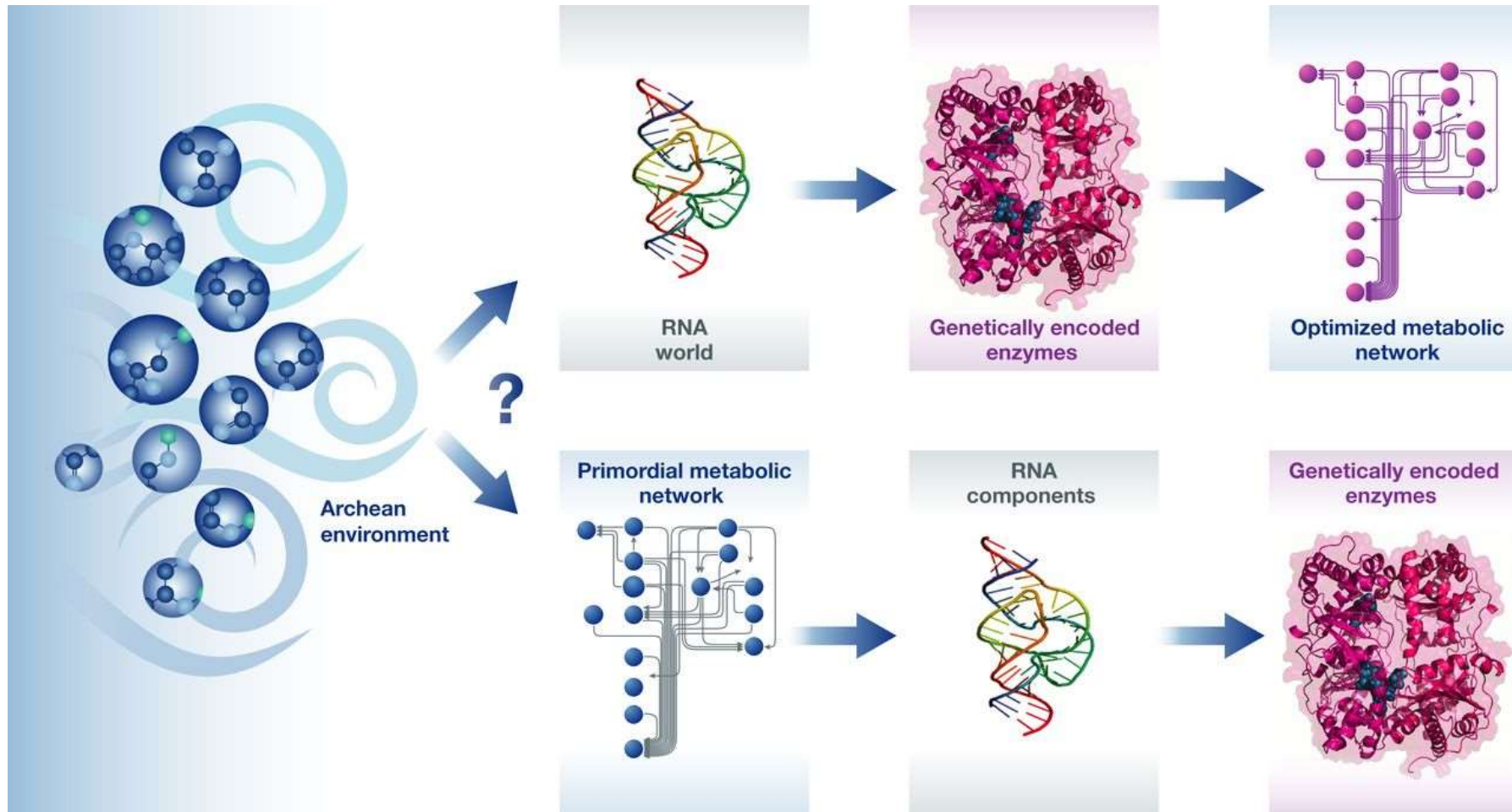


## Route to life by chemical networks



## *Metabolism-first vs. Genes-first*

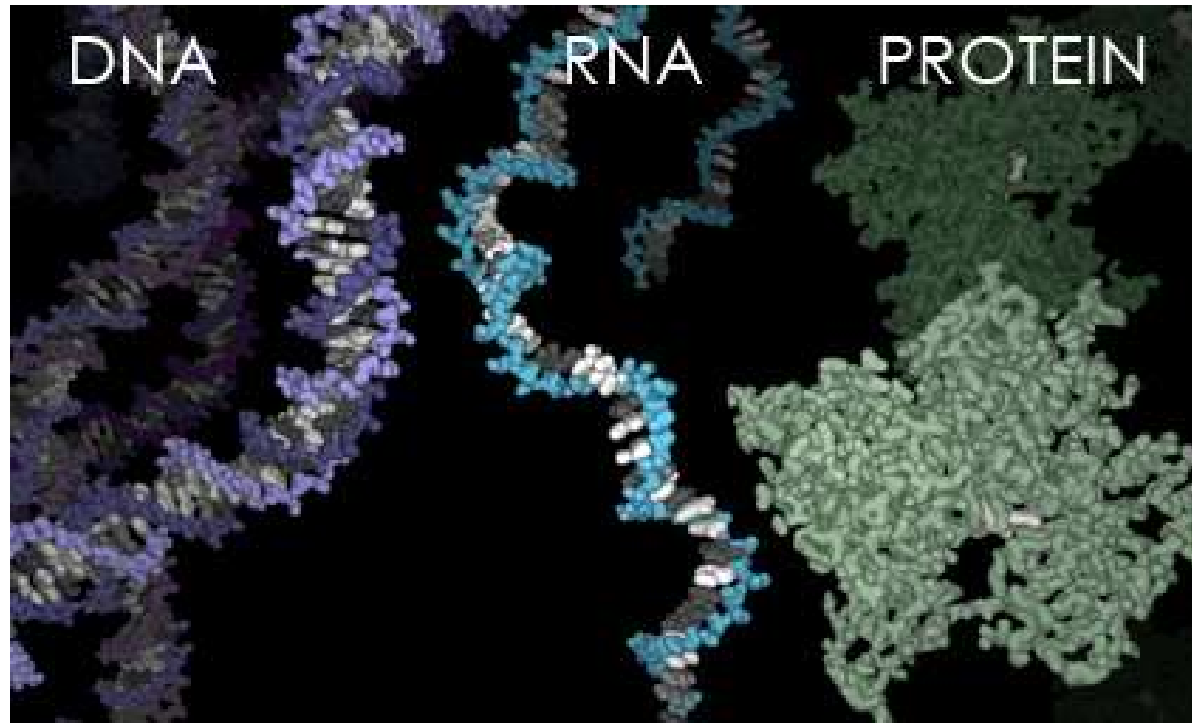
*Genetics/replication-first:* an information-carrying polymer capable of replication (RNA or something simpler) spontaneously arose from available prebiotic molecules available on early Earth. Metabolism incorporated later as a mean to receive energy from the surroundings in a controlled manner.

*Metabolism-first:* primitive metabolic cycles spontaneously assembled from simple prebiotic organic molecules or inorganic carbon sources as CO<sub>2</sub>. And the cycles produced a set or more or less complex molecules needed for the replication process and construction of the genetic apparatus.

The supposed *proto-metabolism* would differ from the currently known one, because the chemical reactions were not catalysed by efficient enzymes, nor were aminoacid and peptide sequences determined by DNA.

The involved reactions were either spontaneous, or catalysed by inorganic catalysts or peptides. Inorganic catalysts would be molecules, or ions, in solutions or on surfaces of solids such as clays or pyrites. Peptides (or peptoids) formed either by random oligomerization or mutual catalysis.

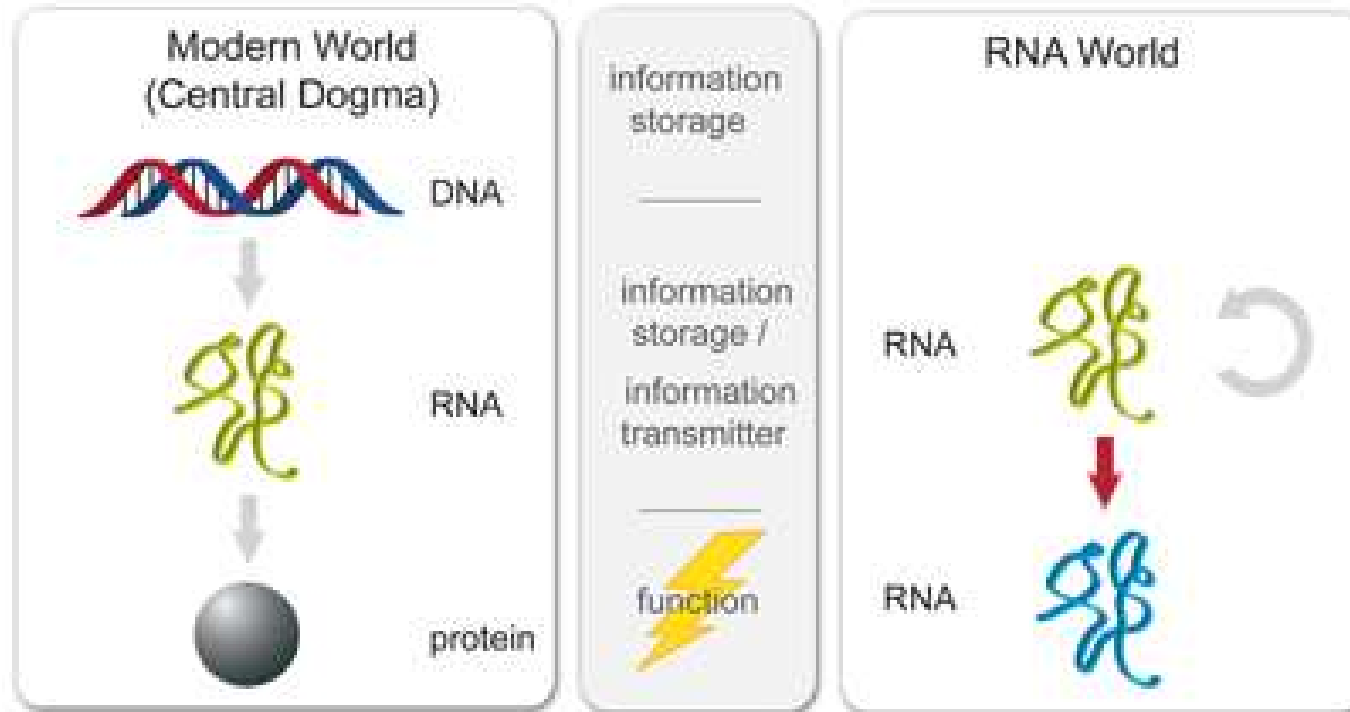
## *„Genes-first”*



*In modern cells, RNA (light blue, center) is made from a DNA template (purple, left) to create proteins (green, right).*

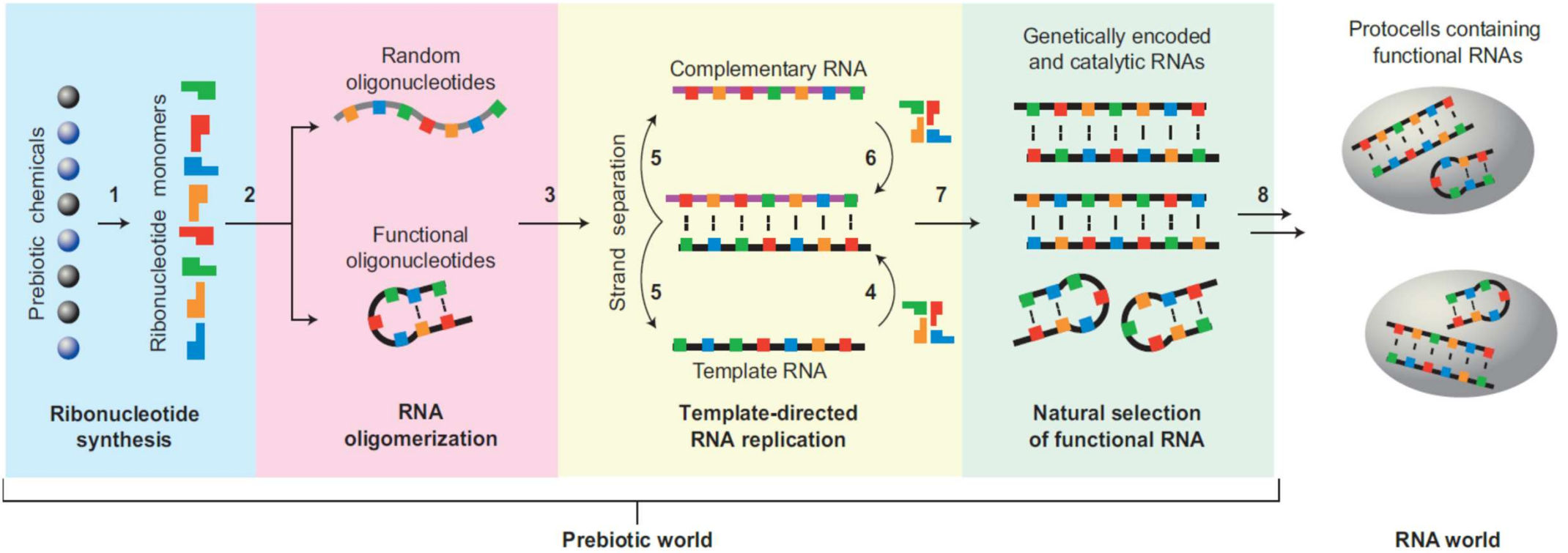
*RNA folding is mediated by base-pairing interactions along different regions of a single-stranded RNA.*

## The RNA world

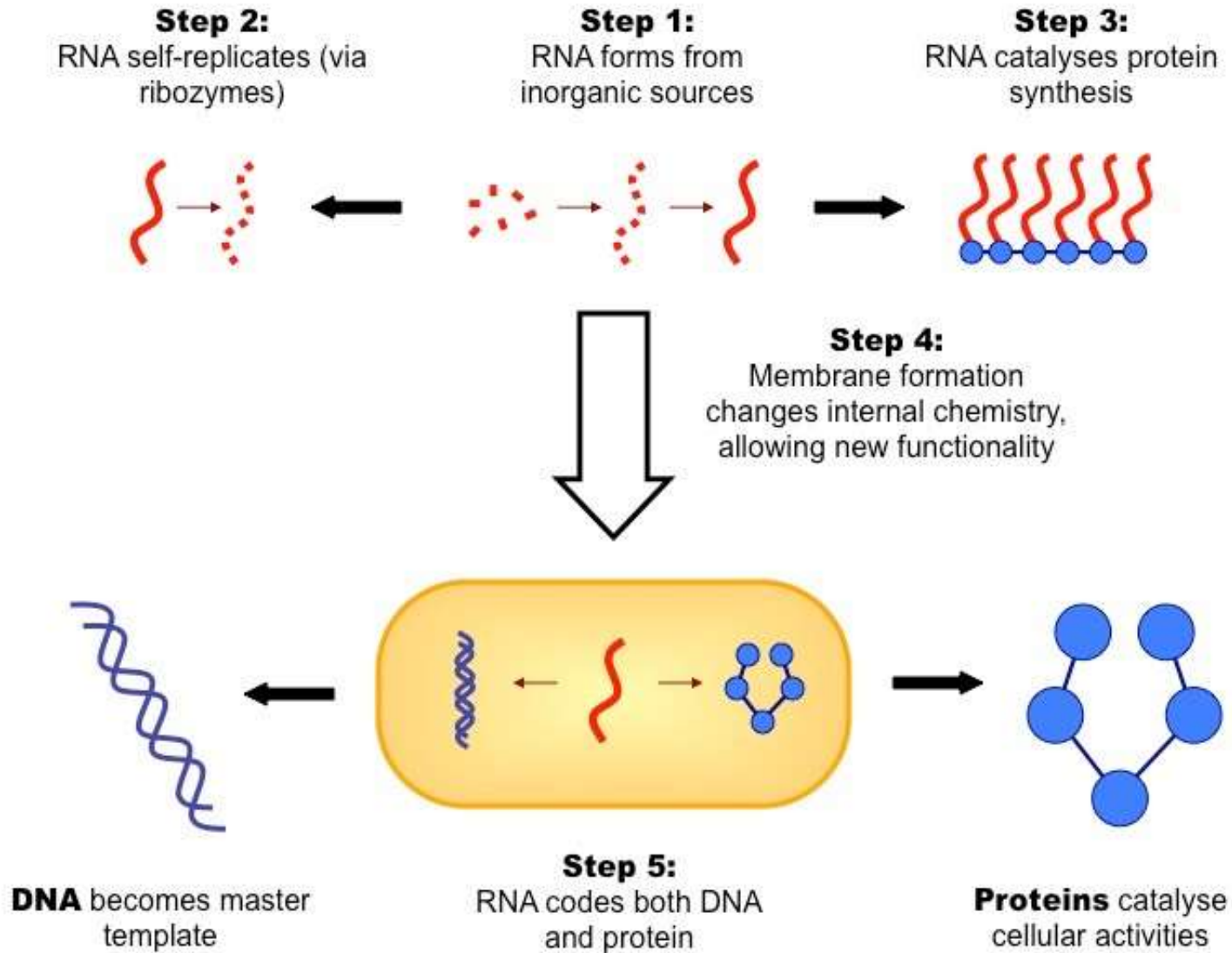


Conceptual idea that there was a period in the early history of life on Earth when RNA (or its structurally simplified analogue) carried out most of the information processing and metabolic transformations needed for biology to emerge from chemistry

# The RNA world



## The RNA world

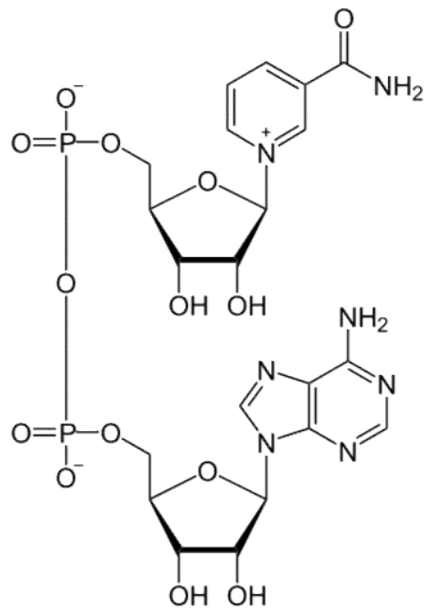


# The RNA world

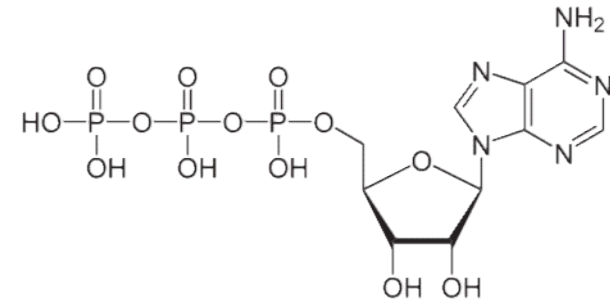
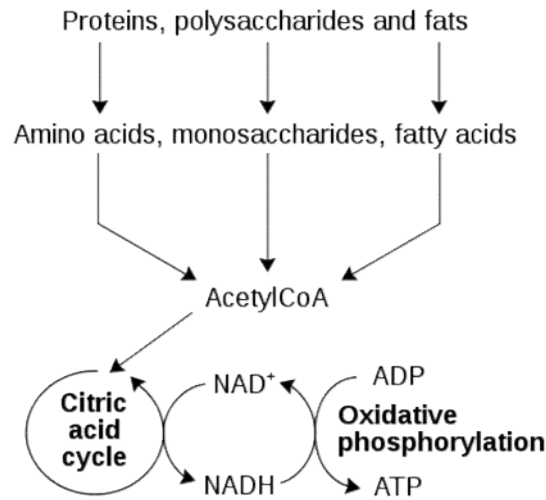
Crick, Orgel and Woese speculated in 1968 that, because RNA can form secondary structures, it has both a genotype and a phenotype and is a good candidate for the emergence of life

F. H. C. Crick *J. Mol. Biol.* **1968**, *38*, 367-379, L. E. Orgel *J. Mol. Biol.* **1968**, *38*, 381-393

Ribonucleotide coenzymes currently used by many proteins may be molecular „fossils” from the primordial RNA-based metabolism



**Nicotinamide adenine dinucleotide (NAD<sup>+</sup>)**

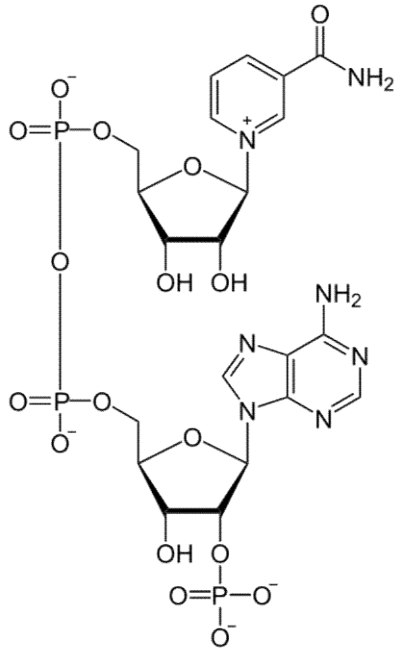


**Adenosine triphosphate (ATP)**

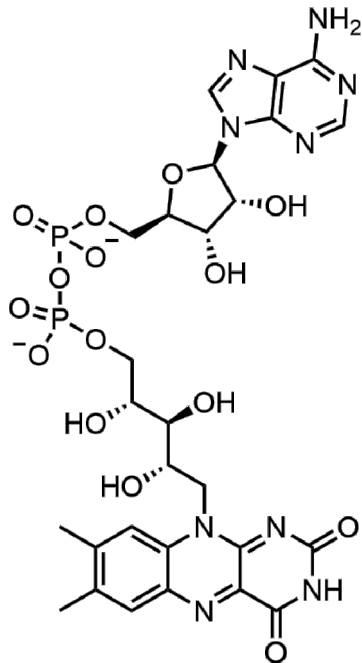
H. B. White III *J. Mol. Evol.* **1976**, *7*, 101-104

## The RNA world

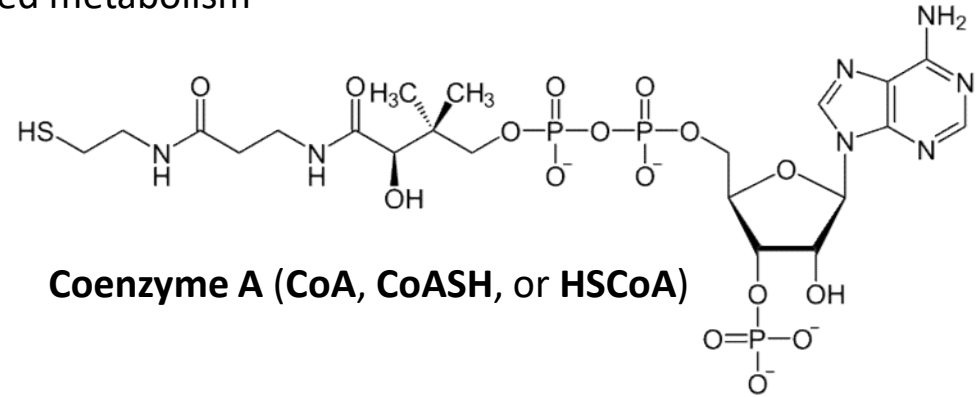
Ribonucleotide coenzymes now used by many proteins may be molecular „fossils” from the primordial RNA-based metabolism



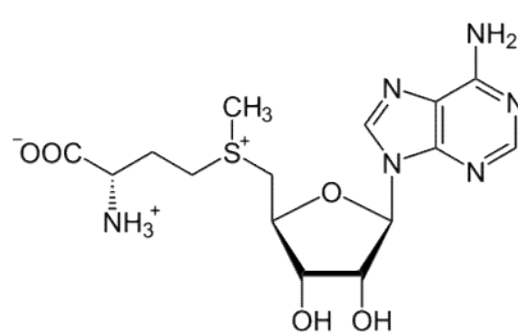
Nicotinamide adenine dinucleotide phosphate (NADP<sup>+</sup>)



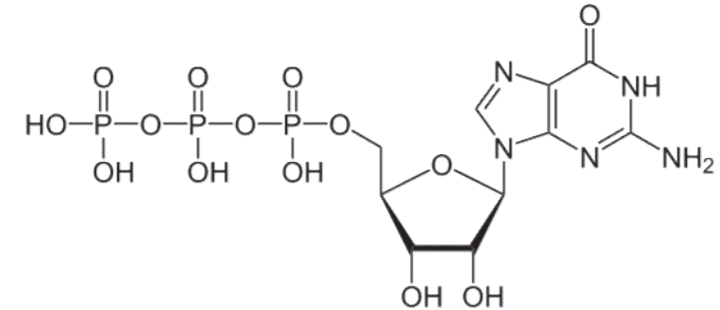
flavin adenine dinucleotide (FAD)



Coenzyme A (CoA, CoASH, or HSCoA)



S-Adenosyl methionine

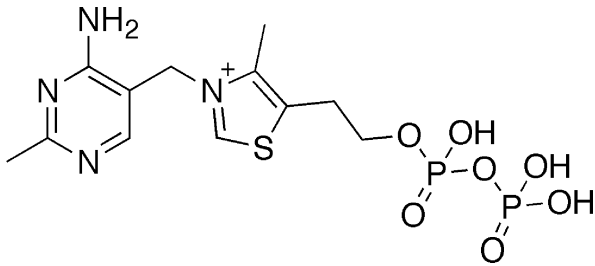


Guanosine-5'-triphosphate (GTP)

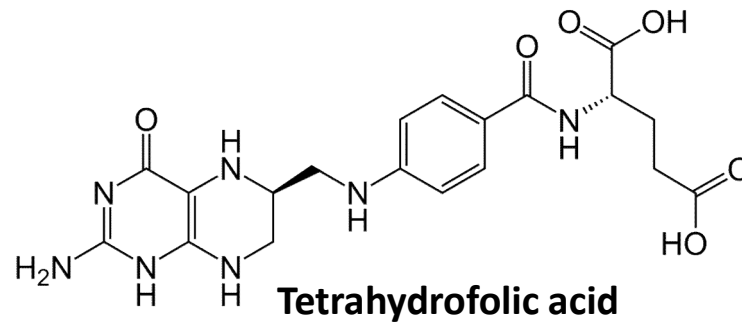


## The RNA world

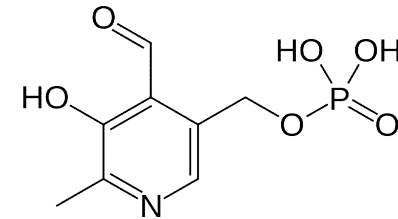
Other coenzymes contain cyclic nitrogen-containing bases that can also derive from nucleotides



**Thiamine pyrophosphate  
(TPP or ThPP) – Vit. B<sub>1</sub>**



**Tetrahydrofolic acid**



**Pyridoxal phosphate  
(PLP) – Vit. B<sub>6</sub>**

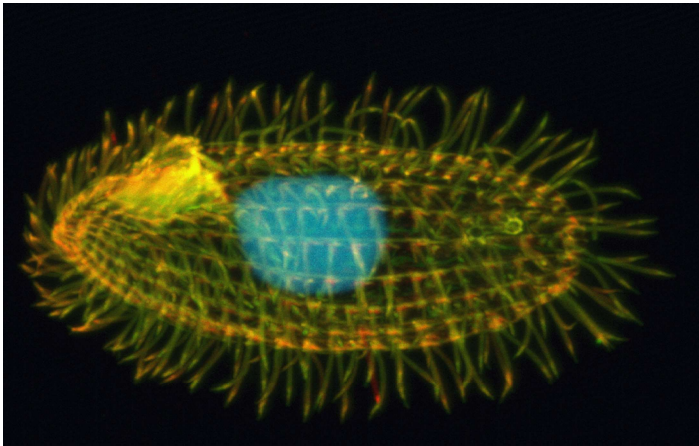
H. B. White III *J. Mol. Evol.* **1976**, 7, 101-104

# *The RNA world*

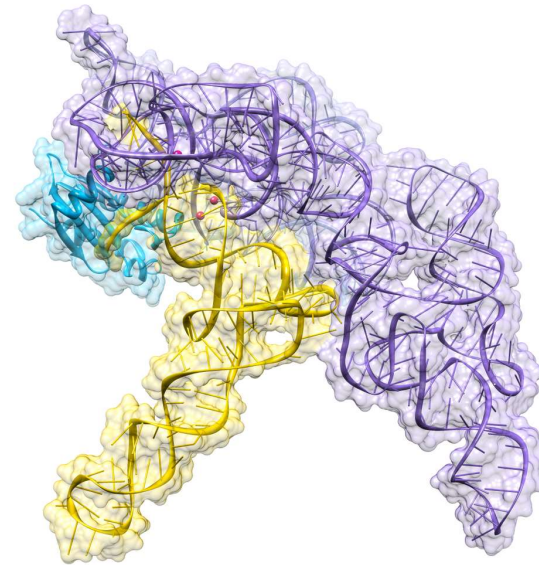
## *Ribozymes – Ribonucleic acid enzymes*

1989 – Thomas Cech and Sidney Altman – Nobel Prize in chemistry for discovery of catalytic RNA

Thomas R. Cech was studying RNA splicing in the ciliated protozoan *Tetrahymena thermophila*  
Sidney Altman and Norman Pace were studying the bacterial RNase P complex.



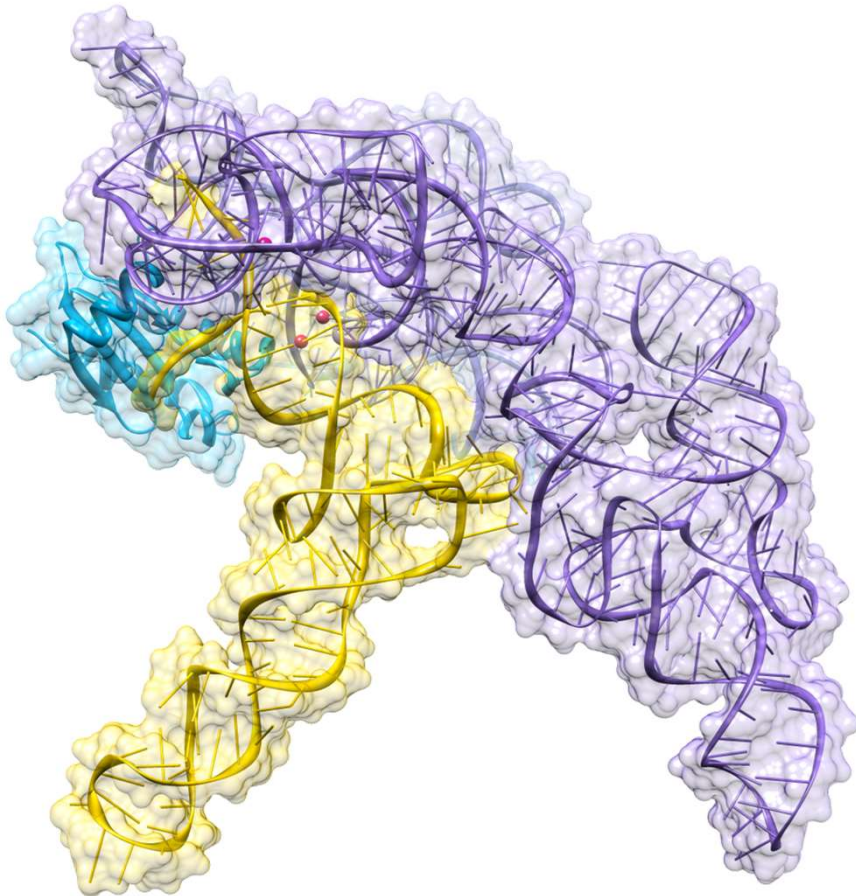
*Tetrahymena thermophila*



*Bacterial RNase P*

## The RNA world

### Ribonuclease P

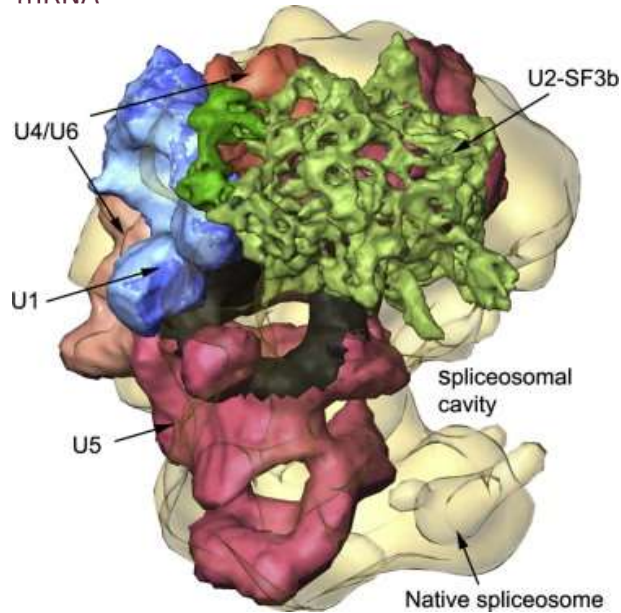
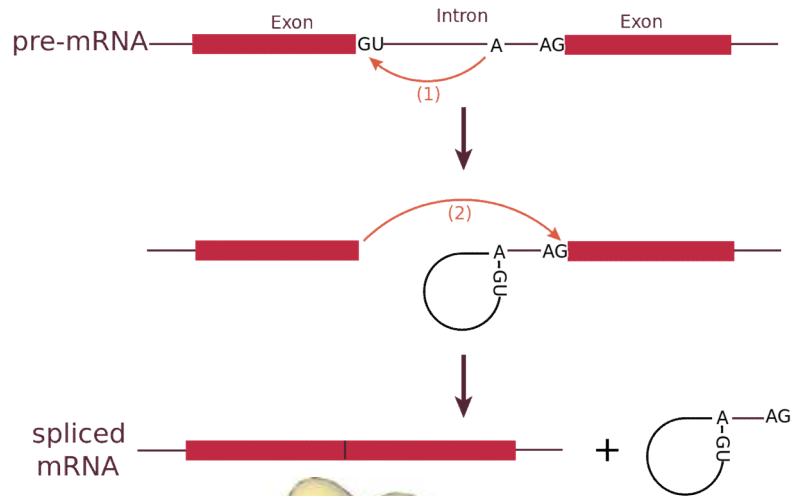


Ribonuclease P (RNase P) is a type of ribonuclease which cleaves RNA. RNase P is unique from other RNases in that it is a ribozyme – a ribonucleic acid that acts as a catalyst in the same way that a protein based enzyme would. Its function is to cleave off an extra, or precursor, sequence of RNA on tRNA molecules.

Bacterial RNase P has two components: an RNA chain, called M1 RNA, and a polypeptide chain, or protein, called C5 protein. *In vivo*, both components are necessary for the ribozyme to function properly, but *in vitro*, the M1 RNA can act alone as a catalyst. The primary role of the C5 protein is to enhance the substrate binding affinity and the catalytic rate of the M1 RNA enzyme probably by increasing the metal ion affinity in the active site.

*Crystal structure of a bacterial ribonuclease P holoenzyme in complex with tRNA (yellow), showing metal ions involved in catalysis (pink)*

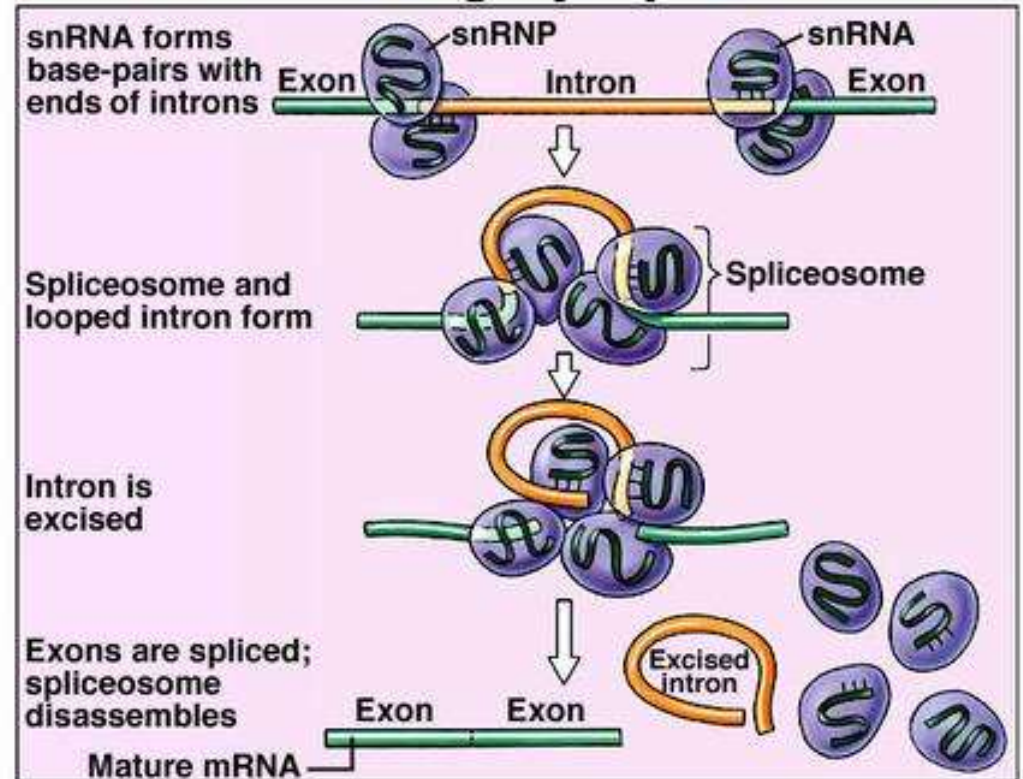
# RNA splicing



*Spliceosome* – a complex of ribonucleoproteins

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

## RNA Processing by Spliceosomes



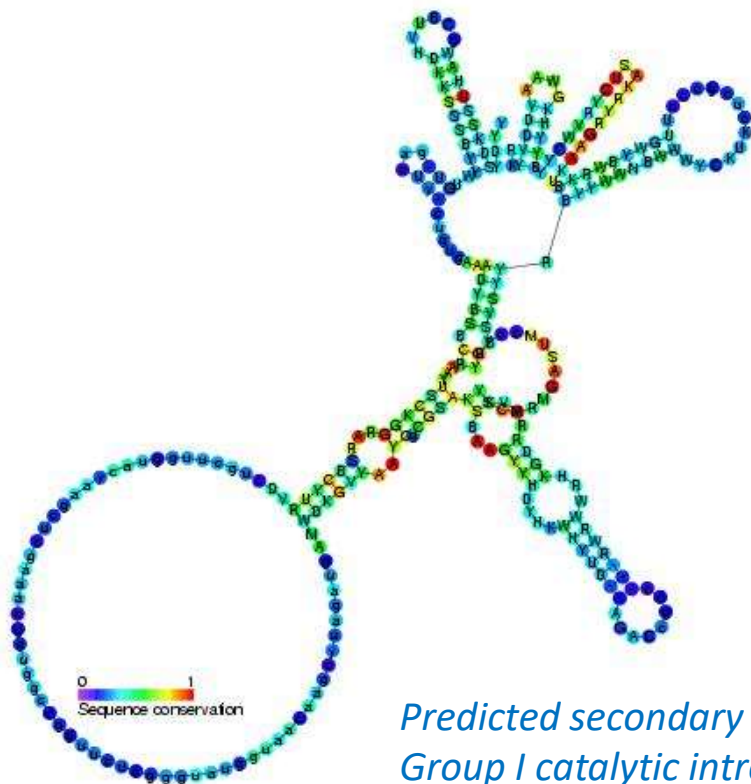
# RNA splicing

## Self-splicing RNA introns

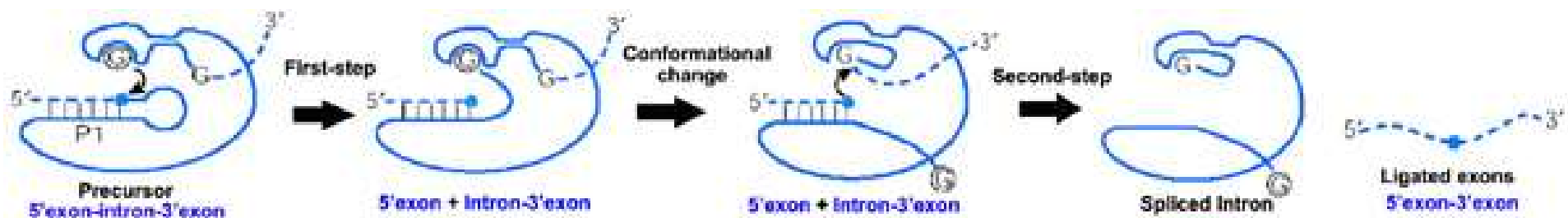
RNA splicing in *Tetrahymena* was taking place also in absence of the spliceosome - the 'negative control' obtained after protease digestion also spliced.

In contrary to the spliceosome, the **catalytic** motif **does not** contain protein part, **only RNA**.

First known example of a **ribozyme** – ribonucleic acid-composed enzyme analogue.



Predicted secondary structure and sequence conservation of Group I catalytic intron



# Ribozymes and riboswitches

## Hammerhead ribozyme

The hammerhead ribozyme is a RNA molecule motif that catalyzes reversible cleavage and joining reactions at a specific site within an RNA molecule (model system; targeted RNA cleavage experiments)

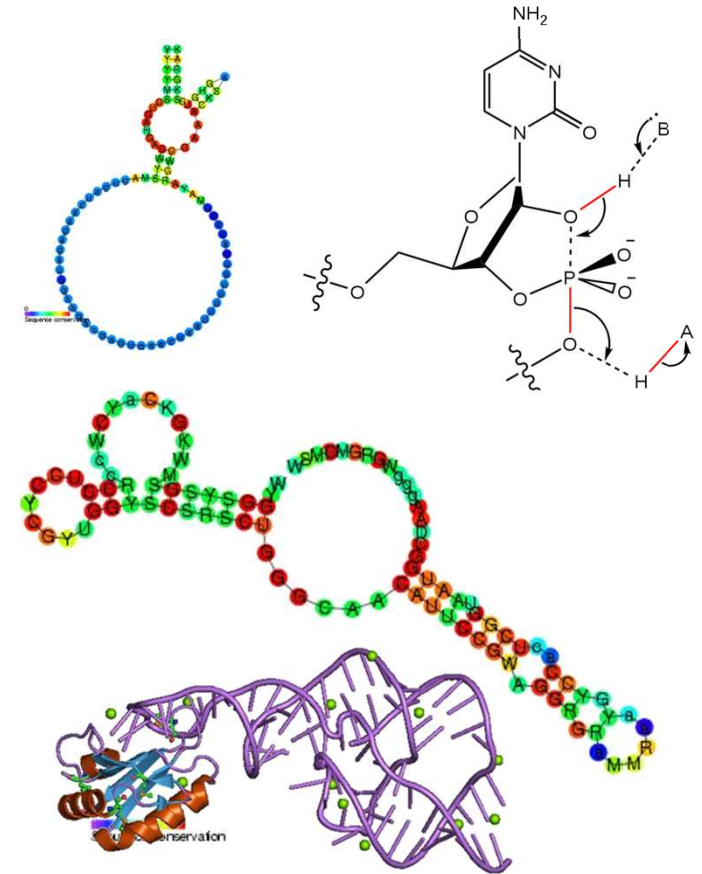
## HDV ribozyme

The hepatitis delta virus (HDV) ribozyme is a non-coding RNA found in the hepatitis delta virus that is necessary for viral replication and is thought to be the only catalytic RNA known to be required for viability of a human pathogen.

The ribozyme acts to process the RNA transcripts to unit lengths in a self-cleavage reaction. The ribozyme is found to be active *in vivo* in the absence of any protein factors and is the fastest known naturally occurring self-cleaving RNA.

## Riboswitches

A riboswitch is a regulatory segment of a messenger RNA molecule that binds a small molecule, resulting in a change in production of the proteins encoded by the mRNA (bacteria, TPP riboswitch also in plants and fungi)



## Riboswitches

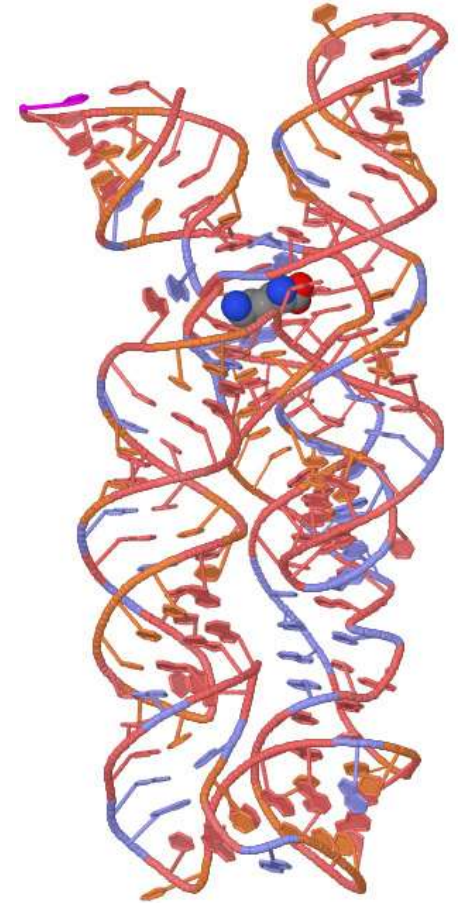
**2002** - (Breaker and Nudler) – discovery of a nucleic acid-based genetic regulatory element – *riboswitch*.

**Riboswitches** - naturally occurring regulatory segments of mRNA that bind small molecules specifically. The binding results in a change in production of the proteins encoded by the mRNA

Before discovery of *riboswitches* only *proteins* were supposed to do so in the biological context.

Most known *riboswitches* occur in bacteria, but functional riboswitches of one type (the TPP riboswitch) have been discovered in archaea, plants and certain fungi.

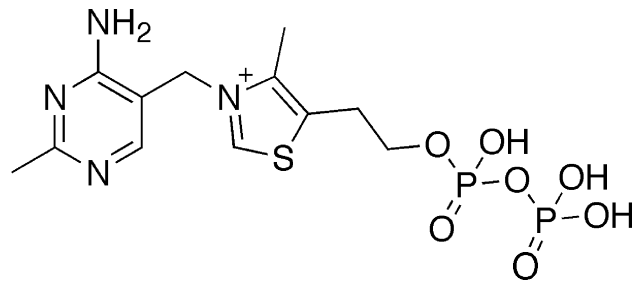
**Riboswitches** exist in all domains of life, and therefore are likely that they might represent ancient regulatory systems or fragments of *RNA-world ribozymes* whose binding domains remained conserved throughout the evolution



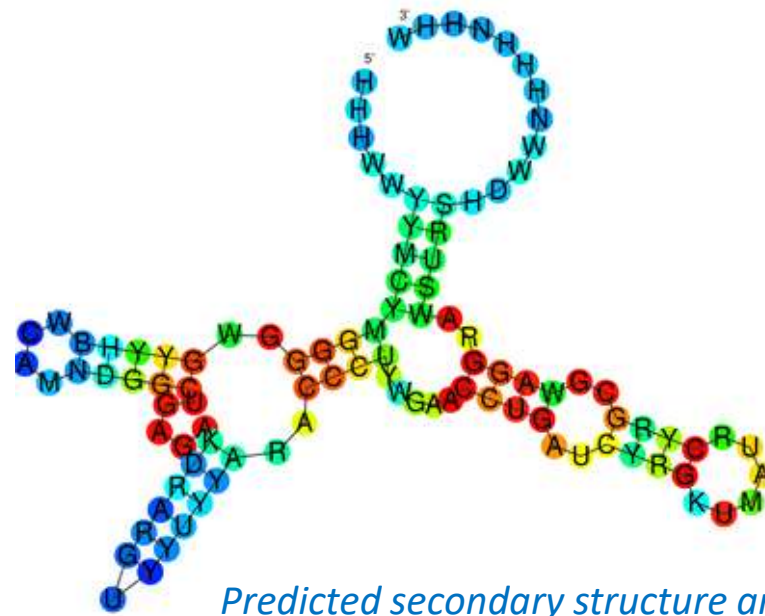
*The lysine riboswitch*

## The TPP Riboswitch

The **TPP riboswitch** (THI element and Thi-box riboswitch), is a highly conserved RNA secondary structure. It binds directly to thiamine pyrophosphate (TPP, a form of the vitamin B1, an essential coenzyme) to regulate gene expression through a variety of mechanisms in archaea, bacteria and eukaryotes.

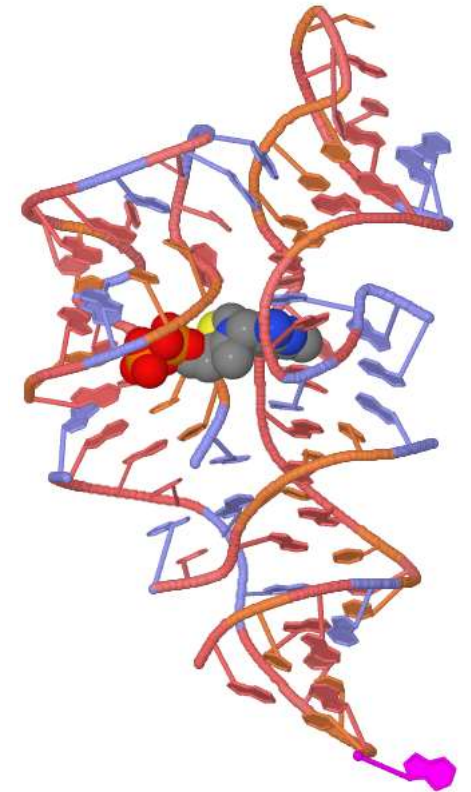


*Thiamine pyrophosphate TPP*



*Predicted secondary structure and sequence conservation of TPP riboswitch*

0 1  
Sequence conservation



*The 3D structure of TPP riboswitch (by Benjamin Schuster-Böckler)*

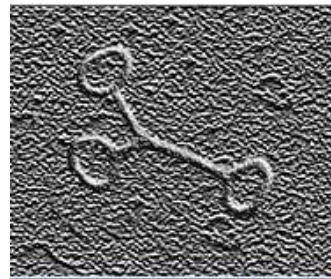


# Viroids

Viroids ("subviral pathogens,,) are mostly plant pathogens, which consist of short stretches of highly complementary, circular, single-stranded, and non-coding RNA without a protein coat. Viroids are extremely small - 246 to 467 nucleobases (genomes of smallest viruses start from 2,000 nucleobases). Viroids are plausible "living relics" of the RNA world.

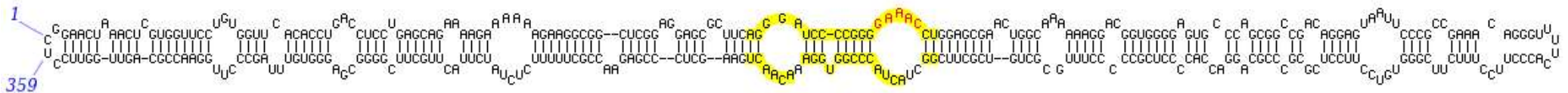
## Viroid properties:

- small size (error-prone replication)
- high G-C content, (stability and replication fidelity)
- circular structure (complete replication without genomic tags)
- lack of protein-coding ability, consistent with a ribosome-free habitat; and replication mediated in some by ribozymes—the fingerprint of the RNA world.



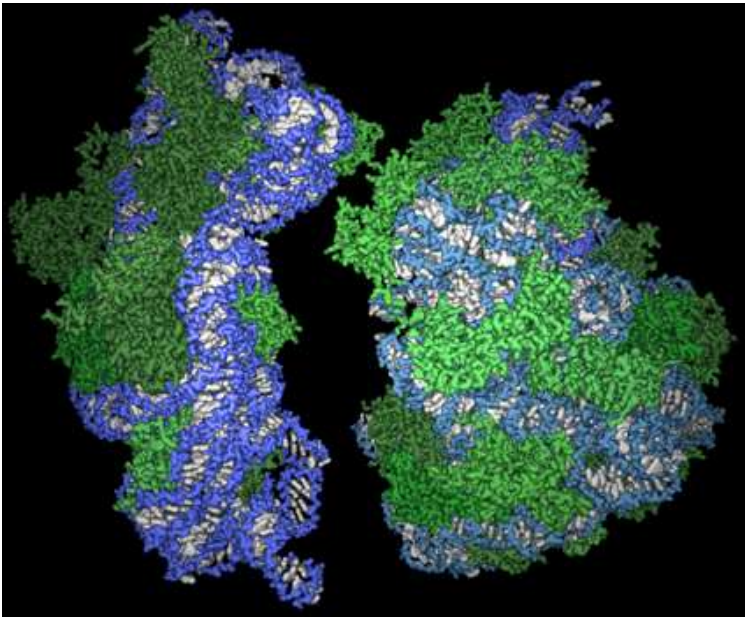
Copyright © 2006 Pearson Education, Inc., publishing as Benjamin Cummings.

*PSTVd-infected potatoes (right)*



*Putative secondary structure of the PSTVd viroid*

## Ribosome – the ,smoking gun’



Ribosome: green - proteins, blue and white - RNA

The **ribosome** is a **molecular machine**, found within all living cells, that serves as the site of biological protein synthesis (translation). Ribosomes link amino acids together in the order specified by messenger RNA (mRNA) molecules.

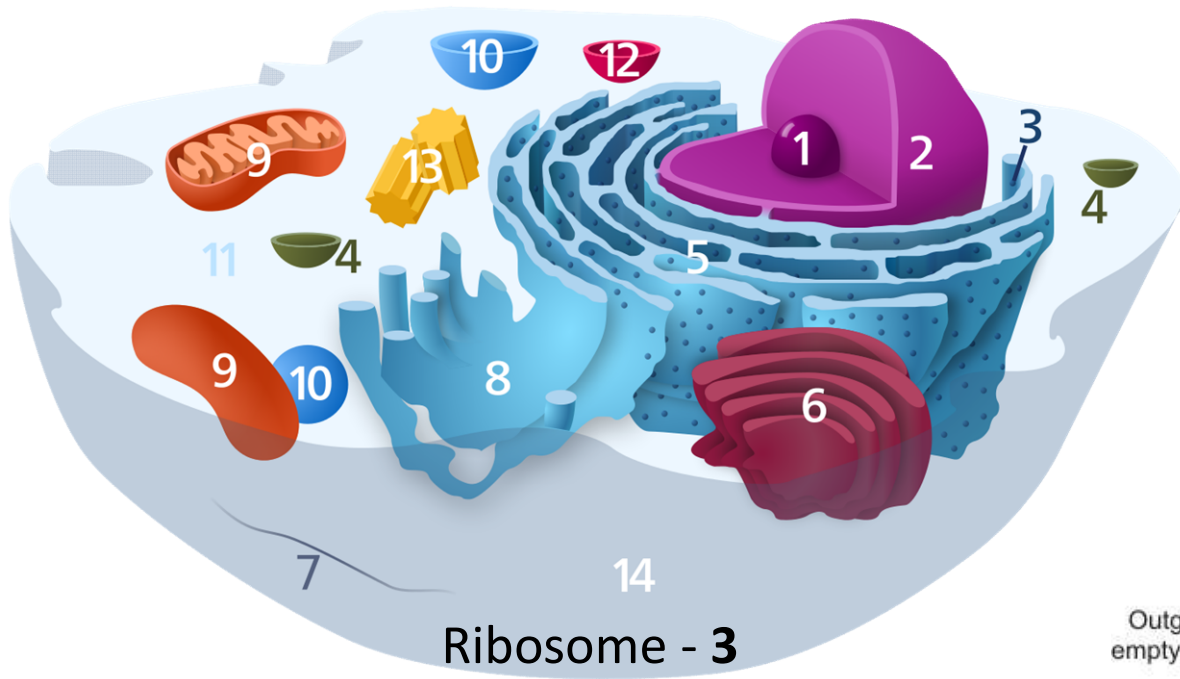
Ribosome is structurally highly conserved among all living species – most likely present in LUCA

### *Ribosomes:*

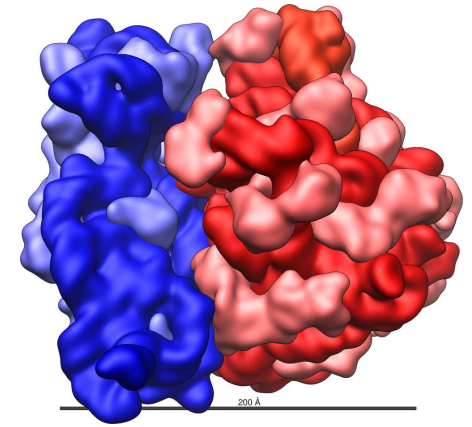
- the **small ribosomal subunit**, which reads the RNA
- the **large subunit**, which joins amino acids to form a polypeptide chain.

Each subunit is composed of one or more ribosomal RNA (rRNA) molecules and a variety of ribosomal proteins.

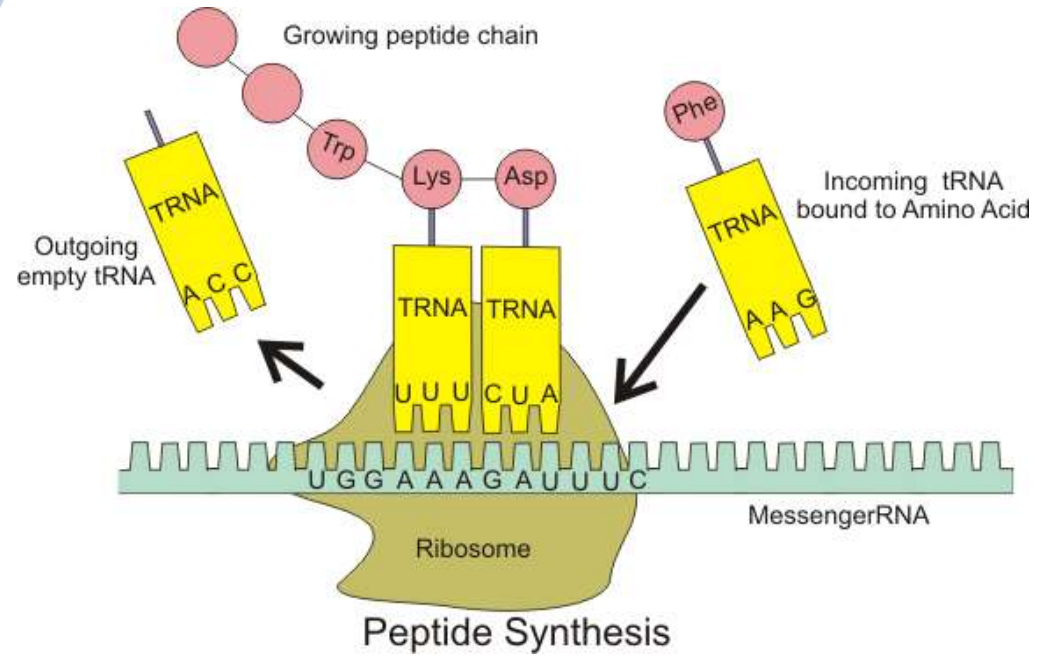
# Ribosome – the ,smoking gun’



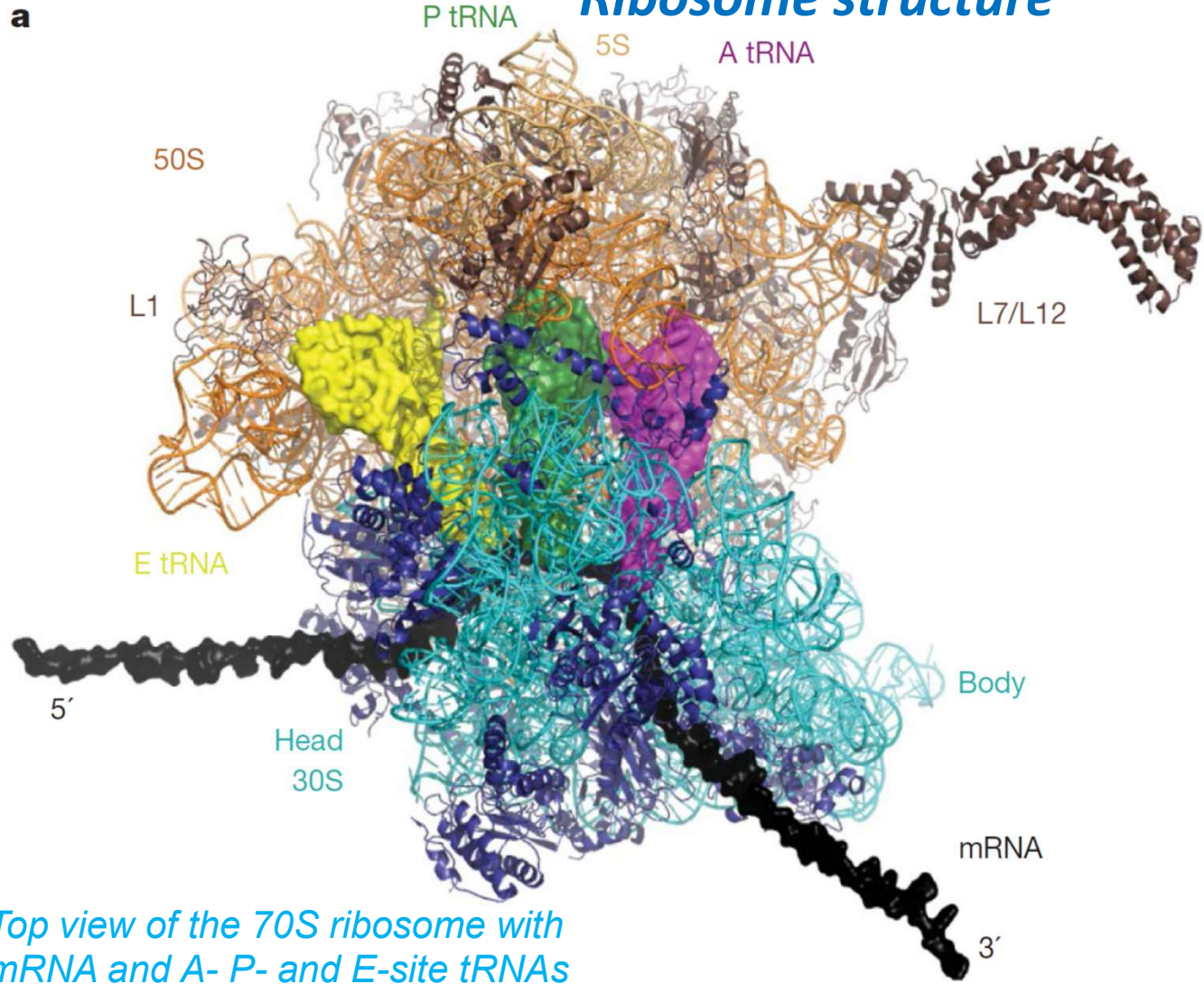
Ribosome - 3



Large and small subunit

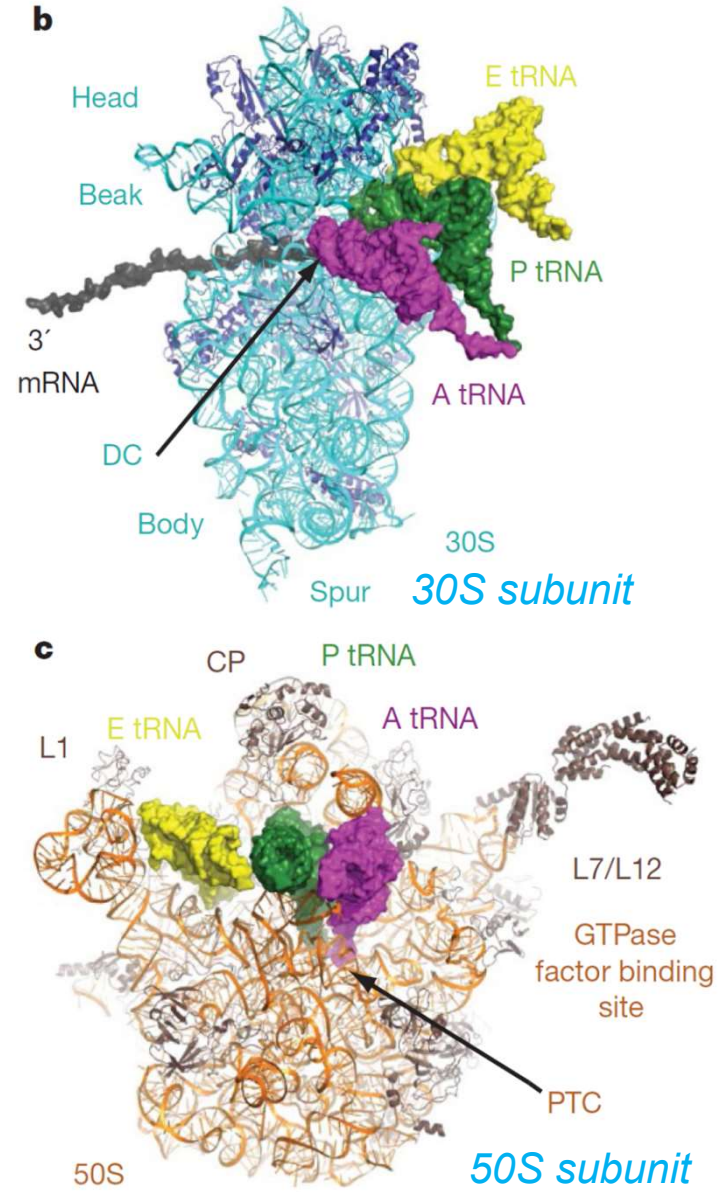


# Ribosome structure



Top view of the 70S ribosome with mRNA and A- P- and E-site tRNAs

T. Martin Schmeing, V. Ramakrishnan *Nature*. 2009, 461, 1234-1242

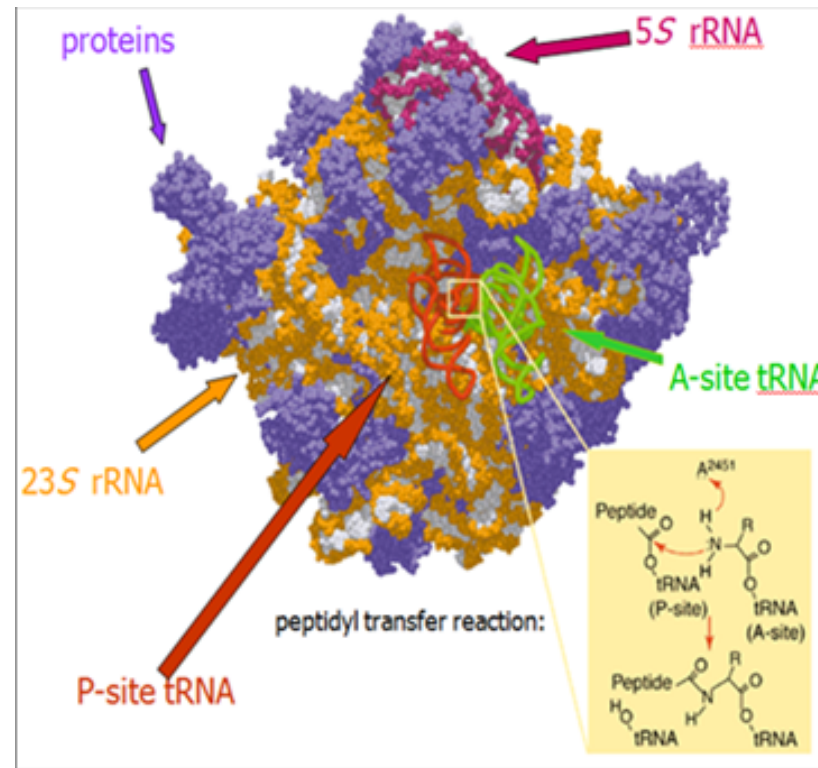


30S subunit

50S subunit

# Ribosome – the ,smoking gun'

*Ribosome is a ribozyme!*



No protein is present within 18 Angstroms from the active site → **proteins** play a structural role, but **DO NOT CATALYZE THE ACYL TRANSFER PROCESS**

T. Cech *Science*. 2000, 289, 878-879

## *Ribosome – the ,smoking gun'*

*Ribosome is a ribozyme!*

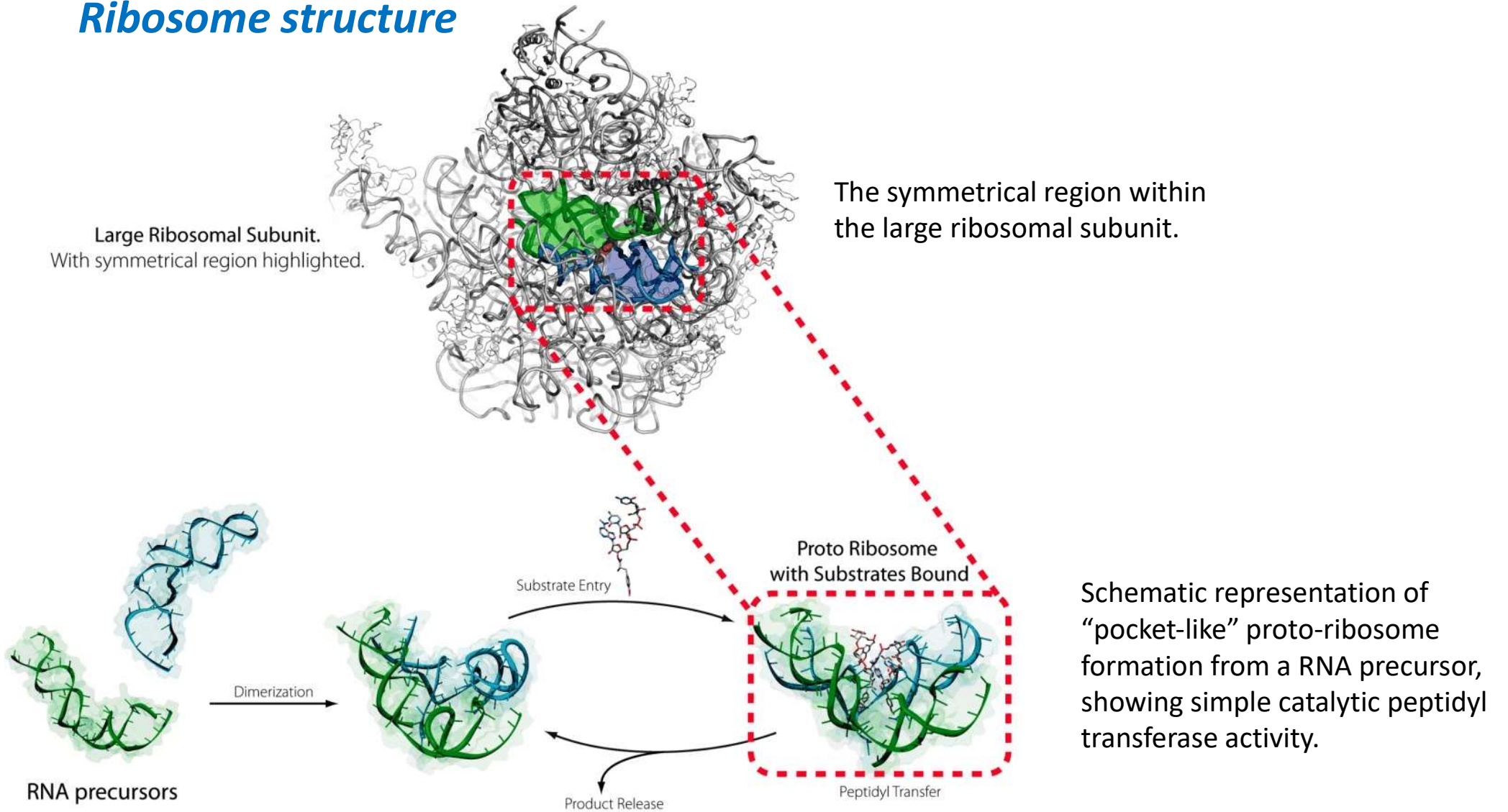
The ribosome may have first originated in an RNA world appearing as a self-replicating complex that only later evolved the ability to synthesize proteins when amino acids began to appear.

Studies suggest that ancient ribosomes constructed solely of rRNA could have developed the ability to synthesize peptide bonds.

In addition, evidence strongly points to ancient ribosomes as self-replicating complexes, where the rRNA in the ribosomes had informational, structural, and catalytic purposes because it could have coded for tRNAs and proteins needed for ribosomal self-replication.

As amino acids gradually appeared in the RNA world under prebiotic conditions, their interactions with catalytic RNA would increase both the range and efficiency of function of catalytic RNA molecules. Thus, the driving force for the evolution of the ribosome from an ancient self-replicating machine into its current form as a translational machine may have been the selective pressure to incorporate proteins into the ribosome's self-replicating mechanisms, so as to increase its capacity for self-replication

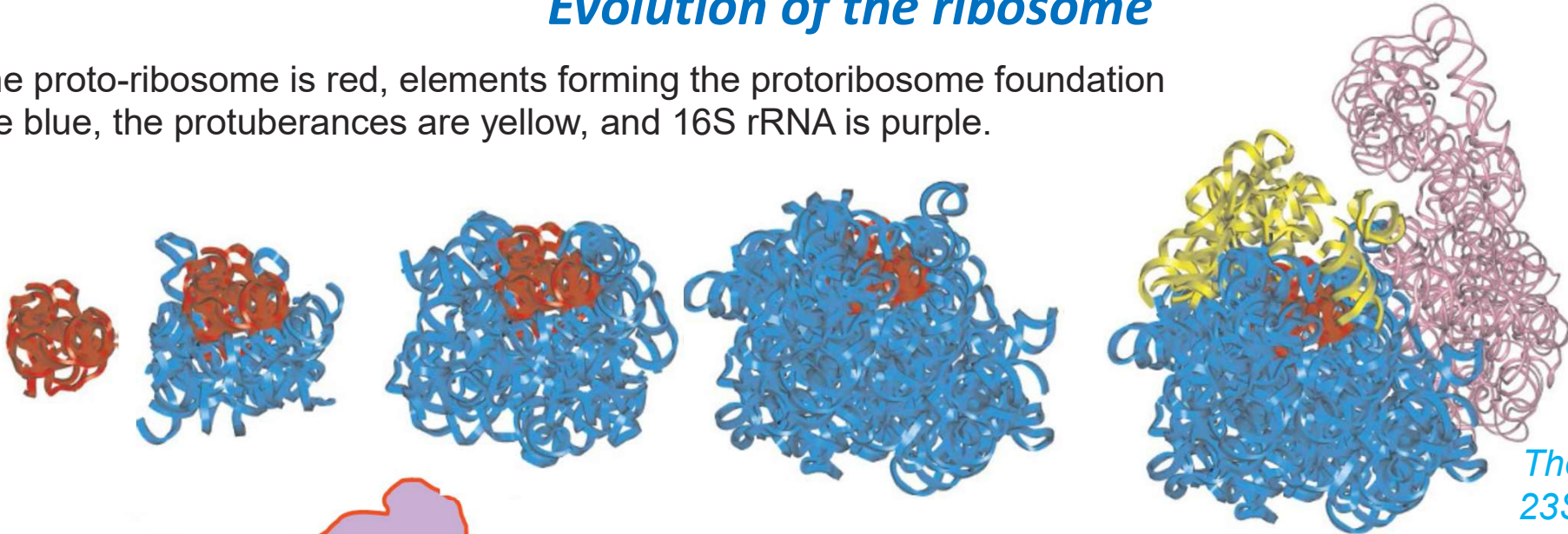
# Ribosome structure



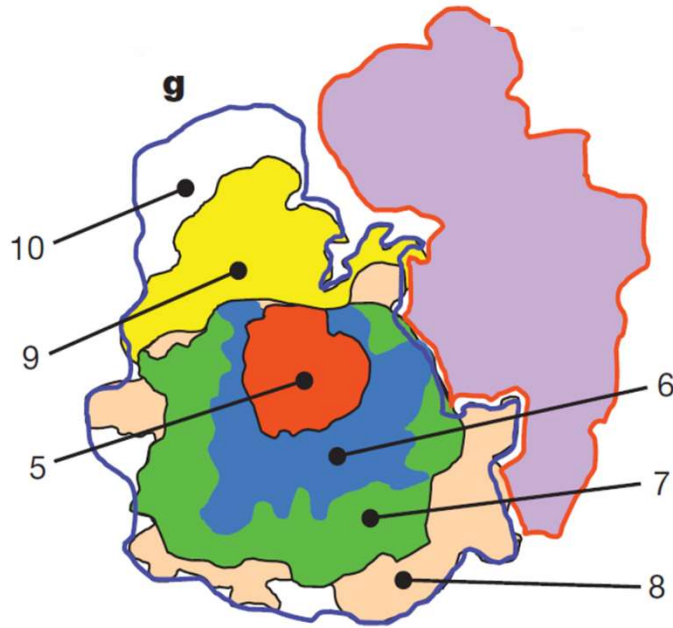
A. Yonath *et al.* *Israel Journal of Chemistry*, 2010, 50, 29-35

## Evolution of the ribosome

The proto-ribosome is red, elements forming the protoribosome foundation are blue, the protuberances are yellow, and 16S rRNA is purple.



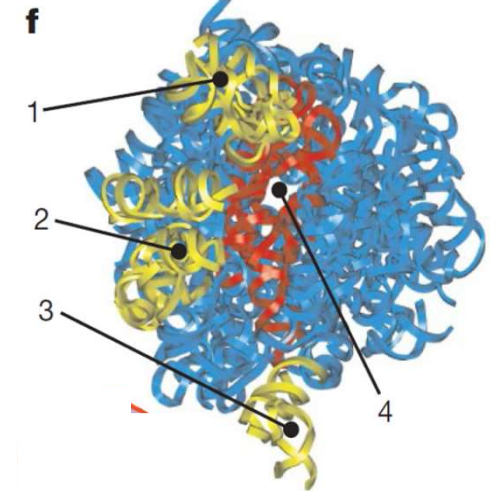
The top view of the 23S rRNA structure shown above



The positions of the parts of 23S rRNA shown above in the context of the whole ribosome. The structures of the 50S and 30S subunits are contoured by the blue and red line, respectively.

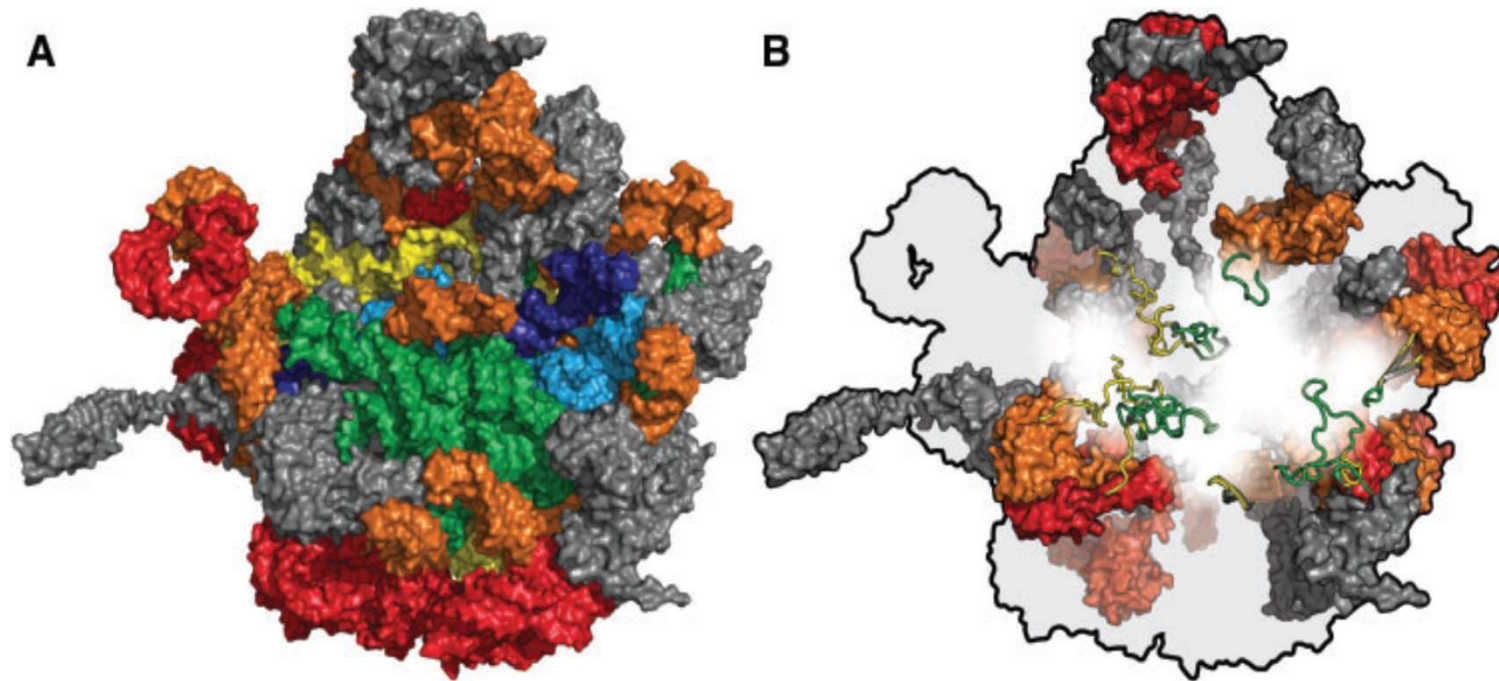
**1–3** are the L7/L12, central and L1 protuberances, respectively; **4** is the exit channel; **5–9** are the structures shown in a–e, respectively; **10** is the part of 50S subunit that does not include 23S rRNA. This part is formed by ribosomal proteins and 5S rRNA.

K. Bokov, S. Steinberg Nature. **2009**, 457, 977-980





## Evolution of the ribosome

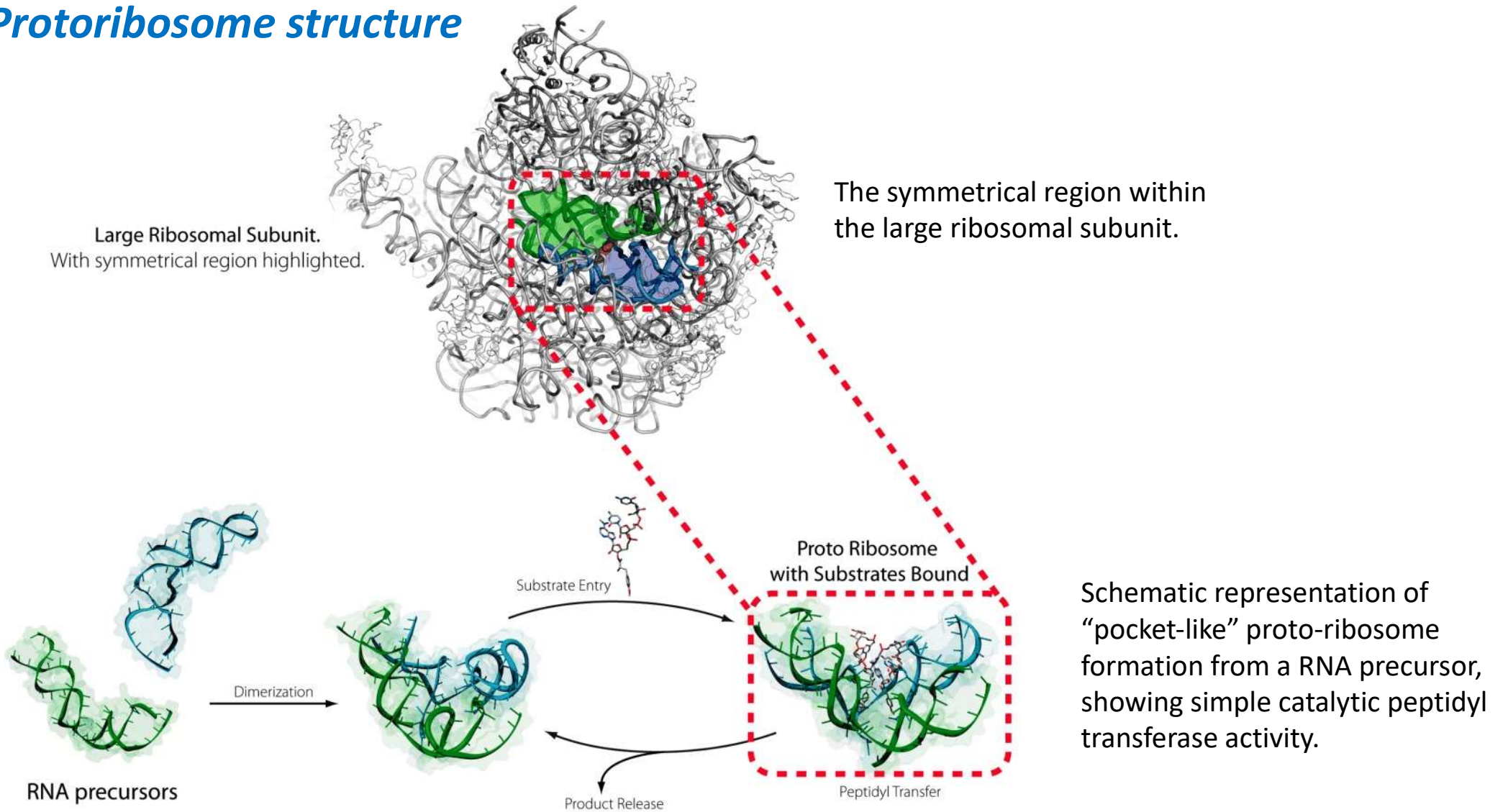


(A) The rRNA of the large subunit of the *T. thermophilus* ribosome colored by relative age. Phase 1, the most ancient phase, is dark blue. Phase 2 is light blue. Phase 3 is green. Phase 4 is yellow. Phase 5 is orange. Phase 6, the most recent prokaryotic phase, is red. rProteins are grey.

(B) The orientation is maintained but rRNA is colored in light grey, universal rProteins are colored by evolutionary phase, and bacterial rProteins are colored dark grey. Phases 3 (green) and 4 (yellow) are shown in cartoon representation. Phases 5 (orange) and 6 (red) are shown in surface representation. From PDB entry 1VY4

*N. A. Kovacs et al. Mol. Biol. Evol.* **2017** 34, 1252–1260.

# Protoribosome structure



A. Yonath *et al.* *Israel Journal of Chemistry*, 2010, 50, 29-35

# *The RNA world*

## *RNA as catalyst*

Currently known co-enzymes

Ribozymes

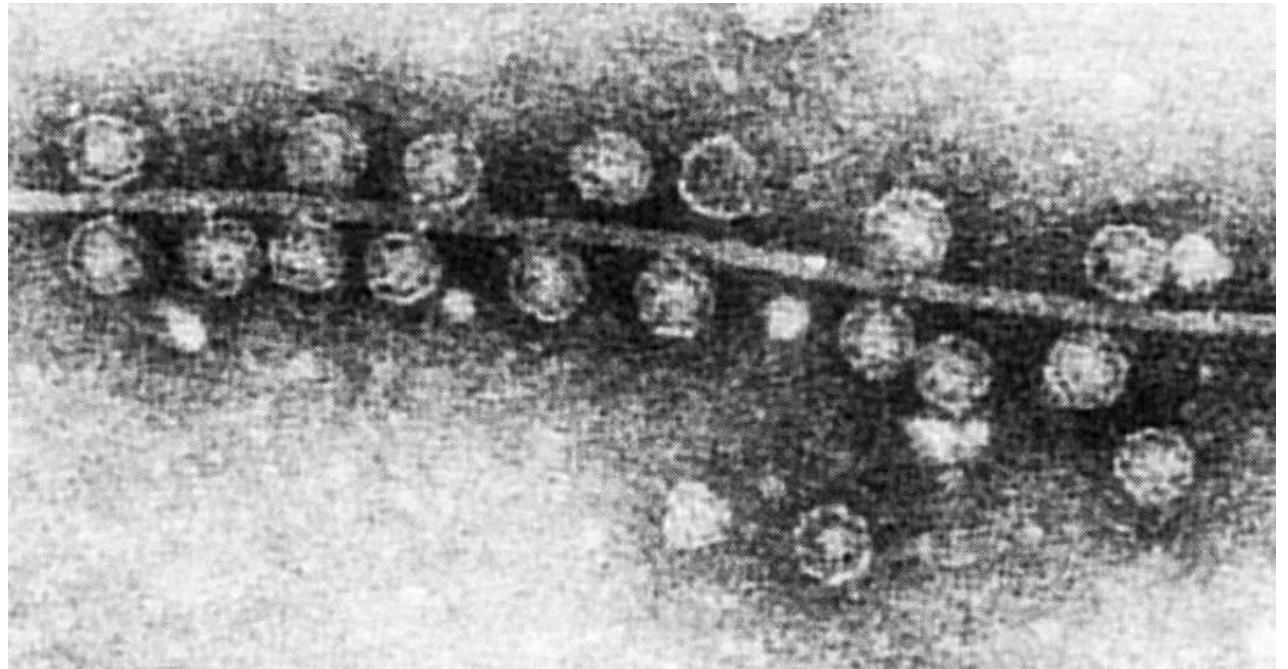
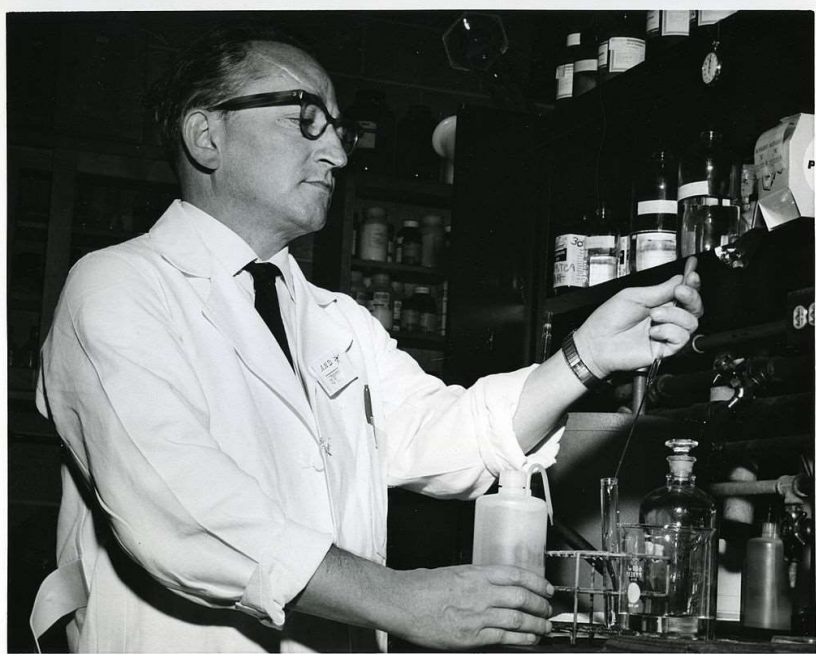
Ribosome

*Can RNA evolve?*

*Can RNA replicate itself?*

## *The RNA world*

*Can RNA evolve?*



*Spiegelman's monster*

## The RNA world

The bacteriophage Q $\beta$  – a virus containing RNA-dependent RNA polymerase (protein, enzymatic replicase)

### *Spiegelman's monster*

Spiegelman mixed the Q $\beta$  RNA, the Q $\beta$  enzymatic replicase, mononucleotides and some salts (buffer). RNA replication begun.

An aliquot was transferred several times to a fresh solution without template.

Shorter RNA chains replicate faster. The selection in this system favors speed.

And no evolutionary pressure on pathogenicity was present anymore.

So the RNA became shorter and shorter due to random mutations during copying.

After 74 passages, the original 4500 nt RNA strand was reduced to 218 nt.

Such a short RNA chain replicated very quickly under these unnatural circumstances.

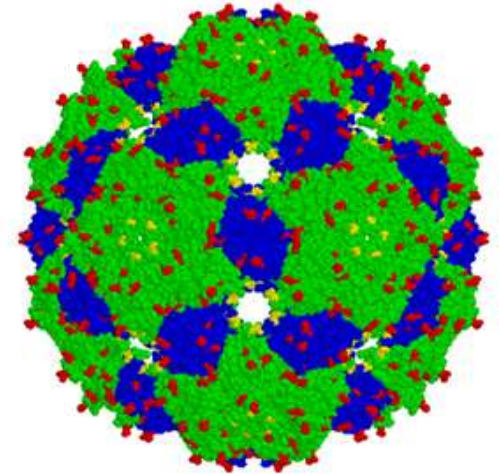
Of course, it lost all its genes and was unable to produce any useful proteins anymore.

### First example of *in vitro* RNA evolution

Kacian D. L., Mills D. R., Kramer F. R., Spiegelman S. *PNAS* **1972**, *69*, 3038-3042.

Spiegelman's monster can be also formed by simple mixing of activated RNA monomers and the Q $\beta$  enzymatic replicase, in absence of any RNA template!

Sumper M., Luce R. *PNAS* **1975**, *72*, 162-166.

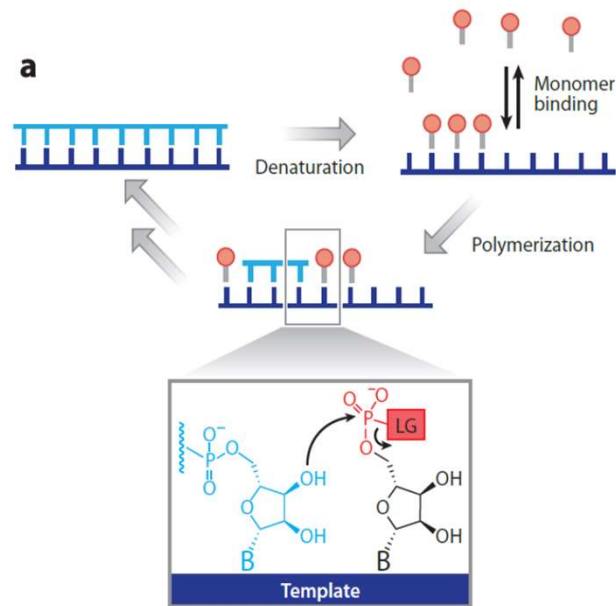


# The RNA world

## RNA self-replication

Nonenzymatic template-directed RNA polymerization

*Maximally 30-50 nt extension, fidelity strongly sequence-dependent*



General RNA polymerase ribozyme (‘replicase’)

Networks of RNA molecules that mutually catalyse their replication – autocatalytic replication of the whole network

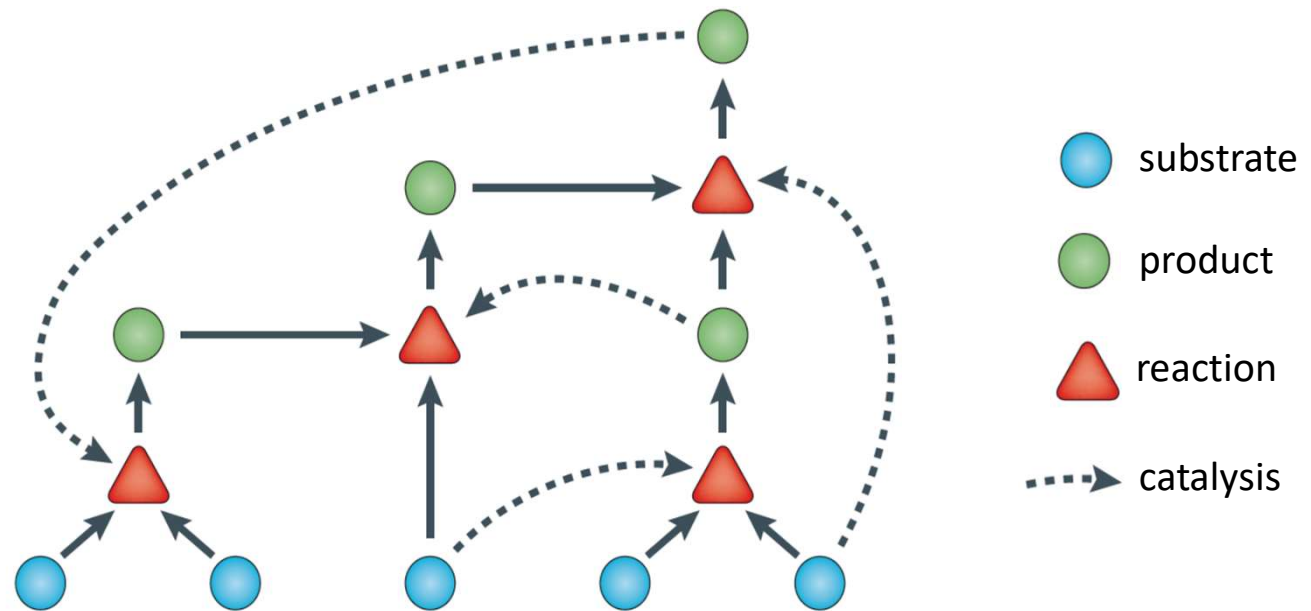


# The RNA world

## Replicase - problem

The replicase most likely needs to be long (> 200 nt) for the efficient replication –  
How could such long functional RNA be spontaneously generated?

## Possible solution – autocatalytic networks

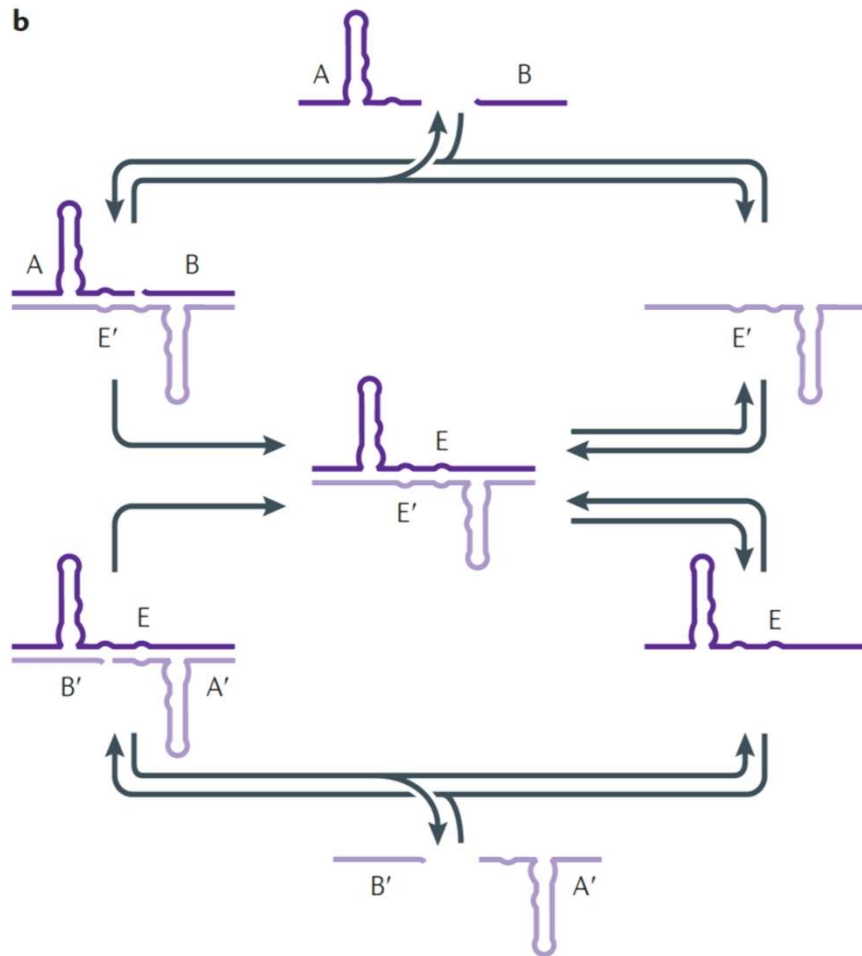


No component can replicate without all the others



# The RNA world

## Mutually autocatalytic RNA networks



An autocatalytic set composed of two cross-catalytic ligases was demonstrated. RNA A and RNA B are ligated together by ribozyme E' to create ribozyme E, which can reciprocate and ligate RNA A' and RNA B' to create ribozyme E'.

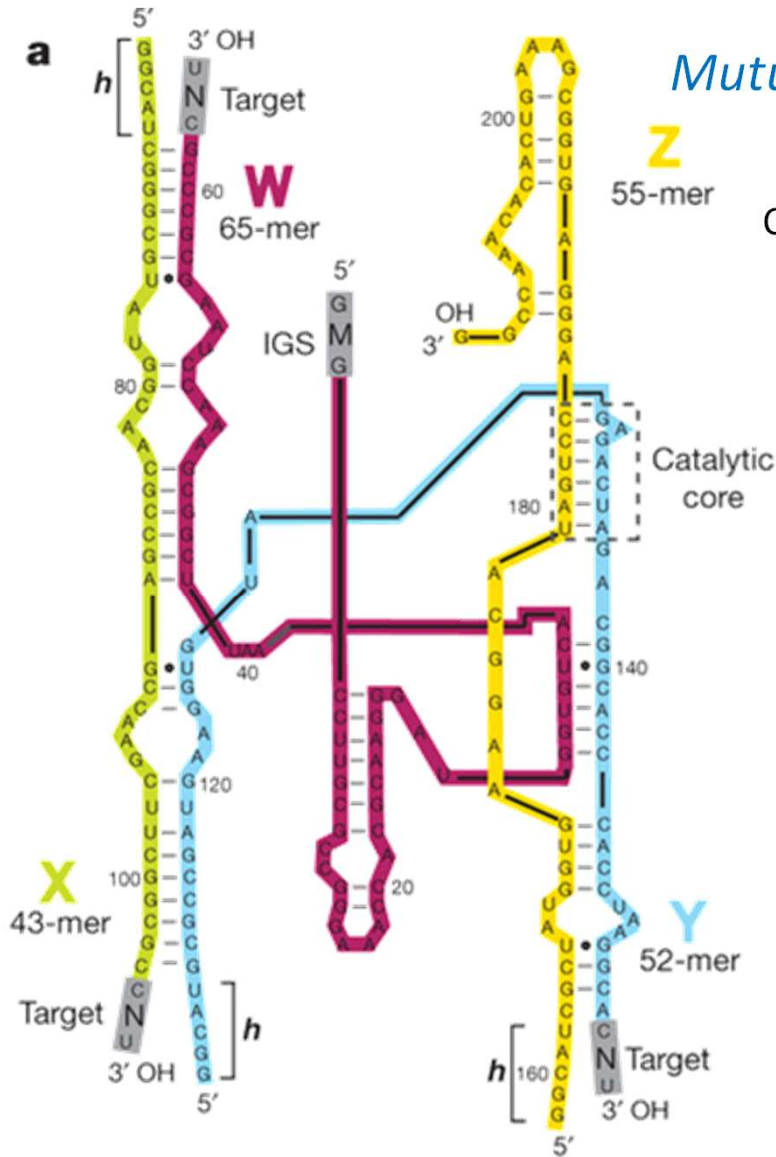
Lincoln, T. A. & Joyce, G. F. *Science* **2009**, 323, 1229–1232.

# The RNA world

## Mutually autocatalytic RNA networks

Cooperation between multiple strands that assemble to perform a single function.

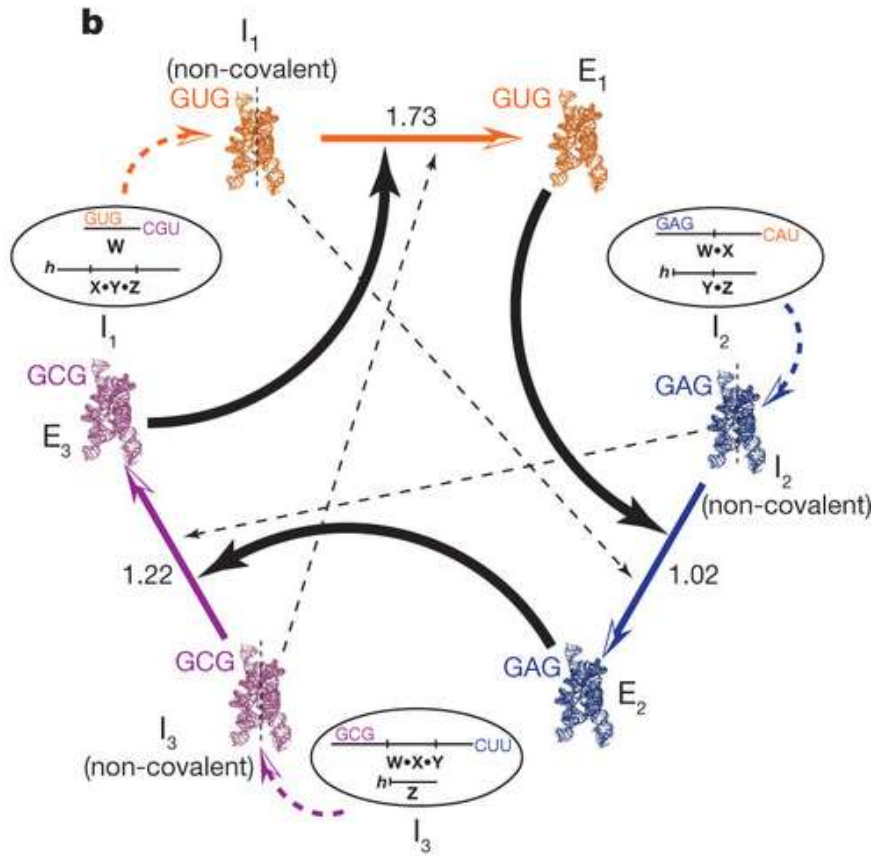
Ribozymes, such as the *Azoarcus* recombinase, can be made from several short strands that assemble as a result of RNA secondary structure formation and information contained in internal guide sequences (IGSs) and complementary targets (grey).



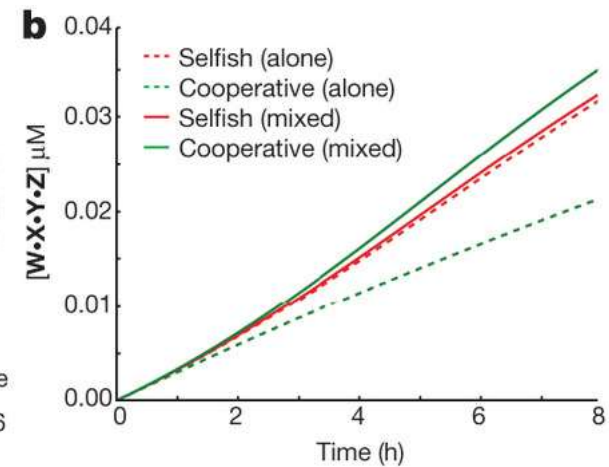
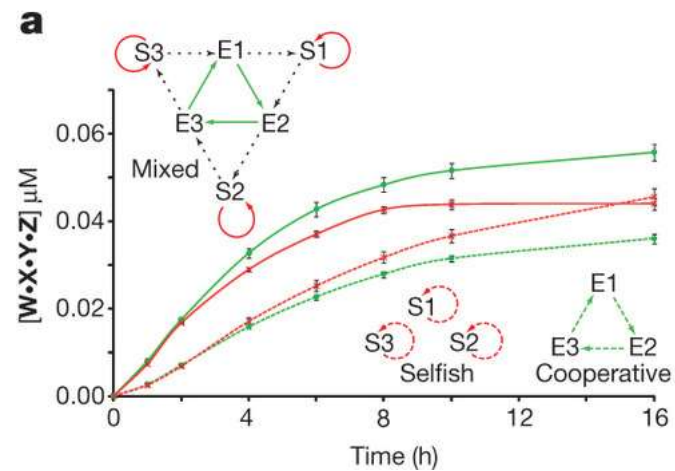
Vadia, N. et al. *Nature* **2012**, 491, 72-77.

# The RNA world

## Mutually autocatalytic RNA networks



mixtures of RNA fragments that self-assemble into self-replicating ribozymes spontaneously form cooperative catalytic cycles and networks.

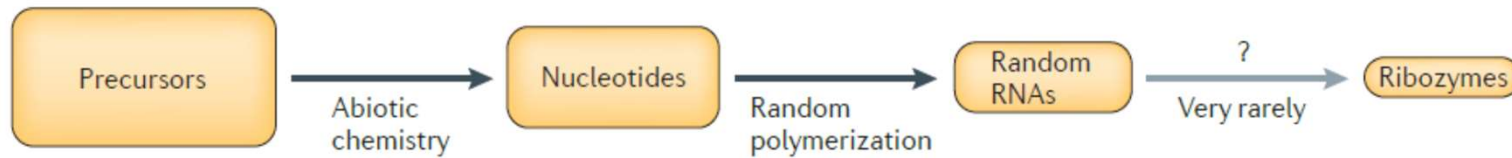


Vadia, N. et al. *Nature* **2012**, 491, 72-77.

# The RNA world

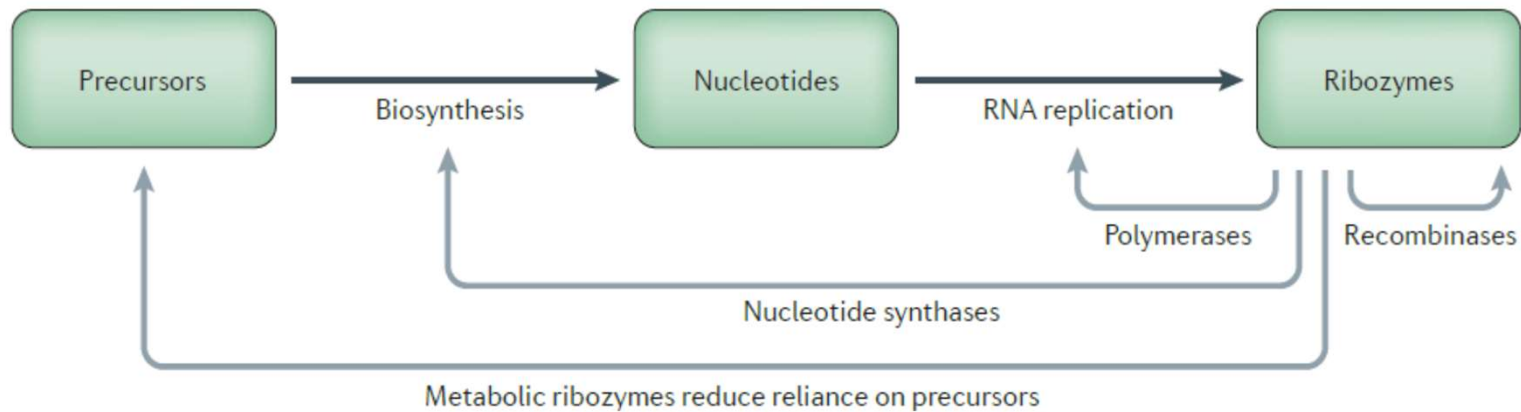
## a Chemistry

The prebiotic world: a dead state



## b Biology

The RNA World: an autocatalytic living state



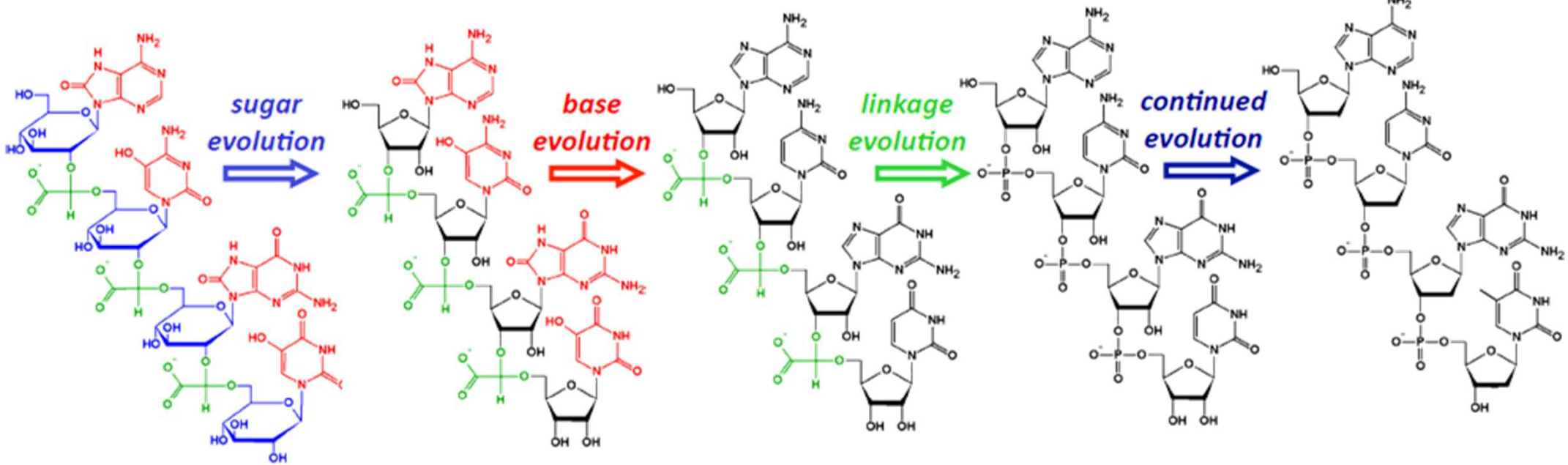
Transition from chemistry to biology involves autocatalytic feedbacks from ribozymes to all stages of the prebiotic chemistry

„RNA-second“

proto-RNA

RNA

DNA



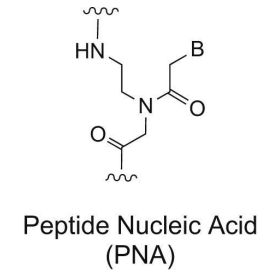
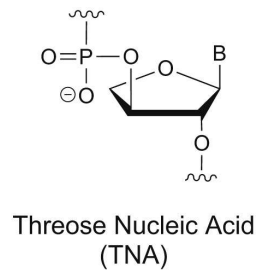
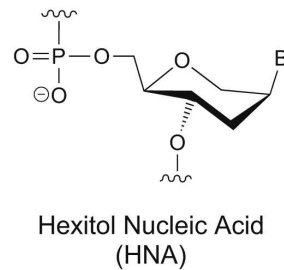
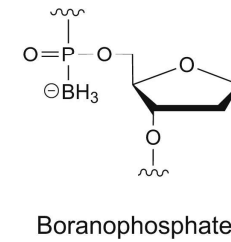
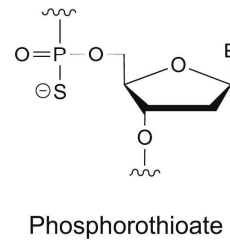
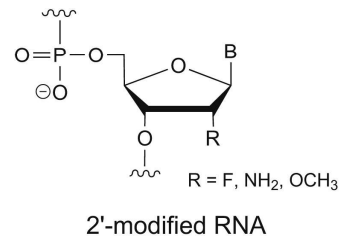
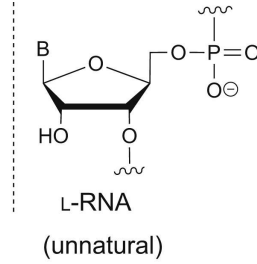
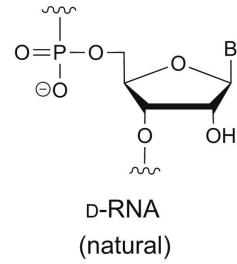
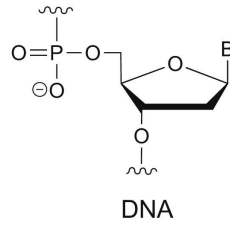
Easy to assemble



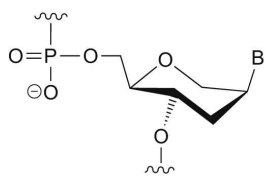
Functionally superior

*Proto-RNA evolution:* According to the protoRNA theory, each of the components of RNA — sugar, base and phosphate backbone — may have originally taken different forms.

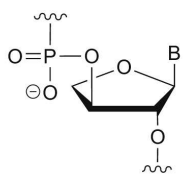
# Artificial genetic polymers



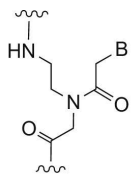
# XNA – Xeno Nucleic Acids



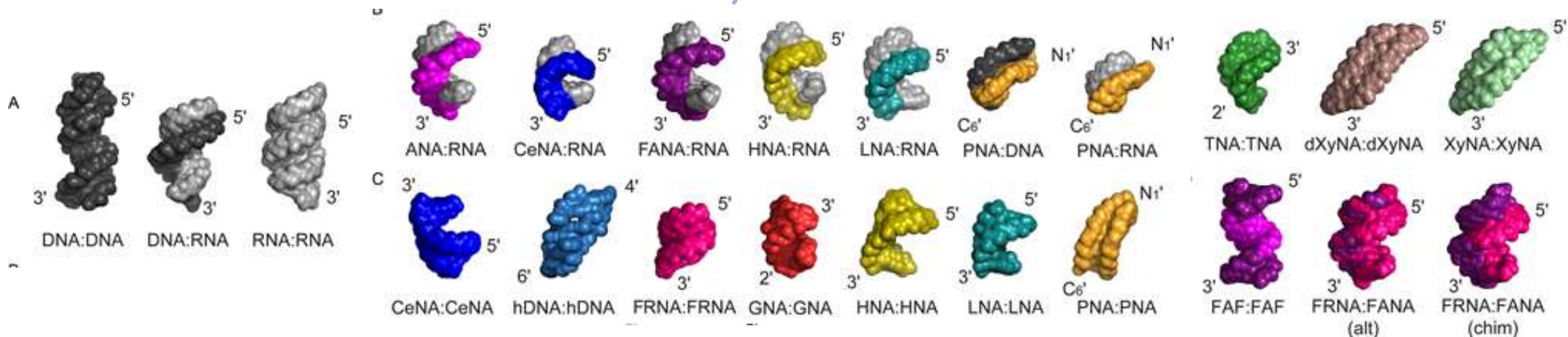
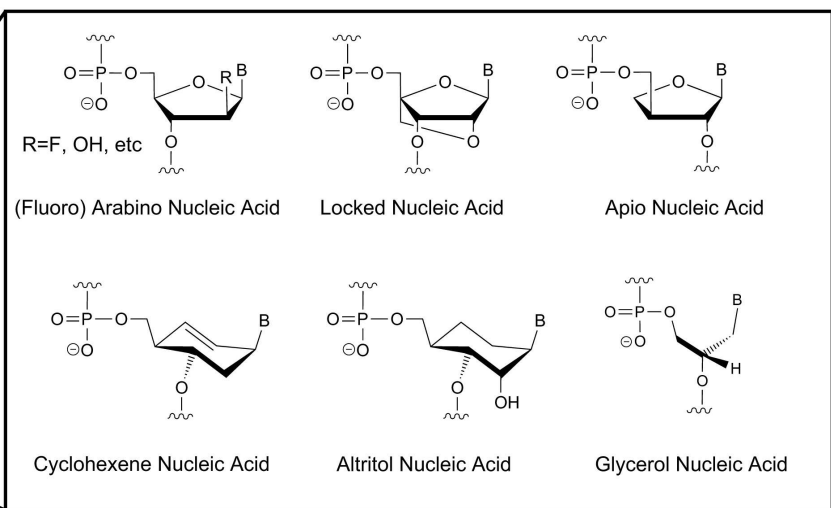
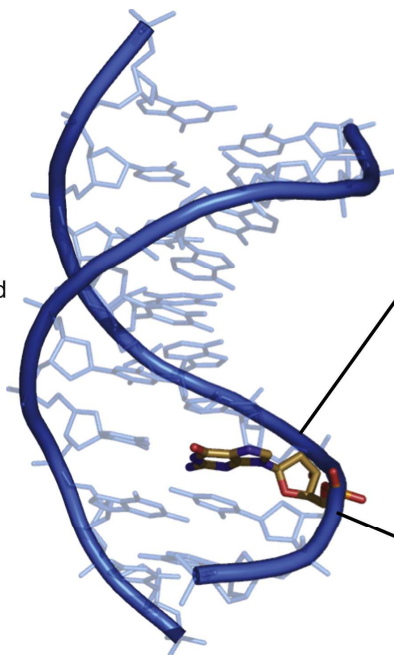
Hexitol Nucleic Acid (HNA)



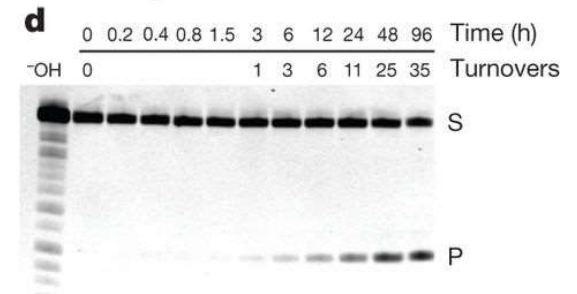
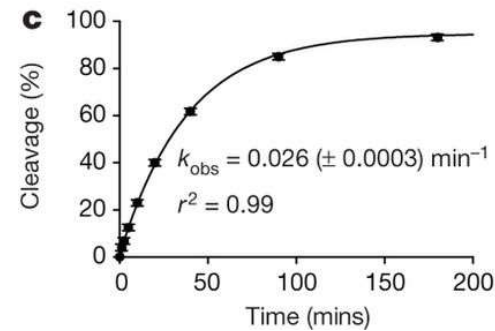
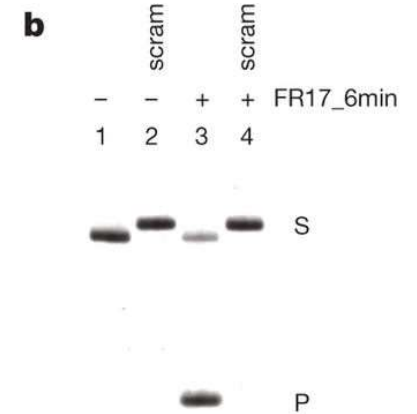
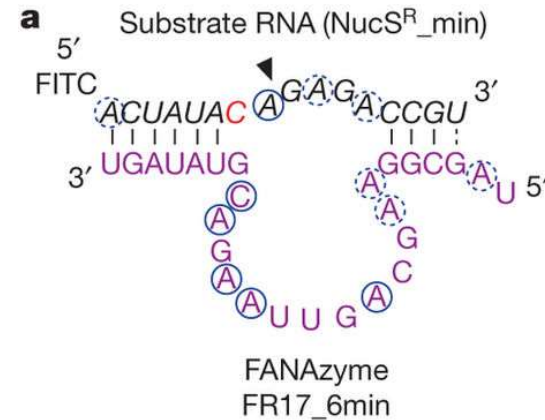
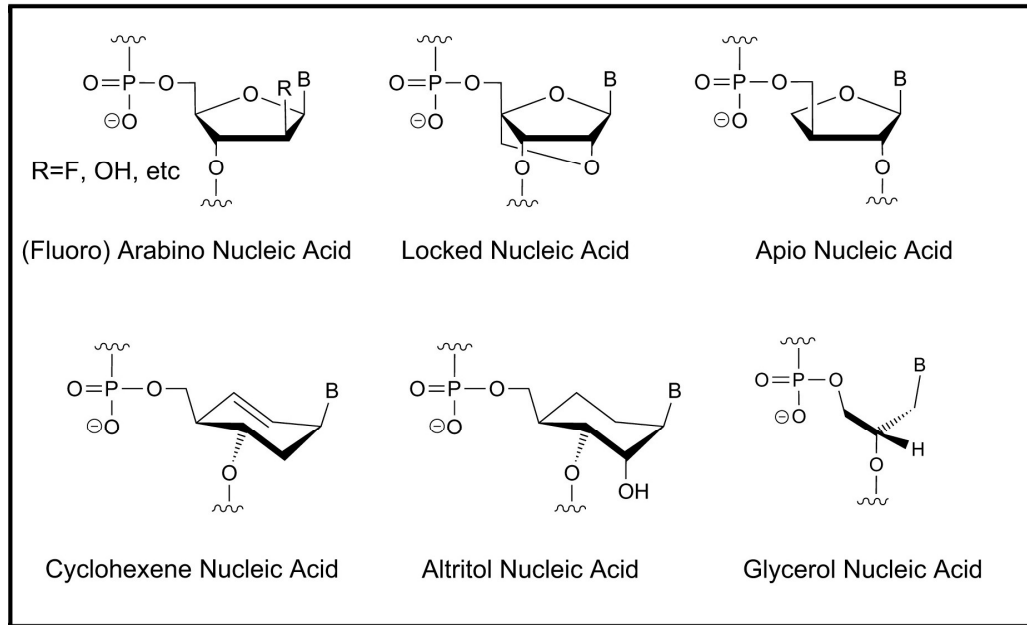
Threose Nucleic Acid (TNA)



Peptide Nucleic Acid (PNA)



# Chemical synthesis yields an active RNA endonuclease XNAzyme



**a**, Secondary structure of truncated FANAzyme FR17\_6 (FR17\_6min, purple)

**b**, FR17\_6min synthesized using FANA phosphoramidites cleaves cognate RNA substrate (NucSR<sub>min</sub>; lanes 1 and 3), but not a scrambled RNA (NucSR SCRAM2; lanes 2 and 4), with...

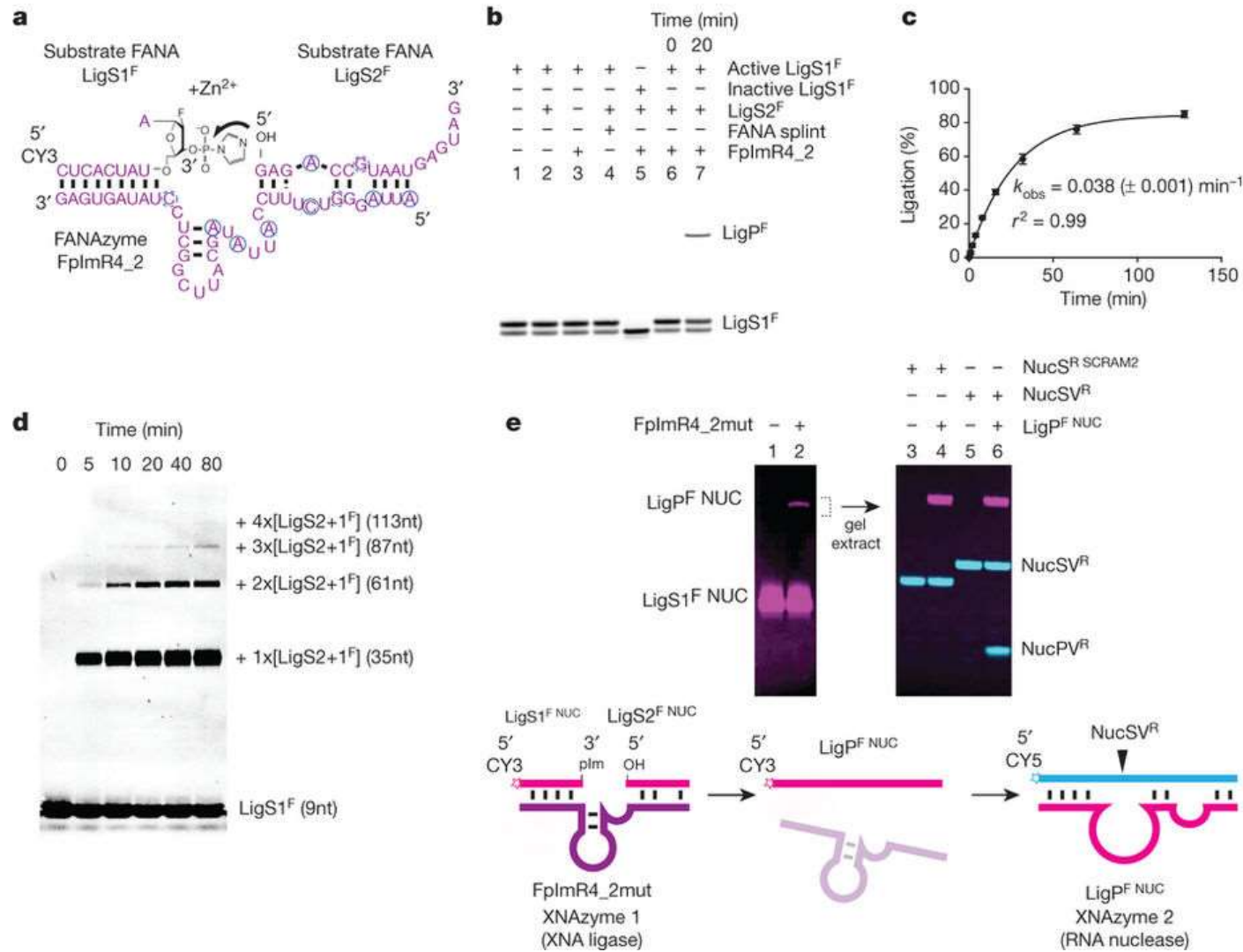
**c**, essentially unchanged catalytic rate ( $k_{obs}$ ) at 25 °C.

**d**, FR17\_6min (10 nM) can perform multiple turnover cleavage of RNA NucSR<sub>min</sub> (1 μM).

P. Herdewijn, P. Holliger, *et al. Nature* **2015**, *518*, 427-430



# XNA-XNA ligase XNAzyme (FANA): catalysis without natural nucleic acids



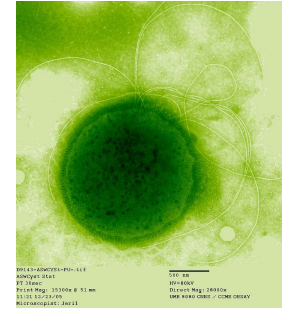
P. Herdewijn, P. Holliger, *et al.* *Nature* **2015**, *518*, 427-430

# Engineering XNA polymerases

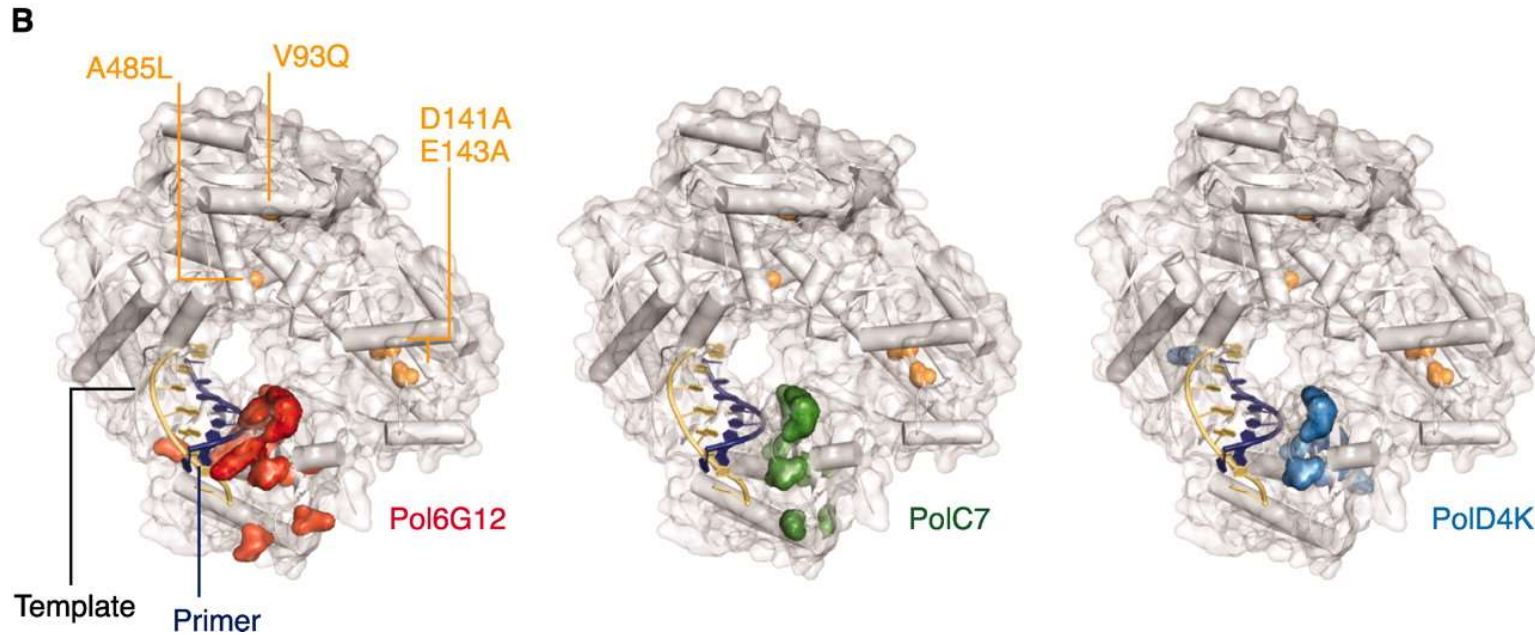
TgoT, a variant of the replicative polymerase of *Thermococcus gorgonarius*

**A**

	402	404	588	590	608	611	653	682	703	710	729	731
TgoT	YLD	..	FVT	..	LEIV	..	YEVPPPEKLVIIYEQITRDLKDYKATGPHVAV	..	VLKGS	GRI	..	AEY
Pol6G12	YLD	..	F <b>AT</b>	..	L <b>KMV</b>	..	YEVPPPEQLVIY <b>QPITKQL</b> HDY <b>RARGPHVSV</b>	..	V <b>PKGS</b>	GRI	..	<b>AGY</b>
PolC7	YLD	..	FVT	..	LEIV	..	YQVPP <b>QQLAIYQPITRALQ</b> DYK <b>AKGPHVAV</b>	..	VLKGS	<b>GKI</b>	..	AEY
PolD4K	Y <b>PD</b>	..	FVT	..	LEIV	..	YEVPT <b>QHLVIHKQITRALN</b> DYKA <b>IGPHVAV</b>	..	VLKGS	GRI	..	AEY



*Thermococcus gorgonarius*  
(Angels Tapias)



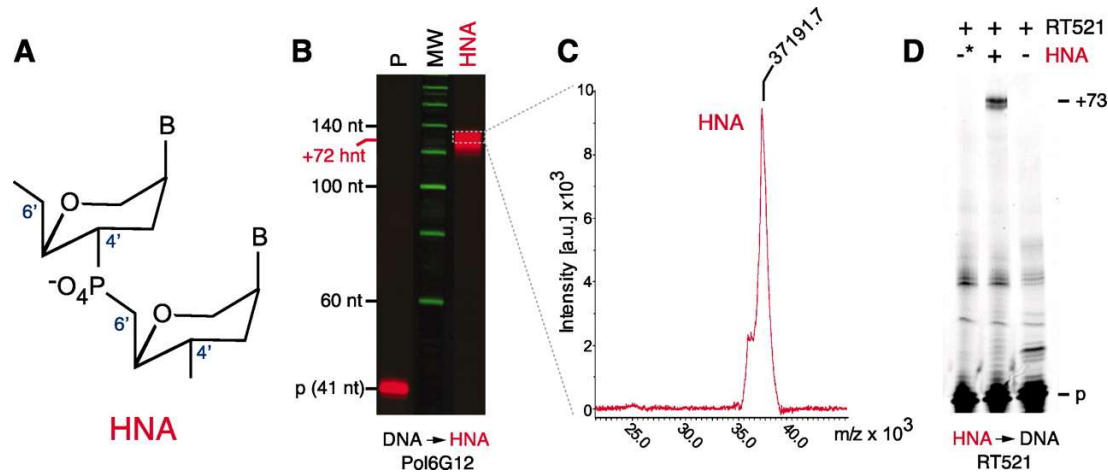
(A) Sequence alignments showing mutations from wtTgo in polymerases Pol6G12 (red), PolC7 (green), and PolD4K (blue).

(B) Mutations are mapped on the structure of Pfu (PDB: 4AIL).

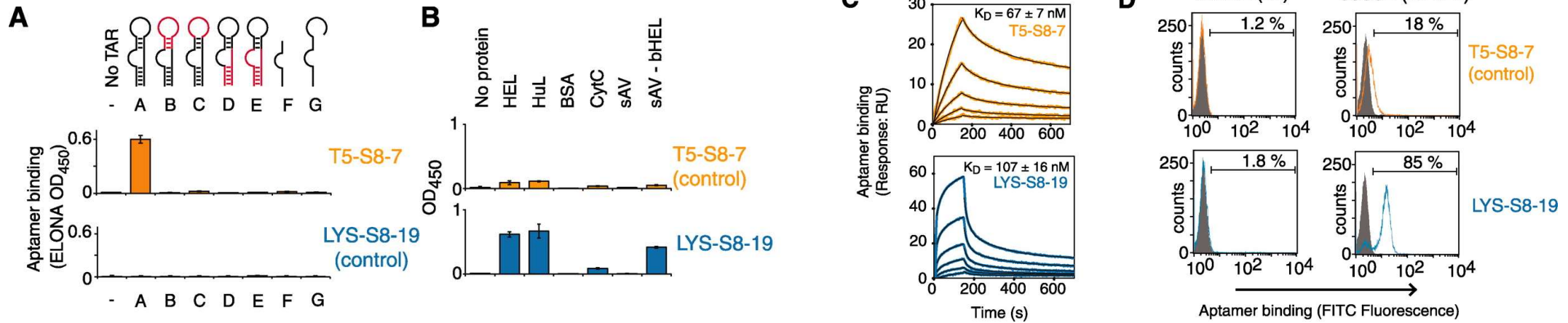
Yellow - template; dark blue - primer; orange - mutations present in the parent polymerase TgoT

P. Herdewijn, P. Holliger, *et al. Science* **2012**, *336*, 341-344

# DNA-templated HNA synthesis and HNA-templated DNA synthesis



## HNA aptamers



P. Herdewijn, P. Holliger, *et al.* *Science* **2012**, *336*, 341-344