Emergence of a dual-catalytic RNA with metal-specific cleavage and ligase activities: The spandrels of RNA evolution

(2′,5′-link/prebiotic/manganese/RNA editing/ribozyme)

LAURA F. LANDWEBER* AND IRINA D. POKROVSKAYA

Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544

Communicated by Thomas E. Shenk, Princeton University, Princeton, NJ, October 26, 1998 (received for review August 7, 1998)

ABSTRACT In vitro selection, or directed molecular evolution, allows the isolation and amplification of rare sequences that satisfy a functional-selection criterion. This technique can be used to isolate novel ribozymes (RNA enzymes) from large pools of random sequences. We used in vitro evolution to select a ribozyme that catalyzes a novel template-directed RNA ligation that requires surprisingly few nucleotides for catalytic activity. With the exception of two nucleotides, most of the ribozyme contributes to a template, suggesting that it is a general prebiotic ligase. More surprisingly, the catalytic core built from randomized sequences actually contains a 7-nucleotide manganese-dependent self-cleavage motif originally discovered in the Tetrahymena group I intron. Further experiments revealed that we have selected a dual-catalytic RNA from random sequences: the RNA promotes both cleavage at one site and ligation at another site, suggesting two conformations surrounding at least one divalent metal ion-binding site. Together, these results imply that similar catalytic RNA motifs can arise under fairly simple conditions suggesting two conformations surrounding at least one divalent metal ion-binding site. Together, these results imply that similar catalytic RNA motifs can arise under fairly simple conditions and that multiple catalytic structures, including bifunctional ligases, can evolve from very small preexisting parts. By breaking apart and joining different RNA strands, such ribozymes could have led to the production of longer and more complex RNA polymers in prebiotic evolution.

The in vitro selection of functional RNA molecules from random sequence pools has successfully isolated far more classes of ribozymes than have yet been found in nature. The classic experiment (1) revealed an abundance of RNA ligase ribozymes in a random sampling of $10^{15}$ unique sequences. The first experiments to probe in vitro evolution of ribozymes searched for improved or altered versions of the naturally occurring Tetrahymena group I intron. For example, Green et al. (2) used in vitro evolution to restore and to improve activity of a pool of molecules based on the Tetrahymena group I ribozyme, whereas Joyce and colleagues subsequently isolated variants of the group I ribozyme that performed subtly different tasks, such as cleavage of a DNA substrate (3) or acquisition of a novel dependency on a specific divalent metal cation (4). The ability of directed evolution and ribozyme engineering to evoke secondary activities of group I introns (5) suggests that RNA has an intrinsic capacity to adapt to novel contexts or substrates, although in these cases the acquisition of new catalytic activity was aided by careful design of the substrate to mimic the natural phosphodiester bond.

In this paper we report the in vitro selection and evolution of a very small ribozyme that catalyzes the ligation of two substrate RNAs but also possesses the unexpected ability to undergo a separate self-cleavage reaction. Substitution of a divalent metal ion in the absence of further sequence evolution triggers the switch from selected ligase to unselected self-cleaving RNA. The facile construction of a minimal RNA ligase ribozyme from random sequences and the simple conversion between two RNA metalloenzymes (9) by a single metal switch imply a surprising plasticity of RNA evolution.

METHODS

Nucleic Acids. The initial DNA pool was prepared as a 132-base oligonucleotide (16 nucleotides of constant region flanking 100 bases of random sequence) on a Millipore Expeditie synthesizer, confirmed for even degeneracy by cycle sequencing, gel purified, and amplified by using large-scale PCR (1) with the 5′ and 3′ primers to extend the constant regions (5′ PCR primer including T7 promoter 5′-TCTAATAGCTCTCATATAGGCTCACACTTGTATAGTCTACT-3′; 3′ constant region PCR primer: 5′-(CUA)$_2$GGGATCCTAATGAGC-3′). Pool RNA for each round was gel-purified and eluted by a crush-and-soak method after T7 transcription [with T7 RNA polymerase either purchased from Epicentre Technologies (Madison, WI) or prepared as described in ref. 11] from phenol-extracted, ethanol-precipitated PCR templates. Double-stranded amplified DNA was also gel-purified after the first round of amplification. Approximately 20 nmol (1.5 mg) of pool RNA was used in the first round, 0.1 nmol was used in the second round, and 20 pmol was used in subsequent rounds.

One of three substrate RNAs was used in each alternating selection round (RNA1, 5′-biotin-(cat)$_4$-guGUAACAAGUGUAGAAAUCAGUCUUUU-3′; RNA2, 5′-biotin-(cat)$_4$-gtauguaUUUAACAGAAAGAUACUCUUUUUU-3′; RNA3, 5′-ggga(cau)agguCAAGAAAGAUACUCUUUUU-3′; tag sequences are lowercase. RNA1 and RNA2 were constructed on a Millipore Expedite synthesizer [with DNA sequence tags italicized and biotin-dT (Glen Research, Sterling, VA) at the 5′ end] and purified by anion exchange HPLC followed by gel purification. RNA3, used in largest quantity in the first round (1 mg, 90 nmol) was prepared by T7 transcription (11, 12) and gel purified. Other RNA substrates and small ribozymes were either prepared by T7 transcription of oligonucleotide or PCR templates or purchased from Yale University (Keck Oligonucleotide Facility).

In Vitro Selection Strategy. The selection protocol we followed was similar to the one described in refs. 1 and 10. In the first seven rounds, 1 µM pool RNA and 4-fold excess substrate were incubated overnight at high concentrations of K$^+$ and Mg$^{2+}$ (400 mM KCl/60 mM MgCl$_2$/30 mM Tris, pH 7.4) at 21–25°C to facilitate binding and catalysis with the substrate usually tethered to a solid support to prevent loss of pool RNA caused by precipitation at elevated Mg$^{2+}$ concentrations (1). Mg$^{2+}$ was always added last. Round 1 was performed with the substrate RNA3 tethered to streptavidin agarose via a 5-fold excess of a biotinylated DNA oligonucleotide complementary to the sequence tag, and rounds 2, 4, and 5 were performed with biotinylated substrate directly attached to magnetic beads. Use of a...
solid support allowed specific retention of covalently ligated ribozymes, and selective PCR with the tag sequence as primer amplified only ligated sequences. Each round achieved >1000-fold enrichment of reacted RNA molecules. We introduced different guide sequences as well as DNA sequence tags on the substrate RNA in alternate rounds of selection to select for a general ligase (instead of one with restricted sequence requirements) and to increase the efficiency of the selection by preventing sequence dependency on the substrate.

In round 1, pool RNA and substrate RNA3 were annealed and reacted on a 7.5-ml streptavidin agarose column (Pierce), and unligated pool RNA was released from the solid support in 25–100 mM NaCl, followed by elution of any reacted (ligated) RNA in 0–5 mM NaCl. Radiola beled RNA3 was used to monitor the presence of fractions containing covalently ligated RNA. These were pooled and ethanol-precipitated with glycerol as a carrier. Round 2 was performed similarly by using biotinylated RNA1 and 0.25 ml of M280 Streptavidin (Dynal, Great Neck, NY) magnetic beads. RNA collected from round 1 was affinity-purified on a second column (2.5 ml of streptavidin agarose) containing 2 nmol of a 20-base oligonucleotide complementary to the 3′ constant region to remove excess substrate (oligo 20.99 in ref. 1). Retained RNA was reverse-transcribed (SuperScript RNase H-). Nonspecific RNA binding to the streptavidin support lose their affinity after reverse transcription and conversion to double-stranded molecules. These RNA–DNA hybrids were washed off the solid support while specifically captured onto streptavidin-coated magnetic beads after reverse transcription when biotinylated substrates were used.

In the first five rounds, 10–15 cycles of selective PCR (1) with the tag sequence as primer were followed by up to 35 cycles of nested PCR until a product was strongly visible on an ethidium bromide-stained agarose gel. Selective PCR (25–35 cycles) was used in later rounds when a product was readily visible, with only five cycles of nested PCR used to introduce the T7 promoter. Nonspecific RNA binding to the streptavidin-coated beads (Pierce, Dynal) was minimal (<0.01%) in most rounds. Progressively shorter incubation times, lower concentrations of Mg2+, and mutagenic PCR were introduced in later rounds to allow refinement of the catalytic pool.

Beginning with round 6, biotinylated RNA:DNA hybrids were captured onto streptavidin-coated magnetic beads after reverse transcription (because rogue RNA molecules that bind the streptavidin support lose their affinity after reverse transcription and conversion to double-stranded molecules). These RNA–DNA hybrids were washed off the solid support while specifically retaining 5′-biotinylated duplexes in TE (10 mM Tris-HCl/1 mM EDTA) buffer. After DNA restriction-fragment and sequence analyses revealed the presence of a single class of closely related molecules in round 7, the addition of mutagenic PCR (7 mM MgCl2/0.2 mM each purine nucleotide/1 mM each pyrimidine nucleotide/0.5 mM MnCl2; see refs. 13 and 14) for mixtures of approximately 30, 60, and 90 doublings (1) with 200,000-fold dilutions between double sets of PCR introduced more variation. In each subsequent round, 30 cycles of nested PCR were performed with mutagenic conditions on 200,000-fold dilutions of DNA recovered after selective PCR. The MgCl2 concentration during ligation was reduced to 10 mM in rounds 8–11 and 5 mM in rounds 12–15. Incubation time was reduced from overnight to 3 hr in round 10, to 30 min in round 11, to 15 min in round 12, and to 2 min (after preincubation of RNA with substrate in the absence of MgCl2) in rounds 13 and 14. All reactions were quenched by adding EDTA. Selective PCR products were cloned into a UDG cloning vector (BRL).

Minimal Base-Pairing Requirements. One round of selection was performed to determine the pairing requirements for ligation by ligating overnight 1 μM randomized hairpin RNA (Fig. 1d) to 3-fold excess RNA substrate containing the 10-nt sequence 5′-pppGACUCACACU fused to the 22-nt sequence tag from RNA2 in a 0.9-ml reaction volume (30 mM Tris-HCl, pH 8.5/0.4 M KCl/100 mM MgCl2). Product ligated to the 10-nt substrate (∼12%) was visualized by autoradiography and isolated on a 12% polyacrylamide gel containing 8.3 M urea, reverse transcribed, and amplified by rTth DNA polymerase (Perkin-Elmer) by using the 3′ sequence tag and a 5′-PCR primer containing the T7 promoter fused to the first eight nucleotides of the RNA 5′ end (GGUGUGGU).

RNA Labeling. RNA molecules were either 32P-labeled at their 3′ termini with 3′-α-[32P]ATP (NEN) and poly(A) polymerase (BRL) or 5′-labeled either with T4 polynucleotide kinase (New England Biolabs) and [γ-32P]ATP or [γ-33P]ATP (NEN) or by T7 transcription in the presence of [α-35P]GTP and unlabeled ATP, CTP, and UTP (for RNAse T2 digestion). T7 transcripts were treated with phosphatase (calf intestinal alkaline phosphatase, Boehringer Mannheim), phenol extracted, and precipitated before labeling with T4 kinase.

Ligation Reactions. Reactions were performed in 5-μl volumes (20 μl for time course reactions) with 1 μM template ribozyme, 1 μM 3′ substrate (for trans-ligations), and trace amounts of 5′-end-labeled substrate in 400 mM KCl, 30 mM buffer, 0.4 units/μl RNasin (Promega) at the indicated pH and with the indicated divalent cation. Tris buffer was used for pH 7.5–9.0 and 2-(N-cyclohexylamino)ethanesulfonic acid (Ches) buffer was used for pH 8.5–10.0 (both were obtained from Sigma). Metal salts [MgCl2 (99.995%), MnCl2 (99.99%), CaCl2 (99.99%+), SrCl2 (99.995%), BaCl2 (99.999%), and CdCl2 (99.99%)] were purchased from Aldrich and of the highest purity available. Solutions were prepared in diethyl pyrocarbonate-treated H2O (Research Genetics, Huntsville, AL) and filtered through 0.22-μm Millipore filters before use. To avoid oxidation, 1 M stock solutions were stored at −20°C and mixed with buffer before use. All reactions were incubated at 22–25°C for between 0 hours and 7 days. Reactions were stopped by addition of an equal volume of EDTA in formamide loading buffer (95% deionized formamide/10 mM EDTA/0.02% xylene cyanol and bromophenol blue).

RNAse T2 Digestion. RNAse T2 reactions were carried to completion by incubation for 30 min at 37°C in 5 μl containing 50 mM sodium acetate (pH 4.5), 2 mM EDTA, and 10–20 units of RNAse T2 (BRL). RNA for this experiment was labeled at the phosphate of the ligation junction by ligating a 10-base substrate RNA containing a 32P label at the α-phosphate of the 5′-triphosphate to the 22-nt RNA (Fig. 1c) and gel-purifying the ligated product. Digestion products were analyzed on a 20% acrylamide, 8.3 M urea gel.

Manganese-Dependent, Site-Specific Cleavage. RNA5 (15) was a gift from S. Kazakov (Somagenizo, Palo Alto, CA). Reaction volumes (10 μl) containing 10 mM MnCl2, 100 mM NaCl, and 50 mM Trisacetate (pH 7.5) were incubated at 45°C for 1 hr. Reactions were stopped by addition of an equal volume of formamide loading buffer, incubated for 15 min at 37°C (to avoid nonspecific hydrolysis of RNA), and loaded directly onto a 20% acrylamide, 8.3 M urea gel.

RESULTS AND DISCUSSION

Selection for a General Ligase. Inspired by the biological process of RNA editing, which uses multiple cycles of cleavage, U addition, and ligation to rewrite up to 90% of the coding-sequence information in kinetoplastid mitochondrial transcripts (16), we designed a substrate RNA and a pool 5′ constant region that contained a truncated NADH dehydrogenase subunit 7 (ND7) guide RNA from Leishmania tarentolae and 18 nucleotides of the 5′ editing domain in the ND7 mRNA, respectively (17). A side activity of RNA editing that has been observed both in vivo (18) and in vitro (17) is the covalent attachment of the guide RNA to the mRNA molecule. We sought to test whether this activity is an intrinsic property of the RNA molecules themselves. The one unifying feature among all guide RNA molecules is the presence of a 3′ oligo(U) tail (18); hence, we initially constructed all 5′ substrate molecules with a short (5–6 nt) oligo(U) tail. From a pool of $1.6 \times 10^{12}$ unique sequences containing 100 nucleotides of random sequence flanked by constant regions, a single class of RNA molecules emerged after six cycles of selection for the ability to ligate each of three tagged substrate RNAs to their 5′
end. Covalent attachment to the substrate marks the pool RNA molecules with either biotin or a specific sequence tag (1), and the strategy of recovering biotin-tagged ribozymes as RNA-zcDNA hybrids avoided recovery of single-stranded RNAs that bind only streptavidin.

All ribozymes that survived the selection contained a sequence within the random region that can anneal to the 5'-9 constant region of the pool molecules as well as to multiple substrate RNAs, forming a long imperfect stem, which positions the 3'-terminus of the oligo(U) tail of the substrate RNA near the 5'-triphosphate of the pool RNA (Fig. 1a). The 5'-end ligation catalyzed by this RNA involves nucleophilic attack by the small substrate RNA on the 5'-triphosphate (1). We found that this ligation requires the presence of a 5'-triphosphate, or activation of the α-phosphate by pyrophosphate, because activity is lost with phosphatase-treated or kinase-treated RNA (not shown) that contains either a 5'-hydroxyl or 5'-monophosphate, respectively.

Because of the difficulty of sampling such a large sequence space [fewer than $10^{16}$ molecules of a possible $4^{100} > 10^{60}$], the ribozymes selected initially were expected to be suboptimal representatives of a class of catalytically active molecules (1). Seven more rounds of in vitro selection were performed by using heavy mutagenic PCR on the pool of ribozymes recovered after the seventh round of selection. This improved the reaction by reducing the required magnesium concentration and reaction time.

Alignment of 51 clones revealed a sequence consensus for this ribozyme with only a conserved 29-nt core region (Fig. 2). By assaying different clones, we verified that the extent of base pairing between ribozyme and substrate in the conserved region contributes to the rate of ligation. In addition, most sequence covariation allowing G-U pairs preserves pairing to the substrate in the core region.

Remarkably, this ligase ribozyme has the fewest sequence requirements of any selected ribozyme, with only two nucleotides.

**Fig. 1.** (a) Possible secondary structure of one of the ligase ribozymes selected from round 13, shown ligated to the substrate guide RNA (gRNA). The 5'- and 3'-constant regions of the pool are shaded (and sequences derived from the 5'-constant region are shown in blue throughout Fig. 1). The 3'-U6 tail (underlined and in green) and the 5'-biotin of the guide RNA substrate are marked. (CAT), a DNA sequence tag, and other sequences in red are not needed. The core region of the ribozyme is boxed in the center. The * arrow indicates the 3'-terminal base needed for efficient ligation. Dashed lines and arrows indicate sites of allowed breakage for trans-ligation or truncation experiments. U–A pairs in magenta are weakly supported by sequence covariation. Substitutions in black increase activity, and mutations in light gray abolish activity, following a grayscale. The small arrow indicates attack of the 2'-hydroxyl of the 3'-U6 tail on the 5'-triphosphate of the G at the end of the RNA. (b) Very small cis-ligation with the same substrate as in a, $k_{obs} = 3.4 \times 10^{-4}$ min$^{-1}$. (c) Very small trans-ligation with a 10-nt 3'-substrate and a 29-nt ribozyme (boxed), $k_{obs} = 5.4 \times 10^{-4}$ min$^{-1}$ at 25°C and $3 \times 10^{-4}$ min$^{-1}$ at 45°C. (d and e) Two very simple catalytic motifs still ligate with up to 10% (d) and 4% (e) the efficiency of the ligation in b or c, demonstrating remarkable versatility of the small ligase motif. Randomized nucleotides (N) are not intended to be base paired.

**Fig. 2.** Alignment of 12 representative functional ribozyme clones from a total of 51 such clones. A dash indicates identity with the top sequence. A . indicates a deletion. The most conserved region is boxed. The constant regions are shaded in the reference sequence.
directly implicated in a non-template role. The only known smaller ribozyme promotes cleavage rather than ligation (15) and is derived from an RNA hairpin found in the Tetrahymena group I intron (19). The ligation catalyzed by this ribozyme thus differs from the ligation catalyzed by other RNA ligase ribozymes (10) by requiring a G-G juxtaposition at the site of ligation and very few selected nucleotides for activity (Fig. 1).

We constructed a minimal version of this RNA-ligase ribozyme. Deletion of sequences 3′ of the boxed conserved region (Fig. 1a) led to a 2-fold increase in activity; however, deletion of some sequences 5′ of the boxed region produced a slight decrease in activity, indicating that some of these bases may contribute modestly to catalysis. The smaller ribozyme contains only substrate nucleotides and a central 27-nt core derived from the originally selected ribozyme. This represents the smallest such ribozyme isolated by in vitro selection, with only two potentially unpaired nucleotides required for catalysis (Fig. 1b).

**Ligation in Trans.** The ribozyme also functions efficiently in trans (Fig. 1c) but without turnover because of the slow dissociation of the enzyme-substrate complex (consisting of three annealed strands that together form the catalytic RNA system). The pseudo-first-order rate constant $k_{\text{obs}}$ for this reaction was $2.7 \times 10^{-4} \text{ min}^{-1}$ for the time course shown in Fig. 3. The 29-nt “ribozyme” is essentially as active as one containing the full 5′ end. To our surprise, we discovered that a general ligase ribozyme stripped of its template either upstream or downstream of the ligation site still ligates several sequences, albeit with approximately 4% the original efficiency (Fig. 1d and e). An alignment of 36 successful ligase clones derived from a single round of selection (as in Fig. 1d) exhibited no marked sequence consensus, suggesting that the catalytic structures are versatile. In fact, the number of Watson–Crick G-U base pairs found in the randomized region of these sequences was not significantly greater than expected for any pair of random hexamers ($\bar{n}_{\text{observed}} = 2-3 \text{ bp}$; $\bar{n} = 6(\%)$ was 2.25 expected at random).

**Creation of 2′-5′ Linkages.** The regioselectivity of bond formation is exclusively 2′-5′-phosphodiester linkages as assayed by resistance of 2′,5′-linkages to ribonuclease T2 (Fig. 4). This is consistent with the absence of extended base pairing at the ligation junction, as a Watson–Crick duplex favors 3′,5′-linkages (20) but considerably destabilizes 2′,5′-linkages (20–22). Although nonbiological in extant systems, 2′,5′ linkages are favored in nonenzymatic oligoribonucleotide synthesis because of the increased nucleophilicity of the 2′-hydroxyl toward activated phosphate esters (23). However, recent experiments demonstrate that both 2′,5′ and mixed linkages are active in template-directed oligomerization of mononucleotides (24) and oligonucleotides (25–26), and this has strengthened their possible role in primitive biochemistry.

**Fig. 4.** Ribonuclease T2 digestion to determine regiospecificity of bond formation. Lane 2 is the RNase T2 digestion product of the ligation in Fig. 1c (see Methods). Lane 1 is the RNase T2 digestion product of a ligated RNA of identical sequence, but with a labeled 3′,5′ linkage generated by T4 DNA ligase and a complementary DNA splint (42). Digestion products were analyzed on a 20% denaturing acrylamide gel. Lanes 3 and 4 are a time course of alkali hydrolysis at 5 and 10 min, respectively, of the ligated product digested in lane 1.

We examined the dependency of the ligation on time, Mg$^{2+}$ concentration, and pH (Fig. 5). The reaction exhibits approximately linear dependence on Mg$^{2+}$ concentration below 50 mM and a saturable Mg$^{2+}$ binding site with half-maximal activity above 50 mM over a pH range of 7.0 to 9.0 (data not shown; ref. 27). The logarithm of $k_{\text{obs}}$ increases linearly with a slope of approximately 1 as a function of pH, indicating a single critical deprotonation step—probably at the 2′-hydroxyl. The pH plateau above pH 9.0 is consistent with an attacking 2′-hydroxyl (Fig. 5). pH >9.5 leads to degradation, deprotonation, and helix disruption and so could not be validated). Deletion analysis mapped the 3′ and 5′ terminal positions required for activity to a small conserved core, which provides mostly a template. Only two nucleotides are directly implicated in a non-template role: the G opposite the ligation site and a single bulged pyrimidine (Fig. 1).

**Mutational Analysis and Rate Enhancement.** Many single point mutations (summarized in Fig. 1a) abolish activity below our threshold of detection. For example, deletion of the single bulged U or introduction of any purine base opposite (and potentially paired with) this U results in complete loss of activity, as does mutation of the G opposite the cleavage site (Fig. 1). However, recent experiments demonstrate that both 2′,5′ and mixed linkages are active in template-directed oligomerization of mononucleotides (24) and oligonucleotides (25–26), and this has strengthened their possible role in primitive biochemistry.
ground reaction containing the C-G pair may be hampered by the preferential hydrolysis of 2′,5′-linkages in a paired duplex (22); however, we estimated the same upper boundary for the single mutant lacking only the bulged U, which preserves the rest of the ligation site. This reaction also is 3,000-fold faster than a template-dependent uncatalyzed ligation with a similar 3′-terminal U-A base pair ($k_{obs}$ $8.5 \times 10^{-5}$ min$^{-1}$; ref. 20).

**Ligation of Hexaauridylate.** Because the guide RNA substrate in the original experiments contains an oligo(U) sequence at its 3′ end, we tested a 6-nt substrate, 5′,3′-(UUUUUU). Even this small RNA is a substrate for both cis- and trans-ligation (Fig. 6), albeit with 0.5% the efficiency of the 22-nt substrate shown in Fig. 1 c and e. As this sequence is small enough to be synthesized by mineral-catalyzed RNA condensation (28), this result suggests the possibility that we have uncovered an RNA ligation reaction with general prebiotic utility.

**Manganese Dependency.** Surprisingly, the catalytic core built from random sequences actually contains the very small manganese-dependent cleavage motif formed during autocatalysis of the Tetrahymena group I intron (refs. 15 and 19; Fig. 7). The trinucleotide UUU in a complex with GAAA (the smallest known catalytic RNA system) acts as a template catalyst (15, 29) to promote cleavage between the G and the A (15).

Further experiments revealed that we had in fact selected a dual-catalytic RNA from random sequences. This RNA promotes both cleavage at one site (Fig. 7) and ligation at another (Fig. 1), with possibly two conformations surrounding a divalent-metal ion binding pocket (or pockets). We therefore tested the metal-ion dependence of the original ligation: at pH 7.0, the ligation proceeds $10$-fold faster with Mn$^{2+}$ than with Mg$^{2+}$ (Fig. 5), but at higher pH the required concentration of Mn$^{2+}$ leads to RNA degradation.

We subsequently confirmed that the selected ligase also undergoes self-cleavage in the presence of Mn$^{2+}$ (Fig. 7) with $k_{obs} = 0.01$ min$^{-1}$ for self-cleavage of RNAs (15), compared with 0.0036 min$^{-1}$ for cleavage of GAAA (23). $k_{obs}$ varied from 5–8 $\times 10^{-4}$ min$^{-1}$ for cleavage of the 29-nt core ribozyme (template strand in Fig. 1c) in the presence of either 5–50 ng/µl poly(U), 0.5 mM 5′,3′-UUU, or 0.5 µM 22-nt substrate RNA shown in Figs. 1 c and e as template catalyst. In addition, the ligation shown in Fig. 1c proceeds remarkably well in optimal cleavage buffer conditions, with $k_{obs}$ $= 0.003$ min$^{-1}$ at 45°C.

**Other Divalent Metal Ions.** Cd$^{2+}$ also supports G ↓AAA cleavage (15) but supports ligation only weakly. Ca$^{2+}$, Sr$^{2+}$, and Ba$^{2+}$ also catalyze the ligation. Overall the ligation rates (Mn$^{2+}$ > Mg$^{2+}$ > Ca$^{2+}$ > Sr$^{2+}$ > Ba$^{2+}$) vary inversely with the $pK_a$ of their bound water molecules and follow the same order as hammerhead cleavage and nonenzymatic ligation (20), implicating a metal hydroxide or directly coordinated metal ion in the catalytic mechanism.

---

**The Roles of RNA and the Metal.** The simplicity of the ligation depicted in Fig. 1 d and e suggests that this motif can be generalized even further. The GNRA tetraloops at the ends may act as terminal clamps, thus reducing the need for base pairing to juxtapose substrates along the ribozyme template. The role of divalent metal ions is particularly relevant, as they are generally thought to have been important prebiotic catalysts (31). Evoking the notion of a “Cheshire catalyst” (32), in which an RNA scaffold mainly positions a catalytic metal ion (9), we suggest that the core nucleotides—most notably the bulged pyrimidine (Fig. 1)—serve two roles: (i) to bring together the 2′ and 5′ termini in a configuration that resembles the transition state for ligation (33) and (ii) to properly position the metal ion(s) in a favorable geometry, whereas the conserved flanking nucleotides (or tetraloop) provide intrinsic binding energy that leads to the closed state required for catalysis (34).

**A Case for RNA Preadaptation.** The self-cleaving GAAA/ UUU motif arose in this case as an unintended consequence of selection for ligation at an independent site. In other words, it appears as a “spandrel” (35–36), an offshoot or structural byproduct of the design of an RNA ligase ribozyme from random sequences. A spandrel has been defined as “any geometric configuration of space inevitably left over as a consequence of other architectural decisions” (36). Because the ability to cleave within G ↓AAA clearly did not emerge as an adaptation itself during the selection experiment—as it was neither selected for nor feasible in the magnesium-containing buffer—this additional activity presumably results from the
need to construct a binding site for one or more divalent metal ions. In vitro evolution therefore provides an unambiguous example of preadaptation, in the sense that the product of evolution was ultimately suited for a catalytic function other than the one for which it was selected. It could thus be co-opted to specialize in its new role, such as cleavage, which would make this feature an exaptation (37). In a separate selection experiment, Faulhammer and Famulok (38) unveiled a DNA enzyme that was preadapted for calcium binding. The pool molecules had never been exposed to calcium during the selection; however, the optimal solution for magnesium-dependent cleavage of RNA coincidentally produced a good solution to the problem of calcium-dependent cleavage as well, as though the geometry of binding the metal hydroxide anticipated the best fit for binding calcium.

Conclusions. The results presented here suggest that bifunctional catalytic RNA motifs can arise under fairly simple conditions and that multiple catalytic structures, including ligases, can evolve from basic preexisting building blocks. By extension, we predict that the combinatorial shuffling of catalytic modules (such as metal binding sites) into random sequence pools will greatly increase the density and representation of nucleic acid catalysts and that this will lead to a tremendous acceleration of in vitro evolution. Along similar lines, Burke and Willis (39) recently introduced recombination between RNA libraries to build aptamers (RNA ligands) that can bind two specific targets. Our finding that RNA structures as simple as a few unpaired nucleotides in a short helix may possess two catalytic activities suggests that similar metalloribozymes (i) arise surprisingly easily and would undoubtedly have been present in prebiotic evolution (as well as at high frequency in random pools) and (ii) could have contributed to the generation of RNA diversity by breaking apart and joining new sequence combinations.

Processes such as kinetoplastid RNA editing that may be primitive (40) require repeated cycles of both cleavage and ligation (41). These experiments demonstrate that a set of RNA molecules based initially on a Leishmania tarentolae guide RNA–mRNA combination are capable of performing both reactions. The full-length ND7 guide RNA is even a substrate for the ligase ribozyme \( k_{\text{obs}} = 1.7 \times 10^{-2} \text{ min}^{-1} \); however, \( k_{\text{obs}} \) increases by a factor of 2 by extending the “anchor” region of pairing between ribozyme strand and guide RNA. The ribozyme-catalyzed ligation employs a different leaving group, however, than does kinetoplastid editing. These observations do hint at another possible role of the poly(U) tails on kinetoplastid guide RNAs: they could assist in the cleavage at frequent G\(^1\)AAA editing sites. This reaction might have played a more significant role early in the history of RNA editing.

We thank Sergei Kazakov for advice, Anne Torjussen, Caroline Fichtenberg, Monika Pfunder, and David Kutzler for assistance, Rob Knight, Drew Ronneberg, and two anonymous reviewers for comments, and Jack Szostak and members of his laboratory for suggestions and RNA oligonucleotides. This research was supported by National Science Foundation Grants MCB-9520253 and MCB-9604377 and a Burroughs Wellcome Fund New Investigator Award in Molecular Parasitology to L.F.L.