

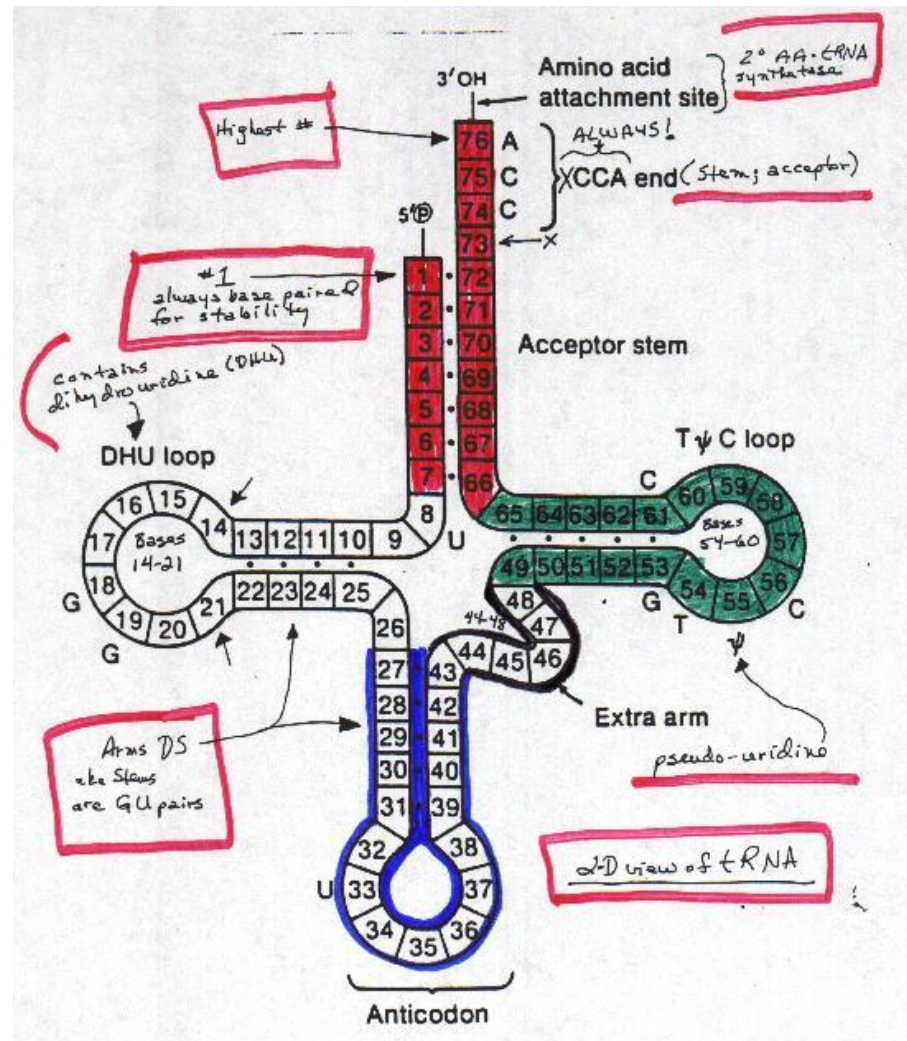
RNA: Transcription and Triplet Code

There are Five Kinds of RNA, All of Which are Templated from DNA.

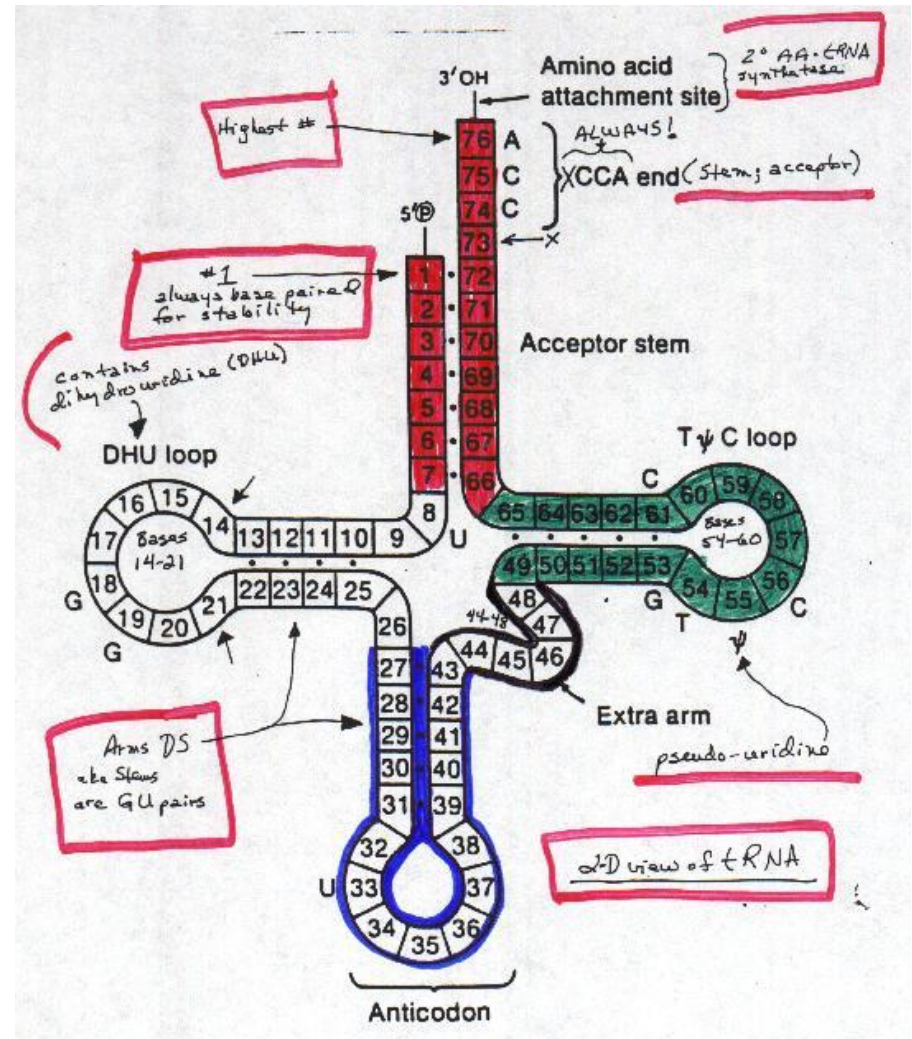
- The first type of RNA is tRNA.
- The "t" stands for "transfer".
- This RNA is the RNA that transfers amino acids to the growing peptide as it is elongating on the ribosome.
- It is single-stranded (SS) and found in the cytosol of the cell.
- tRNA makes up about 15% of all RNA's.

There are at least 20 different kinds of tRNA's as there are at least 20 different amino acids. tRNA has a cloverleaf shape:

1. Bases 1-7 are paired with bases 66-72 to form a double stranded (DS) region in the tRNA that makes it stable/stronger. This region extends through bases 73-76. The whole "arm" is known as the acceptor stem. Note that the 3' -OH is the site of attachment of the amino acid under the direction/catalysis of aminoacyl-tRNA synthetase.
2. Bases 10-13 are paired with bases 22-25 in the DHU loop (on left in graphic). The "H" stands for dihydrouridine (DHU).
3. Bases 27-31 are paired with bases 39-43 to form the anticodon loop of the tRNA (bottom of graphic). Bases 34, 35 and 36 make up the triplet that is the anti-codon.

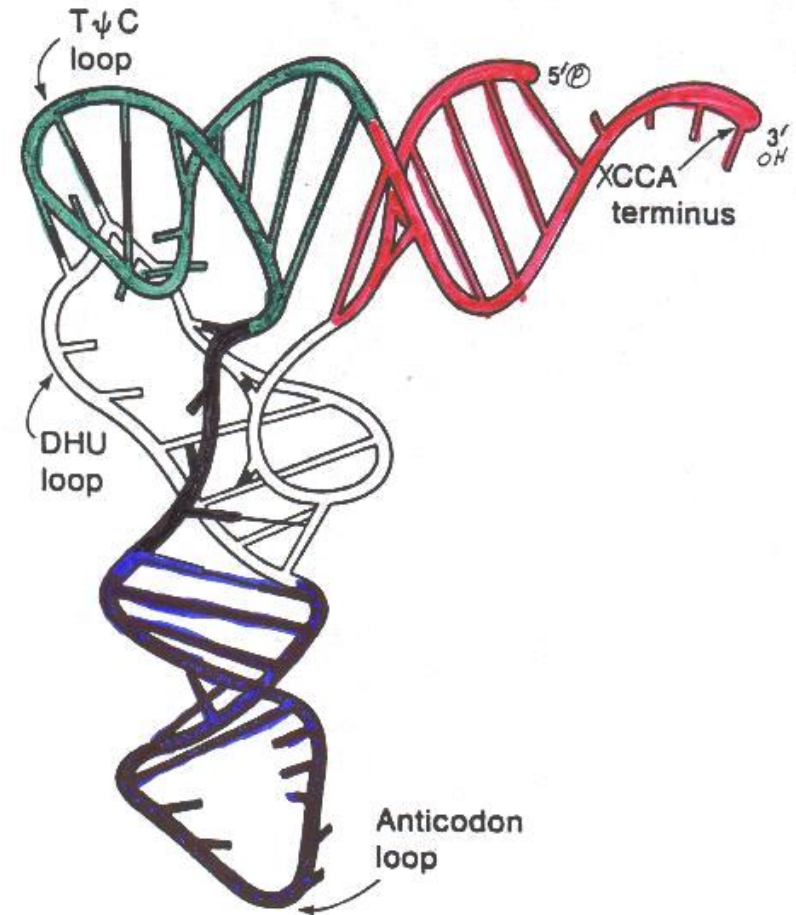


4. Although not DS, bases 44-48 make up the "extra arm" of tRNA.
5. Bases 49-53 pair up with bases 61-65 in the T ψ C loop. ψ is pronounced "sigh" and looks like a three-pronged pitchfork; it represents the presence of pseudo-uridine in this loop.
6. The arms or stems of the loops seem to be primarily GU pairs, a rather odd combination, as we typically think about G and C pairing and A and U pairing.
7. The amino acid acceptor end ALWAYS ends with XCCA, where X is any nucleotide, so that A is always attached to the binding/transferrable amino acid.



tRNA has a cloverleaf shape ... that folds to a more compact "L" shape:

- The region of the tRNA that bonds with the mRNA is called the anti-codon.
- Typically, the tRNA's are identified by which amino acid they transfer, e.g., alanyl tRNA would be represented as "tRNA^{Ala}".
- As with ALL RNA's, the concentration of adenine is not equal to the concentration of uracil ([A] [U]), nor are the concentrations of guanine and cytosine equal ([G] [C]).
- This is due to the fact that RNA is single stranded (SS).



- The second type of RNA is "rRNA" or ribosomal RNA.
- It, too is SS.
- It is found in the ribosomes in the cell.
- rRNA makes up about 80% of the RNA's in the cell.
- It is the most stable of the RNA's and is synthesized only when the cell needs more ribosomes.
- The code for the rRNA is found in nuclear DNA.

Third Type of RNA

- Messenger RNA (mRNA) is also SS.
- It is synthesized in the nucleus, is sent to the cytosol of the cell and binds with ribosomes.
- It makes up less than 5% of the RNA's as it doesn't "survive" long enough to make up much of the RNA's.
- It has a half-life ($t_{1/2}$) of 4-24 hours.
- mRNA is around only long enough to drive the synthesis of its specific protein, then it is recycled.
- mRNA is synthesized from a single gene unlike t and rRNA.
- mRNA may range from 70 nucleotides in length up to 20000 nucleotides in length.
- The 3' terminus carries a poly-A "tail" that consists of 20-200 adenosine residues that are added after mRNA is synthesized.

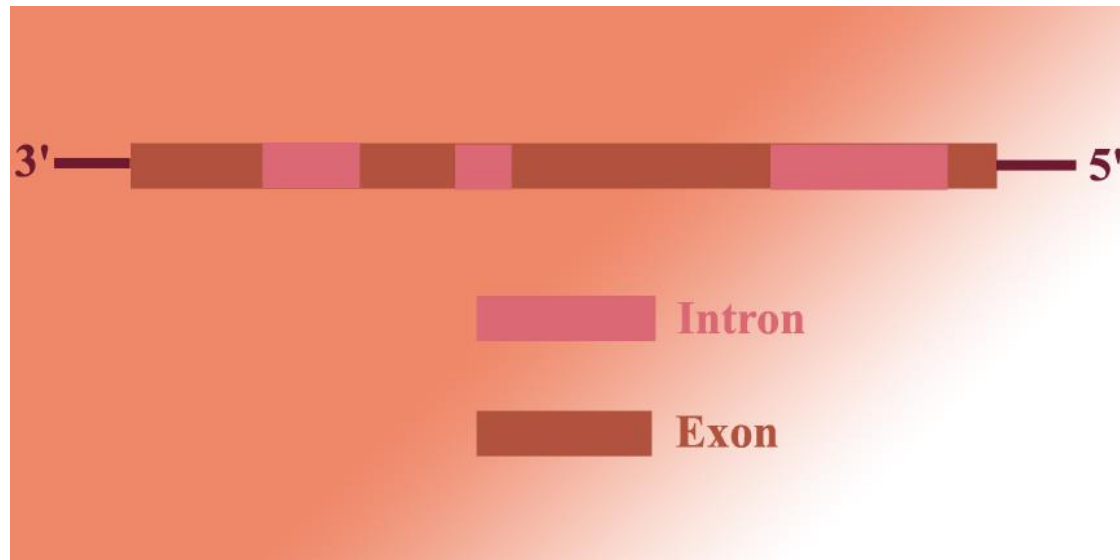
Fourth Kind of RNA

- The mitochondrion makes some of its own RNA, as well, called mitoRNA or mtRNA.
- It is SS and found in, believe it or not, the mitochondrion.
- mitoRNA may be t OR rRNA-type.
- It is utilized in the synthesis of mitochondrial protein.
- Remember, though, that the mitochondrion requires nuclear-coded proteins to function, as well.

Fifth Type of RNA

- The fifth type of RNA is called small nuclear RNA or snRNA (called "snurps").
- It is found in the nucleus of the cell.
- Some snurps are involved in/with RNA processing.
- It consists of and interacts with ribonucleoprotein.
- They are typically named with a "U" followed by a number., e.g., "1", then completed with RNA: U1RNA, U2RNA, U3RNA, *ad nauseum*.

A Generic Gene “Schematic” on An SS Piece of DNA



- Each gene contains units called **exons** (these contain coding information that takes up a very small space in the gene) and **introns** (contains non-coding information; they are not well understood and they take up a lot of space in the gene).
- Hence, not all of the DNA in a single gene is required for the synthesis of a protein.

Enhancers - 1



- This is a gene-specific sequence of DNA, 50-500 nucleotides long and may be repeated a number of times.
- Enhancers are located 100 base pairs (bp's) to thousands of bp's (kbp's) up OR downstream of the DNA sequence that will be read or transcribed.
- Enhancers are conditional, i.e., not every thing "turns them on".
- Enhancers are also known as recognition sites.
- Since enhancers are DNA-coded, they are called "cis" acting.
- Factors that bind to enhancers are called "trans" acting.

Enhancers - 2



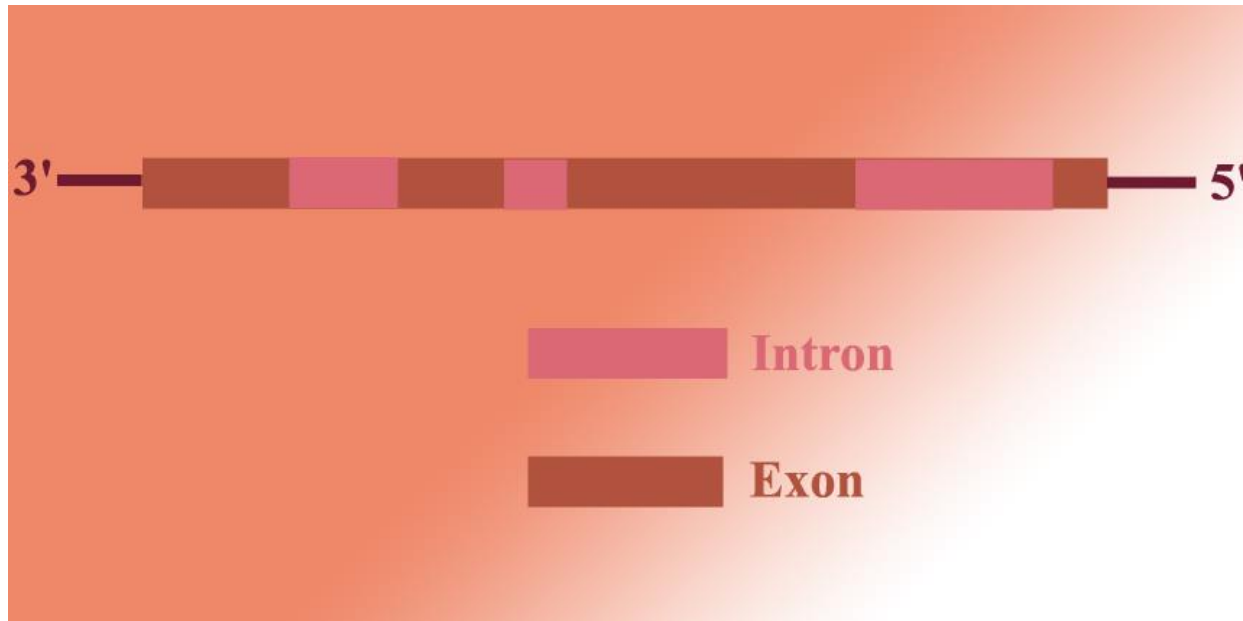
- Three other regions are observable:
 - a sequence of GC's,
 - a region of CAAT and
 - a TATA region.
- All three of these regions are called promoter sites. These are promoters for RNA Pol II (Slide coming). The region between 75-110 bases downstream of the transcription start site (more coming on this later; written as -75 - -110) is called the GC box.
- This is observed only every now and again, i.e., this is rare.
- Between -40 and -75 is a region of four nucleotides, CAAT, called the cat box. This is observed a bit more frequently than the GC box.

Enhancers & Start Box



- The promoter site almost always present is the TATA region between the start site (S) and -30.
- This latter region is called the Hogness box and is about 25 bp up from the transcription start site.
- Data indicates that the TATA box, as it's also called, is the "code" for the first base to be transcribed.

- Enhancers increase transcription more than 100-fold. The properties of a good enhancer are as follows:
 1. They must not possess promoting activity.
 2. They can be 1 kbp up OR down from the promoter.
 3. Some work only in 1 kind of cell, 1 kind of tissue and/or 1 kind of cell type.
 4. Some work where ever they "desire".
 5. They may be spliced into other regions of DNA.
 6. Their removal slows transcription to one-hundredth the activity that it was.
 7. They may be cut out and inserted backwards into the same spot without loss of genetic activity. This latter characteristic says that the enhancers are read front-to-back OR back-to-front for full activity.

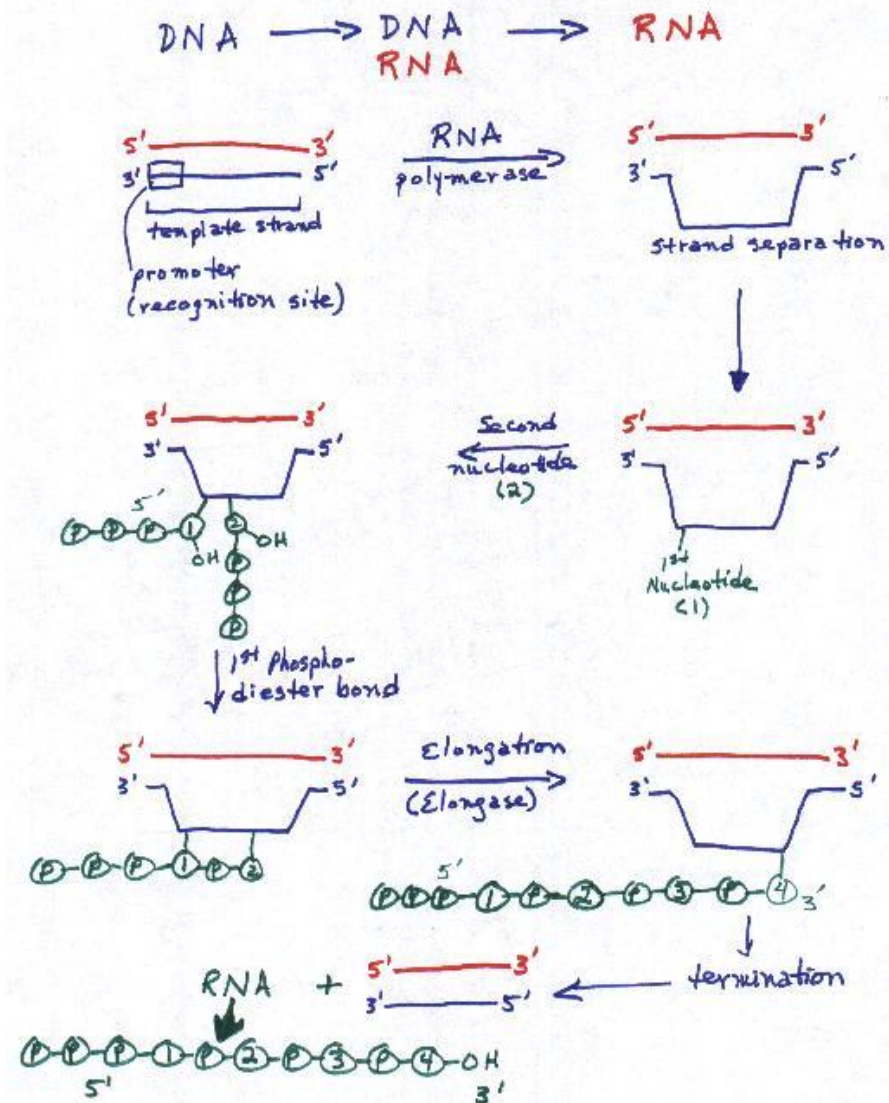


- In order for transcription to continue, the DNA must be unwound.
- Once the DNA is unwound and opened up, RNA Pol II may begin reading this SS strand of DNA

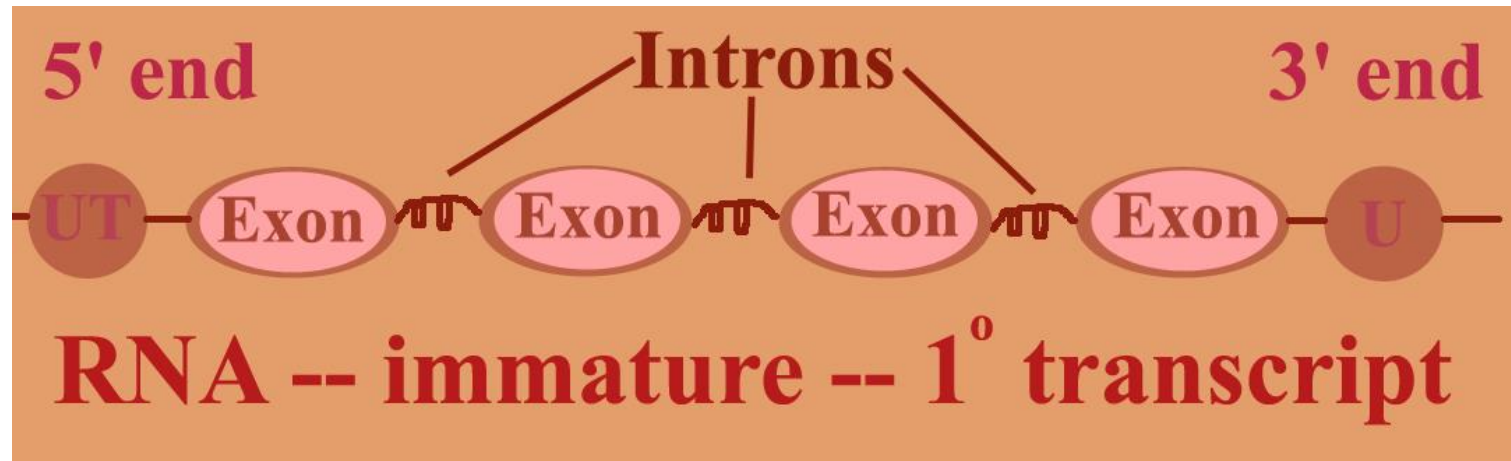
RNA Polymerases (RNA Pol) – An ASIDE

RNA Pol I	RNA Pol II	RNA Pol III
--in nucleolus --for small rRNA synthesis	--requires transcription factors to bind to its promoters (<i>opens up chromatin to less organized DNA for transcription ; remember that histones reduce the availability of DNA to DNA'ses making the DNA more stable and less prone to breakage; RNA Pol knocks out this property to perpetuate transcription</i>) --synthesizes all mRNA does not bind to promoters, but around transcription start site --reaches ahead to locate the transcription start site --in nucleus	--in nucleolus and cytosol --for tiny rRNA and tRNA synthesis --promoters for III are downstream from the transcription start site and within the region being transcribed --reaches behind to locate the transcription start site

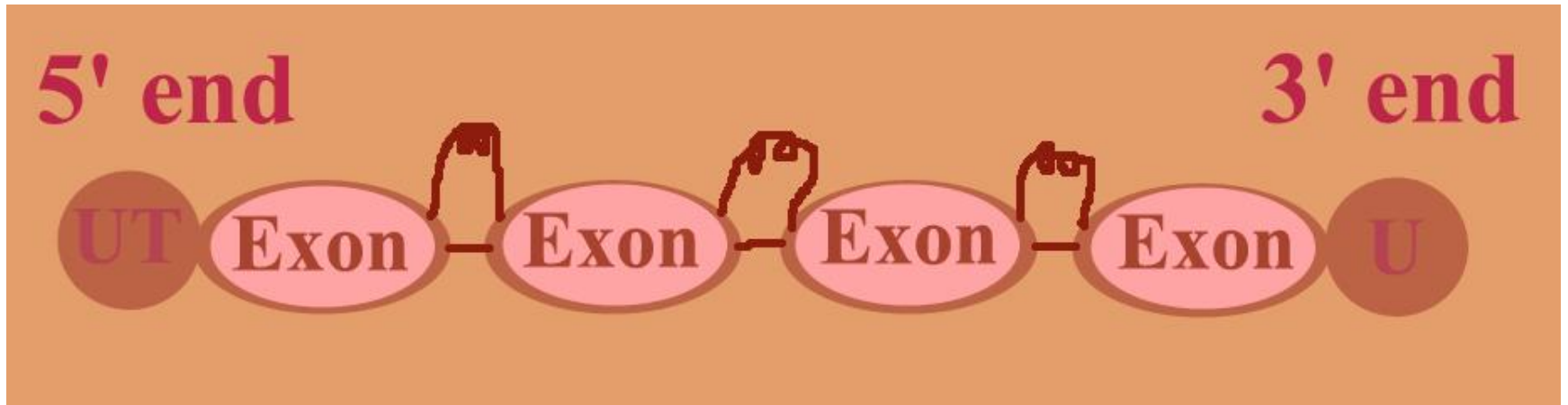
- This strand that is opened up for transcription is called the **template strand**, the **noncoding strand** or the **nonsense strand**. It is so called because the RNA that will be synthesized is complimentary to this DNA strand.
- The NONIDENTICAL strand of DNA is complimentary to the template strand and is identical to the transcript *non obstante* T to U variations and is called the **coding or SENSE strand**.
- As opposed to replication, RNA Pol's require no primer or primase to initiate transcription.



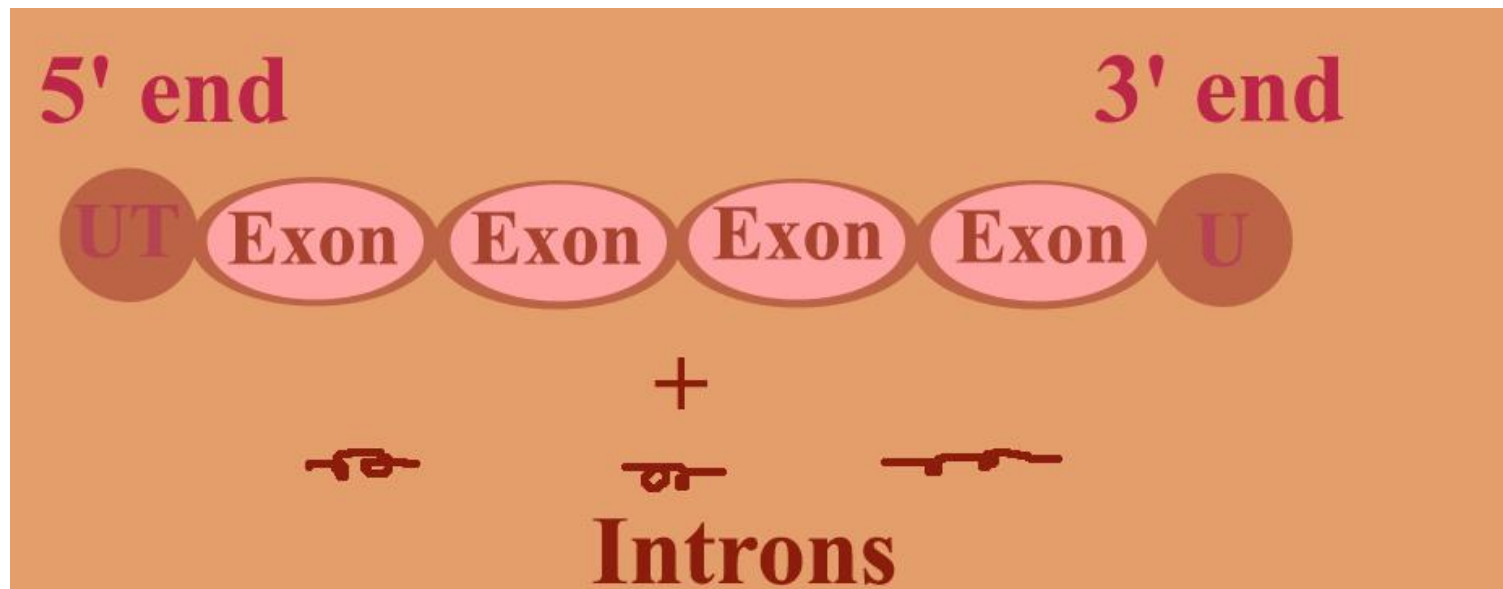
A Simplified Version of Transcription Followed by Post-Transcriptional Modification



- The 5' region of the immature transcript contains the untranslated leader (UT) that contains NO amino acid information.
- The 3' region contains the untranslated region (U) that, likewise, codes for no amino acid information (the UT and U regions MAY provide the structure that the ribosome uses to bind the mRNA - do NOT confuse UT and U regions with the snurps: See Below).
- The bottom line in this process is to remove the non-coding introns and splice the coding exons together to bring them into closer proximity for ribosomal-reading.



- The first step in post-transcriptional modification is to isolate the introns. Once they are isolated, they are spliced out.



- This process is catalyzed by the RNA, itself. When RNA acts as an enzyme, it is called a ribozyme. This is one of several examples/exceptions of enzymes that do not fit the traditional definition of enzymes, i.e., they must be proteins; RNA breaks this "rule". Once the exons are spliced together, the introns are removed (30-50% of the primary transcript).

5' end

3' end

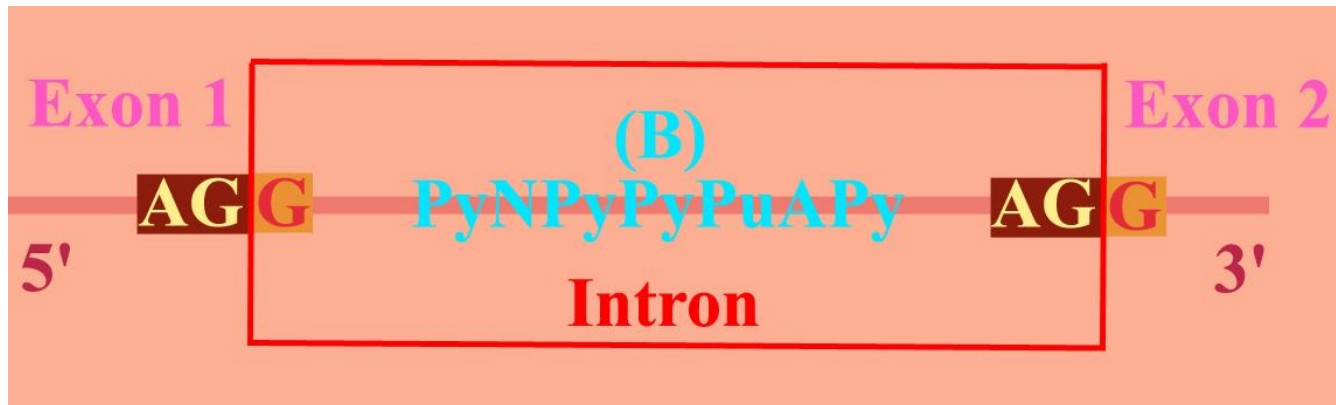


"grown up" RNA

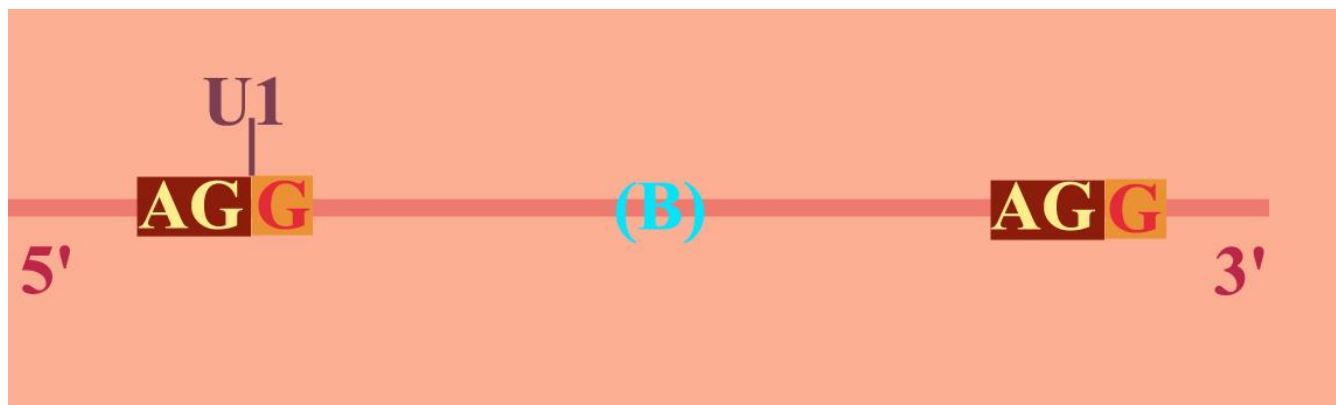
- The third processing, the addition of the poly-A tail, is directed by poly-A polymerase - no DNA template is required for this step.
- The mature RNA is now readable.

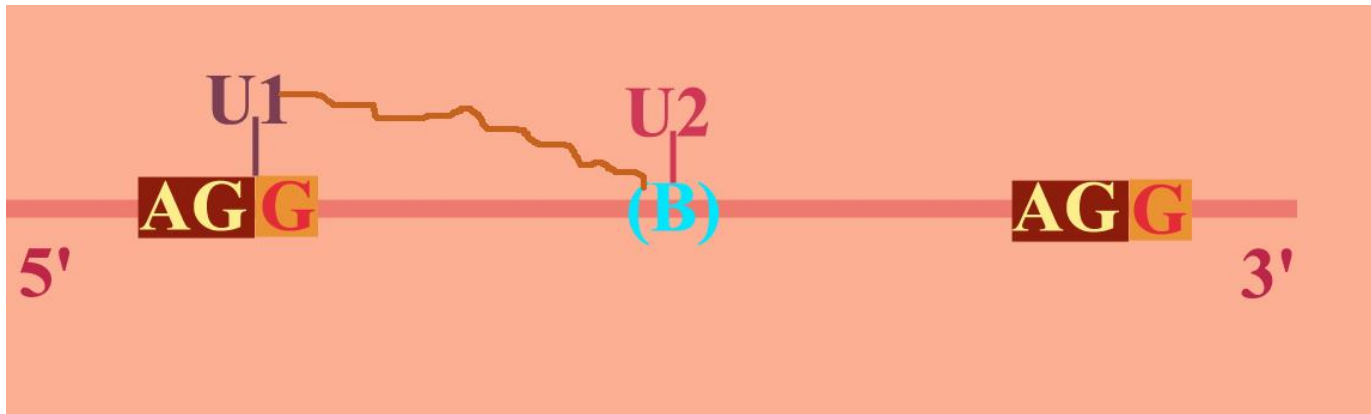
How, then, does all of this splicing occur?

- Splicing requires the primary transcript (also known as hnRNA: heterogeneous nuclear RNA), snurps (U1, U2, U4, U5, U6) and an unknown amount of proteins (perhaps they are used as scaffolding, for binding or for both).
- When all of these are present, they make up a **spliceosome**.

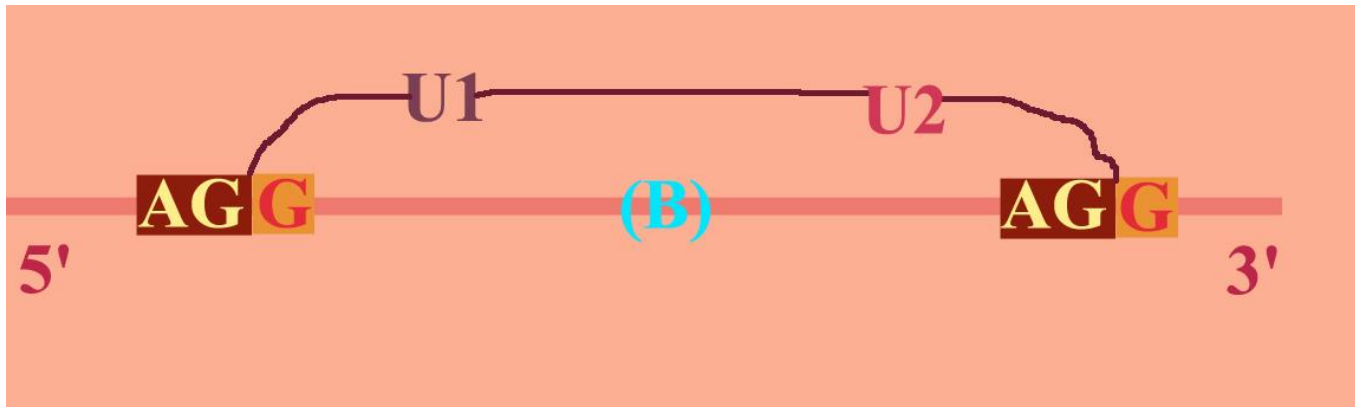


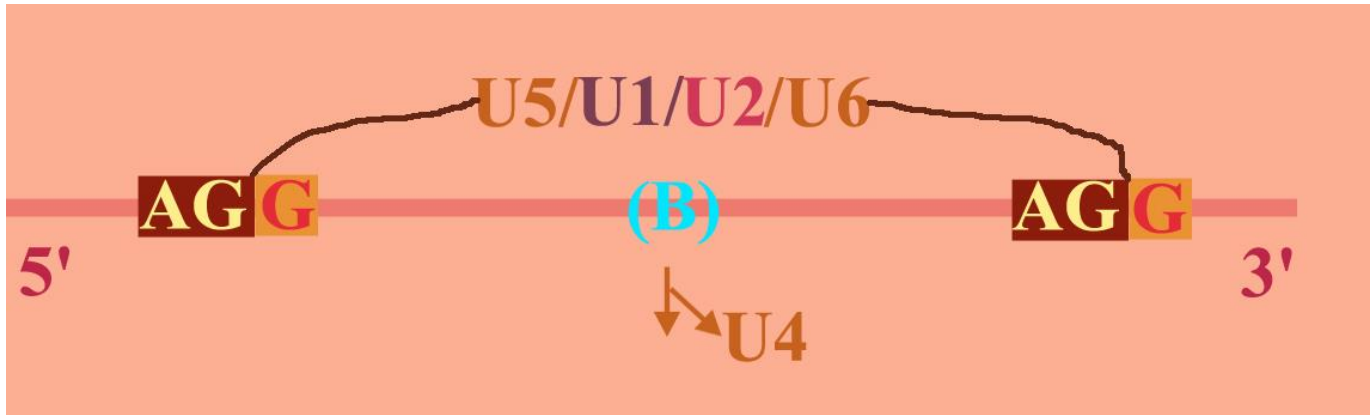
- The ends of exons seem to be AGG from 5' to 3'. Inside the intron is a sequence of pyrimidines (Py) and purines (Pu) with any nucleotide (N) and adenosine. The sequence is PyNPyPyPuAPy, again, from 5' to 3'. This sequence is called a consensus sequence, a conserved sequence that seems to act as the identifying feature of the splice site. **Sometimes this consensus sequence is called a branch site (B).**
- The primary transcript reacts with U1. U1 binds to the 5' end of the intron.



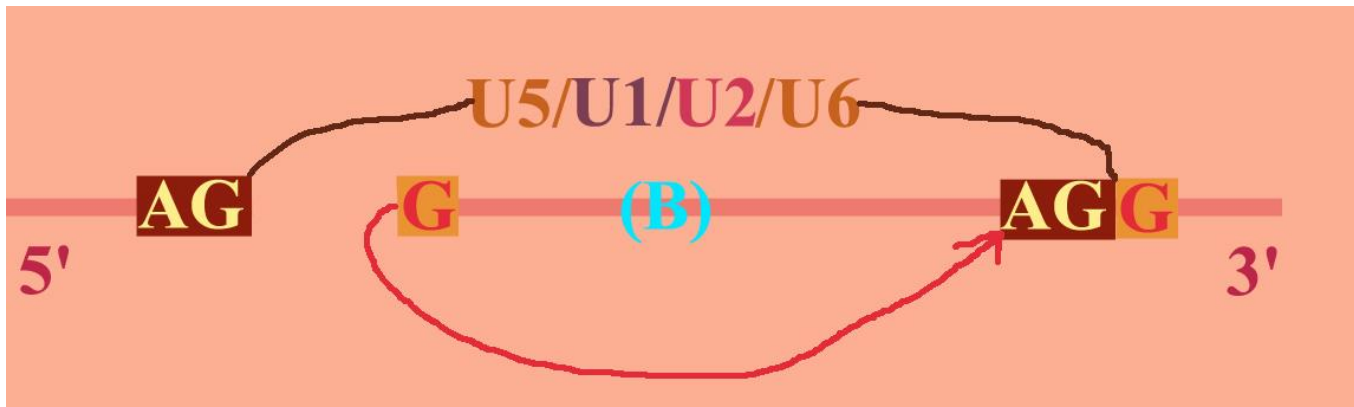


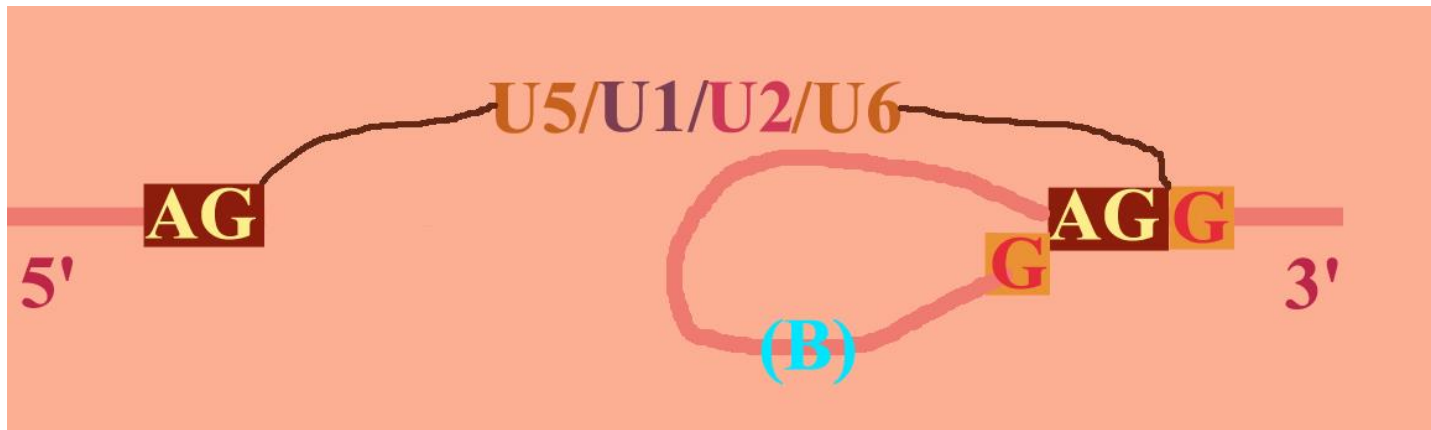
- U2 binds to the branch site (B) and also directs U1 to bind to B, as well as to the 5' end of the intron.
- U1 and U2 rearrange so that U1-U2 spans the intron, i.e., the U2 end binds to the 3' end of the intron.



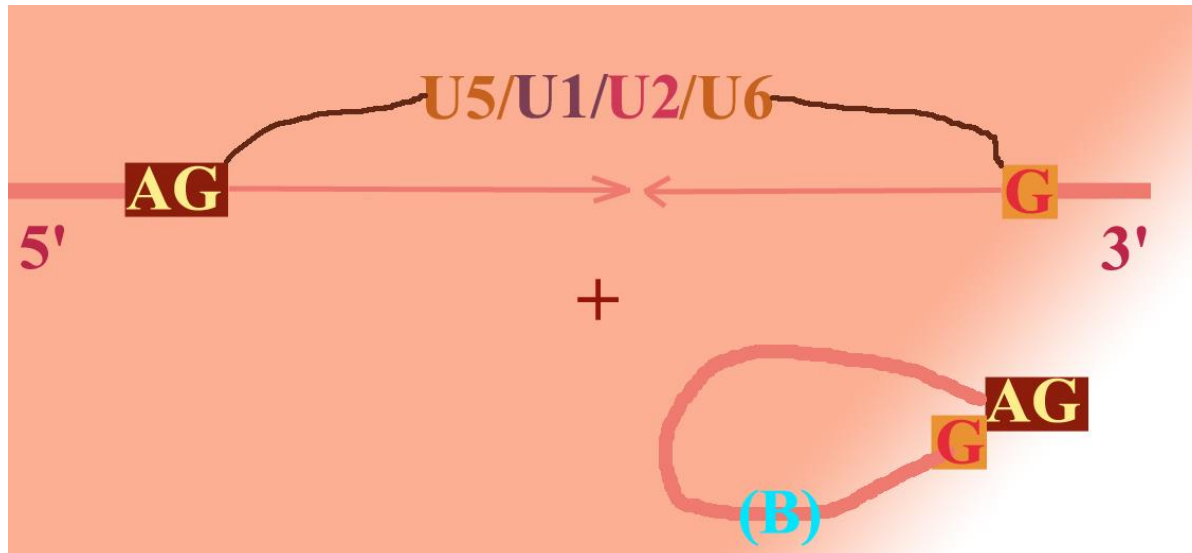


- U4, U5 and U6 then bind together to form a catalytically active complex that causes U4 to be released with the formation of a U1· U2· U5· U6 complex that causes the formation of a "lariat" between the 5' region of the partially "clipped" transcript and the 3' region of the transcript.

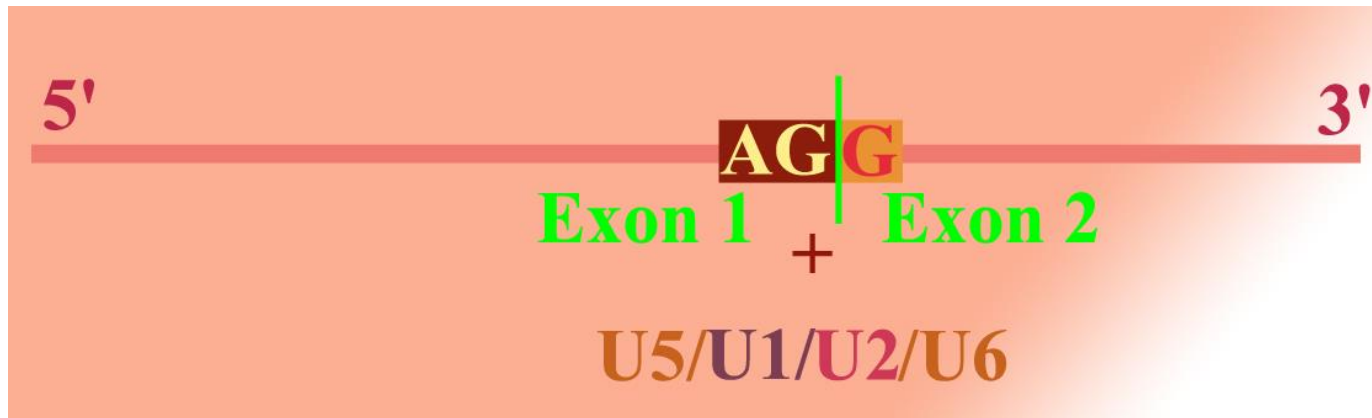




- Once the lariat (derived from the intron) is removed, the snurps catalyze the rearrangement of the exons, butting them up against one another.
- As the ends of the exons are annealed, the snurps are released, the lariat is metabolized and the nucleotides are recycled.
- This process is called the "lariat hypothesis".
- Another form of the "lariat splicing mechanism" has to do with a very unusual 5' to 2' lariat formation.



FINALLY!



- The termination of transcription is brought about by polyadenylation (poly-A). This is a sequence of A's about 200 bp's long.
- While there is no DNA template for poly A, there is a signal sequence in mRNA that "triggers" poly A addition. An RNA endonuclease cleaves the poly-A recognition site and poly-A is added.
- The AAUAAA sequence is a required base sequence for enzymatic activity.
- At this point, the hnRNA is matured and is ready to be "read" by ribosomes.
- mRNA processing also includes capping and methylating: these processes may protect RNA from RNA'ses OR may be recognition sites for ribosomes OR may be a transport sequence to get the mRNA out of the nucleus.

- Although gene regulation is complex and not fully understood, there are some initial regulations of transcription about which we do have some knowledge.
- Transcription may be positively regulated
 - 1) hormonally at the level of the DNA; with RNA Pol and with proteins necessary for Pol interaction. This probably does not occur in man;
 - 2) hormonally at the level of the DNA that causes conformational changes of DNA so RNA Pol may bind to it;
 - 3) hormonally where the hormone binds to the transcriptional factor that has to bind to DNA site before RNA Pol may bind;
 - 4) cAMP is even involved: it increases tyrosine aminotransferase, PEPCK and prolactin syntheses.

- Conversely, transcription may be negatively regulated
 - 1) hormonally where the hormone acts as an inducer that turns off repressors and/or
 - 2) hormonally where the hormone binds to the DNA to cause a conformational change of chromatin to make the DNA susceptible to RNA Pol.