

Improved deoxyribozymes for synthesis of covalently branched DNA and RNA

Christine S. Lee, Timothy P. Mui and Scott K. Silverman*

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

Received July 8, 2010; Revised August 8, 2010; Accepted August 9, 2010

ABSTRACT

A covalently branched nucleic acid can be synthesized by joining the 2'-hydroxyl of the branch-site ribonucleotide of a DNA or RNA strand to the activated 5'-phosphorus of a separate DNA or RNA strand. We have previously used deoxyribozymes to synthesize several types of branched nucleic acids for experiments in biotechnology and biochemistry. Here, we report *in vitro* selection experiments to identify improved deoxyribozymes for synthesis of branched DNA and RNA. Each of the new deoxyribozymes requires Mn^{2+} as a cofactor, rather than Mg^{2+} as used by our previous branch-forming deoxyribozymes, and each has an initially random region of 40 rather than 22 or fewer combined nucleotides. The deoxyribozymes all function by forming a three-helix-junction (3HJ) complex with their two oligonucleotide substrates. For synthesis of branched DNA, the best new deoxyribozyme, 8LV13, has k_{obs} on the order of 0.1 min^{-1} , which is about two orders of magnitude faster than our previously identified 15HA9 deoxyribozyme. 8LV13 also functions at closer-to-neutral pH than does 15HA9 (pH 7.5 versus 9.0) and has useful tolerance for many DNA substrate sequences. For synthesis of branched RNA, two new deoxyribozymes, 8LX1 and 8LX6, were identified with broad sequence tolerances and substantial activity at pH 7.5, versus pH 9.0 for many of our previous deoxyribozymes that form branched RNA. These experiments provide new, and in key aspects improved, practical catalysts for preparation of synthetic branched DNA and RNA.

INTRODUCTION

Covalently branched nucleic acids play important roles in biochemistry and biotechnology. A covalently branched

nucleic acid molecule has a branch-site ribonucleotide with three emerging strands (Figure 1A). The two strands attached to the 5'- and 3'-positions of the branch-site nucleotide constitute one segment of the branch, and the third strand attached to the 2'-position is the second segment. This conceptual division corresponds to one practical synthesis strategy, which involves attack by the 2'-hydroxyl of the branch-site ribonucleotide of one substrate molecule (the first segment) into a suitably activated 5'-phosphorus, which is part of a separate substrate molecule (the second segment). The electrophilic 5'-phosphorus of the second substrate is readily activated either as a 5'-adenylate (for DNA) or a 5'-triphosphate (for RNA) (1,2). The products are designated as covalently branched, which distinguishes them from the noncovalently 'branched' complexes based on double-helical DNA that are frequently applied in DNA nanotechnology (3), found in natural DNA recombination (4), and observed in functional RNAs such as many ribozymes (5).

In accord with the overall synthesis strategy shown in Figure 1A, we recognize four distinct kinds of covalently branched nucleic acid, where the first and second segments may independently be either DNA or RNA. Of these four possibilities, two are the focus of the present work (Figure 1B). First, branched DNA—in which both of the two segments are DNA, and only the branch-site ribonucleotide has a 2'-oxygen atom (6)—has been used for signal amplification in quantification of nucleic acids (7–10) and for templating of multiple chemical reactions (11). Multiply branched DNA has also been named 'comb-branched DNA'. Second, branched RNA—in which both of the segments are entirely RNA—is the natural intermediate of RNA splicing (12,13) and therefore of considerable biochemical interest. Improving the ability to synthesize both branched DNA and branched RNA in an efficient and sequence-general fashion will expand the range of accessible experiments that require such molecules.

Deoxyribozymes are well-defined DNA sequences with catalytic activity, also called DNA enzymes or DNAzymes

*To whom correspondence should be addressed. Tel: +1 217 244 4489; Fax: +1 217 244 8024; Email: scott@scs.illinois.edu

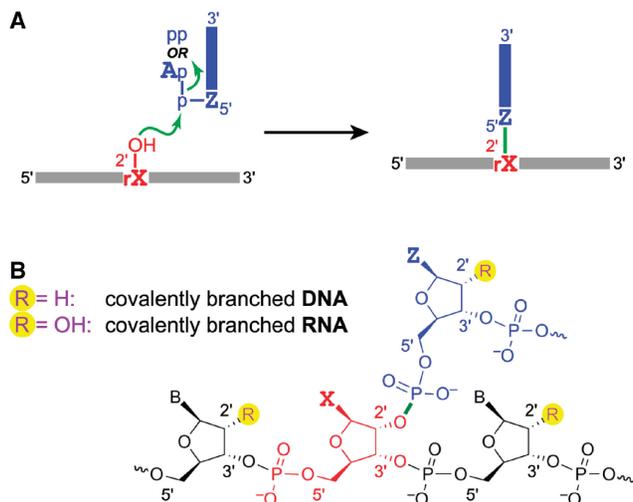


Figure 1. Covalently branched nucleic acids. **(A)** General structure and synthesis strategy for covalently branched nucleic acids. Two substrates are ligated: the left-hand and right-hand substrates, or L and R. The branch-site ribonucleotide, rX, of the L substrate is colored red. The R substrate (with 5'-adenylated or 5'-triphosphorylated nucleotide Z) is colored blue. B = nucleobase. **(B)** Chemical structure of covalently branched DNA (R = H) and RNA (R = OH). The goal of this study is to identify improved deoxyribozymes for synthesis of these branched nucleic acids.

(14–18). We have previously used *in vitro* selection (19) to identify a number of deoxyribozymes that catalyze formation of branched nucleic acids including both branched DNA (6) and branched RNA (20–26), as well as RNA–DNA chimeras (27) that allow imposition of double-stranded DNA constraints on RNA conformation (27,28). We have also reported several experiments in which artificial branched RNA synthesized by deoxyribozymes was used to explore specific aspects of biochemical systems (26,29,30). In most cases, our prior efforts have not identified deoxyribozymes with optimal characteristics. For example, certain branched DNA sequences were made in only poor yields or required very long incubation times (72–96 h) (6). Here, we report systematic experiments that identify improved deoxyribozymes for synthesis of covalently branched DNA and RNA.

MATERIALS AND METHODS

Oligonucleotides

All oligonucleotides were purified by denaturing PAGE with running buffer 1× TBE (89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3) as described previously (21). DNA oligonucleotides, including those incorporating a single branch-site ribonucleotide, were prepared by solid-phase synthesis at IDT (Coralville, IA, USA). The 5'-phosphate for each right-hand (R) substrate was either included during solid-phase synthesis or introduced after synthesis using ATP and T4 polynucleotide kinase. RNA oligonucleotides were prepared by *in vitro* transcription using a DNA template and T7 RNA polymerase (31) or, in specific cases for left-hand

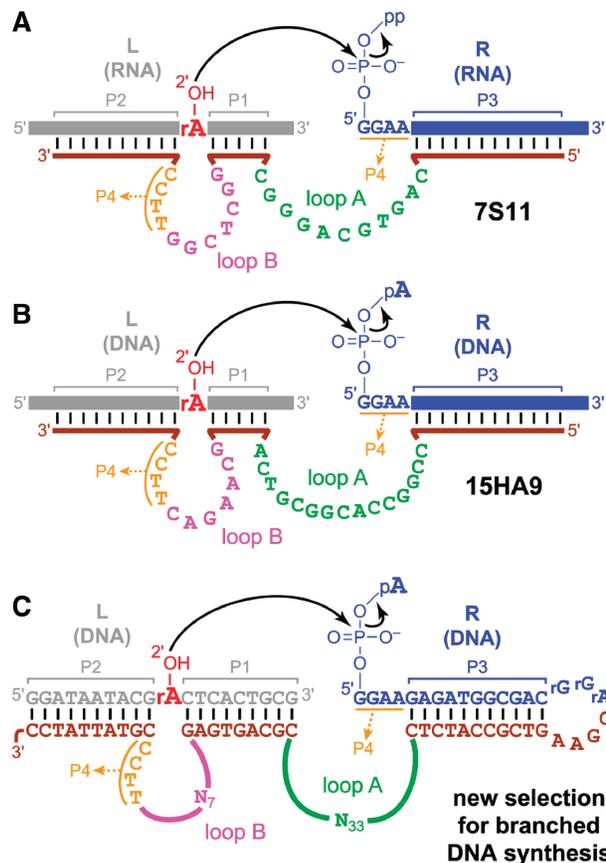


Figure 2. Three-helix-junction (3HJ) complexes between deoxyribozymes and their left-hand and right-hand (L and R) substrates for synthesis of covalently branched nucleic acids. Loops A and B are initially random at the outset of *in vitro* selection but have well-defined sequences for each particular deoxyribozyme. Paired regions P1 through P4 have Watson–Crick base pairs between deoxyribozyme and substrate. Note that formation of P4 draws together the nucleophile (branch-site nucleotide) and electrophile (5'-adenylate or 5'-triphosphate) such that they are juxtaposed at the intersection of a 3HJ composed of P1, P2, and P4. **(A)** The 7S11 deoxyribozyme that forms branched RNA (23,33). **(B)** The 15HA9 deoxyribozyme that forms branched DNA (6). **(C)** The new *in vitro* selection experiment for branched DNA synthesis, with expanded N₃₃ loop A and P1 region (9 bp) relative to 15HA9. The new selection for branched RNA synthesis was similar, except the L substrates were DNA rather than RNA, and the R substrate was 5'-triphosphorylated. The covalent linkage between deoxyribozyme and R substrate as required during selection is shown explicitly (the 3'-terminal rGrGrA allows ligation by T4 RNA ligase), whereas for all of the deoxyribozyme assays performed after selection, this linkage was absent. Not depicted fully is a 3'-extension (...AGTGTATCCTGATGG-3') on the deoxyribozyme, to enable PCR primer binding after each selection step.

(L) RNA substrates that were found to be transcribed poorly, by solid-phase synthesis at IDT. All R DNA substrates were activated by 5'-adenylation using our previously described procedure (1), with additional details as described (6).

The sequences of the parent L and R DNA substrates used during *in vitro* selection are shown in Figure 2C. The R substrate used in assays after selection was terminated with ...dGdG-3' rather than ...rGrGrA-3'. Systematic sequence variations for all substrates were made relative to the parent (Par) sequences as follows: transitions

(T_{sn} = A↔G, C↔T), transversions-1 (Tv1 = A↔C, G↔T) and transversions-2 (Tv2 = A↔T, G↔C). The parent RNA substrate sequences were analogous to the DNA sequences, as were the systematic variations.

In vitro selection, cloning and initial screening

The selection experiments were performed essentially as described previously (32). Figure 2C shows the deoxyribozyme pool sequence. The two polymerase chain reaction (PCR) primers used in each round were 5'-pCG AAGTCGCCATCTC-3' (coding direction) and 5'-(AAC)₄XC CATCAGGATCAGCTGGATAATACGGGAA-3' (noncoding direction), where X denotes a nonamplifiable PEG spacer, IDT Spacer 18, through which Taq polymerase cannot read. In each selection experiment, the initial deoxyribozyme pool was 200 pmol (~10¹⁴ molecules). In the key selection step as shown in Figure 2C, the DNA pool, which was attached to the R substrate using T4 RNA ligase, was heated to 95°C for 3 min in 5 mM Tris, pH 7.5, 15 mM NaCl and 0.1 mM EDTA, and then placed on ice for 5 min. The sample was adjusted to 50 mM HEPES, pH 7.5, 20 mM MnCl₂, 150 mM NaCl and 2 mM KCl, incubated for 2 h at 37°C, and separated by 8% denaturing PAGE. After an appropriate number of selection rounds, individual deoxyribozymes were cloned and surveyed for branch formation activity.

Cloning and initial screening of individual deoxyribozyme clones was performed using DNA strands prepared by PCR from miniprep DNA derived from individual *E. coli* colonies. The miniprep DNA samples were first checked by digestion with EcoRI to confirm the presence of the expected insert. The concentration of each PAGE-purified deoxyribozyme strand was estimated from the UV shadowing intensity relative to suitable standards. Each screening assay was performed using the procedure described below, with timepoints at 0, 1 and 2 h.

Assay procedure for branch formation

The 5'-³²P-radiolabeled L substrate was the limiting reagent relative to the deoxyribozyme (E) and the R substrate. A 6 μl sample containing 0.3 pmol of L, 3 pmol of E and 15 pmol of R was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl and 0.1 mM EDTA by heating at 95°C for 3 min and cooling on ice for 5 min. The sample was adjusted to final concentrations of 50 mM HEPES, pH 7.5, 150 mM NaCl and 2 mM KCl by addition of 2 μl of an appropriate stock solution. The ligation reaction was initiated by addition of 2 μl of 100 mM MnCl₂ (final total volume of 10 μl; final Mn²⁺ concentration of 20 mM), and the sample was incubated at 37°C. Aliquots of 1–2 μl were removed at desired timepoints, quenched onto 8 μl of stop solution (80% formamide, 1× TBE, 50 mM EDTA, 0.025% bromophenol blue, 0.025% xylene cyanol), and analyzed by 20% PAGE. The yield versus time data were fit directly to first-order kinetics; i.e. $yield = Y \cdot (1 - e^{-kt})$, where $k = k_{obs}$ and $Y =$ final yield. For reactions performed on the 0.5 nmol scale, the final volume was 200 μl containing 500 pmol of L, 600 pmol of E and 700 pmol of R.

RESULTS

One of our previously identified deoxyribozymes, 7S11, creates branched RNA by holding its two RNA substrates in a three-helix-junction (3HJ) complex (Figure 2A) (22,23,33). We found that this 3HJ architecture can be implemented for new *in vitro* selection experiments and maintained by the emergent deoxyribozymes, both for branched RNA formation (26,27) and for other chemical reactions such as nucleopeptide linkage formation (34,35), and we have shown that a 3HJ can be engineered to use GTP as a small-molecule substrate (36). Here, we employed the 3HJ architecture for all of our new *in vitro* selection experiments, which were performed to identify new deoxyribozymes for synthesis of both branched DNA and branched RNA.

New deoxyribozymes for synthesis of branched DNA

We recently reported the *in vitro* selection of a deoxyribozyme, 15HA9, which synthesizes branched DNA by ligating the internal 2'-OH group of one DNA substrate (the 'left-hand' substrate, denoted L and containing a single embedded ribonucleotide) to the 5'-phosphorus of a second, 5'-adenylated DNA substrate (the 'right-hand' substrate, designated R) (Figure 2B) (6). 15HA9 was identified in the 3HJ architecture with the following key features: the branch-site ribonucleotide was rA; the provided divalent metal ion was Mg²⁺ (40 mM); and initially random loops A and B were 15 and 7 nt long, respectively. These combined 22 random nucleotides allowed complete sampling of sequence space in a single selection experiment (4²² = 2 × 10¹³ possible sequences; selection was initiated with 200 pmol ≈ 10¹⁴ molecules). The best resulting deoxyribozyme, 15HA9, forms branched DNA with k_{obs} on the order of 0.1 h⁻¹ (requiring 72–96 h incubation for complete reaction) at pH 9.0, 120 mM Mg²⁺ and 37°C. Despite these suboptimal characteristics of long incubation time, elevated pH and high Mg²⁺ concentration, 15HA9 could be used to synthesize multiply branched DNA that has up to four 'addition' strands attached to particular branch sites along a common 'foundation strand', all with useful tolerance for various DNA sequences near each branch site (6). Because DNA (unlike RNA) is not particularly susceptible to nonspecific degradation upon incubation with Mg²⁺ at pH 9.0, elevated pH and high Mg²⁺ concentrations are tolerable in practice for branched DNA synthesis (multiple ribonucleotides in the foundation strand do pose a challenge in this regard). However, the incubation time of 3–4 days to achieve good yield is a practical frustration.

Here, we performed new *in vitro* selection experiments to identify improved deoxyribozymes for branched DNA synthesis. A previous selection effort to form chimeric DNA-RNA branches was successful when loop A was expanded from N₁₅ to N₃₃ (34). We therefore considered that deoxyribozyme-catalyzed branched DNA formation might be improved if loop A were similarly lengthened to 33 nt, even though this means sampling only ~10⁻¹⁰ (i.e., 10¹⁴ out of 4⁽³³⁺⁷⁾ ≈ 10²⁴) of sequence space. In addition, we noted that many of our previous

deoxyribozyme selection efforts have been successful with Mn^{2+} rather than Mg^{2+} as the divalent metal ion cofactor (37,38), as have selection efforts from several other labs for both deoxyribozymes (39–44) and ribozymes (45). Natural Mn^{2+} -dependent ribozymes have also been identified (46–48). Therefore, we considered that improved catalysis might be possible if Mn^{2+} were present during selection (15HA9 itself is inactive with Mn^{2+}). On the basis of these considerations, a new selection experiment was performed, using 20 mM Mn^{2+} at pH 7.5 (higher pH cannot be used due to Mn^{2+} oxidation) along with expansion of loop A to N_{33} ; loop B was maintained as N_7 . The P1 region of the L substrate was correspondingly lengthened from 5 to 9 bp to encourage stronger deoxyribozyme–substrate binding (as shown in Figure 2, the four paired regions involving each deoxyribozyme and its two substrates are denoted P1 through P4). The branch-site ribonucleotide in L was maintained as rA. The particular sequences used in this new selection experiment are shown in Figure 2C.

In each selection round, during the key step of branched DNA formation the sample was incubated in 50 mM HEPES, pH 7.5, 20 mM $MnCl_2$, 150 mM NaCl and 2 mM KCl at 37°C for 2 h. After merely three selection rounds, ligation activity was detectable (0.8% ligation yield for the pool), and after seven rounds the activity

rose to 57%. Selection rounds were restarted beginning at round 5 with only a 10 min incubation, leading to only 5% ligation yield for the repeated round 5, and after eight total rounds the ligation yield for the pool had leveled off at 26%. Individual deoxyribozymes were cloned from this round 8 pool and assayed. From 18 active clones, 12 unique sequences were identified (Table 1).

The nine most active clones were evaluated for their substrate sequence generality, using deoxyribozymes prepared by solid-phase synthesis. DNA-catalyzed formation of multiply branched DNA would be most enabled by deoxyribozymes that tolerate essentially any P2 sequence within the ribonucleotide-containing L substrate as long as the deoxyribozyme–substrate Watson–Crick base pairing is maintained. Therefore, we analyzed individual deoxyribozymes by evaluating branched DNA synthesis using substrates for which the P2 sequence element in L (immediately to the 5' side of the branch-site rA; Figure 2C) was altered from the parent sequence by systematic A↔C and G↔T changes; i.e. ‘transversions-1’ (Tv1) as denoted in our prior report (6). Branched DNA synthesis was evaluated when either the entire P2 region was changed; one nucleotide adjacent to the branch site was retained with the parent identity; or two nucleotides were retained. The results revealed that none of the nine

Table 1. Deoxyribozyme sequences (all unique sequences from both selection experiments)

Deoxyribozyme	Enzyme region sequence, 5' to 3' ^a
Deoxyribozymes for branched DNA synthesis	
8LV1	GCCGGTGAAGGTTCTCTGCCAAAAGTGGTTCGTCGCCAGTGTAG ACTTTAG
8LV2	ATGACTGTAGGAGCGGACGCTAAGCCTTTAGACTCGCAGTGTAG GTTCCG
8LV3	CCATGGACGTATGGTTTGGTATGCCGACCGCGACCGCAGGTGAGGCAGGGC
8LV7	CAGGGTGTAGAGGACCGTAAGTGTAGTTGACACACCGCAGTGTAG CTTCGAT
8LV9	GTCGGGCGTGTGGTTGCAAAGGACGTCCTGGAGTGTGCCAGTGTAG ACCGCCA
8LV11	CCTAGGGAATGCGGAGGCTTGAGGTGTGTCTTGGCGCAGGAG CGTAACC
8LV12	CAGTAGACGCCGCATAGGGCACTGTTTACTGCTCCGAAGTGTAG AGGAGGG
8LV13	CGAGAGCGTGAATGTTGGTTGAGACACAGTGTGA CGCAGTGTAG AGGGAAC
8LV15	CAGGGAGAGAGCGGTACGCGGAGATATACGCGGTGCGCAGTGTAG ATTCGAT
8LV17	CACGAGTAAGTGTCTCGATTGGGTTACTACTGAGGCGCGCAGTGTAG CAAGGAG
8LV21	CAGGGTGTAGAGGACGAAAGCATGGGGTTAGCTGT CGCAGTGTAG ATTCGAT
8LV22	CAGGGTGAAGTGAAGCTACATTGTATGATTGCGCGCAGTGTAG TTTCCGAT
Deoxyribozymes for branched RNA synthesis	
8LX1	GACGCCACCGAAGTTCGCCATCTCCCGTAGGTGAA GGGCGTGTAGGGTTC
8LX3	CCCACCGTAGGGCTACGGGCAAGGTCAACATGCCCGCAAGTGTAGGGGTTCGA
8LX5	CAGGCCATGGATTATTGGAGGAATGGGCGGTCA CGCGGTGTAGGTGTGGA
8LX6	CAGTGCAGGGCGTGTAGGGCTCCATCCCGAGTGCA GGGCGTGTAGGGCTCGG
8LX8	CCGACGAGGTGTAGGGGATGCGGTACACGCGCA TATAGTGTAGGGGTCCA
8LX12	CCCCGAGGTGTGGACATAGCGGGCTGGTGTGGCG CGCAGTGTAGCCTAGTG
8LX13	CATCGGTGTAGCGATGACCGGGCAAAGATACATT CGCAGTGTAGGGTGTGCG
8LX14	CCACGTGCGAGGTGTAGACGTACGTGGCTGGTGTGTT CGCAGTGTAGCCTATGT
8LX15	CCGACGCGTCCAGGAGGCAAGGGCTATATGCAC CGCAGTGTAGGGCTCGG
8LX16	CACGAGGTAGGGAGGGGACACTGCTATGCTCAG CGCAGTGTAGAGTATGC
8LX17	CACATGTCTAGCGCGTGTGGGAGAAGTGTCAAG CGCTGTGTAGGGTGTG
8LX18	CCAGGACAGCGTAGCCAGACAGCGGGAGGTT CGCAGTGTAGTGAATGG
8LV21	GTAGCCACATTAGTGCCTGCAACTGCTATGCAA CGCAGTGTAGAGGGTGC

^aThe originally 33 nt (loop A) and 7 nt (loop B) of the enzyme regions are shown, with the intervening parent-sequence nucleotides underlined (see Figure 2C).

Presumably due to errors by Taq polymerase during selection, many of the deoxyribozymes have insertions or deletions in the enzyme region; several have an insertion, deletion or mutation in the intervening region.

For each deoxyribozyme, the full sequence also includes 5'- and 3'-binding arms that are not depicted. For the parent substrate sequences, the 5'-binding arm is 5'-CCGTCGCCATCTC-3' (note that the 5'-CGAA... of Figure 2C is replaced with 5'-CC... for the intermolecular assays), and the 3'-binding arm is 5'-TTCCCGTATTATCC-3', where the four parent-sequence P4 nucleotides are underlined. For other substrate sequences, the deoxyribozyme binding arms were covaried as appropriate to maintain Watson–Crick base pairs with the substrates.

deoxyribozymes functioned well when sequence changes were made throughout P2, but all nine tolerated sequence changes when two nucleotides were retained, as exemplified by the deoxyribozyme named 8LV13 (Figure 3). The three most active deoxyribozymes, 8LV1, 8LV9 and 8LV13, also tolerated P2 changes when just 1 nt adjacent to the branch site was retained.

For each of 8LV1, 8LV9 and 8LV13, nucleotides of the P1 region in L were also examined (data not shown). Using an L substrate in which Tv1 changes were made starting at the third nucleotide away from the branch site rA (i.e. retaining two P1 nucleotides as present in the parent sequence), no ligation activity was observed. Therefore, we decided to keep all P1 substrate nucleotides to the 3'-side of the branch site unchanged from the parent sequence, which still allowed orthogonal formation of multiply branched DNA by 15HA9 in our previous study (6).

We then turned our attention to changes in the sequence of the 5'-adenylated DNA substrate, R. When nucleotides beginning at the seventh position of R were changed by

Tv1 (i.e. retaining all four P4 nucleotides as well as the first two P3 nucleotides, for a total of six unchanged nucleotides in R), substantial ligation activity by each of 8LV1, 8LV9 and 8LV13 was maintained (Figure 3). These changes in P3 were the same as those tolerated by 15HA9 while still permitting formation of multiply branched DNA (6).

The above findings were combined to design comprehensive assays of sequence changes in both DNA substrates L and R. These systematic sequence changes were either Tv1 as described above; transitions (Tsn; $A \leftrightarrow G$, $C \leftrightarrow T$); or transversions-2 (Tv2; $A \leftrightarrow T$; $G \leftrightarrow C$). Because 8LV13 had the highest activity of the three 8LV deoxyribozymes when the P3 changes were made, only this deoxyribozyme was assayed in comprehensive fashion. All four systematic combinations of substrate sequences (with changes in both P2 and P3) were ligated well by 8LV13, with k_{obs} between 0.07 and 0.19 min^{-1} and with maximum yield reached in <1 h (Figure 4). Therefore, 8LV13 is a highly general deoxyribozyme for synthesis of branched DNA, with $\geq 90\%$ yield in <1 h for a wide range of substrate sequences.

Finally, the dependence of 8LV13 ligation activity on branch-site nucleotide identity was investigated, using the parent substrate sequences for all other nucleotides.

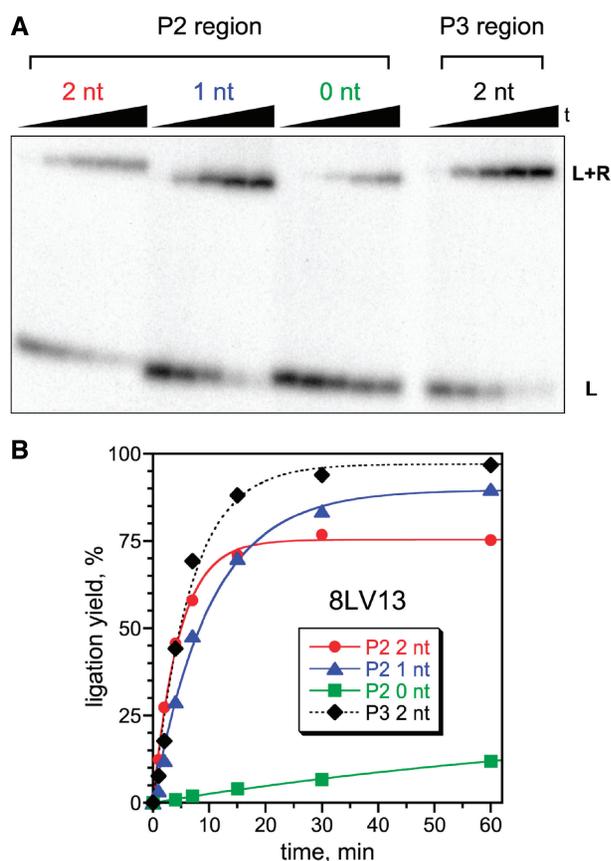


Figure 3. The 8LV13 deoxyribozyme for branched DNA synthesis. (A) PAGE images for assays performed with systematic sequence changes in either the P2 region of the L substrate (left) or the P3 region of the R substrate (right). The specified number of nucleotides adjacent to the branch-site ribonucleotide were unchanged, whereas all other nucleotides of the indicated region were changed by Tv1. Assays were performed in 50 mM HEPES, pH 7.5, 20 mM MnCl_2 , 150 mM NaCl and 2 mM KCl at 37°C. $t = 0, 2, 7, 30$ and 60 min. (B) Kinetic plot for the assays in (A). Data for 8LV1 and 8LV9 were similar (data not shown).

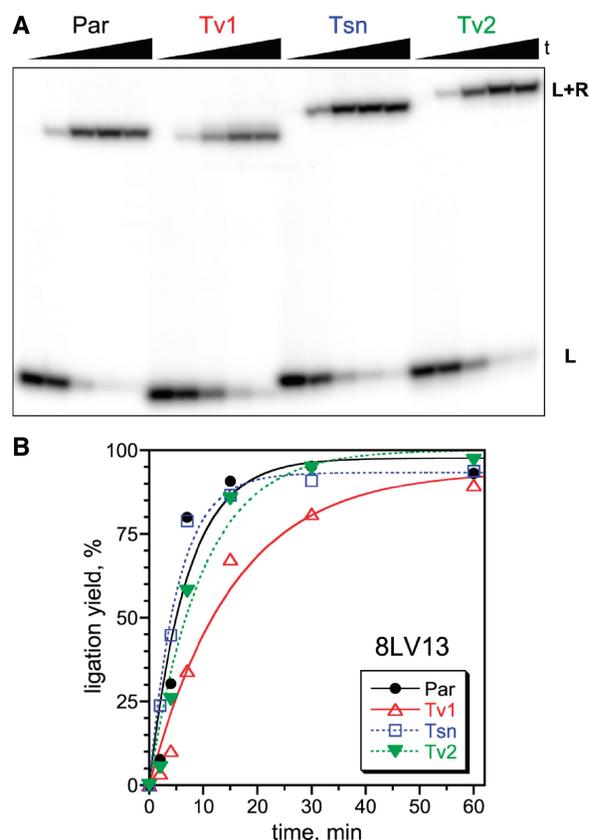


Figure 4. Establishing generality of 8LV13 for branched DNA synthesis by comprehensive sequence changes in both the L and R DNA substrates. Assays were performed as in Figure 3. (A) PAGE image, with timepoints at $t = 0, 2, 7, 30$ and 60 min. (B) Kinetic plots. k_{obs} values (min^{-1}): Par 0.14, Tv1 0.07, Tsn 0.19, Tv2 0.10.

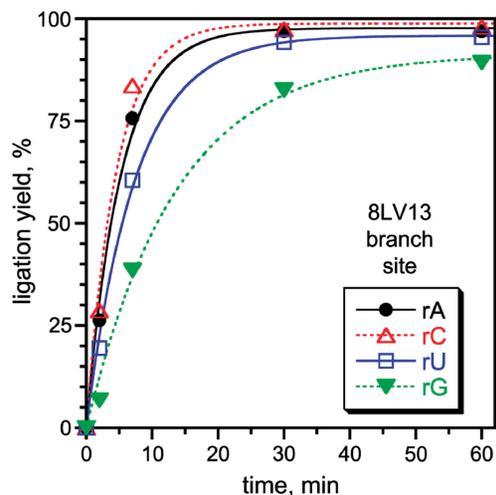


Figure 5. Branch-site nucleotide tolerance of 8LV13. Substrate sequences were as shown in Figure 2C except for variation in the branch-site nucleotide of the L substrate.

When the parent branch-site rA within the L substrate was changed to any of the other three ribonucleotides, substantial ligation activity was observed, with rG only about fourfold slower than the other 3 nt (Figure 5). In contrast, the previously reported 15HA9 was several-fold more active with rA and rU as compared with rG and rC (6).

New deoxyribozymes for synthesis of branched RNA

Our previous efforts toward synthesis of branched RNA led to several deoxyribozymes that create the desired branches and that have useful generality for RNA sequences. In particular, both the 7S11 deoxyribozyme (22,23,33) and an improved variant, 10DM24 (26), tolerate a variety of RNA substrate sequences and create the branched RNA product in good yield, albeit with a substantial range of sequence-dependent rates. Therefore, here we sought additional deoxyribozymes for synthesis of branched RNA. As was done for the new deoxyribozymes that create branched DNA, we used Mn^{2+} as the cofactor, along with N_{33} loop A and P1 expanded from 5 to 9 bp. None of our previous efforts toward branched RNA had used N_{33} as loop A. The substrate sequences were the same as shown in Figure 2C, except with RNA instead of DNA and with 5'-triphosphate instead of 5'-adenylate on R.

The incubation conditions during selection were the same as those used above for branched DNA: 50 mM HEPES, pH 7.5, 20 mM $MnCl_2$, 150 mM NaCl and 2 mM KCl at 37°C for 2 h. Ligation activity was first detectable (0.3% ligation yield for the pool) at round 3, increasing to 40% at round 5 and 48% at round 7. Returning to round 5 and incubating each selection reaction for only 10 min led to 8% ligation yield for the pool, increasing to a plateau of 29% after a total of eight rounds. Cloning led to 18 active sequences, of which 13 were unique (Table 1).

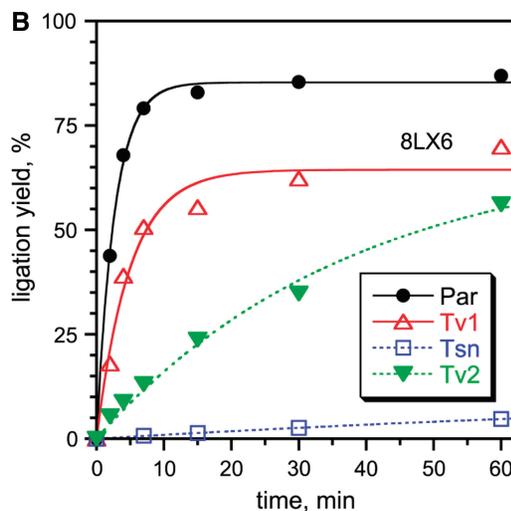
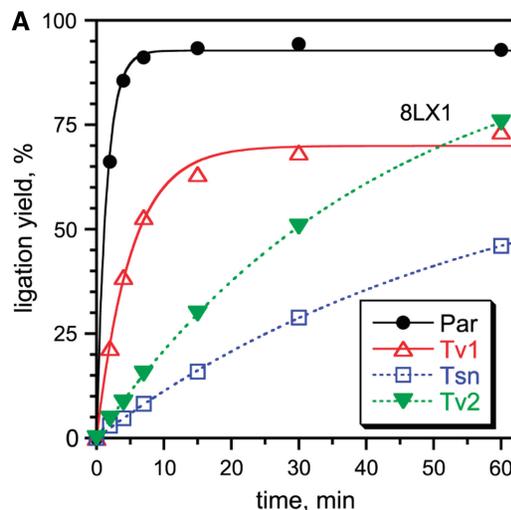


Figure 6. Assays of 8LX deoxyribozymes for branched RNA synthesis. (A) Kinetic plots for assays of 8LX1, with systematic sequence changes in the P1 and P2 regions of the L substrate. (B) Kinetic plots for assays of 8LX6. Data for 8LX8 (not shown) was similar to that shown for 8LX6.

Because many biochemically interesting branched RNAs are derived from naturally occurring sequences, general deoxyribozyme-catalyzed synthesis of branched RNA faces a relatively steep challenge: ideally, all positions of both RNA substrates should tolerate any nucleotide identity. We evaluated the new clones beginning with the L RNA substrate that donates the 2'-hydroxyl group. Three of these deoxyribozymes—named 8LX1, 8LX6 and 8LX8—functioned well when systematic Tv1 changes were made throughout L, in both P1 and P2 except for the branch site itself (Figure 6). Furthermore, 8LX1 maintained substantial activity when Tsn and Tv2 changes were made, whereas 8LX6 and 8LX8 worked appreciably with Tv2 changes but very poorly with Tsn changes. The sequences of all three 8LX deoxyribozymes revealed that mutations accumulated in the P1 region of the deoxyribozyme, such that P1 is reduced from the initially designed 9 bp down to only 5 bp (8LX1 and 8LX6) or 6 bp (8LX8).

8LX1 and 8LX6 were then assayed with systematic Tv1 sequence variants of both RNA substrates, simultaneously changing nucleotides in all paired regions P1 through P4. 8LX1 was found to retain substantial activity when Tv1 changes were made not only in P1 and P2, but also in the R substrate beginning at the ninth position (i.e. retaining all four P4 nucleotides as well as the first four P3 nucleotides (Figure 7A, top). Unfortunately, changing additional nucleotides closer to the R substrate's 5'-terminus suppressed 8LX1 activity. In contrast, 8LX6 was much more tolerant of different R substrate sequences, successfully ligating the two RNA substrates when Tv1 changes were made in P1, P2, all of P3 and all except the first nucleotide (G) of P4 (Figure 7A, bottom).

The broad sequence generality of 8LX6 was further explored by evaluating detailed Tsn changes in the L substrate, revealing that good activity was retained when two parent nucleotides in the substrate were left unchanged on the P2 side; the activity was nearly as high when two parent nucleotides were instead kept on the P1 side (Figure 7B). Then, the sequence generality of 8LX6 was established by comprehensive testing with L and R substrates that were both systematically changed by Tsn, Tv1 or Tv2 changes (Figure 8), retaining the first two P2 nucleotides in the special case of Tsn. With Tv1 changes, the rate was decreased by 36-fold relative to the parent sequences, although high yield (even higher than for the parent sequences) was achieved in 6–12 h. For both Tsn and Tv2 changes, the rate was decreased by 100- to 200-fold relative to the parent sequences, but a useful yield of $\geq 60\%$ was still obtained in 24 h. For all sequence combinations, good yields were maintained on the preparative (0.5 nmol) scale using nonradiolabeled substrates.

Finally, the branch-site nucleotide dependences of both 8LX1 and 8LX6 were evaluated (Figure 9). Each deoxyribozyme displayed modest reactivity with branch-site rG and essentially no activity with rC or rU. For comparison, the original 7S11 deoxyribozyme showed a similar reactivity pattern (somewhat more activity with rG) (23), whereas 10DM24 and several other deoxyribozymes identified in the same effort favored rA but were considerably more reactive than 8LX1, 8LX6 and 7S11 with branch-site rC and rU (26).

Dependence of the deoxyribozymes on substrate identity (DNA or RNA)

An underlying assumption of our experimental strategy is that distinct deoxyribozymes are required to synthesize branched DNA versus branched RNA. This assumption has not been rigorously tested for our previously reported deoxyribozymes, although we did observe that 7S11 is inactive when L is DNA and highly active but with ~ 10 -fold reduced k_{obs} when R is 5'-adenylated DNA (49). Here, we evaluated the previously identified 15HA9 and 10DM24 deoxyribozymes along with the newly found 8LV13 and 8LX6, each with various combinations of DNA and RNA substrates (Supplementary Figure S1). For the deoxyribozymes that were selected to form branched DNA, 15HA9 had essentially no activity when

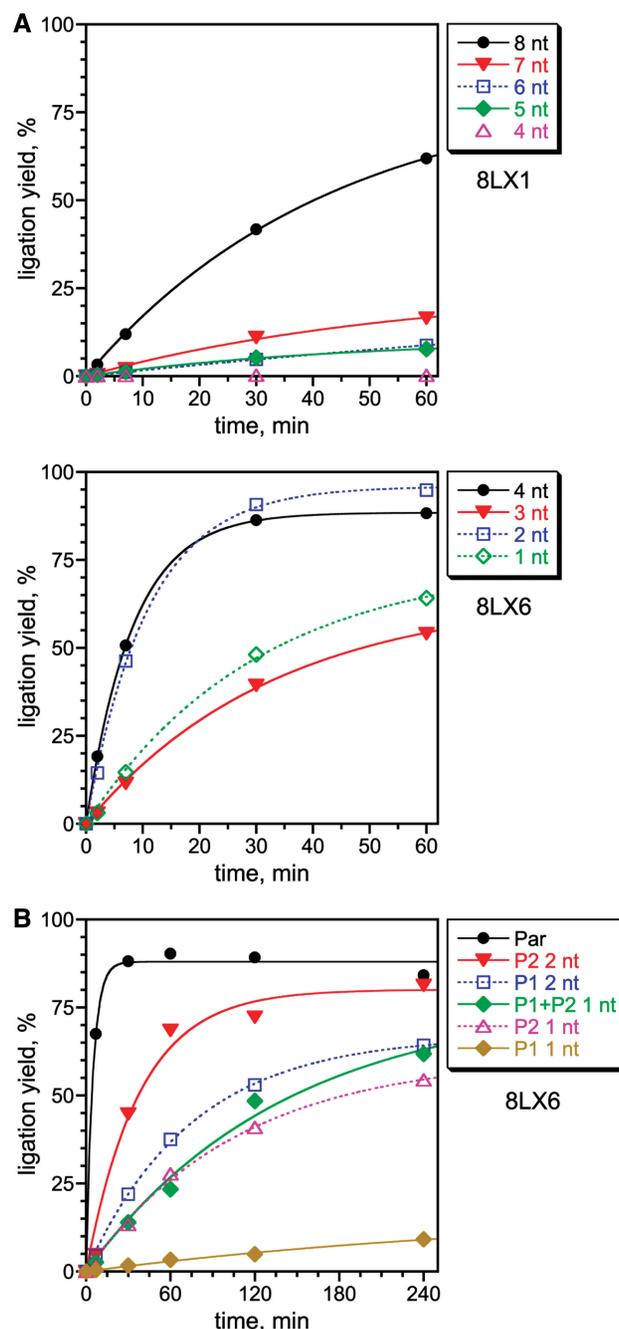


Figure 7. Assays of 8LX1 and 8LX6 with nucleotide changes in the L and R substrates, to determine the maximum allowed changes. (A) Assays of 8LX1 and 8LX6 to determine the maximum allowed changes to the R substrate. Each deoxyribozyme was assayed with Tv1 changes to L (all nucleotides of P1 and P2) as well as R (keeping the specified number of nucleotides unchanged beginning at the 5'-terminus). (B) Assays of 8LX6 with Tsn changes in the L substrate. Both P1 and P2 of the L substrate were changed by Tsn, except for the indicated number of nucleotides in the P1 and P2 regions. In all cases, full Watson-Crick base pairing was retained in both P1 and P2 between L and the deoxyribozyme.

either the L or R substrate was changed to RNA (Supplementary Figure S1A). 8LV13 showed a trace of activity with L RNA, but no activity with R RNA (Supplementary Figure S1B). For the deoxyribozymes

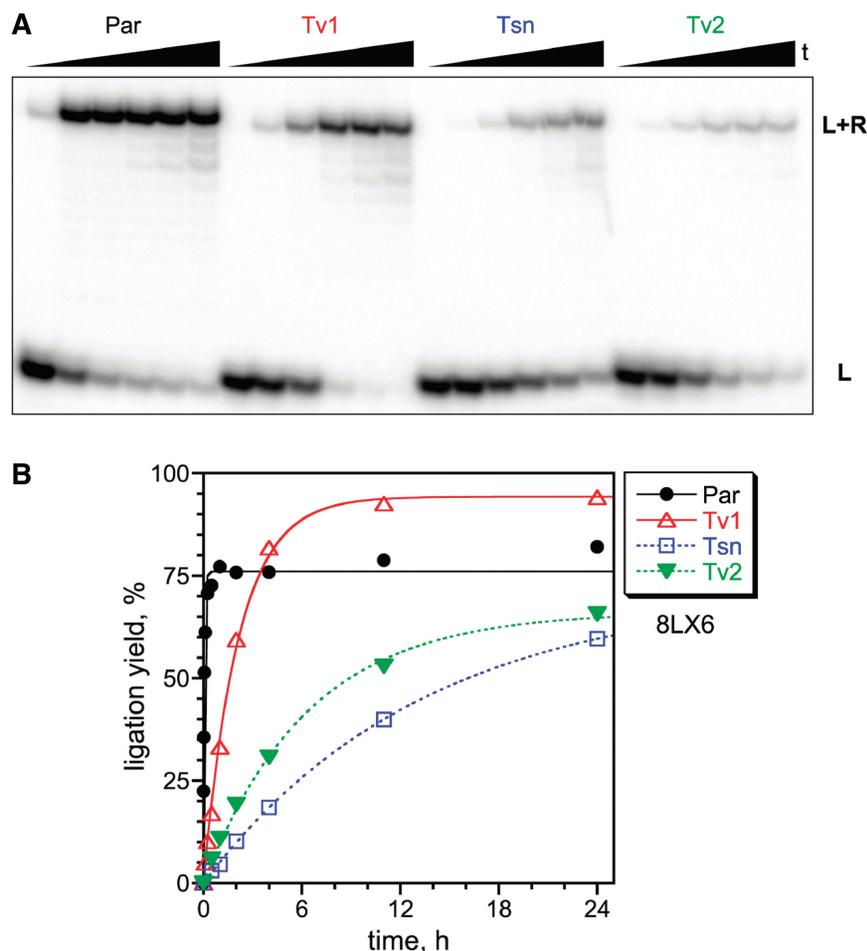


Figure 8. Establishing generality of 8LX6 for branched RNA synthesis by comprehensive sequence changes in both the L and R RNA substrates. For Tv1 and Tv2 changes, all nucleotides in L (P1 + P2) were changed, and all nucleotides except the 5'-G in R (P3 + P4) were changed. For Tsn changes, all L and R nucleotides were similarly changed, except the first two nucleotides adjacent to the branch site of P2 were unmodified. (A) PAGE image, with timepoints at $t = 0, 4, 30$ min; 4, 11 and 24 h. Sequence-dependent loss of signal due to 5'-phosphate hydrolysis is evident and likely unavoidable. Comparable yields were observed when the reactions were performed on the 0.5 nmol scale using nonradiolabeled substrates. (B) Kinetic plots. k_{obs} values: Par 17.1 h^{-1} (0.29 min^{-1}), Tv1 0.47 h^{-1} , Tsn 0.07 h^{-1} , Tv2 0.16 h^{-1} .

that were selected to form branched RNA, the previously studied 10DM24 lost all activity when the L substrate was DNA, but maintained strong activity when the R substrate was 5'-adenylated DNA (Supplementary Figure S1C). Similarly, the newly identified 8LX6 was inactive when L was DNA. However, unlike 10DM24, 8LX6 was much slower and lower yielding when the R substrate was 5'-adenylated DNA (Supplementary Figure S1D).

DISCUSSION

In this study, we performed new *in vitro* selection experiments with the goal of identifying improved deoxyribozymes for synthesizing covalently branched DNA and RNA. In both cases, we indeed identified new deoxyribozymes with desirable characteristics. Figure 10 shows a summary of the substrate sequence requirements and practical considerations for the best newly identified deoxyribozymes. For branched DNA synthesis, 8LV13

has a markedly improved rate constant relative to our previously reported 15HA9 (6), enabling high yield of branched DNA in 1 h at pH 7.5 instead of 3–4 days at pH 9.0 (Figure 4). 8LV13 has broad substrate sequence tolerance, similar to that of 15HA9. For branched RNA synthesis, our previous deoxyribozymes such as 7S11 (22,23,33) and 10DM24 (26) already have reasonably broad sequence tolerance. The newly found 8LX1 and 8LX6 deoxyribozymes also have wide sequence generality and, significantly, they function well at pH 7.5 rather than 9.0. 8LX6 is likely to be more useful than 8LX1 because 8LX6 tolerates a much greater number of nucleotide changes in the right-hand substrate, while providing good yield within 24 h and more quickly in many cases (Figure 8).

All of the new deoxyribozymes (8LV13, 8LX1 and 8LX6) were identified after making key changes to the selection design relative to our previous efforts with DNA-catalyzed synthesis of branched nucleic acids. In particular, here Mn^{2+} rather than Mg^{2+} was used as the metal ion cofactor, and loop A was expanded from 15 (or

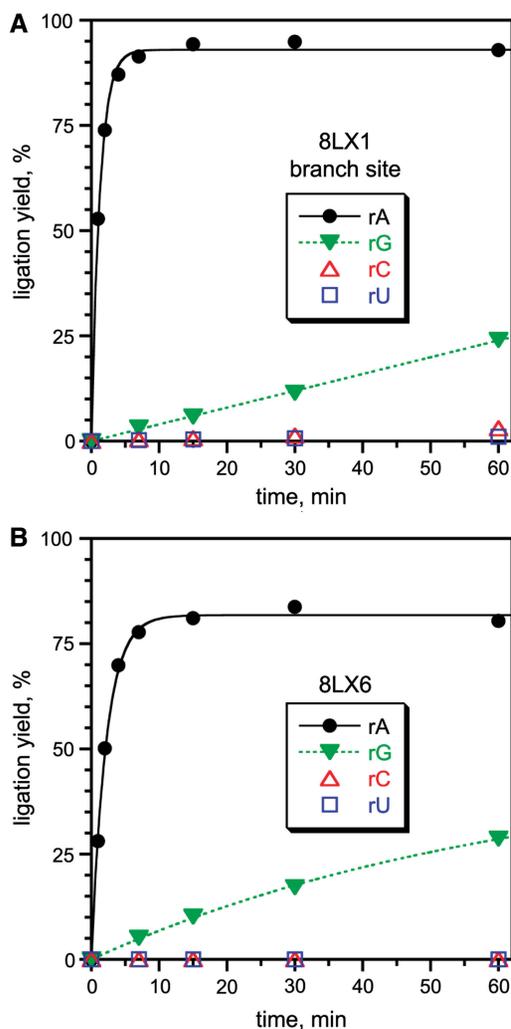


Figure 9. Branch-site nucleotide tolerance of (A) 8LX1, and (B) 8LX6. Substrate sequences were unchanged except for variation at the branch-site nucleotide of the L substrate.

fewer) to 33 nt. Because these changes were made simultaneously, our data cannot specifically allow assignment of the improved rate constant of 8LV13 (relative to 15HA9) to either the use of Mn^{2+} or the expansion of loop A. The data of Supplementary Figure S1 indicate that performing separate selection experiments to achieve synthesis of branched DNA and branched RNA was warranted, because 8LV13 does not function well to create branched RNA, and 8LX6 is poor at forming branched DNA.

For all of our deoxyribozymes that synthesize branched nucleic acids, we have typically analyzed their substrate sequence requirements by seeking boundaries beyond which nucleotide changes can be freely made, as long as Watson–Crick covariation is maintained (e.g. Figure 10 for the new deoxyribozymes identified here). In some cases, this assumption of a strict boundary may not be correct. For example, the data for sequence changes to the R substrate of 8LX6 reveal that retaining 3 nt at the 5'-end of the substrate leads to lower catalytic activity

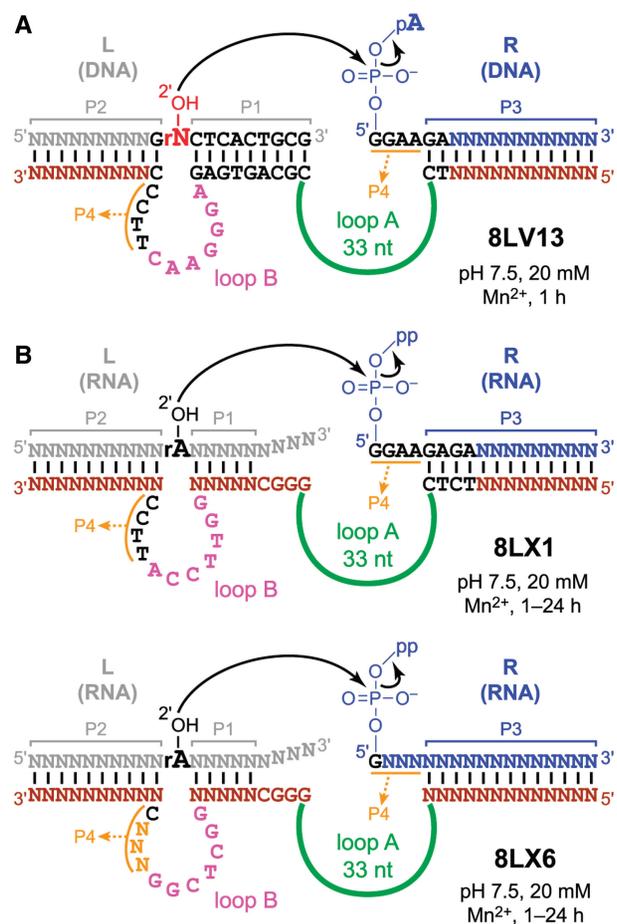


Figure 10. Substrate sequence requirements for the new deoxyribozymes. (A) 8LV13, for synthesis of branched DNA. (B) 8LX1 and 8LX6, for synthesis of branched RNA. For all three deoxyribozymes, see Table 1 for the 33 nt loop A sequences. The same color scheme as in Figure 2 is used, except that substrate nucleotides depicted here in black have restricted identities. Those nucleotides in color (P1 and P2, gray; P3 and P4, blue; branch site, red) can be changed freely as long as Watson–Crick covariation is maintained. For both 8LX1 and 8LX6, the 4 nt immediately following loop A (5'-GGGC-3' in both cases) are different from those at the outset of *in vitro* selection (Figure 2C) and are fixed even when the L substrate sequence is changed. For 8LX6, there may be some restriction on the L sequence for specific nucleotide changes involving transitions relative to the parent sequence (see Figure 7). For 8LX1 and 8LX6, branch-site rG is tolerated with modest efficiency (see Figure 9).

than retaining either 2 or 4 nt (Figure 7A), which is inconsistent with a strict boundary for generality. Such an observation suggests that the actual situation is more complex, perhaps due to discontinuous three-dimensional interactions involving the catalytic loops of the deoxyribozyme and the substrate nucleotides. Such explanations are difficult to validate without high-resolution structural information, which is currently unavailable for any catalytically active deoxyribozyme of any kind. Moreover, systematically analyzing a deoxyribozyme without making the first-order assumption of simple boundaries for sequence generality would require a great deal of complicated experimentation.

The covalently branched DNA and RNA molecules synthesized by deoxyribozymes such as 15HA9, 8LV13,

7S11, 10DM24, 8LX1 and 8LX6 all have only standard phosphodiester linkages linking all parts of the branched product. In contrast, other synthetic methods to prepare covalently branched nucleic acids do not provide standard phosphodiester linkages in the products (9,10,50,51), with some exceptions that all require specialized solid-phase synthesis of key substrates, and in many cases substantial organic synthesis of unusual phosphoramidites (52–54). We hope that the straightforward application of deoxyribozymes for synthesis of branched DNA and RNA will find utility in experiments for which preparation of covalently branched nucleic acids is desired.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

National Institutes of Health (R01 GM065966 to S.K.S.); and the David and Lucile Packard Foundation (fellowship to S.K.S.). The open access publication charge for this paper has been waived by Oxford University Press—NAR Editorial Board members are entitled to one free paper per year in recognition of their work on behalf of the journal.

Conflict of interest statement. None declared.

REFERENCES

- Patel, M.P., Baum, D.A. and Silverman, S.K. (2008) Improvement of DNA adenylation using T4 DNA ligase with a template strand and a strategically mismatched acceptor strand. *Bioorg. Chem.*, **36**, 46–56.
- Chioman, W. and Li, Y. (2002) Making AppDNA using T4 DNA ligase. *Bioorg. Chem.*, **30**, 332–349.
- Seeman, N.C. (2003) Biochemistry and structural DNA nanotechnology: an evolving symbiotic relationship. *Biochemistry*, **42**, 7259–7269.
- Grindley, N.D.F., Whiteson, K.L. and Rice, P.A. (2006) Mechanisms of site-specific recombination. *Annu. Rev. Biochem.*, **75**, 567–605.
- Lilley, D.M. (2008) Analysis of branched nucleic acid structure using comparative gel electrophoresis. *Q. Rev. Biophys.*, **41**, 1–39.
- Mui, T.P. and Silverman, S.K. (2008) Convergent and general one-step DNA-catalyzed synthesis of multiply branched DNA. *Org. Lett.*, **10**, 4417–4420.
- Horn, T. and Urdea, M.S. (1989) Forks and combs and DNA: the synthesis of branched oligodeoxyribonucleotides. *Nucleic Acids Res.*, **17**, 6959–6967.
- Collins, M.L., Irvine, B., Tyner, D., Fine, E., Zayati, C., Chang, C., Horn, T., Ahle, D., Detmer, J., Shen, L.P. *et al.* (1997) A branched DNA signal amplification assay for quantification of nucleic acid targets below 100 molecules/ml. *Nucleic Acids Res.*, **25**, 2979–2984.
- Horn, T., Chang, C.A. and Urdea, M.S. (1997) An improved divergent synthesis of comb-type branched oligodeoxyribonucleotides (bDNA) containing multiple secondary sequences. *Nucleic Acids Res.*, **25**, 4835–4841.
- Horn, T., Chang, C.A. and Urdea, M.S. (1997) Chemical synthesis and characterization of branched oligodeoxyribonucleotides (bDNA) for use as signal amplifiers in nucleic acid quantification assays. *Nucleic Acids Res.*, **25**, 4842–4849.
- Eckardt, L.H., Naumann, K., Pankau, W.M., Rein, M., Schweitzer, M., Windhab, N. and von Kiedrowski, G. (2002) Chemical copying of connectivity. *Nature*, **420**, 286.
- Peebles, C.L., Perlman, P.S., Mecklenburg, K.L., Petrillo, M.L., Tabor, J.H., Jarrell, K.A. and Cheng, H.L. (1986) A self-splicing RNA excises an intron lariat. *Cell*, **44**, 213–223.
- Burge, C.B., Tuschl, T. and Sharp, P.A. (1999) In Gesteland, R.F., Cech, T.R. and Atkins, J.F. (eds), *The RNA World*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 525–560.
- Peracchi, A. (2005) DNA catalysis: potential, limitations, open questions. *ChemBioChem*, **6**, 1316–1322.
- Baum, D.A. and Silverman, S.K. (2008) Deoxyribozymes: useful DNA catalysts in vitro and in vivo. *Cell. Mol. Life Sci.*, **65**, 2156–2174.
- Silverman, S.K. (2008) Catalytic DNA (deoxyribozymes) for synthetic applications—current abilities and future prospects. *Chem. Commun.*, 3467–3485.
- Schlosser, K. and Li, Y. (2009) Biologically inspired synthetic enzymes made from DNA. *Chem. Biol.*, **16**, 311–322.
- Silverman, S.K. (2010) DNA as a versatile chemical component for catalysis, encoding, and stereocontrol. *Angew. Chem. Int. Ed.*, **49**, in press.
- Joyce, G.F. (2004) Directed evolution of nucleic acid enzymes. *Annu. Rev. Biochem.*, **73**, 791–836.
- Wang, Y. and Silverman, S.K. (2003) Deoxyribozymes that synthesize branched and lariat RNA. *J. Am. Chem. Soc.*, **125**, 6880–6881.
- Wang, Y. and Silverman, S.K. (2003) Characterization of deoxyribozymes that synthesize branched RNA. *Biochemistry*, **42**, 15252–15263.
- 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.
- 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.
- Wang, Y. and Silverman, S.K. (2005) Directing the outcome of deoxyribozyme selections to favor native 3′-5′ RNA ligation. *Biochemistry*, **44**, 3017–3023.
- 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.
- Zelin, E., Wang, Y. and Silverman, S.K. (2006) Adenosine is inherently favored as the branch-site RNA nucleotide in a structural context that resembles natural RNA splicing. *Biochemistry*, **45**, 2767–2771.
- Zelin, E. and Silverman, S.K. (2007) Allosteric control of ribozyme catalysis using DNA constraints. *ChemBioChem*, **8**, 1907–1911.
- Silverman, S.K. (2007) Control of macromolecular structure and function using covalently attached double-stranded DNA constraints. *Mol. Biosyst.*, **3**, 24–29.
- Wang, Y. and Silverman, S.K. (2006) Experimental tests of two proofreading mechanisms for 5′-splice site selection. *ACS Chem. Biol.*, **1**, 316–324.
- Pratico, E.D. and Silverman, S.K. (2007) Ty1 reverse transcriptase does not read through the proposed 2′,5′-branched retrotransposition intermediate in vitro. *RNA*, **13**, 1528–1536.
- 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.
- 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.
- 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.
- Pradeepkumar, P.I., Höbartner, C., Baum, D.A. and Silverman, S.K. (2008) DNA-catalyzed formation of nucleopeptide linkages. *Angew. Chem. Int. Ed.*, **47**, 1753–1757.
- Sachdeva, A. and Silverman, S.K. (2010) DNA-catalyzed serine side chain reactivity and selectivity. *Chem. Commun.*, **46**, 2215–2217.

36. Höbartner, C. and Silverman, S.K. (2007) Engineering a selective small-molecule substrate binding site into a deoxyribozyme. *Angew. Chem. Int. Ed.*, **46**, 7420–7424.
37. Silverman, S.K. (2009) Deoxyribozymes: selection design and serendipity in the development of DNA catalysts. *Acc. Chem. Res.*, **42**, 1521–1531.
38. Chandra, M., Sachdeva, A. and Silverman, S.K. (2009) DNA-catalyzed sequence-specific hydrolysis of DNA. *Nat. Chem. Biol.*, **5**, 718–720.
39. Cruz, R.P.G., Withers, J.B. and Li, Y. (2004) Dinucleotide junction cleavage versatility of 8-17 deoxyribozyme. *Chem. Biol.*, **11**, 57–67.
40. Wang, W., Billen, L.P. and Li, Y. (2002) Sequence diversity, metal specificity, and catalytic proficiency of metal-dependent phosphorylating DNA enzymes. *Chem. Biol.*, **9**, 507–517.
41. Achenbach, J.C., Jeffries, G.A., McManus, S.A., Billen, L.P. and Li, Y. (2005) Secondary-structure characterization of two proficient kinase deoxyribozymes. *Biochemistry*, **44**, 3765–3774.
42. Sreedhara, A., Li, Y. and Breaker, R.R. (2004) Ligating DNA with DNA. *J. Am. Chem. Soc.*, **126**, 3454–3460.
43. Feldman, A.R. and Sen, D. (2001) A new and efficient DNA enzyme for the sequence-specific cleavage of RNA. *J. Mol. Biol.*, **313**, 283–294.
44. Feldman, A.R., Leung, E.K., Bennet, A.J. and Sen, D. (2006) The RNA-cleaving bipartite DNAzyme is a distinctive metalloenzyme. *ChemBioChem*, **7**, 98–105.
45. Huang, F., Bugg, C.W. and Yarus, M. (2000) RNA-catalyzed CoA, NAD, and FAD synthesis from phosphopantetheine, NMN, and FMN. *Biochemistry*, **39**, 15548–15555.
46. Dange, V., Van Atta, R.B. and Hecht, S.M. (1990) A Mn²⁺-dependent ribozyme. *Science*, **248**, 585–588.
47. Kazakov, S. and Altman, S. (1992) A trinucleotide can promote metal ion-dependent specific cleavage of RNA. *Proc. Natl Acad. Sci. USA*, **89**, 7939–7943.
48. Kolev, N.G., Hartland, E.I. and Huber, P.W. (2008) A manganese-dependent ribozyme in the 3'-untranslated region of *Xenopus* Vgl mRNA. *Nucleic Acids Res.*, **36**, 5530–5539.
49. Coppins, R.L. (2005) Deoxyribozymes that prepare branched and linear RNA by ligation. *Ph.D. Thesis*. University of Illinois at Urbana-Champaign.
50. Ogasawara, S. and Fujimoto, K. (2005) A novel method to synthesize versatile multiple-branched DNA (MB-DNA) by reversible photochemical ligation. *ChemBioChem*, **6**, 1756–1760.
51. Chandra, M., Keller, S., Gloeckner, C., Bornemann, B. and Marx, A. (2007) New branched DNA constructs. *Chem. Eur. J.*, **13**, 3558–3564.
52. Grötli, M., Eritja, R. and Sproat, B. (1997) Solid-phase synthesis of branched RNA and branched DNA/RNA chimeras. *Tetrahedron*, **53**, 11317–11346.
53. Damha, M.J. and Braich, R.S. (1998) Synthesis of a branched DNA/RNA chimera similar to the msDNA molecule of *Myxococcus xanthus*. *Tetrahedron Lett.*, **39**, 3907–3910.
54. Mendel-Hartvig, M., Kumar, A. and Landegren, U. (2004) Ligase-mediated construction of branched DNA strands: a novel DNA joining activity catalyzed by T4 DNA ligase. *Nucleic Acids Res.*, **32**, e2.