A PURSUIT OF SELF-CLEAVING RNASE-A MIMICKING DNAZYMES WITHOUT AN IMIDAZOLE FUNCTIONALITY AND THEIR FURTHER CHARACTERIZATION

by

Somdeb Paul

B.Tech., Vellore Institute of Technology, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Genome Science and Technology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

February 2022

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, a thesis entitled:

A pursuit of self-cleaving RNAse-A mimicking DNAzymes without an imidazole functionality and their further characterization

submitted by Somdeb Paul in partial fulfillment of the requirements for

the degree of Master of Science

in Genome Science and Technology

Examining Committee:

David M. Perrin, Professor, Department of Chemistry, UBC
Supervisor

Philip Hieter, Professor, Department of Medical Genetics, UBC
Supervisory Committee Member

Stephen G. Withers, Professor, Department of Chemistry, UBC
Supervisory Committee Member

Abstract

DNAzymes are catalytic DNA molecules that resemble ribozymes and, to a lesser extent, protein-based enzymes. Many of the known DNAzymes have been selected in the presence of metal ions (M^{2+}) that are present at non-physiological levels. As a result, there has been much less interest in developing a candidate DNAzyme for therapeutic use. M^{2+}-independent DNAzymes are intriguing as they can function even under the low magnesium (Mg^{2+}) concentrations usually found in an intracellular environment. Previous works in our lab relied on the use of DNAzymes containing three or two modified nucleosides such as an amine (lysine side chain), a guanidine (arginine side chain) and an imidazole (histidine side chain) with an emphasis on the imidazole group to select metal-independent DNAzymes resembling the RNase A mechanism. Using the functional group modifications they have been able to select several fast cleaving DNAzymes. However, to date there has not been a systematic study on the necessity of all three modifications (in particular the imidazole) as the value of each nucleoside have been contextualized at the beginning of the selection. I hypothesized that two of the positively charged dNTP analogs, excluding the imidazole, i.e. 5-aminoallyl-2'-deoxycytidine triphosphate **(dC^{aa}TP)** and 5-guanidinoallyl-2'-deoxyuridine triphosphate **(dU^{ga}TP)**, are sufficient to select *in-vitro* self-cleaving DNAzymes in the absence of Mg^{2+} ions in a pH-independent fashion. I was able to identify two DNAzyme candidates that performed the best (Dz11-23 and Dz11-33). The DNAzymes appeared to show a monophasic rate constant (~0.03 min^{-1}) on-par with previous selections without the imidazole; when the DNAzyme species were inquired using a biphasic kinetics model two phases were identified with one of the phases showing a rate constant that was similar to the fast phase in the imidazole containing DNAzymes (k_{cat} = 0.7-0.9 min^{-1}). When studied over different pH values and Mg^{2+} concentration the DNAzymes also appeared to show a self-cleavage activity that was

independent of pH and Mg^{2+} with an optimal temperature around 25 °C and sufficiently high cleavage observed even at 37 °C. The results in the thesis support my hypothesis that the imidazole moiety may be redundant in selecting fast cleaving metal-independent DNAzymes.

Lay Summary

In the last two years, the impact of COVID-19 has been severe with a global tally of 276 million confirmed cases and 5.37 million deaths according to the World Health Organization (WHO). The RNA-based vaccines developed by Pfizer-BioNTech and Moderna have been effective in preparing the immune system against the virus. Modifications on the RNA has helped it evade nucleic acid degradation by RNA-degrading enzymes (ribonucleases). RNA-cleaving DNAzymes are an effective next-generation therapeutic that can be used to cleave target sites on a viral RNA and can be extended towards other viral RNAs. Our lab has previously selected a DNAzyme candidate that can cleave with multiple turnovers at a fast rate constant under similar conditions that exist in a human body (0.5 mM Mg^{2+}, pH 7.45). I show here that we can select a DNAzyme with a comparable rate constant without one of the functional groups that was initially perceived to be crucial.

Preface

The work was initially developed by Dr. David Perrin during his Post-doctoral work at the Museum National d'Histoire Naturelle, Paris which was subsequently furthered by Dr. Marcel Hollenstein and Dr. Yajun Wang and has now been extended in our research. The overall project was overseen by Dr. David Perrin. One of the modified nucleotides in our work, 5-guanidinoallyl-2'-deoxyuridine triphosphate **(dU^{ga}TP)**, was synthesized by Antonio A.W.L. Wong using a procedure originally developed by Dr. Curtis Lam. Antonio A.W.L Wong and Leo. T. Liu helped optimize the synthesis of the modified nucleotide to maximize its yield.

The initial oligonucleotide sequence for the construct was adopted from Dr. Marcel Hollenstein (2008). The selection cycle, the PCR amplification and the primer extension were optimized by me from Hollenstein *et al*. (2008). All stages of the selection procedure were completely performed by myself. All parts of the thesis were solely written by me. A part of this thesis has also been accepted for publication in ChemBioChem as "Selection of M^{2+}-independent RNA-cleaving DNAzymes with side chains mimicking arginine and lysine". All characterization experiments for the DNAzyme candidates and their subsequent analysis – investigating enzymatic parameters, the elucidation of catalytic properties, initial structure prediction, and substrate specificity – were all performed by myself. Only the part of the experiments from the manuscript that have been done by me and those experiments from my research that have been successful have been elucidated in the thesis.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

A	adenine
ACTH	adrenocorticotropic hormone
ARMD	Age-Related Macular Degenerative (disorder)
ATP	adenosine triphosphate
BTT	5-benzylthio-1*H*-tetrazole
C	cytosine
CDMB	Central Dogma of Molecular Biology
cDNA	complementary DNA
dA^{his}	8-histaminyl-deoxyadenosine
$dA^{his}TP$	8-histaminyl-deoxyadenosine triphosphate
dATP	2’-deoxyadenosine triphosphate
DCA	dichloroacetic acid
dC^{aa}	5-aminoallyl-2'-deoxycytidine
$dC^{aa}TP$	5-aminoallyl-2'-deoxycytidine triphosphate
DCM	dichloromethane
dCTP	2’-deoxycytidine triphosphate
ddH_2O	double distilled H_2O
DEPC	diethylpyrocarbonate
dGTP	2’-deoxyguanosine triphosphate
DMT	dimethoxytrityl
DNA	deoxyribonucleic acid
dNTP	2’-deoxyribonucleoside triphosphate

D-PAGE	denaturing polyacrylamide gel electrophoresis
dsDNA	double stranded DNA
dTTP	2’-deoxythymidine triphosphate
dU^{aa}	5-aminoallyl-2'-deoxyuridine
dU^{ga}	5-guanidinoallyl-2'-deoxyuridine
$dU^{aa}TP$	5-aminoallyl-2'-deoxyuridine triphosphate
$dU^{ga}TP$	5-guanidinoallyl-2'-deoxyuridine triphosphate
dUTP	2’-deoxyuridine triphosphate
dU^{y}	phenol-deoxyuridine
EDTA	Ethylene Diamine Tetraacetic Acid
ETT	5-ethylthio-1*H*-tetrazole
FNA	functional nucleic acid
G	guanine
HDV	hepatitis delta virus
HPLC	High Performance Liquid Chromatography
M^{2+}	divalent metal cations (Mg^{2+}, Hg^{2+}, Ca^{2+} etc.)
mRNA	messenger RNA
N_{40}	40 random nucleosides in a DNA strand
NGS	Next Generation Sequencing
NTA	nitrilotriacetic acid
PAGE	polyacrylamide gel electrophoresis
PCR	Polymerase Chain Reaction
rA	riboadenosine

rC	ribocytidine
rG	riboguanidine
RNA	ribonucleic acid
rU	ribouridine
SELEX	Systematic Evolution of Ligands through EXponential enrichment
SOMAmer	Slow Off-rate Modified Aptamer
SPOC	Solid Phase Organic Coupling
ssDNA	single stranded DNA
T	thymine
TBDMS	tert-butyldimethylsilyl
TDP-43	TAR-DNA binding protein – 43
TEN	Tris/EDTA/sodium chloride
THF	tetrahydrofuran
tRNA	transfer RNA
U	uracil

Acknowledgements

I would like to thank all the faculty and staff at the University of British Columbia, Vancouver for providing me a platform for facilitating an intellectual discourse and for inculcating in me the necessary skills that inspired my work. I would like to thank my graduate supervisor Dr. David Perrin for the endless support that he has provided me during this arduous journey. His guidance and mentorship has been crucial in developing my skills and persistence as a researcher. I would also like to extend my gratitude to my thesis advisory committee, Dr. Stephen G. Withers and Dr. Phil Hieter, for providing my project a sense of direction.

I would like to thank all past and present Perrin group members for helping me grasp the fundamentals of organic chemistry and helping me understand the results of some of my synthesis. I would especially like to thank Jerome Lozada and Antonio Wong for their help in this regard, especially during my initial days as a graduate student. I would also like to thank my fellow graduate students Leo Liu and Alla Pryyma for their insightful advice during frustrating times: be it while trying to figure out the synthetic yield of the modified nucleotide or when my selections did not perform as was expected.

The thesis has been proofread by Antonio Wong, Alla Pryyma, and Dr. David Perrin. Thank you all for suggesting corrections and making my thesis coherent and understandable.

I would also like to thank Teresa Law, the undergraduate that I worked with when I started graduate school, for helping me carry out some crucial experiments even though it did not materialize into my thesis.

My gratitude also goes to all the non-teaching staff in the shops and services of the Chemistry department. I am particularly thankful to George Kamel who is in-charge of inventory management and shipping at the chemistry stores for his prompt response whenever I required to procure materials especially radioactive isotopes. I also appreciate Dr. Elena Polishchuk and Jessie Chen at the BioServices Laboratory for their advice on molecular biology techniques.

I am ultimately thankful to all my friends, co-workers and networks for their support, encouragement and entertainment that made my graduate school life more bearable.

Finally, I would like to thank my parents and my brother for always being there for me throughout my years of education, for all the mental and emotional support, through their encouragement.

To my parents and brother

Chapter 1: Introduction

1.1 Nucleic acids and *in vitro* Selection

Nucleic acids are a cost-effective and an efficient therapeutic solution which has been gaining interest among the research community since the discovery of aptamers in the early 1990s, ribozymes in the late 1980s and antisense oligonucleotides in the early 1980s. The mRNA based vaccines recently approved for clinical use have assured the safety and efficacy of functional nucleic acids (viz., therapeutic mRNA) in a diverse population. This thesis focuses on the development of one type of functional nucleic acid (FNA) that is explored in the context of *in-vitro* selection of a self-cleaving DNAzyme (introduced later in the text) that can cleave a single ribonucleotide using two modified nucleotides (explained further in the chapter). Chapter 1 introduces the relevance of nucleic acids including their structures; it further illustrates the different modes of nucleic acid synthesis, and how they are used in an *in-vitro* selection experiment; and finally, the chapter also highlights the discovery of certain important modified nucleotides that have been used by various research groups.

1.1.1 Relevance of nucleic acids

There exist two major classifications of nucleic acids – deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Nucleic acids as one of the fundamental components in living cells are responsible for the expression of proteins, and the regulation of DNA and RNA. DNA which is the primary carrier of genetic information is localized within the nucleus as a condensed structure known as a chromosome, whereas a free form RNA is mostly localized in the cytosol. The primary structure of DNA comprises two polynucleotide chains intertwined in the form of a double helix, while RNA which is considered to be a linear structure may adopt diverse secondary structures

that are responsible for biological properties including ligand binding and even catalysis (*vide infra*).

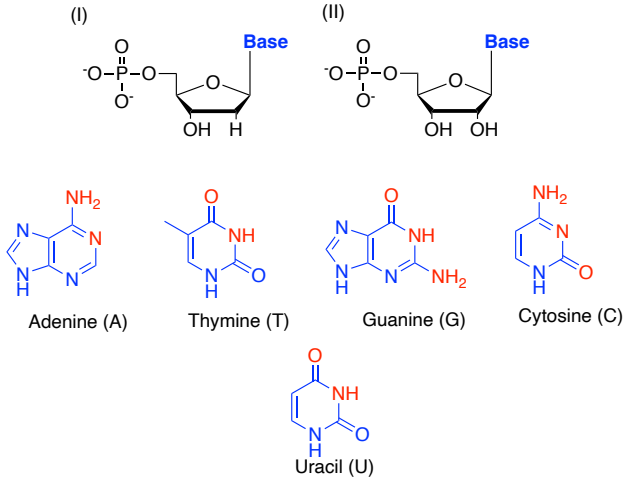


Figure 1.1 Nucleotide monomers of DNA and RNA polymers.
The nitrogenous bases A, T, G and C (figure above) represent the base group on the deoxynucleoside monophosphate (I) whereas the bases A, U, G and C are part of the ribonucleoside monophosphate represented by figure (II). The atoms/groups marked in Red take part in Watson and Crick hydrogen bonding interactions with their complementary counterpart, and the opposite side of the nitrogenous base is the Hoogsteen face. Created with ChemDraw Prime 19.1

The DNA strand is made of nucleotide building blocks (also known as oligonucleotide) that contain nitrogenous bases – adenine, thymine, guanine and cytosine – and sugar which is deoxyribose (Fig 1.1). Canonical base-pairing interactions between the two strands are conventionally between adenine and thymine (through two hydrogen bonds) and guanine and cytosine (through three hydrogen bonds). While the nitrogenous bases are involved in hydrogen bonding, the DNA backbone comprises successive sugar and phosphate groups linked together by 3'-5'-phosphodiester bonds.[1] In RNA, on the other hand, the nucleotides contain the nitrogenous bases – adenine, uracil, guanine and cytosine, whereas the sugar is ribose (Fig 1.1). Unlike DNA which is a double-stranded polymer, RNA is primarily single-stranded and can self-associate (also

referred to as hybridization)[2]. This gives RNA the flexibility to form various secondary structures (will be described further).

The helical structure of a double-stranded DNA had been contemplated for a long time until the x-ray diffraction results by Wilkins and Franklin made it possible for Crick and Watson to develop a stereochemically accurate model of the DNA[3,4]. The photographs also suggested that the structure of DNA was made of multiple polynucleotide chains (at least two). These strands are held together by weak hydrogen bonds, and the hybridization pattern demonstrated by the nitrogenous bases can be predicted by Chargaff's rules[5], which have also been enumerated in Watson and Crick's publication on the DNA double helix. In a double-stranded DNA, for every adenine residue that is present on one strand, there is also a thymine residue on the second; similarly, guanine and cytosine are also equally distributed on either strands[2]. This signifies a complementary relationship between the two strands of double-stranded DNA. Therefore, if one of the polynucleotide strands is identified in the 5' → 3' direction the complementary strand is aligned in an antiparallel orientation (3' ← 5').

The cognate hydrogen bonding interaction between the bases, that keeps the two antiparallel strands together, are not the only stabilizing interactions. The relatively water-insoluble nitrogenous bases that stack above each other also contribute toward stabilizing the double helix. The pi-conjugated system (refer to Fig 1.1 for structure) in the stacked bases interact with each other through Van der Waals interaction because of transient, induced dipoles that exist between the electron clouds[2]. These hydrophobic bases sequester themselves inside a co-axially stacked

and base-paired double-helix minimizing exposure to solvent (water) which can destabilize the structure.

While DNA is a carrier of genetic information, proteins cannot be directly synthesized from DNA. Protein synthesis occurs in a separate compartment (cytoplasm) from the DNA (localized in the nucleus) which is separated by a nuclear membrane. Therefore, there must be another intermediate nucleic acid that carries the information from the gene (viz., DNA) to the site of protein synthesis. This carrier nucleic acid is RNA. One of the strands of the chromosomal DNA (localized in the nucleus) acts as a template for RNA synthesis that is used for protein production.

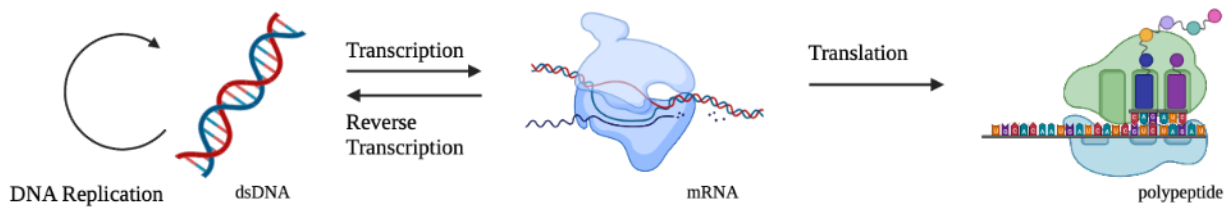


Figure 1.2 Central Dogma of Molecular Biology (CDMB) flowchart.
CDMB describes the flow of genetic information in a cell from the nucleus to the synthesis of a polypeptide in the cytosol. Created with BioRender.com

This flow of information in a cell is governed by the Central Dogma of Molecular Biology (CDMB). A simplification of the CDMB (Fig 1.2) describes that DNA is transcribed into messenger RNA (mRNA) in the nucleus which is then translated into polypeptides within the cytoplasm. The DNA itself undergoes replication inside the nucleus during cell division to maintain the number of chromosomes between daughter cells. In rare instances like retroviruses (a class of viruses), the RNA can also be copied back into complementary DNA (cDNA) by reverse

transcription, however, RNA cannot be synthesized using the polypeptide as a template. The polypeptide chains further undergo post-translation modifications such as addition of sugars (glycosylation), ubiquitin (ubiquitylation), phosphate group (phosphorylation), nitrosyl group (nitrosylation), methyl group (methylation), acetyl group (acetylation), and lipid groups (lipidation) to form mature protein products. Some common uses of such proteins are as cell signal receptors, skeletal proteins (like microtubules) and other extracellular entities (such as hormones, enzymes and antibodies).

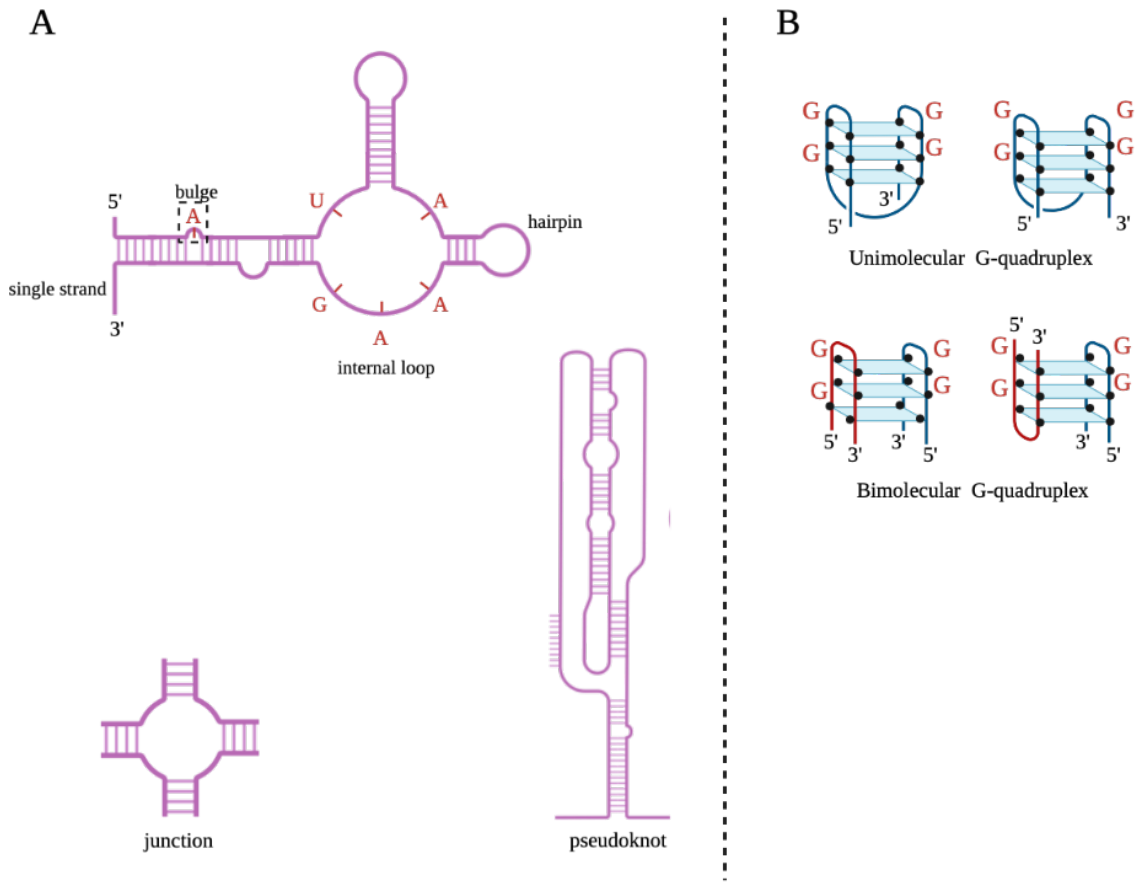


Figure 1.3 Common Secondary Structures found in RNA and DNA.
(A) illustrates bulge, hairpin, internal loop, junction and pseudoknot structures commonly found in an RNA whereas (B) shows various kinds of G quartets – unimolecular quartets found in artificially designed systems and bimolecular quartets in natural telomeric regions. Black dots in the G-quartet (or quadruplex) denote the guanine residues in the quartet. Created with BioRender.com

Apart from the conventional use of nucleic acids in the flow of genetic information, nucleic acids can also have several regulatory functions that are exerted by the strands as they fold into secondary structures. DNA is largely a rigid double-helical structure with less freedom of movement and as such there are limited structural forms that can exist. An important DNA regulatory element is a **G-quadruplex** (a set of 4 guanine residues lying in the same plane) that is usually found in the telomeric regions of chromosomes. The quartets (another term for G quadruplex) are structures that contain two tracts of guanine (G) residues separated by a region of thymine bases that is part of a **hairpin loop**[6] (Fig 1.3). Such quadruplexes are usually further stabilized by monovalent ions (such as Na^+ and K^+)[7,8]. Single-stranded RNA on the other hand is flexible and irregular; it can form a much varied set of secondary structures as a result (Fig 1.3)[2].

When complementary regions of the RNA are adjacent or brought into proximity, they can hybridize with each other with the intervening regions looped outward. This form of structure is considered a **stem loop**. Regions within the stem-loop structure can also have **bulges** (a single base bulging out of the stem with the bases on either side having complementary partners), **internal loops** (when there is a short stretch of unpaired sequence on either side of the stem), and **junctions** (when there are forks in the structure). Distant regions of an RNA can interact with each other in the form of knotted constructs (such as **pseudoknots**). These pseudoknots were first identified in a tobacco mosaic virus (TMV), which is minimally two helical components of the RNA linked through a single-stranded region (Fig 1.2). Pseudoknots are also present in the catalytic regions of various types of ribozymes, self-splicing introns, and telomeric DNA[9–12].

The hepatitis delta virus (HDV) ribozyme (a catalytic RNA), one of the fastest known ribozymes, with a catalytic rate greater than 1 s^{-1} folds into a double pseudoknot conformation[10]. RNA can also adopt a wealth of tertiary structures as it has enormous rotational freedom in the phosphodiester backbone of the non-base-paired (single-stranded) region. Complex tertiary structures that the RNA can fold into may involve unconventional base pairings, such as base triplet and base-backbone interactions (seen in tRNA). Such secondary and tertiary structures in the RNA provide active sites for catalysis as well as ligand binding to various proteins, as described subsequently. With the advent of chemical synthesis for the preparation of single-stranded DNA strands, it was appreciated that many of these RNA structural motifs can be replicated in DNA as well.

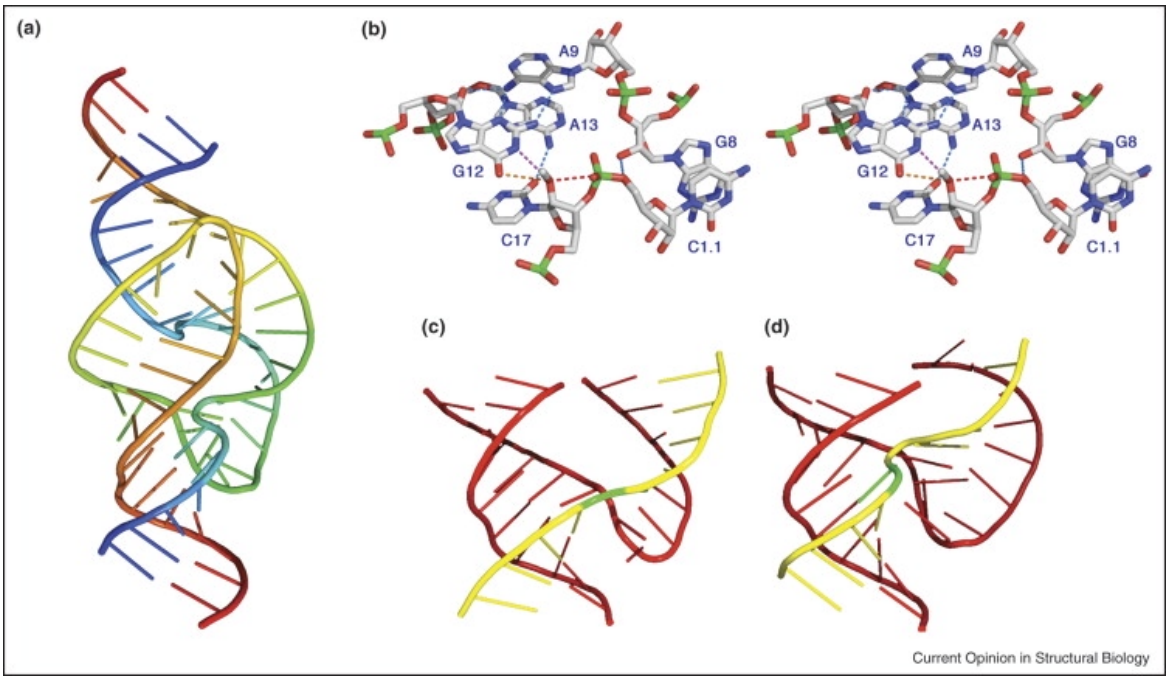


Figure 1.4 Structures of the hammerhead ribozyme.
(A) The global structure of the full-length ribozyme (B) active site of the ribozyme shown in sticks representation (C) comparison between the minimal ribozyme and the full-length ribozyme. Picture from Scott, W. (2007).

Interest in RNA exploded when ribozymes were discovered in the mid-1980s; this discovery provided a logical explanation to the origin of the present-day information-carrying system from a primordial soup. The ribosome is one of the most important ribozymes that was identified; 23S and the 70S ribosomes have now provided us with many structural insights[13]. The structure of the ribosomal RNA is stabilized by proteins that can recognize specific motifs. One of the unique RNA secondary structure motifs found in these ribosomes is the **kink-turn** (K-turn)[14]. The K-turns are kinks in the phosphodiester backbones that cause a sharp turn in the helix (binds to proteins that help maintain the structure). One of the most well-known ribozymes that are used as a benchmark for synthetic deoxyribozymes (DNAzymes) and ribozymes (RNAzymes) is the hammerhead ribozyme.

For a long time, only a partial sequence of the enzyme was studied (minimal ribozyme) as in Fig 1.4 (C). However, a full-length hammerhead ribozyme contains the nucleophilic site which is aligned perfectly with the scissile phosphodiester bond of its target; two invariant (essential to the activity) nucleotide residues G12 (guanine-12) and G8 (guanine-8) are positioned in the active site which participates in an acid-base catalysis (Fig 1.4)[13]. No divalent metals like Mg^{2+} and Mn^{2+} ions have been found to contribute to any bridges with the scissile phosphate and acid-base catalysis was found independent of metals. Therefore, this naturally occurring ribozyme has consistently been used as a gold standard for comparing the activities of many *in vitro* selected metal-independent RNA cleaving DNAzymes (*vide infra*). Other small self-cleaving ribozymes have also been discovered that can function in the absence of divalent metal ions. Such forms of catalytic ribozymes are the Group I family of introns, Diels-Alderase ribozyme, and a hairpin ribozyme.

Apart from RNA as catalysts, there are also RNA that present an affinity towards ligands such as proteins – Qß and other RNA binding proteins (such as structural proteins, translation elongation factors, poly-A binding protein, and transcription/translation initiation factors)[15]. Qß and MS2 are coat proteins of the RNA bacteriophages that recognize specific RNA motifs[16]. They bind distinct RNA stem-loops which provide a useful model of specific RNA-protein interaction.

Such secondary structures (as described above) in nucleic acids have formed the basis for researchers to develop synthetic nucleic acid catalysts (RNA/DNAzymes) and affinity ligands (aptamers) using single-stranded DNA, and RNA. In fact, synthetic aptamers isolated by *in vitro* selection (explained in Section 1.1.3) are based on naturally occurring riboswitches that regulate RNA expression through binding soluble metabolites. Such synthetic oligonucleotides can adopt several of the secondary structures in native nucleic acid structures that have been discussed above.

1.1.2 Common Synthetic Routes of DNA Oligonucleotides

Nucleic acids that are used in biological experiments, including the research contemplated in this thesis, can be synthesized using either enzymatic synthesis or by chemical synthesis on a solid phase using phosphoramidite monomers. The synthesis of oligonucleotides varies based on the application as either method have merits and demerits. Enzymatic synthesis due to its exponential nature usually leads to a large excess of the pure product (with minor truncates) as compared to solid-phase methods. However, enzymatic synthesis may be challenging when using modified nucleotides with side chains that present added xenobiotic and/or protein-inspired functional groups (based on the efficiency of the DNA polymerase used to incorporate them). In such cases, solid-phase organic chemistry (SPOC) can yield much better results. SPOC is always preferred

when synthesizing oligonucleotides in bulk as individual components of enzymatic synthesis can be expensive on a large scale (enzymes, and modified nucleotides). Therefore, during *in vitro* selection (explained in Section 1.1.3) experiments, the oligonucleotides are mostly synthesized by enzymatic methods but afterwards during large scale synthesis for characterization and distribution, researchers prefer making the oligonucleotides using SPOC.

Enzymatic Synthesis of DNA Oligonucleotides

The enzymatic synthesis of DNA is done using polymerase chain reaction (PCR). A standard PCR allows a single short region of a gene or a DNA strand to be copied multiple times by a DNA polymerase (Fig 1.3). A single copy of DNA can provide thousands of copies of the product using simple molecular biology reagents and cycling conditions (temperature and time). PCR comprises three major components – a target oligonucleotide to be amplified (template strand), forward and reverse primers (primer pair), and the polymerase (enzyme). In addition to the template DNA, the polymerase also requires free nucleotides, dNTPs, which implies a mixture of dATP, dTTP, dGTP and dCTP, usually in an equimolar ratio. A PCR is done using a ThermoCycler™ where the DNA is synthesized iteratively over cycling temperatures and each PCR cycle is divided into three stages: the double-stranded DNA (dsDNA) is first **denatured** at a high temperature, the primers **anneal** as the temperature cools down to an optimum temperature, and the enzyme then starts **extending** the primer by incorporating nucleotides (Fig 1.5). This repeats until the reaction ends owing to the depletion of either primers (usually) or dNTPs (sometimes) (viz., at least 25-30 cycles of amplification). One of the primers is identical to the 5' → 3' sense strand and the other identical to its antiparallel strand (in the 3' ← 5' direction). PCR typically results in two identical copies of double-stranded (dsDNA) after the first cycle, each acting as a template for the following round.

Apart from temperature and cycle optimization the addition of denaturing agents (such as betaines and dimethylsulfoxide) usually help to improve yields by minimizing artefact formation (such as primer-dimers and other PCR artefacts due to concentration of the polymerase and long incubation time).

PCR effectively amplifies trace amounts (e.g. fmol) of a given target sequence to give nanomoles of the same. Modified nucleotides, however, are unusually incorporated into the growing strand using a single-sided extension of the primer instead of doing so by PCR. This uses fewer resources in terms of enzyme and modified nucleotide consumption. In brief, a single primer (instead of a primer pair) is annealed to the template strand and extended isothermally over a set time. A more detailed explanation of the procedure is covered in Chapter 2.

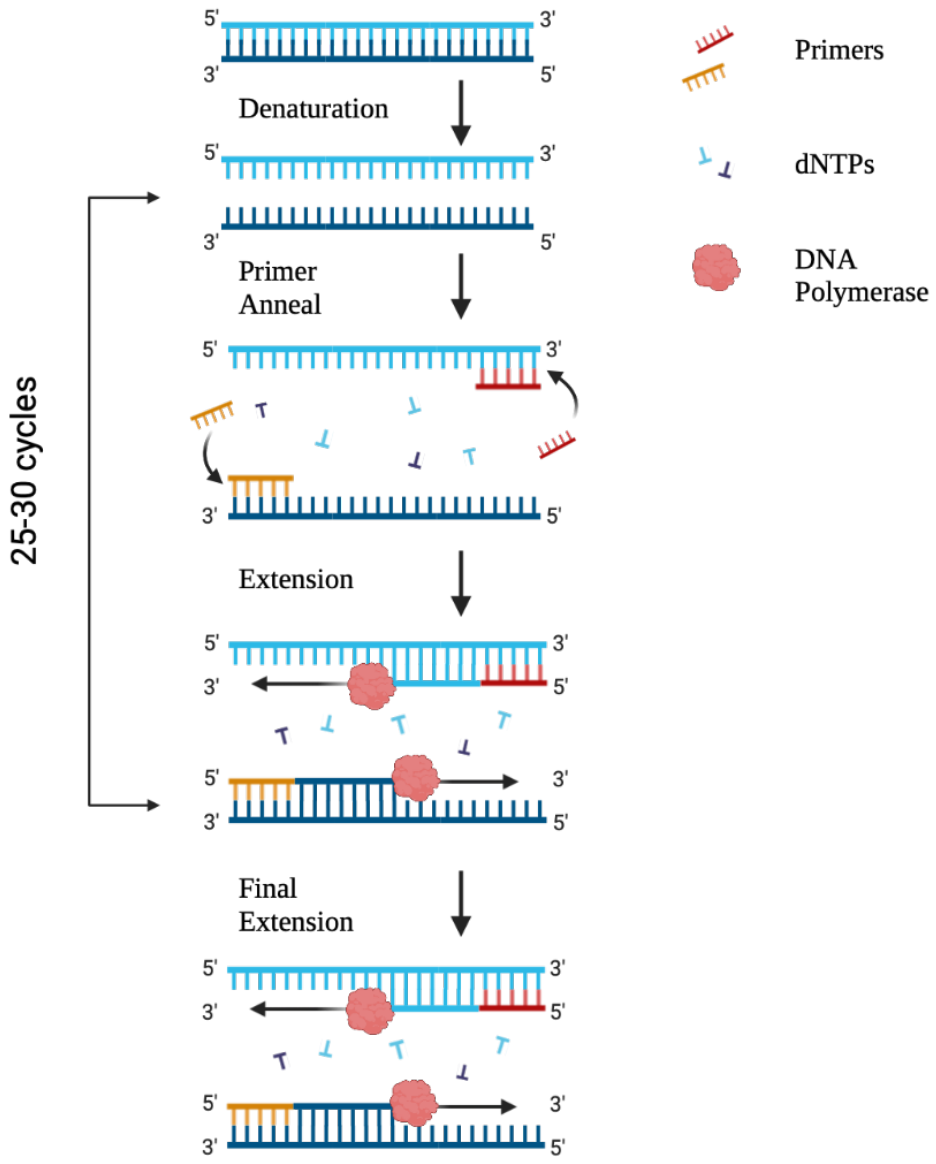


Figure 1.5 Enzymatic Synthesis of DNA oligonucleotides by PCR
This is also known as semiconservative replication. Image created with BioRender.com

Chemical Synthesis of Oligonucleotides

A second way of synthesizing oligonucleotides utilizes solid-phase synthesis. Solid Phase Organic Chemistry (SPOC) has recently become popular in the realm of *in vitro* selection experiments to prepare degenerate libraries of oligonucleotides with combinatorial methods[17–23]. Modern nucleic

acid synthesizers employ stable phosphoramidite monomers to build the growing polynucleotide. Such robust reactions assist both chemists and molecular biologists in the generation of specific ribo- and deoxyribo- oligonucleotides with a variety of labels, modified linkages, and unnatural bases that are attached throughout the chain. A large number of degenerate sites can be incorporated into the chain extension while synthesizing oligonucleotides for *in vitro* selection experiments by pre-mixing nucleotide bases in specific ratios. The protected nucleosides that are used in the synthesis of these oligonucleotides are called phosphoramidites.

In traditional protection schemes, the nucleophilic amino groups on the nitrogenous base were usually protected with N^2-isobutyrylguanine, N^6-benzoyladenine or N^4-acetylcytosine groups. More recently, however, protection of the exocyclic nitrogens using a phenoxyacetyl group on adenosine, dimethylformadine on guanosine, and acyl group on cytosine have been widespread as they were more labile towards aminolysis and can be deprotected rapidly under relatively mild conditions at the end of the oligonucleotide synthesis[18]. The oligonucleotides can be very rapidly synthesized on the solid support; after each extension step, the excess reagents are washed off the column keeping the product attached to a support, which is comparatively of a much higher molecular weight. The synthesis can be staged into four steps – Detritylation, Coupling, Oxidation, and Capping[18]. These stages in the synthesis of the polymer are illustrated in Fig. 1.6.

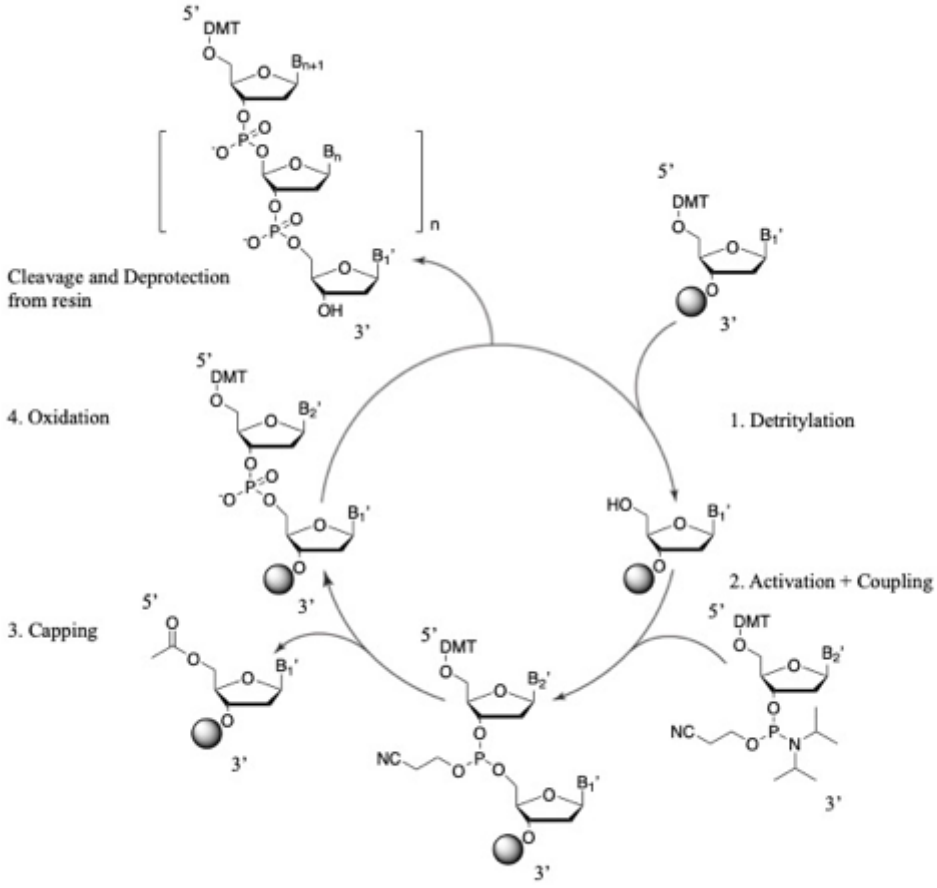


Figure 1.6 Solid Phase Organic Coupling (SPOC) for the synthesis of DNA and RNA.
Schematic has been adapted from Ellington & Pollard, (1998) and created with ChemDraw Prime 19.1.

Unlike enzymatic synthesis, the synthesis on a solid support is done in a 3' → 5' direction. The first nucleoside is attached to the bead at the 3' hydroxyl group while the 5' primary alcohol of the amidite is protected by dimethoxytrityl (DMT). The first step in the synthesis, **Detritylation**, involves the removal of the acid-labile DMT ether from 5'-OH. This is achieved using dichloroacetic acid (DCA) in dichloromethane (DCM). In the subsequent steps, the only available reactive nucleophile is the deprotected 5'-OH. The step must be performed quickly to avoid acid-catalyzed depurination of the nitrogenous base. The reagents are washed away before the next step

to prevent premature detritylation of the incoming nucleoside. The 5'-OH nucleophile can then be **Coupled** after **Activation** of the incoming phosphoramidite. The incoming phosphoramidite is delivered to the reaction system with a molar excess of the weakly acidic (pK_a = 4.8) activator tetrazole. The molar excess of tetrazole ensures a complete activation of the phosphoramidite. The 5'-OH (nucleophile) of the preceding nucleotide attacks the newly added phosphoramidite at the diisopropylamine-protected phosphorus that is activated with a tetrazole. The incoming phosphoramidite is added in molar excess to achieve efficient coupling. Despite the activated phosphoramidite being highly reactive, there is still a small percentage of oligonucleotides that may not be coupled. To avoid these from becoming deletion sequences (truncated library of sequences) such unextended support-bound oligonucleotides are **Capped** with an acetyl group that remains inert throughout the entire synthesis. Acetic anhydride and N-methylimidazole in pyridine and tetrahydrofuran (THF) usually act as the acylating agent. At the last stage of each coupling cycle, the unstable phosphite triester linkages are **Oxidized** to a more stable phosphotriester using 0.02 M Iodine dissolved in water/pyridine/THF. Iodine forms an adduct with the pyridine at the phosphite triester which is subsequently displaced by water. Pyridine (alkaline in nature) also neutralizes the hydrogen iodide byproduct. Since the oxidization step includes water the reaction system must be rinsed with acetonitrile before the following step. The 5' DMT from the newly added monomer is removed before and a new cycle of coupling is started. At the end of the entire synthesis, the last trityl group is removed with a final acid wash and the oligonucleotide is displaced from the support beads by ammonium hydroxide wash (ammonolysis). This treatment also removes protecting groups from the heterocyclic nitrogenous bases.

The synthesized DNA strands are purified to resolve strands of length N from those of length N-1. Most modern oligonucleotide coupling reactions are ~ 99.5% efficient, however, depending on the requirement strands are purified from their truncated byproducts[19]. Polyacrylamide Gel Electrophoresis (PAGE) is by far the easiest option available to separate oligonucleotides of length N from N-1. However, this becomes increasingly much harder for oligonucleotides greater than 30 bases. Increasing the coupling efficiency using a higher concentration of phosphoramidites or longer coupling times may be good options for longer oligonucleotides. However, a high degree of stringency in terms of purity is non-essential for the downstream requirement of oligonucleotides used in most biological applications. Single-residue resolution for oligonucleotides can also be resolved with high purity using High-Performance Liquid Chromatography (HPLC). An advantage of HPLC is speed – an oligonucleotide greater than 40 bases using an optimized solvent system can be purified in under 30 minutes, as supposed to be 120 minutes by PAGE.

In recent times, after synthesis of the oligonucleotide strands, they can be subsequently capped with various functionalities such as 5'-triphosphates (Zlatev et. al., 2012) and 5'-O-triphosphates (Sarac et. al., 2016) for different applications[22,23]. The efficiency of oligonucleotide synthesis has been further improved using different protecting groups and coupling activators. Protecting groups of adenine and cytosine are largely the same as described previously, however guanine bases may also be used as N^2-isopropylphenoxyacetyl protected guanine phosphoramidites. Coupling activators like 5-ethylthio-1*H*-tetrazole (ETT) or 5-benzylthio-1*H*-tetrazole (BTT) are alternatives that improve coupling time and efficiency. In the case of RNA, the 2'-OH is protected using tert-butyldimethylsilyl chloride (TBDMS protection).

1.1.3 Introduction to *in vitro* Selection

The last section talked about how oligonucleotide synthesis has become much simpler over the years. Researchers can now generate large pools of degenerate oligonucleotides with much better yields than what one was used to more than forty years back. However, the requirement of a selection experiment is not huge due to the nature of *in vitro* selection.

Aptamers (RNA, DNA and modified nucleic acids) and RNA/DNAzymes are a promising class of therapeutics, based on single-stranded oligonucleotides, that are isolated using an *in vitro* selection and evolution process. Such an iterative process is known as Systematic Evolution of Ligands through EXponential enrichment (SELEX). Three decades of work has provided sufficient interest in this technique to produce FNAs containing modified nucleotides[24]. A flow chart of a generic selection toward an aptamer has been shown in Fig 1.7. SELEX experiments usually start from an initial library containing over ~10^{15} different sequences[25]. The number of possible sequences of a nucleic acid of length n is 4^n. Therefore, a combinatorial library containing one copy each of a 25mer library would contain about 10^{15} molecules. From the pool of oligonucleotides, molecules that conform to the desired phenotype accumulate over generations of "selection" and those that do not are eliminated from the population. The collected fraction is then amplified using PCR to scale up the size of the population to the original level which provides an exponential amplification[25–27]. The selection process is then repeated until the desired phenotype is retained at the same ratio and no improvement is achieved by further iterations. In conventional selection, usually more than one generation is required since there is a small, yet non-zero probability for inactive sequences to survive. Therefore, to eliminate most inactive sequences and have a population that can competently deliver the desired phenotype several rounds of selection are

conventionally preferred. As the selection proceeds the reaction parameters such as temperature, time, pH and metal concentration are made more stringent to favor the development of molecules that are adapted to the adjusted condition. The selected aptamers from the final conclusive generation are then sequenced. Best of the clones, after bioinformatics and preliminary characterization experiments, are then studied in depth through various assays by varying the temperature, time, pH, metal ion concentration, and substrate concentration to name a few.

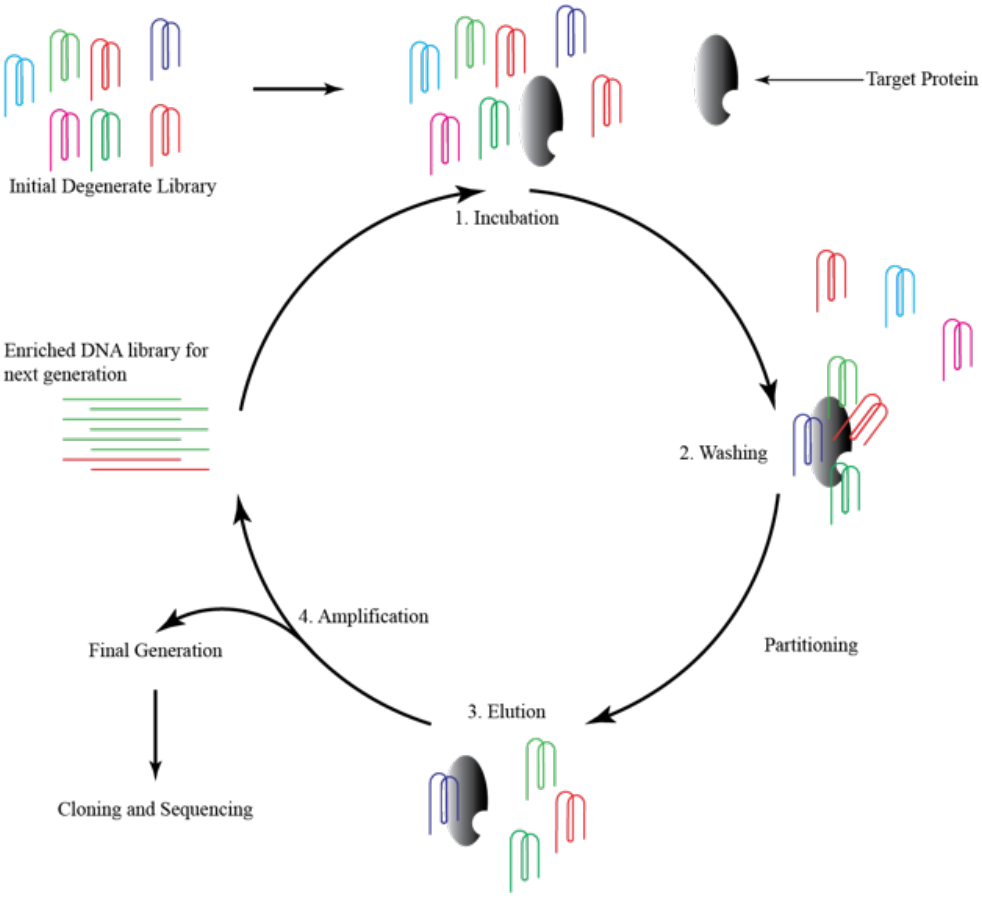


Figure 1.7 A general schema for *in vitro* selection

1.2 Aptamers and DNAzymes

1.2.1 A brief introduction to aptamers

Aptamers are single-stranded nucleic acids (DNA or RNA) that can bind to a range of molecular structures including – small molecules[25], proteins, metal ions, viruses[28], and whole cells by a process known as SELEX as discussed in the previous section. Table 1.1 presents a selected list of aptamers to showcase the range of targets and different types of aptamers that have been selected.

Table 1.1 Aptamers that have been selected under the previously mentioned categories

Aptamer Target	Binding Constant	Aptamer Type	References
Cibacron Blue 3G-A	< 100 μM	RNA	[25]
Tachykinin (Subst. P)	190 nM	RNA	[29]
ATP	5 μM	RNA	[30]
HIV-Tat 1	0.12 nM	RNA	[31]
Immunoglobulin IgE (Human)	30 nM	DNA	[32]
Lead ions (Pb^{2+})	10 nM	DNA	[33]
Ochratoxin A	50 nM	DNA	[34]
Avian Flu (H5N1) particles	4.65 nM	DNA	[35]
Acute Lymphoblastic Leukemia (CCRF-CEM)	0.8 nM	DNA	[36]

1.2.2 Advantages and Limitations of Aptamers

Aptamers have a combined advantage over small molecules and antibodies as they can be discovered and optimized relatively rapidly, and are also chemically synthesized with ease at a lower cost and often enable efficient and homogeneous conjugation at a given base or one of the termini. Further, they also have a much reduced risk of immunogenicity. Although aptamers designed using unmodified DNA and RNA are readily generated using SELEX, they are prone to nuclease mediated degradation and also to a lesser extent toward chemical degradation (particularly RNA-based aptamers)[24]. This may be addressed at least in part by using modified nucleotides. Aptamers that have been generated using such medicinal chemistry approaches may be far more stable, however, the modifications important for nuclease and chemical stability may also weaken desired interactions between the aptamer and the target. Therefore, more recently, functional nucleic acids have been selected with the desired modifications in place (termed modified nucleotide SELEX). Modified nucleotides are also able to produce high affinity and high specificity DNA and RNA aptamers toward a wide variety of targets[37].

1.2.3 Historic and current advancements in the therapeutic potential of aptamers

Qß and MS2 viral coat proteins have been historically some of the earliest targets recognized by RNA hairpin structures that researchers have identified as RNA recognition motifs[16,38]. *In vitro* evolution has been extensively studied in the Qß bacteriophage system[39]. The RNA-directed RNA polymerase of Qß was one of the earliest evolutionary systems studied and there were seven key papers published by Spiegelman *et al*. that highlight the originating idea of Darwinian selection of RNA motifs that could recognize Qß[40–46]. They performed RNA replication experiments *in vitro* using a purified replicase from the Qß phage. The purified Qß replicase yielded small RNA

products that could accurately and rapidly undergo replication. For instance, one of the mutant RNA variants that they isolated (V-2) replicated faster (when initiated with a 3000 times lesser amount of material) than their previously identified variant (V-1) in a 15-minute experiment. They altered the conditions (selective conditions) of the environment to isolate "good" RNA species. In one of their experiments, they started with a small RNA of 218 nucleotides and those that outperformed the "wild-type" species, in replicative ability, when the selection condition (addition of 15 μM ethidium bromide) was altered were selected as the mutant population[46]. Spiegelman's early experiments focused on the binding and the catalytic potential of the purified replicase. They established that Qß replicase was essential for replicating both the mutant and the normal RNA while RNA was the actual instructive agent. Since the RNA polymers satisfied the definition of a self-duplicating entity they served as a precursor to experiments involving Darwinian evolution.
In 1990 Tuerk and Gold used this idea for their selection but separated the binding reaction from the extension[39]. They used the binding as the selective pressure and substituted the extension reaction of the Qß replicase with the combined reaction of reverse transcriptase, Taq DNA polymerase and T7 RNA polymerase to generate efficient RNA aptamers to bacteriophage T4 DNA polymerase[39]. This was a watershed event as this work demonstrated one of the earliest examples of selecting an aptamer toward a protein target. Along the same time, Szostak and Ellington also selected RNA molecules (aptamers) that bind to a variety of organic dyes (small molecule ligands) from a random pool of sequences (RNA library)[26].

Subsequent to these seminal works, a fairly large amount of early DNA and RNA aptamer selections were directed toward the protein thrombin. Thrombin binding has been one of the most commonly used systems to demonstrate a proof-of-concept for aptamer assays. One of the first

aptamers to thrombin was a 15-mer DNA aptamer that can form a stable intramolecular G-quadruplex structure, which has a chair-like conformation in an antiparallel orientation[47]. Several of the clones from the selection had a dissociation constant (K_D) ≈ 200 nM while the original sequence had little affinity toward thrombin. There was another 29-mer thrombin binding aptamer to the heparin-binding exosite of thrombin, which also had a G-quadruplex structure, that had a higher affinity ((K_D) ≈ 0.5 nM)[48]. There have also been two thrombin-specific DNA aptamers (NU172 and ARC183) that have been studied in clinical trials but to little progress toward a market approval[49]. There have been numerous aptamers selected for various targets, yet few are of therapeutic importance, some of which have been highlighted in Table 1.1. Currently, however, the only FDA approved, and commercially available aptamer is Macugen™ which was developed by OSI Pharmaceuticals, USA towards Age-Related Macular Degenerative disorder (ARMD). However, other aptamers are currently in clinical trials like Zimura® and Fovista® for dry ARMD. Nox-A12 for Colorectal and Pancreatic cancer and AS1411 for Acute Myeloid Lymphoma have also been in clinical trials for a while. Table 1.2 presents several of these aptamers that are currently in the process of clinical trials.

Table 1.2 Aptamers currently in Clinical Trial and their targets

Aptamer	Target	Applications	Status
Avacincaptad pegol (Zimura®)	Complement protein C5	Geographic Atrophy Macular Degeneration	Clinical Trial Phase 2
Pegpleranib (Fovista®)	Platelet-derived Growth Factor B	Wet ARMD	Clinical Trial Phase 3 (Terminated)
Olaptesed pegol (Nox-A12)	CXCL 12	Metastatic Colorectal Cancer	Clinical Trial Phase 1
Olaptesed pegol (Nox-A12) Combination Therapy	CXCL 12	Metastatic Pancreatic Cancer	Clinical Trial Phase 2
AS-1411	Nucleolin	Metastatic Renal Cell Carcinoma	Clinical Trial Phase 2
The current status of the drugs have been updated according to the US National Library of Medicine Clinical Trials record			

1.2.4 Historic development of DNAzymes

Many of the nucleic acid enzymes that have been studied are primarily RNA based due to the precedence of naturally occurring ribozymes and the general belief that the hydroxyl group at the 2' position makes them a superior catalyst. According to Cech[50,51] and Rich[52], it is believed that the 2'-hydroxyl group as a hydrogen bond donor and acceptor gives the RNA more versatility in forming secondary and tertiary structures. But, in reality, it has been confirmed that catalytic rate enhancement with DNAzymes lies within the same order of magnitude of ribozymes, which has been summarized in separate reviews on ribozymes[53] and DNAzymes[54] by Silverman.

DNA catalysis was first demonstrated by Breaker and Joyce in 1994; their assumption being DNA can have catalytic activity comparable to that of ribozymes considering it generally has the same functional groups as RNA except for the 2'-hydroxyl group. Their work featured the discovery of the first Pb^{2+}-dependent DNAzyme that was selected to catalytically cleave a ribophosphodiester bond[55]. Since this discovery, there have been several more DNAzymes isolated that catalyze as wide a spectrum of targets as their ribozyme counterparts. Aside from RNA cleavage[56,57] and ligation reactions[58], which are well within the scope of DNAzyme catalysis, there were several DNAzymes selected that include: (i) DNA cleavage reactions by an oxidation mechanism using Cu^{2+} as a cofactor[59,60] and alternative hydrolysis mechanisms using Zn^{2+} or Mn^{2+} as cofactors[61,62], (ii) porphyrin metalation[63,64] by introducing Cu^{2+} and Zn^{2+}, and (iii) phosphorylation of the tyrosine residue of peptides using an ATP bound aptamer[65] (iv) DNA-catalyzed DNA cleavage[62,66] among many more[67–69]. The focus of the thesis is the development of divalent metal ion independent RNA-cleaving DNAzymes with amino acid side chain modifications that are undertaken largely due to a promising application potential in targeted RNA therapy.

1.3 Modified Nucleotides

1.3.1 History of Modified Nucleotides

Modifications introduced in nucleotides during the 1980s were usually made on the phosphate backbone or sugar moiety and rarely on the nitrogenous base. The early applications for such modified nucleotides were mostly in designing affinity probes for photo-crosslinking agents, and fluorescent probes[70–75]. Recalling that researchers often used modified nucleotides to impart chemical and physical stability to nucleic acids, such introductions often weakened the desired interactions between the aptamers and their targets (Section 1.2.2). This can be evident from the

early development of synthetic ribozymes which were mostly backbone modified (amino, fluoro, or an O-methyl group at the 2' position of the ribose sugar) to reduce nuclease mediated degradation (see Fig. 1.8). A significant advantage of DNAzymes over ribozymes is that DNA lacks the 2'-hydroxyl group found in RNA, and consequently is largely inert to cleavage through hydrolysis or transesterification by small-molecule catalysts, with a few notable exceptions[76–80]. More recently functional nucleic acids have been selected with the desired modifications in place (termed modified nucleotide SELEX). Ever since modified nucleotides can produce high affinity and high specificity DNA and RNA aptamers toward a wide variety of targets[37].

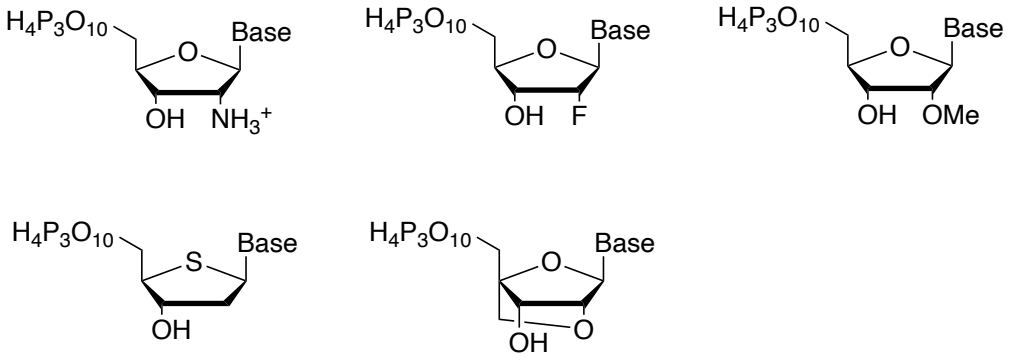


Figure 1.8 Some early sugar modified nucleotides with variations in the ribose sugar.
The 2'-ammonium, 2'-fluoro and the 2'-O-methyl nucleotides (in the top row) have been used in RNA polymerization, whereas the nucleotides in the bottom row are available for DNA and RNA based nucleic acid catalysts.

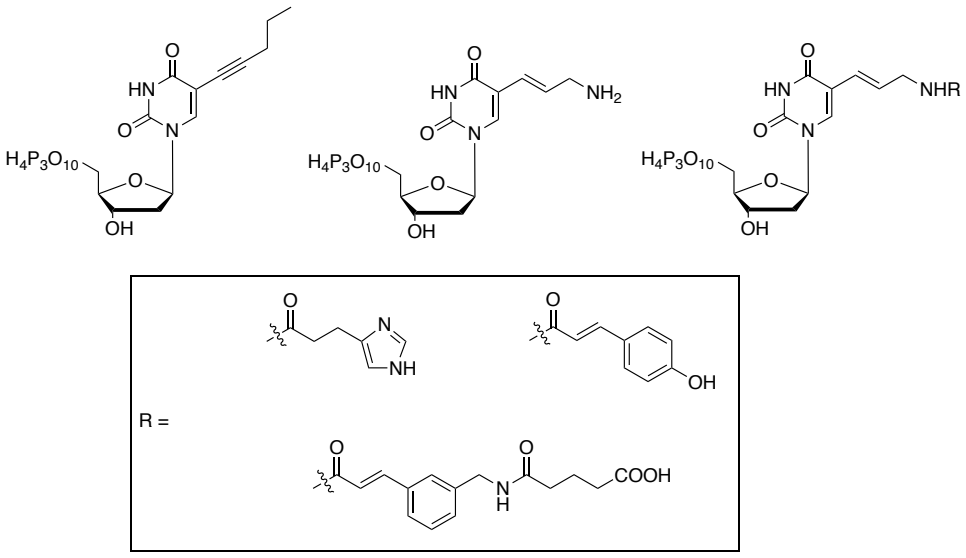


Figure 1.9 Early modified nucleotides on the nitrogenous base.
From left to right, pentynyl dUTP was the first modified nucleotide used in an aptamer developed by Latham et. al., (1998), amino-allyl dUTP ($dU^{aa}TP$) and the nucleotide analogs using it as a starting material were early base modified nucleotides that were developed by Sakthivel and Barbas, containing different positive and negatively charged functional groups.

One of the earliest uses of a modified nucleotide in DNA aptamers was pentynyl-dUTP in the selection of a thrombin binding aptamer[81]. Subsequently, Sakthivel and Barbas III developed a variety of dUTP derivatives with primary amines, imidazole and amide functional groups on the nitrogenous base, intended for developing nucleic acids with a wide array of binding and catalytic opportunities (Fig 1.9)[82]. A key paper for designing high-affinity DNA aptamers was published by Larry Gold's SomaLogic in 2010. There have been several high-affinity RNA and DNA aptamers selected towards various human proteins (as discussed previously). There are also several examples of difficult to select targets. Gold *et al*. (2010) pioneered a new class of aptamers called Slow Off-rate Modified Aptamer (SOMAmer) which could be used to select aptamers against such difficult to identify targets that have even found clinical application[83,84]. This was based on the idea that aptamers could be endowed with protein-like functionalities.

Gawande *et al*. contributed to the development of SOMAmers (aptamers that rely on modified nucleosides) while working at SomaLogic, a company founded by Larry Gold to support research and development of aptamers. Gawande *et al*. further developed several aptamers with two modified bases where they showed large hydrophobic functional groups have been effective against a wide range of protein targets especially among SOMAmers[85–87]. Paralleling this work, Gierlich *et al*., (2007) designed modified nucleotides with an alkyne functional group which finds potential application in "click" chemistry[88]. Günther Mayer *et al*. (2018) developed the alkyne modified nucleotides into a well-established technology that relied on commercially available building blocks to introduce a wide array of modifications including large hydrophobic modifications developed for SOMAmers. This makes Click-SELEX available to laboratories even without access to in-house synthesis[89–91]. With these early research initiatives in modified nucleotides, it opens up possibilities to incorporate many more, novel functionalities that could potentially be useful in the *in vitro* selection of nucleic acid binding and catalysis.

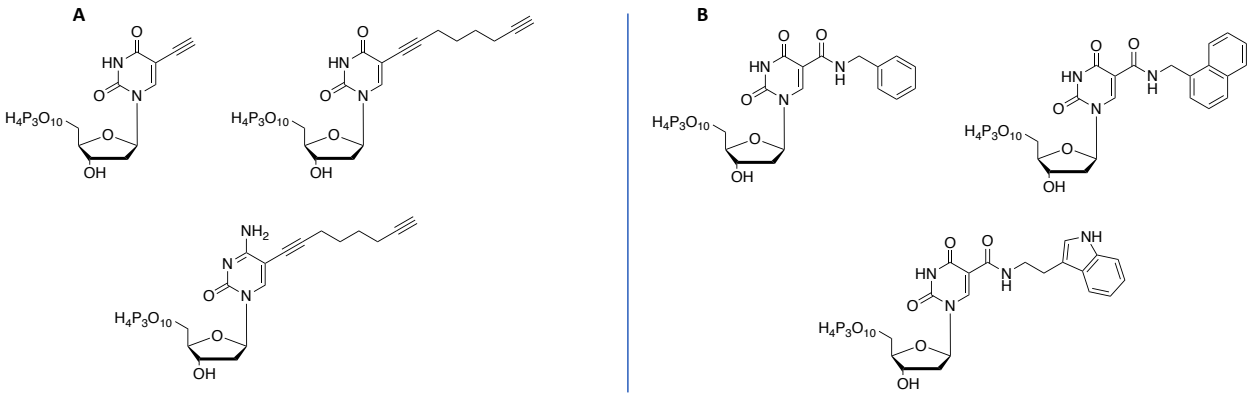


Figure 1.10 Modified nucleotides for Click chemistry and those discovered by SomaLogic.
Crosslinking nucleotides (A) can be attached to novel functional groups through Click Chemistry. Some modified nucleotides developed by SomaLogic (B) to produce high-affinity aptamers (In a clockwise direction, the molecules are benzyl-dU, naphthyl-dU, and tryptamino-dU).

1.3.2 Modified Nucleotides in DNAzymes

The *in vitro* selection of RNA-cleaving DNAzymes was first demonstrated by Breaker and Joyce using natural nucleotides in the presence of metal ions, which has then been widely used by several groups[55,92]. Wiegand *et al*. first selected an RNA-catalyzed DNAzyme with an imidazolyl linked to the 5' terminus to catalyze an RNA amide synthesis, where Cu^{2+} ions caused a change in the affinity of the RNA to the substrate rather than being directly involved in the amide-bond formation[93]. Herein the imidazole group proved to be crucial for its catalytic activity but divalent metal ions like Cu^{2+}, Mg^{2+}, Ca^{2+} and Zn^{2+} were essential for its function. In the absence of metal dependence, an alternative approach was conceptualized that involves co-factor recruitment in the presence of high levels of histamine which served as a cofactor for M^{2+}-free cleavage[57].

An alternative might be that the functionalities can be introduced on the monomer dNTPs during the combinatorial selections, however, two conditions must be satisfied: i) they must be substrates for various polymerases and ii) the modified strands must serve as templates that can be accurately recopied. Most DNAzymes that have been selected, however, require Mg^{2+} as it is critical to their structure and/or catalytic activity in contrast to ribozymes[94–96]. Many DNAzymes unfortunately require at least ~ 10 mM Mg^{2+} for their optimal activity[97]. Few DNAzymes function well using ~ 0.1 mM Ca^{2+}/Zn^{2+} as well. However, intracellular requirement for Mg^{2+} concentration is low (< 0.5 mM, read cytosolic)[98–100]. Therefore, M^{2+}-independent DNAzyme is likely to provide a more efficient route to *in vivo* gene regulation through catalytic RNA cleavage.

Independent efforts by Faulhammer and Famulok[101], Sen and Geyer[102], as well as Benner *et al*. [103] were successful in selecting metal independent DNAzymes, however, all of which had meagre self-cleavage constant between 10^{-4} to 10^{-3} min^{-1}. The catalytic rate of two of the DNAzymes was similar to the hammerhead ribozyme in the absence of any M^{2+} ions except for Benner's DNAzyme which had a very slow turnover (~ 0.02 h^{-1}). Such early findings demonstrate the challenges in the in-vitro selection of metal-independent DNAzyme cleavage and therefore even among the limited number of metal-independent DNAzymes there are only a few enzymes that have shown a rate constant faster than 10^{-3} min^{-1}. The three independent labs deliberately tried to tackle this challenging problem to select self-cleaving DNAzymes under M^{2+}-free conditions. Furthermore, they were only fully operative at high concentrations of monovalent cations (0.25-1 M Li^+ or Na^+). Not long after, Torabi *et al*. reported a DNAzyme that exhibits surprisingly high activity for intra- and intermolecular ribophosphodiester bond cleavage in the presence of Na^+ alone[104]. In 2001, Perrin *et al*., developed a metal independent RNAse A mimicking DNAzyme, $Dz9_{25}$-11, which contained two modified nucleotides – a histaminyl-modified dATP along with amino allyl dUTP[105]. It was observed that using the modified nucleotides it was possible to achieve k_{cat} values 25-100-fold over the unmodified selections. After the selection attempt of their metal independent DNAzyme, the Perrin group has selected DNAzymes Dz9-86 and Dz10-66, which are both modified with three unnatural nucleotides[106,107]. Apart from the histaminyl-dATP and aminoallyl-dCTP it also contained a guanidinium-dUTP (Chapter 2). Using the third modified nucleotide they were able to select a DNAzyme with one of the highest M^{2+} independent self-cleavage which was also able to function under low concentrations of monovalent ions that is usually found in highly efficient protein enzymes like RNase A. With the selection of these fast self-cleaving enzymes the group selected for RNA cleaving DNAzyme which acts as a mercury (Hg^{2+}) sensor[108]. Mercuric

cation (Hg^{2+}) is known to stabilize a DNA duplex by intercalating between two thymine nucleotides, as thymine-Hg^{2+}-thymine, leading to an erroneous DNA replication. This is a property that has allowed a lot of researchers to develop nucleic acid-based sensors for mercury. Further, Perrin *et al*. have also identified a fast biphasic (one phase that has a superior catalytic rate over a population of inactive misfolded species) DNAzyme using the triple modified system that can cleave with multiple turnover[109,110]. The group has also worked with other modified nucleotides for selecting aptamers for whole cells like phenol dUTP[111].

However, this thesis aims to use two of the modified nucleotides (5-(3-guanidinoallyl)-2'-dUTP ($dU^{ga}TP$) and 5-aminoallyl-2'-dCTP ($dC^{aa}TP$), excluding an imidazole) for selecting both aptamers and catalysts. The first objective of my thesis, identifying a metal-independent RNA cleaving DNAzyme using the modified nucleotides agrees with our hypothesis (as will be elucidated in Chapter 2). Further, an allosteric DNAzyme (also known as a self-cleaving DNAzyme in the presence of a binding substrate) and an aptamer were attempted, however, leading to no successful results. A concluding chapter discussing future directions and a brief explanation of the failure of the selection will be attempted in Chapter 4.

1.4 Specific Objectives

1. Develop a self-cleaving DNAzyme with two modified nucleotides (dU^{ga}TP and dC^{aa}TP). (Chapter 2)
2. Characterization of the DNAzyme based on temperature, pH profile, substrate specificity and buffer conditions. (Chapter 2)
3. Concluding chapter discusses future directions from this work and possible reasons why my aptamer selections (not included in the thesis) may not have worked (Chapter 4)

Chapter 2: A self-cleaving DNAzyme with two modified nucleotides

This chapter describes my effort in an early attempt to select an M^{2+}-independent self-cleaving DNAzyme that is modified with cationic amines and guanidines but lacks an imidazole functionality. Sequence-specific cleavage of RNA by nucleic acid catalysts in the absence of a divalent metal cation (M^{2+}) has been elusive for a long time. Due to the lack of functional group diversity in catalytic RNA and DNA, modified nucleotides with amino acid-like side chains have been used to enhance self-cleavage rates at a single embedded ribonucleotide site. Previous efforts have relied on up to three modified nucleotides including an amine, a guanidine and an imidazole functional group. However, there has not been any systematic study on the necessity of all three modifications, as the value of each modified nucleotide is contextualized at the initiation of the selection.

In this thesis, I report the use of two modified dNTPs, excluding the imidazole, i.e. 5-(3-guanidinoallyl)-2'-dUTP ($dU^{ga}TP$) and 5-aminoallyl-2'-dCTP ($dC^{aa}TP$), to select *in-vitro* self-cleaving DNAzymes that cleave in the absence of M^{2+} in a pH-independent fashion while exhibiting biphasic kinetics for self-cleavage with rate constants that are significantly higher than in unmodified DNAzymes and also compared favorably to certain examples involving an imidazole.

2.1 Introduction

In the last chapter, I overviewed the vast landscape of modified nucleotides available to researchers working on different facets of nucleic acid chemistry (aptamers and DNAzymes). This chapter, however, is exclusively focused on DNAzymes. Before delving into the DNAzymes developed by

our laboratory, the following serves as a primer to discuss selected DNAzymes of exceptional chemical importance to date. The focus on RNA cleaving DNAzymes is not a new trend in the field of catalytic nucleic acids research as it has been elucidated in Chapter 1 of the thesis. Besides an enduring interest in DNAzymes, the question of M^{2+}-independence in catalytic DNA species was first addressed with the initial discovery of the 33-nucleotide catalytic DNAzyme (PS5.ST1) in 1997 in the Sen lab at Simon Fraser University[112]. As the field of DNAzymes grew, the catalytic repertoire of DNAzymes has been expanded into various reaction manifolds to include Diels-Alder reactions[50,113], amide bond formation[93], Click chemistry[89,114–116], phosphatase[117], phospho-serine hydrolase[118], tyrosine kinase phosphorylation[119] and aldol condensation reactions[120].

In contrast to the commercial and medical applications of aptamers selected against biological targets[121,122], the selection of DNAzymes has not received as much attention beyond the academic community. One may speculate the reason behind this to be: the high concentration of metal cations used during in-vitro selections of DNAzymes inevitably leads to DNAzymes with metal cation dependencies. DNAzymes have featured prominently as metal sensors for use in environmentally-focused applications. Development of DNAzymes as catalytic beacons, for activity-based sensing of heavy metal cations such as – Hg^{2+}, Pb^{2+}, Co^{2+} $Cu^{2+,}$ and lanthanides, was quite a feat at the time of its first discovery, leading to many follow-up works[123,124]. In 2007, researchers in the Liu lab further designed a catalytic sensor for uranyl ions (UO_2^{2+}) with an ultra-high sensitivity in the order of parts per trillion and high selectivity for uranium[125]. Hg^{2+} was another important heavy metal cation for which researchers have developed sensors, and is important in explaining the thesis due to our lab's prior work on a Hg^{2+} sensing DNAzyme with imidazole as one of the functional groups. The Hg^{2+} ions bind to thymine nucleobases by

intercalating between two DNA strands, which was an early inspiration for scientists to design metal-dependent sensors as a concept[126,127].

While metal cation sensing has been an important endeavor with DNAzymes, the physiological concentration of Mg^{2+} is 0.49 ± 0.03 mM in the cytosol, which begs the question as to the utility of metal-cation dependent DNAzymes in cells. Of particular concern is the question of creating therapeutic DNAzymes for cleaving pathogenic mRNA targets e.g. viral genomes. It is thus an important consideration when selecting DNAzymes under low divalent metal ion concentrations[98–100]. Several RNA cleaving DNAzymes have been identified at high Mg^{2+} concentrations larger than 5-10 mM[92,128–133]; when they were interrogated under Mg^{2+} free conditions, the catalytic rate constants lie in the range 10^{-3} to 10^{-2} min^{-1} [134]. However, there does exist a counterexample: a DNAzyme that was selected under low Mg^{2+} concentration which nonetheless requires a high concentration of monovalent cations i.e. 0.25-1 M Li^{+} or Na^{+}[104].

Ideal modifications for discovering such M^{2+}-independent RNA cleaving DNAzymes would comprise functionalities that are normally found in the active sites of ribonucleases and phosphodiesterases[135,136]. These include amine, guanidine, and imidazole functional groups that can deliver their activity through acid-base catalysis and electrostatic complementarity. Chemical synthesis of modified nucleotides with guanidinium, primary ammonium, and imidazole groups have therefore been crucial to developing catalysts that work under low divalent metal cation concentrations. An early approach in that regard has been cofactor recruitment where high levels of histamine (toxic to cells) have been used to deliver M^{2+}-free cleavage (presumably through a well-positioned imidazole)[137,138]. Important parameters to consider an RNA cleaving DNAzyme

as an ideal therapeutic candidate are based on the fulfilment of the following criteria: (i) must be functional under low Mg^{2+} concentration within a physiological pH range, (ii) it can cleave an all RNA target in a base-specific manner, and (iii) must have a high catalytic turnover rate.

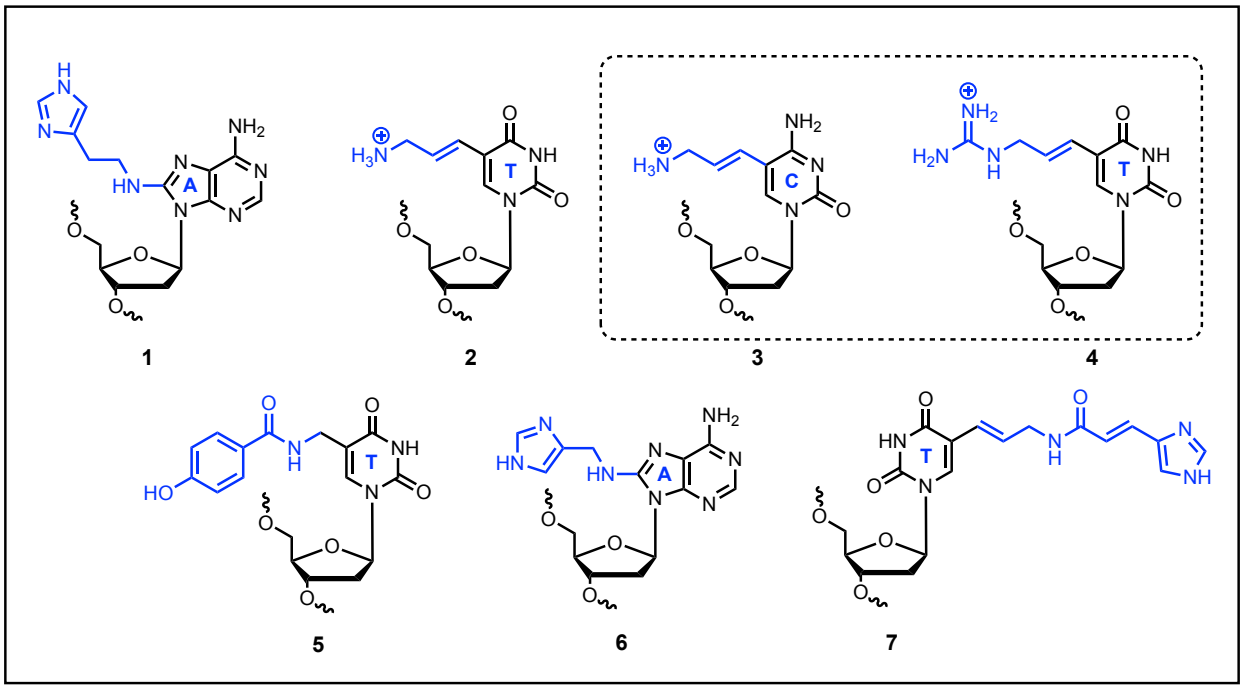


Figure 2.1 Modified nucleosides in previously reported M^{2+}-free DNAzyme for self-cleavage.
They have been used for cleavage at an internal ribose. Compound 1 is 8-histaminyl-deoxyadenosine, compound 2 is 5-aminoallyl-2'-deoxyuridine a precursor of compound 4, compound 3 is 5-aminoallyl-2'-deoxycytidine, compound 4 is 5-guanidinoallyl-2'-deoxyuridine, compound 5 is phenol-deoxyuridine (dU^{y}), compound 6 is 8-imidazolyl-methyleneamino deoxyadenosine and compound 7 is a C5-imidazole functionalized dUTP analogue.

The first set of successful approaches disclosed M^{2+}-free RNA cleaving DNAzymes containing imidazole and an amine, introduced in the form of 8-histaminyl-deoxyadenosine (**dA^{his}, 1)** and 5-aminoallyl-2'-deoxycytidine **(dC^{aa}, 3)** (Figure 2.1), and combination with a guanidine as 8-histaminyl-deoxyadenosine (**dA^{his}, 1),** 5-aminoallyl-2'-deoxycytidine **(dC^{aa}, 3),** and 5-guanidinoallyl-2'-deoxyuridine **(dU^{ga}, 4)** (Figure 2.1) in our lab and other competing labs. The

focus of our lab ever since has been around the development of nucleic acid enzymes containing such nucleotides that have amino acid-like functional groups. One of the earliest cases of M^{2+}-independent, modified RNA-cleaving DNAzymes ($Dz9_{25}$-11) by Perrin *et al*., resulted in a k_{cat} of 0.044 min^{-1} at 37 ºC that had been selected using **dA^{his} (1)** and 5-aminoallyl-2'-deoxyuridine **(dU^{aa}, 2**)[105]. Respectively, these represent mimics for histidine and lysine amino acid residues found within the active site of RNase A. Subsequently, Lermer *et al*., reported the *trans* cleaving potential of a minimal motif of $Dz9_{25}$-11 demonstrating a diminished k_{cat} of 0.015 min^{-1} compared to its intramolecular counterpart[139]. Nevertheless, their experiment showed a high substrate specificity when cleaving an identical template (substrate) that is degenerate for ribose which also cleaves with multiple turnovers.

In 2008, Hollenstein *et al*. identified an M^{2+}-free (without Mg^{2+}) DNAzyme where one of the modified nucleotides was **dA^{his} (1)**, which could self-cleave an internal ribonucleotide site in the presence of Hg^{2+} ions with an equilibrium binding constant $K_D \approx 1$ μM for Hg^{2+}. Lam *et al*., further discovered an M^{2+}-dependent DNAzyme using **dU^{y} (5)**[140]. The different strategies attempted were aimed at unearthing nucleic acid enzymes that provide low cost of development with high catalytic activity. In addition, electrostatic interactions can provide a stabilizing effect through a highly basic alkyl guanidium group as found in the **dU^{ga} (4)**.

The modified nucleosides, **dA^{his} (1)**, **dU^{aa} (2)**, and **dC^{aa} (3)** have been incorporated in different combinations along with **dU^{ga} (4)** in DNAzymes with two or three modifications; these have led to catalysts with multiple catalytic turnovers and also as metal ion sensors.[105,107,108,139,141] Catalytic nucleic acids such as Dz9-86, Dz12-91, Dz11-17PheO, and Dz7-38-32 have been discovered with

decent k_{cat} values; the nucleoside **dU^{ga} (4)** in addition to **dA^{his} (1)** and **dC^{aa} (3)** have dramatically increased the rate constant to > 0.6 min^{-1} at 37-39 ºC [140–142]. Further, Dz7-38-32, which uses **dA^{his} (1)**, **dC^{aa} (3)**, and **dU^{ga} (4)** demonstrate a biphasic kinetic rate plot for self-cleavage which implies a two-phase model involving two conformations. One fraction of the DNAzyme undergoes rapid catalysis while the other lags behind due to what is most likely an inactive conformation. The fast phase in the above-mentioned case could self-cleave an internal ribose site at a k_{cat} of 4.9 min^{-1} and when the multiple turnover cleavage *in-trans* was interrogated it could cleave a therapeutic all-RNA target with multiple turnovers that were over 1 min^{-1} at 25 ºC [141].

Previous studies in our lab have shown that the imidazole group in **dA^{his} (1)** likely plays the role of acid-base catalysis implicated by a characteristic bell-shaped curve around the pKa of the imidazole. When the modified nucleotide was removed, the DNAzyme cleaved poorly. Yet such evidence can only be validated contextually since there could be several other mechanisms like secondary and tertiary folding conformations, long-range H-bond networks, and pKa perturbation that is difficult to distinguish from acid-base catalysis.

In summary, the DNAzymes selected previously have generally followed the important parameters of a DNAzyme as discussed earlier. They however either use all the three modified substituents– or two of them – imidazolyl and aminoallyl. While the **dA^{his} (1)** involves a challenging chemical synthesis including protection and deprotection of nucleophilic groups on the nitrogenous base and the sugar, modified analogs of dUTP and dCTP can be used in a plug-and-play approach while installing amino acid like side-chain modifications.

Our study is a preliminary investigation to determine whether modified DNAzymes with two nucleosides **dC^{aa} (3)** and **dU^{ga} (4)** (Figure 2.1) can produce comparable cleavage to analogous DNAzymes with two/three modifications (inclusive of an imidazole). This project aims to show that the two modified nucleotides could also provide quite active catalysts with k_{cat} values comparable to those of the previously selected DNAzymes in our lab. The project also cites the importance of selecting catalysts for targeting viral RNAs which is of great significance given the imminent danger posed by the SARS-CoV-2 RNA virus around the world.

2.2 Optimizing Primer Extension for full-length synthesis

Unlike in-vitro selection procedures employing canonical nucleotides, modified nucleotides must be incorporated by one-sided primer extension reactions due to the limited availability of the nucleotide and the desire to limit bias that can be exacerbated by standard PCRs. A standard PCR can exponentially amplify any bias that the polymerase might exert against the incorporation of the modified substrates or while reading through the resulting modified DNA. The modified substrate **dU^{ga}TP (4)** is not commercially available and must be synthesized in the laboratory, which involves an expensive precursor and a low final yield. The complete DNAzyme product synthesized by primer extension is not always incorporated perfectly in the case of many modified nucleotides due to such nucleotides being poor substrates for many of the polymerases. This has been addressed in our previous works[143]. As studied by Dr Curtis Lam, a former member of the Perrin lab, **dU^{ga}TP (4)** can be incorporated by Sequenase V2.0, Vent (exo-), Pfu (exo-), and Klenow (exo-) including Taq Polymerase however the latter also yields several shorter truncates. Their study demonstrated a clean full-length extension of **dU^{ga}TP (4)** however with some gel smearing contributed in part by **dA^{his}TP (1)** as it is a poor substrate for many polymerases. Our

selection and the preceding primer extension, however, does not include the nucleoside **dA^{his}TP (1)**. Therefore, we wanted to address that the incorporation of the nucleotides during primer extension was complete (e.g. with fewer truncated artifacts observed).

The concentration of Klenow polymerase was not determined as it had not been extensively used in prior selections. Further, earlier works in our lab used 1 µL polymerase in a 20 µL incorporation reaction and no scale-up concentration had been determined for larger-scale primer extensions. We found a concentration of 0.25 U/µL of Klenow Polymerase was sufficient for the complete utilization of the primer while we found that a concentration of 0.8 U/µL of Klenow Polymerase was the maximum acceptable amount before yields diminished. A concentration of enzyme higher than that led to erroneous incorporation possibly due to non-specific extension. The nucleotide concentration was also adjusted as a parameter for primer extension to maximize the incorporation of starting material while minimizing truncated artefacts. From previous published (and unpublished data in our lab), a nucleoside triphosphate concentration of 50 µM was shown to be an effective concentration of **dU^{ga}TP (4)** and **dC^{aa}TP (3)** for oligonucleotide incorporations up to 99 bases (data in Appendix, Fig B.1). From our optimization studies, we determined that a concentration of at least 30 µM of the unmodified nucleotides was also needed to enable a full-length product (during Klenow primer extension) even for DNA templates of 80 bases in length. If a full-length product is not observed, sequences may be lost as artefacts which could act as false-positive results during selection.

2.3 *In-vitro* selection of DNAzymes

2.3.1 Design of selection scheme

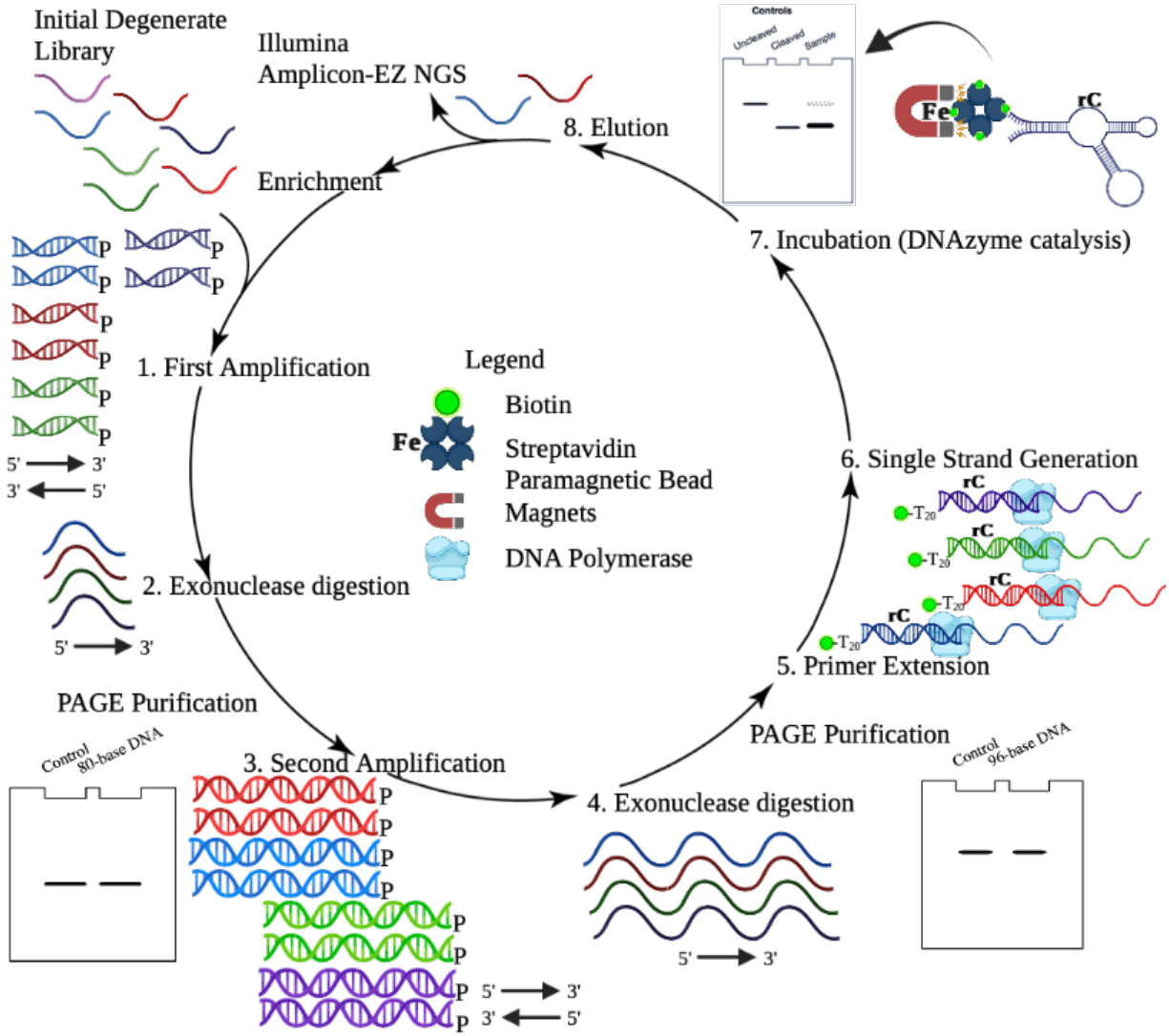


Figure 2.2 *In-vitro* selection of RNA cleaving DNAzyme at an internal ribophosphodiester.
In-vitro selection channels a population of a degenerate library (depicted in different colors) to a narrow population of active sequence (depicted here in red and blue) post elution. In the incubation (step 7) to the elution (step 8) step, only a single species of catalyst has been shown for demonstration. The cleaved product after incubation is collected and purified by PAGE. The internal ribonucleotide cleavage site in the DNA strand is ribocytidine (rC) which is part of the extending primer. Created with BioRender.com

The general scheme of my DNAzyme selection is described in Fig 2.2 that starts from a degenerate library. The template sequence for primer extension (to generate the DNAzyme) is prepared from an initial 80-base degenerate library that is amplified over a set of nested PCRs (first amplification and second amplification). In the first amplification, a phosphorylated 19-base primer and a standard 17-base primer are used to amplify the initial degenerate library (~80 bases). After a biphasic extraction(phenol/chloroform/isoamyl alcohol) to remove impurities like salts, residual nucleotides and enzymes, the double-stranded DNA (dsDNA) is subjected to exonuclease digestion to isolate a single strand. Lambda Exonuclease catalyzes the digestion of the strand bearing a 5'-phosphate of a dsDNA (into its monomers) thus freeing the non-phosphorylated strand. The preferred substrate for Lambda Exonuclease is a 5'-phosphorylated dsDNA although it can degrade a single-stranded DNA (ssDNA) albeit at a much-reduced rate[144]. The exonuclease digested product is purified by denaturing polyacrylamide gel electrophoresis (D-PAGE).

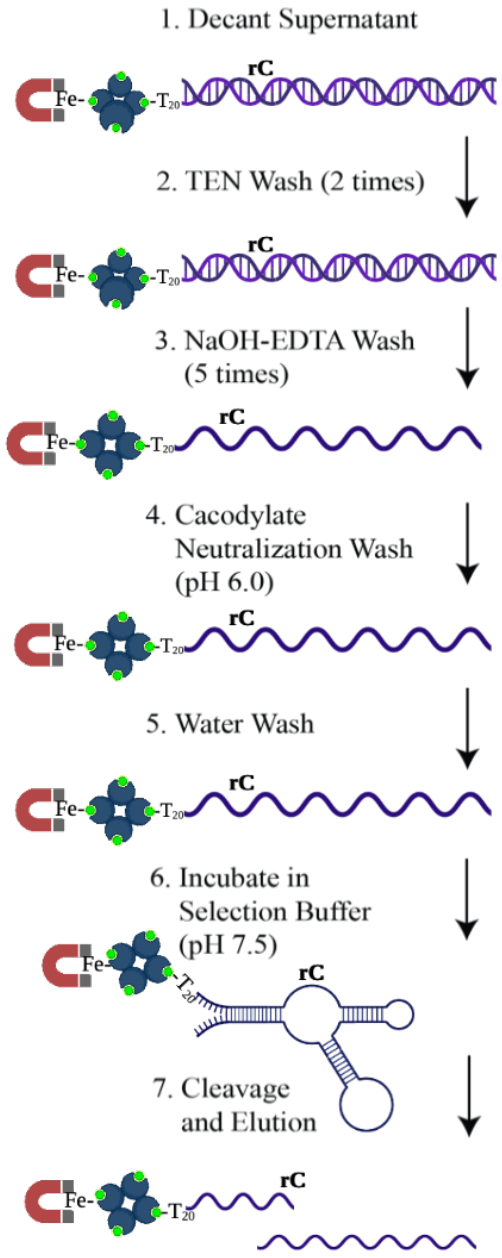


Figure 2.3 Schematic of single-stranded generation of the DNAzyme and incubation.
Single-strand generation (step 6) from Fig 2.2 has been elaborated in this schematic. Image created with BioRender.com

The resulting 80-base library is subjected to a second amplification to generate a 96bp dsDNA (Refer to Table 3.1 Chapter 3 for primers and library) which is further extracted, digested, and then

purified by D-PAGE. The purified DNA strand (template) is extended using a primer containing a 5'-Biotin and an internal ribocytidine to generate the DNAzyme strand that is initially isolated as a heteroduplex containing its corresponding template. The polymerization is quenched and the resulting dsDNA is immobilized on streptavidin magnetic beads (Fig 2.3). After magnetization of the beads, the supernatant is removed. Washes are then performed to remove residual impurities and generate the single-stranded DNAzyme (details in Section 2.3.2 and Chapter 3). This is followed by incubating the isolated DNAzyme in a selection buffer for a stipulated time (discussed in Table 2.1). The cleaved strands are collected and purified by D-PAGE (Fig 2.2). The cleaved and isolated product is again amplified over a set of nested PCR reactions and a new library is prepared for the next generation. After a discernible and fairly sustained enrichment, the selected library is processed and quantified for Next-Generation Sequencing.

For a sufficient sequence diversity, libraries of different lengths (N_n) that range between 20 to 100 bases (n = 20-100) can be considered. A typical selection library usually consists of a degenerate region comprising 40 random nucleotides flanked by two defined sequences that serve as primer binding regions for amplifying the partitioned sequences. The number of nucleotides can provide a theoretical estimate of the number of different sequences that exist in a library pool. As there are four different bases, the existence of an N_{40} library theoretically suggests that there exist 4^{40} different combinations of bases (equivalent to 10^{24} sequences). Due to constraints that will be discussed later, the number of oligonucleotides that are used in a typical modified nucleotide selection is in the order of $10^{11} – 10^{12}$ sequences.

The composition of the degenerate region is another aspect that must be considered. The degenerate library can either be synthesized by absolute randomization of the bases or alternatively generate a library in which each base has an equal probability of occurrence. Both approaches have merits, but in most cases and our instance, a completely randomized degenerate region is used. Having an equal probability of each base could allow for the presence of all possible sequences in the library, which is not required in most cases. As my selection involves modified nucleotide DNAzymes, the modifications should increase chemical diversity even if there may be limitations on sequence diversity.

2.3.2 Preparation of library and *in-vitro* selection

The procedure for my selection has been adapted from Hollenstein *et al*. The selection library was prepared in a buffer of physiological pH to later account for catalysts that could be tested *in-vivo*. The library is a DNAzyme strand that contains a single ribocytidine (rC), which is incorporated using a 3'-terminal extension of a primer that contains the embedded ribonucleotide. In my initial primer extension reactions, I utilized an equimolar concentration of primer and template which I later adjusted to a ratio of 0.5-0.8:1 (primer:template) to ensure complete primer consumption (optimization data in Appendix). The amount of template and primer used was typically less than 20 pmol, which provides about 10^{12} sequences out of a total possible 10^{24} sequences. The primer extension was initially achieved with the polymerase Sequenase v2.0 for the first few generations while in subsequent generations, primer extensions were performed using Klenow (exo-) polymerase, which is sufficient for the modified nucleotides that were used in our current selection.

As described in the schematic (Fig 2.3), the primer consists of a 20 nucleotide thymidine chain (T_{20}) at the 5'-end that is capped with biotin. The extended region provides a suitable mass difference to distinguish the uncleaved nucleic acid from the cleaved product (post cleavage) which can be separated and purified by PAGE. Post primer extension, the resulting dsDNA is immobilized on streptavidin magnetic beads and washed with Tris/NaCl/EDTA (TEN) buffer to remove remnants (primers, unincorporated monomers, and the polymerase) of the primer extension (details in Chapter 3). Before washing the beads, a small amount of sample is collected as an uncleaved control. The bead-bound oligonucleotides are then quickly washed with an NaOH-EDTA solution to separate the template strand from the catalyst and then immediately neutralized with a pH 6.0 cacodylate buffer. The washes are done in quick succession to ensure minimum loss of catalyst due to base mediated cleavage at rC. The oligonucleotide strands are then incubated in a selection buffer (composition elucidated in Chapter 3) for a pre-determined time frame.

The cleaved product was collected in solution after magnetizing the beads. The desired product was then precipitated ($LiClO_4$ in acetone) and washed (with ethanol) to remove excess salt. The precipitated DNA was resolved by 8% D-PAGE. After eluting the cleaved product from the beads, I treated the residual beads with NaOH-EDTA at 95 ºC for 5 minutes (cleaved control). The cleaved band was extracted from the gel with reference to the uncleaved (initial sample collected in TEN buffer) and the cleaved controls. The PAGE purified cleaved strand (79 bases) was then amplified using a set of nested polymerase chain reactions (nested PCRs). The first amplification selectively amplifies the cleaved region to increase the population of the selected species to detectable levels (by UV spectrometry) and the second amplification prepares the full-length template for the next primer extension. The dsDNA resulting from the first amplification is

exonuclease digested to isolate a single strand (the phosphorylated strand is digested into its monomers) that is then purified by 10% D-PAGE (Section 2.3.1) alongside a molecular weight marker which is a DNA of known molecular weight. PAGE is used as a check to avoid any truncates from propagating to the subsequent steps. The band identified as the product of interest is extracted from the gel using $LiClO_4$ in water and then precipitated in ethanol. The second amplification product was again exonuclease digested to isolate the non-phosphorylated strand and then purified by 10% 8 M mini-D-PAGE. The final PAGE purified DNA (after second amplification) was then used as a template for primer extension

For initial generations, the first amplification product was barely visible by UV shadowing hence these PCR mixtures were doped with radioactivity to track the presence of the oligonucleotide. However, in those after the 4^{th} generation, the first amplification was done without radiolabeling as the amount of oligonucleotides was sufficient to detect by UV shadowing the polyacrylamide gel.

Table 2.1 Progress of Selection showing stringency in terms of time and Mg^{2+} concentration

Generation	Time (min)	Mg^{2+} concentration (mM)	% Cleavage
1	60	2.5	0.0
2	60	2.5	0.0
3	60	2.5	0.1
4	60	2.5	0.9
5	60	2.5	13.9
6	60	2.5	14.2
7	60	0.5	9.0
8	60	0.5	11.5
9	60	0.5	18.0
10	5	0.5	2.1
11	5	0.5	4.2
12	5	0.5	3.0
13	5	0.5	1.4

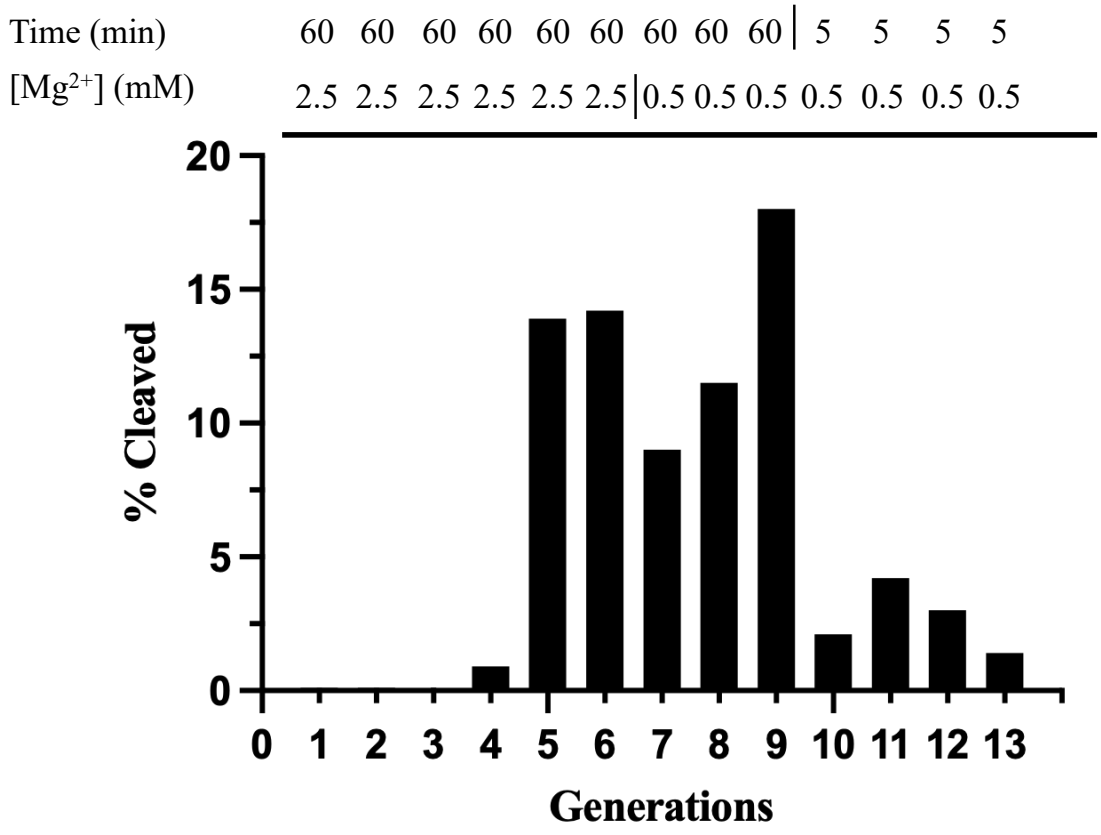


Figure 2.4 Progress of DNAzyme self-cleavage over generation 1 to 13 (G1 – G13).

The Mg^{2+} concentration in the selection buffer was varied across generations (Table 2.1) – starting from a high concentration and then progressively reduced. Initially, the DNAzyme was incubated for 1 hr up to generation 6. The cleavage was nearly unobservable (by dosimetry and autoradiography) until the 4th generation thereafter sharply rising between generations 5 and 6. The self-cleavage was allowed for 1 hr until generation 9 and then reduced to five minutes (to increase stringency) between generations 10-13 as shown in Table 2.1 and as graphically represented in Fig 2.4.

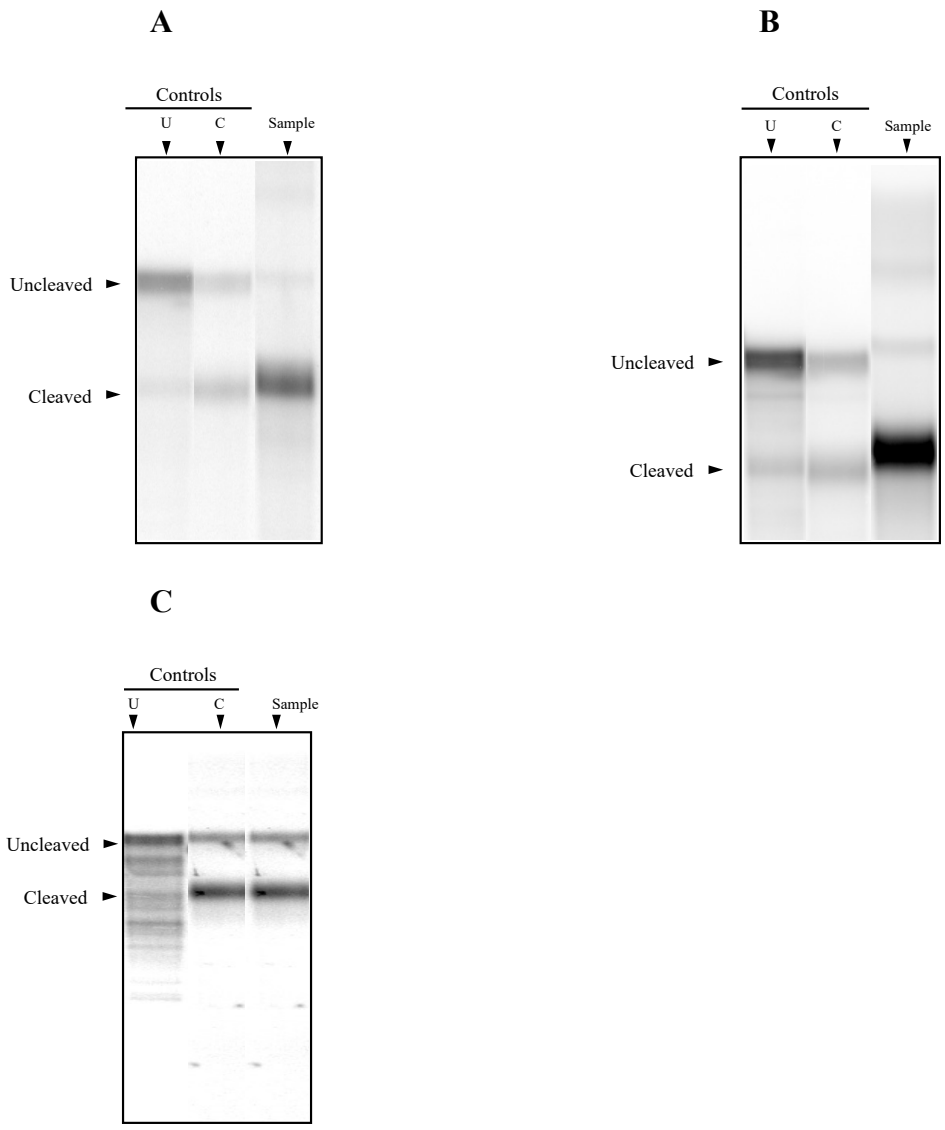


Figure 2.5 Representative gel images showing the results of G6 (A), G7 (B) and G8 (C).
The figure shows the autoradiogram of gels that were run to isolate the cleaved band from any uncleaved material that may have been collected while eluting the DNAzyme (post self-cleavage). All samples were loaded post $LiClO_4$ precipitation and ethanol wash. U stands for the uncleaved control collected in TEN buffer and C stands for the NaOH cleaved control.

The DNAzymes were initially selected under low stringency (Fig 2.4) to allow them to cleave. The stringency was increased (Table 2.1) when the cleavage yield (% cleavage) was above baseline (Fig 2.4) and a relatively strong signal was visible (by autoradiography). This increase in stringency was done to favor fast-cleaving DNAzymes. The enrichment was evident from the

growing intensity of the cleaved band, as seen from the autoradiograms (Fig 2.5). The progress was monitored and a total of 11 generations were done before concluding the selection.

2.4 Cloning the sequences and further characterization

2.4.1 Determining Suitability for Cloning

To determine whether the selected nucleic acids catalysts showed metal independence, generation 9 was repeated but the samples were divided into two treatment groups – one of which was incubated in a selection buffer containing 0.5 mM Mg^{2+} whereas the other of which was allowed to self-cleave in the absence of Mg^{2+}. This was compared to generation 11 that was done in parallel using an identical condition. Each of the treatment groups (from either generation) was assessed using a time-course assay (t = 1, 3, 10, 32, 60, 100, 120 min). A small amount of sample was taken out at each interval and quenched into dye (Formamide/EDTA) which was then resolved by 10% D-PAGE and imaged by autoradiography. This experiment hypothesized that the generation 11 candidates being more superior catalysts (% cleavage yield) would cleave the ribophosphodiester linkage faster than the more inferior generation 9 that was selected under a less stringent regime. The results of the experiment show that the catalysis accelerates to a much greater extent for generation 11 (Fig 2.7). It also attains a higher cleavage maximum (the maximum amount of possible cleavage) as compared to generation 9. This assured a successful selection and determined suitability for library cloning (and identifying active sequences) which is further discussed in the following section.

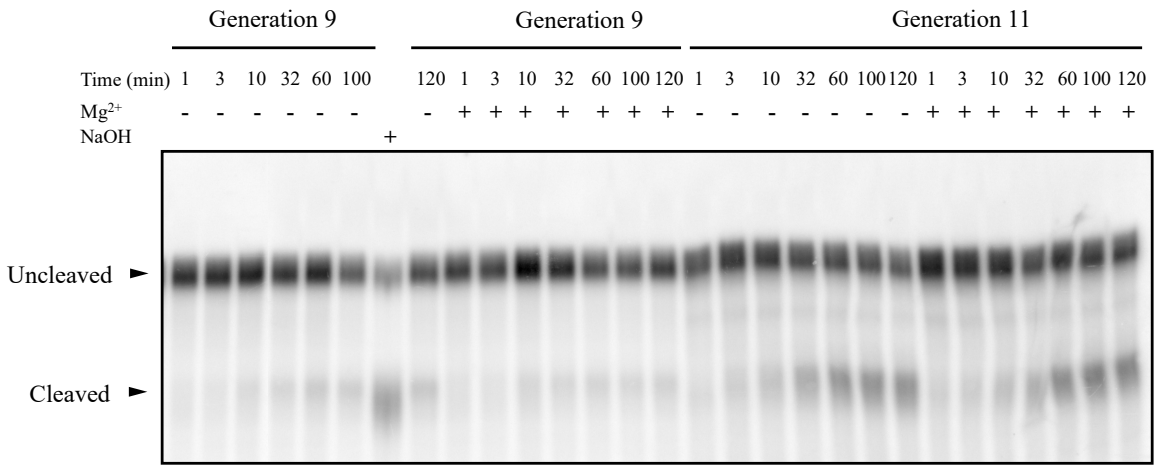


Figure 2.6 Preliminary time course experiment.
This autoradiogram depicts a comparison of DNAzyme activity between generations 9 and 11.

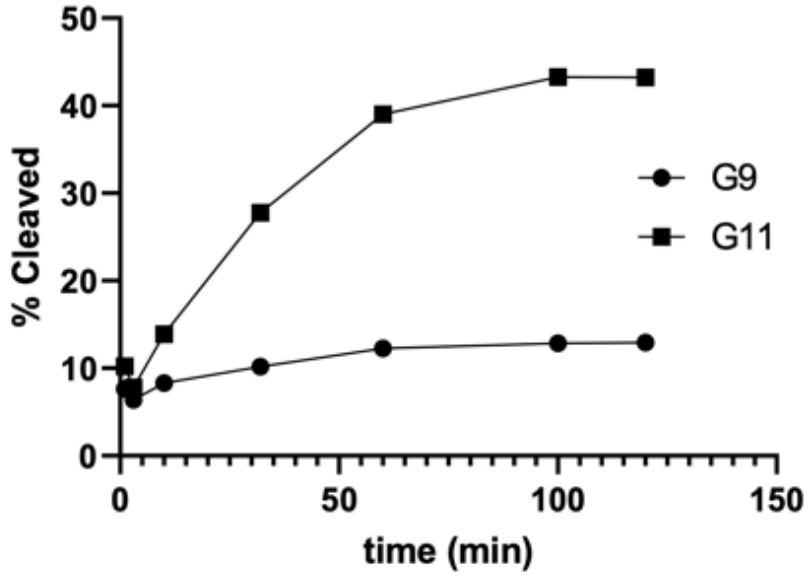


Figure 2.7 Graph depicting the time course profile for G9 and G11.

2.4.2 Next-Generation Sequencing and Structure Prediction

The cleaved sequences from generation 11 were amplified with the sequencing primers – ON5 (Forward Sequencing Primers) and ON6 (Reverse Sequencing Primers) see Table 3.1, Chapter 3. The PCR amplified product was isolated using an agarose gel and purified by an ion-exchange

chromatography column (QIAquick Gel Extraction Kit). The sequencing primers include the corresponding primer binding sites of the library, along with a barcode region to identify either strand and a fragment of the Illumina primer binding region on the 5' ends of each primer. Illumina NGS Sequencing (Amplicon-EZ, Genewiz, LLC) produces more than 200 reads. However, only 30 reads were obtained after cleaning and trimming out the low-quality reads (using tools on Galaxy server v20.09). These 30 reads were analyzed by Multiple Sequence Alignment to group the clones based on similar families. The sequences were grouped based on percentage similarity on a phylogenetic tree (distance between the branches of the tree). A phylogenetic tree clusters sequences together based on the degree of similarity between the composition of nucleotides. For the initial analysis, clones from individual families were selected keeping diversity in mind. Each of these identified clones was then studied using a preliminary time-course assay to determine the clone with the best k_{cat}. Two candidates were identified for further characterization that outperformed the rest in terms of the preliminary kinetics – Dz11-23 and Dz11-33 (Table 2.2).

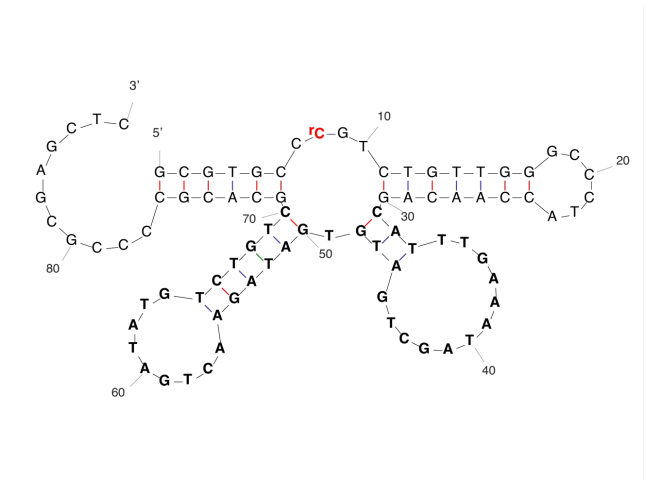
Table 2.2 Best candidates identified for further characterization.

Clone ID	Sequence	%CU
11-23	TTTTTTTTTTTTTTTTTTTTGCGTGCC**rC**GTCTGTTGGGCCCTACCAACAG **CAUUUGAAAUAGCUGAUGUGAUAGAACUGAUAUGUCUGUC**GCACGCCCCGCGAGCTC	45
11-33	TTTTTTTTTTTTTTTTTTTTGCGTGCC**rC**GTCTGTTGGGCCCTACCAACAG **CAUUUGAAAUAGCUGAUGUGAUAGAAACGAUAUGUCUGUC**GCACGCCCCGCGAGCTC	43

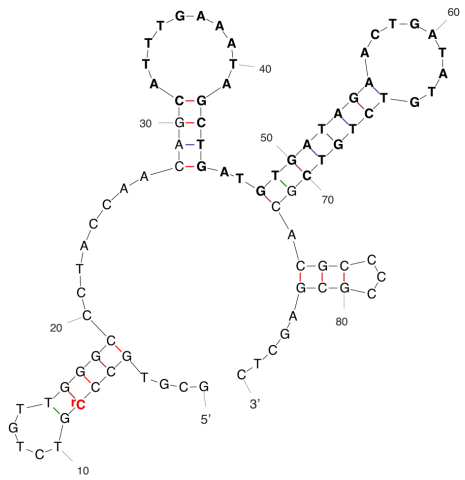
Bold nucleotides are those that have been selected in the degenerate region. U and C represent the modified nucleotides. %CU refers to pyrimidine content in the degenerate region of the library.

The sequences were also studied for preliminary secondary structure analysis (using mFold) wherein we identified two hairpin loops that fold the DNAzyme, positioning the catalytic motif towards the ribonucleotide substrate with the guide arms interspersed on either side (Fig 2.8). The mFold calculations also suggested other possible structures that could be thermodynamically more favorable but most likely these folds were not kinetically active because the substrate site was located very distally from a hypothetical active site (Fig 2.8). Nevertheless, an mFold analysis does not attribute various charge-charge interactions that are possible. Therefore, it cannot completely predict the exact structure of the nucleic acid catalysts particularly with modified nucleosides **dU^{ga} (4)** and **dC^{aa} (3)** (Figure 2.1) presenting positive charges.

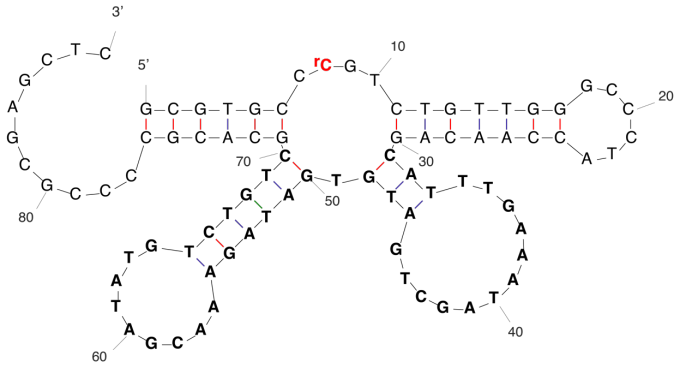
This could, however, serve as a preliminary study towards understanding the secondary structures of the catalysts that would be obtained in future studies. I did not pursue this question here as the positively charged nucleotides could influence various footprinting assays while other analyses (like crystallography, NMR or cryo-EM) could provide a better understanding. Further, secondary structures, using mFold, were determined only for the clones that were identified as the best catalytic molecules (Dz11-23 and Dz11-33) in our preliminary assay. I found two predicted conformations for each candidate (Dz11-23 and Dz11-33), by mFold, corroborating our assertion that there is a fast cleaving conformation that outperforms the misfolded conformation despite it being thermodynamically more favorable (resulting in biphasic kinetics). I consider one of the sequences as a misfolded conformation since the ribophosphodiester linkage in the sequence is part of a double-helix and therefore less likely to be accessible for cleavage.



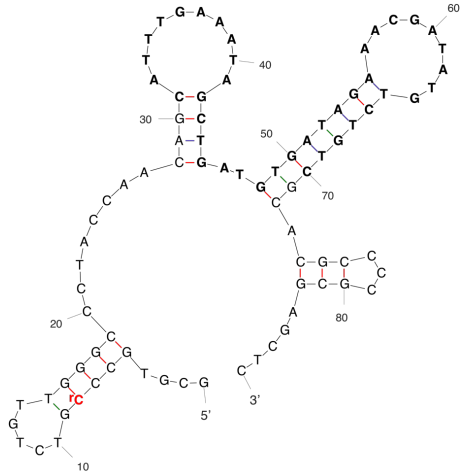
dG = -11.69 kcal/mol (Dz11-23)



dG = -14.27 kcal/mol (Dz11-23)



dG = -11.69 kcal/mol (Dz11-33)



dG = -14.27 kcal/mol (Dz11-33)

Figure 2.8 Potential mFold structural representation of 11-23 and 11-33 DNAzymes.
The thermodynamically stable sequences with dG = -14.27 kcal/mol are depicted as misfolded conformations. (T is used in the mFold instead of dU^{ga} 2). rC represents the cleavable ribonucleotide and the *bold* nucleotides are those that have been selected in the degenerate region.

2.4.3 Preliminary Time Course Kinetics Assay

From the 30 identified sequencing reads, 12 clones were ordered from IDT – Dz11-4, Dz11-5, Dz11-6, Dz11-7, Dz11-10, Dz11-12, Dz11-15, Dz11-23, Dz11-30, Dz11-33, Dz11-39, Dz11-40 –

as their corresponding complementary strands. For an initial time-course assessment, the DNAzymes were prepared by primer extension reaction, as described in Chapter 3 (Section 3.8.2).

The DNAzymes were incubated in a selection buffer for 60 minutes with ionic concentration resembling that used in generation 11. Time-dependent self-cleavage was observed by autoradiography only for Dz11-10, Dz11-15, Dz11-23 and Dz11-33.

2.4.4 Kinetics (cis) of intramolecular cleavage

From the four initial candidates, the most active sequences were identified as Dz11-23 and Dz11-33 after a full time course assessment (Dz11-10 and Dz11-15 data in Appendix). The catalysts Dz11-23 and Dz11-33 were characterized more comprehensively under pseudo-physiological ionic conditions (25 mM cacodylate buffer pH 7.5, 0.5 mM $MgCl_2$, and 200 mM NaCl, 25 ± 1 ºC) and visualized by autoradiography (Fig 2.10). The enzymes were characterized at time points 0.5, 0.67, 1, 2, 3, 5, 10, 15, 30, 60, 90, and 1080 minutes. Samples were taken out at individual time points and quenched in a stop solution containing Formamide/EDTA/Biotin (refer to Chapter 3) before running them on an 8% D-PAGE which was imaged by autoradiography. The samples primarily contain two bands a cleaved product that increases in intensity over the time points, and an uncleaved starting material that subsequently decreases in band intensity. The assay, however, shows a maximum cleavage of 60-70% which was slightly lower than some of the imidazole modified DNAzymes (maximum cleavage about 80-85%) that have previously been selected in our lab. As no further cleavage (t > 1200 min) was observed, this was taken as the maximum amount of observable cleavage.

The two best candidates, Dz11-23 and Dz11-33 were subjected to further characterization. The data was initially fit on a standard first-order rate equation (equation 1).

$$P_t = P_E(1 - e^{-k_{cat}t}) \qquad (Equation\ 1)$$

P_t represents the fraction cleaved at time t, whereas P_E denotes the percentage cleaved at the assay's endpoint. The rate constant of the reaction is given as k_{cat}.

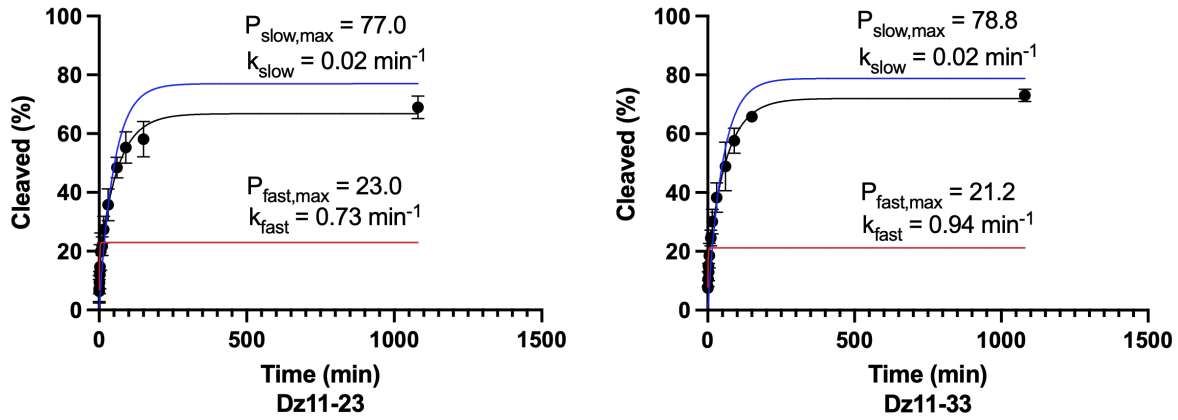


Figure 2.9 Plots for catalytic *cis*-cleavage of Dz11-23 and Dz11-33.
The above plots representing the curve fit for time-dependent self-cleavage was obtained from the band intensity data in Fig 2.10 and then computed using the double-exponential equation (Equation 2) while the red curve (fast phase) and the blue curve (slow phase) were computed using the single-exponential equation (Equation 1) with the amplitude and rate data obtained from the double-exponential equation. The rate constants k_{fast} and k_{slow} represent the constants for their respective phases (fast and slow). The left and the right panels are the plots for catalytic self-cleavage of Dz11-23 (R^2 = 0.95, n = 3 with two sets of separate experimental replicates) and Dz11-33 (R^2 = 0.97, n = 3 with two sets of separate experimental replicates) respectively.

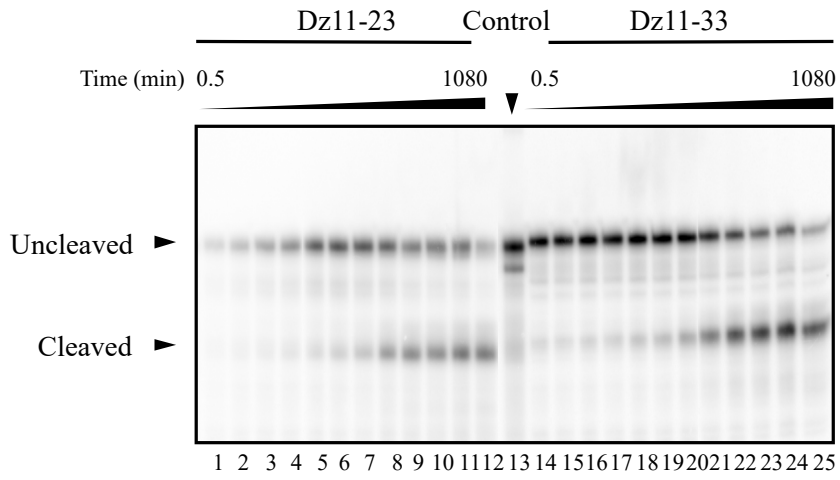


Figure 2.10 Autoradiography of time-dependent self-cleavage of Dz11-23 and Dz11-33.
Time course of self-cleavage of Dz11-23 and Dz11-33 visualized by autoradiography from 0.5 to 1080 min. The time points are – 0.5, 0.67, 1, 2, 3, 5, 10, 15, 30, 60, 90, 1080 minutes performed at 25 ± 1 ºC (n = 5).

The data from this experiment was used to determine the rate constant. As previously mentioned, Equation 1 was initially used to determine the first-order kinetics. Although a simple monophasic self-cleavage rate constant of approximately 0.03 min^{-1} is readily calculated, the self-cleavage data was distinctly biphasic and curve fitting to a biphasic equation (Equation 2) provided a more coherent fit (Fig 2.10).

$$P_t = P_{fast,max}\left(1 - e^{-k_{fast}t}\right) + P_{slow,max}\left(1 - e^{-k_{slow}t}\right) \qquad (Equation\ 2)$$

For the biphasic model, the maximum amplitude of the slow phase was calculated to be approximately 77% with the apparent rate constant ($k_{slow,app}$) of 0.02 min^{-1} whereas that of the fast phase was about 23% with an apparent rate constant ($k_{fast,app}$) of 0.73 min^{-1} in case of Dz11-23. For Dz11-33, the biphasic kinetics resulted in a maximum amplitude of the slow phase at 79% with a similar rate constant ($k_{slow,app}$) of 0.02 min^{-1} and the maximum amplitude of the fast phase was 21% with a slightly higher apparent rate constant ($k_{fast,app}$) of 0.94 min^{-1}. Though the Dz11-33 had

slightly fewer modified nucleosides, the rate constant for the slow phase was much lower than those previously selected by the Perrin group using the modified nucleosides **dA^{his} (1)**, **dC^{aa} (3) and dU^{ga} (4)** (Figure 2.1) however the rate constant of the fast phase was comparable. Biphasic kinetics for self-cleavage have been previously observed in DNAzymes containing all three modifications[109,145]: **dA^{his} (1)**, **dC^{aa} (3) and dU^{ga} (4)** (Figure 2.1), and even in DNAzymes containing no xenobiotic functionality[146,147].

To further address the lower-than-normal cleavage yield, a refolding analysis was performed using Dz11-23 by heat-denaturing the uncleaved DNAzyme. The heated tube was placed over ice (snap cooled) and then left at room temperature for an additional 24 hours (this helps denature and refold the misfolded conformation). This procedure increased the cleavage yield to approximately 80% (Fig 2.11), which suggests that some of the DNAzymes were likely folded in an inactive structure (see discussion).

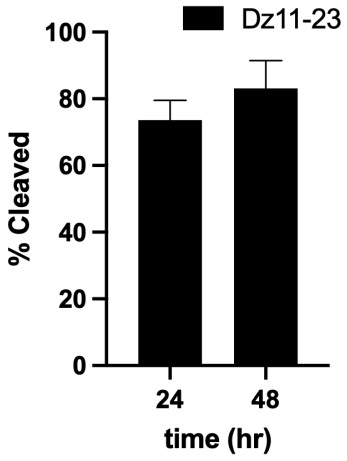


Figure 2.11 Refolding Analysis at 25 ºC to elucidate additional conformations (Dz11-23).
Dz11-23 was sampled at the 24 hr time point and then subject to thermal denaturation at 95 ºC followed by snap cooling on ice. It was then refolded for a further 24 hours, which is depicted as the 48 hr time point. The analysis was replicated twice, presented as an average ±standard deviation.

2.4.5 Essentiality of the modified nucleotide

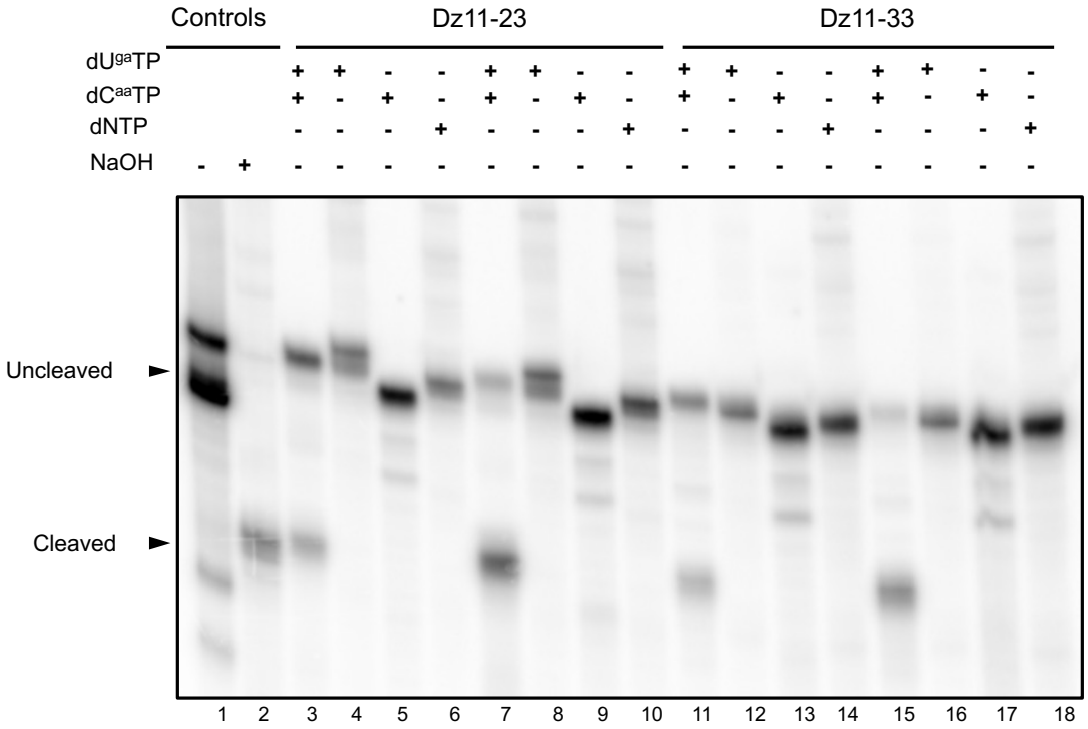


Figure 2.12 Autoradiography of an 8% D-PAGE depicting the importance of the modified nucleotides. Experiments illustrate dU^{ga}TP (4, guanidium), dC^{aa}TP (3, amino-allyl), and dNTP canonical in each lane done in duplicate for both Dz11-23 and Dz11-33. The rest of the nucleotides are canonical bases. The reactions were incubated for 30 minutes.

To demonstrate that the modified nucleotides are crucial to the activity of the catalysts, both the clones were enzymatically resynthesized while varying the composition of each modified nucleotide in the cocktail. Self-cleaving DNAzymes that had either one or both modified nucleotides removed and replaced with an unmodified congener showed a significant loss in activity over 1080 minutes (less than 10% cleavage, Figure 2.12). When the **dU^{ga}TP (4)** was replaced with the commercially available precursor **dU^{aa}TP (2)** (Figure 2.1) it resulted in several truncation artifacts (see Appendix) though the full-length modified DNAzyme seemed to show no observable activity (Figure 2.12).

2.4.6 Effect of temperature and pH on the enzyme catalysis

Further, the DNAzymes were characterized for their activity over different temperatures and pH profiles. While self-cleavage activity was very limited at 5 ˚C and 55 ˚C, I also noted that the extent of fast and slow phases varied with temperature in addition to the rate constants for self-cleavage. However, it should be considered that due to two separate phases existing, the rate constants extracted from each individual phase might not be perfectly reproducible as the clones can vary in terms of their relative fraction with temperature and pH.

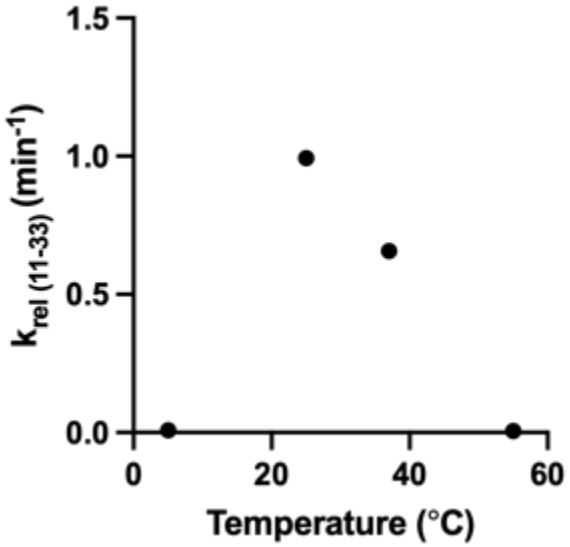


Figure 2.13 Temperature profile for Dz11-33 between 5 ºC and 55 ºC (n = 3).

A preliminary analysis of the rate constants was studied using the fast phase of the biphasic model. The rate constants as observed of the slower phase were distinctly separate from the faster phase (Fig 2.14). At 5 ºC and 55 ºC, however, the rate constants of the faster and the slower phase were nearly identical with almost no activity observed in each of those temperatures. As the selection had been performed at 25 ºC with an error margin of 2 ºC, the temperature optimum of catalysis

was determined to be around 25 ºC with a good amount of cleavage yield still observable at 37 ºC (Fig 2.13) which is comparable to the previously selected DNAzymes containing imidazole-linked-dATP **(1)**.

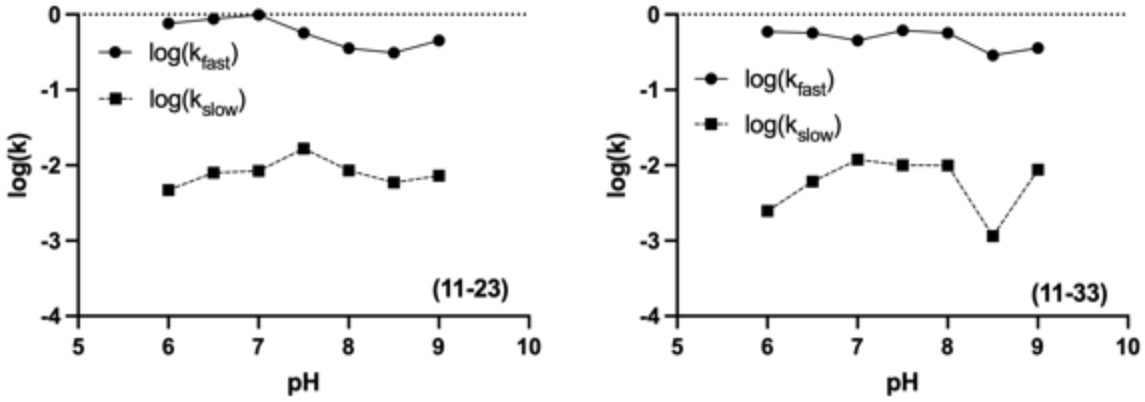


Figure 2.14 Representative pH-rate profiles (Dz11-23 and Dz11-33).
The experiments were investigated in various pH buffers at 25 ± 2 ºC. Errors were computed using standard deviation between the dataset collected from two independent experiments, each with n = 3, pooled. Buffers used in the experiment were sodium phosphate buffer (25 mM) adjusted to pH 6, 6.5, 7, 8, sodium cacodylate buffer at pH 7.5, and Tris (25 mM) at pH 8.5 and 9 in 0.5 mM Mg^{2+}.

Unexpectedly, the pH rate profile did not show a distinctive bell-shaped curve (Fig 2.14) that is otherwise characteristic of many ribozymes, and for both M^{2+}-free and M^{2+}-dependent DNAzymes that have been selected in the past[106,142,148,149]. Though most ribozymes that have been identified seem to distinctly exhibit a bell-shaped pH rate profile, there have been instances such as with variants of hairpin ribozymes that function with only a shallow dependence on pH[150,151]. The absence of a bell-shaped curve concomitant with the absence of the imidazole group may reflect rate-limiting conformational changes or a possible inverse relationship between Brønsted base catalysis favored at higher pH values. Electrostatic or H-bond stabilization of a pentacoordinate phosphorane intermediate may also be disfavored at higher pH in cases where an ammonium cation loses its proton. I admit that the biphasic kinetics of the enzyme may significantly

complicate the ability to address the true temperature and pH-rate profile as the relative amplitude of each phase can be influenced by both temperature and pH. Yet, further studies are needed to determine the possible mechanism by which the DNAzyme cleavage might be occurring in this case.

2.4.7 Effect of monovalent K^+ ion on the cleavage rate

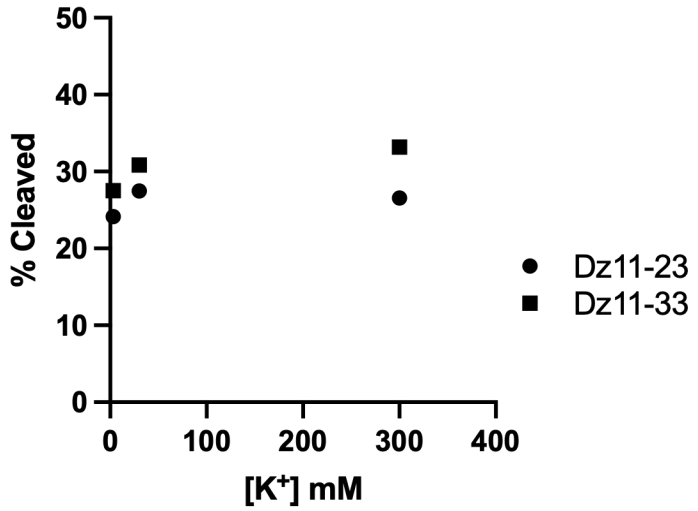


Figure 2.15 Qualitative plot of the effect on [K^+] on the activity (Dz11-23 and Dz11-33).
Percentage cleavage was found to be only marginally affected across a wide range of K^+ concentrations (3-300 mM) for either of the DNAzymes. The DNAzymes were incubated for 30 minutes in sodium cacodylate buffer (25 mM) at pH 7.5 containing three different levels of K^+ - 3 mM, 30 mM, and 300 mM. There is no significant trend observed in the percentage cleavage across the different concentrations of the monovalent ion.

I further evaluated the cleavage rate of the DNAzyme as a function of potassium ion concentration. G-quartet sequences are particularly formed with the help of monovalent ions like Na^+, K^+ and Rb^+. They are observable in highly structured DNA molecules consisting of a stretch of 3-4 guanosine residues in a square planar structure. The size of the central cavity determines which ion can bind tightly in the centre[152]. Potassium ions are an important co-factor that helps in stabilizing such structures though Na^+ and Rb^+ could have some effect[153,154]. Despite an obvious

lack of any G quartet forming sequences, I tested the effect of K^+-concentration, on the rate, over a long-range (3-300 mM). I wanted to test if K^+ ions themselves had any effect on the cleavage, at the ribophosphodiester linkage, on a qualitative scale, at three different levels (as previously mentioned). There was only a little difference in activity in the tested range (Fig 2.15).

2.4.8 Effect of magnesium ions on the activity of Dz11-23 and Dz11-33

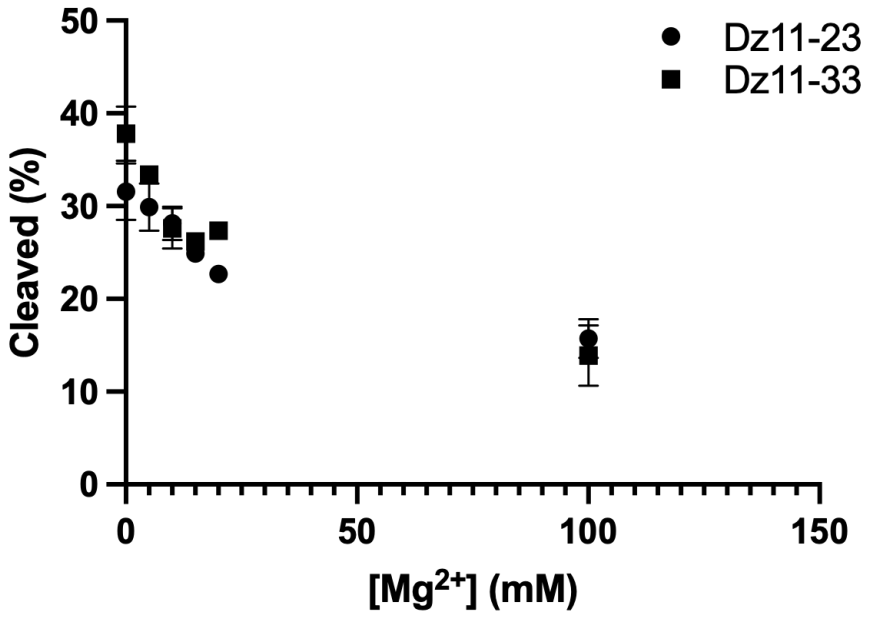


Figure 2.16 Percentage cleavage plotted as a function of Mg^{2+} concentration (0 – 100 mM).
The DNAzymes were incubated for 30 minutes in different pH 7.5 sodium cacodylate (25 mM) buffers containing varying concentrations of Mg^{2+} - 0 mM, 0.5 mM, 5 mM, 10 mM, 15 mM and 100 mM (n = 3).

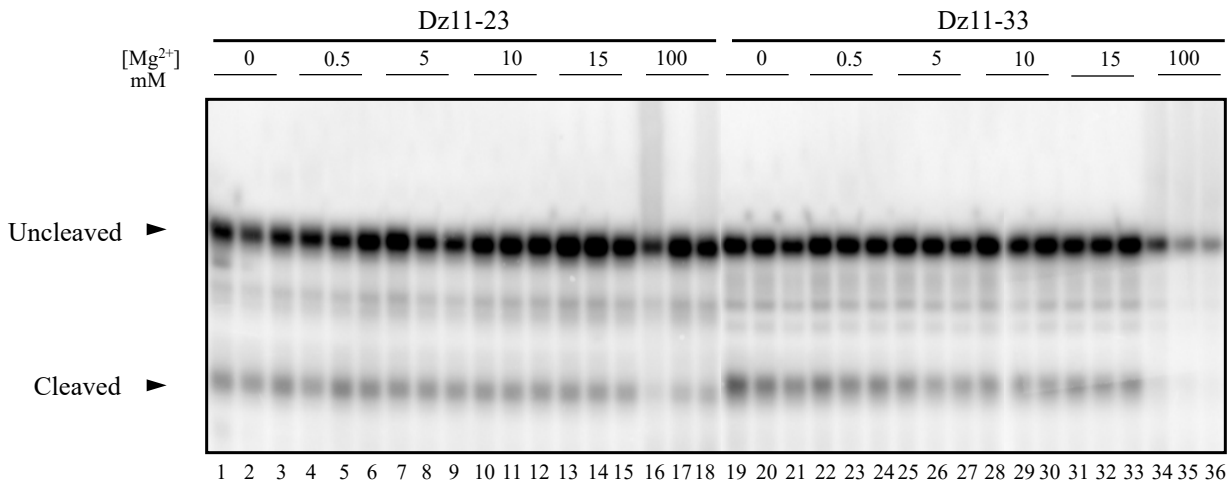


Figure 2.17 Autoradiogram of the magnesium cleavage profile.
Concentration of Mg^{2+} between 0 – 100 mM for the catalysts Dz11-23 and Dz11-33 (n = 3)

I further assessed the role of Mg^{2+} concentration on the rate of self-cleavage. The primary hypothesis of my work is that modified DNAzymes in the absence of an imidazole functionality (**dA^{his} (1)**, Fig 2.1) would have a self-cleavage rate similar to the imidazole containing DNAzymes using **dC^{aa} (3)** and **dU^{ga} (4)** (Fig 2.1), particularly if one or more Mg^{2+} cations could be recruited to support catalysis. Initially, the DNAzymes were selected at 2 mM Mg^{2+} but the stringency was increased by dropping Mg^{2+} concentration to 0.5 mM (Table 2.1). Interestingly, both of the identified clones, Dz11-23 and Dz11-33 were observed to have the highest activity under the Mg^{2+}-free condition (Figure 2.16) as is also evidenced by autoradiography (Fig 2.17). This is in accordance with my hypothesis that the DNAzymes selected without the imidazole functionality could still lead to promising catalysts with self-cleavage rates that are significantly higher than unmodified M^{2+}-free DNAzymes.

2.4.9 Effect of transition metal ions on the activity of Dz11-33

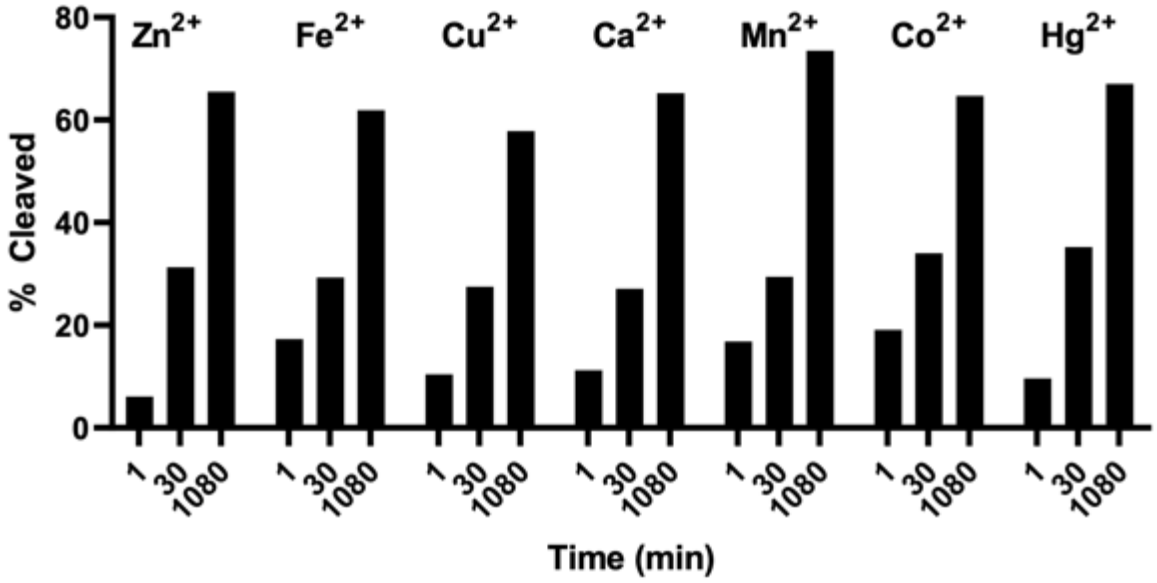


Figure 2.18 Percentage cleavage in the presence of transition metal M^{2+} cations.
The samples were collected over time points - 1, 30, 1080 min.

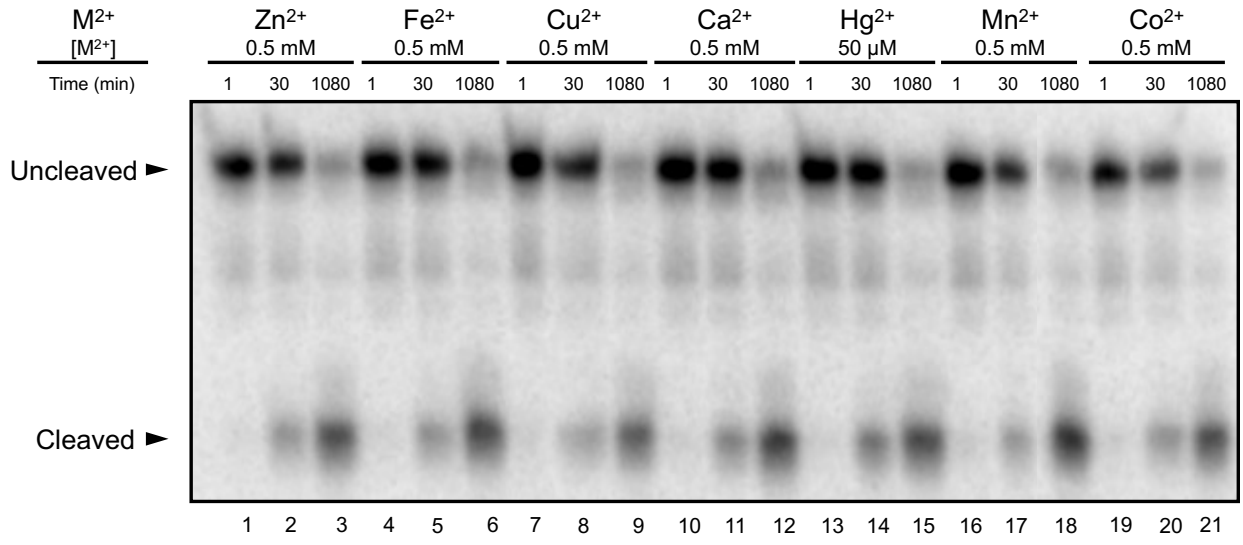


Figure 2.19 Representative autoradiogram of the effect of various transition metals
The effect of various transition metals (500 µM Zn^{2+}, Fe^{2+}, Cu^{2+}, Ca^{2+}, Mn^{2+}, Co^{2+}, 50 µM Hg^{2+}) on the cleavage rate of Dz11-33.

I next examined the cleavage activity of Dz11-33 using different divalent transition metal cations to confirm that the enzyme was metal independent each at 0.5 mM with the exception of Hg^{2+} which was set at 50 μM (Figure 2.18 and Figure 2.19). Hollenstein et. al. (2008) published an allosteric DNAzyme (Dz10-13) that could self-cleave with an apparent dissociation constant ($K_{D,app}$) of 110 nM in the presence of Hg^{2+} (10 μM) showing significant cleavage[108]. While aminophilic metal ions like Zn^{2+} and Cu^{2+} may have been expected to be inhibitory, none of the metal ions seemed to have any significant effect on the self-cleavage of our candidates (Dz11-23 and Dz11-33). This alludes to the fact that the Mg^{2+} independence nature of these catalysts may extend to other transition metal cations as well.

2.4.10 Target specificity of DNAzyme Dz11-33

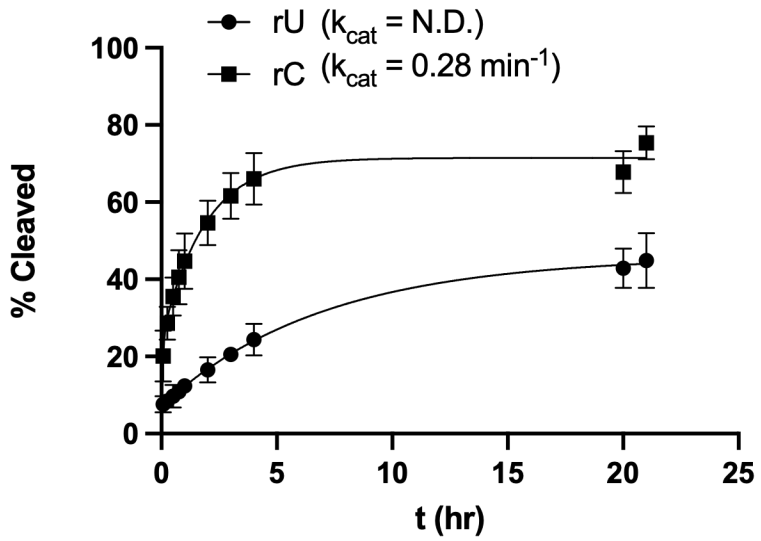


Figure 2.20 Graphical representation of the cleavage kinetics for substrates rC (R^2 = 0.91).
The substrate rU shows a weak time-dependent cleavage trend by the naked eye however to statistical analysis software (Fiji and Prism), the regression of first-order kinetic plots did not converge, with the cleavage yield maximizing at less than 40%. All other ribonucleotide substrates showed unascertainable cleavage.

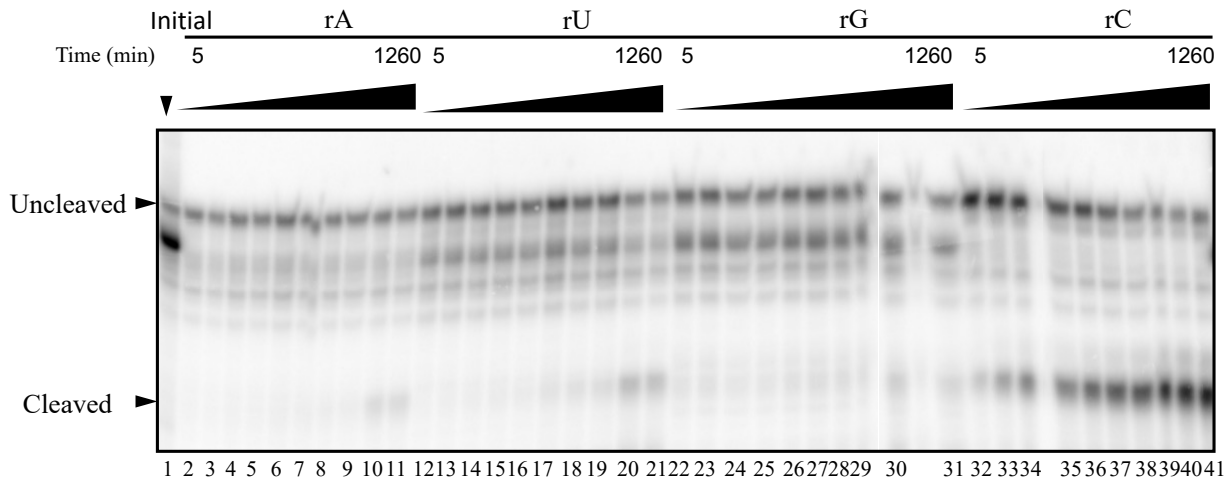


Figure 2.21 Representative autoradiogram obtained for the substrate specificity assay.
Shows substrate specificity of different ribonucleotides (rA, rU, rC, rG) cleavage conditions at 25 °C (t = 5 – 1260 min) in buffer pH 7.5 at 0.5 mM Mg^{2+}. A lower running band was observed in samples prepared with rU and rG, which is attributed to secondary structure in the hairpin loop connecting the guide arms to the target. Lanes 30-35 were edited from the original autoradiogram of the gel to correct for spacing (complete autoradiogram in Appendix).

Finally, I assessed the specificity of Dz11-33 for recognizing a labile ribonucleoside. This serves as a preliminary assessment for future experiments on the ability to engineer this species into a *trans*-acting DNAzyme for a specific RNA target. For this, 4 separate primers were ordered each containing either an rA, rU, rG or rC; each of them were used to direct an enzymatic resynthesis of the Dz11-33 DNAzyme using primer extension. The cleavage experiment was observed in an intra-molecular environment for this work. Qualitatively it was assessed that the self-cleavage which was selected towards a ribocytidine (rC) substrate was more favorable to that substrate comparable to other ribonucleotide counterparts; minimal self-cleavage was observed for the target rA (lanes 2-11, Fig 2.21) and rU (lanes 12-21, Fig 2.21), and no time-dependent cleavage observed for rG. The fast phase of the observed apparent rate constant towards the rC substrate was 0.28 min^{-1} with maximum amplitude at 24.1%. There was no biphasic kinetics decipherable on using Equation 2 (in Prism) for the ribouridine substrate though it showed a minimally observable

cleavage. This could signify a certain preference for pyrimidine nucleotide targets though with a significant specificity towards ribocytidine.

2.5 Discussion

Modified dNTPs mostly, but not always enhance the rate of RNA cleaving DNAzymes. These are predominant in RNA-cleaving DNAzymes with multiple turnovers as they can compensate for the use of high divalent metal cation concentrations which exceed typical intracellular ionic concentrations (Section 2.1). The imidazole functionality (**dA^{his} (1)**, Fig 2.1) was previously identified as a crucial functionality in previous selections of M^{2+}-independent RNA-cleaving DNAzymes, as RNase A mimics (histidine is an important amino acid in the active site of RNase A), through its acid-base catalytic mechanism. The importance of an imidazole has been highlighted earlier in the thesis in the context of post-selection modification and re-engineering, and also as a co-factor in RNA cleavage. In our lab, the use of the modified nucleotide **dA^{his}TP (1)** has resulted in not just DNAzymes with high k_{cat} values but also in multiple turnover DNAzymes and an allosteric DNAzyme toward Hg^{2+} [108,141].

However, **dA^{his}TP (1)** has proven to be a substrate that is difficult to incorporate by many polymerases and the resulting modified DNA is poorly amplified. Nevertheless, selected strands with multiple dA^{his}'s engendered a substantial gain of function, and limitations caused by poor sequence space coverage were evidently compensated by an increase in "chemical space" contributed by the imidazole group. Nevertheless, in one case the imidazole had not performed as expected; the self-cleaving activity of the DNAzyme was highly dependent on an imidazole functionality (8-imidazolyl-methyleneamino deoxyadenosine, **6**) (along with **dC^{aa}TP** (**3**) and

dU^{ga}TP (4)) but its catalytic rate was more characteristic of unmodified M^{2+}-free DNAzymes ($k < 10^{-3}$ min^{-1}).

This was attributed to the observation that **6** is a much poor substrate of many polymerases as compared to **dA^{his}TP (1)**. In another case, a DNAzyme which had been selected in the presence of a modified nucleotide, optimized for an excellent sequence space coverage, yet was not entirely dependent on the incorporation of the modification[155]. Imidazoles may have a superior role to play in DNAzyme self-cleavage, however not all of these DNAzymes exhibited RNaseA-like activity without a second modified nucleotide that presented a cationic amine[156,157]. Non-metal cations such as ammonium and guanidium groups have a possible role in the Coulombic stabilization of the pentacoordinate phosphorane intermediate[158–163]. In addition, covalently attached cations, such as guanidium groups, can demonstrate significant thermostability which can provide extensive support towards target recognition and secondary structure formation that would otherwise be impaired at a low ionic concentration in general[164]. Hence, I hypothesized that the cation-modified nucleotides, without the imidazole, could lead to the discovery of efficient M^{2+}-free DNAzymes.

As a result, this thesis focuses on developing DNAzymes using modified deoxynucleoside triphosphates containing cationic amino acid side chain functional groups (**dC^{aa}TP** (**3**) and **dU^{ga}TP** (**4**)) which were previously used in DNAzymes along with the **dA^{his}TP (1)**. We initially observed a cleavage rate of about 0.03 min^{-1} with the two modified nucleosides. However, as described in Section 2.4.4 the DNAzymes were identified to be biphasic; also from mFold predictions (Section 2.4.2), I identified multiple conformational forms. When the catalyst was heat-denatured to dissociate the thermodynamically favorable slow cleaving conformations and

then refolded they reached a slightly higher cleavage threshold about 80%, from the initial 70% maximum cleavage. Further, relative self-cleavage rates showed almost no cleavage at 5 ºC and 55 ºC with a maximum observable cleavage at 25 ºC. There was also no significant effect observed under increased monovalent ions, notably K^+ ions in the range of 3-300 mM. Based on my hypothesis I also found the Dz11-33 was the most catalytically active in the absence of any M^{2+} ions. When tested over an increasing Mg^{2+} concentration the cleavage activity decreased to 10% in 30 minutes for 100 mM Mg^{2+}. This was contrary to other catalysts which require high Mg^{2+} concentration to function[92,129–131,133,134,165]. In terms of sequence specificity, there was 5-fold higher target specificity towards its target rC compared to its base-paired partner rG; with only minuscule activity towards rA and no discernible activity to rU.

Ahead of the selections, I assumed that in lieu of the imidazole, Mg^{2+} would be essential to i) afford the activation of the 2'-OH for an attack at the phosphorus, and ii) potentially promote Lewis-acid catalysis to collapse the phosphorane intermediate. Though Mg^{2+} was present throughout the selection, the most active sequences that were selected were found to function in the absence of any Mg^{2+} ion suggesting M^{2+}-independence. Assuming monophasic cleavage, a rate constant of 0.03 min^{-1} could be estimated. Disregarding that complex biphasic kinetics was observed, a rate constant of 0.03 min^{-1} is within an order of magnitude of other hyper-modified M^{2+}-independent DNAzymes that have been selected using **dA^{his}TP (1)**, and on par with those selected using **dA^{his}TP (1)** and **dU^{aa} (2)**.

The selected sequences appeared to be biphasic predominated by a slow phase with an apparent rate constant of 0.02 min^{-1} and a fast phase that was between 0.7-0.9 min^{-1}. Such kinetic

mechanisms usually have two different folds where one is slow-to-cleave due to a thermodynamically stable inactive conformer in comparison to a smaller population of the active conformation. This potentially inactive structure (as explained by the alternative mFold) may also be a plausible explanation of lower than normal cleavage yields.

Interestingly, the cleavage rate appears to be independent of pH. The absence of a pH-rate profile that is bell-shaped (which is suggestive of acid-base catalysis) is consistent with the lack of an imidazole group which is normally known to show such a mode of catalysis in the pH range (6-9). The presence of cationic amines and guanidines, which are otherwise not suitable as bases in the pH range due to their normally higher pKa values (cationic amines at 9 and guanidine at 13), could be expected to exhibit pKa perturbation of the unmodified guanine or uridine nucleobases to support acid-base catalysis; also possible are alternative rate-limiting conformations or some form of Brønsted-base catalytic mechanism influenced by a guanidium group. The guanidinium cation, which has a pKa around 13 remains mostly protonated in the pH range 6-9 (as a result arginine side chains in proteins are always predominantly and even remain so at physiological conditions of near neutral pH)[166]. The primary amines (pKa = 9.5), however, can have pH values that are perturbed[167,168] as a result of charge-charge repulsion from the neighboring positively charged guanidium groups. These repulsive forces can drive the amine groups into pockets of hydrophobicity that could keep such amine groups deprotonated which may have roles to play in the catalytic nucleophilicity at the ribophosphodiester.

The modified nucleotides were also shown to be essential for self-cleavage. In contrast to Sidorov's DNAzyme that demonstrated notable cleavage even when the imidazole modified

nucleotide **(6)** was resynthesized with an unmodified dC, our DNAzyme showed a significantly reduced activity when one of the modified nucleotides was removed; the modified nucleotides were thus indelible to our catalysts[155]. When one of the two modified nucleotides or both were replaced with an unmodified congener to prepare DNAzymes Dz11-23 and Dz11-33, there was no activity. Interestingly, when the guanidium group was replaced with the modified nucleotide precursor **dU^{aa}TP (2)** that presents a cationic amine, no activity was observed but also a lot of truncation appeared. It is quite possible, as I earlier stated, that the guanidium groups were essential in maintaining the structure. Thus, we can confidently assume that the modified nucleotide side chains are vital in either maintaining the structure, catalysis or both.

This work was undertaken to evaluate the efficacy of selecting RNA cleaving DNAzymes using two different nucleosides each modified with a cationic amine in the absence of imidazole. Surprisingly, M^{2+}-independent cleavage was selected despite the addition of Mg^{2+} at the outset of the selection. While rate constants were modest and could be improved, it is surprising that the selection did not make use of available Mg^{2+} and instead made use of motifs that were entirely self-sufficient in the absence of Mg^{2+}.

Chapter 3: Materials and Methods

3.1 Materials

Canonical nucleotides (dATP, dGTP, dCTP, and dUTP) were purchased from Invitrogen™, a Thermofisher Scientific entity. The modified nucleotides $dU^{aa}TP$ (precursor for $dU^{ga}TP$) and $dC^{aa}TP$ were obtained from TriLink BioTechnologies. The nucleotide $dU^{ga}TP$ was synthesized in-house after reaction solvent optimization for increased yield (work included in the associated publication and carried out by Antonio A.W.L Wong) as compared to a previously established procedure. All oligonucleotides were synthesized by Integrated DNA Technologies which were purified on a 10-20% PAGE in-house. Vent polymerase, Klenow (exo-) polymerase, yeast inorganic pyrophosphatase (YIPP) and lambda exonuclease were all purchased from New England Biolabs (NEB). Sequenase v2.0 polymerase used in the initial generations was purchased from Applied Biosystems™ also a Thermofisher Scientific entity. Streptavidin Magnetic Beads were sourced from the NEB Biofreezer at UBC. The radioactive nucleotide dGTP α-[^{32}P] was sourced from Perkin Elmer. For some of the characterizations and primer extension reactions, dATP α-[^{32}P] was used instead of dGTP α-[^{32}P] which was also purchased from Perkin Elmer. N,N'-di-Boc-N''-trifluoromethanesulfonylguanidine was purchased from Chem-Impex International. Precoated glass-backed plates containing Silica Gel 60 F254 used in the purification of $d^{ga}UTP$ was purchased from EMD Chemicals. HPLC-grade acetonitrile was purchased from Sigma Aldrich. The 40% Acrylamide/Bis Solution (29:1) and 10x Tris/Boric Acid/EDTA buffer (TBE) were purchased from Bio-Rad. The agarose for visualizing the PCR amplified products was Agarose B, Low EEO which was obtained from Bio Basic. The Quick-Load Purple Low Molecular Weight DNA Ladder was obtained from New England Biolabs. QIAEX II Gel Extraction Kit and

Quant-iT PicoGreen, for NGS sample preparation, were purchased from Qiagen and Invitrogen respectively. All buffer and metal salts were purchased from Sigma Aldrich.

Oligonucleotide sequences

Table 3.1 Table of Oligonucleotide sequences used

Oligonucleotides	Sequences
Library (ON5)	GAGCTCGCGGGGCGTGC(N40)CTGTTGGTAGGGCCCAACAGACG
ON1	Biotin-TTTTTTTTTTTTTTTTTTTTGCGTGCC**rC**GTCTGTTGGGCCC
ON2	Phosphate-CGTCTGTTGGGCCCTACCA
ON3	GAGCTCGCGGGGCGTGC
ON4	Phosphate-ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA
ON6	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGCTCGCGGGGCGTGC
ON7	GACTGGAGTTCAGACGTGTGCTCTTCCGATCTACGACACAGAGCGTGCCCGT CTGTTGGGCCCTACCA
ON8	Biotin-TTTTTTTTTTTTTTTTTTTTGCGTGCC**rG**GTCTGTTGGGCCC
ON9	Biotin-TTTTTTTTTTTTTTTTTTTTGCGTGCC**rU**GTCTGTTGGGCCC
ON10	Biotin-TTTTTTTTTTTTTTTTTTTTGCGTGCC**rA**GTCTGTTGGGCCC

All sequences are in the 5' → 3' direction. The rC, rG, rU, and rA in bold typeface represents the internal scissile ribonucleotide.

3.2 Preparation of buffers and solutions

3.2.1 Common buffers and solutions

DEPC treated water (nuclease free water)

For every 1 L of double-distilled water (ddH_2O), 0.902 mL of diethylpyrocarbonate (DEPC) is added. The solution is then incubated in a 37 °C incubator overnight to destroy any contaminating nucleases and then autoclaved. If preferred can be further filter-sterilized using a 0.22 μm filter. All buffers are prepared using the DEPC treated water.

Gel Eluting Solution

1% $LiClO_4$ is dissolved in a 10 mM Tris-HCl (pH 8.0) in sterile nuclease-free water. The solution is then filter sterilized (0.22 μm) to remove any bioburden that may have been introduced by the Lithium Perchlorate salt.

Gel Loading Solution

A 30 mL solution of the Gel Loading Dye contains 90% v/v of Formamide with 10% $EDTA_{aq}$ (0.5 M) and trace amounts of Bromophenol Blue (0.05%).

Ammonium Persulfate Solution

10 mg of Ammonium Persulfate is weighed and added to the 10 mL of nuclease-free water. Solutions of this must be prepared fresh for perfect gel casts. If it must be stored, it is aliquoted and preserved at 4 °C (for the short term) or -20 °C (for the long term). Solutions of APS in water is not very stable at room temperature even if it is protected from light or air (MSDS Sigma Aldrich).

Urea Diluent (for Polyacrylamide Gels)

To prepare a 7 M urea diluent stock, urea is dissolved at a final concentration of 60.9% w/v with 1X TBE at 14.5% v/v. Urea is used as the denaturant for the polyacrylamide gels while TBE is an

electrolyte. If a desired volume of 1L is prepared, the total solvent volume used must be 690 mL, which is 100 mL of 10 X TBE and 590 mL of water. Must be stored at 2-8 °C.

20% Acrylamide/Bis-acrylamide (29:1) (monomer mixture for Polyacrylamide Gels)

To prepare a 1L solution of Acrylamide/Bis solution (20% dilution), urea is dissolved to a final concentration of 60.9% w/v with 1X TBE 14.5% v/v, and 40% Acrylamide/Bis stock solution which is 1:1 diluted in the final solution. If a desired volume of 1L is prepared the urea is dissolved in a total solvent volume of 690 mL (100 mL of 10 X TBE, 500 mL of Acrylamide/Bis and 90 mL water). Must be stored at 2-8 °C.

Tris/EDTA/NaCl Buffer (TEN Buffer)

Used as a Washing Buffer for streptavidin magnetic beads. Contains 50 mM Tris-HCl (pH 7.5), 200 mM NaCl and 1 mM EDTA.

3.3 Denaturing PAGE (D-PAGE) gels

3.3.1 Preparation of the gels

The standard volume of gel required for a large PAGE (42 x 33 x 0.04 cm, L x W x D) is 60 mL and for a mini PAGE (17 x 16.5 x 0.1 cm, L x W x D) at least 30 mL. To prepare a 10% gel for a large PAGE, a 1:1 solution of the 20% monomer mixture is prepared using the Urea Diluent at a final volume of 60 mL. The ratio can be adjusted for other percentages according to standard buffer dilution. For a gel volume of 60 mL, the amount of TEMED used is 60 µL and the amount of APS (10%) is 600 µL. Large gels must be allowed to polymerize for at least 3 hours (preferentially overnight) and small gels must be allowed to set for at least 1.5 hours.

3.3.2 Running a gel sample and image analysis

Before starting a gel run, the gels must be pre-equilibrated with Running Buffer (1X TBE). This allows the gel to warm up uniformly, and wash the urea away from the wells. Samples are prepared in Gel Loading Solution, with 0.5 M NaOH (NaOH is, however, not used in oligonucleotides with an internal ribonucleotide) and then heated to 95 °C. The sample is then snap cooled at 4 °C on an ice bath, and loaded into wells that have been pre-washed with running buffer. The large gels are run for at least 2 hours on 10% D-PAGE and 3 hours on 20% D-PAGE (the bromophenol blue layer must have run 90-100% of the length of the gel depending on the resolution required). 20% D-PAGE in our experiments have only been used for purification of primers or modified nucleotide incorporation assays. It is sufficient to run 45 minutes for a small gel (usually 8-10% D-PAGE). The oligonucleotides are visualized by either UV shadowing, or by autoradiography depending on the application. The data can be obtained either by directly scanning the gel (in the case of fluorescent oligonucleotides) or by exposing a phosphorescent screen. The gel is usually frozen during exposure and the screen is normally exposed overnight. Data from the phosphorescent screen is obtained using a Typhoon Scanner Model 9200. The image is then further processed using Fiji.

3.4 Gel elution protocol

Subsequently, after visualization, the bands are cut out using a sterile surgical blade that can be detected either by UV shadowing or by cross-referencing with a standard oligonucleotide (uncleaved/cleaved DNAzyme or an oligonucleotide of known molecular weight) in case of radioactivity. In either case, the gel fragments are collected in a 1.5 mL tared Eppendorf tube and the tube is then weighed on an analytical balance to determine the gel mass. A 1.5 mL flame sealed

pipette tip is used to mash the gel shard until a fine consistency is obtained. About 3-4 volumes (in μL) of Gel Eluting Solution is added for every 1 mg of gel mass. It is further agitated on a vortexer and the finely distributed suspension is then frozen at -78 °C (dry ice). The frozen solution is immediately thawed over 95 °C for 5 minutes (the rapid freeze-thaw serves to break up the gel and release the DNA) and spun down at 13krpm. The supernatant is carefully decanted into a fresh Eppendorf tube without disturbing the gel sediment. This is then repeated three more times but it is incubated at 65 °C (the freeze-thaw step is only done during the first iteration). The collected fraction (combined supernatant) is evaporated in-vacuo and then precipitated using 5 volumes of absolute ethanol (or >95%). The supernatant is discarded and the pellet is dissolved in nuclease-free water or Tris buffer depending on the future application. It was then quantified on a SpectraMax QuickDrop Micro-Volume Spectrophotometer.

3.5 Amplification protocol

Table 3.2 PCR amplification protocol for SELEX

			Reaction Volume
Components	**Initial Stock**	**Final Concentration**	**Volume (µL)**
Thermopol Buffer	10X	1X	10
dNTP	20 mM	300 µM	1.5
$MgSO_4$	100 mM	1 mM	1.0
A. ON2 Stock Solution **B. ON3 Stock Solution**	175 µM	7 µM	4.0
A. ON3 Stock Solution **B. ON4 Stock Solution**	51 µM	7 µM	13.8
Template DNA			3.0
Vent DNA Polymerase	2 U/µL	0.08 U/µL	4.0
Nuclease Free Water			61.7
		Final Volume (µL)	**100**

(A) refers to the primer stocks included in the first amplification mixture whereas (B) refers to the second amplification cocktail.

3.5.1 Optimized PCR conditions for thermal cycling

General PCR amplification procedure is as follows: – 95 °C for 3 minutes (Initial Denaturation); 95 °C for 15 seconds (Denaturation), 68.8 °C for 10 seconds (Annealing), 72 °C for 20 seconds (Extension); and 72 °C for 5 minutes (Final Extension). A total of 25 cycles were performed for the first amplification and 35 cycles for the second amplification.

3.5.2 Amplification cocktail for sequencing (5X stock)

Amplification cocktail for sequencing was prepared as a 5X solution containing Thermopol Buffer (5X), dNTP mixture (1.5 mM each nucleotide) and DMSO solution (25%) that was divided into aliquots of 10 µL each. The amplification was done in a final volume of 50 µL.

3.5.3 Primer cocktail for sequencing the selected library

A primer stock (ON6 and ON7) was prepared at a concentration of 25 µM of each primer in nuclease-free water and then divided into aliquots of 1 µL.

3.6 DNA purification

3.6.1 Extraction of the enzyme-treated DNA

After each amplification (1st and 2nd) and their corresponding exonuclease digestion (procedure in Section 3.7), the DNA is extracted from the reaction impurities (monomers, enzyme, and/or primers) by a biphasic separation. The enzyme-treated DNA that is in an aqueous medium is mixed with equal parts of phenol:chloroform:isoamyl alcohol (25:24:1, v/v) and vortexed well. The layers are then separated by spinning them in a mini centrifuge for 5 minutes. The aqueous layer is collected and precipitated with absolute ethanol (to remove residual salts) followed by centrifugation at 13 krpm for 5 minutes. The supernatant is discarded and the pellet is dried in-vacuo.

3.6.2 Extraction of oligonucleotides received from the vendor

The DNA is first extracted from any non-polar impurities (that might have passed on from the solid-phase synthesis) by phenol/chloroform extraction as described previously. It is further precipitated with 9 parts of 3% $LiClO_4$ in acetone by vortexing and then spun down at 13 krpm for 5 minutes. The supernatant is discarded and the pellet is further precipitated using absolute ethanol

which is then dried down in-vacuo. The purified oligonucleotides are stored in 10 mM Tris pH 8.1.

3.7 Lambda exonuclease digestion

The PCR amplified template after clean-up (by phenol/chloroform extraction) is resuspended in 1X Lambda Exonuclease Buffer, 5 units of lambda exonuclease and then brought to 50 μL with nuclease-free water. The reaction is then incubated at 37 °C for 3 hours.

3.8 Materials and methods for selection of DNAzymes

3.8.1 Buffers and solutions

Buffer 1 (Cleavage buffer 1): 25 mM sodium cacodylate (pH 7.5), 2 mM $MgCl_2$ and 200 mM NaCl in DEPC treated H_2O (nuclease free water).

Buffer 2 (Cleavage buffer 2): 25 mM sodium cacodylate (pH 7.5), 0.5 mM $MgCl_2$ and 200 mM NaCl in nuclease free H_2O.

Buffer 3 (Cleavage buffer 3): 25 mM sodium cacodylate (pH 7.5), and 200 mM NaCl in nuclease free H_2O.

Buffer 4 (Neutralization buffer): 25 mM sodium cacodylate buffer in nuclease-free H_2O, pH 6.

pH characterization buffers:

Buffer 5: 25 mM phosphate buffer at pH 6; **Buffer 6:** 25 mM phosphate buffer at pH 6.5; **Buffer 7:** 25 mM phosphate buffer at pH 7; **Buffer 8:** 25 mM phosphate buffer at pH 8; **Buffer 9:** 25 mM Tris buffer at pH 8.5; **Buffer 10:** 25 mM Tris buffer at pH 9.

Metal solutions in buffer (metal ions):

Buffer 6 (Transition metal Solutions): Zn^{2+}, Fe^{2+}, Cu^{2+}, Ca^{2+}, Mn^{2+}, Co^{2+} at 0.5 mM, and Hg^{2+} 50 μM in Buffer 3

Buffer 7 (Monovalent metal solution): K^{+} prepared at 3, 30, and 300 mM in Cleavage buffer 3.

Buffer 8 (Mg^{2+} solutions): Mg^{2+} solution prepared at a final concentration of 0.5, 5, 10, 15, 100 mM in Buffer 3.

Nucleotide Cocktail A: 50 μM of each $dC^{aa}TP$, $dU^{ga}TP$

Nucleotide Cocktail B: 30 μM of dRTP; dRTP comprise of dATP and dGTP.

(All reported concentrations of the buffers and solutions are reported as final cleavage mixture concentration.)

3.8.2 Primer extension reaction

3.8.2.1 Library preparation for SELEX

20 pmol of PAGE-purified template and 20 pmol of primer (ON1) (template and primer final concentrations at 1 μM) are heat denatured at 95 °C for 5 minutes and then cooled to 37 °C over 30 minutes in a water bath (however during the characterization assays this was done in a ThermoCycler as described later). This allows enough time for annealing the primer to the template. The tube is then treated with 1-2 μL of dGTP or dATP α-[^{32}P] (interchangeable based on availability), dithiothreitol (DTT), yeast pyrophosphatase inorganic yeast (YIPP), Klenow (exo-) DNA polymerase, nucleotide cocktail A and nucleotide cocktail B. The reaction is then allowed to extend at 37 °C for 2 hr 15 min. The concentration of the components in the final primer

extension reaction mixture is 5 mM DTT, 5 U/mL YIPP, 0.75 U/μL Klenow Fragment, 30 μM dATP and dGTP each, 50 μM of modified nucleotides **3** and **4** each (Chapter 2, Fig 2.1), 5-10 μCi of α-[^{32}P]-dGTP in 1X NEBuffer 2. To stop the reaction it is quenched with EDTA (25 mM) and then immobilized on 40 μL of streptavidin magnetic beads which is prewashed with TEN buffer. For every 20 pmol of primer (in the primer extension reaction) 40 μL of beads are used and can be scaled accordingly.

3.8.2.2 Library preparation for time course assays

The annealing mixture for the characterization experiments is identically prepared but using 15 pmol of the primer (ON1) and 20 pmol of PAGE-purified template (primer to template ratio at 0.75-0.8:1). The reaction mixture can be scaled according to the requirement as long as the primer:template ratio is maintained. Primer annealing is done in a ThermoCycler preset at 95 °C for 5 minutes followed by a step-wise decrement in temperature from 95-37 °C at a rate of -2 °C/min for 30 mins. The radioactive dGTPα-[^{32}P] or dATPα-[^{32}P], DTT, YIPP, NEBuffer 2, and Nucleotide Cocktails A and B are added identically as above. After the reaction is complete it is quenched with EDTA (25 mM) and then immobilized on prewashed streptavidin magnetic beads. The volume of beads is according to the scale of production (as mentioned above).

3.8.3 Preparation of streptavidin magnetic beads

During in-vitro selection, about 40 μL of streptavidin magnetic beads are used for a final selection volume of 100 μL. The beads are washed three times with TEN Buffer before they are ready for immobilizing the nucleic acids. The beads must not be left dry for long periods.

3.8.4 Single-strand DNA generation

The quenched reaction was immobilized on the washed streptavidin magnetic beads for 30 minutes to allow maximum binding to occur. The immobilized oligonucleotides were separated from other reaction components (residue from primer extension) by decanting and then washed twice with 100 μL TEN buffer. The TEN wash further removes any remaining impurities. It is followed by five rapid washes with 100 μL NaOH-EDTA (0.1M NaOH, 1 mM EDTA) to denature the dsDNA, followed by neutralization in 200 μL Buffer 4, and finally washed with 100 μL nuclease-free water before incubating them in selection buffer. Due to the presence of the ribocytosine group in the DNAzyme strand, this process must be done rapidly to ensure that there is minimum loss of product due to premature cleavage but there must be a sufficient number of washes so that the DNAzyme population is completely denatured from its complementary strand. The selection was carried out in Buffer 1 for the first 5 generations (G1-G5), followed by Buffer 2 from G6 onwards.

3.8.5 Selection Set-Up

The washed DNAzyme population was incubated on a 25-27 °C temperature block with 100 μL Buffer 1 or Buffer 2 (vide supra) for a predetermined duration of time. The self-cleavage was allowed to take place over a temperature block preset at 25 °C (and wrapped in aluminum foil) to minimize fluctuations in the reaction temperature due to unforeseeable changes in the external conditions.

3.8.6 Post Selection Processing and PAGE

Following incubation, the supernatant containing the cleaved DNA is collected by decanting. The decanted elute fraction was then precipitated with 9 parts of 1% $LiClO_4$ in acetone by agitating

over a vortexer and then spun down at 13 krpm for 5-7 minutes (a small amount of ON2-ON3 primer cocktail added as a carrier to help precipitate the radiolabeled sample). The supernatant was discarded and the pellet was washed with 5 volumes of absolute ethanol. After the addition of ethanol, the tube is vortexed to wash away any residual salts from the pellet and then the DNA is precipitated by centrifugation. The ethanol supernatant was removed and the pellet was dried in-vacuo before preparing the sample for D-PAGE. The post-selection sample was resuspended in a solution of Formamide-EDTA at 62.5% v/v. Extra EDTA may be added if the nucleic acid catalysts are excessively structured (as in later generations). The sample was treated at 95 °C for 5 minutes, and then snap cooled before loading it on an 8% 7 M polyacrylamide gel. The gel is run at 35-36 W for about 2 hrs or until the bromophenol blue is more than 75% of the length of the gel.

3.8.7 Amplification of the cleaved DNA

The cleaved fraction was excised from the gel and extracted according to the Gel Elution procedure described previously and further prepared for double nested PCR. The gel-purified material was resuspended in water. A volume of 15 μL of the purified oligonucleotides was amplified using primers ON2 and ON3 (all components in Table 3.2) and brought to 100 μL with nuclease-free water. It is amplified according to Section 3.5.1. The progress of the amplification was monitored using a 4% agarose gel. The amplification product was then purified by Phenol-chloroform extraction and ethanol precipitation (Section 3.7.1). The extracted oligonucleotide was resuspended in nuclease-free water, and 15-20 μL of it was used in the lambda exonuclease reaction mixture (Section 3.6). The exonuclease reaction is allowed to digest for 3 hours and was then stored at 4 °C overnight (in the ThermoCycler) if it was not attended to immediately. After

the strand containing the 5' phosphate had been digested the oligonucleotide is further cleaned up by biphasic extraction (phenol/chloroform extraction) and gel purified by 10% D-PAGE. For a D-PAGE the pellet was resuspended in 15.6 µL of gel loading solution and then made to 25 µL with nuclease-free water (Section 3.3). This sample was run parallel to an 80 base control DNA sample. The DNA fragment was identified based on the 80 base control DNA and then extracted according to the gel extraction protocol in Section 3.4. 15-20 µL of the resulting product was amplified using primers ON3 and ON4 for 35 cycles (to maximize the production of DNA) to generate the DNA library for the following generation. The second amplicon was further exonuclease digested and purified on an 8-10% mini D-PAGE. The eluted product was then carried forward for the subsequent generation. After quantifying the DNA using a SpectraMax QuickDrop™ Micro-Volume Spectrophotometer 20 pmol of it was taken out for Primer Extension reaction (Section 3.8.2). Although 13 generations were done in total, the sequences from G11 were cloned (Chapter 2, Fig 2.4).

3.8.8 Preliminary verification prior to sequencing

The pool from generation 11 (G11) was investigated for confirmation prior to sequencing. Samples from G9 and G11 were prepared using Klenow for primer extension and then the single-stranded DNAzyme was generated by washing and neutralization (Section 3.8.4). The beads were resuspended in Buffer 2 and allowed to incubate for 2.4 hr, a slurry was prepared and samples were collected at regular time points (0.5-144 min). The samples were quenched into aliquots of formamide dye solution containing 25 mM EDTA. The samples were heated to 95 °C and resolved on an 8% D-PAGE. The experiment was at least duplicated for confirmation.

3.8.9 Sequencing of Amplicons by Amplicon-EZ NGS

The DNA sample from G11 was first amplified using the primers ON2 and ON3 (following Table 3.2). After cleaning up the DNA amplicon by phenol/chloroform extraction, 0.5 µL of the sample was further amplified with 1 µL of the primer cocktail (Section 3.5.3) and 10 µL of the amplification cocktail (Section 3.5.2) which was brought to a final reaction volume of 50 µL. A small amount of material was sampled from the amplified product and verified by 4% agarose gel. The remaining amplification reaction was purified on a separate 4% agarose gel. With reference to a DNA ladder, (Quick-Load Purple Low Molecular Weight DNA Ladder) the desired band was excised from the gel, and purified according to the manufacturer's protocol in the QIAquick Gel Extraction Kit. The excised DNA fragment was initially collected in a tared Eppendorf tube and weighed on an analytical balance, followed by incubating in Buffer QG (volume of buffer is 6 µL for every 1 mg of gel) at 50 °C for 10 minutes. The tube must be agitated every 2–3 min to help resuspend the gel. After the gel slice has completely resuspended, the colour of the mixture should be yellow (similar to Buffer QG without dissolved agarose). If the colour of the mixture is orange or violet, 10 µL 3M sodium acetate, pH 5.0 must be added and mixed to make the mixture turn back yellow.

The biological load (DNA) is then precipitated from the residual gel fragments using 1 volume of isopropanol to the mass of gel. The sample is then applied over a QIAquick spin column to allow the DNA to bind to the column. When the tube is centrifuged at 13 krpm for 1 minute the DNA remains bound to the column while the elute passes through. The column is then centrifuged for 1 min after the addition of 500 µL of Buffer QG. To wash the column, it is incubated, with 750 µL Buffer PE, for 5 minutes and then centrifuged at 13 krpm for 1 minute. The flow-through is

discarded and the column is placed on a fresh, clean Eppendorf tube. DNA is eluted into the fresh tube by incubating the column in 50 μL Buffer EB (10 mM Tris-Cl, pH 8.5) for 1 minute before centrifuging for a further 1 minute. The DNA samples were quantified using Quant-iT PicoGreen assay before submitting for sequencing at Genewiz, LLC. The purified oligonucleotide is normalized to 20 ng/μL at about 500 ng. Sequencing is done using Amplicon-EZ (an Illumina Sequencing based platform). The FASTQ files were pre-processed by Genewiz to remove the NGS adapter regions at the 5' terminal. The pre-processed samples were then further filtered using the MultiQC program on Galaxy Version 1.7.1 for sequences with a Phred score around 40 (represents the average quality of NGS read at a particular nucleotide). The sequences were then aligned using the Multiple Sequence Alignment tool in Jalview v1.0 to obtain sequence families.

3.8.10 mFold Analysis for preliminary structure prediction

The preliminary structure prediction is done using mFold (on the UNAFOLD webserver). The structures were aligned using constraints to allow the N40 selected region to be in the vicinity of the scissile phosphate. I used forced constraints to allow certain base pairing to occur. The forced constraint method is the same as forcing a helix to form. The helix is determined by the external closing base pair (which is by the base pair *i.j* where *i* is as small as possible and *j* is as large as possible, *i* and *j* denote the position of the base on the DNA strand). A second parameter used is *k* which denotes the number of consecutive base pairs beginning with *i.j*. The base pairs are therefore: *i.j, i+1.j-1*, *i+2.j-2*, …, *i+k-1.j-k+1*. It is also important that the base pairs must actually be able to form. Further, the ionic and temperature conditions are set according to the selection parameters before allowing mFold to generate the secondary structures.

3.8.11 Preliminary Assay in Identifying Best Clones

10 identified clonal families (all clones mentioned in the Appendix), synthesized by Integrated DNA Technologies, after PAGE Purification, were extended using the primer ON1 according to Section 3.8.2.2. The dsDNA intermediates were then immobilized on washed Streptavidin Magnetic Beads for 30 minutes to allow maximum binding. The generation of single-stranded DNA is similar to the protocol used during selection (Section 3.8.4), except, after the water wash the samples were resuspended in a sufficient volume such that each clone was then separated into three tubes (technical replicates). The clones were allowed to self-cleave for a total of 60 minutes and samples were collected periodically (1, 3, 30, 60 min). 1 μL of the sample was taken out from each experiment at specified time points (Chapter 2, Fig 2.7), and quenched into a Formamide-EDTA solution that was later made to 10 μL with nuclease-free water. While the experiment was in-progress all samples must be preserved over ice to prevent any unwanted cleavage from happening to the previous samples. All the samples were then heated to 95 °C and snap cooled on ice. They were subsequently resolved by 8% D-PAGE run for 2 hours. The gels are then imaged on a phosphoimager screen overnight and checked for cleavage rate using Typhoon Scanner Model 9200 coupled with ImageQuant for the image analysis.

3.8.12 Time-Course Assay to determine the kinetics of the best candidates

Kinetic analysis was performed on clones Dz11-10, Dz11-15, Dz11-23 and Dz11-33; 10 pmol of primer ON1 was annealed with each template sequence (primer to template ratio is however maintained at 0.8:1) as mentioned previously. The initial libraries were prepared similarly to the Primer Extension protocol (Section 3.8.2.2). About 1 μL of sample from each candidate DNAzyme was removed as Control for the uncleaved material. The reactions were quenched with EDTA to

a final concentration of 25 mM. The sequences were immobilized on prewashed streptavidin magnetic beads (concentration of 40 µL beads for 100 µL solution) for 30 minutes and the supernatant was then discarded. The beads were washed twice with TEN buffer (100 µL), followed by five quick washes of 0.1 M NaOH-1 mM EDTA. The beads were then neutralized with Buffer 4 followed by a final water wash before splitting them into individual aliquots each containing 3 pmol of DNAzyme. Kinetics experiments for individual clones were triplicated within the experiment and a total of two sets of experiments were done on separate days.

The self-cleavage reaction was initiated by resuspending the beads in 50 µL Buffer 2 and incubating at 25 °C. Samples were taken out at intervals (0.5 min-15 hr) and quenched in 5 µL formamide, bromophenol blue, 25 mM EDTA and 275 µM biotin. The samples were heat-denatured at 95 °C and snap-cooled on ice before separation by 8% D-PAGE. The gels were autoradiographed overnight and imaged using a phosphorimager (Amersham Typhoon 9200). The dosimetry was completed with Fiji (v2.0.0). Statistical analysis and biphasic kinetics curve fitting were performed using Prism (v8.4.3). The percentage of DNAzymes cleaved (P_t) is expressed as a function of time (t) and a linear combination of the kinetically fast and slow phases with horizontal asymptotes ($P_{fast,max}$ and $P_{slow,max}$) and rate constant (k) such that $k_{fast} > k_{slow}$ (as given in equation 2):

$$P_t = P_{fast,max}\left(1 - e^{-k_{fast}t}\right) + P_{slow,max}\left(1 - e^{-k_{slow}t}\right) \qquad (Equation\ 2)$$

3.8.13 pH-rate profile on DNAzyme self-cleavage

60 pmol of ON1 was annealed with either Dz11-23 and Dz11-33 template for primer extension, worked up and single-strand generation, similarly to section 3.8.4. Aliquots of 3 pmol of DNAzymes were distributed among tubes and allowed to cleave in Buffers 5-10 ranging from pH 6 to 9. Samples were collected at different time points and quenched in formamide, bromophenol blue, 25 mM EDTA and 275 μM biotin before resolving in 8% D-PAGE. These data were used to study the variation in the rate of the DNAzymes over different pH. All data are reported from two independent sets of experiments and each sample was done in triplicate.

3.8.14 Temperature dependence on self-cleavage

An estimated 3-6 pmol of a self-cleaving DNAzyme, e.g. Dz11-33, was prepared (as previously described) and then incubated at each temperature (5, 25, 37, and 55 ºC) to address temperature dependence. Buffer 2 was added to each of the samples and incubated for a desired duration of time. All reactions except the tube at 25 °C were performed in a thermocycler with a heated lid to prevent evaporation of the solvent. First-order rate constants k_{fast} and k_{slow} are obtained by curve fitting to equation (2) above.

3.8.15 Dependence of Mg^{2+} on the rate and the impact of monovalent K^{+} ions

60 pmol of self-cleaving DNAzymes were prepared using the method described above for selection and aliquoted into tubes containing varying amounts of Mg^{2+}. The isolated single-stranded DNAzyme was treated with different concentrations of Mg^{2+} (0-100 mM) and the reaction was quenched after a stipulated time, following separation by 8% D-PAGE. To study the impact of the

monovalent K^+ ions on the cleavage, the samples were run in varying concentrations of KCl (Buffer 7) at room temperature (25 °C).

3.8.16 Effect of transition metal ions

Self-cleaving sequences Dz11-23 and Dz11-33 were prepared by primer extension as described previously starting with up to 60 pmol of template DNA. Following primer extension, template strand removal, and washing, bead-bound self-cleaving strands were aliquoted into tubes such that each contained ca. 3 pmol. The DNAzymes were incubated in Buffer 3 containing divalent transitional metal ions at a final concentration of 0.5 mM except in the case of Hg^{2+} which was present at 50 μM. The reactions were incubated at room temperature for a stipulated amount of time before collecting samples by quenching in a 4 μL stop solution consisting of formamide-EDTA-dyes. Samples were resolved by 8% D-PAGE and exposed on a phosphorimager screen. The images were further analyzed by Fiji after autoradiography.

3.8.17 Effect of modified nucleotide composition on the cleavage profile

The DNAzyme (Dz11-33) was synthesized by primer extension as noted above (Section 3.8.2.2) but different composition of nucleotides were added to each of the four sets of tubes reported to final concentrations – 1) 50 μM $dU^{ga}TP$ 4, 50 μM $dC^{aa}TP$ 3, 30 μM dGTP, 30 μM dATP; 2) 50 μM $dU^{ga}TP$ 4, 30 μM dCTP, 30 μM dGTP, P and 30 μM dATP; 3) 30 μM dTTP, 50 μM $dC^{aa}TP$ 3, 30 μM dGTP, and 30 μM dATP; 4) 50 μM $dU^{aa}TP$ 2, 50 μM $dC^{aa}TP$ 3, 30 μM dGTP and 30 μM dATP; and 5) 30 μM dNTP where dNTP refers to all unmodified deoxynucleoside triphosphates. The reaction was quenched with 25 mM EDTA after 2.25 hr and single-stranded self-cleaving species were prepared in accordance with the selection method. The immobilized

DNAzymes were incubated in Buffer 2 for different intervals of time (1 min, 30 min, and 15 hr). 1 μL of samples were progressively taken out and quenched in the formamide-EDTA solution. Samples are resolved by 8% D-PAGE followed by autoradiography.

3.8.18 Target sequence specificity in self-cleavage reactions

Dz11-33 was prepared as noted previously except that primers ON1, ON8, ON9, and ON10 were used to vary the target ribonucleotide. Following primer extension, the different variants of Dz11-33 were converted to single-stranded self-cleaving species, washed with water and interrogated for self-cleavage in Buffer 2 with aliquots quenched at various time points. The quantity of cleavage yield was calculated using autoradiographic densitometry and k_{obs} of each of the DNAzymes were then calculated using Prism.

Chapter 4: Conclusion

4.1 Summary of thesis

To conclude, the thesis addresses my preliminary hypothesis regarding the pursuit of RNase A-mimicking DNAzymes using only two modified nucleotides notably: 5-aminoallyl-2'-deoxycytidine triphosphate **(dC^{aa}TP, 3, Chapter 2)** and 5-guanidinoallyl-2'-deoxyuridine triphosphate **(dU^{ga}TP, 4, Chapter 2)** without using an imidazole functionality. In **Chapter 2** I initially discussed the optimization of the Klenow primer extension reaction to address the maximum acceptable concentration of Klenow (exo-) polymerase (0.75 U/μL) before yields diminished. I then further discussed the in-vitro selection of a bi-modified DNAzyme selection using the above-mentioned nucleotides. **Chapter 2** also discussed the characterization of two of the identified DNAzymes (Dz11-23 and Dz11-33) which were able to foster on-par fast-cleaving sequence-specific activity (k_{fast} = 0.7-0.9 min^{-1}) using positively charged modifications presented by a cationic amine and a guanidium cation without any imidazole.

Despite selecting them in the presence of Mg^{2+} ions, the self-cleavage activity of the DNAzymes was significantly better at a lower concentration of Mg^{2+} ion (0.5 mM) with the best activity observed under the Mg^{2+}-independent condition. As discussed in the thesis, up to this point there has been only a little development in metal-independent DNAzyme selections and the discovery of such modified DNAzymes improve our ability to develop true metal-independent catalysts as intracellular therapeutic agents. Earlier, I had stated in the thesis that the imidazole group (a core component in the catalytic site of RNase A is an essential group in the catalysis) had been crucial to the activity of many metal-independent DNAzymes as it could extend acid-base catalysis in the pH range (6-9). Many of the metal-independent DNAzymes as a result include imidazole as one

of the functionalities. The imidazole based nucleotides are however a challenging substrate in terms of incorporation using many polymerases.

Herein, experiments also demonstrated that the nucleotides **dC^{aa}TP (3)** and **dU^{ga}TP (4)** were incorporated in the DNAzymes much better with few truncates, however they unexpectedly showed a lack of a relationship between the cleavage activity of DNAzyme and the pH (in the range 6-9) in comparison to their imidazole counterpart. The pH-rate profile normally seen with DNAzymes that contain the imidazole functionality follows a bell-shaped curve in the pH range (6-9) however most likely due to the high pKa of the guanidium ion (~12) and the amine (~10) a normal acid-base profile was not observed. The cationic amines and the guanidium group which may not be readily associated as bases could extend catalysis through various means such as possible Brønsted base catalysis that is favored at higher pKa values. The pKa perturbation of the amine functionality due to the guanidium groups could also have afforded a self-cleavage at the ribophosphodiester. As discussed earlier in the thesis, electrostatic or H-bond stabilization of a pentacoordinate phosphorane intermediate may be disfavored at higher pH values. Further, the pH-dependent effect of our catalysts may have been confounded due to the presence of two operating phases (fast cleaving and slow cleaving species), the existence of which I have alluded to in the refolding study (Fig 2.12) and our mFold analysis. Conformational effects may include folding or an acid-base reaction that is operative at pH values that are outside of the range in which the DNAzyme would normally be folded. Further, a definitive mechanism for the lack of a bell-shaped pH curve is not readily interpreted and may be subject to further mechanistic studies on Dz11-23 and Dz11-33.

The results from these experiments further aid our understanding of our aptamer selections (projects not included as part of the thesis) that failed to result in affinity ligands to adrenocorticotropic hormone (ACTH) and TAR DNA-binding protein (TDP-43) using guanidium and primary amine as functional groups. Notwithstanding selection related parameters such as stringency of negative selections, the aptamers failed to develop specificity towards our target; cationic groups like a guanidium and a primary amine, and the lack of any hydrophobic handles presented less surface for interaction with the peptide/protein targets. Proteins in their native conformation usually sequester polar residues leaving hydrophobic patches on the surface that are protected by a water shell. Further, the higher molecular weight beads (such as streptavidin magnetic bead and agarose Ni-NTA beads) in comparison to the peptides/protein of interest (< 35 kDa) may have been competitively better targets for the aptamers. This could be addressed by precoating the beads (charging) with tRNA to block potential binding sites. This is, however, a hypothesis that needs to be tested further, It is also possible that since my potential modification for targeting is a nucleophilic amine (**dC^{aa}**), the modified nucleoside provide a better handle for catalysis at a scissile ribonucleoside, suggesting that the two modified nucleotides used in this thesis are more suitable in catalytic DNAzymes over affinity ligands (aptamers). It is likely that further evolution of Dz11-33 and the use of a larger library at the outset of the selection will enable the selection of even more proficient self-cleaving species. In addition, one might consider decorating nucleosides with polyamines or involving a third nucleoside that introduces a hydrophobic effect in lieu of imidazole in the expectation that such will further enhance activity by providing localized domains of altered dielectric constants and might also help in selecting better aptamers or allosteric DNAzymes as a result. It is anticipated that Dz11-33 could be

converted into a multiple turnover catalyst with the nucleosides **dC^{aa} (3)** and **dU^{ga} (4)** helping in the *trans*-cleavage of an all-RNA target with multiple turnovers.

4.2 Future Directions

The catalyst Dz7-38-32[109], an all-RNA cleaving, multiple turnover M^{2+}-independent DNAzyme that could cleave a 19 nt RNA showed a second-order rate constant (k_{cat}/K_m) of ~10^6 $M^{-1}min^{-1}$ and k_{cat} of 1.06 min^{-1}. After that discovery, Dz11-23 and Dz11-33 presented in this thesis are subsequent iterations to that research. Since the discovery of DNAzymes, much has been achieved[89,90,92,104,106,108,109,117,119,125,127,130,132,133,137,142,156,158,169–171], yet with little progress on DNAzymes that are M^{2+}-independent and much less initiative on a M^{2+}-independent DNAzyme that also lacks an imidazole group. From a therapeutic perspective, the Dz11-23 and Dz11-33 could therefore be potentially tested further in cell-based assays. As I discussed previously, I was able to identify superior metal independent catalysts with a guanidium and a cationic amine without using an imidazole functional group, this could further the discovery of metal-independent RNA-cleaving DNAzymes without this functionality (which is a poor substrate for many polymerases). Since Dz11-33 also showed a high specificity towards a ribocytidine target, this could further be tested on an all-RNA target as in DNAzyme Dz7-38-32. The DNAzymes in the thesis show on-par fast cleaving activity (k_{fast} = 0.7-0.9 min^{-1}) to many of the imidazole based DNAzymes.

Hence, it might also be wise to perform a “cutdown” or a “boundary” experiment to identify a more catalytically active core[109,172]. This is achieved by truncating a more long-range secondary structure into smaller fragments which might improve catalytic efficiency. There is an agreement

in the research community that not all sequences inside the structure may be essential for cleavage and consensus sequences may be essential to buttress the structure and may not contribute to the activity. Eliminating such extraneous sequences might reduce steric bulk around the active site. Another effort to improve the catalytic efficiency may be the reselection of the isolated clones (of DNAzyme) which has been previously done in our lab[109]. The active clone is reshuffled by introducing mutations in the degenerate region when the library is synthesized on the solid phase; mutagenic PCR can further introduce more sequence diversity to the new library, and as a result, increase the sequence space for isolating faster self-cleavers.

In conclusion, this work is the first report of a self-cleaving M^{2+}-independent DNAzyme that provides a reasonable self-cleaving activity in the absence of imidazole. This might make us wonder if imidazole is truly critical for fostering an efficient RNA cleavage. I anticipate further experiments on these DNAzymes and newer discoveries in providing faster all-RNA cleaving metal-independent activity.

Bibliography

(1) Alberts, B. *Essential Cell Biology*, Fourth edition.; Garland Science: New York, NY, 2013.

(2) *Molecular Biology of the Gene*, Seventh edition.; Watson, J. D., Ed.; Pearson: Boston, 2014.

(3) FRANKLIN, R. E.; GOSLING, R. G. Molecular Configuration in Sodium Thymonucleate. *Nature* **1953**, *171* (4356), 740–741. https://doi.org/10.1038/171740a0.

(4) Crick, F. H. C.; Watson, J. D. The Complementary Structure of Deoxyribonucleic Acid. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1954**, *223* (1152), 80–96.

(5) Chargaff, E. Structure and Function of Nucleic Acids as Cell Constituents; 1951; Vol. 10, pp 654–659.

(6) Črnugelj, M.; Hud, N. V.; Plavec, J. The Solution Structure of d(G4T4G3)2: A Bimolecular G-Quadruplex with a Novel Fold. *Journal of Molecular Biology* **2002**, *320* (5), 911–924. https://doi.org/10.1016/S0022-2836(02)00569-7.

(7) Lyonnais, S.; Hounsou, C.; Teulade-Fichou, M.-P.; Jeusset, J.; Cam, E. L.; Mirambeau, G. G-Quartets Assembly within a G-Rich DNA Flap. A Possible Event at the Center of the HIV-1 Genome. *Nucleic Acids Research* **2002**, *30* (23), 5276–5283.

(8) Pandey, S.; Agarwala, P.; Maiti, S. Effect of Loops and G-Quartets on the Stability of RNA G-Quadruplexes. *J. Phys. Chem. B* **2013**, *117* (23), 6896–6905. https://doi.org/10.1021/jp401739m.

(9) Ten Dam, E.; Pleij, K.; Draper, D. Structural and Functional Aspects of RNA Pseudoknots. *Biochemistry* **1992**, *31* (47), 11665–11676. https://doi.org/10.1021/bi00162a001.

(10) Staple, D. W.; Butcher, S. E. Pseudoknots: RNA Structures with Diverse Functions. *PLoS Biol* **2005**, *3* (6), e213. https://doi.org/10.1371/journal.pbio.0030213.

(11) Ke, A.; Zhou, K.; Ding, F.; Cate, J. H. D.; Doudna, J. A. A Conformational Switch Controls Hepatitis Delta Virus Ribozyme Catalysis. *Nature* **2004**, *429* (6988), 201–205. https://doi.org/10.1038/nature02522.

(12) Rastogi, T.; Beattie, T. L.; Olive, J. E.; Collins, R. A. A Long-Range Pseudoknot Is Required for Activity of the Neurospora VS Ribozyme. *The EMBO Journal* **1996**, *15* (11), 2820–2825. https://doi.org/10.1002/j.1460-2075.1996.tb00642.x.

(13) Scott, W. G. Ribozymes. *Current Opinion in Structural Biology* **2007**, *17* (3), 280–286. https://doi.org/10.1016/j.sbi.2007.05.003.

(14) Steitz, T. A.; Moore, P. B. RNA, the First Macromolecular Catalyst: The Ribosome Is a Ribozyme. *Trends in Biochemical Sciences* **2003**, *28* (8), 411–418. https://doi.org/10.1016/S0968-0004(03)00169-5.

(15) Pichon, X.; A. Wilson, L.; Stoneley, M.; Bastide, A.; A King, H.; Somers, J.; E Willis, A. RNA Binding Protein/RNA Element Interactions and the Control of Translation. *CPPS* **2012**, *13* (4), 294–304. https://doi.org/10.2174/138920312801619475.

(16) Lim, F.; Spingola, M.; Peabody, D. S. The RNA-Binding Site of Bacteriophage Qβ Coat Protein. *Journal of Biological Chemistry* **1996**, *271* (50), 31839–31845. https://doi.org/10.1074/jbc.271.50.31839.

(17) Brownlee, G. G.; Fodor, E.; Pritlove, D. C.; Gould, K. G.; Dalluge, J. J. Solid Phase Synthesis of 5-Diphosphorylated Oligoribonucleotides and Their Conversion to Capped M7Gppp-Oligoribonucleotides for Use as Primers for Influenza A Virus RNA Polymerase in Vitro. **1995**, 7.

(18) Ellington, A.; Pollard, J. D. Synthesis and Purification of Oligonucleotides. *Current Protocols in Molecular Biology* **1998**, *42* (1), 2.11.1-2.11.25. https://doi.org/10.1002/0471142727.mb0211s42.

(19) Pon, R. T. Solid-Phase Supports for Oligonucleotide Synthesis. *Current Protocols in Nucleic Acid Chemistry* **2000**, 3.1.1-3.1.28.

(20) Lebedev, A. V.; Koukhareva, I. I.; Beck, T.; Vaghefi, M. M. PREPARATION OF OLIGODEOXYNUCLEOTIDE 5′-TRIPHOSPHATES USING SOLID SUPPORT APPROACH. *Nucleosides, Nucleotides and Nucleic Acids* **2001**, *20* (4–7), 1403–1409. https://doi.org/10.1081/NCN-100002565.

(21) Zlatev, I.; Lavergne, T.; Debart, F.; Vasseur, J.-J.; Manoharan, M.; Morvan, F. Efficient Solid-Phase Chemical Synthesis of 5′-Triphosphates of DNA, RNA, and Their Analogues. *Org. Lett.* **2010**, *12* (10), 2190–2193. https://doi.org/10.1021/ol1004214.

(22) Zlatev, I.; Manoharan, M.; Vasseur, J.; Morvan, F. Solid-Phase Chemical Synthesis of 5′-Triphosphate DNA, RNA, and Chemically Modified Oligonucleotides. *Current Protocols in Nucleic Acid Chemistry* **2012**, *50* (1). https://doi.org/10.1002/0471142700.nc0128s50.

(23) Sarac, I.; Meier, C. Solid-Phase Synthesis of DNA and RNA 5′- *O* -Triphosphates Using *Cyclo* Sal Chemistry. *Current Protocols in Nucleic Acid Chemistry* **2016**, *64* (1). https://doi.org/10.1002/0471142700.nc0467s64.

(24) Keefe, A. D.; Cload, S. T. SELEX with Modified Nucleotides. *Current Opinion in Chemical Biology* **2008**, *12* (4), 448–456. https://doi.org/10.1016/j.cbpa.2008.06.028.

(25) Joyce, G. F. Directed Evolution of Nucleic Acid Enzymes. *Annu. Rev. Biochem.* **2004**, *73* (1), 791–836. https://doi.org/10.1146/annurev.biochem.73.011303.073717.

(26) Ellington, A. D.; Szostak, J. W. In Vitro Selection of RNA Molecules That Bind Specific Ligands. *Nature* **1990**, *346* (6287), 818–822. https://doi.org/10.1038/346818a0.

(27) Wilson, D. S.; Szostak, J. W. In Vitro Selection of Functional Nucleic Acids. *Annu. Rev. Biochem.* **1999**, *68* (1), 611–647. https://doi.org/10.1146/annurev.biochem.68.1.611.

(28) Nitsche, A.; Kurth, A.; Dunkhorst, A.; Pänke, O.; Sielaff, H.; Junge, W.; Muth, D.; Scheller, F.; Stöcklein, W.; Dahmen, C.; Pauli, G.; Kage, A. One-Step Selection of Vaccinia Virus-Binding DNA Aptamers by MonoLEX. *BMC Biotechnol* **2007**, *7* (1), 48. https://doi.org/10.1186/1472-6750-7-48.

(29) Nieuwlandt, D.; Wecker, M.; Gold, L. In Vitro Selection of RNA Ligands to Substance P. *Biochemistry* **1995**, *34* (16), 5651–5659. https://doi.org/10.1021/bi00016a041.

(30) Sazani, P. L.; Larralde, R.; Szostak, J. W. A Small Aptamer with Strong and Specific Recognition of the Triphosphate of ATP. *J. Am. Chem. Soc.* **2004**, *126* (27), 8370–8371. https://doi.org/10.1021/ja049171k.

(31) Yamamoto, R.; Katahira, M.; Nishikawa, S.; Baba, T.; Taira, K.; Kumar, P. K. R. A Novel RNA Motif That Binds Efficiently and Specifically to the Tat Protein of HIV and Inhibits the Trans-Activation by Tat of Transcription in Vitro and in Vivo. *Genes Cells* **2000**, *5* (5), 371–388. https://doi.org/10.1046/j.1365-2443.2000.00330.x.

(32) Wiegand, T. W.; Williams, P. B.; Dreskin, S. C.; Jouvin, M. H.; Tasset, D. High-Affinity Oligonucleotide Ligands to Human IgE Inhibit Binding to Fc Epsilon Receptor I. **1996**, 11.

(33) Lu, Y.; Liu, J.; Li, J.; Bruesehoff, P. J.; Pavot, C. M.-B.; Brown, A. K. New Highly Sensitive and Selective Catalytic DNA Biosensors for Metal Ions. *Biosensors and Bioelectronics* **2003**, *18* (5–6), 529–540. https://doi.org/10.1016/S0956-5663(03)00013-7.

(34) Cruz-Aguado, J. A.; Penner, G. Determination of Ochratoxin A with a DNA Aptamer. *J. Agric. Food Chem.* **2008**, *56* (22), 10456–10461. https://doi.org/10.1021/jf801957h.

(35) Wang, R.; Zhao, J.; Jiang, T.; Kwon, Y. M.; Lu, H.; Jiao, P.; Liao, M.; Li, Y. Selection and Characterization of DNA Aptamers for Use in Detection of Avian Influenza Virus H5N1. *Journal of Virological Methods* **2013**, *189* (2), 362–369. https://doi.org/10.1016/j.jviromet.2013.03.006.

(36) Shangguan, D.; Li, Y.; Tang, Z.; Cao, Z. C.; Chen, H. W.; Mallikaratchy, P.; Sefah, K.; Yang, C. J.; Tan, W. Aptamers Evolved from Live Cells as Effective Molecular Probes for Cancer Study. *Proceedings of the National Academy of Sciences* **2006**, *103* (32), 11838–11843. https://doi.org/10.1073/pnas.0602615103.

(37) Cload, S. T.; McCauley, T. G.; Keefe, A. D.; Healy, J. M.; Wilson, C. Properties of Therapeutic Aptamers. In *The Aptamer Handbook*; Klussmann, S., Ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, FRG, 2006; pp 363–416. https://doi.org/10.1002/3527608192.ch17.

(38) Johansson, H. E.; Liljas, L.; Uhlenbeck, O. C. RNA Recognition by the MS2 Phage Coat Protein. *Seminars in Virology* **1997**, *8* (3), 176–185. https://doi.org/10.1006/smvy.1997.0120.

(39) Tuerk, C.; Gold, L. Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase. *Science* **1990**, *249* (4968), 505–510. https://doi.org/10.1126/science.2200121.

(40) Mills, D. R.; Peterson, R. L.; Spiegelman, S. An Extracellular Darwinian Experiment with a Self-Duplicating Nucleic Acid Molecule. *Biochemistry* **1967**, *58*, 217–224.

(41) Levisohn, R.; Spiegelman, S. The Cloning of a Self-Replicating RNA Molecule. **1968**, *60*, 866–872.

(42) Levisohn, R.; Spiegelman, S. Further Extracellular Darwinian Experiments with Replicating RNA Molecules: Diverse Variants Isolated under Different Selective Conditions. *Genetics* **1969**, *63*, 805–811.

(43) Saffhill, R.; Schneider-Berloehr, H.; Orgel, L. E.; Spiegelman, S. In Vitro Selection of Bacteriophage Qß Ribonucleic Acid Variants Resistant to Ethidium Bromide. *J. Mol. Biol.* **1970**, *51*, 531–539.

(44) Kacian, D. L.; Mills, D. R.; Kramer, F. R.; Spiegelman, S. A Replicating RNA Molecule Suitable for a Detailed Analysis of Extracellular Evolution and Replication. *Proceedings of the National Academy of Sciences* **1972**, *69* (10), 3038–3042. https://doi.org/10.1073/pnas.69.10.3038.

(45) Mills, D. R.; Kramer, F. R.; Spiegelman, S. Complete Nucleotide Sequence of a Replicating RNA Molecule: The Sequence Suggests How Nucleic Acids Exhibit Phenotypes for Selection and Can Evolve to Greater Complexity. *Science* **1973**, *180* (4089), 916–927. https://doi.org/10.1126/science.180.4089.916.

(46) Kramer, F. R.; Mills, D. R.; Cole, P. E.; Nishihara, T.; Spiegelman, S. Evolution in Vitro: Sequence and Phenotype of a Mutant RNA Resistant to Ethidium Bromide. *Journal of Molecular Biology* **1974**, *89* (4), 719–736. https://doi.org/10.1016/0022-2836(74)90047-3.

(47) Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J. Selection of Single-Stranded DNA Molecules That Bind and Inhibit Human Thrombin. *Nature* **1992**, *355* (6360), 564–566. https://doi.org/10.1038/355564a0.

(48) Tasset, D. M.; Kubik, M. F.; Steiner, W. Oligonucleotide Inhibitors of Human Thrombin That Bind Distinct Epitopes. *Journal of Molecular Biology* **1997**, *272* (5), 688–698. https://doi.org/10.1006/jmbi.1997.1275.

(49) Thiel, K. W.; Giangrande, P. H. Therapeutic Applications of DNA and RNA Aptamers. *Oligonucleotides* **2009**, *19* (3), 209–222. https://doi.org/10.1089/oli.2009.0199.

(50) Seelig, B.; Jäschke, A. A Small Catalytic RNA Motif with Diels-Alderase Activity. *Chemistry & Biology* **1999**, *6* (3), 167–176.

(51) Cech, T. The Chemistry of Self-Splicing RNA and RNA Enzymes. *Science* **1987**, *236* (4808), 1532–1539. https://doi.org/10.1126/science.2438771.

(52) Quigley, G. J.; Rich, A. Structural Domains of Transfer RNA Molecules: The Ribose 2′ Hydroxyl Which Distinguishes RNA from DNA Plays a Key Role in Stabilizing TRNA Structure. *Science* **1976**, *194* (4267), 796–806. https://doi.org/10.1126/science.790568.

(53) Silverman, S. K.; Begley, T. P. Wiley Encyclopedia of Chemical Biology. *Wiley Encyclopedia of Chemical Biology* **2008**.

(54) Baum, D. A.; Silverman, S. K. Deoxyribozymes: Useful DNA Catalysts in Vitro and in Vivo. *Cell. Mol. Life Sci.* **2008**, *65* (14), 2156–2174. https://doi.org/10.1007/s00018-008-8029-y.

(55) Breaker, R. R.; Joyce, G. F. A DNA Enzyme That Cleaves RNA. *Chemistry & Biology* **1994**, *1* (4), 223–229. https://doi.org/10.1016/1074-5521(94)90014-0.

(56) Santoro, S. W.; Joyce, G. F. A General Purpose RNA-Cleaving DNA Enzyme. *Proceedings of the National Academy of Sciences* **1997**, *94* (9), 4262–4266. https://doi.org/10.1073/pnas.94.9.4262.

(57) Roth, A.; Breaker, R. R. An Amino Acid as a Cofactor for a Catalytic Polynucleotide. *Proceedings of the National Academy of Sciences* **1998**, *95* (11), 6027–6031. https://doi.org/10.1073/pnas.95.11.6027.

(58) Purtha, W. E.; Coppins, R. L.; Smalley, M. K.; Silverman, S. K. General Deoxyribozyme-Catalyzed Synthesis of Native 3‘−5‘ RNA Linkages. *J. Am. Chem. Soc.* **2005**, *127* (38), 13124–13125. https://doi.org/10.1021/ja0533702.

(59) Carmi, N.; Shultz, L. A.; Breaker, R. R. In Vitro Selection of Self-Cleaving DNAs. *Chemistry & Biology* **1996**, *3* (12), 1039–1046.

(60) Carmi, N.; Balkhi, S. R.; Breaker, R. R. Cleaving DNA with DNA. *Proceedings of the National Academy of Sciences* **1998**, *95* (5), 2233–2237. https://doi.org/10.1073/pnas.95.5.2233.

(61) Gu, H.; Furukawa, K.; Weinberg, Z.; Berenson, D. F.; Breaker, R. R. Small, Highly Active DNAs That Hydrolyze DNA. *J. Am. Chem. Soc.* **2013**, *135* (24), 9121–9129. https://doi.org/10.1021/ja403585e.

(62) Chandra, M.; Sachdeva, A.; Silverman, S. K. DNA-Catalyzed Sequence-Specific Hydrolysis of DNA. *Nat Chem Biol* **2009**, *5* (10), 718–720. https://doi.org/10.1038/nchembio.201.

(63) Li, Y.; Sen, D. A Catalytic DNA for Porphyrin Metallation. *Nature Structural Biology* **1996**, *3* (9), 743–747.

(64) Li, Y.; Sen, D. Toward an Efficient DNAzyme. *Biochemistry* **1997**, *36*, 5589–5599.

(65) Dokukin, V.; Silverman, S. K. A Modular Tyrosine Kinase Deoxyribozyme with Discrete Aptamer and Catalyst Domains. *Chem. Commun.* **2014**, *50* (66), 9317–9320. https://doi.org/10.1039/C4CC04253K.

(66) Lee, Y.; Klauser, P. C.; Brandsen, B. M.; Zhou, C.; Li, X.; Silverman, S. K. DNA-Catalyzed DNA Cleavage by a Radical Pathway with Well-Defined Products. *J. Am. Chem. Soc.* **2017**, *139* (1), 255–261. https://doi.org/10.1021/jacs.6b10274.

(67) Wang, Y.; Silverman, S. K. Deoxyribozymes That Synthesize Branched and Lariat RNA. *J. Am. Chem. Soc.* **2003**, *125* (23), 6880–6881. https://doi.org/10.1021/ja035150z.

(68) Wang, Y.; Silverman, S. K. Efficient One-Step Synthesis of Biologically Related Lariat RNAs by a Deoxyribozyme. *Angew. Chem. Int. Ed.* **2005**, *44* (36), 5863–5866. https://doi.org/10.1002/anie.200501643.

(69) Brandsen, B. M.; Velez, T. E.; Sachdeva, A.; Ibrahim, N. A.; Silverman, S. K. DNA-Catalyzed Lysine Side Chain Modification. *Angew. Chem. Int. Ed.* **2014**, *53* (34), 9045–9050. https://doi.org/10.1002/anie.201404622.

(70) Langer, P. R.; Waldrop, A. A.; Ward, D. C. Enzymatic Synthesis of Biotin-Labeled Polynucleotides: Novel Nucleic Acid Affinity Probes. *Proceedings of the National Academy of Sciences* **1981**, *78* (11), 6633–6637. https://doi.org/10.1073/pnas.78.11.6633.

(71) Praseuth, Dani.; Perrouault, L.; Doan, T. L.; CHASSIGNOLt, M.; THUONGt, N. Sequence-Specific Binding and Photocrosslinking of a and F8 Oligodeoxynucleotides to the Major Groove of DNA via Triple-Helix Formation. *Proc. Natl. Acad. Sci. USA* **1988**, 5.

(72) Welch, M. B.; Burgess, K. Synthesis of Fluorescent, Photolabile 3′- *O* -Protected Nucleoside Triphosphates for the Base Addition Sequencing Scheme. *Nucleosides and Nucleotides* **1999**, *18* (2), 197–201. https://doi.org/10.1080/15257779908043067.

(73) Ju, J.; Kim, D. H.; Bi, L.; Meng, Q.; Bai, X.; Li, Z.; Li, X.; Marma, M. S.; Shi, S.; Wu, J.; Edwards, J. R.; Romu, A.; Turro, N. J. Four-Color DNA Sequencing by Synthesis Using Cleavable Fluorescent Nucleotide Reversible Terminators. *Proceedings of the National Academy of Sciences* **2006**, *103* (52), 19635–19640. https://doi.org/10.1073/pnas.0609513103.

(74) Kim, T.-S.; Kim, D.-R.; Ahn, H.-C.; Shin, D.; Ahn, D.-R. Novel 3′-O-Fluorescently Modified Nucleotides for Reversible Termination of DNA Synthesis. *Chem. Eur. J. of Chem. Bio.* **2009**, *11* (1), 75–78. https://doi.org/10.1002/cbic.200900632.

(75) Hocek, M.; Fojta, M. Nucleobase Modification as Redox DNA Labelling for Electrochemical Detection. *Chem. Soc. Rev.* **2011**, *40* (12), 5802. https://doi.org/10.1039/c1cs15049a.

(76) Takasaki, B. K.; Chin, J. Cleavage of the Phosphate Diester Backbone of DNA with Cerium (III) and Molecular Oxygen. *J. Am. Chem. Soc.* **1994**, *116*, 1121–1122.

(77) Young, M. J.; Wahnon, D.; Hynes, R. C.; Chin, J. Reactivity of Copper(I1) Hydroxides and Copper(I1) Alkoxides for Cleaving an Activated Phosphate Diester. *J. Am. Chem. Soc.* **1995**, *117*, 9441–9447.

(78) Basile, L. A.; Raphael, A. L.; Barton, J. K. Metal-Activated Hydrolytic Cleavage of DNA. *J. Am. Chem. Soc.* **1987**, *109* (24), 7550–7551. https://doi.org/10.1021/ja00258a061.

(79) Schnaith, L. M.; Hanson, R. S.; Que, L. Double-Stranded Cleavage of PBR322 by a Diiron Complex via a “Hydrolytic” Mechanism. *Proceedings of the National Academy of Sciences* **1994**, *91* (2), 569–573. https://doi.org/10.1073/pnas.91.2.569.

(80) Komiyama, M.; Kodama, T.; Takeda, N.; Sumaoka, J.; Shiiba, T.; Matsumoto, Y.; Yashiro, M. Catalytically Active Species for CeCI3-Induced DNA Hydrolysis1. *The Journal of Biochemistry* **1994**, *115* (5), 809–810. https://doi.org/10.1093/oxfordjournals.jbchem.a124419.

(81) Latham, J. A.; Johnson, R.; Toole, J. J. The Application of a Modified Nucleotide in Aptamer Selection: Novel Thrombin Aptamers Containing -(1 -Pentynyl)-2′-Deoxyuridine. *Nucl Acids Res* **1994**, *22* (14), 2817–2822. https://doi.org/10.1093/nar/22.14.2817.

(82) Sakthivel, K.; Barbas III, C. F. Expanding the Potential of DNA for Binding and Catalysis: Highly Functionalized DUTP Derivatives That Are Substrates for Thermostable DNA Polymerases. **1998**, 4.

(83) Gold, L.; Ayers, D.; Bertino, J.; Bock, C.; Bock, A.; Brody, E. N.; Carter, J.; Dalby, A. B.; Eaton, B. E.; Fitzwater, T.; Flather, D.; Forbes, A.; Foreman, T.; Fowler, C.; Gawande, B.; Goss, M.; Gunn, M.; Gupta, S.; Halladay, D.; Heil, J.; Heilig, J.; Hicke, B.; Husar, G.; Janjic, N.; Jarvis, T.; Jennings, S.; Katilius, E.; Keeney, T. R.; Kim, N.; Koch, T. H.; Kraemer, S.; Kroiss, L.; Le, N.; Levine, D.; Lindsey, W.; Lollo, B.; Mayfield, W.; Mehan, M.; Mehler, R.; Nelson, S. K.; Nelson, M.; Nieuwlandt, D.; Nikrad, M.; Ochsner, U.; Ostroff, R. M.; Otis, M.; Parker, T.; Pietrasiewicz, S.; Resnicow, D. I.; Rohloff, J.; Sanders, G.; Sattin, S.; Schneider, D.; Singer, B.; Stanton, M.; Sterkel, A.; Stewart, A.; Stratford, S.; Vaught, J. D.; Vrkljan, M.; Walker, J. J.; Watrobka, M.; Waugh, S.; Weiss, A.; Wilcox, S. K.; Wolfson, A.; Wolk, S. K.; Zhang, C.; Zichi, D. Aptamer-Based Multiplexed Proteomic Technology for Biomarker Discovery. *PLoS ONE* **2010**, *5* (12), e15004. https://doi.org/10.1371/journal.pone.0015004.

(84) Kraemer, S.; Vaught, J. D.; Bock, C.; Gold, L.; Katilius, E.; Keeney, T. R.; Kim, N.; Saccomano, N. A.; Wilcox, S. K.; Zichi, D.; Sanders, G. M. From SOMAmer-Based Biomarker Discovery to Diagnostic and Clinical Applications: A SOMAmer-Based, Streamlined Multiplex Proteomic Assay. *PLoS ONE* **2011**, *6* (10), e26332. https://doi.org/10.1371/journal.pone.0026332.

(85) Gawande, B. N.; Rohloff, J. C.; Carter, J. D.; von Carlowitz, I.; Zhang, C.; Schneider, D. J.; Janjic, N. Selection of DNA Aptamers with Two Modified Bases. *Proc Natl Acad Sci USA* **2017**, *114* (11), 2898–2903. https://doi.org/10.1073/pnas.1615475114.

(86) Jarvis, T. C.; Davies, D. R.; Hisaminato, A.; Resnicow, D. I.; Gupta, S.; Waugh, S. M.; Nagabukuro, A.; Wadatsu, T.; Hishigaki, H.; Gawande, B.; Zhang, C.; Wolk, S. K.; Mayfield, W. S.; Nakaishi, Y.; Burgin, A. B.; Stewart, L. J.; Edwards, T. E.; Gelinas, A. D.; Schneider, D. J.; Janjic, N. Non-Helical DNA Triplex Forms a Unique Aptamer Scaffold for High Affinity Recognition of Nerve Growth Factor. *Structure* **2015**, *23* (7), 1293–1304. https://doi.org/10.1016/j.str.2015.03.027.

(87) Gupta, S.; Hirota, M.; Waugh, S. M.; Murakami, I.; Suzuki, T.; Muraguchi, M.; Shibamori, M.; Ishikawa, Y.; Jarvis, T. C.; Carter, J. D.; Zhang, C.; Gawande, B.; Vrkljan, M.; Janjic, N.; Schneider, D. J. Chemically Modified DNA Aptamers Bind Interleukin-6 with High Affinity and Inhibit Signaling by Blocking Its Interaction with Interleukin-6 Receptor. *Journal of Biological Chemistry* **2014**, *289* (12), 8706–8719. https://doi.org/10.1074/jbc.M113.532580.

(88) Gierlich, J.; Gutsmiedl, K.; Gramlich, P. M. E.; Schmidt, A.; Burley, G. A.; Carell, T. Synthesis of Highly Modified DNA by a Combination of PCR with Alkyne-Bearing Triphosphates and Click Chemistry. *Chem. Eur. J.* **2007**, *13* (34), 9486–9494. https://doi.org/10.1002/chem.200700502.

(89) Pfeiffer, F.; Tolle, F.; Rosenthal, M.; Brändle, G. M.; Ewers, J.; Mayer, G. Identification and Characterization of Nucleobase-Modified Aptamers by Click-SELEX. *Nat Protoc* **2018**, *13* (5), 1153–1180. https://doi.org/10.1038/nprot.2018.023.

(90) Tolle, F.; Brändle, G. M.; Matzner, D.; Mayer, G. A Versatile Approach Towards Nucleobase-Modified Aptamers. *Angew. Chem. Int. Ed.* **2015**, *54* (37), 10971–10974. https://doi.org/10.1002/anie.201503652.

(91) Tolle, F.; Rosenthal, M.; Pfeiffer, F.; Mayer, G. Click Reaction on Solid Phase Enables High Fidelity Synthesis of Nucleobase-Modified DNA. *Bioconjugate Chem.* **2016**, *27* (3), 500–503. https://doi.org/10.1021/acs.bioconjchem.5b00668.

(92) Breaker, R. R.; Joyce, G. F. A DNA Enzyme with Mg2+-Dependent RNA Phosphoesterase Activity. *Chemistry & Biology* **1995**, *2* (10), 655–660. https://doi.org/10.1016/1074-5521(95)90028-4.

(93) Wiegand, T. W.; Janssen, R. C.; Eaton, B. E. Selection of RNA Amide Synthases. *Chemistry & Biology* **1997**, *4* (9), 675–683. https://doi.org/10.1016/S1074-5521(97)90223-4.

(94) Cochrane, J. C.; Strobel, S. A. Catalytic Strategies of Self-Cleaving Ribozymes. *Acc. Chem. Res.* **2008**, *41* (8), 1027–1035. https://doi.org/10.1021/ar800050c.

(95) Bratty, J.; Chartrand, P.; Ferbeyre, G.; Cedergren, R. The Hammerhead RNA Domain, a Model Ribozyme. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1993**, *1216* (3), 345–359. https://doi.org/10.1016/0167-4781(93)90001-T.

(96) Scott, W. G.; Klug, A. Ribozymes: Structure and Mechanism in RNA Catalysis. **1996**, 5.

(97) Silverman, S. K. In Vitro Selection, Characterization, and Application of Deoxyribozymes That Cleave RNA. *Nucleic Acids Research* **2005**, *33* (19), 6151–6163. https://doi.org/10.1093/nar/gki930.

(98) Murphy, E.; Freudenrich, C. C.; Levy, L. A.; London, R. E.; Lieberman, M. Monitoring Cytosolic Free Magnesium in Cultured Chicken Heart Cells by Use of the Fluorescent Indicator Furaptra. *Proceedings of the National Academy of Sciences* **1989**, *86* (8), 2981–2984. https://doi.org/10.1073/pnas.86.8.2981.

(99) Murphy, E.; Steenbergen, C.; Levy, L. A.; Raju, B.; London, R. E. Cytosolic Free Magnesium Levels in Ischemic Rat Heart. *Journal of Biological Chemistry* **1989**, *264* (10), 5622–5627. https://doi.org/10.1016/S0021-9258(18)83593-1.

(100) Mulquiney, K. Free Magnesium-ion Concentration in Erythrocytes by 31P NMR: The Effect of Metabolite–Haemoglobin Interactions. *NMR IN BIOMEDICINE* **1997**, *10*, 9.

(101) Faulhammer, D.; Famulok, M. Characterization and Divalent Metal-Ion Dependence of in Vitro Selected Deoxyribozymes Which Cleave DNA/RNA Chimeric Oligonucleotides. *Journal of Molecular Biology* **1997**, *269* (2), 188–202. https://doi.org/10.1006/jmbi.1997.1036.

(102) Geyer, C. R.; Sen, D. Evidence for the Metal-Cofactor Independence of an RNA Phosphodiester-Cleaving DNA Enzyme. *Chemistry & Biology* **1997**, *4* (8), 579–593. https://doi.org/10.1016/S1074-5521(97)90244-1.

(103) Carrigan, M. A.; Ricardo, A.; Ang, D. N.; Benner, S. A. Quantitative Analysis of a RNA-Cleaving DNA Catalyst Obtained via in Vitro Selection. *Biochemistry* **2004**, *43* (36), 11446–11459. https://doi.org/10.1021/bi049898l.

(104) Torabi, S.-F.; Lu, Y. Identification of the Same Na+-Specific DNAzyme Motif from Two In Vitro Selections Under Different Conditions. *J Mol Evol* **2015**, *81* (5–6), 225–234. https://doi.org/10.1007/s00239-015-9715-7.

(105) Perrin, D. M.; Garestier, T.; Hélène, C. Bridging the Gap between Proteins and Nucleic Acids: A Metal-Independent RNAseA Mimic with Two Protein-Like Functionalities. *J. Am. Chem. Soc.* **2001**, *123* (8), 1556–1563. https://doi.org/10.1021/ja003290s.

(106) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. A Self-Cleaving DNA Enzyme Modified with Amines, Guanidines and Imidazoles Operates Independently of Divalent Metal

Cations (M 2+). *Nucleic Acids Research* **2009**, *37* (5), 1638–1649. https://doi.org/10.1093/nar/gkn1070.

(107) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. A DNAzyme with Three Protein-Like Functional Groups: Enhancing Catalytic Efficiency of M^{2+} -Independent RNA Cleavage. *ChemBioChem* **2009**, *10* (12), 1988–1992. https://doi.org/10.1002/cbic.200900314.

(108) Hollenstein, M.; Hipolito, C.; Lam, C.; Dietrich, D.; Perrin, D. M. A Highly Selective DNAzyme Sensor for Mercuric Ions. *Angewandte Chemie International Edition* **2008**, *47* (23), 4346–4350. https://doi.org/10.1002/anie.200800960.

(109) Wang, Y.; Liu, E.; Lam, C. H.; Perrin, D. M. A Densely Modified M^{2+} -Independent DNAzyme That Cleaves RNA Efficiently with Multiple Catalytic Turnover. *Chem. Sci.* **2018**, *9* (7), 1813–1821. https://doi.org/10.1039/C7SC04491G.

(110) Wang, Y.; Ng, N.; Liu, E.; Lam, C. H.; Perrin, D. M. Systematic Study of Constraints Imposed by Modified Nucleoside Triphosphates with Protein-like Side Chains for Use in in Vitro Selection. *Org. Biomol. Chem.* **2017**, *15* (3), 610–618. https://doi.org/10.1039/C6OB02335E.

(111) Renders, M.; Miller, E.; Lam, C. H.; Perrin, D. M. Whole Cell-SELEX of Aptamers with a Tyrosine-like Side Chain against Live Bacteria. *Org. Biomol. Chem.* **2017**, *15* (9), 1980–1989. https://doi.org/10.1039/C6OB02451C.

(112) Li, Y.; Sen, D. Toward an Efficient DNAzyme. **1997**, *36*, 5589–5599.

(113) Tarasow, T. M.; Tarasow, S. L.; Eaton, B. E. RNA-Catalysed Carbon–Carbon Bond Formation. *Nature* **1997**, *389* (6646), 54–57. https://doi.org/10.1038/37950.

(114) Ge, C.; Luo, Q.; Wang, D.; Zhao, S.; Liang, X.; Yu, L.; Xing, X.; Zeng, L. Colorimetric Detection of Copper(II) Ion Using Click Chemistry and Hemin/G-Quadruplex Horseradish

Peroxidase-Mimicking DNAzyme. *Anal. Chem.* **2014**, *86* (13), 6387–6392. https://doi.org/10.1021/ac501739a.

(115) Li, D.; Zhou, W.; Chai, Y.; Yuan, R.; Xiang, Y. Click Chemistry-Mediated Catalytic Hairpin Self-Assembly for Amplified and Sensitive Fluorescence Detection of Cu^{2+} in Human Serum. *Chem. Commun.* **2015**, *51* (63), 12637–12640. https://doi.org/10.1039/C5CC04218F.

(116) Winz, M.-L.; Linder, E. C.; André, T.; Becker, J.; Jäschke, A. Nucleotidyl Transferase Assisted DNA Labeling with Different Click Chemistries. *Nucleic Acids Res* **2015**, *43* (17), e110–e110. https://doi.org/10.1093/nar/gkv544.

(117) Chandrasekar, J.; Silverman, S. K. Catalytic DNA with Phosphatase Activity. *Proceedings of the National Academy of Sciences* **2013**, *110* (14), 5315–5320. https://doi.org/10.1073/pnas.1221946110.

(118) Chandrasekar, J.; Wylder, A. C.; Silverman, S. K. Phosphoserine Lyase Deoxyribozymes: DNA-Catalyzed Formation of Dehydroalanine Residues in Peptides. *J. Am. Chem. Soc.* **2015**, *137* (30), 9575–9578. https://doi.org/10.1021/jacs.5b06308.

(119) Walsh, S. M.; Sachdeva, A.; Silverman, S. K. DNA Catalysts with Tyrosine Kinase Activity. *J. Am. Chem. Soc.* **2013**, *135* (40), 14928–14931. https://doi.org/10.1021/ja407586u.

(120) Fusz, S.; Eisenführ, A.; Srivatsan, S. G.; Heckel, A.; Famulok, M. A Ribozyme for the Aldol Reaction. *Chemistry & Biology* **2005**, *12* (8), 941–950. https://doi.org/10.1016/j.chembiol.2005.06.008.

(121) Doggrell, S. A. Pegaptanib: The First Antiangiogenic Agent Approved for Neovascular Macular Degeneration. *null* **2005**, *6* (8), 1421–1423. https://doi.org/10.1517/14656566.6.8.1421.

(122) Gold, L.; Ayers, D.; Bertino, J.; Bock, C.; Bock, A.; Brody, E. N.; Carter, J.; Dalby, A. B.; Eaton, B. E.; Fitzwater, T.; Flather, D.; Forbes, A.; Foreman, T.; Fowler, C.; Gawande, B.;

Goss, M.; Gunn, M.; Gupta, S.; Halladay, D.; Heil, J.; Heilig, J.; Hicke, B.; Husar, G.; Janjic, N.; Jarvis, T.; Jennings, S.; Katilius, E.; Keeney, T. R.; Kim, N.; Koch, T. H.; Kraemer, S.; Kroiss, L.; Le, N.; Levine, D.; Lindsey, W.; Lollo, B.; Mayfield, W.; Mehan, M.; Mehler, R.; Nelson, S. K.; Nelson, M.; Nieuwlandt, D.; Nikrad, M.; Ochsner, U.; Ostroff, R. M.; Otis, M.; Parker, T.; Pietrasiewicz, S.; Resnicow, D. I.; Rohloff, J.; Sanders, G.; Sattin, S.; Schneider, D.; Singer, B.; Stanton, M.; Sterkel, A.; Stewart, A.; Stratford, S.; Vaught, J. D.; Vrkljan, M.; Walker, J. J.; Watrobka, M.; Waugh, S.; Weiss, A.; Wilcox, S. K.; Wolfson, A.; Wolk, S. K.; Zhang, C.; Zichi, D. Aptamer-Based Multiplexed Proteomic Technology for Biomarker Discovery. *PLoS ONE* **2010**, *5* (12), e15004. https://doi.org/10.1371/journal.pone.0015004.

(123) Huang, P.-J. J.; Liu, J. Sensing Parts-per-Trillion Cd^{2+}, Hg^{2+}, and Pb^{2+} Collectively and Individually Using Phosphorothioate DNAzymes. *Anal. Chem.* **2014**, *86* (12), 5999–6005. https://doi.org/10.1021/ac501070a.

(124) Huang, P.-J. J.; Vazin, M.; Liu, J. *In Vitro* Selection of a DNAzyme Cooperatively Binding Two Lanthanide Ions for RNA Cleavage. *Biochemistry* **2016**, *55* (17), 2518–2525. https://doi.org/10.1021/acs.biochem.6b00132.

(125) Liu, J.; Brown, A. K.; Meng, X.; Cropek, D. M.; Istok, J. D.; Watson, D. B.; Lu, Y. A Catalytic Beacon Sensor for Uranium with Parts-per-Trillion Sensitivity and Millionfold Selectivity. *Proceedings of the National Academy of Sciences* **2007**, *104* (7), 2056–2061. https://doi.org/10.1073/pnas.0607875104.

(126) Clever, G. H.; Kaul, C.; Carell, T. DNA-Metal Base Pairs. *Angew. Chem. Int. Ed.* **2007**, *46* (33), 6226–6236. https://doi.org/10.1002/anie.200701185.

(127) Lake, R. J.; Yang, Z.; Zhang, J.; Lu, Y. DNAzymes as Activity-Based Sensors for Metal Ions: Recent Applications, Demonstrated Advantages, Current Challenges, and Future

Directions. *Acc. Chem. Res.* **2019**, *52* (12), 3275–3286. https://doi.org/10.1021/acs.accounts.9b00419.

(128) Brown, A. K.; Li, J.; Pavot, C. M.-B.; Lu, Y. A Lead-Dependent DNAzyme with a Two-Step Mechanism †. *Biochemistry* **2003**, *42* (23), 7152–7161. https://doi.org/10.1021/bi027332w.

(129) Kim, H.-K.; Rasnik, I.; Liu, J.; Ha, T.; Lu, Y. Dissecting Metal Ion–Dependent Folding and Catalysis of a Single DNAzyme. *Nat Chem Biol* **2007**, *3* (12), 763–768. https://doi.org/10.1038/nchembio.2007.45.

(130) Cui, L.; Peng, R.; Fu, T.; Zhang, X.; Wu, C.; Chen, H.; Liang, H.; Yang, C. J.; Tan, W. Biostable L-DNAzyme for Sensing of Metal Ions in Biological Systems. *Anal. Chem.* **2016**, *88* (3), 1850–1855. https://doi.org/10.1021/acs.analchem.5b04170.

(131) Zhou, W.; Zhang, Y.; Ding, J.; Liu, J. In Vitro Selection in Serum: RNA-Cleaving DNAzymes for Measuring Ca^{2+} and Mg^{2+}. *ACS Sens.* **2016**, *1* (5), 600–606. https://doi.org/10.1021/acssensors.5b00306.

(132) Zhang, W.; Feng, Q.; Chang, D.; Tram, K.; Li, Y. In Vitro Selection of RNA-Cleaving DNAzymes for Bacterial Detection. *Methods* **2016**, *106*, 66–75. https://doi.org/10.1016/j.ymeth.2016.03.018.

(133) Huang, P. J.; Rochambeau, D.; Sleiman, H. F.; Liu, J. Target Self-Enhanced Selectivity in Metal-Specific DNAzymes. *Angew. Chem. Int. Ed.* **2020**, *59* (9), 3573–3577. https://doi.org/10.1002/anie.201915675.

(134) Trepanier, J.; Tanner, J. E.; Momparler, R. L.; Le, O. N. L.; Alvarez, F.; Alfieri, C. Cleavage of Intracellular Hepatitis C RNA in the Virus Core Protein Coding Region by Deoxyribozymes. *J Viral Hepat* **2006**, *13* (2), 131–138. https://doi.org/10.1111/j.1365-2893.2005.00684.x.

(135) Campbell, R. L.; Petsko, G. A. Ribonuclease Structure and Catalysis: Crystal Structure of Sulfate-Free Native Ribonuclease A at 1.5- .ANG. Resolution. *Biochemistry* **1987**, *26* (26), 8579–8584. https://doi.org/10.1021/bi00400a013.

(136) Nishikawa, S.; Morioka, H.; Kim, H. J.; Fuchimura, K.; Tanaka, T.; Uesugi, S.; Hakoshima, T.; Tomita, K.; Ohtsuka, E.; Ikehara, M. Two Histidine Residues Are Essential for Ribonuclease T1 Activity as Is the Case for Ribonuclease A. *Biochemistry* **1987**, *26* (26), 8620–8624. https://doi.org/10.1021/bi00400a019.

(137) Hovinen, J.; Guzaev, A.; Azhayeva, E.; Azhayev, A.; Lonnberg, H. Imidazole Tethered Oligodeoxyribonucleotides: Synthesis and RNA Cleaving Activity. *J. Org. Chem.* **1995**, *60* (7), 2205–2209. https://doi.org/10.1021/jo00112a047.

(138) Roth, A.; Breaker, R. R. An Amino Acid as a Cofactor for a Catalytic Polynucleotide. *Proceedings of the National Academy of Sciences* **1998**, *95* (11), 6027–6031. https://doi.org/10.1073/pnas.95.11.6027.

(139) Lermer, L.; Roupioz, Y.; Ting, R.; Perrin, D. M. Toward an RNaseA Mimic: A DNAzyme with Imidazoles and Cationic Amines. *J. Am. Chem. Soc.* **2002**, *124* (34), 9960–9961. https://doi.org/10.1021/ja0205075.

(140) Lam, C. H.; Hipolito, C. J.; Hollenstein, M.; Perrin, D. M. A Divalent Metal-Dependent Self-Cleaving DNAzyme with a Tyrosine Side Chain. *Org. Biomol. Chem.* **2011**, *9* (20), 6949. https://doi.org/10.1039/c1ob05359k.

(141) Wang, Y.; Liu, E.; Lam, C. H.; Perrin, D. M. A Densely Modified M $^{2+}$ -Independent DNAzyme That Cleaves RNA Efficiently with Multiple Catalytic Turnover. *Chem. Sci.* **2018**, *9* (7), 1813–1821. https://doi.org/10.1039/C7SC04491G.

(142) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. Toward the Combinatorial Selection of Chemically Modified DNAzyme RNase A Mimics Active Against All-RNA Substrates. *ACS Comb. Sci.* **2013**, *15* (4), 174–182. https://doi.org/10.1021/co3001378.

(143) Lam, C.; Hipolito, C.; Perrin, D. M. Synthesis and Enzymatic Incorporation of Modified Deoxyadenosine Triphosphates. *Eur. J. Org. Chem.* **2008**, *2008* (29), 4915–4923. https://doi.org/10.1002/ejoc.200800381.

(144) Little, J. W. Lambda Exonuclease. *Gene Amplif. Anal.* **1981**, No. 2, 135–145.

(145) Ting, R.; Thomas, J. M.; Perrin, D. M. Kinetics Characterization of a Cis- and Trans-Acting M2+-Independent DNAzyme That Depends on Synthetic RNase A-like Functionality - Burst-Phase Kinetics from the Coalescence of Two Active DNAzyme Folds. *Can. J. Chem* **2007**, *85*, 313–329.

(146) Chiuman, W.; Li, Y. Evolution of High-Branching Deoxyribozymes from a Catalytic DNA with a Three-Way Junction. *Chemistry & Biology* **2006**, *13* (10), 1061–1069. https://doi.org/10.1016/j.chembiol.2006.08.009.

(147) Schlosser, K.; Lam, J. C. F.; Li, Y. A Genotype-to-Phenotype Map of in Vitro Selected RNA-Cleaving DNAzymes: Implications for Accessing the Target Phenotype. *Nucleic Acids Research* **2009**, *37* (11), 3545–3557. https://doi.org/10.1093/nar/gkp222.

(148) Roychowdhury-Saha, M. Extraordinary Rates of Transition Metal Ion-Mediated Ribozyme Catalysis. *RNA* **2006**, *12* (10), 1846–1852. https://doi.org/10.1261/rna.128906.

(149) Han, J.; Burke, J. M. Model for General Acid−Base Catalysis by the Hammerhead Ribozyme: PH−Activity Relationships of G8 and G12 Variants at the Putative Active Site. *Biochemistry* **2005**, *44* (21), 7864–7870. https://doi.org/10.1021/bi047941z.

(150) Nesbitt, S.; Hegg, L. A.; Fedor, M. J. An Unusual PH-Independent and Metal-Ion-Independent Mechanism for Hairpin Ribozyme Catalysis. *Chemistry & Biology* **1997**, *4* (8), 619–630. https://doi.org/10.1016/S1074-5521(97)90247-7.

(151) Kath-Schorr, S.; Wilson, T. J.; Li, N.-S.; Lu, J.; Piccirilli, J. A.; Lilley, D. M. J. General Acid–Base Catalysis Mediated by Nucleobases in the Hairpin Ribozyme. *J. Am. Chem. Soc.* **2012**, *134* (40), 16717–16724. https://doi.org/10.1021/ja3067429.

(152) Williamson, J. R.; Raghuraman, M. K.; Cech, T. R. Monovalent Cation-Induced Structure of Telomeric DNA: The G-Quartet Model. *Cell* **1989**, *59*, 871–880.

(153) Williamson, J. R. G-Quartet Structures in Telomeric DNA. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 703–730.

(154) Basu, S.; Szewczak, A. A.; Cocco, M.; Strobel, S. A. Direct Detection of Monovalent Metal Ion Binding to a DNA G-Quartet by ^{205}Tl NMR. *J. Am. Chem. Soc.* **2000**, *122* (13), 3240–3241. https://doi.org/10.1021/ja993614g.

(155) Sidorov, A. V. Sequence-Specific Cleavage of RNA in the Absence of Divalent Metal Ions by a DNAzyme Incorporating Imidazolyl and Amino Functionalities. *Nucleic Acids Research* **2004**, *32* (4), 1591–1601. https://doi.org/10.1093/nar/gkh326.

(156) Santoro, S. W.; Joyce, G. F.; Sakthivel, K.; Gramatikova, S.; Barbas, C. F. RNA Cleavage by a DNA Enzyme with Extended Chemical Functionality. *J. Am. Chem. Soc.* **2000**, *122* (11), 2433–2439. https://doi.org/10.1021/ja993688s.

(157) Zinnen, S. P.; Domenico, K.; Wilson, M.; Dickinson, B. A.; Beaudry, A.; Mokler, V.; Daniher, A. T.; Burgin, A.; Beigelman, L. Selection, Design, and Characterization of a New Potentially Therapeutic Ribozyme. *RNA* **2002**, *8* (2), 214–228. https://doi.org/10.1017/S1355838202014723.

(158) Danneberg, F.; Westemeier, H.; Horx, P.; Zellmann, F.; Dörr, K.; Kalden, E.; Zeiger, M.; Akpinar, A.; Berger, R.; Göbel, M. W. RNA Hydrolysis by Heterocyclic Amidines and Guanidines: Parameters Affecting Reactivity. *Eur. J. Org. Chem.* **2021**, ejoc.202100950. https://doi.org/10.1002/ejoc.202100950.

(159) Ullrich, S.; Nazir, Z.; Büsing, A.; Scheffer, U.; Wirth, D.; Bats, J. W.; Dürner, G.; Göbel, M. W. Cleavage of Phosphodiesters and of DNA by a Bis(Guanidinium)Naphthol Acting as a Metal-Free Anion Receptor. *ChemBioChem* **2011**, *12* (8), 1223–1229. https://doi.org/10.1002/cbic.201100022.

(160) Göbel, M. W.; Bats, J. W.; Dürner, G. En Route to Synthetic Phosphodiesterases: Supramolecular Phosphoryl-Transfer Mediated by Amidinium–Phosphate Contact Ion-Pairs. *Angew. Chem. Int. Ed. Engl.* **1992**, *31* (2), 207–209. https://doi.org/10.1002/anie.199202071.

(161) Göbel, M. W.; Roussev, C. D.; Scheffer, U. RNA Cleavage Catalyzed by Amphoteric Bis(Acyl)Guanidinium Derivatives. *HCA* **2014**, *97* (2), 215–227. https://doi.org/10.1002/hlca.201300308.

(162) Scheffer, U.; Strick, A.; Ludwig, V.; Peter, S.; Kalden, E.; Göbel, M. W. Metal-Free Catalysts for the Hydrolysis of RNA Derived from Guanidines, 2-Aminopyridines, and 2-Aminobenzimidazoles. *J. Am. Chem. Soc.* **2005**, *127* (7), 2211–2217. https://doi.org/10.1021/ja0443934.

(163) Fouace, S. Polyamine Derivatives as Selective RNaseA Mimics. *Nucleic Acids Research* **2004**, *32* (1), 151–157. https://doi.org/10.1093/nar/gkh157.

(164) Roig, V.; Asseline, U. Oligo-2‘-Deoxyribonucleotides Containing Uracil Modified at the 5-Position with Linkers Ending with Guanidinium Groups. *J. Am. Chem. Soc.* **2003**, *125* (15), 4416–4417. https://doi.org/10.1021/ja029467v.

(165) Huang, P. J.; Liu, J. In Vitro Selection of Chemically Modified DNAzymes. *ChemistryOpen* **2020**, *9* (10), 1046–1059. https://doi.org/10.1002/open.202000134.

(166) Fitch, C. A.; Platzer, G.; Okon, M.; Garcia-Moreno E., B.; McIntosh, L. P. Arginine: Its p K_a Value Revisited: P K_a Value of Arginine. *Protein Science* **2015**, *24* (5), 752–761. https://doi.org/10.1002/pro.2647.

(167) Harris, T. K.; Turner, G. J. Structural Basis of Perturbed PKa Values of Catalytic Groups in Enzyme Active Sites. *IUBMB Life (International Union of Biochemistry and Molecular Biology: Life)* **2002**, *53* (2), 85–98. https://doi.org/10.1080/15216540211468.

(168) Isom, D. G.; Castaneda, C. A.; Cannon, B. R.; Garcia-Moreno E., B. Large Shifts in PKa Values of Lysine Residues Buried inside a Protein. *Proceedings of the National Academy of Sciences* **2011**, *108* (13), 5260–5265. https://doi.org/10.1073/pnas.1010750108.

(169) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. A DNAzyme with Three Protein-Like Functional Groups: Enhancing Catalytic Efficiency of M^{2+} -Independent RNA Cleavage. *ChemBioChem* **2009**, *10* (12), 1988–1992. https://doi.org/10.1002/cbic.200900314.

(170) Lam, C. H.; Hipolito, C. J.; Hollenstein, M.; Perrin, D. M. A Divalent Metal-Dependent Self-Cleaving DNAzyme with a Tyrosine Side Chain. *Org. Biomol. Chem.* **2011**, *9* (20), 6949. https://doi.org/10.1039/c1ob05359k.

(171) Zhou, W.; Zhang, Y.; Ding, J.; Liu, J. In Vitro Selection in Serum: RNA-Cleaving DNAzymes for Measuring Ca^{2+} and Mg^{2+}. *ACS Sens.* **2016**, *1* (5), 600–606. https://doi.org/10.1021/acssensors.5b00306.

(172) Uhlenbeck, O. C. Less Isn't Always More. *RNA* **2003**, *9* (12), 1415–1417. https://doi.org/10.1261/rna.5155903.

Appendices

Appendix A Sequence Information

A.1 Trimmed NGS Data

Seq Name	DNAzyme clones
Sequence1 (Dz11-1)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**CAGTATCGCAGGGAGCACGCTTCACTAAGCACACCTGCGAAAA**GCACGCCCCGCGAGCTC
Sequence2 (Dz11-2)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACCGCCAATAAGCAGGTTCGACCAACAGTAGCCATAACTTGTG**GCACGCCCCGCGAGCTC
Sequence4 (Dz11-4)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**GCGGCGAAGTTCAGTCTGACGCACAACAGTAGTCATAGCTTGAG**GCACGCCCCGCGAGCTC
Sequence5 (Dz11-5)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAAGCGGTGGAGCAAGCGATAACGCGATAGACTCGAAGTGGAA**GCACGCCCCGCGAGCTC
Sequence6 (Dz11-6)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAGTAGTCATATTTGCGGCCAAATACAGCTTGGCCACCTTGAG**GCACGCCCCGCGAGCTC
Sequence7 (Dz11-7)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAAGCGGTGGAGCAAGCGACAACGCGATAGACTCGAAGTGGAA**GCACGCCCCGCGAGCTC
Sequence9 (Dz11-9)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCTACCA**ATAGCCCCCAAGGTACTGGACAGGCAAGTATATCAACGCATGCG**GCACGCCCCGCGAGCTC
Sequence10 (Dz11-10)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ATAGCCCCTAAGGTACTGGCCTGACACAGATGCGTGGAAGATTC**GCACGCCCCGCGAGCTC
Sequence11 (Dz11-11)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACGACCAATAAGCAAGTTCGACCAACAGTAGCCATAACTTGTG**GCACGCCCCGCGAGCTC
Sequence12 (Dz11-12)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAGTGTTGGATATACCAGCTGCAGTCAGTTACGGCACGTCCTG**GCACGCCCCGCGAGCTC
Sequence14 (Dz11-14)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAGTAGTCATATTTGCGGCCGAGTACAGCTTGGCCACCTTGAG**GCACGCCCCGCGAGCTC
Sequence15 (Dz11-15)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAAGCGGTGGACCAAGCGACAACGCGATAGACTCGAAGTGGAA**GCACGCCCCGCGAGCT

Sequence16 (Dz11-16)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**GCAACGAAGTTCAGTCTGACGCACAACGGTAGTCATAGCTTGTG**GCACGCCCCGCGAGCTC
Sequence18 (Dz11-18)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAAGCGGTGGAGCAAGCGACAACGCGGTAGACTCGAAGTGGAA**GCACGCCCCGCGAGCTC
Sequence19 (Dz11-19)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**GCAGCGAAGTTCAGTCTGACGCACAACGGTAGTCATAGCTTGTG**GCACGCCCCGCGAGCTC
Sequence20 (Dz11-20)	ACGACACAGAGCGTGCCGTCTGTTGGGCCCTACCA**ACGAGCGGTGGAACAAGCGACAACGCGATAGACTCGAAGTGGAA**GCACGCCCCGCGAGCTC
Sequence21 (Dz11-21)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAGCCCCTAAGGTACTGGCCACGTTTACAATGGTGACGCATTT**GCACGCCCCGCGAGCTC
Sequence23 (Dz11-23)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAGCATTTGAAATAGCTGATGTGATAGAACTGATATGTCTGTC**GCACGCCCCGCGAGCTC
Sequence25 (Dz11-25)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ATAGCCCCTAAGGTACTGGGTATCGAACGGAGTTACCATACCA**GCACGCCCCGCGAGCTC
Sequence28 (Dz11-28)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAGCCCCCAAGGTACTGGCCACGTTCAAAATGGTGATGCATTT**GCACGCCCCGCGAGCTC
Sequence29 (Dz11-29)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAACGAAGTTCAGTCTGACGCACAACAGTAGTCATAGCTTGTG**GCACGCCCCGCGAGCTC
Sequence30 (Dz11-30)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAGCCCCTAAGGTACTGGTCTGACACAGATATGAGGAAGATCT**GCACGCCCCGCGAGCTC
Sequence31 (Dz11-31)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAAGCGGTGGAACAAGCGACAACGCGATAGACTCGAAGTGGAA**GCACGCCCCGCGAGCTC
Sequence33 (Dz11-33)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAGCATTTGAAATAGCTGATGTGATAGAAACGATATGTCTGTC**GCACGCCCCGCGAGCTC
Sequence34 (Dz11-34)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ATAGCCCCTAAGGTACTGGGTATGGGATGGAGTTATCATACCA**GCACGCCCCGCGAGCTC

Sequence35 (Dz11-35)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACAACGAAGTTCACTCTGACGCACAACAGTA GTCATAGCTTGAG**GCACGCCCCGCGAGCTC
Sequence36 (Dz11-36)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ATAGTACATATCCTCGAATCACCAAACGTCA TTCGTTCGGTAA**GCACGCCCCGCGAGCTC
Sequence39 (Dz11-39)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**ACTGCGAAGTTCAGTCTGAAGCACAACAGTA GTCATAGCTTGAG**GCACGCCCCGCGAGCTC
Sequence40 (Dz11-40)	ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA**GCAGCCCCTAAGGTACTGGCCTGACACAGAT ACGTGGAAGATCC**GCACGCCCCGCGAGCTC
Sequence41 (Dz11-41)	ACGACACAGAGCGTGCCCGTCTGTTGGCCCTACCA**ATAGCCCCTAAGGTACTGGGTATGGGATAGAG TTACCATACCA**GCACGCCCCGCGAGCTC

A.2 Sequence Ancestry

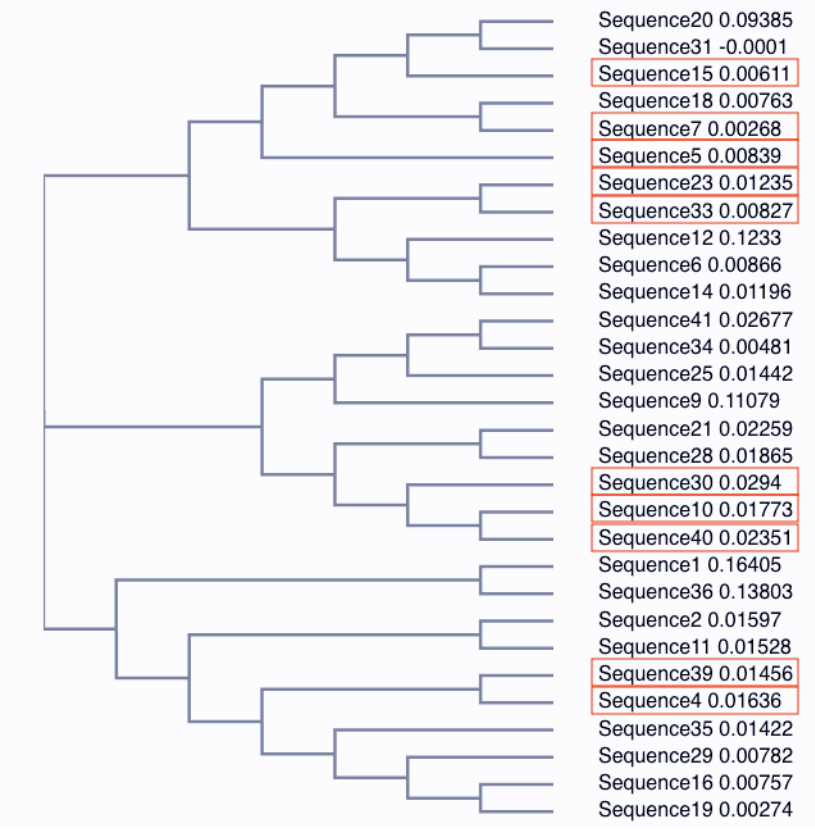


Figure A1.1 Ancestry information of the trimmed sequences obtained using ClustalW.

Appendix B Optimization and unedited gels

B.1 Primer extension optimization

Sequences:
Template
5′
TGCGTGC**AA**GCTGCGTGC**AA**GCTCTGTTG**A**CGG**A**CG**A**GTG**A**CG**AA**C**A**CG**A**C**A**GTT**A**G**A**C**A**GCAC
GCCGCGCGAGCGC 3′

Primer
5′ /56-FAM/GCGCTCGCGCGGCGTGC 3′

The **bold** A represent the regions complementary to the GadUTP.

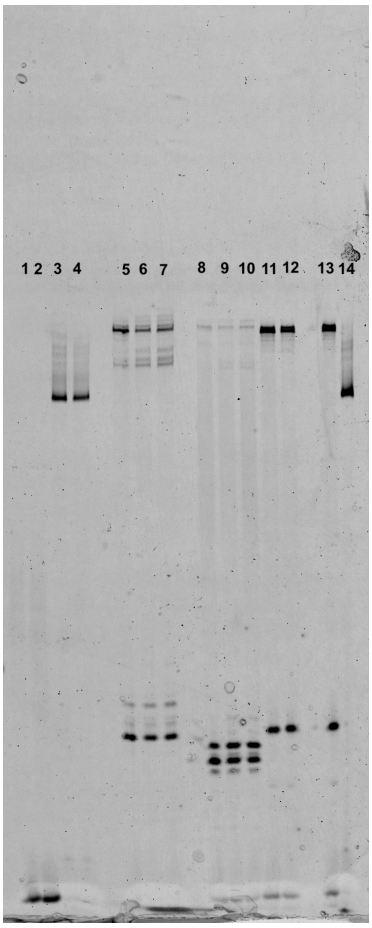


Figure B1.2 Modified nucleotide incorporation assay.
It was run in 1X TBE at 35W on a 10% PAGE. Lane 1-2 Negative control, Lanes 3,4,14 Positive controls, Lanes 5-7=dU^{ga}TP, dGTP and dCTP, Lanes 8-10=dC^{aa}TP, dTTP, dGTP and Lanes 11-13=dU^{ga}TP, dC^{aa}TP, dGTP. Positive control was extended with dU^{ga}TP, dC^{aa}TP, dGTP, dATP to show the full length extension.

B.2 Unedited gel data

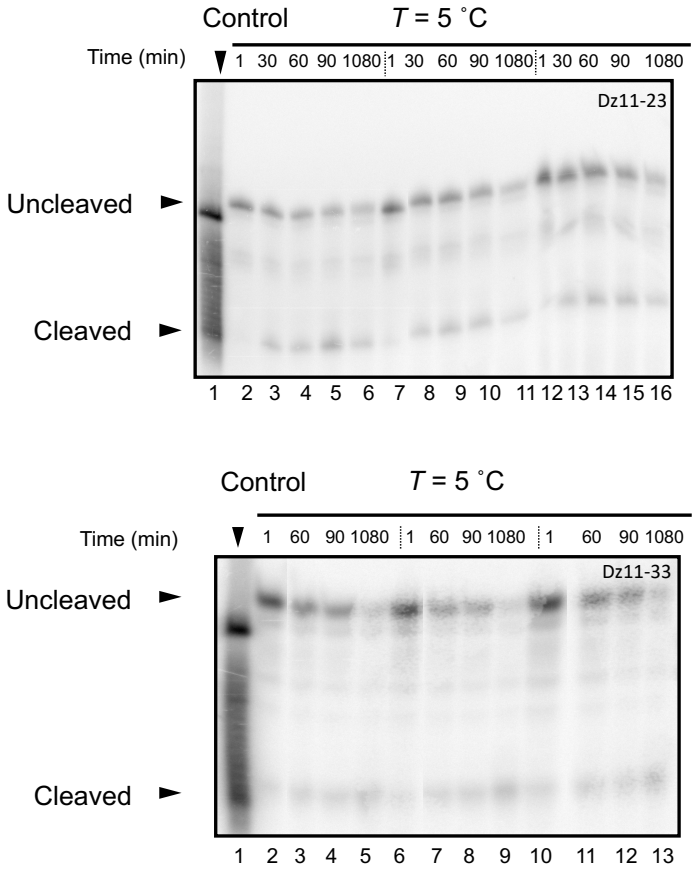


Figure B2.3 Quantification of cleavage activity of Dz11-23 and Dz11-33 at 5ºC.

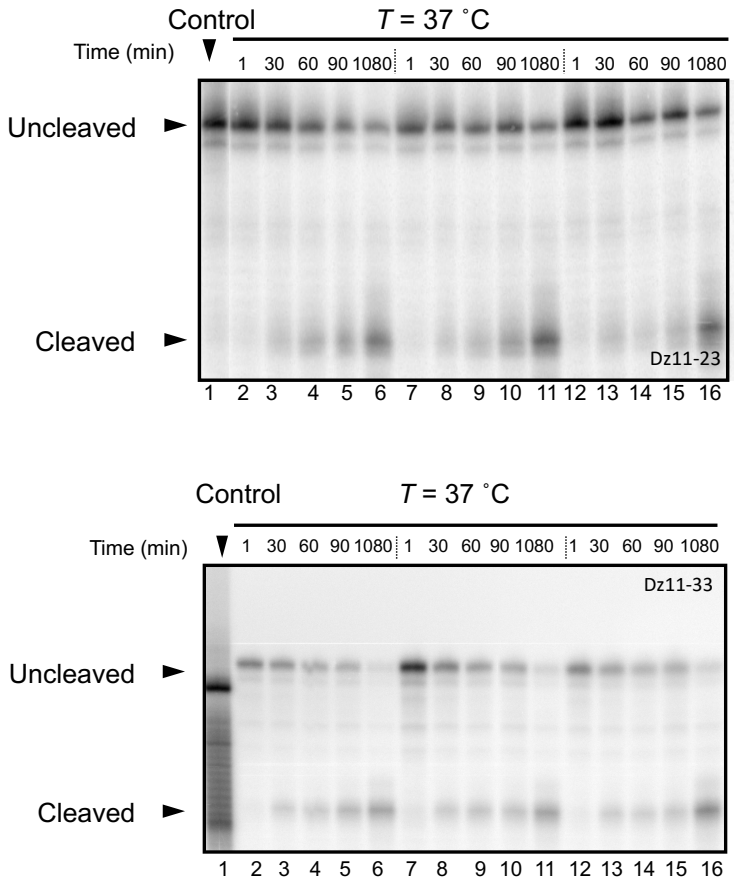


Figure B3.4 Quantification of cleavage activity of Dz11-23 and Dz11-33 at 37ºC.

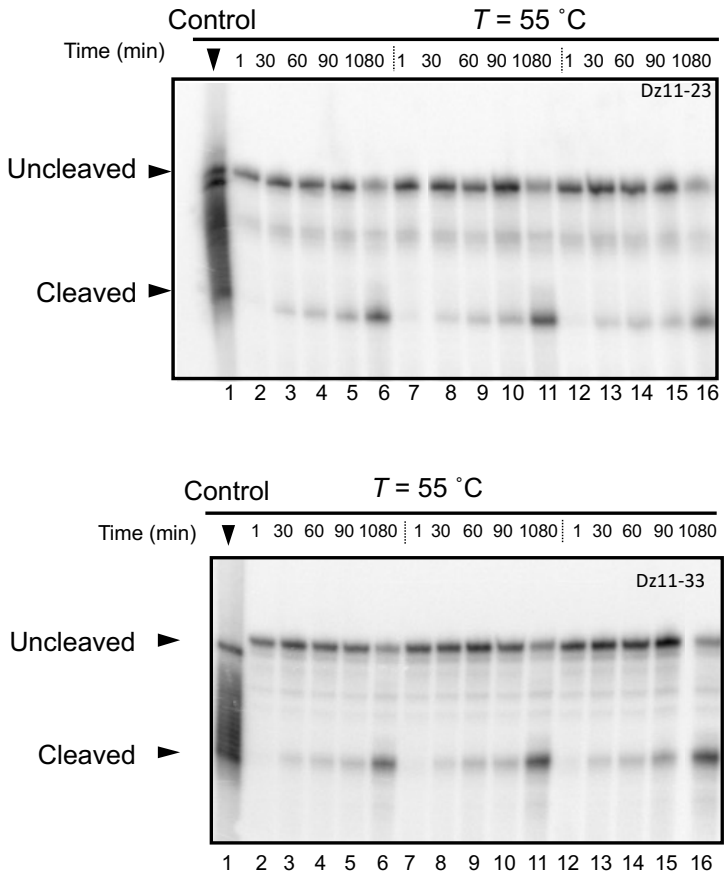


Figure B4.5 Quantification of the cleavage activity of Dz11-23 and Dz11-33 at 55ºC.

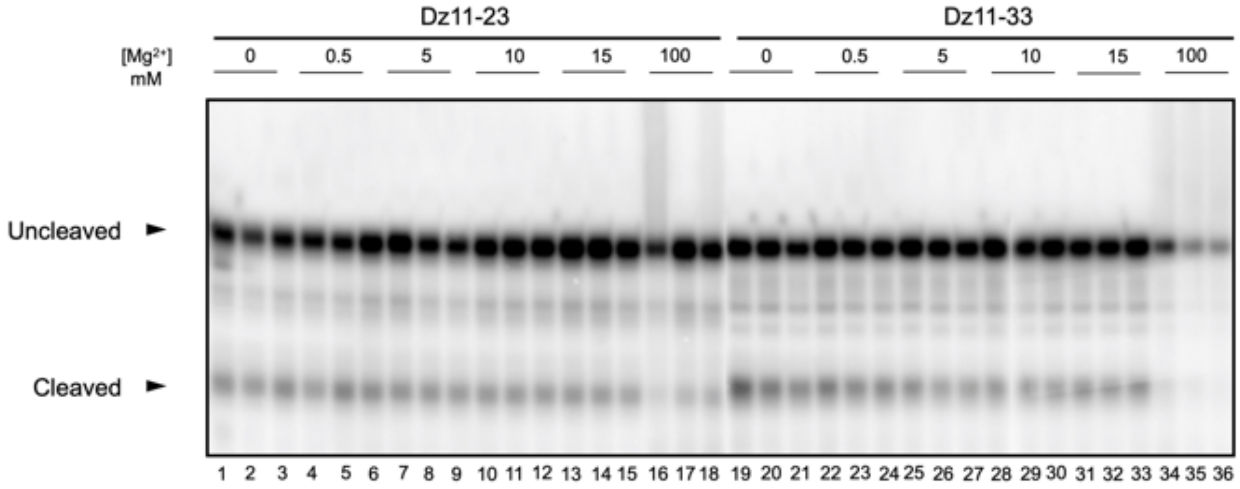


Figure B5.6 Magnesium cleavage profile observed between 0 mM and 100 mM.
Dz11-23 and Dz11-33 were incubated for 30 minutes in separate pH 7.5 cacodylate (25 mM) buffers each containing varying [Mg2+] - 0 mM, 0.5 mM, 5 mM, 10 mM, 15 mM and 100 mM (n = 3)