Evolutionary Optimization of the Catalytic Properties of a DNA-Cleaving Ribozyme†

Joyce Tsang and Gerald F. Joyce*

Departments of Chemistry and Molecular Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037

Received December 22, 1993; Revised Manuscript Received March 7, 1994®

ABSTRACT: In a previous study [Beaudry, A. A., & Joyce, G. F. (1992) *Science 257,* 635–641], an *in vitro* evolution procedure was used to obtain variants of the *Tetrahymena* ribozyme with 100-fold improved ability to cleave a target single-stranded DNA under physiologic conditions. Here we report continuation of the *in vitro* evolution process to achieve 10^5-fold overall improvement in DNA-cleavage activity. In addition, we demonstrate that, by appropriate manipulation of the selection constraints, one can optimize specific catalytic properties of the evolved ribozymes. We first reduced the concentration of the DNA substrate 50-fold to favor ribozymes with improved substrate binding affinity. We then reduced the reaction time 12-fold to favor ribozymes with improved catalytic rate. In both cases, the evolving population responded as expected, first improving substrate binding 25-fold, and then improving catalytic rate about 50-fold. The population of ribozymes has undergone 27 successive generations of *in vitro* evolution, resulting in, on average, 17 mutations relative to the wild type that are responsible for the improved DNA-cleavage activity.

In vitro selection and *in vitro* evolution techniques allow new catalysts to be isolated without *a priori* knowledge of their composition or structure. Such methods have been used to obtain RNA enzymes with novel catalytic properties. Ribozymes that undergo autolytic cleavage with lead cation have been derived from a randomized pool of $tRNA^{Phe}$ molecules (Pan & Uhlenbeck, 1992). Group I ribozyme variants have been isolated that can cleave DNA (Beaudry & Joyce, 1992) or that have altered metal dependence (Lehman & Joyce, 1993a). Starting with a pool of random RNA sequences, molecules have been obtained that catalyze a polymerase-like reaction (Bartel & Szostak, 1993). In the present study, we describe how alteration of the selection constraints during an *in vitro* evolution procedure can lead to refinement of specific catalytic properties of an evolved enzyme.

We began this study with derivatives of the self-splicing group I intron of *Tetrahymena thermophila*, a ribozyme that is able to catalyze sequence-specific cleavage of single-stranded RNA via a phosphoester transfer mechanism (Zaug & Cech, 1986; Zaug et al., 1986). The ribozyme contains a template region, referred to as the internal guide sequence (IGS), which lies at the 5′ end of the molecule and forms Watson–Crick base pairs with the target RNA substrate. The 3′-OH of guanosine, including a guanosine residue that lies at the 3′ end of the ribozyme, is directed to attack a particular phosphoester bond within the ribozyme-bound substrate. A phosphoester transfer reaction ensues, resulting in cleavage of the substrate at a position immediately downstream from the region of base pairing, and concomitant ligation of the 3′ portion of the substrate to the 3′ oxygen of the attacking guanosine. The wild-type *Tetrahymena* ribozyme can cleave a single-stranded DNA substrate with low efficiency under conditions of high magnesium concentration (50 mM $MgCl_2$) and/or high temperature (50 °C) (Herschlag & Cech, 1990a; Robertson & Joyce, 1990). Under more physiologic conditions (e.g., 37 °C, 10 mM $MgCl_2$, pH 7.5), however, the DNA-cleavage reaction is almost undetectable.

Beaudry and Joyce (1992) have described an *in vitro* evolution procedure that was used to obtain variants of the *Tetrahymena* ribozyme that can cleave DNA under physiologic conditions with improved efficiency compared to the wild type (Figure 1). At the beginning of this procedure, a population of ribozyme variants was generated by partially randomizing the phylogenetically conserved portions of the molecule that are known to be essential for catalytic activity. Superior DNA-cleaving ribozymes were distinguished from less active molecules on the basis of the likelihood of attachment of the 3′ portion of the substrate to the 3′ end of the ribozyme. A DNA primer was hybridized across the ligation junction of successful reaction products and used to initiate a selective isothermal amplification reaction (Figure 1, bottom). The selectively amplified molecules then served as templates for cDNA synthesis, the resulting cDNA was amplified by the polymerase chain reaction (PCR) (Saiki et al., 1985, 1988), and the PCR products were transcribed to generate a new pool of RNAs. The entire process, beginning with the cleavage reaction, and followed by selective isothermal amplification, cDNA synthesis, PCR amplification, and *in vitro* transcription, constitutes one generation of the *in vitro* evolution procedure.

Beaudry and Joyce (1992) carried out 10 successive generations, starting with a pool of 10^{13} variants of the *Tetrahymena* ribozyme (Figure 2). After the 9th generation (G9), individual ribozymes were isolated from the population and shown to catalyze the cleavage of a DNA substrate 100-fold more efficiently compared to the wild-type enzyme. This modest improvement in catalytic efficiency resulted from both an increased catalytic rate (k_{cat}) and a decreased value for the Michaelis constant (K_M). The outcome, however, was somewhat dissatisfying because the ribozymes were still inefficient catalysts in an absolute sense, with k_{cat}/K_M on the order of 10^3 M^{-1} min^{-1}.

For each generation, the evolving population was provided with 10 μM DNA substrate and allowed 1 h to carry out the DNA-cleavage reaction. By G9, K_M had improved from 6

†Supported by Research Grant AI 30882 from the National Institutes of Health.

* Author to whom correspondence should be addressed.

® Abstract published in *Advance ACS Abstracts,* April 15, 1994.

0006-2960/94/0433-5966$04.50/0

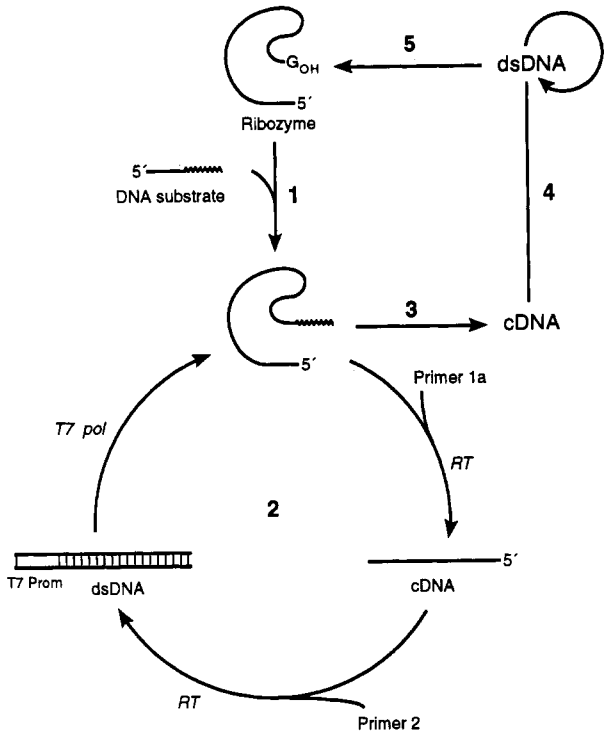

FIGURE 1: *In vitro* evolution procedure. Steps: (1) Cleavage of the DNA substrate via phosphoester transfer results in ligation of the 3′ portion of the substrate to the 3′ end of the ribozyme. (2) Selective isothermal amplification of DNA-cleaving ribozymes: first, selective primer 1a hybridizes to the extended 3′ terminus of active molecules and initiates cDNA synthesis in the presence of reverse transcriptase *(RT)*; next, primer 2, which contains a T7 promoter sequence, hybridizes to the cDNA and initiates second-strand DNA synthesis; finally, T7 RNA polymerase *(T7 pol)* produces multiple copies of the selected RNA, each of which can enter a new round of amplification. (3) Selective cDNA synthesis employing primer 1a and reverse transcriptase. (4) PCR amplification employing nonselective primers 1b and 2 restores the original terminus of the ribozyme-encoding gene and introduces occasional mutations. (5) *In vitro* transcription to produce the progeny population of ribozymes.

μM for the wild type to about 2 μM for the evolved individuals (Beaudry & Joyce, 1992). Accordingly, we felt that the population was no longer under stringent selection pressure to drive further improvement of K_M. Individual cleavage rates, on the other hand, were on the order of 0.007 min^{-1} by G9, still slow enough to be constrained by the 1-h incubation period. However, if the reaction rate continued to improve, then the selection constraints would eventually become insufficient to favor further improvement of the catalytic rate. We realized that additional generations of *in vitro* evolution, under different selection constraints, would be necessary to obtain substantially greater DNA-cleavage activity.

In this study, we apply *in vitro* evolution techniques with a higher level of sophistication and control. Because the outcome of an *in vitro* evolution experiment depends on the nature of the selection constraints, specific catalytic properties of a ribozyme, such as substrate binding affinity, catalytic rate, substrate specificity, and turnover, might be improved by appropriate manipulation of the reaction conditions. With this in mind, we set out to optimize two catalytic properties of the DNA-cleaving ribozymes, namely, substrate binding affinity and catalytic rate. We reasoned that ribozymes with the greatest affinity for the substrate would enjoy a selective advantage when the substrate is presented at low concentrations. Under saturating conditions, ribozymes with the fastest first-order rate of reaction would be favored when the reaction time is very short.

We resurrected the previously-characterized G9 population of DNA-cleaving ribozymes and carried out 18 additional generations of *in vitro* evolution under somewhat different reaction conditions. From generations 10 through 18, we reduced the substrate concentration 50-fold, from 10 to 0.2 μM. From generations 19 through 27, we maintained the lower substrate concentration and reduced the reaction time 12-fold, from 1 h to 5 min. On the basis of binding and kinetic studies, the population of ribozymes responded to each alteration of the selection constraints as predicted, becoming enriched with tighter substrate binders during generations 10–18, and then with faster catalysts during generations 19–27.

MATERIALS AND METHODS

Materials. Unlabeled nucleoside triphosphates (NTPs) and deoxynucleoside triphosphates (dNTPs) were purchased from Pharmacia, and dideoxynucleoside triphosphates (ddNTPs) were from U.S. Biochemical. [α-^{32}P]GTP, [γ-^{32}P]ATP, and [^{3}H]UTP were from ICN Radiochemicals. Synthetic oligodeoxynucleotides were obtained from Operon Technologies and purified by polyacrylamide gel electrophoresis and subsequent chromatography on Sephadex G-25. Restriction enzymes and T4 polynucleotide kinase were from New England Biolabs, calf intestine phosphatase was from Boehringer, AMV reverse transcriptase was from Life Sciences, MoMLV reverse transcriptase and Sequenase 2.0 (modified T7 DNA polymerase) were from U.S. Biochemical, and *Taq* DNA polymerase was from Cetus. T7 RNA polymerase was prepared as previously described (Davanloo et al., 1984) and purified according to a procedure originally developed for SP6 RNA polymerase (Butler & Chamberlain, 1982).

Preparation of Wild-Type Ribozyme. The L-21 form of the *Tetrahymena* ribozyme was prepared by *in vitro* transcription of *Hin*dIII-digested pT7L-21 plasmid DNA (Zaug et al., 1988). The transcription reaction mixture contained 0.1 μg/μL of cut plasmid, 15 mM $MgCl_2$, 2 mM spermidine, 50 mM Tris (pH 7.5), 5 mM DTT, 2 mM each NTP, 0.005 unit/μL inorganic pyrophosphatase, and 25 units/μL T7 RNA polymerase, incubated at 37 °C for 2 h. The 23-nucleotide 3′ exon sequence was removed by RNA-catalyzed site-specific hydrolysis (Inoue et al., 1986): RNA was incubated in the presence of 50 mM CHES (pH 9.0) and 10 mM $MgCl_2$ at 42 °C for 1 h. The resulting RNA was isolated by electrophoresis in a 5% polyacrylamide/8 M urea gel, visualized by UV shadowing, eluted from the gel overnight at room temperature in a buffer containing 200 mM NaCl, 10 mM Tris (pH 7.5), and 0.5 mM EDTA, and purified by affinity chromatography on DuPont Nensorb. The concentration of ribozyme was determined spectrophotometrically, based on $\epsilon_{260} = 3.2 \times 10^6$ M^{-1} cm^{-1} (Zaug et al., 1988).

In Vitro Evolution Procedure. In vitro evolution was carried out as described previously (Beaudry & Joyce, 1992) and as depicted in Figure 1. We began with the population of DNA-cleaving ribozymes obtained after 9 generations of *in vitro* evolution in a previous study (Beaudry & Joyce, 1992). Ribozymes (0.1 μM) and DNA substrate (0.2 μM) were incubated at 37 °C for 1 h in a 100 μL volume containing 10 mM $MgCl_2$ and 30 mM EPPS (pH 7.5). After ethanol precipitation, a portion of the reaction products (10–50%) was added to a 20-μL isothermal amplification reaction mixture (Guatelli et al., 1990), containing 10 mM $MgCl_2$, 80 mM KOAc, 50 mM Tris (pH 7.5), 5 mM DTT, 2 mM each NTP, 0.2 mM each dNTP, 4 μCi [α-^{32}P]GTP, 12.5 units/μL MoMLV reverse transcriptase, 50 units/μL T7 RNA poly-

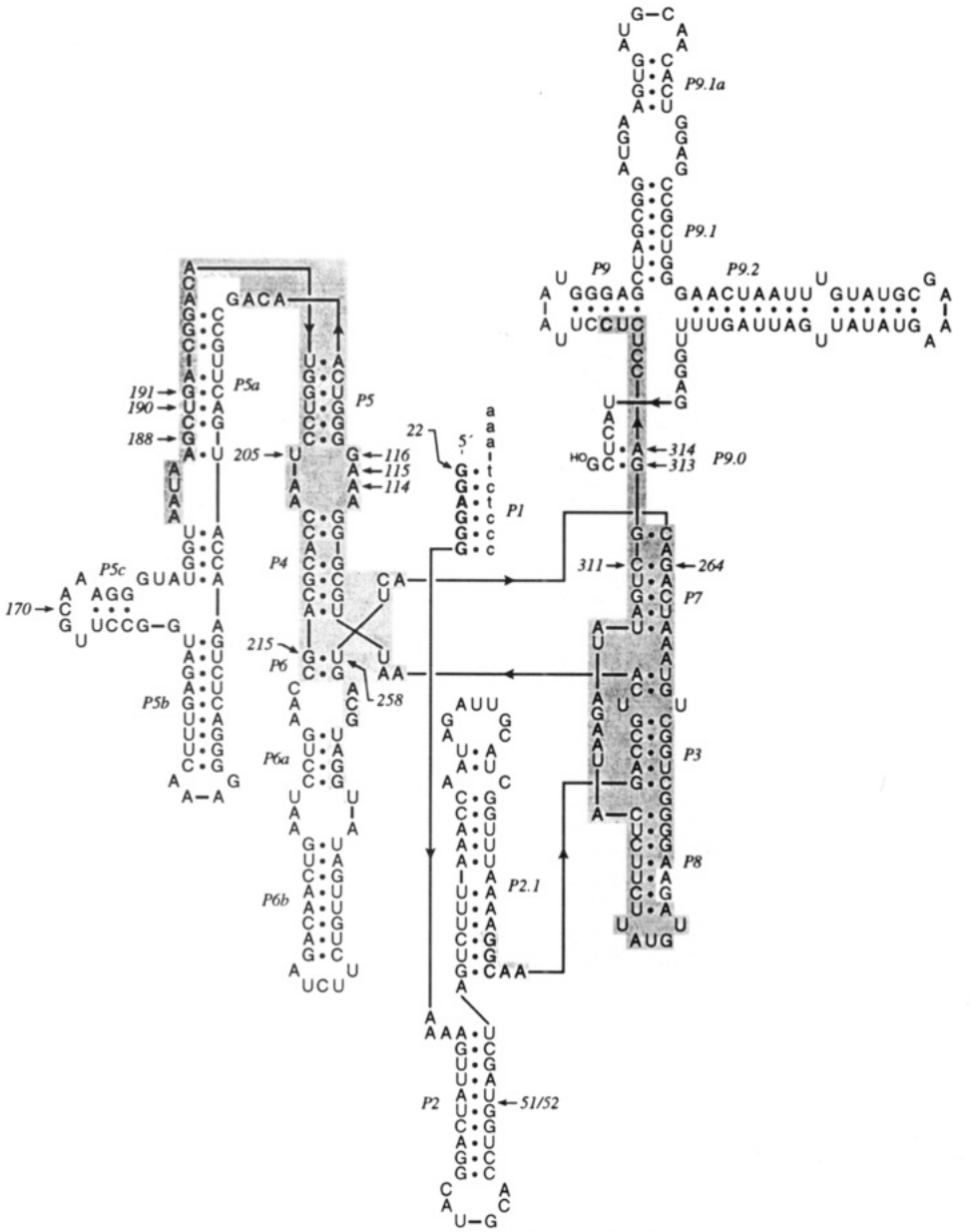

FIGURE 2: Secondary structure of the *Tetrahymena* ribozyme (Cech et al., in press). Paired structural elements are designated by P*i*. Joining regions between paired elements *i* and *j*, referred to as J*i*/*j*, are not labeled. Nucleotide positions that were partially randomized in the initial population are indicated by shaded regions. The internal guide sequence (IGS) is shown in boldface, and the DNA substrate is shown in lowercase letters. Nucleotide positions discussed in the text are labeled.

merase, and 20 pmol each of 5′-TTTATTTATTTATTTC-3′ (primer 1a) and 5′-CTGCAGAATTCTAATACGACTCAC-TATAGGAGGGAAAAGTTATCAGGC-3′ (primer 2), which was incubated at 37 °C for 2 h. Primer 1 hybridizes to the 3′ portion of the substrate that becomes attached to the 3′ end of the ribozyme. Primer 2 hybridizes to the 3′ end of the resulting cDNA and introduces the T7 promoter sequence. Twenty-five percent of the isothermal amplification products were used to generate cDNA in a 20-μL reaction mixture containing 10 mM $MgCl_2$, 50 mM Tris (pH 7.5), 5 mM DTT, 2 mM each NTP, 0.2 mM each dNTP, 0.2 units/μL AMV reverse transcriptase, and 20 pmol of primer 1a, incubated at 37 °C for 1 h. Approximately 5–10% of the resulting cDNA was amplified by the PCR in a 100-μL reaction mixture containing 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris (pH 8.3), 0.1% gelatin, 0.2 mM each dNTP, 20 pmol 5′-CGAGTACTCCAAAACTAATC-3′ (primer 1b), 20 pmol of primer 2, and 2.5 units of Taq DNA polymerase, carried out for 30 cycles of 92 °C for 1 min, 45 °C for 1 min, and 72 °C for 1 min, and 1 cycle of 72 °C for 10 min. Primer 1b is complementary to the 3′ end of the ribozyme, allowing regeneration of its original, active form. PCR DNA (~250–500 ng, 5–10% of the total) then served as template in an *in vitro* transcription reaction, carried out in a 25–50 μL volume. The transcribed RNA was isolated by polyacrylamide gel electrophoresis, visualized by UV shadowing, cut and eluted from gel, purified on DuPont Nensorb, and quantified spectrophotometrically, as described above. The entire process was repeated 18 times, the first 9 as described above and the second 9 with the incubation time for the cleavage reaction reduced from 1 h to 5 min. Occasionally, the cDNA was purified to improve the quality of the PCR amplification. To do so, cDNA was synthesized as above except in the presence of 25–50 μCi of [α-^{32}P]dATP. Labeled cDNA was isolated by electrophoresis in a 5% polyacrylamide/8 M urea gel, visualized by autoradiography, cut and eluted from gel, and purified on DuPont Nensorb.

Shotgun Cloning, Sequencing, and Preparation of Individual Ribozymes. The G18 subclones were obtained as previously described (Beaudry & Joyce, 1992). The G27 subclones were obtained using the Invitrogen TA Cloning Kit. The PCR DNA at G27 was ligated into a linearized plasmid, and the resulting DNA was used to transform competent INVαF′ cells, which were grown on ampicillin/X-gal plates. Individual colonies containing the insert were identified by their white color, chosen at random, and grown overnight in liquid media. Plasmid DNA was prepared by the boiling, lysis method (Holmes & Quigley, 1981) and screened for the presence of insert by restriction digestion. Cloned individuals were sequenced by the dideoxy chain-termination method, as previously described (Sanger et al., 1977; Beaudry & Joyce, 1992). Complete sequences of individual subclones are available upon request. Individual ribozymes were prepared as follows: the gene encoding the ribozyme was amplified by the PCR using primers 1b and 2; the resulting DNA was used as a template for *in vitro* transcription; the RNA products were isolated by polyacrylamide gel electrophoresis and were purified and quantified as described above.

Preparation of Substrate and Product Oligonucleotides. The DNA substrate 5′-GGCCCTCTAAATAAATA-3′ and DNA product 5′-GGCCCTCT-3′ were (5′-^{32}P)-labeled in a 20 μl reaction mixture containing 20 pmol of oligonucleotide, 10 pmol of (4.5 μCi/pmol) [γ-^{32}P]ATP, 5 mM $MgCl_2$, 25 mM CHES (pH 9.0), 3 mM DTT, and 1.25 units/μL T4 polynucleotide kinase, incubated at 37 °C for 1 h. Labeled oligonucleotide was isolated by electrophoresis in a 20% polyacrylamide/8 M urea gel, visualized by autoradiography, eluted from the gel, and purified on DuPont Nensorb.

The RNA substrate 5′-GGCCCUCUAAAUAAAUA-3′ was prepared by *in vitro* transcription using a partially single-stranded synthetic DNA template (Milligan et al., 1987), as described previously (Beaudry & Joyce, 1990). The RNA transcript was dephosphorylated with calf intestine phosphatase, extracted with phenol and chloroform, and then (5′-^{32}P)-labeled and purified as described above.

Kinetics Analysis. All cleavage reactions were carried out at 37 °C in 10 mM $MgCl_2$, 30 mM EPPS (pH 7.5), and 40 μg/μL BSA, using (5′-^{32}P)-labeled substrate. BSA was added to prevent oligonucleotides from adhering to the walls of the 500-μL Eppendorf tubes and did not affect the course of the reaction. Ribozyme and substrate were preincubated separately for 15 min at 37 °C and then mixed to initiate the reaction. If instead the preincubation was carried out at 50 °C for 20 min (Herschag & Cech, 1990b), there was no significant difference in the observed catalytic efficiency of the ribozyme. Typically, 5 aliquots of 3–10 μL each were removed from the reaction mixture at specified times and quenched by addition to 1–2 volumes of an ice-cold mixture containing 8 M urea, 50–100 mM EDTA, 0.05% xylene cyanol, 0.05% bromophenol blue, 10% SDS, 9 mM Tris–borate (pH 8.3), and 20% sucrose. Substrate and product were separated by electrophoresis in a 20% polyacrylamide/8 M urea gel, visualized by autoradiography, excised from the gel, and quantified by Cerenkov counting.

K_M and k_{cat} values were determined in experiments with substrate (S) in excess over ribozyme (E). Initial rates of reaction (v_0), over a range of substrate concentrations, were estimated from the initial linear phase, generally the first 5% or less of the reaction. Typically, 8 data points were fit by a least-squares method to a theoretical line given by the equation: $v = -K_M(v_0/[S]) + V_{max}$.

Single-turnover experiments were performed with ribozyme in excess of substrate (Herschlag & Cech, 1990b). Initial rates (k_{obs}) were obtained using no more than the first 5% of the reaction. Given that $k_{cat}/K_M = k_{obs}/[E]$, each k_{obs} value, obtained at different ribozyme concentrations, provided an estimate of k_{cat}/K_M. Eight or more measurements of k_{obs} were obtained at 4 or more different concentrations of ribozyme; the data were fit by a least-squares method to the equation: $k_{obs} = (k_{cat}/K_M)[E]$.

Determination of Binding Constants. The equilibrium dissociation constant, K_D, of the complex between ribozyme and DNA product (P) was determined by gel-shift analysis in a native polyacrylamide gel (Pyle et al., 1990). Ribozyme at twice the final concentration was preincubated at 37 °C for 15 min in 10 mM $MgCl_2$ and 30 mM EPPS (pH 7.5) before mixing with an equal volume of 0.05–1 nM (5′-^{32}P)-labeled DNA product in 10 mM $MgCl_2$, 30 mM EPPS (pH 7.5), 0.05% xylene cyanol, 3% glycerol, and 80 μg/μL BSA. In all cases, the concentration of ribozyme was at least 20-fold greater than that of DNA product. The mixture was allowed to equilibrate at 37 °C min for 15–60 min before loading on a 10% polyacrylamide gel containing 10 mM $MgCl_2$ and 30 mM EPPS (pH 7.5). Longer equilibration times did not alter the binding profile. The electrophoresis buffer also contained 10 mM $MgCl_2$ and 30 mM EPPS (pH 7.5). The gel was run at 6 mA in a 37 °C room until the sample had entered the gel (~10 min), and then moved into a 4 °C cold room where the current was increased to 30 mA. This was done to prevent the temperature of the gel from rising above 37 °C. The ribozyme–product complex and free product were visualized by autoradiography, cut from the gel, and quantified by Cerenkov counting.

A binding curve was generated by plotting the percentage of product bound to ribozyme (% bound) over a range of ribozyme concentrations. K_D was determined by fitting the data to a theoretical binding curve using a least-squares method. Because ribozyme was in vast excess over product, the theoretical binding curve could be represented by the equation: % bound = $[E]/([E] + K_D)$, where $K_D = [E]$ when half of the total product is bound to the ribozyme.

RESULTS

Evolution in Vitro. Beginning with the 9th generation (G9) population of ribozymes obtained in a previous study (Beaudry & Joyce, 1992), we carried out 18 additional generations of *in vitro* evolution. Variation in the population was maintained by PCR amplification, which introduces mutations at a rate of ~0.1% per nucleotide position per generation. Because mutation is ongoing, evolution based on Darwinian principles can occur. Progeny ribozymes have the opportunity to acquire new mutations that confer favorable attributes not possessed by the parent molecules. This phenomenon is reflected by the steadily increasing frequency of accepted mutations over the 27 generations.

Sequence data were obtained from 50 randomly-chosen subclones, isolated from the evolving population at G9, G18, and G27 (Figure 3). The mean number of mutations per subclone rose from 5.9 at G9, to 12.7 at G18, and to 16.5 at G27. Most of the mutations occurred within the phylogenetically conserved portions of the ribozyme that were randomized in the initial population (Figure 2). However, 26% of the total mutations at G18, and 38% at G27, occurred in peripheral regions as a result of ongoing mutagenesis. Most of the commonly-occurring mutations (>30% frequency) that occur in the G18 subclones (Figure 3B) were not observed at

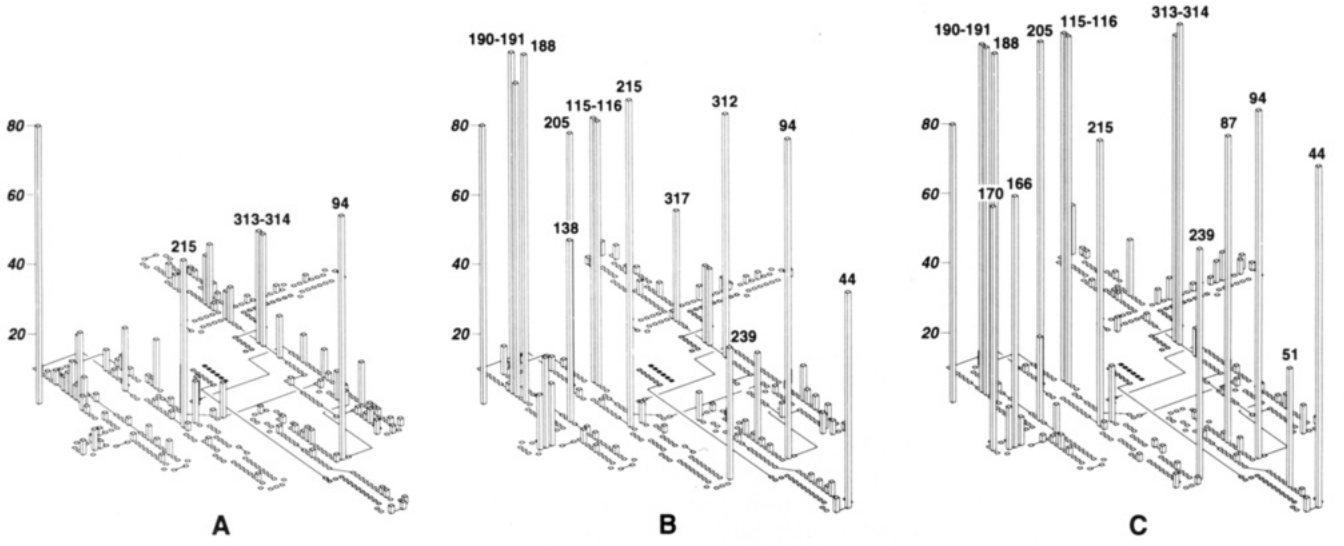

FIGURE 3: Sites at which mutations occurred over the course of evolution, superimposed on the secondary structure of the *Tetrahymena* ribozyme. Box height corresponds to the frequency of mutations (%) at each nucleotide position, on the basis of 50 subclones sequenced at G9, G18, and G27. Nonmutable primer binding sites are shaded; substrate is shown in black. Commonly-occurring mutations (>30% frequency) are labeled. (A) G9 (Beaudry & Joyce, 1992); (B) G18; (C) G27.

G9 (Figure 3A), suggesting that these mutations arose in response to the increased selection pressure designed to enhance substrate binding affinity. Between G18 and G27, nearly all of the most commonly-occurring mutations continued to increase in frequency (Figure 3C). However, two significant mutations, the NGAA insertion between positions 51 and 52 and the C → U change at position 170, first appeared during this interval, suggesting that these mutations arose in response to the increased selection pressure designed to enhance the catalytic rate.

Concerted and Mutually-Exclusive Mutations. As demonstrated in a previous study (Lehman & Joyce, 1993b), RNA populations display tendencies similar to naturally evolving systems, including the occurrence of mutations that are concerted (and presumably synergistic) or mutally exclusive. For example, the changes at nucleotide positions 188, 190, and 191 in the P5a region (Figure 2) co-occur in 90% of subclones, while mutations in the J4/5 and J5/4 internal loop at positions 115, 116, and 205 co-occur in 68% of the subclones at G18. Interestingly, the J4/5 and J5/4 mutations co-occur only if the set of P5a mutations is also present ($\chi^2 = 110$, $p < 0.001$), suggesting an interaction between these two regions.

The 313: G → Y and 314: A → G mutations nearly always occur together. These mutations co-occur in 16 of 50 subclones at G9, 11 of 50 subclones at G18, and 44 of 50 subclones at G27. Only two G27 subclones contain the mutation at position 313 but lack the mutation at position 314. At G9 and G18, the 313 mutation always occurs as a G → U change. At G27, however, the 313 mutation occurs primarily as a G → C change, with the G → U change occurring only once. The GA sequence normally present at positions 313 and 314 is thought to form a short duplex structure (P9.0) that brings the 3′-terminal guanosine residue of the ribozyme into the catalytic core (Michel et al., 1989, 1990). The 3′-OH of this guanosine serves as the nucleophile in the RNA-catalyzed phosphoester transfer reaction. Although the 313–314 mutation would prevent the P9.0 duplex from forming, the 313–314: GA → UG change confers selective advantage with respect to the DNA-cleavage reaction, as demonstrated by site-directed mutagenesis studies (Beaudry & Joyce, 1992). The appearance of the 313–314: GA → CG change, between G18 and G27, suggests that this altered form of the 313–314 mutation may contribute to the improved catalytic rate of the DNA-cleavage reaction.

The 312: G → A mutation occurs only if the 313–314: GA → YG mutations are not present. The 312: G → A change is present in 4 of 25 subclones at G3, 8 of 25 subclones at G6, and 5 of 50 subclones at G9 (Beaudry & Joyce, 1992). There is a dramatic rise in the frequency of the 312: G → A mutation between G9 and G18, followed by an equally dramatic drop between G18 and G27 (Figure 3). As the frequency of the 312: G → A mutation declines, the 313–314: GA → YG mutations become more abundant.

The 215: G → A mutation, present at high frequency in all of the studied populations, putatively allows a Watson–Crick base pair to form with the U at position 258 (Figure 2). This change is present in nearly all of the subclones at G18 and G27. Of the 12 individuals that lack this mutation, 11 carry a U → C change at position 258, which would allow a Watson–Crick pair to form with the wild-type G at position 215. Thus, in 99 of 100 subclones from G18 and G27, a Watson–Crick base pair is expected to form between positions 215 and 258.

Improvement of DNA Binding Affinity. Beginning with G10, we lowered the concentration of DNA substrate employed during the RNA-catalyzed reaction from 10 to 0.2 μM to impose increased selection pressure favoring individuals with enhanced substrate binding affinity. In order to assess the impact of this change, K_D values for the complex between ribozyme and DNA product (GGCCTCT) were determined for the population of ribozymes at every third generation over the 27 generations (Figure 4B; see Materials and Methods). We employed the DNA product rather than substrate to avoid a cleavage reaction during the gel-shift analysis. The binding affinity for the product is assumed to be similar to that of the substrate, on the basis of previous studies showing that the wild-type ribozyme binds the RNA substrate with roughly the same affinity as it binds the product (Pyle et al., 1990; Herschlag & Cech, 1990b).

Binding data for each studied population were fit to a theoretical binding curve, an example of which is shown in Figure 4A for the G27 population. As expected, the greatest improvement in binding affinity occurred between G9 and G18 (Figure 4B), subsequent to tightening of the selection

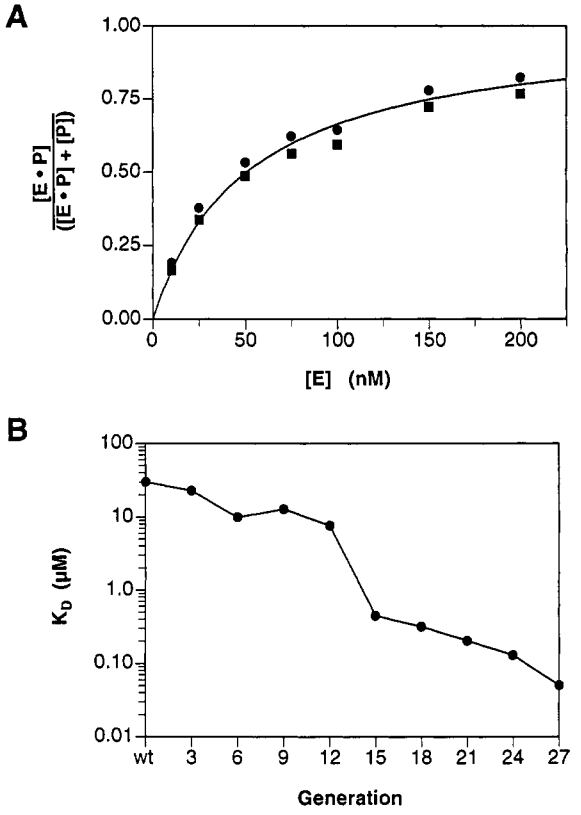

FIGURE 4: Improvement in substrate binding affinity over 27 successive generations of *in vitro* evolution. (A) A typical binding curve showing data obtained for the G27 population of ribozymes. ● and ■ indicate data from two different gel-shift experiments. Data were fit by a least-squares method to a theoretical binding curve (indicated by solid line), given by the equation: $y = [E]/([E] + K_D)$, where y is the fraction of product (P) bound to ribozyme (E). In this case, K_D = 51 (±2) nM. (B) K_D for the population of ribozymes at every third generation. Standard errors averaged 11%.

constraints. After G18, the population became saturated with ribozymes having a K_D of less than 0.2 μM, accounting for the slow but continued improvement between G18 and G27.

Kinetic Analysis. Beginning with generation 19, we reduced the reaction time from 1 h to 5 min to favor selection of ribozymes with increased k_{cat} values. To study the effect of this change, we chose two individuals isolated from the population at G9, G18, and G27 for formal kinetic analysis (Table 1). These ribozymes are representative of the population from which they were isolated because they contain most of the prominent mutations that occur in their respective populations. In addition, the total number of mutations in each of the studied individuals coincides with the mean number of mutations per subclone in the corresponding population. We emphasize that the k_{cat} and K_M values of the studied individuals are not equivalent to the average k_{cat} and K_M values for the entire population. It is likely that the catalytic efficiencies of the studied ribozymes are somewhat higher than the average because these ribozymes possess a greater fraction of the dominant mutations than a typical individual in the population. Nevertheless, the relative differences in k_{cat} and K_M values between representative pairs of individuals should be comparable. As expected, the improvement in k_{cat} is greatest between the G18 and G27 ribozymes (Table 1), while the improvement in K_M is greatest between the G9 and G18 ribozymes.

RNA-Cleavage Activity of G27 Ribozymes. In order to assess the effect of the evolution procedure on RNA-cleavage activity, we compared the efficiency of RNA cleavage by

Table 1: Catalytic Parameters of DNA-Cleaving Ribozymes

ribozyme	mutations	k_{cat}[b] (min^{-1})	K_M[b] (μM)	k_{cat}/K_M (M^{-1} min^{-1})
wild type[a]	0	2.4 (±0.2) × 10^{-4}	6.0 ± 1.7	4.0 × 10^1
G9 #23[a]	7	5.1 (±0.2) × 10^{-3}	1.8 ± 0.3	2.8 × 10^3
G9 #29[a]	6	7.1 (±0.3) × 10^{-3}	1.9 ± 0.3	3.8 × 10^3
G18 #13[c]	12[f]	1.7 (±0.1) × 10^{-2}	0.24 ± 0.04	7.1 × 10^4
G18 #66[c]	13[g]	1.1 (±0.1) × 10^{-2}	0.32 ± 0.08	3.5 × 10^4
G27 #48[d]	17[h]	7.0 (±0.6) × 10^{-1}	0.31 ± 0.05	2.3 × 10^6
G27 #61[e]	15[i]	3.3 (±0.7) × 10^{-1}	0.11 ± 0.06	2.9 × 10^6

[a] Data obtained previously (Beaudry & Joyce, 1992), modified slightly as a result of subsequent statistical analysis. [b] Measurements were carried out as described in Materials and Methods with the following: [c] 0.025 μM ribozyme and 0.125, 0.25, 0.5, and 1.0 μM DNA substrate; [d] 0.02 μM ribozyme and 0.1, 0.2, 0.4, and 0.8 μM DNA substrate; or [e] 0.02 μM ribozyme and 0.05, 0.1, 0.2, and 0.4 μM DNA substrate. [f] 44: G → A, 94: A → U, 115: A → U, 116: G → A, 138: C → A, 188: G → A, 190: U → A, 191: G → U, 205: U → C, 215: G → A, 312: G → A, and 317: U → G. [g] 44: G → A, 94: A → U, 115: A → U, 116: G → A, 138: C → A, 167: U → G, 188: G → A, 190: U → A, 191: G → U, 205: U → C, 215: G → A, 239: U → A, 312: G → A. [h] 44: G → A, 51/52: insert AGAA, 87: A → del, 94: A → U, 115: A → U, 116: G → A, 166: C → A, 170: C → U, 188: G → A, 190: U → A, 191: G → U, 205: U → C, 215: G → A, 239: U → A, 312: G → A, 350: C → U, and 364: C → U. [i] 44: G → A, 51/52: insert AGAA, 87: A → del, 94: A → U, 115: A → U, 116: G → A, 166: C → A, 170: C → U, 188: G → A, 190: U → A, 191: G → U, 205: U → C, 215: G → A, 313: G → C, and 314: A → G.

both the G27 #48 and G27 #61 ribozymes to that of the wild type. Single-turnover kinetic experiments revealed that the G27 ribozymes have slightly enhanced RNA-cleavage activity: k_{cat}/K_M values are 2.8 (±0.1) × 10^7 and 2.1 (±0.3) × 10^7 M^{-1} min^{-1} for clones G27 #48 and G27 #61, respectively, compared to 9.7 (±1.9) × 10^6 M^{-1} min^{-1} for the wild type. Thus, the 27 generations of *in vitro* evolution resulted in a 10^5-fold improvement of DNA-cleavage activity and a 2- to 3-fold enhancement of RNA-cleavage activity. Similarly, gel-shift experiments revealed a significantly greater improvement in DNA binding affinity compared to RNA binding affinity. Ribozymes G27 #48 and G27 #61 bind the DNA product with a K_D of 4 nM and 1 nM, respectively, compared to 30 μM for the wild type, and bind the RNA product with a K_D of 0.5 nM and 0.4 nM, respectively, compared to 1.5 nM for the wild type. Thus, the G27 ribozymes exhibit a 10^4-fold improvement in DNA binding affinity and a 3- to 4-fold improvement in RNA binding affinity.

DISCUSSION

We have shown that specific catalytic properties of a DNA-cleaving ribozyme can be optimized by appropriate manipulation of the selection constraints during an *in vitro* evolution procedure. Beginning with a heterogeneous population of ribozymes, enriched for modest DNA-cleavage activity, 18 additional generations were carried out to obtain DNA-cleaving ribozymes that have a catalytic rate of 0.7 min^{-1} and a substrate binding affinity of 10^{-9} M. These catalytic parameters are improved 10^3-fold and 10^4-fold, respectively, compared to the wild type. The greatest improvement in K_D and K_M (Figure 4B; Table 1) occurred between G9 and G18 in response to alteration of the selection constraints to favor ribozymes with enhanced affinity for the DNA substrate. Likewise, on the basis of k_{cat} values for representative individuals (Table 1), the greatest improvement in k_{cat} occurred between G18 and G27, following alteration of the selection constraints to favor a faster rate of catalysis.

The DNA binding affinity of the G27 #48 and G27 #61 ribozymes is comparable to the RNA-binding affinity of the wild-type ribozyme. Previous studies have suggested that the wild-type ribozyme binds RNA more strongly than DNA as a result of interactions between the 2′-OH groups of the RNA substrate and specific nucleotides within the catalytic core of the ribozyme (Pyle & Cech, 1991; Pyle et al., 1992; Herschlag et al., 1993). In binding the DNA substrate with nanomolar affinity, the evolved ribozymes must compensate for the lack of substrate 2′-OH groups by forming alternative interactions that provide an additional 5 kcal mol^{-1} of binding energy (at 37 °C).

The *Tetrahymena* ribozyme binds its substrate through a two-step process involving, first, Watson–Crick base pairing between the internal guide sequence (IGS) and the substrate and, second, docking of the IGS/substrate duplex (P1 helix) into the catalytic core of the ribozyme via tertiary interactions (Herschlag, 1992; Bevilacqua et al., 1992). Because the sequence of both the IGS and substrate is unchanged throughout the *in vitro* evolution procedure and because the first step is not rate-limiting under our selection conditions, it is unlikely that we have evolved compensatory mutations that operate at the first step of binding. Instead, the 5 kcal mol^{-1} of additional binding energy is likely to result from additional tertiary interactions that affect the second step of binding.

We can learn more about such interactions by examining the specific mutations that arose in response to the increased selection pressure aimed to improve substrate binding. For example, mutations at positions 115, 116, and 205 in the J4/5 and J5/4 internal loop of the ribozyme (Figure 2) became prominent in the population between G9 and G18. Both a tertiary structural model of the wild-type ribozyme (Michel & Westhof, 1990) and experimental evidence suggest that residues in the J4/5 region may interact with the IGS. On the basis of cross-linking data, Wang et al. (1993) concluded that A114 and A115 lie in close proximity to G22 of the IGS (the 5′-terminal residue of the L-21 form of the ribozyme) when P1 is docked into the ribozyme core. The mutations at positions 115 and 116 may enhance P1 docking by allowing new contacts to be made with the IGS, compensating for the lack of 2′-OH groups in the substrate. Such interactions would strengthen binding of both DNA and RNA substrates and might account for the slight improvement in RNA-cleavage activity. This may also explain our observation that the G27 #61 ribozyme efficiently cleaves a modified RNA substrate that has an arabinose sugar at the cleavage site (data not shown).

Co-occurring mutations at positions 188, 190, and 191 in the P5a region also became prominent in the population between G9 and G18. The correlation between these mutations and the co-occurring mutations in the J4/5 and J5/4 internal loop (see Results) suggests a possible interaction between these two regions. It has been proposed that the adenosine-rich bulge in P5a (Figure 2; positions 183–186) interacts with P4 (Flor et al., 1989) by bending at the J5/5a internal loop (Murphy & Cech, 1993), which would place residues G188, U190, and G191 in close proximity to the J4/5 and J5/4 internal loop. Thus, the mutations in P5a may facilitate the contact between residues in the J4/5 region and the IGS.

The evolved ribozymes might compensate for the absent substrate 2′-OH groups by forming new tertiary interactions with the bases, phosphates, and/or sugars of the DNA substrate. It has been shown, for example, that a residue in the J8/7 region of the wild-type ribozyme interacts with the 2′-OH group at position −3(u) of the RNA substrate (Pyle & Cech, 1991; Pyle et al., 1992). However, frequent mutations did not occur in the J8/7 region of the evolved ribozymes, suggesting that new contacts are not made in the vicinity of position −3(t) of the DNA substrate. In addition, specific base contacts seem unlikely, on the basis of the observation that a variety of DNA substrates of different sequence can be cleaved efficiently by the ribozyme, provided the IGS has been changed in a complementary manner to maintain Watson–Crick base pairing (Raillard and Joyce, unpublished results). It is possible that the improvement in DNA binding affinity results from an evolved ability to accommodate the general structural features of an RNA·DNA heteroduplex within the catalytic core of the ribozyme.

The 10^3-fold improvement in k_{cat} over the 27 generations is more difficult to rationalize. k_{cat} reflects all first-order rate constants along the reaction pathway, including those related to P1 docking and substrate cleavage. At least part of the enhancement in k_{cat} thus may be attributed to additional tertiary interactions between P1 and the catalytic core, which would favorably affect the docking rate. The cleavage step of the reaction depends on the appropriate positioning of the 3′-terminal guanosine of the ribozyme for attack on the target phosphoester bond of the substrate. This is accomplished by the formation of a base triple involving the attacking guanosine and the G264·C311 base pair within the P7 region of the ribozyme (Michel et al., 1989). Mutations at positions 312–314 all lie in close proximity to the binding site for the attacking guanosine and may play some role in facilitating the chemical step of the reaction. However, new mutations that became frequent in the population between G18 and G27 in response to the shorter reaction time occurred in peripheral regions, at positions 51/52 and 170. Such mutations may increase first-order reaction rates indirectly through long-range effects or by facilitating folding of the ribozyme into its active conformation.

It is important to note that some mutations may confer no selective advantage with respect to catalysis, but instead enhance the ability of the polymerase enzymes (i.e., reverse transcriptase, T7 RNA polymerase, and *Taq* polymerase) to operate efficiently during the amplification procedure. Future studies, relying on site-directed mutagenesis analysis, will enable us to assess the contribution made by various mutations, in either the conserved core or the peripheral regions, to substrate binding, first-order reaction rates, and ribozyme folding.

Having demonstrated that substrate binding and first-order rate constants can be specifically enhanced by *in vitro* evolution, we are attempting to optimize other catalytic properties of the DNA-cleaving ribozymes, including turnover and substrate specificity. Turnover might be improved by evolving ribozymes that can carry out a site-specific hydrolysis reaction, subsequent to DNA cleavage, which removes the attached 3′ portion of the DNA substrate from the 3′ end of the ribozyme, returning the molecule to its original form. Until now, our evolution procedure has selected against site-specific hydrolysis, as discussed previously (Beaudry & Joyce, 1992). Specificity of the ribozymes for DNA versus RNA substrates might be increased by selecting for DNA cleavage in the presence of RNA that acts as a competitive inhibitor. Our aim is to develop DNA-cleaving ribozymes that have high catalytic efficiency, undergo rapid turnover, and operate in a highly specific manner. Such molecules will contribute to our understanding of the catalytic potential of RNA. In

addition, they may have utility as sequence-specific DNA endonucleases and as therapeutic agents directed against viral pathogens.

ACKNOWLEDGMENT

We thank Bob Niculescu for analysis of the arabinose-containing substrate analogue and Luc Jaeger for helpful comments on the manuscript.

REFERENCES

Bartel, D., & Szostak, J. (1993) *Science 261*, 1411–1418.
Beaudry, A. A., & Joyce, G. F. (1990) *Biochemistry 29*, 6534–6539.
Beaudry, A. A., & Joyce, G. F. (1992) *Science 257*, 635–641.
Bevilacqua, P. C., Kierzek, R., Johnson, K. A., & Turner, D. H. (1992) *Science 258*, 1355–1358.
Butler, E. T., & Chamberlin, M. J. (1982) *J. Biol. Chem. 257*, 5772–5778.
Cech, T. R., Damberger, S. H., & Gutell, R. R. (1994) *Nature Struct. Biol.* (in press).
Davanloo, P., Rosenberg, A. H., Dunn, J. J., & Studier, F. W. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 2035–2039.
Guatelli, J. C., Whitfield, K. M., Kwoh, D. Y., Barringer, K. J., Richman, D. D., & Gingeras, T. R. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1874–1878.
Herschlag, D. (1992) *Biochemistry 31*, 1386–1399.
Herschlag, D., & Cech, T. R. (1990a) *Nature 344*, 405–409.
Herschlag, D., & Cech, T. R. (1990b) *Biochemistry 29*, 10159–10171.
Herschlag, D., Eckstein, F., & Cech, T. R. (1993) *Biochemistry 32*, 8299–8311.
Holmes, D. S., & Quigley, M. (1981) *Anal. Biochem. 114*, 193–197.
Inoue, T., Sullivan, F. X., & Cech, T. R. (1986) *J. Mol. Biol. 189*, 143–165.
Lehman, N., & Joyce, G. F. (1993a) *Nature 361*, 182–185.
Lehman, N., & Joyce, G. F. (1993b) *Curr. Biol. 3*, 723–734.
Michel, F., & Westhof, E. (1990) *J. Mol. Biol. 216*, 585–610.
Michel, F., Hanna, M., Green, R., Bartel, D. P., & Szostak, J. W. (1989) *Nature 342*, 391–395.
Michel, F., Netter, P., Xu, M.-Q., & Shub, D. A. (1990) *Genes Dev. 4*, 777–788.
Milligan, J. F., Groebe, D. R., Witherell, G. W., & Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783–8798.
Murphy, F. L., & Cech, T. R. (1993) *Biochemistry 32*, 5291–5300.
Pan, T., & Uhlenbeck, O. C. (1992) *Biochemistry 31*, 3887–3895.
Pyle, A. M., & Cech, T. R. (1991) *Nature 350*, 628–631.
Pyle, A. M., McSwiggen, J. A., & Cech, T. R. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8187–8191.
Pyle, A. M., Murphy, F. L., & Cech, T. R. (1992) *Nature 358*, 123–128.
Robertson, D. L., & Joyce, G. F. (1990) *Nature 344*, 467–468.
Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., & Arnheim, N. (1985) *Science 230*, 1350–1354.
Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., & Erlich, H. A. (1988) *Science 239*, 487–491.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Wang, J., Downs, W. D., & Cech, T. R. (1993) *Science 260*, 504–508.
Zaug, A. J., & Cech, T. R. (1986) *Science 231*, 470–475.
Zaug, A. J., Been, M. D., & Cech, T. R. (1986) *Nature 324*, 429–433.
Zaug, A. J., Grosshans, C. A., & Cech, T. R. (1988) *Biochemistry 27*, 8924–8931.