CHARACTERIZATION AND APPLICATION OF DEOXYRIBOZYMES

CHARACTERIZATION AND APPLICATION OF SELF-PHOSPHORYLATING DEOXYRIBOZYMES

By

SIMON ANTHONY McMANUS, BSc.

A Thesis Submitted to the School of Graduate Studies in
Partial Fulfillment of the Requirements for the Degree
Doctor of Philosophy

McMaster University Doctor of Philosophy (2012) Hamilton, Ontario

TITLE: Characterization and application of self-phosphorylating deoxyribozymes

AUTHOR: Simon Anthony McManus, B.Sc.

SUPERVISOR: Professor Yingfu Li

NUMBER OF PAGES: xvi; 202

ABSTRACT

The process of in vitro selection has led to the isolation of many catalytic DNA molecules, called deoxyribozymes, which can catalyze a range of biologically-relevant reactions. Despite these advances, questions still remain as to why DNA, which seems more suited to information storage than catalysis can efficiently catalyze chemical reactions. In this thesis, a group of deoxyribozymes that can catalyze their own phosphorylation using NTP substrates are used a model system to study how DNA is able to fold into complex structures necessary for catalysis. Using a variety of structural probing techniques, these studies elucidated a common secondary structure shared by three deoxyribozymes, which do not appear to share a common ancestor sequence. This suggests that this motif may be most efficient motif to catalyze self-phosphorylation by DNA. It also more generally demonstrates that DNA can undergo convergent evolution to reach the same complex folding arrangement. A fourth deoxyribozyme was found to fold into a complex tertiary structure containing a novel quadruplex-helix pseudoknot motif. The finding of this pseudoknot and comparison with other quadruplexes found in other functional nucleic acids led us to investigate whether these stable motifs could be incorporated into nucleic acid libraries to improve the process of in vitro selection and give researchers a better chance of isolating functional nucleic acids. Design and characterization of structured libraries revealed that DNA libraries could be made in which the majority of sequences are folded into quadruplex arrangements. The incorporation of this quadruplex scaffold into DNA sequence libraries may ease the isolation of functional nucleic acids that contain this useful structural motif. In the final part of this thesis, a self-phosphorylating deoxyribozyme was converted from a tool for study of DNA structure to a sensor for GTP and Mn^{2+}, demonstrating that deoxyribozyme substrates can be converted into targets for biosensors.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor Dr. Yingfu Li for giving me the opportunity to work in his innovative and dynamic lab. His advice and guidance has been indispensible in the completion of my PhD studies. I have learned much from him at both the professional and personal levels.

I would also like to thank my supervisory committee members Drs. Murray Junop and Giuseppe Melacini for their useful discussions and suggestions over the course of my studies. Additionally, I would like to thank Dr. Junop for his guidance when I was a young teaching assistant trying to simultaneously learn and teach protein purification. I would also like to thank Dr. Melacini for his technical advise during our attempts to perform NMR studies. I also thank Dr. Alexander Romanschin for allowing me to gain industrial experience in the midst of my doctoral studies. I would also like to thank Drs. Raquel and Richard Epand for the use of their CD spectrometer and their technical advice.

Over the course of my studies, members of the Li lab have provided me with a friendly and helpful working atmosphere. I would particularly like to thank J.C. Achenbach for this mentorship in the early part of my scientific career. I would also like to thank Casey Fowler, Naveen Kumar, Monsur Ali and Razvan Nutiu for all of their advice and friendship during my time in the lab.

On a personal level, I would like to thank my family for their love and support. Finally, I would like to thank my fiancée Wendy for her reassurance, encouragement, and patience as I strived to complete my degree.

Dedicated to my mother, Pat

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

LIST OF ABBREVIATIONS

AppDNA	adenylated DNA
ATMND	2-amino-5,6,7-trimethyl-1,8-naphthyridine
ATP	adenosine 5′- triphosphate
CD	circular dichroism
CEM	crude extracellular mixture
CIAP	calf intestine alkaline phosphatase
CTP	cytidine 5′- triphosphate
DABCYL	4-[4′-(dimethylamino)phenylazo]benzoic acid
Dk1–5	deoxyribozyme kinase 1–5
DNA	deoxyribonucleic acid
dGTP	deoxyguanosine 5′-triphosphate
DMS	dimethyl sulfate
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
FMN	flavin mononucleotide
FRET	Förster resonance energy transfer
GAG	glycoaminoglycan
GTP	guanosine 5′- triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
ITP	inosine 5′-triphosphate
k_{cat}	catalytic rate constant of enzyme
mRNA	messenger RNA

m(7)GTP	7-methylguanosine 5′-triphosphate
N7-dGTP	7-deazaguanosine 5′-triphosphate
NTP	nucleoside 5′-triphosphate
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
pKa	acid dissociation constant
RCA	rolling circle amplification
RNA	ribonucleic acid
T4 PNK	T4 polynucleotide kinase
TAMRA	6-carboxytetramethylrhodamine
TBR	tris(2,2′-bipyridine)-ruthenium
tRNA	transfer RNA
UTP	uridine 5′- triphosphate
VEGF	vascular endothelial growth factor
VEGFR	vascular endothelial growth factor receptor

Chapter 1:
General Introduction

1.1 Organization of Introduction

This thesis spans the structural characterization of a group of deoxyribozymes, the use of some of this structural knowledge to improve in vitro selection, and the conversion of characterized deoxyribozymes into novel biosensors. The introduction will provide background of these areas of research chronologically, starting with the isolation of functional nucleic acids through in vitro selection, descriptions of known deoxyribozymes, structural characterization of deoxyribozymes, and the applications of deoxyribozymes including known deoxyribozyme sensors. The introduction will be concluded with a description of the model self-phosphorylating deoxyribozyme system that formed the foundation of this dissertation.

1.2 Nucleic Acids: From Passive Templates to Functional Molecules

Ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) are the major players in the storage and transfer of genetic information of all life forms on earth. Long strands of double-helical DNA contain the code for all enzymes and structural proteins necessary for cellular life. This code is transcribed to messenger RNA molecules, which is decoded by amino acid loaded transfer RNA molecules to create proteins. When the roles of these molecules were originally deduced, DNA and RNA were envisioned to act as passive information carriers, with proteinacous enzymes being responsible for their creation and replication. These roles seemed logical as nucleic acids are ideal for carrying information as they have an inherent copying mechanism using specific base pairs, while proteins contain a diverse array of amino acid building blocks that allows them to fold into

complex structures necessary for binding other molecules and performing catalysis. This distinct assignment of roles for the major macromolecules was altered with the discovery that RNA was capable of performing catalysis (Kruger et al., 1982), with efficiencies similar to protein enzymes (Cech, 1993). The first of these catalytic RNA molecules discovered were self-splicing introns, RNAs capable of catalyzing their own removal through a cleavage and ligation reaction (Zaug et al., 1983). Following this, several other natural catalytic RNA molecules, termed ribozymes, were found that cleave RNA phosphoester linkages (Forster et al., 1987 ; Feldstein et al., 1989 ; Beattie et al., 1995 ; Been, 1994 ; Guerrier-Takada et al., 1983). These ribozymes were found in divergent organisms from all kingdoms of life. Viral ribozymes that perform self-cleavage during genome replication have been found in plant pathogens. A trans acting ribozyme, RNase P, which processes premature tRNAs by removal of a precursor sequence has been found in bacteria, archea and eukaryotes. In recent years RNase P has also been found to process other RNA molecules and act as a transcription factor for RNA polymerase III (Jarrous et al., 2007).

Apart from catalysis, RNA molecules have also been found to act as regulators. One class of functional nucleic acids found across various bacterial genomes are small RNAs (sRNAs) (Masse et al., 2003), which are important players in regulatory circuits of prokaryotes. These RNAs are trans-acting regulators that regulate gene expression through binding to mRNA or directly to the protein being regulated. Cis-acting RNAs have also been found that act as gene expression regulators (Tucker et al., 2005). These motifs, called riboswitches, are appended to the mRNA of the gene being regulated and

alter either transcription or translation of the associated gene upon binding of a specific metabolite.

These discoveries demonstrated that the paradigm of DNA to RNA to protein was not sufficient to describe the major processes in biological systems and that RNA was capable of performing many of the functions originally thought to be monopolized by proteins. While no naturally occurring functional DNAs have been found to date, DNA does contain similar chemical functionality as RNA, suggesting that is possible that some functional DNA molecules could be found or engineered.

1.3 In Vitro Selection: On-demand Isolation of Functional Nucleic Acids

1.3.1 Early Nucleic Acid Selection

With the discovery that natural RNA molecules served protein-like functions in the cell, researchers sought to develop methods to engineer nucleic acid molecules to perform desired functions. An interesting feature of functional nucleic acid sequences is that they contain phenotypic information by folding into active structures, as well as genotypic information by serving as a template for their own replication. The fact that nucleic acids contain both these properties allows for the direct isolation and amplification of nucleic acid sequences that can perform a specific task. As nucleic acid replication contains a degree of mutation, this allows for the possibility of evolution of a sequence. Early work in this area was performed by Sol Spiegelman on the Qβ RNA phage (Spiegelman et al., 1965). The Qβ phage can be regarded as a functional nucleic acid as it needs to fold into a specific structure in order to be replicated by its native polymerase (Axelrod et al., 1991). By incubating Qβ RNA with purified phage

polymerase and serially transferring the amplified RNA into fresh polymerase, Spiegelman was able to select for sequences that were optimized for replication by truncation of sequences that were not required for polymerase recognition and binding (Klein, 1990). This demonstrated that nucleic acids with improved capacities to perform a desired function could be generated by multiple rounds of selection and amplification. This set the groundwork for the development of in vitro selection. Spiegelman's system is limited in the sense that the selection and replication reactions are performed as a single step meaning this system cannot be used to select for nucleic acids that perform other functions. To select for other types of functional nucleic acids, it is necessary to decouple the selection step from the amplification step.

1.3.2 In Vitro Selection of Functional RNAs

As opposed to Sol Spiegelman's work, which started with a functional RNA that was optimized through the acquisition of mutations during replication, the challenge for isolation of *de novo* functional nucleic acids is that there is no starting sequence to provide a basal level of activity in the first round of selection. To overcome this, modern in vitro selection is initiated using very large sequence libraries of up to 10^{16} individual sequences. These initial libraries are used in selection schemes in which sequences that perform the desired function are separated from inactive sequences and amplified by polymerase chain reaction (PCR). This ability of separating active sequences gives in vitro selection a major advantage over other techniques, such as screening, especially in the initial round of selection when there is likely only a small number of sequences that can perform the activity of interest in a vast pool of inactive sequences. This allows the

use of very large sequence libraries, which is not feasible with screening techniques that rely on assessing the activity of sequences individually. The first uses of this type of in vitro selection strategy were performed for the isolation of RNA aptamers, RNA molecules that can bind specific ligands (Ellington et al., 1990 ; Tuerk et al., 1990). These schemes involved the binding of sequences from RNA libraries to target molecules that were immobilized to a solid support. These binding sequences were extracted and converted to DNA by reverse transcription before being amplified by PCR and transcribed to RNA for the next round of selection. Variations of this selection scheme have been used to isolate numerous functional RNAs, which can bind a plethora of ligands (Stoltenburg et al., 2007) and catalyze a wide range of chemical reactions (Chapman et al., 1994).

1.3.3 In Vitro Selection of Functional DNA Molecules

In vitro selection also provided the first evidence that DNA molecules could form functional nucleic acids. Jack Szostak's group produced the first functional DNA through in vitro selection for DNA aptamers for several ligands (Ellington et al., 1992). Two years later, the first catalytic DNA, or deoxyribozyme, was isolated after selection for DNA molecules that could cleave a phosphate linkage adjacent to a single RNA residue embedded in a DNA substrate (Breaker et al., 1994). This selection scheme is shown in Figure 7.1A. Isolation of deoxyribozymes was carried out on a substrate-library chimera modified with a biotin moiety at the end of the substrate domain, which bound the library to a streptavidin-coated column. Selection buffer was added and sequences that were capable of cleaving the RNA linkage could be eluted from the column. This

allowed for the separation of a small number of catalytic DNA sequences from the majority of inactive sequences that remain bound to the column. Once isolated, these sequences were amplified by PCR, which allowed the generation of multiple copies of the initial catalytic sequences. The procedure was then repeated multiple times, enriching the pool for catalytic sequences *via* iterative rounds of selection and amplification. Many subsequent in vitro selection schemes have used this method of coupling catalysis to the addition or removal of a tag during the selection step. Catalytic molecules can then be isolated through affinity chromatography, as described above, or based on the size difference between the tagged and untagged sequences. This technique is also tuneable such that the components of the buffer used in the selection step will interact with the DNA library and can become incorporated into the active structures of the selected deoxyribozymes. This typically leads to deoxyribozymes that are dependent on the buffer components, such as metal ions, for catalysis. Over the past three decades, in vitro selection has been employed to isolate countless deoxyribozymes that can catalyze a fascinating array of chemical reactions.

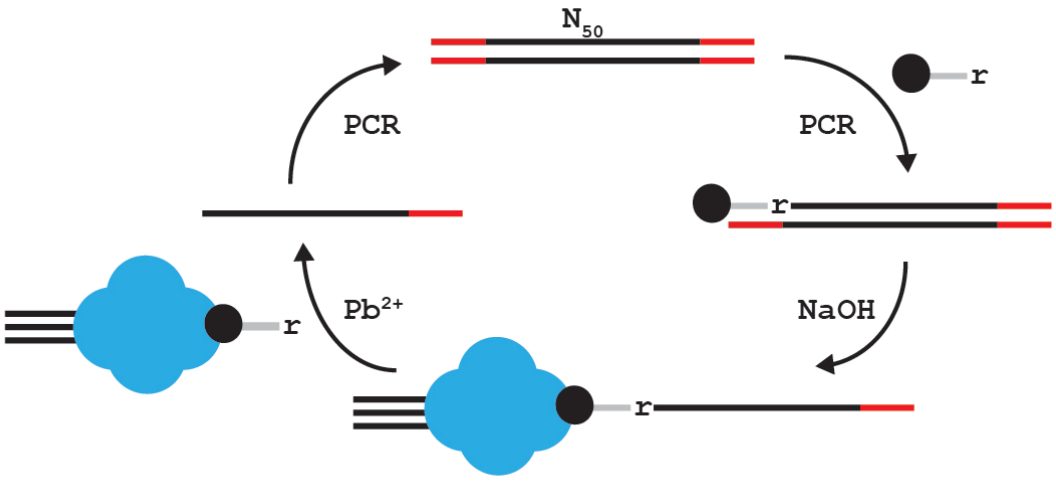


Figure 1.1 In vitro selection of deoxyribozymes. In vitro selection scheme used for the isolation of the first DNAzyme. PCR is used to incorporate a primer (grey) containing a biotin label (black circle) and a ribonucleotide (r) into a library with a 50-nucleotide random region. After binding of the library to a streptavidin column (blue complex), the antisense sequences are removed by NaOH and the bound sequences are incubated in buffer containing Pb^{2+}. Sequences that undergo RNA cleavage are removed from column and amplified by PCR. PCR primer binding sites are shown in red.

1.4 Existing Classes of Deoxyribozymes

1.4.1 RNA-Cleaving Deoxyribozymes

The first in vitro selection for catalytic DNA isolated a deoxyribozyme that can cleave RNA. It is logical that this is the first reaction for which deoxyribozymes were isolated, as RNA cleavage is a reaction for which DNA is well suited to catalyze. Firstly, RNA is a natural binding partner of DNA, and they can form base pairs as well as other interactions such as base stacking. Furthermore, RNA is inherently prone to cleavage through intramolecular attack of the 2′ hydroxyl group of the ribose on the adjacent bridging phosphate within the RNA chain. DNA needs only to bind and position the RNA in the necessary in-line orientation for cleavage to take place. This has led to the

isolation of a variety of deoxyribozymes capable of cleaving RNA substrates, or chimeric DNA-RNA substrates, under a wide range of conditions (Breaker et al., 1994 ; Breaker et al., 1995 ; Faulhammer et al., 1996 ; Santoro et al., 1997 ; Geyer et al., 1997 ; Roth et al., 1998 ; Feldman et al., 2001 ; Perrin et al., 2001 ; Lermer et al., 2002 ; Cruz et al., 2004).

As mentioned earlier, the first deoxyribozyme selection was performed by the laboratory of Gerald Joyce, and led to the isolation of an RNA-cleaving deoxyribozyme that could cleave an RNA linkage embedded in a DNA substrate in a buffer containing Pb^{2+} and monovalent metal ions (Breaker et al., 1994). After five rounds of selection, a sequence that could cleave the substrate with a rate enhancement of 10^5 over the spontaneous cleavage reaction was identified. Structural analysis revealed that the deoxyribozyme formed two helical interactions with the substrate on either side of the cleavage site, with 15 additional nucleotides necessary for catalysis. This two binding arm structure with a central catalytic domain has subsequently been observed in the majority of RNA-cleaving deoxyribozymes and its significance will be discussed further in the deoxyribozyme structure section. This deoxyribozyme was also found to be able to perform bimolecular catalysis, similar to a true enzyme, using a substrate and deoxyribozyme strand, with a turnover rate of 1 min^{-1}. These rates and enhancements over spontaneous reaction are comparable to many natural ribozymes. Other early selection endeavours by the Joyce (Breaker et al., 1995) and Famulok (Faulhammer et al., 1996) groups also isolated deoxyribozymes that can cleave an embedded RNA linkage in the presence of Mg^{2+} and Ca^{2+}, respectively.

After these initial in vitro selection experiments demonstrated that DNA could perform catalysis under different reaction conditions, the Joyce group selected for deoxyribozymes that could cleave an all-RNA substrate under physiological conditions (Santoro et al., 1997). Deoxyribozymes that could function under these conditions have the potential to target cellular RNAs for research and therapeutic applications. Using a similar protocol as their initial selection, two classes of deoxyribozymes that cleaved an RNA substrate at two distinct cleavage sites were isolated. These two deoxyribozymes were named 8-17 and 10-23 as they represented the 17^{th} and 23^{rd} clones isolated in the eighth and tenth round of selection, respectively. These two constructs, shown in Figure 1.2 have become the best-studied representatives of all deoxyribozymes. Characterization of the 10-23 deoxyribozyme revealed that it has a catalytic rate of 0.1 min^{-1} under physiological conditions and a catalytic efficiency of 10^9 $M^{-1}min^{-1}$. By changing the sequences of its binding arms, 10-23 could be made to cleave almost any RNA with a purine-pyrimidine junction at the cleavage site. The other deoxyribozyme isolated from this selection, 8-17, has also proved to be a versatile deoxyribozyme. Reselection and mutagenesis analysis revealed its small catalytic core contained a small three base pair stem and 8 essential nucleotides. The 8-17 motif gained recognition due it being isolated multiple times in several independent in vitro selections (Santoro et al., 1997 ; Faulhammer et al., 1996 ; Li, J. et al., 2000 ; Cruz et al., 2004 ; Schlosser et al., 2004). Recent mutational analysis using abasic and carbon spacer mutants of 10-23, have even provided evidence that 10-23 is a specialized version of the 8-17 motif (Wang et al., 2010). The repeated isolation of this motif from different DNA libraries suggests that it

may be the simplest structural motif for DNA-based RNA-cleavage. Variations of the 8-17 motif were also shown to use different metal ion co-factors such as Mg^{2+} (Santoro et al., 1997), Ca^{2+} (Faulhammer et al., 1996), and Zn^{2+} (Li, J. et al., 2000). 8-17 also emerged to be flexible in substrate recognition. In a study to isolate deoxyribozymes that could cleave all sixteen possible dinucleotide junctions, deoxyribozymes containing the 8-17 motif were found to cleave fourteen of the possible sixteen junctions (Cruz et al., 2004). The small size and versatility of 8-17 has led to it being utilized in a number of different sensor schemes.

1.4.2 Deoxyribozymes that Catalyze other Reactions

In addition to RNA-cleaving deoxyribozymes, deoxyribozymes that catalyze many other chemical reactions have been discovered, as listed in Table 1.1. Several deoxyribozymes have been isolated by the Silverman group that catalyze the reverse reaction of RNA cleavage, the ligation of two RNA substrates. Some of these deoxyribozymes can catalyze the ligation of an RNA to the end of another RNA molecule creating a 2′-5′ or 3′-5′ linkage and producing a longer RNA (Flynn-Charlebois et al., 2003 ; Purtha et al., 2005). Alternatively, others catalyze the ligation between the 5′ end of one RNA with an internal 2′-hydroxyl of another RNA producing branched RNA (Pratico et al., 2005). Many deoxyribozymes have been isolated that process DNA substrates. An early study by the Szostak group selected for DNA molecules that could catalyze DNA ligation. DNA ligase deoxyribozymes were isolated that could mediate the ligation of one DNA strand to a second modified DNA molecule by the condensation of a 5′-hydroxyl to a 3′ phosphorimidazolide (Cuenoud et al., 1995). The deoxyribozyme was

found to function in the presence of Zn^{2+} or Cu^{2+} and was capable of performing ligation in a multiple turnover fashion. The Ellington group has also isolated ligase deoxyribozymes that can ligate two DNA oligonucleotides containing a 3′-phosphorothioate and a 5′-iodine group (Levy et al., 2001). A deoxyribozyme that ligates natural DNA substrates by mimicking the second reaction catalyzed by T4 DNA ligase has also been isolated by the Breaker group (Sreedhara et al., 2004). This deoxyribozyme catalyzes the ligation of adenylated DNA (AppDNA) to an acceptor DNA oligonucleotide containing a 3′-hydroxyl group. A deoxyribozyme that mimics the other reaction catalyzed by T4 DNA ligase, DNA adenylation, has also been isolated (Li, Y. et al., 2000). This deoxyribozyme catalyzes the addition of an AMP cap to a 5′ phosphorylated DNA molecule and found practical use in the preparation of substrates for the aforementioned ligase deoxyribozyme. Deoxyribozymes that mimic the reaction catalyzed by T4 polynucleotide kinase (PNK), by transferring the γ-phosphate from an NTP molecule to the 5′ end of the deoxyribozyme have also been isolated (Li et al., 1999 ; Wang, W. et al., 2002). The characterization and utilization of a group of these self-phosphorylating, or kinase, deoxyribozymes will be the focus of this thesis. The Breaker group performed a selection for DNA molecules capable of self-cleavage (Carmi et al., 1996). As DNA is much less labile than RNA, the selection scheme was designed to induce hydroxyl radical based DNA cleavage. The selection was carried out in the presence of Cu^{2+} with ascorbate acting as the reducing agent. An isolated deoxyribozyme was later optimized to act as a restriction endonuclease that can target a variety of different single-stranded DNA molecules (Carmi et al., 1998). Deoxyribozymes have

also been recently isolated to cleave a DNA substrate via a hydrolytic mechanism as opposed to oxidative cleavage (Chandra et al., 2009). Another deoxyribozyme, which cleaves DNA through site-specific depurination of a DNA substrate, has also been identified (Sheppard et al., 2000). This deoxyribozyme catalyzes hydrolysis of the N-glycosidic bond of a deoxyguanosine residue leading to subsequent strand scission by β-elimination. Another deoxyribozyme has been designed to process a DNA substrate that has been damaged by UV radiation (Chinnapen et al., 2004). This deoxyribozyme repairs a thymine cyclobutane dimer in the presence of 305 nm light using a mechanism similar to natural photolyase enzymes.

Deoxyribozymes are also capable of processing non-nucleic acid substrates. One early deoxyribozyme study by the Sen group led to the isolation of DNA molecules that could catalyze porphyrin metalation (Li et al., 1996). This deoxyribozyme was obtained by performing a selection for aptamers that can bind a N-meythlmesoporphyrin IX, a transition state analogue of the metalation of mesoporphyrin IX. Following selection, aptamer sequences were found to be able to catalyze metalation. Guanine quadruplex-forming DNA molecules from this selection were also found to be aptamers for hemin and enhance its low peroxidase activity by 250-times (Travascio et al., 1998). This aptamer was subsequently used in many sensing applications (Willner et al., 2008 ; Moshe et al., 2009 ; Teller et al., 2009 ; Pelossof et al., 2010 ; Liu, X. et al., 2011). Another deoxyribozyme has been isolated that can catalyze carbon-carbon bond formation through the Diels-Alder reaction (Chandra et al., 2008). Another bond-forming deoxyribozyme has been isolated that catalyzes the formation of nucleopeptide

bonds between amino acid side chains of tyrosine or serine and triphosphorylated RNA substrates (Pradeepkumar et al., 2008 ; Sachdeva et al., 2010 ; Wong et al., 2011).

1.4.3 Deoxyribozymes with Modified Nucleotide Building Blocks

One of the major criticisms of DNA and its potential as a catalyst is the fact that it is limited in the number of different interactions it can form, as it is composed of only four building blocks. Although numerous deoxyribozymes that can perform efficient catalysis are currently available despite this constraint, the addition of chemical moieties with catalytic functionality could allow for more efficient catalysts. For example, DNA modified with functional groups that mimic the side chains of amino acids involved in catalysis by hydrolytic enzymes could allow for the selection of nucleases and proteases with protein enzyme like efficiencies. In vitro selection with these modified DNAs would allow coupling of the ability to isolate nucleic acids with the proper folding from large DNA libraries and the efficient catalysis achievable with particular amino acids.

It has been shown that the modified nucleoside triphosphates (5-(3-amino-allyl)-2′-deoxyribouridine-5′-triphosphate and (8-(2-(4-imidazolyl)ethylamino)- 2′-deoxyriboadenosine-5′-triphosphate can be incorporated into DNA using DNA polymerase. This modified DNA can also be amplified by PCR (Latham et al., 1994 ; Perrin et al., 1999). This ability to synthesize (by polymerization) and amplify modified DNA sequences makes them amenable to in vitro selection. One early example showcasing this strategy was used to develop an RNA-cleaving deoxyribozyme modified with C5-imidozole deoxyuridine residues (Santoro et al., 2000). Moreover, using DNA libraries polymerized with two nucleotides modified with imidazole and cationic amine-

groups, RNase-mimicking deoxyribozymes that could perform acid-base catalysis in the absence of any metal ion co-factors were isolated (Perrin et al., 2001 ; Lermer et al., 2002 ; Sidorov et al., 2004). Characterization of one of these deoxyribozymes isolated by the Perrin group showed that two imidozoles were involved in general acid and base catalysis with a cationic amine stabilizing the transition state, mimicking the reaction catalyzed by the RNaseA enzyme (Ting et al., 2004b ; Thomas et al., 2009). Modified deoxyribozymes also have the potential to bind to co-factors and substrates that are not well-suited to interact with DNA. A previously isolated imidazole-modified deoxyribozyme has been developed into an Hg^{2+} sensor, exploiting interactions known to form between imidazole groups and Hg^{2+}. These interactions severely inhibit deoxyribozyme activity, allowing the correlation of increased inhibition of activity with increased concentration of Hg^{2+} (Thomas et al., 2004). Imidazole-modified deoxyribozymes that are dependent on Hg^{2+} for activity have also been isolated (Hollenstein et al., 2008). These two deoxyribozymes have the potential to be developed into "turn on" and "turn off" sensors for Hg^{2+}.

Modified deoxyribozymes that catalyze other reactions have also been isolated. A deoxyribozyme containing both primary amines and imidazole groups have been found to perform DNA cleavage by forming a Schiff base with an abasic site on a DNA substrate (May et al., 2004). Modification of deoxyribozymes has also been used to create a deoxyribozyme that is activated by light (Ting et al., 2004a). Incorporation of a light sensitive nucleoside, 8-(2-(4-imidazolyl)ethyl-1-thio)-2'-deoxyriboadenosine, into the 8-17E deoxyribozyme renders it inactive. Activity is restored through photoinduced reversion of the modified nucleoside to adenosine. In the search for deoxyribozymes

modified with many functional groups, deoxyribozymes that contain nucleotides modified with amines, guanidines and imidazoles have been developed (Hollenstein et al., 2009). The inclusion of guanidine nucleotides allows for activity at higher temperature, adding more functionality to the deoxyribozymes enhanced catalytically with amine and imidazole groups. Current study in this area involves the investigation of whether modifications could enhance the activity of known deoxyribozymes such as 10-23 (Lam et al., 2010). The types of modifications incorporated into deoxyribozymes has also recently been expanded to groups such as a phenol mimicking a tyrosine side chain (Lam et al., 2011).

Table 1.1 Reactions catalyzed by deoxyribozymes

Reaction Catalyzed	Co-factors	Reference
RNA Cleavage	Pb^{2+}, Mg^{2+}, Ca^{2+}, Hg^{2+}, UO_2^{2+}, Zn^{2+}, co-factor independent	(Breaker et al., 1994; Breaker et al., 1995; Faulhammer et al., 1996; Santoro et al., 1997; Geyer et al., 1997; Roth et al., 1998; Feldman et al., 2001; Perrin et al., 2001; Lermer et al., 2002; Cruz et al., 2004; Santoro et al., 2000; Sidorov et al., 2004; Hollenstein et al., 2009; Liu et al., 2007a; Hollenstein et al., 2008)
RNA ligation	Mg^{2+}, Mn^{2+}, Zn^{2+}	(Flynn-Charlebois et al., 2003; Purtha et al., 2005; Pratico et al., 2005)
DNA ligation	Cu^{2+}, Zn^{2+}, Mg^{2+}, Ca^{2+}, Mn^{2+} co-factor independent	(Cuenoud et al., 1995; Levy et al., 2001; Levy et al., 2002b; Sreedhara et al., 2004)
DNA adenylation	Cu^{2+}	(Li, Y. et al., 2000)
DNA phosphorylation	$Ca^{2+,}$ Cu^{2+}, Mg^{2+}, Mn^{2+}	(Li et al., 1999; Wang, W. et al., 2002)
DNA cleavage	Cu^{2+}, ascorbate	(Carmi et al., 1996; Carmi et al., 1998; May et al., 2004)
DNA hydrolysis	Mn^{2+}, Zn^{2+}	(Chandra et al., 2009)
N-glycosylation	Ca^{2+}, Mg^{2+}, Mn^{2+}, Ba^{2+}	(Sheppard et al., 2000)
Thymine dimer repair	Serotonin, co-factor independent	(Chinnapen et al., 2004)
Carbon-carbon bond Formation	Ca^{2+}, Mg^{2+}, Mn^{2}	(Chandra et al., 2008)
Nucleopeptide bond formation	Mg^{2+}, Mn^{2+}	(Pradeepkumar et al., 2008; Sachdeva et al., 2010; Wong et al., 2011)

Porphyrin metalation	Cu^{2+}, Zn^{2+}	(Li et al., 1996)

1.5 Structures Adopted by Deoxyribozymes

1.5.1 The Two Binding Arm Motif

To date, much study has been done to characterize deoxyribozyme structure and structural models of many deoxyribozymes have been deduced. We have produced an extensive review on the current state of deoxyribozyme structural study (McManus et al., 2010). The simplest and most common structural motif used by deoxyribozymes is the two binding arm motif. This structural arrangement contains two regions that are complementary to two regions of a nucleic acid substrate flanking the reaction site. The catalytic core of the deoxyribozyme is located between the two binding arms opposite the reaction site. This structure is most commonly seen in deoxyribozymes that cleave RNAs. This type of structure is also observed with many deoxyribozymes that catalyze the reverse RNA ligation reactions. Here, the two binding arms base pair with the two RNA substrates that are to be ligated. The two most studied RNA-cleaving deoxyribozymes, 10-23 and 8-17, both contain the two-binding arm motif but show very different structural arrangements in their catalytic core despite being isolated in the same in vitro selection (Santoro et al., 1997). Extensive structural study on 8-17 has shown its small catalytic core to contain six required nucleotides and a three base-pair stem that can accommodate any variety of sequence providing that base pairing is maintained (Peracchi et al., 2005 ; Cruz et al., 2004). In contrast, the catalytic core of 10-23 has not been found to contain any Watson-Crick base-pairing with most of the individual nucleotides within the catalytic core being very intolerant to mutations (Zaborowska et al., 2005 ;

Zaborowska et al., 2002). This suggests that many of the nucleotides within the catalytic core of 10-23 must be involved in tertiary interactions, allowing the catalytic domain to fold into its active conformation and perform catalysis. These two different variations on the two-binding arm framework demonstrate that significant variations within the core of this motif are possible. To look for a higher degree of structural diversity, or deoxyribozymes with different structural scaffolds, it is necessary to look at deoxyribozymes that catalyze more complex reactions. It appears that the simplicity of the RNA cleavage reaction has allowed the smallest, simplest structures to dominate in vitro selections, as smaller motifs are more likely to occur in random DNA libraries used to initiate the selections and hijack in vitro selections at the expense of larger sequences that may be efficient catalysts but are sampled less in random sequence spaces. To catalyze more complex reactions, DNA requires more complex structures to bind exogenous substrates and co-ordinate additional co-factors.

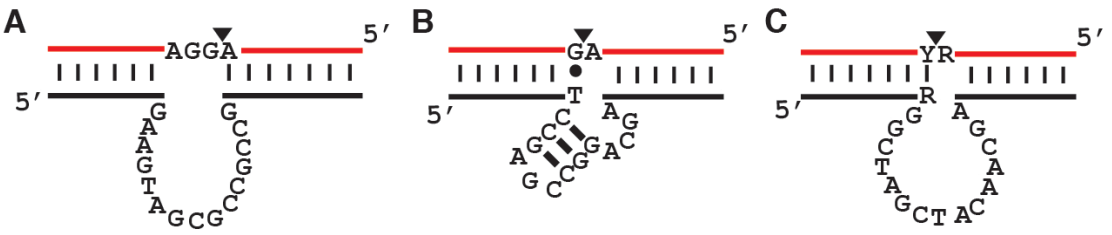


Figure 1.2 RNA-cleaving deoxyribozymes using two binding arm arrangements. Three deoxyribozymes that can cleave RNA. The DNAzymes are shown in black and the substrates are shown in red. The left structure is the first isolated Pb^{2+}-dependent DNAzyme. The center and right structure are the 8-17 and 10-23 DNAzymes, respectively. R: purine based nucleotide; Y: pyrimidine based nucleotide.

1.5.2 Complex Helical Arrangements

Although the two binding arm motif seems to be the favoured structural arrangement for RNA-cleaving deoxyribozymes as well as some deoxyribozymes that catalyze the reverse RNA ligation reaction (Flynn-Charlebois et al., 2003), other systems appear to use different structural arrangements. Other structures have been found for deoxyribozymes that catalyze other reactions. DNA ligase deoxyribozymes, though catalyzing similar reactions to RNA ligase deoxyribozymes, have been found to have more complex secondary structures. The first ligase deoxyribozyme, named E47, which ligates the 5′ hydroxyl of one DNA substrate to a second activated DNA molecule and creates a phosphodiester bond (Cuenoud et al., 1995) was found to have a secondary structure containing a binding arm for one substrate at the 3′ end of the deoxyribozyme with a TTT bulge near the ligation site while the 5′ end binds the other substrate, as shown in Figure 1.3a. Interestingly, this 5′ portion also contains an internal stem, creating a three-way helical junction. With this structure, the nucleobases across from the ligation site are involved in base-pairing interactions with the substrate. Therefore, this DNA-ligating deoxyribozyme must be using residues from distal regions to perform catalysis.

Another DNA-ligating deoxyribozyme that ligates an adenylated DNA to a second DNA substrate also contained a unique structural arrangement (Sreedhara et al., 2004). As shown in Figure 1.3b, the AppDNA substrate is held in place through six base pairs with the deoxyribozyme, while the other substrate was found to have eight base-pair interactions with the deoxyribozyme. Notably, this secondary substrate contains eleven unpaired nucleotides adjacent to the ligation site. This suggests that the deoxyribozyme

uses base pairing in combination with other unidentified interactions to bind this substrate and place it in the correct orientation for catalysis. Another distinct feature of this deoxyribozyme in relation to the previously discussed deoxyribozymes is the large amount of essential sequence outside the substrate-binding region on the 5′ end of the molecule. This sequence is probably involved in undefined tertiary interactions, adding to the complexity of the structure.

Another deoxyribozyme with a distinct secondary structural arrangement is the 10-28 deoxyribozyme, which catalyzes the N-glycosylation of a specific guanine residue within a DNA substrate (Sheppard et al., 2000). As shown in Figure 1.3c, this deoxyribozyme has a complex secondary structure containing several entwined helices. The model contains two stem-loops at the 5′ and 3′ ends, as well as two stems in the center of the deoxyribozyme. The stems in the center are entwined in a type of double pseudoknot with part of each stem being in the loop at the end of the other stem. The observation that 10-28 can use such a complex arrangement of interconnected double-helical interactions illustrates that deoxyribozymes can have substantial structural complexity, at least on the secondary structural level.

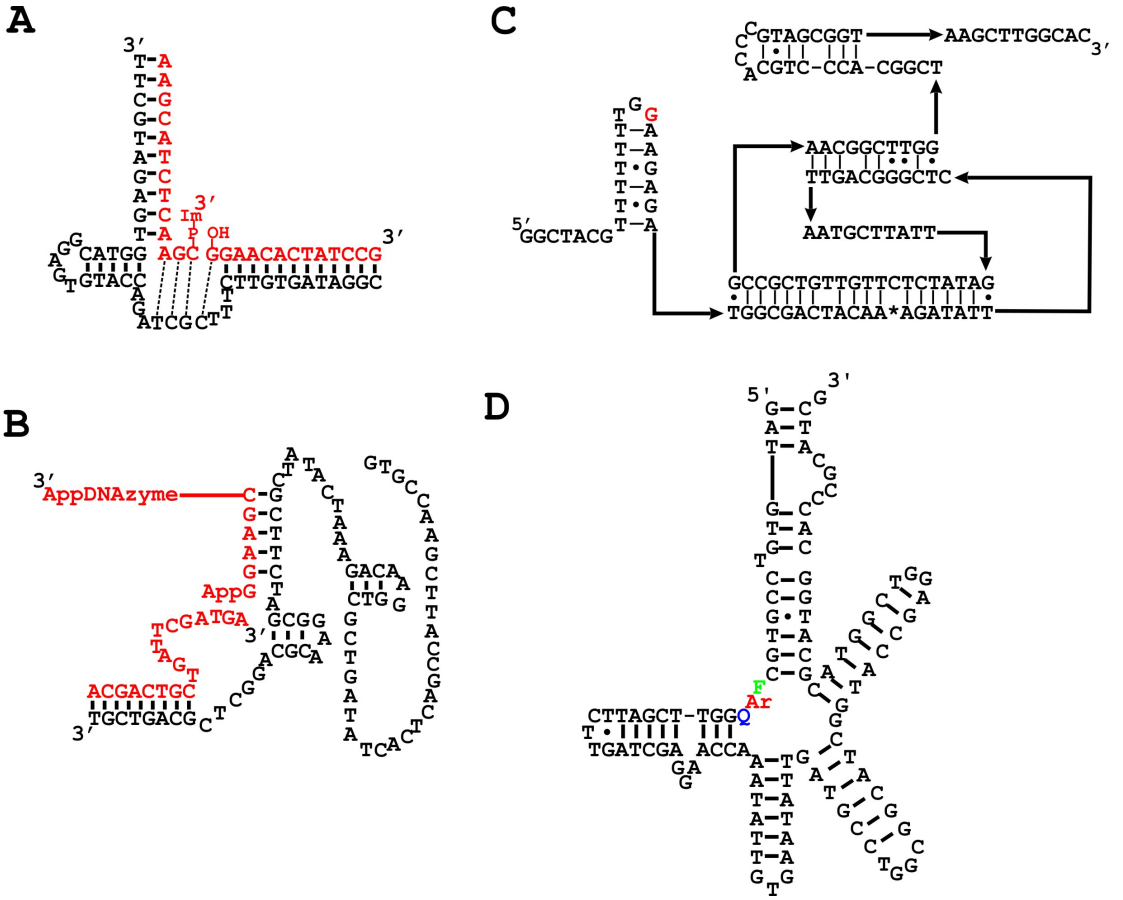


Figure 1.3 Complex helical arrangements used by Deoxyribozymes. A) Secondary structure of a DNA ligase deoxyribozymes that ligates a substrate with a 5′ imidazole to a second substrate with a 3′ hydroxyl. The substrate strands are shown in red. B) Secondary structural model for a DNA ligase deoxyribozyme that can ligate a substrate with an adenylated guanine (AppG) with a substrate with a 3′ hydroxyl. The two substrate strands are shown in red. C) Proposed secondary structure model for the 10-28 N-glycosylase DNAzyme. The site of depurination is shown in red. The * shows the point where the deoxyribozyme and substrate portions can be separated to form a trans construct. D) The secondary structure of a fluorogenic RNA-cleaving deoxyribozyme containing a five-way helical junction. The riboadenosine cleavage site is shown in red. F and Q represent fluorescein-dT and dabcyl-dT, respectively.

In addition to pseudoknots, deoxyribozymes that utilize other complex helical arrangements have also been isolated. In one study, an in vitro re-selection experiment was performed to increase the efficiency of an RNA-cleaving fluorescence-generating deoxyribozyme that utilized a three-way helical junction in its active structure (Chiuman

et al., 2006). After sequencing the population from the fifth round of the re-selection, two sequence classes containing a five-way helical junction arranged in a star-like pattern in their active structure were found. After additional rounds of selection under stringent selection conditions, the star-structured sequences out-competed the original three-way junction motif and became the dominant structural class, suggesting that this complex structure is more efficient at performing this particular reaction.

1.5.3 Non-helical Arrangements

It has long been known that DNA is capable of forming structures other than the Watson-Crick double helix commonly seen in biological systems. Two prominent tertiary structures are guanine quadruplexes and triple helices. These structures have also been found in the active structures of deoxyribozymes. Triple-stranded DNA structures are formed by stretches of pyrimidine-purine-pyrimidine base triples such as T˙AT and C^{+}˙GC (Morgan et al., 1968 ; Beal et al., 1991 ; Wells et al., 1988). A triple helix has been found in the active structure of a deoxyribozyme that cleaves a DNA substrate (Carmi et al., 1998). During secondary structural analysis, they noticed that a string of unpaired pyrimidines in the substrate had the correct sequence to form a triple-helical complex with a polypurine-polypyrimidine stem in the deoxyribozyme, as shown in Figure 1.4a. Through alterations of nucleotides, they were able to show that the deoxyribozyme was active with different variations of the triple helix, but became inactive if the triplex was disrupted, showing that this DNA-cleaving deoxyribozyme indeed contains a triplex in its active structure.

Guanine quartets consist of four guanines arranged in a planar square arrangement with two hydrogen bonds formed between each adjacent guanine (Gellert et al., 1962 ; Arnott et al., 1974 ; Sen et al., 1988). Stacked quartets called quadruplexes have been found to be very stable and have been isolated in telomeric sequences at the end of chromosomes, as well as in the promoter regions of certain genes (Aboul-ela et al., 1992 ; Williamson et al., 1989 ; Wang et al., 1993 ; Huppert et al., 2007). Investigations have shown that several deoxyribozymes also contain guanine quadruplexes in their active structures (Li et al., 1997 ; Chinnapen et al., 2004 ; Li, Y. et al., 2000 ; Li et al., 1999). One interesting example is a thymine dimer repair deoxyribozyme with a two-tiered guanine quadruplex directly across from the dimer repair site, as shown in Figure 1.4c. It is hypothesized that apart from a structural role, this quadruplex may also act as an antenna to harness light energy used to repair the dimer. If this is the case, it demonstrates that deoxyribozymes may be capable of using different structural folds to alter their properties, possibly allowing DNA to catalyze complex reactions.

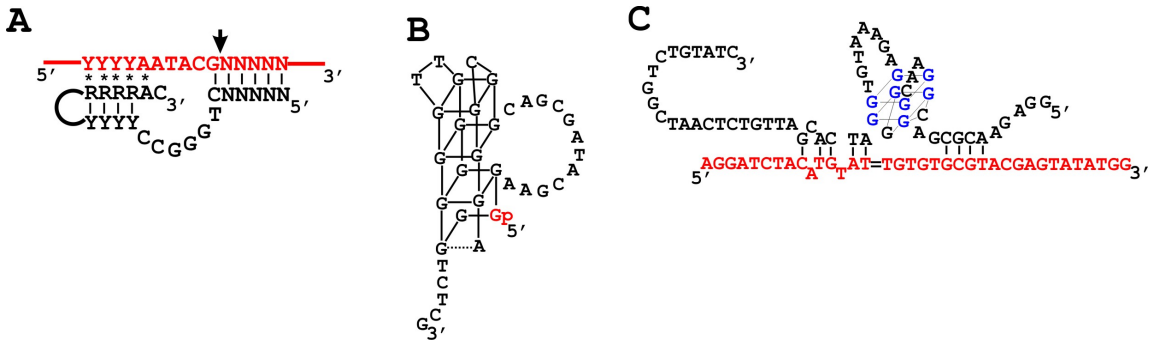


Figure 1.4 Non-Watson Crick interactions in deoxyribozymes. A) The secondary structure of a Cu^{2+}-dependent DNA-cleaving deoxyribozyme. The substrate is shown in red. Y, R, and N represent a pyrimidine, purine, and any nucleotide, respectively. The major cleavage site is shown with an arrow. Triplex interactions are shown with *. B) A secondary structure model for a self-adenylation deoxyribozyme containing a three and a half tier quadruplex. The phosphorylated guanine that is adenylated is shown in red. C) The secondary structure of a thymine dimer repair deoxyribozyme containing a two-tiered guanine quadruplex. The substrate is shown in red and guanines involved in quadruplex interactions are shown in blue.

Intriguingly, while several deoxyribozymes have been shown to contain quadruplexes in their active structures, quadruplexes are not observed in ribozymes. The known classes of natural ribozymes, whose structures have been intensively studied (reviewed in (Lilley, 2005)), use extensive Watson-Crick base-pairing in their structural scaffolds. Artificial ribozymes isolated by in vitro selection have also been found to contain primarily structural scaffolds built with Watson-Crick helical elements (Chen et al., 2007). Thus, it appears that deoxyribozymes are more prone to incorporating quadruplex arrangements in their active structures than ribozymes. The reasons for this difference in structural preference are unclear at present. Quadruplexes have been rationally designed into ribozymes to act as regulatory components (Beaudoin et al., 2008; Wieland et al., 2006), and a quadruplex-containing deoxyribozyme has been converted into a ribozyme (Travascio et al., 1999), showing that quadruplexes can be

utilized by RNA catalysts. The presence of quadruplexes in many deoxyribozyme structures may reflect the fact that a higher number of strand conformations are available for DNA quadruplexes than for RNA quadruplexes. RNA quadruplex-forming sequences have been shown to fold exclusively into all-parallel conformations due to orientation of the ribose pucker (Joachimi et al., 2009), while DNA quadruplexes have been shown to adopt parallel, antiparallel and mixed strand orientations (Neidle et al., 2006). This increase in the number of possible conformations may translate into more functional (catalytic) quadruplex scaffolds being present in a random-sequence DNA pool than an RNA pool.

The deoxyribozyme structures discussed thus far are based on interactions that are prominent at neutral pH. At low pH (high acidity), potential hydrogen-bond acceptors, such as the N3 of cytidine (pKa = 3.5) and the N1 of adenosine (pKa = 4.2), become protonated, which disfavours the formation of Watson-Crick base pairs. With these protonations, different interactions, such as base pairs involving protonated cytosines (C^{+}· C), adenines (A^{+}· A^{+}) or triple-base interactions (C^{+}· CG) become favourable (Bloomfield et al., 2000).

To test whether deoxyribozymes could form different types of structures in acidic conditions, a series of studies were carried out to isolate and characterize RNA-cleaving deoxyribozymes that function at a range of acidic pHs. Several rounds of in vitro selection were performed at pH 4, and the active population was subsequently subjected to five parallel selections performed at pH 3, 4, 5, 6 and 7 (Liu, Z. et al., 2003). Several catalytically active sequences were found in each reaction pool. Overall, most sequences

were found in only one pH population and no common sequences were found in all five pools. The lack of a common sequence across the five populations suggests that different structures are necessary to perform RNA-cleavage at different pHs. Structural characterization revealed that the deoxyribozymes functioning near neutral pH utilized Watson-Crick base-pairing, while those functional at lower pH used different interactions. Deoxyribozyme pH7DZ1, functional at pH 7, was found to contain two Watson-Crick base-paired regions in its active structure (Liu, Z. et al., 2003). Similarly, pH6DZ1, which displayed optimal activity at pH 6, was also found to be highly helical, containing a four-way junction (Shen et al., 2005). In contrast, pH5DZ1, a deoxyribozyme that was functional at pH 5, was not found to contain any Watson-Crick base-pairing interactions after exhaustive mutational analysis (Kandadai et al., 2009). Similar results were obtained with deoxyribozymes that function at pH 4 (pH4DZ1) (Kandadai et al., 2005) and pH 3 (pH3DZ1) (Ali et al., 2007a), suggesting that these deoxyribozymes must be folding into different types of structures mediated by non-Watson-Crick interactions. Chemical footprinting analysis also suggested the presence of different structural interactions in the low pH deoxyribozymes. For instance, methylation of certain guanine residues was found to lead to an increase in activity, contrasting with the usual observation of methylation interfering with deoxyribozyme activity. High-resolution study of these deoxyribozymes will be of particular value as it can reveal the exact nature of these interactions and show which structural folds are preferable to deoxyribozymes that function at low pH.

1.6. Using the 10-23 Deoxyribozyme as a therapeutic agent

One of the goals of early in vitro selections was to develop deoxyribozymes that could be used in therapeutic applications. DNA contains many attributes that make it a good candidate as a therapeutic agent. DNA is highly stable both chemically and thermodynamically. It can also be synthesized quickly and relatively inexpensively. Selections to isolate deoxyribozymes that cleave all RNA substrates, as opposed to the original selections that targeted DNA-RNA chimeras, were carried out in part to isolate deoxyribozymes with the potential to cleave mRNA molecules leading the gene knockdown. Of the known deoxyribozymes, the 10-23 RNA-cleaving deoxyribozyme seems to have the most therapeutic potential for several reasons. First it is the fastest of all known deoxyribozymes, with a k_{cat} of 10 min^{-1}, and it retains a reasonable cleavage rate under physiological conditions. Secondly, its binding arms, which recognize its RNA substrate through Watson-Crick base pairing, have been shown to be very tolerant of nucleotide changes, allowing it to cleave any small RNA molecule with an unpaired purine-pyrimidine linkage. When trying to design a 10-23 construct for a therapeutic application, some factors need to be considered. First is selection of an appropriate cleavage site in the RNA of interest. Although 10-23 is known to cleave any purine-pyrimidine junction in small RNA molecules, larger RNAs that are present *in vivo* such as mRNAs can form secondary and tertiary interactions that can inhibit 10-23 binding and cleavage of a target. Appropriate cleavage sites can be selected *in vitro* using longer length RNA transcripts along with structural prediction to find regions of RNA lacking significant higher-order structure. Another factor is that while DNA is more stable than

RNA and proteins, it is still susceptible to degradation during and after administration in serum and in cellular environments. Modifications have been utilized in some cases to increase the stability of deoxyribozymes, such as 3′ terminal nucleotide inversions (Sioud et al., 2000) and modifications to nucleic acids within the binding arms (Vester et al., 2002). To date, several 10-23 variants have been isolated to catalyze cleavage of biologically relevant RNAs (see Table 1.2, updated from The Aptamer Handbook (Klussmann, 2006), Chapter 10 by Schlosser, McManus and Li). These deoxyribozymes can target viral RNAs, bacterial RNAs, and endogenous mRNAs involved in disease processes. Viral RNAs that have been targeted include many HIV protein coding RNA regions. One study has created deoxyribozymes that target a stem-loop of the TAR region of the HIV-1 genome (Chakraborti et al., 2003). TAT and other cellular proteins bind to the TAR region and regulate transcription of viral genes, making TAR an attractive target as cleavage at this site could down-regulate expression of all viral genes. *In vitro* tests against RNA transcripts followed by tests using cultured cells co-transfected with HIV-1 and deoxyribozyme showed that a 10-23 variant was capable of a greater than 80% reduction of HIV-1 replication.

Looking at examples of cleavage of endogenous mRNAs, a deoxyribozyme was designed to inhibit angiogenesis through cleavage of the mRNA translated into vascular endothelial growth factor receptor 2 (VEGFR2) (Zhang et al., 2002). Angiogenesis is the process of creation of new blood vessels and is critical for tumour growth. VEGF receptor 2 is usually expressed in embryonic cells and down-regulated in adult cells. During induction of angiogenesis, VEGFR2 is up-regulated. A 10-23 deoxyribozyme

targeting the mRNA of VEGFR2 was designed with the cleavage site between the U and G residues of the AUG initiation codon. When its activity was tested *in vitro,* the deoxyribozyme capable of cleaving an RNA oligonucleotide and an *in vitro* RNA transcript, with the cleavage fragments corresponding in size to cleavage at the UG site of the initiation codon. When tested in BAEC cells, treatment with the deoxyribozyme produced a 90% reduction in VEFGR2 mRNA production 24 hours after treatment. Treatment of deoxyribozyme in null mice with pre-established human MDA-MB-435 breast carcinomas resulted in a 75% reduction in tumour size, compared to a 35% reduction when treated with control DNA. Conformation of inhibition of angiogenesis is shown by reduction of blood vessel density in deoxyribozyme treated mice tumours in comparison with untreated mice. Another deoxyribozyme targeting the c-Jun mRNA displayed similar angiogenic inhibition properties (Zhang et al., 2004). This particular study, showed that this deoxyribozyme targeting c-Jun mRNA was capable of reducing the size of certain tumours and that c-Jun has a pro-angiogenic role, which was not previously characterized. Another deoxyribozyme has been found that aids in ganglia axon regeneration at lesions in the spinal cord (Grimpe et al., 2004). This is achieved by cleavage of an mRNA for glycosaminoglycan (GAG), an inhibitory proteoglycan. These examples demonstrate the wide range of RNA targets for which 10-23 have been developed and show the potential of these deoxyribozymes to possibly treat a variety of different diseases and affliction.

Table 1.2 deoxyribozymes targeting mRNAs.

Viral and Bacterial RNAs	**Results/effects of RNA cleavage**	**Reference(s)**
HIV-1 HXBR	Inhibition of HIV-1 replication by cleavage of envelope protein mRNA	(Zhang et al., 1999)
HIV-1 TAT/Rev	Inhibition of mRNAs for regulatory proteins, resulted in inhibition of HIV-1 replication	(Unwalla et al., 2001)
HIV-1 gag	Inhibition of viral gene expression/replication by cleavage of *gag* mRNA	(Basu et al., 2000)
HIV-1 env	Cleavage of HIV-1 *env* RNA and inhibition of HIV-1 envelope-CD4 mediated cell fusion	(Dash et al., 1998)
HIV-1 TAR	Inhibition of viral gene expression by cleavage of HIV-1 TAR stem-loop RNA	(Chakraborti et al., 2003)
Influenza A PB2 mRNA	Inhibition of viral replication by cleavage of the PB2 mRNA at translation initiation site	(Takahashi et al., 2004)
RSV nucleocapsid protein	Inhibition of viral replication by cleavage of conserved genomic RNA of RSV nucleocapsid protein	(Zhou et al., 2007)
EBV latent membrane protein	Inhibition of cellular signalling transduction pathways by cleavage of latent membrane protein	(Lu et al., 2008; Lu et al., 2005)
Bacterial ftsZ mRNA	Inhibition of bacterial cell proliferation	(Tan et al., 2004)
Eukaryotic mRNAs		
Protein kinase Cα	Induction of apoptosis	(Sioud et al., 2000)
TWIST protein	Induction of apoptosis	(Hjiantoniou et al., 2003)
Human Platelet-Type 12-Lipoxygenase mRNA	Inhibition of tumour angiogenesis	(Liu et al., 2001)
LG-BCR-ABL fusion	Cleavage at fusion site of CML specific	(Warashina et

mRNAs	hybrid mRNA	al., 1999)
Integrins	Inhibition of endothelial cell capillary tube formation	(Cieslak et al., 2002)
c-myc	Suppression of smooth muscle cell proliferation	(Sun et al., 1999)
Huntington protein	Suppression of HD protein expression	(Yen et al., 1999)
Vanilloid receptor subtype 1 mRNA	Cleavage of vanilloid receptor subtype 1 mRNA	(Kurreck et al., 2002)
Laminins	Cleavage resulting in inhibition of mossy fibre axon regeneration	(Grimpe et al., 2002)
Early growth response factor 1 mRNA	Inhibition of vascular smooth muscle cell proliferation, inhibition of in-stent restenosis, inhibition of neointima formation, and inhibition of breast carcinoma proliferation, migration, angiogenesis and solid tumour growth.	(Fahmy et al., 2003; Mitchell et al., 2004; Santiago et al., 1999; Lowe et al., 2001; Lowe et al., 2002)
TNF-α mRNA	Improve hemodynamic performance	(Iversen et al., 2001)
C-Raf kinase	Reduction of C-raf mRNA by up to 36% and inhibition of apoptosis	(Chen et al., 2000)
VEFG receptor 2	Inhibition of angiogenesis	(Zhang et al., 2002)
c-JUN	Inhibition of angiogenesis, inhibition of vascular smooth muscle cell proliferation, induction of caspase-2 activation	(Khachigian et al., 2002; Fahmy et al., 2006; Dass et al., 2010)
Raf-1	Inhibition of myelomoncytic leukemia	(Iversen et al., 2002)
xylosyltransferase-1	Inhibition of glycosylation of proteoglycans	(Grimpe et al., 2004; Hurtado et al., 2008)
GATA-3	Inhibition of airway inflammation	(Sel et al., 2008)

1.7 Deoxyribozymes as Sensors

1.7.1 Deoxyribozyme-Fluorophore Conjugates

With the isolation of many deoxyribozymes that function in the presence of different co-factors (Table 1.1), efforts have been made to convert some of these catalysts into sensors. In particular RNA-cleaving deoxyribozymes have been exploited in many sensor designs in which the cleavage of the substrate into two smaller oligonucleotides is detected. Fluorescence detection platforms are especially popular, as they offer excellent sensitivity and many different fluorescent dyes are available for DNA modification (Cox et al., 2004 ; van Gijlswijk et al., 2002). Several methods exist to label DNA with fluorophores, including enzymatic addition of fluorophore-labeled dNTPs (Awonusonu et al., 1999) or chemical addition during solid-phase synthesis of oligonucleotides using dye-derived phosphoramidites. Doubly labeled nucleic acids capable of Forster resonance energy transfer (FRET) (Förster, 1948) have been widely used to detect other nucleic acids (Tyagi et al., 1996) as well as other biomolecules (reviewed in (Nutiu et al., 2005)). The most common utilization of fluorescence in deoxyribozyme-based sensing is done via close contact quenching mechanisms between fluorophore donors and quencher acceptors. The key feature of this type of sensor is that the fluorescence signal is quenched before substrate cleavage, with subsequent deoxyribozyme-mediated cleavage resulting in a loss of quenching and a high level of fluorescence. This strategy requires the fluorophore and quencher to be within a few nucleotides of each other (Chiuman et al., 2007). Labeling of the substrate with a fluorophore and quencher near the cleavage site, however, can lead to a loss of deoxyribozyme activity as the bulky fluorophore and

quencher may disrupt the interaction of the deoxyribozyme with the substrate. To prevent labeling from causing a loss of activity, deoxyribozymes and substrates have been labeled at sites distal to the cleavage site, with deoxyribozyme-substrate separation upon cleavage used to create an increase in fluorescence intensity. The general scheme for this sensor design is shown in Figure 1.5a and involves placing a fluorophore and quencher on the 5′ end of the substrate and the 3′ end of the deoxyribozyme in an extended stem region.

1.7.2 Fluorescent sensors using existing deoxyribozymes

The first sensor developed using this scheme was carried out by the Lu group using an 8-17 deoxyribozyme variant that can cleave an RNA substrate in the presence of Pb^{2+} with an observed rate of 6.5 min^{-1} (Li, J. et al., 2000). To convert this deoxyribozyme into a fluorescent sensor, a TAMRA fluorophore (6-carboxytetramethylrhodamine) was attached to the 5′ end of the substrate and a dabcyl quencher (4-[4′-(dimethylamino)phenylazo] benzoic acid) on the 3′ end of the deoxyribozyme (Li, Y. et al., 2000). In the absence of Pb^{2+}, the substrate is bound to the deoxyribozyme, placing the fluorophore in proximity to the quencher and quenching fluorescence. In the presence of Pb^{2+}, the substrate is cleaved and thus separated from the deoxyribozyme resulting in loss of quenching and an observed enhancement of fluorescence. This sensor was found to be able to detect Pb^{2+} over a concentration range of 10 nM to 4 μM with an 80-fold specificity for Pb^{2+} over other metal ions. This sensor was improved in a follow-up study where a second quencher moiety was used to reduce background fluorescence, resulting in a five-fold improvement in signalling compared

with the original design (Liu, J. et al., 2003b). The sensitivity of this sensor was further enhanced by the Tan group by connecting the substrate and deoxyribozyme with a poly-thymine linker (Wang et al., 2009), as shown in Figure 1.5b. This results in increased hybridization between deoxyribozyme and substrate, which enhances the catalytic efficiency and reduces background fluorescence. This modified deoxyribozyme sensor exhibited a much improved detection limit of 3 nM. In a recent effort to create a fluorescent deoxyribozyme sensor for Pb^{2+} with high specificity, the Lu group created a similarly designed sensor using the first deoxyribozyme ever isolated (Breaker et al., 1994), which was shown to be 40,000 times more specific than the previously designed deoxyribozyme-based Pb^{2+} sensors (Lan et al., 2010).

Sensors for other metal ions have also been developed using other deoxyribozymes as the recognition elements. One sensor for uranium ions developed by the Lu group utilized *in vitro* selected deoxyribozymes specific for UO_2^{2+}, which were modified with a quencher and a substrate containing a second quencher and a fluorophore (Liu et al., 2007a). This sensor has a detection limit of 45 pM and a selectivity of over one million-fold over other metal ions, making it one of the most sensitive uranium sensors available. Using the same deoxyribozyme, the Lu group applied a rational design approach to convert their uranium sensor into a sensor for Hg^{2+} (Liu et al., 2007a). It has been shown that Hg^{2+} can bind pairs of thymines creating T-T base pairs. To make the deoxyribozyme responsive to Hg^{2+}, T-T mismatches were first added into a stem of the deoxyribozyme to deactivate the deoxyribozyme as shown in Figure 1.5c. When Hg^{2+} is present, the activity of the deoxyribozyme is restored producing an increase of

fluorescence. Using this system, it is possible to detect 4 nM of Hg^{2+}, which is below the Environmental Protection Agency limit for toxicity. Fluorescent deoxyribozyme sensors have also been developed using other DNA catalyzed reactions. A fluorescent sensor for Cu^{2+} was developed using a copper-dependent DNA-cleaving deoxyribozyme (Carmi et al., 1996) with a fluorophore and quencher placed at the ends of an extended helical region (Liu et al., 2007c). The sensor was found to be specific for Cu^{2+} with a detection limit of 35 nM.

In addition to dyes that quench fluorescence, other types of quenching have been employed to simplify sensor design. Graphene is a two-dimensional nanomaterial consisting of a single carbon-atom layer that is homogenously arranged into a honeycomb crystal lattice. An interesting derivative of graphene, graphene oxide, has been used in a variety of biological applications due to its unique surface and electronic properties (Lu et al., 2009 ; Zhou et al., 2009 ; Dong et al., 2010 ; He et al., 2010). In addition to being water-soluble and having quenching capacities, graphene oxide can interact with oligonucleotides through hydrophobic and π stacking interactions. Interestingly, due to the super quenching effects of graphene oxide, the fluorescence generating deoxyribozymes do not require a quenching dye partner. Selected biosensing applications that previously used fluorescence-generating deoxyribozymes have been converted to graphene oxide-based systems. Graphene immobilized fluorescence generating deoxyribozymes have been applied to detect metal ions such as Cu^{2+} and Pb^{2+} with detection limits ranging between 0.3 to 2.0 nM (Zhao et al., 2011 ; Wen et al., 2011 ; Liu, M. et al., 2011).

Fluorescent deoxyribozyme sensors that do not require substrate and deoxyribozyme labeling have also been developed. In one system, an abasic site is introduced into the deoxyribozyme's substrate binding region (Xiang et al., 2009). As shown in Figure 1.5d, when the deoxyribozyme is bound to the substrate, the abasic site is capable of binding 2-amino-5,6,7-trimethyl-1,8-naphthyridine (ATMND), which is non-fluorescent when bound. Upon substrate cleavage, the deoxyribozyme-substrate complex is disassembled releasing the ATMND, which becomes fluorescent in its unbound state. When this system is used with a Pb^{2+}-utilizing deoxyribozyme, it is capable of detecting as little as 4 nM of Pb^{2+}, comparable with its labeled deoxyribozyme sensor counterparts. This method has been generalized and used to detect UO_2^{2+} and Hg^{2+}, using relevant deoxyribozymes mentioned above, with detection limits of 3 nM and 30 nM being achieved, respectively (Xiang et al., 2010).

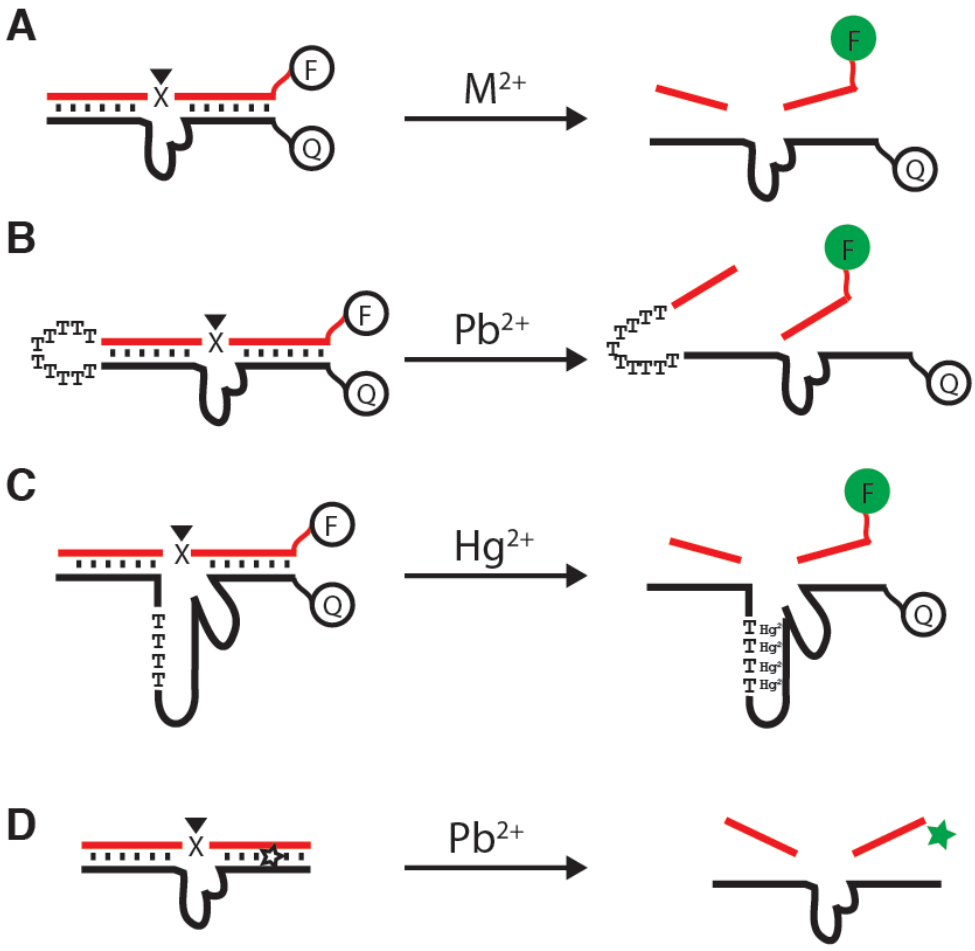


Figure 1.5 Design of fluorescent-based sensors based on RNA-cleaving deoxyribozymes. (A) Sensors with labeled substrates and deoxyribozymes. Substrate and deoxyribozyme are modified with a fluorophore (F) and quencher (Q), respectively. In the presence of the correct metal-ion cofactor, the RNA substrate is cleaved, releasing the substrate and leading to an increase in fluorescence. (B) Unimolecular Pb^{2+} detection. By linking the DNAzyme and substrate with a polythymine linker, hybridization is increased and background fluorescence is reduced. (C) A DNAzyme sensor for Hg^{2+}. By placing thymine-thymine mismatches in an essential stem, an RNA-cleaving DNAzyme is deactivated. The binding of Hg^{2+} to thymines causes thymine-thymine base pairs to form, activating the DNAzyme. (D) Label-free fluorescent sensors. ATMND (star) binds to an abasic site in the DNAzyme-substrate complex, which quenches its fluorescence. Upon Pb^{2+}-dependent substrate cleavage, ATMND is released and it fluorescence is increased.

1.7.3 Isolation of de novo fluorescent deoxyribozyme sensors through in vitro selection

One drawback of the aforementioned systems is that in order to convert the RNA-cleaving deoxyribozymes into fluorescent sensors it is necessary to modify the substrate and deoxyribozymes with fluorescent dyes. As the residues near the cleavage site cannot usually be modified without a loss of enzymatic activity, the fluorophore and quencher dyes typically need to be placed distal from the active site. This can lead to issues such as loss of sensitivity and high background fluorescence. To prevent these problems, a selection for deoxyribozymes that can cleave substrates containing an RNA linkage flanked by a fluorophore and quencher have been performed (Mei et al., 2003). The substrate is designed with the fluorophore and quencher separated by only one residue resulting in strong quenching in the uncleaved substrate and low background fluorescence. Upon cleavage, two portions containing the fluorophore and quencher are released and separated, resulting in an increase in fluorescence. After twenty-two rounds of selection, a deoxyribozyme was obtained that cleaves the fluorogenic RNA substrate at a rate of 7 min^{-1}, placing it amongst the fastest deoxyribozymes known to date. This demonstrated that placement of a fluorophore and quencher near the active site did not hinder the selection of efficient deoxyribozymes. This deoxyribozyme was found to have strong activity in the presence of $Co^{2+,}$ Ni^{2+}, and Mn^{2+}, suggesting that this type of sensor system could be used in detection of transition metals. A subsequent selection experiment was performed to isolate fluorescence-generating deoxyribozymes that could cleave an RNA linkage at various ranges of pH (Liu, Z. et al., 2003 ; Shen et al., 2005 ;

Kandadai et al., 2009 ; Kandadai et al., 2005 ; Ali et al., 2007b). By performing parallel selections in buffers with different acidity settings, deoxyribozymes that cleave fluorogenic substrates at pH of 3, 4, 5, 6 and 7 were isolated. Deoxyribozymes from each pH class were found to perform efficient catalysis with reaction rates around 1 min^{-1}.

1.7.4 Fluorescence-generating Aptazymes: Linking Target Recognition to Catalytic Activity

To use deoxyribozymes to sense molecules larger than metal ions, it is necessary for sensors to have recognition elements in addition to catalytic motifs. *In vitro* selection has been very effective in isolating DNA and RNA aptamers that can recognize and bind a target molecule with high sensitivity and specificity. There are many natural and artificially made sensors that function by binding induced structural changes. By designing constructs and assessing the equilibrium constants between conformations that can and cannot bind targets, it is possible to engineer structure-switching sensors with low detection limits and large dynamic range (Vallee-Belisle et al., 2009). Attempts have been made to examine whether the catalytic signalling capacity of deoxyribozymes could be coupled with the specific binding of DNA aptamers to generate catalytic sensors that are responsive to target binding. Early demonstrations of this type of system were made by the Breaker group using RNA aptamers appended to ribozymes. By rational design, aptamer-linked ribozymes or 'RNA aptazymes' were created in which ATP binding to an ATP aptamer domain altered the folding of an attached hammerhead ribozyme domain, causing it to adopt an inactive conformation and inhibiting the catalytic activity of the ribozyme (Tang et al., 1997 ; Tang et al., 1998). Similar to In another study, the Breaker

group was able to isolate FMN-responsive aptazymes with 270-fold rate enhancements by performing an in vitro selection with a randomized linker region (Soukup et al., 1999). The first DNA aptazyme was designed by the Sen group using a well-studied anti-ATP DNA aptamer (Huizenga et al., 1995) as a recognition element and the 10-23 deoxyribozyme as the catalytic reporter (Wang, D. Y. et al., 2002). As shown in Figure 1.6a, their design used an aptazyme construct with an altered substrate-binding domain and an ATP aptamer domain flanked by a short region complementary to the substrate. In the absence of ATP, the rate of catalysis is reduced, as the deoxyribozyme cannot bind the substrate, as the two shortened complementary regions on the aptazyme are far apart in the unstructured ATP-binding region. When ATP is added, it binds to the aptamer domain, rearranging the aptazyme structure, permitting additional base pairing with the substrate. To show the versatility of this design, a similar system was constructed using the 8-17 deoxyribozyme as a reporter. The second DNA aptazyme sensor was developed by the Ellington group, which utilized the same anti-ATP DNA aptamer and a previously isolated DNA ligase (Levy et al., 2002a), as shown in Figure 1.6b. Binding of ATP to the recognition domain was found to increase the rate of ligase activity by up to 460-fold. The conformation of the aptazyme in the absence of ATP was also found to reduce background ligation levels below that of the template reactions. Another approach, reported by our group and modeled after a structure-switching aptamer design (Nutiu et al., 2003), used the same ATP aptamer sequence linked to an RNA-cleaving deoxyribozyme (Liu, Z. et al., 2003) and a regulatory oligonucleotide that binds part of the substrate-binding region of the deoxyribozyme and part of the aptamer domain

(Achenbach et al., 2005b), as shown Figure 1.6c. In the absence of ATP, the regulatory oligonucleotide blocks substrate binding to the deoxyribozyme portion. When ATP is added, binding to the aptamer domain disrupts the aptamer's interaction with the regulatory oligonucleotide, resulting in its release. The deoxyribozyme portion is now free to bind and cleave its RNA substrate. This system was shown to achieve a 30-fold rate enhancement in the presence of ATP. A unimolecular structure-switching mechanism was also developed using a fluorescence generating RNA-cleaving deoxyribozyme functional at pH 6 and the ATP aptamer (Shen et al., 2006). The sensor was designed by sequestering catalytically essential nucleotides in a based-paired stem within the ATP aptamer domain. Upon addition of ATP, the aptamer domain switches from a helical interaction with the deoxyribozyme to form a complex with ATP. The deoxyribozyme can now cleave its RNA substrate, producing a fluorescence increase.

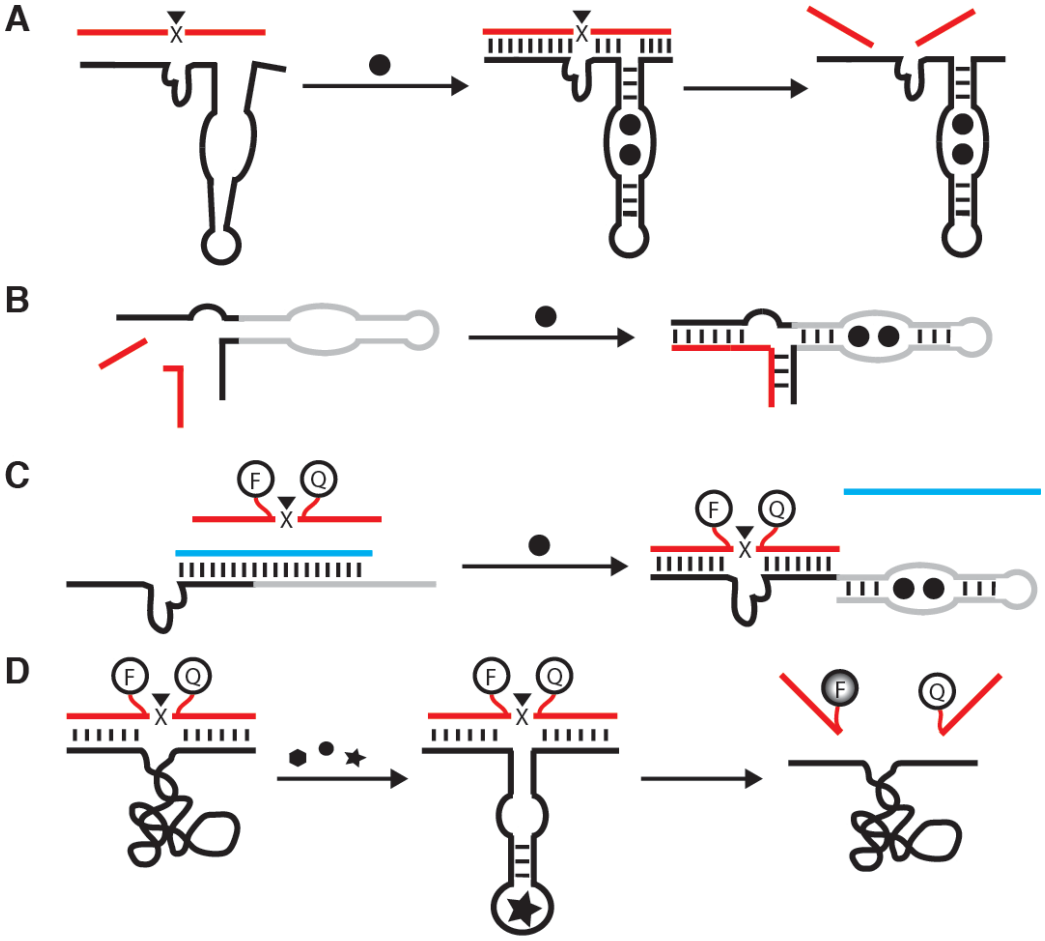


Figure 1.6 Aptazyme-based Sensors. A) A DNA aptazyme constructed using an ATP aptamer and the 8-17 DNAzyme. In the absence of ATP (black circle), the aptazyme cannot interact with the substrate (red) due to a lack of base pairing. ATP binding alters the structure of the aptazyme, allowing substrate binding and RNA cleavage. (B) Aptazyme containing an ATP aptamer (grey) and a ligase DNAzyme (black). In the absence of ATP (black circle), the aptazyme forms a conformation that does not allow it to bind its substrates shown in red. ATP binding changes the conformation, allowing substrate binding and ligation. (C) An aptazyme using a structure-switching mechanism. In the absence of ATP, a regulatory DNA (blue) binds to the aptazyme, blocking binding of a fluorogenic substrate (red). Upon ATP (black circle) binding to the aptamer domain, the aptazyme switches from a duplex with the regulatory oligonucleotide to a complex with ATP. This relieves the DNAzyme to bind its substrate and cleave its fluorogenic substrate. (D) Aptazyme sensor for *E. coli* crude extracellular mixture (CEM, represented by three different shapes). In the presence of *E. coli* CEM, the aptazyme can bind its target and cleave a fluorogenic substrate (red).

Another method to develop aptazymes, as opposed to rational design is to use in vitro selection to directly isolate sequences that have catalytic activity that is responsive to a target of interest. This technique has potential for designing sensors for targets with no existing aptamers. Target dependence can be obtained by adding negative selection steps to the in vitro selection process in order to remove sequences from the selection pool that are catalytic in the absence of target. By monitoring the level of activity in the absence of target, the stringency of the negative selection can be increased if needed, resulting in an aptazyme with very low background fluorescence. In one study, a fluorescent deoxyribozyme sensor was developed for the detection of bacteria (Ali et al., 2011). As shown in Figure 1.6d, a DNA library containing a 70-nt random region attached to a DNA substrate with a fluorophore and quencher flanking a ribonucleotide was used to select for a deoxyribozyme responsive to crude extracellular mixture (CEM) of *Escherichia coli*. Specificity was attained by performing counter-selection using CEM from *Bacillus subtilis* and discarding the population that could cleave the substrate under these conditions. After twenty rounds of selection, a sequence was isolated that could cleave the substrate and generate a fluorescence signal in the presence of *E. coli* CEM. The sensor was shown to be specific, as cleavage and fluorescence were not seen in the presence of several other Gram-negative and Gram-positive bacteria. With the addition of a culturing step, the sensor was shown to be able to detect a single seeding *E. coli* cell. It is also noteworthy that this *E. coli* sensor was developed using a complex mixture, the CEM, and was not directed for a specific target molecule. This allowed the in vitro selection procedure to choose the best target from the mixture. This avoids the laborious

steps of finding a suitable target from *E. coli* (or another specific bacterium of interest), but not other bacteria, and purifying the target while maintaining its native structure. Also, there is no guarantee that the purified targets will be suitable ligands in an aptamer selection. The direct aptazyme in vitro selection process on the other hand will intrinsically target molecules for which it has a natural affinity.

1.7.5 Colorimetric Deoxyribozyme-based Sensors

Colorimetric sensors offer advantages over many other detection systems as they can report the presence of a target without the need for expensive and bulky instrumentation. Gold nanoparticles in particular have gained much attention and been used extensively for detection as they produce an intense red color with higher extinction coefficients than most organic dyes. A change in the surface plasmon resonance upon aggregation of the nanoparticles results in a change in color, which can be coupled to a target-responsive event to allow for sensing. Following the reports of using DNA oligonucleotide conjugated gold nanoparticles for DNA detection by the Mirkin group (Elghanian et al., 1997), they have been used in an array of DNA sensing schemes (Taton et al., 2000 ; Sato et al., 2003). By coupling oligonucleotide aptamers to gold nanoparticle aggregation, sensors have been made for proteins (Wei et al., 2007) and small molecules (Liu et al., 2005a ; Liu et al., 2006 ; Wang et al., 2006 ; Zhao et al., 2007 ; Zhao et al., 2008a) (reviewed in (Liu et al., 2004c)). In a series of publications, the Lu group has used gold nanoparticle aggregation as a detection system for deoxyribozyme-based sensors. In one study, a sensor was created by designing a deoxyribozyme substrate that contains identical 5′ and 3′ regions that hybridizes to oligonucleotides covalently-bound to gold

nanoparticles (Liu, J. et al., 2003a). When the system is heated above the melting temperature of the complementary regions and cooled, the substrate forms a bridge between nanoparticles, leading to nanoparticle aggregation. Under these conditions the sensing solution appears purple. When substrate and Pb^{2+}-dependent deoxyribozyme are incubated with Pb^{2+} before melting and cooling, the substrate is cleaved, preventing nanoparticle aggregation. Consequently, a red colour is seen. This sensor was found to be able to detect lead ions with a dynamic range between 100 nM and 4 μM. By using a mutant deoxyribozyme that can bind the substrate without causing cleavage, it is possible to tune the range of the sensor by altering the ratio of active and mutant deoxyribozyme. In a follow-up study, the sensor was improved by changing the orientation of the nanoparticle-substrate complex and the size of the nanoparticles (Liu et al., 2004a). This allows the sensor to function at room temperature and removes the melting and cooling steps. These improvements reduced the time of detection from 2 hours to 10 minutes. This system has also been modified with nanoparticle disassembly, instead of assembly, being monitored in a 'light up' format (Liu et al., 2005b). Another system for the detection of lead was designed that utilized properties of gold nanoparticles that cause the particles to aggregate in high concentration of salt. Substrate-labeled nanoparticles are incubated with a Pb^{2+}-dependent deoxyribozyme and the maximum amount of salt at which the nanoparticles are still stable and not aggregated (Zhao et al., 2008b). Upon the addition of Pb^{2+}, the substrate is cleaved, and a portion is released from the nanoparticles. These shorter oligonucleotides on the nanoparticles are not sufficient to stabilize the nanoparticles, resulting in aggregation and a color change from red to purple. Aptazyme-

based colorimetric sensors using gold nanoparticles have also been developed (Liu et al., 2004b). The aptazyme is divided into two oligonucleotides, one with an 8-17 catalytic domain and one arm of an ATP aptamer domain, and a second oligonucleotide containing the second half of the aptamer domain. Like the previous studies, the system contains a substrate capable of bridging oligonucleotide-modified nanoparticles leading to nanoparticle aggregation and a purple colour. In the presence of adenosine, the aptazyme is activated, cleaving the substrate, and preventing gold particle aggregation. Gold nanoparticle aggregation has also been used as a sensing system with a DNA ligase deoxyribozyme. A sensor for Cu^{2+} was designed using a copper-dependent deoxyribozyme and two substrates capable of hybridizing to oligonucleotide modified gold nanoparticles (Liu et al., 2007b). In the presence of Cu^{2+}, the deoxyribozyme is active and ligates the two substrates. This ligated substrate is then capable of bridging two gold nanoparticles, leading to aggregation and a red to purple color change.

1.7.6 Electrochemical-based Sensors

Surface-based detection has gained popularity in recent years. Electrochemical sensors in particular offer several advantages over solution-based systems. Electroactive labels used for detection are typically more insensitive to environmental changes than fluorescent or colorimetric dyes, and background signal from contaminants is far less common than that seen with optical techniques. Due to these factors, steps have been made to utilize the molecular recognition capacities of deoxyribozymes to develop sensitive electrochemical sensors. The first study was carried out by the Plaxco group to develop a Pb^{2+} sensor using a methylene blue-modified, lead-sensitive deoxyribozyme

attached to a gold electrode (Xiao et al., 2007). When bound to its substrate, the double-helical complex separates the methylene blue from the electrode. In the presence of Pb^{2+}, the substrate is cleaved and dissociates from the deoxyribozyme. In this unbound form, the deoxyribozyme becomes flexible and the attached methylene blue can contact the electrode, resulting in electron transfer. The sensor was shown to be able to detect as little as 300 nM Pb^{2+}. No signal was observed using 10 μM of several other divalent metal ions, showing that the sensor is very specific for Pb^{2+}. Another sensor design utilizes an electrode-bound Pb^{2+}-dependent deoxyribozyme and $Ru(NH_3)_6^{3+}$, which binds to the negative phosphate groups of DNA, as the signal transducer (Shen et al., 2008). Upon the addition of Pb^{2+}, the substrate is cleaved, resulting in deoxyribozyme-substrate complex dissociation and a reduction of amount of $Ru(NH_3)_6^{3+}$ confined to the electrode surface. This system was found to be able to detect 5 nM of Pb^{2+}. By coupling the substrate to oligonucleotide-modified gold nanoparticles, the sensitivity of the sensor was increased to 1 nM Pb^{2+}. This signal amplification is accomplished by $Ru(NH_3)_6^{3+}$ being bound to multiple oligonucleotides on the nanoparticle. Another sensor also used gold nanoparticle loaded oligonucleotides to amplify an electrochemical signal. In this sensor design, gold nanoparticles are modified with Pb^{2+}-dependent deoxyribozymes which can hybridize to substrate DNA/RNAs (Yang et al., 2010). In the presence of Pb^{2+}, the substrate is cleaved and released from the deoxyribozyme. The solution is then added to an electrode modified with single-stranded DNA complementary to the deoxyribozyme sequence. The deoxyribozyme can hybridize to the electrode-bound DNA, bringing the gold nanoparticle labeled with many oligonucleotides to the electrode resulting in a large

increase in electrochemical signalling in the presence of $Ru(NH_3)_6^{3+}$. This sensor has been found to detect as little as 28 pM of Pb^{2+}. Another sensor with subnanomolar detection capacity has been developed by combining the benefits of electrochemical and chemiluminescence detection. The sensor utilizes electrochemiluminescence, or the detection of light emitted from the excited state of an electrochemical signalling compound, to detect Pb^{2+} (Zhu et al., 2009). This sensor contains an electrode-bound deoxyribozyme and substrate labeled with tris(2,2′-bipyridine)-ruthenium(II) (TBR). Upon Pb^{2+} induced substrate cleavage, the substrate-bound TBR is separated from the electrode, leading to a reduction in electrochemiluminescence intensity. Using this system has allowed for the detection of Pb^{2+} at a concentration of 100 pM.

Another example of the potential of deoxyribozyme-based electrochemical sensor was shown in a study that coupled deoxyribozyme activity to the personal glucose meter. The personal glucose meter is an electrochemical sensor that detects glucose by monitoring the redox reaction catalyzed by glucose oxidase. This platform has been optimized over the last thirty years into very efficient pocket-sized sensors. To utilize this well-established system to detect other molecules, a beta-fructofuranosidase, known as invertase, was used along with sucrose. The invertase was conjugated to an oligonucleotide using the maleimide-thiol reaction and hybridized to a complex of a UO_2^{2+}-dependent RNA-cleaving deoxyribozyme and its substrate bound to a magnetic bead (Xiang et al., 2011). In the presence of UO_2^{2+}, the substrate is cleaved, releasing the DNA-invertase conjugate. Once released, the invertase can hydrolyze sucrose to glucose, which can be oxidized by glucose oxidase and detected by the meter with a detection

limit of 9 nM UO_2^{2+}. This system was also shown to be versatile and was able to selectively detect cocaine, adenosine, and interferon-gamma using structure-switching aptamers for each of these targets. Some of these sensors have since been commercialized and demonstrate that deoxyribozymes can be converted into real world sensors using creativity to couple deoxyribozyme activity to practical detection systems.

1.8 Kinase Deoxyribozymes: A Tool for the Study of DNA Structure, in vitro Selection and a Platform for Development of Novel Biosensors

Our lab uses a class of deoxyribozymes as a tool for the study different aspects of DNA based catalysis. These deoxyribozymes, which catalyze the transfer of a γ-phosphate from a nucleoside 5′-triphosphate (such as ATP) to the 5′ end of a DNA substrate, and the related deoxyribozymes are called kinase deoxyribozymes (or self-phosphorylating deoxyribozymes). A group of kinase deoxyribozymes were first isolated by Li and Breaker in a study to investigate deoxyribozyme substrate specificity (Li et al., 1999). Previous work in the lab has involved parallel in vitro selection experiments for the isolation of self-phosphorylating deoxyribozymes under different divalent metal ion conditions to test the effects of differing co-factors on DNA-based catalysis (Wang, W. et al., 2002).

We decided to use some of the kinase deoxyribozymes from this second study as a model system to examine structural diversity of functional DNA molecules. We chose five unique sequences that were observed multiple times in the pool of sequences retrieved from the sixteenth round of a sub-selection experiment performed under stringent reaction conditions. In this selection, reaction time was progressively decreased

in order to obtain efficient catalysts with fast reaction rates. These sequences were named Deoxyribozyme kinase 1 to 5 (Dk1 to Dk5) in order of abundance. The secondary structures of the two most abundant sequences Dk1, which is ATP-dependent, and Dk2, which is GTP-dependent, were elucidated and both were showed to contain a common secondary structure (Achenbach et al., 2005a). Both deoxyribozymes appear to have a central stem-loop motif that acts to anchor two arms that are responsible for substrate binding and catalysis. This stem-loop motif was shown to be purely structural and not involved in catalysis, as its base-pair sequence could be altered without loss of activity. In fact, it was found that the stem-loops from each deoxyribozyme could be interchanged between the two deoxyribozymes without change in function. Further characterization revealed that Dk2 has other base-pair interactions in its active structure, while Dk1 does not. The methylation interference patterns for the two deoxyribozymes were also different, indicating they use differing tertiary interactions in order to fold into their respective active structures.

1.9 Specific Goals of this Study

The projects carried out in this thesis use self-phosphorylating deoxyribozymes as a model system for deoxyribozyme structural characterization, improvement of in vitro selection, and development of novel deoxyribozyme-based sensor systems. The body of this thesis is divided into four chapters, detailing the four major projects carried out during my doctoral studies.

1. To determine whether there are preferred structural motifs for self-phosphorylation under specific conditions.

- My initial characterization of kinase deoxyribozymes Dk3 and Dk4 showed that they share common co-factor and substrate requirements as Dk2, yet all three deoxyribozymes have distinct primary sequences are arose independently during their selection. This study demonstrated that the three deoxyribozymes share a common secondary and tertiary structure suggesting a preferred structural motif for self-phosphorylation under certain conditions.

2. To investigate whether novel DNA structures can be found through the structural study of deoxyribozymes.
   - The characterization of Dk5 revealed that it contained a two-tiered quadruplex entwined with a Watson-Crick base-paired stem in a pseudoknot arrangement never seen before in natural or artificially selected DNA molecules.
3. To apply the knowledge obtained through our deoxyribozyme structural study to improve the in vitro selection process.
   - DNA libraries were designed to be enriched for two-tiered guanine quadruplexes, a motif shown to be present in several DNA aptamers and deoxyribozymes including the aforementioned Dk5. Characterization of several libraries revealed that some contain a mixture of quadruplexes of differing topologies, ideal for an in vitro selection experiment.
4. Convert self-phosphorylating deoxyribozymes into biosensors.
   - The self-phosphorylation activity of Dk2 was coupled to circular ligation and rolling circle amplification to increase the sensitivity of detection of

self-phosphorylation by 1000-fold over the previous ligation detection method. This amplification was linked to fluorogenic activity of an appended RNA-cleaving deoxyribozyme. This system allowed for the fluorescent detection of the GTP substrate and Mn^{2+} co-factor of Dk2.

1.10 References

Aboul-ela, F., A. I. Murchie, and D. M. Lilley (1992). NMR study of parallel-stranded tetraplex formation by the hexadeoxynucleotide d(TG4T). Nature 360: 280-282.

Achenbach, J. C., G. A. Jeffries, S. A. McManus, L. P. Billen, and Y. Li (2005a). Secondary-structure characterization of two proficient kinase deoxyribozymes. Biochemistry 44: 3765-3774.

Achenbach, J. C., R. Nutiu, and Y. Li (2005b). Structure-switching allosteric deoxyribozymes. Analytica Chimica Acta 534: 41-51.

Ali, M. M., S. D. Aguirre, H. Lazim, and Y. Li (2011). Fluorogenic DNAzyme probes as bacterial indicators. Angew Chem Int Ed Engl 50: 3751-3754.

Ali, M. M., S. A. Kandadai, and Y. Li (2007a). Characterization of pH3DZ1 - An RNA-cleaving deoxyribozyme with optimal activity at pH 3. Canadian Journal of Chemistry 85: 261-273.

Ali, M. M., S. A. Kandadai, and Y. Li (2007b). An RNA-cleaving deoxyribozyme with optimal activity at pH 3. Canadian Journal of Chemistry 85: 261-273.

Arnott, S., R. Chandrasekaran, and C. M. Marttila (1974). Structures for polyinosinic acid and polyguanylic acid. Biochem J 141: 537-543.

Awonusonu, F., S. Srinivasan, J. Strange, W. Al-Jumaily, and M. C. Bruce (1999). Developmental shift in the relative percentages of lung fibroblast subsets: role of apoptosis postseptation. Am J Physiol 277: L848-859.

Axelrod, V. D., E. Brown, C. Priano, and D. R. Mills (1991). Coliphage Q beta RNA replication: RNA catalytic for single-strand release. Virology 184: 595-608.

Basu, S., B. Sriram, R. Goila, and A. C. Banerjea (2000). Targeted cleavage of HIV-1 coreceptor-CXCR-4 by RNA-cleaving DNA-enzyme: inhibition of coreceptor function. Antiviral Res 46: 125-134.

Beal, P. A., and P. B. Dervan (1991). Second structural motif for recognition of DNA by oligonucleotide-directed triple-helix formation. Science 251: 1360-1363.

Beattie, T. L., J. E. Olive, and R. A. Collins (1995). A secondary-structure model for the self-cleaving region of Neurospora VS RNA. Proc Natl Acad Sci U S A 92: 4686-4690.

Beaudoin, J. D., and J. P. Perreault (2008). Potassium ions modulate a G-quadruplex-ribozyme's activity. RNA 14: 1018-1025.

Been, M. D. (1994). Cis- and trans-acting ribozymes from a human pathogen, hepatitis delta virus. Trends Biochem Sci 19: 251-256.

Bloomfield, V. A., D. M. Crothers, and I. Tinoco, (2000). *Nucleic Acids: Structures, Properties, and Functions*. University Science Books.

Breaker, R. R., and G. F. Joyce (1994). A DNA enzyme that cleaves RNA. Chem Biol 1: 223-229.

Breaker, R. R., and G. F. Joyce (1995). A DNA enzyme with Mg(2+)-dependent RNA phosphoesterase activity. Chem Biol 2: 655-660.

Carmi, N., S. R. Balkhi, and R. R. Breaker (1998). Cleaving DNA with DNA. Proc Natl Acad Sci U S A 95: 2233-2237.

Carmi, N., L. A. Shultz, and R. R. Breaker (1996). In vitro selection of self-cleaving DNAs. Chem Biol 3: 1039-1046.

Cech, T. R. (1993). The efficiency and versatility of catalytic RNA: implications for an RNA world. Gene 135: 33-36.

Chakraborti, S., and A. C. Banerjea (2003). Inhibition of HIV-1 gene expression by novel DNA enzymes targeted to cleave HIV-1 TAR RNA: potential effectiveness against all HIV-1 isolates. Mol Ther 7: 817-826.

Chandra, M., A. Sachdeva, and S. K. Silverman (2009). DNA-catalyzed sequence-specific hydrolysis of DNA. Nat Chem Biol 5: 718-720.

Chandra, M., and S. K. Silverman (2008). DNA and RNA can be equally efficient catalysts for carbon-carbon bond formation. J Am Chem Soc 130: 2936-2937.

Chapman, K. B., and J. W. Szostak (1994). In vitro selection of catalytic RNAs. Curr Opin Struct Biol 4: 618-622.

Chen, X., N. Li, and A. D. Ellington (2007). Ribozyme catalysis of metabolism in the RNA world. Chem Biodivers 4: 633-655.

Chen, Y., Y. J. Ji, R. Roxby, and C. Conrad (2000). In vivo expression of single-stranded DNA in mammalian cells with DNA enzyme sequences targeted to C-raf. Antisense Nucleic Acid Drug Dev 10: 415-422.

Chinnapen, D. J., and D. Sen (2004). A deoxyribozyme that harnesses light to repair thymine dimers in DNA. Proc Natl Acad Sci U S A 101: 65-69.

Chiuman, W., and Y. Li (2007). Efficient signaling platforms built from a small catalytic DNA and doubly labeled fluorogenic substrates. Nucleic Acids Res 35: 401-405.

Chiuman, W., and Y. Li (2006). Evolution of high-branching deoxyribozymes from a catalytic DNA with a three-way junction. Chem Biol 13: 1061-1069.

Cieslak, M., J. Niewiarowska, M. Nawrot, M. Koziolkiewicz, W. J. Stec, and C. S. Cierniewski (2002). DNAzymes to beta 1 and beta 3 mRNA down-regulate expression of the targeted integrins and inhibit endothelial cell capillary tube formation in fibrin and matrigel. J Biol Chem 277: 6779-6787.

Cox, W. G., and V. L. Singer (2004). Fluorescent DNA hybridization probe preparation using amine modification and reactive dye coupling. Biotechniques 36: 114-122.

Cruz, R. P., J. B. Withers, and Y. Li (2004). Dinucleotide junction cleavage versatility of 8-17 deoxyribozyme. Chem Biol 11: 57-67.

Cuenoud, B., and J. W. Szostak (1995). A DNA metalloenzyme with DNA ligase activity. Nature 375: 611-614.

Dash, B. C., T. A. Harikrishnan, R. Goila, S. Shahi, H. Unwalla, S. Husain, and A. C. Banerjea (1998). Targeted cleavage of HIV-1 envelope gene by a DNA enzyme and inhibition of HIV-1 envelope-CD4 mediated cell fusion. FEBS Lett 431: 395-399.

Dass, C. R., S. J. Galloway, and P. F. Choong (2010). Dz13, a c-jun DNAzyme, is a potent inducer of caspase-2 activation. Oligonucleotides 20: 137-146.

Dong, H., W. Gao, F. Yan, H. Ji, and H. Ju (2010). Fluorescence resonance energy transfer between quantum dots and graphene oxide for sensing biomolecules. Anal Chem 82: 5511-5517.

Elghanian, R., J. J. Storhoff, R. C. Mucic, R. L. Letsinger, and C. A. Mirkin (1997). Selective colorimetric detection of polynucleotides based on the distance-dependent optical properties of gold nanoparticles. Science 277: 1078-1081.

Ellington, A. D., and J. W. Szostak (1990). In vitro selection of RNA molecules that bind specific ligands. Nature 346: 818-822.

Ellington, A. D., and J. W. Szostak (1992). Selection in vitro of single-stranded DNA molecules that fold into specific ligand-binding structures. Nature 355: 850-852.

Fahmy, R. G., C. R. Dass, L. Q. Sun, C. N. Chesterman, and L. M. Khachigian (2003). Transcription factor Egr-1 supports FGF-dependent angiogenesis during neovascularization and tumor growth. Nat Med 9: 1026-1032.

Fahmy, R. G., A. Waldman, G. Zhang, A. Mitchell, N. Tedla, H. Cai, C. R. Geczy, C. N. Chesterman, M. Perry, and L. M. Khachigian (2006). Suppression of vascular permeability and inflammation by targeting of the transcription factor c-Jun. Nat Biotechnol 24: 856-863.

Faulhammer, D., and M. Famulok (1996). The Ca^{2+} ion as a cofactor for a novel RNA-cleaving deoxyribozyme. Angew Chem Int Ed Engl 35: 2837-2841.

Feldman, A. R., and D. Sen (2001). A new and efficient DNA enzyme for the sequence-specific cleavage of RNA. J Mol Biol 313: 283-294.

Feldstein, P. A., J. M. Buzayan, and G. Bruening (1989). Two sequences participating in the autolytic processing of satellite tobacco ringspot virus complementary RNA. Gene 82: 53-61.

Flynn-Charlebois, A., Y. Wang, T. K. Prior, I. Rashid, K. A. Hoadley, R. L. Coppins, A. C. Wolf, and S. K. Silverman (2003). Deoxyribozymes with 2'-5' RNA ligase activity. J Am Chem Soc 125: 2444-2454.

Forster, A. C., and R. H. Symons (1987). Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites. Cell 49: 211-220.

Förster, T. (1948). Intermolecular energy transference and fluorescence. Ann Physik 2: 55-57.

Gellert, M., M. N. Lipsett, and D. R. Davies (1962). Helix formation by guanylic acid. Proc Natl Acad Sci U S A 48: 2013-2018.

Geyer, C. R., and D. Sen (1997). Evidence for the metal-cofactor independence of an RNA phosphodiester-cleaving DNA enzyme. Chem Biol 4: 579-593.

Grimpe, B., S. Dong, C. Doller, K. Temple, A. T. Malouf, and J. Silver (2002). The critical role of basement membrane-independent laminin gamma 1 chain during axon regeneration in the CNS. J Neurosci 22: 3144-3160.

Grimpe, B., and J. Silver (2004). A novel DNA enzyme reduces glycosaminoglycan chains in the glial scar and allows microtransplanted dorsal root ganglia axons to regenerate beyond lesions in the spinal cord. J Neurosci 24: 1393-1397.

Guerrier-Takada, C., K. Gardiner, T. Marsh, N. Pace, and S. Altman (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell 35: 849-857.

He, Shijiang, Bo Song, Di Li, Changfeng Zhu, Wenpeng Qi, Yanqin Wen, Lihua Wang, Shiping Song, Haiping Fang, and Chunhai Fan (2010). A Graphene Nanoprobe for Rapid, Sensitive, and Multicolor Fluorescent DNA Analysis. Advanced Functional Materials 20: 453-459.

Hjiantoniou, E., S. Iseki, J. B. Uney, and L. A. Phylactou (2003). DNazyme-mediated cleavage of Twist transcripts and increase in cellular apoptosis. Biochem Biophys Res Commun 300: 178-181.

Hollenstein, M., C. J. Hipolito, C. H. Lam, and D. M. Perrin (2009). A self-cleaving DNA enzyme modified with amines, guanidines and imidazoles operates independently of divalent metal cations (M2+). Nucleic Acids Res 37: 1638-1649.

Hollenstein, M., C. Hipolito, C. Lam, D. Dietrich, and D. M. Perrin (2008). A highly selective DNAzyme sensor for mercuric ions. Angew Chem Int Ed Engl 47: 4346-4350.

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

Huppert, J. L., and S. Balasubramanian (2007). G-quadruplexes in promoters throughout the human genome. Nucleic Acids Res 35: 406-413.

Hurtado, A., H. Podinin, M. Oudega, and B. Grimpe (2008). Deoxyribozyme-mediated knockdown of xylosyltransferase-1 mRNA promotes axon growth in the adult rat spinal cord. Brain 131: 2596-2605.

Iversen, P. O., P. D. Emanuel, and M. Sioud (2002). Targeting Raf-1 gene expression by a DNA enzyme inhibits juvenile myelomonocytic leukemia cell growth. Blood 99: 4147-4153.

Iversen, P. O., G. Nicolaysen, and M. Sioud (2001). DNA enzyme targeting TNF-alpha mRNA improves hemodynamic performance in rats with postinfarction heart failure. Am J Physiol Heart Circ Physiol 281: H2211-2217.

Jarrous, N., and R. Reiner (2007). Human RNase P: a tRNA-processing enzyme and transcription factor. Nucleic Acids Res 35: 3519-3524.

Joachimi, A., A. Benz, and J. S. Hartig (2009). A comparison of DNA and RNA quadruplex structures and stabilities. Bioorg Med Chem 17: 6811-6815.

Kandadai, S. A., and Y. Li (2005). Characterization of a catalytically efficient acidic RNA-cleaving deoxyribozyme. Nucleic Acids Res 33: 7164-7175.

Kandadai, S. A., W. W. Mok, M. M. Ali, and Y. Li (2009). Characterization of an RNA-cleaving deoxyribozyme with optimal activity at pH 5. Biochemistry 48: 7383-7391.

Khachigian, L. M., R. G. Fahmy, G. Zhang, Y. V. Bobryshev, and A. Kaniaros (2002). c-Jun regulates vascular smooth muscle cell growth and neointima formation after arterial injury. Inhibition by a novel DNA enzyme targeting c-Jun. J Biol Chem 277: 22985-22991.

Klein, George, (1990). *The atheist and the holy city : encounters and reflections*. 1st MIT Press ed. Cambridge, Mass.: MIT Press.

Klussmann, Sven, (2006). *The aptamer handbook : functional oligonucleotides and their applications*. Weinheim: Wiley-VCH.

Kruger, K., P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, and T. R. Cech (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell 31: 147-157.

Kurreck, J., B. Bieber, R. Jahnel, and V. A. Erdmann (2002). Comparative study of DNA enzymes and ribozymes against the same full-length messenger RNA of the vanilloid receptor subtype I. J Biol Chem 277: 7099-7107.

Lam, C. H., C. J. Hipolito, M. Hollenstein, and D. M. Perrin (2011). A divalent metal-dependent self-cleaving DNAzyme with a tyrosine side chain. Org Biomol Chem 9: 6949-6954.

Lam, C. H., and D. M. Perrin (2010). Introduction of guanidinium-modified deoxyuridine into the substrate binding regions of DNAzyme 10-23 to enhance

target affinity: implications for DNAzyme design. Bioorg Med Chem Lett 20: 5119-5122.

Lan, T., K. Furuya, and Y. Lu (2010). A highly selective lead sensor based on a classic lead DNAzyme. Chem Commun (Camb) 46: 3896-3898.

Latham, J. A., R. Johnson, and J. J. Toole (1994). The application of a modified nucleotide in aptamer selection: novel thrombin aptamers containing 5-(1-pentynyl)-2'-deoxyuridine. Nucleic Acids Res 22: 2817-2822.

Lermer, L., Y. Roupioz, R. Ting, and D. M. Perrin (2002). Toward an RNaseA mimic: A DNAzyme with imidazoles and cationic amines. J Am Chem Soc 124: 9960-9961.

Levy, M., and A. D. Ellington (2002a). ATP-dependent allosteric DNA enzymes. Chem Biol 9: 417-426.

Levy, M., and A. D. Ellington (2002b). In vitro selection of a deoxyribozyme that can utilize multiple substrates. J Mol Evol 54: 180-190.

Levy, M., and A. D. Ellington (2001). Selection of deoxyribozyme ligases that catalyze the formation of an unnatural internucleotide linkage. Bioorg Med Chem 9: 2581-2587.

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

Li, Y., and R. R. Breaker (1999). Phosphorylating DNA with DNA. Proc Natl Acad Sci U S A 96: 2746-2751.

Li, Y., Y. Liu, and R. R. Breaker (2000). Capping DNA with DNA. Biochemistry 39: 3106-3114.

Li, Y., and D. Sen (1996). A catalytic DNA for porphyrin metallation. Nat Struct Biol 3: 743-747.

Li, Y., and D. Sen (1997). Toward an efficient DNAzyme. Biochemistry 36: 5589-5599.

Lilley, D. M. (2005). Structure, folding and mechanisms of ribozymes. Curr Opin Struct Biol 15: 313-323.

Liu, C., R. Cheng, L. Q. Sun, and P. Tien (2001). Suppression of platelet-type 12-lipoxygenase activity in human erythroleukemia cells by an RNA-cleaving DNAzyme. Biochem Biophys Res Commun 284: 1077-1082.

Liu, J., A. K. Brown, X. Meng, D. M. Cropek, J. D. Istok, D. B. Watson, and Y. Lu (2007a). A catalytic beacon sensor for uranium with parts-per-trillion sensitivity and millionfold selectivity. Proc Natl Acad Sci U S A 104: 2056-2061.

Liu, J., and Y. Lu (2004a). Accelerated color change of gold nanoparticles assembled by DNAzymes for simple and fast colorimetric Pb2+ detection. J Am Chem Soc 126: 12298-12305.

Liu, J., and Y. Lu (2004b). Adenosine-dependent assembly of aptazyme-functionalized gold nanoparticles and its application as a colorimetric biosensor. Anal Chem 76: 1627-1632.

Liu, J., and Y. Lu (2004c). Colorimetric biosensors based on DNAzyme-assembled gold nanoparticles. J Fluoresc 14: 343-354.

Liu, J., and Y. Lu (2007b). Colorimetric Cu2+ detection with a ligation DNAzyme and nanoparticles. Chem Commun (Camb) 4872-4874.

Liu, J., and Y. Lu (2003a). A colorimetric lead biosensor using DNAzyme-directed assembly of gold nanoparticles. J Am Chem Soc 125: 6642-6643.

Liu, J., and Y. Lu (2005a). Fast colorimetric sensing of adenosine and cocaine based on a general sensor design involving aptamers and nanoparticles. Angew Chem Int Ed Engl 45: 90-94.

Liu, J., and Y. Lu (2003b). Improving fluorescent DNAzyme biosensors by combining inter- and intramolecular quenchers. Anal Chem 75: 6666-6672.

Liu, J., and Y. Lu (2006). Preparation of aptamer-linked gold nanoparticle purple aggregates for colorimetric sensing of analytes. Nat Protoc 1: 246-252.

Liu, J., and Y. Lu (2007c). Rational design of "turn-on" allosteric DNAzyme catalytic beacons for aqueous mercury ions with ultrahigh sensitivity and selectivity. Angew Chem Int Ed Engl 46: 7587-7590.

Liu, J., and Y. Lu (2005b). Stimuli-responsive disassembly of nanoparticle aggregates for light-up colorimetric sensing. J Am Chem Soc 127: 12677-12683.

Liu, M., H. Zhao, S. Chen, H. Yu, Y. Zhang, and X. Quan (2011). Label-free fluorescent detection of Cu(II) ions based on DNA cleavage-dependent graphene-quenched DNAzymes. Chem Commun (Camb) 47: 7749-7751.

Liu, X., R. Freeman, E. Golub, and I. Willner (2011). Chemiluminescence and chemiluminescence resonance energy transfer (CRET) aptamer sensors using catalytic hemin/G-quadruplexes. ACS Nano 5: 7648-7655.

Liu, Z., S. H. Mei, J. D. Brennan, and Y. Li (2003). Assemblage of signaling DNA enzymes with intriguing metal-ion specificities and pH dependences. J Am Chem Soc 125: 7539-7545.

Lowe, H. C., C. N. Chesterman, and L. M. Khachigian (2002). Catalytic antisense DNA molecules targeting Egr-1 inhibit neointima formation following permanent ligation of rat common carotid arteries. Thromb Haemost 87: 134-140.

Lowe, H. C., R. G. Fahmy, M. M. Kavurma, A. Baker, C. N. Chesterman, and L. M. Khachigian (2001). Catalytic oligodeoxynucleotides define a key regulatory role for early growth response factor-1 in the porcine model of coronary in-stent restenosis. Circ Res 89: 670-677.

Lu, C. H., H. H. Yang, C. L. Zhu, X. Chen, and G. N. Chen (2009). A graphene platform for sensing biomolecules. Angew Chem Int Ed Engl 48: 4785-4787.

Lu, Z. X., X. Q. Ma, L. F. Yang, Z. L. Wang, L. Zeng, Z. J. Li, X. N. Li, M. Tang, W. Yi, J. P. Gong, L. Q. Sun, and Y. Cao (2008). DNAzymes targeted to EBV-encoded latent membrane protein-1 induce apoptosis and enhance radiosensitivity in nasopharyngeal carcinoma. Cancer Lett 265: 226-238.

Lu, Z. X., M. Ye, G. R. Yan, Q. Li, M. Tang, L. M. Lee, L. Q. Sun, and Y. Cao (2005). Effect of EBV LMP1 targeted DNAzymes on cell proliferation and apoptosis. Cancer Gene Ther 12: 647-654.

Masse, E., N. Majdalani, and S. Gottesman (2003). Regulatory roles for small RNAs in bacteria. Curr Opin Microbiol 6: 120-124.

May, J. P., R. Ting, L. Lermer, J. M. Thomas, Y. Roupioz, and D. M. Perrin (2004). Covalent Schiff base catalysis and turnover by a DNAzyme: a M2+ -independent AP-endonuclease mimic. J Am Chem Soc 126: 4145-4156.

McManus, S. A., and Y. Li (2010). The structural diversity of deoxyribozymes. Molecules 15: 6269-6284.

Mei, S. H., Z. Liu, J. D. Brennan, and Y. Li (2003). An efficient RNA-cleaving DNA enzyme that synchronizes catalysis with fluorescence signaling. J Am Chem Soc 125: 412-420.

Mitchell, A., C. R. Dass, L. Q. Sun, and L. M. Khachigian (2004). Inhibition of human breast carcinoma proliferation, migration, chemoinvasion and solid tumour growth by DNAzymes targeting the zinc finger transcription factor EGR-1. Nucleic Acids Res 32: 3065-3069.

Morgan, A. R., and R. D. Wells (1968). Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences. J Mol Biol 37: 63-80.

Moshe, M., J. Elbaz, and I. Willner (2009). Sensing of UO22+ and design of logic gates by the application of supramolecular constructs of ion-dependent DNAzymes. Nano Lett 9: 1196-1200.

Neidle, S., and S. Balasubramanian, (2006). *Quadruplex Nucleic Acids*. Royal Society of Chemistry.

Nutiu, R., and Y. Li (2005). Aptamers with fluorescence-signaling properties. Methods 37: 16-25.

Nutiu, R., and Y. Li (2003). Structure-switching signaling aptamers. J Am Chem Soc 125: 4771-4778.

Pelossof, G., R. Tel-Vered, J. Elbaz, and I. Willner (2010). Amplified biosensing using the horseradish peroxidase-mimicking DNAzyme as an electrocatalyst. Anal Chem 82: 4396-4402.

Peracchi, A., M. Bonaccio, and M. Clerici (2005). A mutational analysis of the 8-17 deoxyribozyme core. J Mol Biol 352: 783-794.

Perrin, D. M., T. Garestier, and C. Helene (2001). Bridging the gap between proteins and nucleic acids: a metal-independent RNAseA mimic with two protein-like functionalities. J Am Chem Soc 123: 1556-1563.

Perrin, D. M., T. Garestier, and C. Helene (1999). Expanding the catalytic repertoire of nucleic acid catalysts: simultaneous incorporation of two modified deoxyribonucleoside triphosphates bearing ammonium and imidazolyl functionalities. Nucleosides Nucleotides 18: 377-391.

Pradeepkumar, P. I., C. Hobartner, D. A. Baum, and S. K. Silverman (2008). DNA-catalyzed formation of nucleopeptide linkages. Angew Chem Int Ed Engl 47: 1753-1757.

Pratico, E. D., Y. Wang, and S. K. Silverman (2005). A deoxyribozyme that synthesizes 2',5'-branched RNA with any branch-site nucleotide. Nucleic Acids Res 33: 3503-3512.

Purtha, W. E., R. L. Coppins, M. K. Smalley, and S. K. Silverman (2005). General deoxyribozyme-catalyzed synthesis of native 3'-5' RNA linkages. J Am Chem Soc 127: 13124-13125.

Roth, A., and R. R. Breaker (1998). An amino acid as a cofactor for a catalytic polynucleotide. Proc Natl Acad Sci U S A 95: 6027-6031.

Sachdeva, A., and S. K. Silverman (2010). DNA-catalyzed serine side chain reactivity and selectivity. Chem Commun (Camb) 46: 2215-2217.

Santiago, F. S., H. C. Lowe, M. M. Kavurma, C. N. Chesterman, A. Baker, D. G. Atkins, and L. M. Khachigian (1999). New DNA enzyme targeting Egr-1 mRNA inhibits vascular smooth muscle proliferation and regrowth after injury. Nat Med 5: 1264-1269.

Santoro, S. W., and G. F. Joyce (1997). A general purpose RNA-cleaving DNA enzyme. Proc Natl Acad Sci U S A 94: 4262-4266.

Santoro, S. W., G. F. Joyce, K. Sakthivel, S. Gramatikova, and C. F. Barbas, 3rd (2000). RNA cleavage by a DNA enzyme with extended chemical functionality. J Am Chem Soc 122: 2433-2439.

Sato, K., K. Hosokawa, and M. Maeda (2003). Rapid aggregation of gold nanoparticles induced by non-cross-linking DNA hybridization. J Am Chem Soc 125: 8102-8103.

Schlosser, K., and Y. Li (2004). Tracing sequence diversity change of RNA-cleaving deoxyribozymes under increasing selection pressure during in vitro selection. Biochemistry 43: 9695-9707.

Sel, S., M. Wegmann, T. Dicke, S. Sel, W. Henke, A. O. Yildirim, H. Renz, and H. Garn (2008). Effective prevention and therapy of experimental allergic asthma using a GATA-3-specific DNAzyme. J Allergy Clin Immunol 121: 910-916 e915.

Sen, D., and W. Gilbert (1988). Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature 334: 364-366.

Shen, L., Z. Chen, Y. Li, S. He, S. Xie, X. Xu, Z. Liang, X. Meng, Q. Li, Z. Zhu, M. Li, X. C. Le, and Y. Shao (2008). Electrochemical DNAzyme sensor for lead based on amplification of DNA-Au bio-bar codes. Anal Chem 80: 6323-6328.

Shen, Y., J. D. Brennan, and Y. Li (2005). Characterizing the secondary structure and identifying functionally essential nucleotides of pH6DZ1, a fluorescence-signaling and RNA-cleaving deoxyribozyme. Biochemistry 44: 12066-12076.

Shen, Y., W. Chiuman, J. D. Brennan, and Y. Li (2006). Catalysis and rational engineering of trans-acting pH6DZ1, an RNA-cleaving and fluorescence-signaling deoxyribozyme with a four-way junction structure. Chembiochem 7: 1343-1348.

Sheppard, T. L., P. Ordoukhanian, and G. F. Joyce (2000). A DNA enzyme with N-glycosylase activity. Proc Natl Acad Sci U S A 97: 7802-7807.

Sidorov, A. V., J. A. Grasby, and D. M. Williams (2004). Sequence-specific cleavage of RNA in the absence of divalent metal ions by a DNAzyme incorporating imidazolyl and amino functionalities. Nucleic Acids Res 32: 1591-1601.

Sioud, M., and M. Leirdal (2000). Design of nuclease resistant protein kinase calpha DNA enzymes with potential therapeutic application. J Mol Biol 296: 937-947.

Soukup, G. A., and R. R. Breaker (1999). Engineering precision RNA molecular switches. Proc Natl Acad Sci U S A 96: 3584-3589.

Spiegelman, S., I. Haruna, I. B. Holland, G. Beaudreau, and D. Mills (1965). The synthesis of a self-propagating and infectious nucleic acid with a purified enzyme. Proc Natl Acad Sci U S A 54: 919-927.

Sreedhara, A., Y. Li, and R. R. Breaker (2004). Ligating DNA with DNA. J Am Chem Soc 126: 3454-3460.

Stoltenburg, R., C. Reinemann, and B. Strehlitz (2007). SELEX--a (r)evolutionary method to generate high-affinity nucleic acid ligands. Biomol Eng 24: 381-403.

Sun, L. Q., M. J. Cairns, W. L. Gerlach, C. Witherington, L. Wang, and A. King (1999). Suppression of smooth muscle cell proliferation by a c-myc RNA-cleaving deoxyribozyme. J Biol Chem 274: 17236-17241.

Takahashi, H., H. Hamazaki, Y. Habu, M. Hayashi, T. Abe, N. Miyano-Kurosaki, and H. Takaku (2004). A new modified DNA enzyme that targets influenza virus A mRNA inhibits viral infection in cultured cells. FEBS Lett 560: 69-74.

Tan, X. X., K. Rose, W. Margolin, and Y. Chen (2004). DNA enzyme generated by a novel single-stranded DNA expression vector inhibits expression of the essential bacterial cell division gene ftsZ. Biochemistry 43: 1111-1117.

Tang, J., and R. R. Breaker (1998). Mechanism for allosteric inhibition of an ATP-sensitive ribozyme. Nucleic Acids Res 26: 4214-4221.

Tang, J., and R. R. Breaker (1997). Rational design of allosteric ribozymes. Chem Biol 4: 453-459.

Taton, T. A., C. A. Mirkin, and R. L. Letsinger (2000). Scanometric DNA array detection with nanoparticle probes. Science 289: 1757-1760.

Teller, C., S. Shimron, and I. Willner (2009). Aptamer-DNAzyme hairpins for amplified biosensing. Anal Chem 81: 9114-9119.

Thomas, J. M., R. Ting, and D. M. Perrin (2004). High affinity DNAzyme-based ligands for transition metal cations - a prototype sensor for Hg2+. Org Biomol Chem 2: 307-312.

Thomas, J. M., J. K. Yoon, and D. M. Perrin (2009). Investigation of the catalytic mechanism of a synthetic DNAzyme with protein-like functionality: an RNaseA mimic? J Am Chem Soc 131: 5648-5658.

Ting, R., L. Lermer, and D. M. Perrin (2004a). Triggering DNAzymes with light: a photoactive C8 thioether-linked adenosine. J Am Chem Soc 126: 12720-12721.

Ting, R., J. M. Thomas, L. Lermer, and D. M. Perrin (2004b). Substrate specificity and kinetic framework of a DNAzyme with an expanded chemical repertoire: a putative RNaseA mimic that catalyzes RNA hydrolysis independent of a divalent metal cation. Nucleic Acids Res 32: 6660-6672.

Travascio, P., A. J. Bennet, D. Y. Wang, and D. Sen (1999). A ribozyme and a catalytic DNA with peroxidase activity: active sites versus cofactor-binding sites. Chem Biol 6: 779-787.

Travascio, P., Y. Li, and D. Sen (1998). DNA-enhanced peroxidase activity of a DNA-aptamer-hemin complex. Chem Biol 5: 505-517.

Tucker, B. J., and R. R. Breaker (2005). Riboswitches as versatile gene control elements. Curr Opin Struct Biol 15: 342-348.

Tuerk, C., and L. Gold (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 249: 505-510.

Tyagi, S., and F. R. Kramer (1996). Molecular beacons: probes that fluoresce upon hybridization. Nat Biotechnol 14: 303-308.

Unwalla, H., and A. C. Banerjea (2001). Inhibition of HIV-1 gene expression by novel macrophage-tropic DNA enzymes targeted to cleave HIV-1 TAT/Rev RNA. Biochem J 357: 147-155.

Vallee-Belisle, A, F. Ricci, and K. W. Plaxco (2009). Thermodynamic basis for the optimization of binding-induced biomolecular switches and structure-switching biosensors. Proc Natl Acad Sci U S A 106: 13802-13807.

van Gijlswijk, R. P., E. G. Talman, I. Peekel, J. Bloem, M. A. van Velzen, R. J. Heetebrij, and H. J. Tanke (2002). Use of horseradish peroxidase- and fluorescein-modified cisplatin derivatives for simultaneous labeling of nucleic acids and proteins. Clin Chem 48: 1352-1359.

Vester, B., L. B. Lundberg, M. D. Sorensen, B. R. Babu, S. Douthwaite, and J. Wengel (2002). LNAzymes: incorporation of LNA-type monomers into DNAzymes markedly increases RNA cleavage. J Am Chem Soc 124: 13682-13683.

Wang, B., L. Cao, W. Chiuman, Y. Li, and Z. Xi (2010). Probing the function of nucleotides in the catalytic cores of the 8-17 and 10-23 DNAzymes by abasic nucleotide and C3 spacer substitutions. Biochemistry 49: 7553-7562.

Wang, D. Y., B. H. Lai, and D. Sen (2002). A general strategy for effector-mediated control of RNA-cleaving ribozymes and DNA enzymes. J Mol Biol 318: 33-43.

Wang, H., Y. Kim, H. Liu, Z. Zhu, S. Bamrungsap, and W. Tan (2009). Engineering a unimolecular DNA-catalytic probe for single lead ion monitoring. J Am Chem Soc 131: 8221-8226.

Wang, L., X. Liu, X. Hu, S. Song, and C. Fan (2006). Unmodified gold nanoparticles as a colorimetric probe for potassium DNA aptamers. Chem Commun (Camb) 3780-3782.

Wang, W., L. P. Billen, and Y. Li (2002). Sequence diversity, metal specificity, and catalytic proficiency of metal-dependent phosphorylating DNA enzymes. Chem Biol 9: 507-517.

Wang, Y., and D. J. Patel (1993). Solution structure of a parallel-stranded G-quadruplex DNA. J Mol Biol 234: 1171-1183.

Warashina, M., T. Kuwabara, Y. Nakamatsu, and K. Taira (1999). Extremely high and specific activity of DNA enzymes in cells with a Philadelphia chromosome. Chem Biol 6: 237-250.

Wei, H., B. Li, J. Li, E. Wang, and S. Dong (2007). Simple and sensitive aptamer-based colorimetric sensing of protein using unmodified gold nanoparticle probes. Chem Commun (Camb) 3735-3737.

Wells, R. D., D. A. Collier, J. C. Hanvey, M. Shimizu, and F. Wohlrab (1988). The chemistry and biology of unusual DNA structures adopted by oligopurine.oligopyrimidine sequences. FASEB J 2: 2939-2949.

Wen, Y., C. Peng, D. Li, L. Zhuo, S. He, L. Wang, Q. Huang, Q. H. Xu, and C. Fan (2011). Metal ion-modulated graphene-DNAzyme interactions: design of a nanoprobe for fluorescent detection of lead(II) ions with high sensitivity, selectivity and tunable dynamic range. Chem Commun (Camb) 47: 6278-6280.

Wieland, M., and J. S. Hartig (2006). Turning inhibitors into activators: a hammerhead ribozyme controlled by a guanine quadruplex. Angew Chem Int Ed Engl 45: 5875-5878.

Williamson, J. R., M. K. Raghuraman, and T. R. Cech (1989). Monovalent cation-induced structure of telomeric DNA: the G-quartet model. Cell 59: 871-880.

Willner, I., B. Shlyahovsky, M. Zayats, and B. Willner (2008). DNAzymes for sensing, nanobiotechnology and logic gate applications. Chem Soc Rev 37: 1153-1165.

Wong, O. Y., P. I. Pradeepkumar, and S. K. Silverman (2011). DNA-catalyzed covalent modification of amino acid side chains in tethered and free peptide substrates. Biochemistry 50: 4741-4749.

Xiang, Y., and Y. Lu (2011). Using personal glucose meters and functional DNA sensors to quantify a variety of analytical targets. Nat Chem 3: 697-703.

Xiang, Y., A. Tong, and Y. Lu (2009). Abasic site-containing DNAzyme and aptamer for label-free fluorescent detection of Pb(2+) and adenosine with high sensitivity, selectivity, and tunable dynamic range. J Am Chem Soc 131: 15352-15357.

Xiang, Y., Z. Wang, H. Xing, N. Y. Wong, and Y. Lu (2010). Label-free fluorescent functional DNA sensors using unmodified DNA: a vacant site approach. Anal Chem 82: 4122-4129.

Xiao, Y., A. A. Rowe, and K. W. Plaxco (2007). Electrochemical detection of parts-per-billion lead via an electrode-bound DNAzyme assembly. J Am Chem Soc 129: 262-263.

Yang, Xiaorong, Juan Xu, Xuemei Tang, Huixiang Liu, and Danbi Tian (2010). A novel electrochemical DNAzyme sensor for the amplified detection of Pb2+ ions. Chemical Communications 46: 3107.

Yen, L., S. M. Strittmatter, and R. G. Kalb (1999). Sequence-specific cleavage of Huntingtin mRNA by catalytic DNA. Ann Neurol 46: 366-373.

Zaborowska, Z., J. P. Furste, V. A. Erdmann, and J. Kurreck (2002). Sequence requirements in the catalytic core of the "10-23" DNA enzyme. J Biol Chem 277: 40617-40622.

Zaborowska, Z., S. Schubert, J. Kurreck, and V. A. Erdmann (2005). Deletion analysis in the catalytic region of the 10-23 DNA enzyme. FEBS Lett 579: 554-558.

Zaug, A. J., P. J. Grabowski, and T. R. Cech (1983). Autocatalytic cyclization of an excised intervening sequence RNA is a cleavage-ligation reaction. Nature 301: 578-583.

Zhang, G., C. R. Dass, E. Sumithran, N. Di Girolamo, L. Q. Sun, and L. M. Khachigian (2004). Effect of deoxyribozymes targeting c-Jun on solid tumor growth and angiogenesis in rodents. J Natl Cancer Inst 96: 683-696.

Zhang, L., W. J. Gasper, S. A. Stass, O. B. Ioffe, M. A. Davis, and A. J. Mixson (2002). Angiogenic inhibition mediated by a DNAzyme that targets vascular endothelial growth factor receptor 2. Cancer Res 62: 5463-5469.

Zhang, X., Y. Xu, H. Ling, and T. Hattori (1999). Inhibition of infection of incoming HIV-1 virus by RNA-cleaving DNA enzyme. FEBS Lett 458: 151-156.

Zhao, W., W. Chiuman, M. A. Brook, and Y. Li (2007). Simple and rapid colorimetric biosensors based on DNA aptamer and noncrosslinking gold nanoparticle aggregation. Chembiochem 8: 727-731.

Zhao, W., W. Chiuman, J. C. Lam, S. A. McManus, W. Chen, Y. Cui, R. Pelton, M. A. Brook, and Y. Li (2008a). DNA aptamer folding on gold nanoparticles: from colloid chemistry to biosensors. J Am Chem Soc 130: 3610-3618.

Zhao, W., J. C. Lam, W. Chiuman, M. A. Brook, and Y. Li (2008b). Enzymatic cleavage of nucleic acids on gold nanoparticles: a generic platform for facile colorimetric biosensors. Small 4: 810-816.

Zhao, X. H., R. M. Kong, X. B. Zhang, H. M. Meng, W. N. Liu, W. Tan, G. L. Shen, and R. Q. Yu (2011). Graphene-DNAzyme based biosensor for amplified fluorescence "turn-on" detection of Pb2+ with a high selectivity. Anal Chem 83: 5062-5066.

Zhou, J., X. Q. Yang, Y. Y. Xie, X. D. Zhao, L. P. Jiang, L. J. Wang, and Y. X. Cui (2007). Inhibition of respiratory syncytial virus of subgroups A and B using deoxyribozyme DZ1133 in mice. Virus Res 130: 241-248.

Zhou, M., Y. Zhai, and S. Dong (2009). Electrochemical sensing and biosensing platform based on chemically reduced graphene oxide. Anal Chem 81: 5603-5613.

Zhu, Xi, Zhenyu Lin, Lifeng Chen, Bin Qiu, and Guonan Chen (2009). A sensitive and specific electrochemiluminescent sensor for lead based on DNAzyme. Chemical Communications 6050.

Chapter 2:
Multiple Occurrences of an Efficient Self-Phosphorylating Deoxyribozyme Motif

2.1 AUTHOR'S PREFACE

The research project described in this chapter reports the identification and characterization of a complex secondary structure isolated repeatedly in the terminal population of a previous in vitro selection experiment. This structural characterization described in this chapter was completed using mutational analysis, site-specific mutation testing, substrate and co-factor activity assays, and chemical probing of intramolecular interactions by dimethyl sulfate interference assays. The major finding of this work was that three deoxyribozymes independently evolved a common secondary structure to perform self-phosphorylation. This was a significant contribution to the field of nucleic acid selection, as it demonstrated that while DNA libraries typically do not contain many sequences folded into complex structural arrangements, an efficient complex motif could still find different means of surviving an in vitro selection experiment. This is also the largest and most complex motif repeatedly isolated by in vitro selection.

This chapter is a modified version of a published research article. I am the first author on this publication and I conducted all of the experiments presented within and wrote the manuscript. I also took a leading role in the conceptualization, experimental design and interpretation of the results obtained. Dr. Yingfu Li is listed as an author of this paper and provided many fruitful suggestions and guidance as well as assisting in editing of the final version of this manuscript. This citation for the original publication is as follows:

McManus, S.A. and Li, Y. (2007) Multiple occurrences of an efficient self-phosphorylation motif. *Biochemistry*, 46, 2198-2204.

2.2 ABSTRACT

The catalytic and structural characteristics of two new self-phosphorylating deoxyribozymes (referred to as deoxyribozyme kinases), denoted “Dk3” and “Dk4”, are compared to those of Dk2, a previously reported deoxyribozyme kinase. All three deoxyribozymes not only utilize GTP as the source of activated phosphate and Mn(II) as the divalent metal cofactor but also share a common secondary structure with significant sequence variations. Multiple Watson–Crick helices are identified within the secondary structure, and these helical interactions confine three extremely conserved sequence elements of 8, 5, and 14 nucleotides in length, presumably for the formation of the catalytic core for GTP binding and the self-phosphorylating reaction. The locations of the conserved regions suggest that these three deoxyribozymes arose independently from in vitro selection. The existence of three sequence variants of the same deoxyribozyme from the same in vitro selection experiment implies that these catalytic DNAs may represent the simplest structural solution for the DNA self-phosphorylation reaction when GTP is used as the substrate.

2.3 INTRODUCTION

DNA is inherently stable, and its propensity for double-helical structure formation makes it the ideal material for storage of genetic information. In single-stranded form, DNA behaves similarly to RNA and has the ability to perform functions such as ligand binding and catalysis. Although no naturally occurring enzymes made of DNA have been found to date, many catalytic DNA molecules (deoxyribozymes or DNAzymes) have been isolated from single-stranded random-sequence DNA libraries by in vitro selection (1–4). Deoxyribozymes are now known to catalyze an increasing number of reactions including RNA cleavage (5), DNA cleavage (6), DNA phosphorylation (7), DNA adenylation (8), DNA ligation (9), RNA ligation (10), RNA branching (11), depurination (12), thymine dimer repair (13), and porphyrin ring metalation (14). The existence of such a large array of deoxyribozymes implies that DNA is capable of folding into complex structures needed for catalysis.

Deoxyribozyme kinases are single-stranded DNA molecules capable of self-phosphorylation at their 5′ ends. The reaction scheme is illustrated in Figure 2.1A. These deoxyribozymes have previously been isolated in experiments designed to investigate catalytic DNA sequence diversity, metal ion cofactor specificity, and DNA's molecular discrimination ability toward small-molecule substrates (7, 15). We are currently using this model system to examine the level of structural diversity and complexity capable by DNA catalysts. Specifically, we are studying the structures of several highly efficient kinase deoxyribozymes obtained from an in vitro selection experiment conducted under stringent reaction conditions. We have previously characterized two of these

deoxyribozymes (named Dk1 and Dk2) and reported their probable secondary structures (16). After further analyzing the terminal population of this selection, we found two other molecules that are significantly different in primary sequences from Dk1 and Dk2. Interestingly, the results from the characterization and structural analysis to be described in this report show that these two additional deoxyribozymes share a common secondary structure with Dk2. This secondary structure contains multiple short helical regions and 26 absolutely conserved nucleotides. The multiple observation of this structural motif suggests that it may represent the simplest structural arrangement to perform DNA self-phosphorylating reaction under the in vitro selection conditions. These results were obtained through substrate and metal ion cofactor characterization, structural probing via base-pair alteration, and methylation interference to identify guanosine residues that are essential to the deoxyribozyme activity.

2.4 EXPERIMENTAL PROCEDURES

2.4.1 Oligonucleotides and Other Materials

DNA was prepared using standard phosphoramidite chemistry (Mobix Laboratory, McMaster University; Integrated DNA Technologies, Coralville, IA). Oligonucleotides were purified by 10% denaturing (8 M urea) PAGE. Purified oligonucleotides were dissolved in water and their concentrations determined spectroscopically as previously described (15). T4 PNK, T4 DNA ligase, and CIAP were purchased from MBI Fermentas. [γ-32P]ATP was purchased from GE Healthcare. All other chemicals were purchased from Sigma.

2.4.2 Catalytic Assays

Dk3 and Dk4 constructs used in secondary-structural studies were assembled by ligation of two oligonucleotides with the donor oligonucleotide labeled with a ^{32}P at its 5′ end. This procedure generated deoxyribozyme constructs internally labeled with ^{32}P to allow the detection of these molecules. Ligation was performed as follows. Donor DNA was first phosphorylated using PNK in the presence of [γ-32P]ATP. After incubation for 30 min at 37 °C, 1 mM nonradiolabeled ATP was added and the solution incubated for an additional 10 min to ensure complete phosphorylation. The solution was then incubated at 90 °C for 5 min to inactivate PNK. Template and acceptor oligonucleotides were then added directly to this solution, followed by T4 DNA ligase buffer and T4 DNA ligase. The solution was incubated at room temperature for 1 h for ligation. The mixture was then precipitated with ethanol and the ligated DNA isolated by 10% denaturing PAGE. All assembled constructs were treated with CIAP to remove any 5′-phosphate. DNA was heated to 90 °C for 1 min and cooled rapidly to room temperature to prevent any structural folding. CIAP buffer (10×) and CIAP were added, and the solution was incubated at 37 °C for 1 h. CIAP was then removed with two rounds of equal-volume phenol/chloroform extractions. For self-phosphorylation, reaction volumes were 10 μL. The phosphorylation trials were carried out as follows. DNA was dissolved in water at a concentration of 200 nM. Reaction buffer (2×) was added (800 mM NaCl, 200 mM KCl, 20 mM $MnCl_2$, 100 mM HEPES, pH 7.0 at 23 °C), and the solution was incubated for 5 min at room temperature. The reactions were initiated by the addition of a nucleoside 5′-triphosphate (such as GTP) to the final concentration of 1 mM. After incubation for a set

length of time, the reactions were quenched with Na_2EDTA (at the final concentration of 15 mM) as well as 10 pmol of the template DNA and 15 pmol of the acceptor DNA. Solutions were immediately precipitated with ethanol. The DNA was resuspended and ligated using 10× T4 DNA ligase buffer and 5 units of T4 DNA ligase. The reactions were incubated at room temperature for 1 h. The solutions were precipitated with ethanol before being subjected to analysis by 10% denaturing PAGE. The rate constants were determined using Graphpad Prism software by plotting percent ligation versus reaction time. All reported rates are the average of at least two independent time–course experiments with a variation of less than 20%.

For Dk3 and Dk4 mutants, sequence differences are as follows. Dk3: covaried P1, T8 to A and A15 to T; mismatch P1, G7 to C; covaried P2, G21 to C and C55 to G; mismatch P2, G54 to C; covaried P3, G27 to C and C40 to G, mismatch P3, G27 to C. Dk4: covaried P1, G9 to C and C19 to G; mismatch P1, G9 to C; covaried P2, A25 to T and T58 to A; mismatch P2, C26 to G; covaried P3, G31 to C and C43 to G; mismatch P3, G31 to C.

2.4.3 Methylation Interference Assays

Each reaction started with 200 pmol of DNA. For the control reactions, each relevant deoxyribozyme was incubated with 1 mM GTP, as described above, to initiate self-phosphorylation. After being quenched and precipitated, the DNA was dissolved in water and methylated by adding an equal volume of freshly made 0.4% DMS (v/v) and incubated at room temperature for 40 min. The DNA was recovered by precipitation with ethanol and two 70% ethanol washes. For the test reactions, the same procedure was

followed except that the DMS treatment was performed before the self-phosphorylation reaction. The DNA from these reactions was then ligated to the acceptor DNA as described for the catalytic assays. The ligated DNA was isolated by 10% denaturing PAGE. The deoxyribozyme-acceptor chimeras were then treated with PNK and [γ-32P]ATP to label them at their 5′ ends. The reactions were precipitated with ethanol and purified by 10% denaturing PAGE. The DNA was then subjected to methylation-dependent cleavage by resuspending in 50 μL of 10% piperidine (v/v). After incubation for 30 min at 90 °C, the cleavage products were precipitated with ethanol and resolved by 10% denaturing PAGE.

2.5 RESULTS

2.5.1 Sequence Comparison of Dk2, Dk3, and Dk4

The sequences of Dk2, Dk3, and Dk4 are shown in Figure 2.1C. It should be noted that the DNA library used for the original in vitro selection experiment (15) contained the fixed sequence of 5′-**GGAAGAGAT**GGCGAC (the first nine nucleotides are in bold for comparison to the mutated sequence below). This sequence was designed to allow hybridization to the template DNA for the isolation of active sequences by the DNA ligase-mediated DNA ligation reaction (see Figure 2.1B). However, upon the completion of the in vitro selection experiment, all three deoxyribozymes were found to have a common, but altered, eight-nucleotide (nt) sequence element at their 5′-ends, GGAΔGAG*G*T (blue letters in Figure 2.1C), with one deletion (indicated by Δ) and one base mutation (italicized G). In addition to this common motif, stretches of 5 and 14 nt (red and green letters, respectively, in Figure 2.1C) are found to be conserved in the

interior of the deoxyribozyme sequences (with a one-base mutation in Dk4, which is shown in gray). It is noteworthy that these internal sequence motifs are found at different locations within the three deoxyribozymes and are sandwiched between different nucleotide content on either side in each deoxyribozyme. This, coupled with the selection strategy used, strongly suggests that the three deoxyribozymes arose independently from the original random-sequence DNA library and will be discussed later.

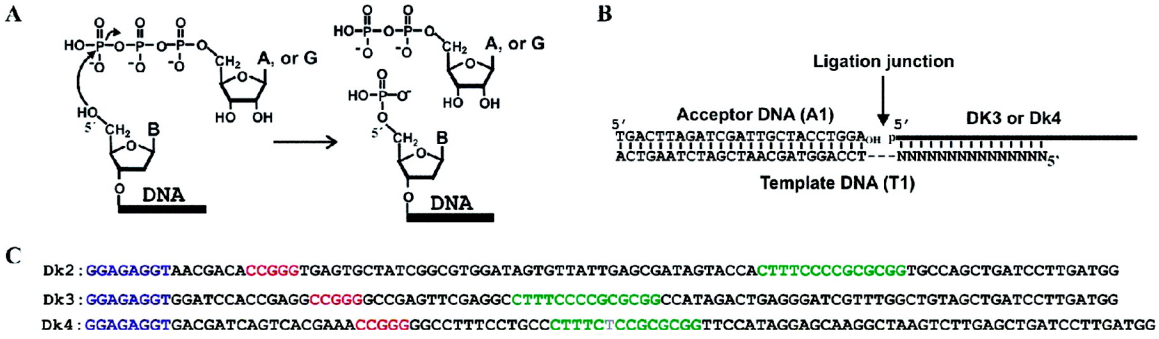


Figure 2.1 Mechanism of DNA-catalyzed self-phosphorylation (A) Proposed reaction mechanism for DNA self-phosphorylation. Nucleophilic attack by the 5′-hydroxyl group on the γ-phosphate of ATP or GTP results in phosphate transfer to the 5′ end of the deoxyribozyme. (B) Indirect method for detection of phosphorylation. After phosphorylation, the DNAzyme is ligated to an acceptor DNA in the presence of T4 DNA ligase. Only deoxyribozymes that have reacted (therefore acquired a 5′-phosphate) will be able to ligate. The increased size of the catalytic DNA molecules allows easy identification and isolation by PAGE separation. N's in T1 refer to the antisense of the 5′ nucleotides of Dk2, Dk3, or Dk4, as shown in part C of this figure. (C) Sequences of Dk2, Dk3, and Dk4. Conserved domains are colored.

2.5.2 Characterization of Dk3 and Dk4

Dk3 and Dk4 were tested to see whether they shared the same substrate and metal ion cofactor requirements as Dk2, which was shown to use GTP as the source of phosphate and was dependent on Mn2+ for activity (15). Through the indirect ligation assay illustrated in Figure 2.1B, it is evident that both Dk3 and Dk4 indeed use GTP as a substrate and Mn2+ as a divalent metal ion cofactor (Figure 2.2A). These identical

characteristics, along with the conserved sequence motifs, suggest that Dk2, Dk3, and Dk4 may use the same structural arrangement to bind GTP and perform catalysis.

To further characterize the substrate requirements for Dk2, Dk3, and Dk4, we conducted a specificity test using several GTP analogues (their chemical structures are given in Supporting Information, Figure 2.1). As shown in Figure 2.2B, Dk3 and Dk4 can accept both GTP and dGTP as the substrate but not 7-methylguanosine 5′-triphosphate [m(7)GTP] and 7-deazaguanosine 5′-triphosphate (N7-dGTP). Dk3 exhibited very low activity (1% yield in 2 h) when inosine 5′-triphosphate (ITP) was tested as a substrate while Dk2 and Dk4 showed no activity toward ITP. These data indicate that all three deoxyribozymes recognize guanine highly specifically and the amine group at the C2 position on guanine is important for the deoxyribozyme activity.

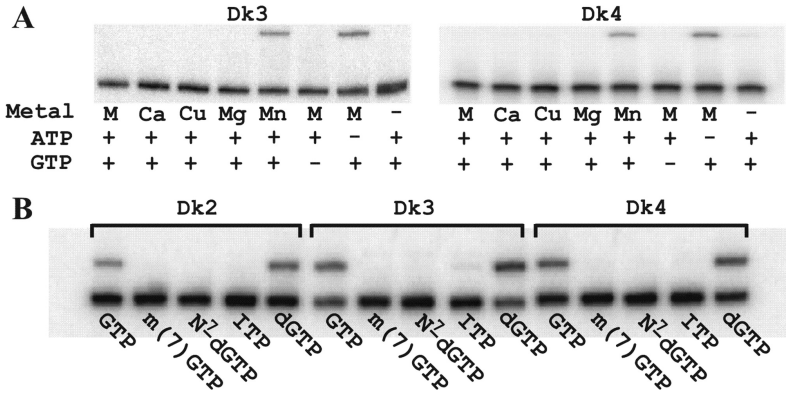


Figure 2.2 Characterization of Dk3 and Dk4. The upper bands in the gels represent reacted deoxyribozyme ligated to an acceptor DNA. The lower bands represent unreacted DNA. (A) Both deoxyribozymes were tested using different divalent metal ions and with ATP and GTP, as shown. M stands for a combination of Ca2+ (10 mM), Cu2+ (50 μM, plus 10 mM Mg2+), Mg2+ (10 mM), and Mn2+ (10 mM). (B) Test of substrate specificity for Dk3 and Dk4. 1 mM of respective substrate was used for each reaction. N7-dGTP represents 7-deazaguanosine 5′-triphosphate and m(7)GTP represents 7-methylguanosine 5′-triphosphate.

2.5.3 Secondary-Structure Study of Dk3 and Dk4

The secondary structures of Dk3 and Dk4 were analyzed using a variety of synthetic deoxyribozyme constructs and compared with the previously reported structure of Dk2 (16). Predicted structures for Dk3 and Dk4 using mfold software (17) are shown in Figure 2.3. The two structures are very similar, both containing four stem–loop regions, with the 14-nt internal conserved sequence acting as a junction between two of the stems. The secondary structures of Dk3 and Dk4 were confirmed and refined through the testing of mutant constructs with specific nucleotides mutated or removed. The key results are summarized in Table 2.1. First, truncations were made from the 3′ end to determine how many nucleotides could be removed without the loss of activity. For both Dk3 and Dk4, it was found that truncations could be made up to the end of the stem P2 (M1, Figure 2.3) without significant loss of activity. Further truncation resulted in inactive constructs (such as M2, Figure 2.3).

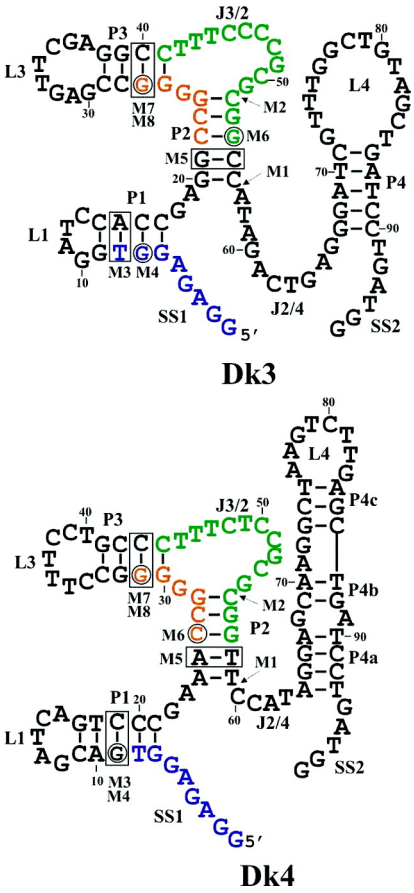


Figure 2.3 Secondary-structure prediction and confirmation for Dk3 and Dk4. Secondary structures of Dk3 and Dk4 predicted by the mfold program are shown separately. SS represents a single-stranded region, P a base-paired region, L a loop region, and J a junction between two base-paired regions. Lines with arrowheads indicate 3′-truncations up to that site. Rectangles around base pairs indicate sites where constructs were made with "flipped" base pairs (i.e., a C-G to a G-C). Circled bases were mutated to create a mismatch in the predicted stem. The exact changes are shown in Experimental Procedures. Reaction rates and percent yield for each mutant construct are shown in Table 2.1.

To test the three remaining predicted stems of both deoxyribozymes, constructs were made with different base pairs in each stem. As shown in Table 2.1, all covariation mutants for each stem were active for both Dk3 and Dk4, suggesting that these are indeed base-paired regions. Notably for both deoxyribozymes, the P2 stem shows a significant

drop in activity and reaction yield when covariations are introduced. This suggests that the wild-type base pairs are preferred over different base-pair combinations. One possible explanation is that this region is involved in other tertiary interactions as well as being a stem. To further confirm these three stems, constructs were made in which mismatches were introduced into these stem regions. If these stems are present in the active structure and necessary for folding, then adding these mismatches should weaken the stem and reduce the catalytic activity. As shown in the table, both catalytic rate and reaction yield were significantly reduced or absent in all but one of the constructs tested, further confirming that these predicted base-pair regions are present and important in the active structure of Dk3 and Dk4. The only exception is P3 of Dk4, which showed catalytic rate and yield comparable to its corresponding covariation mutant. This could mean either that P3 is not necessary for proper folding of Dk4 or that a single mismatch is not enough to disrupt this stem. In addition to the mfold-predicted stems, we found a potential stem between the 5′ region and the large central conserved sequence element in both Dk3 and Dk4. Several constructs were designed to test this possibility, and the data (Supporting Information, Figure 2.2) suggest that this helical region may not exist or that the base content is essential to catalysis and could not be altered.

Table 2.1 Reaction Rates and Percent Yields of Truncated and Mutant Constructs Listed for Dk3 and Dk4. Abbreviations are the same as Figure 2.3.

	Dk3		Dk4	
	Rate (min-1)	Percent yield	Rate (min-1)	Percent Yield
Wild type	0.19	76	0.28	51
M1 (3' truncation up to P3)	0.11	80	0.12	80
M2 (3' truncation + half of P3)	No activity		No activity	
M3 (Co-varied base P1)	0.09	75	0.02	78
M4 (Mismatched P1)	0.01	56	0.0058	26
M5 (Co-varied base P2)	0.03	35	0.018	10
M6 (Mismatched P2)	0.002	5	No activity	
M7 (Co-varied P3)	0.06	69	0.014	64
M8 (Mismatched P3)	0.002	30	**0.012**	**62**

2.5.4 Structural Comparison of Dk2, Dk3, and Dk4

After revealing that Dk3 and Dk4 contained a common structural arrangement, we reexamined the previously reported structural data for Dk2 to look for structural similarities. We found that, indeed, the secondary structure of Dk2 is quite similar to those of Dk3 and Dk4. As shown in Figure 2.4, Dk2 contains two stems (corresponding to P2 and P3 in Dk3 and Dk4) flanking the conserved 14-nt sequence motif. In fact, when comparing the structures, we were able to identify a 3-bp extension of the central stem in Dk2 (P′ in Figure 2.4) that was not identified previously. This stem was confirmed through identification of covarying mutations present in the previously reported Dk2 reselection data (16). The confirmed stem P1 in Dk3 and Dk4 is absent in Dk2.

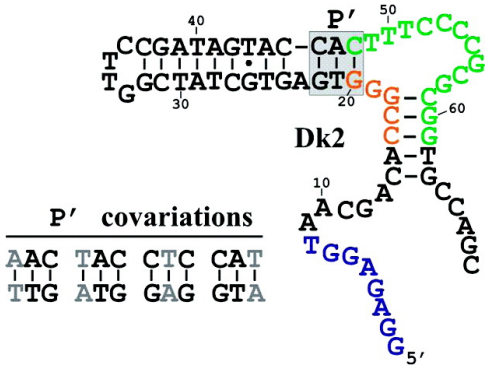


Figure 2.4 Secondary-structural model of Dk2. The newly observed stem addition P′ is shown in a gray box. Observed covariations of newly discovered P′ are shown with gray letters at the bottom left of this figure.

2.5.5 Methylation Interference Studies

Methylation interference assays have previously been used with several deoxyribozymes as a method of identifying guanine residues that are involved in tertiary interactions (7, 8, 18–21). The method used on deoxyribozyme kinases has been previously described in detail (16). Briefly, dimethyl sulfate (DMS) is used to methylate guanine residues specifically at their N7 position. These molecules are screened for activity, cleaved by piperidine at methylated guanine sites, and resolved by denaturing PAGE. This procedure can reveal sites where methylation interferes with deoxyribozyme activity. The results of the methylation interference experiments for Dk2, Dk3, and Dk4 are shown in Figure 2.5. It should be noted that methylation interference analysis has been performed previously for Dk2 (16) but was redone here for direct comparison with Dk3 and Dk4, as a different labeling procedure was used previously. The interference profiles for all three deoxyribozymes show a similar pattern with four guanines exhibiting obvious interference when methylated in both cases. In the context of the common secondary-structure model, these guanines are conserved between the deoxyribozymes,

with one guanine near the 5′ end, one guanine located between the two conserved stems, and two guanines within the large internal conserved sequence motif. These guanines likely play important catalytic or structural roles in the active structure of these deoxyribozymes.

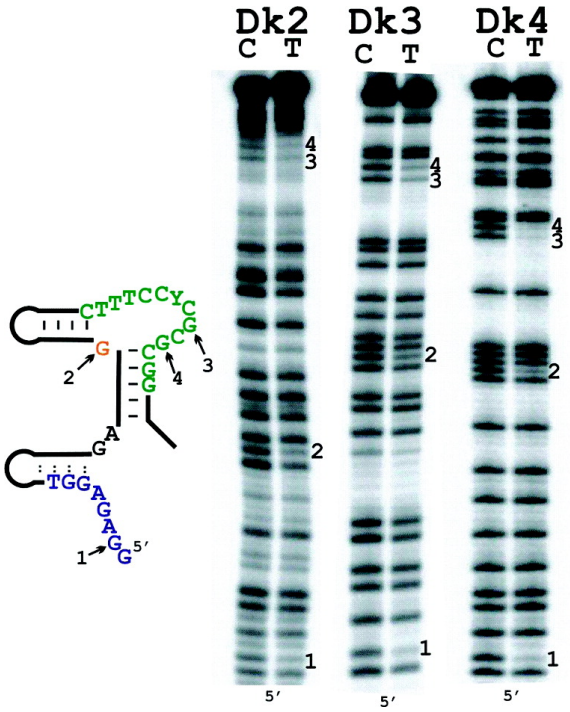


Figure 2.5 Methylation interference of Dk2, Dk3, and Dk4. "C" represents control samples, and "T" represents test samples. A common model for the three deoxyribozymes is shown on the left of the figure with conserved residues shown. Y represents a C in Dk2 and Dk3 and a T in Dk4. Dotted lines indicate base pairs present in Dk3 and Dk4 but not in Dk2. The numbers 1–4 represent guanines that show significant interference and are labeled on the gels for Dk2, Dk3, and Dk4 as well as in the corresponding model.

2.6 DISCUSSION

In this study we have characterized two new deoxyribozyme kinases and compared them to a previously reported deoxyribozyme. We found that they all contained three common sequence motifs as well as identical substrate and divalent metal ion

cofactor requirements. Using all of the structural data obtained for the three deoxyribozymes, we can construct a common secondary-structural model as shown in Figure 2.6. This model is complex and unique in comparison to other studied kinase deoxyribozymes (15, 16). The common structural features are the two helices flanking the 14-nucleotide conserved sequence. An additional stem was found near the 5′ end of Dk3 and Dk4 but not Dk2.

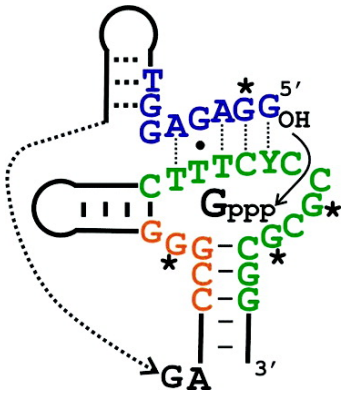


Figure 2.6 Common secondary-structural model for Dk2, Dk3, and Dk4. Thick dotted lines represent base pairs present in Dk3 and Dk4 but not in Dk2. Thin dotted lines (and a dot) represent suspected base pairs that could not be confirmed by mutational analysis. Conserved nucleotides in Dk2, Dk3, and Dk4 are shown in color as per their alignment in Figure 2.1. Y represents a cytosine in Dk2 and Dk3 and a thymine in Dk4. Asterisks represent guanines that showed methylation interference for all three deoxyribozymes. The dotted arrow indicates the chain direction at that location. The substrate GTP is shown as "Gppp".

As shown in the model, it is likely that the 5′ region interacts in some way with the conserved 14-nucleotide internal sequence, simply based on the observation that these two regions are the only two large conserved sequence domains in the deoxyribozymes that are not implicated in the helical interactions. A putative 5-bp stem is proposed as a possible way for these two single-stranded motifs to interact with each other. Although this helical interaction could not be proven using covariation mutational testing

(Supporting Information, Figure 2.2), it is possible that there are helical interactions present, but the identity of these base pairs cannot be changed as the concerned bases may be involved in additional interaction with other residues. This proposed helix could also explain the C to T difference in the internal conserved region between Dk2 (as well as Dk3) and Dk4 (C53/C46 in Dk2 and Dk3, T49 in Dk4), as both C and T residues could bind with a guanine in a Watson–Crick or wobble base pair, respectively.

Another possible scenario for the variance at this specific base site is that this residue may be involved in base-pair-like interaction with the GTP substrate and, when mutated to T, can still form a wobble pair with GTP. If this is the case, it is possible that this mutation might broaden the substrate specificity of the deoxyribozyme to include ATP as a substrate (as the mutated thymine might be able to form a base pair with ATP). We can rule out this possibility because Dk4 contains a T at this site and can only use GTP but not ATP as a substrate (Figure 2.2A). By the same thinking, we noted that another cytosine (C54 in Dk2) in this region was shown to mutate to a thymine during the Dk2 reselection experiment (16). To investigate the substrate specificity of this mutation, we tested a construct of Dk3 containing this mutation (C47 to T in Dk3). As shown in Supporting Information, Figure 2.3, the mutated deoxyribozyme did not show altered specificity, indicating that this cytosine site is not involved in simple base-pair interaction with the substrate.

Even though we were not able to identify specific molecular interactions involved in these two highly conserved regions, it is highly likely that two important motifs are responsible for the binding of the GTP substrate and the catalysis of the phosphate

transfer and may use molecular interactions beyond simple Watson–Crick base pairings. Hopefully, future developments in the high-resolution study of deoxyribozymes (which has so far remained elusive) will shed light on the nature of these unknown interactions.

As mentioned earlier, it is our belief that these three deoxyribozymes containing a similar secondary structure and multiple conserved regions arose independently from the initial DNA library and are not descendants from a common sequence. The main piece of evidence for this is the fact that the internal conserved domains are located at different locations within the deoxyribozymes. Since the in vitro selection strategy used to select these deoxyribozymes included a PAGE-based size-selection step designed to isolate catalysts having a length of ~100 nt and all selected deoxyribozymes indeed contain 98–100 nt (15), it is unlikely that polymerase pausing or jumping could be responsible for the observation of these sequence motifs at different locations as these events would yield sequences of variable lengths that would be lost during the size-exclusion step. The most likely explanation for the repeated observation of these conserved sequence motifs is that they were repeatedly sampled in the initial library or were introduced independently through mutations during PCR. This type of independent isolation of catalytic motifs by nucleic acids has been observed previously. The first reported case is the hammerhead ribozyme motif. The hammerhead ribozyme is a small naturally occurring ribozyme that has been found in a wide range of organisms (22–25). Several groups have isolated this motif repeatedly during independent in vitro selection experiments designed to isolate RNA-cleaving ribozymes (26–28). This is in contrast to other naturally occurring RNA-cleaving ribozymes such as the hairpin ribozyme (29–31), Neurospora VS motif (32), and

the hepatitis delta ribozyme (33), which have not been observed from in vitro selection experiments for RNA-cleaving ribozymes. Another well-documented case is a small RNA-cleaving deoxyribozyme known as "8-17", which has also been isolated several times from independent in vitro selection experiments. 8-17 was initially isolated during an in vitro selection for DNA molecules that can cleave an all-RNA substrate (34). The same catalytic motif has since been identified by several other groups in independent in vitro selection experiments (35–39). As is seen with the common motif in Dk2, Dk3, and Dk4, many of these in vitro selection experiments led to the isolation of many different 8-17 variants that appear to have evolved independently. Also, there are many other efficient deoxyribozymes that cleave RNA, such as 10-23 (34), the bipartite deoxyribozyme (40), and the E2 deoxyribozyme (41), that have not shown repeated occurrences during in vitro selection experiments. It has been suggested that the recurrence of the 8-17 deoxyribozyme indicates that it is the simplest structural solution to the RNA-cleavage reaction for DNA catalysts (28, 38). Similar to 8-17, multiple occurrences of the common deoxyribozyme kinase seem to suggest that it could be the simplest design for the DNA self-phosphorylation reaction. In comparison to the 8-17 motif, however, the kinase motif has some important distinctions. While the common 8-17 secondary structure only contains four absolutely conserved nucleotides and a three-base-pair stem (38), the common kinase motifs share 28 conserved nucleotides (colored nucleotides in Figure 2.6) and two stems containing at least nine base pairs. The need for more conserved bases and structural elements may be due to the increased complexity of phosphorylation as opposed to RNA cleavage, as phosphorylation requires the binding of

an exogenous small-molecule substrate as well as catalysis. The other difference is that while 8-17 has been found to be the dominant deoxyribozyme in several in vitro selection experiments, this kinase motif was found to be less abundant than a faster deoxyribozyme (Dk1) in the original in vitro selection experiment (15). This observation can be rationalized by the fact that Dk1 requires ATP for catalysis instead of GTP. Since both of these substrates were provided during the in vitro selection process, Dk2 (as well as Dk3 and Dk4) did not have to compete with Dk1 because they use different substrates. In fact, the deoxyribozymes containing this common kinase motif were found to be the only catalysts that are dependent on GTP in the selected deoxyribozyme pool, providing more evidence that this structural arrangement is likely the simplest arrangement to carry out the phosphorylation reaction using GTP as a substrate.

2.7 ACKNOWLEDGEMENTS

This work was supported by a research grant from the Canadian Institutes for Health Research. Y.L. is a Canada Research Chair. S.A.M. is an OGSST/David Prosser

2.8 REFERENCES

(1) Joyce, G. F. (2004) Directed evolution of nucleic acid enzymes, Annu. Rev. Biochem. 73, 791–836.

(2) Achenbach, J. C., Chiuman, W., Cruz, R. P., and Li, Y. (2004) DNAzymes: from creation in vitro to application in vivo, Curr. Pharm. Biotechnol. 5, 321–336.

(3) Silverman, S. K. (2005) In vitro selection, characterization, and application of deoxyribozymes that cleave RNA, Nucleic Acids Res. 33, 6151–6163.

(4) Peracchi, A. (2005) DNA catalysis: potential, limitations, open questions, ChemBioChem 6, 1316–1322.

(5) Breaker, R. R., and Joyce, G. F. (1994) A DNA enzyme that cleaves RNA, Chem. Biol. 1, 223–229.

(6) Carmi, N., Shultz, L. A., and Breaker, R. R. (1996) In vitro selection of self-cleaving DNAs, Chem. Biol. 3, 1039–1046.

(7) Li, Y., and Breaker, R. R. (1999) Phosphorylating DNA with DNA, Proc. Natl. Acad. Sci. U.S.A. 96, 2746–2751.

(8) Li, Y., Liu, Y., and Breaker, R. R. (2000) Capping DNA with DNA, Biochemistry 39, 3106–3114.

(9) Cuenoud, B., and Szostak, J. W. (1995) A DNA metalloenzyme with DNA ligase activity, Nature 375, 611–614.

(10) Flynn-Charlebois, A., Wang, Y., Prior, T. K., Rashid, I., Hoadley, K. A., Coppins, R. L., Wolf, A. C., and Silverman, S. K. (2003) Deoxyribozymes with 2′–5′ RNA ligase activity, J. Am. Chem. Soc. 125, 2444–2454.

(11) Wang, Y., and Silverman, S. K. (2003) Deoxyribozymes that synthesize branched and lariat RNA, J. Am. Chem. Soc. 125, 6880–6881.

(12) Sheppard, T. L., Ordoukhanian, P., and Joyce, G. F. (2000) A DNA enzyme with N-glycosylase activity, Proc. Natl. Acad. Sci. U.S.A. 97, 7802–7807.

(13) Chinnapen, D. J., and Sen, D. (2004) A deoxyribozyme that harnesses light to repair thymine dimers in DNA, Proc. Natl. Acad. Sci. U.S.A. 101, 65–69.

(14) Li, Y., and Sen, D. (1996) A catalytic DNA for porphyrin metallation, Nat. Struct. Biol. 3, 743–747.

(15) Wang, W., Billen, L. P., and Li, Y. (2002) Sequence diversity, metal specificity, and catalytic proficiency of metal-dependent phosphorylating DNA enzymes, Chem. Biol. 9, 507–517.

(16) Achenbach, J. C., Jeffries, G. A., McManus, S. A., Billen, L. P., and Li, Y. (2005) Secondary-structure characterization of two proficient kinase deoxyribozymes, Biochemistry 44, 3765–3774.

(17) Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction, Nucleic Acids Res. 31, 3406–3415.

(18) Li, Y., Geyer, C. R., and Sen, D. (1996) Recognition of anionic porphyrins by DNA aptamers, Biochemistry 35, 6911–6922.

(19) Shen, Y., Brennan, J. D., and Li, Y. (2005) Characterizing the secondary structure and identifying functionally essential nucleotides of pH6DZ1, a fluorescence-signaling

and RNA-cleaving deoxyribozyme, Biochemistry 44, 12066–12076.

(20) Kandadai, S. A., and Li, Y. (2005) Characterization of a catalytically efficient acidic RNA-cleaving deoxyribozyme, Nucleic Acids Res. 33, 7164–7175.

(21) Chiuman, W., and Li, Y. (2006) Revitalization of six abandoned catalytic DNA species reveals a common three-way junction framework and diverse catalytic cores, J. Mol. Biol. 357, 748–754.

(22) Forster, A. C., and Symons, R. H. (1987) Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites, Cell 49, 211–220.

(23) Ferbeyre, G., Smith, J. M., and Cedergren, R. (1998) Schistosome satellite DNA encodes active hammerhead ribozymes, Mol. Cell. Biol. 18, 3880–3888.

(24) Rojas, A. A., Vazquez-Tello, A., Ferbeyre, G., Venanzetti, F., Bachmann, L., Paquin, B., Sbordoni, V., and Cedergren, R. (2000) Hammerhead-mediated processing of satellite pDo500 family transcripts from Dolichopoda cave crickets, Nucleic Acids Res. 28, 4037–43.

(25) Pabon-Pena, L. M., Zhang, Y., and Epstein, L. M. (1991) Newt satellite 2 transcripts self-cleave by using an extended hammerhead structure, Mol. Cell. Biol. 11, 6109–6115.

(26) Conaty, J., Hendry, P., and Lockett, T. (1999) Selected classes of minimised hammerhead ribozyme have very high cleavage rates at low Mg2+ concentration, Nucleic Acids Res. 27, 2400–2407.

(27) Tang, J., and Breaker, R. R. (2000) Structural diversity of self-cleaving ribozymes, Proc. Natl. Acad. Sci. U.S.A. 97, 5784–5489.

(28) Salehi-Ashtiani, K., and Szostak, J. W. (2001) In vitro evolution suggests multiple origins for the hammerhead ribozyme, Nature 414, 82–84.

(29) Feldstein, P. A., Buzayan, J. M., and Bruening, G. (1989) Two sequences participating in the autolytic processing of satellite tobacco ringspot virus complementary RNA, Gene 82, 53–61.

(30) Hampel, A., and Tritz, R. (1989) RNA catalytic properties of the minimum (–)sTRSV sequence, Biochemistry 28, 4929–4933.

(31) Haseloff, J., and Gerlach, W. L. (1989) Sequences required for self-catalysed cleavage of the satellite RNA of tobacco ringspot virus, Gene 82, 43–52.

(32) Beattie, T. L., Olive, J. E., and Collins, R. A. (1995) A secondary-structure model

for the self-cleaving region of Neurospora VS RNA, Proc. Natl. Acad. Sci. U.S.A. 92, 4686–4690.

(33) Been, M. D. (1994) Cis- and trans-acting ribozymes from a human pathogen, hepatitis delta virus, Trends Biochem. Sci. 19, 251–256.

(34) Santoro, S. W., and Joyce, G. F. (1997) A general purpose RNA-cleaving DNA enzyme, Proc. Natl. Acad. Sci. U.S.A. 94, 4262–4266.

(35) Faulhammer, D., and Famulok, M. (1996) The Ca2+ ion as a cofactor for a novel RNA-cleaving deoxyribozyme, Angew. Chem., Int. Ed. Engl. 35, 2809–2813.

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

(37) Peracchi, A. (2000) Preferential activation of the 8-17 deoxyribozyme by Ca2+ ions. Evidence for the identity of 8-17 with the catalytic domain of the Mg5 deoxyribozyme, J. Biol. Chem. 275, 11693–11697.

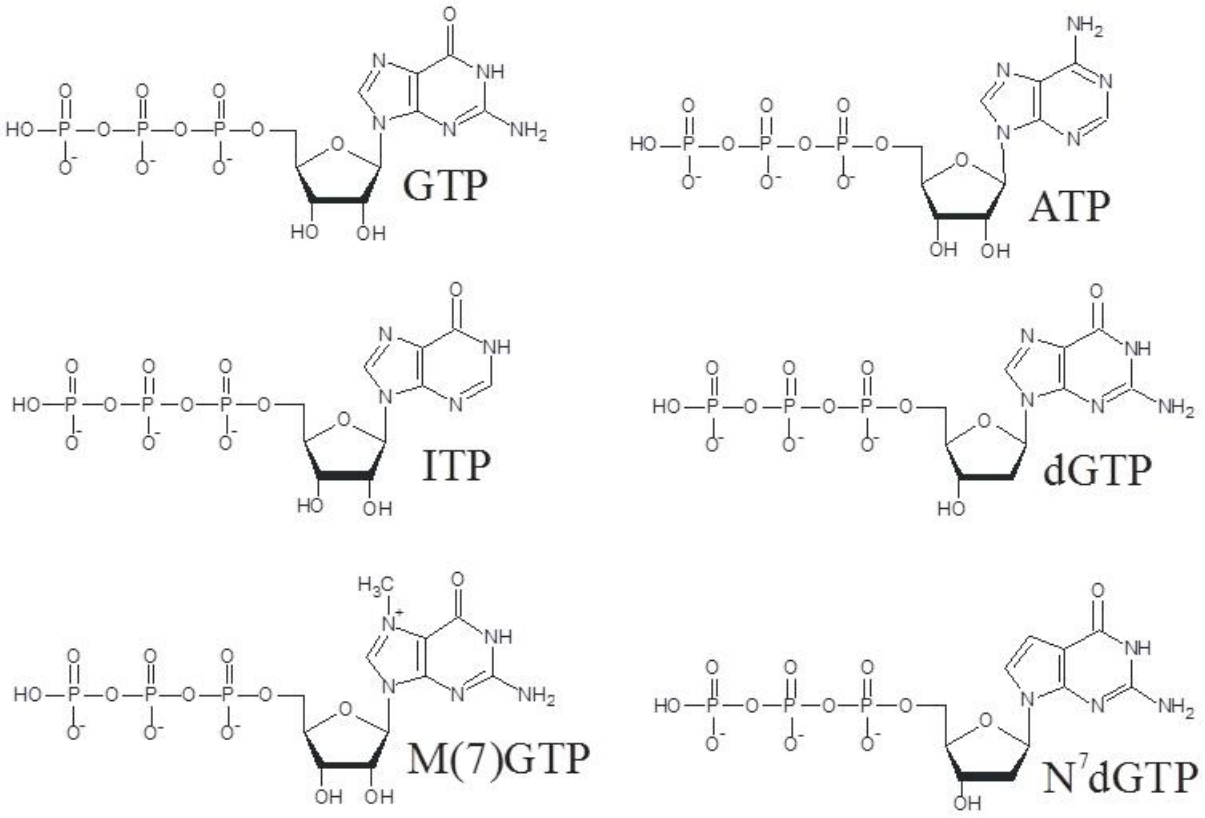
(38) Cruz, R. P., Withers, J. B., and Li, Y. (2004) Dinucleotide junction cleavage versatility of 8-17 deoxyribozyme, Chem. Biol. 11, 57–67.

(39) Schlosser, K., and Li, Y. (2004) Tracing sequence diversity change of RNA-cleaving deoxyribozymes under increasing selection pressure during in vitro selection, Biochemistry 43, 9695–9707.

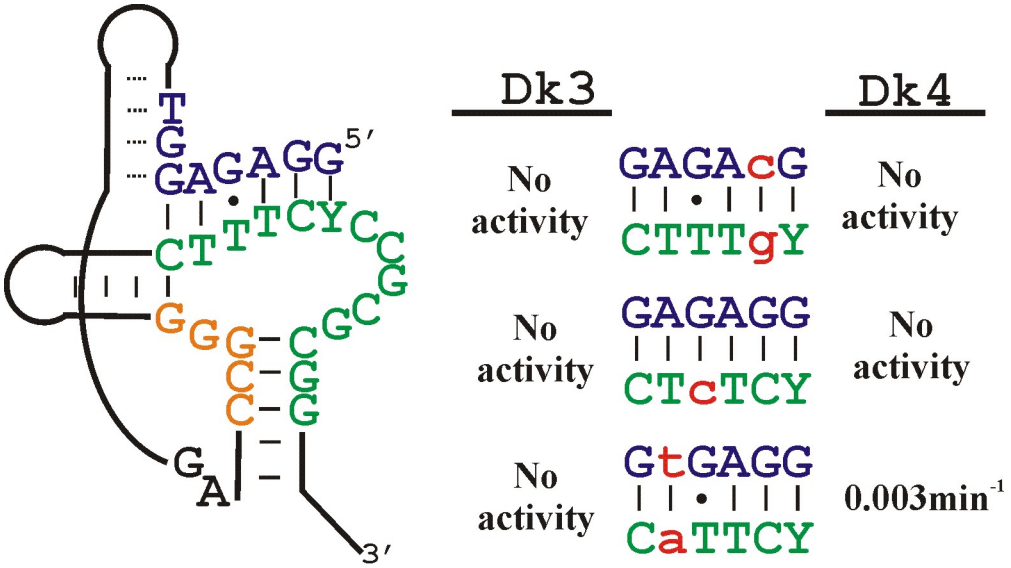
(40) Feldman, A. R., and Sen, D. (2001) A new and efficient DNA enzyme for the sequence-specific cleavage of RNA, J. Mol. Biol. 313, 283–294.

(41) Breaker, R. R., and Joyce, G. F. (1995) A DNA enzyme with Mg2+-dependent RNA phosphoesterase activity, Chem. Biol. 2, 655–660.

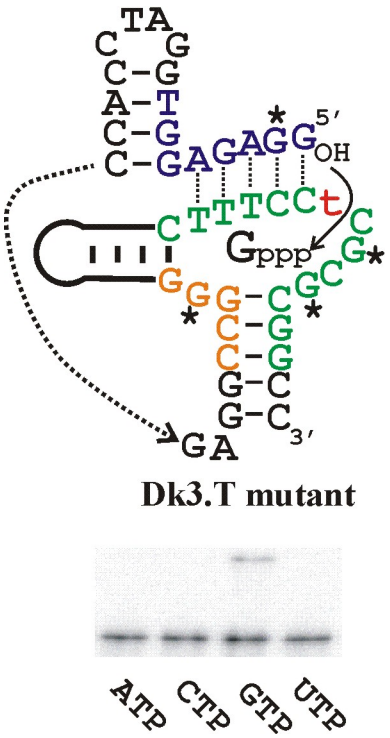
2.9 SUPPLEMENTARY INFORMATION



Supplementary Figure 2.1 Chemical structures of nucleoside 5′-triphosphates tested as potential substrates of Dk2, Dk3 and Dk4.



Supplementary Figure 2.2 Test of putative helical interaction between the 5′ region (blue) and the internal conserved sequence element (green) of Dk2 and Dk3. Constructs tested are shown on the right, with co-variation mutations shown in lowercase red letters. Rates (or no activity) for Dk3 and Dk4 are shown for each construct.



Supplementary Figure 2.3 Substrate specificity of a C-to-T mutant of Dk3. C47T (red t) mutant was tested using 1 mM ATP, CTP, GTP and UTP as shown. This mutant did not show altered specificity, as it was only active in the presence of GTP similar to wild-type Dk3.

Chapter 3:
A Deoxyribozyme with a Novel Guanine Quartet-Helix Pseudoknot Structure

3.1 AUTHOR'S PREFACE

The research project describes the discovery and characterization of a novel deoxyribozyme structure containing a quadruplex-helix pseudoknot. This study was carried using biophysical techniques to establish both secondary and tertiary structure elements of this deoxyribozyme. Secondary structural elements were determined using in vitro reselection and site-specific mutagenesis. Tertiary interactions were investigated using dimethyl sulfate interference assays, circular dichroism and metal ion specificity tests. The major finding from this study was the identity of a novel quadruplex arrangement in the active structure of this deoxyribozyme. This structure is unique both in terms of loop lengths, which are much larger than known quadruplexes, and the fact that is entwined with a base-paired element in a never before seen quadruplex-helix pseudoknot. The discovery of this motif has implications for the future of quadruplex identification and the future of the design of DNA libraries for in vitro selection.

As with the previous chapter, this chapter is a modified version of a published research article. I am the first author on this publication and conducted all of the experiments presented within and wrote the manuscript. I also took a leading role in the conceptualization, experimental design and interpretation of the results obtained. Dr. Yingfu Li is listed as an author of this paper and provided many fruitful suggestions and guidance as well as assisting in editing of the final version of this manuscript. This citation for the original publication is as follows:

McManus, S.A. and Li, Y. (2008) A deoxyribozyme with a novel quartet-helix pseudoknot structure. *Journal of Molecular Biology*, 375, 960-968.

3.2 ABSTRACT

Here we report a deoxyribozyme with a unique structure that contains a two-tiered guanine quadruplex interlinked to a Watson-Crick duplex. Through in vitro selection, sequence mutation, and methylation interference, we show the presence of both the two-tiered guanine quadruplex and two helical regions contained in the active structure of this self-phosphoryloating deoxyribozyme. Interestingly, one GG element of the quadruplex is part of a hairpin loop within one of the identified helical regions. Circular dichroism analysis showed that antiparallel quadruplex formation was dependent on this helix. To our knowledge, this is the first report of a pseudoknot nucleic acid structure that involves a guanine quadruplex. Our finding indicate that guanine quadruplexes can be a part of complex structural arrangements, increasing the likelihood of finding more complex guanine quadruplex arrangements in biological systems.

3.3 INTRODUCTION

Deoxyribozymes are small single stranded DNA molecules capable of performing catalysis. To date, numerous deoxyribozymes have been isolated that catalyze a wide range of chemical reactions.[1; 2; 3; 4] Previous structural studies of deoxyribozymes have revealed numerous interesting structural arrangements, many of which mimic structures observed in DNA or RNA in biological systems. These include different types of helical junctions,[5; 6] triple helix,[7] and guanine quadruplex structures.[3; 8; 9; 10] These findings have demonstrated that it is possible to reveal novel nucleic acid structures that are biologically relevant by studying deoxyribozymes isolated by in vitro selection. We use a family of self-phosphorylating deoxyribozymes (or kinase deoxyribozymes) as a model system to study the capacity of single-stranded DNA to form diverse and complex structures.[11; 12; 13] Although high resolution study of deoxyribozymes has remained elusive, we can use a series of biochemical experiments to deduce many structural elements of a deoxyribozyme to create a detailed structural model. Here we report the characterization of a self-phosphorylating DNAzyme that reveals a new structural arrangement containing a helical element interacting with a two-tiered guanine quadruplex.

Guanine-rich DNA sequences are widely known for their ability to form a unique form of DNA structure known as guanine-quartets (G-quartets) in which (1) four guanines interact with each other through cyclic Hoogsteen base-pairs and (2) two (or more) consecutive tiers of these hydrogen-bonded guanines are stacked on top of each other forming complexes called quadruplexes.[14; 15] G-quadruplex-forming sequences are found in the telomere of many organisms[16] (a region that is of great interest because its

maintenance has been linked to cell immortality and cancer[17]) as well as in promoters of some oncogenes.[18; 19; 20; 21] Because of their uniqueness and biological relevance, G-quadruplexes have been suggested as potential drug targets.[22]

Various forms of G-quadruplexes have been observed. Some DNA sequences have been shown to fold into a four-stranded structure called G4 DNA.[23] Dimeric G-quadruplex complexes are also known where two strands containing folded-back Hoogsteen-paired guanines hybridize to complete the quadruplex.[24] Unimolecular guanine quadruplexes have also been observed in which three hairpin loops are present between four strings of quadruplex-forming guanines within a single strand of DNA.[15; 25] Examinations of the loop regions of these intrastrand G-quadruplexes have shown that the sequence and length of connecting loops can be important for quadruplex formation.[26; 27] What has not been observed to date is whether the non-G-quadruplex elements in a G-quadruplex structure can engage in Watson-Crick interactions with other loops or surrounding sequence to aid in the formation and/or stability of a G-quadruplex structure. A demonstration of such a structural arrangement could lead to the discovery of new biologically relevant G-quadruplexes by expanding the parameters used to predict which genomic sequences have the potential to form guanine quadruplexes. In this report we show that the active structure of a nucleic acid enzyme (more specifically an artificial deoxyribozyme) contains two-tiered intrastrand G-quadruplexes with one pair of guanines located within the loop of a Watson-Crick stem, creating a novel pseudoknot interaction. We also demonstrate that formation of the quartet is dependent on this helical interaction.

3.4 RESULTS

3.4.1 Substrate and cofactor requirements for Dk5

The deoxyribozyme under examination in this report is Dk5, a previously isolated, but not characterized, 98-nucleotide (nt) catalytic DNA that catalyzes the transfer of γ-phosphate from a nucleoside 5′-triphosphate to the 5′-hydroxyl group of its own sequence (Figure 3.1A). Dk5 is the fifth in a series of efficient kinase deoxyribozymes isolated by an in vitro selection experiment.[11] Substrate characterization via an indirect ligation assay (Figure 3.1B) revealed that Dk5 can use ATP, GTP, dATP, dGTP or ITP as a source of phosphate, but has little activity in the presence of any of the tested pyrimidine triphosphates (Figure 3.1C). This indicates that the purine ring of the substrate is involved in binding to the deoxyribozyme. In terms of metal ion cofactor requirements, Dk5 can use any of the tested divalent metal ions (such as Ca^{2+}, Cu^{2+}, Mg^{2+}, or Mn^{2+}) as a divalent metal ion cofactor (Figure 3.1D). This finding was intriguing as the other four reported kinase deoxyribozymes (Dk1-Dk4), were all shown to be highly specific in terms of cofactor requirements.[12; 13] This suggests that perhaps Dk5 uses a distinct structural arrangement when performing catalysis. Dk5's requirements for monovalent metal ion are analyzed in a separate section.

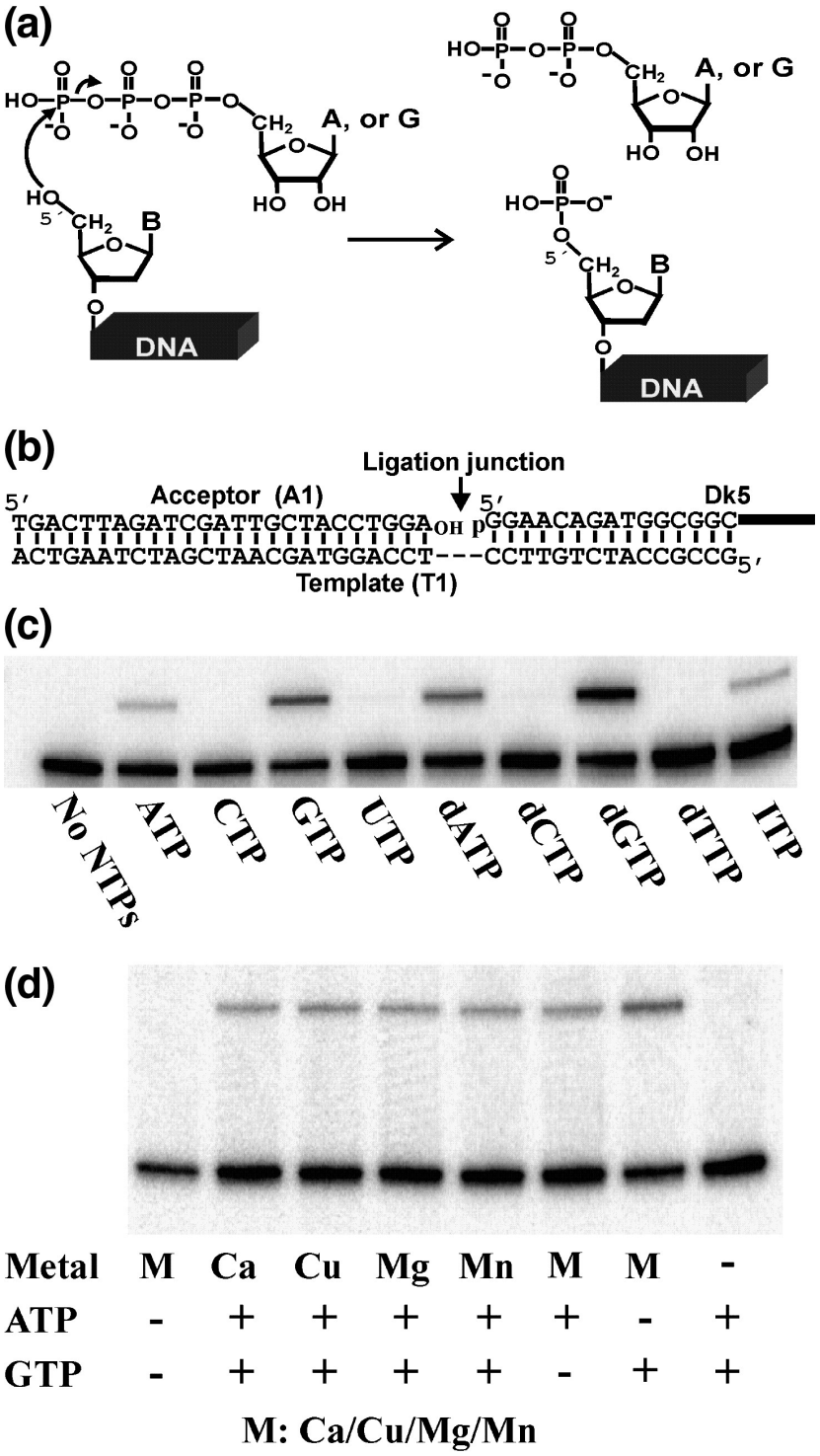

(a)
A, or G
DNA
A, or G
DNA
(b)
Ligation junction
5′
Acceptor (A1)
5′
Dk5
TGACTTAGATCGATTGCTACCTGGA OH pGGAACAGATGGCGGC
ACTGAATCTAGCTAACGATGGACCT---CCTTGTCTACCGCCG 5′
Template (T1)
(c)
No NTPs
ATP
CTP
GTP
UTP
dATP
dCTP
dGTP
dTTP
ITP
(d)
Metal M Ca Cu Mg Mn M M -
ATP - + + + + + - +
GTP - + + + + - + +
M: Ca/Cu/Mg/Mn

Figure 3.1 Catalytic reaction and Characterization of Dk5. (A) proposed reaction mechanism for self-phosphorylation of Dk5. Nucleophilic attack by the 5′ hydroxyl of Dk5 on the γ-phosphate of ATP or GTP results in 5′-phosphorylated Dk5 and an NDP molecule. (b) Indirect ligation assay used for detection of kinase deoxyribozyme activity. After incubation with ATP and GTP, Dk5 is incubated with an acceptor DNA molecule and a template DNA complementary to the 5′ end of Dk5 and the acceptor. Buffer and T4 DNA ligase are added. Only molecules with a 5′-phosphate (meaning they have undergone self-phosphorylation) can be ligated by T4 DNA ligase. The ligated DNA can be resolved on a 10% PAGE gel. (C) Dk5 was tested using five nucleotide triphosphates and four nucleotide phosphates as shown. (D) Dk5 was tested using different divalent metal ions as cofactors. The top bands represent the ligated molecules as described above.

3.4.2 Generation of catalytically active mutants of Dk5 through reselection *in vitro*

To assess the importance of each base within the sequence of Dk5, we generated active Dk5 mutants from a partially randomized Dk5 pool using a previously described in vitro selection protocol.[12]. This process of in vitro 'reselection' can be used to determine the essentiality of particular bases within the deoxyribozyme sequence by observing if this base is changed in active sequence variants found at the end of the selection. Through analysis of 28 unique sequences identified from the fifth round of this selection (shown in Supplementary Figure 3.1), we were able to look at the observed mutation rate at each base site and compare it to the expected mutation rate if there was no specificity for a particular base at that site (in our case 24% mutation rate was introduced during the chemical synthesis of the partially randomized Dk5 library, translating to ~7 mutations per position expected in 28 isolated clones). In our analysis we were able to identify many invariable nucleotides (no mutations seen at this base site; shown as letters in black circles in Figure 3.2), highly conserved nucleotides (with 3 or less mutations; shown as letters in white circles), and nucleotides that are prone to mutation (4 or more mutations; shown as non-circled letters). Inspection of the nucleotide conservation pattern reveals

two important points. First, the central region from G_{20} to A_{70} as well as G_1-G_{10} motif at the 5′ end must play important roles in the catalysis or the folding of Dk5 because almost all invariable and highly conserved nucleotides are located within these two sequence elements. Second, the remaining nucleotides may not be crucial to the catalytic function of Dk5, with the possible exception of G_{82}, and to a lesser extent, C_{15}. It is noteworthy that 17 of the 23 absolutely conserved nucleotides were deoxyguanine residues, and several are clustered into groups of two or more. This pattern has been seen in many oligonucleotides containing G-quadruplexes and suggests that the active structure of Dk5 may utilize a variation of G-quadruplex structure. This point will be investigated further later in this paper.

As well as assessing base essentiality, reselection data are also useful in identification and confirmation of helical regions present in the active structure of the deoxyribozyme. Using *mfold* software[28] we predicted that Dk5 had three stem-loops P1-P3 illustrated in Figure 3.3. The observation of many base co-variations in P2 supported the existence of this 3-bp stem. The existence of the larger, 8-bp stem P3 is also consistent with the G_{71}-to-T_{71} mutation found in some of mutant Dk5 sequences that corrected the G_{71}•A_{89} mismatch to a T-A base-pair (it should be noted that the 15 nucleotides at the 3′ end, shown as italicized letters, were used as the primer binding site and not subjected to mutations). The other helical region predicted by *mfold*, the 4-bp P1, is neither supported nor refuted by the reselection data as no base co-variation was observed.

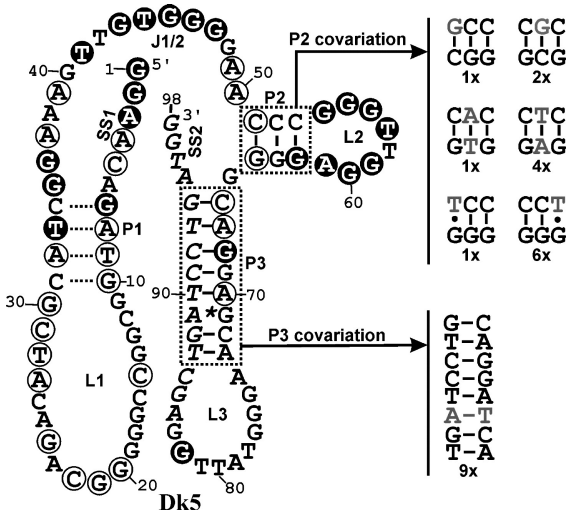


Figure 3.2 Secondary structural model for Dk5 based on *mfold* prediction and in vitro selection data. Co-variations seen in reselection are shown on left. Asterisk represents a mismatch that was observed to be mutated into a base pair during reselection. Dotted base-pair lines indicate base-pairs that were predicted by *mfold* but did not show co-variation during reselection. Letters in black circles represent bases that were found to be invariable during reselection. Letters in hollow circles represent bases that were highly conserved during reselection. P represents a putative stem, L a loop, SS a single stranded region, and J a junction between stems.

3.4.3 Using sequence alteration to probe the secondary structure of Dk5

The reselection data confirmed the central stem P2, but did not show whether or not P1 or P3 were present in the active structure. The level of variability in certain putative loop regions also suggested that some nucleotides in these regions may be removable without loss of activity. Several synthetic deoxyribozyme constructs, some of which are shown in Figure 3.3, were therefore made to provide additional evidence. These mutant deoxyribozymes were assessed for both catalytic rate constant (k_{obs}) and maximal reaction yield (Y_{max}) after two hours as indicators of their catalytic activity (the original data are provided in Supplementary Figure 3.2).

We first produced two constructs to examine the importance of nucleotides in the proposed loop L1. The fact that the many nucleotides in this region are prone to mutation is a strong indication that some of the nucleotides in L1 play an insignificant role in the catalytic activity of Dk5. When 13 mostly variable nucleotides were removed from L1 (Dk5.1a, Figure 3.3), the k_{obs} value was only reduced by half while the Y_{max} value remained unchanged, suggesting that these nucleotides do not contribute significantly to the function of the deoxyribozyme. Further truncation of the remaining seven nucleotides in L1 into a tetraloop (Dk5.1b) resulted in a construct that was still active but with a reaction yield reduced from 88% to 7%. This result appears to suggest that this sequence element, which contains five highly conserved nucleotides, is required for proper folding of Dk5.

Three constructs were made to examine the existence of P1. In the first two constructs, Dk5.2a and Dk5.2b, the content of P1 was altered while Watson-Crick interactions were maintained. The third construct, Dk5.2c, contained an A•A mismatch within the P1 region. While Dk5.2a and Dk5.2b were completely inactive, Dk5.2c exhibited an activity that was similar to the original Dk5. These results indicate P1 either does not exist in the active structure or exists but is dispensable, which is consistent with the reselection results.

According to the reselection data, all but one nucleotide (G_{82}) of L3 were highly tolerant of mutation. Therefore, it is very likely that L3 is a dispensable element. As expected, the shortened deoxyribozyme Dk5.3a in which the original L3 was replaced with an arbitrarily chosen tetraloop (Dk5.3a) was fully active. The complete

dispensability of L3 was finally confirmed by the strong activity and high reaction yield of the bimolecular deoxyribozyme construct Dk5.4a in which L3 was completely eliminated. The strong activity of Dk5.4a also indicates that the role of P3 is purely structural because the base-pairing content of this construct was arbitrarily altered from the original P3. Also, the bimolecular system was found to be able to tolerate one (Dk5.4b) but not two mismatching base pairs (Dk5.4c) without significant loss of activity. This is not surprising as the wild-type Dk5 contains one mismatch in P3. We also made a *cis*-acting construct, Dk5.3c, in which the entire P3-L3 region was reduced into two nucleotides (G and C). This mutant had severely reduced activity (10-fold decrease in k_{obs} and 6-fold reduction in reaction yield), pointing to the importance (but non-essentiality) of P3 to the catalytic function of the deoxyribozyme. Taken together, our data support the following conclusions regarding P3 and L3: (1) P3 does exist and plays an important structural role; (2) the base-pair content of P3 is not important; (3) a stable P3 is required for the optimal activity of the deoxyribozyme; (4) L3 is completely expendable.

Interestingly, the truncation of four single-stranded nucleotides at the 3′ end of the deoxyribozyme (as the SS2 element) resulted in very significant loss of activity (Dk5.3b). This finding reveals that Dk5 has incorporated these four terminal nucleotides (which are part of the 3′ primer-binding site) into its active structure.

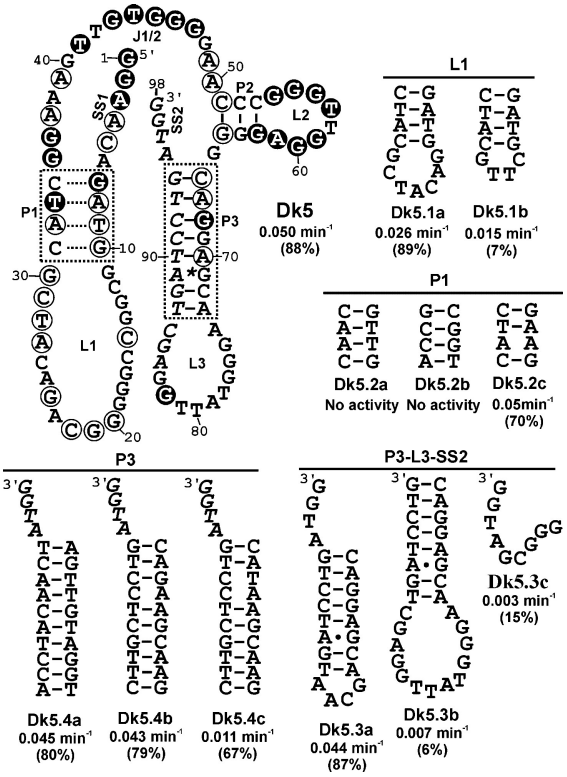


Figure 3.3 Sequence alteration experiments. Labeling of Dk5 is the same as in Figure 3.2. The rate constants are shown beneath each construct, as well as the reaction yield in parentheses.

3.4.4 Identification of guanine residues that are intolerant of methylation at N7

SS1, J1/2 and L2 within the proposed secondary structure of Dk5 (Figures 3.2 and 3.3) are the home of many highly conserved nucleotides. These sequence elements contain a total of 30 nucleotides, 18 of which are absolutely conserved and 6 of which are highly conserved. Many of the conserved nucleotides were guanine residues and often occur as a group of 2-4 consecutive Gs that is often indicative of a G-quadruplex structure. We therefore sought to use a methylation interference technique to probe the possibility of the existence of such a structure in Dk5.

It is widely known that dimethyl sulfate (DMS) can be used to methylate guanine residues within a nucleic acid sequence specifically at their N7 position. Such methylated guanosine nucleotides are prone to piperidine-assisted cleavage, leading to cleavage products that can easily be resolved on a sequencing gel as previously described.[12] Since the N7 atoms of guanines within a guanine quadruplex act as hydrogen bond acceptors, guanine residues participating in the formation of a catalytically essential G-quadruplex structure cannot tolerate methylation. It is also noteworthy that the N7 position is also involved in other interactions such as triple helices, but not in Watson-Crick base pairs.

Specifically in our setup, Dk5 was treated under such conditions that one guanine per DNA sequence was methylated. This 'pool' of methylated Dk5 molecules was then tested for self-phosphorylation and active molecules isolated through the indirect ligation assay shown in Figure 3.1B. The active molecules were then 5′ terminally labeled with ^{32}P and treated with piperidine in basic conditions which leads to depurination of methylated guanine residues and strand scission. When run on a 10% PAGE gel a ladder of DNA fragments will be seen corresponding in size to fragments containing the ligated oligonucleotide and the 5' part of the DNAzyme up to the site of cleavage at a methylated guanine residue. This can be compared to a control lane in which self-phosphorylation was done before methylation to identify the level to which methylation at a particular guanine interferes with the reaction rate of the DNAzyme.

The above experiment led to the identification of eight guanines that were completely intolerant of methylation (Figure 3.4). Interestingly, these guanines occurred as four pairs (G_1G_2, $G_{35}G_{36}$, $G_{45}G_{46}$ and $G_{59}G_{60}$), pointing to a possibility that Dk5 may

adopt a two-tiered G-quadruplex as an essential structural element. Even more interesting is the observation that one pair of guanines, G59G60, is located within the loop of the confirmed P2 stem. If the formation of G-quartet could be confirmed, this would represent the first observation of a special pseudoknot interaction that involves a guanine quadruplex and a helix.

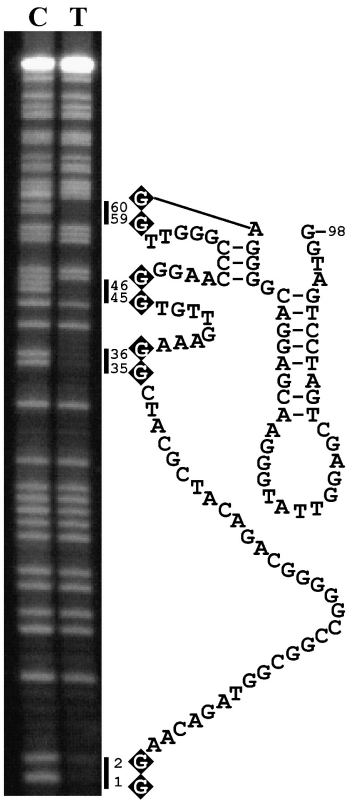


Figure 3.4 Methylation interference of Dk5. (A) 10% PAGE gel of Dk5 after methylation interference and piperidine cleavage. Four pairs of guanines showing significant interference when methylated are highlighted along with their base number. C refers to control samples in which self-phosphorylation reactions were carried out before methylation of DNA with DMS. T refers to test samples in which self-phosphorylation reactions were carried out after methylation.

3.4.5 Investigation of G-quadruplex structure and pseudoknot interaction by circular dichroism

The technique of circular dichroism (CD) was employed to provide evidence for the existence of the putative G-quadruplex element in the structure of Dk5 as well as for the analysis of whether the G-quadruplex adopted a parallel or anti-parallel sequence orientation.[29] CD measures the difference of a molecule in its absorption of left and right plane-polarized light. It has been demonstrated that parallel and antiparallel guanine quadruplexes give distinct CD spectra.[30] A parallel quadruplex exhibits a positive peak at 260 nm and a negative peak at 240 nm, while antiparallel guanine quadruplexes are characterized by a positive peak at 295 nm.

The CD spectrum of Dk5 is shown in Figure 3.5A (dark blue thick line). It contains a positive peak at 295 nm suggesting the presence of an antiparallel guanine quadruplex element within its structure. In this spectrum the 295-nm peak is extended further towards lower wavelengths than is usually seen with pure antiparallel quadruplex structures. This could be due to the influence of a competing four-stranded parallel structure, due to the fact that there are several runs of guanines within the sequence of Dk5. Similar spectra have been observed previously with other DNA oligonucleotides that can fold into a mixture of parallel and antiparallel guanine quadruplexes,[18; 30] with the 240 nm and 260 nm peaks appearing due to the parallel guanine quartets, and the shoulder being contributed by the 295 nm signal observed in antiparallel quadruplexes. The base composition of Dk5 suggests that Dk5 can fold into two distinct populations. One population will contain molecules in their unimolecular catalytically active

conformation with a two-tiered guanine quadruplex. A second structural population is also likely to form due to the fact that Dk5 contains many stretches of multiple guanines (up to 5 guanines in a row). Sequences of this type are known to form four stranded complexes containing parallel guanine quadruplexes[23; 31].

To test the conformation of the active structure, four constructs were made in which the four pairs of guanines thought to be involved in quadruplexes in the active structure were mutated to thymines (g1tg2t, g35tg36t, g45tg46t, g59tg60t). These sequences should not be able to form unimolecular guanine quaduplexes and should only produce the four-stranded parallel quadruplex structures. The spectrum of these sequences contained strong positive peaks at 260 nm (Figure 3.5A) indicative of parallel G-quadruplex formation. They did not contain the shoulder peak of higher wavelength as seen with Dk5, suggesting that antiparallel quadruplexes are not present within these constructs. As a negative control a construct was tested in which a pair of guanines in a loop region in the putative active structure were mutated to thymines (g10tg11t). This sequence showed a spectrum similar to the wild-type sequence with a strong positive peak at 295 indicative of antiparallel quadruplex formation. The fact that antiparallel quadruplexes are only seen in the constructs capable of folding into the active structure of Dk5 indicates that Dk5's active structure contains antiparallel quadruplexes.

Having examined the orientation of the two-tiered quadruplex in the active structure of Dk5, we sought to use CD to investigate whether the entwined helix is essential for intramolecular antiparallel quadruplex formation. To accomplish this, a construct was made with three cytosines within P2 were mutated to adenines

(c51ac52ac53a), thereby eliminating all base pairs in P2. This sequence showed a negative peak at 240 nm and a positive peak at 260 nm (Figure 3.5A), similar to the four previously scanned inactive mutants. This indicates that this construct cannot form unimolecular antiparallel guanine quadruplexes and only forms four stranded parallel quadruplex structures. This demonstrates that P2 is essential for antiparallel guanine quadruplex formation in the active structure of Dk5.

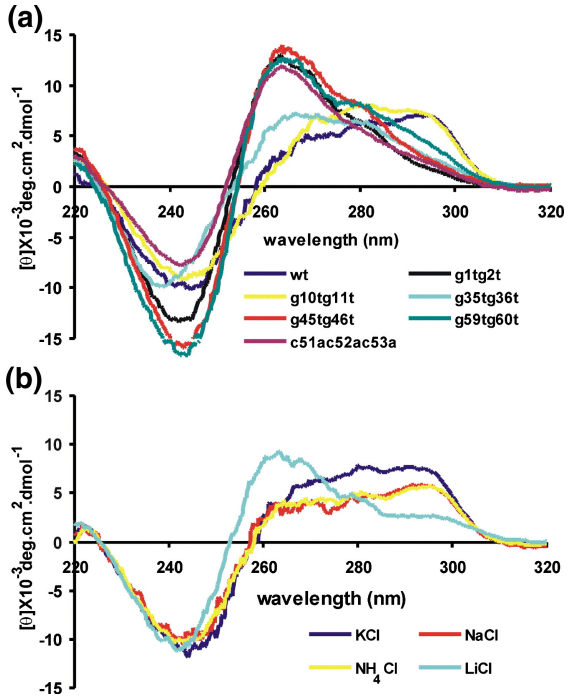


Figure 3.5 CD analysis of Dk5. (A) CD analysis of Dk5 and several sequence variants. (b) CD analysis of Dk5 using buffers containing different monovalent metal ions.

3.4.6 Effects of monovalent metal ions on the activity and CD spectrum of Dk5

As further verification of the formation of antiparallel quadruplexes in its active structure, Dk5's activity was tested in buffers containing different monovalent metal ions. The effect of monovalent metal ions on guanine quadruplex formation has been extensively studied. A particular study by Sen and Gilbert specifically looked at the effect

of different monovalent metal ions on the formation of parallel versus antiparallel guanine quadruplex structures in telomeric DNA.[31] They found that in the presence of potassium ions the G-rich sequences tended to form antiparallel quadruplex structures at the expense of the four-stranded parallel G4 DNA seen when the DNA is incubated in sodium buffer. We tested Dk5 in buffers containing potassium, sodium, ammonium and lithium as a monovalent metal ion cofactor. The rate constants and percent yields of Dk5 in the different buffer conditions are shown in Table 3.1 (the original data are provided in Supplementary Figure 3.3). Looking at the reaction yields, Dk5 has a high yield in potassium buffer indicating that it is folded correctly into its catalytically active structure. In contrast, when reacted in sodium or ammonium buffers, Dk5 has a much lower reaction yield. This is indicative of a high amount of misfolding into inactive structures in these solutions. This is consistent with the active structure containing antiparallel guanine quadruplexes, which are supported well by potassium but to a lesser extent by sodium or ammonium. Dk5 incubated in lithium was found to be inactive. This is not surprising, as lithium is known not to support guanine quadruplex formation.[15; 26] It is noteworthy that the rate constant in sodium is slightly higher than seen with potassium. Since they were both present in the selection buffer it is possible that sodium has an important role that is not completely compensated for by potassium. To further investigate the effects of monovalent metal ions on the structure of Dk5, the CD spectrum of Dk5 was tested in buffers containing the above-mentioned monovalent metal ions with the results shown in Figure 3.5B. As shown in the Figure 3.5B, the spectrum in potassium is similar to that seen in wild-type buffer. A similar spectrum was seen with

sodium and ammonium except with a 295-nm peak of lower intensity. In lithium-containing buffer, the spectrum lacked a positive peak at 295 nm, which is similar to what was observed with the inactive mutants. These results give further support to antiparallel quadruplex formation in potassium, sodium and ammonium but not lithium. The lower intensity of the 295-nm peak in sodium and ammonium also validates the speculation that the lower level of reaction yield in these buffers is due to a lower percentage of the deoxyribozyme being folded in the correct structure that contains a two-tiered antiparallel guanine quadruplex.

Table 3.1 Catalytic activity of Dk5 in the presence of different monovalent metal ions

Monovalent Ion	Rate constant (min^{-1})	Yield (%)
Na^+/K^+ (400 mM/100 mM)	0.050	88
K^+ (500 mM)	0.022	71
Na^+ (500 mM)	0.066	27
NH_4^+ (500 mM)	0.011	33
Li^+ (500 mM)	No activity	Not applicable

3.5 DISCUSSION

We set out to characterize the kinase deoxyribozyme Dk5. What we found was a deoxyribozyme with an active structure that contained an interlinked guanine quadruplex-helix structure that has not been observed previously for any functional nucleic acids. The helical interactions within these models were shown through reselection and sequence alteration experiments. The antiparallel guanine quadruplex was supported

through methylation interference, circular dichroism, and metal ion dependency experiments. The secondary structure was shown to contain two helices of three and eight base pairs. Three possible structural models are shown in Figure 3.6. Antiparallel quadruplexes have several different orientations including alternating polarities (Figure 3.6A), two strands on the same face being parallel but antiparallel to the two strands on the opposing face (Figure 3.6B), and one strand with one polarity and three with the opposite polarity (Figure 3.6C). There are also three known loop types that are possible (edgewise, diagonal and double chain reversal) that can be present in different combinations, adding to the number of different structural possibilities (three different loop combinations are shown in Figure 3.6). Without high-resolution studies (which to date have not been amenable to deoxyribozymes) it is not possible to determine the exact orientation of the two-tiered quadruplex within the active structure of Dk5. However, the structural data obtained here does allow for us to make some important conclusions about the structure of Dk5.

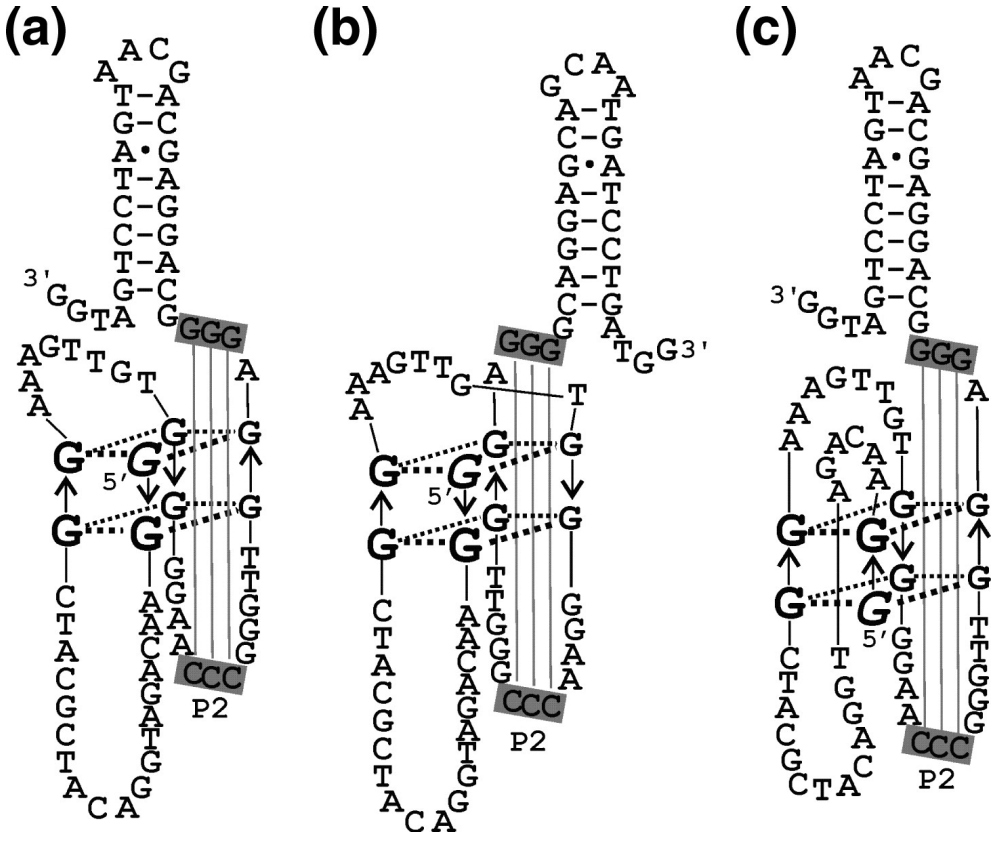


Figure 3.6 Structural models for minimized Dk5. Interactions within two-tiered G-quartet are represented by dotted lines. P2 is outlined in grey. Loop conformations are: (A) three edgewise loops; (B) edgewise loop, diagonal loop, edgewise loop; (C) double chain reversal loop, edgewise loop, edgewise loop.

The Dk5 structure differs from all known G-quadruplex structures in two major ways. First, it has larger loop sequences than commonly observed in quadruplex structures. This is likely due to unidentified tertiary interactions within or between these loops, as reselection showed the sequences of these loops could not be significantly altered. Second, one of its loops is involved in a helical interaction with another segment of adjacent sequence, and this helical interaction is essential for G-quardruplex formation.

Finally, our findings above also have implications on the prediction of G-quadruplex structures within the genomes of different organisms because it shows that G-quadruplex based structures can be more complex than previously known. In studies to date, bioinformatic analyses have been performed by searching sequences with multiple guanines with short intervening loops [32; 33]. The fact that a unimolecular G-quadruplex can involve longer loops suggests that perhaps looking only at regions with short sequences between runs of guanines could potentially miss some G-quadruplex forming sequences. Also, since we have shown a loop region to be involved in a helical interaction that is essential for G-quadruplex formation, it may be possible to design future algorithms to identify low-energy structures containing this type of entwined G-quadruplex/helix motifs from sequenced genomes.

3.6 MATERIAL AND METHODS

3.6.1 Oligonucleotides and Materials

DNA oligonucleotides were prepared using standard phosphoramidite chemistry (HHMI-Keck Biotechnology Resource Laboratory, Yale University; Mobix Lab, McMaster University; Integrated DNA Technologies, Coralville, IA). The partially randomized library for DK5 reselection was synthesized using the following mixtures: 76% wild-type nucleotide and 8% of each of the remaining three nucleotides. Each oligonucleotide was purified by 10% denaturing (8 M urea) polyacrylamide gel electrophoresis (PAGE) and its concentration was determined spectroscopically. T4 polynucleotide kinase (PNK), T4 DNA ligase and calf intestinal alkaline phosphatase

(CIAP) were purchased from MBI fermentas. [γ-^{32}P]ATP and [α-^{32}P]dGTP were acquired from GE Healthcare. All other chemicals were obtained from Sigma.

3.6.2 Buffer Conditions

For the reselection, sequence alteration, methylation interference and circular dichroism experiments, the self-phosphorylation reactions were performed under the following conditions; 400mM NaCl, 100mM KCl, 5mM $CaCl_2$, 5mM $MgCl_2$, 5mM $MnCl_2$, 25mM HEPES pH=7. Each reaction was initiated by adding ATP and GTP to a final concentration of 1mM each. For the divalent metal ion tests, 10mM of the specified divalent metal ion was used as the only divalent metal ion in the reaction, with the monovalent metal ion concentrations remaining the same as listed above. 1mM of each NTP or dNTP was used in the substrate-specificity tests. For the monovalent metal ion tests, the monovalent ion components of the buffer were replaced with 500mM of the selected monovalent metal ion.

3.6.3 Reselection procedures

The reselection procedures were adapted from a previously published method that has been described in detail [11]. The reselection library was made as per a previously reported protocol [3] with a 24% mutation rate at each base position except for 15 nucleotides (nt) at the 3′ end that were kept constant to act as a primer binding site during PCR. The sequences of the template and acceptor oligonucleotides for the ligation step were 5′GCCGCCATCTGTTCCTCCAGGTAGCAATC GATCTAAGTCA and 5′TGACTTAGATCGATTGCTACCTGGA, respectively. The underlined nucleotides were partially mutated with a 24% mutation rate to provide some variations to

accommodate mutations in the donor sequence. For PCR, a ribo-terminated primer (5′TGACTTAGATCGATTGCTACCTGGAr) as well as an all-DNA primer (5′CCATCAGGATCAGCT) were used. In total, four rounds of in vitro reselection were performed. The sequences from the fourth round were used for cloning and sequencing.

3.6.4 Cloning and sequencing of DNA populations

DNA sequences after the 4th round were amplified by PCR and ligated into a pUC57 derivative using the T/A cloning method. Plasmids containing individual catalysts were isolated using the 96-well PerfectPrep kit (Eppendorf) as per the manufacture's instructions. Inserts were unidirectionally sequenced using the -47 primer on a CEQ2000XL DNA sequencer (Beckman-Coulter) using CEQ Quick Start Kit (Beckman-Coulter) as per the manufacture's instructions.

3.6.5 Catalytic assays

The protocol for the catalytic assays used for Dk5 has been described in detail previously [13]. Briefly, Dk5 constructs used in secondary structural and monovalent metal ion studies were assembled by ligation of two oligonucleotides with the donor oligonucleotide labeled with a ^{32}P at its 5′ end.

Deoxyribozyme constructs were tested for phosphorylation using an indirect ligation method. The rate constants were determined using Graphpad prism software by plotting percent ligation versus reaction time. All reported rates are the average of at least two independent time course experiments.

3.6.6 Methylation interference assays

DMS Methylation interference assays were performed on Dk5 as previously described [13] . Briefly, a 'test' sample in which methylated Dk5 was put under phosphorylation conditions, was compared to a 'control' sample which was methylated after phosphorylation. After piperidine cleavage and resolution on 10% PAGE, guanines that interfere with activity when methylated can be identified through absence of bands in the test lane versus the control lane.

3.6.7 Circular dichroism (CD) analysis

CD studies were carried out using a JASCO J-600 instrument as per the manufacturer's instructions. All constructs were scanned from 320 nm to 220 nm at 10 μM DNA concentration in a 0.1 cm quartz cuvette. All DNA samples were incubated for 30 min at room temperature in the 1× selection buffer before scanning. Each curve in Figure 3.4 is the average of ten scans.

3.7 ACKNOLWEDGEMENTS

This work was supported by research grants from the Canadian Institutes for Health Research, the Canada Foundation for Innovation and the Ontario Innovation Trust. We thank the members of our lab for comments on the manuscript. S.A.M. is an OGSST/David Prosser Scholarship recipient. Y.L. is a Canada Research Chair.

3.8 REFERENCES

1. Joyce, G. F. (2004). Directed evolution of nucleic acid enzymes. *Annu Rev Biochem* 73, 791-836.

2. Peracchi, A. (2005). DNA catalysis: potential, limitations, open questions. *Chembiochem* 6, 1316-22.

3. Li, Y. & Breaker, R. R. (1999). Deoxyribozymes: new players in the ancient game of biocatalysis. *Curr Opin Struct Biol* 9, 315-23.

4. Achenbach, J. C., Chiuman, W., Cruz, R. P. & Li, Y. (2004). DNAzymes: from creation in vitro to application in vivo. *Curr Pharm Biotechnol* 5, 321-36.

5. Santoro, S. W. & Joyce, G. F. (1997). A general purpose RNA-cleaving DNA enzyme. *Proc Natl Acad Sci U S A* 94, 4262-6.

6. Cuenoud, B. & Szostak, J. W. (1995). A DNA metalloenzyme with DNA ligase activity. *Nature* 375, 611-4.

7. Carmi, N., Balkhi, S. R. & Breaker, R. R. (1998). Cleaving DNA with DNA. *Proc Natl Acad Sci U S A* 95, 2233-7.

8. Li, Y. & Sen, D. (1997). Toward an efficient DNAzyme. *Biochemistry* 36, 5589-99.

9. Chinnapen, D. J. & Sen, D. (2004). A deoxyribozyme that harnesses light to repair thymine dimers in DNA. *Proc Natl Acad Sci U S A* 101, 65-9.

10. Li, Y., Liu, Y. & Breaker, R. R. (2000). Capping DNA with DNA. *Biochemistry* 39, 3106-14.

11. Wang, W., Billen, L. P. & Li, Y. (2002). Sequence diversity, metal specificity, and catalytic proficiency of metal-dependent phosphorylating DNA enzymes. *Chem Biol* 9, 507-17.

12. Achenbach, J. C., Jeffries, G. A., McManus, S. A., Billen, L. P. & Li, Y. (2005). Secondary-structure characterization of two proficient kinase deoxyribozymes. *Biochemistry* 44, 3765-74.

13. McManus, S. A. & Li, Y. (2007). Multiple Occurrences of an Efficient Self-Phosphorylating Deoxyribozyme Motif. *Biochemistry*.

14. Gellert, M., Lipsett, M. N. & Davies, D. R. (1962). Helix formation by guanylic acid. *Proc Natl Acad Sci U S A* 48, 2013-8.

15. Williamson, J. R., Raghuraman, M. K. & Cech, T. R. (1989). Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell* 59, 871-80.

16. Zakian, V. A. (1989). Structure and function of telomeres. *Annu Rev Genet* 23, 579-604.

17. Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L., Coviello, G. M., Wright, W. E., Weinrich, S. L. & Shay, J. W. (1994). Specific association of human telomerase activity with immortal cells and cancer. *Science* 266, 2011-5.

18. Dai, J., Dexheimer, T. S., Chen, D., Carver, M., Ambrus, A., Jones, R. A. & Yang, D. (2006). An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution. *J Am Chem Soc* 128, 1096-8.

19. Rankin, S., Reszka, A. P., Huppert, J., Zloh, M., Parkinson, G. N., Todd, A. K., Ladame, S., Balasubramanian, S. & Neidle, S. (2005). Putative DNA quadruplex formation within the human c-kit oncogene. *J Am Chem Soc* 127, 10584-9.

20. Siddiqui-Jain, A., Grand, C. L., Bearss, D. J. & Hurley, L. H. (2002). Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc Natl Acad Sci U S A* 99, 11593-8.

21. Han, H. & Hurley, L. H. (2000). G-quadruplex DNA: a potential target for anti-cancer drug design. *Trends Pharmacol Sci* 21, 136-42.

22. Cogoi, S., Quadrifoglio, F. & Xodo, L. E. (2004). G-rich oligonucleotide inhibits the binding of a nuclear protein to the Ki-ras promoter and strongly reduces cell growth in human carcinoma pancreatic cells. *Biochemistry* 43, 2512-23.

23. Sen, D. & Gilbert, W. (1988). Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* 334, 364-6.

24. Oka, Y. & Thomas, C. A., Jr. (1987). The cohering telomeres of Oxytricha. *Nucleic Acids Res* 15, 8877-98.

25. Panyutin, I. G., Kovalsky, O. I., Budowsky, E. I., Dickerson, R. E., Rikhirev, M. E. & Lipanov, A. A. (1990). G-DNA: a twice-folded DNA structure adopted by single-stranded oligo(dG) and its implications for telomeres. *Proc Natl Acad Sci U S A* 87, 867-70.

26. Risitano, A. & Fox, K. (2004). Influence of loop size on the stability of intramolecular DNA quadruplexes. *Nucleic Acids Res* 32, 2598-2606.

27. Hazel, P., Huppert, J., Balasubramanian, S. & Neidle, S. (2004). Loop-length-dependent folding of G-quadruplexes. *J Am Chem Soc* 126, 16405-15.

28. Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31, 3406-15.

29. Neidle, S. & Parkinson, G. N. (2003). The structure of telomeric DNA. *Curr Opin Struct Biol* 13, 275-83.

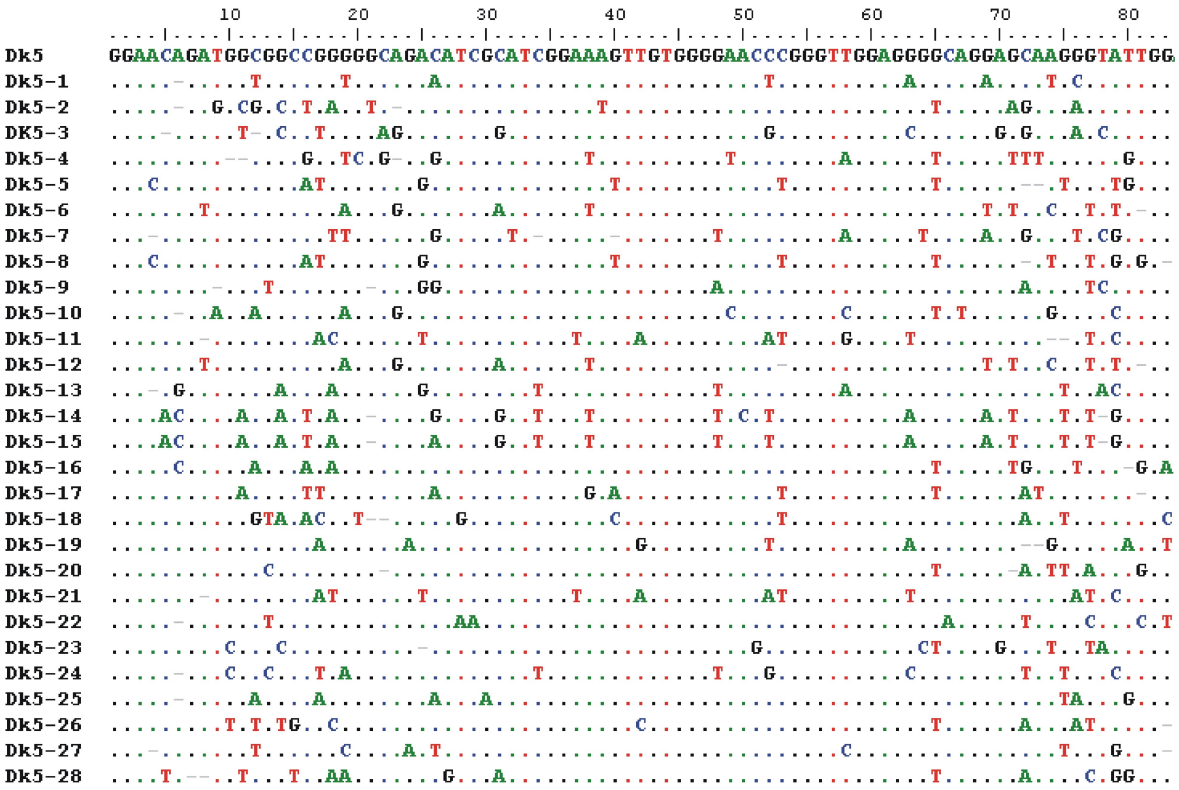
30. Balagurumoorthy, P., Brahmachari, S. K., Mohanty, D., Bansal, M. & Sasisekharan, V. (1992). Hairpin and parallel quartet structures for telomeric sequences. *Nucleic Acids Res* 20, 4061-7.

31. Sen, D. & Gilbert, W. (1990). A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature* 344, 410-4.

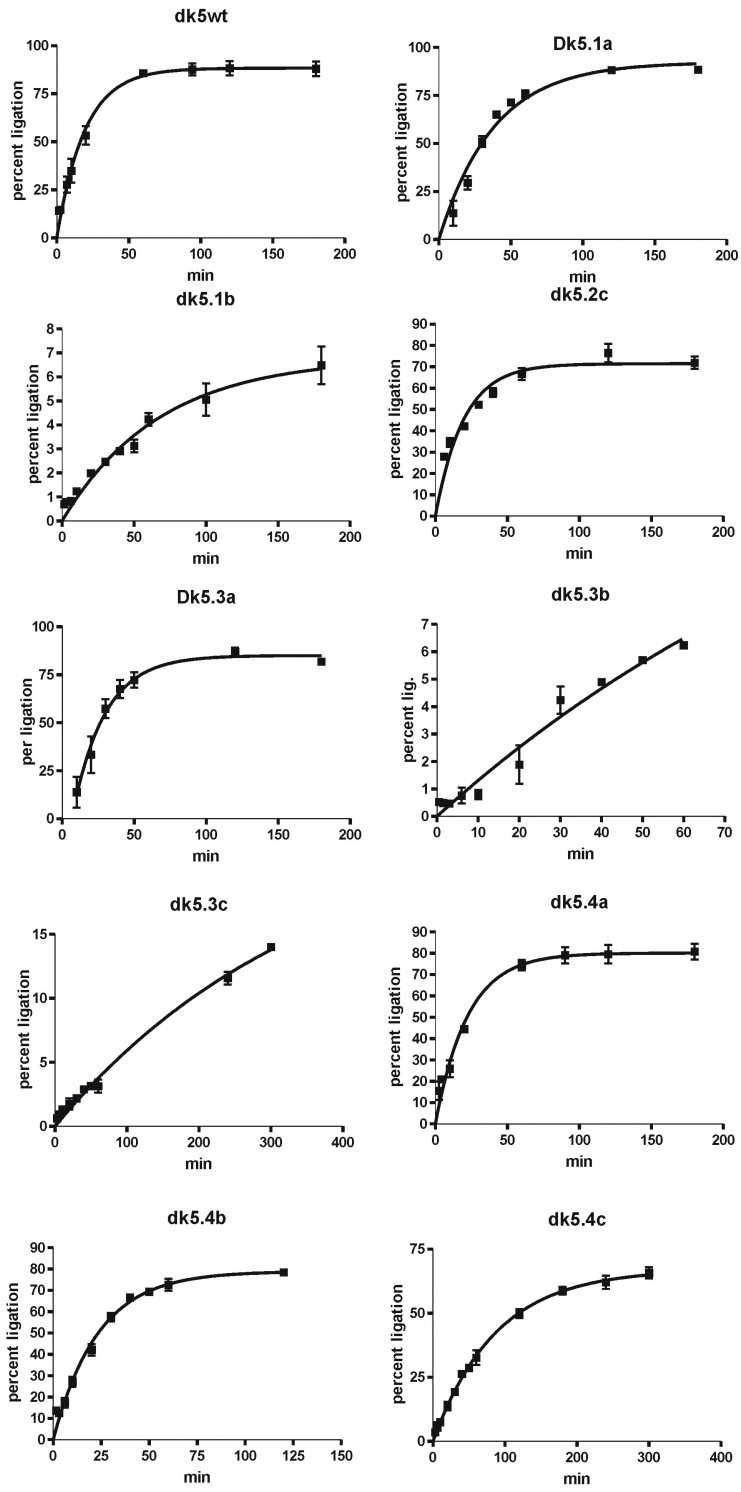
32. Todd, A. K., Johnston, M. & Neidle, S. (2005). Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res* 33, 2901-7.

33. Huppert, J. L. & Balasubramanian, S. (2005). Prevalence of quadruplexes in the human genome. *Nucleic Acids Res* 33, 2908-16.

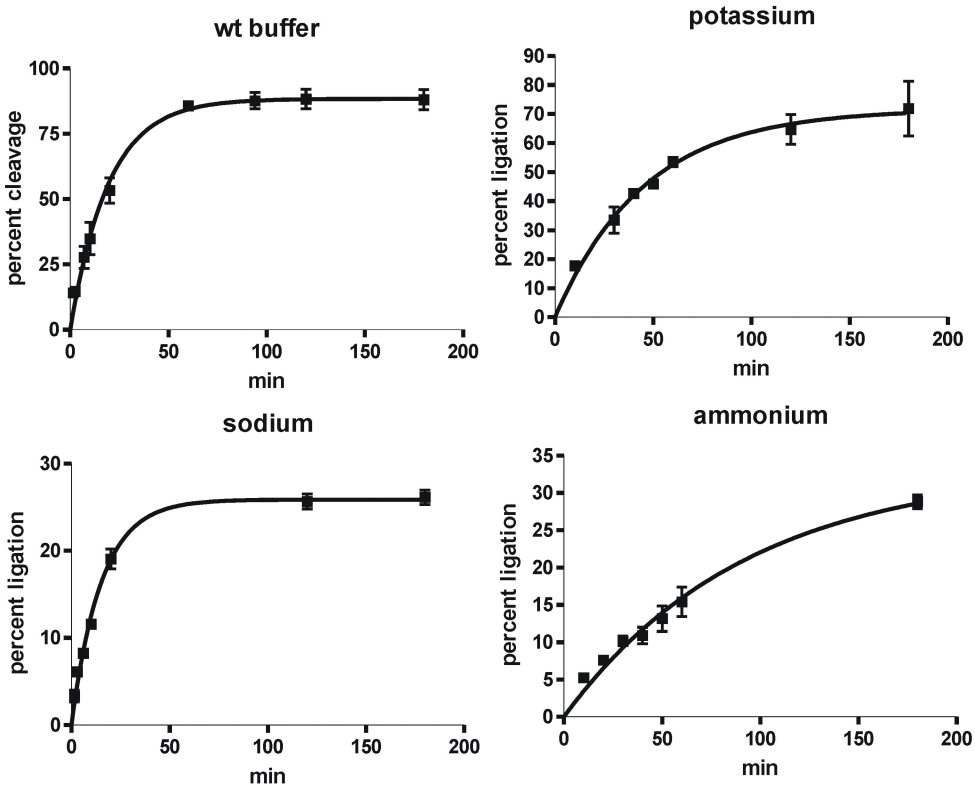
3.9 SUPPLEMENTARY INFORMATION



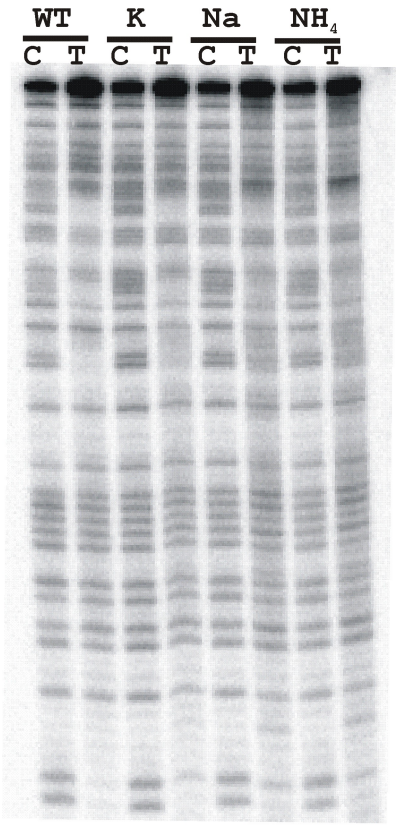
Supplementary Figure 3.1 Alignment of sequences found in the fifth round of the in vitro reselection of Dk5. Dots in each sequence indicate non-mutated nucleotides in reference to the original Dk5; "-": a deletion. The 3' constant sequence, AGCTGATCCTGATGG, is omitted for clarity.



Supplementary Figure 3.2 Graphs used to calculate reaction rates for Dk5 mutants shown in Figure 3.3 of the main manuscript.



Supplementary Figure 3.3 Graphs used to calculate reaction rates for Dk5 under different buffer conditions shown in Table 3.1 of the main manuscript.



Supplementary Figure 3.4 DMS interference patterns for Dk5 under the buffer conditions used in Table 3.1. A similar pattern was seen as in the wild-type buffer (left two lanes). DMS protocol performed the same as in Figure 3.4 of the main manuscript.

Chapter 4:
Assessing the Amount of Quadruplex Structures Present within Diguanine Repeat Libraries

4.1 AUTHOR'S PREFACE

The research project described in this chapter describes the design and validation of DNA sequence libraries that are composed primary of DNA molecules folded into quaduplex structures. These libraries were modelled on two-tiered quadruplex motifs elucidated in chapter 3 of this thesis as well in other aptamers and deoxyribozymes. Biophysical analysis of DNA libraries containing four diguanine repeats and intervening random sequences of differing lengths revealed that short DNA libraries could be designed that contained the majority of their sequences folded into a variety of different quadruplex arrangements and can be used in amplification schemes suitable for in vitro selection. This should be of great value to the field of nucleic acid selection, as it will give another tool to researchers trying to develop aptamers against elusive target molecules.

This chapter is a modified version of a manuscript that has been submitted for publication. I will be the first author on this publication and I conducted all of the experiments presented within and wrote the manuscript. I also took a leading role in the conceptualization, experimental design and interpretation of the results obtained. Dr. Yingfu Li is listed as an author of this paper and provided many fruitful suggestions and guidance as well as assisting in editing of the final version of this manuscript.

4.2 ABSTRACT

The process of in vitro selection has led to the discovery of many aptamers with potential to be developed into inhibitors and biosensors, but problems in isolating aptamers against certain targets with desired affinity and specificity still remain. One possible improvement is to use libraries enhanced for motifs repeatedly isolated in aptamer molecules. One such frequently observed motif is the two-tiered guanine quadruplex. In this study we investigated whether DNA libraries could be designed to contain a large fraction of molecules capable of folding into two-tiered DNA quadruplexes. Using circular dichroism and UV melting analysis, we found that DNA libraries could be designed to contain a large proportion of sequences that adopt quadruplex structures. Analysis of individual sequences from a small library revealed a mixture of quadruplexes of different topologies providing the diversity desired for an in vitro selection. We also investigated the effect of primer binding sites on the folding of a quadruplex library and devised a method for the amplification of quadruplex libraries. With the development of guanine quadruplex enriched DNA libraries it will be possible to improve the chances of isolating aptamers that utilize a quadruplex scaffold and enhance the success of in vitro selection experiments.

4.3 INTRODUCTION

Since its development over twenty years ago (Tuerk et al., 1990; Ellington et al., 1990), the process of in vitro selection has led to the isolation of numerous functional nucleic acids that bind a wide range of target molecules. These nucleic acids, called aptamers, have been used in applications such as enzyme inhibition and biosensing. The power of in vitro selection comes from the ability to select for aptamers from very large DNA libraries, allowing an aptamer sequence to be isolated from up to 10^{16} unique DNA sequences. Despite this, the process of in vitro selection is enigmatic and is in part responsible for the prevention of aptamers from reaching the mainstream. While sensitive and specific aptamers have been isolated for many targets, selections against certain target ligands have failed to isolate aptamers, or have yielded aptamers that lack the desired affinity or specificity. Currently, in vitro selection is a lengthy process requiring iterative rounds of selection and amplification with no guarantee of success. As such, any means of increasing the chances of isolating potent aptamers by improving an aspect of the current in vitro selection procedure would be of great value. One possible approach to improving in vitro selection is to design an initial library that in some way increases the number of potential aptamer sequences while maintaining library diversity. It has been observed that some structural motifs arise repeatedly in isolated aptamers and nucleic acid catalysts, such as an ATP aptamer motif (Huizenga et al., 1995; Nutiu et al., 2005), the hammerhead ribozyme motif (Salehi-Ashtiani et al., 2001), the 8-17 RNA-cleaving DNAzyme motif (Faulhammer et al., 1996; Santoro et al., 1997; Li, J. et al., 2000; Schlosser et al., 2004), and a self-phosphorylating DNAzyme motif (McManus et

al., 2007). As certain structural motifs seem to be favored, one possible method to improve library design is to bias DNA libraries with sequences that fold into these structures. One particular motif that arose repeatedly during in vitro selections, and is of interest to us, is the guanine quadruplex motif.

Guanine quadruplexes are four-stranded structures composed of stacks of quartets of guanines (Gellert et al., 1962; Williamson et al., 1989). They are found naturally in the form of telomeric DNA (Zakian, 1989) and in the promoter regions of several proto-oncogenes (Dai et al., 2006; Rankin et al., 2005; Siddiqui-Jain et al., 2002; Han et al., 2000). Due to their proposed roles in cell immortality and gene regulation these sequence motifs have received much attention as possible cancer-drug targets. These motifs have also repeatedly been identified in functional nucleic acids isolated from numerous in vitro selection experiments. This is particularly true in selections for DNA aptamers and deoxyribozymes (Bock et al., 1992; Mazumder et al., 1996; Li et al., 1997; Li, Y. et al., 2000; Chinnapen et al., 2004; McManus et al., 2008). The propensity of DNA to form these types of structures may be due to conformational varieties of quadruplexes available to DNA in which the loop residues can be in a number of different arrangements allowing many possible interactions with target molecules by one or more loops. Also noteworthy is that as opposed to the majority of quadruplexes found in biological systems, many of the quadruplexes obtained through in vitro selection contain quadruplexes with only two tiers of guanine quartets (Bock et al., 1992; Mazumder et al., 1996; Chinnapen et al., 2004; McManus et al., 2008). We speculate that this may be due to two-tier quadruplexes being stable enough to provide a structural

scaffold, while still being small enough, with only eight nucleotides, to be sampled frequently in a typical random sequence library.

We set out to assess the frequency of two-tiered quadruplex structures within a library containing four pairs of guanines interspersed with random sequences. Studies of this type have been carried out computationally to investigate the types of helical structures present within random sequence libraries (Gevertz et al., 2005), and libraries have been designed to include desired amounts of particular helical structural motifs (Davis et al., 2002; Ruff et al., 2010; Luo et al., 2010). Assessing the level of quadruplex formation with a random sequence DNA library by computation methods is more problematic, as folding programs, such as M-fold (Zuker, 2003), were developed using the free energy of known secondary structures comprised of Watson-Crick base-pairs and do not account for competing structures such as guanine quadruplexes. For quadruplex structural predictions, algorithms do exist to predict the folding and stability of quadruplex-forming sequences (Stegle et al., 2009) but these methods are designed and modelled based primarily on known quadruplexes which are far fewer in number than helical-based structures used for the aforementioned secondary structure prediction algorithms. As there is no connection between these servers, it is difficult to deduce the influence of random sequences on guanines predicted to fold in quadruplex structures. As eight guanines are needed for form a two-tiered quadruplex structure, formations of a stable helix that forms a base pair with even one guanines in a designed library could disrupt quadruplex formation. To investigate the quadruplex folding of DNA libraries we undertook an experimental approach to assess the fraction of sequences within partially

designed DNA libraries that fold into quadruplex arrangements and determine whether libraries can be designed that contain a large number of sequences folded into two-tiered quadruplexes.

4.4 RESULTS

4.4.1 Design of the DNA libraries

We set out to test whether DNA libraries could be designed in which a large fraction of the DNA sequences fold into two-tiered quadruplex scaffolds. As shown in Figure 4.1A, our design consisted of libraries with four pairs of guanines interspersed with three random nucleotide regions. This arrangement should allow sequences to form two-tiered quadruplexes, with the four diguanine repeats forming the two tiers of the quadruplex and the random nucleotides serving as the loop residues which fold back the DNA strands between each diguanine of the quadruplex. The importance of loop length for quadruplex formation has been studied in higher tier number quadruplexes (Risitano et al., 2004; Hazel et al., 2004; Rachwal et al., 2007; Bugaut et al., 2008). These studies have shown that quadruplexes with small loops tend to form defined structures of either parallel quadruplexes with all four strands oriented in the same direction or antiparallel quadruplexes with strands lined up in opposite orientations. With increased loop length, mixtures of multiple types of quadruplexes are found in solution. Several different quadruplex topologies are possible when regarded from the orientation of the loops, with some representative examples shown in Figure 4.1b. In total, 26 different quadruplex topologies are theoretically possible for quadruplexes containing three loops (Webba da Silva, 2007). This mixed set of structures would be ideal for an in vitro selection library

as different types of quadruplexes require different sets of loop orientations, such as edgewise loops, diagonal loops, and double-chain reversal loops. These different loops would be located at different positions relative to the quadruplex, which allows for different possible interactions with the target ligand. This study will examine which loop lengths promote the most quadruplex formation and produce the most variety of quadruplex structures in the DNA library. A lower of limit of three random nucleotides per loop was chosen as any fewer random nucleotides would result in a library of less than 15 nucleotides, which is likely to be too small for isolation of sequences that can bind to a target molecule. The upper limit of seven nucleotides per loop was chosen as it has been found that loops longer than seven nucleotides typically reduce stability and typically do not support quadruplex formation (Hazel et al., 2004; Bugaut et al., 2008).

It has also been demonstrated the choice of monovalent metal ion co-factor can affect the folding of guanine rich sequences into quadruplexes. K^+ is typically found to be favored over Na^+ and Li^+ in quadruplex formation, due to its larger ionic radius which allows K^+ to coordinate better with carbonyl oxygen atoms within the quadruplex (Neidle et al., 2006). It has also been found that K^+ requires less energy for dehydration than smaller ions like Na^+ and Li^+ (Hud et al., 1996). To increase the chances of quadruplex formation within the library, buffers containing K^+ will be used during biophysical studies.

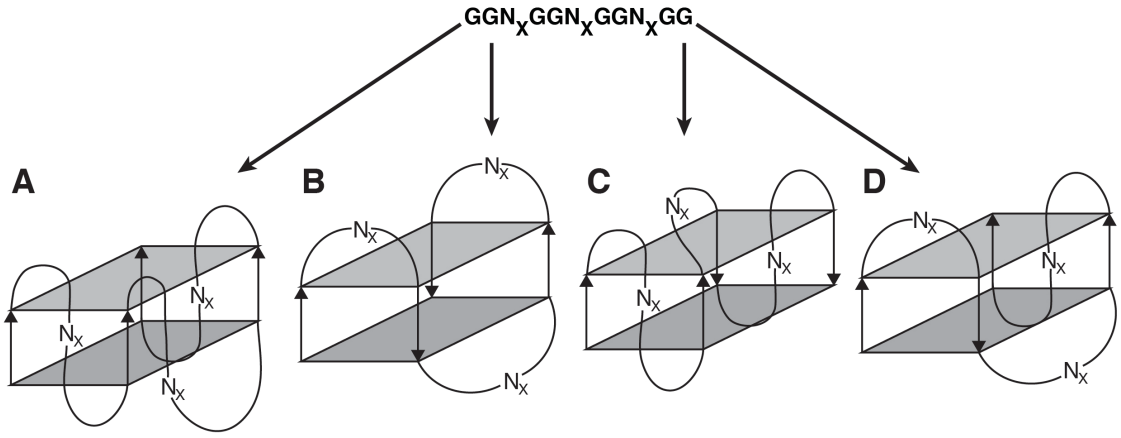


Figure 4.1 Diversity of potential folding patterns of diguanine repeat libraries. Sequences with four pairs of diguanine (GG) sequences and intervening random regions (N_X) have the potential to form different types of quadruplex structures such as (A) A quadruplexes with three double chain reversal loops, (B) A quadruplex with three edgewise loops, (C) A quadruplex with containing two edgewise loops and a double chain reversal loop, or (D) A quadruplex with a diagonal loop, and two double chain reversal loops.

4.4.2 Testing the effects of loop length on folding of diguanine repeat sequences

Before examining the effects of random loop regions on the folding of diguanine repeat libraries, we tested whether diguanine repeat sequences with fixed loop sequences of different loop lengths could form quadruplexes and whether these structures could be analyzed empirically. The core of quadruplexes contain overlapping molecular oribitals of interacting guanines (Webba da Silva et al., 2009) making them amenable to spectroscopic techniques. Circular dichroism (CD) analysis between 320 and 220 nm has been shown to be a powerful technique for the characterization of quadruplex-forming nucleic acids (Vorlickova et al., 2012). CD can be used to determine whether DNA sequences are folding into quadruplex arrangements, and even yield information about

the topology of the quadruplex formed. Through CD of several quadruplexes of known structures, it have been shown that quadruplexes with different arrangements of glycosidic bond angles of the interacting guanines, resulting in different quadruplex topologies, produce distinct spectra (Karsisiotis et al., 2011). These quadruplex topologies were classified into three types. Type I quadruplexes are parallel quadruplexes with all glycosidic bond angles in the *anti* configuration and were found to produce spectra with positive ellipticity around 265 nm and minimal ellipticity at 295 nm. Type II and III quadruplexes are antiparallel quadruplexes, in which the glycosidic bond angles of stacked guanines are in same or the opposite orientation, respectively. Type II quadruplexes show positive ellipticity peaks at 295 nm and 265 nm. Type III quadruplexes also show positive ellipticity at 295 nm but show a negative ellipticity peak at 265 nm. To test the effect of loop length on sequences with the potential to form two-tiered quadruplexes, we subjected DNA molecules with four pairs of guanines interspaced with different numbers of thymines ($GGT_XGGT_XGGT_XGG$) to CD analysis. As shown in Figure 4.2A, we found that DNA molecules with two thymines in each of their loops folded into predominantly type I quadruplex structures, as indicated by a large positive peak at 264 nm and a negative peak at 240 nm. In contrast, DNA molecules with three, four, and five thymines between guanines displayed spectra characteristic of type III quadruplexes with strong positive peaks at 295 nm and negative peaks at 265 nm. We also found that sequences with different numbers of thymines in between diguanines behaved based on the longest loop. As shown in Supplementary Figure 4.1, all sequences with at least one loop of three nucleotides, and one or two loops of two nucleotides,

showed spectra indicative of type III quadruplexes, similar to the sequence with three thymines interspacing all pairs of guanines. This demonstrated that this technique could be used to detect the presence of two-tiered quadruplex structures sequences containing four diguanine repeats.

4.4.3 Determining folding of diguanine repeat libraries by CD

To examine the folding that would be present in a library containing four pairs of guanine residues we performed circular dichroism analysis on libraries containing three random regions flanked by four pairs of guanines ($GGN_XGGN_XGGN_XGG$). Circular dichroism is an appropriate technique in this endeavor as the concentrations necessary to obtain a readable spectrum (μM range) are similar to those used in the first round of selection where it is desirable to screen as many sequences as possible. We chose to test libraries with loop lengths between three and seven residues for a number of reasons. First, the majority of efficient aptamers fall within this size range. Also, using these libraries will allow for the complete sampling of available sequence space, which would not be possible using larger libraries. Finally, as libraries increase in size and the random sequences represent higher proportions of the sequences compared to the fixed guanines, they will likely become more involved in the structural folding and the two-tiered quadruplex arrangement may be lost in favor of other structural arrangements. The CD spectra of the GGN_3 to GGN_7 are shown in Figure 4.2B. Each scan contains a bimodal spectrum with positive ellipticity peaks at 264 and 295 nm of differing intensities. This could be explained by two scenarios. One scenario is that the majority of sequences are folded into type II quadruplexes, which would explain the positive ellipticity at 264 and

295 nm. In the second scenario a mixture of type I, II and III quadruplexes could be present, with their cumulative spectra averaging to create the positive peaks at 264 and 295 nm. This has been observed previously with sequences known fold into heterogenous mixtures of different quadruplex structures. This second scenario would also explain the differences in 264/295 nm ratios seen when comparing the spectra of the different loop libraries, as these libraries could contain different amounts of each type of quaduplex. For instance, the spectrum of $GGN_3GGN_3GGN_3GG$ has a peak in ellipticity at 264 nm about twice the intensity of a second peak seen at 295 nm. When one random nucleotide is added to each loop to create the $GGN_4GGN_4GGN_4GG$ the spectrum is reversed with a peak in ellipticity at 295 nm that is twice as intense as the peak at 264 nm. It could be reasoned that this shift in spectra could be due to more sequences in the $GGN_4GGN_4GGN_4GG$ library being folding into type III quadruplexes and less being folded into type I quadruplexes. This would explain the lowering of 264 nm ellipticity in $GGN_4GGN_4GGN_4GG$ as less type I quadruplexes would be present to contribute to positive ellipticity at 264 nm and the increased number of type III quadruplexes would provide negative ellipticity lessening the ellipticity at 264 nm. This analysis is somewhat speculative based solely on analysis of these libraries and the question of the folding of the individual sequence in the libraries will be addressed by analysis of a small library, later in the manuscript.

In contrast to the diguanine repeat libraries, a random library of twenty-three nucleotides showed a spectrum with a single positive peak around 280 nm and no peaks at 264 and 295 nm. This suggests that the majority of sequences in this library are folded

into other non-quadruplex structures. The lack of other peaks, such as the 280 nm peak seen with the completely random library, shows that there are not enough sequences folded into other structures to appear in the CD spectra of the diguanine libraries.

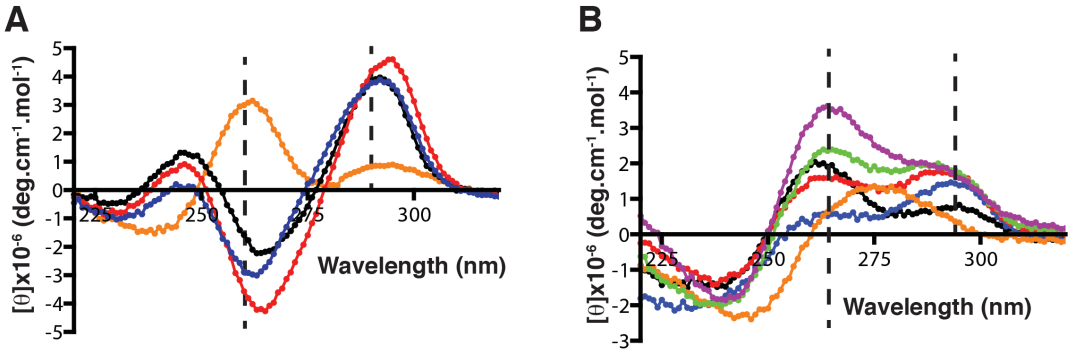


Figure 4.2 Effects of loop length on diguanine repeat libraries. (A) Diguanine repeat sequences ($GGT_XGGT_XGGT_XGG$) with differing number of intervening thymine residues were analyzed by circular dichroism between 320 and 220 nm. Series are colored as follows: orange, X = 2; black, X = 3; red, X=4; blue, X=5. Dotted lines are placed at 264 and 295 nm as ellipticities at these wavelengths indicate the presence of of different subtypes of guanine quadruplexes. (B) $GGN_XGGN_XGGN_XGG$ were scanned for X = 3 to X = 7. Black is X = 3, blue is X = 4, red is X = 5, green is X = 6, and purple is X = 7. A random 23 nucleotide was scanned for comparison and is shown in orange.

4.4.4 Testing stability by UV melting profile

We next set out to assess the stability of the folded sequences within the libraries. It has been shown that quadruplexes absorb 50 to 80% more light at 295 nm compared with their unfolded counterparts (Mergny et al., 1998) and A_{295} melting profile have been used to train quadruplex stability algorithms (Stegle et al., 2009). Observing the change in absorbance at 295 nm during heating and cooling can assess the stability of a quadruplex. The absorbance of each diguanine repeat library was recorded between 15 and 85°C. As shown in Figure 4.3, each library exhibited a decrease in absorbance as the

temperature increased. This is consistent with the molecules within the libraries being folded into quadruplex structures. The melting profile of each library is broader than that typically observed with a single DNA sequence, likely due to the multiple quadruplex types present and interaction between different loop sequences increasing or decreasing the stability of the quadruplexes formed. The melting temperature (T_m) of each library was determined by taking the minimum of the first derivative of the melting profile (Figure 4.3B). The melting temperatures for all diguanine libraries with loops from three to seven nucleotides were all found to be between 50 and 60^{o}C. This demonstrates that these libraries are comprised of quadruplex scaffolds at temperatures typically used for in vitro selection (between 16 and 37^{o}C). It is also noteworthy that the scans did not show hysteresis during cycles of heating and cooling. This indicates that the solutions contain mostly unimolecular quadruplexes, as four stranded quadruplexes have very slow folding kinetics that results in hysteresis during repeated heating and cooling. The dominance of unimolecular quadruplexes would be beneficial for an in vitro selection as most selection schemes are designed to select for unimolecular functional nucleic acids. There is also a decreasing trend in the difference in absorbance between the folded and melted states as the number of the nucleotides in the random loops increases, with the $GGN_7GGN_7GGN_7GG$ profile having a significantly lower difference in absorbance than $GGN_3GGN_3GGN_3GG$. This suggests that as the number of nucleotides increases, there are less DNA molecules folded into unimolecular quadruplexes.

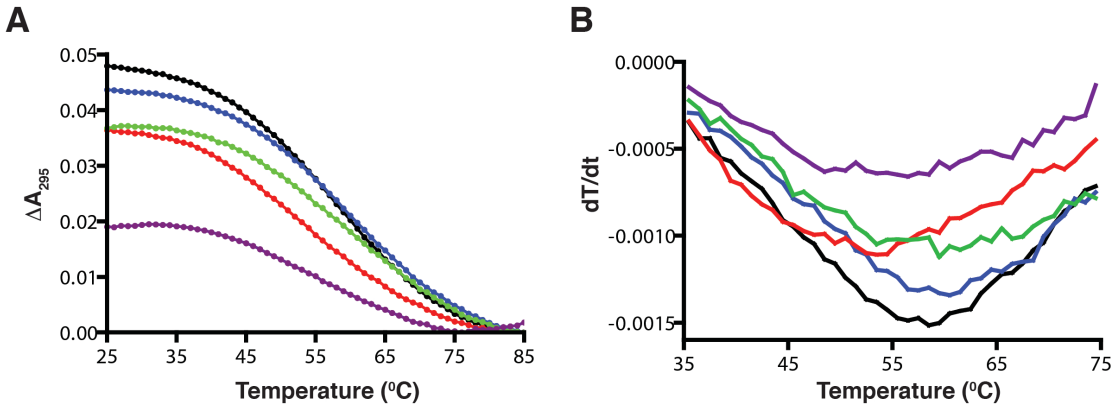


Figure 4.3 UV melting profile of diguanine repeat libraries. (A) The stability of diguanine repeat libraries with random loop regions between three and seven nucleotides was assessed by thermal denaturation analysis. The absorbance of each library was measured at 295 nm between 15 and 85°C. The difference in absorbance was measured as the difference between the highest and lowest A_{295} measurements for each sample. The color scheme is the same as in Figure 4.2. (B) The first derivative of the melting profiles was taken for each library. The most negative derivative value for each library represents the T_m.

4.4.5 Examining the folding of a small library

The work described so far has been performed on a mixture of a library of 10^{13} sequences with the data presented representing an accumulation of the properties of each sequence within the library. To investigate the behaviour of individual sequences in a library, we analyzed 50 diguanine repeat sequences with randomly generated loop regions by CD. The sequences were designed based on the $GGN_5GGN_5GGN_5GG$ library, containing three random regions of five nucleotides each. This library was chosen for further analysis as its CD spectrum indicated that it showed nearly equal amounts of ellipticity at 264 and 295 nm suggesting it may contain a mixture of molecules folded different types of quadruplexes. By analyzing individual sequences, it can be determined whether this represents a mixture of type I type II and III arrangements or a homogenous

population of type II quadruplexes that would also explain this spectra. The 50 sequences were generated using rsat (van Helden, 2003)(Supplementary Table 1) and each were individually scanned by CD between 320 and 220 nm. 44 of the 50 sequences gave a profile that was indicative of a quadruplex structure. Each sequence was classified by the its CD profile, with sequences with a 264 nm maximum classified as type I quadruplexes, sequences with positive ellipticities at 264 and 295 nm being classified as type II quadruplexes sequences and sequences with positive ellipticitity at 295 nm and negative ellipticity at 264 nm being classified as type III quadruplexes. As seen in Figure 4.4, 14 of the sequences adopted a type II conformation (Figure 4.4A), 19 folded into a type I arrangement (Figure 4.4B), and 11 sequences showed spectra indicative of a type III conformation (Figure 4.4C). The presence of all three of the possible quadruplex structural topologies in roughly equal amounts suggest that this library has the variety of arrangements necessary for it to be useful for an in vitro selection. In contrast, in a screen of 50 completely random sequences of the same length (Supplementary Table 2), only 6 sequences displayed a peak maximum near 264 nm indicative of a type I quadruplex, while the other 44 sequences had a maximum around 280 nm. This suggests that the majority of randomly generated sequences are not folding into quadruplex structures (Figure 4.4D). For comparison with the library scans, the spectra of the 50 diguanine repeat sequences were summed and divided by the total number of sequences. As shown in Supplementary Figure 4.2, the profile contains a peak at 295 nm with a shoulder at 265 nm similar to that observed when with the random sequence libraries. When the spectra of the 50 random sequences were combined they yielded a profile with

a peak around 280 nm, also similar to their random library counterpart. The observation that both the combined spectra of the 50 diguanine repeat sequences and of the 50 random sequences resembled the spectra generated with their respective libraries validates the utility of CD scans of libraries as an indicator of the percentages of different types of quadruplex structures present within each pool of sequences.

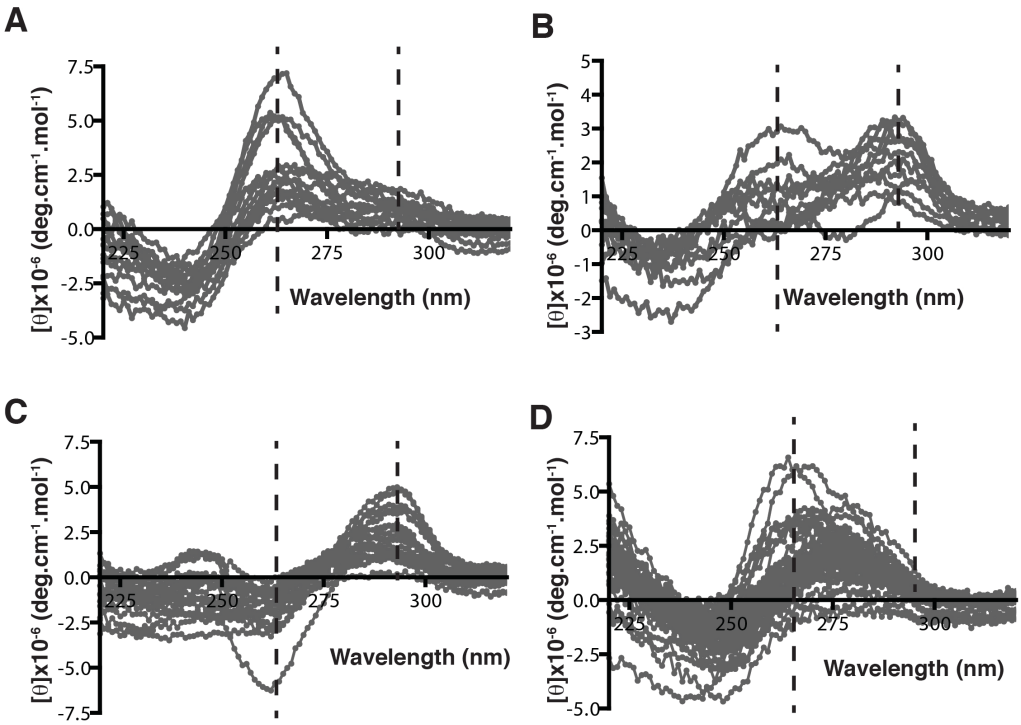


Figure 4.4 Analysis of small diguanine repeat library. Circular dichroism analysis was carried out on 50 randomly generated diguanine repeat sequences based on the library $GGN_5GGN_5GGN_5GG$. (A) 14 sequences had spectra with a positive peak near 264 nm indicating they fold into type I quadruplex structures. (B) 11 sequences had positive peaks at 264 and 295 nm suggesting they fold into type II quadruplex structures. (C) 19 sequences had positive peaks at 295 nm and negative peaks at 265 nm indicating they fold into type III quadruplexes. (D) A random 23-nucleotide library contained six sequences with positive peaks around 264 nm, and 44 sequences no peaks at 264 and 295 nm indicating that these sequences are not forming quadruplex structures.

4.4.6 Assessing the utility of diguanine libraries for in vitro selection

While CD and UV melting analyses suggest that libraries containing four diguanine repeats are comprised primarily of DNA molecules folded into quadruplex structures, the libraries need to possess other attributes to be useful for in vitro selection. First and foremost, as in vitro selection contains multiple rounds of selection and amplification, sequences from the library must be capable of being amplified after each selection step to allow for the enrichment of active sequences. This is typically achieved by the addition of fixed sequences to the 5′ and 3′ ends of the library. These sequences serve as primer binding sites, allowing for the amplification of active sequences by PCR. We first tested whether addition of fixed sequences at the 5′ and 3′ ends affected the folding of the diguanine repeat libraries. As shown in Figure 4.5A, 10 nucleotides were added to the 5′ and 3′ of the $GGN_5GGN_5GGN_5GG$ library. CD spectrometry was performed on this library and compared to the original $GGN_5GGN_5GGN_5GG$ library, which lacked primer binding arms. As seen in Figure 4.5B, when the fixed arms are added the spectrum shifts from the mixed quadruplex profile to a profile similar to the completely random library. This suggests that the majority of sequences within the library are forming interactions between the fixed sequence domains and the random nucleotides with the library. If this library were to be used for in vitro selection, many of the quadruplex scaffolds that were present in the library without the fixed sequences would not form and be lost during the selection step. As an alternative, we considered the possibility of ligating primer binding sequences to the 5′ and 3′ ends of sequences within $GGN_5GGN_5GGN_5GG$ library. This would allow the primer binding sites to be

added after the selection step and prevent them from disrupting the folding of sequences into quadruplex structures during the selection steps. The ligation scheme is shown in Figure 4.5C. The specific ligation of the primer binding sites to sequences in the $GGN_5GGN_5GGN_5GG$ library is achieved using two DNA oligonucleotides containing regions complementary to the primer binding sites along with two dicytosines flanking a five nucleotide random region as templates. While performing template directed ligation with randomized regions might seem problematic, in this case the random regions only need to hybridize to five base pairs. Therefore there are only 1024 possible complementary sequences, meaning there are 6×10^{11} complementary sequences present in a one picomole reaction. The specificity of the templates for the library is also conferred by the terminal dicytosine in the template, which can bind to the central diguanines within the library. To test whether ligation of a sequence was possible using partially randomized templates, the ligation of a sequence from the randomly generated library was incubated with 5′ and 3′ adapter sequences and two partially randomized templates as shown in Figure 4.5C. The sequence (DGR36, Supplementary Table 1) was selected as its CD spectrum suggested it folded into a type III quadruplex formation similar to those structures seen with known quadruplex aptamers. The results of the ligation are shown in Figure 4.5D. Well-defined products are seen when one or two oligonucleotides are ligated to the library, showing that specific ligation can be carried out with a randomized region. The efficiency of the ligation is lower than that achieved when the completely complementary template is used, as seen in lanes 3 and 5 of the figure. As the library contains only 10^9 sequences, one nanomole of library will contain

10^6 copies of each individual sequence, so this lower efficiency of ligation should still allow for the amplification of all unique sequences that survive the selection step of the first round of a selection. The presence of a ligation product also shows that this sequence with a strong tertiary quadruplex structure can successfully hybridize to the template in an orientation that allows for ligation.

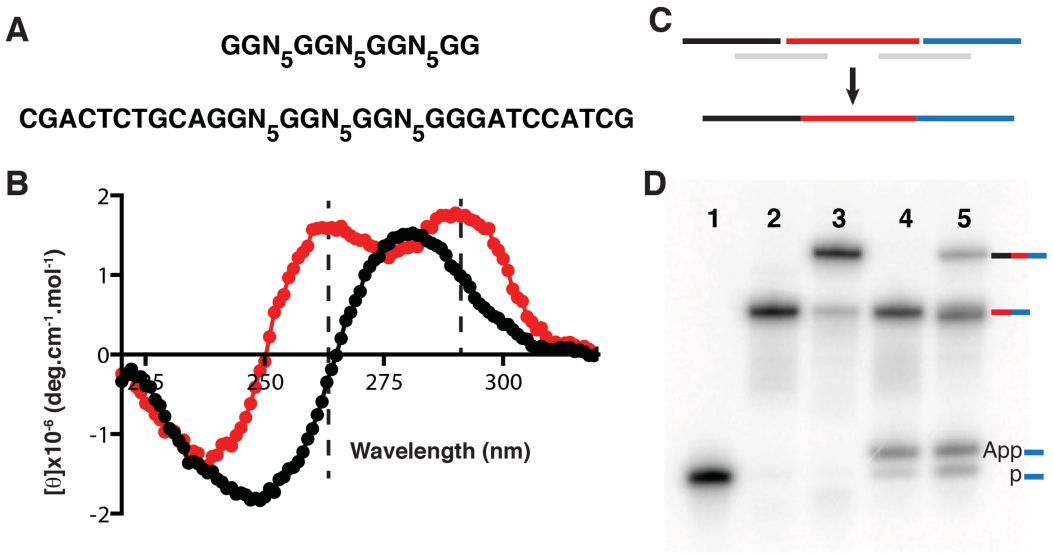


Figure 4.5 Effects of primer binding sites on diguanine repeat library folding. (A) Sequences of the $GGN_5GGN_5GGN_5GG$ library and a library with fixed 5′ and 3′ regions. (B) CD analysis of $GGN_5GGN_5GGN_5GG$ (red) and $GGN_5GGN_5GGN_5GG$ with the two fixed sequences (black). (C) Ligation scheme for the ligation of the representative sequence DGR36 (red) to 5′ (black) and 3′ (blue) adapters, with two template oligonucleotides (grey). (D) Ligation of one and two adapters to DGR36. The 3′ adapter is labelled with ^{32}P (lane 1) and incubated with T4 DNA ligase, DGR36 and fully complementary template (lane 2), as well as all these oligonucleotides plus a 5′ adapter and fully complementary template (lane 3). The same reactions were performed with the fully complementary templates replaced with partially randomized 3′ templates (lane 4) and both 3′ and 5′ templates (lane 5). The p and App on the 3′ adapters in lanes 4 and 5 represent 5′ phosphorylated and 5′ adenylated adapters, respectively.

4.4.7 Amplification and regeneration of diguanine libraries

The ligated DNA molecules were amplified by PCR, using two primers that hybridize to the fixed region of the sense and antisense strand of the diguanine repeat sequence (Figure 4.6A). For successive rounds of selection to be undertaken, it is crucial that the segment corresponding to the initial population be separated from its complementary strand. Moreover, in the case of a diguanine repeat population, the fixed regions at the 5′ and 3′ ends must be removed prior to the next selection step as they will interfere with the folding of the active sequences. In typical aptamer selections, the library strand is separated from its antisense strand by methods such as affinity chromatography of a biotin labeled antisense primer. Methods exist for removing a 5′ sequence from an amplified library, such as incorporation of a ribonucleotide into the reverse PCR primer that can be cleaved by sodium hydroxide treatment (Li et al., 1999). Removing the 3′ fixed region, which is at the end of the amplified product, is typically not performed and requires a novel method for removal.

To accomplish the removal of both the 5′ and 3′ fixed ends as well as the separation of the library strand from its antisense counterpart after PCR, restriction sites were incorporated into the sequences as shown in Figure 4.6A. The 5′ primer contains a PstI site (CTGCAΔG) which encompasses the 5′ terminal guanine of the library and five nucleotides of the fixed 5′ region; the 3’ primer consists of a BamHI site (GΔGATCC), which contains the 3′ terminal guanine and five nucleotides of the 3′ fixed region. Digestion with PstI and BamHI following PCR amplification leads to cleavage of the product into several fragments. Since the restriction enzymes cut each strand at different

sites, the 23 nucleotide library can be separated from the antisense strand by denaturing PAGE as the library will be seven nucleotides shorter. The band distribution observed following PstI and BamHI digestion of the amplified DGR36 sequence is shown in Figure 4.6B. When the PCR product is digested with PstI or BamHI individually, the bands corresponding to the fragments of a single digestion are seen. When the PCR product is digested with both restriction enzymes, a band corresponding to the library is seen. This resulting population of selected sequences from the $GGN_5GGN_5GGN_5GG$ library can then be used to initiate the next round of selection.

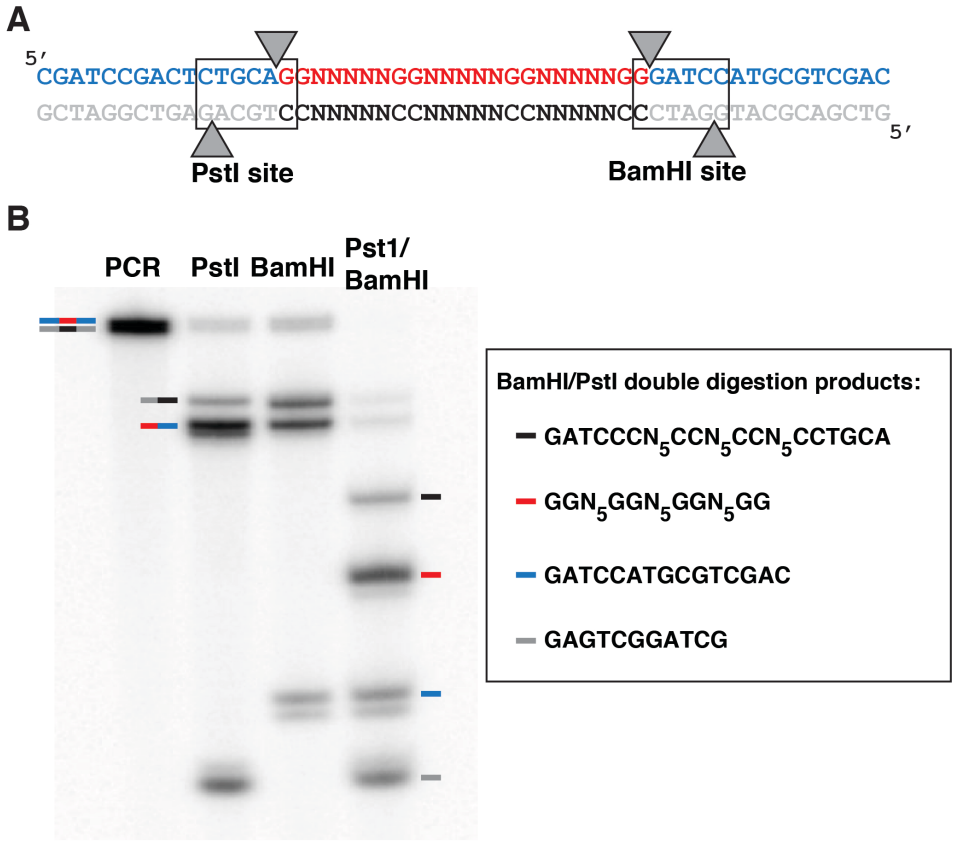


Figure 4.6 Regeneration of $GGN_5GGN_5GGN_5GG$ population from PCR products. (A) Sequence of the PCR product. PstI and BamHI restriction sites are shown in boxes. Restriction enzyme cleavage sites are represented by arrows. (B) Digestion of amplified DGR36 sequence with BamHI and PstI individually (lanes 2 and 3) and in combination

(lane 4). Coloured arrows indicate fragments comprised of sequences of the same colour in part a. Fragments consisting of only the first fifteen nucleotides of the 5′ regions of the PCR products in the single enzyme digestion experiments cannot be seen as only polymerized segments acquire ^{32}P during amplification.

4.5 DISCUSSION

In vitro selection has established itself as a powerful technique for the isolation of a wide range of functional nucleic acids. However, despite this progress, there has been difficulty in isolating aptamers for certain targets or obtaining aptamers of a desired affinity or being able to bind under specific conditions. By understanding which structural arrangements are effectively used by aptamers, it may be possible to modify in vitro selection protocols to favor these types of arrangements. One type of structure, the guanine quadruplex has been seen repeatedly in the core of a number of functional DNA molecules. By analyzing DNA oligonucleotide libraries with four pairs of diguanine sequence elements we were able to demonstrate that the majority of molecules fold into quadruplex structures that are stable at temperatures typically used for in vitro selections. The libraries were also found to have features of all three types of known quadruplexes, suggesting a mixture of these structures in the libraries. Analysis of 50 representative sequences from the $GGN_5GGN_5GGN_5GG$ library revealed that roughly equal number of sequences folded in type I, type II, and type III quadruplexes. This diversity of quadruplex folds translates to a diversity of loop arrangements in quadruplexes within the libraries. In an in vitro selection, these loop residues represent potential binding partners with the target ligand. The more loop arrangements and combinations of loop arrangements present within molecules within the library, the greater the chance of a sequence being present in the library that can bind a target ligand. We have also devised

a scheme for ligation of primer binding arms after a selection step to avoid these fixed sequence elements being present during selection, as we found these sequences to be detrimental to quadruplex formation. Providing libraries with a structural scaffold shall reduce the burden of potential aptamers to use their random residues for intramolecular folding and allow the random nucleotides to be used solely for target binding, which in turn shall allow for the selection of highly selective, high affinity aptamers.

4.6 MATERIAL AND METHODS

4.6.1 Oligonucleotides and materials

DNA oligonucleotides were prepared using standard phosphoramidite chemistry (MOBIX Lab, McMaster University, Hamilton, ON, Canada; and Integrated DNA Technologies, Coralville, IA, USA). The partially randomized diguanine repeat libraries were synthesized using a 25% mixture of each of the four nucleotides at the random positions. Each oligonucleotide was purified by 10% denaturing (8 M urea) PAGE, and its concentration was determined spectroscopically. T4 polynucleotide kinase (PNK) and T4 DNA ligase were purchased from MBI Fermentas. Taq polymerase, BamHI and PstI were purchased from New England Biolabs (Pickering, ON, Canada). [γ-^{32}P]ATP and [α-^{32}P]deoxy-GTP were acquired from Perkin Elmer. All other chemicals were obtained from Sigma.

4.6.2 Buffer conditions

For the CD and UV melting experiments, the analyses were performed under the following conditions: 500 mM KCl, 10 mM $MgCl_2$, and 25 mM HEPES, pH 7.

4.6.3 CD analysis

CD studies were carried out using an AVIV model 410 CD spectrometer (Lakewood, NJ, USA) as per the manufacturer's instructions. All constructs were scanned from 320 to 220 nm at 10 μM DNA concentration in a 0.1 cm quartz cuvette. All DNA samples heated to 90^{o}C in 1× buffer and allowed to cool to room temperature. All shown spectra are the average of three individual scans.

4.6.4 UV melting profiles

Thermal denaturation profiles were carried out using a Cary 100-bio UV spectrometer (Mississauga, ON, Canada) as per the manufacturer's instructions. All libraries were incubated in 1× buffer at a concentration of 4 μM. Oligonucleotides were heated and cooled at a rate of 0.5^{o}C per min and the absorbance was measured at 1^{o}C intervals. First derivative curves were calculated using Graphpad Prism software.

4.6.5 PNK, Ligation, PCR and digestion reactions

Oligonucleotides for the ligation of DGR36 to 5′ and 3′ PCR adapters are shown in Supplementary Figure 4.3A. Before ligation, 50 pmol of DGR36 and 3′ adapter sequence were phosphorylated at their 5′ ends with PNK. The sequences were heated to 90^{o}C and cooled to room temperature before PNK treatment. DGR36 was incubated with 10× PNK buffer A (MBI Fermentas), 1 mM ATP and 10 U of PNK at 37^{o}C for 30 min. The reaction was then heated at 90^{o}C for 5 min to deactivate the PNK. The 3' adapter was treated in the same way except 1 μL of [γ-^{32}P]ATP was used instead of 1 mM ATP to allow for radiolabeling of the adapter. Ligations were carried out using 50 pmol of each oligonucleotide shown in Supplementary Figure 4.3A. Oligonucleotides were heated to

90^{o}C for 1 min and cooled to room temperature. 10× T4 ligase buffer and 1 U of T4 ligase were added and the reaction was incubated for 1 hr at room temperature. Each reaction was precipitated with ethanol and the DNA was resuspended in 10μL 1× loading buffer and resolved by 10% denaturing PAGE. Ligated products were extracted from the gel for subsequent PCR amplification. PCR primers used for amplification of ligated $GGN_5GGN_5GGN_5GG$ library are shown in Supplementary Figure 4.3B. 50 pmol of forward and reverse primers were mixed with a 1000× dilution of the ligated product, 200 μM dNTPs, 1 μL of [α-^{32}P]deoxy-GTP, 10× Taq polymerase buffer with $MgCl_2$ and 1 U of Taq polymerase. Twenty cycles of PCR were performed. The reaction was precipitated with ethanol and the PCR products isolated by 10% denaturing PAGE. Restriction double-digest of the PCR product was carried out using 20 U of BamHI and PstI in NEB Buffer 3 supplemented with BSA.

4.7 ACKNOWLEDGEMENTS

We would like to thank Rachel and Richard Epand for use of their CD spectrometer.

4.8 FUNDING

This work was supported by the Natural Science and Engineering Research Council of Canada and the Canada Foundation for Innovation (CFI). Funding for open access charge: the Natural Science and Engineering Research Council of Canada.

4.9 REFERENCES

Bock, L. C., L. C. Griffin, J. A. Latham, E. H. Vermaas, and J. J. Toole (1992). Selection of single-stranded DNA molecules that bind and inhibit human thrombin. Nature 355: 564-566.

Bugaut, A., and S. Balasubramanian (2008). A sequence-independent study of the influence of short loop lengths on the stability and topology of intramolecular DNA G-quadruplexes. Biochemistry 47: 689-697.

Chinnapen, D. J., and D. Sen (2004). A deoxyribozyme that harnesses light to repair thymine dimers in DNA. Proc Natl Acad Sci U S A 101: 65-69.

Dai, J., T. S. Dexheimer, D. Chen, M. Carver, A. Ambrus, R. A. Jones, and D. Yang (2006). An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution. J Am Chem Soc 128: 1096-1098.

Davis, J. H., and J. W. Szostak (2002). Isolation of high-affinity GTP aptamers from partially structured RNA libraries. Proc Natl Acad Sci U S A 99: 11616-11621.

Ellington, A. D., and J. W. Szostak (1990). In vitro selection of RNA molecules that bind specific ligands. Nature 346: 818-822.

Faulhammer, Dirk, and Michael Famulok (1996). The Ca2+ Ion as a Cofactor for a Novel RNA-Cleaving Deoxyribozyme. Angew Chem Int Ed Engl 35: 2837-2841.

Gellert, M., M. N. Lipsett, and D. R. Davies (1962). Helix formation by guanylic acid. Proc Natl Acad Sci U S A 48: 2013-2018.

Gevertz, J., H. H. Gan, and T. Schlick (2005). In vitro RNA random pools are not structurally diverse: a computational analysis. RNA 11: 853-863.

Han, H., and L. H. Hurley (2000). G-quadruplex DNA: a potential target for anti-cancer drug design. Trends Pharmacol Sci 21: 136-142.

Hazel, P., J. Huppert, S. Balasubramanian, and S. Neidle (2004). Loop-length-dependent folding of G-quadruplexes. J Am Chem Soc 126: 16405-16415.

Hud, N. V., F. W. Smith, F. A. Anet, and J. Feigon (1996). The selectivity for K+ versus Na+ in DNA quadruplexes is dominated by relative free energies of hydration: a thermodynamic analysis by 1H NMR. Biochemistry 35: 15383-15390.

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

Karsisiotis, A. I., N. M. Hessari, E. Novellino, G. P. Spada, A. Randazzo, and M. Webba da Silva (2011). Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. Angew Chem Int Ed Engl 50: 10645-10648.

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

Li, Y., and R. R. Breaker (1999). Phosphorylating DNA with DNA. Proc Natl Acad Sci U S A 96: 2746-2751.

Li, Y., Y. Liu, and R. R. Breaker (2000). Capping DNA with DNA. Biochemistry 39: 3106-3114.

Li, Y., and D. Sen (1997). Toward an efficient DNAzyme. Biochemistry 36: 5589-5599.

Luo, X., M. McKeague, S. Pitre, M. Dumontier, J. Green, A. Golshani, M. C. Derosa, and F. Dehne (2010). Computational approaches toward the design of pools for the in vitro selection of complex aptamers. RNA 16: 2252-2262.

Mazumder, A., N. Neamati, J. O. Ojwang, S. Sunder, R. F. Rando, and Y. Pommier (1996). Inhibition of the human immunodeficiency virus type 1 integrase by guanosine quartet structures. Biochemistry 35: 13762-13771.

McManus, S. A., and Y. Li (2008). A deoxyribozyme with a novel guanine quartet-helix pseudoknot structure. J Mol Biol 375: 960-968.

McManus, S. A., and Y. Li (2007). Multiple occurrences of an efficient self-phosphorylating deoxyribozyme motif. Biochemistry 46: 2198-2204.

Mergny, J. L., A. T. Phan, and L. Lacroix (1998). Following G-quartet formation by UV-spectroscopy. FEBS Lett 435: 74-78.

Neidle, Stephen, and Shankar Balasubramanian, (2006). *Quadruplex nucleic acids*. RSC biomolecular sciences. Cambridge: RSC Pub.

Nutiu, R., and Y. Li (2005). In vitro selection of structure-switching signaling aptamers. Angew Chem Int Ed Engl 44: 1061-1065.

Rachwal, P. A., I. S. Findlow, J. M. Werner, T. Brown, and K. R. Fox (2007). Intramolecular DNA quadruplexes with different arrangements of short and long loops. Nucleic Acids Res 35: 4214-4222.

Rankin, S., A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian, and S. Neidle (2005). Putative DNA quadruplex formation within the human c-kit oncogene. J Am Chem Soc 127: 10584-10589.

Risitano, A., and K. R. Fox (2004). Influence of loop size on the stability of intramolecular DNA quadruplexes. Nucleic Acids Res 32: 2598-2606.

Ruff, K. M., T. M. Snyder, and D. R. Liu (2010). Enhanced functional potential of nucleic acid aptamer libraries patterned to increase secondary structure. J Am Chem Soc 132: 9453-9464.

Salehi-Ashtiani, K., and J. W. Szostak (2001). In vitro evolution suggests multiple origins for the hammerhead ribozyme. Nature 414: 82-84.

Santoro, S. W., and G. F. Joyce (1997). A general purpose RNA-cleaving DNA enzyme. Proc Natl Acad Sci U S A 94: 4262-4266.

Schlosser, K., and Y. Li (2004). Tracing sequence diversity change of RNA-cleaving deoxyribozymes under increasing selection pressure during in vitro selection. Biochemistry 43: 9695-9707.

Siddiqui-Jain, A., C. L. Grand, D. J. Bearss, and L. H. Hurley (2002). Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. Proc Natl Acad Sci U S A 99: 11593-11598.

Stegle, O., L. Payet, J. L. Mergny, D. J. MacKay, and J. H. Leon (2009). Predicting and understanding the stability of G-quadruplexes. Bioinformatics 25: i374-382.

Tuerk, C., and L. Gold (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 249: 505-510.

van Helden, J. (2003). Regulatory sequence analysis tools. Nucleic Acids Res 31: 3593-3596.

Vorlickova, M., I. Kejnovska, J. Sagi, D. Renciuk, K. Bednarova, J. Motlova, and J. Kypr (2012). Circular dichroism and guanine quadruplexes. Methods.

Webba da Silva, M. (2007). Geometric formalism for DNA quadruplex folding. Chemistry 13: 9738-9745.

Webba da Silva, M., M. Trajkovski, Y. Sannohe, N. Ma'ani Hessari, H. Sugiyama, and J. Plavec (2009). Design of a G-quadruplex topology through glycosidic bond angles. Angew Chem Int Ed Engl 48: 9167-9170.

Williamson, J. R., M. K. Raghuraman, and T. R. Cech (1989). Monovalent cation-induced structure of telomeric DNA: the G-quartet model. Cell 59: 871-880.

Zakian, V. A. (1989). Structure and function of telomeres. Annu Rev Genet 23: 579-604.

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

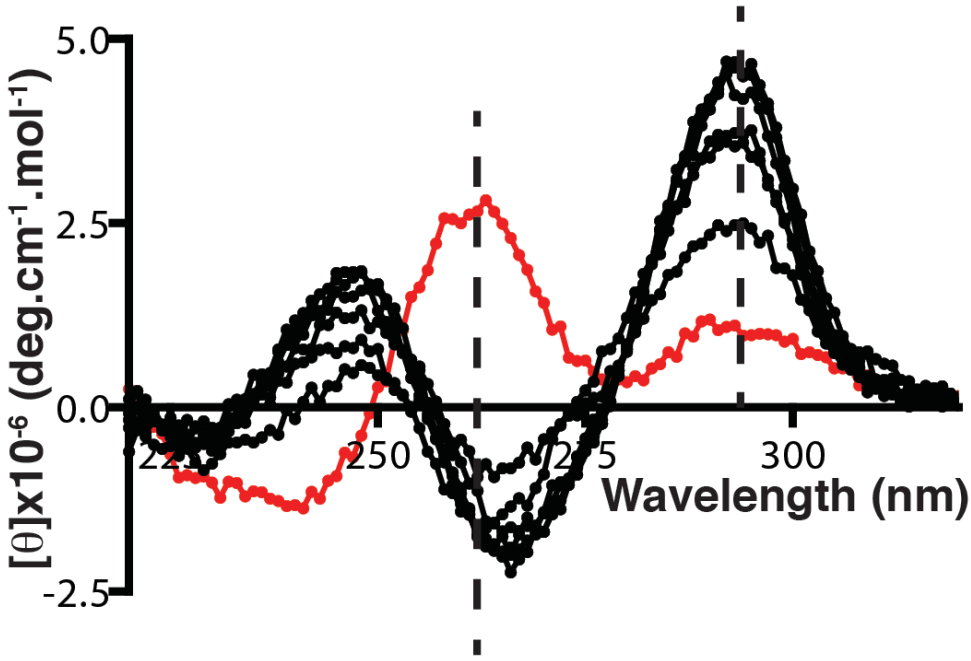
4.10 SUPPLEMENTARY INFORMATION

Supplementary Table 4.1 50 Randomly generated diguanine repeat (dgr) sequences based on $GGN_5GGN_5GGN_5GG$ used for CD analysis. Fixed diguanine motifs are shown in red.

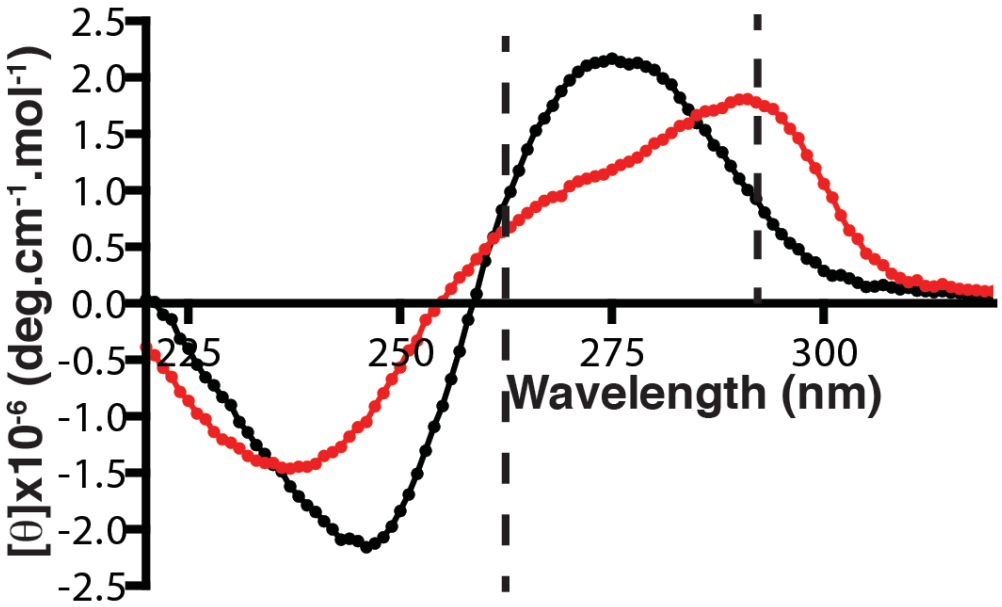
Name	Sequence	Name	Sequence
DGR1	GGCAATAGGCGTAAGGTCTACGG	DGR26	GGGTAGGGGCCCTGGCAGGGGG
DGR2	GGAATAGGGGGGTCGGTGCCGGG	DGR27	GGATAACGGCGTTAGGAATCGGG
DGR3	GGCTCCAGGTTCGCGGATCACGG	DGR28	GGTAGACGGTCCTCGGTATGAGG
DGR4	GGCGAAGGGGAACTGGGGTGGGG	DGR29	GGTAAAAGGCAAATGGATACTGG
DGR5	GGGTTATGGACTGGGGTTCGAGG	DGR30	GGACGTAGGCTGTGGGCCTTTGG
DGR6	GGGGGCGGGGTCACGGAGTTGGG	DGR31	GGATGACGGTGGCAGGCGTGCGG
DGR7	GGAGATAGGGAACGGGTCACTGG	DGR32	GGCATAAGGAGAAGGGGGTAAGG
DGR8	GGGCGCGGGGAAACGGAAGTTGG	DGR33	GGTAGTGGGCGCAAGGCTTCAGG
DGR9	GGCTGACGGGACCGGGAAGAAGG	DGR34	GGGGATGGGCACTCGGAACTAGG
DGR10	GGATCTTGGGTCACGGTTGAGGG	DGR35	GGCCGCCGGTTGCGGGGAATAGG
DGR11	GGATACGGGCGTTGGGCGCGTGG	DGR36	GGAAAATGGCATAAGGAGTTGGG
DGR12	GGTATCTGGCCGGAGGGCCTAGG	DGR37	GGTGAACGGTCCGTGGAATCCGG
DGR13	GGGGATTGGAAGGCGGCCGGGGG	DGR38	GGAATAGGGTGTGAGGGGACAGG
DGR14	GGGTGTCGGACGCCGGTCTTGGG	DGR39	GGGCAAAGGTAACGGGCCCCCGG
DGR15	GGCACATGGGTTAAGGTCGAGGG	DGR40	GGTGGCAGGGGAATGGCAGACGG
DGR16	GGATGTCGGAGATTGGCCAGTGG	DGR41	GGGGTCACGGTCTGTGGCCTCGG
DGR17	GGGCCGCGGCACCGGGGCTTAGG	DGR42	GGGCTCGGGGGATTGGGGTTAGG
DGR18	GGCCGGAGGTTGTTGGGGTGTGG	DGR43	GGGGCGAGGTCAGAGGCAAGGGG
DGR19	GGAAAGGGGAACCCGGTGCGCGG	DGR44	GGTTTAGGGCGCCCGGAGGCTGG
DGR20	GGATGTAGGTGTCTGGTAAACGG	DGR45	GGGAACCGGTGTGAGGTAGACGG
DGR21	GGAAAGAGGCTGGCGGGAACCGG	DGR46	GGCTTTAGGCGAACGGGCTCTGG
DGR22	GGTCAGCGGTATGAGGTGTATGG	DGR47	GGCCATGGGATTTAGGCCTGTGG
DGR23	GGGGGACGGAATAAGGTTTGGGG	DGR48	GGGATTCGGTAGCTGGCTCCTGG
DGR24	GGTAAGAGGAAGATGGTTCGTGG	DGR49	GGTTGGGGGAATACGGCCGAGGG
DGR25	GGCACTTGGGGTGCGGGAGGGG	DGR50	GGCTCATGGTGTGCGGAATCCGG

Supplementary Table 4.2 50 randomly generated 23 nucleotide sequences used for CD analysis.

Name	Sequence	Name	Sequence
RANDOM1	GAATTGGCATGGGGTCTGACTCC	RANDOM26	GCACAGGTTCGTGGATGTCTTGT
RANDOM2	GCCCGGCGAAGAGCGGAGCAGCA	RANDOM27	GGGGTGGTTACTCGGAGTAGTTT
RANDOM3	TCCACGGCGGAAATATTTCATAA	RANDOM28	TGCTGGTTCAGTCCGTAACAGAT
RANDOM4	TCAAGTACCAGTTTAGCTGAGCA	RANDOM29	AAATGGAAACCGAACCTAAATCC
RANDOM5	TGAATAGCAGATCCGCCCTGGTC	RANDOM30	CACCATAGTTATAGTTACCTGAT
RANDOM6	CACCCTCACCCCCGAGTATTCGT	RANDOM31	AGGCCCTAAAATGCGGTAGGCCT
RANDOM7	GATAGTAAGTCCTAAACTTTCTG	RANDOM32	AATGAACGACACTTAACGCCTAT
RANDOM8	ACATGTCTATGATTATGTATGAA	RANDOM33	TGGTAATCATCGACACGGGTCAA
RANDOM9	TAGGATACAAATGCCGATAGCTC	RANDOM34	AATTTGTCAGCGAAGCGTGCTCA
RANDOM10	TAGGATACAAATGCCGATAGCTC	RANDOM35	TGACCCATATCTATCTGTGTCCC
RANDOM11	GATTGGATTAGTATGTCCCCTCT	RANDOM36	ATACGGGGGTTATCTGGTCTGTC
RANDOM12	TAAGTCTAGCGGGGTTGCACTCG	RANDOM37	ACAGATTGCGATTAATGCTAACG
RANDOM13	GGGAAGGAAACACCCGCACAGTC	RANDOM38	CTCGCGGTTAAGCCACCCCGGGT
RANDOM14	GGTTGCGGCGCAGAAACGAACTC	RANDOM39	CCGGCACCCGGACCTGTTACGTC
RANDOM15	AAATTCCTCGCTCCTGAACGTAC	RANDOM40	ATCGGGTCACCGTCCACTGCTTG
RANDOM16	GTATAAGGACAACAAAGCGTGCT	RANDOM41	AGACGCTCTTGAGACTTCTAACA
RANDOM17	CGGCTTGAAGTGAATTGCGTGAA	RANDOM42	GGTATTGGGAGTCGTATATTACT
RANDOM18	ATACCGGTTGCGCACACGGCATC	RANDOM43	ATTGTCCTGACAAGATTCGGCTC
RANDOM19	AGTCGCGGCCCGAGGGCTCGATC	RANDOM44	TTGTACAGTGTCGTATGCACCTG
RANDOM20	CTACGATTCACGCGGTCTAGCTT	RANDOM45	TTGTGTTTCCCCTTGACCGGAGT
RANDOM21	GCTCGCACGAGTGGTCTATTTTG	RANDOM46	CGAGGCGACGCTCCTTTTACGCT
RANDOM22	AAAGTCCTTGTTGACGGTAGAGC	RANDOM47	GTGATGCGGCCGAAGAATGAAAC
RANDOM23	CCGGGAAGAGAGGATGGGGCGCT	RANDOM48	ACGCGATAACCGGTTGAGTGTCT
RANDOM24	GTTTTATCGTCGGCGAAGTAGAA	RANDOM49	GGAACTCAGGTGTCTATAGCGAG
RANDOM25	CATCGGGTGGTGAACAAGGGTGA	RANDOM50	TCGAGAGAGATTTCAAGACGCAT



Supplementary Figure 4.1 Circular dichroism (CD) analysis of diguanine repeat sequences ($GGT_XGGT_XGGT_XGG$) with X=2 or 3. CD of sequence with two thymines in each loop region ($GGT_2GGT_2GGT_2GG$) is shown in red. Scans of all other combinations of two or three thymines in each loop are shown in black. Dotted lines are placed at 264 and 295 nm as ellipticities at these wavelengths indicate the presence of different arrangements of quadruplex structures.



Supplementary Figure 4.2 Average CD of 50 randomly generated diguanine repeat sequences. The CD spectra of 50 randomly generated diguanine repeat sequences based on $GGN_5GGN_5GGN_5GG$ were summed and divided by 50 to generate an average spectrum shown in red. 50 randomly generated 23 nucleotide sequences were scanned by CD and averaged for comparison and shown in black.

Supplementary Figure 4.3 Oligonucleotides used for ligation and PCR of DGR36. (A) Oligonucleotides used for ligation of DGR36 to 5′ and 3′ adapters (5PA and 3PA). Oligos 5TM and 3TM are used as templates. (B) Oligonucleotides used for PCR of ligated GGN_5 library.

Chapter 5:
Turning a Kinase Deoxyribozyme into a Sensor

5.1 AUTHOR'S PREFACE

The research project described in this chapter reports the conversion of a self-phosphorylating deoxyribozyme into a biosensor for GTP and Mn^{2+}. Chapters 2 and 3 of this thesis have described the characterization of a group of deoxyribozymes with different NTP substrate requirements. While these deoxyribozymes have been used to answer questions about in vitro selection and DNA structure, they also have properties that make them amenable to some practical applications. For instance, these deoxyribozymes are prime candidates for conversion to sensors for different NTP and dNTPs as they are dependent on specific NTPs for activity. Conversion of one deoxyribozyme, Dk2, into a GTP sensor was achieved through amplification of self-phosphorylation by rolling circle amplification (RCA), resulting in a lowering of detection limit for GTP from micromolar to low nanomolar. Optical detection was achieved through inclusion of the antisense sequence of a fluorogenic deoxyribozyme into the circular template used for RCA. The major finding of this work was that self-phosphorylating deoxyribozymes could be converted into NTP biosensors through amplification and coupling to a detection system. In a general sense, this shows that substrates for a deoxyribozyme can be converted to target for a sensor if a method is developed to couple deoxyribozyme activity to amplification and detection systems. These types of creative schemes will be necessary in the future to develop sensors for important targets.

This chapter is a modified version of a submitted research article. I am the first author on this publication and conducted all of the experiments presented within and wrote the manuscript. I also took a leading role in the conceptualization, experimental design and interpretation of the results obtained. Dr. Yingfu Li is listed as an author of this paper and provided many fruitful suggestions and guidance as well as assisting in editing of the final version of this manuscript.

5.2 ABSTRACT

Deoxyribozyme-based sensors are currently limited to a few DNA-catalyzed reactions whose targets are co-factors or allosteric regulators of catalytic activity. Creating sensors using deoxyribozymes whose substrates act as targets would increase the number of molecules that can be detected by this class of sensors. Self-phosphorylating deoxyribozymes process nucleoside triphosphates and have the potential to act as sensors for molecules such as ATP, CTP, GTP and UTP. We report a sensitive and selective deoxyribozyme-based GTP sensor consisting of a self-phosphorylating deoxyribozyme coupled to signal amplification by rolling circle amplification (RCA). With this coupling, the detection limit of this deoxyribozyme for GTP was increased from high micromolar to low nanomolar concentrations. By appending the antisense sequence of a fluorogenic RNA-cleaving deoxyribozyme to the self-phosphorylating deoxyribozyme, the RCA product could be used to generate a fluorescent signal. This sensor was further found be highly selective for GTP, over ATP, CTP and UTP. Utilizing the co-factor requirements of a self-phosphorylating deoxyribozyme, specific detection of Mn^{2+} over other divalent ions was also achieved. With several self-

phosphorylating deoxyribozymes available with different NTP and divalent metal ion co-factor requirements, there is potential to implement this sensor design in the development of sensors for many targets. This study further demonstrates that deoxyribozyme activity can generally be coupled to amplification and detection to convert substrates for deoxyribozymes into targets for sensors.

5.3 INTRODUCTION

Catalytic DNA molecules, known as deoxyribozymes, have been developed into sensors for targets as wide ranging as metal ions (Zhang et al., 2011) and specific species of bacteria (Ali et al., 2011). In most systems, deoxyribozyme-based sensors detect targets that act as co-factors or allosteric regulators of deoxyribozyme activity of a few well-known deoxyribozymes by inducing conformational changes. Background and false positive signals can arise in these systems due to folding of the deoxyribozyme into its active structure in the absence of target. Other molecules in the sensing solution can also mimic the deoxyribozyme reaction, especially in the case of RNA-cleaving deoxyribozymes, also leading to a false positive signal. An alternate method of detection would be to use substrates from deoxyribozyme-catalyzed reactions as targets for new sensors. This would allow for reduced background, as contaminant molecules are not likely to induce substrate processing and substrate specificity can be tuned by *in vitro* selection. Currently, many deoxyribozymes are available that catalyze different reactions modifying a range of different substrates and novel deoxyribozymes for other targets can be isolated by *in vitro* selection. If the activity of these deoxyribozymes can be coupled to

efficient reporting systems, they may serve as the foundation for the development of a plethora of new sensors.

Self-phosphorylating deoxyribozymes, or deoxyribozyme kinases, are catalytic DNA molecules that can transfer a γ-phosphate from an NTP substrate to their own 5′ ends. Through *in vitro* selection experiments, numerous self-phosphorylating deoxyribozymes with wide ranging properties have been isolated and in some cases well-characterized (Li et al., 1999; Wang et al., 2002; Achenbach et al., 2005; McManus et al., 2007; McManus et al., 2008). Such deoxyribozymes possess very specific substrate and co-factor requirements. These deoxyribozyme kinases, which can discriminate between a particular NTP and closely related analogues, appear to be ideal candidates for the development of sensors for NTP molecules. One such deoxyribozyme, Dk2, can transfer a phosphate from guanosine-5′-triphosphate (GTP) to its 5′ end at a rate of 0.2 min^{-1} in presence of Mn^{2+} (Wang et al., 2002; Achenbach et al., 2005). It was found that this deoxyribozyme was highly specific for GTP as a substrate and could not process other purine-based nucleoside triphosphates, including several closely related GTP analogues (McManus et al., 2007). As this deoxyribozyme can specifically process GTP, it is a prime candidate for conversion into a GTP sensor.

GTP plays many critical roles in the cell. In addition to being a building block for RNA, GTP also plays a major role in energy utilization. During protein translation, GTP provides the energy for binding of aminoacyl tRNAs and for translocation of the ribosome (Nyborg et al., 1996). GTP created during the Krebs cycle is also readily converted to ATP (Sanadi et al., 1954). Furthermore, GTP is important for intercellular

communication, coordination of cellular processes and the maintenance of structural stability. Signal transduction through G-protein coupled receptors is mediated by the conversion of GTP to GDP through coupled GTPases (Hilgenfeld, 1995). GTP is also essential in microtubule stability, as hydrolysis of GTP to GDP at the end of a microtubule leads to depolymerisation (Wade, 2007). Early assays for the detection of this important triphosphate molecule involved the chromographic separation of guanine nucleotides from the more abundant adenosine nucleotides followed by enzymatic conversion of GTP to ATP, which could be measured by luciferase-based detection (Pogson et al., 1979). In addition to enzymatic detection, GTP sensors have also been designed using artificial receptors. In these systems, binding of GTP to the receptor displaces or changes the environment of an associated fluorescent dye leading to a change in fluorescence. Receptors used in GTP sensors include guanidinium tripeptide receptors (McCleskey et al., 2003), benzimidazolium receptors (Kwon et al., 2004; Wang et al., 2006), and cyclophane receptors (Neelakandan et al., 2006; Ahmed et al., 2011). These sensors can typically detect GTP in the micromolar to millimolar range. With regards to functional nucleic acids, aptamers that can bind to GTP have been isolated (Davis et al., 2002; Huang et al., 2003). Ribonucleopeptide-based sensors have been developed for GTP, in which an RNA aptamer domain binds GTP, changing the environment of an associated fluorescently labeled peptide to generate a signal (Hagihara et al., 2006).

Currently, self-phosphorylating activity of Dk2 using GTP is indirectly detected through a ligation assay as 5′ phosphorylated Dk2 can serve as a substrate for T4 DNA ligase and ligated to an acceptor DNA molecule, while inactive deoxyribozymes lacking

the 5′ phosphate will not be ligated. The active sequences can then be electrophoretically separated from unreacted sequences based on their increased size. Using this system, it is possible to detect micromolar amounts of GTP (K_M = 550 μM) (Wang et al., 2002). While this method of detection has been useful for structural and kinetic characterization of this deoxyribozyme, and has enabled the detection of micromolar to millimolar amounts of GTP, we do not believe that the lowest detection limit for GTP has been achieved using this system. During the early rounds of the *in vitro* selection for the isolation of this self-phosphorylating deoxyribozyme, no ligation product was visible, yet the region where the ligated product would have appeared on the gel could be eluted and amplified by PCR. This demonstrates that though sequences are being ligated, the key factor impeding their detection is the inability to visualize trace amounts of ligated products. To improve the sensitivity and thereby increase the detection limit, we have modified our ligation assay rendering its product suitable for signal enhancement through rolling circle amplification (RCA). RCA is a process that involves continuous polymerization of a primer from a circular template, producing long single-stranded DNA molecules. Originally discovered in replication of certain viruses, this technique has been adapted for the amplification of small single-stranded circular templates (Fire et al., 1995). This system has been exploited to detect many DNA and RNA targets with the development of padlock probes (reviewed in (Nilsson et al., 2006)). In this system, linear probes are designed such that their 3′ and 5′ ends will hybridize to a target sequence. In the presence of a ligase enzyme, these probes will ligate into a closed circle, which can be used as a template for RCA by a suitable polymerase such as ϕ29

DNA polymerase. A modified version of this padlock system has been used to detect ATP. In this case, a ligase deoxyribozyme allosterically controlled by ATP was used, and the linear probe would circularize following ATP binding (Cho et al., 2005).

As RCA has been shown to be a powerful method for amplification, we examined the possibility of employing this technique to sense low levels of Dk2 self-phosphorylation and increase the detection limit for GTP. The scheme depicting our system is outlined in Figure 5.1. First, self-phosphorylation is carried out in the presence of GTP and self-phosphorylation buffer (50 mM HEPES, 400 mM NaCl, 100 mM KCl, 10 mM $MnCl_2$, pH = 7.0). This is followed by the addition of a splint DNA oligonucleotide that is complementary to the 5′ and 3′ end of the deoxyribozyme. This will bring the phosphorylated 5′ end into proximity to the 3′ end of the same deoxyribozyme strand. Addition of T4 DNA ligase will lead to self-ligation of the deoxyribozyme into a single-stranded circle. This circularized DNA can then be used as a template for RCA upon the addition of ϕ29 DNA polymerase and 0.5 mM of each dNTP with the splint now acting as a primer for the RCA reaction. The resulting RCA reaction will produce long repetitive single-stranded products of more than 10 kilobases from phosphorylated deoxyribozyme templates of less than 100 nucleotides, thus greatly amplifying the initial self-phosphorylation.

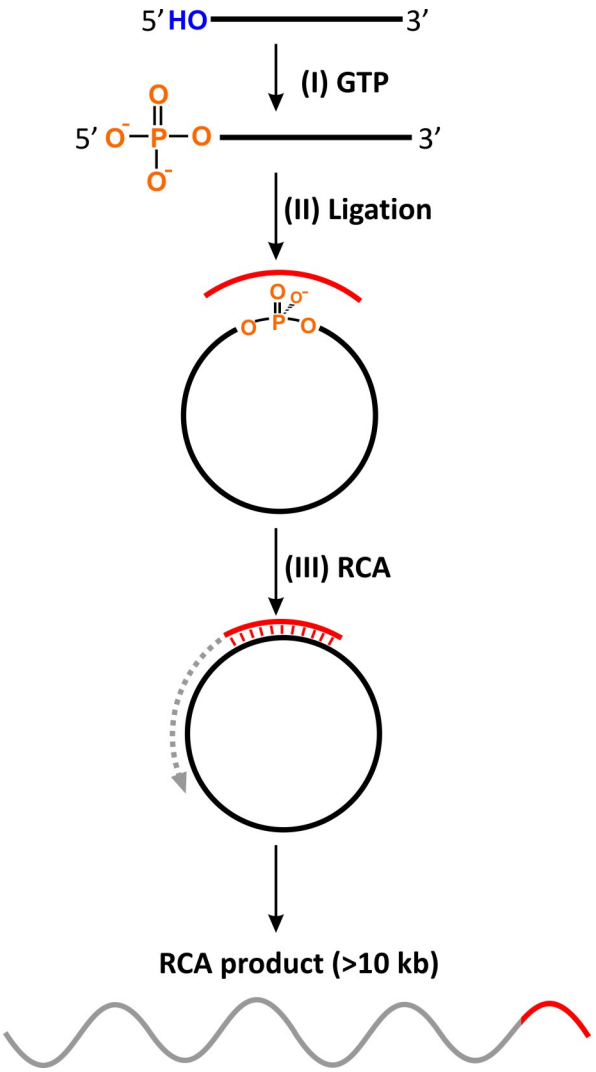


Figure 5.1 Amplifying self-phosphorylation signal through circular ligation and rolling circle amplification.

5.4 EXPERIMENTAL SECTION

5.4.1 Oligonucleotides and Materials

DNA oligonucleotides were prepared using standard phosphoramidite chemistry (W. M. Keck Facility, Yale University, New Haven, CN, USA; and Integrated DNA Technologies, Coralville, IA, USA). Each oligonucleotide was purified by 10% denaturing (8 M urea) PAGE, and its concentration was determined spectroscopically. T4 DNA ligase and ϕ29 DNA polymerase were purchased from MBI Fermentas, Burlington, ON, Canada. [α-^{32}P]deoxy-GTP was acquired from Perkin Elmer, Woodbridge, ON, Canada. All other chemicals were obtained from Sigma-Aldrich, Oakville, ON, Canada.

5.4.2 Self-phosphorylation Reactions

Self-phosphorylation assays were carried out using 500 nM DNA. DNA was first heated to 90 ^{o}C for 1 minute and cooled to room temperature. Reactions were initiated by mixing equal amounts of DNA and 2x self-phosphorylation buffer containing 100 mM HEPES, 800 mM NaCl, 200 mM KCl, 20 mM $MnCl_2$ and 200 μM GTP. For the divalent metal ion tests $MnCl_2$ was replaced with 20 mM $CaCl_2$, $MgCl_2$ or $ZnCl_2$. For the GTP sensitivity tests, the amount of GTP used in the 2X buffer was from 2 nM to 4 μM. For the NTP specificity tests GTP was replaced with 200 μM ATP, CTP and UTP. Reactions were incubated at room temperature for 30 minutes and used immediately in the ligation reaction.

5.4.3 Circular Ligation Reactions

Self-phosphorylation reactions were diluted 5x for ligation reactions. Ligation reactions contained 3 μM circular ligation splint DNA (SPDK2 for Dk2, SPDKMgZ for

Dk2/MgZ′, see Supplementary Table 5. 1), 1x T4 DNA ligase buffer (MBI Fermentas), and 0.01 U/μL T4 DNA ligase. DNA was heated to 90 ^{o}C and cooled to room temperature. Buffer and ligase were then added and the reaction was incubated for 45 min at room temperature. The reaction was heated to 90 ^{o}C for 5 min to deactivate the ligase. These reactions were then immediately used for rolling circle amplification.

5.4.4 RCA Assays

Ligation reactions were diluted 5x for rolling circle amplification (RCA) reactions. Reactions were performed in 1x ϕ29 DNA polymerase buffer (MBI fermentas) with 500 μM ATP, GTP, CTP, UTP and 0.01 U/ μL ϕ29 DNA polymerase. For Figure 5.2, 1 μL of [α-^{32}P]deoxy-GTP was also added to a 50 μL RCA reaction. RCA reactions were incubated at 30 ^{o}C for 3 h. Reactions were then incubated at 70 ^{o}C for 20 min to deactivate the ϕ29 DNA polymerase.

5.4.5 Fluorescent Detection of RCA with MgZ and RS28

After heat inactivation of ϕ29 DNA polymerase, RCA product created from Dk2/MgZ circular DNA template was incubated with 2 μM RS28 substrate (see Supplementary Table 5. 1 for Dk2/MgZ′ and RS28 sequences) at room temperature. Fluorescence changes over time were monitored on a BioRad CFX96 Optical Reaction Module, as per the manufacturer's instructions. Fluorescence was normalized by subtracting the fluorescence of uncleaved RS28.

5.5 RESULTS AND DISSCUSSION

We first tested whether an RCA product could be formed following self-phosphorylation and circular ligation of Dk2 using our experimental scheme. The sequence used for our study was a truncated 63-nucleotide construct previously shown to be the shortest variant of Dk2 capable of robust self-phosphorylating activity. Self-phosphorylation was carried out by incubating Dk2 with 100 μM GTP in self-phosphorylation buffer. Following phosphorylation, circular ligation was carried out using T4 DNA ligase and a splint oligonucleotide complementary to the 5′ and 3′ ends of Dk2. RCA was carried out in presence of ϕ29 DNA polymerase, dNTPs and [α-^{32}P]deoxy-GTP to allow for detection of the RCA product. The products were subsequently resolved on a 10% denaturing PAGE gel. The results of this assay are shown in Figure 5.2. Carrying out these steps produced a large product, as seen in lane 4. In contrast, RCA product is not observed in the absence of GTP (lane 1), or if ligase is omitted (lane 2), only a short faint product resulting from polymerization of linear Dk2 is apparent. No product is seen when ϕ29 DNA polymerase is not added during the polymerization step (lane 3).

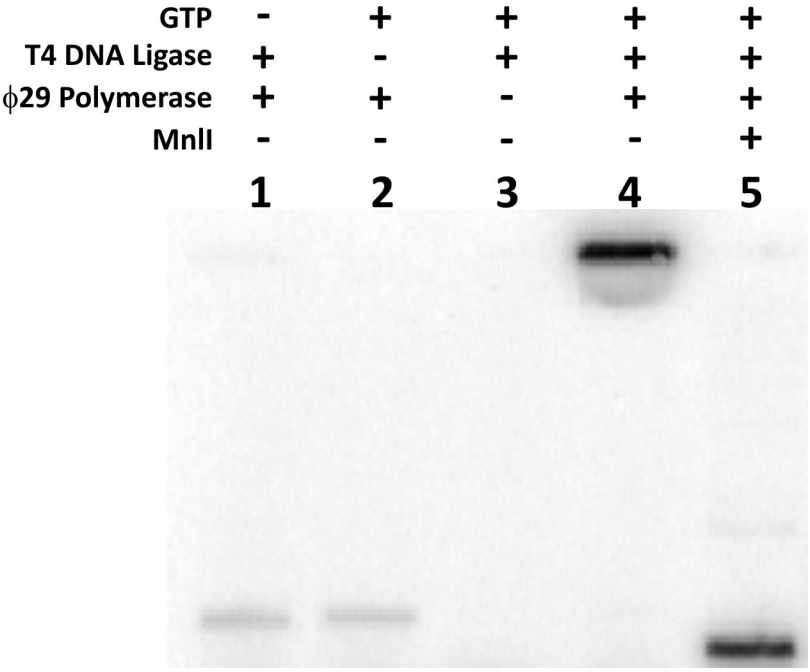


Figure 5.2 Rolling circle amplification of Dk2. Dk2 was incubated sequentially with GTP, T4 DNA ligase and ϕ29 DNA polymerase (lane 4). The identity of the RCA product was confirmed by MnlI digestion (lane 5).

To confirm the identity of the product as a concatameric complementary copy of the circularized Dk2, the RCA product was subjected to restriction enzyme digestion. Addition of a nucleotide sequence identical to a portion of Dk2 created an MnlI restriction site in the RCA product. If the RCA product is a repetitive complement of Dk2, digestion with MnlI will cleave the RCA product into monomeric units of 63 nucleotides. As seen in lane 5 of Figure 5.2, MnlI did indeed result in the loss of the large RCA product and the formation of a monomeric product.

To assess the sensitivity of our system at low concentrations of GTP, we tested our system over a range of 1 nM to 2 μM GTP. As seen in Figure 5.3, a RCA product larger

than 10 kilobases is visible at as low as 5 nM GTP and increases in intensity to 2 μM GTP. Taking into account the reaction volume of as little as 2 μL, 1 nM of GTP corresponds to attomole quantities demonstrating that this is a sensitive sensor for GTP.

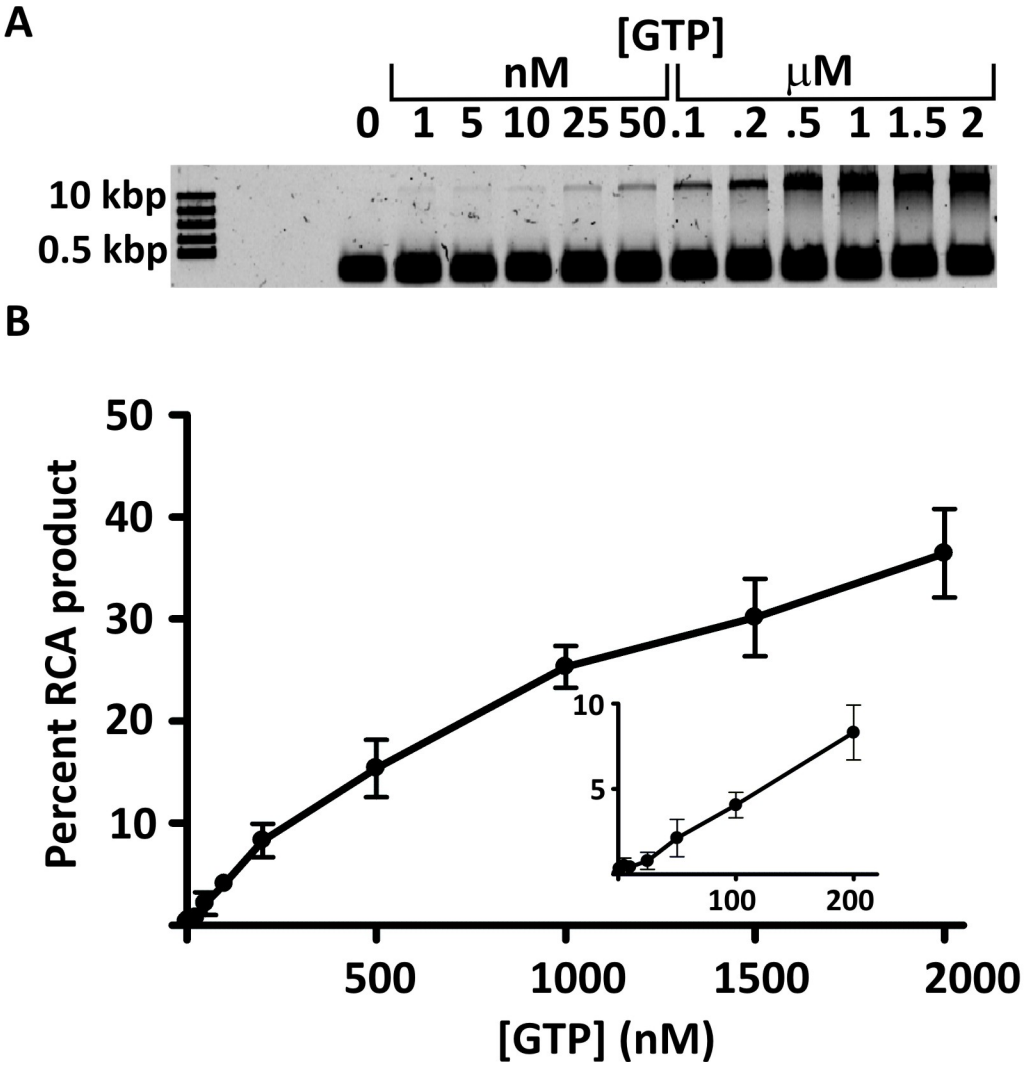


Figure 5.3 GTP sensitivity after RCA. (A) Agarose gel electrophoresis of RCA reactions after self-phosphorylation with differing amounts of GTP and circular ligation. (B) Plot of the percent of RCA product observed compared to the concentration of GTP used for self-phosphorylation. Enlargement of graph from 0 to 200 nM GTP shown in insert.

To convert this RCA amplified self-phosphorylating deoxyribozyme into a sensor for GTP, it is necessary to couple it to a detectable readout in lieu of relying on gel electrophoresis. Several methods are available for detection of an RCA product. Fluorescent detection can be achieved with the use of molecular beacons (Nilsson et al., 2002; Smolina et al., 2004), while colorimetric detection has been made possible with deoxyribozyme-modulated peroxidase activity assays (Cheglakov et al., 2007) and organic dyes (Ali et al., 2009). For this study, we chose to use a novel detection method incorporating a fluorogenic deoxyribozyme into the RCA product as shown in Figure 5.4. This method takes advantage of both the sensitivity of fluorescent detection and the processivity of a DNA catalyzed reaction. The deoxyribozyme we chose for fluorescent detection was MgZ, a fluorescence generating RNA-cleaving deoxyribozyme (Chiuman et al., 2007). This deoxyribozyme cleaves a chimeric DNA/RNA substrate containing a fluorescein and DABCYL dye flanking an RNA residue that acts as the cleavage site. In the uncleaved state the DABCYL quenches the fluorescence of the fluorescein. Once cleaved, portions containing the fluorescein and DABCYL are separated and the quenching effect of DABCYL is lost resulting in an increase in fluorescence. To allow for expression of MgZ in the RCA product of Dk2 circularization, a Dk2 construct was made in which the antisense of MgZ (MgZ′) was appended to the to the 3′ end of Dk2. The 3′ end was chosen for adding MgZ′ as 40 additional nucleotides were present at the 3′ end in the originally selected Dk2 sequence so addition of this extra sequence should not effect the folding of Dk2 into its active structure.

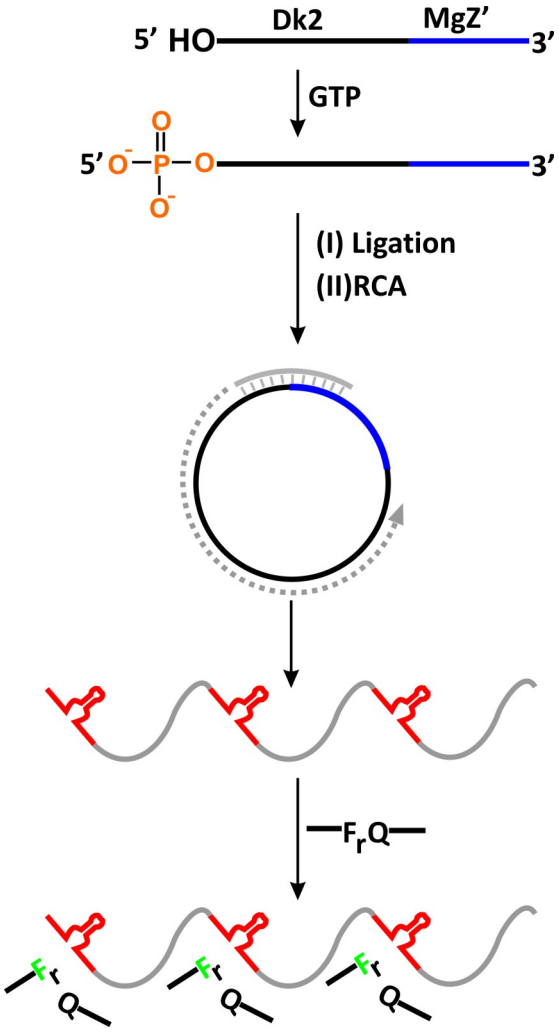


Figure 5.4 Fluorescent reporting of amplification using the MgZ fluorogenic deoxyribozyme. Scheme for generation of RCA products containing multiple MgZ RNA-cleaving fluorescence-generating deoxyribozymes.

Similar to the Dk2 construct in used in Figure 5.2, this construct was able to produce an RCA product after incubation with GTP and ligation, as shown in Figure 5.5a. This RCA product, which is a concatamer of the antisense of the Dk2/MgZ′, will contain multiple copies of the catalytic MgZ sequence. After RCA and heat inactivation, a

substrate for MgZ can be added and will be cleaved by the multiple copies of MgZ in the RCA product producing an increase in fluorescence. As shown in Figure 5.5b, an increase in fluorescence is observed over time when GTP is added during the self-phosphorylation step (red line), while no fluorescence increase is observed when no GTP is provided or when no ligase or polymerase is added during the necessary reaction steps. This is confirmed in the inserted gel of Figure 5.5b, as substrate cleavage is only seen when all reaction components are present (lane 4).

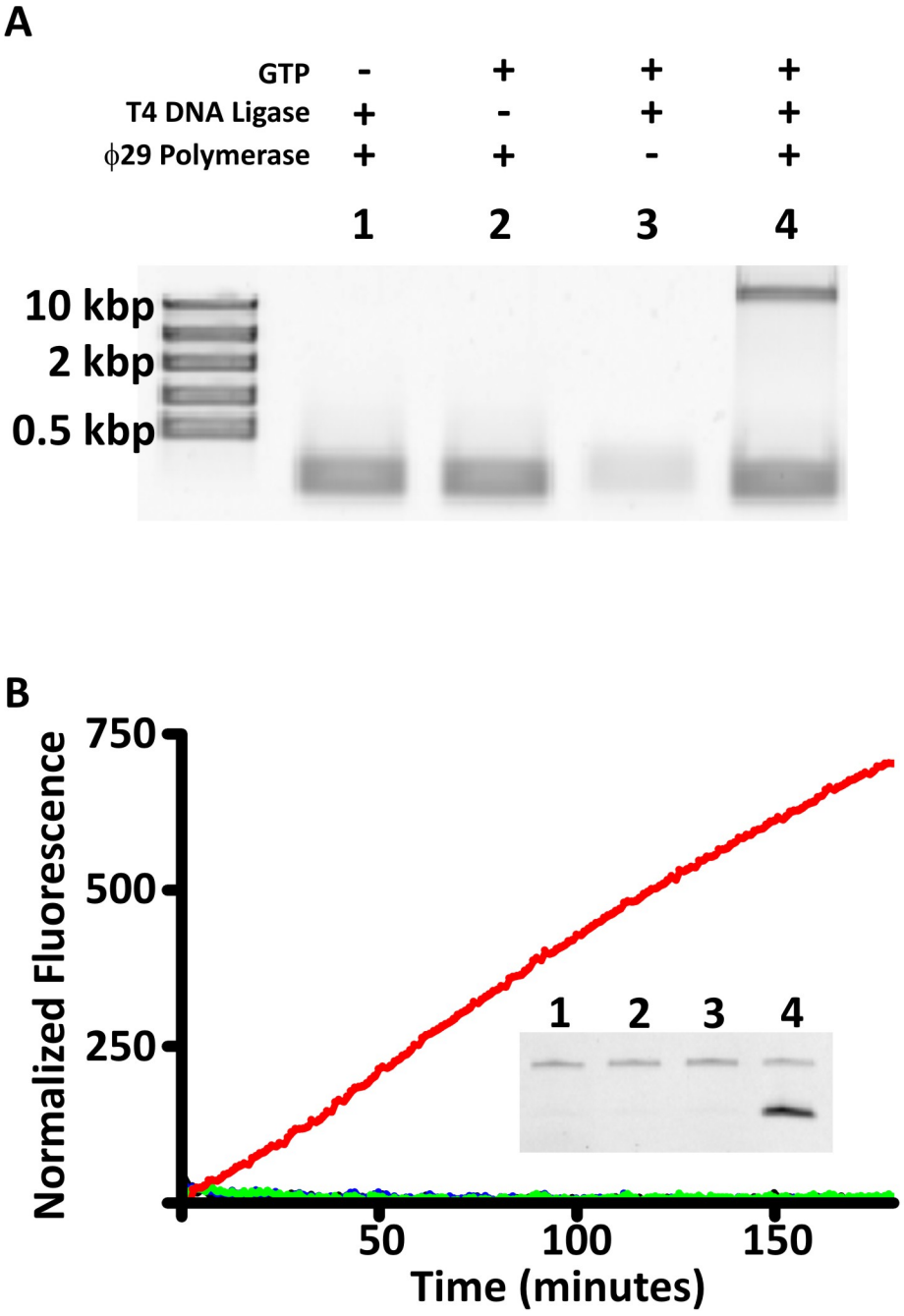


Figure 5.5 (A) RCA test after self-phosphorylation and circular ligation of Dk2/MgZ′. (B) Fluorescence generation upon incubation of RS28 substrate with Dk2/MgZ RCA product after self-phosphorylation and ligation. A time-dependent fluorescence increase is seen when all components are added (red line), but not in the absence of GTP (black

line), ligase (blue line), or ϕ29 DNA polymerase (green line). RS28 cleavage is shown in inserted gel (lanes are labeled as in part a).

After establishing that RCA and fluorogenic deoxyribozyme activity allowed for the detection of GTP though self-phosphorylation, we assessed the selectivity of our system between the biologically relevant ribonucleoside triphosphates: ATP, CTP, GTP and UTP. We implemented the same reaction protocol described above. An RCA product was only observed when DK2/ MgZ′ was incubated with GTP and not with any of the other three NTPs (Supplementary Figure 5.1a). When the product of the RCA reactions are incubated with RS28 substrate an increase in fluorescence over time is observed with GTP, while no increase in fluorescence is seen when Dk2/MgZ′ is incubated with ATP, CTP or UTP, as shown in Figure 5.6a. Gel analysis of the RS28 substrate after incubation showed this fluorescence increase to be the result of RS28 cleavage as shown in Supplementary Figure 5.1b.

While the system described has been used as an NTP sensor, self-phosphorylating deoxyribozymes also have specificity for divalent metal ion cofactors used during catalysis, suggesting they could also be used as sensors for divalent metal ions. Dk2, which was selected using Mn^{2+} in its selection buffer, was found to be specific for manganese ions over other divalent metal ion cofactors (Wang et al., 2002). To test whether our system could be used to detect the presence of Mn^{2+}, Dk2/MgZ′ was incubated in reaction buffer with Ca^{2+}, Mg^{2+}, Mn^{2+}, or Zn^{2+}, prior to circular ligation, and RCA. Consistent with the previous ligation-based detection, an RCA product was only observed in the presence of Mn^{2+} and not with the other divalent metal ions tested

(Supplementary Figure 5.2a). When incubated with RS28, fluorogenic substrate cleavage was observed (Supplementary Figure 5.2b) and could be monitored over time, as shown in Figure 5.6b, when self-phosphorylation is performed in the presence of Mn^{2+}, but not in the presence of the other divalent metal ions tested. As several self-phosphorylating deoxyribozymes have been isolated with different divalent metal cofactor requirements, this system has the potential to be used for detection of divalent metal ions of choice.

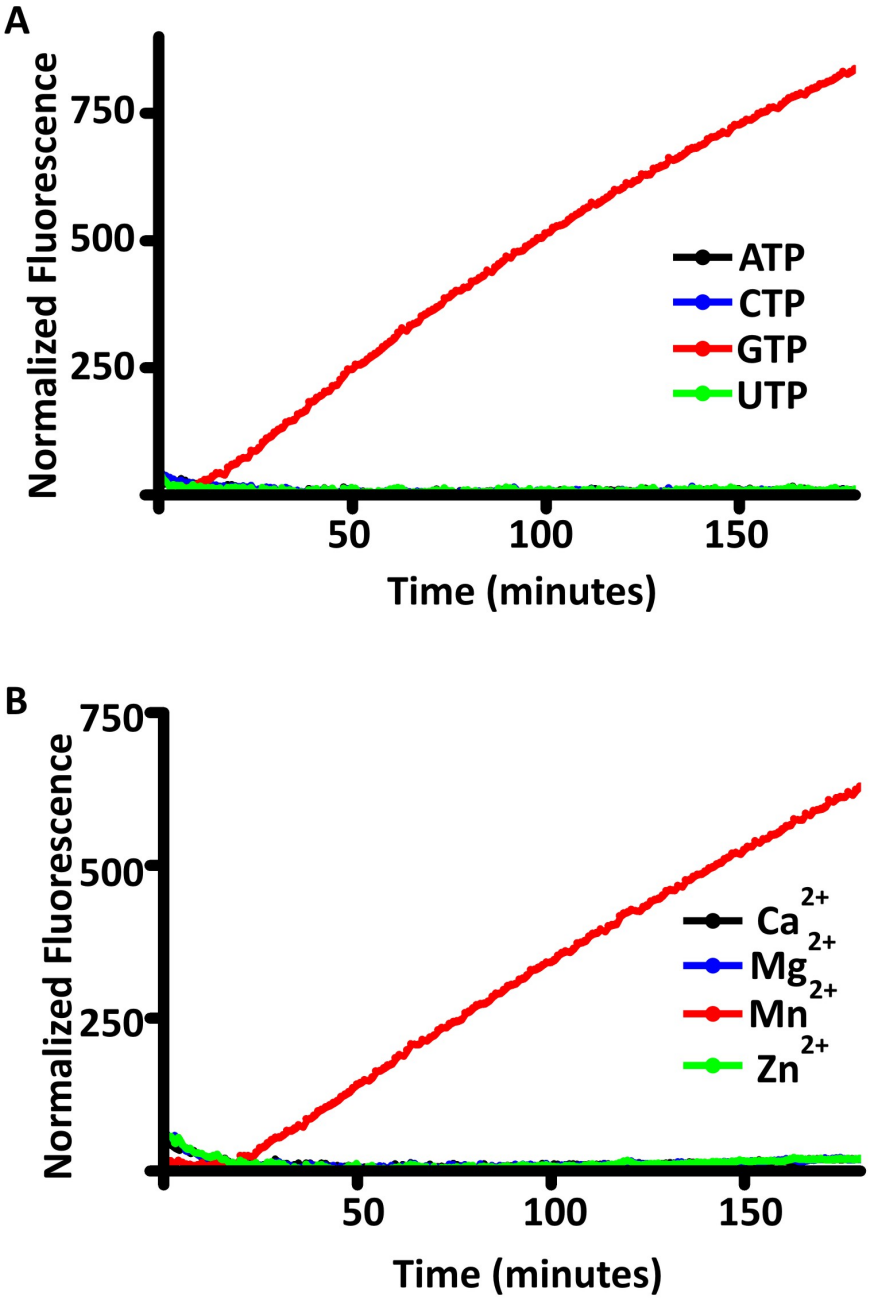


Figure 5.6 Specificity of fluorescent sensor. (A) NTP specificity test. Self-phosphorylation of Dk2/MgZ′ was performed in the presence of different NTPs followed by circular ligation, RCA amplification and RS28 cleavage. (B) Divalent metal ion specificity test. The same procedure was followed using reaction buffer containing different divalent metal ions during the self-phosphorylation of Dk2/MgZ′.

5.6 CONCLUSIONS

Our study has shown that it is possible to convert a deoxyribozyme substrate into a target for a novel sensor system. With the sensitivity and selectivity of our system, it has the potential to be used in many applications, both for research and practical purposes. While the RCA methodology was being developed, methods for reporting the polymerization of an RCA product were also being diversified and improved. Our current system can be readily coupled to these and other detection platforms, as the application dictates, in order to produce an efficient and facile sensor. It is also noteworthy that our system was performed without any purification between any of the steps from self-phosphorylation to fluorescent detection. Thus, this sensor system has the potential to be developed into a kit format with the five step procedure being carried out by repeated dilutions after each reaction with the components of the next reaction, thereby making this system amenable for real-world applications. More generally, we have demonstrated that deoxyribozyme substrates can be converted into targets for sensors by amplification of the deoxyribozyme reaction and coupling to a detection system. As it is becoming more and more apparent that deoxyribozymes can catalyze any reaction for which a selection scheme can be developed, the number of possible deoxyribozyme-based sensors is only dependent on the imagination of the researcher.

5.7 REFERENCES

Achenbach, J. C., G. A. Jeffries, S. A. McManus, L. P. Billen, and Y. Li (2005). Secondary-structure characterization of two proficient kinase deoxyribozymes. Biochemistry 44: 3765-3774.

Ahmed, N., B. Shirinfar, I. Geronimo, and K. S. Kim (2011). Fluorescent imidazolium-based cyclophane for detection of guanosine-5'-triphosphate and I(-) in aqueous solution of physiological pH. Org Lett 13: 5476-5479.

Ali, M. M., S. D. Aguirre, H. Lazim, and Y. Li (2011). Fluorogenic DNAzyme probes as bacterial indicators. Angew Chem Int Ed Engl 50: 3751-3754.

Ali, M. M., and Y. Li (2009). Colorimetric sensing by using allosteric-DNAzyme-coupled rolling circle amplification and a peptide nucleic acid-organic dye probe. Angew Chem Int Ed Engl 48: 3512-3515.

Cheglakov, Z., Y. Weizmann, B. Basnar, and I. Willner (2007). Diagnosing viruses by the rolling circle amplified synthesis of DNAzymes. Org Biomol Chem 5: 223-225.

Chiuman, W., and Y. Li (2007). Simple fluorescent sensors engineered with catalytic DNA 'MgZ' based on a non-classic allosteric design. PLoS One 2: e1224.

Cho, E. J., L. Yang, M. Levy, and A. D. Ellington (2005). Using a deoxyribozyme ligase and rolling circle amplification to detect a non-nucleic acid analyte, ATP. J Am Chem Soc 127: 2022-2023.

Davis, J. H., and J. W. Szostak (2002). Isolation of high-affinity GTP aptamers from partially structured RNA libraries. Proc Natl Acad Sci U S A 99: 11616-11621.

Fire, A., and S. Q. Xu (1995). Rolling replication of short DNA circles. Proc Natl Acad Sci U S A 92: 4641-4645.

Hagihara, M., M. Fukuda, T. Hasegawa, and T. Morii (2006). A modular strategy for tailoring fluorescent biosensors from ribonucleopeptide complexes. J Am Chem Soc 128: 12932-12940.

Hilgenfeld, R. (1995). Regulatory GTPases. Curr Opin Struct Biol 5: 810-817.

Huang, Z., and J. W. Szostak (2003). Evolution of aptamers with a new specificity and new secondary structures from an ATP aptamer. RNA 9: 1456-1463.

Kwon, J. Y., N. J. Singh, H. N. Kim, S. K. Kim, K. S. Kim, and J. Yoon (2004). Fluorescent GTP-sensing in aqueous solution of physiological pH. J Am Chem Soc 126: 8892-8893.

Li, Y., and R. R. Breaker (1999). Phosphorylating DNA with DNA. Proc Natl Acad Sci U S A 96: 2746-2751.

McCleskey, S. C., M. J. Griffin, S. E. Schneider, J. T. McDevitt, and E. V. Anslyn (2003). Differential receptors create patterns diagnostic for ATP and GTP. J Am Chem Soc 125: 1114-1115.

McManus, S. A., and Y. Li (2008). A deoxyribozyme with a novel guanine quartet-helix pseudoknot structure. J Mol Biol 375: 960-968.

McManus, S. A., and Y. Li (2007). Multiple occurrences of an efficient self-phosphorylating deoxyribozyme motif. Biochemistry 46: 2198-2204.

Neelakandan, P. P., M. Hariharan, and D. Ramaiah (2006). A supramolecular ON-OFF-ON fluorescence assay for selective recognition of GTP. J Am Chem Soc 128: 11334-11335.

Nilsson, M., F. Dahl, C. Larsson, M. Gullberg, and J. Stenberg (2006). Analyzing genes using closing and replicating circles. Trends Biotechnol 24: 83-88.

Nilsson, M., M. Gullberg, F. Dahl, K. Szuhai, and A. K. Raap (2002). Real-time monitoring of rolling-circle amplification using a modified molecular beacon design. Nucleic Acids Res 30: e66.

Nyborg, J., and M. Kjeldgaard (1996). Elongation in bacterial protein biosynthesis. Curr Opin Biotechnol 7: 369-375.

Pogson, C. I., S. V. Gurnah, and S. A. Smith (1979). A sensitive and specific assay for GTP and GDP in tissue extracts. Int J Biochem 10: 995-1000.

Sanadi, D. R., M. Gibson, and P. Ayengar (1954). Guanosine triphosphate, the primary product of phosphorylation coupled to the breakdown of succinyl coenzyme A. Biochim Biophys Acta 14: 434-436.

Smolina, I. V., V. V. Demidov, C. R. Cantor, and N. E. Broude (2004). Real-time monitoring of branched rolling-circle DNA amplification with peptide nucleic acid beacon. Anal Biochem 335: 326-329.

Wade, R. H. (2007). Microtubules: an overview. Methods Mol Med 137: 1-16.

Wang, S., and Y. T. Chang (2006). Combinatorial synthesis of benzimidazolium dyes and its diversity directed application toward GTP-selective fluorescent chemosensors. J Am Chem Soc 128: 10380-10381.

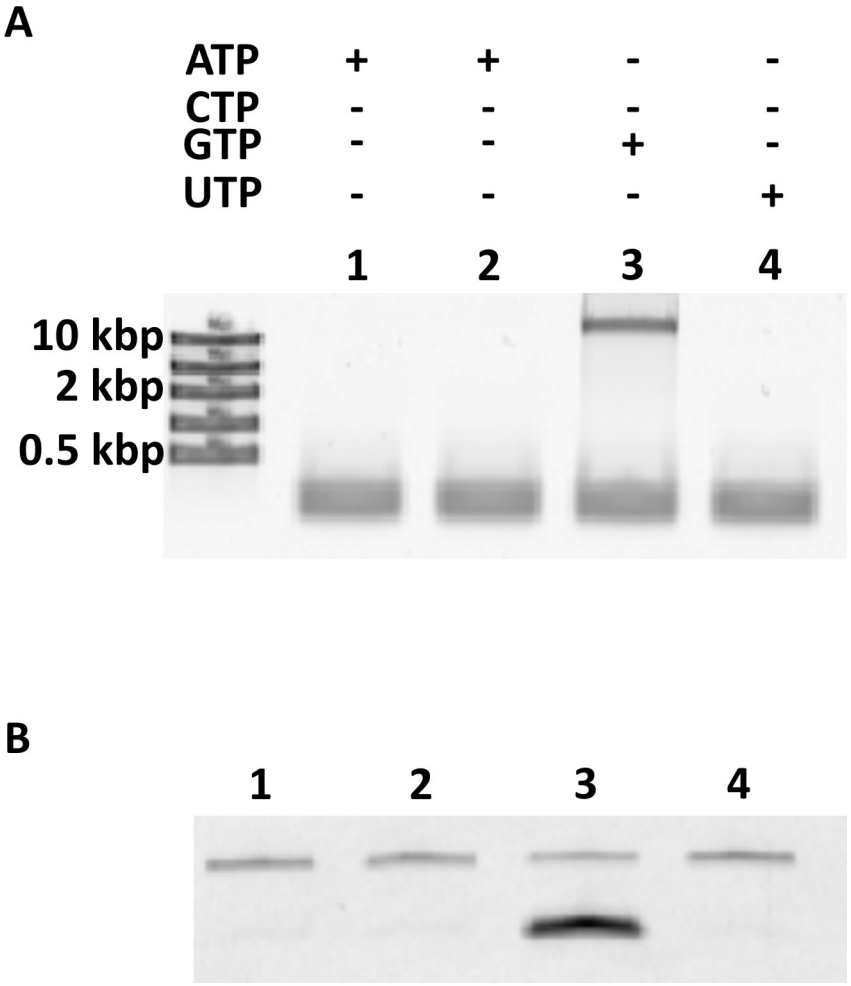
Wang, W., L. P. Billen, and Y. Li (2002). Sequence diversity, metal specificity, and catalytic proficiency of metal-dependent phosphorylating DNA enzymes. Chem Biol 9: 507-517.

Zhang, X. B., R. M. Kong, and Y. Lu (2011). Metal ion sensors based on DNAzymes and related DNA molecules. Annu Rev Anal Chem (Palo Alto Calif) 4: 105-128.

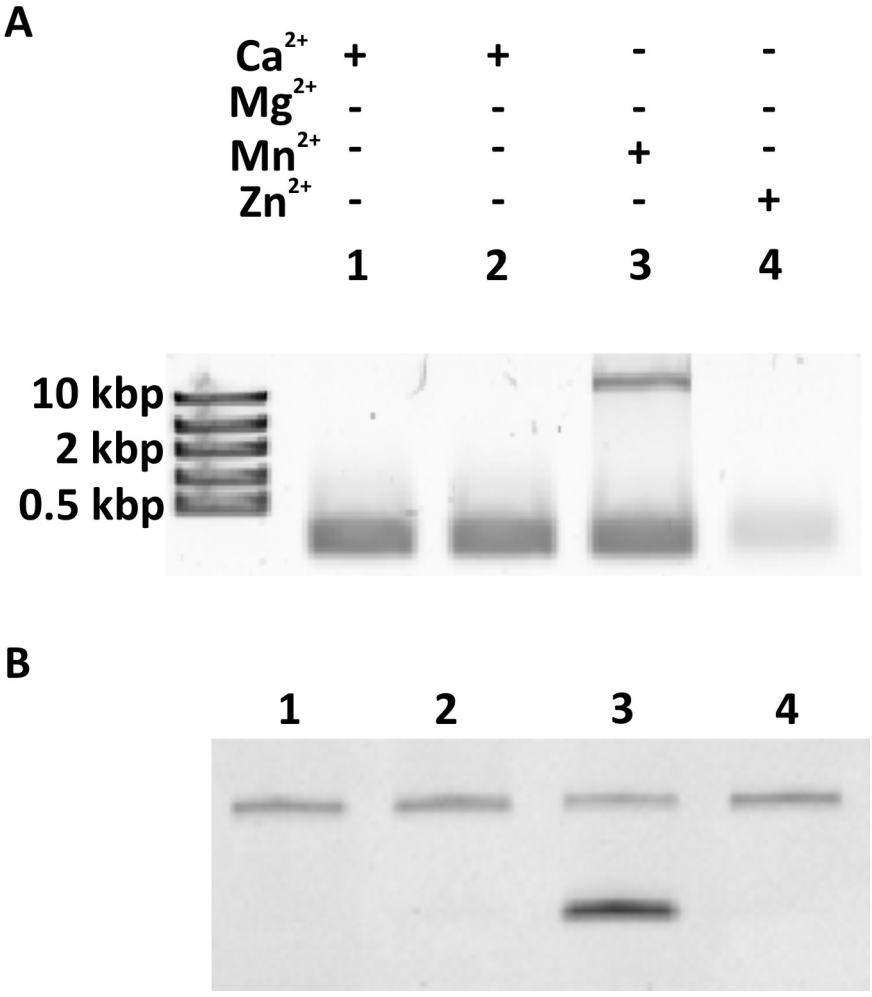
5.8 SUPPLEMENTARY INFORMATION

Supplementary Table 5.1 Sequences used.

Sequence Name	Sequence	Description
Dk2	5′GGAGAGGTAACGACACCGGGTGAG TGCTATCGCTTGCGATAGTACCACTTT CCCCGCGCGGTG3′	Self-phosphorylating deoxyribozyme active in the presence of GTP and $MnCl_2$
SPDk2	5′TTACCTCTCCCACCGCGCGG3′	Splint oligonucleotide used for circular ligation of phosphorylated Dk2
Dk2/MgZ′	5′GGAGAGGTAACGACACCGGGTGAG TGCTATCGCTTGCGATAGTACCACTTT CCCCGCGCGGTGCCTAGCATAGCCTC CCAAAATATCCTATATTCGGCCCCGA CCTGGTTC3′	Chimera of Dk2 and the antisense sequence of MgZ RNA-cleaving deoxyribozyme
SPDK2/MgZ′	5′AGGCACCGCGCGGGGAAAGT3′	Splint oligonucleotide used for circular ligation of phosphorylated DK2/MgZ′
RS28	5′ACTCTTCCTAGCFrAQGGTTCGATCA AGA3′	Fluorogenic substrate, F= fluorescein, rA= riboadenosine, Q= DABCYL
COMMNLI	CGGTGCCAGCGGAGAGGTAA	Reverse complement of a section of RCA product of circularized Dk2, creating a restriction site for MnlI



Supplementary Figure 5.1 NTP test and circular ligation. A) RCA test after self-phosphorylation with ATP, CTP, GTP or UTP and circular ligation of Dk2/MgZ′. B) RS28 cleavage after RCA test. Lanes are labeled as in part a.



Supplementary Figure 5.2 RCA and RS28 cleavage following divalent metal ion test and circular ligation. A) RCA test after self-phosphorylation with reaction buffer containing Ca^{2+}, Mg^{2+}, Mn^{2+}, Zn^{2+} and circular ligation of Dk2/MgZ ′. B) RS28 cleavage after RCA test. Lanes are labeled as in part a.

Chapter 6:
Discussion and Future Directions

This thesis described the characterization and utilization of a group of self-phosphorylating deoxyribozymes. The projects included the structural alignment of three unique deoxyribozyme sequences, the characterization of a novel deoxyribozyme structural motif, the development of a DNA library enriched in two-tiered quadruplex motifs, and the creation of NTP sensors using amplified self-phosphorylation activity. These studies have demonstrated that DNA can be used as an efficient catalyst and that DNA can be versatile and use different structural arrangements to perform the same function.

The first project described in this dissertation investigated three self-phosphorylating deoxyribozymes with similar substrate and cofactor requirements for structural similarities. The data from several structural probing techniques revealed that all three deoxyribozymes contained a common secondary structure with a two-way helical junction and a putative helix between the reaction site and a central conserved sequence. Based on differing location of the conserved structural motifs within each of the sequences, these deoxyribozymes likely arose independently in the initial library, as the selection scheme did not allow for insertions and deletions of nucleotides that would be necessary for the relocation of a sequence motif. In fact, on the primary sequence level, these deoxyribozyme did not appear to have many striking similarities and some of the smaller conserved sequence motifs were only identified after structural characterization. This repeated independent isolations of the same functional nucleic acid sequence has

been recorded previously for a DNA aptamer for ATP (Huizenga and Szostak 1995; Nutiu and Li 2005), the hammerhead ribozyme (Conaty et al. 1999; Tang and Breaker 2000; Salehi-Ashtiani and Szostak 2001), and an RNA-cleaving deoxyribozyme (Faulhammer and Famulok 1996; Santoro and Joyce 1997; Li et al. 2000; Cruz et al. 2004; Schlosser and Li 2004).

As has been proposed for these systems, it is likely that multiple occurrences of the self-phosphorylating deoxyribozyme structure suggest that this is the simplest structural solution for DNA-based self-phosphorylation using GTP as a substrate and Mn^{2+} as a divalent metal ion cofactor. In contrast to these other examples this motif is larger in size and more complex. These larger deoxyribozyme motifs are less likely to be present multiple times in the initial pool used to start the in vitro selection. A possible explanation is that the original sequences contained less efficient versions of these motifs and these sequences evolved towards the common secondary structure with acquired mutations during the PCR steps that took place during each round of selection.

One of the goals of the characterization of these self-phosphorylating deoxyribozymes was to use what structural information we obtained to aid in future in vitro selections. The structural characterization of Dk2, Dk3 and Dk4 revealed a complex secondary structure with three to four stems, and they potentially contain stems formed between distal regions of the deoxyribozyme. The fact that this structure is more complex than that seen with some other deoxyribozymes, such as RNA-cleaving deoxyribozymes, can be rationalized as the reaction being catalyzed is more complex, consequently requiring a more complex structure. As opposed to RNA cleavage, where

the deoxyribozyme can form base pairs with its substrate that is linked to the deoxyribozyme during in vitro selection, self-phosphorylating deoxyribozymes must bind an exogenous NTP molecule and position their 5′ end in the correct orientation for catalysis. This understandably will require more interactions and a more complex structure. This phenomenon has been observed with GTP aptamers, with higher affinity aptamers that need to encode more interactions with GTP having more complex structures than aptamers with lower affinity (Carothers et al. 2004).

A major question for in vitro selections for more complex reactions is whether they contain enough sequences folded into these complex structures to allow for the initial rounds of in vitro selection to isolate and amplify catalysts for these reactions. Currently, the libraries used to select functional nucleic acids typically contain relatively short random-sequence regions of 40 to 100 nucleotides. While these libraries have been very effective in isolating aptamers and deoxyribozymes with small motifs that exhibit some level of structural complexity, they may be less effective in isolating sequences that fold into complex structural arrangements. A study analyzing the folding of sequences in the Ellington lab aptamer database (Lee et al. 2004) showed that few aptamers isolated through in vitro selection contained complex secondary structures (Luo et al. 2010). This observation contrasts the structures elucidated for natural aptamers found in riboswitches, which tend to contain more complex structures with three and four-way helical junctions (Schwalbe et al. 2007) and even five-way junctions as seen in the lysine riboswitch (Serganov et al. 2008).

One method to increase complexity will be to increase the length of DNA libraries. In a study of the effect of random-sequence length on ribozyme selection, it was found that larger structural motifs were much more readily found in libraries with larger random regions (Sabeti et al. 1997). Computational analysis suggests that completely randomized RNA libraries mainly contain sequences that fold into simple structures such as low-branching and linear folds, while more complex structures such as higher-order helical junctions are underrepresented (Gevertz et al. 2005). As DNA has similar folding propensities as RNA, these conclusions can also be applied to random-sequence DNA pools.

In order to obtain libraries that are not biased towards simple structures, it may be necessary to incorporate some rationality into the library design at the start of an *in vitro* selection experiment, such as designing libraries with a mixture of pre-engineered structural folds. Some work has been done with these libraries by introducing stretches of purine-pyrimidine repeats predicted to increase secondary structure (Ruff et al. 2010) as well as computation approaches to increase the number of complex secondary structures in nucleic acid libraries (Luo et al. 2010). It may also be worthwhile to incorporate structural motifs known to support catalysis, such as the common secondary structure of Dk2, Dk3 and Dk4, into DNA libraries. With such modifications to the powerful *in vitro* selection technique, it should be possible to isolate larger and more efficient deoxyribozymes with very complex structures, allowing for the isolation of deoxyribozymes that catalyze complex chemical reactions and increase the repertoire of reactions beyond what has been currently demonstrated. As well as base pairs,

deoxyribozymes can form other interactions in their active structures that are not accounted for in many of the aforementioned computation studies. An attempt to increase these types of structures in DNA libraries was the basis for Chapter 4 of this thesis and will be discussed in a later section.

The second project, detailed in Chapter 3, investigated another self-phosphorylating deoxyribozyme, Dk5, which displayed some unusual properties. Initial characterization revealed it could use ATP or GTP as a substrate and Ca^{2+}, Cu^{2+}, Mg^{2+}, or Mn^{2+} as a divalent metal ion cofactor. This was in contrast to the other deoxyribozymes isolated from the same high stringency selection that were all specific for a particular substrate and divalent metal ion cofactor. The difference in divalent cofactor specificity versus the lack of specificity seen with Dk5 could be due to different interactions between the deoxyribozyme and its divalent metal ion cofactor. Dk1, Dk2, Dk3, and Dk4 all are specific for Mn^{2+}. Although there could be multiple factors involved, one possible explanation is that Dk1, Dk2, Dk3, and Dk4 are forming interaction with Mn^{2+} that are more covalent in nature. As Ca^{2+} and Mg^{2+} have lower covalent indices that Mn^{2+}, they would not be able form the same type of interactions. Dk5, on the other hand, could be forming bonds with its divalent metal ion cofactors that are more ionic in nature, allowing it to interact equally with Ca^{2+}, Mg^{2+}, and Mn^{2+} as they all have ionic indices of similar value and ionic interactions are based on electrostatic attraction rather than the geometry of the ion. This suggests that Dk5 could be utilizing more ionic interactions than Dk1 to Dk4. Also, the base ring nitrogens of nucleobases bind most strongly to the d orbitals of the transition metal ions (Izatt et al. 1971) and the presence of this interaction could

explain the specificity for Mn^{2+} over the alkaline earth metals Mg^{2+} and Ca^{2+} if these interactions are taking place in the active structures of Dk1, Dk2, Dk3 and Dk4.

We examined whether the different interactions with cofactors and substrates translated to a different structural arrangement for Dk5. Secondary structure analysis revealed that Dk5 contained very little secondary structure, with evidence for only a small three base pair stem in the center of the sequence and an eight base pair stem at the 3′ end of the deoxyribozyme incorporating the 3′ primer binding site. Thus, 50 nucleotides were found not to be involved in secondary interactions. This is in contrast to the structure adopted by Dk2, Dk3 and Dk4 in which most nucleotides are involved in secondary structural interactions. Following tertiary structural analysis to investigate other possible types of interactions in Dk5, it was discovered that Dk5 contained a two-tiered guanine quadruplex in its active structure. Adding more complexity to its structure, when the active structure is modeled with the quadruplex and helix motifs, it is apparent that a pseudoknot arrangement is present with one arm of the three base pair stem being located in one of the loops from the quadruplex. Quadruplex formation was also found to be dependent on this stem for correct folding. This structure differed from other known quadruplexes in terms of both its long loop size and its dependence on this entwined helical element for quadruplex formation. This motif was also found to be stable as Dk5 has a high reaction yield, suggesting that most Dk5 sequences are folded into the correct active structure for catalysis. This stability could explain why Dk5 was able to flourish in a highly stringent selection scheme that favoured fast catalysts despite having a reaction rate that was about 10 times lower than that of the other Dk catalysts

isolated. The discovery of this structure could have implications on both the study of naturally occurring quadruplexes and quadruplexes derived from in vitro selection.

Many studies have been devoted to discovering new quadruplexes in the chromosomes and RNA sequences of the human genome. Currently, computational searches look for sequences that could form quadruplexes with more than two-tiers of guanine quartets and small intervening loops of up to seven nucleotides (Huppert and Balasubramanian 2005; Todd et al. 2005; Todd and Neidle 2011). Our discovery of a functional complex quadruplex structure containing two-tiers of guanines and large loops of up to 32 nucleotides suggests that current search parameters may miss some genome encoded quadruplex sequences. As the structure of Dk5 also shows that helices can interact with and stabilize quadruplexes, it may also be possible to design algorithms that incorporate helical elements between different stretches of guanines into their search parameters and identify naturally occurring quadruplex-helix pseudoknot motifs. For artificially derived quadruplexes, the discovery that complex quadruplex arrangements can be used in functional nucleic acids indicates that it may be beneficial to design DNA libraries that are enriched with sequences that have the potential to form quadruplex scaffolds.

Summarizing the kinase deoxyribozyme characterization projects, it was found that five self-phosphorylating utilized three distinct catalytic structures. The discovery that DNA can use different structural arrangements for phosphorylation is a significant finding, as it shows that DNA can provide multiple solutions for the same chemical problem. This contrasts the common finding that different proteins catalyzing the same

reaction typically contain similar structural arrangements. For instance, the catalytic subunits of the majority of protein kinases contain very similar structures across different species (Adams 2001). Common structures are also seen between protein kinases that phosphorylate different substrate sequences and target different amino acid residues for phosphorylation. The discovery that DNA uses many different structural arrangements to form active structures for the same chemical reaction suggests that DNA may be more versatile than previously thought. Having more structural arrangements available may offset DNA's lack of chemical variability as compared to proteins, adding to the advantages of DNA (low cost, high stability, amenability to engineering) over proteins in the development of future molecular tools.

The third study of this dissertation investigated whether DNA libraries could be designed to contain a large fraction of sequences that could fold into two-tiered quadruplexes. This motif was investigated as, in addition to being in the active structure of Dk5, it has been found in several other DNA aptamer and deoxyribozyme sequences (Bock et al. 1992; Mazumder et al. 1996; Chinnapen and Sen 2004) with diverse ligands and substrates, suggesting that this motif may have the ability to support a wide array of functional nucleic acids. As mentioned above, many studies have been undertaken to access the frequency of different secondary structural arrangements present in nucleic acid libraries. Trying to assess the frequency of sequences folded into quadruplex structures is more difficult. Unlike well-established secondary structure prediction algorithms that predict the lowest energy secondary structures of sequences (Hofacker 2003; Zuker 2003), prediction of quadruplex structure formation based on sequence is

still in its infancy. Current quadruplex prediction tools are limited by a low number of quadruplex sequences characterized, especially those that form two-tiered quadruplexes (Stegle et al. 2009).

To investigate the possibility of designing DNA libraries enriched in two-tiered quadruplex motifs, we took an empirical approach using biophysical techniques to determine the percentage of sequences in certain libraries that adopt quadruplex conformations. Libraries were designed to contain four diguanine repeats interspaced by stretches of random nucleotides that have the potential to form the loops in sequences folded into quadruplexes. It was found that libraries with loops of three and seven nucleotides all contained large numbers of sequences that folded into stable quadruplex structures. In contrast, a library containing 25 random nucleotides did not have any features associated with quadruplex formation, suggesting that few sequences in this library are folded into quadruplex structures. Of the diguanine repeat libraries, one library containing three loops of five nucleotides was shown to fold into a mixture of quadruplex structures. This mixture of structures would be particularly useful in an DNA library as different quadruplex structures would utilize different type of loop arrangements of edgewise, diagonal and double-chain reversal loops. These different topologies place the loops in different positions relative to the quadruplex scaffold. As the nucleotides of these loop residues are responsible for forming interactions with the ligand or substrate, having diversity of loop topologies in a DNA library would give a greater chance of finding sequences that have the stability of a quadruplex scaffold with loop nucleotide positions that can form interactions with a target molecule. Another

important finding from this study was that addition of primer binding sites to tested libraries changed the folding of the libraries such that sequences no longer folded into quadruplex structures, but rather behaved like a completely random sequence library that folds mostly into helical type arrangements. This would mean that performing in vitro selection with quadruplex-based libraries with primer binding sites would nullify the advantage gained by having a structured library, as the quadruplex structures observed in the primer site free quadruplexes would not fold into quadruplex arrangements. To overcome this we designed a scheme in which oligonucleotides were ligated to 5′ and 3′ ends of the diguanine repeat library. This was achieved using partially randomized template sequences that bind to part of the library and one of the oligonucleotides. Using this system it would be possible to perform a selection step using a diguanine repeat library and ligate isolated sequences to the two oligonucleotides that would act as primer binding sites for PCR amplification. Using this scheme, it will be possible to use these diguanine repeat libraries for in vitro selection of DNA aptamers against any target of choice. Since this new selection scheme will be using a library dominated by sequences folded into quadruplex structures, unlike conventional random sequence libraries, the chances of isolating aptamers that use quadruplexes as a structural scaffold should be greatly increased.

The final chapter of the dissertation describes the conversion of self-phosphorylating deoxyribozymes into NTP biosensors. Several self-phosphorylating deoxyribozymes have been isolated, including those having high specificity for a specific NTP substrate as well as universal kinases that can use several different NTPs as substrates (Li and

Breaker 1999; Wang et al. 2002). As each self-phosphorylation event results in the dephosphorylation of an NTP, catalytic activity of these deoxyribozymes can be equated to NTP concentration. We designed a scheme to amplify self-phosphorylation activity and couple this amplification event to a detection system. Through coupling of self-phosphorylation activity to circular ligation and RCA, Dk2 was used to detect nanomolar concentrations of GTP. The RCA product of circularized Dk2 can be coupled to a number of different detection systems (Nilsson et al. 2002; Smolina et al. 2004; Cheglakov et al. 2007; Ali and Li 2009), allowing for many means of reporting the presence of GTP. As a demonstration, we appended the antisense sequence of a fluorescence-generating deoxyribozyme to Dk2 to allow for fluorescent detection of the RCA product produced after self-phosphorylation, circular ligation and RCA. Through this we were able to show that Dk2 could be used as a specific sensor that reported the presence of GTP, but not ATP, CTP or UTP. We also showed that this sensor system could be used to detect divalent ions, which are required cofactors for all known self-phosphorylating deoxyribozymes. The RCA product of Dk2, which is dependent on Mn^{2+} for activity, could be detected only in the presence of Mn^{2+}, but not when other divalent metal ions were used during the self-phosphorylation step. This system could easily be modified to detect other NTP and divalent ions on which other self-phosphorylating deoxyribozymes are dependent. As these deoxyribozymes have been shown to be readily isolated by *in vitro* selection, selections could be used to improve the selectivity of a self-phosphorylating deoxyribozyme for a substrate or cofactor.

Additionally, self-phosphorylating deoxyribozymes could be selected *de novo* enabling the detection of other substrates or cofactors for which no current deoxyribozyme exists.

To apply this more generally to deoxyribozyme-based sensors, we have shown that the substrate of a deoxyribozyme can be converted into a target for a sensor through amplification and coupling to a detection system. This type of system holds some advantages over many of the current deoxyribozyme-based sensors that detect the presence of cofactors or allosteric regulator for RNA-cleaving deoxyribozymes. These systems have common problems, including deoxyribozyme activity in the absence of target and cleavage of the RNA substrate by contaminant molecules in the sensing solution. Using deoxyribozyme substrates as the targets of sensors some of these problems may be avoided. Background reaction in the absence of target should not occur, as the target is the substrate of the reaction. Stringent *in vitro* selection can also be used to ensure the specificity of the deoxyribozyme for the target and not for other molecules that might be present in the solution being tested.

As in vitro selection has been shown to be a powerful technique to isolate deoxyribozymes for most reactions for which a selection scheme can be developed, the number of substrate and hence targets that could be detected by these kinds of systems could be very large. Complex in vitro schemes could also be used to directly isolate deoxyribozymes that process a substrate and yield a detectable readout, akin to those employed with structure-switching signalling aptamers (Nutiu and Li 2005) and bacteria-sensing RNA-cleaving deoxyribozymes (Ali et al. 2011). The future of deoxyribozyme-based sensors also depends on developing sensors that outcompete other systems either in

sensitivity, specificity or ease of use. Many of the current deoxyribozyme sensors have detection limits and specificities that rival competing sensor techniques. Coupled with the recent advances in signal amplification and surface-based sensors that offer low background and are amenable to miniaturization, deoxyribozyme-based sensors have the potential to emerge as viable options for many sensing applications.

6.1 REFERENCES

Adams, J.A. (2001). "Kinetic and Catalytic Mechanisms of Protein Kinases." Chem Rev 101: 2271-2290

Ali, M. M., S. D. Aguirre, H. Lazim and Y. Li (2011). "Fluorogenic DNAzyme probes as bacterial indicators." Angew Chem Int Ed Engl 50(16): 3751-3754.

Ali, M. M. and Y. Li (2009). "Colorimetric sensing by using allosteric-DNAzyme-coupled rolling circle amplification and a peptide nucleic acid-organic dye probe." Angew Chem Int Ed Engl 48(19): 3512-3515.

Bock, L. C., L. C. Griffin, J. A. Latham, E. H. Vermaas and J. J. Toole (1992). "Selection of single-stranded DNA molecules that bind and inhibit human thrombin." Nature 355(6360): 564-566.

Carothers, J. M., S. C. Oestreich, J. H. Davis and J. W. Szostak (2004). "Informational complexity and functional activity of RNA structures." J Am Chem Soc 126(16): 5130-5137.

Cheglakov, Z., Y. Weizmann, B. Basnar and I. Willner (2007). "Diagnosing viruses by the rolling circle amplified synthesis of DNAzymes." Org Biomol Chem 5(2): 223-225.

Chinnapen, D. J. and D. Sen (2004). "A deoxyribozyme that harnesses light to repair thymine dimers in DNA." Proc Natl Acad Sci U S A 101(1): 65-69.

Conaty, J., P. Hendry and T. Lockett (1999). "Selected classes of minimised hammerhead ribozyme have very high cleavage rates at low Mg^{2+} concentration." Nucleic Acids Res 27(11): 2400-2407.

Cruz, R. P., J. B. Withers and Y. Li (2004). "Dinucleotide junction cleavage versatility of 8-17 deoxyribozyme." Chem Biol 11(1): 57-67.

Faulhammer, D. and M. Famulok (1996). "The Ca2+ Ion as a Cofactor for a Novel RNA-Cleaving Deoxyribozyme." Angew Chem Int Ed Engl 35(2324): 2837-2841.

Gevertz, J., H. H. Gan and T. Schlick (2005). "In vitro RNA random pools are not structurally diverse: a computational analysis." RNA 11(6): 853-863.

Hofacker, I. L. (2003). "Vienna RNA secondary structure server." Nucleic Acids Res 31(13): 3429-3431.

Huizenga, D. E. and J. W. Szostak (1995). "A DNA aptamer that binds adenosine and ATP." Biochemistry 34(2): 656-665.

Huppert, J. L. and S. Balasubramanian (2005). "Prevalence of quadruplexes in the human genome." Nucleic Acids Res 33(9): 2908-2916.

Izatt, R. M., J. J. Christensen and J. H. Rytting (1971). "Sites and thermodynamic quantities associated with proton and metal ion interaction with ribonucleic acid, deoxyribonucleic acid, and their constituent bases, nucleosides, and nucleotides." Chem Rev 71(5): 439-481.

Lee, J. F., J. R. Hesselberth, L. A. Meyers and A. D. Ellington (2004). "Aptamer database." Nucleic Acids Res 32(Database issue): D95-100.

Li, J., W. Zheng, A. H. Kwon and Y. Lu (2000). "In vitro selection and characterization of a highly efficient Zn(II)-dependent RNA-cleaving deoxyribozyme." Nucleic Acids Res 28(2): 481-488.

Li, Y. and R. R. Breaker (1999). "Phosphorylating DNA with DNA." Proc Natl Acad Sci U S A 96(6): 2746-2751.

Luo, X., M. McKeague, S. Pitre, M. Dumontier, J. Green, A. Golshani, M. C. Derosa and F. Dehne (2010). "Computational approaches toward the design of pools for the in vitro selection of complex aptamers." RNA 16(11): 2252-2262.

Mazumder, A., N. Neamati, J. O. Ojwang, S. Sunder, R. F. Rando and Y. Pommier (1996). "Inhibition of the human immunodeficiency virus type 1 integrase by guanosine quartet structures." Biochemistry 35(43): 13762-13771.

Nilsson, M., M. Gullberg, F. Dahl, K. Szuhai and A. K. Raap (2002). "Real-time monitoring of rolling-circle amplification using a modified molecular beacon design." Nucleic Acids Res 30(14): e66.

Nutiu, R. and Y. Li (2005). "In vitro selection of structure-switching signaling aptamers." Angew Chem Int Ed Engl 44(7): 1061-1065.

Ruff, K. M., T. M. Snyder and D. R. Liu (2010). "Enhanced functional potential of nucleic acid aptamer libraries patterned to increase secondary structure." J Am Chem Soc 132(27): 9453-9464.

Sabeti, P. C., P. J. Unrau and D. P. Bartel (1997). "Accessing rare activities from random RNA sequences: the importance of the length of molecules in the starting pool." Chem Biol 4(10): 767-774.

Salehi-Ashtiani, K. and J. W. Szostak (2001). "In vitro evolution suggests multiple origins for the hammerhead ribozyme." Nature 414(6859): 82-84.

Santoro, S. W. and G. F. Joyce (1997). "A general purpose RNA-cleaving DNA enzyme." Proc Natl Acad Sci U S A 94(9): 4262-4266.

Schlosser, K. and Y. Li (2004). "Tracing sequence diversity change of RNA-cleaving deoxyribozymes under increasing selection pressure during in vitro selection." Biochemistry 43(30): 9695-9707.

Schwalbe, H., J. Buck, B. Furtig, J. Noeske and J. Wohnert (2007). "Structures of RNA switches: insight into molecular recognition and tertiary structure." Angew Chem Int Ed Engl 46(8): 1212-1219.

Serganov, A., L. Huang and D. J. Patel (2008). "Structural insights into amino acid binding and gene control by a lysine riboswitch." Nature 455(7217): 1263-1267.

Smolina, I. V., V. V. Demidov, C. R. Cantor and N. E. Broude (2004). "Real-time monitoring of branched rolling-circle DNA amplification with peptide nucleic acid beacon." Anal Biochem 335(2): 326-329.

Stegle, O., L. Payet, J. L. Mergny, D. J. MacKay and J. H. Leon (2009). "Predicting and understanding the stability of G-quadruplexes." Bioinformatics 25(12): i374-382.

Tang, J. and R. R. Breaker (2000). "Structural diversity of self-cleaving ribozymes." Proc Natl Acad Sci U S A 97(11): 5784-5789.

Todd, A. K., M. Johnston and S. Neidle (2005). "Highly prevalent putative quadruplex sequence motifs in human DNA." Nucleic Acids Res 33(9): 2901-2907.

Todd, A. K. and S. Neidle (2011). "Mapping the sequences of potential guanine quadruplex motifs." Nucleic Acids Res 39(12): 4917-4927.

Wang, W., L. P. Billen and Y. Li (2002). "Sequence diversity, metal specificity, and catalytic proficiency of metal-dependent phosphorylating DNA enzymes." Chem Biol 9(4): 507-517.

Zuker, M. (2003). "Mfold web server for nucleic acid folding and hybridization prediction." Nucleic Acids Res 31(13): 3406-3415.