Assembly and Abilities of RNA Polymerase Ribozymes

SEMINAR: PHYSICS OF EARLY LIFE

PRESENTED ON THE 29TH OF JANUARY 2021, BY VALENTIN STRAUB

LUDWIG MAXIMILIAN UNIVERSITY MUNICH

Content

1. Introduction - RNA World
2. How did RNA polymerase ribozymes evolve?
3. Are RNA polymerase ribozymes able to copy functional RNA molecules and the polymerase?
4. Conclusion

1. Introduction - RNA World

The RNA world is a hypothetical stage in the evolutionary history of life on Earth, in which self-replicating RNA molecules proliferated before the evolution of DNA and proteins.

Like protein enzymes, RNA enzymes (ribozymes) can catalyse chemical reactions that are critical for life. [1]

RNA polymerase ribozymes (RPR) are credible candidates for an RNA molecule capable of replicating its own sequence and therefore initiating Darwinian evolution.[2]

1. Introduction - RNA World

Question 1:

How did RNA polymerase ribozymes evolve?

Question 2:

Are RNA polymerase ribozymes able to copy functional RNA molecules and the polymerase?

2. How did RNA polymerase ribozymes (RPR) evolve?

Non-Enzymatic Assembly of a Minimized RNA Polymerase Ribozyme

by Dr. Falk Wachowius and Dr. Philipp Holliger [2]

Initial Question
Could RNA polymerase ribozymes have emerged from pools of short RNA oligomers (accessible from prebiotic chemistry, ergo < 30 nts) through their assembly by iterative ligation reactions?
Approach
Exploration of non-enzymatic approach, based on the widely used phosphorimidazolide (Imp) activation chemistry (Enables non-templated and templated RNA polymerization).
Goal
Connecting enzymatic (ribozyme catalyzed) RNA polymerization to the simpler non-enzymatic RNA polymerization reactions that must have preceded it.

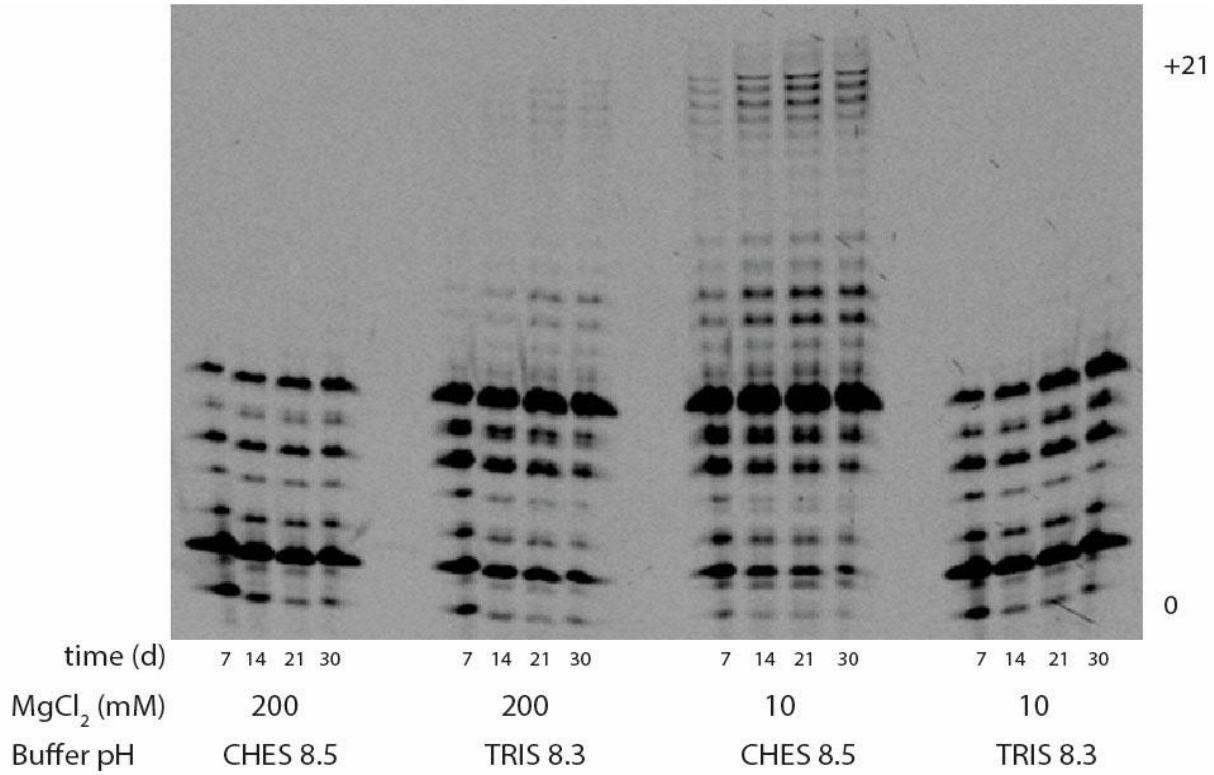
2.1 Optimizing reaction conditions

Investigation of pH vs Mg^{2+} concentration dependence of RPR activity in ice → switch of buffer system.

Also enhanced activity at higher pH values with decreasing Mg^{2+}.

Initial: -7°C, 200mM $MgCl_2$, TRIS pH 8.3

New: -7 °C, 10mM $MgCl_2$, CHES pH 9

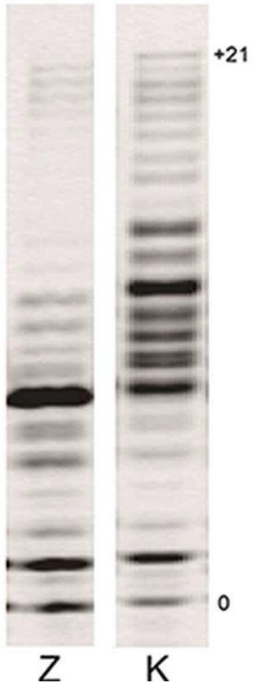


Optimisation of in-ice RNA polymerase ribozyme activity [2]

2.2 Optimizing RPR activity while minimalizing RPR size

1. RPR variants Z, W, Y with random linker lengths → 8 rounds of CBT selection → only 12nt linker region variants remained → among these 3 clones with improved activity (C1,C2,K) were identified.
2. All three variants derived from the W variant pool.
3. Clone K showed improved RNA polymerase activity in comparison to the best RPR variant (Z) & could also be shortened to 180nt (Z = 195nt);
4. The K ribozyme accessory domain truncated is further and results in the RPR variant F with 150nt.
5. Variant F retained robust polymerase activity, but only in ice at higher RPR concentrations (7µM).

→ Improved polymerization activity on challenging templates.



New K RPR with improved activity compared to the Z RPR[2]

2.3 Can the F ribozyme be assembled from short activated RNA oligomers?

Non-enzymatic polymerization of RNA oligomers from 5’-phosphorimidazole-activated monomers in ice usually generates short oligomers with maximum lengths of ~20nt and a mixed base composition.

Optimized templated non-enzymatic RNA replication using Imp (or similar phosphate activation) chemistry is increasingly able to generate mixed sequence oligomers.

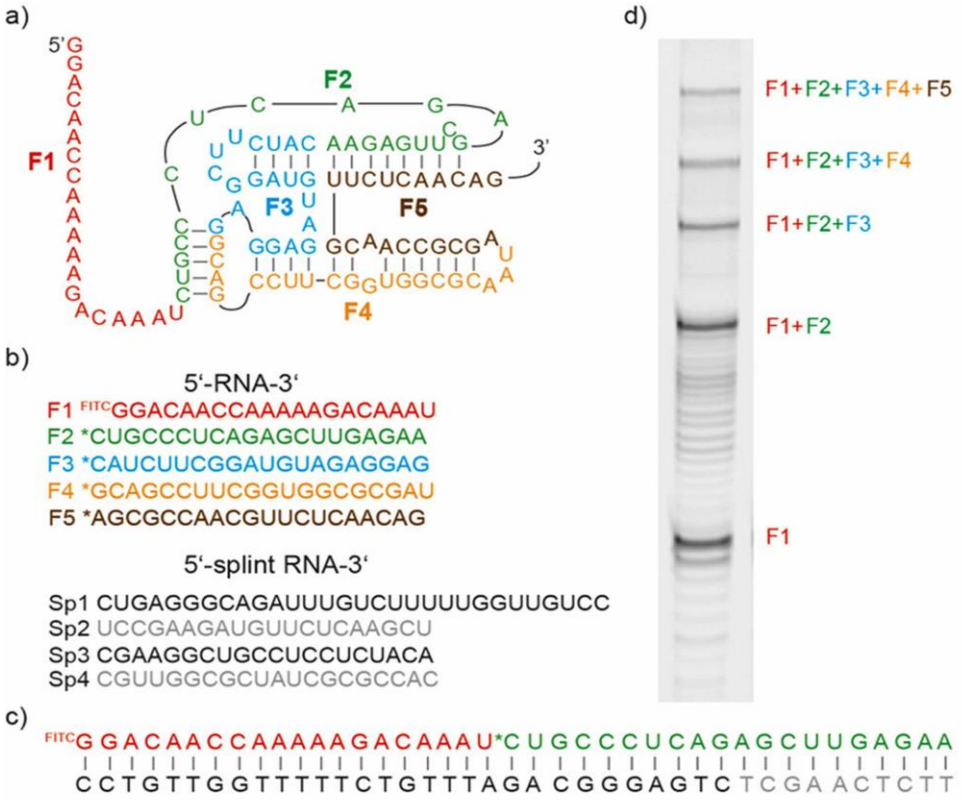
This suggests that mixed RNA oligomer sequence pools in the size range ≤ 30nt may be prebiotically plausible.

However as Imp activated monomers are preferentially introduced, ligation of oligomers is much less efficient, which hindered polymerisation of 5’-Imp activated RNA oligomers so far.

2.3.1 Non-enzymatic assembly of 5’-phosphor-2-MeImp activated RNA eicosamers

1. Dividing of class I ligase derived catalytic core of the RPR (100nt) into five 20-mer segments
2. Addition of complementary RNA splints (20nts)

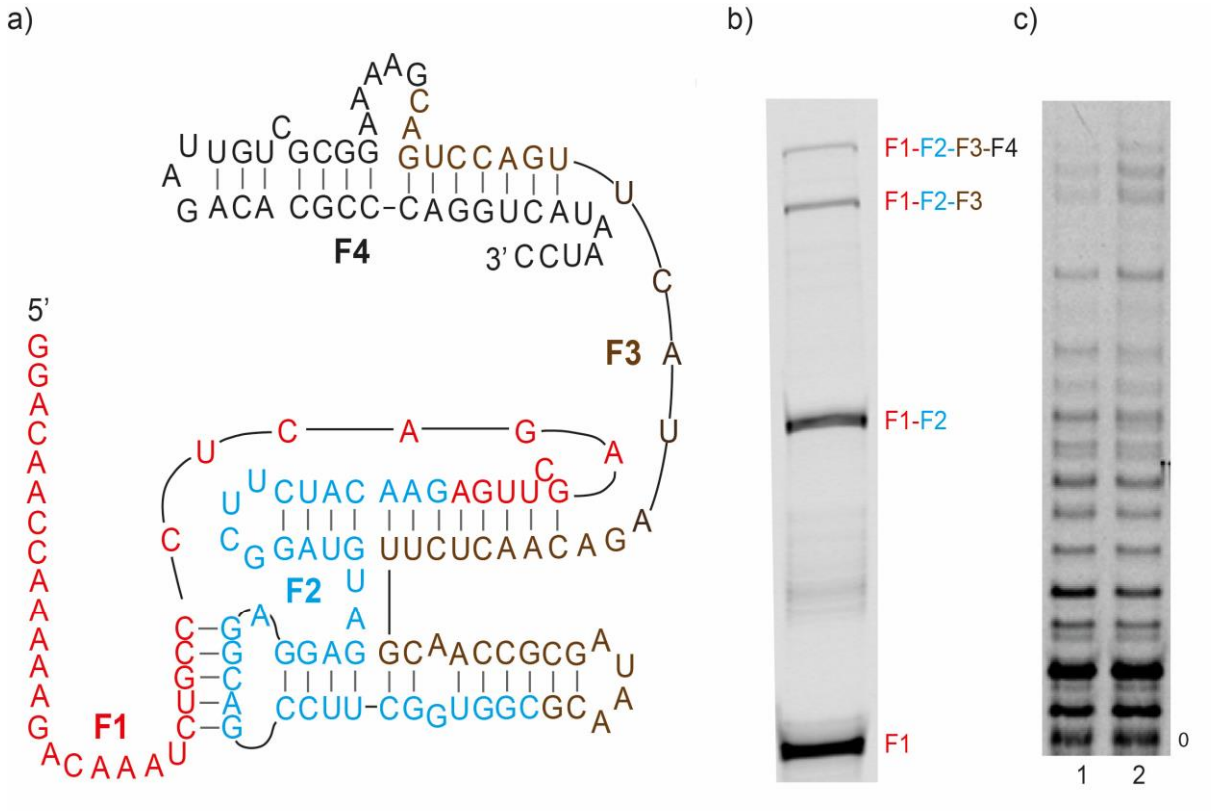
→ Observation of spontaneous assembly of a full length catalytic core sequence with 5.4% yield (24 h, 37°C)



Non-enzymatic templated ligation of the F ribozyme catalytic core [2]

2.3.2 Assembly of full length RPR (variant F) in same manner

1. Division of F RPR sequence in four parts (38nt).
2. Addition of four complementary RNA splints (24nt).
3. → Assembly of whole functional F ribozyme sequence (24 h at 37°C) with a total yield of 1.7%.
4. → RNA oligomers, should they become available from non-enzymatic polymerization reactions, can be assembled in analogous fashion despite their increased tendency for secondary structure formation.



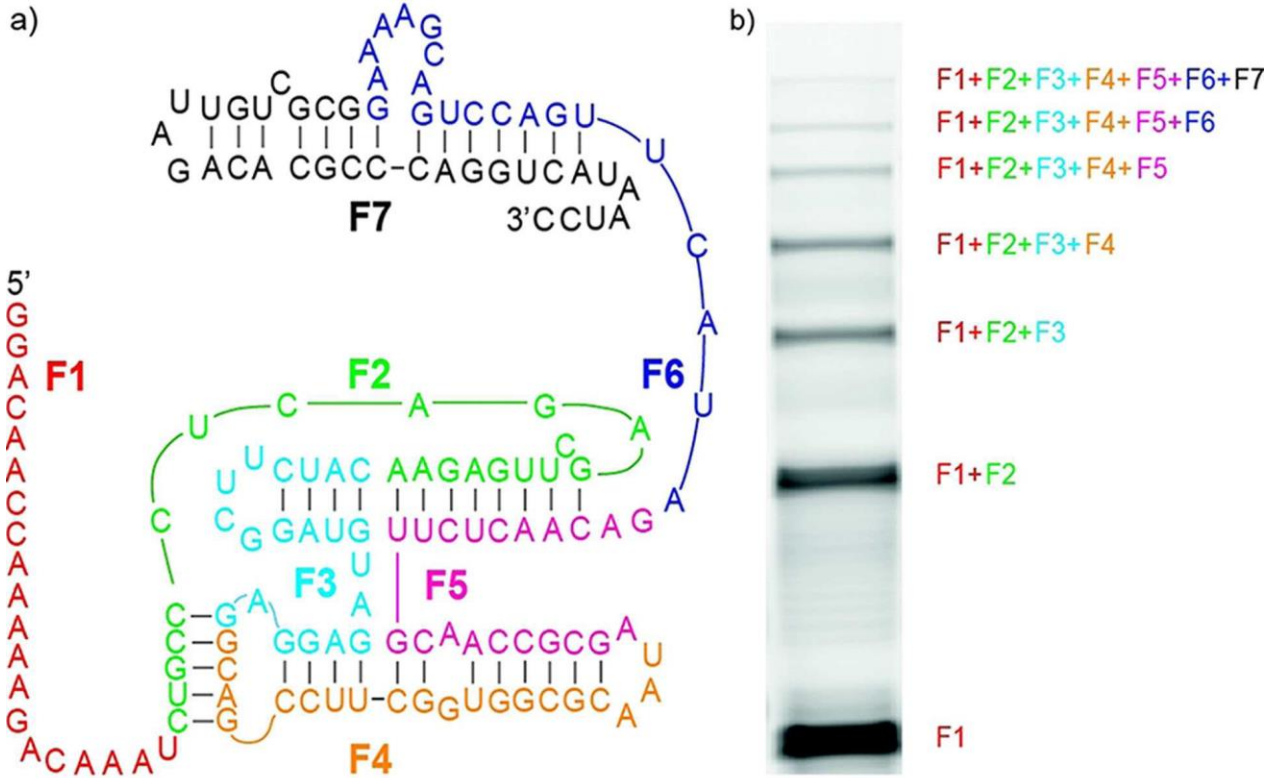
Non-enzymatic RNA polymerase ribozyme assembly [2]

2.3.3 Assembly of whole F RNA polymerase ribozyme from RNA oligomers

Requires the definition of six mutually exclusive (orthogonal) assembly sites in the RPR sequence, with a steep increase in the potential for mis-assembly.

1. Splitting of the F ribozyme sequence of 150nt into seven pieces with arbitrary ligation sites (6×20nt, 1×30nt).
2. Activation of the 5’-phosphate termini with 2-MeImp.
3. Addition of the corresponding six complementary RNA splints (5×20nt, 1×30nt).

→ Observation of the ligation of full-length ribozyme sequence.

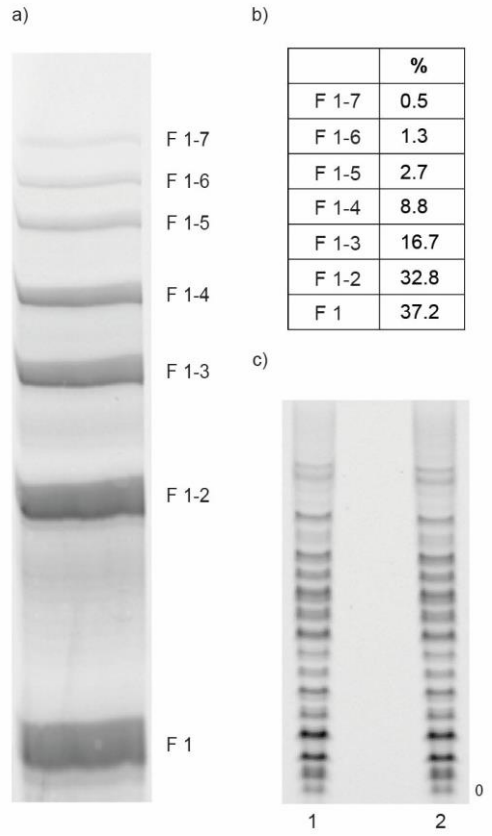


Functional RPR assembly from short RNA fragments[2]

2.3.4 Verification of the F RNA polymerase assembly

Correct assembly verified by sequencing and the full RNA polymerase ribozyme activity is confirmed by primer extension in ice (with optimized buffer conditions).

The final yield of the full length ribozyme from seven pieces is 0.5%.



	%
F 1-7	0.5
F 1-6	1.3
F 1-5	2.7
F 1-4	8.8
F 1-3	16.7
F 1-2	32.8
F 1	37.2

Activity of assembled ribozyme[2]

2.4 Results

Experiments show, how complex RNA structures could have been generated non-enzymatically in a plausible prebiotic setting by the iterative ligation of multiple Imp (phosphorimidazole) activated RNA oligomers.

A proof-of-principle demonstration for the spontaneous assembly of a functional 150nt RNA polymerase ribozyme (F1234567) from up to 7 short (20–30nt), non-functional 5’-2-methylimidazole-activated RNA oligonucleotides via splint assisted non-enzymatic ligation in a “one pot” reaction.

2.4 Results

The final yield of the full-length ribozyme from 7 pieces is currently relatively modest (0.5 %). Reasons are:

1. Chemical instability and short half-life of the phosphorimidazole activating groups in aqueous solution.
2. The incomplete formation of ligation junctions due to RNA misfolding and the incomplete ligation of assembled junctions.
3. The chemical reactivity of 5’-phosphorimidazole activated RNA oligonucleotides is in general reduced, compared to the equivalent activated RNA monomers. This is related to the fact, that the latter is assisted by beneficial interactions of the imidazole leaving groups of the incoming monomer and the next downstream monomer that are absent in the oligonucleotide ligation approach.

2.5 Summary and Outlook

The new minimized RNA polymerase ribozyme (F) (comprising only 150nt) represents one of the shortest RPRs, with the ability to show robust RNA polymerase activity on difficult template sequences in the preferential medium of water-ice at -7°C.

This could pave the way for future RPR catalyzed RNA self-replication, that is based on

1. a combination of RNA oligomer synthesis by enzymatic NTP polymerization and
2. a non-enzymatic templated ligation of these oligonucleotides.

3. Are RNA polymerase ribozymes able to copy functional RNA molecules and the polymerase?

An RNA polymerase ribozyme that synthesizes its own ancestor

by Katrina F. Tjhung, Maxim N. Shokhirev, David P. Horning, and Gerald F. Joyce [3]

Initial Assumption
• RNA based organisms (ancestors of modern life) must have depended on an RNA polymerase ribozyme, which is able to copy functional RNA molecules (incl. polymerase itself). • This polymerase must have been highly efficient and have a high fidelity to maintain genetic information across generations.
Approach
Evolving Class I RNA polymerase ribozyme in vitro for the ability to synthesize functional ribozymes, resulting in the markedly improved ability to synthesize complex RNAs using nucleoside 5’-triphosphate (NTP) substrates.
Goal
Using directed evolution to obtain a highly improved class I polymerase ribozyme with the ability to synthesize a functional copy of its own evolutionary ancestor (class I ligase).

3.1 Evolution of Polymerase Ribozymes that Synthesize Functional Ribozymes (24-3 → 38-6)

1. A Class I RNA polymerase ribozyme that has been evolved in vitro for the ability to synthesize functional RNA aptamers (“24-3” variant) is taken as a basis.
2. In Vitro evolution procedure is continued for an additional 14 rounds.
3. Cloning and sequencing of individuals from the population. → 3 distinct sequence families.
4. Most active individual named “38-6”; The most notable structural change is disruption of the distal portion of the P7 stem, opening the L7 loop that closes the stem.

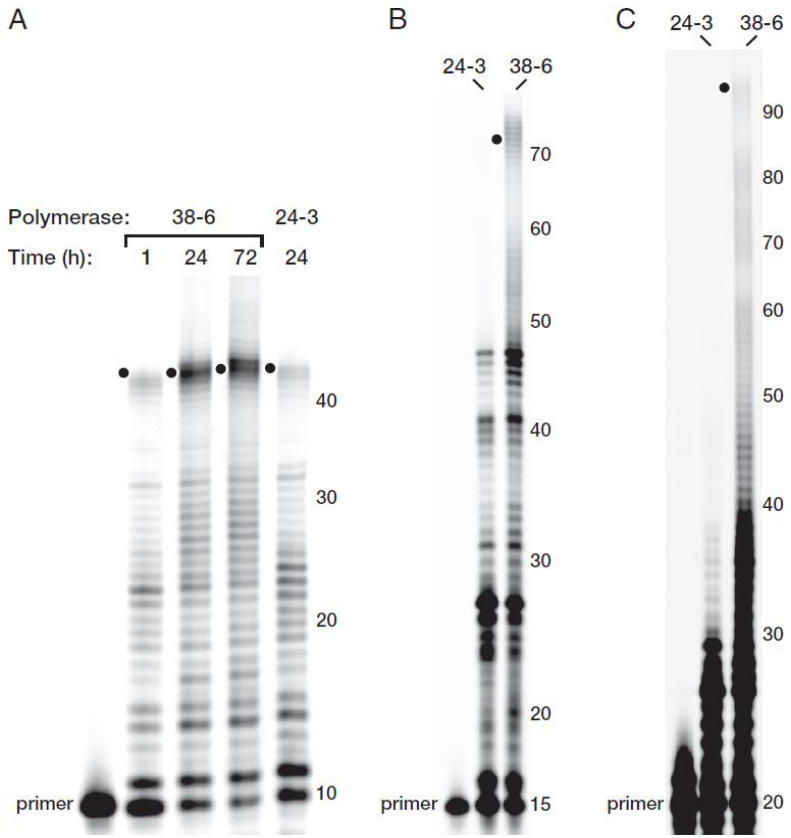
3.2 The Evolved Polymerase Can Synthesize Complex RNAs

Improvement of RNA-dependent RNA polymerase activity (compared to 24-3)

A. Synthesizing the hammerhead ribozyme:
   - 24-3: 2.4% yield after 24 h
   - 38-6: 2.0% yield in 1 h and a 24% yield in 24 h

B. Synthesizing a more complex RNA:
   - 24-3: barely detectable yield after 5 d
   - 38-6: 2.4% yield after 5 d

C. Synthesizing the b1-207t variant of the class I ligase:
   - 24-3: no detectable yield
   - 38-6: 0.12% after 5 d



Synthesis of functional RNA molecules by the 24-3 and 38-6 polymerases [3]

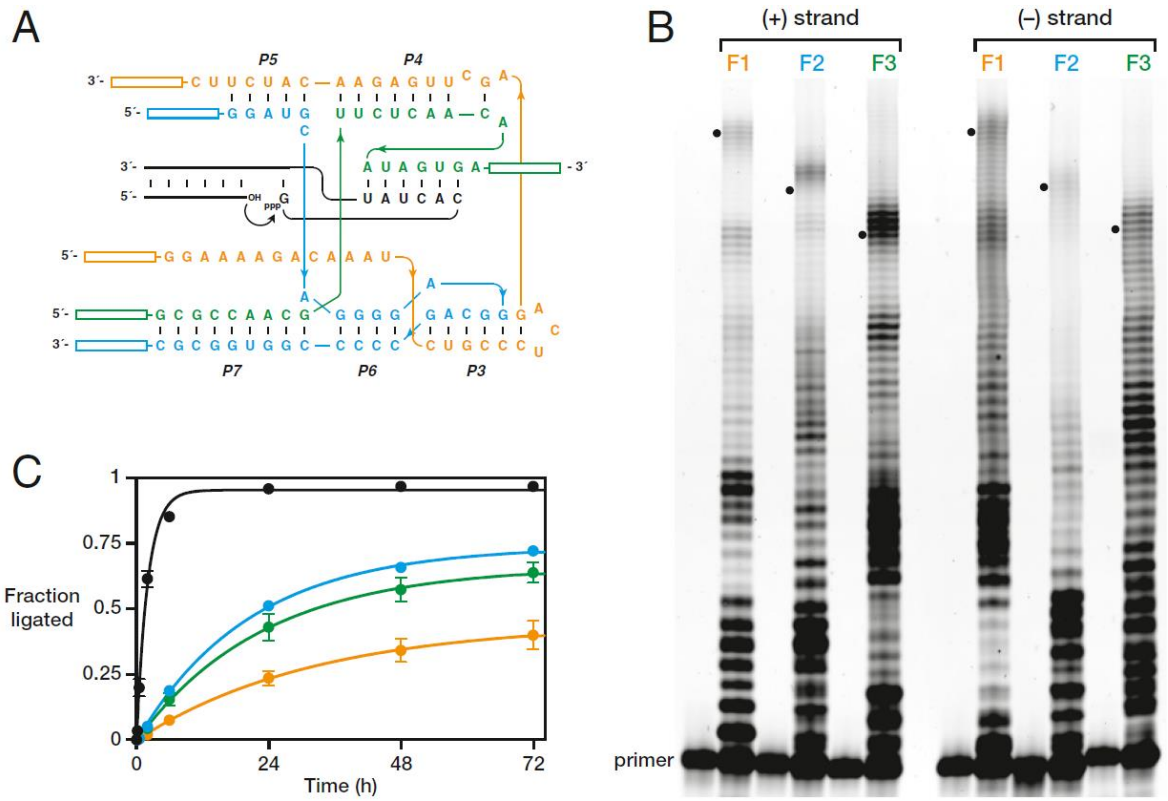
3.3 Polymerization of own ancestor in three segments

1. Splitting of the ligase into three frements
2. Addition of external sites to both ends of each fragment to prevent the primers from including any nucleotides of the ribozyme
3. After 3-d incubation

Result:

Fragments 1, 2, and 3 were generated from their corresponding RNA templates in yields of 1.1%, 1.0%, and 5.7%, respectively.

The polymerase synthesizes the complement of each fragment, in yields of 0.53%, 0.91%, and 2.0%, respectively.



Synthesis of the three-fragment form of the class I ligase ribozyme by the 38-6 polymerase[3]

3.4 Results (Fidelity of evolved polymerase)

Partial length products that contain a mutation are less likely to be extended to full length compared with accurately synthesized intermediates.

Average fidelities when polymerizing full length products:

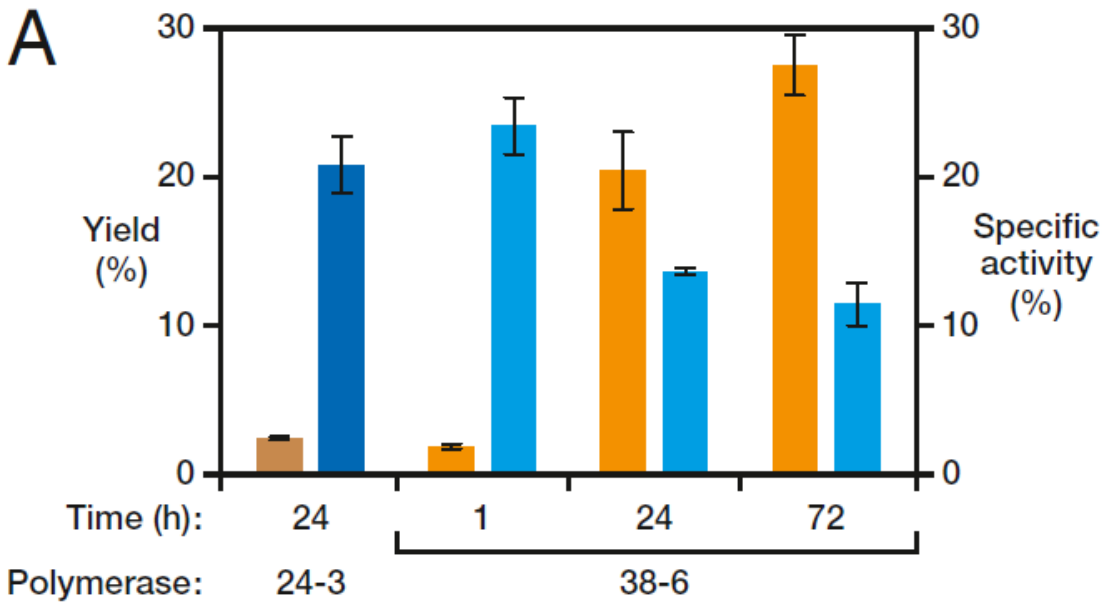
- 24-3: 97.1%
- 38-6: 96.2%

Pol	Expected	Observed (%)					
		A	G	C	U	Del	Ins
24-3	A	**91.7**	6.9	0.0	0.0	1.4	0.0
	G	0.0	**100.0**	0.0	0.0	0.0	0.0
	C	0.0	0.0	**98.6**	1.4	0.0	0.0
	U	0.9	0.9	0.0	**98.1**	0.0	0.0
38-6	A	**98.6**	0.0	0.0	0.0	1.4	0.0
	G	0.0	**97.3**	0.5	0.0	0.0	2.2
	C	0.0	0.0	**94.4**	5.6	0.0	0.0
	U	0.9	2.8	0.9	**94.4**	0.9	0.0

Fidelity of the 24-3 and 38-6 polymerases [3]

3.4 Results (Investigation of the relationship between yield and fidelity)

At longer incubation times, the 38-6 polymerase is able to continue the extension of partial-length products that contain a mutation, improving the yield of full-length products but resulting in decreased fidelity of those full-length products.

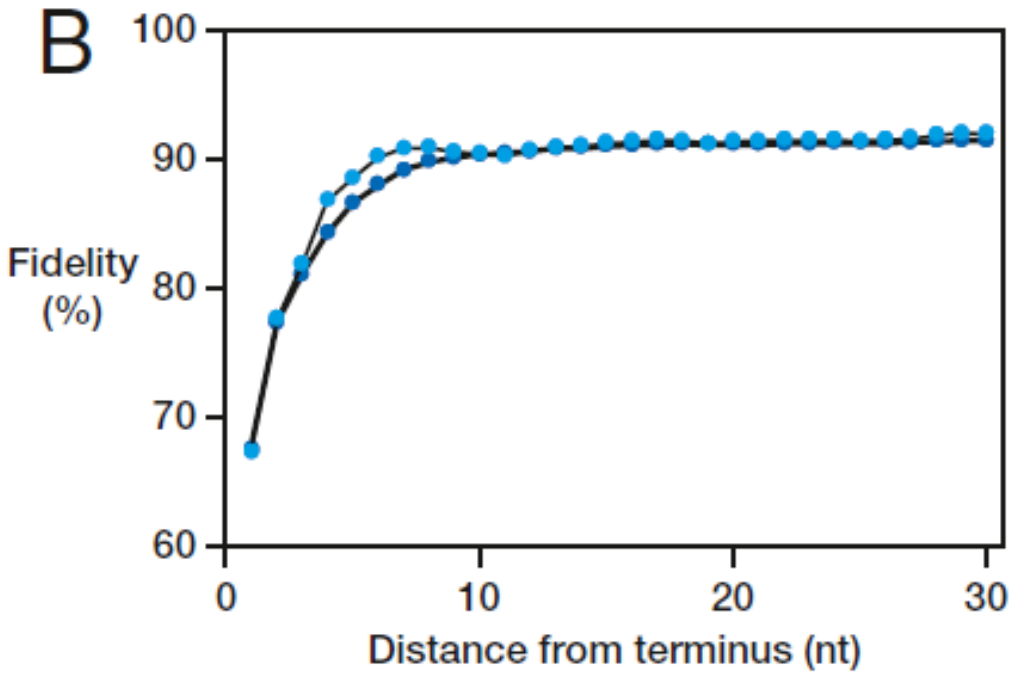


Yield and specific activity of synthesized products.[3]
Yield (orange) is the percent primer extended to give the full-length hammerhead.
Specific activity (blue) is the initial rate of reaction, under multiple-turnover conditions, of the ribozyme-synthesized hammerhead relative to that of the protein-synthesized hammerhead.

3.4 Results (Similarity of the two polymerases)

Fidelity is lowest for the last added nucleotide (~67%) and increases monotonically for nucleotides that are increasingly further upstream from the 3’ terminus.

→ Similarity between 24-3 and 38-6

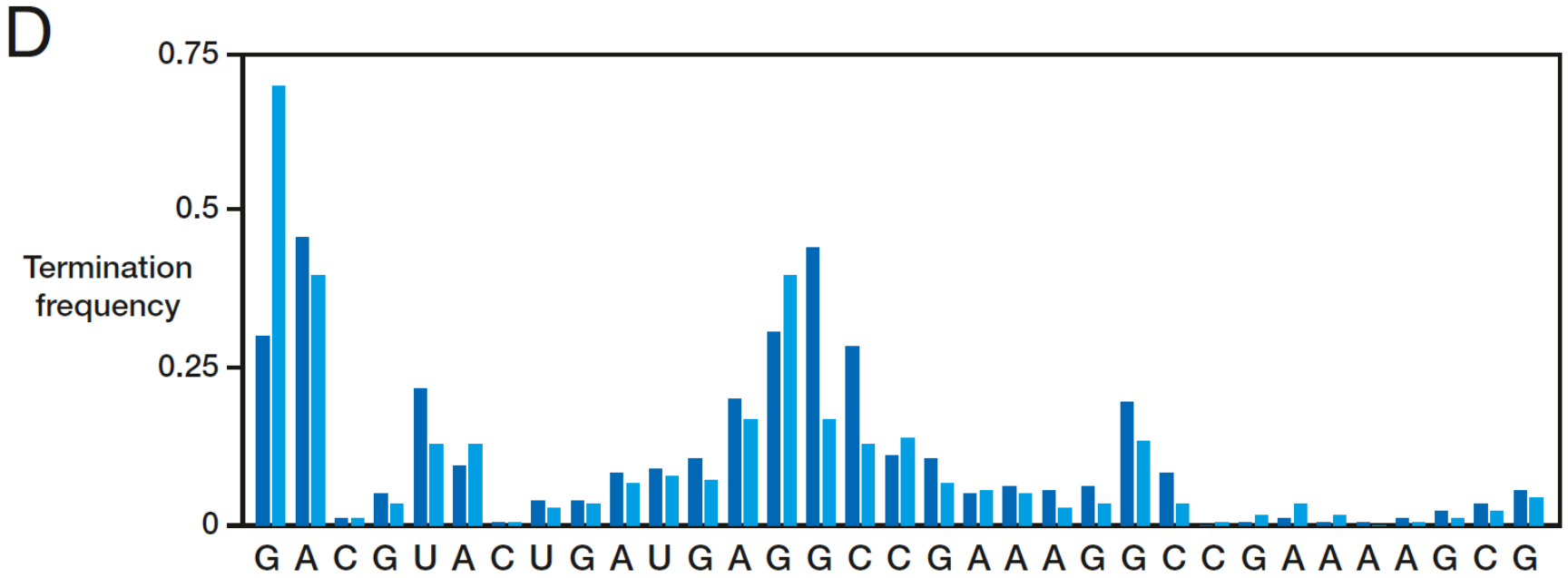


Fidelity as a function of distance from the 3’-terminal nucleotide. [3]

- 24-3 in 24 h (dark blue)
- 38-6 in 1 h (light blue)

3.4 Results (Difference between the two polymerases)

Difference in propensity for chain termination that results in partial-length products, with 38-6 being substantially less prone to early termination, especially at the most challenging template positions.



Positional termination frequencies across the hammerhead sequence for 24-3 (dark blue) and 38-6 (light blue).[3]

3.5. Summary and Outlook

1. 38-6 polymerase was selected based on its ability to synthesize functional ribozymes from NTP substrates.
2. 38-6 is able to synthesize its own ancestor (class I ligase) in the form of three fragments.
3. These fragments can self-assemble to give an active complex.

38-6 is the most complex ribozyme ever synthesized by a ribozyme from mononucleotide substrates.

This suggests that further improvements in activity might enable the polymerase to synthesize itself.

Even though the fidelity of polymerization, especially on the most challenging templates, severely limits the amount of functional information that can be copied.

This limitation would prevent the propagation of heritable information across successive generations of RNAs.

Going forward, it will be important to impose greater pressure on fidelity by requiring the polymerase to synthesize more complex RNAs that contain a larger number of nucleotides that are critical for function.

4. Conclusion

Question 1: How did RNA polymerase ribozymes evolve?

The results show how complex RNA structures could have been generated non-enzymatically in a plausible prebiotic setting by the iterative ligation of multiple Imp activated RNA oligomers.

This shows how complex RNA structures could have emerged from pools of activated RNA oligomers.

Question 2: Are RNA polymerase ribozymes able to copy functional RNA molecules and the polymerase?

The experiments show, how improved polymerase can synthesize a functional copy of its own evolutionary ancestor (class I ligase) by generating three RNA fragments.

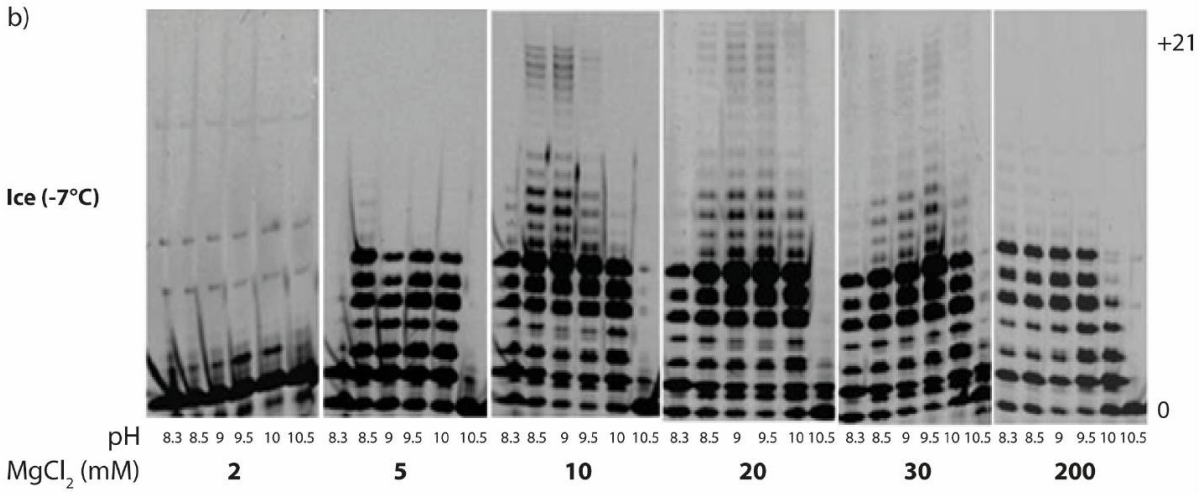
This hints at a possible way of propagating heritable information across successive generations.

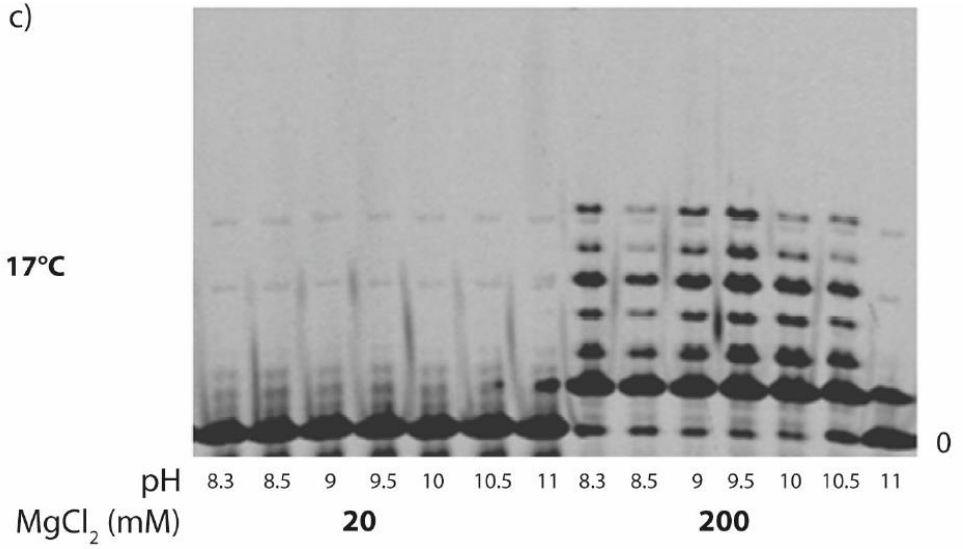
Thank you!

References

1. Wikipedia; 2021; RNA world; https://en.wikipedia.org/wiki/RNA_world
2. Dr. Falk Wachowius, Dr. Philipp Holliger; Non-Enzymatic Assembly of a Minimized RNA Polymerase Ribozyme; CHemSystemsChem. 2019 July; DOI:10.1002/syst.201900004
3. Katrina F. Tjhung, Maxim N. Shokhirev, David P. Horning, and Gerald F. Joyce; An RNA polymerase ribozyme that synthesizes its own ancestor; Proceedings of the National Academy of Sciences Feb 2020, 117 (6) 2906-2913; DOI: 10.1073/pnas.1914282117
4. Wikipedia; In vitro compartmentalization; 2021; https://en.wikipedia.org/wiki/In_vitro_compartmentalization

2.1 Optimizing reaction conditions





PH dependence (pH 8.3-11) of the RPR (Z) catalysed extension on primer/template (P HP/T HP) in ice at -7°C (10d) using different MgCl2 conc. (2 mM, 5 mM, 10 mM, 20 mM, 30 mM, 200 mM) [2]

PH dependence (pH 8.3-11) of the RPR (Z) catalysed extension on primer/template (P HP/T HP) in ice at at 17°C (4d) applying two different MgCl2 concentrations (20 mM, 200 mM). [2]

2.2 Optimizing RPR activity while minimalizing RPR size

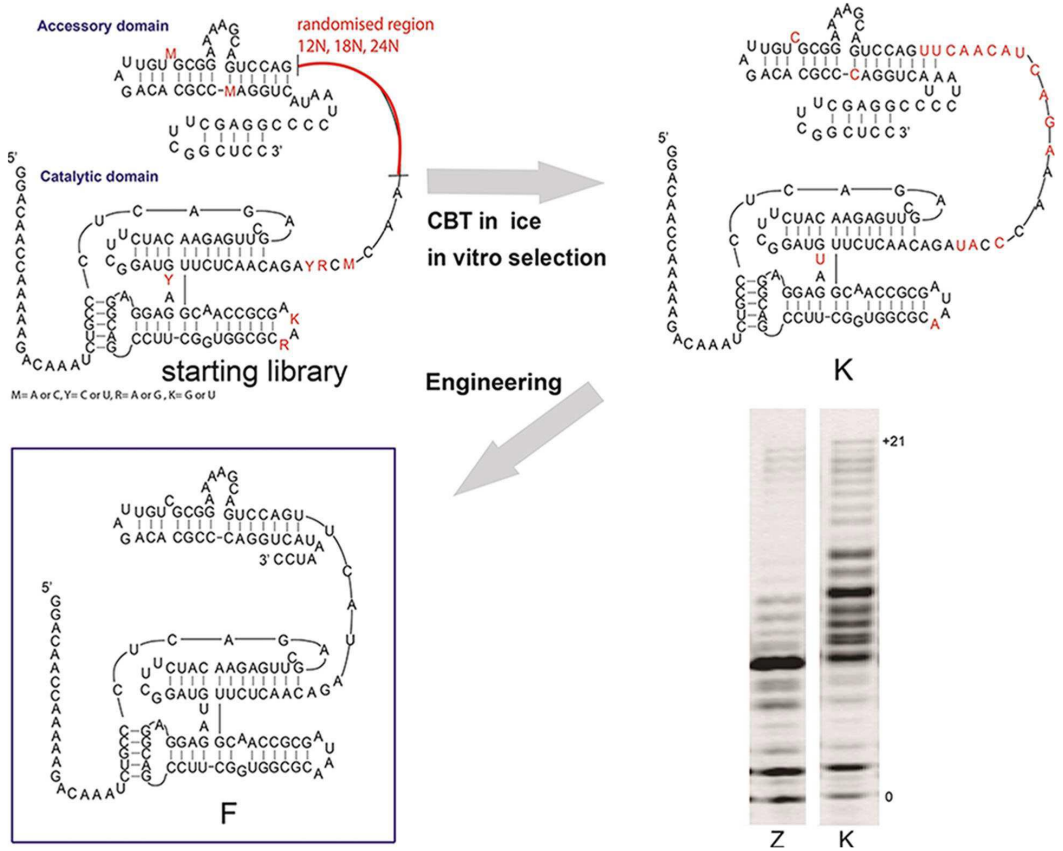
Generation of the F RNA polymerase ribozyme. [3] In-ice selection ofdifferent linker domain lengths (12–24 N) and defined mutations (red) resulted in new K RPR with improved activity compared to the Z RPR (bottom right panel). The K linker and accessory domain could be further truncated without loss of activity to yield the minimized F RPR.

CBT = Compartmentalized bead-tagging

In vitro compartmentalization (IVC) is an emulsion-based technology that generates cell-like compartments in vitro.

These compartments are designed such that each contains no more than one gene.

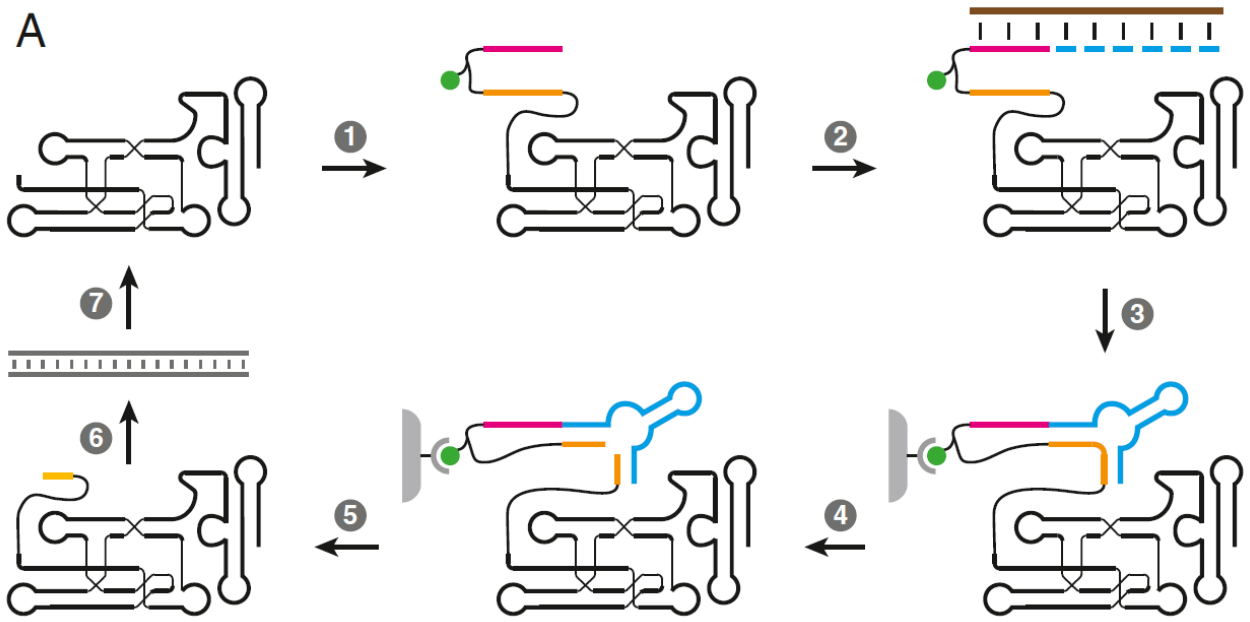
When the gene is transcribed and/or translated, its products (RNAs and/or proteins) become 'trapped' with the encoding gene inside the compartment. By coupling the genotype (DNA) and phenotype (RNA, protein), compartmentalization allows the selection and evolution of phenotype. [4]



3.1 In vitro evolution of the 38-6 RNA polymerase ribozyme

Scheme for selective amplification of polymerase ribozymes that synthesize a functional hammerhead ribozyme.

1. Attachment to the polymerase of an RNA primer (magenta), biotin (green), and the RNA substrate (orange) to be cleaved by the hammerhead.
2. Hybridization of the primer to an RNA template (brown) that encodes the hammerhead.
3. Extension of the primer by polymerization of NTPs (cyan), followed by biotin capture on streptavidin magnetic beads (grey).
4. Cleavage of the attached RNA substrate by the hammerhead, releasing the polymerase from the beads.
5. Recovery of functional polymerases.
6. Reverse transcription and PCR amplification.
7. Transcription to generate progeny polymerases.

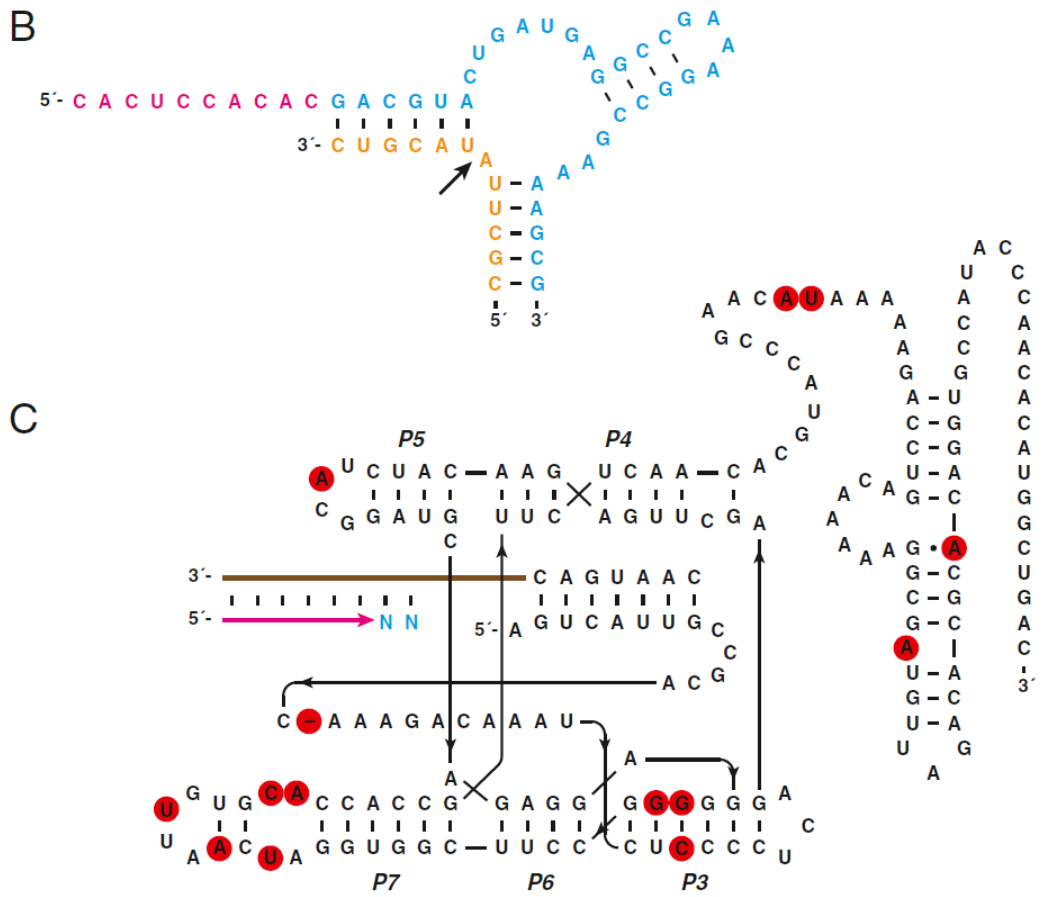


In vitro evolution of the 38-6 RNA polymerase ribozyme[3]

3.1 In vitro evolution of the 38-6 RNA polymerase ribozyme

B Sequence and secondary structure of the hammerhead ribozyme (cyan), together with the primer used to initiate its synthesis and the RNA substrate. The arrow indicates the site of cleavage.

C Sequence and secondary structure of the 38-6 RNA polymerase. Red circles indicate mutations relative to the 24-3 polymerase. Stem elements P3–P7 within the core domain are labelled.



3.3 Polymerization of own ancestor in three segments

A divide-and-conquer strategy was adopted to enable the 38-6 polymerase to synthesize the class I ligase in its entirety, splitting the ligase into three fragments that can assemble noncovalently to form a functional ribozyme.

The resulting three-part class I ligase is only approximately sevenfold less active than the contiguous form of the b1-207t ribozyme.

To enable the RNA-catalyzed synthesis of the entire class I ligase and its complement, the primers used to initiate synthesis must not include any nucleotides of the ribozyme.

→ Addition of external sites to both ends of each fragment → class I ligase assembled from the three fragments is approximately eightfold less active compared with an assembled three-fragment ligase that does not include the added primer regions.

Following a 3-d incubation, fragments 1, 2, and 3 were generated from their corresponding RNA templates in yields of 1.1%, 1.0%, and 5.7%, respectively. The polymerase was also able to synthesize the complement of each fragment, in yields of 0.53%, 0.91%, and 2.0%, respectively.

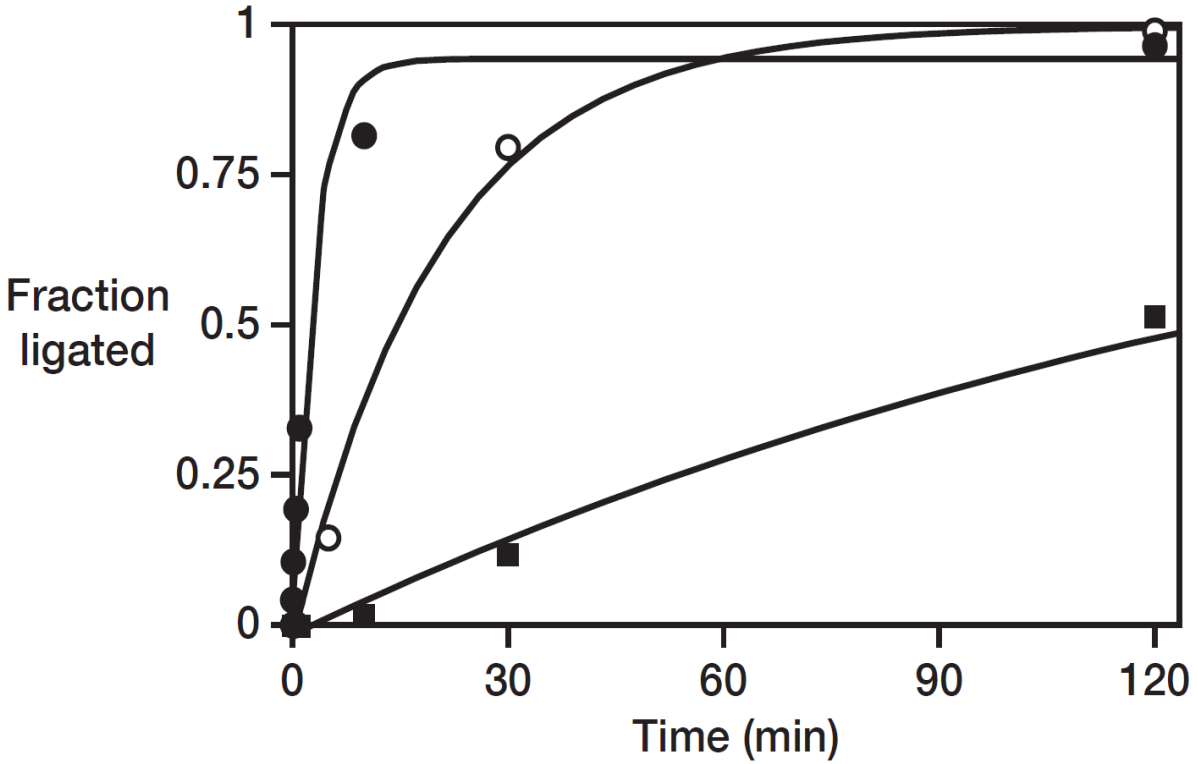
ligase is assembled from three fragments that have been synthesized by the 38-6 ribozyme, then there is an ∼8,000-fold reduction of ligase activity, with an observed rate of only $6.1 \pm 0.2 \times 10{-}5\ h{-}1$. This rate is only sevenfold faster than that of the uncatalyzed, RNA-templated reaction. Thus, while the improved polymerase ribozyme has the ability to synthesize large complex RNAs, including a three-fragment form of its own ancestor, the fidelity of synthesis is insufficient to enable a significant fraction of the products to retain the information necessary for catalytic function.

3.3 Polymerization of own ancestor in three segments

Time course of RNA ligation catalyzed by various forms of the class I ligase, all prepared by in vitro transcription using T7 RNA polymerase.

The ligase was provided as a contiguous strand (black circles), 3 separate fragments (white circles), or 3 separate fragments with added primer regions at both ends of each fragment (black squares).

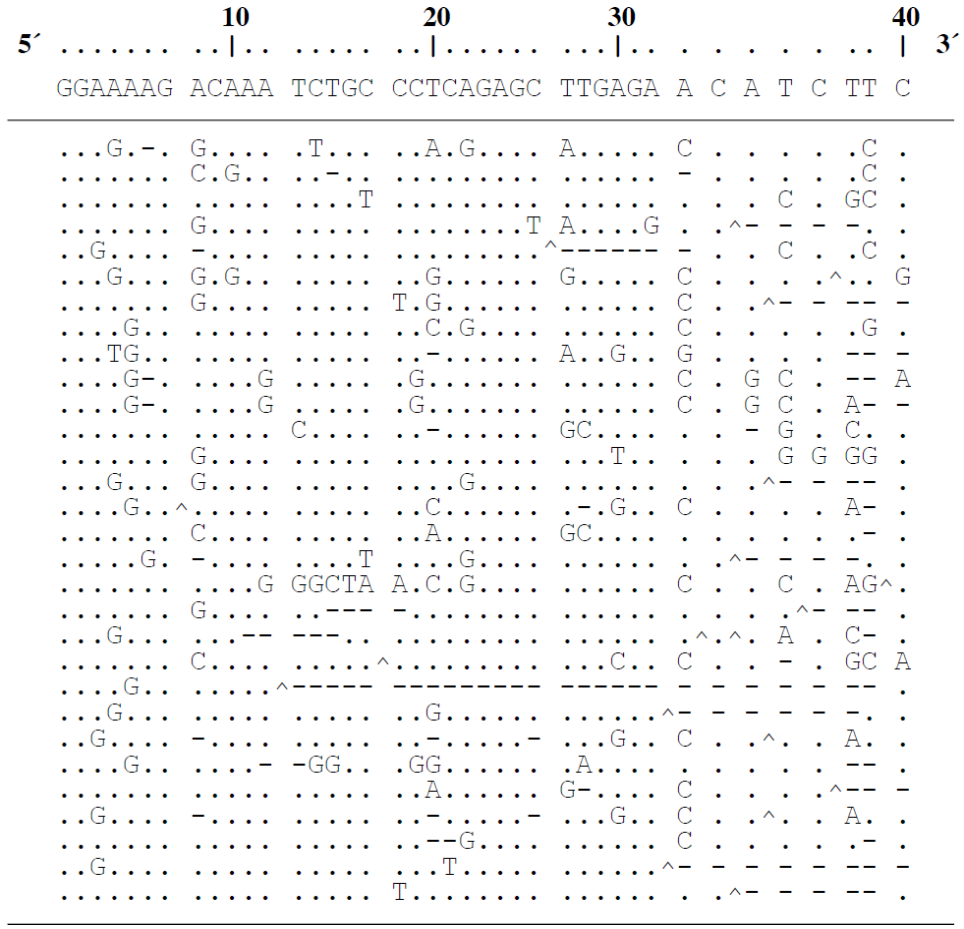
Multiple-turnover, first-order rate constants are 0.33, 0.050, and 0.0062 min–1, respectively. Reaction conditions: 1 µM ligase (either contiguous or each fragment), 20 µM 5´-substrate (S2), 80 µM 3´-substrate with attached template (S3), 60 mM MgCl2, 200 mM KCl, and 0.6 mM EDTA at pH 8.3 and 23 °C.



Time course of RNA ligation catalyzed by various forms of the class I ligase, all prepared by in vitro transcription using T7 RNA polymerase.[3]

3.4 Results (38-6 polymerase ribozyme generates far more complex RNA - at an expense)

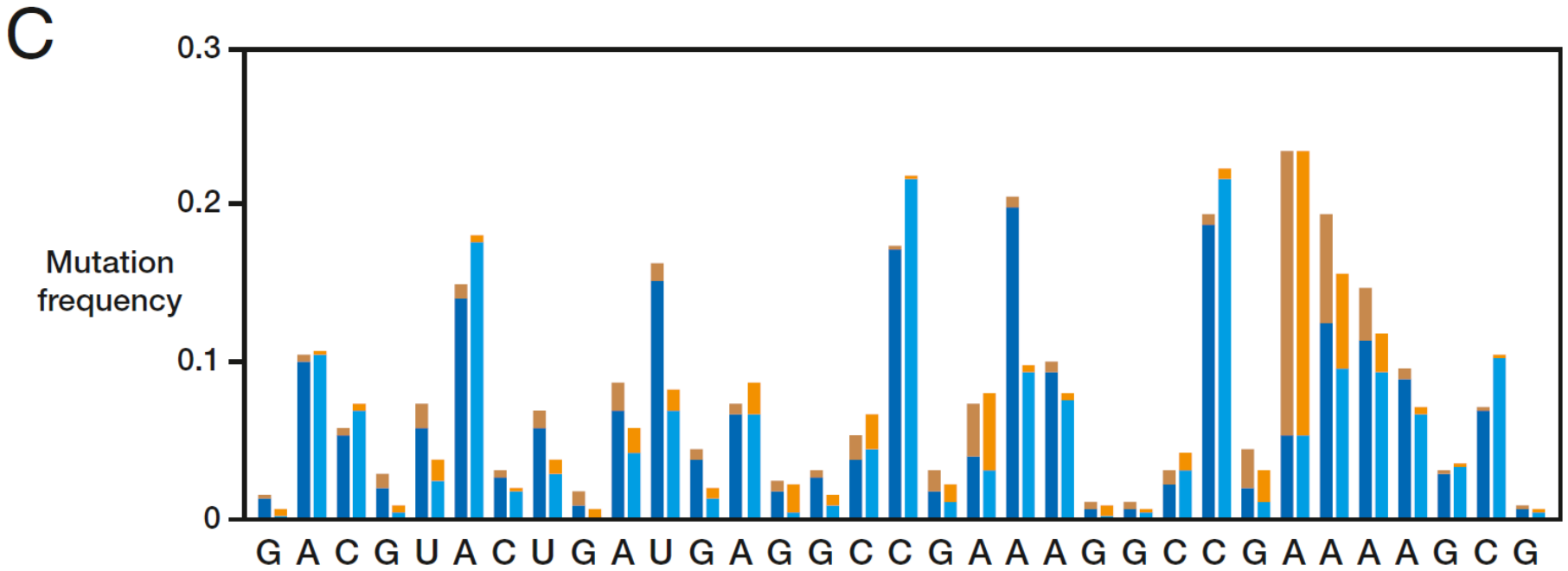
The 38-6 polymerase ribozyme can generate far more complex RNA products than could be achieved previously, but only at the expense of reduced fidelity for the most challenging cases.



Sequences of cloned individuals obtained from synthesis of fragment 1 of the class I ligase by the 38-6 polymerase [3]
Sequences of 30 individuals, aligned in reference to the correct sequence of fragment 1. The flanking primer regions are not shown. Dots indicate nucleotides that match the correct sequence, dashes indicate sites of deletion, and carat marks indicate sites of insertion of one or more nucleotides.

3.4 Results (Similarity between the two polymerases)

Similarity between the two polymerases with regard to the position-specific frequency of mutation, which ranges from 76% to 99%, depending on the difficulty of copying various template positions.



Positional mutation frequencies across the entire hammerhead sequence, with substitutions shown in blue and insertions/deletions shown in orange, for 24-3 (darker colors) and 38-6 (lighter colors). [3]

3.4 Results (Investigation of the relationship between yield and fidelity)

The frequency of mutation is greatest at positions that are susceptible to wobble mutation.

Pol	Expected	Observed (%)					
		A	G	C	U	Del	Ins
24-3	A	**88.0**	6.7	2.7	0.8	1.8	0.2
	G	0.5	**97.5**	0.7	0.7	0.6	0.1
	C	0.4	0.5	**93.2**	5.6	0.4	0.1
	U	1.5	1.9	5.4	**90.1**	1.1	0.1
38-6	A	**88.5**	7.0	1.2	0.4	3.0	0.2
	G	0.3	**98.5**	0.4	0.2	0.6	0.1
	C	0.3	0.7	**89.7**	8.8	0.6	0.1
	U	0.7	1.1	2.6	**94.7**	1.0	0.1

Fidelity of the 24-3 and 38-6 polymerases[3]

The average fidelities of 24-3 and 38-6 were 92.1 and 92.8%. Wobble mutations are underlined.