J|A|C|S
ARTICLES
Published on Web 11/20/2004

Rational Modification of a Selection Strategy Leads to Deoxyribozymes that Create Native 3′-5′ RNA Linkages

Rebecca L. Coppins and Scott K. Silverman*

Contribution from the Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801

Received July 12, 2004; E-mail: scott@scs.uiuc.edu

Abstract: We previously used in vitro selection to identify several classes of deoxyribozymes that mediate RNA ligation by attack of a hydroxyl group at a 5′-triphosphate. In these reactions, the nucleophilic hydroxyl group is located at an internal 2′-position of an RNA substrate, leading to 2′,5′-branched RNA. To obtain deoxyribozymes that instead create linear 3′-5′-linked (native) RNA, here we strategically modified the selection approach by embedding the nascent ligation junction within an RNA:DNA duplex region. This approach should favor formation of linear rather than branched RNA because the two RNA termini are spatially constrained by Watson−Crick base pairing during the ligation reaction. Furthermore, because native 3′-5′ linkages are more stable in a duplex than isomeric non-native 2′-5′ linkages, this strategy is predicted to favor the formation of 3′-5′ linkages. All of the new deoxyribozymes indeed create only linear 3′-5′ RNA, confirming the effectiveness of the rational design. The new deoxyribozymes ligate RNA with k_{obs} values up to 0.5 h^{-1} at 37 °C and 40 mM Mg^{2+}, pH 9.0, with up to 41% yield at 3 h incubation. They require several specific RNA nucleotides on either side of the ligation junction, which may limit their practical generality. These RNA ligase deoxyribozymes are the first that create native 3′-5′ RNA linkages, which to date have been highly elusive via other selection approaches.

Deoxyribozymes (DNA enzymes)[1] are catalysts that mediate a variety of reactions of nucleic acid substrates.[2] Our laboratory has focused on deoxyribozymes for RNA ligation, using one of two fundamental chemical reactions (Figure 1). First, a 5′-hydroxyl group nucleophile can attack a 2′,3′-cyclic phosphate, forming either a 2′-5′ (non-native) or 3′-5′ (native) linear phosphodiester linkage (Figure 1A, paths *i* and *ii*).[3−5] Alternatively, a 2′- or 3′-hydroxyl group nucleophile can attack a 5′-triphosphate, forming either 2′-5′ linear RNA, 3′-5′ linear RNA, or 2′,5′-branched RNA (Figure 1B, paths *i*−*iii*).[6−8] Among the possible products, 3′-5′-linked linear RNA is a particularly important practical goal because synthesis of such native linkages will allow preparation of large RNAs by ligation of smaller fragments.[9] Significantly, this ligation will occur in a traceless fashion; that is, with no evidence remaining in the RNA itself that a ligation reaction was performed and, therefore, no nonstandard linkage that could have unintended structural or functional effects. Because one or more of the RNA substrates may be prepared by solid-phase synthesis, the ligated RNAs may incorporate site-specific modifications[10] to enable structure−function studies.[11] However, our efforts reported to date have not provided deoxyribozymes that synthesize 3′-5′ RNA, using either set of substrates from Figure 1. For the deoxyribozymes that mediate reaction of a 5′-triphosphate (Figure 1B), the nucleophilic hydroxyl group has always been an internal 2′-hydroxyl, leading to biochemically relevant 2′,5′-branched RNA (path *iii*).[6,7] In this study, we sought to modify the selection strategy specifically to obtain 3′-5′ linkages (path *ii* of Figure 1B).

The strategy that we previously pursued for selecting RNA ligase deoxyribozymes using the 5′-triphosphate substrate combination is shown in Figure 2A as strategy 1. In this approach, the ligation junction is formed between two RNA substrates that each have four-nucleotide RNA overhangs between the RNA:DNA duplex regions. The resulting deoxy-

(1) (a) Breaker, R. R.; Joyce, G. F. *Chem. Biol.* **1994**, *1*, 223−229. (b) Breaker, R. R. *Nat. Biotechnol.* **1997**, *15*, 427−431. (c) Breaker, R. R. *Science* **2000**, *290*, 2095−2096.
(2) (a) Lu, Y. *Chem.*−*Eur. J.* **2002**, *8*, 4589−4596. (b) Emilsson, G. M.; Breaker, R. R. *Cell. Mol. Life Sci.* **2002**, *59*, 596−607. (c) Silverman, S. K. *Org. Biomol. Chem.* **2004**, *2*, 2701−2706.
(3) Flynn-Charlebois, A.; Wang, Y.; Prior, T. K.; Rashid, I.; Hoadley, K. A.; Coppins, R. L.; Wolf, A. C.; Silverman, S. K. *J. Am. Chem. Soc.* **2003**, *125*, 2444−2454.
(4) (a) Flynn-Charlebois, A.; Prior, T. K.; Hoadley, K. A.; Silverman, S. K. *J. Am. Chem. Soc.* **2003**, *125*, 5346−5350. (b) Ricca, B. L.; Wolf, A. C.; Silverman, S. K. *J. Mol. Biol.* **2003**, *330*, 1015−1025.
(5) Prior, T. K.; Semlow, D. R.; Flynn-Charlebois, A.; Rashid, I.; Silverman, S. K. *Nucleic Acids Res.* **2004**, *32*, 1075−1082.
(6) (a) Wang, Y.; Silverman, S. K. *J. Am. Chem. Soc.* **2003**, *125*, 6880−6881. (b) Wang, Y.; Silverman, S. K. *Biochemistry* **2003**, *42*, 15252−15263.
(7) Coppins, R. L.; Silverman, S. K. *Nat. Struct. Mol. Biol.* **2004**, *11*, 270−274.
(8) Coppins, R. L.; Silverman, S. K. **2004**, submitted for publication.
(9) (a) Moore, M. J.; Sharp, P. A. *Science* **1992**, *256*, 992−997. (b) Moore, M. J.; Query, C. C. In *RNA-Protein Interactions: A Practical Approach*; Smith, C. W. J., Ed.; Oxford University Press: Oxford, 1998; pp 75−108. (c) Moore, M. J. *Methods Mol. Biol.* **1999**, *118*, 11−19. (d) Moore, M. J.; Query, C. C. *Methods Enzymol.* **2000**, *317*, 109−123. (e) Bain, J. D.; Switzer, C. *Nucleic Acids Res.* **1992**, *20*, 4372. (f) Sherlin, L. D.; Bullock, T. L.; Nissan, T. A.; Perona, J. J.; Lariviere, F. J.; Uhlenbeck, O. C.; Scaringe, S. A. *RNA* **2001**, *7*, 1671−1678.
(10) (a) Scaringe, S. A.; Wincott, F. E.; Caruthers, M. H. *J. Am. Chem. Soc.* **1998**, *120*, 11820−11821. (b) Earnshaw, D. J.; Gait, M. J. *Biopolymers* **1998**, *48*, 39−55. (c) Scaringe, S. A. *Methods Enzymol.* **2000**, *317*, 3−18. (d) Scaringe, S. A. *Methods* **2001**, *23*, 206−217.

10.1021/ja045817x CCC: $27.50

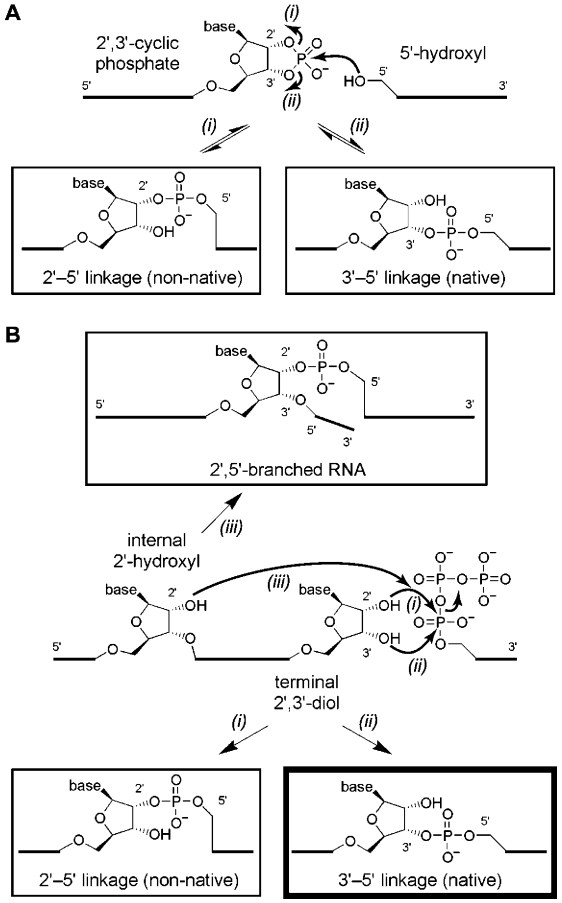

Figure 1. Possible RNA ligation reactions catalyzed by deoxyribozymes. (A) Attack of a 5′-hydroxyl group onto a 2′,3′-cyclic phosphate, leading to either 2′-5′-linked RNA (*i*) or 3′-5′-linked RNA (*ii*). (B) Attack of a hydroxyl group onto a 5′-triphosphate, forming either linear RNA or 2′,5′-branched RNA. The product depends on whether the nucleophilic hydroxyl group is from the terminal 2′,3′-diol (as in *i* and *ii*) or an internal nucleotide (as in *iii*). The reactions of 5′-triphosphorylated RNA are shown with a nonreversible reaction arrow because we have never observed the reverse reaction for any of our deoxyribozymes, although reversibility is theoretically possible.

ribozymes used the indicated internal nucleotide 2′-hydroxyls (solid arrows) as nucleophiles, providing 2′,5′-branched RNA rather than linear RNA.[6] Strategy 1 was subsequently repeated after extending the base-pairing of the left-hand (L) RNA substrate to leave just one overhanging RNA nucleotide (not

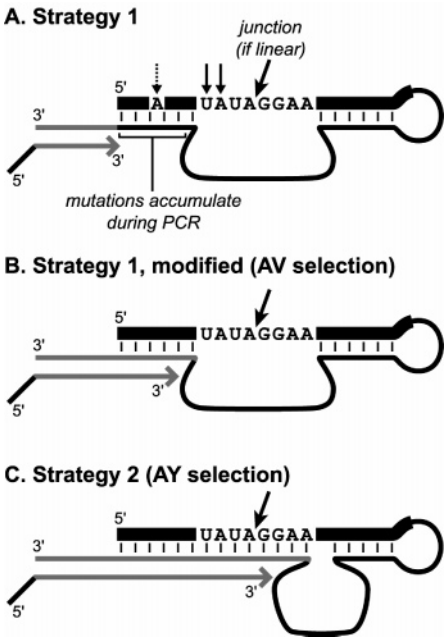

Figure 2. Selection strategies for deoxyribozyme-mediated RNA ligation using the 5′-triphosphate reactions of Figure 1B. RNA is shown as a thick bar and DNA as a thin line. The number of RNA:DNA base pairs is not depicted quantitatively. (A) The original approach (strategy 1), which leaves four-nucleotide RNA overhangs flanking the linear ligation junction. In practice, this strategy leads to formation of 2′,5′-branched RNA, using the indicated U or A 2′-hydroxyl groups (solid arrows) as the nucleophiles.[6] When the overhang on the left is reduced to just one nucleotide by creating three additional base pairs with the overhanging UAU, branch formation is observed using an adenosine 2′-hydroxyl from within the left-hand RNA (dashed arrow). This is accompanied by loss of RNA:DNA base pairing due to accumulation of mutations within the indicated portion of the DNA binding arm.[7] The PCR primer that binds to the 3′-end of the DNA strand is shown explicitly. (B) A modified version of strategy 1 that prevents DNA mutations from accumulating by using a longer PCR primer. Because four overhanging RNA nucleotides are retained, branched RNA is still the expected outcome; the picture shows a linear connectivity of the RNA product to facilitate comparison with the other panels. This strategy was implemented in the selection designated here as AV. (C) A new approach (strategy 2) in which the enzyme region is offset from the ligation junction, such that the new RNA linkage is formed within a duplex region. This is anticipated to favor linear RNA and specifically 3′-5′ linkages, which are more stable than 2′-5′ linkages in the context of a duplex.[12] This strategy was performed in the selection designated here as AY. Note that the illustrated PCR primer is further extended to cover all DNA nucleotides that are complementary to the RNA substrate sequences, including those nucleotides immediately around the ligation junction.

shown explicitly in Figure 2).[7] In this effort, we reasoned that constraining the overhanging nucleotides in a duplex should restrict their reactivity and favor formation of linear RNA by using a hydroxyl group from the terminal 2′,3′-diol as a nucleophile. While the reactivity of the formerly overhanging internal nucleotides was indeed suppressed, accumulation of PCR-derived mutations in the left-hand DNA binding arm led to substantial alteration of the RNA:DNA interactions and emergence of a deoxyribozyme that creates branched RNA.[7,8] This reaction occurs essentially in the manner of Figure 2A using the indicated adenosine 2′-hydroxyl from deep within the L RNA substrate (dashed arrow). Although the resulting deoxyribozyme has proved to be a very interesting model system for nucleic acid catalysis, an important question remains: how can we synthesize 3′-5′-linked RNA using deoxyribozymes?

The results of these previous studies suggested two rational modifications of the strategy that together would achieve

(11) (a) Musier-Forsyth, K.; Usman, N.; Scaringe, S.; Doudna, J.; Green, R.; Schimmel, P. *Science* **1991**, *253*, 784−786. (b) Pieken, W. A.; Olsen, D. B.; Benseler, F.; Aurup, H.; Eckstein, F. *Science* **1991**, *253*, 314−317. (c) SantaLucia, J., Jr.; Kierzek, R.; Turner, D. H. *Science* **1992**, *256*, 217−219. (d) Paolella, G.; Sproat, B. S.; Lamond, A. I. *EMBO J.* **1992**, *11*, 1913−1919. (e) Herschlag, D.; Eckstein, F.; Cech, T. R. *Biochemistry* **1993**, *32*, 8299−8311. (f) Grasby, J. A.; Butler, P. J. G.; Gait, M. J. *Nucleic Acids Res.* **1993**, *21*, 4444−4450. (g) Tuschl, T.; Ng, M. M.; Pieken, W.; Benseler, F.; Eckstein, F. *Biochemistry* **1993**, *32*, 11658−11668. (h) Fu, D.-J.; Rajur, S. B.; McLaughlin, L. W. *Biochemistry* **1993**, *32*, 10629−10637. (i) Hamy, F.; Asseline, U.; Grasby, J.; Iwai, S.; Pritchard, C.; Slim, G.; Butler, P. J.; Karn, J.; Gait, M. J. *J. Mol. Biol.* **1993**, *230*, 111−123. (j) Strobel, S. A.; Cech, T. R.; Usman, N.; Beigelman, L. *Biochemistry* **1994**, *33*, 13824−13835. (k) Abramovitz, D. L.; Friedman, R. A.; Pyle, A. M. *Science* **1996**, *271*, 1410−1413. (l) Liu, Q.; Green, J. B.; Khodadadi, A.; Haeberli, P.; Beigelman, L.; Pyle, A. M. *J. Mol. Biol.* **1997**, *267*, 163−171. (m) Hamm, M. L.; Piccirilli, J. A. *J. Org. Chem.* **1997**, *62*, 3415−3420. (n) Verma, S.; Eckstein, F. *Annu. Rev. Biochem.* **1998**, *67*, 99−134. (o) Silverman, S. K.; Cech, T. R. *Biochemistry* **1999**, *38*, 8691−8702. (p) Silverman, S. K.; Cech, T. R. *Biochemistry* **1999**, *38*, 14224−14237. (q) Silverman, S. K.; Deras, M. L.; Woodson, S. A.; Scaringe, S. A.; Cech, T. R. *Biochemistry* **2000**, *39*, 12465−12475. (r) Silverman, S. K.; Cech, T. R. *RNA* **2001**, *7*, 161−166. (s) Young, B. T.; Silverman, S. K. *Biochemistry* **2002**, *41*, 12271−12276.

synthesis of the desired 3′-5′-linked RNA. Modifying strategy 1 slightly by extending the appropriate PCR primer to encompass all of the left-hand DNA binding arm, instead of merely part of this arm, was anticipated to prevent DNA mutations from accumulating (Figure 2B, strategy 1 modified). However, in this approach, we would still expect 2′,5′-branched RNA as the product because the overhanging nucleotides of the left-hand RNA substrate remain present and likely reactive. Therefore, the strategy was additionally modified by offsetting the ligation junction away from the DNA enzyme region (Figure 2C, strategy 2). With this change, the newly formed RNA linkage is embedded within a continuous Watson−Crick DNA:RNA duplex. Formation of linear instead of branched RNA should be strongly favored because the two terminal nucleotides are held closely together by the base pairing, whereas branch formation would require disruption of at least one duplex element. Furthermore, 3′-5′ RNA linkages are more stable in an RNA:DNA duplex than the isomeric 2′-5′ linkages,[12,13] and a substantial fraction of this stability difference is predicted to be present in the transition state for ligation, thereby favoring the synthesis of 3′-5′-linked RNA. Here, we demonstrate that this rationally modified approach indeed provides 3′-5′-linked RNA, revealing unprecedented control over the linkage formed via deoxyribozyme-mediated RNA ligation. In addition to increasing our control over the outcome of deoxyribozyme-catalyzed reactions, the successful identification of deoxyribozymes that create native 3′-5′ RNA linkages is an important step to improving synthetic access to site-specifically modified large RNAs for studies of structure, folding, and catalysis.

Results

Selection Procedure and Progression. The strategies shown in Figure 2B,C were performed here as previously described for our other deoxyribozyme selections,[3−7] using the 5′-triphosphate RNA substrate combination that can perform any of the reactions of Figure 1B. As part of our laboratory's ongoing selection nomenclature, the two new efforts were designated AV and AY, respectively. In the key selection step (illustrated in Figure 3A for strategy 2, corresponding to the AY selection), the left-hand and right-hand (L and R) RNA substrates were joined in a reaction mediated by the deoxyribozyme's enzyme region, which initially comprised 38 random nucleotides (N_{38}). The incubation conditions were 50 mM CHES, pH 9.0, 150 mM NaCl, 2 mM KCl, and 40 mM $MgCl_2$ at 37 °C for 2 h. Prior to the key selection step, the R substrate was covalently attached to the DNA, which permits polyacrylamide gel electrophoresis (PAGE) separation of active from inactive deoxyribozymes. Following the key selection step, we amplified the "winning" deoxyribozyme sequences by PCR. After seven and eight rounds for the AV and AY selections, respectively, the selection activities (i.e., the fraction of each deoxyribozyme pool that forms ligated RNA during the selection step) leveled off at ∼51 and 38% (data not shown), respectively, signifying the end of selection.

Linkages Formed by the New Deoxyribozymes. Before cloning individual deoxyribozymes, we assessed the linkage selectivity during RNA ligation for both selections by assaying

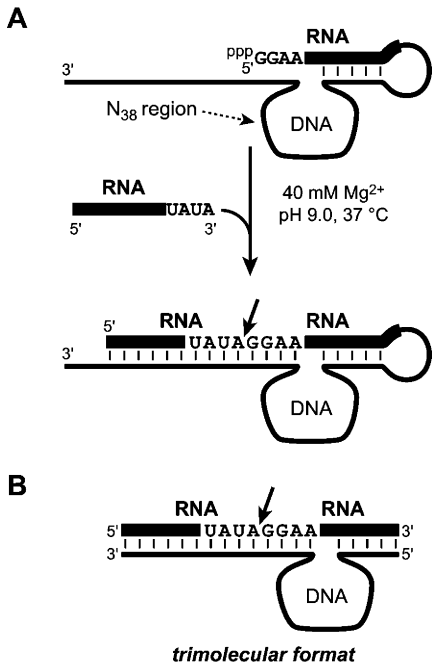

Figure 3. (A) Key step of each selection round, illustrated for strategy 2 of Figure 2C (the AY selection). The number of RNA:DNA base pairs is not depicted quantitatively. (B) Trimolecular format in which the new deoxyribozymes can be used to join two separate RNA substrates.

the pool products from round 7 for selection AV and round 8 for selection AY. This was achieved by partial alkaline hydrolysis (Figure 4A), which generates an unbroken ladder for linear RNA but a ladder with a characteristic gap for 2′,5′-branched RNA. For branched RNA, the location of the gap indicates the branch point, as calibrated by an RNase T1 digestion ladder run in parallel on the same gel. The 7AV products are mostly branched. The branch site is the adenosine two nucleotides from the 3′-terminal residue (i.e., the A of ...UAUA-3′), although contributions from other branch sites cannot be rigorously excluded based on this whole-pool assay. In contrast, the 8AY products are largely linear, as evidenced by the unbroken alkaline hydrolysis ladder.

Further characterization of the products required cloning of individual deoxyribozymes, which was performed as described previously.[3,5] Individual 8AY clones were tested for ligation activity (data not shown), and six promising deoxyribozymes were prepared independently by solid-phase synthesis. See the Supporting Information for sequences, which show regions of some similarity but no clear consensus. The RNA ligation products of these six deoxyribozymes (see below for kinetic analyses) were assayed by partial alkaline hydrolysis (Figure 4B), which confirmed that each is linear, as expected based on the whole-pool assay. To distinguish native 3′-5′ linkages from non-native 2′-5′ linkages, two previously described assays were performed.[3] In the first assay, the 8-17 deoxyribozyme[14] selectively cleaves 3′-5′ RNA linkages, and in the second assay, incubation with the exactly complementary DNA splint and 100 mM Mg^{2+} at pH 9 selectively cleaves 2′-5′ RNA linkages.[12] The 3′-5′-selective assay performed on the uncloned round 8AY pool showed that >95% of the pool products are 3′-5′-linked (data not shown). Consistent with this, all six individual 8AY

(12) Rohatgi, R.; Bartel, D. P.; Szostak, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 3340−3344.
(13) Usher, D. A.; McHale, A. H. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 1149−1153.
(14) Santoro, S. W.; Joyce, G. F. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4262−4266.

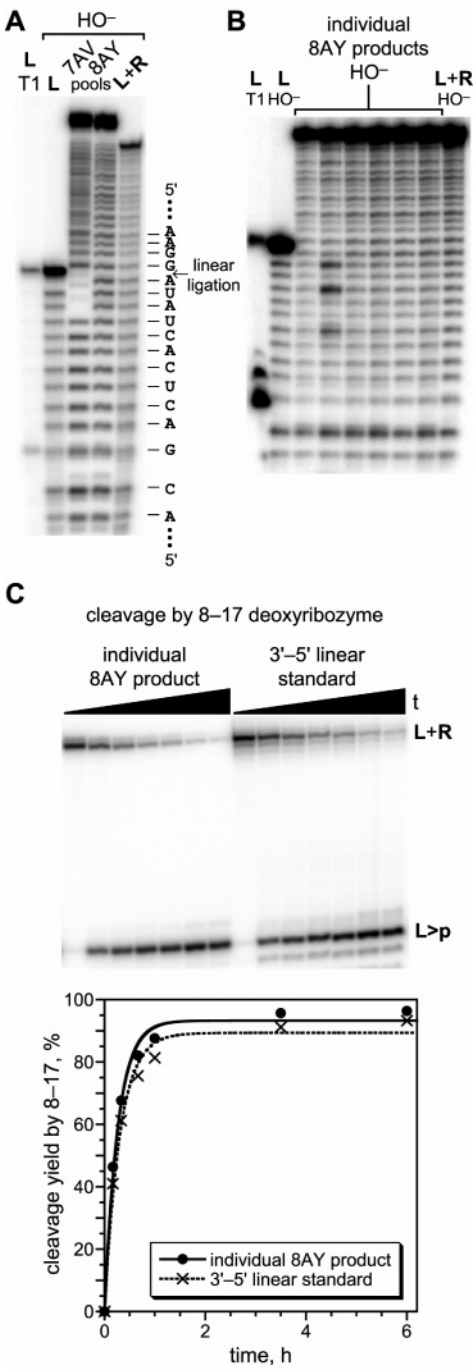

Figure 4. Analysis of the branched versus linear nature of the ligation products. (A) Determining the linkages formed by the pools of the AV and AY selections (after rounds 7 and 8, respectively) using a partial alkaline hydrolysis assay (HO^-). The 7AV products are branched, as indicated by the gap in the ladder, whereas the 8AY products are linear, as indicated by the unbroken ladder. (B) Partial alkaline hydrolysis of the ligation products from the six individual 8AY deoxyribozymes, demonstrating that all products are linear. (C) Cleavage of the new linkages with the 8-17 deoxyribozyme, indicating that the products are 3′-5′-linked. The products from each of the six 8AY deoxyribozymes were treated with the appropriate 8-17 deoxyribozyme, which selectively cleaves 3′-5′ linkages;[3] a representative data set for one deoxyribozyme is shown (all six products and the uncloned 8AY pool gave equivalent results). The linear 3′-5′-linked standard RNA was prepared independently by solid-phase synthesis.

deoxyribozymes unambiguously create 3′-5′ linkages (Figure 4C). Similarly, the 2′-5′-selective assays performed alongside a positive control clearly show that the new linkages are not 2′-5′ (data not shown).

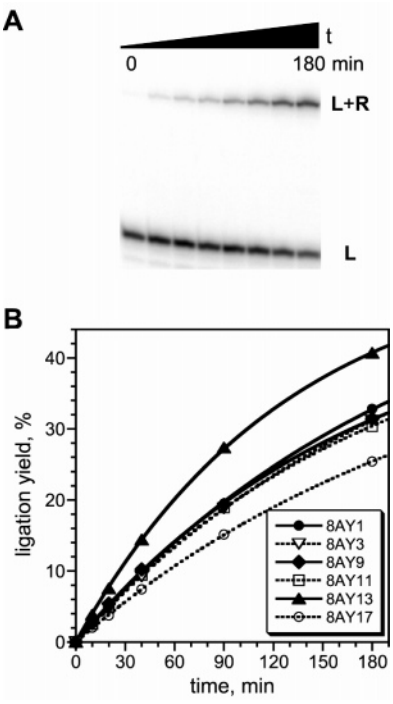

Figure 5. Kinetic analyses of the 8AY deoxyribozymes using the trimolecular format of Figure 3B. (A) Gel image for RNA ligation by 8AY13. Incubation conditions: 50 mM CHES, pH 9.0, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, 37 °C. (B) Kinetic data for all six 8AY deoxyribozymes. For 8AY13, $k_{obs} = 0.0077 \pm 0.0004$ min^{-1} ($n = 4$); for the other five deoxyribozymes, k_{obs} ranges from 0.004 to 0.006 min^{-1}. At timepoints >180 min under these conditions, small amounts of nonspecific degradation are often evident (not shown), and therefore the yields at these long timepoints are unreliable.

Secondary Structures of the 8AY Deoxyribozymes. The predicted secondary structures of the six representative 8AY deoxyribozymes were analyzed using the mfold program.[15] Without exception, the deoxyribozymes and their RNA substrates are predicted by mfold to adopt structures that maintain the designed arrangement of strategy 2 (see Supporting Information for representative secondary structures). In particular, the four nucleotides on either side of the ligation junction are continuously base-paired with the left-hand DNA binding arm. This confirms that the extended PCR primer (Figure 2) functions as intended to prevent the accumulation of PCR-derived mutations, which were previously observed at the analogous nucleotide positions in our earlier selection efforts.[7]

Kinetic Analyses of the 8AY Deoxyribozymes. The six representative 8AY deoxyribozymes were assayed to determine their rates and yields for RNA ligation. This was done in the trimolecular format of Figure 3B, in which neither of the RNA substrates are covalently linked to the DNA enzyme. For the "best" of the six DNA enzymes, 8AY13, the observed rate constant is $k_{obs} = 0.008$ min^{-1} (0.5 h^{-1}) at 40 mM Mg^{2+} and pH 9.0, with 41% yield of ligated RNA at 3 h (Figure 5). At longer timepoints at pH 9.0, nonspecific RNA degradation becomes a substantial concern (data not shown). At pH 7.5, the yield was about the same, 37% at 24 h incubation, with $k_{obs} = 0.06$ h^{-1} ($t_{1/2} = 12$ h). For all six deoxyribozymes, when the Mg^{2+} concentration was quadrupled from 40 to 160 mM at pH 9.0, the k_{obs} values increased by about 3-fold, indicating that

(15) (a) SantaLucia, J., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1460−1465. (b) Zuker, M. *Nucleic Acids Res.* **2003**, *31*, 3406−3415.

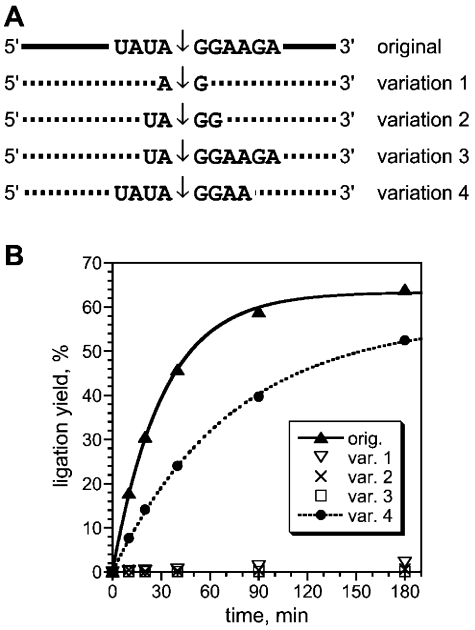

Figure 6. Assaying the generality of the 8AY13 deoxyribozyme with varying RNA substrate sequences. (A) Variations of the substrate sequences. The solid lines denote the original RNA sequences used during the selection procedure, and the arrow marks the ligation site. The dashed lines denote systematic RNA transversions A ↔ C and G ↔ U, with maintenance of Watson−Crick complementarity for all corresponding DNA nucleotides. (B) Kinetic assays using the modified substrate sequences. These were performed as in Figure 5 except with 160 mM Mg^{2+}, which led to higher k_{obs} values. For the original substrate sequences, 8AY13 showed $k_{obs} = 0.027 \pm 0.005$ min^{-1} ($n = 3$) under these conditions.

the deoxyribozymes require several tens of millimolar Mg^{2+} for full activity. Finally, because other RNA ligase deoxyribozymes have been effective with Mn^{2+} instead of Mg^{2+}, we tested 20 mM Mn^{2+} as the metal ion cofactor at pH 7.5. The yields are reduced several-fold (data not shown), indicating that Mn^{2+} supports catalytic activity much less effectively than does Mg^{2+} for this particular deoxyribozyme. Because the leaving group in the ligation reaction is pyrophosphate (Figure 1), we tested if inclusion of pyrophosphatase increases the ligation activity. No increase in rate or yield was observed (data not shown).

Assessing the Generality of the 8AY Deoxyribozymes for Ligating Varying RNA Substrate Sequences. An important question regarding the practical utility of RNA ligase deoxyribozymes is their generality for joining RNA substrates of varying sequence. In the arrangement of strategy 2 (Figure 2C), all of the RNA substrate nucleotides are initially base-paired with DNA nucleotides of the deoxyribozyme, as supported by the mfold analysis. Changing the RNA sequence therefore prompts alteration of the corresponding DNA nucleotide(s) to maintain Watson−Crick interactions, which are presumed to be required. For the most active of the new deoxyribozymes, 8AY13, we investigated its generality by systematically covarying the nucleotides of the RNA substrates and the complementary DNA binding arms via transversions A ↔ C and G ↔ U/T. An initial experiment in which only the two RNA nucleotides A↓G directly at the ligation junction were retained and all others were changed (Figure 6, variation 1) led to nearly complete loss of 8AY13 ligation activity. This indicates a more substantial RNA sequence requirement than is ideal for a general RNA ligase enzyme. Additional experiments in which more nucleotides were left unchanged (variations 2−4) are consistent with a requirement minimally at the five AUA↓GG positions, although the data allow that the requirement may extend as far as the eight UAUA↓GGAA positions. The precise requirement was not pursued experimentally in greater detail because regardless of the detailed outcome, none of the various possible requirements are sufficiently permissive for 8AY13 to be useful in the general preparative sense.

For the other five 8AY deoxyribozymes, analogous experiments also showed that retaining only A↓G is insufficient to support activity, and we established that the RNA sequence requirements are no more restrictive than UAUA↓GGAAGAGA (data not shown). However, we have not performed the large number of experiments necessary to determine more precisely the substrate sequence requirements for these other five deoxyribozymes.

Analysis of the 7AV Deoxyribozymes. We briefly examined the 7AV deoxyribozymes, which were obtained from the modified version of strategy 1 (Figure 2B) and collectively created 2′,5′-branched RNA (Figure 4A). Several 7AV clones were tested individually at 40 mM Mg^{2+} and pH 9.0, with $k_{obs} \approx 0.01$ min^{-1} and 60−70% ligation yield (data not shown). With 20 mM Mn^{2+} at pH 7.5, the k_{obs} values were also ~0.01 min^{-1} with ~80% yield despite the lower pH, indicating that Mn^{2+} is a more effective cofactor than Mg^{2+} for the 7AV deoxyribozymes, unlike for the 8AY deoxyribozymes. Such improvement in ligation activity upon replacing Mg^{2+} with Mn^{2+} has been observed previously with our unrelated branch-forming deoxyribozymes;[6,7] the detailed explanation for the Mn^{2+} preference in any of these cases is not yet known. Because the focus of the present study was on synthesis of linear and not branched RNA, the 7AV deoxyribozymes were not examined further.

Discussion

In vitro selection is typically used to obtain functional nucleic acid sequences.[16] However, our understanding of the selection process itself is vague, and undesired or unexpected outcomes are common (e.g., ref 17). Selection has been described as an "irrational" approach to function.[18] While one can debate the merits of this description, it is clear that rational control over the outcome of a selection process is an important practical objective. Our previous in vitro selection efforts have led only to linear 2′-5′-linked RNA (using the substrate combination of Figure 1A)[3−5] or to 2′,5′-branched RNA (using the substrate combination of Figure 1B);[6−8] we have never formed linear 3′-5′-linked RNA using deoxyribozymes. In sharp contrast, via careful modification of the selection strategy (Figure 2) and using the substrates of Figure 1B, all of the new 8AY deoxyribozymes (each of the six tested DNA enzymes and >95% of the pool before cloning) create solely 3′-5′ linkages (Figure 4). Therefore, we have successfully and rationally modified the selection procedure, leading to the desired outcome of linear 3′-5′-linked RNA. The ligated RNA products from the new deoxyribozymes are sought largely for structure−function studies of RNA folding and catalysis. Significantly, the ligation

(16) (a) Breaker, R. R. *Chem. Rev.* **1997**, *97*, 371−390. (b) Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1999**, *68*, 611−647. (c) Joyce, G. F. *Annu. Rev. Biochem.* **2004**, *73*, 791−836.
(17) Tuschl, T.; Sharp, P. A.; Bartel, D. P. *EMBO J.* **1998**, *17*, 2637−2650.
(18) Breaker, R. R.; Joyce, G. F. *Trends Biotechnol.* **1994**, *12*, 268−275.

event is traceless because the newly formed 3′-5′ linkage is indistinguishable from all of the other native linkages in the RNA product; a non-native linkage could lead to undesired structural or functional effects in the ligated RNA. In efforts that will be described elsewhere, we are also seeking to use the 2′,3′-cyclic phosphate substrate combination of Figure 1A to obtain native 3′-5′ linkages.

Catalytic Characteristics of the 8AY Deoxyribozymes. While the new 8AY deoxyribozymes achieve our principal objective of 3′-5′-linked RNA, several other goals have not been fully met. Specifically, the rates, yields, and generality of the deoxyribozymes are not optimal. The best new deoxyribozyme, 8AY13, has $k_{obs} = 0.06\ h^{-1}$ and 37% yield in the Figure 3B trimolecular format under typical RNA incubation conditions (pH 7.5, 40 mM Mg^{2+}, 24 h), with higher k_{obs} at elevated pH (Figure 5) and Mg^{2+} concentration (Figure 6). As we have observed for other deoxyribozyme selections,[5] we anticipate that further experimentation (e.g., using in vitro evolution) would result in catalysts with higher rates and yields.

The separate issue of generality with respect to the RNA substrate sequences can also be addressed. 8AY13 requires at most the RNA nucleotides UAUA↓GGAA at the ligation site; the requirement minimally includes the five AUA↓GG positions (Figure 6). However, any of these possibilities are rather restrictive requirements for practical purposes. For example, the probability of finding a particular five-nucleotide RNA sequence at a candidate ligation site is only about 1 in $4^5 \approx 0.1\%$ (assuming a random RNA sequence). Therefore, more work is necessary to obtain general 3′-5′ RNA ligase deoxyribozymes. As the mechanistic origin of the sequence requirement, we presume that the DNA nucleotides of the 8AY13 enzyme region interact directly with the required RNA nucleobases, which would be particularly accessible in the minor groove near the ligation site; high-resolution structural information is needed to assess this hypothesis. Ongoing efforts in our laboratory are checking if changes to the nucleotides directly at the ligation junction (A↓G) at the outset of selection provide more general RNA ligase deoxyribozymes, which may operate via different mechanisms. An untested possibility is that DNA enzymes with greater generality could be obtained by varying the RNA nucleotides that surround the ligation junction at each selection round. However, this would be technically challenging within the framework of this particular selection approach because the complementary DNA nucleotides would also have to be changed, which may interfere with the PCR step during each selection round. In particular, note that the PCR primer shown in Figure 2C would have to be changed, disrupting the required complementarity with the 3′-end of the deoxyribozyme product from the previous selection round.

Comparison to Other Nucleic Acid Enzyme Selections. The rational modification embodied in strategy 2 places the ligation junction within an RNA:DNA duplex region. Some of the earliest reported artificial ribozymes ligate RNA substrates in a roughly analogous fashion, although one of the two flanking duplexes was merely three base pairs long.[19] More importantly, two of the three classes of these RNA ligase ribozymes create non-native 2′-5′-linked RNA from a 5′-triphosphate, and the well-studied class I ligase ribozyme that creates 3′-5′ linkages has not retained the duplex arrangement at all.[19,20] In a different study, 3′-5′ linkages were obtained from a ribozyme built upon a *Tetrahymena* group I intron scaffold;[21] as found here, ligation occurred within a duplex region. The observation that the 8AY deoxyribozymes all create 3′-5′ linkages while retaining the designed duplex arrangement around the ligation site indicates that this general approach is critical for full control over the RNA ligation regioselectivity.

Conclusions

We have rationally modified our previously established deoxyribozyme selection strategy to obtain native 3′-5′-linked RNA by attack of a 3′-hydroxyl at a 5′-triphosphate. The approach successfully provides 3′-5′ linkages, which have eluded synthesis in all of our previous studies. For each of the new deoxyribozymes, sequence requirements for the RNA substrates are found near the ligation junction, and these requirements restrict the general practical application of the deoxyribozymes. Further development of this and other selection approaches should lead to deoxyribozymes with substantial practical utility for RNA ligation with formation of native 3′-5′ linkages.

Experimental Section

Selection Procedure and Cloning. The selection procedure was described previously, including sources of material.[3] The left-hand (L) RNA substrate was 5′-UAAUACGACUCACUAUA-3′, preceded by 5′-GGA... when appropriate to permit in vitro transcription. The right-hand (R) substrate was 5′-GGAAGAGAUGGCGACGG-3′, prepared by in vitro transcription with a 5′-triphosphate. Modified versions of the RNA substrate sequences were prepared by in vitro transcription. For the AV selection, the 83-mer deoxyribozyme strand was identical to that used previously.[3] For the AY selection, the 89-mer deoxyribozyme strand was 5′-CGAAGTCGCCATCTC-N_{38}-TTCCTATAGT-GAGTCGTATTA-AGCTGATCCTGATGG-3′. The underlined regions are complementary to the RNA substrates according to the arrangement of Figure 2C; N denotes an equal probability of incorporating any of the four deoxynucleotides, corrected for unequal coupling efficiencies of the four phosphoramidites. In each round, the key selection step of Figure 3A was performed in 50 mM CHES, pH 9.0, 150 mM NaCl, 2 mM KCl, and 40 mM $MgCl_2$ at 37 °C for 2 h. Individual deoxyribozymes were cloned as described,[3] except that the PCR primers for cloning additionally incorporated TAA stop codons in all three frames to avoid false negatives.[22] The survey of kinetic activities was performed using the procedure described previously,[5] pursuing those deoxyribozymes with ligation yields >20% at the longest timepoint.

Assaying Branched Versus Linear RNA Products. The uncloned selection pools were assayed either by partial alkaline hydrolysis (Figure 4A), by cleavage with the 8-17 deoxyribozyme, or by cleavage with 100 mM Mg^{2+} at pH 9.0 in the presence of the DNA complement as described.[3,6] Ligation products created by individual deoxyribozymes (Figure 4B,C) were assayed in the same way.

Kinetic Analyses. The kinetic analyses (Figures 5 and 6) were performed in the trimolecular assay format of Figure 3B, using deoxyribozymes prepared by solid-phase synthesis. The ^{32}P-radiolabeled left-hand RNA substrate L was the limiting reagent relative to the deoxyribozyme E and right-hand substrate R. The ratio of L:E:R was 1:5:15, with the concentration of E equal to ~0.5 μM. The incubation conditions were 50 mM CHES, pH 9.0, 150 mM NaCl, 2 mM KCl, and 40 mM $MgCl_2$ at 37 °C (or 50 mM HEPES, pH 7.5 with all other

(19) (a) Bartel, D. P.; Szostak, J. W. *Science* **1993**, *261*, 1411−1418. (b) Ekland, E. H.; Szostak, J. W.; Bartel, D. P. *Science* **1995**, *269*, 364−370.

(20) Bergman, N. H.; Lau, N. C.; Lehnert, V.; Westhof, E.; Bartel, D. P. *RNA* **2004**, *10*, 176−184.

(21) Jaeger, L.; Wright, M. C.; Joyce, G. F. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14712−14717.

(22) Langner, J.; Klussmann, S. *BioTechniques* **2003**, *34*, 950−954.

components unchanged). See our earlier report for a detailed description of the methods of sample preparation and ligation analysis.[3] In all cases, values of k_{obs} and final yields were obtained by fitting the yield versus time data directly to first-order kinetics; that is, yield $= Y(1 - e^{-kt})$, where $k = k_{obs}$ and Y = final yield. Increasing the concentration of E or R had no effect on the observed kinetics, confirming that all L is bound in the ternary L:E:R complex.

Acknowledgment. This research was supported by the Burroughs Wellcome Fund (New Investigator Award in the Basic Pharmacological Sciences), the March of Dimes Birth Defects Foundation (Research Grant 5-FY02-271), the National Institutes of Health (GM-65966), the American Chemical Society Petroleum Research Fund (38803-G4), and the UIUC Department of Chemistry (all to S.K.S.), and by a Sigma Xi Grant-in-Aid of Research (to R.L.C). S.K.S. is the recipient of a fellowship from The David and Lucile Packard Foundation. We thank members of the Silverman lab for discussions.

Supporting Information Available: Sequences and mfold-predicted secondary structures of the 8AY deoxyribozymes. This material is available free of charge via the Internet at http://pubs.acs.org.

JA045817X