A FLUORESCENCE SIGNALLING DNA ENZYME

CREATION AND CHARACTERIZATION

OF

A FLUORESCENCE SIGNALLING DNA ENZYME

By

SHIRLEY HSIN-JU MEI, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

In Partial Fulfillment of the Requirements

for the Degree

Master of Science

McMaster University

MASTER OF SCIENCE (2003) McMaster University

(Biochemistry) Hamilton, Ontario

TITLE: Creation and Characterization of a Fluorescence DNA Enzyme

AUTHOR: Shirley Hsin-Ju Mei, B.Sc. (Simon Fraser University)

SUPERVISOR: Professor Yingfu Li

Number of PAGES: viii, 111

ABSTRACT

In vitro selection has been widely used to isolate single-stranded DNA molecules from large random-sequence pools that are able to perform a desired catalytic reaction, or bind to a target molecule. Using this technique, we have created a DNA enzyme, named DET22-18, which has a uniquely linked chemical catalysis/real-time signaling capability. It is a true enzyme with a k_{cat} of ~7 min^{-1} – the second fastest rate ever reported for a DNA enzyme. The DNA enzyme cleaves a substrate containing a single ribonucleotide linkage embedded in a DNA chain and sandwiched between a fluorophore-labeled deoxyribonucleotide and a quencher-modified deoxyribonucleotide. The ability of DET22-18 to generate a large fluorescence signal provides a useful tool to engineer potential allosteric deoxyribozyme biosensors for real-time detection of important biological targets. To provide a proof of concept that a reporter system can be built from the above signaling DNA enzyme, the *cis*-acting version of this enzyme, DEC22-18A, was engineered into an allosterically regulated deoxyribozyme biosensor that can report ATP. A preliminary investigation was also conducted to determine a possible secondary structure of the DNA enzyme. This study lays a foundation for pursuing novel signaling DNA enzymes for biological detection directed applications.

AKNOWLEDGEMENTS

I would like to thank my project director/producer, Yingfu, for his guidance, insight, and endless enthusiasm. It is truly my honour to become his first M.Sc. graduate. Thanks to my committee members, John Brennan and Eric Brown, for their constructive comments and ample advice, and Murray Junop for being the best internal examiner a girl can have.

I would like to express my gratitude to the entire Li lab, but special thanks goes to Zhongjie for his invaluable contribution to every aspect of this project (and free organic chemistry tutorials), and Erick (a.k.a. Ericka the helper elf) for his active role in the final chapter of this project. I also have to thank my desk neighbor Sylva (a.k.a. my long lost half-sister) for her always-helpful big-sister advices, my knowledge "data" base J.C., my lunch buddies Simon, Rachael, and Jill (for making an ordinary FUN day in the lab, absolutely extraordinary), and Ken (a.k.a. the grand master) for helping me survive through the entire thesis writing process with his enormous support and patience.

I also want to thank my friend Margaret and people in Gupta's lab, Emma, Tim, and Jae (specifically for Figure 10) for being the coolest and most genius people in biochemistry (after the Li Lab, of course).

Finally, I would like to dedicate this thesis to my family, to my mom Nancy, and brother Alven. Without their unlimited support and encouragement, I could not have accomplished this work. Last, but certainly not least, I wish to express my eternal gratitude to the man who inspired me to pursue a career in science, my father James.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 Background

In nature, double-stranded DNA is best known as the material that stores and transmits genetic information from generation to generation. Compared with RNA and proteins, double-stranded DNA with its inherent structural and chemical constraints, is not capable of forming complex tertiary structures for executing intricate enzymatic reactions. The sugar-phosphate backbone of DNA is negatively charged and highly flexible, while the limited number of side chains restricts the formation of a compact catalytic core due to the lack of structural interconnections (Peracchi, 2000). At least in the case of RNA, the extra 2'-hydroxyl group permits the formation of more sophisticated structures to support catalysis, as seen in naturally occurring ribozymes (Cate et al., 1996; Hermann and Patel, 1999). Nonetheless, using the technique of *in vitro* selection, Breaker and Joyce successfully derived a DNA enzyme in 1994 (Breaker and Joyce, 1994).

In vitro selection has now been widely used to isolate single-stranded DNA molecules that are able to perform a desired catalytic reaction (Emilsson and Breaker, 2002) or bind to a target molecule (Gold, 1995), from large random-sequence pools. The DNA molecules which can catalyze chemical reactions are referred to as DNA enzymes, deoxyribozymes or even simply DNAzymes, while the ones that can interact specifically with a given target molecule (or ligand) are referred to as aptamers. A scheme

demonstrating *in vitro* selection concept is shown in Figure 1. The total number of deoxyribozyme or aptamer molecules in an original DNA pool is usually very small. In order to perform the identification and isolation of these rare DNA molecules, repetitive selection-amplification steps have to be performed. After screening for catalysis or binding, the selected DNA molecules will be amplified by the polymerase chain reaction (PCR) and used for the next cycle of selection and amplification. This repetitive effort yields an increase in the new population of successful molecules derived from the previous pool. Typically, several rounds are required before the population size of the desirable DNA molecules begins to peak. In the end, catalytic or binding molecules will be isolated. The method of *in vitro* selection has been used successfully in isolating DNA enzymes that cleave RNA (Breaker and Joyce, 1994; Santoro and Joyce, 1997), *N*-glycosidic bonds (Sheppard el al., 2000), or DNA (Carmi et al., 1996); modify (Li and Breaker, 1999a; Li and Breaker, 2000) and ligate DNA (Cuenoud and Szostak, 1995); and metalate porphyrin rings (Li and Sen, 1996), among other reactions.

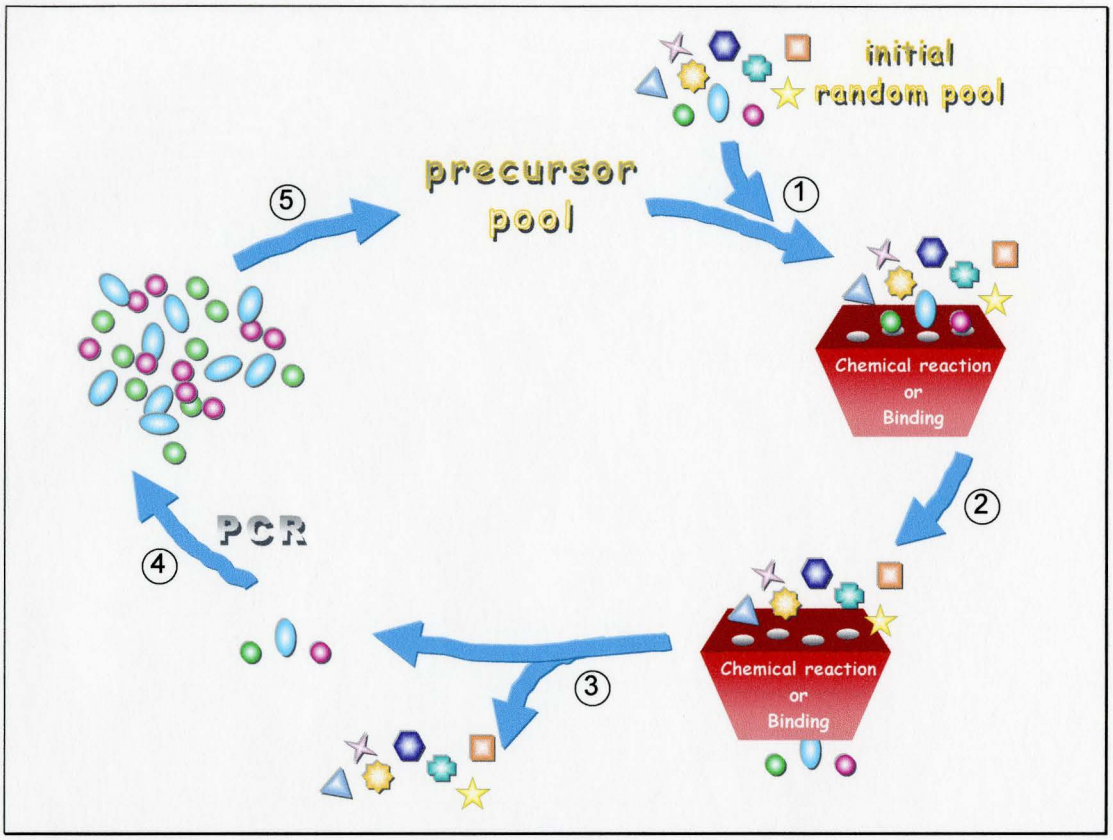


Figure 1: A general scheme of *in vitro* selection. A population of random-sequence pool is (1) subjected to a given task (either a chemical reaction for enzymes or a target binding for aptamers). The molecules that (2) can perform the given task are selected, while the molecules that fail are (3) eliminated. The selected molecules are then (4) enriched by PCR amplification to (5) generate a new precursor pool to initiate the next round of selection. The condition of the given task usually becomes increasingly stringent as the rounds of selection progress. The selection is repeated until the activity (catalysis for enzymes or binding for aptamers) can be detected through a biochemical assay (Emilsson and Breaker, 2002).

Aside from catalyzing an ever-widening array of chemical reactions, DNA enzymes have proven their ability to perform efficient catalysis (Santoro and Joyce, 1997; Santoro et al., 2000) and to provide significant rate enhancement (Li and Breaker, 1999a; Li and Breaker, 2000). With respect to the efficiency, an RNA-cleaving DNA enzyme (referred as "10-23" by the authors) was demonstrated to catalyze a multiple-turnover RNA cleavage (via RNA transesterification) with a k_{cat} of ~10 min^{-1} in the presence of the cofactor Mn^{2+} (Santoro and Joyce, 1997; Santoro and Joyce, 1998). When compared with RNA or protein enzymes that catalyze the same type of reaction, the 10-23 enzyme exhibits a superior catalytic efficiency (catalytic efficiency = specificity constant = k_{cat} / K_M) at 4.5 x 10^9 $min^{-1} \cdot M^{-1}$, compared with its protein rival RNase A at 1.4 x 10^8 $min^{-1} \cdot M^{-1}$. Thus, the 10-23 DNA enzyme is categorized as a perfect enzyme since its catalytic efficiency falls within the range of 10^8 to 10^{10} $sec^{-1} \cdot M^{-1}$ (Li and Breaker, 1999b). With respect to the rate enhancement, DNA enzymes with DNA phosphorylation (Li and Breaker, 1999a) and DNA adenylation (Li and Breaker, 2000) activities were shown to generate rate enhancements (k_{cat} / k_{uncat}) approximately 10^9-fold over the rate of spontaneous ATP hydrolysis, comparable to a self-phosphorylating ribozyme (Li and Breaker, 1999b). As a new member of the enzyme family, single-stranded DNA has certainly established itself to be as proficient a catalyst as RNA or proteins.

Protein enzymes are known to recruit cofactors to expand the spectrum of reactions they catalyze, to increase the catalytic efficiency, and to enhance the reactivity for complicated reactions. With a more limited chemical repertoire, DNA appears to be

in greater need of assistance from cofactors. Among those cofactors, divalent metal ions are certainly one of the most important groups necessary for the folding and transition state stabilization of most DNA enzymes. While nearly all naturally occurring ribozymes tend to utilize alkaline-earth metal ions for catalysis, the majority of DNA enzymes seem to use both alkaline earth metal and transition metal ions (Lu, 2002). However, transition metal ions appear to promote some DNA-enzyme-catalyzed reactions more efficiently than alkaline earth metal ions, as in the case of RNA cleavage. One example is the enzyme 10-23 (Santoro and Joyce, 1997) that catalyzes its RNA-cleaving reaction ~3-fold higher in the presence of saturating Mn^{2+} rather than Mg^{2+}. The divalent metal ion has been postulated to act either as a general base to support metal-assisted deprotonation of the 2′-hydroxyl adjacent to the cleavage site, or as a Lewis acid to coordinate directly to the 2'-hydroxyl to enhance its acidity (Figure 2) (Santoro and Joyce, 1998). Transition metal ions are generally better Lewis acids and their hydroxides are more basic than alkali and alkaline earth metal hydroxides (Li et al., 2000). These properties make the transition metal ion a seemly candidate for assisting in the deprotonation event that has been proposed as part of the RNA cleavage mechanism for the hammerhead ribozyme (Blount and Uhlenbeck, 2002), and most RNA-cleaving DNA enzymes (Emilsson and Breaker, 2002). There have been a few cases of divalent-metal ion-independent DNA enzymes being reported which also catalyze RNA cleavage (Geyer and Sen, 1997; Faulhammer and Famulok, 1997). The isolation of cofactor-independent DNA enzymes indicates that DNA alone is sufficient in supplying the chemical groups responsible for

catalysis. However, those enzymes are approximately 100-fold slower than those that depend on divalent metal ions (Sen and Geyer, 1998).

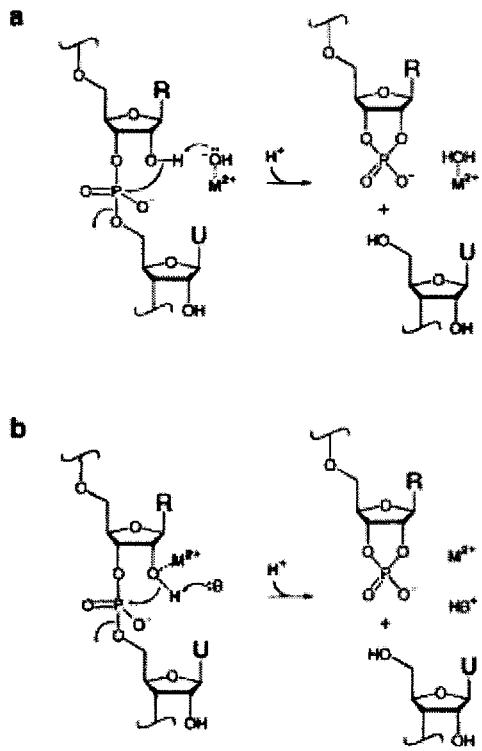


Figure 2: Two divalent-metal assisted mechanisms proposed for catalysis of RNA-cleaving DNA enzymes. (A) A divalent metal hydroxide functions as a general base. (B) A divalent metal ion functions as a Lewis acid (Santoro and Joyce, 1998).

1.2 Applications of DNA Enzymes and Aptamers

To date, various DNA enzymes have been utilized in different biotechnological and pharmaceutical applications (Breaker, 2000; Emilsson and Breaker, 2002; Lu, 2002). The RNA-cleaving activity of DNA enzymes has been exploited as endonucleases for selective target destruction of messenger RNAs (mRNAs) both *in vitro* and *in vivo* (Emilsson and Breaker, 2002). Deoxyribozyme-based logic gates were constructed as one of the first few steps toward the engineering of molecular-scale computation elements (Stojanovic et al., 2002; Stojanovic and Stefanovic, 2003). The immobilized deoxyribozyme was used as a tool to distinguish higher-ordered structures of RNAs (Okumoto et al., 2002). Finally, several groups have attempted to use DNA enzymes as analytical tools, considering the many advantages that DNA possesses over RNA or protein.

Both DNA enzymes and aptamers have great potential to be engineered as biosensors for analytical purposes. A biosensor is an analytical tool that generates a signal following the detection of a specific target. There are two classes of biosensors: a qualitative biosensor simply identifies the target of interest, while a quantitative biosensor provides information regarding the concentration of a given analyte in a sample (Brennan, 2000). DNA is known to be exceedingly stable and relatively cost-effective when compared with RNA or proteins. There have been reports which demonstrated that some DNA-based sensors can withstand extreme experimental conditions, such as autoclaving or sonication, and still remain active for up to one year (Piunno et al., 1995). Last but not the least, DNA is amenable to molecular design by rational design and *in*

vitro selection methods. Both methods undoubtedly facilitate the process of biosensor engineering, towards the development of DNA microarrays.

Efforts have been made in the development of DNA-based biosensors that can directly transduce molecular recognition to fluorescent signals. In sensor development, one needs to convert the interaction between the analyte and the enzyme into a measurable signal. By directly coupling a fluorescence-reporting group onto a biosensor system, no additional reagent has to be added (Brennan, 1999). Several advantages for a fluorescence-based system include rapid response time, real-time detection capability, high selectivity, and the potential to be developed into multi-analyte sensors utilizing optic arrays (Brennan, 2000). Signaling aptamers capable of generating fluorescence signals in the presence of ATP (Jhaveri et al., 2000; Nutiu and Li, 2003), thrombin (Potyrailo et al., 1998; Nutiu and Li, 2003; Barezovski et al., 2003), or IgE (German et al., 1998) have been reported. Some of these signaling aptamers behave as quantitative biosensors that can monitor the concentration of an analyte in solution, without binding or washing steps (Berezovski et al., 2003; Nutiu and Li, 2003).

Similarly, sensor molecules made of DNA enzymes have been described in a few cases (Li and Liu, 2000; Stojanovic et al., 2000; Todd et al., 2000; Stojanovic et al., 2001). Thus far, all of the reported systems have adopted the concept of fluorescence-resonance-energy-transfer (FRET) and can be grouped into two general designs. One design involves labeling one fluorophore at the 5'-end and one quencher at the 3'-end of the substrate (Figure 3A). The fluorophore and quencher pair are placed at least 15 nucleotides apart to avoid interfering with enzymatic catalysis. The cleavage of the

ribonucleotide in the middle of the substrate results in the separation of the fluorophore from quencher and thus generates a fluorescence enhancement. The second strategy involves labeling one fluorophore at the 5'-end of the substrate and one quencher at the 3'-end of the enzyme (Figure 3B). When the substrate strand is hybridized to the enzyme strand, the fluorescence is quenched. Upon RNA cleavage, the substrate and enzyme strands are denatured, separating the fluorophore away from the quencher. There are inherent flaws for both designs. With the first design, the relatively long distance (15 to 29 nucleotides) between the fluorophore and quencher pair results in less efficient quenching, due to the distance dependency of FRET. With the second design, false signaling is possible if less than 100% of the substrate strands are annealed to the enzyme strands. Furthermore, all of those catalytic DNA sensors were constructed based on existing DNA enzymes, neither the fluorescence enhancement nor catalytic activity were optimized.

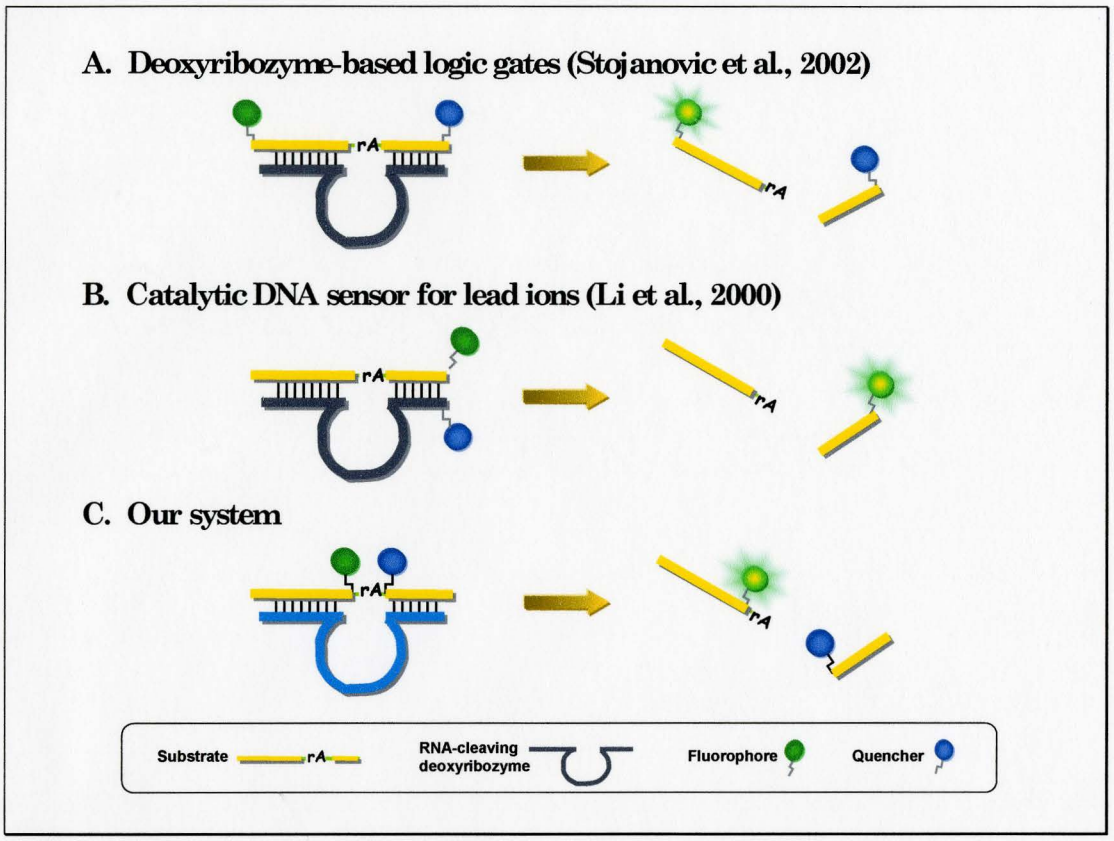


Figure 3: Sensor molecules made of DNA enzymes. A general RNA-cleaving deoxyribozyme is illustrated with the proposed secondary structure of enzyme "10-23" (Santoro et al., 1997). (A) The substrate is labeled with one fluorophore at the 5'-end and one quencher at the 3'-end. The fluorophore and quencher pair is placed at least 15 nucleotides apart to avoid interfering with enzyme catalysis. The cleavage of the ribonucleotide in the middle of the substrate results in the separation of fluorophore from quencher and thus generates fluorescence enhancement. (B) The 5'-end of the substrate is labeled with one fluorophore, while the 3'-end of the enzyme is labeled with one quencher. When the substrate strand is hybridized to the enzyme strand, the fluorescence is quenched. Upon RNA cleavage, the substrate and the enzyme strands are denatured, separating fluorophore away from quencher. (C) In our system, the substrate is labeled in a way that the ribonucleotide to be cleaved is sandwiched by a fluorophore-modified nucleotide and a quencher-modified nucleotide.

To achieve a system with low background fluorescence and minimize possibility of false signaling, we designed our substrate in such a way that the ribonucleotide to be cleaved was sandwiched between a fluorophore-modified nucleotide and a quencher-modified nucleotide (Figure 3C). The close arrangement of fluorophore and quencher relative to the ribonucleotide implied that the existing RNA-cleaving DNA enzymes, which had been selected to cleave an RNA junction without nearby modified groups, might not be able to cleave our substrate. This obstacle, however, can be easily overcome with the technique of *in vitro* selection to isolate new enzymes capable of cleaving our substrate.

Once we obtained an enzyme that could catalyze RNA cleavage in our modified substrate, we could proceed to the next step and engineer our biosensor system with rational design. The basis of such a system is a DNA enzyme under allosteric regulation. In this system, the DNA enzyme only catalyzes a desired reaction under the influence of a specific target, or the enzyme remains otherwise inactive. Tang and Breaker first demonstrated an allosteric ribozyme by appending the ATP aptamer (Huizenga and Szostak, 1995) to the hammerhead ribozyme (Tang and Breaker, 1997). Since then, various groups have developed allosterically regulated ribozymes through rational design, *in vitro* selection, or a combination of both approaches (reviewed in Soukup and Breaker, 1999; Soukup and Breaker, 2001; Breaker, 2002; Silverman, 2003). Allosteric regulation has also been achieved with a DNA ligase enzyme that can be activated by ATP (Levy and Ellington, 2002). By coupling allosteric regulation with a fluorescence signalling

enzyme, we can create a reporter system that catalyzes substrate conversion to generate fluorescence only when activated by the target (Figure 4).

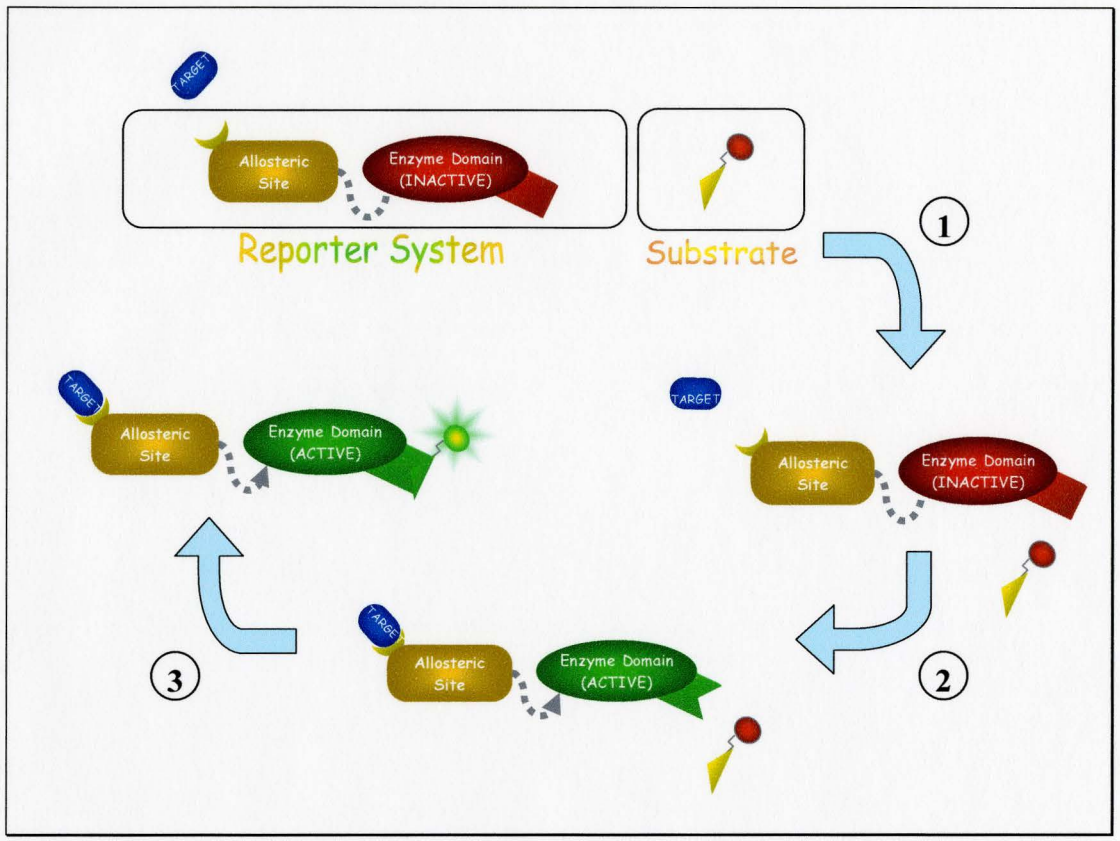


Figure 4: A DNA-enzyme based reporter system. The reporter system consists of an allosteric site (yellow) covalently linked to an enzyme domain (red when inactive), and a fluorescence generating substrate. The substrate does not emit fluorescence when not processed by the enzyme. The enzyme domain remains inactive in the absence of target. When the target is (1) added to the system, it (2) binds specifically to the allosteric site. This binding induces a conformational change in the adjoining enzyme domain and activates the enzyme (enzyme domain turns green). The activated enzyme then (3) catalyzes the conversion of substrate to its fluorescence generating state. Even though the substrate here is shown separately from the reporter system, it can also be covalently linked to the reporter system.

1.3 Project Objective

The goal of this project was to use the strategy of *in vitro* selection to isolate efficient DNA enzymes that could couple RNA cleavage with fluorescence signaling. To ensure the isolation of efficient catalysts, the selection was carried out in the presence of divalent metal ions. The substrate used in this selection consisted of a ribonucleotide sandwiched between a fluorophore and a quencher, and thus, cleavage of the ribonucleotide junction resulted in the separation of the fluorophore from the quencher. This system can provide the lowest background fluorescence under any condition (due to the close proximity of fluorophore and quencher), and generate large fluorescence enhancement upon RNA cleavage. The first section of my thesis describes the isolation and optimization of this fluorescence-signalling, RNA-cleaving DNA enzyme (referred to as DEC22-18). A proof-of-concept system that converted this DEC22-18 enzyme into an allosterically regulated biosensor is demonstrated as well. Secondly, an attempt to optimize this allosteric DNA enzyme is described. Finally, a comprehensive mutational analysis for DEC22-18 was carried out to derive a possible secondary structure model.

CHAPTER TWO

MATERIALS & METHODS

2.1 Materials

Standard oligonucleotides were prepared by automated DNA synthesis using cyanoethylphosphoramide chemistry (Keck Biotechnology Resource Laboratory, Yale University; Central Facility, McMaster University; Sigma Genosys). Random-sequence DNA libraries were synthesized using an equimolar mixture of the four standard phosphoramidites (unless indicated otherwise). DNA oligonucleotides were purified by 10% preparative denaturing (8 M urea) polyacrylamide gel electrophoresis (PAGE), and their concentration were determined spectroscopically and calculated using the Biopolymer Calculator program (http://paris.chem.yale.edu/extinct.frames.html.).

Fluorescein and 4-(4-dimethylaminophenylazo)benzoic acid (DABCYL) labels were incorporated into the DNA substrate during automated DNA synthesis using Fluorescein-dT amidite and DABCYL-dT amidite (Glen Research, Sterling, Virginia). The adenine ribonucleotide linkage was introduced during solid-state synthesis using A-TOM-CE Phosphoramidite (Glen Research). Fluorescein- and DABCYL-modified oligonucleotides were purified by reverse phase liquid chromatography (HPLC) performed on a Beckman Coulter HPLC System Gold with a 168 diode array detector. The HPLC column used was an Agilent Zorbzx ODS C18 column with dimensions of 4.6 mm × 250 mm and a 5-μm bead diameter. Elution was achieved using a two-solvent system with solvent A being 0.1 M triethylammonium acetate (TEAA, pH 6.5) and

solvent B being pure acetonitrile. The best separation results were achieved using a nonlinear elution gradient (0% B for 5 minutes, 10% B to 30% B over 95 minutes) at a flow rate of 0.5mL/min. The main peak was found to have very strong absorption at both 260 and 491 nm.

The TOM protective group on the 2′-hydroxyl group of the RNA linkage was removed by incubation with 150 μL of 1M tetrabutylammonium fluoride (TBAF) in THF at 60°C with shaking for 6 hours, followed by the addition of 250 μL of 100 mM Tris (pH 8.3) and further incubation with shaking for 30 min at 37°C. The DNA was recovered using ethanol precipitation and dissolved in water containing 0.01% SDS; the tetrabutylammonium salt was removed by centrifugation using a spin column (Nanosep 3K Omega, Pall Corp., Ann Arbor, Michigan).

Nucleotide 5′-triphosphates, [γ-^{32}P]ATP and [α-^{32}P]dGTP were purchased from Amersham Pharmacia. *Taq* DNA polymerase, T4 DNA ligase, and T4 polynucleotide kinase (PNK) were purchased from MBI Fermentase. All other chemical reagents were purchased from Sigma.

2.2 *In Vitro* Selection Procedures

5′-phosphorylated, gel-purified, 86-nt random-sequence DNA L1 (300 pmol) was mixed in an equimolar ratio with template T1 and acceptor A1 (all sequences shown in Figure 6A), heated to 90°C for 30 seconds, cooled to room temperature, and combined with ligase buffer and T4 DNA ligase for DNA ligation to introduce the modified DNA domain. The ligation mixture (50 μL) contained 50 mM Tris-HCl (pH 7.8 at 23°C), 40

mM NaCl, 10 mM $MgCl_2$, 1 mg/mL BSA, 0.5 mM ATP, and 0.1 units (Weiss) μL^{-1} T4 DNA ligase. The solution was incubated at 23°C for 1 hour, and the ligated 109-nt DNA was purified by 10% denaturing PAGE.

The 109-nt DNA population constructed as above was used as the initial pool (denoted generation 0 or G0), which was heated to 90°C for 30 seconds, cooled to room temperature, and then combined with a 2X selection buffer (100 mM HEPES pH 6.8 at 23°C, 800 mM NaCl, 200 mM KCl, 15 mM $MgCl_2$, 10 mM $MnCl_2$, 2 mM $CoCl_2$, 0.5 mM $NiCl_2$, and 2.5 mM $CdCl_2$) to a final DNA concentration of 0.05 μM. The mixture was incubated for self-cleavage at 23°C for 5 hours.

The cleavage reaction was stopped by the addition of EDTA (pH 8.0) to a final concentration of 30 mM. The cleaved DNA was isolated by 10% denaturing PAGE. To increase the yield of DNA recovery and to track the status of 94-nt cleaved product, 0.25 pmol of highly radioactive 94-nt DNA marker, made by alkaline digestion of the 109-nt construct, was used as the "carrier DNA". The isolated cleavage product was amplified by PCR in 5 × 100 μL reaction volume using primers P1 and P2 (Figure 6A). The PCR reaction was monitored in real-time using SYBR Green (Molecular Probes). The amplified DNA product (2%) was used as the DNA templates for a second PCR reaction in a 10 × 100 μL reaction volume using primer P1 and ribo-terminated primer P3 (Figure 6A). The reaction mixture also contained 30 μCi of [α-^{32}P]dGTP for DNA labeling.

The DNA product in the second PCR was recovered by ethanol precipitation, resuspended in 90 μL of 0.25 M NaOH and incubated at 90°C for 10 minutes to cleave the single embedded RNA linkage. The cleavage solution was neutralized by adding 10

μL of 3 M NaOAc (pH 5.2 at 23°C), and the 86-nt single-stranded DNA fragments were isolated by denaturing 10% PAGE. The recovered DNA molecules were incubated with 10 units of PNK at 37°C for 1 hour for DNA phosphorylation in a 100-μL reaction mixture containing 50 mM Tris-HCl (pH 7.8 at 23°C), 40 mM NaCl, 10 mM $MgCl_2$, 1 mg/mL BSA, and 0.84 μM [γ-^{32}P]ATP (70μCi). The 5′-phosphorylated DNA was used for the second round of selection using the same procedure described for the first round of selection.

A total of 22 rounds of selection and amplification were carried out. The third and subsequent rounds of selection were conducted in the same way as described for the first round except that the scale of the self-cleavage reaction, PCR, DNA phosphorylation, and DNA ligation reactions was all reduced by 5-fold. Also, the self-cleavage time was shortened from 5 hours (G0-G11), to 10 minutes (G12), to 1 minute (G13), to 30 seconds (G14 and G15), to 5 seconds (G16 and G17), and finally to 1 second (G18-G22). The "carrier DNA" was not used after G2.

2.3 Cloning and Sequencing of Selected Deoxyribozymes

DNA sequences recovered from the selection were amplified by PCR and cloned into a vector by the TA cloning method. The plasmids containing individual catalysts were prepared using Promega Wizard *Plus* SV Miniprep and Eppendorf Perfectprep Plasmid 96 Vac Direct Bind. DNA sequencing was performed on a CEQ 8000 capillary DNA sequencer (Beckman Coulter), following the procedures recommended by the manufacture.

2.4 Fluorescence Measurements

All measurements were taken from 400-μL solutions on a Cary Eclipse fluorescence spectrophotometer (Varian). The excitation was set at 490 nm, and emission at 520 nm.

2.5 Kinetic Analysis

A typical reaction involved the following steps. Before each reaction, DNA was first denatured by heating at 90°C for 30 seconds, and left in room temperature for 5 minutes before proceeding further. The reaction was started by the addition of 2X reaction buffer, and stopped by the addition of EDTA to a final concentration of 30 mM. The cleavage products were then separated by denaturing 10% PAGE, and quantitated using a PhosphoImager and ImageQuant software (Molecular Dynamics). The rate constant for the reaction was determined by plotting the natural logarithm of the fraction of DNA that remained unreacted versus the reaction time. The negative slope of the line, obtained from points within the first 10% of the reaction and produced by a least-square fit, was taken as the rate constant.

CHAPTER THREE

RESULTS & DISCUSSION

3.1 Isolation of an Efficient RNA-Cleaving and Fluorescence Signalling DNA Enzyme

We sought to create a reporter system based on an RNA-cleaving DNA enzyme that would generate a large fluorescence signal only when triggered by a specific target. In its unbound inactive state, the RNA-cleaving DNA enzyme would maintain a low level of background fluorescence. A single RNA linkage was embedded within a stretch of DNA and flanked by a covalently linked fluorophore-quencher pair. Cleavage of this single RNA linkage releases the quencher from the fluorophore, which subsequently leads to an enhanced fluorescent signal. The close juxtaposition of these three moieties--a fluorophore, an RNA junction, and a quencher--leads to efficient fluorescence quenching due to the close proximity of the fluorophore and the quencher. Furthermore, the close arrangement of these three moieties minimizes the likelihood of false positives, since the quencher cannot be physically separated from the fluorophore until RNA cleavage occurs.

Creating this unique reporter system was a multi-step process. One of the earliest steps was the isolation of a DNA enzyme capable of cleaving an RNA linkage sandwiched between two nucleotides, one labeled with a fluorophore and the other with a quencher. We adopted the technique of *in vitro* selection, which had been proven to be a

powerful method for isolating DNA molecules that could serve as catalysts or aptamers. The DNA enzyme was selected in the presence of divalent metal ions, which play an important role in facilitating catalysis.

Once a DNA enzyme that can efficiently cleave our designed substrate had been isolated, we then proceeded to the next step of our reporter system engineering and subjected the isolated enzyme under allosteric control, using the strategy of rational design. Various allosterically regulated ribozymes (reviewed in Soukup and Breaker, 1999), along with one DNA ligase enzyme (Levy and Ellington, 2002) have been successfully engineered through this approach.

3.1.1 Substrate design

To achieve our goal of establishing an efficient reporter system, it was critical to design a substrate that can provide a low level of background fluorescence, and a high degree of fluorescent signalling upon RNA-cleavage. These favourable characteristics would be determined largely by the respective positions of each of the three key moieties: the fluorophore, the quencher, and the ribonucleotide junction. To determine the optimal distance between the fluorophore and quencher, a series of DNA oligonucleotides were synthesized with the modifications shown in Figure 5A (F: fluorescein-dT; Q: dabcyl-dT; Ar: adenine ribonucleotide). The RNA cleavage-dependent signalling behaviour of these molecules was assessed by treatment with 0.25 M NaOH (Figure 5B), where F_0 and F were the fluorescence intensities of a relevant solution measured immediately after the addition of 0.25 M NaOH (RNA cleavage yet to occur) and after incubation for 20 hours, respectively (full RNA cleavage, Li and Breaker, 1999c). F1QDNA displayed the most significant fluorescence change (with an increase in intensity of ~70-fold), followed by F2QDNA (~30-fold). In contrast, F3QDNA produced a much smaller fluorescence enhancement of only 4-fold. The decrease in fluorescence enhancement that was observed when the fluorophore and quencher labels were moved further apart on the molecule was due to increasing background fluorescence (F_0), which may be attributed to less efficient quenching. All FxQDNA systems (x = 1~3) reached final intensity values similar to that of FDNA. The small increase in intensity seen with FQDNA (no RNA linker) likely resulted from a slow base-catalyzed hydrolysis reaction that removed either the F or Q moiety from the DNA backbone.

Since F1QDNA exhibited the largest fluorescence intensity increase, a cleavage system based on this type of modification should be the most sensitive in terms of signal generation. For this reason, a ribonucleotide immediately flanked by a fluorophore-modified nucleotide and a quencher-modified nucleotide was incorporated into the initial random-sequence pool to be used for the creation of our DNA enzymes.

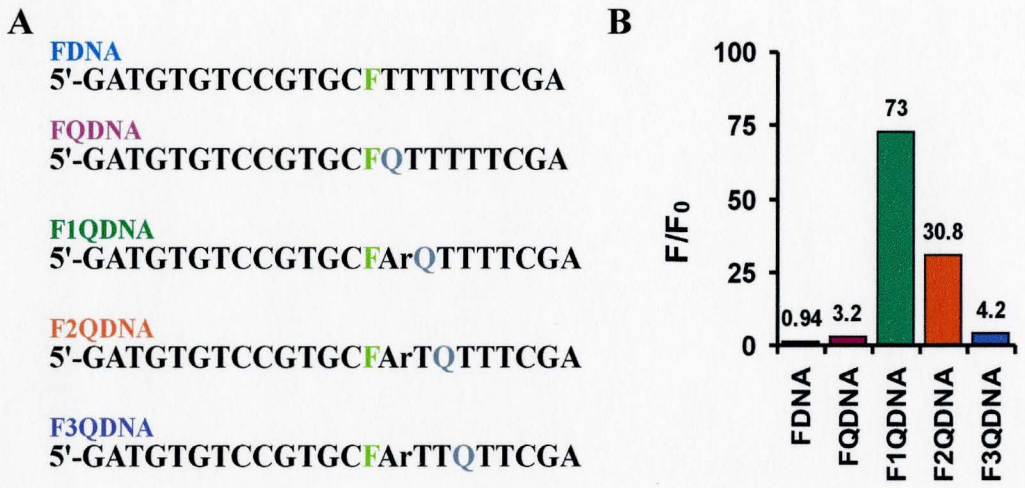


Figure 5: Examination of the signalling capability of DNA oligonucleotides modified with a fluorophore, a quencher, or a ribonucleotide. (A) DNA sequences used. F: fluorescein-dT; Q: dabcyl-dT; Ar: adenine ribonucleotide. (B) Fluorescence enhancement of each molecule following 20-hours incubation in the presence of 0.25 M NaOH. 0.5M NaOH was combined with an equal volume of a solution containing 80 nM of FDNA, FQDNA, F1QDNA, F2QDNA, or F3QDNA. The fluorescence intensity of each mixture was recorded twice: one measurement was taken immediately following the addition of NaOH (F_0) and the other after 20 hours of incubation at room temperature (F).

3.1.2 Library design

A successful *in vitro* selection is governed by many factors, one of which is an effective library design. The process of *in vitro* selection is based on the sheer probability that some DNA molecules in a large random-sequence library will fold into the appropriate tertiary structure and perform some desired function. Consequently, the length of the random-sequence domain is a critical decision, as it will largely determine the extent of sequence diversity in the initial DNA pool. The choice is not readily apparent, however, as there are advantages and disadvantages associated with both long and short random-sequence libraries. A library of DNA molecules with long random-sequence domains would provide a vast reservoir of sequence diversity. For instance, a library with a random-sequence domain of 80 nucleotides can theoretically generate 4^{80} or approximately 10^{48} unique sequence variants. Furthermore, a large random-sequence catalytic domain may be necessary to form complex, multifaceted structures, which in turn may be a prerequisite for performing complex, multifaceted reactions.

On the other hand, DNA libraries typically consist of no more than 10^{15} different molecules, due to the inherent limitations associated with chemical synthesis. Therefore, a DNA library with an 80-nt random-sequence domain would only cover a small fraction of the total conceivable sequence space. By contrast, a DNA library based on a 23-nt random-sequence domain would cover the entire sequence space, such that at least one copy of each possible sequence variant would be present in the initial pool. A short random-sequence domain also helps to preclude the formation of non-specific interactions between adjacent molecules within the library. In addition, a short random-

sequence domain provides an attractive economic incentive by reducing the cost of synthesis.

Ultimately, we decided to include some libraries with small random-sequence regions (20 to 23 nt) as well as one library with a relatively large random-sequence region (43 nt) in our starting pool. In this way we were able to accommodate both arguments associated with the random-sequence length, and used it as an opportunity to investigate how the design and size of a starting library would affect the success of an *in vitro* selection.

Six DNA libraries were constructed for our *in vitro* selection effort. Each single-stranded DNA library was composed of molecules containing a total of 86 nucleotides, including a variable zone and two primer binding sites necessary for PCR amplification (Figure 6). The schematic drawing in Figure 6 illustrates each of the six libraries in their putative folded structures as obtained from the folding program, *mfold* (it can be accessed at http://www.bioinfo.rpi.edu/applications/mfold/old/dna/; Zuker, 2003). *Mfold* is a computer program that predicts optimal and sub-optimal secondary structures based on the stability parameters of Turner and colleagues, and the energy minimization method of Zuker (Riesner, 1997). For library A to E, regions within the precursor libraries were designed in such a way that the predicted secondary structures would yield two binding arms (stem 1 and 2 in Figure 6B and 6C) to the invariable substrate strand (illustrated in green, Figure 6), while the random-sequence domain in between would act as the catalytic core for performing catalysis. The proposed binding arms would serve to bring

the substrate in proximity with the putative catalytic domain represented by the random-sequence region.

Libraries A, B, C, and D each contained 23 random nucleotides, while library E contained only 20 random nucleotides. These random nucleotides were distributed on either side of an incorporated ACT stem-loop. The random-sequence regions of libraries B, C, and D also included a random-sequence bulge consisting of 3, 6, or 9 nucleotides, whereas library E consisted of a TGAT bulge. These engineered bulges are intended for providing more catalytic structure options. The ACT stem-loop (stem 3 in Figure 6B and 6C) was introduced so that future allosteric enzymes could be facilely engineered, simply by attaching an aptamer domain in place of the ACT loop (Breaker, 2002).

A DNA library based on a 23-nt random-sequence domain gives rise to approximately 10^{14} different sequence variants (from 4^n, where n = # of random nucleotides). DNA libraries are typically synthesized on a 1 μmol-scale to yield a total of 10^{15} different sequences (after PAGE gel purification). We believed 23 random nucleotides in each library would afford sufficient variation in the initial pool providing approximately 10 copies of each unique sequence. Another library (Library F, Figure 6) consisting of 43 random nucleotides and lacking pre-designed features was also used. This particular library was employed in an effort to circumvent any unforeseen limitations associated with the pre-designed stem-loops or short random-sequence domains used in the other libraries, which might hinder the emergence of potential catalysts. Each library was tested for its ability to be amplified by PCR method, and to be ligated to the substrate prior to proceeding with the *in vitro* selection.

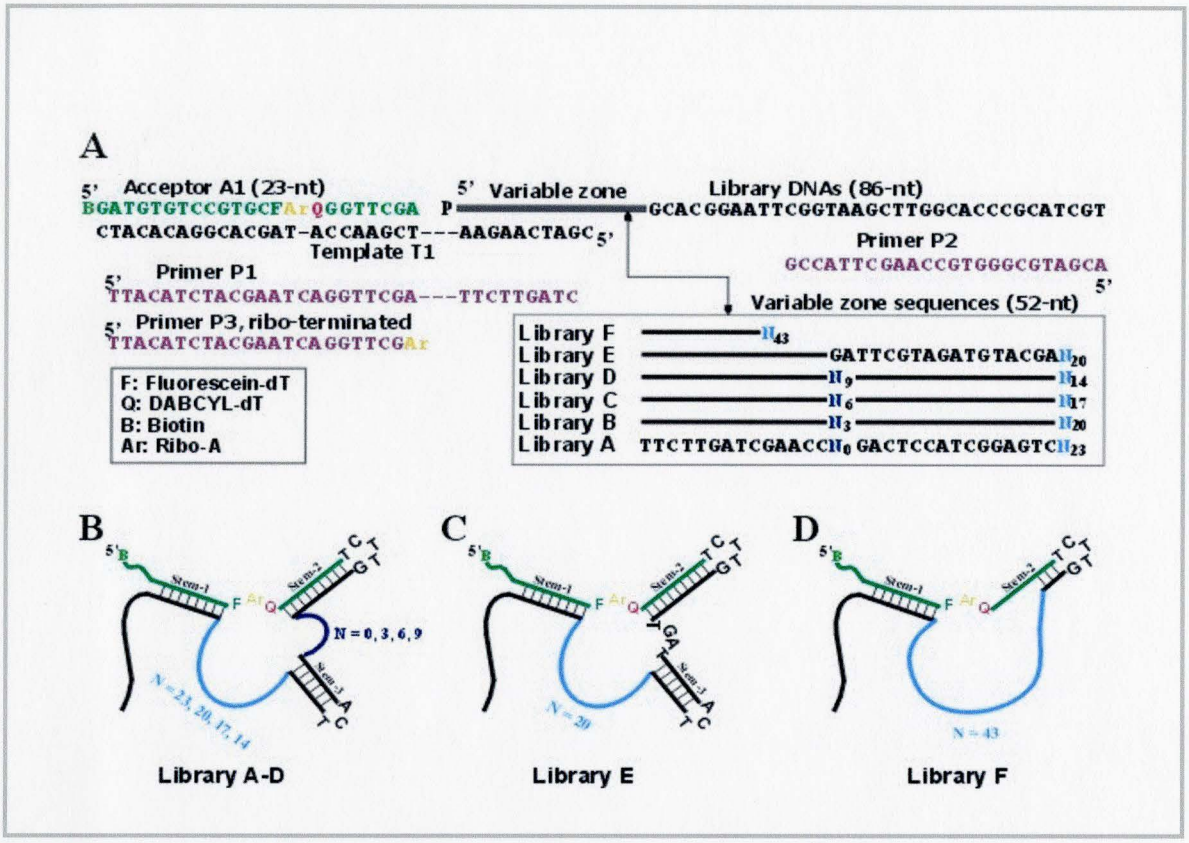


Figure 6: Six pre-designed libraries engineered with known stem-loop regions. (A) Library and PCR primer sequences. (B) Putative secondary structure of library A to D. All libraries, except library A, possess an ACT stem-loop in the middle of the randomized region. (C) Putative secondary structure of library E. A TGAT bulge was introduced in place of the short randomized region 5' to the ACT stem loop. (D) Library F contained 43 consecutive random nucleotides, and lacked pre-designed stem-loop or bulges.

3.1.3 In vitro selection of a fluorescence-generating, RNA-cleaving DNA enzyme

The enzyme DEC22-18 was isolated from a pool of 109-nt DNA molecules, which were constructed by ligating the designed 23-nt substrate to the 5' end of the 86-nt DNA library molecules containing the putative catalytic random-sequence domain. Each 109-nt molecule in the starting pool (denoted G0) contained the three key moieties engineered in the 5'-fixed sequence region (see section 3.1.1). Twenty-two rounds of *in vitro* selection were performed in the presence of the monovalent ions, Na^+ or K^+, as well as a group of divalent metal ions including Mg^{2+}, Mn^{2+}, Co^{2+}, Ni^{2+}, and Cd^{2+}. The type and concentration of each metal ion used in the selection buffer was rather arbitrary, based largely on some unpublished data obtained from previous studies in our laboratory (Achenback and Li, unpublished data). In each round of selection (Figure 7A), the ribonucleotide junction on the substrate was cleaved to yield 94-nt products, which were subsequently separated from the 109-nt precursor molecules by 10% denaturing polyacrylamide gel electrophoresis (PAGE). Those molecules capable of self-cleavage into the 94-nt products were subjected to a two-step PCR amplification: the first PCR was used to amplify cleaved fragments into multiple copies of double-stranded 109-nt molecules, while the second PCR served to introduce the ribonucleotide-terminated primer (step VI, Figure 7A). An alkaline treatment cleaved the ribonucleotide at the sense strand, yielding the original-length 86-nt single-stranded DNA library molecules. Those 86-nt oligonucleotides were then ligated with the 23-nt A1 to regenerate the 109-nt precursor molecules to initiate the next round of selection (Figure 7A).

The reaction time used in the first eleven rounds of selection was 5 hours, which was progressively reduced to 10 minutes for the twelfth round (G12), 1 minute for the thirteenth round (G13), 30 seconds for the fourteenth (G14) and fifteenth rounds (G15), 5 seconds for the sixteenth (G16) and seventeenth rounds (G17), and finally 1 second from the eighteenth round (G18) onward to yield the most catalytically efficient sequence(s) (Figure 7B). The decision to cease further rounds of *in vitro* selection was made after observing a plateau in the self-cleavage activity for five rounds of selection. Populations from G10 (5 hours reaction time), G17 (5 seconds reaction time), and G22 (used 1 minute reaction time pool, refer to section 3.1.4 for explanation) were selected for cloning and subsequent sequence analysis.

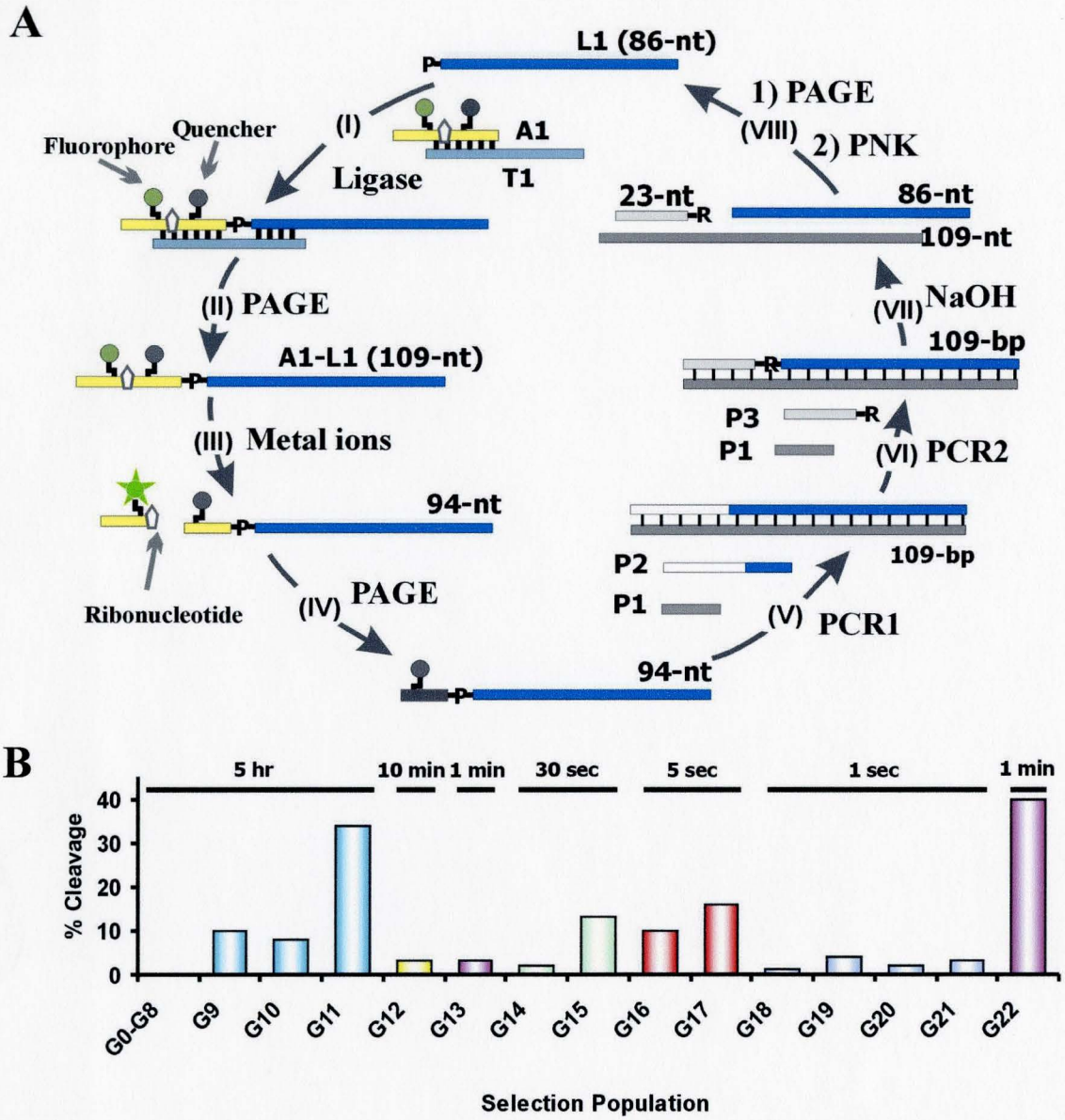


Figure 7. *In vitro* selection of a fluorescence-signalling, RNA-cleaving DNA enzyme. (A) The enzyme DEC22-18 was isolated from a pool of 10^{13} single-stranded 109-nt DNAs, each containing the three key moieties engineered in the 5'-fixed sequence region (see figure 6A for the construct design). (B) 22 rounds of *in vitro* selection were carried out in the presence of monovalent ions, Na^{+} or K^{+}, and a group of divalent metal ions including Mg^{2+}, Mn^{2+}, Co^{2+}, Ni^{2+}, and Cd^{2+}. A reaction time of 5 hours was used for the first twelve rounds of selection, and was gradually reduced to yield the most catalytically efficient sequence(s). Reaction times are given on the top of the graph.

3.1.4 Cloning and Sequencing

Rather than using the G21 (generation 21) one-second population for cloning and sequencing, the catalytic molecules from this population were allowed to self-cleave for an extended period of time (one minute reaction time) during G22 selection (Figure 7B), for the purpose of yielding a more diverse pool for analysis. Sequencing revealed only one class of sequence in the terminal population (Figure 8D), and this sequence was named DEC22-18 (DE stands for DNA enzyme, C stands for *cis*-acting enzymes, and 22-18 indicates that this clone was identified as the 18^{th} clone in generation 22). Generation 10 and 17 were also cloned and sequenced at the same time (Figure 8B & 8C). The final clone, DEC22-18, appears to have dominated the selection at or before generation 17, since all sequences analyzed from both the G17 and G22 pools were variants of DEC22-18 (Figure 8C & 8D). The pool of generation 10 yielded at least 14 unique sequence classes, which were identified from a total of 53 successfully sequenced clones. For each class in generation 10, one representative sequence was shown in Figure 8B, where the total number of clones belonging to the same class was listed in brackets beside the class number (in Roman numerals). For instance, the sequences designated as class I appeared in 14 out of a total of 53 clones. Since the selection stringency was relatively mild at a reaction time of 5 hours, it was not surprising to observe 14 unique sequence classes in the G10 pool.

Even though six different libraries were used in the selection (refer to Figure 6B), the majority (13 out of 14 classes) of sequences recovered from G10 appeared to be descendents of library F (containing no engineered stem-loop or bulge). Only class V of

the G10 sequences (Figure 8B) appeared to be derived from one of the partially randomized libraries (library D) containing an engineered stem-loop (Figure 6B). However, as the selection stringency was increased by decreasing the reaction time, the class V molecules were ultimately eliminated from the pool of surviving DNA enzymes. As it were, only one class of catalytic DNA molecules survived the rigorous selection pressure – the enzyme referred to as DEC22-18. It seems that library F played a predominant role in our selection, as reflected by the preponderance of library F descendants observed in generation 10, and in addition to the fact that the most efficient catalyst was derived from this library. The primary difference between library F and the libraries A to E was the length of the random-sequence domain. The predominance of library F, therefore, may suggest that the smaller random-sequence domain that characterized libraries A to E may not provide sufficient structural complexity for performing efficient catalysis.

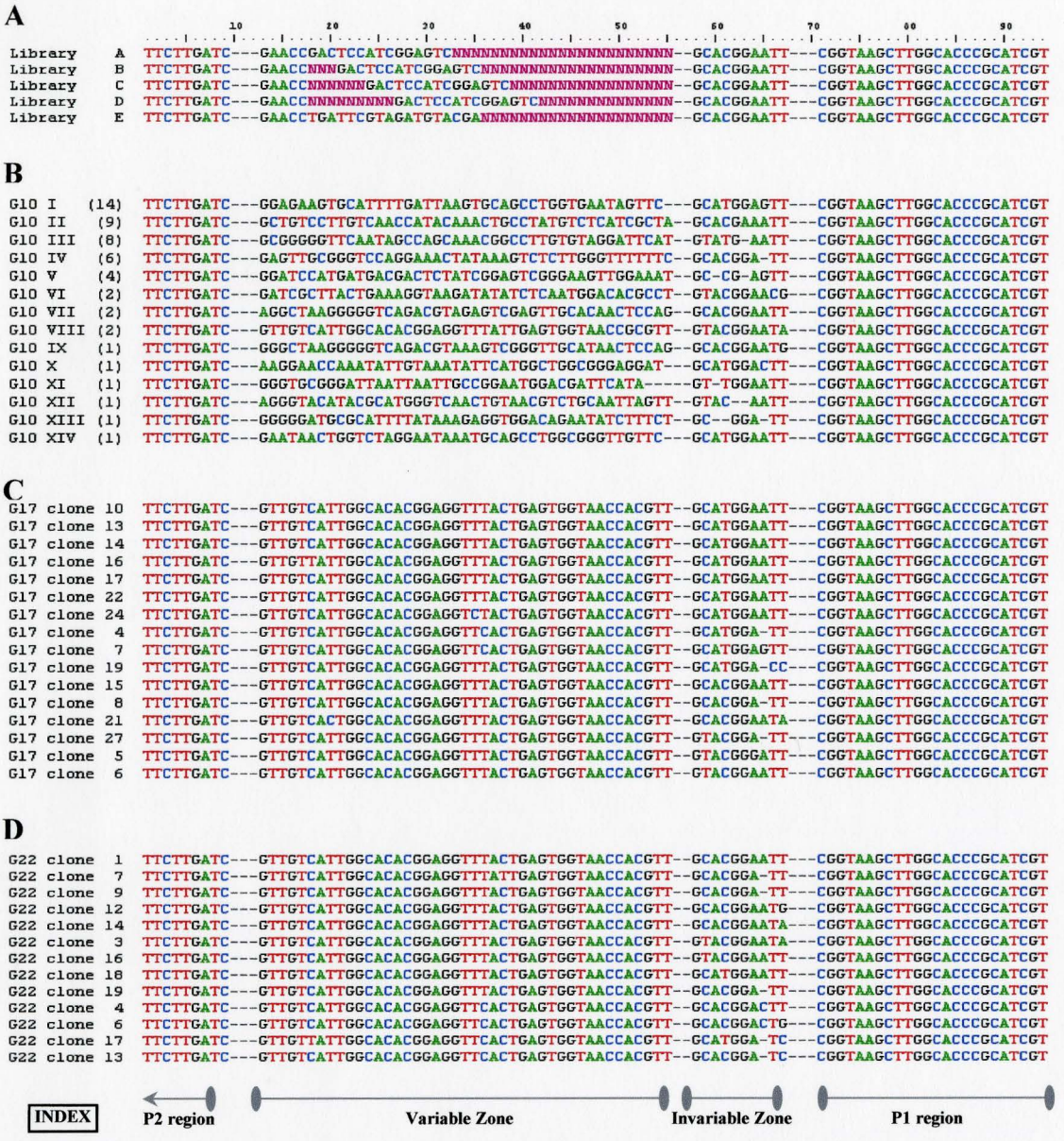


Figure 8: Sequencing results on clones obtained from generations 10, 17 and 22. (A) Sequences of randomized libraries A to F used for *in vitro* selection. Pink letters represent the randomized nucleotides. (B) For G10 pool, every sequence listed here represents one unique sequence class. The members of each class differ only by point mutations or deletions. A total of 14 classes were identified from 53 clones sequenced. The representative sequence of each class was given a class number in Roman numeral with the number of clones identified for this particular class in brackets. Only one class was identified from generations 17 (C) and 22 (D).

3.1.5 Metal specificity and kinetic analysis of DEC22-18

Chemical reactions catalyzed by ribozymes and deoxyribozymes require the participation of metal ions, in executing chemical actions and/or stabilizing secondary and tertiary structures. The correct positioning of metal ion cofactors and substrates are essential in stabilizing high-energy transition states (Yarus, 1993). Divalent metal ions can have a direct or indirect hydrolytic effect on nucleic acids depending on the p*K*a of the hydrated metal ion, due to their ability to polarize functional groups (Faulhammer and Famulok, 1997). Specific binding of the metal ions with nucleic acid enzymes plays a critical role in the activation of bound water or cofactors (Cech, 1993). Monovalent ions such as Na^+ and K^+ serve to stabilize the active conformation of the deoxyribozymes (Carmi and Breaker, 2001).

DEC22-18 exhibited a k_{obs} of 1.0 min^{-1} under the original selection buffer conditions (50 mM HEPES pH 6.8 at 23°C, 400 mM NaCl, 100 mM KCl, 7.5 mM $MgCl_2$, 5 mM $MnCl_2$, 1 mM $CoCl_2$, 0.25 mM $NiCl_2$, and 1.25 mM $CdCl_2$). It self-cleaved to release the fluorescent 5' segment (15-nt) and the radioactive 3' segment (94-nt) from the original 109-nt full-length enzyme (Figure 9A). Two types of monovalent ions, along with five types of divalent metal ions, including four from the transition metal group, were used in the selection buffer. Most *in vitro* selected deoxyribozymes make use of divalent metal ions to assist with catalysis (Lu, 2002), with few exceptions that require only monovalent ions (Geyer and Sen, 1997; Faulhammer and Famulok, 1997). To appreciate the full catalytic potential of DEC22-18, therefore, it was important to

investigate the type and concentration of metal ions that would confer the greatest catalytic ability.

When DEC22-18 was assessed for metal ion requirements, its catalytic behaviour appeared to be dependent on the presence of some particular divalent metal ions, mainly Co^{2+}, Ni^{2+}, and Mn^{2+} (Figure 9C). The transition metal ion Cd^{2+} (13.75 mM $MgCl_2$ / 1.25 mM $CdCl_2$) and divalent metal ion Mg^{2+} (15 mM $MgCl_2$) did not support catalysis, with the particular concentrations that were tested. Each metal ion was used in its respective concentration as found in the original selection buffer, and a corresponding concentration of $MgCl_2$ was used simultaneously to maintain an overall concentration of divalent metal ions of 15 mM. For instance, DEC22-18 was tested for its dependence on Mn^{2+} in a buffer that contained 50 mM HEPES pH 6.8 at 23°C, 400 mM NaCl, 100 mM KCl, 10 mM $MgCl_2$ and 5 mM $MnCl_2$. DEC22-18 did not cleave in H_2O (lane 2 on Figure 9C) or with monovalent ions in 50 mM HEPES buffer at pH 6.8 (lane 3 on Figure 9C). The radioactive intensity of the 3' product band depicted in the phosphorimage (P2 band on Figure 9C), is proportional to the fraction of self-cleaved molecules. Conversely, the fluorescence intensity of the 5' product band depicted in the fluorimage (P1 on Figure 9C), is not proportional to the fraction of cleaved molecules, and so the intensity appears several folds higher than the uncleaved 109-nt precursor band. This is because the separation of the quenching agent, dabcyl, from the fluorophore, fluorescein, results in de-quenching or the abolition of FRET and the subsequent fluorescence enhancement.

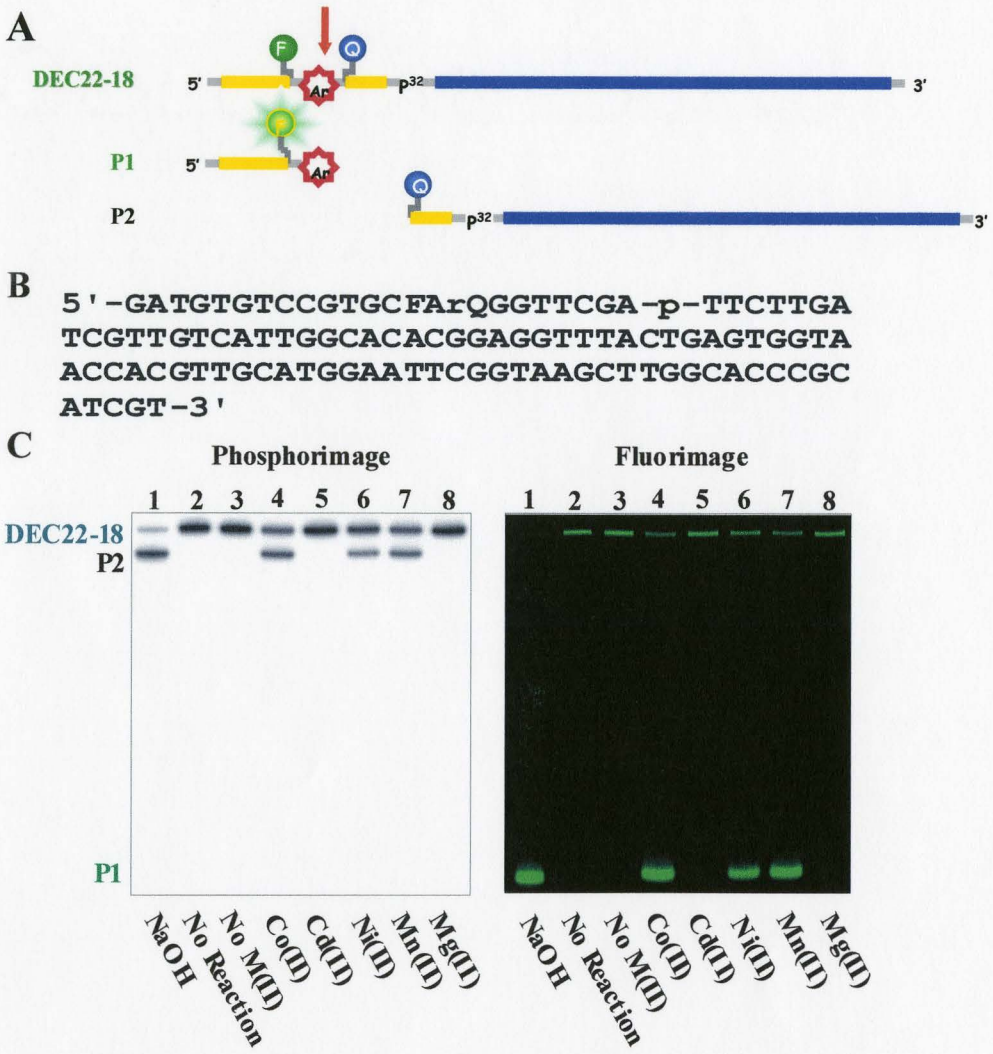


Figure 9: Catalysis and metal specificity of DEC22-18. (A) The construct of the full-length, *cis*-acting enzyme. Three key moieties were engineered in the 5'-fixed sequence region (F: fluorescein-dT; Q: DABCYL-dT; Ar: adenine ribonucleotide). Prior to the cleavage reaction, the full-length enzyme was radioactive with its fluorescence partially quenched. Upon self-cleavage, two product bands were generated. Cleavage occurred 3' adjacent to the adenine ribonucleotide, thus releasing fluorescent P1 and radioactive P2. (B) Sequences of DEC22-18. (C) DEC22-18 was assessed for catalytic activity and divalent metal ion requirement. Two product bands, P1 and P2, were obtained by partial NaOH treatment and used as the control (lane 1). No cleavage product was seen when DEC22-18 was treated with water (lane 2), monovalent ions (lane 3), Cd^{2+} (lane 5), or Mg^{2+} (lane 8). When DEC22-18 was treated with Co^{2+} (lane 4), Ni^{2+} (lane 6), or Mn^{2+} (lane 7), it self-cleaved into P1 and P2.

Further investigation revealed Co^{2+} to be the preferred divalent metal ion for highly efficient catalysis, while the monovalent ions, Na^+ and K^+, slightly inhibited the activity. Various concentrations of the transition metal ions required for DEC22-18 catalysis (Mn^{2+}, Ni^{2+}, and Co^{2+}) were used to assess their effects on the enzyme's self-cleaving capability. A $MnCl_2$ concentration as high as 40 mM was required to promote the self-cleavage of DEC22-18 to approach 20% in 10 seconds (Figure 10A), while around 4 mM $NiCl_2$ was adequate for 25% self-cleavage within the same reaction time (Figure 10B). The enzyme DEC22-18 exhibited the highest activity in the presence of 10 mM $CoCl_2$, where more than 40% of the enzyme self-cleaved in 10 seconds (Figure 10C). A Co^{2+} concentration higher than 10 mM, however, did not significantly improve the catalytic efficiency. Experiment with 0 to 1500 mM Na^+ (400 mM NaCl was the concentration used during *in vitro* selection) in the presence of 100 mM K^+ and 10 mM of Co^{2+} showed that removing Na^+ actually improved the catalytic efficiency in the presence of Co^{2+} (Figure 11). Experiment with 0 to 300 mM K^+ (100 mM KCl was the concentration used during *in vitro* selection) in the presence of 400 mM Na^+ and 10 mM of Co^{2+} showed a similar trend, except the activity was slightly higher at 10 mM rather than at 0 mM (Figure 11). Nevertheless, the removal of both Na^+ and K^+ from the selection buffer, in the presence of 10 mM Co^{2+}, improved the activity of DEC22-18 up to 5-fold (45% of self-cleavage in 5 seconds) over the activity observed in the original selection buffer (9% of self-cleavage in 5 seconds).

The monovalent ion independence of DEC22-18 suggests that this enzyme has a very stable structure. The fact that DEC22-18 does not require any monovalent ions,

suggests that the enzyme can fold into its conformation without the assistance of monovalent ions. Monovalent ions have also been shown to assist in catalysis when used in high concentration, as in the case of hammerhead ribozyme (Murray et al., 1998). The use of divalent but not monovalent metal ions in DEC22-18 indicates that the cations are probably used for inner sphere co-ordination rather than electrostatic stabilization of the folded state (Blount et al., 2002). The metal ions could be required for the specific folding of DEC22-18 or be directly involved in the catalytic mechanism. A detailed analysis has to be conducted in order to elucidate the specific role played by the divalent metal ions, and to determine whether they contribute to catalysis or the structural stabilization of DEC22-18.

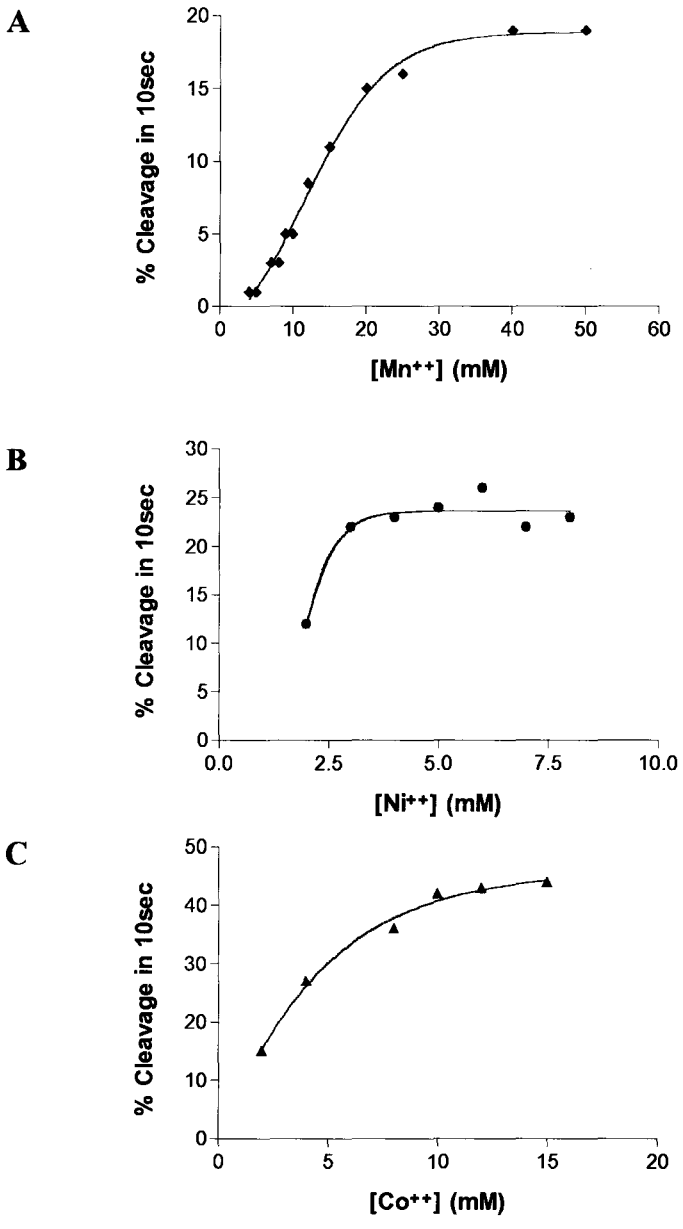


Figure 10: Effect of various divalent metal ion concentrations on DEC22-18 activity. (A) DEC22-18 was subjected to various concentrations of $MnCl_2$ in 400 mM NaCl, 100 mM KCl and 50 mM HEPES (pH 6.8 at 23°C) in the presence of a corresponding concentration of $MgCl_2$ to maintain an overall 15 mM divalent metal ion concentration (except in the case of the ones exceeded 15 mM, in which no $MgCl_2$ was added). Each reaction was stopped after 10 seconds with the addition of quenching buffer (30 mM EDTA and loading dye), loaded onto 10% PAGE gel, and quantified using a PhosphorImager and ImageQuant software. The same procedure for Mn^{2+} concentration assay was used for the Ni^{2+} curve in (B) and Co^{2+} curve in (C).

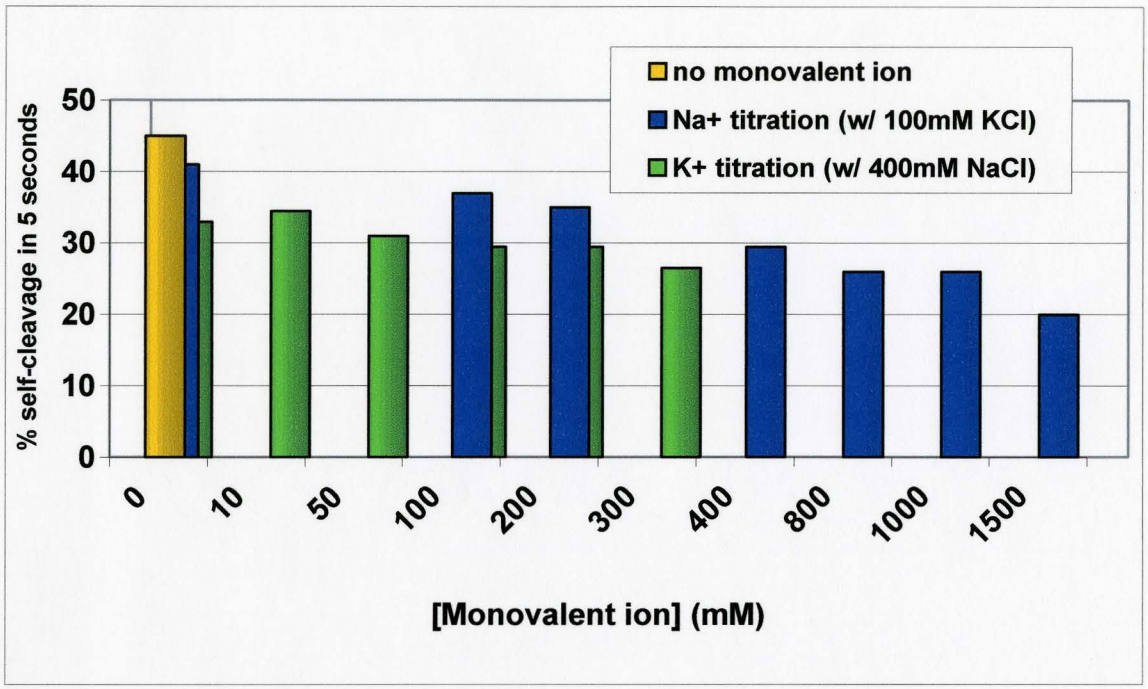


Figure 11: Effect of various monovalent ion concentrations on DEC22-18 activity. Blue bars represent percent cleavage of DEC22-18 in 5 seconds reaction time under various concentrations of NaCl buffer (range from 0 to 1500 mM,, along with 100 mM KCl, 10 mM $CoCl_2$, and 5 mM $MgCl_2$ in 50 mM HEPES, pH 6.8 at 23°C). Green bars represent percent self-cleavage of DEC22-18 in 5 seconds reaction time under various concentrations of KCl buffer (ranging from 0 to 300 mM, along with 400 mM NaCl, 10 mM $CoCl_2$ and 5 mM $MgCl_2$ in 50 mM HEPES, pH 6.8 at 23°C). The yellow bar represents percent cleavage of DEC22-18 in the presence of 10 mM $CoCl_2$ and 5 mM $MgCl_2$ in 50 mM HEPES (pH 6.8 at 23°C) without any monovalent ion.

3.1.6 The pH profile of DEC22-18

DEC22-18 was selected in a buffer composed of 50 mM HEPES (pH 6.8 at 23°C), 400 mM NaCl, 100 mM KCl, 7.5 mM $MgCl_2$, 5 mM $MnCl_2$, 1 mM $CoCl_2$, 0.25 mM $NiCl_2$, and 1.25 mM $CdCl_2$. Kinetic studies performed on previous RNA-cleaving DNA enzymes and ribozymes (Santoro and Joyce, 1997; Breaker et al., 2003) demonstrated that the catalytic rates generally increase as the pH of the reaction buffer increases. This trend is especially true if the mechanism of RNA-cleavage uses either a metal-bound water or the 2'-hydroxyl group of RNA to perform an S_N2 reaction (Dahm et al., 1993). Dahm and Uhlenbeck also proposed that a single deprotonation event occurs on the 2'-hydroxyl of the RNA adjacent to the cleavage site as part of the rate-limiting step of hammerhead ribozyme-catalyzed reactions.

The fluorophore used in this study was fluorescein, which is pH sensitive and emits greater fluorescence at an elevated pH, particularly above pH 7.0 (Lakowicz, 1999). Consequently, determining the optimal pH range for DEC22-18 was an important step towards fully realizing the full catalytic aptitude and signalling capacity of this enzyme.

The pH dependence of DEC22-18 was studied in the presence of either 6 mM Ni^{2+} or 10 mM Co^{2+} under different pH conditions at 23°C, while an appropriate concentration of $MgCl_2$ was used to maintain a minimum overall divalent metal ion concentration of 15 mM (Figure 12). The three buffers used were MES for pH 6.4, HEPES for pH 6.8 and 7.6, and Tris-HCl for pH 7.4, 7.7, 8.0, and 8.5. The Tris-HCl buffer was specifically chosen to accommodate the pH preference of fluorescein, which

is above pH 7.0. Certain divalent metal ions, including Ni^{2+} and Co^{2+}, precipitate at a pH higher than 7.5 due to their relatively low solubility product constant. This can be demonstrated in the following equation:

$$M^{2+}{}_{(aq)} + 2(OH^-) \leftrightarrow M(OH)_{2\,(s)}$$

$$\text{Solubility Product Constant} = K_{sp} = [M^{2+}][OH^-]^2$$

The M^{2+} here represents any heavy divalent metal ion. The metal ion combines with hydroxide ion to form an insoluble metal hydroxide solid. As the pH of the buffer increases, more metals react with hydroxide ions to form precipitates. Therefore, a buffer with pH higher than 8.5 was not included in this study.

DEC22-18 showed the highest activity in HEPES buffer at pH 6.8. As it were, this buffer also happened to be the original selection buffer used to isolate the enzyme. Contrary to the trend typically observed with RNA-cleaving deoxyribozymes, the catalytic rate of DEC22-18 actually decreased as the pH of the solution was increased. Only Ni^{2+} and Co^{2+} were tested in this set of pH profiling experiments, since they had previously demonstrated a greater capacity for promoting DEC22-18 catalysis (please refer to results in section 3.1.5). Each reaction buffer was made with either 10 mM $CoCl_2$ and 5 mM $MgCl_2$, or 6 mM $NiCl_2$ and 9 mM $MgCl_2$, left at room temperature overnight, and then centrifuged at 14,000 rpm for 5 minutes to check for precipitation before proceeding with the assay. The solubility constants for $Ni(OH)_2$ and $Co(OH)_2$ are 2.0×10^{-15} and 1.6×10^{-15}, respectively. Consequently, 10 mM Co^{2+} and 6 mM Ni^{2+}

would precipitate at pH 7.6 and pH 7.8, respectively. This precipitation may account for the poor performance exhibited by DEC22-18 at a pH higher than 7.7. Furthermore, the rates obtained from the high pH buffer may be underestimated. Despite the fact that DEC22-18 does not exhibit a typical RNA-cleaving deoxyribozyme pH-profile, there is still insufficient data to conclude that DEC22-18 also does not utilize the mechanism adopted by most RNA-cleaving deoxyribozymes and ribozymes. Additional experiments will be required to elucidate the precise mechanism by which DEC22-18 performs RNA-cleavage.

Through monovalent and divalent metal ion concentration studies and pH profiling, the optimal buffer condition for catalysis was ultimately determined. The reaction was most effective in a solution containing 50 mM HEPES (pH 6.8 at 23°C), 10 mM $CoCl_2$, 5 mM $MgCl_2$, and no monovalent ion.

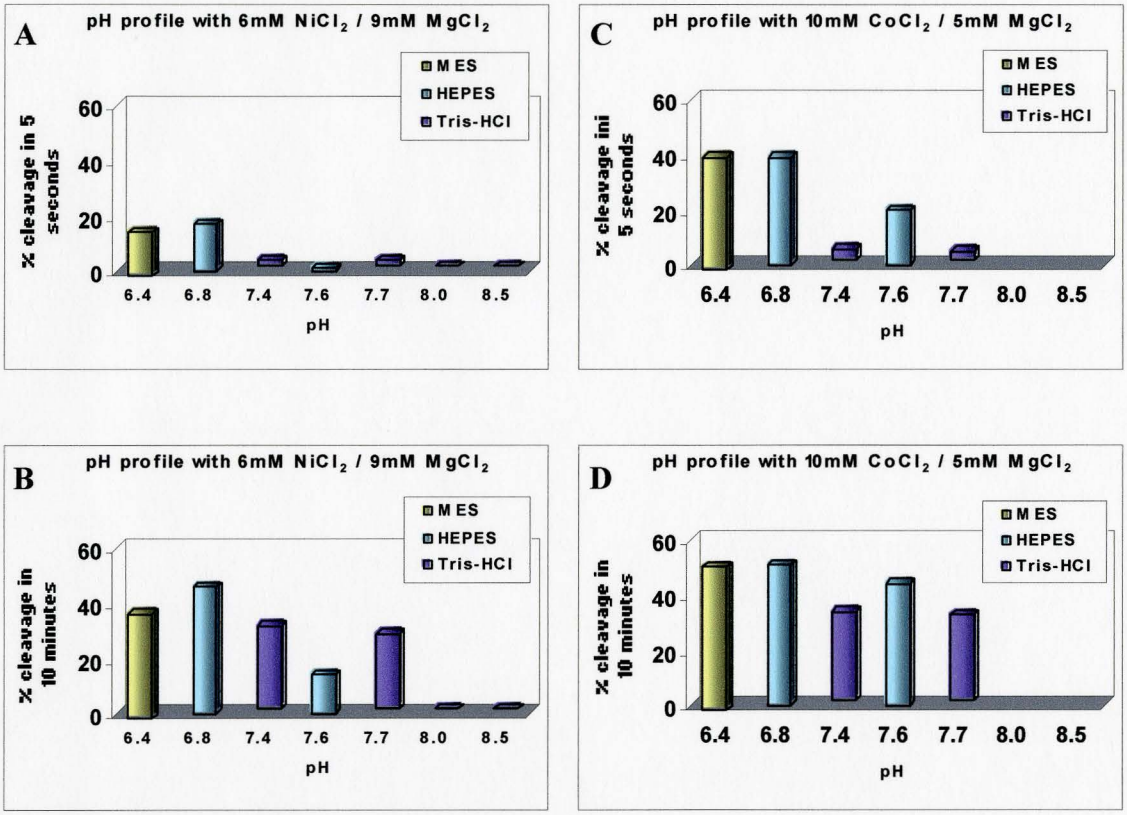


Figure 12: The pH profile of DEC22-18, in the presence of Ni^{2+} or Co^{2+}. The effect of pH on Ni^{2+}-assisted DEC22-18 reaction was investigated. Each buffer contained 6 mM $NiCl_2$ and 9 mM $MgCl_2$, in its respective pH buffer. The percent cleavage in 5 seconds (A) or in 10 minutes (B) is shown. The effect of pH on Co^{2+}-catalyzed DEC22-18 reaction was investigated. Each buffer contained 10 mM $CoCl_2$ and 5 mM $MgCl_2$, in its respective pH buffer. The percent cleavage in 5 seconds (C) or in 10 minutes (D) is shown.

3.1.7 Truncation of the full-length enzyme from the 3' end

Given the cost of synthesizing oligonucleotides and our desire to improve the versatility of DEC22-18 so that it may ultimately function as a useful reporter system, the next logical step was to perform truncation studies. DEC22-18 consists of 109 nucleotides, but not all of them will be absolutely required for catalysis. We sought to establish the minimum number of nucleotides required for full catalytic activity. A series of deletion mutants were tested to determine the essential catalytic sequence within the 109-nt, full-length DEC22-18 (Figure 13). Mutants were chemically synthesized with 10-nt, 20-nt, 30-nt, 40-nt, 50-nt, and 58-nt deletions from the 3' end of DEC22-18, and tested for self-cleavage activities. These DNA molecules were used to initially approximate the location of the catalytically essential sequence from the 3' terminus of the enzyme. From the initial experiment, we established a specific range for further truncation experiments (between 20-nt to 30-nt from 3'-end of DEC22-18), and consecutive single-nucleotide deletions were made from the 21st to the 29th nucleotide positions. These deletion mutants were synthesized and subsequently assayed to define the maximum number of nucleotides that could be deleted without compromising the catalytic activity.

Deletions anywhere between 20 to 29 nucleotides from the 3'-end of DEC22-18 resulted in an increase in the enzyme activity by as much as two-fold over the full-length wild type DEC22-18. It was not surprising that the fixed-sequence region used simply as the 3' primer-binding region (for primer P2, Figure 6A) during *in vitro* selection was neither utilized for binding nor for catalysis. In fact, the results suggest that this region

was actually inhibiting the full catalytic prowess of DEC22-18, perhaps by limiting favourable interactions between the substrate and catalytic domains. Similarly, an increase in activity was observed in an RNA ligase reported previously (Ekland et al., 1995) after the deletion of some nucleotides from an *in vitro* selected ribozyme. Presumably, these nucleotides were making unfavorable contacts with the catalytic domain or substrate domain.

Deletion of more than 30 nucleotides from the 3' end completely abolished the enzymatic activity emphasizing the catalytic importance of the 4 nucleotides located 79 to 82 nucleotides from the 5'-end (marked by purple arrows and labels, Figure 13). After determining the k_{obs} for each mutant construct, relative activity was plotted with the wild-type activity set at 100. Ultimately, 26 nucleotides were deleted from the 3'-end to produce an 83-nt enzyme (Figure 13). This shortened version of DEC22-18 was named DEC22-18A, and has a k_{obs} of 1.7 min^{-1} under the original selection conditions (50 mM HEPES pH 6.8 at 23°C, 400 mM NaCl, 100 mM KCl, 7.5 mM $MgCl_2$, 5 mM $MnCl_2$, 1 mM $CoCl_2$, 0.25 mM $NiCl_2$, and 1.25 mM $CdCl_2$). A secondary structure obtained through the *mfold* program was proposed and is illustrated in Figure 15A (see section 3.1.10). Under optimal buffer conditions (10 mM $CoCl_2$, and 5 mM $MgCl_2$ in 50 mM HEPES at pH 6.8 and 23°C) more than 50% of DEC22-18A self-cleaved in 5 seconds, which is remarkably fast by deoxyribozyme standards. Due to the exceedingly efficient nature of this optimized enzyme, it was difficult to obtain accurate measurements of the rate constant using conventional manual quenching techniques. The chemical-quench flow technique was subsequently used in an attempt to alleviate this problem. However,

this method also failed to produce reliable measurements due to unresolved technical problems. Ultimately, the creation of a *trans* substrate-enzyme system rectified the problem, and will be addressed in section 3.1.10.

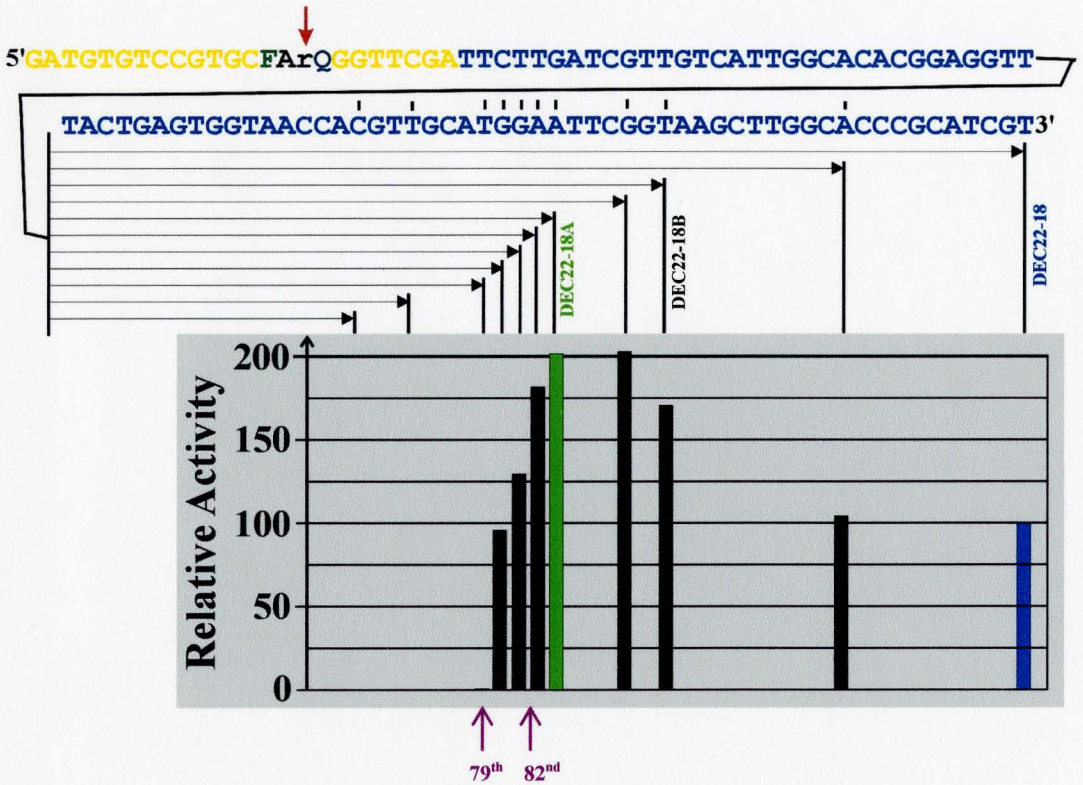


Figure 13: Sequence truncation of DEC22-18. All 3' truncated mutants of DEC22-18 were assessed for relative cleavage capabilities under the following reaction conditions: 400 mM NaCl, 100 mM KCl, 7.5 mM $MgCl_2$, 5 mM $MnCl_2$, 1 mM $CoCl_2$, 0.25 mM $NiCl_2$, and 1.25 mM $CdCl_2$ in 50 mM HEPES pH 6.8 at 23°C. Approximately 5 nM of each candidate deoxyribozyme was treated for self-cleavage at room temperature. The rate constants of the mutants were normalized against the wild-type DEC22-18 (which was taken as 100) and plotted in the embedded graph. The sequence in yellow is the substrate portion with the fluorophore-labeled nucleotide in dark green and the quencher-labeled nucleotide in dark grey. The mutants with deletions of 26 and 20 nucleotides from the 3'-end are named DEC22-18A and DEC22-18B, respectively.

3.1.8 Truncation of the substrate from the 5'-end

We wanted to determine if the entire length of the 23-nt substrate was required for catalysis. Two substrates were chemically synthesized bearing 5' deletions. One of the substrates contained a 3-nt deletion and consisted of 20 nucleotides (designated SGA20 in Figure 14), while the other substrate contained a 6-nt deletion and consisted of only 17 nucleotides (designated STA17 in Figure 14). Each substrate was ligated to a 66-nt enzyme portion (DEC22-18B, refer to Figure 13) to test their respective enzymatic activities. Although the structure depicted in Figure 14 shows a base-pairing interaction between the substrate and enzyme at stems 1 and 2, there was no evidence at the time to confirm this structure. The 66-nt DEC22-18B contains six more nucleotides at its 3'-end than the DEC22-18A construct, and also exhibits slightly less activity (see Figure 13). The decision to use DEC22-18B instead of DEC22-18A was a rather arbitrary one, influenced mainly by the ready availability of DEC22-18B at the time of the experiment. The activity assay was conducted in optimal buffer conditions (10 mM $CoCl_2$ and 5 mM $MgCl_2$ in 50 mM HEPES at pH 6.8 and 23°C) with one short time point (30 seconds) and one long time point (5 minutes). The enzyme construct containing the SGA20 substrate exhibited 52% self-cleavage in 30 seconds, while the enzyme construct containing the STA17 substrate exhibited 38% self-cleavage in the same time. By comparison, DEC22-18A containing the full-length 23-nt substrate exhibited 45% self-cleavage in 5 seconds in the optimal buffer conditions. These results suggest that the first six nucleotides at 5'-end of the substrate do not play a significant role in the structural interaction with the catalytic domain. This study was performed with the DEC22-18B construct using optimal

buffer conditions (each experiment was performed once). Since most of the data from other kinetic experiments were obtained with either the DEC22-18 or DEC22-18A construct using the selection buffer conditions, the catalytic activities are not strictly comparable. However, experiments are currently being conducted to provide more relevant comparisons.

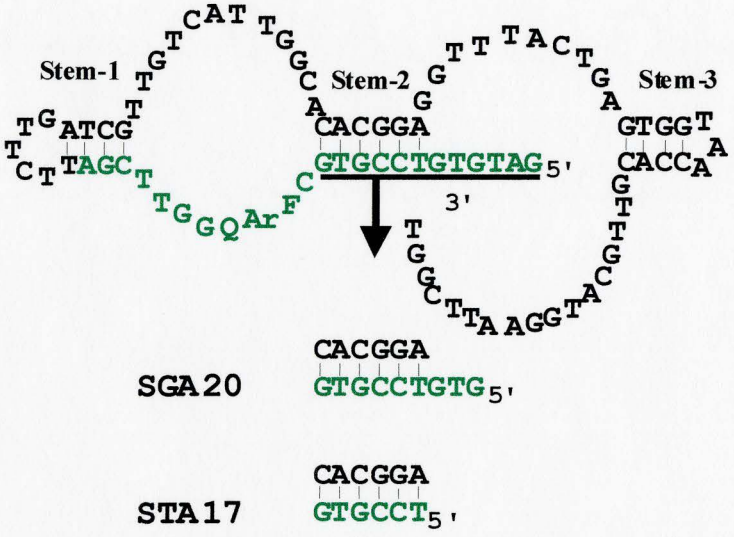


Figure 14: Truncation of the substrate from the 5'-end. Two shortened versions of the substrate (sequences in green) were chemically synthesized and ligated to the enzyme DEC22-18B (sequence in black). The substrate SGA20 contains a 3-nt deletion from the 5'-end while the substrate STA17 contains a 6-nt deletion. Each construct was subjected to self-cleavage under the optimal buffer condition (10 mM $CoCl_2$ and 5 mM $MgCl_2$ in 50 mM HEPES at pH 6.8 and 23°C) for both 30 seconds and 5 minutes. The structure shown above, including the base-paired stems 1, 2, and 3, is based on the putative secondary structure obtained from the *mfold* program.

3.1.9 Rate enhancement by DEC22-18

DEC22-18 catalyzes the self-cleavage of a single ribonucleotide located between a fluorophore (fluorescein-dT) and a quencher (dabcyl-dT). Thus, it is able to synchronize catalysis with fluorescence signalling. The k_{obs} of DEC22-18A under the selection buffer conditions (400 mM NaCl, 100 mM KCl, 7.5 mM $MgCl_2$, 5 mM $MnCl_2$, 1 mM $CoCl_2$, 0.25 mM $NiCl_2$, and 1.25 mM $CdCl_2$ in 50mM HEPES at pH 6.8 and 23°C) was determined to be 1.7 min^{-1}, while the rate under optimal buffer conditions (10 mM $CoCl_2$, and 5 mM $MgCl_2$ in 50 mM HEPES at pH 6.8 and 23°C) was too fast to be accurately measured by manual quenching (nearly 50% of DEC22-18A self-cleaved in 3 seconds).

A rough estimation of the k_{obs} of DEC22-18A under optimal buffer conditions would be ~ 14 min^{-1} (calculated from 50% self-cleavage in 3 seconds; therefore first-order rate constant = $k = 0.693 \div t_{1/2} = 0.693 \div 0.05$ min = 14 min^{-1}), though this rate was not even estimated from the linear range of the reaction. The uncatalyzed cleavage rate (6×10^{-7} min^{-1}) was determined by incubating the substrate by itself, in sub-optimal buffer conditions (10 mM $CoCl_2$, 5 mM $MgCl_2$, and 100 mM KCl in 50 mM HEPES at pH 6.8 and 23°C). The rate was obtained from the negative slope of a graph in which the natural logarithm of the fraction of unreacted DNA was plotted against the reaction time. This experiment was conducted prior to our having knowledge of the inhibitory effect of monovalent ions on the catalysis of DEC22-18 (see section 3.1.5) and hence 100 mM of

KCl was included with the other constituents that actually comprise the optimal buffer (10 mM $CoCl_2$ and 5 mM $MgCl_2$ in 50 mM HEPES at pH 6.8 and 23°C).

By taking 6 x 10^{-7} min^{-1} as the uncatalyzed rate of substrate cleavage (k_{uncat}), and 14 min^{-1} as the approximate catalyzed rate by DEC22-18A, the rate enhancement generated by this DNA enzyme under optimized conditions approaches twenty-three million fold (2.3 x 10^7 fold).

3.1.10 Construction and kinetic analysis of the *trans-acting* enzyme, DET22-18

Currently, the process of *in vitro* selection is only feasible with *cis*-acting enzymes, and so DEC22-18 was selected for its ability to perform an *intramolecular* self-cleavage reaction. As a *cis*-acting enzyme or pseudo-enzyme as they are commonly referred, DEC22-18 can only process a single substrate that has been covalently linked to itself. The next logical step in our pursuit of a "true enzyme" was to engineer a *trans*-acting enzyme able to perform cleavage in an intermolecular fashion. By utilizing results gathered from preliminary structural analysis, coupled with a secondary structure generated from the *mfold* program, we were able to rationally separate the enzyme domain from the substrate domain, to create a true *trans*-acting enzyme referred to as DET22-18.

DET22-18 has the ability to act as a true catalyst by processing multiple substrate molecules. Kinetic studies have indeed shown that the cleavage reaction may be described by Michaelis-Menten kinetics. The catalytic turnover rate, k_{cat}, and the Michaelis constant, K_M, were determined under multiple-turnover conditions in which an excess of substrate was used against a fixed concentration of enzyme. Based on the putative secondary structure of DEC22-18A (Figure 15A), a 34-nt substrate (S1 in Figure 15B) with an extended 3'-end was synthesized for DET22-18. The additional nucleotides were added to form a longer binding arm with the 5'-end of DET22-18 (stem 1, Figure 15B), with expectations that this would facilitate binding interactions between the newly separated substrate and enzyme domains. The concentration of DET22-18 was fixed at 5 nM, while the concentration of substrate was varied from 0 to as high as 4000 nM.

Nevertheless, the rates obtained at high substrate concentrations were somewhat inconsistent. Therefore, only rates from substrate concentrations between 0 and 2000 nM were used in plotting the Michaelis-Menten graph. The inconsistent rates observed beyond 2000 nM may be due to the presence of a high concentration of self-aggregated single-stranded DNAs. This self aggregation of single-stranded DNA may serve to prevent the enzyme from forming its catalytically active conformation, thus distorting the observed rate. The curve of best fit on the Michaelis-Menten graph was used to determine a k_{cat} of ~7.2 ± 0.7 min^{-1} and a K_M of 0.94 ± 0.19 μM for DET22-18 (Figure 15C).

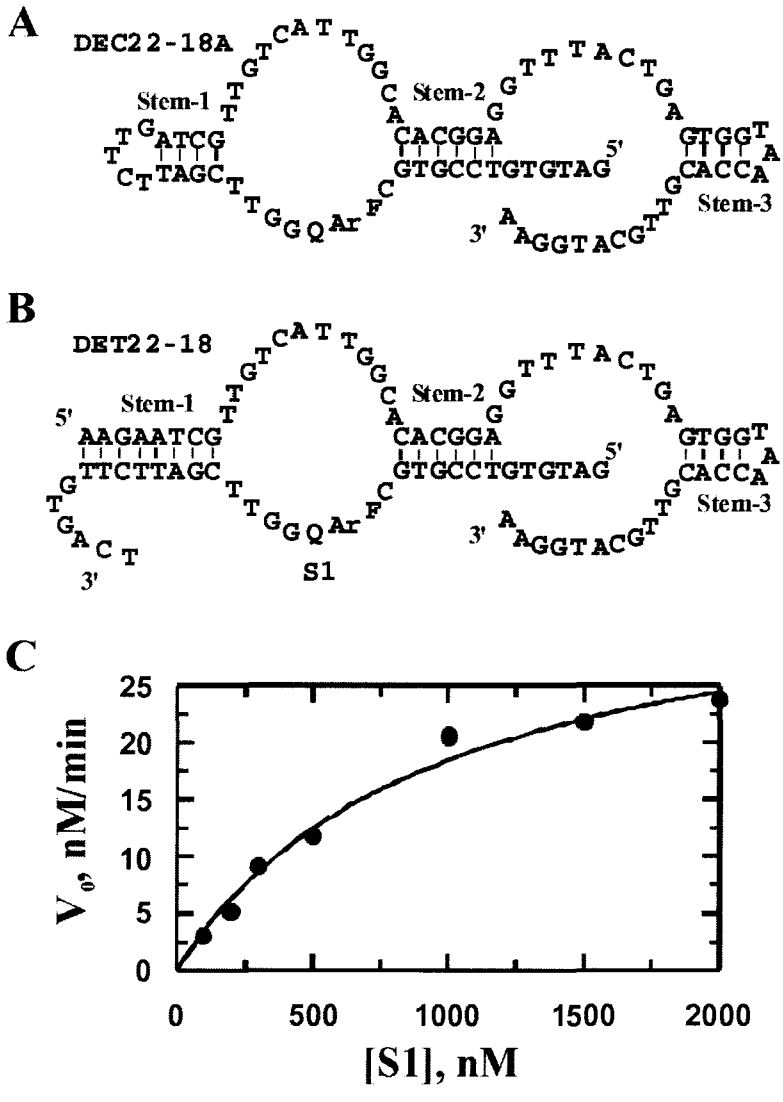


Figure 15: Structure and catalytic parameters of DET22-18. (A) A secondary structure of DEC22-18A predicted by the *mfold* program. (B) A *trans*-acting DNA enzyme, DET22-18, was designed based on the predicted secondary structure. (C) Kinetic analysis of DET22-18. The 34-nt substrate S1 was labeled at the 5'-end with [γ-^{32}P]ATP and PNK. 5 nM of DET22-18 was incubated with S1 at 100, 200, 300, 500, 1000, 1500, and 2000 nM in a reaction buffer containing 10 mM $CoCl_2$ and 5 mM $MgCl_2$ in 50 mM HEPES at pH 6.8 and 23°C. Aliquots were collected at different time points and loaded onto 10% PAGE gel for rate determinations. The rate constants at different substrate concentrations were fit to the Michaelis-Menten equation using the GraFit software. Experiments were conducted in triplicate, and the average velocity at each substrate concentration is shown in the graph.

3.1.11 Real-time signalling of DET22-18

Initial studies on the fluorescence signaling capability of DET22-18 revealed only a 4-fold enhancement in 23 hours under optimal buffer conditions (10 mM $CoCl_2$, and 5 mM $MgCl_2$ in 50 mM HEPES at pH 6.8 and 23°C) (represents in red circle in Figure 16). As it turns out, both Co^{2+} and Ni^{2+} exhibited a substantial quenching effect on fluorescein (the fluorophore used in our construct), as suggested by Kowinder et al (unpublished data). Equipped with this new data, we decided to investigate the effect of Co^{2+} on the fluorescence signalling capability of our enzyme. Reducing the concentration of Co^{2+} (1 mM) resulted in a significant fluorescence enhancement while maintaining a relatively fast catalytic rate. Interestingly, a lower concentration of Co^{2+} (0.5 mM) actually exhibited a reduced fold enhancement, suggesting that fine-tuning of the Co^{2+} concentration may play an important role in balancing efficient catalysis with large fluorescence enhancements (Figure 16). Additional experiments are required to confirm this hypothesis.

Using identical reaction buffers, the real-time signalling ability of DET22-18 was assessed under two types of conditions: an excess of enzyme over substrate (Figure 17A), and an excess of substrate over enzyme (Figure 17B). In the complete absence of DET22-18, the fluorescence system displayed steady background fluorescence (first 10 minutes of incubation). However, once the enzyme was introduced into the reaction cuvette, the fluorescence intensity increased 10-fold in 3 minutes. The reaction was monitored for 20 hours and a 17-fold enhancement was ultimately reached. Even under conditions where there was an excess of substrate over enzyme (multiple turnover), a

fluorescence intensity increase of 10-fold was observed in 10 minutes and a 15.7-fold increase observed at the end of 20 hours.

The advantage of using the enzyme under single turnover conditions is two-fold, in that it allows for rapid detection in addition to reducing the amount of substrate required in each assay. When a large fluorescence signal is desired, the enzyme can be used under multiple turnover conditions with more substrate.

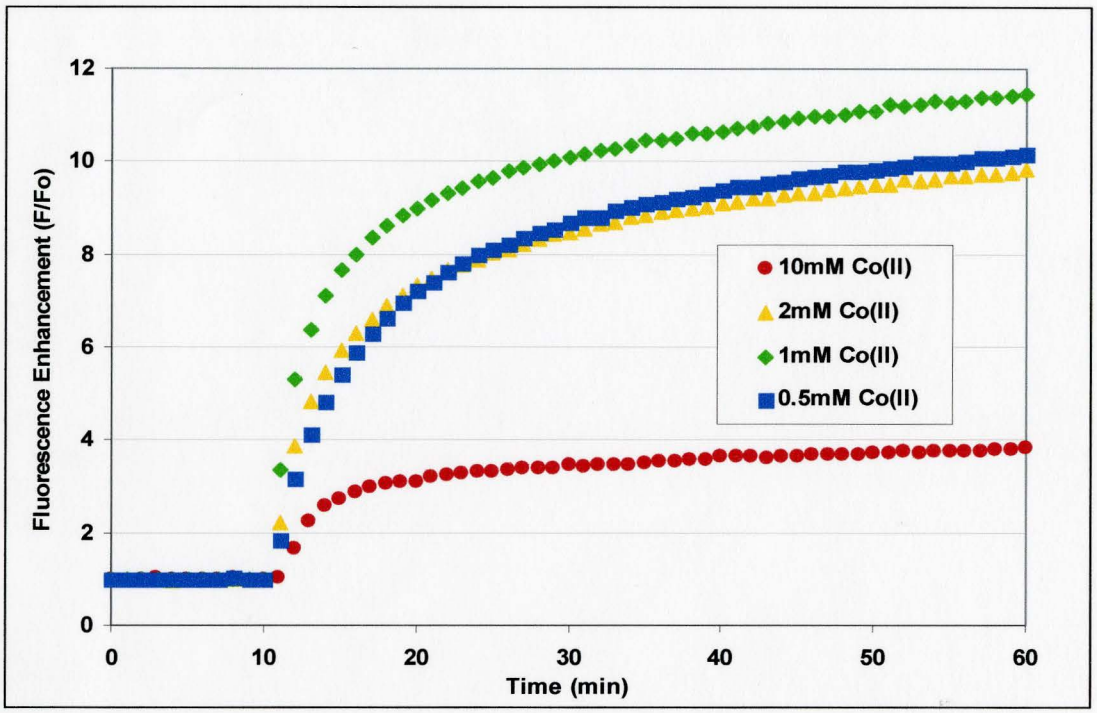


Figure 16: Real-time signalling of DET22-18 with various Co^{2+} concentrations. Fluorescence readings of each reaction mixture in the absence of DET22-18, were taken for 10 minutes. The addition of enzyme DET22-18 catalyzed RNA cleavage separating the fluorophore from the quencher, leading to an increase in fluorescence. Various concentrations of Co^{2+} were used to determine the optimal buffer conditions to yield the largest fluorescence enhancement while maintaining a fast catalytic rate.

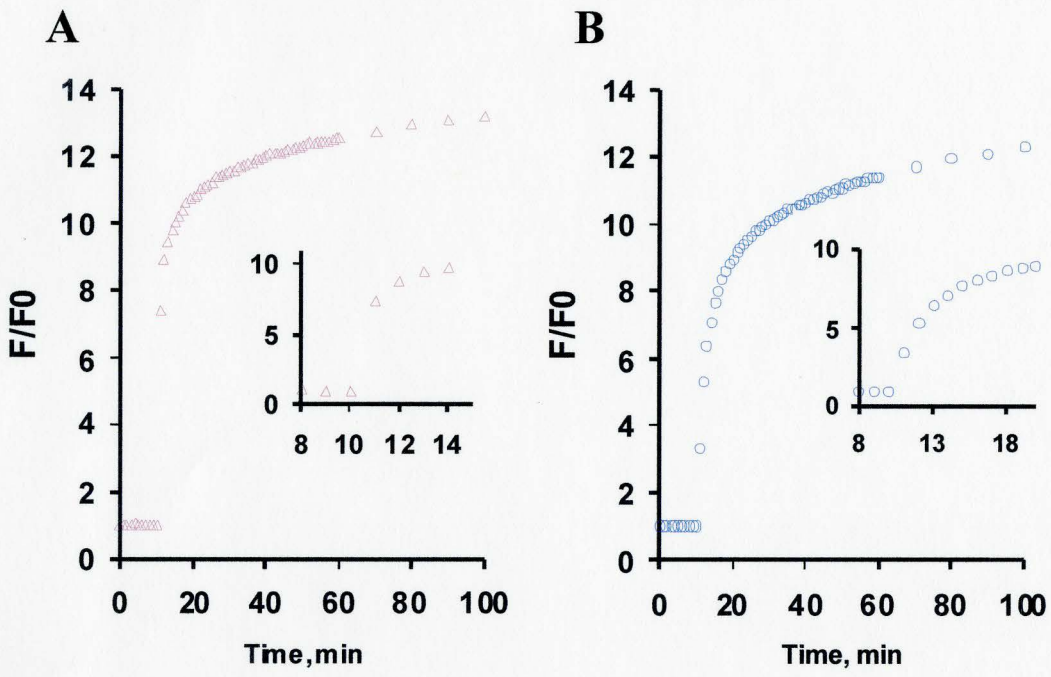


Figure 17: Real-time signalling of DET22-18 / S1 system. (A) Substrate cleavage under single turnover conditions (DET22-18:S1 = 10:1). In the absence of the enzyme DET22-18, the system exhibited steady background fluorescence (first 10 minutes of incubation). Once the enzyme was introduced to the reaction cuvette, the fluorescence intensity increased 10-fold in 3 minutes. The fluorescence intensity was recorded every minute for the first 60 minutes and every 10 minutes thereafter. The reaction was monitored for 20 hours and a 17-fold enhancement was ultimately reached. (B) Substrate cleavage under multiple turnover conditions (DET22-18:S1 = 1:10). The same protocol was used as in the single turnover reaction. The fluorescence intensity increased to 10-fold in 10 minutes and a 15.7-fold enhancement was reached at the end of 20 hours. The DNA concentrations were used as follows: S1 at 0.5 μM, DET22-18 at 5 μM in Figure 17A and 0.05 μM in Figure 17B. The reaction mixture also contained 1 mM $CoCl_2$, 14 mM $MgCl_2$ in 50 mM HEPES at pH 6.8 and 23°C.

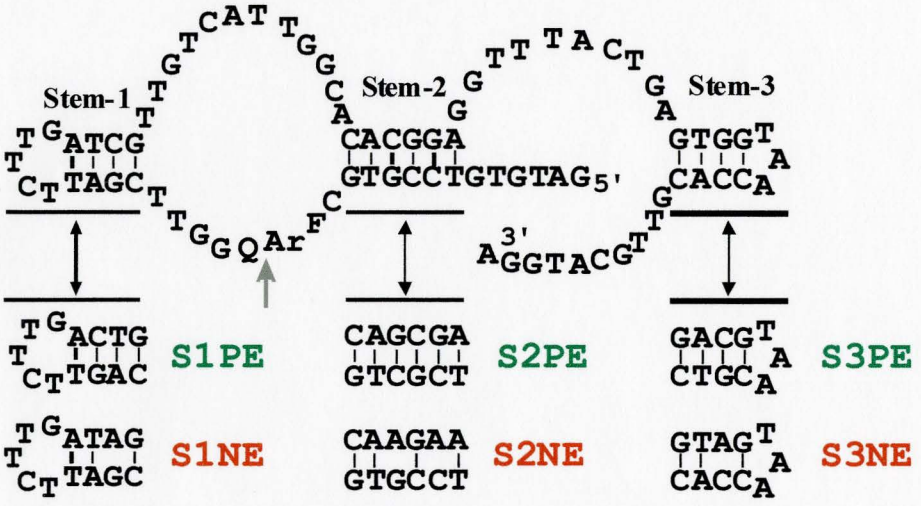
3.1.12 Preliminary examination on the predicted secondary structure of DEC22-18A

To test the predicted secondary structure of DEC22-18A obtained through *mfold*, six site-directed mutagenic constructs were chemically synthesized and subsequently assayed for their respective activities. The predicted secondary structure consisted of three stems, with base-pairing interactions occurring between the enzyme and substrate at stems 1 and 2, as illustrated in Figure 18A. Mutants were designed to test those putative pairing regions by either exchanging the original sequences for different base-pair combinations (e.g. changing A-T to G-C), or flipping the base-pair (e.g. changing A-T to T-A). Presumably, a true stem (a stand alone stem that does not significantly interact with other structural elements) would be able to tolerate these base-pair mutations, so long as the overall number of base-pairs remains constant. Conversely, if the proposed stem is not a true stem or is involved in a more complex tertiary structure like a triplex, we would expect the activity of the mutant to decrease or disappear entirely. In this experiment, mutant constructs are referred to in such a way that each name indicates the specific stem that was mutated and the predicted outcome from the introduced mutation. For example, the mutant construct designated as S1PE was mutated at putative stem 1 without compromising the overall base-pairing interactions (S1 stands for stem 1); therefore, the activity should still be maintained (P stands for positive mutation and E stands for enzyme). However, if intentional mismatches were introduced at one of the proposed stems, the mutant construct was given the name NE (N stands for negative mutation) and would be expected to have little or no enzymatic activity. The assay was

conducted in the selection buffer with 10-seconds and 30-minutes reaction times. The activity of each mutant was compared with the activity of DEC22-18A in the same selection buffer. Mutant constructs that displayed similar activities to DEC22-18A were assigned two plus signs (+ +), while mutants showing less activity than DEC22-18A were designated with a single plus sign (+). Mutants exhibiting no activity at all were assigned a minus sign (–). The data from this experiment is illustrated in a table in Figure 18B.

The results from this study were rather inconclusive. The mutants S2PE and S3PE showed no activity despite having viable base-pairing mutations, while S1NE and S3NE remained active despite having mismatch mutations. The constructs containing either base-pairing or mismatch mutations at putative stem 1 did not appear to differ in their respective activities compared with DEC22-18A. This suggests that stem 1, if it indeed exists, may not play a crucial role in facilitating the formation of the enzyme's active conformation. Conversely, any mutation introduced on either of stems 2 or 3 seemed to completely abolish the activity. This implies that the residues at stems 2 and 3, and particularly the specific residues mutated in this study, may play an integral role in catalysis. The sum of this evidence seems to suggest that the secondary structure proposed by the *mfold* program is not an accurate depiction of the true secondary structure of DEC22-18A. Further experiments were subsequently conducted to elucidate the correct secondary structure and will be discussed in section 3.3.4.

A



B

Construct	Relative Activity* 10 seconds	30 minutes
S1PE	+	+
S1NE	++	++
S2PE	-	-
S2NE	-	-
S3PE	-	-
S3NE	-	+

* ++ indicate comparable activity to the DEC22-18A
+ indicate activity less than the DEC22-18A
– indicate no activity.

Figure 18: Preliminary secondary structural study of DEC22-18A. (A) The predicted secondary structure obtained from the *mfold* program. It showed the enzymatic part of the DEC22-18A forms a stem-loop at stem 3 and interacts with the substrate through stems 1 and 2. The sequence changed for each construct is shown above, with the respective name given on the right. Each construct was ligated with the substrate, and then tested for its respective activity in the selection buffer condition (400 mM NaCl, 100 mM KCl, 7.5 mM $MgCl_2$, 5 mM $MnCl_2$, 1 mM $CoCl_2$, 0.25 mM $NiCl_2$, and 1.25 mM $CdCl_2$ in 50 mM HEPES at pH 6.8 and 23°C) with reaction time at 10 seconds or 30 minutes. (B) Percent cleavage obtained from each construct is compared with the activity from the DEC22-18A in the same selection buffer, and listed as relative activity.

3.1.13 Design of an allosteric deoxyribozyme

The preliminary structural analysis of DEC22-18A suggested that perfect sequence complementation in stem 1 (section 3.1.12) might not be necessary to maintain the catalytic activity (however this stem was later proved to be rather insignificant, refer to section 3.3.4). By contrast, stems 2 and 3 cannot be manipulated without adversely affecting the activity of the enzyme. These findings provided us with an opportunity to attach a known aptamer domain onto DEC22-18A via a destabilized stem 1 (since results in section 3.1.12 showed that stem 1 can tolerate mismatch), which was achieved by eliminating one base-pair and introducing a mismatch. Presumably, binding of the ligand to the aptamer domain would induce a conformational change that would subsequently lead to the formation of a compact structure in the aptamer region. The re-stabilization of stem 1 would then be achieved to activate catalysis.

Several versions of rationally designed constructs, each with a previously isolated DNA aptamer against ATP (Huizenga and Szostak, 1995), were tested. These constructs achieved activations ranging from 5- to 11-fold. One of the constructs (Figure 19A) was assessed for its fluorescent signalling capability under similar buffer conditions (1 mM $CoCl_2$, and 14 mM $MgCl_2$, 400 mM NaCl, and 100 mM KCl, 50 mM HEPES at pH 6.8 and 23°C) as was used for the real-time signalling of DET22-18. The rate of fluorescence intensity increase was enhanced 4-fold upon the introduction of 1 mM ATP (Figure 19B). A relatively small rate activation was observed mainly because of the rapid, ligand-independent self-cleaving nature of allosteric DEC22-18A. At room temperature, the destabilization caused by the mutation in stem 1 was insufficient to

ensure that the allosteric DEC22-18A remained inactive in the absence of ATP. It is generally known, however, that the incidence of base-pairing occurs less frequently at elevated temperatures. The higher reaction temperature should promote a less stable active conformation, and therefore lead to a lower degree of ATP-independent catalysis, that was in fact, experimentally observed. Further optimization with various Co^{2+} concentrations and reaction temperatures ultimately showed that greater rate activation, nearly 20-fold by ATP (versus GTP as a control), could be obtained with a decreased Co^{2+} concentration (0.25 mM $CoCl_2$, and 14 mM $MgCl_2$) at a temperature of 37°C (Figure 19C). However, there was a trade-off: the overall catalytic rate of the enzyme decreased, since the enzyme was less effective in maintaining its active conformation at the elevated temperature.

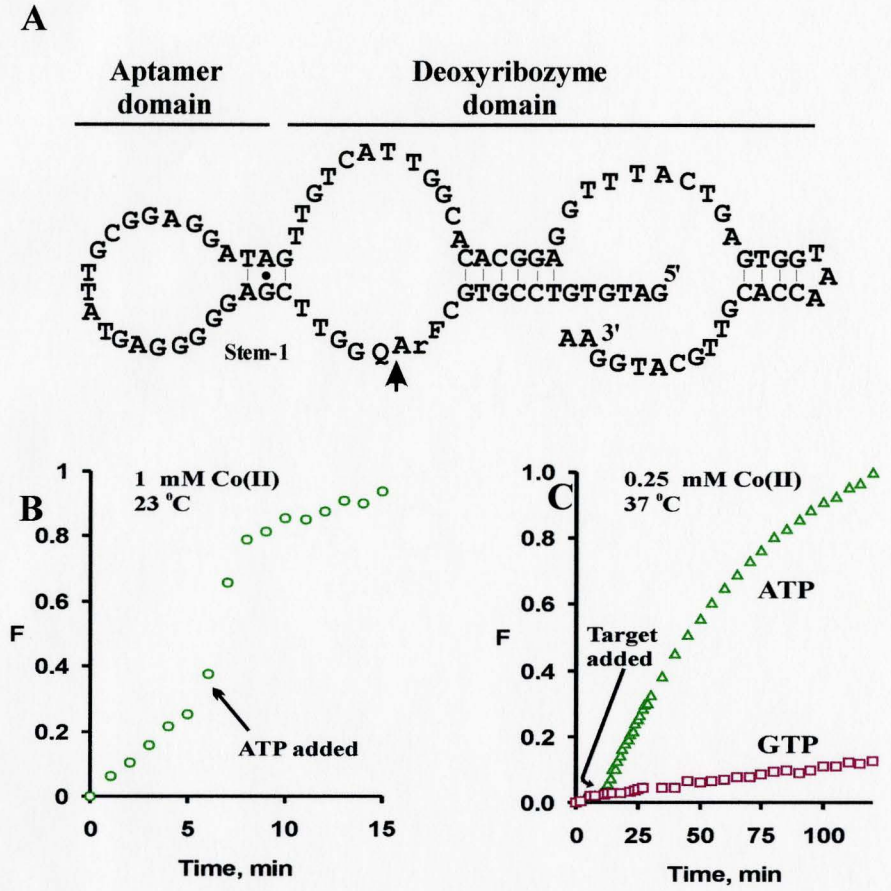


Figure 19: ATP-dependent allosteric DNA enzyme. (A) A DNA enzyme conjugated with a known ATP aptamer and the signalling DNA enzyme through a weakened stem-1. (B) ATP activation assay. The conjugated DNA was incubated at room temperature in the absence of ATP for 5 minutes, followed by the addition of ATP (final concentration = 1 mM), and a further incubation at the same temperature for 10 minutes. The reaction mixture contained 40 nM of DNA molecules in the buffer (1 mM $CoCl_2$, 14 mM $MgCl_2$, 400 mM NaCl, and 100 mM KCl in 50 mM HEPES at pH 6.8 and 23°C). (C) Activation by ATP and GTP. The DNA molecule was incubated at 37°C in the absence of ATP and GTP for 10 minutes, followed by the addition of 1 mM ATP (green triangle) or GTP (red square) and incubated for an additional 110 minutes at the same temperature. The fluorescence intensity was recorded every minute for the first 30 minutes and then every 5 minutes thereafter. The reaction conditions were the same as in (B) except for the concentration of $CoCl_2$, which was used at 0.25 mM. The fluorescence intensity was normalized as $(F - F_0) / (F_{max} - F_0)$ where F is the fluorescence reading of a solution at any given time point, F_0 is the reading of each sample made at the starting point, and F_{max} is the reading made either at 15 minutes for the sample used in (B) or at 120 minutes for the ATP-containing sample used in (C).

In summary, we have created a special RNA-cleaving DNA enzyme, DET22-18, that has a uniquely linked chemical catalysis/real-time signalling capability. It is a true enzyme with a k_{cat} of ~7 min^{-1} – the second fastest rate ever reported for a DNA enzyme. The ability of DET22-18 to generate a large fluorescence enhancement provides a useful tool to engineer potential allosteric deoxyribozyme biosensors with real-time detection capabilities. Currently one version of the signalling allosteric DNA enzyme, which specifically recognizes ATP, has been constructed. The ATP reporter has demonstrated the ability to distinguish between ATP and GTP, and performs real-time fluorescent signalling upon addition of the ligand. The success in constructing this prototype allosteric deoxyribozyme serves as a major milestone in our continuing pursuit of a deoxyribozyme-based reporter system.

3.2 Optimization of the Signalling Allosteric Deoxyribozyme by Re-selection

An ATP-dependent, allosteric deoxyribozyme was previously engineered by appending a known ATP aptamer (Huizenga and Szostak, 1995) onto the destabilized stem 1 of our original DEC22-18A enzyme (Figure 19A), as discussed in section 3.1.13. In our preliminary optimization effort, various Co^{2+} concentrations and reaction temperatures were evaluated, and a rate activation approaching 20-fold was ultimately achieved (Figure 19C). Nevertheless, the large 20-fold rate activation was acquired at the expense of the overall catalytic rate, which decreased as a result of the elevated temperature. An enzyme that had once required only a few minutes to complete the reaction, now required close to two hours.

To circumvent this problem an alternative optimization approach was adopted. Instead of trying to optimize the metal ion concentration or the reaction temperature, we directly targeted the ATP-independent catalysis for elimination. The process involved a re-selection on a newly designed DNA library, based on our existing allosteric deoxyribozyme, in which the sequences of the stem joining the aptamer and deoxyribozyme had been completely randomized. By subjecting this new DNA library to the same selection pressure used in our original selection effort (see Figure 7A), we hoped to evolve a new class of ATP- dependent allosteric DEC22-18A. To aid in this process, negative selections (explained in detailed in section 3.2.1) were judiciously incorporated as a means of removing DNA molecules that could self-cleave in the absence of ATP. Various groups have already employed this strategy with demonstrated success in reducing the level of target-independent catalysis, while increasing the rate

activation (Soukup and Breaker, 1999; Robertson and Ellington, 2000; Levy and Ellington, 2002). The combination of rational design and *in vitro* selection makes for a very powerful strategy in allosteric deoxyribozyme engineering.

3.2.1 Library design and re-selection

The re-selection was performed on a newly designed DNA library, in which 3 nucleotides in the stem joining the aptamer and deoxyribozyme regions were 100% randomized (sequence in blue, Library A1, Figure 20). Since the substrate side of stem 1 was composed of a non-variable sequence (due to the exorbitant cost of both randomized and fluorophore/quencher labeled oligonucleotides), two additional libraries were chemically synthesized to increase the diversity of the starting DNA pool. Library A2 consisted of five 100% randomized nucleotides located on the enzyme side of stem 1, while library A3 had two more randomized nucleotides extending into the enzymatic core. To increase the diversity of the starting pool even further, partial mutations at a frequency of 24% per nucleotide were incorporated into the proposed catalytic core (sequence in red, Figure 20).

Prior to each round of positive selection, a 20-hour negative selection was carried out in the selection buffer (400 mM NaCl, 100 mM KCl, 7.5 mM $MgCl_2$, 5 mM $MnCl_2$, 1 mM $CoCl_2$, 0.25 mM $NiCl_2$, and 1.25 mM $CdCl_2$ in 50 mM HEPES at pH 6.8 and 23°C) in the absence of the effector molecule, ATP. This negative selection step removes the majority of molecules that can self-cleave without ATP. A temperature ramping protocol was also introduced, in which the DNA molecules were heated up to 60°C for 30 seconds followed by incubation at 25°C for 30 minutes, before being heated up to 60°C once again. This temperature cycling profile was repeated 40 times in each round of the negative selection. Generally, most deoxyribozymes fail to catalyze substrate conversion to 100% completion, because there is some tendency for the enzyme to fold into an

inactive conformation and/or to base-pair with other DNA molecules rather than self-fold. The temperature ramping was implemented as a means to combat this occurrence, by denaturing any incorrectly folded molecules and allowing them to refold correctly. This protocol should help to eliminate any ATP-independent self-cleaving molecules from fortuitously surviving the entire selection. If for example, the starting pool was composed of 50% of the unwanted ATP-independent catalysts, but only 40% of those molecules folded properly with one heating/cooling process, then a conventional negative selection would only eliminate 80% of those molecules (40% out of the 50%). As a result, the remaining 20% of unwanted ATP-independent catalysts would survive along side our desired ATP-dependent molecules, and interfere with their specific isolation. If, however, the heating/cooling process is performed 40 times per round of selection, the elimination process would also be repeated 40 times, and approximately 100% of the unwanted molecules would be eliminated (1st round of heating/cooling gets rid of 80%, 2nd round of heating/cooling gets rid of 80% more of the left over and results in 96%, 3rd round of heating/cooling gets rid of 80% more of the left over from the 2nd round and results in 99.2%, etc...).

The uncleaved population, consisting of molecules that were either ATP-dependent catalysts or simply non-catalytic in nature, was then subjected to a positive selection in the presence of 1 mM ATP (final concentration). Aside from the changes mentioned above, most of the *in vitro* selection process was very similar to the one described in Figure 7A, section 3.1.3. A total of nine rounds of selection were performed, with each round consisting of both a negative and positive selection from

generation 4 onwards. Figure 21A is a 10% PAGE gel that shows the ATP selectivity of the pools before and after going through the G9 negative selection. The precursor band in Figure 21A consisted of three libraries of two different sizes, 115-nt and 117-nt, from which the catalytic DNA molecules capable of self-cleavage yielded the 100-nt and 102-nt product bands (however the gel shown here was not run long enough to resolve those double bands). The DNA pool from the G8 positive selection (equivalent to "before G9 negative selection" in Figure 21A) already appeared to possess the ability to selectively self-cleave, only in the presence of ATP (the first and second lanes in Figure 21A). Nevertheless, an additional round of positive selection (G9) was performed to corroborate the results from G8. As shown in Figure 21A, the positive selection (third to sixth lanes in Figure 21A) exhibits the same selectivity and percent-cleavage as the pool derived from the G8 positive selection (first and second lanes in Figure 21A). Results of the percent cleavage obtained from the G4 to G9 selections were plotted in Figure 21B. The percent cleavage of a 20-hour negative selection rapidly decreased from round to round, while the percent cleavage of a 2-minute positive selection gradually increased and seemed to plateau after G8. To obtain a better perspective on the selection progress, the k_{obs} (in min^{-1}) of both negative and positive selections were estimated and their ratio was defined as the fold ATP activation (blue line, Figure 21B). As much as 800-fold activation was achieved by the 9^{th} round of selection. Even though the re-selection seemed to eliminate the majority of the ATP-independent self-cleaving molecules, the resulting ATP-dependent molecules failed to yield self-cleavage beyond 10%, even when allowed to proceed overnight. Due to the time constraints imposed on this project,

restarting the re-selection or step-by-step trouble shooting was simply not feasible, and so the decision was made to proceed with the cloning and sequence analysis of the selected population.

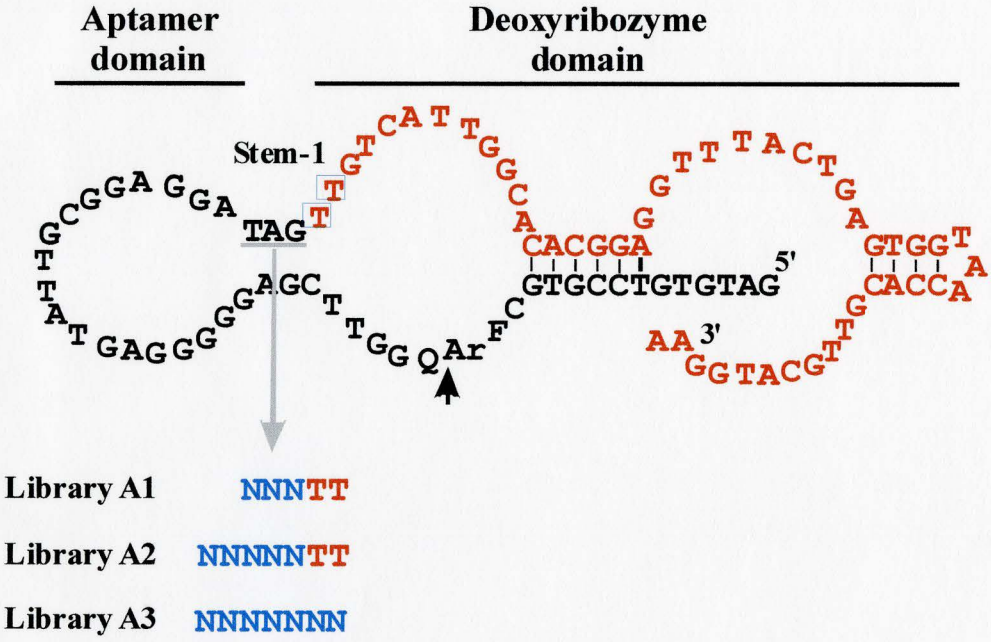


Figure 20: Library design for the ATP-dependent allosteric deoxyribozyme re-selection. Library A1 consisted of three 100% randomized nucleotides at the proposed stem 1, library A2 of five 100% randomized nucleotides, and library A3 of two more 100% randomized nucleotides into the proposed catalytic core in addition to the five as in library A2. The completely randomized nucleotides at stem region are shown in blue, while the partial mutations with a frequency of 24% per nucleotide at part of the purposed catalytic core are shown in red.

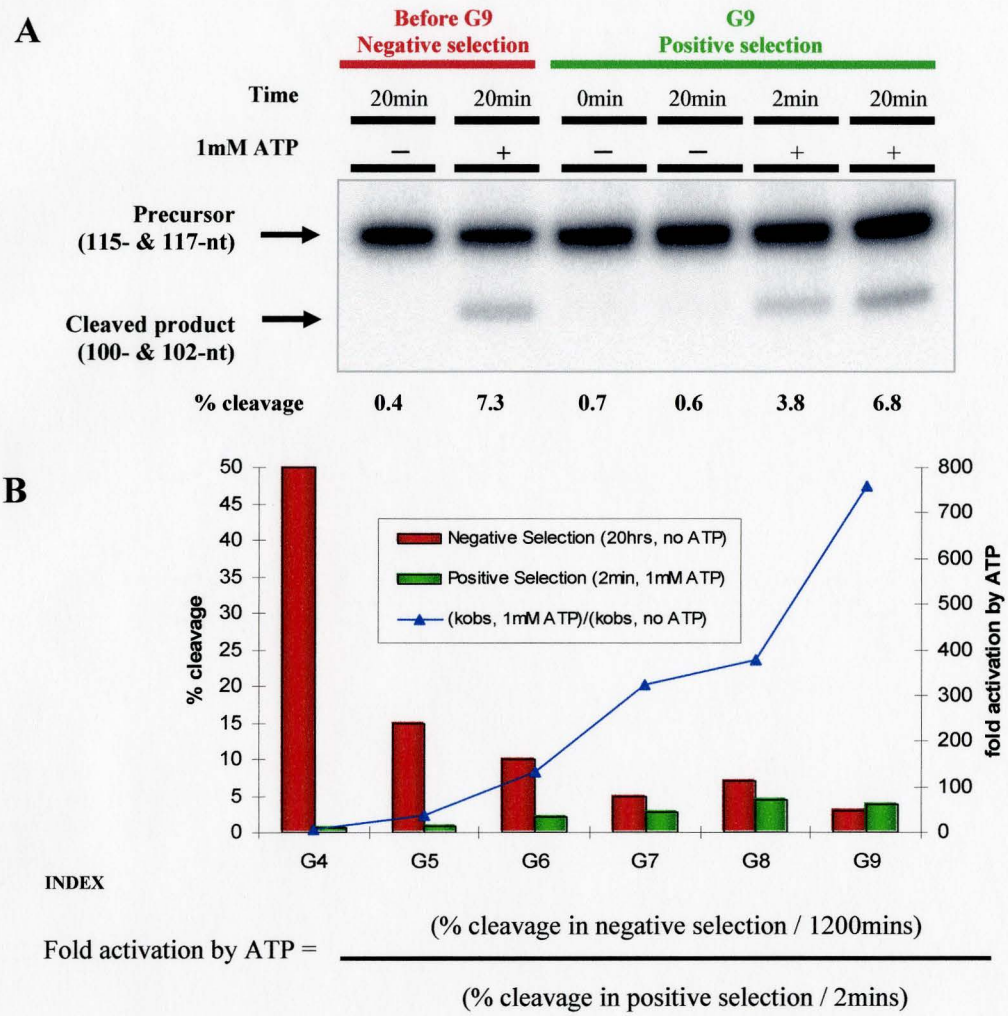


Figure 21. *In vitro* selection of ATP-dependent allosteric deoxyribozyme. (A) The selection gel with negative and positive selection pools from G9 population showing their self-cleaving activities. The precursor band consisted of three libraries in two different lengths (115-nt and 117-nt); thus resulted in two different sizes of the cleaved products (100-nt and 102-nt). The reaction mixture contained the corresponding DNA population in selection buffer (400 mM NaCl, 100 mM KCl, 7.5 mM $MgCl_2$, 5 mM $MnCl_2$, 1 mM $CoCl_2$, 0.25 mM $NiCl_2$, and 1.25 mM $CdCl_2$ in 50 mM HEPES at pH 6.8 and 23°C), in the presence or absence of 1 mM ATP. Each reaction was stopped at the time specified with the addition of EDTA to reach 30 mM. (B) A total of 9 rounds of *in vitro* selection were carried out, including both the negative (20 hours in the absence of ATP) and the positive (2 minutes in the presence of 1 mM ATP) selections for every round of selection from generation 4 onward. The red bars represent the % cleavage at each round of negative selection, while the green bars represent the positive selection. An approximated k_{obs} of each selection was calculated. The fold-activation by ATP was determined by the above formula and represented by a blue line.

3.2.2 Cloning and Sequencing

The G9 two-minute population was chosen for cloning and sequence analysis. Figure 22A shows three randomized libraries that were used in the starting pool of the re-selection. The clones from generation 9 yielded 25 sequences, which exhibited no readily distinguishable features, making the alignment and classification of these clones into individual classes extremely difficult (Figure 22B). One sequence (sequences G9a I-8) was identified as a possible descendant from the original library A1 (with the possibility of G9a II-13), while two sequences (sequences G9a I-15 & G9a I-1) were derived from library A2. The majority of sequences seemed to be derived from library A3, based on the way those sequences were aligned in Figure 22B. There seems to be a random insertion and deletion of nucleotide(s) occurring throughout the entire length of the randomized region.

The original purpose of mutating the nucleotides that connected the aptamer to the enzymatic domain, was to identify a possible "communication module" similar to the ones described by Breaker's group for the hammerhead ribozyme (Soukup et al., 2000). Through *in vitro* selection, their group was able to identify specific bridging sequences, a.k.a. the communication module, which "transduced" the signal of ligand-binding from the aptamer, to the enzymatic domain. Once a communication module is identified, one can easily utilize this pre-existing module to simplify the engineering process required to connect an aptamer domain onto the existing enzyme (Breaker, 2002). In our case, however, there was no common bridging sequence that could be easily identified from the alignment. Either more clones have to be sequenced to identify a pattern of possible

communication module(s), or a new library containing mutated bridging sequences in an otherwise identical library design will have to be used in another re-selection effort. Further characterization of individual clones would also have to be carried out in order to identify potential communication modules.

The activity of the re-selected allosteric deoxyribozymes decreased significantly. The parent sequence (construct shown in Figure 19A) was capable of 15% self-cleavage within 30 minutes in the selection buffer. By contrast, only about 10% self-cleavage was observed with an overnight reaction time for the G9 population. The faster reaction observed for the parent sequence may be due to ATP-independent self-cleavage, which is consistent with the high background cleavage observed in the absence of ATP (Figure 19B). Another possible explanation for G9's low cleavage rate is that the mutations acquired during the re-selection process were crucial for allosteric regulation, but had an adverse effect on catalysis. Although these mutations may have served to reduce the ATP-independent self-cleavage, they also may have altered the optimally active conformation of the original DEC22-18A enzyme leading to the observed decrease in rate. This problem may have been preventable, had the catalytic core not been partially randomized, in addition to the bridging sequences. However, the starting pool would have had very limited sequence diversity. A library with three 100% randomized nucleotides gives rise to only 64 different sequences, which may have been inadequate to evolve a selective allosteric deoxyribozyme. A third possibility that may account for G9's low rate is that the re-selected allosteric deoxyribozymes employ a self base-paring strategy as a mean to block the enzymatic domain when ATP is not present. Base pairing

between strands is certainly the most prevalent interaction utilized by nucleic acid molecules. Unfortunately, this strategy will have adverse consequences on the catalytic rate. It would be more efficient for the enzyme to employ a strategy based on an alternating tertiary structure, between active and inactive conformations. Not surprisingly, several folded structures obtained using the *mfold* program showed long stretches of base-pairing, either at the substrate cleavage side or at the 3'-end of the catalytic domain (this part was later proven to be crucial for catalysis, see discussion in section 3.3.4). Their respective melting temperatures (T_m) were calculated and some are as high as 56°C. This implies that for catalysis to occur, the allosteric enzyme not only has to capture the ATP but also discards its self-annealing sequences in order to have a chance to form an active conformation. A higher reaction temperature would facilitate the departure of these self-annealing, inhibiting sequences. However, a higher reaction temperature would also lead to the reappearance of ligand-independent background activity or a reduction in the kinetic performance, and so one must endeavour to balance these conflicting factors.

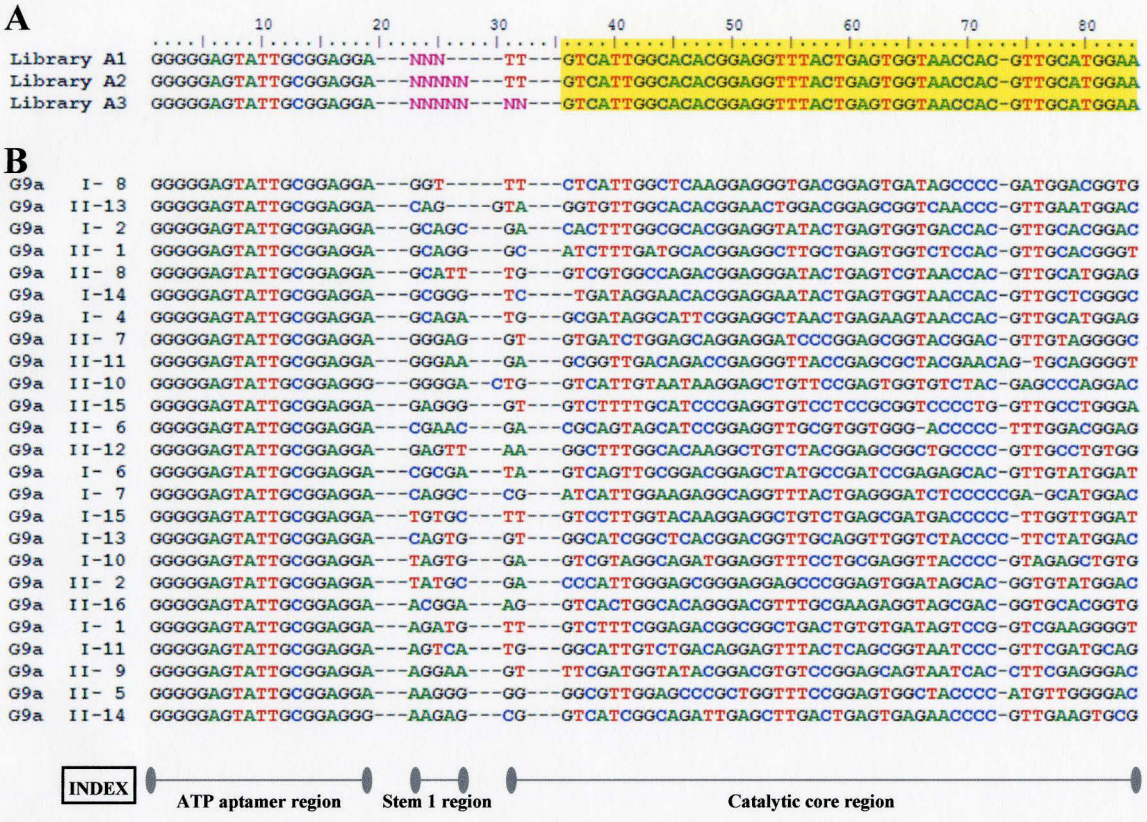


Figure 22: Sequencing results on clones obtained from generation 9 of re-selection for the ATP-dependent allosteric deoxyribozymes. (A) Sequences of randomized libraries A1, A2 and A3 used in the *in vitro* selection. Letters in pink represent the randomized nucleotide. Library A1 consisted of three 100% randomized nucleotides at the proposed stem 1, library A2 of five 100% randomized nucleotides, and library A3 of two more 100% randomized nucleotides into the proposed catalytic core in addition to the five as in library A2. The yellow shaded sequences were partially randomized with a frequency of 24% per nucleotide. (B) For G9 pool, every sequence listed here is unique. A total of 25 sequences were identified. All sequences were aligned according to the index given.

3.3 Refinement of Catalytic Motifs with Re-selection

One of the many objectives we wished to accomplish through this re-selection was to further define sequences essential for catalysis, with prospects of improving the kinetic rate and lowering the Co^{2+} ion requirement (it is now known that high Co^{2+} concentration causes significant fluorescence quenching). With a folding program like *mfold*, which merely provides a secondary structure based on base-pairing and thermodynamic energy minimization, it is sometimes difficult to have a real grasp on the true structure. One approach adopted by various groups studying deoxyribozymes is to construct a partially randomized library, based on a previously selected parental deoxyribozyme, to repeat the selection. The new library consists of many molecules that closely resemble the original enzyme, but have random point mutations throughout the entire sequence. Any molecule with a mutation at a position that is important for catalysis would exhibit low or zero activity, and thus, be unable to survive the selection process. This effect can then be verified through sequence analysis, which would show little or no tolerance for mutation at this particular position. Sequence alignment allows one to get a better perspective on which nucleotides and which positions are essential for catalysis. By utilizing this information a more meaningful secondary structure model can be constructed.

One of the problems with *in vitro* selection is that selection alone does not provide the opportunity to search all sequence space, but rather enriches those active molecules already present within the starting pool of sequences (Bittker et al., 2002). The term sequence space refers to all possible sequence variants in a given library. This led to our

second and third objectives, which were to evolve more efficient catalysts than the parental DEC22-18, and to isolate possible variants with a lower requirement for Co^{2+} ions. By constructing a partially mutagenized library based on the 60-nt DEC22-18A (the optimized version of the DEC22-18), we were able to explore more of the sequence space of this 60-nt oligonucleotide and also expedite the selection process.

A speed-selection strategy was carried out on this new library, in which random mutations were introduced at a frequency of 24% per nucleotide position throughout the entire enzyme (Figure 23). In order to increase the selection pressure, a shorter reaction time (started from 10 minutes) was used at the beginning of the re-selection, along with a reduction in Co^{2+} ion concentration once the product band was visible. As carried out in the original selection (section 3.1.3), the incubation time was gradually reduced to yield the most catalytically efficient sequence(s). Sequence analysis of the terminal population was used to define the essential sequences in the catalytic core in addition to revealing more structural requirements for this enzyme.

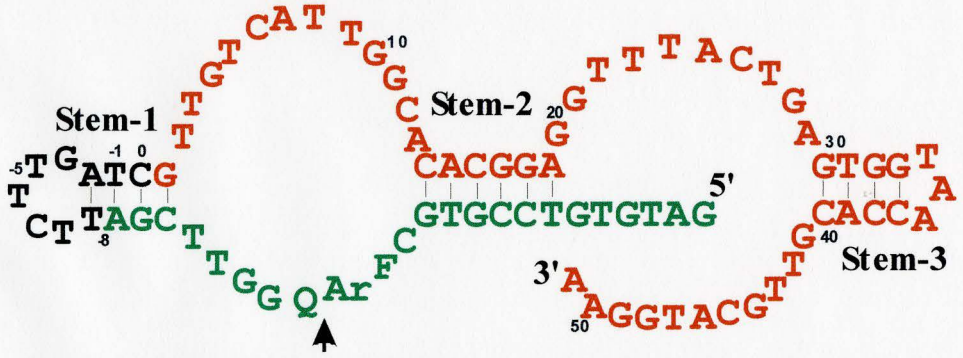


Figure 23: Library design based on the original DEC22-18A for the re-selection. Sequence shown in red, from position 1 to 51 of the enzyme region, were partially randomized in a frequency of 24% per nucleotide position throughout the proposed catalytic core. Total of 51 out of 60 nucleotides were partially randomized. Sequence in green represents the substrate region.

3.3.1 Re-selection on a partially randomized library based on DEC22-18A

Most of the procedures employed during the re-selection were similar to the ones described in section 3.1.3, except for the use of a different 3' primer to avoid cross contamination with the original selection, as well as the changes mentioned in the following paragraph. In order to increase the selection pressure, the re-selection started with a 10 minute incubation in the selection buffer (400 mM NaCl, 100 mM KCl, 7.5 mM $MgCl_2$, 5 mM $MnCl_2$, 1 mM $CoCl_2$, 0.25 mM $NiCl_2$, and 1.25 mM $CdCl_2$ in 50 mM HEPES at pH 6.8 and 23°C). The majority of sequences in this new library closely resembled the sequences of DEC22-18A, since each nucleotide position had a 76% probability of remaining the same as the wild-type. We predicted that a few rounds of selection would be sufficient to allow for the emergence of catalysts. As expected, a 2% self-cleavage was observed at the second round of selection (G2) after 10 minutes. When 8% self-cleavage could be seen within 1 minute in G3, this pool was split into two different subsets of selection. One subset, referred to as the speed selection, was conducted in the same selection buffer with the incubation time gradually reduced from 10 minutes to 1 second (Figure 24A). When no significant progress was observed for four rounds after the reaction time had been reduced to 1 second, the selection was stopped and populations from G4 (1 minute reaction time), and G8 (1 second reaction time) were cloned and sequenced for further analysis.

The other selection subset, referred to as the lower Co^{2+} selection, was conducted under a reduced Co^{2+} ion concentration (400 mM NaCl, 100 mM KCl, 0.25 mM $CoCl_2$, and 14.75 mM $MgCl_2$ in 50 mM HEPES at pH 6.8 and 23°C). Although Co^{2+} is the most

favourable divalent metal ion for catalysis by DEC22-18A, it adversely quenches fluorescence when used in high concentration (> 1mM). Nevertheless, the enzyme DEC22-18A does not reach its full capacity unless the concentration of Co^{2+} ions reaches 10 mM (see the titration curve with $CoCl_2$ on Figure 10C). As discussed briefly in section 3.1.11, a fine-tuning in Co^{2+} concentration could play an important role in balancing efficient catalysis with large fluorescence enhancement. By performing the re-selection under a lower concentration of Co^{2+} ions (0.25 mM instead of 1 mM), we anticipated the isolation of DEC22-18A variants that would perform efficient catalysis with a reduced requirement for Co^{2+} ions. After 8 more rounds of selection, the self-cleavage activity seemed to reach a plateau and hence the selection was terminated. The population from G11 (1 second reaction time) was cloned and sequenced for further analysis.

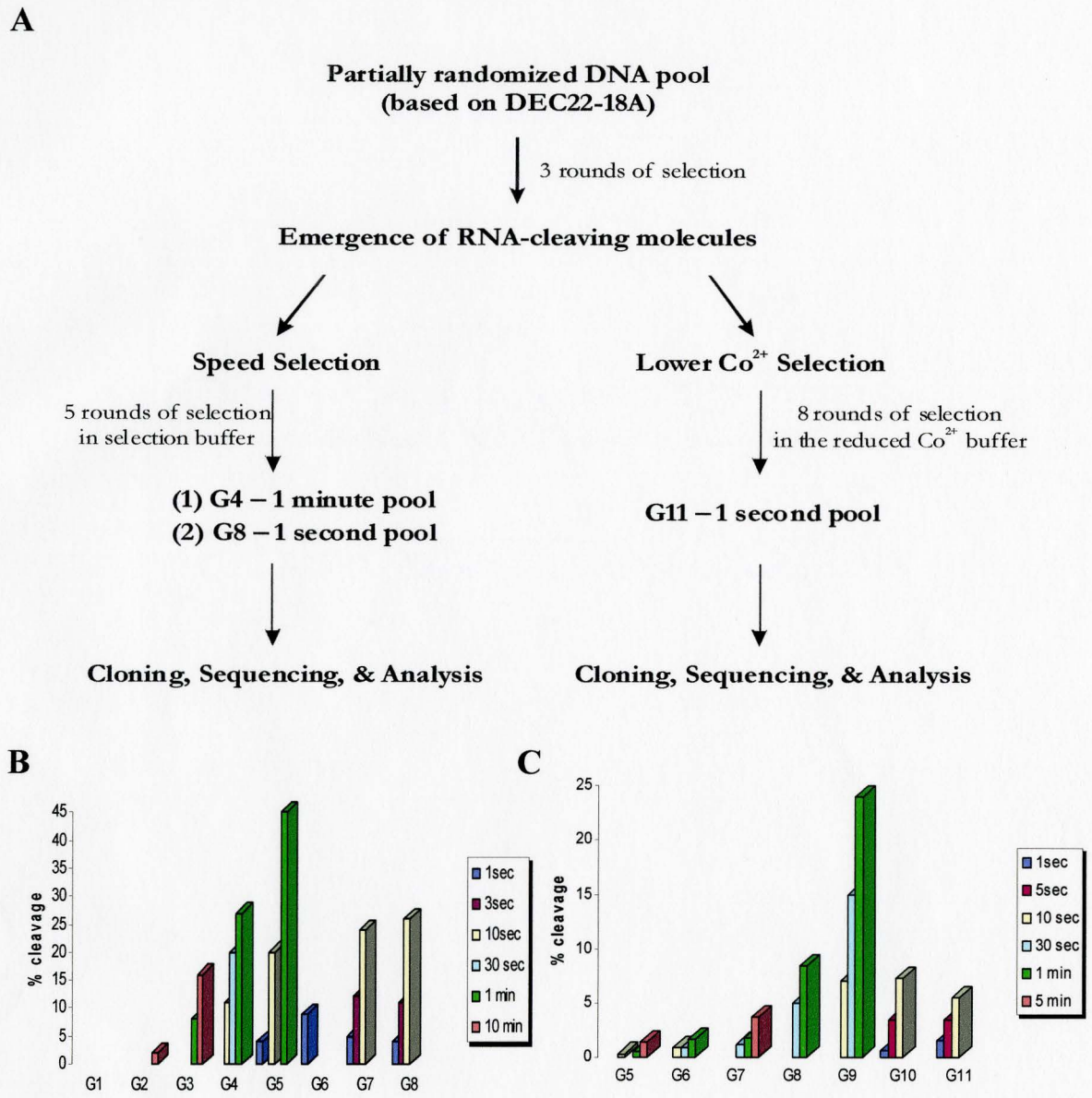


Figure 24: Progress of the re-selection. (A) The scheme of re-selection. A partially randomized pool, based on parental DEC22-18A, was used to start the re-selection. After three rounds of the selection, the pool was split into two subsets of re-selection: speed selection and lower Co^{2+} selection. (B) The progress of the speed selection. A total of 8 rounds of selections were carried out in 400 mM NaCl, 100 mM KCl, 7.5 mM $MgCl_2$, 5 mM $MnCl_2$, 1 mM $CoCl_2$, 0.25 mM $NiCl_2$, and 1.25 mM $CdCl_2$ in 50 mM HEPES at pH 6.8 and 23°C,. A reaction time of 10 minutes was used for the first three rounds of selection, and was reduced gradually to yield the most catalytically efficient sequence(s). (C) The progress of the lower Co^{2+} selection. A total of 11 rounds of selections were carried out in the presence of 400 mM NaCl, 100 mM KCl, 0.25 mM $CoCl_2$, and 14.75 mM $MgCl_2$ in 50 mM HEPES at pH 6.8 and 23°C. The selection was continued until the activity plateau.

3.3.2 Cloning and sequencing on the speed selection and lower Co^{2+} selection

The populations from G4 (1 minute reaction time), and G8 (1 second reaction time) were cloned and sequenced for further analysis. The clones from generation 4 yielded 35 sequences (Figure 25B), while the ones from generation 8 yielded 16 (Figure 25C). When compared with the parental DEC22-18A, all isolates from both populations showed remarkable similarity among sequences, considering that nucleotides from position 1 to 51 (nucleotides numbered and shown in red in Figure 23) had been partially mutagenized prior to the selection. Sequence alignment was performed on isolates from the G4 population and a more detailed analysis is discussed in the following section (section 3.3.3).

As for the lower Co^{2+} selection, the G11 one-second population was chosen for sequence analysis. The goal of this subset of re-selection was to isolate a variant of the DEC22-18A which could perform efficient catalysis with a lower requirement for Co^{2+}, by conducting the selection in a reduced Co^{2+} environment (0.25 mM instead of 1 mM). The activity stopped improving and reached plateau from G9 to G11, all exhibited barely 1% self-cleavage at one second. Two possible outcomes had been speculated, before the sequencing results became available. One was that even under this "harsh" environment (lower Co^{2+} means lower activity) the same or similar variants of DEC22-18A would survive through re-selection. The other possibility was that the partially mutagenized library would yield different sequences that were capable of performing efficient catalysis with lower Co^{2+} requirement.

Sequencing of those clones from generation 11 yielded 19 sequences. Those sequences were subsequently grouped into 4 classes based on their sequence homology (Figure 26B, C, D, and E). To be grouped in a class, sequences had to possess at least 80% homology among them. Sequences in class I (Figure 26B) appeared to be very similar to the parental DEC22-18A. However, sequences in class II (Figure 26C), III (Figure 26D), and IV (Figure 26E) not only appeared to be different from DEC22-18A, but also had quite a few guanine nucleotide substitutions at various positions. Preliminary examination using buffers with different Co^{2+} concentration seemed to indicate that there is no difference in self-cleavage activity of representative clones from each class. Even though further study is necessary to confirm this result, there is a possibility that our attempt to isolate sequence variants with lower Co^{2+} requirement had not been as successful as we had anticipated. In the mean time, we decided to focus our effort to study the structural aspect of DEC22-18A. A thorough understanding of the secondary structure will greatly facilitate our effort in engineering future allosteric deoxyribozymes using DEC22-18A.

A

```
                    0         10        20        30        40        50
                    ...|....|....|....|....|....|....|....|....|....|....|.
G22.18A(WT)         TTCTTGATCGTTGTCATTGGCACACGGAGGTTTACTGAGTGGTAACCACGTTGCATGGAA
```

B

```
G4s    II- 9        TTCTTGATCGGTGTCATTGGCACACGGAGGTTTACTGAGTGGTAACCACGTTGCAGGGAA
G4s    II-13        TTCTTGATCGTTGTCATTGGCACACGGAGGTTTACTGAGTGGTAACCACGTTGCATGGAA
G4s    II- 4        TTCTTGATCGTTGGCATTGGCACTCGGAGGTTTAATGAGAGGTAACCACGTTGCACGGAA
G4s    II- 2        TTCTTGATCGTTGTCATTGGTCTAAGCATGTTTCCGGGGCGGTAAACGCGTTGGATGGAA
G4s    II- 3        TTCTTGATCGTTGCAATTGGCACACGGAGGGATACCGAGTGGTAACCACGTTGCATGGAA
G4s    II-11        TTCTTGATCGGTGGGATGGGCACACGGAGGTTCACCGAGTGGTAACCACGTTGCACGGCA
G4s    II-12        TTCTTGATCGTTGTCTTTGGCACACGGAGGTTCACTGAGTGGTAACCACGTTGCAGGGCT
G4s    II-14        TTCTTGATCGTTGTCATTGGCACACGGAGGTTTACCGAGTGGTAACCACGTTGCATGGAC
G4s    II-15        TTCTTGATCGGTGGCAATGGCACACGGAGGTTGACCGAGTGGTAACCATGTTGCATGGCC
G4s    II-17        TTCTTGATCGTTGGCTTGGGCACACGGAGGTGTACCGAGTGGTAACCACGTTGCATGGCA
G4s    II-18        TTCTTGATCGTTGCCTTGG-CGCACGGAGGTGTACCGAGTGGTAACCACGTTGCATGGGA
G4s    II-19        TTCTTGATCGTTGTCATTGGCACACGGAGGTGGACTGAGTGGTAACCACGTTGCGCGGAA
G4s    II-21        TTCTTGATCGTTGCTCTTGGCACGCGGAGGCTTACTGAGCGGTAACCACGTTGCATGAAA
G4s    II-22        TTCTTGATCGTTGTCAATGGCACACGGAGGTTTACTGAGTGGTAACCACGTTGCATGGAC
G4s    II-23        TTCTTGATCGTTGTCAATGACACACGGAAGTTTACTGAGTGGTAACCACGTTGCAGGTCA
G4s    II-26        TTCTTGATCGTTGTCTTCGGCACACGGAGGTTTACTGAGTGGTAACTACGTTGCATGGGA
G4s    II-27        TTCTTGATCGTTGCCATTGGCACGCGGAGGTTCAATGAGCGGTAACCACGTTGCACGGAA
G4s    II-28        TTCTTGATCGTCGGCATTGGCGCACGGAGGGTTACTGAGTGGTAACCACGTTGCATGGTA
G4s    II-30        TTCTTGATCGTTGGTATTGCCACACGGAGGCTTGCTGAGTGGTAACCACGTTGCATCGGA
G4s    II-31        TTCTTGATCGTTGTCACTGGCACACGGAGGTTTACTGAGTGGTAACCACGTTGCATGGCG
G4s    II-32        TTCTTGATCGCTGACTTTGGAACACGGAGGTTTACTGAGTGGTAACCACGTTGCATGTAG
G4s    II-33        TTCTTGATCGGTGACATAGGCAGACGGAGGTTTACTGAGTCGTAACCACGTTGCATGGTA
G4s    II-35        TTCTTGATCGTTGTTATTGGCACACGGAGGTATACCGAGTGGTAACCACGTTGCATGCAA
G4s    II-36        TTCTTGATCGTTGGCATTGGCACACGGAGGCTCACCGAGTGGTAACCACGTTGCACGGAA
G4s    II-37        TTCTTGATCGTTGTCCTTGGCACACGGAGGGTTACTGAGTGGTAACCACGTTGCACGGTG
G4s    II-38        TTCTTGATCGTTGGCTCTGGCACACGGAGGTTTTCTGAGTGGTAACCACGTTGCATGGAA
G4s    II-42        TTCTTGATCGGTGACATGGGCACACGGAAGTTAACTGAGTGGTAACCACGTTGCACGGAA
G4s    II-45        TTCTTGATCGTTGCAACTGCTACACGGAGGTTTACCGAGTGGTAACCACGGTGCGCGGAG
G4s    II-46        TTCTTGATCGTTGACATTGGCACACGGAGGTTTACTGAGTGGTAACCACGTTGCAGGGAA
G4s    II-47        TTCTTGATCGGTGTCATAGGCACACGGAGGTTTACCGAGTGGTAACCACGTTGCATGGAC
G4s    II-48        TTCTTGATCGTTGCTAATGGCACACGGAGGTTTACTGAGTGGTAACCACGCTGCAGGGAG
G4s    II-50        TTCTTGATCGTTGGCCCCGGCACACGGAGGTTTACCGAGTGGTAACCACGTTGTATGGAA
G4s    II-52        TTCTTGATCGCTGCCTGTGGCACACGGAGGTATAGCGAGTGGTAACCACGTTGCATGGAT
G4s    II-54        TTCTTGATCGTTGGCATTGGCACACGGAGGTTAGCTGAGTGGTAACCACGTTGCACGGAA
G4s    II-55        TTCTTGATCGTTGTCATTGGCACACGGAGGTTTACAGAGTGGTAACCACGTTGCACGGAA
```

C

```
G8s     I- 1        TTCTTGATCGTTGGTATTGCCACACGGAGGTTCACTGAGTGGTAACCACGTTGCGCGGCA
G8s     I- 2        TTCTTGATCGTTGGCAGTGCCACACGGAGGGTTGCCGAGTGGTAACCACGTTGCGCGGAA
G8s     I- 5        TTCTTGATCGTCGGCATTGGCGCTCGGAAGTGTGCTGAGAGGTAACCACGTTGTAGGGCA
G8s     I- 6        TTCTTGATCGTTGGCATTGGCACACGGAGGTTTACCGAGTGGTAACCACGTTGCACGGGA
G8s     I- 7        TTCTTGATCGTTGGCATTGCCACTCGGAGGTTTACTGAGAGGTAACTACGTTGCGCGGGT
G8s     I- 8        TTCTTGATCGTTGACATTGGCACACGGAGGTCTACCGAGTGGTAACCACGTTGCATGGAC
G8s     I- 9        TTCTTGATCGTTGGCAATGGCACACGGAGGTTTACCGAGTGGTAACCACGTTGCATGGCA
G8s     I-10        TTCTTGATCGTTGACATTGGCACACGGAGGTTTACTGAGTGGTAACCACGTTGCATGGAG
G8s     I-11        TTCTTGATCGTTGTCATTGGCACACGGAGGTTGACCGAGTGGTAACCACGTTGCACGGGA
G8s     I-13        TTCTTGATCGTTGGCATTGCCACACGGAGGTTTACCGAGTGGTAACCACGTTGCGCGGAG
G8s     I-14        TTCTTGATCGTTGGCTTTGCCACACGGAGGTTTACTGAGTGGTGACCACGTTGCATGGCA
G8s     I-15        TTCTTGATCGTTGGCATTGGCACACGGAGGTTTACTGAGTGGTAACCACGTTGCACGGCG
G8s     I-16        TTCTTGATCGTTGGCCTTGGCACTTGAAGGTCGAGTTAGCGGAATCCACCTTGCCTTGAT
G8s    II- 1        TTCTTGATCGTTGACATAGGCACACGGAGGTTCACCGAGTGGTAACCACGTTGCAGGGTA
G8s    II-11        TTCTTGATCGTTGGCATTGCCACACAGAAGTTTACTGAGTGGTAACCACGTTGCAGGGAC
G8s    II-12        TTCTTGATCGGTGACATTGGCACACGGAGGTTCACCGAGTGGTAACCACGTTGCATGGAA
```

Figure 25: Sequences isolated from generation 4 (labeled as G4s) and 8 (labeled as G8s) of speed selection. (A) Parental DEC22-18A sequence. (B) A total of 35 sequences were identified from G4 speed selection (G4s). (C) A total of 16 sequences were identified from G8 speed selection (G8s).

A

```
                    0         10        20        30        40        50
                    ...|....|....|....|....|....|....|....|....|....|....|.
G22.18A(WT)         TTCTTGATCGTTGTCATTGGCACACGGAGGTTTACTGAGTGGTAACCACGTTGCATGGAA
```

B

```
G11c   I- 2         TTCTTGATCGTTGGCATTGCCACACGGAGGTATACTGAGTGGTAACCACGTTGCACGGTA
G11c   I-14         TTCTTGATCGTTGGCATTGCCACGCGGAGGGGCGCTGAGCGGTAACCACGTTGCAGGGGC
G11c   I-16         TTCTTGATCGTTGGCAATGCCACACGGAGGTTGACTGAGTGGTAACCACGCTGCATGGGA
G11c  II- 8         TTCTTGATCGTTGGCATTGCCACACGGAGTGTTAATGAGTGGTAACCACGTTGCCGGGCA
G11c  II-10         TTCTTGATCGTTGGCACTGCCACACGGAGGTGTACCGAGTGGTAACCACGTTGCATGGAC
G11c  II- 2         TTCTTGATCGTTGGCATTGCCACTCGGAGGGG-ACTGAGAGGTAACCACGTTGCACGGTC
G11c  II- 3         TTCTTGATCGTTGGCTGTGCCATCCGGAGGTTAACTGAGGAGTAACCACGCTGCAGGGAG
```

C

```
G11c   I- 5         TTCTTGATCGTGGGGAGTGGAACACGGGGGTGTGGGGGGAGGTTACCATGTAGCATCGGA
G11c   I- 6         TTCTTGATCGTGGGGAGTGGAACACGGGGGTGTGGGGGGAGGTTACCATGTAGCATCGGA
G11c   I-10         TTCTTGATCGTGGGCAGTGGAACACGAGGGTGTGGCGGGTGGTAACAAAGTTGCATGTAA
G11c   I-12         TTCTTGATCGTGGGCAGTGGAACACGAGGGTGTGGCGGGTGGTAACAAAGTTGCATGTAA
G11c   I-13         TTCTTGATCGTGGGGAGTGGAACACGGGGGTGTGGGGGGAGGTTATTATGTAGCATCGGA
G11c   I-15         TTCTTGATCGTGGGGAGTGGAACACGGGGGTGTGGGGGGAGGTTATTATGTAGCATCGGA
G11c  II- 1         TTCTTGATCGTGGGGAGTGGAACACGGGGGTGTGGGGGGAGGTTACCATGTAGTATCGGA
G11c  II- 4         TTCTTGATCGTGGGGAGTGGAACACGGGGGTGTGGGGGGAGGTTATCATGTACCATCGGA
G11c  II- 5         TTCTTGATCGTGGGGAGTGGAACACGGGGGTGTGGGGGGAGGTTACCATGTAGCATCGGA
```

D

```
G11c   I- 8         TTCTTGATCGCGGGCAGTGGAACACGAGGGTGTGGCGGGTGGTGACAAAGTTGCATGTAA
G11c   I-11         TTCTTGATCGCGGGCAGTGGAACACGAGGGTGTGGCGGGTGGTAACAAAGTTGCCTGGTG
```

E

```
G11c   I- 9         TTCTTGATCGGTGGCCTGGGAACGGGGAGGTGGAATGAGCTGTGATCACGAGGAGGGGAT
```

Figure 26: Sequences isolated from generation 11 (labeled as G11c) of lower Co^{2+} selection. (A) Parental DEC22-18A sequence. A total of 19 sequences were identified from G11 lower Co^{2+} selection (G11c). Those sequences were subsequently grouped into 4 classes based on their sequence homology: (B) seven sequences in class I, (C) nine sequences in class II, (D) two sequences in class III, and (E) one sequence in class IV.

3.3.3 Sequencing alignment and proposal of a new secondary structure

To study the structure-function relationship of DEC22-18A, the G4 1-minute population (Figure 25B) from the speed selection was chosen for cloning and subsequent sequence analysis. The G4 1-minute population, with a longer, less stringent reaction time, was expected to have a more diverse sequence pool and therefore a wider range of variants. Comparative analysis of the 35 clones with the parental DEC22-18A identified conserved residues that were distributed more in the 3'-half of the original enzyme (Figure 27). The sequence analysis, along with evidence of the covarying residues (Figure 28), suggested a new secondary structure model (structure B in Figure 27B).

Model structure A, which was predicted from the *mfold* program, had been proposed before the re-selection was conducted (STA, Figure 27B). It contained three base-pairing regions (labeled as P1, P2, and P3), five single-stranded stretches (labeled SS1, SS2, SS3, SS4, and SS5), and two loop regions (labeled L1 and L2). Mutational analysis based on the STA structure turned out to be inconclusive and left us with more questions than answers (section 3.1.12). With the availability of the isolated sequence variants from the speed selection, we were able to compare those sequences with the parental DEC22-18A (Figure 27A). The analysis showed that many residues from the isolates were highly conserved, despite the fact that nucleotide positions 1 to 51 had been mutagenized at a level of 24% per residue. A survey of the complete length of the mutagenized region showed that only 12 out of the 51 positions could tolerate mutations in at least 24% of the 35 isolates (assigned grey color index in Figure 27A). There were 15 residues that were strictly conserved in all 35 isolates (assigned purple color index in

Figure 27A). Most of those conserved residues must participate in some way in the catalysis of DEC22-18A.

Analysis of the conserved and covarying residues led to model structure B presented in Figure 27B. It contains four base-pairing regions (labeled as P1, P2, P3, and P4), five single-stranded stretches (labeled SS1, SS2, SS3, SS4, and SS5), and two loop regions (labeled L1and L2). In this structure, there are three pairs of covarying residues (C14:G32, A15:T31, C16:G30) which give rise to the P4 region. Five cases of covariation were observed at this region (Figure 28A). Covariation means that mutation of one residue leads to a mutation at its corresponding pairing residue to maintain Watson-Crick base-pairing. The residue G30, T31, and G32 were also proposed to be part of a base-pairing region (P3) in the original model structure A (Figure 28B). However, mutation in the proposed P3 region of structure A did not confer a corresponding mutation in its proposed pairing partner. For example, mutation at residue T31 resulted in only 1 out of 4 covariations in its pairing sequence (A39). This evidence, along with the inconclusive results gathered from a previous structural study on model structure A (section 3.1.12), led us to pursue model structure B (Figure 27B). We sought to generate a set of mutants suitable for testing the significance of those conserved residues and the results of those mutants are discussed in the following section.

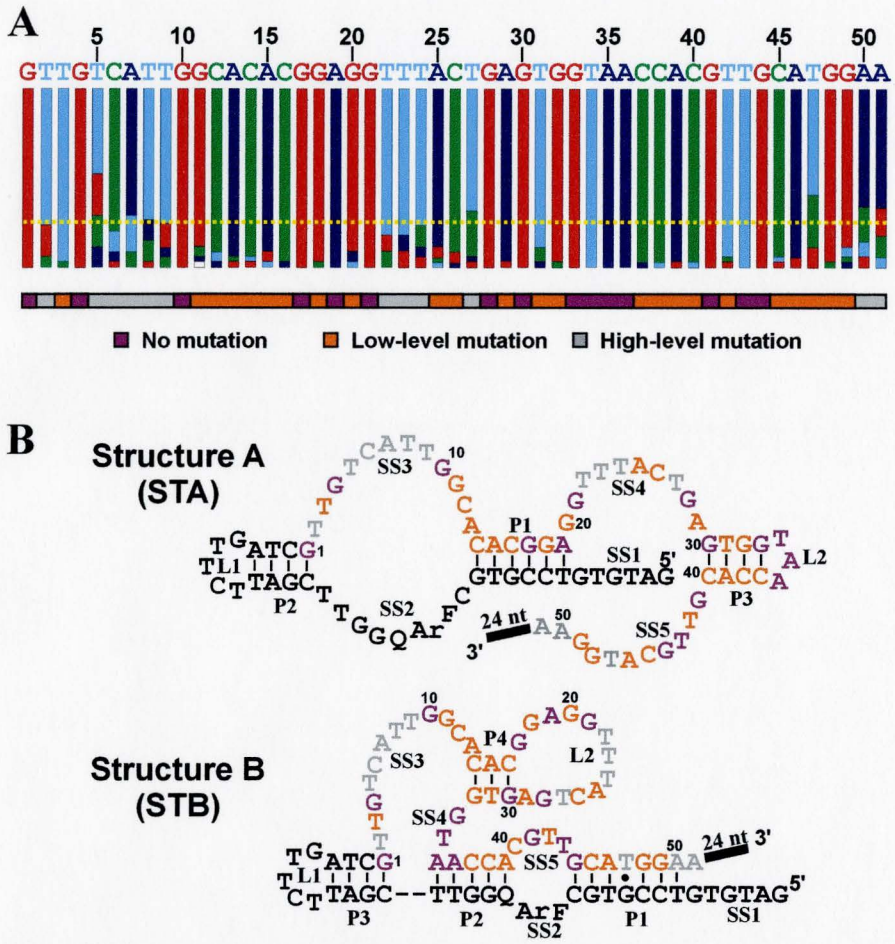


Figure 27: Proposal of a new secondary structure model. (A) All 35 sequences isolated from the speed selection were aligned and illustrated in different colors to show their sequence homology. Each nucleotide is represented by one unique colour (G: red; T: light blue; C: green; A: dark blue). The length of each color bar on the chart represents the frequency of occurrence of that nucleotide. A threshold of 24% was set to define the frequency of mutation at that position (represented by the yellow dash line). The projected maximum number of mutations for each nucleotide position is 8 in a total of 35 sequences (total # of isolates × % randomization = 35 × 24% = 8.4). For a given position, if at least 8 out of the 35 isolates showed mutation, it was identified to have a high level of mutation and assigned a grey color index. If a given position has less than 8 isolates showing mutation, it was identified to have a low level of mutation and assigned an orange color index. An absolutely conserved nucleotide was assigned a purple color index. The index is illustrated at the bottom of the color coded sequences. (B) Model structure A (STA) is the originally proposed secondary structure for DEC22-18A. Model structure B (STB) is the revised version of DEC22-18A secondary structure. The short stretch that is labeled "24 nt" at the 3'-end of STB represents the 3' primer sequence used during the re-selection.

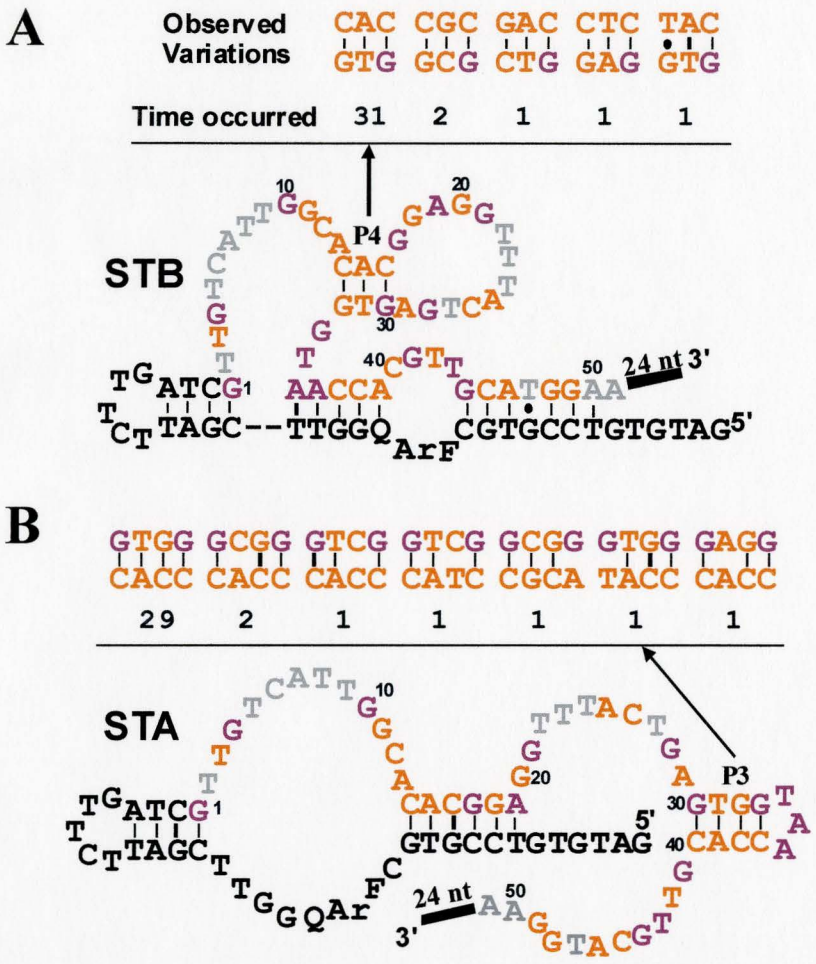


Figure 28: Identification of the conserved and covarying residues from sequence analysis. A detailed comparison of all speed selection sequence isolates revealed the conserved and potentially covarying residues. (A) Five combinations of covarying base-pairs had been identified at one site that was proposed as the P4 region on model structure B. Except for one instance, all isolates showed perfect sequence covariation at this region. (B) The previously proposed secondary structure model, model structure A (STA), was used for the same analysis to validate the base-pairing at P3 region. At least 7 out of 36 isolates exhibited imperfect base-pairing.

3.3.4 Structural study on DEC22-18A (based on the re-selection data)

The analysis on the sequence alignment from the speed selection revealed that 15 nucleotides in DEC22-18A were strictly invariant (nucleotides in purple, Figure 27B), whereas various degrees of nucleotide substitutions were tolerated at all other positions (nucleotides in orange and grey, Figure 27B). It was not clear how mutations at either the tolerant or intolerant positions would influence DEC22-18A catalysis. In the end, we would like to identify which specific residue or even possibly a structural feature that can provide the highest catalytic efficiency. We decided to conduct a comprehensive mutational analysis of the DEC22-18A in a step-wise manner to address these questions.

In the original substrate-enzyme construct, the 3'-end of the substrate was ligated to the 5'-end of the enzyme to build the *cis*-acting DEC22-18A. However, this configuration limited our ability to probe the sequence relevance from the 5'-end of the enzyme region (sequence in blue in Figure 13, excluding the substrate sequence in yellow). We decided to switch the physical linkage between the substrate and enzyme regions from the 3'-end substrate with the 5'-end enzyme (substrate-enzyme configuration), to the 3'-end enzyme with the 5'-end substrate (enzyme-substrate configuration, STB3, Figure 29). Construct STB3 showed activity only slightly less than the wild-type DEC22-18A. The high activity of the alternative configuration suggests that the 3'-end of substrate does not have to be physically attached to the 5'-end of the enzyme for catalysis to happen, such as in the case of DET22-18 (however, DET22-18 does have extended base-paring to compensate for the lack of physical attachment). Nonetheless the attachment between the 3'-end of the substrate and the 5'-end of the

enzyme does have the advantage to facilitate the catalysis. Ultimately, the establishment of this alternative configuration did provide us the freedom to perform truncation from the 5'-end of the enzyme region, and possibility to truncate the enzyme further.

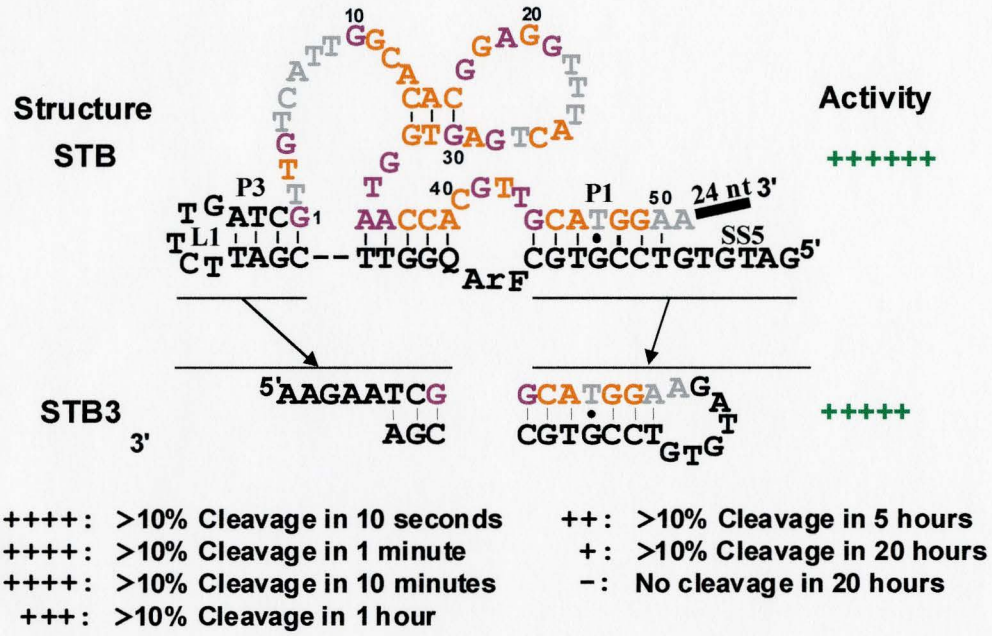


Figure 29: DEC22-18A topology experiment. STB stands for STructural model B. The wild-type, *cis*- acting version of the enzyme was constructed by linking the 3'-end of the substrate with the 5'-end of the enzyme (substrate-enzyme configuration) as shown above in STB. To change the topology of the enzyme, the 3'-end of the enzyme was linked to the 5'-end of the substrate (enzyme-substrate configuration), as shown in STB3. The activity of STB3 was then assessed and compared with wild-type enzyme using the index listed.

The P2 region in model structure A is the same as the P3 region in model structural B in Figure 27B. This stem-loop was one of the pre-designed library features and was proposed to connect and to bring together the substrate region to the enzymatic region (Figure 6B, see discussion in section 3.1.2). The necessity of this stem-loop could not be confirmed in previous structural study (discussed in section 3.1.12). In addition, the result gathered from construct STB3 did suggest the possibility to eliminate those sequences. To test the dispensability of this stem-loop, we decided to replace it with a short linker sequence (Figure 30). It had been established that the substrate region and enzyme region can be ligated in either the substrate-enzyme (STB in Figure 29) or the enzyme-substrate configuration (STB3 in Figure 29). The mutant for this study was therefore constructed and tested in both configurations (STB4 = substrate-enzyme configuration; STB5 = enzyme-substrate configuration). Construct STB4 showed activity close to wild-type, while construct STB5 showed slightly less. The high activity of both constructs not only suggests that the first nine residues of the original enzyme (residue numbered -8 to 0, Figure 23) are not absolutely required for catalysis, but also verifies the results from the study performed on constructs S1PE and S2PE in section 3.1.12. The decreases in activities observed in both mutants also implies that modification on the original stem-loop sequences might cause slight distortion from the enzyme's most active conformation and consequently led to a slight drop in activities.

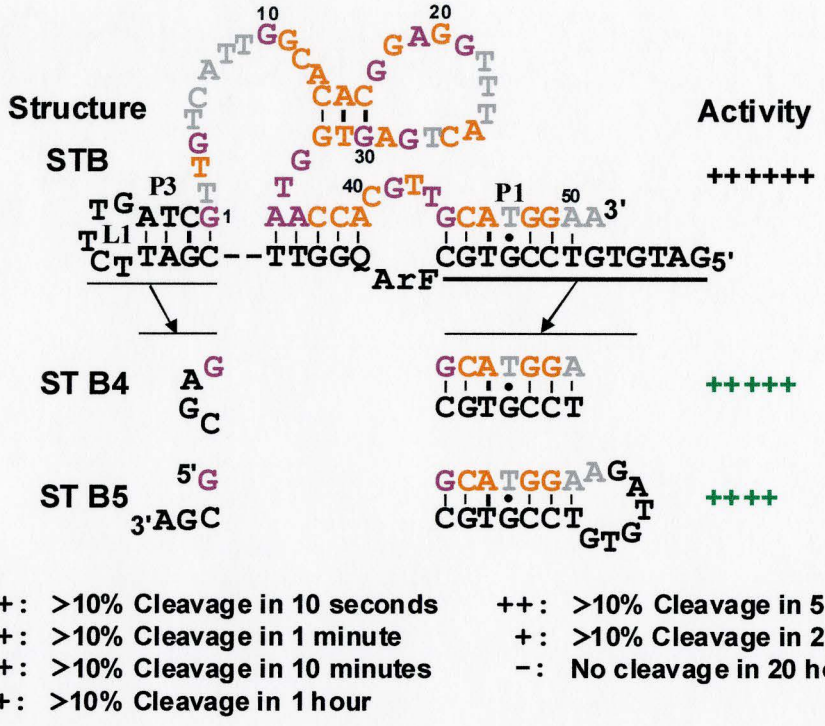


Figure 30: The dispensability of stem-loop L1-P3 in structure B. To test if these sequences were essential for catalysis, the stem (P3) and loop (L1) were replaced with a short linker sequence shown above. The same construct was tested in both the substrate-enzyme configuration (STB4) and enzyme-substrate configuration (STB5). Their activities were then assessed and compared with wild-type enzyme using the index listed.

From the sequence alignment, we noticed a non-continuous stretch of highly mutated residues in the SS3 region at the 5'-end of the enzyme (grey sequence in structure B, Figure 27B). Except for residues G1, T3, G4, and G10, the first ten nucleotides showed a high level of mutation with a particular preference for nucleotide T. Additionally, nucleotides from position -8 to 0 had already been proven to be dispensable. Therefore, we decided to make a series of constructs, which had various nucleotides from position 1 to 10 replaced by nucleotide T, to test the catalytic relevance of sequence on the SS3 region. When we tested construct STB5.1, in which residues C6 and A7 were mutated to T, its activity remained the same as the wild type (Figure 31). The full activity observed from mutant STB5.1 did not surprise us, considering several of the 35 isolates had nucleotide T substitutions occurring at residue C6 and A7. This observation led us to question the role of those T residues might play in DEC22-18A catalysis. We hypothesized that this T stretch could act either as a bridge to position residues that participate in catalysis, or as participants forming interactions with neighboring nucleotides. A construct with nucleotide substitutions from T to A at position 2, 3, 5, 6, 7, 8, and 9 was made to assess the bridge hypothesis (STB5.2 in Figure 31). If these T residues were merely used in positioning the catalytically relevant residues, presumably the highly conserved G1, G4, and G10 residues, to proper geometry for catalysis, the activity of the enzyme would still remain. On the other hand, if those T residues were indeed involved in interaction with neighboring nucleotides, any possible wobble pairing between G and T would be disrupted with the nucleotide A substitutions. The result showed a drastic reduction in STB5.2 activity and thus invalidated the bridge

hypothesis. Furthermore, the hypothesis of those T residues interacting with the neighboring residues was reinforced when the activity was rescued with the re-substitution of two T residues (T2 and T3) back in construct STB5.3. We intentionally synthesized an additional construct that had an eight base-paired duplex introduced in the original SS3 region (STB8 in Figure 31). Interestingly, the activity of STB8 was successfully maintained, even when the original single-stranded region was replaced with a duplex. This provides us an opportunity to engineer future allosteric deoxyribozymes through this duplex, possibly adopting the same approach discussed in section 3.1.13.

When we started to examine the importance of the conserved residues G1 and G4, their roles were utterly confirmed with the loss of activities in construct STB5.4, STB5.5, STB5.6, STB5.7, and STB5.8 (Figure 31). Even a single deletion made on residue G1 (STB5.6) or a single mutation made on G4 (STB5.5) enabled a complete abolishment in activity, which reemphasized the catalytic significance of both highly conserved G residues. It is known that nucleotide G can form wobble-base-pair with nucleotide T that, along with other non-Watson-Crick base-paring, are crucial in forming higher ordered structure in nucleic acids. The residues from position 1 to 10 are thus thought to be involved in non-Watson-Crick interactions with the neighboring nucleotides. Nevertheless, it is difficult to confirm this hypothesis simply using the mutational analysis, while the NMR or X-ray crystal structure is not available.

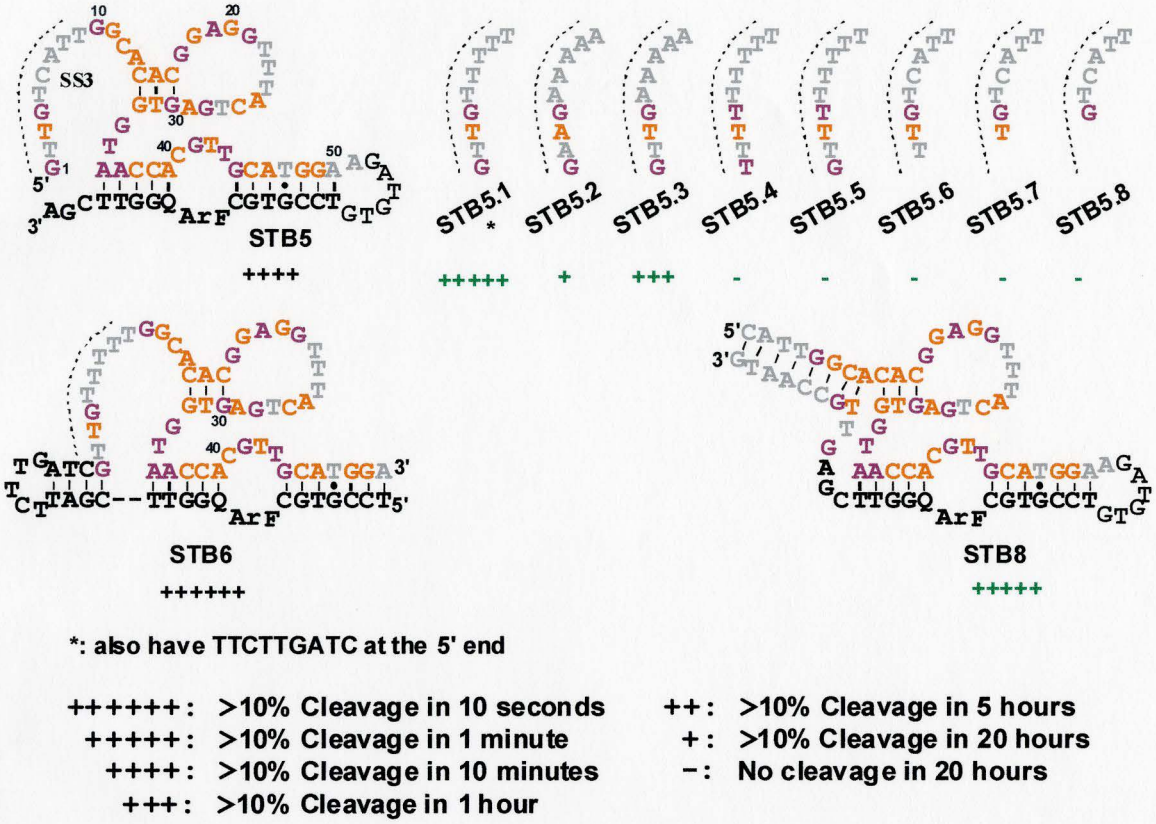


Figure 31: Examination of sequence in the SS3 region of structure model B. Constructs with various mutations and deletions in the SS3 region were synthesized and then tested for their self-cleaving activities. Constructs STB5.1, 5.2, 5.3, 5.4 and 5.5 consisted of various mutations on the SS3 region, while constructs STB5, 5.6, 5.7, and 5.8 consisted of nucleotide deletions from the 5'-end of the enzyme. Construct 5.1 and 6 had the same sequence, except that STB5.1 was ligated in the enzyme-substrate configuration and STB6 was ligated in the substrate-enzyme configuration. Construct STB8 had its SS3 region replaced with an eight base-paired duplex. Their activities were assessed and compared with wild-type enzyme using the index listed.

Next we investigated the region L2 on model structure B (STB in Figure 27B). This loop (L2) consists of nucleotides from position 17 to 29, with 4 absolutely conserved residues, 5 moderately conserved residues, and 4 highly mutated T residues. When these conserved residues were mutated to T (STB9: mutated G20, G21, A25, C26) or deleted (STB13: deletion from G20 to T27), no activity from either construct was observed (Figure 32). Even a breakage in the L2 region between two of the conserved residues (between A19 and G20) resulted in a total loss of activity in construct STB16 (Figure 32).

When probing the relevance of those highly mutated T residues at position 22, 23, and 24 (Figure 32), intriguing results were obtained from constructs STB10 (one nucleotide T deletion), STB11 (two nucleotides T deletion), and STB12 (three nucleotides T deletion). The deletion of three T residues completely abolished the enzyme's ability to self-cleave (STB12 in Figure 32). The deletion of one T residues substantially lowered the activity to barely above background (~6% in 1 hour) and the construct failed to cleave more than 7% even after 20 hours (STB10 in Figure 32). In contrast, the deletion of two T residues sustained the enzyme activity (STB11 in Figure 32). Construct STB17, which was made with a breakage between T22 and T23, also displayed a similar activity level as construct STB11. Those results imply again the importance of positioning critical residues, presumably those that were highly and moderately conserved (purple and orange sequences in Figure 32), in the right geometry for catalysis. This hypothesis was further supported by the fact that substitutions of nucleotide A at position 22, 23, and 24 resumed the activity as observed in construct

STB14. The identity of those three residues was probably not as important as for their role to provide proper geometry for catalysis. Nevertheless, it is difficult to define the precise tertiary interaction between the L2 region and the active site with this study.

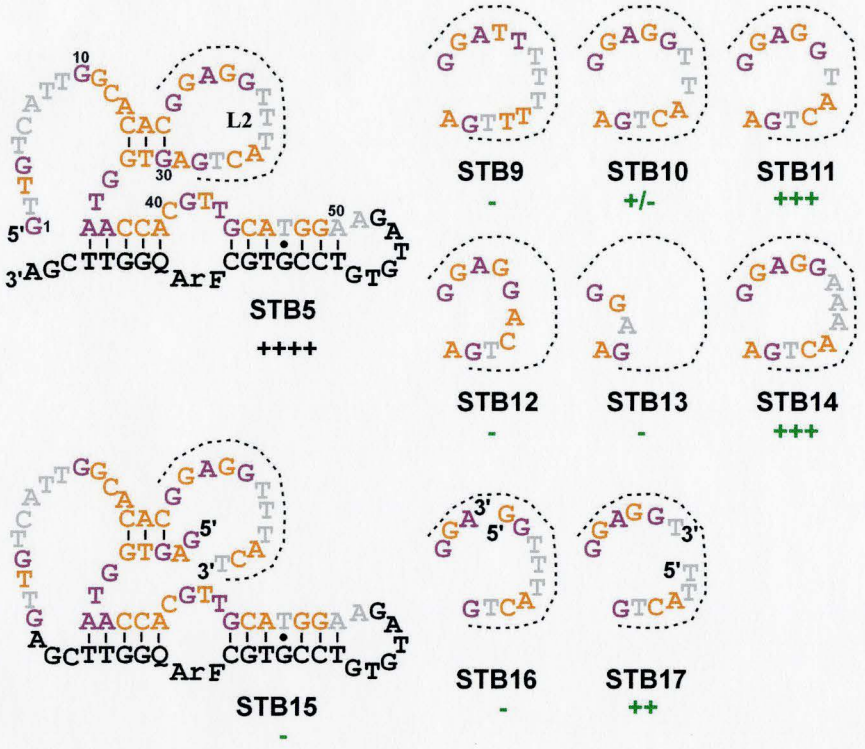


Figure 32: Examination of sequence in the L2 region of model structure B. Constructs with various mutations and deletions on the L2 region were synthesized and then tested for their self-cleaving activities. Constructs STB9 and 14 contained mutations in the L2 region, while constructs STB10, 11, 12, 13 contained nucleotide deletions within the L2 region. Construct STB16 contained a break point between A19 and G20, while construct 17 contained a break point between T22 and T23. Their activities were assessed and compared with wild-type enzyme using the index listed. As for construct STB10, even though ~6 % cleavage was achieved after 1 hour of reaction, it failed to cleave more than 7% after 20 hours. Therefore, construct STB10 was assigned a "+/-" sign as defined by the index.

Last but not the least, we provided evidence for the base-pairing regions, the P1 and P2 in model structure B, to consolidate our revised model (Figure 33). In model structure B, the P1 and P2 regions had been proposed to form the substrate binding arms on each side of the RNA cleavage site. The P1 region, which is near the 3'-end of the enzyme region (nucleotides 44 to 50, Figure 33), consists of the pre-designed sequence that was meant to base-pair with the invariable substrate strand (sequence illustrated in green, Figure 6). Even though those nucleotides were not in the primer binding region, they remained highly conserved throughout the original selection and the re-selection (those nucleotides were partially mutagenized before the re-selection). One of the most frequently occurring mutations in the P1 region was at nucleotide position 47 in Figure 33 (the residue was labeled as position 61 instead in Figure 8). The nucleotide at position 45 (Figure 33) showed a high mutation rate among clones that had been selected from the original selection (residue was labeled as position 59 instead in Figure 8). There seem to be an approximately equal likelihood of those two residues being a T or a C among all the sequences we isolated from the first selection. Both thymine and cytosine are known to base pair with guanine, which respectively is the base-pairing partner for both positions at the substrate region. As for region P2, nucleotides at position 35 to 39 in Figure 33 (these residues were labeled as position 47 to 51 instead in Figure 8) were also well conserved throughout both selections.

We constructed four mutants in the same manner as the ones that were used in the preliminary structural study (section 3.1.12), in which intentional base-pairing partners were swapped with either different base-pair or mismatch. Construct STB18 had two

mutations made in the substrate to introduce mismatch, and construct STB19 had two Watson-Crick covarying mutations (Figure 33). The mismatches completely abolished the activity of STB18, while the swapped base-pairs still enabled STB19 to exhibit a moderate level of activity compared to the wild-type. The reduction in STB19 activity might be due to the mutation made to one of the strictly conserved residue (A36 to C), which was possibly involved with other interactions than simply forming base-pair. Nevertheless, the fact that STB19 retained the activity provides the evidence for the proposed base-pairing at P2 region. The same strategy was also employed to probe the integrity of P1 region. Construct STB20 had two mutations made in the substrate to introduce mismatch, and construct STB21 had three Watson-Crick covarying mutations (Figure 33). STB20 exhibited a complete loss of activity, while STB21 displayed only a slight decrease in activity compared to the wild-type. These results demonstrate that nucleotides in the P1 region indeed base-pair with the substrate. The requirement for those seven base-pairs in the P1 region further confirmed the results from the DEC22-18 truncation study (refer to Figure 13, section 3.1.7), in which deletions made from nucleotide 82 to 79 (correspond to nucleotide 50 to 47 of STB in Figure 27B) progressively decreased the activity to zero. The abolishment in activity from four nucleotide deletion was probably due to the decrease in melting point of this stem (P1) from 22°C (7 base pairs) to 8°C (3 base pairs), and consequently led to the denaturing of P1 region.

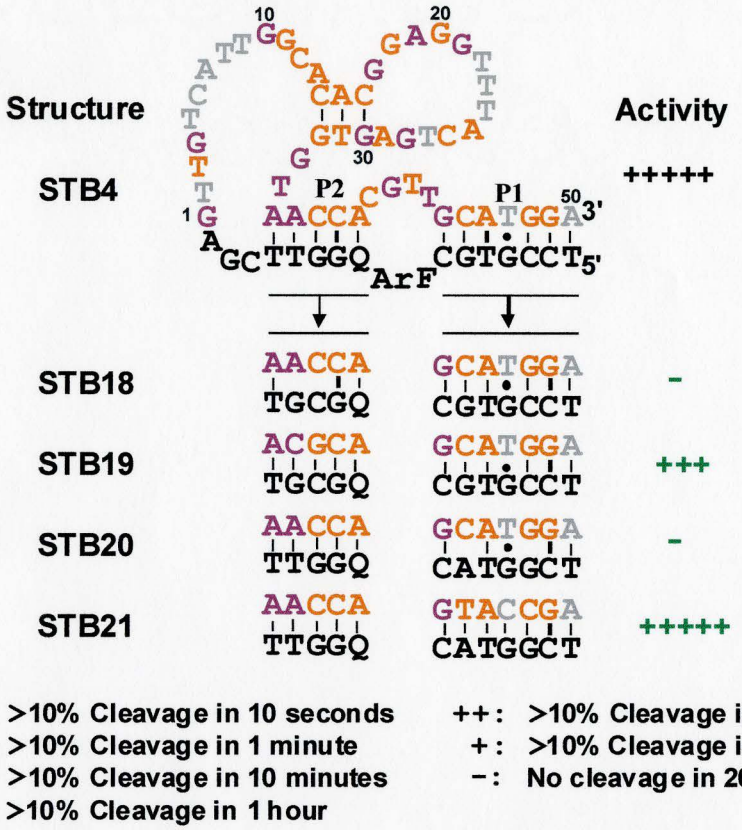


Figure 33: Examination of the proposed base-paring regions between enzyme and substrate. Constructs STB18 and 19 were made to investigate the proposed base-pairing at region P2 on model structure B. Construct STB18 had two mutations made at the substrate to introduce mismatch, and construct STB19 had two Watson-Crick covarying mutations. Construct 20 and 21 were made to investigate the proposed base-pairing at region P1 on model structure B. Construct STB20 had two mutations made at the substrate to introduce mismatch, and construct STB21 had three Watson-Crick covarying mutations. Their activities were assessed and compared with wild-type enzyme using the index listed.

In this study, we provide additional evidence for the secondary structural model of DEC22-18A. Sequence analysis allowed us to propose a revised secondary structure, which was named model structure B (STB in Figure 27B). Based on this model, we conducted mutational analysis investigating the significance of certain nucleotides, which were either highly conserved or highly mutated. Mutations made to the conserved-sequence domain resulted in dramatic or even complete loss of catalytic function. As for nucleotides that showed high tolerance for mutations, majority of them seem to play roles in maintaining proper geometry for catalysis rather than actually participating in reaction. Mutational study enabled us to gain an insight into the structural basis of DEC22-18A catalysis and defined certain features that might be essential. More efforts will be spent to further characterize this intriguing enzyme, since a thorough understanding of DEC22-18A will definitely benefit our future effort for utilization.

CHAPTER FOUR

CONCLUDING REMARKS

We are interested in developing a fluorescence-generating biosensor system utilizing DNA enzymes. The criteria for our system include low background fluorescence when inactivated with minimal false signalling, and a fast response upon activation. To achieve these goals, we designed our substrate to have a fluorophore and quencher physically linked to one DNA chain, sandwiching the proposed site of RNA cleavage. We then sought a DNA enzyme that could efficiently cleave this specially arranged RNA junction. Finally, we wanted to prove our conceptual biosensor system using the enzyme we specifically isolated, along with an existing DNA aptamer. By adjoining the enzyme domain with an aptamer domain, we aimed to rationally engineer our DNA enzyme into one that could be allosterically regulated, as many others have demonstrated with ribozymes (Breaker, 2002).

We successfully isolated a DNA enzyme that is capable of coupling RNA cleavage with fluorescence signalling (Mei et al., 2003). The enzyme (named DEC22-18) isolated directly from selection performs catalysis in a *cis*-acting manner, which means it can only process a single substrate that has been covalently linked to it. The selection buffer used to isolate DEC22-18 contained 400 mM NaCl, 100 mM KCl, 7.5 mM $MgCl_2$, 5 mM $MnCl_2$, 1 mM $CoCl_2$, 0.25 mM $NiCl_2$, and 1.25 mM $CdCl_2$ in 50 mM HEPES (pH 6.8 at 23°C). Monovalent ion titration, divalent ion titration, and preliminary pH profiling studies permitted us to further optimize the catalytic

performance of DEC22-18. DEC22-18 was found to be most effective with 10 mM of Co^{2+} (along with 5 mM $MgCl_2$) in 50 mM HEPES at pH 6.8 without monovalent ion. The full-length enzyme (109 nt) was subsequently truncated from its 3'-end yielding an even more efficient, 83-nt (enzyme region alone consists of only 60 nucleotides) enzyme (named DEC22-18A). Under optimal conditions (10 mM $CoCl_2$, and 5 mM $MgCl_2$ in 50 mM HEPES at pH 6.8 and 23°C), the kinetic parameters of DEC22-18A were too fast to be accurately measured by manual quenching (nearly 50% of the DEC22-18A self-cleaved in 3 seconds).

Preliminary structural data obtained from our experimental approach and computer folding program enabled us to separate the enzyme region from the substrate region. A true enzyme that is able to perform cleavage in an intermolecular fashion was therefore engineered. The *trans*-acting version of the enzyme was named DET22-18 and its k_{cat} was determined to be ~7.2 ± 0.7 min^{-1} and a K_M at 0.94 ± 0.19 mM. DET22-18 is the second fastest DNA enzyme reported to date, its speed closely matches the k_{cat} (~10 min^{-1}) reported for the "10-23" enzyme (Santoro and Joyce, 1998).

Both *cis*- and *trans*-versions of the enzyme can generate fluorescence as we anticipated. Most promisingly, we were able to adjoin DEC22-18A with a known DNA aptamer that binds specifically to ATP (Huizenga & Szostak, 1995) by the strategy of rational design. This proof-of-concept construct is able to distinguish ATP and GTP, and to perform real-time fluorescence signalling upon addition of the ligand. After optimization, our ATP reporter system showed a rate activation of nearly 20-fold by ATP

versus GTP (0.25 mM $CoCl_2$, and 14 mM $MgCl_2$ in 50mM HEPES, pH 6.8) at a temperature of 37°C.

To assist in our future efforts to productively utilize DEC22-18A, we synthesized a partially randomized library, based on the original sequence of DEC22-18A, and carried out a re-selection. The re-selection enabled us to isolate sequences that closely resembled DEC22-18A but contained random point mutations. By aligning and comparing all sequences from the re-selection, we were able to identify nucleotides that seemed to be absolutely conserved and possibly relevant for catalysis. Following the sequence alignment, we went on to conduct a comprehensive mutational analysis to derive a more meaningful secondary structure model, rather than simply obtaining one using a computer folding program. While most of our data supported model structure B (STB in Figure 27B), further investigation is still underway, including chemical probing experiments as well as an effort to define the substrate specificity for DEC22-18A.

The revised model STB (Figure 34B) shows some similarities with the secondary structure of the hairpin ribozyme. Hairpin ribozyme is a naturally occurring RNA-cleaving RNA that was first discovered in tobacco ringspot virus satellite RNA. Enormous efforts have been made towards understanding the structure and function of the hairpin ribozyme (reviewed in Fedor, 2000), especially with the recent report of a high-resolution crystal structure (Rupert and Ferre-D'Amare, 2001). Both hairpin ribozyme and DEC22-18A seemed to share analogous features, based mainly on the essential helix-loop-helix elements in both proposed secondary structures. For example, stems P1/P2 and P4 in model STB appear to resemble stem A and B in the hairpin

ribozyme secondary structure, respectively. The functional structure of the hairpin ribozyme has been proposed to consist of a docking event of loop A and B domains into a non-coaxial orientation that brings loop nucleotides into proximity (Fedor, 2002). As for now, we have no evidence indicating that DEC22-18A adopts the same docking pathway for its catalysis. Nevertheless, future investigation will be conducted for the determination of the structural dynamics and function of DEC22-18A.

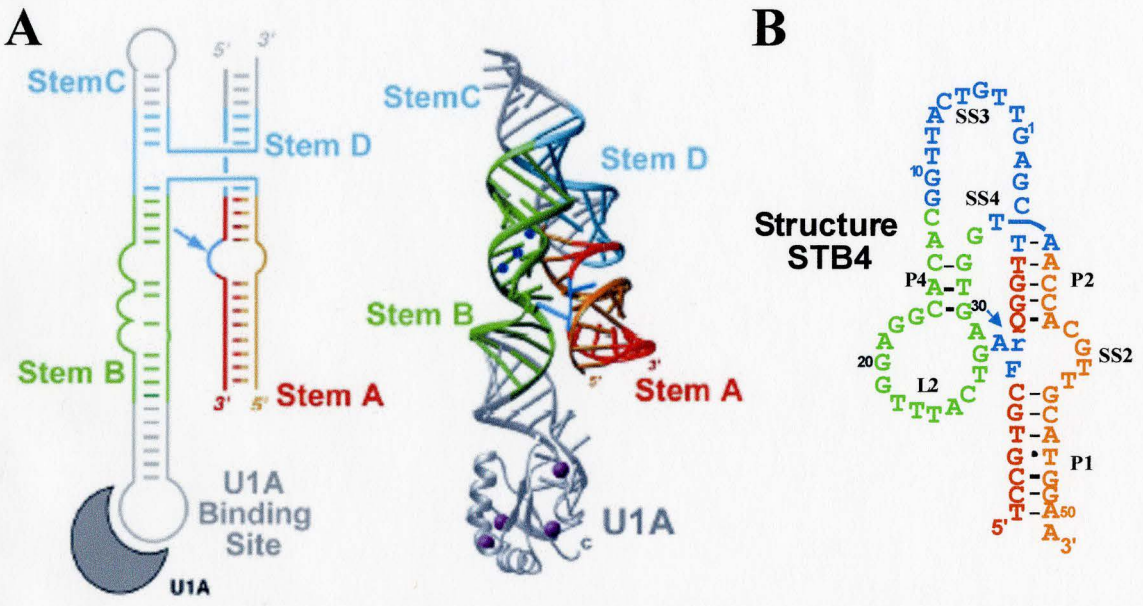


Figure 34: Comparison between the hairpin ribozyme overall structure and DEC22-18A proposed structure. (A) The secondary structure of the hairpin ribozyme is on the left. In the red substrate strand, the two nucleotides adjacent to the cleavage site are light blue. The arrow indicates the cleavage site. Three-dimensional structure of the hairpin ribozyme is on the right. The RNA backbone is represented schematically as ribbons, the nucleotide bases as sticks. Two bound calcium ions are represented as blue spheres (figure obtained from www.als.lbl.gov/als/compendium/AbstractManager/uploads/00009.pdf; Rupert and Ferré-D'Amaré, 2001). (B) Proposed secondary structure of DEC22-18A, also referred to as structure B (STB4).

One of the key features in our biosensor system was the use of a rationally designed substrate. The fluorophore and the quencher used in the substrate, which are fluorescein and dabcyl respectively, can be easily obtained through chemical synthesis for a reasonable cost. Fluorescein was chosen due to its relatively high absorptivity, excellent fluorescence quantum yield, and good water solubility. However, there are a few shortcomings with this substrate that are worth mentioning. Fluorescein is inherently pH and light sensitive; both properties make the quantitative measurement of this fluorophore problematic. A decrease in pH from 8.0 to 6.5 can result in a 40% loss of the fluorescence signal. Special care has to be taken to avoid photo bleaching of fluorescein, which leads to the irreversible destruction of the excited fluorophore.

The fluorophore and quencher labeled substrate was synthesized offsite; therefore, it is challenging to ensure both labels were actually attached to the substrate strand. If some of the substrate oligonucleotides were missing the fluorophore modification, a much lower fluorescence reading will be observed. If some of the substrate oligonucleotides were missing the quencher modification, a much higher background reading can be expected. In addition, we speculated that those two closely situated fluorophore and quencher groups, adjacent to the RNA cleavage site, might actually participate in the catalysis of DEC22-18. Missing one or both modified groups on the substrate can potentially jeopardize the enzyme's ability to cleave those substrates. The necessity of the fluorophore and quencher groups for catalysis will be revealed upon completion of the substrate specificity experiment. Purification of the synthesized substrates can eliminate those substrates that are missing the modifications. As for our

future sensor system, a far more stable fluorophore label that can be used over a wide pH range will be a more suitable candidate.

In addition to DEC22-18, which had been discussed extensively in this thesis, our group has also successfully created a set of DNA enzymes with pH optima covering 5 pH units (pH 3 to 8) utilizing the same biosensor concept (Liu et al., 2003). These DNA enzymes, with their unique pH-reporting properties, have great potential to be developed into fluorogenic probes that can follow pH changes in a chemical or enzymatic reaction mixture. There is an ongoing effort to develop more signalling DNA enzymes that can be engineered into "reagentless" sensors. By creating a collection of DNA enzymes carrying different fluorophores and quenchers having different recognition and reporting features, we will be able to set up various forms of multiplexed assays for the use of high-throughput or massive detection of biological samples (Breaker, 2002).

REFERENCES

Barezovski, M., Nutiu, R., Li, Y. & Krylov, S.N. 2003. Affinity analysis of a protein-aptamer complex using nonequilibrium capillary electrophoresis of equilibrium mixtures. *Anal. Chem.* **75,** 1382-1386.

Bittker, J.A., Le, B.V. & Liu, D.R. 2002. Nucleic acid evolution and minimization by nonhomologous random recombination. *Nature Biotechnology.* **20,** 1024-1029.

Breaker R.R. & Joyce, G.F. 1994. A DNA enzyme that cleaves RNA. *Chem. Biol.* **1,** 223-229.

Breaker, R.R. 2000. Making catalytic DNAs. *Science.* **290,** 2095-2096.

Breaker, R.R. 2002. Engineering allosteric ribozymes as biosensor components. *Current Opinion in Biotechnology.* **13,** 31-39.

Breaker, R.R., Emilsson, G.M., Lazarev, D., Nakamura, S., Puskarz, I.J., Roth, A. & Sudarsan, N. 2003. A common speed limit for RNA-cleaving ribozymes and deoxyribozymes. *RNA.* **9,** 949-957.

Blount, K.F. & Uhlenbeck, O.C. 2002. The hammerhead ribozyme. *Biochemical Society Transactions.* **30,** 1119-1122.

Brennan, J.D. 2000. Fluorescence-based biosensors. In *Encyclopedia of Analytical Chemistry,* pp. 383-403, John Wiley and Sons, Ltd., Chichester.

Brennan, J.D. 1999. Preparation and entrapment of fluorescently-labeled proteins for the development of reagentless optical biosensors. *Journal of Fluorescence.* **9,** 295-312.

Carmi, N. & Breaker, R.R. 2001. Characterization of a DNA-cleaving deoxyribozyme. *Bioorganic & Medicinal Chemistry.* **9,** 2589-2600.

Carmi, N., Shultz, L. & Breaker, R.R. 1996. In vitro selection of self-cleaving DNAs. *Chem. Biol.* **3,** 1039-1046.

Cate J.H., Gooding A.R., Podell E., Zhou K., Golden, B.L. & Kundrot C.E. 1996. Crystal structure of a group I ribozyme domain: Principles of RNA packing. *Science.* **273,** 1678-1685.

Cech, T. R. 1993. Structure and mechanism of the large catalytic RNAs: group I and group II introns and ribonuclease P. In *The RNA World* (Gesteland, R.F. & Atkins, J.F., eds), pp. 239-269, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Cuenoud, B. & Szostak, J.W. 1995. A DNA metalloenzyme with ligase activity. *Nature.* **375,** 611-614.

Dahm, S.C., Derrick, W.B. & Uhlenbeck, O.C. 1993. Evidence for the role of solvated metal hydroxide in the hammerhead cleavage mechanism. *Biochemistry.* **32,** 13040-13045.

Ekland, E.H., Szostak, J.W. & Bartel, D.P. 1995. Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science.* **269,** 364-370.

Emilsson, G.M. & Breaker, R.R. 2002. Deoxyribozymes: new activities and new applications. *Cell. Mol. Life Sci.* **59,** 596-607.

Faulhammer, D. & Famulok, M. 1997. Characterization and divalent metal-ion dependence of *in vitro* selected deoxyribozymes which cleave DNA/RNA chimeric oligonucleotides. *J. Mol. Biol.* **269,** 188-202.

Fedor, M.J. 2000. Structure and function of the hairpin ribozyme. *J. Mol. Biol.* **297,** 269-291.

Fedor, M.J. 2002. The catalytic mechanism of the hairpin ribozyme. *Biochemical Society Transactions.* **30,** 1109-1115.

German, I., Buchanan, D.D. & Kennedy, R.T. 1998. Aptamers as ligands in affinity probe capillary electrophoresis. *Anal. Chem.* **70,** 4540-4545.

Geyer, C.R. & Sen, D. 1997. Evidence of the metal cofactor-independence of a RNA phosphodiester cleaving DNAzyme. *Chem. Biol.* **4,** 579-593.

Gold, L. 1995. Oligonucleotides as research, diagnostic, and therapeutic agents. *The Journal of Biological Chemistry.* **270,** 13581-13584.

Hermann, T. & Patel, D.J. 1999. Stitching together RNA tertiary architectures. *J. Mol. Biol.* **294,** 829-849.

Huizenga, D.E. & Szostak, J.W. 1995. A DNA aptamer that binds adenosine and ATP. *Biochemistry.* **34,** 656-665.

Jhaveri, S., Rajendran, M. & Ellington A.D. 2000. In vitro selection of signalling aptamers. *Nature Biotechnology.* **18,** 1293-1297.

Lakowicz, J.R. 1999. Fluorescence sensing. In *Principles of Fluorescence Spectroscopy*, pp. 546-547, Kluwer Academic/Plenum Publishers, New York, NY.

Levy, M. & Ellington A.D. 2002. ATP-dependent allosteric DNA enzymes. *Chemistry & Biology.* **9,** 417-426.

Li, J. & Liu Y. 2000. A highly sensitive and selective catalytic DNA biosensors for Lead Ions. *J. Am. Chem. Soc.* **122,** 10466-10467.

Li, J., Zheng, W., Kwon, A.H. & Lu, Y. 2000. *In vitro* selection and characterization of a highly efficient Zn(II)-dependent RNA-cleaving deoxyribozyme. *Nucleic Acids Research.* **28,** 481-488.

Li, Y. & Breaker, R.R. 1999a. Phosphorylation DNA with DNA. *Proc. Natl. Acad. Sci.* **96,** 2746-2751.

Li, Y. & Breaker, R.R. 1999b. Deoxyribozymes: new players in the ancient game of biocatalysis. *Current Opinion in Structural Biology.* **9,** 315-323.

Li, Y. & Breaker, R.R. 1999c. Kinetics of RNA degradation by specific base catalysis of transesterication involving the 2'-hydroyl group. *JACS.* **121,** 5364-5372.

Li, Y. & Breaker, R.R. 2000. Capping DNA with DNA. *Biochemistry.* **39,** 3106-3114.

Li, Y. & Sen, D. 1996. A catalytic DNA for porphyrin metallation. *Nat. Struct. Biol.* **3,** 743-747.

Liu, Z., Mei, S.H.J., Brennan, J.D. & Li, Y. 2003. Assemblage of signalling DNA enzymes with intriguing metal-ion specificities and pH dependences. *JACS.* **125,** 7539-7545.

Lu, Y. 2002. New transition-metal-dependent DNAzymes as efficient endonuclease and as selective metal biosensors. *Chem. Eur. J.* **8,** 4589-4596.

Mei, S.H.J., Liu, Z., Brennan, J.D. & Li, Y. 2003. An efficient RNA-cleaving DNA enzyme that synchronizes catalysis with fluorescence signaling. *JACS.* **125,** 412-420.

Murray, J.B., Seyhand, A.A., Walter, N.G., Burke, J.M. & Scott, W.G. 1998. The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone. *Chem. Biol.* **5,** 587-595.

Nutiu, R. & Li, Y. 2003. Structure-switching signalling aptamers. *JACS.* **125,** 4771-4778.

Okumoto, Y., Ohmichi, T. & Sugimoto, N. 2002. Immobilized small deoxyribozyme to distinguish RNA secondary structures. *Biochemistry.* **41,** 2769-2773.

Paracchi, A. 2000. Preferential activation of the 8-17 deoxyribozyme by Ca^{2+} ions. *JBC.* **275,** 11693-11697.

Piunno, P.A., Krull, U.J., Hudson, R.H., Damha, M.J. & Cohen, H. 1995. Fiber-optic DNA sensor for fluorometric nucleic acid determination. *Anal. Chem.* **67,** 2635-2643.

Potyrailo, R.A., Conrad, R.C., Ellington A.D. & Hieftje, G.M. 1998. Adapting selected nucleic acid ligands (aptamers) to biosensors. *Anal. Chem.* **70,** 3419-3425.

Riesner, D. Nucleic acid structures. In: Lichtenstein, C & Nellen, W., editors. Antisense Technology. New York: Oxford University Press; 1997. p. 1-23.

Robertson, M.P. & Ellington, A.D. 2000. Design and optimization of effector-activated ribozyme ligases. *Nucleic Acids Res.* **28,** 1751-1759.

Rupert, P.B. & Ferré-D'Amaré, A. 2001. Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature.* **410,** 780-786.

Santoro, S.W. & Joyce, G.F. 1997. A general purpose RNA-cleaving DNA enzyme. *Proc. Natl. Acad. Sci.* **94,** 4262-4266.

Santoro, S.W. & Joyce, G.F. 1998. Mechanism and utility of an RNA-cleaving DNA enzyme. *Biochemistry.* **37,** 13330-13342.

Santoro, S.W., Joyce, G.F., Sakthivel K., Gramatikova, S. & Barbas, C.F. 2000. RNA cleavage by a DNA enzyme with extended chemical functionality. *JACS.* **122,** 2433-2439.

Sen, D. & Geyer, C.R. 1998. DNA enzymes. *Current Opinion in Chemical Biology.* **2,** 680-687.

Sheppard, T.L., Ordoukhanian, P. & Joyce, G.F. 2000. A DNA enzyme with *N*-glycosylase activity. *Proc. Natl. Acad. Sci.* **97,** 7802-7807.

Silverman, S.K. 2003. Rube Goldberg goes (ribo)nuclear? Molecular switches and sensors made from RNA. *RNA.* **9,** 377-383.

Soukup, G.A. & Breaker, R.R. 1999. Engineering precision RNA molecular switches. *Proc. Natl. Acad. Sci.* **96,** 3584-3589.

Soukup, G.A. & Breaker, R.R. 2001. Allosteric nucleic acid catalysts. *Nucleic Acids Research.* **29,** 1815-1834.

Soukup, G.A., Emilsson G.A.M. & Breaker, R.R. 2000. Altering molecular recognition of RNA aptamer by allosteric selection. *J. Mol. Biol.* **298,** 623-632.

Stojanovic, M.N., de Prada, P. & Landry D.W. 2000. Homogeneous assays based on deoxyribozyme catalysis. *Nucleic Acids Research.* **28,** 2915-2918.

Stojanovic, M.N., de Prada, P. & Landry, D.W. 2001. Catalytic molecular beacons. *Chembiochem.* **2,** 411-415.

Stojanovic, M.N., Mitchell, T.E. & Stefanovic, D. 2002. Deoxyribozyme-based logic gates. *JACS.* **124,** 3555-3561.

Stojanovic, M.N. & Stefanovic, D. 2003. Deoxyribozyme-based half adder. *JACS.* **125,** 6673-6676.

Tang, J. & Breaker, R.R. 1997. Rational design of allosteric ribozymes. *Chem. Biol.* **4,** 453-459.

Todd, A.V., Fuery, C.J., Impey, H.L., Applegate, T.L. & Haughton, M.A. 2000. DzyNA-PCR: Use of DNAzymes to detect and quantify nucleic acid sequences in a real-time fluorescence format. *Clinical Chemistry.* **46,** 625-630.

Yarus, M. 1993. How many catalytic RNAs? Ions and the Cheshire cat conjecture. *FASEB J.* **7,** 31-39.

Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31,** 3406-3415.