DNA-CATALYZED COVALENT MODIFICATION OF AMINO ACID SIDE CHAINS IN TETHERED AND FREE PEPTIDE SUBSTRATES

BY

ON YI WONG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Professor Scott K. Silverman, Chair
Assistant Professor Ryan C. Bailey
Professor Paul J. Hergenrother
Professor Susan A. Martinis

Abstract

Scientists thought of RNA as passive messenger molecules until the discovery of catalytic RNA (ribozymes) in the 1980s, which revolutionized thinking toward the multifaceted functions of RNA. The discovery of catalytic RNA also sparked an intense interest in exploring the catalytic functions of nucleic acids. As a close analog of RNA, the catalytic ability of DNA was evaluated and the first artificial DNA catalyst (deoxyribozyme) discovered in 1994. Since then, deoxyribozymes have been shown to catalyze a wide variety of chemical reactions, including phosphodiester bond cleavage, nucleic acid ligation, thymine dimer photoreversion, and the Diels-Alder reaction. However, the majority of deoxyribozyme-catalyzed reactions involve oligonucleotide substrates. To expand the scope of substrates that can be catalyzed by deoxyribozymes, the main theme of the thesis is to seek DNA catalysts that covalently modify the side chains of amino acids.

The thesis describes the numerous selection efforts that led to the first discovery of a particular deoxyribozyme to catalyze the reactivity of a free tripeptide. This unprecedented discovery establishes a future potential use of DNA to modify proteins. In addition, this thesis presents the unexpected discovery of DNA-catalyzed reductive amination that may be applicable to site-specific labeling of nucleic acids. Finally, the thesis summarizes the optimization of the incubation conditions using click chemistry for direct free peptide selection. From these studies, we gained insights on how to direct our future selection efforts to identify DNA-catalyzed reactivity of amino acid side chains.

Acknowledgements

I would to like to thank my research advisor, Professor Scott K. Silverman, for his support and guidance in research. Scott has been a great mentor and I am very grateful for the freedom and trust that Scott has given to his graduate students in conducting research. It has been a great pleasure to work in the Silverman laboratory.

I am grateful for the help and support from Dr. P. I. Pradeepkumar and Dr. Claudia Höbartner during the beginning of my graduate career. Their enthusiasms toward research were very inspiring. The daily discussions with Dr. Madhavaiah Chandra and Dr. Amit Sachdeva were also crucial to drive the research forward.

I am also grateful to Professor Anne M. Baranger and my thesis committee members: Professor Ryan C. Bailey, Professor Paul J. Hergenrother, Professor Susan A. Martinis, and Professor Scott K. Silverman. All have offered outstanding support and advice during my graduate career.

I thank my parents and family for their support, encouragement, and unconditional love through all these years.

Table of Contents

List of Figures

List of Schemes

List of Tables

Chapter 1: Introduction

1.1 Nucleic acids

Nucleic acids, DNA and RNA, are important biological molecules. Most genomes are made up of DNA and a few viruses have RNA genomes. In an adult human, each of the ~10^{13} somatic cells has its own DNA genome that comprises 46 chromosomes (~3.2 × 10^9 base pairs). The DNA genetic material in chromosomes is transcribed to messenger RNA, which encodes information for protein translation. RNA has multiple biological roles beyond acting as a genetic material. The unexpected discovery of catalytic RNA by Thomas Cech and Sidney Altman in the 1980s revealed that RNA molecules can function as biocatalysts.[1] Over the last two decades, the discoveries of RNA interference[1] and regulatory RNA[2] further illustrate the extensive biological role of RNA and intensify our ongoing research interest in nucleic acids. In addition, the powerful method of in vitro selection enables the identification of artificial functional nucleic acids.

Chemical structure of nucleic acids. Nucleic acids are polymers of nucleotide monomers, connected by phosphodiester bonds. A nucleotide monomer is a phosphate ester of a pentose, which is covalently linked to a nitrogenous base at the C′1 position of the sugar. For a deoxyribonucleotide, it consists of DNA monomers in which the pentose is 2′-deoxy-D-ribose. For a ribonucleotide, the pentose is D-ribose. The nitrogenous bases are planar, heterocyclic derivatives of purine or pyrimidine. In nucleic acids, adenine and guanine are the purine bases; cytosine, thymine, and uracil (found in RNA) are the pyrimidine bases.

1.1.1 DNA structure and biological role

Key features of B-form DNA double helical structure. The double helical structure was elucidated by Watson and Crick in 1953, which marked a new era of modern molecular biology.[3] The Watson–Crick structure of B-DNA has several notable features.

First, the two oligonucleotides are antiparallel with a right-handed twist to form a ~20 Å diameter helix. A helical turn consists of 3.6 base pairs (bp). Second, each base is hydrogen bonding to a base on the opposite strand. These hydrogen bond interactions between the purine and pyrimidine base between the two strands of DNA, also known as complementary base pairings. The hydrogen bonds, ~1–10 kcal mol^{-1} in bond energy, are crucial in stabilizing the DNA double helix. The π–π stacking interaction, which arises from the hydrophobic interaction of the nucleobases, also assists the formation of the DNA double helix. Successful cooperative pairings of hydrogen bonding allow two DNA strands to assemble spontaneously into the double helical structure. Third, major grooves and minor grooves are observed on the surface spiraling along the helical axis. The major groove (11.7 Å in width, 8.8 Å in depth) is larger than the minor groove (5.7 Å in width, 7.5 Å in depth).

Other DNA structures. The A-form DNA is observed in a low humidity and high salt environment, whereas the B-form DNA is found in a high humidity and low salt environment. A-form DNA is shorter (11 bases in each turn of 28 Å, 2.6 bp per turn) than the B-form DNA. In A-form DNA, the 3′-carbon lies above the plane of the other four sugar carbons, which is known as C-3′ endo. In B-form DNA, the 2′-carbon is in the endo position (**Fig 1.1**).

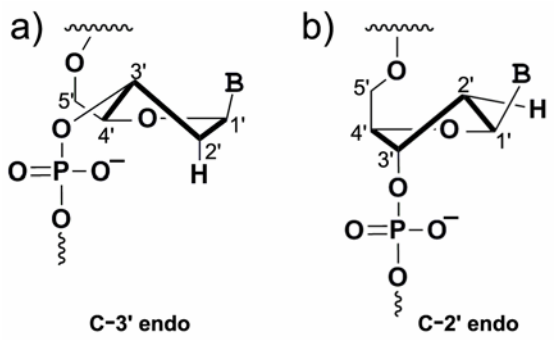


Figure 1.1 The sugar puckering of DNA. (a) The C-3′ endo structure. (b) The C-2′ endo structure.

Genes are made up of DNA. DNA is the building block of the genomes of almost all life forms, with the exception that some viruses have a RNA genome. A gene is a DNA sequence that encodes the information to produce proteins or functional RNA. In eukaryotes, genes consist of introns and exons. The chromosome consists of a collection of genes. The human genome (3.2×10^9 bp) locates in 22 different autosomes and 2 sex chromosomes in the nucleus. Genes are not evenly spaced in the chromosome. The average density of a human chromosome is between 0–64 genes per 100 kbp.

DNA is tightly packed in a chromosome. In a chromosome, DNA is associated with DNA-binding histone proteins to form nucleosomes. The nucleosome contains eight histone proteins, two molecules each of H2A, H2B, H3 and H4. These eight proteins together form a barrel-shaped core octamer. 140–150 bp DNA wrap around the core octamer. The nucleosomes stack on top of each other to form a 30 nm fiber called chromatin.

Only a fraction DNA encodes for functional units production. It was originally thought that the human genome contains around 80000–100000 genes. The completion of the Human Genome Project[4] in 2003 revealed that the human genome contains around 30000–40000 genes, which is only about twice the number of genes of a nematode worm

Caernoharbditis elegans. The exons of the human genome make up to 48 × 10^3 kbp, which is only 1.5% of the total genome. In contrast, 44% of the genome contains repetitive DNA sequences.

The small number of genes in the human genome is explained by the diversification of the proteome by three processes as follows. First, at the mRNA level, a ribosome can bind to alternate promoter sequences upstream of mRNA to initiate translation, generating a few variants of gene products. Second, alternative splicing occurs in more than half of the RNAs in vertebrates. A particular mRNA is spliced in multiple ways, which results in distinct proteins after translation. Third, proteins can be covalently modified after translation, which is known as posttranslational modification. The use of alternate promoter sequences, alternative splicing, and posttranslational modifications increases the complexity of the human proteome.

1.1.2 RNA structure and biological role

Structural versatility of RNA. In biology, RNA are more structural versatile than DNA as they fold into different secondary structures such as bulges, hairpin loop and stem junction (**Fig. 1.2**). RNA also fold into complex tertiary structures for functions. For example, tRNA fold into three-dimensional structures with different functional domains.

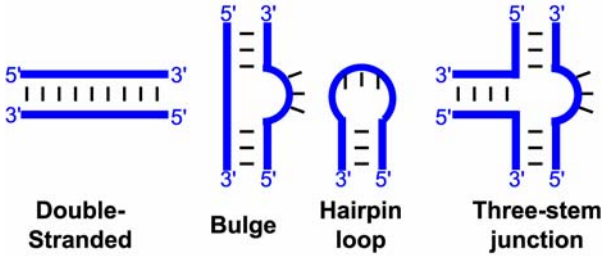


Figure 1.2 Examples of secondary structure of RNA.

RNA in protein translation. The series of contiguous triplet nucleotide mRNA codons encode the sequence of amino acids in a protein. In protein translation, the hydrogen bonding of the anticodon of a tRNA molecule and the codon of mRNA directs the addition of specific amino acids to the growing polypeptide in the ribosome.

RNA in post-transcriptional regulation of gene expression

Silencing RNA. In 1998, Andrew Fire and Craig Mellow observed that the injection of the double-stranded RNA duplex to *Caenorthabditis elegans* results in a twicking phenotype.[1] This observation led to the discovery that the double-stranded RNA was responsible for gene silencing, known as RNA interference. In RNA interference, the enzyme DICER recognizes and cleaves invading double-stranded RNA viruses to ~ 20 nt small interference RNA (siRNA). SiRNA then associate with another protein RISC. RISC guides a specific siRNA to its complementary mRNA. The complementary base-pairings leads to the degradation of the entire mRNA and hence represses the expression of a specific protein (**Fig 1.3a**). The discovery of RNAi greatly expedites the gene annotation process. Using RNAi, researchers can generate loss-of-function phenotypes within days.[5] In comparison, it can take years to prepare knockout mice. Fire and Mellow were awarded the Nobel Prize in Physiology or Medicine in 2006 for their discovery.

MicroRNA. In 1994, *lin*-4 microRNA[2a] (miRNA), a gene that regulates the timing of *Caenorthabditis elegans* larval development, was discovered by Victor Ambros and colleagues. miRNA are small ~ 22 nt non-coding RNA that post-transcriptionally repress protein expression by imperfect base-pairing of the mature miRNA with the 3′ UTR region of mRNAs. Numerous miRNA were identified in flies, worms, vertebrates and plants. miRNA are important regulators in developmental processes and are linked to

malignancy. More than 500 miRNA genes have currently been identified in humans and it was estimated that 30% of human genes are regulated by miRNA.[6]

The process of miRNA maturation[7] starts from the transcription of miRNA gene by RNA polymerase II in the nucleus to generate capped and polyadenylated transcripts called primary miRNA (pri-miRNA). (**Fig 1.3b**) The pri-miRNA are cleaved by RNase III endonuclease Drosha with its double-stranded RNA-binding protein cofactor Pasha (or DGCR8) to form ~70 nt stem loop precursor miRNA known as precursor miRNA (pre-miRNA). The stem loop pre-miRNA are actively exported by RAN–GTP and exportin 5 to the cytoplasm. In the cytoplasm, another RNase III endonuclease Dicer recognizes the double-stranded portion of the pre-miRNA and cleaves the pre-miRNA at about two helical turns away from the stem loop to generate transient ~22 nt miRNA:miRNA* duplexes. The miRNA:miRNA* duplexes are unwound and loaded into the Argonaute protein of the silencing complex (RISC) to yield ~22 nt mature miRNA. The miRNA are then hybridized to the 3′ UTR of the target mRNA by recognition through a 6 nt seeding region, which leads to transcription repression.

RNA interference (RNAi) primarily silences protein expression through mRNA degradation. In contrast, miRNA, depending on the complementarity with 3′ UTR of the target transcript, work by either blocking the protein translation machinery or by mRNA degradation by similar mechanism as in RNAi.

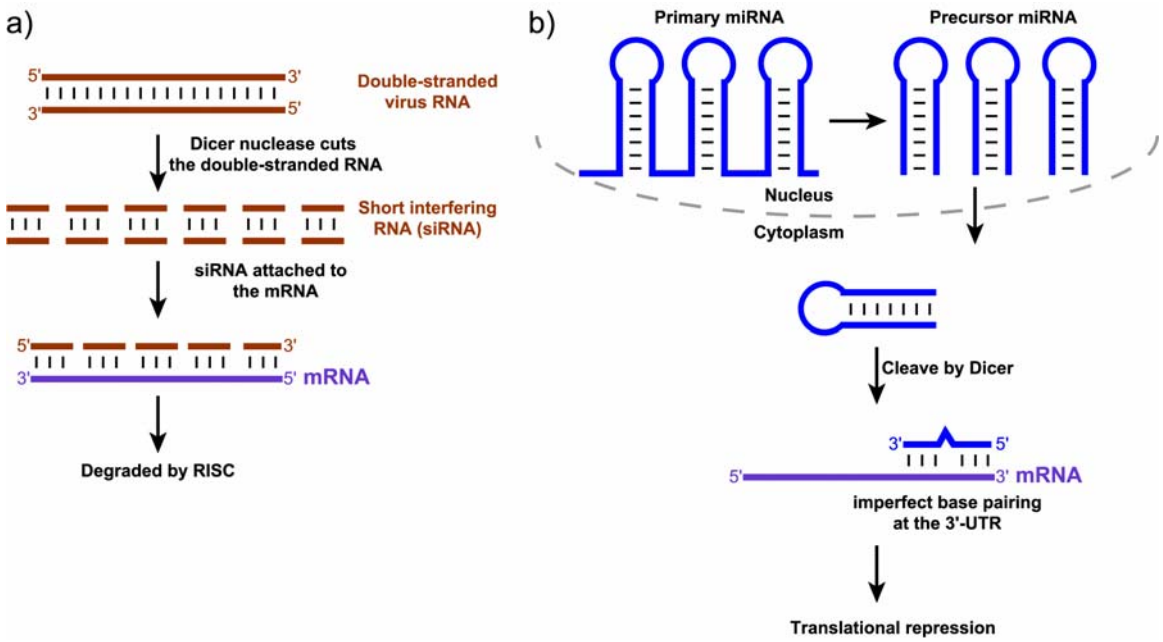


Figure 1.3 The siRNA and miRNA biogenesis pathway. (a) The siRNA pathway. (b) The miRNA pathway.

Riboswitches. Riboswitches are structural noncoding mRNA motifs that sense cellular metabolites and allosterically control gene expression.[7] Each riboswitches contain an aptamer domain and an expression platform. The aptamer domain for a specific metabolite is well conserved, located upstream from the expression platform. Transcription termination, translation initiation and splicing control (only in eukaryotes) are the most common mechanisms utilized by riboswitches to regulate gene expression.[7]

1.2 Ribozymes

Ribozymes are single-stranded catalytic RNA molecules. The first two ribozymes were independently discovered by Thomas Cech and Sidney Altman in the 1980s.[8] Thomas Cech discovered that the ribosomal RNA of *Tetrahymena Thermophila* undergoes autocatalytic splicing of the introns and joining of the flaking extrons in the absence of protein. Sidney Altman discovered that the RNA component of Ribonuclease P (RNase P) is responsible of its ribonuclease activity in tRNA processing. Since then,

many natural ribozymes and artificial ribozymes were identified. Together, the growing repertoire of the reaction catalyzed by RNA molecules illustrates the catalytic ability of nucleic acids in general. In this section, natural ribozymes and artificial ribozymes will be discussed.

1.2.1 Natural ribozymes

Ribozymes can be found widely in nature such as in the genomes of viruses, fungi, and plants. The known natural ribozymes catalyze three types of reaction: transesterfication[9-10], hydrolysis[11] and amide bond formation[12] (**Fig 1.4**). Examples of natural ribozymes are summarized in **Table 1.1**.

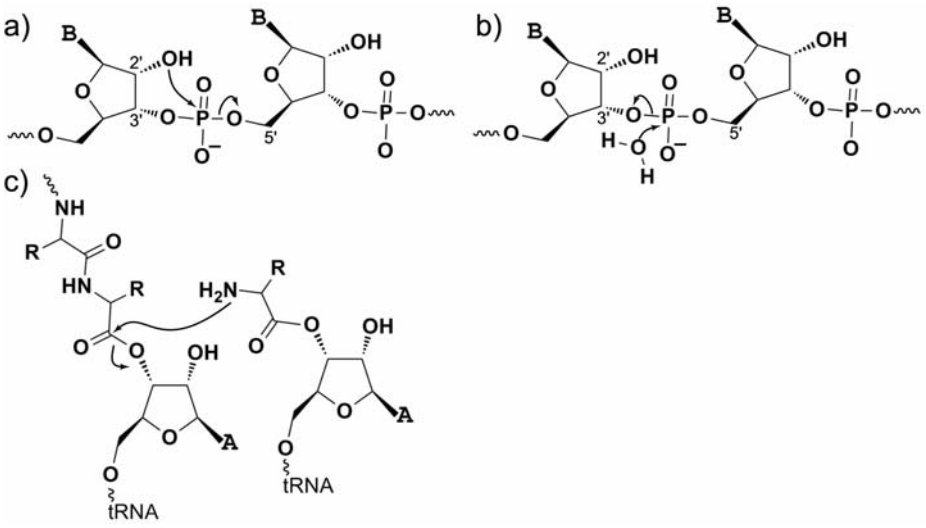


Figure 1.4 Examples of reactions catalyzed by known ribozymes. (a) RNA cleavage by ribozyme-catalyzed tranesterification. (b) RNA cleavage by ribozyme-catalyzed hydrolysis. (c) Peptide bond formation catalyzed by rRNA in the ribosome.

Table 1.1 Examples of known natural ribozymes

Ribozyme	Size (nt)	Function	Reaction	Products
Self-Splicing RNAs				
Group I introns	200–1500	Splicing	Transesterification	5′-guanosine 3′-OH introns, ligated exons
Group II introns	300–3000	Splicing	Transesterification	2′-5′ lariats and 3′-OH, ligated exons
Self-cleaving RNAs				
Hammerhead ribozyme	~40	Replication	Transesterification	5′-OH, 2′,3□-cyclic phosphate products
Hairpin ribozyme	~60	Replication	Transesterification	5′-OH, 2′,3′-cyclic phosphate products
HDV	~80	Replication	Transesterification	5□-OH, 2′,3′-cyclic phosphate products
RNase P	140–490	tRNA processing	Hydrolysis	5′-OH, 3′-OH products
Peptidyl transferase RNAs				
Ribosomal RNA (23S *E.coli* rRNA, for example)	~2900	Peptidyl transfer	Peptide bond formation	A new peptide bond formation of the peptidyl-tRNA at the A site, deacylated tRNA at the P site

Adapted from Ref (13).

Ribozyme-catalyzed transesterification. The majority of natural ribozymes catalyzes transesterfication. Self-cleaving hammerhead ribozymes, found in plant viroids and satellite RNA[9], catalyze the site-specific cleavage of a phosphodiester bond by the nucleophilic attack of the 2′-hydroxyl of the ribose to the adjacent phosphodiester bond, and generate a 2′,3′-cyclic phosphate product and a 5′-hydroxyl product.

In RNA splicing, the Group I and II introns catalyze two transesterfication reactions. Group I introns catalyzes the attack of an exogenous guanosine 3′-hydroxyl group to the 5′ end of the intron (**Fig 1.5a**).[10] In the subsequent transesterfication, 5′-exon is ligated to the 3′-introns. In Group II introns, the 2′-hydroxyl of branch-site adenosine attacks the 5′

end of the intron followed by another transesterfication reaction to ligate the 5′-exon and the 3′-intron (**Fig 1.5b**).

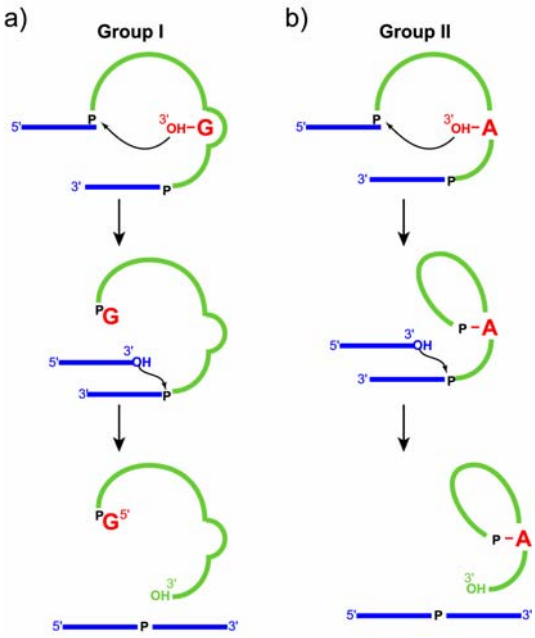


Figure 1.5 The Group I and II intron splicing pathway. (a) The Group I intron. (b) The Group II intron.

Ribozyme-catalyzed hydrolysis. RNAse P, found in eukaryotes and prokaryotes, consists of a RNA component and at least one protein.[11] In tRNA processing, RNase P catalyzes the cleavage of the 5′-end of pre-tRNA to generate the mature tRNA. The RNA component of the RNase P catalyzes the site-specific nucleophilic attack of a water molecule to a phosphodiester bond, which produces a 3′-hydroxyl and a 5′-phosphate product.

Ribozyme-catalyzed peptide bond formation. One of the most important ribozyme-catalyzed reactions is the peptide bond formation in the ribosome. The ribosome catalyzes the attack of the α-amino group of the aminoacyl-tRNA at the A site to the carbonyl carbon of the peptidyl-tRNA at the P site. The nucleophilic attack leads to a

single peptide bond formation of the peptidyl-tRNA at the A site and the deacylation of tRNA at the P site.

In bacteria, the ribosome consists of three rRNA molecules and two subunits: 50S and 30S. The peptidyl transferase center is located in the 50S subunit. Biochemical studies[12] provide evidence that the peptide bond formation is catalyzed by RNA. The notion of RNA-catalyzed peptide bond formation is further supported by the determination of high-resolution X-ray crystal structures[14] of the *Haloarcula Marismortui* ribosome with aminoacyl-tRNA substrate analogs (CCdA-phosphate-puromycin, and the mini-helix 3′ puromycin). The crystal structures show that the substrate analogs contact exclusively with the 23S rRNA in the peptidyl transferase center in the 50S subunit. The all-RNA active center establishes that 23S rRNA catalyzes the peptide bond formation while the protein components properly orient 23S rRNA for catalysis.

1.2.2 Artificial ribozymes

The discovery of natural ribozymes sparked an intense interest in identifying ribozyme in the laboratory. The method of in vitro selection enables researchers to identify many ribozymes that encompasses a wide range of reaction scopes.

In vitro selection as a means to identify functional nucleic acids. In vitro selection is a powerful method that allows researchers to identify functional nucleic acids that are not found in nature.[15] In vitro selection comprises iterative rounds of separation and amplification of the catalytically active functional nucleic acids. The separation method is either by size difference, affinity capture, or both. Polymerase chain reaction (PCR) is performed to amplify functional nucleic acids. The general processes of in vitro selection to identify artificial ribozymes and deoxyribozymes are similar. For ribozyme in vitro

selection, RT-PCR followed by transcription is required in the amplification step. Typically, nucleic acids with the desired catalytic function are selected from a pool of ~10^{14} nucleic acid molecules. The nucleic acid molecules contain a random region flanked by constant sequences, which act as primer-binding regions for PCR amplification. Typically, the random region is 40–80 nt in size, denoted as N_{40}–N_{80}. The nucleic acid pool with a random region is physically ligated to one of its substrates. In the key selection step, the pool of nucleic acid is incubated with another substrate and divalent metal ions. Successful catalysis would result in either a size difference of the product and the substrate-ligated nucleic acid molecules which can be separated by gel purification, or incorporation of a chemical moiety that can be isolated by affinity capture. The isolated material contains a population of functional nucleic acids which is then PCR-amplified (for ribozyme in vitro selection, RT-PCR and transcriptions are performed at this stage), ligated to one of the substrates, and subjected to another selection step. These iterated rounds are continued to enrich the population of functional nucleic acids with the intended activity.

Using in vitro selection, researchers are able to identify ribozymes with new catalytic functions beyond the typical natural ribozyme-catalyzed reactions such as transesterfication, hydrolysis and amide bond formation. The first artificial ribozymes were reported in 1990.[16] Notably, artificial ribozymes can catalyze C–C bond formations such as in the aldol reaction[17] (**Fig 1.6a**) and the Diels–Alder reaction[18] (**Fig 1.6b**), which are considered the prototypical organic reactions. **Table 1.2** exemplifies the wide scope of reactions that are catalyzed by ribozymes.

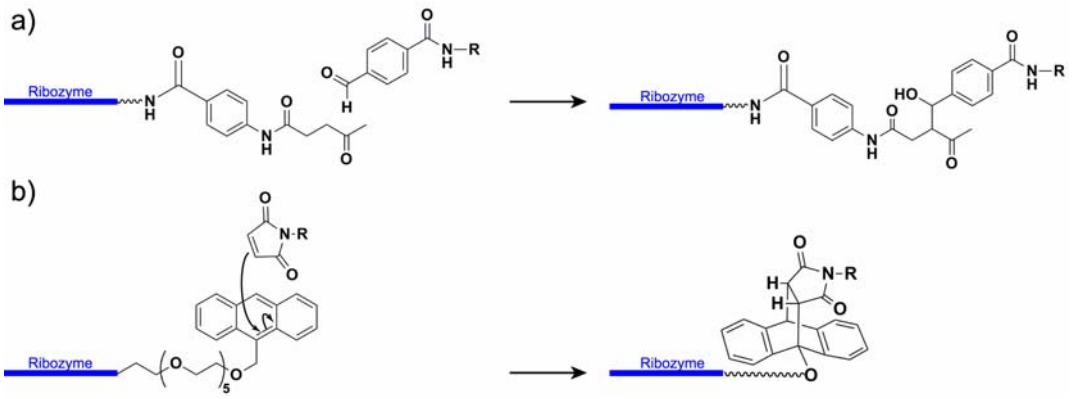


Figure 1.6 Examples of artificial ribozyme-catalyzed reaction. a) The aldol reaction. [17] **b) The Diels-Alder reaction.**[18]

Table 1.2 Types of reactions catalyzed by artificial ribozymes

Reaction	Reference	Reaction	Reference
RNA cleavage	(19)	*S*-alkylation	(20)
RNA ligation	(21)	Carbonate hydrolysis	(22)
RNA Phosphorylation	(23)	Aldol reaction	(17)
RNA Branching	(24)	Peptide bond formation	(25)
RNA Capping	(26)	Amino acid adenylation	(27)
Mononucleotide polymerization	(28)	Diels-Alder cycloaddition	(18)
Aminoacyl transfer	(29)	Biphenyl isomerization	(30)
Aldehyde reduction	(31)	Porphyrin metallation	(32)

Note: Adapted from (33). This table is not a comprehensive list of known ribozymes.

Comparison of directed evolution and in vitro selection. Directed evolution is an in vitro method, useful for the development of highly active and specific biocatalysts, that shortens Darwinian evolution down to a period of months or even weeks.[34] Directed evolution involves an iterative process that consists of mutant library generation, recombination, and screening or selection. The success of directed evolution to optimize catalyst activity is based on the efficiency and sensitivity of the screening or selection method, because mutant libraries can be generated by many methods, such as random mutagenesis by error-prone PCR, chemical mutagenesis and DNA shuffling.

There are several key differences between in vitro selection and directed evolution.

(1) Directed evolution requires a known functional catalyst to begin the process. In directed evolution, a mutant library was generated from the wild-type to initiate the rounds of screening or selection. In contrast, in vitro selection of nucleic acid catalysts does not require prior knowledge of functional nucleic acid catalysts for a particular reaction. Directed evolution is very well suited for improving existing functions of proteins as well as functional nucleic acids. However, it is still a great challenge to identify new functions for protein catalysts without the amino acid sequence blueprints from known wild-type function protein catalysts.

(2) Directed evolution comprising a screening or selection step followed by mutant library generation of the clones that exhibit desired activities. Therefore, directed evolution increases the diversity of the functional molecules throughout the round of screening or selections. Whereas, in the process of in vitro selection, the diversity of the nucleic acids molecules originated from the pool of the nucleic acids molecules in the beginning.

In vitro selection is not amenable for de novo identification of new proteins. One of the obstacles is that amino acids monomers are not amenable for PCR amplification. The amplification obstacle can be overcome by the process of phage or mRNA display, which physically links the genetic information to the phenotype of the protein.

However, more importantly, the 20 amino acids monomer makes up a vast collection of possible sequences that the sequence coverage of in vitro selection for protein is exceedingly low. For a small peptide of 100 amino acids in size, there are $20^{100} \sim 1 \times$

10^{130} possible amino acid sequences, which corresponds to the mass of 1.4×10^{134} daltons.

Screening and selection are distinct methods. Screening is the examination of catalytic activity by the phenotype of individual clones, typically by fluorescence and UV absorbance methods. In contrast, selection acts on a pool of genes such that only those that display the desired phenotype are propagated.[35] Selection is the method of choice for larger libraries $\geq 10^7$, because screening is applicable only to much smaller library sizes, typically only a few thousands clones (up to 10^6 library size).

1.3 Deoxyribozymes

Deoxyribozymes are single-stranded catalytic DNA molecules. Deoxyribozymes do not exist in nature and are identified by in vitro selection. The first deoxyribozyme was reported by Breaker and Joyce in 1994.[36] The 10–23 and 8–17 deoxyribozyme[37], reported by Santoro and Joyce, are the first examples of deoxyribozymes that have the ability to cleave all-RNA substrates (**Fig 1.7**). The 10–23 and 8–17 deoxyribozyme were named after the selection round and the clone number at which each of the deoxyribozyme was identified. The 10–23 deoxyribozyme cleaves all purine-pyrimidine junctions (robust activity at A↓U ; reduced activity at A↓C, and G↓C). The 8–17 deoxyribozyme cleaves at the A↓G junction. Variants of the 8–17 deoxyribozymes cleave all N↓G junctions, where N stands for any nucleotide.

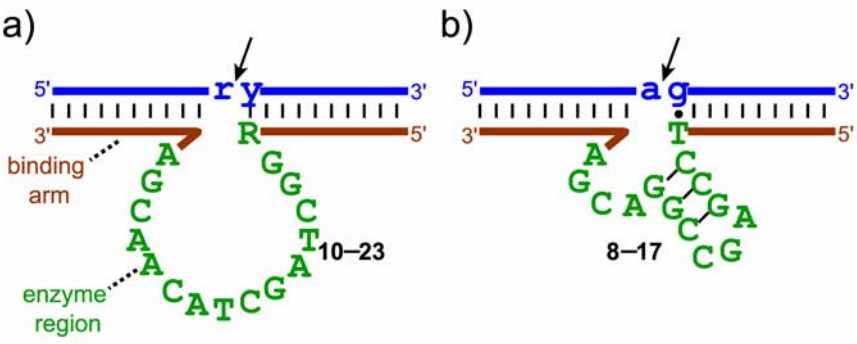


Figure 1.7 RNA-Cleaving deoxyribozymes. (a) The 10–23 deoxyribozyme. (b) The 8–17 deoxyribozyme. Adapted from Ref (37).

Many deoxyribozymes have been identified after the first reported deoxyribozymes. Representative deoxyribozyme-catalyzed reactions include phosphodiester bond cleavage[36] and ligation[38,39], DNA depurination[40], and Diels-Alder reaction[41]. Although nucleotides in DNA have less diverse functional groups than that of amino acids in protein, deoxyribozymes catalyze a wide variety of chemical reactions and display robust catalytic abilities with rate enhancement[42] ranging from 10^3 to 10^{12}.

Table 1.3 Representative types of reactions catalyzed by deoxyribozymes

Reaction	Reference	Reaction	Reference
RNA cleavage	(36)	DNA deglycosylation	(43)
RNA ligation	(39)	Phosphoramidate Cleavage	(44)
RNA branching	(45)	Nucleopeptide linkage formation	(46)
RNA lariat formation	(47)	Reductive Amination	Chapter 3
DNA phosphorylation	(48)	Thymine dimer photoreversion	(49)
DNA adenylation	(50)	DNA depurination	(40)
Oxidative DNA cleavage	(51)	Diels-Alder cycloaddition	(41)
DNA hydrolysis	(52)	Porphyrin metalation	(53)

Note: Adapted from (33). This table is not a comprehensive list of known deoxyribozymes.

Practical applications of nucleic acid catalysts. Beyond fundamental studies of the catalytic potential of nucleic acids, researchers are applying deoxyribozymes and ribozymes for practical applications.

RNA-cleaving deoxyribozyme as alloteric biosensors. Metal ions or small molecules dependent RNA-cleaving deoxyribozymes can be utilized as sensors. The Lu laboratory reported several RNA-cleaving deoxyribozymes that sense specific metal ions. For instance, the 39E deoxyribozyme detects UO_2^{2+} at concentrations as low as 45 pM. [54] The 39E deoxyribozyme cleaves a single ribonucleotide embedded in a DNA substrate with a fluorophore and a quencher at the 5′ end and 3′ end, respectively. Another quencher is covalently attached at the 3′ end of the 39E deoxyribozyme to suppress the fluorescence of the fluorophore. In presence of UO_2^{2+}, the 39E deoxyribozyme cleaves at the ribonucleotide, which results in a fluorescence signal (**Fig 1.8**).

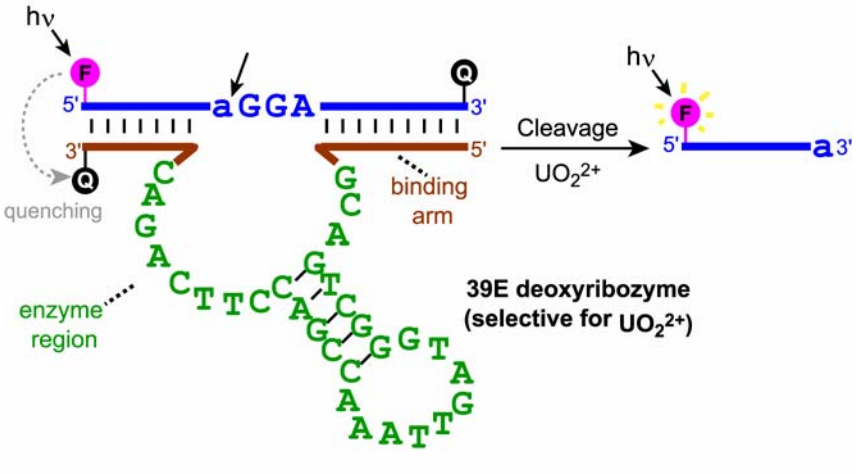


Figure 1.8 The 39E deoxyribozyme senses uranium. Adapted from ref (38).

Deoxyribozyme-catalyzed labeling of RNA. The Silverman laboratory has developed a RNA labeling approach, known as Deoxyribozyme-catalyzed Labeling (DECAL).[55] The 10DM24 deoxyribozyme catalyzes the nucleophilic attack of an internal 2′-hydroxyl of an adenosine of a specific RNA substrate to the α-phosphate of the 5′-triphosphate of another RNA substrate, which results in a branched RNA formation. In DECAL, a 5′-triphosphate RNA is tagged with fluorophores or biotin molecules. The 10DM24

deoxyribozyme catalyzed the branched ligation of the tagged RNA to the internal adenosine. To demonstrate the applicability of the DECAL approach, specific sites of the P4–P6 domain of the group I introns were labeled with fluorophores using DECAL, which enables studies of the P4–P6 folding event using fluorescence resonance energy transfer (**Fig 1.9**).

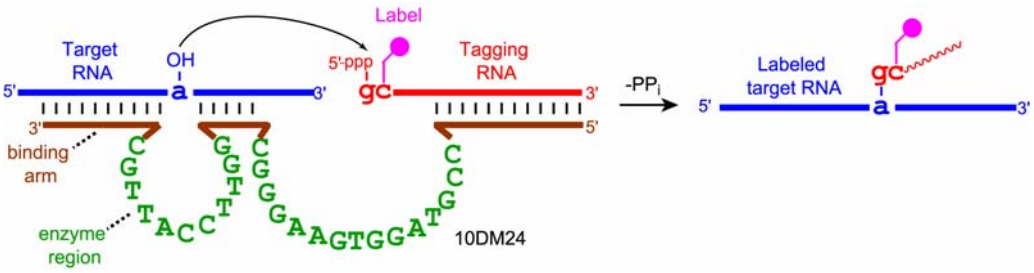


Figure 1.9 The DECAL approach to site-specifically label a RNA. Adapted from ref (37).

1.4 Protein side chains modifications

Protein side chains modifications are prevalent in biology. The importance of protein side chain modifications provides a scientific impetus for researchers to develop different in vitro methods to modify protein side chains. In the following section, protein side chains modifications in biology will be discussed briefly.

Amino acids. Proteins are macromolecules made of amino acid monomers. The size range of proteins varies from about 70 amino acids to 15000 amino acids. Directed by its specific amino acid sequence, proteins are folded into three-dimensional structures that enable functions. There are 22 genetically encoded amino acids. The 21st amino acid, selenocysteine, is found in prokaryotes and eukaryotes (**Fig 1.10a**). The 22nd amino acid is pyrrolysine, found in an archaebacterium (**Fig 1.10b**).

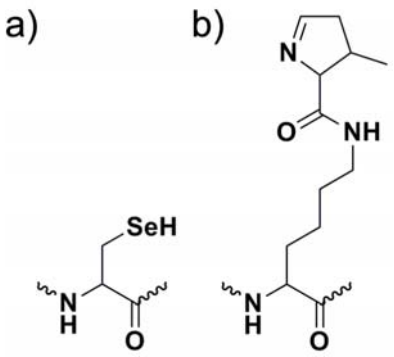


Figure 1.10 The 21st and 22nd amino acid. (a) The selenocysteine side chain. (b) The pyrrolysine side chain.

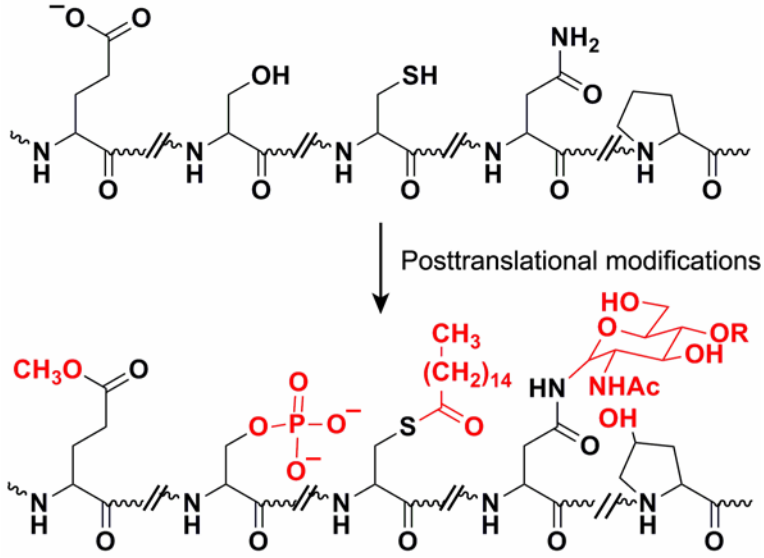


Figure 1.11 Posttranslational modifications of protein side chains. Adapted from Ref (56).

In biology, the process of modifications of protein after translation is known as posttranslational modification (PTM; **Fig 1.11)**. PTM is essential to almost all cellular processes. PTM is broadly divided in two categories: (1) the covalent addition of chemical moieties to amino acids side chains, (2) the proteolytic modifications of precursor proteins (for example, insulin is first translated to an inactive single-chain form followed by the cleavage by a specific protease into its active form).[56] **Table 1.4** illustrates representaivte types of PTM.

In general, PTM occurs at the nucleophilic amino acids side chains, in which enzymes catalyze the covalent addition of electrophilic chemical groups. Exceptions to

the rule that PTM occurs at the nucleophilic amino acids are exemplified as follows: (1) the hydroxylation of proline by reactive oxygen species that is mediated by Fe(II)-dependent hydroxylases, and (2) the *N*-glycosylation at the low nucleophilic amide nitrogen of asparagines.[56]

Table 1.4 Representative types of protein side chains modifications

Residue	Reaction	Example
Glu	Methylation	Chemotaxis receptor proteins
Ser	Phosphorylation	Serine kinases and phosphatases
	O-glycosylation	Notch *O*-glycosylation
Thr	Phopshorylation	Threonine kinases and phosphatases
Tyr	Phosphorylation	Serine kinases and phosphatases
	Sulfation	CCR5 receptor maturation
	Ortho-nitration	Inflammatory responses
His	Aminocarboxypropylation	Diphthamide formation
	Phosphorylation	Sensor protein kinases
Lys	*N*-methylation	Histone methylation
Cys	*S*-Acylation	Ras
	Phosphorylation	PTPases
Met	Oxidation to sulfoxide	Met sulfoxide reductase
Arg	*N*-ADP-ribosylation	$G_{S\alpha}$

Adapted from Ref (55).

One important example PTM is the phosphorylation of the hydroxyl side chains of serine, threonine, and tyrosine by kinases, which are essential elements in modulating signaling cascades in cells. The majority of phosphorylation occurs at serine and threonine residues. Adenosine triphosphate is the major phosphoryl donor, in which the γ-phosphate is transferred to the hydroxyl side chains by kinases. In a small subset of kinases, guanosine triphosphate is the phosphoryl donor.

The biological importance of kinases can be illustrated by the transmembrane receptor tyrosine kinases (RTK). The binding with ligands such as hormones and growth factors at the receptors at the extracellular region induces RTK to dimerize and autophorphorylate specific tyrosine residues in the cytosolic domain. The autophosphorylation concomitantly turns on a myriad of signaling cascades. For instance,

the Ras-activated MAP kinase cascade begins with the binding of its cognate growth factor at the RTK (**Fig 1.12**). The binding triggers the dimerization and autophosphorylation of tyrosine at the RTK. Grb2/Sem-5 binds to the phosphorylated tyrosine domain at the RTK and activates multiple downstream effectors such as Ras. Activated Ras binds to Raf, a Ser/Thr kinase that activates MEK. The MEK activates MAPK. MAPK activates several transcription factors. Given the prevalent examples and significance of side chain modifications in biology, novel methods to modify protein side chains are of great scientific interest.

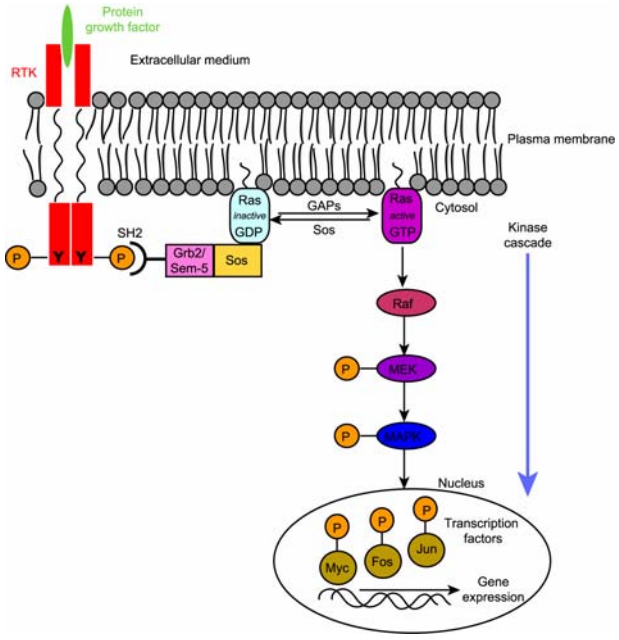


Figure 1.12 The Ras-activated MAP kinase cascade. Adapted from Ref (57).

1.5 Thesis Research Focus

The Silverman laboratory first reported a deoxyribozyme that catalyzes a nucleopeptide ligation reaction using tyrosine hydroxyl as a nucleophile.[46] The broad implication of this initial work is that nucleic acid catalysts have the ability to catalyze reactions involving protein side chains. This thesis reports the development of

deoxyribozymes that catalyze covalent modification of peptide substrates. Achieving this research objective will allow potential future DNA-catalyzed synthesis of biologically relevant protein variants.

Chapter 2 details the comprehensive in vitro selection effort to identify novel deoxyribozymes that covalently modify the tyrosine or serine side chains of DNA-tethered tripeptide substrates as well as the tyrosine side chain of a free tripeptide. Chapter 3 presents the unexpected discovery of DNA-catalyzed reductive amination that may be applicable to site-specific labeling of nucleic acids. Chapter 4 outlines the current efforts using click chemistry for direct free peptide selection.

1.6 References

1 (a) Cech, T. R.; Zaug, A. J.; Grabowski, P. J., In vitro splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell* **1981**, *27*, 487–496; (b) Guerriertakada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S., The RNA Moiety of Ribonuclease P Is the Catalytic Subunit of the Enzyme. Cell **1983**, *35*, 849–857.

2 Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C., Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **1998**, *391*, 806–811.

3 (a) Lee, R. C.; Feinbaum, R. L.; Ambros, V., The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **1993**, *75* (5), 843–854; (b) Reinhart, B. J.; Slack, F. J.; Basson, M.; Pasquinelli, A. E.; Bettinger, J. C.; Rougvie, A. E.; Horvitz, H. R.; Ruvkun, G., The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* **2000**, *403*, 901–906.

4 Watson, J. D.; Crick, F. H., Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **1953**, *171* (4356), 737–738.

5 Finishing the euchromatic sequence of the human genome. *Nature* **2004**, *431* (7011), 931–945.

6 Dorsett, Y.; Tuschl, T., siRNAs: Applications in functional genomics and potential as therapeutics. *Nat. Rev. Drug Discov.* **2004**, *3*, 318–329.

7 Bartel, D. P., MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297.

8 Breaker, R. R., Riboswitches and the RNA World. *Cold Spring Harb. Perspect Biol.* **2010**.

9 Tanner, N. K., Ribozymes: the characteristics and properties of catalytic RNAs. *FEMS Microbio.l Rev.* **1999**, *23*, 257–275.

10 Saldanha, R.; Mohr, G.; Belfort, M.; Lambowitz, A. M., Group I and group II introns. *FASEB J.* **1993**, *7*, 15–24.

11 (a) Altman, S.; Kirsebom, L.; Talbot, S., Recent studies of ribonuclease P. *FASEB J.* **1993**, *7*, 7–14; (b) Kurz, J. C.; Fierke, C. A., Ribonuclease P: a ribonucleoprotein enzyme. *Curr. Opin. Chem. Biol* **2000**, *4*, 553–558.

12 (a) Green, R.; Samaha, R. R.; Noller, H. F., Mutations at nucleotides G2251 and U2585 of 23 S rRNA perturb the peptidyl transferase center of the ribosome. *J. Mol. Biol.* **1997**, *266*, 40–50; (b) Moazed, D.; Noller, H. F., Interaction of tRNA with 23S rRNA in the ribosomal A, P, and E sites. *Cell* **1989**, *57*, 585–597; (c) Noller, H. F.; Green, R.; Switzer, C., Ribosome-catalyzed peptide-bond formation with an a-site substrate covalently linked to 23S ribosomal RNA. *Science* **1998**, *280*, 286–289; (d) Noller, H. F.; Hoffarth, V.; Zimniak, L., Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* **1992**, *256*, 1416–1419; (e) Samaha, R. R.; Green, R.; Noller, H. F., A base pair between tRNA and 23S rRNA in the peptidyl transferase centre of the ribosome. *Nature* **1995**, *377*, 309–314.

13 Lehmann, K.; Schmidt, U., Group II introns: structure and catalytic versatility of large natural ribozymes. *Crit. Rev. Biochem. Mol. Biol.* **2003**, *38*, 249–303.

14 Nissen, P.; Hansen, J.; Ban, N.; Moore, P. B.; Steitz, T. A., The structural basis of ribosome activity in peptide bond synthesis. *Science* **2000**, *289*, 920–930.

15 a) Breaker, R. R., In Vitro Selection of Catalytic Polynucleotides. *Chem. Re.* **1997,** *97*, 371–390; (b) Ellington, A. D.; Szostak, J. W., In vitro selection of RNA molecules that bind specific ligands. *Nature* **1990**, *346*, 818–822; (c) Tuerk, C.; Gold, L., Systematic Evolution of Ligands by Exponential Enrichment –RNA Ligands to Bacteriophage-T4 DNA-Polymerase. *Science* **1990**, *249*, 505–510.

16 Robertson, D. L.; Joyce, G. F., In vitro Selection of an RNA Enzyme That Specifically Cleaves Single-Stranded DNA. *Nature* **1990**, *344*, 467–468.

17 Fusz, S.; Eisenfuhr, A.; Srivatsan, S. G.; Heckel, A.; Famulok, M., A ribozyme for the aldol reaction. *Chem. Biol.* **2005**, *12*, 941–950.

18 Tarasow, T. M.; Tarasow, S. L.; Eaton, B. E., RNA-catalysed carbon-carbon bond formation. *Nature* **1997**, *389*, 54–57.

19 Williams, K. P.; Ciafre, S.; Tocchini-Valentini, G. P., Selection of novel Mg^{2+}-dependent self-cleaving ribozymes. *EMBO J.* **1995**, *14*, 4551–4557.

20 Wecker, M.; Smith, D.; Gold, L., In vitro selection of a novel catalytic RNA: characterization of a sulfur alkylation reaction and interaction with a small peptide. *RNA* **1996**, *2*, 982–994.

21 Bartel, D. P.; Szostak, J. W., Isolation of new ribozymes from a large pool of random sequences. *Science* **1993**, *261*, 1411–1418.

22 Yu, J.; Chun, S. M.; Jeong, S. J.; Kim, J. M.; Chong, B. O.; Park, Y. K.; Park, H., Cholesterol esterase activity by in vitro selection of RNA against a phosphate transition-state analogue. *J. Am. Chem. Soc* **1999**, *121*, 10844–10845.

23 Lorsch, J. R.; Szostak, J. W., In vitro evolution of new ribozymes with polynucleotide kinase activity. *Nature* **1994**, *371*, 31–36.

24 Tuschl, T.; Sharp, P. A.; Bartel, D. P., Selection in vitro of novel ribozymes from a partially randomized U2 and U6 snRNA library. *EMBO J* **1998**, *17*, 2637–2650.

25 Zhang, B.; Cech, T. R., Peptide bond formation by in vitro selected ribozymes. *Nature* **1997**, *390*, 96–100.

26 Chapman, K. B.; Szostak, J. W., Isolation of a Ribozyme with 5'-5' Ligase Activity. *Chem. Biol.* **1995**, *2*, 325–333.

27 Yarus, M.; Kumar, R. K., RNA-catalyzed amino acid activation. *Biochemistry* **2001**, *40*, 6998–7004.

28 Johnston, W. K.; Unrau, P. J.; Lawrence, M. S.; Glasner, M. E.; Bartel, D. P., RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science* **2001**, *292*, 1319–1325.

29 Illangasekare, M.; Yarus, M., A tiny RNA that catalyzes both aminoacyl-RNA and peptidyl-RNA synthesis. *RNA* **1999**, *5*, 1482–1489.

30 Prudent, J. R.; Uno, T.; Schultz, P. G., Expanding the scope of RNA catalysis. *Science* **1994**, *264*, 1924–1927.

31 Tsukiji, S.; Pattnaik, S. B.; Suga, H., Reduction of an aldehyde by a NADH/Zn^{2+}-dependent redox active ribozyme. *J. Am. Chem. Soc.* **2004**, *126*, 5044–5045.

32 Conn, M. M.; Prudent, J. R.; Schultz, P. G., Porphyrin metalation catalyzed by a small RNA molecule. *J. Am. Chem. Soc.* **1996**, *118*, 7012–7013.

33 S. K. Silverman, Artificial functional nucleic acids: aptamers, ribozymes and deoxyribozymes identified by in vitro selection. *Functional Nucleic Acids for Analytical Applications*, Y. Li and Y. Lu, eds.; Springer Science + Business Media, LLC: New York, **2009**.

34 Zhao, H.; Chockalingam, K.; Chen, Z., Directed evolution of enzymes and pathways for industrial biocatalysis. *Curr. Opin. Biotechnol.* **2002**, *13*, 104–10.

35 Olsen, M.; Iverson, B.; Georgiou, G., High-throughput screening of enzyme libraries. *Curr. Opin. Biotech* .**2000,** *11*, 331–337.

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

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

38 Baum, D. A.; Silverman, S. K., Deoxyribozymes: useful DNA catalysts in vitro and in vivo. *Cell. Mol. Life. Sci.* **2008,** *65*, 2156–2174.

39 Silverman, S. K.; Flynn-Charlebois, A.; Wang, Y. M.; Prior, T. K.; Rashid, I.; Hoadley, K. A.; Coppins, R. L.; Wolf, A. C., Deoxyribozymes with 2□-5□ RNA ligase activity. *J. Am. Chem. Soc.* **2003,** *125*, 2444–2454.

40 Hobartner, C.; Pradeepkumar, P. I.; Silverman, S. K., Site-selective depurination by a periodate-dependent deoxyribozyme. *Chem. Commun.* **2007**, (22), 2255–2257.

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

42 Silverman, S. K., DNA as a versatile chemical component for catalysis, encoding, and stereocontrol. *Angew. Chem., Int. Ed.* **2010**, *49*, 7180–7201.

43 Sheppard, T. L.; Ordoukhanian, P.; Joyce, G. F., A DNA enzyme with N-glycosylase activity. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (14), 7802–7807.

44 Burmeister, J.; vonKiedrowski, G.; Ellington, A. D., Cofactor-assisted self-cleavage in DNA libraries with a 3'–5' phosphoramidate bond. *Angew. Chem., Int. Ed.* **1997**, *36*, 1321–1324.

45 Wang, Y. M.; Silverman, S. K., Deoxyribozymes that synthesize branched and lariat RNA. *J. Am. Chem. Soc.* **2003**, *125*, 6880–6881.

46 Pradeepkumar, P. I.; Hobartner, C.; Baum, D. A.; Silverman, S. K., DNA-catalyzed formation of nucleopeptide linkages. *Angew. Chem., Int. Ed.* **2008**, *47*, 1753–1757.

47 Wang, Y.; Silverman, S. K., Efficient one-step synthesis of biologically related lariat RNAs by a deoxyribozyme. *Angew. Chem., Int. Ed.* **2005**, *44*, 5863–5866.

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

49 Chinnapen, D. J. F.; Sen, D., A deoxyribozyme that harnesses light to repair thymine dimers in DNA. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 65–69.

50 Li, Y.; Liu, Y.; Breaker, R. R., Capping DNA with DNA. *Biochemistry* **2000**, *39*, 3106–3114.

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

52 Chandra, M.; Sachdeva, A.; Silverman, S. K., DNA-catalyzed sequence-specific hydrolysis of DNA. *Nat. Chem. Biol.* **2009**, *5*, 718–720.

53 Li, Y. F.; Sen, D., A catalytic DNA for porphyrin metallation. *Nat. Struct. Biol.* **1996**, *3*, 743–747.

54 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. *Proc. Nat.l Acad. Sci. U. S. A.* **2007**, *104*, 2056–2061.

55 Baum, D. A.; Silverman, S. K., Deoxyribozyme-catalyzed labeling of RNA. *Angew. Chem., Int. Ed.* **2007**, *46*, 3502–3504.

56 Walsh, C., *Posttranslational modification of proteins: expanding nature's inventory*. **2006** Roberts and Co. Publishers.

57 Egan, S. E.; Weinberg, R. A., The pathway to signal achievement. *Nature* **1993**, *365*, 781–783.

Chapter 2: DNA-Catalyzed Covalent Modification of Amino Acid Side Chains in Tethered and Free Peptide Substrates[a]

2.1 Introduction

In this project, we performed a comprehensive in vitro selection study to identify deoxyribozymes that covalently modify amino acid side chains, with the long-term goal of achieving DNA-catalyzed covalent protein modifications. Using small-molecule reagents to covalently modify protein side chains is challenging because they are not site-specific.[1-3] On the other hand, engineering protein enzymes to obtain desired functions may be difficult.[4]

In 2002, Baskerville and Bartel reported a ribozyme that catalyzes ligation of the N-terminus of a polypeptide to the RNA.[5] However, catalysis using potentially reactive amino acid side chains as nucleophiles has remained a challenge for nucleic acid catalysts. As a first step toward this goal, our laboratory has previously reported deoxyribozyme, Tyr1[6], which covalently modifies a tyrosine side chain to form a nucleopeptide linkage (**Fig. 2.1**).

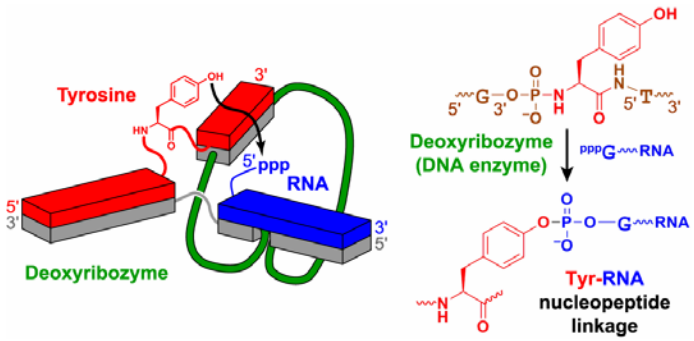


Figure 2.1 The Tyr1 deoxyribozyme catalyzes ligation of the tyrosine side chain to RNA.

[a] The material described in this chapter has been published.
Wong, O. Y.; Pradeepkumar, P. I; Silverman, S. K., DNA-Catalyzed Covalent Modification of Amino Acid Side Chains in Tethered and Free Peptide Substrates. *Biochemistry* **2011**, *50*, 4741–4749.

Examples of covalent connections between the protein and RNA are found in biology, exemplified by the RNA genome of poliovirus, in which it is covalently linked to a viral protein.[5-15] In order to achieve the proper orientation for catalysis, the Tyr1 deoxyribozyme requires that the Tyr nucleophile to be embedded in two DNA oligonucleotide segments (**Fig. 2.1**). The broad implication of this result is that deoxyribozymes are able to use the side chain of an amino acid as a nucleophile. To circumvent this substrate requirement, this project sought deoxyribozymes that catalyze amino acid side chain reactivity in the context of a more open architecture, which requires that the deoxyribozyme simultaneously interact with and spatially organize the peptide substrate to achieve catalysis (**Fig. 2.2**).

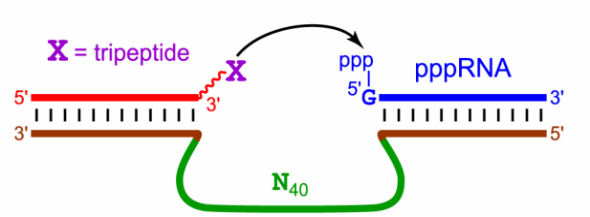


Figure 2.2 The selection efforts reported in this chapter sought deoxyribozymes that catalyze amino acid side chain reactivity in the context of a more open architecture.

More importantly, such an open architecture could be amenable to catalytic function with large protein substrates. To expand the catalytic repertoire of deoxyribozymes, the Silverman laboratory is addressing the question of whether deoxyribozymes can achieve catalysis of all potential nucleophilic side chains of polypeptides. To answer this, we began with investigating the capability of deoxyribozymes in catalyzing reactions using tripeptide substrates. We performed selections using tripeptide substrates, Cys-Xaa-Ala, where Xaa denotes the potentially reactive amino acid Tyr, Ser, or Lys (**Fig. 2.3**). The nucleophile was presented with less preorganization by varying presentations of the tripeptide substrate to DNA in the open architecture. This chapter focuses on the in vitro

selection of deoxyribozymes that catalyze the ligation of the potentially reactive amino acid side chain in a tripeptide to the 5′-terminus of RNA in a non-preorganized format.

2.2 Results and Discussion

2.2.1 Selection experiments using tethered tripeptides (KC Selection)

We began with selections using tripeptide substrates, Cys-Xaa-Ala, where Xaa denotes the potentially reactive amino acid Tyr, Ser, or Lys. Also, we include a negative control (Xaa = Ala, lacking any nucleophile). The substrates each had a short three-carbon tether linking the 3′-terminal oxygen atom of the DNA anchor to the cysteine of the tripeptide via a disulfide bond, denoted DNA-C_3-CXA (**Fig. 2.3**). Along with the negative control, we also include a positive control, called the dT selection (all DNA substrate, 3′-hydroxyl as the nucleophile). The DNA-C_3-CXA conjugates were prepared by reaction with a DNA oligonucleotide with a 3′- terminal thiol and pyridyl disulfide activated Cys-Xaa-Ala substrates (**Fig. 2.4**, see Materials and Methods for preparation details).

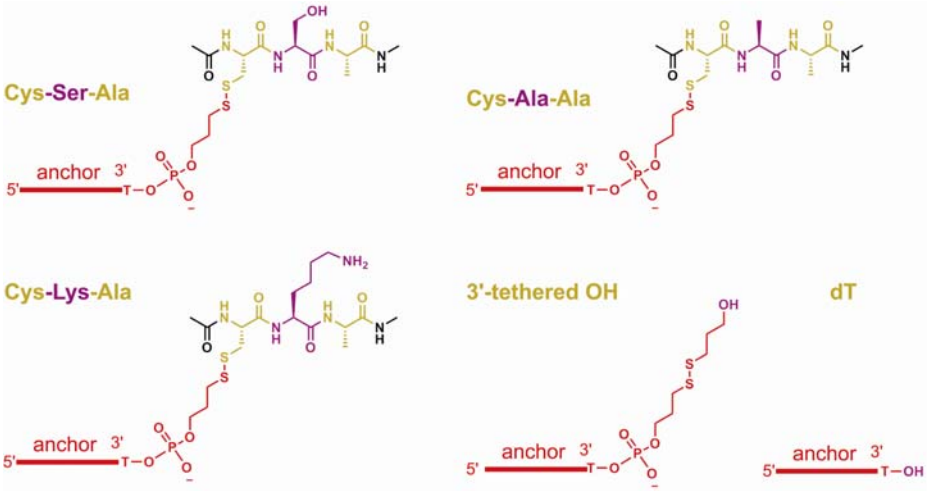


Figure 2.3 DNA-C_3-Cys-Xaa-Ala substrates with Xaa = Ser, Tyr or Lys as the corresponding potential nucleophile. dT and Ala substrates were used as the positive and negative control selection, respectively. A 3′-tethered OH substrate was used as well.

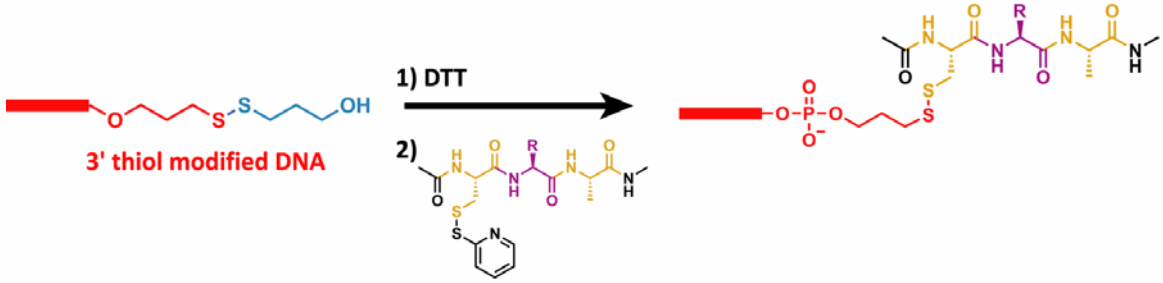


Figure 2.4 DNA-Cys-Xaa-Ala conjugation reactions. The 3′-thiol modified DNA was first treated with DTT to produce a 3′-terminal thiol, followed by reaction with a pyridyl disulfide-activated Cys-Xaa-Ala.

To initiate selections, a pool of N_{40} DNA was ligated to the 5′-triphosphate of the RNA substrates by T4 RNA ligase. In the key selection step, we incubated the DNA-tripeptide with the 5′-triphosphate of the RNA substrates in 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC for 2 h. The product was isolated by PAGE and subsequently amplified by PCR to go on to the next round. The amplified material containing deoxyribozymes with the intended catalytic activity was then ligated to the 5′-triphosphate of the RNA substrates. We performed the next selection step by incubating the ligated product with the DNA-tripeptide. The process was iterated for several rounds to enrich the population of deoxyribozymes with the ligation activity.

2.2.2 KC selection progression

For the dT positive control, we observed detectable DNA-catalyzed ligation between the 3′-OH and the 5′-triphosphate of the RNA in round 6. The pool activity of nucleopeptide-ligating deoxyribozymes was increased to 75% at round 11. The dT positive control validated the selection efforts were preformed properly. In the Cys-Tyr-Ala selection, we first observed ligation activity in round 8. The ligation activity reached a plateau at 47% in round 10 (**Fig. 2.5**). To select for fast deoxyribozymes, we applied a time pressure in the selection step by lowering the incubation time from 2 h to 10 min.

However, after three rounds at increased time pressure, the activity was not increased. No activity was observed for the Cys-Ser-Ala, Cys-Lys-Ala, or 3′-tethered OH selections.

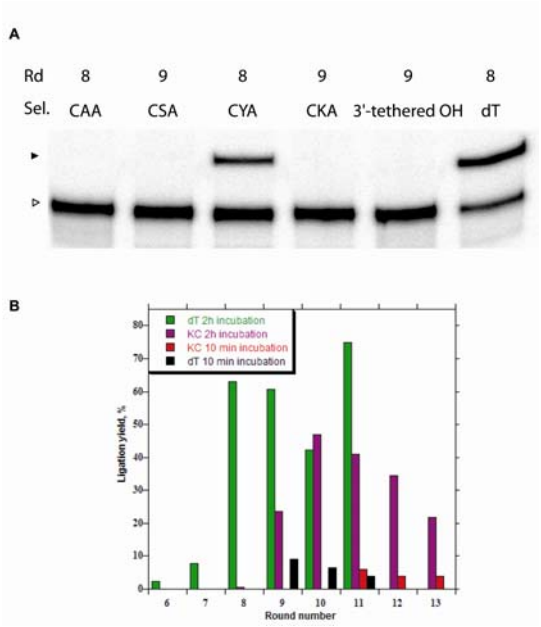


Figure 2.5 KC selection progression. (A) 8% PAGE image of key rounds for the individual selection experiments with DNA-tripeptide substrates. The open triangle marks the DNA-tripeptide substrate, and the filled triangle marks the ligation product. (B) Selection progression.

2.2.3 Cloning and sequencing of KC deoxyribozymes

We surveyed the clones of the DNA-C_3-CYA selection in round 10 when activity was the most robust. Of the 24 clones surveyed, 22 showed catalytic activity in a trimolecular format, in which the DNA-C_3-CXA and 5′-triphophate RNA were not connected to the DNA enzyme. **Fig. 2.6** shows the results from the kinetic screening. We designated KC to represent the DNA-C_3-CYA selection. Individual deoxyribozymes from round 10 of the KC pool were cloned.

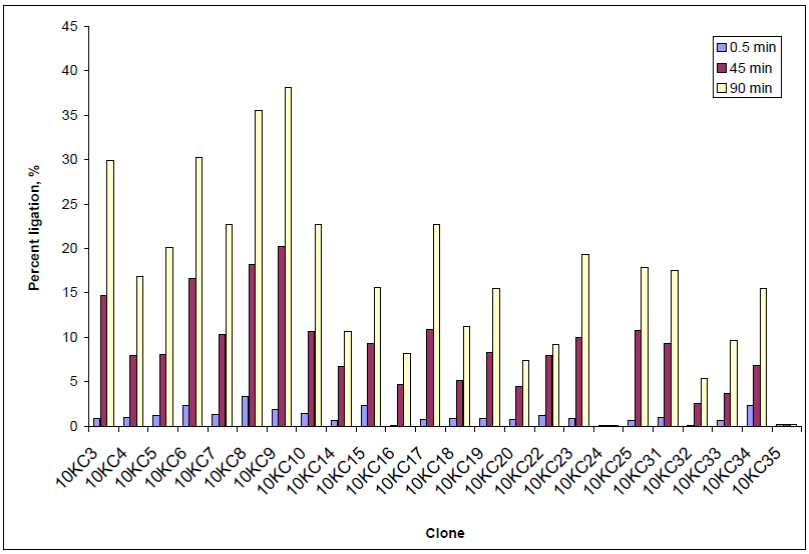


Figure 2.6 Kinetic screening of individual KC clones. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. Time point: 30 s, 45 min, 90 min. From the results of this assay, 10KC clones 3, 6, 8, 9, 10, 17, 19, 23, 31 were chosen for sequencing.

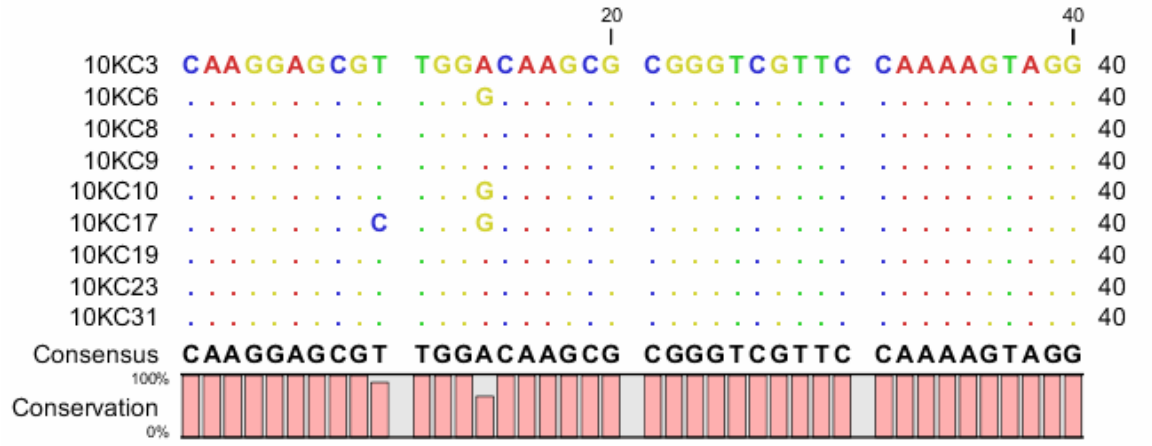


Figure 2.7 Sequence alignment of 10KC sequences. A single dominant sequence, 10KC3, was identified.

Sequences were obtained for 10 of the active clones. The 10 sequences were in consensus, except two transitions at position 10 and 14 were observed (**Fig. 2.7**). It is likely that the transitions were non-essential to the catalytic activity. Thus, we obtained a single sequence 10KC3 (5′-CAAGGAGCGTTGGACAAGCGCGGGTCGTTCCAAAAGTAGG-3′) for a detailed characterization.

To investigate the capability of the 10KC3 deoxyribozyme to use the Cys-Tyr-Ala (CYA) substrate remotely placed from the anchor DNA, we assayed the 10KC3 deoxyribozyme in three different formats: C_3-tethered, HEG-tethered, and untethered CYA (**Fig. 2.8**). The three formats present the Tyr nucleophile in different spatial orientations. In the HEG-tethered format, we placed a hexa(ethylene glycol) (HEG) linker between the tripeptide and the anchor DNA (DNA-HEG-CYA). To prepare DNA-HEG-CYA, hexa(ethylene glycol) spacer phosphoramidite was first coupled to the 3′-thiol modifier solid support during solid-phase synthesis. After the solid-phase synthesis, DNA-HEG with a 3′-thiol modifier was reduced and coupled to the PyS-S-Cys-Tyr-Ala to yield a DNA-HEG-Cys-Tyr-Ala conjugate (see Materials and Methods for details). We tested the DNA-HEG-Cys-Tyr-Ala as a substrate for the 10KC3 catalysis. 10KC3 was able to catalyze the ligation of the DNA-HEG-Cys-Tyr-Ala and the 5′-triphosphate of the RNA (**Fig. 2.8B**). In the untethered CYA format, the CYA is not linked to the DNA anchor, and the cysteine thiol function group is terminated with a *tert*-butyl disulfide (**Fig. 2.8C**).

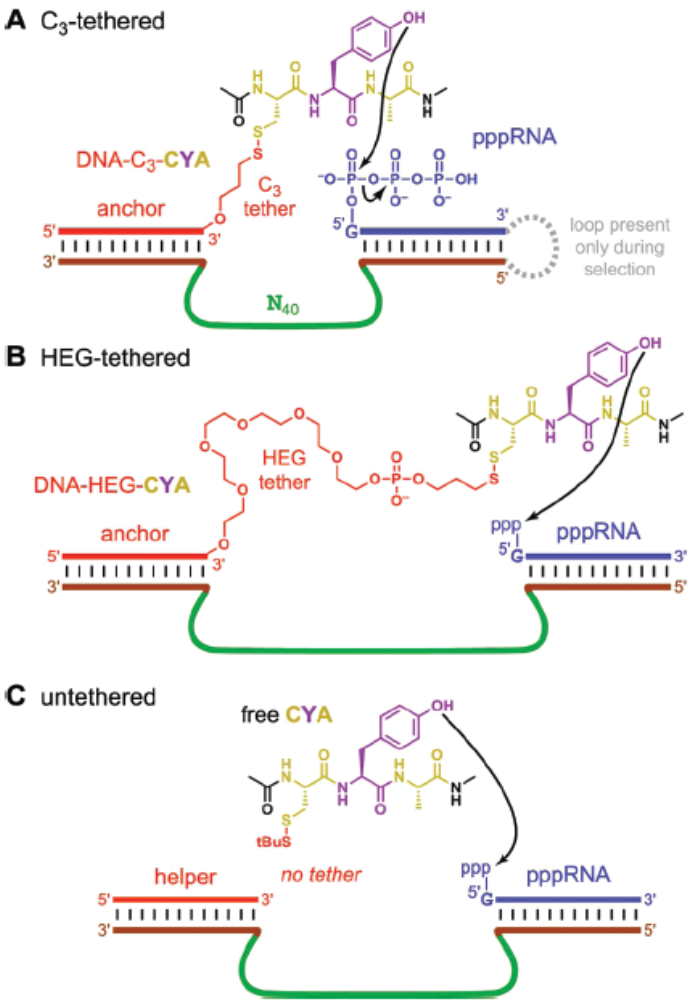


Figure 2.8 Various spatial presentations of the Tyr nucleophile. (A) In DNA-C_3-tethered substrates, the tripeptide was connected to the DNA anchor via a propyl linker. (B) In DNA-HEG-tethered substrates, the tripeptide was covalently attached to the DNA anchor with a hexa(ethylene glycol) linker. (C) In free CYA substrates, the cysteine thiol moiety was terminated with *tert*-butyl disulfide.

10KC3 showed excellent reactivity toward C_3-tethered and HEG-tethered CYA substrates with 70% yield (k_{obs} = 0.30 h^{-1}) and 62% (k_{obs} = 0.10 h^{-1}) yield in 60 h, respectively. However, 10KC3 displayed modest activity toward the free CYA substrates with ~10% yield at 60 h (k_{obs} = 0.0017 h^{-1}, linear fit; **Fig. 2.9**).

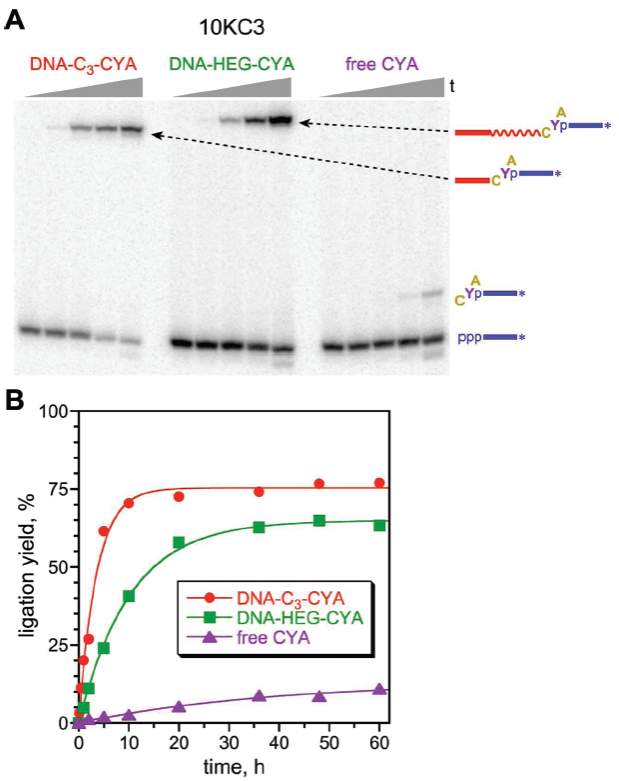


Figure 2.9 Kinetic assay of the 10KC3 deoxyribozyme. (A) 8% PAGE image of kinetic assay with DNA-C3-CYA, DNA-HEG-CYA, and free CYA. Time points were taken at 30 s, 10 min, 2 h, 10 h and 60 h. (B) Kinetic plot of the data from panel A. The data was fitted using first-order kinetics, i.e. yield = $Y \cdot (1 - e^{-kt})$, where $k = k_{obs}$. See Materials and Methods for kinetic assay procedures.

Despite the low overall yield of covalently modification of free CYA catalyzed by the 10KC3 deoxyribozyme, the results demonstrate that deoxyribozymes can use peptide substrates that are placed remotely away from the anchor DNA. Therefore, further optimization of 10KC3 and new selection designs are required to achieve activity toward free CYA.

2.2.4 Follow-up selection with the 10KC3 25% partially randomized pool using DNA-C_3-CYA (NG selection)

The identification of 10KC3 established the potential of deoxyribozymes to catalyze covalent modification of the Tyr hydroxyl of a free CYA substrate. The initial KC

selection effort covered only 10^{-10} sequence space (we started with a pool of 10^{14} deoxyribozyme out of 10^{24} possible sequences). We reasoned that by exploring the sequence space that is closely related to the original 10KC3 sequence would allow us to obtain deoxyribozymes that displayed improved activity toward the free CYA substrate. To explore sequence space that is closely related to the 10KC3 deoxyribozyme, we designed a selection with DNA-C_3-CYA substrate using the DNA pool with the enzyme region partially randomized to the extent of 25% per nucleotide position (i.e., 75% related to the original 10KC3 sequence). The average number of mutations was 10 out of 40 nucleotides as a result of the partial randomization. We denoted the selection effort using the 25% partially randomized pool from the 10KC3 sequence as the NG selection.

2.2.5 NG Selection progression

We observed a detectable activity with the RNA substrate in round 2 with 2% yield. The pool activity of nucleopeptide-ligating deoxyribozymes was increased to 57% in round 6 (**Fig. 2.10**). To select for faster deoxyribozymes, we applied a time pressure in the selection step by lowering the incubation time from 2 h to 10 min in round 7. However, after four rounds at increased time pressure, no enrichment was observed as the activity remained at ~26%. We applied an additional time pressure by decreasing the incubation time from 10 min to 1 min starting at round 9 and observed 5% activity. We continued the selection with 1 min incubation for an additional round and observed no enrichment of activity (8 %). The pool of deoxyribozymes was cloned and sequenced at round 9 with 1 min incubation.

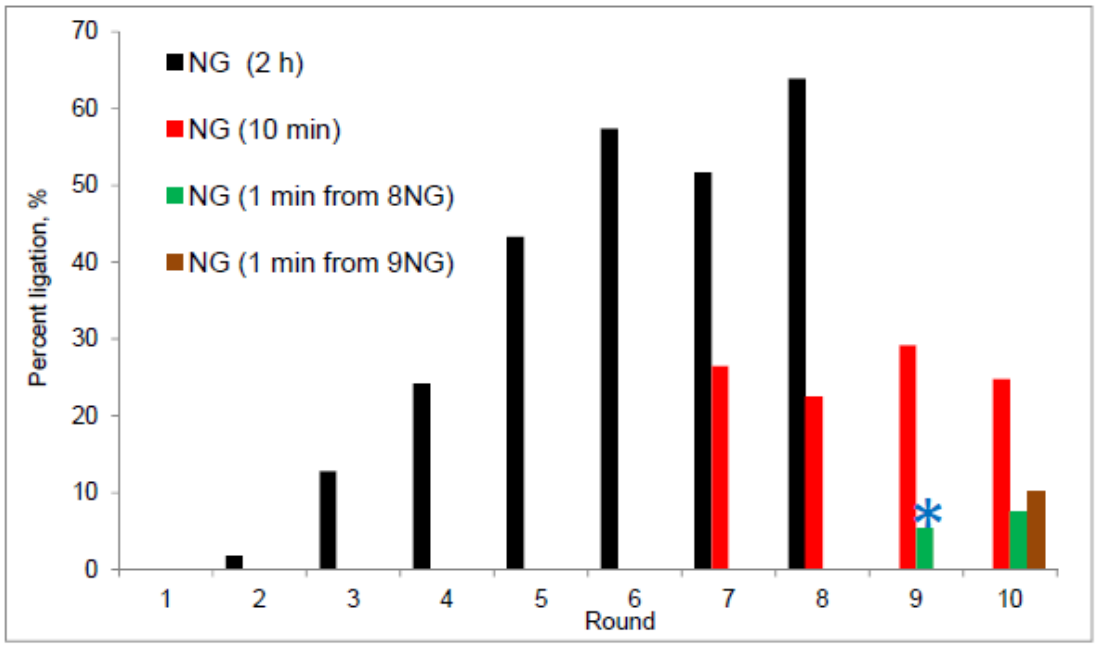


Figure 2.10 NG selection progression. The NG selection was initiated with the 10KC3 25% partially randomized pool using the DNA-C_3-CYA substrate. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. To identify fast deoxyribozymes, a time pressure was applied. Beginning in round 6, the reaction was incubated at 10 min instead of the default 2 h incubation. In addition, the effect of the timing for applying the time pressure was studied. We compared the effect of applying the time pressure from 10 min to 1 min starting in round 8 or in round 9. It was found that applying the time pressure at different rounds in a rather late stage of selection yielded comparable results. After round 9 at 1 min incubation, the pool of deoxyribozymes was cloned and sequenced. The blue asterisk marks the round and activity where the NG selection was cloned and sequenced.

2.2.6 Cloning and sequencing of NG deoxyribozymes

Out of all the 22 clones surveyed (**Fig. 2.11**), 11 of the clones were sequenced. From the sequence alignment, three conserve motifs were observed (**Fig. 2.12**).We synthesized six sequences (9NG2, 3, 5, 6, 14, 15) for further kinetic studies.

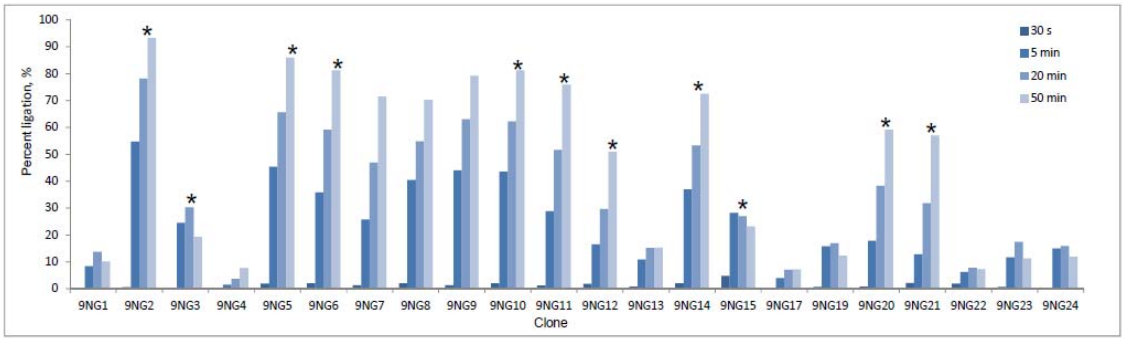


Figure 2.11 Kinetic screening of individuals 9NG clones. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. Time points: 30 s, 5 min, 20 min, 50 min. From the results of this assay, 9NG clones 2, 3, 5, 6, 10, 11, 12, 14, 15, 20, 21 were chosen for sequencing. The asterisks highlight the clones that were chosen for sequencing.

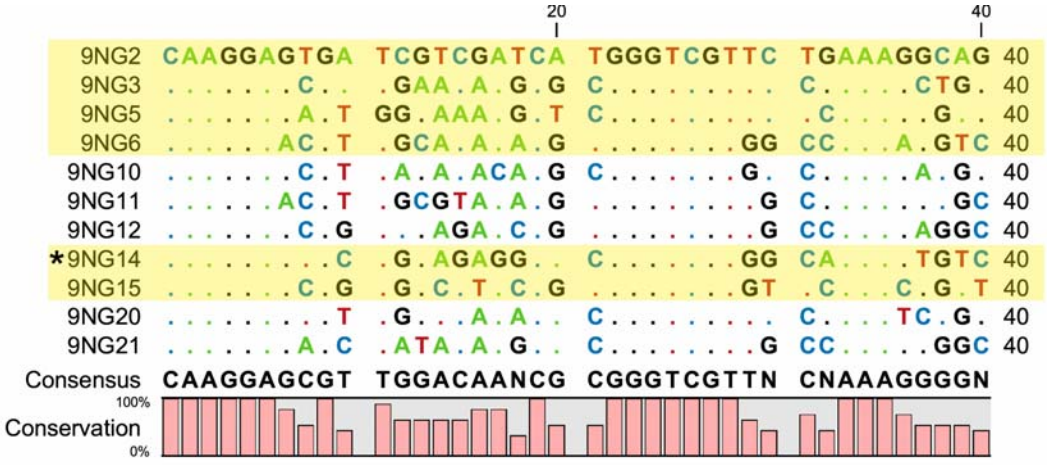


Figure 2.12 Sequence alignment of 9NG sequences. Highlighted sequences were synthesized by DNA solid-phase synthesis for kinetic characterizations. 9NG14, highlighted by an asterisk, was a representative deoxyribozyme from the 9NG selection effort.

9NG14, a representative deoxyribozymes identified in the NG selection, showed excellent reactivity toward C_3-tethered and HEG-tethered CYA substrates with 77% yield (k_{obs} = 5.6 h^{-1}) and 75% yield (k_{obs} = 1.6 h^{-1}), respectively. 9NG14 displayed a 20-fold increase in k_{obs} with the C_3-tethered substrates compared to the 10KC3 deoxyribozyme. For the HEG-tethered, a 16-fold increase in k_{obs} compared to the 10KC3 deoxyribozyme that was observed. 9NG14 was not selected to work with the free CYA substrate yet

9NG14 showed great improvement toward the free CYA with 40% yield in 60 h (k_{obs} = 0.063 h^{-1}, $t_{1/2}$ of 11h, ~30-fold faster and with a 4-fold improvement in yield compared to 10KC3(**Fig. 2.13**). Five other deoxyribozymes were also studied (9NG2, 3, 5, 6, 15).They displayed similar kinetic profiles as 9NG14.

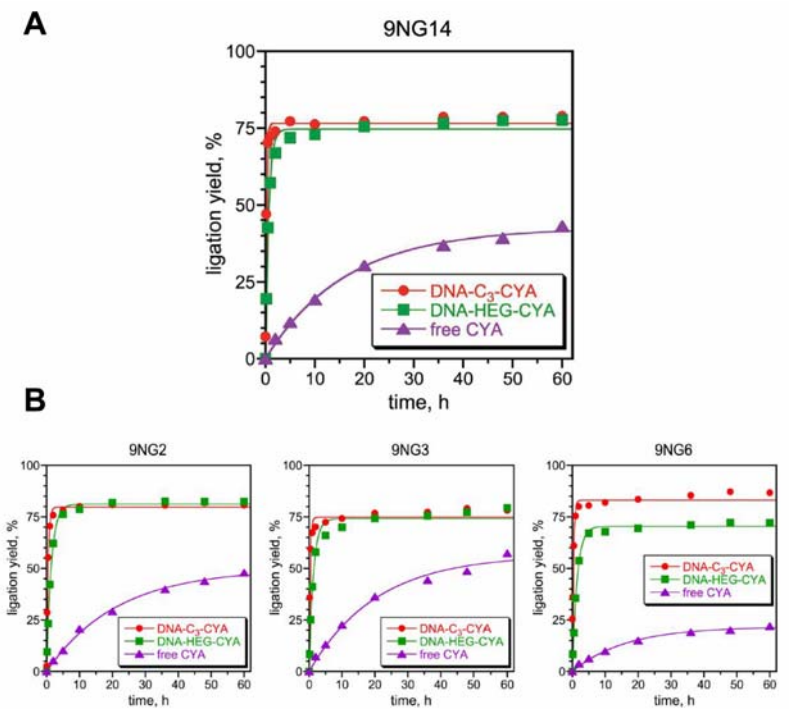


Figure 2.13 Kinetic assays of 9NG deoxyribozymes. (A) Kinetic plot of 9NG14, one of the representative deoxyribozymes from the NG selection. (B) Kinetic plot of 9NG2, 3 and 6. The data was fitted using the first-order kinetics equation. 9NG5 and 15 also showed similar activities (data not shown).

2.2.7 Selection with DNA-HEG-CYA (MN selection) and selection progression

We performed selection with DNA-HEG-CYA using unbiased N_{40} pool, denoted as the MN selection. The HEG tether is longer than the C_3 tether. Therefore, the HEG tethered substrate imposes a more spatially challenging requirement for the deoxyribozymes to achieve catalysis compared to the C_3 tethered-substrates. We expected that the deoxyribozymes from the new DNA-HEG-CYA selection effort would

yield deoxyribozymes that function better with free CYA substrate than from the KC selection effort.

The MN selection first showed discernable activity in round 7 and the activity was increased to 38% at round 10 and reached its plateau in round 11 (46%). To select for fast enzymes, we applied a 10 min incubation time pressure starting in round 12 (**Fig. 2.14**). However, with the applied time pressure for three addition rounds, the activity remained at ~12%. The deoxyribozymes from the MN pool in round 11 with 2 h incubation were cloned.

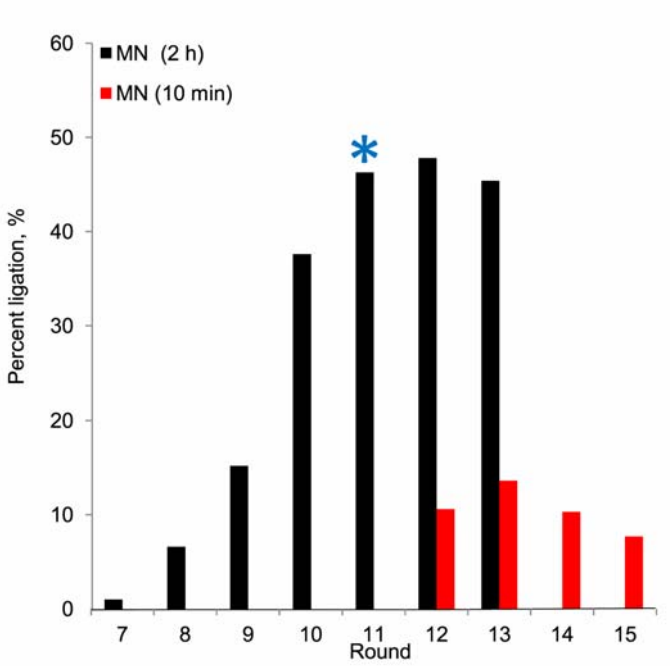


Figure 2.14 MN selection progression. The MN selection was initiated with the unbiased N_{40} pool using the DNA-HEG-CYA substrate. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. To identify fast deoxyribozymes, time pressure was applied. Starting from round 12, the reaction was incubated at 10 min instead of the default 2 h incubation. Three additional selection rounds with the applied time pressure did not enrich activity. After round 11 with 2 h incubation, deoxyribozymes were cloned and sequenced. The blue asterisk denotes the round and activity where the deoxyribozymes were cloned and sequenced.

2.2.8 Cloning and sequencing of MN deoxyribozymes and their kinetics

Round 11 of the MN selection was cloned. We surveyed 11 clones and 9 of them were active without including two clones which had less than 6% activity (**Fig. 2.15**). Of the 9 actives clones, 6 of the clones were sequenced (**Fig. 2.16**). 11MN5, a representative deoxyribozymes identified in the NG selection, showed excellent reactivity toward C_3-tethered and HEG-tethered CYA substrates with 77% yield (k_{obs} = 3.1 h^{-1}) and 76% yield (k_{obs} = 0.86 h^{-1}), respectively. 11MN5 displayed a 20-fold increase in k_{obs} with the C_3-tethered substrates compared to the 10KC3 parent. For the HEG-tethered, a 16-fold increase in k_{obs} compared to the 10KC3 deoxyribozyme was observed. 11MN5 was not selected to work with the free CYA substrate. Remarkably, 11MN5 showed a significant improvement toward the free CYA with 55% yield in 60 h (k_{obs} = 0.05 h^{-1}, $t_{1/2}$ of 14 h, ~30-fold faster and 5 fold improvement in yield compared to 10KC3, **Fig. 2.17A**). Two other deoxyribozymes were also studied (11MN10, 11MN19).They displayed similar kinetic profiles as that of 11MN5 (**Fig. 2.17B**).

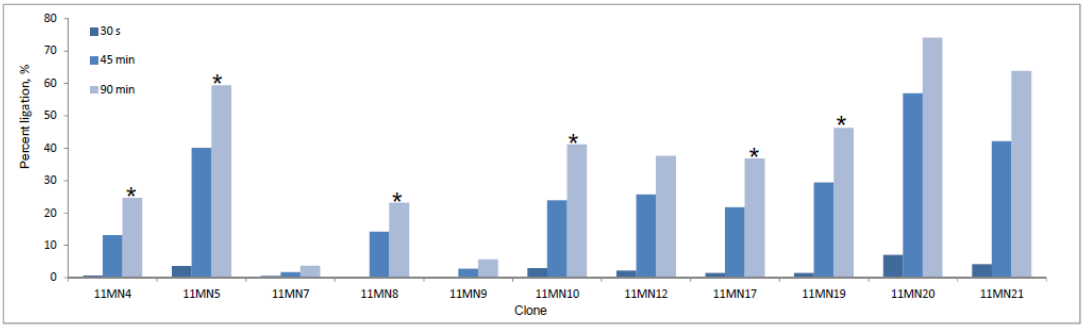


Figure 2.15 Kinetic screening of individual 11MN clones. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. Time points: 30 s, 45 min, 90 min. From the results of this assay, 11MN clones 4, 5, 7, 8, 9, 10, 12, 17, 19, 20, 21 were chosen for sequencing. The asterisks highlight the clones that were chosen for sequencing.

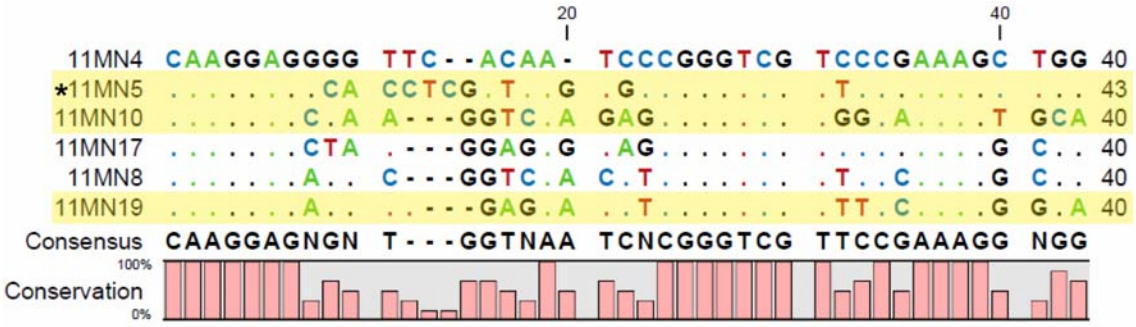


Figure 2.16 Sequence alignment of 11MN sequences. Highlighted sequences were prepared by solid-phase synthesis for kinetic characterization. 11MN5, marked by an asterisk, was a representative deoxyribozyme from the MN selection effort.

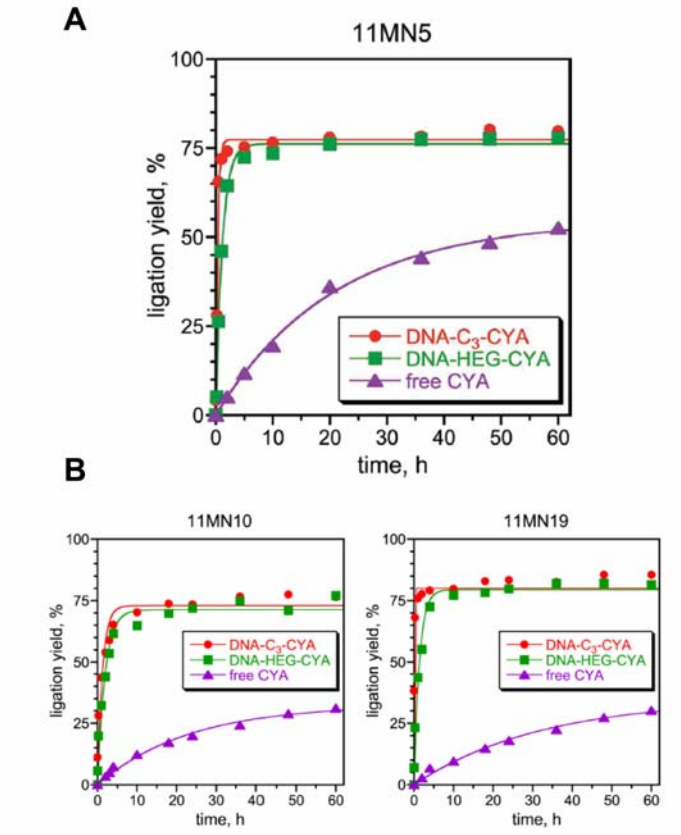


Figure 2.17 Kinetic assays of 11MN deoxyribozymes. (A) Kinetic plot of 11MN5. 11MN5 was one of the representative deoxyribozymes from the MN selection. (B) Kinetic plot of 11MN10 and 11MN19. The kinetic profile of 11MN10 was as follows: with C_3-tethered CYA (73%, k_{obs} = 0.81 h^{-1}), HEG-tethered CYA (73%, k_{obs}= 0.81 h^{-1}), free CYA (32%, k_{obs} = 0.044 h^{-1}). The kinetic profile of 11MN19 was as follows: with C_3-tethered CYA (80%), HEG-tethered CYA (80%, k_{obs} = 0.73 h^{-1}), free CYA (35%, k_{obs} = 0.032 h^{-1}).

2.2.9 Selection using DNA-HEG-CYA (MP selection, redirection from 4KC)

We performed a selection with DNA-HEG-CYA by redirecting from the round 4 of the KC selection, denoted as the MP selection. In the MP selection, we started with the pool of uncloned deoxyribozymes from the round 4 of the DNA-C3-CYA (KC) selection, and redirected the pool to catalyze reaction involving the DNA-HEG-CYA substrate. As part of the comprehensive effort to study the optimal strategy of in vitro selection for DNA–catalyzed covalent modification of free CYA, we reasoned that the redirection using the 4KC pool with HEG-tethered CYA could yield deoxyribozymes that function better with free CYA. Second, the MP selection effort would provide us insights on whether redirection is a viable strategy for future selection efforts.

Fig. 2.18 depicts the MP selection progression. The MP selection first showed detectable activity at round 8 and the activity was increased to 23% in round 11 and reached its plateau in round 12 (36%). A 10 min incubation time pressure was applied starting in round 12. However, with the applied time pressure for three additional rounds, the activity remained at ~11%. The MP pool in round 12 with 2 h incubation (36% activity) was cloned.

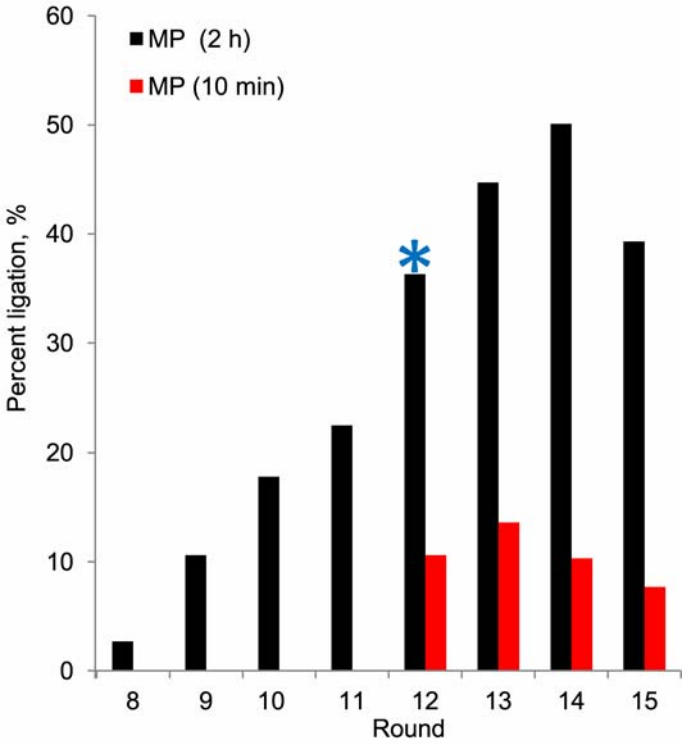


Figure 2.18 MP selection progression. The MP selection was initiated with 4KC pool using the DNA-HEG-CYA substrate. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. To identify fast deoxyribozymes, a time pressure was applied. Starting in round 12, the reaction was incubated at 10 min instead of the default 2 h incubation. Three additional selection rounds with the applied time pressure did not enrich activity. In round 12 with 2 h incubation, deoxyribozymes were cloned and sequenced. The blue asterisk denotes the round and activity (36%) where deoxyribozymes were cloned and sequenced.

2.2.10 Cloning and sequencing of MP deoxyribozymes and their kinetics

Of the 13 clones surveyed, 12 of them were active (**Fig. 2.19**). Sequences were obtained for 10 of the active clones (**Fig. 2. 20**). Comparison of the MP sequences reveals that 12MP4, 8, 16, and 20 are identical. Also, 12MP5, 19, 17 are identical. 12MP5 and 12MP10 differ by one nucleotide; 12MP5 and 12MP15 also differ by one single nucleotide. 12MP4, 12MP9 (identical to 12MP5), and 12MP11 were prepared by solid-phase synthesis for further kinetic analyses.

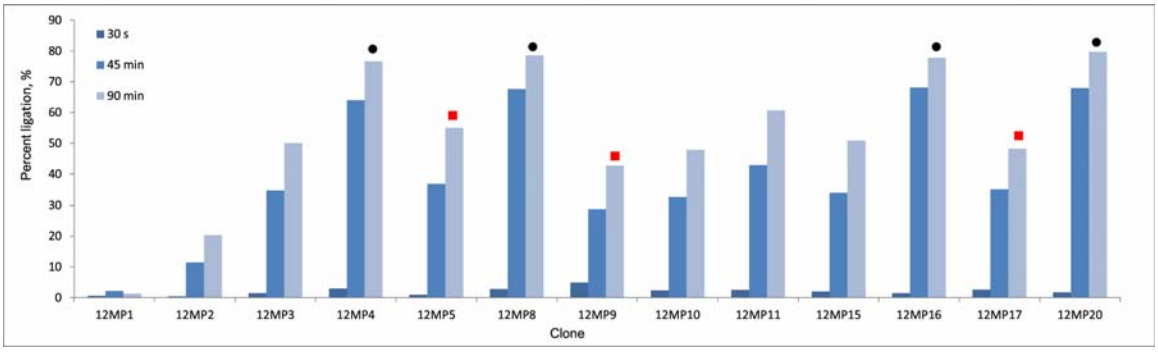


Figure 2.19 Kinetic screening of individuals 12MP clones. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. Time points: 30 s, 45 min, 90 min. From the results of this assay, all of the 12 were chosen for sequencing. The black circles denote identical sequences (12MP4, 8, 16, 20). The red rectangles also mark identical sequences (12MP5, 9, and 17).

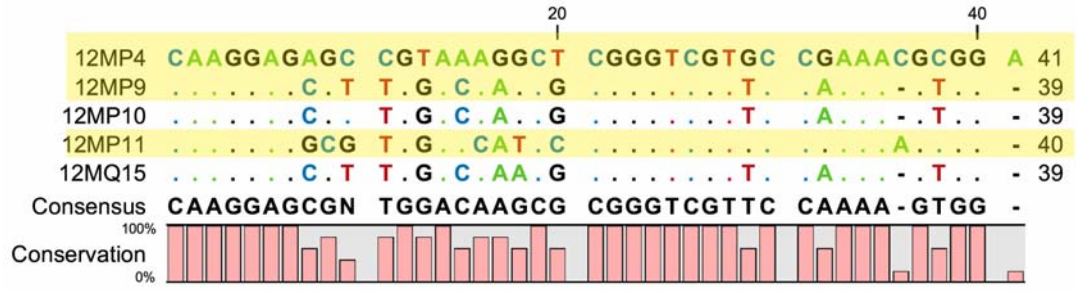


Figure 2.20 Sequence alignment of 12MP sequences. Highlighted sequences were prepared by solid-phase synthesis for kinetic characterization.

The activity of 12MP4, 12MP9, and 12MP11 were assayed with DNA-HEG-CYA. 12MP4, 12MP9 and 12MP11 showed excellent reactivity with the parent DNA-HEG-CYA substrate in 24 h incubation with 70% yield (k_{obs} =1.6 h^{-1}), 86% yield (k_{obs} = 0.44 h^{-1}) and 91% (k_{obs} = 0.66 h^{-1}), respectively (**Fig. 2.21**). Their activities toward free CYA were also assayed. 12MP4, 12MP9, 12MP11 catalyze the free CYA reactivity with 40% yield, 29% yield, and 34% yield, respectively. The detailed kinetic profile of 12MP4, 9 and 11 were not determined. Compared to the 10KC3 deoxyribozyme, 12MP4 was 16-fold faster; 12MP9 was 4-fold faster and 12MP11 showed 7-fold improvement in rate compared to the 10KC3 deoxyribozyme. From these kinetic data, 12MP deoxyribozymes showed a comparable rate and yield as the 11MN5 deoxyribozyme.

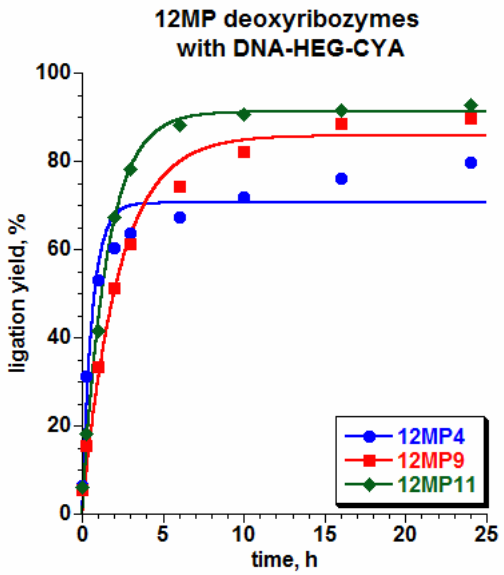


Figure 2.21 Kinetic assays of 12MP deoxyribozymes. 12MP4, 12MP9, 12MP11 were assayed with the parent substrate DNA-HEG-CYA. The data was fitted using first-order kinetics.

2.2.11 Reselection using DNA-HEG-CYA (MQ selection, partially 25% randomized 10KC3)

We performed reselection with DNA-HEG-CYA from the partially randomized 10KC3 pool, denoted as the MQ selection. We expected that the reselection using the partially randomized 10KC3 pool with HEG-tethered CYA would yield deoxyribozymes that function well with free CYA. The MQ selection first showed detectable activity in round 2 and the activity was increased to 29% in round 5 and reaches its plateau in round 7 (47%). To select for fast enzymes, we applied a 10 min incubation time pressure starting in round 8 (**Fig. 2.22**). However, with the applied time pressure for three addition rounds, the activity remained at ~14%. The deoxyribozymes from the MP pool in round 8 with 10 min incubation (13% activity) were cloned.

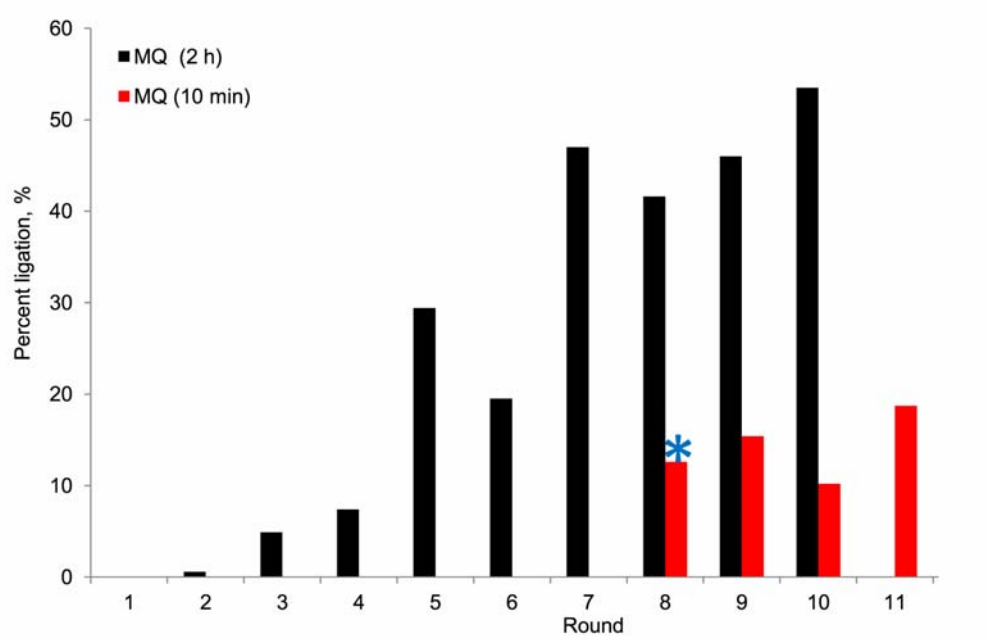


Figure 2.22 MQ selection progression. The MQ selection was initiated with the 25% 10KC3 partially randomized pool using the DNA-HEG-CYA substrate. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. To identify fast deoxyribozymes, time pressure was applied. Starting at round 8, the reaction was incubated at 10 min instead of the default 2 h incubation. Three additional selection rounds with the applied time pressure did not enrich activity. The pool of deoxyribozymes was cloned and sequenced at round 8. The blue asterisk denotes the round and activity (13%) where the MQ selection was cloned and sequenced.

2.2.12 Cloning and sequencing of MQ deoxyribozymes and their kinetics

Of the 15 clones surveyed, 13 of them were active (**Fig. 2.23**). Sequences were obtained for seven of the clones. 8MQ5, 8MQ7, and 8MQ11 were prepared by DNA solid-phase synthesis for further kinetic analyses (**Fig. 2.24**). The activity of 8MQ5, 8MQ7, and 8MQ11 were assayed with DNA-HEG-CYA. 8MQ5, 8MQ7, and 8MQ11 show excellent reactivity with the parent DNA-HEG-CYA substrate after 24 h incubation with 79% yield (k_{obs} = 0.20 h^{-1}), 72% yield (k_{obs} = 0.72 h^{-1}) and 76% (k_{obs} = 0.37 h^{-1}), respectively (**Fig. 2.25**). Their activities toward free CYA were also studied. 8MQ5, 8MQ7, and 8MQ11 catalyzed the free CYA reactivity in 60 h incubation with 36% yield, 41% yield, and 21% yield, respectively. Their detailed kinetic profiles were not determined. Compared to the 10KC3 deoxyribozyme, 8MQ5 reacted with HEG-CYA at a similar rate; 8MQ7 was 7-fold faster and 8MQ11 showed 4-fold improvement in rate

compared to the 10KC3 deoxyribozyme. From these kinetic data, 8MQ deoxyribozymes show comparable rate and yield as the 11MN5 deoxyribozyme.

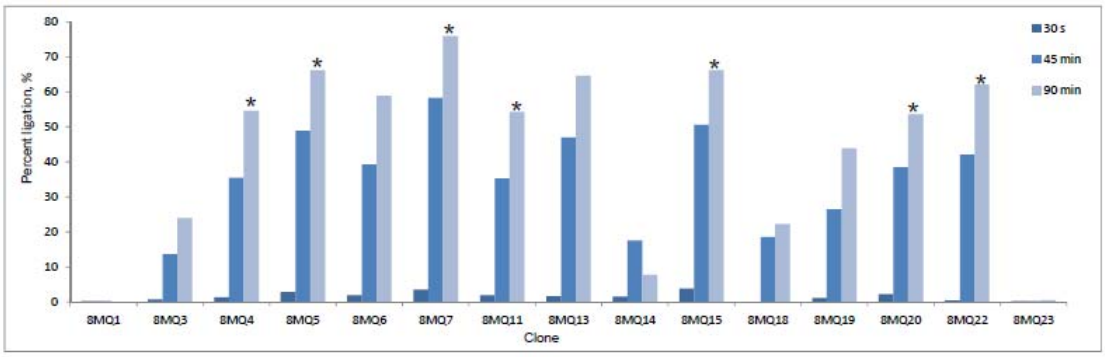


Figure 2.23 Kinetic screening of individual 8MQ clones. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. Time point: 30 s, 45 min, 90 min. Sequences were obtained and highlighted by the asterisks.

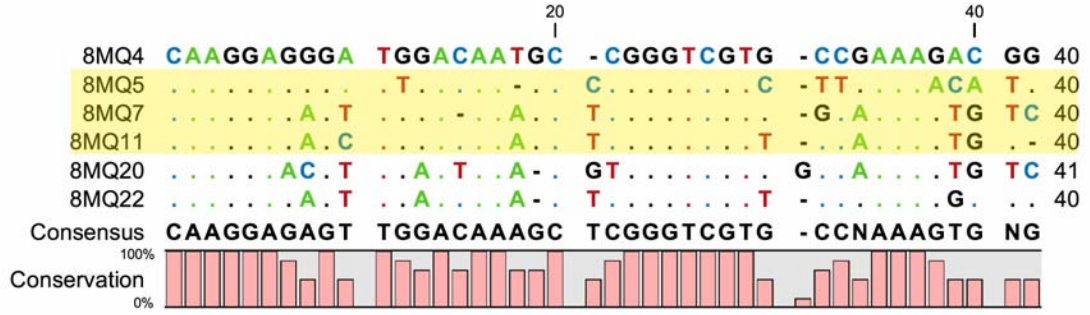


Figure 2.24 Sequence alignment of 8MQ sequences. Highlighted sequences were prepared by solid-phase synthesis for kinetic characterization.

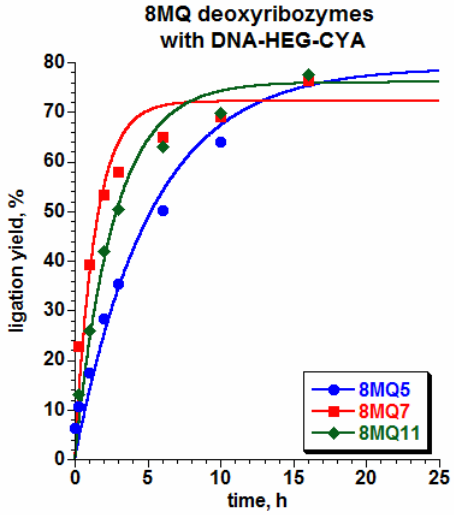


Figure 2.25 Kinetic assays of 8MQ deoxyribozymes. 8MQ5, 8MQ7 and 8MQ11 were assayed with the parent substrate DNA-HEG-CYA. The data was fitted using first-order kinetics.

2.2.13 Seeking serine reactivity: alternative selection approach with a tethered hydroxyl substrate (MY to MZ selection) and selection progression

All of the previous experiments have shown promising DNA-catalyzed reactivity toward the Tyr side-chain. Previous results from this laboratory have also demonstrated that the Ser side chain is much more challenging than analogous reactivity of Tyr, even in a highly preorganized architecture. Because the selection effort with the DNA-C_3-CSA substrate gave no activity, we pursued an alternative approach. We initiated a selection effort with a DNA substrate conjugated at its 3′-terminus to a tri(ethylene glycol) (TEG) tether that terminates in an aliphatic hydroxyl group (DNA-TEG-OH). Upon seeing activity with the DNA-TEG-OH, we redirected the selection with we redirected the selection effort using DNA-C_3-CSA as a substrate (**Fig. 2.26**).

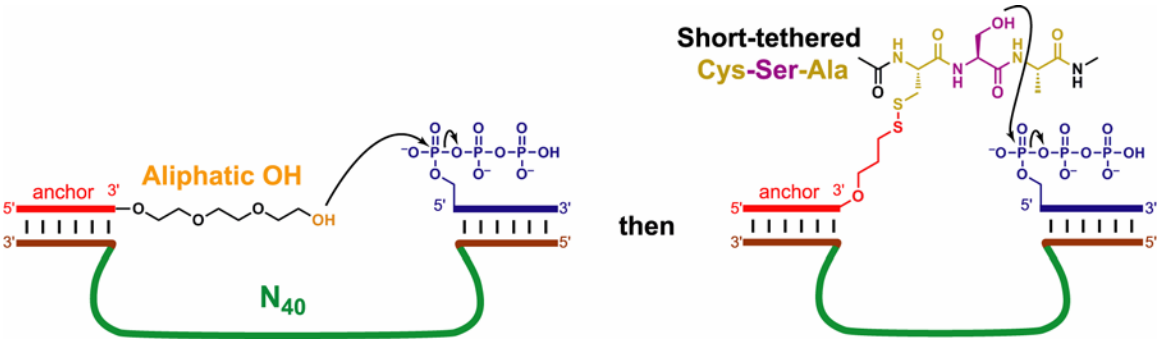


Figure 2.26 Alternate approach to obtain DNA-catalyzed serine modification. The selection effort was initiated first with a DNA substrate conjugated with an aliphatic hydroxyl group (DNA-TEG-OH) and redirected with the DNA-C_3-CSA substrate.

2.2.14 Cloning and sequencing of MZ deoxyribozymes and their kinetics

We observed DNA-catalyzed ligation between the DNA-TEG-OH with the RNA in round 11 with 8% yield. We redirected the selection effort using DNA-C_3-CSA from the round 11 pool. The pool's ligation activity was 6% at round 15, where individual deoxyribozymes were identified by cloning (**Fig. 2.27**).

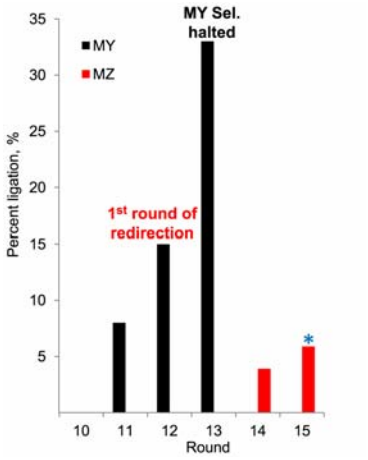


Figure 2.27 MY and MZ selection progression. The MY selection was initiated with the N_{40} pool using the DNA-TEG-OH substrate. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. The selection effort was redirected with DNA-C_3-CSA starting in round 12, denoted as the MZ selection. The pool of MZ deoxyribozymes was cloned and sequenced at round 15. The blue asterisk denotes the round and activity (6%) where the MZ selection was cloned.

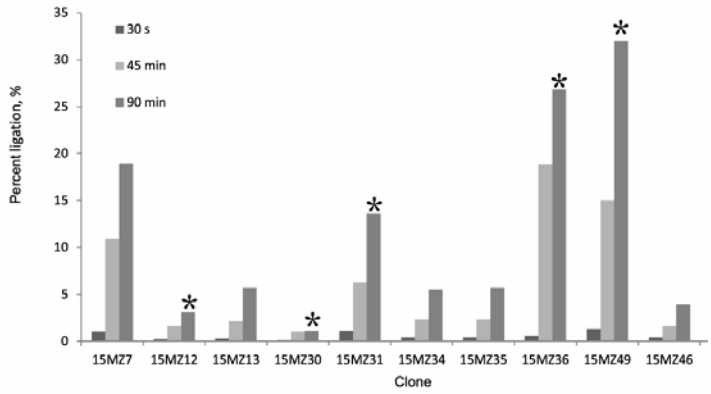


Figure 2.28 Representative kinetic screening of individual 15MZ clones. Only fraction of the clones assayed was presented here. The grey color scheme was used because clones contain double-inserts were also assayed. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. Time points: 30 s, 45 min, 90 min. Sequences were obtained and highlighted by the asterisks.

Deoxyribozymes in round 15 of the 15MZ selection were cloned. Of the 56 clones surveyed, the majority of the clones (36 clones) contain two of the N_{40} sequences (double insert), as a result of two sequences that were ligated to the plasmid during cloning. Only 5 out of the 20 of the clones with the corrected insert were active. The number of the

correct insert clones was low therefore we also surveyed the clones with double insert (**Fig. 2.28**).

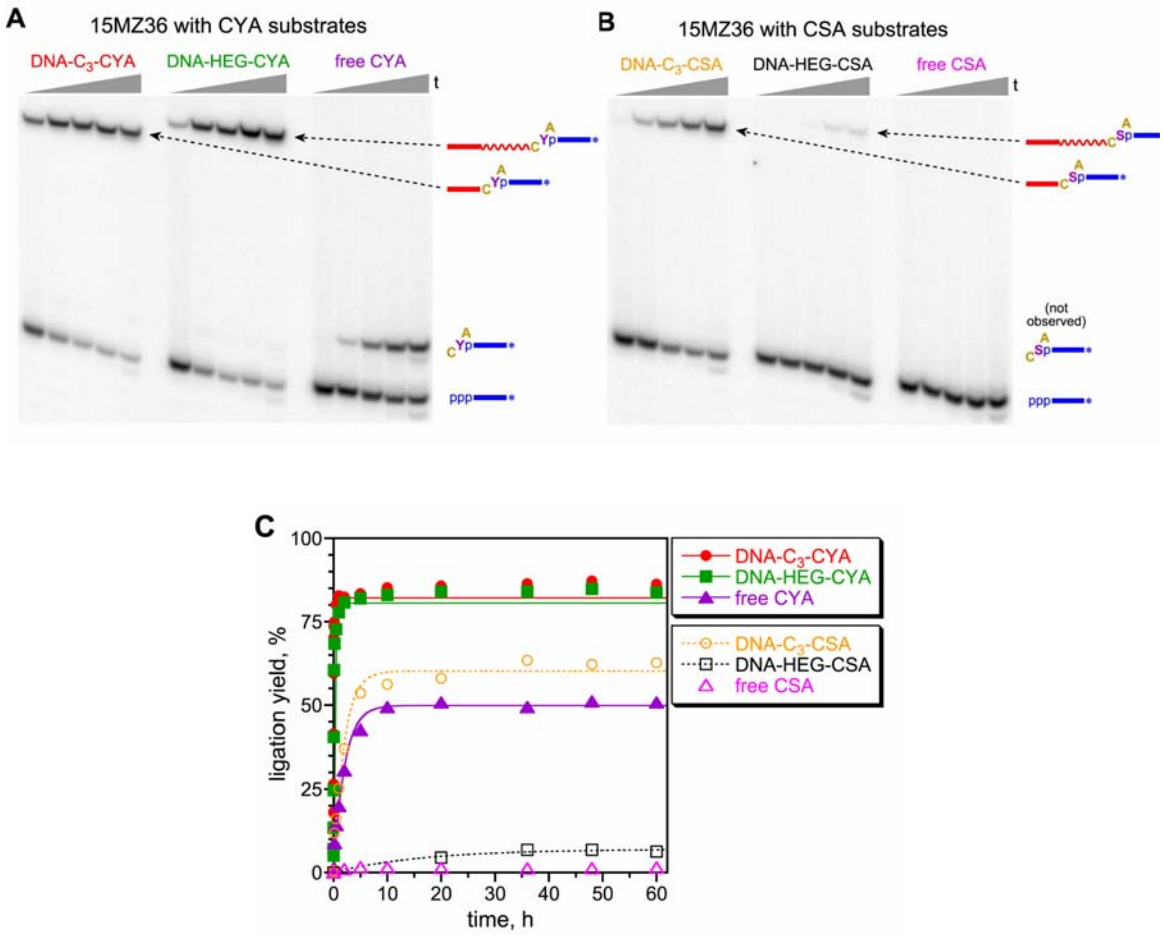


Figure 2.29 Kinetic assays of the 15MZ36 deoxyribozyme. (A) PAGE image for assays with the three CYA substrates. Time points were taken at t = 30 s, 10 min, 2, 10 and 60 h. (B) PAGE image for reaction of the three analogous CSA substrates. (C) Kinetic plots for data from panels A and B.

The best deoxyribozyme that identified from this effort 15MZ36, had $k_{obs} = 0.52\ h^{-1}$ and 62% yield after 60 h incubation with the DNA-C_3-CSA substrate. However, 15MZ36 had only around 6% yield after 60 h incubation with the DNA-HEG-CSA substrate and no detectable activity with free CSA. The three analogous CYA substrates reactivity were assayed with 15MZ36 (**Fig. 2.29**). 15MZ36 displayed robust activities with the tethered

variants. For DNA-C_3-CSA, 15MZ36 had k_{obs} = 80 h^{-1} and 82% yield after 60 h incubation (~270-fold faster than 10KC3, and ~26-fold faster than 11MN5). For DNA-HEG-CYA, 15MZ36 had k_{obs} = 18.7 h^{-1} and 81% yield after 60 h incubation (~ 60-fold faster than 10KC3, and ~22-fold faster than 11MN5). For free CYA, 15MZ36 had k_{obs} = 0.5 h^{-1} and 50% yield after 60 h incubation (~290-fold faster than 10KC3, and 10-fold faster than 11MN5). As negative controls, no reactions with DNA-C_3-CAA and free CAA were observed.

Two other deoxyribozymes, 15MZ30 and 15MZ49 were also assayed to assess the improvement in reactivity compared to 15MZ36 (**Fig. 2.30**). For simplicity, three illustrative substrates CYA, DNA-C_3-CSA and DNA HEG-CSA were used for the assays. 15MZ30 and 15MZ49 were not faster than 15MZ36.

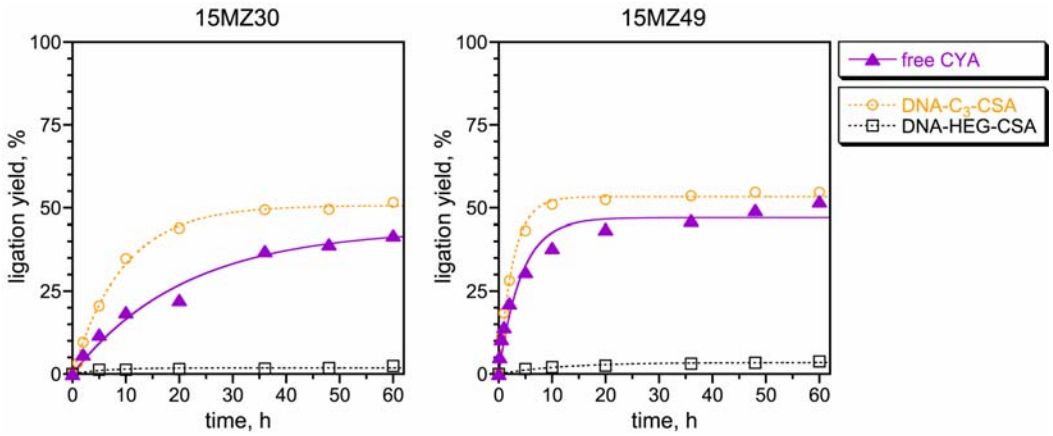


Figure 2.30 Kinetic plots of the 15MZ30 and 15MZ49 deoxyribozyme.

2.2.15 Reselection using DNA-C_3-CYA

(QG selection, partially 25% randomized 15MZ36)

15MZ36 had robust activity toward all the Tyr variants and the tethered Ser variants but not with the free CSA substrate. To select for free CSA reactivity, we designed a new selection, aiming to explore sequences closely related to 15MZ36. We initiated a selection effort using the partially (25%) randomized 15MZ36 with DNA-C_3-CYA, denoted as the QG selection. We anticipated that the QG selection would allow us to obtain deoxyribozymes with improved reactivity with free CYA and CSA.

In the QG selection effort, we also applied a time pressure and redirected the selection using the DNA-HEG-CYA substrate. However, the redirection with DNA-HEG-CYA was not fruitful as only 1% activity was detected. The deoxyribozymes in the QG selection with DNA-C3-CYA substrate in round 6 with 10 min incubation (13%) were cloned (**Fig. 2.31**).

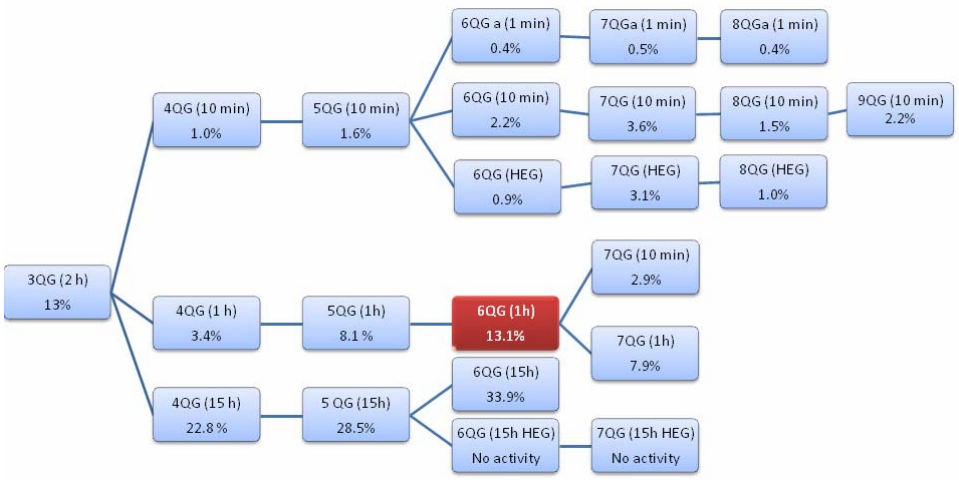


Figure 2.31 QG selection progression. The QG selection was initiated with 15MZ36 partially (25%) randomized pool using the DNA-C_3-OH substrate. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. To select for faster deoxyribozymes, 10 min and 1 min incubation time pressures were applied. To select for deoxyribozyme that can function better with free CYA, the selection effort was redirected with DNA-HEG-CYA starting in round 12, denoted as 6QG (HEG). The pool of QG deoxyribozymes was cloned and sequenced at round 6. The red box highlights the round and activity (13%) where the QG selection was cloned.

2.2.16 Cloning and sequencing of QG deoxyribozymes and their kinetics

6QG clones were assayed with the parent DNA-C_3-CYA substrate and all 14 clones were active (**Fig. 2.32**). Sequences were obtained for all of the clones (**Fig. 2.33**). The kinetic characteristics of 6QG deoxyribozymes were studied to compare with the 15MZ36 and 11MN5 deoxyribozyme. For the parent DNA-C_3-CSA substrate, 6QG6 and 6QG8 had k_{obs} = 0.40 h^{-1} and 50% yield and k_{obs} = 0.40 h^{-1} and 41% yield after 60 h incubation, respectively. However, none of the 6QG deoxyribozymes were active with DNA-HEG-CSA.

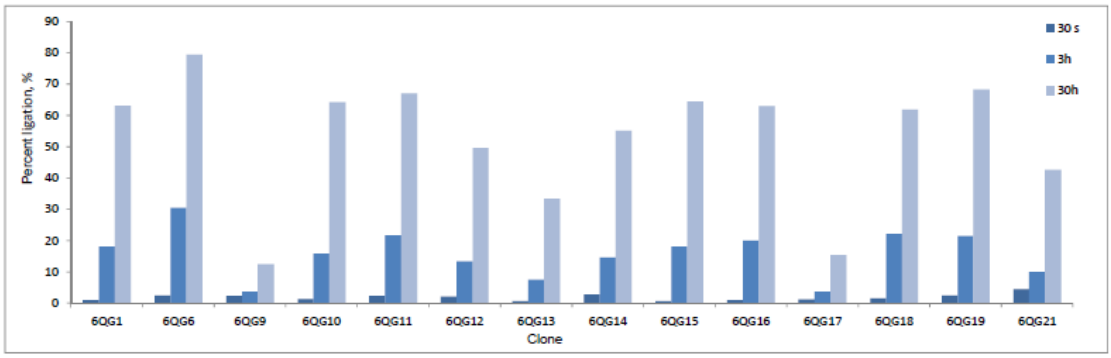


Figure 2.32 Kinetic screening of individuals 6QG clones. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. Time points were taken at 30 s, 3 h, 30 h.

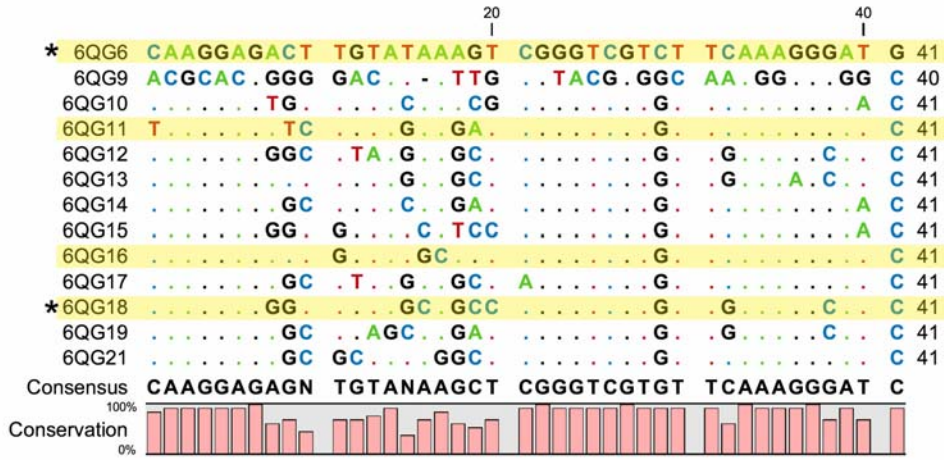


Figure 2.33 Sequence alignment of 6QG sequences. Highlighted sequences were prepared by solid-phase synthesis for kinetic characterization. Asterisks denote sequences that were studied in greater details.

For simplicity, two 6QG deoxyribozymes, 6QG6 and 6QG18, were assayed with the free CYA substrate (**Fig. 2.34**). For the free CYA, 6QG6 and 6QG18 had k_{obs} = 0.22 h^{-1} and 41% yield and k_{obs} = 0.30 h^{-1} and 45% yield after 60 h incubation. Compared to the previously obtained deoxyribozymes 11MN5, 6QG6 was 4-fold faster and 6QG18 was 6-fold faster. However, the 6QG deoxyribozymes showed no better reactivity with all the Tyr and Ser variants than the 15MZ36 deoxyribozyme.

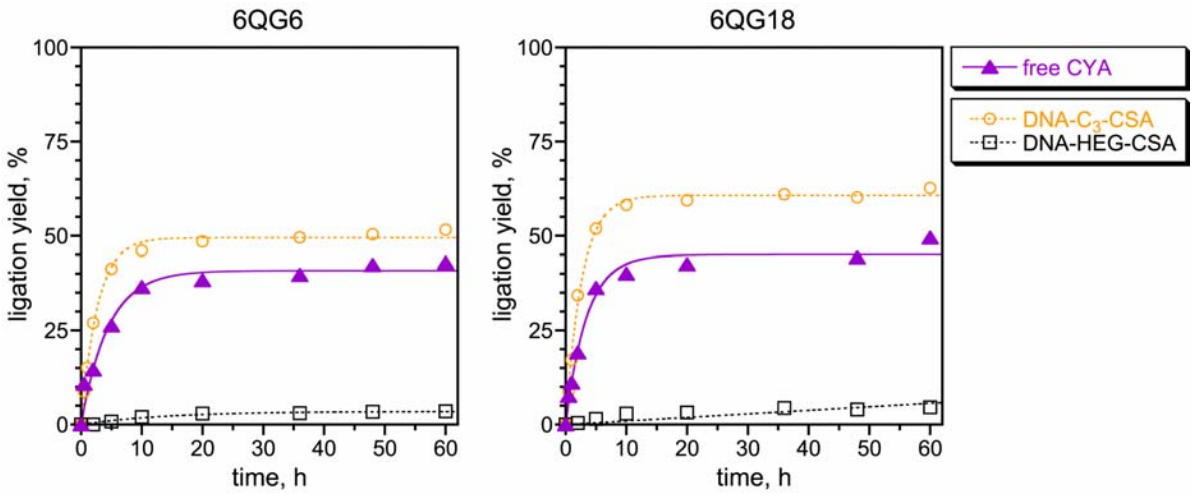


Figure 2.34 Kinetic plots of the 6QG6 and 6QG18 deoxyribozyme. They showed comparable activities as 15MZ36.

2.2.17 Redirection from the DNA-TEG-OH selection using DNA-C_3-CYA (NZ selection)

As part of the comprehensive effort to achieve CYA and CSA reactivity, we redirected the MY selection effort (DNA-TEG-OH selection) to use DNA-C_3-CYA. The redirection selection was called the NZ selection. The round 11 pool from the selection with DNA-TEG-OH (MY selection) was continued for two rounds with the DNA-C_3-CYA substrate with 2 h incubation. A time pressure was applied by decreasing the incubation time from 2 h to 10 min. Two additional rounds were carried out in 10 min

incubation. The activity reached to 30% (10 min incubation) in round 15 and the pool of deoxyribozyme was cloned and sequenced (**Fig. 2.35**).

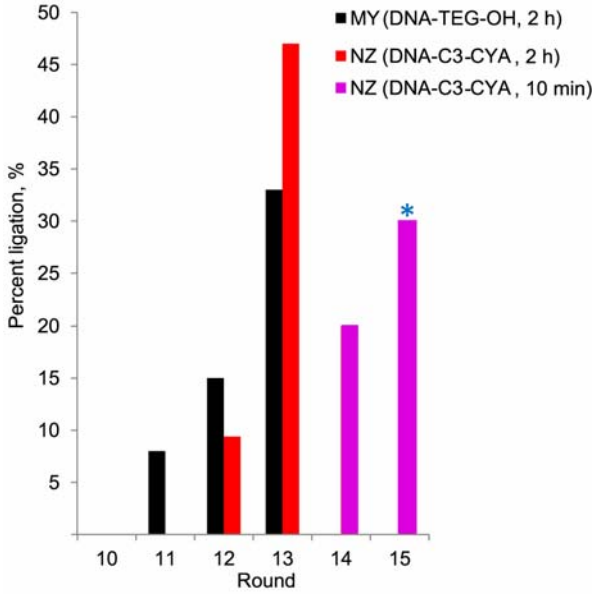


Figure 2.35 MY and NZ selection progression. The MY selection was initiated with N_{40} pool using the DNA-TEG-OH substrate. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. The selection effort was redirected with DNA-C_3-CYA using the round 11 of the MY selection. The redirection effort from DNA-TEG-OH to DNA-C_3-CYA was denoted as the NZ selection. After two rounds of redirection, the incubation time was decreased to 10 min. The pool of NZ deoxyribozymes was cloned and sequenced at round 15. The blue asterisk denotes the round and activity (30%) of where the NZ selection was cloned.

2.2.18 Cloning and sequencing of NZ deoxyribozymes and their kinetics

Of the 23 clones surveyed, 22 clones were active (>5 % activity; **Fig. 2.36**). Sequences were obtained for 10 of the active clones (**Fig. 2.37**). We examined the kinetic activity of 15NZ11 and 15NZ16 with DNA-C_3, HEG-CYA substrates as well as free CYA. For the DNA-C_3-CSA substrate, 15NZ11 and 15NZ16 had $k_{obs} = 0.34\ h^{-1}$ and 57% yield and $k_{obs} = 0.10\ h^{-1}$ and 47% yield after 60 h incubation, respectively. However, none of the 15NZ deoxyribozymes were active with DNA-HEG-CSA. For the free CYA,

15NZ11 and 15NZ16 had $k_{obs} = 0.24\ h^{-1}$ and 44% yield and $k_{obs} = 0.01\ h^{-1}$ and 50% yield after 60 h incubation, respectively (**Fig. 2.38**).

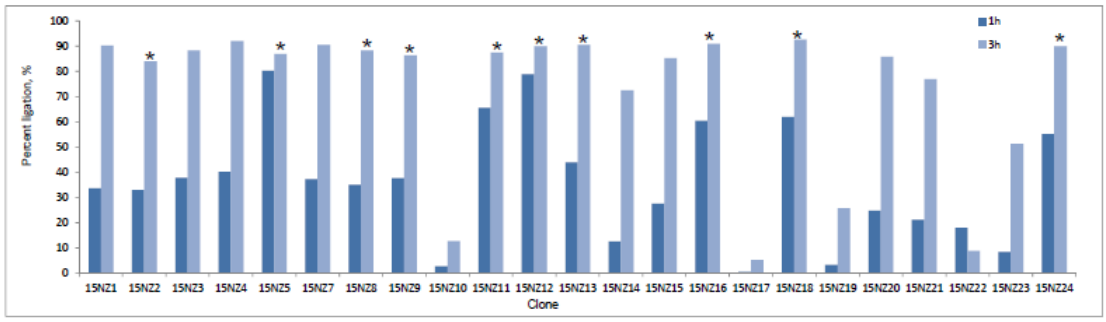


Figure 2.36 Kinetic screening of individuals 15NZ clones. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC. Time points were taken at 1 h, 30 h. The sequences were obtained with the clones marked with an asterisk.

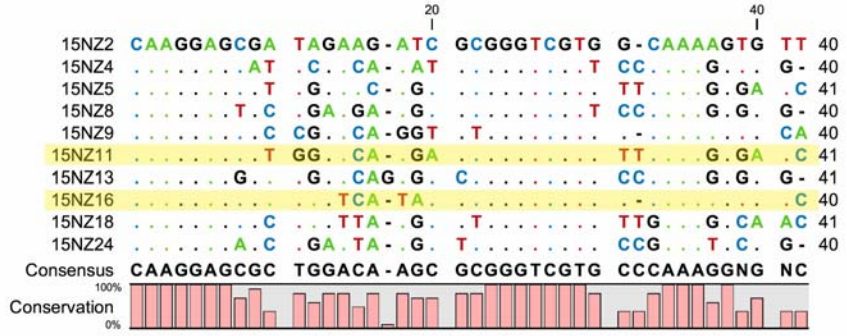


Figure 2.37 Sequence alignment of 15NZ sequences. Highlighted sequences were prepared by solid-phase synthesis for kinetic characterization.

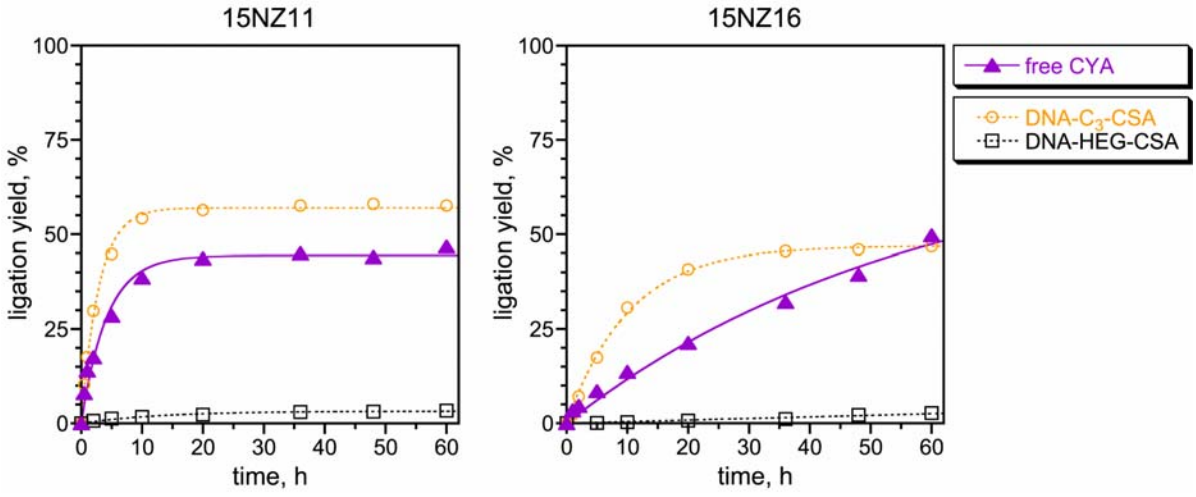


Figure 2.38 Kinetic plots of the 15NZ11 and 15NZ16 deoxyribozyme.

2.2.19 Dependence of 15MZ36 free CYA reactivity on the 3′-terminal composition of the deoxyribozymes

15MZ36 was identified from an in vitro effort that was employed entirely using a tethered tripeptide substrate. The 3′-terminal of the deoxyribozyme was always Watson-Crick basepaired to the substrate either with the C_3 or HEG-tethered tripeptide. Therefore, we evaluated the 15MZ36 activity with the free CYA substrates with changes to the composition of the 3′-terminal of the deoxyribozyme. First, we demonstrated that truncating the 3′-terminal of the deoxyribozymes lead to substantial decrease of the reactivity. Second, we showed that changing the helper oligonucleotide to either DNA terminated TEG, or HEG tether, rather than having a free 3′-OH terminus did not affect the activity of 15MZ36 toward free CYA substrate (**Fig. 2.39**). Two conclusions can be made from these studies. First, the 15MZ36 deoxyribozyme required the 3′-terminal to be double-stranded DNA. Second, the tether atoms did not contribute to the reactivity of 15MZ36. In related assays, similar observations were made for 9NG14 and 11MN5.

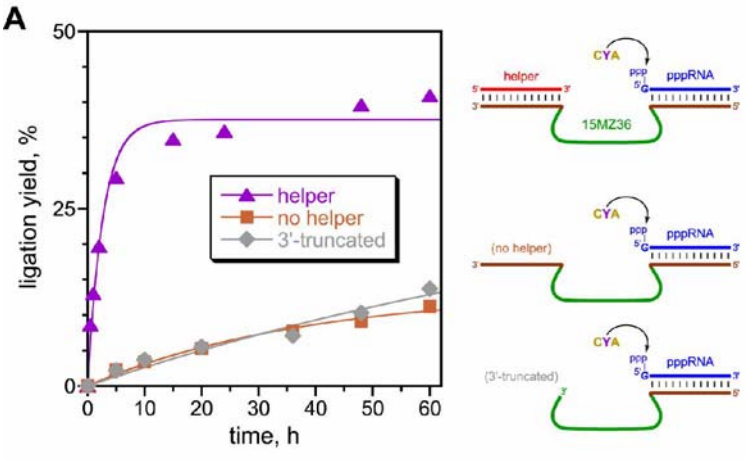


Figure 2.39 Dependence of 15MZ36 activity on the 3′-terminal composition of the 15MZ36.

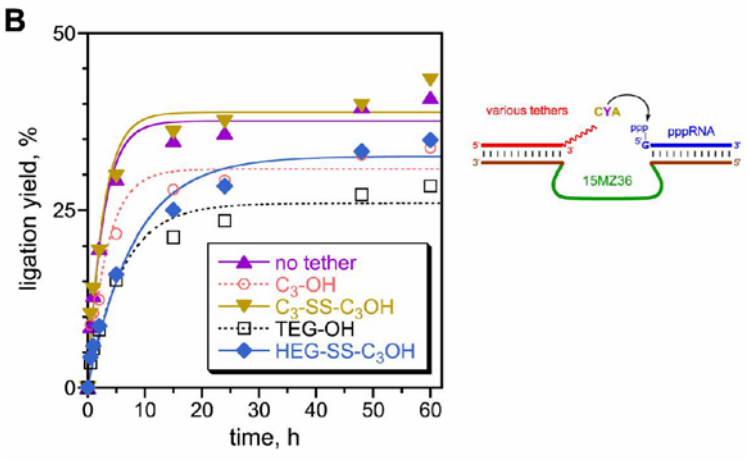


Figure 2.39 (cont.) Dependence of 15MZ36 activity on the 3′-terminal composition of the 15MZ36. 15MZ36 was used as an illustrative example. (A) Kinetic assays were performed in two formats: without the helper oligonucleotide or with truncating the 3′-terminal of 15MZ36. Both the omission of the helper oligonucleotide and the truncation of the 3′-terminal of 15MZ36 led to substantial decrease of activity. (B) Various helper oligonucleotides terminating with different tethers were used to assay the activity of 15MZ36 with free CYA. The comparable activity of 15MZ36 with different helper oligonucleotides terminating with different tether illustrates that the tether atoms do not contribute to the activity of 15MZ36.

2.2.20 Dependence of 15MZ36 activity on divalent metal cofactors

We examined the divalent metal cofactor requirements for 10KC3, 11MN5 and 15MZ36-catalyzed reactivity with the DNA-C_3-CYA substrate. Metal cofactors 40 mM $MgCl_2$, and 20 mM $MnCl_2$ were employed during the selection. 10KC3, 11MN5 and 15MZ36 depend on Mn^{2+} ions only. The concentration of $MnCl_2$ requirement was assayed. From these assays, the $K_{d,app}$ with $MnCl_2$ of 10KC3 was estimated to be ~5 mM (**Fig. 2.40**). 11MN5 had a $K_{d,app}$ with $MnCl_2$ ~2 mM. 15MZ36 had a $K_{d,app}$ with $MnCl_2$ < 1 mM (**Fig. 2.41**).

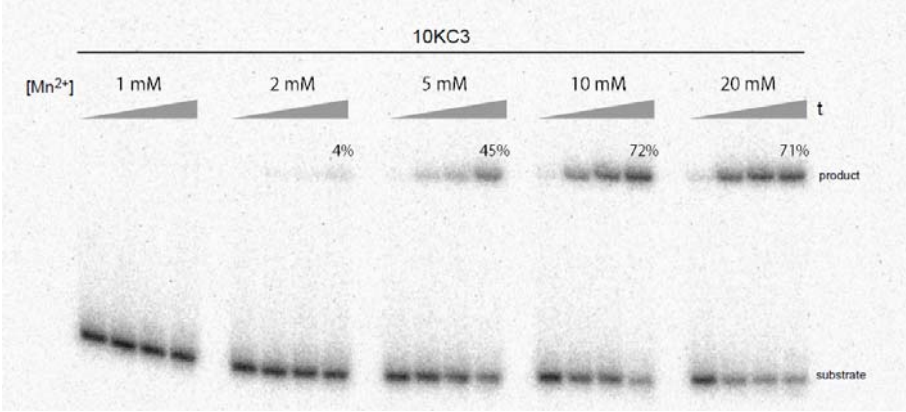


Figure 2.40 Kinetic assay of 10KC3-catalyzed tethered CYA reactivity with various concentrations of Mn^{2+}. Metal cofactors 40 mM $MgCl_2$, and 20 mM $MnCl_2$ were employed during the selection. 10KC3-catalyzed Tyr reactivity only depends on Mn^{2+} ions. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and various concentrations of $MnCl_2$ at 37 ºC. Time points were taken at 5 min, 3 h, 25 h. From these assays, the $K_{d,app}$ with $MnCl_2$ of 10KC3 was estimated to be less than 5 mM. Experiments were performed by Sarah Kwon (SCK).

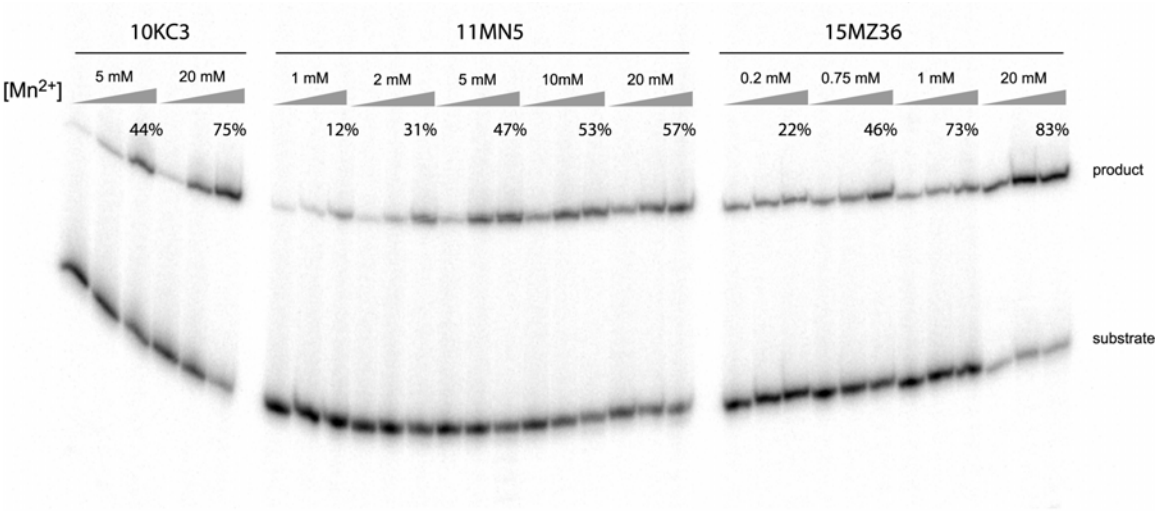


Figure 2.41 Kinetic assay of 10KC3, 11MN5, 15MZ36-catalyzed tethered CYA reactivity with various concentrations of Mn^{2+}. Metal cofactors 40 mM $MgCl_2$, and 20 mM $MnCl_2$ were employed during the selection. 10KC3, 11MN5 and 15MZ36 only depend on Mn^{2+} ions. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and various concentrations of $MnCl_2$ at 37 ºC. Time points were taken at 5 min, 3 h, 21 h. From these assays, the $K_{d,app}$ with $MnCl_2$ of 10KC3 was estimated to be ~5 mM. 11MN5 had a $K_{d,app}$ with $MnCl_2$ ~2 mM. 15MZ36 had a $K_{d,app}$ with $MnCl_2$ < 1 mM. Experiments were performed by SCK.

2.2.21 Effect of lanthanides on 10KC3 catalysis

Triavalent lanthanide ions have been used as alternative catalytic cofactors for a RNA-cleaving deoxyribozyme.[16] To assess the effect of lanthanides on 10KC3-catalyzed short-tethered CYA ligation to the RNA substrate, kinetic assays were performed with lanthanide ions, Eu^{3+} or Yb^{3+}, along with addition of 5 mM Mn^{2+}. It was found that both lanthanide ions inhibit the reactivity of 10KC3 (**Fig. 2.42**).

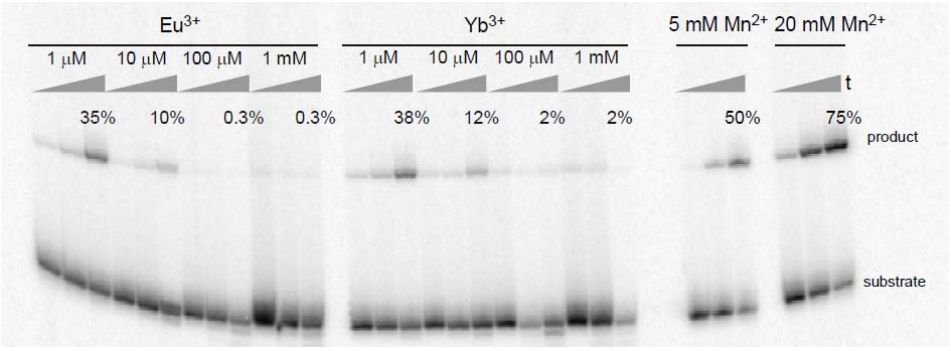


Figure 2.42 The effect of lanthanide ions (Eu^{3+} or Yb^{3+}) on 10KC3. Reaction conditions: 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 5 m $MnCl_2$ at 37 ºC. The 4-fold decrease of Mn^{2+} (from 20 mM to 5 mM) allows us to readily discern the effect of the lanthanide ions on 10KC3 catalysis. Time points were taken at 5 min, 3 h, 24 h. Lanthanide ions inhibit 10KC3 catalysis, leading to a lanthanide concentration-dependent decrease in product formation. Experiments were performed by SCK.

2.2.22 Determination of $K_{d,app}$ for 15MZ36 with free CYA tripeptide

We examined the dependence of 15MZ36 activity on the concentration of the free CYA substrate. The $K_{d,app}$ value was found to be on the order of 1 mM, with 65% yield at 60 h with 3 mM CYA (**Fig. 2.43**). The result of the $K_{d,app}$ value in the order of mM was not unexpected as the various versions of tethered CYA substrates were used in all the selection efforts.

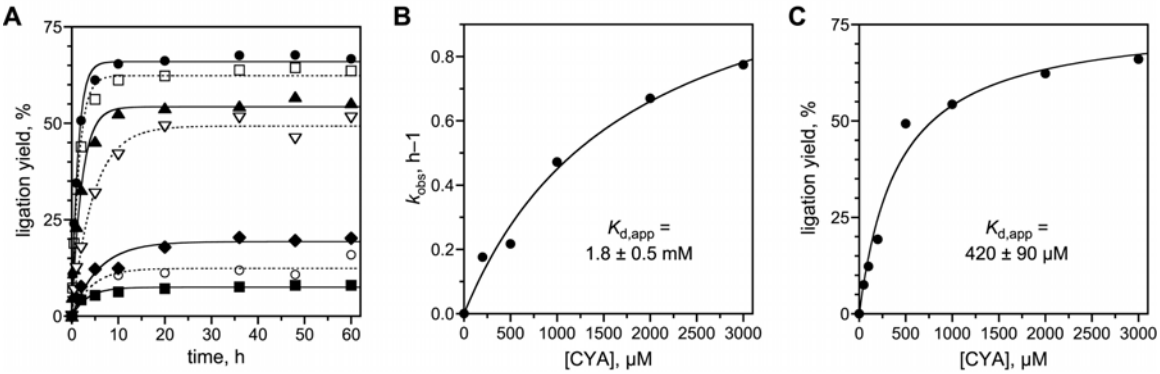


Figure 2.43 Determination of $K_{d,app}$ for 15MZ36 with the free CYA tripeptide substrate. (A) 15MZ36 ligation yield was evaluated with CYA concentrations of 50, 100, 200, and 500 μM as well as 1, 2, and 3 mM. Data were fit to the standard equation $Y = Y_{max} \cdot C/(K_d + C)$, where C is the tripeptide concentration. (B) From the k_{obs} data, the $K_{d,app}$ value for the free CYA tripeptide substrate was 1.8 ± 0.5 mM. (C) From the yield data, the $K_{d,app}$ value was 420 ± 90 μM.

2.2.23 Multiple turnover assay of 15MZ36

We studied 15MZ36 activity under multiple turnover conditions. To do this, we assayed the DNA-catalyzed reaction with 10% (0.1:1; deoxyribozyme:RNA) of the deoxyribozymes relative to the 3′-^{32}P-radiolabeled 5′-triphosphorylated RNA substrate, along with excesses of the helper oligonucleotide and free CYA substrate. When the assay was performed with 100% deoxyribozyme (1:1; deoxyribozyme: RNA) with the RNA substrate the ligation yield was ≤8%, providing no evidence for turnover (**Fig. 2.44A**). It was reasoned that reducing the number of base pairs between the deoxyribozyme and the RNA substrate would encourage the product dissociation. Therefore, a set of deoxyribozyme was designed with a 5′-binding arm forming between 7–11 base pairs with 5′-triphosphorylated RNA substrate to modulate potential product inhibition (which was frequently observed with many deoxyribozymes). However, again no evidence for multiple turnover was observed (**Fig. 2.44B**).

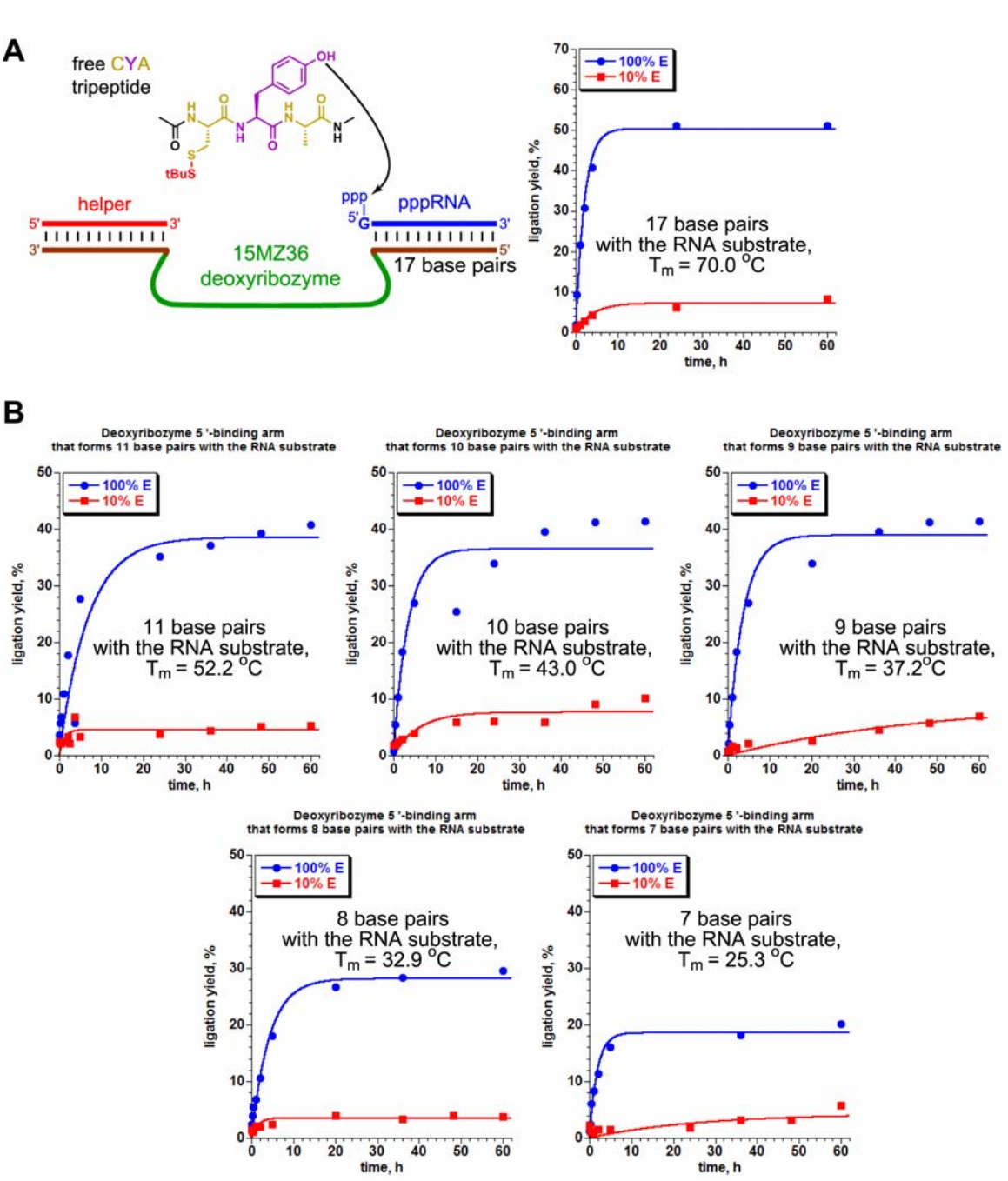


Figure 2.44 Multiple turnover assays with 15MZ36. (A) 15MZ36 forms 17 base pairs with the RNA substrate. Kinetic plots of 100% or 10% of the deoxyribozyme was used relative to the 3′-^{32}P-radiolabeled 5′-triphosphorylated RNA substrate. No evidence for multiple turnover was observed. (B) Kinetic plot of deoxyribozymes with reduced number of base pairs with the RNA substrate to foster product dissociation.

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

To characterize the product, MALDI mass spectrometry was obtained for the representative ligations products, as depicted in **Fig. 2.45** with representative data for the 15MZ36 DNA-C_3-CSA–RNA product. To further support with the product assignment, DTT treatment and RNase T1 digestions were separately performed on the ligation product. DTT treatment led the reduction of the disulfide bond, and RNase T1 cleaves specifically every 3′ end of single-stranded G residues. Both the DTT treatment and RNase T1 digestion led to the expected masses. Analogous mass analyses were obtained for several other ligation products, including the CYA–RNA product formed by 15MZ36.**Table 2.1** shows the results of the masses of several other ligation products.

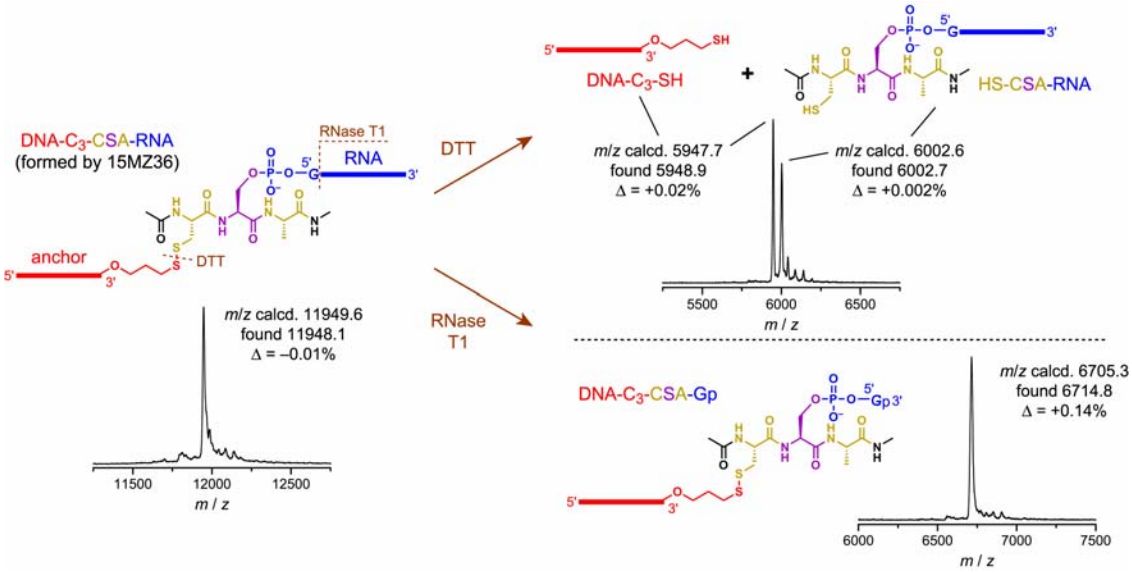


Figure 2.45 MALDI mass spectrometry analyses of the 15MZ36 DNA-C_3-CSA–RNA ligation product and its DTT and RNase T1 digestions.

deoxyribozyme and substrate	mass calcd. [a]	mass found	error, % (found – calcd.)
10KC3 with **DNA-C_3-CYA**			
ligation product	12026.2	12025.6	–0.01
DTT digestion (L product) [b]	5947.7	5950.2	+0.04
DTT digestion (R product) [b]	6078.9	6080.2	+0.02
RNase T1 digestion	6782.9	6781.2	+0.03
11MN5 with **DNA-HEG-CYA**			
ligation product	12371.0	12369.2	–0.02
DTT digestion (L product) [b]	6293.0	6295.2	+0.04
DTT digestion (R product) [b]	6079.0	6079.3	+0.01
RNase T1 digestion	7126.7	7126.8	0.00
15MZ36 with **DNA-C_3-CSA**			
ligation product	11949.6	11948.1	–0.01
DTT digestion (L product) [b]	5947.7	5948.9	+0.02
DTT digestion (R product) [b]	6002.6	6002.7	0.00
RNase T1 digestion	6705.3	6714.8	+0.14
15MZ36 with **untethered CYA**			
ligation product	6167.2	6165.9	–0.02
DTT digestion [c]	6078.0	6078.7	+0.01
RNase T1 digestion	923.8 [d]	924.5 [d]	+0.08

Table 2.1. MALDI mass spectrometry analyses of key deoxyribozyme products and their DTT and RNase T1 digestions. All MALDI mass spectra were obtained in the mass spectrometry laboratory of the UIUC School of Chemical Sciences. See the Experimental Procedures for reaction details.

[a] Calculated masses are for $[M–H]^-$

[b] "L product" refers to the left-hand product (DNA anchor oligonucleotide + tethered thiol), and "R product" refers to the right-hand product (tripeptide + RNA), as shown in Fig. 2.44.

[c] When the substrate is a free tripeptide, the DTT digestion leads only to the tripeptide-RNA product.

[d] Mass spectrum obtained in positive ion mode, $[M+H]^+$.

2.2.25 Sequence alignments and relationships among the new deoxyribozymes

The sequence similarities between of each of the sequences were compared (**Fig. 2.46**). Each of the more active deoxyribozymes 9NG14, 11MN5, and 15MZ36 differ from the original 10KC3 by numerous mutations (10 for 9NG14; 9 as well as three

inserted nucleotides for 11MN5; and 8 as well as one inserted nucleotide for 15MZ36). Upon aligning the representative sequences of the selection efforts, three entirely or largely conserved nucleotide blocks could be observed. The many shared nucleotides suggest the emergence in this study of one overall motif for DNA-catalyzed reactivity of Tyr/Ser side chains in peptide substrates.

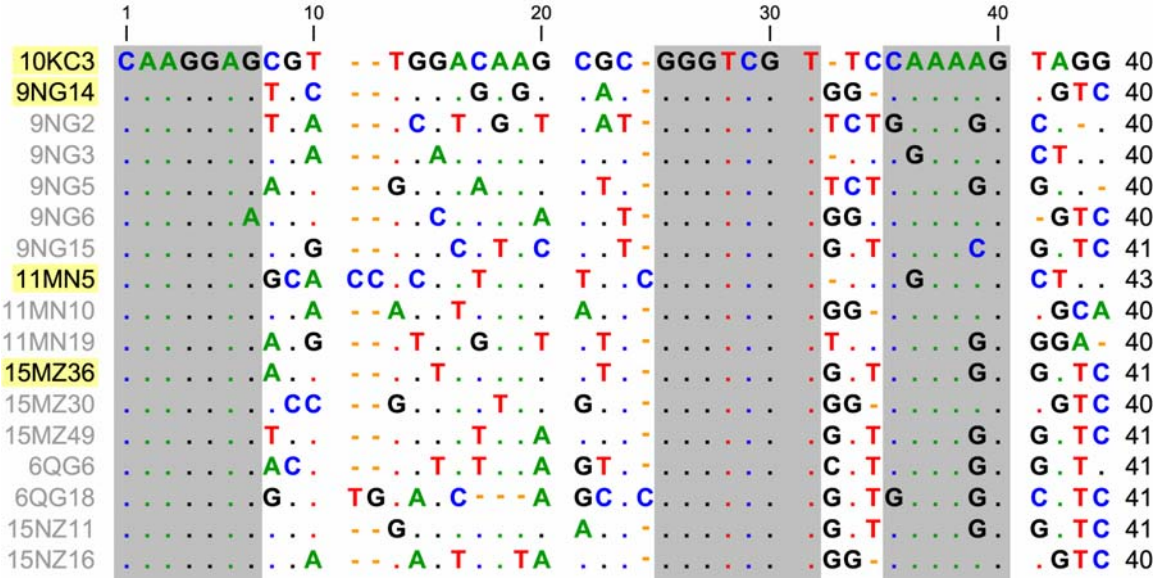


Figure 2.46 Sequence alignment of the representative deoxyribozyme from the Ser/Try selection efforts. Three nucleotides blocks (in grey) were observed, suggesting a common motif for DNA-catalyzed reactivity of Tyr/Ser side chains in peptide substrates.

2.2.26 Nucleotide mutation analysis of 15MZ36 and 10KC3

15MZ36 differs in 8 as well as one more inserted nucleotide compared to the 10KC3, yet the reactivity is drastically different between 15MZ36 and 10KC3. We examined the role of the 8 and one inserted nucleotide of 15MZ36. To achieve this, we studied the effect of loss-of-function 15MZ36 mutants, i.e., mutating the nucleotides of 15MZ36 to the nucleotides of 10KC3 in a specific position) as well as the effect of gain-of-function 10KC3 mutants, i.e., mutating the nucleotides of 10KC3 to the nucleotides of 15MZ36 in

a specific position. From these studies (performed by Yoonhee Anh, an undergraduate student), the following was concluded.

First, the identity of position 29 is important for 15MZ36 as well as 10KC3. The kinetics studies of 15MZ36: G29Δ, 10KC3: Δ29G, and 15MZ36: G29A with C_3 and HEG-CYA substrates revealed that the nucleotide in position 29 is essential for the deoxyribozyme. Both the identity and presence of the nucleotide at this position play a significant role. The importance of G in position 29 was shown through the fact that 15MZ36 completely lost its activity when guanine was deleted and when guanine was replaced with adenine. This suggested that adenine, although it is similar in size to guanine, was not adequate to retain the activity of the deoxyribozyme. Also, the importance of the guanine was verified with the completely loss of activity with the 15MZ36: G29Δ, in which the guanine was absent. In a similar manner, position 29 of 10KC3 seems to be critical because it lost its activity when guanine was substituted at position 29 (in the original 10KC3 sequence, G is absent). Thus, it was concluded that 15MZ36 requires guanine at position 29 and 10KC3 requires that the position remained unoccupied.

Second, the identity of position 41 is critical for the function of 15MZ36. We examined the kinetics of 15MZ36:C41G and 10KC3:G41C. It was found that 15MZ36:C41G completely lost its activity when C was replaced with G, indicating that the transversion was not tolerated. Whereas, 10KC3:G41C mutant is tolerated with a slight decrease in activity compared to the original 10KC3 sequence.

2.3 Summary

In this chapter, deoxyribozymes were found to catalyze Tyr-RNA or Ser-RNA nucleopeptide linkage formation when the peptide was either tethered to a DNA anchor oligonucleotide or, in the case of Tyr, untethered (free). Several new deoxyribozymes for Tyr modification (and several for Ser modification as well) were identified via in vitro selection. The best new deoxyribozyme from these efforts, 15MZ36, catalyzes covalent Tyr modification of a free Cys-Tyr-Ala substrate with k_{obs} = 0.5 h^{-1} ($t_{1/2}$ 83 min) with up to 65% yield.

Fig. 2.47 illustrates the key selection efforts that were performed in this comprehensive study. The KC selection was initiated with the DNA-C_3-CYA using a N_{40} pool, which lead to the discovery of the 10KC3 deoxyribozymes. To seek deoxyribozymes that function better with the free CYA substrate, different selection efforts were performed. A reselection using the DNA-C_3-CYA substrate with the 25% partially randomized 10KC3 pool led to the discovery of 9NG14 with improved activity with the CYA substrate 9NG14: 40% yield in 60 h (k_{obs} = 0.063 h^{-1}, $t_{1/2}$ of 11h), ~ 30 fold faster and with a 4-fold improvement in yield compared to 10KC3. MN selection was performed with DNA-HEG-CYA from using a N_{40} pool; and 11MN5 showed significant improvement toward the free CYA with 55% yield in 60 h (k_{obs} = 0.05 h^{-1}, $t_{1/2}$ of 14 h, ~ 30-fold faster and 5-fold improvement in yield compared to 10KC3). To seek DNA-catalyzed Ser reactivity, we initiated the MZ selection effort with the DNA-TEG-OH using a N_{40} pool. Upon seeing activity with the DNA-TEG-OH, we redirected the selection effort using DNA-C_3-CSA as a substrate. From the MZ selection effort, we identified 15MZ36 which was the best among all the selection efforts described in this

chapter. To explore sequences closely related to 15MZ36. We initiated a selection effort using the partially (25%) randomized 15MZ36 with DNA-C_3-CYA, denoted as the QG selection. Compared to the previously obtained deoxyribozymes 11MN5, 6QG6 was 4-fold faster and 6QG18 was 6-fold faster in catalyzing free CYA reactivity. Furthermore, the DNA-TEG-OH selection was redirected to use DNA-C_3-CYA, which was the NZ selection. 15NZ11 showed comparable activity (~2-fold slower with CYA, similar in yield) compared to that of 15MZ36.

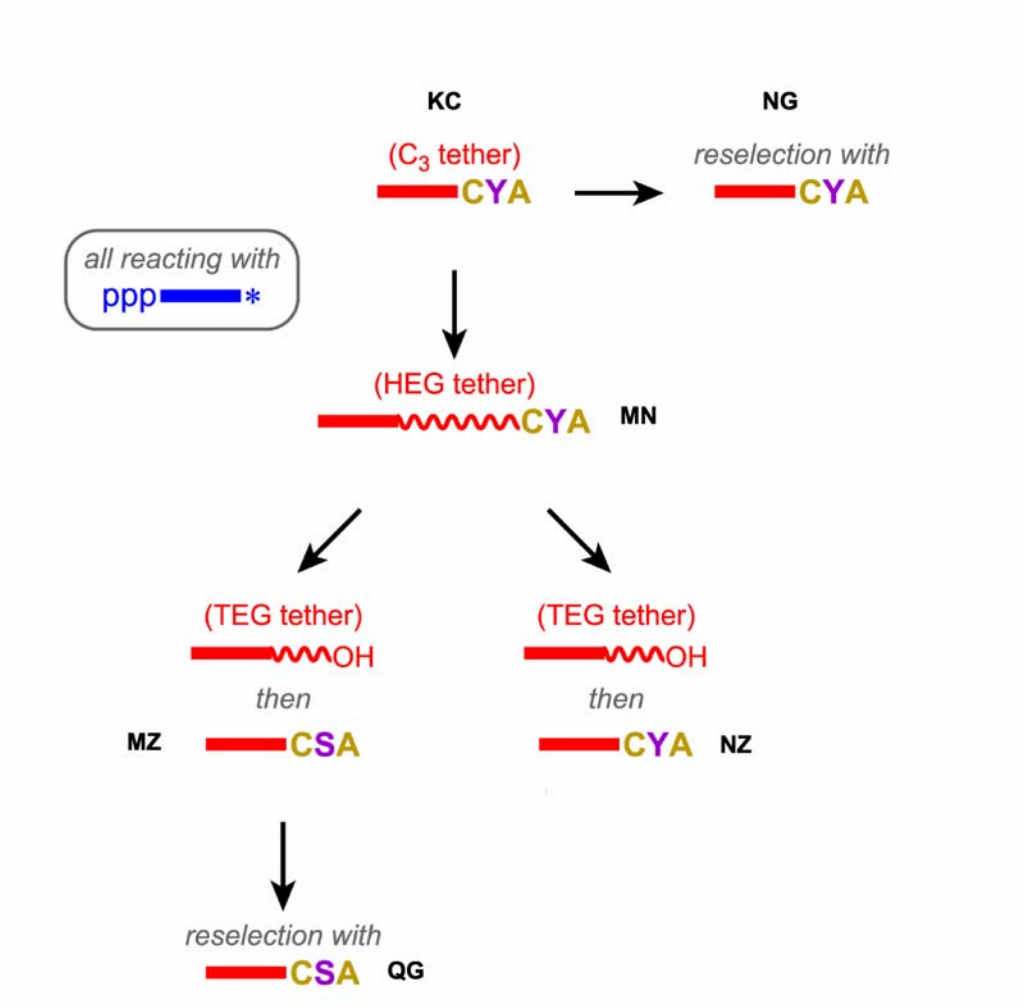


Fig. 2.47 Comprehensive study using various tethered substrates during the selection.

From all the selection efforts reported in this chapter, several implications could be drawn to direct the design of future selection efforts. First, 15MZ36 was identified from the selection effort starting with the DNA-TEG-OH substrate and redirected using the DNA-C_3-CSA as its substrate. For future experiments, redirection maybe a good strategy to obtain activity. Second, including a DNA substrate terminating with a hexa(ethylene) glycol and covalently linked to a peptide (DNA-HEG-CYA) in the selection led to the identification of deoxyribozymes with improved kinetics in the context of DNA-catalyzed nucleopeptide ligation. In the future, a substrate with a polyethylene glycol linker conjugated with a peptide could be employed during selection. Third, selection efforts in which the peptide is present free in solution should be employed, which should potentially leads to the identification of deoxyribozymes with higher affinity to the peptide.

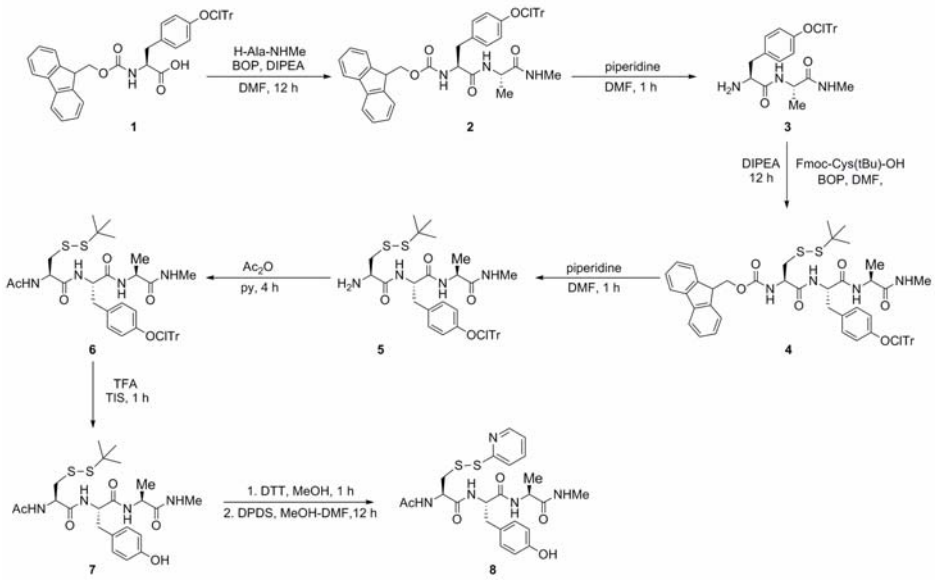
2.4 Materials and Methods

2.4.1 Preparation of oligonucleotide, DNA-Cys-Xaa-Ala conjugates, Cys-Xaa-Ala small molecules, and 25% partially randomized pool

Oligonucleotides. Standard DNA oligonucleotides were prepared at IDT (Coralville, IA). The 5′-triphosphorylated RNA substrate sequence, 5′-GGAAGAGAUGGCGACGG-3′, was prepared by in vitro transcription with T7 RNA polymerase and a chemically synthesized DNA template. The 3′-thiol modifier solid support was purchased from Glen Research. Denaturing PAGE with running buffer 1 × TBE (89 mM Tris, 89 mM boric acid and 2 mM EDTA, pH 8.3) was employed to purify all oligonucleotides including DNA-Cys-Xaa-Ala. N_{40} DNA with primer-binding regions: 5′-CGAAGTCGCCATCTCTTC-N_{40}-ATAGTGAGTCGTATTAAGCTGATCCTGATGG-3′ was used. Two primers were used in PCR. Primer 1 was 5′-$(AAC)_4$XCCATCAGGATCAGCTTAATACGACTCACTAT-3′ [X = hexa(ethylene glycol) to stop Taq polymerase] and Primer 2 was 5′-CGAAGTCGCCATCTCTTC-3′. Primer 2 is phosphorylated to allow ligation to the RNA substrate.

Preparation of DNA-Cys-Xaa-Ala conjugates. Four small molecules, PyS-S-Cys-Xaa-Ala, where Xaa denotes the amino acid, Tyr, Ser, Lys, or Ala were previously synthesized in our laboratory by a solution-phase method1. (Scheme 1.1). Fmoc-protected Xaa **1** were coupled to *N*-methylalanine using the benzotriazol-1-yl-oxytris(dimethylamino)phosphonium hexafluorophosphate (BOP) coupling reagent to generate the corresponding Fmoc-dipeptide **2**. The Fmoc was removed with piperidine and the dipeptide **3** was coupled to Fmoc-Cys(tBu)-OH. Fmoc-tripeptide **4** was treated with piperidine to yield tripeptide **5**. The N-terminus was acylated to yield **6**. Subsequent TFA treatment deprotected the side chain of the Xaa generated **7**. The *tert*-butyl disulfide was reduced, and the reaction with pyridyl disulfide yielded PyS-S-Cys-Xaa-Ala **8**. PyS-S-Cys-Xaa-Ala small molecules were coupled to the anchor DNA via a disulfide bond.

Thus, DNA with a 3'-thiol was required for conjugation with the PyS-S-Cys-Xaa-Ala small molecules. The Cys t-butyl disulfide was exchanged for a pyridyl disulfide by reduction with DTT and activation with 2,2′-dipyridyl disulfide. The final PyS-CXA tripeptides were analyzed by 1H NMR and 13C NMR spectroscopy as well as by ESI mass spectrometry.

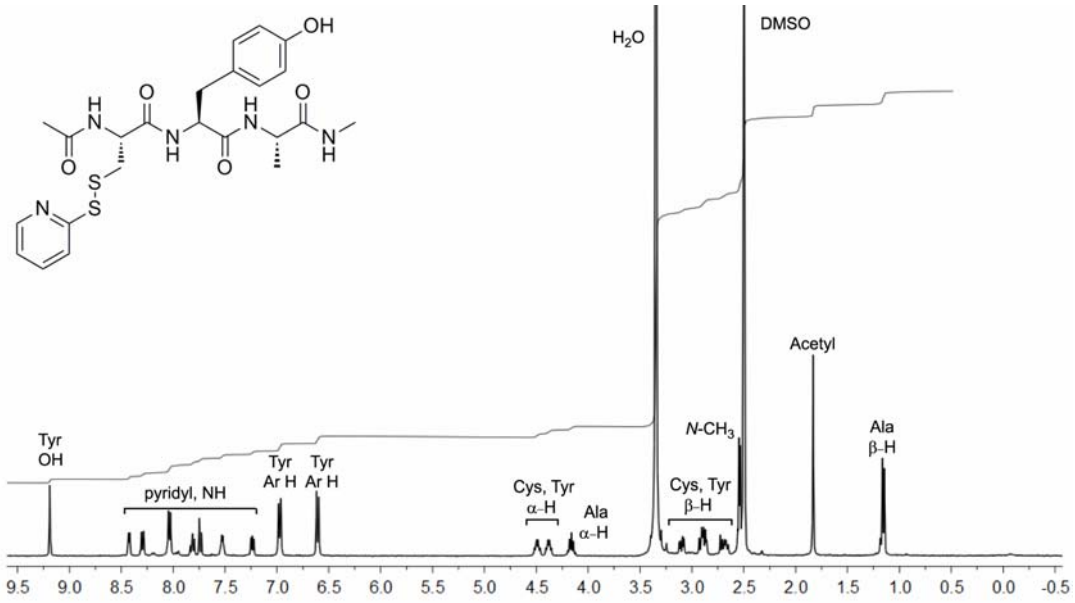


Scheme 2.1 Synthesis of PyS-S-Cys-Tyr-Ala. The 2-chlorotrityl group is abbreviated as ClTr. Tripeptides with Xaa = Ser, Lys or Ala were prepared using the same scheme.

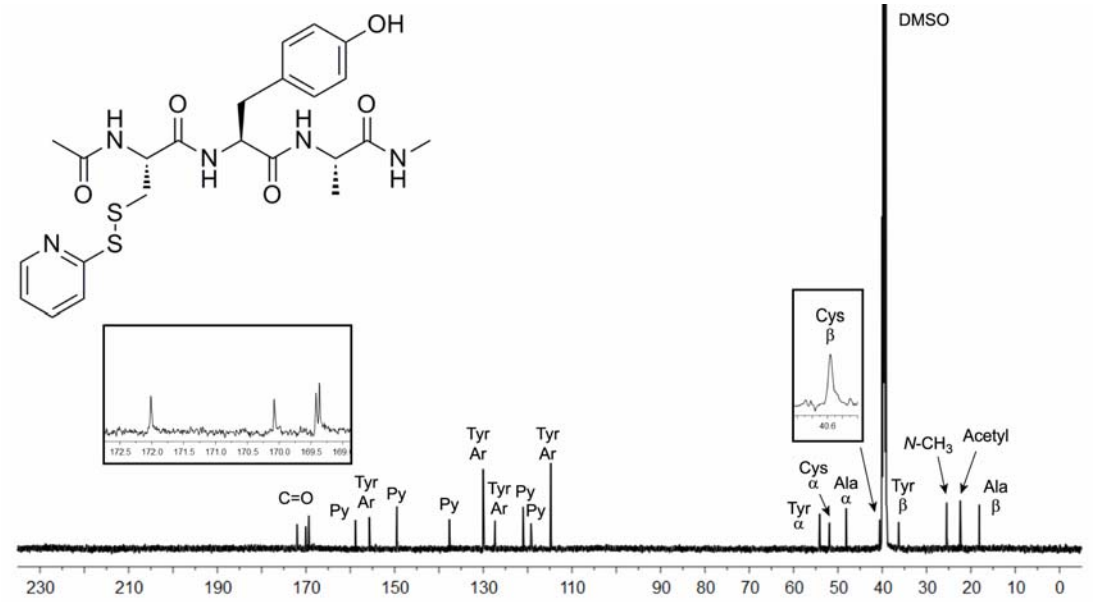
Each PyS-CXA tripeptide was conjugated to DNA to form DNA-C_3-CXA or DNA-HEG-CXA. DNA-C_3-SS-C_3-OH or DNA-HEG-SS-C_3-OH was reduced with DTT (50 mM HEPES, pH 7.5, 50 mM DTT, 37 °C, 2 h) and precipitated with ethanol. The conjugation reaction was performed using 3 nmol of DNA-C_3/HEG-SH in 20 µL of 50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

Characterization of the tripeptide substrates CYA, CSA, and CAA

CYA tripeptide

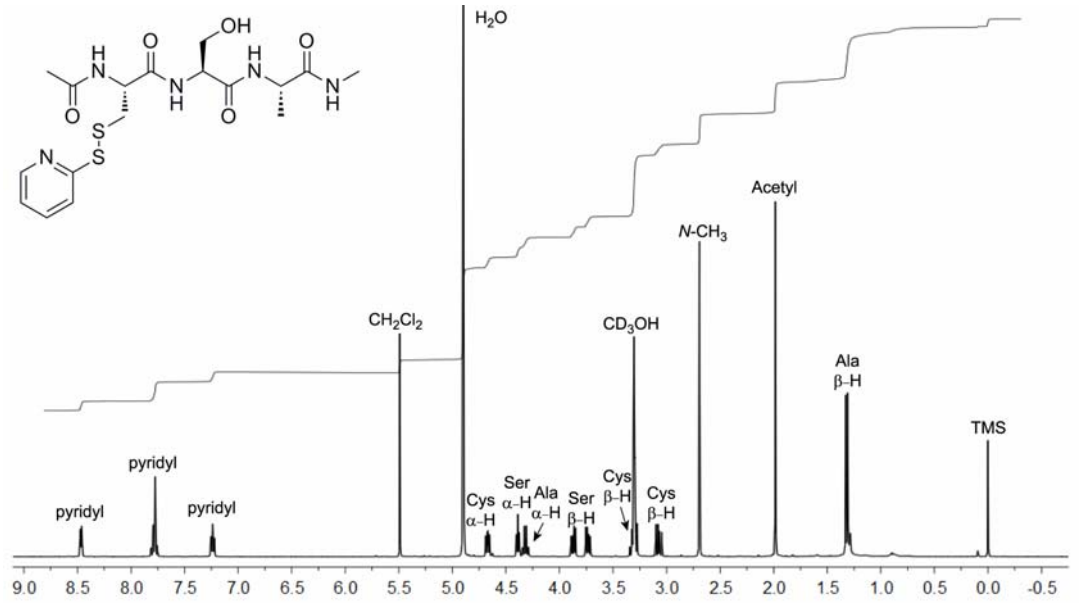


^{1}H NMR (400 MHz, DMSO-d_6): δ 9.19 (s, 1H), 8.43 (d, *J* = 4.8 Hz, 1H), 8.30 (d, *J* = 8.0 Hz, 1H), 8.04 (d, *J* = 7.7 Hz, 2H), 7.81 (td, *J* = 7.7 Hz, <2 Hz, 1H), 7.74 (d, *J* = 8.1 Hz, 1H), 7.53 (m, 1H), 7.24 (dd, *J* = 6.5, 5.0 Hz, 1H), 6.97 (d, *J* = 8.4 Hz, 2H), 6.61 (d, *J* = 8.4 Hz, 2H), 4.49 (td, *J* = 8.5, 5.1 Hz, 1H), 4.38 (td, *J* = 8.1, 5.0, 1H), 4.16 (quintet, *J* = 7.1 Hz, 1H), 3.10 (dd, *J* = 13.3, 4.9 Hz, 1H), 2.93-2.87 (m, 2H), 2.68 (dd, *J* = 13.9, 8.8 Hz, 1H), 2.54 (d, *J* = 4.6 Hz, 3H), 1.83 (s, 3H), 1.15 (d, *J* = 7.2 Hz, 3H) ppm.

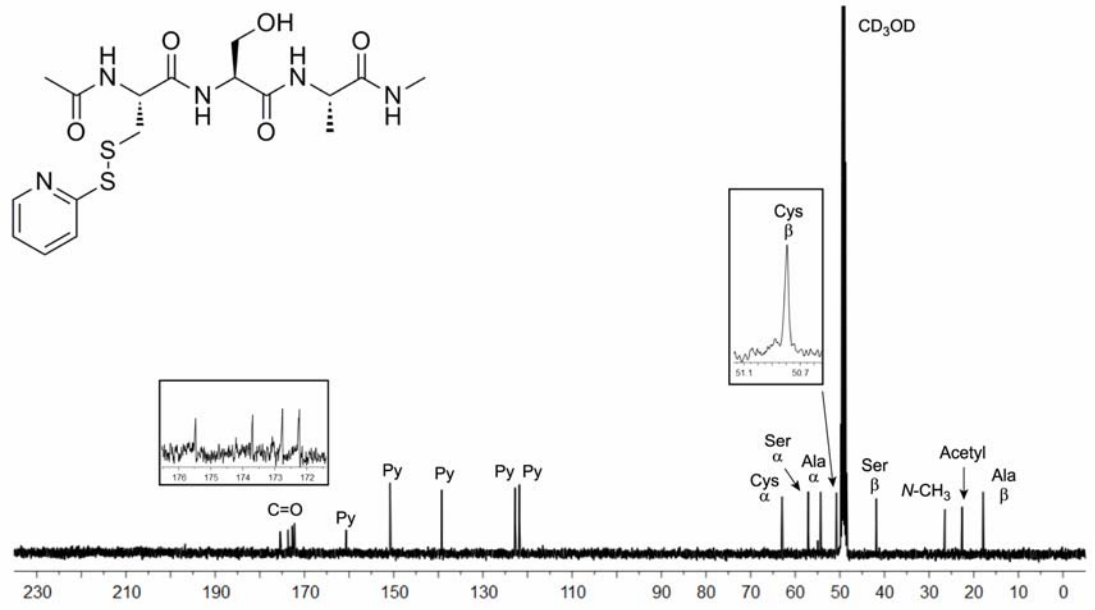


^{13}C NMR (126 MHz, DMSO-d_6): δ 172.01, 170.08, 169.42, 169.37, 158.85, 155.70, 149.45, 137.63, 130.04, 127.30, 121.05, 119.20, 114.73, 54.12, 51.87, 48.12, 40.58, 36.30, 25.41, 22.35, 18.06 ppm.
ESI-HRMS: *m/z* calcd. for $C_{23}H_{29}N_5O_5S_2$ $[M+H]^+$ 520.1688, found 520.1705 (Δm +0.0017, error +3.3 ppm).

CSA tripeptide



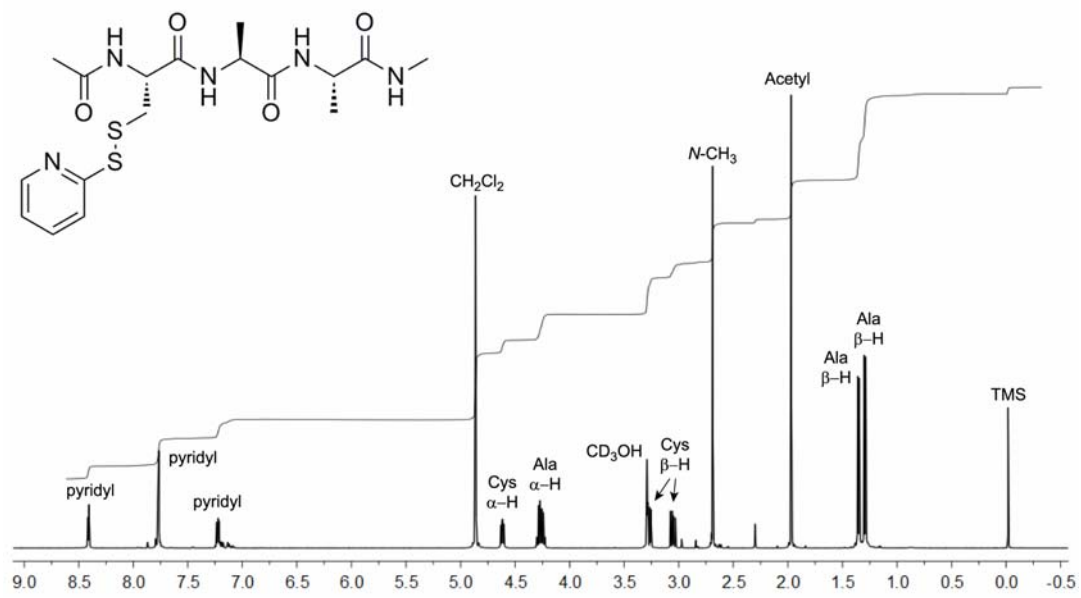
^{1}H NMR (400 MHz, CD_3OD): δ 8.47 (ddd, *J* = 5.0, 1.6, 1.2 Hz, 1H), 7.82-7.75 (m, 2H), 7.24 (ddd, *J* = 6.5, 4.9, 1.9 Hz, 1H), 4.67 (dd, *J* = 8.5, 5.5 Hz, 1H), 4.39 (t, *J* = 5.7 Hz, 1H), 4.30 (q, *J* = 7.3 Hz, 1H), 3.87 (dd, *J* = 10.9, 5.2 Hz, 1H), 3.73 (dd, *J* = 10.8, 6.0 Hz, 1H), 3.29 (dd, *J* = 13.8, 5.6 Hz , 1H), 3.06 (dd, *J* = 13.9, 8.5 Hz, 1H), 2.69 (s, 3H), 1.99 (s, 3H), 1.32 (d, *J* = 7.3 Hz, 3H) ppm.



^{13}C NMR (101 MHz, CD_3OD): δ 175.46, 173.68, 172.75, 172.22, 160.66, 150.82, 139.27, 122.79, 121.84, 62.95, 57.09, 54.35, 50.79, 41.83, 26.49, 22.62, 17.93 ppm.

ESI-HRMS: *m/z* calcd. for $C_{17}H_{25}N_5O_5S_2$ $[M+Na]^+$ 466.1185, found 466.1187 (Δm +0.0002, error +0.4 ppm).

CAA tripeptide



^{1}H NMR (500 MHz, CD_3OD): δ 8.43 (ddd, *J* = 4.9, 1.7, 1.0 Hz, 1H), 7.83-7.77 (m, 2H), 7.24 (ddd, *J* = 6.0, 4.9, 2.5 Hz, 1H), 4.64 (dd, *J* = 8.6, 5.4 Hz, 1H), 4.30 (q, *J* = 7.2 Hz, 1H), 4.27 (q, *J* = 7.2 Hz, 1H), 3.30 (dd, *J* = 13.9, 5.4Hz, 1H), 3.07 (dd, *J* = 13.9, 8.6 Hz, 1H), 2.71 (s, 3H), 1.99 (s, 3H), 1.36 (d, 7.2 Hz, 3H), 1.29 (d, 7.2 Hz, 3H) ppm.

^{13}C NMR (101 MHz, CD_3OD): δ 175.40, 174.52, 173.62, 172.59, 160.69, 150.62, 139.24, 122.74, 121.77, 54.20, 51.10, 50.60, 41.69, 26.51, 22.66, 18.19, 17.69 ppm.

ESI-HRMS: *m/z* calcd. for $C_{17}H_{25}N_5O_4S_2$ $[M+H]^+$ 428.1426, found 428.1422 (Δm –0.0004, error –0.9 ppm).

25% *Partially randomized N_{40} pool preparation.* Pre-mix solutions were made in which 75% of the oligonucleotide in that position was the same base as in the original nucleotide, and in the remaining 25% of the oligonucleotide, that position had an equal chance of being any of the other three bases.

All the standard stock solutions of phosphoramidite were 0.1 M. In general, an upper estimate of 200 μL of the phosphoramidite solution was required for each coupling by an ABI 394 DNA synthesizer. Also, a minimum 1000 Å size pore size controlled pore glass (CPG) was used to minimize the steric effect of the long nucleotide synthesis. Below is the method to prepare 1 mL of four 25% partially randomized phosphoramidite stock solution. The calculation took into account the relative coupling efficiencies of the four phosphoramidites (A = 0.2, G = 0.26, C = 0.24. T = 0.30).

N* denotes the 25 % partially randomized solution, and N denotes the 0.1 M phosphoramidite stock.

A* = A (799 μL) + G (68 μL) + C(74 μL) + T (59 μL) = 1000 μL

G* = A (106 μL) + G (735 μL) + C(88 μL) + T (71 μL) = 1000 μL

C* = A (101 μL) + G (77 μL) + C(755 μL) + T (67 μL) = 1000 μL

T* = A (116 μL) + G (89 μL) + C(97 μL) + T (697 μL) = 1000 μL

2.4.2 In vitro selection procedure

Selection Initiation. To prepare a ligated pool (RNA ligated to the pool), 200 pmol of the pool of deoxyribozymes prepared by solid-phase synthesis was ligated to the RNA with T4 RNA ligase. In the first round, 400 pmol of L was incubated with 200 pmol of the ligated pool in a total volume of 40 μL in 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ at 37 ºC for 2 h. The sample was quenched by 40 μL stop solution (80% formamide, 1X TBE [89 mM each Tris and boric acid, pH

8.3, 50 mM EDTA], and 0.025% each xylene cyanol and bromophenol blue). The sample was purified by 20% PAGE.

PCR amplification. The ligated product was excised from PAGE, extracted, and precipitated with ethanol. The ligated product was resuspended in 83 μL of water and PCR- amplified (94 ºC, 30 s; 47 ºC, 30 s; 72 ºC, 30 s) by *Taq* polymerase with 0.25 μM of primer 1 and 1 μM of primer 2 in a total volume of 100 μL for 10 cycles. Phenol-chloroform-isoamyl alcohol extraction was performed. 1 μL of the PCR product was used as the templates for a 30-cycle PCR in a total volume of 50 μL. The PCR product was purified by 8% PAGE. The shorter PCR product was excised, extracted and precipitated by ethanol. PCR material was ligated to the RNA substrate with T4 RNA ligase and purified by 8% PAGE.

Selection. The ligated material was incubated with 100 pmol of L in a total volume of 20 μL in 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 37 ºC for 2 h and purified by 8% PAGE. PCR amplification, ligation and selection procedures were performed iteratively until ligation activity reached a plateau.

2.4.3 Cloning DNA pools, screening clones for ligation activity and sequencing

Plasmid inserts preparation by PCR. Once ligation activity reached a plateau, cloning procedures were performed to isolate individual deoxyribozymes. First, a 1000-fold dilution was performed on the 10-cycle PCR product. Second, a 30-cycle PCR in 100 μL total volume was performed with 1 μL of the diluted 10-cycle PCR product by *Taq* polymerase with 1 μM cloning primers that encode the stop condons (in bold). Two Cloning primers are listed as follows: `5'-TAATTAATTAATTAC`**`CCATCAGGATCAGCT`**`-3',` and

5′-TAATTAATTAATTA**CGAAGTCGCCATCTC**-3′

The PCR products were isolated on a 2% agarose gel, and extracted by Fermentas GeneJET extraction kit. Invitrogen TOPO TA cloning kit was used to insert the *Taq* polymerase-amplified PCR product to the plasmid vector. The detailed protocol can be obtained from the manual. In the following, only the key procedures were mentioned. The subcloning reaction was initiated with the addition of 1 μL of the extracted PCR product, 1 μL TOPO vector plasmid, and 1 μL of salt solution (Invitrogen) in a 6 μL total volume. The subcloning reaction was incubated at room temperature for 10 min. To initiate transformation, 2 μL of the subcloning reaction were added to the One Shot® Chemically Competent *E. coli.* cells and incubated at room temperature for 10 min. The cells were heat shock for 30 seconds at 42 ^{o}C, followed by addition of 250 μL SOC and shaken horizontally at 200 rpm for 37 ^{o}C for 1 h. Before plating, 40 μL of X-gal (40 mg/mL in dimethylformamide) and 40 μL of 100 mM of IPTG were spread on the pre-warmed Luria-Bertani (LB) plates containing 100 μg/mL of ampicillin. The cells were spread on the LB plates and incubated at 37 ^{o}C for 18 h.

After 18 h, 20–40 white colonies were chosen and cultured in 5 mL culture tubes containing 3 mL of LB medium containing 40 μg/mL of ampicillin. The cells were incubated while shaken at 220 rpm in 37 ^{o}C for 18 h. The cells were pelleted and the plasmids were isolated by the Fermentas Miniprep Kit.

Kinetic screening of individual clones. 30-cycle PCR was performed for individual clones using 0.25 μM of primer 1 and 1 μM of primer 2 in a total volume of 200 μL. The PCR product was ethanol precipitated and isolated by 8% PAGE. The extracted product should have a lower limit of 10 pmol after extraction. 0.5 pmol of 3′-^{32}P-radiolabeled

RNA substrate, 5 pmol of the PCR product, and 10 pmol of the DNA-C3/HEG-CXA was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 1 min and cooling on ice for 5 min. The ligation reaction was initiated by addition of stock solutions to a final volume of 10 μL containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$. At an appropriate time point, 2 μL portion of the reaction was taken out and quenched by 6 μL of stop solution, followed by 20% PAGE.

Sequencing. Sequencing was preformed with the reverse primer, M13Rev-24, 5′-AACAGCTATGACCATG-3′ (T_m = 45 °C). When the initial sequencing result in ambiguous nucleotides, the forward primer was used instead. The sequence of the forward primer M13For-21 was 5′-GTAAAACGACGGCCAGT-3′ (T_m = 58 °C). The chromatogram was analyzed by Chromas and the sequence alignment was obtained by CLC Sequence Viewer 6.

2.4.4 Deoxyribozyme activity assays

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

U of T4 RNA ligase (Fermentas) in 20 μL of 50 mM Tris, pH 7.5, 5 mM $MgCl_2$, 10 mM DTT, 0.05 mM ATP at 37 °C for 12 h, followed by 20% PAGE.

Kinetic assays with DNA-Anchored tripeptide substrates. The DNA-tripeptide substrate is designated as the L (left-hand) substrate, and the 5′-triphosphorylated RNA substrate is designated as the R (right-hand) substrate. In each deoxyribozyme assay, R was the limiting reagent relative to the deoxyribozyme (E) and the L substrate. A 14 μL sample containing 1 pmol of partially 3′-^{32}P-radiolabeled, 10 pmol of E, and 20 pmol of L was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 1 min and cooling on ice for 5 min. The ligation reaction was initiated by addition of stock solutions to a final volume of 20 μL containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$. Final concentrations were 50 nM R, 0.5 μM E, and 1 μM L (1:10:20 R:E:L; single-turnover conditions). At appropriate times, 1.4 μL aliquots were quenched with 6 μL of stop solution. Before PAGE, 100 pmol of a "decoy oligonucleotide" was added to each sample. The decoy oligonucleotide was a 60-mer complementary to the enzyme region (~40 nt) along with 10 nt of binding arm on either side. This decoy oligonucleotide was added to displace the deoxyribozyme from the substrate and ligation product. Without the decoy oligonucleotide, the gel bands were noticeably smeared, which inhibited proper quantification.

Kinetic assays with free tripeptide substrates. Similar to the procedures above, with the exception that that the 20 pmol of DNA-tripeptide substrate was replaced with 20 pmol of DNA helper oligonucleotide and 20 nmol of free tripeptide was added after the annealing step. 50 mM stock solutions of tripeptide were prepared in formamide. Final

concentrations were 50 nM R, 0.5 μM E, 1 μM helper oligonucleotide, and 1 mM tripeptide substrate (1:10:20:20,000 R:E:helper:tripeptide; single-turnover conditions), with 2% formamide by volume. For determining $K_{d,app}$ of the CYA tripeptide for the 15MZ36 deoxyribozyme, the concentration of tripeptide was varied between 0 and 3 mM, with 2% formamide in the final sample in all cases. For experiments in which multiple turnover was examined, the concentrations were 0.5 μM R, 50 nM E, 10 μM helper oligonucleotide, and 1 mM tripeptide substrate (10:1:200:20,000 R:E:helper:tripeptide).

Multiple turnover assays with various 5′-truncated 15MZ36. The kinetic assays were performed in a similar manner as the above procedure. The key differences are as follows:

For 100% E reaction, the reaction involved 1 pmol of radiolabeled R along with 9 pmol of R (10 pmol of R in total), and 10 pmol of E (100%) and 200 pmol of the helper oligonucleotide in 1 mM CYA (1:1:20:2000 R:E:helper:tripeptide). 10 pmol of R instead of 1 pmol of R was used to avoid making the deoxyribozyme too diluted for the assay using 10% E.

For 10% E reaction, the reaction involved 1 pmol of radiolabeled R along with 9 pmol of R (10 pmol in total), and 1 pmol of E (10%) and 200 pmol of the helper oligonucleotide in 1 mM CYA (10:1:200:20,000 R:E:helper:tripeptide).

RNase T1 digestion of Ser-RNA and Tyr-RNA products. RNase T1 digestion of DNA-C_3/HEG-CXA–RNA, 100 pmol of ligation product was incubated in 50 mM Tris, pH 7.5, 2 mM EDTA, and 20 U of RNase T1 (Ambion) in 20 μL total volume at 37 °C for 4 h, extracted by phenol/chloroform, precipitated by ethanol, and desalted with C_{18} ZipTip.

For RNase T1 digestion of untethered CYA–RNA, the same procedure was used, except extraction and precipitation were omitted. All MALDI mass spectra were obtained in the mass spectrometry laboratory of the UIUC School of Chemical Sciences.

DTT treatment of Ser-RNA and Tyr-RNA products. For DTT reduction, 100 pmol of the ligation product was incubated in 50 mM HEPES, pH 7.5, and 50 mM DTT in 20 μL total volume at 37 °C for 2 h, precipitated with ethanol, and desalted with C_{18} ZipTip (Millipore).

2.5 References

1 Joshi, N. S.; Whitaker, L. R.; Francis, M. B., A three component Mannich-type reaction for selective tyrosine bioconjugation. *J. Am. Chem. Soc.* **2004**, *126*, 15942–15943.

2 Antos, J. M.; McFarland, J. M.; Iavarone, A. T.; and Francis, M. B., Chemoselective tryptophan labeling with rhodium carbenoids at mild pH. *J. Am. Chem. Soc.* **2009**, *131*, 6301–6308.

3 Ban, H.; Gavrilyuk, J.; Barbas, C. F., III., Tyrosine bioconjugation through aqueous ene-type reactions: A click-like reaction for tyrosine. *J. Am. Chem. Soc.* **2010**, *132*, 1523–1525

4 Jackel, C.; Kast, P.; and Hilvert, D., Protein design by directed evolution. *Annu. Rev. Biophys.* **2008**, *37*, 153–173.

5 Baskerville, S.; and Bartel, D. P. A ribozyme that ligates RNA to protein. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9154–9159.

6 Pradeepkumar, P. I.; Hobartner, C.; Baum, D. A.; Silverman, S. K., DNA-Catalyzed Formation of Nucleopeptide Linkages. *Angew. Chem., Int. Ed.* **2008**, *47*, 1753–1757.

7 Tse, Y.-C.; Kirkegaard, K.; Wang, J. C., Covalent bonds between protein and DNA. Formation of phosphotyrosine linkage between certain DNA topoisomerases and DNA. *J. Biol. Chem.* **1980**, *255*, 5560–5565.

8 Deweese, J. E., and Oshero□, N., The DNA cleavage reaction of topoisomerase II: Wolf in sheep's clothing. *Nucleic Acids Res.* **2008**, *37*, 738–748.

9 Baker, N. M., Rajan, R., and Mondragon, A., Structural studies of type I topoisomerases. *Nucleic Acids Res.* **2009**, *37*, 693–701.

10 Ledesma, F. C.; El Khamisy, S. F.; Zuma, M. C.; Osborn, K.; Caldecott, K. W., A human 5′-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage. *Nature* **2009**, *461*, 674–678.

11 Grindley, N. D. F.; Whiteson, K. L.; and Rice, P. A. Mechanisms of site-specific recombination. *Annu. Rev. Biochem*. **2006**, *75*, 567–605.

12 Ambros, V.; Baltimore, D. Protein is linked to the 5′ end of poliovirus RNA by a phosphodiester linkage to tyrosine. *J. Biol. Chem.* **1978**, *253*, 5263–5266.

13 Hermoso, J. M.; Salas, M. Protein p3 is linked to the DNA of phage φ29 through a phosphoester bond between serine and 5′-dAMP. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 6425–6428.

14 Penalva, M. A.; Salas, M. Initiation of phage φ29 DNA replication in vitro: Formation of a covalent complex between the terminal protein, p3, and 5′-dAMP. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 5522–5526.

15 Samad, A.; Carroll, R. B. The tumor suppressor p53 is bound to RNA by a stable covalent linkage. *Mol. Cell. Biol.* **1991**, *11*, 1598–1606.

16 Geyer, C. R.; Sen, D. Lanthanide probes for a phosphodiester-cleaving, lead-dependent, DNAzyme. *J. Mol. Biol.* **1998**, *275*, 483-489.

Chapter 3: DNA-Catalyzed Reductive Amination[a]

3.1 Introduction

This chapter describes a new in vitro selection method that sought DNA catalysts which modify the side chain of a free peptide substrate. First, the free peptide was allowed to react with the 5′-triphosphorylated RNA. Second, active deoxyribozymes were captured by reaction of the peptide N-terminus amino group with an oligonucleotide terminating with aldehyde functional groups. Despite the in vitro selection design, we identified deoxyribozymes that catalyzes reductive amination reaction using a weakly nucleophilic N^2-amine of a guanine of the RNA substrate, instead of the originally intended N-terminus amino nucleophile. Reductive amination involves in key biochemical processes, for example, for amino acid biosynthesis and catabolism using amino acid dehydrogenases (oxidases)[1] and transaminases (aminotransferases)[2-3]. In vitro, enzymatic reductive amination is important in both laboratory-scale[4-7] and industrial-scale organic synthesis[8-9]. Reductive amination may also have been important in RNA World scenarios.[10-11] The unexpected discovery of DNA-catalyzed reductive amination expands the catalytic repertoire of deoxyribozymes. DNA-catalyzed reductive amination may have a practical application as a tool for labeling nucleic acids site-specifically.

---

[a] The material described in this chapter has been published.

Wong, O. Y.; Mulcrone, A. E.; Silverman, S. K., DNA-Catalyzed Reductive Amination. *Angew. Chem. Int. Ed.* **2011**, *50*, 11679–11684.

3.2 Results and Discussion

3.2.1 Selection design using free tripeptide

In the key selection step of in vitro selection, the active deoxyribozymes are physically separated from the inactive DNA sequences. The separation can be done by affinity capture, size separation or both. In the Silverman laboratory, we employ the size separation method by PAGE. The size separation method was chosen to avoid artifacts which could be associated with bead-based selection methods. For bead-based selection methods, noncatalytic nucleic acid sequences could be bound to the beads even after several stringent washes. As described by previous efforts in Chapter 2, tethered tripeptide substrates were used during selection to provide gel-shift selection products, which could be readily separated by PAGE (**Fig. 3.1A**). In this project, a direct selection method was designed to find deoxyribozymes that catalyze the ligation of the potential nucleophile of a tripeptide to the electrophile of a 5′-triphosphorylated RNA. During the selection step, the catalytic active DNA sequences catalyze the ligation of a particular peptide to the RNA substrate. However, the additional of the tripeptide to the RNA substrate would not have lead to a substantial PAGE shift. Therefore, an alternate strategy was developed to enable the isolation of the ligation products after selection. We designed a two-stage selection strategy which involved a selection step followed by a capture step. In the selection step, deoxyribozymes appended themselves to the free CYA tripeptide substrate via reactions between the Tyr nucleophile with the 5′-triphosphate of the RNA. In the capture step, an oligonucleotide with 3′ dialdehyde reacts with the N-

terminus of the free peptide of the selection product via reductive amination, which leads to a gel-shift product that could be isolated by PAGE (**Fig. 3.1B**).

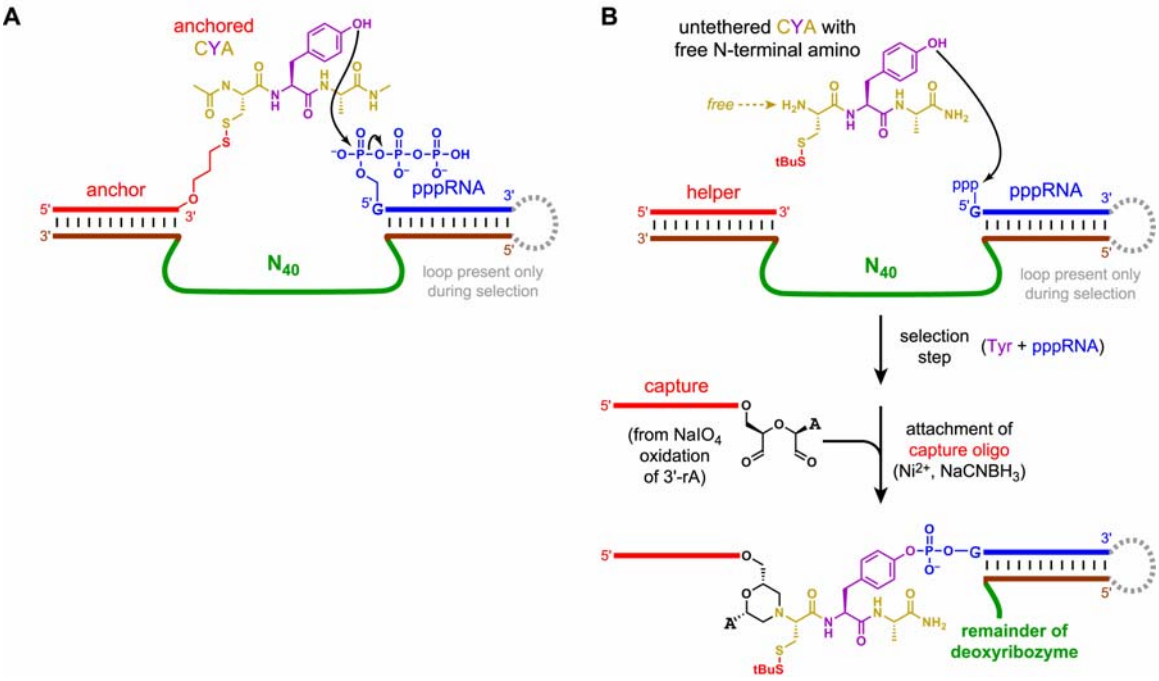


Figure 3.1 Selection strategy for direct selection using free CYA tripeptide. (A) Previous selection efforts used DNA-C_3/HEG-CXA as substrates. Although several deoxyribozymes were found to function well with CYA, they require CYA in mM concentration. A selection that uses CYA directly instead of its tethered version may yield deoxyribozymes that function better with the CYA tripeptide than the previous selection efforts described in Chapter 2. (B) The new selection design that employed an additional capture step to physically isolate the DNA-catalyzed peptide-RNA ligation product. The capture step involves a reductive amination of the N-terminus peptide-RNA ligation product with an oligonucletiode terminating with two aldehyde functional groups. The oligonucleotide containing aldehyde functional groups is denoted as the capture oligonucleotide.

To establish a reliable reductive amination method for the capture step, the reductive amination was optimized by using the peptide-RNA-DNA substrate (see Material and Methods for preparation of the CYA-RNA-DNA ligation product). The reductive amination method was optimized to yield ~50% of the product.

3.2.2 Direct selection using free tripeptide (QA selection)

We initiated selection with a free CYA tripeptide to identify deoxyribozymes that catalyze the nucleophilic attack of the Tyr-OH to the RNA substrate, denoted as the QA selection. We first observed discernable activity (2%) in round 4 and the activity reached a plateau in round 8 (35.6%), where the individual clones were sequenced and cloned (**Fig. 3.2**).

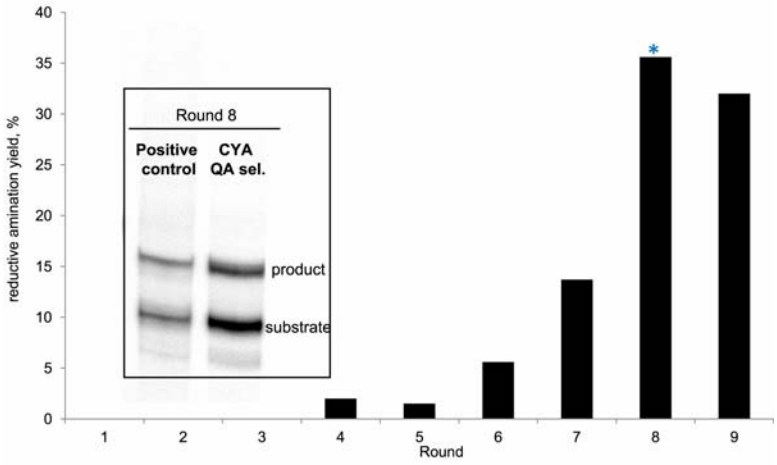


Figure 3.2 Direct selection using CYA, denoted as the QA selection. The inset illustrates the key round of the QA selection and the independent positive control. The asterisk marks the round where deoxyribozymes were cloned and sequenced.

3.2.3 Tripeptide-independent reductive amination catalyzed by QA deoxyribozymes

To test the reductive amination product was a direct result of the DNA-catalyzed free peptide ligation to the RNA substrate, we first assayed the uncloned pool with or without the free CYA tripeptide substrates, and also with a peptide substrate where its N-terminus was blocked with an acetyl group (**Fig. 3.3A**). The omission of CYA and the absence of a primary amine of the free peptide still led to substantial reductive amination product. We concluded that the activity did not involve the tripeptide substrate. To make sure the observed activity was due to reductive amination, we tested the uncloned pool with or without key reactants of the reductive amination. **Fig. 3.3** illustrates that CYA was not involved in the DNA-catalyzed reductive amination.

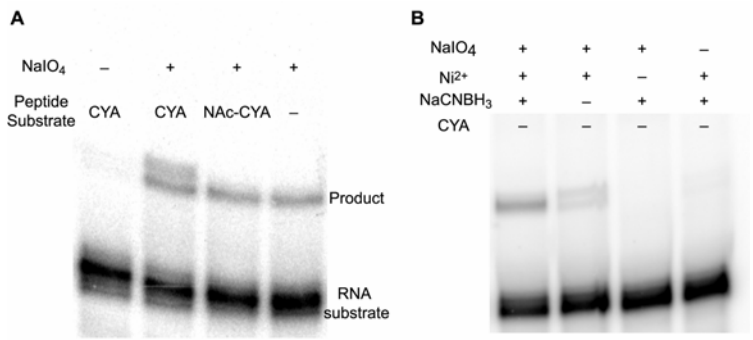


Figure 3.3 CYA was not involved in the reductive amination reaction. (A) The 7QA uncloned pool was assayed in the following conditions: with *N*-acetylated CYA, with or without CYA, and with or without the aldehyde-containing oligonucleotide. The absence of aldehyde (omitting the oxidation treatment using sodium periodate to the oligonucleotide terminating with a ribonucleotide at the 3′ end) abolished the product formation, demonstrating the aldehydes were mandatory in the reaction. Reaction with the *N*-acetylated CYA led to substantial activity, showing that the CYA was not involved in the reaction as the N-terminus of the CYA was blocked with an acetyl group and thus unavailable for the reaction. In addition, in absence of CYA, a significant amount of reductive amination product was observed; further supporting CYA was not involved in the reaction. (B) The 7QA uncloned pool-catalyzed reaction was further confirmed to be reductive amination as omitting key reactants led to no product formation. Without the nickel cations or sodium cyanoborohydride, the activity of the uncloned pool was greatly suppressed.

3.2.4 Cloning and sequencing of 8QA deoxyribozymes and their kinetics

The reductive amination product did not involve the tripeptide substrate yet interesting as we postulated that the reaction was between the nucleophilic amine of nucleobase and the capture oligonucleotide. The 8QA selection was cloned and kinetic screening of the clones was performed (**Fig. 3.4**). We sequenced all the active clones. **Fig. 3.5** illustrates the unique sequences from the 8QA selection effort. It was found that 8QA108 was identical to 8QA114; 8QA113 had 1 nt more than 8QA115.

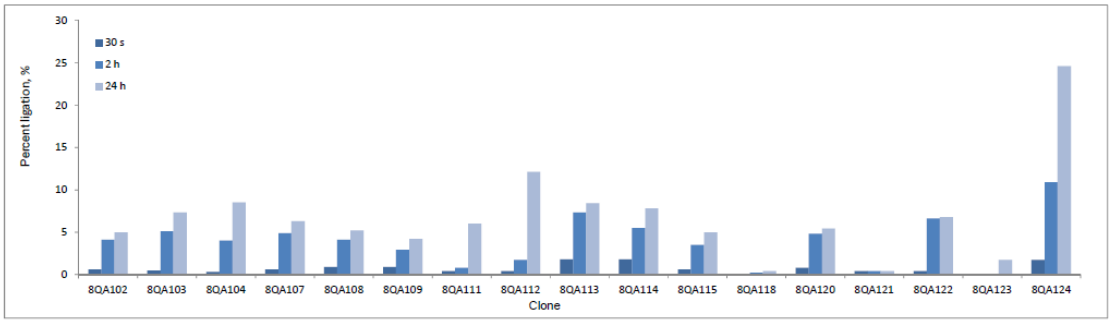


Figure 3.4 Kinetic screening of 8QA clones. Time points: 30 s, 2 h, 24 h. All the active clones were sequenced.

```
                1                  10                  20                  30                  40
                |                  |                   |                   |                   |
8QA124  AAGGGGCAGT  GCTAAACGTG  GACAGGAATC  ATAAGTGTCA
8QA102  ......T...  ..........  A.G.T.GT.T  GGGTAATCTC
8QA103  ..T..CAGTG  AGA..CGTAT  .GTCAC.TGT  ..GT.GTCAC
8QA107  ......T...  ..C.......  ..A.CCTG.G  GAT.AG...C
8QA111  GG...C.GCG  AT....T..T  .T..T.C.G.  TCTCTGAGGC
8QA115  .G...CAGTG  CTA..CGTGT  ACAGACT.AA  ..GGTCTGTC
8QA122  ........CG  TGCT..ACGT  .GA....C.T  CGGTTCTC.C
8QA104  .G...TAGTG  C.A..CGTG.  CCTG.TTTGA  TG.TTC.G.C
8QA108  ......T...  ..C......A  CGA...CT..  ..T.CC.A.C
8QA109  ...A....CG  TGCT..ACGT  .GA....C.T  CGGTTCTC..
```

Figure 3.5 Sequences of new deoxyribozymes that catalyze reductive amination. The first seven deoxyribozymes were studied in greater details. The highlighted deoxyribozyme, 8QA124, had the best rate and yield compared to the others. The last three entries were not studied in details. The originally randomized (N_{40}) catalytic region is shown. In each deoxyribozyme, the sequence is 5′-CCGTCGCCATCTCTTC-(N_{40})-ATAGTGAGTCGTATTA-3′.

One of the new deoxyribozymes, 8QA124, catalyzed the reductive amination between the RNA substrate and the $NaIO_4$-oxidized capture oligonucleotide in 60–90% yield with k_{obs} of 0.37 h^{-1} ($t_{1/2}$ 1.7 h; **Fig. 3.6**). The other deoxyribozymes had yields ranging from 18–55% after 22 h (**Fig. 3.7**). To confirm the reaction was a deoxyribozyme-catalyzed reaction, control experiments were performed with a noncatalytic template strand omitting the N_{40} region, or replacing the N_{40} region with unpaired T nucleotides (T_0, T_1, T_2), the yield in 22 h was only 0.6%, 0.4% and 0.2%. From the T_0 results, a conservative lower limit of the rate enhancement for 8QA124 was calculated to be 1400 (8QA124 k_{obs} of 0.37 h^{-1} / T_0 k_{obs} of 0.00027 h^{-1}).

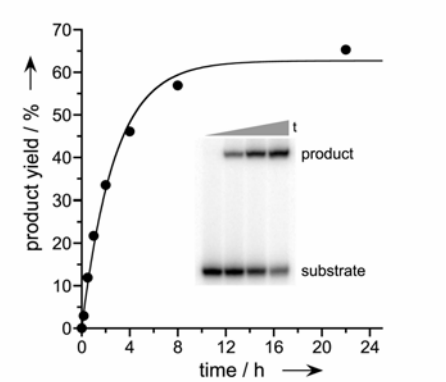


Figure 3.6 Kinetic plot of the activity of 8QA124 deoxyribozyme. The RNA substrate was 3′-^{32}P-radiolabeled. Incubation conditions: 100 mM NaOAc, pH 5.2, 50 mM $NiCl_2$, and 10 mM $NaCNBH_3$ at 37 °C.

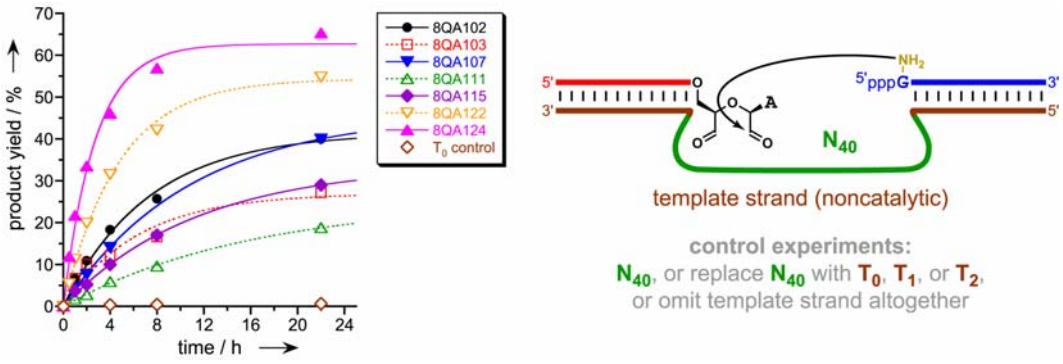


Figure 3.7 Catalytic activities of the seven new deoxyribozymes studied in this chapter. The assay procedure is described in the Materials and Methods section. k_{obs} values (h^{-1}): 8QA124 0.37; 8QA102 0.14; 8QA103 0.16; 8QA107 0.092; 8QA111 0.062; 8QA115 0.088; 8QA122 0.21.

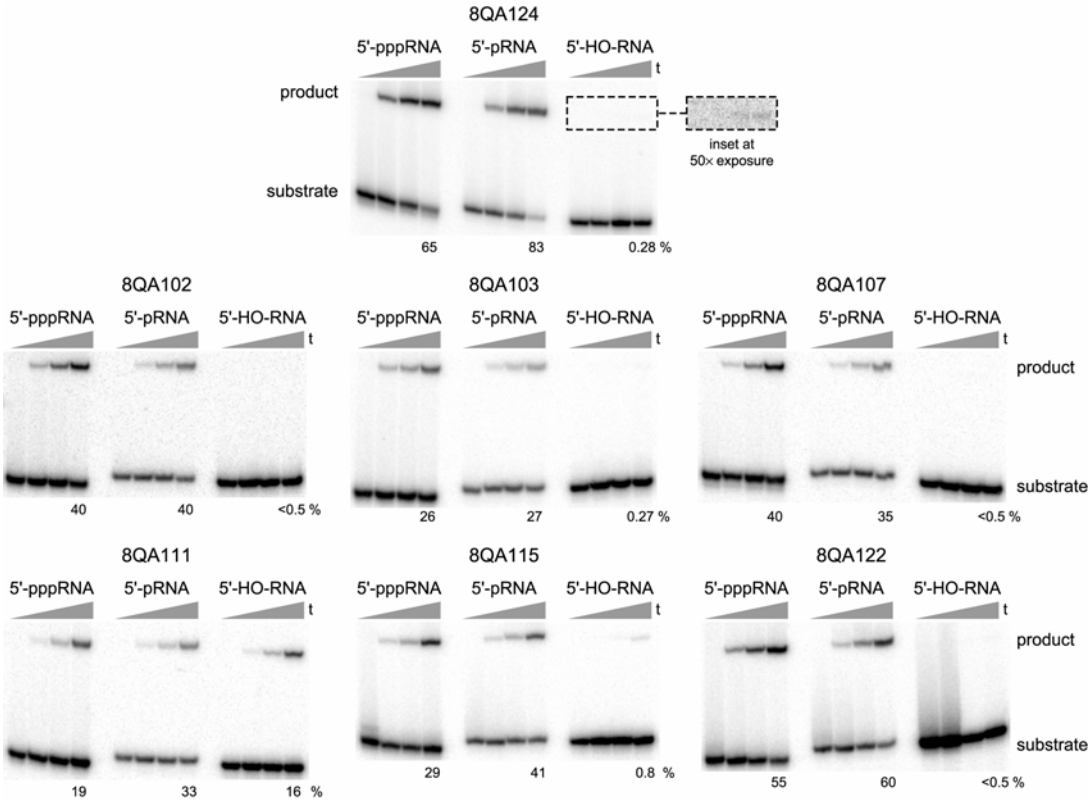


Figure 3.8 Dependence of the catalytic activities on the RNA 5′-terminus. Each of the seven new deoxyribozymes was assayed with RNA substrates that have either 5′-triphosphate, 5′-monophosphate, or 5′-hydroxyl terminus. The results of these assays showed that all seven deoxyribozymes function well with 5′-monophosphorylated RNA. In contrast, reactivity varied widely involving the 5′-hydroxyl RNA substrate. 8QA111 has relatively robust activity with 5′-hydroxyl RNA; three of the deoxyribozymes (8QA124, 8QA103, and 8Q115) have low but detectable activity with 5′-hydroxyl RNA; and the remaining three deoxyribozymes (8QA102, 8QA107, and 8QA122) have no detectable activity with 5′-hydroxyl RNA.

3.2.5 Dependence of the catalytic activities on the RNA 5′-terminus

We examined the 5′ terminus requirement for all the seven 8QA deoxyribozymes. All of the deoxyribozymes functioned well with the 5′-triphosphate RNA as well as 5′-monophosphate RNA, demonstrating that the originally intended electrophilic triphosphate was not involving in the reaction. Six of the seven deoxyribozymes do not catalyze the reaction with the 5′-hydroxyl RNA substrate (**Fig. 3.8**).

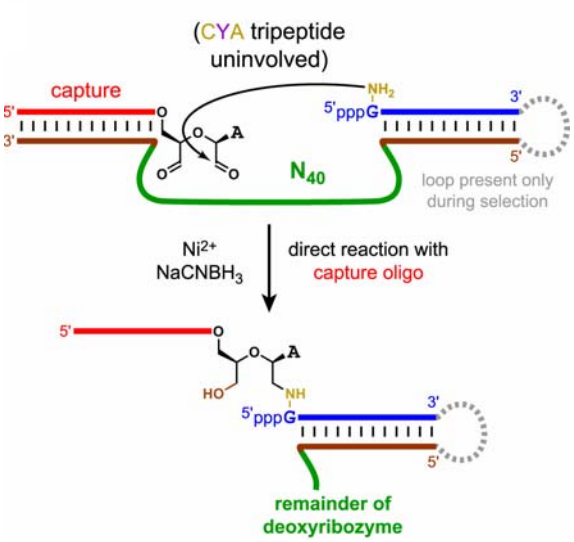


Figure 3.9 Proposed reductive amination catalyzed by 8QA124 as well as the other 8QA deoxyribozymes.

3.2.6 Dependence of the catalytic activities on $NaIO_4$, Ni^{2+} and $NaCNBH_3$

To further confirm that the reaction catalyzed by the 8QA deoxyribozymes was reductive amination, we assayed each of the deoxyribozymes to find if they required $NaIO_4$, Ni^{2+} and $NaCNBH_3$. $NaIO_4$ was used for oxidation of the precursor capture oligonucleotide to the aldehyde-containing capture oligonucleotide. All of the 8QA deoxyribozymes strictly required the aldehyde functional groups and Ni^{2+} for reductive amination. We proposed that one of the amino group of a nucleobase on the RNA substrate was catalyzed by 8QA deoxyribozymes to react with the aldehyde-containing capture oligonucleotide to from an adduct product (**Fig. 3.9**). Mass spectrometry analyses and biochemical studies were performed to provide evidence for the proposed reaction.

We assayed each of the deoxyribozymes for their requirement for $NaIO_4$ (used for oxidation of the precursor capture oligonucleotide to the aldehyde-containing capture oligonucleotide), Ni^{2+}, and $NaCNBH_3$. The 8QA deoxyribozymes strictly require $NaIO_4$, Ni^{2+}, and $NaCNBH_3$. The stringent requirement of key reductive amination component

supported that the 8QA deoxyribozymes catalyzed reductive amination. **Fig. 3.10** provides an illustrative example of how 8QA deoxyribozymes catalysis require $NaIO_4$, Ni^{2+}, and $NaCNBH_3$.

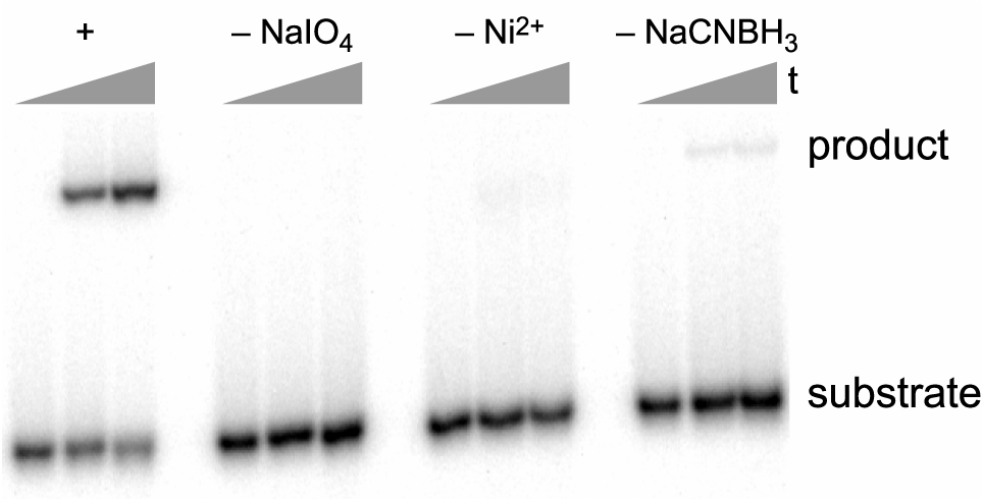


Figure 3.10 Dependence of the catalytic activities on $NaIO_4$, Ni^{2+}, and $NaCNBH_3$, illustrated with data for 8QA124. Data for each of the other six studied deoxyribozymes were similar. Incubation conditions: 100 mM NaOAc, pH 5.2, 50 mM $NiCl_2$, and 10 mM $NaCNBH_3$ at 37 °C (omitting Ni^{2+} or $NaCNBH_3$ as indicated). The "– $NaIO_4$ experiment" was performed by omitting $NaIO_4$ during preparation of the capture oligonucleotide. Time points were taken at 30 s, 4 h, and 22 h. The yield for the "– $NaCNBH_3$" experiment at the 22 h time point was 1.3%.

3.2.7 Characterization of the reductive amination product by RNase T1 treatment

To assign and support our proposed product structure, we obtained the mass of the product of the reductive amination product by MALDI-TOF mass spectrometry. The mass was consistent with a single reductive amination event, in which an exocyclic primary amine attacks one of the two available aldehyde functional groups. The subseqent reaction which involved the secondary amine attack of the second aldehyde to form a morpholine ring was not supported by the mass of the product. To identify the nucleophilic amine in the RNA substrate, we digested the reductive amination product with RNase T1, followed by MALDI-TOF mass spectrometry analysis. RNase T1 specifically cleaves every 3′ of single-stranded G residues of the RNA. After RNase T1 treatment, the truncated product contained two G nucleotides, indicating that the first G nucleotide counting from the 5′-end was not cleaved by RNase T1. With the mass data, we proposed that the exocylic amine of the first G nucleotide (G1) attacks one of the aldehydes of the capture oligonucleotide. RNAse T1 was not able to cleave G1, which was likely due to the covalent linkage between the amine of the G nucleotide base and the capture oligonucleotide (**Fig. 3.11**). **Table 3.1** shows the mass of the reductive amination product of the rest of the 8QA deoxyribozymes and also their masses after RNase T1 treatment.

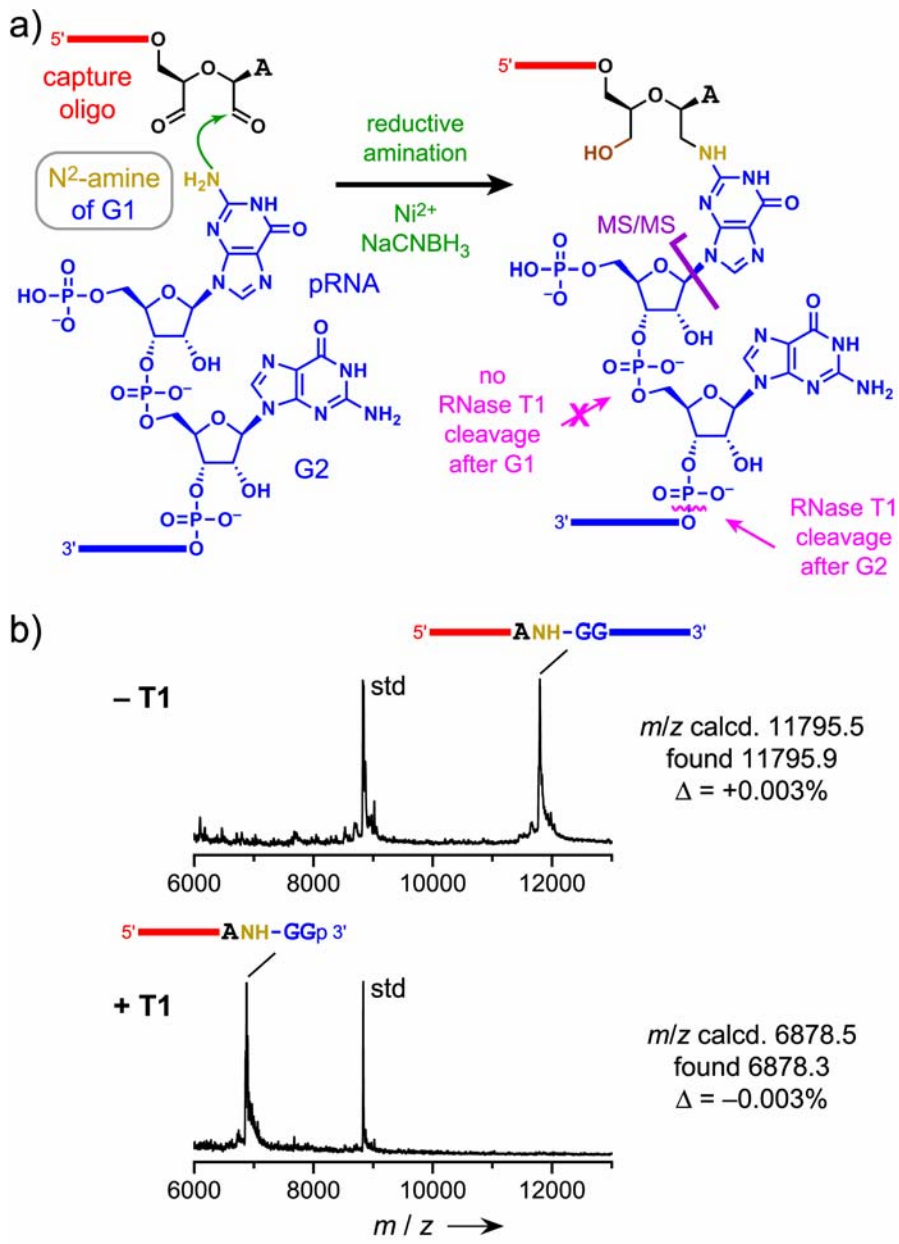


Figure 3.11 Proposed structure of the reductive amination product catalyzed by 8QA124 as well as the other 8QA deoxyribozymes. (a) RNase T1 cleaved after G2 of the product. (b) MALDI-TOF mass analysis of the product, which supported the chemical structure assignment.

deoxyribozyme	RNase T1	mass calcd.	mass found	error, % (found – calcd.)
8QA124	– T1	11795.5	11795.9	+0.003
	+ T1 [a]	6878.5	6878.3	–0.003
	+ T1 [b]	6896.5	6896.6	+0.001
8QA102	– T1	11795.5	11793.5	–0.02
	+ T1 [b]	6896.5	6894.5	–0.03
8QA103	– T1	11795.5	11794.6	–0.008
	+ T1 [b]	6896.5	6896.4	–0.001
8QA107	– T1	11795.5	11793.7	–0.02
	+ T1 [b]	6896.5	6894.4	–0.03
8QA111	– T1	11795.5	11793.2	–0.02
	+ T1 [b]	6896.5	6895.3	–0.02
8QA115	– T1	11795.5	11793.6	–0.02
	+ T1 [b]	6896.5	6892.0	–0.07
8QA122	– T1	11795.5	11794.4	–0.009
	+ T1 [b]	6896.5	6895.8	–0.01

Table 3.1. MALDI mass spectrometry analyses of the reaction products before and after RNase T1 digestion. RNase T1 initially creates a 2′,3′-cyclic phosphate terminus that can be hydrolyzed to 3′-monophosphate, the ratio of these two products depends on the incubation details.[12-14] Both products were typically observed in our assays. In the table, data is shown for both products (for 8QA124) and for the 3′-monophosphate, which was the more abundant of the two products (for the other deoxyribozymes).
[a] 2′,3′-cyclic phosphate product.
[b] 3′-monophosphate product.

3.2.8 Characterization of the reductive amination product by MS/MS experiment

For additionally support of our assigned structure, we acquired MS/MS data of a suitably digested reductive amination product by MALDI-TOF mass spectrometry. First, the reductive amination product was treated with RQ1 RNase-free DNase, which digested only DNA but not RNA, followed by phenol-chloroform extraction and PAGE purification. Three bands were observed. One of three prominent bands was excised, extracted, and mass spectrometry was acquired. The mass of the excised band was consistent with the sequence 5′-ATa-ggaagagauggcgacgg-3′. The RQ1 DNase-treated product was further digested with RNase T1. The product was desalted and mass analysis revealed that the product was 5′-ATa-gg-3′. MS/MS experiment was performed with the

pentamer product. The diagnostics fragment of the G nucleobase with the G1 ribose as well as others fragments that support the assignment were observed (**Fig. 3.12**).

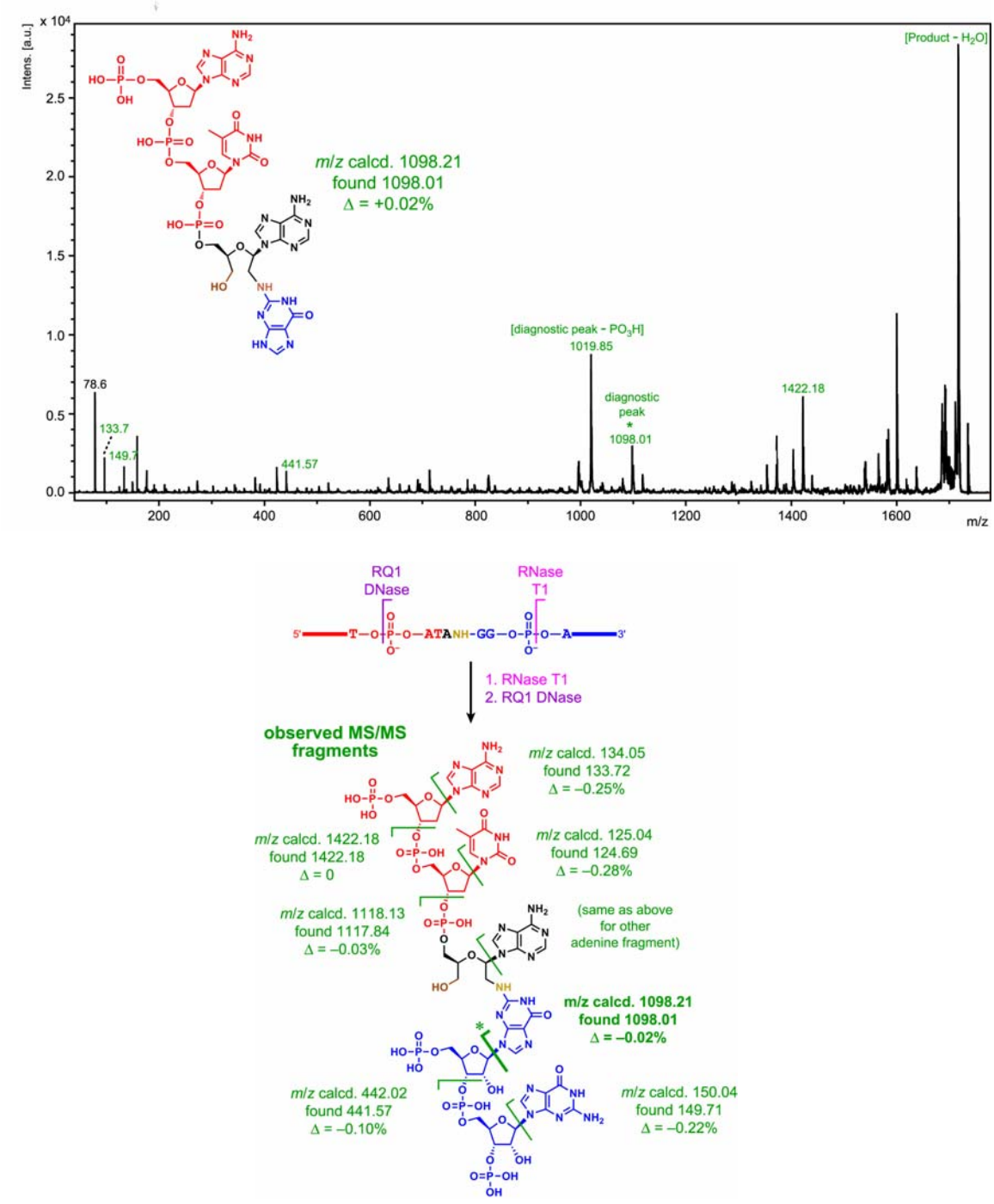


Figure 3.12 MS/MS experiment on a suitably digested fragment of the reductive amination product of the 8QA124 deoxyribozyme. The diagnostic fragment, highlighted with an asterisk, corresponds to the cleavage between the G1 ribose and the G1 nucleobase. Other fragments were also observed that support the structural assignment.

3.2.9 The reduction of the second aldehyde group during selection

Our assigned chemical structure of the reductive amination product depicts that the second aldehyde group of the capture oligonucleotide was reduced to an alcohol. The alcohol assignment is based on three independent experiments, as follows. First, the DNA-catalyzed reductive amination product was not observed to further oxidized to the corresponding carboxylic acid by sodium chlorite, $NaClO_2$. The reaction between an aldehyde and $NaClO_2$ is known as Pinnick or Lindgren oxidation. In the corresponding positive control experiment, the capture oligonucleotide was treated with $NaClO_2$, and oxidation was observed by MALDI mass spectrometry analysis. Second, the DNA-catalyzed reductive amination product was not observed to further react via reductive amination with several versions of oligonucleotides terminating with an amine. Whereas, the corresponding positive control which involved the capture oligonucleotide readily reacted with several versions of oligonucleotides terminating with an amine to yield the reductive amination adduct. In addition, the DNA-catalyzed reductive amination product was treated with AAASAA hexapeptide with a free N-terminus, and no reductive amination adduct was observed. On the contrary, the dialdehyde-containing oligonucleotide reacted with the hexapeptide via reductive amination reaction. Third, the DNA-catalyzed reductive amination product was treated with biotin-hydrazide, but no adduct was observed. In the corresponding positive control, the capture oligonucleotide reacted with the biotin-hydrazide readily and adducts were observed. Together, all these data establish that the DNA-catalyzed product does not have an aldehyde group. It was determined in our assignment that the second aldehyde is an alcohol.

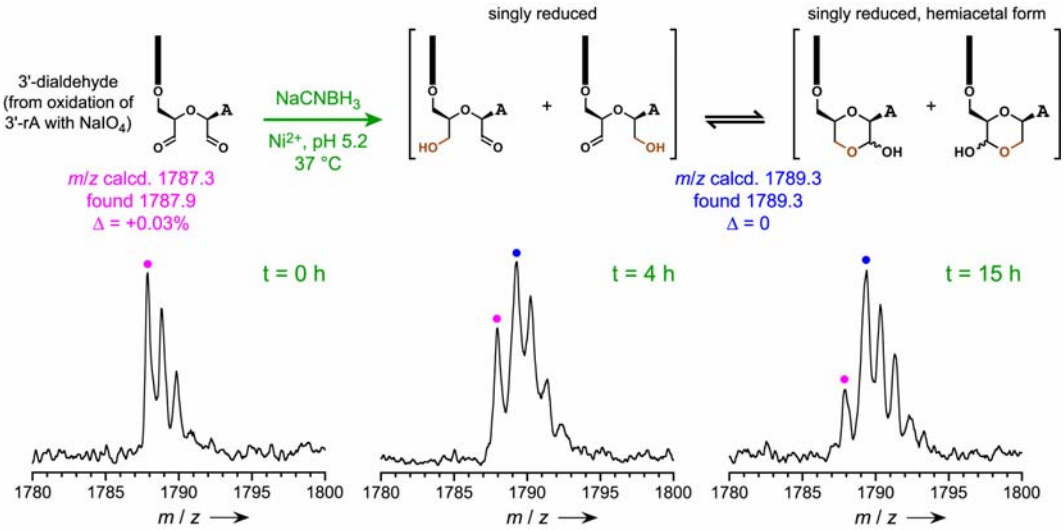


Figure 3.13 Direct observation of the reduction of the aldehyde functional groups by $NaCNBH_3$ in the capture step incubation condition. The isotopic distributions partially obscure the peaks but nonetheless a clear pattern can be observed. The pink dot denotes the mass of the oligonucleotide with dialdehyde functional groups and the blue dot indicates the mass of the oligonucleotide with the singly reduced alcohol

A reducing equivalent must be present to reduce the second aldehyde of the capture oligonucleotide to an alcohol. It was reasoned that the $NaCNBH_3$ could be the reducing agent. Under neutral conditions, $NaCNBH_3$ would not readily reduce an aldehyde. However, in acidic conditions, $NaCNBH_3$ can be used as reducing agent for aldehydes or ketones. In fact, $NaCNBH_3$ is synthetically employed as a reducing agent for aldehydes and ketones at pH ~3–4.[15-16] For example, it was reported that cyclohexanone was reduced to cyclohexanol by $NaCNBH_3$ with 88% yield at pH 4 for 1 h.[17] Moreover, there were examples[7] of aldehyde reduction at pH as high as 5.5. Therefore, we hypothesized that the unreacted aldehyde was reduced by $NaCNBH_3$ during the capture step at pH 5.2. To show that this scenario is plausible, we validated the hypothesis by incubating a hexamer oligonucleotide terminating with aldehyde functional groups. To prepare the aldehyde-terminating hexamer oligonucleotide, a hexamer oligonucleotide ending with a 3′ ribonucleotide A was first treated with $NaIO_4$. The hexamer oligonucleotide

terminating with aldehyde functional groups were used instead of the 20 nt long capture oligonucleotide because the small hexamer enabled a reliable assignment by MALDI-TOF mass analysis. Indeed, we observed the reduction of the aldehyde of the hexamer in the incubation condition (**Fig. 3.13**).

3.2.10 8QA deoxyribozymes use alternative borane reducing agents

We examined the effect of using alternative reducing agents on the DNA-catalyzed reductive amination reaction. We replaced the initially used reducing agent $NaCNBH_3$ with dimethylamine-borane or 2-picoline-borane. In each case of the 8QA deoxyribozymes-catalyzed reductive amination, the replacement of other borane reducing agents was tolerated (**Fig. 3.14**). The results show that 8QA deoxyribozymes catalyzed poorly or are not involve in the imine reduction. Therefore, the 8QA deoxyribozymes are likely to catalyze the first amine nucleophilic attack of carbonyl.

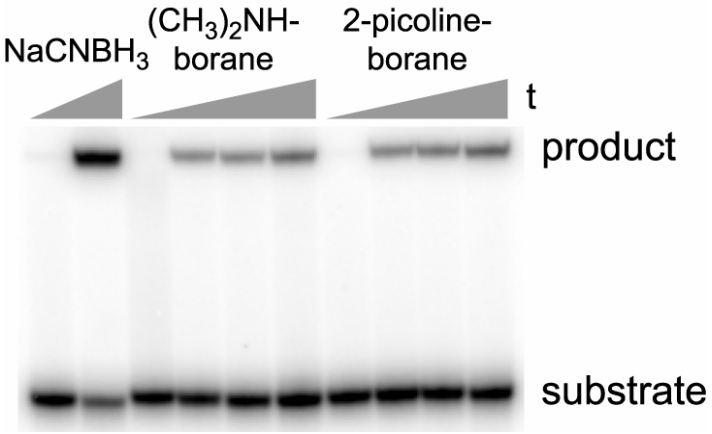


Figure 3.14 Sodium cyanoborohydride reducing agent can be substitued by other borane reducing agent. Two borane reducing agents were assayed with all 8QA deoxyribozymes. In all cases, the 8QA deoxyribozymes tolerate the reducing agent replacement relatively well. Shown here are the representative assays performed with 8QA124 to illustrate that the replacement of reducing agent is tolerated. The key point of this experiment was not to compare the absolute yield of the reductive amination product but to note the substantial product formation when alternative reducing agents were used.

3.2.11 Dependence of the catalytic activities on divalent metal cofactors

We examined the effect of using various divalent metal ions on the DNA-catalyzed reductive amination reaction. We substituted the Ni^{2+} with other divalent metal ions Mg^{2+}, Ca^{2+} or Mn^{2+}. In each cases of 8QA deoxyribozyme-catalyzed reductive amination, the reaction was suppressed when Ni^{2+} was replaced with other divalent metal ions (**Fig. 3.15**). The results were not unexpected as the 8QA deoxyribozymes were identified under a selection condition that requires Ni^{2+}. Therefore, the 8QA deoxyribozymes requires Ni^{2} to be catalytic active. Also, Ni^{2+} may serve as a Lewis acid to activate the carbonyl group of the aldehyde for the subsequent catalyzed attack of the amine of the G1. However, the mechanism on how Ni^{2+} assists the DNA-catalyzed reductive amination requires extensive studies that involve obtaining the high resolution structural information of the deoxyribozyme, which is beyond the scope of this thesis work.

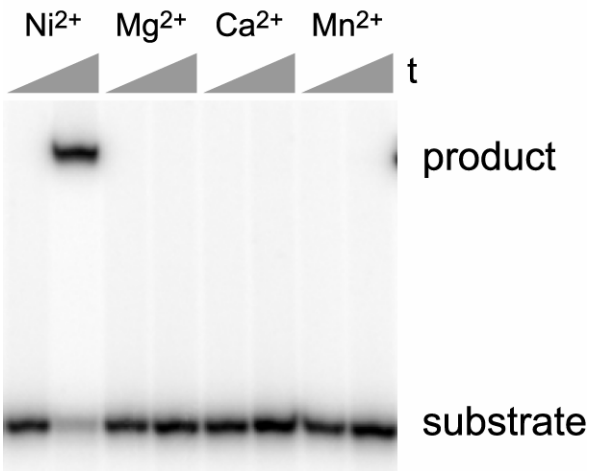


Figure 3.15 Ni^{2+} cannot be replaced by other divalent metal cations. The illustrative PAGE of 8QA124 assayed with different divalent metal ions are shown here. Other 8QA deoxyribozymes also show the same results. These experiments establish that Ni^{2+} is mandatory for the proper folding of the deoxyribozymes and also may serve as a Lewis acid to activate the carbonyl functional group of the aldehyde.

3.3 Summary

In this chapter, we described the selection effort to identify free CYA reactivity. We unexpectedly found deoxyribozymes that catalyze reductive amination reaction using a weak nucleophilic N^2-amine of a guanine of the RNA substrate as the nucleophile, instead of the originally intended N-terminus amino nucleophile. One of the deoxyribozymes, 8QA124, catalyzes the reductive amination involving the N^2-amine of a guanine with the RNA substrate with yield ranging from 60–90% and k_{obs} = 0.37 h^{-1} ($t_{1/2}$ 1.7 h). This unprecedented discovery expands the scope of reactions catalyzed by DNA and is currently being developed as a tool for site-specific labeling nucleic acids.

3.4 Materials and Methods

3.4.1 Solid phase peptide synthesis

CYA tripeptide was prepared by standard methods using Fmoc Rink mide MBHA resin. The solid phase peptide synthesis (SSPS) procedures can be found in ref. 18. The key procedures are described below.

Materials. Anhydrous DCM and DMF solvents (Fisher) were used. The Fmoc-protected amino acids, HATU coupling reagent, and the Rink amide MBHA resin were purchased from either Novabiochem or Chem-Impex International. A coarse Pyrex fritted funnel (40–60 μm pore size) was used for efficient agitation with nitrogen. Deprotection solvent contained 20% piperidine and 80% DMF, and it was prepared freshly prior its use. SSPS couples Fmoc amino acid from the C-terminus to the N-terminus. To prepare CYA tripeptide, the alanine was first coupled, followed by the tyrosine and cysteine. Fmoc-Ala-OH (CAS no. 35661-39-3), Fmoc-Tyr(tbu)-OH(CAS no. 71989-38-3), and

Fmoc-Cys(tbuthio)-OH (CAS no. 73724-43-3) were used to prepare the CYA tripeptide. Beacause SSPS required iterative rinse cycle and deprotection cycle. Therefore, the procedure of the rinse cycle and deprotection cycle were described first as follows.

The rinse cycle consisted of following steps:

1. 10 mL of DMF was added to the resins and agitated with nitrogen gas.
2. The solvent was removed by filtration.
3. The resins were resuspended in 10 mL DMF from a squirt bottle. DMF was removed by filtration.
4. Step 1–3 were repeated for a total four times.

The deprotection procedure consisted of the following steps:

1. 10 mL of deprotection solvent was added to the resins and agitated with nitrogen gas for 2 min.
2. The solvent was removed by filtration.
3. A rinse cycle was performed.
4. Step 1–3 were repeated for a total three times.

The Kaiser test was performed, as follows.

Three solutions were prepared. (Solution A: 5 g ninhydrin in 100 mL of ethanol, Solution B: 80 g of phenol in 20 mL of ethanol, Solution C: 2 mL of 1 mM KCN in 98 mL pyridine)

1. ~10 beads were taken from the fritted funnel and placed in a 10 × 75 mm test tube. The resins were washed by repeatedly addition and removal of ~1 mL methanol with a glass pipette.
2. Three drops of each solution (A, B and C) were added to the test tube.
3. The test tube was heated using a heat gun. The tube was shaken gently to avoid overheating.

4. After 0.5 min of heating, a positive test resulted in resins with purple color, which indicates the presence of primary amine functional groups on the resins. A negative result would result in a pale yellow to orange color, which indicates there were no primary amine function groups on the resins.

Swell resins. 0.54 mg (0.2 mmol, 0.37 mmol/g) of resin was measured, followed by swelling of the resin in a 10 mL of DMF/DCM (v/v; 1:1) for 1 h (3 × bed volume). The resins were filtered and washed thoroughly with DMF.

Resins deprotection. The Fmoc-protecting group of the Rink amide resin was removed by the deprotection step, followed by a rinse cycle. The Kaiser test was performed and a positive result was observed, indicating the successful removal of the Fmoc-protecting group by the base treatment.

Coupling of Fmoc amino acid. Fmoc protected amino acid (1 mmol, 5 eq.) was first activated with HATU peptide coupling reagent (0.98 mmol, 4.9 eq.) and diisopropylethylamine (DIPEA, 2 mmol, 10 eq.) for 10 min. The mixture was added to the resin, agitated with nitrogen for 1 h. After 1 h, a rinse cycle followed by a Kaiser test was performed. A successful coupling would yield a negative result with the Kaiser test.

Fmoc deprotection. The N-terminus Fmoc protected group was removed by the deprotection step, followed by a rinse cycle. A Kaiser test was performed and purple resins were observed. The second amino acid coupling was performed after the deprotection step. The process was iterated until all of the amino acids were coupled.

Peptide removal from the solid support by TFA. Resin was transferred to a 20 mL disposable scintillation vial containing a small stir bar. Scavenger cocktail (125 μL of water, 125 μL of ethanedithiol, and 50 μL of triisopropylsilane) was added to scavenger

the t-butyl cations. 5 mL of trifluoroacetic acid (TFA) was added to the vial. The vial was capped and the mixture was stirred at room temperature for 3 h. The cleavage mixture was filtered through a clean coarse fritted filter. The clear pale yellow cleavage solution was transferred to a round-bottom flask and the excess TFA was evaporated by a rotary evaporator. 20 mL of cold diethyl ether was added . White precipitates were observed. The ether layer was carefully removed by a glass pipette. The ether wash procedures were repeated. The residual ether in the round-bottom flask was removed by a rotary evaporator, followed by drying under high vacuum for several hours. The peptide was purified by HPLC. For the CYA tripeptide, ESI $[M+H]^+$ calcd: 443.6, found: 443.4.

3.4.2 Preparation of the CYA-RNA-DNA for capture step optimization

Preparation of CYA-RNA. To test the efficacy of the capture step, a sample containing free N-terminus CYA ligated to the RNA and the pool was generated. To prepare a CYA-RNA product on the preparative scale, four reactions were performed using the following procedure. For each reaction, 1 nmol of 5′-triphosphate RNA, 1 nmol of the DNA helper oligonucleotide, and 700 pmol of 15MZ36 were annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 1 min and cooling on ice for 5 min. 2 μL of CYA (from 50 mM stock solution in DMF) were added, and the ligation reaction was initiated by addition of stock solutions to a final volume of 100 μL containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$. The reaction was incubated at 37 °C for 48 h, precipitated with ethanol, and purified by 20% PAGE.

Ligation of CYA-RNA to 5′ ^{32}P- radiolabeled DNA pool. 10 pmol of CYA-RNA was ligated to 1 pmol 5′-^{32}P-radiolabeled DNA pool using 1 U of T4 RNA ligase (Fermentas)

in 20 μL of 50 mM Tris, pH 7.5, 5 mM $MgCl_2$, and 0.05 mM ATP at 37 °C for 12 h, followed by 8% PAGE. In this ligation reaction, DTT was omitted to avoid the reduction of the *S*-tbutyl group of the cysteine.

3.4.3 In vitro selection procedure

In the key selection step during each round of selection, the PCR-amplified deoxyribozyme pool (estimated 10 pmol) and 50 pmol of helper oligonucleotide (complementary to the 3′-binding arm of the deoxyribozyme) were annealed in 14 μL of 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 1 min and cooling on ice for 5 min. The sample was brought to final volume of 20 μL containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, 20 mM $MnCl_2$, and 1 mM CYA tripeptide (added from 50 mM stock solution in DMF) and incubated at 37 °C for 14 h, followed by ethanol precipitation. For the subsequent capture step, 200 pmol of capture oligonucleotide was added, and the sample was annealed in 10 μL of 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 1 min and cooling on ice for 5 min. The sample was then brought to final volume of 20 μL containing 100 mM NaOAc, pH 5.2, 50 mM $NiCl_2$, and 10 mM $NaCNBH_3$. The $NaCNBH_3$ was added from a 200 mM aqueous stock solution. The reaction was incubated at 37 °C for 14 h, and separated by 8% PAGE.

3.4.4 Deoxyribozyme activity assays

The 3′-dialdehyde-terminated capture oligonucleotide was prepared from 5′-GGATAATACGACTCACTAT(rA)-3′ by oxidation of 20–5000 pmol of oligonucleotide in 20–50 μL of 100 mM HEPES, pH 7.5, and 10 mM $NaIO_4$ at room temperature for 1 h followed by ethanol precipitation.

The general assay procedure for each deoxyribozyme was as follows. A 10 μL sample containing 1 pmol of 3′-^{32}P-radiolabeled RNA substrate, 10 pmol of deoxyribozyme, and 20 pmol of capture oligonucleotide was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 1 min and cooling on ice for 5 min. The reductive amination reaction was initiated by bringing the sample to 20 μL total volume containing 100 mM NaOAc, pH 5.2, 50 mM $NiCl_2$, and 10 mM $NaCNBH_3$ and incubating at 37 °C. The $NaCNBH_3$ was added from a 200 mM aqueous stock solution. At appropriate times, 2 μL aliquots were quenched with 6 μL of stop solution (80% formamide, 1 × TBE [89 mM each Tris and boric acid, 2 mM EDTA, pH 8.3], 50 mM EDTA, and 0.025% each bromophenol blue and xylene cyanol). Before PAGE, to each sample was added 100 pmol of a decoy oligonucleotide, which was a 60-mer complementary to the 40 nt catalytic region of the deoxyribozyme (to displace deoxyribozyme from substrate and product). Samples were separated by 20% PAGE and quantified with a PhosphorImager. k_{obs} values were obtained by fitting the yield versus time data directly to first-order kinetics.

3.4.5 DNA-catalyzed reductive amination products preparation

To generate the reductive amination product for MS analysis, reactions were performed on a preparative scale. First, a 5′-hydroxyl RNA was phosphorylated with T4 PNK, precipitated by ethanol, and desalted with a 3000 MWCO Amicon column. The extra desalting step using an Amicon column was found to be essential for the preparative scale reaction. Significant precipitation was observed in the reductive amination reaction when the Amicon desalting step was omitted.

To prepare the reductive amination product on a preparative scale, two reactions were performed for each individual deoxyribozyme. In each reaction, 1 nmol of 5′-phosphorylated RNA, and 1.2 nmol of capture oligonucleotide, and 800 pmol of deoxyribozyme were annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA to a 30 μL sample by heating at 95 °C for 1 min and cooling on ice for 5 min. The reductive amination reaction was initiated by bringing the sample to 40 μL total volume containing 100 mM NaOAc, pH 5.2, 50 mM $NiCl_2$, and 10 mM $NaCNBH_3$ and incubating at 37 °C for overnight. The $NaCNBH_3$ was added from a 200 mM aqueous stock solution. The reactions were desalted by ethanol precipitation and purified by 20% PAGE.

3.4.6 DNase and RNAseT1 treatments

A 900 pmol portion of the 8QA124 reductive amination product was treated with 10 U of RQ1 RNase-free DNase (Promega) in 40 mM Tris, pH 8.0, 10 mM $MgSO_4$, and 1 mM $CaCl_2$ in 100 μL total volume at 37 °C for 14 h. The sample was extracted with phenol/chloroform, precipitated with ethanol, and separated by 20% PAGE. Three predominant bands were observed; the middle of these three was excised from the gel, extracted with TEN, and precipitated with ethanol. This DNase-digested product was treated with 20 U of RNase T1 (Ambion) in 50 mM Tris, pH 7.5, and 2 mM EDTA in 40 μL total volume at 37 °C for 14 h. The sample was concentrated to 10 μL by SpeedVac, desalted by C_{18} ZipTip, and analyzed by MALDI mass spectrometry, revealing its identity as the pentamer 5′-pATrA–GGp-3′, where the new reductive amination linkage is between the rA and the G (calcd. 1735.2, found 1733.6, Δ = –0.09%). MS/MS analysis was performed using an UltrafleXtreme TOF/TOF mass spectrometer (Bruker).

3.5 References

1 Ohshima, T.; Soda, K., Biochemistry and biotechnology of amino acid dehydrogenases. *Adv. Biochem. Eng. Biotechnol.* **1990**, *42*, 187–209.

2 Braunstein, A.; Kritzmann, M., Biologically inspired synthetic enzymes made from DNA. *Nature* **1937**, *140*, 503–504.

3 Eliot, A. C.; Kirsch, J. F., Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annu. Rev. Biochem.* **2004**, *73*, 383–415.

4 Stewart, J. D., Dehydrogenases and transaminases in asymmetric synthesis. *Curr. Opin. Chem. Biol.* **2001**, *5*, 120-129.

5 Faber, K.; Kroutil, W., New enzymes for biotransformations. *Curr. Opin. Chem. Biol.* **2005**, *9*, 181–187.

6 Höhne, M.; Bornscheuer, U. T., Biocatalytic routes to optically active amines. *ChemCatChem* **2009**, *1*, 42–51.

7 Ward, J.; Wohlgemuth, R., High-yield biocatalytic amination reactions in organic synthesis. *Curr. Org. Chem.* **2010**, *14*, 1914–1927.

8 De Wildeman, S. M.; Sonke, T.; Schoemaker, H. E.; May, O., Biocatalytic reductions: from lab curiosity to "first choice". *Acc. Chem. Res.* **2007**, *40*, 1260-1266.

9 Zhu, D.; Hua, L., Biocatalytic asymmetric amination of carbonyl functional groups – a synthetic biology approach to organic chemistry. *Biotechnol. J.* **2009**, *4*, 1420–1431.

10 Hud, N. V.; Jain, S. S.; Li, X.; Lynn, D. G., Addressing the problems of base pairing and strand cyclization in template-directed synthesis. *Chem. Biodivers.* **2007**, *4*, 768–783.

11 Kumar, A.; Sharma, S.; Maurya, R. A., Single nucleotide-catalyzed biomimetic reductive amination. *Adv. Synth. Catal.* **2010**, *352*, 2227–2232.

12 Uchida, T.; Egami; F., (1971) *The Enzymes*, New York: Academic Press, 205–250.

13 Herschlag, D.; Piccirilli, J. A.; Cech, T. R., Ribozyme-catalyzed and nonenzymic reactions of phosphate diesters: rate effects upon substitution of sulfur for a nonbridging phosphoryl oxygen atom. *Biochemistry* **1991**, *30*, 4844–4854.

14 Kirpekar, F; Douthwaite, S.; Roepstorff, P., Mapping posttranscriptional modifications in 5S ribosomal RNA by MALDI mass spectrometry. *RNA* **2000**, *6*, 296–306.

15 Borch, R. F.; Bernstein, M. D.; Durst, H. D., Cyanohydridoborate anion as a selective reducing agent. *J. Am. Chem. Soc.* **1971**, *93*, 2897–2904.

16 Lane, C. F., Sodium cyanoborohydride - a highly selective reducing agent for organic functional groups. *Synthesis* **1975**, 135-146.

17 Domagala, J.; Wemple, J., Biomimetic formation of tropic acid esters. *Tetrahedron Lett.* **1973**, *14*, 1179-1182.

18 Chan, W. C.; White, P. D., (2000) Fmoc Solid Phase Peptide Synthesis: A Practical Approach, Oxford University Press, Oxford.

Chapter 4: Optimization of Click Chemistry Incubation Conditions for Direct Free Peptide Selection

4.1 Introduction

Chapter 2 and 3 describe our various selection efforts to achieve DNA-catalyzed covalent modification of a short peptide. In chapter 3, the selection effort with an additional reductive amination step was originally intended to select for free peptide activity; however, the selection effort gave rise to an unexpected DNA-catalyzed reductive amination that involves the exocyclic amino group of G1.[1] Therefore, in this chapter, an alternative strategy was sought to avoid undesired side reactions in the capture step. The bioorthogonal Cu(I)-catalyzed Azide-Alkyne Cycloaddition[2-3] (CuAAC), the most commonly known example of click chemistry, has been widely developed for use in organic syntheses and bioconjugation reactions.[3-8] CuAAC provides an excellent option for the capture step as CuAAC is highly specific to the alkyne and the azide functional groups. In addition, templated click reactions that involved oligonucleotide substrates have been reported.[9-10] In this chapter, the optimization of CuAAC for direct selection using an azido-modified peptide is reported.

4.2 Results and Discussion

4.2.1 In vitro selection design

We designed a two-stage strategy to seek DNA-catalyzed free peptide reactivity, which is shown in **Fig. 4.1**. During the selection step, deoxyribozymes append an azido bearing-peptide to themselves via reaction between the nucleophilic amino acid side chain with the 5′-triphosphate of the RNA. In the capture step, a capture oligonucleotide

that terminates with a 3′-alkyne reacts with the azido-bearing selection product via CuAAC reaction, forming a stable triazole linkage joining the capture oligonucleotide to the selection product. The capture step leads to a substantial gel-shift of the selection product which allows the separation of the selection product by PAGE.

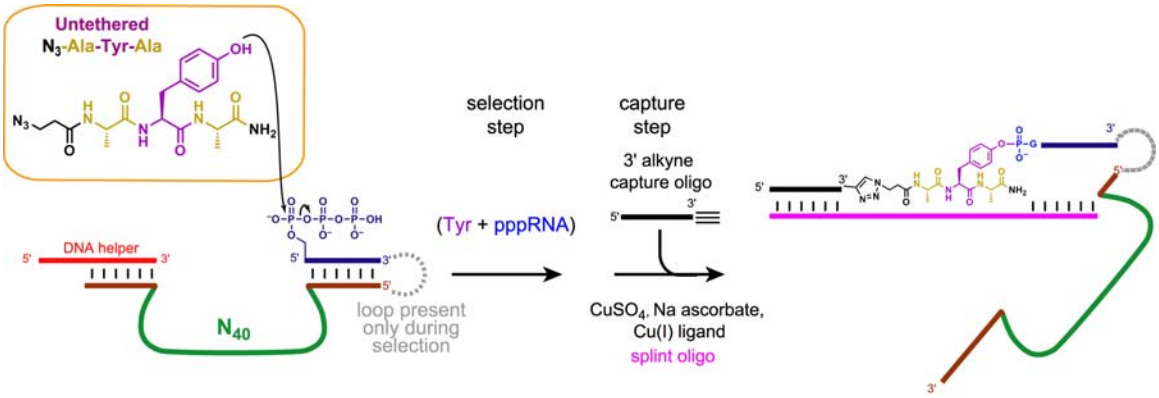


Figure 4.1 A two-stage design aims to select for DNA-catalyzed free peptide reactivity. The selection step allows deoxyribozymes append themselves to the peptide in which the N-terminal is tethered to an azido propyl group via an amide linkage. The capture step employs CuAAC reaction between the azido-bearing peptide and a 3′-alkyne capture oligonucleotide, resulting in a gel-shift product which can be readily separated by PAGE.

4.2.2 Click chemistry capture step optimization

CuAAC capture step optimization in a trimolecular format. To ensure an efficient capture step, we sought the incubation conditions for the CuAAC reaction. The Cu(I) catalyst was prepared in situ from Cu(II) sulfate using sodium ascorbate as the reducing agent. In CuAAC reaction, severe degradation of oligonucleotide was observed.[10] Therefore, a minimum 5-fold excess of an aqueous Cu(I) stabilizing ligand tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) to the $CuSO_4$ should be employed.[10] We examined the optimal ratio of $CuSO_4$: THPTA: sodium ascorbate as well as the optimal concentration for each key reactant in the capture step. It was previously reported[10] that a final concentrations of 0.2 mM $CuSO_4$, 1.38 mM THPTA, and 2 mM sodium ascorbate (1:7:10 CuSO4:THPTA:sodium ascorbate) were used in a CuAAC reaction between a 3′-

azido modified oligonucleotide and an oligonucleotide with an alkyne at its 5′-end in a splinted fashion. To optimize the capture step, we first covalently ligated azido modified AYA to the RNA substrate via 15MZ36-catalyzed ligation between the Tyr of the azido modified AYA and the 5′-triphosphate RNA (see Materials and Methods). We performed CuAAC optimization in a trimolecular format, in which the azido modified AYA-RNA substrate was not covalently linked to the DNA pool. From the initial assay using a final concentrations of 0.4 mM $CuSO_4$, 2.8 mM THPTA, and 4.0 mM sodium ascorbate in the CuAAC reaction, 73% of product was formed in 30 min (**Fig. 4.2**, condition C).

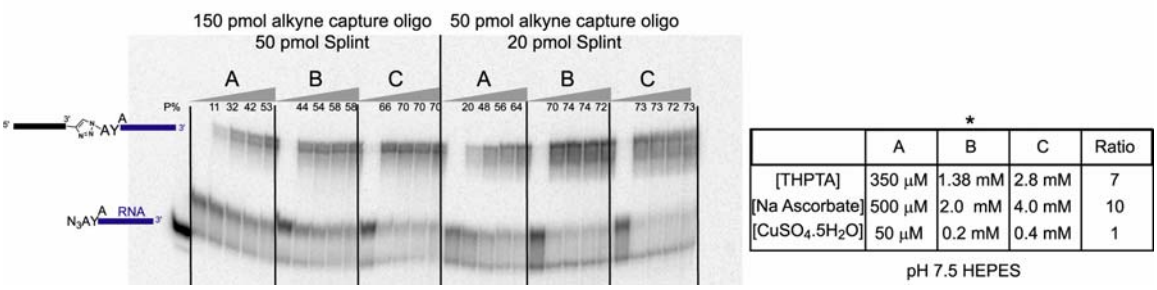


Figure 4.2 Assays of azido modified AYA-RNA with the alkyne capture oligonucleotide and various concentrations of THPTA, sodium ascorbate and $CuSO_4$. Condition B, highlighted by an asterisk, was a particular reaction condition that has been reported.[4b] The 3′-^{32}P-radiolabeled azido modified AYA-RNA, the alkyne capture oligonucleotide, and the splint oligonucleotide were incubated in 50 mM HEPES, pH 7.5, 150 mM NaCl, followed by sequential addition of THPTA, sodium ascorbate, and $CuSO_4$. Each of the latter three was prepared as 10× stock solutions. The reaction was incubated at 37°C. Time points were taken at 30 s, 30 min, 2 h, 4 h, 20 h.

Having established a reliable capture step in a trimolecular format, we examined the CuAAC chemistry in cis, in which the DNA pool was covalently ligated to the azido modified-AYA-RNA substrate (see Materials and Methods for preparation). In order to mimic the selection conditions, we first incubated the internally radiolabeled azido modified-AYA-RNA-DNA pool substrate with 60 pmol of RNA-DNA pool in pH 7.5 HEPES with 20 mM Mn^{2+} and 40 mM Mg^{2+}, followed by ethanol precipitation. [Note: in a regular selection round, we estimated an upper limit of 20 pmol RNA-DNA pool was

present. 60 pmol of RNA-DNA pool was assayed in this optimization step as 200 pmol of RNA-DNA pool was used in the selection initiation round. We reasoned that a 60 pmol of the RNA-DNA pool (3-fold less than the initial round, 3-fold in excess of a regular round of selection) should be a good choice for the optimization.] After ethanol precipitation, the azido modified-AYA-RNA-DNA pool substrate was incubated in the optimized CuAAC condition which was described previously in the trimolecular format assay. To ensure the observed gel-shifted product was due to the Cu(I) catalyzed reaction, a negative control was performed by omitting $CuSO_4$. Indeed, complete suppression of product formation was observed when CuSO4 was omitted. Also, THPTA was found to be necessary to minimize oligonucleotide degradation. We have established a capture step using CuAAC, in which ~50% of product was observed in 2 h (**Fig. 4.3**).

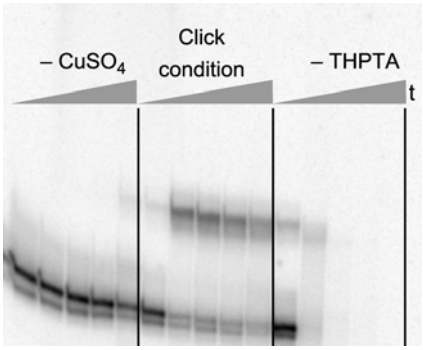


Figure 4.3 Assays of azido modified AYA-RNA-DNA pool with the alkyne capture oligonucleotide in selection conditions. ^{32}P-radiolabeled *Azido modified* AYA-RNA-DNA pool, 60 pmol RNA-DNA pool, 150 pmol of alkyne capture oligonucleotide, and 50 pmol of splint were incubated in 50 mM HEPES, pH 7.5, 150 mM NaCl, followed by sequential addition of THPTA, sodium ascorbate, and $CuSO_4$ to a final volume of 20 µL. Each of the latter three was prepared as 10× stock solutions. A final concentrations of 0.4 mM $CuSO_4$, 2.8 mM THPTA, and 4.0 mM sodium ascorbate were used in the reaction. The reaction was incubated at 37°C. Time points were taken at 30 s, 30 min, 1 h, 2 h, and 18 h.

4.3 Summary

In this chapter, a new two-stage in vitro selection method to identify deoxyribozymes that catalyze peptide covalent modifications was developed. The novel in vitro selection

method involved the key selection step, followed by a capture step that employs a bioorthogonal CuAAC reaction. The newly developed capture step should be amenable to large peptides that are modified with an azide handle.

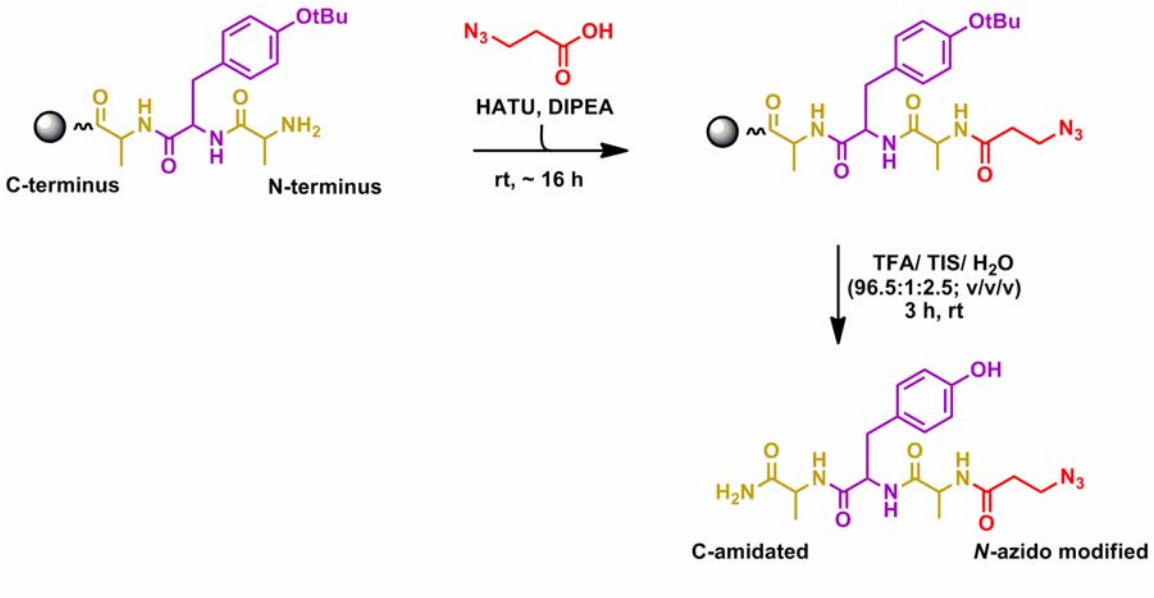
4.4 Materials and Methods

4.4.1 3′-Alkyne-modified capture oligonucleotide, splint, and helper oligonucleotides

The 3′-alkyne-modified DNA was prepared by using the 3′-alkyne serinol CPG (Glen Research) using an ABI 394 DNA solid-phase synthesizer. RNA oligonucleotides were prepared by in vitro transcription using T7 RNA polymerase.

3′-Alkyne:	5′-GCCGCATCAGACAGCG-3′Alkyne
Splint:	5′-TTCGCCGTCGCCATCTCTTCCCGCTGTCTGATGCGGC-3′
RNA:	5′-ggaagagauggcgacgg-3′
Sel. Helper:	5′-AACAACAACAACAACAACAACGGATAATACGACTCACTAT-3′
15MZ36 Helper:	5′-GGATAATACGACTCACTAT-3′

4.4.2 Azido modified peptide preparation



Scheme 4.1 Solid phase preparation of azido modified AYA. The last coupling and the deprotection step are illustrated**.**

The azido modified AYA peptide was prepared by 0.2 mmol-scale Fmoc solid-phase peptide synthesis with Rink amide resin. Standard procedures (as described in chapter 2) were performed. For the last coupling (**Scheme 4.1**), 3-azidopropanoic acid (1 mmol, 5 eq.) was activated by HATU coupling reagent (0.98 mmol, 4.9 eq) and DIPEA (2 mmol, 10 eq) in 5 mL of DMF for 30 min before adding to the resin. The reaction was stirred under nitrogen at room temperature overnight. A Kaiser test was performed. The peptide was cleaved off the resins by TFA, triturated with ether, dried, and purified by HPLC. ESI-Mass calcd $[M+H]^+$: 420.44; found:420.40.

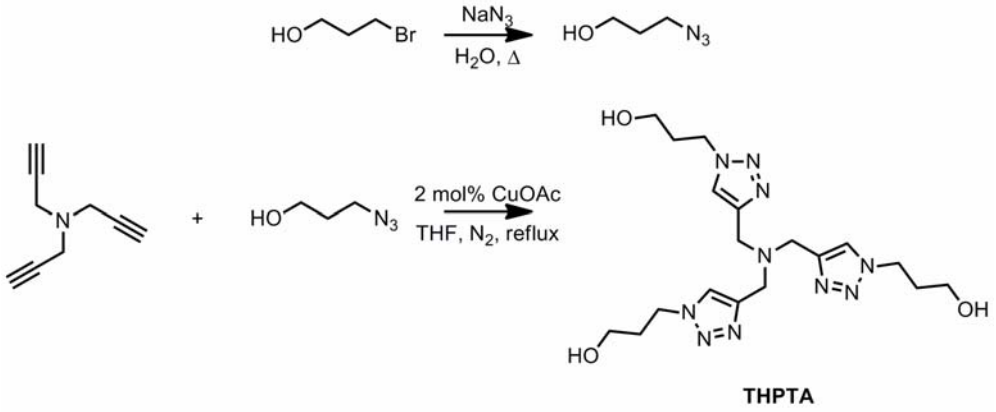
Preparation of 3-azidopropanoic acid.

Scheme 4.2 3-azidopropanoic acid preparation.

Caution: Workup of reaction involving inorganic azide should avoid acid, because HN_3 is volatile, explosive and toxic. Do not use halogenated solvents for reactions that involve inorganic azides. Halogenated solvent such as methylene chloride and chloroform could result in the formation of potentially explosive diazidomethane.[11]

The 3-azidopropanoic acid was prepared according to the published procedure (**Scheme 4.2**).[12] 3-Bromoproanoic acid was dried under vacuum for 20 min before use. Into a 50 mL round-bottom flask, 2.7 g of 3-bromopropanoic acid (17.7 mmol, 1 eq.) was dissolved in 10 mL of anhydrous DMF. The solution was stirred at room temperature under argon. Sodium azide (3.48 g, 53.5 mmol, 3 eq.) was slowly added to the solution. The sodium azide was not soluble in DMF. The mixture was heated to 60 °C in an oil bath with a condenser for 48 h. A ESI mass spectrometry was taken with the crude compound and the desired product mass was observed. The reaction was stopped by diluting the reaction with 150 mL of acetonitrile, and filtering by suction filtration to remove excess sodium azide. The solution was concentrated by a rotary evaporator at room temperature behind a blast shield. Low res. ESI $[M–H]^-$ cacld: 114.09, found: 114.20.

4.4.3 Aqueous Cu(I) ligand tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) preparation



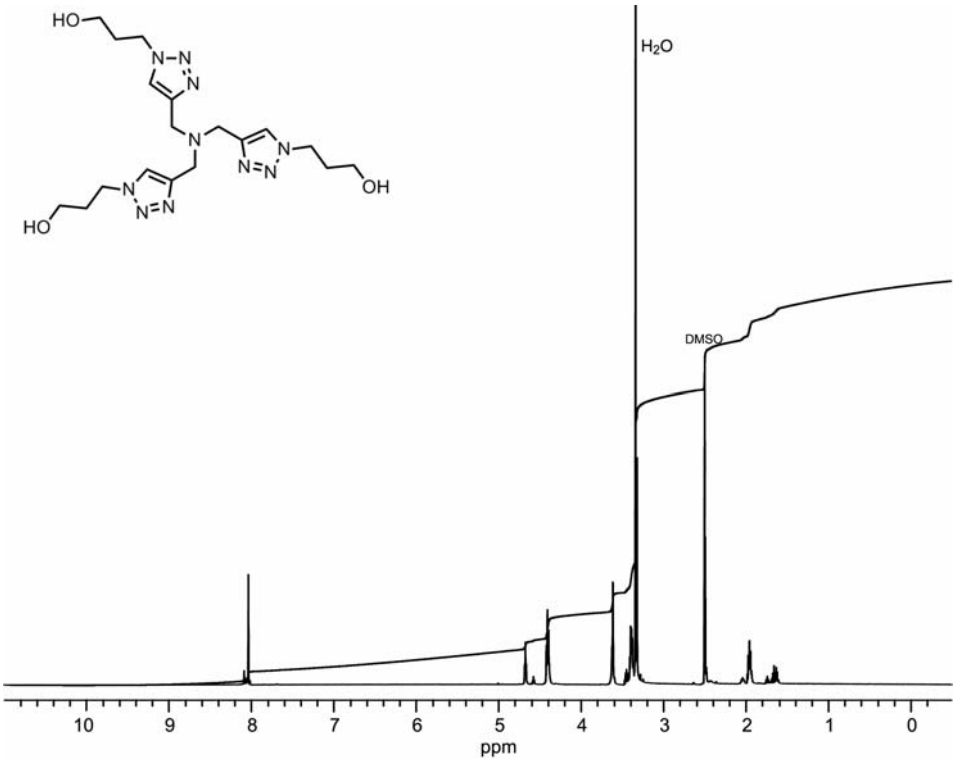
Scheme 4.3 Aqueous Cu(I) ligand tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) preparation.

The aqueous Cu(I) THPTA ligand was prepared according to published procedures in the supportive information in ref (8) (**Scheme 4.3**). Note that the starting material, cuprous acetate, should not be green in color. The green color indicates a significant oxidation to Cu(II) has occurred.

3-azido-1-propanol preparation. 4 g of 3-bromo-1-propanol (28.8 mmol, 1 eq.) was dissolved in 30 mL of water in a round-botton flask at room temperature. Sodium azide (3.74 g, 57.6 mmol, 2 eq.) was added to the round-botton flask, followed by refluxing at 90 °C for 20 h behind a blast shield. The reaction was cooled to room temperature. The reaction was extracted with 150 mL (3 × 50 mL) acetonitrile. The combined organic layers were dried over $MgSO_4$ and concentrated by rotary evaporation behind the blast shield at room temperature to pale yellow oil (2.5 g). Storage of the concentrated 3-azido-1-propanol oil should be avoided.

THPTA ligand preparation. To a stirred solution of 3-azido-1-propanol (2.5g, 27.7 mmol, 3.7 eq) and tripropargylamine (0.97g, 7.4 mmol, 1 eq.) in 25 mL THF, cuprous

acetate (18.2 mg, 0.15 mmol, 0.02 eq) was added under argon. The resulting solution was refluxed overnight. The mixture was concentrated and stirred with Cuprisorb resin (0.04 g) to remove the copper ions. The solution was filtered and concentrated to yellow oil.



^{1}H NMR (500 MHz, DMSO-d_6): δ 8.03 (s, 3H), 4.68 (t, J = 4.95 Hz, 3H), 4.41 (t, J = 7.08 Hz, 3H), 3.61 (s, 6H), 3.40 (m, 6H), δ 1.96 (s, 6H) ppm.

4.4.4 Preparation of azido modified AYA-RNA

Azido modified AYA-RNA product was prepared both on radiolabeled and preparative scale. For radiolabeled-scale preparation. 1 pmol of 3′-^{32}P-5′- triphosphate of RNA substrate, 10 pmol of 15MZ36, and 20 pmol of 15MZ36 helper oligonucleotide were annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 1 min and cooling on ice for 5 min. 0.4 μL of azido modified AYA (from 50 mM stock solution in DMF) was added. The ligation reaction was initiated by addition of stock solutions to a final volume of 20 μL containing 50 mM HEPES, pH 7.5, 150 mM

NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$. Final concentrations were 50 nM R, 0.5 μM E, 1 μM helper oligonucleotide, and 1 mM tripeptide substrate (1:10:20:20,000 R:E:helper:tripeptide). The reaction was incubated at 37 °C for 24 h, and purified on a 20 % PAGE (**Fig. 4.4**). The excised azido modified AYA-RNA product was used for the capture step optimization.

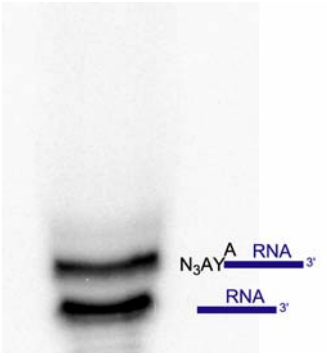


Figure 4.4 3′-^{32}P Azido-AYA-RNA prepared by the 15MZ36 deoxyribozyme.

Preparative-scale preparation. The procedures were similar to the radiolabeled-scale detailed above. 1 nmol of 5′-triphosphate RNA, 1 nmol of the DNA helper oligonucleotide, and 700 pmol of 15MZ36 were annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 1 min and cooling on ice for 5 min in 50 μL. 2 μL of of *Azido modified* AYA (from 50 mM stock solution in DMF) was added and the ligation reaction was initiated by addition of stock solutions to a final volume of 100 μL containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, and 20 mM $MnCl_2$ and incubated at 37 °C for 48 h. The reaction was ethanol precipitated, and purified on a 20 % PAGE. MALDI-TOF $[M–H]^-$calcd: 6089.9, found: 6084.4, Δ: –0.090% (**Fig. 4.5**). The excised azido modified-AYA-RNA sample was further ligated to a 5′-^{32}P- radiolabeled DNA pool for capture step optimization in cis.

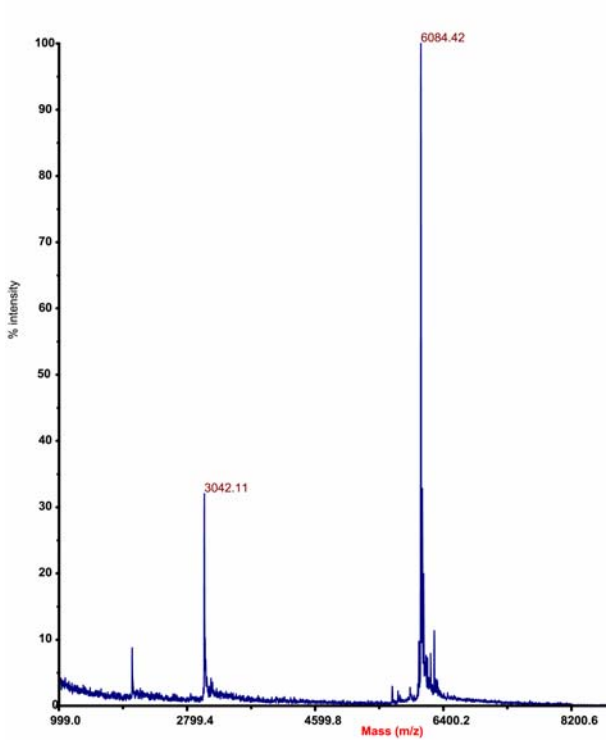


Figure 4.5 MALDI-TOF mass analysis of azido modified-AYA-RNA prepared by the 15MZ36 deoxyribozyme.

4.4.5 In vitro selection procedure using CuAAC chemistry

Selection step. In the key selection step during each round of selection, the PCR-amplified deoxyribozyme pool and 50 pmol of helper oligonucleotide (complementary to the 3′-binding arm of the deoxyribozyme) were annealed in 14 μL with 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 1 min and cooling on ice for 5 min. The sample was brought to a final volume of 20 μL containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, 20 mM $MnCl_2$, and 1 mM *Azido modified* AYA tripeptide (added from 50 mM stock solution in DMF) and incubated at 37 °C for 14 h, followed by ethanol precipitation.

Capture Step. To the selection product, 150 pmol of capture oligonucleotide and 50 pmol of splint (complementary to the capture oligonucleotide, the RNA substrate and 4 nt of the 5′ end of the DNA pool) were added to a final volume of 20 μL containing 50 mM

HEPES, pH 7.5, 150 mM NaCl, 2.8 mM THPTA, 4.0 mM sodium ascorbate, and 4 mM $CuSO_4 \cdot 5H_2O$. The latter three components were added sequentially, i.e., THPTA, then sodium ascorbate, followed by $CuSO_4 \cdot 5H_2O$. The reaction was incubated at 37 °C for 2 h and separated on 8 % PAGE.

4.5 References

1 Wong, O. Y.; Mulcrone, A. E; Silverman, S. K., DNA-Catalyzed Reductive Amination. *Angew. Chem. Int. Ed.* **2011**, *50*, 11679–11684.

2 Tornøe, C. W.; Christensen, C.; Meldal, M., Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **2002**, *67*, 3057–3062.

3 Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B., Stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596–2599.

4 Tron, G. C.; Pirali, T.; Billington, R. A.; Canonico, P. L.; Sorba, G.; Genazzani, A. A., CuAAC chemistry reactions in medicinal chemistry: applications of the 1,3-dipolar cycloaddition between azides and alkynes. *Med. Res. Rev.* **2008**, *28*, 278–308.

5 Lutz, J.-F.; Zarafshani, Z., Efficient construction of therapeutics, bioconjugates, biomaterials and bioactive surfaces using azide-alkyne "click" chemistry. *Adv. Drug Delivery Rev.* **2008**, *60*, 958–970.

6 Moses, J. E.; Moorhouse, A. D., The growing applications of click chemistry. *Chem. Soc. Rev*. **2007**, *36*, 1249–1262.

7 Fokin, V. V., Click imaging of biochemical processes in living systems. *ACS Chem. Biol.* **2007**, *2*, 775–778.

8 Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. G., Analysis and Optimization of Copper-Catalyzed Azide–Alkyne Cycloaddition for Bioconjugation. *Angew. Chem. Int. Ed.* **2008**, *48*, 9879–9883. [THPTA preparation procedure can be found in the SI.]

9 El-Sagheer, A.; Brown, T., Click chemistry with DNA. *Chem. Soc. Rev.* **2010**, *39*, 1388–1405.

10 Kumar, R.; El-Sagheer, A.; Tumpane, J.; Lincoln, P.; Wilhelmsson, L. M.; Brown, T., Template-directed oligonucleotide strand ligation, covalent intramolecular DNA circularization and catenation using click chemistry. *J. Am. Chem. Soc.* **2007**, *129*, 6859–6864.

11 Bräse, S.; Gil, C.; Knepper, K.; Zimmermann, V.,Organic Azides: An Exploding Diversity of aUnique Class of Compounds. *Angew. Chem. Int. Ed*. **2005**, *44*, 5188–5240.

12 Zhou, Y.; Wang, S.; Xie, Y.; Guan, W.; Ding, B.; Yang, Z.; Jiang, X., 1, 3-dipolar cycloaddition as a general route for functionalization of Fe_3O_4 nanoparticle. *Nanotechnology* **2008**, *19*, 175601–175605.