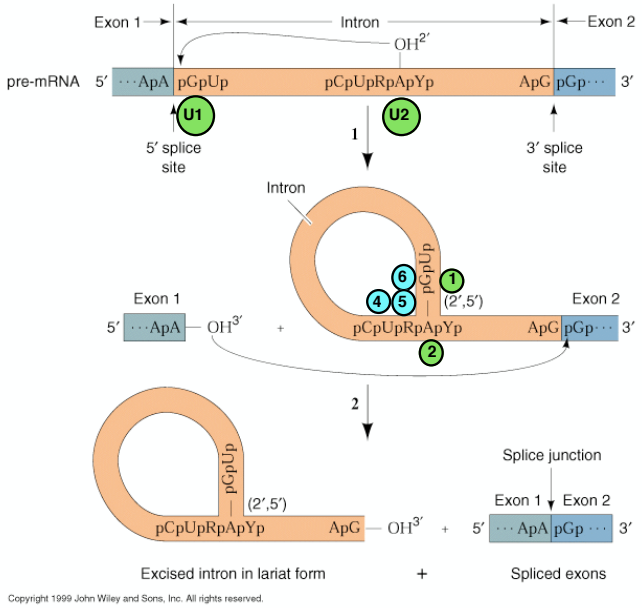
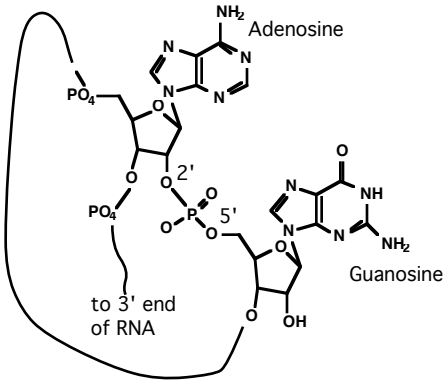
RNA processing
RNAs in both prokaryotes and eukaryotes are post-transcriptionally modified in some way:

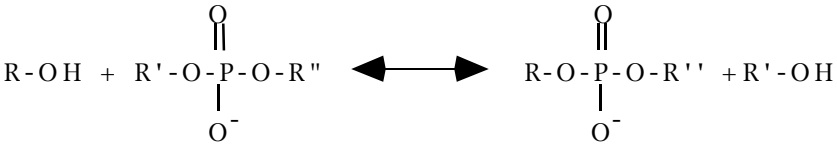
mRNA: splicing, editing, capping, polyadenylation, some chemical modifications

splicing: occurs in both prokaryotes and eukaryotes
genomic DNA sequence contains regions not found in mRNA
genomic sequence transcribed and extra sequences (called introns) spliced out leaving exons of mature mRNA
most eukaryotic splicing performed by spliceosomal complex:
splice sites determined by sequences at the ends of introns "GTAG rule"

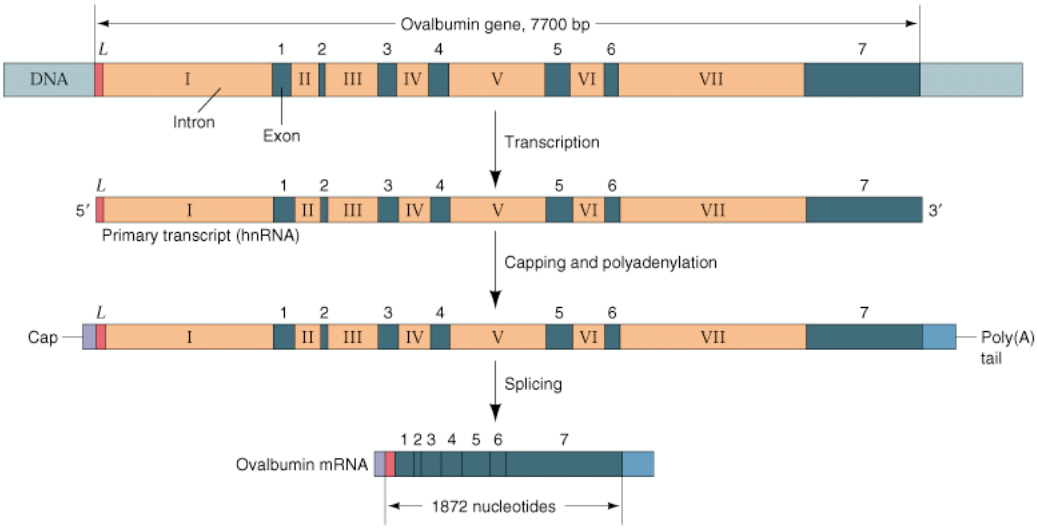




5' end of intron bound by U1 ribonucleoprotein **(commitment complex)**
Promotes U2 rnp recognition of internal A residue of intron
U4-U5-U6 complex docks with U1-U2 complex
Base Pairing interactions between RNAs rearrange:
5' splice site moves from U1 to U6
U6 pairs with U2 instead of U4
U6 catalyzes attack by 2'OH of internal A to PO4 of 5' end of intron
➔ free 5' exon (but held in place by spliceosome) and lariat-shaped intermediate
attack of 3'OH of 5' exon at PO4 of 3' exon
➔ spliced exons and lariat-shaped intron

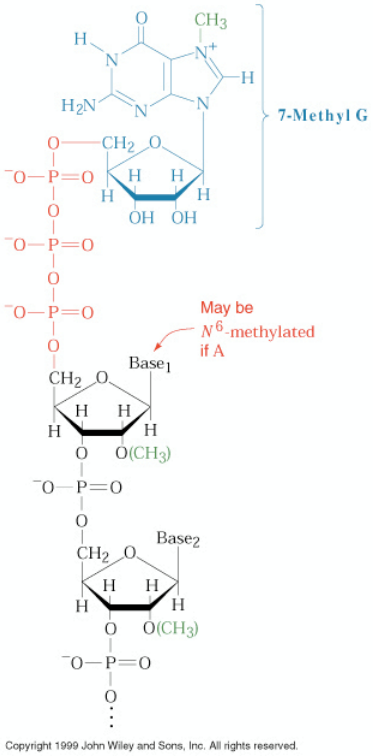


reactions are transesterifications:
concerted breaking of one bond while forming a new one
for first step: R = Adenosine2', R' = 5'exon, R" = intron+3'exon
for second step: R= 5'exon, R'= intron, R"=3'exon



usually occur in 5' to 3' direction, but some introns not spliced under certain developmental conditions: **alternative splicing**

some splicing in eukaryotes and prokaryotes doesn't fit rules of spliceosomal splicing. These are the self-splicing group I and group II introns (see below)



capping: 7methylGppp

added by guanyltransferase to 5' end of eukaryotic mRNA during transcription

polyadenylation: many eukary mRNAs have polyA tail added during the process of transcription termination.

AAUAAA sequence recognizec cleavage.
~250 A residues added by pol polymerase

Both are important for mRNA stability

editing: trypanosomes- some genomic sequences lack U residues found in mature message. Guide RNAs with polyU tails bind to mRNA, targeting it for editing. The U residues of the tail are spliced in to yield mature mRNA.

modifications: some messages are chemically modified at specific residues
cytosine deamination of apoB mRNA:
truncated form of apolipoproteinB found in intestine
CAA (Gln) codon deaminated to UAA (stop)

rRNA: cleavage, chemical modification

16S, 23S, 5S rRNA plus some tRNAs synthesized in a single transcript
tRNAs cleaved out by RNase P, rest is processed by RNase III, which recognizes stem-loops between the rRNA sequences
modifications many 2´-OMethyl, plus others similar to tRNA

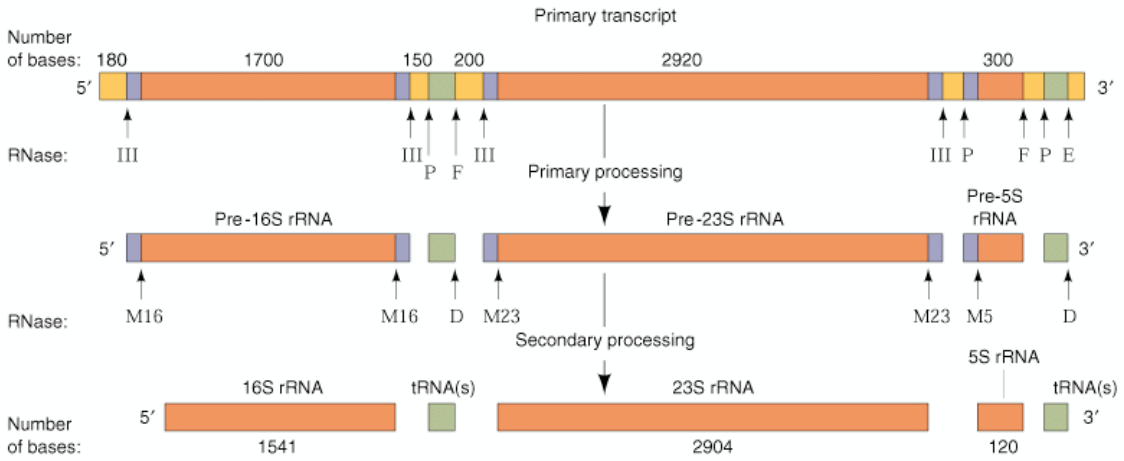


Figure 25-21. The posttranscriptional processing of *E. coli* rRNA. [After Apiron, D., Ghora, B.K., Plantz, G., Misra, T.K., and Gegenheimer, P., *in* S¨ll, D., Abelson, J.N., and Schimmel, P.R. (Eds.), *Transfer RNA: Biological Aspects*, p. 148, Cold Spring Harbor Laboratory (1980).]

Bacterial rRNA processing

Eukaryotic rRNA processing roughly similar.

Methylation occurs before further processing.

tRNA: cleavage, chemical modification, some introns

cleavage: many tRNAs are transcribed in groups on a single message or interspersed between rRNAs

a variety of exonucleases and endonucleases are responsible for production of mature tRNA:

RNase P: endonuclease cleaves off extra 5' sequences to yield mature tRNA 5' end

3' end from endonuclease cleavage (RNase P or others) followed by a variety of exonucleases. (including RNase D)

In some cases the 3' CCA sequence is added post transcriptionally by tRNA nucleotidyl transferase

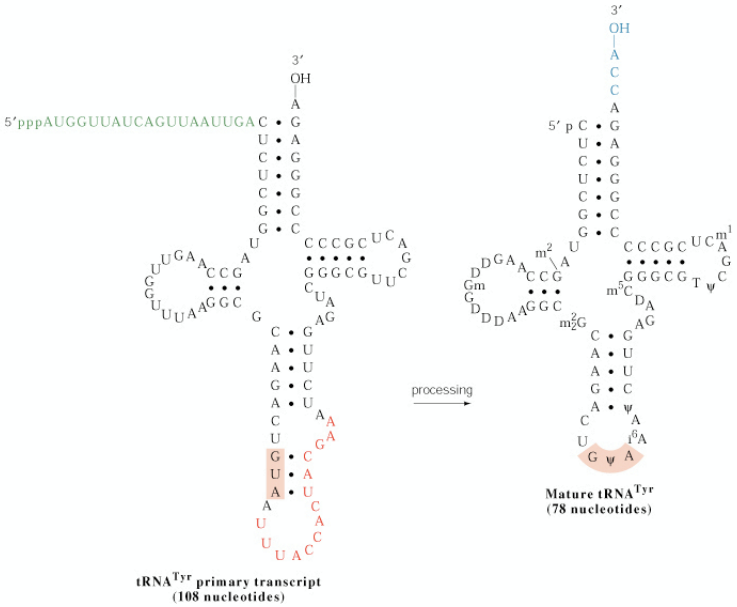


Figure 25-26. The posttranscriptional processing of yeast tRNA^{Tyr}. [After DeRobertis, E.M. and Olsen, M.V., *Nature* 278, 142 (1989).]

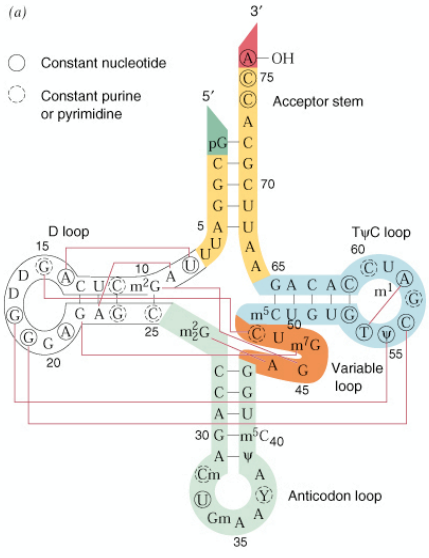


Figure 26-6a. Key to Structure. The structure of yeast tRNA^{Phe}. [Courtesy of Mike Carson, University of Alabama at Birmingham.]

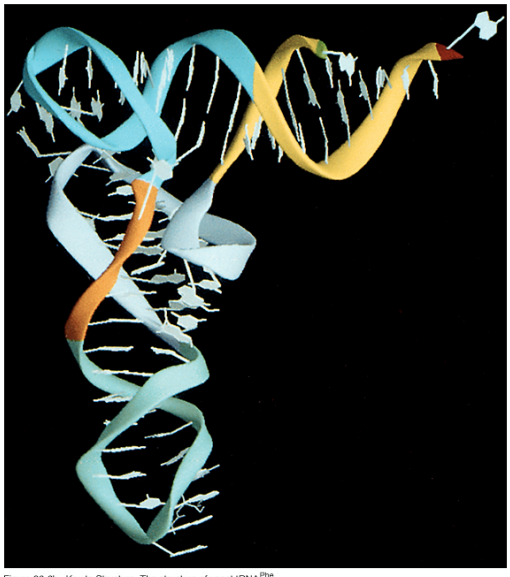


Figure 26-6b. Key to Structure. The structure of yeast tRNA^{Phe}. [Courtesy of Mike Carson, University of Alabama at Birmingham.]

mature tRNA

73 - 93 nt in length, most of sequence is variable, but secondary and tertiary structure essentially the same

2° structure - "cloverleaf"

four base-paired helical regions: anticodon stem, D, T and acceptor stems

T and acceptor stem stack into one continuous helix

D and anticodon stack too

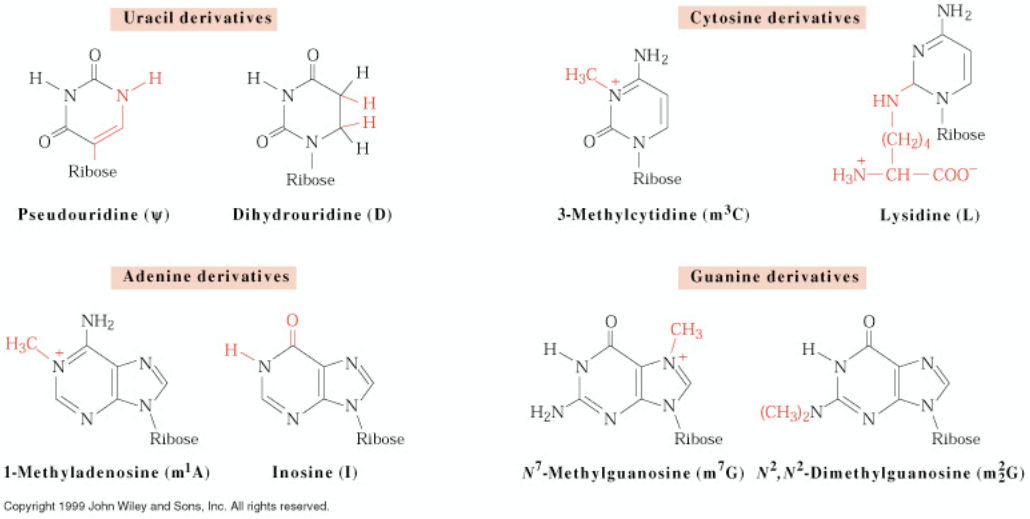
4 loops: anticodon, D, T and variable loops. D and T loops interact to form ➔

3° structure - elbow shaped

the anticodon loop, which base pairs with mRNA codon is 80Å away from acceptor stem 3' end, where amino acid is attached

3' end always CCA

modifications: a variety of chemical modifications are made to specific bases of tRNA. These help the tRNA fold into proper tertiary structure and aid in the recognition of tRNA by enzymes such as the ribosome



introns: removed by yet another pathway-
cleaved by specific endonuclease and later ligated together

Catalytic RNAs (Ribozymes)			
Intron splicing	**Satellite RNAs**	**RNase P**	**Others**
Group I	Hammerhead	Bacterial	Ribosome!
Group II	Hairpin	Other species?	Spliceosome
Spliceosome?	Hepatitis (HDV)		Do-it-yourself

Most involved in making/breaking phosphodiester bonds

Most require Mg^{++} or other divalent cation for activity

Many are associated with proteins *in vivo* to increase efficiency.

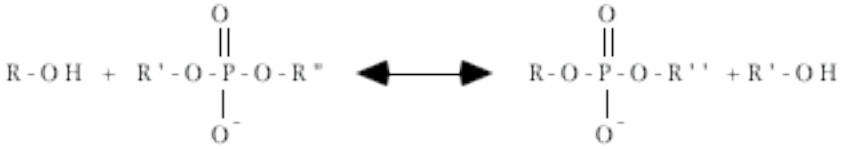
•Group I intron - 1st ribozyme discovered

- Tetrahymena 26S rRNA gene has ~400nt intron (IVS).

- Purified unspliced RNA to study splicing
  Added back cell extracts to catalyze reaction,
  but reaction occurred in control reaction without adding any proteins!
  - Protein covalently linked to RNA?
  - Purified RNA using SDS, phenol, protease digestion, urea gels.
    Still worked.
  - Cloned into E. coli using lac promoter plasmid - transcribed in vitro, purify --> RNA free of Tetrahymena cellular factors.
  - Still worked.
- Intron RNA itself is responsible for self-splicing
- Requires Mg^{++} (mM)and guanosine (uM) co-factors
  Reaction occurs under physiological conditions

Mechanism: transesterification, concerted making/breaking of phosphodiester bonds - bond energy conserved.

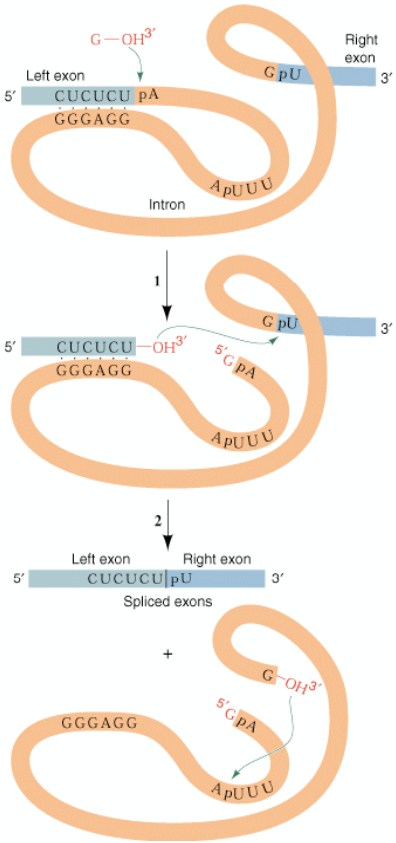
•G**OH** attacks 3' end of exon1.



for first step: R = guanosine, R' = 5'exon, R" = IVS+3'exon

for second step: R= 5'exon, R'= IVS, R"=3'exon

Isn't acting as a catalyst in vivo, intron is changed during the course of the reaction.



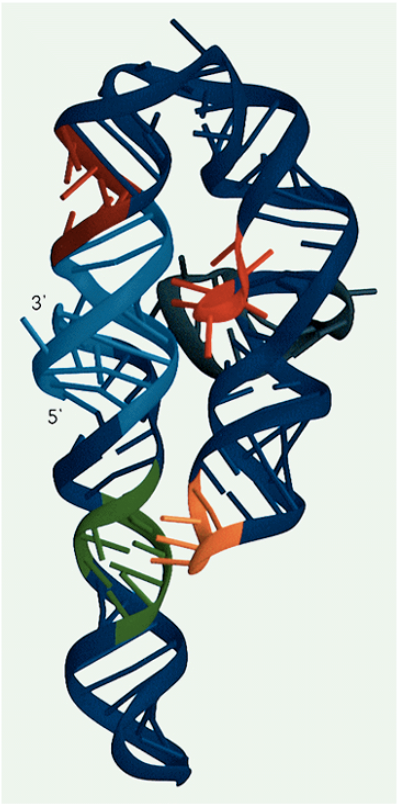


Figure 25-24. The X-ray structure of a portion of the catalytic domain of a self-splicing intron from *Tetrahymena thermophila*. [Courtesy of Jennifer Doudna, Yale University.]

But "L19" (intron final product) is able to perform a variety of catalytic reactions in vitro:

RNA endonuclease

RNA "polymerase" actually a disproportionation

Phosphatase

Turnover 2/min (slow) but k_{chem}= 400/min

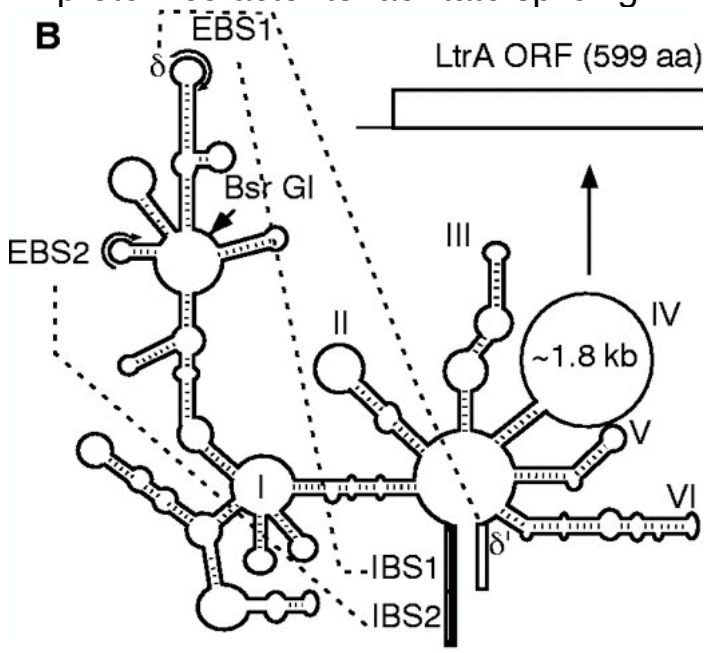
Group II Intron

- •Large RNA (6 domains)
- •Domain I binds exon1
- •Domain V, catalytic
  Catalyzes attack of A-2'OH of domain VI at exon1 site.
  --> lariat intermediate analogous to that of spliceosomal splicing.

Homology to spliceosome ==> RNA may be catalytic moiety of spliceosome.

Requires high salt in vitro

Some encode reverse transcriptase necessary for transposition, or protein cofactor to facilitate splicing.

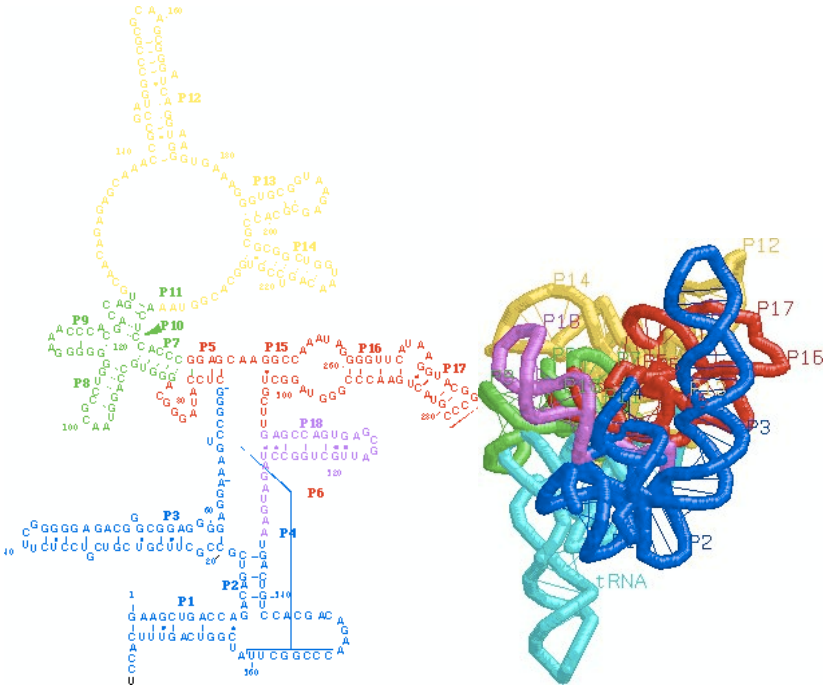


Ribonuclease P (RNase P)

- Cleaves 5'-precursor sequences from tRNA transcripts.
- Bacterial enzyme:
  Long known to have large RNA and small protein subunits- thought RNA was just playing structural role
  recognizes tRNA structure, not sequence
  Dissociate subunits: neither active alone in normal buffer (100 mM monovalent, 10 mM Mg^{++})
  But at 10X salt concentration RNA alone is active- protein isn't. RNA binds tRNA as well as holoenzyme

	K_M	k_{cat}
RNA alone	20 nM	0.5/min
holoenzyme	20 nM	10

==> RNA alone has all necessary groups for binding substrate specifically.



Cofactors:

Mg++
Salt to overcome repulsive forces of charged RNAs
Protein required in vivo.

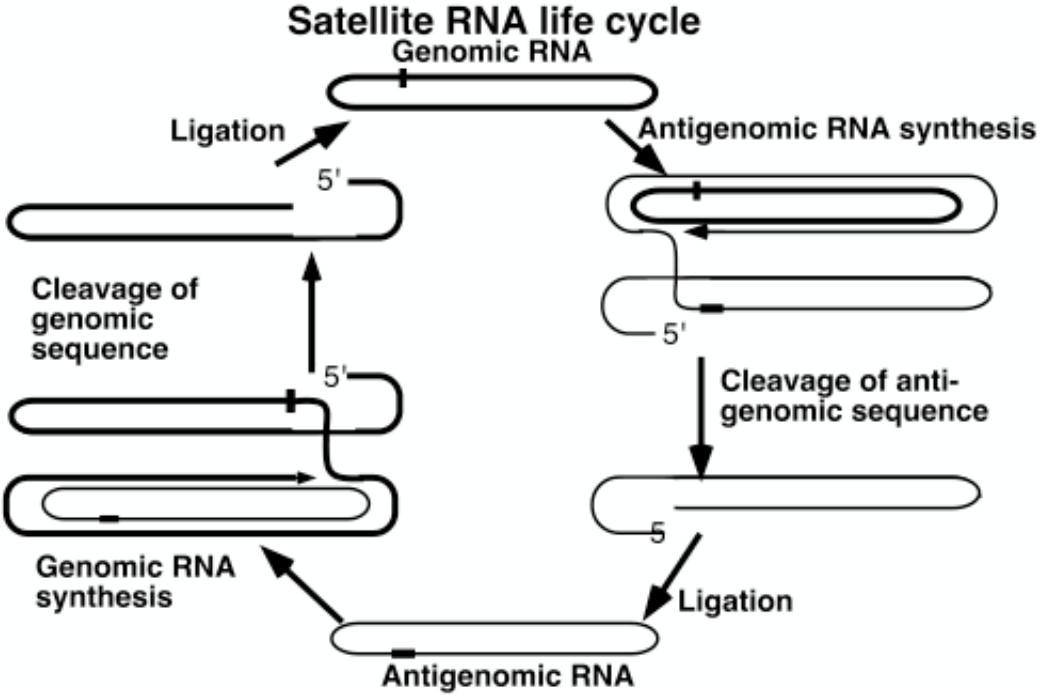
- Eukaryotes:
  protein required in vitro (10 subunits known in yeast).
  perhaps some binding or catalytic functions transferred to protein subunit - or -
  protein required for proper folding of RNA.
  cannot dissociate and reassociate and retain activity.

- ***Satellite RNA ribozymes:***

Viroids - Naked RNAs which infect plants (ASBV-avocado sunblotch)
Virusoid- satellite of plant RNA viruses encapsidated by viral proteins (sTRSV-tobacco ringspot; LSTV-lucerne transient streak).

Both have ssRNA genomes-

can isolate genomic and antigenomic monomers and linear multimers. ==> rolling circle replication.



Found that multimer RNA can self-cleave to form linear monomers with 5'-OH and 2'-3' cyclic phosphate ends.

Linears can self-ligate to form circles.

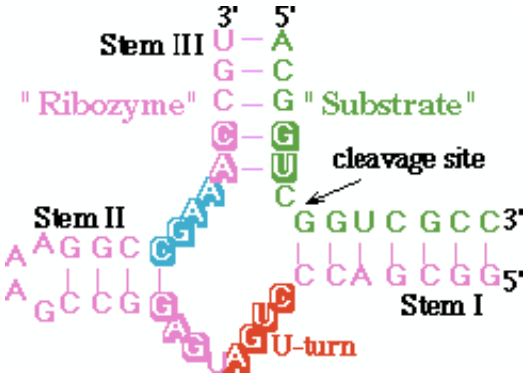
In vivo conformation not very active, but heat denature, then quick cool --> good activity.

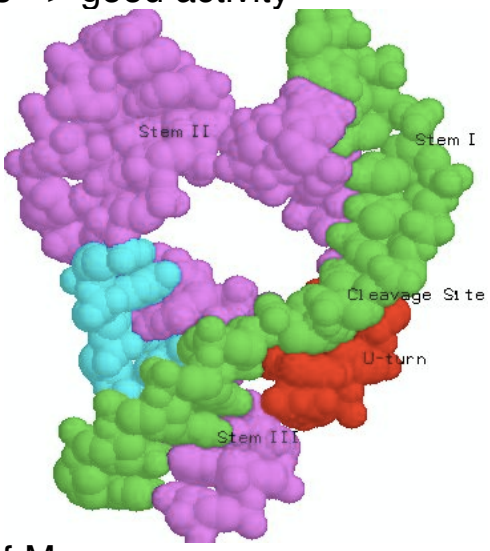
" " slow cool --> low activity.

==> most stable conformation inactive.

Hammerhead

Found similar sequence motif around cleavage site (hammerhead).
trimmed away excess sequences --> good activity





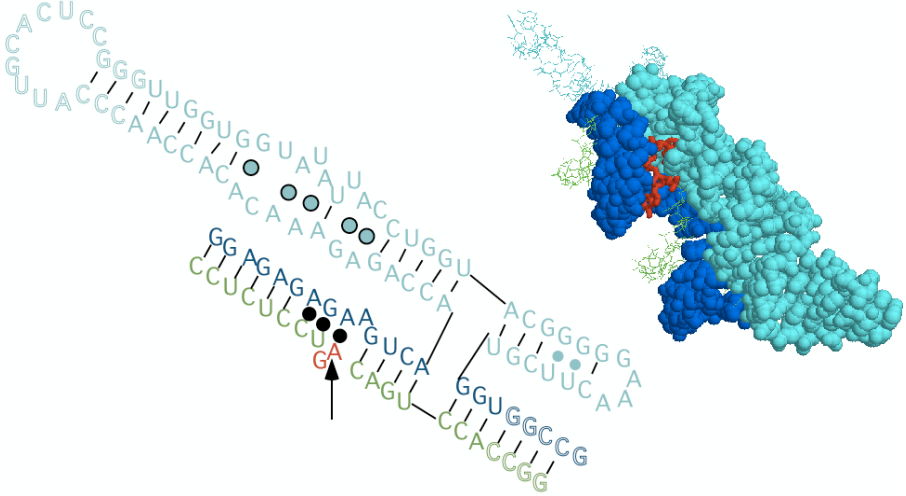
Cleaves in presence of low conc. of Mg
(can cleave during transcription)

Connectivity of strands unimportant for reaction.
Can reconfigure into "substrate" strand and "enzyme" strand.

--> true catalyst with few sequence constraints on substrate
smallest natural ribozyme, can use to cleave any sequence , in theory.
Heavily studied as possible therapeutic molecule to knock out
undesirable mRNA such as retroviruses, oncogenes.

Hairpin

Antigenomic strand of TRSV self-cleaves, but doesn't fit hammerhead motif. Looks more like a hairpin.
Can be dissected into two domains:
Loop A - substrate strand and SBS (substrate binding site)
Loop B - catalytic domain (light blue).



Hepatitis □ virus

satellite of Hepatitis B in humans, causes accelerated liver damage in hep B carriers.
Life cycle, reaction products similar to hammerhead and hairpin, but doesn't fit either motif.
Can fold genome and antigenome into similar structure.
Self-religation hasn't been demonstrated.

Mg and ribozyme active site

Most ribozymes coordinate Mg at active site (hairpin doesn't require Mg)
Mg stabilizes transition state in transesterification or hydrolysis reaction

Why aren't there more ribozymes?

Nucleic acids: information poor (4 possible sidechains, not 20)
info on inside of the molecule, not outside
==> proteins can present greater variety of environments for catalysis

Using ribozyme (or aptamers) to cure your favorite disease

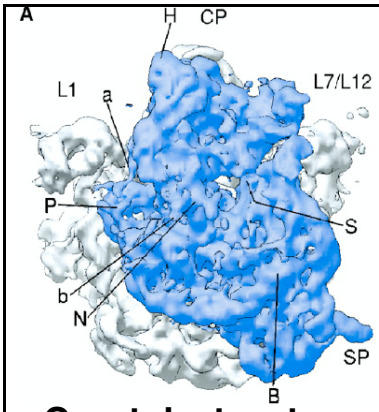
Problems:
Delivery to cells
Stability of ribozyme
Accessibility of target
Turnover

The "RNA World"

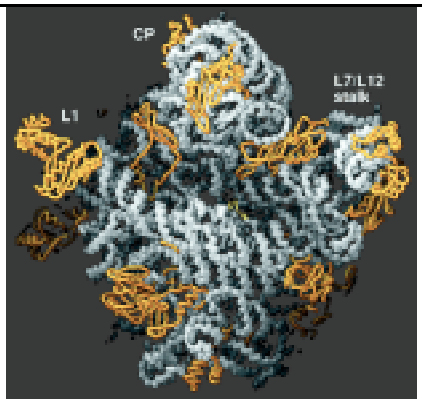
How did life begin?
RNA is a good candidate for first molecule that could self-replicate
Has genetic info and catalytic functions on same molecule.
Once ribosome available for translation, then could replicate proteins.

Ribosome

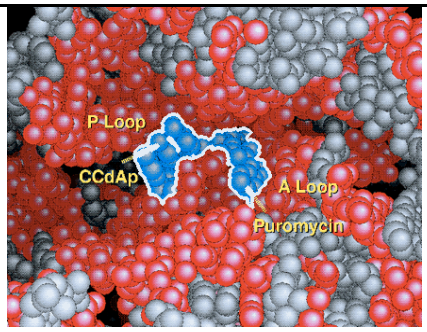
Large, highly conserved RNAs plus many proteins.
Cannot isolate pure rRNA and get any activity,
But protease treat ribosome --> some aa-tRNA cleaving activity remains.
But at least one protein still intact. Are proteins necessary for activity or for proper folding?



Crystal structure of intact ribosome
Large subunit – gray
Small subunit – blue



Structure of Large (catalytic) subunit
Proteins – gold
RNA – gray, active site in green in center.



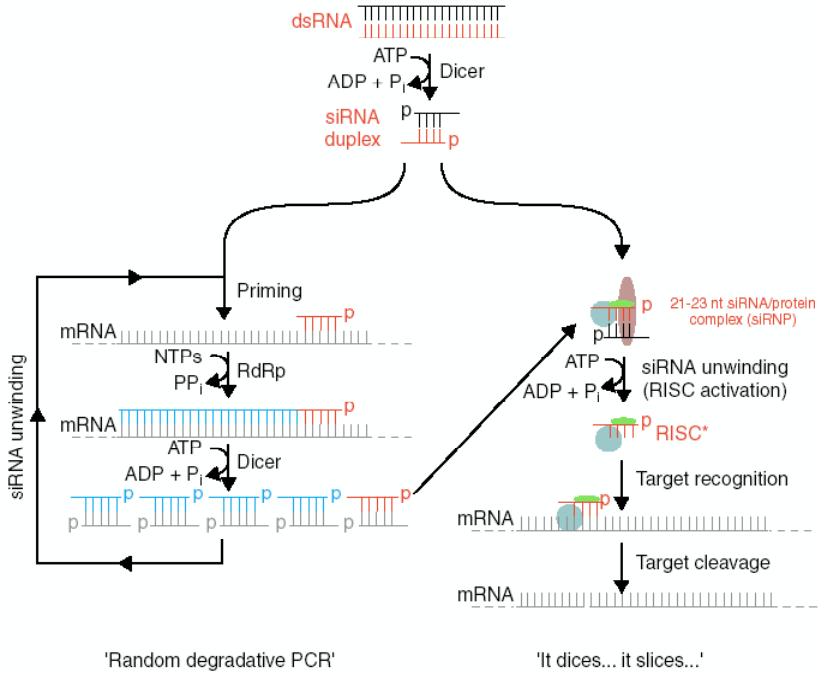
Close-up of active site
Substrate analog in blue,
Conserved RNA sequences in red

Crystal structure of active site with substrate analog -> no protein near active site!
So the ribosome is a ribozyme. First ribosome was made of RNA, proteins added later.

RNA interference (RNAi)

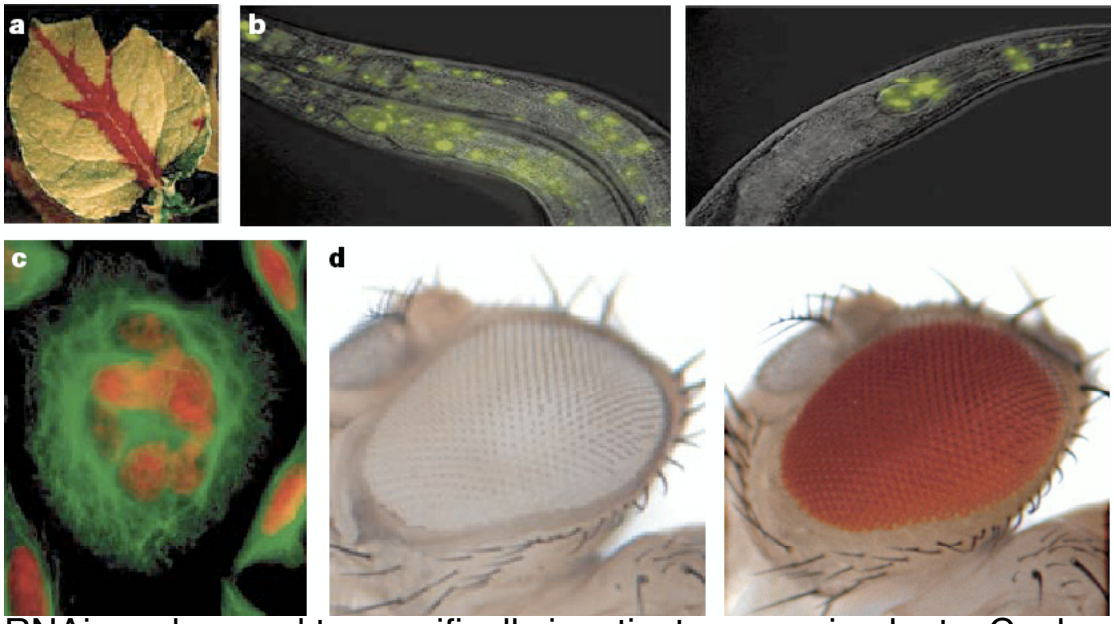
Ability to synthesize and express RNAs led to attempts to inactivate genes by expression of *antisense* RNA (opposite strand from mRNA coding)

In some cases, this worked. But once again, control reactions gave an important, surprising result: Using *sense* RNA also inactivated!
Others noted that overexpression in plants of mRNA resulted in gene inactivation.
Using dsRNA resulted in 10-fold more inactivation.
So it appeared that dsRNAs could direct inactivation of mRNAs containing the same sequence.



Current Opinion in Genetics & Development

The dsRNA is digested into 21-25 nt long RNAs by the Dicer protein a type of RNase III). These RNAs (small interfering RNAs – siRNA) are unwound and can pair with complementary mRNA.
In some organisms, at least, this paired RNA serves as a primer for an RNA-dependent RNA polymerase (RdRP).
The RNA is copied into dsRNA, which serves as substrates for further cleavage.
This “degradative PCR” amplifies the original input RNA, so the cleavage can be propagated and inherited in C. elegans.



RNAi can be used to specifically inactivate genes in plants, C. elegans, mammalian cells, and Drosophila. RNAi can be either synthesized *in vitro* and applied to cells (worms eat it), or expressed endogenously *in vivo* using a promoter.

Why do cells do this??

The natural role for Dicer and other enzymes involved in RNAi appears to be in regulation of expression of specific mRNAs.

Cells produce short hairpin RNAs called stRNAs (small temporal RNAs), which use the same protein complex to bind to mRNAs and inactivate them, although they are not cleaved.

Cells lacking key enzymes in the RNAi pathway have multiple developmental abnormalities, due to misregulation of the target RNAs.