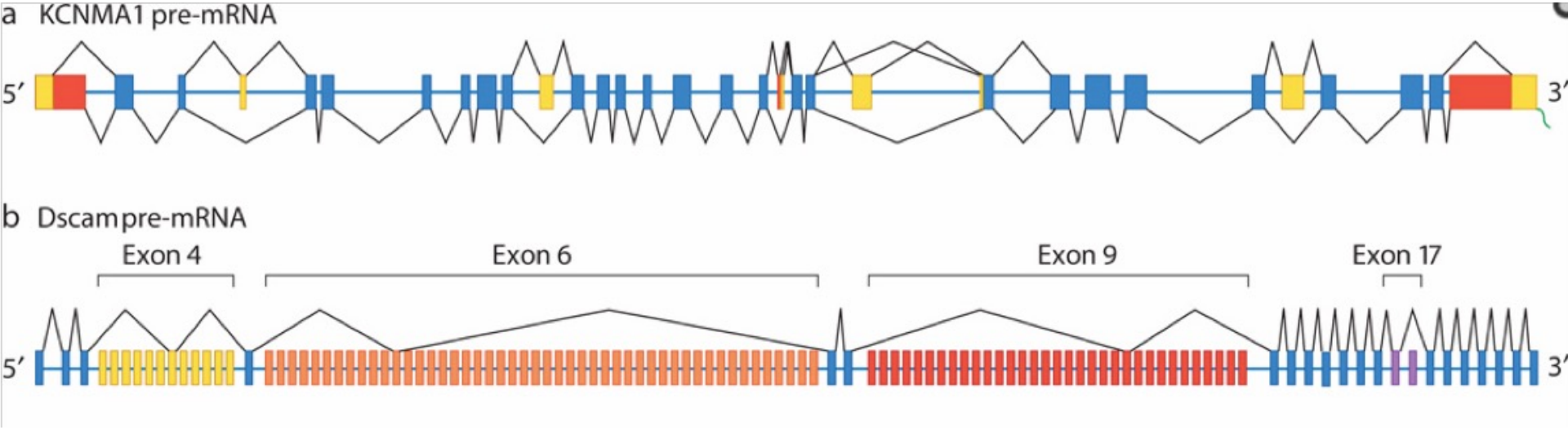
Generating Diversity through alternative splicing

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

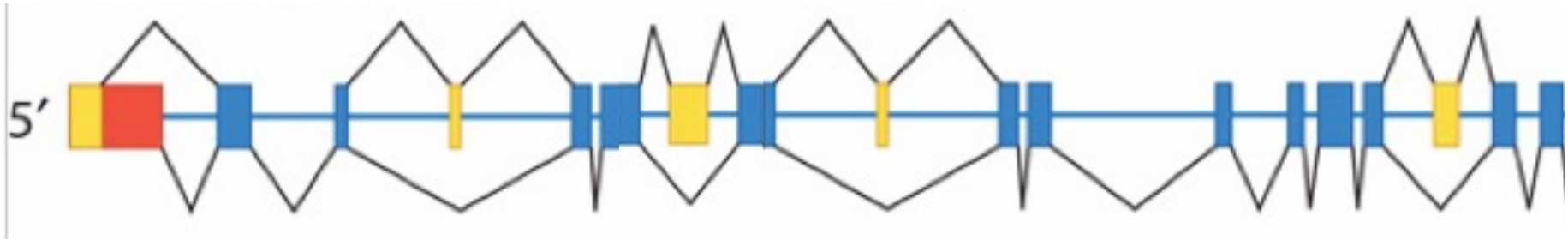
Total possible combinations =12 X 48 X 33 X 2= 38,016

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.

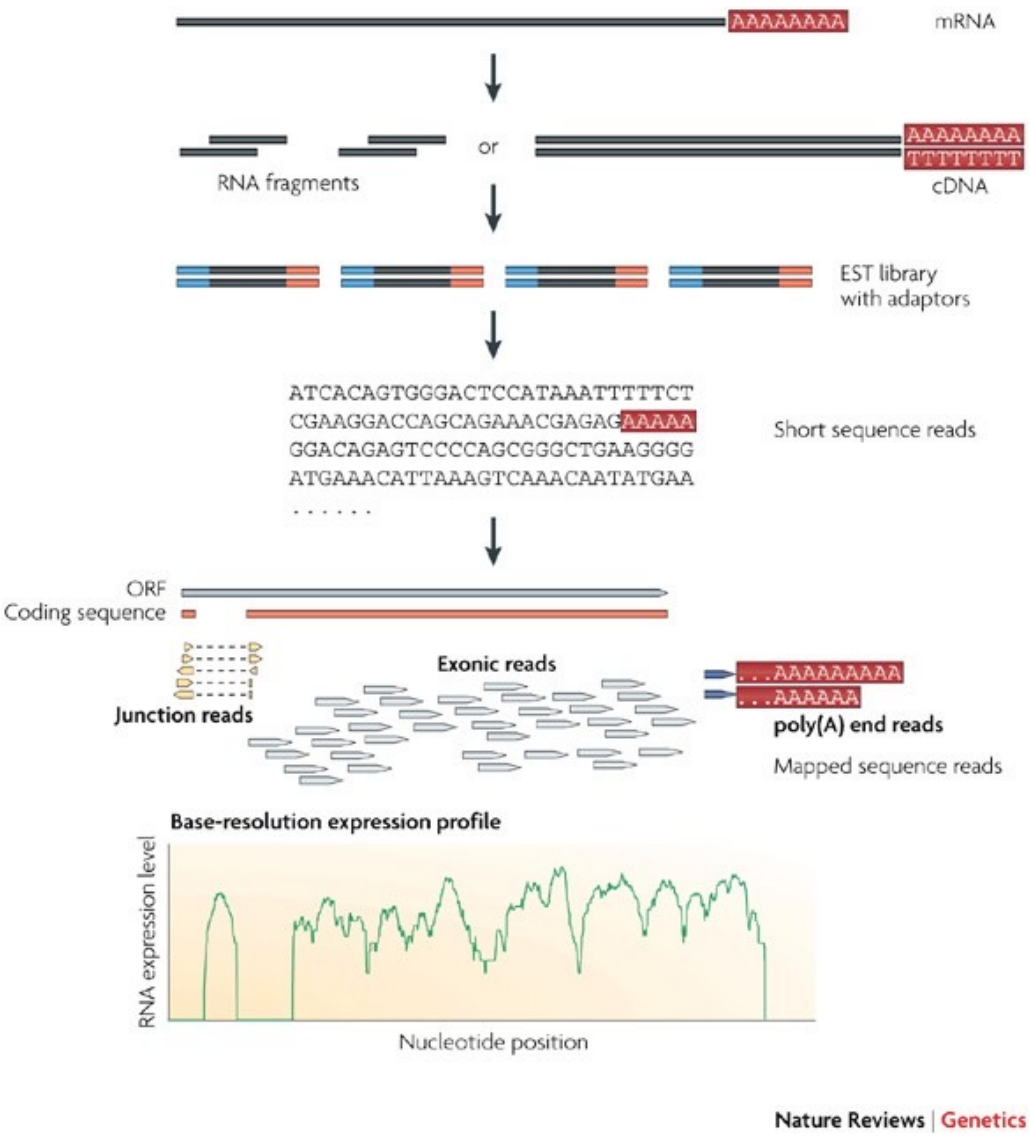
18 exons

- **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?**

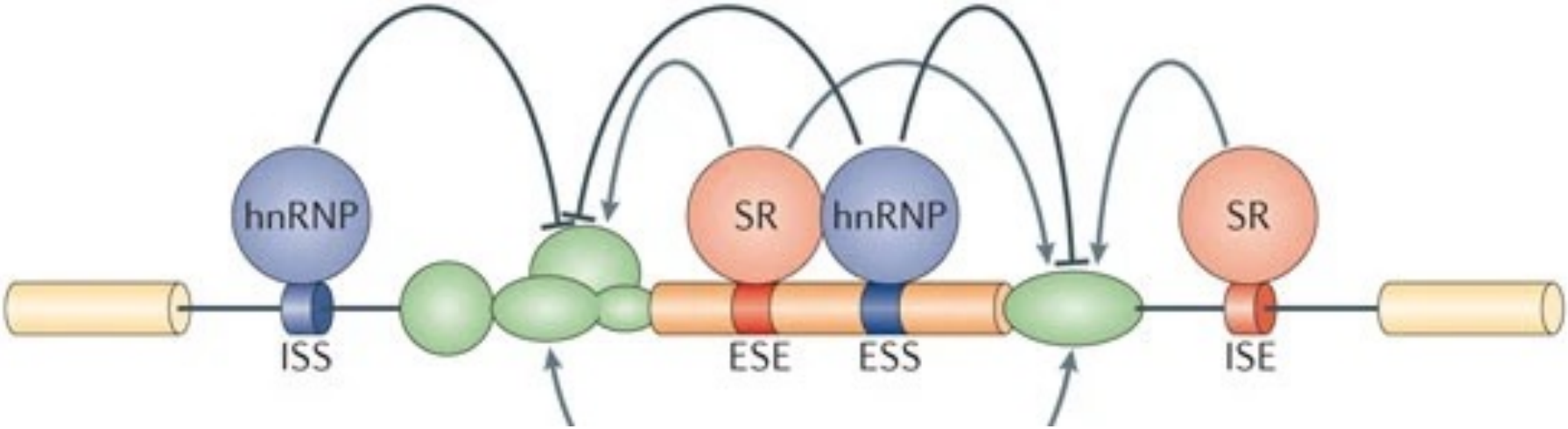
Snyder et al., Nat Rev Genetics 2009

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

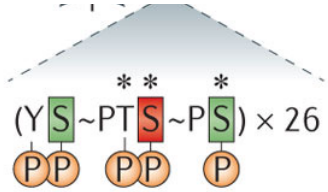
ESE/ESS = Exonic Splicing Enhancer/Silencer Sequence

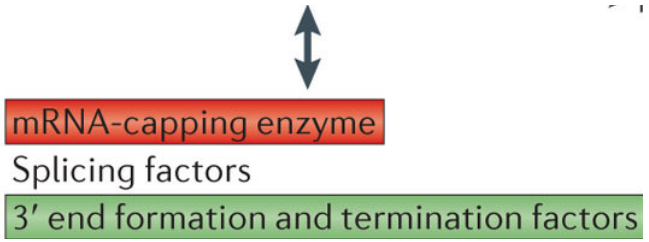
Spliceosome Components

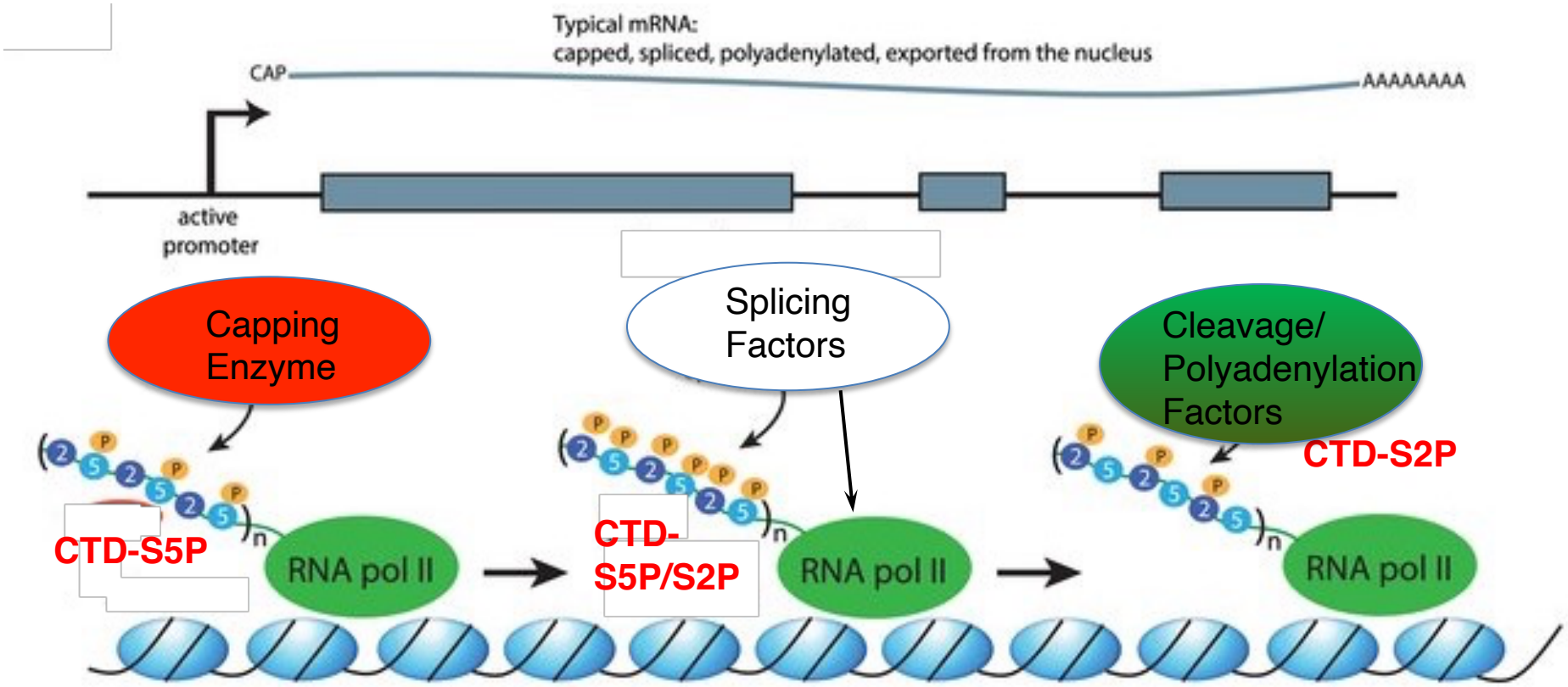
ISE = Intronic Splicing Enhancer Sequence

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

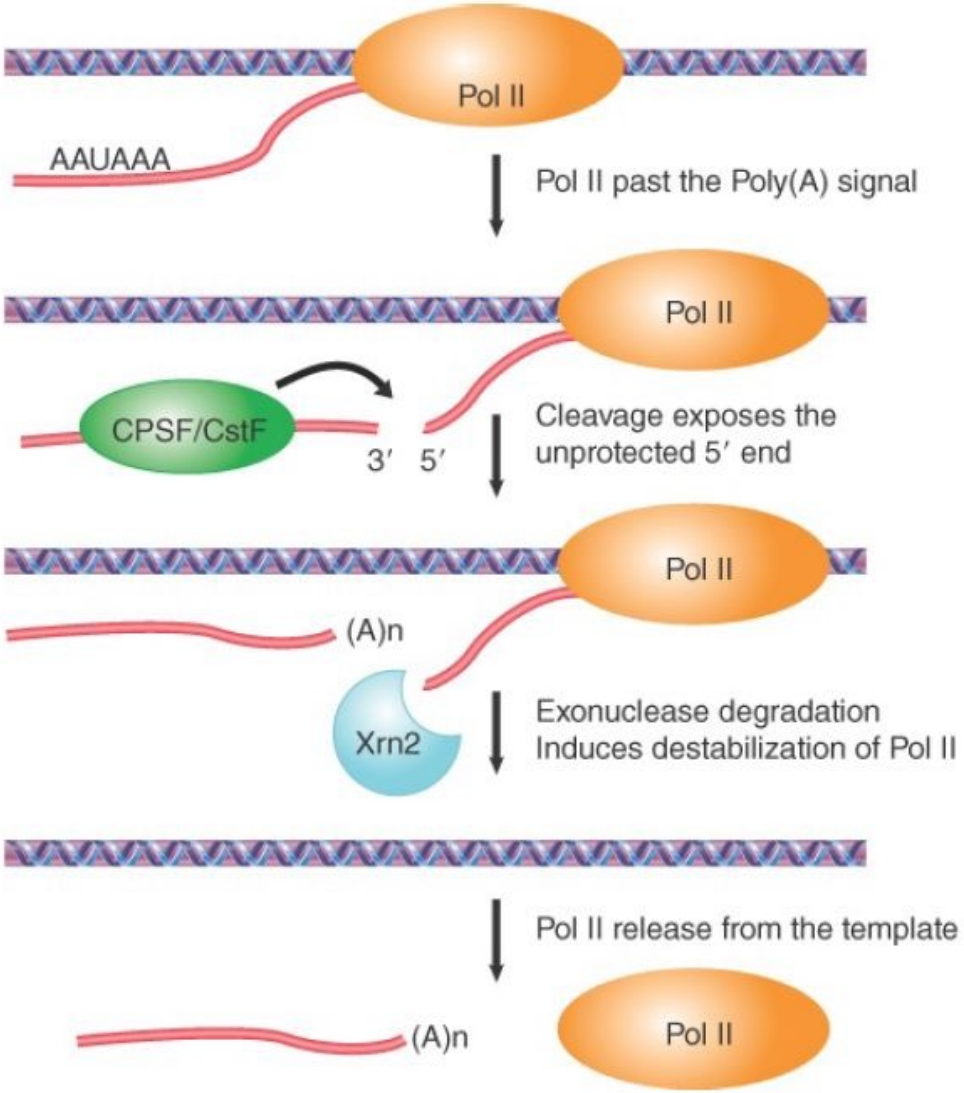






Adapted from Ukaegbu & Deitsch PLoS Pathogens 2017

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



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

Different splicing mechanisms exist for different types of introns:

Intron type	Enzyme(s) involved	(Genomic) localization
Nuclear Genes	Spliceosome	>99% of nuclear genes encoding proteins in eukaryotes
Group I Introns	self-splicing = intron Ribozyme NO Proteins!	• Nuclear rRNA Tetrahymena • genes of organelles in fungi and plants • some bacteriophages/ bacterial genes
Group II introns	self-splicing = intron Ribozyme splicing pathway is identical to nuclear pre-mRNA introns	• organelles fungi/plants/animals • bacteria • bacteriophages
tRNAs	Endonuclease+ligase = protein catalyzed	tRNA genes

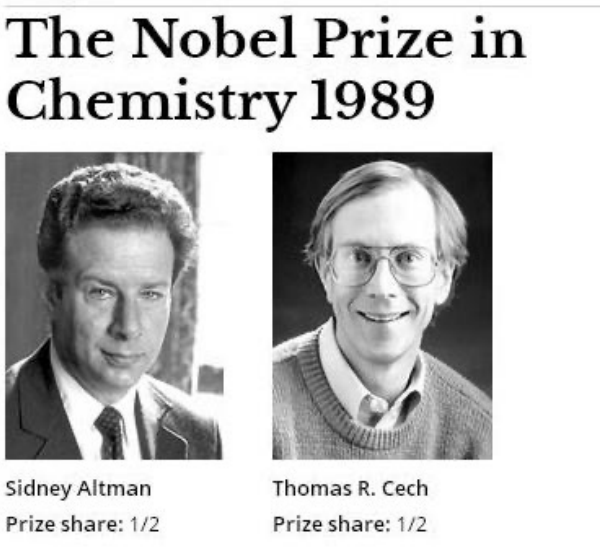
T.Cech

ARTICLES

RNA-catalysed synthesis of complementary-strand RNA

Jennifer A. Doudna & Jack W. Szostak

Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA



The Nobel Prize in Chemistry 1989

Sidney Altman

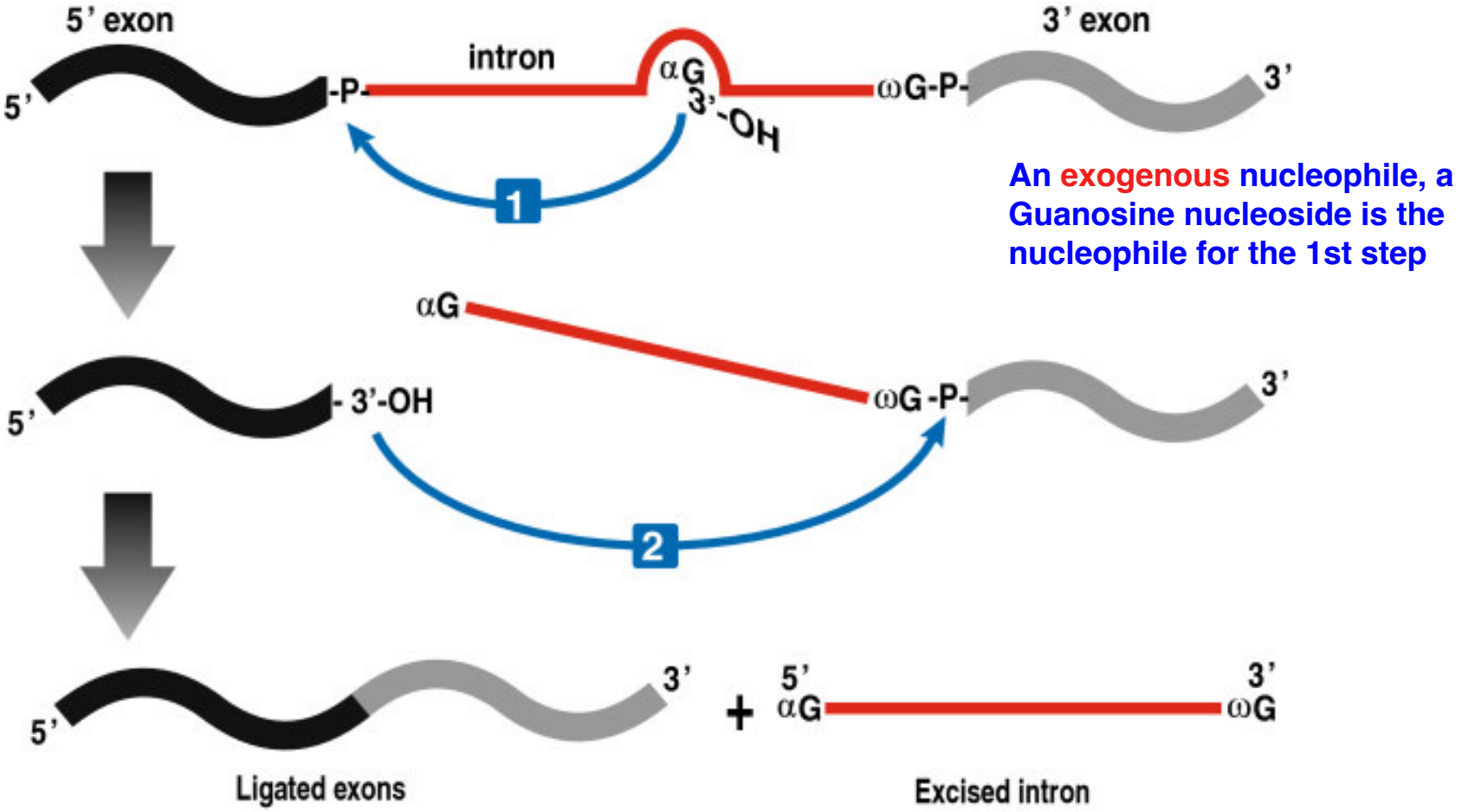
Prize share: 1/2

Thomas R. Cech

Prize share: 1/2

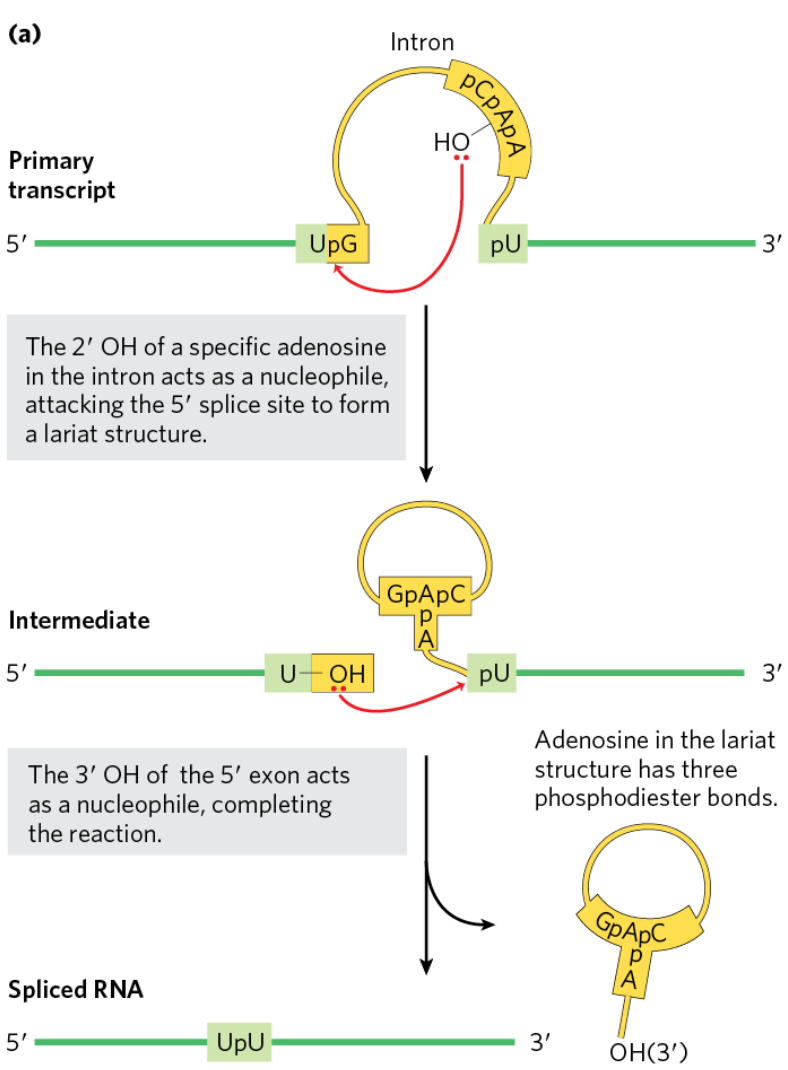
Splicing of some precursor RNAs do not require the spliceosome: the splicing reaction is catalyzed by the RNA intron itself without proteins
= self-splicing = splicing ribozymes

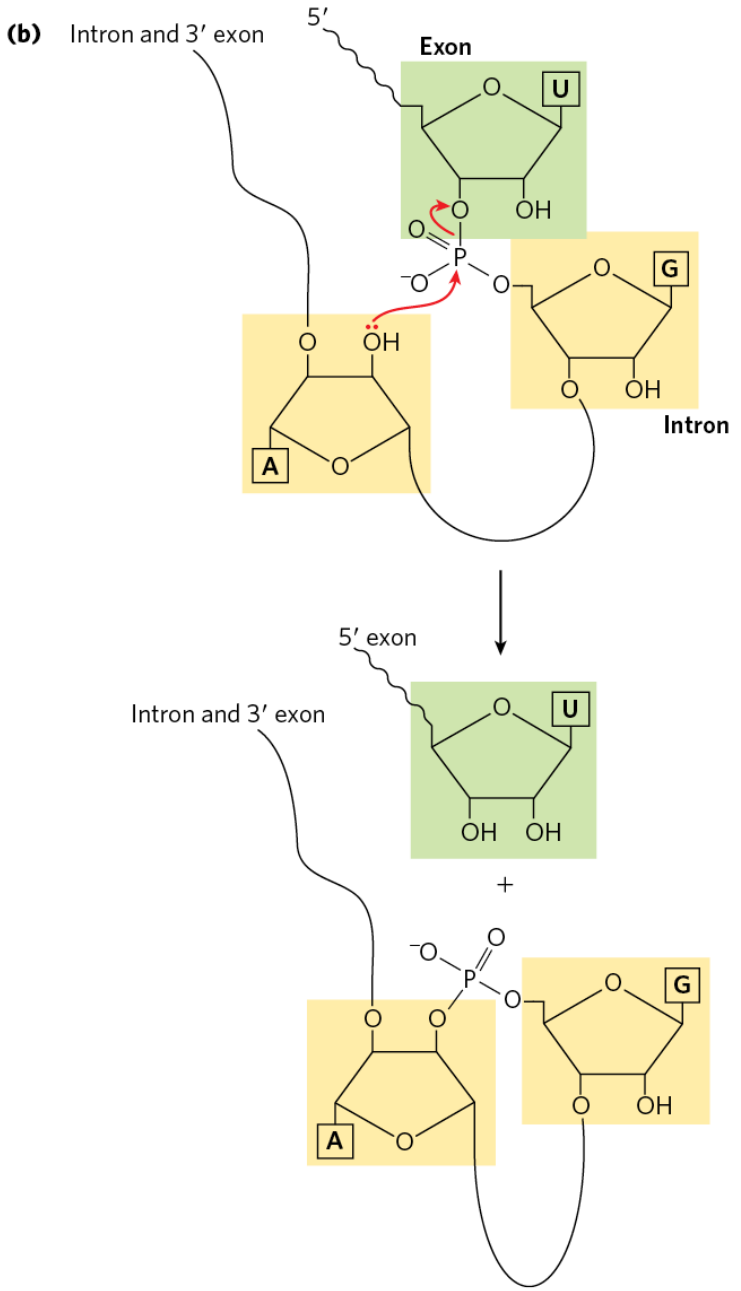
Group I intron self-splicing mechanism



Hausner, Mob DNA, 2014.

Group II intron self-splicing mechanism





Why do most eukaryotic introns require the spliceosome rather than self-splicing like Group I or Group II introns?

A: Eukaryotic introns are too short to self-splice

B: Eukaryotic introns lack the catalytic RNA structures required for self-splicing

C: The splicesome provides the energy required for RNA cleavage

D: Self-splicing reactions require DNA

Group I Ribozymes catalyze splicing reactions by forming a Complex 3D structure -like protein enzymes-

- **Monovalent and Divalent Cations required to prevent electrostatic repulsion between strands**

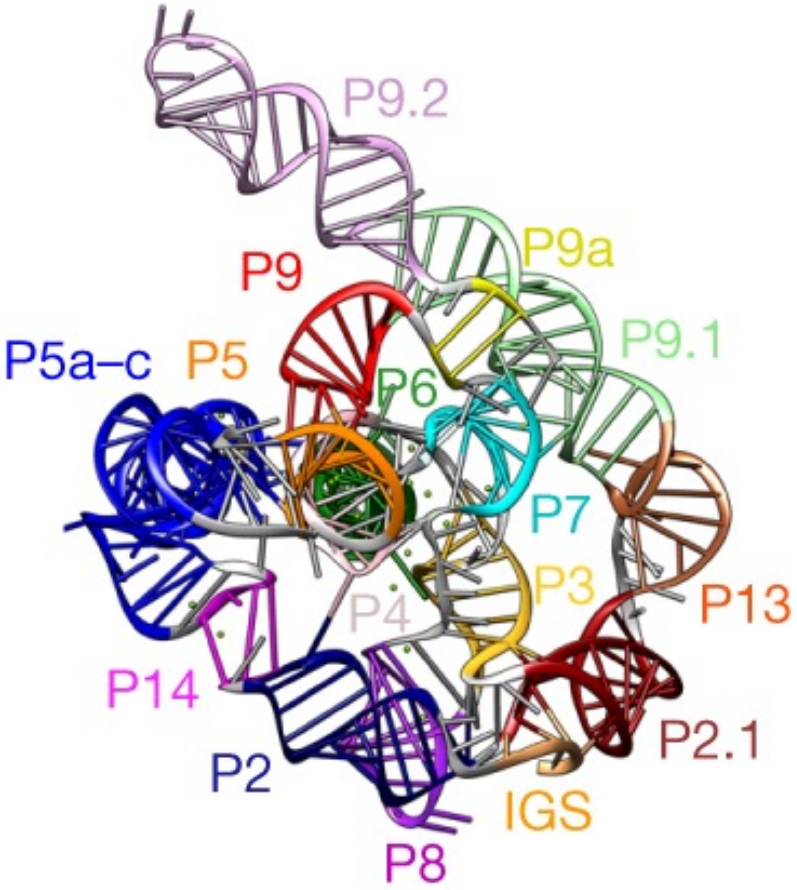
(Purple Sphere = K+)

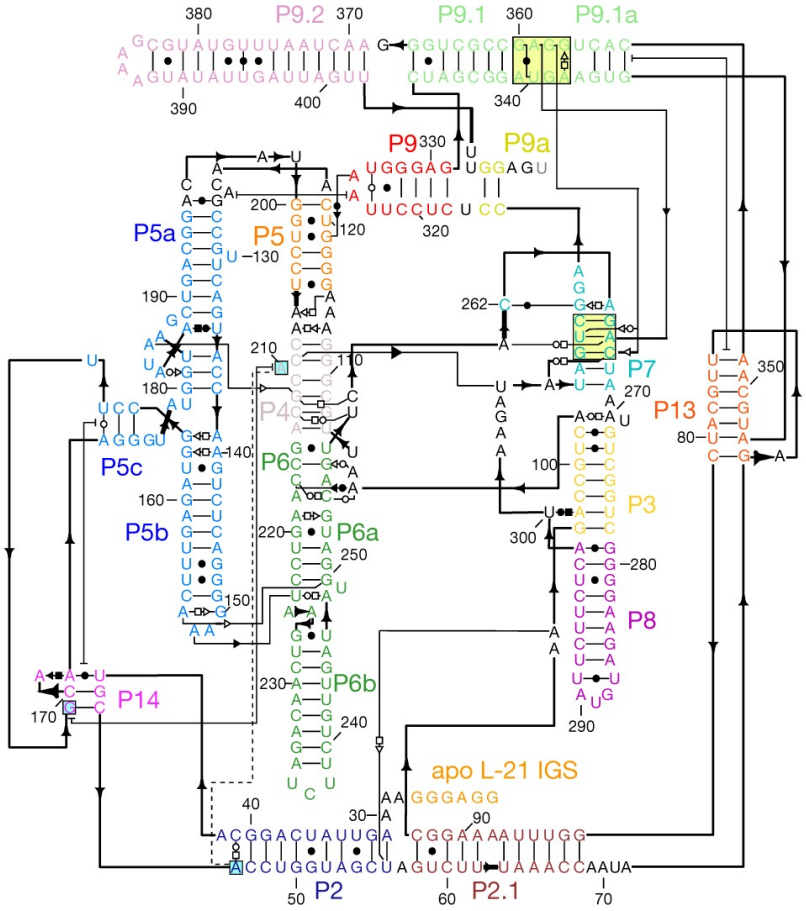
(Green Spheres = Mg^{++})

Also see tertiary interactions from nucleic acids structures chapter

- **Ribose Zippers**
- **Pseudoknots**
- **A-minor interactions**

The complete structure of the Tetrahymena group I self-splicing intron





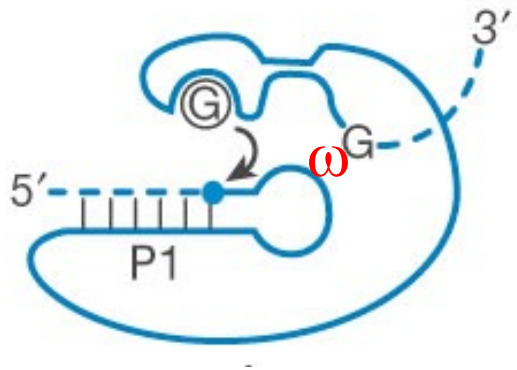
Su*, Zhang*, Kappel*, et al., Nature, 2021.

Self-splicing of group I Ribozymes: major steps

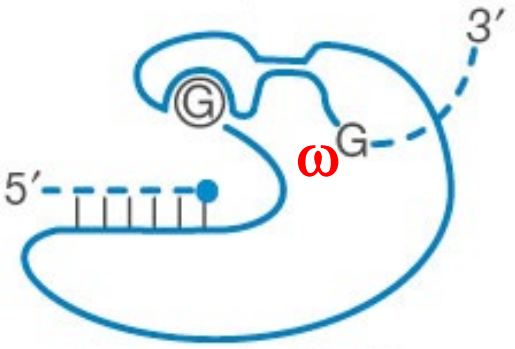
G = Guanosine cofactor

ωG = last intronic nucleotide

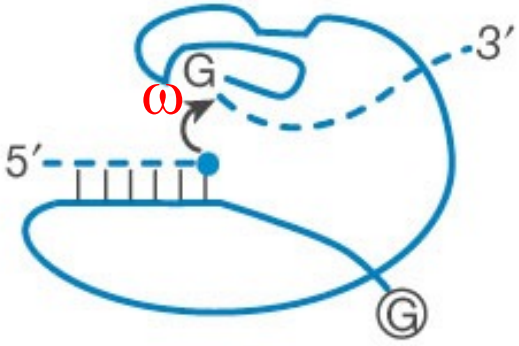
1-Intron folding; formation of the P1 RNA duplex defines the 5´ splice site; Binding of Guanosine cofactor in the G binding site



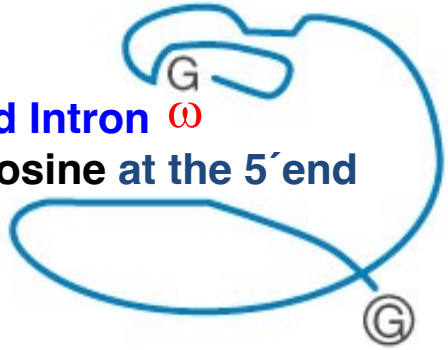
2- First Transesterification



3- Conformational Change: Binding of the ωG In the G binding site



Second Transesterification

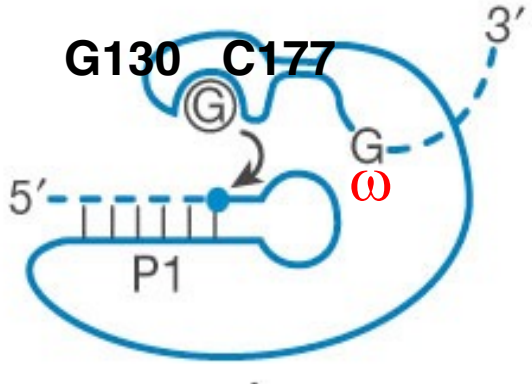


Excised Intron ω
+ guanosine at the 5´end

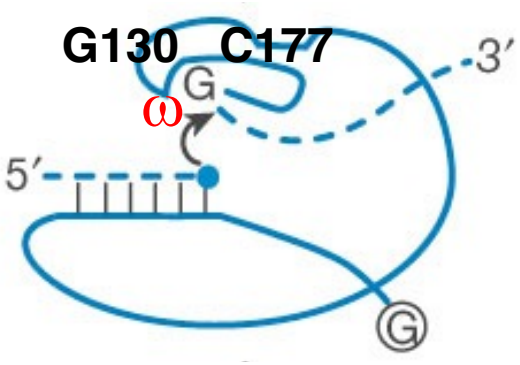
Spliced Exons

The ribozyme uses a single active site that catalyzes both reactions

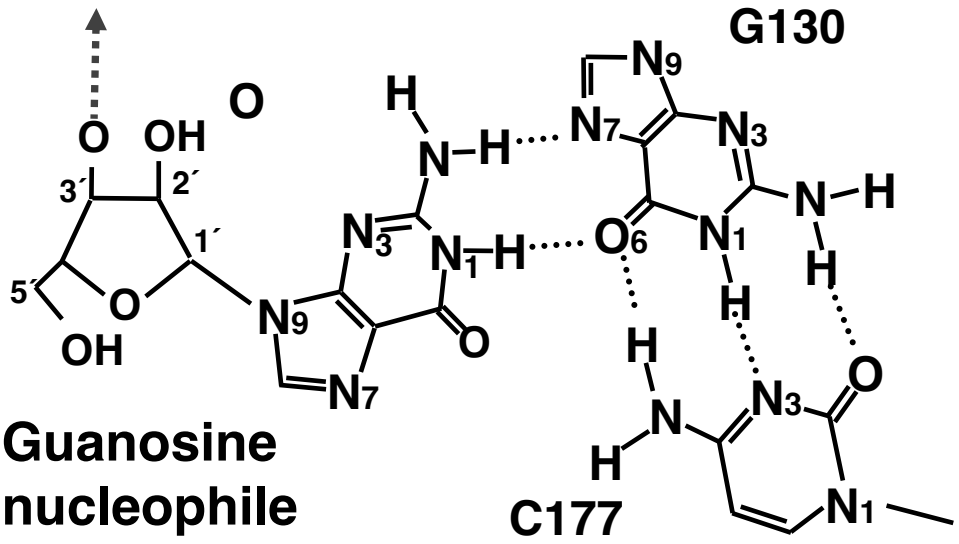
Binding of the G nucleosides (G cofactor and ωG of the 3´ splice site) by a single binding site in the ribozyme structure



G-binding site =G130-C177 intronic base pair

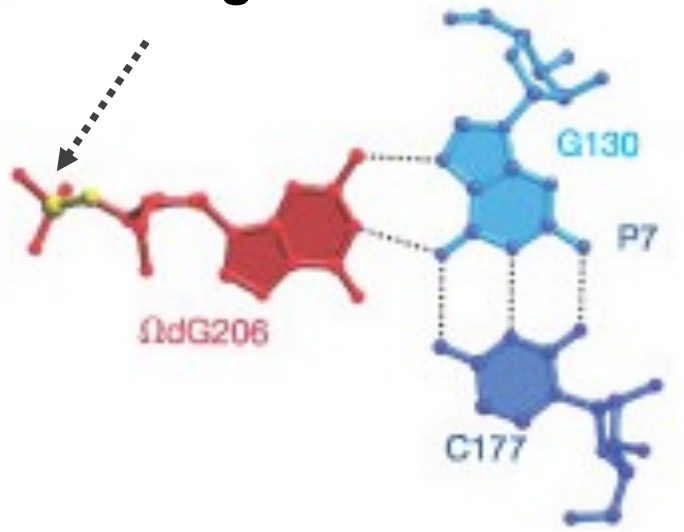


5´ splice site cleavage



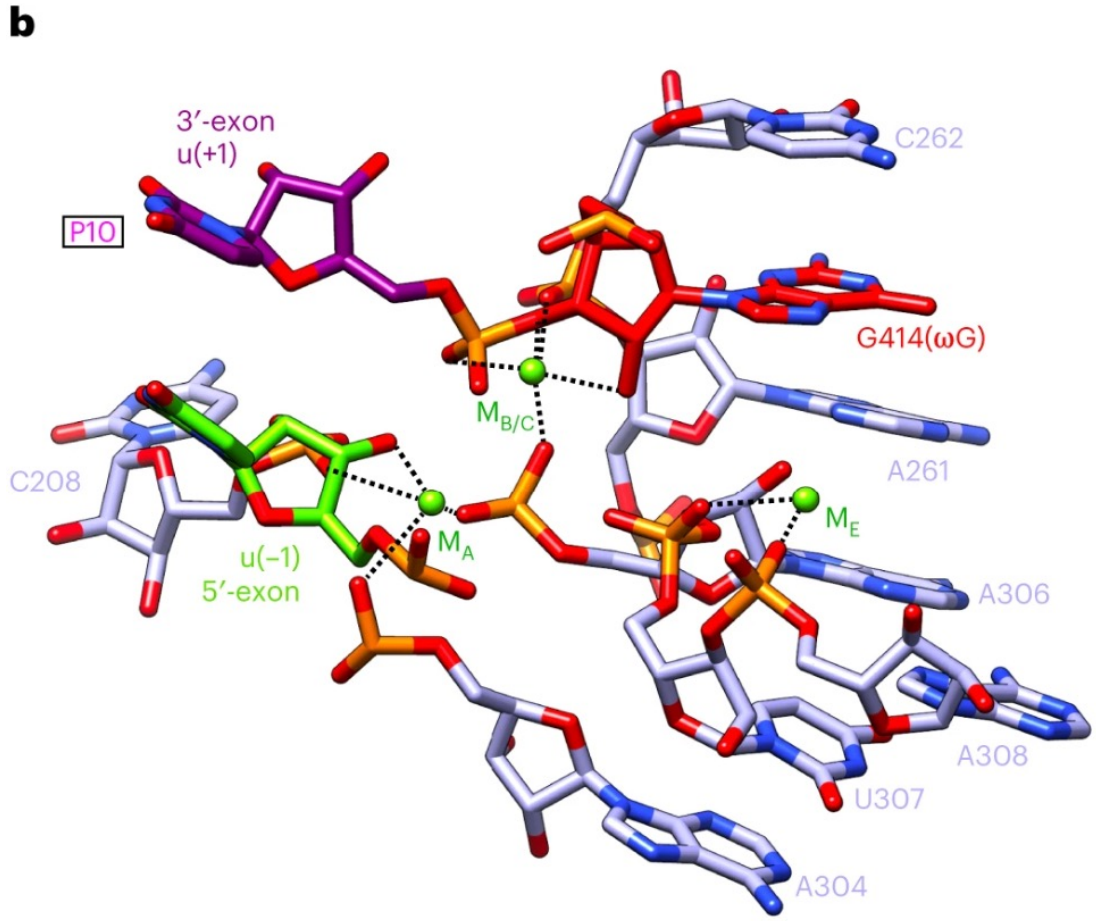
Guanosine nucleophile

3´splice site cleavage



The G-binding site
binds both the guanosine cofactor and the ωG
during the successive steps ->structural evidence for a single active site

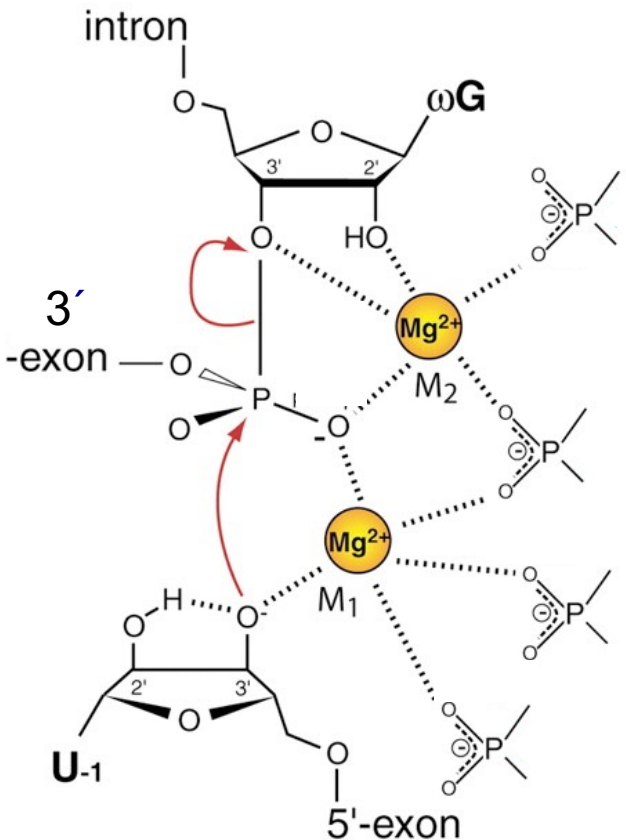
Metal ions are found close to the reactive substrates in Group I introns



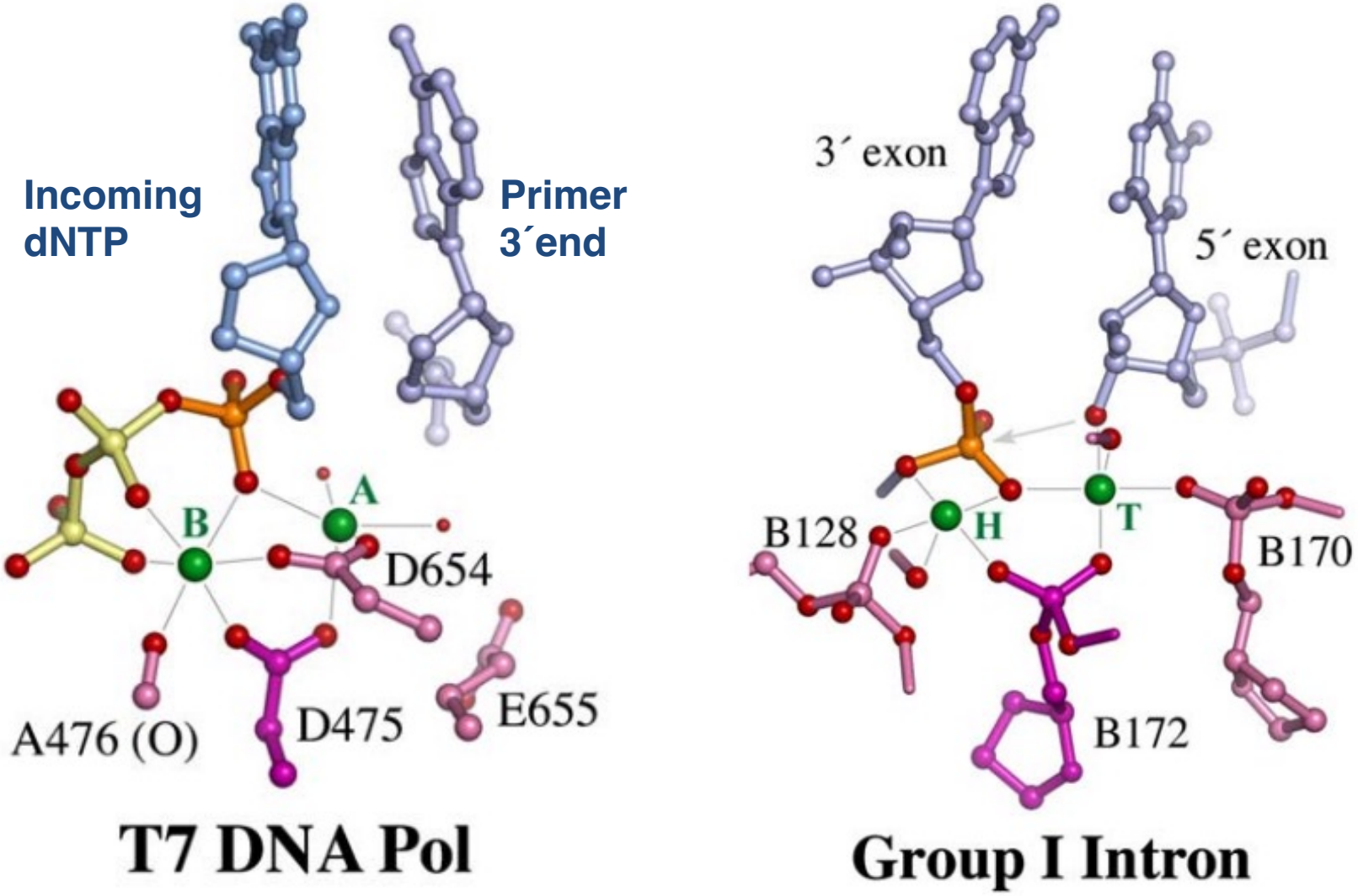
CryoEM Structure

Luo et al. Nature Catalysis 2023

Catalysis Model (2nd step)

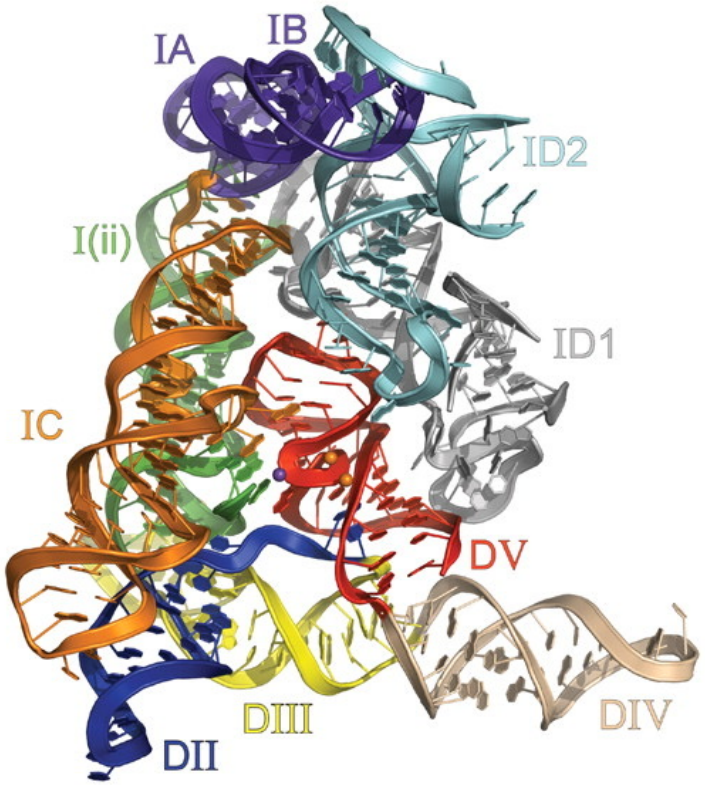


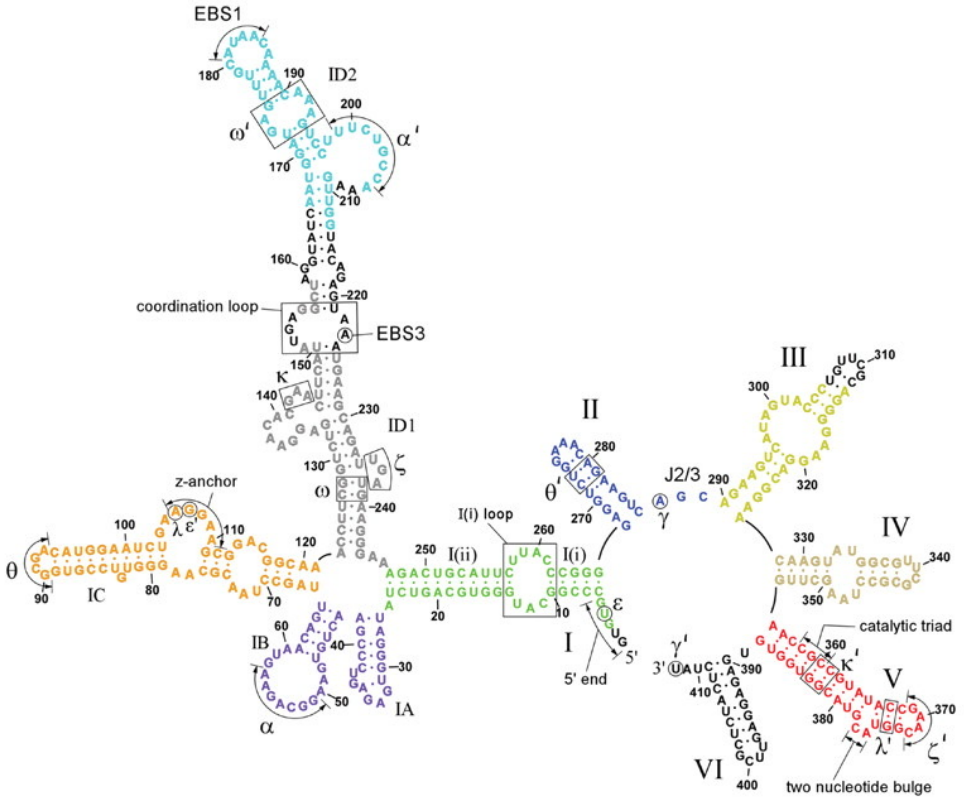
Two metal ion catalysis model in Group I introns similar to the Catalytic mechanism of DNA/RNA Polymerases



Yang et al. Mol.Cell 2006

Group II introns also have a complex tertiary fold

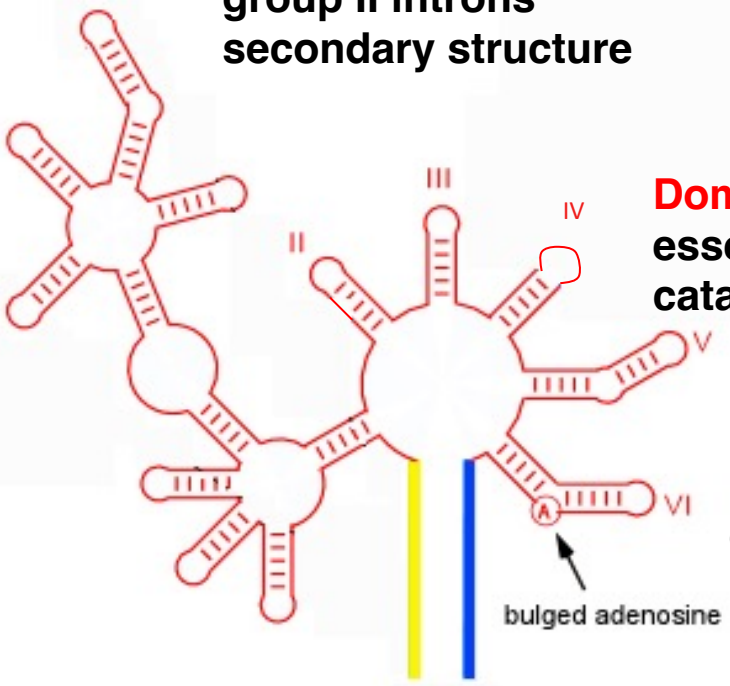




Toor et al., Science, 2008.

Self-splicing group II introns ribozymes and similarities with spliceosome-mediated splicing

group II introns
secondary structure



Domain V: essential for catalysis

Domain VI: carries the branchpoint bulged A

+monovalent cations
+Mg^{++}

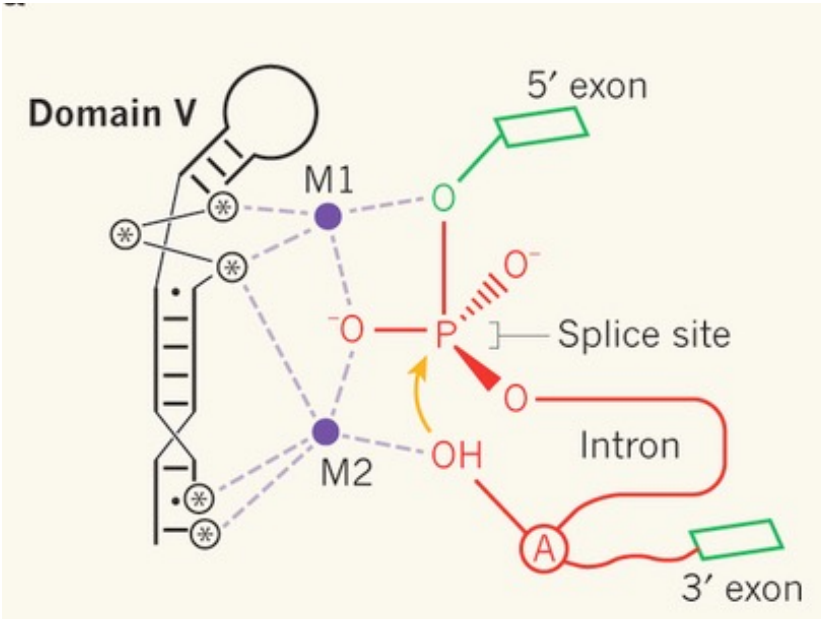
↓

Tertiary Structure
Ribozyme activity

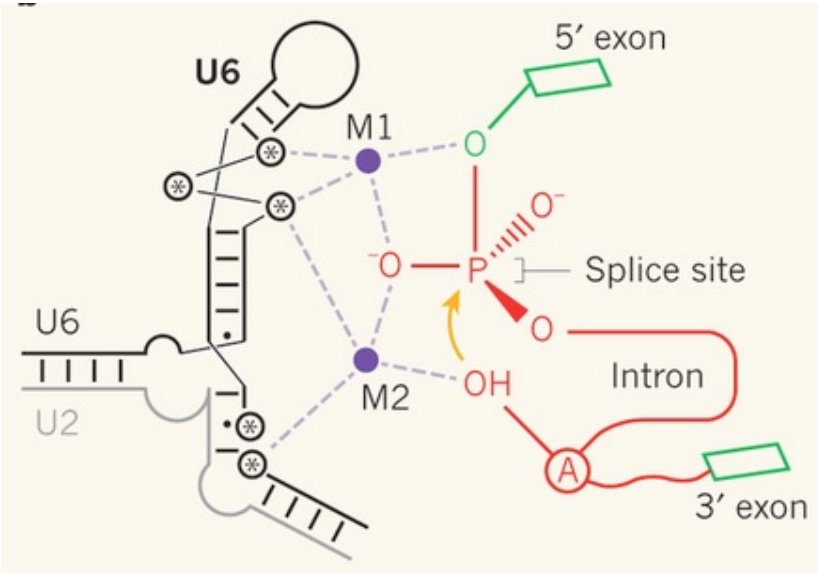
↓

Self-splicing in the absence of proteins or snRNAs

<u>The two splicing steps are *identical* to those of nuclear splicing</u>

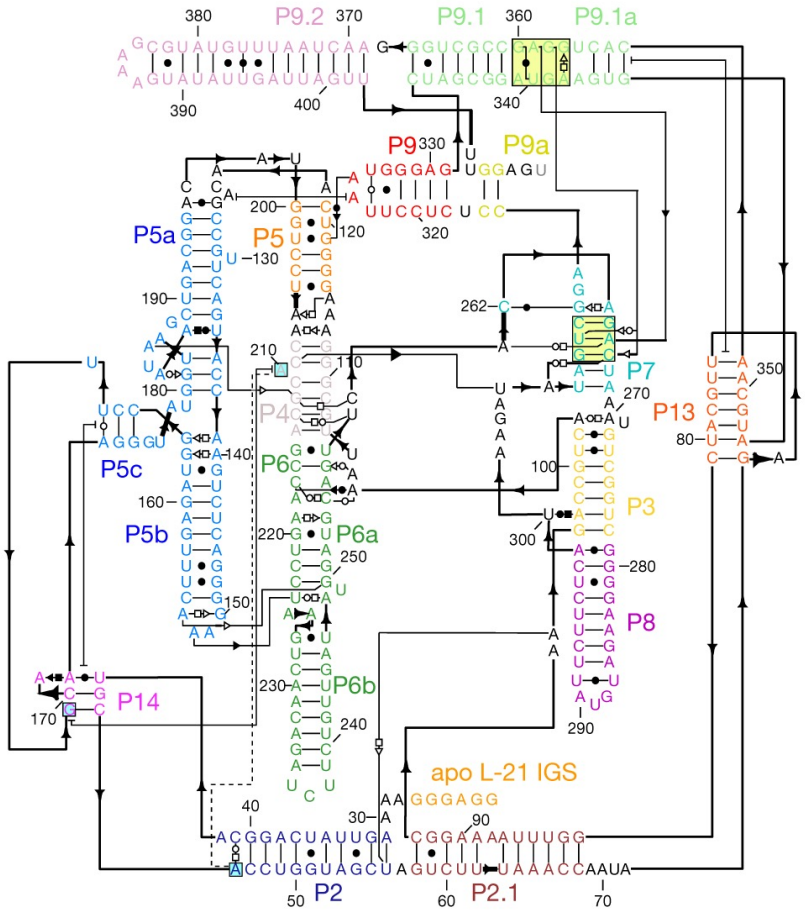


Domain V binds two divalent cations involved in catalysis -> same 2-metal ion catalysis mechanism shown for the spliceosome



Shared features between ribozymes and protein enzymes

- They can vary dramatically in size
- 3D structure is important for function
- Ribozymes can be inactivated by heating above their melting temperature
- Mutations can affect folding and catalytic activity

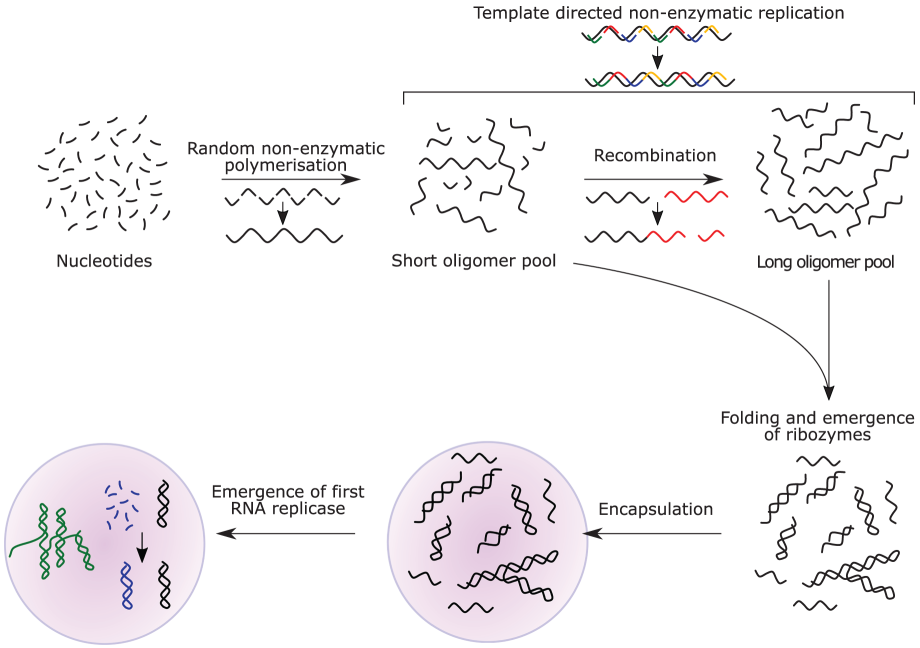


The RNA world hypothesis

How might a self-replicating polymer come to be?

<u>RNA world hypothesis</u>: life on Earth originated with RNA molecules capable of storing and replicating genetic information in addition to catalyzing biochemical reactions

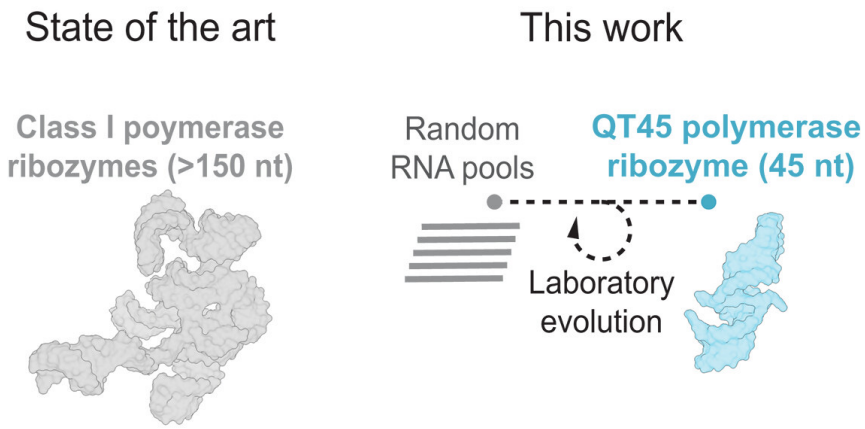
- **Proposed in the 1960s by Carl Woese, Francis Crick, and Leslie Orgel**
- **Key evidence:**
  - **Existence and diversity of catalytic RNAs**
  - **Structure of the ribosome (see next Unit!)**



A small polymerase ribozyme that can synthesize itself and its complementary strand

EDOARDO GIANNI, SAMANTHA L. Y. KWOK, CHRISTOPHER J. K. WAN, KEVIN GOEIJ, BRYCE E. CLIFTON, ENRICO S. COLIZZI, JAMES ATTWATER, AND PHILIPP HOLLIGER Authors Info & Affiliations

De novo evolution of a small polymerase ribozyme



Synthesis of itself and its template