

Enrichment for RNA molecules that bind a Diels–Alder transition state analog

KEVIN N. MORRIS*, THEODORE M. TARASOW†‡, CAROL M. JULIN*, SHAUNA L. SIMONS*, DONALD HILVERT†, AND LARRY GOLD*‡

*Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO 80309; †Department of Chemistry and Molecular Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, CA 92037; and ‡NeXagen, Inc., 2860 Wilderness Place, Suite 200, Boulder, CO 80301

Communicated by Richard A. Lerner, August 22, 1994 (received for review April 27, 1994)

ABSTRACT RNA molecules that bind a transition state analog for a Diels–Alder reaction ($K_d = 0.35 \pm 0.05$ mM) were isolated from a starting pool of $\approx 10^{14}$ sequences by affinity chromatography. After the initial rise and plateau of the amount of RNA that eluted with soluble analog, a step gradient elution was used to further enrich the pool for sequences with higher affinities for the target. To our knowledge, the isolation of RNA molecules that bind either a nonplanar or a hydrophobic ligand has not been reported previously. A conserved nucleotide sequence and secondary structure present in many of the RNA molecules are necessary but not sufficient for binding the analog. No catalysts of the targeted Diels–Alder reaction were found among the binders. The absence of catalysis contrasts with previous successful experiments with antibodies and suggests that other strategies may be needed to identify oligonucleotides with diverse catalytic activities.

Theories of evolution based on RNA as the progenitor molecule were fueled by the discovery of RNA molecules capable of catalyzing transesterification reactions (1, 2). Recently, a derivative of the *Tetrahymena* ribozyme was shown to hydrolyze an aminoacyl ester bond with a modest rate acceleration (3), and there is strong evidence that the peptidyltransferase activity of the ribosome is largely due to RNA (4). In addition, ribozymes with novel catalytic properties have been isolated from pools of mutagenized *Tetrahymena* self-splicing introns (5–7) and tRNAs (8, 9) and from pools of random RNA molecules (10). However, in an “RNA world” these primitive oligonucleotide catalysts must have been able to interact with various substrates and to accelerate a wide range of chemical transformations.

RNA molecules that bind a variety of proteins (11–16) and small molecules (17–21) have been isolated from large pools of random sequences through a reiterative selection and amplification process called SELEX (systematic evolution of ligands by exponential enrichment) (11). We are seeking to utilize this technology to develop a general scheme for surveying the catalytic repertoire of RNA and to examine the variety of ligands that can be specifically bound by nucleic acids. Antibodies that catalyze an extensive list of chemical reactions have been elicited by transition state analogs (22–24). Similarly, RNAs that tightly bind such analogs might also catalyze the corresponding reactions.

Here, we report the isolation of RNA molecules that bind a transition state analog for a Diels–Alder reaction (Fig. 1). This analog successfully elicited antibodies capable of catalyzing the cycloaddition of tetrachlorothiophene dioxide with *N*-ethylmaleimide (25). This target was chosen for three reasons: (i) the bimolecular Diels–Alder reaction is an excellent candidate for RNA catalysis since the condensation of

the substrates is strongly driven by proximity effects, and the large conformational changes that accompany the reaction can be exploited to minimize problems due to product inhibition; (ii) a catalytic antibody elicited by this analog is available for direct comparison of the catalytic efficiency of protein and any RNA isolated; and (iii) it tests the limits of RNA–ligand recognition by demanding that the nucleic acid bind a molecule that is nonplanar and hydrophobic.

MATERIALS AND METHODS

Target Synthesis. The *N*-hydroxysuccinimidyl ester of the transition state analog was prepared as described (25). Twelve milliliters of preswollen EAH Sepharose (Pharmacia) was washed with 1 liter of 0.5 M NaCl and then equilibrated with 1:1 dioxane/100 mM NaHCO₃, pH 8.0. The activated ester was dissolved in 1.5 ml of dioxane and added to the resin at room temperature. The resulting slurry was incubated at 5°C for 16 hr with occasional agitation. Remaining free amines were acylated by treatment of the resin with 1 ml of acetic anhydride at 5°C for 2 hr. The resin was then washed with 150 ml of 1:1 dioxane/100 mM NaHCO₃, pH 8.0, followed by 500 ml of phosphate-buffered saline (pH 7.5).

SELEX. The random RNA molecules for selection were transcribed *in vitro* from a double-stranded DNA template containing a random region 80 nucleotides long flanked by regions with fixed sequences as described (11). One nanomol of unlabeled RNA and 500,000 cpm of radioactively labeled RNA were denatured at 70°C for 5 min in 50 mM Tris·HCl, pH 7.5/0.5 M NaCl. The MgCl₂ concentration was adjusted to 5 mM and the sample was immediately incubated at 0°C for 10 min. The RNA solution was allowed to equilibrate to room temperature and applied to 0.1 ml of the derivatized EAH Sepharose, which had been prewashed with 1 ml of wash buffer (50 mM Tris·HCl, pH 7.5/0.5 M NaCl/5 mM MgCl₂). The RNA was incubated with the column matrix for 15 min before washing with 1 ml of wash buffer. Bound RNA was eluted with wash buffer containing various concentrations of soluble transition state analog (see *Results*). A cDNA copy of the eluted RNA molecules was synthesized with avian myeloblastosis virus reverse transcriptase and amplified by PCR using standard methods (11).

The binding affinities of selected RNA molecules were determined by analytical affinity chromatography (19) with 1 mM and 10 mM soluble analog.

Cloning and Nucleotide Sequence Determination. Individual molecules were isolated from selected pools by ligating *Bam*HI- and *Hind*III-digested PCR products into pUC18. The ligated DNA was introduced into *Escherichia coli* strain SURE (Stratagene) by electroporation, plasmids were isolated, and the nucleotide sequences in the inserted DNAs

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: SELEX, systematic evolution of ligands by exponential enrichment.

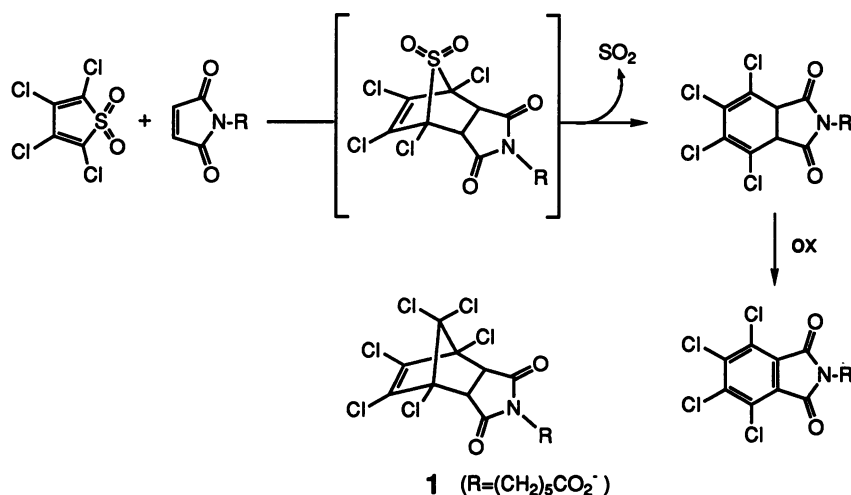


FIG. 1. The Diels-Alder reaction between tetrachlorothiophene dioxide (TCTD) and *N*-ethylmaleimide (NEM) (25). The high-energy intermediate of the reaction is shown in brackets with the soluble form of the transition state analog (hexachloronorbornene derivative 1) below.

were determined by standard dideoxynucleotide methods. The sequences were searched for patterns in their primary structures (gap penalty of 2) (26) and in their possible secondary structures (27, 28).

Truncated RNAs. RNA molecules containing only the consensus secondary structure were transcribed from synthetic DNA templates whose complementary strand had been made by using standard PCR techniques. Deletions were introduced into the 15-B-9 and 15-B-11 PCR templates using primers that hybridized within the double-stranded DNA. The transcription products were purified, applied to the derivatized column, and eluted with 1 ml of 50 mM transition state analog as described above. Each RNA was tested in at least two independent trials.

Assays for Enzymatic Activity. The starch/iodide assay was conducted in 96-well microtiter plates containing (100 μ l per well) 10 μ M RNA, 500 μ M TCTD, 5 mM NEM, 250 μ M I₂, 0.084% starch, 5 mM MgCl₂, 0.5 M NaCl, 50 mM Tris-HCl (pH 7.5), and 3% MeCN. The reactions were monitored at 605 nm with a microtiter plate reader (Molecular Devices) at room temperature for at least 1 hr. The antibody 1E9 (24) was used as a positive control.

The catELISA assay was carried out as described (29, 30).

RESULTS

Selection of RNA Molecules That Bind a Transition State Analog. RNA molecules that bind to the hexachloronorbornene derivative 1 (Fig. 1) were isolated from a pool of $\approx 10^{14}$ sequences containing 80 random nucleotides flanked by 25-base fixed regions. The RNA pool was folded in 50 mM Tris-HCl, pH 7.5/0.5 M NaCl/5 mM MgCl₂ and applied to a Sepharose column to which the analog was covalently attached at a concentration of ≈ 10 mM. The column was washed extensively before bound RNA molecules were eluted with 50 mM soluble analog. cDNA copies of the eluted RNA molecules were synthesized by reverse transcriptase and amplified by PCR. The PCR product was used as template for transcription of RNA for the next round of selection.

In the first six rounds, <1% of the RNA molecules bound to the column. Binding increased to 2.9% in round 7 and 20% in round 8. Two additional rounds of selection and amplification did not significantly increase the amount of RNA bound to the column. Moreover, reapplication of the round 9 RNA that was not retained by the column again resulted in the same percent binding (21%), indicating that the maximum

degree of enrichment attainable by this method had been reached.

Analysis of the nucleotide sequence in 125 clones from the round 9 pool showed that it still consisted of a complex, but not random, mixture. Twelve sequences were repeated one or more times. In addition, a conserved sequence appeared to be present at the 5' end of the random region in 11 of the molecules (see below).

Rounds of selection using concentration gradient elutions were used to further decrease the complexity of the selected RNA pool. In round 11, a step gradient of four column volumes each of 5 mM, 10 mM, 25 mM, and finally 50 mM analog was used to elute bound RNA from the column. RNAs with the weakest affinity for the column-bound analog elute with the lowest concentrations of soluble analog, whereas the higher affinity RNAs require increased concentrations of ligand for elution. RNA molecules that eluted in the 5 mM analog fractions (population A) and 50 mM analog fractions (population B) were collected and amplified. Starting with round 12, population A was applied to the column and eluted with 5 mM analog. In a separate experiment, population B was eluted with the four-step gradient, and those RNAs that dissociated from the column at the highest concentration of analog (50 mM) were collected. Over five rounds of selection and amplification, the elution profile and the total amount of RNA eluting from the column changed according to which part of the gradient was collected and amplified for the next round of selection (Fig. 2). In round 11, 26% of the RNA bound to the column and eluted with a bell-shaped profile. By round 15, the total amount of population A RNA that bound to the column remained unchanged (27%), but the majority eluted with the lowest concentrations of analog. In contrast, the total amount of population B RNA that bound to the column dramatically increased to 41% by round 15 and required higher concentrations of the analog for elution.

Various gradients were used in an attempt to further purify the tightest binding RNA species from the round 15-B pool. However, four rounds using a two-step gradient of 25 mM and 50 mM analog and a fifth round using a third step of 84 mM analog elution did not have an effect on the amount of RNA that bound to the column. A determination of the nucleotide sequences of members of several pools (15-B, 19-B, and 20-B) showed that the gradient elution process had significantly narrowed the complexity from that of the round 9 pool. There were only 82 unique sequences among 153 clones. However, no single sequence dominated the pool.

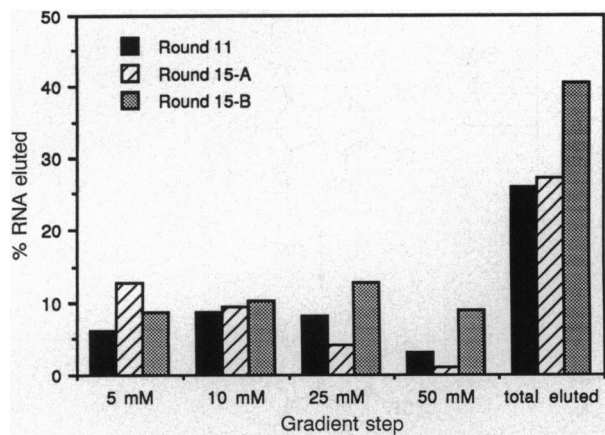


FIG. 2. Elution profiles of the step gradient elutions of round 11, round 15-A, and round 15-B as percent of the total amount of RNA applied to the column. The total percent of RNA eluted is also shown.

A Conserved Sequence. Analysis of the sequence of nucleotides in RNA molecules (26) from the round 9, round 15-B, round 19-B, and round 20-B pools led to the identification of a 21-nucleotide conserved sequence (Fig. 3A) that was present in 10% of the 113 unique round 9 RNAs and in 29% of the 82 unique sequences from later rounds. In all but one case, this sequence occurred at the 5' end of the random region of the RNAs. No sequence homology was found downstream of the conserved element. Folding of these RNA sequences by a computer algorithm (27, 28) consistently yielded a bulged stem-loop structure formed by interaction of the 7 nucleotides at the 3' end of the consensus sequence with the 5' end fixed region (Fig. 3B). Strong evidence for this structure was provided by a clone from round 15-B (15-B-9),

which had a portion of the consensus sequence and was capable of forming the same secondary structure entirely from nucleotides in its random region (Fig. 3C). The stem is composed of different base pairs in 15-B-9, providing covariational evidence for the helix. Surprisingly, even the sequence of the bulge contributed by the fixed sequence in the other RNAs is conserved in 15-B-9, apart from the addition of a single U. An additional G-U base pair leaving a bulged C in the top stem would make the large bulge in 15-B-9 even more similar to the consensus structure (Fig. 3C).

Deletions were introduced into the template encoding 15-B-9 and into a clone with the consensus structure at the 5' end of the molecule (15-B-11) (Fig. 4). The 15-B-9 RNA lacking 35 nucleotides from its 3' end [15-B-9 (95-129)] bound to the column slightly better than the full-length molecule. This deleted region may either destabilize the structure of the full-length RNA or interfere with its recognition of the analog. All of the truncated RNAs in which the sequence of the consensus structure was intact retained at least modest capacity (>10% of the full-length RNA) to bind to the column, whereas those in which the sequence was disrupted no longer bound the analog (<1% of the full-length RNA). In a second experiment, an RNA containing the consensus sequence with the ability to form an 8-base-pair helix (Fig. 3D) was transcribed from a synthetic DNA template and applied to the column. This 38-nucleotide RNA was not retained by the column-bound analog, nor was a similar RNA in which the helix was extended by an additional 6 base pairs. Therefore, the consensus sequence is essential but not sufficient for RNA binding to the column.

Determination of Binding Affinities. The affinity of the selected RNA molecules for the hexachloronorbornene derivative 1 in solution was determined by analytical affinity chromatography (19). The mean K_d for the round 15-A and round 15-B pools was measured to be 1.5 ± 0.5 mM and 0.45

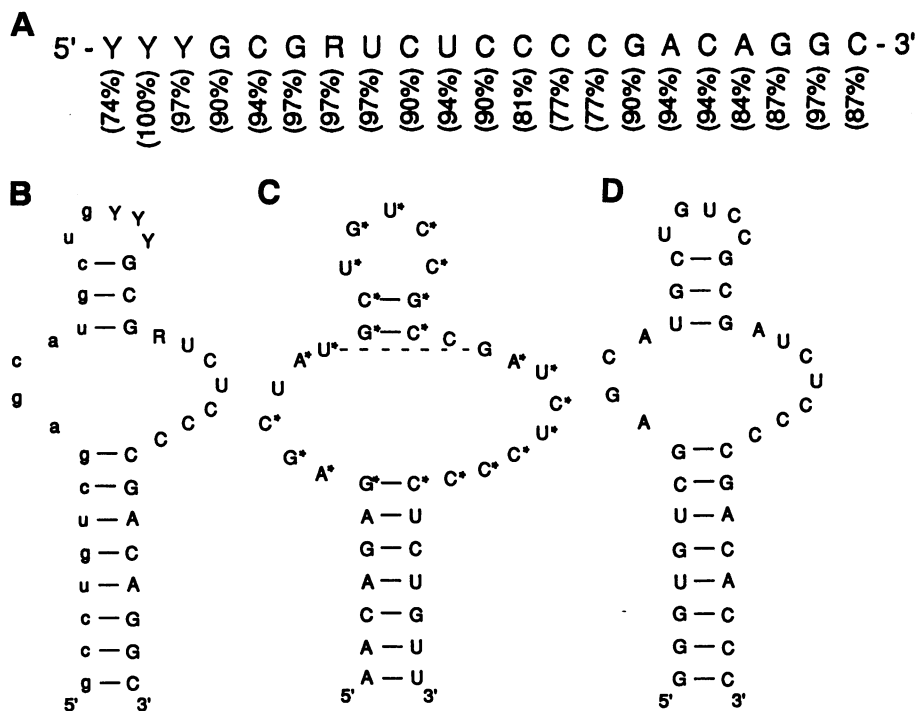


FIG. 3. (A) Consensus sequence identified in the RNA pools that recognize the hexachloronorbornene derivative 1 (Y = pyrimidine; R = purine). The percent conservation of each nucleotide after alignment of the sequences (26) is shown below. (B) Proposed secondary structure (27, 28) formed by a portion of the 5' fixed region (lowercase letters) and the consensus sequence (uppercase letters) at the 5' end of the random region. Base-pairing interactions are shown as lines connecting the nucleotides. (C) Proposed secondary structure formed by the partial consensus sequence in the RNA clone 15-B-9. Nucleotides that are conserved between this structure and that in A are denoted by an asterisk. An additional possible base pair that would make these two structures more similar is shown with a dashed line. (D) The 38-mer RNA that has the consensus sequence and secondary structure but failed to bind to the column derivatized with 1.

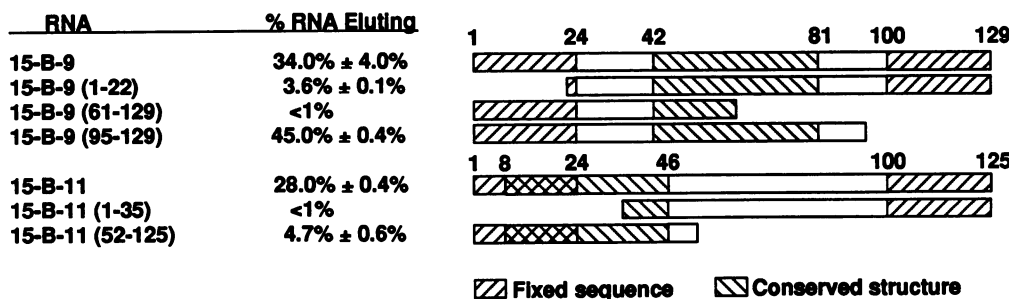


Fig. 4. Effect of deletions on the ability of the RNA clones 15-B-9 and 15-B-11 to bind to the hexachloronorborene column. The percentage of a homogeneous population of each RNA that binds to the column and elutes with 50 mM soluble analog is given. The truncated RNA molecules are designated by the nucleotides that were deleted and are depicted graphically. The regions of fixed sequences and nucleotides that form the proposed conserved secondary structure are designated by the nucleotides in the primary sequence and are shaded with diagonal lines. The fixed sequence that contributes to the secondary structure in 15-B-11 is shown as cross-hatched.

± 0.05 mM, respectively, and that of the 15-B-11 clone, which has the conserved sequence at its 5' end, was determined to be 0.35 ± 0.05 mM.

Screening for Catalysts. Two methods were used to assess the ability of the RNA molecules that bind the transition state analog to catalyze the Diels–Alder reaction. In the first assay, the bleaching of a starch/iodide solution by the SO₂ liberated following the initial cycloaddition of the substrates is monitored spectroscopically. This method was used to screen 117 clones from round 9. A more sensitive catELISA assay (30) was developed and used to screen 18 unique clones from round 15-B, 50 unique clones from round 19-B, and 15 unique clones from round 20-B. In addition, 64 clones from round 9 were rescreened for catalytic activity by this method. In all, 196 unique clones were analyzed by one, or both, of the methods without the identification of an RNA molecule that could accelerate the rate of the Diels–Alder reaction.

DISCUSSION

RNA molecules that bind a transition state analog of a Diels–Alder reaction were isolated from a starting pool of ≈10¹⁴ sequences. After the initial rise and plateau of the amount of RNA retained by the column, the use of a step gradient elution further enriched the pool for RNAs with higher affinities for the target. Variations of this technique could be used to isolate nucleic acids with enhanced affinity for any target from large pools of sequences.

To our knowledge, there have been no previous reports of the isolation of RNA molecules that bind either a nonplanar or a hydrophobic ligand. Other reports have involved ligands that are planar and/or positively charged (17–21). RNA presumably binds the later molecules through base stacking interactions and favorable charge–charge contacts with the phosphate backbone. RNA is apparently capable of forming a relatively hydrophobic binding pocket that complements the nonplanar topology of the hexachloronorborene 1 and may take advantage of the few hydrogen bonding opportunities presented by the imide functionality. A fairly large number of very different RNA sequences appear to form structures capable of binding this molecule.

The conserved nucleotide sequence present in many of the selected RNA molecules (Fig. 3) provides a starting point for examining how RNA can bind this unusual ligand. The proposed secondary structure formed by this conserved sequence and the 5' fixed region is strongly supported by its occurrence entirely within the random region of a clone. The step gradient elution enriched the consensus in the population, suggesting that this sequence contributes to the structure of the RNA molecules with the highest affinities for the analog. The *K_d* of one of these RNAs, 15-B-11 (0.35 ± 0.05 mM), is somewhat lower than the mean of the round 15-B

pool (0.45 ± 0.05 mM). The nucleotide deletion experiments demonstrate that the region of the RNA that contains the structure is essential for binding the analog. However, the conserved sequence alone is not retained by the column. It may require other sequence elements to fold into the correct conformation or to bind the target.

It is discouraging that none of the 196 unique RNA molecules had any detectable catalytic activity. This transition state analog elicited an antibody (1E9) that is able to enhance the reaction with an effective molarity greater than 100 M (25). However, the affinities of the selected RNA molecules are almost five orders of magnitude lower than that of 1E9 (*K_i* = 8 nM) (24). Therefore, it is likely that the lack of catalysis reflects an insufficient amount of binding energy available for increasing the effective concentrations of the two reactants with respect to one another or to stabilize the transition state for the cycloaddition reaction.

We are left to contemplate the general prospects for isolating catalytic RNA molecules through transition state analog methodology. The highest affinities of RNA for small ligands are in the high nanomolar range (20, 21). The relatively sparse chemical features and lack of rotational freedom of nucleotides when compared to amino acids may prevent RNAs from tightly binding the transition state without also having an inhibitory affinity for the substrates or products of the reaction. Therefore, although RNA molecules have been isolated that recognize small molecules with considerable affinity and specificity (17–21), the large, rigid bases may not be able to form binding pockets that complement the chemical topology of small ligands with the same accuracy of a protein binding pocket that is asked to “see” tiny differences in shape. Although RNA is clearly capable of catalyzing certain reactions (1–10, 31), we predict that the difficulties in binding small transition state analogs with higher affinity than substrates may limit the chemical transformations amenable to rate enhancement by nucleic acids.

The limited catalytic repertoire of nucleic acids might be enlarged through the use of several changes and additions to the SELEX procedure described in this report. The SELEX itself might have missed the highest affinity binders (were they present) since column SELEX unavoidably presents the target molecules at high density. Counter-SELEX (21) with substrates or products to enrich those rare sequences that may discriminate the transition state from the ground states of the reaction. In addition, there is compelling evidence that primitive RNA catalysts contained a larger number of bases than the four utilized in modern biological systems (32–34). Toward this end, many pyrimidines with additions at the 5-position are substrates for RNA polymerases and reverse transcriptases (refs. 35 and 36; Bruce Eaton, personal communication), allowing nucleotides that provide additional chemical groups to be incorporated into RNA for use in the

SELEX procedure. In an extension of this, RNA molecules with one of the Diels–Alder substrates covalently attached to pyrimidines could be incubated with the second substrate followed by purification of sequences that have product covalently attached. This approach might allow direct selection for RNA-mediated catalysis of the cycloaddition reaction.

We thank the members of our laboratories for numerous helpful comments and suggestions and Terri R. Gibson for preparation of the figures. This work was supported by National Institutes of Health Grants GM 28685 and GM 19963 to L.G. and GM 38275 to D.H., by the Army Research Office and the Office of Naval Research (N00014-90-J-4089 to D.H.), and by a Damon Runyon–Walter Winchell Cancer Research Fund Fellowship, DRG-1139, to K.N.M. We are pleased to acknowledge the W. M. Keck Foundation and NeXagen for their generous support of RNA science on the Boulder campus.

1. Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E. & Cech, T. R. (1982) *Cell* **31**, 147–157.
2. Guerrier-Takado, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. (1983) *Cell* **35**, 849–857.
3. Piccirilli, J. A., McConnell, T. S., Zaug, A. J., Noller, H. F. & Cech, T. R. (1992) *Science* **256**, 1420–1424.
4. Noller, H. F., Hoffarth, V. & Zimniak, L. (1992) *Science* **256**, 1416–1419.
5. Beaudry, A. A. & Joyce, G. F. (1992) *Science* **257**, 635–641.
6. Green, R. & Szostak, J. W. (1992) *Science* **258**, 1910–1915.
7. Lehman, N. & Joyce, G. F. (1993) *Nature (London)* **361**, 182–185.
8. Pan, T. & Uhlenbeck, O. C. (1992) *Biochemistry* **16**, 3887–3895.
9. Pan, T. & Uhlenbeck, O. C. (1992) *Nature (London)* **358**, 560–563.
10. Bartel, D. P. & Szostak, J. W. (1993) *Science* **261**, 1411–1418.
11. Tuerk, C. & Gold, L. (1990) *Science* **249**, 505–510.
12. Bartel, D. P., Zapp, M. L., Green, M. R. & Szostak, J. W. (1991) *Cell* **67**, 529–536.
13. Schneider, D., Tuerk, C. & Gold, L. (1992) *J. Mol. Biol.* **228**, 862–869.
14. Tuerk, C., MacDougall, S. & Gold, L. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 6988–6992.
15. Schneider, D., Gold, L. & Platt, T. (1993) *FASEB J.* **7**, 201–207.
16. Jellinek, D., Lynott, C. K., Rifkin, D. B. & Janjic, N. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 11227–11231.
17. Ellington, A. D. & Szostak, J. W. (1990) *Nature (London)* **346**, 818–822.
18. Famulok, M. & Szostak, J. W. (1992) *J. Am. Chem. Soc.* **114**, 3990–3991.
19. Connel, G. J., Illangeskare, M. & Yarus, M. (1993) *Biochemistry* **32**, 5497–5502.
20. Lorsch, J. R. & Szostak, J. W. (1994) *Biochemistry* **33**, 973–982.
21. Jenison, R. D., Gill, S. C., Pardi, A. & Polisky, B. (1994) *Science* **263**, 1425–1429.
22. Schultz, P. G. & Lerner, R. A. (1993) *Acc. Chem. Res.* **26**, 391–395.
23. Stewart, J. D., Liotta, L. J. & Benkovic, S. J. (1993) *Acc. Chem. Res.* **26**, 396–404.
24. Hilvert, D. (1993) *Acc. Chem. Res.* **26**, 552–558.
25. Hilvert, D., Hill, K., Nared, K. D. & Auditor, M.-T. M. (1989) *J. Am. Chem. Soc.* **111**, 9261–9262.
26. Hogeweg, P. & Hesper, B. (1984) *J. Mol. Evol.* **21**, 323–333.
27. Jaeger, J. A., Turner, D. H. & Zuker, M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7706–7710.
28. Zuker, M. (1989) *Science* **244**, 48–52.
29. Tawfik, D. S., Green, B. S., Chap, R., Sela, M. & Eshhar, Z. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 373–377.
30. MacBeath, G. & Hilvert, D. (1994) *J. Am. Chem. Soc.* **116**, 6101–6106.
31. Prudent, J. R., Uno, T. & Schultz, P. G. (1994) *Science* **264**, 1924–1927.
32. White, H. B. (1982) in *The Pyridine Nucleotide Coenzymes*, eds. Everse, J., Anderson, B. & You, K. (Academic, New York), pp. 1–17.
33. Switzer, C. Y., Moroney, S. E. & Benner, S. A. (1989) *J. Am. Chem. Soc.* **111**, 8322–8323.
34. Piccirilli, J. A., Krauch, T., Moroney, S. E. & Benner, S. A. (1990) *Nature (London)* **343**, 33–37.
35. Aurup, H., Williams, D. M. & Eckstein, F. (1992) *Biochemistry* **31**, 9636–9641.
36. Crouch, G. J. & Eaton, B. (1994) *Nucleosides Nucleotides* **13**, 939–944.