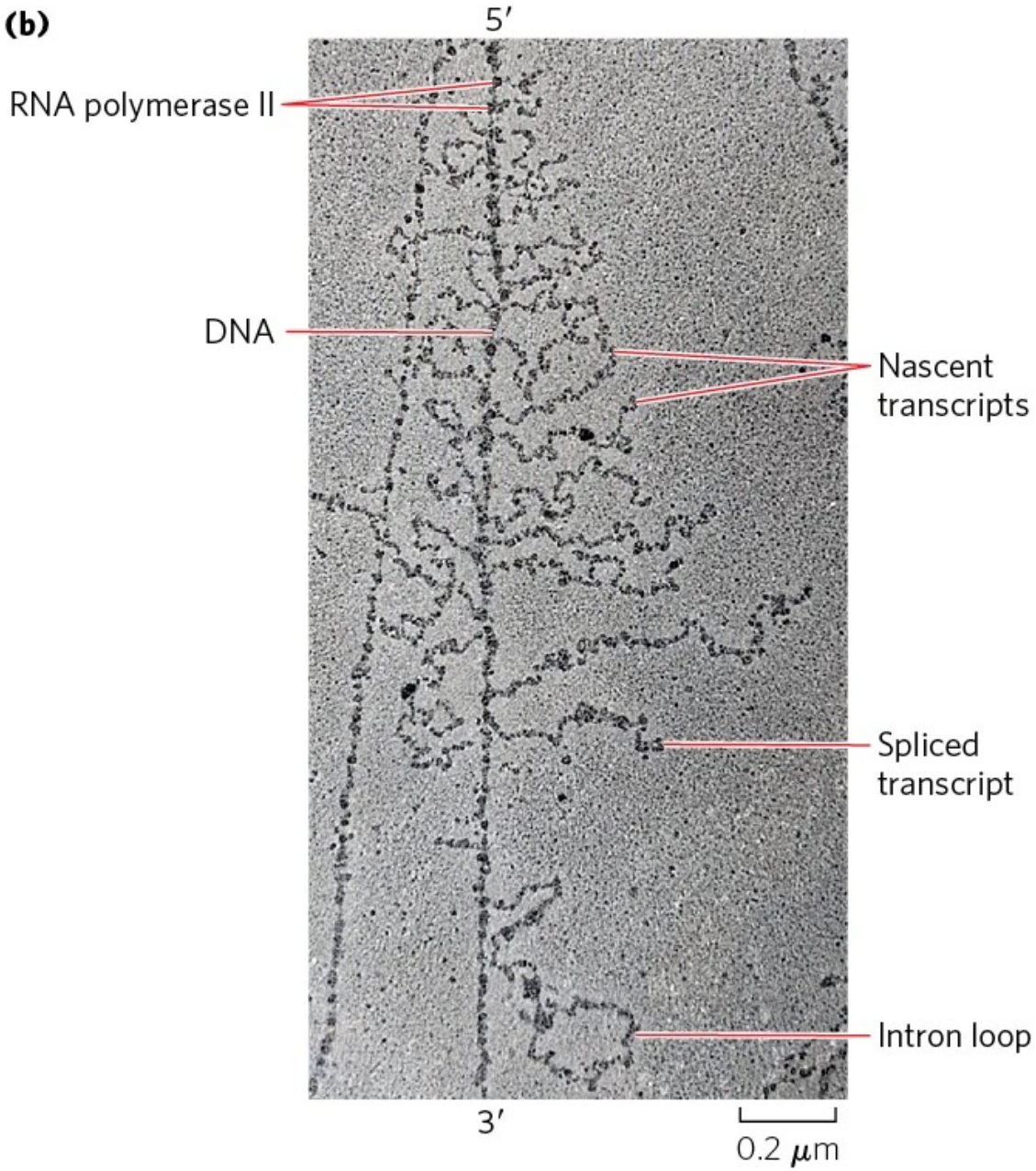
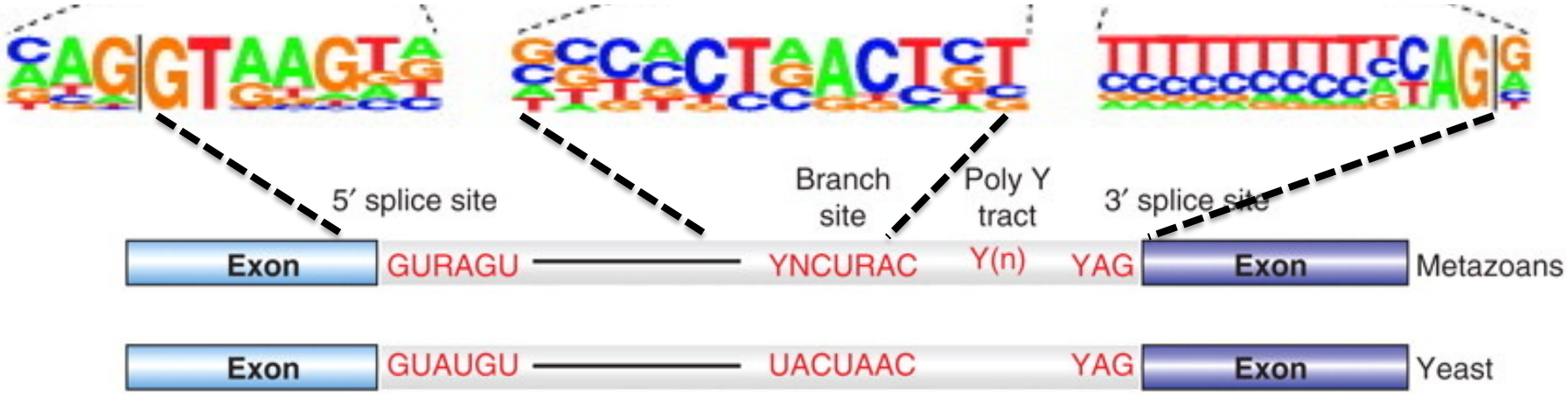
Splicing happens co-transcriptionally



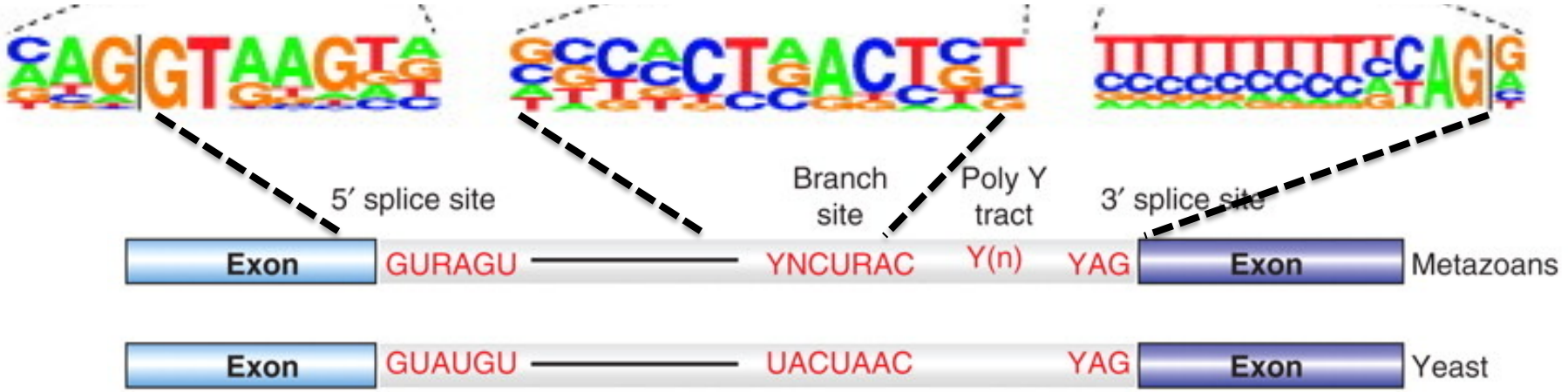
Exon/ introns junctions and consensus sequences of nuclear introns

"Sequence logo" representation of consensus sequences



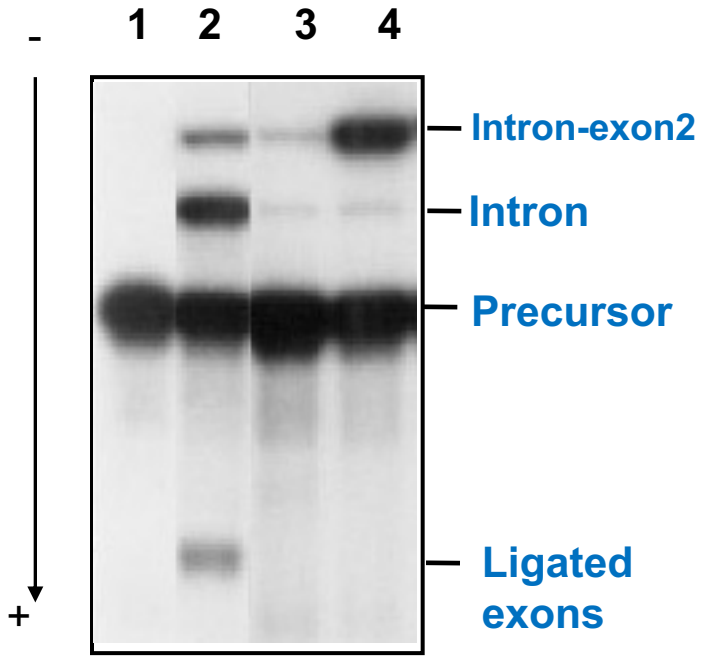
Exon/ introns junctions and consensus sequences of nuclear introns

"Sequence logo" representation of consensus sequences

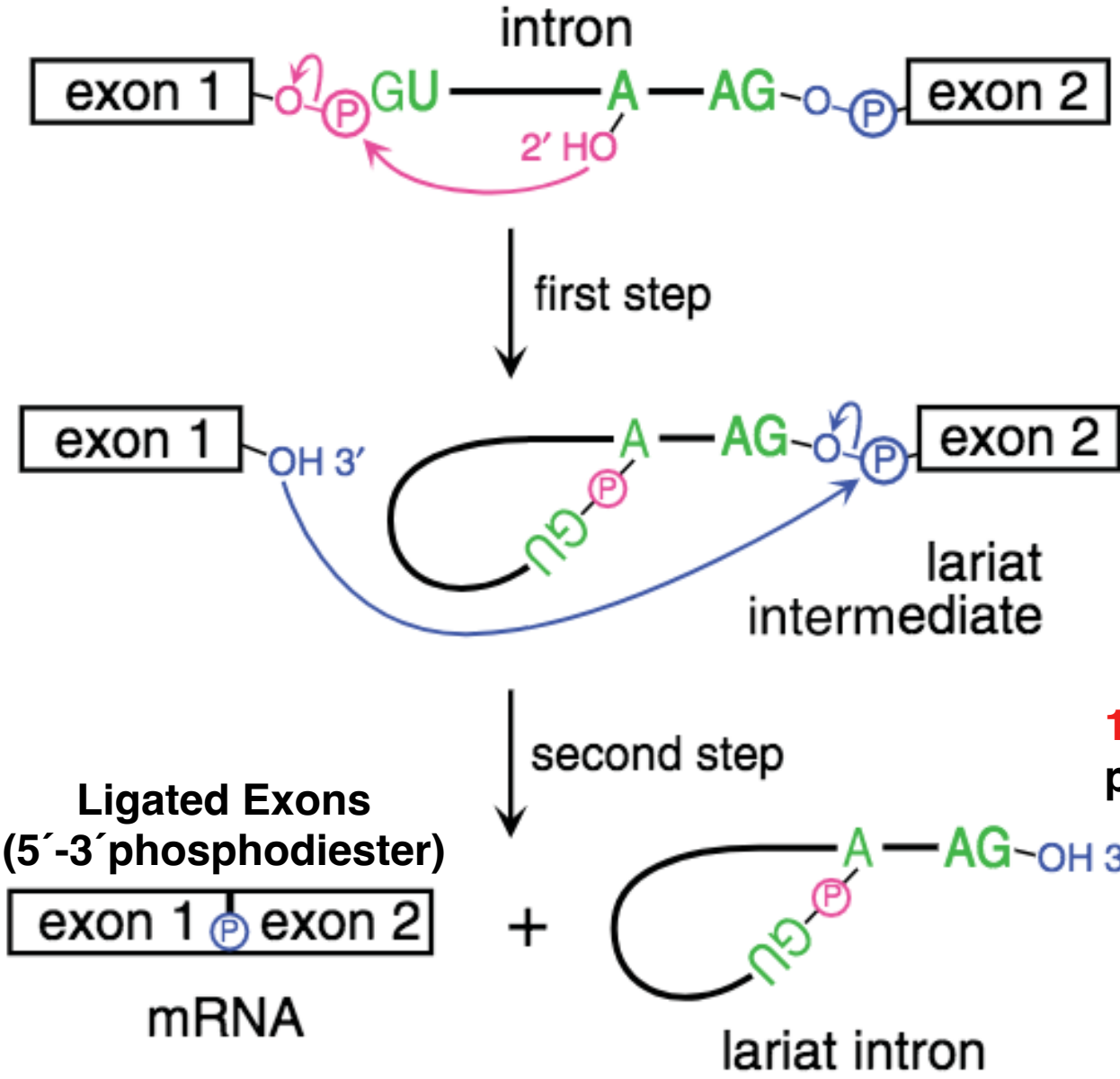


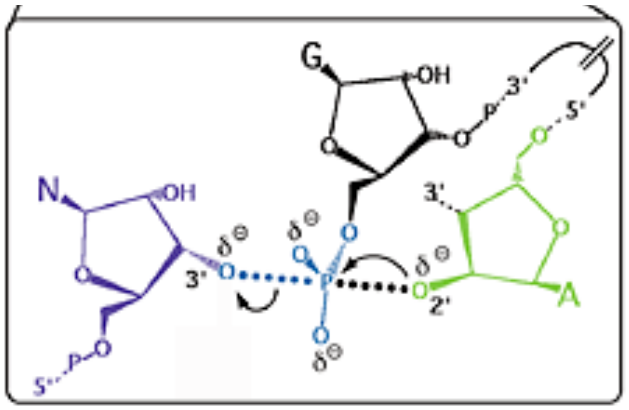
In vitro pre-mRNA splicing reaction:

- **synthesis of an RNA substrate containing exon1, intron, exon2 and splicing signals**
- **Incubation with nuclear extract with $MgCl_2$ and ATP**
- **Fractionation of RNA products by denaturing gel electrophoresis**
- **Visualization of products by autoradiography**



1 - Extract, $MgCl_2$, no ATP ; 2- Extract, ATP and $MgCl_2$; 3 - RNA substrate with a 5´ splice site mutation 4 - RNA substrate with a 3´ splice site mutation





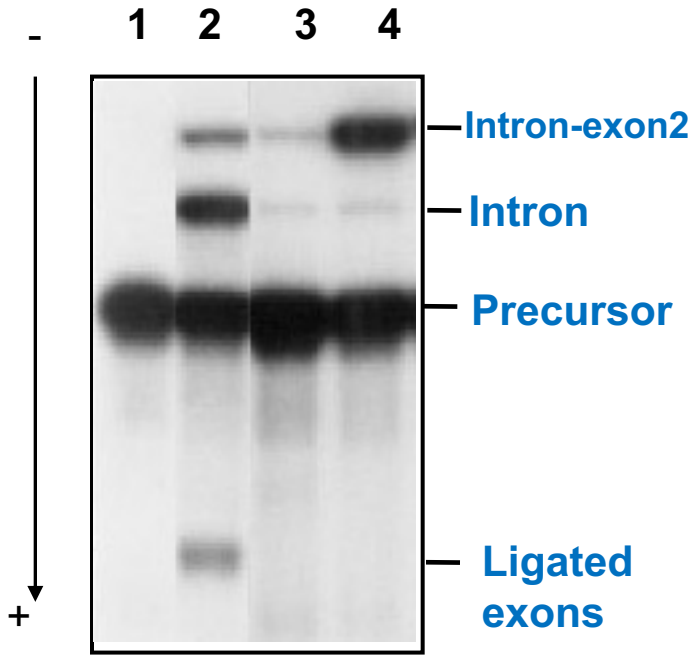
Each chemical step is a *transesterification*

1st step: formation of a 2′-5′ phosphodiester link

2nd step: formation of a 5′-3′phosphodiester link

The chemical steps per se do not require energy
The ATP is required for steps other than the chemical steps

Why do the intron-exon2 and intron migrate slower than the unspliced precursor?



A: They are longer than the precursor

B: They are not linear molecules and thus do not migrate according to their mass in this type of gel

C: They become polyadenylated which adds some extra nucleotides compared to the unspliced precursor

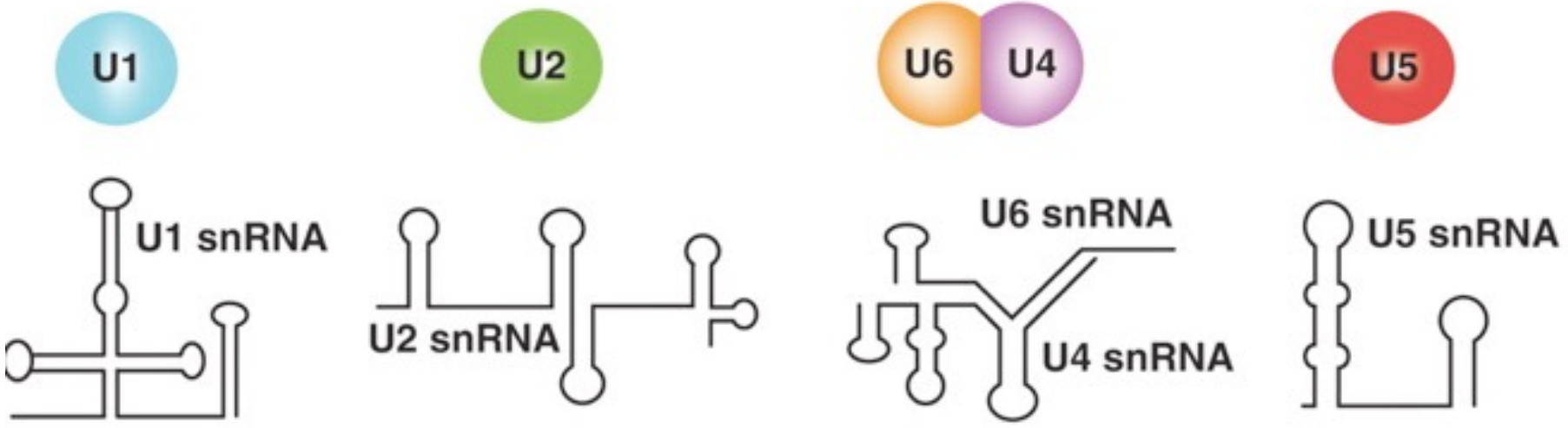
D: They are bound by splicing factors which slows down their migration in the gel compared to the precursor

Pre-mRNA splicing is catalyzed by a complex macromolecular machine: the Spliceosome

~145 spliceosomal proteins

Some of them assembled into complexes with snRNAs: U1, U2, U4, U5, U6 snRNPs

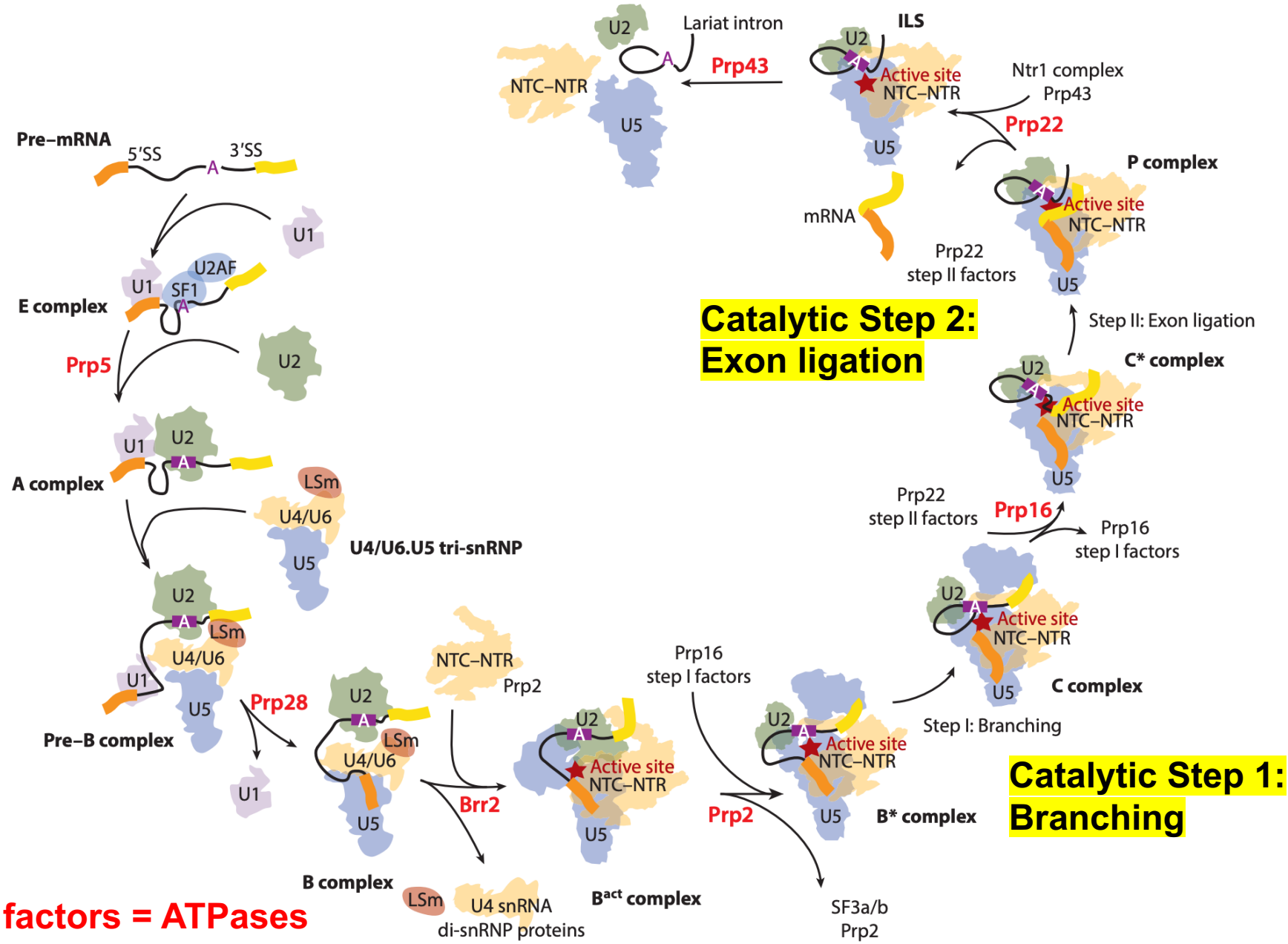
U4, U5, U6 snRNPs can associate with each other to form di/tri-snRNPs U4/U6 and U4/U5/U6



Each snRNP contains:

- **1 small nuclear RNA (snRNA: U1, U2, U4, U5 or U6)**
- **Several proteins associated with the snRNA**

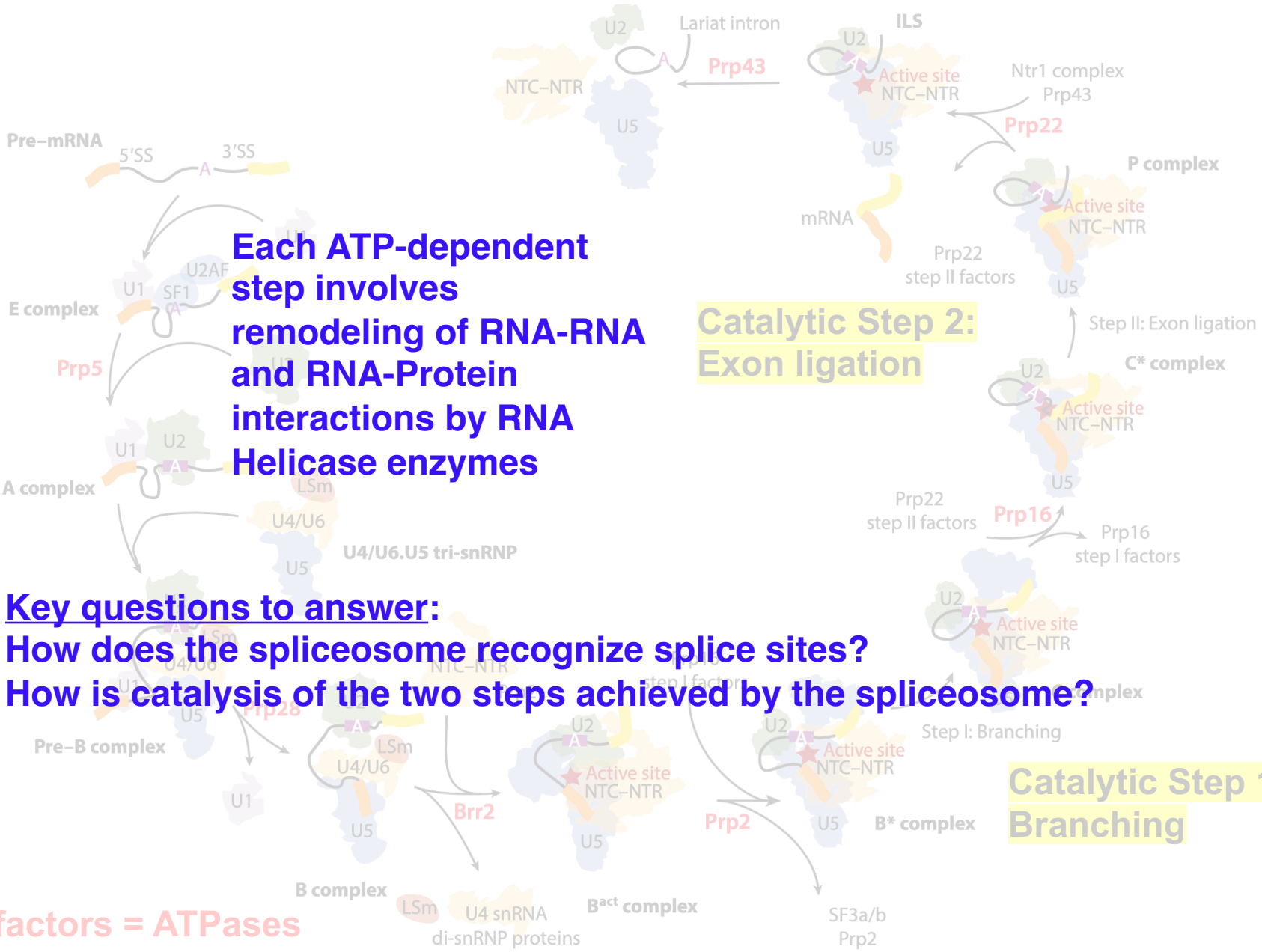
Simplified model of Spliceosome Assembly and Catalysis of Splicing



Red factors = ATPases

Wilkinson, Charenton, Nagai, Annu. Rev. Biochem., 2020.

Simplified model of Spliceosome Assembly and Catalysis of Splicing



Each ATP-dependent step involves remodeling of RNA-RNA and RNA-Protein interactions by RNA Helicase enzymes

Key questions to answer:

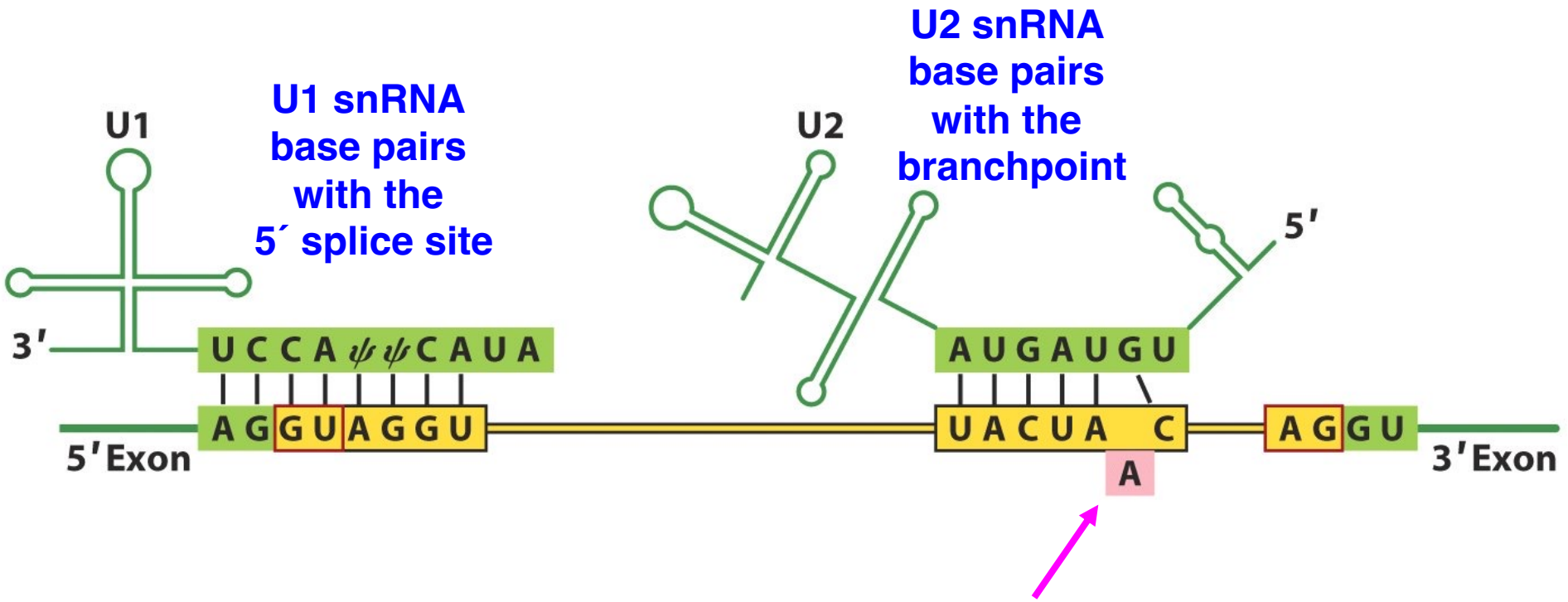
How does the spliceosome recognize splice sites?

How is catalysis of the two steps achieved by the spliceosome?

Red factors = ATPases

Wilkinson, Charenton, Nagai, Annu. Rev. Biochem., 2020.

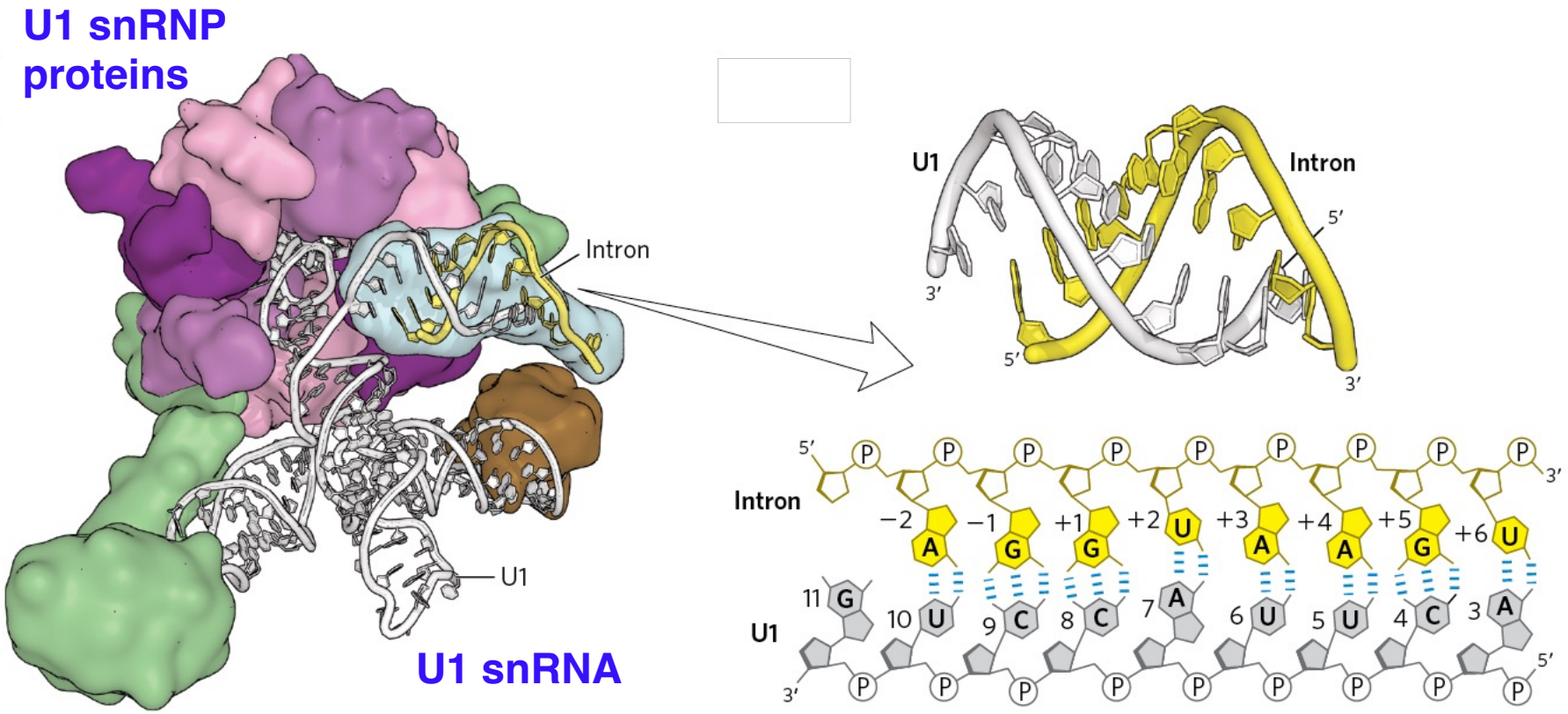
Recognition of the 5' splice site and branch point A by base-pairing with the snRNAs within snRNPs



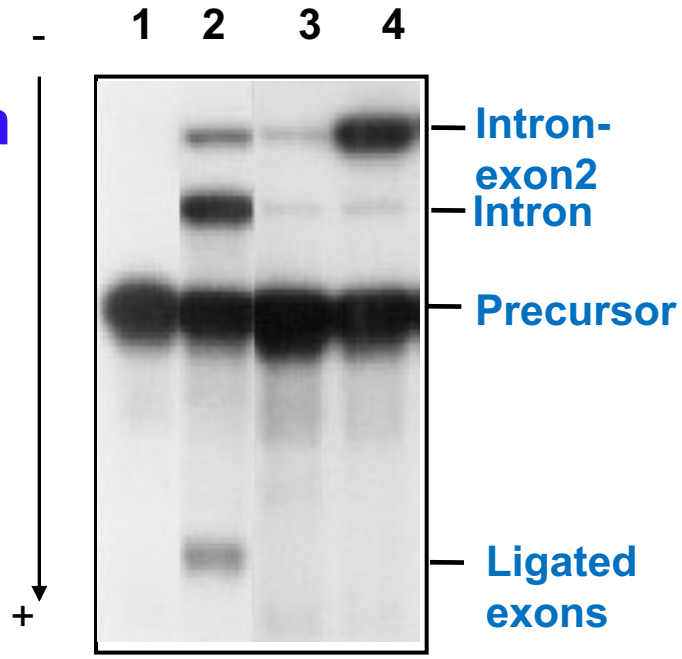
- the branchpoint nucleoside must be "bulged" (unpaired and extruded from the duplex) to act as the nucleophile

- this base pairing occurs in the context of RNA-Protein complexes; the RNAs are not "naked" as shown here)

Recognition of the 5´ splice site of the intron by base pairing with the U1 snRNA in the U1 snRNP



Why is splicing inhibited In lane #3 with a mutation of the 5´ splice site?



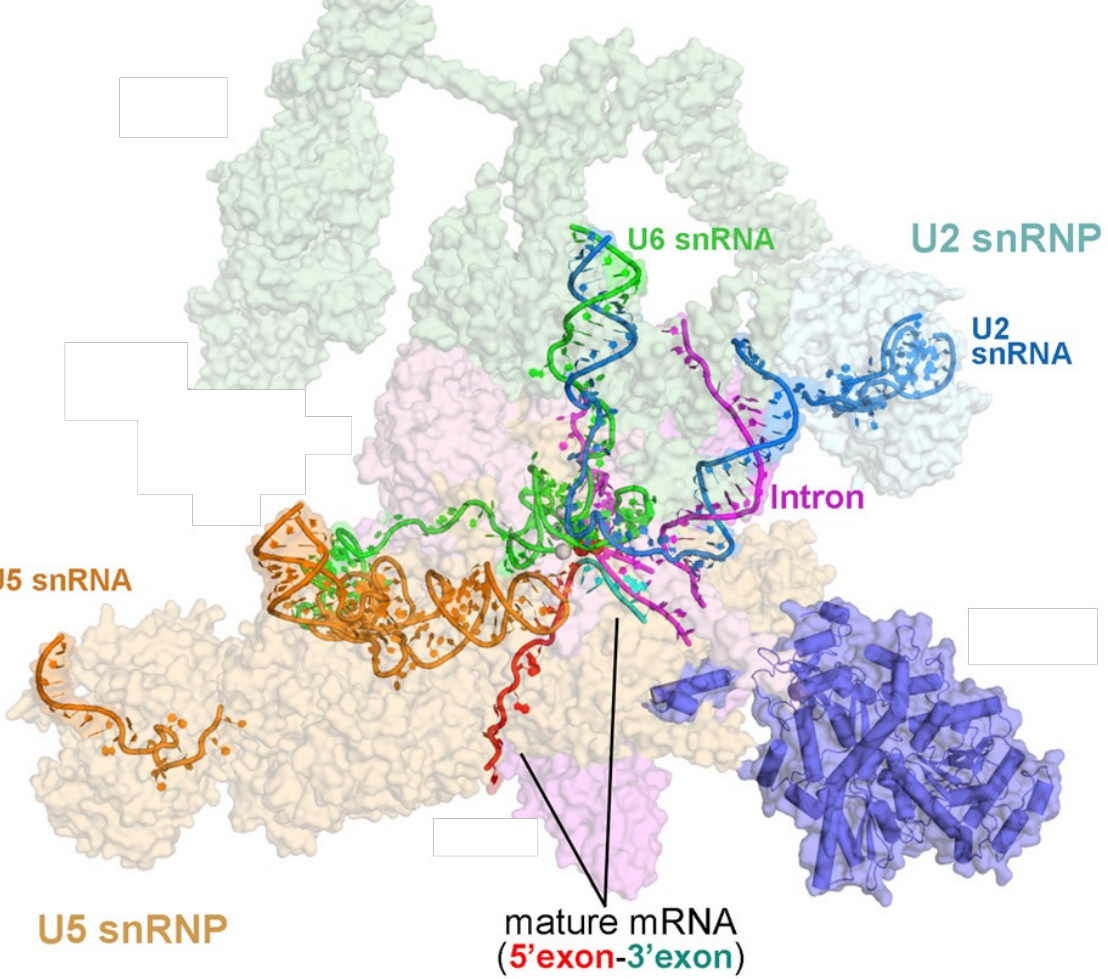
A: It inhibits the first catalytic step

B: It inhibits recognition of the 5´ splice site by the U1 snRNA and the early step of spliceosome assembly

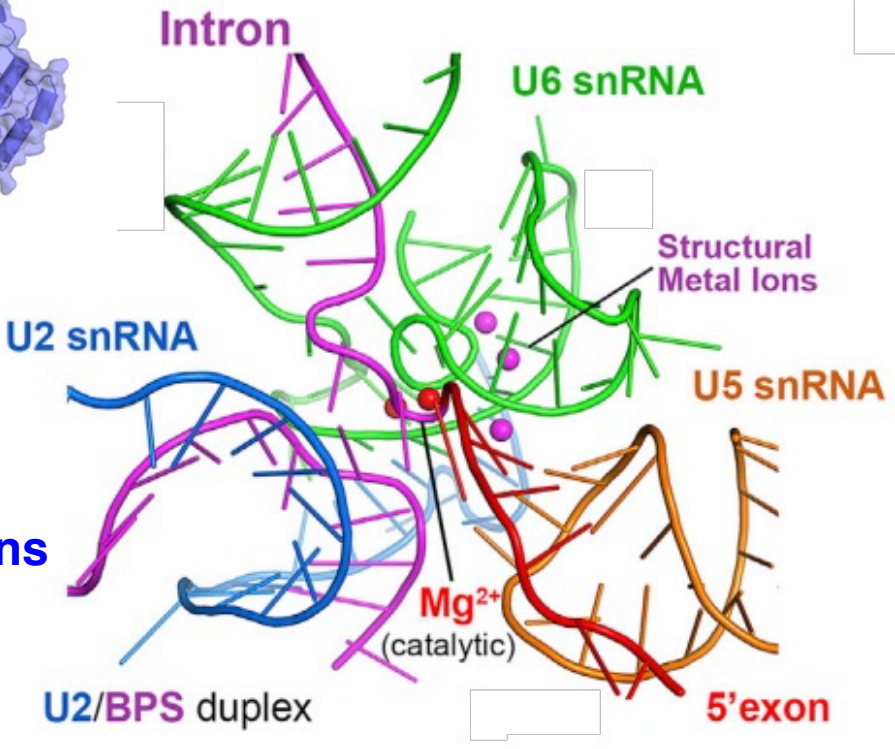
C: It inhibits recognition of the branchpoint by the U2 snRNA and the early step of spliceosome assembly

D: It inhibits the second catalytic step

Spliceosomes structures reveal mechanisms for catalysis



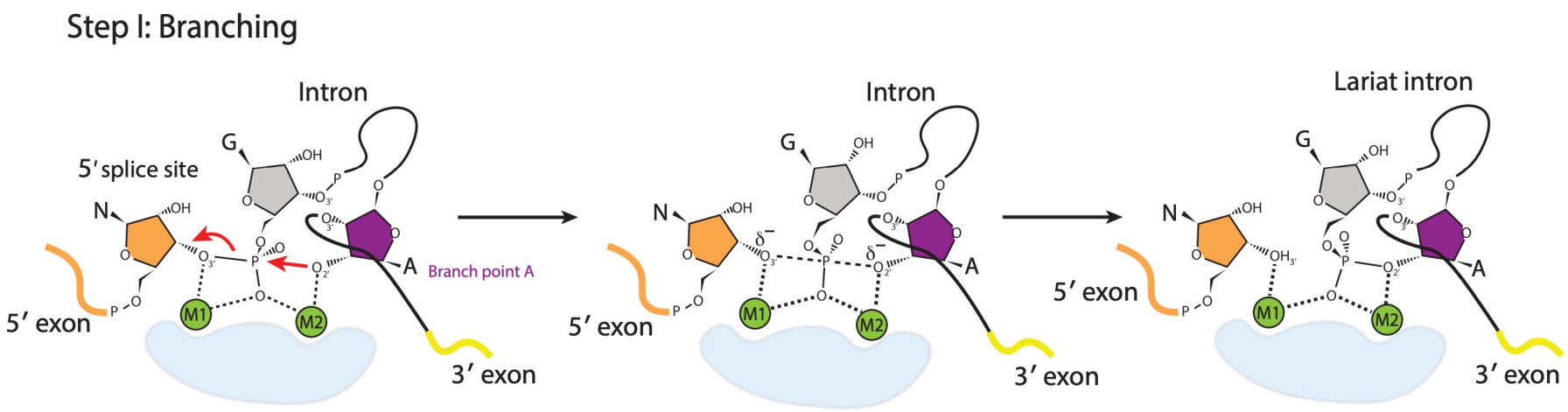
Bai et al. Cell 2017



The active site of the spliceosome is primarily made of U6 snRNA and metal ions close to the splice sites and branchpoint

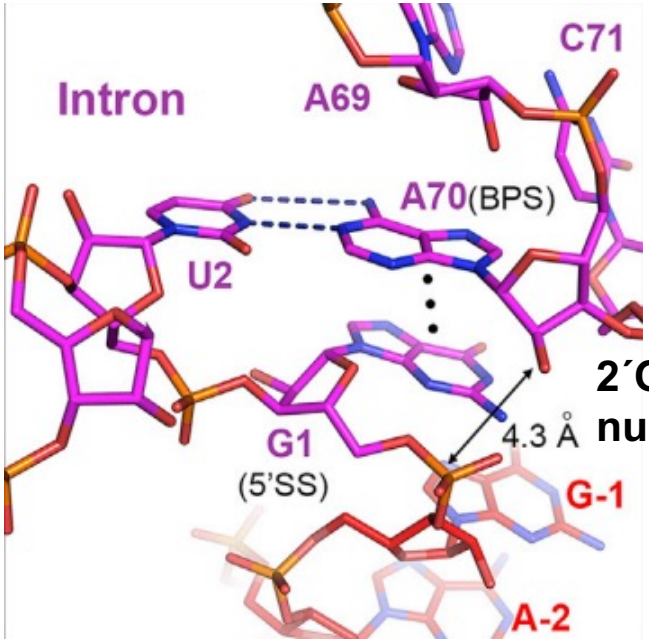
Wan et al. Cell 2019

Mechanism for Catalytic Step I: Branching



- **2-metal-ion mechanism (metal ions: M1 and M2)**
- **The branchpoint A 2’ O acts as the nucleophile and attacks the phosphorus of the 5’ splice site**
- **The pentavalent transition state is stabilized by the two metal ions**
- **M1: activates the 2’ OH of the BP A; M2: stabilizes the leaving group**
- **Result:**
  - **free 5’ exon**
  - **lariat-3’ exon intermediate (phosphorus atom of first intron nucleotide is linked to the 2’ O of the branchpoint A)**

Interaction between the branchpoint nucleophile A and the 5´ splice site GU sequence aligns substrates for 1st catalytic step

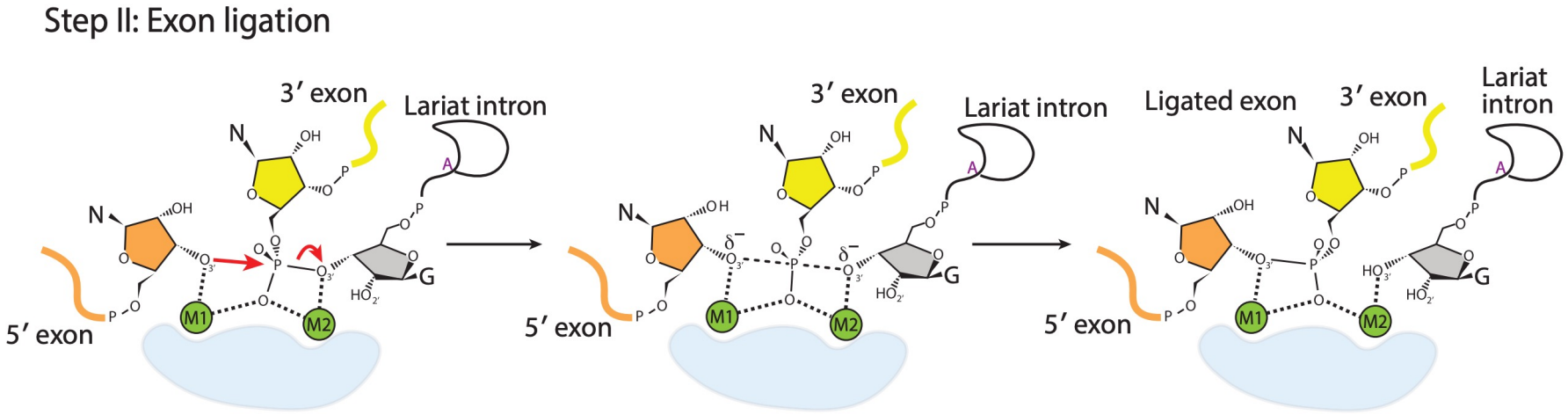


Branch point A

2´OH nucleophile

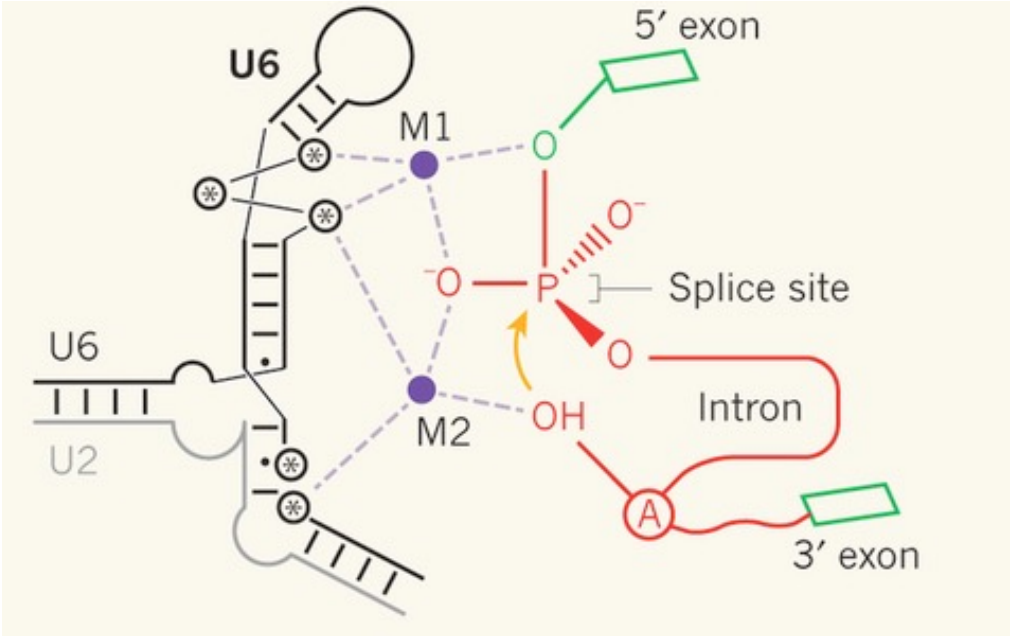
Wan et al. Cell 2019

Mechanism for Catalytic Step II: Exon ligation



- **2-metal-ion mechanism <u>in the same active site</u> (metal ions: M1 and M2)**
- **5’ exon remained in the active site from step I, but the branchpoint A moved away and the 3’ splice site has entered the active site**
- **3’ OH group acts as the nucleophile and attacks the phosphorus atom of the 3’ splice site**
- **The pentavalent transition state is stabilized by the two metal ions**
- **M1: activates the 3’ OH of the 5’ exon; M2: stabilizes the leaving group**
- **Result:**
  - **Ligated 5’ and 3’ exons**
  - **Lariat intron**

What binds the catalytic metal ions?



A: They are bound by water molecules as described in some DNA/RNA polymerases

B: They are bound by Aspartate residues like in RNA/DNA polymerases

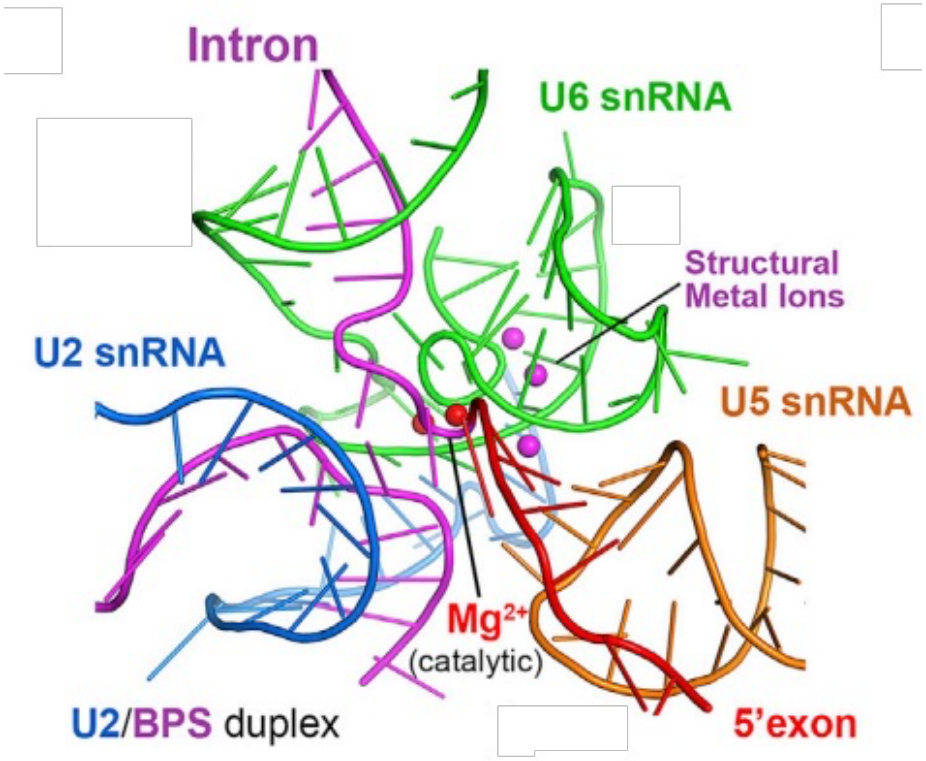
C: Divalent cations can form stacking interactions with bases of nucleic acids

D: They are bound by phosphate oxygens as those carry negative charges and can make electrostatic interactions.

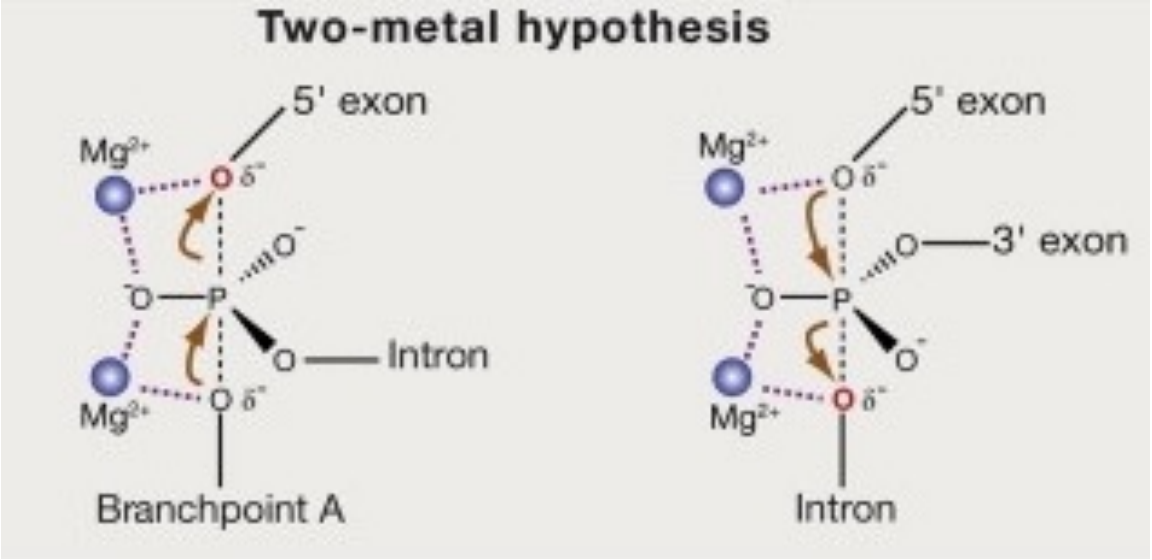
The spliceosome is a ribozyme!

- Catalysis is carried out through a two metal ion mechanism
- Both catalytic metal ions are coordinated by RNA

The active site:

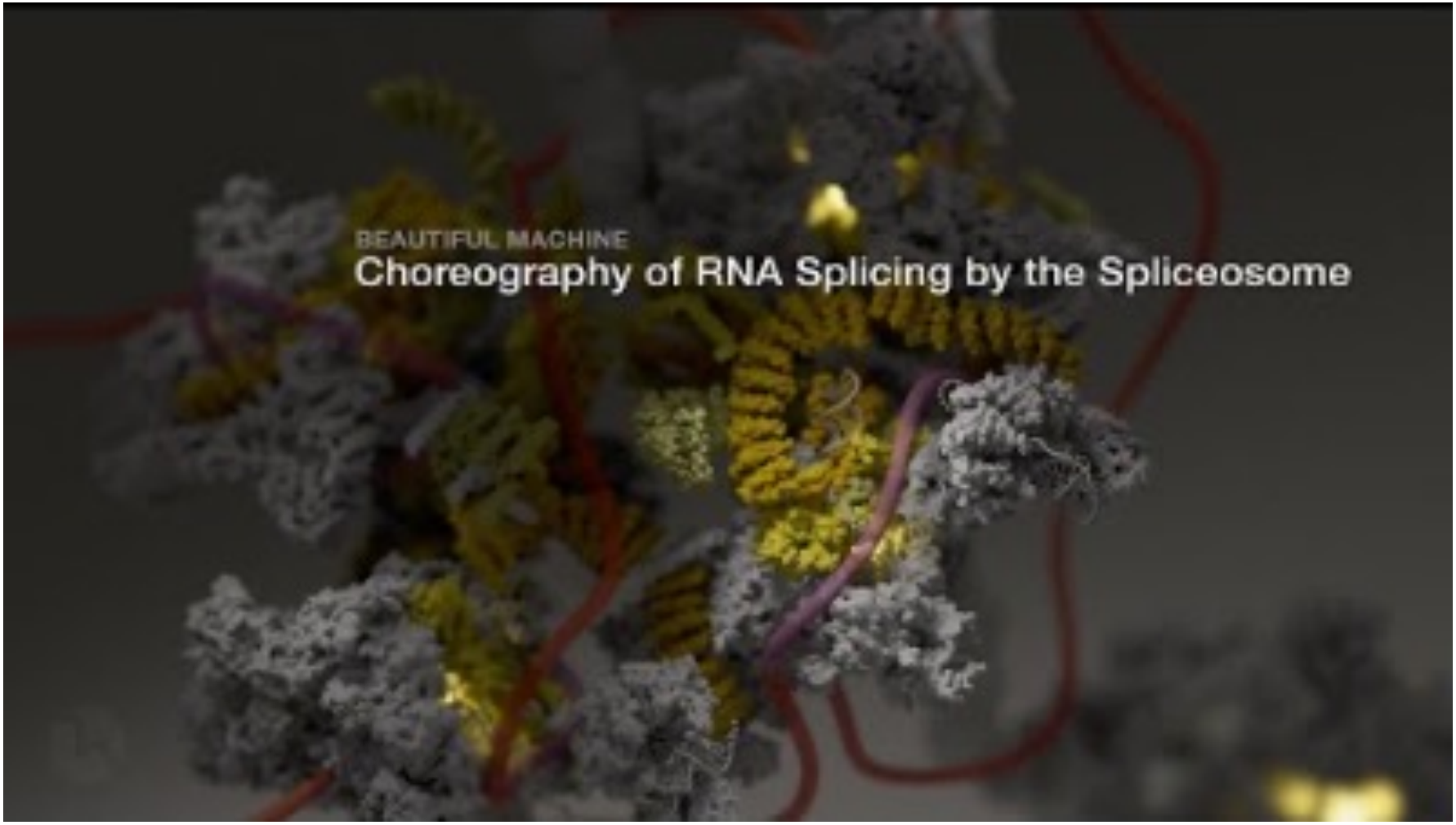


History of science – this was predicted in 1993 by Tom & Joan Steitz…



Madhani (Cell 2013),
from T. Steitz and J. Steitz PNAS 1993

Visualizing the process of splicing by the spliceosome



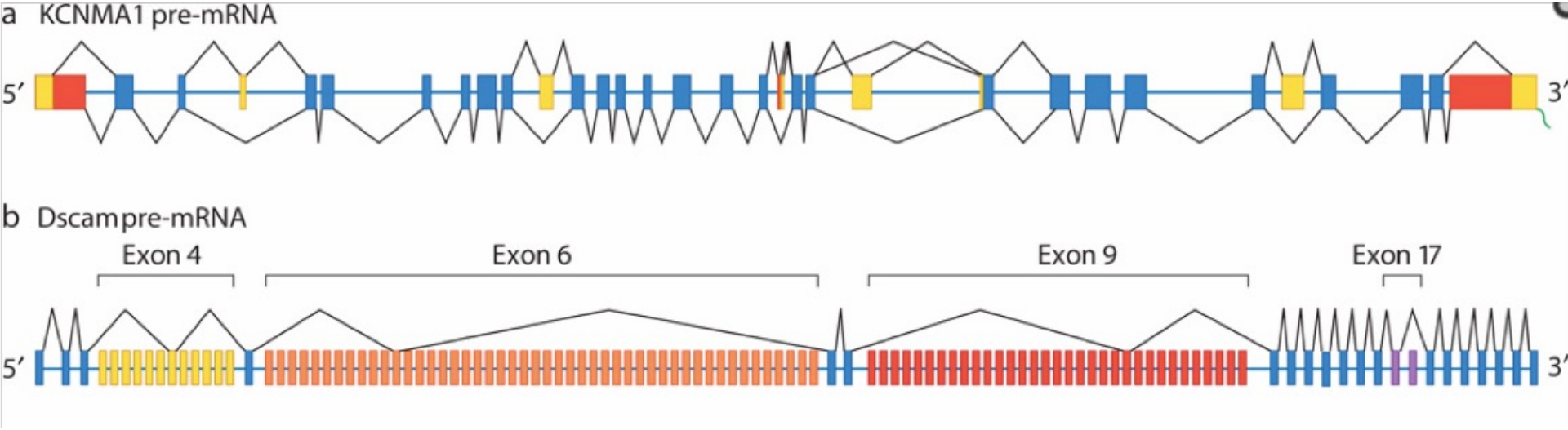
https://www.youtube.com/watch?v=w0WY0pA1cMM

Generating Diversity through alternative splicing

Two examples: KCNMA1 and DSCAM genes

Blue exons are always included in the mature mRNA
Other colors indicate alternative splice sites/ exons

KCNMA1 gene: ~500 possible mRNAs



Dscam gene, an extracellular receptor required for axon guidance in *Drosophila*

Exon 4: 12 possible exons sequences
Exon 6: 48 possible exons sequences
Exon 9: 33 possible exons sequences
Exon 17: 2 possible exons sequences

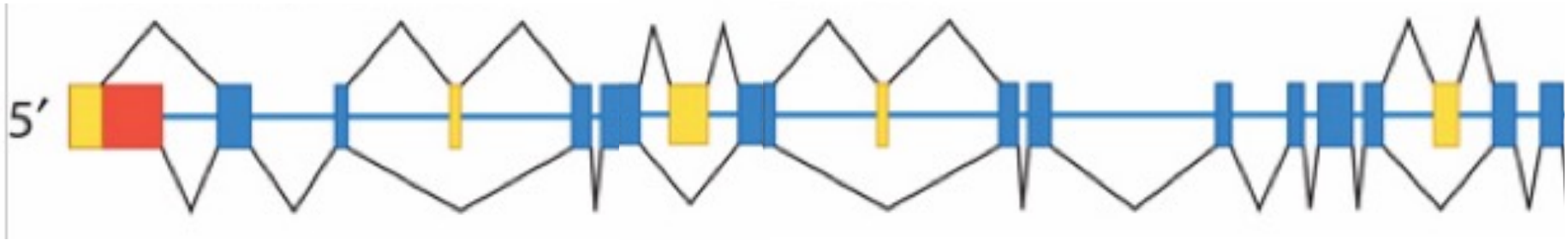
Total possible combinations =12 X 48 X 33 X 2= 38,016

Number of genes in the Drosophila genome = 20,000-50,000….

How many splice isoforms are possible from this gene? Assume each splicing event is independent from the others.

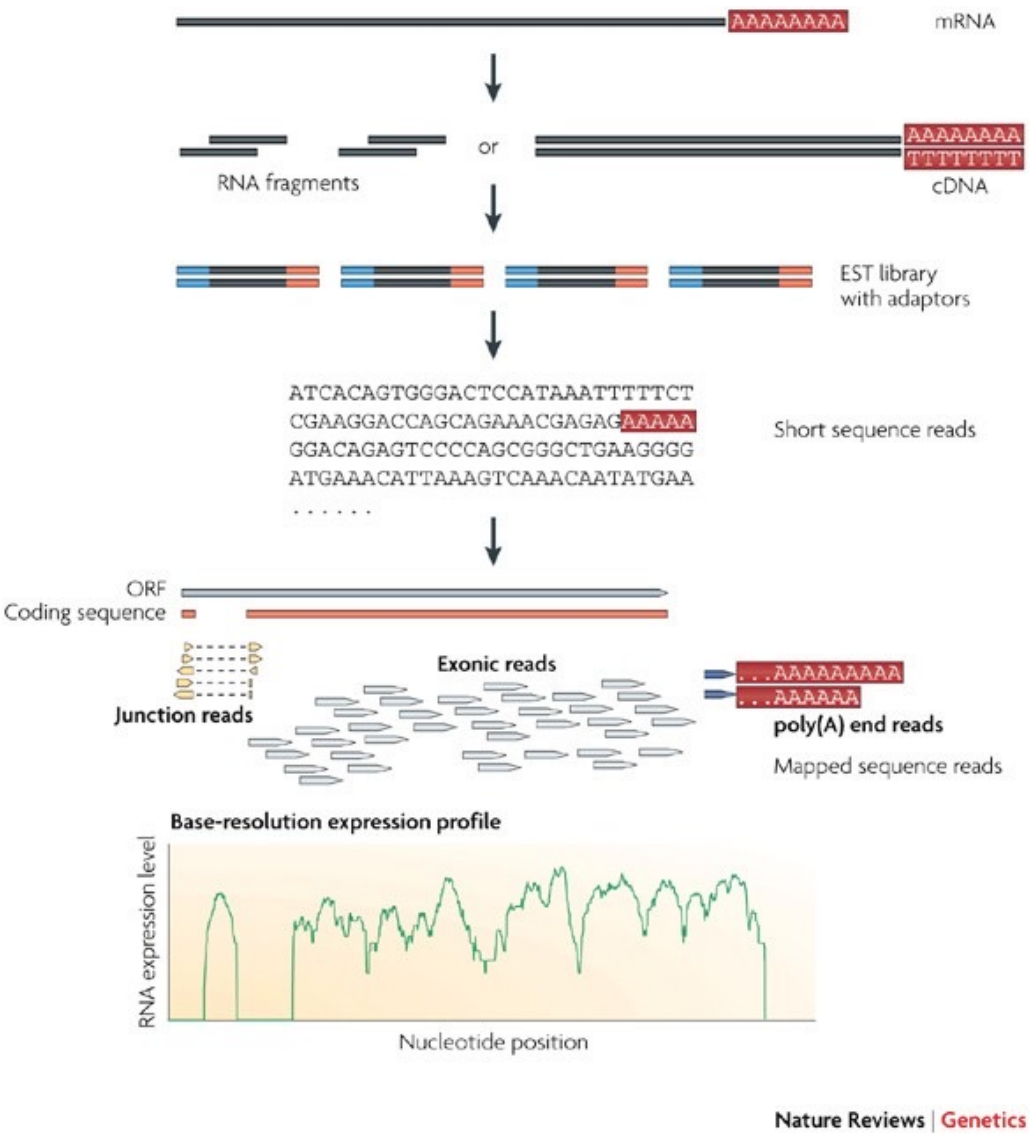
18 exons

- **2 alternative 5´ splice sites for exon1**
- **4 possible exon skipping events**



A: 10 **B: 32** **C: 18** **D: 36**

Measuring splicing using RNA sequencing



- Reports sequence of mRNAs
- Reports abundance
- Multiple techniques to do this (next-gen sequencing shown)
- **Can also use this to assess splicing – do reads contain intron/exon junctions or exon/exon?**

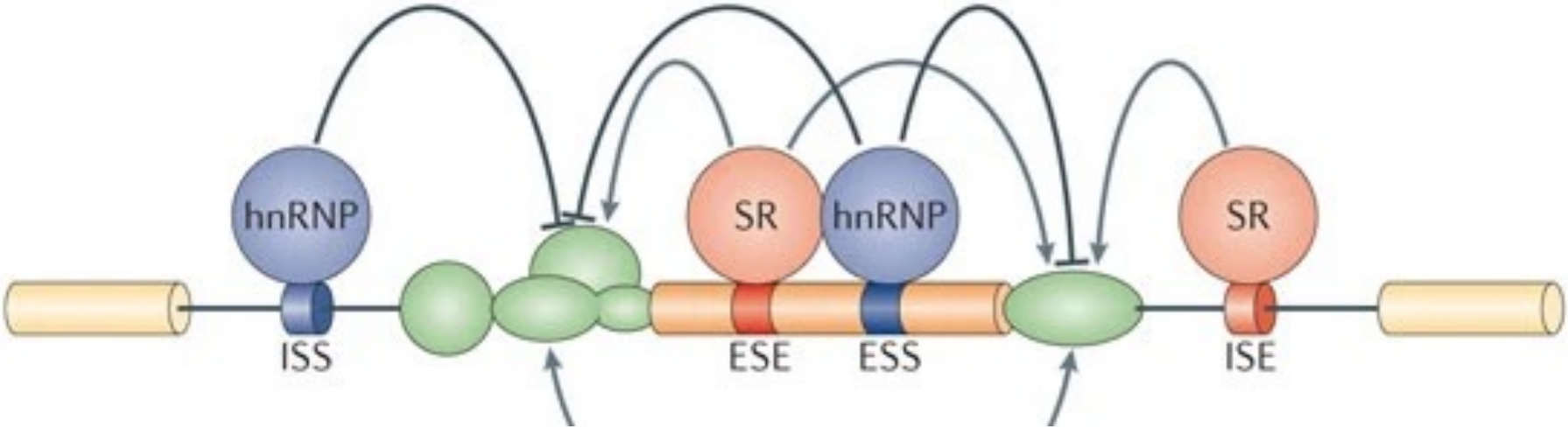
Snyder et al., Nat Rev Genetics 2009

How do cells choose which exons to splice ?

- Exons and introns possess regulatory sequences (ESS, ESE, ISS, ISE) that splicing regulatory proteins can bind

- Splicing regulatory proteins (SR proteins, hnRNP,) bind to specific enhancers or silencers RNA sequences

- Splicing regulators (SR proteins or others) promote or repress recruitment of U1 or U2 snRNPs at nearby splice sites resulting in inclusion or skipping of exons



ISS = Intronic Splicing Silencer Sequence

Spliceosome Components

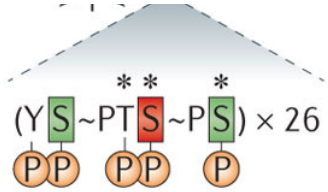
ESE/ESS = Exonic Splicing Enhancer/Silencer Sequence

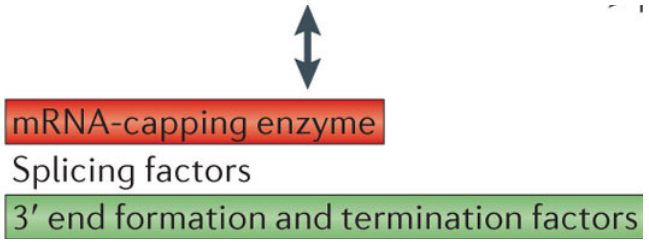
Spliceosome Components

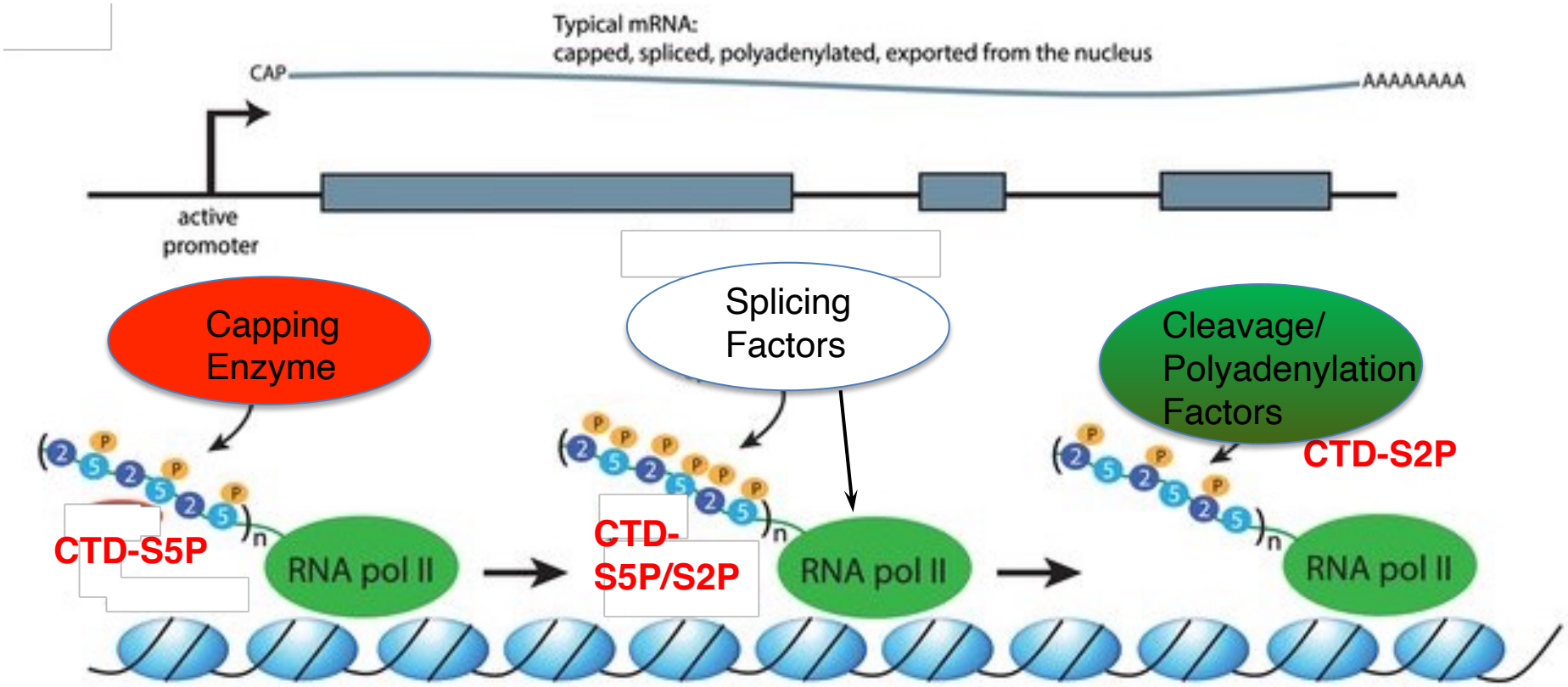
ISE = Intronic Splicing Enhancer Sequence

Phosphorylation of the CTD of the largest subunit of RNA Polymerase II recruits RNA Processing factors/ couples RNA processing to transcription in vivo

Amino acids repeats in the C-terminal Domain (CTD) of Pol.II

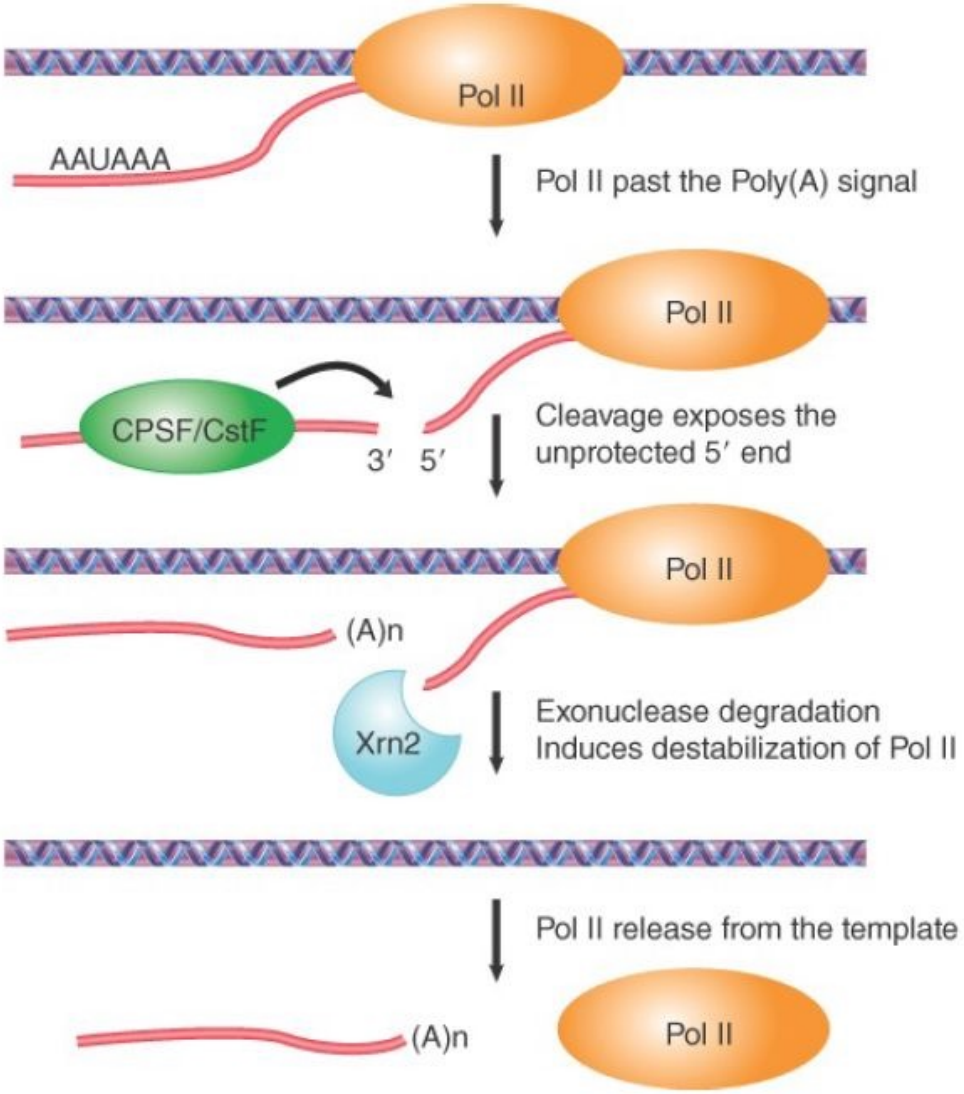






Adapted from Ukaegbu & Deitsch PLoS Pathogens 2017

Co-transcriptional Cleavage and Polyadenylation mediates termination of RNA Polymerase II transcription for most mRNA genes



https://mail.almerja.com/reading.php?idm=151667