

Topic 9

Transcription

Nelson & Cox (2016)

Chapter 26 pp. 1057-1087

Nelson & Cox (2017)

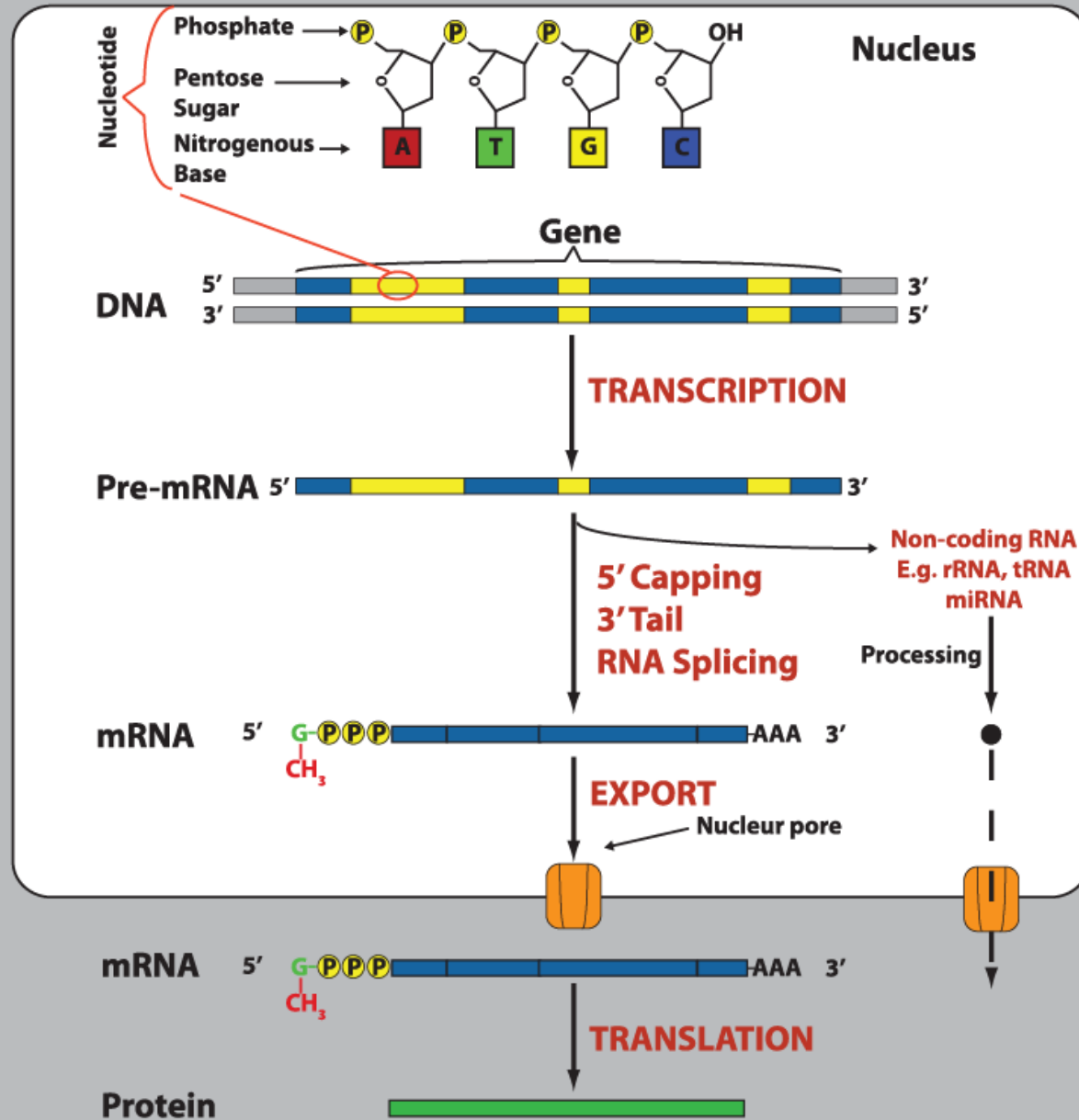
Chapter 26 pp. 1035-1065

Central Dogma

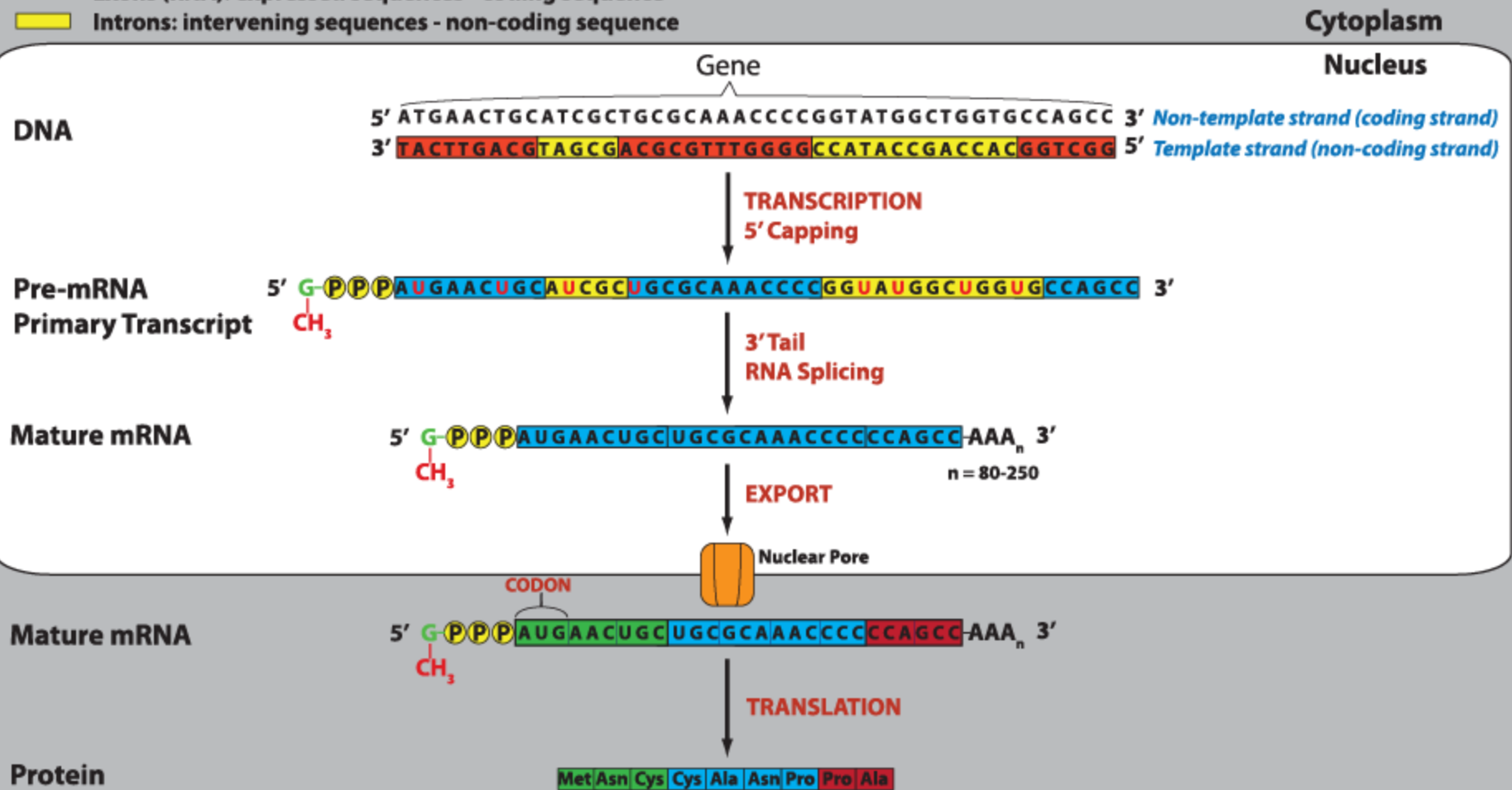
- DNA gives rise to new DNA via replication
- DNA is converted to RNA via transcription
 - HOWEVER: RNA can be converted to DNA via reverse transcription
- RNA is converted to protein via translation

- Exons: expressed sequences - coding sequence
- Introns: intervening sequences - non-coding sequence

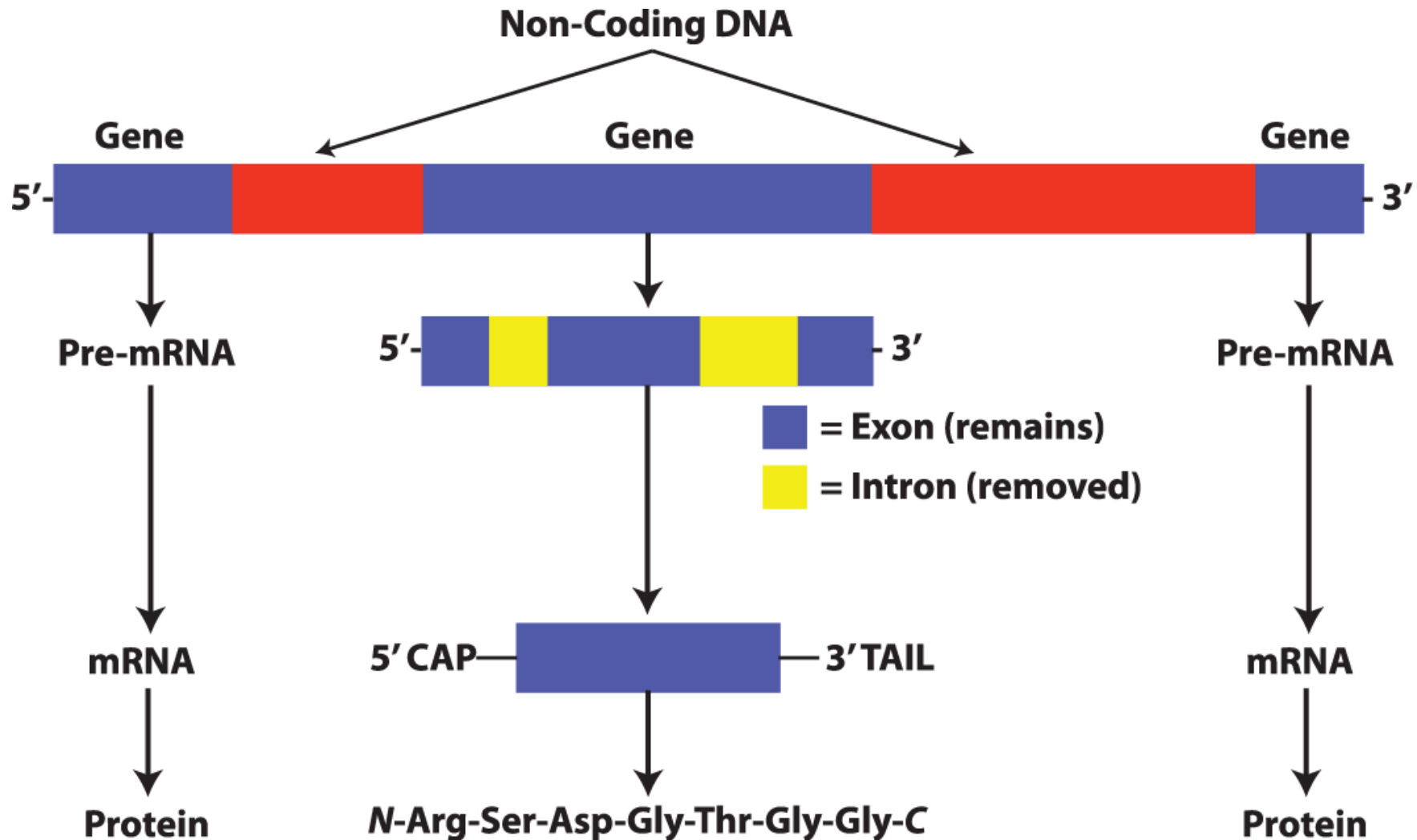
Cytoplasm



- Exons (DNA): expressed sequences - coding sequence
- Exons (RNA): expressed sequences - coding sequence
- Introns: intervening sequences - non-coding sequence



A Genome is Comprised of Thousands of Genes



Approx. 19,000 to 22,000 protein-coding genes in the human genome

Terminology

- **Genome:** entirety of the hereditary genetic information of a cell
- **Transcriptome:** sum of all RNA molecules produced in a cell at any given time
- **Gene:** a chromosomal segment of DNA that CAN encode for a specific biologically functional polypeptide or RNA molecule, or set of closely related isoforms
- **Transcription:** is the process of creating a complementary RNA copy of a sequence of DNA
- **Translation:** the process by which an mRNA sequence is decoded (converted) into a specific amino acid sequence in the construction of a polypeptide chain

Transcription Vs Replication

7

- **Similarities with Replication:**
 - DNA template
 - Direction of synthesis is $5' \rightarrow 3'$
 - 3 phases: initiation, elongation, termination
- **Differs from Replication:**
 - Only 1 DNA strand acts as a template
 - No primers used
 - Only small areas of DNA are used as a template
 - Newly synthesised strand is RNA not DNA

Numbering Base-pairs in Genes

8

Start sequence for transcription (e.g. ATG)

5'-T G C A C **A T G** C T C A G T C A C-3'



← -1, +1, +2 →

Left of start
= upstream

Right of start
= downstream

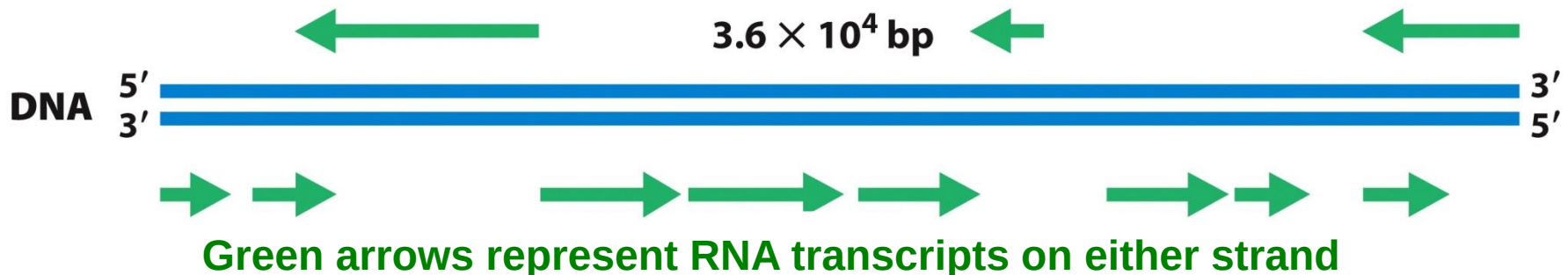
Strand Terminology

- Template (non-coding) DNA strand: serves as template for RNA synthesis
- Non-template (coding) DNA strand: identical in sequence to the RNA transcribed (**U replaces T**)
- A genes template can be located on either strand

(5') C G C T A T A G C G T T T (3') DNA nontemplate (coding) strand

(3') G C G A T A T C G C A A A (5') DNA template strand

(5') C G C U A U A G C G U U U (3') RNA transcript



Initiation

- Occurs within the “Promoter” of a gene
 - Specific sequence of DNA
 - *E. coli* promoters typically span a region of -70 to +30 bp's
 - Contain consensus sequences
 - @-10bp: 5'-TATAAT-3' (TATA Box)
 - @-35bp: 5'-TTGACA-3'
 - @-40 to -60bp: AT rich region (upstream promoter (UP)- highly expressed genes)
 - The spacing is particularly important, as this is associated with RNA polymerase binding



Promotor Region



Consensus
sequence



rrnB P1



trp



lac



recA



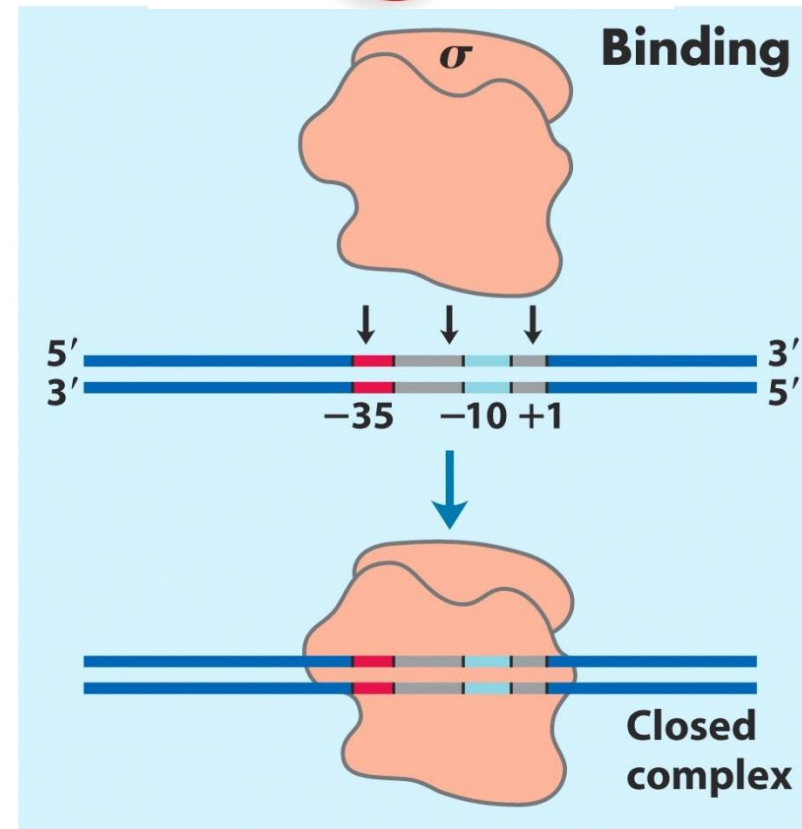
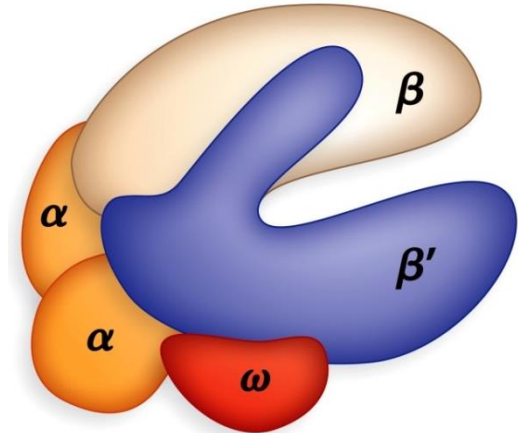
araBAD



DNA-Dependent RNA Polymerase

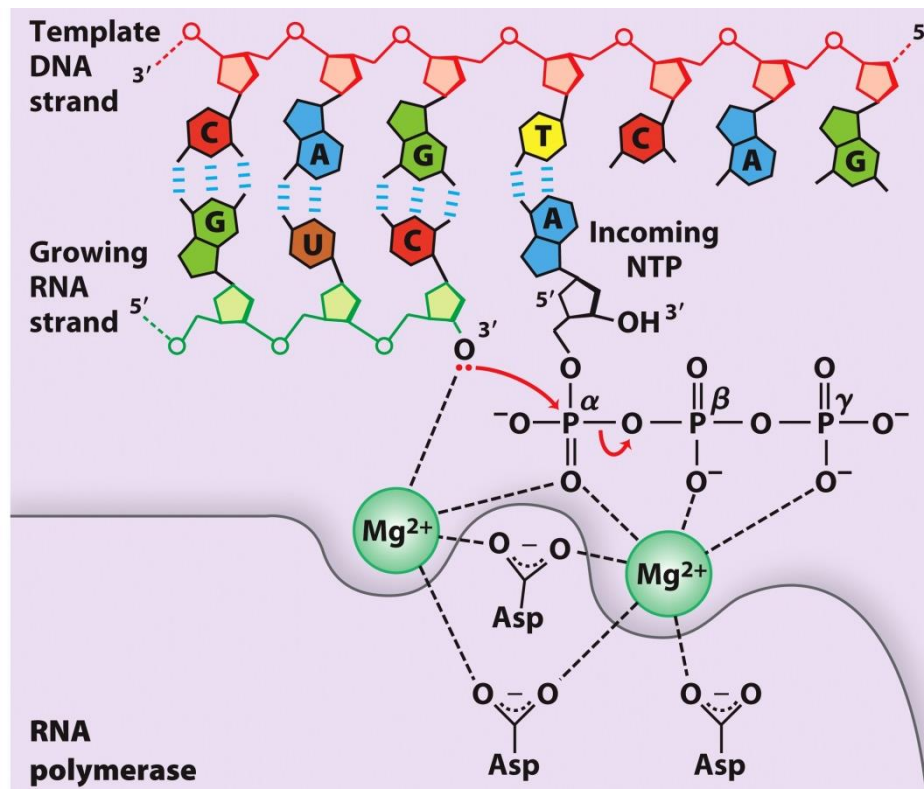
- 6 subunits
 - 1 x σ (sigma)
 - Binds temporarily to core subunits
 - Binding directs RNA Pol complex to promoter sites
 - Different sigma subunits recognise different promoter regions, most common subunit is σ^{70}
 - 5 core subunits: α_2 , β , β' , ω (omega)
 - Catalytic activity without σ
 - No true proofreading, higher mistake rate (10^4 - 10^5)

Why is the error rate not as critical for RNA Pol?



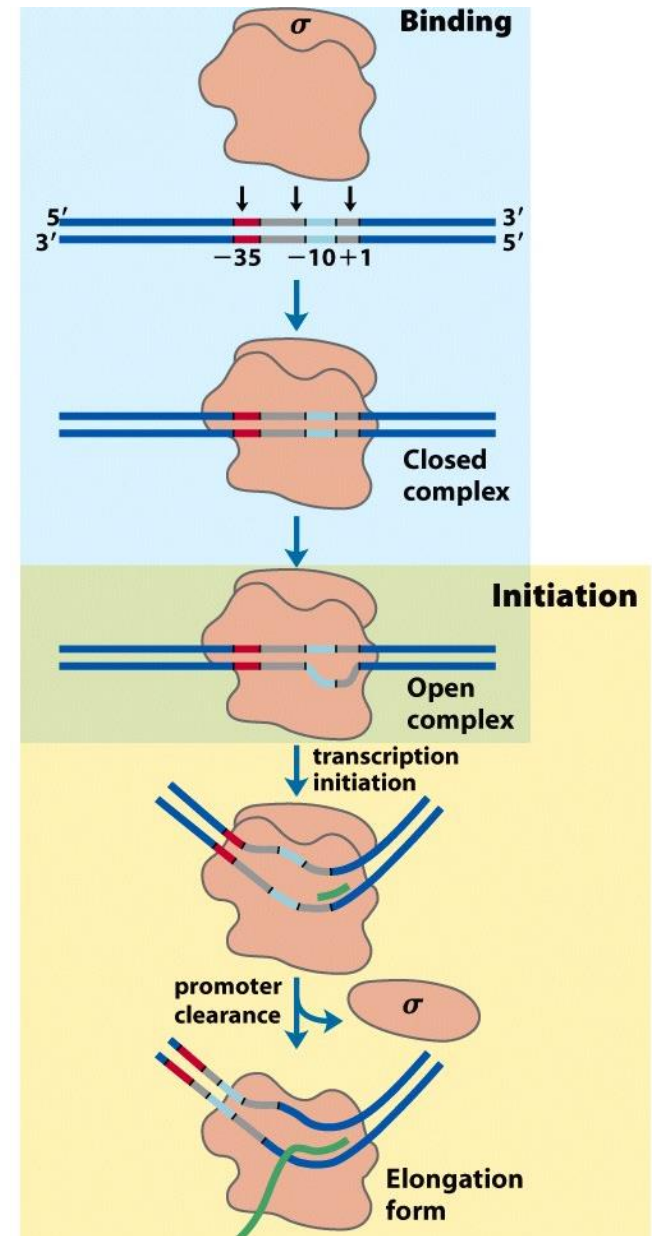
DNA-Dependent RNA Polymerase

- Attaches successive nucleotides (**NTPs NOT dNTPs**) in the same manner as DNA polymerase (**but does not cleave first triphosphate – 5' cap**)
- Same polarity of growth $5' \rightarrow 3'$, but adds U instead of T
- Uses DNA as template
- Does not require:
 - Primer, Helicase (is part of polymerase), DNA binding proteins.



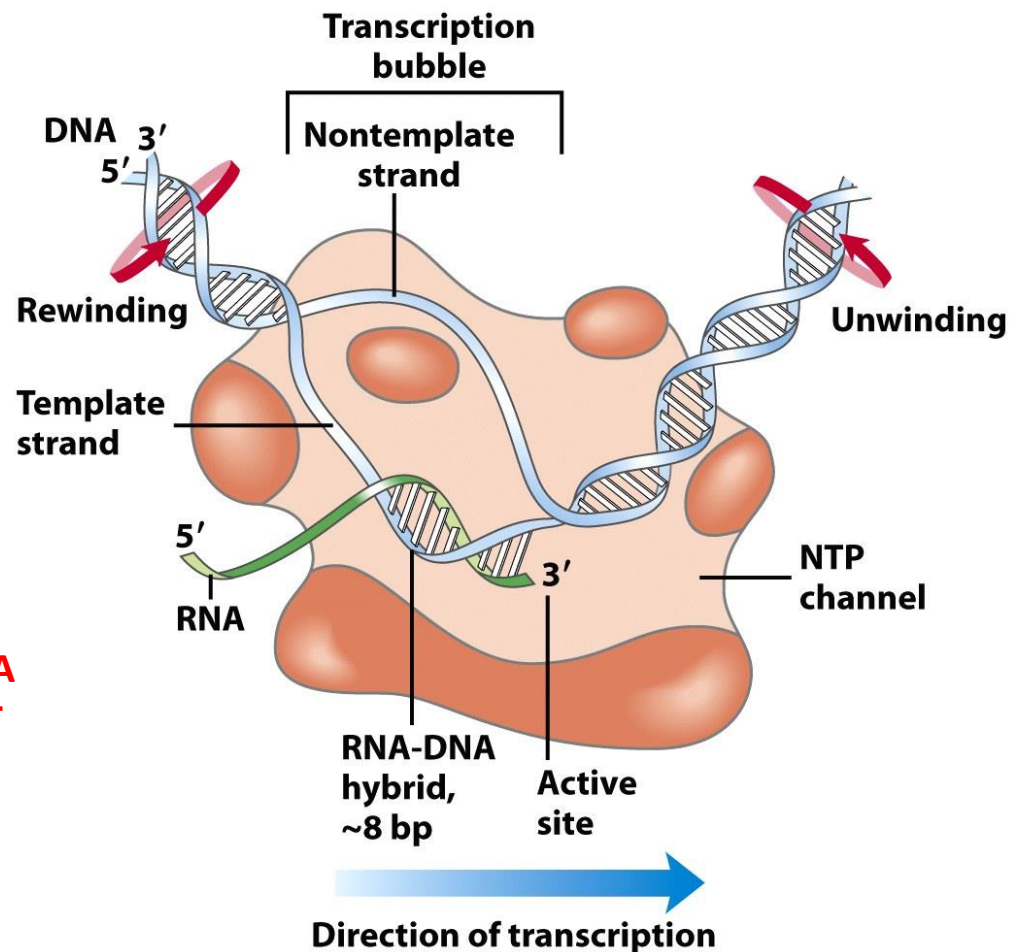
Initiation of Transcription

- Major steps
 - Binding to promoter (closed complex)
 - Binding (open complex – unwound 12-15bp)
 - Initiation (conformational change - σ subunit detaches)
 - Initiation (promoter clearance – movement of transcription complex away from the promoter)



Elongation

- Nucleotides added in 5'→3' direction to free 3'-OH group
 - NTPs not dNTPs**
- Polymerase unwinds 17bp's at a time
- An 8 bp RNA-DNA hybrid double helix forms at any one time
- RNA Pol moves at a rate of 50-90 nucleotides/sec
 - ATGGCGTATTATTGCCGGGGGCCGTGAATA**
AATGTCCAAAAACCGGTTCCGGCCGGTTGT
GTGCCACTTCCCGCGGGGTATATAGGCCT
- Topoisomerase prevents supercoiling



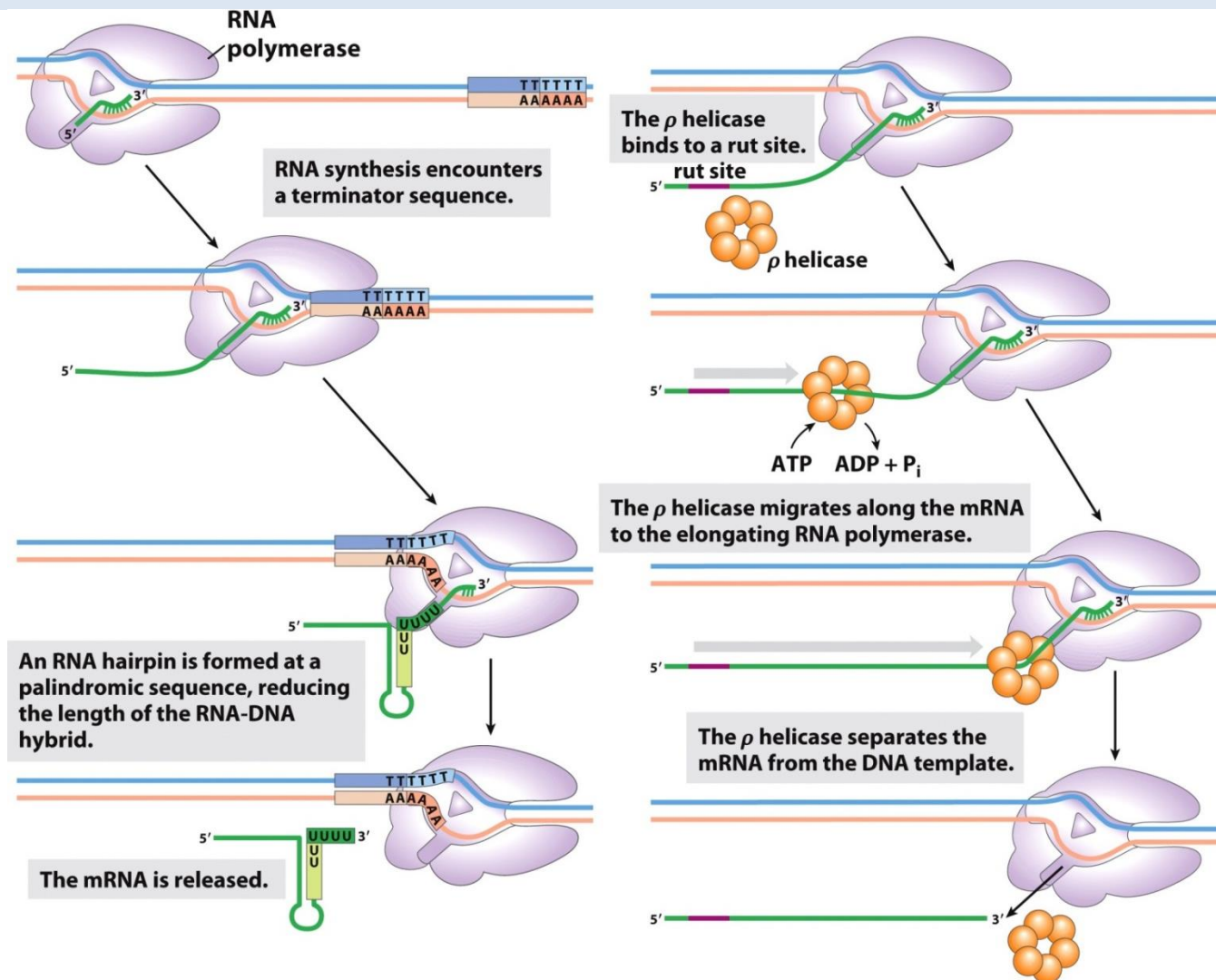
DNAnts 5'TACGGCAT3'

DNAts 3'ATGCCGTA5'

RNA 5'UACGGCAU3'

U in RNA, T in DNA

Termination

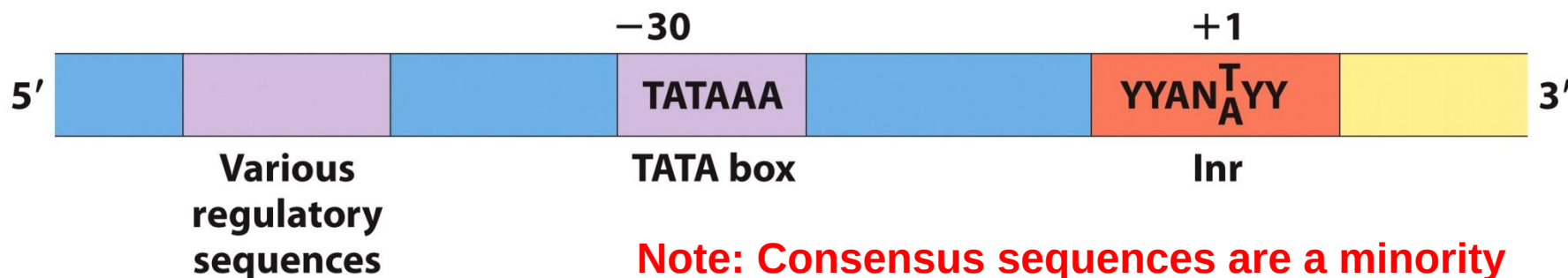


Rho-independent:
Presence of hairpin releases RNA
from RNA Pol.

Rho-dependent:
Involves binding of Rho protein,
initiated at the *rut* element

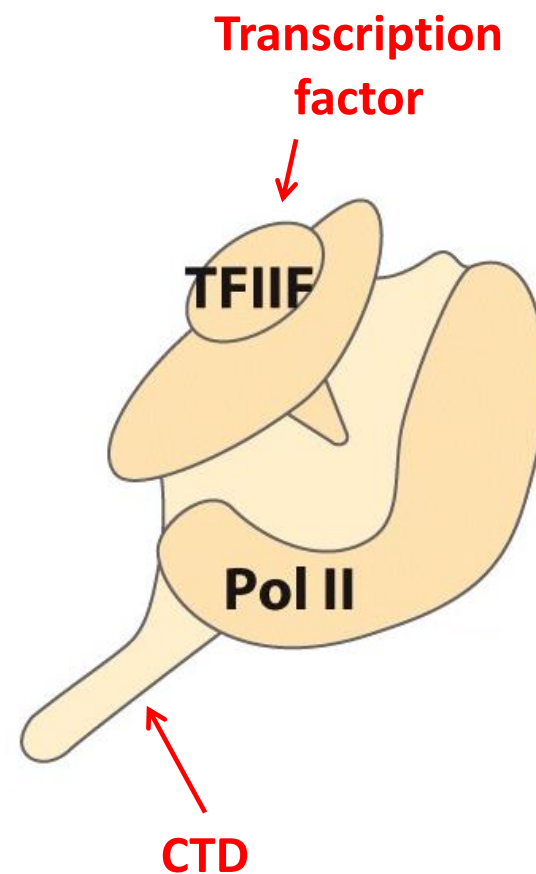
Eukaryotic Transcription

- The general process is the same, simply more complex
- 3 RNA pol's (I, II, III)
- RNA Pol I: produces pre-rRNA (18S, 5.8S and 28S)
- RNA Pol II: produces mRNA and specialised RNAs
 - Large variety of promoters recognised
 - Some common consensus sequences – e.g. TATA (TATA box, -30bp) and initiator (Inr) sequence (+1bp)
- RNA Pol III: produces tRNA, 5S rRNA, specialised RNAs



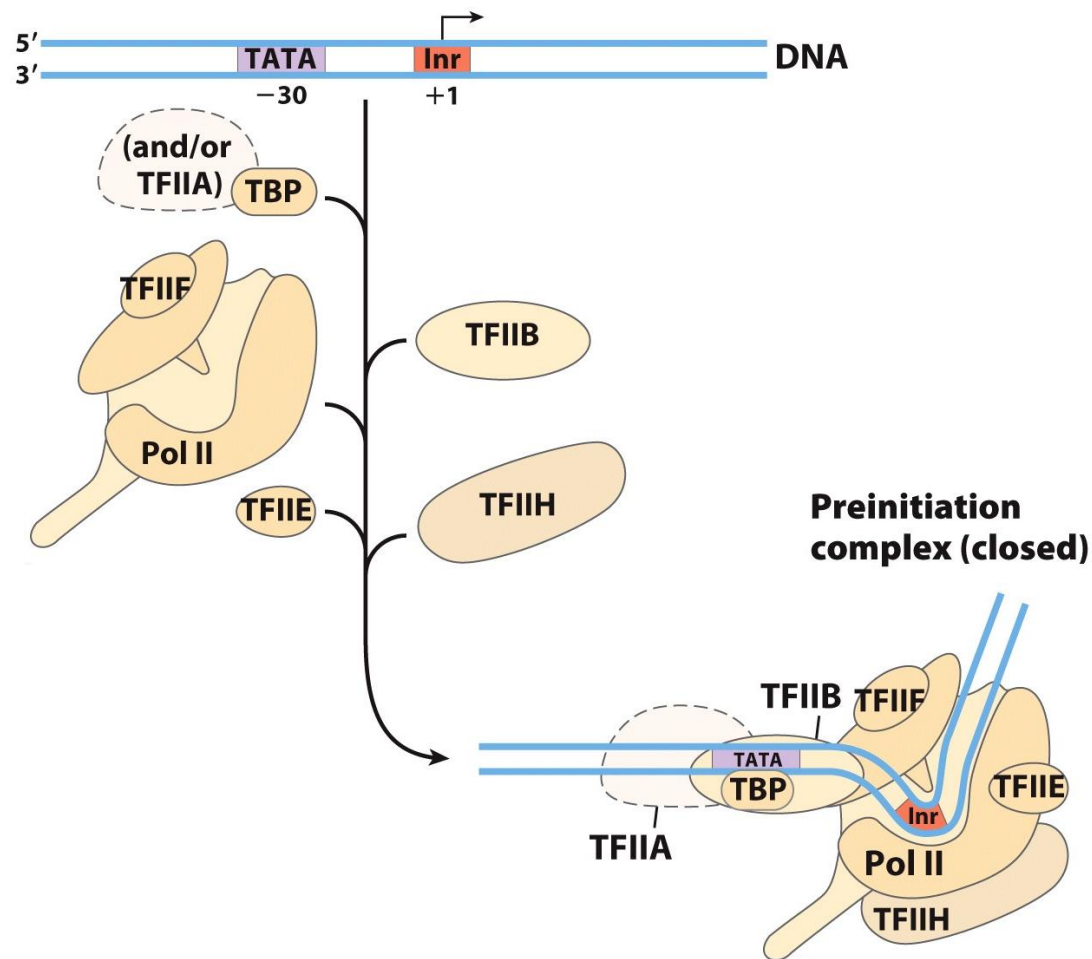
RNA Polymerase II

- Contains 12 subunits (RPB 1-12); of note, RPB1 contains a long carboxyl-terminal domain (CTD) separate from the main body that is important for Pol II functioning
- Greater complexity, due to the complexity of genome and its packaging/association with numerous proteins
- RNA Pol II requires numerous transcription factors for proper functioning (formation of an active transcription complex)
- Naming transcription factors for Pol II: **TFII**, followed by an additional identifier e.g. **TFIIF**



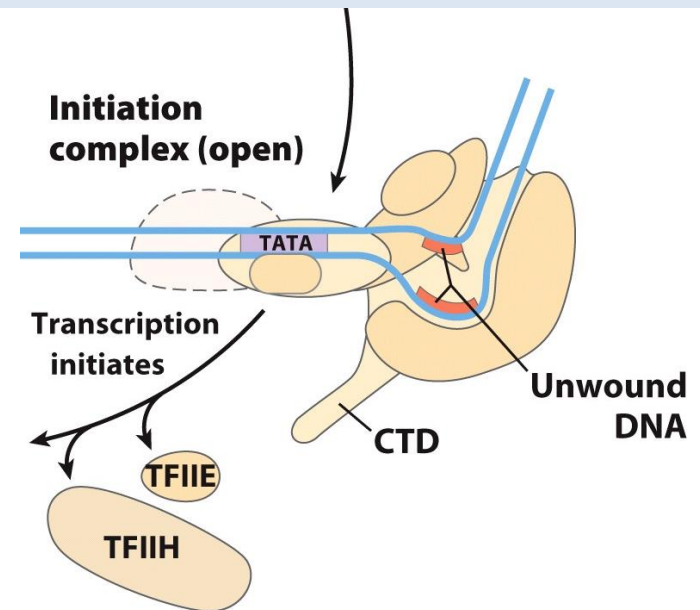
RNA Polymerase II – Pre Initiation

1. Closed complex formation is initiated by the binding of TATA-binding protein (**TBP**) to the **TATA box**.
2. If lacking a TATA box, TBP arrives as part of the complex **TFIID**.
3. Both **TFIIA** and **TFIIB** bind TBP, stabilising its interactions with the dsDNA
4. TFIIB provides a connection point for the attachment of **Pol II**, which has **TFIIF** (helps target promoter) bound to it.
5. Finally **TFIIE** and **TFIIH** (helicase activity –unwinds DNA at promoter) bind to Pol II to create the **closed complex**

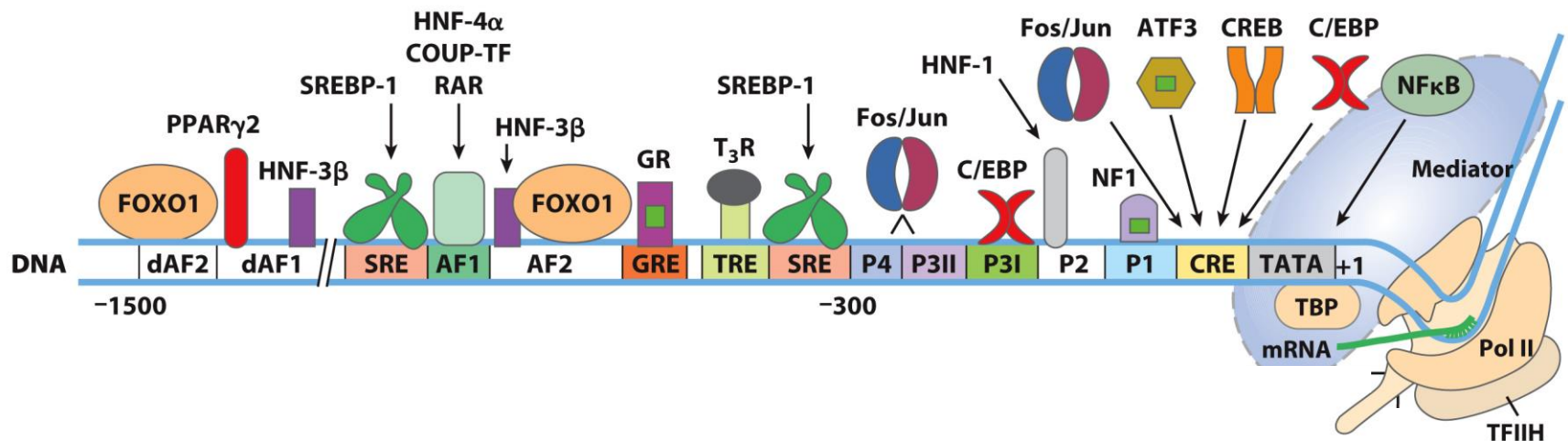


RNA Polymerase II - Initiation

1. Bound **TFIIE/TFIH** unwinds 12-15 nucleotides of the DNA helix
2. A subunit of **TFIIH** phosphorylates the **CTD** inducing a conformational change which initiates transcription.
3. During the initial synthesis of 60-70 nucleotides, **TFIIE** and **TFIIH** are released.



Complexity of Regulation – PEP Carboxykinase Promoter



Transcription

RNA Poly II – Elongation and Termination

1. During elongation, **TFIIF** remains bound to Pol II
2. Polymerisation (elongation) is enhanced by the binding of **elongation factors** (important in preventing pausing of transcription and in allowing post-transcriptional modifications in association with the **CTD**).
3. Once elongation has terminated, Pol II is de-phosphorylated and recycled.

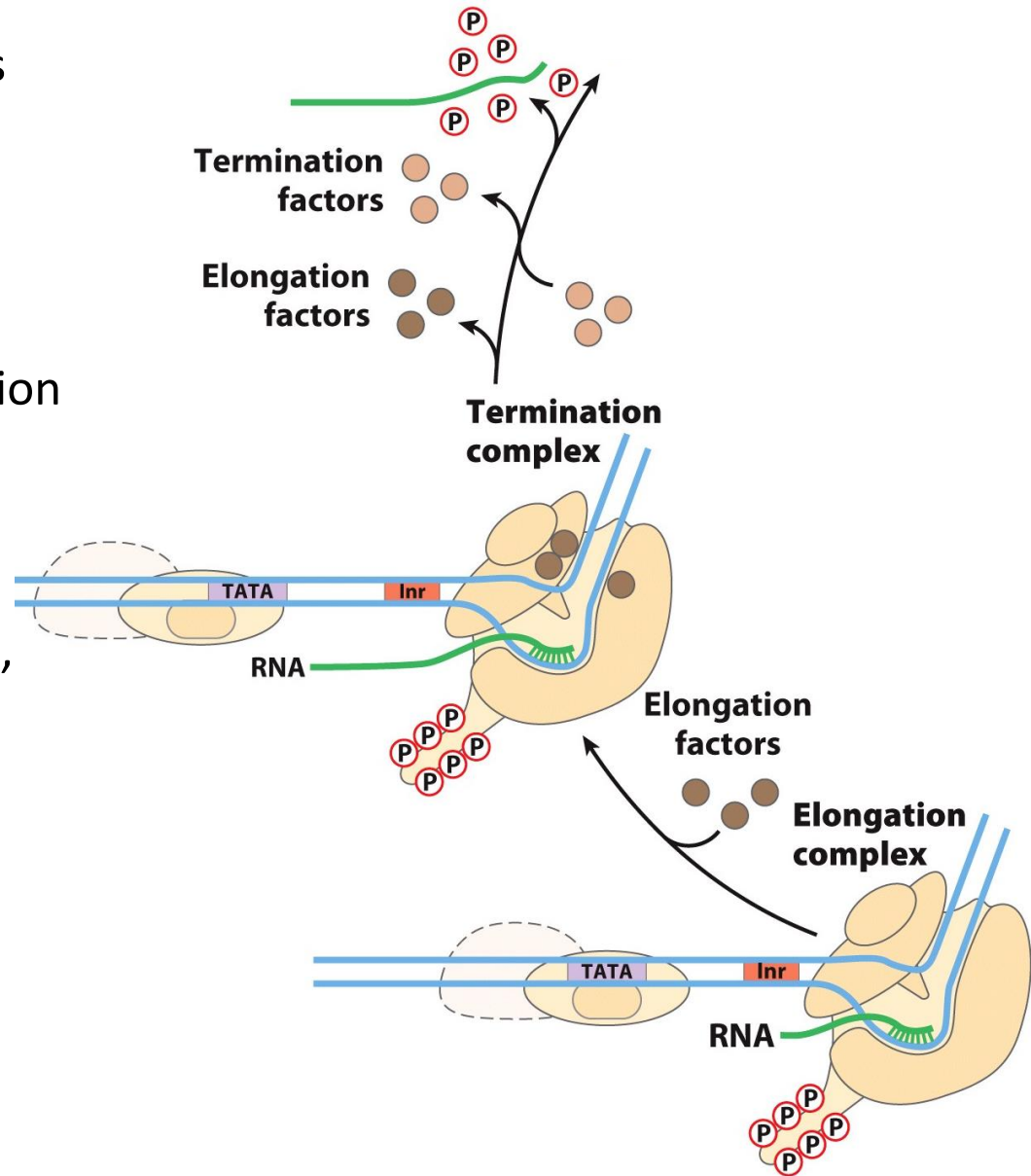


TABLE 26-2

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

Transcription protein	Number of different subunits	Subunit(s) M_r^a	Function(s)
Initiation			
Pol II	12	7,000–220,000	Catalyzes RNA synthesis
TBP (TATA-binding protein)	1	38,000	Specifically recognizes the TATA box
TFIIA	2	13,000, 42,000	Stabilizes binding of TFIIB and TBP to the promoter
TFIIB	1	35,000	Binds to TBP; recruits Pol II–TFIIF complex
TFIID ^b	13–14	14,000–213,000	Required for initiation at promoters lacking a TATA box
TFIIE	2	33,000, 50,000	Recruits TFIIH; has ATPase and helicase activities
TFIIF	2–3	29,000–58,000	Binds tightly to Pol II; binds to TFIIB and prevents binding of Pol II to nonspecific DNA sequences
TFIIH	10	35,000–89,000	Unwinds DNA at promoter (helicase activity); phosphorylates Pol II (within the CTD); recruits nucleotide-excision repair proteins
Elongation^c			
ELL ^d	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	

^a M_r reflects the subunits present in the complexes of human cells. Some components differ somewhat in size in yeast.

^bThe presence of multiple copies of some TFIID subunits brings the total subunit composition of the complex to 21–22.

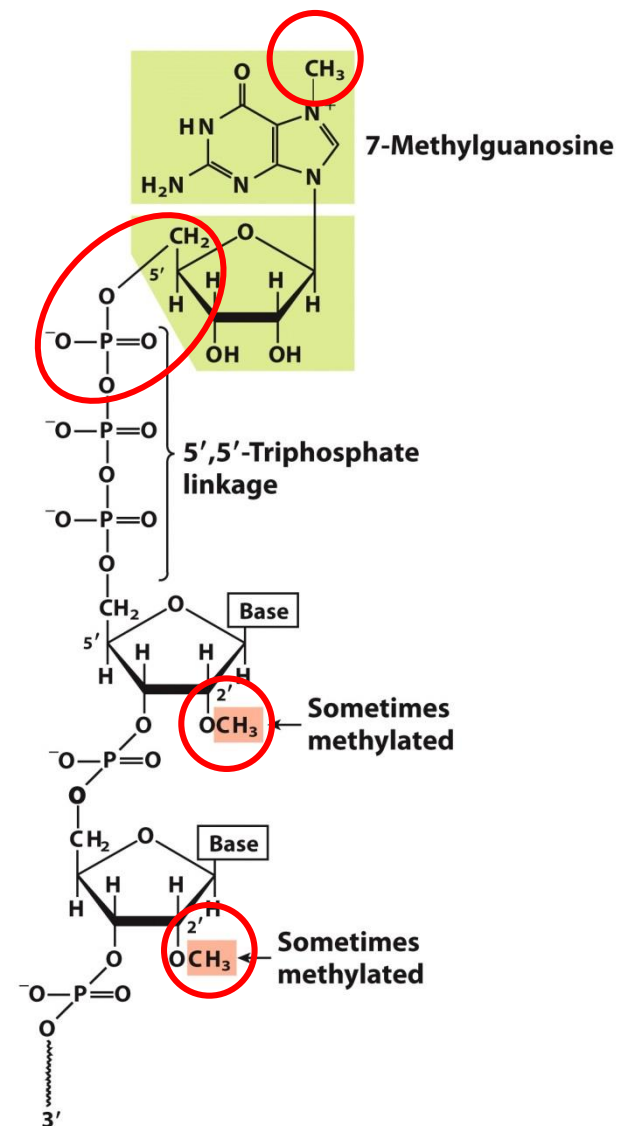
^cThe function of all elongation factors is to suppress the pausing or arrest of transcription by the Pol II–TFIIF complex.

^dName derived from eleven-nineteen /lysine-rich /leukemia. ELL is the site of chromosomal recombination events frequently associated with acute myeloid leukemia.

RNA Processing (Post-Transcriptional Modifications)

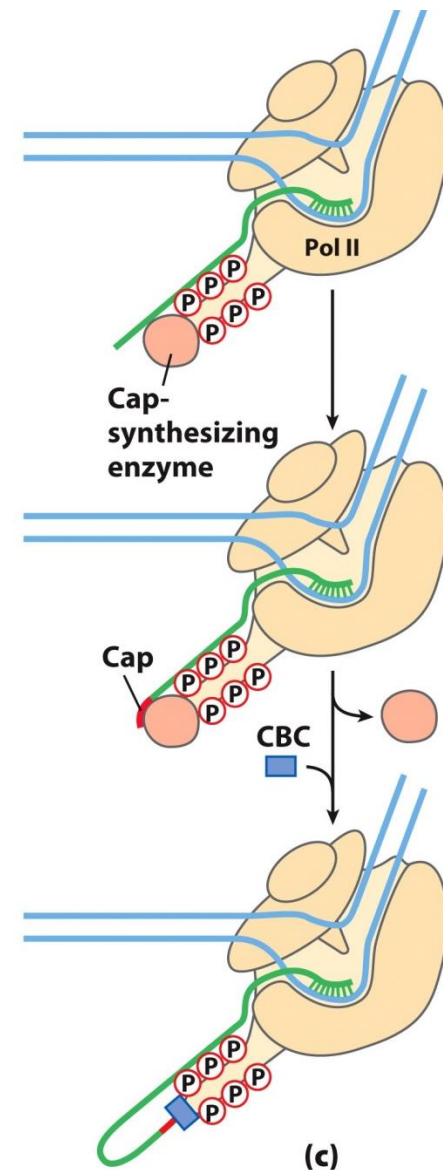
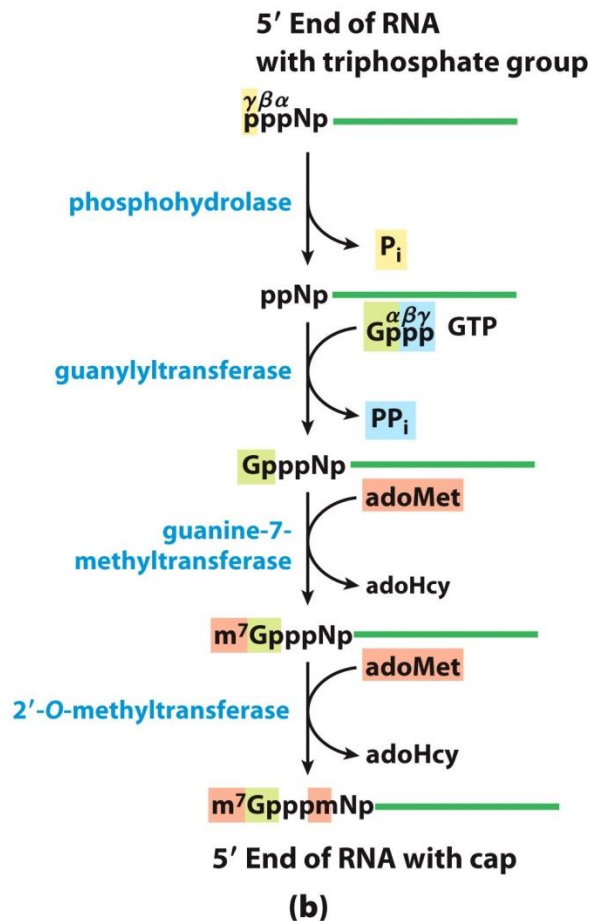
(Eukaryote mRNA)

- 5' Capping
 - Addition of 7-methyl guanosine to 5' terminal
- 3' poly(A) tail
 - Addition of 80-250 A's to 3' terminal mRNA
 - Polyadenylate polymerase
- Splicing
 - Removal of introns
 - Can be self-splicing or require snRNPs



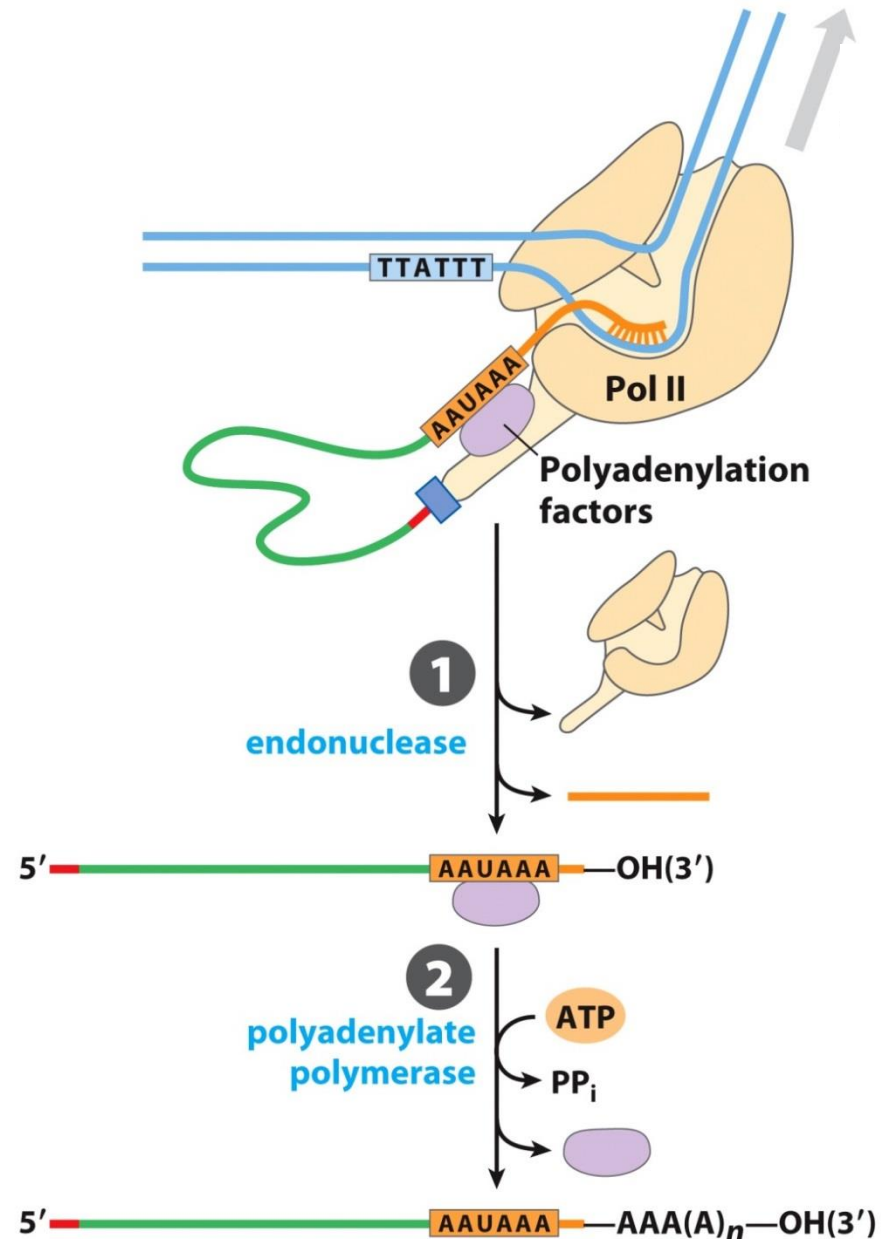
RNA 5' Capping

- A 7-methylguanosine molecule is added to the 5' end of the RNA via a **5',5'-triphosphate** linkage
- Is required for protection against ribonucleases
- Participates in binding to ribosome to initiate translation
- 5' capping occurs at CTD
 1. Condensation of GTP with 5' triphosphate of the RNA chain
 2. Methylation at N-7 of the guanine base
 3. Additional methylation of the first and second nucleotides at 2' hydroxyl group sometimes occurs



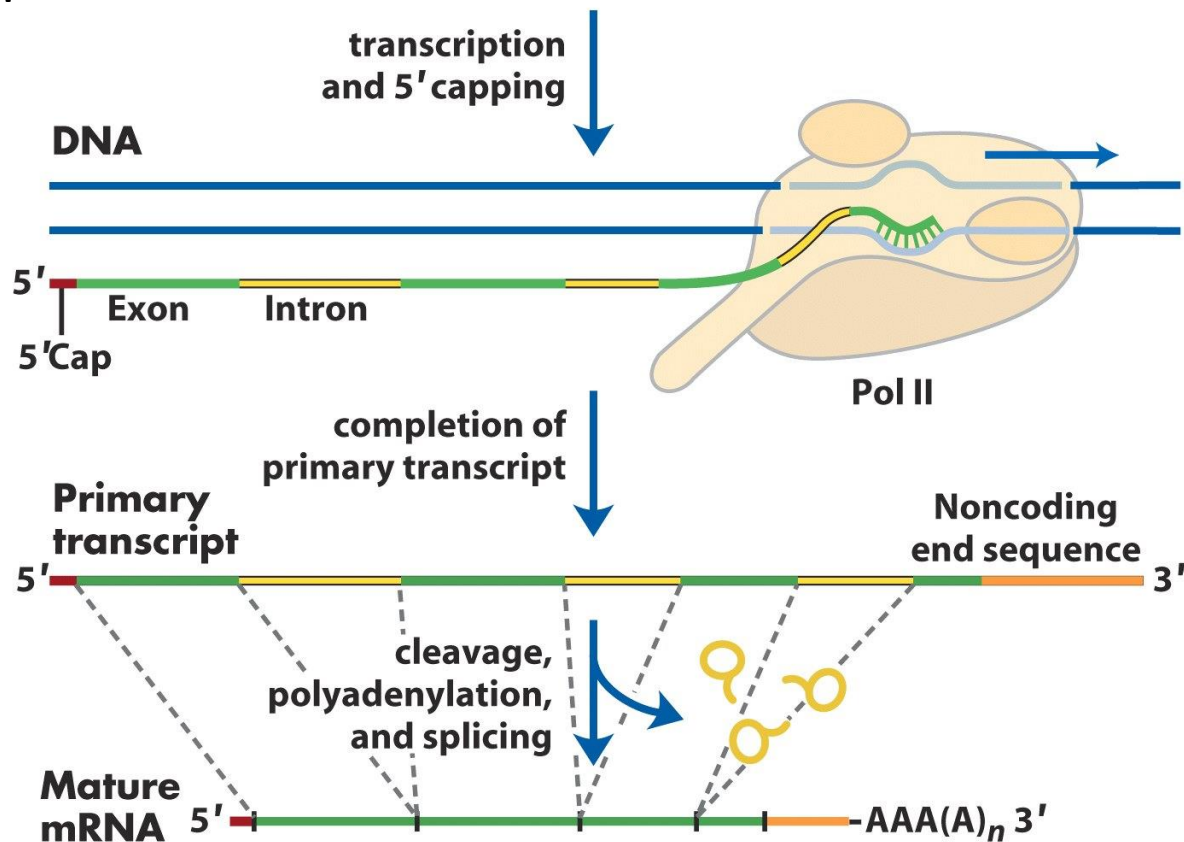
RNA Poly A Tail

- Transcript is extended beyond the site where the poly(A) tail will be added.
- Cleavage of RNA and addition of poly(A) tail is marked by sequence elements (e.g. AAUAAA)
- Addition of 80-250 adenosine residues
- Occurs at the CTD
- Helps protect against enzymatic degradation

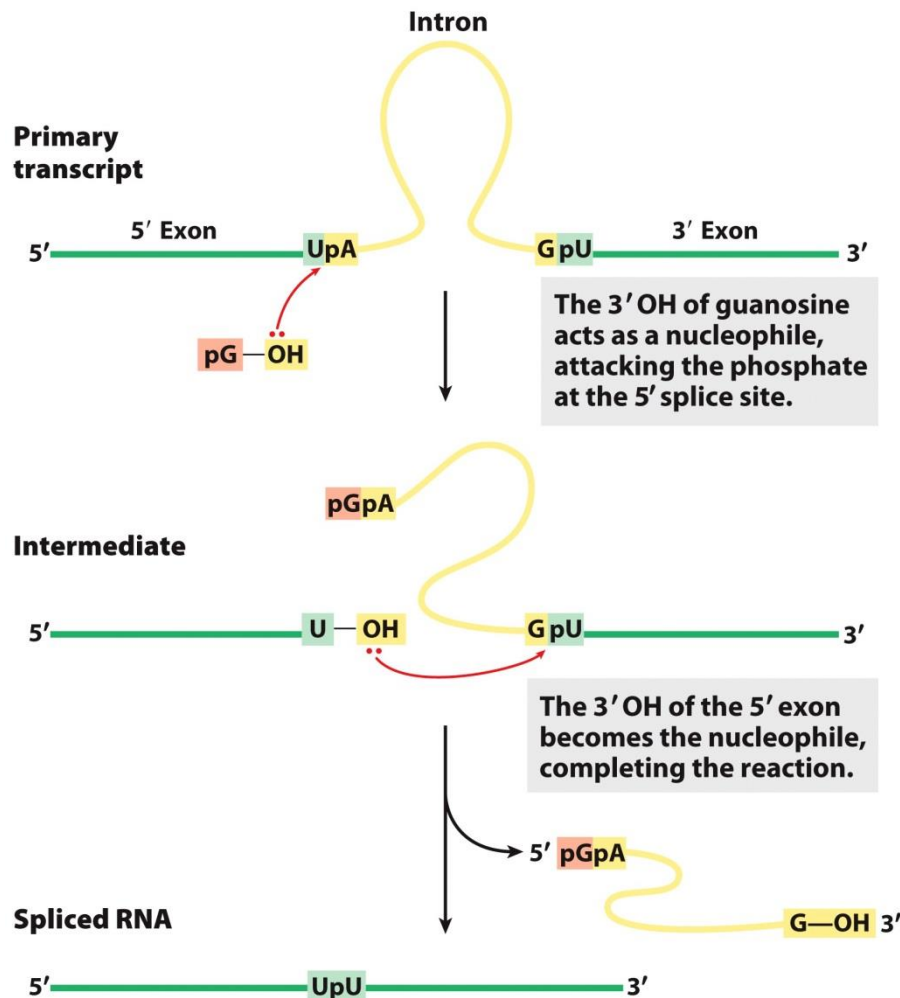


Splicing of Introns

- Self-splicing
 - Group I and II
- Require the enzyme spliceosome (more common)
 - Group III and IV

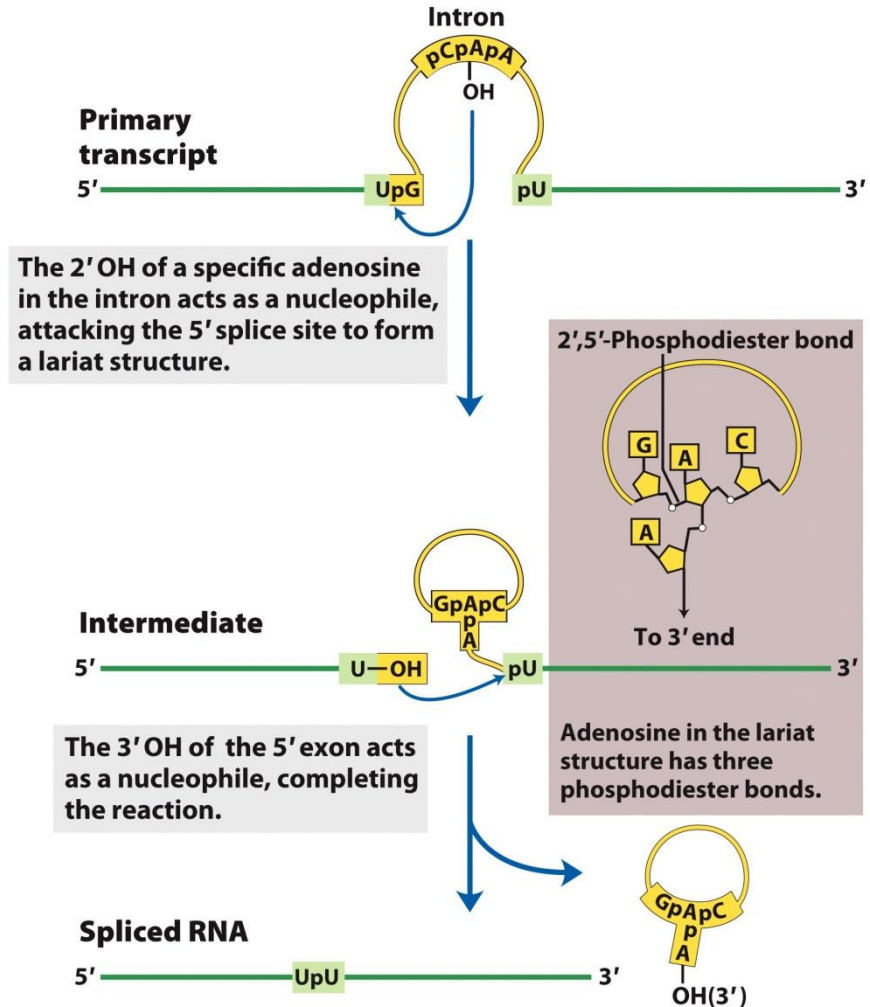


Group I



- Requires a guanine nucleotide
- Two trans-esterification reactions
- Formation of new phosphodiester bonds

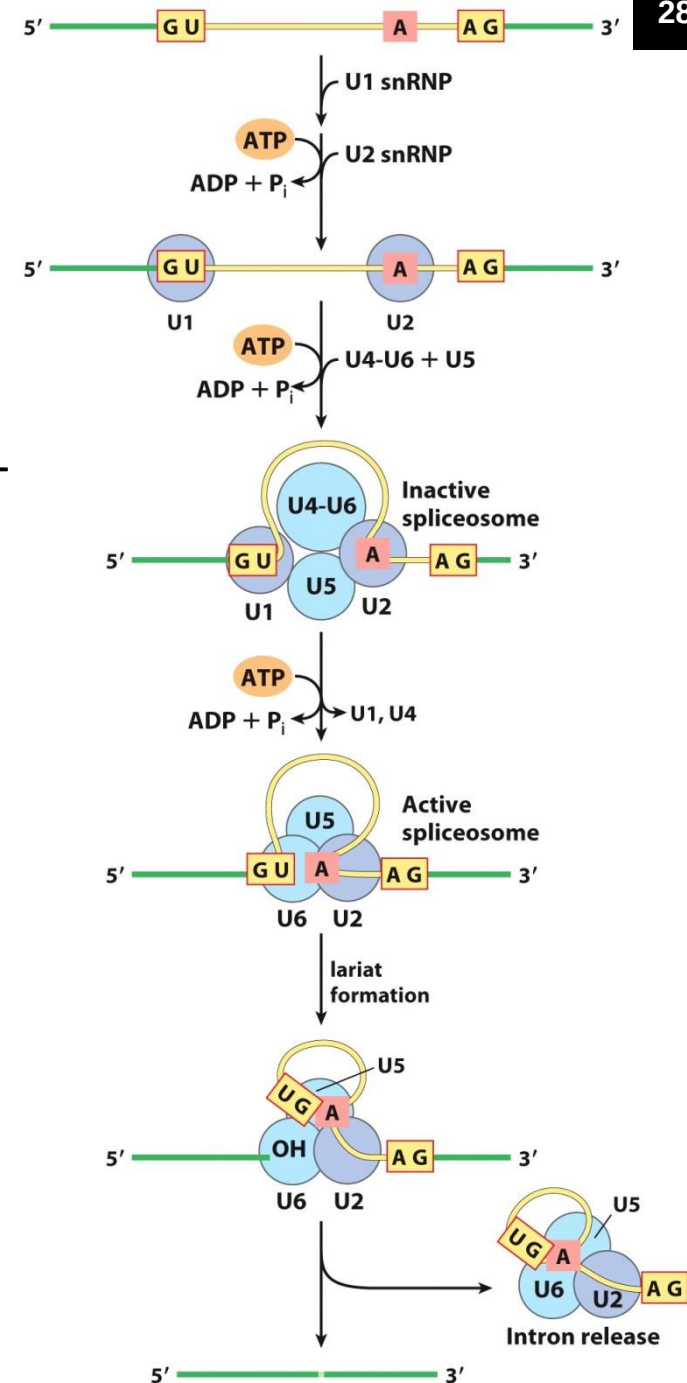
Group II



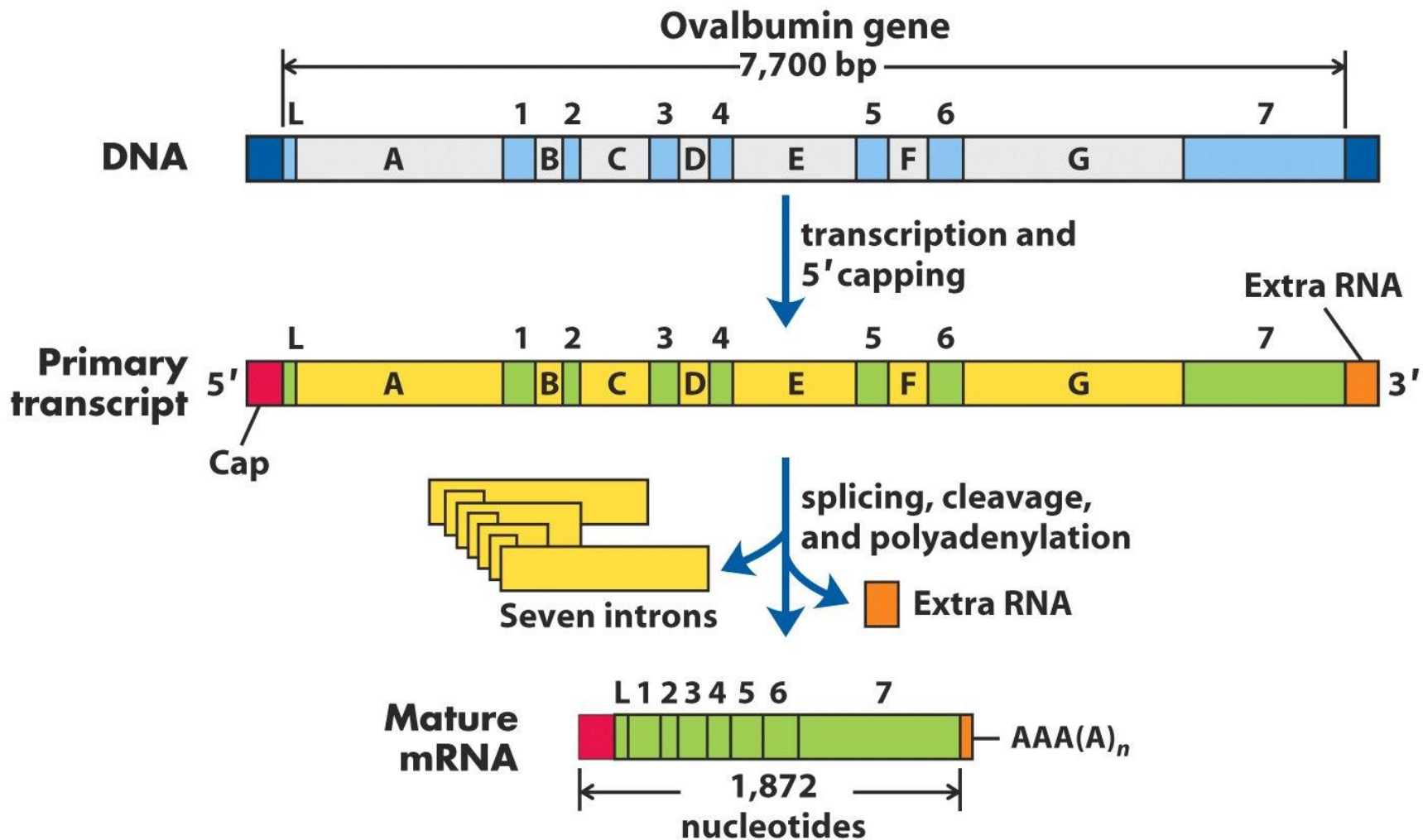
- Requires 2'OH group of an A in the intron
- Two trans-esterification reactions
- Formation of new phosphodiester bonds

Spliceosome

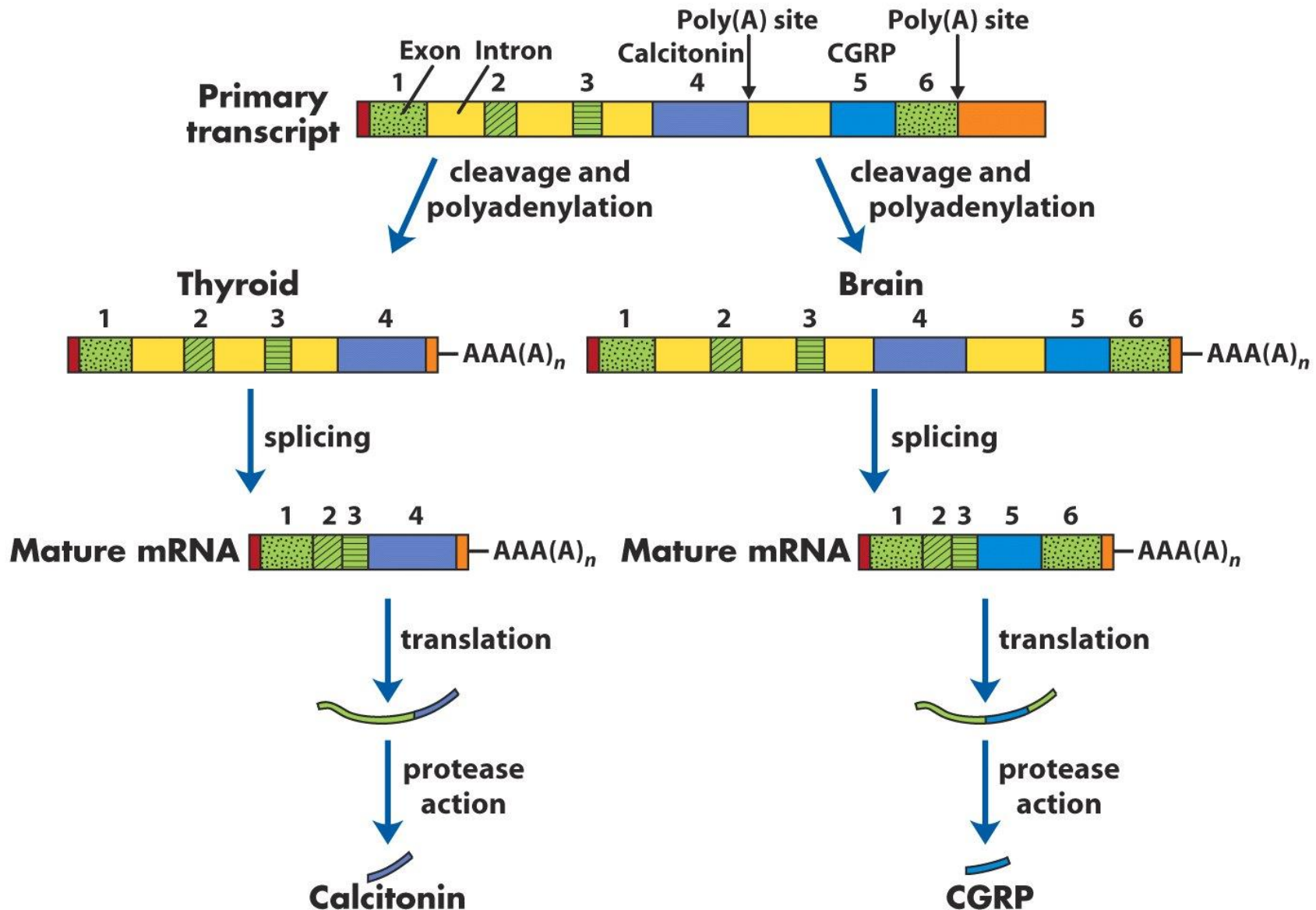
- Specialised RNA-protein complex - small nuclear ribonucleoproteins (snRNP), pronounced “snurps”
- Each snRNP contains one class of RNA (100-200 nucleotides long): small nuclear RNAs (snRNAs)
 - Five major snRNAs involved in splicing (U1, U2, U4, U5 & U6)
- Spliceosomal introns normally have the sequence GU at the 5' end and AG at the 3' end
- ATP energy is required to form the spliceosome
- Splicing method is similar to group 2



Overview



Alternative Splicing



RNA-Dependent Polymerase

- Central dogma: DNA → RNA → Protein
- However, some systems can use RNA → DNA (important in viral replication of retroviruses)
 - The enzyme reverse transcriptase can catalyse the synthesis of DNA from an RNA template.
 - First, catalyses the formation of reverse complementary DNA strand to that of the RNA template strand.
 - Second, degrades the RNA strand in the hybrid RNA-DNA complex and replaces it with a second DNA strand.
 - The double stranded DNA can be incorporated into the hosts genome, allowing for transcription and translation of the viral genes by the hosts cellular machinery.
 - Reverse transcriptase has a high error rate, leading to a high mutation rate and faster viral evolution.
- Some viral systems also synthesis RNA from RNA
 - RNA-dependent RNA polymerase (RNA replicase) is involved in genome replication from single stranded RNA genomes.

You should be able to...

- Describe and use the terms: transcription, messenger RNA, transfer RNA, ribosomal RNA, template strand, coding strand, promoter, initiation, elongation, termination, transcriptional start site;
- Appreciate the complexity of the transcription process;
- Describe the transcription process in both eukaryotes and prokaryotes in detail;
- Understand the difference in transcription complexity between eukaryotes and prokaryotes;
- Appreciate the function of base-pairing in transcription;
- Understand and describe the process of RNA post-transcriptional processing in detail;
- Appreciate the existence of introns and exons.

What's Next?

- ✓ • Topic 1 – Foundations in Biochemistry
- ✓ • Topic 2 – Amino Acids
- ✓ • Topic 3 – Proteins
- ✓ • Topic 4 – Lipids and Membranes

- ✓ • Topic 5 – Signal Transduction
- ✓ • Topic 6 – Enzymes
- ✓ • Topic 7 – Nucleotides and Nucleic Acids
- ✓ • Topic 7 – DNA Replication
- ✓ • Topic 8 – Transcription
- Topic 9 – Translation

↑
MSE
FE
↓