

Controlling the direction of site-selectivity and regioselectivity in RNA ligation by Zn²⁺-dependent deoxyribozymes that use 2',3'-cyclic phosphate RNA substrates†

Diana M. Kost,‡ Joseph P. Gerdt,‡ P. I. Pradeepkumar§ and Scott K. Silverman*

Received 5th August 2008, Accepted 1st September 2008

First published as an Advance Article on the web 27th October 2008

DOI: 10.1039/b813566e

Our previous efforts have used *in vitro* selection to identify numerous Zn²⁺-dependent deoxyribozymes that ligate two RNA substrates with reaction at a 2',3'-cyclic phosphate. Each deoxyribozyme creates one of several different RNA linkages, including native 3'-5' and non-native 2'-5' phosphodiester bonds as well as many unnatural linkages. In this report, we describe experiments to reveal design aspects of the selection strategy that favor site-selective and regioselective synthesis of native 3'-5' RNA linkages. The results also reveal that an explicit selection pressure for RNA substrate sequence generality must be developed if the deoxyribozymes are to have practical generality.

Introduction

We previously established an *in vitro* selection strategy to identify deoxyribozymes (DNA enzymes)^{1,2} that ligate two RNA substrates.³ These substrates are one of two readily obtained combinations: 2',3'-cyclic phosphate with 5'-hydroxyl (Scheme 1),³⁻⁵ or 2',3'-diol with 5'-triphosphate.⁶⁻⁸ For both substrate combinations, multiple ligation junctions are possible, leading to several selectivity considerations. Through the experiments reported here, we sought to understand how to control both site-selectivity and regioselectivity in the deoxyribozyme-mediated RNA ligation reaction involving a 2',3'-cyclic phosphate substrate. Our findings indicate that several subtle aspects of the selection design are important to provide native 3'-5' RNA linkages in the ligation products.

Our first report on RNA-ligating deoxyribozymes that use a 2',3'-cyclic phosphate substrate revealed only formation of non-native 2'-5' linkages rather than native 3'-5' linkages.³ Subsequent experiments showed that the non-native connectivity is favored in several different contexts, for mechanistic reasons that remain unclear.^{4,5} All of these early findings used Mg²⁺ as the metal ion cofactor for the DNA enzymes. We later discovered that using Zn²⁺ in place of Mg²⁺ leads to deoxyribozymes that collectively create the various ligation products of Scheme 1, including formation of native 3'-5' linkages (Scheme 1b).⁹ Such linkages are desirable for practical RNA ligation because no evidence remains in the RNA product that a ligation event had taken place.

In all of these cases, each deoxyribozyme creates only one particular RNA linkage and is therefore both site-selective and regioselective. The direction of site-selectivity for an individual

deoxyribozyme is controlled by which functional group acts as the attacking nucleophile: the 5'-OH, the 2'-OH of the 5'-terminal nucleotide, or an internal 2'-OH of the second substrate. The direction of regioselectivity is determined by which O atom of the cyclic phosphate, 2' or 3', acts as the leaving group. During *in vitro* selection of our previously reported Zn²⁺-dependent deoxyribozymes, the directions of both site-selectivity and regioselectivity were controlled poorly, because examples were found of individual deoxyribozymes that collectively form at least five out of the 2 × 3 = 6 possible ligation products. This includes all four of the linkages formed by nucleophilic attack of either the 5'-OH or 2'-OH of the 5'-terminal nucleotide combined with both possible openings of the 2',3'-cyclic phosphate (Scheme 1b and 1c).

These Zn²⁺-dependent deoxyribozymes were identified using two selection strategies whose key steps are shown in Fig. 1a and b.⁹ The selections were designated as 'J' and 'BB' according to our laboratory's ongoing alphabetical nomenclature. Two key aspects of the J selection design were as follows. (1) The presence of four-nucleotide RNA overhangs on both sides of the ligation junction, beyond which DNA:RNA base pairs were formed between the fixed deoxyribozyme binding arms and the RNA substrates. (2) An unmodified 'right-hand' (R) RNA substrate that reacts with the 2',3'-cyclic phosphate of the 'left-hand' (L) RNA substrate. The J selection led to deoxyribozymes that use only the 5'-terminal guanosine 2'-OH group as the nucleophile, with two unnatural ligation products formed *via* both possible cyclic phosphate opening reactions (*i.e.*, linear 3'-2' and 2'-2' linkages; Scheme 1c). Thus, the J selection was site-selective (albeit for the undesired 5'-terminal 2'-OH nucleophile) but not regioselective.

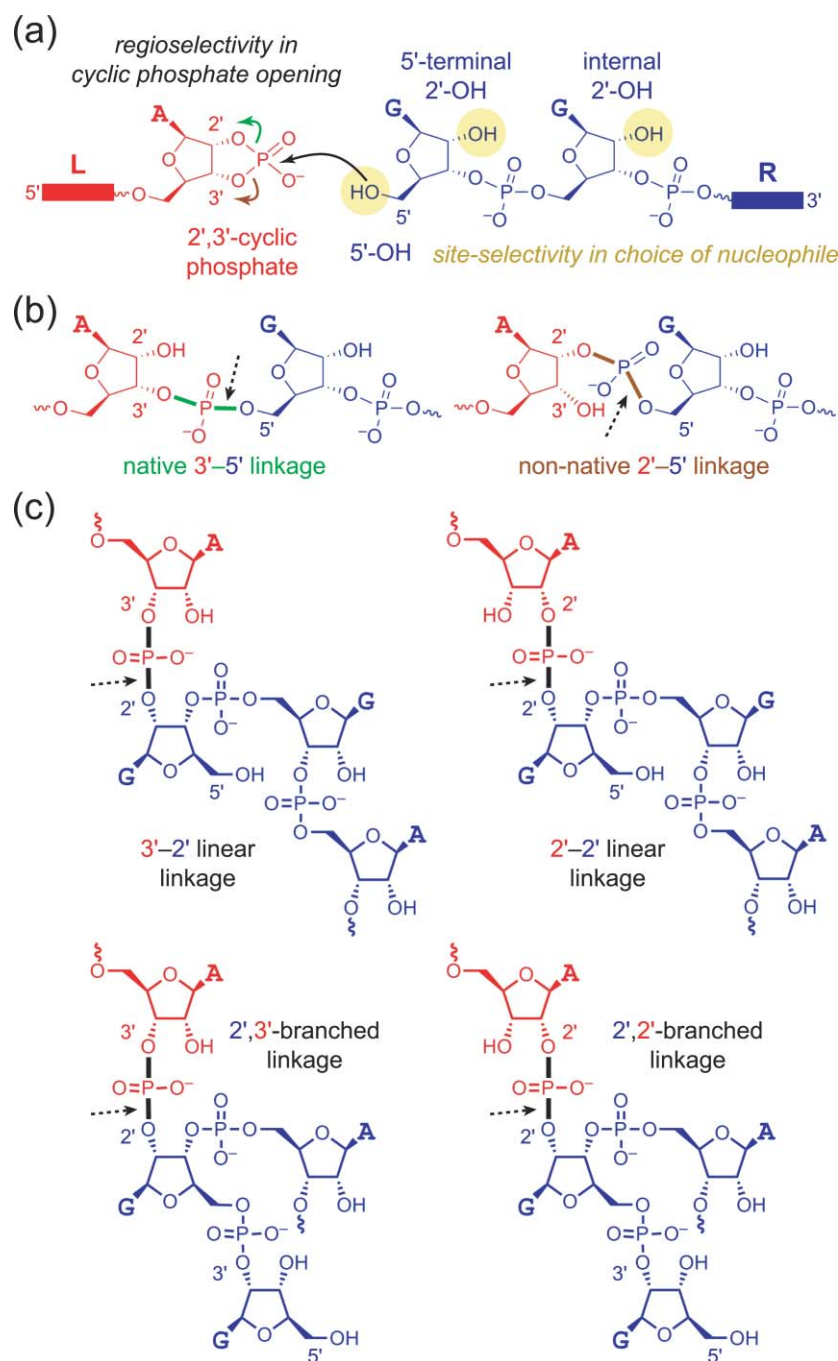
Because both of the products formed by deoxyribozymes from the J selection were unnatural RNA linkages, we changed the selection strategy in an attempt to favor formation of native 3'-5' linkages. The ensuing BB selection was identical to the J selection except that the R substrate had a 2'-deoxy-G nucleotide at its 5'-terminus rather than a standard G ribonucleotide (*i.e.*, 2'-dG rather than 2'-G). This change was made to prevent emergence of deoxyribozymes that require nucleophilic reactivity of the 5'-terminal 2'-OH group of R (as was observed in the J selection),

Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois, 61801, USA. E-mail: scott@scs.uiuc.edu; Fax: 217 244 8024; Tel: 217 244 4489

† Electronic supplementary information (ESI) available: Deoxyribozyme sequences. See DOI: 10.1039/b813566e

‡ D.M.K. and J.P.G. are co-first authors of this manuscript.

§ Present address: Department of Chemistry, Indian Institute of Technology, Bombay, India.



Scheme 1 Potential DNA-catalyzed ligation reactions of a 2',3'-cyclic phosphate RNA substrate. (a) The two RNA reaction partners. The strand providing the cyclic phosphate is designated as the 'left-hand' (L) substrate (red), and the strand providing the -OH nucleophile is designated as the 'right-hand' (R) substrate (blue). The 2',3'-cyclic phosphate can be opened in either of two ways (green and brown arrows); controlling the regioselectivity of this opening is one main focus of this study. In addition, three potential -OH group nucleophiles are highlighted (yellow) on the R substrate; controlling the site-selectivity among these competing nucleophiles is another main focus of this study. (b) Nucleophilic attack by the 5'-OH of R with the 2'-O of L acting as the leaving group leads to the native 3'-5' RNA linkage. When the 3'-O of L is instead the leaving group, the non-native 2'-5' RNA linkage is formed. In each structure, the dashed arrow marks the bond formed between the 5'-OH and the P atom. (c) Various unnatural linkages are formed when either the 5'-terminal 2'-OH or the internal 2'-OH of R attacks the cyclic phosphate.

because this functional group was consistently absent during the BB selection process. Indeed, the BB selection led to deoxyribozymes that use either the 5'-terminal 5'-OH or the internal 2'-OH of the second R nucleotide (but not the 5'-terminal 2'-OH) as the nucleophile, again with both cyclic phosphate openings. Although one particular product among this collection was

the desired native 3'-5' linkage, overall the redesigned selection experiment did not control either the site-selectivity or the regioselectivity of the RNA ligation process.

In the present work, we explored how the selection design can be further modified to control the direction of site-selectivity and regioselectivity during DNA-catalyzed RNA ligation. We

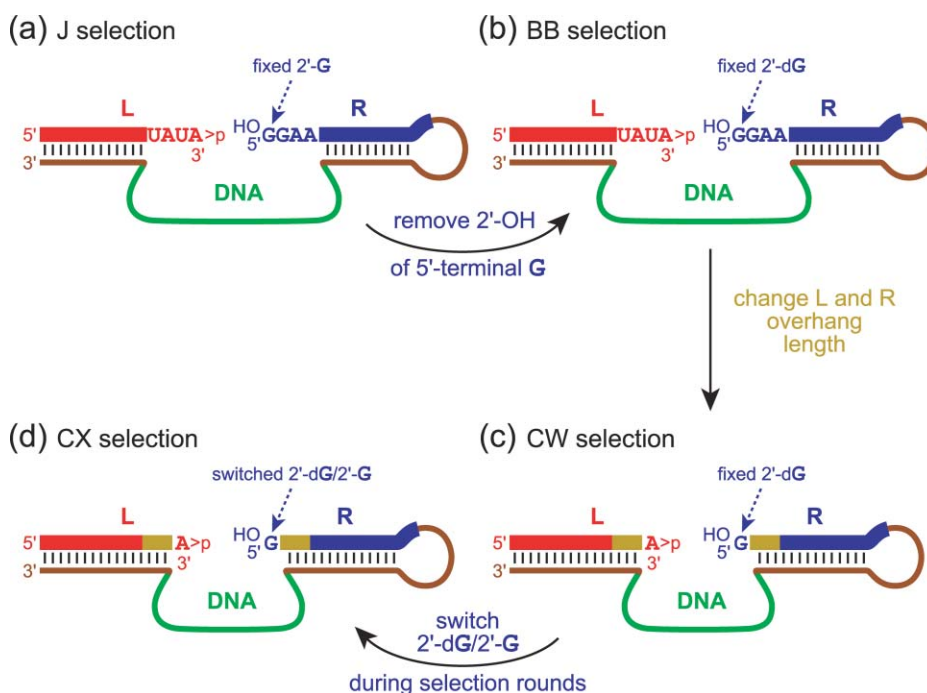


Fig. 1 Selection designs for deoxyribozyme-catalyzed RNA ligation using a 2',3'-cyclic phosphate RNA substrate. In each case, the two RNA substrates (L and R) and the deoxyribozyme are shown poised for the key selection step (see Fig. 2). (a) and (b) The previously reported J and BB selections, in which four overhanging RNA nucleotides were present on both the L and R substrates.⁹ The 5'-terminal nucleotide of R was fixed as either G with a 2'-OH (2'-G; J selection) or G with a 2'-H (2'-dG; BB selection). (c) The new CW selection, which is the same as BB except only one overhanging RNA nucleotide was present on both L and R. (d) The new CX selection, which differs from BB *via* both the overhang length and in switching of 2'-dG/2'-G in various selection rounds.

focused on strategically changing two particular aspects of the J/BB selection design. First, in the new CW selection (Fig. 1c) we altered the RNA overhang length from four nucleotides to only one nucleotide. The reduced overhang length should help to suppress nucleophilic reactivity by the internal 2'-OH of R, because this nucleotide is now physically constrained within a DNA:RNA Watson–Crick base pair; this change should therefore contribute to controlling the site-selectivity of ligation. Another benefit of reducing the RNA overhang length on both substrates is an anticipated increase in RNA substrate sequence generality of the resulting deoxyribozymes, because we have observed that long single-stranded regions of the RNA substrates often lead to deoxyribozymes that have undesired requirements for specific RNA nucleotides in these non-duplex regions.^{3,6,9–11} RNA substrate generality is an important practical feature of useful RNA-ligating deoxyribozymes.

The second altered aspect of the selection design was to repeatedly switch the 5'-terminal nucleotide of the R substrate between 2'-dG and 2'-G during the iterated rounds of the new CX selection (Fig. 1d). By switching this nucleotide, we expected to continue to suppress nucleophilic reactivity of the corresponding 5'-terminal 2'-OH of R. Importantly, we also expected this switching to provide a direct pressure for the resulting deoxyribozymes to retain substantial ligation activity when the 2'-OH is present on the RNA substrate. The previous BB deoxyribozymes, which were identified by selection with the fixed 2'-dG substrate, catalyzed RNA ligation relatively poorly with the 2'-G substrate when compared with the 2'-dG substrate.⁹

By assaying a collection of the Zn²⁺-dependent deoxyribozymes that emerged from both of the new CW and CX selection experiments, we determined how both altered aspects of the selection design influenced the site-selectivity and regioselectivity of DNA-catalyzed RNA ligation. We were pleased to find that changing both design aspects simultaneously (in the CX selection) led to control of both site-selectivity and regioselectivity, because nearly all of the resulting deoxyribozymes create solely native 3'–5' RNA linkages. We also examined the RNA substrate sequence generality of deoxyribozymes that form native 3'–5' linkages from both new selections. The observation of nontrivial sequence requirements near the ligation junction suggests that further experiments to obtain highly general deoxyribozymes will require the introduction of a selection pressure specifically designed to induce such generality.

Results

In vitro selection of new deoxyribozymes

For both of the new selection experiments designated CW and CX (Fig. 1), the procedure was analogous to that employed previously for the J and BB selections.⁹ The R substrate was attached covalently to the deoxyribozyme strand using T4 RNA ligase. Then, in the key selection step (Fig. 2a), the L and R substrates were provided the opportunity to become ligated in 70 mM Tris, pH 7.5, 150 mM NaCl, 2 mM KCl, and 1 mM ZnCl₂ by incubation for 3 h at 23 °C. Those DNA sequences that

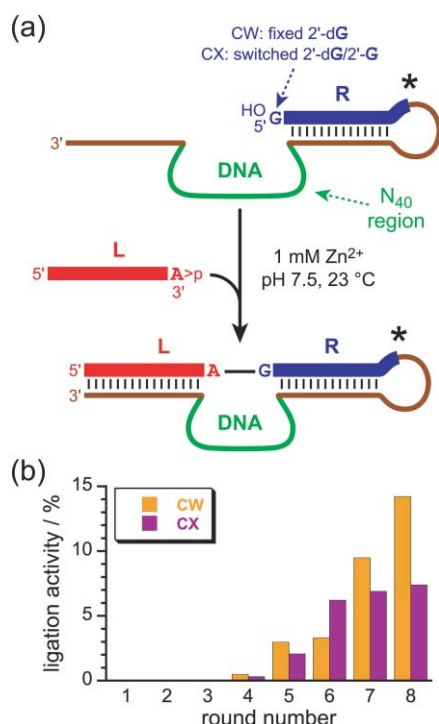


Fig. 2 Key selection step and progression of the new selection experiments. (a) The key selection step, in which the L and R substrates are joined by the action of the deoxyribozyme that is attached to R (asterisk). Note that this covalent attachment is required only during the selection process itself; L and R can be provided to the deoxyribozyme as separate substrates for analysis of ligation activity after the selection process is completed. (b) Ligation activities of the CW and CX selections by round number.

successfully joined L and R were separable by 8% polyacrylamide gel electrophoresis (PAGE) due to the pre-existing covalent connection between R and the deoxyribozyme (*i.e.*, L+R+DNA was readily separated from R+DNA). Finally, the DNA pool (now enriched in catalytically active sequences) was regenerated by PCR, and the two single strands were separated by denaturing PAGE; the undesired strand was longer than the desired strand due to the presence of a nonamplifiable spacer within the DNA primer that initiates synthesis of the undesired strand. This process was iterated for eight rounds, with the activity progression shown in Fig. 2b.

For the CW selection, the only difference from the prior BB selection was that only one overhanging RNA nucleotide (rather than four) was present on both L and R. As in the BB selection, the 2'-dG R substrate was used in all CW rounds. For the CX selection, only one overhanging nucleotide was present on both L and R; in addition, the 5'-terminal nucleotide of the R substrate was switched between 2'-dG and 2'-G in various rounds. The CX selection began in round 1 with the 2'-dG substrate and used this substrate in rounds 2 and 4; the alternative 2'-G substrate was used in rounds 3 and 5–8.

Deoxyribozymes from the CX selection

We first analyzed individual deoxyribozymes from the CX selection, in which both aspects of the selection design (overhang length and switching of 2'-dG/2'-G) were altered relative to the

previous BB selection. The ligation activity of the uncloned CX pool was *ca.* 7% (Fig. 2b), which is relatively low. Nonetheless, we have obtained highly active deoxyribozymes from such pools in many past cases (note that individual DNA enzymes can have much higher ligation yield than the pool average). We cloned deoxyribozymes from round 7CX using standard procedures that were adapted in straightforward fashion from our past selection experiments.^{3,12} Briefly, these procedures involved (1) insertion of the double-stranded PCR product (purified on agarose gel) into a cloning vector; (2) screening of 'miniprep' DNA from individual *E. coli* colonies by restriction digestion to verify the presence of the expected insert; (3) PCR using the miniprep DNA as a template to prepare individual deoxyribozyme clones; and (4) trimolecular RNA ligation assays for each clone using the L and R substrates, with the L substrate bearing a 5'-³²P radiolabel.

Seven 7CX clones with promising ligation activities were sequenced,[†] and the deoxyribozymes were prepared independently by solid-phase DNA synthesis. Quantitative RNA ligation assays (Fig. 3) revealed that six out of seven 7CX deoxyribozymes have ligation yield in the 45–60% range, reaching maximal yield with $t_{1/2}$ of 20–60 min (k_{obs} 0.01–0.03 min⁻¹). Because the ligation reaction of a 2',3'-cyclic phosphate is inherently reversible, the reverse cleavage reaction can occur, and this likely reduces the net ligation yield below the maximal 100%.

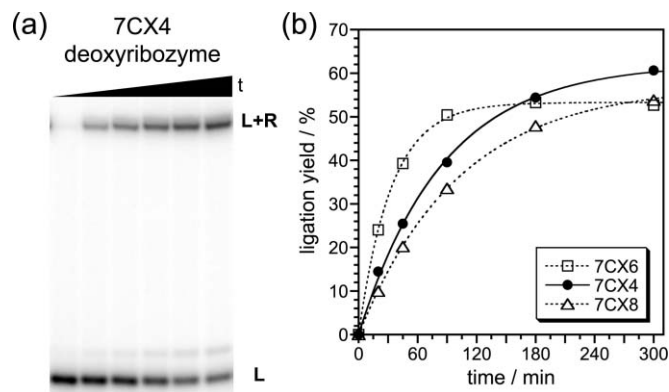


Fig. 3 Catalytic activities of the 7CX deoxyribozymes. (a) PAGE image using 7CX4 as a representative example. (b) Kinetic plot for 7CX6, 7CX4, and 7CX8 (k_{obs} = 0.030, 0.012, and 0.010 min⁻¹).

Our key interest in the new deoxyribozymes is in the junction within their RNA ligation products. Because of the reversible nature of 2',3'-cyclic phosphate ring opening, the principle of microscopic reversibility enables the rapid determination of the ligation junction formed by each new DNA enzyme *via* assaying for cleavage of suitably linked RNA 'product' strands. For example, any 7CX deoxyribozyme that joins the L and R substrates with formation of a native 3'-5' RNA linkage should also be observed to cleave the 3'-5' linkage of a separately synthesized product strand. Similarly, if the same deoxyribozyme does NOT create the non-native 2'-5' linkage (*i.e.*, is highly selective for 3'-5' linkage formation), then it should not be observed to cleave the analogous product strand that has a 2'-5'-linkage at the cleavage site. The opposite experimental outcome, cleavage of the 2'-5' but not 3'-5' product strand, would indicate a deoxyribozyme that creates 2'-5' linkages. Finally, cleavage of neither the 3'-5' nor

2'-5' product strand would suggest a deoxyribozyme that creates an unnatural RNA linkage formed by attack of a nucleophile other than the 5'-OH of the R substrate (*i.e.*, the terminal or internal 2'-OH; see Scheme 1a). Because this third outcome would constitute negative data (unlike for successful cleavage of either the 3'-5'-linked or 2'-5'-linked product strands), further experimental confirmation of the unnatural RNA linkage assignment would be necessary.

We tested each of the seven 7CX deoxyribozymes for their ability to cleave product strands that have either a 3'-5' or a 2'-5' linkage directly at the ligation site (Fig. 4). In all seven cases, the 7CX deoxyribozyme cleaved only the 3'-5'-linked strand, as illustrated for 7CX4 along with assays of the previously reported⁹ 12BB1 (here a positive control for cleaving the 2'-5'-linked product) and 12BB5 (a positive control for cleaving the 3'-5'-linked product). On the basis of these data, we concluded that all seven of the tested 7CX deoxyribozymes create native 3'-5' RNA linkages. This finding is in sharp contrast to the results from the previous BB selection, in which only two out of six tested 12BB deoxyribozymes formed 3'-5' linkages.⁹ Therefore, changing both aspects of the selection design (overhang length and switching of 2'-dG/2'-G; Fig. 1) lead to a striking extent of control over both the site-selectivity and regioselectivity in the DNA-catalyzed RNA ligation process, in favor of synthesizing the native 3'-5' linkage.

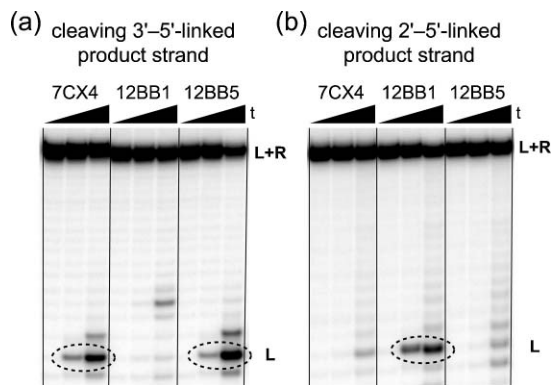


Fig. 4 Cleavage assays to rapidly establish the RNA ligation junction connectivities created by the new deoxyribozymes. The diagnostic bands are circled. Microscopic reversibility requires that deoxyribozymes which are able to cleave a given linkage (*e.g.*, 3'-5' or 2'-5') must form that linkage. Assays used either (a) the 3'-5'-linked 'product' strand, or (b) the 2'-5'-linked 'product' strand, with incubation in 70 mM Tris, pH 7.5, 150 mM NaCl, 2 mM KCl, 1 mM ZnCl₂ at 23 °C (t = 0, 1, 20 h). In both cases illustrated here, the 7CX4 deoxyribozyme was tested along with the 12BB1 and 12BB5 DNA enzymes from our previous report.⁹ These latter deoxyribozymes form the non-native 2'-5'-linked and native 3'-5'-linked product, respectively, and therefore serve as positive and negative controls for these cleavage assays. The data unambiguously indicate that 7CX4 creates a native 3'-5' RNA linkage. Other 7CX, 8CW, and 12BB clones were tested in similar fashion.

With an eye towards efficiently assaying additional DNA enzyme clones, we showed that the linkages in the RNA products from individual deoxyribozymes can be determined without prior sequencing and solid-phase synthesis of each clone. For this purpose, we chose eleven additional 7CX clones for which single-stranded DNA was available by PCR using miniprep DNA isolated during the cloning process. Without knowing the sequences

of these eleven additional 7CX clones, the catalytically active DNA strands were obtained by PCR, purified by PAGE, and used in the 'product' strand cleavage assays of Fig. 4. Ten of these additional eleven 7CX clones create native 3'-5' linkages, and one clone creates non-native 2'-5' linkages. Sequencing then showed that of the ten additional deoxyribozymes that create 3'-5' linkages, six unique sequences are represented; the lone deoxyribozyme that forms 2'-5' linkages is also unique. Therefore, in the final analysis we identified a total of 14 unique 7CX deoxyribozymes, 13 of which form native 3'-5' linkages. Sequences of all clones are listed in the supplementary information.†

To verify the 3'-5' and 2'-5' linkage assignments made for 7CX deoxyribozymes on the basis of the Fig. 4 assays, in several cases we used the deoxyribozymes to synthesize their RNA ligation products. We then directly determined the linkages using our previously established approaches, in which 3'-5' linkages are cleaved selectively by the 8-17 deoxyribozyme whereas 2'-5' linkages are cleaved in the presence of the complementary DNA template at 100 mM Mg²⁺ and pH 9.0.³ In all tested cases, the results were clearly in accord with the Fig. 4 assay (data not shown), confirming the linkage assignments.

Deoxyribozymes from the CW selection

We also examined DNA enzymes from the CW selection, for which only the overhang length was changed relative to the BB selection (Fig. 1). The ligation activity of the uncloned CW pool was *ca.* 14% at round 8 (Fig. 2b), and the 8CW clones were examined. The ligation junctions formed by individual 8CW clones were assayed by the same approach as used for the final eleven 7CX clones; *i.e.*, as PCR products from miniprep DNA without prior sequencing or solid-phase synthesis. For the 8CW deoxyribozymes, 20 active clones were analyzed, leading to ten clones that form native 3'-5' linkages and ten that form non-native 2'-5' linkages.¶ Sequencing revealed that these 20 clones can be grouped into 16 unique classes, of which eight create native 3'-5' linkages and eight form non-native 2'-5' linkages.† The lack of unnatural linkages in the products indicates that the CW selection design led to improved control of site-selectivity in favor of 5'-OH nucleophilicity when compared with the BB selection. In contrast, the CW selection design (unlike the CX selection design) led to very poor control over regioselectivity in 2',3'-cyclic phosphate opening during DNA-catalyzed RNA ligation.

Analysis of additional 12BB deoxyribozymes

Only six 12BB clones were examined in our previous report, leading to two deoxyribozymes that create native 3'-5' linkages, two that form non-native 2'-5' linkages, and two that form unnatural linkages by nucleophilic attack of the internal 2'-OH of the R substrate.⁹ Here we analyzed 21 additional 12BB clones to improve the statistical coverage of the distribution of ligation products. Of these additional clones, seven, nine, and five form 3'-5', 2'-5', or unnatural linkages, respectively. Sequencing then

¶ In a small number of cases for which the assays of Fig. 4 were not entirely conclusive, the ligation product was separately synthesized, and the 8-17 deoxyribozyme,³ complement/100 mM Mg²⁺/pH 9.0,³ and partial alkaline hydrolysis⁹ assays were used to clarify the nature of the linkage (data not shown).

indicated that these 21 clones can be grouped into 13 unique classes, of which five each create 3′–5′ and 2′–5′ linkages, and three form unnatural linkages.[†] Combining these clones with the six unique 12BB deoxyribozymes identified previously,⁹ the final ratio of the three possible types of products (native, non-native, unnatural) is 7:7:5. Therefore, the BB selection was confirmed to have poor control over the direction of both site-selectivity and regioselectivity.

Evaluating identities of unnatural linkages

The identities of the unnatural linkages created by three of the newly identified 12BB deoxyribozymes (12BB23, 30, and 32) were probed by additional experiments. In all three cases, partial alkaline hydrolysis of the ligation product verified that the linkage is unnatural (neither native 3′–5′ nor non-native 2′–5′), because unusually rapid cleavage at the ligation site was observed upon brief incubation under partial alkaline hydrolysis conditions of pH 9.2 and 90 °C (data not shown; similar to the findings for unnatural linkages in our previous report⁹). In contrast, several of the products from clones that form either 3′–5′ or 2′–5′ linkages led to conventional cleavage ladders upon partial alkaline hydrolysis. Also in all three cases, the ligation reaction was completely suppressed by using an R substrate that has a 2′-deoxy modification at the second nucleotide position, consistent with reaction of this 2′-OH as a nucleophile.⁹ Finally, cleavage of each new unnatural product by 12BB8 proceeded with both k_{obs} and yield that are approximately the same as for 12BB8-catalyzed cleavage of the 12BB8 product itself.⁹ Considered together, these data strongly suggest that all three of the new deoxyribozymes synthesize the same unnatural linkage, which (as is the case for 12BB8) is either the 2′,3′-branched or 2′,2′-branched linkage (Scheme 1c).

RNA substrate generality of the 7CX and 8CW deoxyribozymes

One important motivation for pursuing RNA-ligating deoxyribozymes is their potential utility for practical RNA ligation. For the 2′,3′-diol and 5′-triphosphate RNA substrate combination, we have reported several deoxyribozymes that join two substrates *via* native 3′–5′ linkages and that simultaneously have useful rates, yields, and substrate sequence generalities.⁸ However, for the 2′,3′-cyclic phosphate and 5′-OH substrate combination, no such deoxyribozymes have yet been found. Here we performed initial assays on several of the new 7CX and 8CW deoxyribozymes to explore their generality for the two RNA substrate sequences.

On both RNA substrates, the selection design leaves either four overhanging unpaired RNA nucleotides (J and BB selections) or one overhanging unpaired RNA nucleotide (CW and CX selections; Fig. 1). Because RNA-ligating deoxyribozymes that require numerous unpaired RNA substrate nucleotides typically have substantial sequence requirements for those nucleotides,^{3,6,9–11} we did not consider any of the 12BB deoxyribozymes that create native 3′–5′ linkages as likely candidates for high RNA substrate generality. In contrast, because a smaller number of overhanging RNA nucleotides has been more compatible with substrate sequence generality, the new 7CX and 8CW deoxyribozymes (which have only one-nucleotide overhangs) are more reasonable candidates for generality. We therefore examined systematic changes in both

the L and R RNA substrate nucleotides, while always making corresponding changes to the deoxyribozyme binding arms to maintain Watson–Crick complementarity.

We focused our attention on the set of 7CX and 8CW deoxyribozymes that had potentially useful (45–60%) ligation yields with the ‘parent’ RNA substrate sequences and also form native 3′–5′ linkages. This set constituted all 13 of the relevant 7CX clones and six out of eight of the relevant 8CW clones. The generality assays were performed in two phases, in all cases using deoxyribozymes prepared by solid-phase synthesis. We first tested RNA substrate combinations that have systematic sequence changes outside of the fixed UA↓GG nucleotides at the ligation junction (denoted with the arrowhead). All remaining nucleotides in both RNA substrates were changed by systematic transversions, A↔C and G↔U (the deoxyribozyme binding arm nucleotides were altered to maintain base pairing). Eight of the 7CX clones and five of the 8CW clones had relatively high ligation yields with these alternative substrates. Of these, we then tested four of the 7CX clones and five of the 8CW clones with RNA substrates that included the sequence elements GA↓GU, where the underlined nucleotides one position away from the ligation junction (*i.e.*, the penultimate nucleotides) were altered by transversion from the parent sequence UA↓GG, and all remaining nucleotides were unchanged from the original substrates. None of these nine clones had high ligation yields (<1%; the remaining four 7CX clones were not tested with these substrates). We conclude that whereas the 7CX and 8CW deoxyribozymes have promising RNA substrate sequence generality outside of the ligation site nucleotides UA↓GG, sequence requirements that prohibit very high generality are found closer to the ligation site. The precise nature of these requirements was not investigated.

Discussion

Influence of selection design on direction of site-selectivity and regioselectivity during DNA-catalyzed RNA ligation

One of the primary challenges for any *in vitro* selection experiment is to devise a selection design that leads to the desired catalytic outcome.^{2,13} In the case of DNA-catalyzed ligation of a 2′,3′-cyclic phosphate RNA substrate, the principal difficulty revealed by our past experiments has been identifying the best combination of selection design aspects that induce only native 3′–5′ linkages in the RNA products. Here we have found that when Zn²⁺ is the metal ion cofactor, deoxyribozymes that uniformly have the direction of site-selectivity and regioselectivity necessary to synthesize native 3′–5′ linkages (Scheme 1) can be identified *via* an appropriate combination of selection design aspects. This is a significant improvement over our previous work, in which native 3′–5′ linkages were formed by only a small subset of the deoxyribozymes.⁹

In the present work, both the number of overhanging RNA substrate nucleotides (one rather than the previous four) and the identity of the 5′-terminal nucleotide on the right-hand RNA substrate R (switching between 2′-dG and 2′-G during selection, rather than only 2′-dG as used previously) were critical to obtain the desired selectivity (CX selection; Fig. 1d). In contrast, when only the overhang length was changed (CW selection; Fig. 1c), control over the direction of regioselectivity in

2',3'-cyclic phosphate opening remained poor, leading to both native 3'-5' and non-native 2'-5' linkages. However, control over the direction of site-selectivity was very high, favoring the 5'-OH as the nucleophile.

In the CW and CX selections, the two design changes that were made to the original J/BB selection strategy⁹ were largely intended to control the direction of site-selectivity. In both cases, reducing the number of overhanging RNA nucleotides (from four in J/BB to one in CW/CX) predictably did result in increased control over site-selectivity, likely due to physically constraining the internal nucleotides of R. In combination with the use of the 2'-deoxy substitution at the 5'-terminal nucleotide of R, which suppresses reactivity of the 2'-OH, this led to complete control over site-selectivity, favoring nucleophilic attack of the 5'-OH.

In contrast to this rational and successful control of the direction of site-selectivity, it is very difficult to envision how to engineer logical control of the direction of regioselectivity during 2',3'-cyclic phosphate opening. We do not know the mechanistic reason that native 3'-5' linkages (reflecting control over both site-selectivity and regioselectivity) were dramatically favored in the CX selection but not the CW selection, which differed only in whether the 5'-terminal nucleotide of R was switched between 2'-dG and 2'-G (CX) or always 2'-dG (CW). Similarly, we do not know the mechanistic reason that non-native 2'-5' linkages were favored in our earlier efforts with Mg²⁺-dependent deoxyribozymes.³⁻⁵ Collectively, these results demonstrate that control over the final outcome of a selection procedure depends on relatively subtle aspects of the selection design.

One might consider if we can control the linkage formation using our previously introduced 3'-5'-selective pressure that is based on the 8-17 deoxyribozyme, which only cleaves 3'-5' linkages.¹⁴ In this approach, the 8-17 is used during each selection round immediately after the step depicted in Fig. 2, and only those deoxyribozymes that are attached to a 3'-5'-linked RNA product (which is efficiently cleaved by 8-17) are separated by PAGE and taken into the next round. However, the reversibility of RNA ligation with a 2',3'-cyclic phosphate L substrate means that the deoxyribozyme attached to the R substrate during a given selection round can cleave the ligation junction that it just formed, regardless of whether this linkage is 3'-5' or 2'-5' (or, for that matter, any unnatural linkage) As a consequence, this 3'-5'-selective pressure cannot readily be used with the 2',3'-cyclic phosphate RNA substrate.

Potential for general utility of deoxyribozymes for RNA ligation involving a 2',3'-cyclic phosphate

In contrast to the successful control of both site-selectivity and regioselectivity, the generality of the new deoxyribozymes for various RNA substrate sequences was suboptimal, because sequence requirements were found in at least one nucleotide not immediately at the ligation junction. This observation highlights the need to impose a selection pressure specifically in favor of generality, at least for RNA-ligating deoxyribozymes. By reducing the number of overhanging nucleotides on each RNA substrate from four to one, we did successfully limit the sequence requirements to nucleotides near the ligation junction, unlike in our previous efforts where longer single-stranded substrate regions led to more

extensive requirements.^{3,6,9-11} Nevertheless, reducing even further the number of substrate nucleotides that have sequence limitations is important for practical utility.

In other efforts with 5'-triphosphate RNA substrates, we have used a mutant Taq polymerase¹⁵ that tolerates mismatches between the template and the 3'-terminus of the PCR primer. This mutant polymerase allows alternation of the penultimate RNA substrate nucleotides near the ligation junction in successive selection rounds, always while maintaining Watson-Crick base pairing between the RNA and the deoxyribozyme (D. A. Baum, S.K.S., and coworkers, data not shown). Such a strategy is required to reliably achieve generality with 5'-triphosphate substrates, because the generality observed in a small number of cases without applying a specific selection pressure⁸ has not been found in follow-up selection experiments with different RNA substrate sequences (data not shown). We expect that this alternation strategy or a related approach will be necessary to identify deoxyribozymes that ligate 2',3'-cyclic phosphate RNA substrates with high generality while also creating native 3'-5' linkages.

The need for a selection pressure specifically to enforce generality during DNA-catalyzed RNA ligation is in sharp contrast to the observations with RNA-cleaving deoxyribozymes such as 10-23, 8-17, and numerous variants,¹⁶⁻¹⁹ which require at most one nucleotide on either side of the cleavage site despite no special selection pressure in favor of such generality. This intriguing distinction between the generality of deoxyribozymes that ligate and cleave RNA suggests a fundamental difference in the difficulty of RNA ligation *versus* cleavage by a DNA catalyst. Understanding the origin of this difference, which may relate to different active-site dynamics for ligation *versus* cleavage substrates, will require detailed experiments to reveal the mechanisms of catalysis.

Experimental section

Oligonucleotide preparation

DNA oligonucleotides were prepared at IDT (Coralville, IA). The L RNA substrates (bearing a 2',3'-cyclic phosphate) were synthesized by *in vitro* transcription using T7 RNA polymerase and a suitable DNA template that encodes a self-cleaving HDV or hammerhead ribozyme at the 3'-terminus of the desired transcript. The R RNA substrates were prepared by solid-phase synthesis at IDT or by *in vitro* transcription and subsequent CIP dephosphorylation. All DNA and RNA oligonucleotides were purified by denaturing PAGE with running buffer 1× TBE (89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3) as described previously.¹⁰ The identities of the RNA transcripts were confirmed by MALDI mass spectrometry. For kinetic assays (Fig. 3), the 5'-³²P-radiolabeled L substrate with a 2',3'-cyclic phosphate was prepared by 5'-³²P-radiolabeling (γ -³²P-ATP, 3'-phosphatase-free T4 PNK from New England BioLabs) an oligoribonucleotide that was previously cleaved with a 10-23 deoxyribozyme to provide the cyclic phosphate terminus. For ligation junction assays (Fig. 4), the L+R 'product' strands were prepared by solid-phase synthesis, using a 2'-5'-adenosine phosphoramidite monomer (ChemGenes) for the 2'-5'-linked strand, and then 5'-³²P-radiolabeled.

In vitro selection procedure

We previously described the generic *in vitro* selection strategy that is based on separation of catalytically active deoxyribozymes by PAGE.³ For both of the selections performed here (CW and CX), the initial deoxyribozyme pool was 200 pmol (*ca.* 10¹⁴ molecules). In the key selection step, the DNA pool—previously attached to the R substrate using T4 RNA ligase—was heated to 95 °C for 3 min in 5 mM Tris, pH 7.5, 15 mM NaCl and 0.1 mM EDTA, then placed on ice for 5 min. The sample was adjusted to 70 mM Tris, pH 7.5, 150 mM NaCl, 2 mM KCl, and 1 mM ZnCl₂, then incubated for 3 h at 23 °C. The metal Zn²⁺ was added from a 10× stock solution containing 10 mM ZnCl₂, 20 mM HNO₃, and 200 mM Tris at pH 7.5; this stock solution was freshly prepared from a 100× stock of 100 mM ZnCl₂ in 200 mM HNO₃. After incubation with 1 mM Zn²⁺, the samples were separated by 8% denaturing PAGE.

Cloning and initial screening of individual deoxyribozyme clones was performed using DNA strands prepared by PCR from miniprep DNA derived from individual *E. coli* colonies. The miniprep DNA samples were first assayed by digestion with EcoRI to ascertain the presence of the expected insert. The concentration of each PAGE-purified deoxyribozyme strand was estimated from the intensity (UV shadowing) relative to suitable standards. Each screening assay used *ca.* 0.1 pmol of 5'-³²P-radiolabeled L substrate, 2 pmol of the deoxyribozyme strand, and 20 pmol of R substrate. Samples were incubated at 23 °C with timepoints at 0, 1, and 7–15 h, using the procedure described below for ligation activity assays.

Ligation activity assays (Fig. 3)

Individual deoxyribozymes prepared by solid-phase synthesis were assayed for ligation activity as follows. A sample was prepared in 7 μL containing 0.5 pmol 5'-³²P-radiolabeled L substrate, 2 pmol deoxyribozyme, and 10 pmol R substrate in 5 mM Tris, pH 7.5, 15 mM NaCl, 0.1 mM EDTA. (Alternatively, some of the generality assays used 0.5 pmol 3'-³²P-radiolabeled R substrate [³²P-pCp, T4 RNA ligase], 2 pmol deoxyribozyme, and 10 pmol L substrate.) The sample was annealed by heating at 95 °C for 3 min and cooling on ice for 5 min. Stock solutions were then added to provide a total volume of 10 μL containing 70 mM Tris, pH 7.5, 150 mM NaCl, 2 mM KCl, and 1 mM ZnCl₂ (the ZnCl₂ stock solution was as described above for the *in vitro* selection procedure). The sample was incubated at 23 °C, with timepoints taken between 0–300 min and separated by 20% denaturing PAGE.

Ligation junction assays (Fig. 4)

The ligation junctions created by each deoxyribozyme were assayed by testing for cleavage of the 3'–5'-linked and 2'–5'-linked 'product' strands. The 5'-³²P-radiolabeled product strand (0.5 pmol) was incubated with the deoxyribozyme (2 pmol) using the procedure described above for ligation activity assays, with timepoints taken at 0, 1, and 7–20 h and separated by 20% denaturing PAGE. The deoxyribozyme strand was prepared either by solid-phase synthesis (some 7CX clones) or by PCR from miniprep DNA followed by PAGE purification (remaining 7CX clones; all 8CW clones; and all new 12BB clones).

Acknowledgements

This research was supported by the US National Institutes of Health (GM-65966 to S.K.S.), the David and Lucile Packard Foundation (fellowship to S.K.S.), and Sigma Xi (Grant-in-Aid of Research to J.P.G.).

Notes and references

- 1 A. Peracchi, *ChemBioChem*, 2005, **6**, 1316–1322.
- 2 S. K. Silverman, *Chem. Commun.*, 2008, 3467–3485.
- 3 A. Flynn-Charlebois, Y. Wang, T. K. Prior, I. Rashid, K. A. Hoadley, R. L. Coppins, A. C. Wolf and S. K. Silverman, *J. Am. Chem. Soc.*, 2003, **125**, 2444–2454.
- 4 A. Flynn-Charlebois, T. K. Prior, K. A. Hoadley and S. K. Silverman, *J. Am. Chem. Soc.*, 2003, **125**, 5346–5350.
- 5 D. R. Semlow and S. K. Silverman, *J. Mol. Evol.*, 2005, **61**, 207–215.
- 6 Y. Wang and S. K. Silverman, *J. Am. Chem. Soc.*, 2003, **125**, 6880–6881.
- 7 R. L. Coppins and S. K. Silverman, *Nat. Struct. Mol. Biol.*, 2004, **11**, 270–274.
- 8 W. E. Purtha, R. L. Coppins, M. K. Smalley and S. K. Silverman, *J. Am. Chem. Soc.*, 2005, **127**, 13124–13125.
- 9 K. A. Hoadley, W. E. Purtha, A. C. Wolf, A. Flynn-Charlebois and S. K. Silverman, *Biochemistry*, 2005, **44**, 9217–9231.
- 10 Y. Wang and S. K. Silverman, *Biochemistry*, 2003, **42**, 15252–15263.
- 11 B. L. Ricca, A. C. Wolf and S. K. Silverman, *J. Mol. Biol.*, 2003, **330**, 1015–1025.
- 12 T. K. Prior, D. R. Semlow, A. Flynn-Charlebois, I. Rashid and S. K. Silverman, *Nucleic Acids Res.*, 2004, **32**, 1075–1082.
- 13 G. F. Joyce, *Annu. Rev. Biochem.*, 2004, **73**, 791–836.
- 14 Y. Wang and S. K. Silverman, *Biochemistry*, 2005, **44**, 3017–3023.
- 15 F. J. Ghadessy, N. Ramsay, F. Boudsocq, D. Loakes, A. Brown, S. Iwai, A. Vaisman, R. Woodgate and P. Holliger, *Nat. Biotechnol.*, 2004, **22**, 755–759.
- 16 S. W. Santoro and G. F. Joyce, *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 4262–4266.
- 17 R. P. G. Cruz, J. B. Withers and Y. Li, *Chem. Biol.*, 2004, **11**, 57–67.
- 18 K. Schlosser, J. Gu, L. Sule and Y. Li, *Nucleic Acids Res.*, 2008, **36**, 1472–1481.
- 19 S. K. Silverman, *Nucleic Acids Res.*, 2005, **33**, 6151–6163.

Electronic Supplementary Information

Controlling the direction of site-selectivity and regioselectivity in RNA ligation by Zn²⁺-dependent deoxyribozymes that use 2',3'-cyclic phosphate RNA substrates

Diana M. Kost, Joseph P. Gerdt, P. I. Pradeepkumar and Scott K. Silverman*

Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801 (USA)

Deoxyribozyme sequences

Unique sequences of all deoxyribozyme enzyme regions are listed below. Some of the enzyme regions are 39 or 41 nt rather than 40 nt due to a nucleotide deletion or addition during the selection process. For two pairs of the 8CW deoxyribozymes, the sequences could not be determined confidently due to overlap within one sequencing run between the two members of the pair, presumably due to overlapping *E. coli* colonies harboring plasmids with the two different deoxyribozyme sequences. Nonetheless, in each such case, the RNA linkage assignment is clear because partial alkaline hydrolysis assay of the mixture of products clearly revealed that no unnatural linkage was present, and the 8–17 and complement/100 mM Mg²⁺/pH 9.0 assays were similarly unequivocal. For each 7CX and 8CW deoxyribozyme, the binding arm to the 5'-side of the enzyme region is 5'-CGAAGTCGCCATCTCTTC..., and the binding arm to the 3'-side of the enzyme region is ...ATAGTGAGTCGTATTA-3'. For each 12BB deoxyribozyme, the binding arm to the 5'-side of the enzyme region is 5'-CGAAGTCGCCATCTC..., and the binding arm to the 3'-side of the enzyme region is ...GTGAGTCGTATTA-3'. Ligation yields are shown for deoxyribozymes prepared by solid-phase synthesis using either the parent (Par) or transversion (Tv) RNA substrate sequences outside of the unchanged UA↓GG nucleotides near the ligation site. Yields were determined only for deoxyribozymes that form 3'–5' linkages. n.d., yield not determined using a deoxyribozyme prepared by solid-phase synthesis.

7CX deoxyribozymes

<u>Name</u>	<u>Enzyme region sequence, 5' to 3'</u>	<u>Linkage</u>	<u>Yield (Par)</u>	<u>Yield (Tv)</u>
7CX2	GTGAGAGCCATTGTGCATTGTGGTGCAACTGTACGCGCCT	3'–5'	55%	46%
7CX3	GTGAGAGCAGGTTTGGCACGCGTGCCACTAACCCGCGCCT	3'–5'	n.d.	50%
7CX4	GTGGGAGCCTGATGGAAGAATGGGAGAACAGCACGCGCCT	3'–5'	56%	11%
7CX5	GGGAGAGCTAGGAGGCGCTGAGTCTCGGCTCCTCGCGCCT	3'–5'	n.d.	35%
7CX6	GTGAGAGCAAGCGAGGGATACATGAACCGAAACCGCGCCT	3'–5'	44%	43%
7CX7	GTGAGAGCAGTAAACGGAAGGCCCTCTTCCTGTCCGCGCCT	3'–5'	n.d.	50%
7CX8	GTGAGAGCATGTGGCACCCCTGGCAGGAACCTAACGCGCCT	3'–5'	56%	19%
7CX9	ACTGAGAGCCTCGATGAATGCCTAATTATATCGTCGCGCC	3'–5'	57%	32%
7CX10	GTGAGAGCAATACGTATGAAGATATTGAAGTTTCGCGCCT	3'–5'	55%	57%
7CX12	GTGAGAGCATCACTGTAACCGCCGTGCGGTGGACGCGCCT	3'–5'	53%	10%
7CX13	GGGAGAGCAGTTATGATGCATGAGCCATTGACCGCGCCT	3'–5'	n.d.	34%
7CX14	GGGAGAGCAGCGATGCAAGACATTGGGGCGGAGCCGCGCT	3'–5'	n.d.	15%
7CX21	TGAGAGCATTTCCGGCGGAGCTCTACGGGGCCGCGCCT	3'–5'	n.d.	<1%
7CX11	TTTAGTCAGGAGTAGACATCGATGATTGCTAATCGAACCT	2'–5'	n.d.	n.d.

8CW deoxyribozymes

<u>Name</u>	<u>Enzyme region sequence, 5' to 3'</u>	<u>Linkage</u>	<u>Yield (Par)</u>	<u>Yield (Tv)</u>
8CW1	GTGAGAGAACC GCGGTGTAGACACAGATCGCGGCGCGCCT	3'-5'	n.d.	45%
8CW8	TACAAGGTGGGAGGAGGGAGCACCCGATGCGGCATATCGTG	3'-5'	n.d.	50%
8CW11	GGGAGAGCAGTTATGACGCATAAGCTATTGACCGCGCCT	3'-5'	n.d.	15%
8CW15	GTGAGAGCAATCACTTTGGGTAGGTACGGGTGAACGCGCT	3'-5'	n.d.	52%
8CW16	GTTAGAGCCAAACACGTTTGTGTGTCAGCGGGTTTCGCGCCT	3'-5'	n.d.	40%
8CW20	GGAAGAGCAATGTACTCCGACGTCAGAGGATATCGCGCCT	3'-5'	n.d.	39%
8CW17	-	3'-5'	n.d.	n.d.
8CW22	-	3'-5'	n.d.	n.d.
8CW3	ACGCCCACCATTTAAGAGCATCGCCGGAATAGCGGGGACT	2'-5'		
8CW4	GCACGGCCGGATTGGGGGGCCCAATGACTTGATCATTGCT	2'-5'		
8CW9	ACACGCGTTACGGACTGAGCAGATTCAGGCCAAACTGCCT	2'-5'		
8CW12	ACATGTGCGGTAATGAGGCGTAGACAAATAGAGATACCT	2'-5'		
8CW13	GGGGAAGACAAGATAGTCGAGGGGGACGCGCTCTCAGCCT	2'-5'		
8CW14	GCGGGGTTAAGCCACGAATGCGGGGCAAAGCGTACCCCT	2'-5'		
8CW19	-	2'-5'		
8CW21	-	2'-5'		

12BB deoxyribozymes (sequences of 12BB1, 2, 5, 6, 8, and 12 are from ref. 9)

<u>Name</u>	<u>Enzyme region sequence, 5' to 3'</u>	<u>Linkage</u>	<u>Yield (Par)</u>	<u>Yield (Tv)</u>
12BB5	CAGCGCGATTGAGTGCGTGATTGAAGCTCGGGGTTGGTTA	3'-5'	n.d.	n.d.
12BB12	CAGAGCCCCCTTACGTACAGCCTTTTTAGGTAACCGGGGAG	3'-5'	n.d.	n.d.
12BB24	CAGCGCGATTGGGGGCGTGATTGAAGCTCGGGGTTGGTTA	3'-5'	n.d.	n.d.
12BB26	ACTATACCACCGAGATTCGAATTGGAGCAGTAGTGGCTTG	3'-5'	n.d.	n.d.
12BB33	TGTAGCCTTCTGAAGGTTGGCTGGTTCGGCGAGGTGGGAA	3'-5'	n.d.	n.d.
12BB36	ACCGGTCCCTGTCGGTCAAGGAGTGGCACTGAGGAGAAT	3'-5'	n.d.	n.d.
12BB41	ACGCGAGGGGAGTTCAATCGCTTGTTCGGCAAGGTCGGGA	3'-5'	n.d.	n.d.
12BB1	CCGCTCCGATTGGTGGAGTCTATTGGGGCTGTAGGCGAG	2'-5'		
12BB2	ACCGCGGGAGCTACGTTAGTGGTAACTGCTTGTAGGCGAG	2'-5'		
12BB21	AAGGCAGCAGGCGGCATTTTTGTCCGTACGTTCTCCTATA	2'-5'		
12BB22	ACCGCGGGAGTTTCGTTAGTGGTACTGCTTGTAGGCGAG	2'-5'		
12BB25	GTAGAGTGGTTCGGTTATCGCCTTAGATAACCATTGCTGT	2'-5'		
12BB27	CCGCTCCGATTGGTGAAGTCTATTGGGGCTGTAGGCGAG	2'-5'		
12BB38	CCGCTGCGGAGGTTGTACGCGTCTGTGGCTTGTAGGCGAG	2'-5'		
12BB6	TACACCTATTATGGTTTTCGTGAGGGGTGTGGCTGGTGCTG	unnat		
12BB8	CACGCTGACTAGCTTCGTGAGGGGTGTGATAGATGCGG	unnat		
12BB23	TACACCTATAATGGTTTTCGTGAGGGGTGTGGCTGATGCTG	unnat		
12BB30	CCTGTACTGCGTGTCTCAAATCAGCCGGGTGTGTGAACTC	unnat		
12BB32	TGTATGTGGGGGTGTGTATCAGTCTACTGTGGCTTAAGC	unnat		