

Emergence of a fast-reacting ribozyme that is capable of undergoing continuous evolution

Sarah B. Voytek and Gerald F. Joyce*

Departments of Chemistry and Molecular Biology, The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037

Contributed by Gerald F. Joyce, August 8, 2007 (sent for review July 9, 2007)

It is possible to evolve RNA enzymes in a continuous manner by employing a simple serial-transfer procedure. This method was previously applied only to descendants of one unusually fast-reacting RNA enzyme with RNA ligase activity. The present study establishes a second continuously evolving RNA enzyme, also with RNA ligase activity, but with a completely independent evolutionary origin. Critical to achieving the fast catalytic rate necessary for continuous evolution, development of this enzyme entailed the addition and evolutionary maturation of a 35-nucleotide accessory domain and the application of highly stringent selection pressure, with reaction times as short as 15 ms. Once established, continuous evolution was carried out for 80 successive transfers, maintaining the population against an overall dilution of 10^{207} -fold. The resulting RNA enzymes exhibited $\approx 10^5$ -fold improvement in catalytic efficiency, compared with the starting molecules, and became dependent on a structural feature of the substrate that previously conferred no selective advantage. This adaptation was eliminated by deleting the substrate feature and then carrying out 20 additional transfers of continuous evolution, which resulted in molecules with even greater catalytic activity. Now that two distinct species of continuously evolving enzymes have been established, it is possible to conduct molecular ecology experiments in which the two are made to compete for limited resources within a common environment.

in vitro evolution | RNA catalysis | RNA ligase | selection | serial transfer

Darwinian evolution has been applied to populations of RNA molecules in the laboratory to generate RNA enzymes (ribozymes) that catalyze a variety of reactions (1–3). These experiments involve repeated rounds of selective amplification and mutagenesis to drive the evolving population toward desired catalytic behaviors. Of the many ribozymes that have been obtained by *in vitro* evolution, six different classes catalyze the RNA-templated joining of an oligonucleotide 3'-hydroxyl and an oligonucleotide 5'-triphosphate, forming a 3',5'-phosphodiester and releasing inorganic pyrophosphate (4–9). This reaction is similar to that carried out by protein polymerases that catalyze the synthesis of biological RNAs.

All six of the known 3',5'-ligase ribozymes achieve rate enhancements of several orders of magnitude, compared with the uncatalyzed rate of templated RNA ligation, which is $\approx 10^{-7} \text{ min}^{-1}$ (10). Notably, only derivatives of the class I ligase, first described by Bartel, Szostak, and colleagues (4, 11), have been able to achieve a catalytic rate of $>1 \text{ min}^{-1}$. Because of the superior catalytic rate of this ribozyme, it has been chosen as the starting point for several subsequent *in vitro* evolution experiments. One series of studies led to the development of a sequence-general, RNA-dependent RNA polymerase ribozyme that can add as many as 20 NTPs (12, 13). This effort involved several structural modifications of the class I ligase, including the addition of a large random-sequence accessory domain, followed by 24 rounds of stepwise evolution and extensive screening of cloned individuals.

Another line of investigation beginning with the class I ligase resulted in the establishment of a system for the continuous *in vitro* evolution of ribozymes (14). This system operates at a constant temperature within a common reaction vessel. It employs a pro-

moter-containing oligonucleotide substrate and two proteins, reverse transcriptase and RNA polymerase, to bring about the selective amplification of reactive ribozymes (Fig. 1). The RNA population expands exponentially and can be diluted serially to allow sustained exponential growth. The serial transfers also provide an opportunity to change the selection pressures applied to the evolving population, for example, lowering the Mg^{2+} concentration (15), raising or lowering the pH (16), or introducing a ribozyme-cleaving DNA enzyme (17).

Continuous evolution is the most efficient and powerful method for evolving functional molecules *in vitro*, but thus far has only been applied to descendants of the class I ligase. It has been hypothesized that continuous evolution could be used to evolve other ligase ribozymes and perhaps RNA molecules that catalyze other bond-forming reactions (18). However, no other ligase ribozyme has a catalytic rate as fast as that of the class I ligase. A fast catalytic rate is essential for continuous evolution because the ribozyme must react before it becomes reverse transcribed. Only if the ribozyme reacts before it is reverse transcribed will it acquire a promoter sequence that allows the resulting cDNA to give rise to progeny ribozymes.

The goal of the present study was to develop a second ribozyme that can evolve in a continuous manner. This goal is a challenge for RNA catalysis and an opportunity to develop a distinct species of continuously evolving ribozymes that might be made to operate in competition or cooperation with the class I ligase. The DSL ligase ribozyme, described previously by Inoue and colleagues (8), was chosen as the starting material for this effort. The DSL ligase was obtained by using a combination of rational design and *in vitro* selection. The framework for this ribozyme is a structural scaffold based on a naturally occurring RNA motif, to which was added a duplex region for templated ligation and a potential catalytic domain of 30 random-sequence nucleotides. A large pool of these RNA molecules was subjected to 10 rounds of *in vitro* selection, resulting in a variety of ligase ribozymes. The most active catalyst identified is the *cis*-DSL-1S ribozyme, which contains 140 nucleotides and has a k_{cat} of 0.066 min^{-1} and a K_m of 940 nM, measured in the presence of 50 mM MgCl_2 and 200 mM KCl at pH 7.5 and 37°C (8, 19).

The catalytic activity of the DSL ribozyme is noteworthy for a ligase obtained after only 10 rounds of selective amplification and derived from only 30 random-sequence nucleotides. The small size and modular construction of the DSL ligase were regarded as advantageous features for the development of a second continuously evolving ribozyme. In the present study, the DSL ribozyme was structurally modified to accommodate the promoter-containing substrate needed for continuous evolution, and an accessory domain of 35 random-sequence nucleotides was added to

Author contributions: S.B.V. and G.F.J. designed research; S.B.V. performed research; S.B.V. and G.F.J. analyzed data; and S.B.V. and G.F.J. wrote the paper.

The authors declare no conflict of interest.

*To whom correspondence should be addressed. E-mail: gjoyce@scripps.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0707490104/DC1.

© 2007 by The National Academy of Sciences of the USA

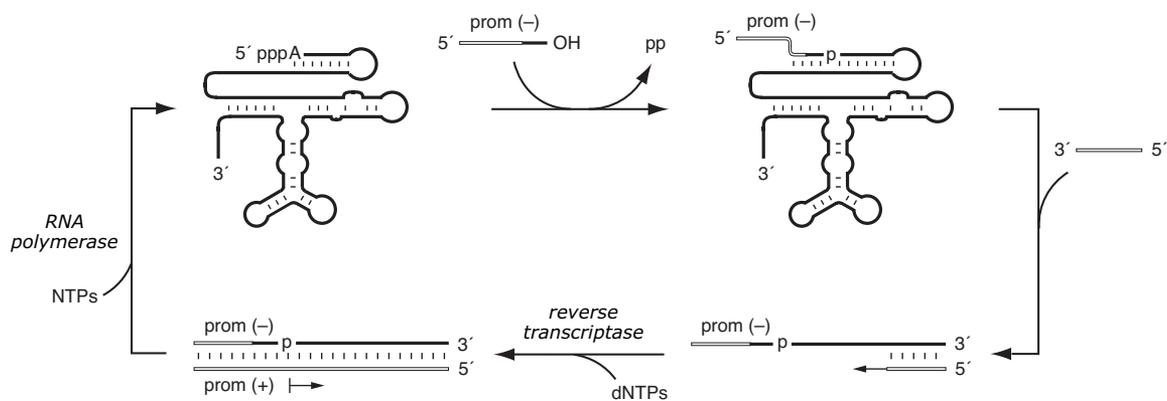


Fig. 1. Scheme for continuous evolution of ligase ribozymes. RNA and DNA are depicted as solid and open lines, respectively. Ligation of a chimeric DNA–RNA substrate by the ribozyme results in attachment of the T7 promoter sequence (prom) to the 5' end of the ribozyme. Reverse transcription of reacted, but not unreacted, ribozymes results in the formation of a cDNA that serves as a template for transcription of multiple progeny RNAs.

provide the opportunity for substantial improvement in catalytic rate. A population of these modified ribozymes was subjected to vigorous stepwise evolution, with reaction times as short as 15 ms, to drive the improvement of catalytic function. This stringent selection pressure ultimately resulted in a population of DSL-derived ribozymes that were capable of undergoing continuous evolution. Once established, continuous evolution can be continued indefinitely, resulting in further improvement in the catalytic activity of the ribozymes. This work demonstrates that continuous evolution can be applied to ribozymes in addition to the class I ligase, setting the stage for molecular ecology studies in which two or more species evolve within a common environment.

Results

Stepwise Evolution. The DSL ribozyme was modified in several ways to render it compatible with the continuous evolution system (Fig. 2A). Continuous evolution requires the RNA-catalyzed ligation of a chimeric DNA–RNA substrate that has the sequence of an RNA polymerase promoter. This substrate is more AU-rich at its 3' end, compared with the preferred substrate for the DSL ribozyme, making it necessary to extend the template region of the ribozyme so that it could bind the substrate with sufficient affinity. Extension of the template region necessitated removal of a triplex structural element, which was replaced by a flexible oligouridylyate linker (Fig. 2A). Additionally, the primer binding site at the 3' end of the ribozyme was replaced by nucleotides of a different sequence that were known to be compatible with continuous evolution.

The modified DSL ligase had substantially diminished activity with the promoter-containing substrate, exhibiting a k_{cat} of $1.76 \pm 0.02 \times 10^{-2} \text{ min}^{-1}$ and a K_m of $4.0 \pm 1.0 \mu\text{M}$ when the reaction was carried out in the presence of 25 mM MgCl_2 at pH 8.5 and 37°C and almost no detectable activity ($k_{\text{obs}} < 5 \times 10^{-4} \text{ min}^{-1}$) when the reaction was carried out under the conditions necessary for continuous evolution. The continuous evolution mixture contains 8.8 mM NTPs and dNTPs, which reduce the concentration of free Mg^{2+} to 16.2 mM, accounting for the reduction in activity under these conditions. An improvement in catalytic rate of >10,000-fold under continuous evolution conditions would be needed for the modified DSL ribozyme to have a rate comparable to that of the class I ligase.

A starting pool of $\approx 10^{14}$ variants was prepared based on partial randomization of the modified form of the DSL ligase ribozyme. Random mutations were introduced at a frequency of 9% per nucleotide position at all positions except the tetraloop, tetraloop receptor, and 3' primer binding site (Fig. 2A). This pool was subjected to eight rounds of stepwise evolution, in which the molecules were required to react in progressively shorter times,

eventually reaching 30 s. The ligated RNAs, representing catalytically competent molecules, were isolated by gel electrophoresis and then amplified to generate progeny RNAs to begin the next round of evolution. During the first three rounds, the reaction mixture contained only ribozymes, substrate, 25 mM MgCl_2 , and 50 mM

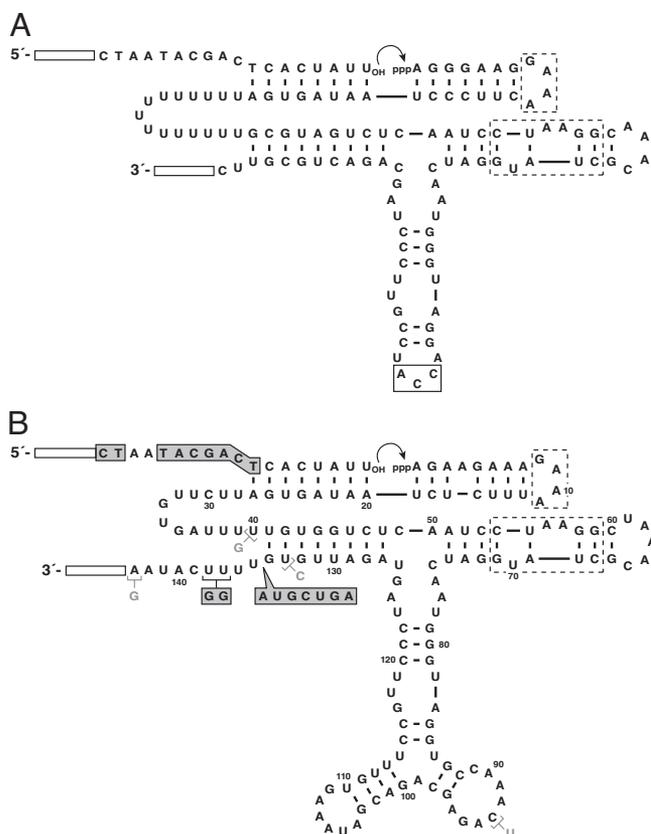


Fig. 2. Sequence and secondary structure of the starting and evolved DSL ribozymes. Open rectangles indicate primer binding sites, which are immutable during the course of evolution. Dashed boxes indicate the tetraloop and tetraloop receptor. (A) Modified DSL ligase that was used as the basis for constructing the starting pool of ribozymes. Boxed residues were replaced by 35 random-sequence nucleotides. (B) T80-1 ligase, which was obtained after transfer 80 of continuous evolution. Ribozyme residues in shaded boxes indicate changes present in all clones obtained after transfer 100. These changes confer complementarity to substrate residues shown in shaded boxes. Additional mutations present in the T100-1 ligase are shown in gray.

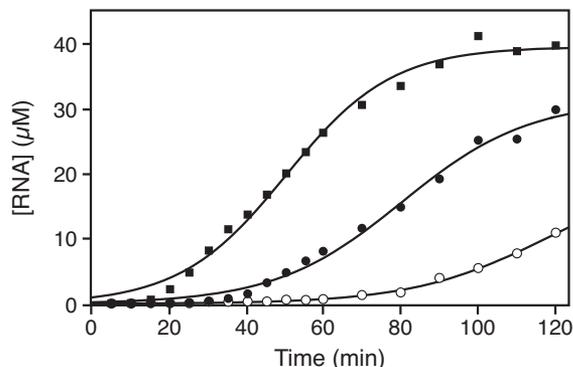


Fig. 5. Amplification profiles of individual ribozymes in the continuous evolution mixture. Filled and open circles, T80-1 ligase using either the full-length or truncated substrate, respectively; filled square, T100-1 ligase using the truncated substrate. The concentration of RNA was determined at various times, and the data were fit to the logistic equation: $[RNA] = a/(1 + be^{-ct})$, where a is the maximum extent of amplification and c is the exponential growth rate. Curvilinear regression coefficients were in the range of 0.996–0.999.

35-nucleotide accessory domain, including one highly conserved point mutation, compared with the consensus type I sequence identified after transfer 15. In addition, two clones isolated after transfer 50 exhibited a seven- or eight-nucleotide deletion near the 5' end of the accessory domain. This deletion became fixed within the population by transfer 80. Two clones contained a point mutation at the 3' end of the accessory domain that also became fixed after transfer 80.

There was little sequence variation among 10 clones that were isolated after the 80th transfer (SI Tables 1 and 2). A typical clone, designated as T80-1, whose sequence conforms to the consensus sequence, was chosen for further study. The amplification profile of the T80-1 ribozyme is consistent with a logistic growth curve, indicative of exponential growth subject to a limiting constraint (Fig. 5). Here the constraint is the finite amount of substrate and other resources present in the continuous evolution mixture.

Kinetic and Secondary Structure Analyses. The catalytic activity of the T80-1 ligase ribozyme was examined under continuous evolution conditions (excluding the protein enzymes), exhibiting a k_{cat} of $1.39 \pm 0.06 \text{ min}^{-1}$ and a K_m of $0.14 \pm 0.04 \mu\text{M}$ (SI Fig. 7). This result corresponds to a catalytic efficiency, k_{cat}/K_m , of $1 \times 10^7 \text{ M}^{-1}\text{min}^{-1}$, which is a 2,000-fold improvement compared with the behavior of the R8-11 ligase and a >80,000-fold improvement compared with the starting DSL ligase ribozyme. The parent *cis*-DSL-1S ribozyme has a catalytic efficiency of $7 \times 10^4 \text{ M}^{-1}\text{min}^{-1}$ under its preferred reaction conditions (19).

To verify that the continuously evolving ribozyme retained the ability to form a 3',5'-phosphodiester, the ligation junction was probed by using RNase T2, which cleaves 3',5'- but not 2',5'-phosphodiester linkages of RNA and does not cleave DNA. An unlabeled ribozyme was allowed to react with a 5'-³²P-labeled substrate that contained a single ribonucleotide at its 3' end. The ligated products were isolated by polyacrylamide gel electrophoresis and then digested with RNase T2 (SI Fig. 8). Because the substrate only contained a single 3'-terminal ribonucleotide, the labeled product of RNase T2 digestion would be the same length as the product of complete alkaline digestion if a 3',5'-phosphodiester had formed, which was indeed the case.

Enzymatic and chemical probing studies were carried out to analyze the secondary structure of the T80-1 ligase, especially the newly evolved accessory domain. 5'-³²P-labeled ligated products were partially digested with either RNase A or RNase T1 under either native conditions (mimicking the conditions of continuous

evolution) or denaturing conditions (SI Fig. 9A). Unlabeled ligated products were subjected to chemical modification, employing either dimethyl sulfate or kethoxal under either native or denaturing conditions, followed by reverse transcription using a 5'-³²P-labeled DNA primer (SI Fig. 9B). These two methods were used in combination to confirm that the overall secondary structure of the DSL ribozyme had not changed during the 80 transfers of continuous evolution. The data also were used to establish a secondary structural model for the newly evolved accessory domain, which appears to form two short stem-loop elements that join to the remainder of the ribozyme by a three-way junction (Fig. 2B).

Substrate Dependency of the Evolved Ribozyme. The T80-1 ligase was challenged to ligate a 5'-truncated substrate that contains the entire RNA polymerase promoter sequence, but lacks the upstream primer binding site that was used for PCR amplification and cloning. The ribozyme reacted poorly with this truncated substrate, with a k_{obs} of only $0.032 \pm 0.001 \text{ min}^{-1}$. This finding corresponds to a rate reduction of ≈ 50 -fold compared with the reaction with the full-length substrate. The T80-1 ribozyme was tested for its ability to undergo exponential amplification by using the truncated substrate and demonstrated much slower exponential growth compared with its behavior with the full-length substrate (Fig. 5).

The population of ribozymes obtained after transfer 80 was subjected to mutagenic PCR and then used to initiate 20 additional transfers of continuous evolution employing the truncated substrate (Fig. 4B). The evolving population quickly regained the ability to undergo rapid exponential growth. After transfer 100, individuals were cloned from the population and sequenced. There was little sequence variation among 10 clones that were analyzed (SI Tables 1 and 2), and all contained a seven-nucleotide insertion (5'-AGUCGUA-3') and two U \rightarrow G changes located near the 3' end of the ribozyme (Fig. 2B). These mutations confer complementarity between this region of the ribozyme and 11 nucleotides at the 5' end of the truncated substrate.

One of the cloned individuals, designated as T100-1, whose sequence conforms to the consensus of all sequenced clones, was found to undergo rapid exponential growth in the presence of the truncated substrate, even exceeding that of the T80-1 ribozyme with the full-length substrate (Fig. 5). The k_{obs} of the T100-1 ribozyme with the truncated substrate is $2.9 \pm 0.3 \text{ min}^{-1}$, which is about twice as fast as the rate of the T80-1 ribozyme with the full-length substrate.

Mutations were introduced within the seven-nucleotide insertion of the T100-1 ribozyme, replacing the sequence 5'-AGUCGUA-3' by 5'-ACAGCAA-3' (mutated residues in italics). The mutant ribozyme was unable to ligate the truncated substrate (SI Fig. 10). Full ligation activity was restored, however, by employing a different substrate that contained five compensatory mutations. This result indicates that the T100-1 ribozyme binds the truncated substrate through a newly evolved Watson-Crick pairing interaction that augments the preexisting substrate-binding domain (Fig. 2B). To exclude the possibility that the two pairings engage two different substrate molecules, the mutant ribozyme was tested with an equimolar mixture of the standard truncated substrate and an all-DNA version of the substrate bearing the compensatory mutations. The latter cannot undergo ligation, but in principle could satisfy the requirement for a second pairing interaction. There was no detectable activity under these conditions, consistent with the hypothesis that the same substrate molecule is involved in both pairing interactions.

Discussion

The Need for Speed. Continuous evolution is the most powerful method currently available for the *in vitro* evolution of catalytic function. However, since the development of this method a decade ago (14), its application has been limited to variants of the class I ligase ribozyme. With the present study, the DSL ribozyme be-

comes the second continuously evolving functional macromolecule. It is desirable to have multiple classes of ribozymes that are capable of undergoing continuous evolution, not only to facilitate the evolution of each catalyst individually, but also to enable more complex experiments in which two or more molecular species evolve within a common environment.

The most significant obstacle to achieving the continuous evolution of a ribozyme is the need to attain a sufficiently fast reaction rate. The ribozyme must catalyze ligation of the promoter-containing substrate before becoming inactivated by reverse transcription. The minimum catalytic rate needed depends on the sequence and structure of the ribozyme and various aspects of the reaction environment. For the class I and DSL ribozymes under the conditions of continuous evolution, the k_{obs} for reverse transcription was determined to be $4.1 \pm 0.3 \text{ min}^{-1}$ and $2.4 \pm 0.2 \text{ min}^{-1}$, respectively. Thus, it appears that a catalytic rate for the ribozyme of $\approx 1 \text{ min}^{-1}$, in addition to other context-dependent adaptive features, is required for it to cross the threshold to continuous evolution. The class I ligase was preadapted for continuous evolution because its catalytic rate was already $>10 \text{ min}^{-1}$ under the reaction conditions for which it was originally optimized (4, 11). Yet many rounds of stepwise evolution were required to adapt the ribozyme to the promoter-containing substrate and to the reaction conditions necessary for continuous evolution (14). In contrast, the slow-reacting DSL ligase required substantial improvement in its catalytic rate. This goal was achieved by using a quench-flow apparatus to select molecules that could react in times as short as 15 ms. This technique could be applied more broadly, apart from the goal of continuous evolution, to obtain other fast-reacting catalysts.

A critical step in the development of the continuously evolving DSL ribozyme was the addition and evolutionary maturation of the 35-nucleotide accessory domain, which greatly augments the behavior of the catalytic core. The added nucleotides provided an opportunity to evolve peripheral elements that result in enhanced catalytic performance. A similar situation exists for various biological ribozymes, where the catalytic core alone provides a modest level of activity and peripheral elements bring about a substantial increase in the catalytic rate. The hammerhead and *Neurospora* Varkud satellite ribozymes, for example, both catalyze a self-cleavage reaction that yields products bearing a 2',3'-cyclic phosphate and 5'-hydroxyl. Minimized forms of these ribozymes that contain only the phylogenetically conserved residues necessary for catalysis have markedly reduced activity, compared with the full-length motifs that occur in nature (21, 22). In both cases, the additional peripheral elements allow formation of tertiary contacts that stabilize the active conformation of the ribozyme and enhance its catalytic rate.

Adaptation to Substrate Features. The substrate molecule that is used in continuous evolution contains the sequence of an RNA polymerase promoter. For convenience, additional nucleotides often are placed upstream of the promoter sequence to serve as a primer binding site for PCR amplification and subsequent cloning. In the present study, 18 nucleotides were placed upstream of the promoter, having a sequence that was irrelevant to the substrate-recognition properties of the ribozyme. During 80 transfers of continuous evolution, however, the ribozyme became dependent on the added nucleotides. The activity of the T80-1 ribozyme is reduced ≈ 50 -fold when provided a substrate that lacks the added nucleotides, and its rate of exponential amplification in the continuous evolution system decreases accordingly (Fig. 5). Thus, what began as an irrelevant appendage was coopted by evolution to confer enhanced selective advantage to the ribozyme.

It is not uncommon in biological evolution for traits to arise that initially have no selective advantage, but subsequently become coopted in a way that enhances fitness (23). Evolutionary biologists are challenged to distinguish between a primary adaptation and a

neutral change that subsequently was "exapted" (24) for selective advantage. Laboratory evolution, in contrast, generates a complete historical record in the form of the molecules that are archived at each step of the evolution process. One can return to any point along an evolutionary pathway and determine which molecular traits are present and whether they provide a selective advantage.

The added substrate nucleotides, which initially had no effect on catalysis, led to an exaptation, seen in the strong dependence of the activity of the T80-1 ribozyme on this feature of the substrate. When the basis for this exaptation was removed, the ribozymes exhibited a substantial decrease in fitness. After 20 additional transfers of continuous evolution, however, the population adapted in a new way. The ribozymes obtained after the 100th transfer had acquired a seven-nucleotide insertion and a two-nucleotide substitution that improved their ability to bind the truncated substrate (Fig. 2B). These changes enhanced the catalytic activity of the ribozymes and their corresponding rate of amplification in the continuous evolution mixture (Fig. 5), an example of how continuous evolution can be used as a tool to address questions pertaining to the biochemical basis for evolutionary adaptation. By developing a second ribozyme that is capable of continuous evolution, the scope of these experiments can be broadened considerably.

Materials and Methods

Stepwise Evolution. DNA templates for the starting pool were generated by a 10-cycle PCR employing two synthetic oligonucleotides having the sequence 5'-GAAGTAATACGACTCACTAT-TAGGGAAGGAACTTCCCTAATAGTGATTTTTTTTTT-TTTTTTGCCTAGTCTCAATCCTAAGGCAAACGCTATG-3' and 5'-CGTGACCGAGCCACTAGGAACGCAGTCTGCTAG-GGAACGGATGGTCTACCTATTGATCCATAGCCGTTTG-CCTTAG-3' (T7 RNA polymerase promoter sequence underlined; nucleotides randomized at 9% degeneracy in italics). The PCR products, consisting of $\approx 10^{14}$ different sequences, were used as templates for *in vitro* transcription. The resulting RNAs were purified by denaturing PAGE. A starting population of ≈ 20 copies each of 10^{14} different RNA molecules (3.3 nmol) were challenged to ligate a chimeric DNA-RNA substrate having the sequence 5'-CTTGACGTCAGCCTGGACTAATACGACTCACUAUU-3' (T7 promoter sequence underlined; RNA residues in bold). During rounds 1–3 of *in vitro* evolution, the ligation reaction was carried out in the presence of 1 μM ribozyme, 2.5 μM substrate, 25 mM MgCl_2 , and 50 mM EPPS (pH 8.5) at 37°C. During rounds 4–8, the reaction mixture also contained 2.5 μM cDNA primer having the sequence 5'-CGTGACCGAGCCACTAG-3', 2 mM each NTP, 0.2 mM each dNTP, 50 mM KCl, 2 mM spermidine, and 4 mM DTT. The ligated products were purified by PAGE, eluted from the gel, reverse transcribed, PCR amplified, and then forward transcribed to produce progeny RNAs, which again were gel purified.

DNA templates for the second pool (after round 8) were generated by a single-cycle extension reaction employing two oligonucleotides having the sequence 5'-GAAGTAATACGACTCACTATTAGAGAAGGAACTTCCCTAATAGTGATTT-TTGTTGTTTTTGCCTAGTCTCAATCCTAAGGCAAACGCTATGG-3' and 5'-GGATGGCACGGAGTCAGGGAG-CGAGTCTGCTAGGGAACGGA-N₃₅-TCCTACCCATTGATC-CATAGCGTTTTGCCTTA-3' (T7 promoter sequence underlined; nucleotides randomized at 9% degeneracy in italics; random nucleotides denoted as N). The resulting double-stranded DNAs, consisting of $\approx 10^{14}$ different sequences, were transcribed to produce RNAs, which were gel purified. A starting population of ≈ 20 copies each of 10^{14} different RNA molecules (3.0 nmol) were challenged to ligate a substrate having the sequence 5'-CATCGTGCCTTGCTGCTCAATACGACTCACUAUU-3'. The reaction conditions for rounds 9–16 were as in rounds 4–8, except that a different cDNA primer was used, having the sequence 5'-GGATGGCACGGAGTCAG-3'.

Rounds 17–21 were carried out by using a KinTek model RQF-3

