4768–4777 *Nucleic Acids Research, 2008, Vol. 36, No. 14* *Published online 21 July 2008*
doi:10.1093/nar/gkn396

In vitro selection of small RNA-cleaving deoxyribozymes that cleave pyrimidine–pyrimidine junctions

Kenny Schlosser[1], Jimmy Gu[1], Jeffrey C. F. Lam[1] and Yingfu Li[1,2,*]

[1]Department of Biochemistry and Biomedical Sciences and [2]Department of Chemistry, McMaster University, Hamilton, Ontario, Canada L8N 3Z5

Received April 15, 2008; Revised and Accepted June 5, 2008

ABSTRACT

Herein, we sought new or improved endoribonucleases based on catalytic DNA molecules known as deoxyribozymes. The current repertoire of RNA-cleaving deoxyribozymes can cleave nearly all of the 16 possible dinucleotide junctions with rates of at least 0.1/min, with the exception of pyrimidine–pyrimidine (pyr–pyr) junctions, which are cleaved 1–3 orders of magnitude slower. We conducted four separate *in vitro* selection experiments to target each pyr–pyr dinucleotide combination (i.e. CC, UC, CT and UT) within a chimeric RNA/DNA substrate. We used a library of DNA molecules containing only 20 random-sequence nucleotides, so that all possible sequence permutations could be sampled in each experiment. From a total of 245 clones, we identified 22 different sequence families, of which 21 represented novel deoxyribozyme motifs. The fastest deoxyribozymes exhibited k_{obs} values (single-turnover, intermolecular format) of 0.12/min, 0.04/min, 0.13/min and 0.15/min against CC, UC, CT and UT junctions, respectively. These values represent a 6- to 8-fold improvement for CC and UC junctions, and a 1000- to 1600-fold improvement for CT and UT junctions, compared to the best rates reported previously under identical reaction conditions. The same deoxyribozymes exhibited ~1000-fold lower activity against all RNA substrates, but could potentially be improved through further *in vitro* evolution and engineering.

INTRODUCTION

Deoxyribozymes (or DNAzymes) are enzymes made from single-stranded DNA (1). Although they are yet to be found in nature, many different deoxyribozymes have been made in laboratories through a process known as *in vitro* selection (2). Enzymes made from DNA can offer several practical advantages over those made from protein or RNA, including simplicity/convenience, greater stability and lower production costs. RNA-cleaving deoxyribozymes represent a very important class (3), which has been used for a variety of purposes including (i) biosensors (4–7), (ii) DNA computational elements with Boolean logic functions (8,9), (iii) nanomachines (10), (iv) antiviral or gene control agents (11–13) and (v) the *in vitro* manipulation of RNA (14).

We are interested in understanding the functional limits of DNA-mediated catalysis, using RNA-cleaving deoxyribozymes as a model system. These efforts should expand their utility, and provide insight to better assess future applicative potential. One specific question that we have sought to answer is how many different dinucleotide junctions can be cleaved efficiently by RNA-cleaving deoxyribozymes?

This is an important question because the ability to efficiently cleave at any location in a given RNA molecule would be very useful for RNA structural and/or modification studies. Furthermore, improved cleavage site versatility would benefit applications involving long RNA molecules, which often possess higher order structures that can interfere with DNAzyme binding, and therefore limit the number and types of susceptible target sites (15).

We have previously established that all 16 dinucleotide junctions are in fact susceptible to deoxyribozyme-mediated cleavage (16), but efficient cleavage of all junctions has not been demonstrated. Existing RNA-cleaving deoxyribozymes such as the 10-23 (17), bipartite (18) and 8-17 (19) serve as a standard baseline, to which activity can be compared. These deoxyribozymes can achieve reaction rates from ~0.1 to 10/min for the cleavage of various purine–purine, purine–pyrimidine and/or pyrimidine–purine junctions (Table 1). However, comparable rates for cleavage of any of the four different pyrimidine–pyrimidine (pyr–pyr) junctions have not been reported. So far, only the 8-17 deoxyribozyme has confirmed

*To whom correspondence should be addressed. Tel: +9055259140, ext. 22462; Fax: +9055229033; Email: liying@mcmaster.ca

activity against pyr–pyr junctions, with rates that are several orders of lower magnitude (19).

We have hypothesized that (i) efficient deoxyribozyme-mediated cleavage of all four pyr–pyr junctions is possible, and (ii) these deoxyribozymes can be isolated through a focused *in vitro* selection effort using a small (i.e. 20 nt) random library approach. These hypotheses are largely motivated by the recent discovery that 8-17 can cleave all 16 dinucleotide junctions (19), and different sequence variants can show different dinucleotide selectivities. Because only 4 of the 15 nt that comprise the 8-17 catalytic core are absolutely conserved, there may still be more proficient variants among the thousands of possible sequence permutations. A small random-sequence library would thereby ensure that all 8-17 sequence variations are represented in the initial pool. The identification of more efficient 8-17 sequence variants is an attractive prospect, because it would avoid the time-consuming characterization experiments that are often necessary before novel motifs can be used productively.

Toward this end, we conducted four separate *in vitro* selection experiments to target the four different types of chimeric (RNA/DNA) pyr–pyr junctions (i.e. CC, UC, CT and UT). Two additional selection experiments were also performed as controls, targeting AC and AT junctions, for which small and efficient RNA-cleaving deoxyribozymes are known to exist.

Table 1. Reaction rates of various RNA-cleaving DNAzymes against different dinucleotide junctions

DNAzyme	10-23	Bipartite	8-17	Various
Reference	17	18	19	This work
Size (nt)	15	20	15	20
Dinucleotide junction				
GG			9.2	
AG		0.0001	3.1	
CG			0.63	
UG			0.20	
GA		ND	3.4	
AA		0.1	1.4	
CA		0.0001	0.17	
UA		<0.0001	0.47	
GC	0.10		0.36	
AC	0.015	0.01	0.11	
CC			0.019	0.12
UC			0.0047	0.04
GU/T	0.38		0.0037	
AU/T	0.31	0.01	0.00089	
CT			0.00014	0.13
UT			0.000095	0.15

All k_{obs} (min^{-1}) values were determined under single-turnover conditions in an intermolecular format. Reaction conditions for 10-23: 10 mM $MgCl_2$, pH 7.5, 37°C, all-RNA substrate. Reaction conditions for bipartite: 30 mM $MgCl_2$, pH 7.4, 23°C, all-RNA substrate. Reaction conditions for 8-17: 100 mM KCl, 400 mM NaCl, 7.5 mM $MgCl_2$, 7.5 mM $MnCl_2$, pH 7.0, 23°C, chimeric RNA/DNA substrate. The 8-17 rates are derived from several different 8-17 sequence variants. Reaction conditions for various DNAzymes from this work: 100 mM KCl, 400 mM NaCl, 7.5 mM $MgCl_2$, 7.5 mM $MnCl_2$, pH 7.0, 23°C, chimeric RNA/DNA substrate. Size refers to the length of the catalytic core. ND: no detectable cleavage activity. Blank squares indicate no reported cleavage activity.

MATERIALS AND METHODS

Oligonucleotides and reagents

Oligonucleotides were prepared by automated DNA synthesis using cyanoethylphosphoramidite chemistry (Integrated DNA Technologies; Mobix Central Facility, McMaster University). Oligonucleotides were purified by 10% preparative denaturing (8 M urea) polyacrylamide gel electrophoresis (PAGE) and their concentrations were determined by spectroscopic methods. Nucleoside 5′-triphosphates, [γ-^{32}P]ATP and [α-^{32}P]dGTP were purchased from Perkin Elmer (Boston, MA, USA). T4 DNA ligase and T4 polynucleotide kinase (PNK) were purchased from MBI Fermentas (Burlington, ON, CAN). *Tth* DNA polymerase was purchased from Biotools (Madrid, Spain). All chemical reagents were purchased from Sigma (Oakville, ON, CAN).

In vitro selection procedure

The library design and *in vitro* selection procedure are illustrated in Figures 1A and 2, respectively.

Step 1 and 2. The full-length 73-nt library was constructed by splint-directed ligation of the 43-nt enzyme domain to the 30-nt substrate domain. Each selection experiment used the same enzyme domain, but different substrate domains and appropriate ligation splints (i.e. 22-nt DNA oligos complementary to adjacent regions in both the substrate and enzyme domains). A total of 1000 pmol of 5′-phosphorylated enzyme domain, 1100 pmol of ligation splint, 1200 pmol of substrate

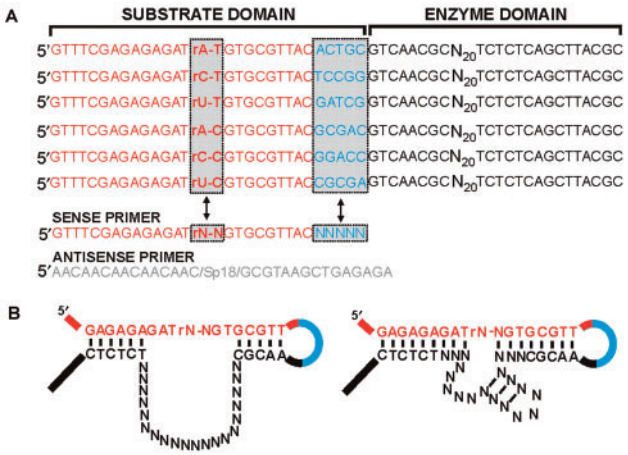

Figure 1. Library design. (**A**) Sequences of libraries and PCR primers. Regions that differ in sequence are highlighted by gray boxes, which include the dinucleotide cleavage junction (denoted as rN–N in the sense primer) and the five terminal nucleotides of the substrate domain (denoted as NNNNN in the sense primer). It should be noted that six different sense primers were used to minimize PCR cross-contamination, although only one is shown for simplicity. N_{20} refers to 20 random nucleotides, where N = A, G, C or T deoxyribonucleotides. Sp18 refers to an 18-atom hexa-ethyleneglycol spacer. (**B**) Two hypothetical secondary structure arrangements showing location of preengineered base-pairing interactions between substrate and enzyme domains. A 2-nt frameshift in base-pairing interactions on the 5′-side of the cleavage junction is possible. The structure on the right shows how an 8-17 motif might form.

domain and ~1 pmol of 5′-^{32}P-labelled enzyme domain were combined, supplemented with DEPC-H_20 to a final volume of 85 µl, and then heated at 90°C for 1 min. After cooling at RT for 10 min, 10 µl of 10× T4 DNA ligase buffer (manufacturer supplied) and 5 µl of T4 DNA ligase (5 Weiss units/µl) were added to initiate the reaction. The ligation mixture was incubated at RT for 2 h, then ethanol precipitated, and purified by 10% denaturing PAGE. The final yield of the 73-nt ligation product was ~600 pmol, which was used for the first round of selection. Therefore, on average ~300 copies of every possible sequence should have been represented in the initial pool.

Step 3. The RNA cleavage reaction was initiated by adding a 2× selection buffer (200 mM KCl, 800 mM NaCl, 15 mM $MgCl_2$, 15 mM $MnCl_2$, 100 mM HEPES pH 7.0 at 23°C) to an equal volume of DEPC-H_20 in

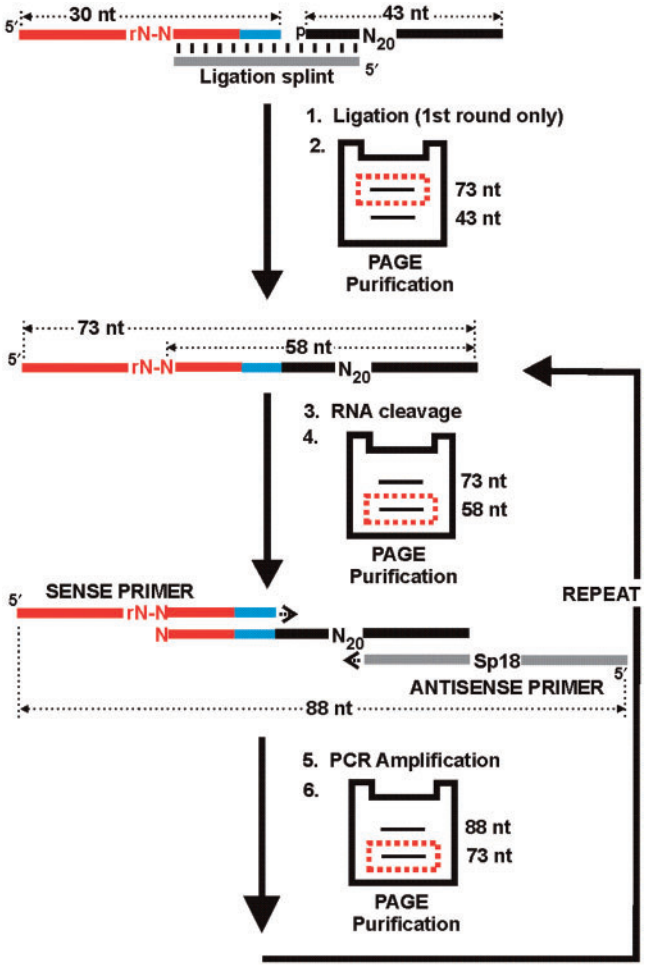

Figure 2. *In vitro* selection scheme. Step 1 and 2: ligation of substrate and enzyme domains, followed by PAGE purification (first round only). Step 3: RNA cleavage reaction is conducted in presence of suitable buffer and designated period of time. Step 4: the 58-nt cleavage product is separated from the 73-nt uncleaved precursor molecules. Step 5: PCR amplification. Sp18 refers to an 18-atom hexa-ethyleneglycol spacer that blocks DNA polymerization, leading to a dsDNA product composed of two strands of unequal length. The sense primer contains the target ribonucleotide cleavage junction. Step 6: separation of the 73-nt sense strand from the 88-nt antisense strand. Steps 3–6 are then repeated to enrich the population for active sequences.

which the 73-nt library had been resuspended and preheated (90°C for 30 s, followed by 10 min cooling at RT), to a final concentration of ~1 µM. The reaction was incubated at RT for a designated period of time, quenched with the addition of EDTA (pH 8.0) to 30 mM, and ethanol precipitated. In rounds 1–4, a 100 min reaction time was used. In round 5, the population was divided into three subpopulations and subjected to 1, 10 and 100 min reaction times. In rounds 6–9, reaction times of 0.1, 1 and 10 min were used. From rounds 5 to 8, the 1 min subpopulation was used to seed the next round of selection.

Step 4. The 58-nt cleavage product was purified by 10% denaturing PAGE, excised from the gel and recovered by ethanol precipitation.

Step 5. The cleavage product was PCR amplified in a 100 µl reaction volume, containing 0.2 mM of each dNTP, 10 µl of 10× PCR buffer (Biotools), 0.5 µM each of sense and antisense primers, 5 U *Tth* DNA polymerase, 2 µl of 50× SYBR green (Invitrogen, Burlington, ON, CAN) and 20 µCi of [α-^{32}P]dGTP. Reactions were conducted on a Cepheid SmartCycler (Sunnyvale, CA, USA) and monitored in real time using SYBR green as the reporter. Each PCR cycle consisted of 94°C for 30 s, 50°C for 45 s and 72°C for 40 s. PCR amplification reintroduces the substrate domain, via the sense primer which contains the target ribonucleotide. The antisense primer contains an internal 18-atom hexa-ethyleneglycol spacer that blocks further DNA polymerization, leading to a dsDNA product of unequal length.

Step 6. The 73-nt sense strand was separated from the 88-nt antisense strand by 10% denaturing PAGE, excised from the gel and recovered by ethanol precipitation. Steps 3–6 were repeated for nine rounds.

Cloning and sequencing of selected DNA populations

Deoxyribozymes from the 100 min subpopulation of round 5, and the 10 min subpopulation of round 9 were cloned with the InsTAclone PCR cloning kit (Fermentas). Approximately 60 clones from each selection experiment were chosen at random and sequenced by Functional Biosciences Inc. (Madison, WI, USA).

Kinetic analyses

All rates were determined under single turnover conditions, in which a large excess of deoxyribozyme (~1.7 µM) was used with a trace amount of 5′-^{32}P-labeled substrate (~0.002 µM). Substrate and deoxyribozyme were heated together at 90°C for 30 s and allowed to cool at RT for ~10 min. An equal volume of 2× reaction buffer (200 mM KCl, 800 mM NaCl, 15 mM $MgCl_2$, 15 mM $MnCl_2$, 100 mM HEPES pH 7.0 at 23°C) was added to the deoxyribozyme/substrate mixture to initiate the reaction. The reaction was terminated after a designated period of time by the addition of quenching buffer containing 60 mM EDTA, 7 M urea and loading dye solution. The cleavage products from a reaction timecourse

were separated by denaturing 10% PAGE, and quantitated using PhosphorImager and ImageQuant software. A graph of fraction cleaved versus time was plotted for each timecourse, and the experimental data fit to a single exponential equation $Y = Y_{max}(1 - e^{(-kobs*t)})$ using nonlinear regression analysis in GraphPad Prism 4, from which the observed rate constant (k_{obs}) and maximum cleavage yield (Y_{max}) were determined. Kinetic parameters were determined from at least two independent experiments that differed by <30%.

Metal ion dependency

Simple assays based on single-timepoint measurements were used to determine the relative activity of a given deoxyribozyme under various 'nonwild-type' buffer conditions. The same general reaction conditions and protocol were used as described earlier; however, rates were approximated using the equation $k_{obs} = -\ln(1-\text{fraction cleaved})/\text{time}$. We have previously found that this method provides a reasonable estimation of the rate, when appropriate timepoints are chosen (i.e. prior to reaction plateau). Therefore, any reactions that exceeded 75% (typical reaction endpoints ranged from ~75% to 90%) cleavage product at an initial timepoint were subsequently assayed with a shorter timepoint(s) for accuracy. The amount of error incurred by this simplified assay is small relative to the overall range of activity being assayed, which spans several orders of magnitude. The relative activity was calculated as $k_{\text{non-WT buffer}}/k_{\text{WT buffer}}$. When necessary, slower reactions were followed for at least 72 h. If a distinct cleavage signal was not observed in this period of time, the rate was assumed to be the background rate of $\sim 10^{-7}$/min.

RESULTS

Library design

Six independent *in vitro* selection experiments were conducted in parallel using the library design illustrated in Figure 1A. The substrate domain contains a single ribonucleotide embedded in an otherwise all-DNA sequence. Chimeric (RNA/DNA) substrates were used instead of all-RNA substrates, so that specific dinucleotide junctions could be unambiguously targeted. Whereas the phosphodiester bonds in an all-RNA substrate are equally susceptible to hydrolytic cleavage, the target ribonucleotide in a chimeric substrate is ~100 000-fold more susceptible than the surrounding deoxyribonucleotides (20). The four types of pyr–pyr junctions (i.e. CC, UC, CT and UT) were our primary targets of interest, because efficient cleavage of these junctions has not been demonstrated. However, two additional selection experiments were also performed as controls, targeting AC and AT junctions, for which efficient RNA-cleaving deoxyribozymes (i.e. bipartite and 8-17) are known to exist (Table 1). The recurrence of these known RNA-cleaving deoxyribozymes would provide a measure of confidence in the effectiveness of our selection strategy and the reproducibility of our results.

The last five nucleotides at the 3′-end of the substrate domain differed between experiments, in order to minimize the possibility of cross-contamination during PCR amplification. The enzyme domain was designed with 20 random sequence nucleotides. The length of this random domain is not only short enough to cover all sequence permutations for a typical 8-17 motif measuring 14–15 nt, but also long enough to accommodate atypical 8-17 variants containing 1–2 nt bulges/insertions. All selection experiments started with a library of $\sim 3 \times 10^{14}$ molecules, so every possible sequence (i.e. 4^{20} or $\sim 10^{12}$ sequences) should be represented by ~300 copies on average. Multiple copy representation of each sequence should minimize the risk of losing deoxyribozymes through random experimental losses. Substrate recognition arms measuring 5 and 6 bp in length were designed into the sequence of the substrate and enzyme domains according to the scheme shown in Figure 1B. These base-pairing interactions are located several nucleotides downstream and upstream from the cleavage site, to provide some versatility in how potential deoxyribozymes engage the substrate. For instance, an 8-17 motif (Figure 1B, right structure) could use some of the excess random nucleotides to form additional stabilizing base-pairs with the substrate.

In vitro selection strategy

The *in vitro* selection scheme used in this study is depicted in Figure 2. In step 1, the sense primer (which serves as the substrate domain) is ligated to the enzyme domain, followed by PAGE purification in step 2. In step 3, the 73-nt full-length library molecules are incubated with the reaction buffer (7.5 mM $MgCl_2$, 7.5 mM $MnCl_2$, 100 mM KCl, 400 mM NaCl, 50 mM HEPES pH 7.0 at 23°C) for a designated period of time. In step 4, PAGE is used to separate the 58-nt cleavage products from the uncleaved precursor molecules. PCR amplification is conducted in step 5, which reintroduces the substrate domain via the sense primer. The antisense primer contains an internal 18 atom spacer that cannot be extended by DNA polymerase, leading to a dsDNA product composed of two strands of unequal length. In step 6, the 73-nt sense strand (containing the ribonucleotide cleavage junction) is separated from the 88-nt antisense strand through PAGE. Steps 3–6 are then repeated to enrich the population for functional deoxyribozyme sequences.

Selection progress

The percent cleavage product observed at each round of selection is illustrated in Figure 3. A total of nine rounds of selection were conducted in each experiment, and a distinct cleavage product (from 5% to 40%) was first observed by round 4. During the first four rounds of selection, a single reaction time of 100 min was used. In subsequent rounds, the selection stringency was increased by reducing the reaction time, in an effort to isolate the fastest deoxyribozymes within each population of sequences. In round 5, the main population was divided into three subpopulations that were subjected to 1, 10 and 100 min reaction times. Multiple reaction times were used to better gauge the response/progress of each population to a given time pressure. In rounds 6–9, reaction times of 0.1, 1 and 10 min were used. However, no appreciable cleavage

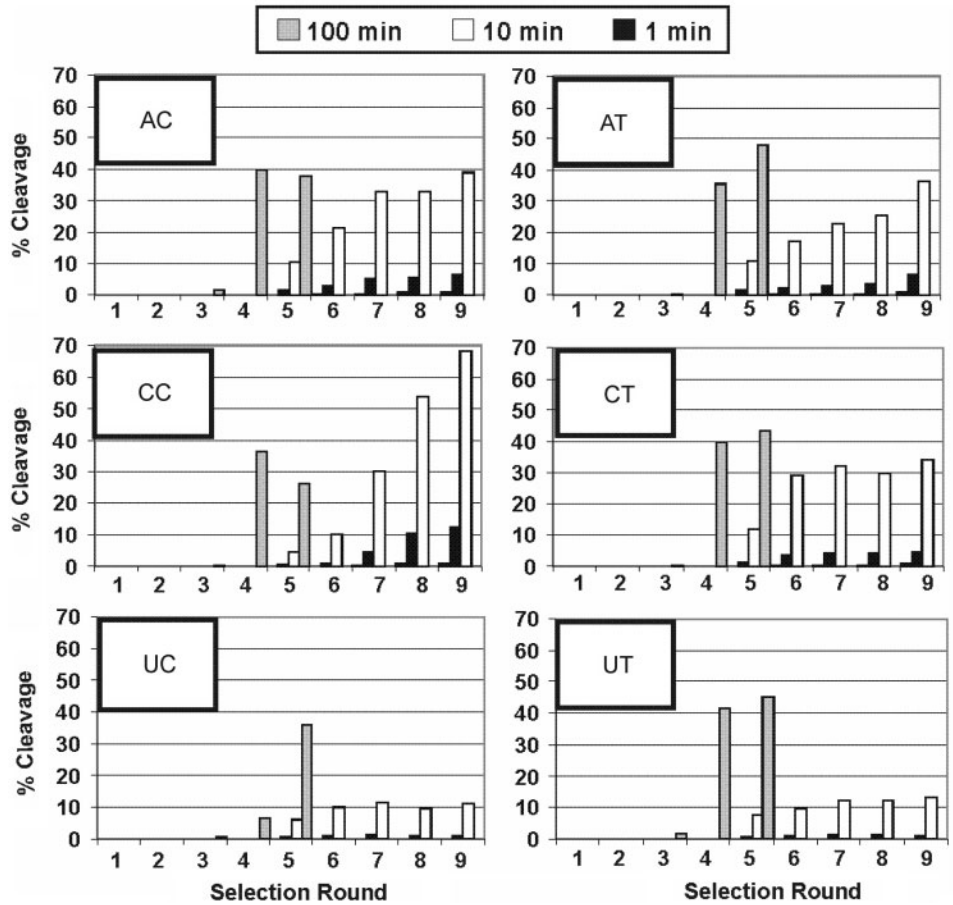

Figure 3. *In vitro* selection progress. The percent cleavage product is shown for each selection round. A 100 min selection time was used for the first four rounds of selection. In round 5, the population from each selection experiment was divided into subpopulations and subjected to a 1, 10 or 100 min selection time. In rounds 6–9, selection times of 0.1, 1 and 10 min were imposed. Less than 1% cleavage product was observed in the 0.1 min subpopulations. From rounds 5–9, the 1 min subpopulation from each selection experiment was used to seed each subsequent round of selection.

product was observed under the 0.1 min reaction time in any of the selection experiments. From rounds 5 to 9, the 1 min subpopulation in each selection experiment was used to seed (i.e. provide input molecules for) the next round of selection. By round 9 (or earlier), the composite population from each selection experiment could cleave its target dinucleotide junction with a rate of ~0.01–0.1/min. All experiments were terminated at round 9 after no appreciable increase in activity was observed for several rounds.

Deoxyribozyme sequence diversity and distribution

Approximately 60 clones from each selection experiment were randomly chosen for sequence analysis. Approximately half of these clones were taken from round 5 (100 min subpopulation) and the other half were taken from the terminal round 9 (10 min subpopulation). Clones from the 10 min subpopulation of round 9 were sequenced instead of the 1 min subpopulation, in order to minimize the likelihood of background contaminating sequences due to the low level of cleavage observed in some populations under the latter selection time. Twenty-six different sequence families were identified in total, of which only 12 survived until round 9. The number of sequence families per population varied from 3 to 10, and several sequence families were identified in two or three different selection experiments. The frequency and distribution of clones are summarized in Figure 4. The sequences of all clones are available in Supplementary Figure 1.

Recurrence of the 8-17 motif

We identified 28 different 8-17 sequence variants from 34 clones and grouped them into one sequence family (S24) for simplicity. Based on previous characterization studies (16,19,21), we have defined an 8-17 motif by the presence of four absolutely conserved bases and 4–6 co-varying nucleotides that form a putative 2–3 bp stem. The 8-17 variants were observed in the AC, CC and CT populations, with the majority (22 sequence variants from 27 clones) found in the AC population. Only three 8-17 sequence variants survived until round 9, and all were found in the AC control selection experiment, consistent with expectations. Eight sequence variants were chosen for kinetic analysis in *trans*, under single-turnover conditions. Three sequence variants from the round 9 AC population exhibited comparable rates to the best NC-cleaving 8-17 variant (i.e. '8-17NC') reported previously by our lab (19).

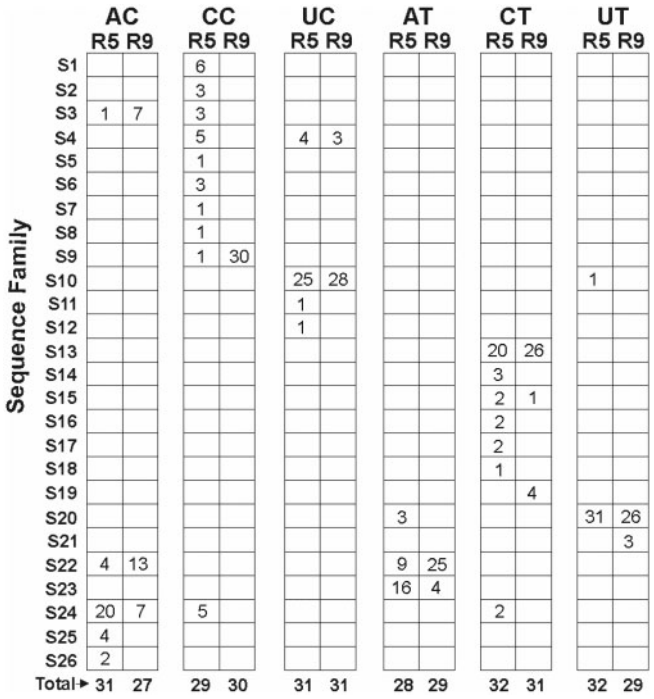

Figure 4. Frequency and distribution of *in vitro* selected deoxyribozymes. A total of 360 clones were sequenced and grouped into 26 different sequence families. The number of clones in each sequence family is shown, along with their distribution among the six selection experiments. Each selection experiment is denoted by a specific dinucleotide junction (AC, CC, UC, AT, CT or UT), and divided into round 5 (R5) and round 9 (R9). S24 contains the 8-17 motif. S22 contains the bipartite motif.

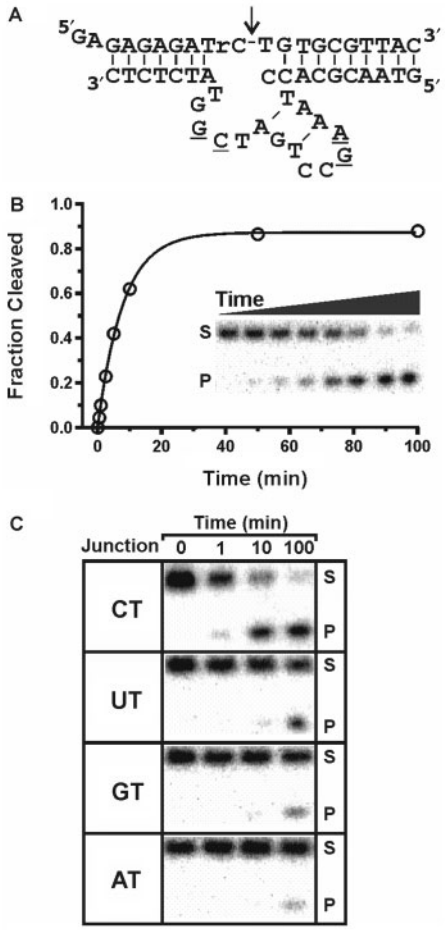

Figure 5. Characterization of the 8-17CT sequence variant. (**A**) Sequence and proposed secondary structure. Underlined nucleotides are absolutely conserved in the 8-17 motif and used to identify sequence variants. (**B**) Reaction time course conducted under single-turnover conditions. The $k_{obs} = 0.13$/min, $Y_{max} = 87\%$. Circles represent the average value of duplicate trials, for which standard error bars are smaller than circle diameter. Corresponding autoradiogram shows increase in cleavage product (P) as a function of time. Cleavage products (P) were separated from uncleaved substrates (S) by 10% denaturing PAGE. (**C**) Cleavage site selectivity. Autoradiogram depicting cleavage activity against CT, UT, GT and AT dinucleotide junctions.

It should be noted that the exact 8-17NC sequence variant was not isolated herein, but this is not surprising since many sequence variants are known to be phenotypically equivalent. The activities of two 8-17 sequence variants from each of the CC and CT populations (round 5) were also determined. Interestingly, while the rates of three of these 8-17 variants were comparable to our previous report, one 8-17 variant (denoted as 8-17CT) actually exhibited a ~1000-fold improvement with a rate of 0.13/min ($Y_{max} = 87\%$). The sequence and kinetic analysis of 8-17CT is shown in Figure 5. The 8-17CT contains a distorted stem (due to a mismatch in the central position), and a tetra-loop instead of a tri-loop, which are features also shared in common with 8-17NC described previously. The 8-17 variants have been shown to exhibit robust activity against groups of related dinucleotide junctions (e.g. NG group, where N = G, A, C or U) rather than any single junction. Contrary to expectations, however, 8-17CT was quite selective (from ~25- to 100-fold) for a CT junction over the related GT, AT and UT junctions (Figure 5). The 8-17CT was also tested against GC, AC, CC and UC junctions, but showed no activity in 100 min (data not shown).

Recurrence of the bipartite motif and identification of novel deoxyribozymes

Consistent with expectations, we identified the bipartite deoxyribozyme (S22) in both the AT and AC control selection experiments. It should be noted that the 10-23 motif was not reisolated in either control experiment, because it does not cleave chimeric substrates as proficiently as all-RNA substrates. The activity of 10-23 against chimeric substrates has not been previously reported; therefore, we tested two different 10-23 motifs optimized for cleavage of either AC or AU junctions (17), under our reaction conditions and with our chimeric AC and AT substrates. Only 2 and 8% cleavage product was observed in 100 min against the AC and AT junctions, respectively. These results indicate that the 10-23 deoxyribozyme would not be competitive under the imposed selection time pressure, and therefore unlikely to survive multiple rounds of selection.

The remaining 10 sequence families identified in round 9 represent novel deoxyribozymes that have not been previously reported. A representative sequence from each

Table 2. Sequence and rate constants of select deoxyriboyzmes

Family	5′ arm	N_{20} region	3′ arm	Junction	k_{obs} (min^{-1})	Y_{max} (%)
S3	GTAACGC---	AAATCCAGGGTTGGCCGACA	---TCTCTC	AC	0.03	75
S22	GTAACGC---	C**A**GGTAGGGGT**C**CCG**G**TC**A**	----TCTCTC	AT	0.09	88
S23	GTAACGC---	GGGTTCATCAAGGGGTATGG	---TCTCTC	AT	0.05	83
S9	GTAACGC---	TACTGCTTTACTGGCGGCCA	---TCTCTC	CC	0.12	90
S13	GTAACGC---	GGACAAGAGAAAGAGGTTGA	---TCTCTC	CT	0.05	87
S19	GTAACGC---	TTGGGGGGGAGAGGTGGGGA	---TCTCTC	CT	0.04	74
S15	GTAACGC---	AGTATATCAAGTGAATGGCA	---TCTCTC	CT	0.10	85
S10	GTAACGC---	CGGGGTTTGGGATGAAGGGG	---TCTCTC	UC	0.009	85
S4	GTAACGC---	TAGGGTGGGGTTAGAGTGGA	---TCTCTC	UC	0.04	76
S20	GTAACGC---	TAGGGATAGGGGATAGGGGG	---TCTCTC	UT	0.007	89
S21	GTAACGC---	GGGTGGATGGCGAGTGGCAG	---TCTCTC	UT	0.15	80
Sub.	3′	CATTGCGTG---	YrH	---TAGAGAGAG	5′	

The k_{obs} values represent the average of at least two independent trials, which differed by <30%. Reactions were conducted under single-turnover conditions in 100 mM KCl, 400 mM NaCl, 7.5 mM $MgCl_2$, 7.5 mM $MnCl_2$, 50 mM HEPES pH 7.0 at 23°C. The underlined region of S22 is identical to the bipartite consensus sequence, with exceptions denoted in boldface. The sequence of the substrate (sub.) is also provided, written 3′ to 5′ in opposite orientation. YrH represents the dinucleotide cleavage junction, where Y = C or T, and H = A, C or U.

family is illustrated in Table 2. Secondary structure formation within the random-sequence domain of each family appears limited, with contiguous base-pairing up to 3 bp (Supplementary Figure 2). The preengineered substrate binding arms made for the facile separation of substrate and enzyme domains into an intermolecular format, which was further simplified with the removal of excess sequence elements peripheral to the core region. The random sequence domain also did not harbor any strong sequence complementarity to the substrate that might compete with the preengineered binding arms. Each deoxyribozyme sequence was confirmed to be active in *trans* under single-turnover conditions against their primary (i.e. most abundant) target dinucleotide junction (Table 2). The reaction rate (k_{obs}) varied from 0.007 to 0.15/min, with maximum cleavage yields (Y_{max}) between 74% and 90%. From among these new sequence families, deoxyribozymes S9, S4, S15 and S21 exhibited the fastest k_{obs} values against CC (0.12/min), UC (0.04/min), CT (0.1/min) and UT (0.15/min) junctions, respectively. In comparison to the fastest rates previously reported and determined under identical reaction conditions (Table 1), S9 and S4 provide a 6- to 8-fold improvement, while S15 and S21 provide a 700- to 1600-fold improvement.

Clones from several sequence families (including S3, S4, S10, S20 and S22) were identified in two different selection experiments (Figure 4). The activity of each of these sequence families was therefore verified against the relevant secondary dinucleotide junction. Each family was indeed active, but generally slower, by up to ~20-fold in some cases.

Metal ion dependency

The 'wild-type' (WT) reaction buffer used during *in vitro* selection contained several monovalent and divalent metal ions including 100 mM KCl, 400 mM NaCl, 7.5 mM $MgCl_2$ and 7.5 mM $MnCl_2$ (in 50 mM HEPES pH 7.0). The effect of each buffer component on the activity of the new deoxyribozymes was systematically examined (Figure 6). Monovalent ions were generally unnecessary, although the rate of three deoxyribozymes (S4-UC, S10-UC and S20-UT) decreased by 10- to 100-fold when both K^+ and Na^+ were removed. None of the deoxyribozymes exhibited a strict dependence on Mg^{2+}. In contrast, a strong Mn^{2+}-dependency was observed for the majority of deoxyribozymes including (S3-AC, S23-AT, S9-CC, S13-CT, S15-CT, S19-CT, S4-UC and S21-UT), which was reflected by a 10^2- to 10^6-fold decrease in activity upon Mn^{2+} removal. In several cases (S22-AT, S10-UC and S20-UT), the presence of either Mg^{2+} or Mn^{2+} could support near WT levels of activity, although Mn^{2+} was always preferred by ~10-fold. Interestingly, the presence of monovalent ions alone could support some level of activity (up to ~100-fold) above background for S4-UC, S10-UC and S20-UT. The S10-UC deoxyribozyme exhibited a near WT level of activity (only ~10-fold less) in the presence of 2 M NaCl. This deoxyribozyme bears little sequence homology to the G3 metal-independent deoxyribozymes reported by Sen and colleagues (22).

Cleavage activity against all-RNA substrates

The cleavage activity of each pyr–pyr cleaving deoxyribozyme from Table 2 was evaluated against all-RNA substrates, using the same reaction conditions and format described previously. Each substrate was identical to the substrate sequence shown in Table 2, except for the addition of two G residues on the 5′-end (to increase RNA transcript yield). Deoxyribozymes S9, 8-17CT, S4, S20 and S21 showed detectable but severely reduced (~1000-fold lower) activity against all-RNA substrates (Supplementary Figure 3). The remaining deoxyribozymes did not show detectable cleavage activity within a 72 h time period.

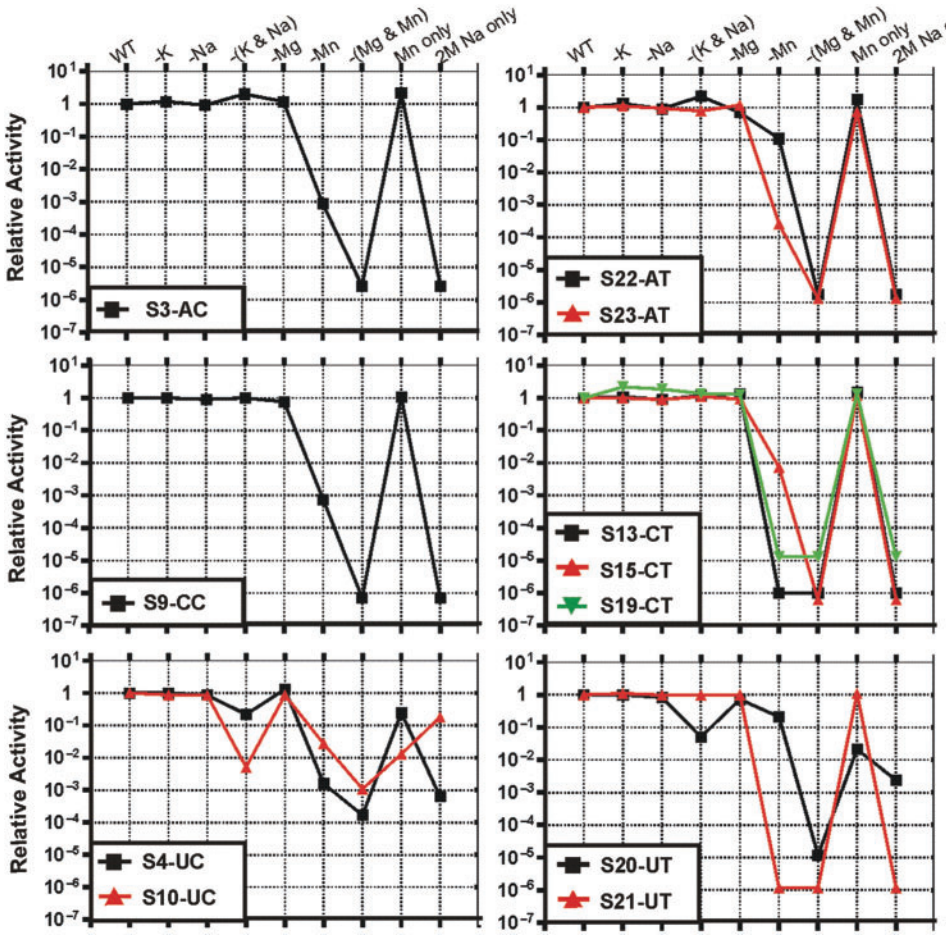

Figure 6. Metal ion dependency of select deoxyribozymes. The sequence class and target dinucleotide junction is provided in each graph. Relative activity = $k_{Mutant,buffer}/k_{WT,buffer}$. WT buffer contains 100 mM KCl, 400 mM NaCl, 7.5 mM $MgCl_2$, 7.5 mM $MnCl_2$ (and 50 mM HEPES pH 7.0 at 23°C). Each metal ion was systematically removed individually, or in combination, as indicated above each graph. 'Mn only' refers to a buffer containing only 7.5 mM $MnCl_2$. '2M Na only' refers to a buffer containing only 2 M NaCl.

DISCUSSION

Herein, we have demonstrated that all four chimeric RNA/DNA, pyr–pyr junctions, can be efficiently cleaved by small deoxyribozymes. From a total of 245 deoxyribozyme clones isolated in four separate *in vitro* selection experiments, we identified 22 different sequence families that can cleave one or more pyr–pyr junctions. Twenty-one of these sequence families have not been previously described. The best deoxyribozymes (S9, S4, 8-17CT and S21) can collectively cleave all four pyr–pyr junctions with rates ranging from 0.04 to 0.15/min, which are ~1–3 orders of magnitude faster than the best deoxyribozymes reported previously (Table 1). From among 115 clones isolated in two additional control experiments, targeting AC and AT junctions, we identified seven different deoxyribozyme sequence families including variants of the 8-17 and bipartite motifs. The remaining five sequence families represent novel motifs, but they do not provide any significant rate improvements over 8-17 or bipartite.

All of the deoxyribozymes isolated herein are 10- to 100-fold slower than other deoxyribozymes of comparable size (Table 1) that catalyze the cleavage of purine–purine junctions. Therefore, this study strengthens our previous suggestion that pyr–pyr junctions may be inherently more difficult to cleave (16,19). We have speculated that base-stacking interactions at the dinucleotide junction may play an important role in the cleavage mechanism of small deoxyribozymes, since purine bases can form stronger stacking interactions than pyrimidine bases (23–25).

This study also provides new insight into the functional versatility of the well known 8-17 deoxyribozyme. We have previously speculated that the structure of this deoxyribozyme might be flexible enough to support fast cleavage of all dinucleotide junctions, since only 4 of the 15 nucleotides that form its structural scaffold are catalytically essential. Because there are thousands of possible sequence variations (with unknown functional implications), a systematic and comprehensive mutational analysis would not be practical. However, by using a library containing only 20 random-sequence positions, we were able to sample every 20-nt sequence permutation in each selection experiment, including all 8-17 sequence variations. Consistent with expectations, we found 8-17 variants in the AC and CC populations that exhibited comparable rates to those described previously. No 8-17 variants were found in the AT, UC or UT

selection experiments. However, we did identify one 8-17 variant (8-17CT) in the CT population that was ~1000-fold faster than the best 8-17 variant previously reported for the cleavage of this junction.

These results should be considered in the design of future *in vitro* selection experiments. A substantial effort has been expended to search for new and more useful RNA-cleaving deoxyribozymes, and in at least five independent studies (16,26–30), the recurrence of 8-17 has hampered this objective. Since UC and UT junctions are the least susceptible to 8-17-mediated cleavage, these dinucleotide junctions represent the best target cleavage sites to search for novel motifs, using larger random sequence domains. The likelihood of recurrence of other motifs described herein will depend on several factors (31), but in general, is expected to be lower due to the higher degree of sequence conservation in their putative catalytic cores (Supplementary Figure 1).

Limitations

It is important to recognize that the best possible solutions to a given selection experiment (typically presumed to be the most abundant sequences) might not always have the desired target phenotype. For instance, we did not always observe a direct correlation between the abundance of a given clone and its catalytic rate, as in the case of S20 and S21. Interestingly, S20 was about ~20-fold more abundant than S21, but ~20-fold slower than S21, contrary to expectations. This discrepancy may be due to the fact that rates were determined using a simplified intermolecular format (in which some fixed sequence regions were removed), and therefore may not reflect the activities of the full-length *cis* constructs that actually competed during selection. We tested this possibility, and determined that the activity of each full-length *cis* construct (S21 $k_{obs} = 0.056$/min, $Y_{max} = 46\%$; S20 $k_{obs} = 0.036$/min, $Y_{max} = 22\%$) differs from the activity determined by the simplified intermolecular format (Table 2). The difference in activity between the full-length *cis* constructs (~1.5-fold) is smaller than the difference observed between the simplified *trans* constructs (~20-fold); however, S21 still exhibits superior cleavage activity, which does not readily explain the discrepancy in their relative abundance.

Our previous study of *in vitro* selection population dynamics has suggested how secondary fitness constraints imposed unintentionally on the population, can unpredictably influence the relative abundance of different catalytic species (32). This scenario can enable sequences with non-target phenotypes (e.g. better PCR amplification efficiency than catalytic efficiency) to increase in abundance at the expense of those with target phenotypes. However, we have used several general strategies to try and mitigate the effects of this scenario including: (i) conducting control experiments under identical conditions; (ii) applying stringent time pressure to favor the target phenotype; (iii) performing extensive sequence analysis of each population at two different stages in the selection to minimize sampling artifacts and (iv) determining the rates of both major and minor sequence classes.

The practical utility of the deoxyribozymes described herein is currently limited by their inability to cleave all-RNA substrates as effectively as the chimeric substrates against which they were selected. Nevertheless, these catalytic DNA molecules can potentially serve as good starting points from which to evolve and/or engineer improved activity against pyr–pyr junctions within all-RNA substrates.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This work was supported by a research grant from the Canadian Institutes of Health Research (grant number MOP37964). Y.L. is a Canada research chair. K.S. holds an Natural Sciences and Engineering Research Council Doctoral Canada Graduate Scholarship. Funding to pay the Open Access publication charges for this article was provided by the Canadian Institutes of Health Research.

Conflict of interest statement. None declared.

REFERENCES

1. Peracchi,A. (2005) DNA catalysis: potential, limitations, open questions. *Chembiochem.*, **6**, 1316–1322.
2. Achenbach,J.C., Chiuman,W., Cruz,R.P. and Li,Y. (2004) DNAzymes: from creation in vitro to application in vivo. *Curr. Pharm. Biotechnol.*, **5**, 321–336.
3. Silverman,S.K. (2005) In vitro selection, characterization, and application of deoxyribozymes that cleave RNA. *Nucleic Acids Res.*, **33**, 6151–6163.
4. Liu,J., Brown,A.K., Meng,X., Cropek,D.M., Istok,J.D., Watson,D.B. and Lu,Y. (2007) A catalytic beacon sensor for uranium with parts-per-trillion sensitivity and millionfold selectivity. *Proc. Natl Acad. Sci. USA*, **104**, 2056–2061.
5. Liu,J. and Lu,Y. (2007) A DNAzyme catalytic beacon sensor for paramagnetic Cu2+ ions in aqueous solution with high sensitivity and selectivity. *J. Am. Chem. Soc.*, **129**, 9838–9839.
6. Liu,J. and Lu,Y. (2003) A colorimetric lead biosensor using DNAzyme-directed assembly of gold nanoparticles. *J. Am. Chem. Soc.*, **125**, 6642–6643.
7. Chiuman,W. and Li,Y. (2007) Efficient signaling platforms built from a small catalytic DNA and doubly labeled fluorogenic substrates. *Nucleic Acids Res.*, **35**, 401–405.
8. Stojanovic,M.N., Mitchell,T.E. and Stefanovic,D. (2002) Deoxyribozyme-based logic gates. *J. Am. Chem. Soc.*, **124**, 3555–3561.
9. Stojanovic,M.N. and Stefanovic,D. (2003) Deoxyribozyme-based half-adder. *J. Am. Chem. Soc.*, **125**, 6673–6676.
10. Chen,Y., Wang,M. and Mao,C. (2004) An autonomous DNA nanomotor powered by a DNA enzyme. *Angew. Chem. Int. Ed. Engl.*, **43**, 3554–3557.
11. Dass,C.R., Choong,P.F. and Khachigian,L.M. (2008) DNAzyme technology and cancer therapy: cleave and let die. *Mol. Cancer Ther.*, **7**, 243–251.
12. Khachigian,L.M. (2000) Catalytic DNAs as potential therapeutic agents and sequence-specific molecular tools to dissect biological function. *J. Clin. Invest*, **106**, 1189–1195.
13. Peracchi,A. (2004) Prospects for antiviral ribozymes and deoxyribozymes. *Rev. Med. Virol.*, **14**, 47–64.
14. Pyle,A.M., Chu,V.T., Jankowsky,E. and Boudvillain,M. (2000) Using DNAzymes to cut, process, and map RNA molecules for structural studies or modification. *Methods Enzymol.*, **317**, 140–146.

15. Cairns,M.J., Hopkins,T.M., Witherington,C., Wang,L. and Sun,L.Q. (1999) Target site selection for an RNA-cleaving catalytic DNA. *Nat. Biotechnol.*, **17**, 480–486.
16. Cruz,R.P., Withers,J.B. and Li,Y. (2004) Dinucleotide junction cleavage versatility of 8-17 deoxyribozyme. *Chem. Biol.*, **11**, 57–67.
17. Cairns,M.J., King,A. and Sun,L.Q. (2003) Optimisation of the 10-23 DNAzyme-substrate pairing interactions enhanced RNA cleavage activity at purine-cytosine target sites. *Nucleic Acids Res.*, **31**, 2883–2889.
18. Feldman,A.R. and Sen,D. (2001) A new and efficient DNA enzyme for the sequence-specific cleavage of RNA. *J. Mol. Biol.*, **313**, 283–294.
19. Schlosser,K., Gu,J., Sule,L. and Li,Y. (2008) Sequence-function relationships provide new insight into the cleavage site selectivity of the 8-17 RNA-cleaving deoxyribozyme. *Nucleic Acids Res.*, **36**, 1472–1481.
20. Li,Y. and Breaker,R.R. (1999) Kinetics for specific base catalysis of RNA degradation by transesterification involving the 2′-hydroxyl group. *J. Am. Chem. Soc.*, **121**, 5364–5372.
21. Peracchi,A., Bonaccio,M. and Clerici,M. (2005) A mutational analysis of the 8-17 deoxyribozyme core. *J. Mol. Biol.*, **352**, 783–794.
22. Geyer,C.R. and Sen,D. (1997) Evidence for the metal-cofactor independence of an RNA phosphodiester-cleaving DNA enzyme. *Chem. Biol.*, **4**, 579–593.
23. Solie,T.N. and Schellman,J.A. (1968) The interaction of nucleosides in aqueous solution. *J. Mol. Biol.*, **33**, 61–77.
24. Nakano,N.I. and Igarashi,S.J. (1970) Molecular interactions of pyrimidines, purines, and some other heteroaromatic compounds in aqueous media. *Biochemistry*, **9**, 577–583.
25. Guckian,K.M., Schweitzer,B.A., Ren,R.X.-F., Sheils,C.J., Tahmassebi,D.C. and Kool,E.T. (2000) Factors contributing to aromatic stacking in water: evaluation in the context of DNA. *J. Am. Chem. Soc.*, **122**, 2213–2222.
26. Faulhammer,D. and Famulok,M. (1996) The Ca2+ ion as a cofactor for a novel RNA-cleaving deoxyribozyme. *Angew. Chem. Int. Ed. Engl.*, **35**, 2809–2813.
27. Li,J., Zheng,W., Kwon,A.H. and Lu,Y. (2000) In vitro selection and characterization of a highly efficient Zn(II)-dependent RNA-cleaving deoxyribozyme. *Nucleic Acids Res.*, **28**, 481–488.
28. Peracchi,A. (2000) Preferential activation of the 8-17 deoxyribozyme by Ca(2+) ions. Evidence for the identity of 8-17 with the catalytic domain of the Mg5 deoxyribozyme. *J. Biol. Chem.*, **275**, 11693–11697.
29. Santoro,S.W. and Joyce,G.F. (1997) A general purpose RNA-cleaving DNA enzyme. *Proc. Natl Acad. Sci. USA*, **94**, 4262–4266.
30. Schlosser,K. and Li,Y. (2004) Tracing sequence diversity change of RNA-cleaving deoxyribozymes under increasing selection pressure during in vitro selection. *Biochemistry*, **43**, 9695–9707.
31. Lehman,N. (2004) Assessing the likelihood of recurrence during RNA evolution in vitro. *Artif. Life*, **10**, 1–22.
32. Schlosser,K. and Li,Y. (2005) Diverse evolutionary trajectories characterize a community of RNA-cleaving deoxyribozymes: a case study into the population dynamics of in vitro selection. *J. Mol. Evol.*, **61**, 192–206.