Deoxyribozymes that recode sequence information

Jeffrey J. Tabor¹, Matthew Levy¹ and Andrew D. Ellington¹,2,*

¹Center for Systems and Synthetic Biology and Institute for Cell and Molecular Biology and ²Department of Chemistry and Biochemistry, University of Texas at Austin, Austin, TX 78712, USA

Received January 25, 2006; Revised and Accepted March 21, 2006

ABSTRACT

Allosteric nucleic acid ligases have been used previously to transform analyte-binding into the formation of oligonucleotide templates that can be amplified and detected. We have engineered binary deoxyribozyme ligases whose two components are brought together by bridging oligonucleotide effectors. The engineered ligases can ‘read’ one sequence and then ‘write’ (by ligation) a separate, distinct sequence, which can in turn be uniquely amplified. The binary deoxyribozymes show great specificity, can discriminate against a small number of mutations in the effector, and can read and recode DNA information with high fidelity even in the presence of excess obscuring genomic DNA. In addition, the binary deoxyribozymes can read non-natural nucleotides and write natural sequence information. The binary deoxyribozyme ligases could potentially be used in a variety of applications, including the detection of single nucleotide polymorphisms in genomic DNA or the identification of short nucleic acids such as microRNAs.

INTRODUCTION

Allosteric nucleic acid enzymes have been generated previously by both design and selection. Enzymes whose activities are modulated by small molecules have been generated by fusing ribozyme domains with allosteric domains [aptamers, yielding aptazymes; (1–10)], while enzymes whose activities are modulated by oligonucleotides have been generated by the strategic insertion of hybridization sites (11–19). In most instances, the modulation of catalytic function has relied on the analyte-dependent re-organization of secondary or tertiary structure. In contrast, a ‘maxizyme’ has been developed in which the modulation of catalytic function relied on the analyte-dependent formation of a specific quartenary structure (20). In this design, two half-ribozymes were brought together by a bridging oligonucleotide in order to form an active hammerhead ribozyme (Figure 1).

We have engineered a DNA ligase to function in a manner similar to the maxizyme, in that the half-deoxyribozymes can be activated by a bridging oligonucleotide to carry out a ligation (rather than a cleavage) reaction. The engineered deoxyribozyme can recode nucleic acid information by ‘reading’ one sequence through hybridization and then ‘writing’ a separate sequence by ligation (Figure 1). Since the newly ligated sequence can be a unique template for amplification, the ligase maxizyme can potentially find use in recoding short, hard-to-detect sequences (such as antisense oligonucleotides or microRNAs) into longer templates that can be readily detected by PCR.

As a proof-of-principle, we show that the ligase maxizyme is highly and specifically activated by cognate oligonucleotides, functions faithfully against a background of genomic DNA, and can even read oligonucleotides containing modified nucleotides.

MATERIALS AND METHODS

All of the ligase maxizymes were composed of two oligonucleotides, which we designate as the left (L) and right (R) subunits. The sequences of the oligonucleotides are as follows: dR8(L), 5’-CGAGACAGGTTGGCCGCAT-TAAAA-3’, dR8(R), 5’-AAAAACGTTGGACCTGTGCTG-TAGTCT-3’; dR3(L), 5’-CGAGACAGGTTGGGAGGTGTGGCCGATTTAAAA-3’; dR3(R), 5’-AAAAACGTTGGCCTTCTACACCAACACGCTTAGTGC-3’. The sequence of the forward effector oligonucleotide, 18N, is 5’-TGTGGGCTTCTTTCCTTACTGCTTAGTC-3’. The sequence of the reverse effector oligonucleotide, 19N, is 5’-GACTAAGCACCCTGTC-3’. There are two substrate oligonucleotides for the ligation reactions. The 5’ substrates are terminated with a 3’-phosphorothioate (PS) group while the 3’ substrates carried a 5’-iodine (I) group. The sequences of the substrates for the ‘forward’ reaction are K3S2, 5’-TACATGTCTATCGATCTGACACAC-PS-3’, and 51.8.c14.m1, 5’-TGCT-TCTTCG-3’. The sequences of the substrates for the ‘reverse’ reaction are K3S2, 5’-TACATGTCTATCGATCTGACACAC-PS-3’, and 51.8.c14.m1, 5’-TGCT-TCTTCG-3’.

*To whom correspondence should be addressed. Tel: 1 512 471 6445; Fax: 1 512 471 7014; Email: andy.ellington@mail.utexas.edu

© The Author 2006. Published by Oxford University Press. All rights reserved.
reaction are 3PS.18N.ol1, 5'-TTTTAATGCCG-PS-3'; and 5I.18N.ol2, 5'-TTTTTTTACACTTACGAACGT-3'.

The non-natural ligase maxizyme is composed of the oligonucleotides dR8b(L), 5'-CGAAGACAGGTTGTGGCC-GCATTiCAAA-3'; and dR8b(R), 5'-AAAiCAAACGTT-GACCTCTGCTTAGTC-3' where iC indicates 5-methyl isocytosine. The sequence of the forward effector for dR8B, 18Nb, is 5'-TTTiGAATGCCGTTTiGTTT-3', where iG indicates isoguanosine.

All natural and non-natural oligonucleotides were synthesized by Integrated DNA Technologies (Coralville, IA). All phosphorothioate- and iodine-bearing oligonucleotide substrates were synthesized in our laboratory as described in Ref. (3). All oligonucleotides were purified by denaturing PAGE prior to use.

**Ligation assays**

All ligation reactions were conducted in a volume of 10 μl at 25°C unless otherwise noted. Reactions were assembled by the addition of 10 pmol of each deoxyribozyme subunit, and 20 pmol of DTT treated phosphorothioate-bearing substrate to 1× reaction buffer [500 mM NaCl, 50 mM Tris–HCl (pH 7.4), 10 mM MgCl₂ and 500 μM DTT]. The reactions were denatured for 3 min at 70°C and then cooled to 25°C at 0.2°C/s. An aliquot of 10 pmol effector oligonucleotide or H₂O (in the effector-independent reactions) was then added, and the reaction was finally initiated by the addition of 20 pmol of the iodine-bearing substrate. Reactions were terminated by the addition of 4 vol of 95% formamide containing bromophenol blue.

Prior to the ligation reaction, the 3'-substrate was radiolabeled using 3'-terminal deoxynucleotidyl transferase (Invitrogen, Carlsbad, CA) and deoxyadenosine 5'-[α-³²P]triphosphate (Amersham Pharmacia Biotech, Piscataway, NJ). After the ligation reaction, the ligated and unligated species were separated on a denaturing 8% polyacrylamide gel containing 7 M urea. Ligation was quantitated using a Phosphorimager (Molecular Dynamics, Sunnyvale, CA) and ImageQuant software.
Reactions were initially linear as a function of time, and pseudo-first order rates were calculated from a best fit line passing through at least three data points taken at <10% total ligation.

Real-time PCR detection

To accommodate real-time PCR-based detection, the forward substrates were modified. KSS2 was modified to contain a primer-binding site at its 5' end, followed by a region whose reverse complement could bind to a TaqMan probe (described below), while oligonucleotide 5l.8.c14.m1 was modified to contain a primer-binding site at the 3' end. The sequence of the resulting 5' substrate 3PS.RTs1 was 5'-GTGACTTCCGTGAACACTATCTAAGCGTTGATCGTGAGTGGCGCATGTAGCAGAGGGACTAACCCAC-PS-3', and the 3' substrate 5l.RTs2 was 5'-1-TGTCTTCGTGCTACATTGGAATCTGACTGCAATCGGTTATT-3'. The enzyme dR8b-(R) was also modified to carry a 3'-amine modification in order to stop nucleotide extension on 5l.RTs2 during PCR. The sequence of the TaqMan probe PLA.TqMnPb was 5'-6FAM-TGTACGTGATGGGACATTGACAAAGGB-HQ1-BHQ1-3' where BHQ1 indicates Black Hole Quencher™ (IDT, Coralville, IA). The sequences for real-time PCR were adapted from Ref. (21).

For real-time PCR detection, ligation reactions were conducted for 5 min as described above and then directly diluted 1:50 into a real-time PCR mix. Real-time PCR was performed on an MJ DNA Engine Opticon (Bio-Rad, Hercules, CA). The reaction conditions were 20 mM Tris–HCl (pH 8.3), 50 mM KCl, 0.2 mM dNTPs, 500 nM 5' and 3' primers, 75 nM PL.A.TqMnPb, 0.5x SmartCycler additive [0.1 mg/ml non-acetylated BSA, 75 mM trehalose and 0.1% Tween-20 in 8.5 mM Tris–HCl (pH 8.0)] and 1.5 U of Platinum Taq DNA polymerase (Invitrogen). All real-time PCR products were carried out in a volume of 50 μl. The samples were heated at 92°C for 5 min then cycled 50 times at 92°C for 1 min, 50°C for 1 min and 72°C for 1 min. The fluorescence intensity was measured at the end of each 72°C extension step. Amplification was controlled for using 1 pm full-length template bearing both primer-binding sites and a TaqMan probe-binding region as in Ref. (22).

RESULTS

Design of a binary deoxyribozyme ligase

We have previously used *in vitro* selection to evolve a deoxyribozyme ligase that can catalyze the formation of internucleotide phosphorothioester linkages (22,23). This enzyme relies upon chemistry pioneered by Xu and Kool (24), in which a 3' phosphorothioate displaces a 5' iodide group, resulting in a phosphorothioester. The deoxyribozyme has a small hairpin stem that functions as a catalytic domain, and two single-stranded arms that can base pair with DNA substrates. At around the same time, the hammerhead ribozyme was re-engineered to act as an allostERIC ribozyme or 'maxizyme' that could be activated by hybridization to a nucleic acid effector (20). We noted that since the structure of the deoxyribozyme ligase was superficially similar to that of the hammerhead ribozyme it should be possible to engineer a nucleic acid sequence-dependent ligase. The ligase maxizyme was generated by fusing two catalytic domains via a common stem structure (Figure 1a).

The resultant binary deoxyribozyme is composed of two half-molecules, and the catalytic ability of either ligase domain should depend upon the association of the half-molecules. DNAs that bridge the substrate-binding arms should also template the association and folding of the half-molecules, and thus could act as sequence-specific effectors of catalysis. Since there are two catalytic domains, either end of the ligase can bind either substrate or effector DNAs. This duality in turn allows one piece of sequence information to act indirectly upon another, via the intervening enzyme (Figure 1b).

**Effector-dependent catalysis**

A number of different deoxyribozyme constructs were designed in which the specificity domains (binding arms) accompanying the two catalytic units were different in order to create distinct reactions on each side the maxizyme. For convenience, we will discuss reactions in the ‘bottom’ direction (ligation on the ‘bottom’ of the construct) and in the ‘top’ direction (ligation on the ‘top’) (Figure 1b). The sequence and stability of the stem structure connecting the top and bottom catalytic domains was then systematically varied.

Initially, the deoxyribozymes were assayed for dependence on an effector oligonucleotide that spanned the hybridizing arms on the top side and activated ligation of substrates that bound the bottom side. Of the designs that were assayed, several proved to be constitutively ‘on’ (required no effector), while others were constitutively ‘off’ (could not be activated by effector). For the remainder, activation varied significantly (Supplementary Data). Two constructs (dR8 and dR3) showed strong effector-dependence (Figure 2).

While it should be possible for the ligase maxizymes to be activated in either the bottom or top direction (i.e. with a full-length effector oligonucleotide at either end), activation was found to occur predominantly in one direction, the bottom direction. Activation required a full-length effector oligonucleotide; the two oligonucleotide substrates which compose the effector did not stabilize the maxizyme structures (data not shown).

The construct dR8 had a relatively fast rate of bottom ligation (0.13 h⁻¹) and showed the largest effector-dependence (1300-fold). This construct, however, showed no detectable ligase activity in the top direction. Construct dR3 showed approximately the same rate of ligation as dR8 in the bottom direction, but also had a higher rate of effector-independent ligation. This resulted in an overall lower effector-dependence (71-fold). Construct dR3 also displayed modest effector-dependent ligation in the top direction (0.0017 h⁻¹; 17-fold activation).

We hypothesized that the differences between ligation rates in the bottom and top directions might be due to the inherent asymmetry of the maxizymes, and that a more symmetrical construct might be activated in both directions. Therefore, all bulge sequences in the internal stem structure of dR3 were mutated to T–T mismatches, creating dR3.1.5. This modification had no significant effect on top ligation, but did increase activation in the bottom direction from 71- to 740-fold. The background activity of dR3.1.5 was also significantly lower.
than dR3, presumably because T–T base-pairing is inherently less favorable than A–A or A–G base pairing.

We attempted to measure substrate turnover by dR3 and dR8. When monitoring ligation as a function of time, no accumulation of ligation products was observed in the absence of effector, consistent with previous results. In the presence of the effector, some accumulation of ligation products was observed, but it was extremely slow, yielding a maximum turnover number of \( \frac{1}{C_24} \) > 100 h (data not shown). The lack of turnover in our system can most likely be attributed to the increased stability of base pairing between the ligated bottom product and the deoxyribozyme (19 bp with the ligation product, as opposed to 8 and 11 bp with the corresponding substrates).

In an attempt to optimize turnover, we shortened the lengths of the hybridizing arms of dR3 and dR8. However, even the deletion of 1 bp completely abolished effector-dependent ligation. Conversely, increasing the number of base pairs between the hybridizing arms and substrates resulted in higher overall rates of catalysis but also less effector-dependent activation, owing to the greater stability of the effector-independent enzyme-substrate complex at room temperature. These results confirm the extraordinary sensitivity of the ligase maxizyme to even small changes in binding energies.

Recognition of non-natural nucleotides

In order to expand the range of information that could be recoded by the ligase maxizyme we incorporated two non-natural nucleotides, 5-methyl-isocytidine (isoMeC) and isoguanosine (isoG) (25) into the effector oligonucleotide and the substrate-binding arms of the maxizyme. A variant of construct dR8 was first synthesized that included isoMeC residues at two positions in its substrate-binding arms. The resultant construct, dR8b, was assayed for its ability to specifically recognize an effector oligonucleotide that contained two complementary isoG substitutions (18Nb). While construct dR8b was highly activated by the non-natural effector 18Nb, it was much less active with effectors bearing residues that were predicted to form mismatches, such as A–isoMeC, C–isoMeC, G–isoMeC and T–isoMeC (Figure 3). The extent of activation correlated well with the previously observed stabilities of the mispairs (26), except for a greater than expected activation by the T–isoMeC mispairs.

To increase the fidelity of recognition against the T–isoMeC mispair, we increased the temperature of the reaction. At 25°C, dR8b showed 5-fold discrimination for its cognate effector versus the effector bearing the T–isoMeC mispair, while at 32.5°C the discrimination increased to 21-fold (Figure 4). The increase in discrimination was accompanied by an expected decline in the rate of ligation, since the originally selected (22) ligase was most active near room temperature. The decrease in activity at higher temperatures was also consistent with the previous finding that the deoxyribozyme was so finely poised that the deletion of even a single base pair led to loss of activation.

Detection of recoded sequence information

While we have demonstrated that a deoxyribozyme can be used to recode sequence information, the utility of this method...
will rely on being able to detect the transformed product. To this end, we attempted to recode short oligonucleotide sequences that could not be readily amplified by PCR into different, longer amplicons. Moreover, we wished to recode and detect one sequence against a background of non-specific sequence information. As a proof-of-principle, we used the ligase maxizyme dR8b that recognizes non-natural base pairs to detect non-natural oligonucleotides obscured within vast amounts of genomic DNA.

Oligonucleotide substrates were prepared for dR8b that, upon ligation, would yield an amplicon that could be readily detected by real-time PCR by extending the 5' and 3' substrates so that they contained primer and TaqMan probe-binding sites (Figure 5a). Ligation reactions were carried out with the oligonucleotide effector, 18Nb, in the presence of a 1000-fold mass excess of genomic DNA (Materials and Methods). When dR8b-mediated ligation was allowed to proceed for only 5 min, the presence of the effector 18Nb in the genomic DNA resulted in a 12 cycle threshold \((C_T)\) real-time PCR advantage for the ligated product relative to amplification in the absence of the effector (Figure 5b). This is a huge \(C_T\) value relative to typical, diagnostically relevant real-time PCR signals (21), and corresponds to a nearly 4100-fold activation of the non-natural enzyme by its cognate effector. Importantly, the presence of 1000-fold excess genomic DNA also decreased the extent of effector-independent ligation (5 cycles or 32-fold). This effect is likely due to genomic DNA sequestering either substrate or half-enzyme molecules (by non-specific hybridization) or magnesium (by chelation). Either mechanism would decrease effective reactant concentrations, increase the overall stringency of the reaction and thereby increase reaction specificity.

Figure 3. Expanded base pairing in an allosteric deoxyribozyme. Construct dR8 was re-engineered to carry a non-natural nucleotide, 5-methyl-isocytidine \((\text{i}^\text{Me}\text{C})\), at two positions (blue), generating dR8b. The effector was then synthesized with compensatory isoguanosine \((\text{iG})\) substitutions, generating 18Nb. The adjacent gel shows the ligation activity of dR8b with different substrates. The lower band on the gel is a radiolabeled, unligated substrate, while the higher band is the ligated product. The lane labeled ‘0’ is time 0 of the reaction and all subsequent lanes are 3.5 h reactions with effector variants engineered to make the indicated base-pairings. The extent of ligation is normalized to the most active pairing, iG–iC.

Figure 5a. Ligation reactions were carried out with the oligonucleotide effector, 18Nb, in the presence of a 1000-fold mass excess of genomic DNA (Materials and Methods). When dR8b-mediated ligation was allowed to proceed for only 5 min, the presence of the effector 18Nb in the genomic DNA resulted in a 12 cycle threshold \((C_T)\) real-time PCR advantage for the ligated product relative to amplification in the absence of the effector (Figure 5b). This is a huge \(C_T\) value relative to typical, diagnostically relevant real-time PCR signals (21), and corresponds to a nearly 4100-fold activation of the non-natural enzyme by its cognate effector. Importantly, the presence of 1000-fold excess genomic DNA also decreased the extent of effector-independent ligation (5 cycles or 32-fold). This effect is likely due to genomic DNA sequestering either substrate or half-enzyme molecules (by non-specific hybridization) or magnesium (by chelation). Either mechanism would decrease effective reactant concentrations, increase the overall stringency of the reaction and thereby increase reaction specificity.
DISCUSSION

The ligase maxizyme is a catalyst that recodes nucleic acid sequence information: it ‘reads’ one sequence and ‘writes’ another. There are a variety of potential applications for such a simple reagent. For example, in many instances the detection of specific alleles by PCR suffers from background problems due to primer binding to closely related alleles or to other related sequences in genomic DNA (27). The ligase maxizyme could transform individual alleles into amplicons that were unrelated to most sequences in a genome, and thereby reduce problems associated with background amplification.

Nucleic acid enzymes have been engineered previously to detect biologically relevant nucleic acids such as microRNAs and regulatory untranslated regions (10,18,19). As the ligase maxizyme converts short oligonucleotide sequences into longer amplicons, it could also prove useful as a tool for detecting therapeutic nucleic acids such as antisense, aptamers or siRNA molecules by real-time PCR. Such a technique might be especially useful if the therapeutic contained modified or non-standard nucleotides that were difficult to detect or amplify by normal methods. In addition, short, natural sequences, such as microRNA molecules that are usually difficult to detect, could be transformed into amplicons and then read by PCR.

The ligase maxizyme can also be viewed as a simple machine or ‘part’ that would have a unique function in the nascent fields of DNA computation and synthetic biology. To date most of the various implementations of nucleic acid computation have involved what can be described as ‘hybridization logic,’ assembling and analyzing answers based upon pre-encoded patterns of hybridization (14–17, 28–32). As might be expected from the error-prone nature of DNA hybridization and polymerization, large-scale nucleic acid computation based on hybridization logic is inherently infeasible (33). For example, the best nucleic acid computer built to date demonstrated an error rate of ~1 in 2500 (29), far greater than even a poor silicon device. However, the use of nucleic acid enzymes as ‘silicomimetic’ devices that can interface with biology and make decisions has been demonstrated brilliantly by Stojanovic and Stefanovic, who encoded an algorithm for playing Tic-Tac-Toe into a series of deoxyribozymes (16). More recently, Shapiro and co-workers have extended these ideas to the creation of self-diagnosing and self-actuating nucleic acid therapeutics (34). A deoxyribozyme that can transform information might be extremely useful in the design and implementation of such therapeutic nanomachines, especially since the unidirectional activation observed with most of the constructs examined would allow ‘reading’ and ‘writing’ to be uniquely specified.

Irrespective of the application, this is a further example of how allosteric nucleic acid enzymes can be generated by engineering at the level of secondary structure, and stands as one of the first demonstrations that DNA can act not only as a molecule to carry information but as a machine that can recode information.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

Figure 5. Real-time PCR detection of the ligation product formed by dR8b. (a) Each of the substrates for dR8b was extended to carry one primer-binding site. When ligated, the product becomes a template for PCR. The complementary DNA strand (cDNA, gray) formed by the first primer extension bears a site complementary to a Taqman probe. When the cDNA strand is replicated the Taqman probe is digested, liberating the fluorescent molecule from the quencher. The fluorescent readout is quantified by real-time PCR. (b) Quantification of the ligation product of dR8b by real-time PCR. dR8b was allowed to react with the modified substrates for 5 min and the product was detected as in (A). Minus and plus indicate the absence and presence of the effector molecule, 18 Nb in the reaction. C T shift indicates the decrease in number of cycles required to reach exponential amplification relative to the effector-independent background reaction plus 1000x excess DNA (left-most bar), which had an absolute C T value of 37 ± 0.15. Error bars represent 2 SDs.
ACKNOWLEDGEMENTS
This work was supported by the United States Department of Defense, University of Florida (UF03010), the National Institute of Health and the Texas Higher Education Coordinating Board, TDT-003658-0611-2003. Funding to pay the Open Access publication charges for this article was provided by Texas Higher Education Coordinating Board.

Conflict of interest statement. None declared.

REFERENCES