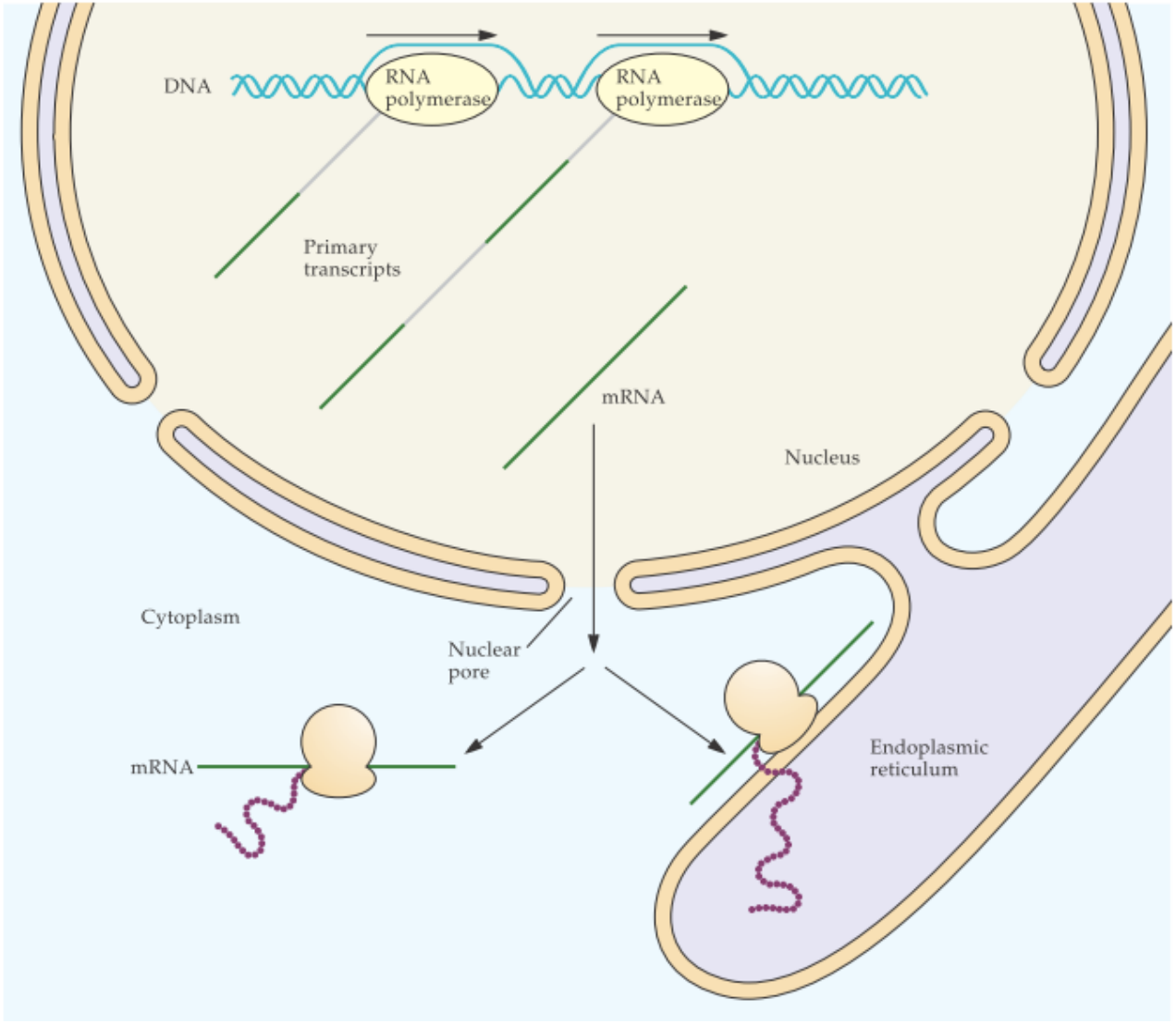
SEMESTER – III (THEORY)

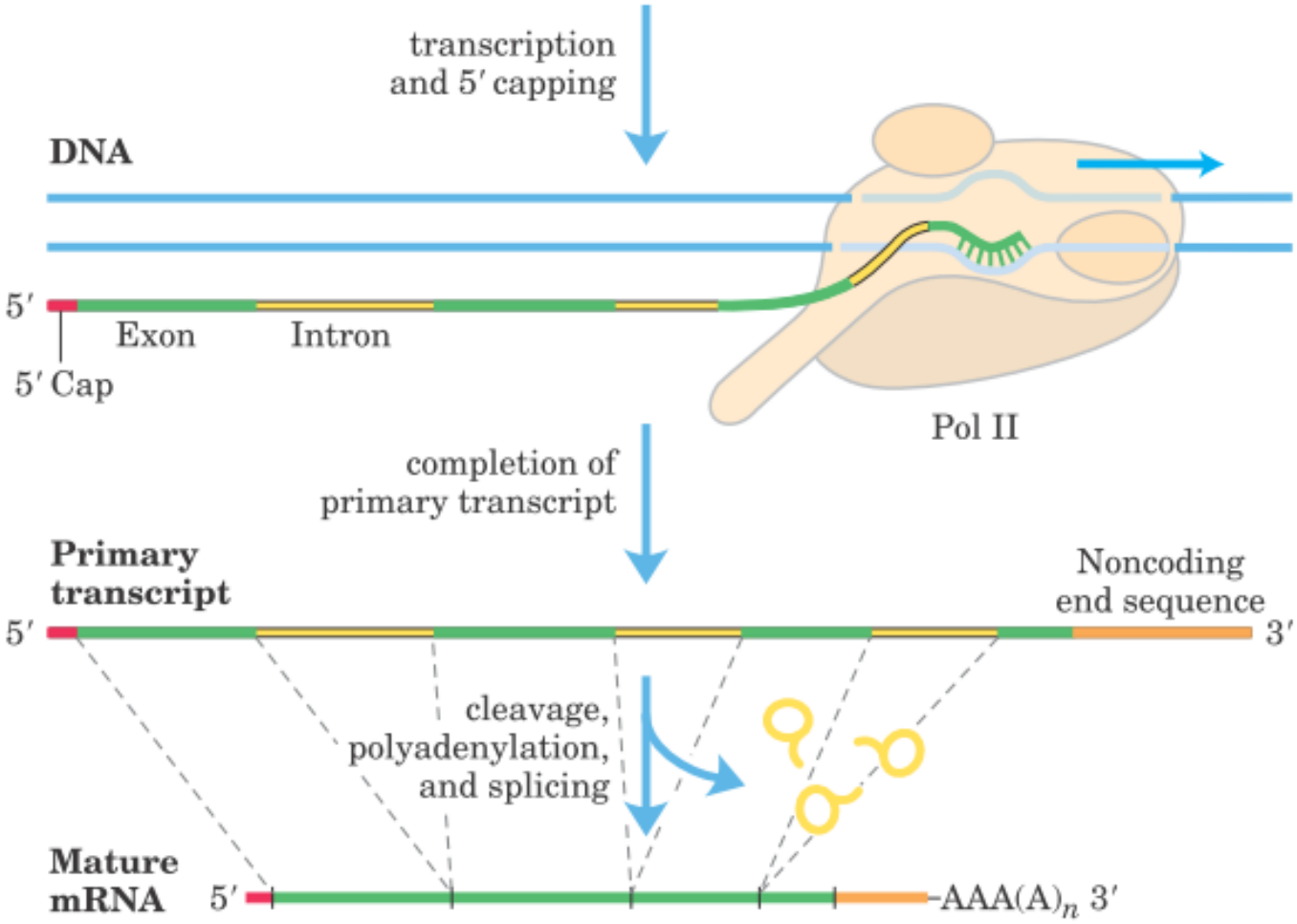
C-7: MOLECULAR BIOLOGY

Unit 4 Post-Transcriptional Processing

Transcription and translation are spatially separated in eukaryotes

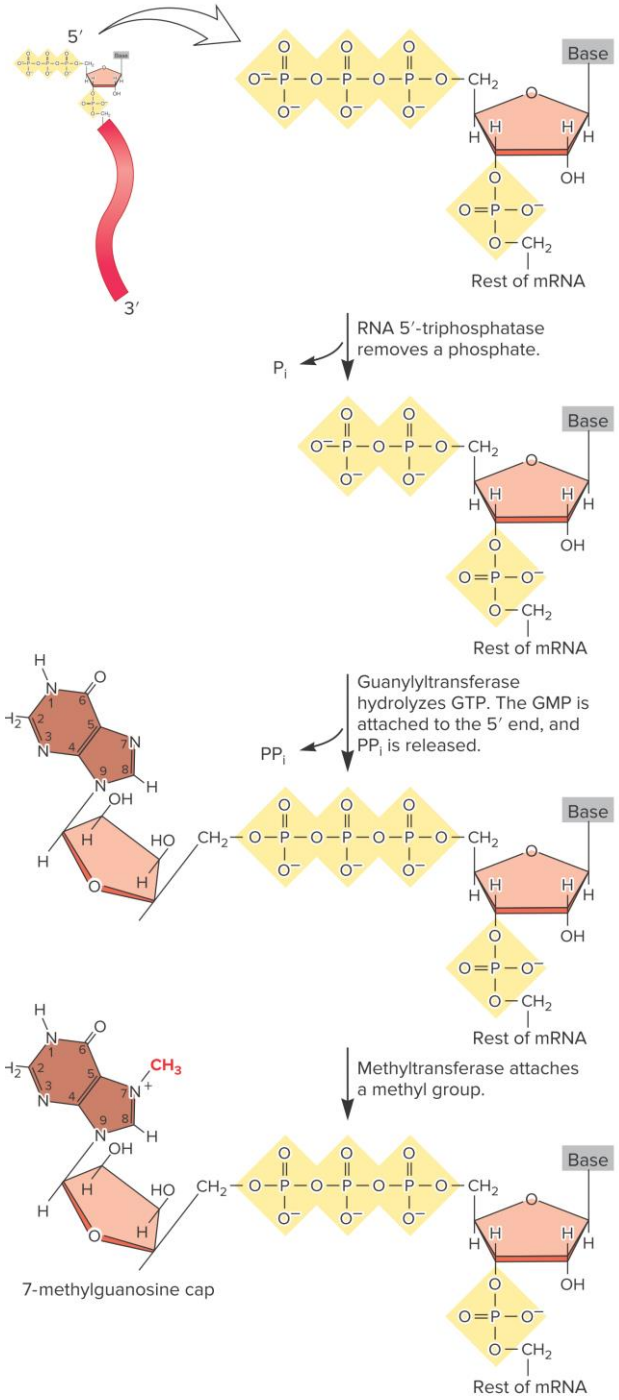


Posttranscriptional modifications: Splicing, Capping and Polyadenylation



- ✓ In the nucleus, mRNA capping, splicing and polyadenylation occur at the same time and place, in close proximity.
- ✓ Post transcriptional modifications of pre m-RNA are Co-transcriptional, rather than post-transcriptional
- ✓ The transcription elongation complex (TEC) comprising the growing nascent RNA, RNA polymerase and proteins associated with it are targeted by a plethora of factors involved in RNA processing.

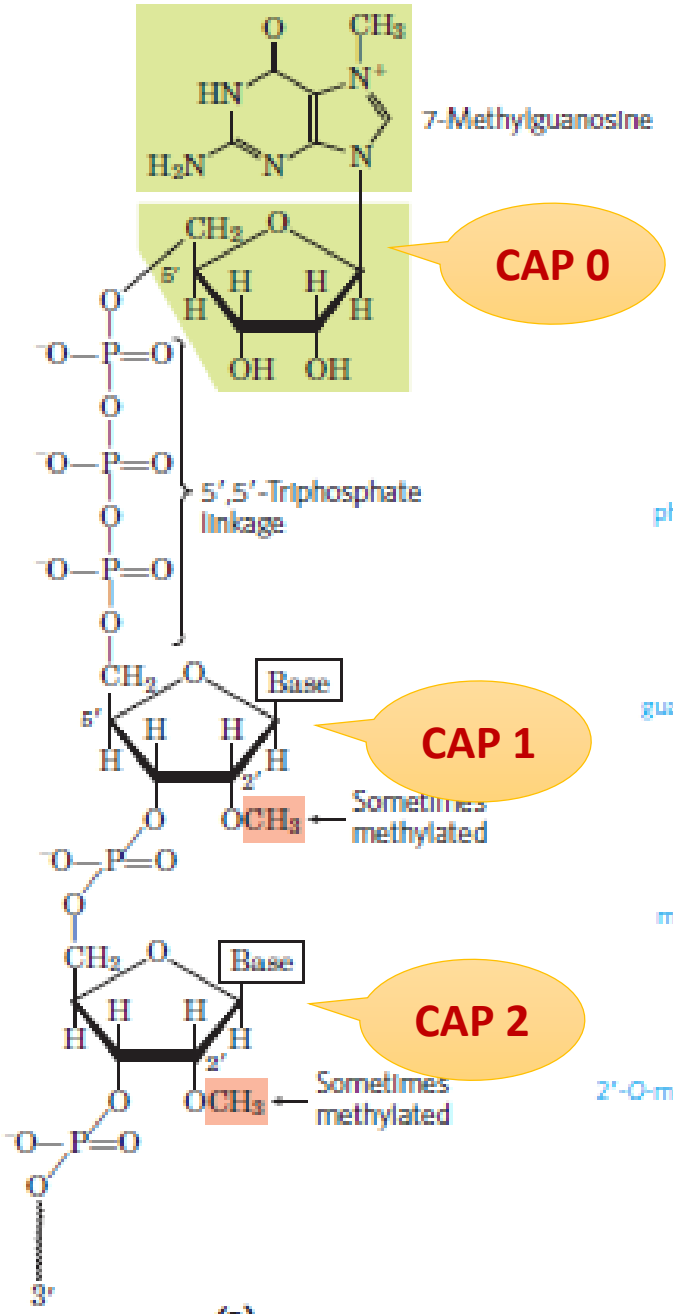
Capping

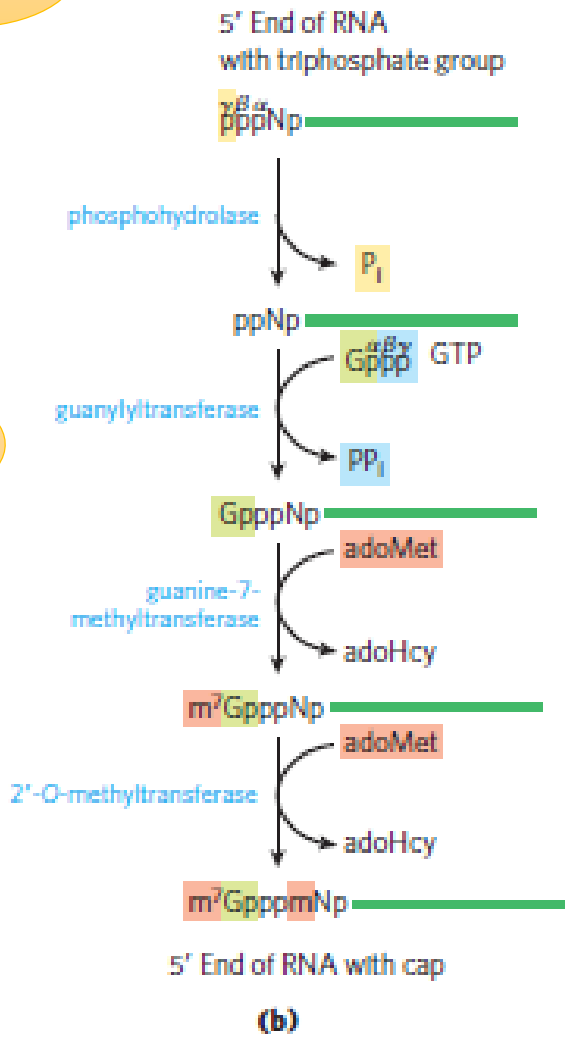


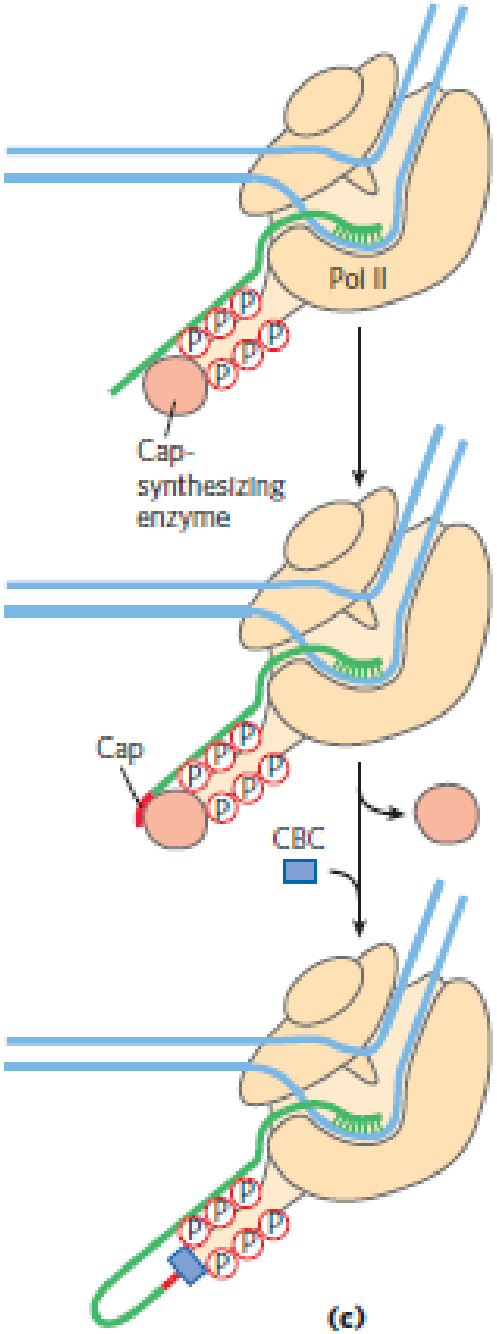
- At the 5′ end, most mature mRNAs have a 7-methylguanosine covalently attached—an event known as capping. Capping occurs while the pre-mRNA is being made by RNA polymerase II, usually when the transcript is only 20–25 nucleotides in length.

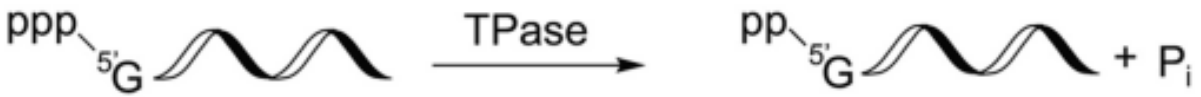
- The nucleotide at the 5′ end of the transcript has three phosphate groups. First, an enzyme called **RNA 5′-triphosphatase** removes the **γ-phosphate**, and then a second enzyme, **guanylyltransferase**, hydrolyzes **guanosine triphosphate (GTP) to attach a guanosine monophosphate (GMP) to the 5′ end**. Finally, a **methyltransferase** attaches a methyl group to a nitrogen at **position 7 in the base guanine**.

Capping

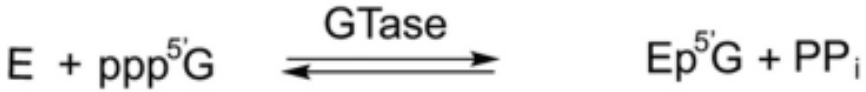

7-Methylguanosine
CAP 0
5′,5′-Triphosphate linkage
Base
CAP 1
Sometimes methylated
Base
CAP 2
Sometimes methylated
3′
(a)


5′ End of RNA with triphosphate group
pppNp
phosphohydrolase
P_i
ppNp
GTP
guanylyltransferase
PP_i
GpppNp
adoMet
guanine-7-methyltransferase
adoHcy
m^7GpppNp
adoMet
2′-O-methyltransferase
adoHcy
m^7GpppmNp
5′ End of RNA with cap
(b)

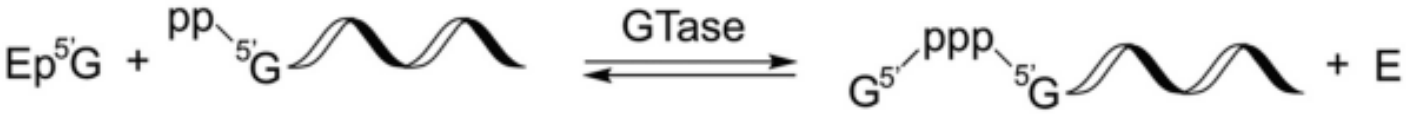

Pol II
Cap-synthesizing enzyme
Cap
CBC
(c)



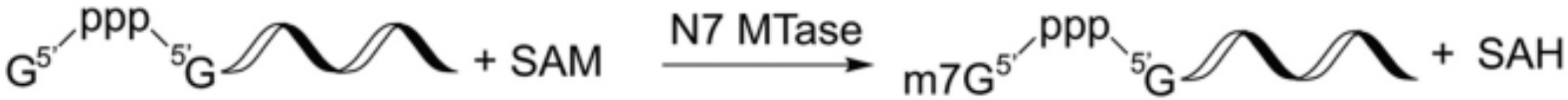
RNA triphosphatase



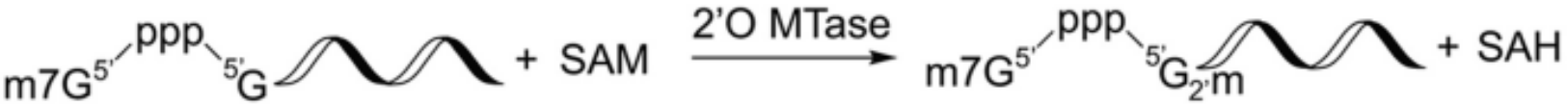
guanylyltransferase



GTase transfers a GMP group from GTP to the 5′-diphosphate via a lysine-GMP covalent intermediate



guanine-N7 methyltransferase



m7G cap-specific 2′O methyltransferase

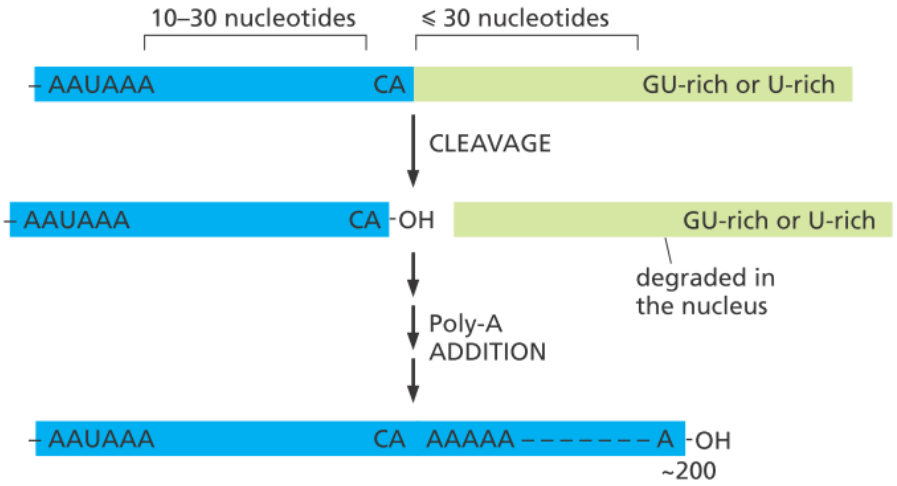
Function

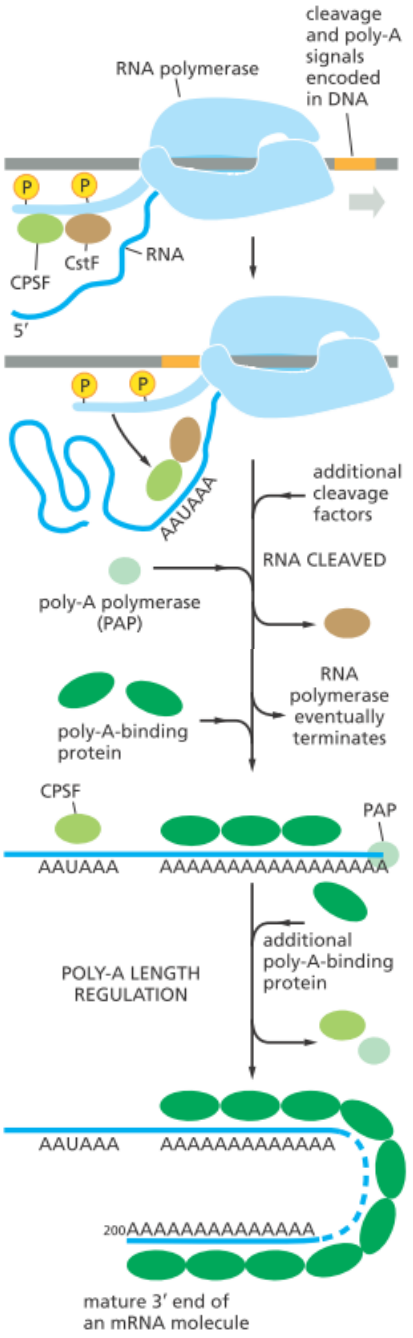
- ✓ The cap might be expected to protect the mRNA from attack by RNases that begin at the 5′-end of their substrates and that cannot cleave triphosphate linkages.

- ✓ The cap also appears to facilitate the transport of a mature RNA out of the nucleus.

- ✓ It also enhance translatability of the mRNA's. The majority of cellular mRNA translation is initiated by the cap-dependent mechanism. Upon exit into the cytoplasm, Cap-binding complex (CBC) stays bound to the mRNA cap and recruits eukaryotic translation initiation factors (eIFs) to the 5′ end of the mRNA. The eIF4E subunit of multisubunit eIF4 complex binds the 5′ cap on mRNAs. This interaction is necessary for translation initiation of eukaryotic mRNA.

Polyadenylation





Consensus nucleotide sequences that direct cleavage and polyadenylation to form the 3′ end of a eukaryotic mRNA. The enzyme in nuclei called poly(A) polymerase (PAP) that adds AMP residues one at a time to mRNA precursors.

$$\text{RNA} + n\text{ATP} \longrightarrow \text{RNA–(AMP)}_n + n\text{PP}_\text{i}$$

where n = 80 to 250. This enzyme does not require a template but does require the cleaved mRNA as a primer.

Cleavage/polyadenylation specificity factor (CPSF),
Cleavage stimulation factor (cstf),
Two cleavage factors, CFI and CFII,
Poly(a) polymerase (PAP)
Poly(a) binding protein II (PABII)

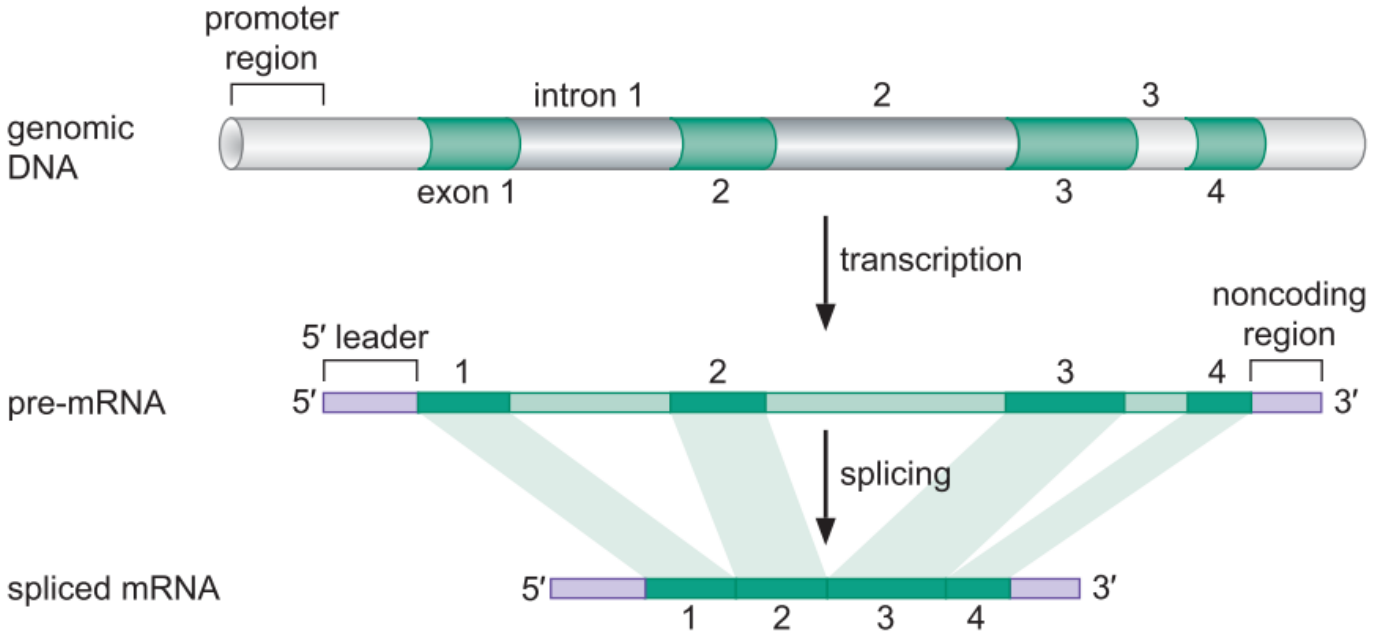
Function of Poly (A) tail

1. It helps to protect mRNAs from **degradation**

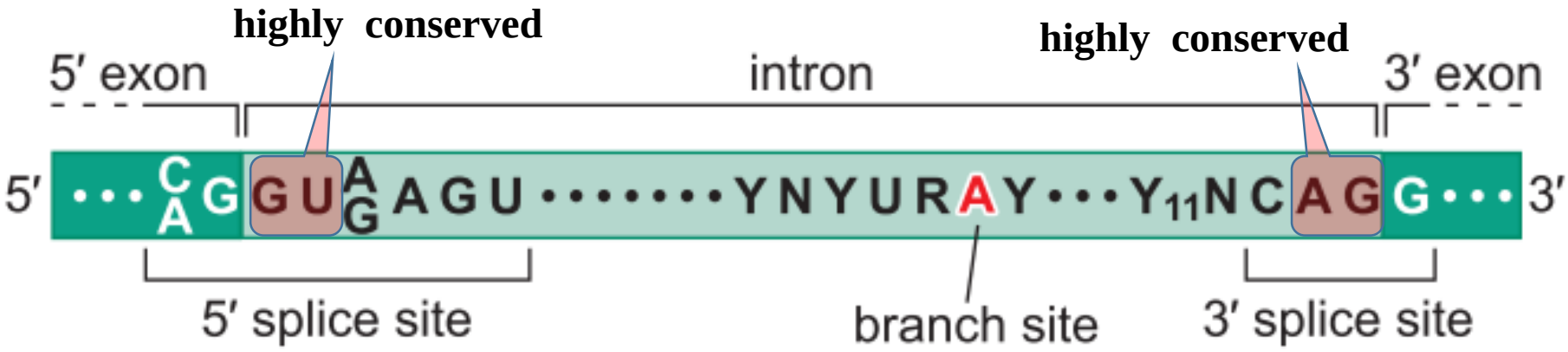
2. Poly(A) seems to **enhance translatability** by helping to recruit mRNA to polysomes, thereby promoting initiation of translation. During translation poly(A)-binding protein I (PAB I) binds to a eukaryotic mRNA . Binding to this protein seems to boost the efficiency with which an mRNA is translated.

Splicing

- Many eukaryotic genes are thus mosaics, consisting of blocks of coding sequences separated from each other by blocks of non-coding sequences. The **coding sequences** are called **exons** and the **intervening** sequences are called **introns.**
- At the molecular level, different RNA splicing mechanisms have been identified. The splicing of **group I** and **II introns** is relatively uncommon. By comparison, **pre-mRNA splicing by spliceosome** complex is a widespread phenomenon among complex eukaryotes.



Sequences at intron–exon boundaries

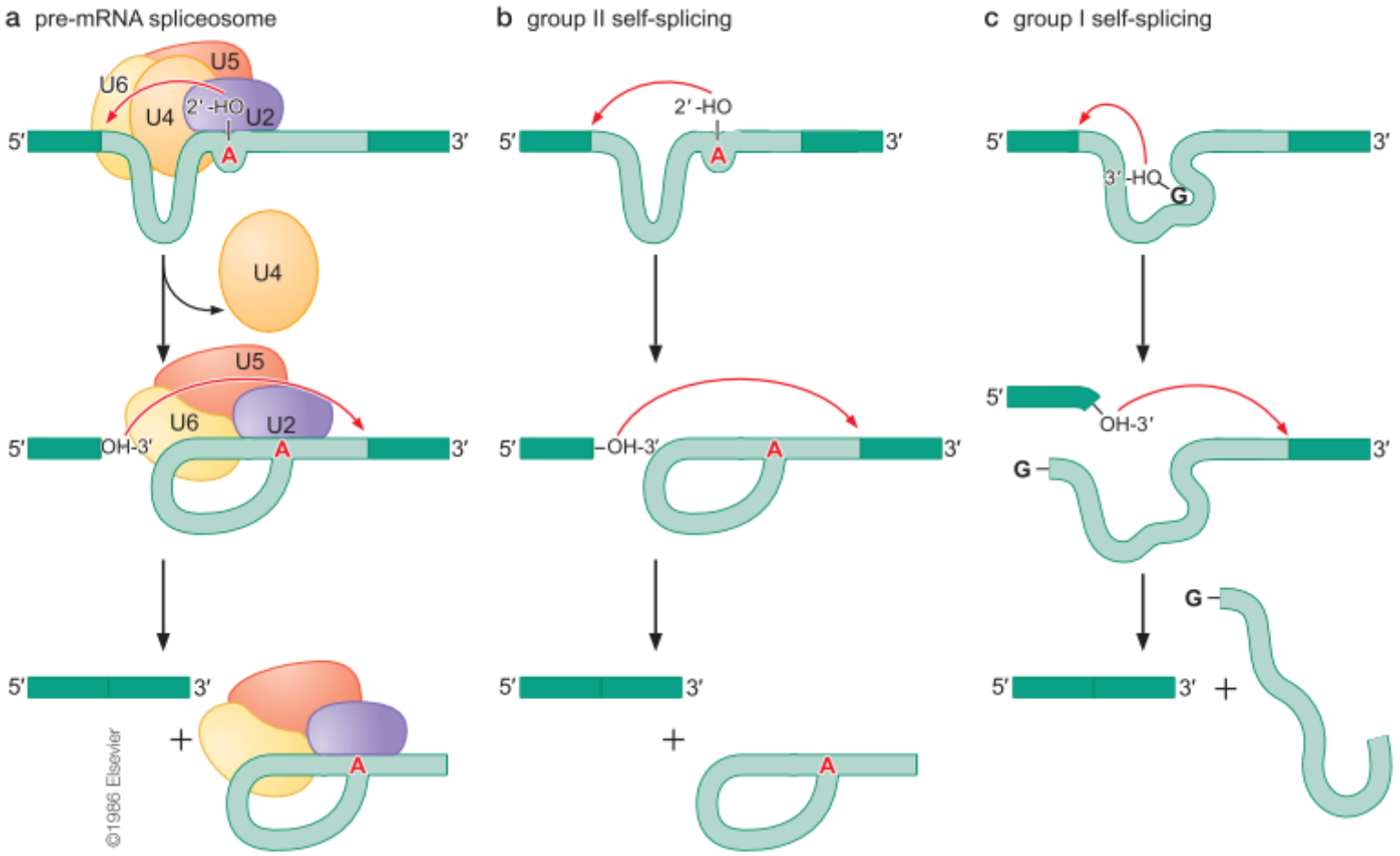


- Intron RNA is defined by particular sequences within the intron and at the intron-exon boundaries.

- The splicing of **group I and group II introns occurs via self-splicing**—splicing that does not require the aid of other catalysts. Instead, the RNA **functions as its own ribozyme**.

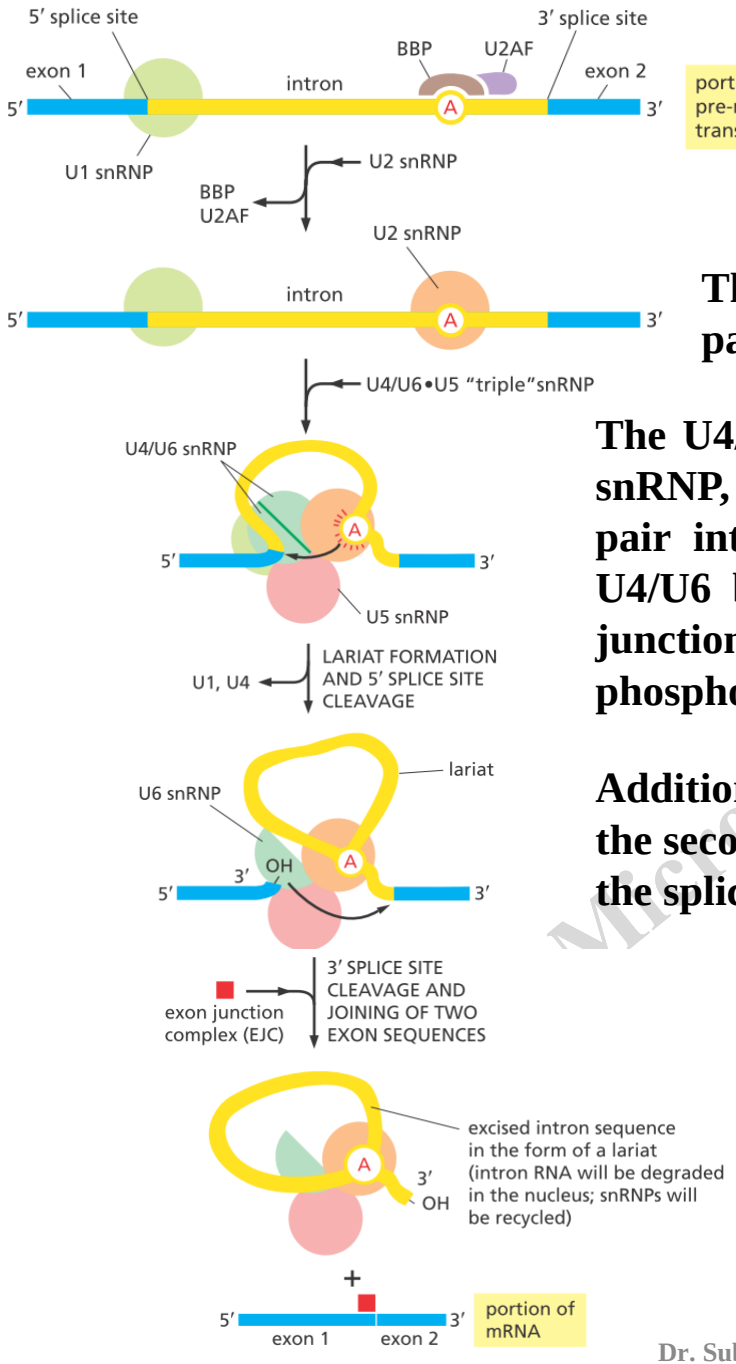
- Groups I and II differ in the ways that the introns are removed and the exons are connected.

- In group I intron, the RNA folds in a way that forms a **guanine-binding pocket**, which allows the molecule to bind a **free guanine nucleotide** and use that to initiate splicing.
- In case of group II introns, a highly reactive **adenine within the intron** initiating splicing and leading to the formation of a **lariat product**.



Occurrence of Introns

Type of Intron	Mechanism of Removal	Occurrence
Group I	Self-splicing	Found in rRNA genes within the nucleus of *Tetrahymena* and other simple eukaryotes. Found in a few protein-encoding, tRNA, and rRNA genes within mitochondrial DNA (in fungi and plants) and in chloroplast DNA. Found very rarely in tRNA genes within bacteria.
Group II	Self-splicing	Found in a few protein-encoding, tRNA, and rRNA genes within mitochondrial DNA (in fungi and plants) and in chloroplast DNA. Also found rarely in bacterial genes.
Pre-mRNA	Spliceosome	Very commonly found in protein-encoding genes within the nucleus of eukaryotic cells.



The U1 snRNP forms base pairs with the 5′ splice junction and the BBP (branch-point binding protein) and U2AF (U2 auxilliary factor) recognize the branch-point site.

The U2 snRNP displaces BBP and U2AF and forms base pairs with the branch-point site consensus sequence.

The U4/U6•U5 "triple" snRNP enters the reaction. In this triple snRNP, the U4 and U6 snRNAs are held firmly together by base-pair interactions. Subsequent rearrangements break apart the U4/U6 base pairs, allowing U6 to displace U1 at the 5′ splice junction. This creates the active site that catalyzes the first phosphoryl- transferase reaction.

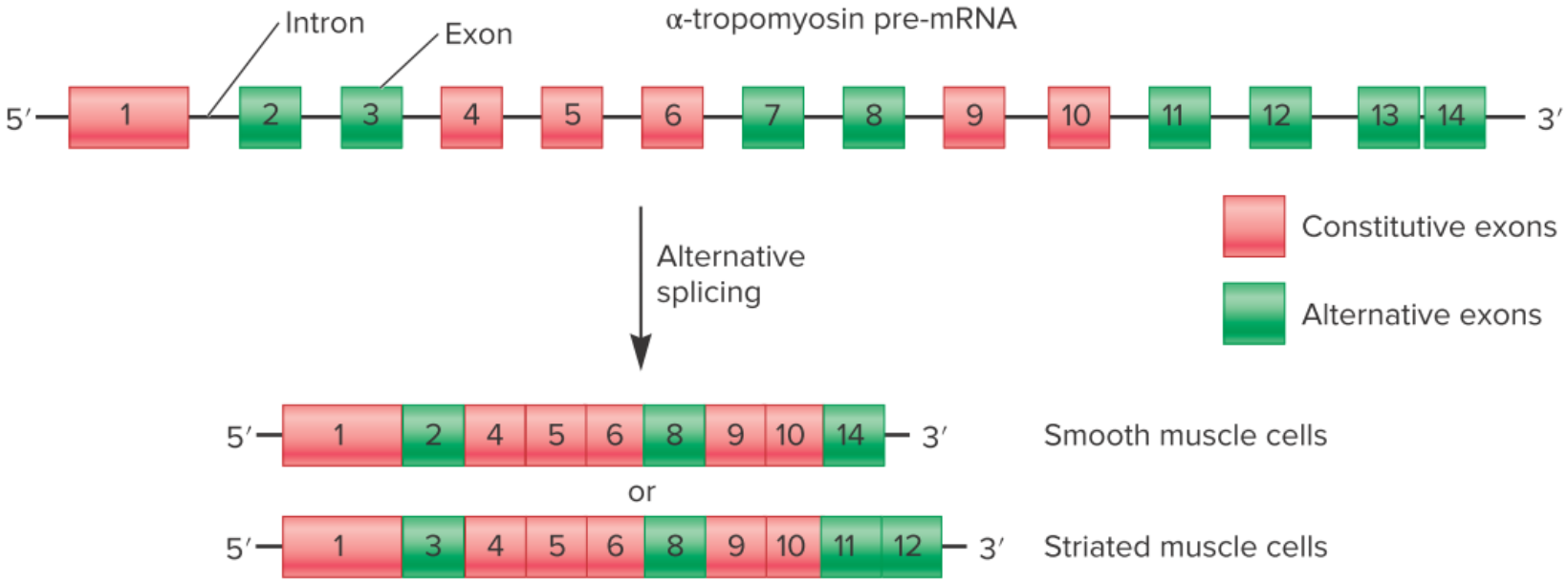
Additional RNA–RNA rearrangements create the active site for the second phosphoryl-transferase reaction, which then completes the splice.

Catalysis within the Spliceosome: The Splicing Pathway

Alternative splicing

- **Alternative splicing (AS) is a phenomenon in which a pre-mRNA can be spliced in more than one way.** In this process exons can be either excluded or included in or from a pre-mRNA resulting in multiple mRNA isoforms.

- In violation of the **'one gene, one polypeptide**' rule, **alternative splicing** allows **individual genes to produce multiple protein isoforms** — thereby playing a central part in generating complex proteomes.



The advantage of alternative splicing

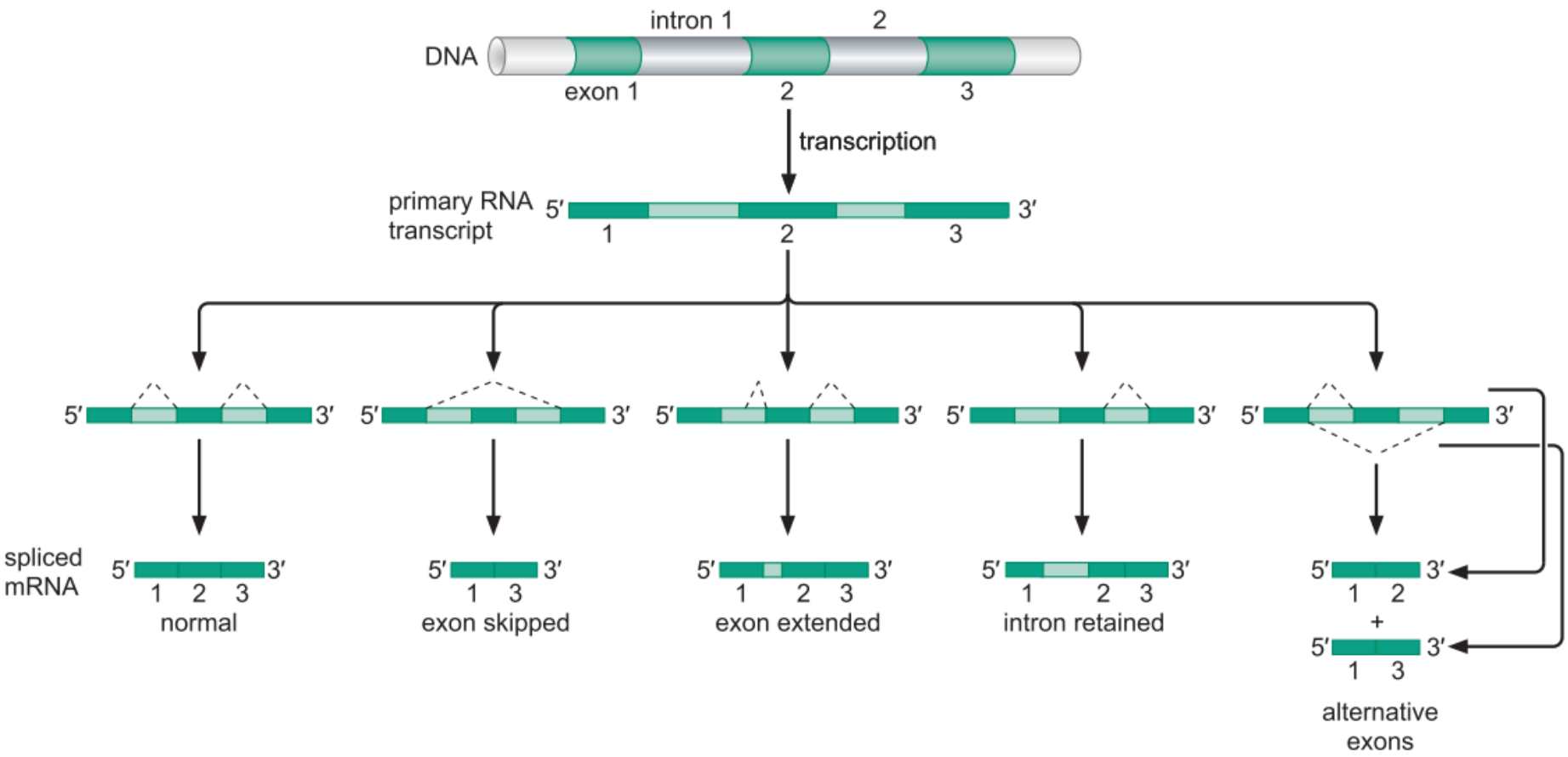
- Alternative splicing produces two or more polypeptides from the same gene that have differences in their amino acid sequences, leading to possible changes in their functions.

- Because alternative splicing allows two or more different polypeptide sequences to be derived from a single gene, some geneticists have speculated that an important advantage of this process is that it allows **an organism to carry fewer genes in its genome**.

- Different combinations of exons from the same gene may be spliced into a mature RNA, producing distinct forms of **a protein for specific tissues**, **developmental stages, or signaling pathways**.

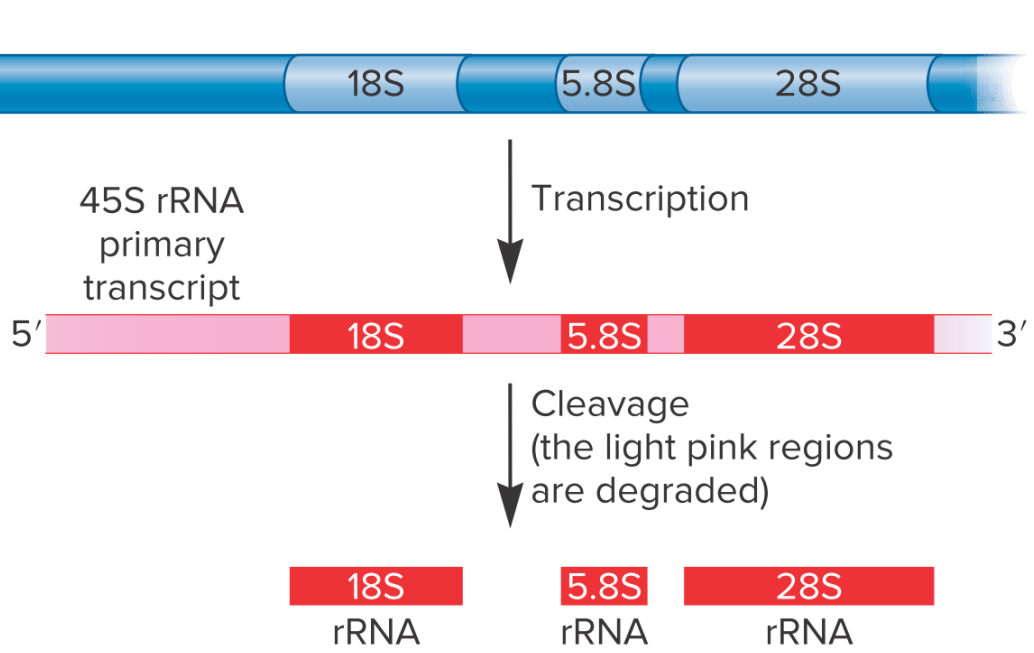
- Most alternative splicing leads to changes in the coding sequence, resulting in proteins with different functions. **Alternative splicing provides a powerful mechanism for expanding the versatility of genomic sequences** through combinatorial control.

Alternative splicing increases the coding capacity of the genome

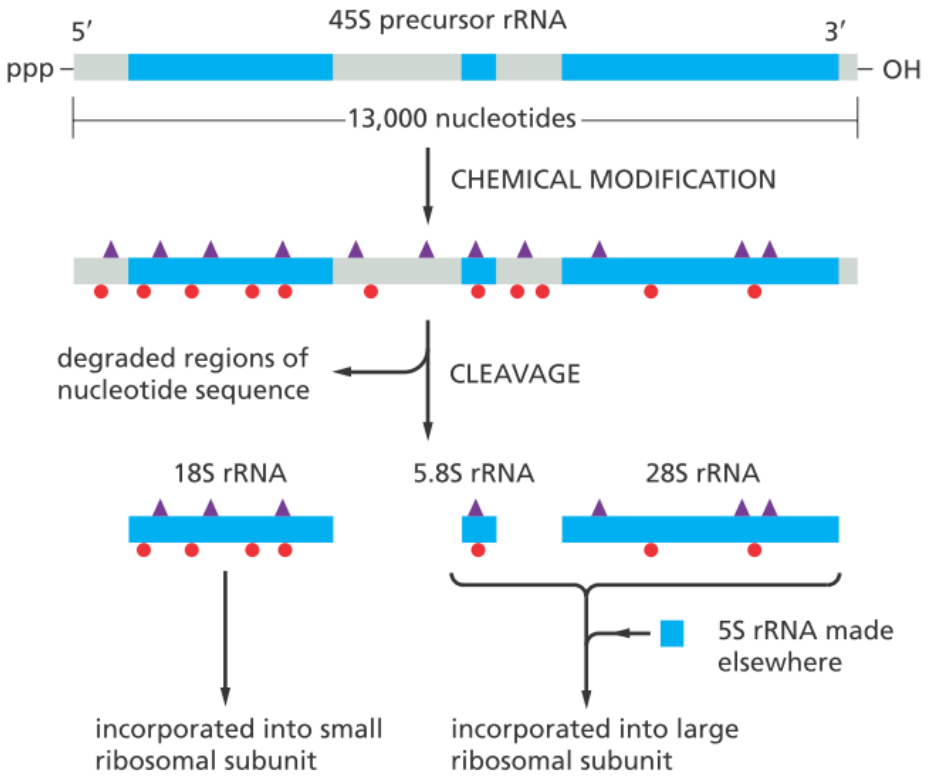


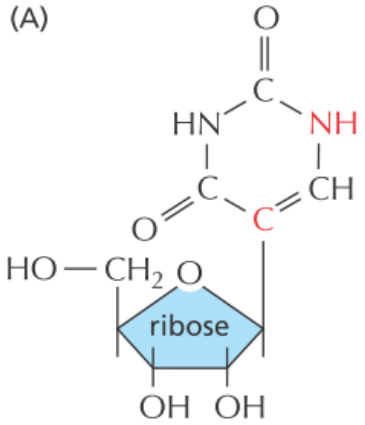
The processing of ribosomal RNA in eukaryotes

The ribosomal RNA gene is transcribed by RNA polymerase I, resulting in a long primary transcript known as 45S rRNA. Following synthesis of the 45S rRNA, cleavage occurs at several points to produce three fragments, termed 18S, 5.8S, and 28S rRNA. These functional rRNA molecules play a key role in forming the structure of the ribosome. In eukaryotes, the cleavage of 45S rRNA into smaller rRNAs and the assembly of ribosomal subunits occur in a site within the cell nucleus known as the nucleolus.

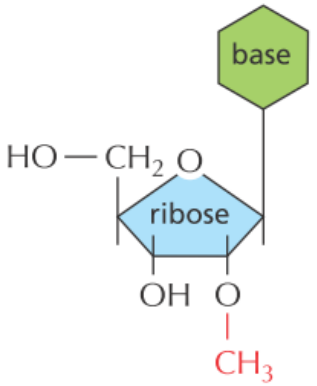


- There are four types of eukaryotic rRNAs, each present in one copy per ribosome. Three of the four rRNAs (18S, 5.8S, and 28S) are made by chemically modifying and cleaving a single large precursor rRNA; the fourth (5S RNA) is synthesized from a separate cluster of genes by RNA polymerase III and does not require chemical modification.
- Extensive chemical modifications occur in the 13,000-nucleotide-long precursor rRNA before the rRNAs are cleaved out of it and assembled into ribosomes. These include about 100 methylations of the 2′-OH positions on nucleotide sugars and 100 isomerizations of uridine nucleotides to pseudouridine.
- Each modification is made at a specific position in the precursor rRNA, specified by "guide RNAs." Other guide RNAs promote cleavage of the precursor rRNAs into the mature rRNAs, probably by causing conformational changes in the precursor rRNA that expose these sites to nucleases. All of these guide RNAs are members of a large class of RNAs called small nucleolar RNAs (or snoRNAs), so named because these RNAs perform their functions in a sub-compartment of the nucleus called the nucleolus.

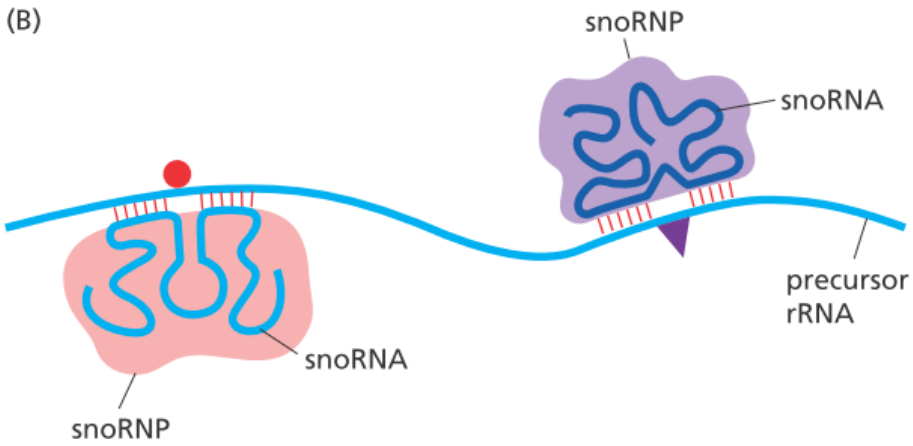




pseudouridine



2′-*O*-methylated nucleotide



Modifications of the precursor rRNA by guide RNAs.

small nuclear RNA (snRNA): Small RNA molecules that are involved in RNA splicing in the nucleus of eukaryotic cells.

small nucleolar RNA (snoRNA): Small RNA molecules that are involved in ribosomal RNA base modification in the nucleolus of eukaryotic cells.

ribonuclease P: A ribonuclease involved in processing tRNA in bacteria that consists of an RNA ribozyme plus an accessory protein.

Spliceosome: Complex of proteins and small nuclear RNA molecules that removes introns during the processing of messenger RNA.

guide RNA (gRNA): Small RNA used to locate sequences on a longer mRNA during RNA editing.

Pseudouridine: An isomer of uridine that is introduced into some RNA molecules by post-transcriptional modification

Inosine: An unusual modified nucleoside derived from guanosine

polycistronic mRNA: mRNA carrying multiple coding sequences that may be translated to give several different protein molecules; only found in prokaryotic (bacterial) cells

monocistronic mRNA: mRNA carrying the information of a single cistron, which is a coding sequence for only a single protein.

RNA interference

- **RNA interference** (**RNAi**) or **Post-Transcriptional Gene Silencing** (**PTGS**) is defined as a mechanism of specific post-transcriptional gene-silencing mediated by small RNAs, including endogenous microRNA (miRNA) and exogenous small interfering RNA (siRNA). This natural mechanism for sequence-specific **gene silencing** promises to revolutionize experimental biology and may have important practical applications in functional genomics, therapeutic intervention, agriculture and other areas.

- **MicroRNAs** (miRNAs) are **Non-coding RNAs** (ncRNAs) that are transcribed from **endogenous eukaryotic genes**—genes that are normally found in the genome. They play key roles in regulating gene expression, particularly during embryonic development in animals and plants.

- By comparison, **small-interfering RNAs** (siRNAs) are ncRNAs that usually originate from sources that are **exogenous**, which means they are not normally made by cells. The sources of siRNAs can be **viruses that infect a cell**, or researchers can make siRNAs to study gene function experimentally.

Mechanism of RNA interference

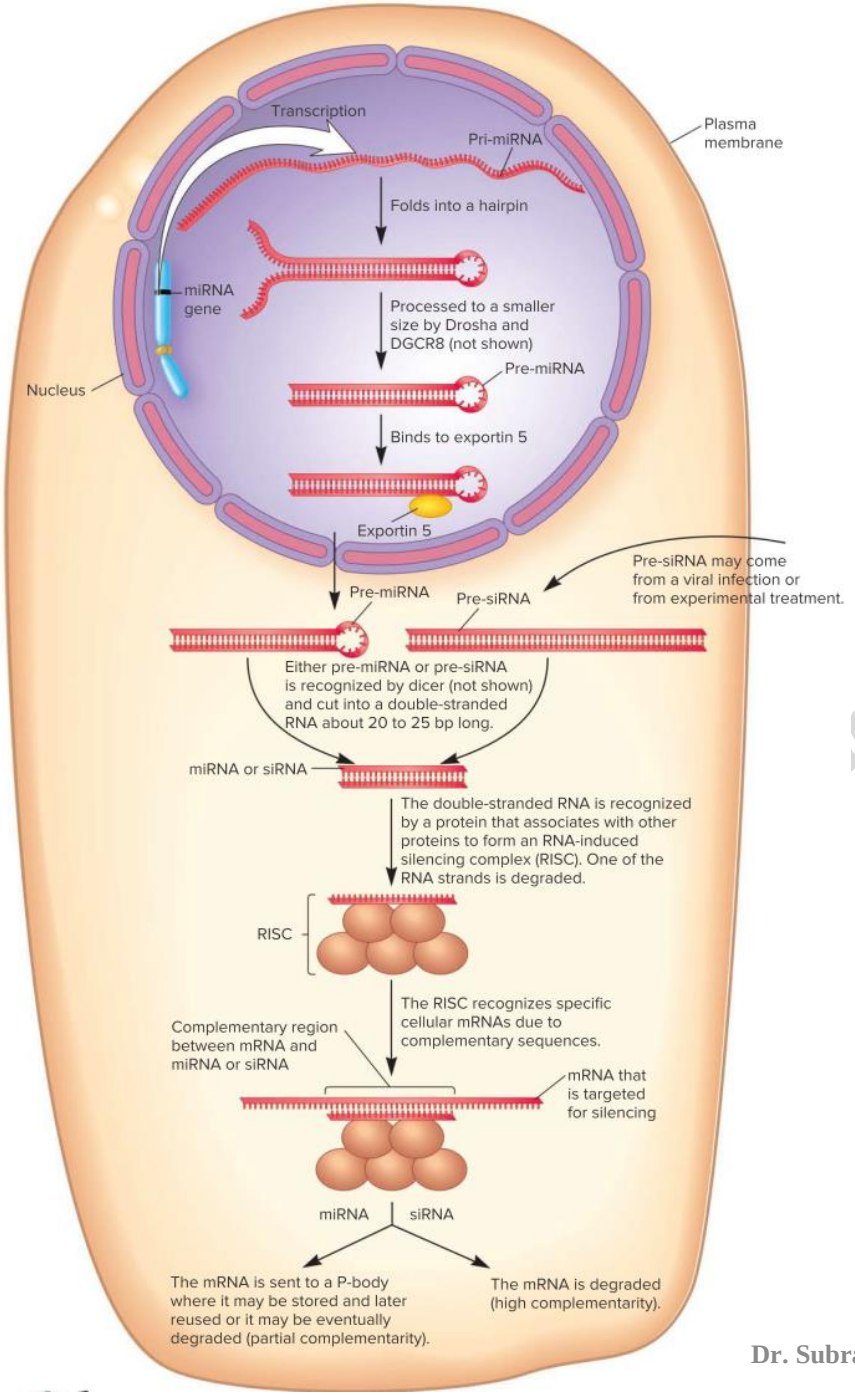
- An miRNA is first synthesized as a pri-miRNA (for primary-miRNA) in the nucleus. Due to complementary base pairing, the pri-miRNA forms a hairpin structure (a stem-loop) with long single-stranded 5′ and 3′ ends.
- The pri-miRNA is recognized in the nucleus by two proteins, **Drosha** and **DGCR8**, and is cleaved at both ends. The result is a 70-nucleotide RNA molecule that is called a pre-miRNA (for precursor-miRNA).
- The pre-miRNA is then exported from the nucleus with the aid of a protein called **exportin 5**.
- [*siRNAs do not go through the processing events that occur in the nucleus. Instead, precursor-siRNAs (pre-siRNAs) are usually derived from viral RNAs or may be made by researchers and taken up by cells*.]
- In the cytosol, both pre-miRNAs and pre-siRNAs are cut by an endonuclease called **dicer**, releasing a double-stranded RNA molecule that is 20–25 bp long.
- This double-stranded RNA associates with proteins to form a complex called the **RNA-induced silencing complex** (RISC). One of the RNA strands is degraded. The remaining single-stranded miRNA or siRNA is complementary to specific mRNAs that will be silenced.

After RISC binds to an mRNA, any of the following three outcomes may result:

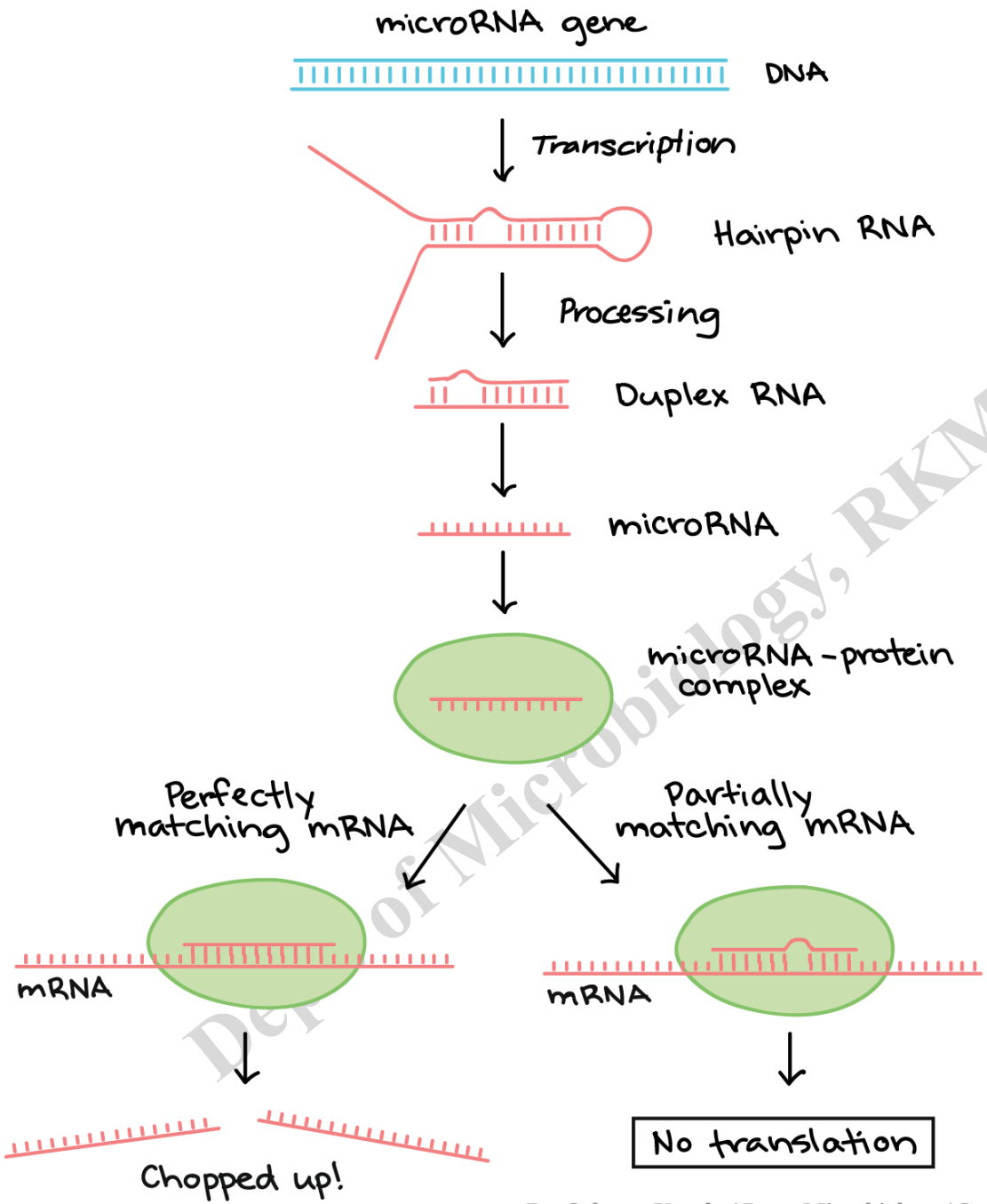
1. RISC may inhibit translation **without degrading the mRNA**. This is more common for miRNAs, which often are only partially complementary to their target mRNAs.

2. The RISC-mRNA complex may remain in **a cellular structure called a processing body** (P-body), where it can be stored and later reused. In this case, the inhibition of translation by an miRNA is only temporary.

3. RISC may direct the **degradation of the mRNA**. One of the proteins in RISC, which is called **Argonaute**, can cleave the mRNA. This outcome usually occurs with siRNAs, which are typically a perfect match to their target mRNAs.



Mechanism of RNA interference



Mechanism of RNA interference (simplified form)