

10 - RNA Modifications

After the RNA molecule is produced by transcription (Part 9), the structure of the RNA is often modified prior to being translated into a protein. These modifications to the RNA molecule are called **RNA modifications** or **posttranscriptional modifications**. Most RNA modifications apply only to eukaryotic RNA transcripts.

Key Questions

- Which group of organisms modify their RNA transcripts?

RNA Modifications

The modifications to eukaryotic RNA transcripts include the following:

- **RNA processing.** RNA processing involves cutting large precursor RNA transcripts into smaller ones. RNA processing can involve both **exonucleases** (removing nucleotides from the ends of the RNA transcript) or **endonucleases** (cleaving the RNA transcript at internal sites). The ribosomal RNA (rRNA) molecules that are essential components within ribosomes (see Part 11) commonly experience RNA processing.
- **RNA splicing.** Most eukaryotic genes are **split genes**, being composed of both **intron** sequences (or intervening sequences) and **exon** sequences (or expressed

sequences). For split genes, initial transcription in the nucleus produces a **precursor mRNA (pre-mRNA)**. This pre-mRNA is then spliced, meaning that **introns** are removed and discarded (see **Figure 10.1**). The remaining exon RNA segments are spliced together to produce a **mature mRNA** molecule that is transported to the cytoplasm of the cell for translation.

- **5' end capping.** The 5' end capping process involves the attachment of a modified nucleotide called **7-methylguanosine (7-mG)** to the 5' end of pre-mRNA molecules. The added 7-mG is sometimes called the **5' cap**.
- **3' end polyadenylation.** 3' end polyadenylation involves the addition of a string of adenine (A) nucleotides to the 3' end of the pre-mRNA molecule. The added sequence of A nucleotides is called the **polyA tail**.
- **RNA editing.** RNA editing involves changing the nucleotide sequence of a mRNA molecule prior to translation.
- **Base modification.** Occasionally, nitrogenous bases within the RNA transcript are covalently modified by the addition of chemical groups, including methyl groups.

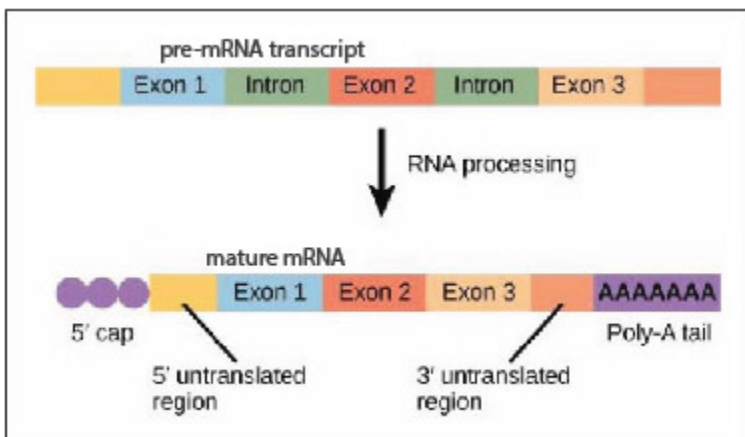


Figure 10.1 RNA Modifications Overview --- Image used from OpenStax (access for free at <https://books.byui.edu/-vuzA>)

Key Questions

- What is the difference between a pre-mRNA and a mature mRNA?
- What is the difference between an intron and an exon?
- What is meant by 5' end capping?
- What is meant by 3' end polyadenylation?

5' End Capping

The 5' end of the pre-mRNA molecule is modified by the addition of a **7-methylguanosine (7-mG)** nucleotide. The process of adding the 7-mG to the pre-mRNA is **5' end capping**. 5' end capping occurs while the transcription of the pre-mRNA occurs and is the first RNA modification. 5' end capping (see **Figure 10.2**) involves the following enzymes:

1. **RNA 5'-triphosphatase.** RNA 5'-triphosphatase removes one of the three phosphates from the nucleotide at the 5' end of the pre-mRNA transcript.
2. **Guanylyltransferase.** Guanylyltransferase cleaves GTP to produce GMP and pyrophosphate (PP_i). Guanylyltransferase then attaches the 5' carbon of the GMP molecule to the 5' carbon on the nucleotide at the 5' end of the pre-mRNA transcript. It is important to note that an unusual 5' to 5' covalent bond is formed. There are now three phosphate groups between these two adjacent nucleotides.
3. **Methyltransferase.** Methyltransferase attaches a methyl group to the added guanine base, producing the 7-mG cap.

The 7-mG cap does the following:

- Serve as a binding site for proteins that transport the mRNA from the nucleus to the cytoplasm of the cell.
- Serve as a recognition site for translation factor proteins that

help the ribosome bind to the mRNA. Once the ribosome binds to the mRNA, translation begins (see Part 11).

- Protect the 5' end of the mRNA transcript from degradation by exonucleases.
- Regulate RNA splicing.

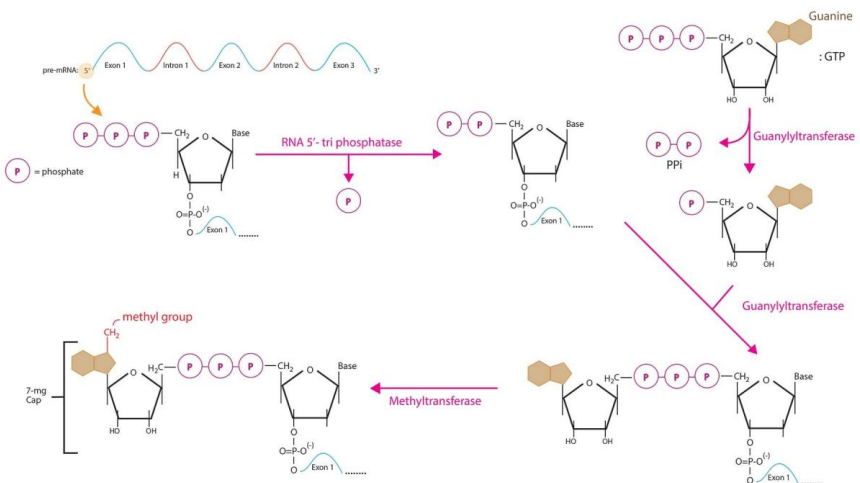


Figure 10.2 5' end capping mechanism --- Image created by JET

Key Questions

- How does the 7-mG structure contribute to translation?
- Which nucleotide provides the energy for 5' end capping?
- What is unusual about the covalent bond between 7-mG and the rest of the pre-mRNA molecule?

3' End Polyadenylation

The 3' end of the pre-mRNA is modified by the addition of a **polyA tail**, a string of approximately 250 adenine (A) nucleotides. The process of adding a polyA tail to the mRNA transcript (see **Figure 10.3**), called **3' end polyadenylation**, involves:

1. The detection of two recognition sequences, called **polyadenylation signal sequences**. Both polyadenylation signal sequences are located near the 3' end of the pre-mRNA.
 - The first polyadenylation signal sequence is 5'-AAUAAA-3'. This polyadenylation signal sequence is recognized by the endonuclease **cleavage and polyadenylation specificity factor (CPSF)**.
 - The second polyadenylation signal sequence, enriched in guanine and uracil bases, is called the **GU-rich sequence**. This GU-rich sequence is the binding site for the **cleavage stimulatory factor (CstF)** protein.
 - CstF activates CPSF.
2. CPSF cleaves the pre-mRNA 10-35 nucleotides downstream (farther towards the 3' end) of the 5'-AAUAAA-3' sequence. This free 3' end of the pre-mRNA is then available for the addition of adenine nucleotides.
3. **Poly(A)-polymerase (PAP)** attaches as many as 250 adenine nucleotides to the 3' end of the pre-mRNA transcript.

The polyA tail added by 3' end polyadenylation functions as follows:

- The polyA tail protects the 3' end of the pre-mRNA transcript from degradation by exonucleases.
- The polyA tail facilitates the transport of the mRNA from the nucleus to the cytoplasm of the cell.
- The polyA tail helps the ribosome bind to the mRNA to initiate translation.

The 3' end polyadenylation process occurs after 5' end capping, but prior to RNA splicing. In fact, 3' end polyadenylation assists in terminating transcription in eukaryotes by the torpedo model (see Part 9).

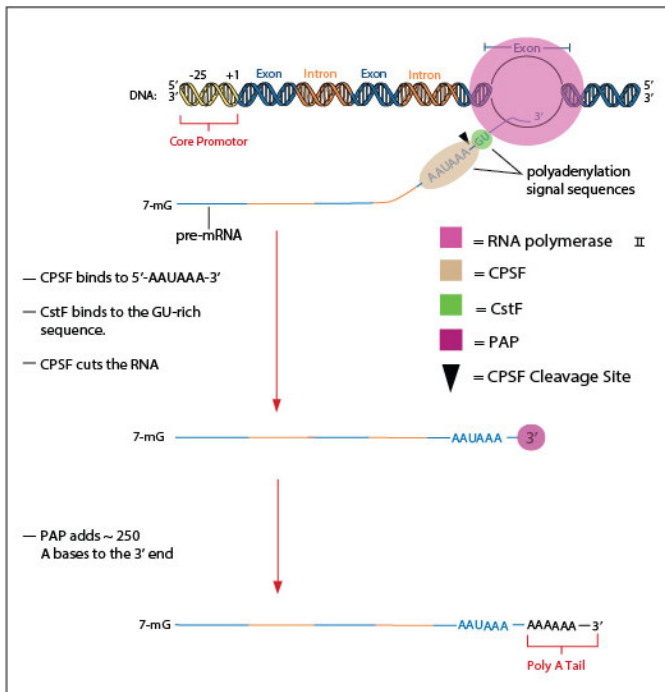


Figure 10.3 3' end polyadenylation mechanism --- Image created by SL

Key Questions

- How does 3' end polyadenylation contribute to transcription termination in eukaryotes?
- Describe the functions of CPSF, CstF, and Poly(A)-polymerase.

Splicing of Group I and Group II Introns

There are three general mechanisms to remove introns from RNA molecules. The **group I** and **group II** mechanisms are limited to certain eukaryotes or certain organelles in the cell. For example, the group I mechanism removes the introns found in ribosomal RNA (rRNA) molecules in certain protozoa. The group II mechanism

removes the introns found in the mRNA and transfer RNA (tRNA) transcripts produced by mitochondrial and chloroplast genes. The **spliceosome mechanism** is the major mechanism to remove introns from pre-mRNA transcripts. The spliceosome mechanism removes introns from the pre-mRNAs produced by most structural genes in the nucleus of eukaryotic cells.

- **Removing group I introns.** RNA splicing of group I introns occurs by **self-splicing**, meaning that the precursor RNA molecule catalyzes the removal of its own intron (see **Figure 10.4**). These catalytic precursor RNAs molecules are also called **ribozymes**. The self-splicing of group I introns involves the following mechanism:

1. A free **guanosine nucleoside** (guanine nitrogenous base covalently linked to a ribose sugar) binds to a pocket within the intron. The guanosine nucleoside bound to the intron serves as a enzyme cofactor (i.e., assists the ribozyme in catalysis) for the remaining steps in the reaction.
2. A break forms at the junction between the 3' end of the first exon and the 5' end of the intron. The guanosine nucleoside becomes attached to the 5' end of the intron.
3. The released exon cleaves the junction between the 3' end of the intron and the 5' end of the second exon.
4. A phosphodiester bond is formed that links the first and second exons, generating a mature RNA transcript. The intron is released and degraded.

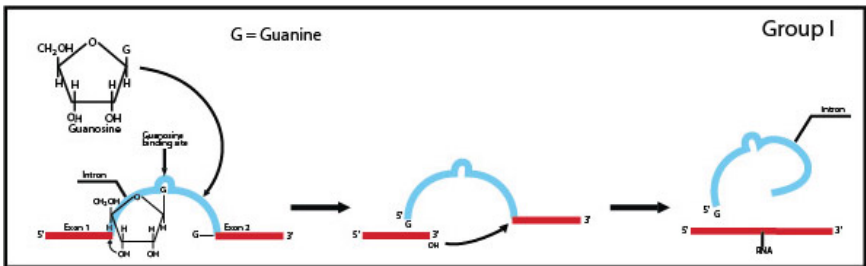


Figure 10.4 Removing group I introns --- Image created by SL

- Removing group II introns.** RNA splicing of group II introns also occurs by self-splicing, meaning that the precursor RNA removes its own intron (see **Figure 10.5**). In other words, pre-mRNAs containing group II introns are also ribozymes. The self-splicing of group II introns involves the following:

1. The 2'-OH group of an **adenine nucleotide** within the intron helps to cleave the junction between the 3' end of the first exon and the 5' end of the intron. In this reaction, the adenine nucleotide serves as an enzyme cofactor for the reaction.
2. The released exon cleaves the junction between the 3' end of the intron and the 5' end of the second exon.
3. A phosphodiester bond is formed that links the first and second exons, generating a mature RNA transcript. The intron is released and degraded.

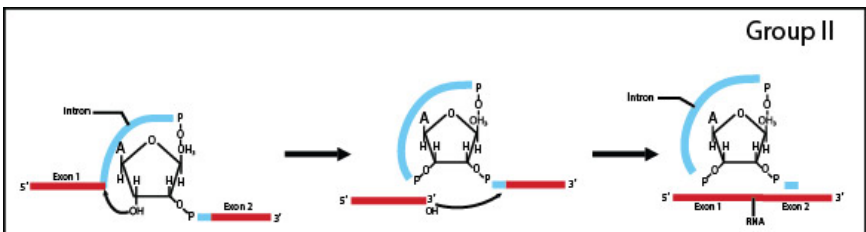


Figure 10.5 Removing group II introns --- Image created by SL

Key Questions

- What is a ribozyme?
- Describe the major events that occur in the group I and group II splicing mechanisms.
- What molecules serve as enzyme cofactors in the group I and group II splicing mechanisms?

Removal of Introns by Spliceosomes

Transcription of most structural genes in the nucleus of eukaryotic cells produces pre-mRNA molecules; the removal the introns within these pre-mRNA molecules involves a large multi-subunit complex called the **spliceosome**. The spliceosome binds to recognition sequences within the intron RNA sequence (see **Figure 10.6**). These recognition sequences within the intron include:

- **The 5' splice site.** The 5' splice site (also called the **donor sequence**) is a 5'-GU-3' at the 5' end of the intron RNA sequence.
- **The branch site.** The branch site is an adenine nucleotide (A) near the middle of the intron RNA sequence.
- **The 3' splice site.** The 3' splice site (also called the **acceptor sequence**) is a 5'-AG-3' at the 3' end of the intron RNA sequence.

The subunits of the spliceosome are called **small nuclear ribonucleoproteins** or **snRNPs** ("snurps"). snRNPs are composed of uracil-rich **small nuclear RNAs (snRNAs)** that act as RNA enzymes (ribozymes) to remove the introns from nuclear pre-mRNA molecules. snRNPs are also composed of protein subunits that function to stabilize the spliceosome.

The spliceosome splicing mechanism occurs as follows:

1. The **U1** snRNP binds to the 5' splice site, and the **U2** snRNP binds to the branch site adenine within the intron RNA sequence.
2. Additional snRNPs called **U4**, **U5**, and **U6** bind to the intron. These five snRNPs (U1, U2, U4, U5, and U6) assemble to form the spliceosome complex.
3. The intron loops out bringing the two exon sequences close together.
4. The 5' splice site is cut by U1, and the 5' end of the intron is covalently linked to the 2'-OH group of the branch site adenine, forming an RNA loop structure called a **lariat**.
5. The U1 and U4 snRNPs are released.
6. The 3' splice site is cut by the U5 snRNP.
7. A phosphodiester bond is formed that links the two exons together to form the mature mRNA. The intron is released along with U2, U5, and U6.

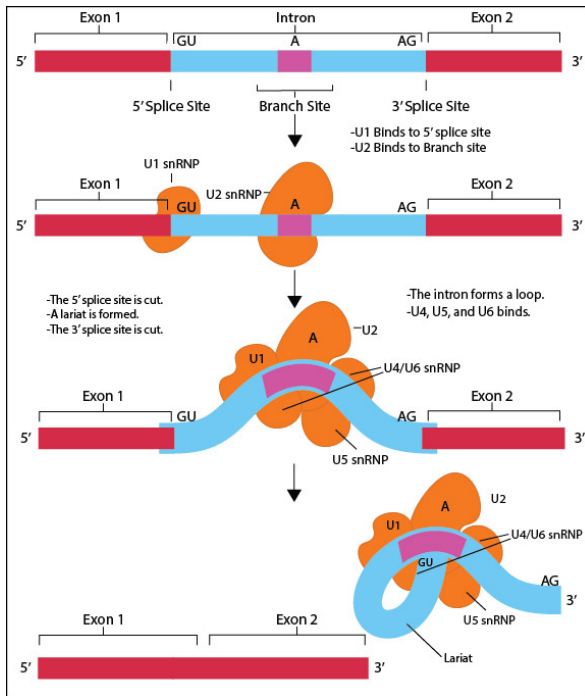


Figure 10.6 *Spliceosome splicing* --- Image created by SL

Key Questions

- Which two splicing mechanisms are found in human cells?
- What are the names of the three sequences found within spliceosome introns?
- Which spliceosome components are ribozymes?
- What are the functions of the U1 and U5 snRNPs?

Identifying Introns Using R Loop Experiments

Introns were initially identified within the β -globin and ovalbumin genes by performing **R-loop (hybridization) experiments**. These experiments relied on denaturing an isolated DNA molecule that

contains a gene, allowing a mRNA molecule to form hydrogen bonds (hybridize) with the template DNA strand, adding the coding strand DNA, which attempts to form hydrogen bonds with the template DNA strand, and finally, examining the resulting nucleic acid structure in an electron microscope. What would such a molecule look like at the end of the R-loop (hybridization) experiment?

- **Gene hybridized to the pre-mRNA.** The pre-mRNA that has formed hydrogen bonds with the template DNA strand prevents the coding DNA strand from binding. Because the coding DNA strand fails to bind to the template DNA strand, the coding DNA strand loops out from the RNA-DNA hybrid region. This loop where the coding DNA strand cannot bind to the template DNA strand is called an RNA displacement loop or **R loop** (see **Figure 10.7**).
- **Gene hybridized to the mature mRNA.** When the DNA contains introns, hybridization between the template DNA strand and the mature mRNA forces the intron DNA sequences from the template DNA strand to loop out, because the mature mRNA lacks a corresponding sequence that can bind to the template strand introns. Adding the coding DNA strand produces R-loops with an intervening region of double-stranded DNA (i.e., the intron sequences within the template and the coding DNA strands from hydrogen bonds) called an **intron loop**.

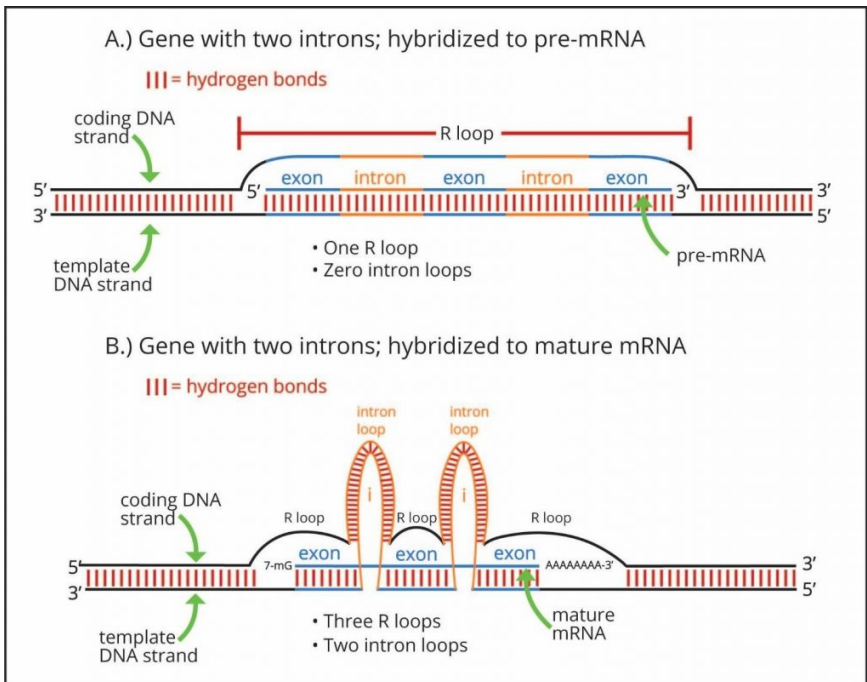


Figure 10.7 R-Loop Results --- Image created by SL

Key Questions

- Suppose a gene contains four introns and is hybridized with its mature mRNA. How many R loops would be observed in the electron microscope at the end of an R loop experiment? How many intron loops would be observed?

Identifying Introns by Comparing gDNA with cDNA

Introns within genes can also be identified by comparing the length of a **genomic DNA (gDNA)** version of a gene to the **complementary DNA (cDNA)** version of the same gene. gDNA is the version of a gene found in the genome; the gDNA version of a gene contains both introns and exons. cDNA is produced in the laboratory by **reverse**

transcription (see Part 8). Reverse transcription converts mature mRNA into a cDNA molecule using the viral enzyme **reverse transcriptase**. Since the cDNA molecule is produced from the mature mRNA, cDNA molecules contain exons but lack introns. The gDNA version of the gene, which contains introns, will be longer than the cDNA version of the same gene, which lacks introns.

The **polymerase chain reaction (PCR)** technique (see Part 8) can be used to make billions of copies of the gDNA and the cDNA versions of any gene of interest. The gDNA and cDNA PCR products are then separated by size using **agarose gel electrophoresis** (see Part 8). The size difference between the gDNA and the cDNA copy of the gene can be easily observed on an agarose gel (see **Figure 3.2**).

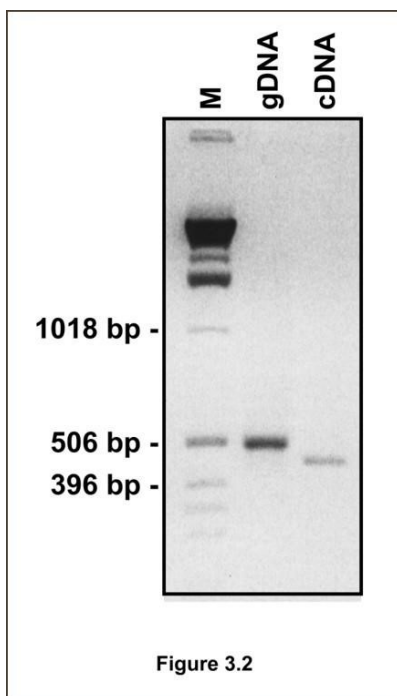


Figure 10.8 Comparing cDNA to gDNA to identify introns --- Image provided by K. Mark DeWall

Key Questions

- What is the difference between gDNA and cDNA?
- How can comparing gDNA to cDNA on an agarose gel help you determine that a gene contains introns?

Alternative Splicing

Alternative splicing involves splicing a single type of pre-mRNA molecule in various ways to produce different mature mRNA molecules (see **Figure 10.9**). Each of these mature mRNAs can then produce a slightly different protein upon translation. These distinct, yet related protein **isoforms**, all derived from a single gene, can have specialized functions.

Alternative splicing is beneficial in that it allows eukaryotes to carry fewer genes in the genome, permitting a relatively small number of genes the flexibility to encode a vast array of proteins. In humans, it is estimated that 30–60% of the genes in the genome are alternatively spliced. As a result, the human genome, which contains approximately 23,000 structural genes can produce at least ten times that number of protein products.

One example of alternative splicing involves the splicing of the pre-mRNA molecule for α -tropomyosin, a protein involved in muscle contraction. The α -tropomyosin gene contains 14 exons (13 introns). There are two types of exons within the α -tropomyosin pre-mRNA:

- **Constitutive exons.** Constitutive exons are exons that are always included in the mature α -tropomyosin mRNAs products of alternative splicing. These exons likely encode amino acid sequences that maintain the general three-dimensional structure of the encoded protein.
- **Alternative exons.** Alternative exons vary between mature α -tropomyosin mRNAs. In one cell type, one combinations of

alternative exons are spliced together, in another cell type, a different combination of alternative exons are spliced together. The result is two related proteins that have slightly different functions to meet the unique needs of these two cell types.

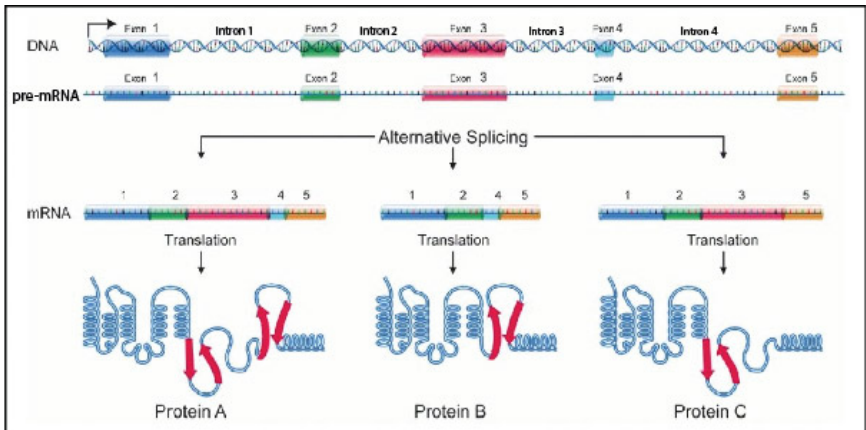


Figure 10.9 Alternative splicing allows one gene to produce three proteins. --- [DNA Alternative Splicing](#) by National Human Genome Research Institute and is used under [CC0](#)

Key Questions

- Why is alternative splicing advantageous to eukaryotic cells?
- What are protein isoforms?
- What is the difference between a constitutive exon and an alternative exon?

Patterns of Alternative Splicing

Alternative splicing is regulated by **splicing factor** proteins. These splicing factor proteins help the spliceosome choose which splice sites to use during RNA splicing. Different cell types have different splicing factor proteins, allowing different RNA splicing patterns to occur in each cell type. The **SR proteins** are an example of a group of splicing factor proteins found in animals, including humans.

Here are some common alternative splicing patterns observed in eukaryotic cells:

- **Exon Skipping.** Some splicing factor proteins act as **splice repressors**. Splice repressor proteins prevent the spliceosome from recognizing a particular splice site within an intron (see Figure 10.10). For example, when a splice repressor protein blocks a 3' splice site within an intron, the 3' splice site in the next intron is chosen instead, and the intervening exon is removed during RNA splicing (**exon skipping**).
- **Alternative 5' and 3' Splice Sites.** In addition to the 5' splice site, the branch site, and the 3' splice sites discussed earlier, there are other pre-mRNA sequences involved in splicing. These additional sequence elements, often located within a nearby exon, can promote the use of a particular 5' or 3' splice site. For example, some potential 5' or 3' splice sites in the pre-mRNA are normally poorly recognized by the spliceosome. In certain cell types, the binding of a **splice activator protein** to a **splice enhancer sequence** within a nearby exon promotes the use of these otherwise poorly recognized 5' or 3' splice sites (see Figure 10.10). When a splice activator protein binds to a splice enhancer sequence, a nearby exon is found in the mature mRNA (i.e., the exon is not skipped).
- **Mutually Exclusive Exons.** In some cases, splicing events are coordinated between different cells to ensure that unique protein isoforms are produced in different cell types. For example, suppose there are four exons (three introns) in a pre-mRNA molecule. During splicing in one cell type, exon two is consistently retained in the mature mRNA, while exon three is always spliced out. In a different cell type, exon two is always spliced out, while exon three is always retained in the mature mRNA. Exons one and four are found in the mature mRNAs in both cell types and are thus constitutive exons.

Scientists know little about the true complexity of alternative splicing; however, alternative splicing patterns are cell-type and developmental stage specific. Moreover, mutations in human genes often lead to aberrant splicing patterns. This aberrant splicing produces abnormal protein isoforms and ultimately, disease phenotypes.

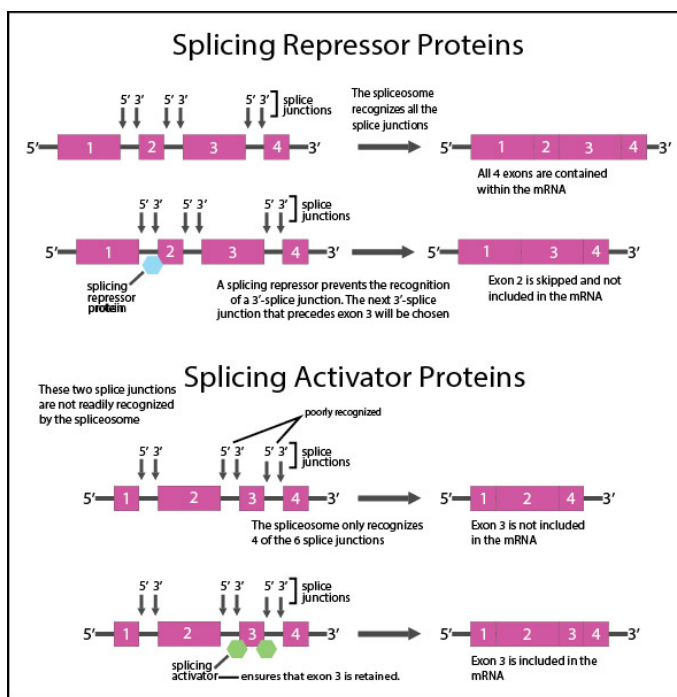


Figure 10.10 Splicing repressor and activator proteins --- Image created by SL

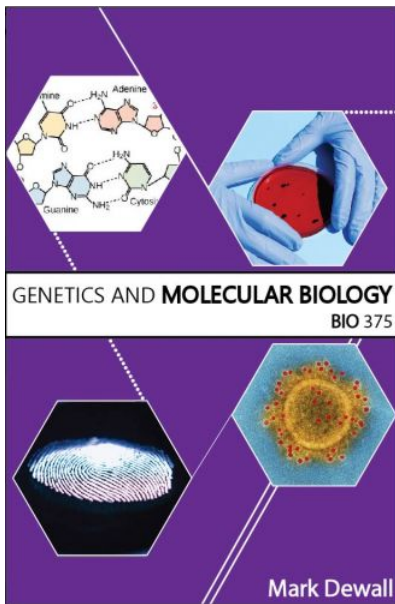
Key Questions

- What happens when a splice repressor protein binds to the 3' splice site within an intron?
- What effect would a splicing activator protein binding to a splicing enhancer sequence have on alternative splicing?
- What is meant by the term mutually exclusive exons?

Review Questions

Fill in the blank:

1. _____ is an endonuclease that releases the pre-mRNA from RNA polymerase II during transcription.
2. _____ is an enzyme that attaches two nucleotides together via a 5' to 5' covalent bond.
3. One function of the 7-mG cap is to _____.
4. A _____ protein prevents the spliceosome from binding to a 3' splice site.
5. _____ is an enzyme that adds adenine nucleotides to the 3' end of a pre-mRNA. These are added in the 5' to 3' direction.
6. The Group I intron splicing mechanism uses the nucleoside _____ as a cofactor during catalysis, while the _____ intron splicing mechanism uses an adenine nucleotide as a cofactor during catalysis.
7. The U2 snRNP binds to the _____ site of the pre-mRNA.
8. Spliceosome subunits are composed of two components: proteins and _____.
9. _____ is a pattern of alternative splicing where one exon is always retained in one cell while that same exon is always skipped in another cell.



Dewall, M. (n.d.). *BIO 375: Genetics and Molecular Biology*. BYU-I Books.

https://books.byui.edu/genetics_and_molecul