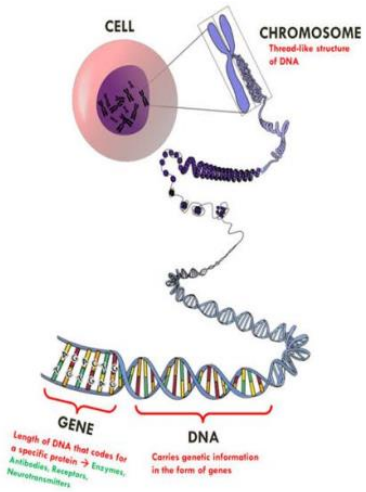
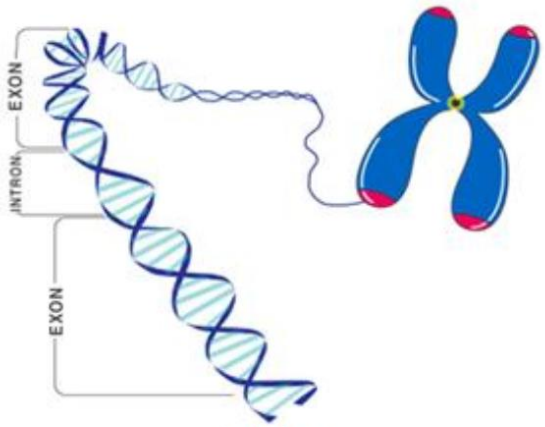
Processing and modification of RNA

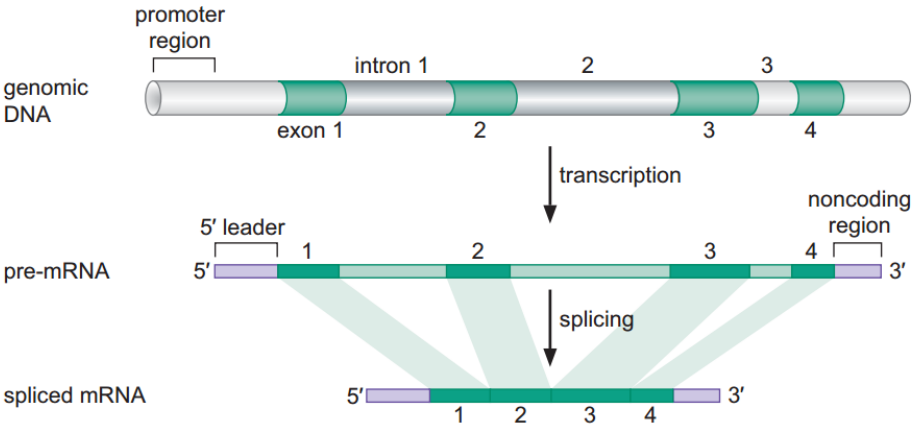
Split Genes: A split gene (or interrupted gene) is a gene that contains sections of DNA called exons, which are expressed as RNA and protein, interrupted by sections of DNA called introns, which are not expressed.





Split Gene

Most of the portion of a gene in higher eukaryotes consists of noncoding DNA that interrupts the relatively short segments of coding DNA. The coding sequences are called exons. The noncoding sequences are called introns.



Introns:

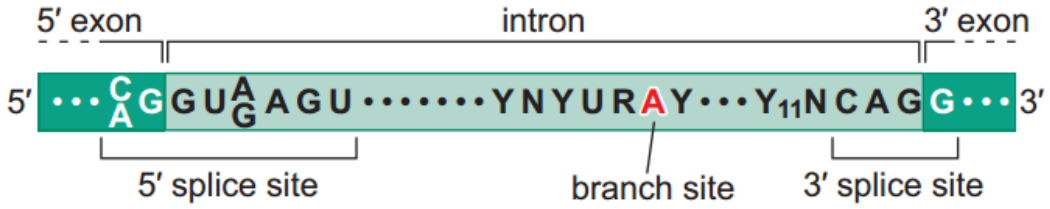
- An intron is a portion of a gene that does not code for amino acids.
- Represented in the primary transcript of the gene, but absent in the final processed form.
- Introns are noncoding regions of an RNA transcript which are eliminated by splicing before translation.
- Introns are very large chunks of RNA within a mRNA that interfere with the code of the exons.
- These introns get removed from the RNA molecule to leave a string of exons attached to each other so that the appropriate amino acids can be encoded for.
- They are not expressed in the protein.
- Introns are rare in genes of prokaryotes.

Exons:

- The coding sequences are called Exon.
- An exon is the portion of a gene that codes for amino acids.
- They are expressed in the protein.

RNA Splicing:

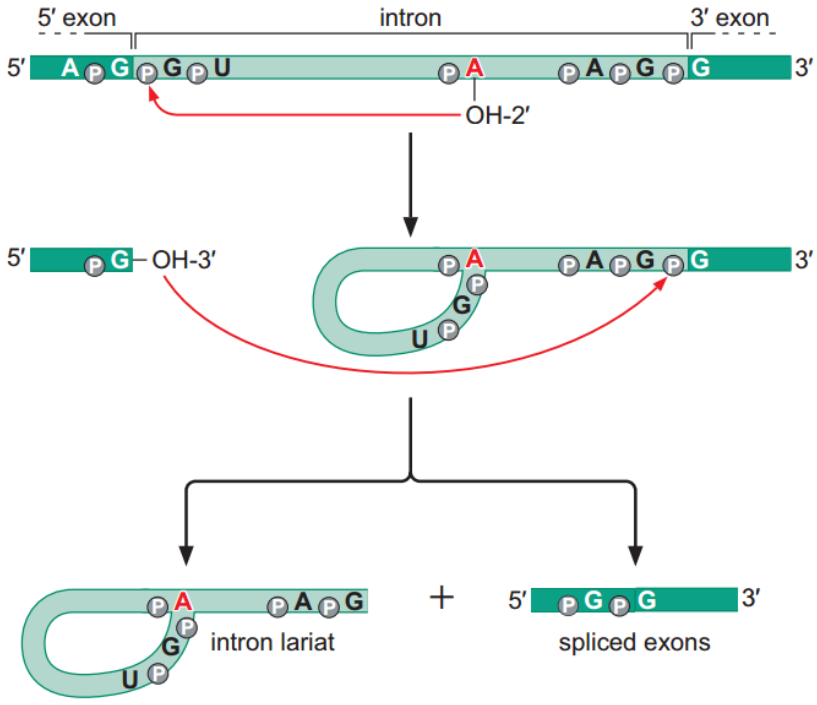
- Following transcription, new, immature strands of mRNA, called pre-mRNA, may contain both introns and exons. These pre-mRNA molecules go through a modification process in the nucleus called splicing during which the noncoding introns are cut out and only the coding exons remain.
- During splicing, introns are removed and exons are joined together.
- Splicing produces a mature mRNA molecule that is then translated into a protein.
- Sequences within the RNA Determine where Splicing Occurs.



Y- pyrimidine, R- purine, N- any nucleotide

The exon-intron boundary

- The boundary at the 5' end of the intron is marked by a sequence called the **5' splice site** and at the 3' end of the intron is marked by the **3' splice site**.



An intron is removed through two successive transesterification reactions in which phosphodiester linkages within the pre-mRNA are broken and new ones are formed.

First reaction:

- The 2'-OH group of the conserved **A** at the branch site acts as a *nucleophile* to attack the phosphoryl group of the conserved **G** in the 5' splice site.
- The phosphodiester bond between the sugar and the phosphate at the 5' junction between the intron and the exon is cleaved.
- The freed 5' end of the intron is joined to the **A** within the branch site.

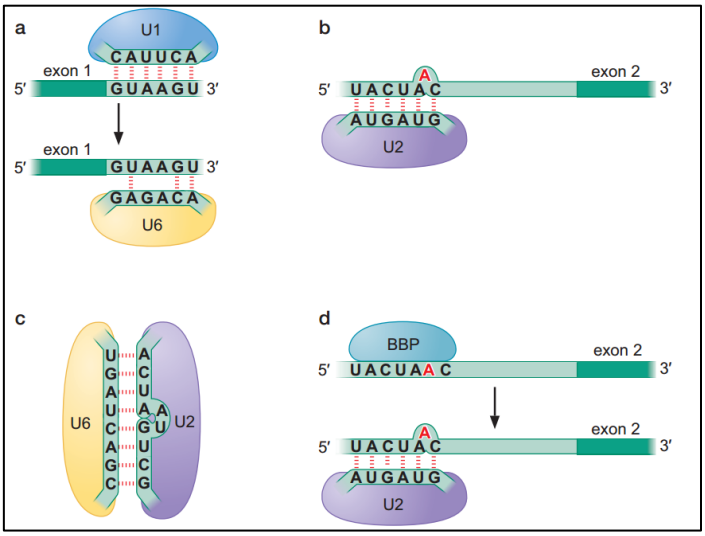
Second reaction:

- The 5' exon (3'-OH group) reverses its role and becomes a nucleophile that attacks the phosphoryl group at the 3' splice site. Then it joins the 5' and 3' exons.
- In the two reaction steps, there is no net gain in the number of chemical bonds-
- two phosphodiester bonds are broken, and two new ones made.

THE SPLICEOSOME MACHINERY

- RNA Splicing Is Performed by a Large Complex Called the Spliceosome.
- This complex comprises about 150 proteins and five RNAs and is similar in size to a ribosome
- The five RNAs (U1, U2, U4, U5, and U6) are collectively called small nuclear RNAs (snRNAs). These RNAs are complexed with several proteins. These RNA–protein complexes are called small nuclear ribonuclear proteins (snRNPs).
- The spliceosome is the large complex made up of these snRNPs.
- The snRNPs have 3 roles in splicing: they recognize the 5' splice site and the branch site; they bring those sites together as required; and they catalyze the RNA cleavage and joining reactions.
- To perform these functions, RNA–RNA, RNA–protein, and protein–protein interactions are all important.

Some RNA–RNA hybrids formed during the splicing reaction.



(a) snRNP U1 and U6 recognizing the 5' splice site.
(b) snRNP U2 recognizing the branch site.
(c) The RNA:RNA pairing between U2 and U6.
(d) The same sequence within the pre-mRNA is recognized by a protein (not part of an snRNP) at one stage and displaced by an snRNP at another.

Each of these changes accompanies the arrival or departure of components of the spliceosome and a structural rearrangement that is required for the splicing reaction to proceed.

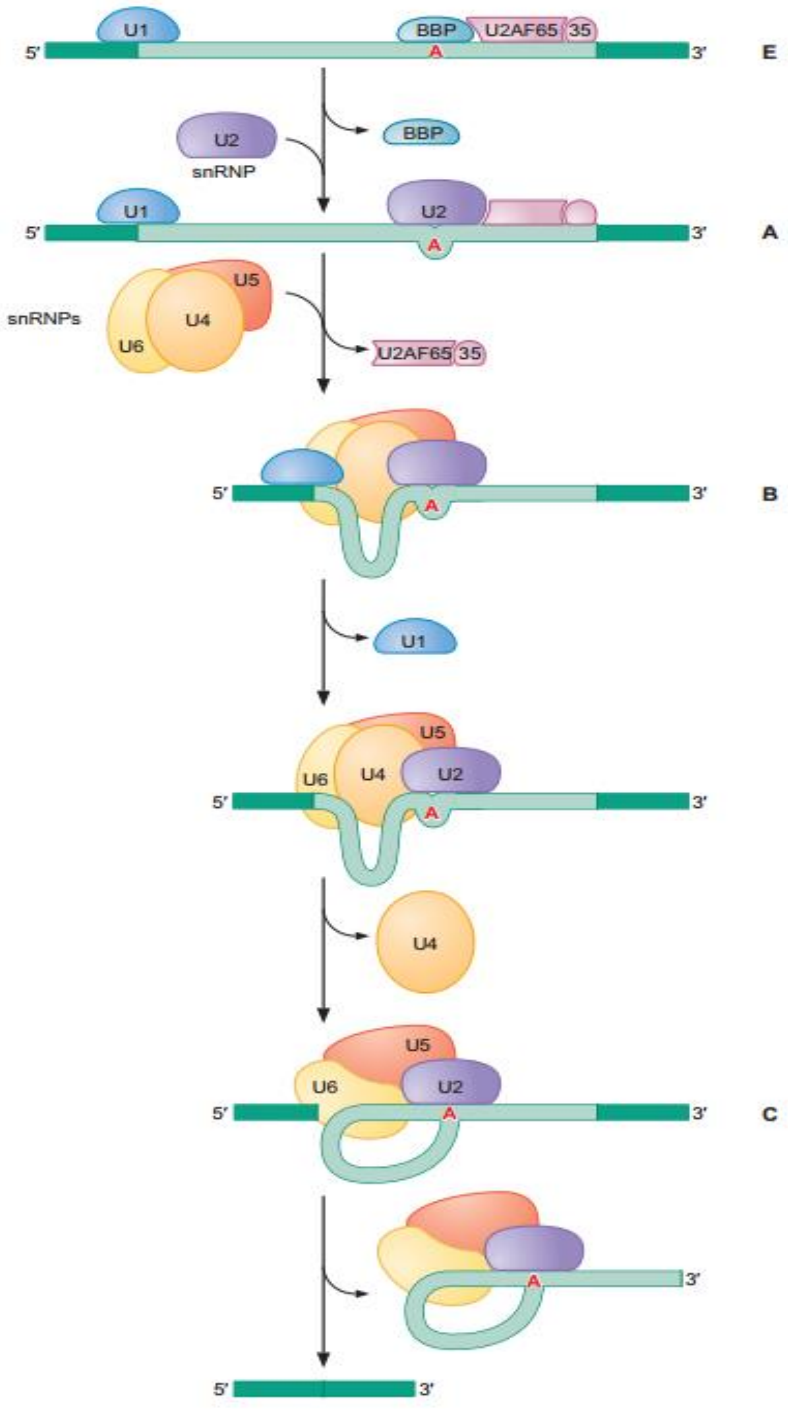
Other proteins involved in the splicing reaction include

- RNA-annealing factors, which help load snRNPs onto the mRNA, and

- DEAD-box helicase proteins, that use their ATPase activity to dissociate given RNA–RNA interactions.

SPLICING PATHWAYS:

Assembly, Rearrangements, and Catalysis within the Spliceosome

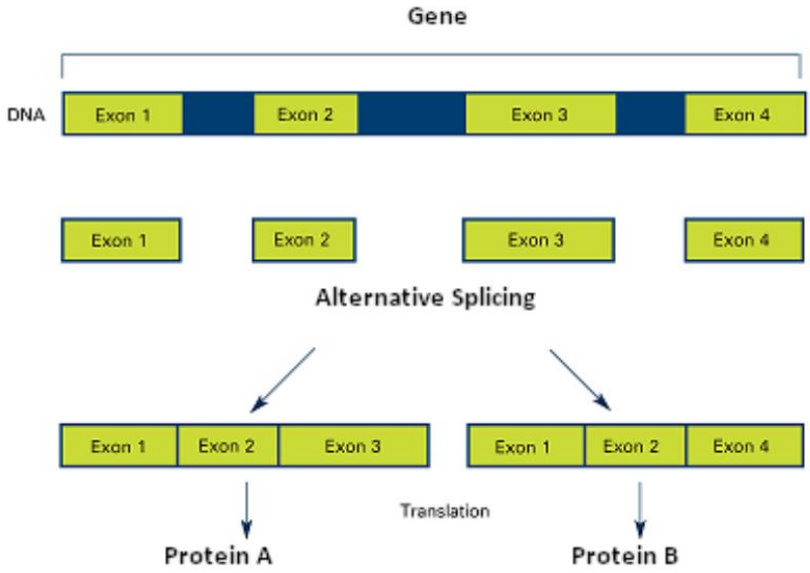


- **Early (E) complex:** U1 snRNP recognizes the 5′ splice site. U2AF binds to the Py tract and 3′ splice site. U2AF interacts with BBP and helps binding the branch site. This arrangement is called the E complex.
- **A complex:** U2 snRNP then binds to the branch site displacing BBP. This arrangement is called the A complex.
- **B complex:** All three splice sites brought together by tri-snRNP particle (U4, U6 and U5 snRNP). This arrangement is called the B complex.

- Then U1 leaves the complex, and U6 replaces it.
- **C complex:** Then U4 is released, and U6 interact with U2. 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 mRNA product and the snRNPs.

Alternative Splicing:

- Alternative splicing is a common posttranscriptional process in eukaryotic organisms, by which multiple distinct functional transcripts are produced from a single gene.
- Alternative splicing is a process that enables a mRNA to direct synthesis of different protein variants (isoforms) that may have different cellular functions.
- It occurs by rearranging the pattern of intron and exon elements that are joined by splicing to alter the mRNA coding sequence.
- During RNA splicing exons are either retained in the mRNA or targeted for removal in different combinations to create diverse mRNA from a single pre mRNA. This process is known as alternative RNA splicing.

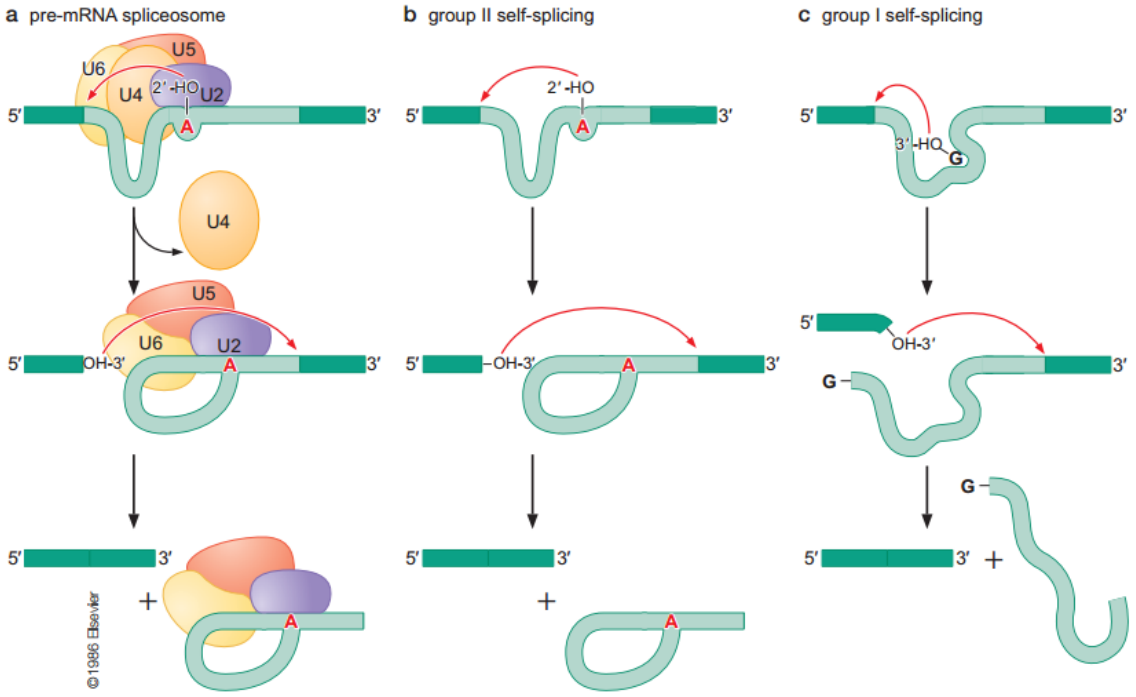


Self-splicing Intron:

- The introns that itself folds into a specific conformation within the precursor RNA and catalyzes the chemistry of its own release not requiring any other proteins.
- The self-splicing introns are grouped into two classes on the basis of their structure and splicing mechanism

Class	Abundance	Mechanism	Catalytic Machinery
Nuclear pre-mRNA	Very common; used for most eukaryotic genes	Two transesterification reactions; branch site A	Major and minor spliceosomes
Group II introns	Rare, some eukaryotic genes from organelles and prokaryotes	Two transesterification reactions; branch site A	RNA enzyme encoded by intron (ribozyme)

Group I introns	Rare, nuclear rRNA in some eukaryotes, organelle genes, and a few prokaryotic genes	Two transesterification reactions; branch site G	RNA enzyme encoded by intron (ribozyme)



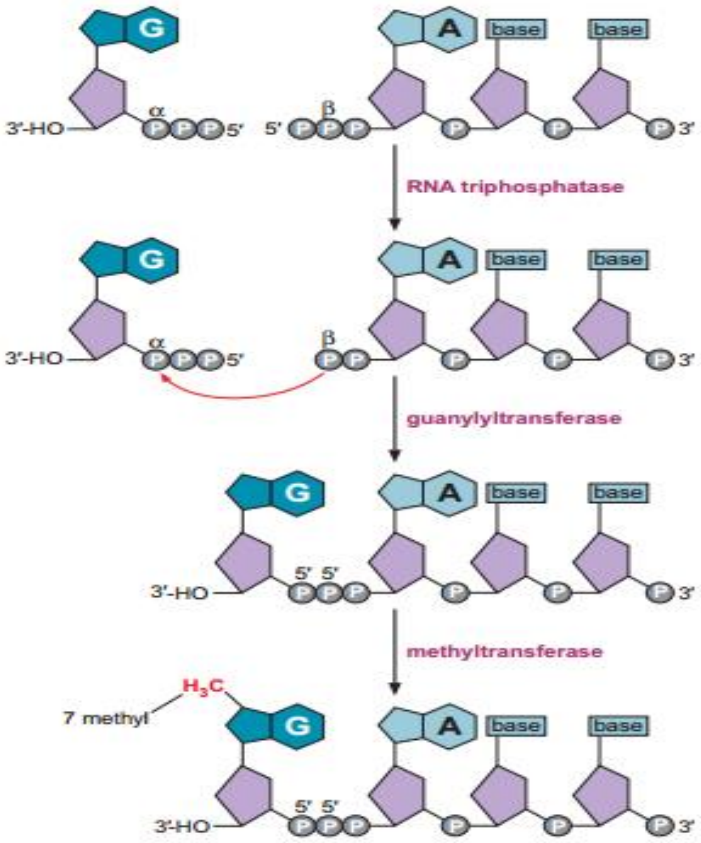
Group I vs Group II intron splicing

- The chemistry in the case of group II introns is essentially the same as in the spliceosome case, with a highly reactive adenine (A) within the intron initiating splicing and leading to the formation of a lariat product.
- In the case of the 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 this case, the intron is linear rather than a lariat structure.

5′ Capping of RNA:

Capping involves the addition of a modified guanine (G) base to the 5′ end of the RNA. It is a methylated guanine, and is joined to the RNA transcript by an **unusual 5′-5′** linkage involving three phosphates.

- **First step:** the γ-phosphate at the 5′ end of the RNA is removed by an enzyme called RNA triphosphatase.
- **Next step:** the enzyme guanyltransferase adds a GMP moiety to the resulting terminal β-phosphate. First, an enzyme–GMP complex is generated from GTP with release of the β and γ- phosphates of that GTP, and then the GMP from the enzyme is transferred to the β -phosphate of the 5′ end of the RNA.
- The newly added guanine at the original 5′ end of the mRNA is 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.

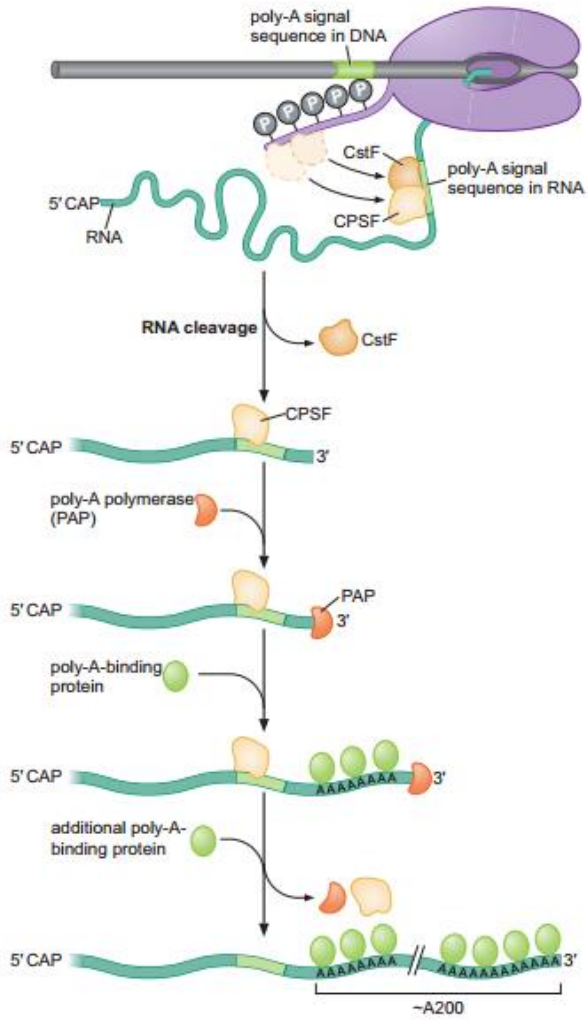


Functions of 5′ cap

- Protection of mRNA from degradation.
- Transport of the mRNA from nucleus to cytoplasm.
- Binding of ribosome with mRNA

3′ poly A tail of RNA:

- All eukaryotic mRNAs have a series of upto 250 adenosines at their 3′ end called poly (A) tail.
- Poly A tail plays an important role in mRNA stability, nucleocytoplasmic export and translation.
- Polyadenylation is mediated by an enzyme called **poly-A polymerase (PAP)**, which adds approximately 250 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.
- PAP is a template independent polymerase.



Ribozymes:

- The ribozyme (Ribonucleic acid enzymne) is an RNA molecule that is capable of performing specific biochemical reactions, similar to the action of protein enzyme.
- It contains an active site that consists entirely of RNA

RNA editing:

- RNA editing can change the sequence of an RNA after it has been transcribed.
- The protein produced upon translation is different from that predicted from the gene sequence.
- During editing, individual bases are either inserted, deleted, or changed.
- That is, the coding information in the RNA is altered.

Two mechanisms of RNA editing:

(a) site-specific deamination of adenines (A) or cytosines (C)

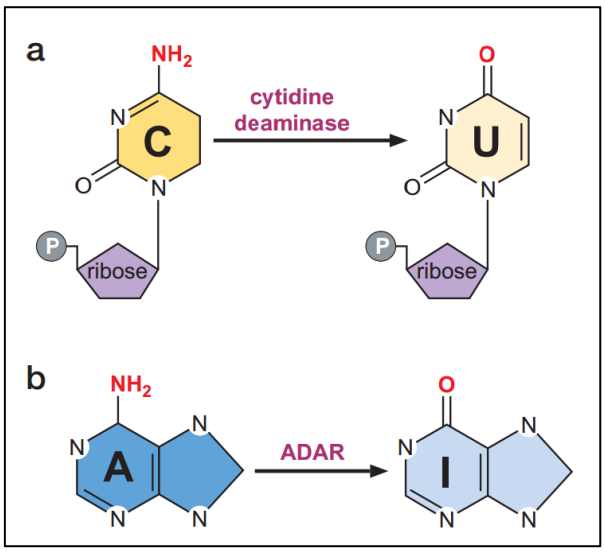
(b) Guide RNA–directed uridine (U) insertion or deletion.

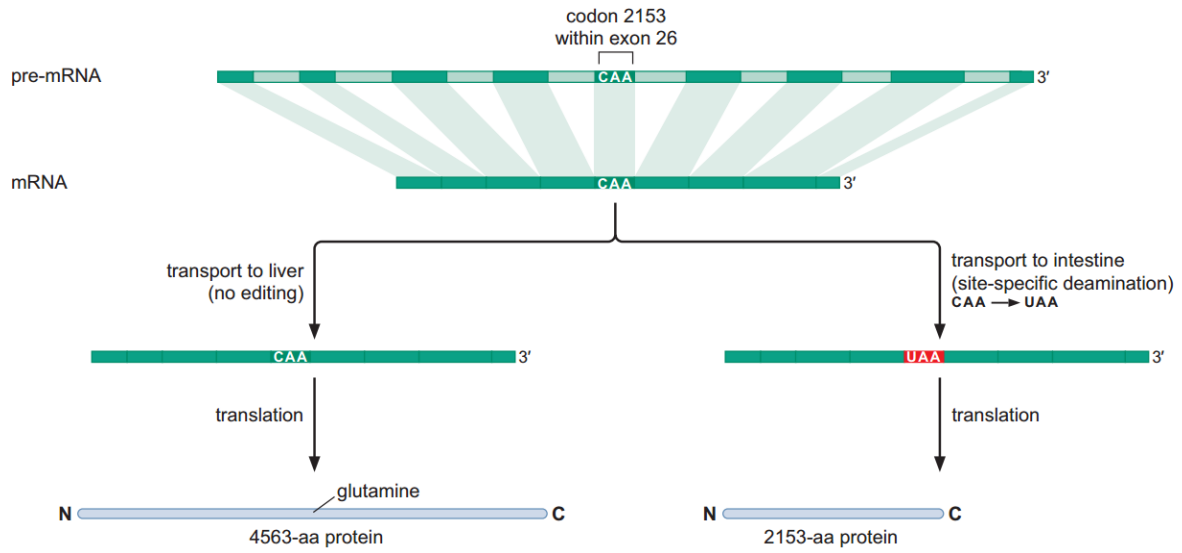
C deamination:

- A specifically targeted cytosine (C) residue within mRNA is converted into uridine by deamination, performed by the enzyme **cytidine deaminase**.

A deamination:

- Converts A into I (inosine) by deamination.
- Enzyme **ADAR** (adenosine deaminase acting on RNA)—produces inosine.
- Inosine can base-pair with cytosine, and thus can alter the protein sequence.





Transport endogenously synthesized cholesterol and triglycerides

Transport dietary lipids to various tissues.

Guide RNA–directed uridine (U) insertion or deletion.

- Multiple Us are inserted by guide RNAs (gRNAs) into specific regions of mRNAs after transcription.
- gRNAs range from 40 to 80 nucleotides in length.

- The gRNA and mRNA form an RNA–RNA duplex with looped-out single-stranded regions opposite where Us are inserted.
- An endonuclease recognizes and cuts the mRNA opposite these loops.
- Editing involves the transfer of Us into the gap in the message.
- This process is catalyzed by the enzyme 3′ terminal uridylyl transferase (TUTase).
- The two halves of the mRNA are joined by an RNA ligase, and the "editing" region of the gRNA continues its action along the mRNA in a 3′ to 5′ direction.

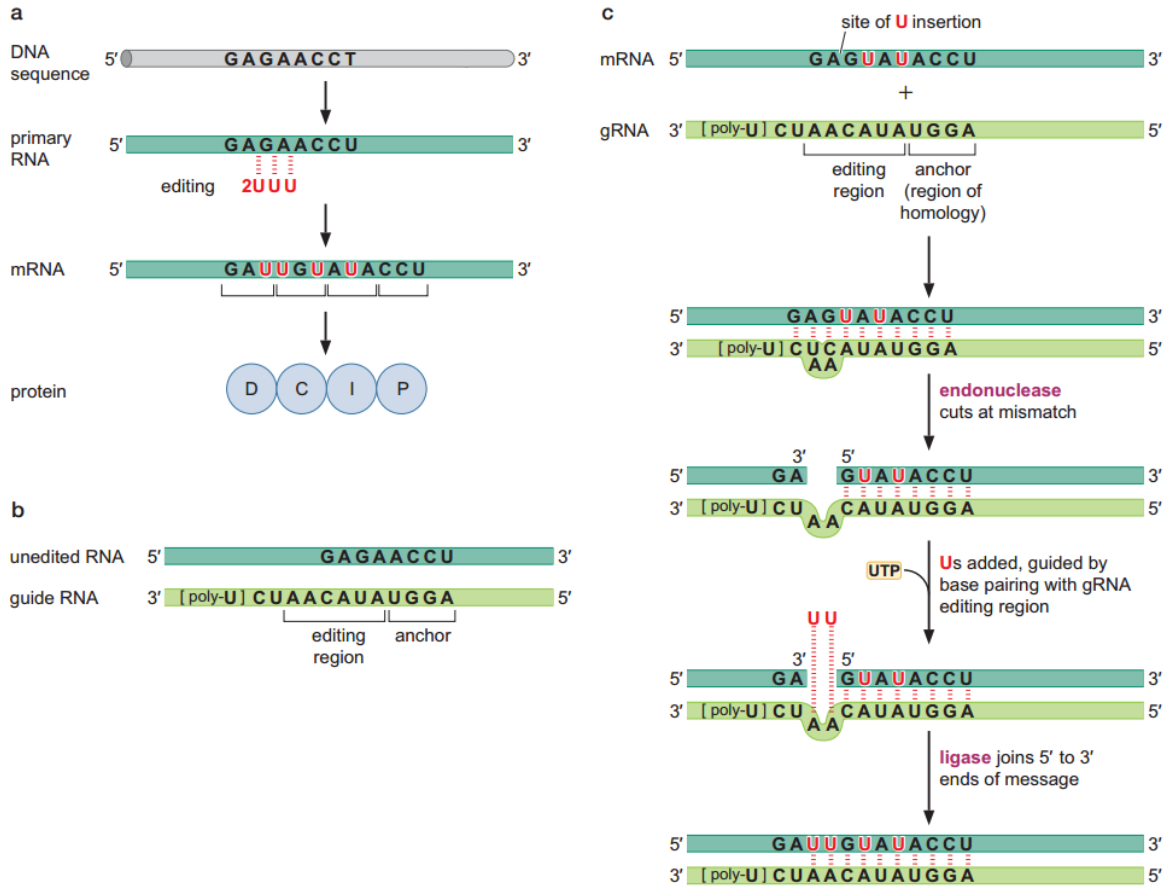


Fig: Guide RNA–directed uridine (U) insertion or deletion

mRNA transport:

- Once mRNA has been fully processed (capped, spliced, and polyadenylated), it is transported out of the nucleus and into the cytoplasm, where it is translated to give its protein product.
- Export takes place through the nuclear pore complex.
- RNA export from the nucleus is an active process, that requires energy.
- Some proteins associated with the RNA carry nuclear export signals that are recognized by export receptors that guide the RNA out through the pore.

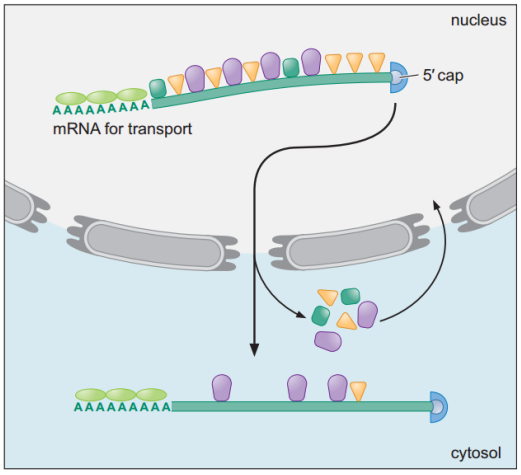


Fig: mRNA transport