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

‡*Department of Biochemistry, University of Illinois*

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 into a 5′-triphosphate or 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 ~50-fold lower k_{obs}, and neither uridine nor cytidine are 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

[1]Abbreviations:

PAGE
  polyacrylamide gel electrophoresis

CHES
  2-(*N*-cyclohexyl)amino-ethanesulfonic acid

EDTA
  ethylenediaminetetraacetic acid

k_{obs}
  observed rate constant

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.

EXPERIMENTAL PROCEDURES

General Considerations

DNA oligonucleotides were prepared at IDT (Coralville, IA) or at the UIUC 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 $1\times$ TBE [89 mM each Tris and boric acid, 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 arrangment 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 underlined are the four DNA nucleotides that interact with the 5′-end of the R substrate to form P4. 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 versus round number) is shown in Figure 2. See the Results 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 1-nt 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 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 for

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 ratio L:E:R 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 = $Y \cdot (1 - e^{-kt})$, where $k = k_{obs}$ and Y = 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 timepoints (10-180 min as appropriate).

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 an RNA substrate for which uridine 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 the 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 the 7-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 $4^{22} = 2 \times 10^{13}$ possible 22-mer sequences, all of 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 were 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%; grey 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 Supporting Information). Unexpectedly, a clear branch-site purine preference was evident; the trend of ligation rates and yields was A>G>U≈C (Figure 4), 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 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 an 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 as originally used with branch-site U (maximal pool activity ~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 instead of 2-h incubation, the selection activity was substantially lower (grey bars in Figure 2B), suggesting that maximal G-branching activity had been achieved. Cloning from round 7 after 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 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 <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 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% versus ~4% yield in 3 h). It was possible that the 4-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-nucleotide length as 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 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% versus <1%; Y.M.W. and S.K.S., unpublished results). On the basis of these observations, we recommend that others use 10DM24 instead of 7S11 for 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 commentary accompanying that manuscript (15), our recent identification of a deoxyribozyme unrelated to 7S11 that uses any branch-site nucleotide (14) was cited as supporting 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 on several loop B mutants, which retain branch-site A selectivity despite the possibility of such a base pair (see 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 $10^{(0.3-0.4)} = 2.0\text{-}2.5$-fold 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.
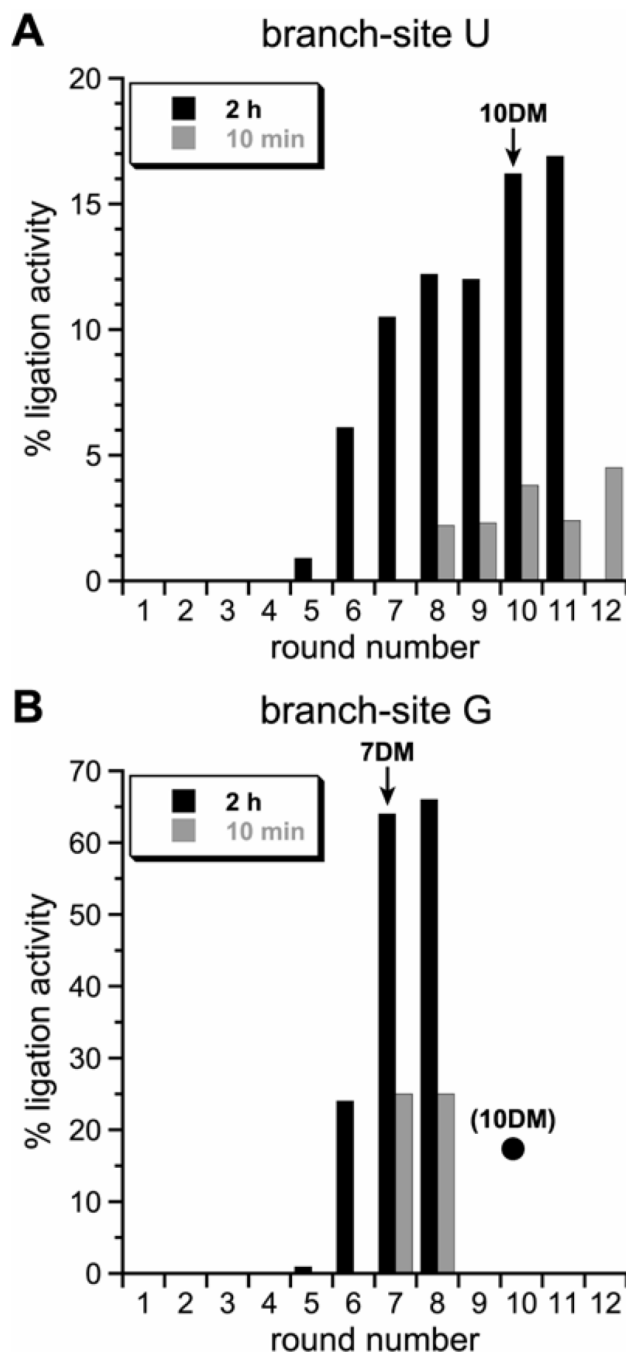
Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

REFERENCES

1. Joyce GF. Directed Evolution of Nucleic Acid Enzymes. Annu. Rev. Biochem 2004;73:791–836. [PubMed: 15189159]

2. Silverman SK. Deoxyribozymes: DNA catalysts for bioorganic chemistry. Org. Biomol. Chem 2004;2:2701–2706. [PubMed: 15455136]

3. Coppins RL, Silverman SK. A DNA Enzyme that Mimics the First Step of RNA Splicing. Nat. Struct. Mol. Biol 2004;11:270–274. [PubMed: 14758353]

4. Coppins RL, Silverman SK. A Deoxyribozyme That Forms a Three-Helix-Junction Complex with Its RNA Substrates and Has General RNA Branch-Forming Activity. J. Am. Chem. Soc 2005;127:2900–2907. [PubMed: 15740125]

5. Coppins RL, Silverman SK. 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 2005;44:13439–13446. [PubMed: 16216067]

6. Michel F, Ferat JL. Structure and activities of group II introns. Annu. Rev. Biochem 1995;64:435–461. [PubMed: 7574489]

7. Jacquier A. Group II introns: elaborate ribozymes. Biochimie 1996;78:474–487. [PubMed: 8915537]

8. Burge, CB.; Tuschl, T.; Sharp, PA. The RNA World. 2nd ed.. Gesteland, RF.; Cech, TR.; Atkins, JF., editors. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY: 1999. p. 525-560.

9. Chu VT, Liu Q, Podar M, Perlman PS, Pyle AM. More than one way to splice an RNA: branching without a bulge and splicing without branching in group II introns. RNA 1998;4:1186–1202. [PubMed: 9769094]

10. Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic Acids Res 1987;15:8783–8798. [PubMed: 3684574]

11. Flynn-Charlebois A, Wang Y, Prior TK, Rashid I, Hoadley KA, Coppins RL, Wolf AC, Silverman SK. Deoxyribozymes with 2′-5′ RNA Ligase Activity. J. Am. Chem. Soc 2003;125:2444–2454. [PubMed: 12603132]

12. Wang Y, Silverman SK. Characterization of Deoxyribozymes That Synthesize Branched RNA. Biochemistry 2003;42:15252–15263. [PubMed: 14690435]

13. Nielsen H, Westhof E, Johansen S. An mRNA is capped by a 2′,5′ lariat catalyzed by a group I-like ribozyme. Science 2005;309:1584–1587. [PubMed: 16141078]

14. Pratico ED, Wang Y, Silverman SK. A Deoxyribozyme That Synthesizes 2′,5′-Branched RNA with Any Branch-Site Nucleotide. Nucleic Acids Res 2005;33:3503–3512. [PubMed: 15967808]

15. Pyle AM. Capping by branching: a new ribozyme makes tiny lariats. Science 2005;309:1530–1531. [PubMed: 16141065]

16. Wang Y, Silverman SK. Deoxyribozymes That Synthesize Branched and Lariat RNA. J. Am. Chem. Soc 2003;125:6880–6881. [PubMed: 12783536]

17. Wang Y, Silverman SK. Efficient One-Step Synthesis of Biologically Related Lariat RNAs by a Deoxyribozyme. Angew. Chem. Int. Ed 2005;44:5863–5866.

18. Acharya S, Foldesi A, Chattopadhyaya J. The pK_a of the Internucleotidic 2′-Hydroxyl Group in Diribonucleoside (3′→5′) Monophosphates. J. Org. Chem 2003;68:1906–1910. [PubMed: 12608809]

19. Liu Q, Green JB, Khodadadi A, Haeberli P, Beigelman L, Pyle AM. 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 1997;267:163–171. [PubMed: 9096215]

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.

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, then either 2 h or 10 min as indicated for round 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.

```
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 (on the left) is four nucleotides shorter than for the other deoxyribozymes (see Figure 1).
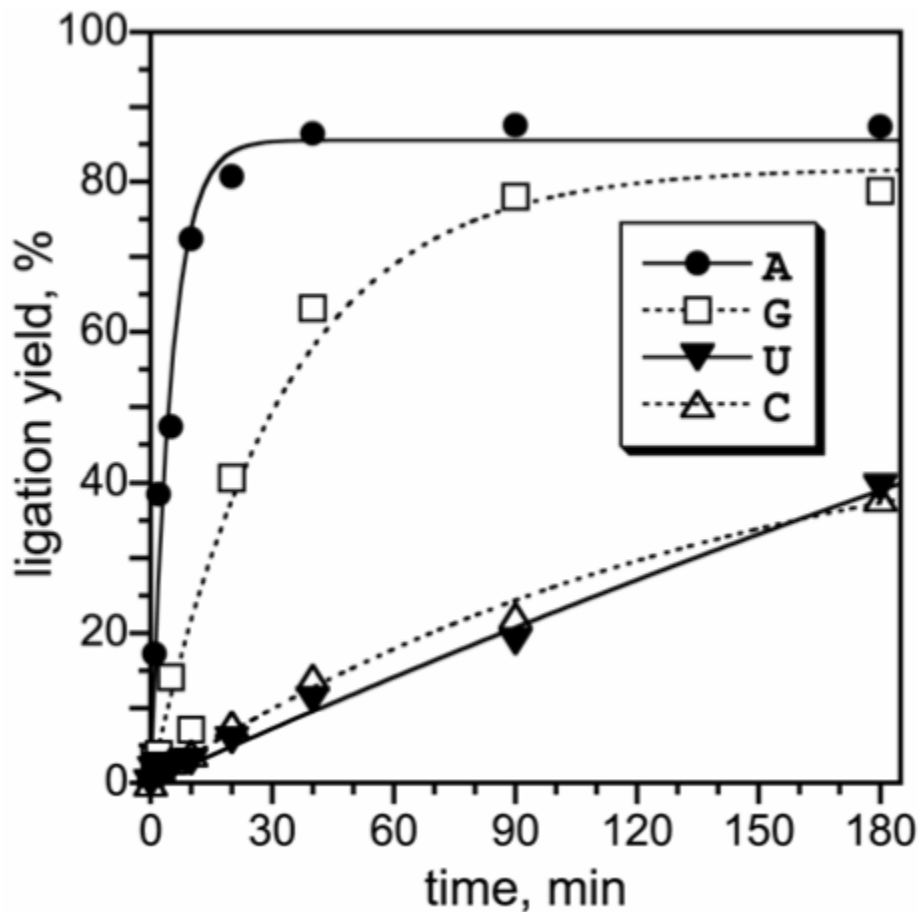
FIGURE 4.
Branched RNA formation by the 10DM24 deoxyribozyme shows a clear preference for branch-site adenosine. k_{obs} values (min^{-1} ± sem; $n = 3$ in each case): A 0.26 ± 0.02; G 0.035 ± 0.002; U 0.0035 ± 0.0011; C 0.0030 ± 0.0007 ($k_A/k_U = 73$). For comparison, 7S11 synthesizes A-branched RNA with $k_{obs} \approx 0.5$ min^{-1} and $k_A/k_U \approx 3000$ (3,4). See Table 1 for similar data on all new deoxyribozymes

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 analogously to those in Figure 4. See Supporting Information for full data sets. The three columns towards the right show the adenosine selectivity as quantified by 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 (4).