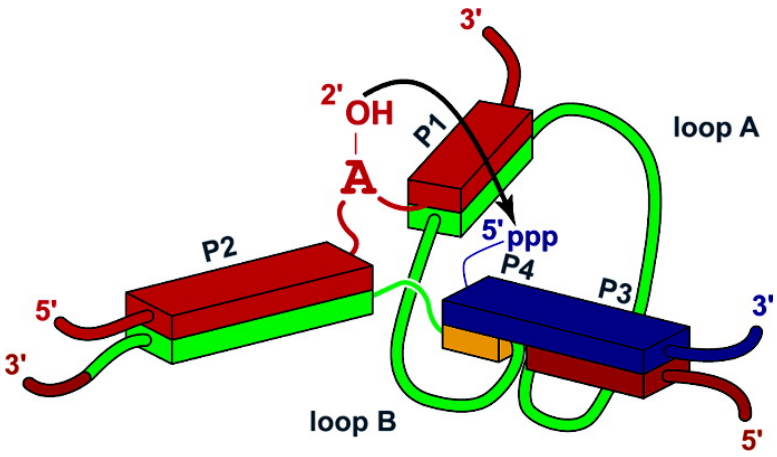
Article

A Deoxyribozyme that Forms a Three-Helix-Junction Complex with its RNA Substrates and has General RNA Branch-Forming Activity

Rebecca L. Coppins, and Scott K. Silverman

J. Am. Chem. Soc., **2005**, 127 (9), 2900-2907• DOI: 10.1021/ja044881b • Publication Date (Web): 12 February 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 8 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 02/12/2005

A Deoxyribozyme that Forms a Three-Helix-Junction Complex with its RNA Substrates and has General RNA Branch-Forming Activity

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 August 24, 2004; E-mail: scott@scs.uiuc.edu

Abstract: We recently used in vitro selection to identify 7S11, a deoxyribozyme that synthesizes 2′,5′-branched RNA. The 7S11 DNA enzyme mediates the nucleophilic attack of an adenosine 2′-hydroxyl group at a 5′-triphosphate, forming 2′,5′-branched RNA in a reaction that resembles the first step of in vivo RNA splicing. Here, we describe 7S11 characterization experiments that have two important implications for nucleic acid chemistry and biochemistry. First, on the basis of a comprehensive analysis of its substrate sequence requirements, 7S11 is shown to be generally applicable for the synthesis of a wide range of 2′,5′-branched RNAs. Strict substrate sequence requirements are found at the two RNA nucleotides that directly form the branched linkage, and these requirements correspond to those nucleotides found most commonly at these two positions in natural spliced RNAs. Outside of these two nucleotides, most substrate sequences are tolerated with useful ligation activity, although rates and yields vary. Because chemical synthesis approaches to branched RNA are extremely limited in scope, the deoxyribozyme-based route using 7S11 will enable many experiments that require branched RNA. Second, comprehensive nucleotide covariation experiments demonstrate that 7S11 and its RNA substrates adopt a three-helix-junction structure in which the branch-site nucleotide is located at the intersection of the three helices. Because many natural ribozymes have multi-helix junctions, 7S11 is an interesting model system for catalytic nucleic acids.

RNA molecules with an internal 2′,5′-branch (Figure 1A) play important roles in biochemistry. Branched RNAs are intermediates in RNA splicing by group II introns and the spliceosome,[1,2] and they have recently been proposed to participate in Ty1 retrotransposition.[3] For investigating these processes, a significant chemical challenge is the synthesis of branched RNAs without using natural splicing enzymes, which have many specific sequence requirements. These requirements may be incompatible with synthetic targets, such as mis-spliced versions of intron RNAs. Various chemical approaches to artificial branched RNA are based on solid-phase synthesis,[4−7] but these methods are often severely constrained by practical requirements. For example, the synthetic method often demands that two of the three RNA strands extending from the branch-site nucleotide have identical sequences,[5] which restricts subsequent application of the branched RNA. Circumventing this limitation requires either separation of complex mixtures of branched RNAs[6] or tedious orthogonal protecting-group strategies.[7] In

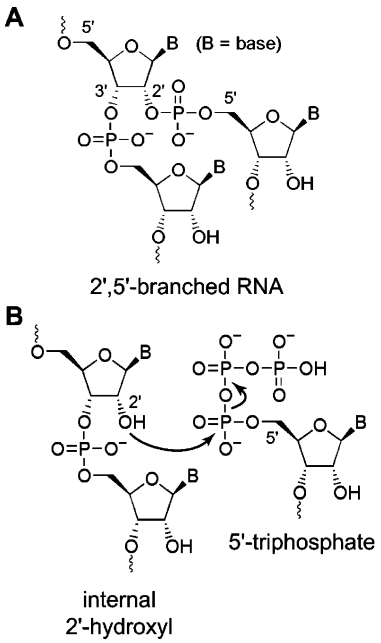


Figure 1. Structure and synthesis of 2′,5′-branched RNA. (A) Structure of 2′,5′-branched RNA. (B) Synthesis of branched RNA by attack of an internal 2′-hydroxyl on a 5′-triphosphate.

(1) Domdey, H.; Apostol, B.; Lin, R. J.; Newman, A.; Brody, E.; Abelson, J. *Cell* **1984**, *39*, 611−621.
(2) Peebles, C. L.; Perlman, P. S.; Mecklenburg, K. L.; Petrillo, M. L.; Tabor, J. H.; Jarrell, K. A.; Cheng, H. L. *Cell* **1986**, *44*, 213−223.
(3) (a) Cheng, Z.; Menees, T. M. *Science* **2004**, *303*, 240−243. (b) Perlman, P. S.; Boeke, J. D. *Science* **2004**, *303*, 182−184.
(4) Damha, M. J.; Ogilvie, K. K. *J. Org. Chem.* **1988**, *53*, 3710−3722.
(5) (a) Damha, M. J.; Zabarylo, S. *Tetrahedron Lett.* **1989**, *30*, 6295−6298. (b) Damha, M. J.; Ganeshan, K.; Hudson, R. H.; Zabarylo, S. V. *Nucleic Acids Res.* **1992**, *20*, 6565−6573.
(6) Ganeshan, K.; Tadey, T.; Nam, K.; Braich, R.; Purdy, W. C.; Boeke, J. D.; Damha, M. J. *Nucleosides Nucleotides* **1995**, *14*, 1009−1013.
(7) Reese, C. B.; Song, Q. *Nucleic Acids Res.* **1999**, *27*, 2672−2681.

10.1021/ja044881b CCC: $30.25

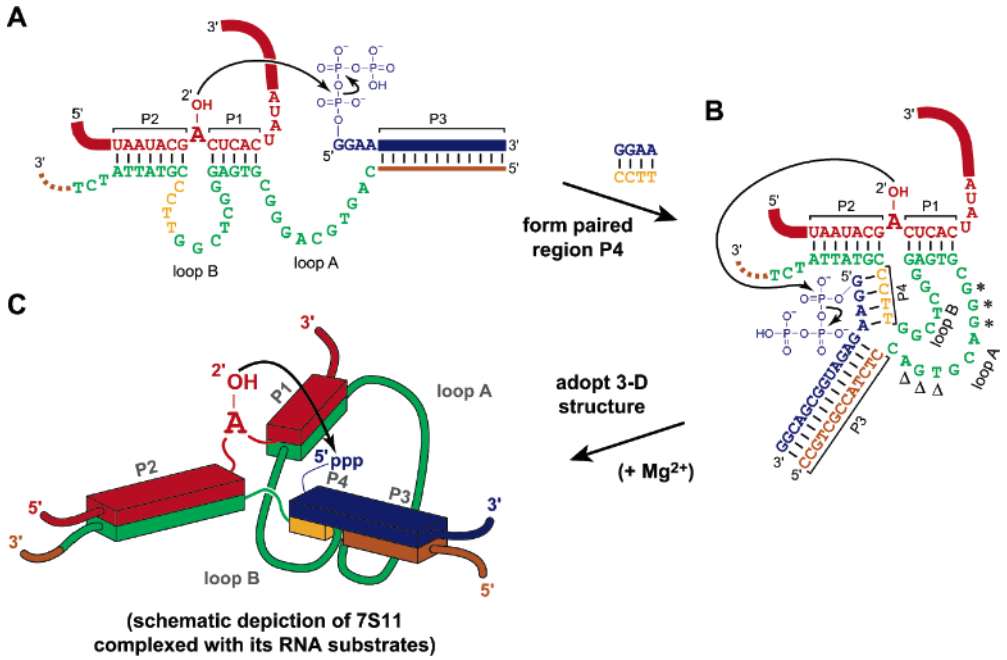


Figure 2. Sequence and structure of the 7S11 deoxyribozyme. (A) Secondary structure of 7S11 complexed with its RNA substrates. The bulged branch-site adenosine 2′-hydroxyl attacks the 5′-triphosphorylated guanosine, forming a 2′,5′-branch. Note paired regions P1, P2, and P3, as well as enzyme loops A and B. (B) Three-helix-junction (3HJ) structure of 7S11 complexed with its RNA substrates, obtained from the structure in panel A by formation of paired region P4. The symbols next to individual nucleotides of loop A are Δ = may be deleted; ∗ = cannot be mutated or deleted (see Table 1). (C) Schematic three-dimensional representation of the 3HJ structure of 7S11. The helical regions P1−P4 are depicted in a highly simplified form as rectangular slabs so that the spatial relationships among the nucleic acid components may be more clearly visualized. The 3HJ is formed by convergence of paired regions P1, P2, and P4, each of which is experimentally established (see text). The precise spatial disposition of the helices is unknown; coaxial stacking of P4 and P3 is possible but not demanded by the available data. In our initial report, P1 and P2 were termed "block A" and "block B", respectively.[11]

the context of these considerations, qualitatively new approaches to branched RNA synthesis are clearly desirable.

Our laboratory has initiated substantial efforts toward the in vitro selection of deoxyribozymes (DNA enzymes) that catalyze RNA ligation reactions.[8−12] In a recent report, we described the identification of 7S11, a deoxyribozyme that synthesizes branched RNA by mediating the reaction of an internal 2′-hydroxyl on one RNA substrate with a 5′-triphosphate on a second RNA substrate (Figure 1B).[11] The 7S11 deoxyribozyme positions the branch-site RNA 2′-hydroxyl nucleophile on a bulged adenosine nucleotide (Figure 2A) in an arrangement that resembles the structures found in the reaction pathways of both group II introns and the spliceosome. On this basis, we concluded that 7S11 mimics structurally the first step of natural RNA splicing.[11]

In our initial report on 7S11,[11] the base pairs surrounding the branch-site adenosine were covaried in limited Watson−Crick fashion using one particular set of nucleotide changes. Efficient RNA ligation activity was maintained, suggesting that 7S11 may be generally useful for synthesis of 2′,5′-branched RNA. Here, we comprehensively demonstrate that 7S11 is indeed applicable to prepare a wide range of branched RNA sequences. Strict sequence requirements are found only at the two RNA nucleotides that are directly involved in the covalent branch formation. Outside of these two nucleotides, most substrate sequences are tolerated with preparatively useful ligation activity, although 7S11 favors certain sequence combinations. The general synthetic ability of 7S11 is validated by preparing the branched core of ai5γ, a representative group II intron RNA that is often used as a model system for RNA splicing.[2,13]

Our interest in 7S11 also focuses upon the structure and function of the deoxyribozyme itself, the study of which is expected to reveal general principles of nucleic acid catalysis. We designate the two base-paired regions surrounding the branch-site adenosine as P1 and P2 (Figure 2A), by analogy to the paired (P) regions of many natural ribozymes. A third paired region, P3, comprises the interaction between the deoxyribozyme and most nucleotides of the 5′-triphosphate RNA substrate. Upon inspection of the 7S11 secondary structure shown in Figure 2A, we noted that the two indicated four-nucleotide segments (orange and blue) have the potential to form an additional, short base-pairing interaction, designated P4 (Figure 2B). Formation of P4 along with P1 and P2 would draw the nucleic acid strands together into a three-helix junction (3HJ) involving these three base-paired regions, with the branch-site adenosine located at the intersection of the three helices. Many natural ribozymes

(8) 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.
(9) (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. (c) Prior, T. K.; Semlow, D. R.; Flynn-Charlebois, A.; Rashid, I.; Silverman, S. K. *Nucleic Acids Res.* **2004**, *32*, 1075−1082. (d) Wang, Y.; Silverman, S. K. *J. Am. Chem. Soc.* **2003**, *125*, 6880−6881. (e) Coppins, R. L.; Silverman, S. K. *J. Am. Chem. Soc.* **2004**, *126*, 16426−16432. (f) Wang, Y.; Silverman, S. K. *Biochemistry* **2005**, *44*, in press. (g) Semlow, D. R.; Silverman, S. K. *J. Mol. Evol.* **2005**, in press.
(10) Wang, Y.; Silverman, S. K. *Biochemistry* **2003**, *42*, 15252−15263.
(11) Coppins, R. L.; Silverman, S. K. *Nat. Struct. Mol. Biol.* **2004**, *11*, 270−274.
(12) Silverman, S. K. *Org. Biomol. Chem.* **2004**, *2*, 2701−2706.
(13) (a) Bonitz, S. G.; Coruzzi, G.; Thalenfeld, B. E.; Tzagoloff, A.; Macino, G. *J. Biol. Chem.* **1980**, *255*, 11927−11941. (b) Jacquier, A.; Michel, F. *Cell* **1987**, *50*, 17−29. (c) Jarrell, K. A.; Peebles, C. L.; Dietrich, R. C.; Romiti, S. L.; Perlman, P. S. *J. Biol. Chem.* **1988**, *263*, 3432−3439.

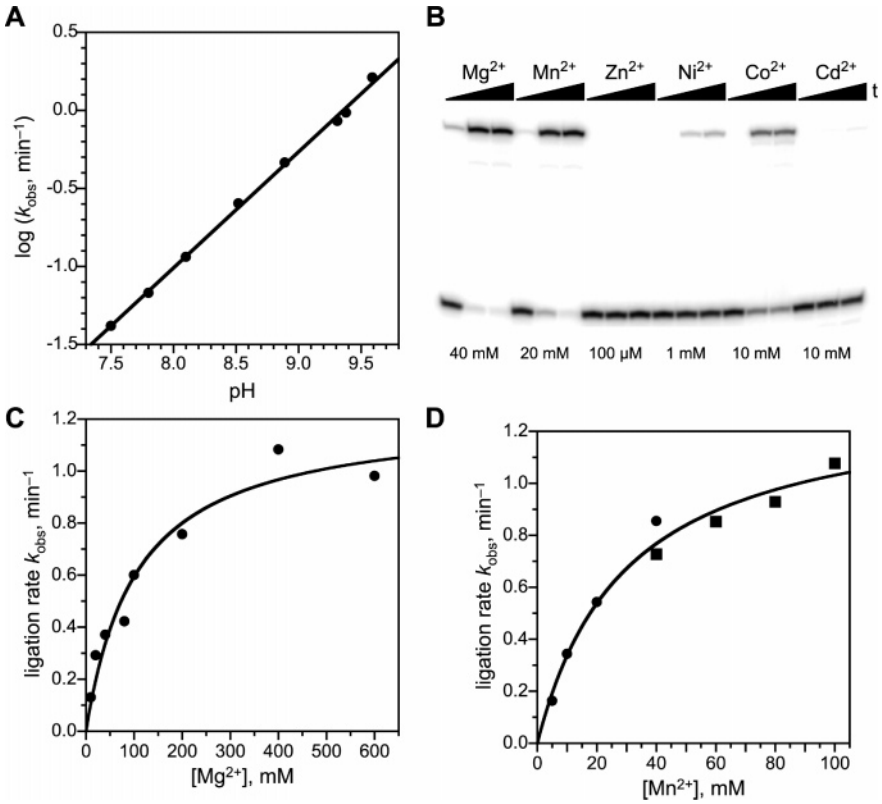


Figure 3. Optimizing incubation conditions for branched RNA formation by 7S11. (A) pH dependence determined at 40 mM Mg^{2+} and 37 °C. The slope of the line is 0.74; the fit value of k_{obs} at pH 9.0 is 0.54 min^{-1}. (B) Metal ion dependence. The assays were performed at 37 °C and pH 9.0 (Mg^{2+}) or 7.5 (other metals). Time points are 20 s, 5 min, and 20 min for Mg^{2+} and Mn^{2+}; 0, 12, and 24 h for the other ions. (C) Determining $K_{d,app}$ for Mg^{2+} at 37 °C. The exact pH of the buffer for this particular experiment was 8.91. (D) Determining $K_{d,app}$ for Mn^{2+} at pH 7.5, 37 °C. The different symbols represent data sets collected in two separate experiments.

adopt structures in which 2−4 helices converge,[14] and the possibility that 7S11 has a similar structure would reveal a useful model system for multi-helix junctions in catalytic nucleic acids. Here, we present compelling evidence that 7S11 indeed adopts the 3HJ structure shown schematically in Figure 2C.

Results

Optimizing the Incubation Conditions for 7S11-Mediated Branched RNA Formation. Before we explored the sequence tolerance and substrate interactions of 7S11, its pH and metal ion dependence were examined. These experiments ensured that optimal incubation conditions were used in the later assays. The pH dependence of 7S11 was assayed at 40 mM Mg^{2+} and 37 °C (Figure 3A). A plot of log(k_{obs}) versus pH was linear between pH 7.5 and 9.6 with a slope of 0.74, suggesting a net loss of one proton during the ligation reaction. From the plot, the k_{obs} with 40 mM Mg^{2+} at pH 9.0 and 37 °C is 0.54 min^{-1}. All remaining assays in this study were performed at pH 9.0, which facilitated the rapid assessment of ligation activities.

In addition to Mg^{2+}, the other divalent metal ions Mn^{2+}, Zn^{2+}, Ni^{2+}, Co^{2+}, and Cd^{2+} were tested (Figure 3B). Substantial ligation activity was observed with 20 mM Mn^{2+} ($k_{obs} \approx 0.5$ min^{-1} at pH 7.5 and 37 °C) and to a lesser degree with Co^{2+} (~50% maximal yield). Both Ni^{2+} and Cd^{2+} supported only limited activity (5−10% and 1−2% yield, respectively, in 24 h), whereas Zn^{2+} led to ≤0.1% yield in 24 h. The $K_{d,app}$ for Mg^{2+} was 106 ± 27 mM (Figure 3C), and the $K_{d,app}$ for Mn^{2+} was 30 ± 6 mM (Figure 3D). In subsequent assays, we routinely incubated samples at 40 mM Mg^{2+} and 37 °C, which permits a substantial k_{obs} while suppressing nonspecific RNA degradation at long incubation times.

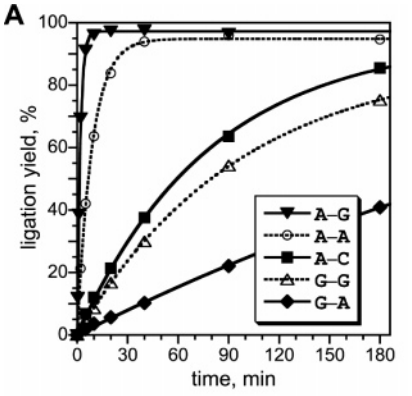
The Two Nucleotides Directly at the Ligation Junction. The two nucleotides directly involved in the ligation reaction are the adenosine (A) that provides the 2′-hydroxyl nucleophile and the 5′-triphosphorylated guanosine (pppG) that bears the reactive branch-site phosphorus atom. For preparing branched RNAs of varying sequence, it is important to know the extent to which these two nucleotides may be changed. We previously noted that the nucleotide providing the 2′-hydroxyl group may be G instead of A with a 30-fold reduced k_{obs} while retaining high ligation yield, whereas neither C nor U supports any ligation activity at this branch-site position.[11] As expected, a 2′-deoxy-A at the branch site abolishes all activity.[11]

Here, we comprehensively examined all combinations of the two nucleotides involved directly in branch formation (Figure 4). A 5′-triphosphate-G nucleotide is typically incorporated into RNA by in vitro transcription using T7 RNA polymerase,[15] and 5′-triphosphate-A may also be obtained in this fashion.[16] However, the 5′-triphosphate-pyrimidines cannot be incorporated in this manner. We therefore used 5′-adenylated RNA (5′-

(14) (a) Lilley, D. M.; Clegg, R. M.; Diekmann, S.; Seeman, N. C.; von Kitzing, E.; Hagerman, P. *Eur. J. Biochem.* **1995**, *230*, 1−2. (b) Lilley, D. M. J. *Q. Rev. Biophys.* **2000**, *33*, 109−159.

(15) (a) Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C. *Nucleic Acids Res.* **1987**, *15*, 8783−8798. (b) Milligan, J. F.; Uhlenbeck, O. C. *Methods Enzymol.* **1989**, *180*, 51−62.

(16) (a) Coleman, T. M.; Wang, G.; Huang, F. *Nucleic Acids Res.* **2004**, *32*, e14. (b) Huang, F.; Wang, G.; Coleman, T.; Li, N. *RNA* **2003**, *9*, 1562−1570.



B

2′-OH nucleotide \ 5′-adenylated nucleotide	G	A	C	U
A	97%	95%	86%	5%
G	75%	41%	<1%	nr
C	5%	<1%	nr	nr
U	3%	<1%	nr	nr

Figure 4. Varying the two nucleotides that directly form the 2′,5′-branched linkage. (A) The successful substrate combinations (those with >5% ligation yield) at 40 mM Mg^{2+}, pH 9.0, and 37 °C. Values of k_{obs} (min^{-1}, top to bottom): 0.56, 0.11, 0.013, 0.010, and 0.0021. In each entry of the legend, the nucleotides that provide the 2′-hydroxyl nucleophile and 5′-adenylated electrophile are shown in that order. In all assays, the DNA nucleotide opposite the 5′-adenylated RNA nucleotide was Watson−Crick complementary to the RNA, thereby maintaining P4 with four base pairs (see Figure 2B). More comprehensive experiments in which the 5′-terminal RNA nucleotide and its DNA counterpart were systematically covaried indicate that a Watson−Crick base pair is optimal for these nucleotides (see Supporting Information). (B) Summary of ligation yields at 3 h incubation for the various nucleotide combinations, with the parent A−G combination shaded gray; nr = no reaction detected (<0.1%). For yields shown as <1%, discrete but very faint product bands were observed. This and all subsequent experiments used the standard incubation conditions of 50 mM CHES, pH 9.0, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 37 °C.

AppRNA) as a functional surrogate for 5′-triphosphate RNA (5′-pppRNA),[17] having previously demonstrated that 5′-AppG instead of 5′-pppG does not disrupt 7S11 function.[11] The new experiments demonstrate that the 2′-hydroxyl of the branch-site A may be used as the nucleophile with any of 5′-adenylated G, A, or C as the electrophile with high rate and yield, although the rates for 5′-AppA and especially 5′-AppC are diminished relative to those of 5′-AppG. A branch-site G permits modest rate and yield only with 5′-AppG or A. Finally, a branch-site C or U reacts very poorly with any 5′-adenylated RNA substrate, regardless of the identity of the adenylated nucleotide.

Substrate Sequence Tolerance of 7S11 and Evidence for a 3HJ Structure. The proposed 7S11 secondary structure shown in Figure 2 has four RNA:DNA duplex regions denoted P1, P2, P3, and P4. Duplex regions P3 and P4 additionally have the potential for coaxial stacking, as is observed in other nucleic acid structures (e.g., ref 18). To provide firm experimental evidence for paired regions P1−P4 and to demonstrate the generality of 7S11 for synthesizing branched RNA, we performed comprehensive nucleotide covariation experiments (Figure 5). In these assays, appropriate regions of the RNA substrate sequences were changed relative to the parent sequences either by systematic nucleotide transitions (A ↔ G and U ↔ C) or by transversions (A ↔ C and G ↔ U, here abbreviated "transv-1", or A ↔ U and G ↔ C, abbreviated "transv-2"). In all cases, the corresponding DNA nucleotides were also changed to retain Watson−Crick pairing. When pairing was not maintained in any one of the four regions, all ligation activity was lost (ref 11 for P1−P3 and Supporting Information for P4). In these assays, the nucleotides directly at the branch site were maintained as A (providing the 2′-hydroxyl nucleophile) and pppG, which were the original and optimal combination (Figure 4). Overall, the data in Figure 5 demonstrate that each of the four double-stranded regions of P1−P4 may be varied widely in sequence while retaining substantial branch-forming activity, as long as RNA:DNA base-pairing is maintained.

A more detailed examination of the covariation data reveals quantitative differences among the ligation efficiencies for certain RNA substrate sequences. For the P1 and P2 regions (Figure 5A), systematic transversions of both regions simultaneously (either transv-1 or transv-2) led to a modest 4-fold or 12-fold decrease in k_{obs}; the latter still leaves 7S11 with a useful k_{obs} on the order of 2 h^{-1}. In contrast, transitions in both regions led to a larger 110-fold drop in k_{obs}; this effect was separable into a 21-fold decrease for transitions in P1 only and a 2.7-fold drop for P2 only (Figure 5B). For P3 and P4, systematic changes in either region are generally tolerated well (Figure 5C,D). Only for one of three sets of P4 substitutions was a decrease in k_{obs} observed (a 14-fold drop for transv-2; Figure 5D). Mutations of individual nucleotides within P4 showed that no specific P4 residues are responsible for the bulk of this effect (Figure 6).

The P4 duplex can be considered the key element of the 3HJ structure because its formation draws together the three helices P1, P2, and P4 (Figure 2). Therefore, the existence of P4 along with P1−P3 is particularly compelling evidence for the three-helix-junction structure adopted by 7S11 with its RNA substrates. Even if P3 is not stacked coaxially with P4, as depicted schematically in Figure 2C, paired regions P1, P2, and P4 necessarily form a three-helix junction for the following reasons. (i) P1 and P2 are covalently linked through the branch-site adenosine; (ii) P2 and P4 are covalently linked by the DNA backbone; and (iii) the branch-site adenosine between P1 and P2 must be near the 5′-triphosphate of P4 during the ligation reaction itself.

To test a "worst-case scenario" for branch formation with various RNA substrate sequence combinations, we combined the most poorly performing sequences, which were transitions for P1/P2 and transv-2 for P3/P4. Not surprisingly, the resulting ligation activity was low (~4.2% yield in 3 h, equivalent to $k_{obs} \approx 0.00023$ min^{-1}; data not shown). The observed drop in activity was quantitatively predictable from the contributions of each substrate component. Specifically, the predicted decrease in k_{obs} is $110 \times 0.9 \times 14 = 1400$-fold relative to the original parent sequences, assuming additivity of interactions. The observed decrease is 2300-fold (i.e., less than a factor of 2 difference between predicted and observed). Although these data reveal that 7S11 is sluggish with particularly poor combinations of RNA substrate sequences, it should be emphasized that a

(17) Silverman, S. K. *RNA* **2004**, *10*, 731−746.
(18) Cate, J. H.; Gooding, A. R.; Podell, E.; Zhou, K.; Golden, B. L.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A. *Science* **1996**, *273*, 1678−1685.

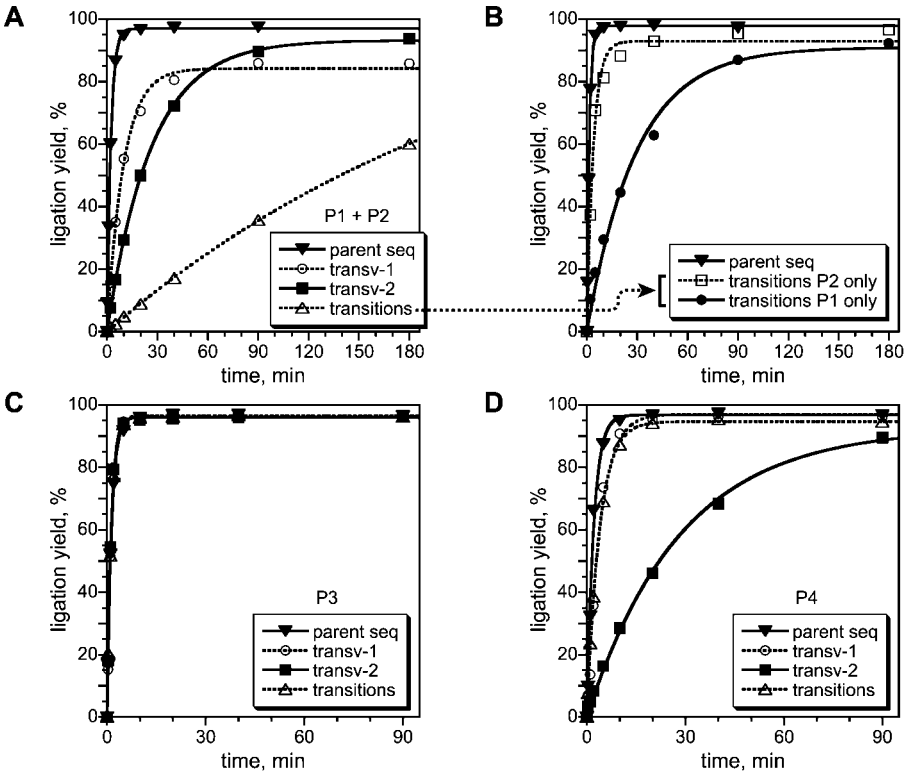


Figure 5. Nucleotide covariation experiments that demonstrate the structure and generality of 7S11 for branched RNA synthesis. (A) Covariations within both P1 and P2. Values of k_{rel} (parent = 1.0) are transitions, 0.0091; transv-1, 0.23; and transv-2, 0.085. (B) Covariations within P1 and P2 separately. Values of k_{rel} are P1, 0.048 and P2, 0.37. (C) Covariation within P3. Values of k_{rel} are transitions, 1.06; transv-1, 1.05; and transv-2, 1.11. (D) Covariation within P4. Values of k_{rel} are transitions, 0.54; transv-1, 0.51; and transv-2, 0.073. For mismatched combinations of RNA and DNA, which have no activity, see ref 11 for P1−P3 and Supporting Information for P4.

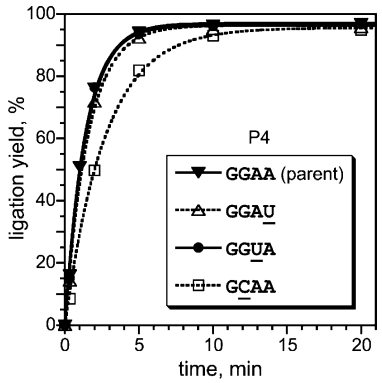


Figure 6. Assaying the effects of individual mutations within P4. Each of the three indicated RNA nucleotides was separately changed by mutation (transv-2; underlined nucleotides) along with the complementary mutation in the DNA, and the effect on k_{obs} was determined. Values of k_{rel} (top to bottom): 1.0, 0.87, 0.96, and 0.50.

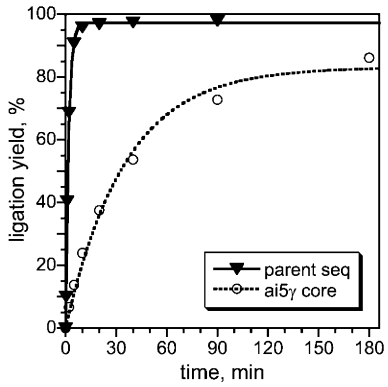


Figure 7. Successful synthesis of the 46-nt branched core of the ai5γ group II intron by the 7S11 deoxyribozyme. Values of k_{obs}: parent sequences, 0.55 min^{-1}; ai5γ core sequences, 0.028 min^{-1} (k_{rel} = 0.051). Despite the lower rate, the ai5γ branch can be obtained in high yield using 7S11.

wide variety of sequences are readily joined with rates and yields that are useful for practical purposes (Figure 5 and also Figure 7 below).

Preparation of the Branched Core of the ai5γ Group II Intron. As a direct demonstration of 7S11's practical utility, we prepared the branched RNA core of the ai5γ group II intron.[2,13] This representative group II intron RNA[19] is often used as a model system in structure−function studies, and its sequence is unrelated to the substrates used during the selection procedure that led to 7S11. Despite many sequence differences, the ai5γ branched core is successfully synthesized in high yield and reasonable rate by 7S11 (Figure 7). In contrast, our earlier-reported 9F7 deoxyribozyme that creates branched RNA is unable to synthesize the ai5γ core due to 9F7's stringent sequence requirements near the ligation site.[10]

Length Requirements for the RNA Substrates. The structures depicted in Figure 2 suggest that the RNA substrate strands may be extended indefinitely at three places: the 3′-tail of P1, the 5′-end of P2, and the 3′-end of P3. Successful strand extension would mean that 7S11 is broadly useful for synthesizing branched RNAs such as the full-length ai5γ group II intron, which is several hundred nucleotides long.[2,13] Previous data showed that lengthening the 3′-tail of P1 by up to 18 nt led

(19) (a) Saldanha, R.; Mohr, G.; Belfort, M.; Lambowitz, A. M. *FASEB J.* **1993**, *7*, 15−24. (b) Michel, F.; Ferat, J. L. *Annu. Rev. Biochem.* **1995**, *64*, 435−461. (c) Jacquier, A. *Biochimie* **1996**, *78*, 474−487.

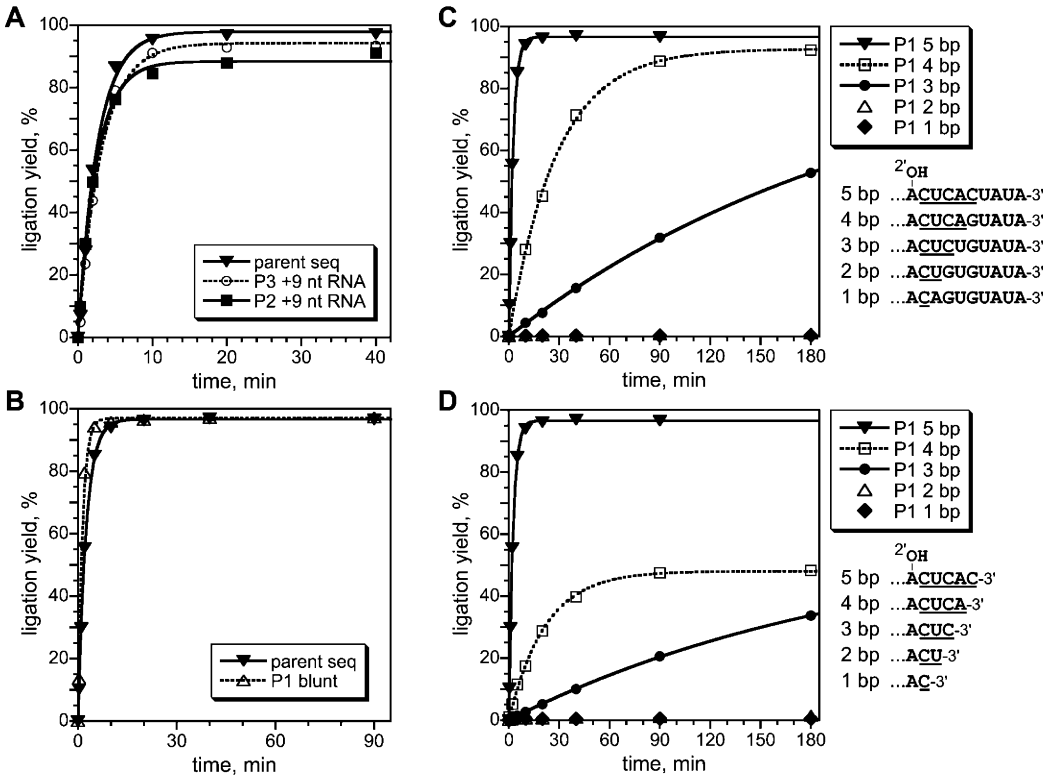


Figure 8. Length requirements for the RNA substrates. (A) Lengthening the RNA strands that extend from P2 or P3 by 9 nucleotides (see Experimental Section for sequences). (B) Entirely truncating the 3′-tail that extends from P1, leaving a blunt P1 duplex. For this truncated substrate, $k_{rel} = 1.83$, compared with that of the parent sequence. (C) Shortening the P1 duplex by mutating the L substrate while retaining its length. Partial sequences of the L substrates that provide the 2′-hydroxyl group nucleophile are shown adjacent to the graph; the RNA portion of the P1 duplex is underlined. Values for $k_{rel} = 0.087$ and 0.011 for the 4-bp and 3-bp substrates compared with those of the parent sequence, which has a 5-bp P1. (D) Shortening the P1 duplex by truncating the L substrate. Values for $k_{rel} = 0.11$ and 0.011 for the 4-bp and 3-bp substrates.

only to a modest (1.6-fold) drop in k_{obs} with no decrease in ligation yield.[11] Similarly, we found here that the respective ends of P2 and P3 may be extended by 9 nt without decreasing the effectiveness of branch formation (Figure 8A). Separately, we shortened the four-nucleotide 3′-tail of P1, which in Figure 2 is depicted as not interacting with any other nucleotides. Indeed, this tail can be truncated entirely, leaving the P1 duplex with no overhanging RNA nucleotides, without lowering the ligation rate or yield (Figure 8B).

Finally, the extent to which the 5-bp P1 duplex may be shortened was examined. This was done in two ways. First, P1 was shortened by mutating its RNA sequence to disrupt one or more RNA:DNA Watson−Crick base pairs while maintaining the length of the entire RNA substrate (Figure 8C). When P1 was 4 or 3 bp, the ligation rate (k_{rel}) relative to a 5-bp P1 was only ∼0.09 or 0.01, respectively. When P1 was 2 bp or shorter, no ligation was observed. Alternatively, P1 was shortened simply by truncating the RNA substrate sequence (Figure 8D). When P1 was 4 or 3 bp via this approach, the ligation rate (k_{rel}) relative to a 5-bp P1 was only ∼0.11 or 0.01, respectively, and the yield also decreased. Again, when P1 was 2 bp or shorter, no ligation was observed. These data confirm that the P1 duplex is an important (albeit short) contributor to the multi-helix-junction structure of 7S11. The data also define the limitations on using 7S11 to synthesize a branch with a short tail emerging from the 3′-position of the branch-site nucleotide. Specifically, the tail must be at least 3 nt long to observe any ligation activity, and 5 nt or longer for maximal activity.

Modification of 7S11 Enzyme Loops A and B. We briefly examined the two 7S11 enzyme region loops A and B of Figure 2. Individually, DNA nucleotides of the loops were either deleted or mutated (via A ↔ G or T ↔ C transitions), and the corresponding deoxyribozymes were assayed for ligation rate and yield. In loop A, three adjacent nucleotides could be deleted individually without effect; these are marked with a Δ in Figure 2B (deletion of all three nucleotides at once did lead to loss of activity). Conversely, individually deleting or mutating three other loop A nucleotides almost completely abolished the ligation activity; these are marked with a ∗ in Figure 2B. The remaining loop A nucleotides and all loop B nucleotides showed intermediate tolerance to deletion or mutation (Table 1).

Discussion

Application of the 7S11 Deoxyribozyme to Synthesize Branched RNA. The new data establish that 7S11 is broadly useful for the in vitro synthesis of 2′,5′-branched RNA using practical incubation conditions (40 mM Mg^{2+} and 37 °C; Figure 3). Although most of the assays in this report were conducted at pH 9.0 for expediency, we routinely use 7S11 at pH 7.5 for preparative purposes, with incubation times up to 24 h without substantial RNA degradation.

In the application of 7S11 for branched RNA synthesis, only a modest number of sequence restrictions apply to the RNA substrates. Directly at the branch site, A is preferred as the donor of the 2′-hydroxyl nucleophile and G is acceptable with lower k_{obs} and high yield, but C and U do not support branch formation

Table 1. Modification of 7S11 Enzyme Loops A and B by Deletion and Mutation[a]

	Deletions		Mutations	
	k_{obs} (min^{-1})	yield (%)	k_{obs} (min^{-1})	yield (%)
	Loop A Nucleotides[b]			
C1	0.0019	35	0.015	90
A2	0.42	97	nd	nd
G3	0.42	97	nd	nd
T4	0.38	97	nd	nd
A2 + G3 + T4	0.0011	97	nd	nd
G5	0.024	92	0.044	95
C6	0.0030	54	0.094	96
A7	0.0022	39	0.41	97
G8	0.00006	10	0.011	19
G9	0.00006	10	*c*	<1
G10	0.00006	10	0.00022	4
C11	0.0016	29	0.20	97
	Loop B Nucleotides[b]			
G12	0.00083	15	0.0032	58
G13	0.00083	15	0.0087	74
C14	0.012	82	0.55	98
T15	0.00078	14	0.020	90
C16	*c*	1	0.0031	55
G17	0.0088	63	0.19	96
G18	0.0088	63	0.041	96

[a] Tabulated are the k_{obs} and yield for each deoxyribozyme as modified by deletion or mutation. In all cases, the mutations were by transitions (A ↔ G and T ↔ C); nd = not determined. Ligation assays were performed with 50 mM CHES, pH 9.0, with 150 mM NaCl, 2 mM KCl, and 40 mM $MgCl_2$ at 37 °C with the parent RNA substrate sequences. Under these conditions, k_{obs} for the unmodified 7S11 deoxyribozyme is 0.54 min^{-1} with >95% yield (Figure 3). In all cases, the yield shown is that observed at the last recorded time point of $t = 180$ min. For yields <60%, the k_{obs} value was estimated from a linear fit to the last recorded time point because little, if any, curvature was observed in the data set. [b] Nucleotides of loop A are numbered 1−11, starting from C1 (which is closest to the P3 duplex as shown in Figure 2B) and ending with C11 (closest to P1). Nucleotides of loop B are numbered 12−18, starting from G12 (closest to P1) and ending with G18 (closest to P4). [c] The observed rate constant could not be determined with confidence because the intensity of the product band was too low.

(Figure 4). As the 5′-triphosphorylated or 5′-adenylated RNA partner, A, G, and to a lesser extent C are tolerated, whereas U is not. The combination of 2′-hydroxyl-A and a 5′-AppC is the only situation for which a pyrimidine is tolerated at either nucleotide position with substantial ligation activity. In contrast to these observations, all other nucleotides within the RNA substrates have much greater sequence tolerance, although a range of rates and yields is observed (Figures 5 and 6). This generality is highlighted by the preparation of the branched core of the ai5γ group II intron RNA (Figure 7), the overall sequence of which is unrelated to that used during the selection that led to 7S11. Because most group II introns and spliceosomal mRNA substrates have A and G nucleotides forming the 2′,5′-branched linkage,[20] 7S11 should be particularly useful for preparation of branched RNAs that are relevant to RNA splicing. In other studies, we are pursuing the identification of deoxyribozymes that tolerate other branch-site nucleotides for practical RNA ligation (E.D. Pratico, Y. Wang, and S.K.S., manuscript in preparation).

Application of 7S11 to "real" structured RNAs (e.g., those with hundreds of nucleotides) requires two additional considerations. First, are very long RNA substrate strands tolerated by 7S11, and second, does RNA secondary structure near the branch site interfere with ligation? Regarding the first question, our data directly demonstrate that ligation activity is maintained when each of the substrate strands is lengthened (Figure 8 and ref 11). In addition, we have successfully applied 7S11 to several natural RNAs that are >150 nt in length (Y. Wang, C.V. Miduturu, and S.K.S., unpublished results). Addressing the second question, the latter experiments used disruptor DNA oligonucleotides to sequester inhibitory base pairing within the RNA, thereby permitting the deoxyribozyme to bind the substrates. More details on the use of disruptors for branch formation with very long RNA substrates will be reported elsewhere.

The 7S11 deoxyribozyme does not show multiple turnover behavior (data not shown), similar to our findings for other RNA ligase deoxyribozymes.[8] We hypothesize that this is due to product inhibition as is observed for other deoxyribozymes and protein enzymes that ligate RNA,[21] for which a stoichiometric amount of the enzyme must be used. For branched RNA synthesis using 7S11, this is not generally problematic because the RNA substrates are typically more expensive than the DNA in terms of both time and money. Moreover, the DNA enzyme can often be recovered during purification of the branched RNA product and subsequently reused, as is routinely done in our laboratory.

Advantage of a Deoxyribozyme-Based Synthesis Approach. The 7S11 DNA enzyme mediates the reaction of a single RNA 2′-hydroxyl functional group in a highly selective and precise structural context, while many other similar 2′-hydroxyl functional groups do not react (Figure 2). This high selectivity without detectable side reaction (i.e., branch formation at other sites) is typical behavior for an enzyme. In contrast, a preparative approach to branched RNA based on traditional organic synthesis with the same selectivity characteristics would require substantial use of protecting groups, as in solid-phase approaches.[4−7] Because of the simplicity and the wide sequence tolerance of 7S11, the deoxyribozyme-based synthetic approach is a significant advance toward the preparation of the topologically interesting and biochemically relevant branched RNAs. We consider that 7S11 exemplifies the advantage of a (deoxy)ribozyme-based approach to appropriate bioorganic reactions such as branched RNA synthesis, for which enzymatic selectivity is highly advantageous.[12]

The 7S11 Three-Helix Junction as a Model System for Nucleic Acid Structure. The unambiguous nucleotide covariation data for P1, P2, and especially P4 (Figure 5) firmly establish the three-helix-junction structure for 7S11. Because many natural ribozymes adopt multi-helix junctions, 7S11 can be viewed as a model system for such ribozymes. We anticipate that more detailed structural biology approaches to investigate 7S11 (e.g., experiments based on native gel electrophoresis, fluorescence resonance energy transfer,[22] X-ray crystallography, and NMR spectroscopy) will prove useful in understanding how the three-dimensional structure of this deoxyribozyme permits highly selective branched RNA formation.

(20) (a) Michel, F.; Umesono, K.; Ozeki, H. *Gene* **1989**, *82*, 5−30. (b) Burge, C. B.; Tuschl, T.; Sharp, P. A. In *The RNA World*, 2nd ed.; Gesteland, R. F., Cech, T. R., Atkins, J. F., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1999; pp 525−560.

(21) 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.

(22) Lafontaine, D. A.; Norman, D. G.; Lilley, D. M. *EMBO J.* **2002**, *21*, 2461−2471.

The experiments in which the two enzyme loops A and B of 7S11 were mutated provide insight into the activity of the deoxyribozyme. The observation that certain portions of these loops have strict nucleotide requirements (as marked with ∗ in Figure 2B) suggests that these loops are not merely passive connectors, whereas the observation that certain nucleotides can be deleted (Δ in Figure 2B) indicates that the loops do have some flexibility. The metal ion dependence of 7S11 (Figure 3) suggests that specific contacts are made with metal ions for either structural or mechanistic reasons. Further analyses along these lines, in conjunction with direct structural methods, should be informative regarding the structure and mechanism of 7S11 branch-forming activity.

Relationship to Biological RNA Splicing. In light of the new data demonstrating a 3HJ structure for 7S11, the relationship of this deoxyribozyme to biological RNA splicing can be addressed. We originally concluded that 7S11 structurally mimics the first step of RNA splicing[11] because the bulged nucleotide that provides the reactive 2′-hydroxyl group is straddled by two Watson−Crick duplex regions (denoted here P1 and P2), as found in group II introns and the spliceosome. The involvement of a third helical region (P4 with possibly coaxially stacked P3) adds an interesting structural feature to this picture. The discovery that the relatively simple 7S11 deoxyribozyme enables branch formation by drawing the reactive functional groups to the center of a 3HJ suggests that a similar catalytic strategy may be employed by natural ribozymes. Further structural characterization of 7S11 and other natural and artificial nucleic acid enzymes will certainly provide further insights into their mechanisms.

Experimental Section

Materials. DNA oligonucleotides were prepared at IDT (Coralville, IA). 5′-Triphosphorylated RNA oligonucleotides were prepared by in vitro transcription using T7 RNA polymerase with an appropriate double-stranded DNA template[15] that was prepared by annealing two DNA oligonucleotides. 5′-Adenylated RNA substrates were prepared using T4 RNA ligase and a DNA template oligonucleotide as described,[17] or by a newer approach that uses T4 DNA ligase and a DNA template (Y. Wang and S.K.S., manuscript in preparation). All DNA and RNA oligonucleotides were purified by PAGE as described previously.[8,10] For convenience, the substrate providing the 2′-hydroxyl is denoted the left-hand (L) substrate, and the substrate providing the 5′-activated group is the right-hand (R) substrate. The parent sequence of the L RNA substrate was 5′-GGAUAAUACGACUCACUAUA-3′, where the underlined adenosine is the branch point. The parent sequence of the R RNA substrate was 5′-GGAAGAGAUGGCGACGG-3′. For the experiment of Figure 8A in which P2 was extended, the L sequence was 5′-GGA*CACCUAAGA*UAAUACGACUCACUAUA-3′ (A = branch point). For the experiment in which P3 was extended, the R sequence was 5′-GGAGAGAUGGCGACGG*UGGGUGCGA*-3′. For both substrates, the italicized nucleotides were added relative to the parent sequence. For the experiment of Figure 7, the RNA sequences used to prepare the ai5γ core were L 5′-GGAGUAAGGUCUACCUAUCGGGAU-3′ and R 5′-GAGCGGUCUGAAAGUUAUCAU-3′. For the experiment of Figure 8B in which the 3′-tail from P1 was truncated, the L sequence was 5′-GGAUAAUACGACUCAC-3′, which is 4 nt shorter at its 3′-end than the parent sequence.

Kinetic Assays. As described in our initial report on 7S11,[11] all of the kinetic assays used a trimolecular format in which the two RNA substrates were incubated together with the deoxyribozyme. The dephosphorylated and subsequently 5′-^{32}P-radiolabeled L substrate was the limiting reagent relative to the deoxyribozyme E and the R substrate. The ratio L:E:R was 1:5:15, with the concentration of E equal to ∼0.5 μM. Values of k_{obs} and final yield 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. See our earlier report for a detailed description of the methods of sample preparation and ligation analysis.[8]

The standard incubation conditions for the kinetic assays were 50 mM CHES buffer, pH 9.0, containing 150 mM NaCl, 2 mM KCl, and 40 mM $MgCl_2$ at 37 °C. When [NaCl] was varied between 0 and 300 mM or [KCl] was varied between 0 and 100 mM, essentially no effect on ligation rate or yield was observed (data not shown). For the experiment of Figure 3A, the buffers were HEPES (pH 7.5), EPPS (pH 7.8−8.5), and CHES (pH 8.9−9.6). For the experiment of Figure 3B, each metal ion other than Mg^{2+} was tested at 10 μM, 100 μM, 1 mM, and 10 mM in 50 mM Tris, pH 7.5; shown are the experiments with highest ligation yield (and least degradation in the case of Zn^{2+}). When Mn^{2+} was used in place of Mg^{2+} (Figure 3D), higher concentrations of Mn^{2+} (>100 mM) led to substantial nonspecific degradation.

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 No. 5-FY02-271), the National Institutes of Health (GM-65966), the American Chemical Society Petroleum Research Fund (38803-G4), 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 are grateful to Imran Rashid, Teresa Fraterman, Sarah Carter, and Elena Zelin for assistance with various experiments.

Supporting Information Available: Details of experiments not fully described in the text (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA044881B