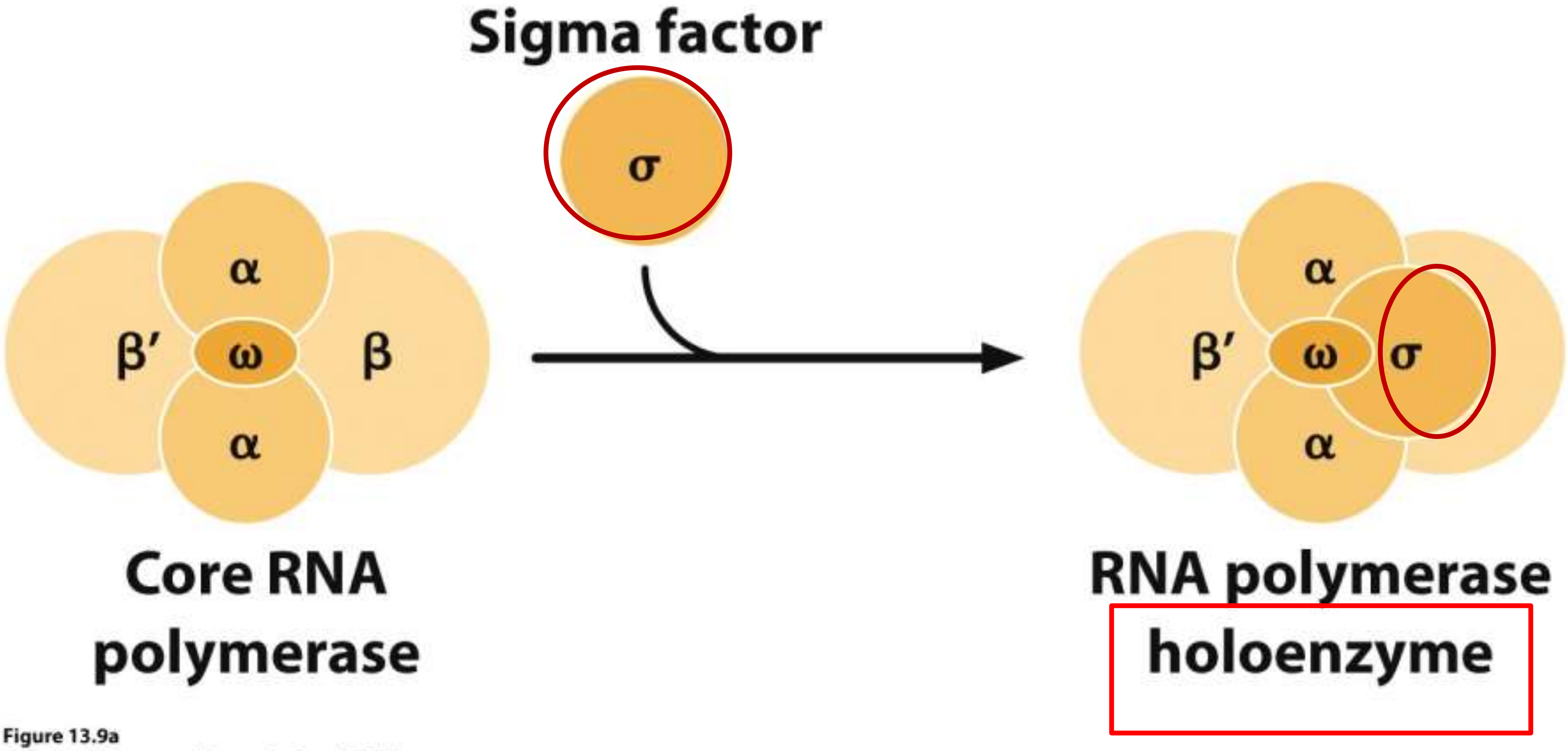


Figure 13.9a
Genetics: A Conceptual Approach, Fourth Edition
© 2012 W. H. Freeman and Company

RNA Polymerase of *E. Coli*

The holoenzyme of RNA-polymerase in *E.coli*

consists of 5 different subunits: $\alpha_2\ \beta\ \beta'\ \omega\ \sigma$

Subunit	MW	Function
α	36.5 KD	Determines the DNA to be transcribed
β	150 KD	Catalyzes polymerization
β'	155.5 KD	Binds & open DNA template
σ	70 KD	Recognizes the promoter for synthesis initiation
ω	11 KD	Subunit packing

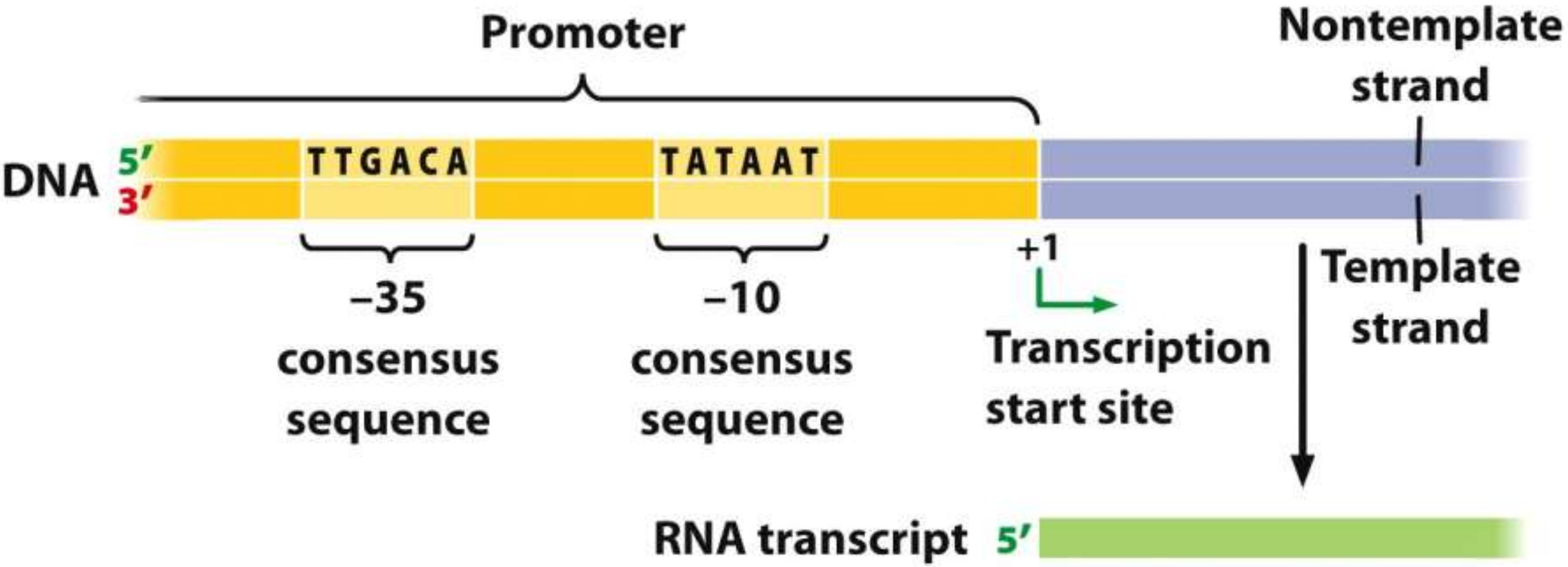
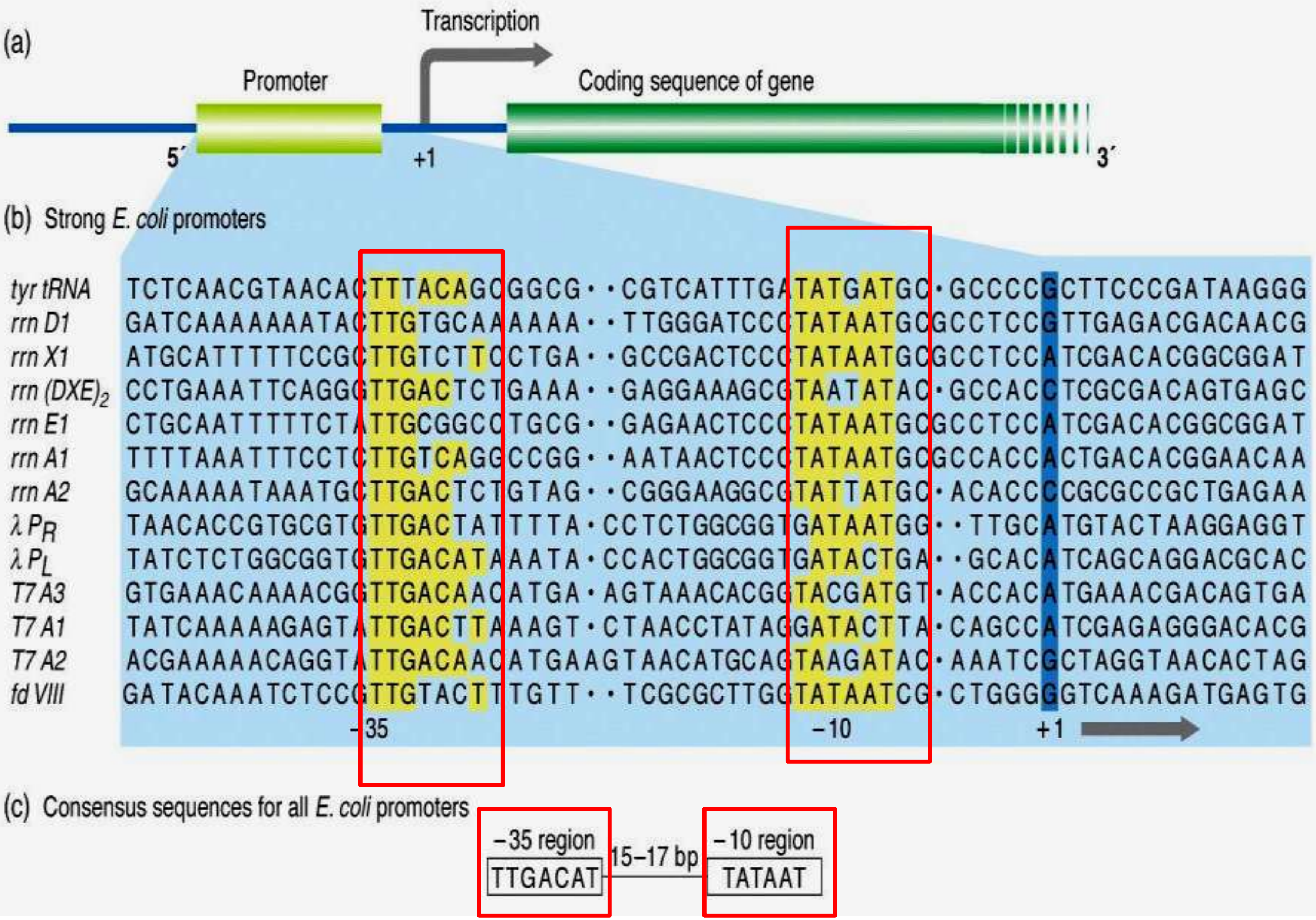


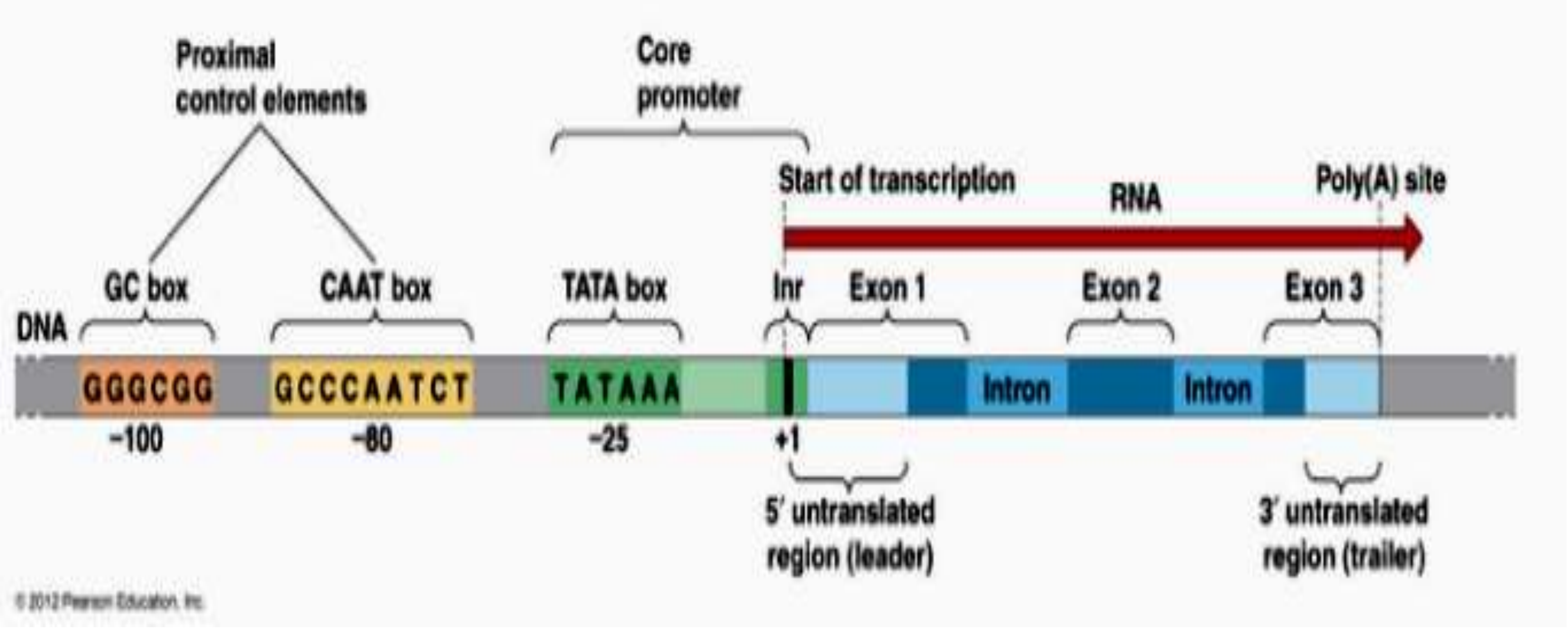
Figure 13.11

Genetics: A Conceptual Approach, Fourth Edition


(a)
Transcription
Promoter
Coding sequence of gene
5′
+1
3′
(b) Strong E. coli promoters
tyr tRNA TCTCAACGTAACACTTTACAGCGGCG··CGTCATTTGATATGATGC·GCCCCGCTTCCCGATAAGGG
rrn D1 GATCAAAAAAATACTTGTGCAAAAAA··TTGGGATCCCTATAATGCGCCTCCGTTGAGACGACAACG
rrn X1 ATGCATTTTTCCGCTTGTCTTCCTGA··GCCGACTCCCTATAATGCGCCTCCATCGACACGGCGGAT
rrn (DXE)2 CCTGAAATTCAGGGTTGACTCTGAAA··GAGGAAAGCGTAATATAC·GCCACCTCGCGACAGTGAGC
rrn E1 CTGCAATTTTTCTATTGCGGCCTGCG··GAGAACTCCCTATAATGCGCCTCCATCGACACGGCGGAT
rrn A1 TTTTAAATTTCCTCTTGTCAGGCCGG··AATAACTCCCTATAATGCGCCACCACTGACACGGAACAA
rrn A2 GCAAAAATAAATGCTTGACTCTGTAG··CGGGAAGGCGTATTATGC·ACACCCCGCGCCGCTGAGAA
λ PR TAACACCGTGCGTGTTGACTATTTTA·CCTCTGGCGGTGATAATGG··TTGCATGTACTAAGGAGGT
λ PL TATCTCTGGCGGTGTTGACATAAATA·CCACTGGCGGTGATACTGA··GCACATCAGCAGGACGCAC
T7 A3 GTGAAACAAAACGGTTGACAACATGA·AGTAAACACGGTACGATGT·ACCACATGAAACGACAGTGA
T7 A1 TATCAAAAAGAGTATTGACTTAAAGT·CTAACCTATAGGATACTTA·CAGCCATCGAGAGGGACACG
T7 A2 ACGAAAAACAGGTATTGACAACATGAAGTAACATGCAGTAAGATAC·AAATCGCTAGGTAACACTAG
fd VIII GATACAAATCTCCGTTGTACTTTGTT··TCGCGCTTGGTATAATCG·CTGGGGGTCAAAGATGAGTG
−35
−10
+1
(c) Consensus sequences for all E. coli promoters
−35 region
TTGACAT
15–17 bp
−10 region
TATAAT

Eukaryotic Promoter Sequences

- Promoter
- Enhancers
- Activators

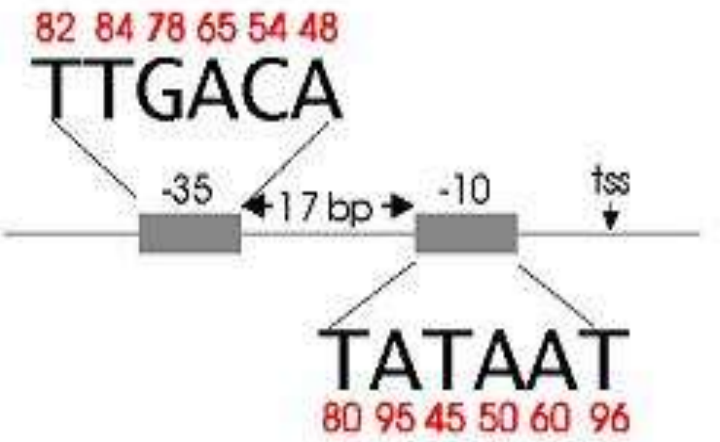


© 2012 Pearson Education, Inc.

Structure of bacterial prokaryotic promoter region

TATA-Box / Pribnow box

- This is a stretch of 6 nucleotides (5'- TATAAT-3') centered about 8-10 nucleotides to the left of the transcription start site.



-35 Sequence

- A second consensus nucleotide sequence (5'- TTGACA-3'), is centered about 35 bases to the left of the transcription start site.

Transcription process

1. Promoter binding
2. DNA unwinding
3. RNA chain initiation
4. RNA chain elongation
5. RNA chain termination

Transcription of Prokaryotes

- **Initiation phase: RNA-polymerase recognizes the promoter and starts the transcription.**
- **Elongation phase: the RNA strand is continuously growing.**
- **Termination phase: RNA-polymerase stops synthesis and the nascent RNA is separated from the DNA template.**

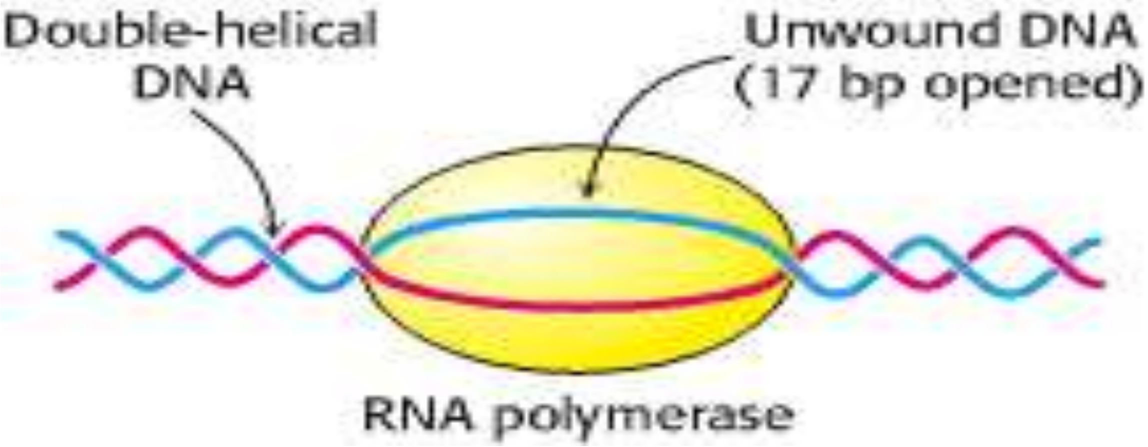
Initiation of Transcription at Promoters

Transcription is divided into three steps for both prokaryotes and eukaryotes. Initiation, Elongation and Termination.

The process of elongation is highly conserved between prokaryotes and eukaryotes, but initiation and termination are somewhat different.

Initiation

- **RNA-polymerse recognizes the *TTGACA* region (-35 sequence), and slides to the *TATAAT* region (-10 sequence), then opens the DNA duplex.**
- **The unwound region is about 17 bp.**



- The first nucleotide on RNA transcript is always purine triphosphate. GTP is more often than ATP.
- The pppGpN-OH structure remains on the RNA transcript until the RNA synthesis is completed.
- The three molecules form a transcription initiation complex.

RNA-pol ($\alpha_2\beta\beta'\sigma$) - DNA - pppGpN- OH 3′

- **No primer is needed for RNA synthesis.**

- **The σ subunit falls off from the RNA-polymerase once the first 3′,5′-phosphodiester bond is formed.**

- **The core enzyme moves along the DNA template to enter the elongation phase.**

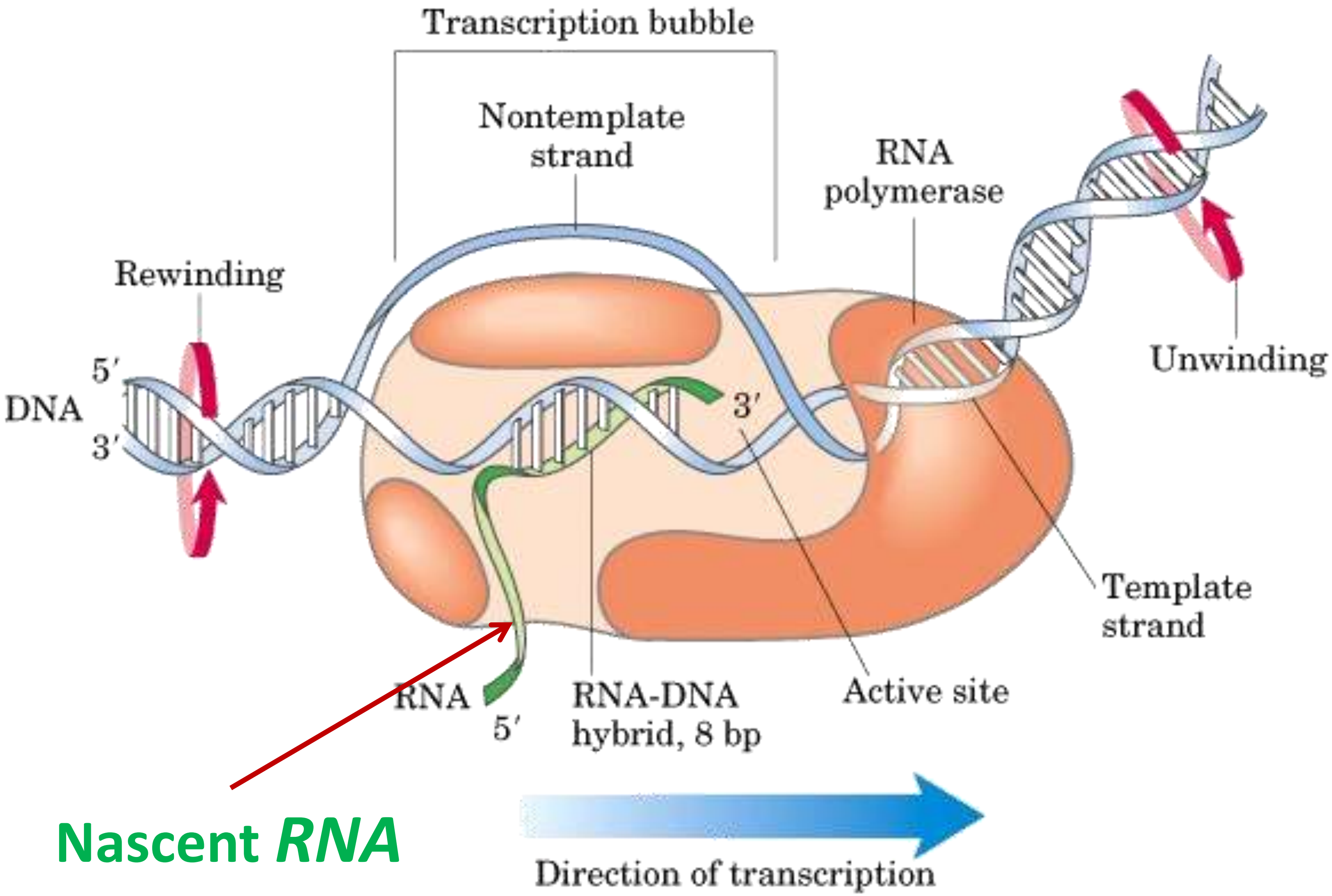
Elongation

- The release of the σ subunit causes the conformational change of the core enzyme. The core enzyme slides on the DNA template toward the 3′ end.
- Free NTPs are added sequentially to the 3′ -OH of the nascent RNA strand.

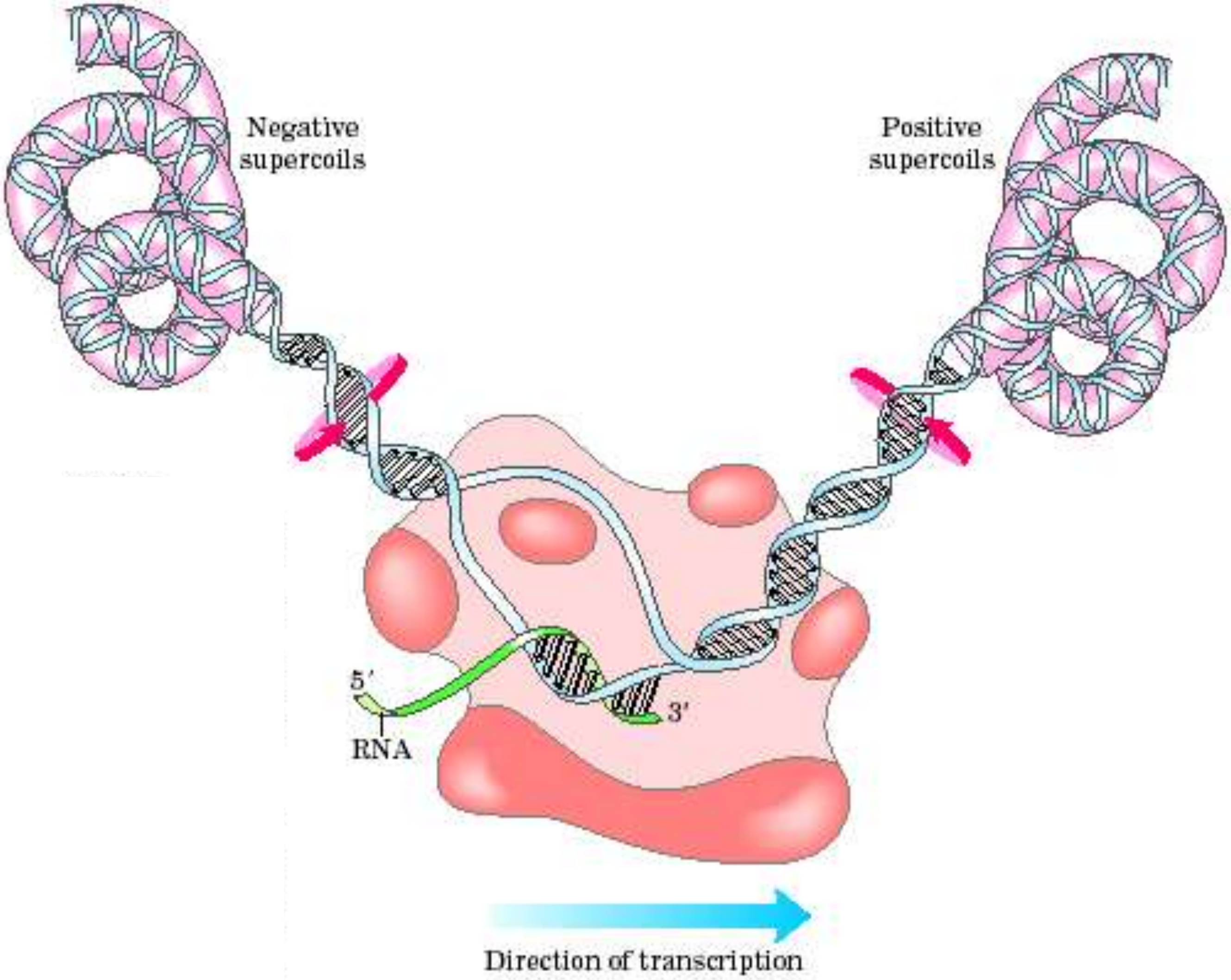
$$\underset{\text{RNA strand}}{(\mathrm{NMP})_n} + \underset{\text{substrate}}{\mathrm{NTP}} \longrightarrow \underset{\substack{\text{elongated}\\ \text{RNA strand}}}{(\mathrm{NMP})_{n+1}} + \mathrm{PPi}$$

- **RNA-polymerase, DNA segment of ~40nt and the nascent RNA form a complex called the transcription bubble.**
- **The 3′ segment of the nascent RNA hybridizes with the DNA template, and its 5′ end extends out the transcription bubble as the synthesis is processing.**

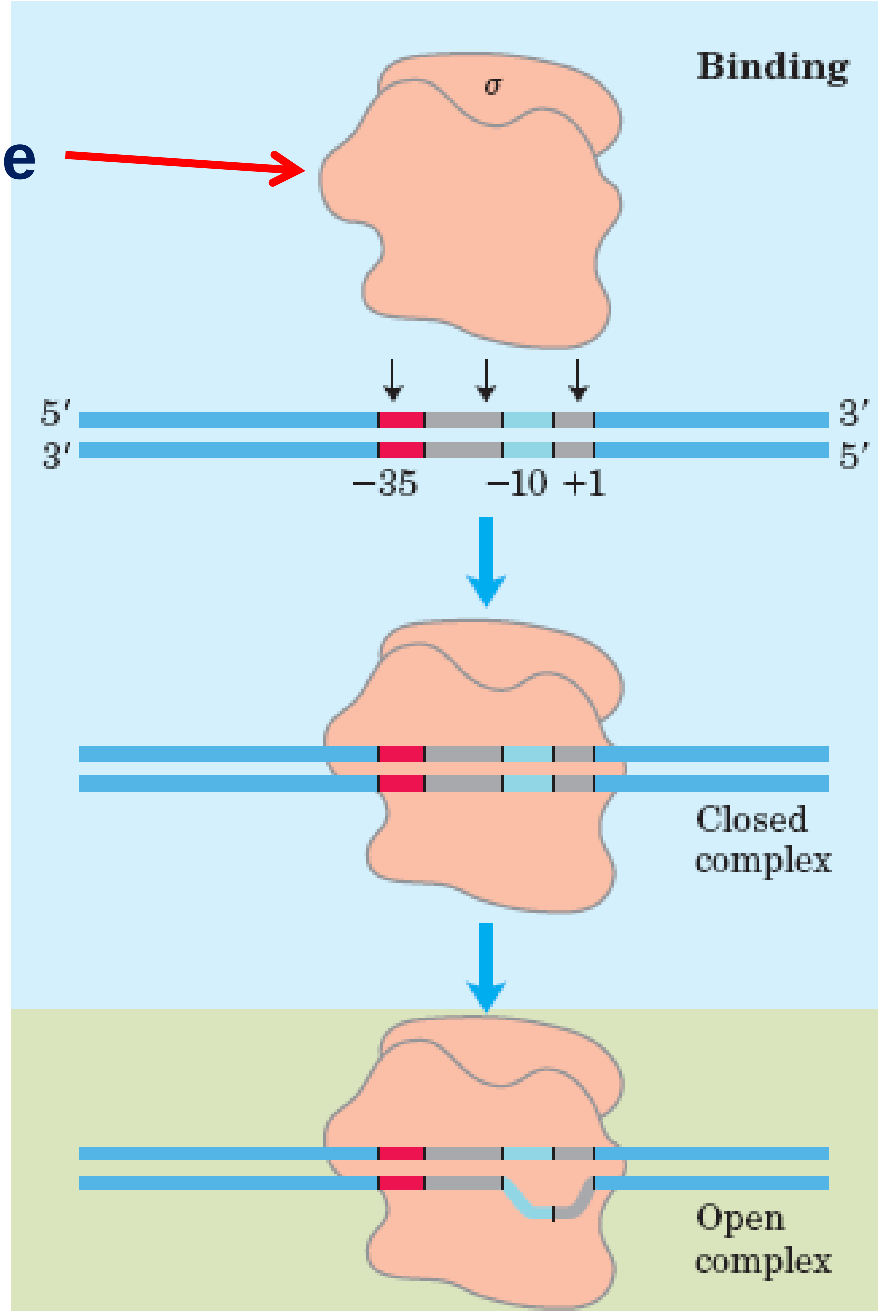
Transcription bubble

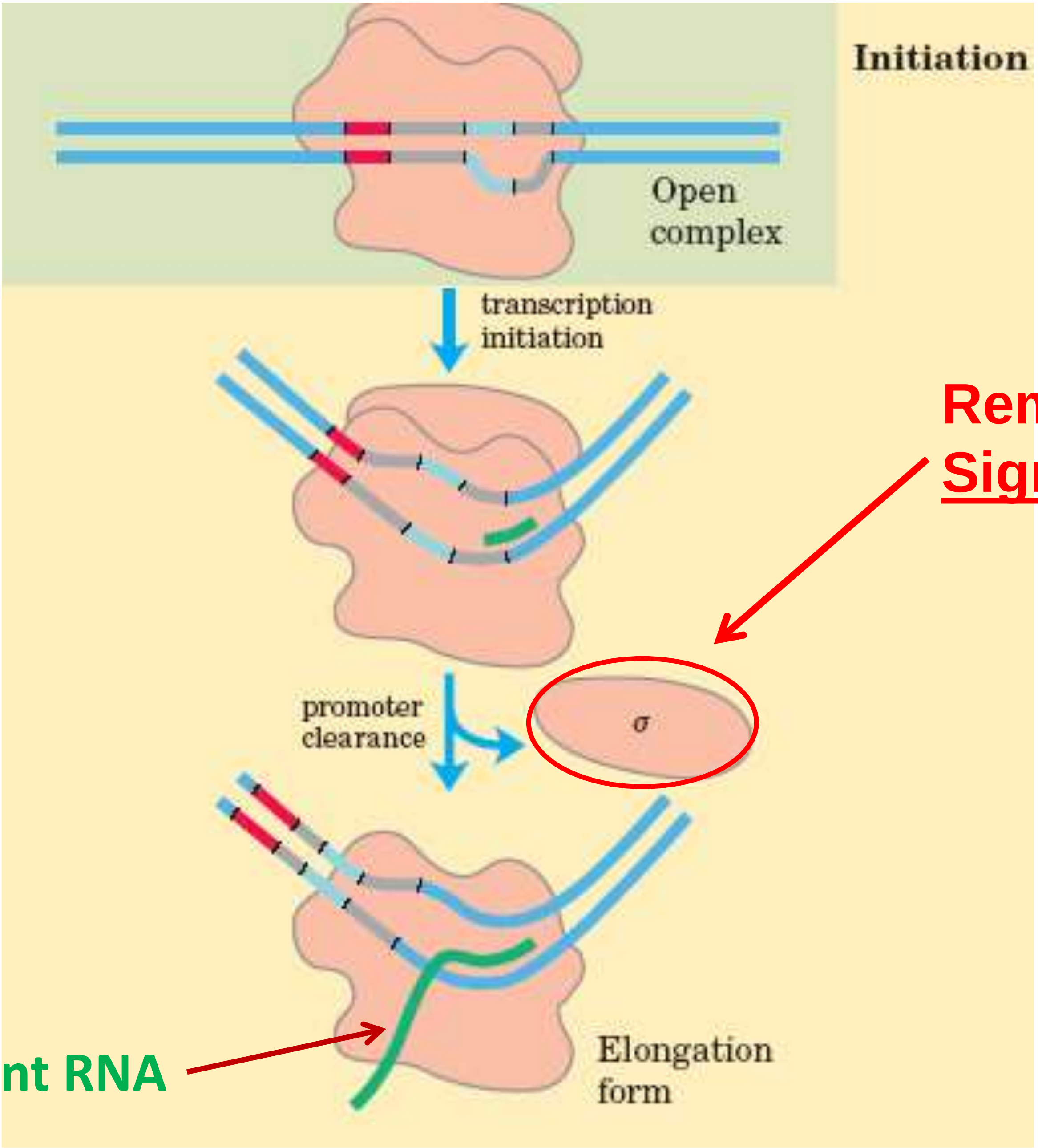


Nascent *RNA*


Negative supercoils
Positive supercoils
5′
RNA
3′
Direction of transcription

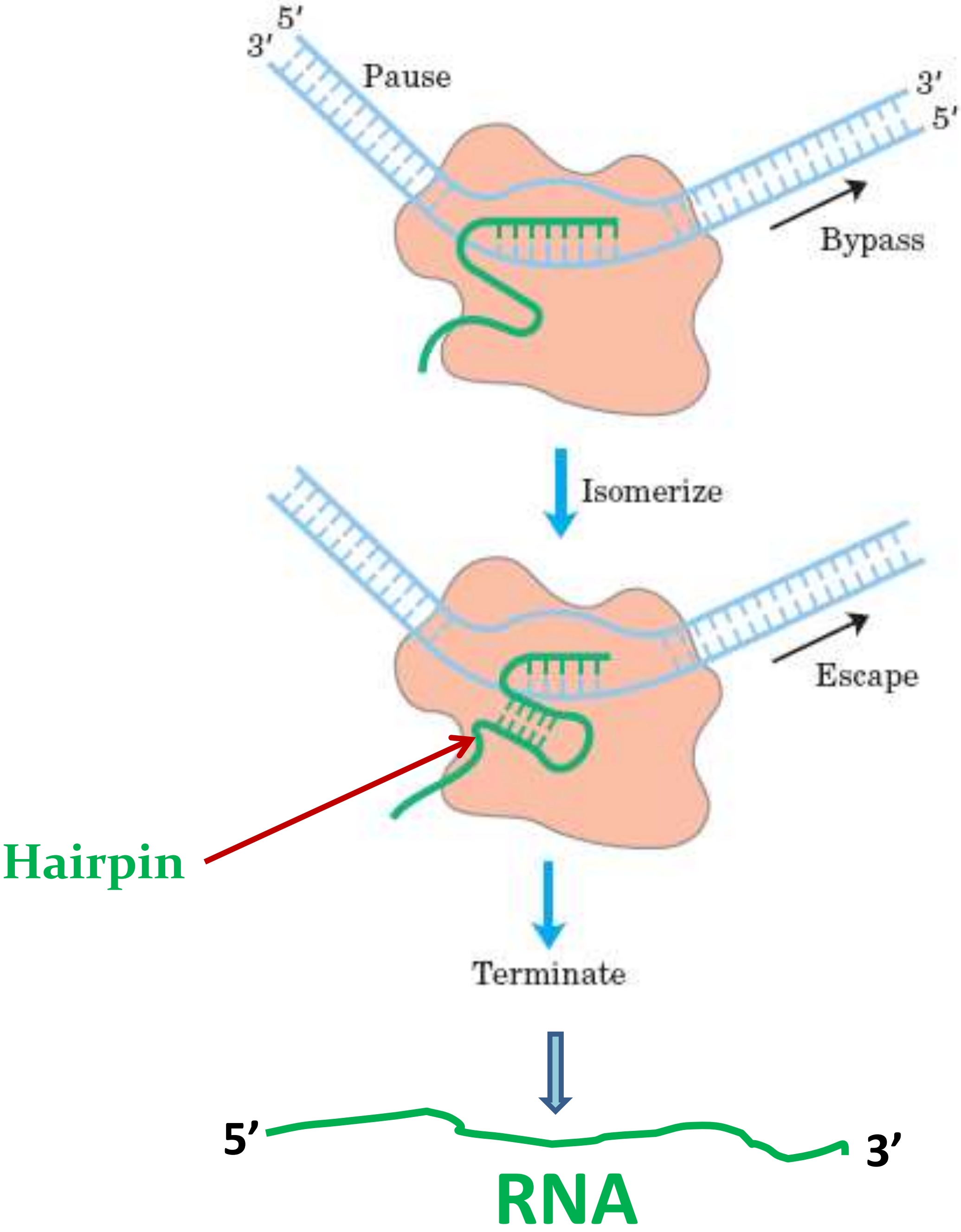
RNA Polymerase





Removal of Sigma subunit

Nascent RNA



Hairpin

RNA

Termination

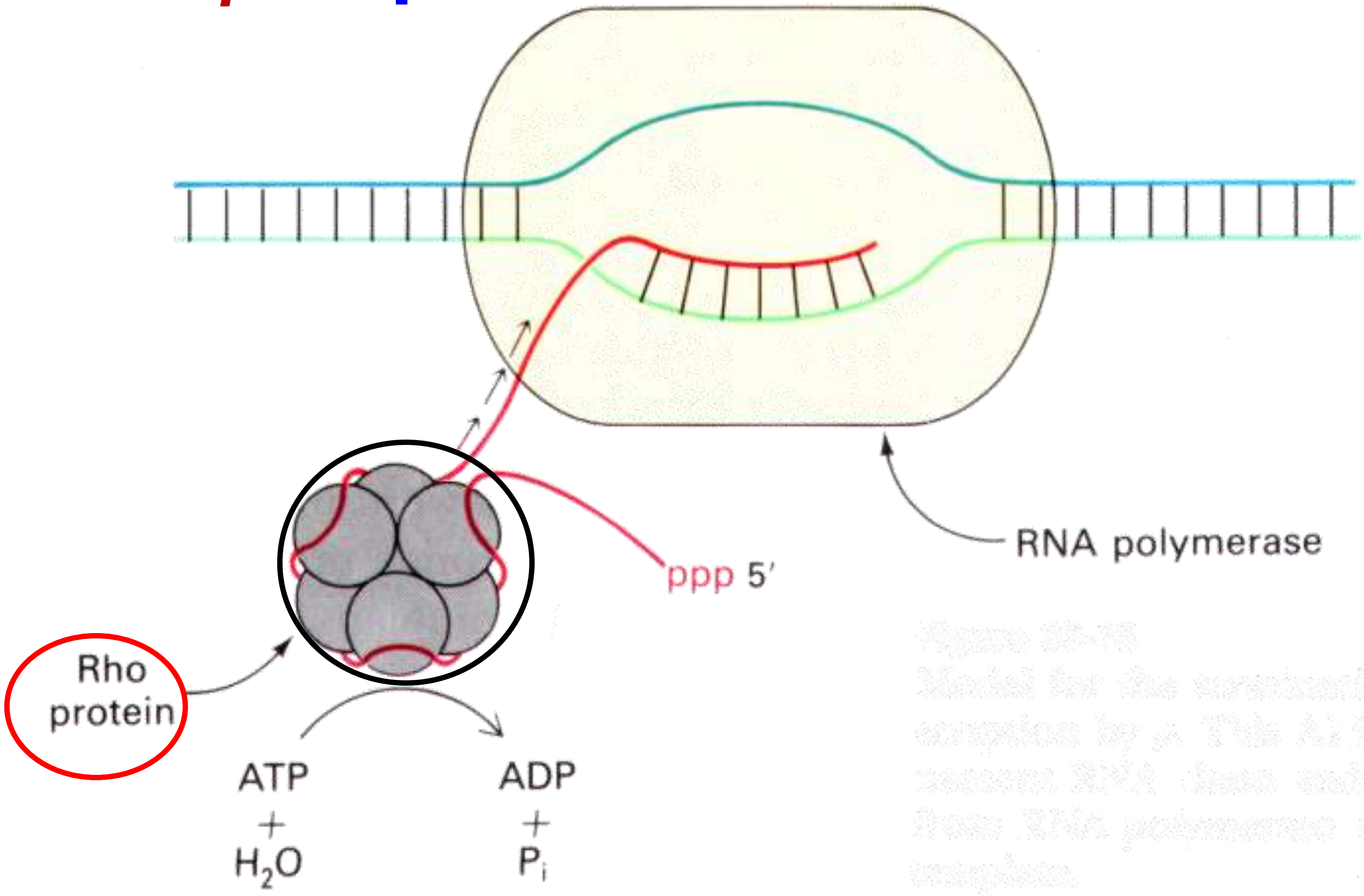
- The **RNA Polymerase** stops moving on the DNA template. The RNA transcript falls off from the transcription complex.

- The termination occurs in either

ρ -dependent or ρ -independent manner.

Termination function of ρ factor

ρ-dependent termination

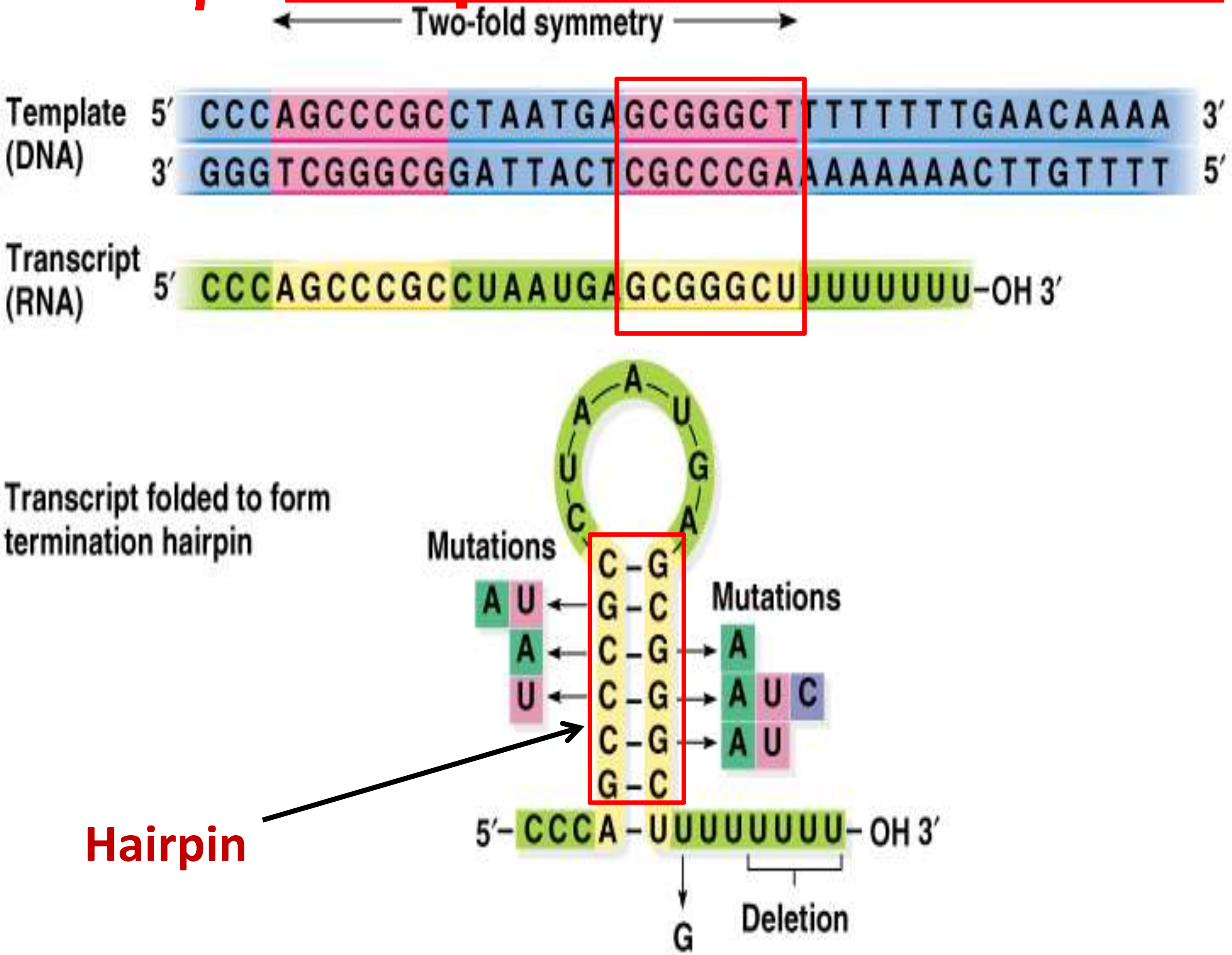


The ρ factor, a hexamer, is a ATPase and a Helicase.

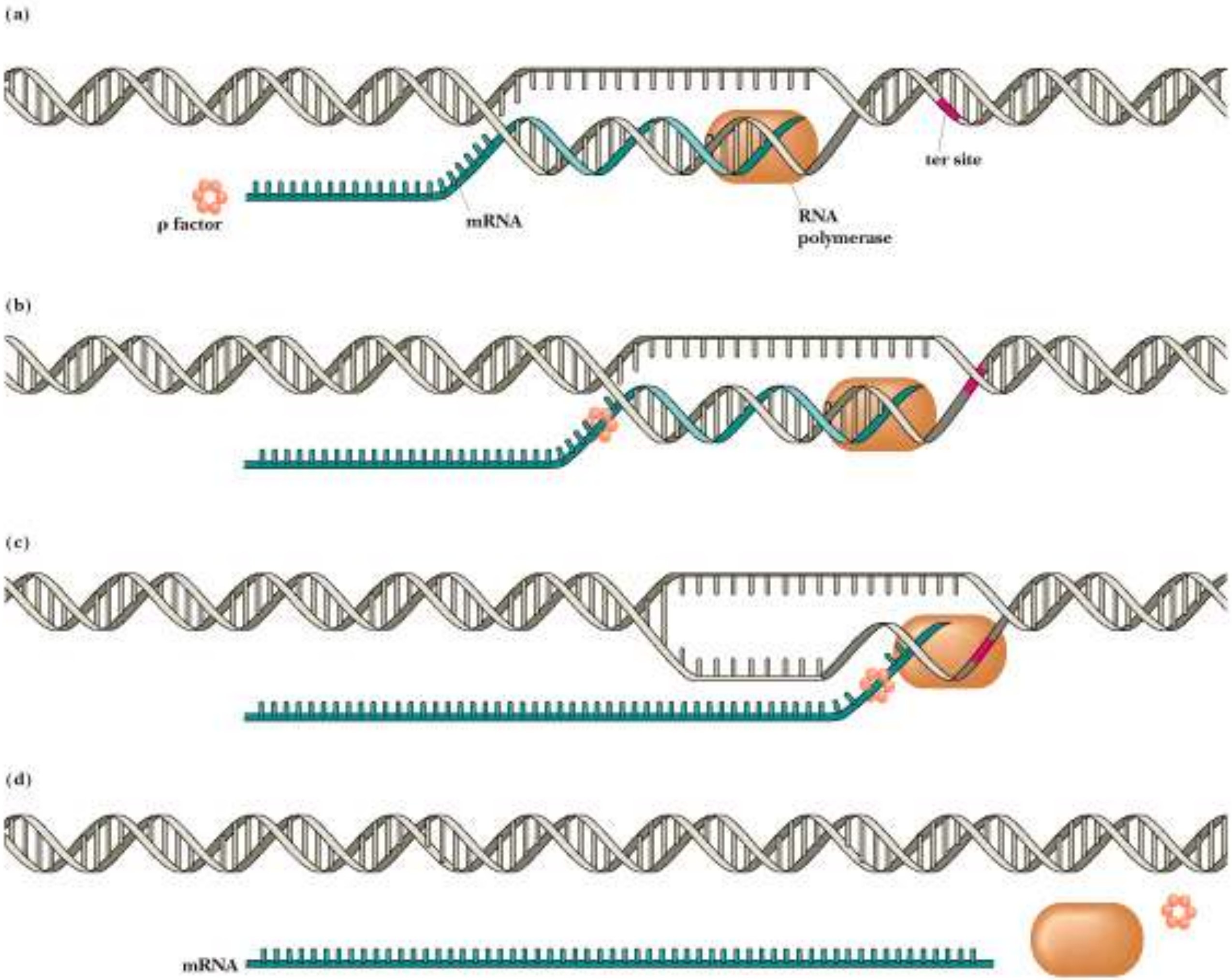
ρ-independent termination

- The termination signal is a stretch of 30-40 nucleotides on the RNA transcript, consisting of many GC followed by a series of U.
- The sequence specificity of this nascent RNA transcript will form particular stem-loop structures to terminate the transcription.

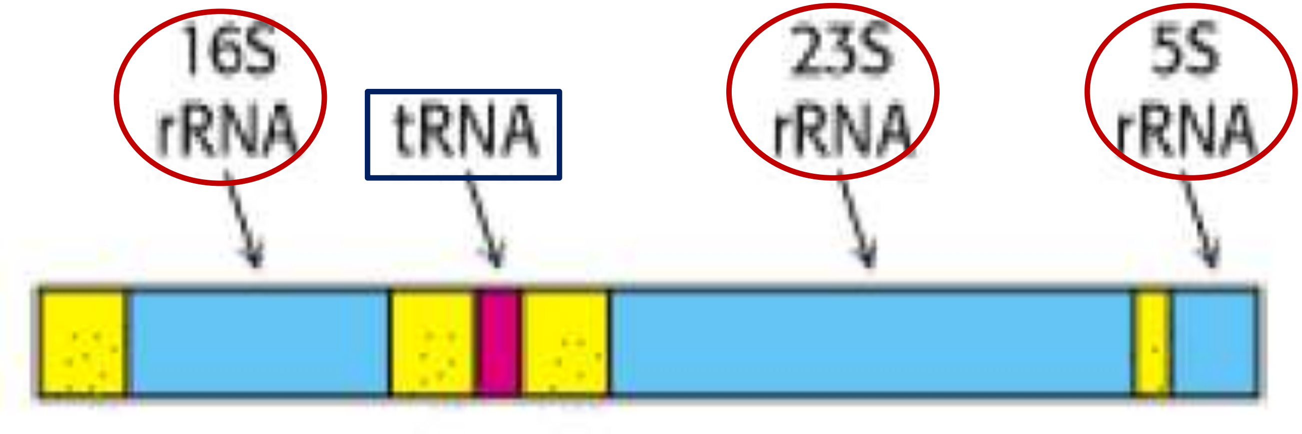
ρ -independent termination



Hairpin

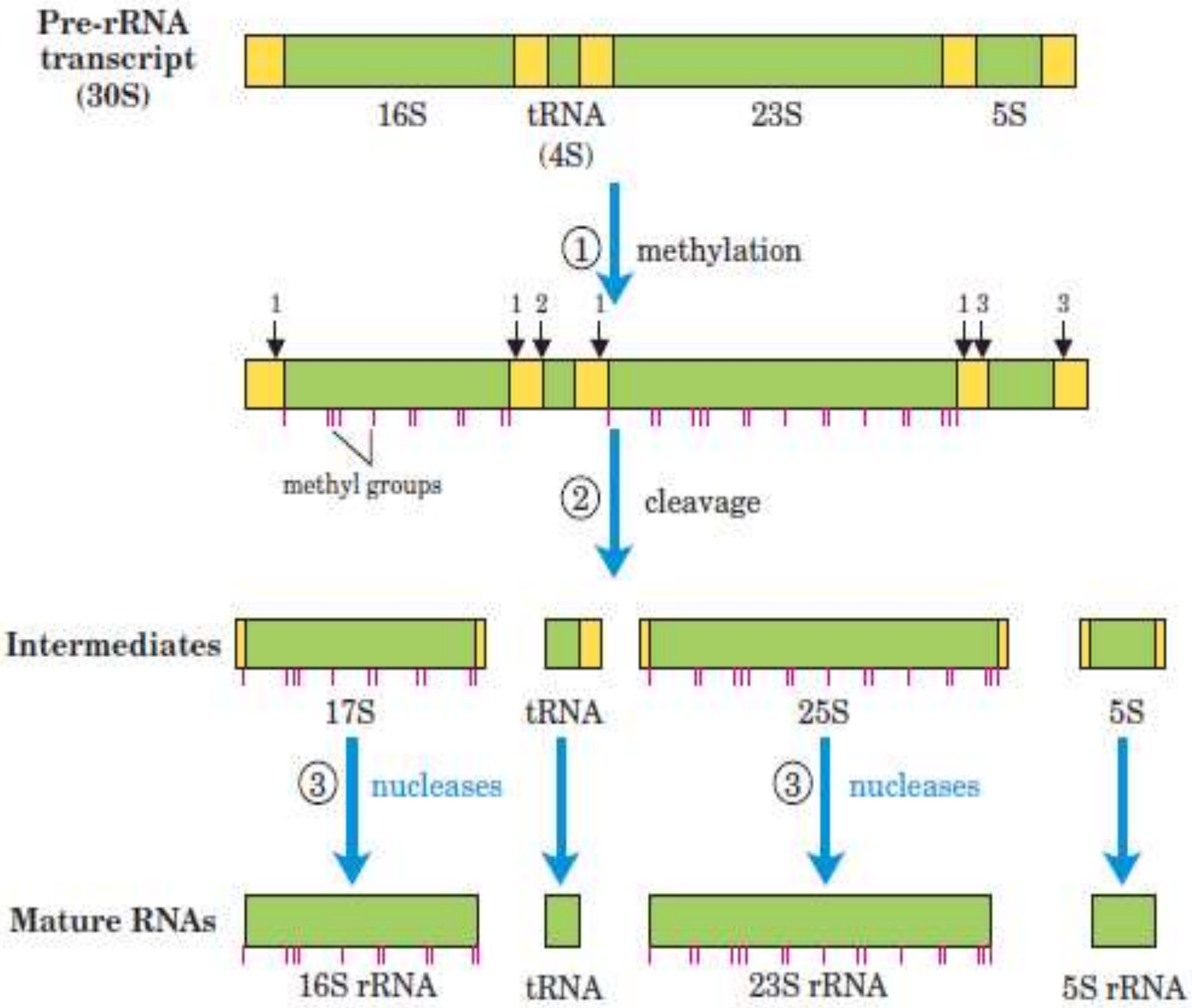

(a)
ρ factor
mRNA
RNA polymerase
ter site
(b)
(c)
(d)
mRNA

Synthesis of rRNA and tRNA

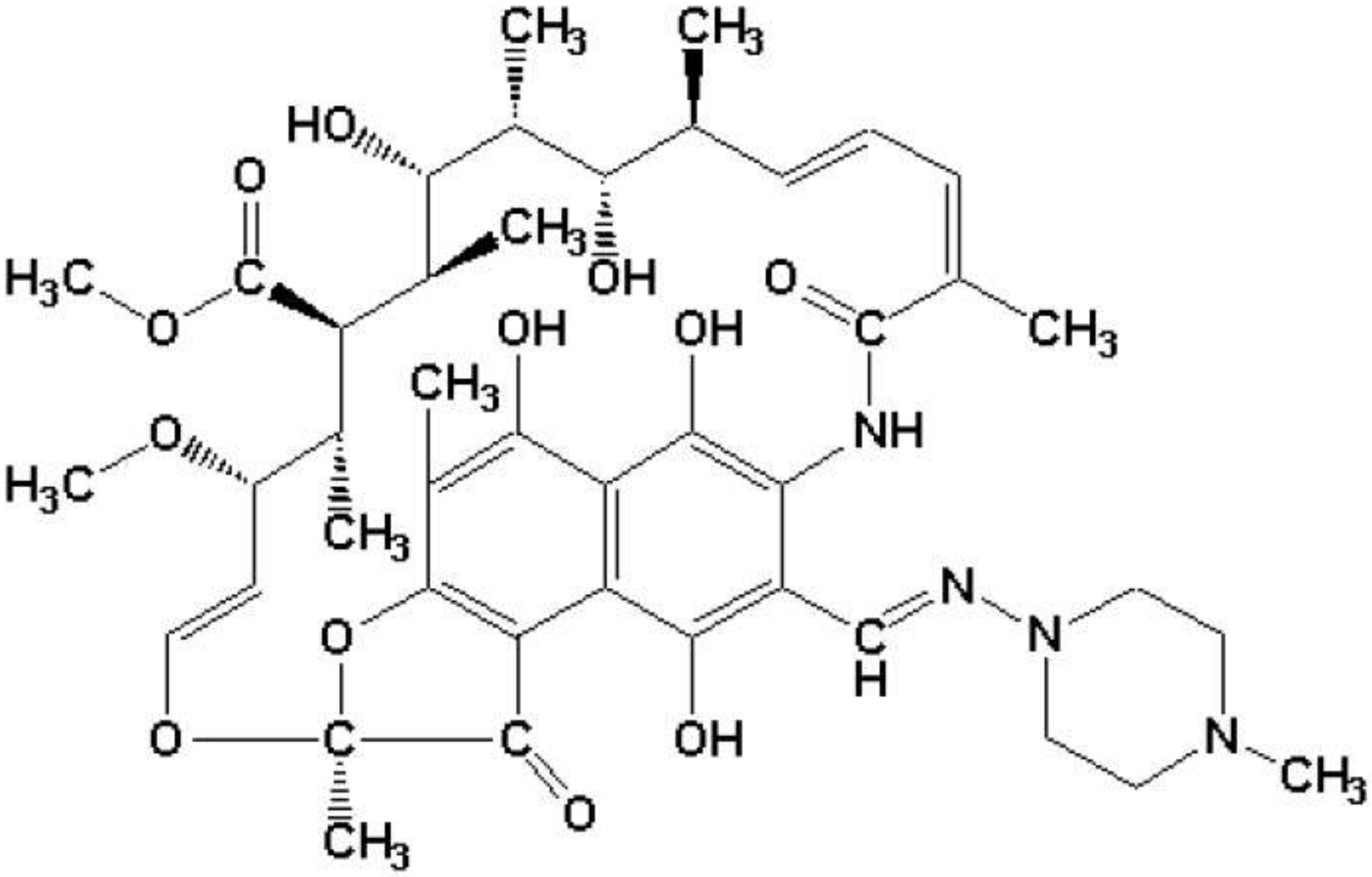


Cleavage of this transcript produces 5S, 16S, and 23S rRNA molecules and a tRNA molecule.

Spacer regions are shown in yellow.

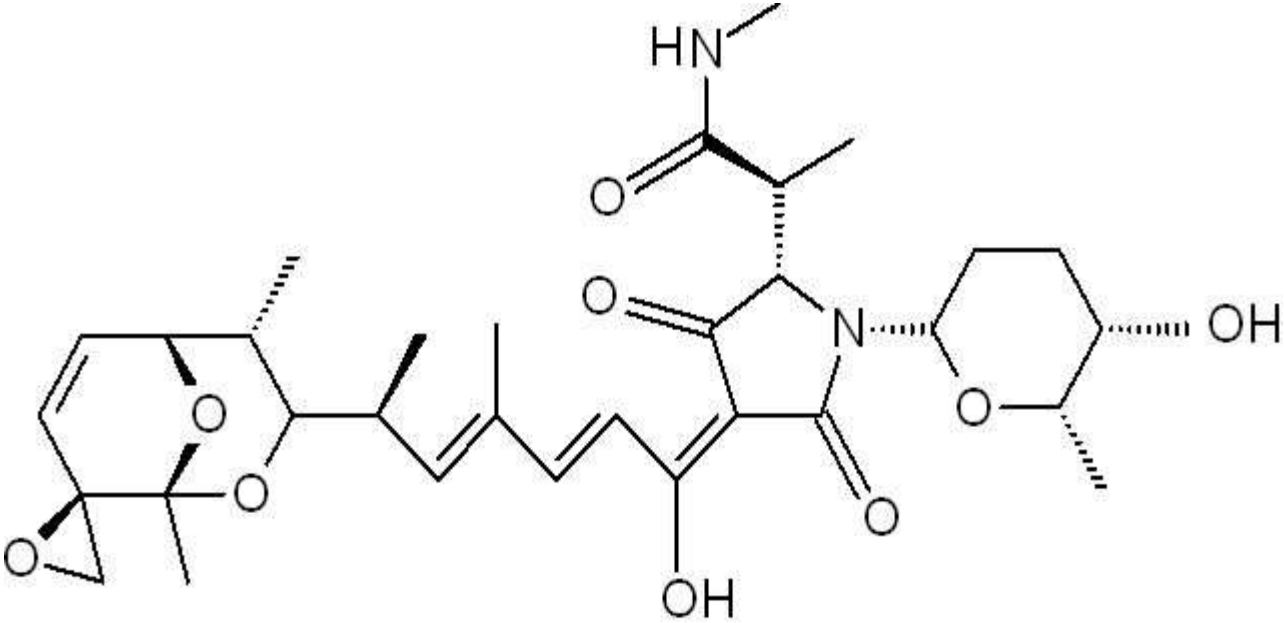

Pre-rRNA transcript (30S)
16S
tRNA (4S)
23S
5S
① methylation
1
1 2
1
1 3
3
methyl groups
② cleavage
Intermediates
17S
tRNA
25S
5S
③ nucleases
③ nucleases
Mature RNAs
16S rRNA
tRNA
23S rRNA
5S rRNA

INHIBITORS OF TRANSCRIPTION



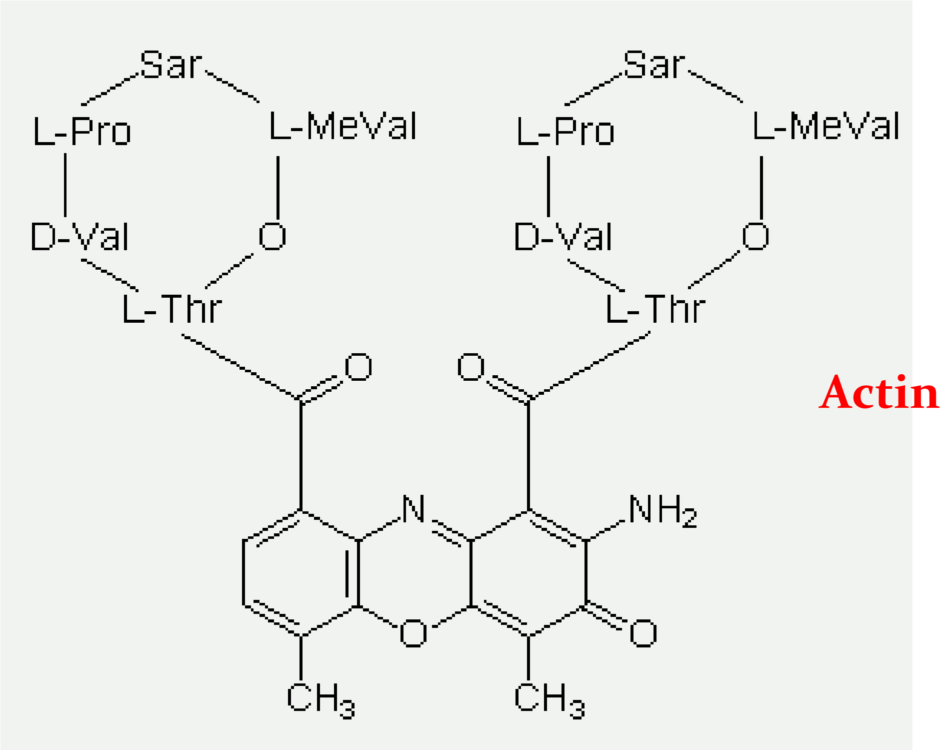
Rifampicin

It binds to the β-subunit of the RNA Polymerase to block the initiation of transcription.



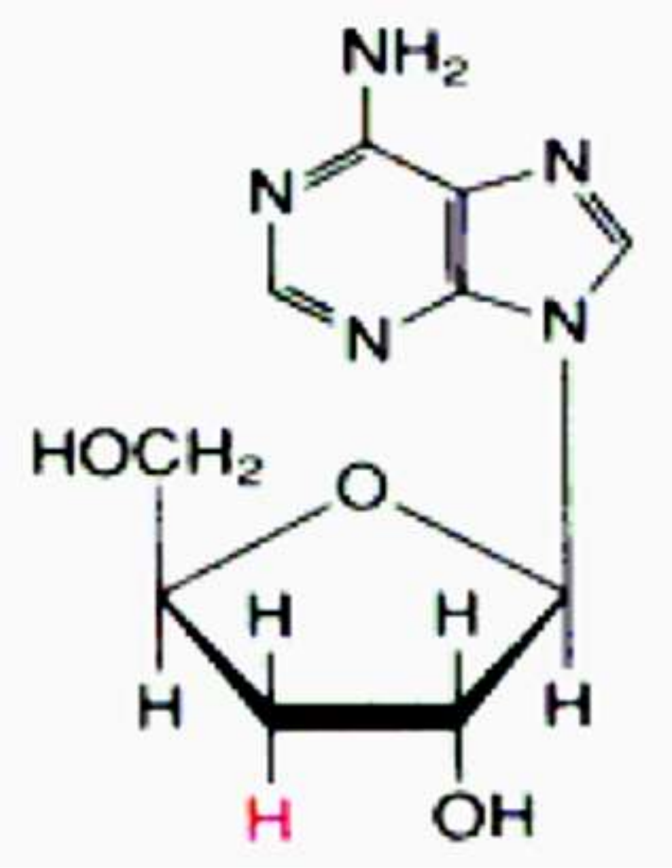
STREPTOLYDIGIN

It binds with the **β-subunit** of prokaryotic **RNA Polymerase** and thus inhibits the **Elongation phase of Transcription.**



Actinomycin D

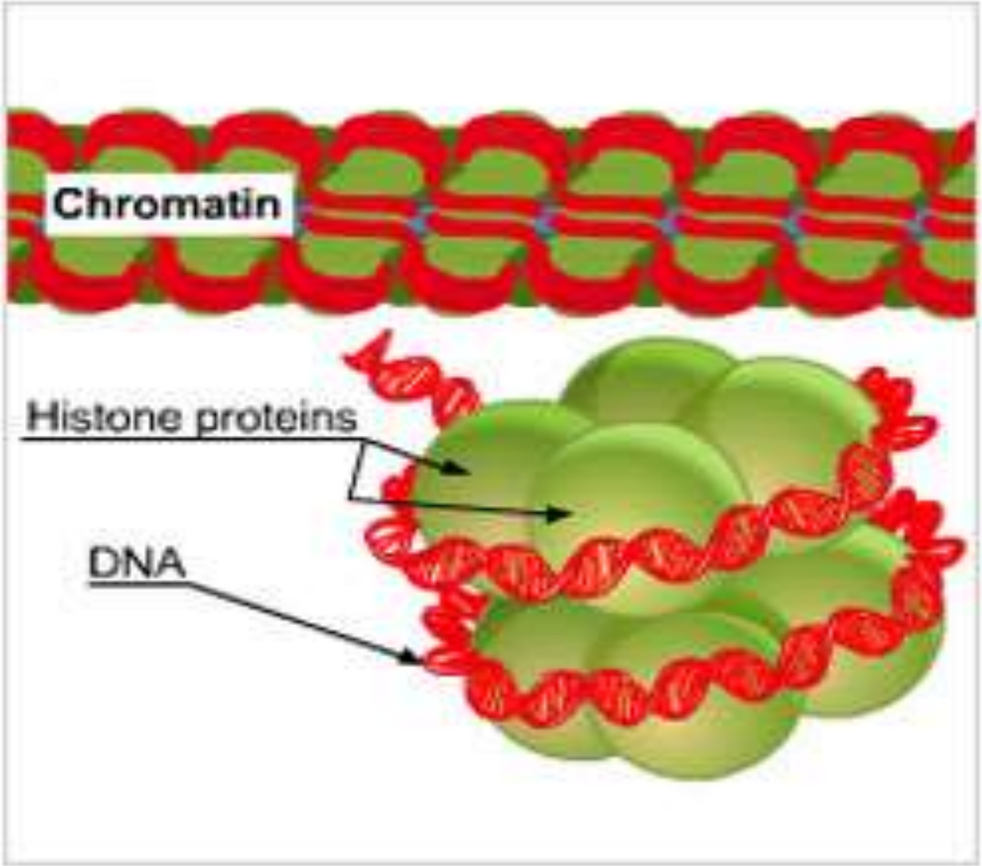
The tricyclic ring system (phenoxazone) of Actinomycin D intercalates between adjacent G-C base pairs, and the cyclic polypeptide arms fill the nearby narrow groove and inhibits Elongation phase of Transcription.

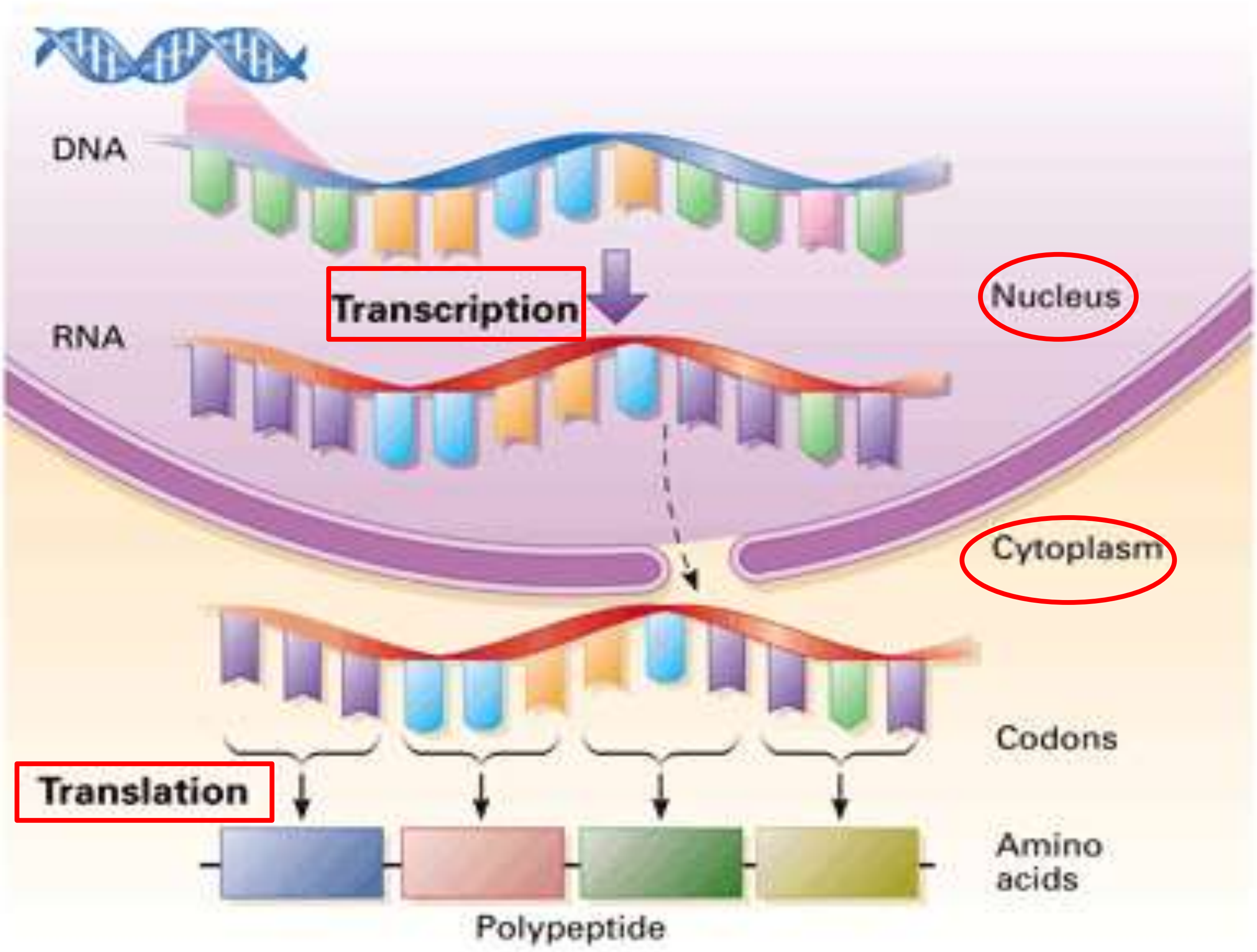


Cordycepin (3-deoxy Adenosine)

It inhibits the Elongation phase of Transcription

EUKARYOTIC TRANSCRIPTION




DNA
Transcription
RNA
Nucleus
Cytoplasm
Codons
Translation
Amino acids
Polypeptide

Eukaryotic RNA polymerases

RNA Polymerase-I: Transcribes / Synthesizes

- 28s rRNA
- 18s rRNA
- 5.8s rRNA

RNA Polymerase-II: Transcribes

- m-RNA
- Some sn-RNA

RNA Polymerase-III: Transcribes

- t-RNA
- 5s rRNA
- Some sn-RNA

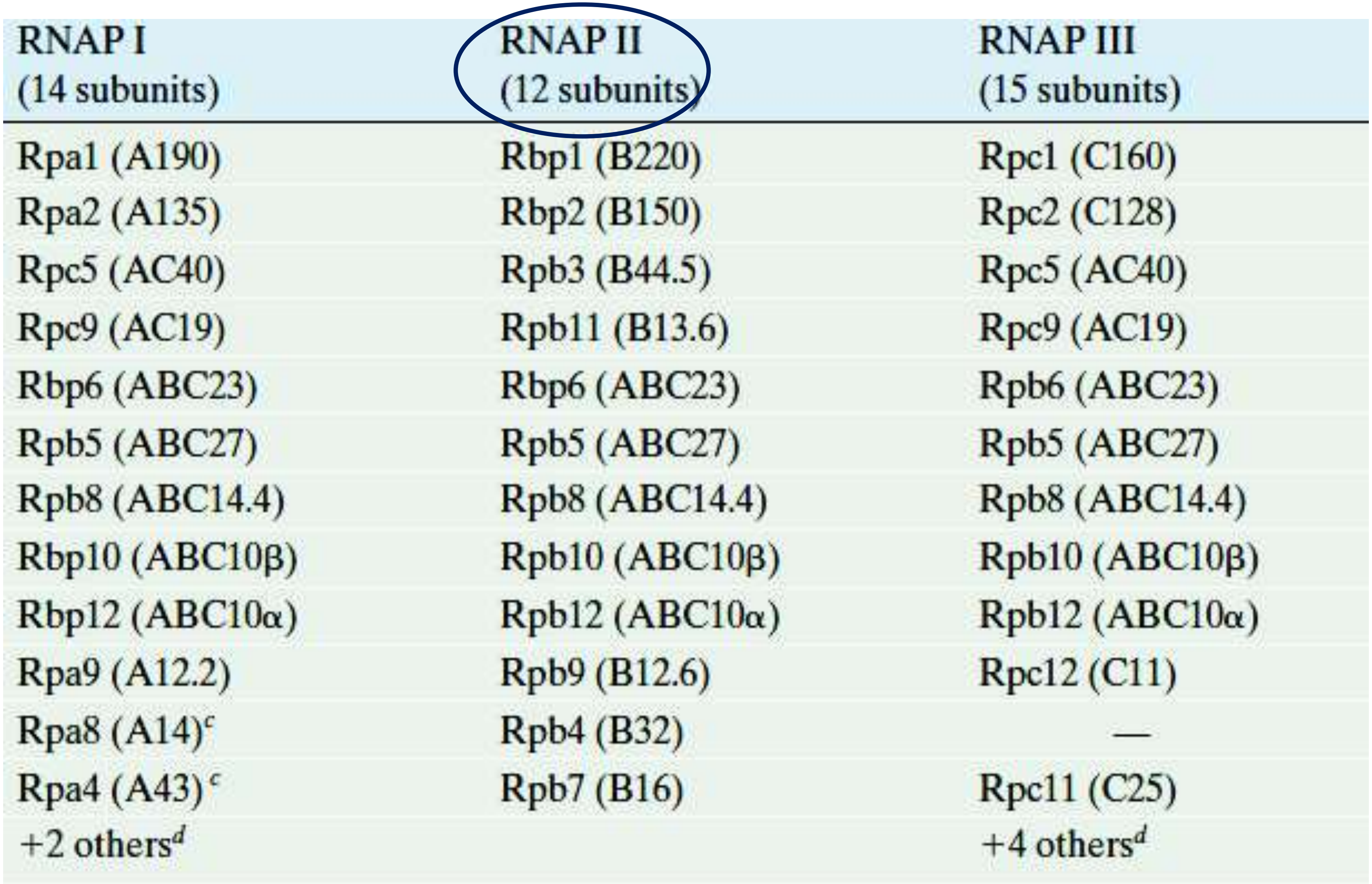
RNA POLYMERASE-II

- **RNA polymerase II** is central to eukaryotic gene expression and has been studied extensively.

- **RNA polymerase II** is a multi subunit enzyme with **12 subunits**.

- **RNA polymerase II** requires an array of other proteins, called **transcription factors (TF II)** in order to form the active transcription complex.

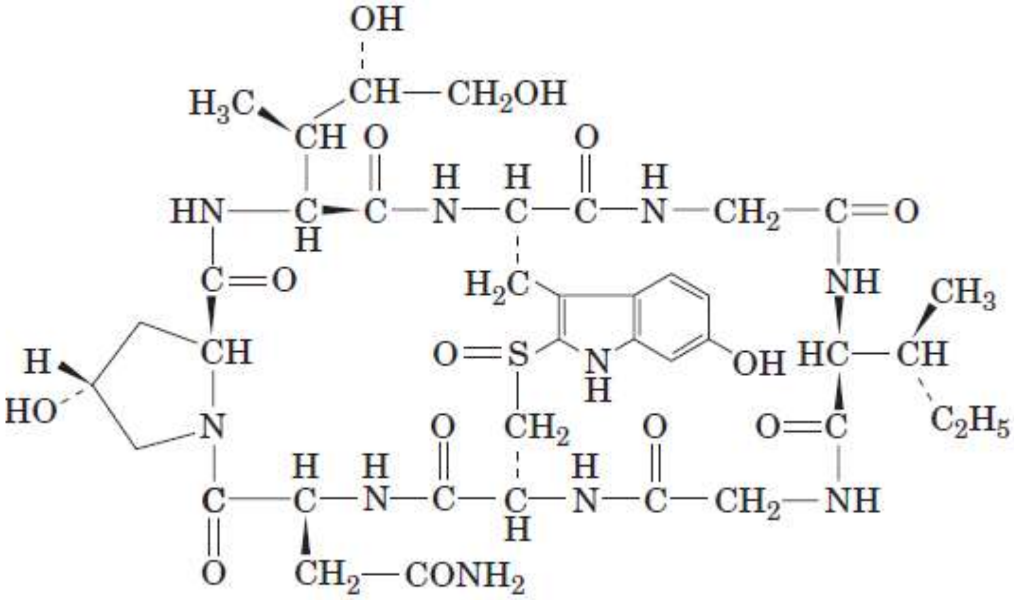
Eukaryotic RNA polymerases



RNAP I (14 subunits)	RNAP II (12 subunits)	RNAP III (15 subunits)
Rpa1 (A190)	Rbp1 (B220)	Rpc1 (C160)
Rpa2 (A135)	Rbp2 (B150)	Rpc2 (C128)
Rpc5 (AC40)	Rpb3 (B44.5)	Rpc5 (AC40)
Rpc9 (AC19)	Rpb11 (B13.6)	Rpc9 (AC19)
Rbp6 (ABC23)	Rbp6 (ABC23)	Rpb6 (ABC23)
Rpb5 (ABC27)	Rpb5 (ABC27)	Rpb5 (ABC27)
Rpb8 (ABC14.4)	Rpb8 (ABC14.4)	Rpb8 (ABC14.4)
Rbp10 (ABC10β)	Rpb10 (ABC10β)	Rpb10 (ABC10β)
Rbp12 (ABC10α)	Rpb12 (ABC10α)	Rpb12 (ABC10α)
Rpa9 (A12.2)	Rpb9 (B12.6)	Rpc12 (C11)
Rpa8 (A14)[c]	Rpb4 (B32)	—
Rpa4 (A43)[c]	Rpb7 (B16)	Rpc11 (C25)
+2 others[d]		+4 others[d]

α-Amanitin (Fungal toxin from Amanita phalloides) - cyclic octapeptide with unussual amino acids.

Chapter 13 Opener
Genetics: A Conceptual Approach, Fourth Edition
© 2012 W. H. Freeman and Company



α-Amanitin

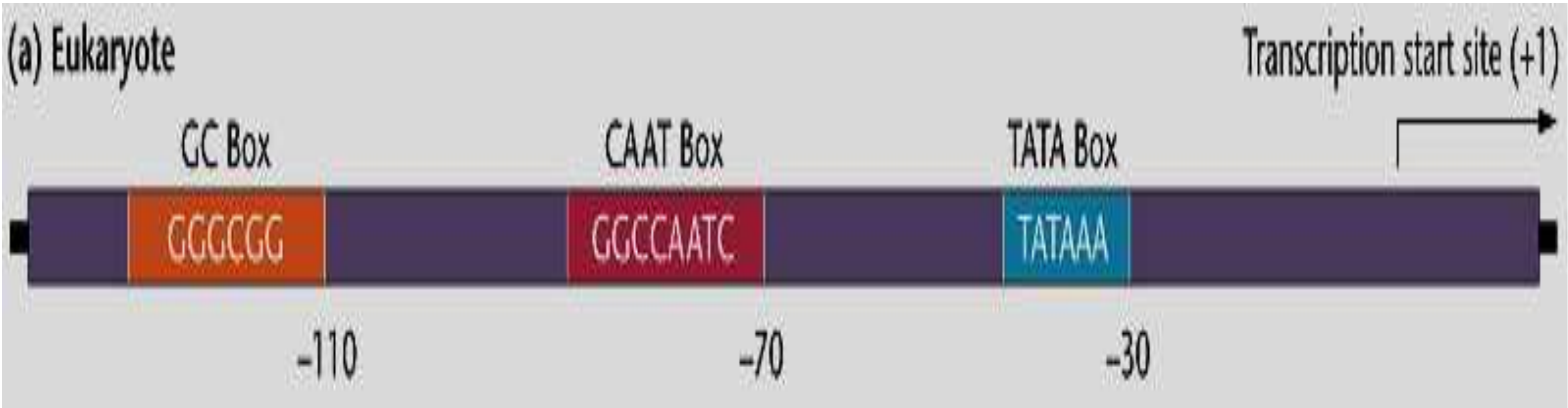
Inhibitor of eukaryotic RNA polymerase (mainly of type II)

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

Eukaryotic Transcription

Promoters

- ✓ **Much more complex than those found in bacteria.**
- ✓ **These are consensus sequences located at the upstream regions of Coding strand.**
- ✓ **Mutation of this region usually significantly lowers the rate of transcription.**



1) TATA box (Hogness Box)

Very similar to the prokaryotic TATA box, except the sequence is slightly different (TATAAA) and it is located in between -25 to -30.

2) CAAT box

Located in between -70 to -80.

Always contains CCAAT.

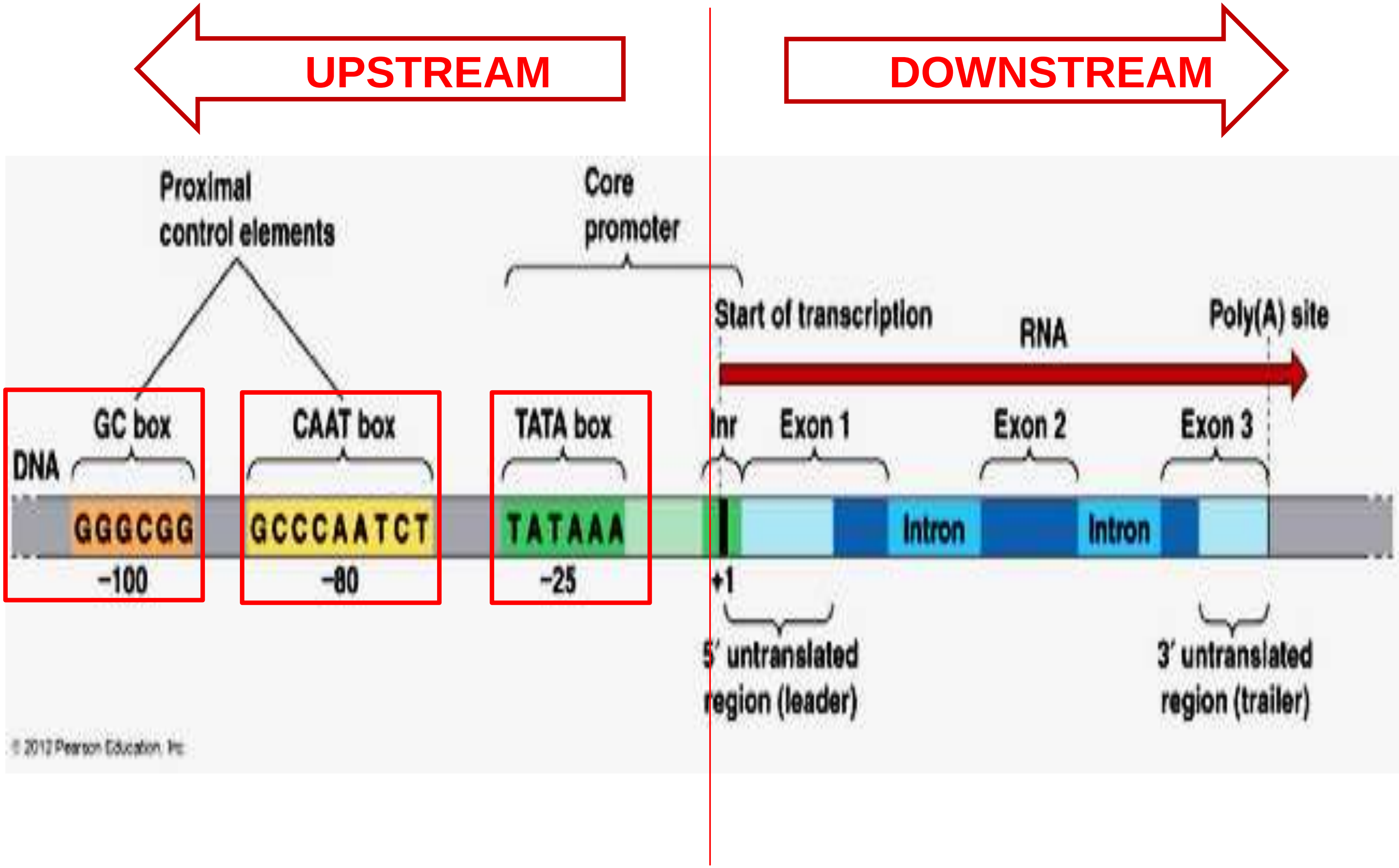
3) GC box

Usually has the sequence GGGCGG and is typically found at -110.

ENHANCERS :

- Enhancers elements are the sequences located in a variety of regions of a gene both upstream and downstream of the transcription start site and even within the transcribed portions of some genes.

- Enhancers increases the transcription rate by several folds.



UPSTREAM

DOWNSTREAM

© 2012 Pearson Education, Inc.

Transcription factors

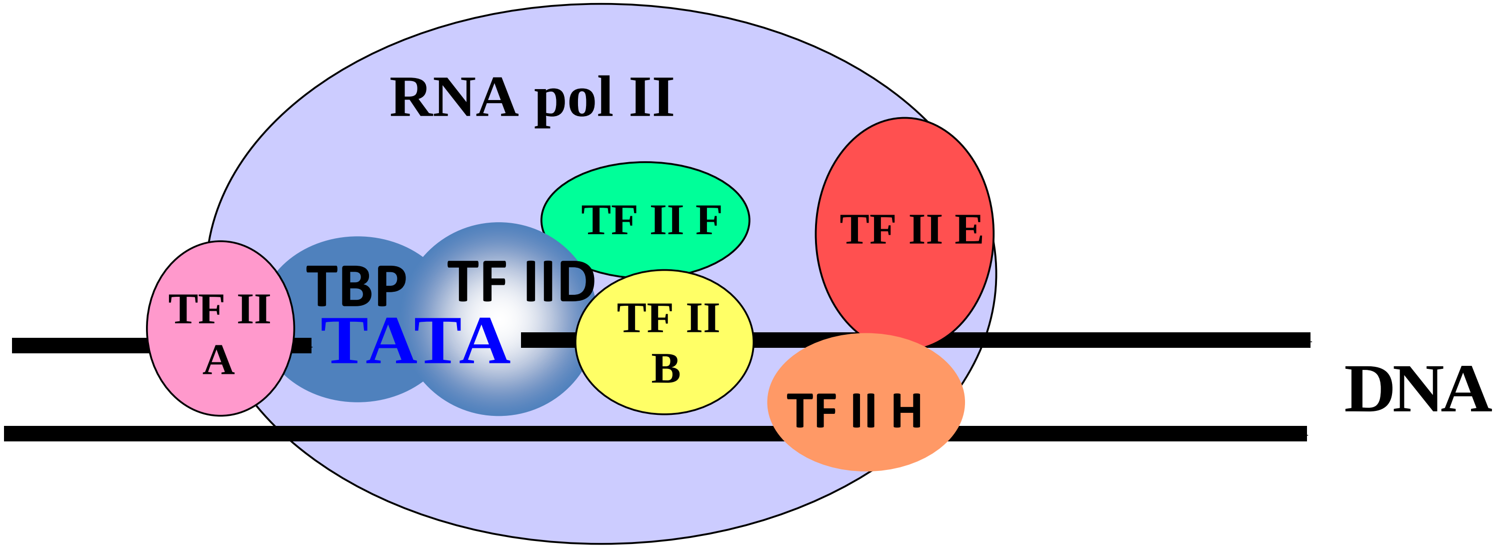
- RNA-pol II does not bind to the promoter sequences directly.
- RNA-pol II associates with six transcription factors.
- TFII A, TFII B, TFII D, TFII E, TFII F and TFII H

Proteins Required for Initiation of Transcription at the RNA Polymerase II (Pol II)

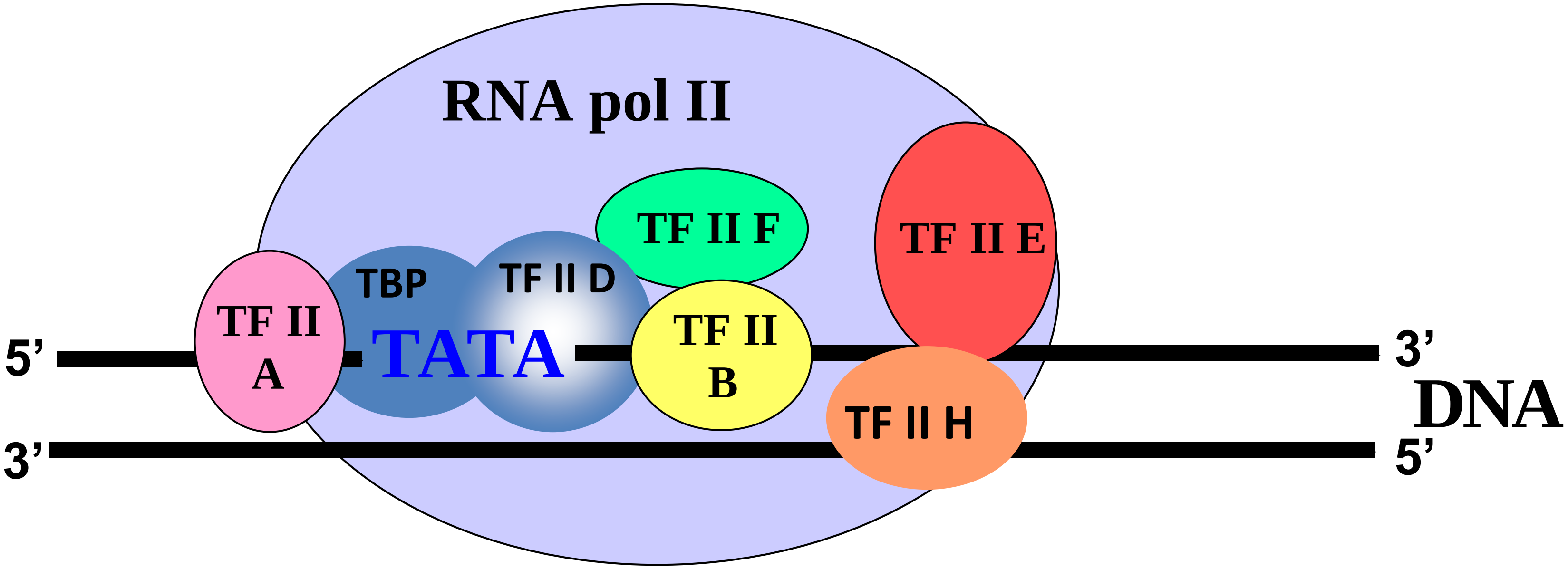
Transcription protein	*Number of subunits*	*Subunit(s) M_r*	*Function(s)*
Initiation			
Pol II	12	10,000–220,000	Catalyzes RNA synthesis
TBP (TATA-binding protein)	1	38,000	Specifically recognizes the TATA box
TFIIA	3	12,000, 19,000, 35,000	Stabilizes binding of TFIIB and TBP to the promoter
TFIIB	1	35,000	Binds to TBP; recruits Pol II–TFIIF complex
TFIIE	2	34,000, 57,000	Recruits TFIIH; has ATPase and helicase activities
TFIIF	2	30,000, 74,000	Binds tightly to Pol II; binds to TFIIB and prevents binding of Pol II to nonspecific DNA sequences
TFIIH	12	35,000–89,000	Unwinds DNA at promoter (helicase activity); phosphorylates Pol II (within the CTD); recruits nucleotide-excision repair proteins
Elongation[*]			
ELL[†]	1	80,000	
p-TEFb	2	43,000, 124,000	Phosphorylates Pol II (within the CTD)
SII (TFIIS)	1	38,000	
Elongin (SIII)	3	15,000, 18,000, 110,000	

Pre-initiation complex (PIC)

- TBP of TFII D binds TATA–Box(-10 sequence)
- TFII A and TFII B bind TFII D
- TFII F- RNA-pol complex binds TFII B
- TFII F and TFII E open the dsDNA (helicase and ATPase)
- TFII H: completion of PIC



Pre-initiation complex (PIC)



DNA + RNA Poly-II + TBP + Transcription Factors (TF)

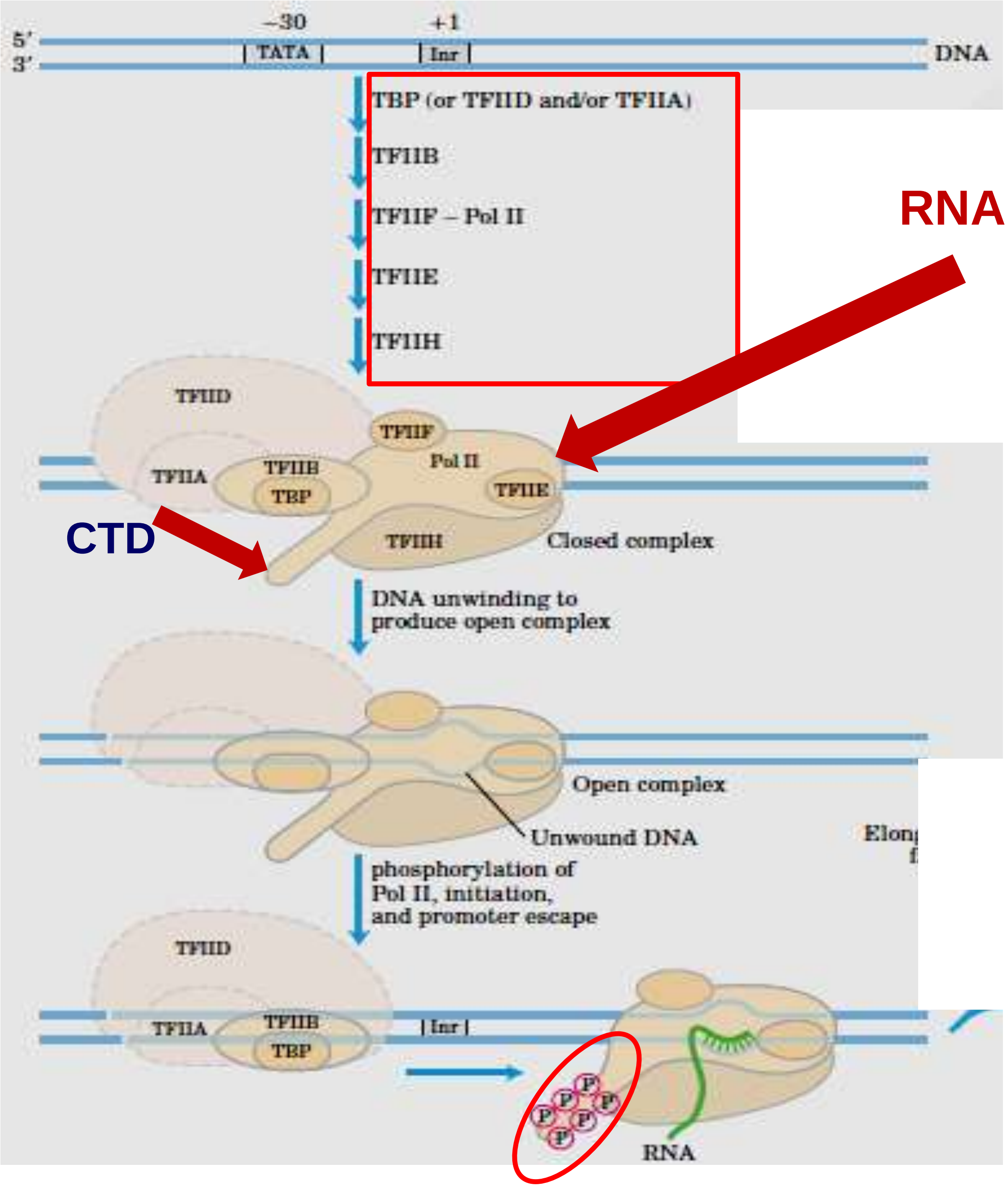
Phosphorylation of RNA-Polymerase-II

- **TF II H is of protein kinase activity to phosphorylate CTD of RNA pol-II.**

(CTD is the C-terminal domain of RNA pol-II)

- **Only the RNA Polymerase can move toward the downstream, starting the elongation phase.**

- **Most of the Transcription Factors fall off from PIC during the elongation phase.**



RNA Poly-II

CTD

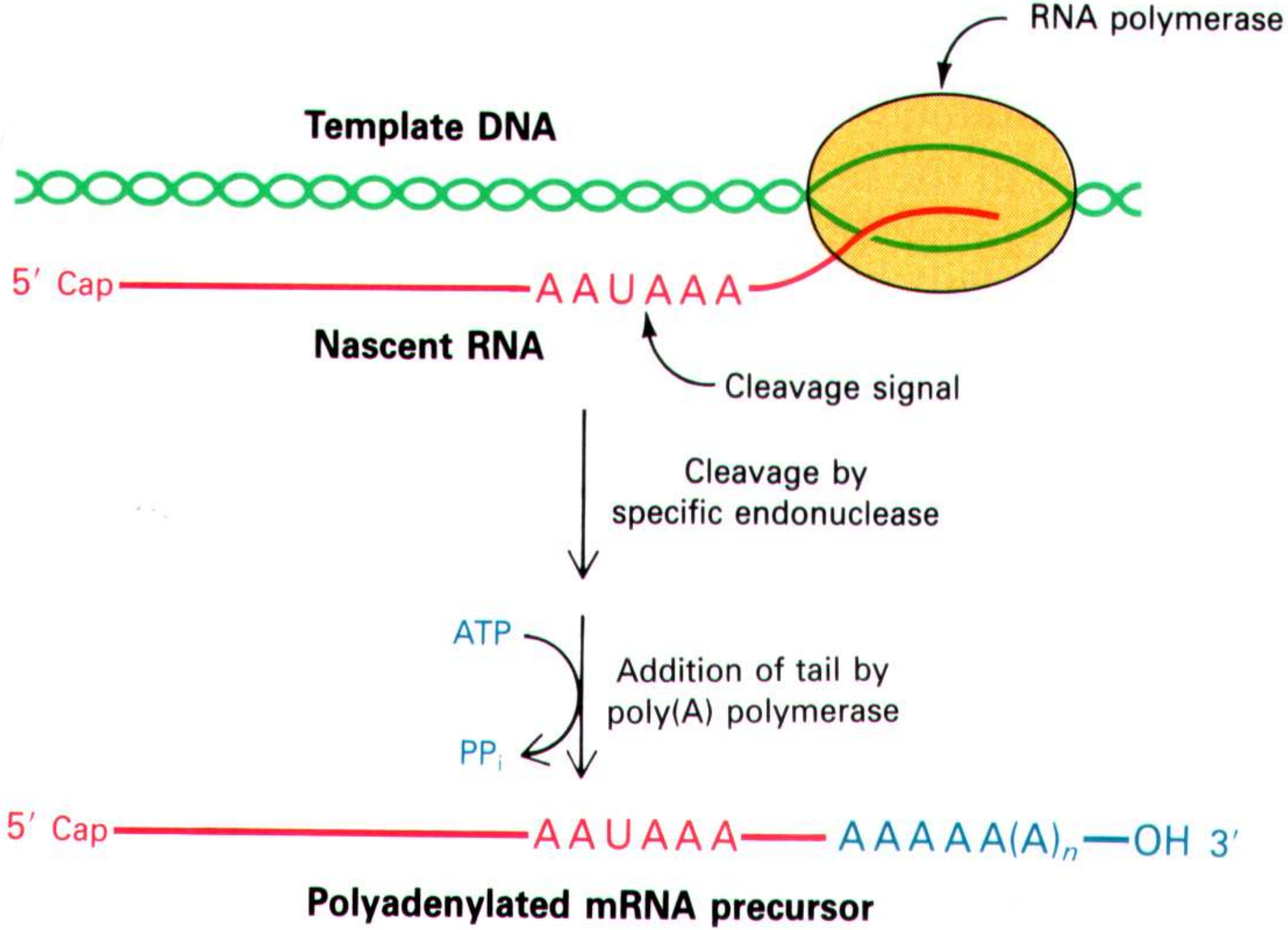
Termination

- **When the RNA Polymerase transcribes the terminator region of the DNA, the polymerase releases the mRNA**
- **The termination sequence is AATAAA followed by GT repeats.**

Elongation

TFIIF remains associated with RNA Pol-II throughout elongation.

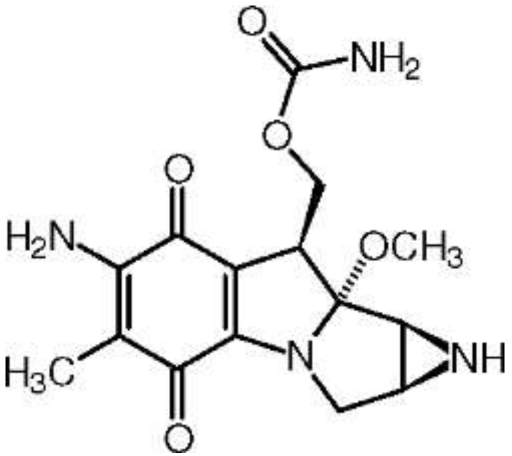
The activity of the RNA poly-II is greatly enhanced by proteins called Elongation factors


RNA polymerase
Template DNA
5′ Cap
AAUAAA
Nascent RNA
Cleavage signal
Cleavage by
specific endonuclease
ATP
PP$_i$
Addition of tail by
poly(A) polymerase
5′ Cap
AAUAAA
AAAAA(A)$_n$—OH 3′
Polyadenylated mRNA precursor

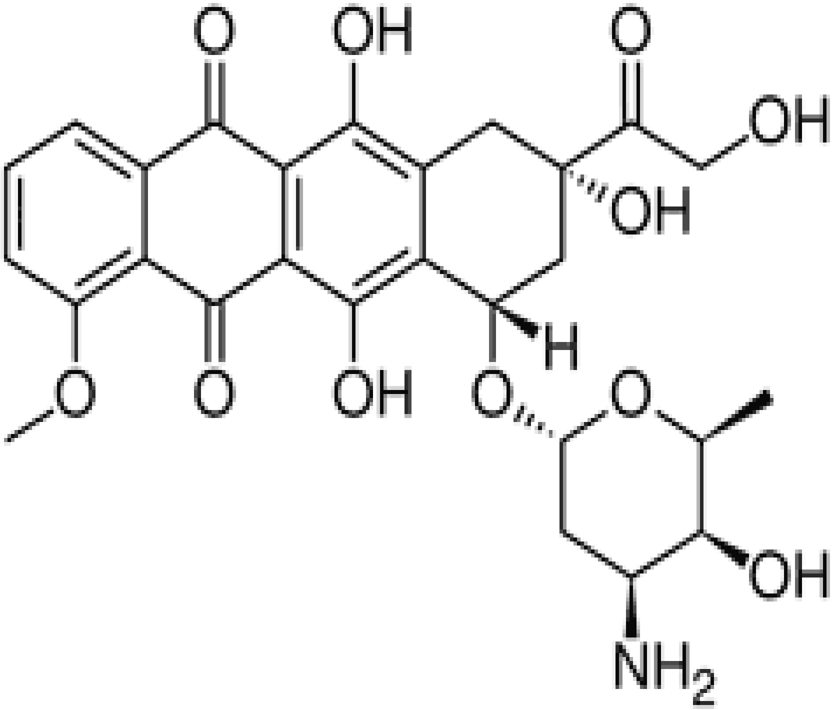
INHIBITORS
OF
EUKARYOTIC TRANSCRIPTION

Mitomycin

- Mitomycin- **Intercalates with DNA strands**
- Blocks transcription
- Used as **anticancer drug**

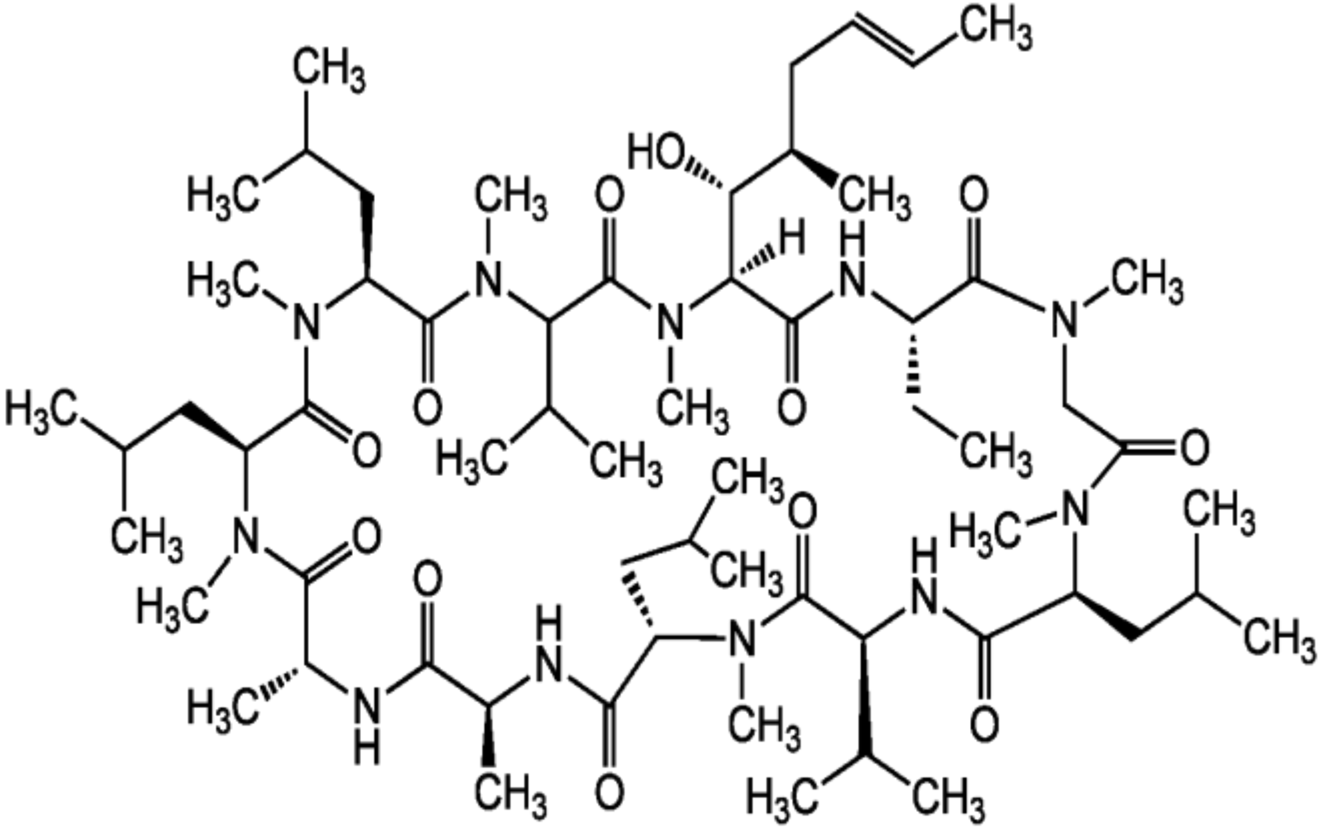


ADRIAMYCIN



Inhibits the Initiation phase by preventing the interaction of TF-IID with RNA-Poly-II and DNA complex.

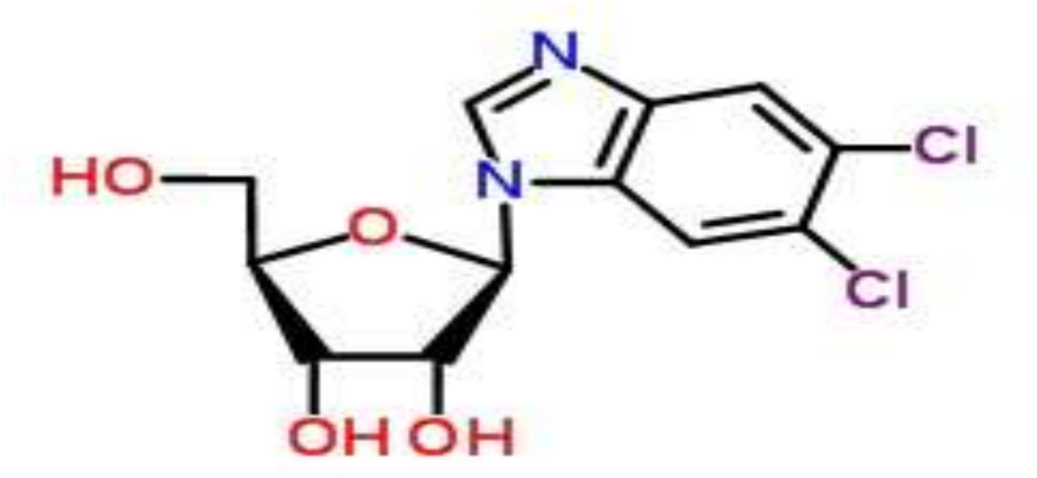
CYCLOSPORIN -A



Immunosuppressant Drug inhibits Transcription in T-Cells.

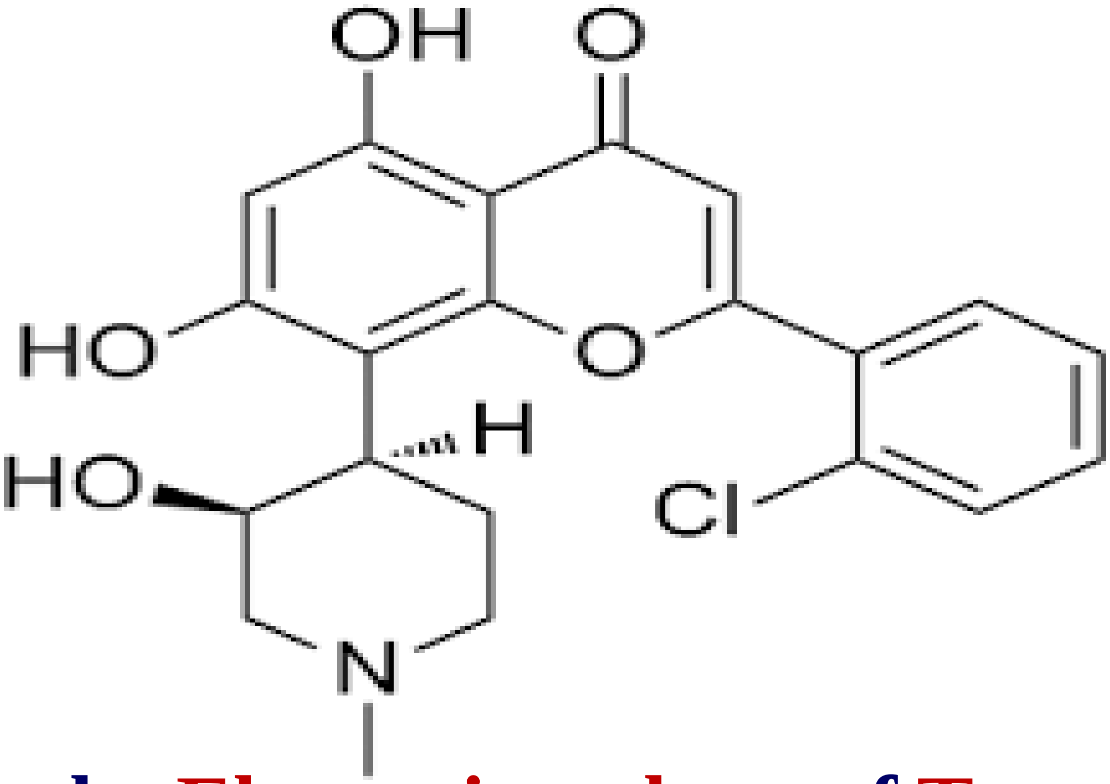
DRB

(5,6-dichlorobenzimidazone-1-β-D-ribofuranoside)



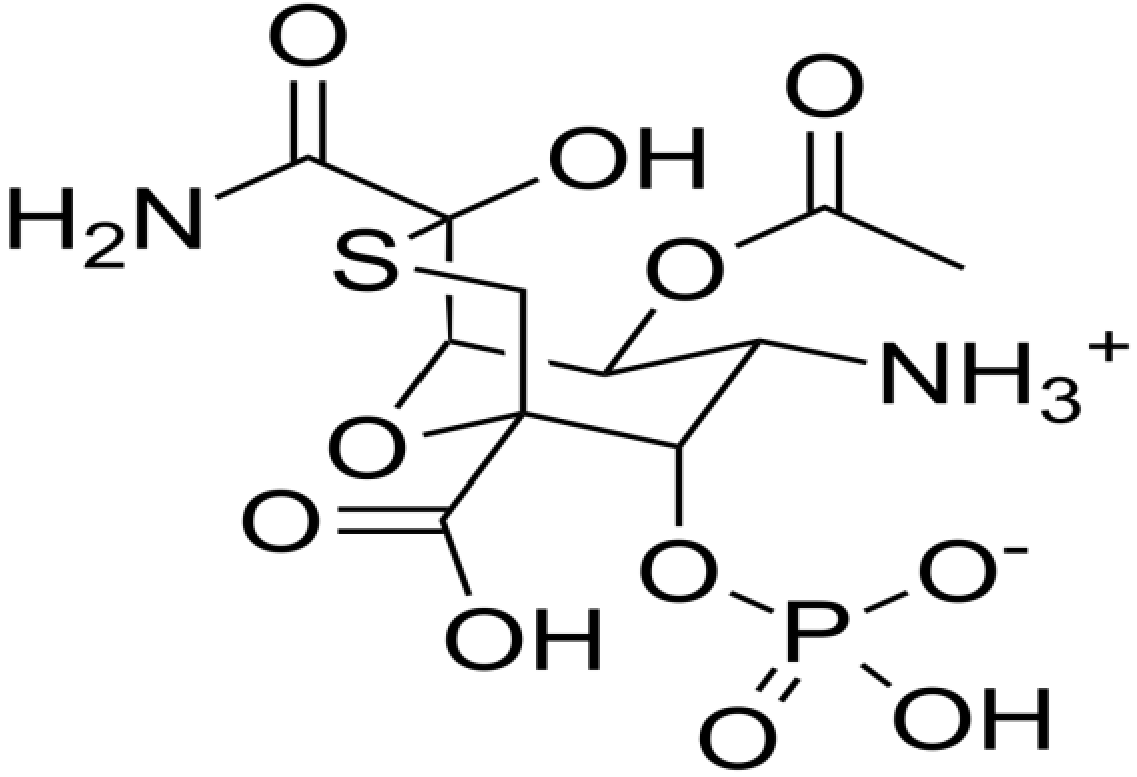
Inhibits the Elongation phase of Transcription by selectively inhibiting RNA Poly-II.

Flavopyridol (Alvocidib)



Inhibits the Elongation phase of Transcription by selectively inhibiting RNA Poly-II.

Tagetitoxin

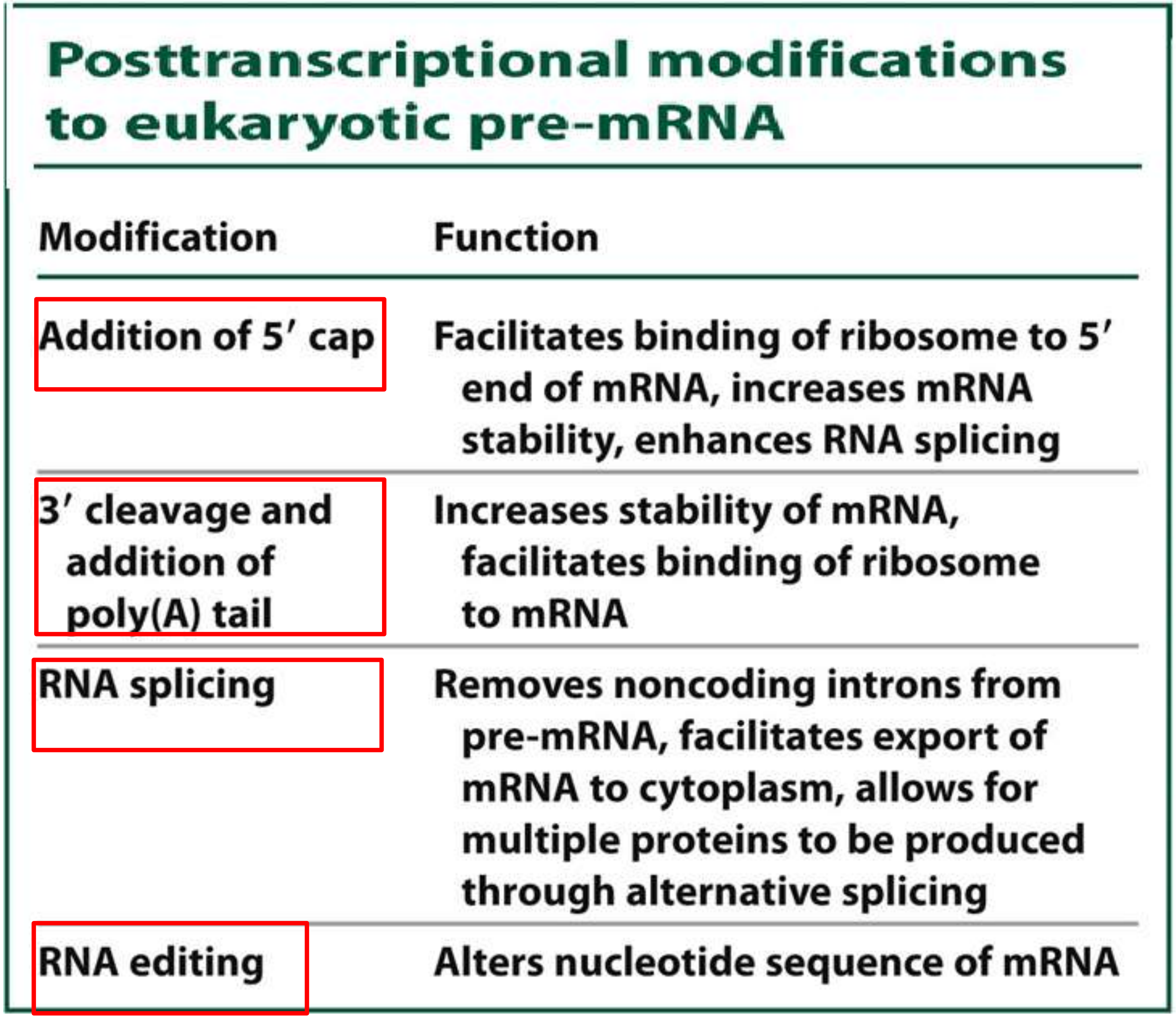


Inhibits tRNA synthesis by binding to RNA Poly-III.

Post-Transcriptional Modifications

- **The nascent RNA, also known as Primary transcript, needs to be modified to become functional, mRNAs, tRNAs and rRNAs.**

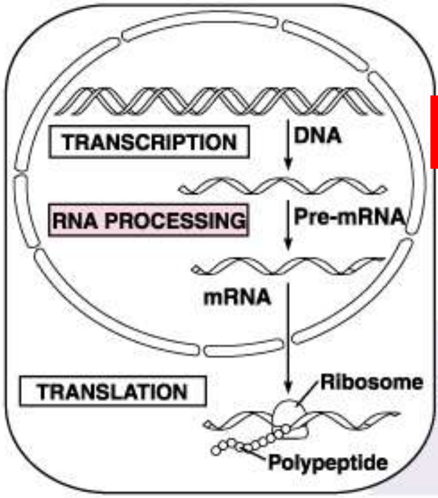
- **These modification is critical to eukaryotic systems.**

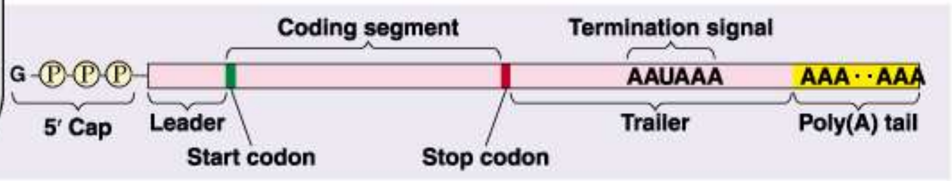


Posttranscriptional modifications to eukaryotic pre-mRNA

Modification	Function
Addition of 5′ cap	Facilitates binding of ribosome to 5′ end of mRNA, increases mRNA stability, enhances RNA splicing
3′ cleavage and addition of poly(A) tail	Increases stability of mRNA, facilitates binding of ribosome to mRNA
RNA splicing	Removes noncoding introns from pre-mRNA, facilitates export of mRNA to cytoplasm, allows for multiple proteins to be produced through alternative splicing
RNA editing	Alters nucleotide sequence of mRNA

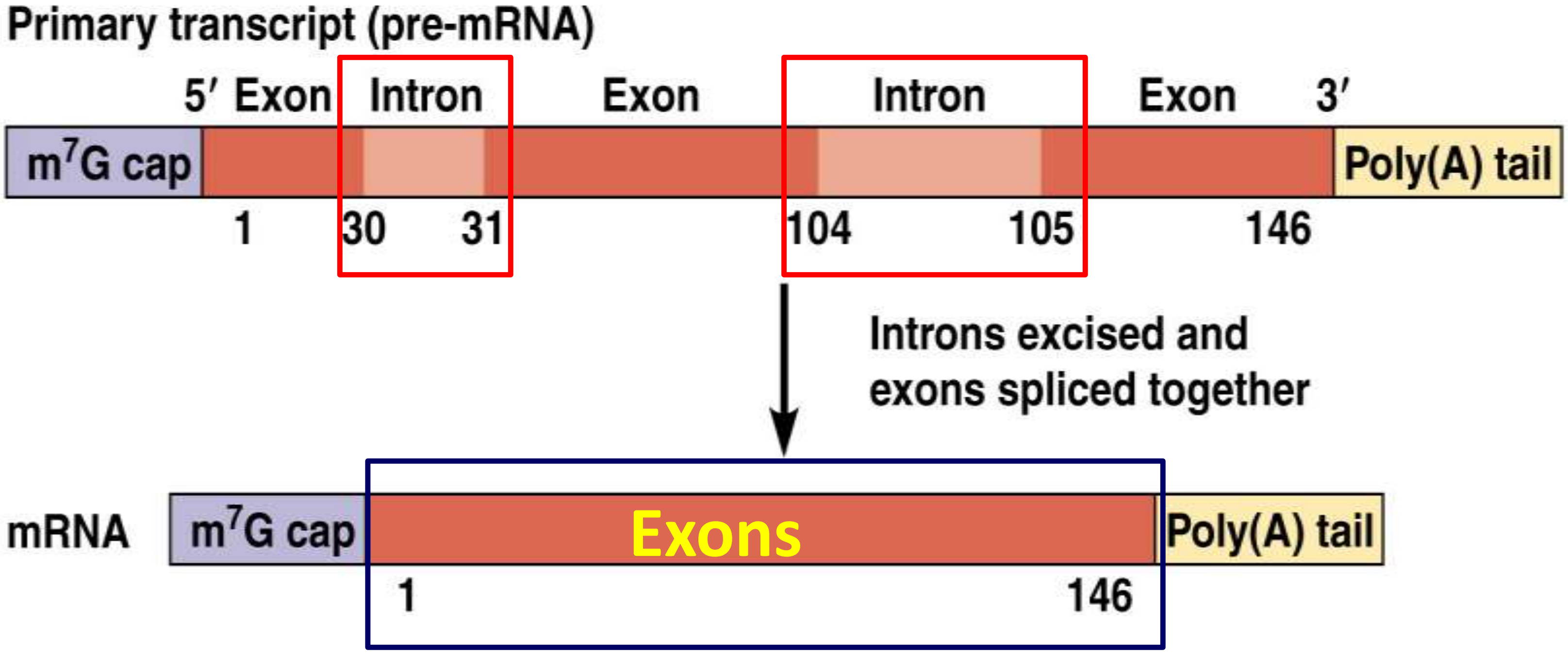
Modification of hnRNA





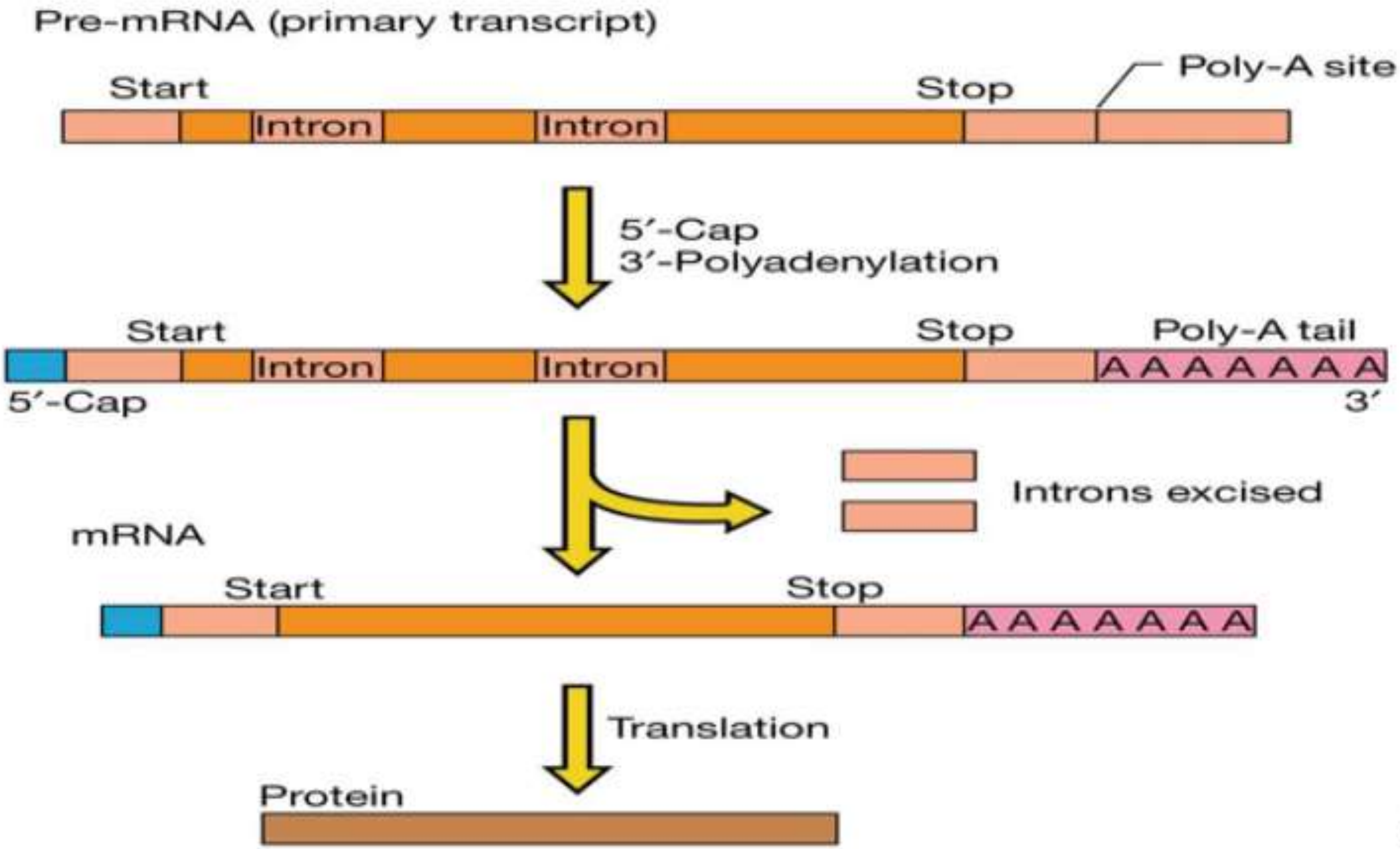
- Primary transcripts of mRNA are called as **heteronuclear RNA (hnRNA)**.
- **hnRNA** are larger than **matured mRNA** by many folds.
- Modification includes
  - Capping at the 5′- end
  - Tailing at the 3′- end
  - mRNA splicing
  - RNA editing

Post Transcriptional modifications of Pre-mRNA (or) hnRNA

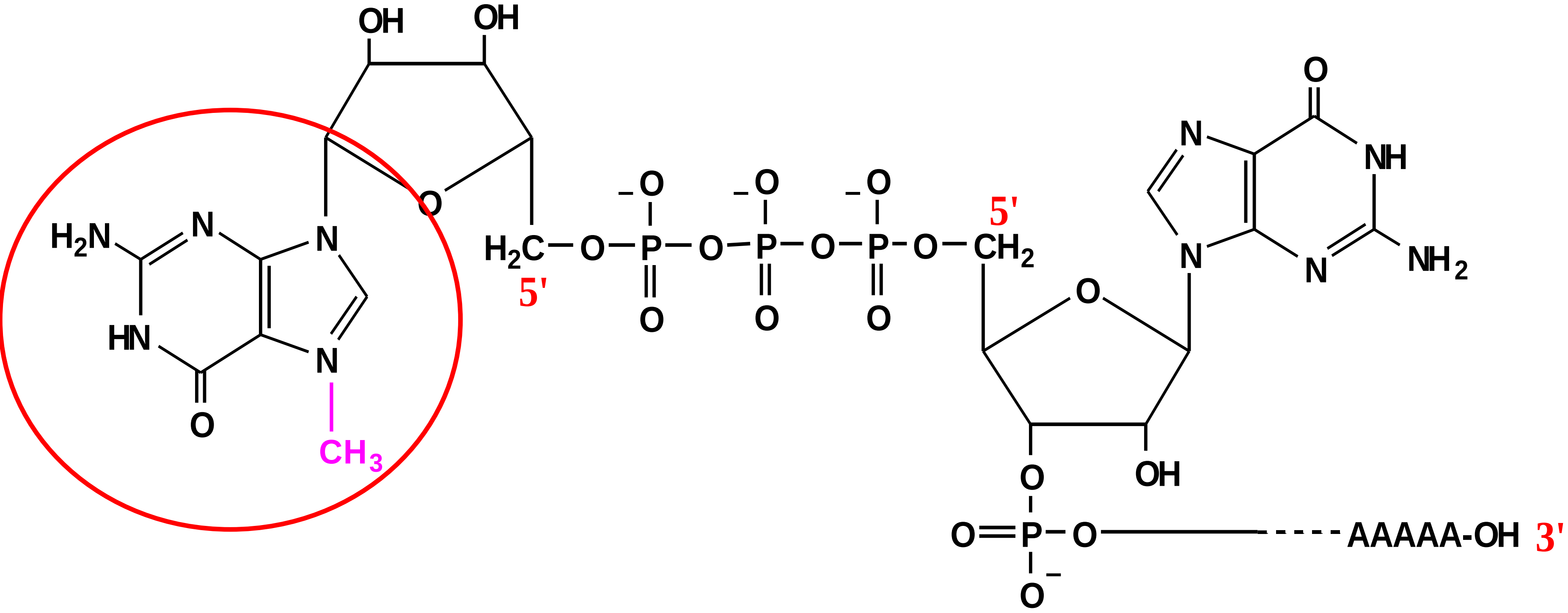


- **Introns** are removed from the primary transcript in the nucleus, **exons (coding sequences)** are ligated to form the mRNA molecule.

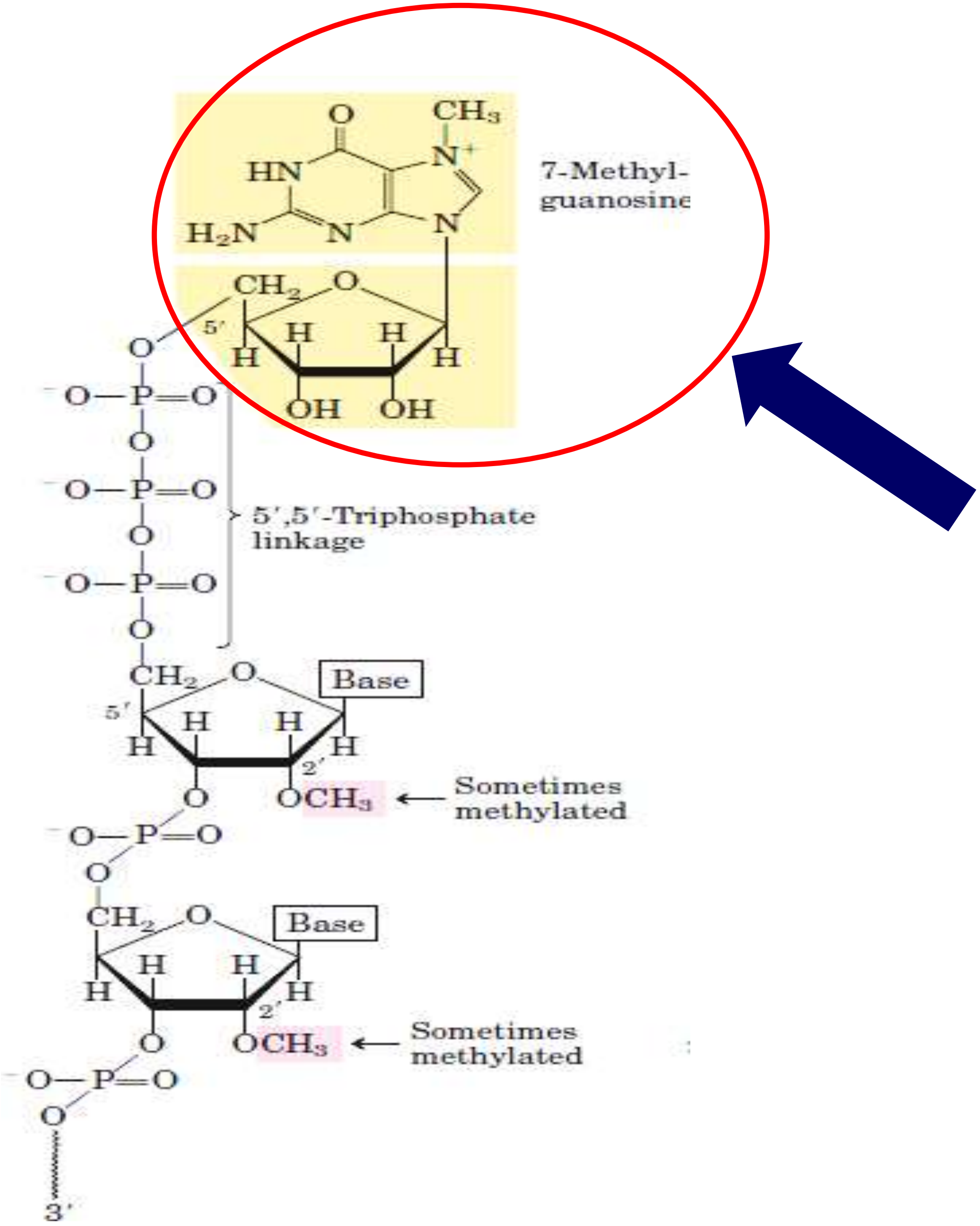
pre-mRNA maturation



Capping at the 5′- end



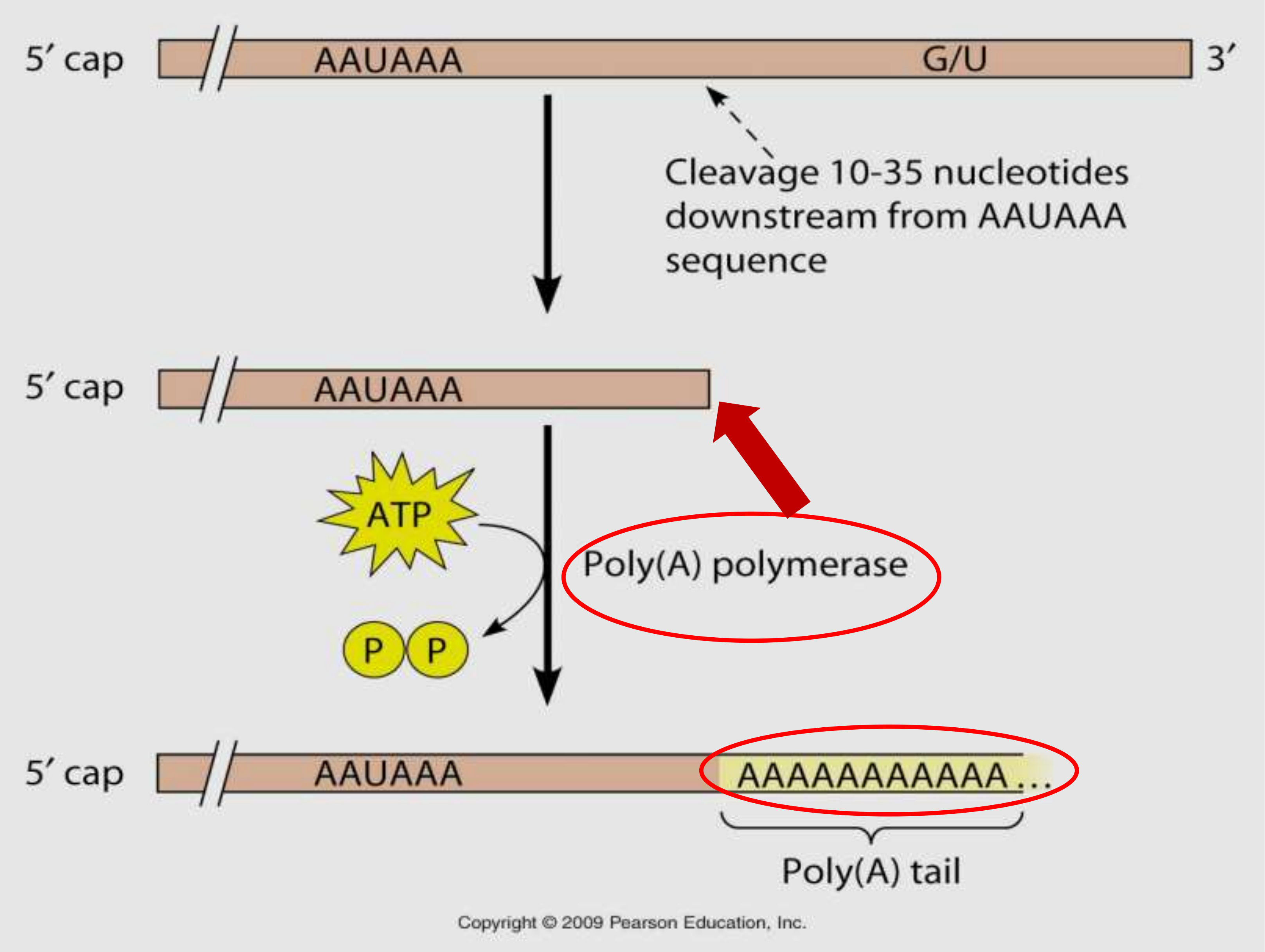
m^{7}GpppGp----


7-Methyl-guanosine
5′,5′-Triphosphate linkage
Base
Sometimes methylated
Base
Sometimes methylated
3′

5' Cap

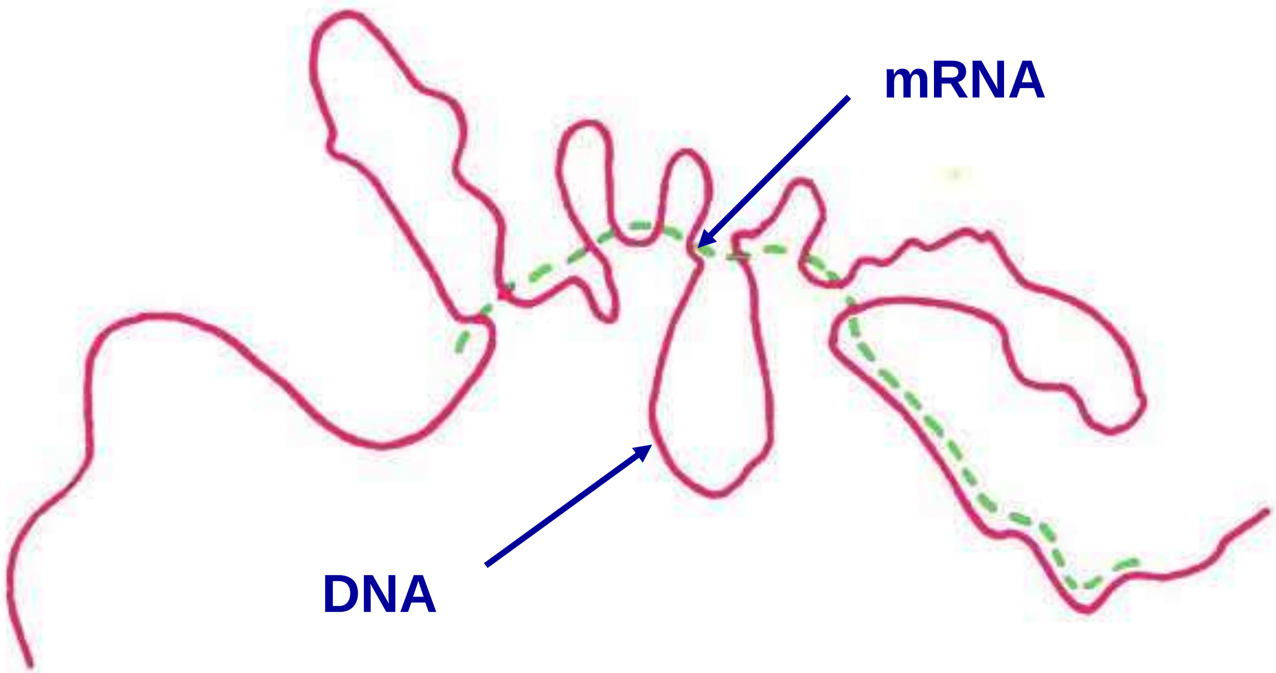
- The 5′- cap structure is found on hnRNA too. ⇒ The capping process occurs in nuclei.
- The cap structure of mRNA will be recognized by the cap-binding protein required for translation.
- The capping occurs prior to the splicing.

Poly-A tailing at 3′ - end



Copyright © 2009 Pearson Education, Inc.

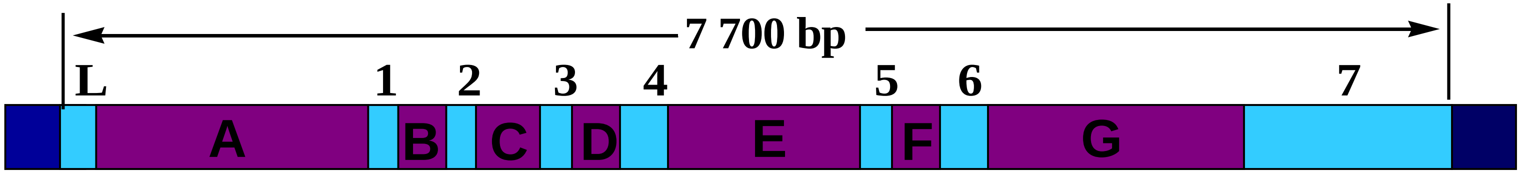
mRNA splicing



The matured mRNAs are much shorter than the DNA templates.

Split gene

The structural genes are composed of Coding (Exons) and Non-coding (Introns) regions that are alternatively separated.



A~G no-coding region

1~7 coding region

Splicing of hnRNA / pre-mRNA

- **Introns (or) intervening sequences are the RNA sequences which do not code for the proteins.**
- **Introns usually start with 5´-GU.**
- **Introns usually end with 3´-AG.**
- **RNA splicing involves the removal introns from pre-mRNA and is carried out by small nuclear complexes Spliceosomes.**

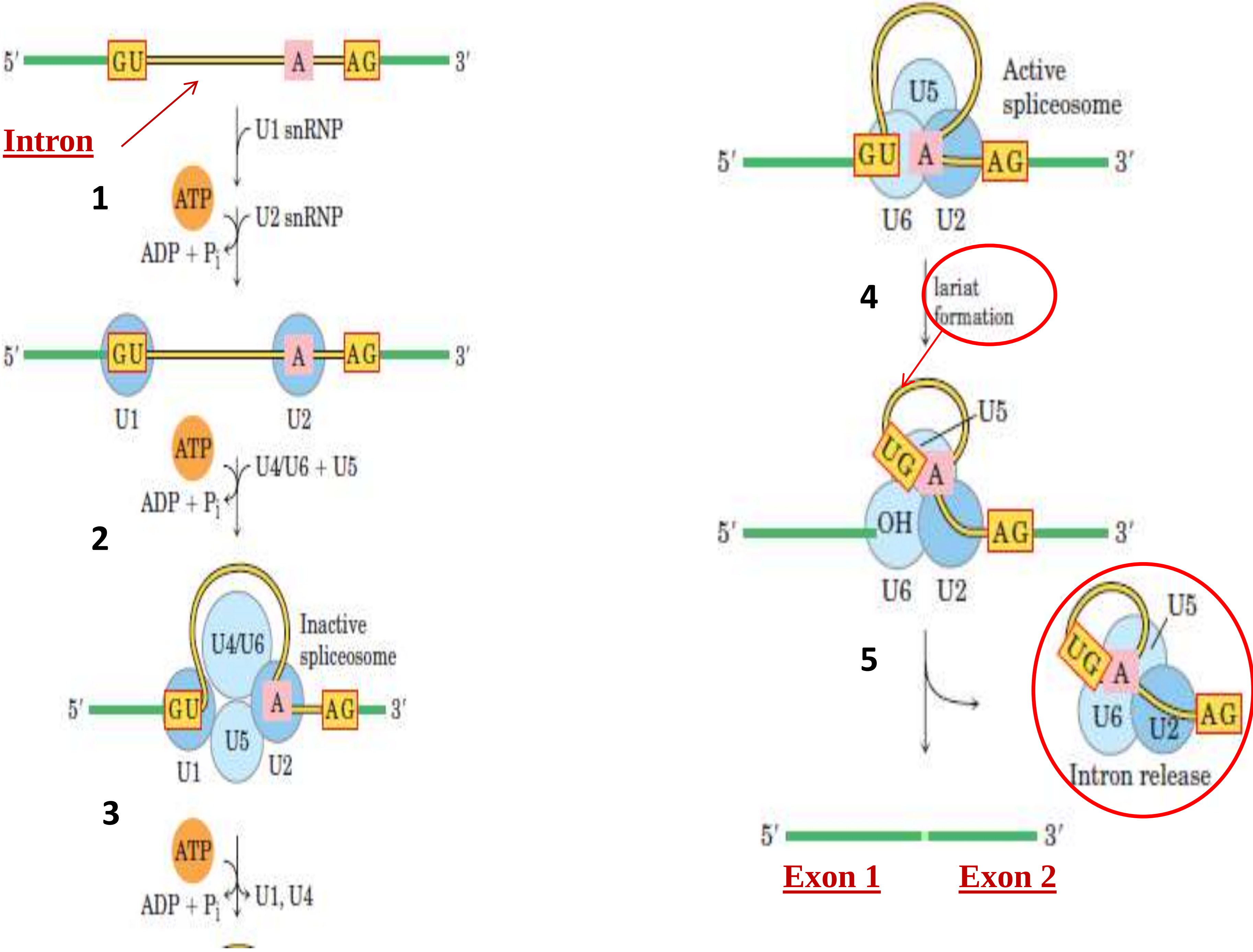
Spliceosome

- **The spliceosome is a large Protein-RNA complex in which splicing of pre-mRNAs occurs.**

- **The spliceosome is made up of specialized RNA and Protein complexes called small nuclear RiboNucleoProteins (snRNPs, often pronounced "snurps").**

- **Each snRNP contains RNAs with 100 to 200 nucleotides long, known as small nuclear RNAs (snRNAs).**

- **Five snRNAs (U1, U2, U4, U5, and U6) involved in splicing reactions are generally found in abundance in eukaryotic nuclei.**
- **Splice sites of Introns are recognized by snRNPs.**



Intron

1

2

3

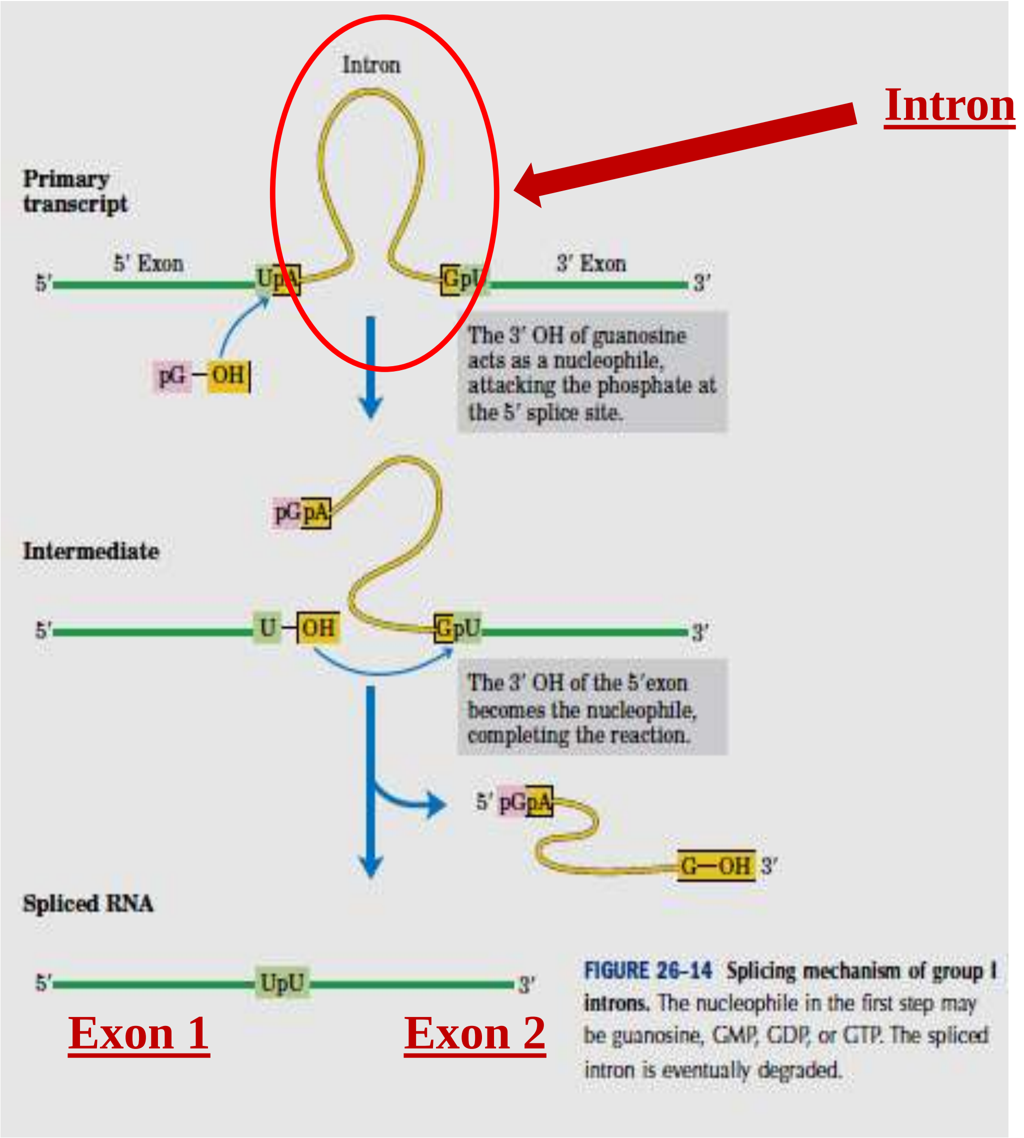
4

5

Exon 1

Exon 2

Self splicing Introns (Group –I Introns)



The 3′ OH of guanosine acts as a nucleophile, attacking the phosphate at the 5′ splice site.

The 3′ OH of the 5′exon becomes the nucleophile, completing the reaction.

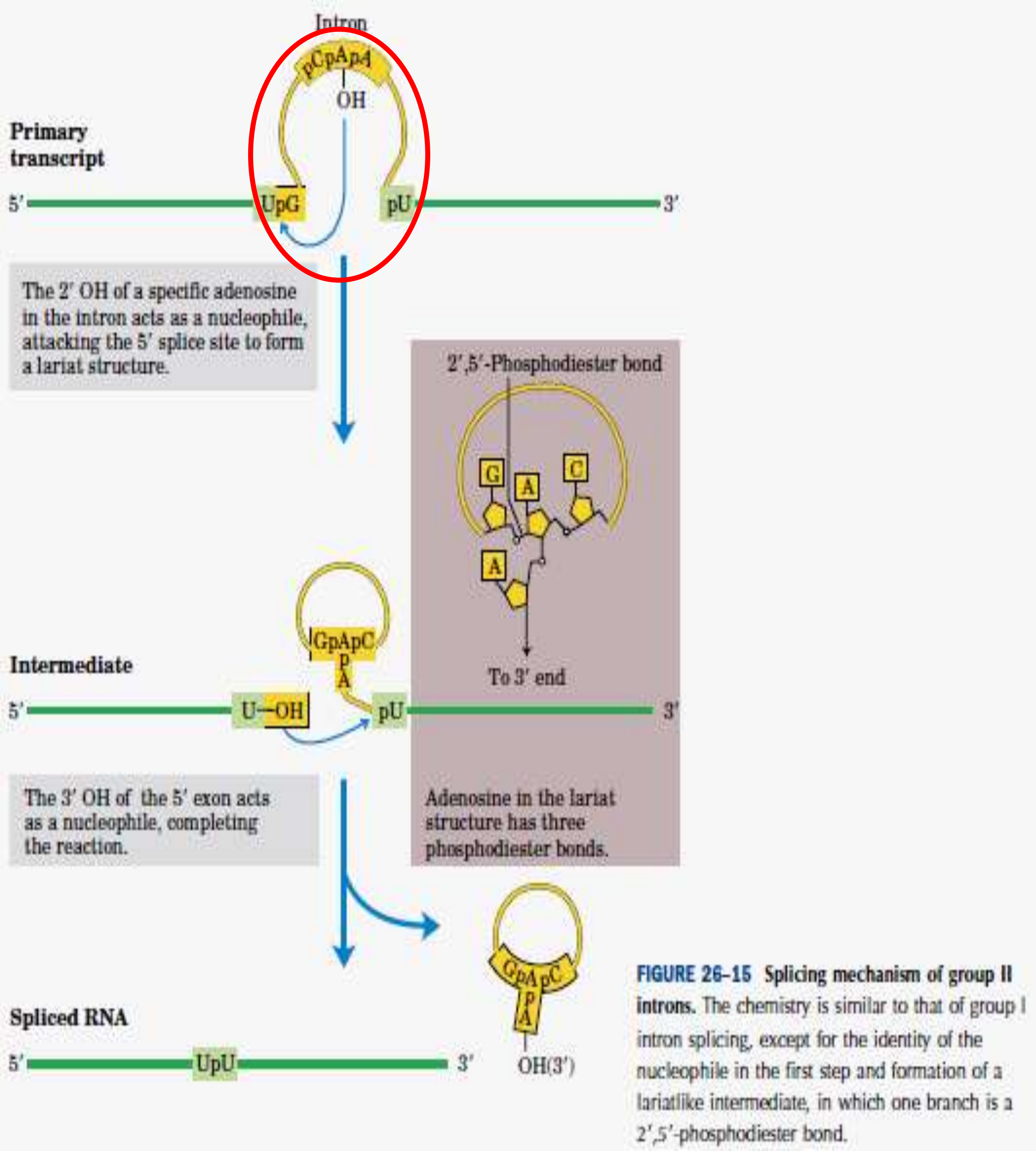
FIGURE 26–14 Splicing mechanism of group I introns. The nucleophile in the first step may be guanosine, GMP, GDP, or GTP. The spliced intron is eventually degraded.

Intron

Exon 1

Exon 2

Self splicing Introns (Group –II Introns)



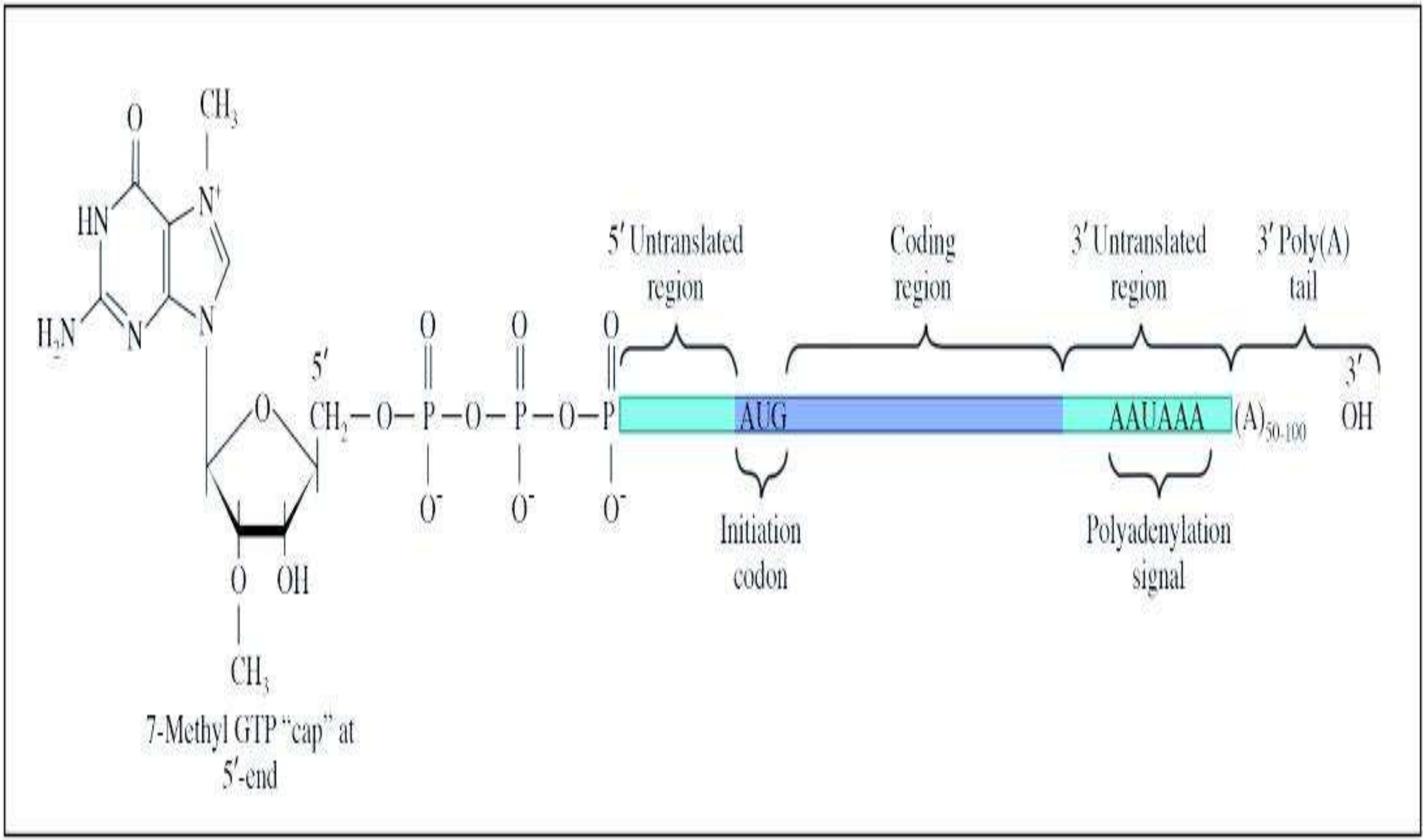
The 2′ OH of a specific adenosine in the intron acts as a nucleophile, attacking the 5′ splice site to form a lariat structure.

Adenosine in the lariat structure has three phosphodiester bonds.

The 3′ OH of the 5′ exon acts as a nucleophile, completing the reaction.

FIGURE 26–15 Splicing mechanism of group II introns. The chemistry is similar to that of group I intron splicing, except for the identity of the nucleophile in the first step and formation of a lariatlike intermediate, in which one branch is a 2′,5′-phosphodiester bond.

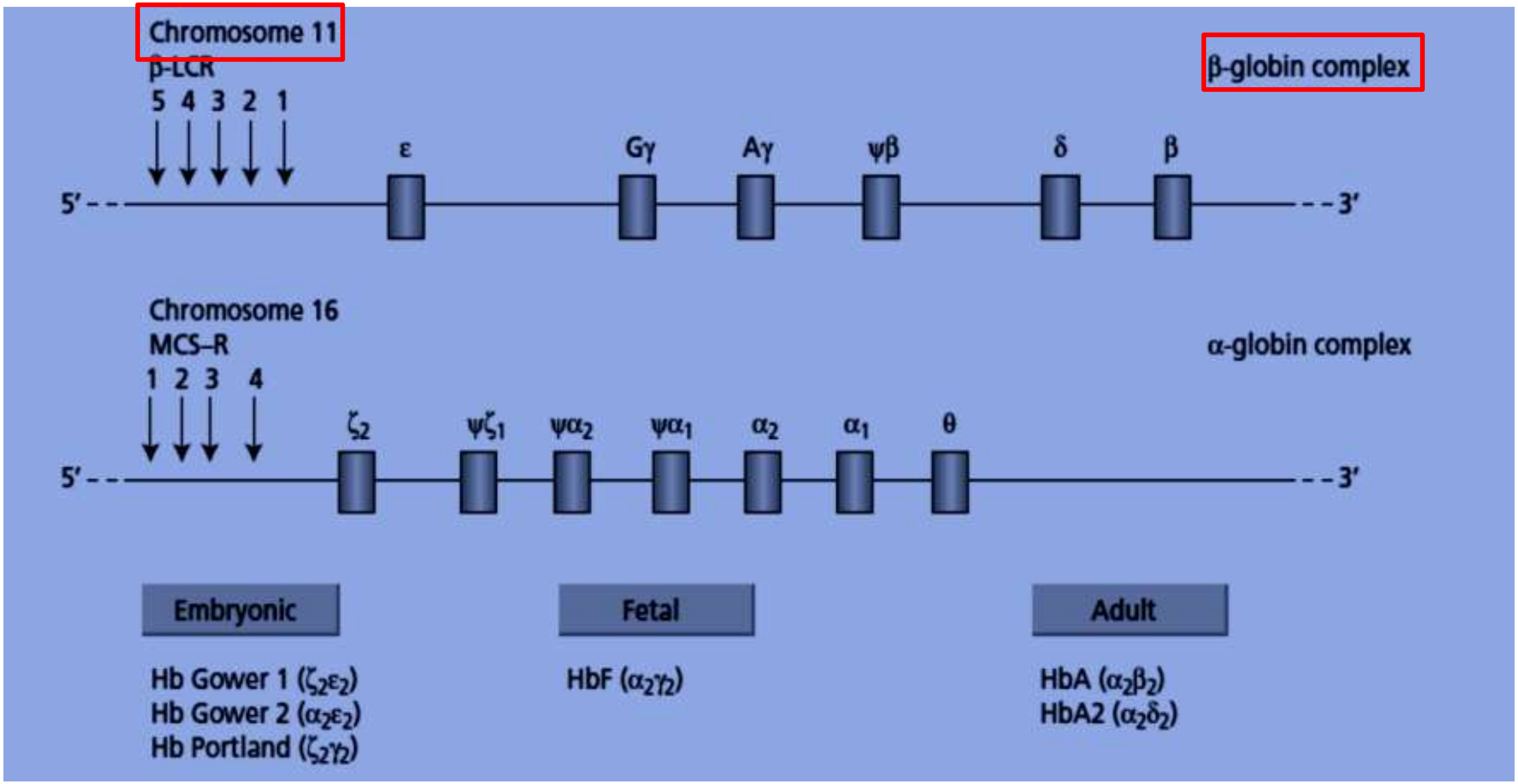
Eukaryotic m-RNA after modifications



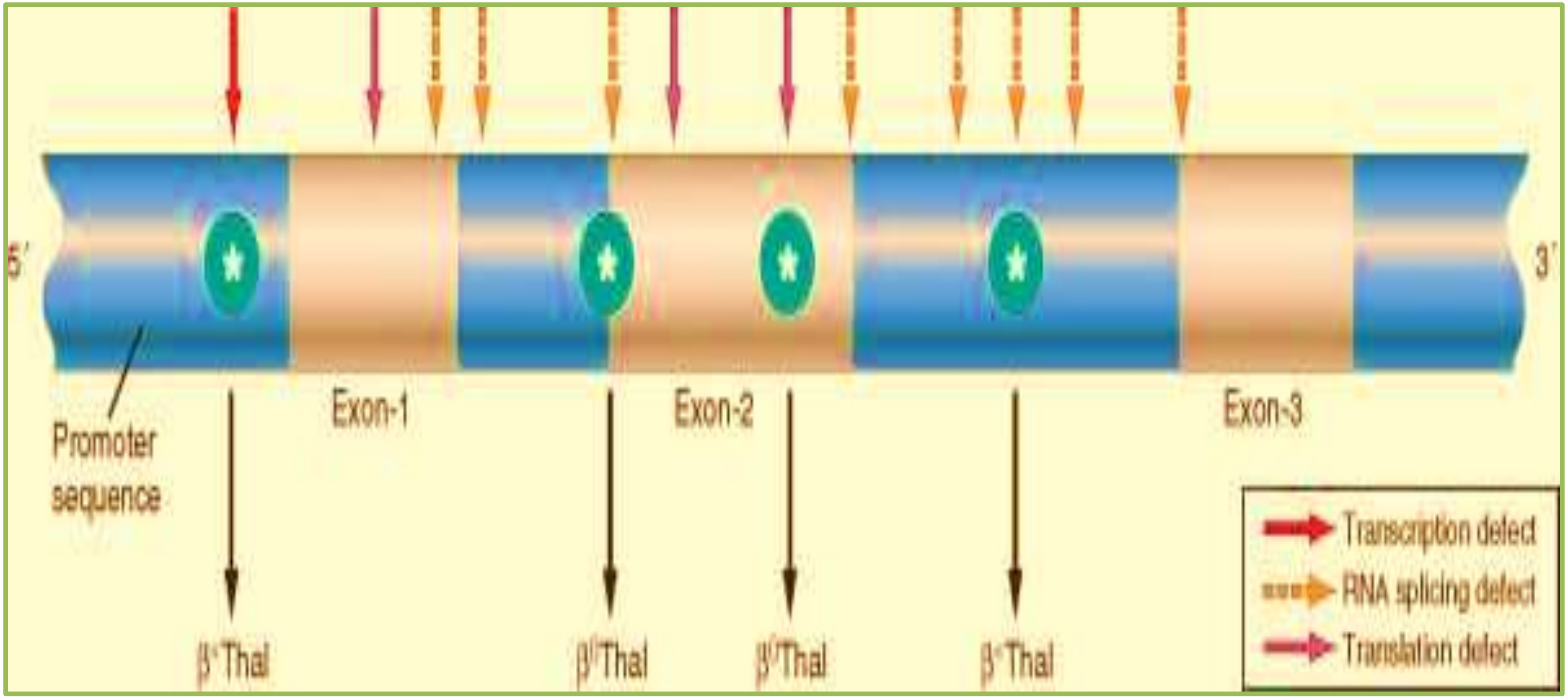
Splice site mutations

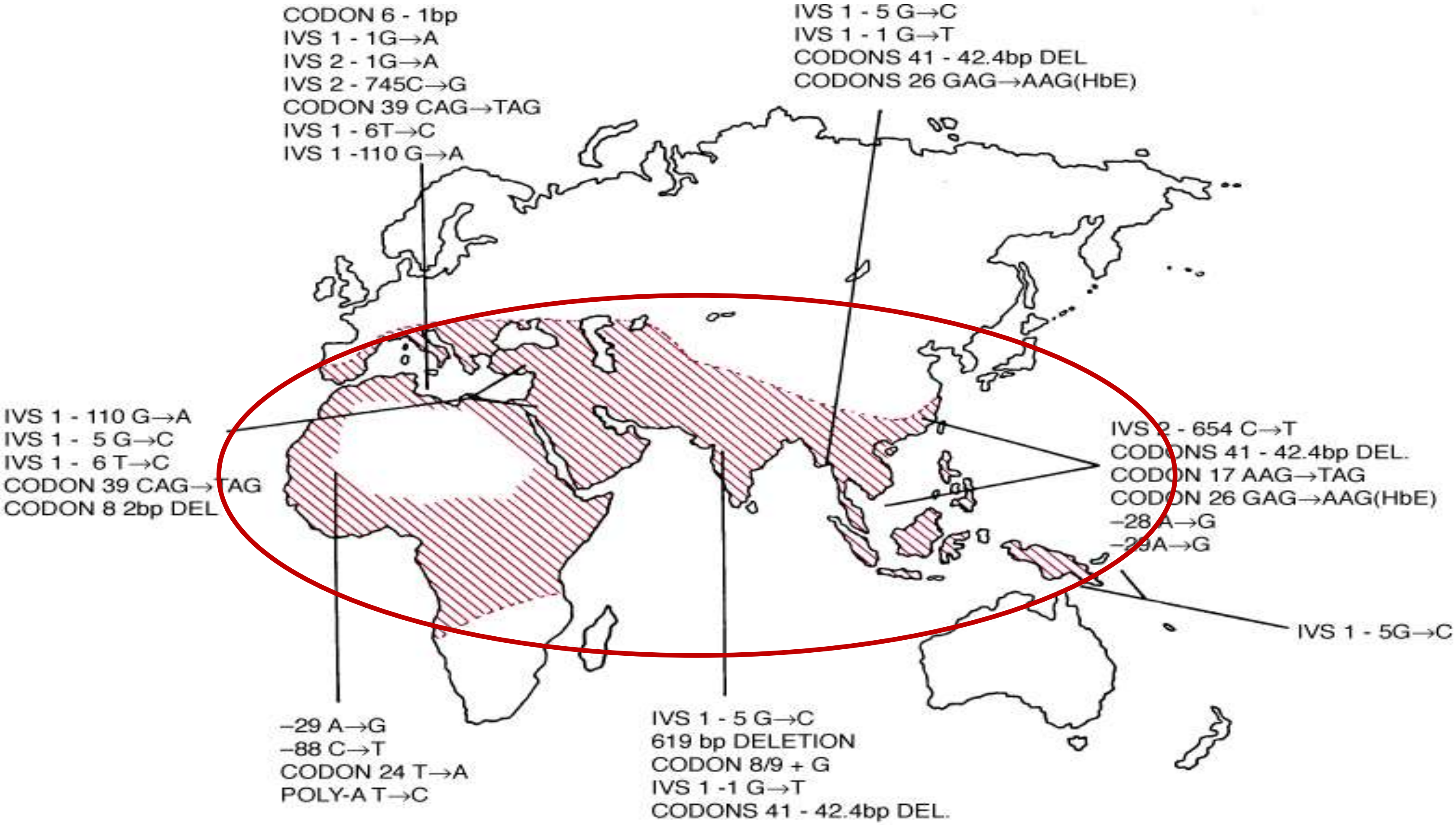
- Mutation at splice sites can lead to improper splicing and production of abberant proteins
- Eg: β - thalassemia
- β-subunit of hemoglobin is not formed in sufficient amount.
- It results from point mutation in β-globin gene where the G→A mutation occurs.
- This creates a new splice acceptor site nineteen nucleotides upstream from the normal splice acceptor
- A faulty beta-globin protein is made, leading to severe anemia.

Location of Globin genes



Distribution of β-globin gene mutations associated with β-thalassemia.





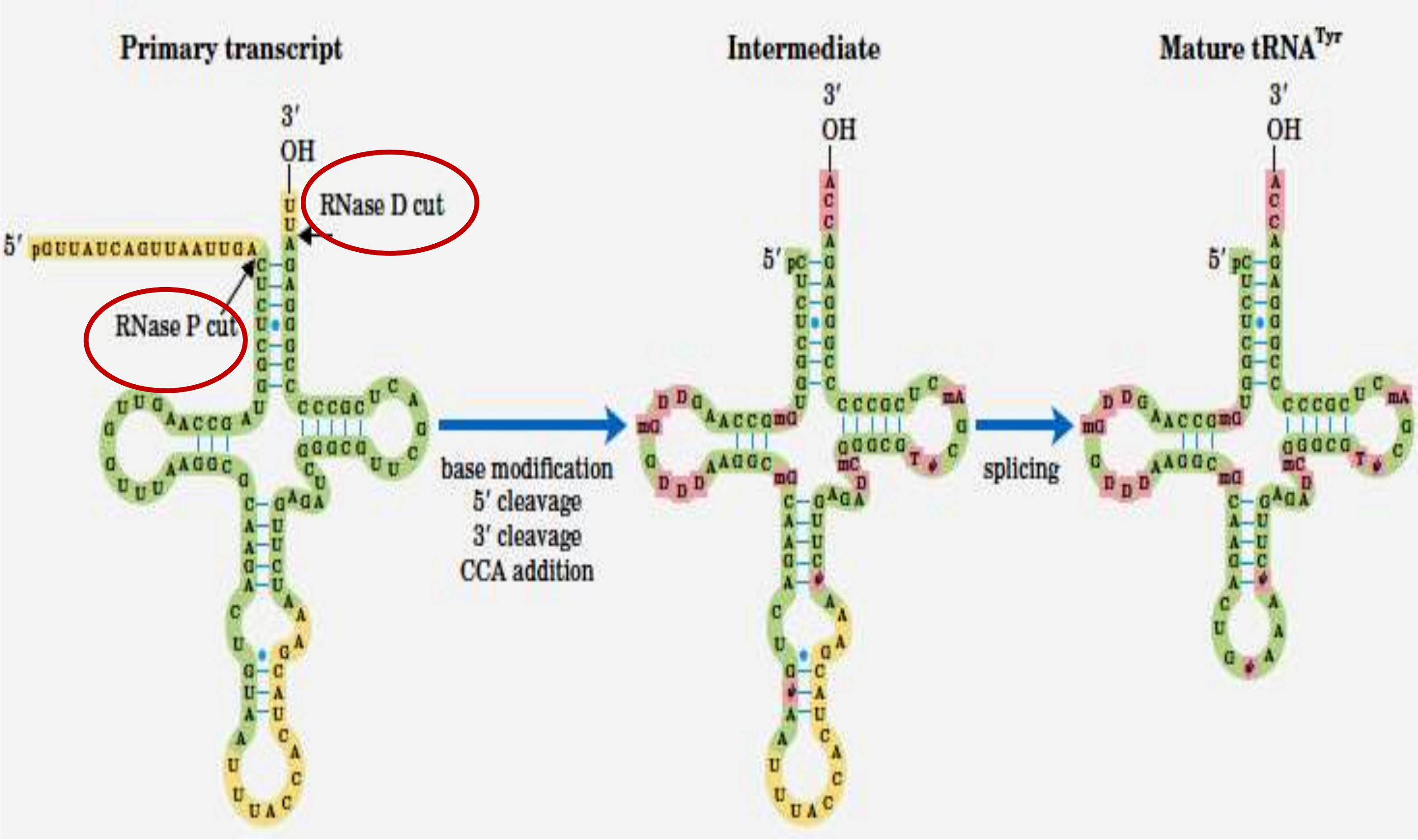
Source: Lichtman MA, Kipps TJ, Seligsohn U, Kaushansky K, Prchal JT: *Williams Hematology, 8th Edition*: http://www.accessmedicine.com

World distribution of **β-thalassemia**

Clinical syndromes in β-THALASSEMIAS

β-Thalassemia major	**Severe; requires blood transfusions**
β-Thalassemia intermedia	**Severe but does not require regular blood transfusions**
β-Thalassemia minor	**Asymptomatic with mild or absent anemia; red cell abnormalities seen**

Modification of tRNA

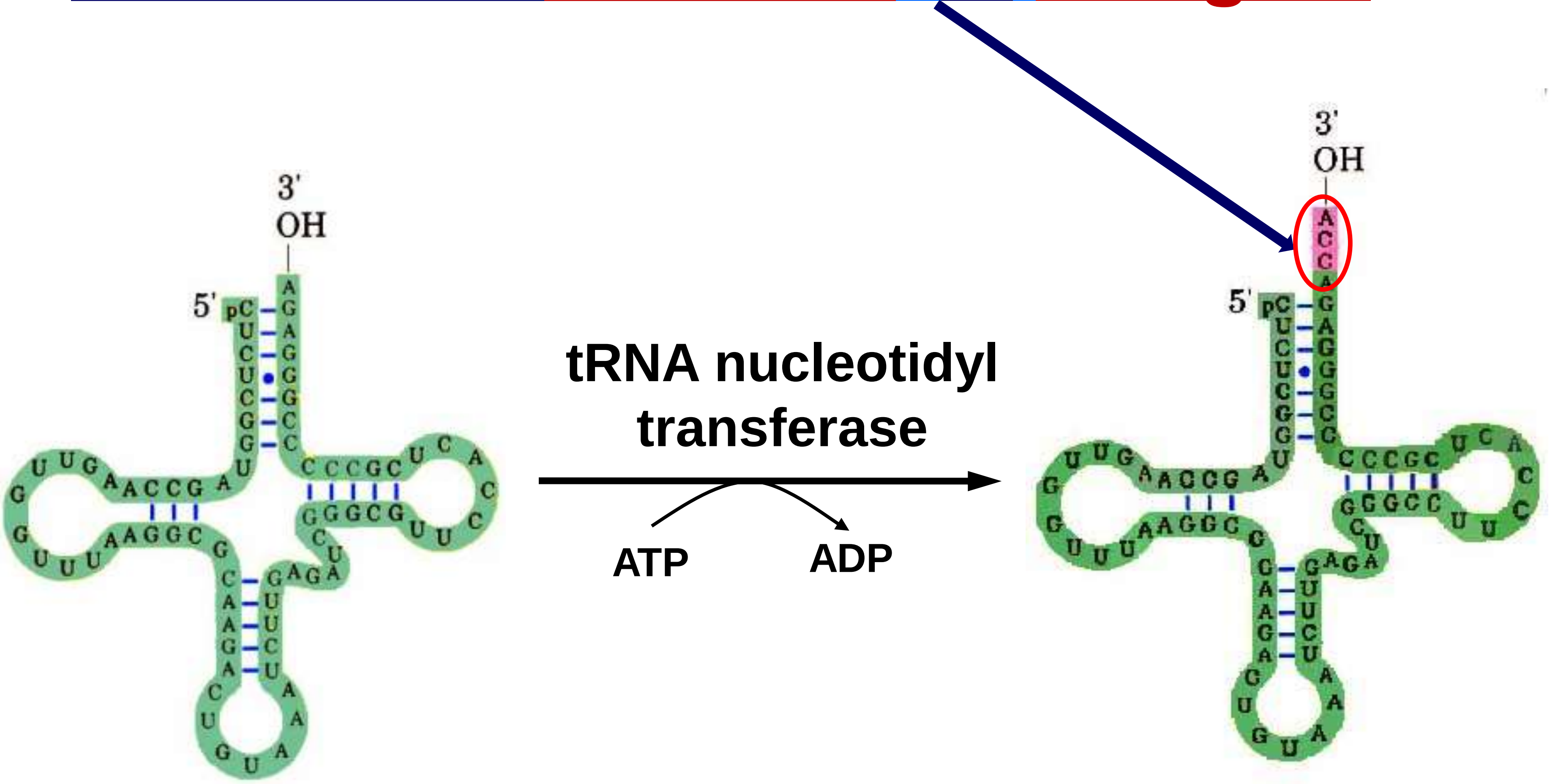


Endo- and exonucleases to generate ends of tRNA

- Endonuclease **RNase P** cleaves to generate the **5´ end**.

- Exonuclease **RNase D** trims 3′ to 5′, leaving the mature **3´ end**.

Addition of -CCA-OH at 3′ region



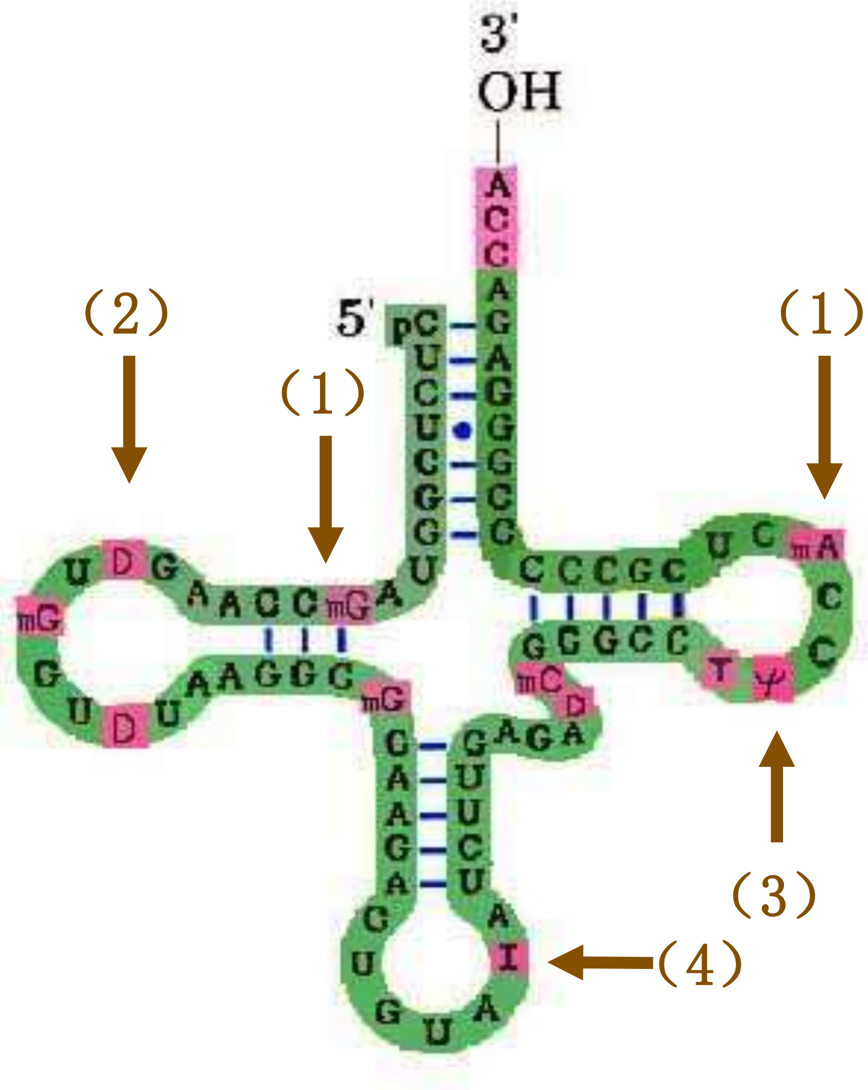
tRNA nucleotidyl transferase

ATP

ADP

At 3´-CCA region of tRNA an activated Amino acid will be attached during Protein Synthesis.

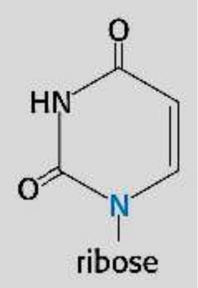
Base modifications



1. **Methylation**
   A→mA, G→mG
2. **Reduction**
   U→DHU
3. **Transversion**
   U→ψ
4. **Deamination**
   A→I

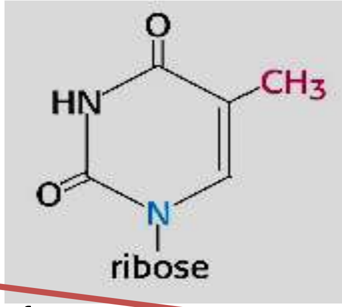
Modification of some bases

uridine

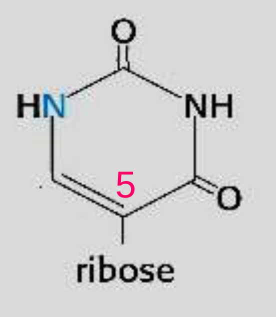


methylation

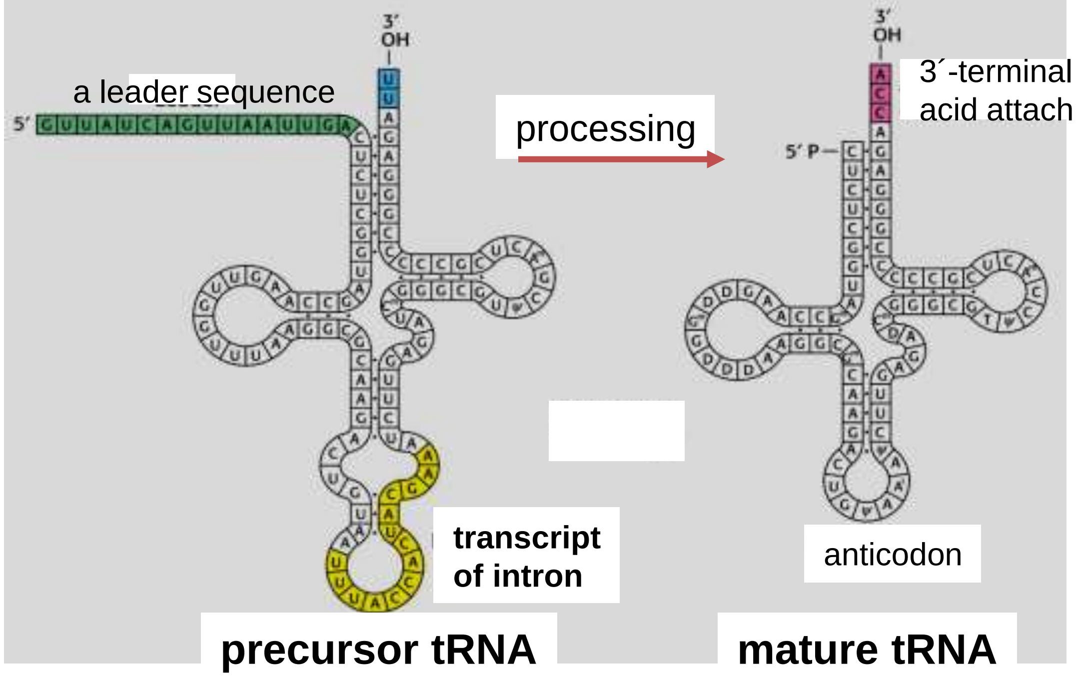
transformation of
the linkage to ribosyl



ribosyl thymine



pseudouridine (φ)



a leader sequence

processing

3´-terminal UU replaced by amino acid attachment site CCA-3´-OH

transcript of intron

anticodon

precursor tRNA

mature tRNA

Modification of rRNA

- 45S Pre-rRNA transcript in nucleus is the precursor of 3 kinds of rRNAs.
- The matured rRNA will be assembled with ribosomal proteins to form ribosomes that are exported to cytosolic space.