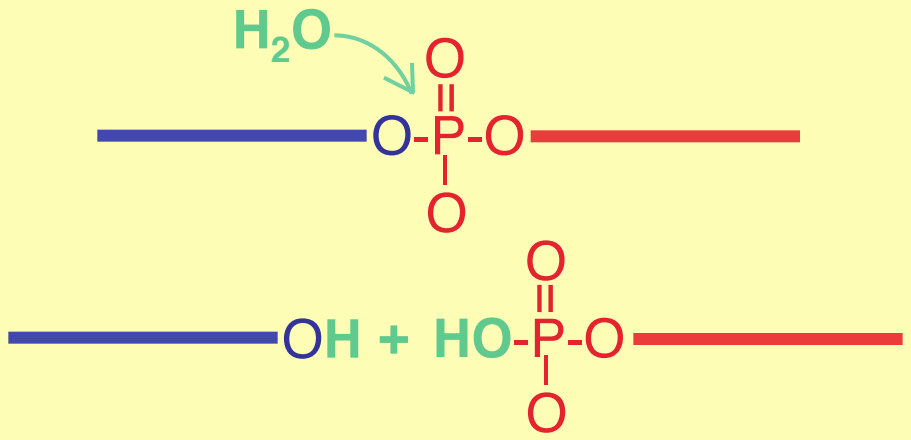
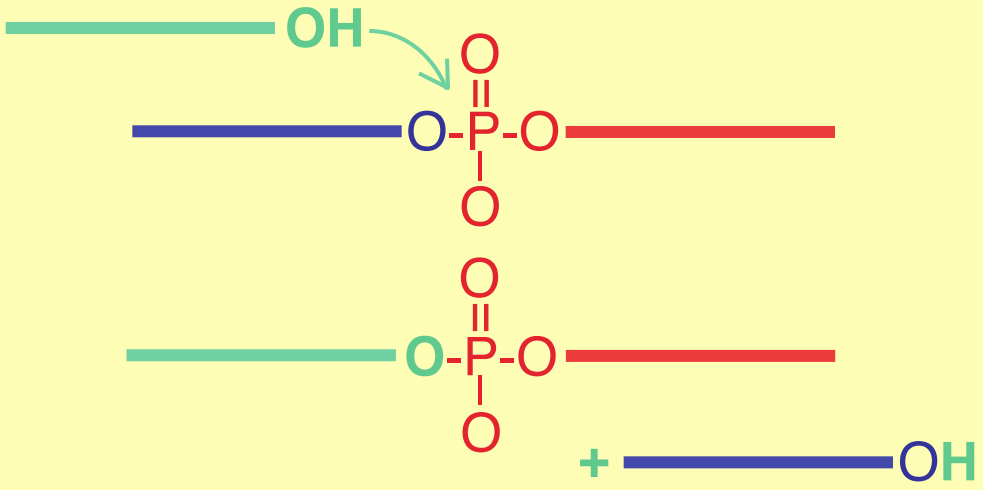

Alan Weiner
BIOCHEM 530
Wednesday, MEB 248
October 28, 2015
RNA splicing,
RNA interference,
miRNAs and lncRNAs,
RNA therapeutics

All natural ribozymes (*except the ribosome!*) use the same mechanism: the attacking hydroxyl group expels the leaving hydroxyl group, making a new phosphoester bond as it breaks the old one

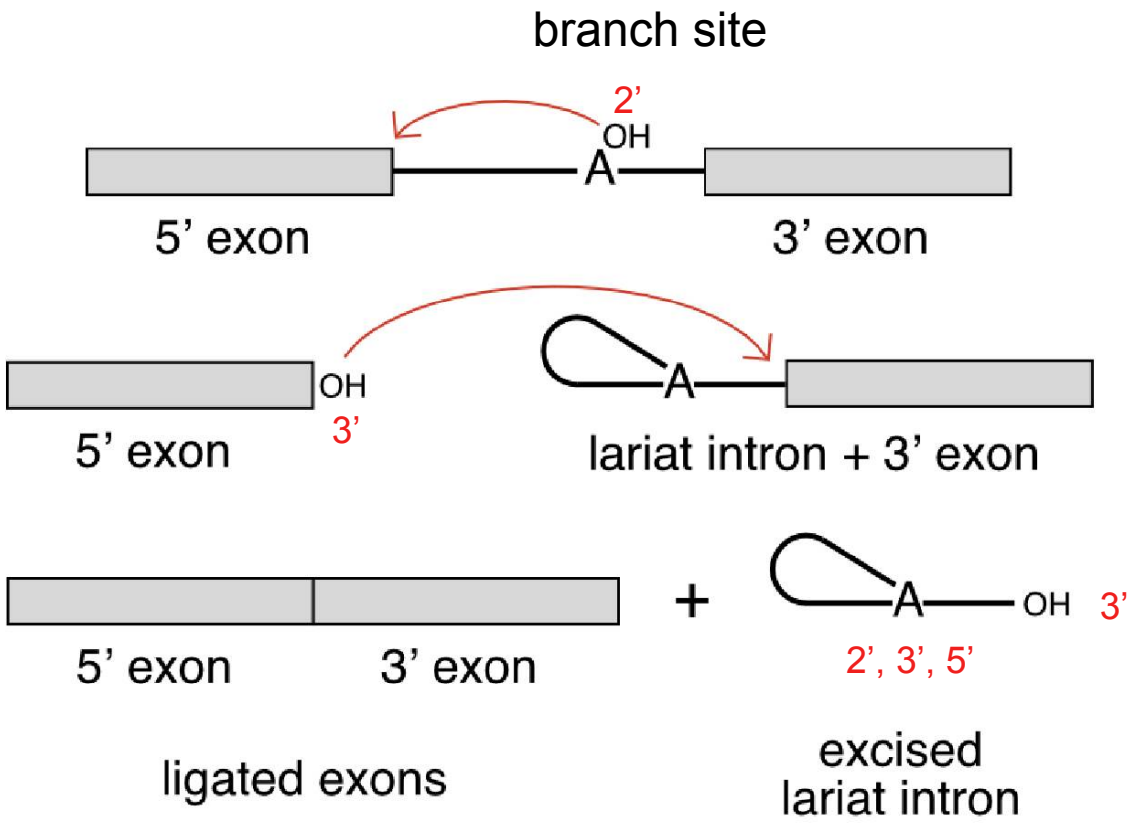
If the attacking hydroxyl belongs to H_2O, the RNA is hydrolyzed (cleaved).



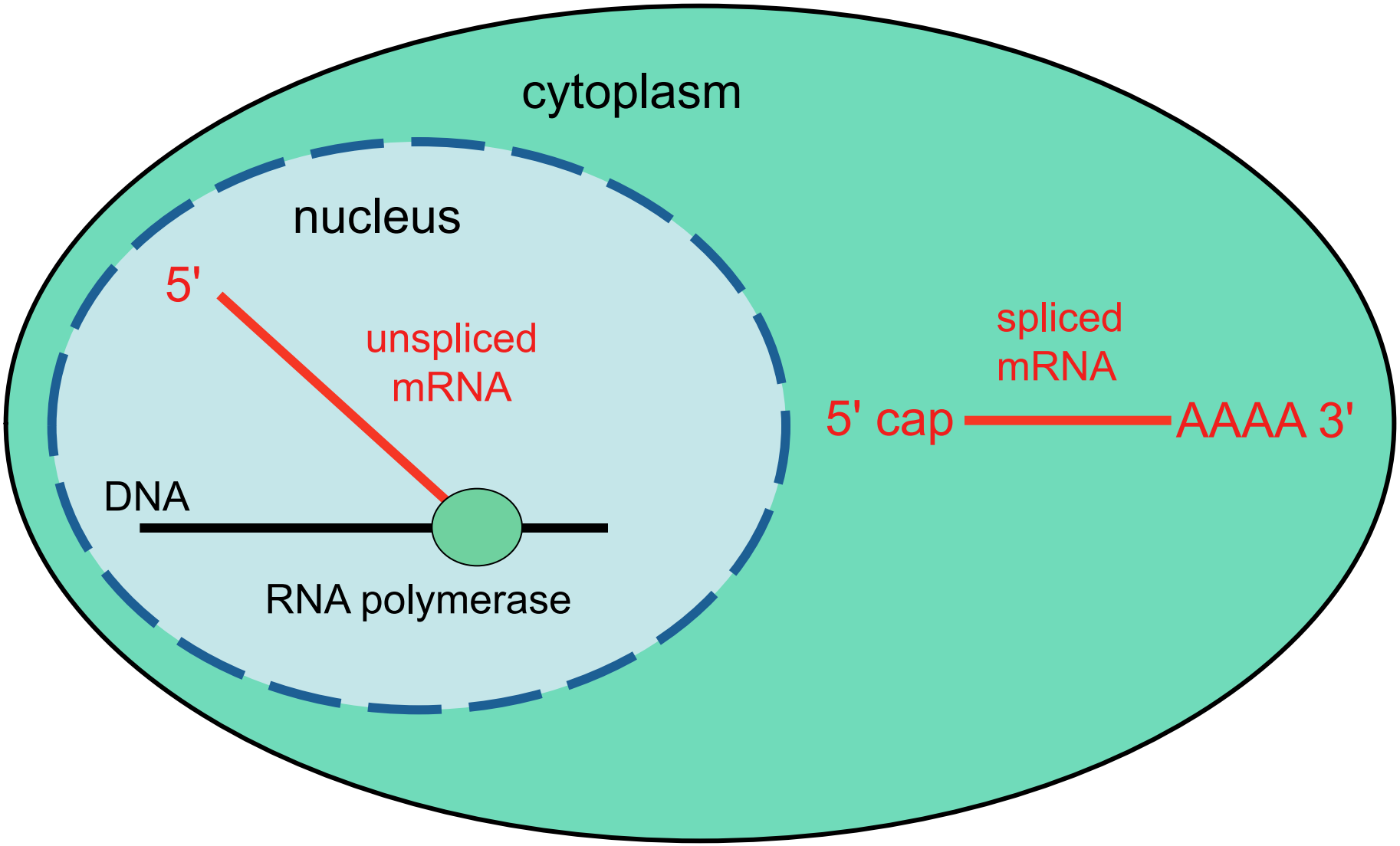
If the attacking 2' or 3' hydroxyl belongs to RNA, the RNAs are isomerized.



mRNA splicing is just an isomerization…
a new bond is made for every bond broken

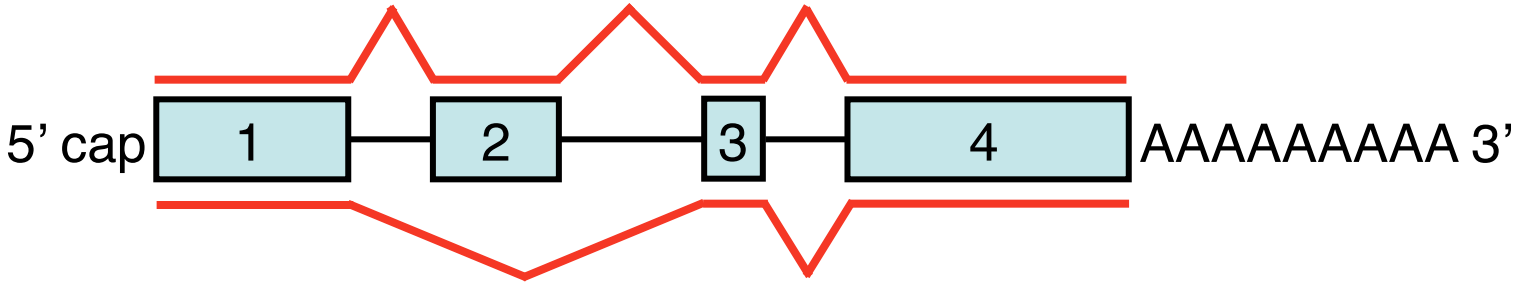


mRNA life history is regulated at every step
from synthesis to destruction

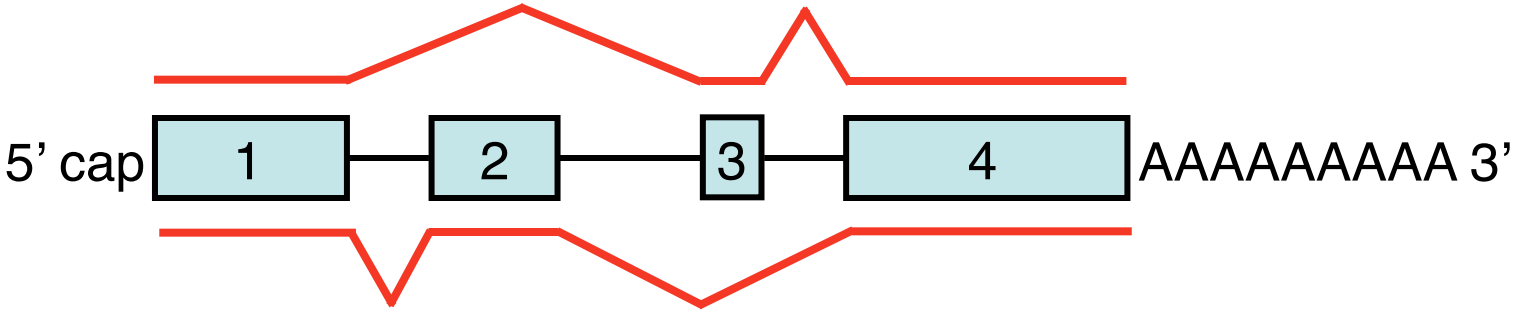


Flavors of alternative mRNA splicing

default splicing pattern

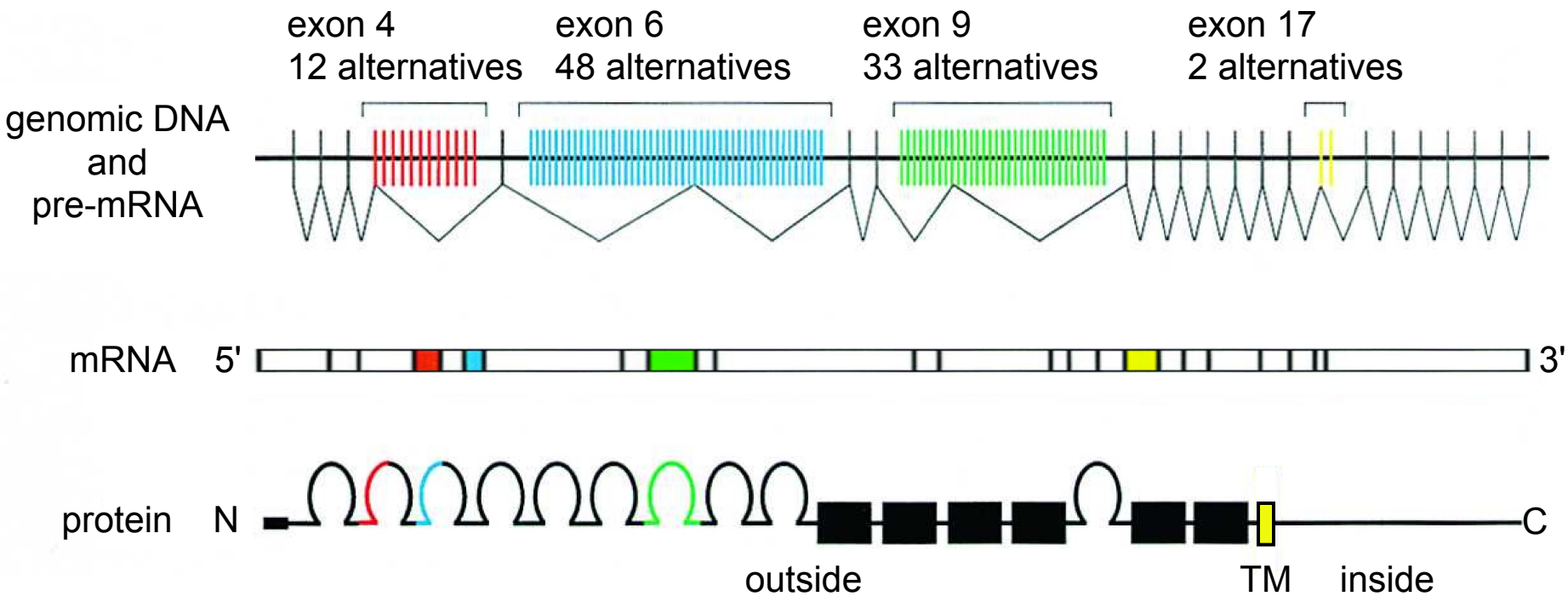


exon skipping or exclusion (+/– exon 2)



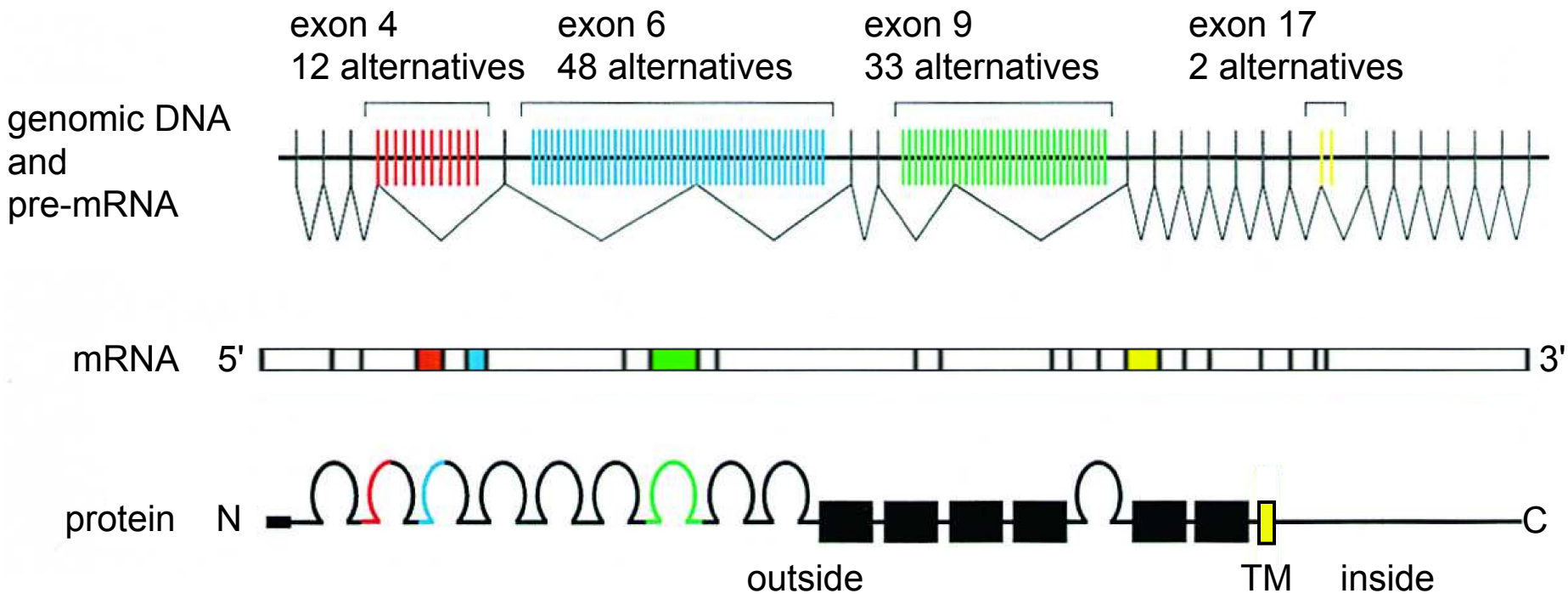
alternative exons (include exon 2 or 3)

Current world record holder for alternative mRNA splicing is the Drosophila DSCAM gene, whose protein products function in axon guidance and innate immunity in the fly



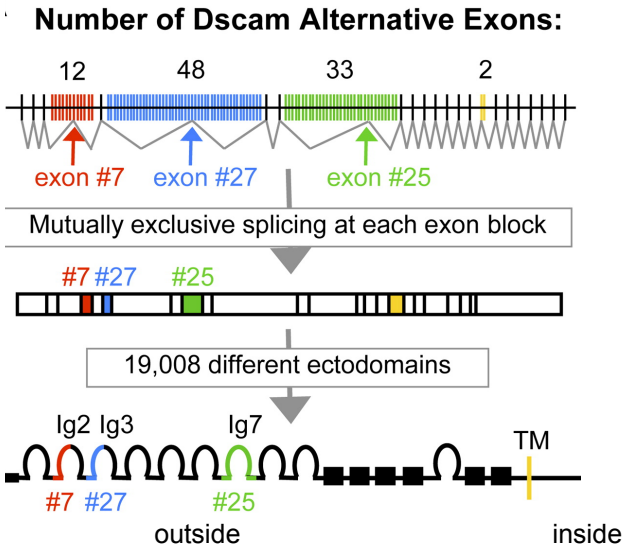
The 61 kb DSCAM gene generates an 8 kb mRNA containing 24 exons. Exons 4, 6, 9, and 17 are encoded as tandem arrays of mutually exclusive alternative exons, so this one gene could in principle generate as many as 12 × 48 × 33 × 2 or 38,016 different mRNAs and proteins. McManus and Graveley (2011) Curr Op Genet Dev 21, 373-379

Current world record holder for alternative mRNA splicing is the Drosophila DSCAM gene, whose protein products function in axon guidance and innate immunity in the fly

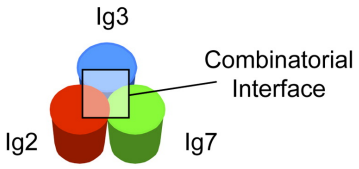


Immunoglobulin folds with alternative regions (color) and constant regions (black): homophilic interactions between similar DSCAM variants play a role in axon guidance; alternative regions may function in innate immunity against microbial pathogens

Schmucker et al. (2000) Drosophila Dscam Is an Axon Guidance Receptor Exhibiting Extraordinary Molecular Diversity. Cell 101, 671-684; Science 344, 1182-1186



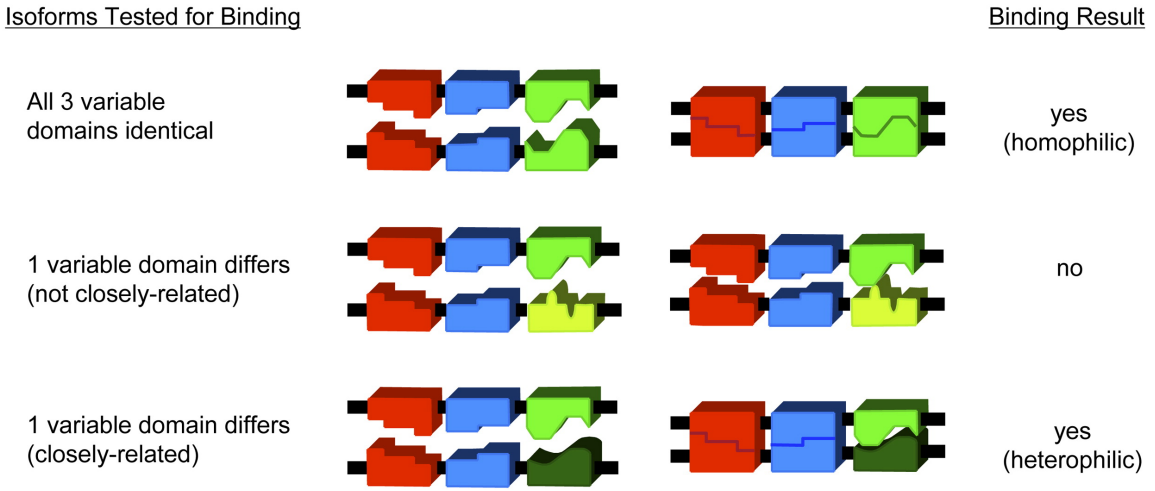
New answers pose new questions. We assumed that transcriptional control was the major determinant of complexity, and we asked ***What wires transcription?*** Now it turns out that nerves are wired by alternative mRNA splicing, and so we must now ask ***What wires alternative splicing?*** Maybe not an endless hall of mirrors, but certainly no end in sight...



Context-Dependent Model: ***one interface*** formed jointly by 3 variable Ig domains.

Modular Model: a ***separate interface*** at each of the 3 variable Ig domains.

Hattori et al. (2007) Dscam diversity is essential for neuronal wiring and self-recognition. Nature 449, 223-228; Meijers et al. (2007) Structural basis of Dscam isoform specificity. Nature 449, 487-491; Du Pasquier (2005) Insects diversify one molecule to serve two systems. Science 309, 1826-1827.

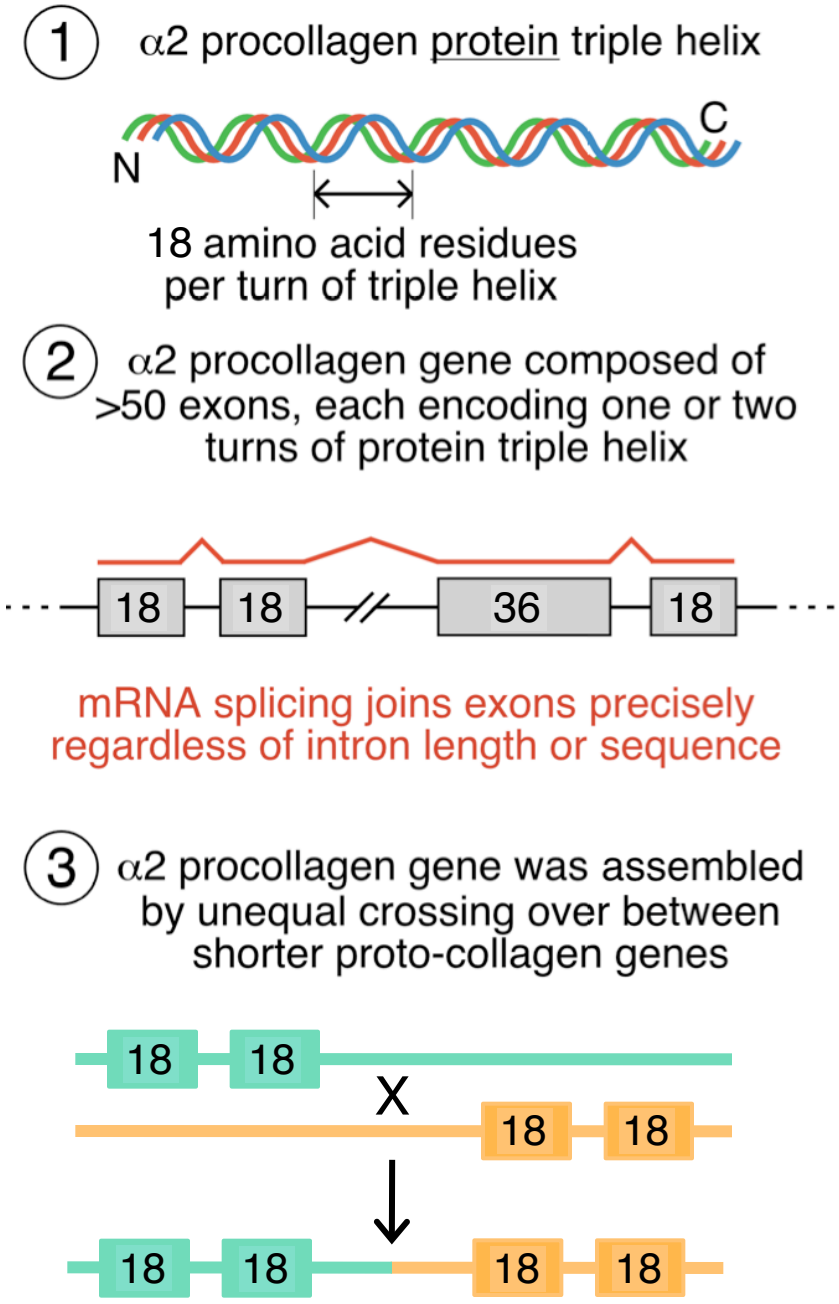


The Exon Theory of Genes: complex genes were assembled from small coding exons, and the assembled exons were then joined together by mRNA splicing.

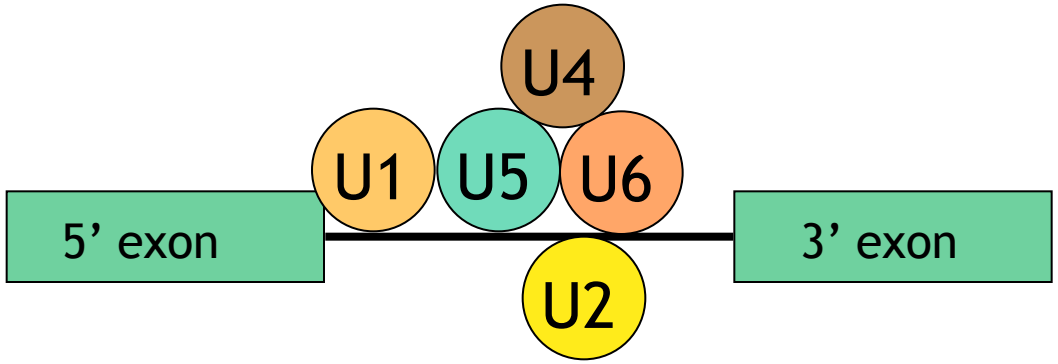
Gilbert, 1986

Accidental DNA recombinations occuring over evolutionary time are random and sloppy…

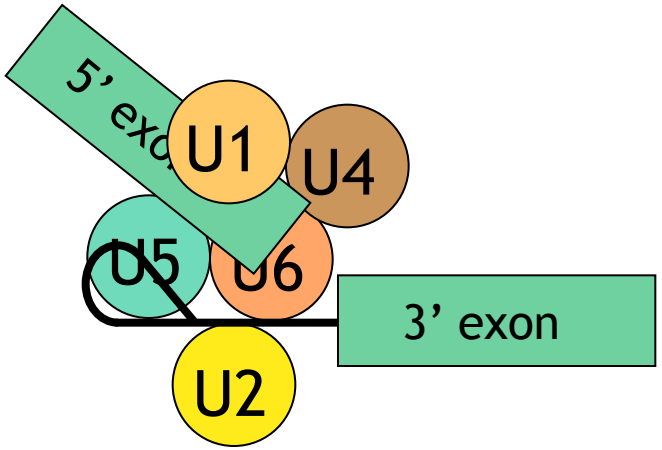
but mRNA splicing is smart, and precisely joins each coding region to the next.



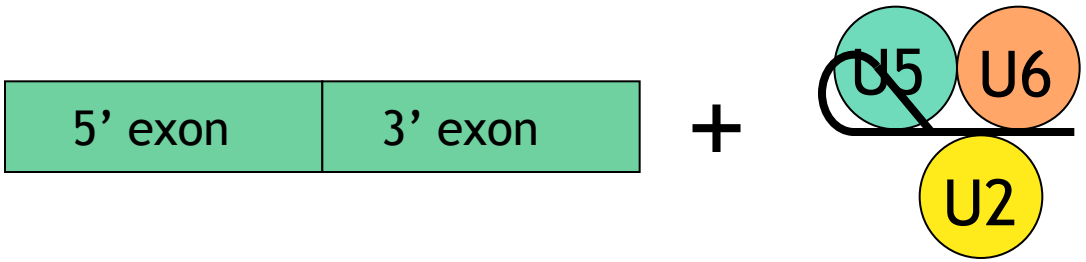
U1, U2, U4, U5, and U6 small nuclear RNA (snRNAs) collaborate to splice out mRNA introns



U snRNPs assemble into a spliceosome on mRNA precursor

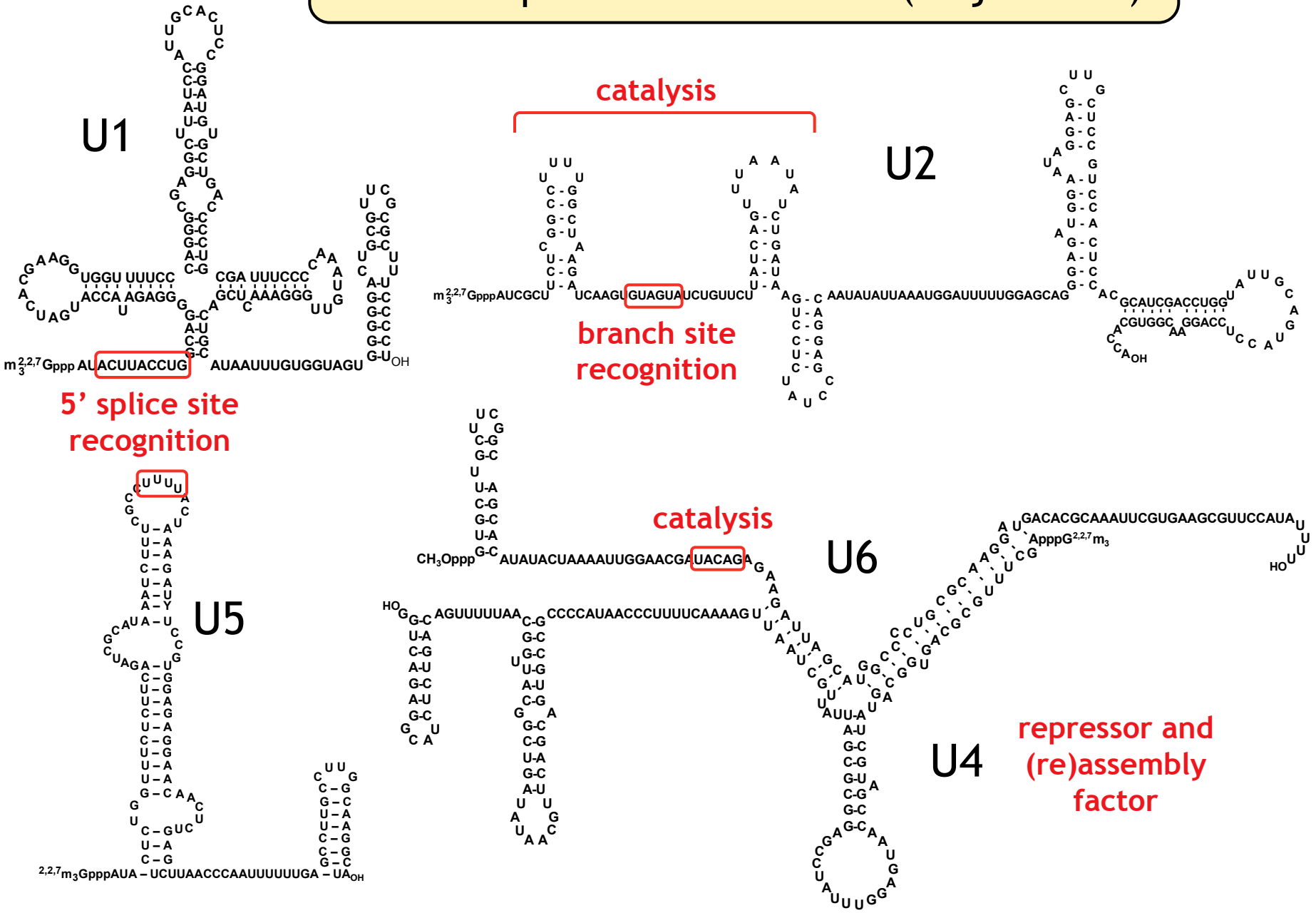


first catalytic step generates an RNA lariat

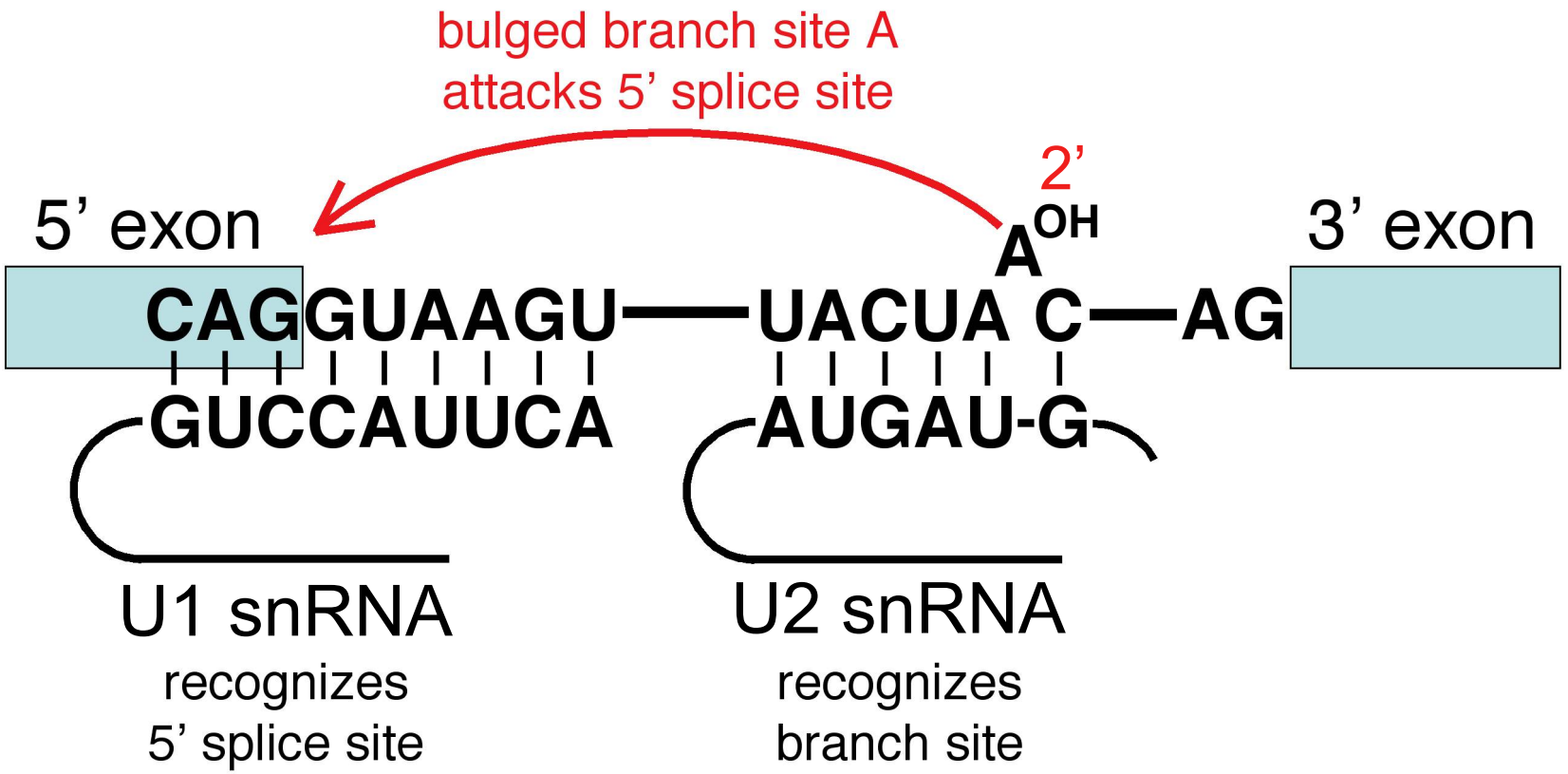


second catalytic step ligates exons and releases lariat intron with U snRNPs

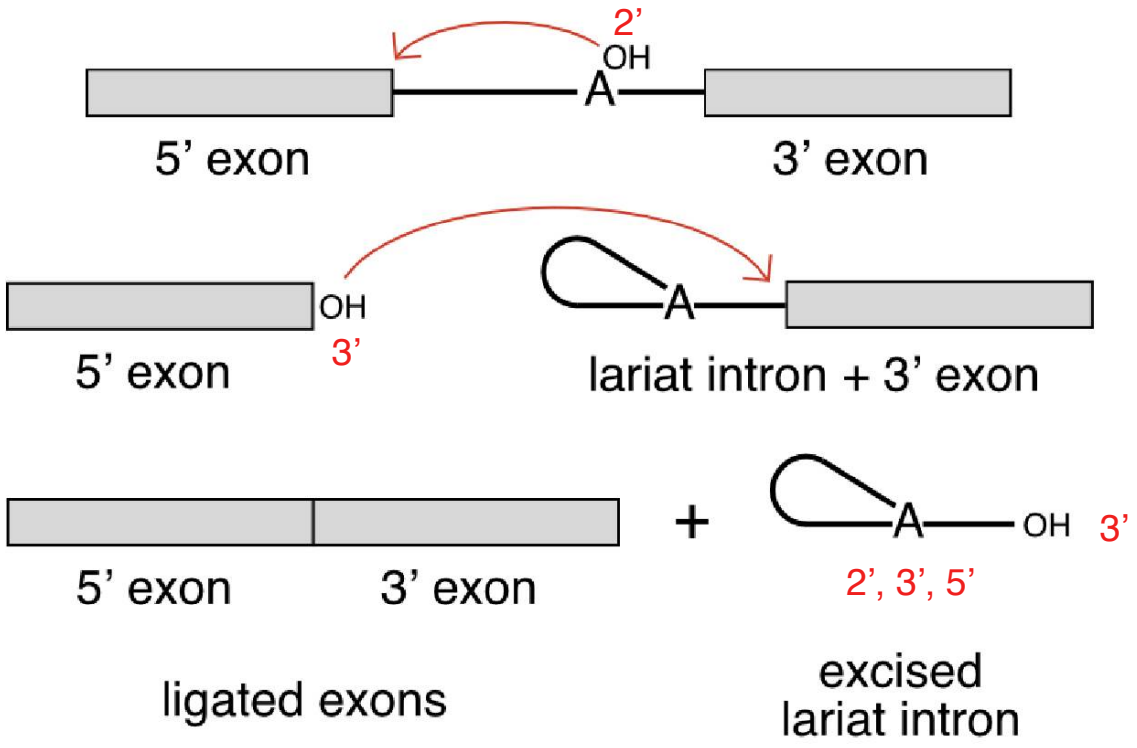
The five spliceosomal snRNAs (major class)



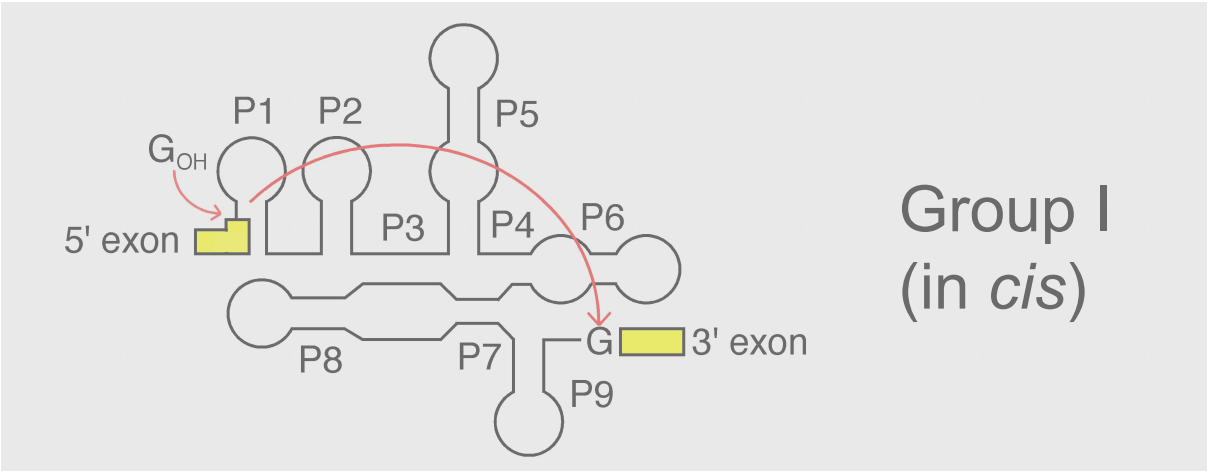
The snRNA components of the snRNPs recognize the 5’ and 3’ splice junctions and catalyze the mRNA splicing reaction



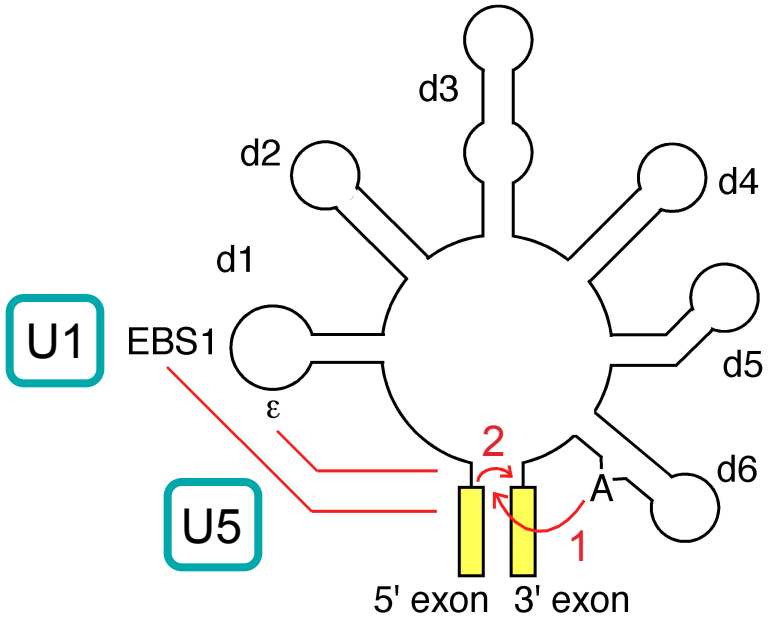
mRNA splicing is mechanistically identical to "Group II" autocatalytic self-splicing



Three kinds of RNA splicing



Group I
(in *cis*)

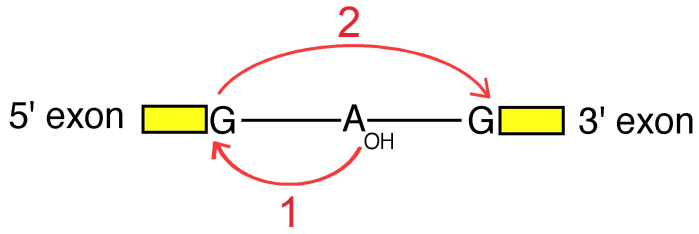


Group II
(in *cis*)

U2/U6

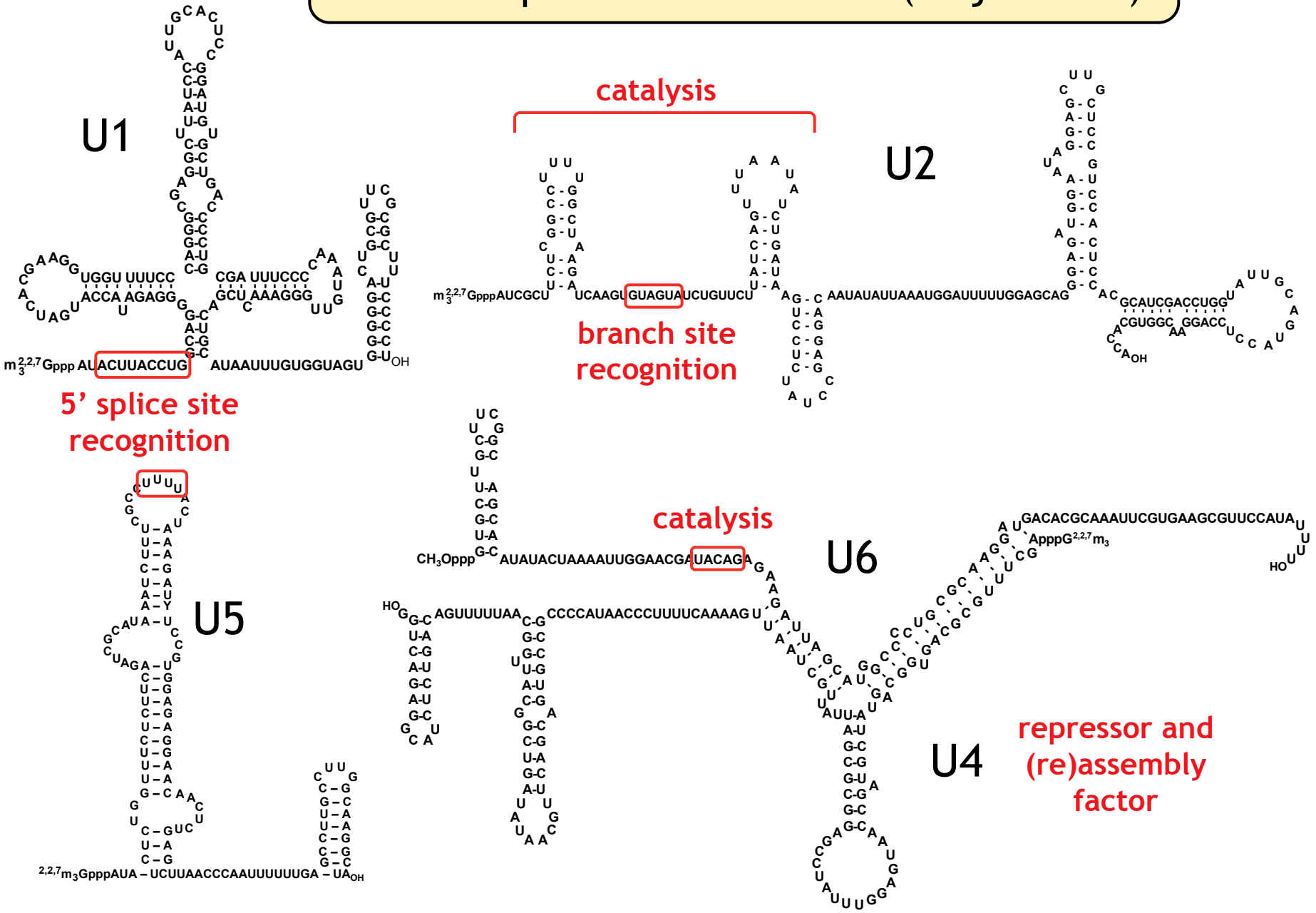
U2/branch site

?

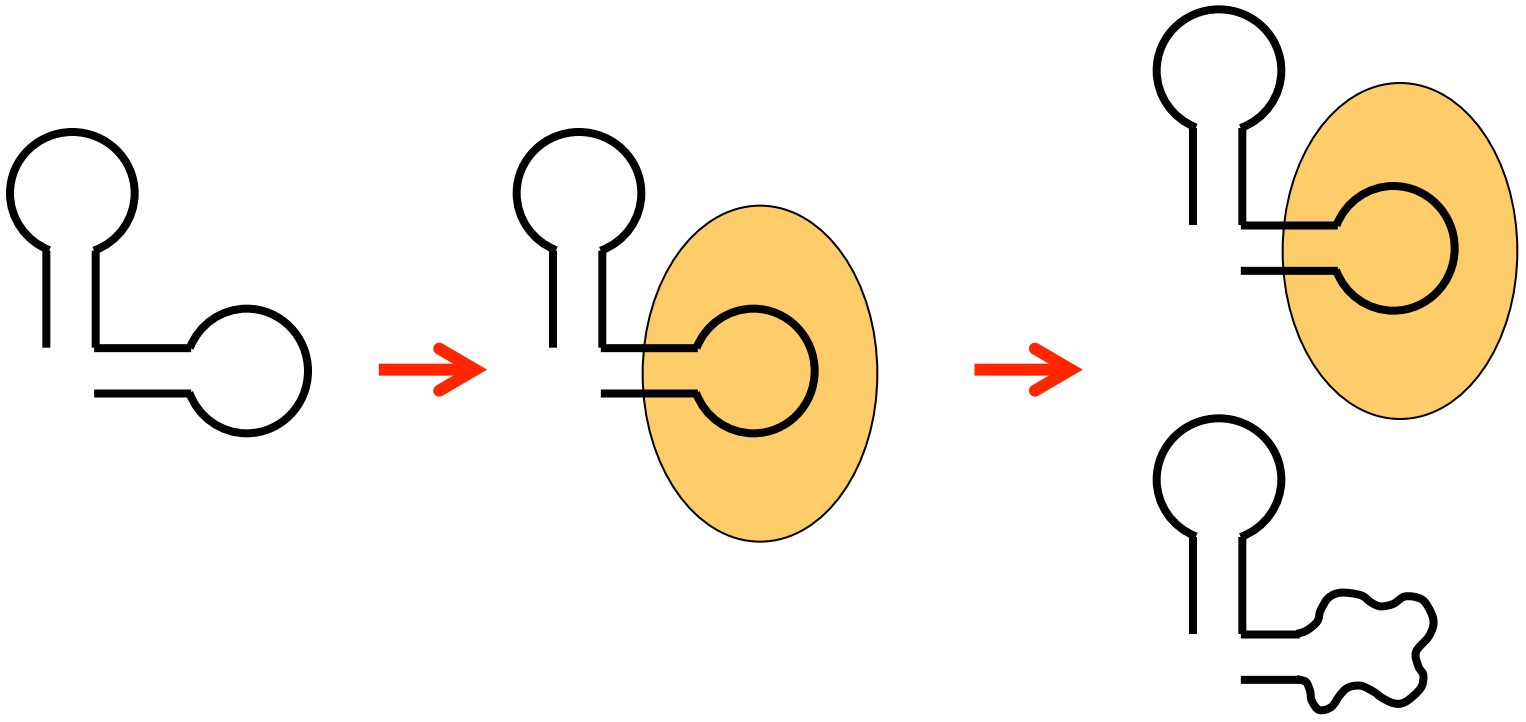


mRNA
(in *trans*)

The five spliceosomal snRNAs (major class)



How the RNP got its protein… an offer it couldn't refuse?



independent RNA

protein makes offer
RNA cannot refuse

RNA relaxes, now
protein-dependent

White HB (1976) Coenzymes as fossils of an earlier metabolic state. *J Mol Evol* 7, 101-104; White HB (1982) Evolution of coenzymes and the origin of pyridine nucleotides. In *The pyridine coenzymes* (ed Everse et al.), pp. 1-17, Academic Press; Alberts BM (1986) The function of the hereditary materials: Biological catalyses reflect the cell's evolutionary history. *Am Zool* 26, 781-796.

RNA is more than a structure or informational sequence…
key steps in the "Central Dogma" are *catalyzed and regulated* by RNAs

mRNA transcription regulated by long noncoding RNAs (**lncRNAs**)

mRNA translation regulated by microRNAs (**miRNAs**) and competing endogenous RNAs (**ceRNAs**)

chromatin structure regulated by small interfering RNAs (**siRNAs** and **piRNAs**)

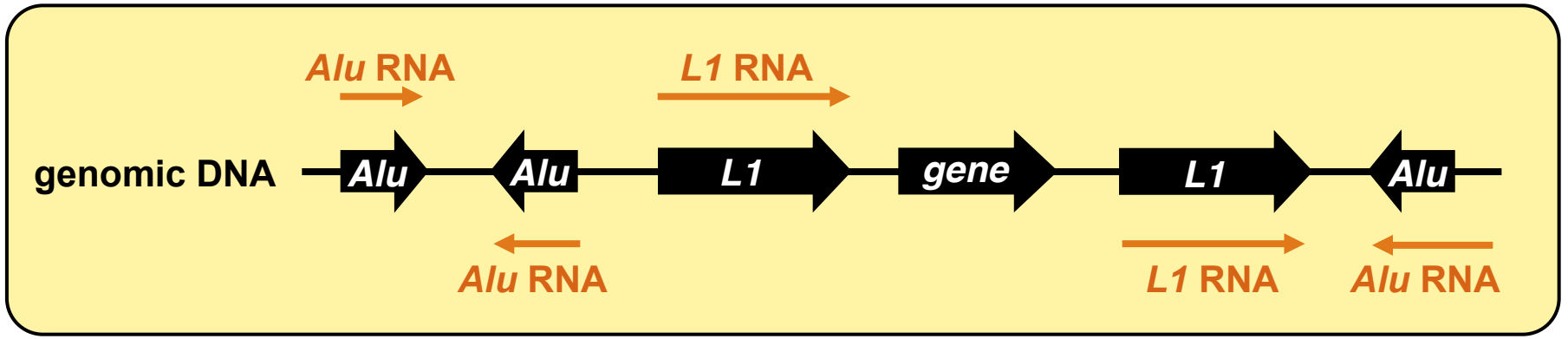
DNA ⟶ mRNA ⟶ protein

mRNA splicing catalyzed by small nuclear RNAs (**snRNAs**)

translation of **mRNAs** catalyzed by ribosomal RNAs (**rRNAs**) and **tRNAs**

defensive RNAs

biosynthetic and regulatory RNAs



The human genome is chock full of two dispersed, highly repeated families of **retroelements** called **L1's** or "**LINEs**" (long interspersed repeated elements) and ***Alu's*** or "**SINEs**" (short interspersed repeated elements). L1 and *Alu* elements both contain promoters, and RNA transcripts of these elements are reverse transcribed into DNA copies which insert into new, almost random sites throughout our genome. L1 elements encode a reverse transcriptase and DNA integrase that enables them to move to new sites, whereas *Alu* elements — which have no protein coding capacity — must borrow the L1 reverse transcriptase and integrase in order to retrotranspose.

The big question: Why aren't our genomes overun or destroyed by this infestation of mobile, parasitic DNA elements? Use RNA to fight RNA!

Surprisingly few diseases are caused by spontaneous retroposition of SINEs and LINEs

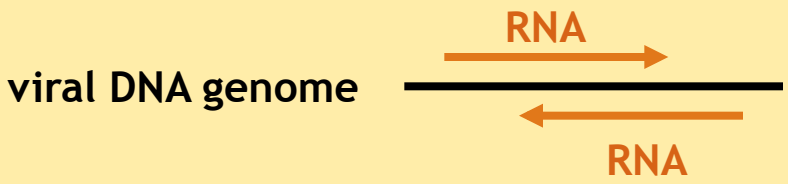
- Although 1,500,000 SINE insertions and 850,000 LINE insertions account for >34% of the complete human genome sequence, SINEs are responsible for only 25 or 0.05% of all known disease-causing mutations, and LINEs for only 5 or 0.01% of the total.
- Most of these 1,500,000 SINE insertions and 850,000 LINE insertions have degenerated in place by neutral mutation and are no longer mobile. The few mobile SINEs and LINEs that have caused spontaneous disease are "young," i.e. they have inserted recently and have not yet had time to degenerate.
- THE BIG QUESTION: What protects our genomes from being blasted to bits by retroposition of SINEs and LINEs? The answer is that ***RNAi is our DNA genome's immune system!***

Callinan and Batzer (2006) Retrotransposable elements and human disease. Genome Dyn 1, 104; Plasterk (2002) RNA Silencing: The Genome's Immune System. Science 296, 1263

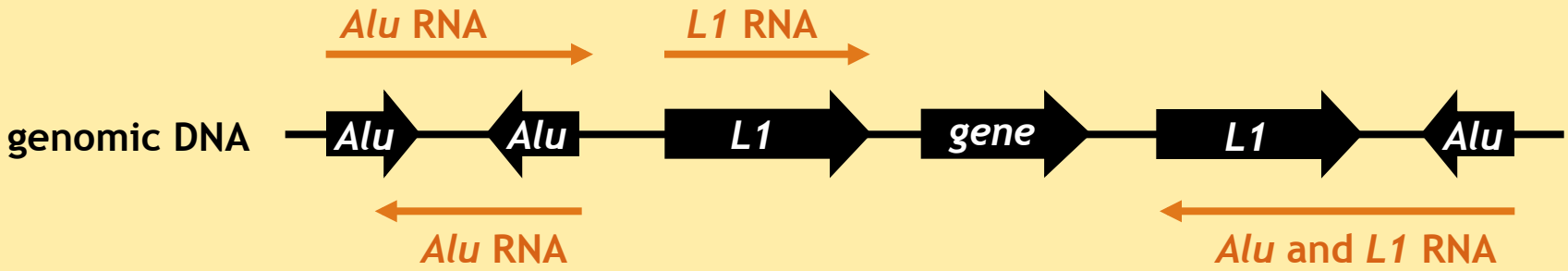
Six easy steps to understanding RNAi

Step One. A perfect RNA duplex is usually a sign of danger.

bidirectional transcription of a compact viral genome

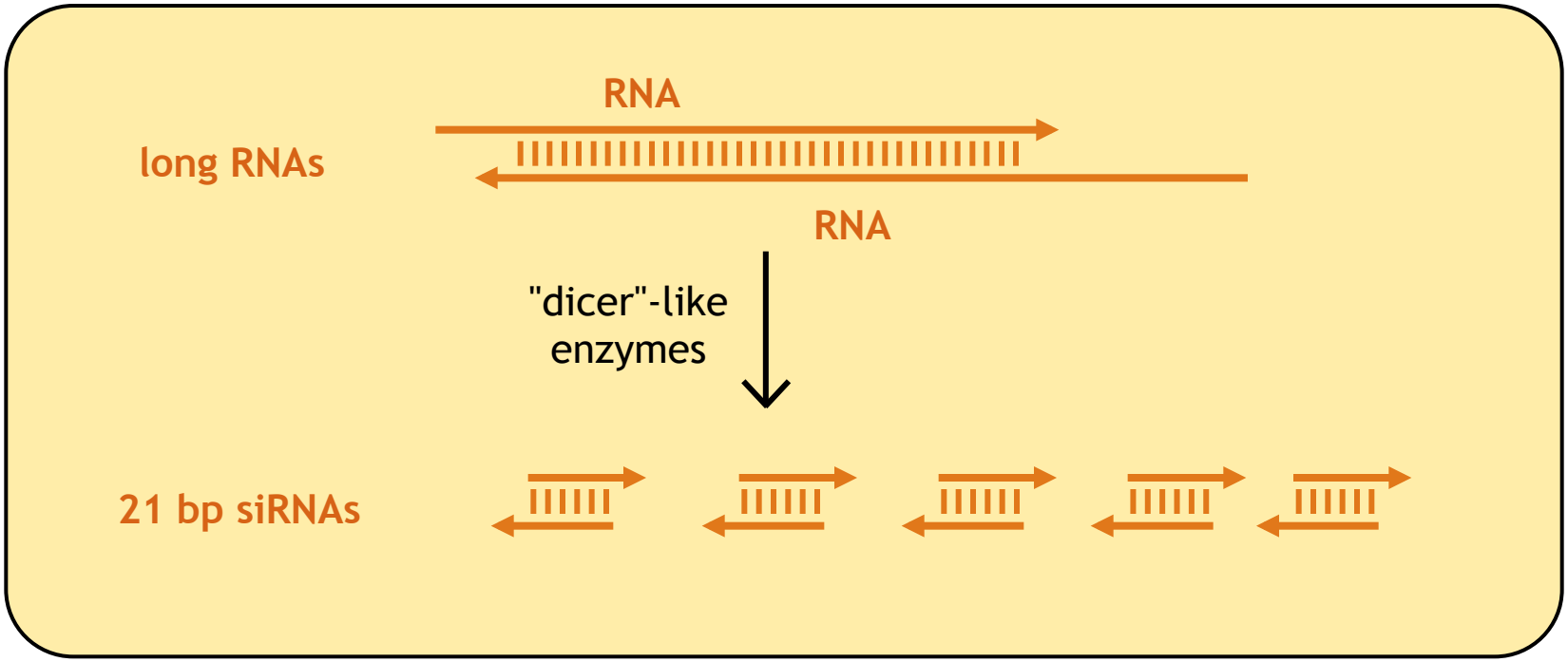


transcription of dispersed repeated retroelements ("genomic parasites")



Six easy steps to understanding RNAi

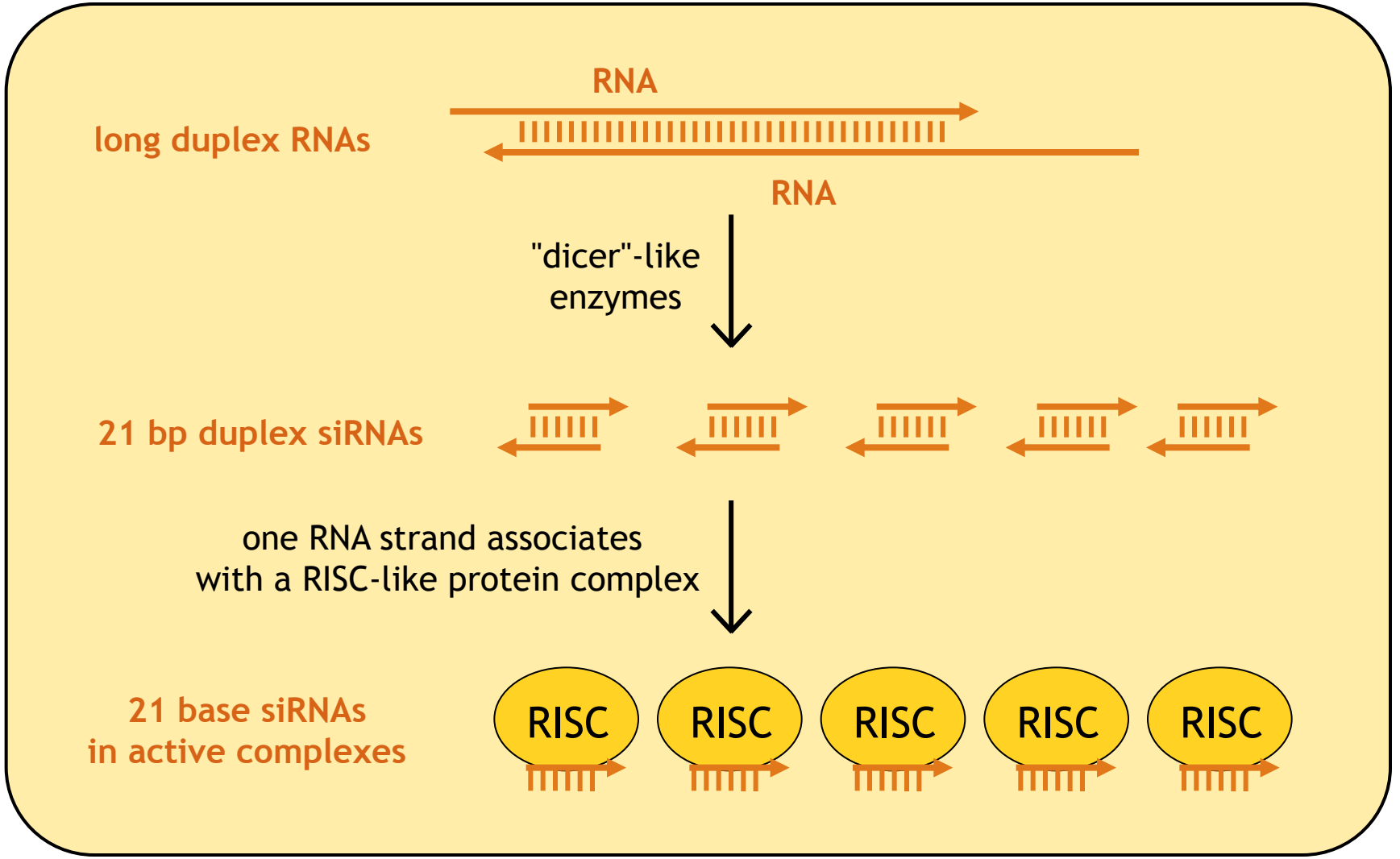
Step Two. The RNA duplex is "diced" into 21 bp fragments.



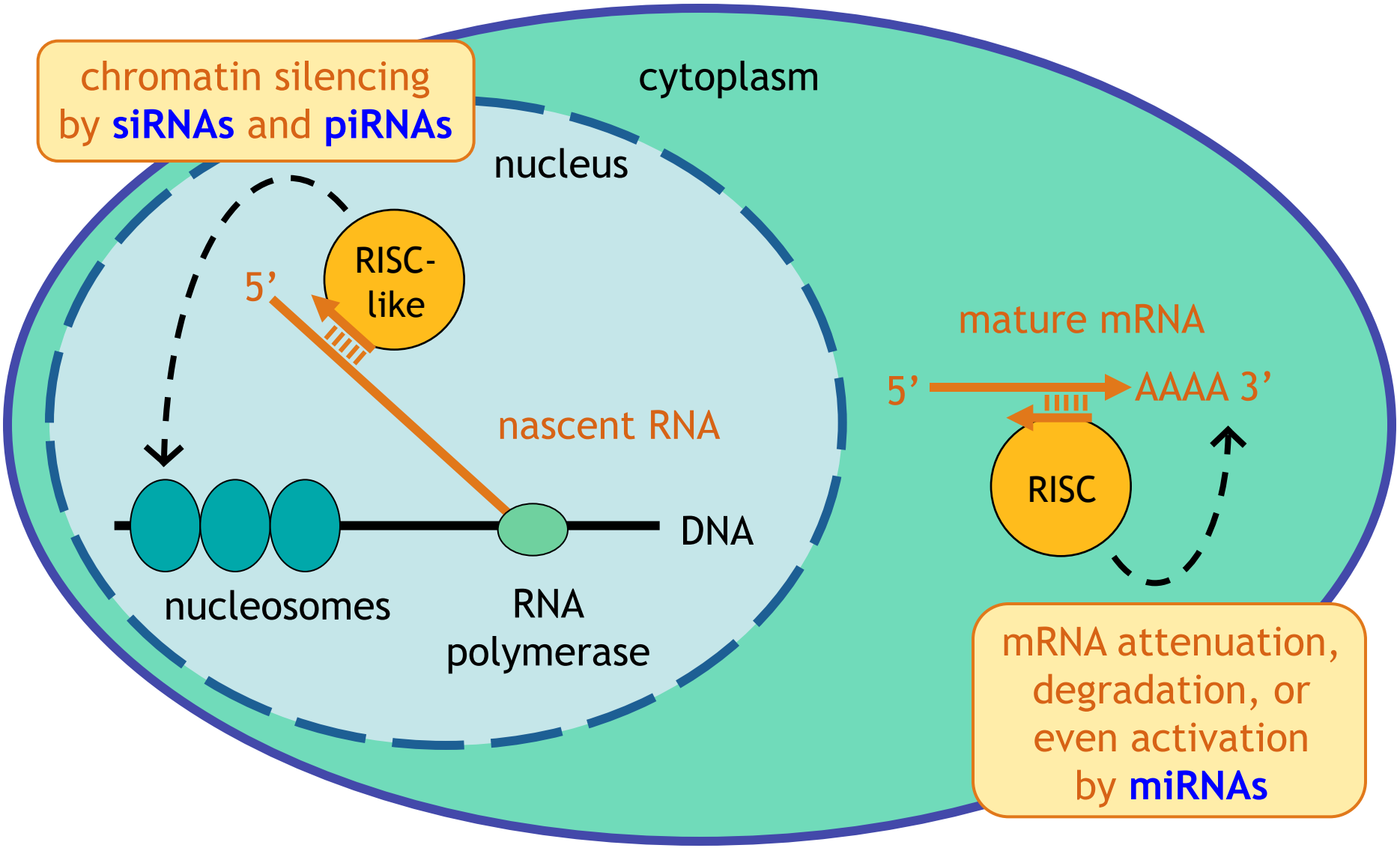
Six easy steps to understanding RNAi

Step Three. One RNA strand associates with a RISC-like complex

(RISC = RNA-Induced Silencing Complex)

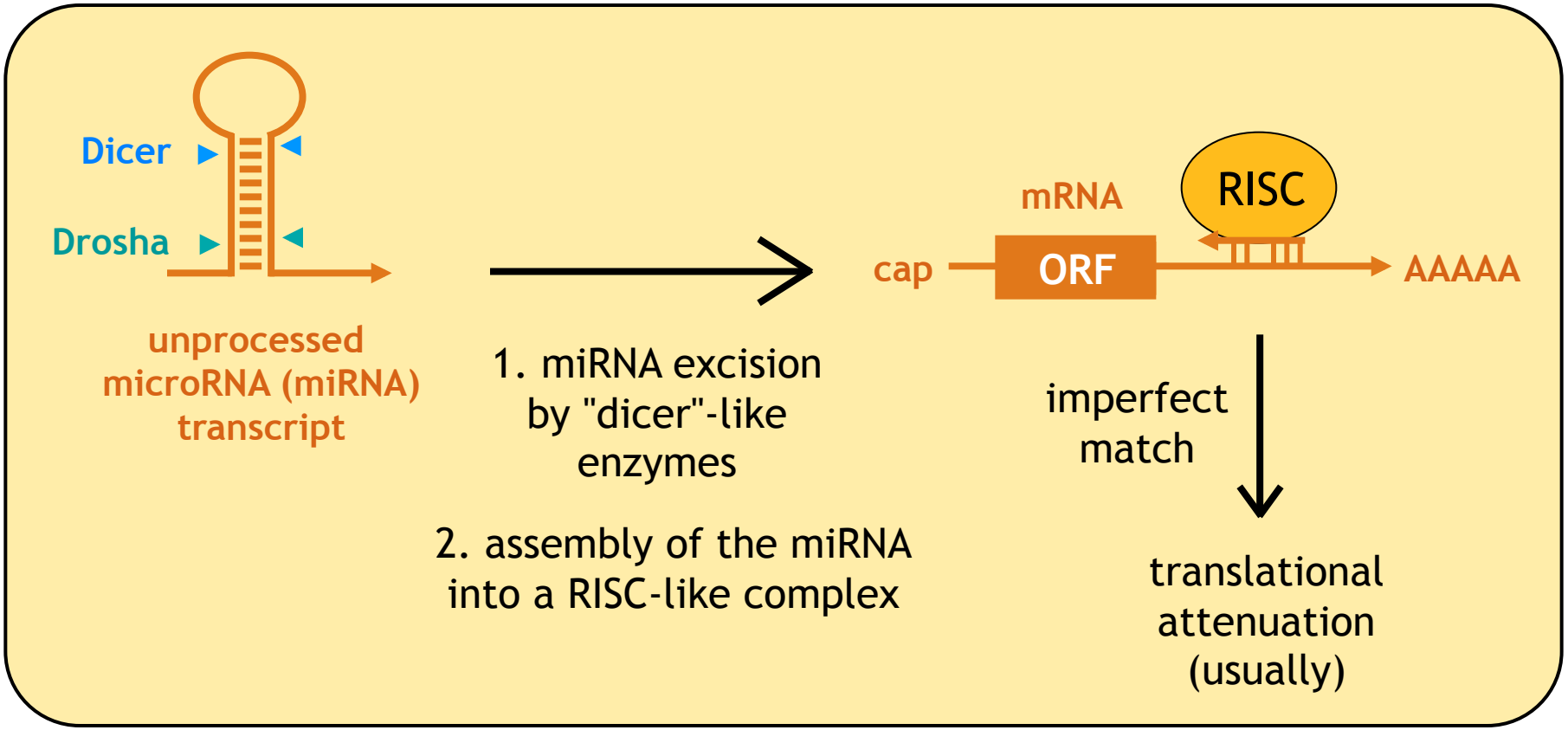


Step Four. A RISC-like complex anneals with complementary RNA, regulating or degrading the RNA and/or silencing the corresponding chromatin.



Six easy steps to understanding RNAi

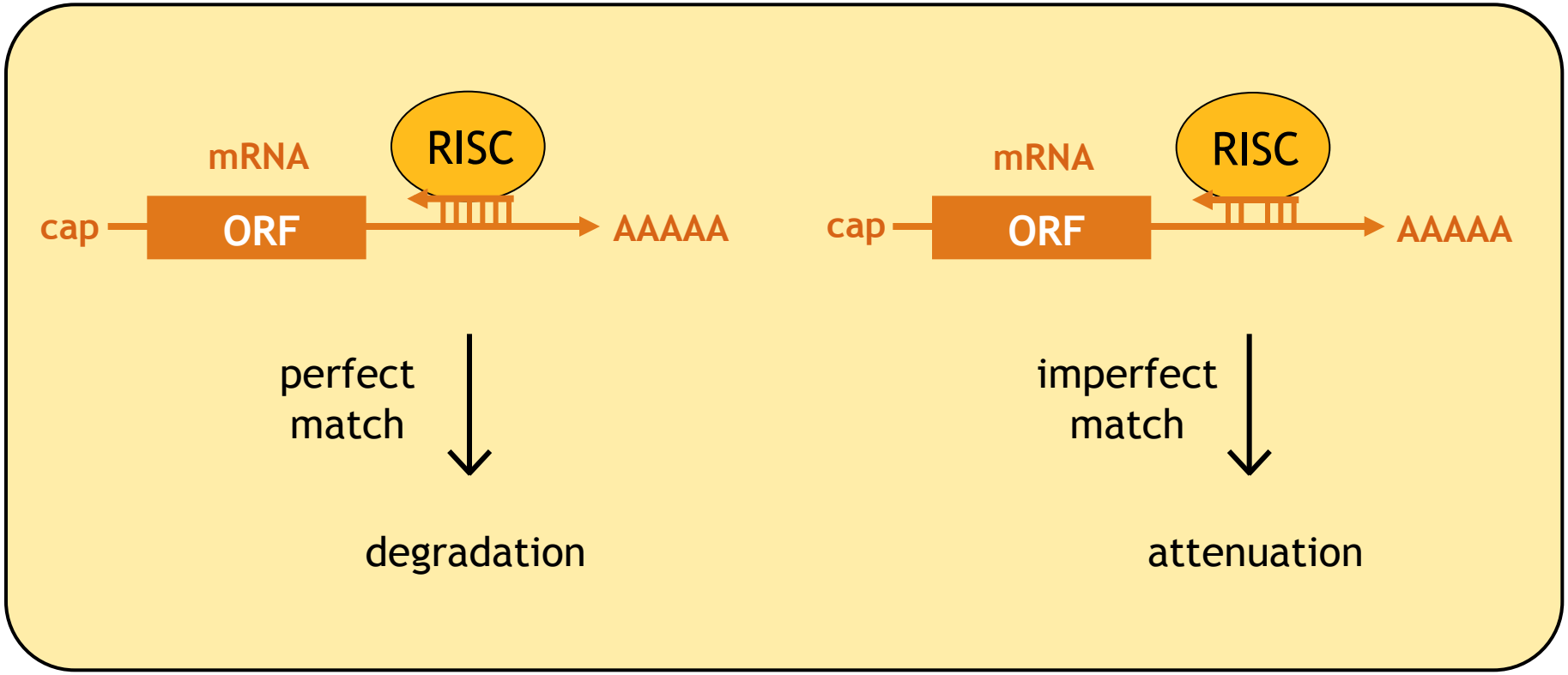
Step Five. Our genome encodes >1200 regulatory miRNAs.



Guo et al. (2010) Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature 466, 835-840; Friedman et al (2009) Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19, 92-105; Bartel (2009) MicroRNAs: target recognition and regulatory functions. Cell 136, 215-233.

Six easy steps to understanding RNAi

Step Six. Match with mRNA determines degradation or attenuation.



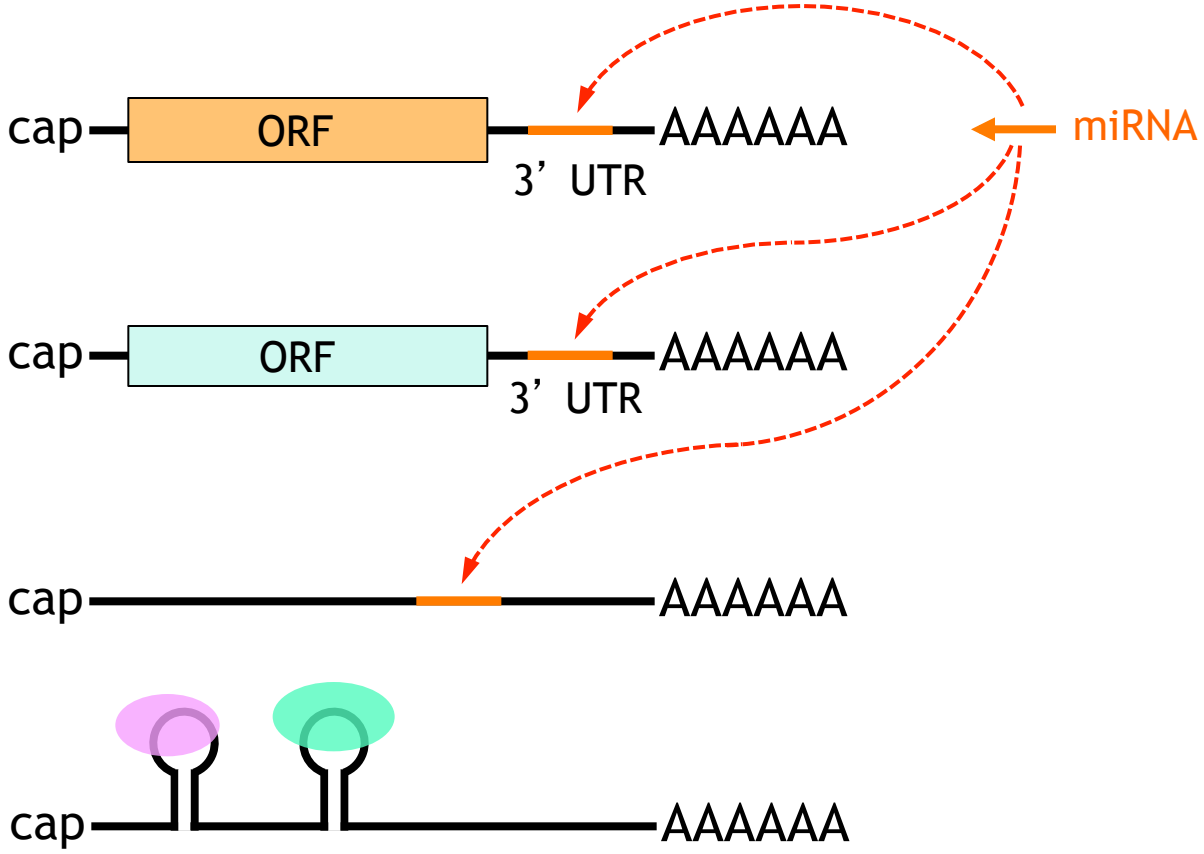
So, naturalists observe, a flea
Has smaller fleas that on him prey;
And these have smaller still to bite 'em;
And so proceed *ad infinitum*.

Jonathan Swift (1667-1745)
Poetry, A Rhapsody

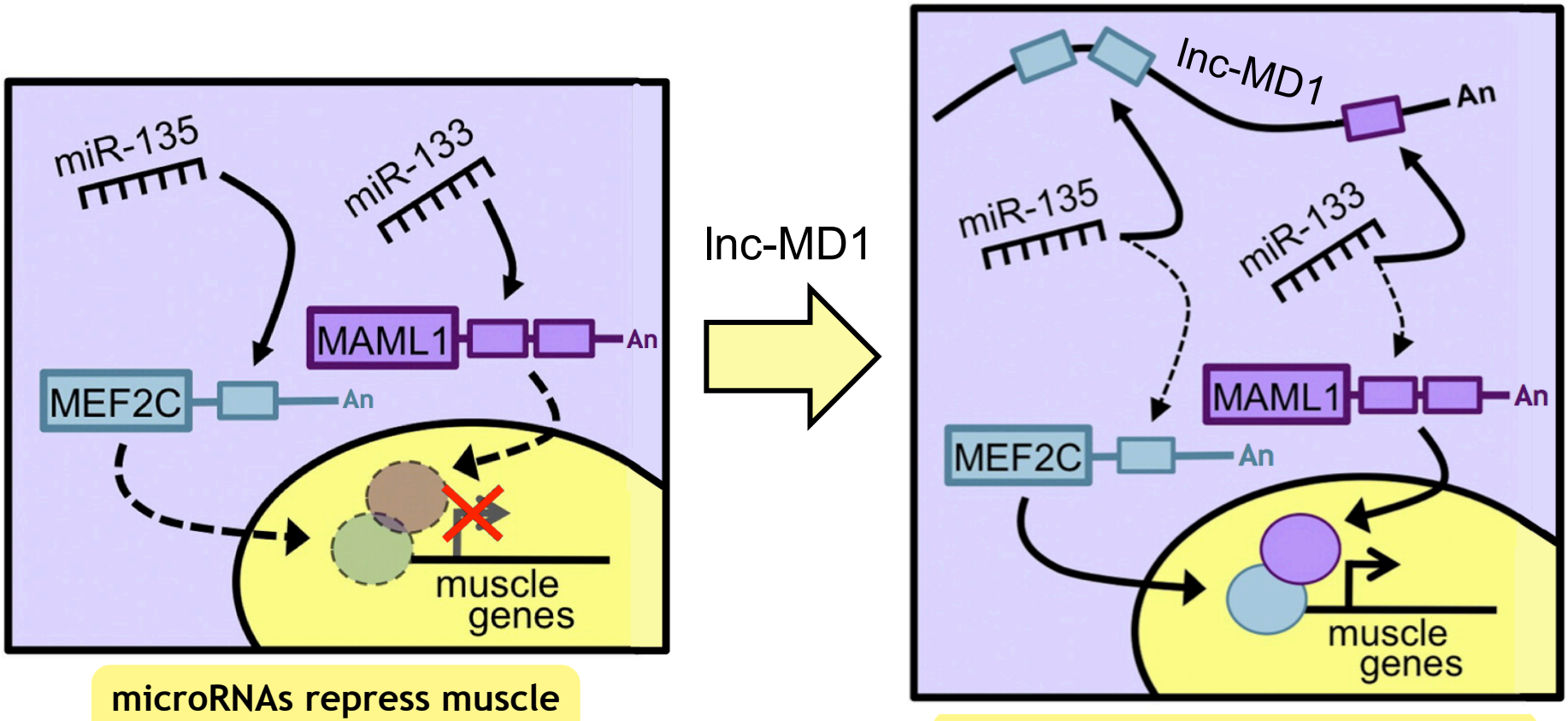
regulated mRNA

coding ceRNA (competing endogenous RNA)

noncoding ceRNAs (aka lncRNA for long noncoding RNA) may bind miRNAs to linear sequences or multiple proteins to 3D structures



Molecular circuitry by which lnc-MD1, miR-135, and miR-133 regulate muscle differentiation



microRNAs repress muscle mRNAs through 3' UTR

lnc-MD1 RNA relieves repression by soaking up miRNAs

Cesana et al. (2011) A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell 147, 358-369

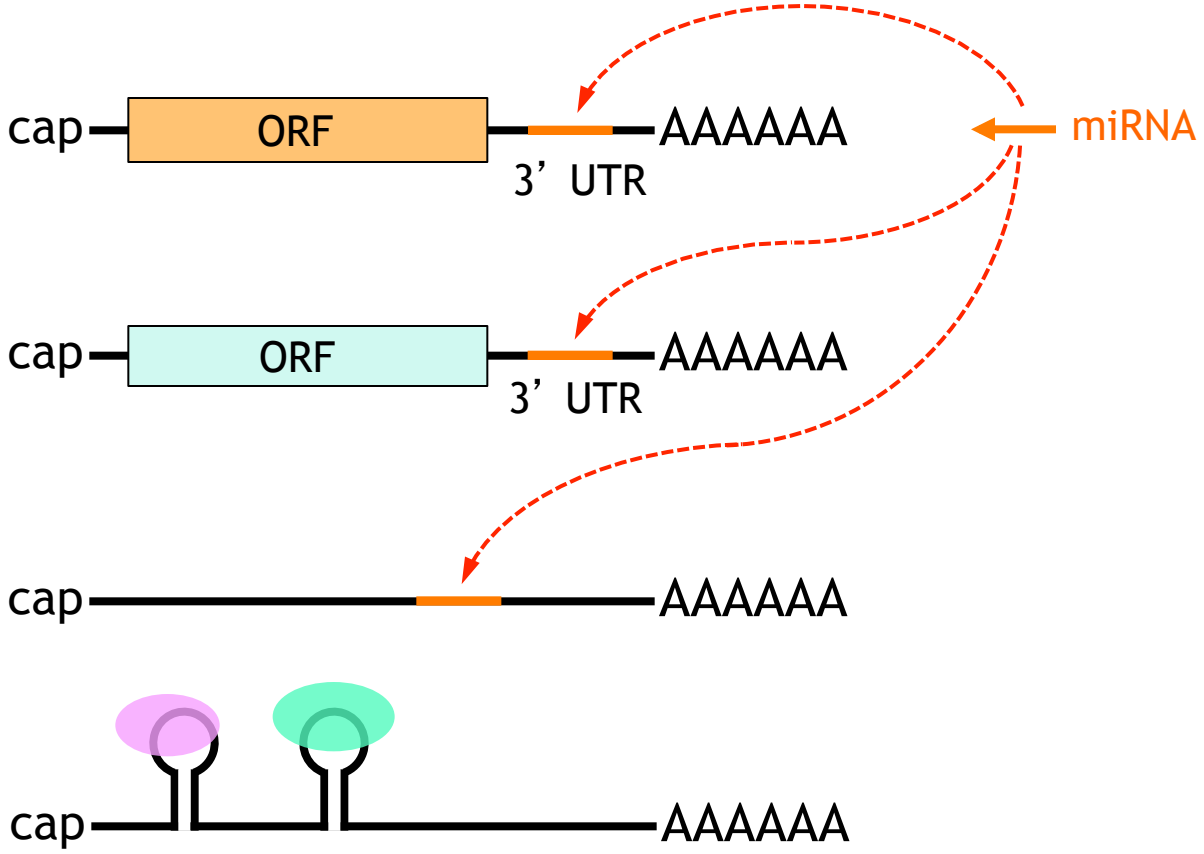
So, naturalists observe, a flea
Has smaller fleas that on him prey;
And these have smaller still to bite 'em;
And so proceed *ad infinitum*.

Jonathan Swift (1667-1745)
Poetry, A Rhapsody

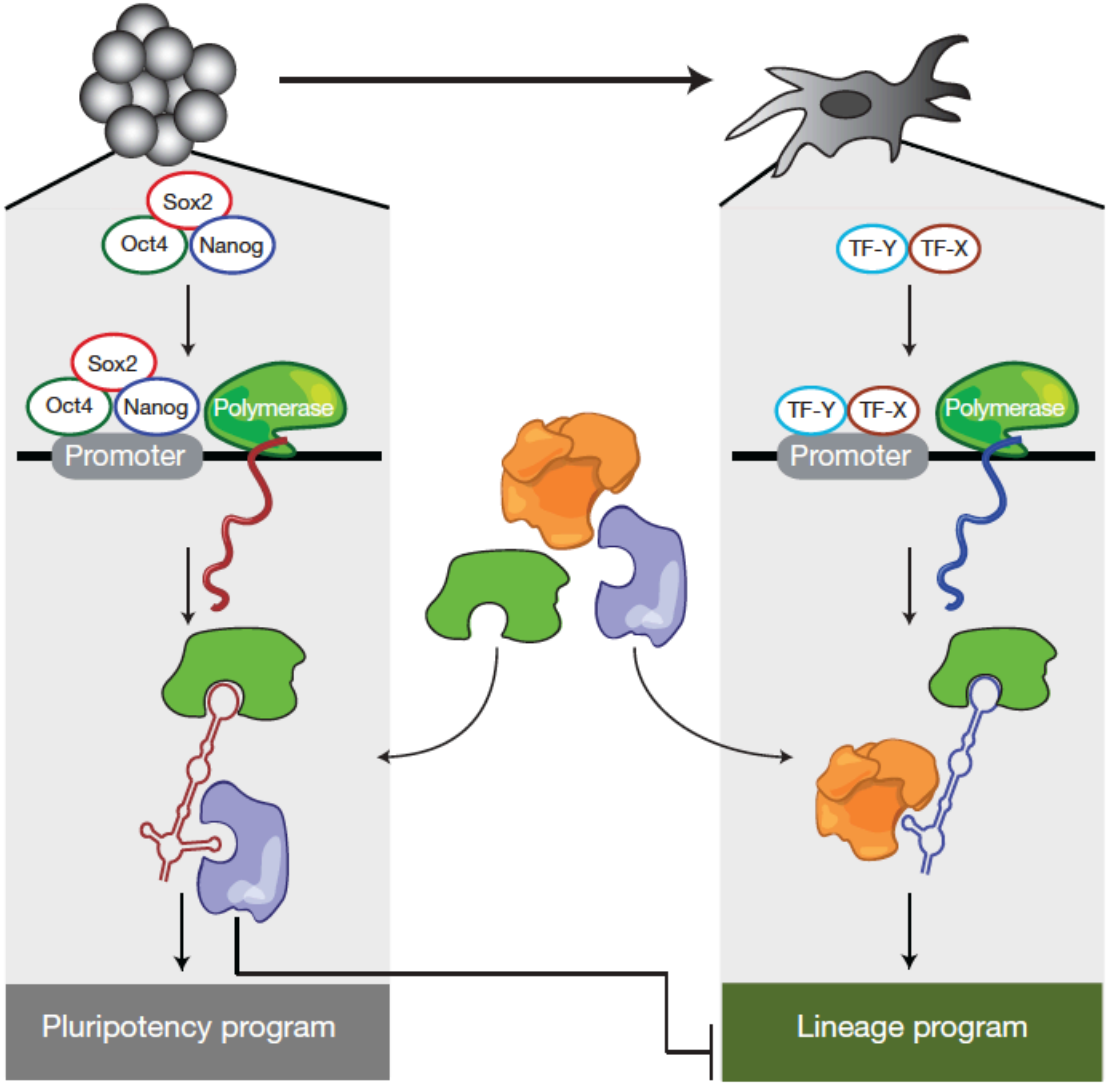
regulated mRNA

coding ceRNA (competing endogenous RNA)

noncoding ceRNAs (aka lncRNA for long noncoding RNA) may bind miRNAs to linear sequences or multiple proteins to 3D structures



lncRNAs act in the circuitry controlling pluripotency and differentiation [Guttman et al. (2011) Nature 477, 295]



A model for **lncRNA (long noncoding RNA)** function in cellular circuitry. ***(left)*** Embryonic Stem Cell (ESC) specific transcription factors (such as Oct4, Sox2 and Nanog) bind to the lncRNA promoter and drive transcription. The lncRNA binds to ubiquitous regulatory proteins, giving rise to cell-type specific RNA-protein complexes. Through various combinations of protein interactions, the lncRNA-protein complex gives rise to unique transcriptional programs. ***(right)*** A similar process may work in other cell types with specific transcription factors regulating lncRNAs, which form cell type-specific RNA-protein complexes that regulate cell-type-specific expression programs. Also see Ramos et al (2015) The long noncoding RNA Pnky regulates neuronal differentiation of embryonic and postnatal neural stem cells. Cell Stem Cell 16, 439-447; Bergmann et al (2015) Regulation of the ESC transcriptome by nuclear long noncoding RNAs. Genome Res 25, 1336-1346.

Leading Edge
Review

Cell

The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones

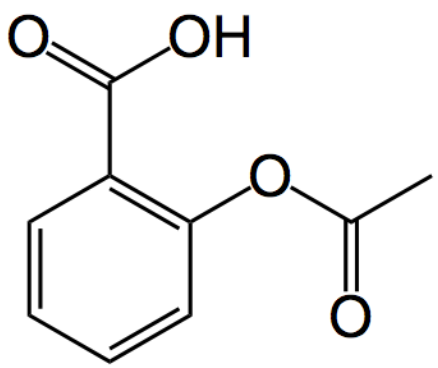
Thomas R. Cech[1,2,*] and Joan A. Steitz[1,3]
[1]Howard Hughes Medical Institute
[2]Department of Chemistry & Biochemistry, BioFrontiers Institute, University of Colorado Boulder, Boulder, CO 80309, USA
[3]Department of Molecular Biophysics and Biochemistry, Boyer Center for Molecular Medicine, Yale University School of Medicine, New Haven, CT 06536, USA
*Correspondence: thomas.cech@colorado.edu
http://dx.doi.org/10.1016/j.cell.2014.03.008

Noncoding RNAs (ncRNAs) accomplish a remarkable variety of biological functions. They regulate gene expression at the levels of transcription, RNA processing, and translation. They protect genomes from foreign nucleic acids. They can guide DNA synthesis or genome rearrangement. For ribozymes and riboswitches, the RNA structure itself provides the biological function, but most ncRNAs operate as RNA-protein complexes, including ribosomes, snRNPs, snoRNPs, telomerase, microRNAs, and long ncRNAs. Many, though not all, ncRNAs exploit the power of base pairing to selectively bind and act on other nucleic acids. Here, we describe the pathway of ncRNA research, where every established "rule" seems destined to be overturned.

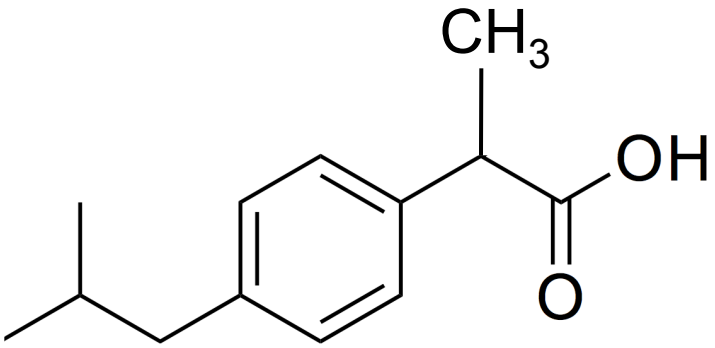
Cech TR and Steitz JA (2014) Cell 157, 77-94

If RNA is everywhere, and does everything, why is it such a challenging drug target?

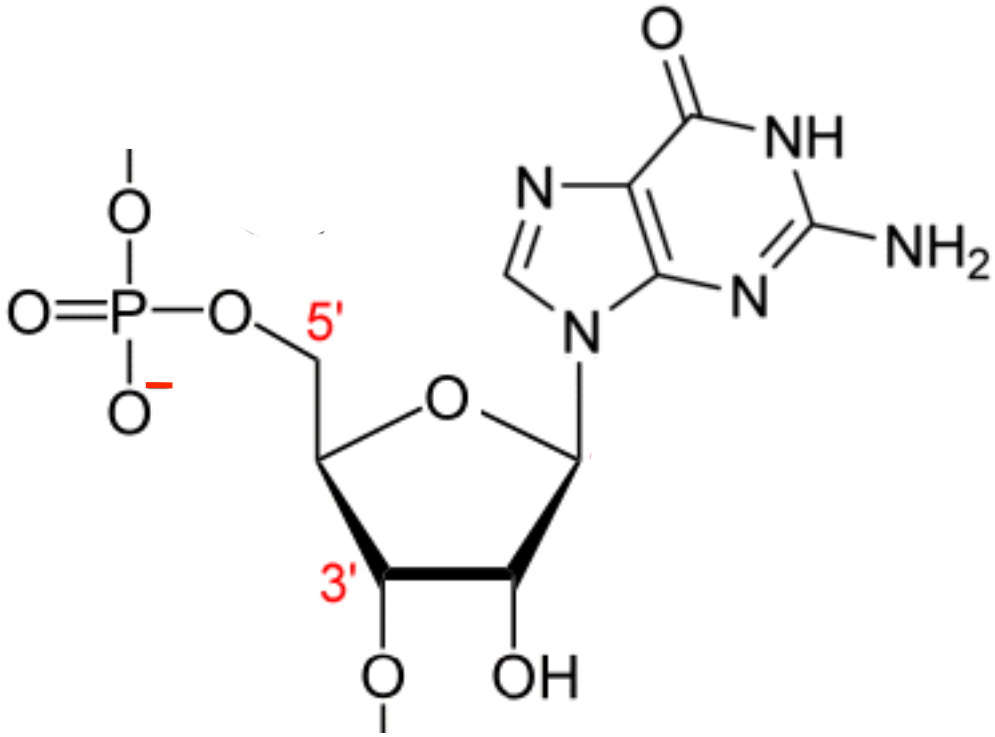
Why RNA doesn't look "druggable" at first glance



aspirin, 180 Da

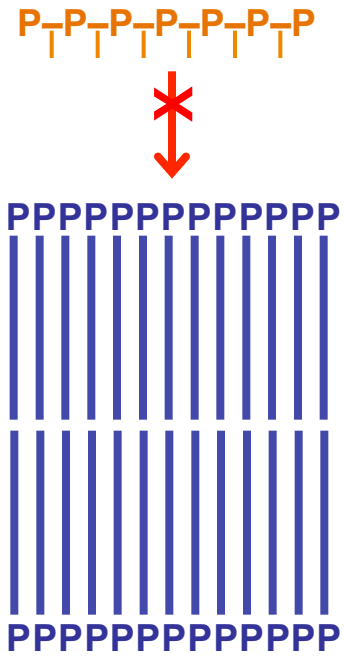


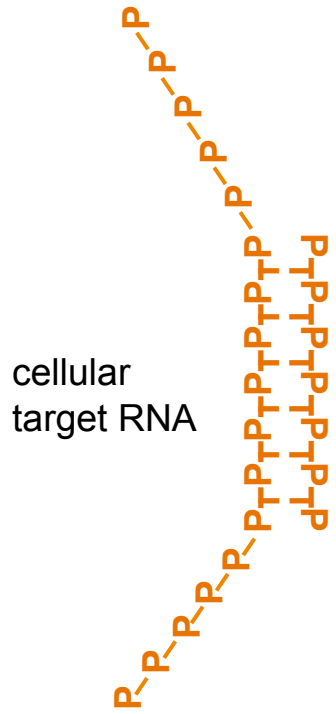
ibuprofen, 206 Da



The monomer unit of RNA is a 330 Da ribonucleotide; an RNA polymer of 15 monomer units is called an oligonucleotide (or “15-mer”) and would be 5,000 Da

Why *unmodified* oligonucleotides are not good drugs

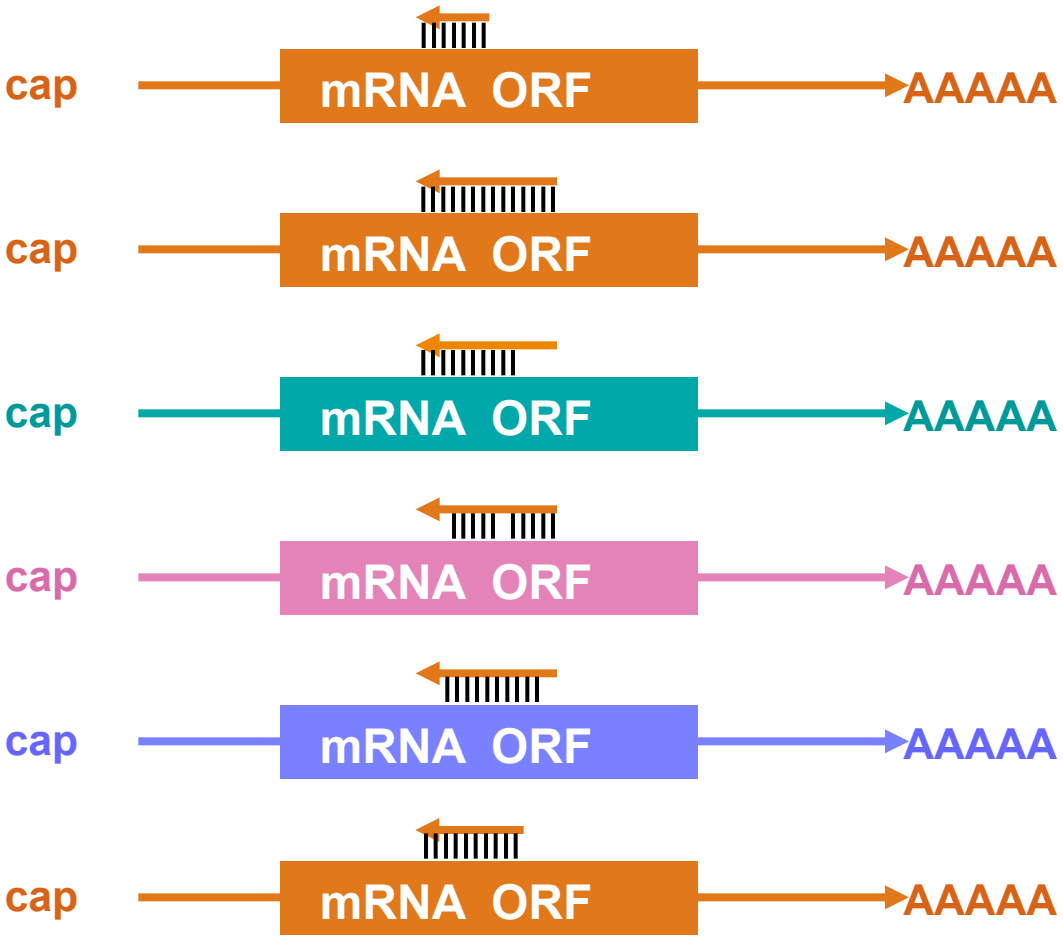




polyanions cannot penetrate the negatively charged membrane bilayer

repulsion between polyanionic backbones reduces stability despite energetically favorable base pairing and stacking

The antisense specificity problem is huge: there is no escape from "off-target" effects

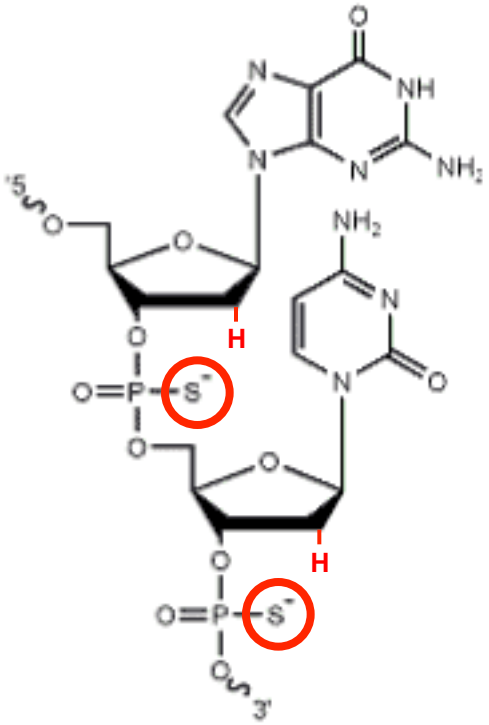


oligo too short, binds too weakly

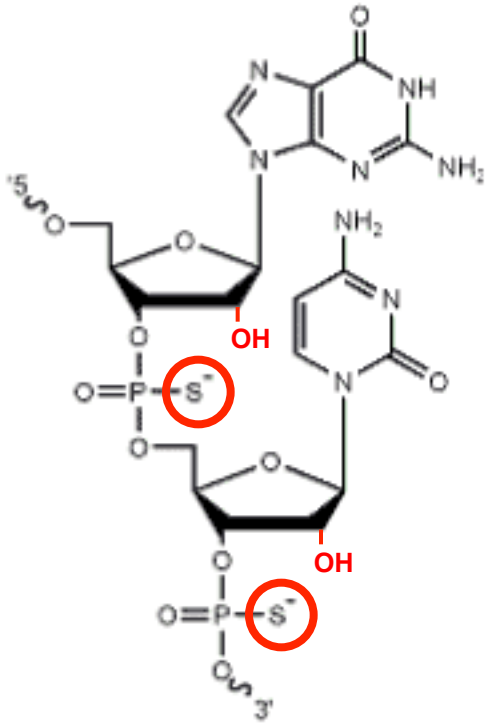
oligo too long, binds to unintended targets ("off-target effects")

"magic bullet" may not exist, biology is too complicated!

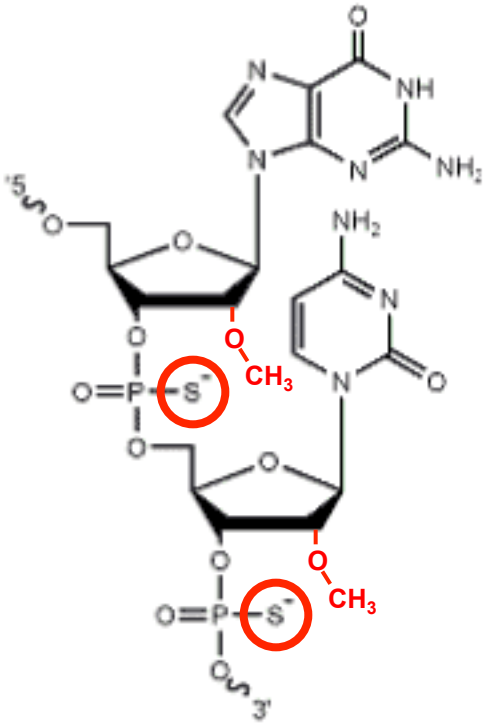
siRNAs that cannot be digested by exonucleases and endonucleases, and/or cannot cleave themselves



DNA with phosphorothioate linkages



RNA with phosphorothioate linkages

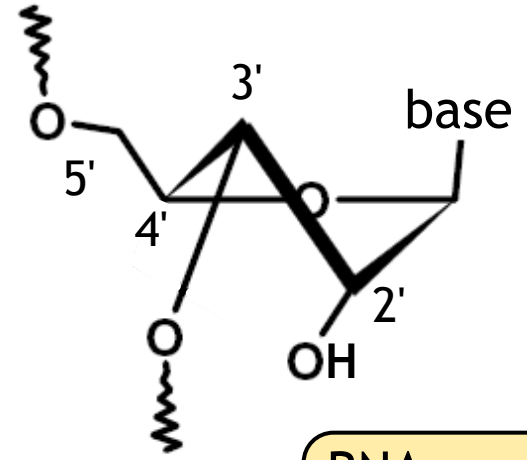


RNA with phosphorothioate linkages and 2'-O-methyl modifications

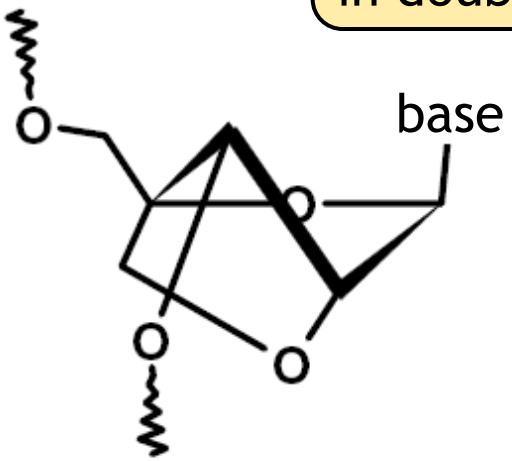
LNAs ("locked nucleic acids") enable shorter RNAs to achieve specificity and resist cellular nucleases

A locked nucleic acid (LNA) is an RNA monomer in which a bridge connecting the 2' oxygen and 4' carbon locks the ribose sugar into the conformation it would normally adopt in an RNA double helix. This conformation enhances base stacking, backbone pre-organization, and increases the melting temperature of the duplex; however, it also prevents LNAs from functioning as a ribozymes. LNA oligomers can be synthesized chemically.

In a therapeutic proof-of-principle, cardiac function and survival both improved when LNA oligomers complementary to miR-208a were injected into the tail vein of hypertensive rats [Montgomery et al. (2011) Circulation 124, 1537].



RNA monomer in double helix

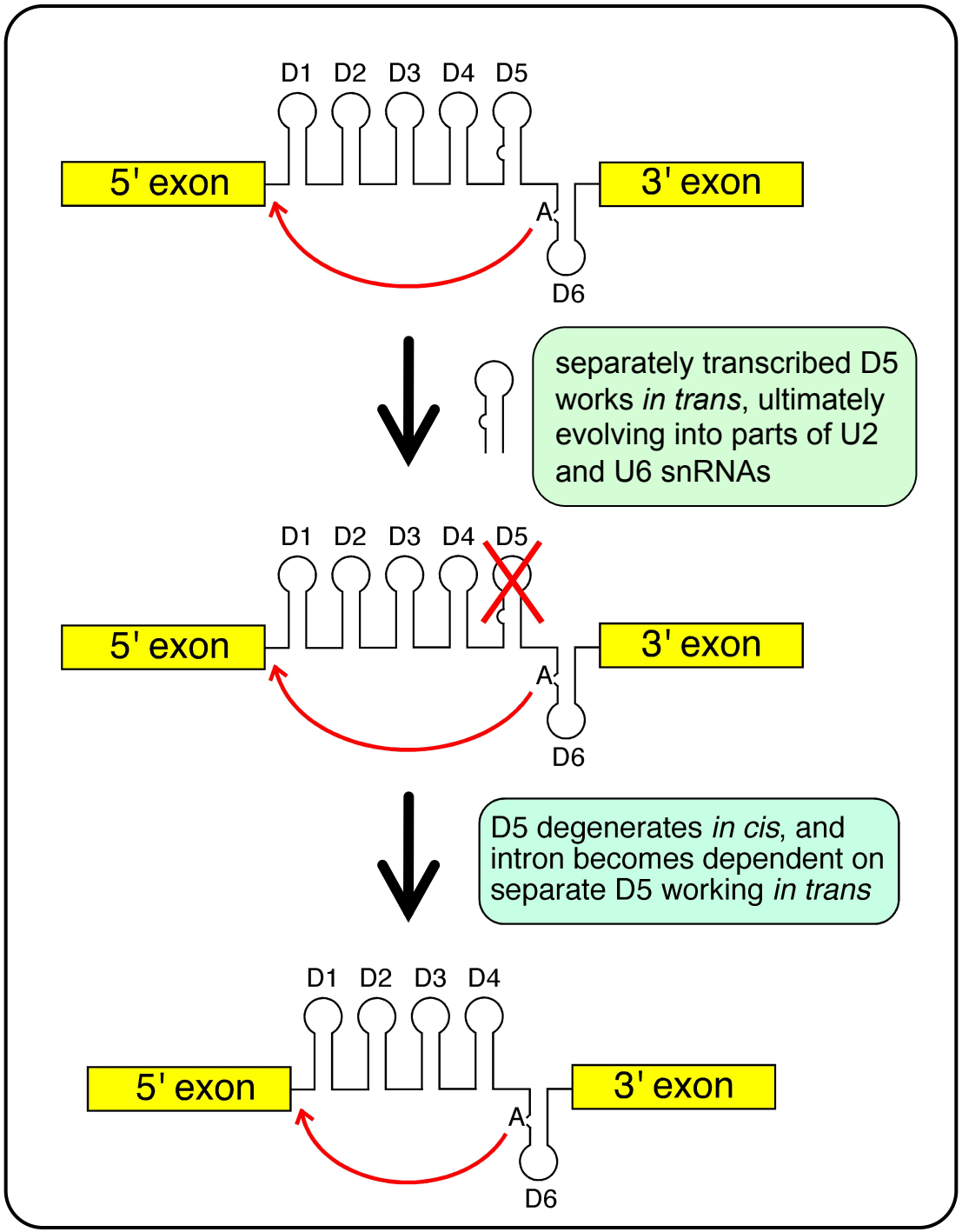


locked nucleic acid (LNA)

ENJOY RNA!

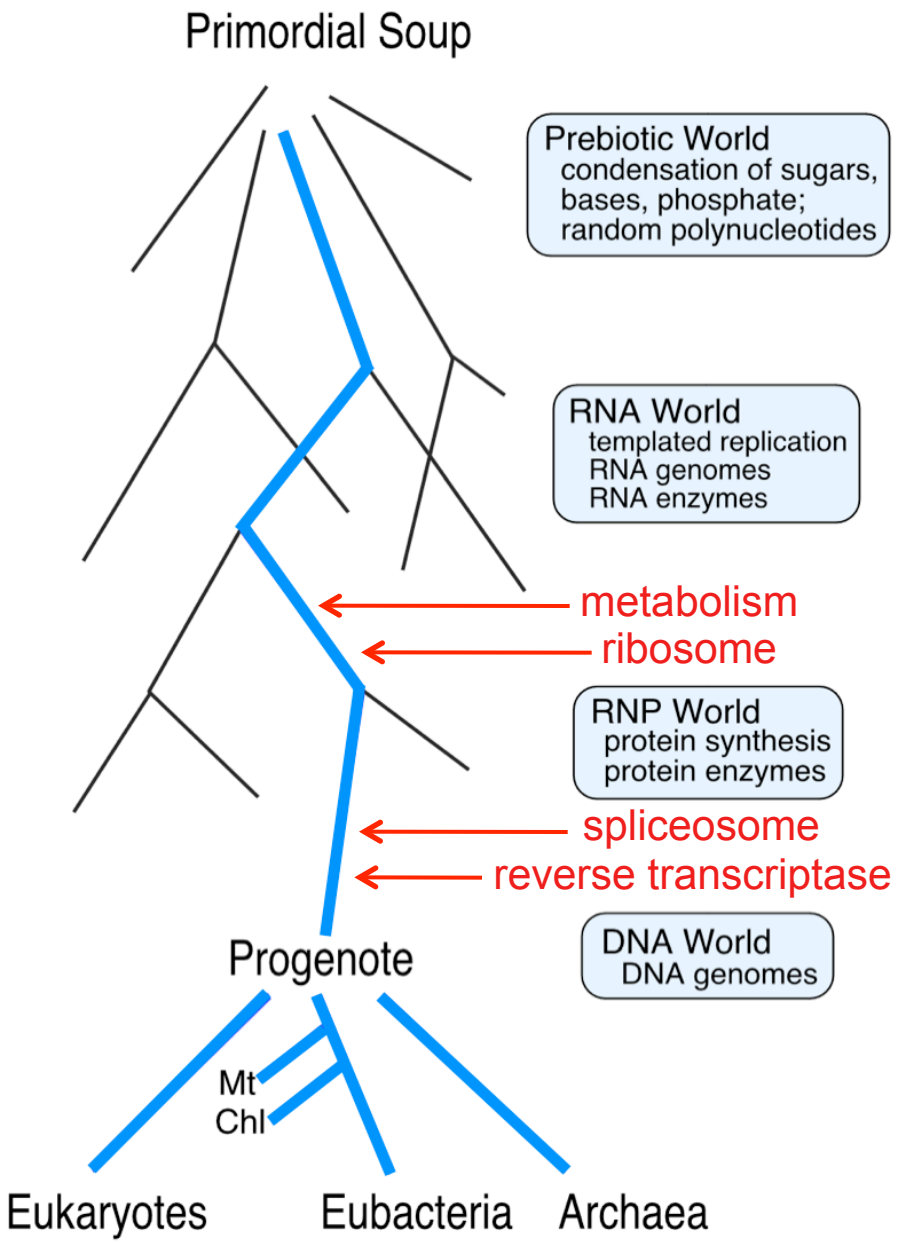
some of the slides I deleted
may interest you...

Group II introns in ancient protein coding genes could have evolved *in situ* into modern mRNA introns by sequential replacement of Group II domains with snRNAs working *in trans*



Jarrell, Dietrich, Perlman (1988)
Mol Cell Biol 8, 2361-2366

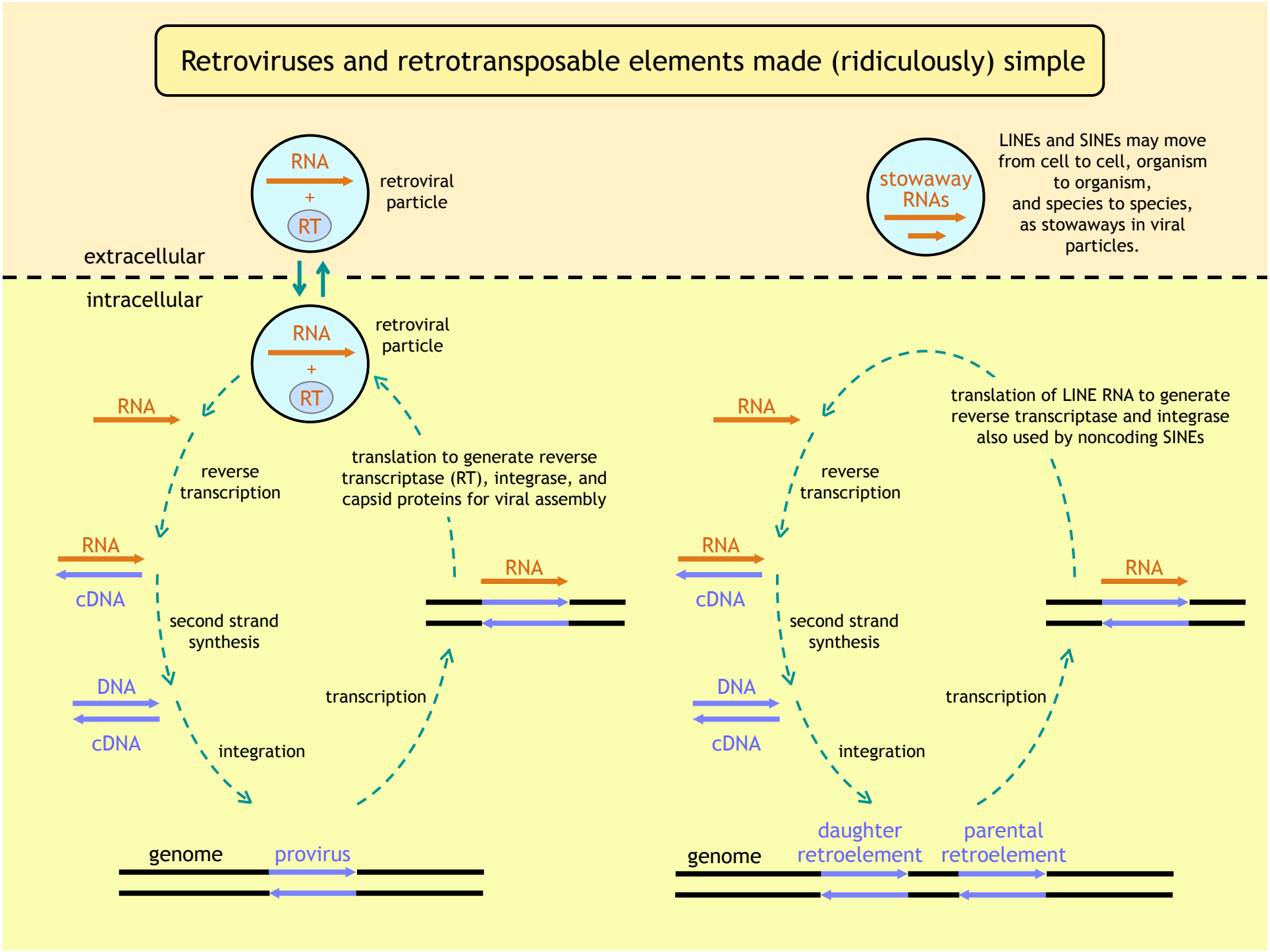
- Catalytic RNA could have replicated itself before the advent of protein synthesis!
- Invention of the ribosome created a ribonucleoprotein (RNP) world composed of RNA + proteins.
- Invention of the spliceosome created complicated proteins by exon shuffling.
- Invention of reverse transcriptase began the transition from RNA genomes to DNA genomes, with retroviral-like elements as transitional genomes that persist today.
- The RNP World is alive and well in our cells, but we are distracted by the vast genetic storage capacity of DNA!



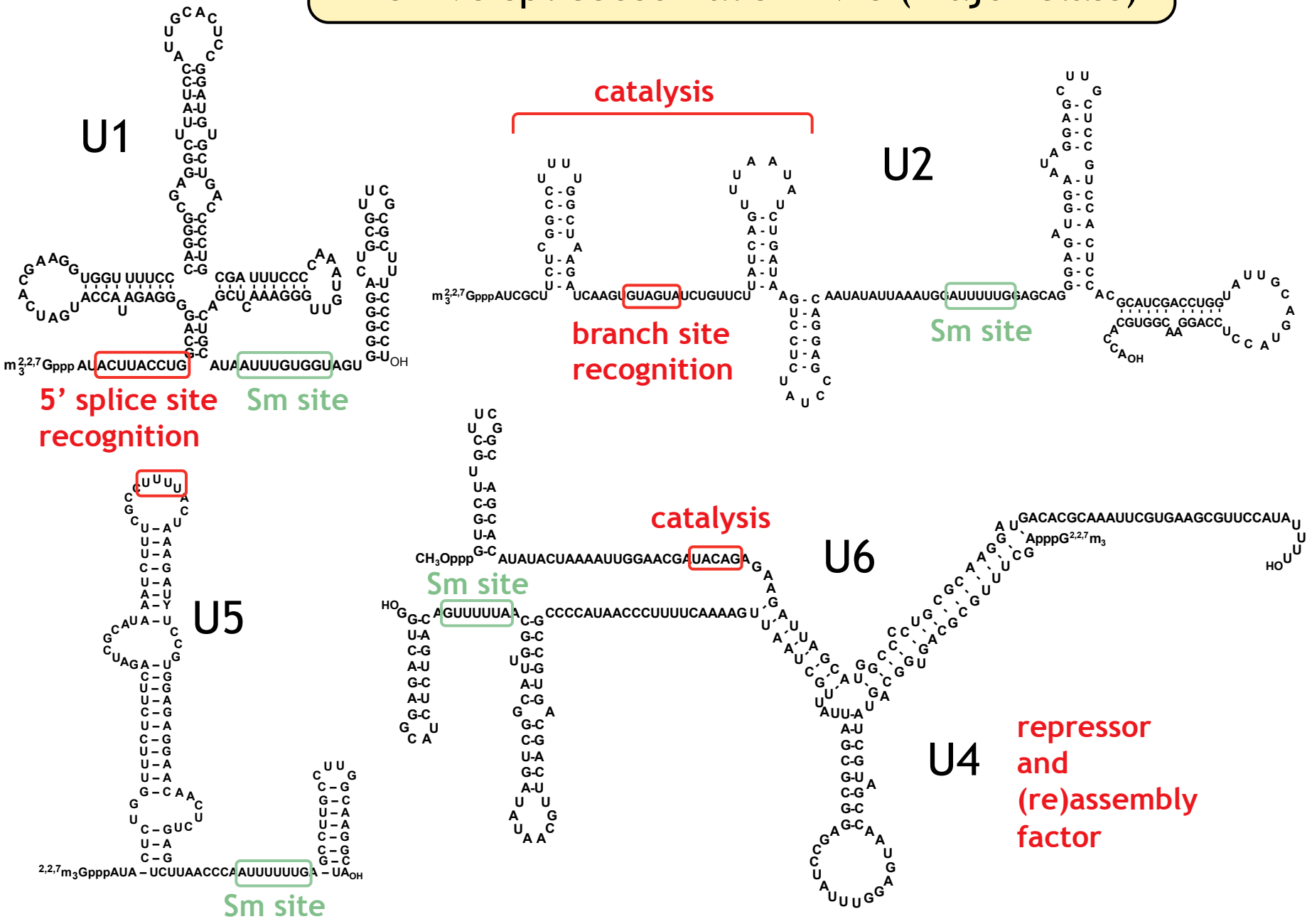
The human genome is chock full of parasitic retrotransposable elements ("retrotransposons" aka "retroposons") that use a reverse transcriptase to replicate and transpose through RNA intermediates

class	autonomy	structure of retroelement	length	copy number	percentage of genome
LINEs	autonomous	ORF1 RT-INT AAA	6-8 kb	850,000	21%
SINEs	non-autonomous	AAA	100-300 bp	1,500,000	13%
endogenous retrovirus-like elements	autonomous	gag RT-INT env	6-11 kb	450,000	8%
	non-autonomous	gag	1.5-3 kb		
DNA transposon fossils	autonomous	transposase	2-3 kb	300,000	3%
	non-autonomous		80-3,000 bp		

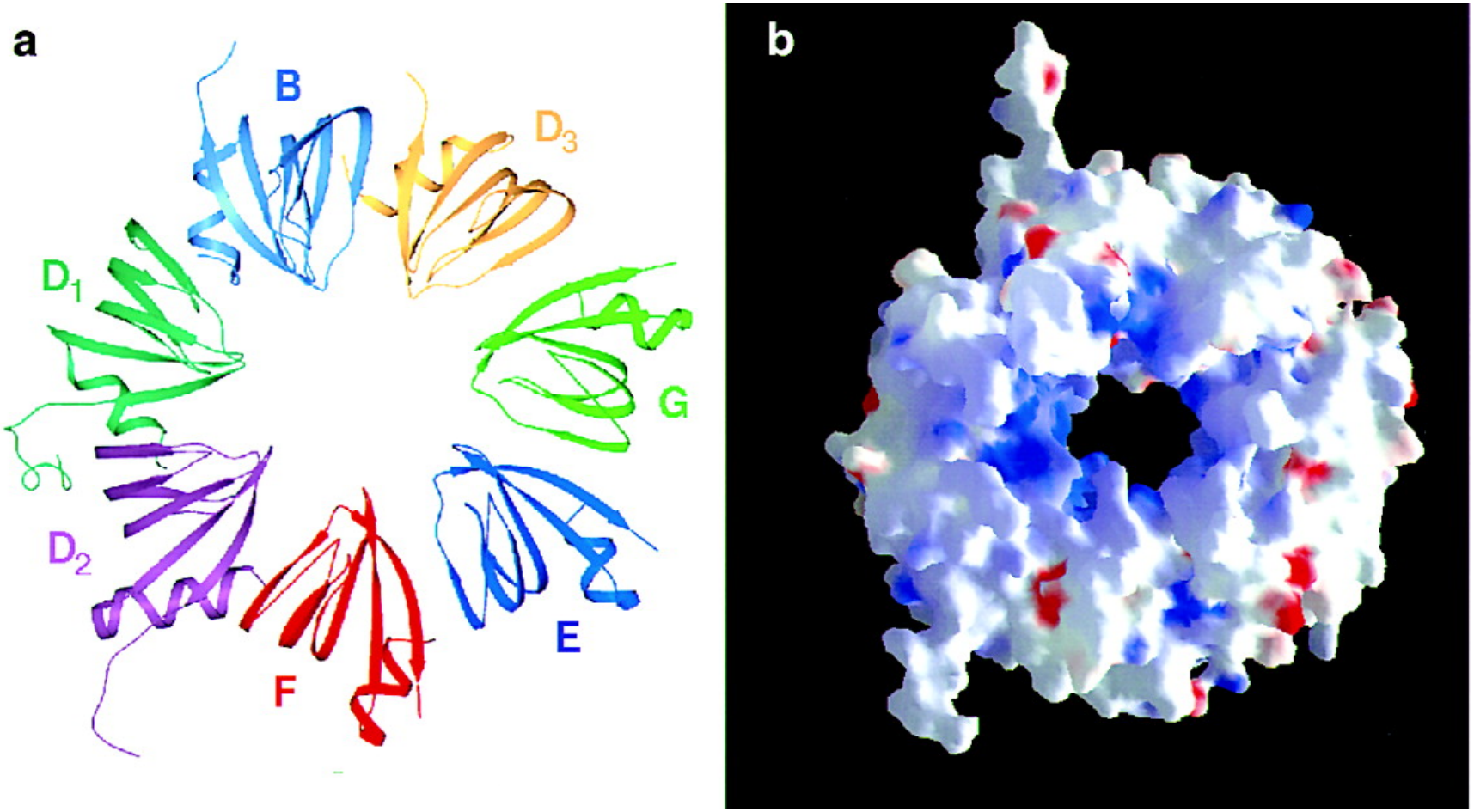
Retroviruses and retrotransposable elements made (ridiculously) simple



The five spliceosomal snRNAs (major class)

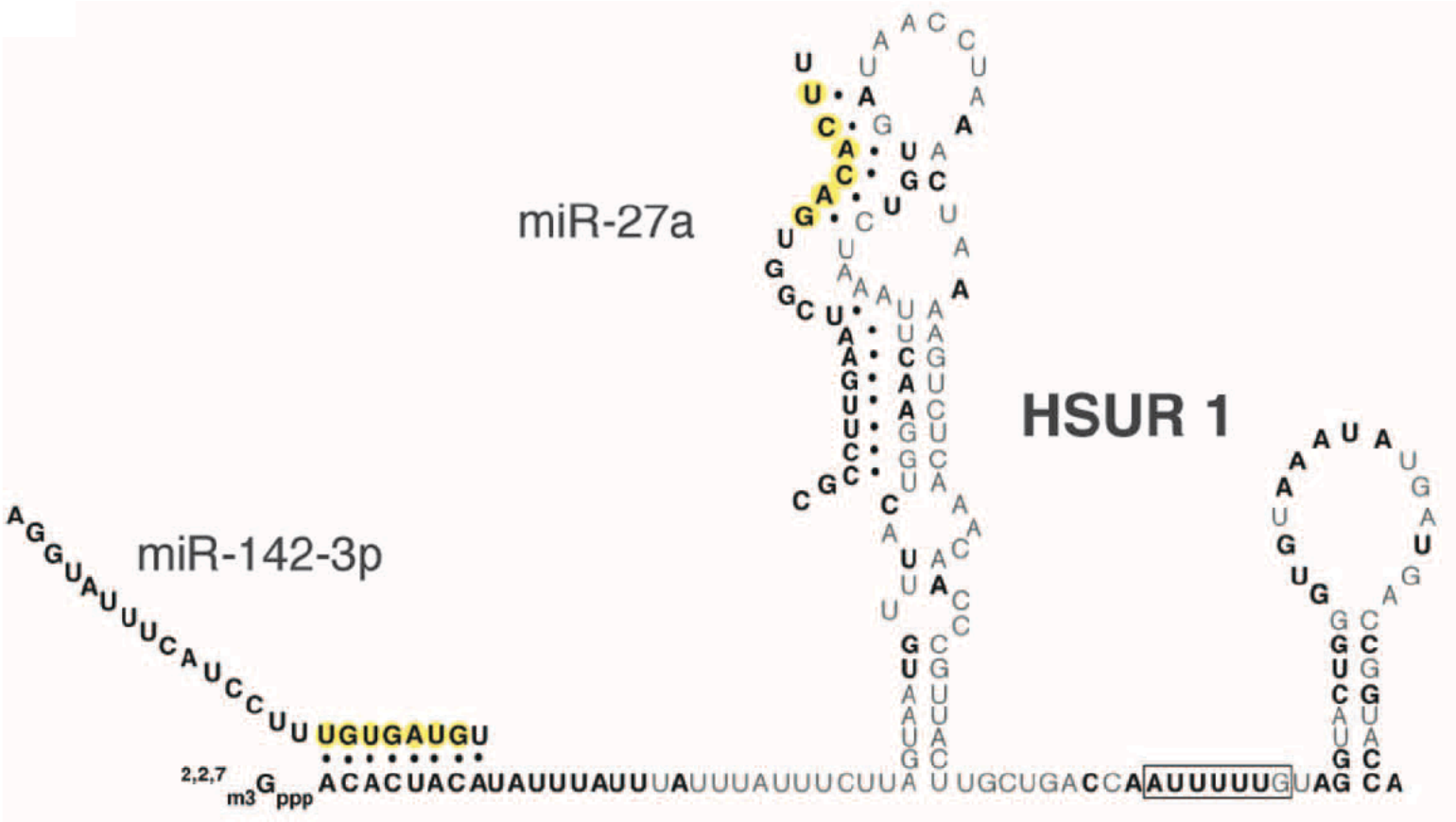


Sm proteins form a heptameric ring, and are found in Archaea!



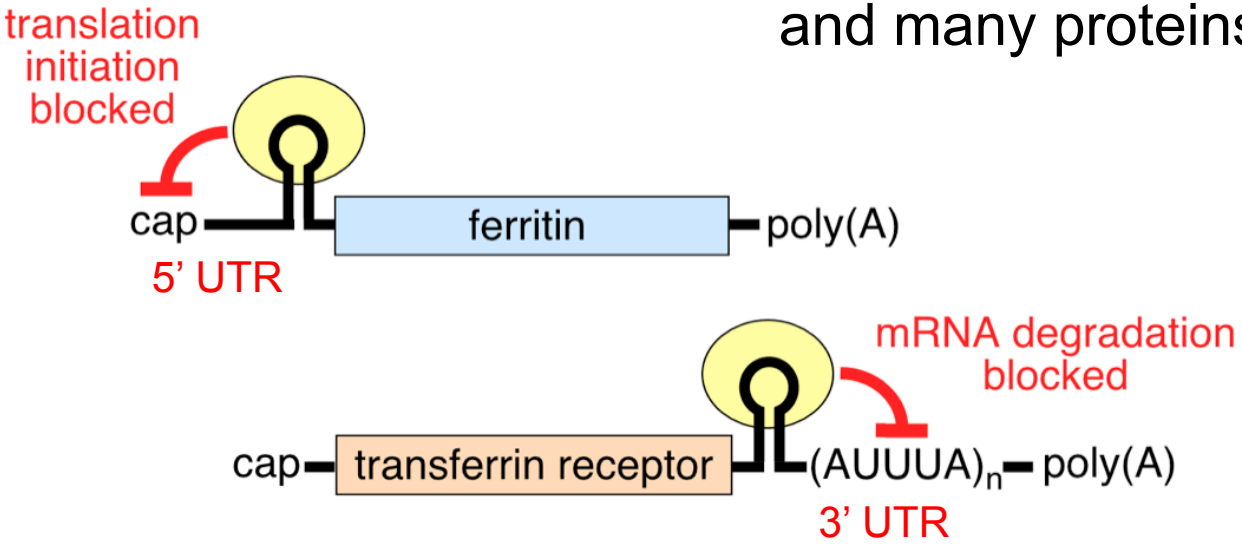
Kambach, Walke, Young, Avis, de la Fortelle, Raker, Lührmann, Li, and Nagai (1999) Cell 96, 375.

Here is Herpesvirus saimiri HSUR1 titrating out host miRNAs miR-142-3p and miR-27a



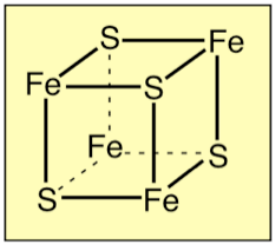
Cazalla et al. (2011) A Primate Herpesvirus Uses the Integrator Complex to Generate Viral MicroRNAs. Mol Cell 43, 982.

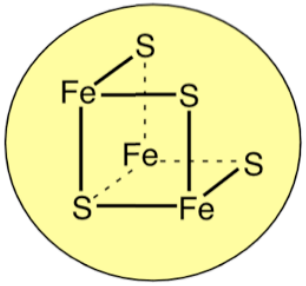
Translational control is powerful, and many proteins have "second jobs"



excess iron sequestered by ferritin

scarce iron imported by transferrin receptor





when iron is replete, cubic cluster forms in IBP (aconitase); protein is enzymatically active in metabolism but cannot bind IRE in mRNAs

when iron is scarce, cubic cluster cannot form in IBP (aconitase); protein is enzymatically inactive in metabolism but now binds IRE in mRNAs

Rouault and Klausner (1996) TiBS 21, 174; Beutler (2004) Science 306, 2051; Du et al. (2008) Science 320, 1088.