POST TRANSCRIPTIONAL MODIFICATIONS OF RNA

CAPPING OF THE 5' END OF THE RNA

Once transcribed eukaryotic RNA has to be processed in various ways before being exported from the nucleus where it can be translated. These processing events include capping of the 5' end of the RNA, splicing, and polyadenylation of the 3' end of the RNA. The RNA is capped as soon as it emerges from the RNA exit channel of polymerase.

The first RNA processing event is capping. This involves the addition of a modified Guanine, and it is joined to the RNA transcript by an unusual 5'-5' linkage involving three phosphates.

Steps for the formation of the 5' cap:

- In the first step, the ϒ- phosphate at the 5' end of the RNA is removed by an enzyme called RNA triphosphatase.
- In the next step, the enzyme guanylyl transferase adds a GMP moiety to the resulting terminal ϒ-phosphate.
- This is a two – step process: first, an enzyme -GMP is generated from GTP with release of the β and ϒ-phosphates of that GTP, and then the GMP of the enzyme is transferred to the β-phosphate of the 5' end of the RNA.
- Once this linkage is made the newly added guanine and the purine at the original 5' end of the mRNA is further modified by the addition of methyl groups by methyltransferase.
- The resulting 5' cap structure subsequently recruits the ribosome to the mRNA for translation to begin.

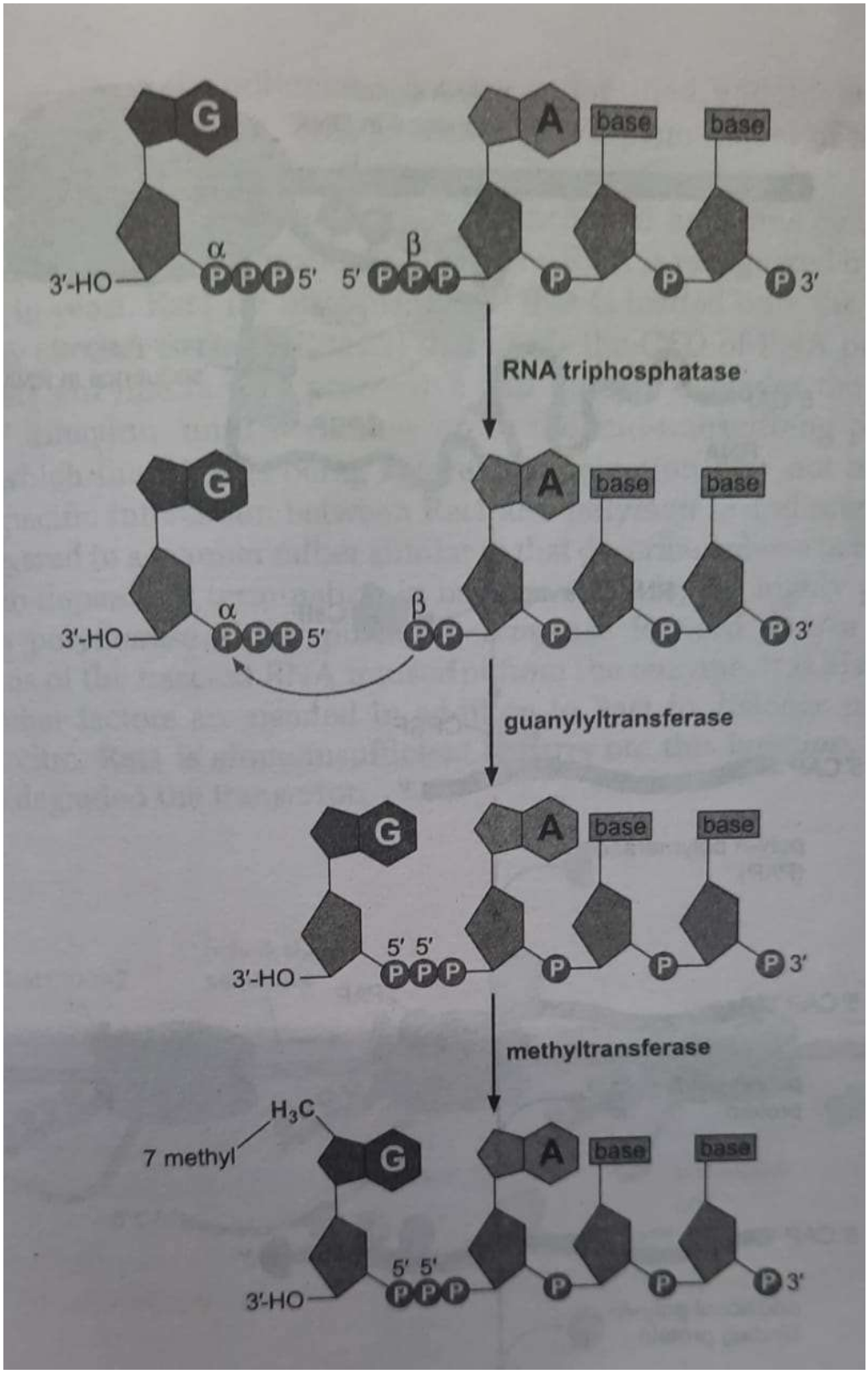

FIG: Image representing the 5’ capping of the mRNA

SPLICING

In the chemistry of splicing, an intron is removed through two successive transesterification reactions in which phosphodiester linkages within the pre mRNA are broken and new ones are formed. The transesterification reactions are mediated by a large molecular machine called the **spliceosome**. The spliceosome is a large complex made up of these small nuclear Ribonuclear proteins- $snRNP_s$ (the five RNA_s- U_1, U_2, U_4, U_5, U_6, complexed with several proteins).

The $snRNP_s$ have three roles in splicing:

1. They recognize the 5' splice site and the branch site
2. They bring those sites together as required
3. They catalyze the RNA cleavage and joining reactions.

THE SPLICING PATHWAY

- Initially, the 5' splice site is recognized by the U_1 snRNP.
- U2AF (U2 auxiliary factor) is made up of two sub-units, the larger of which binds to the Py tract and the smaller binds to the 3' splice site.
- The former subunit interacts with the BBP (branchpoint-binding protein) and helps that protein bind to the branch site. This arrangement of proteins is and RNA is called the **early (E) complex.**
- U2 snRNP then binds to the branch site, aided by U2AF and displacing BBP. This arrangement is called the **A complex**.
- The next step is the rearrangement of the A complex to bring to bring together all three splice sites. This is achieved as follows: the U4 and U6

snRNPs, along with the U5 snRNP, join the complex. Together these three snRNPs are called the tri-snRNP particle.

- The A complex is converted to **B complex**.
- In the next step, U1 leaves the complex and U6 replaces it at the 5' splice site.
- The next rearrangement triggers catalysis and occurs as follows" U4 is released from the complex, allowing U6 to interact with U2 through the RNA: RNA base pairing. This arrangement called the C complex produces the active site.
- Formation of the active site juxtaposes the 5' splice site of the pre-mRNA and the branch site, facilitating the first transesterification reaction.
- The second reaction between the 5' and 3' splice sites, is aided by the U5 snRNP, which helps to bring the two exons together.
- The final step involves release of the final mRNA product and the snRNPs. The snRNPs are initially still bound to the lariat, but they get recycled after rapid degradation of that piece of RNA.

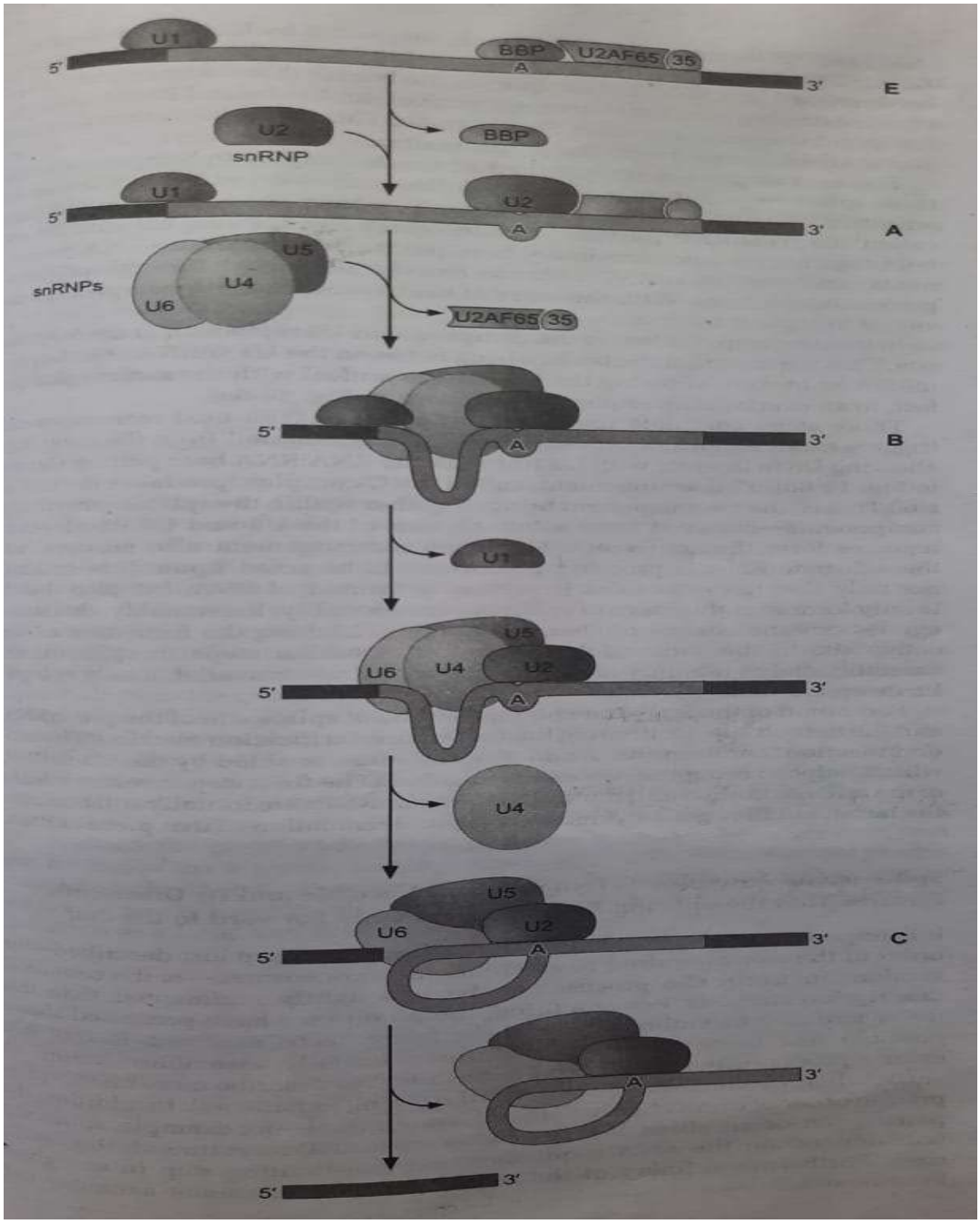

FIG: Image representing function of splicing.

POLYADENYLATION OF THE 3' END OF THE RNA

The final RNA processing event, polyadenylation of the 3' end of the mRNA, is intimately linked with the termination of transcription. Just as with capping and splicing, the polymerase CTD (Carboxy Terminal Domain) tail is involved in recruiting some of the enzymes necessary for polyadenylation. Once polymerase has reached the end of a gene, it encounters specific sequences that , after being transcribed in to RNA, trigger the transfer of the polyadenylation enzymes to that RNA, leading to four events:

- Cleavage of the message
- Addition of many adenine residues to its 3' end
- Degradation of the RNA remaining associated with RNA polymerase by a 5'- 3' ribonuclease
- Termination of transcription.

  The series of events are as follows:

  1. Two protein complexes are carried by the CTD of polymerase as it approaches the end of the genes: CPSF (cleavage and polyadenylation specificity factor) and CSTF (cleavage stimulating factor).
  2. The sequences that once transcribed in to RNA, trigger transfer of these factors to the RNA are called **poly-A signals**.
  3. Once CPSF and CSTF are bound to the RNA, other proteins are recruited as well, leading initially to RNA cleavage and then to polyadenylation.

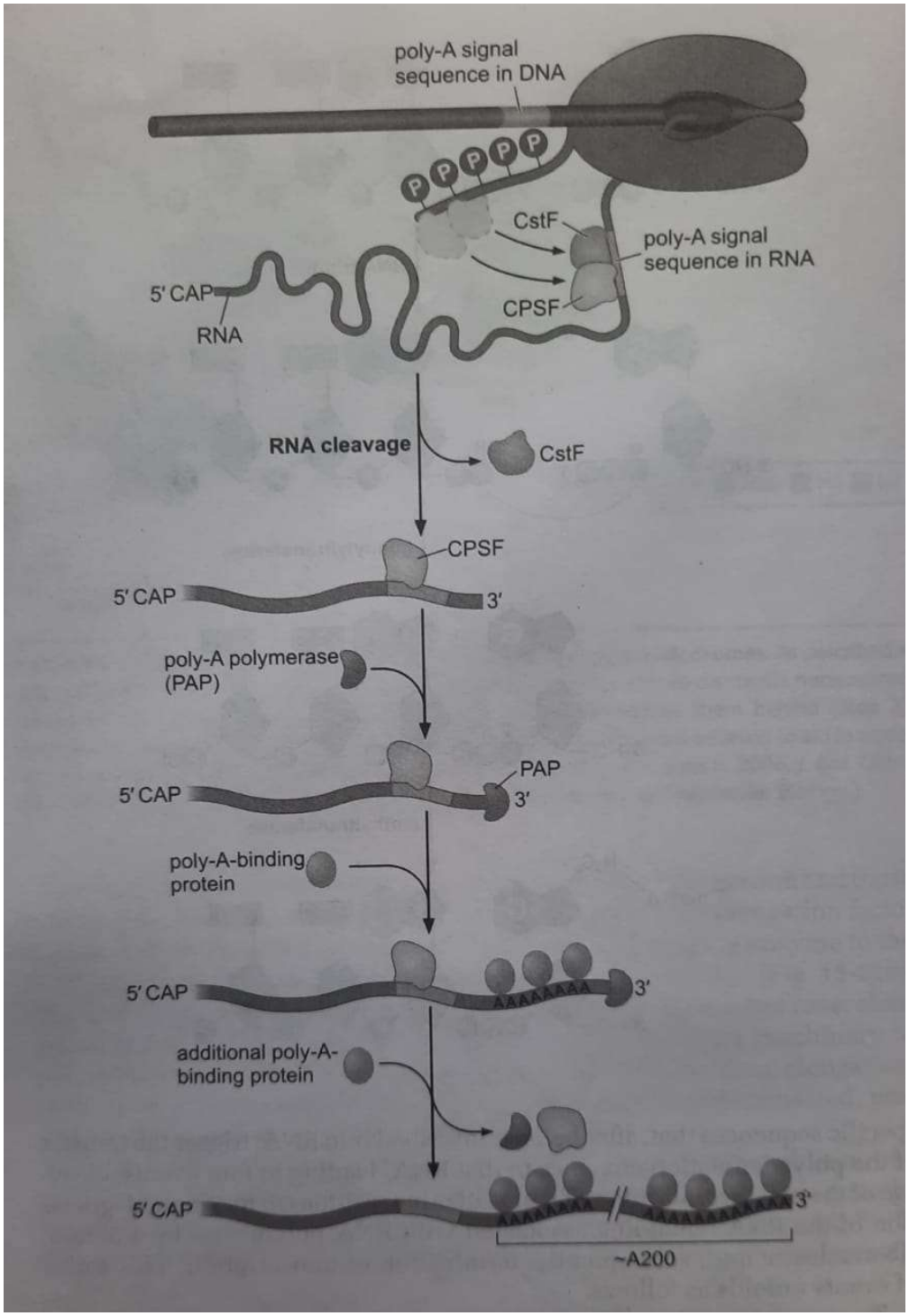

FIG: Image representing the 3' polyadenylation of the mRNA

4. Polyadenylation is mediated by an enzyme called poly-A polymerase, which adds approximately 200 adenines to the RNA's

3' end produced by the cleavage. This enzyme uses ATP as a precursor and adds the nucleotides using the same chemistry as RNA polymerase. But it does so without a template.

GROUP I and GROUP II self-splicing introns

By self- splicing, we mean that the intron itself folds in to a specific confirmation within the precursor RNA and catalyzes the chemistry of its own release.

GROUP II INTRONS

The chemistry of splicing and the RNA intermediates produced are the same as those for nuclear pre-mRNAs.

The intron uses an A residue within the branch site to attack the phosphodiester bond at the boundary between the 5' end and the end of the 5' exon- that is at the 5' splice site. This reaction produces the branch **lariat**, and is followed by a second reaction in which the newly freed 3'-OH of the exon attacks the 3' splice site, releasing the intron as a lariat and fusing the 3' and 5' exons.

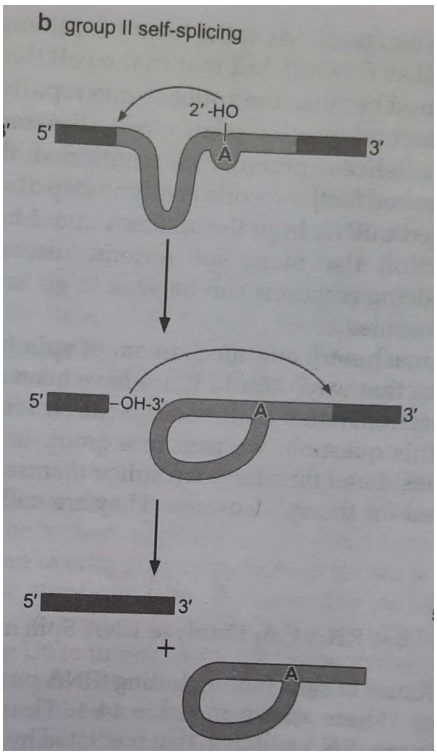

Fig: Group II introns forming a lariat

GROUP I INTRONS

Group I introns, which are smaller than group II introns, share a conserved secondary structure. The structure of group I introns includes a binding pocket that will accommodate any Guanine nucleotide or nucleoside as long as it is a ribose form.

In addition to the nucleotide binding pocket, group I introns contain an “internal guide sequence” that base pairs with the 5’ splice site sequence and thereby determines the precise site at which nucleophilic attack by the G nucleotide takes place.

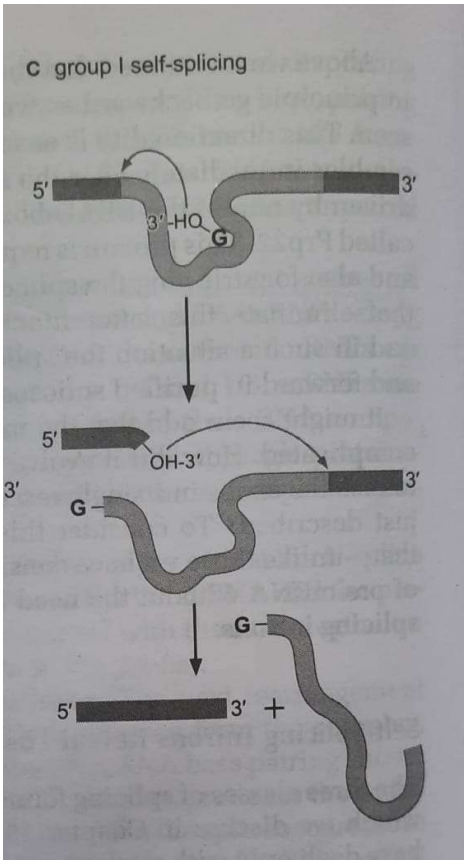

Fig: Splicing by Group I introns

RIBOZYMES

RNA enzymes are known as ribozymes, and they exhibit many of the features of the classical enzyme such as an active site, substrate binding site, and a co-factor binding site, such as a metal ion. Ribozymes perform trans-esterification reactions involved in the removal of intervening sequences known as introns from precursors to certain mRNAs, tRNAs, and ribosomal RNAs in a process known as RNA splicing.

One of the first ribozymes to be discovered was the RNase P, an endoribonuclease that is involved in generating tRNA molecules from larger, precursor RNAs.

Another example is the Hammerhead ribozyme found in *Tetrahymena* (a Protista).

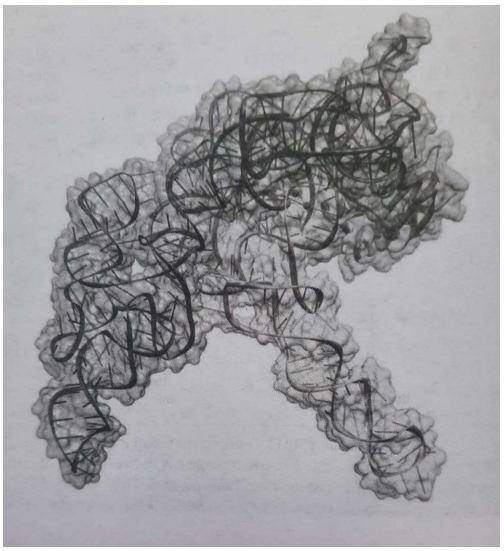

FIG: Image of Hammerhead ribozyme.

Reference:

Watson, J, D. Molecular Biology of The Gene. Pearson Publication. Fourth edition 2019.