In Vitro Selection of Catalytic Polynucleotides

Ronald R. Breaker

Department of Biology, Yale University, New Haven, Connecticut 06520-8103

Received September 26, 1996 (Revised Manuscript Received November 27, 1996)

Contents

Ronald R. Breaker was born in Tigerton, WI, in 1964. After receiving the B.S. degree from the University of Wisconsin—Stevens Point in 1987, he continued his education at Purdue University, where he earned a Ph.D. degree in 1992. His doctoral thesis research was directed by Dr. P. T. Gilham and focused on the chemical and enzymatic synthesis of nucleic acids, nucleic acid analogs, and catalytic RNAs. From 1992 to 1995, he was a Postdoctoral Research Fellow in the laboratory of Dr. G. F. Joyce at The Scripps Research Institute, La Jolla, CA. Here he developed a variety of *in vitro* selection methods to create novel polynucleotides with new or enhanced functions. In 1994, he used *in vitro* selection to create an RNA-cleaving catalytic DNA that displays enzyme-like properties. Dr. Breaker joined the Department of Biology at Yale University in 1995 as an Assistant Professor. His current research efforts are focused primarily on the engineering of new RNA and DNA enzymes.

I. Introduction

The finding that RNA can perform enzyme-like functions in cells overturned the view that biological catalysis is the exclusive realm of protein enzymes. The list of naturally occurring RNA catalysts, begun with the discovery of self-splicing[1] and RNA-cleaving[2] "ribozymes", has continued to expand through the identification of additional self-splicing[3] and also self-cleaving[4−7] ribozymes. In addition, it is likely that the catalytic components of the spliceosome and the ribosome are made of RNA.[8,9] Most of these ribozymes perform indispensable catalytic tasks in modern metabolism, yet natural ribozymes catalyze only a narrow set of chemical transformations that occur primarily at phosphorus centers. For example, self-cleaving ribozymes[10] and both group I and group II self-splicing ribozymes[11] can catalyze phosphoester transfer reactions with RNA. Alternatively, the RNA component of RNase P catalyzes RNA phosphoester cleavage by promoting the nucleophilic attack on phosphorus by water as opposed to an alcohol. These observations make critical the following question: Can RNA be made to catalyze reactions that are fundamentally different than those that are catalyzed by natural ribozymes?

Evidence that the catalytic function of polynucleotides is greater than that seen with natural ribozymes comes from a variety of sources. For example, the discovery of "fortuitous" RNA catalysis, such as the divalent metal-dependent cleavage of tRNAs,[12,13] and the Mn^{2+}-dependent cleavage of short RNA hairpins[14,15] lends support to the notion that, with help from cofactors, RNA can catalyze a broad range of chemical reactions.[16−18] The group I, group II, and RNase P ribozymes, which originally were found to use RNA substrates, can be made to function with DNA substrates.[19−22] In fact, the group II ribozyme has been shown to function as a catalytic component of a transposable element and catalyzes a reverse-splicing reaction to facilitate integration

S0009-2665(96)00008-8 CCC: $28.00

into double-stranded DNA.[23,24] Moreover, a truncated group I ribozyme can be made to accelerate carbon ester hydrolysis[25] and a group II ribozyme has been found to promote phosphoester bond formation with concomitant cleavage of a mixed phosphoanhydride.[22] These and related studies have begun to reveal the true structural and catalytic versatility of RNA.

The notion that RNA has extensive catalytic capabilities is also be consistent with the "RNA World" theory,[26] which holds that ribozymes catalyzed a more varied metabolic chemistry prior to the invention of protein enzymes.[27] In many ways, this theory has been the motivating force behind much of the effort to engineer ribozymes that perhaps represent the reinvention of past biological catalysts.[28] In order for the RNA world to give rise to the protein/DNA world of today, ribozymes alone would have had to catalyze a considerable number of chemical transformations, including reactions that are fundamentally different than phosphoester transfer and phosphoester hydrolysis reactions.

The actual catalytic potential of nucleic acids is being revealed through the engineering of novel ribozymes *in vitro*.[29−40] In particular, the use of combinatorial methods is playing a central role in uncovering ribozyme functions that are likely untapped by modern cellular metabolism. Indeed, "*in vitro* selection" and "*in vitro* evolution" have been used to generate some of the first ribozymes that catalyze reactions other than those that occur at phosphorus centers. Likewise, these methods have been used to isolate the first examples of DNA enzymes, or "deoxyribozymes". These novel catalytic polynucleotides and the methods used to create them are the subject of this review.

II. Ribozyme Engineering

A. Rational Design

Rational methods can be used to reformulate existing ribozymes or to create ribozymes with entirely different catalytic and kinetic features.[34,39] One of the simplest rational design procedures consists of splitting a contiguous polynucleotide chain to create a fragmented ribozyme. The fragments can then be reassembled to form an active multicomponent ribozyme. With similar ease, simple secondary structure elements can be created by following the rules of Watson−Crick base complementation. These strategies are used routinely to create ribozymes that display multiple turnover kinetics or that have altered substrate specificity. Significant progress is now being made in understanding the more complex interactions that guide the tertiary folding of catalytic RNAs.[39,41] However, as with the rational design of proteins,[42] control over the precise positioning of RNA functional groups in three dimensions remains beyond the reach of current rational design techniques. Consequently, the *de novo* design of ribozymes that perform new catalytic functions will remain a difficult task until these properties are understood in greater detail.

B. Selection and Evolution *in Vitro*

An alternative (or complement) to rational design is the use of iterative selection methods that isolate catalytic molecules from mutagenized or random-sequence pools of RNA or DNA. This approach relies on the probability that a given pool of random-sequence molecules will include individuals that can perform the function of interest. For example, the conserved catalytic core (13 nucleotides) of hammerhead self-cleaving ribozymes[10] is expected to occur with a frequency of one in every 67 million random-sequence RNAs. Therefore, in a pool of 10^{15} molecules, approximately 15 million will carry the hammerhead catalytic core, and some are expected to efficiently catalyze RNA cleavage. Engineering new catalysts, then, has been reduced to the synthesis of mutagenized-ribozyme pools or random-sequence pools of nucleic acids.

There is little doubt that a variety of catalytic molecules will be present in a diverse pool of nucleic acids. The more challenging problem then is to identify or isolate those exceedingly rare molecules that have the desired catalytic properties. This problem can be solved if molecules that perform the desired reaction can be separated from the remainder of the pool and subsequently amplified using any of a number of methods for the replication of RNA and DNA. This process of selective amplification or "*in vitro* selection" can be used in an iterative fashion to isolate rare molecules from large pools of random-sequence RNA or DNA. If a significant number of mutations are introduced during selective amplification, then the process is typically termed "*in vitro* evolution" to reflect the similarities between this process and Darwinian evolution.

1. Early in Vitro Evolution Experiments

The directed evolution of nucleic acids dates from the 1960s when Spiegelman and co-workers observed that genomic RNA of bacteriophage Qβ undergoes significant changes upon repeated *in vitro* replication with Qβ replicase.[43] These changes occur during serial transfer amplification, in which a portion of products of the previous RNA amplification reaction are used to initiate the next reaction. After repeated amplification by serial transfer, the dominant templates are truncated RNAs that have deleted nearly 90% of their original nucleotides, thereby eliminating unnecessary sequences and accelerating the rate of replication. Similarly, new template RNAs were created that efficiently amplify under suboptimal conditions or in the presence of ethidium bromide.[44] Presumably, these new RNA templates originate as a result of replication errors made by Qβ replicase during polymerization. From these beginnings, the combinatorial approach for designing new functional nucleic acids has grown into a diverse field of study that uses a variety of methods to direct the evolution of nucleic acids with new or improved functions.

2. Selection and Amplification

A series of critical decisions becomes imminent during the planning and preparation of the pool of nucleic acids to be initially screened. Significant

questions include the following: Is it best to begin with a pool composed of mutagenized versions of a known ribozyme, or begin with random-sequence molecules? If a random-sequence pool is chosen, how many nucleotides should be made random? How many different sequence variants should be screened and should mutations be introduced during the selection process? *In vitro* selection experiments conducted to date have encompassed a variety of goals and selection strategies. In certain experiments, the goal has been to isolate RNAs with novel catalytic function while others have focused the selective effect on a single aspect of the catalytic process (*e.g.* a single step in the kinetic process). The answers to these questions then will differ depending on the desired outcome of the selection. For example, if the goal is to examine the sequence requirements for catalytic activity of an existing ribozyme, then partial or total randomization of the catalytic core within the context of a largely unaltered ribozyme is prudent. However, if a ribozyme with entirely new catalytic function is sought, then perhaps a large pool of random-sequence molecules should be screened. A more detailed description of these issues has been published elsewhere.[34]

Combinatorial libraries of RNA or DNA molecules can be produced in a number of different ways. Pools made of short molecules (<100 nucleotides) can be chemically synthesized. Pools consisting of longer molecules (>100 nucleotides) can more efficiently be prepared by introducing mutations into a preexisting template[45] or can be created by splicing several shorter molecules via enzymatic ligation[46] or by the "overlap extension" method.[47] The pool can be constructed to include only random-sequence domains, or each member of the pool can be "preengineered" to carry specific structural elements that, for example, create substrate binding sites or even crude active sites. A successful strategy employing preengineering makes use of RNA domains that are known to bind specific ligands. This "biased pool" approach combines mutagenized RNA aptamer[48] domains with random-sequence domains to increase the probability of finding molecules that utilize a particular ligand as substrate.

The optimal length for a random-sequence domain remains an unsettled question. There is expected to be a combinatorial advantage when using randomized domains of increasing length. This advantage can be represented by considering the formula $(x - y) + 1$, where x is the number of nucleotides in the random region (N_x) and y is the number of nucleotides in the domain of interest. For example, a 30-nucleotide random-sequence domain can only represent one distinct 30-nucleotide sequence per molecule, but 71 contiguous 30-nucleotide domains of unique sequence can be represented by a single molecule that carries a 100-nucleotide random region. Although the probability for inhibition of the desired function of a 30-mer by another sequence domain within N_{100} is also greater, the overall effect is expected to be beneficial.

In order to screen catalytic nucleic acids, one must have a way of identifying those molecules that facilitate the desired chemical transformation from

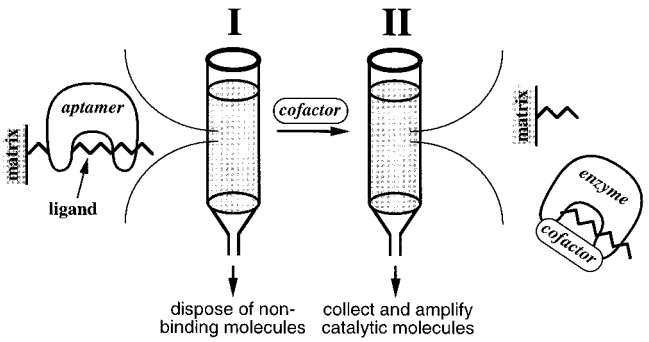


Figure 1. Isolation of RNA and DNA enzymes by "catalytic elution". (I) An immobilized compound acts as a ligand for RNA or DNA aptamers that are subsequently screened (II) for those that catalyze the cleavage of the ligand under permissive incubation conditions.

those that are inactive. Perhaps the most straightforward manner in which to achieve this goal is to identify or isolate nucleic acids that undergo self-modification. However, to isolate catalytic molecules that use non-nucleic acid substrates, or that efficiently function as true enzymes and display multiple turn-over kinetics, it may be more appropriate to employ a selection method that has no requirement for self-modification. Currently, many of the catalytic polynucleotides that have been created by *in vitro* selection carry out reactions with nucleic acid substrates as an artifact of this reasoning.

All selection methods that have been successfully implemented to date can be grouped into the following three categories. *(i) Distinction by self-modification*. A variety of selection protocols exploit self-modification to recover active variants by physical separation or by molecular recognition. Among the most commonly used methods include separation of oligonucleotides by size (differential gel mobility). This method allows one to isolate molecules that promote phosphoester cleavage or ligation reactions. In addition, the interaction between avidin and biotin is used routinely for *in vitro* selection. The equilibrium constant for avidin and its ligand is $\sim 10^{15}$ M^{-1}, making this an efficient method for solid-phase immobilization of nucleic acid pools or for isolating exceedingly rare sequences that acquire a biotin "tag". *(ii) Binding to transition-state analogues*. With analogy to the isolation of "catalytic antibodies", this approach uses transition-state analogues to obtain RNA and DNA aptamers that indirectly yield new catalytic polynucleotides. *(iii) Catalytic elution*. This approach seeks to immobilize nucleic acids on a solid support by covalent means or via the affinity of a catalyst for its substrate and then elute the active molecules by adjusting the solvent conditions to the permissive reaction conditions (Figure 1). A number of alternative selection methods can also be envisaged, but have yet to be widely employed for the selection of catalytic polynucleotides. For example, various tagging and imaging methods like those described by Lam and co-workers[49] could be used to isolate catalysts that are immobilized on beads. Alternatively, an *in vivo* selection might be developed that links the function of a catalytic polynucleotide to the fate of that organism.[50] This selection might be used to isolate ribozymes that are highly active *in vivo*.

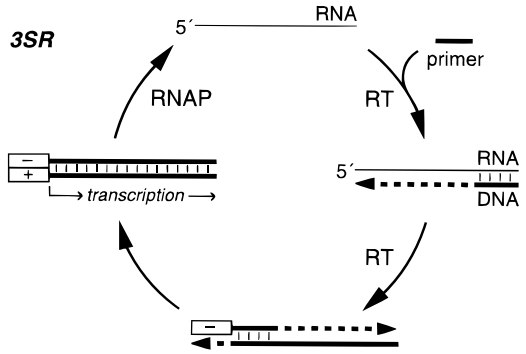


Figure 2. Amplification of RNA and DNA can occur simultaneously with 3SR under isothermal conditions. Template RNAs (thin line) are copied by reverse transcriptase (RT) and the resulting cDNA (thick line) is made double stranded by the same enzyme, using a primer that carries a specific RNA polymerase (RNAP) transcription control element (−). The double-stranded transcription control element (− and +) is used by RNAP to regenerate single-stranded RNA copies.

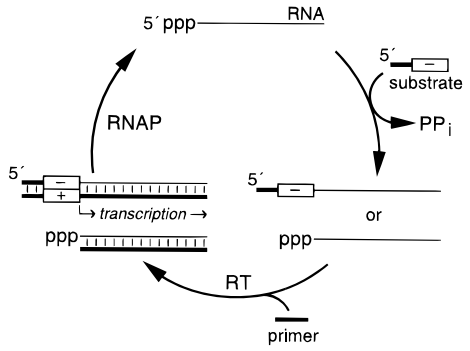


Figure 3. Strategy for the continuous evolution of catalytic RNAs. Pool RNAs (thin line) are incubated in the presence of RT, RNAP, and a substrate oligonucleotide that encodes an RNAP promoter (−). Those RNAs that catalyze the formation of a 3′,5′-phosphodiester link to the substrate acquire the promoter and, when copied by RT, become functional templates for RNAP. The subsequent transcription events are selective for DNA templates (thick line) that are derived from active ribozymes and produce new ribozymes that are immediately available for further selective-amplification. PP_i represents inorganic pyrophosphate.

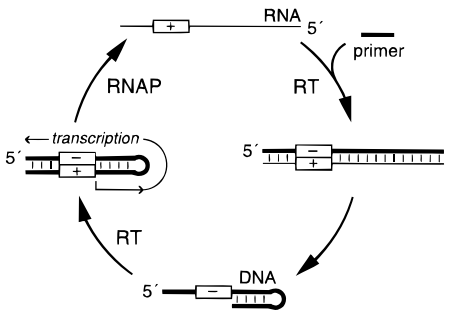


Figure 4. Amplification mechanism of RNA Z. Unlike typical transcription products, RNA Z encodes an RNAP promoter (+) that is copied by RT to produce a single-stranded cDNA. This cDNA refolds to form an incomplete hairpin that is further extended by RT, thereby producing a double-stranded promoter element (+ and −) that can be recognized by RNAP. Transcription proceeds around the hairpin loop, yielding copies of the original RNA Z template. Unlike PCR or 3SR, this amplification process requires only one oligonucleotide primer.

Usually, the number of molecules recovered during each selection cycle is vanishingly small. Almost without exception, experimenters make use of sensitive amplification methods[51] such as the polymerase chain reaction (PCR)[52] or self-sustained sequence replication (3SR)[53] (Figure 2) to amplify the recovered molecules. These methods take advantage of oligonucleotides that can recognize template molecules with exquisite sensitivity and specificity. A bound oligomer can then act as a primer for the enzymatic synthesis of a complementary copy of the selected molecule. Repetitive copying yields a new pool of molecules that are derived from the selected molecules of the previous pool, and new sequence variants can be created by employing mutagenic versions of PCR.[54,55]

3. Selfish Nucleic Acids

Most researchers typically seek a specific outcome for an *in vitro* selection experiment and design the selective-amplification protocol to give the desired products. However, experimenters frequently find that they have underestimated the potential for nucleic acids to adapt to a given set of selection constraints or that alternative solutions to a particular selection challenge are equally adept. Joyce[51] has described a series of selfish RNAs or "minimonsters" that were encountered in the *in vitro* evolution of group I ribozymes. These results are usually troublesome, but in some cases can reveal much about the functional potential of nucleic acids. For example, Breaker and Joyce[56,57] describe the origin and replication mechanism of "RNA Z", a selfish molecule that efficiently amplifies in 3SR reactions. This molecule arose from an *in vitro* evolution experiment that was designed to isolate ribozymes that catalyzed the ligation of an oligonucleotide to their 5′-terminus (Figure 3). Instead, the molecules that were isolated do not perform the desired reaction and can amplify in a 3SR-like reaction with only one (Figure 4) or even no (Figure 5) primer oligonucleotides. Sometimes these selfish molecules can be exploited for alternative uses. For example, randomized molecules that replicate like RNA Z were subsequently used to carry out the continuous evolution of RNA polymerase control elements.[58]

III. Natural Ribozymes

A. Group I Ribozyme

1. DNA Cleavage

RNA splicing is a fundamental part of RNA processing in many organisms. The pre-rRNA of *Tetrahymena thermophila* was found to undergo "self-splicing" *in vitro* without the need for protein catalysts.[59] This and other "group I" self-splicing ribozymes promote two RNA phosphoester transfer reactions (Figure 6) that result in the removal on an intervening sequence and the splicing of adjacent RNA domains. The initial transesterification reaction that is promoted by this ribozyme uses the 3′-hydroxyl of guanosine or one of its 5′-phosphorylated derivatives as the nucleophile for an S_N2 attack on the phosphate of the target internucleotide linkage. The second transesterification uses the newly formed

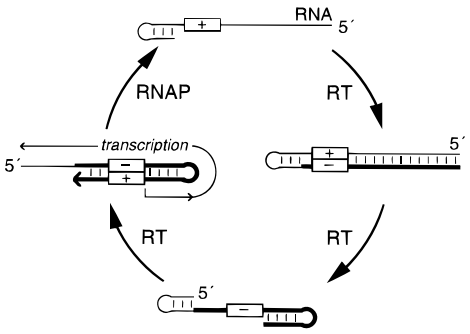


Figure 5. Amplification mechanism of "primerless" RNA Z. Like RNA Z, primerless RNA Z encodes an RNAP promoter. The RNA self-primes cDNA synthesis by RT, thereby eliminating the need for separate primer oligonucleotides. The RNase H-like activity of RT partially degrades the resulting RNA−DNA chimera and ultimately yields a mixed RNA−DNA template that is transcribed by RNAP. The RNA primer domain is regenerated during each amplification cycle by RNAP, which can transcribe through the junction between RNA−DNA to produce copies of the original template RNA.

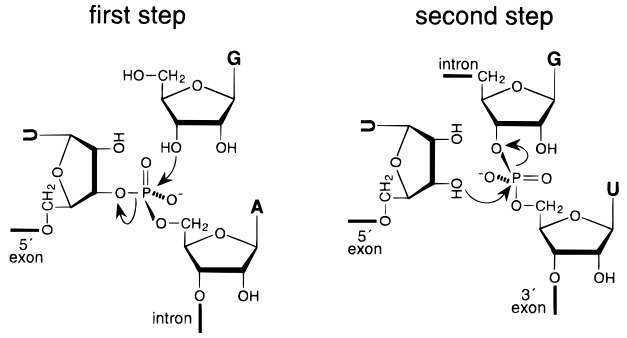


Figure 6. The first and second steps of RNA splicing by group I ribozymes.

3′-hydroxyl of the 5′-exon as the nucleophile in the subsequent attack at the second splice-site junction. Each reaction results in inversion at the chiral phosphate center.[60,61]

Joyce[62] developed an *in vitro* evolution procedure (Figure 7) that exploits the phosphoester transfer activity of a shortened version of the *Tetrahymena* group I intron to create a series of ribozyme variants that efficiently cleave DNA. Briefly, the 5′-terminal nucleotides function as an internal guide sequence, binding and aligning a substrate molecule for site-specific cleavage by the ribozyme. In this case, the nucleophile for the transesterification reaction is supplied by the 3′-terminal guanidyl moiety (G414) of the truncated ribozyme. Upon substrate scission, the 5′-cleavage fragment is released, while the 3′-fragment remains covalently attached to the ribozyme via a 3′,5′-phosphodiester linkage, thereby distinguishing active from inactive ribozyme variants. This "tag" sequence can then be used to specifically amplify (3SR) those ribozymes that have completed the cleavage reaction.

The *Tetrahymena* group I ribozyme can be made to cleave DNA, via a phosphoester transfer reaction, under reaction conditions that include elevated Mg^{2+} concentrations or at elevated reaction temperatures.[19,20] Selection for ribozymes with improved DNA-cleavage activity was first tested[19] using a pool of six deletion variants and was subsequently used[63]

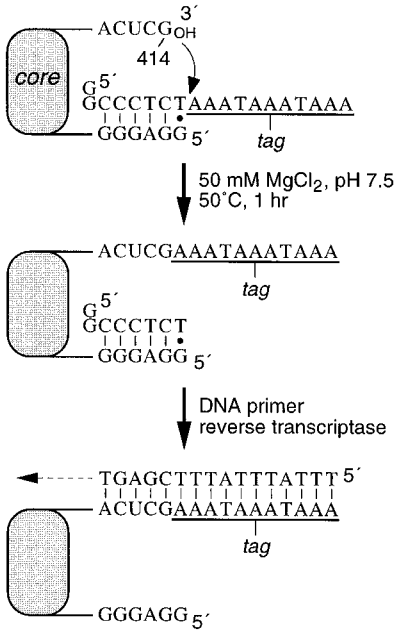


Figure 7. *In vitro* selection of DNA-cleaving ribozymes via self-modification. Shaded area represents the catalytic core of the group I ribozyme.

to screen a pool of $>10^{13}$ ribozyme sequence variants. In the latter study, Beaudry and Joyce[63] mutagenized[45] 140 nucleotides of the catalytic core of the ribozyme (Figure 8) to yield an RNA pool that includes all possible sequence variants with five or fewer mutations. RNAs that acquire a DNA tag sequence at their 3′-terminus are selectively amplified by 3SR, and the resulting cDNAs are amplified by PCR to generate templates for *in vitro* transcription. These amplification processes also generate significant numbers of mutations, some that create advantageous changes that impact ribozyme structure and function. After nine rounds of *in vitro* evolution, the resulting population was ~100-fold more active than the wild-type ribozyme.

Beaudry and Joyce chose only to partially randomize the core of the ribozyme to a degeneracy (*d*) equal to 0.05. For each molecule, there is a probability of 0.95 for the base identity of an individual nucleotide within the mutagenized region to remain wild-type, while there is a 0.05 probability for the wild-type base to change to one of the three remaining bases. This level of mutagenesis allows all possible combinations of ribozyme variants that carry from 0 to 5 mutations to be screened with a pool of 10^{13} molecules.[34,63] Since the group I ribozyme already displays weak DNA-cleaving activity, this partial randomization strategy has a greater chance of yielding efficient DNA-cleaving ribozymes, as opposed to screening a pool comprised of random-sequence molecules. This partial randomization strategy is commonly used when improving the catalytic function of existing ribozymes.

Tsang and Joyce[64] have further improved the DNA-cleaving activity of group I-derived ribozymes by conducting an additional 18 rounds of selection, isolating ribozymes that achieve a 10^5-fold improvement in DNA-cleavage activity over wild type. In this same study, *in vitro* evolution was used to target individual kinetic parameters of the ribozyme for

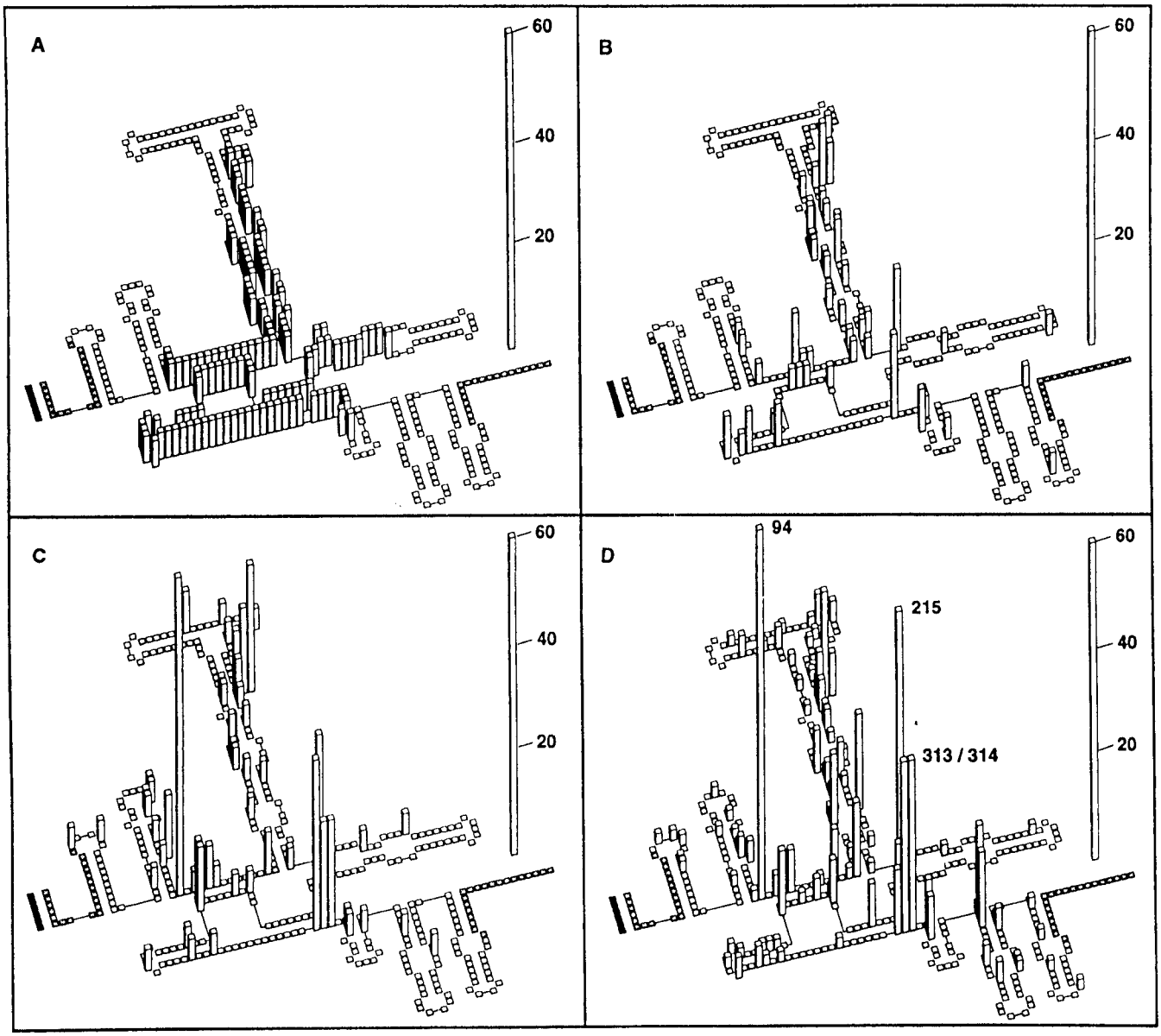


Figure 8. A sequence map of the *in vitro* evolution of the *Tetrahymena* group I ribozyme from 0, 2, 4, and 8 rounds (A−D, respectively) of selection. Mutations (represented by box height) overlay the secondary structure of the ribozyme to identify the location and frequency of particular base changes relative to other mutations. Box heights and positions in A represent the 5% degeneracy that was introduced into the initial RNA pool. (Reprinted from ref 63. Copyright 1992 American Association for the Advancement of Science).

improvement. The dissociation constant (K_d) for the DNA substrate with wild-type ribozyme is ~30 μM. This constant fell to ~10 μM for individuals that were isolated after 10 rounds of selection using 10 μM substrate. An additional nine rounds of selection using 0.2 μM resulted in a rapid change in the measured K_d for the population as a whole, with individual ribozymes at round 18 displaying K_d values of <0.2 μM.

The DNA-cleaving ribozymes show a broadening of substrate specificity and can also cleave RNA with faster rates. In a continuation of this study, Tsang and Joyce[65] performed an additional 36 rounds of selection in the presence of an RNA "product", which acts as an inhibitor of ribozymes that cannot distinguish between RNA substrate and DNA substrate. After a total of 63 rounds of *in vitro* evolution under increasingly stringent conditions, variant ribozymes carry ~28 mutations relative to wild-type and can cleave DNA with an overall rate enhancement estimated to be in excess of 10^{12}-fold over the uncatalyzed rate of DNA hydrolysis.[66] Moreover, variant ribozymes now can discriminate against the corresponding RNA substrate, likely due to a greater affinity for DNA versus RNA substrates.

2. Altered Metal Ion Dependence

An identical *in vitro* selection method was used by Lehman and Joyce[67,68] to isolate group I ribozyme variants that cleave RNA in the presence of Ca^{2+}, an activity that the natural ribozyme displays only in the presence of either Mg^{2+} or Mn^{2+}. Analysis of individual molecules that comprise the populations throughout 12 rounds of selection reveals a significant correlation between catalytic activity and the number of mutations that were acquired relative to wild-type.[68] In fact, the number and locations of mutations that are found in individuals of the later

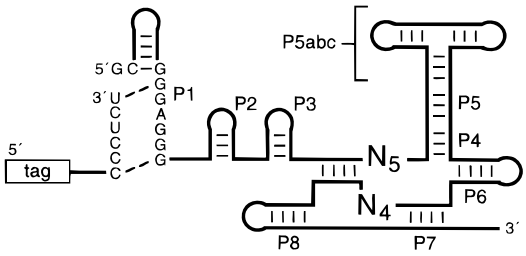


Figure 9. Schematic representation of the group I ribozyme from *Tetrahymena* that has been altered to perform a reversal of the first step of splicing. The 5′-terminal guanosine of the ribozyme is liberated in a transesterification reaction that joins the ribozyme and substrate via a 3′,5′-phosphodiester linkage. N_4 and N_5 represent random-sequence domains of four and five nucleotides respectively.

selection rounds demonstrates the advantage gained by the introduction of mutations during the selection process.

Lehman and Joyce[68] also vividly demonstrate the versatility in handling and analyzing the products of selection experiments. Populations of ribozymes at each stage of the selective-amplification process can be put into storage and accessed at any time, to begin new branches of *in vitro* selection or to rigorously analyze the resulting catalysts. For example, sequence analysis of individual molecules that were isolated at different stages of the Ca^{2+} selection made possible a detailed analysis of the dynamics of *in vitro* evolution. Lehman and Joyce found that advantageous mutations in their molecules fell into three categories, logarithmic, linear, and transient, according to their rise and fall during the selection process. These experiments showcase the usefulness of combinatorial methods to rapidly explore the principles of evolution at the molecular level.

3. RNA Ligation

Szostak and co-workers have also examined the structural and catalytic features of the group I ribozyme, in this case by exploiting the ribozyme's cleavage/ligation activity that is analogous to the reversal of the first step of splicing (Figure 9). For example, Green *et al.*[69] randomized nine nucleotides of the *Tetrahymena* group I ribozyme core and then used three rounds of selection to isolate active ribozymes from the pool of 4^9 different sequence variants. Active ribozymes ligate an RNA substrate to their 5′-terminus with concomitant release of guanosine, and thereby acquire a tag sequence. The tag makes possible the selective amplification by PCR of active ribozymes and the subsequent synthesis of new ribozymes via *in vitro* transcription with T7 RNA polymerase. In this particular study, the wild-type sequence was most commonly encountered upon examination of the selected pool, while individual one- and two-base sequence variants were nearly as active. In similar studies, Green and Szostak have examined the "hinge region" of the sunY group I ribozyme,[70] and have further optimized a deletion mutant of this same ribozyme.[71]

Williams *et al.*[72] have used a different selection strategy that uses a bimolecular group I ribozyme arrangement in which the original unimolecular arrangement is uncoupled in the loop of stem P6. In addition, the RNA strand that comprises the 5′-region of the ribozyme also includes a complete 5′-splice site. Ribozyme activity then results in the covalent coupling of the two domains and facilitates selective reverse transcription and subsequent PCR amplification (RT-PCR) of functional ribozymes. The original pool ($\sim 10^{12}$ variants) was constructed by replacing the 69-nucleotide P5abc domain (Figure 9) with either 20 or 40 random-sequence nucleotides. After six rounds of selection, the activity of the RNA pools nearly matched that of wild-type and individual RNAs from the two pools could be assigned to three different sequence families that differed widely from each other and from the wild-type domain. This selection experiment demonstrates that essential structural subdomains of large ribozymes can be entirely replaced by novel RNA structures.

B. RNase P

In vitro selection for catalytic function is not restricted to selection of the enzyme alone, but can also be applied to the substrates of certain ribozymes. RNase P is a ribonucleoprotein that catalyzes the hydrolysis of precursor tRNAs as well as other RNAs. The RNA component of the complex has been shown to function in the absence of the protein component[2] and acts as a Mg^{2+}-dependent ribozyme. Ribozymes that process RNA or DNA substrates typically bind to their corresponding substrates by Watson−Crick base-pair interactions. Therefore, changing the substrate specificity can be achieved by simply redesigning the substrate binding region of the ribozyme so as to maintain base complementarity with the new substrate (see section C, below). RNase P, however, mostly relies on tertiary contacts to bind substrate RNAs and the guidelines for such interactions are less well defined. Despite these complications, RNase P can be made to cleave any RNA, if the target RNA is complexed with an "external guide sequence" (EGS) that mimics the structure of a natural pre-tRNA substrate.[73] However, cleavage rates for unnatural substrate complexes can be considerably slower than for natural RNA substrates.

Yuan and Altman[74] have employed *in vitro* selection to aid in the design of EGS RNAs that, in conjunction with a target oligoribonucleotide, serve as an improved substrate for RNase P. Selection was achieved by isolating efficiently cleaved substrates from a pool of randomized tRNA-like molecules using polyacrylamide gel electrophoresis. Liu and Altman[75] also employed this strategy of "differential gel mobility" to isolate novel RNase P substrates from random-sequence pools. The RNA substrates derived in this selection show little resemblance to tRNA structures, a finding that is consistent with the fact that RNase P can cleave biological substrates other than tRNAs. Interestingly, selection carried out in the presence of the protein component of RNase P gave a greater variety of substrates compared to the selection conducted with the RNA component only, indicating that the protein component might act to expand the substrate specificity of the ribozyme.

Further evidence for a broad substrate specificity for the ribozyme of RNase P was reported by Pan.[76] Here, a somewhat different selection method to seek

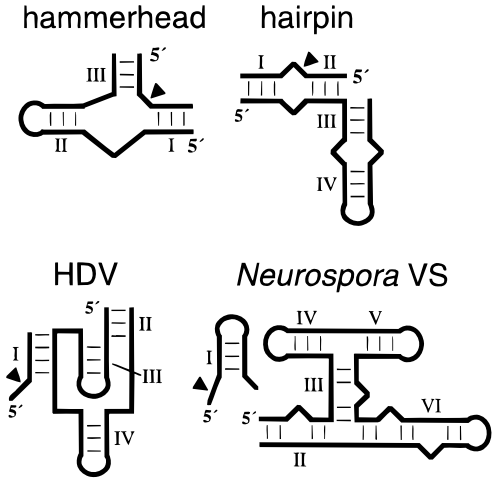


Figure 10. The four natural self-cleaving ribozyme motifs.

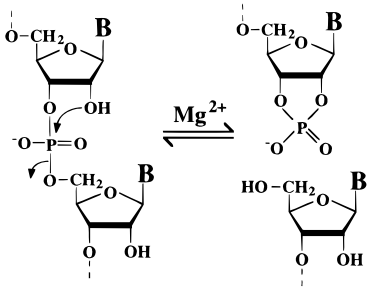


Figure 11. RNA cleavage by the self-cleaving ribozymes.

novel substrates for RNase P (see Figure 18 for a similar selection scheme). A pool of circular RNAs, each carrying a total of 70 random-sequence nucleotides, was incubated with the RNA component of *Bacillus subtilis* RNase P. Cleaved RNAs were isolated by gel electrophoresis, and new RNA substrates were subsequently prepared for the next round of selection. Two classes of substrate RNAs were identified, both of which have little or no similarity to natural pre-tRNA substrates.

C. Hammerhead Self-Cleaving Ribozyme

The hammerhead ribozyme motif[77] is one of four structural classes of natural "self-cleaving" RNA structures[10] (Figure 10) that can accelerate the divalent metal-dependent cleavage of RNA phosphoesters via a cyclizing transesterification mechanism. Different structural classes are defined by characteristic secondary structure folds and by regions of conserved nucleotides. Each motif can fold into a defined tertiary structure that properly orients the phosphate of the labile internucleotide linkage for "in-line" nucleophilic attack by the adjacent 2′-hydroxyl (Figure 11).

Site-directed mutational analysis and sequence data from natural examples of hammerhead ribozymes have been used to derive a consensus sequence and secondary structure (Figure 12). The sequences of the RNA strands that comprise helices I−III can be altered as desired, but must retain complementary base pairing to maintain maximal catalytic activity. In addition, each helix can either exist in "open" (*e.g.* helix I and III) or "closed" (*e.g.* helix II) states, allowing a total of eight different

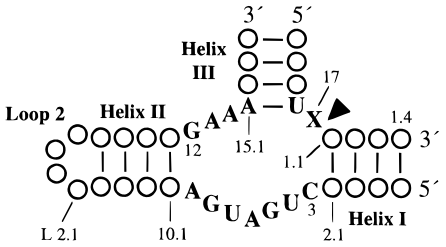


Figure 12. Secondary structure and conserved core of the hammerhead ribozyme.[129] X represents A, U, or C. The sequences and lengths of helices I, II, and III can vary.

arrangements that involve one-, two-, or three-stranded hammerhead complexes.[34] Perhaps the most advantageous arrangement for hammerheads is the bimolecular structure that contains only loop 2 (Figure 12).[78] This results in a complex where most of the conserved residues reside in a ∼30-nucleotide "enzyme" domain and allows the greatest versatility in targeting different "substrate" domains.

Under optimal conditions, various hammerhead ribozymes can attain catalytic rates of >1 min^{-1}, but the actual catalytic rate is slower under ionic and temperature conditions that approximate physiological values.[79] The therapeutic and antiviral potential of the hammerhead and of the other self-cleaving ribozymes has fostered considerable interest in tailor-making ribozyme variants that operate with rapid catalytic turnover. As a result, several attempts[80−82] have been made to assess the catalytic fitness of natural hammerhead ribozymes and to optimize their catalytic function. Although most selections involve the catalytic centers of ribozymes, Lieber and Strauss have reported a method to generate ribozymes that have pairing arms that are selected to efficiently target mRNAs.[83]

A significant problem encountered when studying the hammerhead and other self-cleaving ribozymes comes from the fact that these ribozymes are active as catalysts during preparation by *in vitro* transcription. However, preparation of a mutagenized pool of self-cleaving ribozyme transcripts cannot easily be prepared without losing many of the best catalysts to self-cleavage during transcription. Most *in vitro* selection experiments with the hammerhead and other self-cleaving ribozymes have addressed this issue in two ways. Most frequently, investigators have chosen to allow self-cleavage to occur during transcription and then isolate the active ribozymes for iteration of the selection cycle. A second option is to prepare separate ribozyme and substrate domains and then combine them at a later stage of the selection cycle.

Nakamaye and Eckstein[80] reported the *in vitro* selection of hammerhead ribozymes, from a starting pool of 64 sequence variants that were originally randomized at positions 7, 10.1, and 11.1 (Figure 12). The authors employed a gel-mobility selection that identified processed ribozyme transcripts that had cleaved during transcription, or soon thereafter. This selection yielded hammerhead ribozymes that carried the strictly conserved core sequences, suggesting that the natural ribozyme sequences at these positions are optimal. Two additional pools of RNA that involved complete sequence randomization at either 10 (nucle-

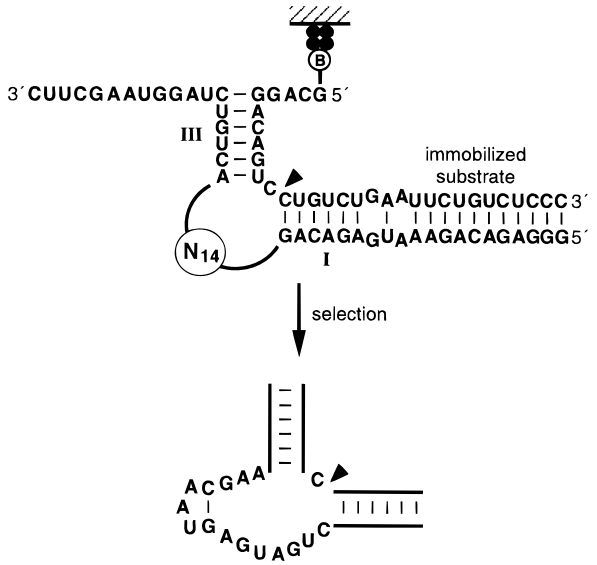


Figure 13. Selection of hammerhead-like self-cleaving ribozymes from a biased RNA pool.

otides 3−9 and 12−14, Figure 12) or 12 (as previously plus nucleotides 15.1 and 16.1) were also screened. Unfortunately, selections from these larger pools yielded shortened RNA transcripts that did not self-cleave, but simply had a gel mobility that matched the mobility of cleaved hammerhead ribozymes.

Long and Uhlenbeck[81] have isolated active hammerhead structural variants, in which stem II has been replaced with a simple loop structure. *In vitro* selection for these truncated ribozymes began with a pool of hammerhead-like RNAs that carried six random-sequence nucleotides in place of stem−loop II, producing a total of 4096 different RNAs. Again, a differential gel mobility was employed to isolate RNAs that cleaved during transcription. A variety of active hammerhead variants were obtained with k_{cat} values that approach ~0.1 min^{-1}, about 10-fold slower than the analogous hammerhead with an intact stem II.

Ishizaka *et al.*[82] reported the *in vitro* selection of hammerhead-like self-cleaving RNAs from a starting pool of RNAs that carried 14 random-sequence nucleotides. For this selection, a bimolecular arrangement consisting of separate "catalyst" and "substrate" domains was constructed (Figure 13). The random-sequence region of the catalytic domain was flanked by regions of defined sequence that served to bind the substrate domain and functioned as primer binding sites for RT-PCR amplification. Most individual catalysts in the pool are expected to base pair both upstream and downstream from the intended cleavage site and position 14 random-sequence nucleotides between the two substrate-pairing domains, thereby creating helices that are analogous to stems I and III of the hammerhead ribozyme. The nucleotide immediately 5′ to the intended cleavage site internucleotide linkage, analogous to nucleotide 17 in the hammerhead, is left unpaired. Each substrate molecule is immobilized on an agarose matrix via a streptavidin−biotin interaction. Individual catalysts are then selected for the ability to bind the substrate molecule in the absence of a divalent metal ion cofactor and for the ability to cleave the substrate to

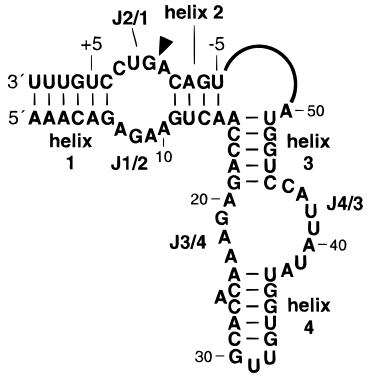


Figure 14. The secondary structure of the hairpin ribozyme derived from the satellite of tobacco ring spot virus. For *in vitro* selection, the hairpin is engineered with an oligocytidine linker to create a unimolecular construct.

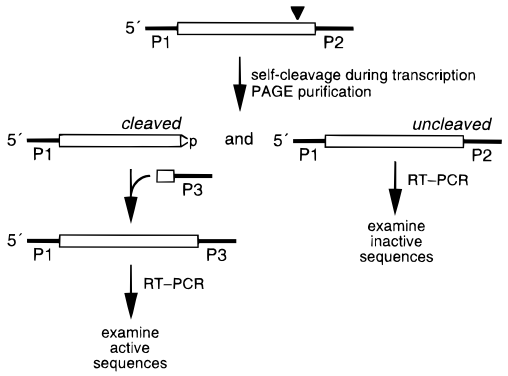


Figure 15. Selection scheme for the isolation of hairpin ribozyme variants that are active for both the forward and reverse transesterification reactions.

free themselves from the matrix by cleaving the substrate upon addition of Mg^{2+} (Figure 1).

The prevailing sequences that were obtained after seven rounds of selection conformed to the consensus hammerhead core (Figure 12). Interestingly, each sequence acquired a single additional nucleotide in the core region, thereby producing a hammerhead motif in which stem−loop II is replaced by five nucleotides instead of the intended four nucleotides. Although truncated hammerheads that have four nucleotides in place of stem II are known to be active,[84,85] presumably certain five-nucleotide loop sequences are more active and most likely a fortuitous nucleotide insertion occurred during the selection process to produce such sequence variants.

D. Hairpin Self-Cleaving Ribozyme

The hairpin ribozyme is another distinct structural class for Mg^{2+}-dependent self-cleaving RNAs. The secondary structure (Figure 14)[86,87] consists of four helical domains (helices 1−4) that are linked by internal bulges. Berzal-Herranz *et al.*[87] have pioneered *in vitro* selection of the hairpin ribozyme using gel-mobility shift coupled with cleavage and ligation selection steps to create an "artificial phylogeny" of ribozymes and to improve the catalytic rate. The selection scheme (Figure 15) again relies on the difference in gel mobility between hairpin ribozyme variants that have cleaved and those that remain uncleaved during preparation by transcription. In addition, the selection scheme exploits the fact that

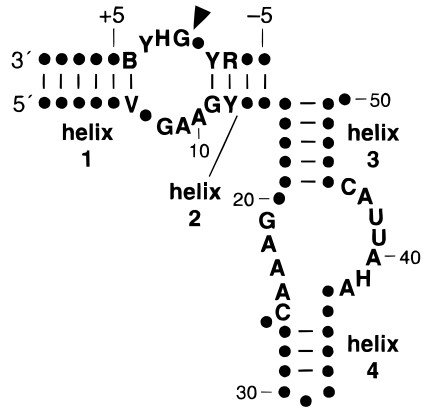


Figure 16. The consensus hairpin ribozyme. Watson–Crick pairing is indicated by dashes, a dot (●) represents any nucleotide; **Y** is U or C; **R** is A or G; **B** is U, C or G; **H** is U, C or A; **V** is A, C or G. A change at position 39 (arrow) from U to C increases k_{cat} by 4-fold.

the hairpin efficiently catalyzes the reverse reaction and is therefore able to append a new RNA sequence to its newly cleaved 3′-terminus. Specifically, a pool of hairpin ribozyme variants, each flanked by primer binding sites (P1 and P2, Figure 15) are produced by *in vitro* transcription, during which the active variants may self cleave. Cleaved ribozymes are separated from uncleaved ribozymes by PAGE and then incubated with an excess of an oligoribonucleotide that is analogous to the 3′-cleavage product, and that carries a new primer binding sequence. Hairpin RNA variants that promote both the cleavage and ligation reactions to retain P1 and to swap P2 for P3 can be selectively amplified by RT-PCR and analyzed by cloning and sequencing. The uncleaved RNAs can also be amplified by RT-PCR, using primers that correspond to P1 and P2, and subsequently examined to confirm the identity of mutations that prevent ribozyme-promoted RNA cleavage.

This selection scheme has been used to probe a number of structural and functional features of the hairpin ribozyme.[87–91] For example, the four nucleotides (+3 to −1, Figure X) A↓GUC that flank the cleavage site were randomized to create a pool of 256 sequence variants of the "J2/1" substrate domain.[87,88] Active variants were then recovered by repeated rounds of selection, revealing a consensus sequence of N↓GHY (H = U, C, or A; Y = C or U) and demonstrating that the G at position +1 is absolutely required for catalysis by the hairpin ribozyme. Similar selections were used to examine the remaining junction regions of the ribozyme,[90] to further define a consensus sequence for cleavable substrates,[89] and to optimize the hairpin ribozyme for cleavage of a new RNA target sequence.[91] The primary and secondary consensus structures, based largely on the *in vitro* selection data, have been constructed without the need to rely on traditional methods of phylogenetic analysis (Figure 16). In addition, several selections have produced ribozymes with a U to C change at position 39 that result in a 4-fold increase in the catalytic rate.

E. HDV Self-Cleaving Ribozyme

The sequence and structural requirements of ribozymes can be rapidly examined through the use of

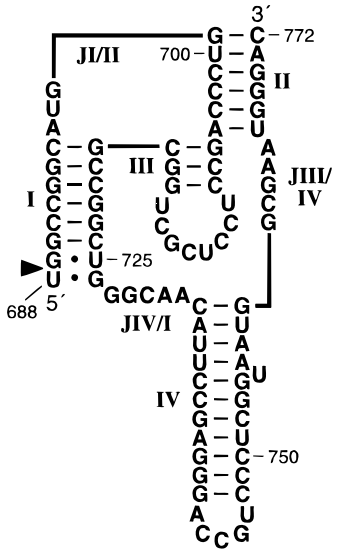


Figure 17. Secondary structure model of the HDV ribozyme as first proposed by Perrotta and Been.[130] The nomenclature follows that used by Tanner *et al.*[131]

in vitro selection. For example, Kawakami *et al.*[92] have investigated the structural and catalytic features of the self-cleaving ribozyme from the genomic strand of the hepatitis δ virus (HDV) (Figure 17) by creating an artificial phylogeny consisting of inactive ribozyme variants. A unique pool of ribozyme variants was constructed that contained all 24 single base changes in the loop that connects helix III (nucleotides 708–715). RNAs that did not cleave during transcription, or that did not cleave after additional incubation with 50 mM $MgCl_2$, were amplified by RT-PCR after treatment by DNase to remove the transcription-template DNA. After three rounds of selection, the RNA population was only 6% as active as the original ribozyme. Sequence analysis of this population revealed that any mutation at C709 significantly reduced catalytic activity, while most base changes at other loop positions had little effect on catalysis.

Nishikawa *et al.*[93] used bimolecular HDV ribozyme and isolated variants in which 44 nucleotides of the core catalytic core were mutagenized. Enzyme and substrate domains were created by separately synthesizing fragments of the ribozyme that was conceptually split in the JI/II region. A starting pool of HDV ribozyme variants was prepared in which the truncated core of the HDV ribozyme (nucleotides 705–718 and 726–771, with nucleotides 737–757 deleted) was mutagenized (d = 0.15). Later, mutagenic PCR was used to introduce new mutations during rounds 11–14. The dominant sequence, isolated after 14 rounds of selection, retained the base identities at those sites recognized previously to be important by site-directed mutational analysis.

IV. Artificial Ribozymes

In addition to investigating the structural and functional features of existing ribozymes, *in vitro* selection can be used to create ribozymes with entirely new structures and catalytic activities. Several fundamental approaches and techniques have

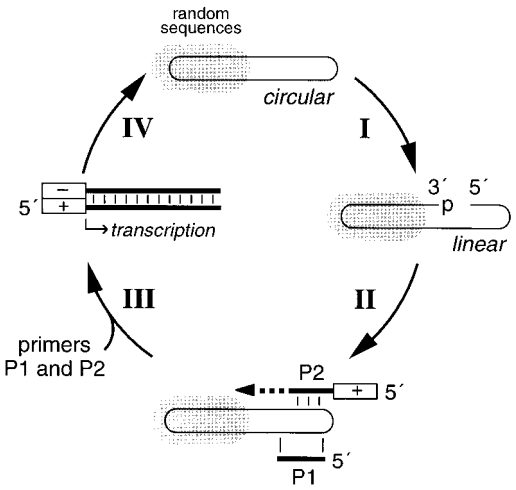


Figure 18. Selection scheme for the isolation of novel self-cleaving ribozymes. To initiate the selection, a pool of RNAs is generated by circularizing linear *in vitro*-transcribed molecules by treatment with RNA ligase. Individuals molecules in the pool carry region(s) of random-sequence nucleotides and two domains of defined sequence to facilitate RT-PCR amplification. Circularized molecules are then incubated under the desired cleavage conditions and the cleaved linear molecules are isolate by their difference in gel mobility using PAGE (I). Isolated molecules are then recircularized by treatment with polynucleotide kinase and RNA ligase to generate templates for amplification (II). RT-PCR with primers P1 and P2 is then employed (III) to introduce an RNA polymerase promoter (boxed "+" and "−") and to create a double-stranded DNA template that is suitable for subsequent transcription using RNA polymerase (IV).

now been established and are represented by the following examples.

A. Self-Cleaving Ribozymes

The four naturally occurring motifs that have been identified to date (Figure 10) likely represent only a small fraction of the number of different self-cleaving RNA motifs that are possible. If we define self-cleaving RNAs simply as RNA structures that correctly position reactive metal-bound hydroxyls to accelerate the rate of divalent metal-catalyzed RNA cleavage, then the vast number of RNAs that would qualify as ribozymes could only be limited by setting some minimum for catalytic rate enhancement. According to this definition, a number of RNAs that fortuitously associate with divalent metal ions could be classified as "self-cleaving". Such examples range from highly structured tRNAs[94] to the simple bimolecular complex formed between the oligomers GAAA and UUU.[15] Indeed, structured RNAs that are known to specifically bind Mg^{2+} can undergo site-specific cleavage under somewhat alkaline conditions.[12,95]

Pan and Uhlenbeck[96] used *in vitro* selection to create new Pb^{2+}-dependent self-cleaving RNAs from pools that were based on the sequence of yeast $tRNA^{Phe}$, an RNA that previously had been found to undergo site-specific cleavage with Pb^{2+}.[94] To achieve this, the authors developed a conceptually powerful selection scheme that allows the isolation of self-cleaving ribozymes that cleave at any individual site within the RNA molecule (Figure 18). Similarly, this

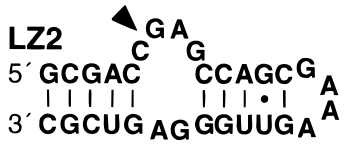


Figure 19. Secondary structure model for LZ2, a Pb^{2+}-dependent self-cleaving ribozyme. Arrow identifies the site of cleavage.

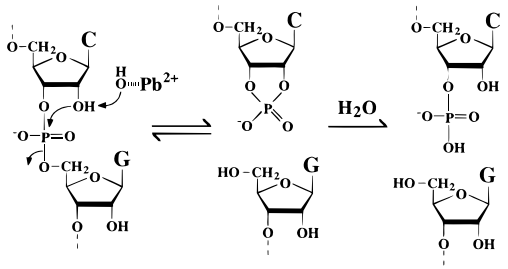


Figure 20. Two-step cleavage mechanism of LZ2. LZ2 cleaves RNA phosphoester presumably by positioning a metal-bound hydroxide ion in proximity to the 2′-hydroxyl of the target internucleotide linkage. The resulting 2′,3′-cyclic phosphate intermediate is subsequently hydrolyzed, primarily yielding the 3′-phosphate isomer. The water nucleophile for this second step is most likely activated by Pb^{2+}. The second step is also stereospecific, as the uncatalyzed hydrolysis of the cyclic intermediate would give an approximately equal mixture of the 2′- and 3′-isomers.

selection scheme has been used to identify tRNA variants that carry significant deletions, yet maintain similar tRNA-like tertiary structures.[97]

To search for novel self-cleaving ribozymes, pools of $tRNA^{Phe}$ variants were generated with either nine or 10 randomized nucleotides that are involved in tertiary structure formation.[95] After six rounds, the parallel selections gave an impressive variety of Pb^{2+}-dependent self-cleaving ribozymes, many that cleave at unique sites. This data demonstrates that a number of small and structurally simple RNAs can accelerate the rate of divalent metal-dependent RNA cleavage by >1000-fold over the rate of RNA phosphoester cleavage in the presence of the metal alone.[98] As discussed earlier, self-cleaving ribozymes operate by transesterification (Figure 11) to yield a product that carries a 2′,3′-cyclic phosphate. An individual self-cleaving RNA (LZ2, Figure 19) that was isolated during this selection was found to promote a two-step hydrolytic mechanism (Figure 20)[98,99] to generate products identical to those produced by certain protein-based enzymes such as pancreatic ribonuclease.

Williams *et al.*[100] have isolated a diverse set of Mg^{2+}-dependent self-cleaving ribozymes from an RNA pool that contained a 100-nucleotide random-sequence domain. The pool was subjected to seven rounds of *in vitro* selection, also using a gel-mobility selection scheme. Each class of ribozymes consist of an internal loop that is flanked by two short helical regions (Figure 21), reminiscent of the self-cleaving structures that were obtained by selection with a Pb^{2+} cofactor.

Catalytic rates for these structures are quite modest, with ribozyme "C45" displaying a k_{obs} of ~0.003 min^{-1}. This equates to an overall rate enhancement of $\sim 10^3-10^4$-fold compared to the rate of RNA cleavage in the presence of Mg^{2+} alone.[101] In comparison,

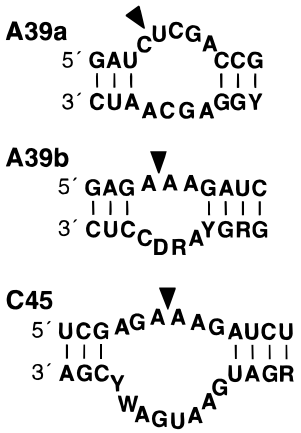


Figure 21. Consensus sequences and secondary structures for three Mg^{2+}-dependent self-cleaving ribozymes: **D** = A, G, or U; **R** = A or G; **W** = A or U; **Y** = U or C.

this k_{obs}, is 100−1000-fold slower than that achieved by examples of the hammerhead ribozyme under similar reaction conditions. Clearly, catalytic motifs such as the hammerhead ribozyme would have enjoyed a considerable selective advantage and would be expected to dominate the final selection pool. More complex ribozymes, however, typically require a greater number of nucleotides to be of defined base identity. The authors of this study speculate that their simple motifs occur 10000-fold more frequently than do the more complex self-cleaving ribozymes that have been found in nature.

B. Other Reactions with Phosphoester and Phosphoanhydride Bonds

1. Phosphoanhydride Formation

Chapman and Szostak[102] have isolated structurally complex ribozymes from a pool of $\sim 10^{15}$ different RNAs that catalyze the formation of phosphoanhydride bonds (Figure 22). The substrate RNA carries an activated phosphate (phosphorimidazole) at its 5′-terminus. Similar nucleotide and oligonucleotide analogues have been made to polymerize on template oligonucleotides in a sequence-specific fashion, by forming new 3′,5′-phosphodiester bonds. In addition, the substrate includes a 3′-terminal biotin to allow the isolation of RNA catalysts via biotin affinity chromatography. The pool RNAs each contained 90 random-sequence nucleotides and were designed to favor the recovery of RNAs that catalyze a self-ligation reaction. The pool RNAs also carried a 3′-terminal hairpin domain that is positioned adjacent to a region that has base complementarity with the substrate hexanucleotide. This initial pool, with its engineered substrate-binding site reacts with its corresponding substrate with a rate of $\sim 7 \times 10^{-6}$ min^{-1}, while ligation to a mismatched substrate was not detected.

The ensemble of RNAs isolated after eight rounds of selection catalyze a joining reaction with a rate of 0.8 min^{-1}. The investigators originally expected to isolate ribozymes that catalyze the formation of a 3′,5′-phosphodiester bond with the activated RNA substrate. However, closer examination of an individual catalyst (Figure 22) revealed that ligation occurs at the 5′-terminus of the ribozyme, thereby

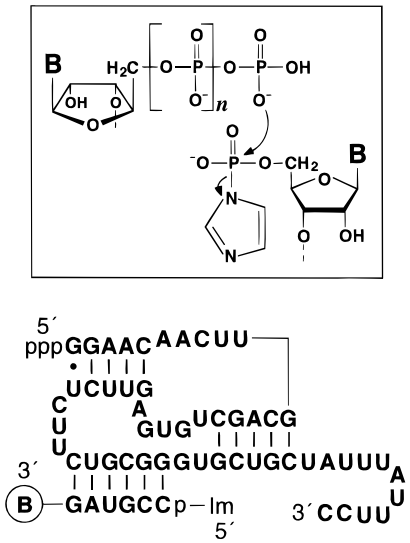


Figure 22. Ribozyme-catalyzed joining of oligoribonucleotides via phosphoanhydride linkages. Sequence and secondary-structure model of an RNA that catalyzes the formation of 5′,5′-oligophosphate linkages between oligoribonucleotides. Encircled **B** represents biotin and **Im** represents imidazole. A ribozyme that carries a 5′-terminal mono-, di-, or triphosphate moiety ($n = 0-2$) catalyzes the formation of a phosphoanhydride bond with a 5′-phosphorimidazole-modified RNA substrate (box).

creating 5′,5′ P1,P4-tetraphosphate linkage. The catalysts have taken advantage of the fact that a γ-phosphate oxygen from the ribozyme's 5′-triphosphate terminus acts as a superior nucleophile compared to the secondary hydroxyls of ribose. Sequential removal of phosphates from the triphosphate terminus of the precursor ribozyme result in the formation of 5′,5′-triphosphate or 5′,5′-pyrophosphate linkages upon catalysis. Interestingly, the 5′-cap of eukaryotic mRNAs is a prominent example of similar linkages that occur in cells.

2. RNA Ligation

Bartel and Szostak[46] demonstrated for the first time that structurally complex ribozymes can be isolated from random-sequence pools of RNA. An RNA pool that carried a total of 220 random-sequence nucleotides was screened (Figure 23) for ribozymes that catalyze the formation of a phosphoester link to another RNA oligomer in a reaction that is analogous to that of an RNA polymerase (Figure 24). The 5′-terminal nucleotides of pool RNAs can form a hairpin structure that brings the 5′-terminal triphosphate of the ribozyme in close proximity to the 3′-terminus of the substrate oligomer. A simple RNA template can also catalyze the condensation of analogous substrates, but only with a rate of $\sim 10^{-8}$ min^{-1}. No detectable coupling occurs in the absence of the template. In effect, a crude active site has been built into the structure of each molecule in the starting pool.

After 10 rounds of selection, the pool catalyzed the ligation reaction with a rate of 0.06 min^{-1}, ~7 million-fold faster than the rate promoted by a simple RNA template. Analysis of the pool by digestion with restriction enzymes demonstrates that the relative abundance of different classes of ribozymes is dynamic and undergoes continuous change as catalysts

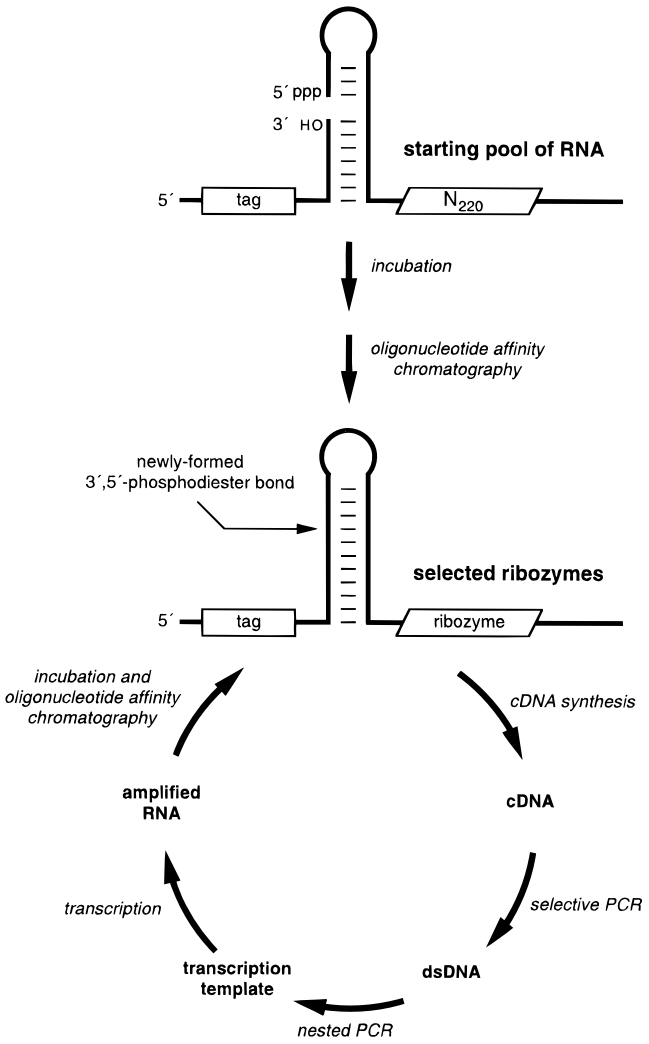


Figure 23. Isolation of "RNA ligase" ribozymes from a random-sequence pool.

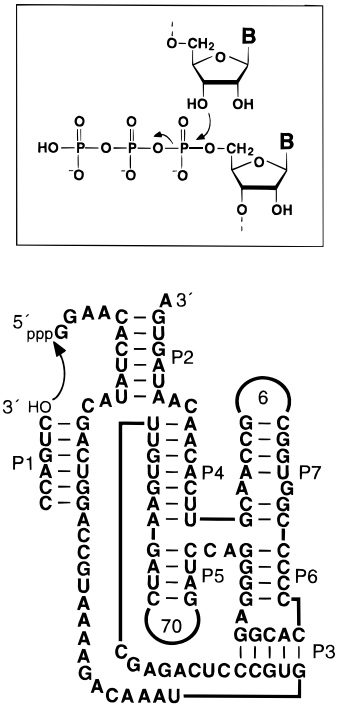


Figure 24. Secondary-structure model of a class I ligase ribozyme from a composite sequence based on comparative sequence analysis of variant catalysts. (box) RNA-catalyzed ligation of RNA involves the nucleophilic attack by the 3′-(2′)-ribose oxygen on the a phosphorus of the ribozyme 5′-triphosphate to form a 3′(2′),5′-phosphodiester linkage at the expense of a phosphoanhydride bond.

that dominate in early selection rounds are subsequently displaced by new catalysts. Ekland *et al.*[103] have identified three distinct classes of RNA ligase ribozymes that are present in this pool. "Class I" ribozymes catalyze the formation of 3′,5′-phosphodiester linkages, while the remaining classes promote the formation of 2′,5′-phosphodiester linkages. In addition, each class has unexpectedly created a new substrate binding site, avoiding the use of the crude active site that was included in the original pool.

The class I molecules include seven base-pairing regions (Figure 24, P1−P7).[103,104] This level of structural complexity approaches that seen with the larger ribozymes that have been isolated from natural sources. Although the random-sequence RNA pool contained $>10^{15}$ different sequences, this only represented a vanishingly small fraction of the total number of RNA sequences possible in a pool of moderate size. The fact that three classes of large ribozymes were isolated supports the hypothesis that RNA is capable of forming a diverse array of higher ordered structures, some that might even match the catalytic achievements of protein enzymes. For example, Ekland and Bartel[105] have altered the template and substrate domains of a class I ligase to create a ribozyme that catalyzes the polymerization of nucleotide triphosphates in a template-directed fashion. This ribozyme is a primitive mimic[40] of the more-refined RNA and DNA polymerase enzymes that are made of protein.

Another important feature of these ribozymes relates to their catalytic speed. *In vitro* analysis of natural ribozymes reveals that most display catalytic rates (k_{cat}) that fall near 1 min^{-1}, an apparent catalytic barrier that unsettles the notion that ribozymes might be capable of more impressive enzyme-like functions. Closer examination of the group 1 ribozyme[106] revealed that this catalyst can cleave RNA with a k_{cat}/K_m of 10^8 M^{-1} min^{-1}, a value that is limited not by the chemical step, but by the rate of RNA duplex formation (substrate binding). Shimayama *et al.*[107] designed a chimeric (mixed RNA/DNA) hammerhead ribozyme that cleaves RNA with rates that approach 100 min^{-1}, providing further evidence that engineered ribozymes might be made to operate at much greater rates. Indeed, the improved ligase ribozymes described by Ekland *et al.*[103] also operate with an estimated k_{obs} of 100 min^{-1}. To improve these rates, one may need to employ more sophisticated selection protocols increase the selective advantage of faster ribozymes, perhaps by isolating tagged ribozymes after millisecond incubation times.

3. RNA Phosphorylation

Sassanfar and Szostak[108] and Connell and Christian[109] have proposed that RNAs of known function could be used as building blocks for the creation of RNAs with entirely new functions. This approach was tested by Lorsch and Szostak,[110] who constructed a pool (Figure 25) of RNAs that was biased in favor of an RNA aptamer specific for adenosine 5′-tri-

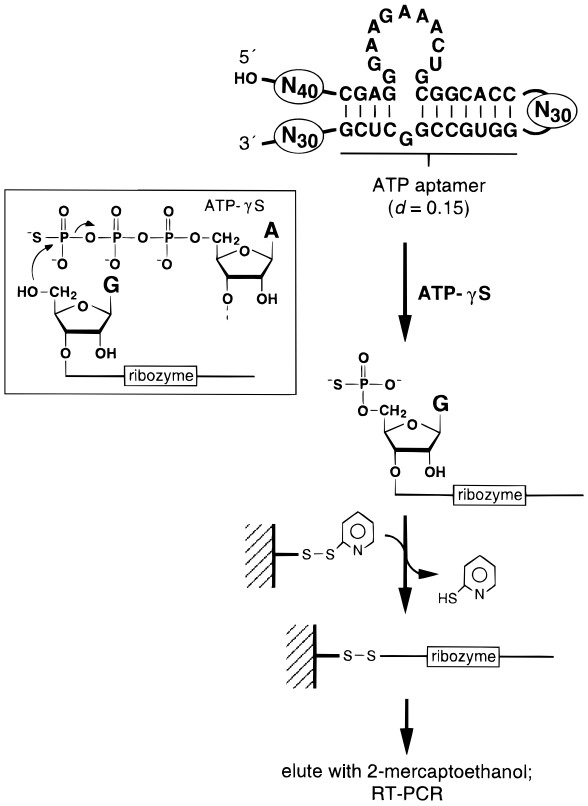


Figure 25. Selection of self-phosphorylating ribozymes from an RNA pool that is biased in favor of an ATP aptamer. The pool is composed of a mutagenized aptamer domain that is flanked by three regions of random sequence. Those RNAs that acquire the γ-thiophosphate from ATP-γS (see inset) are captured on a thiophilic column, then specifically eluted with 2-mercaptoethanol.

phosphate (ATP).[108] This pool, comprised of $>10^{15}$ different RNA molecules, was subjected to 13 rounds of selection for RNAs that acquire a thiophosphate moiety when incubated with γ-thioATP (ATP-γS). The selection yielded an ensemble of RNAs that catalyze self-phosphorylation with an average k_{obs} of $\sim$0.05 min^{-1}. The kinases can also utilize ATP, showing k_{obs} ratios of 50−300-fold for ATP-γS versus ATP. Although these ribozymes achieve an overall rate enhancement of 10^9-fold over the uncatalyzed rate, analogous protein enzymes such as T4 polynucleotide kinase function with catalytic rates that are significantly faster (k_{cat} = 25 000 min^{-1}).

Sequence analysis of individual kinase ribozymes revealed the presence of seven major classes of RNAs, some that display activity like that of polynucleotide kinase. However, two classes of kinases transfer the thiophosphate of ATP-γS to a 2′-hydroxyl located within the ribozyme primary structure. In most instances, the selected ribozymes have largely retained the consensus sequence for the ATP-binding motif, confirming the view that RNA aptamer domains might be an excellent starting point for the evolution of ribozymes with new catalytic activity. However, three classes of catalysts appear to use entirely different structures to bind ATP. This finding suggests that a biased pool was probably not needed to successfully isolate kinase-like ribozymes, but a similar approach may be necessary for other more difficult selection experiments.

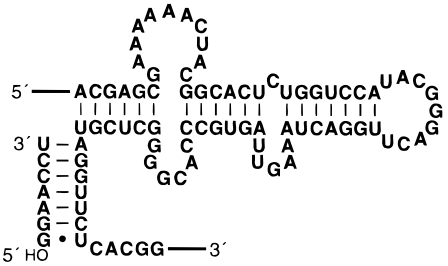


Figure 26. Sequence and secondary structure of a kinase ribozyme. The γ-phosphate of ATP is transferred to the 5′-hydroxyl of the heptamer RNA substrate. IGS is the internal guide sequence.

A more detailed kinetic and thermodynamic analysis[111] was conducted on an individual kinase ribozyme that was converted into two RNAs that correspond to separate enzyme and substrate molecules (Figure 26). The substrate is recognized by the ribozyme via an internal guide sequence (IGS) located within the catalytic RNA. This sequence can be altered to give ribozymes with new substrate specificity as determined by Watson−Crick base-pairing rules. The ribozyme displays a k_{cat} of 0.17 min^{-1} and a K_m of 2 μM for the heptamer substrate. The k_{cat} for this ribozyme is 50-fold slower when ATP is used in place of ATP-γS, consistent with a rate-limiting chemical step that proceeds via a dissociative transition state.

C. Catalytic Aptamers

The application of transition-state theory to the development of new protein catalysts has been quite successful, as demonstrated by studies with catalytic antibodies.[112,113] Several researchers have speculated that nucleic acid catalysts could be developed using similar theoretical approaches.[32,109,114−116] This approach could employ methods that are identical to those used to recover RNA aptamers that have high specificity and affinity for ligand molecules. Hence, if a suitable transition-state analog (TS† analog) is used as the ligand, then tight-binding aptamers may also catalyze the corresponding chemical transformation. Aptamers that bind a given TS† analog are not always catalytic,[117] and individual aptamers must be further screened to determine their ability to catalyze the desired chemical transformation.

Prudent *et al.*[118] provided the first evidence that this approach is applicable to nucleic acids. They began by generating a population of RNA aptamers, some that were specific for the compound shown in Figure 27, a TS† analog of the interconversion between two diasteriomeric biphenyl compounds. Briefly, RNAs were selected for their ability to bind a column matrix that was derivatized with the TS† analog by first incubating the matrix with an RNA pool containing 128 random-sequence nucleotides, removing unbound RNAs by elution with binding buffer, and then specifically eluting the bound RNAs with an excess of the free TS† analog. After seven rounds of *in vitro* selection, several individual RNAs bound the TS† analog, but only one sequence catalyzed the noncovalent isomerization of the bridged biphenyl substrate. A k_{cat} of 2.8×10^{-5} min^{-1} and a K_M of 542 μM were observed. Although modest, the

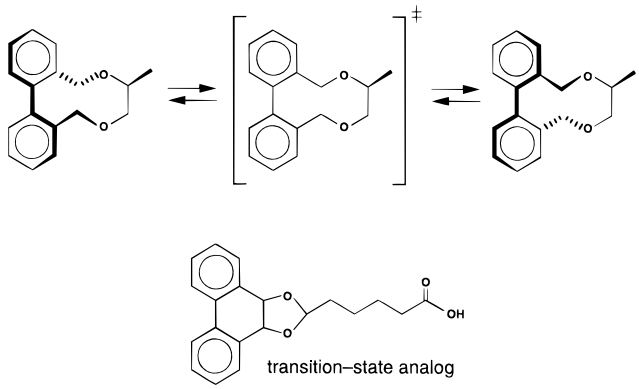


Figure 27. Interconversion between *RS* and *SS* diastereomers of a bridged biphenyl compound and an analog of the transition-state intermediate.

88-fold rate enhancement for this enzyme is likely brought about solely through the use of binding energy, an additional feature that can be exploited in the mechanisms by which nucleic acids promote chemical transformations.

Using a similar approach, Conn *et al.*[119] have screened a pool of $\sim 10^{15}$ different RNA molecules, each carrying a 50-nucleotide random-sequence domain, for aptamers that bind *N*-methylmesoporphyrin (NMM) (Figure 28). NMM is a TS† analog for the metalation of mesoporphyrin and is a potent inhibitor of certain porphyrin chelatases. Individual aptamers that were recovered after 12 rounds of selection were found to accelerate the metalation of mesoporphyrin with Cu^{2+}. One individual displays a k_{cat} of 0.92 min^{-1} and a K_M of 14 μM, giving an overall rate enhancement of 460-fold over the uncatalyzed rate. A 35-nucleotide truncated version displays a k_{cat} of 2 min^{-1}, a rate that corresponds well with rates of natural ribozymes. Similarly, Li and Sen[120] have isolated a catalytic DNA that performs this same reaction (see section V.C). These results demonstrate the feasibility of the indirect approach to the selection of new catalytic polynucleotides.

D. RNA Alkylation

Wilson and Szostak[121] have also used a biased-pool approach to isolate self-alkylating RNAs that show rate enhancements of >1 million-fold over the uncatalyzed reaction. First, biotin-binding RNAs were isolated from a random-sequence pool using affinity-chromatography selection. A 93-nucleotide biotin aptamer, isolated from the first selection, was then randomized ($d = 0.3$) and flanked on each side by 12 nucleotides of random sequence to create a new RNA pool. This pool is expected to contain an abundance of RNAs that still recognize the biotin moiety, but that may also promote chemical reactions with functional groups that are attached to a biotin-like substrate. In this study, a pool of $\sim 10^{14}$ degenerate aptamer sequences was screened for RNAs that catalyze self-alkylation with *N*-biotinyl-*N*′-(iodoacetyl)ethylenediamine (BIE). The RNA pool was constructed such that the 5′-terminal nucleoside of each RNA was 8-mercaptoguanosine, with the intention of isolating RNAs that promote an alkylation of the sulfhydryl group. Selection of catalytic RNAs then proceeds by recovering those RNAs that have acquired a biotin modification via chromatography with streptavidin-modified agarose.

After seven rounds of selective amplification, an individual ribozyme was found to catalyze self-alkylation with a rate of 0.001 min^{-1}. This RNA was further optimized by subjecting a degenerate ($d = 0.3$) ribozyme pool to an additional eight rounds of selection, yielding ribozyme variants that are 50-fold more active. Interestingly, the secondary structure model for an individual ribozyme (Figure 29) does not appear to include a structure similar to the pseudoknot structure found for the biotin aptamer. In addition, alkylation with BIE was found not occur at the 5′-terminus, but at N7 of a guanosine nucleotide that is located in the interior of the molecule (Figure 29). These findings again demonstrate that *in vitro* selection can yield quite unexpected results and that constructing biased pools sometimes may not help or might even hinder the selection process.

E. Reactions with Ester and Amide Bonds

1. Transacylation and Amide Bond Formation

Among the important ester and amide bond reactions in metabolism are tRNA aminoacylation and ribosome-mediated peptidyl transferase reactions, which are the key chemical transformations for the biological synthesis of proteins. In the initial step in the process of encoded protein synthesis, a particular aminoacyl-tRNA synthetase activates its respective amino acid substrate by forming a mixed anhydride with AMP at the expense of ATP. The amino acid is then transferred to the terminus of the corresponding tRNA where it resides as a 2′(3′)-aminoacyl group (Figure 30). Illangasekare *et al.*[122] have demonstrated that this latter reaction can be catalyzed by RNA. A series of RNAs that display self-aminoacylation activity were isolated from a pool of $\sim 10^{14}$ different RNA sequences, each that carried a region of 50 random-sequence nucleotides. The RNA pool was incubated at 0 °C in the presence of activated phenylalanine (phenylalanyl-AMP). RNAs that were acylated were isolated by trapping the unstable aminoacyl RNA via reaction with (naphthoxyacetyl)-*N*-hydroxysuccinimide ester, followed by reverse-phase chromatography to specifically recover those RNAs that carry the hydrophobic naphthyl moiety.

An individual (Figure 30) self-aminoacylating RNA that was recovered after 11 rounds of selection displays a second-order rate of 250 M^{-1} min^{-1}, a rate enhancement of $>10^5$ over that observed for random-sequence RNA. In lieu of selecting the catalysts to operate with a specific metal ion cofactor, the selection reactions were conducted in the presence of several different divalent cations (Mg^{2+}, Ca^{2+}, Mn^{2+}, Zn^{2+}). The strategy of using this "combinatorial mix" of cofactors may be quite helpful by facilitating the isolation of exceedingly rare ribozymes, without regard for the metal ion dependence of the final catalysts. In fact, the ribozyme that was isolated in this selection has requirements for both Mg^{2+} and Ca^{2+}, but its activity is not dependent on Mn^{2+} or Zn^{2+}.

In addition, RNA can also catalyze amide bond formation in a reaction that is analogous to that

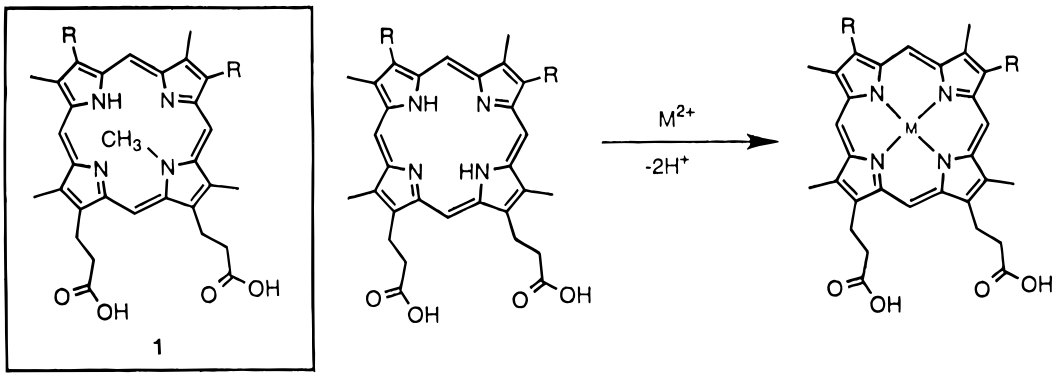
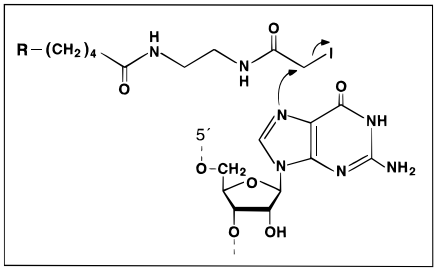
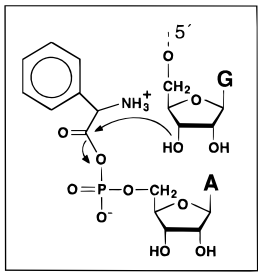


Figure 28. A TS† analogue for the metalation of mesoporphyrin IX is the nonplanar *N*-methylmesoporphyrin (**1**). (Reprinted from ref 113. Copyright 1995 American Association for the Advancement of Science.)





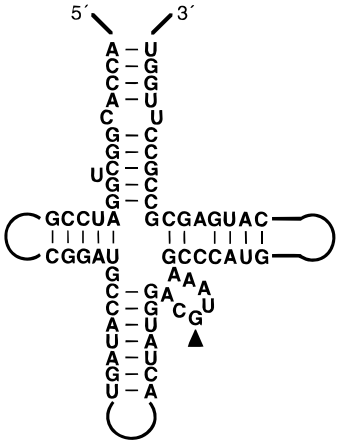


Figure 29. Sequence and secondary structure of the catalytic core of a self-alkylating ribozyme. Arrowhead designates the site of alkylation. (Box) RNA-catalyzed alkylation of N7 of a guanosine nucleotide. R represents biotin.

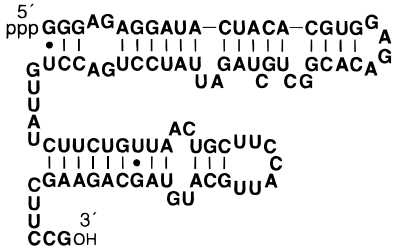


Figure 30. Sequence and proposed secondary structure of a self-aminoacylating ribozyme. (Box) Aminoacylation of RNA using the mixed anhydride of phenylalanine and AMP.

found with natural protein synthesis. Lohse and Szostak[123] have described the isolation of ribozymes that catalyze amino acid transfer reactions, including acyl transfer to an amino group acceptor (Figure 31). This latter reaction is analogous to the peptidyl transferase activity that is carried out by the ribosome. A series of ribozymes was isolated from a pool of 10^{15} different RNA sequences with each molecule carrying a region of 90 random-sequence nucleotides. RNAs that promote acyl transfer from an aminoacyl RNA substrate to their 5′-terminus were recovered by biotin affinity chromatography. Analysis of individuals isolated after 11 rounds of selection revealed a conserved region of 13 nucleotides that likely serves a binding site for the reactants. This template is complementary to the 6-nucleotide DNA substrate and also to the 5′ terminus of the ribozyme, and presumably acts to bring the two reactants in close proximity (Figure 32). Although a simple RNA

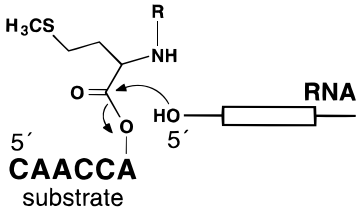


Figure 31. Selection for acyltransferase activity using an RNA substrate that is modified with *N*-biotinylated methionine via a 2′(3′)-ester. R represent biotin.

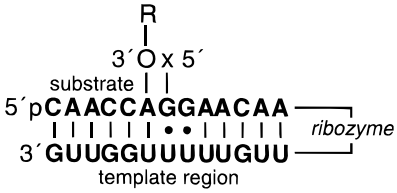


Figure 32. Template region of a ribozyme that catalyzes ester transfer or amide-bond formation. x represents a 5′-hydroxyl or 5′-amino group.

template will accelerate the acyl transfer reaction to a detectable level ($k_{obs} = 1.6 \times 10^{-3}$ min^{-1} at pH 7.3), one individual ribozyme produces a further rate enhancement of 160-fold and operates with a k_{cat}/K_M of 3.8×10^6 min^{-1} M^{-1}. Interestingly, if this same

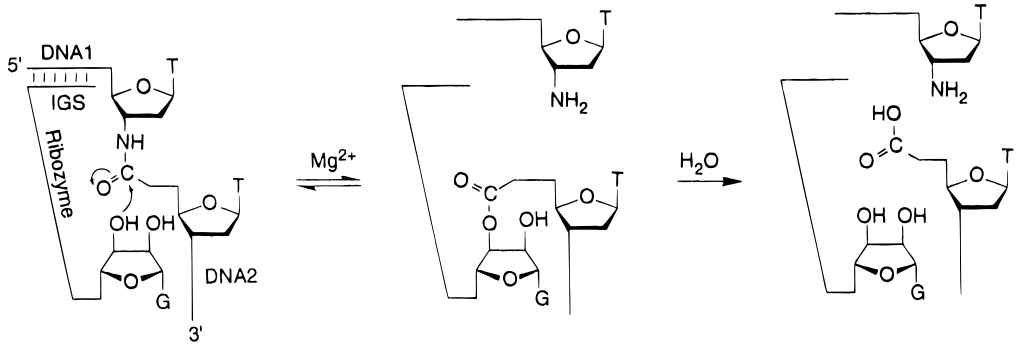


Figure 33. Scheme for ribozyme-catalyzed amide cleavage. The ribozyme binds the amide-containing DNA substrate by base pairing with the internal guide sequence (IGS). The covalent intermediate between the terminal G residue of the ribozyme and the carboxyl group of the cleaved amide is hydrolyzed to regenerate the RNA catalyst. (Reprinted from ref 124. Copyright 1995 American Association for the Advancement of Science.)

ribozyme is modified to carry a 5′-amino in place of the 5′-hydroxyl terminus, then the corresponding acyl transfer occurs in which an amide bond is formed. The k_{obs} values for the hydroxyl and amine nucleophiles are 1.8 and 0.58 min^{-1} at pH 8.4. These findings demonstrate that RNA is capable of catalyzing amide bond formation with rates that approach that of the ribosome-catalyzed peptidyl transferase activity. Moreover, these results are consistent with other data[9] that suggests that ribosomal RNA might form the catalytic center for peptide bond formation in cells.

2. Amide Bond Cleavage

RNA can also catalyze the hydrolysis of amide bonds. Dai *et al.*[124] have characterized a ribozyme that accelerates the hydrolytic cleavage of an amide linkage that resides within an oligodeoxynucleotide (Figure 33). Ribozymes with amidase activity were identified by examining several variant group I ribozymes that were previously selected to cleave a DNA substrate more efficiently.[63] Although the rates reported for the catalysts are modest ($\sim 10^{-7}$ min^{-1}),[124,125] this represents at least 100-fold rate enhancement over the uncatalyzed rate for amide bond hydrolysis. This same ribozyme was also shown to catalyze the cleavage of an amide bond that joined a 3′-deoxy-3′-amino DNA oligomer and the amino acids methionine or arginine, suggesting that ribozymes may be engineered that cleave peptide bonds within the context of proteins.

V. Catalytic DNA

A. RNA Cleavage

Biological catalysis is dominated by enzymes that are made of protein and, to a lesser extent, by catalysts that are composed entirely of RNA. No enzymes made of DNA have been found in nature. However, we can imagine that such catalytic polymers, with functions that are similar to those of ribozymes, could be made via *in vitro* selection. Indeed, the first examples of catalytic DNAs or "deoxyribozymes" have been isolated from pools of single-stranded DNAs. Breaker and Joyce[126] used "catalytic elution" to isolate divalent metal-dependent RNA-cleaving DNA catalysts from a pool of $\sim 10^{14}$ different DNAs (Figure 34). Each molecule in the

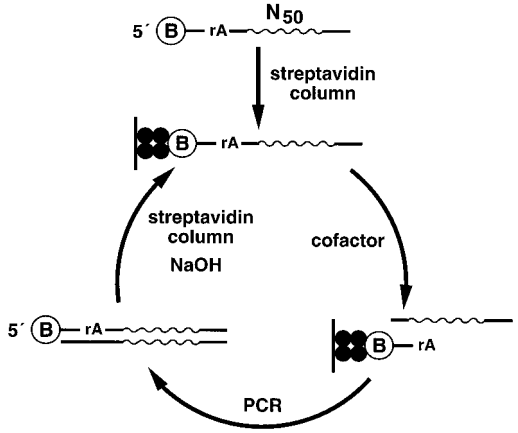


Figure 34. *In vitro* selection of RNA-cleaving DNAs. rA represents the embedded RNA phosphodiester linkage and the encircled B indicates biotin.

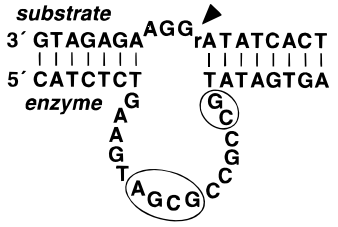


Figure 35. A Pb^{2+}-dependent RNA-cleaving DNA enzyme. Arrowhead identifies the cleavage site. Encircled nucleotides are conserved.

pool contained 50 random-sequence nucleotides and carried a 5′-terminal biotin moiety that allowed each to be immobilized on a streptavidin-derivatized matrix. In addition, a single RNA phosphodiester bond was introduced by enzymatic extension of a DNA analog that included a single 3′-terminal ribonucleotide. Double-stranded DNAs were made single stranded by chemical denaturation, and those that catalyze the cleavage of the embedded RNA linkage were specifically eluted upon inclusion of 1 mM Pb^{2+}. The pK_a of water shifts to ~7.7 when bound to Pb^{2+}, and hence this hydrated metal makes for an excellent general base catalyst for the cleavage of RNA.

Individual DNAs that were recovered after five rounds of selection for Pb^{2+}-dependent RNA phosphoesterase activity display k_{obs} values of 1 min^{-1}. This corresponds to a rate enhancement of $\sim 10^5$-fold over the reaction rate promoted by Pb^{2+} alone. Like ribozymes, these DNA enzymes consist of regions of

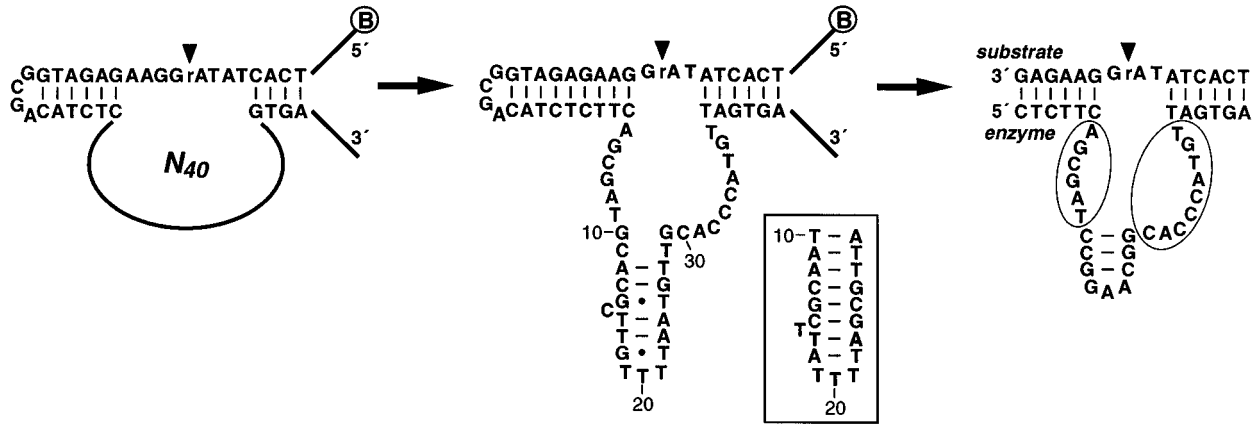


Figure 36. Selection, optimization, and reformulation of a Mg^{2+}-dependent RNA-cleaving DNA. Arrowhead identifies the cleavage site. Boxed sequences are representative of the new hairpin structures that are found in optimized individuals. Encircled nucleotides are conserved.

conserved nucleotide sequence and fold into defined secondary and tertiary structures that facilitate chemical catalysis. The catalysts bind their corresponding substrate by Watson−Crick base pairing and position two conserved regions adjacent to the cleavage site. These self-cleaving catalysts have been divided into separate enzyme and substrate domains (Figure 35) and the enzyme, when supplied with excess substrate, undergoes multiple turn over.

Similarly, selection was used to create three additional pools of RNA phosphoester-cleaving DNAs that require Zn^{2+}, Mn^{2+}, or Mg^{2+} as a cofactor.[101] This selection was begun with a sequence diversity of $\sim 10^{13}$ (Figure 36). The DNA pool was biased such that the 40-nucleotide random-sequence domain was flanked by six and four nucleotides that could base pair either 3′ or 5′, respectively, relative to the target RNA linkage. Active catalysts were isolated after five or six rounds of selection that were dependent on the particular divalent metal that was used for selection. A Mg^{2+}-dependent DNA catalyst (k_{obs} = 0.002 min^{-1}) that was isolated after six rounds of selection was further optimized by preparing a mutagenized pool (d = 0.15) and conducting an additional seven rounds of selection for activity in the presence of 1 mM Mg^{2+}. An individual sequence variant isolated from the reselected pool displays a k_{obs} of 0.02 min^{-1}, a rate enhancement of $\sim 10^5$-fold over the uncatalyzed rate. An engineered form of this catalyst, in which the enzyme domain was separated from the substrate domain (Figure 36), operates with multiple turn over and with a k_{obs} of 0.001 min^{-1} under simulated physiological conditions (2 mM $MgCl_2$, 150 mM KCl, 50 mM HEPES (pH 7.4), 37 °C). These results suggest that enzymes made of DNA may naturally participate in biological catalysis, or at least could be made to do so.

B. DNA Ligation

The catalytic repertoire of DNA enzymes has been expanded by Cuenoud and Szostak,[127] who isolated DNAs that promote the ligation DNA. A pool of $\sim 10^{14}$ different DNAs, each carrying a random-sequence region of 116 nucleotides, was incubated in the presence of a 5′-biotinylated oligonucleotide that also carried a 3′-terminal phosphorimidazole moiety.

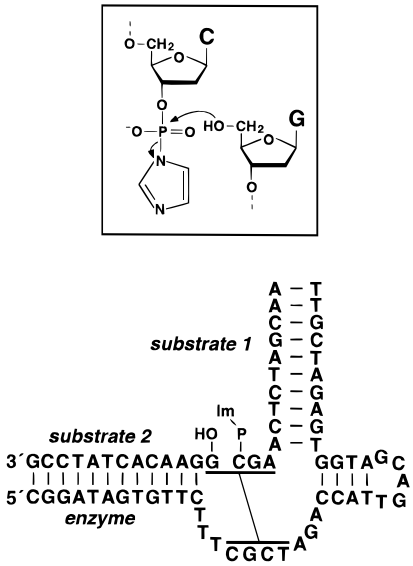


Figure 37. Sequence and secondary-structure model of a DNA enzyme with DNA ligase activity. Bars identify additional base-pairing possibilities and Im indicates imidazole. (Box) Formation of a DNA 3′,5′-phosphodiester bond using an activated DNA phosphoramidate.

DNAs that formed a 3′,5′-phosphodiester bond (Figure 37) were isolated from unreacted DNAs by affinity chromatography using a streptavidin matrix. An individual isolated after nine rounds of selection was reformulated to function with multiple turn over and displays a k_{cat} of 0.07 min^{-1} (Figure 37). By comparison, a simple DNA template provides a rate enhancement of only ~100-fold ($k_{obs} = 2 \times 10^{-5}$ min^{-1} vs $< 2 \times 10^{-7}$ min^{-1} for the uncatalyzed rate) by pairing the two substrates and bringing the reactive groups in close proximity. Although the catalysts were selected in the presence of Zn^{2+}, they also function with Cu^{2+}.

C. Porphyrin Metalation

Deoxyribozymes have also been isolated from random-sequence pools using the "catalytic aptamer" approach. Beginning with a DNA pool of $\sim 10^{15}$ different sequences, Li *et al.*[128] have isolated aptamers to *N*-methylmesoporphyrin IX (Figure 27). A truncated version of an individual aptamer was found to catalyze the metalation of mesoporphyrin IX with

either Cu^{2+} or Zn^{2+}.[120] The catalyst, a 33-nucleotide DNA composed of nearly 50% G residues, displays a k_{cat} of 0.23 min^{-1} and a K_M of ~3 mM, with an overall rate enhancement of ~1400-fold over the uncatalyzed rate of metalation. The catalytic function of the DNA requires K^+, indicating that the abundant guanosine residues in this molecule may be forming a G quartet.

VI. Concluding Remarks

The catalytic repertoire of natural ribozymes, as currently demonstrated, is limited to RNA and DNA phosphoester transfer and hydrolysis reactions. But catalysis by RNA is clearly not restricted to this set of reactions. A variety of ribozymes with enhanced or entirely new catalytic functions have been created by the synthesis of mutagenized and random-sequence pools of RNA and have been isolated by employing repeated rounds of selective amplification. Indeed, a greater number of reactions are now catalyzed by ribozymes derived through *in vitro* selection than are catalyzed by natural ribozymes.

These findings suggest that it may be practical to create "designer" RNA therapeutics and biological reagents. These new catalytic RNAs and DNAs, and the methods used to create them, open entirely new directions in nucleic acids research. Continued discoveries of catalytic RNA lend support to the RNA World theory, whose legitimacy in large part is linked to RNAs potential for robust and diverse chemical catalysis. In addition, DNA can fold into defined tertiary structures, some that can mimic the function of RNA and protein enzymes in their ability to facilitate chemical catalysis.

Whether enzymes made of RNA or DNA can ever rival those that are made of protein remains an unanswered question. The properties of existing artificial ribozymes and deoxyribozymes probably have been shaped more by the limits of the selection methods used to create them than by the catalytic prowess of nucleic acids. It is now known that the chemical rate enhancements of both natural and artificial ribozymes can be substantial, supporting the view that nucleic acids can indeed make efficient catalysts. It will require the implementation of new and refined selection methods and the advancement of rational design methods to continue to expand the catalytic capabilities of RNA and DNA. Most likely, the true limits of catalysis by RNA and DNA have yet to be reached.

VII. References

(1) (1) Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. *Cell* **1982**, *31*, 147−157.
(2) Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. *Cell* **1983**, *35*, 849−857.
(3) Peebles, C. L.; Perlman, P. S.; Mecklenburg, K. L.; Petrillo, M. L.; Tabor, J. H.; Jarrell, K. A.; Cheng, H.-L. *Cell* **1986**, *44*, 213−223.
(4) Prody, G. A.; Bakos, J. T.; Buzayan, J. M.; Schneider, I. R.; Bruening, G. *Science* **1986**, *231*, 1577−1580.
(5) Buzayan, J. M.; Gerlach, W. L.; Bruening, G. *Nature* **1986**, *323*, 349−353.
(6) Sharmeen, L.; Kuo, M. Y. P.; Dinter-Gottlieb, G.; Taylor, J. *J. Virol.* **1988**, *62*, 2674−2679.
(7) Saville, B. J.; Collins, R. A. *Cell* **1990**, *61*, 685−696.
(8) Sharp, P. *Science* **1991**, *253*, 663.
(9) Noller, H. F.; Hoffarth, V.; Zimniak, L. *Science* **1992**, *256*, 1416−1419.
(10) Symons, R. H. *Annu. Rev. Biochem.* **1992**, *61*, 641−671.
(11) Cech, T. R.; Bass, B. L. *Annu. Rev. Biochem.* **1986**, *55*, 599−629.
(12) Wintermeyer, W.; Zachau, H. G. *Biochem. Biophys. Acta* **1973**, *299*, 82−90.
(13) Werner, C.; Krebs, B.; Keith, G.; Dirheimer, G. *Biochim. Biophys. Acta* **1976**, *432*, 161−175.
(14) Dange, V.; Van Atta, R. B.; Hecht, S. M. *Science* **1990**, *248*, 585−588.
(15) Kazakov, S.; Altman, S. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 7939−7943.
(16) Yarus, M. *FASEB J.* **1993**, *7*, 31−39.
(17) Pyle, A. M. *Science* **1993**, *261*, 709−714.
(18) Breaker, R. R.; Joyce, G. F. *J. Mol. Evol.* **1994**, *40*, 551−558.
(19) Robertson, D. L.; Joyce, G. F. *Nature* **1990**, *344*, 467−468.
(20) Herschlag, D.; Cech, T. R. *Nature* **1990**, *344*, 405−409.
(21) Perreault, J.-P.; Altman, S. *J. Mol. Biol.* **1992**, *226*, 399−409.
(22) Mörl, M.; Niemer, I.; Schmelzer, C. *Cell* **1992**, *70*, 803−810.
(23) Zimmerley, S.; *et al. Cell* **1995**, *83*, 529−538.
(24) Yang, J.; Zimmerly, S.; Perlman, P. S.; Lambowitz, A. M. *Nature* **1996**, *381*, 332−335.
(25) Piccirilli, J. A.; McConnell, T. S.; Zaug, A. J.; Noller, H. F.; Cech, T. R. *Science* **1992**, *256*, 1420−1423.
(26) Gilbert, W. *Nature* **1986**, *319*, 618.
(27) Benner, S. A.; Ellington, A. D.; Tauer, A. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 7054−7058.
(28) Hirao, I.; Ellington, A. D. *Curr. Biol.* **1995**, *5*, 1017−1022.
(29) Green, R.; Ellington, A. D.; Bartel, D. P.; Szostak, J. W. *Methods* **1991**, *2*, 75−86.
(30) Szostak, J. W. *Trends Biochem. Sci.* **1992**, *17*, 89−93.
(31) Joyce, G. F. *Pure Appl. Chem.* **1993**, *65*, 1205−1212.
(32) Szostak, J. W.; Ellington, A. D. *The RNA World* **1993**, 511−533.
(33) Famulok, M.; Szostak, J. W. In *Nucleic Acids and Molecular Biology* **1993**, *7*, 271−284.
(34) Breaker, R. R.; Joyce, G. F. *Trends Biotechnol.* **1994**, *12*, 268−275.
(35) Joyce, G. F. *Curr. Opin. Struct. Biol.* **1994**, *4*, 331−336.
(36) Chapman, K. B.; Szostak, J. W. *Curr. Opin. Struct. Biol.* **1994**, *4*, 618−622.
(37) Burgstaller, P.; Famulok, M. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 1189−1192.
(38) Kumar, P. K. R.; Ellington, A. D. *FASEB J.* **1995**, *9*, 1183−1195.
(39) Breaker, R. R. *Curr. Opin. Biotechnol.* **1996**, *7*, 442−448.
(40) Hager, A. J.; Pollard, J. D. Jr.; Szostak, J. W. *Chem. Biol.* **1996**, *3*, 717−725.
(41) Pyle, A. M.; Green, J. B. *Curr. Biol.* **1995**, *5*, 303−310.
(42) Corey, M. J.; Corey, E. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 11428−11434.
(43) Mills, D. R.; Peterson, R. L.; Spiegelman, S. *Proc. Natl. Acad. Sci. USA* **1967**, *58*, 217−224.
(44) Spiegelman, S. *Q. Rev. Biophys.* **1971**, *4*, 213−253.
(45) Joyce, G. F.; Inoue, T. *Nucleic Acids Res.* **1989**, *17*, 711−722.
(46) Bartel, D. P.; Szostak, J. W. *Science* **1993**, *261*, 1411−1418.
(47) Horton, R. M.; Hunt, H. D.; Ho, S. N.; Pullen, J. K.; Pease, L. R. *Gene* **1989**, *77*, 61−68.
(48) Gold, L.; Polisky, B.; Uhlenbeck, O.; Yarus, M. *Annu. Rev. Biochem.* **1995**, *64*, 763−797.
(49) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82−84.
(50) Pagan-Ramos, E.; Lee, Y.; Engelke, D. R. *RNA* **1996**, *2*, 1100−1109.
(51) Joyce, G. F. In *Antisense RNA and DNA*; T. R. Cech Wiley-Liss: New York, 1992; pp 353−372.
(52) Saiki, R. K.; Gelfand, D. H.; Stoffel, S.; Scharf, S. J.; Higuchi, R.; Horn, G. T.; Mullis, K. B.; Erlich, H. A. *Science* **1988**, *239*, 487−491.
(53) Fahy, E.; Kwoh, D. Y.; Gingeras, T. R. *PCR Methods Appl.* **1991**, *1*, 25−33.
(54) Cadwell, R. C.; Joyce, G. F. *PCR Methods Appl.* **1992** *2*, 28−33.
(55) Cadwell, R. C.; Joyce, G. F. *PCR Methods Appl.* **1994** *3*, S136−S140.
(56) Breaker, R. R.; Joyce, G. F. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6093−6097.
(57) Breaker, R. R.; Joyce, G. F. In *Self-production of Supramolecular Structures*; Fleischaker, G. R., Colonna, S. C., Luisi, P. L., Eds.; 1994; pp 127−135.
(58) Breaker, R. R.; Banerji, A.; Joyce, G. F. *Biochemistry* **1994**, *33*, 11980−11986.
(59) Cech, T. R. *Annu. Rev. Biochem.* **1990**, *59*, 543−568.
(60) McSwiggen, J. A.; Cech, T. R. *Science* **1989**, *244*, 679−683.
(61) Rajagopal, J.; Doudna, J. A.; Szostak, J. W. *Science* **1989**, *244*, 692−694.
(62) Joyce, G. F. *Gene* **1989**, *82*, 83−87
(63) Beaudry, A. A.; Joyce, G. F. *Science* **1992**, *257*, 635−641.
(64) Tsang, J.; Joyce, G. F. *Biochemistry* **1994**, *33*, 5966−5973.
(65) Tsang, J.; Joyce, G. F. *J. Mol. Biol.* **1996**, *262*, 31−42.
(66) Guthrie, J. P. *J. Am. Chem. Soc.* **1977**, *99*, 3991−4001.
(67) Lehman, N.; Joyce, G. F. *Nature* **1993**, *361*, 182−185.
(68) Lehman, N.; Joyce, G. F. *Curr. Biol.* **1993**, *3*, 723−734.
(69) Green, R.; Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *347*, 406−408.

(70) Green, R.; Szostak, J. W. *J. Mol. Biol.* **1994** *235*, 140−155.
(71) Green, R.; Szostak, J. W. *Science* **1992** *258*, 1910−1915.
(72) Williams, K. P.; Imahori, H.; Fujimoto, D. N.; Inoue, T. *Nucleic Acids Res.* **1994**, *22*, 2003−2009.
(73) Forster, A. C.; Altman, S. *Science* **1990** *249*, 783−786
(74) Yuan, Y.; Altman, S. *Science* **1994**, *263*, 1269−1273.
(75) Liu, F.; Altman, S. *Cell* **1994**, *77*, 1093−1100.
(76) Pan, T. *Biochemistry* **1995**, *34*, 8458−8464.
(77) Forster, A. C.; Symons, R. H. *Cell* **1987**, *49*, 211−220.
(78) Haseloff, J.; Gerlach, W. L. *Nature* **1988**, *334*, 585−591.
(79) Hertel, K. J.; Herschlag, D.; Uhlenbeck, O. C. *Biochemistry* **1994**, *33*, 3374−3385.
(80) Nakamaye, K. L.; Eckstein, F. *Biochemistry* **1994**, *33*, 1271−1277.
(81) Long, D. M.; Uhlenbeck, O. C. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6977−6981.
(82) Ishizaka, M.; Ohshima, Y.; Tani, T. *Biochem. Biophys. Res. Commun.* **1995**, *214*, 403−409.
(83) Lieber, A.; Strauss, M. *Mol. Cell Biol.* **1995**, *15*, 540−551.
(84) McCall, M. J.; Hendry, P.; Jennings, P. A. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 5710−5714.
(85) Hendry, P.; McCall, M. J.; Santiago, F. S.; Jennings, P. A. *Nucleic Acids Res.* **1995**, *23*, 3922−3927.
(86) Hampel, A.; Tritz, R. *Biochemistry* **1989**, *28*, 4929−4933.
(87) Berzal-Harranz, A.; Joseph, A.; Burke, J. A. *Genes Dev.* **1992**, *6*, 129−134.
(88) Chowrira, B. M.; Berzal-Herranz, A.; Burke, J. M. *Nature* **1991**, *354*, 320−322.
(89) Joseph, S.; Berzal-Herranz, A.; Chowrira, B. M.; Butcher, S. E.; Burke, J. M. *Genes Dev.* **1993**, *7*, 130−138.
(90) Berzal-Herranz, A.; Joseph, S.; Chowrira, B. M.; Butcher, S. E.; Burke, J. M. *EMBO J.* **1993**, *12*, 2567−2574.
(91) Joseph, S.; Burke, J. M. *J. Biol. Chem.* **1993**, *268*, 24515−24518.
(92) Kawakami, J.; Kumar, P. K. R.; Suh, Y.; Nishikawa, F.; Kawakami, K.; Taira, T.; Ohtsuka, E.; Nishikawa, S. *Eur. J. Biochem.* **1993**, *217*, 29−36.
(93) Nishikawa, F.; Kawakami, J.; Chiba, A.; Shirai, M.; Kumar, P. K. R.; Nishikawa, S. *Eur. J. Biochem.* **1996**, *237*, 712−718.
(94) Brown, R. S.; Hingerty, B. E.; Dewan, J. C.; Klug, A. *Nature* **1983**, *303*, 543−546.
(95) Kazakov, S.; Altman, S. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 9193−9197.
(96) Pan, T.; Uhlenbeck, O. C. *Biochemistry* **1992**, *31*, 3887−3895.
(97) Dichtl, B.; Pan, T.; DiRenzo, A. B.; Uhlenbeck, O. C. *Nucleic Acids Res.* **1993**, *21*, 531−535.
(98) Pan, T.; Uhlenbeck, O. C. *Nature* **1992**, *358*, 560−563.
(99) Pan, T.; Dichtl, B.; Uhlenbeck, O. C. *Biochemistry* **1994**, *33*, 9561−9565.
(100) Williams, K. P.; Ciafre, S.; Tocchini-Valentini, G. P. *EMBO J.* **1995**, *14*, 4551−4557.
(101) Breaker, R. R.; Joyce, G. F. *Chem., Biol.* **1995**, *2*, 655−660.
(102) Chapman, K. B.; Szostak, J. W. *Chem., Biol.* **1995**, *2*, 325−333.
(103) Ekland, E. H.; Szostak, J. W.; Bartel, D. P. *Science* **1995**, *269*, 364−370.
(104) Ekland, E. C.; Bartel, D. P. *Nucleic Acids Res.* **1995**, *23*, 3231−3238.
(105) Ekland, E. H.; Bartel, D. P. *Nature* **1996**, *382*, 373−376.
(106) Herschlag, D.; Cech, T. R. *Biochemistry* **1990**, *29*, 10159−10171.
(107) Shimayama, T.; Nishikawa, S.; Taira, K. *FEBS Lett.* **1995**, *368*, 304−306.
(108) Sassanfar, M.; Szostak, J. W. *Nature* **1993**, *364*, 550−553.
(109) Connell, G. J.; Christian, E. L. *Origins Life* **1993**, *23*, 291−297.
(110) Lorsch, J. R.; Szostak, J. W. *Nature* **1994**, *371*, 31−36.
(111) Lorsch, J. R.; Szostak, J. W. *Biochemistry* **1995**, *34*, 15315−15327.
(112) Lerner, R. A.; Benkovic, S. J.; Schultz, P. G. *Science* **1991**, *252*, 659−667.
(113) Schultz, P. G.; Lerner, R. A. *Science* **1995**, *269*, 1835−1842.
(114) Szathmary, E. *Nature* **1990**, *344*, 115.
(115) Abelson, J. *Science* **1990**, *249*, 448−489.
(116) Ellington, A. D.; Szostak, J. W. *Nature* **1992**, *355*, 850−852.
(117) Morris, K. N.; Tarasow, T. M.; Julin, C. M.; Simons, S. L.; Hilvert, D. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 13028−13032.
(118) Prudent, J. R.; Uno, T.; Schultz, P. G. *Science* **1994**, *264*, 1924−1927.
(119) Conn, M. M.; Prudent, J. R.; Schultz, P. G. *J. Am. Chem. Soc.* **1996**, *118*, 7012−7013.
(120) Li, Y.; Sen, D. *Nature Struct. Biol.* **1996**, *3*, 743−747.
(121) Wilson, C.; Szostak, J. W. *Nature* **1995**, *374*, 777−782.
(122) Illangasekare, M.; Sanchez, G.; Nickles, T.; Yarus, M. *Science* **1995**, *267*, 643−647.
(123) Lohse, P. A.; Szostak, J. W. *Nature* **1996**, *381*, 442−444.
(124) Dai, X.; DeMesmaeker, A.; Joyce, G. F. *Science* **1995**, *267*, 237−240.
(125) Dai, X.; De Mesmaeker, A.; Joyce, G. F. *Science* **1996**, *272*, 18−19.
(126) Breaker, R. R.; Joyce, G. F. *Chem., Biol.* **1994**, *1*, 223−229.
(127) Cuenoud, B.; Szostak, J. W. *Nature* **1995**, *375*, 611−614.
(128) Li, Y.; Geyer, R.; Sen, D. *Biochemistry* **1996**, *35*, 6911−6922.
(129) Hertel, K. J.; *et al. Nucleic Acids Res.* **1992**, *20*, 3252.
(130) Perrotta, A. T.; Been, M. D. *Nature* **1991**, *350*, 434−436.
(131) Tanner, N. K.; Schaff, S.; Thill, G.; Petit-Koskas, E.; Crain-Denoyelle, A.-M.; Westhof, E. *Curr. Biol.* **1994**, *4*, 488−498.

CR960008K