Ribozyme-Catalyzed Transcription of an Active Ribozyme

Aniela Wochner, *et al.*

Science **332**, 209 (2011);
DOI: 10.1126/science.1200752

33. G. B. Winkelman *et al.*, *Appl. Phys. Lett.* **87**, 061911 (2005).
34. S. H. Oh, Y. Kauffmann, C. Scheu, W. D. Kaplan, M. Rühle, *Science* **310**, 661 (2005).
35. J. J. Kuna *et al.*, *Nat. Mater.* **8**, 837 (2009).
36. W. D. Kaplan, Y. Kauffmann, *Annu. Rev. Mater. Res.* **36**, 1 (2006).
37. M. Baram, thesis, Technion–Israel Institute of Technology (2010).
38. R. Kobayashi, J. A. Warren, W. C. Carter, *Physica D* **140**, 141 (2000).
39. C. M. Bishop, R. M. Cannon, W. C. Carter, *Acta Mater.* **53**, 4755 (2005).
40. J. W. Cahn, *J. Chem. Phys.* **66**, 3667 (1977).
41. M. Tang, W. C. Carter, R. M. Cannon, *Phys. Rev. Lett.* **97**, 075502 (2006).
42. D. R. Clarke, *J. Am. Ceram. Soc.* **70**, 15 (1987).
43. A. Avishai, W. D. Kaplan, *Acta Mater.* **53**, 1571 (2005).
44. D. M. Lipkin, J. N. Israelachvili, D. R. Clarke, *Philos. Mag. A* **76**, 715 (1997).
45. K. Johnston, M. W. Finnis, *J. Am. Ceram. Soc.* **85**, 2562 (2002).
46. We thank H. Meltzman, P. Wynblatt, and W. C. Carter for fruitful discussions and invaluable comments. This work was supported by the Israel Science Foundation (#163/05) and the Russell Berrie Nanotechnology Institute at the Technion. M.B. acknowledges support from the Israeli Ministry of Science via an Eshkol Fellowship. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement FP7-NMP-2009-CSA-233484.

Supporting Online Material
www.sciencemag.org/cgi/content/full/332/6026/206/DC1
Materials and Methods
Tables S1 and S2
References

13 December 2010; accepted 11 February 2011
10.1126/science.1201596

Ribozyme-Catalyzed Transcription of an Active Ribozyme

Aniela Wochner, James Attwater, Alan Coulson, Philipp Holliger*

A critical event in the origin of life is thought to have been the emergence of an RNA molecule capable of replicating a primordial RNA "genome." Here we describe the evolution and engineering of an RNA polymerase ribozyme capable of synthesizing RNAs of up to 95 nucleotides in length. To overcome its sequence dependence, we recombined traits evolved separately in different ribozyme lineages. This yielded a more general polymerase ribozyme that was able to synthesize a wider spectrum of RNA sequences, as we demonstrate by the accurate synthesis of an enzymatically active RNA, a hammerhead endonuclease ribozyme. This recapitulates a central aspect of an RNA-based genetic system: the RNA-catalyzed synthesis of an active ribozyme from an RNA template.

An earlier, simpler biology might have relied on RNA for both heredity and metabolism. Evidence for such an "RNA world" (*1*) preceding modern life includes the central catalytic and informational roles of RNA in splicing, gene expression, and translation (*2*–*4*), as well as the versatility of RNA in forming specific receptors and catalysts (*5*, *6*). Organisms of the putative RNA world would have required an RNA polymerase ribozyme for both RNA-based heredity and the expression of "RNA genes." Although the ancestral replicase appears to have been lost, key functional aspects of RNA-catalyzed RNA replication can be studied "by proxy" with the use of modern ribozymes generated by in vitro selection, such as the R18 RNA polymerase ribozyme (Fig. 1A) (*7*). R18 was isolated from a random sequence pool by in vitro evolution and stepwise engineering of the initial class I ligase ribozyme (*8*–*10*). Although R18 is a general RNA polymerase, its activity is both sequence-dependent and limited to transcribing stretches of RNA up to 14 nucleotides (nt) long on a favorable RNA template (*7*). Starting from R18, we have used both RNA evolution and engineering to generate new RNA polymerase ribozymes with improved polymerase activity and sequence generality (*11*).

Compartmentalized bead-tagging. Strategies for the in vitro selection of catalytically active RNAs commonly involve the acquisition (or loss) of a capture tag for recovery (*12*, *13*). This restricts direct selection for more advanced enzymatic properties, such as multiturnover catalysis. Directed evolution of RNA polymerase ribozymes is complicated further by their poor primer extension efficiency: Substantial extension of a single primer remains a sporadic event, which reduces the effective size of polymerase ribozyme libraries. Furthermore, ribozyme repertoires must be transcribed by, for example, T7 RNA polymerase, which outperforms likely ribozyme polymerase candidates and must therefore be absent during the selective step. Finally, a sensitive and sequence-specific detection strategy is required to distinguish between simple nucleotide transferase and genuine templated polymerase activities.

We developed a selection strategy, termed compartmentalized bead-tagging (CBT), that harnesses water-in-oil emulsion technology (*14*) to link ribozyme genes to thousands of copies of the corresponding ribozyme via microbead display (fig. S1). CBT separates transcription from subsequent selective primer extension and allows the selection of clonal ribozyme populations rather than individual ribozyme molecules. The sampling of the aggregate activity of thousands of clonal ribozymes and a sensitive fluorescence signal amplification of extended primers by rolling circle amplification (RCA) (figs. S1c, S2, and S3) enable the discovery and isolation of polymerase ribozymes as a function of their primer extension capability. We validated CBT using model selections of R18 ribozyme genes spiked into an excess of inactive R18 genes (R18i) (fig. S1b), which yielded up to a 10^5-fold enrichment of active genes (fig. S1d).

Ribozyme RNA polymerase selection. We sought to improve ribozyme-primer-template interactions, known to be poor in the R18 ribozyme (*7*, *15*, *16*), by appending a random-sequence RNA domain to the 5′-end of R18, a region implicated in interactions with the primer-template duplex (*17*). A short "stem" RNA (table S1) that completes R18 by formation of a helix (P2) (*7*) was omitted from this selection as its functional relevance had been called into question (*16*).

CBT selection was applied to this library (~5×10^7 sequences) (table S2). We developed a microtiter plate–based assay (RPA, ribozyme polymerase plate assay) (fig. S4) to screen for clones of interest. After three rounds of CBT, we observed an increase in polyclonal ribozyme activity (fig. S5a) and, by RPA, identified one ribozyme (C19) with improved RNA polymerase activity among 22 clones screened (Fig. 2, A and B).

C19 differed from the parent R18 ribozyme by the presence of a new 5′-domain, as well as a single point mutation (G93A) (Fig. 1B). The G93A mutation was located in the P2 region of the ribozyme and compensated for the reduction in polymerase activity caused by the absence of the stem oligonucleotide. Indeed, the A93G back-mutation rendered C19 RNA polymerase activity stem-dependent (fig. S5b).

Secondary structure prediction of the new 5′-domain indicated that it comprised a 6-nt single-stranded sequence segment (ss_{C19}) at the 5′ end, followed by a hairpin domain (HP_{C19}). ss_{C19} showed complementarity to the 5′ end of the template (TI) that had been used for both selection and screening. Binding of ss_{C19} to the template was required for C19 activity, as templates containing successive mutations of the ss_{C19} complementary sequence showed a progressive loss in activity (Fig. 2C). Compensating mutations in ss_{C19} (to reconstitute a 6-nt hybridization site) restored activity, although not consistently to C19 levels. This suggests that ss_{C19} enhances C19 activity primarily (but not exclusively) by promoting recognition and binding of template downstream of the primer 3′ end.

Long-range RNA synthesis. Although C19 showed improved polymerase activity on the selection template TI (Fig. 2B), the activity enhancement of C19 compared with R18 was much

Medical Research Council (MRC) Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK.

*To whom correspondence should be addressed. E-mail: ph1@mrc-lmb.cam.ac.uk

less pronounced on longer templates, where the ss_{C19} binding site was located farther downstream of the primer (Fig. 2D). We speculated that C19 was structurally adapted to extending primers close to the ss_{C19} binding site. We therefore engineered a series of 5′-domain variants of C19, where either single secondary-structure elements [sequences upstream of ss_{C19} (R10), ss_{C19}, and HP_{C19}], or combinations thereof, were omitted (fig. S6). One of these truncation variants, tC19 (Fig. 1C), lacking sequences upstream of ss_{C19} and with HP_{C19} replaced by a four-residue adenine (A_4) spacer, showed improved RNA polymerase activity on the longer template TI-3 compared with the parent C19 ribozyme (Fig. 2D), and extended 13% of primers by over 26 nt (table S4).

In the putative RNA world, both RNA replicases and replication templates would have been under selective pressure to coevolve toward maximum replication efficiency. We sought to mimic aspects of this adaptive process by evolving tem-

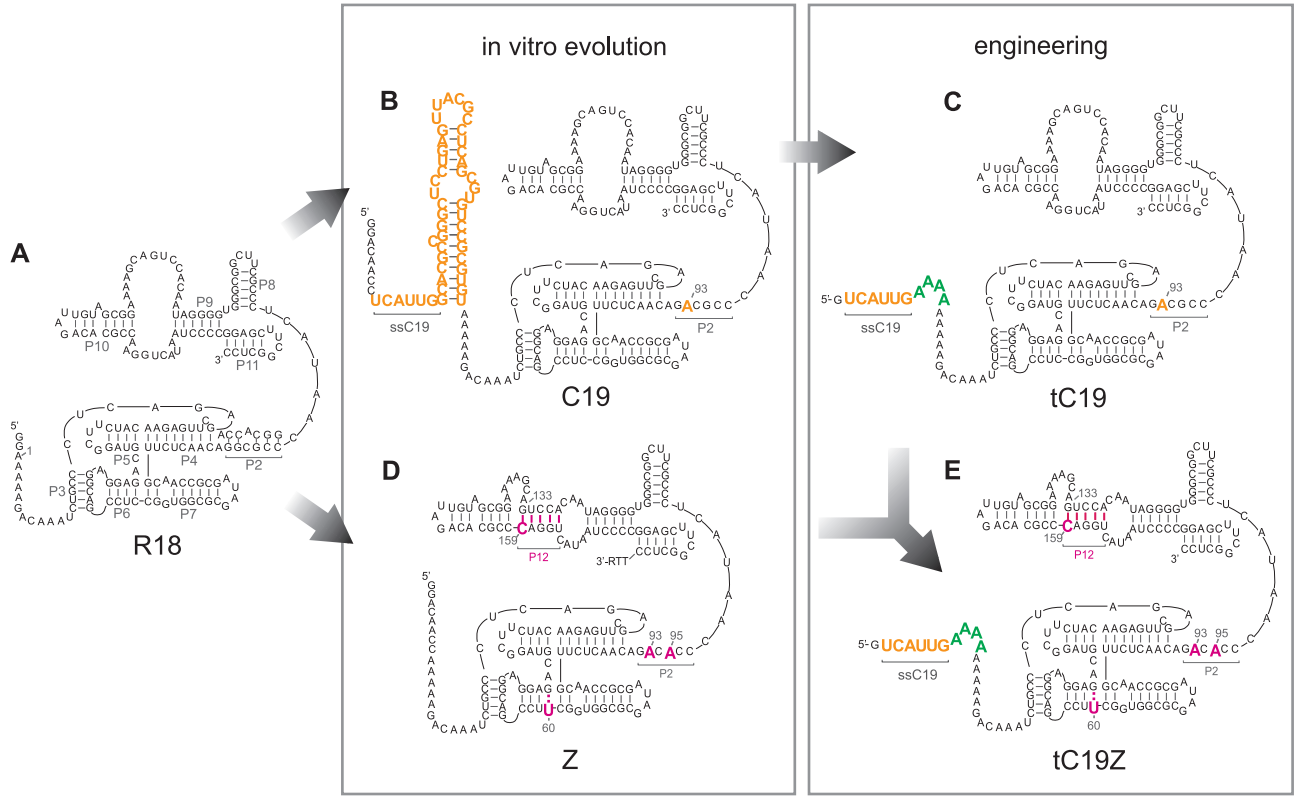

Fig. 1. Evolved and engineered ribozymes. Secondary structures of (**A**) R18 (*7*); (**B**) C19, as predicted by mfold (*27*); (**C**) tC19; and proposed secondary structures for (**D**) Z; and (**E**) tC19Z. Mutations isolated from the Z selection are depicted in magenta, sequences isolated from the C19 selection in orange, and engineered residues in green. The A159C mutation is shown to reshape the processivity domain by stabilizing helix P12. [Run-through transcript, 3′-RTT (table S1).]

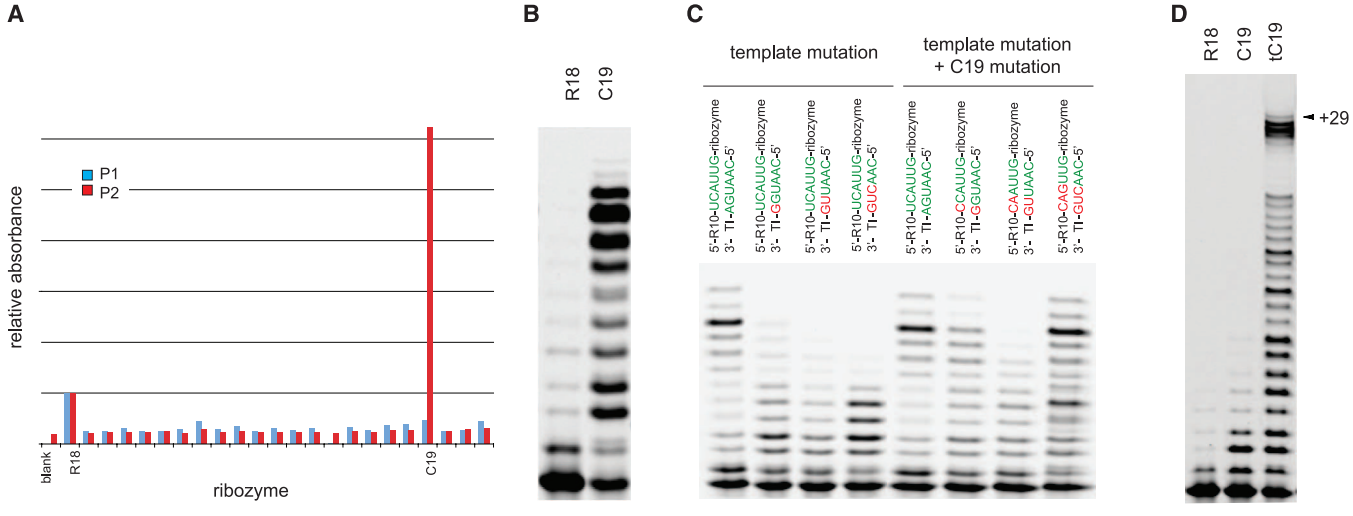

Fig. 2. Selection and engineering of C19. (**A**) RPA of clones from the N48 library selection after three rounds of CBT probed for basic activity (probe P1, blue) and more stringently (probe P2, red). Absorbance at 450 nm (as OD) is shown normalized to R18. (**B**) Primer extension activity of R18 and C19 ribozymes on the selection template TI. (**C**) Primer extension by C19 on templates with successive point mutations (red) in the ss_{C19} binding site ("template mutation"), and with compensating mutations (to reconstitute a 6-nt hybridization site) in C19 ("template mutation + C19 mutation"). ss_{C19} and its respective binding site are depicted in green. (**D**) Primer extension activity of tC19 compared with the R18 and C19 ribozymes on template TI-3.

plates adapted to tC19 using the downstream ss_{C19} recognition site. We applied a template selection scheme (fig. S7) to a library of templates based on TI-3 (table S1). After two rounds of template selection, ~50% of sequences isolated were derived from a single template (TI-5), in which TI-3 had been extended by two additional repeats of the central 11-nt sequence. TI-5 could be transcribed to ≥47 nt by tC19, with 1.5% of primers extended beyond this point (table S4). We generated a series of templates based on the same design principles containing increasing numbers of the 11-nt repeat (TI-1 to TI-10) (table S1). On these templates, tC19 could extend primers by up to 95 nt (Fig. 3) with a fidelity (expressed as the probability of mutation per incorporated nucleotide) of 2.7×10^{-2} as determined by sequencing full-length extension products (fig. S8). Extension products of ≥91 nt were synthesized with a yield of 0.035% of total primer (table S4). Yields of full-length products are limited by a number of factors, including chain termination and ribozyme and product degradation (*18*). Together, these result in a 7% termination probability (on average) per template position (table S4).

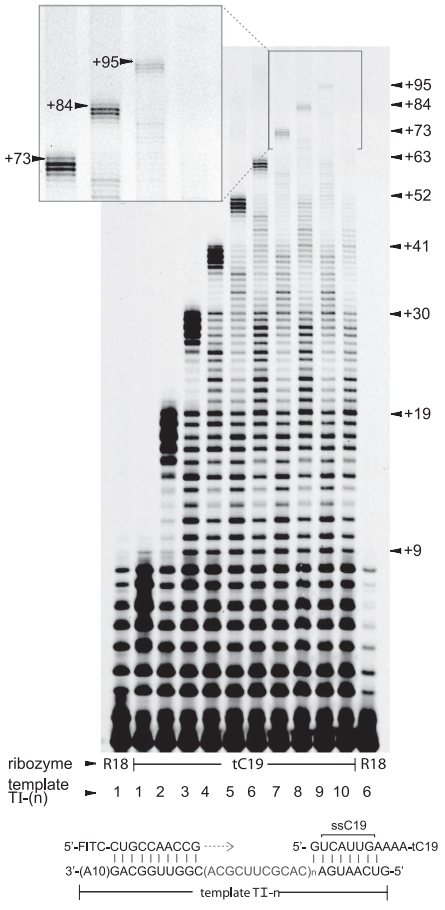

Fig. 3. Long-range RNA synthesis on engineered template series TI-n by the tC19 and R18 ribozymes (7 days). "n" indicates the number of repeats of the central 11-nt sequence between primer and ss_{C19} binding sites. The schematic depicts primer extension by tC19 on TI-n.

Sequence generality. tC19 enabled the synthesis of some long RNAs, but its polymerase activity, like that of its parent R18, remained template-dependent. Although capable of long primer extension on favorable templates, synthesis from a majority of RNA template sequences was limited. To improve generality, we performed a second series of selection experiments starting from a library of 5×10^7 randomly mutated R18 genes. This starting pool was subjected to three rounds of low-stringency CBT selection to deplete it of detrimental mutations. We then generated new combinations of neutral and beneficial mutations through recombination (*19*), before applying five additional rounds of increasingly stringent CBT selection, using two distinct selection templates to encourage the emergence of generality (table S3), after which the activity of this polyclonal pool rose to exceed wild-type R18 activity. Screening by RPA identified a clade of ribozymes (fig. S4c) with substantially improved activity comprising between three and five mutations. We analyzed the contributions of these mutations to disentangle their effects on RNA polymerase activity (fig. S9). Combination of the most favorable mutations (C60U, G93A, G95A, and A159C) in a single RNA molecule yielded Z (Fig. 1D), a ribozyme RNA polymerase with improved sequence generality (fig. S10). Combining Z's core mutations with the 5′ extension of tC19 yielded the hybrid ribozyme tC19Z (Fig. 1E) with improved sequence generality and polymerase activity. Although still not independent of the template sequence, this ribozyme outperformed all of its parent ribozymes (R18, tC19, and Z), synthesizing longer extension products on a range of different primer-template sequences (designed to comprise approximately equal representation of all four nucleobases, as well as all 16 dinucleotide combinations) (Fig. 4A). tC19Z also exhibited a further improvement in fidelity (8.8×10^{-3}), as determined by sequencing of full-length extension products (fig. S8).

Hammerhead ribozyme synthesis by tC19Z. The improvements in RNA polymerase activity and fidelity in the tC19Z ribozyme encouraged us to explore the synthesis of an RNA sequence that encodes a catalytic activity: a hammerhead nuclease ribozyme. To facilitate the synthesis of sufficient amounts of full-length ribozyme for characterization, we chose a minimal version of the hammerhead endonuclease designed for therapeutic applications (*20*) (Fig. 4B). In contrast to R18, the tC19Z RNA polymerase ribozyme could synthesize full-length (>24 nt) hammerhead minizymes (Fig. 4C, left, and fig. S11). Both the +24-nt minizyme and a pool of longer extension products (≥27 nt) exhibited catalytic activity and performed sequence-specific cleavage of a cognate substrate RNA (Fig. 4C, right).

The tC19Z phenotype arises from the contributions of four mutations (C60U, G93A, G95A,

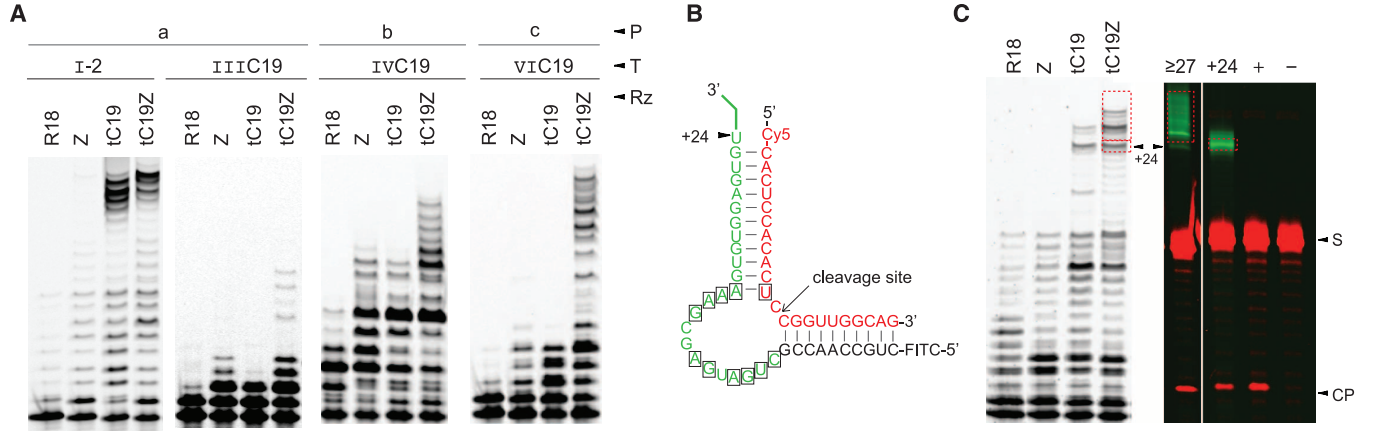

Fig. 4. Generality and hammerhead synthesis. (**A**) Primer extension by R18, Z, tC19, and tC19Z on different primer-template duplexes (primer, P; template, T; ribozyme, RZ). (**B**) Secondary structure of hammerhead endonuclease minizyme with ribozyme-synthesized segment (green) and substrate (red). Essential catalytic residues are boxed. (**C**) Fluorescent primer extension on minizyme templates by ribozymes R18, Z, tC19, and tC19Z (left). tC19Z-synthesized extension products long enough to form a symmetrical minizyme with the substrate (+24 and ≥+27, red boxes) were prepared (fig. S11) and tested for endonuclease activity (right). One to 3% of substrate is cleaved by the control minizyme (+) and by both tC19Z-synthesized minizymes (+24 and ≥+27) with comparable efficiencies (table S5), but not in their absence (–) (substrate, S; cleavage product, CP).

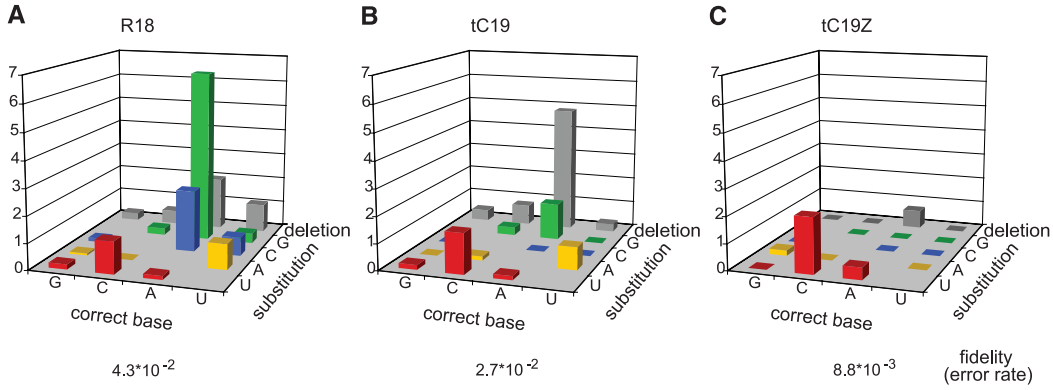

Fig. 5. RNA extension fidelity. Error spectra (defined, on the *y* axes, as the average percentage of incorrect nucleotides inserted, per position) of full-length extension products synthesized by RNA polymerase ribozymes R18 (**A**), tC19 (**B**), and tC19Z (**C**). The error rate denotes the average probability of a mutation per nucleotide position (fig. S8). Fidelity data for R18 is derived from Attwater *et al.* (*18*).

and A159C), as well as a short 5′-extension (5′-GUCAUUGAAAA) to the parental R18 ribozyme. The C60U mutation incrementally enhances RNA polymerase activity; it is the only mutation selected in the catalytic core and changes a G-C base pair to a weaker G-U wobble pair in the central P6 stem. The G93A and G95A mutations in the P2 region disrupt interaction with (and render ribozyme activity independent of) the stem oligonucleotide. G93A and G95A may be necessary to tolerate a single-stranded P2 region in the absence of stem or, alternatively, could promote new interactions with the primer-template duplex or the ribozyme processivity domain. A159C is located in the processivity domain (residues 98 to 187), as part of a large asymmetric internal loop. A159C allows the formation of a new G133:C159 base pair, which augments a potential 4–base pair stem (P12) formed by bulge segments A129 to U132 (5′-ACCU) and A160 to U163 (5′-AGGU) (Fig. 1, D and E). Indeed, G133U, a separate beneficial mutation that was isolated in the selection, would stabilize the P12 stem in an identical way, by promoting formation of an U133:A159 base pair. These two mutations, although beneficial individually [with A159C superior to G133U (fig. S9)], negate one another's effects when combined. This strongly suggests that the selected trait is indeed the formation of a new base pair between positions 133:159, presumably to reshape the structure of the processivity domain. Indeed, a recent study on the processivity domain of the closely related B6.61 RNA polymerase ribozyme also suggested the presence of a 4–base pair P12 stem (*21*).

The ss_{C19} sequence tag emerged rapidly in the selection and appears to mediate interaction with the 5′ end of the template strand. This interaction is reminiscent of the recognition of target mRNAs by the prokaryotic ribosome through the Shine-Dalgarno sequence (*22*) and primordial mechanisms of template recognition as proposed by the genomic tag hypothesis (*23*). Such recognition might have been advantageous in a prebiotic setting, where an RNA polymerase ribozyme would not have evolved in isolation but in the presence of a large number of unrelated RNA oligomers. In such an environment, specific interactions between a replicase and cognate templates via a recognition tag that promotes polymerization (Fig. 2C) could have supported a primitive form of kin selection, which would enable self-replication and Darwinian evolution even in the absence of compartmentalization in protocellular entities.

RNA replication must proceed with a minimum fidelity, as defined by the "error threshold" (*24*, *25*), to avoid corruption of encoded genetic information. tC19Z contains the C60U mutation in the catalytic core, as well as several other mutations that could potentially affect its fidelity. We determined aggregate fidelities (both deletion mutations and nucleotide misincorporations) of the different ribozymes by sequencing full-length extension products (Fig. 5). This method measures the fidelity of the ribozyme for the synthesis of a given sequence (*7*, *18*), however, truncated extension products are not probed for potential errors (see supporting online material text S1). By this measure, we determined tC19Z's fidelity as 8.8×10^{-3} (fig. S8), which suggested an improvement compared with both the parent R18 (4.3×10^{-2}) (*18*) and the intermediate tC19 (2.7×10^{-2}) ribozymes. Some RNA sequences are synthesized by tC19Z with even greater accuracy. Analysis of the hammerhead ribozymes synthesized by tC19Z revealed just two mutations in 37 sequenced clones (2× G→A transitions in 999 nt), which indicates a fidelity of 2×10^{-3}—a 20-fold improvement over the fidelity of the parent R18 ribozyme. Although comparisons are imperfect because of some differences between the sequences sampled for the different ribozymes, there is a notable reduction in A→G transition mutations in the mutation spectrum of tC19Z (and to a lesser extent in tC19), which suggests a reduced susceptibility to misincorporation of G opposite template U. On the other hand, the reverse C→U transition mutation (denoting misincorporation of U opposite template G) is not decreased, indicating an asymmetric, i.e., strand-specific recognition of the G•U wobble pair by tC19Z.

The R18 RNA polymerase ribozyme had been widely considered to be an evolutionary dead-end, trapped in a local activity optimum and largely refractory to further substantial evolution (*26*). Starting from R18, we have shown that the CBT selection method, in combination with RNA engineering, can yield polymerase ribozymes, such as tC19Z, that display enhanced polymerase activity and fidelity, as well as generality, and allow the ribozyme-catalyzed transcription of an active ribozyme from an RNA template.

References and Notes

1. R. Gesteland, T. R. Cech, J. F. Atkins, Eds., *RNA World: The Nature of Modern RNA Suggests a Prebiotic RNA World* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, ed. 3, 2005).
2. S. Valadkhan, J. L. Manley, *Nature* **413**, 701 (2001).
3. M. Mandal, B. Boese, J. E. Barrick, W. C. Winkler, R. R. Breaker, *Cell* **113**, 577 (2003).
4. P. Nissen, J. Hansen, N. Ban, P. B. Moore, T. A. Steitz, *Science* **289**, 920 (2000).
5. C. Tuerk, L. Gold, *Science* **249**, 505 (1990).
6. A. D. Ellington, J. W. Szostak, *Nature* **346**, 818 (1990).
7. W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner, D. P. Bartel, *Science* **292**, 1319 (2001).
8. D. P. Bartel, J. W. Szostak, *Science* **261**, 1411 (1993).
9. E. H. Ekland, J. W. Szostak, D. P. Bartel, *Science* **269**, 364 (1995).
10. E. H. Ekland, D. P. Bartel, *Nature* **382**, 373 (1996).
11. Materials and methods are available as supporting material on *Science* Online.
12. G. F. Joyce, *Annu. Rev. Biochem.* **73**, 791 (2004).
13. D. S. Wilson, J. W. Szostak, *Annu. Rev. Biochem.* **68**, 611 (1999).
14. D. S. Tawfik, A. D. Griffiths, *Nat. Biotechnol.* **16**, 652 (1998).
15. M. S. Lawrence, D. P. Bartel, *Biochemistry* **42**, 8748 (2003).
16. H. S. Zaher, P. J. Unrau, *RNA* **13**, 1017 (2007).
17. D. M. Shechner *et al.*, *Science* **326**, 1271 (2009).
18. J. Attwater, A. Wochner, V. B. Pinheiro, A. Coulson, P. Holliger, *Nat. Commun.* **1**, 10.1038/ncomms1076 (2010).
19. H. Zhao, L. Giver, Z. Shao, J. A. Affholter, F. H. Arnold, *Nat. Biotechnol.* **16**, 258 (1998).
20. M. J. McCall *et al.*, *Mol. Biotechnol.* **14**, 5 (2000).
21. Q. S. Wang, L. K. Cheng, P. J. Unrau, *RNA* **17**, 469 (2011).
22. J. Shine, L. Dalgarno, *Nature* **254**, 34 (1975).
23. A. M. Weiner, N. Maizels, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7383 (1987).
24. M. Eigen, *Naturwissenschaften* **58**, 465 (1971).
25. A. Kun, M. Santos, E. Szathmáry, *Nat. Genet.* **37**, 1008 (2005).
26. G. F. Joyce, *Science* **315**, 1507 (2007).
27. M. Zuker, *Nucleic Acids Res.* **31**, 3406 (2003).

28. We are grateful to M. Daly for assistance with FACS and to our colleagues V. B. Pinheiro, C. Cozens, A. Taylor, and J. Sutherland for helpful discussions. This work was supported by the Medical Research Council (MRC) program grant MC_US_A024_0014.

Supporting Online Material

www.sciencemag.org/cgi/content/full/332/6026/209/DC1
Materials and Methods
SOM Text
Figs. S1 to S11
Tables S1 to S5
References

22 November 2010; accepted 15 February 2011
10.1126/science.1200752