Published in final edited form as:
Biochemistry. 2011 May 31; 50(21): 4741–4749. doi:10.1021/bi200585n.

DNA-Catalyzed Covalent Modification of Amino Acid Side Chains in Tethered and Free Peptide Substrates†

On Yi Wong, **P. I. Pradeepkumar**‡, and **Scott K. Silverman***
Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801, USA

Abstract

This study focuses on the development of DNA catalysts (deoxyribozymes) that modify side chains of peptide substrates, with the long-term goal of achieving DNA-catalyzed covalent protein modification. We recently described several deoxyribozymes that modify tyrosine (Tyr) or serine (Ser) side chains by catalyzing their reaction with 5′-triphosphorylated RNA, forming nucleopeptide linkages. In each previous case, the side chain was presented in a highly preorganized three-dimensional architecture such that the resulting deoxyribozymes inherently cannot function with free peptides or proteins, which do not maintain the preorganization. Here we describe in vitro selection of deoxyribozymes that catalyze Tyr side chain modification of tethered and free peptide substrates, where the approach is potentially generalizable for catalysis involving large proteins. Several new deoxyribozymes for Tyr modification (and several for Ser modification as well) were identified; progressively better catalytic activity was observed as the selection design was strategically changed. The best new deoxyribozyme, 15MZ36, catalyzes covalent Tyr modification of a free tripeptide substrate with $k_{obs} = 0.50\ h^{-1}$ ($t_{1/2}$ 83 min) and up to 65% yield. These findings represent an important advance by demonstrating, for the first time, DNA catalysis involving free peptide substrates. The new results suggest the feasibility of DNA-catalyzed covalent modification of side chains of large protein substrates and provide key insights on how to achieve this goal.

Deoxyribozymes are particular DNA sequences (also called DNA enzymes or DNAzymes) that catalyze a growing repertoire of chemical reactions (1–3). First identified by in vitro selection for RNA cleavage (4, 5), DNA catalysts have been developed for reactions such as RNA ligation (6, 7), DNA depurination (8, 9), thymine dimer photoreversion (10–12), and the Diels-Alder reaction (13). When accompanied by appropriate metal ions, DNA is even capable of catalyzing challenging chemical reactions such as hydrolysis of DNA phosphodiester linkages with very large (10^{12}) rate enhancements (14–16).

An important long-term goal for the development of deoxyribozymes is catalyzing covalent modification of large protein substrates. Among other applications, achieving this goal will enable DNA-catalyzed synthesis of biologically relevant protein variants. Using small-molecule reagents or catalysts for this purpose is difficult, especially with selectivity for noncysteine residues (17–21), and natural protein enzymes cannot always be coaxed or tailored to achieve desired modifications (22, 23), As a first step towards the goal of DNA-catalyzed protein modification, we recently reported a deoxyribozyme, Tyr1, that covalently modifies

†This research was supported by grants to S.K.S. from the National Institutes of Health (R01 GM065966), the Defense Threat Reduction Agency (BRBAA08-L-2-0001), and the National Science Foundation (CHE–0842534).

*To whom correspondence should be addressed. scott@scs.illinois.edu. Phone: (217) 244-489. Fax: (217) 244-8024.
‡Present address: Department of Chemistry, Indian Institute of Technology, Bombay, India

SUPPORTING INFORMATION AVAILABLE: Additional experimental details. This material is available free of charge via the Internet at http://pubs.acs.org.

the tyrosine (Tyr) side chain by reaction with a 5′-triphosphorylated RNA oligonucleotide to form a Tyr-RNA nucleopeptide linkage (Figure 1) (24). Such connections between proteins and nucleic acids are found in several biological contexts (25–34). This previous report constituted the first example of DNA-catalyzed (or, more generally, nucleic acid-catalyzed) reactivity of an amino acid side chain [ribozyme-catalyzed reaction of a peptide *N*-terminal amino group has been reported (35)]. However, at that time we were unable to identify any deoxyribozymes that covalently modify the less reactive serine (Ser) side chain in >0.2% yield.

The Tyr1 deoxyribozyme was identified in the context of a highly preorganized architecture that stringently juxtaposes the Tyr nucleophile and 5′-triphosphate RNA electrophile (see Figure S1 in the Supporting Information) (24). This highly preorganized arrangement obviates the need for the DNA catalyst to spatially organize its substrates, because the spatial organization is already built into the system using Watson-Crick base pairs between the deoxyribozyme and two DNA segments obligatorily attached to the peptide. After our first report, we expanded the reactive peptide moiety from one to three amino acids, and via new selections we found deoxyribozymes that efficiently catalyze Ser-RNA nucleopeptide linkage formation (36). However, all of the reported deoxyribozymes retained the highly preorganized architecture. Notably, it is essentially impossible to expand the applicability of such highly preorganized deoxyribozymes to discrete, free peptide substrates and eventually to larger protein substrates, because a free peptide/protein substrate would not maintain the integrity of the preorganized architecture.

In the present study, we sought DNA-catalyzed amino acid side chain reactivity in the context of a nonpreorganized, inherently more open architecture (Figure 2), which in turn requires that the DNA enzyme itself spatially organizes the peptide substrate to achieve catalysis. Importantly, the nonpreorganized open architecture, unlike the previously used preorganized arrangement (24, 36), should be extendable to catalytic function with large and untethered protein substrates, because the tether in Figure 2 may be dispensable for catalysis. Via several selection approaches, we used a substrate containing a tripeptide moiety and identified deoxyribozymes that catalyze Tyr-RNA or Ser-RNA nucleopeptide linkage formation when the peptide is either tethered to a DNA anchor oligonucleotide or, in the case of Tyr, entirely untethered (free). These efforts used 5′-triphosphorylated RNA as an extensively validated electrophile with DNA catalysts (7), such that our focus here could remain on the peptide nucleophile; later efforts will expand the electrophile identity as well. The results of this study establish the ability of DNA to catalyze covalent modification of an amino acid side chain even without substantial spatial preorganization of the peptide substrate. Moreover, because—unlike in all of our previous work—the present approach is likely to be extendable to large proteins, we anticipate the identification of deoxyribozymes for covalent modification of side chains within large protein substrates.

EXPERIMENTAL PROCEDURES

Oligonucleotides

All DNA substrates and deoxyribozymes that were prepared by solid-phase synthesis were obtained from Integrated DNA Technologies (Coralville, IA) or prepared on an ABI 394 synthesizer and purified by 20% or 8% denaturing PAGE. The DNA anchor sequence was `5 -GGATAATACGACTCACTAT-3` , terminating either with a C_3-thiol linker (for DNA-C_3-CXA substrates) or with a hexa(ethylene glycol) spacer as well as a C_3-thiol linker (for DNA-HEG-CXA substrates). The DNA helper sequence for use with free CXA substrates was identical to the DNA anchor sequence, except terminating in a standard 3′-hydroxyl group. The 5′-triphosphorylated RNA substrate sequence was `5 -GGAAGAGAUGGCGACGG-3` , prepared by in vitro transcription using a double-stranded DNA template and T7 RNA polymerase and purified by 20% PAGE.

Synthesis of Peptide Substrates

For preparing the tethered substrates, the tripeptides CYA, CSA, and CAA were synthesized by standard solution-phase coupling methods starting with *N*-methylalanine and using Fmoc protection for the main chain along with 2-chlorotrityl (Tyr), trityl (Ser), and *t*-butyl disulfide (Cys) side chain protection. The Cys *t*-butyl disulfide was exchanged for a pyridyl disulfide by reduction with DTT and activation with 2,2′-dipyridyl disulfide. The final PyS-CXA tripeptides were analyzed by ^{1}H NMR and ^{13}C NMR spectroscopy as well as ESI mass spectrometry, with details provided in the Supporting Information.

Each PyS-CXA tripeptide was conjugated to DNA to form DNA-C_3-CXA or DNA-HEG-CXA. DNA-C_3-SS-C_3-OH or DNA-HEG-SS-C_3-OH was reduced with DTT (50 mM HEPES, pH 7.5, 50 mM DTT, 37 °C, 2 h) and precipitated with ethanol. The conjugation reaction was performed using 3 nmol of DNA-C_3/HEG-SH in 20 μL of 50 mM triethylammonium acetate, pH 7.0, 10 nmol of PyS-CXA in 5 μL of formamide, 15 μL of water, and 10 μL of additional formamide (total volume 50 μL) at 37 °C for 4 h, followed by 20% PAGE.

In Vitro Selection

For DNA-anchored tripeptide substrates, the selection procedures, cloning, and initial analysis of individual clones were performed essentially as described previously (6, 37). The random deoxyribozyme pool was `5 -CGAAGTCGCCATCTCTTC–N`$_{40}$`–ATAGTGAGTCGTATTAAGCTGATCCTGATGG-3` . Partially (25%) randomized pools were prepared using phosphoramidite ratios computed as described previously (38). PCR primers were `5 -(AAC)`$_4$`XCCATCAGGATCAGCTTAATACGACTCACTAT-3` (where `X` = HEG spacer to stop Taq polymerase) and `5 -CGAAGTCGCCATCTCTTC-3` (5′-phosphorylated to enable ligation using T4 RNA ligase to the 3′-terminus of the 5′-triphosphorylated RNA substrate). Each selection or reselection experiment was initiated with 200 pmol of pool. In each round, the key selection step used 100 pmol of DNA-anchored tripeptide substrate in a final volume of 20 μL containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 °C for 2 h.

Radiolabeling of 5′-Triphosphorylated RNA Substrate

5′-^{32}P-radiolabeled cytidine 3′,5′-bisphosphate (pCp) was prepared by incubating 60 pmol cytidine 3′-monophosphate (Cp), 40 pmol of γ-^{32}P-ATP, and 10 U of T4 PNK (Fermentas) in 10 μL of 1× T4 PNK buffer (50 mM Tris, pH 7.6, 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine, and 0.1 mM EDTA) at 37 °C for 2 h. The PNK was inactivated by heating the sample at 95 °C for 5 min and cooling on ice for 10 min. The resulting sample was assumed to contain 40 pmol of 5′-^{32}P-pCp. The 3′-^{32}P-radiolabeled RNA substrate was prepared by ligating 40 pmol of unradiolabeled R to 40 pmol of 5′-^{32}P-pCp using 10 U of T4 RNA ligase (Fermentas) in 20 μL of 50 mM Tris, pH 7.5, 5 mM $MgCl_2$, 10 mM DTT, and 0.05 mM ATP at 37 °C for 12 h, followed by 20% PAGE.

Deoxyribozyme Activity Assays with DNA-Anchored Tripeptide Substrates

The general assay approach was described previously (6). The DNA-tripeptide substrate is designated as the L (left-hand) substrate, and the 5′-triphosphorylated RNA substrate is designated as the R (right-hand) substrate. In each deoxyribozyme assay, R was the limiting reagent relative to the deoxyribozyme (E) and the L substrate. A 14 μL sample containing 1 pmol of partially 5′-^{32}P-radiolabeled R (as described above), 10 pmol of E, and 20 pmol of L was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 1 min and cooling on ice for 5 min. The ligation reaction was initiated by addition of stock solutions to a final volume of 20 μL containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$. Final concentrations were 50 nM R, 0.5 μM

E, and 1 μM L (1:10:20 R:E:L; single-turnover conditions). At appropriate times, 1.4 μL aliquots were quenched with 6 μL of stop solution (80% formamide, 1× TBE [89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3], 50 mM EDTA, 0.025% bromophenol blue, 0.025% xylene cyanol).

Before PAGE, to each sample was added 100 pmol of a "decoy oligonucleotide", which was a 60-mer complementary to the deoxyribozyme's enzyme region (~40 nt) along with ~10 nt of binding arm on either side. This decoy oligonucleotide was added to displace the deoxyribozyme from the substrate and ligation product. When the decoy was omitted, gel bands were noticeably smeared, which inhibited proper quantification.

Samples were separated by 20% PAGE and quantified with a PhosphorImager. Values of k_{obs} were obtained by fitting the yield versus time data directly to first-order kinetics; i.e., yield = Y•$(1 - e^{-kt})$, where $k = k_{obs}$ and Y is the final yield. When k_{obs} was sufficiently low such that an exponential fit was not meaningful, the initial points were fit to a straight line, and k_{obs} was taken as the slope of the line.

Deoxyribozyme Activity Assays with Free Tripeptide Substrates

The procedure was similar to that described above, except that 20 pmol of DNA-anchored tripeptide substrate was replaced with 20 pmol of DNA helper oligonucleotide, and 20 nmol of free tripeptide was added after the annealing step. The tripeptide was prepared as a 50 mM stock solution in formamide. Final concentrations were 50 nM R, 0.5 μM E, 1 μM helper oligonucleotide, and 1 mM tripeptide substrate (1:10:20:20,000 R:E:helper:tripeptide; single-turnover conditions), with 2% formamide by volume. For determining $K_{d,app}$ of the CYA tripeptide for the 15MZ36 deoxyribozyme, the concentration of tripeptide was varied between 0 and 3 mM, with 2% formamide in the final sample in all cases.

Product Characterization Assays

Ligation products were treated with either DTT or RNase T1 and analyzed by MALDI mass spectrometry. For DTT reduction, 100 pmol of ligation product was incubated in 50 mM HEPES, pH 7.5, and 50 mM DTT in 20 μL total volume at 37 °C for 2 h, precipitated with ethanol, and desalted with C_{18} ZipTip (Millipore). For RNase T1 digestion of DNA-C_3/HEG-CXA–RNA, 100 pmol of ligation product was incubated in 50 mM Tris, pH 7.5, 2 mM EDTA, and 20 U of RNase T1 (Ambion) in 20 μL total volume at 37 °C for 4 h, extracted by phenol/chloroform, precipitated by ethanol, and desalted with C_{18} ZipTip. For RNase T1 digestion of untethered CYA–RNA, the same procedure was used, except extraction and precipitation were omitted. All MALDI mass spectra were obtained in the mass spectrometry laboratory of the UIUC School of Chemical Sciences. See Table S1 for data.

RESULTS

Selection Experiments Using Tethered Tripeptides

We began with a set of in vitro selection experiments that placed DNA-anchored tripeptides in an "open" architecture (Figure 2) rather than the previously used highly preorganized architecture, from which generalization to free peptide or protein substrates cannot be achieved because the peptide must be covalently attached on both termini to DNA (Figure S1). Here either a Cys-Tyr-Ala (CYA) or Cys-Ser-Ala (CSA) tripeptide was used, alongside a negative control experiment with Cys-Ala-Ala (CAA, lacking any nucleophile). The substrates each had a short three-carbon tether linking the 3′-terminal oxygen atom of the DNA anchor to the cysteine of the tripeptide via a disulfide bond, denoted "C_3-tethered" in Figure 2A.

For all three CXA substrates, the in vitro selection experiment was performed essentially as described previously for both RNA ligation (6) and nucleopeptide linkage formation (24, 36). An N_{40} DNA pool oligonucleotide was covalently joined at its 5′-terminus to the 3′-terminus of a 5′-triphosphorylated RNA strand using T4 RNA ligase. (N_{40} was chosen as a compromise between shorter random region lengths, in which sequence space is well covered but structural complexity may be insufficient for catalysis, and longer random region lengths, in which sequence space is very sparsely sampled but more complex structures can be accessed.) Then, individual DNA sequences capable of joining the tripeptide moiety of the DNA-anchored tripeptide to the RNA under various incubation conditions were separated on the basis of a polyacrylamide gel electrophoresis (PAGE) shift. These DNA sequences were amplified by PCR. The desired single strand was separated by PAGE; the two strands were of unequal length due to a nonamplifiable spacer incorporated into one of the PCR primers. Finally, the DNA pool—now enriched in catalytically active sequences—was ligated to the 5′-triphosphorylated RNA to enable the next selection round.

The standard incubation conditions during each key selection step were 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ for 2 h at 37 °C. Both Mg^{2+} and Mn^{2+} were included because both divalent metal ions have proven effective as cofactors for various deoxyribozymes (1, 2, 7). The CYA selection had 24% activity at round 9 and 47% activity at round 10; additional rounds did not lead to any further increase. In contrast, both the CSA selection and the negative control CAA selection showed undetectable ligation activity (<0.5%) even after 15 rounds.

Deoxyribozymes from round 10 of the CYA selection were cloned. Based on initial screening data, individual deoxyribozymes were sequenced, revealing a single dominant sequence (Figure 3). According to our laboratory's systematic nomenclature, this deoxyribozyme was named 10KC3, for round 10, selection experiment designated KC, clone 3. The 10KC3 deoxyribozyme was independently prepared by solid-phase synthesis and assayed for catalytic activity with each of three different CYA variants: the original DNA-tethered CYA substrate (DNA-C_3-CYA in Figure 2A), a variant DNA-anchored substrate in which the short three-carbon tether was replaced with a longer hexa(ethylene glycol) tether (DNA-HEG-CYA in Figure 2B; "HEG-tethered"), or the free CYA tripeptide in which the cysteine sulfur atom was blocked as the *t*-butyl disulfide rather than connected to a DNA oligonucleotide (free CYA in Figure 2C; "untethered"). In assays with the free CYA tripeptide, the DNA anchor sequence was included as an unmodified "helper oligonucleotide" to suppress unintended interactions that may involve the deoxyribozyme's 3′-terminal binding arm if left single-stranded.

As shown in Figure 4, 10KC3 had relatively high ligation activity with the DNA-C_3-CYA substrate (70% yield in 10 h and 78% yield in 60 h) and lower but still substantial activity with the DNA-HEG-CYA substrate (40% yield in 10 h and 62% yield in 60 h). The k_{obs} values with the two substrates were 0.28 and 0.098 h^{-1}, respectively. With the free CYA substrate, ~10% yield at 60 h (k_{obs} ~0.002 h^{-1}) was reproducibly observed.

Follow-up Selections Using CYA Substrates

The ~10% yield of ligation product formed by 10KC3 with the untethered CYA substrate was modest, but this was nevertheless encouraging as the first example of covalent modification of a free peptide by a deoxyribozyme. We therefore performed two key follow-up efforts using appropriate CYA substrates.

First, we noted that 10KC3 was identified from an entirely random pool in which sequence space is sampled rather sparsely. In particular, because only 10^{-10} of sequence space is explored when using an N_{40} random pool (200 pmol = 10^{14} sequences in the initial pool sample; $4^{40} \approx 10^{24}$ sequences in sequence space) (1), 10KC3 is likely to be suboptimal for catalysis. With

this consideration, we performed reselection (i.e., in vitro evolution) by partially randomizing the 10KC3 sequence to the extent of 25% per nucleotide position and reselecting for nucleopeptide linkage formation using the original DNA-C_3-CYA substrate. The reselection pool showed detectable (2%) ligation activity at round 2 and strong (57%) activity by round 6. In subsequent rounds, we increased the selection pressure by decreasing the incubation time from 2 h to 10 min and then to 1 min. After round 9 with 1 min incubation, resulting in 5% ligation activity, individual deoxyribozymes were cloned, screened, and sequenced. Six sequences emerged (Figure 3) and were prepared by solid-phase synthesis. The 9NG14 deoxyribozyme had representative activity with the DNA-C_3-CYA substrate (Figure 5A), with k_{obs} = 5.6 h^{-1} and 77% yield; this is a 20-fold increase in k_{obs} relative to the parent 10KC3. 9NG14 was also 16-fold faster than 10KC3 with the DNA-HEG-CYA substrate (k_{obs} = 1.6 h^{-1} and 75% yield). Intriguingly, although 9NG14 was not selected directly for activity with the free CYA substrate, its activity with this untethered tripeptide was also improved relative to 10KC3, with k_{obs} = 0.063 h^{-1} ($t_{1/2}$ 11 h, ~30-fold faster than 10KC3) and 40% yield in 60 h. The other five 9NG deoxyribozymes each showed a similar overall pattern of activity with the three CYA-containing substrates (Figure S2).

Second, we performed a new selection experiment using the DNA-HEG-CYA substrate, where the HEG tether is much longer tether than C_3 (Figure 2). Because the longer tether imposes less preorganization during the selection process, we expected improved activity by the resulting deoxyribozymes with the free peptide substrate, for which no preorganization is possible. Starting from a random N_{40} pool and using DNA-HEG-CYA, 46% activity was observed by round 11. The best individual deoxyribozyme, 11MN5 (see Figure 3 for sequence), indeed had better activity than 10KC3 with all three substrate variants (Figure 6). Although 11MN5 was not selected directly with free CYA, its activity with this untethered substrate was promising, with k_{obs} = 0.048 h^{-1} and 51% yield after 60 h. Indeed, the activity profile of 11MN5 was similar to that of the reselected 9NG14 (Figure 5B). Several other 11MN clones (Figure 3) showed similar overall patterns of activity with the three CYA-containing substrates (Figure S3).

Seeking Serine Reactivity via an Alternative Selection Approach with a Tethered Hydroxyl Substrate

Our previous selection experiments have shown that achieving DNA-catalyzed reaction of the Ser side chain is much more challenging than analogous reactivity of Tyr, even in a highly preorganized architecture (24, 36). Because the selection effort described above with the DNA-C_3-CSA substrate gave no activity, we pursued an alternative approach. A selection experiment was initiated with a DNA substrate conjugated at its 3′-terminus to a tri(ethylene glycol) (TEG) tether that terminates in an aliphatic hydroxyl group, i.e., DNA-TEG-OH. Detectable DNA-catalyzed ligation between DNA-TEG-OH and the 5′-triphosphorylated RNA substrate was observed in round 11 with 8% yield. From the round 11 pool, we continued the selection effort using DNA-C_3-CSA as substrate. The pool's ligation activity was 6% at round 15, at which point individual deoxyribozyme were identified by cloning.

The best new deoxyribozyme, 15MZ36 (Figure 3), had k_{obs} = 0.52 h^{-1} and 62% yield after 60 h with the DNA-C_3-CSA substrate, but only a small amount of activity (~6% yield after 60 h) with DNA-HEG-CSA and no detectable activity (<0.5%) with free CSA (Figure 6). When tested with the three analogous CYA substrates, 15MZ36 had very strong activity with the two tethered variants (k_{obs} = 80 and 19 h^{-1} with C_3- and HEG-tethered CYA, respectively), as well as k_{obs} = 0.50 h^{-1} ($t_{1/2}$ 83 min) and 48% yield after 60 h with free CYA. Other 15MZ clones (Figure 3) showed similar activities (Figure S4). As negative controls, DNA-C_3-CAA and free CAA were unreactive in the presence of 15MZ36 (<0.5% yield; data not shown). We examined the dependence of 15MZ36 activity on the concentration of the free CYA substrate. The

$K_{d,app}$ value was found to be on the order of 1 mM, with 65% yield at 60 h using 3 mM CYA (Figure S5).

Because 15MZ36 has much higher activity with either tethered or free CYA substrates than with the corresponding CSA analogues, we partially randomized and reselected 15MZ36, seeking improved activity directly with the DNA-C_3-CSA substrate. However, no improved activity was identified (Figure S6), even though many mutations were apparent (Figure 3, clones 6QG6 and 6QG18).

15MZ36 provides this study's best overall catalytic rate with the free CYA substrate, despite the fact that 15MZ36 was identified using a CSA substrate in the later selection rounds. We therefore assayed clones from a parallel selection effort in which the substrate was changed from the TEG-tethered hydroxyl group to DNA-C_3-CYA rather than DNA-C_3-CSA. Although numerous sequences related to 15MZ36 were identified, again no improvement in activity was found relative to 15MZ36 itself (Figure S7; Figure 3, clones 15NZ11 and 15NZ16).

Dependence of 15MZ36 Free Peptide Substrate Reactivity on 3′-Terminal Composition of Deoxyribozyme

15MZ36 was identified via in vitro selection rounds that were conducted entirely using a tethered substrate. In particular, the 3′-terminal binding arm of the deoxyribozyme was always Watson-Crick base paired to the substrate's DNA anchor oligonucleotide, which in turn was tethered to the nucleophilic hydroxyl group. We therefore evaluated how 15MZ36 activity with the free CYA substrate depends on the composition of this region of the deoxyribozyme, in two ways. First, we showed that truncating the 3′-terminal binding arm of the deoxyribozyme substantially suppresses but does not entirely eliminate ligation activity (rate decrease ~300-fold; Figure S8A). Omitting the helper oligonucleotide and leaving this region of the deoxyribozyme single-stranded was functionally equivalent to truncating the 3′-terminal binding arm, indicating that a DNA duplex at the 3′-terminus (as present during selection) is required for efficient catalysis. Second, we examined how the reactivity of free CYA is influenced by using a variant of the helper oligonucleotide that is connected to any of a C_3, TEG, or HEG tether, rather than having a free 3′-OH terminus. No tether-containing helper oligonucleotide improved the activity of 15MZ36 relative to its activity in the presence of a non-tether-containing helper oligonucleotide (Figure S8B). Therefore, the tether atoms themselves do not contribute to this deoxyribozyme's catalytic function. In related assays, similar observations were made for 9NG14 and 11MN5 (data not shown).

Characterization of the Tyr-RNA and Ser-RNA Ligation Products

We performed experiments to characterize several of the key ligation products, as shown in Figure 7 with representative data for the 15MZ36 DNA-C_3-CSA–RNA product. The deoxyribozyme ligation product was analyzed by MALDI mass spectrometry. Then, both DTT and RNase T1 digestions were separately performed on the ligation product, in both cases leading to the expected masses based on disulfide cleavage or nonspecific RNA degradation, respectively. Analogous data were obtained for several other ligation products, including the free CYA–RNA product formed by 15MZ36 (see Table S1 in the Supporting Information).

DISCUSSION

This study aimed to advance two major goals in our understanding of DNA catalysis. First, we sought to identify new deoxyribozymes that covalently modify amino acid side chains of free peptide substrates, rather than of substrates in which the peptide moiety is highly preorganized relative to the deoxyribozyme. This objective is important so that DNA may eventually be used to catalyze covalent modifications of large proteins, which is possible only in the open,

nonpreorganized architectures of Figure 2. Second, we sought to understand the in vitro selection process by which such deoxyribozymes are identified, so that our future experimental efforts towards large protein modification can be optimally directed. The findings of this study have provided key advances toward both of these important goals.

Activity Characteristics of the New Deoxyribozymes

We initially focused our efforts on DNA-catalyzed reactivity of Tyr because of its much greater nucleophilic reactivity than Ser, as established for DNA catalysis via our previous work with highly preorganized substrates (24, 36). Here, our first effort used the DNA-C_3-CYA substrate, which has only a short tether from the DNA anchor to the Tyr-containing tripeptide, resulting in the 10KC3 deoxyribozyme. This DNA enzyme catalyzes Tyr-RNA ligation slowly with the C_3- and HEG-tethered substrates and even less efficiently with the free CYA substrate (Figure 4). Reselection to optimize 10KC3 using the same C_3-tethered substrate led to substantially increased (16–30-fold) k_{obs} values with the tethered and free CYA substrates relative to 10KC3 itself (Figure 5A). This outcome clearly indicated that optimization of side chain reactivity with a tethered substrate can lead to better activity with the free substrate as well, even without devising a selection procedure that directly uses the free substrate. Nevertheless, the best reselected 10KC3 variant, 9NG14, was still rather slow with the free CYA substrate ($t_{1/2}$ 11 h). Therefore, selection was performed directly with the HEG-tethered substrate, seeking improved activity with DNA-HEG-CYA and anticipating that the resulting deoxyribozymes would also work well with free CYA. Such improvement was indeed observed (Figure 5B), leading to the 11MN5 deoxyribozyme that has activity comparable to that found for the reselected 10KC3 variants such as 9NG14.

An alternative approach in which the TEG-tethered aliphatic hydroxyl (DNA-TEG-OH) was first presented as the nucleophile to the random-sequence pool, and then DNA-C_3-CSA was used, led to 15MZ36 as the most active deoxyribozyme found in this entire study (Figure 6). The $t_{1/2}$ of 83 min for free CYA reactivity by 15MZ36 is a ca. 300-fold improvement upon that of 10KC3 and about an order of magnitude better than that of 9NG14 and 11MN5. Even though 15MZ36 was selected directly using CSA (in the form of DNA-C_3-CSA), its reactivity was clearly better with the CYA substrate analogues, which likely reflects the much greater general reactivity of Tyr over Ser. A half-life on the order of an hour is encouraging, although practical DNA-catalyzed protein modification will require identifying deoxyribozymes that have improved peptide K_d values (see below). Moreover, discrimination among different identities at a given amino acid position (e.g., C<u>Y</u>A versus C<u>S</u>A) is less of a concern than is selectivity among the same side chains present at different positions along the peptide chain, because any large protein substrate will almost certainly have a well-defined sequence rather than multiple amino acid identities at any one position (36). Investigating such discrimination for these and other deoxyribozymes will be a topic of future experiments.

Sequence Alignments and Relationships Among the New Deoxyribozymes

As a broad objective, we wish to understand deoxyribozyme structure-function relationships. Indeed, our first efforts in this regard for DNA enzymes with a different catalytic activity, hydrolysis of DNA phosphodiester bonds, have recently been reported (16, 39). Although detailed experimental analyses of the new peptide-modifying deoxyribozymes are clearly beyond the scope of the present study, a brief comparison of the various enzyme-region sequences is warranted. As shown in Figure 3, each of the more active deoxyribozymes 9NG14, 11MN5, and 15MZ36 differ from the original 10KC3 by numerous mutations (10 for 9NG14; 9 as well as three inserted nucleotides for 11MN5; and 8 as well as one inserted nucleotide for 15MZ36). In addition, there are three large nucleotide blocks that are entirely or largely conserved. No obvious commonalities are apparent among the nonconserved nucleotides. The many shared nucleotides suggest the emergence of one overall motif for DNA-catalyzed

reactivity of Tyr/Ser side chains in peptide substrates. Further explorations of these deoxyribozymes as well as experiments aimed at identification of entirely new catalytic DNA motifs are the subject of ongoing work.

Secondary structure predictions using mfold (40) suggest the presence of a stem-loop element for each of the new deoxyribozymes, although the folding energies are low and the stem-loop locations are inconsistent among the predicted structures (Figure S9). By themselves, such secondary structure elements—if actually present—neither reveal tertiary structure nor provide insight into catalytic mechanism. Considering the long-standing challenges associated with understanding structures and mechanisms of deoxyribozymes (1–3, 41), addressing these important facets of the newly identified DNA catalysts will require considerable future work.

Implications for Future Experimental Efforts

From the results obtained in this study, we derive several key implications for efforts to identify deoxyribozymes that covalently modify amino acid side chains of peptide and protein substrates.

1. Identification and optimization of deoxyribozyme-catalyzed ligation activity with a tethered peptide substrate can lead to substantial activity with the analogous free peptide substrate. In future work, the limits of this approach should be evaluated by continuing selection efforts with tethered peptide substrates. The lack of catalytic contribution by the tether atoms that connect the tripeptide to the DNA anchor suggest that the DNA-anchored substrates enjoy merely a nonspecific entropic advantage relative to non-anchored substrates; the tether itself is not actively recognized. Selecting for catalytic activity directly with a free peptide substrate may be productive in eliciting better $K_{d,app}$ values for the peptide substrate, because explicit selection pressure can be imposed in this regard. Importantly, new methodology must be developed to implement selections that directly use free peptides, because reaction of 5′-triphosphorylated RNA with a free peptide (lacking a tethered DNA anchor strand) does not lead to an appreciable gel shift that forms the basis for our current selection procedure.
2. All of the current efforts have used N_{40} random pool DNA along with Mg^{2+} and Mn^{2+} as metal ion cofactors. Both the length of the random region and the identity of the metal ions should be evaluated comprehensively in future work.
3. The 15MZ36 deoxyribozyme was identified by selection using a combination of the TEG-tethered hydroxyl substrate and the DNA-C_3-CSA substrate, yet it has the greatest activity with the CYA substrates. The topic of amino acid selectivity should be investigated further, especially with regard to the same amino acid placed at different positions along the peptide substrate (corresponding to the common practical situation in which the peptide/protein target sequence is known, but the same amino acid is found at more than one sequence position). Initial results for unrelated deoxyribozymes that function in a highly preorganized architecture were very encouraging (36), suggesting that useful site selectivity among amino acid side chains is achievable by DNA catalysts.

The successful DNA-catalyzed covalent modification of a free tripeptide bodes well for the prospect of using large proteins as deoxyribozyme substrates. Efforts are currently underway to expand the length of the reactive peptide moiety involved in our selection experiments, as well as to apply the current deoxyribozymes to large proteins. Importantly, we are also expanding our efforts to include reaction partners other than 5′-triphosphorylated RNA, because the long-term general goal of covalent protein modification by DNA catalysts requires use of other electrophiles.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Ying Xiao for technical assistance during the early phase of the MZ selection experiment. We are grateful for advice on solution-phase peptide synthesis from Danica Galonić Fujimori.

REFERENCES

1. Silverman SK. Catalytic DNA (deoxyribozymes) for synthetic applications—current abilities and future prospects. Chem. Commun. 2008:3467–3485.
2. Schlosser K, Li Y. Biologically inspired synthetic enzymes made from DNA. Chem. Biol. 2009; 16:311–322. [PubMed: 19318212]
3. Silverman SK. DNA as a Versatile Chemical Component for Catalysis, Encoding, and Stereocontrol. Angew. Chem. Int. Ed. 2010; 49:7180–7201.
4. Breaker RR, Joyce GF. A DNA enzyme that cleaves RNA. Chem. Biol. 1994; 1:223–229. [PubMed: 9383394]
5. Silverman SK. In vitro selection, characterization, and application of deoxyribozymes that cleave RNA. Nucleic Acids Res. 2005; 33:6151–6163. [PubMed: 16286368]
6. 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]
7. Silverman SK. Deoxyribozymes: Selection Design and Serendipity in the Development of DNA Catalysts. Acc. Chem. Res. 2009; 42:1521–1531. [PubMed: 19572701]
8. Sheppard TL, Ordoukhanian P, Joyce GF. A DNA enzyme with *N*-glycosylase activity. Proc. Natl. Acad. Sci. USA. 2000; 97:7802–7807. [PubMed: 10884411]
9. Höbartner C, Pradeepkumar PI, Silverman SK. Site-selective depurination by a periodate-dependent deoxyribozyme. Chem. Commun. 2007:2255–2257.
10. Chinnapen DJ, Sen D. A deoxyribozyme that harnesses light to repair thymine dimers in DNA. Proc. Natl. Acad. Sci. USA. 2004; 101:65–69. [PubMed: 14691255]
11. Chinnapen DJ, Sen D. Towards elucidation of the mechanism of UV1C, a deoxyribozyme with photolyase activity. J. Mol. Biol. 2007; 365:1326–1336. [PubMed: 17141270]
12. Thorne RE, Chinnapen DJ, Sekhon GS, Sen D. A deoxyribozyme, Sero1C, uses light and serotonin to repair diverse pyrimidine dimers in DNA. J. Mol. BIol. 2009; 388:21–29. [PubMed: 19281822]
13. Chandra M, Silverman SK. DNA and RNA Can Be Equally Efficient Catalysts for Carbon-Carbon Bond Formation. J. Am. Chem. Soc. 2008; 130:2936–2937. [PubMed: 18271591]
14. Chandra M, Sachdeva A, Silverman SK. DNA-catalyzed sequence-specific hydrolysis of DNA. Nat. Chem. Biol. 2009; 5:718–720. [PubMed: 19684594]
15. Fekry MI, Gates KS. DNA-catalyzed hydrolysis of DNA phosphodiesters. Nat. Chem. Biol. 2009; 5:710–711. [PubMed: 19763101]
16. Xiao Y, Chandra M, Silverman SK. Functional Compromises among pH Tolerance, Site Specificity, and Sequence Tolerance for a DNA-Hydrolyzing Deoxyribozyme. Biochemistry. 2010; 49:9630–9637. [PubMed: 20923239]
17. Tilley, SD.; Joshi, NS.; Francis, MB. Wiley Encyclopedia of Chemical Biology. Begley, TP., editor. Hoboken, NJ: John Wiley and Sons; 2009. DOI 10.1002/9780470048672.wecb9780470048493
18. Joshi NS, Whitaker LR, Francis MB. A three-component Mannich-type reaction for selective tyrosine bioconjugation. J. Am. Chem. Soc. 2004; 126:15942–15943. [PubMed: 15584710]
19. Minakawa M, Guo HM, Tanaka F. Imines that react with phenols in water over a wide pH range. J. Org. Chem. 2008; 73:8669–8672. [PubMed: 18844415]
20. Antos JM, McFarland JM, Iavarone AT, Francis MB. Chemoselective tryptophan labeling with rhodium carbenoids at mild pH. J. Am. Chem. Soc. 2009; 131:6301–6308. [PubMed: 19366262]

21. Ban H, Gavrilyuk J, Barbas CF 3rd. Tyrosine bioconjugation through aqueous ene-type reactions: a click-like reaction for tyrosine. J. Am. Chem. Soc. 2010; 132:1523–1525. [PubMed: 20067259]
22. Gamblin DP, Scanlan EM, Davis BG. Glycoprotein synthesis: an update. Chem. Rev. 2009; 109:131–163. [PubMed: 19093879]
23. Jäckel C, Kast P, Hilvert D. Protein design by directed evolution. Annu. Rev. Biophys. 2008; 37:153–173. [PubMed: 18573077]
24. Pradeepkumar PI, Höbartner C, Baum DA, Silverman SK. DNA-Catalyzed Formation of Nucleopeptide Linkages. Angew. Chem. Int. Ed. 2008; 47:1753–1757.
25. Tse Y-C, Kirkegaard K, Wang JC. Covalent bonds between protein and DNA. Formation of phosphotyrosine linkage between certain DNA topoisomerases and DNA. J. Biol. Chem. 1980; 255:5560–5565. [PubMed: 6155377]
26. Deweese JE, Osheroff N. The DNA cleavage reaction of topoisomerase II: wolf in sheep's clothing. Nucleic Acids Res. 2008; 37:738–748. [PubMed: 19042970]
27. Baker NM, Rajan R, Mondragon A. Structural studies of type I topoisomerases. Nucleic Acids Res. 2009; 37:693–701. [PubMed: 19106140]
28. Ledesma FC, El Khamisy SF, Zuma MC, Osborn K, Caldecott KW. A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage. Nature. 2009; 461:674–678. [PubMed: 19794497]
29. Grindley NDF, Whiteson KL, Rice PA. Mechanisms of site-specific recombination. Annu. Rev. Biochem. 2006; 75:567–605. [PubMed: 16756503]
30. Ambros V, Baltimore D. Protein is linked to the 5' end of poliovirus RNA by a phosphodiester linkage to tyrosine. J. Biol. Chem. 1978; 253:5263–5266. [PubMed: 209034]
31. Rothberg PG, Harris TJ, Nomoto A, Wimmer E. O^4-(5'-Uridylyl)tyrosine is the bond between the genome-linked protein and the RNA of poliovirus. Proc. Natl. Acad. Sci. USA. 1978; 75:4868–4872. [PubMed: 217003]
32. Hermoso JM, Salas M. Protein p3 is linked to the DNA of phage phi 29 through a phosphoester bond between serine and 5'-dAMP. Proc. Natl. Acad. Sci. USA. 1980; 77:6425–6428. [PubMed: 6779279]
33. Peñalva MA, Salas M. Initiation of phage phi 29 DNA replication in vitro: formation of a covalent complex between the terminal protein, p3, and 5'-dAMP. Proc. Natl. Acad. Sci. USA. 1982; 79:5522–5526. [PubMed: 6813861]
34. Samad A, Carroll RB. The tumor suppressor p53 is bound to RNA by a stable covalent linkage. Mol. Cell. Biol. 1991; 11:1598–1606. [PubMed: 1705009]
35. Baskerville S, Bartel DP. A ribozyme that ligates RNA to protein. Proc. Natl. Acad. Sci. USA. 2002; 99:9154–9159. [PubMed: 12077317]
36. Sachdeva A, Silverman SK. DNA-Catalyzed Serine Side Chain Reactivity and Selectivity. Chem. Commun. 2010; 46:2215–2217.
37. Kost DM, Gerdt JP, Pradeepkumar PI, Silverman SK. Controlling regioselectivity and site-selectivity in RNA ligation by Zn^{2+}-dependent deoxyribozymes that use 2',3'-cyclic phosphate RNA substrates. Org. Biomol. Chem. 2008; 6:4391–4398. [PubMed: 19005599]
38. Flynn-Charlebois A, Prior TK, Hoadley KA, Silverman SK. In Vitro Evolution of an RNA-Cleaving DNA Enzyme into an RNA Ligase Switches the Selectivity from 3'-5' to 2'-5'. J. Am. Chem. Soc. 2003; 125:5346–5350. [PubMed: 12720447]
39. Xiao Y, Allen EC, Silverman SK. Merely two mutations switch a DNA-hydrolyzing deoxyribozyme from heterobimetallic (Zn^{2+}/Mn^{2+}) to monometallic (Zn^{2+}-only) behavior. Chem. Commun. 2011; 47:1749–1751.
40. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 2003; 31:3406–3415. [PubMed: 12824337]
41. Nowakowski J, Shim PJ, Prasad GS, Stout CD, Joyce GF. Crystal structure of an 82-nucleotide RNA-DNA complex formed by the 10–23 DNA enzyme. Nat. Struct. Biol. 1999; 6:151–156. [PubMed: 10048927]

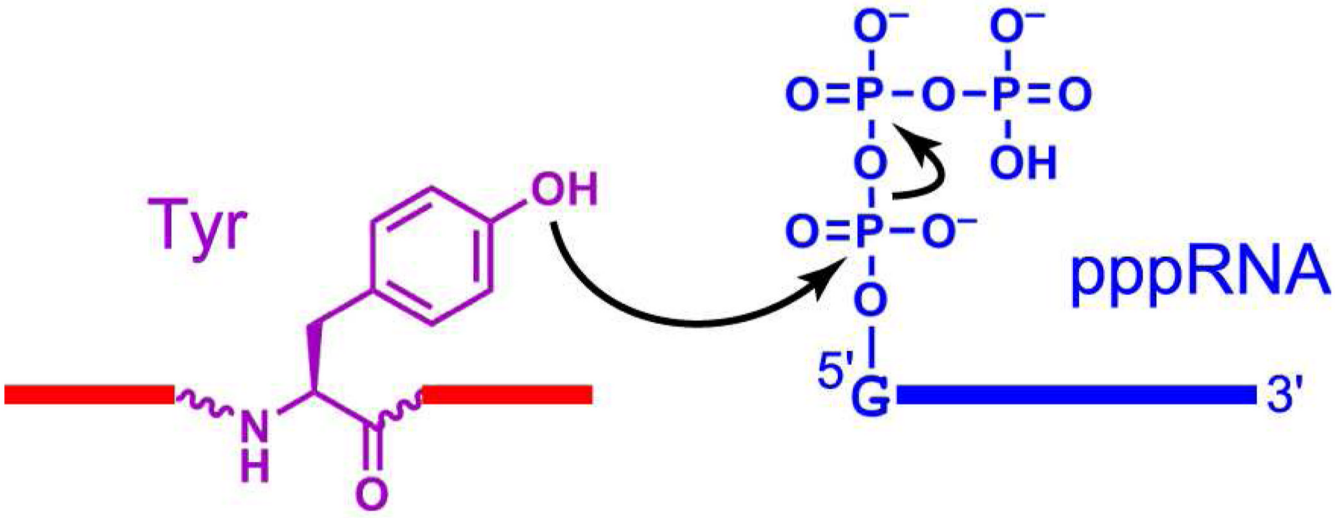

Figure 1.
Formation of a Tyr-RNA nucleopeptide linkage by reaction of a tyrosine side chain with 5′-triphosphorylated RNA.

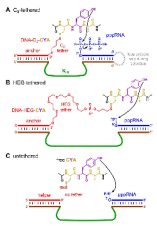

Figure 2.

Varying presentations of the tripeptide substrate to DNA in the open architecture. (A) Presentation with a short three-carbon tether linking the DNA anchor to the CXA tripeptide via a disulfide bond, illustrated with X = Tyr, i.e., CYA tripeptide. The initially random pool is N_{40}, and the loop on the right side is present only during the selection process. (B) Presentation with a longer hexa(ethylene glycol) (HEG) tether to the CYA tripeptide. (C) Presentation with untethered free CYA tripeptide and a DNA helper oligonucleotide in place of the DNA anchor. In all cases, the *N*-acetyl group on the amino terminus of the tripeptide was included to suppress potential amine nucleophilicity.

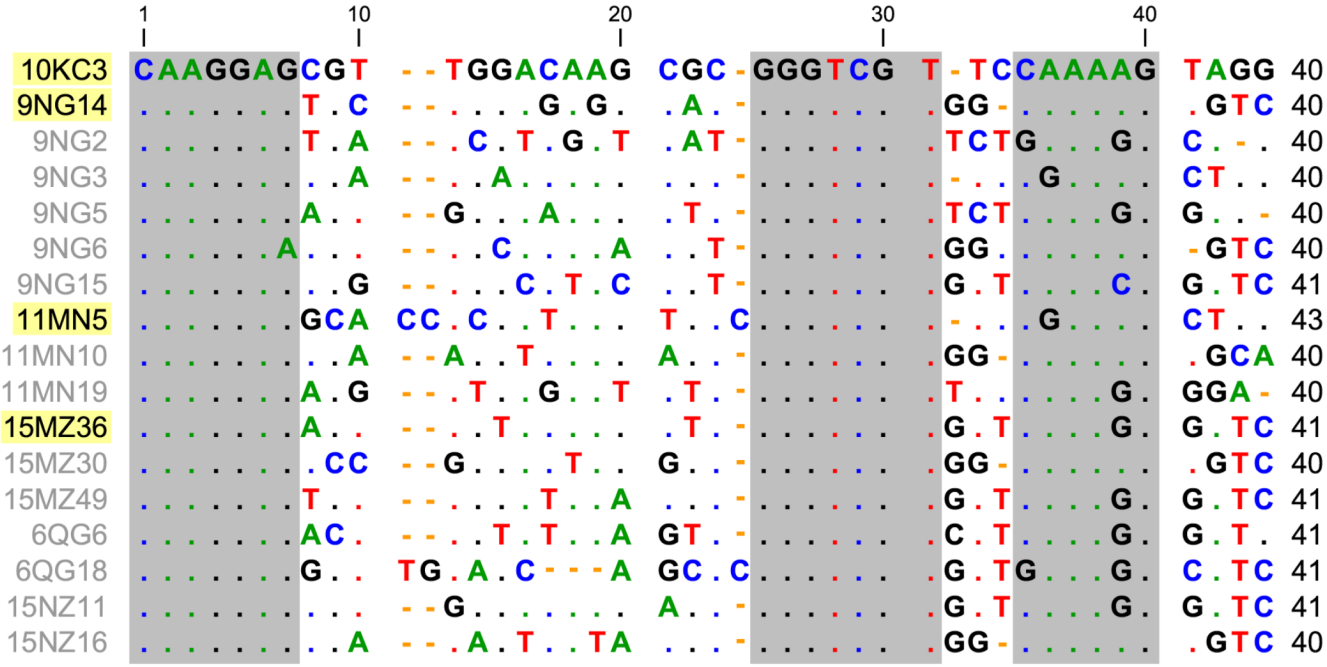

Figure 3.

Aligned sequences of the enzyme regions of the deoxyribozymes identified in this study. The four key deoxyribozymes are highlighted yellow. Note that the enzyme regions of several deoxyribozymes, including 11MN5 and 15MZ36, are longer than the initially random 40 nucleotides, presumably due to insertions by Taq polymerase in one or more selection rounds. The grey boxes denote regions of high sequence conservation. For deoxyribozymes prepared by solid-phase synthesis, the indicated enzyme region was flanked by constant binding arm regions at the 5′-end (`5 -CCGTCGCCATCTCTTC-3 `) and the 3′-end (`5 -ATAGTGAGTCGTATTATCC-3 `). The QG and NZ deoxyribozymes were obtained by reselection starting with 15MZ36 (see text and Figures S6 and S7).

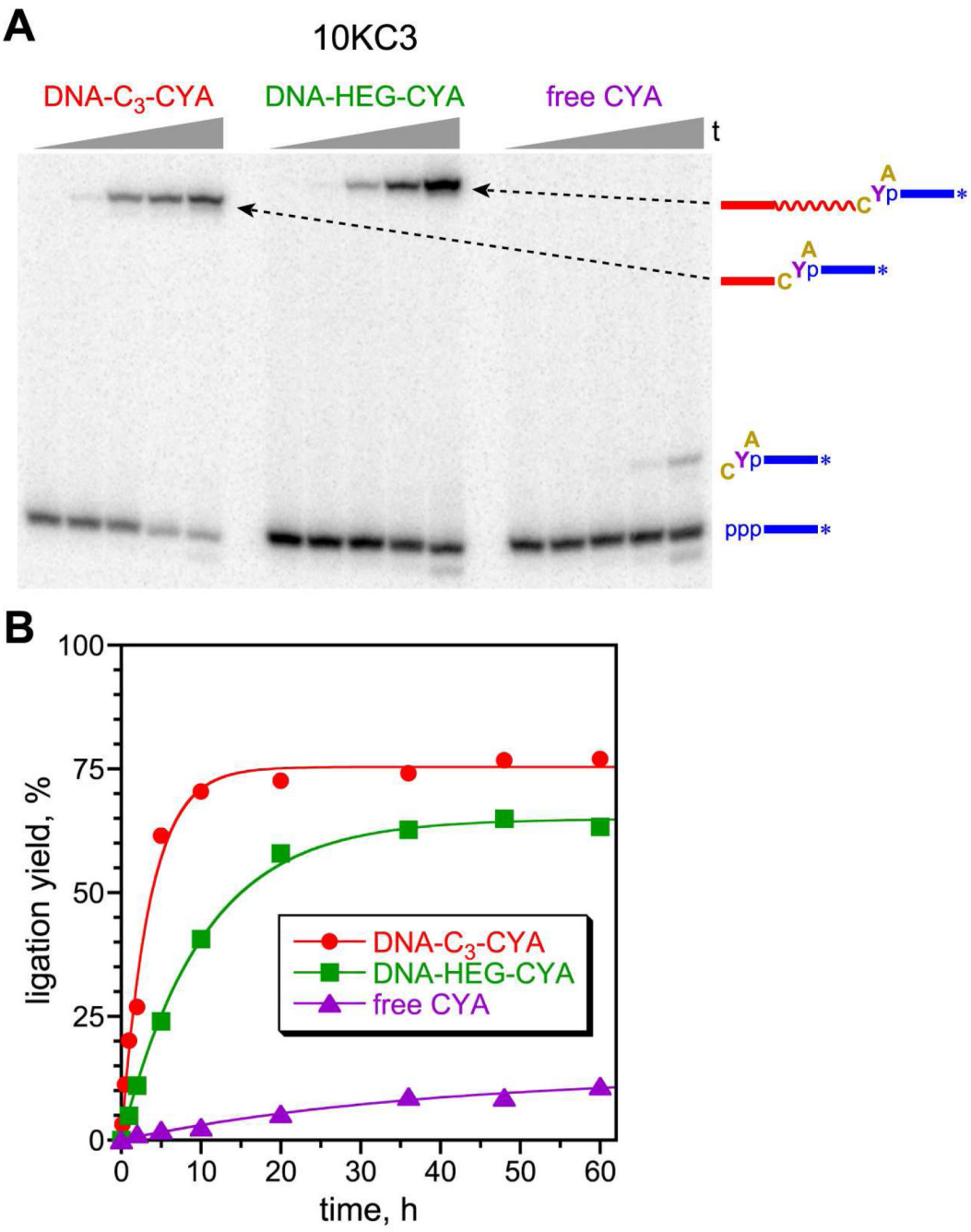

Figure 4.

Assays of the 10KC3 deoxyribozyme, identified by selection from an N_{40} random pool with the DNA-C_3-CYA substrate. (A) PAGE image of assays with the DNA-C_3-CYA, DNA-HEG-CYA, and free CYA substrates. Timepoints at t = 30 s, 10 min, 2, 10 and 60 h using 3′-^{32}P-radiolabeled RNA substrate. (B) Kinetic plot for data from panel A. Data points were fit to first-order kinetics (see Experimental Procedures). See text for k_{obs} values. Similar activities were observed when 5′-^{32}P-radiolabeled DNA-CYA substrates were used (of course, the free CYA cannot readily be radiolabeled).

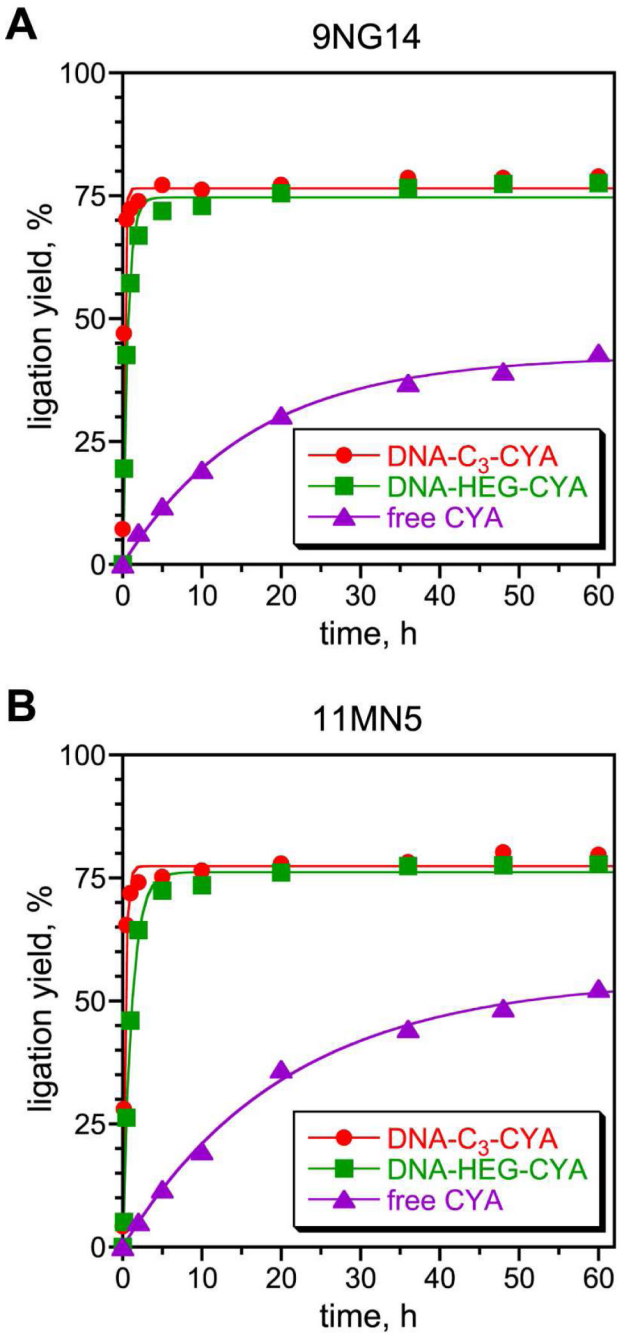

Figure 5.

Kinetic plots for assays of deoxyribozymes from the two follow-up selections. (A) Assays of 9NG14, identified by reselection of 10KC3 with the DNA-C_3-CYA substrate. (B) Assays of 11MN5, identified by selection from N_{40} random pool with the DNA-HEG-CYA substrate.

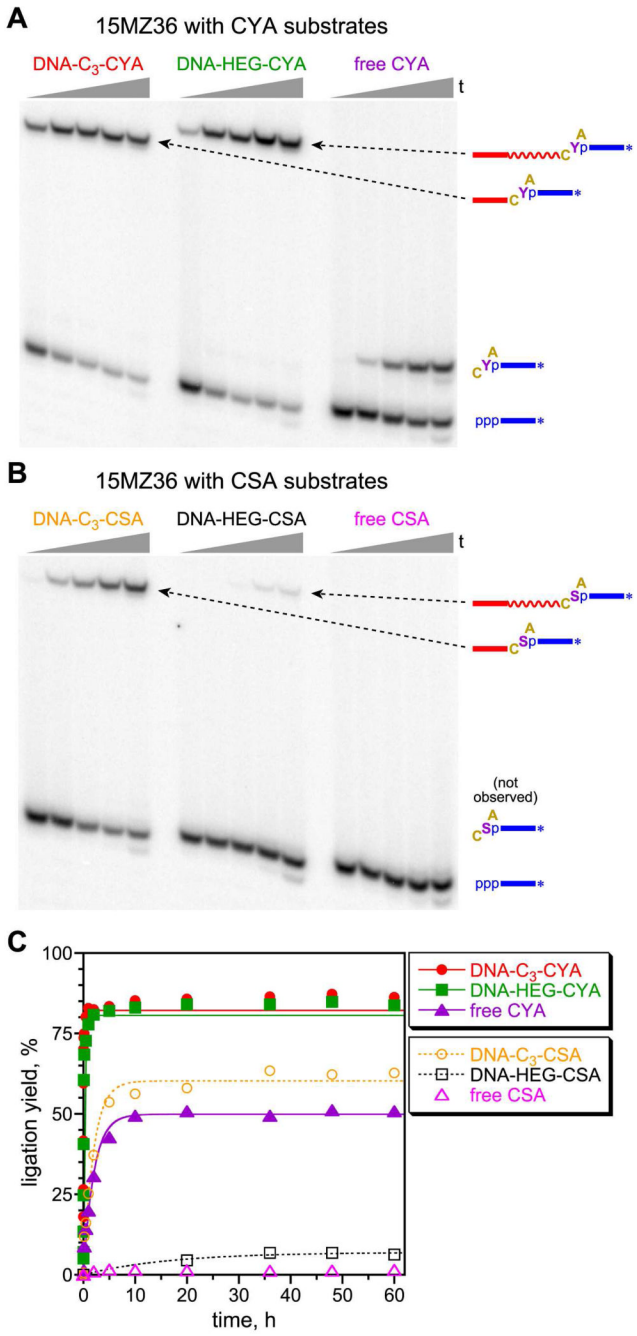

Figure 6.

Assays of the 15MZ36 deoxyribozyme, identified by selection from an N_{40} random pool first with tethered aliphatic hydroxyl substrate and then continuing with the DNA-CSA substrate. (A) PAGE image for assays with the three CYA substrates, with timepoints at t = 30 s, 10 min, 2, 10 and 60 h. (B) PAGE image for reaction of the three analogous CSA substrates (timepoints and labeling as in panel A). (C) Kinetic plots for data from panels A and B.

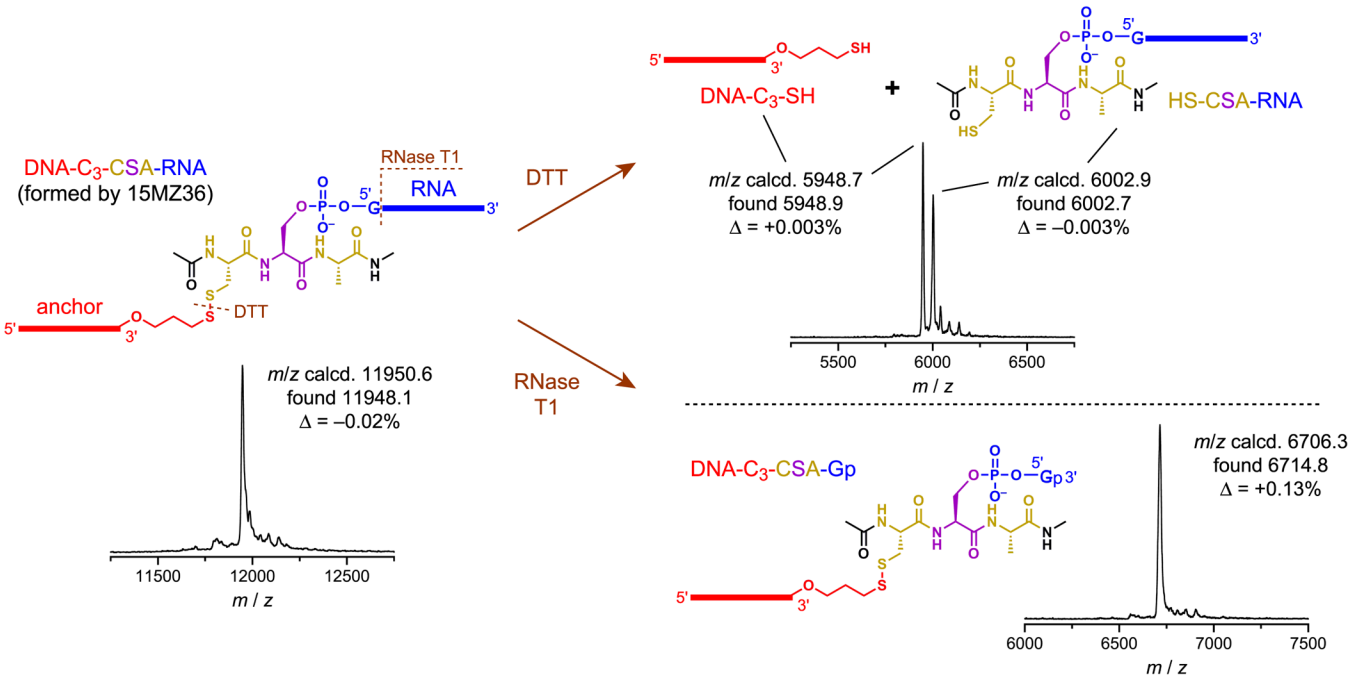

Figure 7.

MALDI mass spectrometry analyses of the 15MZ36 DNA-C_3-CSA–RNA ligation product and its DTT and RNase T1 digestions. Table S1 in the Supporting Information has analogous data for the 10KC3 DNA-C_3-CYA–RNA product, the 11MN5 DNA-HEG-CYA–RNA product, and the 15MZ36 untethered CYA–RNA product.