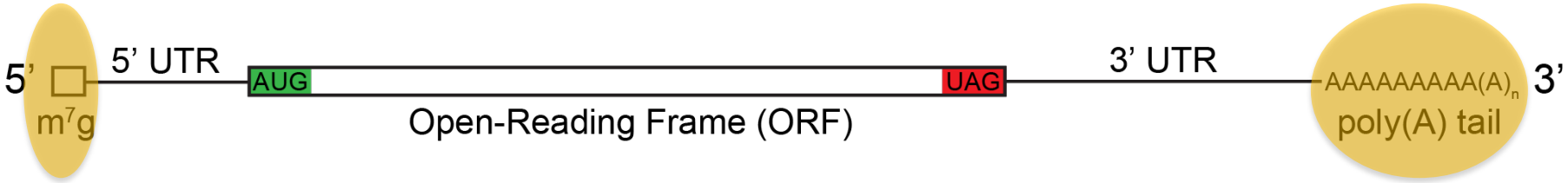
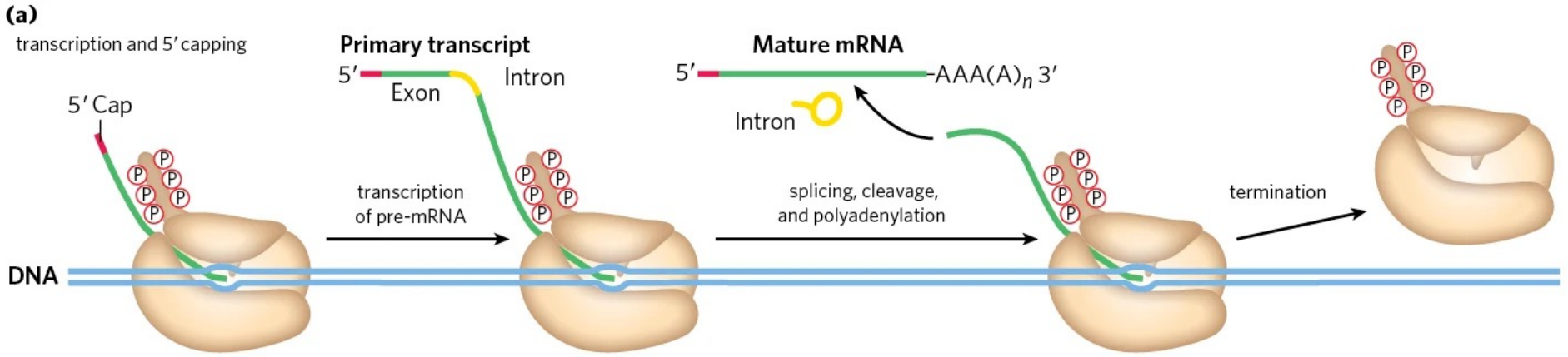
A mature eukaryotic mRNA after processing



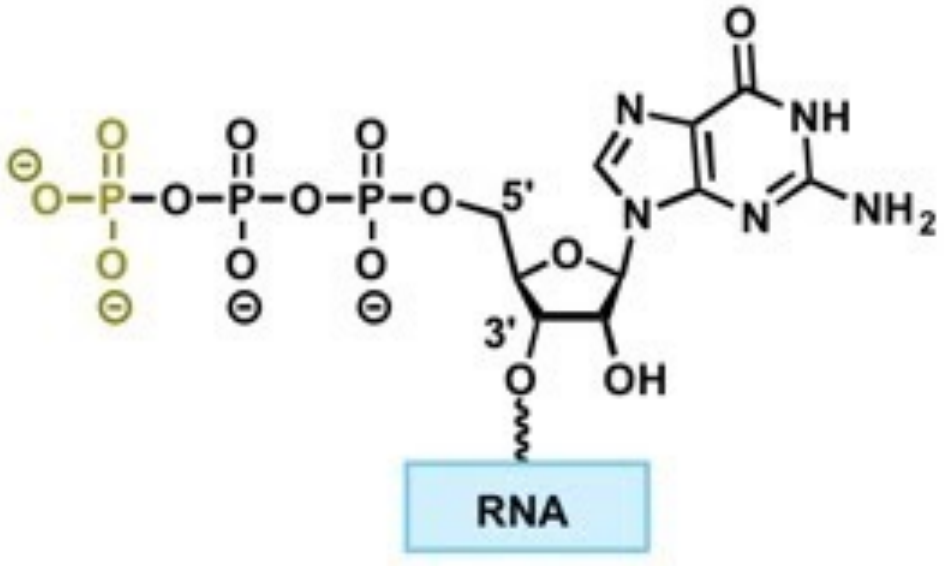
Poly(A) tail length:
70-80 nt in yeast
~200 nt in humans

Highlighted elements are not encoded within the genes: they are added after transcription/co-transcriptionally

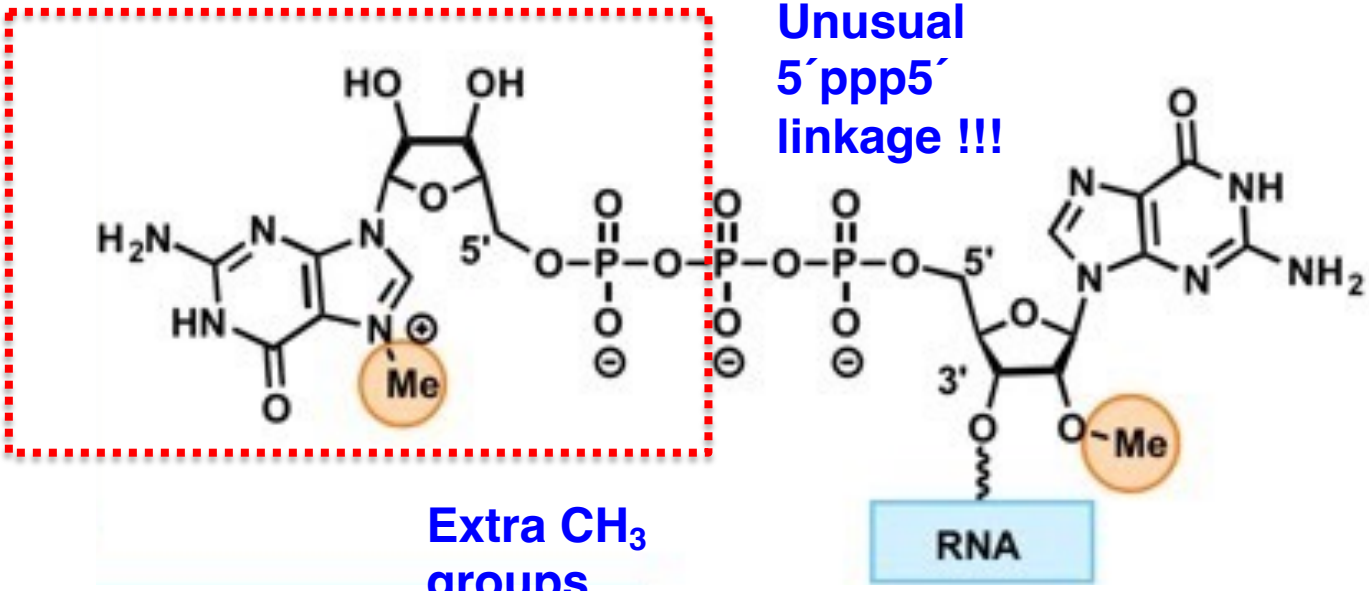
mRNA processing mostly happens co-transcriptionally



RNA 5´-end as it comes out of the RNA polymerase:



The 5´ Cap structure of eukaryotic mRNAs

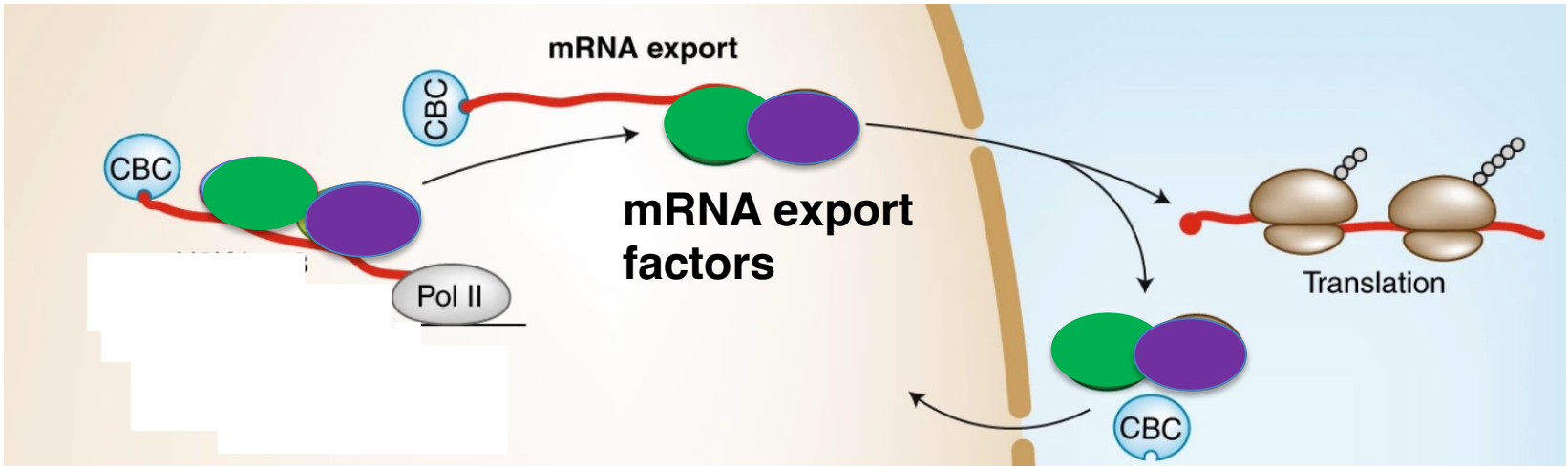


Unusual 5´ppp5´ linkage !!!

Extra CH_3 groups

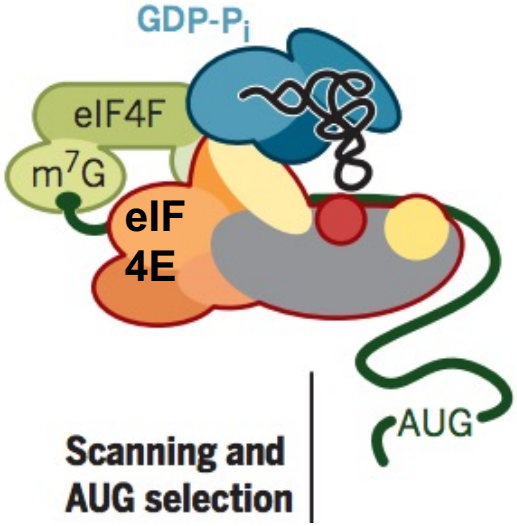
Why a cap structure ?

• Enhances mRNA export out of the nucleus through CBC binding (Cap Binding Complex)



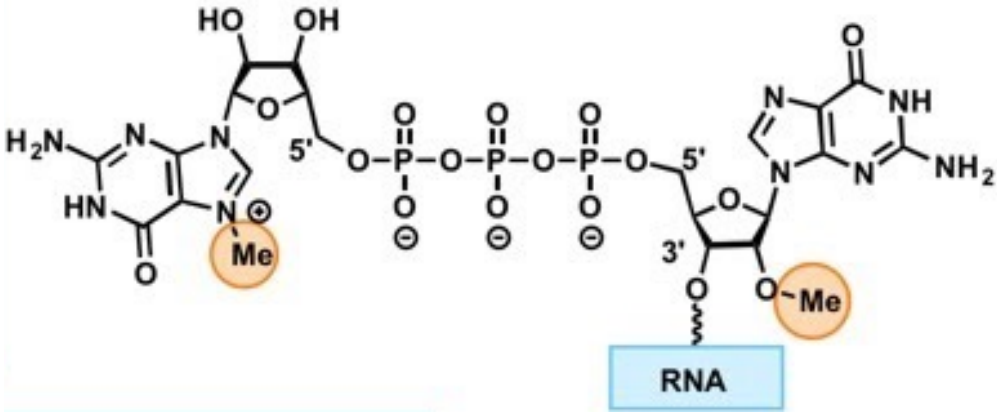
Homolka and Pillai Nat Struct Mol Biol (2019)

• Enhances translation initiation by binding to eIF4E



• Stabilizes mRNA against degradation by 5´->3´ exonucleases

Why does the cap structure protect against degradation by 5´-3´ exonucleases?



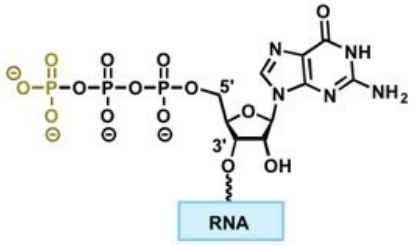
A: The 5´ppp5´ linkage is not a good substrate for exonuclease enzymes which hydrolyze monophosphodiester bonds

B: The methyl group at N7 of the guanine provides a steric hindrance to binding of the exonucleases

C: Positively charged N7 generates an electrostatic repulsion with + charged amino acids of the exonucleases

D: 2´-O-methyl groups interfere with binding of exoribonucleases

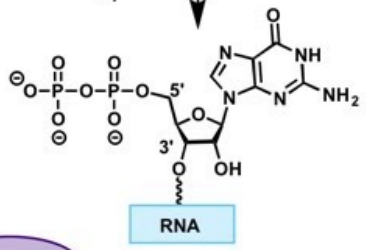
mRNA 5´-end capping



Triphosphatase

RTPase

P_i



PP_i

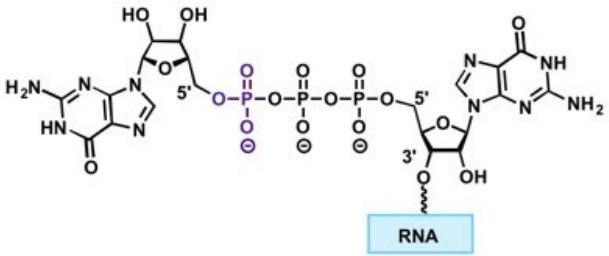
Gp

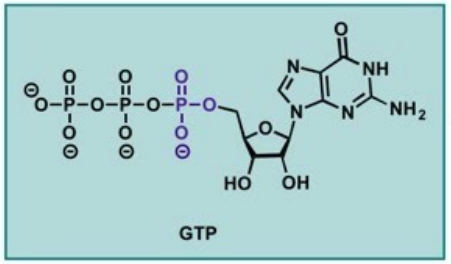
GTase

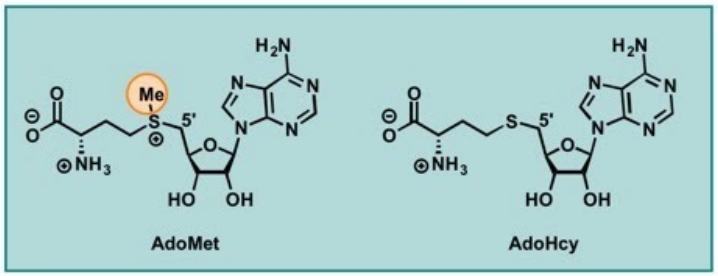
Guanylyltransferase

GTase

GTP





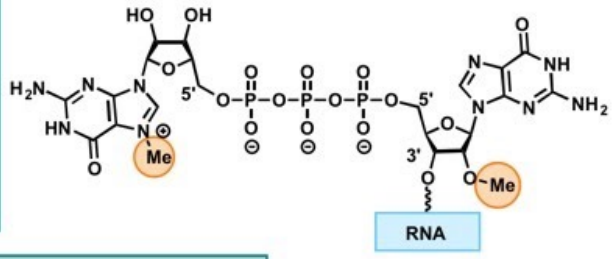


N7-methyl transferase

AdoMet

AdoHcy

N7MTase



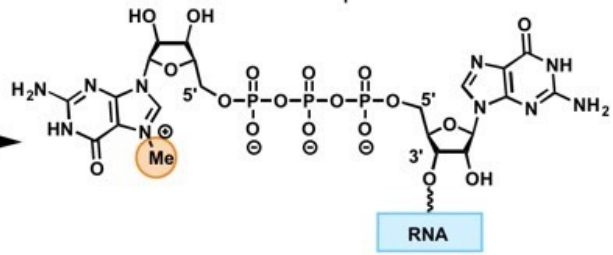
Cap-1

2´-O-methyl transferase

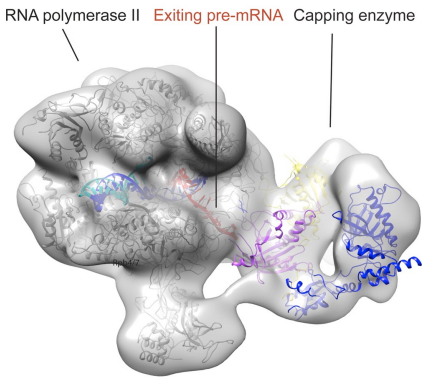
2'OMTase

AdoHcy

AdoMet

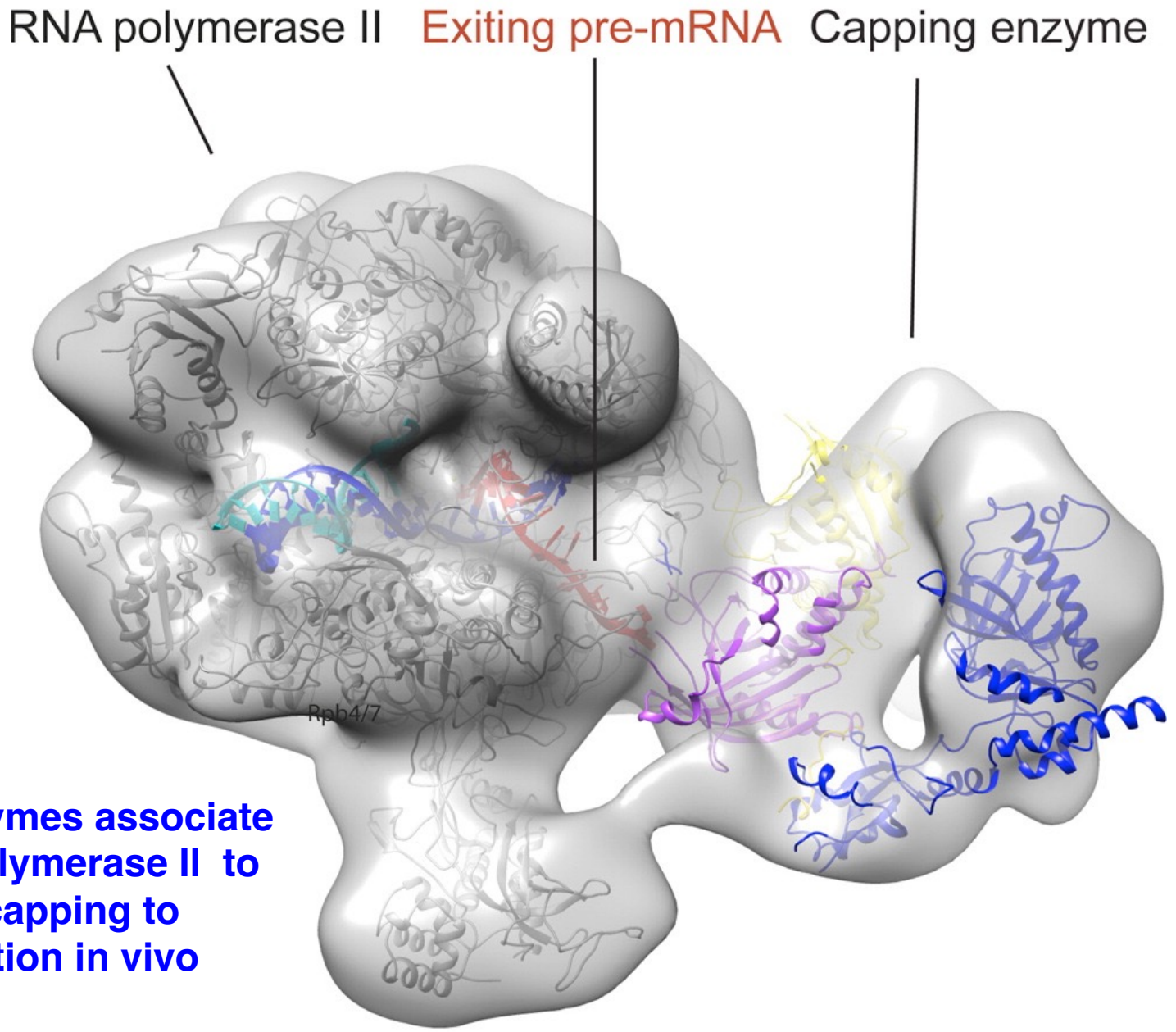


Cap-0



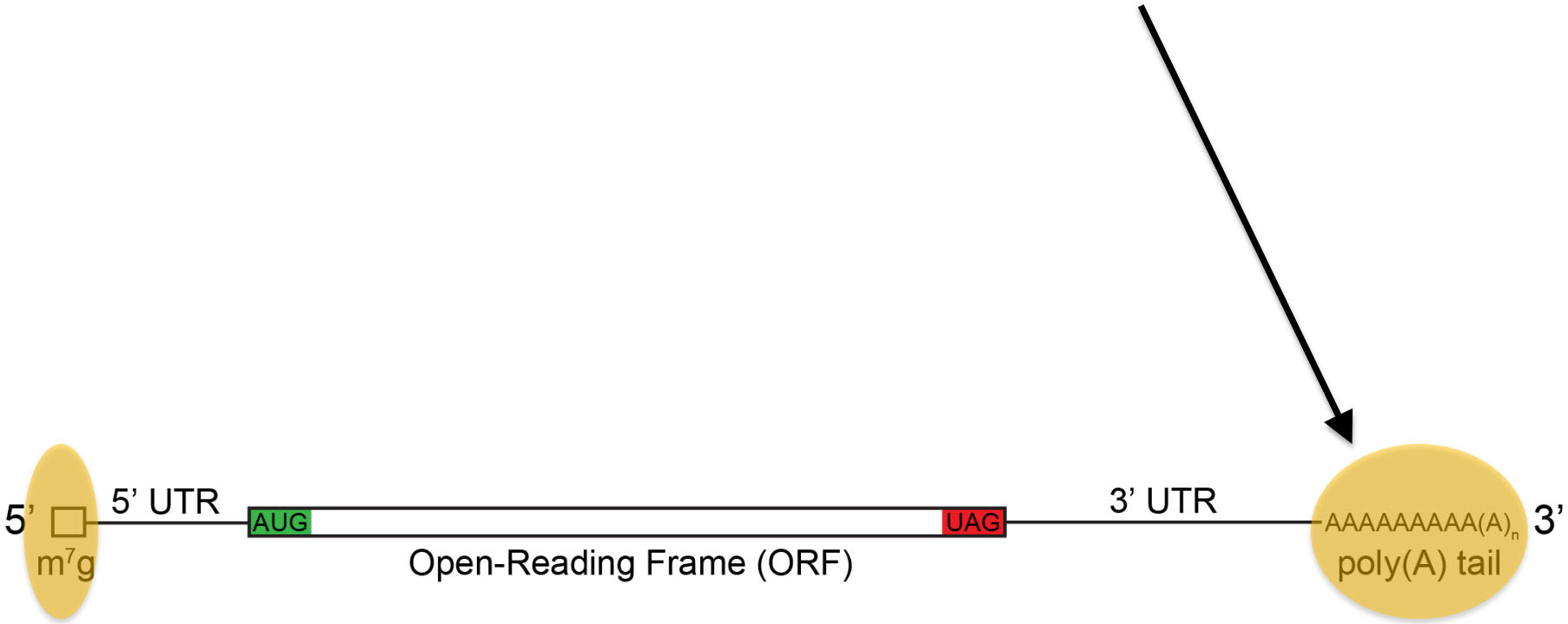
Capping enzymes associate with RNA Polymerase II to couple capping to transcription in vivo

Martinez-Ruboco et al Mol.Cell 2015



Capping enzymes associate with RNA Polymerase II to couple capping to transcription in vivo

At the other end of the transcript: the polyA tail

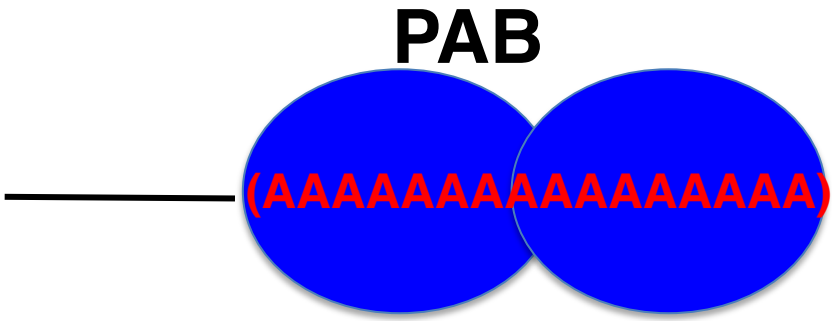


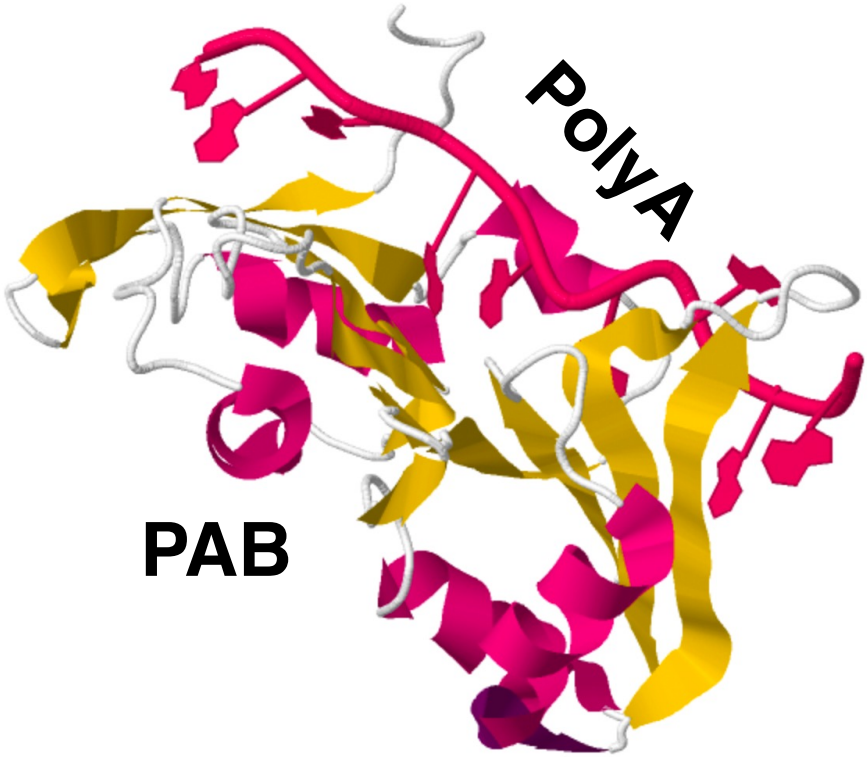
Poly(A) tail length:
70-80 nt in yeast
~200 nt in humans

Highlighted elements are not encoded within the genes: they are added after transcription/co-transcriptionally

Why a poly(A) tail ?

- **Protects mRNA** against degradation by 3´->5´ exonucleases through PAB binding

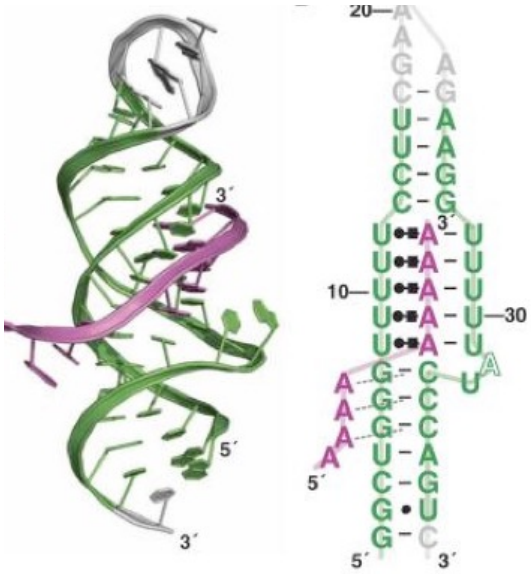




- **Facilitates mRNA export** out of the nucleus

- **Facilitates Translation**: PAB is a component of the translation machinery

Remember the ENE structure from the PAN RNA? Why does it confer stability to the PAN RNA?



A: It inhibits the deadenylation step

B: It inhibits the decapping step

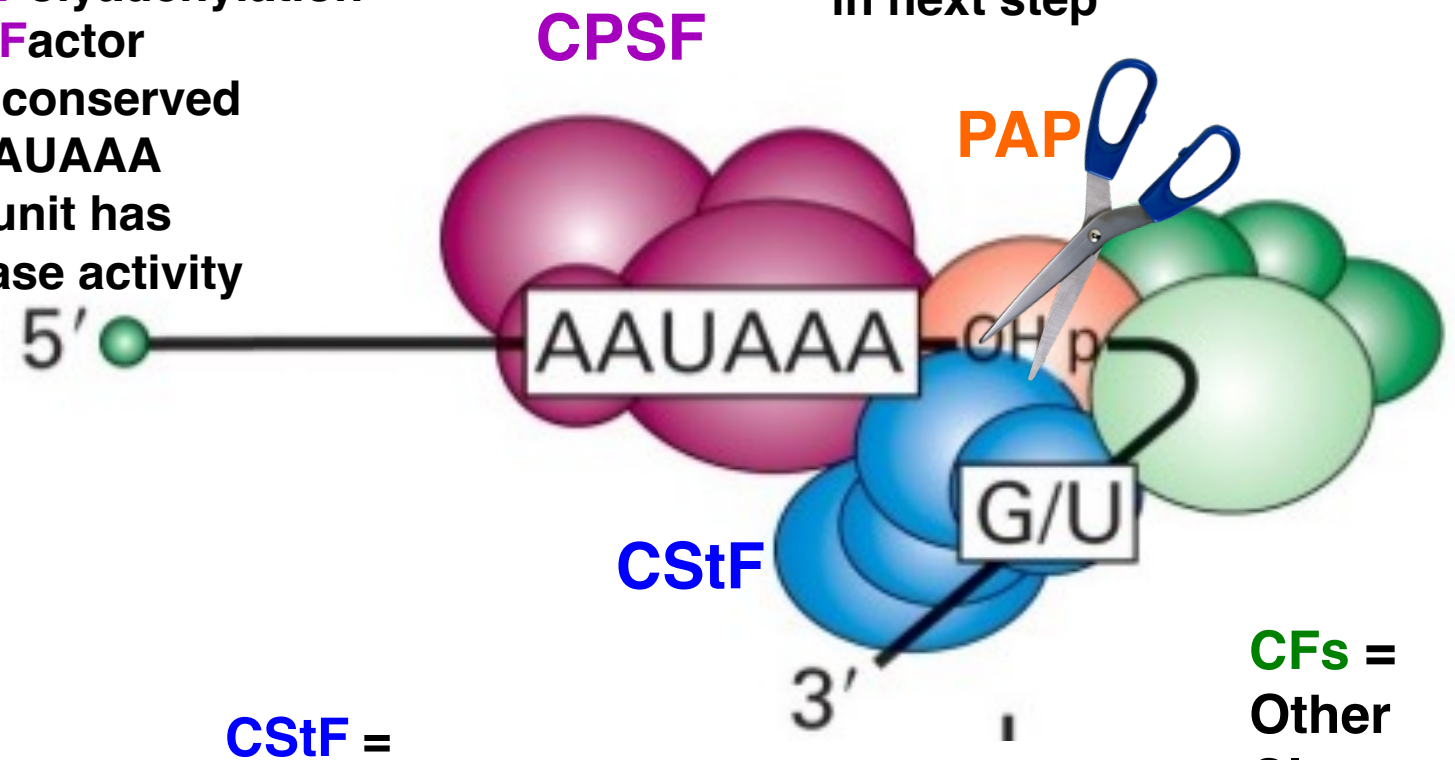
C: It inhibits 5´-3´ degradation by Xrn1/Rat1

D: It inhibits 3´-5´ degradation by the exosome

1st step of polyA addition = mRNA cleavage

CPSF
Cleavage / **P**olyadenylation **S**pecificity **F**actor
- binds the conserved hexamer AAUAAA
- 73kD subunit has endonuclease activity

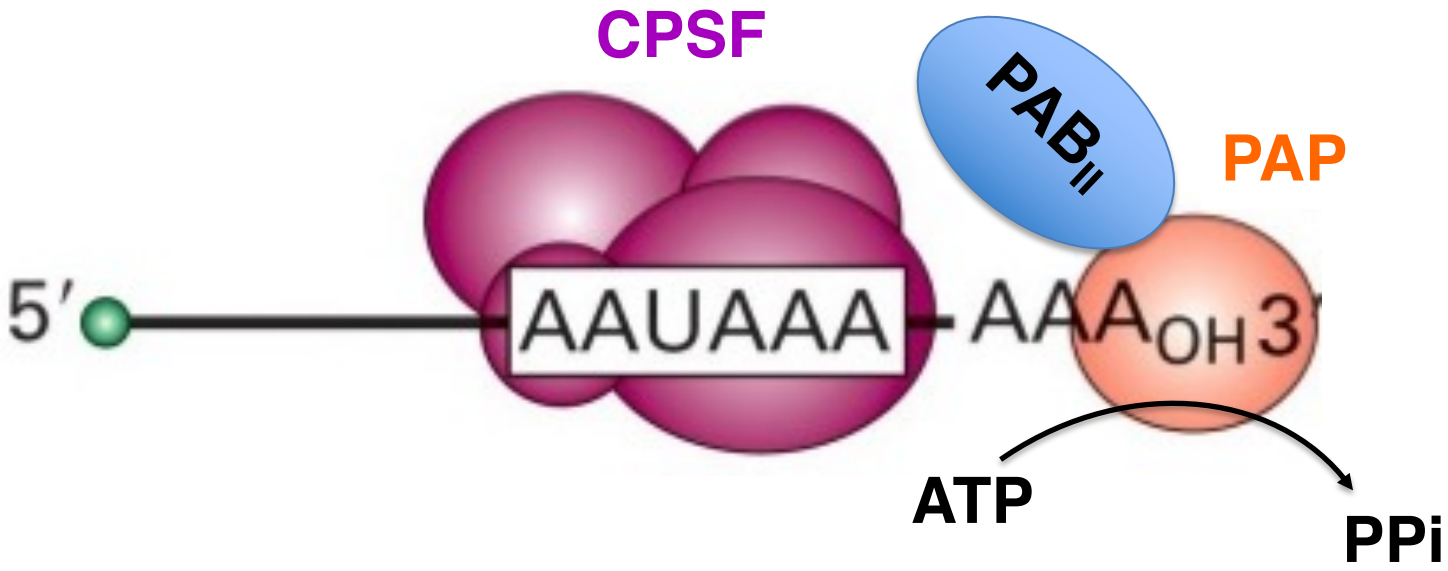
PAP = **P**oly(**A**) **P**olymerase polymerizes poly (A) tail in next step



CStF = **C**leavage **St**imulatory **F**actor recognizes the G/U rich element

CFs = Other Cleavage Factors

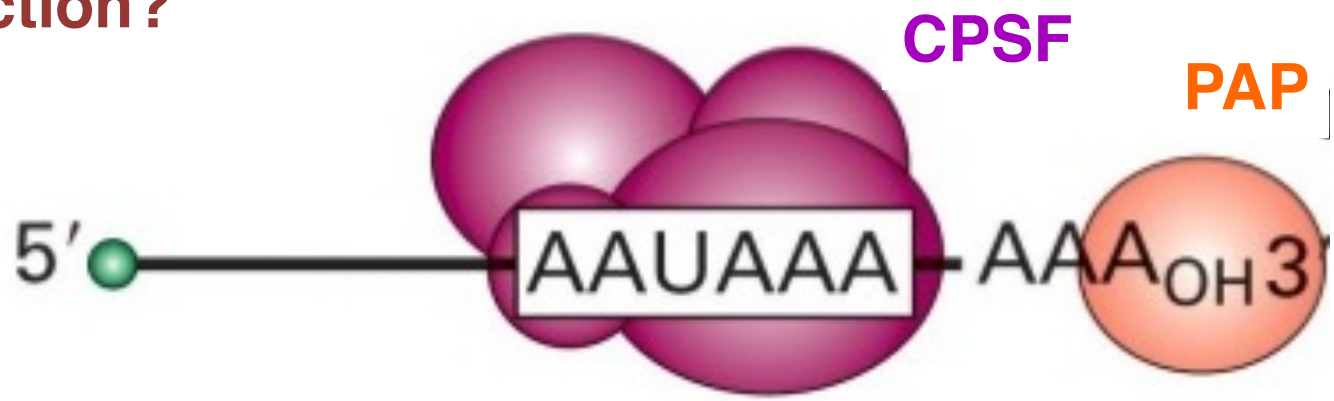
2nd step = Poly(A) addition



PAB$_{II}$ = Poly(A) Binding Protein II
-binds the nascent poly(A) tail when the size is about 10nt
- provides processivity to poly(A) polymerase
- controls the length of the poly(A) tail (70nt in yeast, 200 in mammals)

What catalyzes each step of the reaction?
(1) CPSF 73 -> cleavage
(2) PAP -> Poly A addition

Why is the 3´-RNA fragment derived from cleavage not shown in the 2nd step of the reaction?



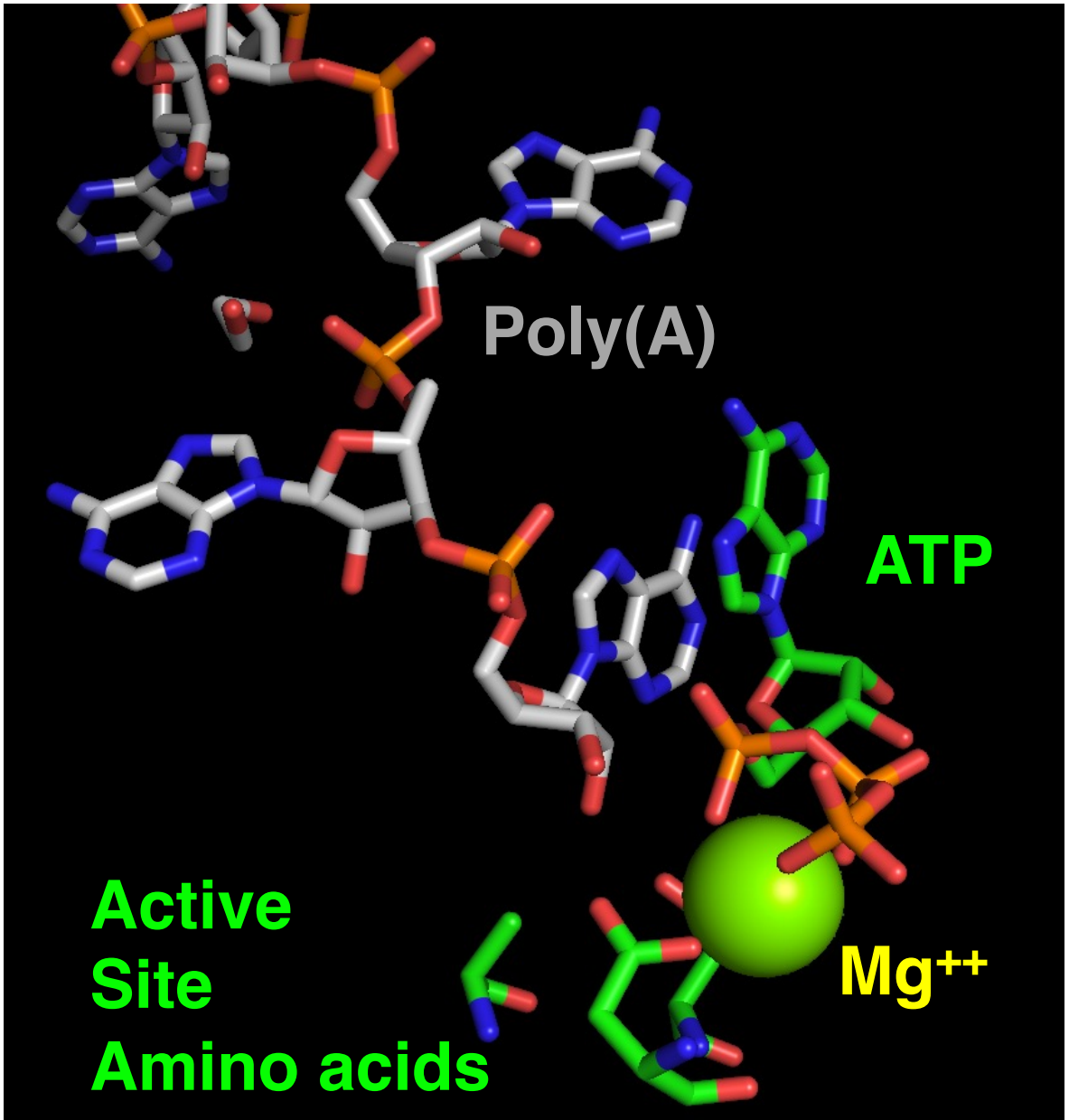
A: It is not involved in the reaction so it is not shown

B: It is immediately bound by PAB and thus protected

C: It is not shown because it is degraded by a 5´-3´ exonuclease as it lacks a cap structure

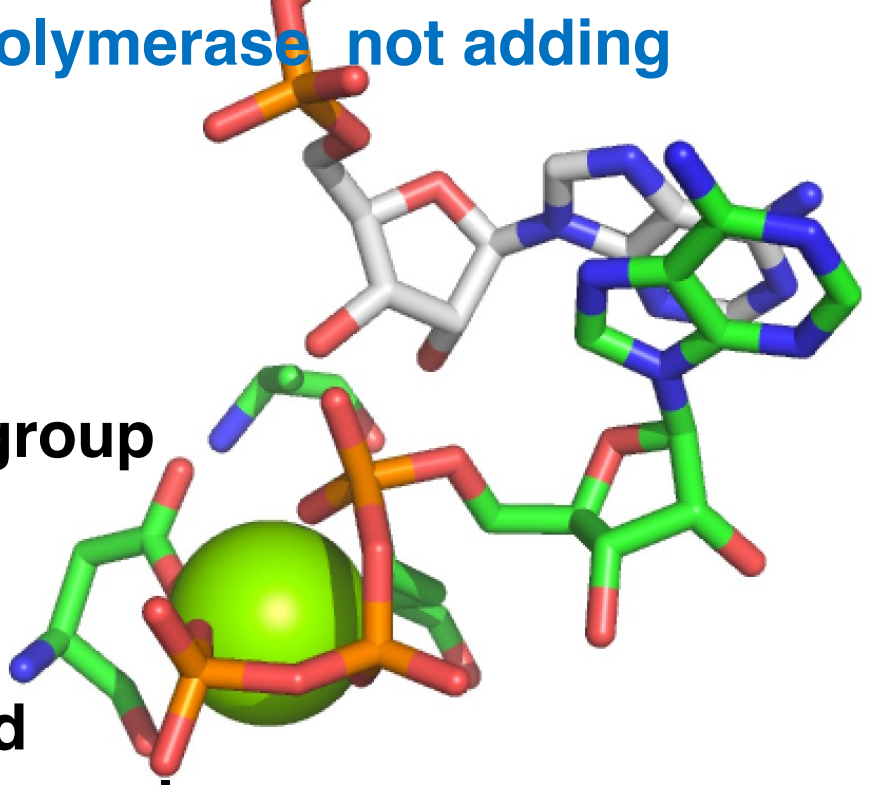
D: It is not shown because it is degraded co-transcriptionally by the 5´-3´ exonuclease activity of RNA polymerase (ie nick translation)

The active site of Poly(A) Polymerase in complex with a poly(A) RNA and ATP



PyMol: PolyAPol.pse

Why is the PolyA polymerase not adding ATP to the 3´-end of the RNA in this structure?



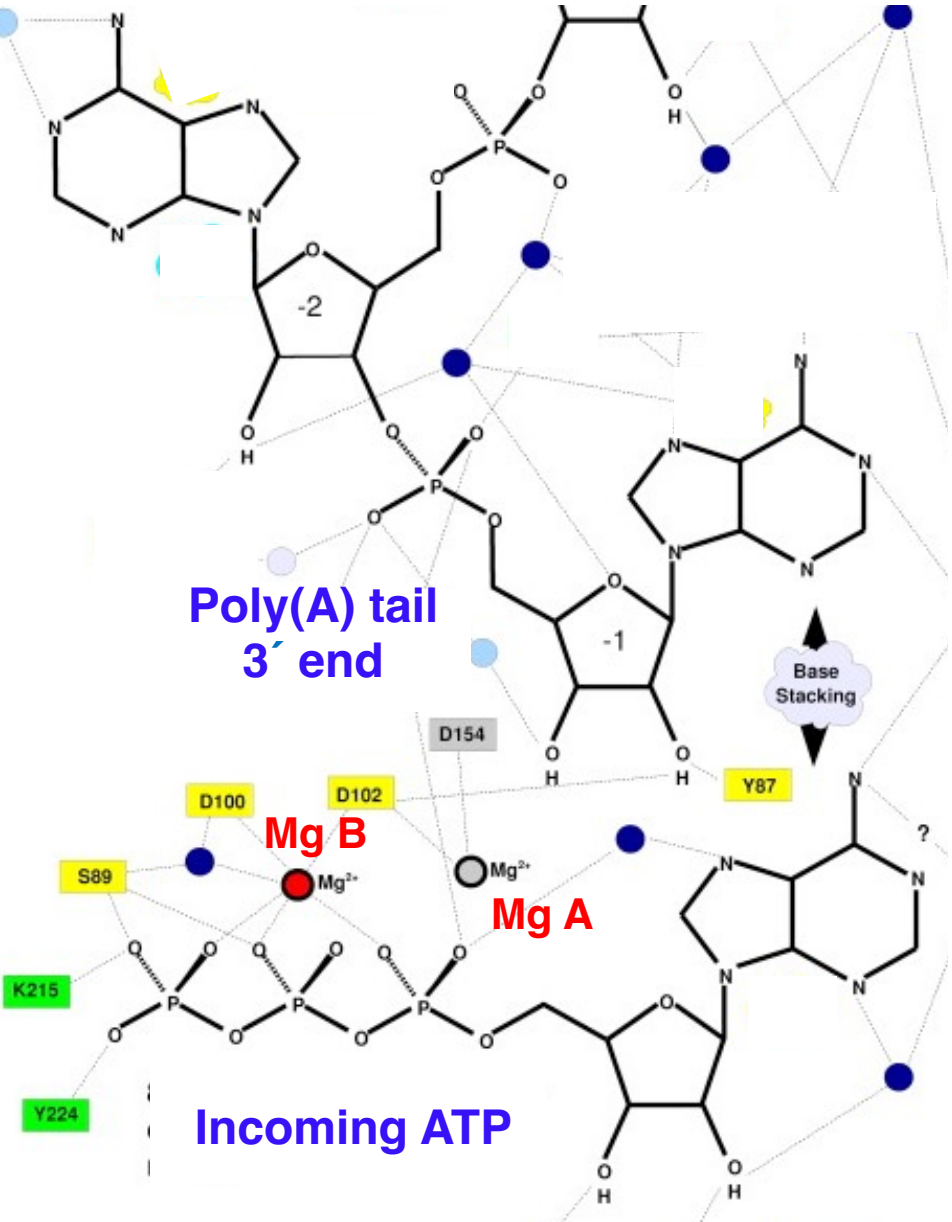
A: The RNA lacks a 3´-hydroxyl group and so it lacks the nucleophile

B: The ATP is chemically modified such as the alpha phosphorus can no longer serve as a leaving group

C: The enzyme has been mutated such as it has disrupted the binding site for the second metal ion and chemistry can no longer happen

D: Chemical reactions do not occur in crystals

PAP uses a 2 metal ion catalytic mechanism similar to DNA/RNA Polymerases…but no template



Specificity of Poly(A) Polymerase for ATP:

- Non-polar interactions above and below the plane of the base
- Water mediated interactions to polar groups in the plane of the base
- N226 participates in recognition of the Watson-Crick face of ATP
- Stacking interactions (A-A) favor purines
- Steric exclusion of Gs

Major mRNA Degradation pathway in eukaryotic cells:

Elements that need to be removed for efficient mRNA degradation:

The cap structure The poly(A) tail	will allow access to exoribonucleases for digestion

Major pathways:

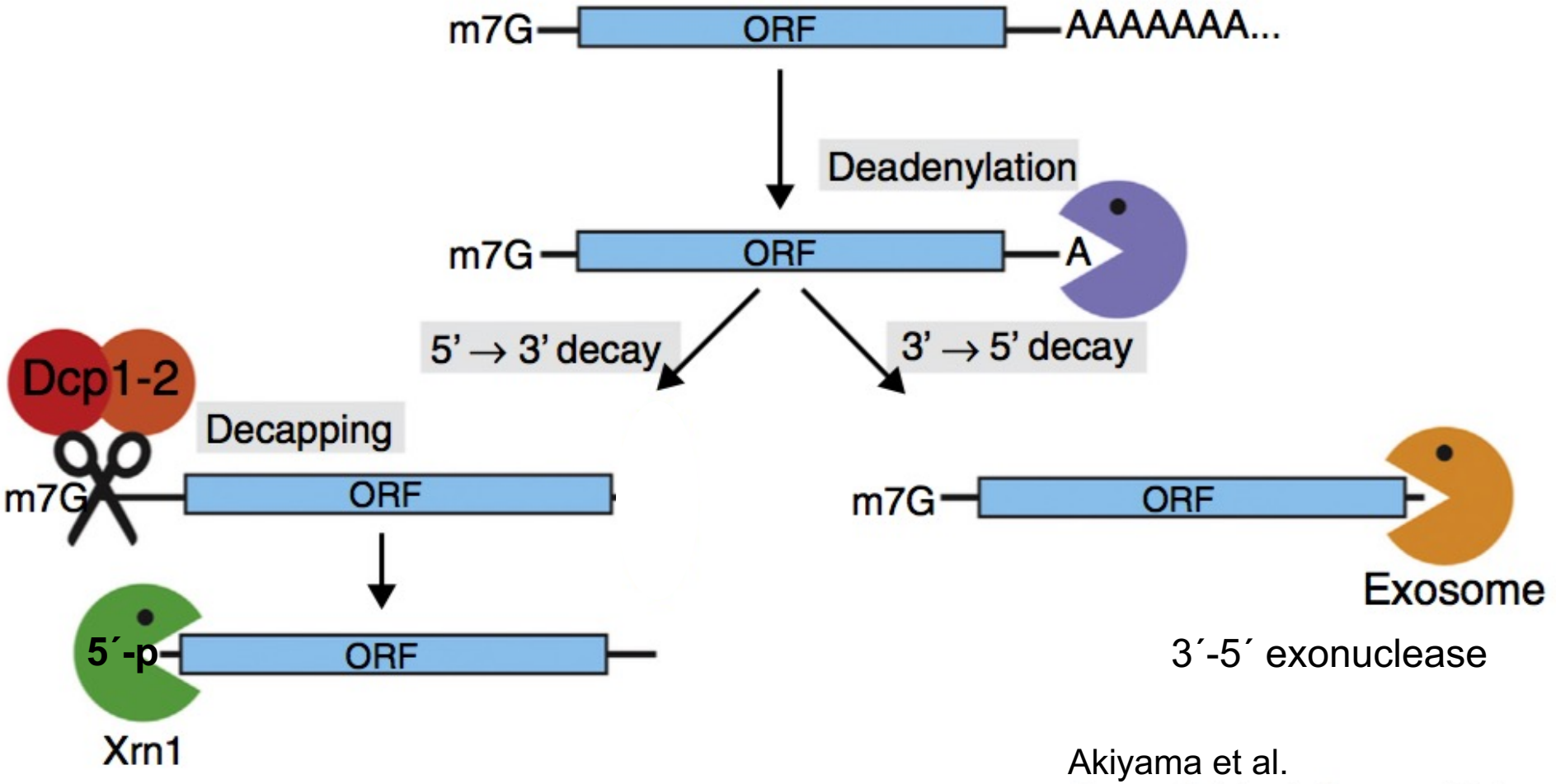
1) Deadenylation — Deadenylases
2) Decapping — Decapping complex (Dcp)

3) Exonuclease digestion of the body of the mRNA

5´->3´ Exonucleases: Xrn1

3´->5´ Exonucleases: The Exosome

Major mRNA Degradation pathway in eukaryotic cells:



5´-3´ exonuclease

3´-5´ exonuclease

Akiyama et al.
Current Opinion in Structural Biology
2016

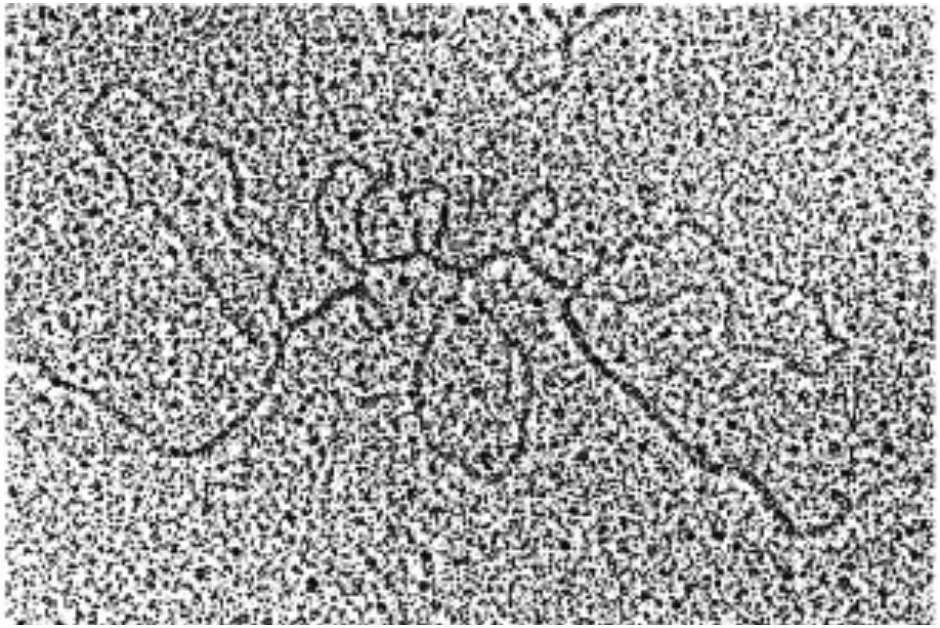
pre-mRNA Splicing

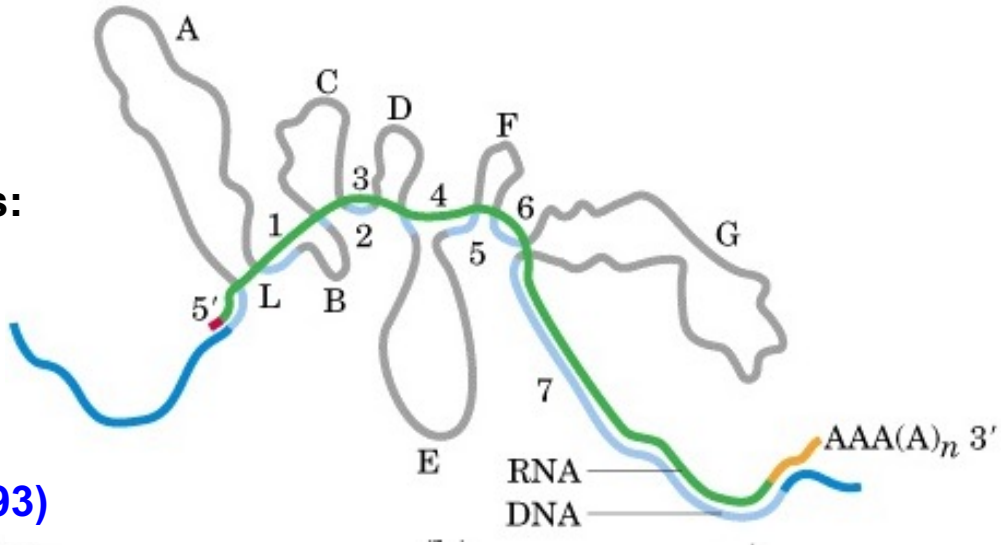
Definitions

Intron = RNA piece between 2 exons that is excised during splicing
Exon = RNA sequence that is spliced with another exon

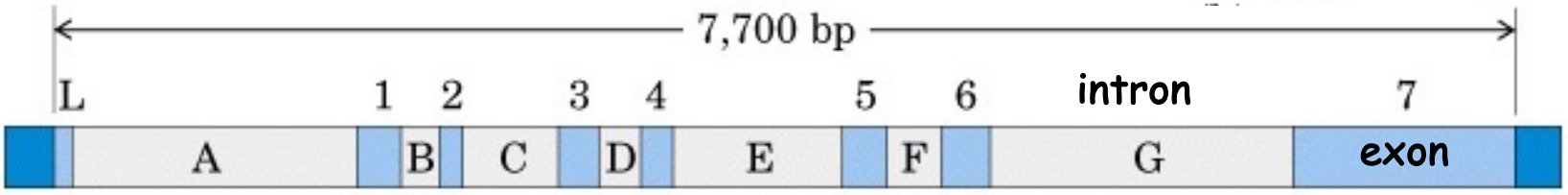
An exon does not always correspond to a coding sequence !
An intron does not always correspond to a non-coding sequence !

First proof of the existence of "split" genes: Hybridization of a gene DNA with the corresponding mature mRNA shows that they are not colinear





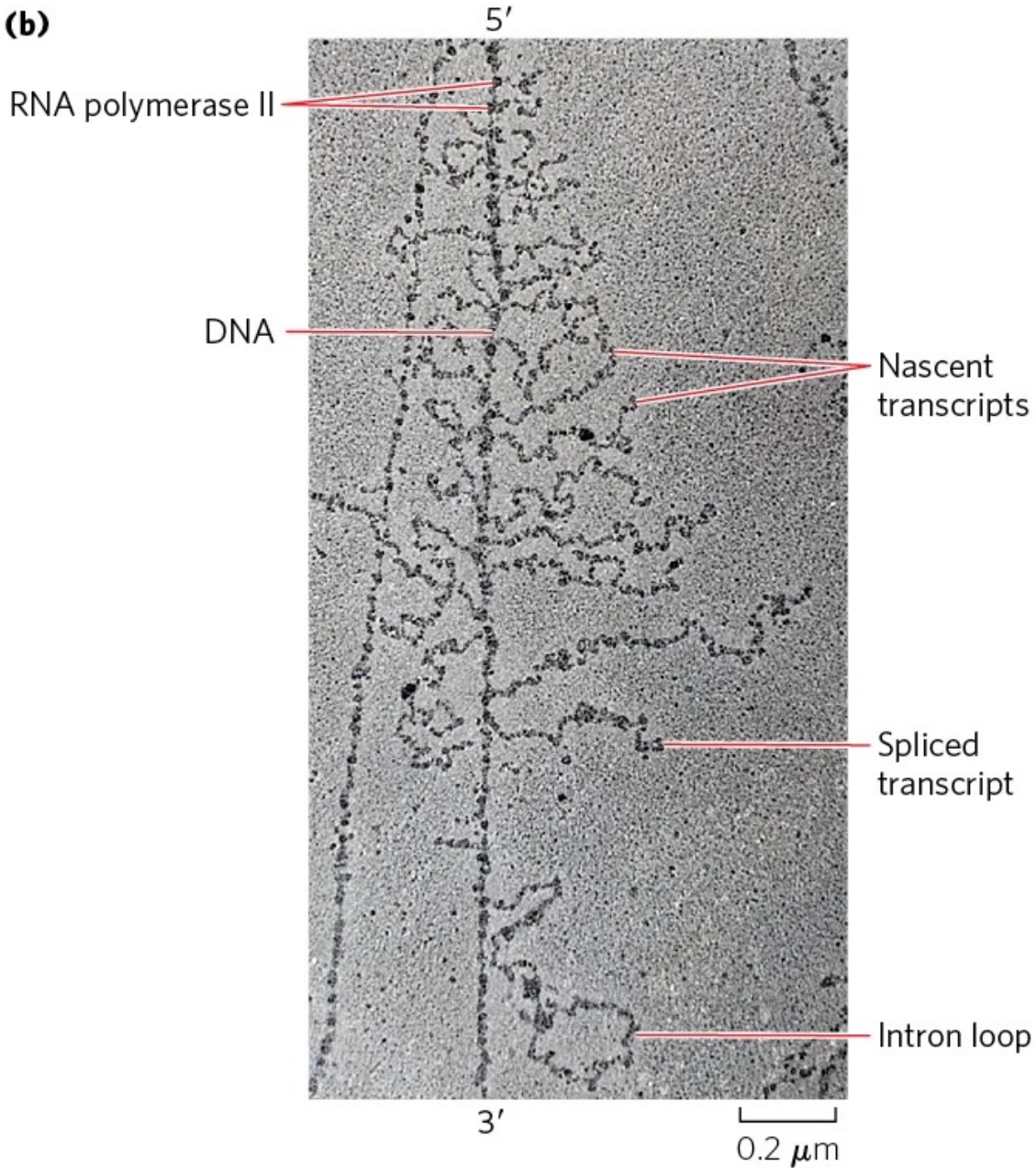
Nobel Prize Medicine (Sharp, Roberts 1993)



Features of introns and exons

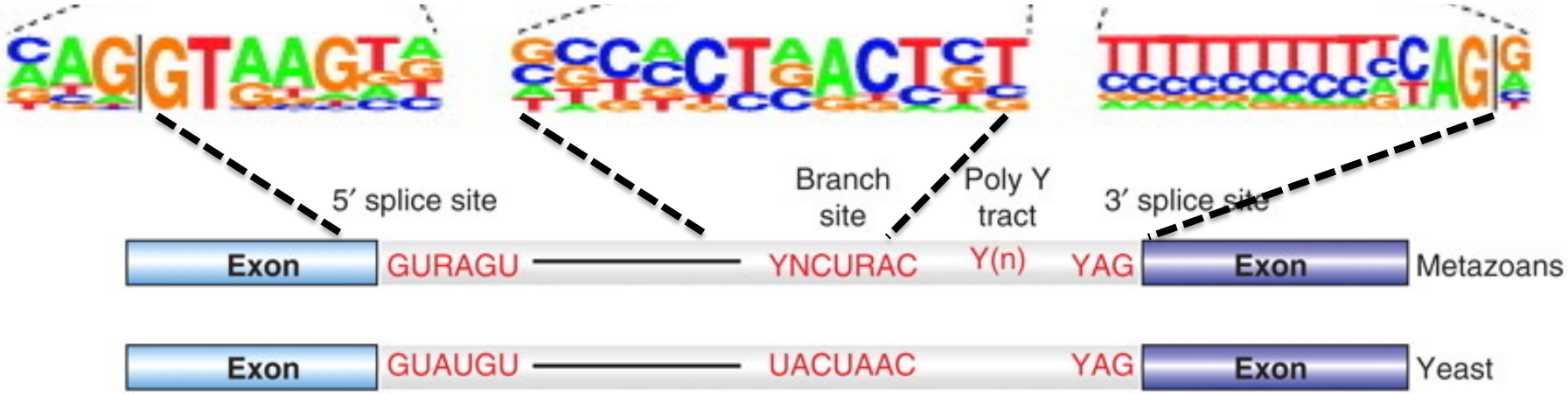
- **Most exons are less than 1,000 nts, many are 100-200nts**
- **Introns vary from ~50 nts to >700,000 nts --> median length = ~1800**
- **~20,000 genes in the human genome include >200,000 introns**

Splicing happens co-transcriptionally



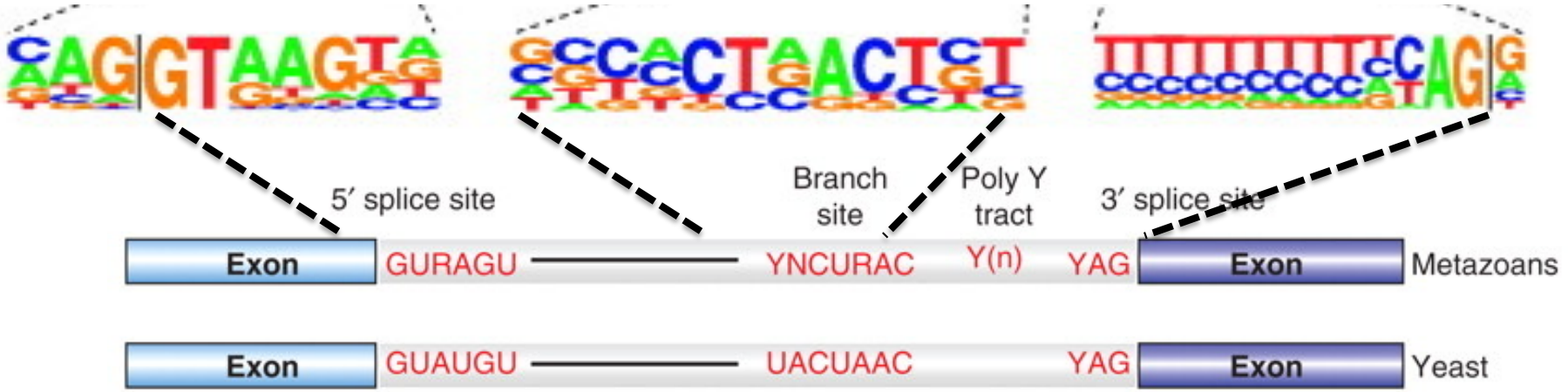
Exon/ introns junctions and consensus sequences of nuclear introns

"Sequence logo" representation of consensus sequences



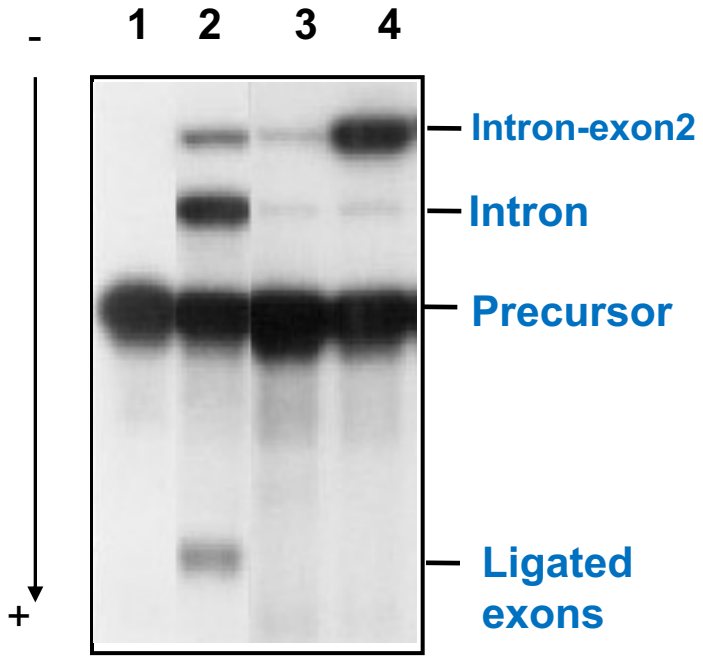
Exon/ introns junctions and consensus sequences of nuclear introns

"Sequence logo" representation of consensus sequences

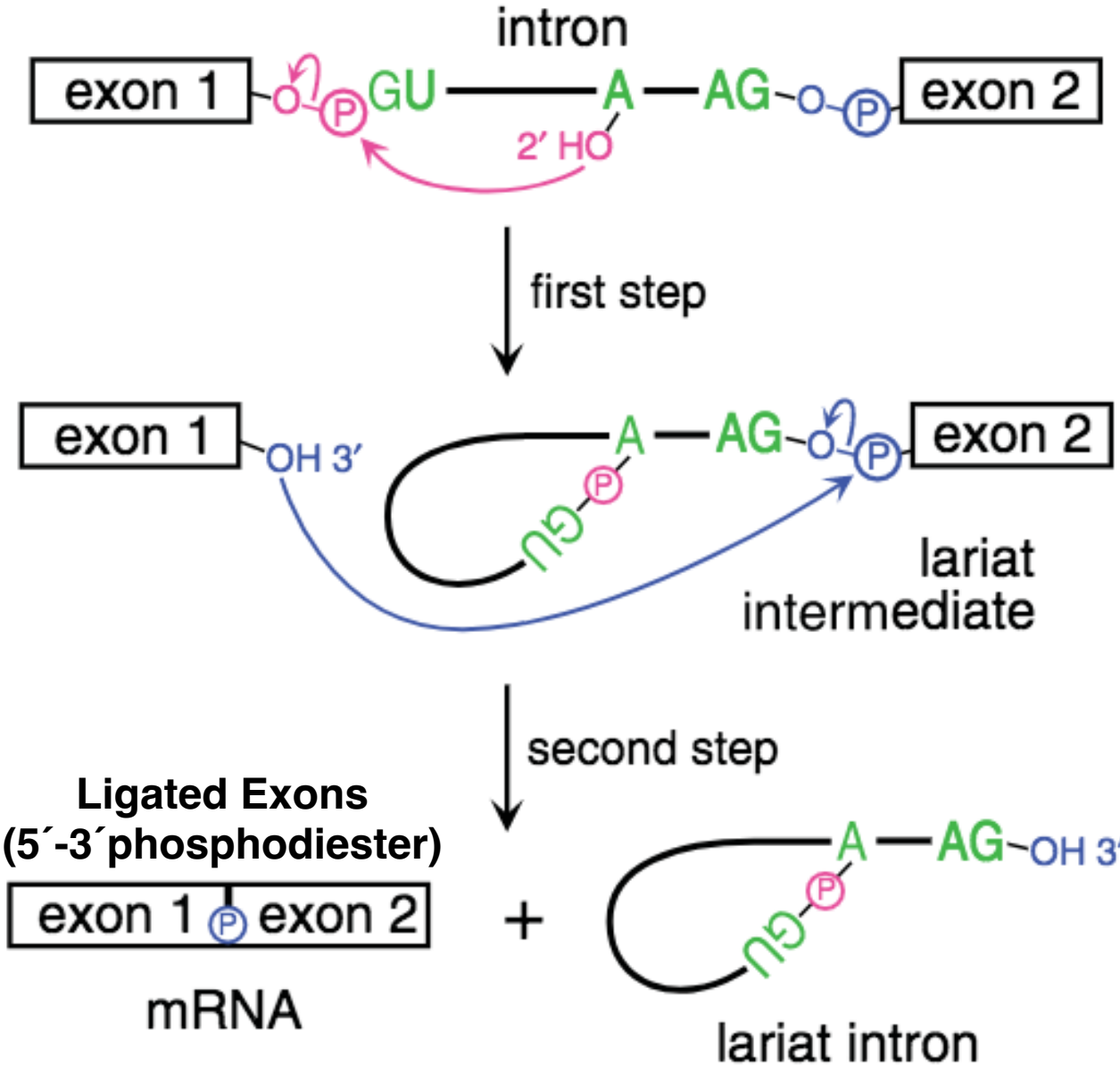


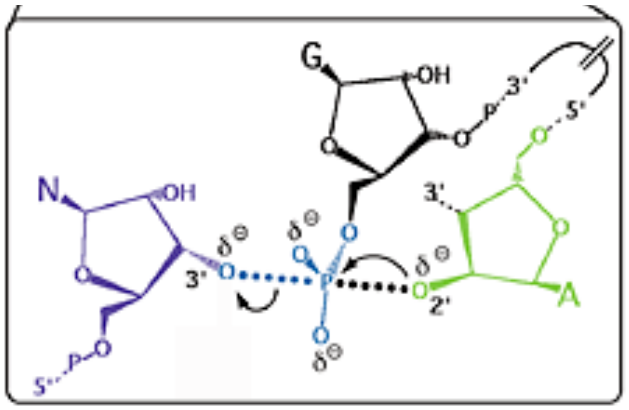
In vitro pre-mRNA splicing reaction:

- **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