Transcription

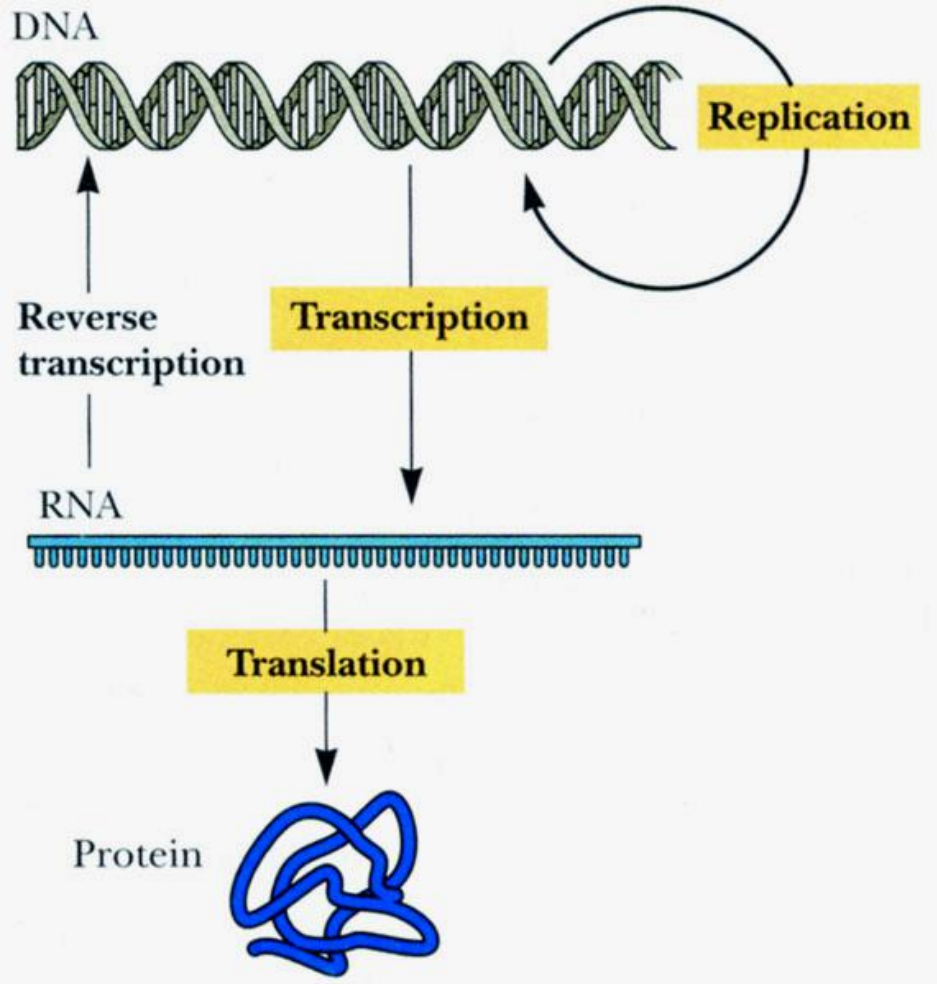
from DNA to RNA

RNA

- Difference between RNA & DNA
- Three major kinds of RNA: **Messenger RNAs (mRNAs), Transfer RNAs (tRNAs), Ribosomal RNAs (rRNAs)**
- Other RNAs: small nucleolar RNAs (snoRNAs), microRNAs (miRNAs), short interfering RNAs (siRNAs), long non-coding RNAs (lncRNA), etc.
- RNA is the only macromolecule known to have a role both in the storage and transmission of information and in catalysis, **ribozyme**.

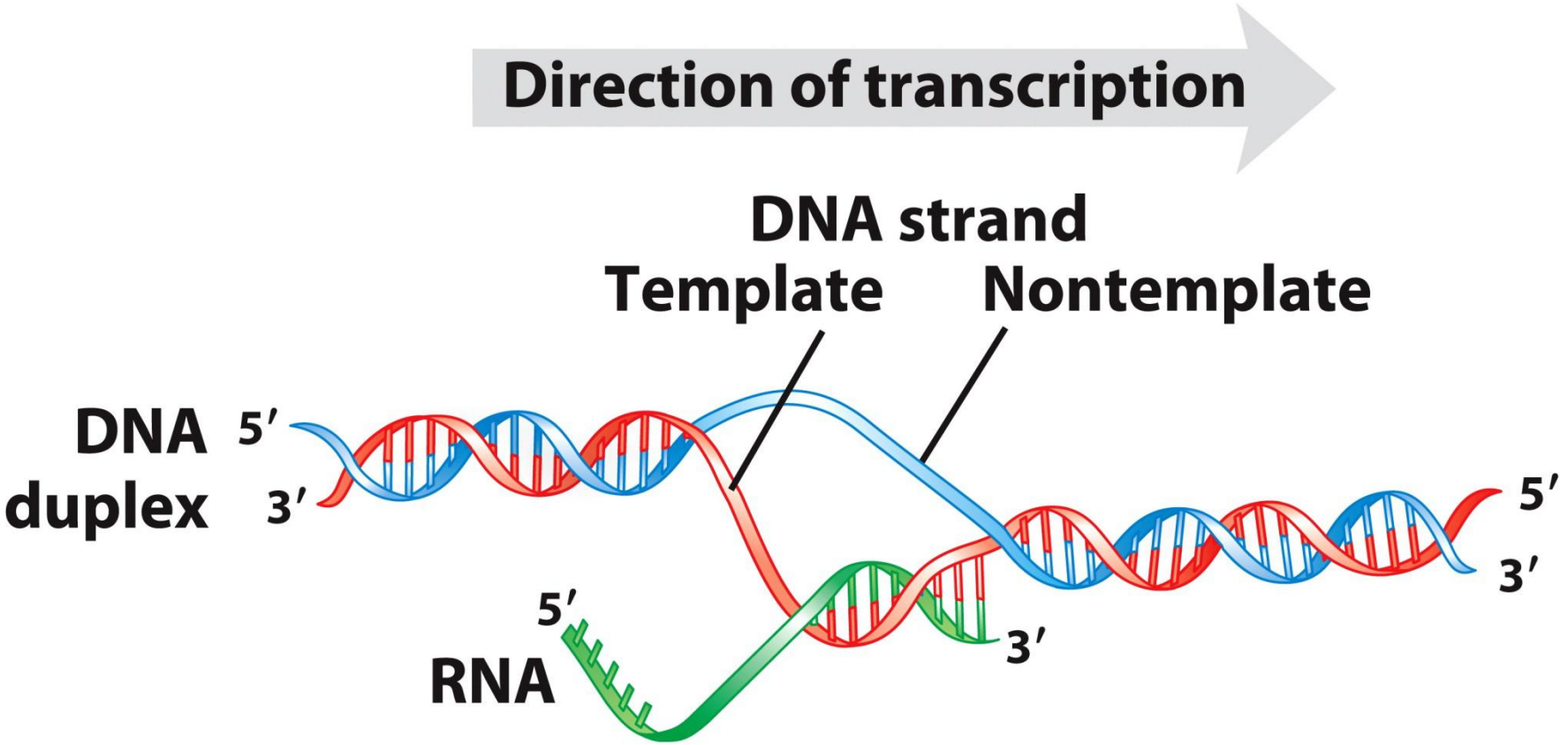
Transcription: From DNA to mRNA

The enzymatic process whereby the genetic information contained in one strand of DNA is used to specify a complementary sequence of bases in an mRNA (messenger RNA) chain.

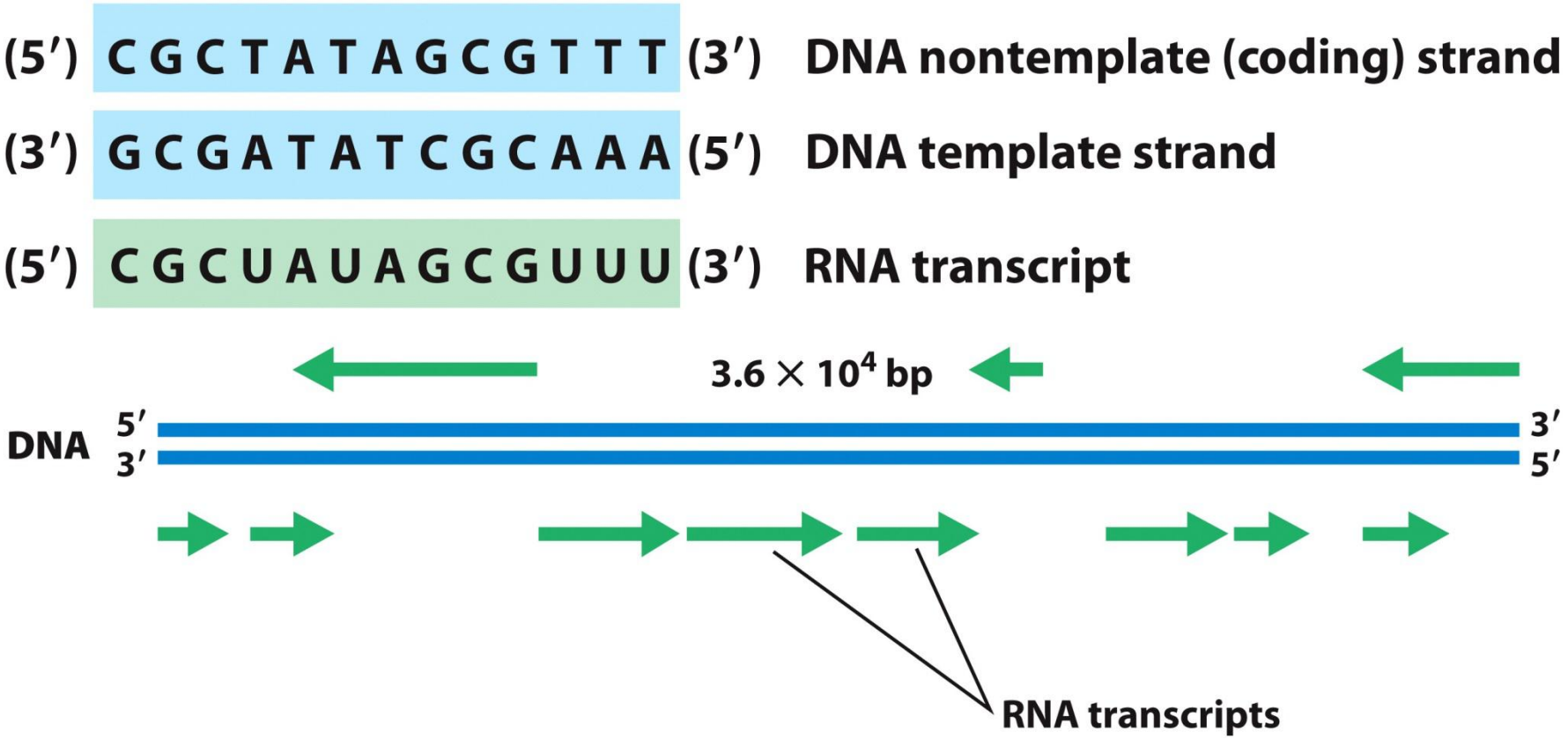


Central dogma:
DNA→RNA→Protein

One of the two strands of DNA serves as template for certain RNA synthesis

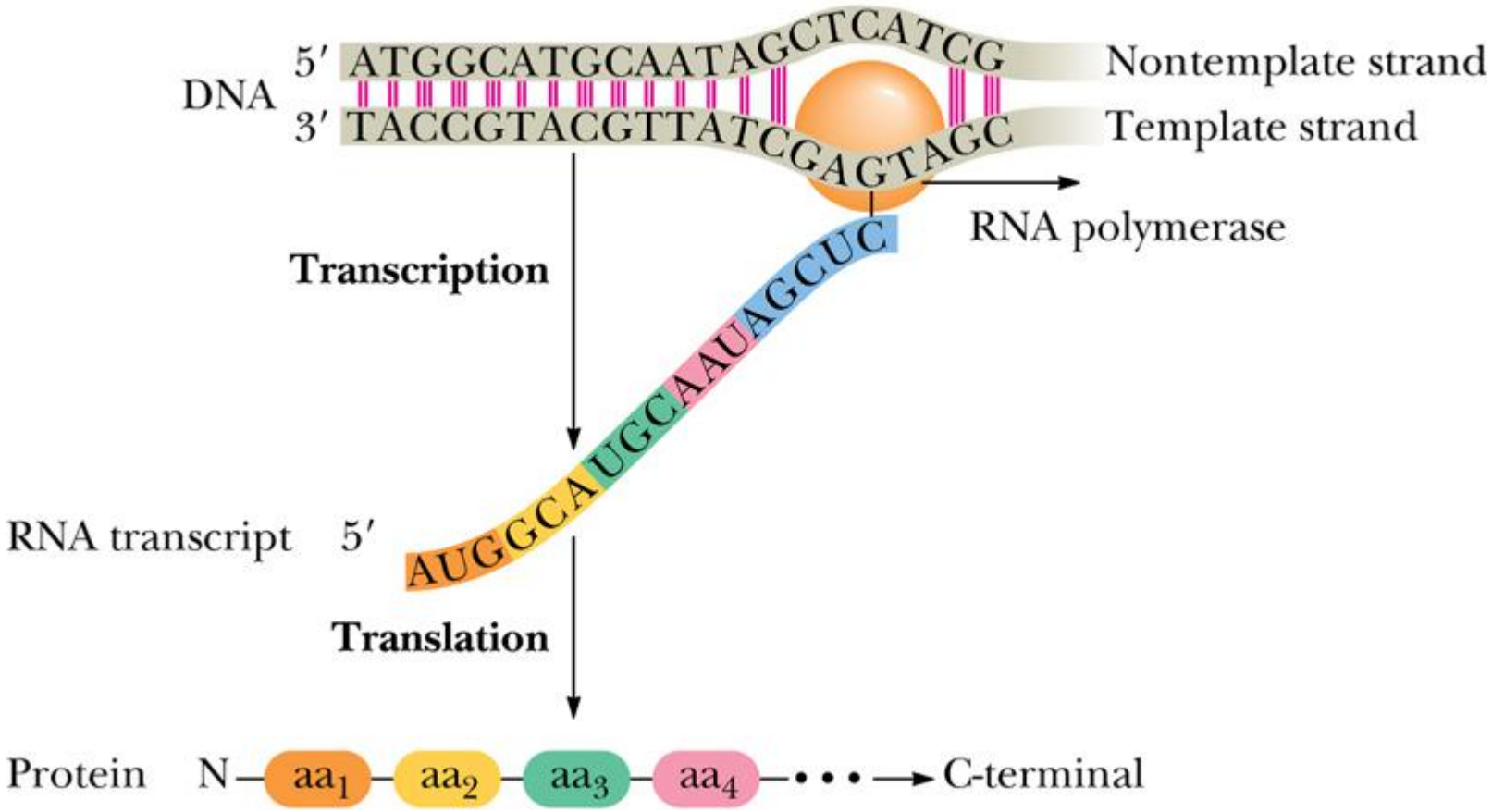


Template and nontemplate (coding) DNA strands



The coding strand for a particular gene may be located in either strand of a given chromosome.

DNA sequence to protein sequence



Transcription in prokaryotes

The phases of transcription in *E. coli*

Binding of RNA polymerase holoenzyme at promoter sites

Initiation of polymerization

Chain elongation

Chain termination

RNA polymerases in prokaryotes

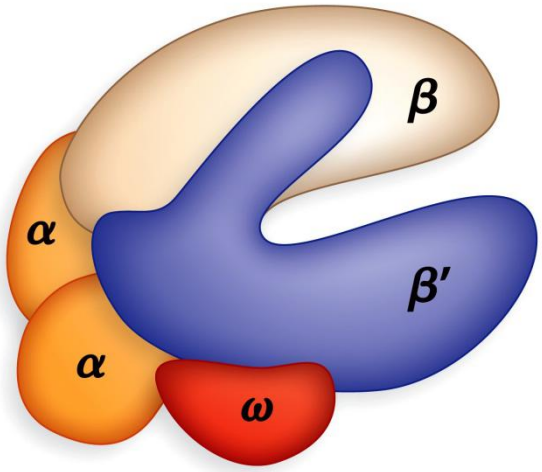
In prokaryotes, there is only one kind of DNA-dependent RNA polymerase.

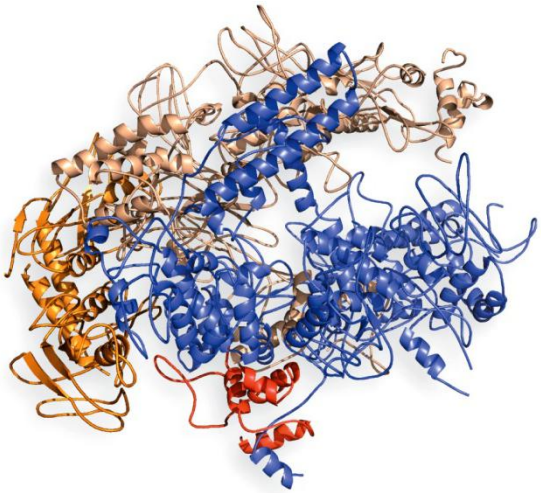
E. Coli RNA polymerase

5 core subunits

$A_2BB'\omega$

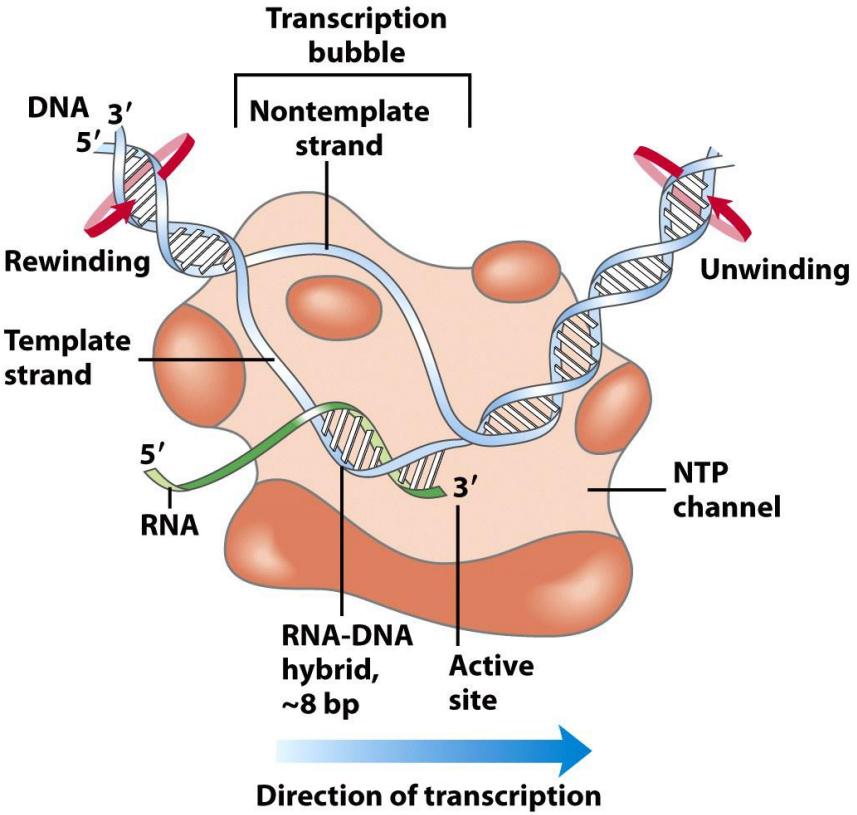
6th subunit: σ



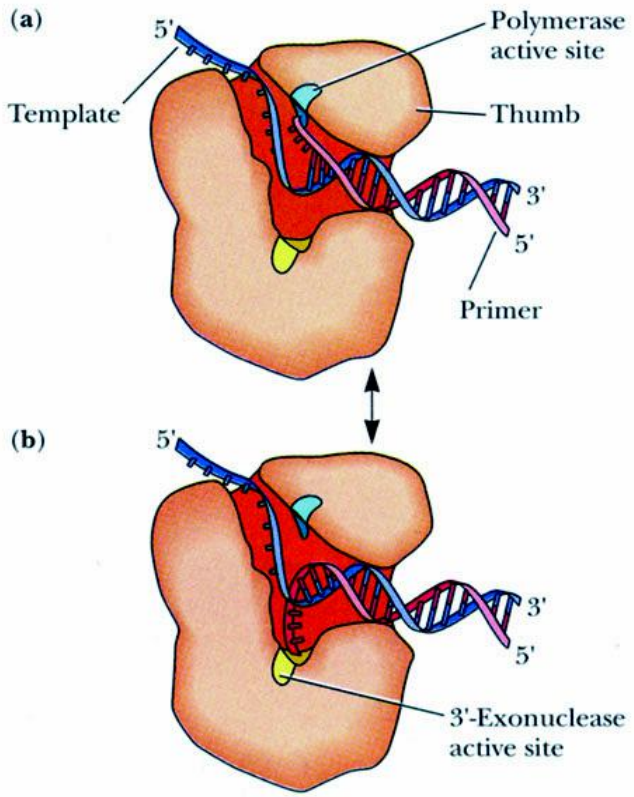


size aa	size (Kd)	gene	function
329	36511	rpoA [α]	required for assembly of the enzyme; interacts with some regulatory proteins; also involved in catalysis
1342	150616	rpoB [β]	involved in catalysis: chain initiation and elongation
1407	155159	rpoC [β′]	binds to the DNA template
613	70263	rpoD [σ]	directs enzyme to the promoter
91	10237	rpoZ [ω]	required to restore denatured RNA polymerase in vitro to its fully functional form

Transcription by RNA polymerase in *E. coli*



RNA polymerase



DNA polymerase

RNA is synthesized by RNA polymerases:

Only one of the two DNA strands serves as a template.

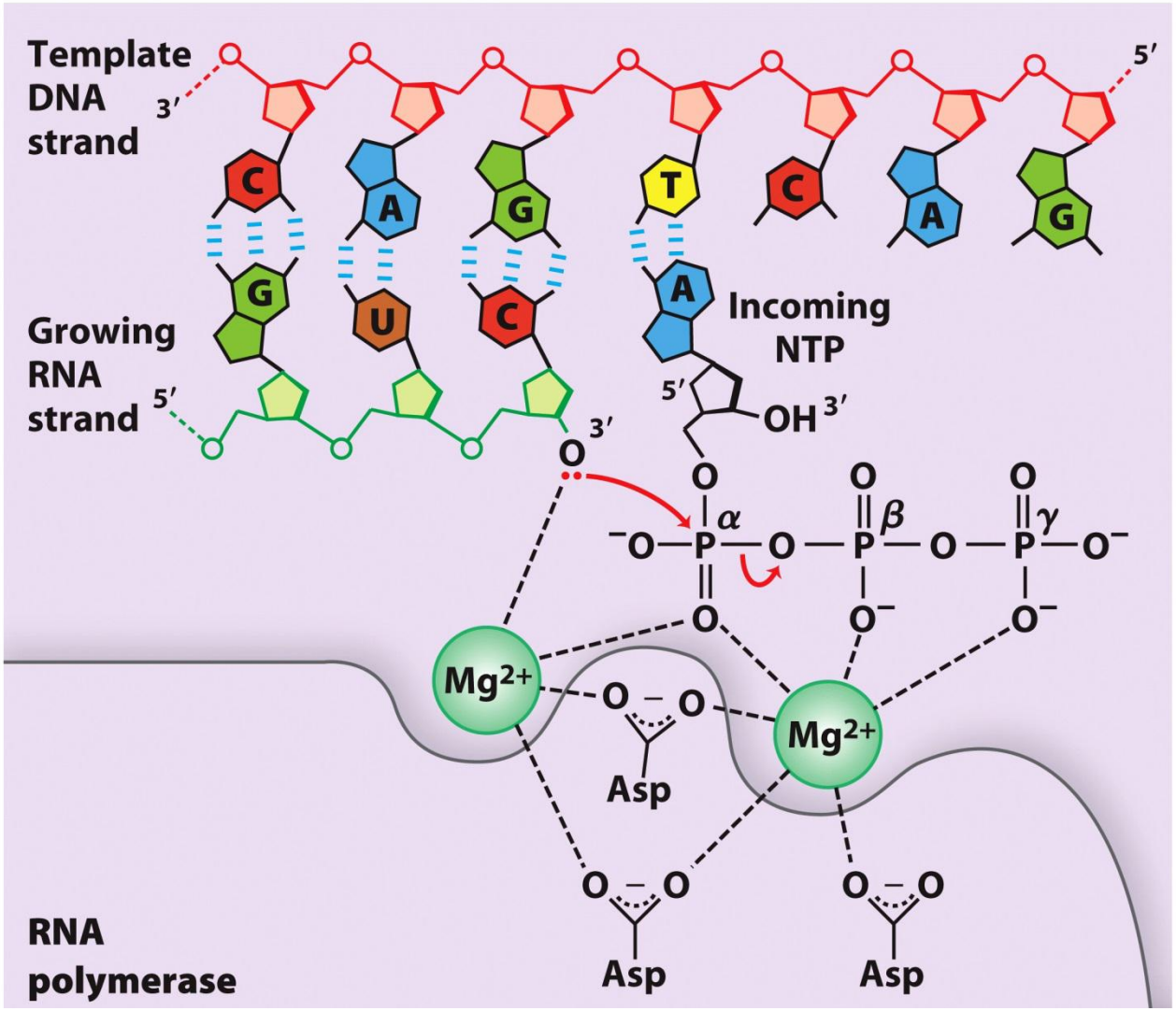
RNA polymerase **does not** require a primer to initiate synthesis.

RNA polymerases lack a separate proofreading 3' →5' exonuclease active site.

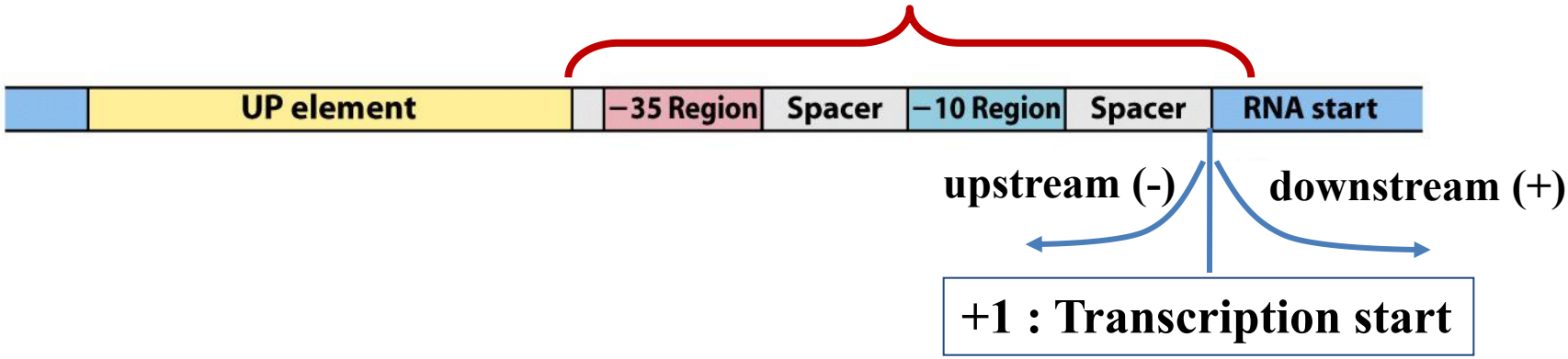
RNA pol: error rate $1/10^4$ to 10^5 ; DNA pol: $2/10^9$

Mechanism of RNA synthesis by RNA polymerase

Incoming NTP is attacked at the α phosphate by the 3’ hydroxyl of the growing RNA chain



Promoter



Promoters: specific sequences in the **DNA that the RNA polymerase holoenzyme can bind and** direct the transcription of adjacent segments of DNA (genes).

The promoter region generally is between positions -70 and +30.

§ Subunits in *E. coli*

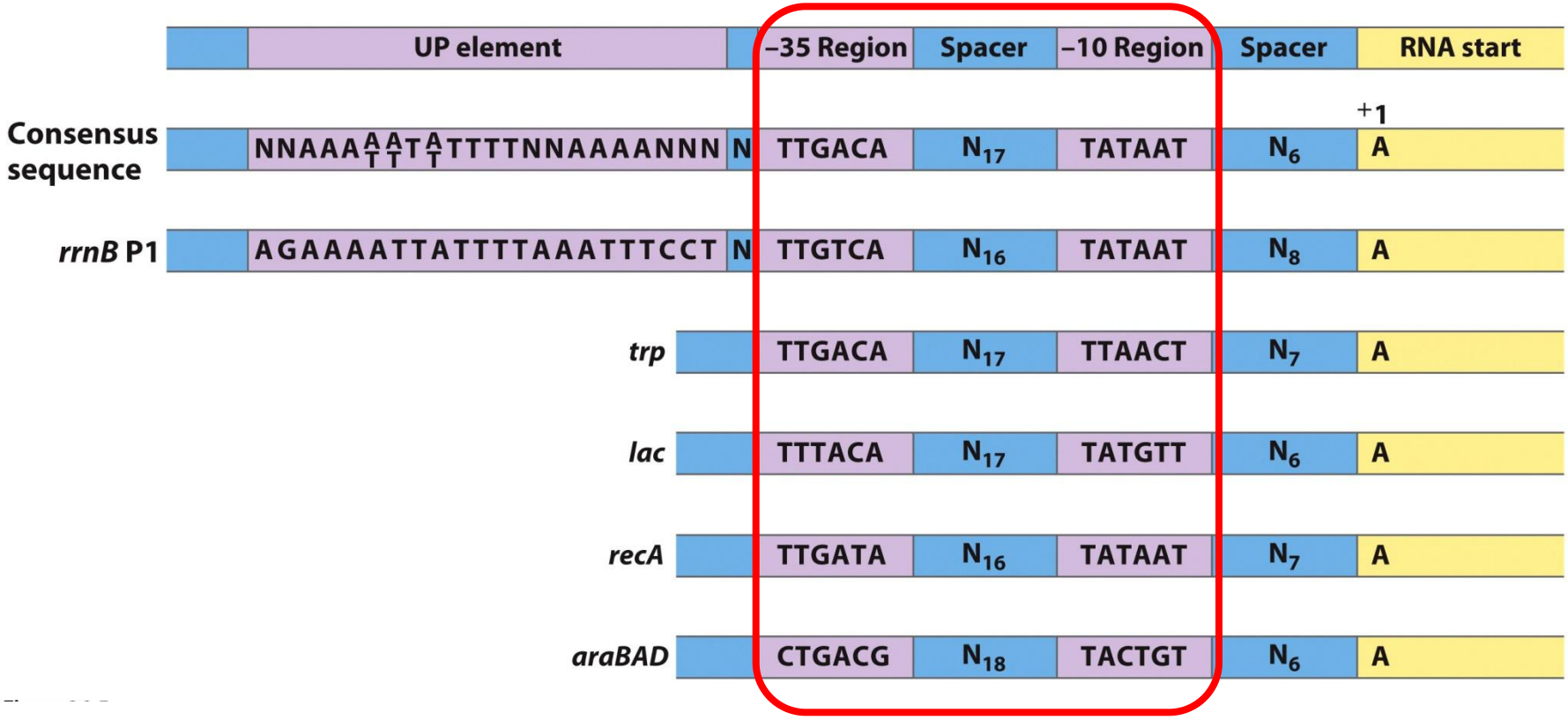
TABLE 26–1 The Seven σ Subunits of *Escherichia coli*

σ subunit	K_d (nM)	Molecules/cell*	Holoenzyme ratio (%)*	Function
σ^{70}	0.26	700	78	Housekeeping
σ^{54}	0.30	110	8	Modulation of cellular nitrogen levels
σ^{38}	4.26	<1	0	Stationary phase genes
σ^{32}	1.24	<10	0	Heat shock genes
σ^{28}	0.74	370	14	Flagella and chemotaxis genes
σ^{24}	2.43	<10	0	Extracytoplasmic functions; some heat shock functions
σ^{18}	1.73	<1	0	Extracytoplasmic functions, including ferric citrate transport

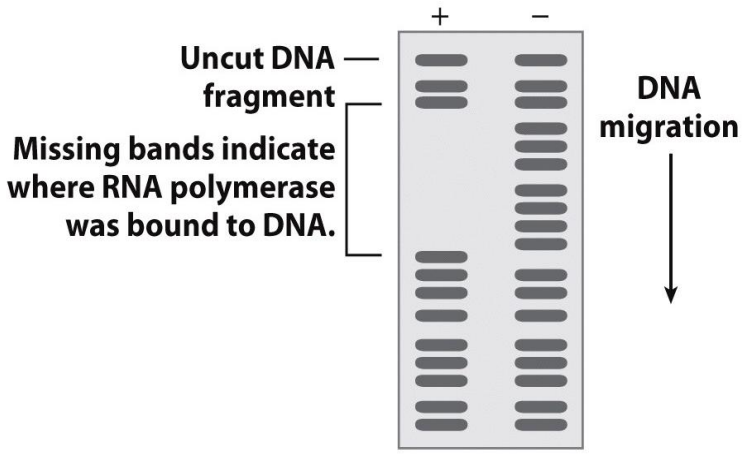
Different § subunits can recognize different promoters.

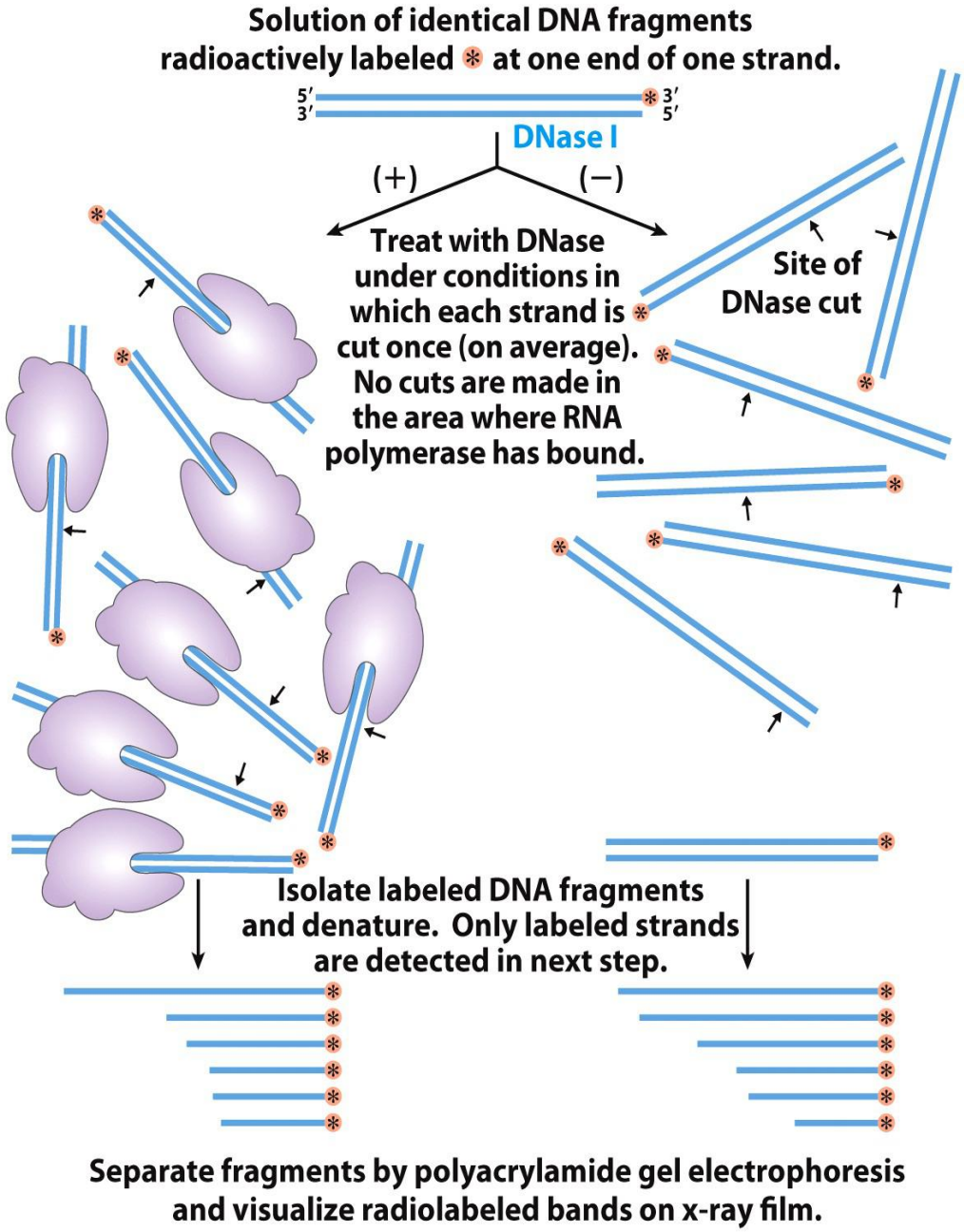
σ^{70} is the most common σ subunit in *E.coli*.

Consensus sequence recognized by *E. coli* σ^{70}



Find Promoters: DNA footprinting to identify protein binding site



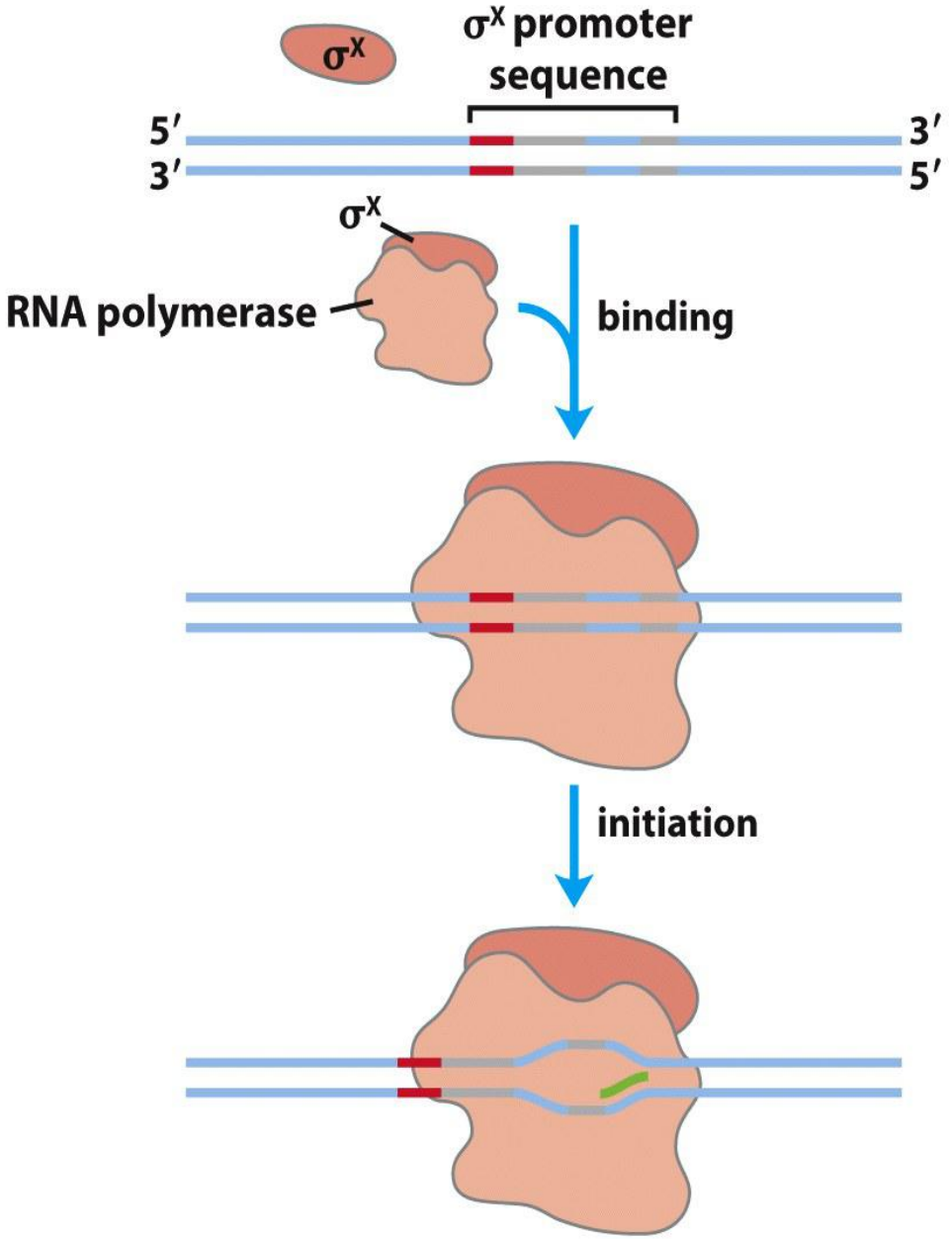


Initiation of transcription

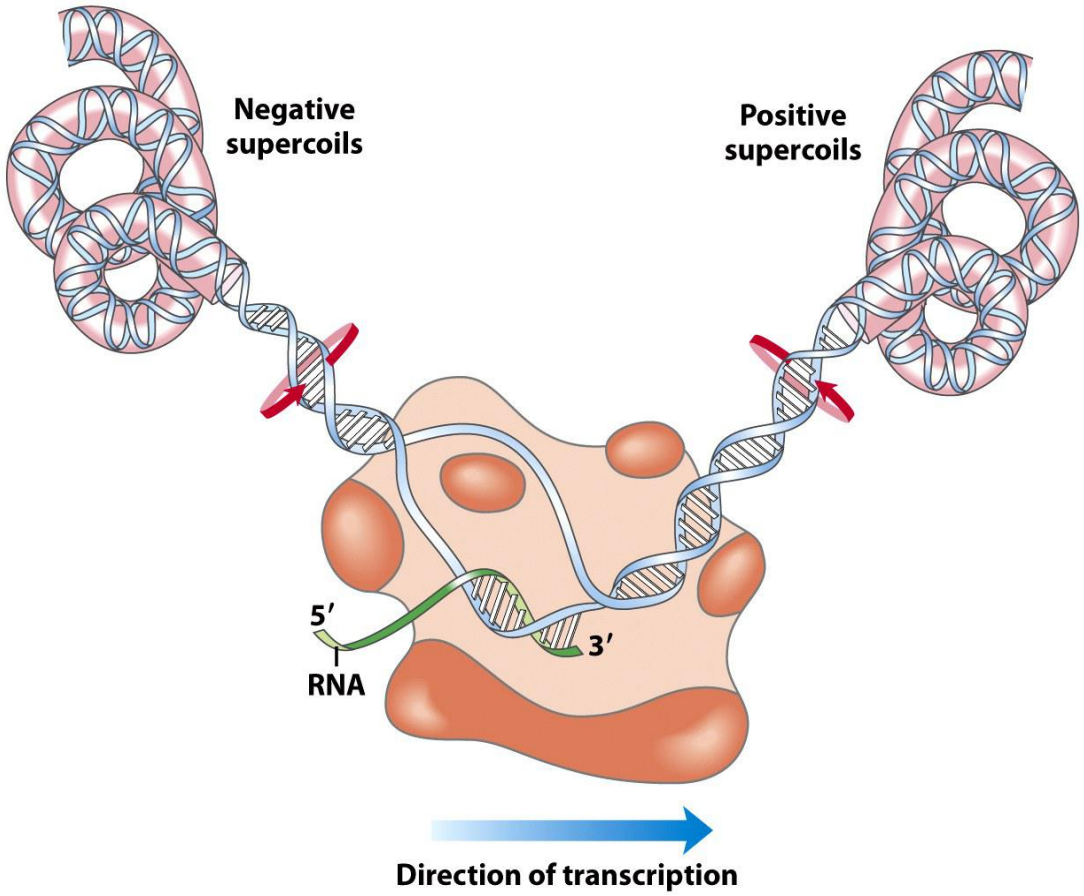
Initiation of transcription contains two phases: **binding** and **initiation**.

Core enzyme ($\alpha_2\beta\beta'\omega$) is the basic transcription machinery, it will lose specificity and can only transcribe nicked DNA template but not intact DNA without σ subunit .

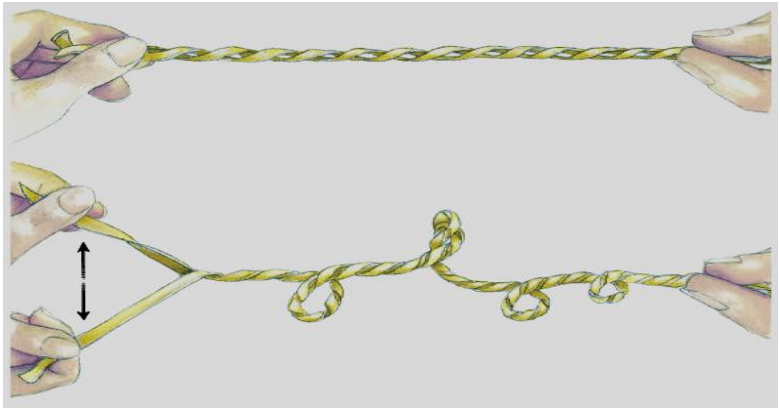
Once the first 8 or 9 nucleotides of a new RNA are synthesized, the σ subunit is released and the polymerase leaves the promoter and becomes committed to elongation of the RNA.



Supercoiling of DNA in transcription

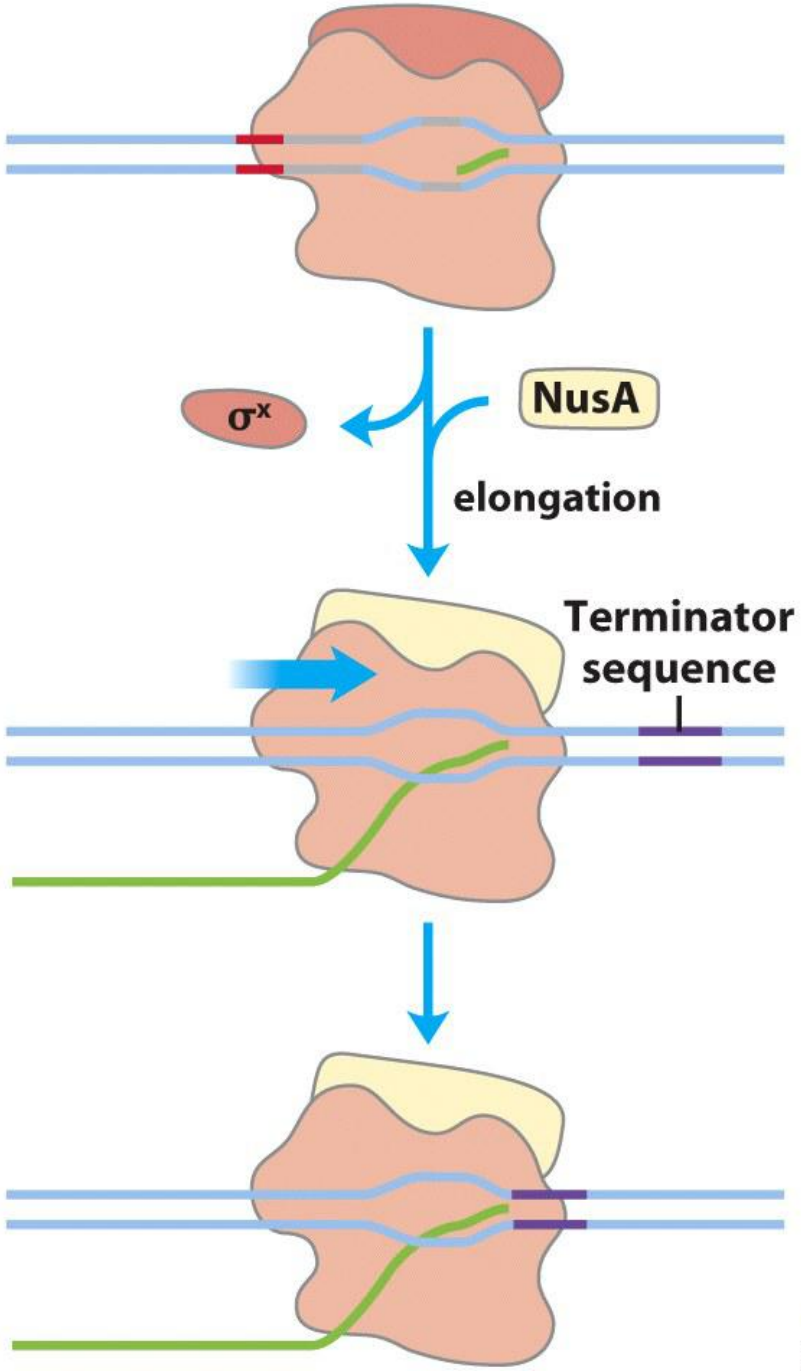


Movement of an RNA polymerase along DNA tends to create positive supercoils (overwound DNA) ahead of the transcription bubble and negative supercoils (underwound DNA) behind it. In a cell, **topoisomerases** rapidly eliminate the positive supercoils and regulate the level of negative supercoiling.



Chain elongation

Elongation of the RNA transcript is catalyzed by the core polymerase without σ subunit.



Chain termination

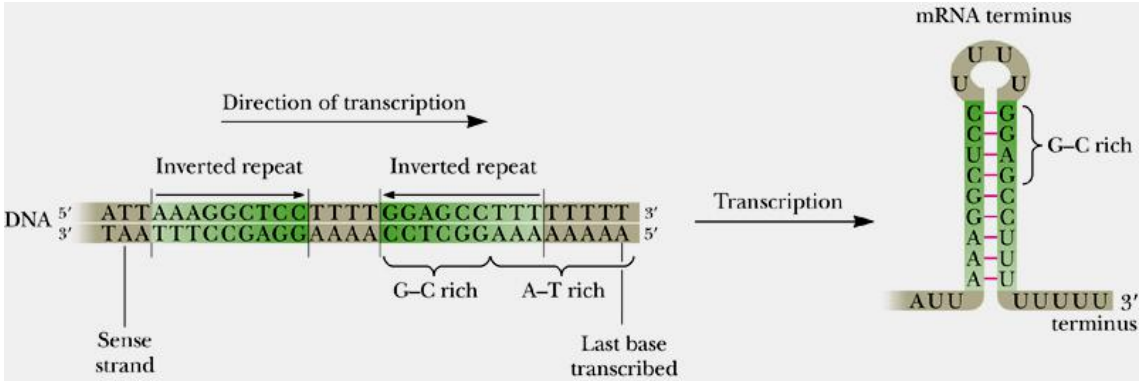
Two types of transcription termination mechanisms in *E. coli*:

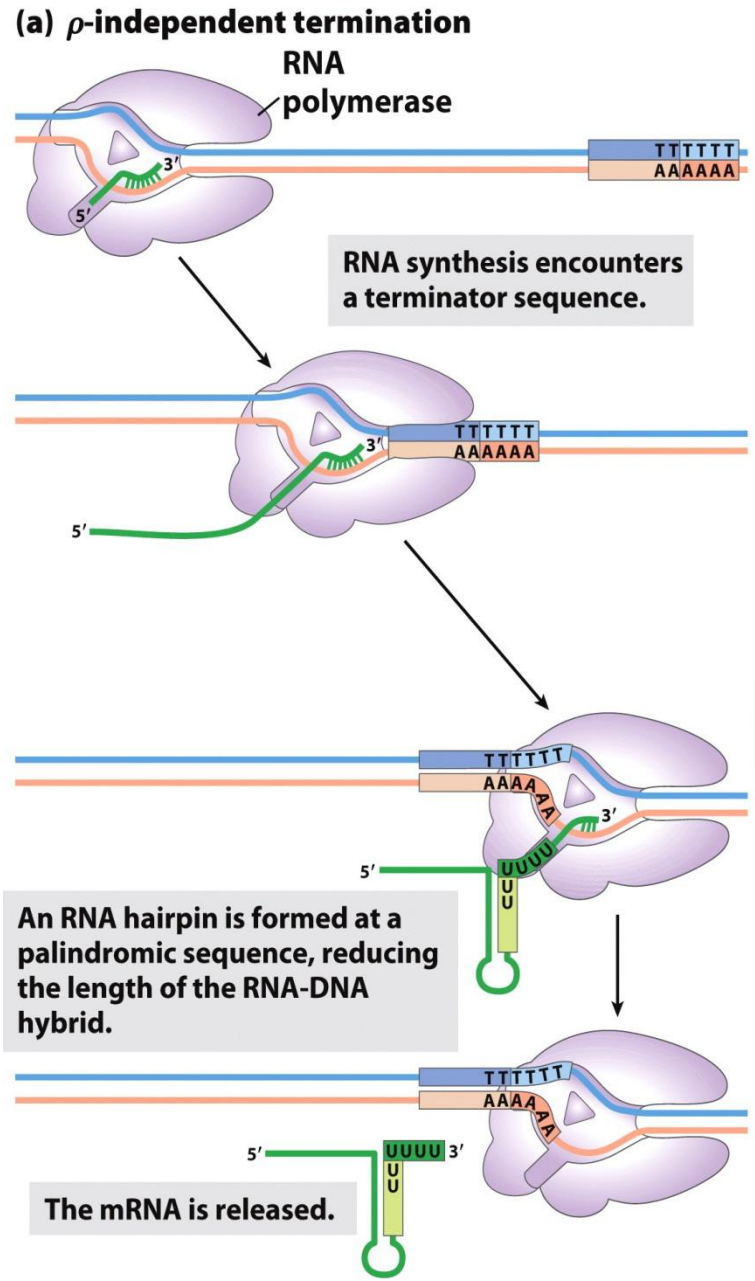
- independent on ρ (rho) termination factor
- dependent on ρ termination factor

ρ-independent termination of transcription in *E. coli.*

Requires a Termination site:

1. Inverted repeats, typically G-C rich and, forms stable stem-loop;
2. A nonrepeating segment punctuates the inverted repeats;
3. A run of 6 to 8 Ts

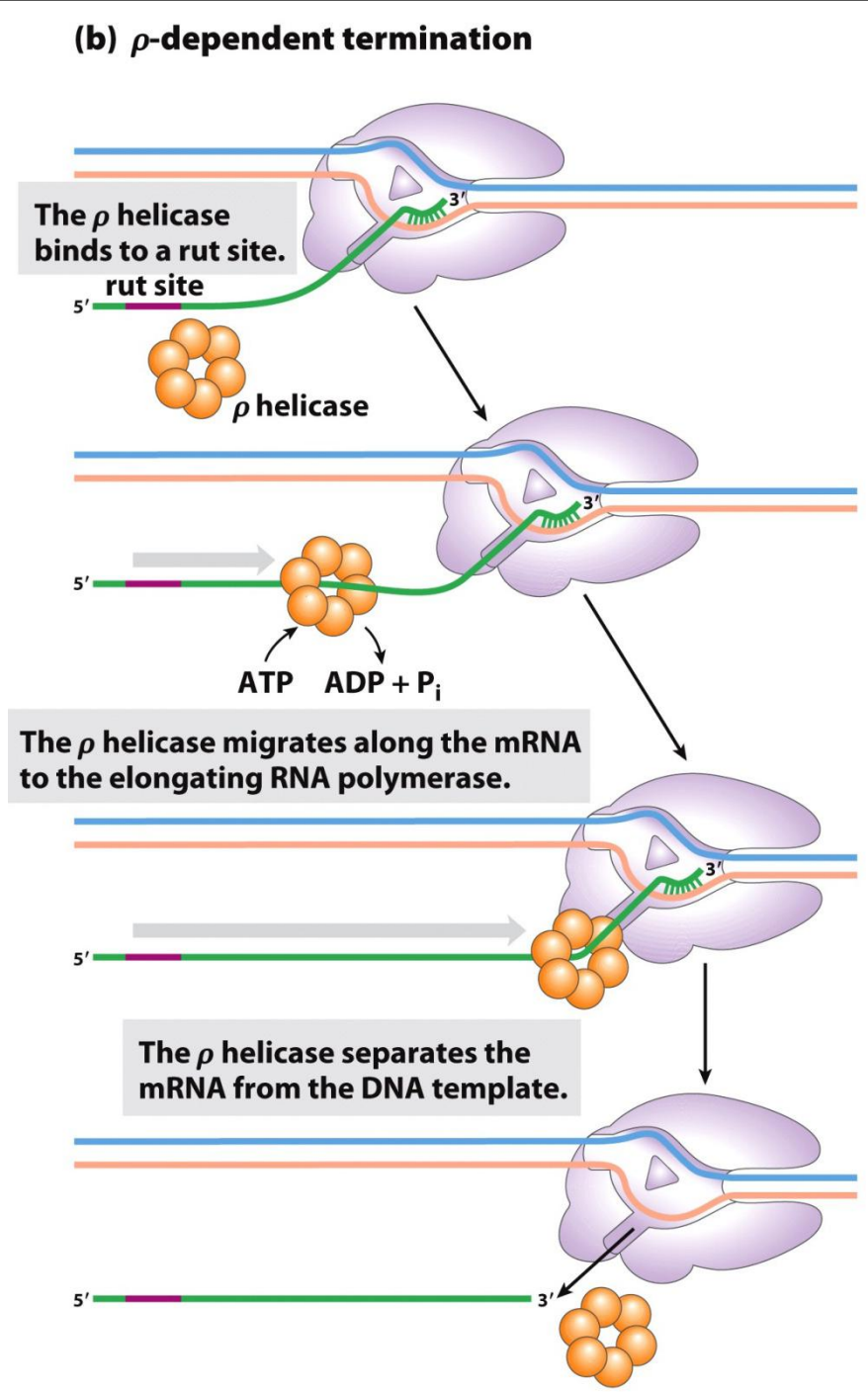




ρ -dependent termination in *E. coli*

rut (*r*ho *ut*ilization) site :

a CA-rich sequence that ρ protein can associate with



Transcription in eukaryotes

RNA polymerases in eukaryotes

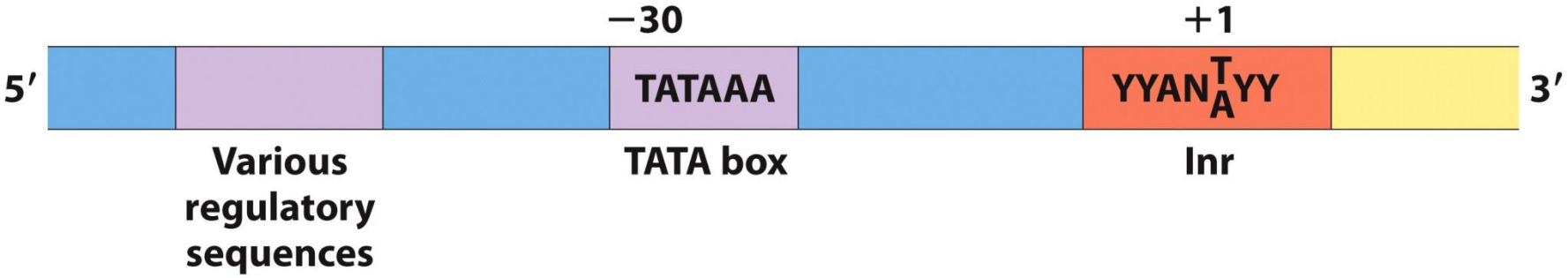
Three classes of RNA polymerase in eukaryotes for synthesis of different class of RNA:

- RNA polymerase I (Pol I) is responsible for the synthesis of preribosomal RNA (or pre-rRNA), which contains the precursor for the 18S, 5.8S, and 28S rRNAs. Pol I promoters vary greatly in sequence from one species to another.

- RNA Polymerase II (Pol II) is for synthesis of mRNAs and some specialized RNAs. This enzyme can recognize thousands of promoters that vary greatly in sequence.

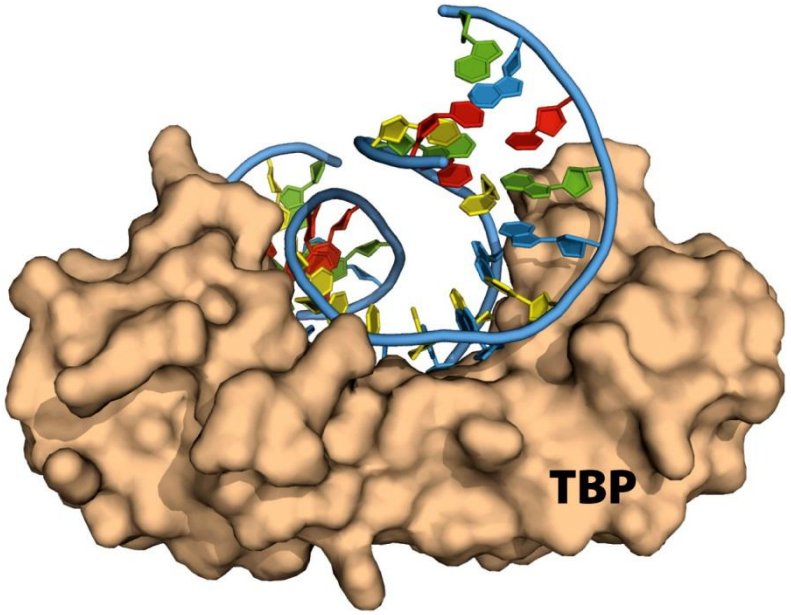
- RNA polymerase III (Pol III) makes tRNAs, the 5S rRNA, and some other small specialized RNAs.

Eukaryotic promoters



- **TATA box:** the major assembly point for the proteins of the preinitiation complexes of Pol II.
- **Initiator sequence (Inr):** Pol II binding site and the transcription start site is usually within or very near this sequence.

TATA-binding protein (TBP)



RNA polymerase II requires many other protein factors for its activity

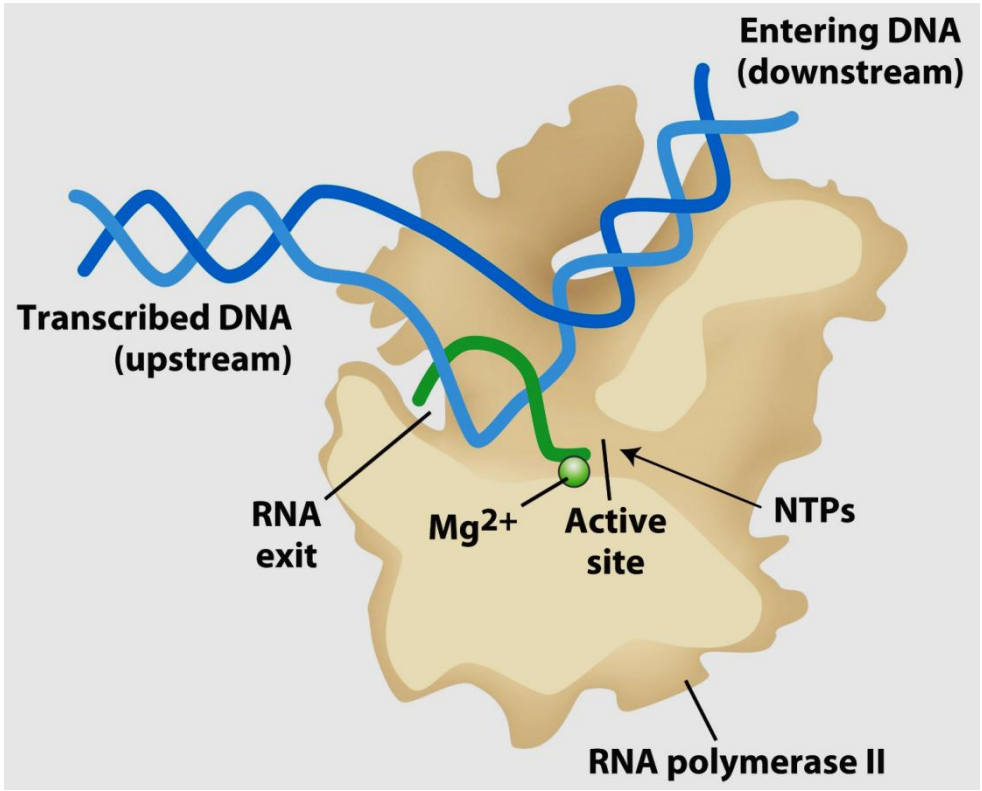
TABLE 26–2 Proteins Required for Initiation of Transcription at the RNA Polymerase II (Pol II) Promoters of Eukaryotes

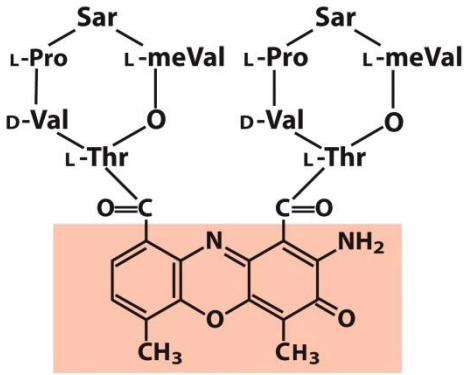
Transcription protein	Number of subunits	Subunit(s) M_r	Function(s)
Initiation			
Pol II	12	10,000–220,000	Catalyzes RNA synthesis
TBP (TATA-binding protein)	1	38,000	Specifically recognizes the TATA box
TFIIA	3	12,000, 19,000, 35,000	Stabilizes binding of TFIIB and TBP to the promoter
TFIIB	1	35,000	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	
pTEFb	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	

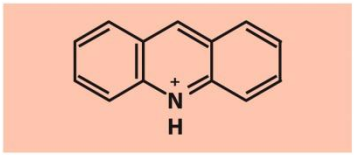
*The function of all elongation factors is to suppress the pausing or arrest of transcription by the Pol II–TFIIF complex.

†Name derived from *e*leven-nineteen *l*ysine-rich *l*eukemia. The gene for ELL is the site of chromosomal recombination events frequently associated with acute myeloid leukemia.

Structure of Pol II core enzyme



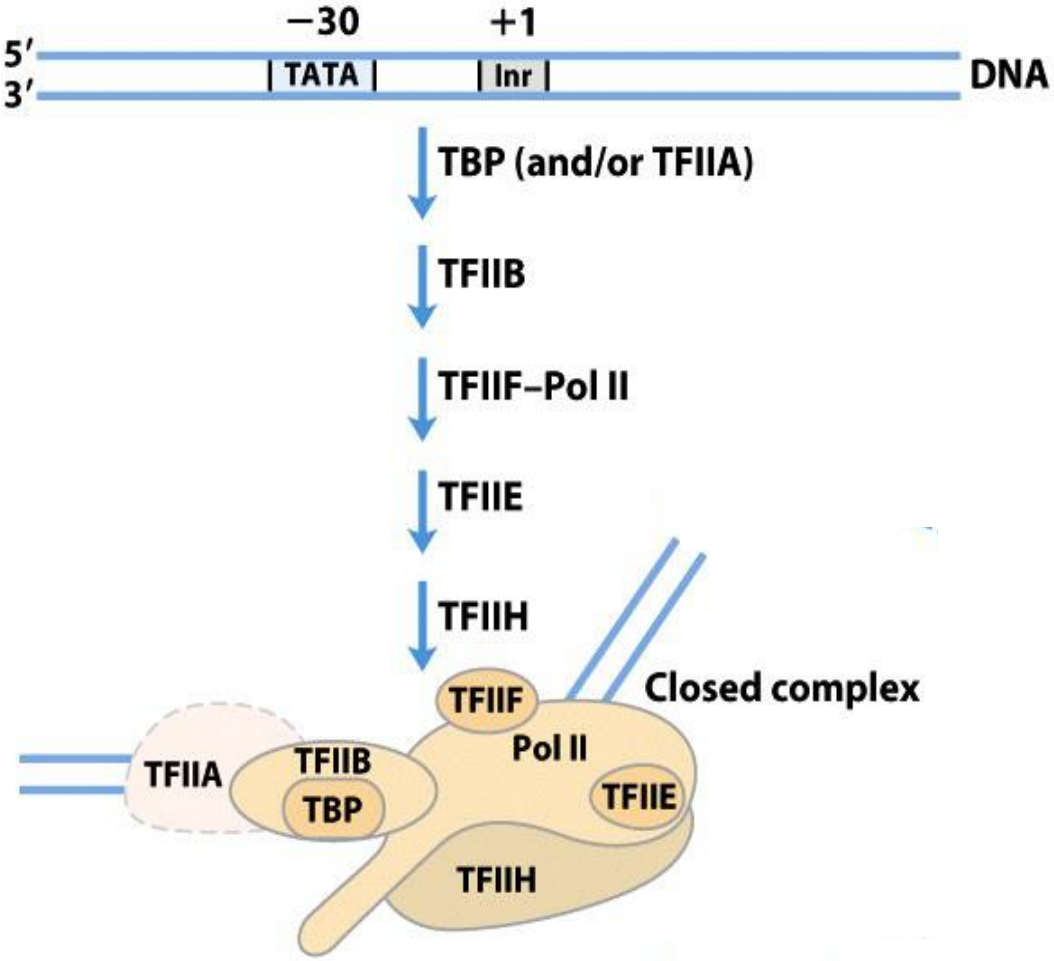




Actinomycin D & Acridine can selectively inhibit DNA-dependent RNA polymerase in both bacteria and eukaryotes by inserting into the double-helical DNA between successive G≡C base pairs to prevent movement of the polymerase along the template.

Assembly of pre-initiation complex

There are more than 30 proteins for this minimal active assembly.



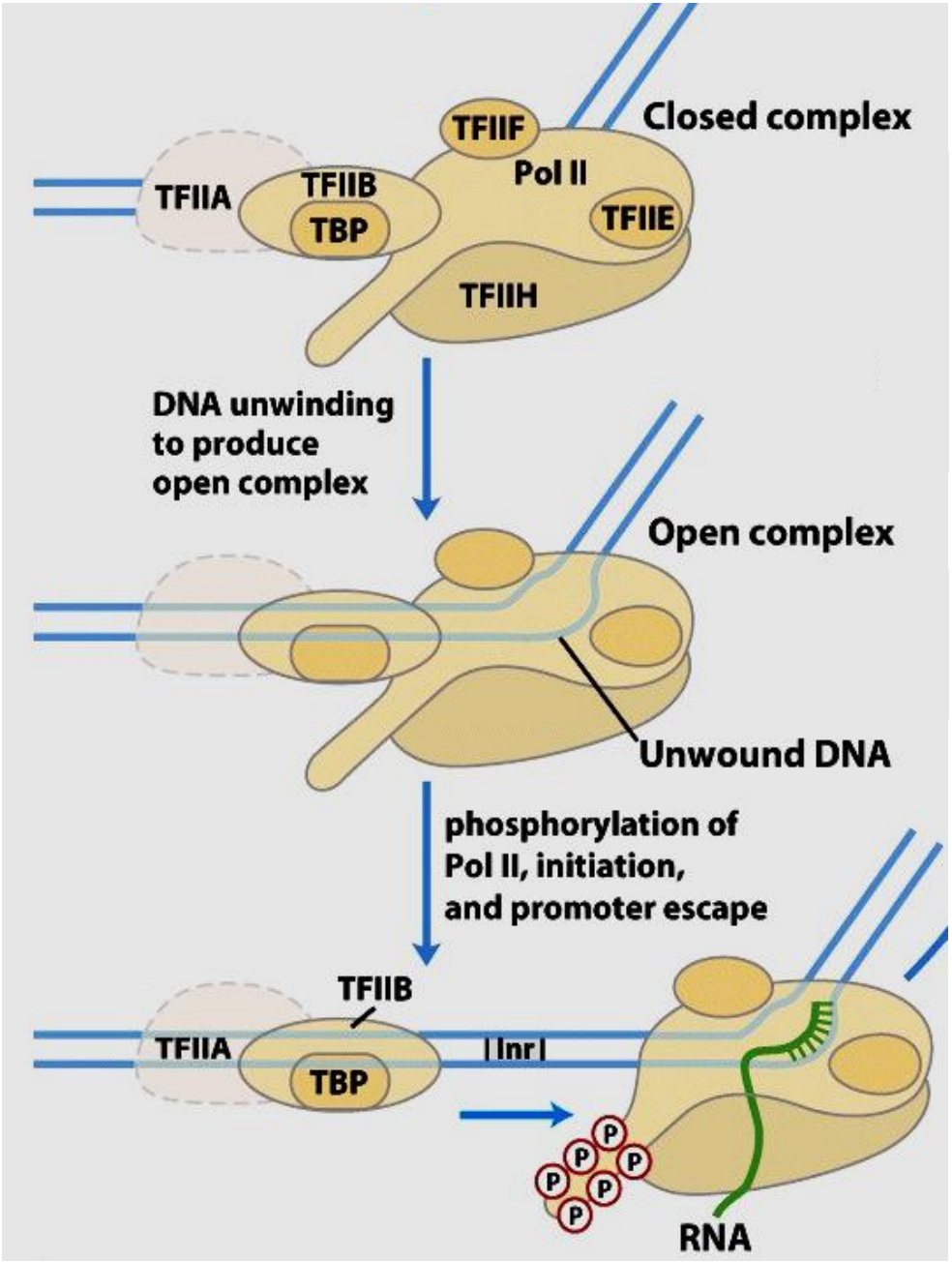
Initiation	
Pol II	Catalyzes RNA synthesis
TBP (TATA-binding protein)	Specifically recognizes the TATA box
TFIIA	Stabilizes binding of TFIIB and TBP to the promoter
TFIIB	Binds to TBP; recruits Pol II–TFIIF complex
TFIIE	Recruits TFIIH; has ATPase and helicase activities
TFIIF	Binds tightly to Pol II; binds to TFIIB and prevents binding of Pol II to nonspecific DNA sequences
TFIIH	Unwinds DNA at promoter (helicase activity); phosphorylates Pol II (within the CTD); recruits nucleotide-excision repair proteins

RNA strand initiation and promoter clearance

DNA is unwound at the Inr (initiator sequence) region by the helicase activity of TFIIH, creating an open complex.

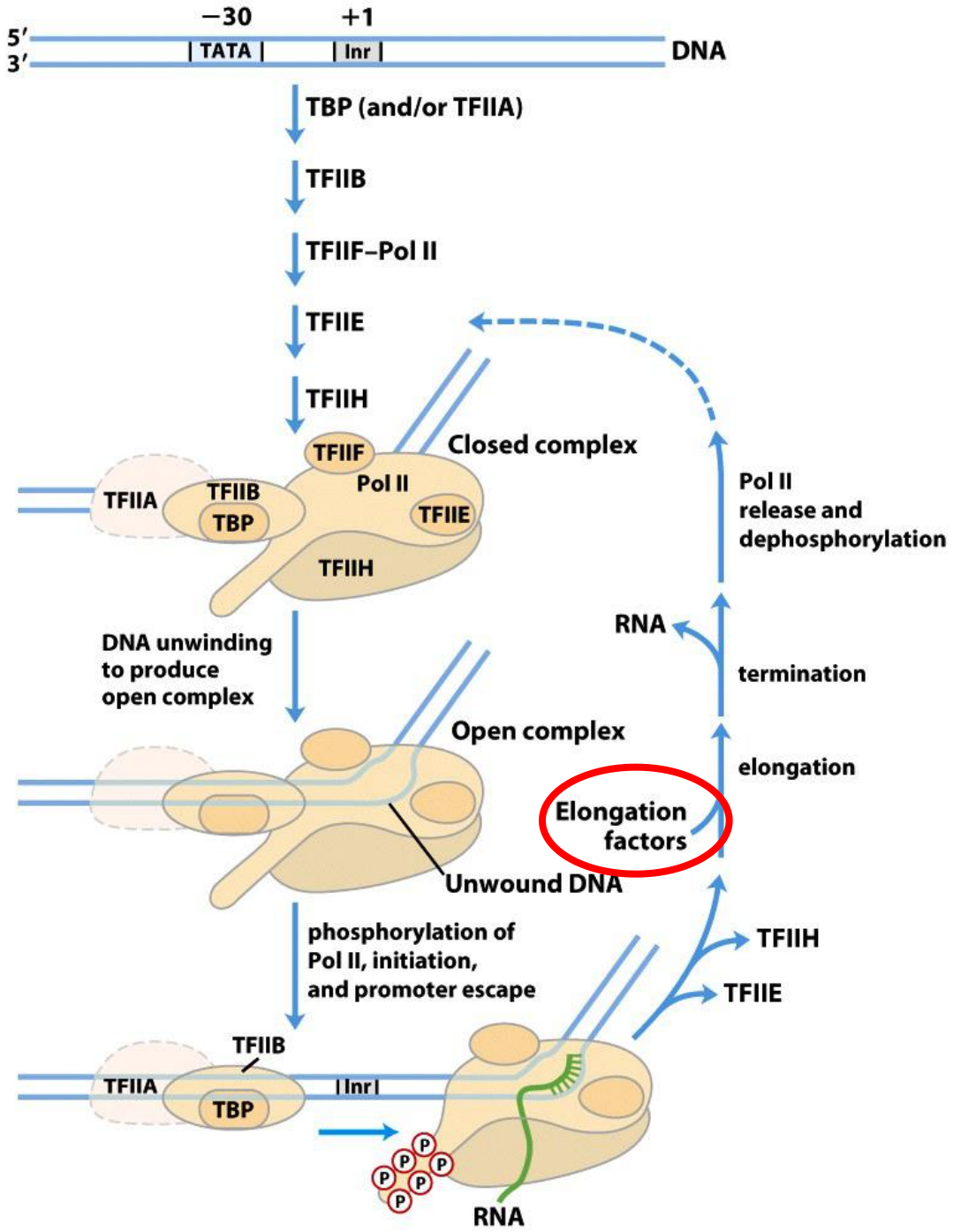
The carboxyl-terminal domain (CTD) of the largest Pol II subunit is phosphorylated, and the polymerase then escapes the promoter and begins transcription.

Elongation is accompanied by the release of many transcription factors and is also enhanced by elongation factors.



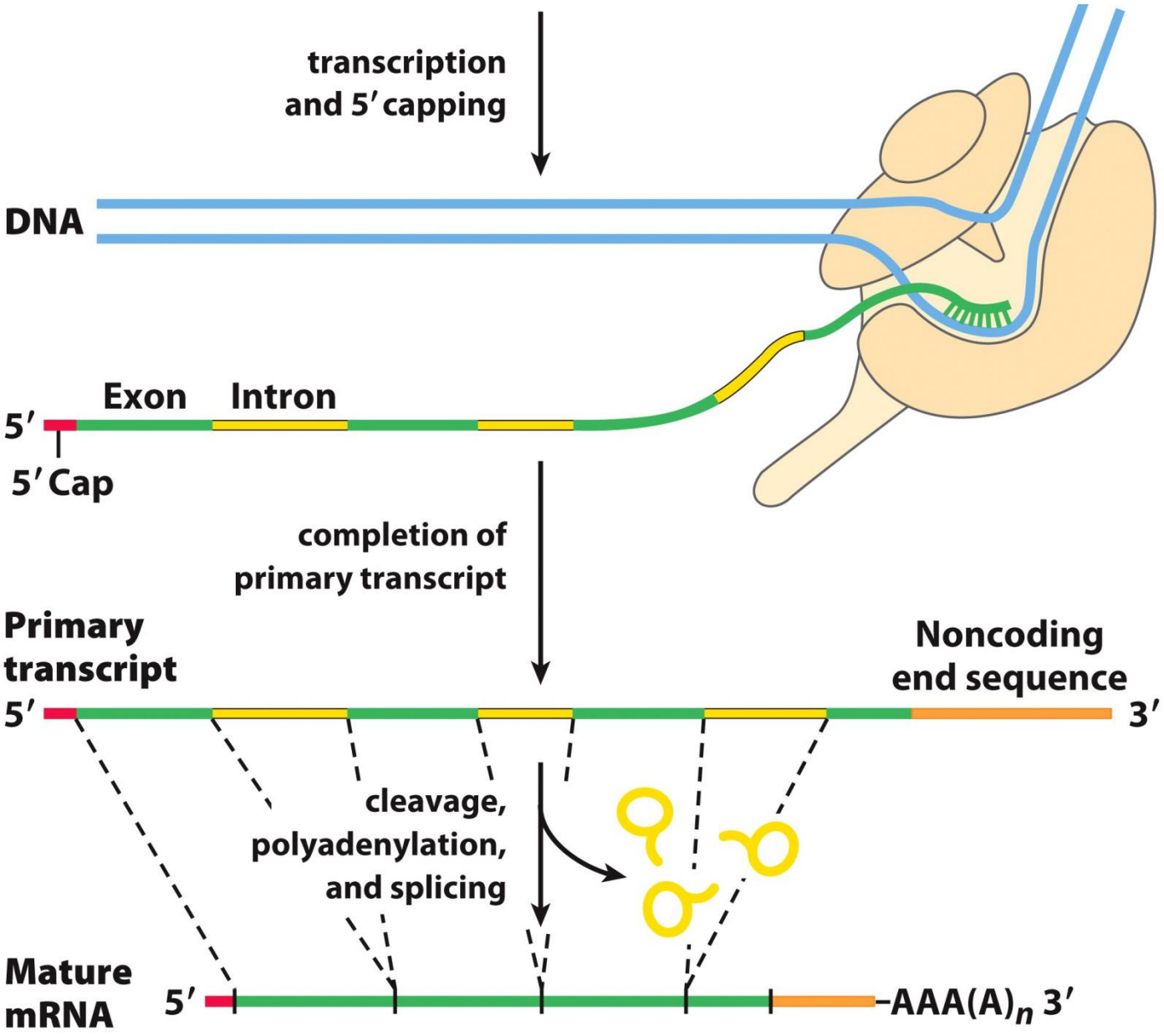
Elongation, termination, & release

The elongation factors bound to Pol II (some to the phosphorylated CTD), suppress pausing during transcription therefore greatly enhance the polymerase activity. They also coordinate interactions between protein complexes involved in the posttranscriptional processing of mRNAs. Once the RNA transcript is completed, transcription is terminated. Pol II is dephosphorylated and recycled, ready to initiate another transcript.

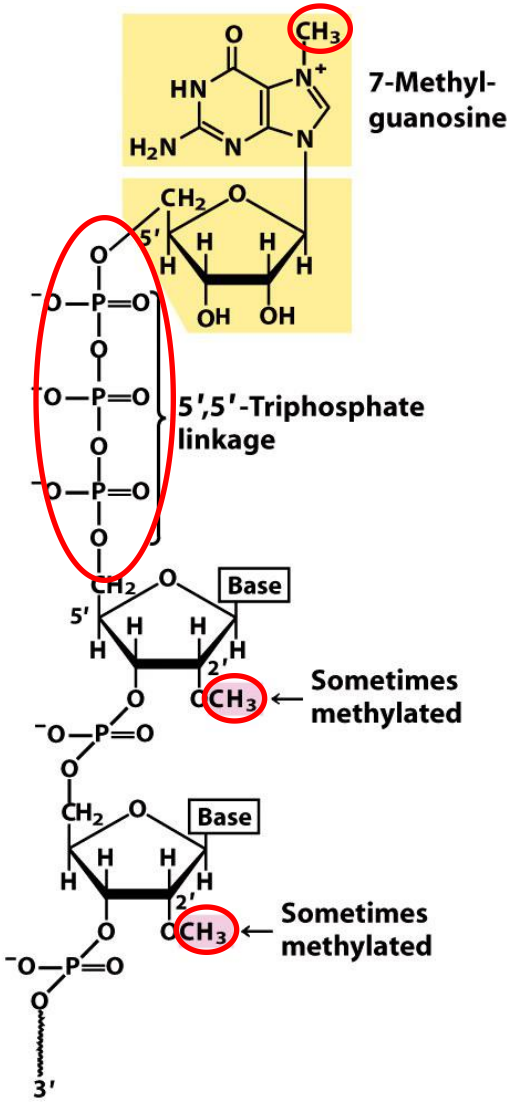


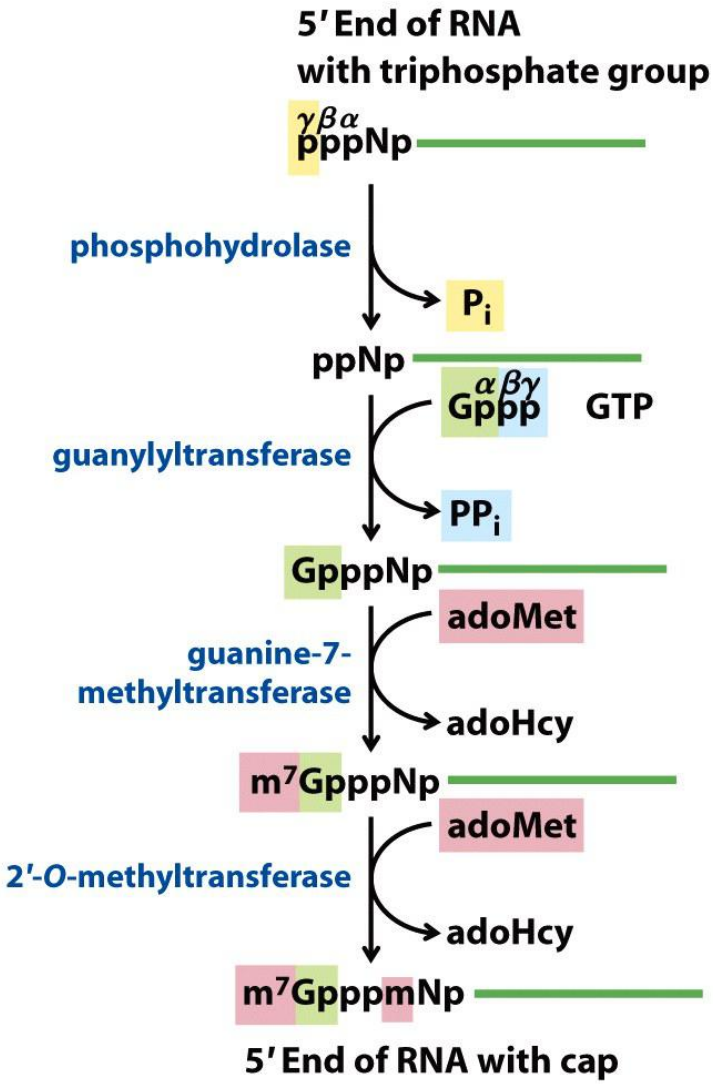

HHMI

Outline of mRNA processing



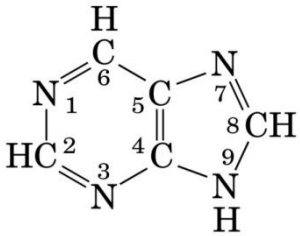
The 5’ cap of mRNA



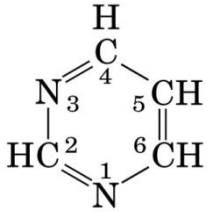


The cap provides:

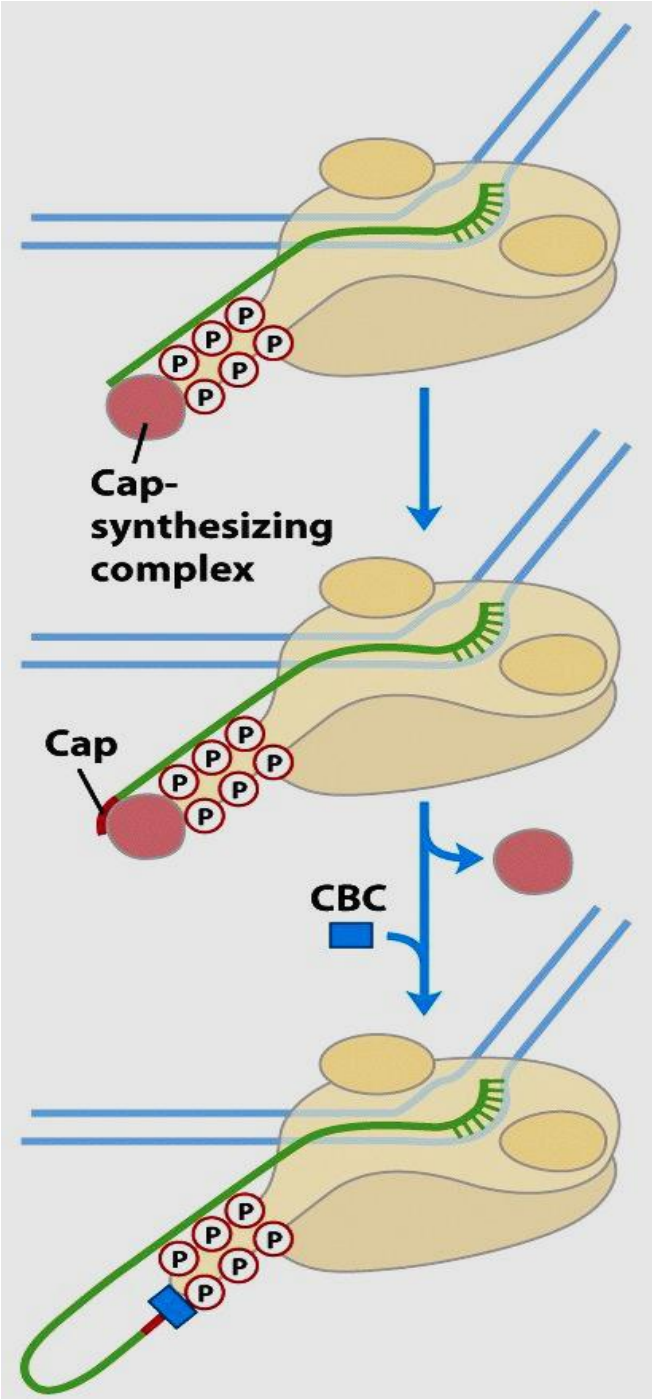
1. protection of the mRNA;
2. transport of the mRNA out of the nucleus;
3. proper



Purine



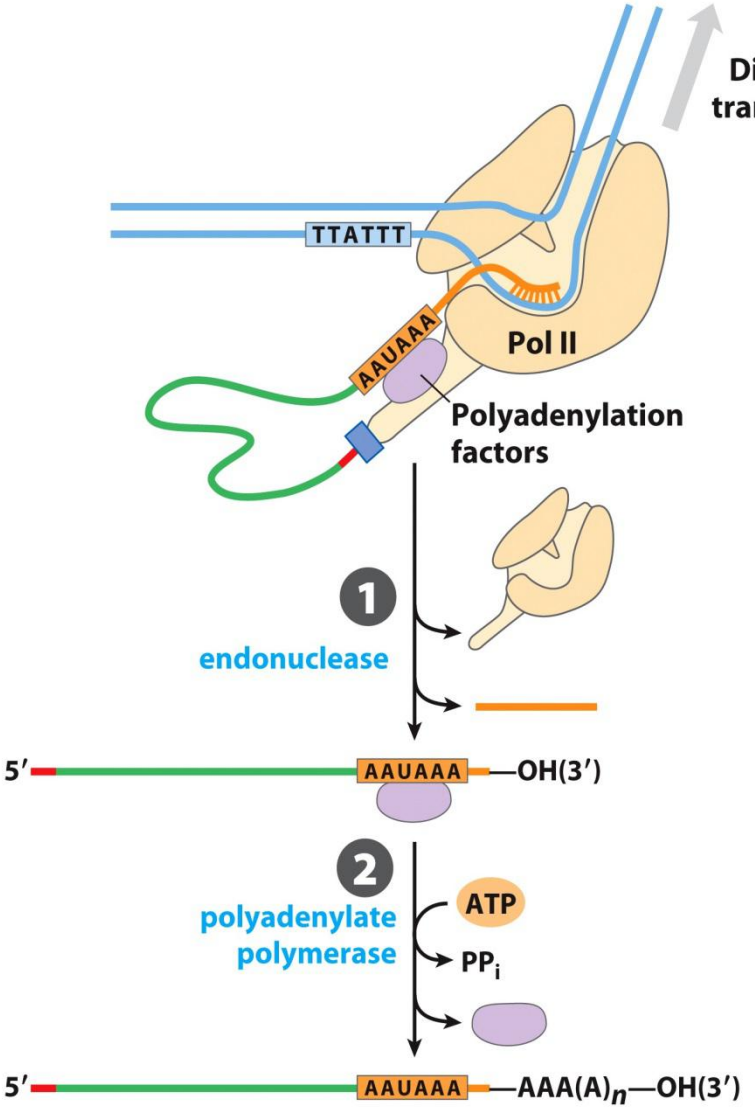
Pyrimidine

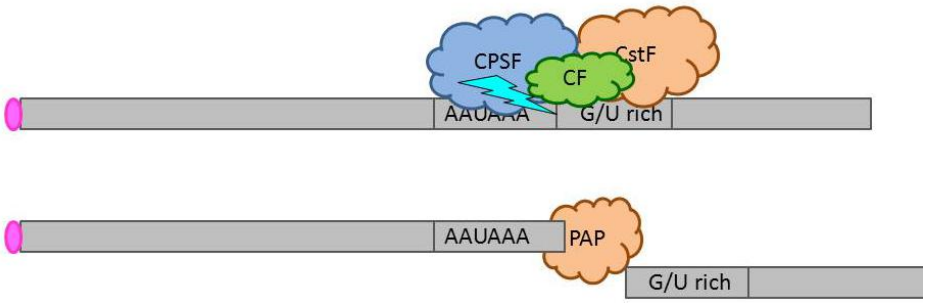


Pol II CTD and 5' cap synthesis

Synthesis of the cap is carried out by enzymes tethered to the **CTD** of Pol II. The cap remains tethered to the **CTD** through an association with the **cap-binding complex (CBC)**.

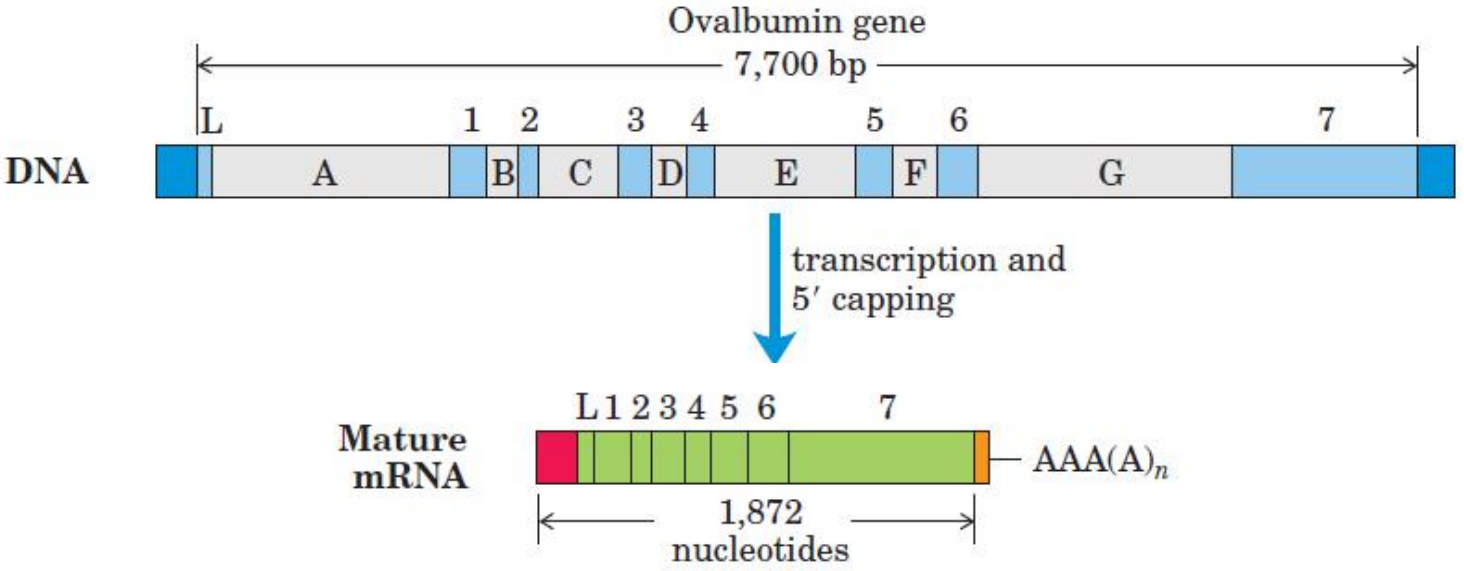
End of the primary RNA transcription in eukaryotes: forming polyA tail





1. CPSF binds AAUAAA
2. CstF binds G/U rich sequence
3. Attract cleavage factor CFs
4. Cleavage
5. PolyA polymerase: add polyA 80-250 nt

Exon & Intron



cDNA (complementary DNA) is DNA synthesized from a messenger RNA (mRNA) template in a reaction catalyzed by the enzymes reverse transcriptase and DNA polymerase.

Four classes of introns

- **Group I and group II introns: *self-splicing* introns**—no protein enzymes are involved and does not require a high energy cofactor (such as ATP) for splicing.

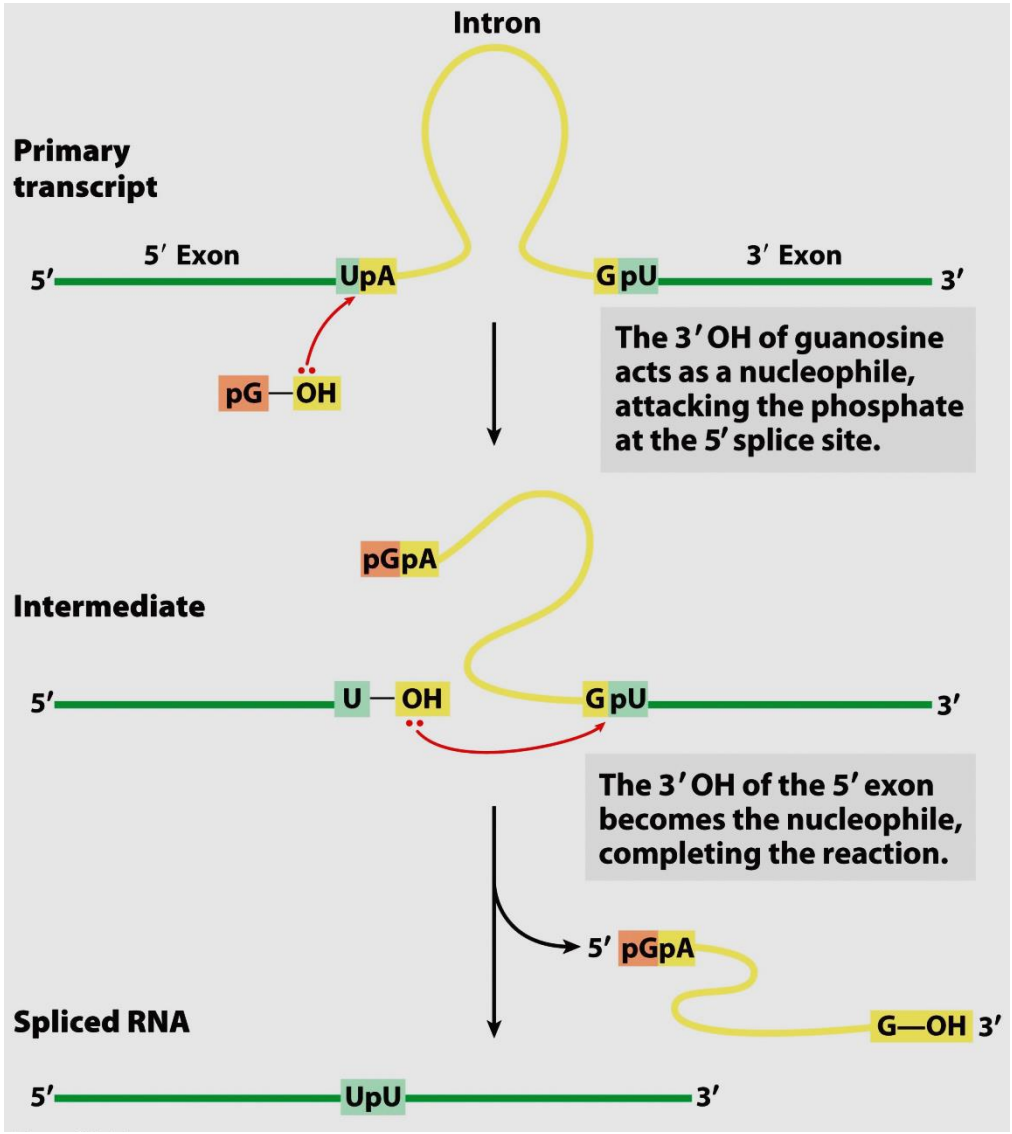
  Group I introns are found in some nuclear, mitochondrial, and chloroplast genes coding for rRNAs, mRNAs, and tRNAs.

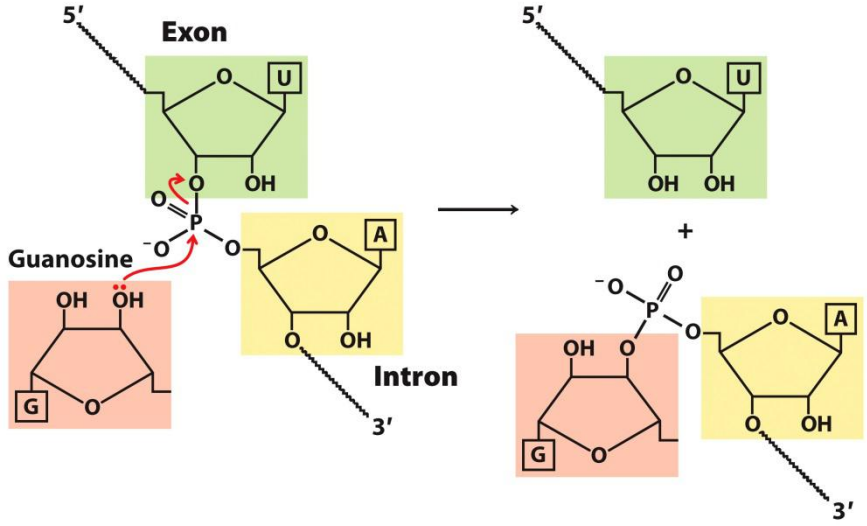
  Group II introns are generally found in the primary transcripts of mitochondrial or chloroplast mRNAs in fungi, algae, and plants.

- **Spliceosomal introns:** the largest class of introns includes those found in nuclear mRNA primary transcripts. Need **spliceosome** (a large protein complex) for splicing.

- **The fourth class of introns**: found in certain tRNAs, splicing requires ATP and an endonuclease. The splicing endonuclease cleaves the phosphodiester bonds at both ends of the intron, and the two exons are joined by a mechanism similar to the DNA ligase reaction.

Splicing mechanism of group I introns





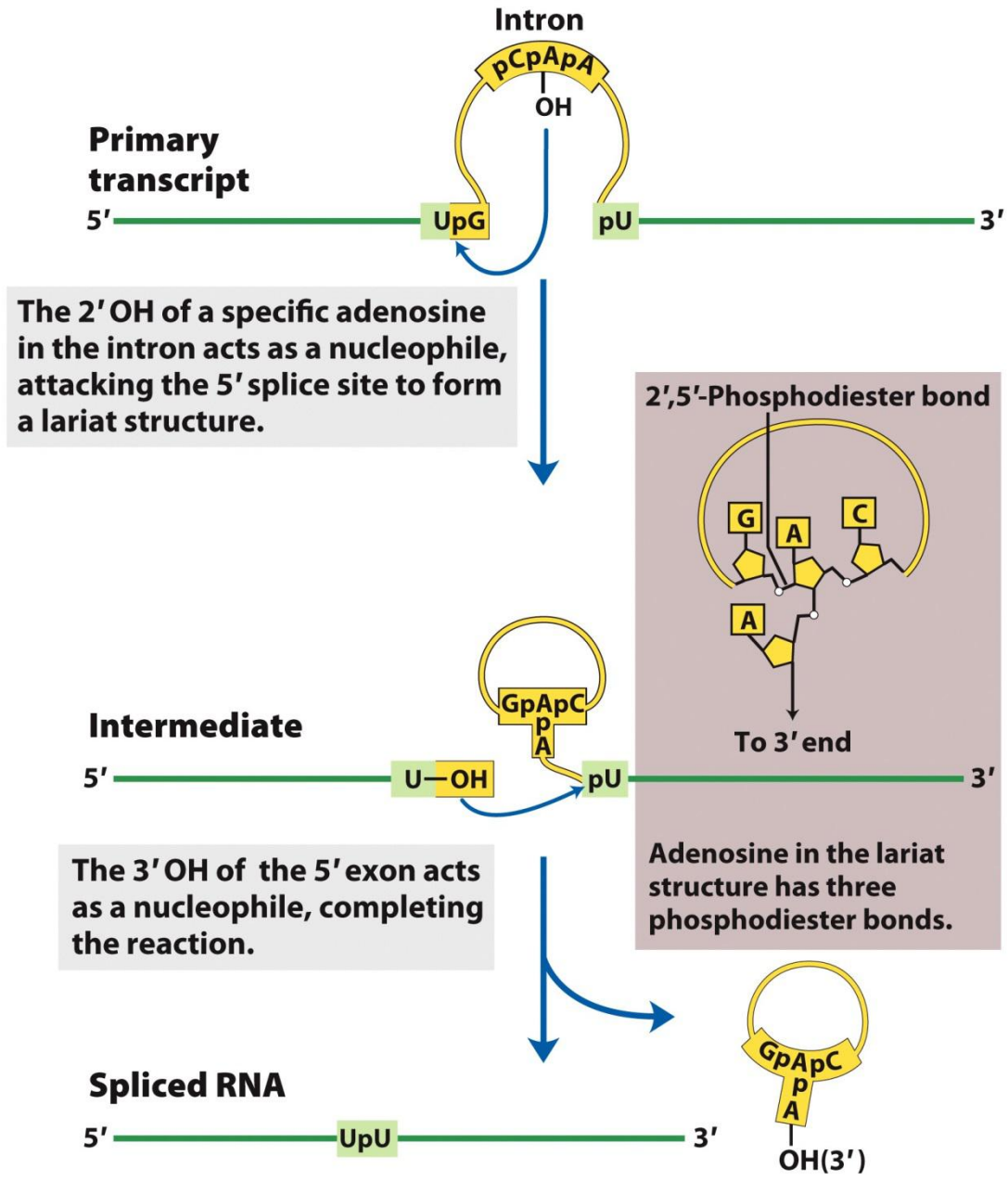
The Nobel Prize in Chemistry 1989

Thomas Cech

Found in *Tetrahymena thermophila* (四膜虫), 1982

Splicing mechanism of group II introns

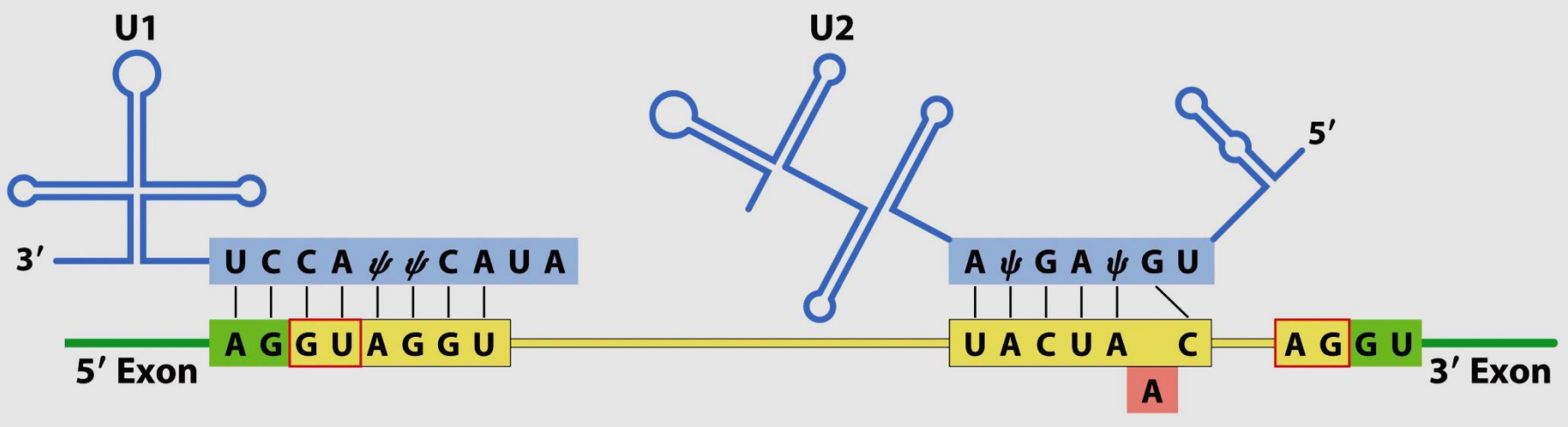
Forming a branched lariat (套索) structure as an intermediate through a 2',5'-phosphodiester bond.



Spliceosomal introns,

which needs spliceosome for the splicing

Consensus sequences at the splicing sites

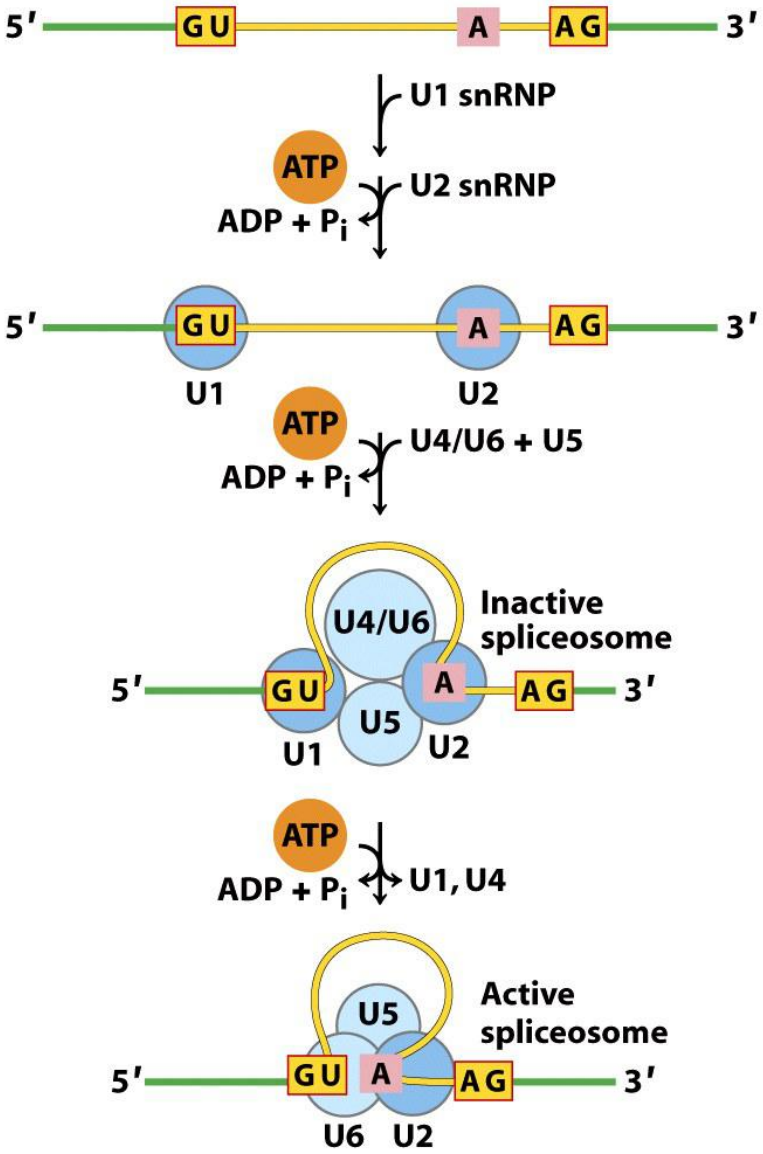


snRNA: small nuclear RNAs, 100-200 nucleotides

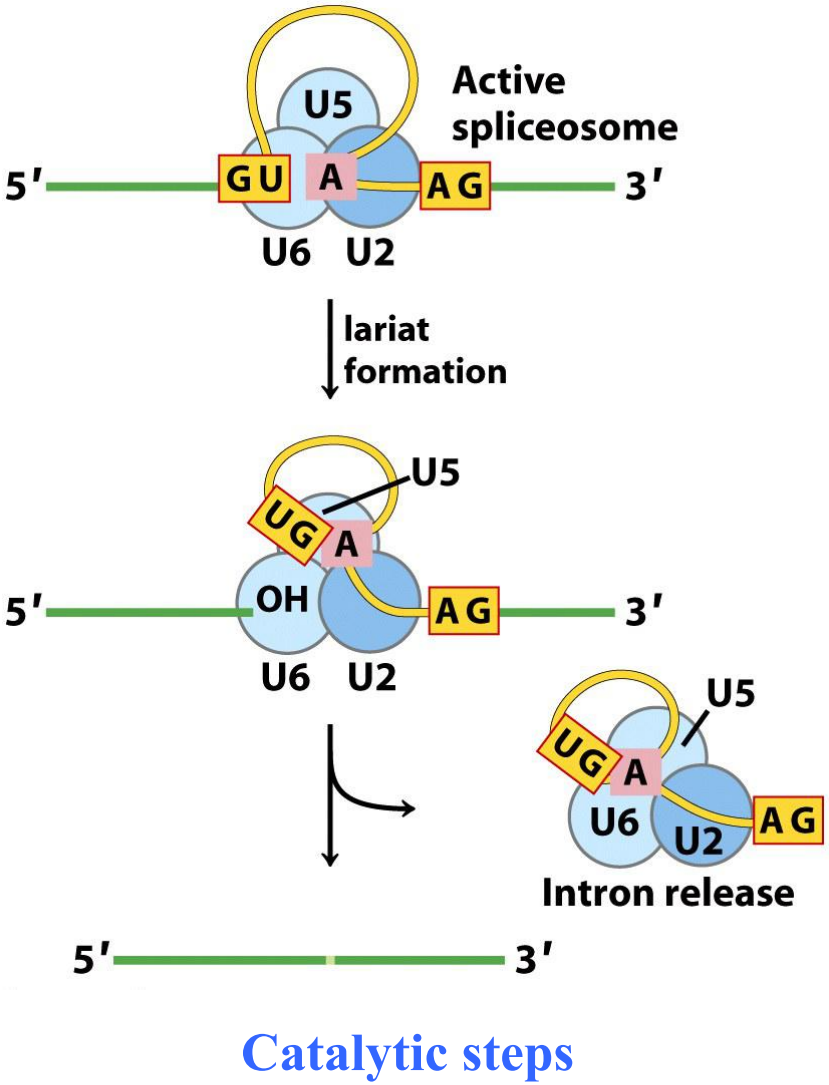
snRNA U1 binds to 5'-splice site binds. ψ is pseudouridine.

snRNA U2 binds to 3'-splice site. The A residue (shaded pink) becomes the nucleophile during the splicing reaction.

Splicing mechanism in mRNA primary transcripts



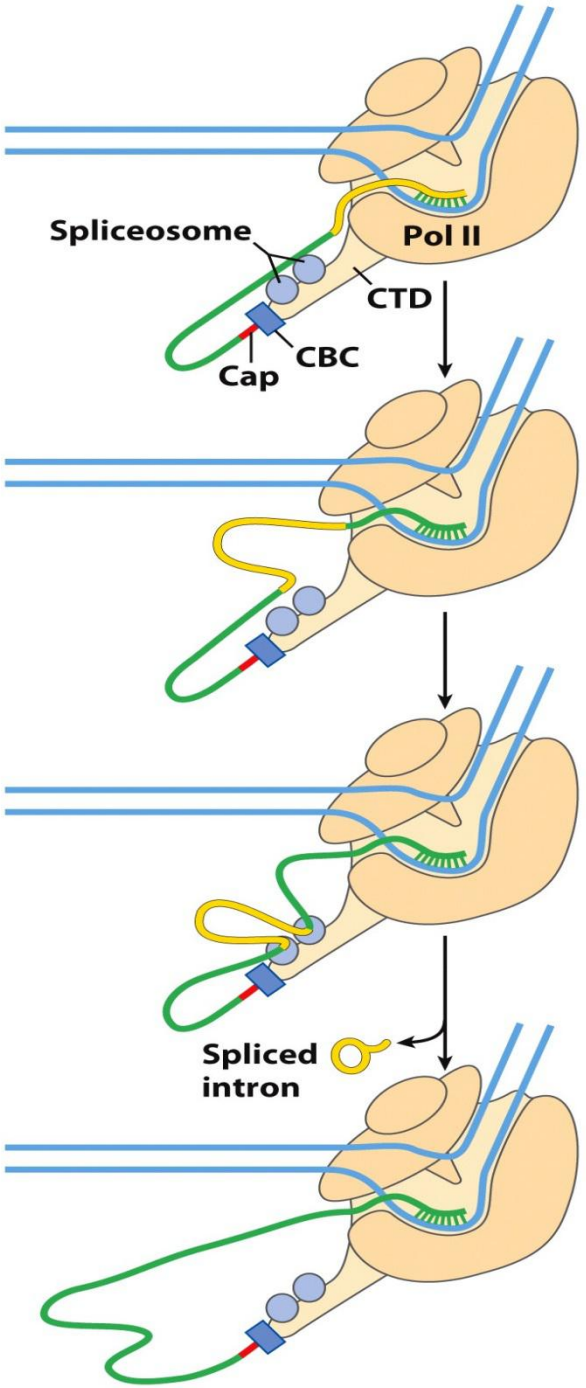
Assembly of spliceosomes



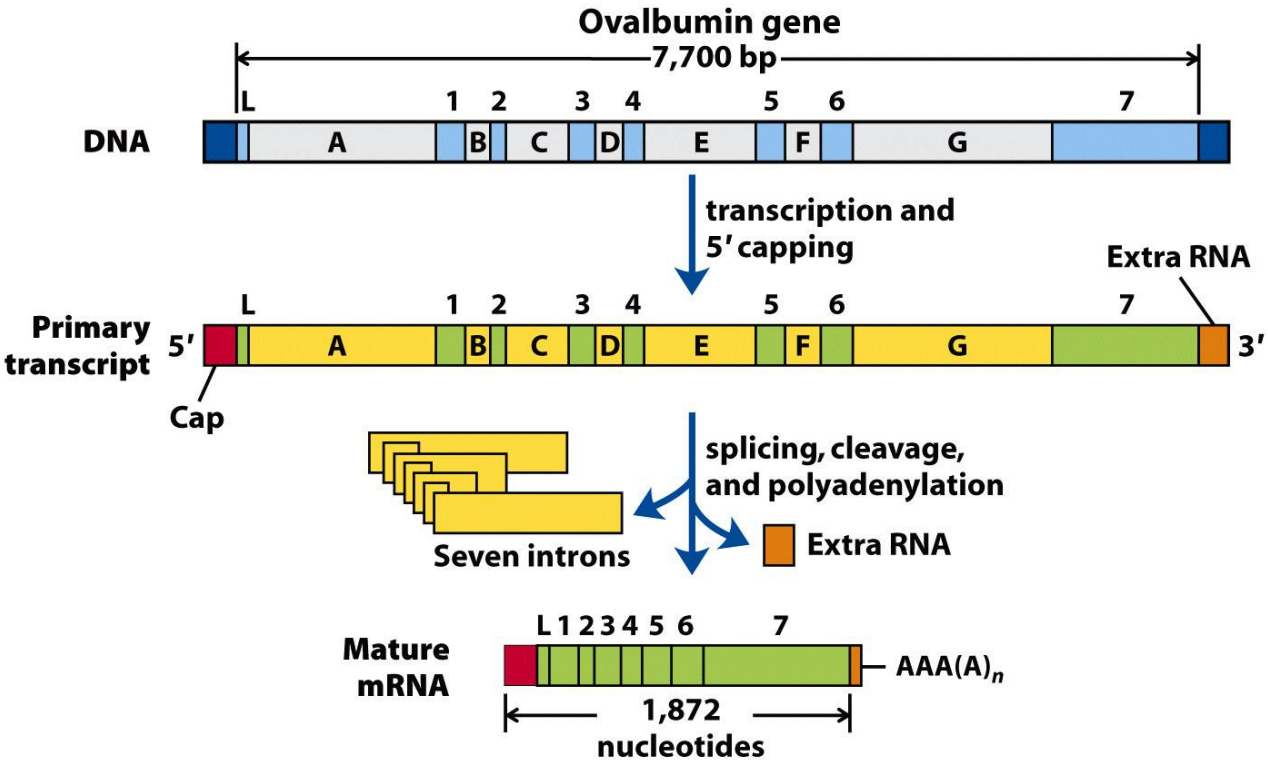
Catalytic steps

Coordination of splicing with transcription

Some components of the splicing apparatus bind are tethered to the CTD of Pol II, suggesting a coordinated mechanism for the splicing reaction and transcription.



Overview of the processing of a eukaryotic mRNA



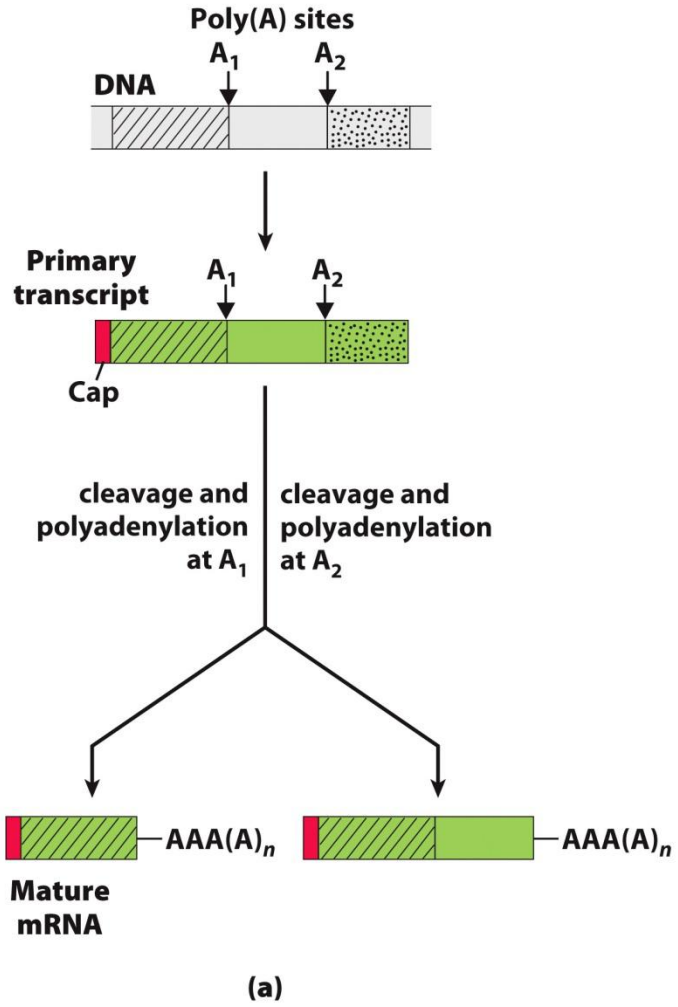
Primary transcript: newly synthesized RNA molecule

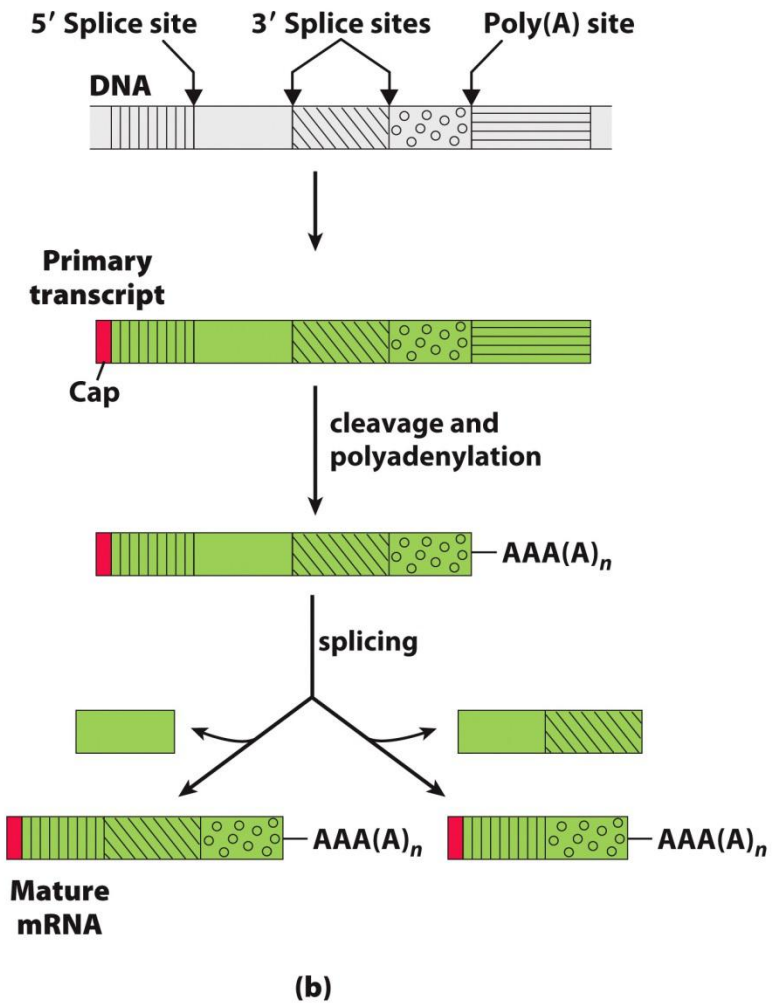
Exon: coding segment

Intron: noncoding sequence

- The **5' cap** is added before synthesis of the primary transcript is complete.
- The 3' end is cleaved, and 80 to 250 A residues are added to create a **poly(A)** "tail."
- Splicing can occur either before or after the cleavage and polyadenylation steps.

Alternative splicing: a gene can give rise to multiple products by differential RNA processing

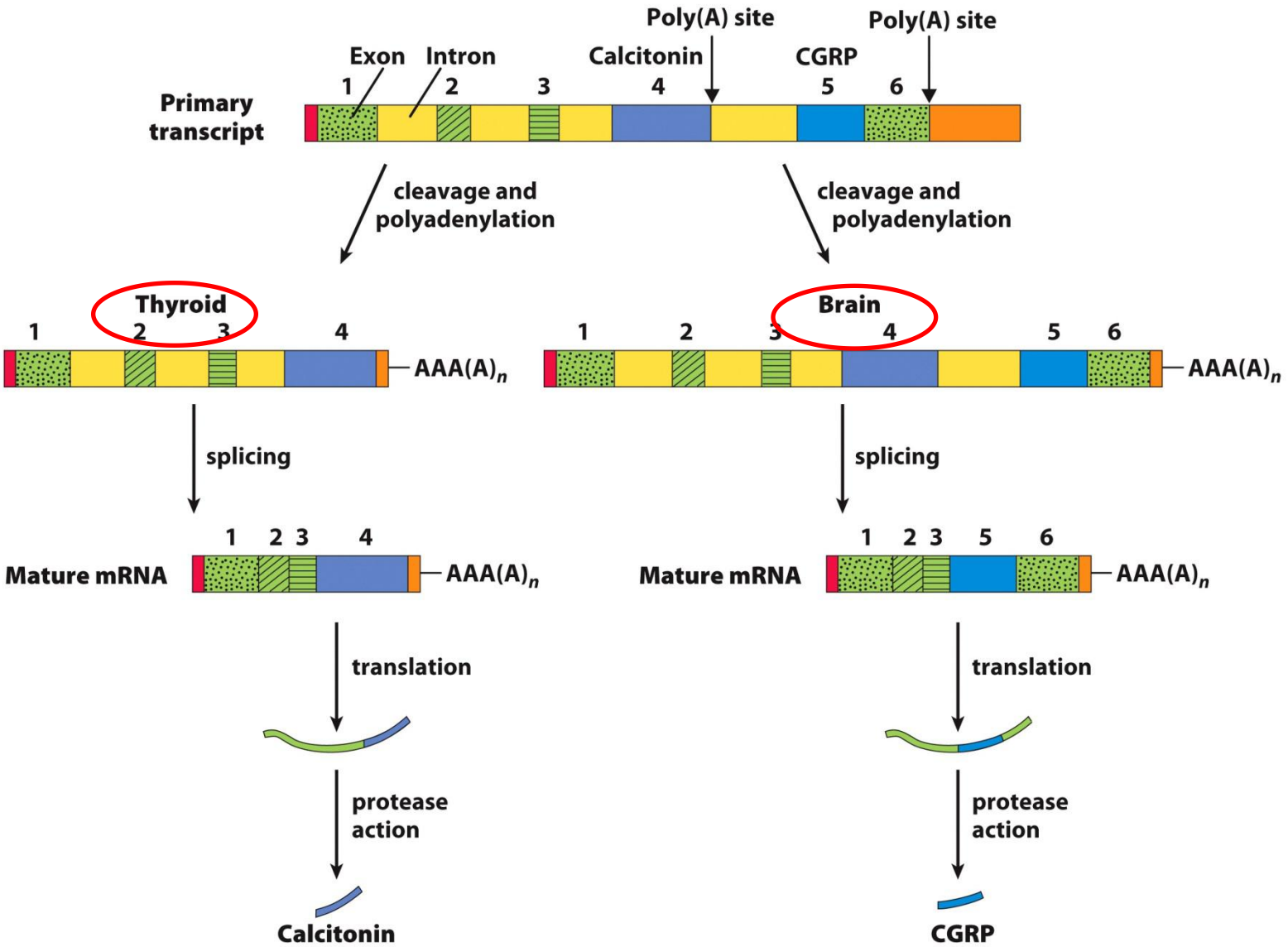




Alternative cleavage and polyadenylation patterns

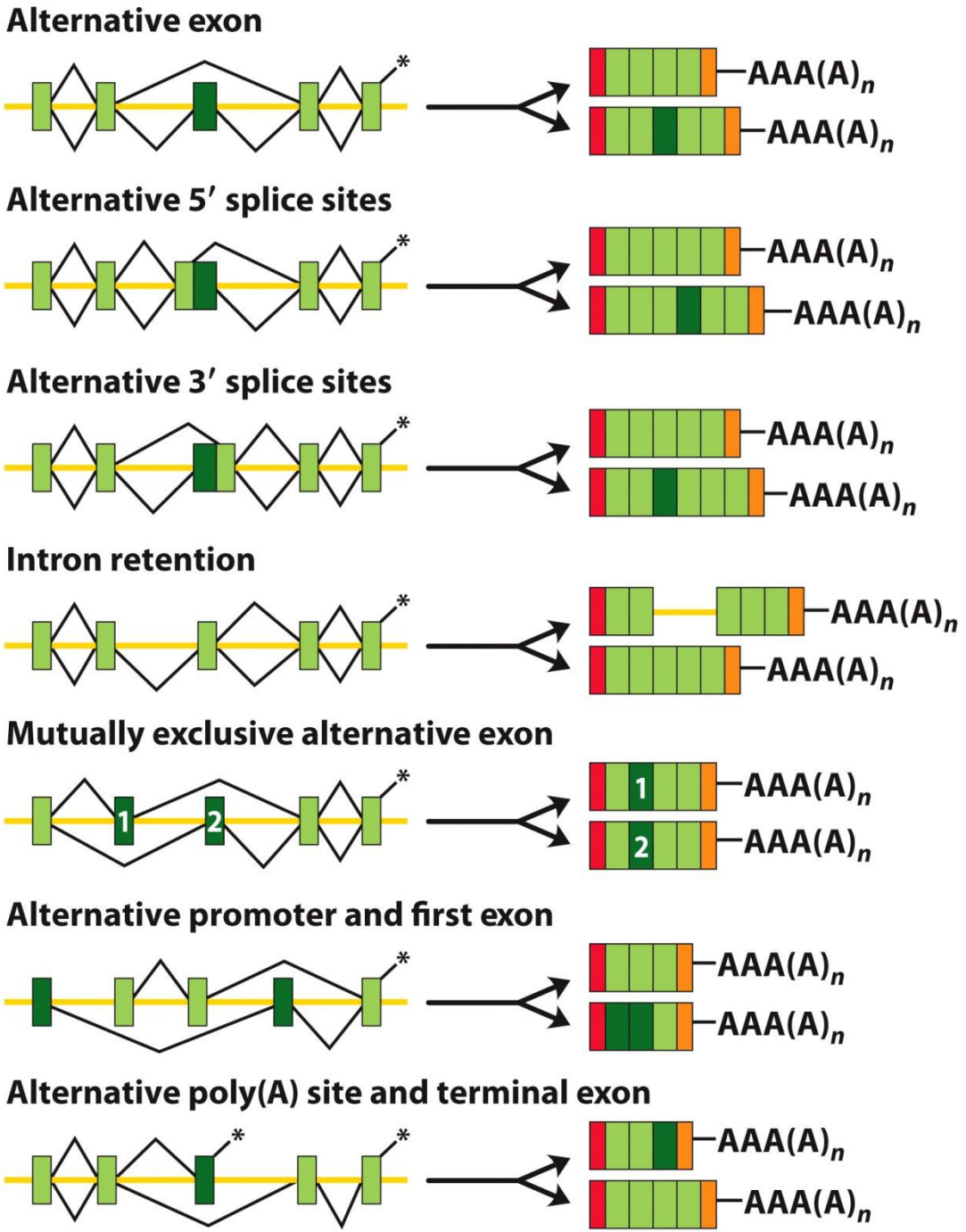
Alternative splicing patterns

Alternative processing of the calcitonin gene transcript

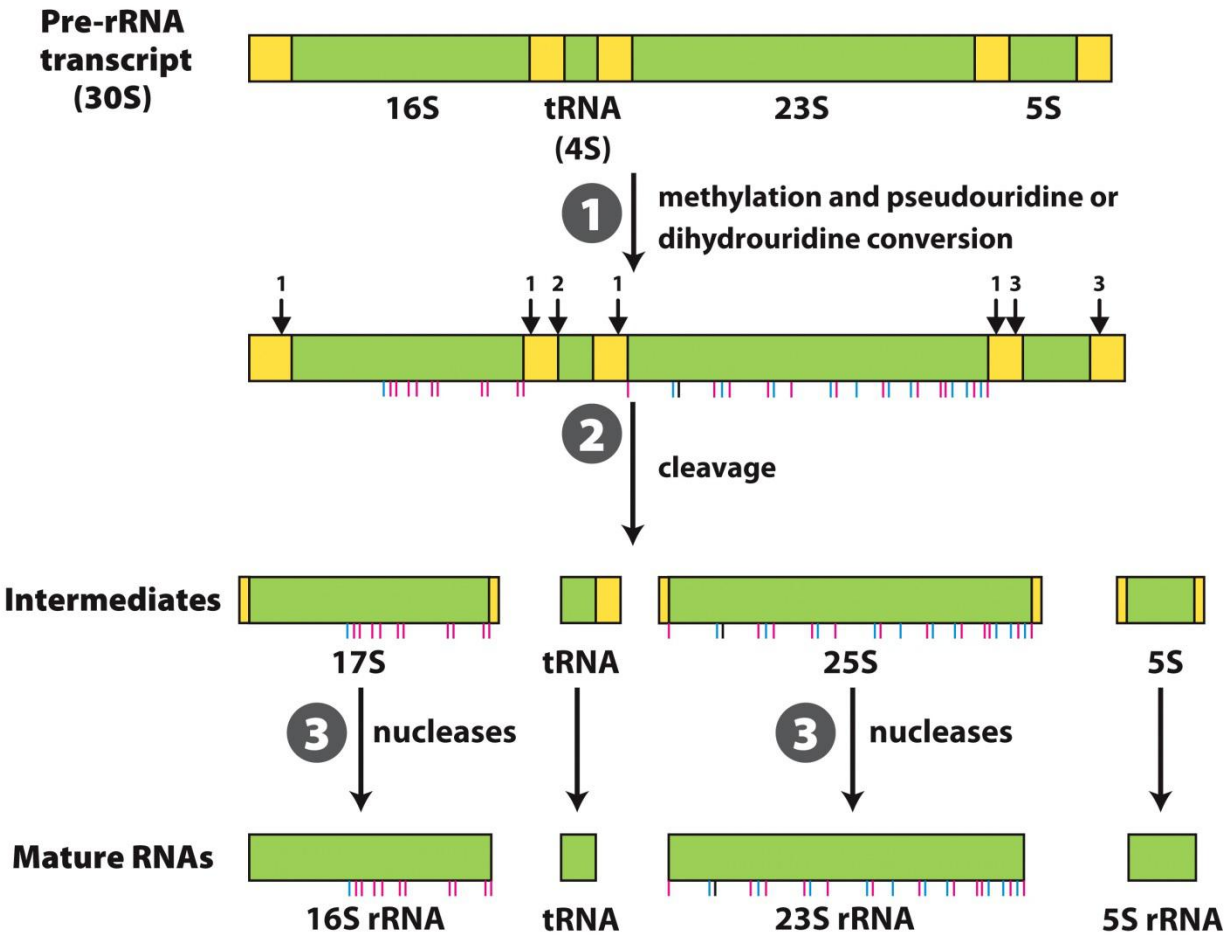


CGRP (calcitonin-gene-related peptide)

Different splicing patterns



Processing of pre-rRNA transcripts in bacteria

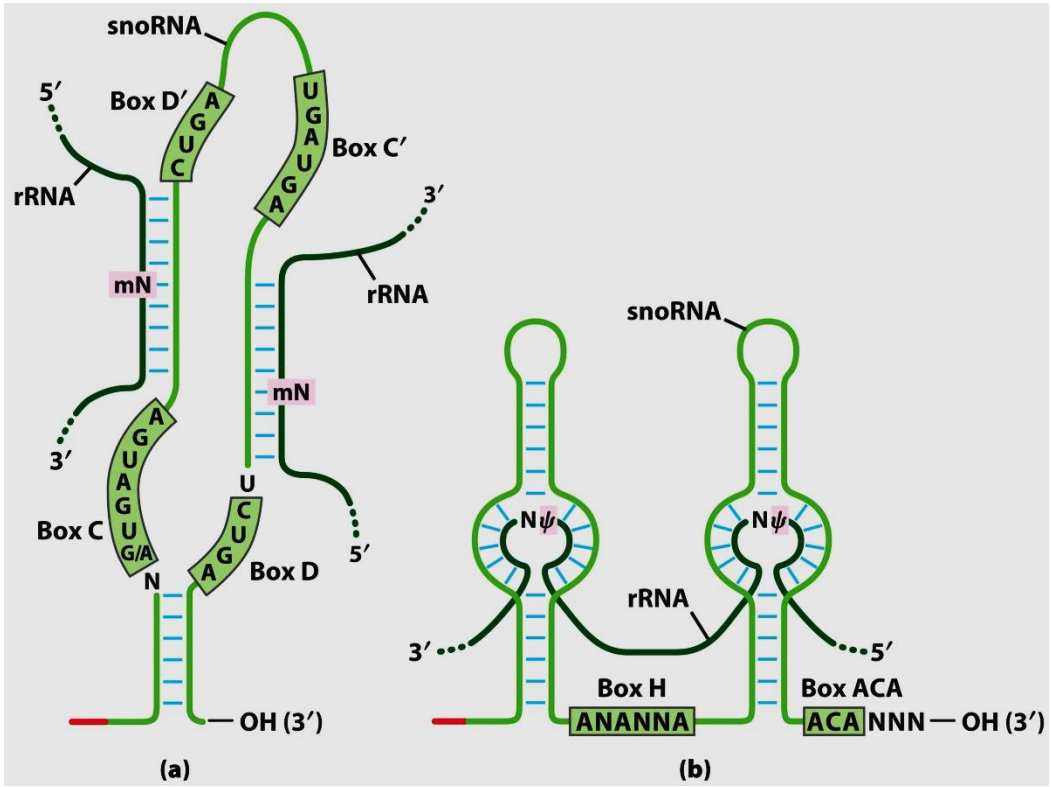


1) Before cleavage, the 30S RNA precursor is methylated at specific bases.

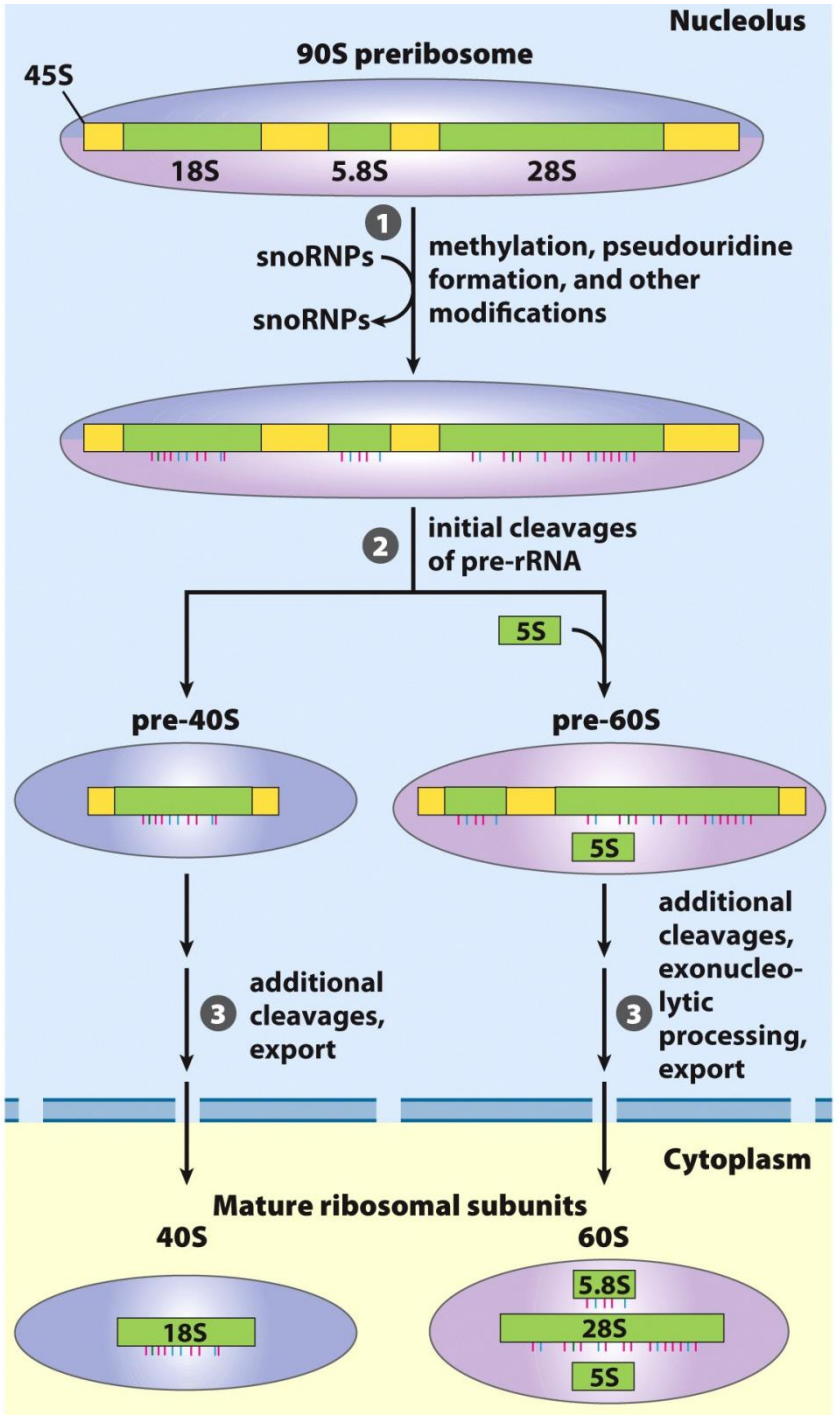
2) Cleavage liberates precursors of rRNA and tRNA.

3) The final 16S, 23S, and 5S rRNA products result from the action of a variety of specific nucleases.

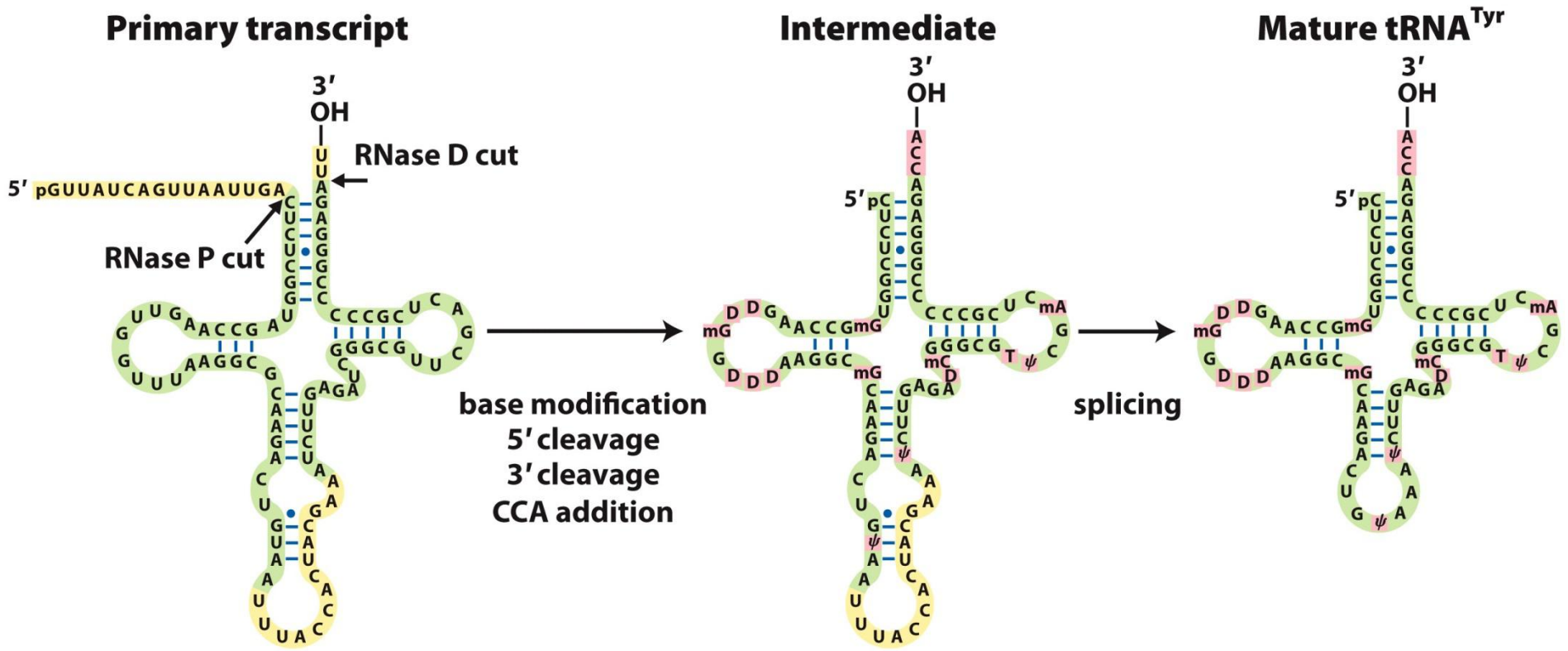
Processing of pre-rRNA transcripts in vertebrates



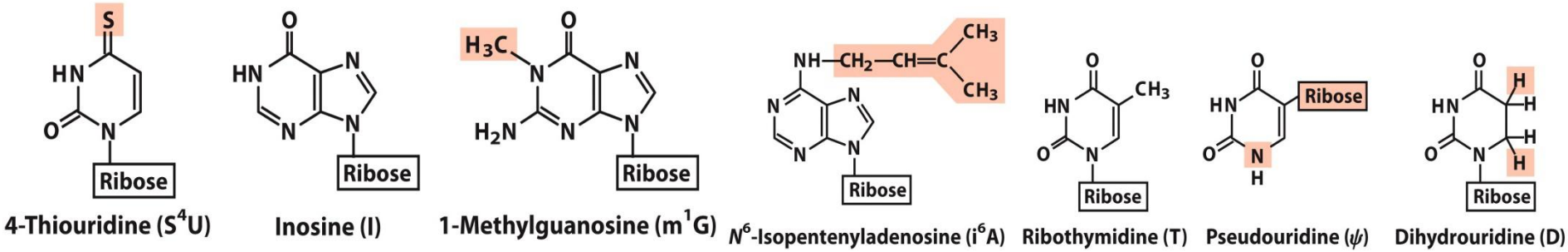
snoRNAs (small nucleolar RNAs) guide necloside modification and some cleavage reactions in the snoRNA-protein complex (snoRNPs).



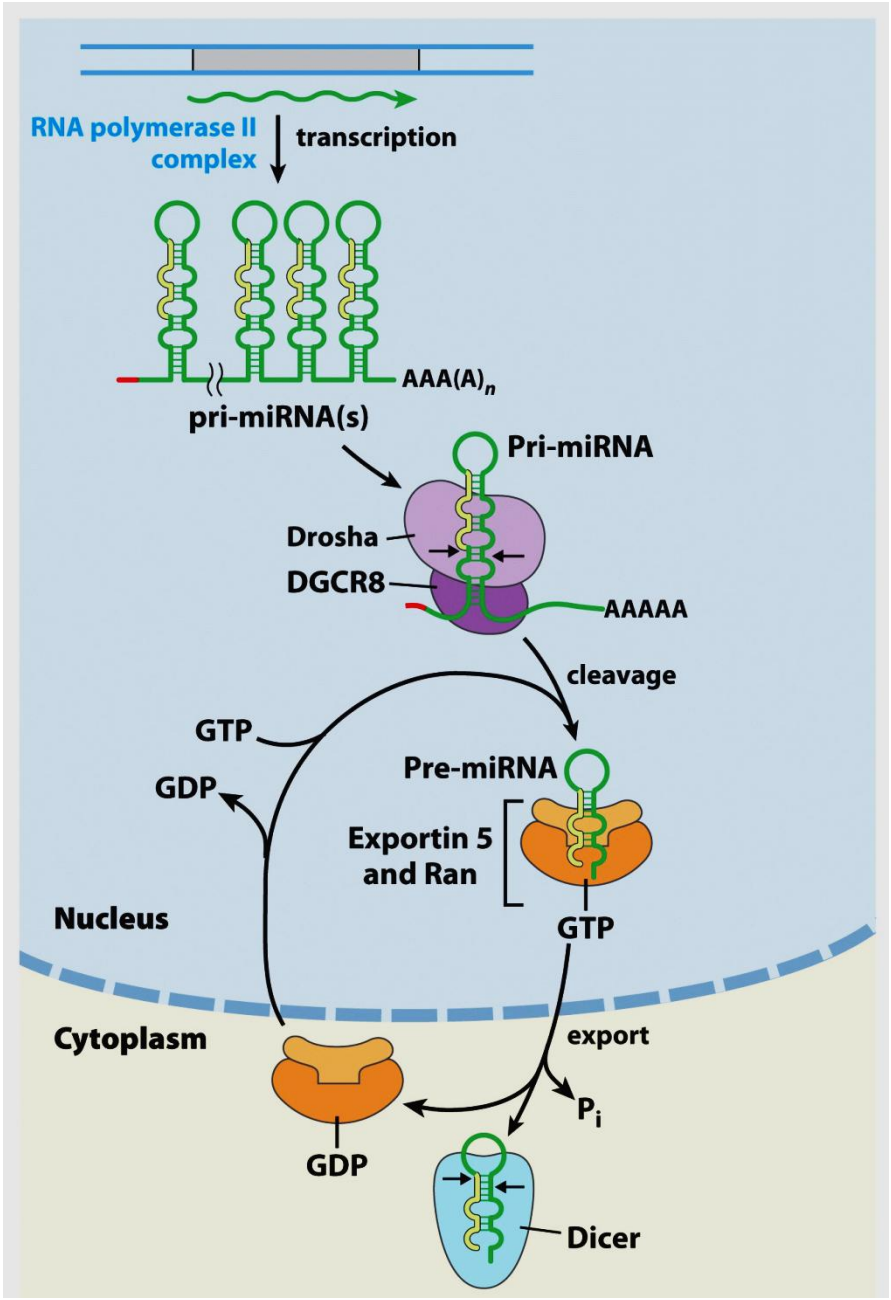
Processing of tRNAs in bacteria and eukaryotes

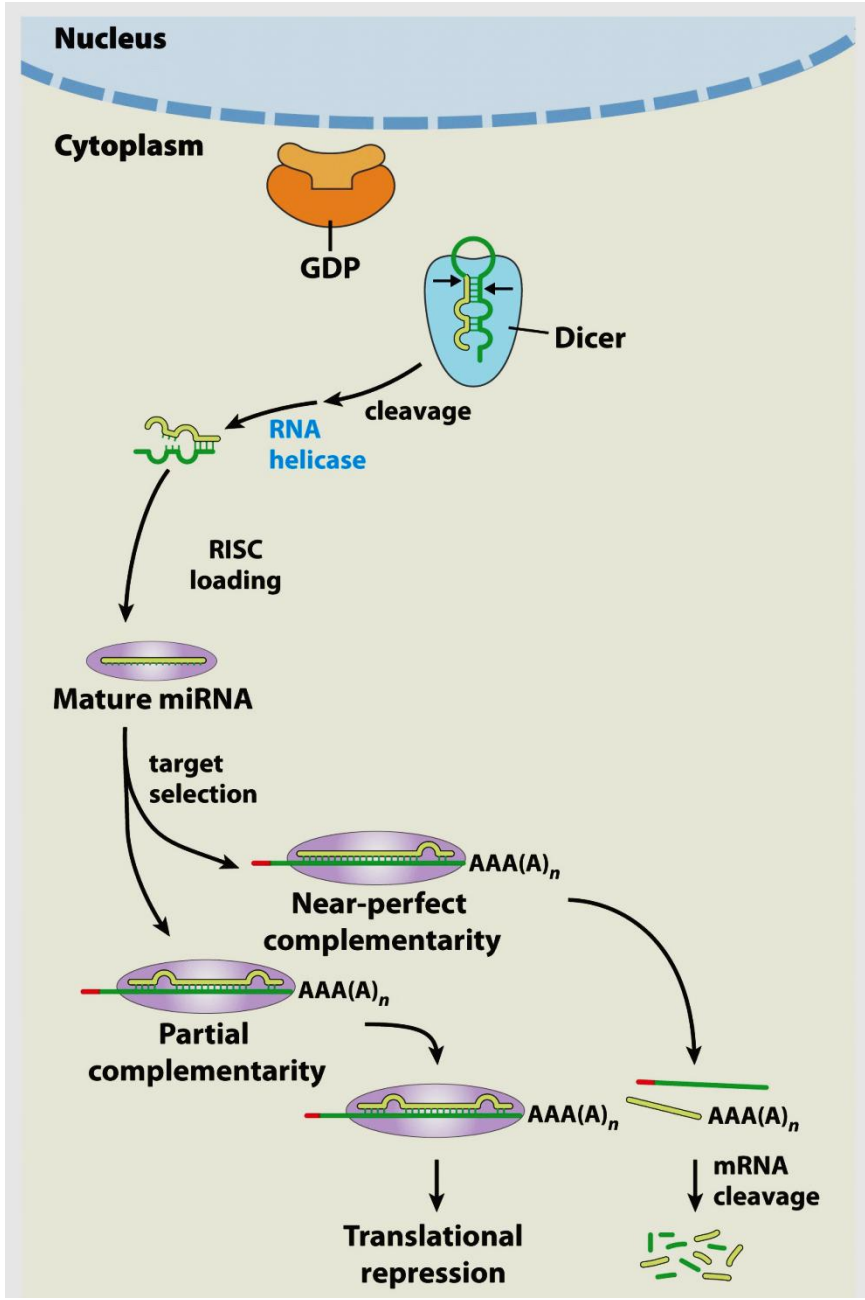


Some modified bases of rRNAs and tRNAs



Synthesis and processing of miRNAs





Degradation of mRNAs in cells

- The rates of degradation vary greatly for mRNAs from different eukaryotic genes.
- Messenger RNA is degraded by ribonucleases present in all cells:

  In *E. coli*, the process begins with one or a few cuts by an endoribonuclease, followed by 3' →5' degradation by exoribonucleases.

  In lower eukaryotes, the major pathway involves first shortening the poly(A) tail, then decapping the 5' end and degrading the mRNA in the 5' → 3' direction.

  A 3' →5' degradative pathway also exists and may be the major path in higher eukaryotes.

- All eukaryotes have a complex of up to 10 conserved 3' →5' exoribonucleases, called the exosome, which is involved in the processing of the 3' end of rRNAs and tRNAs, as well as the degradation of mRNAs.