Adenosine Is Inherently Favored as the Branch-Site RNA Nucleotide in a Structural Context That Resembles Natural RNA Splicing†

Elena Zelin,‡ Yangming Wang,‡ and Scott K. Silverman*

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

Received December 8, 2005; Revised Manuscript Received January 17, 2006

ABSTRACT: We previously used in vitro selection to identify the 7S11 deoxyribozyme, which catalyzes formation of 2′,5′-branched RNA using a branch-site adenosine nucleophile and a 5′-triphosphate electrophile. An unanswered question is whether the use of branch-site adenosine is inherently preferred or a chance event during the particular selection experiment. Here we have found that deoxyribozymes newly selected to use uridine as the branch-site RNA nucleotide in a structural context that resembles natural RNA splicing instead prefer a branch-site adenosine, although adenosine was never available during the selection itself. Our results support a chemical basis for nature's choice of the branch-site nucleotide, which is almost always adenosine in group II introns and the spliceosome.

Studies of artificial RNA and DNA enzymes (ribozymes and deoxyribozymes) reveal fundamental principles of nucleic acid catalysis (*1*). In many cases, nucleic acid enzymes are also useful because of their synthetic abilities (*2*). We recently reported in vitro selection of the 7S11 deoxyribozyme that creates 2′,5′-branched RNA (*3*−*5*), which is an important intermediate in natural RNA splicing (*6*−*8*). 7S11 catalyzes the nucleophilic attack of a branch-site adenosine 2′-hydroxyl group on a 5′-triphosphate or a related electrophile, forming 2′,5′-branched RNA. The branch-site adenosine is an unpaired nucleotide flanked by Watson−Crick duplex regions, which resembles the context of natural RNA splicing by group II introns (*6*, *7*) and the spliceosome (*8*). For 7S11, a branch-site guanosine is tolerated but with an ∼50-fold lower k_{obs},[1] and neither uridine nor cytidine is accepted above trace levels (*3*). Curiously, the preference of 7S11 for branch-site adenosine mimics natural RNA splicing systems, which use branch-site adenosine almost exclusively (*9*). The branch-site adenosine used by 7S11 was not predetermined by the selection design (*3*). Therefore, a key unanswered question is whether this preference is an inherent chemical characteristic of forming branched RNA or a coincidental outcome of the single selection experiment. To resolve this ambiguity, here we have performed a new selection that allows branch-site uridine to react in the same structural context. The results clearly demonstrate an inherent preference in favor of branch-site adenosine, which suggests a chemical basis for nature's choice of the branch-site RNA nucleotide.

† This research was supported by the National Institutes of Health (Grant GM-65966). S.K.S. is the recipient of a fellowship from The David and Lucile Packard Foundation.

* To whom correspondence should be addressed. Phone: (217) 244-4489. Fax: (217) 244-8024. E-mail: scott@scs.uiuc.edu.

‡ Department of Biochemistry, University of Illinois.

[1] Abbreviations: PAGE, polyacrylamide gel electrophoresis; CHES, 2-(*N*-cyclohexyl)aminoethanesulfonic acid; EDTA, ethylenediaminetetraacetic acid; k_{obs}, observed rate constant.

EXPERIMENTAL PROCEDURES

General Considerations. DNA oligonucleotides were prepared at IDT (Coralville, IA) or at the University of Illinois W. M. Keck Center. 5′-Triphosphorylated RNA oligonucleotides were prepared by in vitro transcription using T7 RNA polymerase with an appropriate double-stranded DNA template that was prepared by annealing two DNA oligonucleotides (*10*). All DNA and RNA oligonucleotides were purified by denaturing PAGE with running buffer 1× TBE [89 mM Tris, 89 mM boric acid, and 2 mM EDTA (pH 8.3)] as described previously (*11*, *12*). The substrate providing the 2′-hydroxyl is denoted the left-hand (L) substrate, and the substrate providing the 5′-triphosphate is denoted the right-hand (R) substrate.

Selection Design, Procedure, and Progression. For the selection process, the substrates and DNA were preorganized into the arrangement of the 7S11 deoxyribozyme (*3*−*5*) as shown in Figure 1, except the 3′-end of the right-hand RNA substrate was joined to the 5′-end of the DNA by a small loop during selection as described previously (*11*). The sequence of the left-hand RNA substrate was 5′-GGAUAAUACGUCUCAC-3′, where the branch-site uridine is underlined. The sequence of the right-hand RNA substrate was 5′-GGAAGAGAUGGCGACGG-3′ with a 5′-triphosphate. The sequence of the deoxyribozyme pool strand was 5′-CGAAGTCGCCATCTC-N_{15}-GTGAG-N_7-TTCCCGTATTATCC-3′, where the four DNA nucleotides that interact with the 5′-end of the R substrate to form P4 are underlined. The sequences of the two PCR primers were 5′-$(AAC)_4$-XCCATCAGGATCAGCTGGATAATACGGGAA-3′ (where X denotes Glen spacer 18 to stop Taq polymerase) and 5′-CGAAGTCGCCATCTC-3′. The second primer was 5′-phosphorylated to permit immediate ligation of the right-hand RNA substrate using T4 RNA ligase before the key selection step of the next round (*11*). The selection procedure was performed essentially as described previously (*11*). The selection progression (i.e., ligation activity vs round number)

10.1021/bi052499l CCC: $33.50
Published on Web 02/07/2006

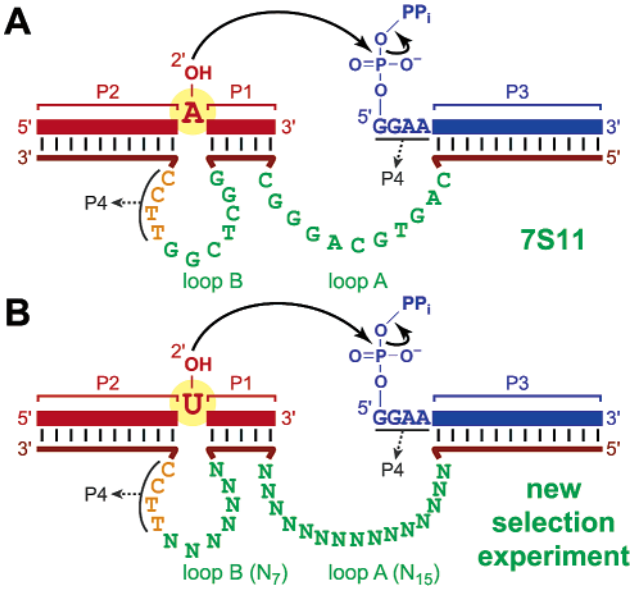

FIGURE 1: Deoxyribozymes for branched RNA synthesis. (A) Secondary structure of the 7S11 deoxyribozyme complexed with its RNA substrates. (B) The new selection experiment to test utilization of branch-site uridine.

is shown in Figure 2. See Results and Discussion for detailed descriptions of the selection experiments.

Sequences and Ligation Activities of the 10DM and 7DM Deoxyribozymes. After cloning of individual round 10 deoxyribozymes but before sequencing, activities were surveyed using the standard incubation conditions of 50 mM CHES (pH 9.0), 150 mM NaCl, 2 mM KCl, and 40 mM $MgCl_2$ at 37 °C (data not shown). On the basis of these data, most of the tested round 10 clones were sequenced, and six unique deoxyribozyme sequences were identified (Figure 3). 10DM21 has a one-nucleotide insertion in loop A (on the left in Figure 3), presumably arising from Taq polymerase during selection. Each of these 10DM deoxyribozymes was prepared independently by solid-phase synthesis, and their ligation activities were assayed with all four branch-site RNA nucleotides (e.g., Figure 4; see the Supporting Information for detailed kinetic data for all new deoxyribozymes). For all 10DM deoxyribozymes, a clear branch-site adenosine preference was observed (Table 1). The 7DM deoxyribozymes were obtained after several selection rounds with branch-site G as shown in Figure 2; they were surveyed, sequenced, synthesized, and assayed in the same fashion as the 10DM deoxyribozymes. Again, in all cases, a clear branch-site adenosine preference was observed (Table 1).

For assays of deoxyribozyme ligation activities using DNA from solid-phase synthesis, the general approach has been described previously (*11*). The 5′-^{32}P-radiolabeled L substrate was the limiting reagent relative to the deoxyribozyme (E) and the R substrate. The L:E:R ratio was 1:5:15, with the concentration of E equal to ~0.5 μM. Values of k_{obs} were obtained by fitting the yield versus time data directly to first-order kinetics; i.e., yield $= Y(1 - e^{-kt})$, where $k = k_{obs}$ and Y is the final yield. However, when insufficient curvature in the data was observed to allow reasonable convergence by the curve fit algorithm or when k_{obs} was sufficiently low (<0.01 min^{-1}), values of k_{obs} were estimated from linear fits to the early time points (10−180 min as appropriate).

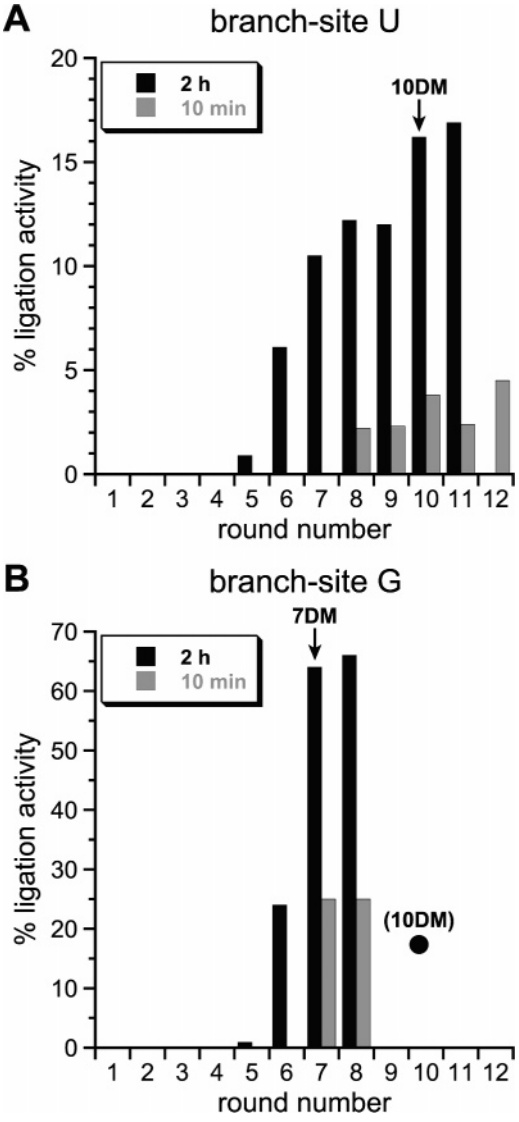

FIGURE 2: Selection progression. (A) Progression using a left-hand RNA substrate with branch-site uridine. During the key selection step, the incubation time was 2 h or 10 min for each round as indicated. The 10DM deoxyribozymes were cloned from the pool indicated by the arrow. (B) Revisiting the selection experiment, restarting at round 5 but switching to a left-hand RNA substrate with branch-site guanosine. The incubation time was 2 h for rounds 5 and 6 and then either 2 h or 10 min as indicated for rounds 7 and 8. Note the difference in *y*-axis values between the graphs in panels A and B. The filled circle in panel B denotes the activity at round 10 from the branch-site U experiments in panel A (the 10DM deoxyribozymes were cloned from that pool). The 7DM deoxyribozymes were cloned from the pool indicated by the arrow.

RESULTS AND DISCUSSION

Design of a New Selection Experiment To Test Usage of Branch-Site Nucleotides. The 7S11 deoxyribozyme binds its two RNA substrates in a three-helix-junction complex with four RNA−DNA duplexes designated P1−P4 (Figure 1A) (*4*). The left-hand RNA substrate that has the branch-site adenosine forms P1 and P2 that surround the branch site, whereas the right-hand RNA substrate forms P3 and P4. Two nominally single-stranded DNA regions (loops A and B) are interspersed among the P1−P4 RNA−DNA duplexes. To determine if branch-site pyrimidine nucleotides can be utilized by deoxyribozymes that are structurally analogous to 7S11, we employed a RNA substrate in which uridine

Table 1: k_{obs} Values and Branch-Site Adenosine Selectivities for the 10DM and 7DM Deoxyribozymes[a]

deoxyribozyme	k_{obs}(A) (min^{-1})	k_{obs}(G) (min^{-1})	k_{obs}(U) (min^{-1})	k_{obs}(C) (min^{-1})	k_{obs}(A)/k_{obs}(G)	k_{obs}(A)/k_{obs}(U)	k_{obs}(A)/k_{obs}(C)
7S11	0.56	0.010	0.0002	0.0003	56	2800	1870
10DM24	0.26	0.035	0.0035	0.0030	7.4	74	87
10DM3	0.18	0.042	0.0014	0.012	4.3	130	15
10DM5	0.32	0.019	0.0019	0.0017	17	170	190
10DM10	0.31	0.026	0.015	0.0050	12	21	62
10DM19	0.27	0.019	0.0013	0.0051	14	210	53
10DM21	0.19	0.036	0.0030	0.0065	5.3	63	29
7DM3	0.11	0.0018	0.00023	0.00015	61	480	730
7DM11	0.13	0.0051	0.00013	0.00012	25	1000	1080
7DM12	0.15	0.015	0.00076	0.0014	10	200	110

[a] Deoxyribozymes were prepared by solid-phase synthesis, and kinetic assays were performed like those described in the legend of Figure 4. See the Supporting Information for full data sets. The three right-most columns show the adenosine selectivity as quantified by a ratio of k_{obs} values; e.g., $k_{obs}(A)/k_{obs}(U) = k_A/k_U$. For representative error estimates of the kinetic values, see Figure 4. Data for 7S11 are from Figure 4 of ref *4*.

7S11 5'-..... ----CAGTGCAGGGCGTGAGGGCTCGGTTCC.....-3'
10DM24 5'-..... CCGTAGGTGAAGGGCGTGAGGGTTCCATTCC.....-3'
10DM3 5'-..... CGGTAAGGCCAGGGCGTGAGGGTCCGCTTCC.....-3'
10DM5 5'-..... CATTATGCGAAGGGCGTGAGGGTTCCGTTCC.....-3'
10DM10 5'-..... CCTGTGGCAAAGGGCGTGAGGGTACTGTTCC.....-3'
10DM19 5'-..... CCTATGGCCCAGGGCGTGAGGGTGCGGTTCC.....-3'
10DM21 5'-.....AGTGTGCTGCTAGGGCGTGAGGGTCCGCTTCC.....-3'
7DM3 5'-..... CCGGCAGGCAAGGGTGTGAGGGCTCGGTTCC.....-3'
7DM11 5'-..... CCGTAGGCAAAGGGCGTGAGGGCTCGGTTCC.....-3'
7DM12 5'-..... CATTATGCCCAGGGCGTGAGGGTGCGGTTCC.....-3'

FIGURE 3: Sequences of the 10DM and 7DM deoxyribozymes, colored as in Figure 1. Pink nucleotides denote sequence differences between each individual deoxyribozyme and 10DM24. As expected, loop A of 7S11 (at the left) is four nucleotides shorter than for the other deoxyribozymes (see Figure 1).

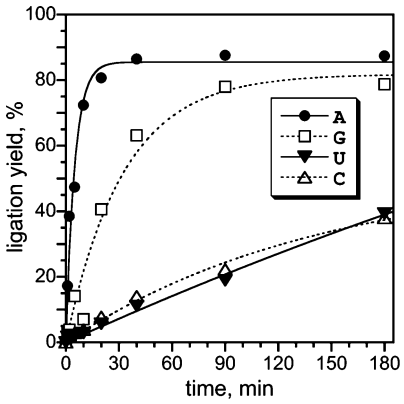

FIGURE 4: Branched RNA formation by the 10DM24 deoxyribozyme shows a clear preference for branch-site adenosine. k_{obs} values (±standard error of the mean; $n = 3$ in each case): 0.26 ± 0.02 min^{-1} for A, 0.035 ± 0.002 min^{-1} for G, 0.0035 ± 0.0011 min^{-1} for U, and 0.0030 ± 0.0007 min^{-1} for C ($k_A/k_U = 73$). For comparison, 7S11 synthesizes A-branched RNA with a k_{obs} of ≈0.5 min^{-1} and a k_A/k_U of ≈3000 (*3*, *4*). See Table 1 for similar data for all new deoxyribozymes.

replaces the branch-site adenosine. A selection pool was synthesized with all DNA nucleotides of 7S11 loops A and B fully randomized, but the other DNA nucleotides base-paired with RNA to maintain P1−P4 (Figure 1B). Because 11-nucleotide loop A of 7S11 tolerates deletions or insertions at certain locations (*4*), we increased its length to 15 nucleotides (N_{15}), which along with seven-nucleotide loop B (N_7) made a total of 22 random DNA nucleotides. Because our starting pool has approximately 200 pmol (10^{14} molecules) whereas there are only 4^{22} (2×10^{13}) possible 22-mer sequences, all of the N_{22} DNA sequence space is examined in this new selection experiment.

Results of the New Selection Experiment. In vitro selection for RNA ligation using the uridine as the branch-site nucleotide was performed with our previous protocol (*11*). Incubation conditions during the key branch-formation step included 50 mM CHES (pH 9.0), 150 mM NaCl, 2 mM KCl, and 40 mM $MgCl_2$ at 37 °C for 2 h. Beginning at round 5, unambiguous RNA ligation activity of the pool was detected, and after a total of 10 selection rounds, ~17% ligation was observed (black bars in Figure 2A). We attempted to improve the branch-site uridine activity by increasing the selection pressure. When rounds 8−12 were performed with an incubation time of 10 min instead of 2 h, the ligation activity of the pool remained low (<5%; gray bars in Figure 2A). This indicated that the longer 2 h incubation time was necessary for the observed U-branching activity.

Individual deoxyribozymes such as 10DM24 (Figure 3) were cloned from the round 10 pool. The new deoxyribozymes were tested for branch-formation activity with RNA substrates having each of the four possible branch-site nucleotides. Synthesis of the expected 2′,5′-branched RNA products was verified by partial alkaline hydrolysis (see the Supporting Information). Unexpectedly, a clear branch-site purine preference was evident; the trend of ligation rates and yields was as follows: A > G > U ≈ C [Figure 4; this is similar to the trend for 7S11 (*3*, *4*]. This was surprising because the DNA pool in the new selection experiment was never presented with an RNA substrate having a branch-site purine. The strong adenosine preference was found for each of the six tested round 10 deoxyribozymes, with a selectivity (as quantified by k_A/k_U) of ~20−200 (Table 1). These new observations provide strong evidence that the branch-site adenosine preference is inherent to the 7S11 structural context and is not a fortuitous outcome of any particular selection experiment.

Revisiting the Selection Experiment with Branch-Site Purine. The branch-site purines supported relatively good ligation activity with the new 10DM deoxyribozymes, even though the selection procedure never used a branch-site purine. We considered that this robust ligation activity with branch-site purines might emerge explicitly (rather than after the fact) if branch-site purines were directly provided in the selection substrate. For this purpose, we revisited the new selection experiment beginning with round 5, where branch-site U activity first appeared, and switched to providing a RNA substrate with branch-site G. (G was used instead of A because 7S11 and the new round 10 clones already work well with branch-site A, and we also sought improved

practical G-branching activity.) Within three selection rounds under the same incubation conditions that were originally used with branch-site U (maximal pool activity of ~17%), the pool activity increased sharply to ~65% (black bars in Figure 2B). This is the largest RNA ligation activity that we have seen in any selection pool to date, encompassing several dozen selection efforts of various kinds. When rounds 7 and 8 were performed in parallel with branch-site G and a 10 min incubation instead of 2 h, the selection activity was substantially lower (gray bars in Figure 2B), suggesting that maximal G-branching activity had been achieved. Cloning from round 7 after a 2 h incubation provided several deoxyribozymes that resemble those from round 10 with regard to DNA sequence, ligation activity, and branch-site preference [A > G > U ≈ C (Figure 3 and Table 1)]. This further confirms that the branch-site adenosine preference is inherent to the 7S11 structural context.

Implications for General Synthesis of Branched RNA. Our results have implications for the practical synthesis of 2′,5′-branched RNA. Although the 7S11 deoxyribozyme is quite general with respect to its RNA substrate sequences, one "difficult" substrate sequence combination gave a <5% ligation yield (*4*). Here, in the context of branch-site adenosine, we examined the ability of 10DM24 to ligate these difficult RNA substrates (see the Supporting Information for detailed data of the following experiments). Although the selection procedure did not demand improved generality in the RNA−DNA binding regions, 10DM24 provided substantially higher yields of branched RNA than 7S11 with the difficult substrates (~60% vs ~4% yield in 3 h). It was possible that the four-nucleotide difference in loop A length between 7S11 and 10DM24 (Figure 3) is the source of this improved generality. Therefore, the difficult RNA substrates were examined with a modified version of 7S11 in which four nucleotides were inserted into loop A, making it the same 15 nucleotides in length as that for 10DM24. However, these additional nucleotides did not improve the generality of 7S11, indicating that merely having a longer loop A does not lead to improved generality. We also identified six specific nucleotides in loops A and B that differ between 7S11 and 10DM24 (Figure 3, pink nucleotides in the 7S11 sequence). Ligation of the difficult RNA substrates was examined using six mutated versions of 7S11 in which five of these six nucleotides were changed to those found in 10DM24. In all cases, no substantial increase in ligation activity was observed. This indicates that all six of the nucleotide differences between 7S11 and 10DM24 (in addition to the four extra nucleotides in loop A) are required for the improved generality of the latter deoxyribozyme. We have already used 10DM24 in our laboratory to prepare several branched RNA products in high yield, where 7S11 gave relatively poor yields (90% vs <1%; Y. Wang and S. K. Silverman, unpublished results). On the basis of these observations, we recommend that others use 10DM24 instead of 7S11 for the practical synthesis of 2′,5′-branched RNA.

Context of the Branch-Site Adenosine Preference. In a recently discovered natural ribozyme, a branch-site uridine creates a three-nucleotide mini-lariat that caps the 5′-terminus of a mRNA (*13*). In an accompanying commentary (*15*), our recent identification of a deoxyribozyme unrelated to 7S11 that uses any branch-site nucleotide (*14*) was cited as evidence that adenosine has no special dispensation for reacting as the branch-site nucleotide. Indeed, several of our independently identified deoxyribozymes use branch-site pyrimidines quite well (*14*, *16*, *17*). However, none of these other deoxyribozymes has the bulged branch-site nucleotide that is characteristic of 7S11 and 10DM24 (Figure 1). It is this structural context, which is shared by group II introns and the spliceosome (*6*−*8*) but apparently not the newly reported capping ribozyme (*13*), in which our results demonstrate an inherent branch-site adenosine preference.

Chemical Basis for Adenosine Preference. The chemical basis for the branch-site adenosine preference is unclear. The possibility of an inhibitory base-pairing interaction between the conserved first G nucleotide of loop B and the branch-site RNA nucleotide is inconsistent with data obtained for several loop B mutants, which retain branch-site A selectivity despite the possibility of such a base pair (see the Supporting Information). An adenosine nucleotide has a lower 2′-hydroxyl pK_a value than either G, U, or C by ~0.3−0.4 units (*18*), which may contribute to adenosine's increased reactivity. However, an increase of merely 2.0−2.5-fold ($10^{0.3-0.4}$) in acidity for adenosine seems insufficient to explain the conserved reactivity trend. Studies of group II introns have identified several interactions that involve the nucleobase of the branch-site adenosine (*19*). However, it is not known if these interactions are responsible for the adenosine selectivity or if they merely reinforce a preference that exists for other reasons. Despite these uncertainties in the underlying chemical explanation, the results reported here unambiguously establish that adenosine is inherently preferred as the branch-site nucleotide in a structural context that closely resembles natural RNA splicing.

SUPPORTING INFORMATION AVAILABLE

Ligation activity data for all new deoxyribozymes, partial alkaline hydrolysis for confirming branch-site nucleotides, assays for investigating 10DM24 generality for the RNA substrate sequences, and data inconsistent with base pairing of the branch-site RNA nucleotide. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Joyce, G. F. (2004) Directed Evolution of Nucleic Acid Enzymes, *Annu. Rev. Biochem. 73*, 791−836.
2. Silverman, S. K. (2004) Deoxyribozymes: DNA catalysts for bioorganic chemistry, *Org. Biomol. Chem. 2*, 2701−2706.
3. Coppins, R. L., and Silverman, S. K. (2004) A DNA Enzyme that Mimics the First Step of RNA Splicing, *Nat. Struct. Mol. Biol. 11*, 270−274.
4. Coppins, R. L., and Silverman, S. K. (2005) A Deoxyribozyme That Forms a Three-Helix-Junction Complex with Its RNA Substrates and Has General RNA Branch-Forming Activity, *J. Am. Chem. Soc. 127*, 2900−2907.
5. Coppins, R. L., and Silverman, S. K. (2005) Mimicking the First Step of RNA Splicing: An Artificial DNA Enzyme Can Synthesize Branched RNA Using an Oligonucleotide Leaving Group as a 5′-Exon Analogue, *Biochemistry 44*, 13439−13446.
6. Michel, F., and Ferat, J. L. (1995) Structure and activities of group II introns, *Annu. Rev. Biochem. 64*, 435−461.
7. Jacquier, A. (1996) Group II introns: Elaborate ribozymes, *Biochimie 78*, 474−487.
8. Burge, C. B., Tuschl, T., and Sharp, P. A. (1999) in *The RNA World* (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) 2nd ed., pp 525−560, Cold Spring Harbor Laboratory Press, Plainview, NY.
9. Chu, V. T., Liu, Q., Podar, M., Perlman, P. S., and Pyle, A. M. (1998) More than one way to splice an RNA: Branching without

a bulge and splicing without branching in group II introns, *RNA 4*, 1186−1202.
10. Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates, *Nucleic Acids Res. 15*, 8783−8798.
11. Flynn-Charlebois, A., Wang, Y., Prior, T. K., Rashid, I., Hoadley, K. A., Coppins, R. L., Wolf, A. C., and Silverman, S. K. (2003) Deoxyribozymes with 2′−5′ RNA Ligase Activity, *J. Am. Chem. Soc. 125*, 2444−2454.
12. Wang, Y., and Silverman, S. K. (2003) Characterization of Deoxyribozymes That Synthesize Branched RNA, *Biochemistry 42*, 15252−15263.
13. Nielsen, H., Westhof, E., and Johansen, S. (2005) An mRNA is capped by a 2′,5′ lariat catalyzed by a group I-like ribozyme, *Science 309*, 1584−1587.
14. Pratico, E. D., Wang, Y., and Silverman, S. K. (2005) A Deoxyribozyme That Synthesizes 2′,5′-Branched RNA with Any Branch-Site Nucleotide, *Nucleic Acids Res. 33*, 3503−3512.
15. Pyle, A. M. (2005) Capping by branching: A new ribozyme makes tiny lariats, *Science 309*, 1530−1531.
16. Wang, Y., and Silverman, S. K. (2003) Deoxyribozymes That Synthesize Branched and Lariat RNA, *J. Am. Chem. Soc. 125*, 6880−6881.
17. Wang, Y., and Silverman, S. K. (2005) Efficient One-Step Synthesis of Biologically Related Lariat RNAs by a Deoxyribozyme, *Angew. Chem., Int. Ed. 44*, 5863−5866.
18. Acharya, S., Foldesi, A., and Chattopadhyaya, J. (2003) The pK_a of the Internucleotidic 2′-Hydroxyl Group in Diribonucleoside (3′→5′) Monophosphates, *J. Org. Chem. 68*, 1906−1910.
19. Liu, Q., Green, J. B., Khodadadi, A., Haeberli, P., Beigelman, L., and Pyle, A. M. (1997) Branch-site selection in a group II intron mediated by active recognition of the adenine amino group and steric exclusion of non-adenine functionalities, *J. Mol. Biol. 267*, 163−171.

BI052499L