KINASE DEOXYRIBOZYMES

BY

SHANNON M. WALSH

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biochemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Professor Scott K. Silverman, Chair
Professor Robert B. Gennis
Professor Susan A. Martinis
Assistant Professor Erik Procko

Abstract

Nature has developed the use of proteins and RNA as enzymes, while DNA is used for the storage and transfer of genetic information. Proteins and RNA are biopolymers that can fold into specific secondary and tertiary structures to enable catalysis. Considering the structural similarity to RNA, single-stranded DNA should also be able to form complex structures capable of catalyzing reactions. DNA catalysts have not been identified in nature, but in vitro selection has led to the identification of DNA catalysts for a variety of chemical reactions. Identification of new catalysts favors the use of DNA for multiple reasons. Amplification of functional DNA sequences is directly possible using natural polymerases, whereas amplification of RNA requires an additional reverse transcription step and amplification of proteins is not possible. The total number of possible sequences is smaller for nucleic acids (4^n, where n is the number of residues) than for proteins (20^n). Within this sequence space a large number of random nucleic acid sequences will fold into secondary and tertiary structures unlike proteins which require specific amino acid sequences to form complex structures. Therefore, in vitro selection experiments to identify DNA catalysts will cover a large portion of sequence space, and a large fraction of the covered space will contain structured DNA sequences with the potential to be catalytically active. The ease of synthesis and stability of DNA compared to RNA or proteins also provides an advantage for its use as a catalyst.

Natural post-translational modifications (PTMs) are important in biological systems. PTMs modulate protein activity resulting in rapid changes to cellular processes. Studying the role of specific PTMs is often limited to the ability to generate site-specific post-translationally modified proteins of interest. Phosphorylation of amino acid side chains is an abundant natural PTM that is essential for cellular function. Protein kinases, which catalyze phosphorylation, are often motif specific. Engineering these natural kinases to change motif requirements is challenging and often results in decreased substrate specificity. To identify new catalysts for the site-specific phosphorylation of a desired protein the use of DNA as a catalytic biomolecule is

advantageous because an initially random population of DNA sequences does not have substrate biases, and DNA is a large biopolymer with the possibility to interact specifically with the substrates.

Both ribozymes and deoxyribozymes have been identified to catalyze the phosphorylation of oligonucleotides. However, previous efforts to identify kinase deoxyribozymes to catalyze the phosphorylation of amino acid side chains were unsuccessful because the γ-thiophosphoryl donor used was not stable in the selection conditions. As described herein, a new in vitro selection method was developed using a previously identified deoxyribozyme to separate the active deoxyribozymes from the inactive DNA sequences. This method led to the identification of the first kinase deoxyribozymes capable of phosphorylating tyrosine residues within a tethered peptide substrate using a bound 5′-triphosphorylated RNA oligonucleotide as the phosphoryl donor. Separate selection experiments were performed using 1 mM GTP as the phosphoryl donor. The identified DNA catalysts are able to phosphorylate tyrosine within a peptide substrate and require only low micromolar concentrations of GTP.

Site-specific modification of proteins is often desired. Most deoxyribozymes identified to modify peptide substrates have been identified using peptide substrates containing the reactive residue flanked by alanine residues. Peptide sequences derived from natural proteins contain a variety of amino acid residues with diverse functional groups that could be a point of interaction between the peptide substrate and DNA catalyst. Selection experiments were performed with biologically derived peptide sequences to identify tyrosine kinase deoxyribozymes with the ability to phosphorylate peptides sequence-specifically. Of the three peptide substrates evaluated the use of one led to deoxyribozymes that are peptide motif-specific, the second peptide led to deoxyribozymes with partial peptide sequence-selectivity, and the third did not lead to the identification of deoxyribozymes. The identification of peptide motif-specific deoxyribozymes demonstrates that DNA catalysts can interact specifically with peptide substrates, and individual DNA enzymes can interact with the same peptide substrate in a different manner.

The ability to phosphorylate substrates that are free in solution is desired. However, previously identified kinase deoxyribozymes are unable to phosphorylate untethered peptide substrates. Original efforts increased the length of the tether between the peptide substrate and DNA anchor to mimic a peptide free in solution. These selection experiments did not lead to deoxyribozymes, and further analysis of other deoxyribozymes with untethered peptide reactivity suggests the long tethers may interfere with catalysis. Further efforts have focused on the incorporation of hydrophobic modifications into the DNA catalysts to improve peptide binding. DNA aptamers containing hydrophobic modifications have improved protein binding. Increased binding affinity between the peptide substrate and DNA catalyst may enable untethered peptide reactivity.

While initial efforts focused on tyrosine phosphorylation, serine phosphorylation is also abundant in nature. Serine kinase deoxyribozymes have been identified to phosphorylate serine within tightly tethered peptide substrates using 5′-triphosphorylated RNA as the phosphoryl donor. Subsequent efforts to increase the tether length or use ATP as the phosphoryl donor were unsuccessful. Efforts to improve DNA catalysts with the ability to phosphorylate serine include using biologically derived peptide sequences to increase interactions between the deoxyribozyme and peptide substrate, and incorporating catalytically participatory modifications into the DNA enzymes.

Acknowledgments

I am grateful for everyone who has helped and supported me throughout my Ph.D. studies. I would like to thank my research advisor, Prof. Scott K. Silverman for his support and guidance. I am grateful for all of the opportunities and mentorship he has provided. His enthusiasm for science is an inspiration.

I am thankful to my thesis committee members, Professor Bob Gennis, Professor Susan Martinis, and Assistant Professor Erik Procko. Thank you for the helpful discussions and for the continued support of my future career plans. I would also like to thank Professor Huimin Zhao for discussions during my qualifying exam and annual seminars.

I am also grateful for the outstanding coworkers in the Silverman Lab who have provided support daily. I would like to thank Ying and Ben for their guidance, patience, and continued encouragement; Jimmy, for the support in and out of lab as we navigated through grad school together; Jag, for reminding me to have fun; Puzhou for the engaging scientific discussions; and Cong for turning me into an organic chemist. Victor, Josh, Peter, and Yves – I am grateful for your helpful discussions and scientific assistance.

I am thankful for all of the undergrads who worked with me: Alison, Stephanie, Nick, Tent, and Tam. I would also like to thank Paul and Yujeong. To all of you – I have learned so much from each of you, and I look forward to seeing where your future takes you.

I would like to thank my parents for their love, support, and understanding. I am grateful for Josh’s love, support, patience, and understanding throughout this adventure.

Table of Contents

Introduction to Natural and Engineered Enzymes

1.1 Enzymatic Catalysis

Enzymes catalyze reactions by adopting three-dimensional structures to bind specifically to substrates and catalyze their modification. Enzymes catalyze reactions by destabilizing ground states and stabilizing transition states.[1] Natural protein enzymes can have rates as high as 10^9 M^{-1} s^{-1}, at which point catalysis is limited by substrate diffusion.[2] The large tertiary structure of macromolecular enzymes enables many interactions between the enzyme and substrate, leading to the high selectivity of enzymes for their reaction substrates. Biochemical reactions need to be specific to ensure proper cellular function, and for this reason enzymes are used in nature to catalyze most biochemical reactions.[3]

Enzymatic catalysts are different from small-molecule catalysts, which are discrete molecules that catalyze specific chemical reactions. Small-molecule catalysts have been identified to catalyze a variety of chemical reactions. Unlike enzymes, small-molecule catalysts are often not site-specific and will react with any relevant functional group present.[4-6]

1.2 Natural Enzymes

1.2.1 Proteins

Protein enzymes are the primary catalysts in cells, and they can catalyze biochemical reactions in a highly specific manner. Proteins are polymers of amino acids connected via amide linkages. The natural amino acids contain a variety of chemical functional groups to promote protein structure and enable catalysis (Figure 1.1). The diversity of the 20 natural amino acids is expanded by unnatural amino acids and post-translational modifications. A protein's functionality is dependent upon its secondary, tertiary, and quaternary structure. The higher-order structure forms based on the primary amino acid sequence. Specific amino acid sequences

are required to form secondary structures such as α-helices and β-sheets. These secondary structures combine to form unique tertiary structures, and quaternary structures containing multiple protein subunits can be formed. Within a structured protein enzyme is a small region, called the active site, where the chemical reaction is catalyzed. Substrate interactions with the amino acid side chains are a result of the protein structure and enable the high selectivity of protein enzymes. Intended substrates have favorable steric and electrostatic interactions within the substrate binding pocket, whereas unintended substrate binding is sterically and electrostatically hindered. The complex three-dimensional structures of protein enzymes enable catalysis by destabilizing the ground state of the substrate and stabilizing the transition state.

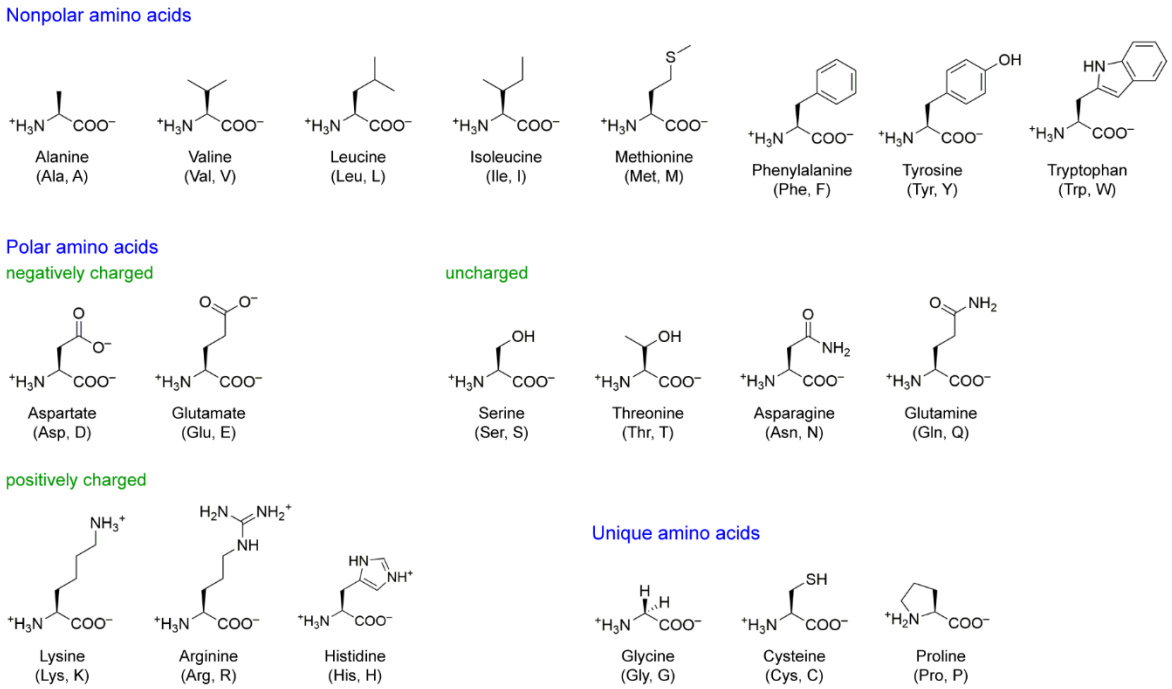


Figure 1.1. Structures of the 20 natural amino acids, whose side chains contain various functional groups. Proteins are specific sequences of amino acids connected via amide bonds.

Many natural protein enzymes catalyze chemical reactions in a highly selective and stereospecific manner. Natural proteins have been evolved to catalyze reactions in physiological conditions, which are relatively mild compared to some synthetic reactions. Most proteins have rates of ~10^5–10^9 M^{-1} s^{-1}. The most efficient enzymes are limited by the rate of diffusion of the

substrates and products from the active site, and these enzymes have rates of 10^8–10^9 M^{-1} s^{-1}. The ability of protein enzymes to catalyze chemical reactions in mild conditions in a highly selective manner is essential in biology. Problems in protein folding, substrate specificity, localization, substrate binding, or catalytic efficiency can result in disease. Protein enzymes have also been used in many industrial roles including in food applications, as detergents, and as therapeutic agents.[7]

1.2.2 Ribozymes

RNA was originally thought to be involved only in the transfer of information between the DNA data storage to the proteins that catalyzed essential cellular functions. However, RNA was discovered to be catalytic in the early 1980s.[8,9] The catalytic ability of RNA was surprising because of the differences between proteins and nucleic acids. Compared to proteins, RNA contains limited functional group diversity (Figure 1.2A). In the secondary structure of RNA, the nucleobases are involved in base-pairing in the interior of the structure. Therefore, the nucleobases were thought to be only partially accessible for substrate binding or catalysis. Additionally, the abundance of available hydrogen bonding interactions was thought to lead to random folding of an RNA oligonucleotide into many different structures, not one discrete three-dimensional structure. Due to the highly negatively charged RNA backbone, it was not clear that compact tertiary RNA structures could form. Despite these challenges, many natural ribozymes have been identified to catalyze chemical reactions.[10,11]

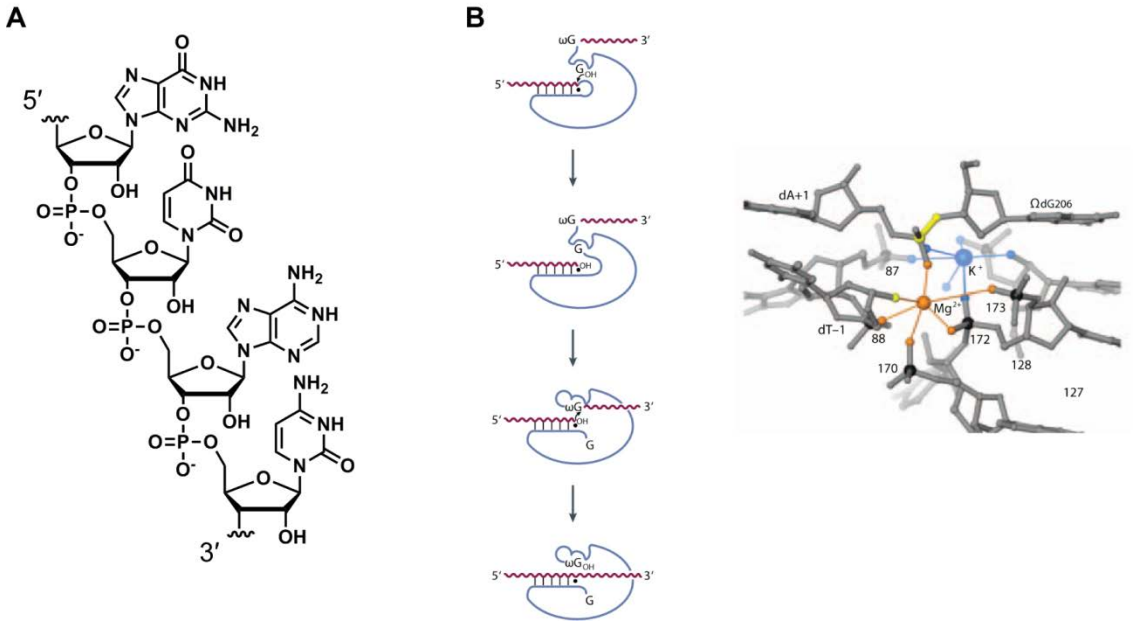


Figure 1.2. Ribozymes. (A) RNA structure. RNA contains ribonucleotides connected via phosphodiester linkages. Each ribonucleotide has one of four nucleobases; from top to bottom is guanine, uracil, adenine, and cytosine. (B) Group I intron from *Tetrahymena thermophila*. The self-splicing pathway for the group I intron is shown. A guanosine residue binds and attacks the 5′-splice site. A conformation change follows and the 3′-hydroxyl group attacks the 5′-end of the exon, resulting in ligation. The active site is shown in the crystal structure of the group I intron. Part B adapted with permission from Ref. 11.

The first ribozyme discovered in the early 1980s was the *Tetrahymena thermophile* group I intron, which catalyzes RNA cleavage as part of intron splicing (Figure 1.2B).[8,12] The group I intron contains a ~200 nucleotide catalytic core that uses the 3′-hydroxyl of guanosine residue or the 3′-hydroxyl of an exogenous guanosine as a nucleophile. This catalytic core is conserved in many other group I introns in different organisms.[13] The ribozyme RNase P was also discovered to catalyze cleavage of the RNA phosphodiester backbone in the early 1980s.[9] However, RNase P catalyzes direct RNA hydrolysis using a water molecule. RNase P is a ribonucleoprotein complex (RNP), with RNA directly catalyzing the reaction.[14] The RNA component alone is sufficient for catalysis in vitro, but in vivo activity requires the additional protein components.

Many other natural ribozymes have since been discovered,[10] including the hammerhead, hairpin, and twister ribozymes that acquired their names based on their structure. The

hammerhead ribozyme, identified in multiple organisms,[15-18]catalyzes RNA self-cleavage via transesterification resulting in a 5′-phosphate and 2′,3′-cyclic phosphate.[19,20] Interestingly, the hammerhead ribozyme has also been shown to catalyze the reverse RNA ligation reaction.[21] Catalysis of both RNA transesterification and RNA ligation is also a feature of the hairpin ribozyme.[22,23] In contrast, the twister ribozyme, which has also been identified in multiple organisms, only catalyzes RNA cleavage by transesterification.[24,25] Another ribozyme, the hepatitis delta virus (HDV) ribozyme, is required for replication of the hepatitis delta virus.[26,27] The HDV ribozyme also catalyzes RNA transesterification and uses a cytosine nucleobase as a general acid for catalysis.

Similar to RNase P, other RNA catalysts function as part of RNP complexes where both the RNA and protein are required, but the RNA performs the catalysis. The ribosome is the most abundant biological RNP in which the RNA is responsible for catalysis.[28-31] The bacterial ribosome is a large complex containing three RNA molecules (rRNA) and over 50 proteins. Within the complex, RNA catalyzes the peptidyl transferase reaction, resulting in the translation of proteins.

Ribozymes catalyze chemical reactions by folding into complex three-dimensional structures using Watson-Crick base-pairing interactions to form secondary structures that can further form complex tertiary structures.[32] RNA misfolding is possible due to the limited number of monomers that can make similar interactions. The negatively charged RNA backbone enables interaction with metal ions, which can be important for structure or catalysis.[33,34] The abundance of natural catalytic RNAs and their involvement in gene expression and protein synthesis indicate the important roles of catalytic RNAs in biological systems.

1.3 Engineered Enzymes

The catalytic efficiency and substrate specificity of natural protein and RNA enzymes result from billions of years of evolutionary pressure. Enzymes have been modified to catalyze

both biological and non-biological reactions for various purposes. Methods such as directed evolution, in vitro selection, and rational design have been developed to both engineer existing enzymes as well as identify or design new enzymes. Natural enzymes that have been modified for a desired purpose or have been artificially developed in the laboratory are referred to as engineered enzymes or artificial enzymes.

1.3.1 Proteins

Directed evolution and in vitro selection methods are based on accelerated evolution methodologies. In contrast, rational design of enzymes relies on the understanding of how enzyme function is related to its structure and mechanism. Using rational design, the earliest protein engineering attempts were site-directed mutagenesis using natural amino acids to site specifically modify enyzmes.[35] Advances in this field have led to rational design assisted by computational modeling.[36]

1.3.1.1 Directed Evolution

Directed evolution enables engineering of proteins with desired properties (Figure 1.3).[37-42] A parent protein enzyme, a known protein enzyme with activity similar to that of the desired new enzyme, is required as a starting point. A library of mutants of this parent enzyme is created. The mutant library can be prepared by various genetic mutational methods including random mutagenesis and genetic recombination. The library can be individually screened or selected for the desired activity.

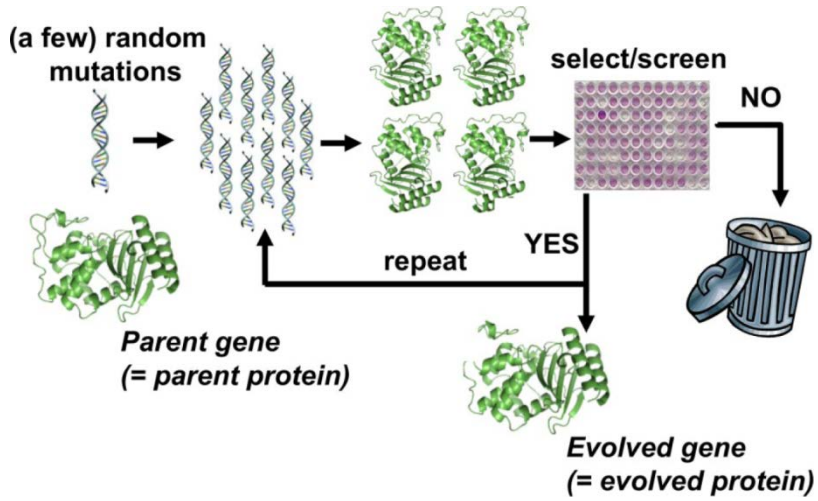


Figure 1.3. General diagram of the directed evolution of proteins. A gene for the parent protein is used as the starting point. The parent gene is subject to random mutations, resulting in a library of mutant genes. These genes are used to express protein mutants that are selected or screened for the desired property. Protein mutants that do not have the desired property are discarded, and the genes for the protein mutants with the desired property undergo mutagenesis to begin the next round. The cycle is iterated until proteins with the desired property are identified. Figure reprinted with permission from Ref. 39.

Random mutagenesis techniques use error-prone DNA replication or unrepaired random DNA damage to induce mutations. Chemical mutagens including ethane methanesulfonate and nitrous acid cause DNA damage and result in errors during replication.[43,44] However, mutations formed by these methods are biased, so they are not commonly used. Error-prone PCR can be used to incorporate genomic mutations, but mutation biases also exist.[45-47] Incorporation of deoxyinosine during PCR results in an increased mutation rate due to the promiscuous base pairing of deoxyinosine.[48] However, the redundancy of the genetic code can limit the protein library diversity generated by individual nucleotide mutations.

Structural information of proteins can be used to aid in focusing mutations to a region of the protein likely to influence function, such as near the active site. These more precise mutations are made by using synthetic DNA to introduce mutations at the desired sites.[49] If mutations are made only in specific regions of the protein, then more of the possible sequence variants can be included. The challenge with focused mutations is that the role of protein structure in catalysis is not fully understood, and mutation of residues outside of the active site

can have significant effects on enzymatic activity.[50] Focused residue mutations often exclude these residues, thus limiting the structures assayed. Gene recombination methods such as homologous recombination can also be used to introduce mutations into a protein library. DNA shuffling is a method for homologous recombination and involves digestion of a gene followed by random reassembly.[51-53] Homologous recombination methods work best when a library is based on a family of related protein sequences.

Evolution of a protein for a desired property requires a method to identify mutants with the desired enzymatic activity. High-throughput screening can be used to identify active enzyme mutants, but this method limits evaluation to 10^2-10^4 mutants. Larger libraries of 10^6-10^{12} mutants can be assayed if a selection method is developed in which protein function enables the active protein mutants to be separated or results in survival of a host organism.[54] Selection methods need to be developed separately for each desired activity. Recent development of phage-assisted continuous evolution (PACE) enables many rounds of evolution on the timescale of days. The PACE method links the desired protein function with the survival and replication of phage.[55] PACE has also been used to evaluate potential organismal drug resistance mechanisms by performing many rounds of evolution in a short time period.[56]

Directed evolution of proteins has led to advances in biology and biotechnology. For example, enzymes have been identified with increased thermostability to increase their utility,[57] and cytochrome P450s have been engineered to catalyze organic chemical reactions that can be synthetically challenging.[41,58-61] However, there are some disadvantages to directed evolution. A known parent protein enzyme is required to begin evolution. Evolution of many proteins has resulted in improved characteristics, but changing the substrate requirements often leads to relaxed selectivity to accommodate the new substrate instead of a change in selectivity.[62]

1.3.1.2 In vitro Selection by mRNA Display

Unlike protein design, in vitro selection of proteins from a random library does not require a known protein, mechanism, or structure as a starting point. While identifying new

proteins from a random population sounds desirable, in vitro selection of proteins cannot generally be performed. The main obstacle is the inability to amplify proteins. In vitro selection requires the separation of active sequences followed by their amplification to enable continued iterations of selection rounds. DNA can be directly amplified by natural polymerases, and RNA can be reverse transcribed into DNA and subsequently amplified. However, it is not possible to directly amplify proteins or reverse translate them into an amplifiable DNA sequence. Additional limitations include the ability to evaluate a large portion of the possible sequences. There are 20 possible amino acids at each position, and therefore, 20^n possible sequences for a protein of length n. For a small protein that is 100 amino acid residues in length, 20^{100} ($\sim 10^{130}$) sequences are possible, even though a much smaller sequence space is sampled. Protein structures require specific amino acid sequences, and unfolded proteins often aggregate and are not functional.

Despite these limitations, several protein enzyme selection methods have been developed, and all link each protein to their genetic information to enable amplification. The most commonly used is mRNA display (Figure 1.4),[63-66] but similar methods of ribosome display[67] and DNA display[68] have been developed. The mRNA display method relies on a covalent linkage between the protein and the mRNA that encodes it. Active proteins are separated, and the linked mRNA sequence is amplified by reverse transcription and PCR. Subsequent selection and amplification rounds are performed. To prevent redundant protein sequences, a limited number of codons are usually chosen, and the initial mRNA library is derived from combinations of these codons. Once the population of proteins is highly active, the DNA sequences are determined, and individual enzymes are evaluated. However, protein enzymes identified from mRNA display have not been identified from a completely random protein sequence library due to the inability to amplify and the limitations in sequence space coverage.

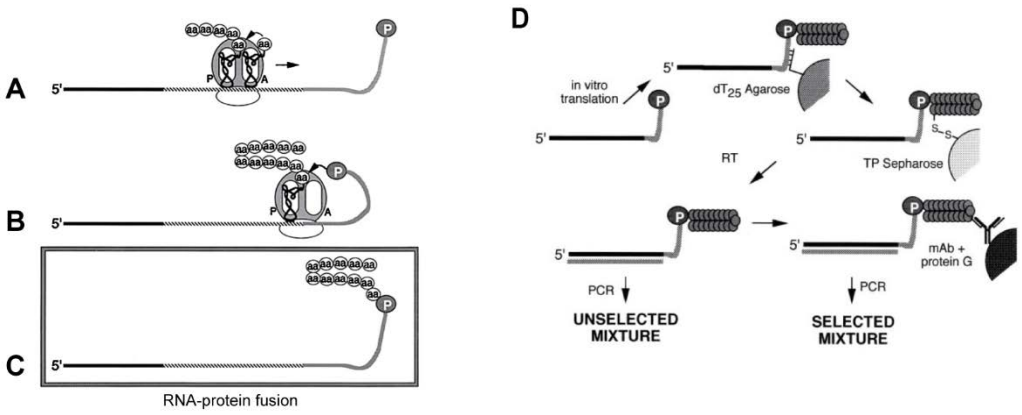


Figure 1.4. The mRNA display method for in vitro selection of proteins. (A) The ribosome synthesizes the protein based on the mRNA sequence. (B) When the ribosome reaches the end, translation stalls because puromycin at the 3′-end of the mRNA binds to the nascent protein. (C) The mRNA-protein fusion is purified. (D) Selection strategy for mRNA display. The mRNA is reverse transcribed, and the DNA product base-pairs to the mRNA to prevent it from interacting during the selection. Next, selection is performed, and the desired products are separated. PCR amplification of the desired products enables subsequent rounds of selection. Figure adapted with permission from Ref. 64.

1.3.1.3 Rational Protein Enzyme Design

De novo protein design uses the knowledge of chemical reaction mechanisms and the functional groups available from amino acid side chains to design active sites in proteins.[36,69,70] A mechanism for the desired reaction is determined, and the location of relevant functional groups within an active site is designed to interact with the transition state (Figure 1.5).[71] This theoretically designed active site, called a theozyme, is incorporated into a protein scaffold and optimized in silico. The optimized protein is expressed, and the enzyme activity is assayed.[72] The computationally designed enzymes usually have modest activity, if at all. Due to the limitations in predicting enzyme active sites and scaffolds, directed evolution of the designed proteins is often performed to improve enzymatic activity. De novo protein design combined with directed evolution has been successfully used to engineer enzymes that catalyze Kemp elimination,[73] a retro-aldol reaction,[74] a Diels-Alder reaction,[75] and ester hydrolysis.[76]

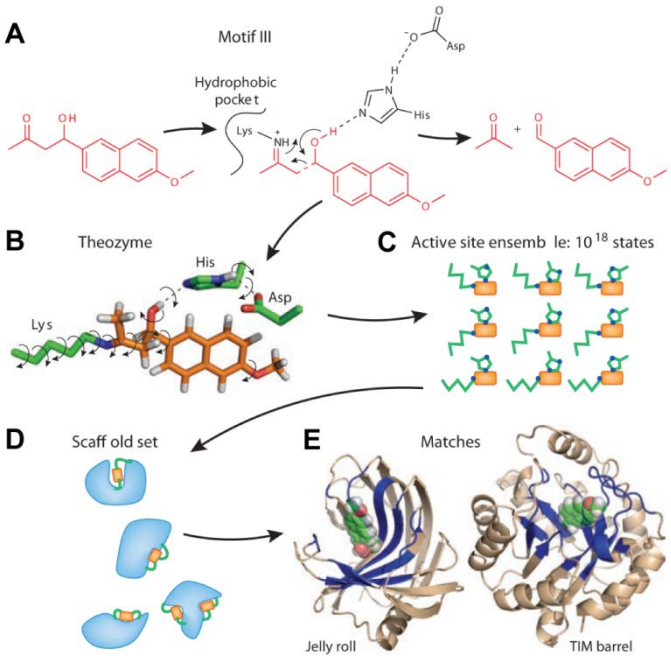


Figure 1.5. Schematic of the de novo design of protein enzymes. (A) The reaction mechanism as catalyzed by amino acid side chains is determined. (B) The theozyme is determined computationally by quantum mechanics optimization of the interactions between the substrate transition state and the catalytic residues. (C) Multiple active-site conformations are determined. (D) The active sites are mapped into potential scaffolds to determine compatible sites. (E) Further optimization is performed on the models with the active site in the scaffold. Figure reprinted with permission from Ref. 36.

After directed evolution of the rationally designed enzyme, the intended structure or mechanism is not always conserved. The engineered active site of a retro-aldolase enzyme was designed to contain a reactive lysine residue. However, after optimization of the designed enzyme via directed evolution, a new catalytic lysine residue and substrate-binding pocket was created.[77] These results highlight the limitations of rational design and the benefits of evolution and selection methods.

While the ability to design active sites as a starting point is encouraging, the need for directed evolution indicates that there are still many challenges associated with designing and predicting enzymatic protein structures. The knowledge of the complex relationships between protein structure and function is incomplete. However, de novo protein design when coupled to

directed evolution can provide a starting point to identify enzymes catalyzing reactions for which there are no known natural enzymes.[54]

1.3.2 Nucleic Acid Catalysts

The discovery of natural ribozymes led to increased interest in the capabilities of nucleic acid catalysts. In the laboratory, both RNA and DNA enzymes have been identified to catalyze a variety of chemical reactions. In vitro selection methods to identify such catalysts have been developed and rely upon the ability to amplify active sequences.

Nucleic acid selections can cover a much larger fraction of sequence space than protein selections. Nucleic acids have 4 monomers and a smaller sequence space compared to proteins which have 20 monomers. Within the sequence space evaluated a large fraction of nucleic acid sequences will fold into complex secondary and tertiary structures, whereas protein structures are formed from specific primary sequences, and most random protein sequences are unlikely to fold and thus aggregate. Therefore, nucleic acid catalysts can be more easily identified from a random population of sequences. Considering nucleic acid catalysts, DNA has some practical advantages compared to RNA in its stability, ease of synthesis, and simplicity in PCR amplification. Artificial nucleic acid enzymes can be used as sensors, catalysts for small-molecule synthesis, or catalysts for biochemical reactions.

1.3.2.1 Ribozymes

Artificial ribozymes are identified by in vitro selection, which begins with a large population of random RNA sequences (~10^{14} sequences). Each sequence in the pool is evaluated simultaneously, and the sequences with the desired activity are separated from the inactive sequences. The active sequences are then amplified via reverse transcription followed by PCR. Additional selection rounds are iterated until sequences with the desired activity dominate the population. This method enables the evaluation of a large number of random sequences at once.

Artificial ribozymes have been identified to modify RNA oligonucleotides including RNA cleavage via transesterification,[78] ligation,[79,80] polymerization,[81,82] and phosphorylation.[83,84] Ribozymes have also been identified to catalyze the Diels-Alder reaction,[85,86] aldol reaction,[87] and acyl transfer reaction.[88,89] Therefore, RNA can catalyze a wide variety of chemical reactions. Despite the large number of artificial ribozymes identified, detailed structural and mechanistic understanding of their function is limited. Crystal structures have been reported for a RNA-cleaving lead-dependent ribozyme[90,91] a ribozyme that catalyzes the Diels-Alder reaction,[92] and class I ligase ribozymes (Figure 1.6).[93,94]

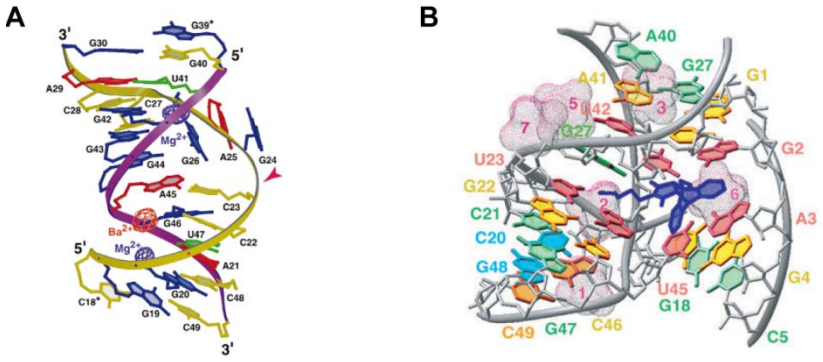


Figure 1.6. Structures of artificial ribozymes. (A) Crystal structure of a lead-dependent RNA-cleaving deoxyribozyme. The ribozyme is shown with the magenta backbone, and the substrate strand is yellow. The red arrowhead indicates the cleavage site. (B) Crystal structure of a ribozyme that catalyzes the Diels-Alder reaction. The structure shows the ribozyme-product complex with the product in blue. The electron density shows the location of bound metal ions. Part A reprinted with permission from Ref. 90. Part B reprinted with permission from Ref. 92.

1.3.2.2 Deoxyribozymes

DNA is structurally similar to RNA. DNA lacks the 2′-hdyroxyl group, and the nucleobase thymidine is present in place of uridine (Figure 1.7A). Natural deoxyribozymes have not been identified. In nature, DNA is primarily double-stranded and is unable to form complex secondary and tertiary structures required for catalysis.[95] However, single-stranded DNA can form complex structures similar to those of RNA. Based on these structural similarities, it was hypothesized that DNA could also catalyze chemical reactions, and in vitro selection

experiments were performed to identify DNA catalysts. In vitro selection to identify deoxyribozymes is similar to that for RNA, but the reverse transcription step is not required. The first DNA catalyst was reported in 1994 and catalyzes the cleavage of a single ribonucleotide in a DNA substrate (Figure 1.7B).[96]

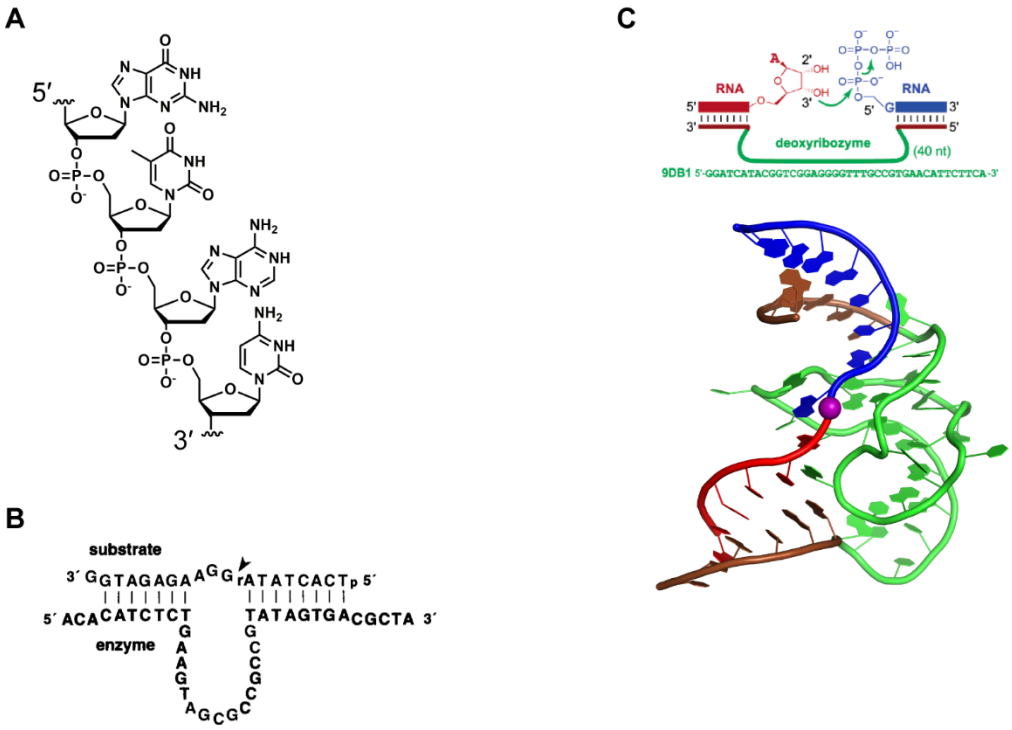


Figure 1.7. Deoxyribozymes. (A) DNA structure. DNA is made up of deoxyribonucleotides connected via phosphodiester linkages. Each deoxyribonucleotide has one of four nucleobases; from top to bottom is guanine, thymine, adenine, cytosine. (B) The first deoxyribozyme discovered catalyzes the cleavage of a ribonucleotide embedded within a DNA substrate. The cleavage site is indicated with an arrowhead, and the deoxyribozymes is Pb^{2+} dependent. (C). The 9DB1 deoxyribozyme catalyzes RNA ligation. The crystal structure of the 9DB1 deoxyribozyme is shown. The deoxyribozyme and substrates are colored to match the diagram. Part B adapted with permission from Ref. 96. Part C adapted with permission from Ref 102 and Ref. 121.

Many DNA catalysts have been identified to catalyze the modification of nucleic acid substrates, because base-pairing interactions can be engineered between the substrate and deoxyribozyme. These reactions include RNA cleavage via transesterification[97,98] or hydrolysis,[99] RNA ligation,[100-102] DNA cleavage via hydrolysis[103,104] or deglycosylation,[105] DNA

ligation,[106] and DNA phosphorylation.[107] More recent efforts have pursued DNA enzymes that catalyze the modification of peptide substrates that are bound to the deoxyribozyme via a DNA anchor and tether.[108,109] Deoxyribozymes have been identified to catalyze nucleopeptide formation between amino acid side chains (tyrosine, serine, lysine, phosphotyrosine, or phosphoserine) and a triphosphorylated or phosphorimidazolidated oligonucleotide.[110-116] DNA catalysts have also been identified to catalyze dephosphorylation of tyrosine and serine,[117] as well as elimination of phosphate from phosphoserine to form dehydroalanine.[118] Some of these deoxyribozymes can also catalyze modification of untethered peptide substrates.[115-120]

Similar to artificial ribozymes, structural and mechanistic understanding of deoxyribozyme function is limited. The first and only crystal structure of a DNA enzyme to date was identified in 2016 for the 9DB1 RNA-ligating deoxyribozyme (Figure 1.7C).[121] The identification of the first deoxyribozyme structure over two decades after the first deoxyribozyme was identified highlights the challenges in obtaining deoxyribozyme structures.

1.4 Kinases

Protein kinases catalyze the phosphorylation of amino acid side chains including tyrosine, serine, threonine, and histidine. Phosphorylation results from the transfer of the γ-phosphate from adenosine triphosphate (ATP). Protein phosphorylation regulates the cell cycle, cell growth, cell death, metabolic pathways, and cell communication.[122] Changes in protein kinase expression have been linked to many diseases including diabetes,[122] cancer,[123,124] cardiovascular disorders,[125] and neurodegenerative disorders.[126]

1.4.1 Protein Kinases

There are over 500 known protein kinases, and their target specificity is determined by interactions with the amino acid sequence surrounding the phosphorylation site. Protein phosphorylation and dephosphorylation are highly regulated and can result in protein structural

changes. The introduction of a phosphate group increases the negative charge and results in conformational changes. Protein kinases are one of the largest protein families in eukaryotes, and genes for kinases make up ~2% of eukaryotic genomes.[127] Approximately 80% of protein kinases are serine/threonine kinases, and the remaining ~20% are primarily tyrosine and histidine kinases.

Despite the different residues that are phosphorylated, protein kinases have similar active-site structures where tyrosine kinases contain a deeper substrate pocket. The active sites of protein kinases contain an ATP-binding site and a protein substrate-binding site. Protein kinases have a conserved catalytic domain of ~250 amino acids with a small N-terminal lobe comprised of β-sheets and a larger C-terminal lobe comprised of α-helices.[127] The ATP-binding pocket is located between the two lobes with the adenosine moiety in a hydrophobic pocket and the phosphates in solution. The substrate-binding pocket structure and amino acid composition impart the substrate specificity of the protein kinase. Surface and noncatalytic active-site residues can differ in charge and hydrophobicity, and these both determine substrate specificity. The protein-binding site is often specific for amino acids surrounding the phosphorylated residue, leading to phosphorylation that is specific to particular peptide sequence motifs. Free amino acids are poor substrates, and on average, protein kinases interact with four amino acids on each side of the phosphorylated residue. Substrate specificity can be further achieved by cellular localization of the kinase.

Technological advances have resulted in the ability to do complex proteomic studies.[128-130] In these studies, phosphorylated residues of many different proteins have been identified. However, the identity of biological kinases responsible for specific phosphorylation events are not determined via this method. In general, identifying kinases responsible for newly identified phosphorylation sites can be challenging.

1.4.2 Nucleic Acid Kinases

Nucleic acid catalysts have been identified to catalyze oligonucleotide phosphorylation (Figure 1.8). Kinase ribozymes catalyze self-phosphorylation of their 5′-terminus or internal 2′-hydroxyl groups,[83,131] with ATPγS as the phosphoryl donor. These oligonucleotide kinase ribozymes have been characterized and related variants identified.[84,132-134] Oligonucleotide kinase deoxyribozymes have also been identified to catalyze self-phosphorylation of their 5′-terminus.[107,135,136] More recently, a deoxyribozyme has been identified to catalyze oligonucleotide 3′-phosphorylation of discrete substrate oligonucleotides.[137] Based on the ability of nucleic acid catalysts to phosphorylate oligonucleotide hydroxyls, attempting amino acid side chain phosphorylation seemed reasonable.

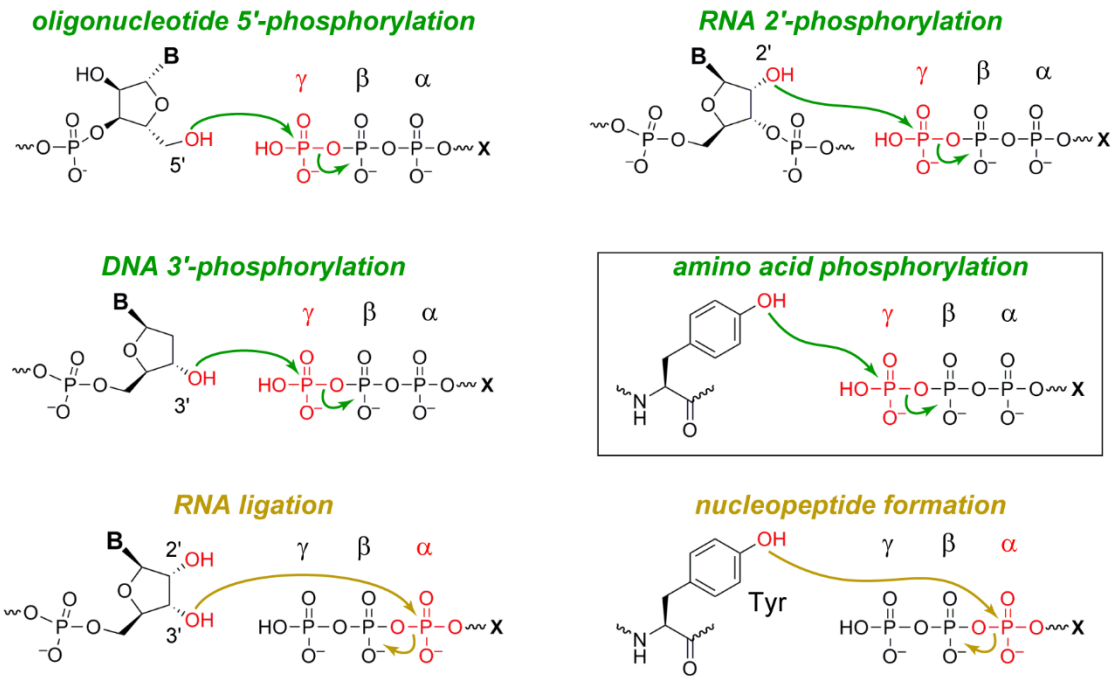


Figure 1.8. Reactions of phosphoryl donor substrates (5′-triphosphorylated RNA or NTP, where X = RNA or N). Phosphorylation results from nucleophilic attack of a hydroxyl group at the γ-phosphate. Ribozymes or deoxyribozymes have been identified to catalyze oligonucleotide 5′-phosphorylation of RNA or DNA, oligonucleotide internal 2′-phosphorylation of RNA, and oligonucleotide 3′-phosphorylation of DNA. Ribozymes or deoxyribozymes have also been identified to catalyze RNA ligation and nucleopeptide formation in which the nucleophilic hydroxyl attacks the α-phosphate of the triphosphoryl donor.

Separately, deoxyribozymes have been identified to catalyze RNA ligation[101] and nucleopeptide formation (Figure 1.8).[110,111,115,119] These DNA enzymes catalyze the reaction of hydroxyl groups with the α-phosphate of a 5′-triphosphorylated oligonucleotide. DNA-catalyzed RNA ligation results when the 3′-hydroxyl group of an RNA oligonucleotide attacks the α-phosphate of a 5′-triphosphorylated RNA oligonucleotide. Nucleopeptide formation occurs when the tyrosine or serine hydroxyl group attacks the α-phosphate of the 5′-triphosphorylated RNA oligonucleotide. Nucleopeptide formation is similar to amino acid phosphorylation, in which the tyrosine or serine side chain reacts with the γ-phosphate of a triphosphoryl donor. Therefore, based on the ability of DNA to catalyze modification of tyrosine and serine side chains with a similar phosphate electrophile, attempting amino acid side chain phosphorylation seemed reasonable.

1.4.3 Preparation of Phosphoproteins

The ability to study the role of a phosphorylated protein is limited by the availability of the desired phosphoprotein. An early method to study the role of protein phosphorylation was site-directed mutagenesis. Residues known or suspected to be phosphorylated can be mutated from their natural amino acid to a negatively charged amino acid such as aspartate or glutamate (Figure 1.9A).[138,139] These negatively charged residues roughly approximate the charge resulting from phosphorylation. This method is not precise as the size and charge of a phosphorylated residue is not accurately mimicked.

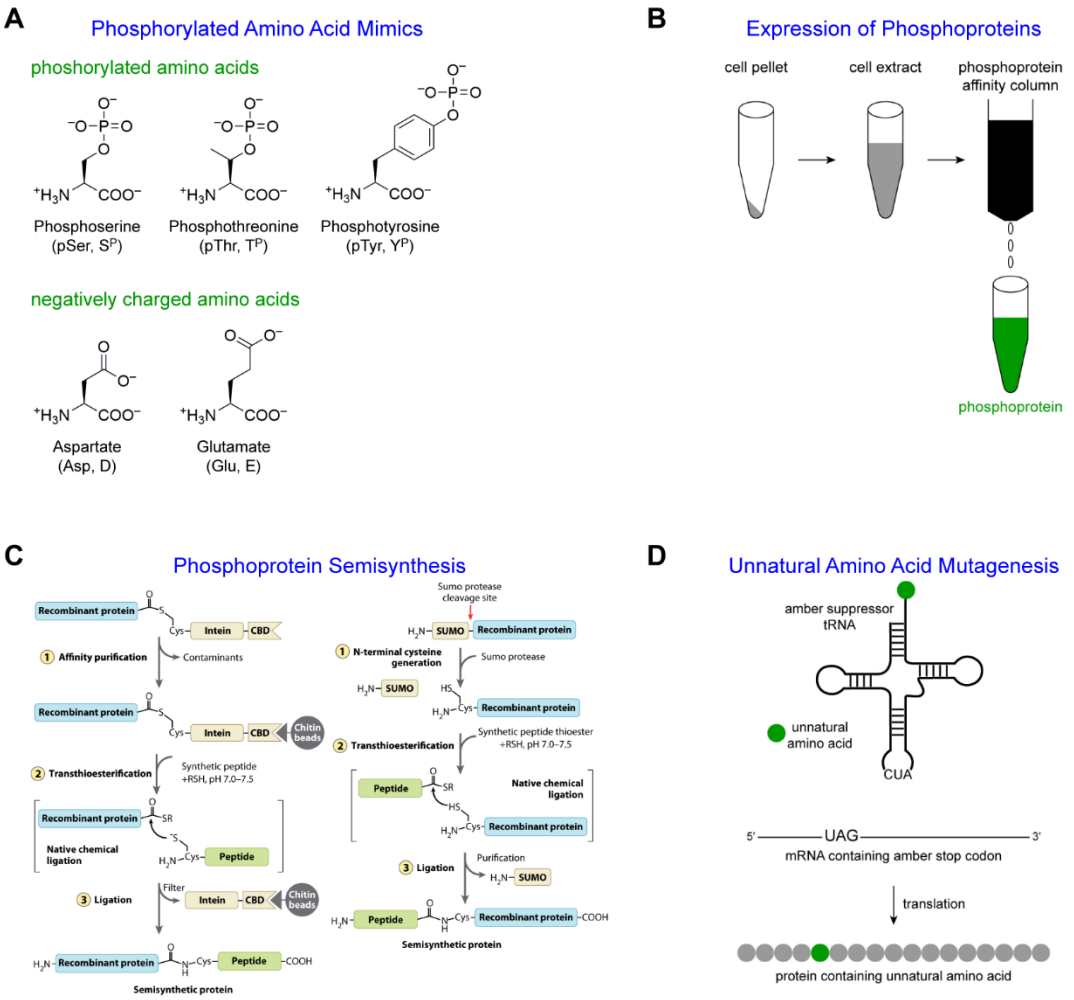


Figure 1.9. Methods to synthesize phosphorylated proteins. (A) Phosphorylated amino acid mimics. The structures of commonly phosphorylated amino acids are shown. Phosphorylated amino acids are often mimicked by site-directed mutagenesis to incorporate the negatively charged amino acids aspartate or glutamate. (B) Expression of phosphoproteins. Overexpression of proteins that are phosphorylated by the host organism can be performed. The phosphorylated proteins can be purified by phosphoprotein-specific chromatography. (C) Phosphoprotein semisynthesis. A small peptide containing a phosphorylated residue is prepared by solid-phase peptide synthesis. Expression and purification of the remainder of the protein of interest is performed. The protein and phosphopeptide fragment are covalently linked by native chemical ligation to form the full-length phosphoprotein. (D) Unnatural amino acid mutagenesis. A gene encoding the phosphoprotein is modified to include an amber stop codon at the position of the phosphorylated amino acid. An engineered host system contains a tRNA that base-pairs to the stop codon charged with the phosphorylated amino acid. Based on the modified gene, the phosphoprotein is expressed and purified. Part C reprinted with permission from Ref. 144.

Phosphoproteins can be overexpressed and isolated from cell culture (Figure 1.9B).[140] This method requires expression of the proteins in a natural system that has the ability to

phosphorylate it. The desired phosphoprotein is purified, but the purified proteins may contain a mixture of phosphorylated residues. The amount of phosphoprotein purified is dependent on the abundance of phosphorylation, and often yields are low. If a kinase is known to phosphorylate the protein of interest, then both the kinase and protein of interest can be expressed in a bacterial host.[141,142] This method requires both the kinase and protein of interest can be expressed and remain active in the host system. The purification of the phosphoprotein is subject to the same challenges.

Phosphorylated proteins can also be prepared by semisynthesis (Figure 1.9C).[143,144] Phosphorylated amino acids can be prepared synthetically and incorporated into peptides via solid-phase peptide synthesis. While proteins are too long to be prepared by solid-phase synthesis, a peptide fragment containing the phosphorylated residue can be synthesized. This peptide fragment is then incorporated into the protein via native chemical ligation, resulting in a site-specifically phosphorylated residue. This method works best when the phosphorylated residue is near one of the termini of the protein of interest. Incorporation of the phosphopeptide fragment into the middle of the protein requires multiple ligation steps. A cysteine residue needs to be included at the N-terminus of the peptide fragment to enable native chemical ligation, and the cysteine residue may not be naturally present in the protein.

Another technique to incorporate phosphorylated amino acids into proteins is unnatural amino acid mutagenesis (Figure 1.9D).[144,145] This method takes advantage of the rarely used amber stop codon in *E. coli* to incorporate a 21st amino acid. A tRNA synthetase is evolved to charge a modified tRNA with the unnatural amino acid. This modified tRNA base-pairs to the amber stop codon, thus incorporating the unnatural amino acid. The unnatural amino acid is incorporated site-specifically based on the location of the amber stop codon within the modified gene. Using this method, phosphotyrosine or a phosphotyrosine analog *p*-carboxymethyl-L-phenylalanine,[146] and phosphoserine or a phosphoserine analog 2-amino-4-phosphonobutyric acid,[147] have been incorporated into proteins. The analogs are non-hydrolyzable by natural phosphatases making them stable, but also not the natural phosphorylated residues. Difficulties

expressing the full-length protein can arise depending on the protein and the location of the unnatural residue within the protein. Incorporation of more than one unnatural nucleotide remains challenging.

The role of protein phosphorylation has been studied by generation of phosphoproteins using all four of these methods. Each method has its own advantages and limitations. However, a method to site-specifically phosphorylate any protein expressed will provide an additional tool to enable the study of phosphorylated proteins and their function in biological systems and disease.

1.5 Thesis Research Focus

The ability to site-specifically modify peptides and proteins will enable biochemical studies of post-translationally modified peptides and proteins. The role of phosphorylation in complex cellular regulation is incompletely understood.[144] With phosphorylation playing a role in 30% of cellular proteins, new technologies are needed to aid in understanding cellular pathways and their regulation.[127] Understanding these signaling pathways and manipulating them can aid in the understanding of biological processes and disease states. New catalysts to phosphorylate tyrosine and serine side chains are sought. For optimal practical utility, these new catalysts will need to phosphorylate untethered peptide and protein substrates using NTPs as phosphoryl donors. Towards this goal, new kinase deoxyribozymes are identified.

Identification of new catalysts for peptide and protein phosphorylation favors the use of DNA as a catalytic biopolymer. In vitro selection experiments to identify nucleic acid catalysts can be performed due to the ability to amplify sequences using natural polymerases, whereas no amplification method is available for proteins. The amplification of RNA catalysts requires an additional reverse transcription step prior to amplification via PCR. Additionally, nucleic acid selection experiments evaluate a significantly larger fraction of the total number of possible sequences than protein selection experiments, and unlike proteins, most of the random nucleic

acid sequences can form secondary and tertiary structures. Compared to RNA, DNA is more stable, easier to synthesize, and directly amplifiable by natural polymerases.

Based on the ability to find new catalysts from random populations of DNA, and the ability of DNA to catalyze oligonucleotide phosphorylation and nucleopeptide formation, it is reasonable to hypothesize that kinase deoxyribozymes can be identified to phosphorylate amino acid side chains within peptide and protein substrates. The work described herein focused on the identification of kinase deoxyribozymes for peptide and protein modification. Chapter 2 reports the identification of the first tyrosine kinase deoxyribozymes using a novel in vitro selection strategy. These DNA enzymes catalyze the phosphorylation of tyrosine residues within a peptide substrate and use either a 5′-triphosphorylated RNA oligonucleotide or NTPs as phosphoryl donors. Chapter 3 discusses the development of a new phosphopeptide capture method and the resulting identification of peptide-sequence-specific tyrosine kinase deoxyribozymes. The DNA catalysts specifically interact with 2-4 amino acids within the peptide substrate. Chapter 4 discusses efforts to identify kinase deoxyribozymes capable of catalyzing the phosphorylation of untethered peptide substrates. These efforts include the use of long tethers between the peptide substrate and DNA anchor oligonucleotide, and the use of hydrophobic modifications to increase the binding affinity between the DNA enzyme and peptide substrate. Chapter 5 focuses on the identification of serine kinase deoxyribozymes. DNA catalysts were identified to phosphorylate serine within a peptide substrate, and efforts for peptide-sequence-selectivity and untethered peptide reactivity were pursued.

1.6 References

(1) Miller, B. G.; Wolfenden, R. Catalytic proficiency: the unusual case of OMP decarboxylase. *Annu. Rev. Biochem.* **2002**, *71*, 847-885.

(2) Wolfenden, R.; Snider, M. J. The depth of chemical time and the power of enzymes as catalysts. *Acc. Chem. Res.* **2001**, *34*, 938-945.

(3) Srere, P. A. Why are enzymes so big? *Trends Biochem. Sci.*, *9*, 387-390.

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

(5) Tilley, S. D.; Francis, M. B. Tyrosine-selective protein alkylation using pi-allylpalladium complexes. *J. Am. Chem. Soc.* **2006**, *128*, 1080-1081.

(6) Stephanopoulos, N.; Francis, M. B. Choosing an effective protein bioconjugation strategy. *Nat. Chem. Biol.* **2011**, *7*, 876-884.

(7) Industrial Enzymes. In *Enzymes in Industry*; Wiley-VCH Verlag GmbH & Co. KGaA, 2007, pp. 99-262.

(8) 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.

(9) Guerrier-Takada, 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.

(10) Doherty, E. A.; Doudna, J. A. Ribozyme structures and mechanisms. *Annu. Rev. Biochem.* **2000**, *69*, 597-615.

(11) Fedor, M. J.; Williamson, J. R. The catalytic diversity of RNAs. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 399-412.

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

(13) Michel, F.; Westhof, E. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* **1990**, *216*, 585-610.

(14) Stark, B. C.; Kole, R.; Bowman, E. J.; Altman, S. Ribonuclease P: an enzyme with an essential RNA component. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 3717-3721.

(15) Prody, G. A.; Bakos, J. T.; Buzayan, J. M.; Schneider, I. R.; Bruening, G. Autolytic processing of dimeric plant virus satellite RNA. *Science* **1986**, *231*, 1577-1580.

(16) Forster, A. C.; Symons, R. H. Self-cleavage of virusoid RNA is performed by the proposed 55-nucleotide active site. *Cell* **1987**, *50*, 9-16.

(17) Uhlenbeck, O. C. A small catalytic oligoribonucleotide. *Nature* **1987**, *328*, 596-600.

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

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

(20) Pley, H. W.; Flaherty, K. M.; McKay, D. B. Three-dimensional structure of a hammerhead ribozyme. *Nature* **1994**, *372*, 68-74.

(21) Canny, M. D.; Jucker, F. M.; Pardi, A. Efficient ligation of the Schistosoma hammerhead ribozyme. *Biochemistry* **2007**, *46*, 3826-3834.

(22) Hampel, A.; Tritz, R. RNA catalytic properties of the minimum (-)sTRSV sequence. *Biochemistry* **1989**, *28*, 4929-4933.

(23) Nesbitt, S.; Hegg, L. A.; Fedor, M. J. An unusual pH-independent and metal-ion-independent mechanism for hairpin ribozyme catalysis. *Chem. Biol.* **1997**, *4*, 619-630.

(24) Liu, Y.; Wilson, T. J.; McPhee, S. A.; Lilley, D. M. Crystal structure and mechanistic investigation of the twister ribozyme. *Nat. Chem. Biol.* **2014**, *10*, 739-744.

(25) Roth, A.; Weinberg, Z.; Chen, A. G.; Kim, P. B.; Ames, T. D.; Breaker, R. R. A widespread self-cleaving ribozyme class is revealed by bioinformatics. *Nat. Chem. Biol.* **2014**, *10*, 56-60.

(26) Perrotta, A. T.; Been, M. D. A pseudoknot-like structure required for efficient self-cleavage of hepatitis delta virus RNA. *Nature* **1991**, *350*, 434-436.

(27) Ferré-D'Amaré, A. R.; Zhou, K.; Doudna, J. A. Crystal structure of a hepatitis delta virus ribozyme. *Nature* **1998**, *395*, 567-574.

(28) Ban, N.; Nissen, P.; Hansen, J.; Moore, P. B.; Steitz, T. A. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science* **2000**, *289*, 905-920.

(29) 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.

(30) Steitz, T. A. A structural understanding of the dynamic ribosome machine. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 242-253.

(31) Hiller, D. A.; Singh, V.; Zhong, M.; Strobel, S. A. A two-step chemical mechanism for ribosome-catalysed peptide bond formation. *Nature* **2011**, *476*, 236-239.

(32) Tian, B.; Bevilacqua, P. C.; Diegelman-Parente, A.; Mathews, M. B. The double-stranded-RNA-binding motif: interference and much more. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 1013-1023.

(33) Lippert, B. Ligand-pKa shifts through metals: potential relevance to ribozyme chemistry. *Chem. Biodivers.* **2008**, *5*, 1455-1474.

(34) Smith, M. D.; Mehdizadeh, R.; Olive, J. E.; Collins, R. A. The ionic environment determines ribozyme cleavage rate by modulation of nucleobase pKa. *RNA* **2008**, *14*, 1942-1949.

(35) Edelheit, O.; Hanukoglu, A.; Hanukoglu, I. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnol.* **2009**, *9*, 61.

(36) Nanda, V. Do-it-yourself enzymes. *Nat. Chem. Biol.* **2008**, *4*, 273-275.

(37) Arnold, F. H.; Volkov, A. A. Directed evolution of biocatalysts. *Curr. Opin. Chem. Biol.* **1999**, *3*, 54-59.

(38) Cherry, J. R.; Fidantsef, A. L. Directed evolution of industrial enzymes: an update. *Curr. Opin. Biotechnol.* **2003**, *14*, 438-443.

(39) Bloom, J. D.; Arnold, F. H. In the light of directed evolution: pathways of adaptive protein evolution. *Proc. Natl. Acad. Sci. USA* **2009**, *106 Suppl 1*, 9995-10000.

(40) Renata, H.; Wang, Z. J.; Arnold, F. H. Expanding the enzyme universe: accessing non-natural reactions by mechanism-guided directed evolution. *Angew. Chem. Int. Ed.* **2015**, *54*, 3351-3367.

(41) Roiban, G. D.; Reetz, M. T. Expanding the toolbox of organic chemists: directed evolution of P450 monooxygenases as catalysts in regio- and stereoselective oxidative hydroxylation. *Chem. Commun.* **2015**, *51*, 2208-2224.

(42) Currin, A.; Swainston, N.; Day, P. J.; Kell, D. B. Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. *Chem. Soc. Rev.* **2015**, *44*, 1172-1239.

(43) Myers, R. M.; Lerman, L. S.; Maniatis, T. A general method for saturation mutagenesis of cloned DNA fragments. *Science* **1985**, *229*, 242-247.

(44) Lai, Y. P.; Huang, J.; Wang, L. F.; Li, J.; Wu, Z. R. A new approach to random mutagenesis in vitro. *Biotechnol. Bioeng.* **2004**, *86*, 622-627.

(45) Leung, D. W.; Chen, E.; Goeddel, D. V. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* **1989**, *1*, 11-15.

(46) Cadwell, R. C.; Joyce, G. F. Randomization of genes by PCR mutagenesis. *PCR Methods Appl.* **1992**, *2*, 28-33.

(47) Gupta, R. D.; Tawfik, D. S. Directed enzyme evolution via small and effective neutral drift libraries. *Nat. Methods* **2008**, *5*, 939-942.

(48) Wong, T. S.; Tee, K. L.; Hauer, B.; Schwaneberg, U. Sequence saturation mutagenesis (SeSaM): a novel method for directed evolution. *Nucleic Acids Res.* **2004**, *32*, e26.

(49) Wells, J. A.; Vasser, M.; Powers, D. B. Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites. *Gene* **1985**, *34*, 315-323.

(50) Marshall, N. M.; Garner, D. K.; Wilson, T. D.; Gao, Y. G.; Robinson, H.; Nilges, M. J.; Lu, Y. Rationally tuning the reduction potential of a single cupredoxin beyond the natural range. *Nature* **2009**, *462*, 113-116.

(51) Stemmer, W. P. Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **1994**, *370*, 389-391.

(52) Stemmer, W. P. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 10747-10751.

(53) Coco, W. M.; Levinson, W. E.; Crist, M. J.; Hektor, H. J.; Darzins, A.; Pienkos, P. T.; Squires, C. H.; Monticello, D. J. DNA shuffling method for generating highly recombined genes and evolved enzymes. *Nat. Biotechnol.* **2001**, *19*, 354-359.

(54) Packer, M. S.; Liu, D. R. Methods for the directed evolution of proteins. *Nat. Rev. Genet.* **2015**, *16*, 379-394.

(55) Esvelt, K. M.; Carlson, J. C.; Liu, D. R. A system for the continuous directed evolution of biomolecules. *Nature* **2011**, *472*, 499-503.

(56) Dickinson, B. C.; Packer, M. S.; Badran, A. H.; Liu, D. R. A system for the continuous directed evolution of proteases rapidly reveals drug-resistance mutations. *Nat. Commun.* **2014**, *5*, 5352.

(57) Kuchner, O.; Arnold, F. H. Directed evolution of enzyme catalysts. *Trends Biotechnol.* **1997**, *15*, 523-530.

(58) Lewis, J. C.; Bastian, S.; Bennett, C. S.; Fu, Y.; Mitsuda, Y.; Chen, M. M.; Greenberg, W. A.; Wong, C. H.; Arnold, F. H. Chemoenzymatic elaboration of monosaccharides using engineered cytochrome P450BM3 demethylases. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16550-16555.

(59) Coelho, P. S.; Brustad, E. M.; Kannan, A.; Arnold, F. H. Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. *Science* **2013**, *339*, 307-310.

(60) Coelho, P. S.; Wang, Z. J.; Ener, M. E.; Baril, S. A.; Kannan, A.; Arnold, F. H.; Brustad, E. M. A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins in vivo. *Nat. Chem. Biol.* **2013**, *9*, 485-487.

(61) Farwell, C. C.; Zhang, R. K.; McIntosh, J. A.; Hyster, T. K.; Arnold, F. H. Enantioselective enzyme-catalyzed aziridination enabled by active-site evolution of a cytochrome P450. *ACS Cent. Sci.* **2015**, *1*, 89-93.

(62) Yoo, T. H.; Pogson, M.; Iverson, B. L.; Georgiou, G. Directed evolution of highly selective proteases by using a novel FACS-based screen that capitalizes on the p53 regulator MDM2. *ChemBioChem* **2012**, *13*, 649-653.

(63) Nemoto, N.; Miyamoto-Sato, E.; Husimi, Y.; Yanagawa, H. In vitro virus: bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Lett.* **1997**, *414*, 405-408.

(64) Roberts, R. W.; Szostak, J. W. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 12297-12302.

(65) Keefe, A. D.; Szostak, J. W. Functional proteins from a random-sequence library. *Nature* **2001**, *410*, 715-718.

(66) Seelig, B.; Szostak, J. W. Selection and evolution of enzymes from a partially randomized non-catalytic scaffold. *Nature* **2007**, *448*, 828-831.

(67) Hanes, J.; Pluckthun, A. In vitro selection and evolution of functional proteins by using ribosome display. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 4937-4942.

(68) Yonezawa, M.; Doi, N.; Kawahashi, Y.; Higashinakagawa, T.; Yanagawa, H. DNA display for in vitro selection of diverse peptide libraries. *Nucleic Acids Res.* **2003**, *31*, e118.

(69) Coxon, J. M.; Thorpe, A. J. Theozymes for intramolecular ring cyclization reactions. *J. Am. Chem. Soc.* **1999**, *121*, 10955-10957.

(70) Simón, L.; Muñiz, F. M.; Sáez, S.; Raposo, C.; Morán, J. R. From theozymes to artificial enzymes: Enzyme-like receptors for Michael additions with oxyanion holes and active amino groups. *Eur. J. Org. Chem.* **2007**, *2007*, 4821-4830.

(71) Tantillo, D. J.; Chen, J.; Houk, K. N. Theozymes and compuzymes: theoretical models for biological catalysis. *Curr. Opin. Chem. Biol.* **1998**, *2*, 743-750.

(72) Azoitei, M. L.; Correia, B. E.; Ban, Y. E.; Carrico, C.; Kalyuzhniy, O.; Chen, L.; Schroeter, A.; Huang, P. S.; McLellan, J. S.; Kwong, P. D.; Baker, D.; Strong, R. K.; Schief, W. R. Computation-guided backbone grafting of a discontinuous motif onto a protein scaffold. *Science* **2011**, *334*, 373-376.

(73) Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. Kemp elimination catalysts by computational enzyme design. *Nature* **2008**, *453*, 190-195.

(74) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F., 3rd; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. De novo computational design of retro-aldol enzymes. *Science* **2008**, *319*, 1387-1391.

(75) Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; St Clair, J. L.; Gallaher, J. L.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L.; Houk, K. N.; Michael, F. E.; Baker, D. Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science* **2010**, *329*, 309-313.

(76) Richter, F.; Blomberg, R.; Khare, S. D.; Kiss, G.; Kuzin, A. P.; Smith, A. J.; Gallaher, J.; Pianowski, Z.; Helgeson, R. C.; Grjasnow, A.; Xiao, R.; Seetharaman, J.; Su, M.; Vorobiev, S.; Lew, S.; Forouhar, F.; Kornhaber, G. J.; Hunt, J. F.; Montelione, G. T.; Tong, L.; Houk, K. N.; Hilvert, D.; Baker, D. Computational design of catalytic dyads and oxyanion holes for ester hydrolysis. *J. Am. Chem. Soc.* **2012**, *134*, 16197-16206.

(77) Giger, L.; Caner, S.; Obexer, R.; Kast, P.; Baker, D.; Ban, N.; Hilvert, D. Evolution of a designed retro-aldolase leads to complete active site remodeling. *Nat. Chem. Biol.* **2013**, *9*, 494-498.

(78) Pan, T.; Uhlenbeck, O. C. A small metalloribozyme with a two-step mechanism. *Nature* **1992**, *358*, 560-563.

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

(80) Hager, A. J.; Szostak, J. W. Isolation of novel ribozymes that ligate AMP-activated RNA substrates. *Chem. Biol.* **1997**, *4*, 607-617.

(81) Ekland, E. H.; Bartel, D. P. RNA-catalysed RNA polymerization using nucleoside triphosphates. *Nature* **1996**, *382*, 373-376.

(82) 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.

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

(84) Curtis, E. A.; Bartel, D. P. Synthetic shuffling and in vitro selection reveal the rugged adaptive fitness landscape of a kinase ribozyme. *RNA* **2013**, *19*, 1116-1128.

(85) Seelig, B.; Jaschke, A. A small catalytic RNA motif with Diels-Alderase activity. *Chem. Biol.* **1999**, *6*, 167-176.

(86) Seelig, B.; Keiper, S.; Stuhlmann, F.; Jaschke, A. Enantioselective ribozyme catalysis of a bimolecular cycloaddition reaction. *Angew. Chem. Int. Ed.* **2000**, *39*, 4576-4579.

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

(88) Lohse, P. A.; Szostak, J. W. Ribozyme-catalysed amino-acid transfer reactions. *Nature* **1996**, *381*, 442-444.

(89) Jenne, A.; Famulok, M. A novel ribozyme with ester transferase activity. *Chem. Biol.* **1998**, *5*, 23-34.

(90) Wedekind, J. E.; McKay, D. B. Crystal structure of a lead-dependent ribozyme revealing metal binding sites relevant to catalysis. *Nat. Struct. Biol.* **1999**, *6*, 261-268.

(91) Wedekind, J. E.; McKay, D. B. Crystal structure of the leadzyme at 1.8 A resolution: metal ion binding and the implications for catalytic mechanism and allo site ion regulation. *Biochemistry* **2003**, *42*, 9554-9563.

(92) Serganov, A.; Keiper, S.; Malinina, L.; Tereshko, V.; Skripkin, E.; Höbartner, C.; Polonskaia, A.; Phan, A. T.; Wombacher, R.; Micura, R.; Dauter, Z.; Jäschke, A.; Patel, D.

J. Structural basis for Diels-Alder ribozyme-catalyzed carbon-carbon bond formation. *Nat. Struct. Mol. Biol.* **2005**, *12*, 218-224.

(93) Shechner, D. M.; Grant, R. A.; Bagby, S. C.; Koldobskaya, Y.; Piccirilli, J. A.; Bartel, D. P. Crystal structure of the catalytic core of an RNA-polymerase ribozyme. *Science* **2009**, *326*, 1271-1275.

(94) Shechner, D. M.; Bartel, D. P. The structural basis of RNA-catalyzed RNA polymerization. *Nat. Struct. Mol. Biol.* **2011**, *18*, 1036-1042.

(95) Mills, J. B.; Hagerman, P. J. Origin of the intrinsic rigidity of DNA. *Nucleic Acids Res.* **2004**, *32*, 4055-4059.

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

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

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

(99) Parker, D. J.; Xiao, Y.; Aguilar, J. M.; Silverman, S. K. DNA catalysis of a normally disfavored RNA hydrolysis reaction. *J. Am. Chem. Soc.* **2013**, *135*, 8472-8475.

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

(101) Coppins, R. L.; Silverman, S. K. A DNA enzyme that mimics the first step of RNA splicing. *Nat. Struct. Mol. Biol.* **2004**, *11*, 270-274.

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

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

(104) Xiao, Y.; Wehrmann, R. J.; Ibrahim, N. A.; Silverman, S. K. Establishing broad generality of DNA catalysts for site-specific hydrolysis of single-stranded DNA. *Nucleic Acids Res.* **2012**, *40*, 1778-1786.

(105) Dokukin, V.; Silverman, S. K. Lanthanide ions as required cofactors for DNA catalysts. *Chem. Sci.* **2012**, *3*, 1707-1714.

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

(107) Li, Y.; Breaker, R. R. Phosphorylating DNA with DNA. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 2746-2751.

(108) Silverman, S. K. Pursuing DNA catalysts for protein modification. *Acc. Chem. Res.* **2015**, *48*, 1369-1379.

(109) Silverman, S. K. Catalytic DNA: Scope, applications, and biochemistry of deoxyribozymes. *Trends Biochem. Sci.* **2016**, *41*, 595-609.

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

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

(112) Wong, O.; 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.

(113) Sachdeva, A.; Chandra, M.; Chandrasekar, J.; Silverman, S. K. Covalent tagging of phosphorylated peptides by phosphate-specific deoxyribozymes. *ChemBioChem* **2012**, *13*, 654-657.

(114) Brandsen, B. M.; Velez, T. E.; Sachdeva, A.; Ibrahim, N. A.; Silverman, S. K. DNA-catalyzed lysine side chain modification. *Angew. Chem. Int. Ed.* **2014**, *53*, 9045-9050.

(115) Chu, C.; Wong, O.; Silverman, S. K. A generalizable DNA-catalyzed approach to peptide-nucleic acid conjugation. *ChemBioChem* **2014**, *15*, 1905-1910.

(116) Wang, P.; Silverman, S. K. DNA-catalyzed introduction of azide at tyrosine for peptide modification. *Angew. Chem. Int. Ed.* **2016**, *55*, 10052-10056.

(117) Chandrasekar, J.; Silverman, S. K. Catalytic DNA with phosphatase activity. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5315-5320.

(118) Chandrasekar, J.; Wylder, A. C.; Silverman, S. K. Phosphoserine lyase deoxyribozymes: DNA-catalyzed formation of dehydroalanine residues in peptides. *J. Am. Chem. Soc.* **2015**, *137*, 9575-9578.

(119) 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.

(120) Chu, C. C.; Silverman, S. K. Assessing histidine tags for recruiting deoxyribozymes to catalyze peptide and protein modification reactions. *Org. Biomol. Chem.* **2016**, *14*, 4697-4703.

(121) Ponce-Salvatierra, A.; Wawrzyniak-Turek, K.; Steuerwald, U.; Höbartner, C.; Pena, V. Crystal structure of a DNA catalyst. *Nature* **2016**, *529*, 231-234.

(122) Martić, S.; Kraatz, H.-B. Chemical biology toolkit for exploring protein kinase catalyzed phosphorylation reactions. *Chem. Sci.* **2013**, *4*, 42-59.

(123) Chen, Y.; Choong, L. Y.; Lin, Q.; Philp, R.; Wong, C. H.; Ang, B. K.; Tan, Y. L.; Loh, M. C.; Hew, C. L.; Shah, N.; Druker, B. J.; Chong, P. K.; Lim, Y. P. Differential expression of novel tyrosine kinase substrates during breast cancer development. *Mol. Cell Proteomics* **2007**, *6*, 2072-2087.

(124) Chong, P. K.; Lee, H.; Kong, J. W.; Loh, M. C.; Wong, C. H.; Lim, Y. P. Phosphoproteomics, oncogenic signaling and cancer research. *Proteomics* **2008**, *8*, 4370-4382.

(125) Denise Martin, E.; De Nicola, G. F.; Marber, M. S. New therapeutic targets in cardiology: p38 alpha mitogen-activated protein kinase for ischemic heart disease. *Circulation* **2012**, *126*, 357-368.

(126) Flajolet, M.; He, G.; Heiman, M.; Lin, A.; Nairn, A. C.; Greengard, P. Regulation of Alzheimer's disease amyloid-beta formation by casein kinase I. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 4159-4164.

(127) Ubersax, J. A.; Ferrell Jr, J. E. Mechanisms of specificity in protein phosphorylation. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 530-541.

(128) Blagoev, B.; Ong, S. E.; Kratchmarova, I.; Mann, M. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat. Biotechnol.* **2004**, *22*, 1139-1145.

(129) Amanchy, R.; Kalume, D. E.; Iwahori, A.; Zhong, J.; Pandey, A. Phosphoproteome analysis of HeLa cells using stable isotope labeling with amino acids in cell culture (SILAC). *J. Proteome Res.* **2005**, *4*, 1661-1671.

(130) Chen, Y.; Low, T. Y.; Choong, L. Y.; Ray, R. S.; Tan, Y. L.; Toy, W.; Lin, Q.; Ang, B. K.; Wong, C. H.; Lim, S.; Li, B.; Hew, C. L.; Sze, N. S.; Druker, B. J.; Lim, Y. P. Phosphoproteomics identified Endofin, DCBLD2, and KIAA0582 as novel tyrosine phosphorylation targets of EGF signaling and Iressa in human cancer cells. *Proteomics* **2007**, *7*, 2384-2397.

(131) Lorsch, J. R.; Szostak, J. W. Kinetic and thermodynamic characterization of the reaction catalyzed by a polynucleotide kinase ribozyme. *Biochemistry* **1995**, *34*, 15315-15327.

(132) Burke, D. H.; Rhee, S. S. Assembly and activation of a kinase ribozyme. *RNA* **2010**, *16*, 2349-2359.

(133) Biondi, E.; Maxwell, A. W.; Burke, D. H. A small ribozyme with dual-site kinase activity. *Nucleic Acids Res.* **2012**, *40*, 7528-7540.

(134) Biondi, E.; Poudyal, R. R.; Forgy, J. C.; Sawyer, A. W.; Maxwell, A. W.; Burke, D. H. Lewis acid catalysis of phosphoryl transfer from a copper(II)-NTP complex in a kinase ribozyme. *Nucleic Acids Res.* **2013**, *41*, 3327-3338.

(135) Li, Y.; Breaker, R. R. In vitro selection of kinase and ligase deoxyribozymes. *Methods* **2001**, *23*, 179-190.

(136) 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.

(137) Camden, A. J.; Walsh, S. M.; Suk, S. H.; Silverman, S. K. DNA oligonucleotide 3'-phosphorylation by a DNA enzyme. *Biochemistry* **2016**, *55*, 2671-2676.

(138) Curmi, P. A.; Maucuer, A.; Asselin, S.; Lecourtois, M.; Chaffotte, A.; Schmitter, J. M.; Sobel, A. Molecular characterization of human stathmin expressed in *Escherichia coli*: site-directed mutagenesis of two phosphorylatable serines (Ser-25 and Ser-63). *Biochem. J.* **1994**, *300*, 331-338.

(139) Huang, W.; Erikson, R. L. Constitutive activation of Mek1 by mutation of serine phosphorylation sites. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 8960-8963.

(140) Wingfield, P. T. Overview of the purification of recombinant proteins produced in *Escherichia coli*. *Curr. Protoc. Protein Sci.* **2003**, *Chapter 6*, Unit 6.1.

(141) Khokhlatchev, A.; Xu, S.; English, J.; Wu, P.; Schaefer, E.; Cobb, M. H. Reconstitution of mitogen-activated protein kinase phosphorylation cascades in bacteria. Efficient synthesis of active protein kinases. *J. Biol. Chem.* **1997**, *272*, 11057-11062.

(142) Ren, Z.; Schaefer, T. S. Isopropyl-beta-D-thiogalactosidase (IPTG)-inducible tyrosine phosphorylation of proteins in *E. coli*. *Biotechniques* **2001**, *31*, 1254, 1256, 1258.

(143) Szewczuk, L. M.; Tarrant, M. K.; Cole, P. A. Protein phosphorylation by semisynthesis: from paper to practice. *Methods Enzymol.* **2009**, *462*, 1-24.

(144) Tarrant, M. K.; Cole, P. A. The chemical biology of protein phosphorylation. *Annu. Rev. Biochem.* **2009**, *78*, 797-825.

(145) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. Expanding the genetic code of *Escherichia coli*. *Science* **2001**, *292*, 498-500.

(146) Rust, H. L.; Subramanian, V.; West, G. M.; Young, D. D.; Schultz, P. G.; Thompson, P. R. Using unnatural amino acid mutagenesis to probe the regulation of PRMT1. *ACS Chem. Biol.* **2014**, *9*, 649-655.

(147) Rogerson, D. T.; Sachdeva, A.; Wang, K.; Haq, T.; Kazlauskaite, A.; Hancock, S. M.; Huguenin-Dezot, N.; Muqit, M. M. K.; Fry, A. M.; Bayliss, R.; Chin, J. W. Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. *Nat. Chem. Biol.* **2015**, *11*, 496-503.

Chapter 2: DNA Catalysts with Tyrosine Kinase Activity[†]

2.1 Introduction

2.1.1 Previous Efforts Towards DNA-Catalyzed Peptide Phosphorylation

Previous in vitro selection experiments to identify deoxyribozymes have relied upon PAGE shift as a means to separate catalytically active DNA sequences from inactive ones.[1-7] A substrate is ligated to each DNA sequence in the pool. Substrate modification catalyzed by the attached DNA sequence leads to a change in mass and subsequent PAGE shift. However, phosphorylation of a peptide substrate results in the addition of only a small phosphoryl group to a tyrosine, serine, or threonine side chain. The relatively small addition of mass and charge from the phosphoryl group added to the peptide substrate connected to the catalytic DNA sequence does not lead to a significant PAGE shift. Therefore, to isolate the active DNA sequences, a method is required to separate phosphopeptide products from unphosphorylated substrates.

2.1.1.1 γ-Thiophosphoryl Donors

The use of a γ-thiophosphoryl donor results in the addition of a thiophosphate to the hydroxyl-containing amino acid side chain within the peptide substrate. The thiophosphorylated product can then be separated from unreacted substrates on a bilayered polyacrylamide gel,

[†] This research has been published:

Walsh, S. M.; Sachdeva, A.; Silverman, S. K. DNA Catalysts with Tyrosine Kinase Activity. *J. Am. Chem. Soc.* **2013**, *135*, 14928-14931.

University of Illinois graduate student Amit Sachdeva performed initial efforts to identify kinase deoxyribozymes using γ-thiophosphoryl donors, developed the 8VP1 capture deoxyribozyme to specifically react with phosphopeptides, and performed the initial XA-XF selections to identify kinase deoxyribozymes using the 8VP1 capture deoxyribozyme.[8]

University of Illinois graduate student Victor Dokukin performed selection experiments to identify modular tyrosine kinase deoxyribozymes containing an aptamer domain for ATP binding.[12]

where the lower layer contains (*N*-acryloylamino)phenyl mercuric chloride (APM). Selection experiments were performed by Amit Sachdeva using a 5′-thiotriphosphorylated RNA oligonucleotide or 5 mM GTPγS as the phosphoryl donor.[8] Activity was observed after 5 rounds in the selection experiments for phosphorylation of 2′-hydroxyls of RNA and the 3′-hydroxyl of DNA. Both selections used 5 mM GTPγS as the phosphoryl donor. Surprisingly, the analogous selection experiments using a Watson-Crick base-paired 5′-thiotriphosphorylated RNA oligonucleotide as the phosphoryl donor were unsuccessful. The base-paired 5′-thiotriphosphorylated RNA phosphoryl donor was thought to be more likely to lead to the identification of DNA catalysts because of the preprogrammed binding to the complementary binding arm next to the initially random region.

Since the selection experiments with the 5′-thiotriphosphorylated RNA phosphoryl donor were unsuccessful, the stability of the phosphoryl donor was evaluated in the selection conditions.[8] The 5′-thiotriphosphorylated RNA oligonucleotide was incubated in selection conditions of 50 mM HEPES, pH 7.5, 20 mM $MnCl_2$, 40 mM $MgCl_2$, 150 mM NaCl, and 2 mM KCl at 37 °C for 12 h. Matrix-assisted laser desorption ionization (MALDI) mass spectrometry analysis of the 5′-thiotriphosphorylated RNA oligonucleotide before and after incubation indicated that the γ-thiotriphosphate hydrolyzed completely, yielding 5′-diphosphorylated RNA. However, when the same experiment was performed with a 5′-triphosphorylated RNA oligonucleotide, only minor hydrolysis to 5′-diphosphorylated RNA was observed. Therefore, the instability of the 5′-thiotriphosphorylated RNA is due to the sulfur in the γ-phosphate of the thiotriphosphate.

In selection experiments with the 5′-thiotriphosphorylated RNA phosphoryl donor, each deoxyribozyme was bound to one phosphoryl donor via Watson-Crick base pairs between the donor and the binding arm adjacent to the deoxyribozyme initially random region. The tightly base-paired thiophosphoryl oligonucleotide was unlikely to exchange with another thiophosphoryl oligonucleotide. Therefore, once the thiophosphoryl donor degraded the bound DNA sequence no longer had a phosphoryl donor available. The inability to exchange and the

rapid degradation of the thiophosphoryl oligonucleotide likely cause the failure of the selection experiments thought to be easier. In contrast, selection experiments with the large concentration of GTPγS phosphoryl donor, the GTPγS molecule bound to the deoxyribozyme was more likely to exchange with others in solution. Despite the increased challenge in binding the phosphoryl donor, the ability to exchange degraded thiophosphoryl donors led to the enrichment of enough active DNA sequences to dominate the pool. Based on their instability, the use of γ-thiophosphoryl donors was discontinued.

2.1.1.2 Development of Phosphotyrosine and Phosphoserine Capture Deoxyribozymes

In many previous in vitro selection experiments, DNA-catalyzed reaction led to a significant mass change, and therefore a PAGE shift.[1-5] To select for DNA catalysis of the desired reaction or at the desired reaction site, some selection experiments have included a subsequent capture step to catalyze a reaction of the desired product and add mass to enable a unique PAGE shift.[6,7] One capture approach uses a deoxyribozyme to catalyze a reaction specifically with the desired product. A capture deoxyribozyme was first used during in vitro selection experiments to identify phosphatase deoxyribozymes.[6] In these selection experiments, the desired product was a dephosphorylated tyrosine residue within the peptide substrate connected to the DNA sequence that catalyzed its dephosphorylation. To capture the dephosphorylated tyrosine products, a previously identified nucleopeptide-forming deoxyribozyme[5] was used to catalyze the addition of an RNA oligonucleotide to the tyrosine residue. The addition of the RNA oligonucleotide led to PAGE shift of the active DNA sequences, and therefore separation of the active DNA sequences from the inactive DNA sequences.

A similar capture deoxyribozyme was developed to enable PAGE-shift separation of phosphorylated peptide substrates attached to active kinase deoxyribozymes. In vitro selection was performed to identify DNA enzymes that covalently modify phosphorylated amino acid side chains of peptides (Figure 2.1).[9] Specifically, DNA enzymes were sought to catalyze

nucleophilic attack of the phosphotyrosine side chain within a peptide substrate at the α-phosphate of a 5′-triphosphorylated RNA oligonucleotide, resulting in nucleopeptide formation. A capture deoxyribozyme must be specific for phosphate to isolate the phosphotyrosine product connected to the active deoxyribozymes from the inactive DNA sequences.

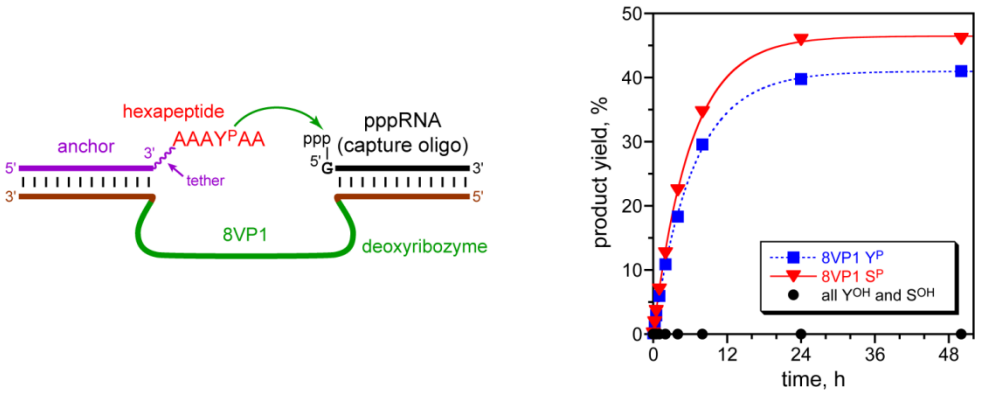


Figure 2.1. DNA-catalyzed covalent modification of phosphopeptide substrates. The 8VP1 deoxyribozyme catalyzes the attachment of a 5′-triphosphorylated RNA oligonucleotide to the phosphorylated amino acid residue within a DNA-anchored hexapeptide substrate. Y^{OH} = Tyr; Y^{P} = pTyr; S^{OH} = Ser; S^{P} = pSer. 8VP1 has k_{obs} of 0.15 h^{-1} with Y^{P}-containing substrates and k_{obs} of 0.17 h^{-1} with S^{P}-containing substrates. No activity was observed with either Y^{OH} or S^{OH}-containing substrates. Figure adapted with permission from Ref. 9.

One deoxyribozyme identified from the selection experiment was 8VP1. The 8VP1 deoxyribozyme has a yield of 40% in 24 h with the AAAY^{P}AA peptide substrate used during the selection. 8VP1 also catalyzes reaction of the analogous phosphoserine-containing peptide, AAAS^{P}AA, with a yield of 46% in 24 h. Nucleic acid enzymes typically do not go to completion at 100% due to misfolding of some of the DNA sequences into inactive forms. This deoxyribozyme is highly selective for phosphorylated peptides; no activity is observed with the nonphosphorylated peptides AAAYAA and AAASAA. 8VP1 will lead to separation of the phosphopeptide products connected to active DNA sequences from the inactive DNA sequences. Therefore, based on its activity and selectivity, 8VP1 can be used as a capture deoxyribozyme during in vitro selection experiments for kinase deoxyribozymes.

2.1.1.3 Initial Kinase Deoxyribozyme Selection Experiments Using Capture Strategy

All in vitro selection experiments are given alphanumeric codes in the following pattern: A-Z, AA-ZZ, AA1-ZZ1, etc. The XA-XF selection experiments were performed to identify tyrosine and serine kinase deoxyribozymes using a 5′-triphosphorylated RNA oligonucleotide (pppRNA) or 5 mM GTP as the phosphoryl donor.[8] After the selection step in which the active DNA sequences catalyzed phosphorylation of the peptide substrate, the sample was purified on a spin column with a 30 kDa mass cutoff to remove salts and the pppRNA or GTP phosphoryl donor. Next, the capture step was performed in which the active DNA sequences connected to their phosphopeptide products were covalently modified by the 8VP1-catalyzed reaction, resulting in PAGE shift.

Activity was observed by round 10 for most selection experiments. However, phosphorylation of the substrate was not observed when the active pools were assayed in trans (i.e. without ligating the DNA-anchored peptide substrate to the DNA pool). The selection activity was also observed if the capture step was not implemented. Therefore, the DNA sequences enriched after the capture step were not the desired kinase deoxyribozymes. The selection method needed to be modified to prevent the enrichment of these undesired, aberrantly migrating DNA sequences.

2.2 Results and Discussion

2.2.1 In Vitro Selection Strategy

In the previous unsuccessful selection experiments to identify kinase deoxyribozymes, in vitro selection led to the enrichment of inactive DNA sequences that aberrantly migrated on the gel at the same position as the desired active DNA enzymes. The aberrant bands observed during capture were also observed after the selection step even if the capture reaction was not performed (Figure 2.2). When the desired selection product was excised and reanalyzed by PAGE, the aberrantly migrating material was successfully removed. Therefore, PAGE purification between

the selection and capture steps was incorporated into the next selection efforts to prevent enrichment of the aberrantly migrating DNA sequences.

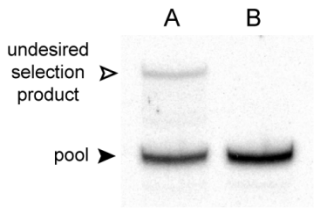


Figure 2.2. Aberrant migration observed in round 9 of the XA selection to identify tyrosine kinase deoxyribozymes using pppRNA as the phosphoryl donor. In the PAGE image, lane A is the product obtained after the selection step (without capture), showing the undesired material migrating higher in the lane. The undesired product migrates in the same position on the gel as the desired product would after the subsequent capture reaction. Lane B shows the selection product after PAGE purification. The undesired selection products are removed by PAGE purification of the pool after the selection step.

The PAGE purification between the selection and capture steps also removes undesired DNA sequences that catalyze nucleopeptide formation (i.e. reaction of the hydroxyl group of the amino acid side chain with the α-phosphate of pppRNA). Nucleopeptide formation has previously been catalyzed by DNA.[2,4,5] Additionally, the length and sequence of the pppRNA capture oligonucleotide was alternated to prevent aberrantly migrating bands from surviving the capture step. The different length of the pppRNA capture oligonucleotide (either 17 nt or 36 nt) results in migration of the capture product to different locations on the capture gel. Thus, the same aberrantly migrating DNA sequences would not be excised in consecutive selection rounds. Changing the sequence of the pppRNA capture oligonucleotide in alternate rounds prevents the DNA sequences that can bind to and react with one capture oligonucleotide from surviving the alternate rounds that uses a different capture oligonucleotide.

With a way to remove the aberrantly migrating DNA sequences that resulted in the previous selection efforts, a modified in vitro selection method was developed to identify kinase deoxyribozymes. In order to isolate active DNA enzymes, the phosphoryl acceptor needs to be connected to the DNA sequence that catalyzed the phosphoryl transfer reaction. The hexapeptide

substrate containing a tyrosine flanked by alanine residues with an N-terminal cysteine (CAAYAA) was synthesized by solid-phase peptide synthesis. The peptide was connected to a DNA-anchor via a disulfide bond and a hexa(ethylene glycol) [HEG] tether. This DNA-anchored peptide substrate was ligated to the initially random DNA pool using T4 DNA ligase. The pppRNA phosphoryl donor was prepared by in vitro transcription using a DNA template and T7 RNA polymerase. The initially random region of the DNA sequences was flanked by fixed-sequence regions to enable PCR amplification of the DNA sequences and substrate binding.

The in vitro selection began with ligation of the initially random DNA pool to the DNA-anchored peptide substrate (Figure 2.3). The ligation product was isolated by PAGE. Next, the selection step was performed in which the DNA sequences were incubated with the phosphoryl donor and divalent metal ions in a buffered solution to enable DNA catalysis. PAGE purification was then performed to remove undesired products and reagents. To isolate the active DNA sequences connected to the desired phosphopeptide products, the capture step was performed to add additional mass, thus providing a PAGE shift. The previously identified deoxyribozyme, 8VP1, was used specifically to add mass to phosphopeptides.[9] PCR was then performed to amplify the catalytic DNA sequences that survived the previous selection and capture steps. The entire selection round was repeated, each time enriching the population with catalytically active sequences, until the catalytically active sequences dominated the population.

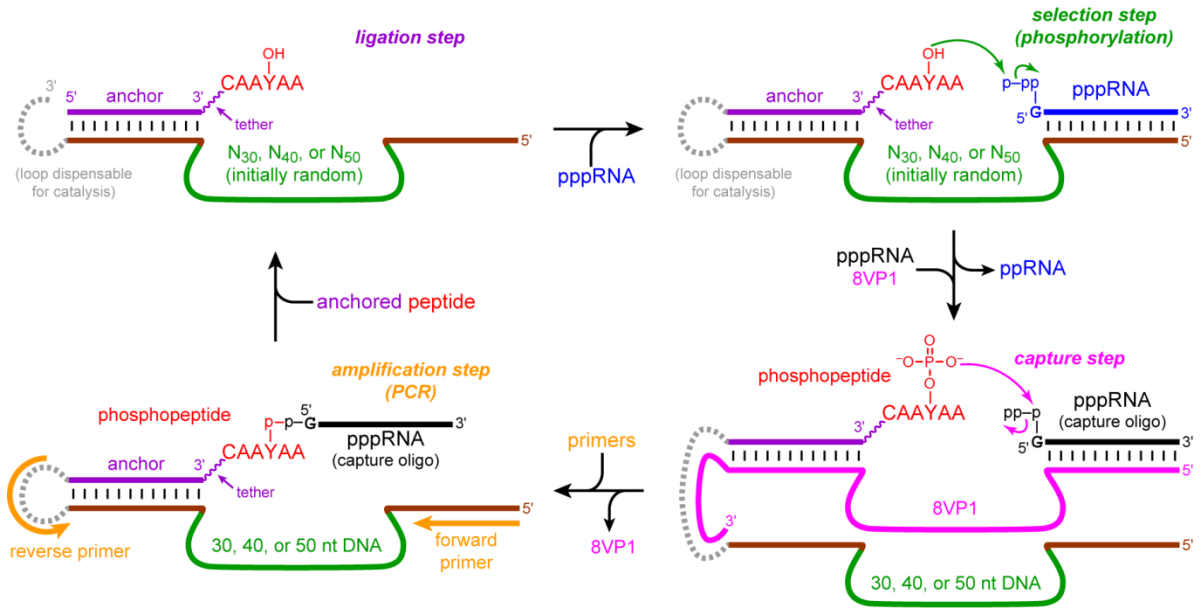


Figure 2.3. Diagram of in vitro selection procedure for kinase deoxyribozymes using the 8VP1 capture deoxyribozyme method. First, the DNA-anchored peptide substrate was ligated to the DNA sequences in the initially random pool. Next, the DNA sequences were incubated with the pppRNA phosphoryl donor and divalent metal ions to enable the DNA enzymes to phosphorylate the peptide substrate. In the subsequent capture step, the 8VP1 capture deoxyribozyme selectively added mass to the phosphopeptide products connected to the active DNA sequences. The surviving catalytic DNA sequences were then PCR-amplified. PAGE purification was performed between each step. The selection process was repeated until the catalytically active DNA sequences dominated the population. Figure adapted with permission from Ref. 11.

New selection experiments AA1-AF1 for kinase deoxyribozymes were performed with the modified selection procedure, including PAGE purification between the selection and capture steps. Selection conditions successful for identifying nucleopeptide-forming deoxyribozymes were used for all selections: 50 mM HEPES, pH 7.5, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h. After six rounds of selection, smeary bands were observed migrating near the capture standards (Figure 2.4A). The smeary bands were also observed if the phosphoryl donor required for substrate phosphorylation was not included in the selection step, indicating the smeary bands were not the desired phosphorylated product. The smearing was likely the result of the enrichment of DNA sequences that tightly interacted with the 8VP1 deoxyribozyme and did not quickly dissociate upon entering the gel, despite the presence of 7 M urea as a denaturant. A decoy oligonucleotide that is complementary to the 8VP1 capture

deoxyribozyme base-pairs to the capture deoxyribozyme and displaces the interactions between 8VP1 and the DNA pool. When the decoy oligonucleotide was added to the capture sample immediately prior to loading on the gel, the undesired smears were removed.

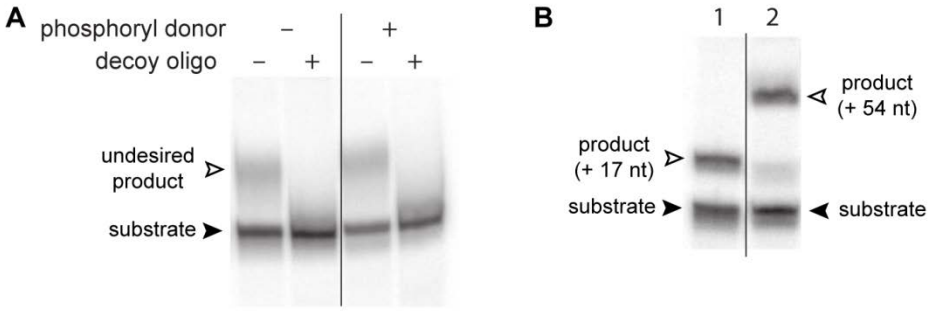


Figure 2.4. Modification to the in vitro selection strategy to identify kinase deoxyribozymes. (A) Decoy oligonucleotide. PAGE image shows the undesired smeary product observed migrating higher than the substrate on the capture gel. The smeary band is observed regardless of whether a phosphoryl donor is included during the selection step, indicating the smear is not related to the desired phosphorylation product. The undesired smear migrates near the position of the desired capture products. Upon addition of a decoy oligonucleotide that is complementary to the 8VP1 capture deoxyribozyme prior to loading the sample on the gel, the undesired products are not observed. (B) Capture oligonucleotide length. To prevent the excision of overlapping regions of the gel during capture, alternation of a 17 nt (lane 1) and 54 nt (lane 2) capture oligonucleotide provides a substantially different PAGE shift of the products.

In addition to adding decoy oligonucleotides, the difference in length between the alternating capture oligonucleotides was increased to prevent overlap of excised regions of the gel in consecutive rounds (Figure 2.4B). Previously, a 17 nt and 36 nt capture oligonucleotide were used in alternating rounds. Increasing the length of the long capture oligonucleotide to 54 nt provided a larger difference in gel migration between the two capture products. The 54 nt capture oligonucleotide is an pppRNA-DNA chimera made by ligating a 17 nt pppRNA oligonucleotide to a 37 nt DNA oligonucleotide using a DNA splint and T4 DNA ligase.

The in vitro selection procedure was modified to include the addition of the 8VP1 decoy oligonucleotide prior to loading the capture sample on the gel and the pppRNA capture oligonucleotide length was alternated between 17 nt and 54 nt. New selection experiments (AY1-BD1) to identify kinase deoxyribozymes were initiated using the same selection conditions

successful for identifying nucleopeptide-forming deoxyribozymes: 50 mM HEPES, pH 7.5, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h. Unfortunately, after 14 rounds of selection no activity was observed in any of the selection experiments, and they were discontinued.

2.2.2 Identification and Characterization of Tyrosine Kinase Deoxyribozymes

New kinase deoxyribozyme in vitro selection experiments CF1-CK1 were designed for tyrosine phosphorylation using pppRNA as the phosphoryl donor with N_{30}, N_{40}, or N_{50} random region lengths with selection conditions (*A*) 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h or (*B*) 50 mM CHES, pH 9.0, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h (Figure 2.5). Multiple initially random region lengths were evaluated because random region length has been shown to be an important variable in previous deoxyribozyme selection experiments.[10] Selection condition *A* is similar to the conditions previously used in kinase deoxyribozyme selections, but also include Zn^{2+} as a divalent metal ion cofactor because Zn^{2+} is required for both DNA-hydrolyzing and peptide-dephosphorylating deoxyribozymes.[3,6] Selection condition *B* include a higher pH, and Mn^{2+} and Zn^{2+} are excluded at this pH because they will oxidize and precipitate, respectively.

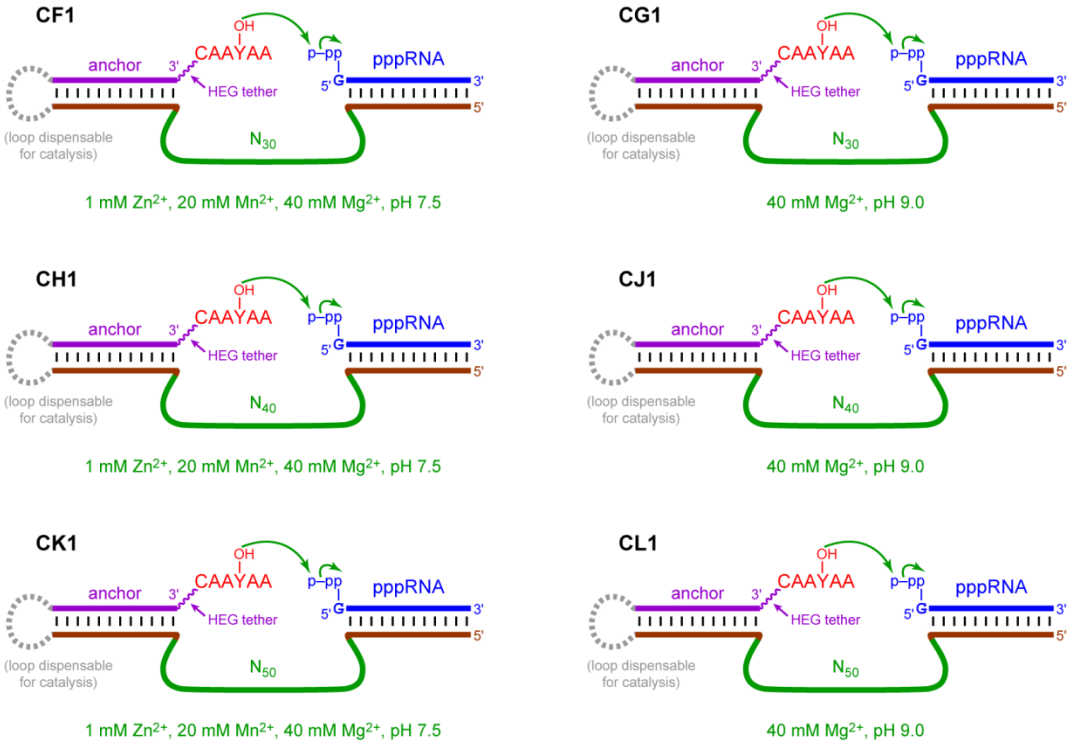


Figure 2.5. Design for the CF1-CL1 in vitro selection experiments to identify tyrosine kinase deoxyribozymes. In all selections, the peptide substrate CAAYAA was connected to a DNA anchor via a disulfide bond and HEG tether, and the phosphoryl donor was pppRNA. CF1-CG1, CH1-CJ1, and CK1-CL1 have N_{30}, N_{40}, and N_{50} initially random regions, respectively. The CF1, CH1, and CK1 selections were performed in condition *A*, and the CG1, CJ1, and CL1 selections were performed in condition *B*.

The CF1, CH1, and CK1 selections led to phosphorylation activity in condition *A* with all three initially random region lengths (Figure 2.6), but no activity was observed for the selections performed in condition *B*.[11] To identify individual deoxyribozyme sequences, active pools were cloned from round 6, 7, and 8 for N_{30}, N_{40}, and N_{50}, respectively. From the N_{30}, N_{40}, and N_{50} selection experiments five, three, and nine unique sequences were identified (Figure 2.7). Sequence alignment revealed a conserved 11-nucleotide segment at the 5′-end of each deoxyribozyme, despite the different lengths of the initially random region. This conserved region is near the pppRNA phosphoryl donor binding site. Changes to the pppRNA sequence or replacement with GTP were not tolerated by the deoxyribozymes.

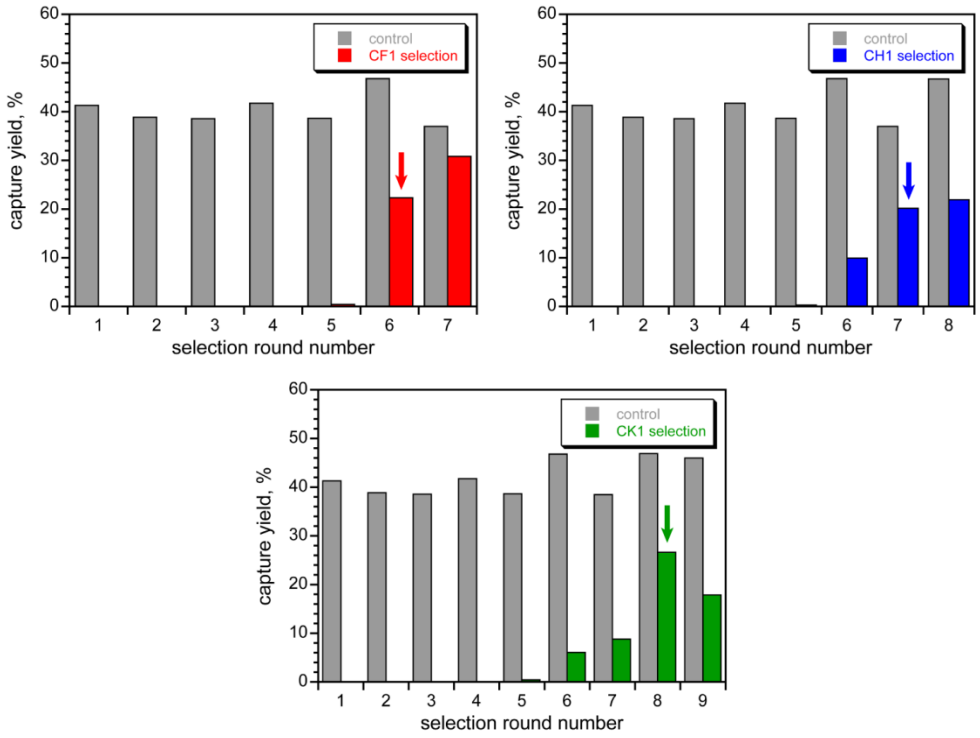


Figure 2.6. Selection progression of the in vitro selection experiments for kinase deoxyribozymes that use pppRNA as the phosphoryl donor. In each round, "control" refers to the yield for the 8VP1-catalyzed capture reaction using a CAAY^{P}AA substrate, and "selection" refers to the yield for the 8VP1-catalyzed capture reaction for the indicated selection experiment. Arrows mark the cloned rounds. The CF1, CH1, and CK1 selections have N_{30}, N_{40}, and N_{50} random region, respectively. Figure adapted with permission from Ref. 11.

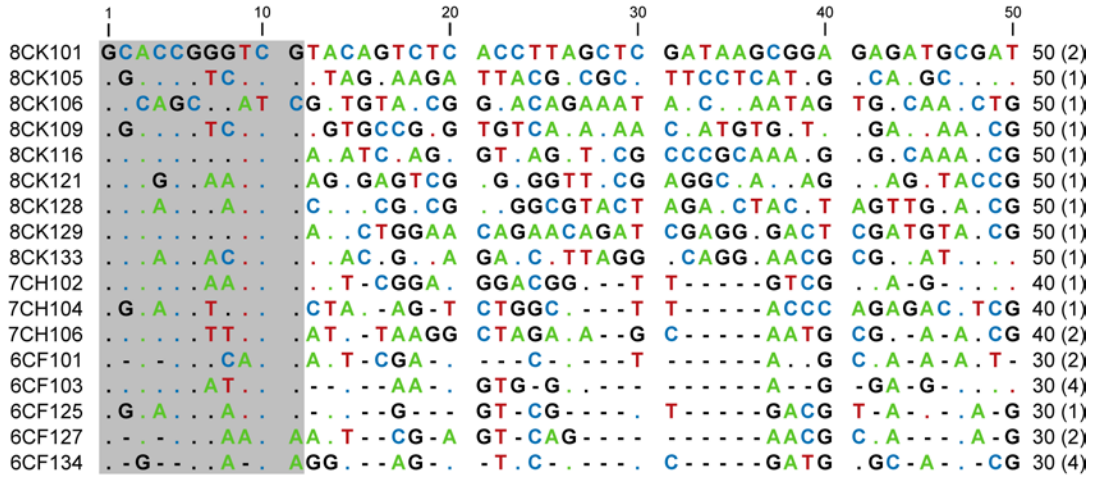
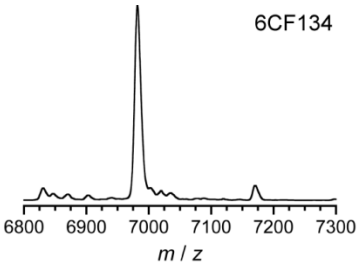


Figure 2.7. Sequences of the initially random region of the tyrosine kinase deoxyribozymes using pppRNA as the phosphoryl donor. Only the initially random region is shown. A dot denotes conservation with the uppermost sequence, and a dash denotes a gap. Next to the sequence length on the right, in parentheses, is the number of times each sequence was found. A strongly conserved 11 nt region is observed at the 5′-end of the sequences, indicated by the grey box, regardless of the deoxyribozyme random region length. Figure reprinted with permission from Ref. 11.

Tyrosine phosphorylation was verified (Figure 2.8). MALDI mass spectrometry of the product confirmed the addition of a phosphoryl group to the DNA-anchored peptide substrate. Additionally, the phosphorylation products were PAGE-purified and subjected to 8VP1-catalyzed reaction with the pppRNA capture oligonucleotide. The same rate and yield of 8VP1 product formation was observed with the CAAY^{P}AA positive control and the phosphorylation products generated by the deoxyribozymes. Therefore, the DNA enzymes catalyze the phosphorylation of tyrosine within the peptide substrate.



deoxyribozyme	[M+H]⁺ calcd.	[M+H]⁺ found	error, % (found – calcd.)
6CF127	6982.9	6982.6	–0.004
6CF134	6982.9	6982.0	–0.01
7CH102	6982.9	6981.7	–0.02
7CH104	6982.9	6984.8	+0.03
8CK105	6982.9	6982.7	–0.003
8CK128	6982.9	6981.4	–0.02

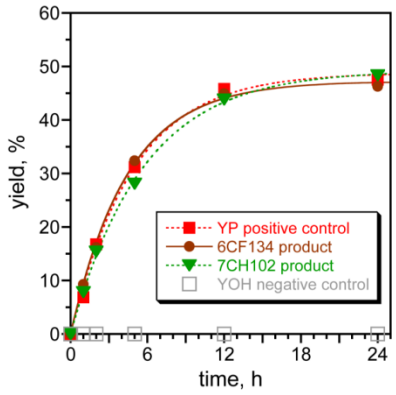


Figure 2.8. MALDI mass spectrometry and 8VP1 reactivity for tyrosine kinase deoxyribozymes. The mass is that of the DNA-HEG-CAAY^{P}AA product. 8VP1-catalyzed reaction of the phosphopeptide product has the same rate and yield as the Y^{P} positive control. This confirms that the phosphoryl group is attached at the tyrosine hydroxyl within the peptide substrate. Figure adapted with permission from Ref. 11.

The 6CF1 (N_{30}), 7CH1 (N_{40}), and 8CK1 (N_{50}) kinase deoxyribozymes were characterized. The 6CF134 deoxyribozyme has k_{obs} of 1.54 ± 0.22 h^{-1} (n = 7) and 71% yield in 12 h (Figure 2.9). All other kinase deoxyribozymes have k_{obs} of 0.1-0.5 h^{-1} and 11-70% yield in 24 h. All kinase deoxyribozymes were identified by using 1 mM Zn^{2+}, 20 mM Mn^{2+}, and 40 mM Mg^{2+} during the selection step. Individual deoxyribozymes have slightly different Zn^{2+} concentration dependencies, but in all cases 0.5 mM was highly effective and was used for characterization. Phosphorylation activity was evaluated in the presence of all combinations of the three metal ions (Figure 2.10). Both Zn^{2+} and Mn^{2+} were required for catalysis, except in the case of 6CF134, which required only Zn^{2+}. In all cases Mg^{2+} was dispensable for catalysis.

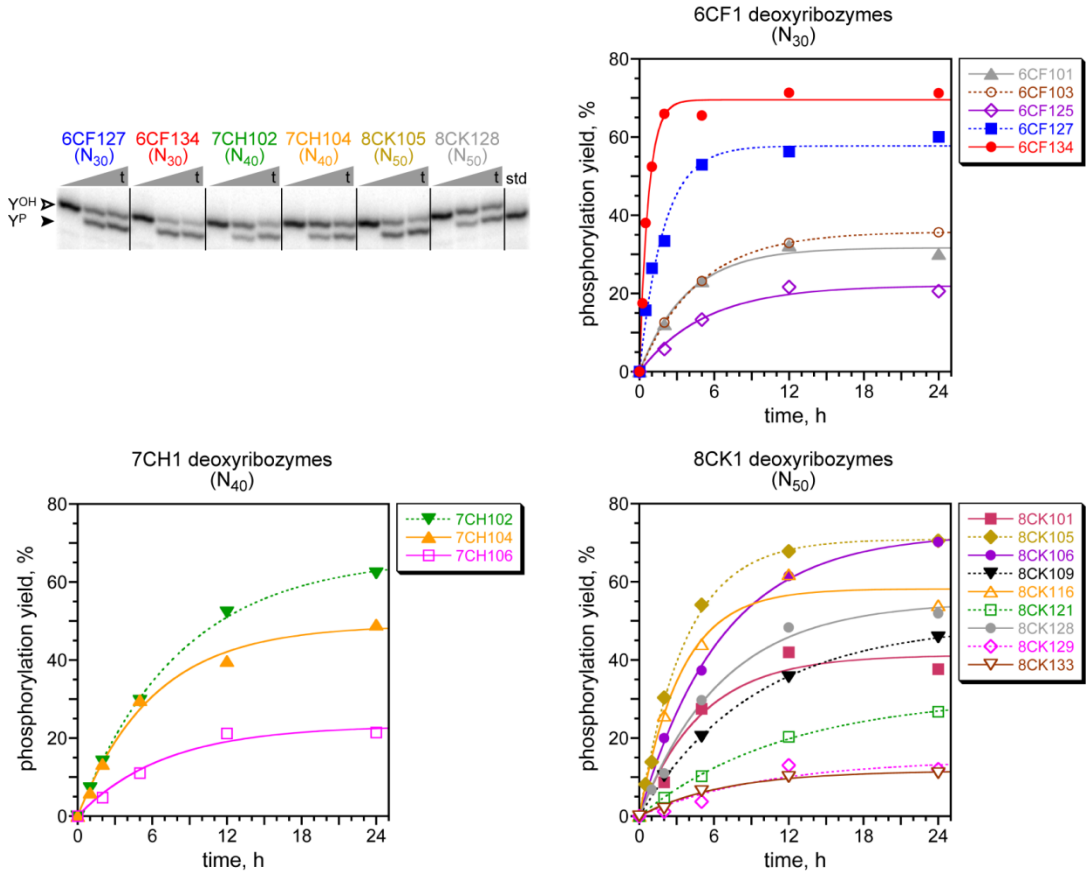


Figure 2.9. Activity of the tyrosine kinase deoxyribozymes that use pppRNA. The PAGE image shows representative time points (t = 30 s, 5 h, 24 h) for six of the deoxyribozymes as indicated. Y^{OH} = substrate; Y^{P} = product. The kinetic plots for all of the identified deoxyribozymes are shown. All assays were performed in trans. Incubation conditions were 70 mM HEPES, pH 7.5, 0.5 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C. Figure adapted with permission from Ref. 11.

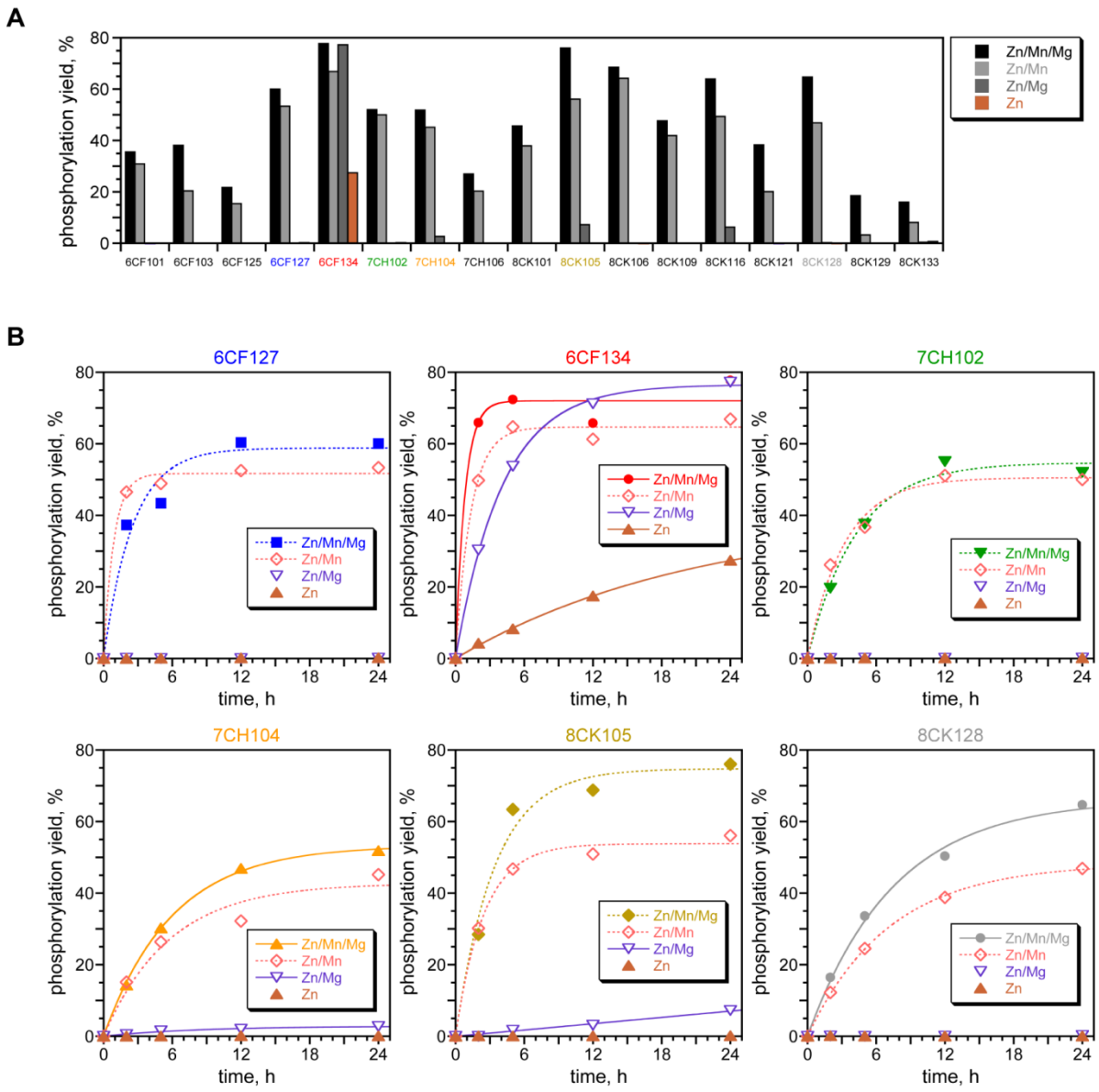


Figure 2.10. Metal dependence of the tyrosine kinase deoxyribozymes. (A) Data at t = 24 h for all deoxyribozymes. Incubation conditions are 70 mM HEPES, pH 7.5, 150 mM NaCl and combinations of 0.5 mM $ZnCl_2$, 20 mM $MnCl_2$, and 40 mM $MgCl_2$ as indicated at 37 °C. The metal ion combinations not shown did not have activity for any deoxyribozyme. (B) Kinetic plots for six representative deoxyribozymes with the indicated metal ion conditions. Figure reprinted with permission from Ref. 11.

Deoxyribozymes were assayed for their peptide sequence, peptide length, and tether length dependence. All of the kinase deoxyribozymes were inactive with serine in place of tyrosine within the peptide substrate and thus specifically are tyrosine kinase deoxyribozymes. The deoxyribozymes were tolerant to changes in the amino acid residues flanking the tyrosine

(Figure 2.11A). Peptide substrates with a hydrophobic (phenylalanine, Phe), negatively charged (glutamate, Glu), or positively charged (lysine, Lys) residue on either side of the tyrosine residue were phosphorylated by the six deoxyribozymes. Additionally, the peptide substrate GPYSGN containing a tyrosine residue surrounded by non-alanine residues was successfully phosphorylated. Therefore, all of the kinase deoxyribozymes are peptide sequence-general.

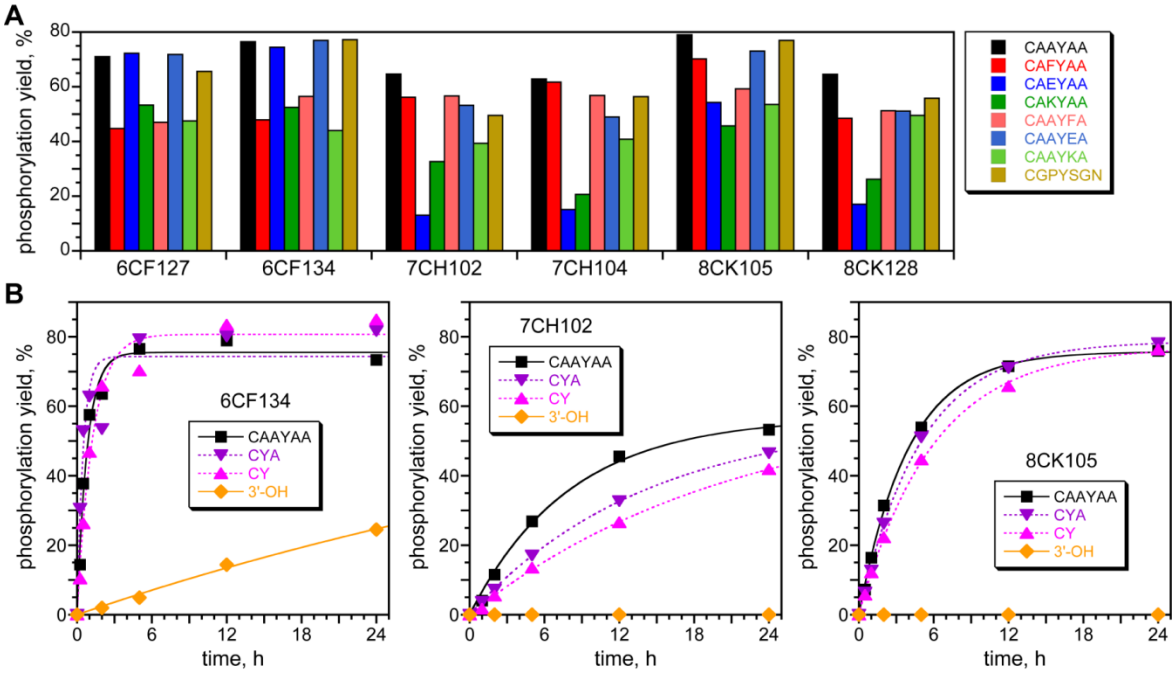


Figure 2.11. Dependence of deoxyribozyme activity on peptide sequence and length. (A) Peptide sequence dependence. The indicated peptide substrate was connected via a disulfide bond and HEG tether to the DNA anchor. Data are shown for six representative deoxyribozymes (t = 24 h). All deoxyribozymes tolerate the incorporation of Phe (F), Glu (E), or Lys (K) at either position flanking the reactive Tyr (Y) residue. The substrate GPYSGN was evaluated as a peptide sequence with Tyr surrounded by non-alanine residues. (B) Peptide length dependence. The substrate included the indicated peptide connected via a HEG tether or only a DNA 3′-hydroxyl group for phosphorylation. Kinetic plots are for three representative deoxyribozymes. Figure reprinted with permission from Ref. 11.

Decreasing the length of the CAAYAA hexapeptide to a shorter CYA tripeptide or CY dipeptide did not affect deoxyribozyme activity (Figure 2.11B). Only 6CF134 had some phosphorylation activity with a DNA 3′-hydroxyl group. All deoxyribozymes maintained activity with a shorter C_3 tether in place of the HEG tether used during the selections (Figure 2.12).

However, no phosphorylation activity was observed with a longer HEG_{10} tether or an untethered, free peptide in solution.

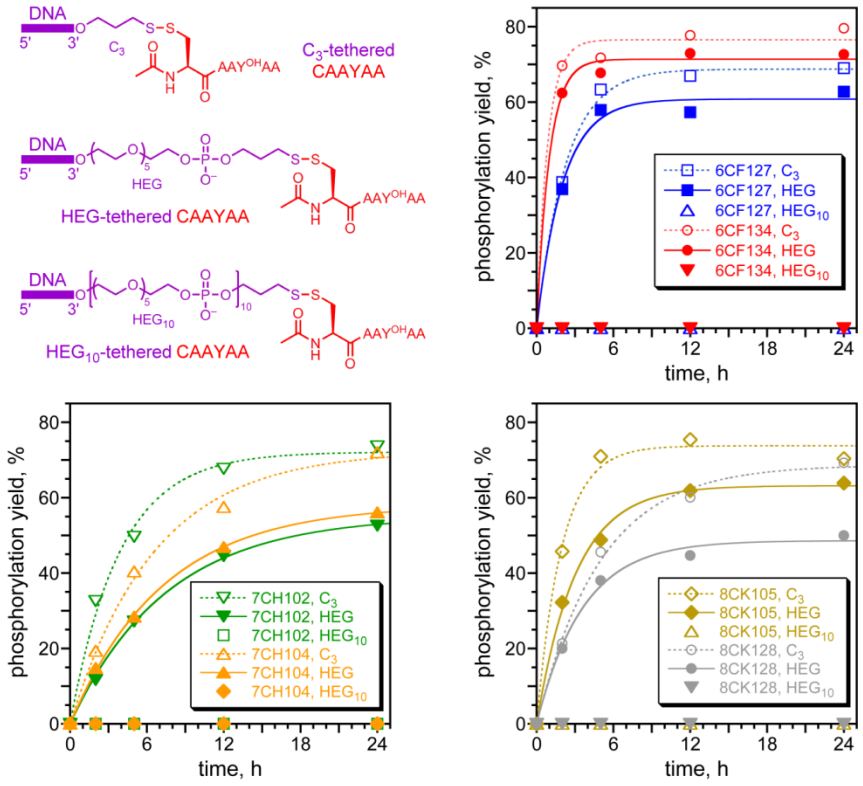


Figure 2.12. Dependence of deoxyribozyme activity on the length of the tether between the DNA anchor and CAAYAA hexapeptide substrate. The structures of the three different tether lengths are shown. The HEG tether was used during the selection experiments. Here, a shorter C_3 and a longer HEG_{10} tether were evaluated. Kinetic plots are for six representative deoxyribozymes. Figure reprinted with permission from Ref. 11.

Reselection of the 6CF134 deoxyribozyme was performed in the EE1 selection by generating a partially randomized pool (25% randomization). The constant binding arm sequences flanking the initially random region were changed to prevent contamination of the reselection experiment from the parent deoxyribozyme. Activity was observed in the reselection and was cloned from round 7 with 17% activity (Figure 2.13A). To compare the parent 6CF134 deoxyribozyme to the reselected variants, all deoxyribozymes were synthesized with both the reselection and 6CF134 parent binding arms. Unfortunately, the change in the binding arm

sequences significantly decreased the activity of the parent deoxyribozyme, 6CF134 (Figure 2.13B). Therefore, the subsequent improvement in activity observed with the reselected variants is difficult to evaluate. In future reselection efforts, deoxyribozyme tolerance of binding arm sequence changes will be evaluated prior to changing the binding arm sequences for reselection experiments.

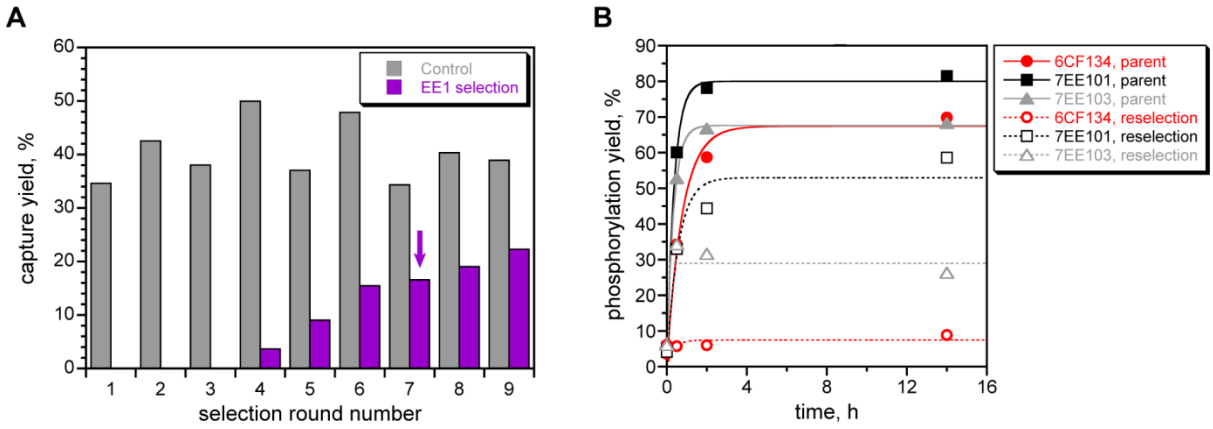


Figure 2.13. EE1 selection experiment, reselection of the 6CF134 deoxyribozyme. (A) Selection progression of the EE1 reselection for improved kinase deoxyribozymes that use pppRNA as the phosphoryl donor. In each round, "control" refers to the yield for the 8VP1-catalyzed capture reaction using a CAAY^{P}AA substrate, and "selection" refers to the yield for the 8VP1-catalyzed capture reaction for the EE1 selection experiment. The arrow marks the cloned round. (B) A kinetic plot for the identified deoxyribozymes is shown. For each deoxyribozyme, "parent" refers to the binding arms used in the initial CF1 selection, and "reselection" refers to the binding arms used during the EE1 selection. All assays were performed in trans. Incubation conditions were 70 mM HEPES, pH 7.5, 0.5 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C.

2.2.3 Tyrosine Kinase Deoxyribozymes that use GTP Phosphoryl Donors

Natural kinases use ATP or in some cases GTP as their phosphoryl donor. Kinase deoxyribozymes would be more easily utilized to phosphorylate proteins if small-molecule NTP donors could be used instead of pppRNA oligonucleotides. Kinase deoxyribozyme selection experiments EA1-EC1 were performed using GTP as the phosphoryl donor (Figure 2.14) instead of pppRNA as used by the previous kinase deoxyribozymes. Successful tyrosine kinase deoxyribozymes need to bind to the phosphoryl donor and catalyze phosphorylation. Whereas the previously used pppRNA phosphoryl donor was bound to the deoxyribozyme binding arm

via preprogrammed Watson-Crick binding interactions. Selection experiments for tyrosine phosphorylation were performed with N_{30}, N_{40}, or N_{50} random region lengths and 1 mM GTP instead of pppRNA as the phosphoryl donor. Condition *A* was used for all selection experiments.

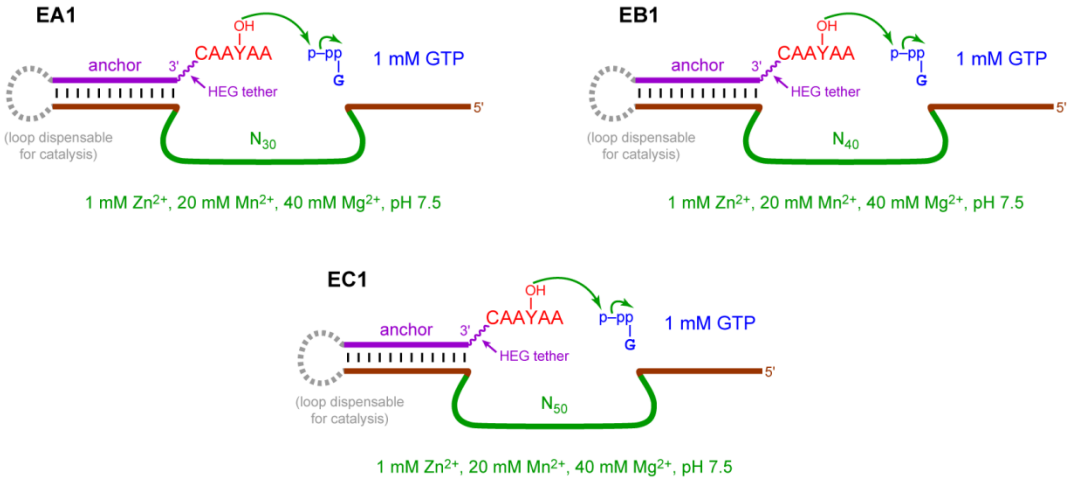


Figure 2.14. Design for the EA1-EC1 in vitro selection experiments to identify tyrosine kinase deoxyribozymes using small-molecule NTP donors. In all selections, the peptide substrate was CAAYAA connected to a DNA anchor via a disulfide bond and HEG tether, and 1 mM GTP was the phosphoryl donor. EA1, EB1, and EC1 have N_{30}, N_{40}, and N_{50} initially random regions, respectively. All selections were performed in selection condition *A*.

Activity was observed in all three selection experiments (Figure 2.15), and a single active DNA sequence was identified from each selection (Figure 2.16). 8EA101 (N_{30}) and 16EC103 (N_{50}) have substantial catalytic activity (Figure 2.17). 8EA101 has $k_{obs} = 0.2$ h^{-1} with 44% phosphorylation yield in 24 h, and 16EC103 has $k_{obs} = 0.2$ h^{-1} with 34% yield in 24 h. 21EB121 (N_{40}) has a modest 5% phosphorylation yield and was not further characterized. Phosphorylation of the tyrosine within the peptide substrate was confirmed by MALDI mass spectrometry and 8VP1 reactivity, as done with the previously identified tyrosine kinase deoxyribozymes (Figure 2.18).

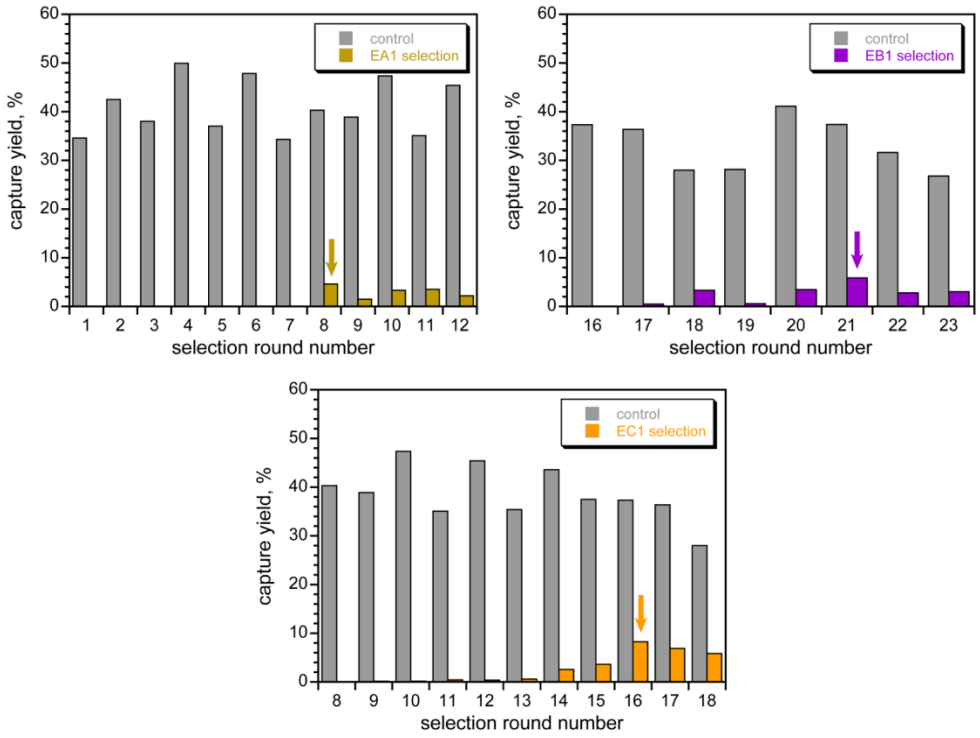


Figure 2.15. Selection progression of the in vitro selection experiments for kinase deoxyribozymes that use GTP as the phosphoryl donor. In each round, "control" refers to the yield for the 8VP1-catalyzed capture reaction using a CAAY^{P}AA substrate, and "selection" refers to the yield for the 8VP1-catalyzed capture reaction for the indicated selection experiment. Arrows mark the cloned rounds. The EA1, EB1, and EC1 selections have a N_{30}, N_{40}, and N_{50} random region, respectively. Figure adapted with permission from Ref. 11.

```
                1         10        20        30        40        50
                |         |         |         |         |         |
16EC103  TGGGAGACGT GTCCAATATG AATAGCGCGC TTCCGAATCA GTTGGACGTT 50 (1)
21EB121  TCAGTCGACT TCGTGTGGCT TTGCGTTTAA AGAGGTAAGC ---------- 40 (1)
8EA101   GTAGGGGGGA CCGTAGCTTC AGCGTGACAG ---------- ---------- 30 (1)
```

Figure 2.16. Sequences of the initially random regions of the tyrosine kinase deoxyribozymes that use GTP as the phosphoryl donor. Only the initially random region is shown. Next to the sequence length on the right, in parentheses, is the number of times each sequence was found. Figure reprinted with permission from Ref. 11.

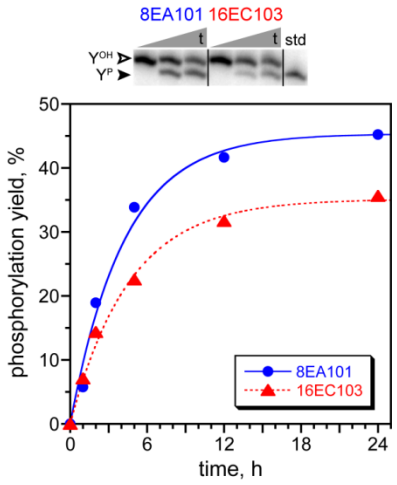
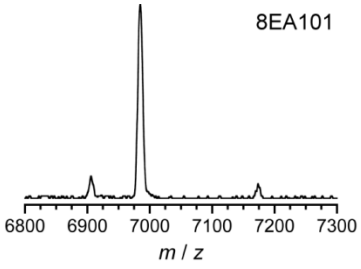


Figure 2.17. Activity of the tyrosine kinase deoxyribozymes 8EA101 (N_{30}) and 16EC103 (N_{50}) that use GTP as the phosphoryl donor. The PAGE image shows representative time points (t = 30 s, 5 h, 24 h). Y^{OH} = substrate; Y^{P} = product. The kinetic plots for the identified deoxyribozymes are shown. All assays were performed in trans. Incubation conditions were 70 mM HEPES, pH 7.5, 100 µM GTP, 0.5 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C. Figure reprinted with permission from Ref. 11.



deoxyribozyme	[M+H]+ calcd.	[M+H]+ found	error, % (found – calcd.)
8EA101	6982.9	6984.7	+0.03
21EB121	6982.9	6982.2	–0.01
16EC103	6982.9	6986.3	+0.05

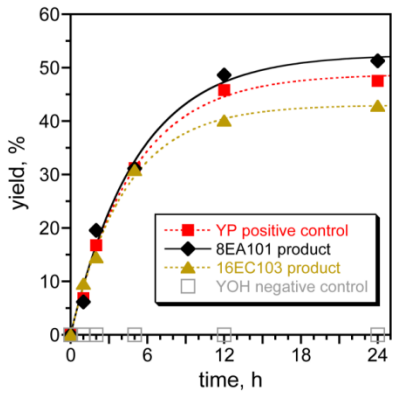


Figure 2.18. MALDI mass spectrometry and 8VP1 reactivity for tyrosine kinase deoxyribozymes that use GTP as the phosphoryl donor. The mass of the DNA-HEG-CAAY^{P}AA product. 8VP1-catalyzed reaction of the phosphopeptide product has the same rate and yield as the Y^{P} positive control. This confirms that the phosphoryl group is attached at the tyrosine hydroxyl within the peptide substrate. Figure adapted with permission from Ref. 11.

The NTP requirement was evaluated for both deoxyribozymes (Figure 2.19). 16EC103 maintains activity with dGTP, but no activity is observed with any other NTP or dNTP, indicating a requirement for guanosine. In contrast, 8EA101 maintains comparable activity with all NTPs, dNTPs, and even inorganic triphosphate as the phosphoryl donor.

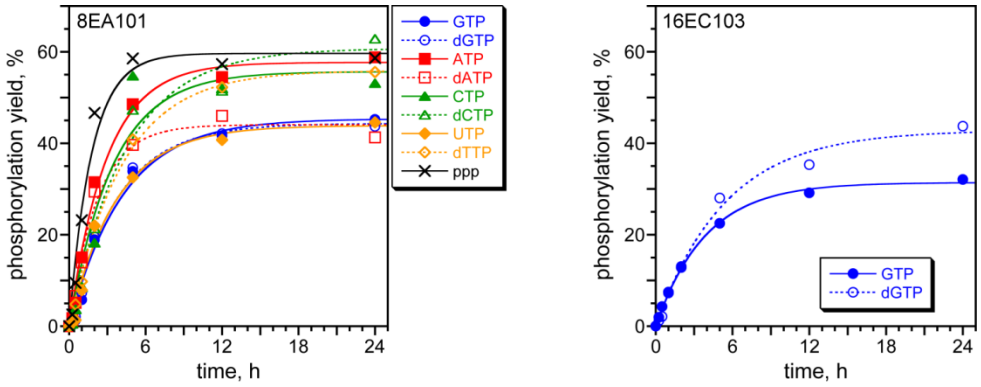


Figure 2.19. NTP requirement of the kinase deoxyribozymes identified using GTP. Both deoxyribozymes were evaluated with all four NTPs and dNTPs. 8EA101 was additionally tested with inorganic triphosphate (ppp). Incubation conditions were 70 mM HEPES, pH 7.5, 100 µM phosphoryl donor, 0.5 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C.

Both 8EA101 and 16EC103 had significant GTP concentration dependence with an apparent K_m(GTP) of ~20 and 42 µM, respectively (Figure 2.20). Substantial inhibition was observed for both deoxyribozymes above 100-200 µM GTP. Despite the increased k_{obs} of 8EA101 with inorganic triphosphate as the phosphoryl donor, the apparent K_m(ppp) was ~19 µM, which is the same as for GTP. These results show that DNA catalysts can use small-molecule NTPs as phosphoryl donors and each deoxyribozyme can interact with the NTP phosphoryl donor in a different way.

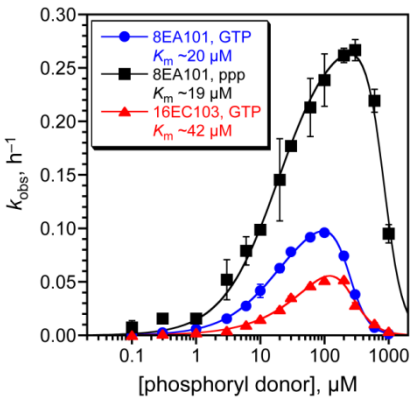


Figure 2.20. Apparent K_m analysis of 8EA101 and 16EC103. The k_{obs} values were determined from initial-rate kinetics (linear fits to plots of yield versus time from 0 to 2 h). The k_{obs} are shown as mean ± standard deviation (n = 3 at each GTP concentration). Data were fit to $k_{obs} = k_{cat} \bullet [C^n/(K_m^n+C^n)] \bullet [1-C^m/(K_i^m+C^m)]$, where k_{cat} is the apparent k_{cat} value, C is the GTP concentration, K_m is the apparent K_m(GTP) value for the productive GTP binding, K_i is the inhibition constant for unproductive GTP binding, and n and m are Hill coefficients for productive and unproductive GTP binding, respectively. For 8EA101, the apparent K_m(GTP) = 20 ± 1 µM and the apparent K_i(GTP) = 240 ± 4 µM. The apparent K_m(ppp) = 19 ± 2 µM and the apparent K_i(ppp) = 841 ± 33 µM. For 16EC103, the apparent K_m(GTP) = 42 ± 7 µM and the apparent K_i = 284 ± 17 µM. Figure adapted with permission from Ref. 11.

2.2.4 Tyrosine Kinase Deoxyribozymes that use ATP Phosphoryl Donors

The EA1, EB1, and EC1 selection experiments used 1 mM GTP as the phosphoryl donor, but the resulting deoxyribozymes are inhibited by GTP concentrations above 100 µM. Additionally, one deoxyribozyme, 8EA101, could use any NTP as the phosphoryl donor. Therefore, new selection experiments HR1-HT1 were designed to identify tyrosine kinase deoxyribozymes using 100 µM ATP during the selection step. The lower NTP concentration seems ideal for kinase deoxyribozymes and ATP is the common biological phosphoryl donor. Selection experiments were performed using N_{30}, N_{40}, and N_{50} random region lengths and condition *A* that were previously found to enable DNA-catalyzed phosphorylation.

Activity was observed for both HR1 (N_{30}) and HS1 (N_{40}) (Figure 2.21A). Selection pressure of 10 µM ATP was implemented at rounds 11 and 9 for HR1b and HS1b, respectively. Activity did not recover after several rounds of selection pressure. Therefore, each selection was cloned from its original population without the ATP concentration pressure applied, at rounds 10 and 8 for HR1 and HS1, respectively. The HR1 and HS1 selections led to the identification of

one and two deoxyribozymes, respectively (Figure 2.21B). No activity was observed for HT1 (N_{50}), and the selection was discontinued after 16 rounds. 10HR114, 8HS119, and 8HS122 have phosphorylation yield of 38%, 47%, and 63%, respectively (Figure 2.21C). None of these deoxyribozymes phosphorylate serine in place of tyrosine.

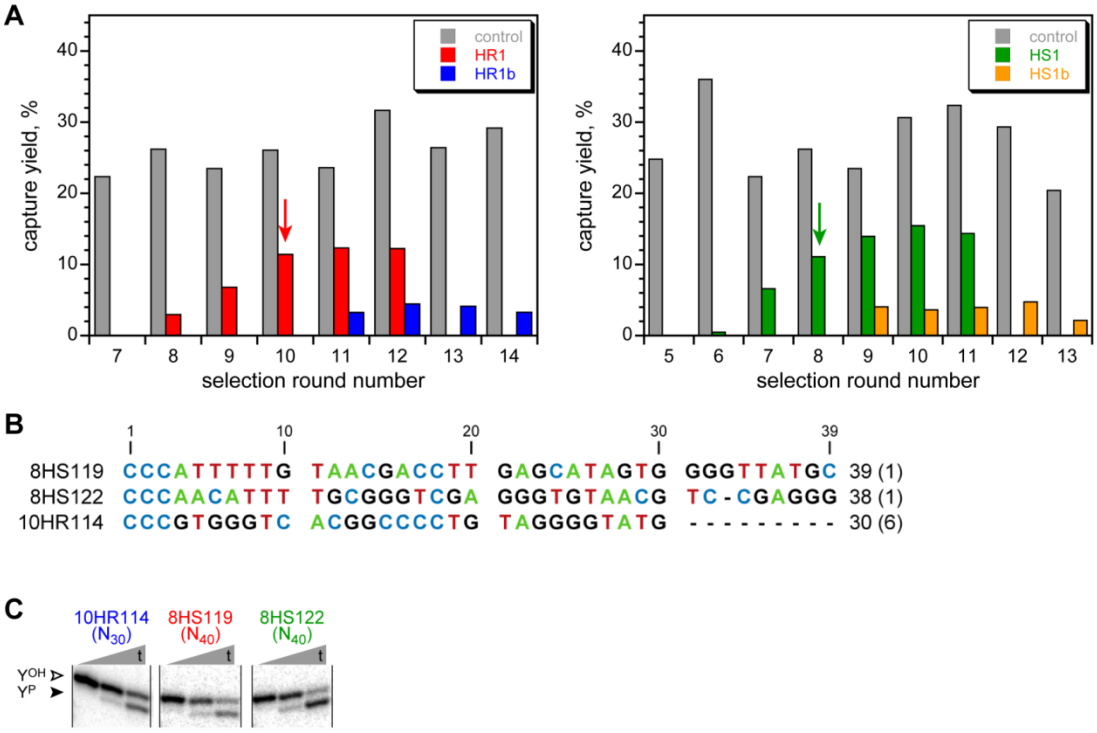


Figure 2.21. Tyrosine kinase deoxyribozymes identified using 100 µM ATP as the phosphoryl donor. (A) Selection progression of the in vitro selection experiments for kinase deoxyribozymes that use ATP as the phosphoryl donor. In each round, "control" refers to the yield for the 8VP1-catalyzed capture reaction using a CAAY^{P}AA substrate, and "selection" refers to the yield for the 8VP1-catalyzed capture reaction for the indicated selection experiment. Arrows mark the cloned rounds. The HR1 and HS1 selections have N_{30} and N_{40} random regions, respectively. HR1b and HS1b are the rounds in which selection pressure was applied by decreasing the ATP concentration to 10 µM. (B) Sequences of the initially random regions. Only the initially random region is shown. Next to the sequence length on the right, in parentheses, is the number of times each sequence was found. (C) The PAGE image shows representative time points (t = 30 s, 2 h, 16 h). Y^{OH} = substrate; Y^{P} = product. Assays were performed in trans. Incubation conditions were 70 mM HEPES, pH 7.5, 100 µM ATP, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C.

Independent selection experiments were performed by Victor Dokukin to evaluate modular deoxyribozymes containing an N_{40} initially random catalytic region and a separate aptamer domain for substrate binding (Figure 2.22).[12] Tyrosine kinase deoxyribozyme selections were performed using the 8VP1 deoxyribozyme capture method. The substrates were a DNA-anchored CAAYAA peptide and 30 μM ATP as the phosphoryl donor. The DNA pool contained an N_{40} initially random region and a separate ATP phosphoryl donor binding site provided by a DNA aptamer region. Selection conditions were 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h. The resulting deoxyribozymes needed to catalyze the phosphoryl transfer reaction by interacting with the bound ATP, but the ATP binding site was preprogrammed via the DNA aptamer.

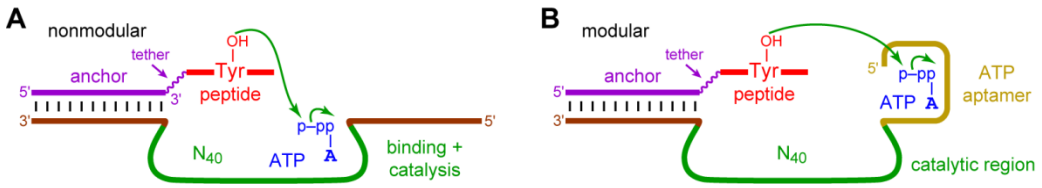


Figure 2.22. Nonmodular and modular tyrosine kinase deoxyribozymes. (A) A nonmodular DNA catalyst. The initially random region both binds the small-molecule ATP substrate and performs catalysis. (B) A modular DNA catalyst. A predefined ATP aptamer region is included adjacent to the initially random region, which is responsible for catalysis. Figure reprinted with permission from Ref. 12.

The tyrosine kinase deoxyribozyme 14JS101 was identified. 14JS101 requires the ATP aptamer domain for activity, suggesting that the ATP phosphoryl donor used for catalysis is bound to the aptamer region. However, 14JS101 does not have improved rate, yield, or K_m(ATP) compared to the tyrosine kinase deoxyribozymes 8EA101 or 16EC103 that were identified without the preprogrammed aptamer domain for NTP substrate binding. These results show that modular deoxyribozymes can be identified. However, the incorporation of the modular ATP binding site does not provide an advantage in identifying tyrosine kinase deoxyribozymes compared to nonmodular tyrosine kinase deoxyribozymes.

2.3 Summary and Future Directions

Deoxyribozymes had previously been identified to catalyze oligonucleotide phosphorylation and nucleopeptide formation. The ability of DNA to catalyze oligonucleotide phosphorylation and separately to react with peptide substrates suggested that DNA-catalyzed peptide phosphorylation was possible. Initial efforts used γ-thiophosphoryl donors to enable separation of active DNA sequences. However, due to the instability of γ-thiophosphoryl donors in selection conditions, their use was discontinued. The 8VP1 capture deoxyribozyme was identified, and a capture method was optimized to separate active DNA sequences from inactive sequences. The kinase deoxyribozyme in vitro selection process was optimized to prevent enrichment of undesired sequences. The first tyrosine kinase deoxyribozymes were identified, and phosphorylate tyrosine residues within DNA-anchored peptide substrates.

To expand upon the capabilities of tyrosine kinase deoxyribozymes, subsequent in vitro selection experiments used small-molecule NTPs as phosphoryl donors. DNA catalysts need to bind to the NTP phosphoryl donor and catalyze phosphorylation of the peptide substrate. Tyrosine kinase deoxyribozymes were identified from selection experiments in which 1 mM GTP was present. Surprisingly, the resulting deoxyribozymes had optimum activity at ~100 μM GTP and were inhibited at higher GTP concentrations. The different NTP requirements of each deoxyribozyme demonstrated that deoxyribozymes can interact with the phosphoryl donor in different ways. Future efforts to identify kinase deoxyribozymes will use ATP for simplicity and relevance as the most common biological phosphoryl donor. Lower ATP concentration of 100 μM will also be used, because the current deoxyribozymes were inhibited GTP concentrations above 100 μM.

Modular tyrosine kinase deoxyribozymes were also identified. These deoxyribozymes had an ATP aptamer domain adjacent to the initially random region. The ATP phosphoryl donor bound to the aptamer domain decreasing the requirements of the initially random region. While a modular tyrosine kinase deoxyribozyme was identified, the incorporation of the ATP binding site

did not provide an advantage compared to the nonmodular tyrosine kinase deoxyribozymes that use GTP as the phosphoryl donor. The ability of DNA catalysts to phosphorylate tyrosine side chains within peptide substrates expands the functionality of DNA to include an important biological regulatory post-translational modification. Some challenges remain, and future efforts are focused on peptide-sequence selectivity, and the phosphorylation of free peptides and proteins.

2.4 Materials and Methods

2.4.1 Substrate Preparation Procedures

Oligonucleotides and peptides. DNA oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA) or prepared by solid-phase synthesis on an ABI 394 instrument using reagents from Glen Research. 5′-Triphoshorylated RNA (pppRNA) oligonucleotides were prepared by in vitro transcription using synthetic DNA templates and T7 RNA polymerase.[13] All oligonucleotides were purified by 7 M urea denaturing PAGE with running buffer 1× TBE (89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3) as described.[14,15] Peptides were synthesized via solid-phase synthesis as described.[6]

Table 2.1: Oligonucleotide sequences of the pools, substrates, and primers used in the selection experiments.

oligonucleotide purpose	oligonucleotide sequence
Selections with pppRNA or 1 mM GTP phosphoryl donors	
DNA-HEG-CAAYAA substrate	GGATAATACGACTCACTAT-HEG-CAAYAA
pppRNA phosphoryl donor	pppGAUGUUCUAGCGCUUCG
forward primer for selection	CGAAGCGCTAGAACAT
reverse primer for selection	$(AAC)_4$-HEG-CCATCAGGATCAGCT
reverse primer for cloning	TAATTAATTAATTACCCATCAGGATCAGCT
random pool for selection	CGAAGCGCTAGAACAT-N_X-ATAGTGAGTCGTATTAAGCTGATCCTGATGG
splint for ligation step during selection	ATAGTGAGTCGTATTATCCTCCATCAGGATCAGCTTAATACGACTCACTAT

Table 2.1 (cont.)

oligonucleotide purpose	oligonucleotide sequence
Reselection of 6CF134 and 7CH102	
DNA-HEG-CAAYAA substrate	GGACTACCTTTATGCGTAT-HEG-CAAYAA
pppRNA phosphoryl donor	pppGGAAGGAGGCUUUCGGG
forward primer for selection	CGAACGAAAGCCTCCTTC
reverse primer for selection	$(AAC)_4$-HEG-CCATCAGGATCAGCT
reverse primer for cloning	TAATTAATTAATTACCCATCAGGATCAGCT
partially randomized pool for selection	CGAACGAAAGCCTCCTTC-N_X-ATACGCATAAAGGTAGAGCTGATCCTGATGG
splint for ligation step during selection	ATACGCATAAAGGTAGTCCTCCATCAGGATCAGCTCTACCTTTATGCGTAT
Selections with 100 μM ATP phosphoryl donor and N_{30} random region	
DNA-HEG-CAAYAA substrate	GGACTACCTTTATGCGTAT-HEG-CAAYAA
forward primer for selection	CGAACGAAAGCCTCCTTC
reverse primer for selection	$(AAC)_4$-HEG-CCATCAGGATCAGCT
reverse primer for cloning	TAATTAATTAATTACCCATCAGGATCAGCT
random pool for selection	CGAACGAAAGCCTCCTTC-N_{30}-ATACGCATAAAGGTAGAGCTGATCCTGATGG
splint for ligation step during selection	ATACGCATAAAGGTAGTCCTCCATCAGGATCAGCTCTACCTTTATGCGTAT
Selections with 100 μM ATP phosphoryl donor and N_{40} random region	
DNA-HEG-CAAYAA substrate	GGAATATCTCGTTTCTTAT-HEG-CAAYAA
forward primer for selection	CGAATTAAGACTGAATTC
reverse primer for selection	$(AAC)_4$-HEG-CCATCAGGATCAGCT
reverse primer for cloning	TAATTAATTAATTACCCATCAGGATCAGCT
random pool for selection	CGAATTAAGACTGAATTC-N_{40}-ATAAGAAACGAGATATAGCTGATCCTGATGG
splint for ligation step during selection	ATAAGAAACGAGATATTCCTCCATCAGGATCAGCTATATCTCGTTTCTTAT

Synthesis of DNA-anchored peptide conjugates. DNA-anchored peptide conjugates were synthesized by disulfide formation between a DNA HEG-tethered 3′-thiol and the N-terminal cysteine side chain of the peptide (Figure 2.23). The DNA anchor oligonucleotide was 5′-GGATAATACGACTCACTAT-HEG-p-C_3-SS-C_3-OH-3′. The 3′-disulfide linker was introduced via standard solid-phase DNA synthesis and reduced to a 3′-thiol by DTT treatment. A 50 μL sample containing 5 nmol of DNA anchor oligonucleotide in 100 mM HEPES, pH 7.5, and 50 mM DTT was incubated at 37 °C for 2 h. The reduced product was precipitated to remove DTT by addition of 50 μL of water, 10 μL of 3 M NaCl, and 300 μL of ethanol. The precipitated product (DNA-HEG-p-C_3-SH) was dissolved in 35 μL of water and 10 μL of 100 mM triethylammonium

acetate, pH 7.0. Activation as the pyridyl disulfide was achieved by adding 5 µL of 100 mM 2,2′-dipyridyl disulfide in DMF and incubating at 37 °C for 2 h. The product (DNA-HEG-p-C_3-SSPy) was precipitated by addition of 50 µL of water, 10 µL of 3 M NaCl and 300 µL of ethanol and dissolved in 35 µL of water. Conjugation to the peptide was performed by adding 10 µL of 100 mM triethylammonium acetate, pH 7.0, and 5 µL of 20 mM peptide (100 nmol, 20 equivalents). The sample was incubated at 37 °C for 12 h, and the DNA-anchored peptide was purified by 20% PAGE.

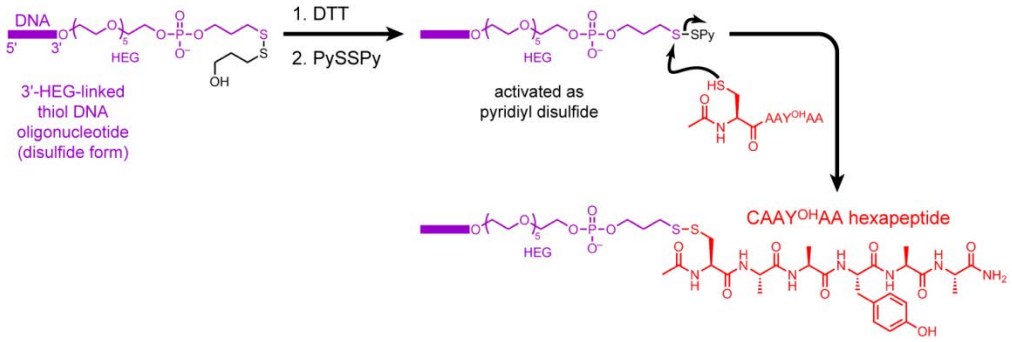


Figure 2.23. Synthesis and structure of the DNA-anchored CAAYAA hexapeptide substrate. A hexa(ethylene glycol), or HEG, tether connects the DNA oligonucleotide anchor to the peptide. The 3′-HEG-linked thiol DNA is synthesized as a disulfide, which is reduced, activated, and then conjugated to the peptide substrate via a disulfide bond with the N-terminal cysteine residue. Figure reprinted with permission from Ref. 11.

2.4.2 In Vitro Selection Procedures

Procedure for ligation step in round 1. A 35 µL sample containing 1 nmol of DNA pool, 800 pmol of DNA splint, and 500 pmol of 5′-phosphorylated DNA-anchored hexapeptide substrate was annealed in 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. To this solution was added 4 µL of 10× T4 DNA ligase buffer (400 mM Tris, pH 7.8, 100 mM $MgCl_2$, and 5 mM ATP) and 1 µL of 5 U/µL T4 DNA ligase (Fermentas). The sample was incubated at 37 °C for 12 h and separated by 8% PAGE.

Procedure for ligation step in subsequent rounds. A 17 µL sample containing the PCR-amplified DNA pool (~5-10 pmol), 25 pmol of DNA splint, and 50 pmol of 5′-phosphorylated DNA-anchored hexapeptide substrate was annealed in 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. To this solution was added 2 µL of 10× T4 DNA ligase buffer and 1 µL of 1 U/µL T4 DNA ligase (Fermentas). The sample was incubated at 37 °C for 12 h and separated by 8% PAGE.

Procedure for selection step in round 1. Each selection experiment was initiated with 200 pmol of the ligated pool. A 20 µL sample containing 200 pmol of ligated pool and 300 pmol of pppRNA phosphoryl donor was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The selection reaction was initiated by bringing the sample to 40 µL total volume containing 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl. The Mn^{2+} was added from a 10× stock solution containing 200 mM $MnCl_2$. The Zn^{2+} was added from a 10× stock solution containing 10 mM $ZnCl_2$, 20 mM HNO_3, and 200 mM HEPES, pH 7.5; this stock solution was freshly prepared from a 100× stock of 100 mM $ZnCl_2$ in 200 mM HNO_3. The metal ion stocks were added last to the final sample. For selections with NTP as the phosphoryl donor, the pppRNA was omitted and 1 mM NTP was added from a 10 mM stock solution after annealing. The sample was incubated at 37 °C for 14 h and separated by 8% PAGE.

Procedure for selection step in subsequent rounds. A 10 µL sample containing ligated pool and 50 pmol of pppRNA phosphoryl donor was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The selection reaction was initiated by bringing the sample to 20 µL total volume containing 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl. For selections with NTP as the phosphoryl donor, the pppRNA was omitted and 1× NTP was added from a 10× stock solution after annealing. The sample was incubated at 37 °C for 14 h and separated by 8% PAGE.

Procedure for capture step in round 1. A 28 µL sample containing ligated pool, 500 pmol of 54 nt 5′-triphosphorylated RNA-DNA chimera substrate, and 300 pmol of 8VP1 capture deoxyribozyme was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The capture reaction was initiated by bringing the sample to 40 µL total volume containing 50 mM HEPES, pH 7.5, 20 mM $MnCl_2$, and 150 mM NaCl. The sample was incubated at 37 °C for 14 h. Before PAGE purification, to the sample was added 500 pmol of a 60 nt decoy oligonucleotide complementary to the 40 nt 8VP1 catalytic region and 10 nt of binding arm on each side, to ensure complete removal of the capture deoxyribozyme from the pool. The sample was separated by 8% PAGE.

Procedure for capture step in subsequent rounds. A 14 µL sample containing ligated pool, 50 pmol of 17 nt pppRNA substrate (in even rounds) or 54 nt 5′-triphosphorylated RNA-DNA chimera substrate (in odd rounds), and 25 pmol of 8VP1 capture deoxyribozyme was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The capture reaction was initiated by bringing the sample to 20 µL total volume containing 50 mM HEPES, pH 7.5, 20 mM $MnCl_2$, and 150 mM NaCl. The sample was incubated at 37 °C for 14 h. Before PAGE purification, to the sample was added 100 pmol of a 60 nt decoy oligonucleotide complementary to the 40 nt 8VP1 catalytic region and 10 nt of binding arm on each side (note that the decoy sequence is different in even-numbered and odd-numbered rounds). The sample was separated by 8% PAGE.

Procedure for PCR in subsequent rounds. In each selection round, two PCR reactions were performed, 10-cycle PCR followed by 30-cycle PCR. First, a 100 µL sample was prepared containing the PAGE-separated capture product, 200 pmol of forward primer, 50 pmol of reverse primer, 20 nmol of each dNTP, and 10 µL of 10× Taq polymerase buffer (200 mM Tris-HCl, pH 8.8, 100 mM $(NH_4)_2SO_4$, 100 mM KCl, 20 mM $MgSO_4$, and 1% Triton X-100), and Taq polymerase. This sampled was cycled 10 times according to the following PCR program: 94 °C for 2 min, 10× (94 °C for 30 s, 47 °C for 30 s, 72 °C for 30 s), 72 °C for 5 min. Taq polymerase was removed by phenol/chloroform extraction. Second, a 50 µL sample was prepared containing

1 μL of the 100 μL 10-cycle PCR product, 100 pmol of forward primer, 25 pmol of reverse primer, 10 nmol of each dNTP, 20 μCi of α-^{32}P-dCTP (800 Ci/mmol), 5 μL of 10× Taq polymerase buffer, and Taq polymerase. This sample was cycled 30 times according to the following PCR program: 94 °C for 2 min, 30× (94 °C for 30 s, 47 °C for 30 s, 72 °C for 30 s), 72 °C for 5 min. Samples were separated by 8% PAGE.

2.4.3 Cloning and Screening

Using 1 μL of a 1:1000 dilution of the 10-cycle PCR product from the end of the selection round being cloned, 30-cycle PCR was performed using the above procedure, omitting α-^{32}P-dCTP and using 25 pmol of forward primer and 25 pmol of reverse cloning primer, where the extensions with TAA stop codons in each frame were included to suppress false negatives in blue-white screening.[16] The PCR product was purified on 2% agarose and cloned using a TOPO TA cloning kit (Invitrogen). Miniprep DNA samples derived from individual *E. coli* colonies were assayed by digestion with EcoRI to ascertain the presence of the expected inserts. Using the miniprep DNA samples, PCR (same conditions as 30-cycle PCR during selection, omitting α-^{32}P-dCTP) was performed to obtain the deoxyribozymes. Half of the PCR product was used to perform screening assays. Each screening assay used ~0.1 pmol of 5′-^{32}P-radiolabeled DNA-anchored hexapeptide substrate, 50 pmol of pppRNA phosphoryl donor, ~20 pmol of deoxyribozyme and followed the single-turnover assay procedure described in the subsequent section.

2.4.4 Deoxyribozyme Activity Assay Procedure

The DNA-anchored hexapeptide substrate was 5′-^{32}P-radiolabeled using γ-^{32}P-ATP and T4 polynucleotide kinase (Fermentas), using 10× buffer that lacks DTT (500 mM Tris, pH 7.6, 100 mM $MgCl_2$, and 1 mM spermidine). A 10 μL sample containing 0.25 pmol of 5′-^{32}P-radiolabeled DNA-anchored peptide substrate, 10 pmol of deoxyribozyme, and 20 pmol of pppRNA phosphoryl donor was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM

EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. For assays with NTP as the phosphoryl donor, the pppRNA was omitted and 2 μL of NTP was added from a 10× stock solution after annealing in an 8 μL volume. The DNA-catalyzed phosphorylation reaction was initiated by bringing the sample to 20 μL total volume containing 70 mM HEPES, pH 7.5, 0.5 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl (or other ion concentrations as appropriate). The Mn^{2+} was added from a 10× stock solution containing 200 mM $MnCl_2$. The Zn^{2+} was added from a 10× stock solution containing 5 mM $ZnCl_2$, 10 mM HNO_3, and 100 mM HEPES at pH 7.5; this stock solution was freshly prepared from a 200× stock of 100 mM $ZnCl_2$ in 200 mM HNO_3. The metal ion stocks were added last to the final sample, which was divided into 2 μL aliquots that were all incubated at 37 °C. At appropriate time points, 2 μL aliquots were quenched with 5 μL stop solution (80% formamide, 1× TBE [89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3], 50 mM EDTA, 0.025% bromophenol blue, 0.025% xylene cyanol). Samples were separated by 20% PAGE and quantified with a PhosphorImager. Values of k_{obs} were obtained by fitting the yield versus time data directly to first-order kinetics; i.e., yield $Y = Y_{max} \cdot (1 - e^{-kt})$, where $k = k_{obs}$ and Y_{max} is the final yield. Errors in k_{obs} values were calculated as the standard deviation from the indicated number of independent determinations.

2.4.5 Mass Spectrometry

The DNA-anchored CAAY^{P}AA phosphorylation product was prepared from a 10 μL sample containing 300 pmol of DNA-anchored HEG-tethered CAAY^{OH}AA substrate, 350 pmol of deoxyribozyme, and 400 pmol of pppRNA phosphoryl donor, which were annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. For samples with GTP as the phosphoryl donor, the pppRNA was omitted and 1× NTP was added from a 10× stock solution after annealing. The DNA-catalyzed phosphorylation reaction was initiated by bringing the sample to 20 μL total volume containing 70 mM HEPES, pH 7.5, 0.5 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl. The sample was incubated at 37 °C for 14 h and separated by 8% PAGE (only the lower one-fourth to one-half of

the product band was excised, to suppress extraneous signals due to unreacted substrate). The PAGE-purified sample was desalted by a Millipore C_{18} ZipTip and analyzed by MALDI mass spectrometry (Bruker UltrafleXtreme; matrix 3-hydroxypicolinic acid).

2.5 References

(1) Höbartner, C.; Silverman, S. K. Recent advances in DNA catalysis. *Biopolymers* **2007**, *87*, 279-292.

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

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

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

(5) 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.

(6) Chandrasekar, J.; Silverman, S. K. Catalytic DNA with phosphatase activity. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5315-5320.

(7) Brandsen, B. M.; Hesser, A. R.; Castner, M. A.; Chandra, M.; Silverman, S. K. DNA-catalyzed hydrolysis of esters and aromatic amides. *J. Am. Chem. Soc.* **2013**, *135*, 16014-16017.

(8) Sachdeva, A. "Deoxyribozymes for Peptide Substrates: Exploring the Landscape of Nucleophiles and Electrophiles". Ph.D. Thesis, University of Illinois at Urbana-Champaign, **2011**.

(9) Sachdeva, A.; Chandra, M.; Chandrasekar, J.; Silverman, S. K. Covalent tagging of phosphorylated peptides by phosphate-specific deoxyribozymes. *ChemBioChem* **2012**, *13*, 654-657.

(10) Velez, T. E.; Singh, J.; Xiao, Y.; Allen, E. C.; Wong, O. Y.; Chandra, M.; Kwon, S. C.; Silverman, S. K. Systematic evaluation of the dependence of deoxyribozyme catalysis on random region length. *ACS Comb. Sci.* **2012**, *14*, 680-687.

(11) Walsh, S. M.; Sachdeva, A.; Silverman, S. K. DNA catalysts with tyrosine kinase activity. *J. Am. Chem. Soc.* **2013**, *135*, 14928-14931.

(12) Dokukin, V.; Silverman, S. K. A modular tyrosine kinase deoxyribozyme with discrete aptamer and catalyst domains. *Chem. Commun.* **2014**, *50*, 9317-9320.

(13) Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **1987**, *15*, 8783-8798.

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

(15) Wang, Y.; Silverman, S. K. Characterization of deoxyribozymes that synthesize branched RNA. *Biochemistry* **2003**, *42*, 15252-15263.

(16) Langner, J.; Klussmann, S. PCR primers containing stop codons reduce the number of false-negatives during blue-white screening. *Biotechniques* **2003**, *34*, 950-954.

Chapter 3: Peptide Sequence-Specific Tyrosine Kinase Deoxyribozymes[†]

3.1 Introduction

3.1.1 Site-Specific Modification of Proteins

Natural proteins are modified by many types of post-translational modifications at specific amino acid residues.[16] Protein kinases phosphorylate specific amino acid residues within their protein substrates based on the local sequence, and many known kinase sequence recognition motifs have been reported.[713] High-throughput screening methods have been developed to determine kinase substrate requirements.[1416] Phosphorylation motifs have been cataloged and incorporated into searchable databases with algorithms to predict phosphorylation sites in proteins of interest.[1724] Despite the large number of known or predicted protein phosphorylation sites, protein kinases that are able to phosphorylate proteins of interest at specific sites are not always known.

Artificial catalysts that catalyze the covalent modification of proteins in a highly selective manner are valuable for biochemical studies. Small-molecule catalysts for protein modification usually modify any accessible reactive side chain without regard to the sequence context.[25] Natural protein enzymes that catalyze the post-translational modification of proteins can be

[†] This research has been published:

Walsh, S. M.; Konecki, S. N.; Silverman, S. K. Identification of Sequence-Selective Tyrosine Kinase Deoxyribozymes. *J. Mol. Evol.* **2015**, *81*, 218-224.

University of Illinois undergraduate student Stephanie N. Konecki synthesized the complex peptide substrates.

University of Illinois graduate student Chih-Chi Chu identified and characterized the nucleopeptide-forming deoxyribozymes with peptide-sequence specificity.[40]

University of Illinois graduate student Puzhou Wang identified and characterized DNA enzymes for tyrosine adenylylation.[52]

subjected to directed evolution[26,27] for a higher rate constant and other improved practical characteristics.[28] However, directed evolution of known enzymes for the desired sequence selectivity is often difficult, and tolerance for new substrates often leads to relaxed substrate specificity.[29] Therefore, identification of new catalysts for site-specific protein modifications is valuable.

3.1.2 Deoxyribozymes with Peptide-Sequence Specificity

Nucleic acid aptamers specifically bind to peptide and protein substrates based on their sequence and structure.[3036] Considering the substrate specificity of protein catalysts and DNA aptamers, it is likely that DNA enzymes that are complex biomolecular catalysts can also interact specifically with peptide and protein substrates.

Deoxyribozymes have been identified to catalyze the modification of peptide substrates[37] including via nucleopeptide formation,[38-41] phosphorylation,[42,43] dephosphorylation,[44] and phosphoserine lyase activity.[45] However, in all of these previous selection experiments, only a few of the resulting deoxyribozymes discriminate among peptide substrates based on the peptide sequence (Figure 3.1).[46] These selective nucleopeptide-forming deoxyribozymes were identified using untethered peptide substrates during the in vitro selection experiments. DNA enzymes that catalyze the modification of untethered peptide substrates may require a stronger interaction between the deoxyribozyme and peptide substrate. All other peptide-modifying deoxyribozymes were identified from selection experiments in which DNA-anchored peptide substrates were used. This includes tyrosine kinase deoxyribozymes that do not discriminate among peptide substrates based on their amino acid sequences.[42,43] The presence of a tethered peptide substrate may not require the deoxyribozyme to interact as specifically with the peptide substrate as the deoxyribozyme may with an untethered peptide substrate.

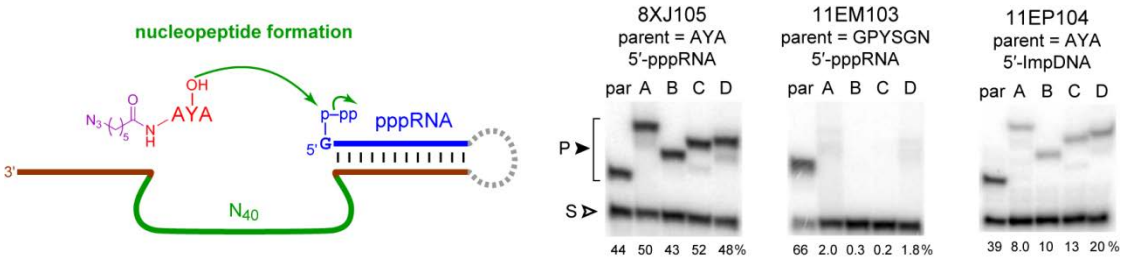


Figure 3.1. Peptide-sequence-selective nucleopeptide-forming deoxyribozymes with untethered peptide reactivity. DNA-catalyzed nucleopeptide formation was evaluated with multiple peptide substrates. The parent (par) peptide sequence with N-terminal azido group used during the selection is indicated. Peptides A-D with an N-terminal acetyl group are as follows: A = RTRYERN; B = FSDYIHP; C = EKIYIHP; D = QAGYKGR. 8XJ105 is peptide sequence-general, 11EM103 is peptide sequence-specific, and 11EP104 is partially peptide sequence-specific. Incubation conditions are 1 mM peptide, 70 mM HEPES, pH 7.5, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1.0 mM (8XJ105) or 0.4 mM (11EM103, 11EP104) $ZnCl_2$, and 150 mM NaCl at 37 °C for 16 h. Figure adapted with permission from Ref. 40.

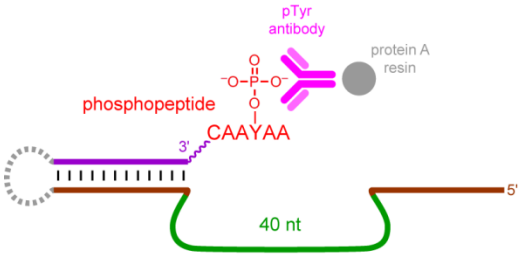
With nucleopeptide-forming deoxyribozymes as the sole exception, all known deoxyribozymes that catalyze peptide modification were identified using peptide substrates containing the reactive amino acid flanked solely by alanine residues (i.e. AAYAA). Incorporation of more diverse amino acids surrounding the reactive residue may increase the interactions between the deoxyribozyme and the peptide substrate.

3.2 Results and Discussion

3.2.1 Development of the Phosphotyrosine Immunoprecipitation Capture Method

The previously used 8VP1 gel-based capture method to identify tyrosine kinase deoxyribozymes enables the isolation of active DNA sequences based on both product reactivity and physical size. Bead-based methods only required reactivity for the product to be separated. Phosphotyrosine (pTyr) antibodies specifically bind to phosphorylated tyrosine residues, and the binding interaction requires only the phosphotyrosine side chain. Therefore, antibody binding of the phosphotyrosine products should be unaffected by the length of the tether between the peptide substrate and the DNA anchor.

A bead-based immunoprecipitation method was developed to capture phosphotyrosine-containing peptide products (Figure 3.2). After the selection step, the sample was precipitated and incubated with a phosphotyrosine antibody (pY20) and protein A immobilized on an agarose resin. Protein A binds to the constant Fc region of an antibody, and the variable region of the pTyr antibody binds specifically to phosphotyrosine. Therefore, the active DNA sequences connected to their phosphorylated peptide product were bound to the resin via the pTyr antibody-protein A interaction. After incubation, the resin was washed, removing the unbound non-phosphorylated peptide substrates connected to the inactive DNA sequences. The bound deoxyribozymes were then eluted from the resin, and scintillation counting of the wash and elution fractions was used to quantify activity in the selection experiment.



CAAY^{P}AA %	capture yield, %	normalized yield, %
0	0.4	1.0
10	7.7	20.2
50	21.4	56.2
90	40.0	105
100	38.1	100

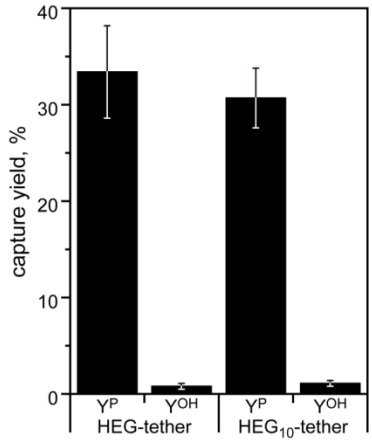


Figure 3.2. Phosphotyrosine immunoprecipitation capture method. The phosphotyrosine-containing peptide product connected to the active DNA sequences were immunoprecipitated by binding to a pTyr antibody immobilized on a protein A-coated agarose resin. The plot shows the capture of DNA-anchored peptides (mean ± sd, n = 3). Y^{P} = CAAY^{P}AA; Y^{OH} = CAAYAA. Phosphotyrosine-containing peptides were captured independently of the length of the tether between the peptide and DNA anchor, but the tyrosine-containing substrate was not captured. The table shows capture results of the mock selection round. During the ligation step, the indicated amount of phosphotyrosine-containing peptide substrate was included. The capture yield is shown, and the yield is normalized to the selection experiment with 100% CAAY^{P}AA.

The phosphotyrosine immunoprecipitation capture method was optimized for use during in vitro selection of kinase deoxyribozymes. To prevent the DNA pool from adhering to the

reaction tube, low-retention tubes were coated with a 5% PEG solution, incubated at room temperature overnight, and rinsed prior to use. To the sample was also added 50 pmol of a sacrificial oligonucleotide $(AAC)_{20}$. Incubation of the DNA pool with the pTyr antibody prior to addition of the protein A resin resulted in the same capture yield as adding both the pTyr antibody and resin to the sample at the same time. Therefore, to simplify the capture method, both the pTyr antibody and protein A resin were added at once. Shaking the sample on a vortexer at the lowest setting was the most effective method to mix the small 35 µL sample. Mixing the sample at room temperature for 1 h led to the same capture yield as mixing at 4 °C for 16 h. The shorter 1 h incubation at room temperature was used. The immunoprecipitation capture method was highly selective for phosphotyrosine with a 30% capture yield, whereas <1% of the tyrosine substrate was captured. A mock selection round was performed to validate the phosphotyrosine immunoprecipitation capture method. Varying amounts of phosphotyrosine-containing peptide product were doped in with the tyrosine-containing substrate during the ligation step. Capture yields using the immunoprecipitation method were consistent with the added amount of phosphotyrosine-containing peptide. Therefore, the phosphotyrosine immunoprecipitation capture method was viable for use during the in vitro selection of tyrosine kinase deoxyribozymes.

One concern with the bead-based capture method is the unintended enrichment of DNA aptamers that bind to the antibody, protein A, or resin. To avoid this, a preclear step was performed between the ligation and selection step. In this step, the DNA sequences ligated to the substrate were subjected to the immunoprecipitation capture. The DNA sequences that did not bind to the pTyr antibody, protein A, or resin were recovered in the wash and taken on to the selection step. During the preclear step only the substrate (Y^{OH}) was present. Therefore, DNA sequences that were aptamers to the antibody, protein A, or resin were removed.

3.2.2 In Vitro Selection for Peptide Sequence-Specific Tyrosine Kinase Deoxyribozymes

In vitro selection experiments were performed to identify peptide sequence-specific tyrosine kinase deoxyribozymes using complex, biologically derived peptide substrates.[47] Peptide substrates were octapeptide sequences derived from biologically phosphorylated proteins, and the peptide sequences were more complex than the previously used CAAYAA peptide substrate (Chapter 2). The more complex peptide sequences may enable more interactions between the DNA enzyme and the peptide substrate, resulting in deoxyribozymes with peptide-sequence selectivity. Three different peptide sequences were used in the selection experiments: CMTGYVAT, CADPYDQS, and CKVIYDFI. The CMTGYVAT and CADPYDQS peptide sequences correspond to amino acids 179-185 and 320-326, respectively, from the MAPK14/p38α protein (mitogen-activated protein kinase 14).[48,49] The CKVIYDFI peptide sequence corresponds to amino acids 253-259 from N-WASP (neuronal Wiskott-Aldrich Syndrome protein).[50] The tyrosine residue within each peptide sequence is naturally phosphorylated in the corresponding protein substrate. Each peptide included an artificial cysteine residue on the N-terminus to enable disulfide conjugation of the peptide to the DNA anchor oligonucleotide.

In vitro selection was performed similarly to the previous kinase deoxyribozyme selection experiments (Chapter 2), except with a different capture step using an immobilized phosphotyrosine (pTyr) antibody to separate the catalytically active DNA sequences in each selection round (Figure 3.3). The octapeptide substrates were prepared by solid-phase peptide synthesis and connected to a DNA anchor via a disulfide bond and a hexa(ethylene glycol) [HEG] tether. The pppRNA phosphoryl donor was prepared by in vitro transcription using a DNA template and T7 RNA polymerase. The DNA-anchored peptide substrate was ligated to the initially random DNA pool using T4 DNA ligase and PAGE-purified. Next, the selection step was performed in which the DNA sequences were incubated with a phosphoryl donor and divalent metal ions in a buffered solution to enable DNA-catalyzed peptide phosphorylation. A subsequent capture step was performed to isolate the desired active DNA sequences connected to

their phosphopeptide product. The capture step was a bead-based immunoprecipitation method to isolate the active DNA sequences. The immunoprecipitation capture method enabled more selection experiments to be performed in parallel than the 8VP1 capture deoxyribozyme method. Immobilized pTyr antibody binds specifically to the phosphotyrosine-containing peptide products connected to the active DNA sequences. The bound, active DNA sequences were separated from the unbound, inactive DNA sequences. PCR was then performed to amplify the catalytic DNA sequences that survived the selection and capture steps. The entire selection round was repeated, each time the population was enriched with catalytically active sequences, until they dominated the population.

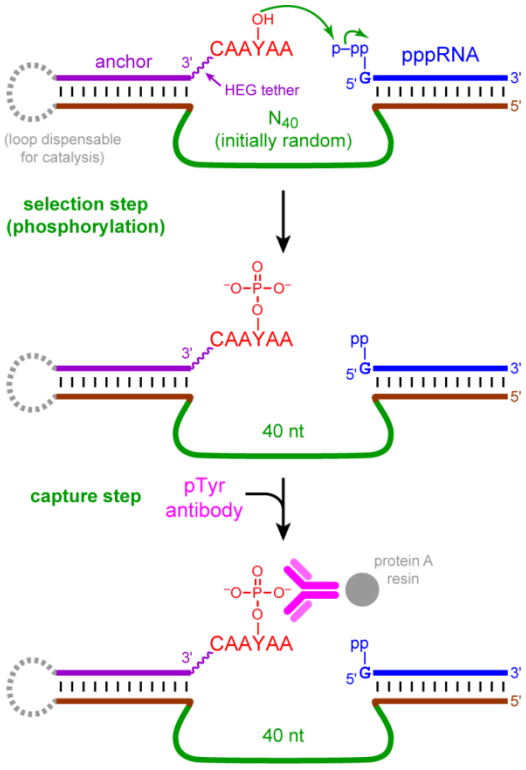


Figure 3.3. Selection strategy for tyrosine kinase deoxyribozymes using an pTyr antibody. After the selection step, the active DNA sequences connected to the phosphopepetide products are captured via immunoprecipitation with an immobilized pTyr antibody. Figure reprinted with permission from Ref. 47.

An additional preclear step was performed in some selection rounds to prevent DNA sequences that bind to the pTyr antibody, protein A, or resin from surviving the capture step and being enriched throughout the selection rounds. In these rounds with the preclear step, the ligation product was incubated with the immobilized antibody and processed with the same procedure as for the capture step. However, in the preclear step the DNA sequences were connected to the tyrosine-containing peptide substrate. Therefore, the DNA sequences that did *not* bind to the resin continued on to the selection step.

In vitro selection experiments RG1-RT1 were designed to identify peptide sequence-specific tyrosine kinase deoxyribozymes (Figure 3.4). Selection experiments with each of the three different biologically derived, complex peptide sequences were performed with an N_{40} or N_{80} initially random region. The length of the initially random region has been shown to be an important variable in deoxyribozyme selection experiments,[51] and tyrosine kinase deoxyribozymes have been identified from N_{30}, N_{40}, and N_{50} initially random region lengths (Chapter 2).[42,43] Larger random regions may enable more complex catalytic DNA structures and may lead to more interactions between the DNA enzyme and the peptide substrate. However, all selection experiments evaluate 10^{14} DNA sequences, and this only a small fraction of the 10^{48} possible N_{80} sequences compared to of the 10^{24} possible N_{40} sequences. Therefore, for the random region length used during selection experiments, a compromise is made between the structural complexity and the sequence-space coverage. Each of these selection experiments were performed with either no preclear step or a preclear step every third round beginning in round 1, for a total of twelve selection experiments. All selections were performed in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h.

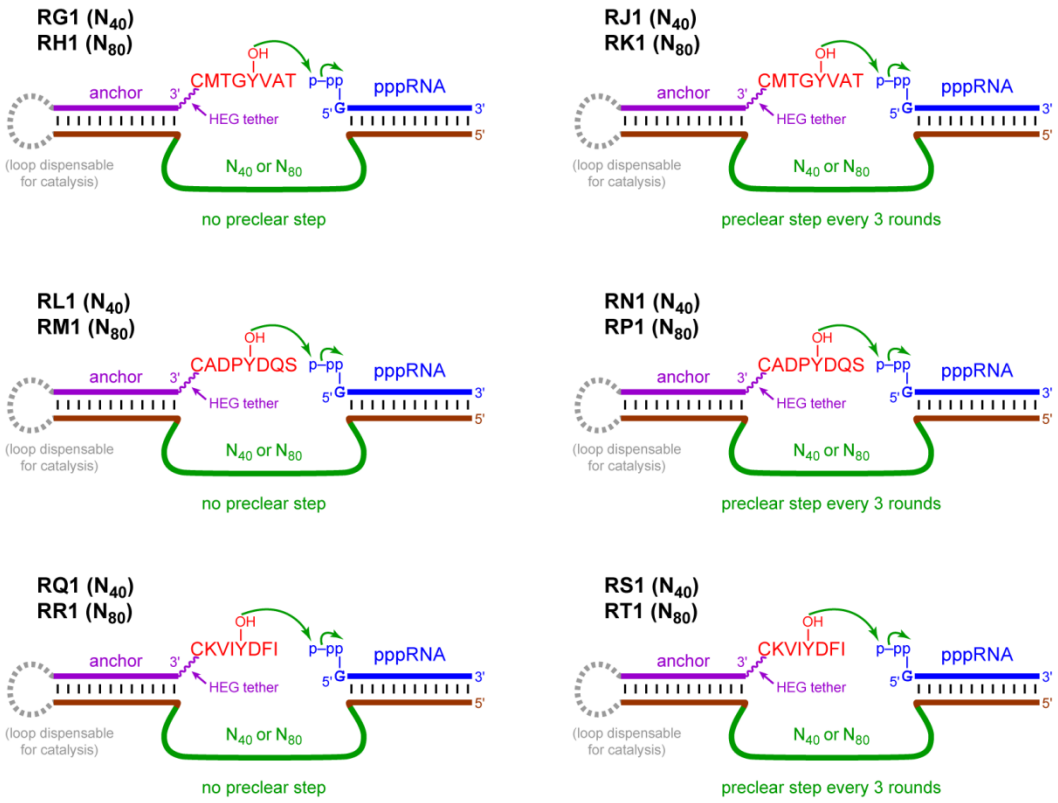


Figure 3.4. Design for the RG1-RT1 in vitro selection experiments to identify peptide sequence-specific tyrosine kinase deoxyribozymes. In each selection, the peptide substrate is connected to a DNA anchor via a disulfide bond and HEG tether, and the phosphoryl donor is pppRNA. The RG1-RK1, RL1-RP1, and RQ1-RT1 selections were with the peptide substrate CMTGYVAT, CADPYDQS, and CKVIYDFI, respectively. The initially random region was N_{40} or N_{80} as indicated. The selections were performed without the preclear step or the preclear step every three rounds beginning in round 1 as indicated. Incubation conditions for each selection was 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h.

Three selection experiments incorporating the preclear step every third round led to the identification of deoxyribozymes with kinase activity: N_{40} with CMTGYVAT (RJ1), N_{40} with CADPYDQS (RN1), and N_{80} with CADPYDQS (RP1) (Figure 3.5). Phosphorylation of the peptide substrate by the active pools was confirmed in trans (without ligating the DNA-anchored peptide substrate to the DNA pool) by directly assessing the DNA-anchored peptide substrate. These three selection experiments were cloned, and individual deoxyribozymes were identified. The other three selection experiments with the preclear step did not lead to activity, and they were discontinued after 11 rounds of selection.

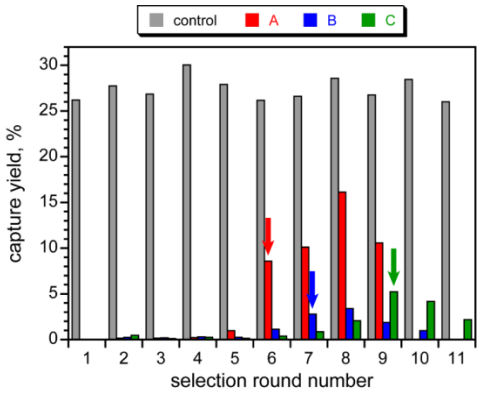


Figure 3.5. Selection progression of the in vitro selection experiments. Control refers to a capture reaction performed each round using the CADPY^{P}DQS peptide (Y^{P} = pTyr). A, B, and C refer to the N_{40} CADPYDQS (RN1), N_{80} CADPYDQS (RP1), and N_{40} CMTGYVAT (RJ1) selection experiments which were halted after round 9, 10, and 11, respectively. Arrows indicate the rounds at which individual deoxyribozymes were cloned from each selection. Figure reprinted with permission from Ref. 47.

Some of the selection experiments in which no preclear step was performed also led to activity. However, phosphorylation of the DNA-anchored peptide substrate was not observed when the pools were evaluated in trans. These results indicate that the enriched sequences were not kinase deoxyribozymes, and the preclear step was required to prevent the enrichment of DNA sequences that bind to the immobilized antibody during the capture step. Therefore, future selection experiments using the bead-based pTyr immunoprecipitation capture method will include the preclear step every third round.

3.2.3 Characterization of Peptide Sequence-Specific Deoxyribozymes

Three unique N_{40} and two unique N_{80} tyrosine kinase deoxyribozymes were identified using the CADPYDQS peptide substrate and designated TyrKinA1-3 and TyrKinB1-2, respectively (Figure 3.6). One unique tyrosine kinase deoxyribozyme, TyrKinC1, was identified from the N_{40} selection with the CMTGYVAT peptide substrate. Matrix-assisted laser desorption ionization (MALDI) mass spectrometry confirmed phosphorylation of the DNA-anchored peptide substrate by each deoxyribozyme (Figure 3.7).

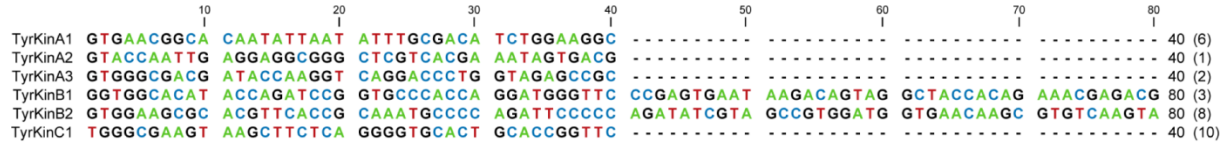
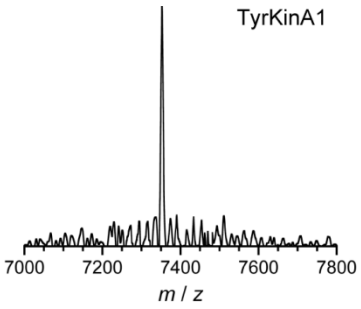


Figure 3.6. Sequences of the initially random regions (N_{40} or N_{80}) of the tyrosine kinase deoxyribozymes. Only the initially random region is shown. Next to the sequence length on the right, in parentheses, is the number of times each sequence was found.



deoxyribozyme	$[M+H]^+$ calcd.	$[M+H]^+$ found	error, % (found – calcd.)
TyrKinA1	7352.2	7352.4	+0.002
TyrKinA2	7352.2	7353.1	+0.009
TyrKinA3	7352.2	7353.0	+0.008
TyrKinB1	7352.2	7351.6	–0.006
TyrKinB2	7352.2	7349.6	–0.026
TyrKinC1	7304.3	7305.0	+0.007

Figure 3.7. MALDI mass spectrometry of the tyrosine kinase deoxyribozyme catalyzed product. The mass is for the DNA-HEG-phosphopeptide product. Parent peptide substrates were used; CADPYDQS for TyrKinA1-3 and TyrKinB1-2 and CMTGYVAT for TyrKinC1. Figure adapted with permission from Ref. 47.

Phosphorylation activity was evaluated with all three peptide substrates from the selection experiments. TyrKinC1 has k_{obs} of 0.09 h^{-1} and 79% phosphorylation yield at 48 h with CMTGYVAT, the peptide substrate used during the selection (Figure 3.8). The TyrKinC1 deoxyribozyme poorly discriminates against the CADPYDQS peptide substrate with k_{obs} of 0.06 h^{-1} and 63% yield at 48 h. However, TyrKinC1-catalyzed phosphorylation of the CKVIYDFI peptide substrate was only 13% at 48 h with k_{obs} of 0.04 h^{-1}. Therefore, TyrKinC1 has partial peptide-sequence selectivity but is not highly selective for the peptide substrate used during the selection experiment.

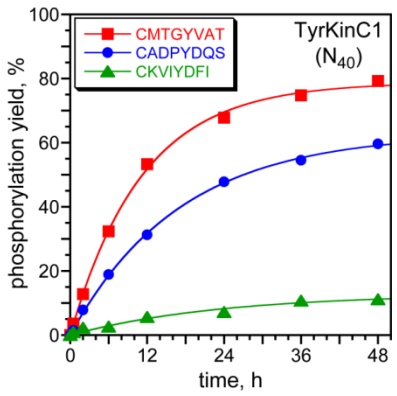


Figure 3.8. Activity of the partially peptide-sequence-selective tyrosine kinase deoxyribozyme TyrKinC1. Phosphorylation activity was evaluated with each of the three peptide substrates CMTGYVAT, CADPYDQS, and CKVIYDFI. The CMTGYVAT peptide substrate was used during the selection. Incubation conditions were 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C.

The TyrKinA1-3 and TyrKinB1-2 kinase deoxyribozymes identified from selection with the CADPYDQS peptide substrate have k_{obs} values ranging from 0.05-0.11 h^{-1} and yields as high as 62% in 48 h (Figure 3.9). Each deoxyribozyme was assayed with the three different peptide substrates from the selection experiments, and in all cases only the peptide substrate used during the selection, CADPYDQS, was phosphorylated. Therefore, all five kinase deoxyribozymes selectively phosphorylate the DNA-anchored CADPYDQS peptide substrate.

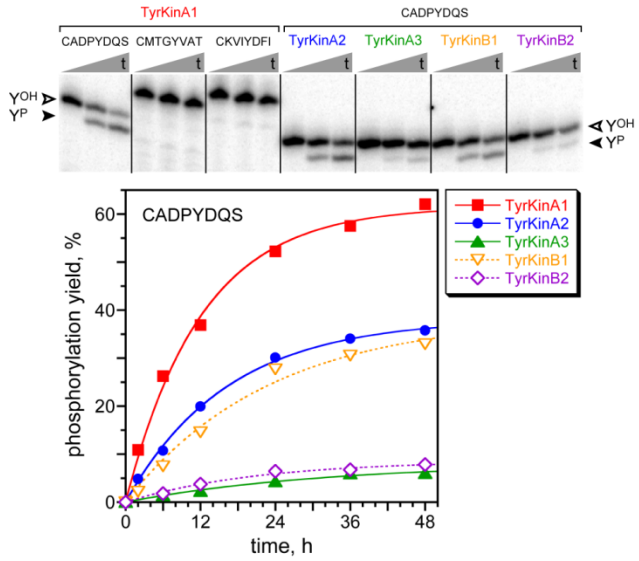


Figure 3.9. Activities of the peptide-sequence-selective tyrosine kinase deoxyribozymes. Each of the five deoxyribozymes was evaluated with each of the three peptide substrates CADPYDQS, CMTGYAT, and CKVIYDFI. The PAGE image shows representative data for TyrKinA1 with all three peptide substrates and for the four other deoxyribozymes with the CADPYDQS peptide substrate used during the selection (t = 30 s, 12 h, 48 h). Y^{OH} = substrate, Y^{P} = product. As with TyrKinA1, the other four deoxyribozymes had no activity with the two peptide substrates not shown. The k_{obs} for TyrKinA1 = 0.11 ± 0.02 h^{-1}, TyrKinA2 = 0.08 ± 0.02 h^{-1}, TyrKinA3 = 0.05 ± 0.02 h^{-1}, TyrKinB1 = 0.05 ± 0.01 h^{-1}, and TyrKinB2 = 0.06 ± 0.01 h^{-1} (n = 4, mean ± sd). The kinetic plot shows the phosphorylation of CADPYDQS by each of the five deoxyribozymes. Incubation conditions were 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C. Figure reprinted with permission from Ref. 47.

The metal dependence was evaluated for the TyrKinA1-3 and TyrKinB1-2 kinase deoxyribozymes (Figure 3.10). The selection experiments were performed in the presence of 1 mM Zn^{2+}, 20 mM Mn^{2+}, and 40 mM Mg^{2+}. Each of the five deoxyribozymes was evaluated with combinations of these metal ions to determine the metal ions required for catalysis. The presence of all three divalent metal ions was optimal for all of the deoxyribozymes.

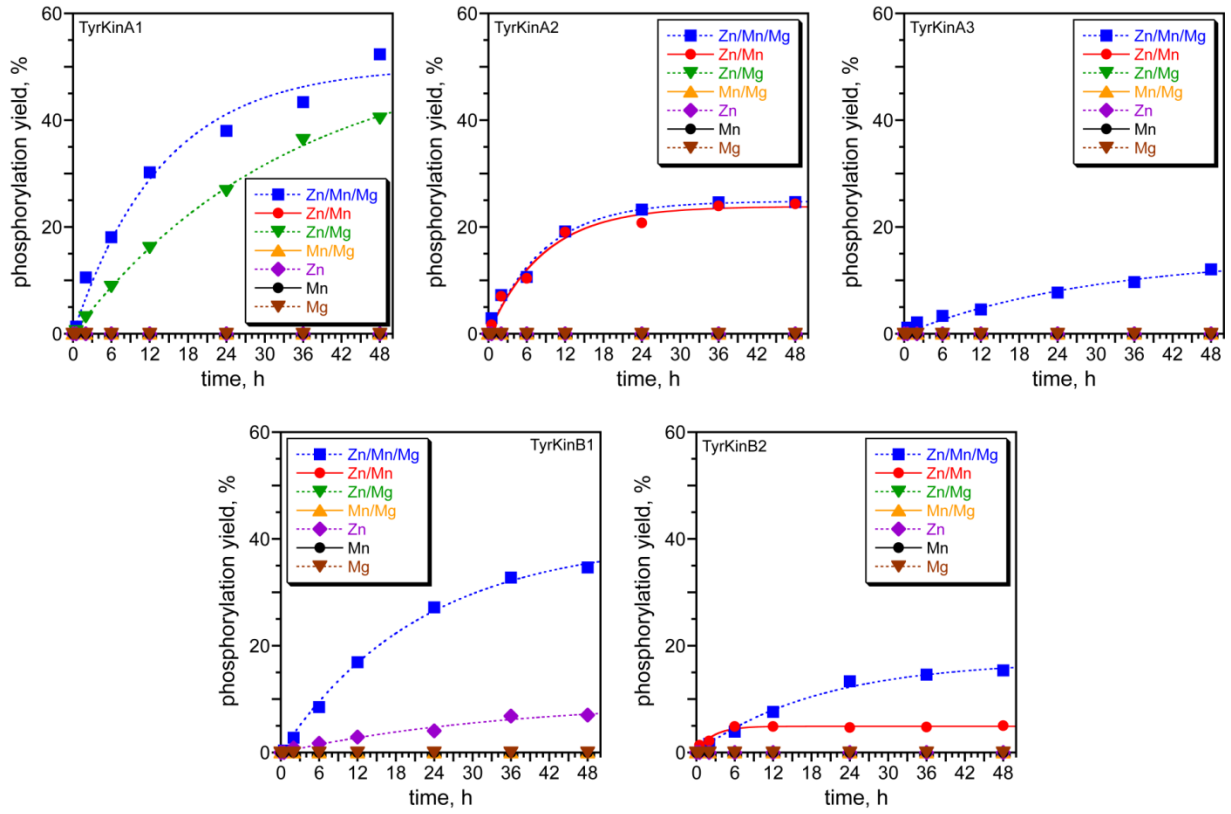


Figure 3.10. Metal dependence of the peptide-sequence-selective tyrosine kinase deoxyribozymes. Each of the five deoxyribozymes was incubated in 70 mM HEPES, pH 7.5, 150 mM NaCl and combinations of 1 mM $ZnCl_2$, 20 mM $MnCl_2$, and 40 mM $MgCl_2$ as indicated at 37 °C. No activity was observed when no divalent metal ions were included.

The peptide-sequence dependence for each deoxyribozyme was analyzed further by testing peptide-sequence mutations (Figure 3.11). As expected for tyrosine phosphorylation, mutation of the tyrosine residue to phenylalanine abolished all kinase activity. Mutation of the tyrosine residue to serine also abolished kinase activity. Therefore, all deoxyribozymes specifically phosphorylate tyrosine. To identify the peptide sequence motif required for kinase activity, each non-alanine amino acid residue flanking the tyrosine within the peptide substrate CADPYDQS was systematically and individually replaced with alanine. Interestingly, for the three highest-yielding deoxyribozymes, three different peptide sequence motifs were identified. For TyrKinA1, TyrKinA2, and TyrKinB1, the respective peptide sequence motifs are DPYD, DPY, and YD. The motifs of the other two deoxyribozymes could not be confidently assigned

due to their low phosphorylation yields with the parent peptide substrate. The three highest-yielding kinase deoxyribozymes were evaluated for their ability to phosphorylate an untethered peptide substrate. Unfortunately, phosphorylation of the CADPYDQS peptide was not observed by MALDI mass spectrometry for any of the deoxyribozymes, despite the tight interaction between the peptide substrate and DNA enzyme leading to peptide-sequence selectivity.

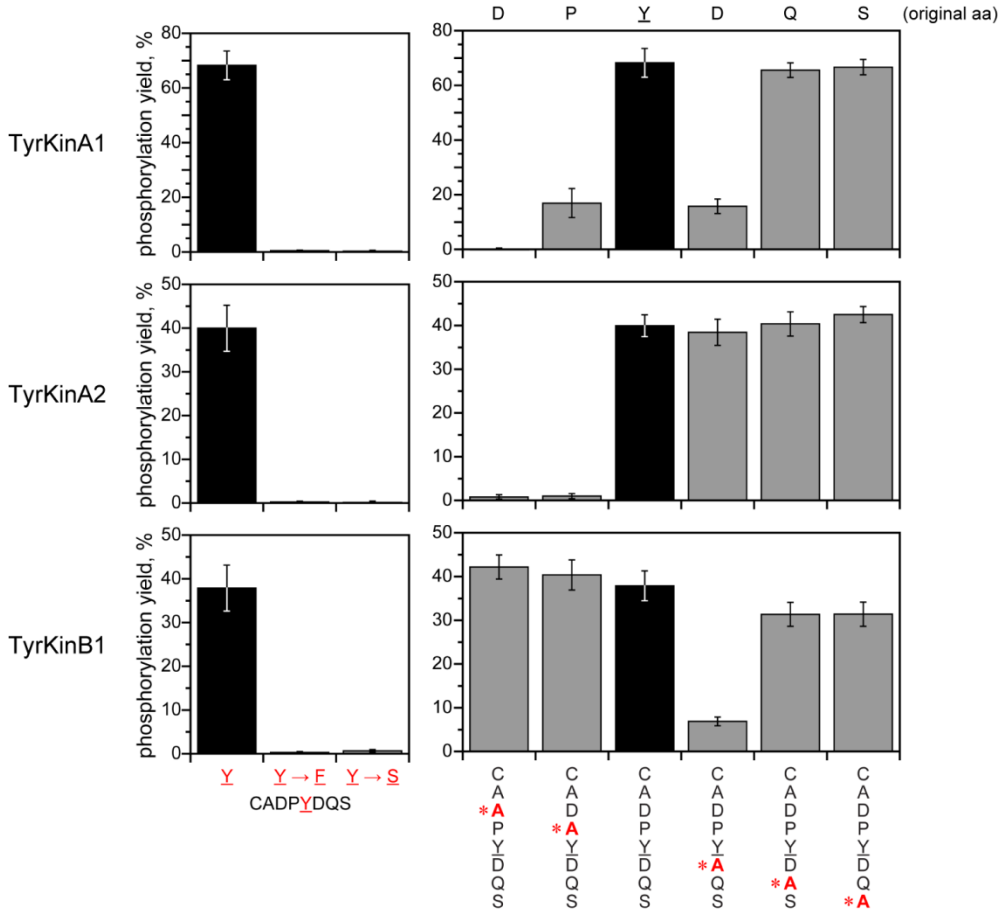


Figure 3.11. Peptide substrate mutations to determine the peptide sequence requirements of the deoxyribozymes. Phosphorylation yields at 48 h are shown for TyrKinA1, TyrKinA2, and TyrKinB1. Each deoxyribozyme was assayed with the series of DNA-anchored peptide substrates indicated (n = 5-8 for original sequence peptides and n = 3 for mutant peptides, mean ± sd). Figure reprinted with permission from Ref. 47.

Efforts by Puzhou Wang for DNA-catalyzed tyrosine adenylylation using diverse peptide sequences during the selection experiments led to the identification of the peptide motif-specific

deoxyribozyme DzAz2.[52] Separate selection experiments were performed with four unique peptide sequences. Selections with three of the peptides led to the identification of only peptide-sequence general deoxyribozymes. Four of the five deoxyribozymes identified with the fourth peptide sequence, CLQTYPRT, were general or partially selective, and only one deoxyribozyme DzAz2 was highly selective for the YPR motif. Similar to the peptide sequence-specific tyrosine kinase deoxyribozymes, the peptide sequence-specific tyrosine adenylylation deoxyribozyme DzAz2 did not catalyze the adenylylation of untethered peptide substrates.

3.3 Summary and Future Directions

Presenting a variety of side chains within the peptide substrate during the selection can lead to deoxyribozymes that are peptide-sequence-selective. In the context of a DNA-anchored peptide substrate, these results show that DNA catalysts can interact tightly with peptide substrates, even though strong interactions are not required for substrate binding. The peptide-sequence selectivity of the previous nucleopeptide-forming deoxyribozymes combined with the current tyrosine kinase and adenylylation deoxyribozymes indicate that using peptide substrates of varied sequence during in vitro selection fosters emergence of peptide-sequence-selective DNA catalysts. The observation of three distinct peptide sequence motifs for the different tyrosine kinase deoxyribozymes demonstrates that DNA enzymes can interact in different ways with the same peptide sequence. The short motifs of 2-4 amino acids suggest the practical utility of peptide-modifying deoxyribozymes.

These results show that DNA enzymes are capable of interacting specifically with peptide substrates, and peptide-sequence selectivity is not limited to one type of DNA-catalyzed reaction. While some peptide-sequence-specific deoxyribozymes were identified, not all peptide sequences used in selection led to the identification of highly selective deoxyribozymes. More experiments are required to understand the relationship between the peptide sequence and the selectivity of the emergent deoxyribozymes. No selection pressure was imposed to require

peptide-sequence selectivity. Therefore, the deoxyribozymes can interact with the peptide substrate, but tight interactions are not required. Kinase deoxyribozymes with peptide-sequence selectivity did not catalyze the phosphorylation of untethered peptide substrates, and this remains a focus for kinase deoxyribozymes.

3.4 Materials and Methods

3.4.1 Substrate Preparation Procedures

Oligonucleotides, peptides, and DNA-anchored peptide conjugates were prepared as described in Chapter 2.

Table 3.1: Oligonucleotide sequences of the pools, substrates, and primers used in the selection experiments.

oligonucleotide purpose	oligonucleotide sequence
Selections with Peptide Sequence CMTGYVAT	
DNA-HEG-CMTGYVAT substrate	GGATCAGGTTACTAATTAT-HEG-CMTGYVAT
pppRNA phosphoryl donor	pppGGAAUGAUAAUCUAUGG
forward primer for selection	CGAAATAGATTATCATTC
reverse primer for selection	$(AAC)_4$-HEG-CCATCAGGATCAGCT
reverse primer for cloning	TAATTAATTAATTACCCATCAGGATCAGCT
random pool for selection	CGAAATAGATTATCATTC-N_X-ATAATTAGTAACCTGAAGCTGATCCTGATGG
splint for ligation step during selection	ATAATTAGTAACCTGATCCCCATCAGGATCAGCTTCAGGTTACTAATTAT
Selections with Peptide Sequence CADPYDQS	
DNA-HEG-CADPYDQS substrate	GGATCCTGGATACAAATAT-HEG-CADPYDQS
pppRNA phosphoryl donor	pppGGAACAGGUUUAUACGG
forward primer for selection	CGAAGTATAAACCTGTTC
reverse primer for selection	$(AAC)_4$-HEG-CCATCAGGATCAGCT
reverse primer for cloning	TAATTAATTAATTACCCATCAGGATCAGCT
random pool for selection	CGAAGTATAAACCTGTTC-N_X-ATATTTGTATCCAGGAAGCTGATCCTGATGG
splint for ligation step during selection	ATATTTGTATCCAGGATCCCCATCAGGATCAGCTTCCTGGATACAAATAT

3.4.2 In Vitro Selection Procedures

Procedures for ligation, selection, and PCR were performed as described in Chapter 2.

Procedure for the preclear step in every third round. To prevent the enrichment of noncatalytic DNA sequences capable of binding either the pTyr antibody or protein A agarose, an additional preclear step was included in every third round beginning with round 1. After the ligation step and prior to the selection step, the preclear step was performed using the same procedure as the capture step (described below), except that the wash fraction was collected and taken onward to the selection step.

Procedure for phosphotyrosine antibody capture step in all rounds. Low-retention microcentrifuge tubes were treated by allowing to stand overnight with 2 mL of 5% (w/v) poly(ethylene glycol) (PEG, average molecular weight 3,350) and rinsing with 3 × 1 mL of water. All centrifugation steps were performed at 1000×*g* for 1 min. After the selection step (37 °C, 14 h), the sample was precipitated with ethanol and redissolved in a PEG-treated tube to 10 µL containing 25 mM Tris, pH 7.2, 150 mM NaCl, 2 mM EDTA, 0.01% Triton X-100, 50 pmol DNA oligonucleotide $(AAC)_{20}$ [as a sacrificial oligo to suppress adhesion of DNA pool molecules], 2 µg pTyr antibody (Thermo Scientific Pierce PY20). To this sample was added 25 µL of protein A agarose suspension (Thermo Scientific Pierce), previously buffer-exchanged to water by centrifugation and washing (3 × 500 µL). Protein A binds to the constant Fc region of the PY20 antibody. The 35 µL sample was shaken on a vortexer at the lowest setting for 1 h at room temperature, transferred to a spin column, centrifuged, and washed with 3 × 50 µL of 25 mM Tris, pH 7.2, and 150 mM NaCl. The washes were combined into a single wash fraction. At this point, the catalytic DNA pool sequences that phosphorylated their attached peptide were bound to the agarose-immobilized antibody, whereas other DNA sequences had been washed away. The bound sequences were centrifugally eluted with 3 × 50 µL of 8 M urea, which were combined into a single elution fraction (each portion of urea solution was allowed to stand in the spin column for 5 min at room temperature before centrifugation). Scintillation counting of the wash and elution fractions was used to quantify selection activity. The elution fraction was precipitated with ethanol and taken onward to the PCR step of the in vitro selection procedure.

3.4.3 Cloning and Screening

Cloning and screening were performed as described in Chapter 2.

Table 3.2: Names and clone numbers of deoxyribozymes identified from the RN1, RP1, and RJ1 selection experiments.

deoxyribozyme	clone
TyrKinA1	6RN101
TyrKinA2	6RN109
TyrKinA3	6RN108
TyrKinB1	7RP108
TyrKinB2	7RP110
TyrKinC1	9RJ101

3.4.4 Deoxyribozyme Activity Assay Procedure

The DNA-anchored hexapeptide substrate was 5′-^{32}P-radiolabeled using γ-^{32}P-ATP and T4 polynucleotide kinase (Fermentas), using 10× buffer that lacks DTT (500 mM Tris, pH 7.6, 100 mM $MgCl_2$, and 1 mM spermidine). A 10 µL sample containing 0.25 pmol of 5′-^{32}P-radiolabeled DNA-anchored peptide substrate, 10 pmol of deoxyribozyme, and 20 pmol of pppRNA phosphoryl donor was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The DNA-catalyzed phosphorylation reaction was initiated by bringing the sample to 20 µL total volume containing 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl. The Mn^{2+} was added from a 10× stock solution containing 200 mM $MnCl_2$. The Zn^{2+} was added from a 10× stock solution containing 10 mM $ZnCl_2$, 10 mM HNO_3, and 100 mM HEPES at pH 7.5; this stock solution was freshly prepared from a 200× stock of 100 mM $ZnCl_2$ in 200 mM HNO_3. The metal ion stocks were added last to the final sample, which was divided into 2 µL aliquots that were all incubated at 37 °C. At appropriate time points, 2 µL aliquots were quenched with 5 µL stop solution (80% formamide, 1× TBE [89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3], 50 mM EDTA, 0.025% bromophenol blue, 0.025% xylene cyanol). For assays with the DNA-anchored CKVIYDFI substrate, in addition to the 5′-^{32}P-radiolabeled substrate, 20 pmol of

unrelated 5′-phosphorylated DNA oligonucleotide and 1 nmol of free peptide were added to suppress loss of signal due to nonspecific binding of the substrate to the microcentrifuge tube. Samples were separated by 20% PAGE and quantified with a PhosphorImager. Values of k_{obs} were obtained by fitting the yield versus time data directly to first-order kinetics; i.e., yield $Y = Y_{max} \cdot (1 - e^{-kt})$, where $k = k_{obs}$ and Y_{max} is the final yield. Errors in k_{obs} values were calculated as the standard deviation from the indicated number of independent determinations.

3.4.5 Mass Spectrometry

The DNA-anchored CAAY^{P}AA phosphorylation product was prepared from a 10 μL sample containing 300 pmol of DNA-anchored HEG-tethered CAAY^{OH}AA substrate, 350 pmol of deoxyribozyme, and 400 pmol of pppRNA phosphoryl donor, which were annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The DNA-catalyzed phosphorylation reaction was initiated by bringing the sample to 20 μL total volume containing 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl. The sample was incubated at 37 °C for 14 h. The sample was desalted by Millipore C_{18} ZipTip and analyzed by MALDI mass spectrometry (Bruker UltrafleXtreme; matrix 3-hydroxypicolinic acid).

3.5 References

(1) Shahbazian, M. D.; Grunstein, M. Functions of site-specific histone acetylation and deacetylation. *Annu. Rev. Biochem.* **2007**, *76*, 75-100.

(2) Tiganis, T.; Bennett, A. M. Protein tyrosine phosphatase function: the substrate perspective. *Biochem. J.* **2007**, *402*, 1-15.

(3) Shi, Y. Serine/threonine phosphatases: mechanism through structure. *Cell* **2009**, *139*, 468-484.

(4) Tarrant, M. K.; Cole, P. A. The chemical biology of protein phosphorylation. *Annu. Rev. Biochem.* **2009**, *78*, 797-825.

(5) Moremen, K. W.; Tiemeyer, M.; Nairn, A. V. Vertebrate protein glycosylation: diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 448-462.

(6) Moore, K. E.; Gozani, O. An unexpected journey: lysine methylation across the proteome. *Biochim. Biophys. Acta* **2014**, *1839*, 1395-1403.

(7) Kemp, B. E.; Pearson, R. B. Protein kinase recognition sequence motifs. *Trends Biochem. Sci.* **1990**, *15*, 342-346.

(8) Flaswinkel, H.; Barner, M.; Reth, M. The tyrosine activation motif as a target of protein tyrosine kinases and SH2 domains. *Semin. Immunol.* **1995**, *7*, 21-27.

(9) Dutta, R.; Qin, L.; Inouye, M. Histidine kinases: diversity of domain organization. *Mol. Microbiol.* **1999**, *34*, 633-640.

(10) Meggio, F.; Pinna, L. A. One-thousand-and-one substrates of protein kinase CK2? *FASEB J.* **2003**, *17*, 349-368.

(11) Drennan, D.; Ryazanov, A. G. Alpha-kinases: analysis of the family and comparison with conventional protein kinases. *Prog. Biophys. Mol. Biol.* **2004**, *85*, 1-32.

(12) Ubersax, J. A.; Ferrell Jr, J. E. Mechanisms of specificity in protein phosphorylation. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 530-541.

(13) Peti, W.; Page, R. Molecular basis of MAP kinase regulation. *Protein Sci.* **2013**, *22*, 1698-1710.

(14) Hutti, J. E.; Jarrell, E. T.; Chang, J. D.; Abbott, D. W.; Storz, P.; Toker, A.; Cantley, L. C.; Turk, B. E. A rapid method for determining protein kinase phosphorylation specificity. *Nat. Meth.* **2004**, *1*, 27-29.

(15) Manke, I. A.; Yaffe, M. B. Screening Kinase Phosphorylation Motifs Using Peptide Libraries. *Cold Spring Harb. Protoc.* **2007**, *2007*, pdb.prot4640.

(16) Leung, G. C.; Murphy, J. M.; Briant, D.; Sicheri, F. Characterization of kinase target phosphorylation consensus motifs using peptide SPOT arrays. *Methods Mol. Biol.* **2009**, *570*, 187-195.

(17) Kreegipuu, A.; Blom, N.; Brunak, S. PhosphoBase, a database of phosphorylation sites: release 2.0. *Nucleic Acids Res.* **1999**, *27*, 237-239.

(18) Yaffe, M. B.; Leparc, G. G.; Lai, J.; Obata, T.; Volinia, S.; Cantley, L. C. A motif-based profile scanning approach for genome-wide prediction of signaling pathways. *Nat. Biotech.* **2001**, *19*, 348-353.

(19) Blom, N.; Sicheritz-Pontén, T.; Gupta, R.; Gammeltoft, S.; Brunak, S. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* **2004**, *4*, 1633-1649.

(20) Beausoleil, S. A.; Villen, J.; Gerber, S. A.; Rush, J.; Gygi, S. P. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat. Biotech.* **2006**, *24*, 1285-1292.

(21) Amanchy, R.; Periaswamy, B.; Mathivanan, S.; Reddy, R.; Tattikota, S. G.; Pandey, A. A curated compendium of phosphorylation motifs. *Nat. Biotech.* **2007**, *25*, 285-286.

(22) Durek, P.; Schmidt, R.; Heazlewood, J. L.; Jones, A.; MacLean, D.; Nagel, A.; Kersten, B.; Schulze, W. X. PhosPhAt: the Arabidopsis thaliana phosphorylation site database. An update. *Nucleic Acids Res.* **2010**, *38*, D828-D834.

(23) Dinkel, H.; Chica, C.; Via, A.; Gould, C. M.; Jensen, L. J.; Gibson, T. J.; Diella, F. Phospho.ELM: a database of phosphorylation sites—update 2011. *Nucleic Acids Res.* **2011**, *39*, D261-D267.

(24) Chen, Y. C.; Aguan, K.; Yang, C. W.; Wang, Y. T.; Pal, N. R.; Chung, I. F. Discovery of protein phosphorylation motifs through exploratory data analysis. *PLoS One* **2011**, *6*, e20025.

(25) Stephanopoulos, N.; Francis, M. B. Choosing an effective protein bioconjugation strategy. *Nat. Chem. Biol.* **2011**, *7*, 876-884.

(26) Renata, H.; Wang, Z. J.; Arnold, F. H. Expanding the enzyme universe: accessing non-natural reactions by mechanism-guided directed evolution. *Angew. Chem. Int. Ed.* **2015**, *54*, 3351-3367.

(27) Roiban, G. D.; Reetz, M. T. Expanding the toolbox of organic chemists: directed evolution of P450 monooxygenases as catalysts in regio- and stereoselective oxidative hydroxylation. *Chem. Commun.* **2015**, *51*, 2208-2224.

(28) Currin, A.; Swainston, N.; Day, P. J.; Kell, D. B. Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. *Chem. Soc. Rev.* **2015**, *44*, 1172-1239.

(29) Yoo, T. H.; Pogson, M.; Iverson, B. L.; Georgiou, G. Directed evolution of highly selective proteases by using a novel FACS-based screen that capitalizes on the p53 regulator MDM2. *ChemBioChem* **2012**, *13*, 649-653.

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

(31) Brody, E. N.; Gold, L. Aptamers as therapeutic and diagnostic agents. *J. Biotechnol.* **2000**, *74*, 5-13.

(32) Nimjee, S. M.; Rusconi, C. P.; Sullenger, B. A. Aptamers: an emerging class of therapeutics. *Annu. Rev. Med.* **2005**, *56*, 555-583.

(33) Shi, H.; Fan, X.; Sevilimedu, A.; Lis, J. T. RNA aptamers directed to discrete functional sites on a single protein structural domain. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 3742-3746.

(34) Cheung, Y. W.; Kwok, J.; Law, A. W.; Watt, R. M.; Kotaka, M.; Tanner, J. A. Structural basis for discriminatory recognition of Plasmodium lactate dehydrogenase by a DNA aptamer. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 15967-15972.

(35) Tan, W.; Donovan, M. J.; Jiang, J. Aptamers from cell-based selection for bioanalytical applications. *Chem. Rev.* **2013**, *113*, 2842-2862.

(36) Walters, R. D.; McSwiggen, D. T.; Goodrich, J. A.; Kugel, J. F. Selection and characterization of a DNA aptamer that can discriminate between cJun/cJun and cJun/cFos. *PLoS One* **2014**, *9*, e101015.

(37) Silverman, S. K. Pursuing DNA catalysts for protein modification. *Acc. Chem. Res.* **2015**, *48*, 1369-1379.

(38) 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.

(39) Sachdeva, A.; Chandra, M.; Chandrasekar, J.; Silverman, S. K. Covalent tagging of phosphorylated peptides by phosphate-specific deoxyribozymes. *ChemBioChem* **2012**, *13*, 654-657.

(40) Chu, C.; Wong, O.; Silverman, S. K. A generalizable DNA-catalyzed approach to peptide-nucleic acid conjugation. *ChemBioChem* **2014**, *15*, 1905-1910.

(41) Brandsen, B. M.; Velez, T. E.; Sachdeva, A.; Ibrahim, N. A.; Silverman, S. K. DNA-catalyzed lysine side chain modification. *Angew. Chem. Int. Ed.* **2014**, *53*, 9045-9050.

(42) Walsh, S. M.; Sachdeva, A.; Silverman, S. K. DNA catalysts with tyrosine kinase activity. *J. Am. Chem. Soc.* **2013**, *135*, 14928-14931.

(43) Dokukin, V.; Silverman, S. K. A modular tyrosine kinase deoxyribozyme with discrete aptamer and catalyst domains. *Chem. Commun.* **2014**, *50*, 9317-9320.

(44) Chandrasekar, J.; Silverman, S. K. Catalytic DNA with phosphatase activity. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5315-5320.

(45) Chandrasekar, J.; Wylder, A. C.; Silverman, S. K. Phosphoserine lyase deoxyribozymes: DNA-catalyzed formation of dehydroalanine residues in peptides. *J. Am. Chem. Soc.* **2015**, *137*, 9575-9578.

(46) Chu, C. C.; Silverman, S. K. Assessing histidine tags for recruiting deoxyribozymes to catalyze peptide and protein modification reactions. *Org. Biomol. Chem.* **2016**, *14*, 4697-4703.

(47) Walsh, S. M.; Konecki, S. N.; Silverman, S. K. Identification of sequence-selective tyrosine kinase deoxyribozymes. *J. Mol. Evol.* **2015**, *81*, 218-224.

(48) Raingeaud, J.; Gupta, S.; Rogers, J. S.; Dickens, M.; Han, J.; Ulevitch, R. J.; Davis, R. J. Pro-inflammatory cytokines and environmental stress cause p38 mitogen-activated protein kinase activation by dual phosphorylation on tyrosine and threonine. *J. Biol. Chem.* **1995**, *270*, 7420-7426.

(49) Salvador, J. M.; Mittelstadt, P. R.; Guszczynski, T.; Copeland, T. D.; Yamaguchi, H.; Appella, E.; Fornace, A. J., Jr.; Ashwell, J. D. Alternative p38 activation pathway mediated by T cell receptor-proximal tyrosine kinases. *Nat. Immunol.* **2005**, *6*, 390-395.

(50) Yokoyama, N.; Lougheed, J.; Miller, W. T. Phosphorylation of WASP by the Cdc42-associated kinase ACK1: dual hydroxyamino acid specificity in a tyrosine kinase. *J. Biol. Chem.* **2005**, *280*, 42219-42226.

(51) Velez, T. E.; Singh, J.; Xiao, Y.; Allen, E. C.; Wong, O. Y.; Chandra, M.; Kwon, S. C.; Silverman, S. K. Systematic evaluation of the dependence of deoxyribozyme catalysis on random region length. *ACS Comb. Sci.* **2012**, *14*, 680-687.

(52) Wang, P.; Silverman, S. K. DNA-catalyzed introduction of azide at tyrosine for peptide modification. *Angew. Chem. Int. Ed.* **2016**, *55*, 10052-10056.

Chapter 4: Efforts towards Untethered Peptide Reactivity[†]

4.1 Introduction

4.1.1 Deoxyribozymes with Untethered Peptide Reactivity

DNA enzymes with the ability to catalyze modification of untethered peptide and protein substrates will expand the applications of deoxyribozymes. These site-specifically modified peptides and proteins can then be studied to determine their biological function. Development of methods to identify DNA catalysts with the ability to catalyze the reaction of untethered peptides will be valuable in pursuit of peptide and protein modification.

4.1.1.1 In Vitro Selection with Tethered Peptide Substrates

Deoxyribozymes with phosphatase activity (hydrolysis of phosphorylated amino acid side chains) have been identified.[1] In vitro selection experiments to identify phosphatase deoxyribozymes were performed using HEG-tethered peptide substrates. Some of the phosphatase deoxyribozymes were able to dephosphorylate untethered peptide substrates. Therefore, they did not require the tether that was present during the in vitro selection. Similarly, nucleopeptide formation,[2] phosphoserine lyase[3] and tyrosine adenylylation[4] deoxyribozymes were identified from selection experiments in which HEG-tethered peptide substrates were used.

[†] University of Illinois graduate student Chi-Chih Chu performed selections for nucleopeptide-forming deoxyribozymes with untethered peptides and performed experiments to phosphorylate untethered, His-tagged tyrosine-containing peptide substrates.[9,33]

University of Illinois graduate student Cong Zhou identified deoxyribozymes containing modified nucleotides for amide hydrolysis.[15]

University of Illinois undergraduate student Stephanie N. Konecki performed selections to evolve nucleopeptide-forming deoxyribozymes into kinase deoxyribozymes and initiated selection experiments to identify kinase deoxyribozymes with carboxyl and amino modifications.

University of Illinois undergraduate student Tiyaporn Tangpradabkul performed selections to identify kinase deoxyribozymes with carboxyl and amino modifications.

For each of these reactions, at least one deoxyribozyme identified was able to catalyze the reaction of untethered peptide substrates.

Unfortunately, the kinase deoxyribozymes identified via a similar in vitro selection method using HEG-tethered peptide substrates all require the DNA-anchored HEG-tethered peptide substrate (Chapter 2, 3).[5] Similarly, deoxyribozymes for lysine modification also require tethered peptide substrates.[6] Therefore, the use of HEG-tethered peptide substrates during selection experiments may lead to deoxyribozymes able to react with untethered peptide substrates in some cases, but untethered peptide reactivity is not required to survive the selection experiment.

4.1.1.2 In Vitro Selection with Untethered Peptides

Initial efforts for DNA-catalyzed peptide modification focused on nucleopeptide formation, particularly the reaction of a nucleophilic tyrosine or serine side chain with the α-phosphate of a 5′-triphosphorylated RNA oligonucleotide, resulting in attachment of the RNA oligonucleotide to the amino acid side chain. The first efforts for DNA-catalyzed amino acid side chain modification used amino acids embedded within a DNA oligonucleotide, and this substrate was bound to the deoxyribozyme via a 3-helix junction.[7,8] Subsequent efforts moved to a less structured, open conformation in which the peptide substrate was tethered to a DNA anchor oligonucleotide.[2] Using this open conformation enabled peptides to be attached via different tether lengths. The resulting deoxyribozymes, while selected with a tethered substrate, were able to modify untethered peptide substrates, albeit only at high peptide concentrations (1 mM).

To improve upon this outcome, subsequent in vitro selection experiments were performed by Chih-Chi Chu using untethered peptide substrates (Figure 4.1).[9] The 5′-triphosphorylated RNA oligonucleotide (pppRNA) was ligated to the pool, and upon DNA-catalyzed nucleopeptide formation the peptide was attached to the RNA oligonucleotide that was connected to the DNA sequence that catalyzed its reaction. The catalytically active DNA sequences were then separated

by reaction of an azide on the N-terminus of the peptide with an alkyne-modified oligonucleotide, enabling a PAGE shift.

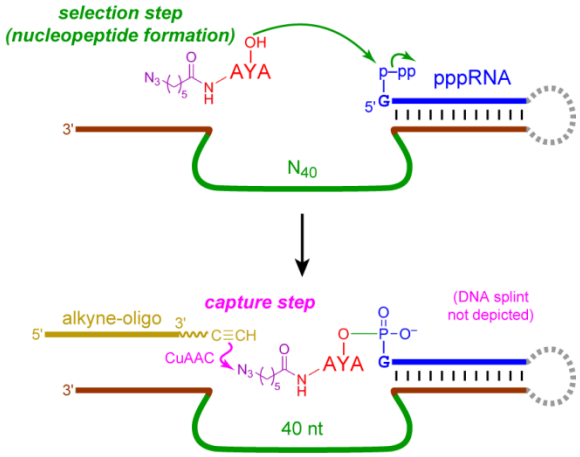


Figure 4.1. In vitro selection to identify deoxyribozymes with untethered peptide reactivity for nucleopeptide formation. The untethered peptide substrate has an N-terminal C_6-tethered azido group. During the selection step, the active DNA sequences catalyzed the attachment of the peptide to the pppRNA. The active DNA sequences were then captured by a DNA-splinted reaction of the azide with a 3′-alkyne oligonucleotide, resulting in PAGE-shift separation from the inactive DNA sequences. Figure adapted with permission from Ref. 9.

This in vitro selection method enabled identification of DNA catalysts with the ability to modify untethered peptides and have lower K_m(peptide) than the previously identified nucleopeptide-forming deoxyribozymes. In vitro selection of DNA enzymes requires that the desired product is attached to the DNA sequence that catalyzed its reaction. Therefore, this selection method using untethered peptide substrates cannot be applied to selection experiments for peptide phosphorylation. If the peptide is free in solution, successful phosphorylation will result in the transfer of a γ-phosphate from a phosphoryl donor to the peptide, and the phosphopeptide product will not remain associated with the catalytic DNA sequence. Therefore, the peptide substrate must be tethered to the DNA pool during selection experiments for kinase deoxyribozymes. The use of loosely tethered peptide substrates may be similar to untethered peptide substrates. While HEG tethers are sufficient to identify deoxyribozymes with untethered

peptide reactivity in some cases, longer tethers may be needed to more reliably mimic untethered peptides.

4.1.2 Modified Nucleotides Incorporated into Deoxyribozymes and Aptamers

4.1.2.1 Modifications for Improved Scope of DNA Catalysis

Proteins are the most common biological catalysts, and they contain diverse functional groups that are not present in nucleic acids. Incorporation of protein-like functional groups into deoxyribozymes has been achieved via modified nucleotides. DNA-catalyzed RNA cleavage via transesterification has been achieved with natural nucleotides. However, the incorporation of modified nucleotides reduces the divalent metal ion requirements of the resulting deoxyribozymes, which is important for biological applications (Figure 4.2).[10-14] DNA catalysts requiring reduced or no divalent metal ions were identified by incorporating modified nucleotides with a primary amino, imidzaolyl, or guanidinium group. Up to three modified nucleotides have been successfully incorporated into a DNA pool and used to identify deoxyribozymes that catalyze RNA cleavage.

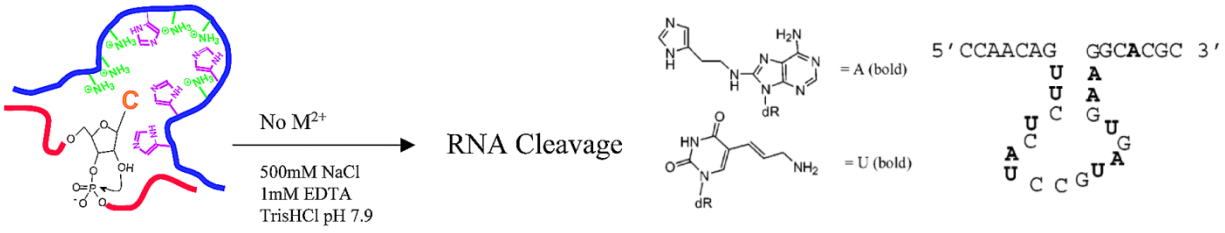


Figure 4.2. The Dz9$_{25}$-11 deoxyribozyme that catalyzes RNA cleavage in a divalent metal ion independent manner. The deoxyribozyme contains imidazolyl-modified dA and amino-modified dU. Figure adapted with permission from Ref. 10.

Separately, the incorporation of modified nucleotides has led to the identification of DNA catalysts that hydrolyze amide bonds, which was previously not possible with only natural DNA nucleotides (Figure 4.3).[15] Modified nucleotides were 5-substituted 2′-deoxyuridine derivatives

with a primary amino, carboxyl, hydroxyl, or imidazolyl group characteristic of lysine, aspartate/glutamate, serine, and histidine residues, respectively. Amide-cleaving deoxyribozymes were identified from separate selection experiments with the amino, carboxyl, and hydroxyl modifications. In this case, the incorporation of modified nucleotides led to the identification of deoxyribozymes to catalyze a reaction that was previously unsuccessful with natural nucleotides. Therefore, the modifications were required to identify DNA enzymes for this reaction.

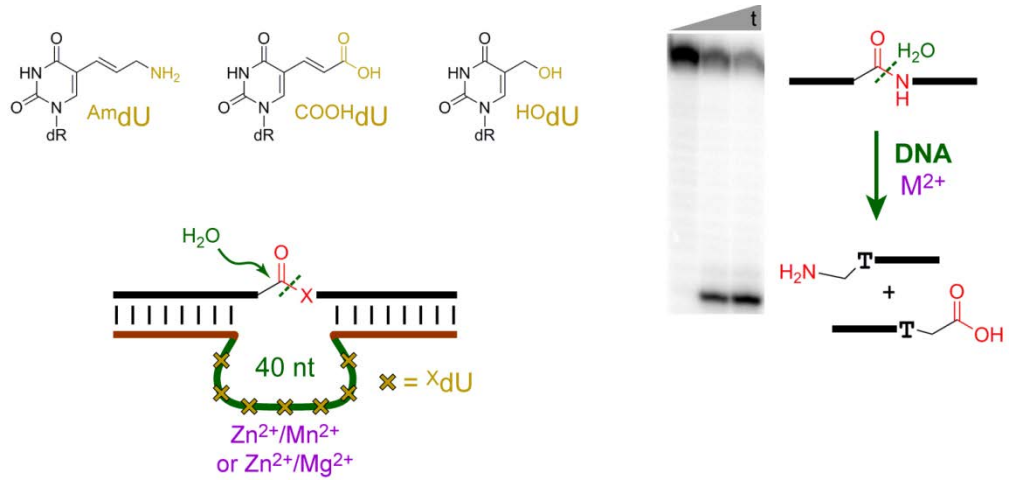


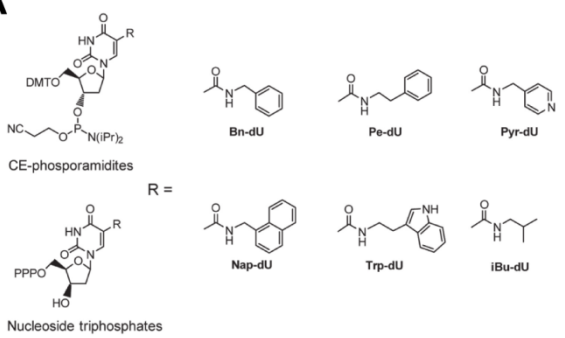
Figure 4.3. Identification of deoxyribozymes containing modified nucleotides for amide bond hydrolysis. The substrate had an amide bond embedded between two DNA oligonucleotide strands that base-paired to the fixed DNA sequences flanking the initially random region. Amino-modified dU (^{Am}dU), carboxyl-modified dU (^{COOH}dU), or hydroxyl-modified dU (^{HO}dU) was incorporated into the initially random region of the DNA pool in place of dT. Amide-cleaving DNA enzymes were identified from each selection experiment. PAGE image shows cleavage of the amide substrate by the ^{Am}dU-modified AmideAm1 deoxyribozyme. Figure adapted with permission from Ref. 15.

4.1.2.2 Modifications for Improved Protein Binding

DNA aptamers have been identified to bind to many protein substrates. To increase the binding affinity between the aptamer and protein substrate, DNA aptamers with hydrophobic modifications have been identified (Figure 4.4A).[16,17] In many cases unmodified aptamers have weak binding affinities (K_d > 100 nM), and upon incorporation of hydrophobic modifications, selection experiments result in the identification of aptamers with stronger binding affinities with

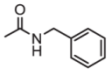
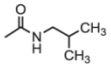
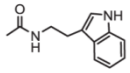
dissociation constants in the low nanomolar range. These aptamers are called slow off-rate modified aptamers (SOMAmers) due to their increased binding affinity. The SOMAmers were optimized by minimizing the sequence to only the required nucleotides and testing the identity of the hydrophobically modified nucleotides. SOMAmers for platelet-derived growth factor BB (PDGF-BB) and Interleukin-6 (IL-6) were identified from N_{40} initially random regions and respectively contained 9 and 10 benzyl-modified nucleotides (^{Bn}dU).[18,19] Optimization resulted in the 24 nt PDGF-BB SOMAmer SL5 with 8 hydrophobically modified residues and K_d = 0.02 nM, while the 32 nt IL-6 SOMAmer SL1025 contains 10 hydrophobically modified residues and K_d = 0.2 nM. SOMAmers SL5 and SL1025 were crystallized bound to their respective protein targets.[18,20] Crystal structures of the SOMAmer-protein complexes revealed that the hydrophobically modified residues within the aptamers are interacting via hydrophobic interactions with their respective protein substrates. The 8 modified residues of SL5 have hydrophobic interactions with 24 PDGF-BB residues, and the 10 modified residues of SL1025 have hydrophobic interactions with 11 IL-6 residues.

A



B

Table 1 SELEX library affinities (K_d, nmol/l) with unmodified and modified nucleotides

Target protein	Unmodified DNA	Bn-dUTP	iBu-dUTP	Trp-dUTP
	CH_3			
GA733-1 [a]	9	3	5	0.5
Osteoprotegerin [a]	40	5	9	0.2
4-1BB [a]	>100 [b]	6	>100	4
B7 [a]	>100	10	>100	7
B7-2 [a]	>100	>100	>100	6
CTLA-4 [a]	>100	>100	>100	1
sE-selectin [a]	>100	>100	>100	2
Fractalkine	>100	>100	>100	0.05
gp130, soluble [a]	>100	6	20	1
HMG-1	>100	>100	20	5
IR	>100	2	10	0.2
PAI-1	>100	0.4	0.9	0.2
P-Cadherin [a]	>100	4	5	3
sLeptin R [a]	>100	2	>100	0.5

[a]The protein used was expressed as a fusion to the F_c of human IgG_1. No detectable binding of the active library to an alternate F_c fusion protein was observed. [b]No detectable binding up to target protein concentration of 100 nmol/l was observed .

Figure 4.4. (A) Hydrophobic modifications at the 5-position of deoxyuridine. Side chain abbreviations: Bn, benzyl; Pe, 2-phenylethyl; Pyr, 4-pyridinylmethyl; Nap, 1-naphthylmethyl; Trp, 3-indole-2-ethyl; iBu, isobutyl. (B) Binding affinities of SELEX libraries selected for binding to the 14 different proteins indicated with either unmodified, Bn-modified, iBu-modified, or Trp-modified DNA. Binding affinities are shown (K_d, nM). Figure adapted with permission from Ref. 17.

Since hydrophobic modifications improve binding between DNA aptamers and protein substrates, incorporating hydrophobic modifications into DNA enzymes may similarly improve the binding affinity between the deoxyribozyme and the peptide or protein substrate. This improved binding may enable deoxyribozymes to catalyze the reaction of untethered peptide

substrates. Selection experiments were performed to identify aptamers for 14 protein targets with various hydrophobic modifications (Figure 4.4B).[17] Using natural nucleotides, aptamers with K_d <100 nM were identified for only 2 of the 14 protein targets. The incorporation of ^{Bn}dU or isobutyl-modified nucleotides (^{iBu}dU) led to tight-binding aptamers for 9 and 7 of the 14 protein targets, respectively. Incorporation of a tryptamine-modified nucleotide (^{Trp}dU) led to the identification of aptamers with K_d values of 0.05-7 nM with all 14 proteins. Based on the evaluation of these different hydrophobic modifications, the incorporation of ^{Trp}dU most often led to the identification of tight-binding aptamers. Therefore, the incorporation of ^{Trp}dU modifications into DNA enzymes may result in improved binding between the deoxyribozyme and peptide or protein substrates.

4.2 Results and Discussion

4.2.1 Use of Long Tethers to Mimic Untethered Peptides

To identify kinase deoxyribozymes, the phosphopeptide product must be connected to the deoxyribozyme that catalyzed the reaction, and therefore a tethered peptide substrate must be used. However, deoxyribozyme-catalyzed untethered peptide reactivity is desired. In vitro selection experiments performed can be performed with a long tether between the peptide substrate and the DNA anchor oligonucleotide base-paired to the binding arm flanking the initially random region of the DNA pool. If a sufficiently long tether is present, then the peptide substrate will essentially be untethered with respect to the deoxyribozyme.

4.2.1.1 Selection for New Phosphopeptide Capture Deoxyribozymes

To enable selection experiments for kinase deoxyribozymes with long tethers, a method to separate the catalytically active deoxyribozymes from the inactive DNA sequences is needed. The capture deoxyribozyme 8VP1 was identified to catalyze nucleopeptide formation between a phosphotyrosine residue in a peptide substrate and an RNA oligonucleotide.[21] The 8VP1-

catalyzed capture yield of HEG-tethered CAAY^{P}AA is 45%, and the 8VP1 capture method was used successfully in selection experiments performed with HEG-tethered peptide substrates.[5,22] The 8VP1 capture deoxyribozyme was evaluated for its ability to catalyze the reaction of phosphopeptides connected to the DNA anchor via longer tethers (Figure 4.5). When the tether length was increased to HEG_3, HEG_6, or HEG_{10}, the capture yield decreased to 20%, 15%, and 5%, respectively. Capture yield <15% is not reliable for use during in vitro selection experiments.

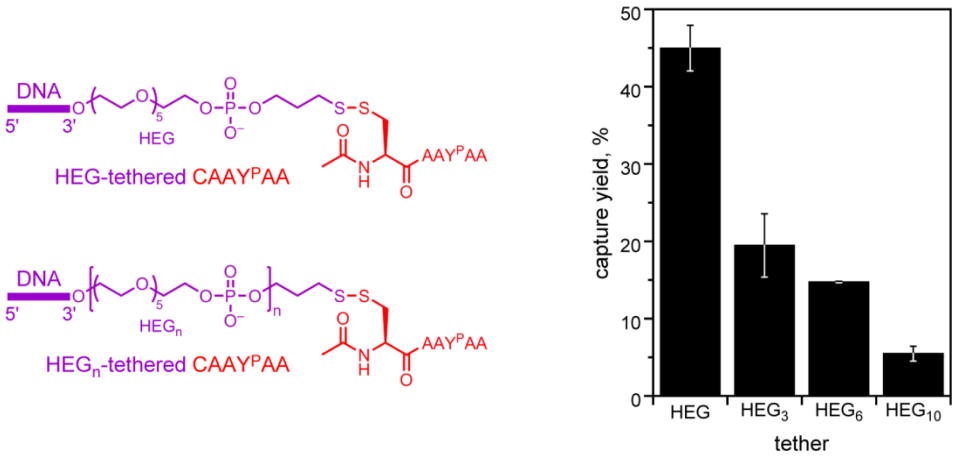


Figure 4.5. Dependence of the 8VP1 deoxyribozyme on the length of the tether between the DNA anchor and CAAY^{P}AA hexapeptide substrate. The structures of the tethers between the DNA anchor and peptide substrate are shown. Capture was evaluated for longer HEG_3, HEG_6, and HEG_{10}-tethered phosphopeptide substrates. The plot shows capture yield at t = 16 h (mean ± sd, n = 3). Incubation conditions were 50 mM HEPES, pH 7.5, 20 mM $MnCl_2$, and 150 mM NaCl at 37 °C.

To enable in vitro selection for tyrosine kinase deoxyribozymes using longer HEG_{10}-tethered peptide substrates, a new capture deoxyribozyme similar to 8VP1 was needed. 8VP1 was identified by in vitro selection using HEG-tethered CAAY^{P}AA with 20 mM $MnCl_2$ and 40 mM $MgCl_2$ at pH 7.5, but had slightly higher yield at 16 h with 20 mM $MnCl_2$ alone. Therefore, in vitro selection experiments were performed to identify capture deoxyribozymes that catalyze nucleopeptide formation between a HEG_6 or HEG_{10}-tethered CAAY^{P}AA peptide substrate and a 5′-triphosphorylated RNA oligonucleotide (Figure 4.6). The initial pool was either a random N_{40}

population or a partially randomized N_{40} population based on 8VP1 (25% randomization). Selection experiments were performed at pH 7.5 with the following conditions: HY1 = 20 mM $MnCl_2$, 40 mM $MgCl_2$; HZ1 = 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$; LA1 = 20 mM $MnCl_2$ alone. Unfortunately, none of these selection experiments led to the identification of deoxyribozymes capable of catalyzing nucleopeptide formation with the increased tether length. Since these selection experiments for new capture deoxyribozymes were unsuccessful, other phosphopeptide capture methods were evaluated.

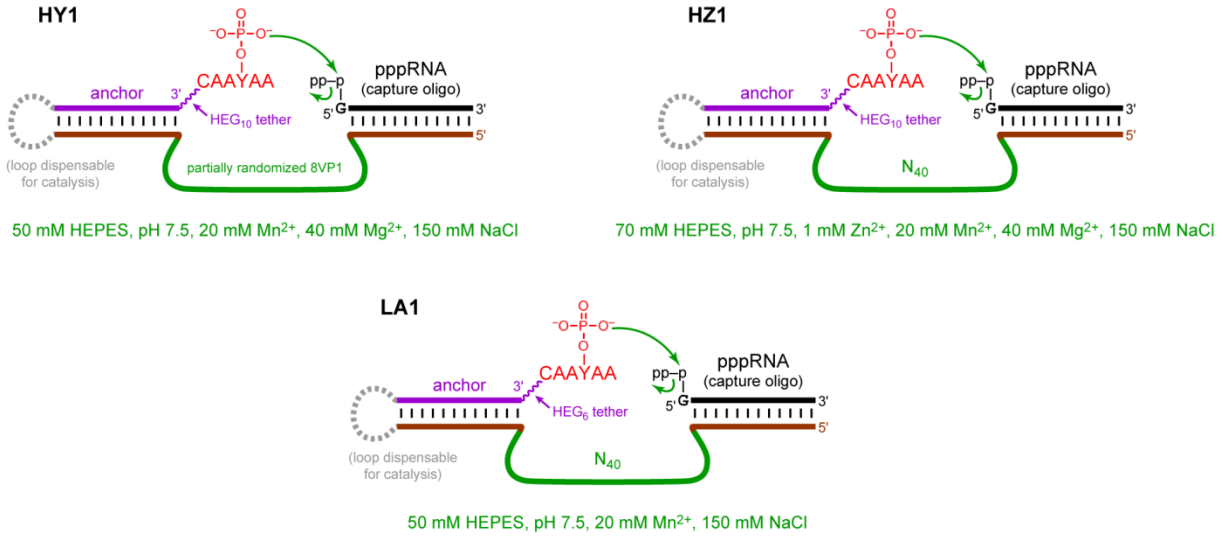


Figure 4.6. Design for the HY1, HZ1, and LA1 in vitro selection experiments to identify phosphotyrosine capture deoxyribozymes. In all selections, the CAAY^{P}AA peptide substrate was connected to a DNA anchor via a disulfide bond and a HEG_6 or HEG_{10} as indicated. The pool was partially randomized based on 8VP1 for HY1, and an N_{40} random pool for HZ1 and LA1. The selection step was performed at 37 °C for 14 h in the indicated conditions.

4.2.1.2 Selection for Tyrosine Kinase Deoxyribozymes Using HEG_6-Tethered Peptides

Based on the limitations of the 8VP1 deoxyribozyme, the HEG_6 tether is the longest that can be used with acceptable capture yield (~15%) during kinase deoxyribozyme selection experiments. Therefore, the KS1-KV1 selection experiments were designed using a HEG_6-tethered peptide substrates with pppRNA or 100 µM ATP as the phosphoryl donor (Figure 4.7). For KT1, the CAKKYKKA peptide contains multiple positively charged lysine (K) residues that

may increase binding interactions between the peptide substrate and the negatively charged DNA backbone of the deoxyribozyme. This increased binding affinity may result in untethered peptide reactivity. All selection experiments were performed at pH 7.5 with 1 mM $ZnCl_2$, 20 mM $MnCl_2$, and 40 mM $MgCl_2$. Selection experiments were continued through round 15, but no activity was observed. Based on these results, the inability to identify kinase deoxyribozymes may be because deoxyribozymes were not able to interact with the peptide substrate or the low 8VP1 capture yield did not enable sufficient enrichment of the catalytic DNA sequences.

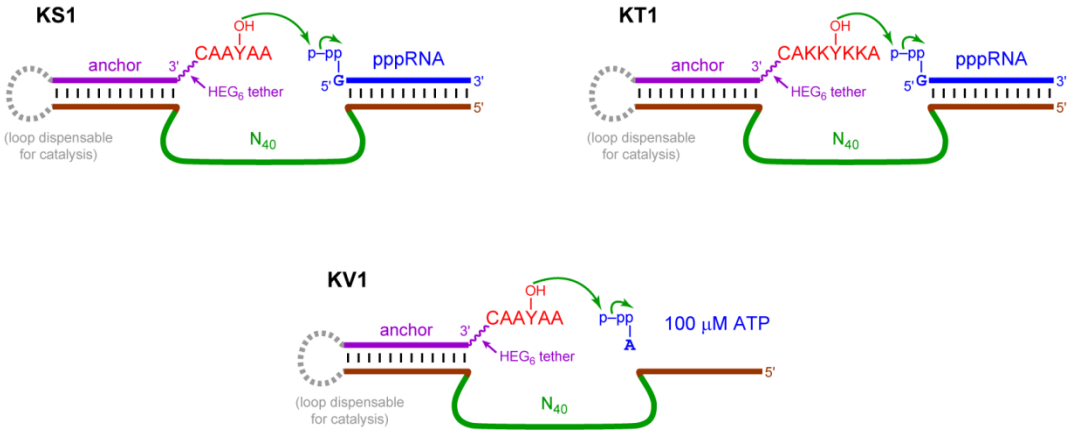


Figure 4.7. Design for the KS1-KV1 in vitro selection experiments to identify tyrosine kinase deoxyribozymes. In all selections, the initially random region was N_{40}, and the peptide substrate was connected to a DNA-anchor via a disulfide bond and a HEG_6 tether. The peptide substrate and phosphoryl donor are indicated. The selection step was performed in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h.

Untethered peptide reactivity is observed with some of the nucleopeptide-forming deoxyribozymes that have been identified.[2,9] While the use of untethered peptides during in vitro selection cannot be applied to the selection of kinase deoxyribozymes, nucleopeptide formation and phosphorylation are similar reactions. Phosphorylation is the reaction of the tyrosine side chain hydroxyl group with the γ-phosphate of a triphosphoryl donor, whereas nucleopeptide formation is reaction with the α-phosphate of the triphosphoryl donor (Figure 4.8). The nucleopeptide-forming deoxyribozymes 15MZ36, 8XJ105, and 11EM103 can catalyze the reaction of untethered peptide substrates (15MZ36 and 8XJ105 with AYA; 11EM103 with

GPYSGN). If these nucleopeptide-forming deoxyribozymes with untethered peptide reactivity are evolved to catalyze the reaction of the γ-phosphate instead of the α-phosphate, thus phosphorylating tyrosine within a peptide substrate, then they may retain their ability to catalyze the modification of untethered peptides.

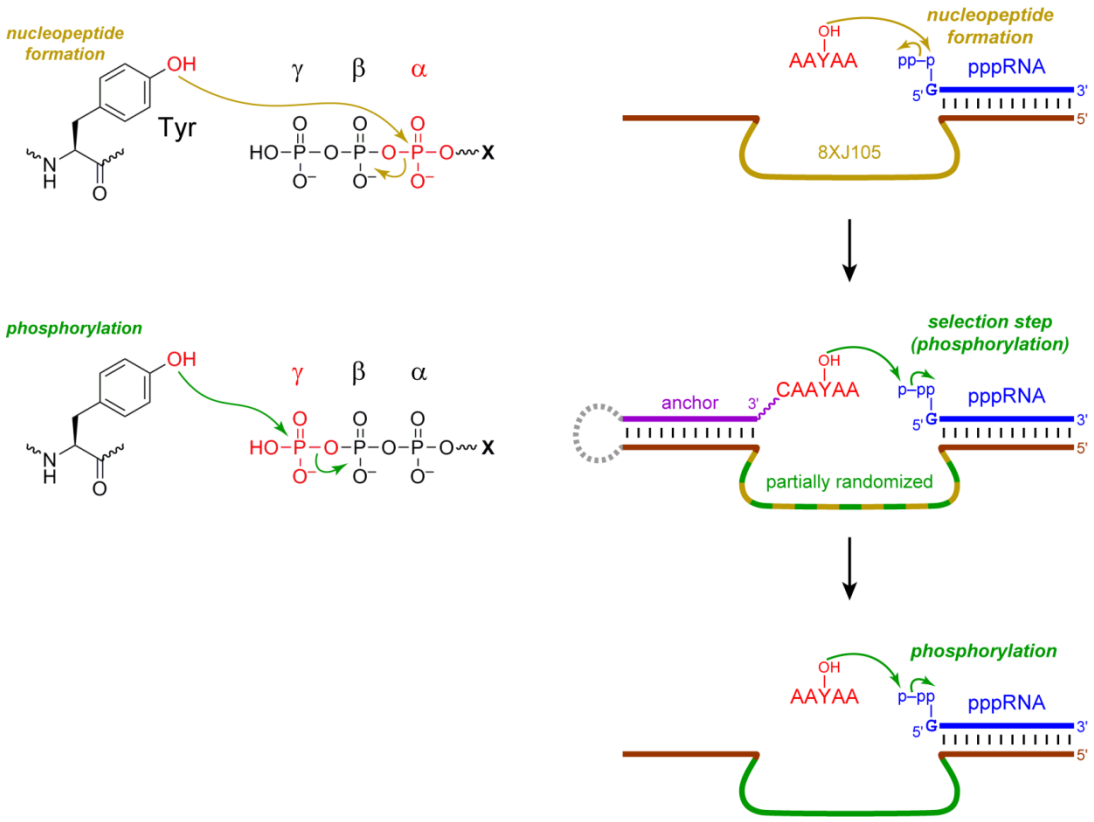


Figure 4.8. Reactions of phosphoryl donor substrates (5′-triphosphorylated RNA or NTP, where X = RNA or N). In nucleopeptide formation, the nucleophilic hydroxyl of the tyrosine side chain attacks the α-phosphate of the triphosphoryl donor. Phosphorylation results from nucleophilic attack of the hydroxyl of the tyrosine side chain at the γ-phosphate. To evolve nucleopeptide-forming deoxyribozymes, a partially randomized pool based on the nucleopeptide-forming deoxyribozyme sequence was prepared. In vitro selection was performed to select for kinase deoxyribozymes from the partially randomized pool.

Selection experiments for kinase deoxyribozymes were performed with a partially randomized pool (25% randomization) based on the sequence of the parent nucleopeptide-forming deoxyribozyme 15MZ36, 8XJ105, or 11EM103. Peptide substrates used by the parent deoxyribozyme were connected to the pool via a HEG_6 tether and a DNA anchor. Selection

conditions were 70 mM HEPES, pH7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h. In each round, active deoxyribozymes connected to the phosphorylated peptide products were captured using the 8VP1 capture deoxyribozyme with ~15% yield. Unfortunately, none of these selection experiments led to the identification of kinase deoxyribozymes. These results indicated one of three likely explanations. First, the population could not be evolved to phosphorylate peptides. Second, the DNA enzymes were unable to phosphorylate peptides connected via a longer HEG_6 tether. Third, the ~15% 8VP1 capture yield did not enable sufficient enrichment of the catalytically active DNA sequences.

4.2.1.3 Selection for Tyrosine Kinase Deoxyribozymes Using HEG_{10}-Tethered Peptides

The 8VP1 capture deoxyribozyme exhibits a decrease in capture yield as the length of the tether between the DNA-anchor and peptide substrate increases (Figure 4.5). Therefore, to use longer tethers in kinase deoxyribozyme selection experiments, a different capture method was required. The phosphotyrosine antibody capture method is independent of the length of the tether between the peptide substrate and DNA anchor (Chapter 3), new kinase deoxyribozyme selection experiments NR1-NZ1 were designed with DNA-anchored HEG_{10}-tethered CAAYAA peptide substrates. Each of N_{40}, N_{60}, and N_{80} initially random region lengths were used with pppRNA as the phosphoryl donor, and N_{60} and N_{80} initially random region lengths were used with 100 μM ATP as the phosphoryl donor (Figure 4.9). In two additional selection experiments with either N_{60} or N_{80} initially random region lengths, the peptide substrate was alternately anchored to either the left or right binding arm in successive selection rounds, and 100 μM ATP was used as the phosphoryl donor. The goal of using the long HEG_{10} tether was to mimic an untethered, free peptide substrate. By alternating the peptide anchoring site, the peptide substrate was the only constant part of the substrate in all rounds. Any resulting deoxyribozymes will hopefully interact only with the peptide and catalyze the phosphorylation of untethered peptide substrates. The preclear step was performed every round. A control selection experiment was performed with a HEG-tethered peptide substrate to validate the new immunoprecipitation selection method. All

selections were performed in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$ and 150 mM NaCl at 37 °C for 14 h.

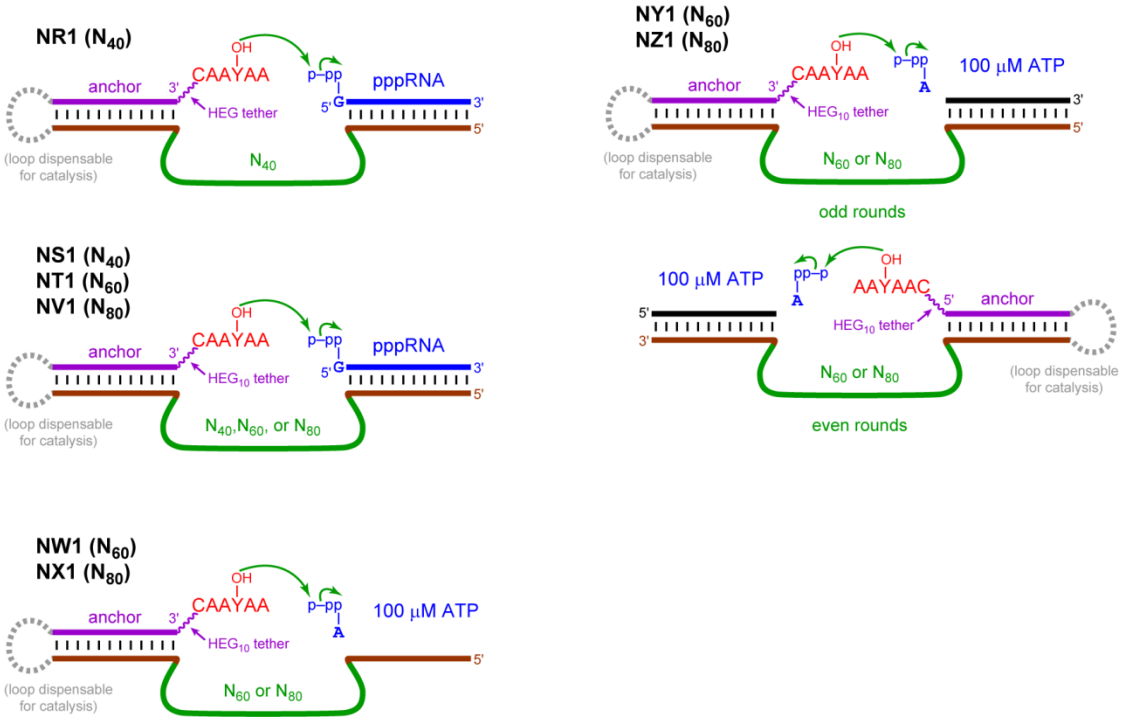


Figure 4.9. Design for the NR1-NZ1 in vitro selection experiments to identify tyrosine kinase deoxyribozymes. In all selections, the peptide substrate CAAYAA was connected to a DNA anchor via a disulfide bond and a HEG_{10} tether. NR1 was the sole exception with a shorter HEG tether instead. Initially random region lengths are indicated. The phosphoryl donor was pppRNA or 100 μM ATP for NR1-NV1 and NW1-NZ1, respectively. For the NY1-NZ1 selections, the DNA-anchored peptide substrate was bound to the opposite binding arm in successive selection rounds. The selection step was performed in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h.

No activity was observed for any of these selection experiments after 12 rounds. This was particularly surprising for the NR1 selection, because the same design had been successful when the 8VP1 capture method was used. To investigate the lack of NR1 activity, after nine rounds of selection, the preclear step was no longer included. The selection was continued eight more rounds, and 6% activity was observed in round 14 (Figure 4.10A). The round 14 and 15 pool was assayed in trans (without the substrate ligated to the pool), and phosphorylation of the DNA-anchored peptide substrate was observed. A parallel selection effort, NR1b, was performed in

which the preclear step was performed only in round 1. After nine rounds of selection 4% activity was observed, and phosphorylation of the DNA-anchored peptide substrate was verified in trans (Figure 4.10B). The active DNA sequences were enriched significantly slower when the preclear step was included for many rounds compared to the selection experiment in which only the first round included the preclear step. These results also show that after the preclear step was no longer included, DNA aptamer sequences did not take over the population and prevent the enrichment of catalytic DNA sequences. Therefore, in subsequent selection experiments using the phosphotyrosine immunoprecipitation capture method would not incorporate the preclear step every round.

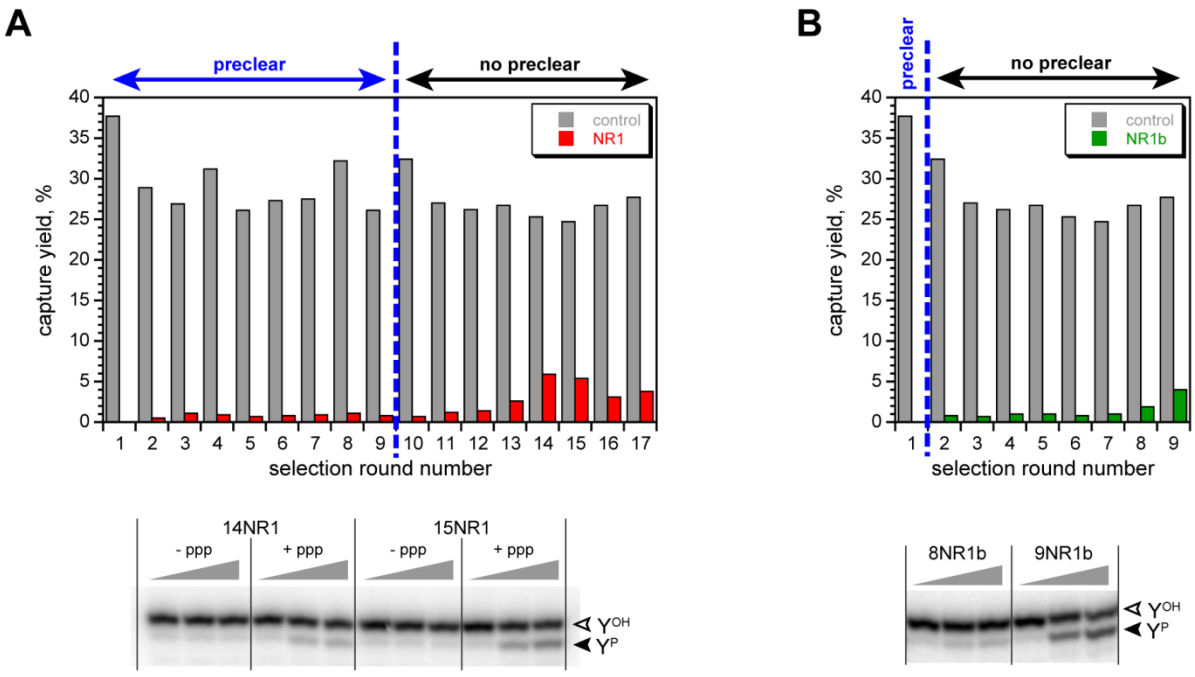


Figure 4.10. Selection progression and in trans assay of the pool for several variants of the NR1 selection experiment. In each round, "control" refers to the yield for the pTyr immunoprecipitation capture reaction using a CAAY^{P}AA substrate, and the selection experiment is indicated. (A) Progression of the initial NR1 selection experiment. The preclear step was performed during rounds 1-9 and not performed during rounds 10-17. The PAGE image shows the in trans assay of the NR1 pool at round 14 and 15 (t = 30 s, 2 h, 16 h). The pool catalyzes phosphorylation of the DNA-anchored peptide substrate only when the phosphoryl donor pppRNA (indicated as "ppp") is present. In 16 h, 14% and 28% phosphorylation yield were observed for the round 14 and 15 pools, respectively. (B) Progression of the NR1b selection experiment. The preclear step was performed only during round 1 and not performed during rounds 2-9. The PAGE image shows the in trans assay of the NR1b pool at round 8 and 9 (t = 30 s, 2 h, 16 h). In 16 h, 5% and 44% phosphorylation yield were observed for the round 8 and 9 pool, respectively.

Deoxyribozymes were not identified from any of the selection experiments in which tethers longer than HEG were used. Therefore, the tether dependence was evaluated for the nucleopeptide-forming 8XJ105 deoxyribozyme that was identified from a selection experiment using untethered peptide substrates (Figure 4.11). Use of a tethered peptide substrate was expected to result in an increased yield compared to an untethered peptide substrate, and the peptide substrate connected via a longer tether was expected to react similarly to the untethered peptide substrate. As expected with the untethered peptide substrate, 8XJ105 catalyzed

nucleopeptide formation in 30% yield. The nucleopeptide formation yield increased to 50% with the DNA-anchored HEG-tethered peptide substrate. Surprisingly, the HEG_6 and HEG_{10}-tethered peptide substrates were modified with only <5% yield. These results suggested that the initial hypothesis was incorrect; and increasing the length of the tether between the DNA anchor and peptide substrate did not mimic an untethered peptide substrate. The long tether was not inert as expected and instead prevented the 8XJ105 deoxyribozyme from catalyzing the reaction of the peptide substrate. These results are similar to those observed for the 15MZ36 deoxyribozyme as well.[2] 15MZ36 catalyzed nucleopeptide formation in 75%, 60%, and 10% yield with C_3-tethered, HEG-tethered, and untethered CYA, respectively. However, HEG_{10}-tethered CYA was not modified by 15MZ36. Combined, these results suggest that the use of tethers longer than HEG interfere with DNA-catalyzed peptide modification. On this basis, the use of tethers longer than HEG during selection experiments was discontinued.

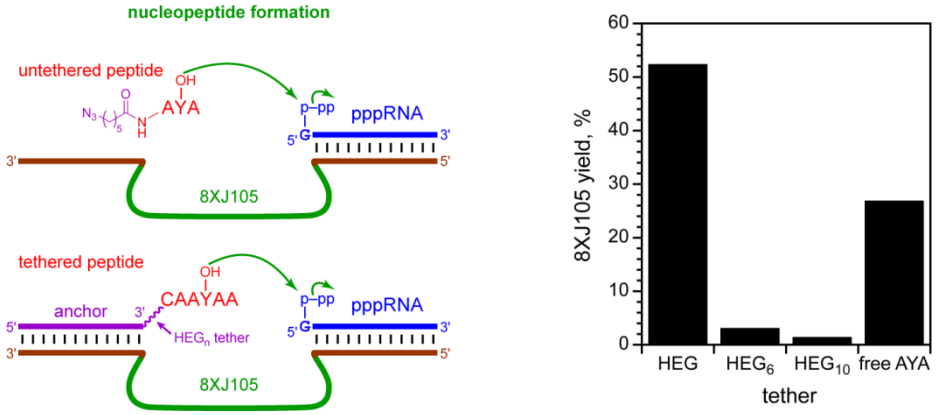


Figure 4.11. The tether dependence of the nucleopeptide-forming 8XJ105 deoxyribozyme. Untethered and tethered peptide substrates were evaluated. 8XJ105 successfully catalyzes the reaction of a free peptide and HEG-tethered peptide substrate. However, peptide substrates connected via a HEG_6 or HEG_{10} tether were not accepted as substrates.

4.2.2 Incorporation of Modified Nucleotides into Kinase Deoxyribozymes

Since long tethers between the peptide substrate and DNA anchor oligonucleotide does not mimic untethered peptide substrates, a new method to find deoxyribozymes with untethered peptide reactivity was needed. Hydrophobic modifications on nucleobases improve aptamer

binding affinity to proteins through an increased number of hydrophobic interactions.[16-20] Incorporation of these hydrophobic modifications into deoxyribozymes may increase the binding interaction between the peptide substrate and deoxyribozyme, leading to untethered peptide reactivity. Based on the hydrophobic aptamer data available, Trp-modified DNA led to aptamers with tight binding affinity to all 14 protein substrates evaluated (Figure 4.4).[17] Therefore, kinase deoxyribozymes with modified ^{Trp}dU were sought.

4.2.2.1 Synthesis and Purification of ^{Trp}dUTP

The Trp-modified dUTP was synthesized to be incorporated into the DNA pool (Figure 4.12). The synthesis of ^{Trp}dUTP was similar to the synthesis of ^{Im}dUTP performed by Cong Zhou. Tryptamine was condensed with the commercially available NHS-activated ^{COOH}dU [(*E*)-5-(2-carbomethoxyvinyl)-2′-deoxyuridine *N*-hydroxysuccinimide ester], and the Trp-modified nucleoside was chromatographically purified. For incorporation into the DNA pool by primer extension, the modified nucleotide triphosphate was needed. The ^{Trp}dU nucleoside was triphosphorylated and purified by HPLC to yield the 5′-triphosphorylated Trp-modified nucleotide. Primer extension incorporating the ^{Trp}dUTP product was successful, and an in vitro selection method to identify tyrosine kinase deoxyribozymes was developed to incorporate ^{Trp}dU into the DNA pool.

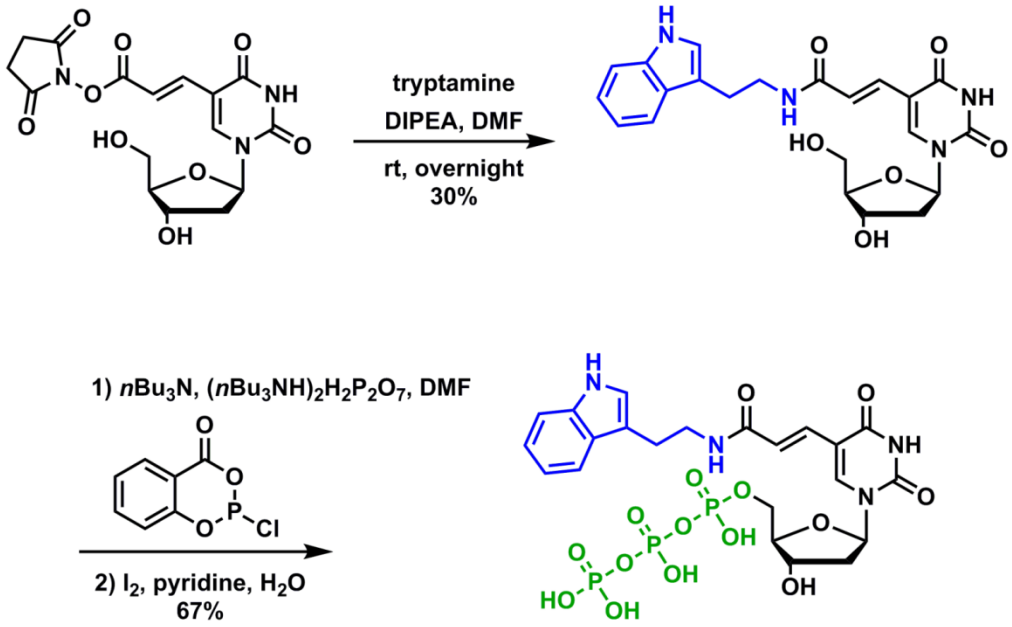


Figure 4.12. Synthesis of ^{Trp}dUTP. Commercially available NHS-activated ^{COOH}dU was condensed with tryptamine. The purified ^{Trp}dU was 5′-triphosphorylated and purified to yield ^{Trp}dUTP for incorporation into the DNA pool via primer extension.

4.2.2.2 Expression of KOD Polymerase

KOD polymerase from *Thermococcus kodakaraensis* has high accuracy and processivity in DNA replication.[23,24] KOD XL (Novagen) is a commercially available mixture of wild-type KOD polymerase and the KOD exonuclease minus (exo$^-$) mutant polymerase. KOD XL has been used during in vitro selection experiments for aptamers with hydrophobic modifications[16] and deoxyribozymes in which modified nucleotides are incorporated.[15] This includes the selection experiments performed in the Silverman lab by Cong Zhou to identify modified deoxyribozymes for amide hydrolysis that incorporated the modified nucleotides ^{Am}dU, ^{Im}dU, ^{COOH}dU, and ^{HO}dU. Of the polymerases tested, KOD XL polymerase also led to the highest primer extension yield with ^{Trp}dU. KOD XL is known to be a mixture of the wild-type KOD polymerase and KOD exo$^-$ polymerase, but the composition of the mixture is unknown. In the literature, KOD Dash polymerase is a 40:1 mixture of KOD exo$^-$ and wild-type KOD polymerase.[24] Therefore, the KOD polymerase and KOD exo$^-$ polymerase genes were synthesized, incorporated into plasmids, expressed, purified, and the optimal polymerase mixture determined for use in primer

extension and PCR during in vitro selection for deoxyribozymes that contain modified nucleotides.

Expression plasmids containing the polymerase genes were synthesized (Figure 4.13). The introns from the full-length KOD polymerase gene (GenBank D29671.2) were combined, and the gene was codon-optimized for expression in *E. coli*.[25] The full-length KOD polymerase gene was constructed from two IDT gBlock gene fragments that were double-stranded, sequence-verified genomic blocks. The first gBlock contained the first half of the gene and a NdeI restriction site at the beginning. The second gBlock contained the second half of the gene and a BamHI restriction site at the end. The two gBlocks were designed with a 30 bp complementary region, and the full-length KOD polymerase gene was prepared by PCR amplification. The KOD exo$^-$ polymerase contains the single point mutation N210D.[24] Site-directed mutagenesis of the full-length KOD polymerase gene was performed using mutagenic primers to generate the KOD exo$^-$ polymerase gene with the single nucleotide mutation A631G. This nucleotide mutation changes the AAT codon for N210 to the GAT codon for D210.

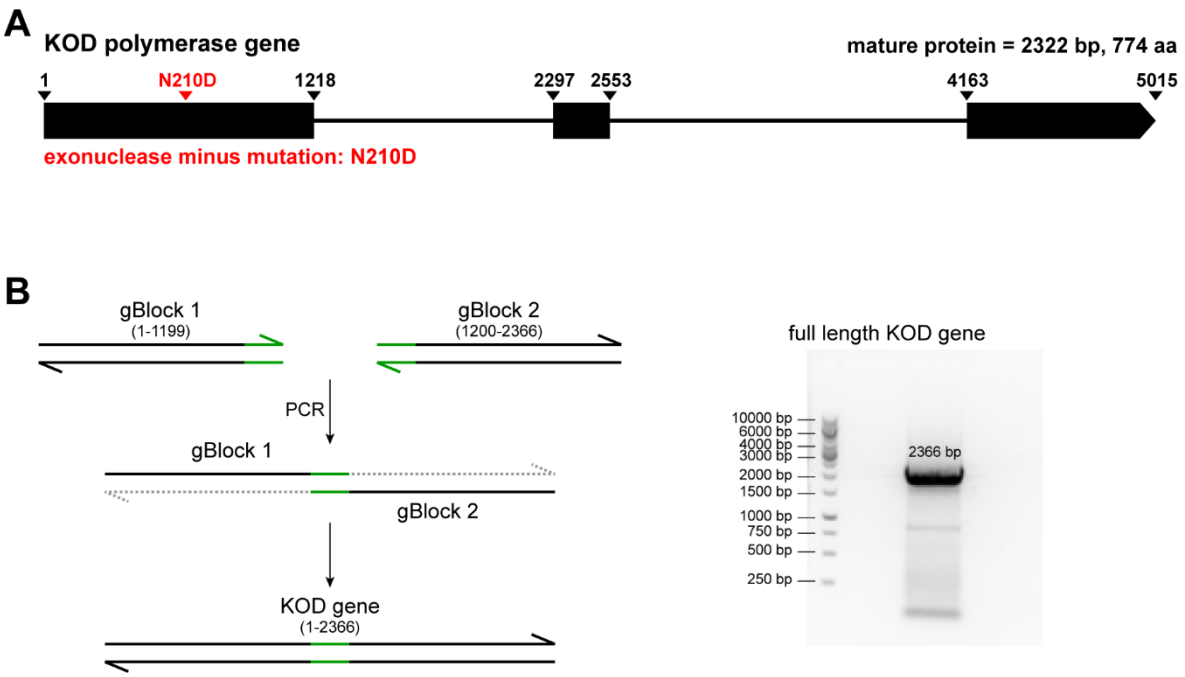


Figure 4.13. Synthesis of the KOD and KOD exo⁻ genes. (A) Gene map of the KOD polymerase gene from *Thermococcus kodakaraensis*. The exons are denoted by the thick boxes and the introns are the lines. The intron sequences were removed, resulting in a 2322 bp gene for the 774 amino acid protein. The KOD exo⁻ polymerase results from the single point mutation N210D. (B) The full-length gene was synthesized from two gBlocks, each containing half of the gene with a 30 bp overlapping region in the middle. PCR amplification resulted in the full-length KOD polymerase gene. The agarose gel shows the full-length PCR product of the KOD gene.

In the Silverman Lab, Pfu polymerase was expressed from a pET-16b vector containing the Pfu gene.[26] The Pfu gene was removed from the vector and the KOD polymerase genes were inserted. The pET-16b vector contains ampicillin and chloramphenicol resistance genes and encodes an N-terminal His_{10} tag on the protein. The Pfu gene was removed from the pET-16b vector by digestion with NdeI and BamHI restriction enzymes. The resulting cut sites were dephosphorylated with alkaline phosphatase to prevent religation, and the empty vector was purified on an agarose gel. The full-length KOD and KOD exo⁻ polymerase genes were digested with the same BamHI and NdeI restriction enzymes and purified on an agarose gel. The KOD polymerase genes were then ligated into the empty pET-16b vector and transformed into *E. coli* strain DH5α by heat shock. Individual colonies were cultured, and the plasmid was purified by

miniprep. Sequencing of the new plasmids verified that the KOD and KOD exo$^-$ polymerase genes were correct.

The KOD and KOD exo$^-$ polymerase plasmids were transformed into *E. coli* strain BL21 (DE3) pLysS by heat shock for expression. The cells were pelleted and lysed. The polymerase was purified by heating the cell lysate to 75 °C for 30 min, followed by centrifugation. The desired polymerase is heat-stable and remains in solution, while the other cell components are precipitated. The polymerases were stored in 50% glycerol at –20 °C. Both KOD and KOD exo$^-$ polymerase were used successfully in 30-cycle PCR with natural nucleotides. Therefore, both polymerases were used in primer extension from a random pool incorporating ^{Am}dU, ^{Im}dU, ^{COOH}dU, ^{HO}dU, or ^{Trp}dU. Each polymerase was used separately or as a mixture of the wild type and exo$^-$ mutant (Figure 4.14A). The 40:1 mixture of the KOD exo$^-$ and wild-type KOD resulted in primer extension yields comparable to the yields from the commercial KOD XL polymerase for all modified nucleotides evaluated. For in vitro selection experiments, the polymerase also needs to be able to copy a modified DNA template into natural nucleotides (Figure 4.14B). Primer extension of DNA templates with each of the five different modifications with wild-type KOD, KOD exo$^-$, or the 40:1 mixture of the KOD exo$^-$ and wild-type KOD resulted in yields similar to the yields from the commercial KOD XL polymerase. Therefore, the 40:1 mixture of the KOD exo$^-$ and KOD wild-type polymerase was comparable to the commercial KOD XL polymerase and was used during in vitro selection experiments for deoxyribozymes containing modified nucleotides.

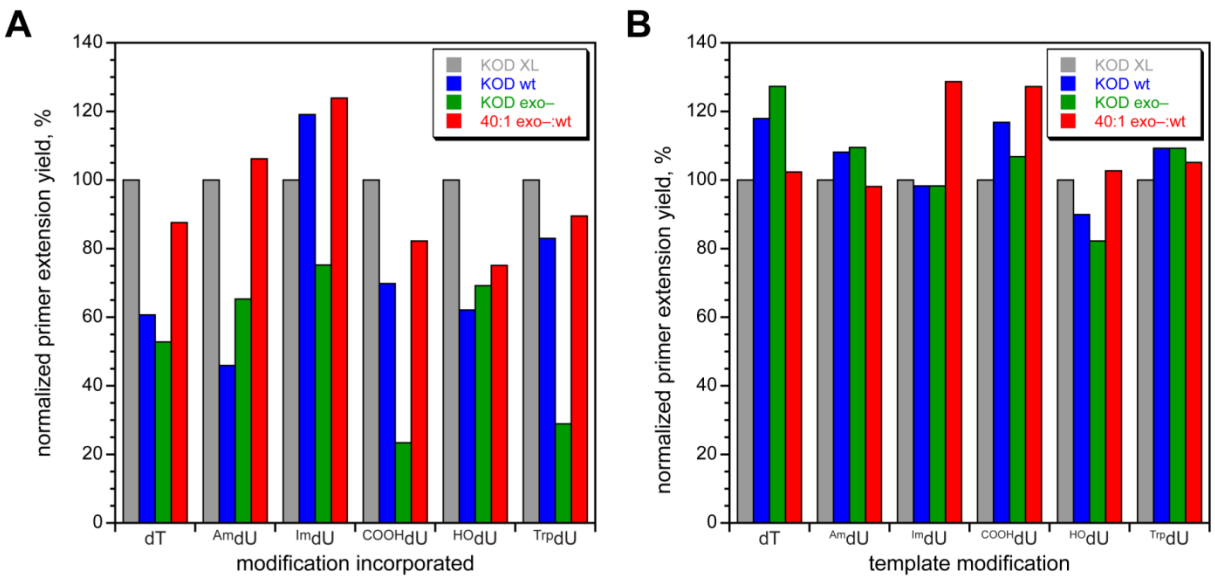


Figure 4.14. Polymerase activity of the purified wild-type KOD and KOD exo⁻ enzymes. KOD XL = commercial polymerase with a mixture of KOD and KOD exo⁻ (Novagen); KOD wt = wild-type KOD polymerase; KOD exo⁻ = KOD exonuclease minus mutant; 40:1 exo⁻:wt = 40:1 mixture of KOD exo⁻ and KOD wt. All yields are normalized to the primer extension yields with the commercial KOD XL polymerase. (A) Primer extension to incorporate the modified nucleotides indicated. (B) Primer extension from the modified DNA template incorporating natural nucleotides.

4.2.2.3 Selection for Kinase Deoxyribozymes Containing Hydrophobic Modifications

The in vitro selection strategy to identify kinase deoxyribozymes (Chapter 2, 3) was modified to enable use of the ^{Trp}dU-modified nucleotide (Figure 4.15). The initially random pool containing ^{Trp}dU was generated by primer extension from a reverse-complement random pool and was PAGE-purified. This pool was ligated to the DNA-anchored HEG-tethered tyrosine-containing peptide substrate, and the product was PAGE-purified. Every three rounds beginning with round 1, the preclear step was performed to remove DNA aptamers that bind to the pTyr antibody, protein A, or resin used during the capture step. DNA sequences that did not bind were separated and precipitated. In the selection step, the DNA sequences were incubated with a phosphoryl donor and divalent metal ions to enable catalysis. The sample was precipitated, and the phosphotyrosine immunoprecipitation capture step was performed to isolate the active DNA sequences connected to their phosphopeptide products. The 10-cycle PCR with natural nucleotides was altered to include an initial 30 min extension step to ensure that the polymerase

copied the modified deoxyribozymes, followed by 10 cycles of PCR. A 30-cycle PCR was performed to amplify the pool using all-natural dNTPs. The resulting reverse-complement strand was PAGE-purified, and primer extension from this template was performed incorporating the modified nucleotides into the pool for the next selection round.

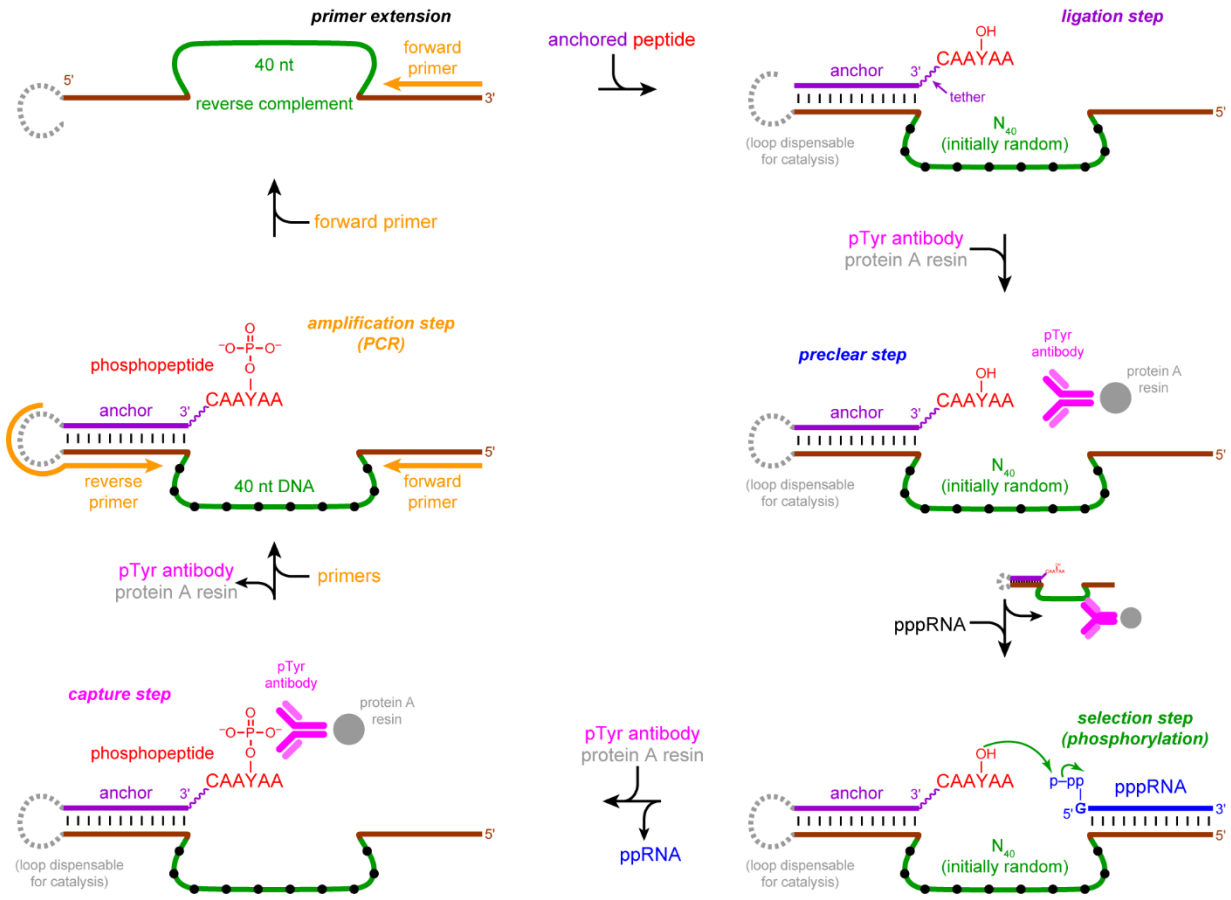


Figure 4.15. Diagram of the in vitro selection procedure to identify kinase deoxyribozymes containing modified nucleotides. The initially random pool was prepared by primer extension from a reverse-complement random pool. The DNA-anchored peptide substrate was ligated to the DNA pool. The subsequent preclear step was performed every third round, beginning in round 1, to remove DNA sequences that bind to components of the bead-based immunoprecipitation capture step. In the selection step, the DNA sequences were incubated with the pppRNA phosphoryl donor and divalent metal ions to enable DNA-catalyzed phosphorylation of the peptide substrate. In the subsequent capture step, the active DNA sequences connected to the phosphopepetide products were captured via immunoprecipitation with an immobilized pTyr antibody. The surviving catalytic DNA sequences were then PCR-amplified. PAGE purification was performed after each of the the PCR, primer extension, and ligation steps. The selection process was repeated until the catalytically active DNA sequences dominated the population.

All steps in the in vitro selection process were validated to ensure that ^{Trp}dU could be incorporated. A reverse-complement template was generated by 30-cycle PCR with natural nucleotides from the previous 1FS1 10-cycle PCR product (the FS1 selection performed by Cong Zhou was for amide bond hydrolysis with a ^{HO}dU-modified pool). Primer extension with ^{Trp}dU from the reverse-complement template was successful using KOD XL polymerase (Figure 4.16A). ^{Trp}dU was incorporated as well as the natural dT or ^{HO}dU control. The Trp-modified pool was purified and ligated to the DNA-anchored peptide substrate. Both an *n* splint (base-pairs exactly to the DNA pool binding arm and the DNA anchor of the substrate) and an *n* + 1 splint (which includes an extra nucleotide in the splint to account for an untemplated nucleotide incorporated at the 3′-end of the DNA pool) were evaluated (Figure 4.16B). KOD XL polymerase is a mixture of the wild-type enzyme with proofreading ability and the exo$^-$ mutant of the enzyme without proofreading capability. Therefore, the composition of the 3′-end of the DNA pool was unknown. Ligation with both splints was successful; and the *n* splint had slightly higher yield and was used during the selection experiments.

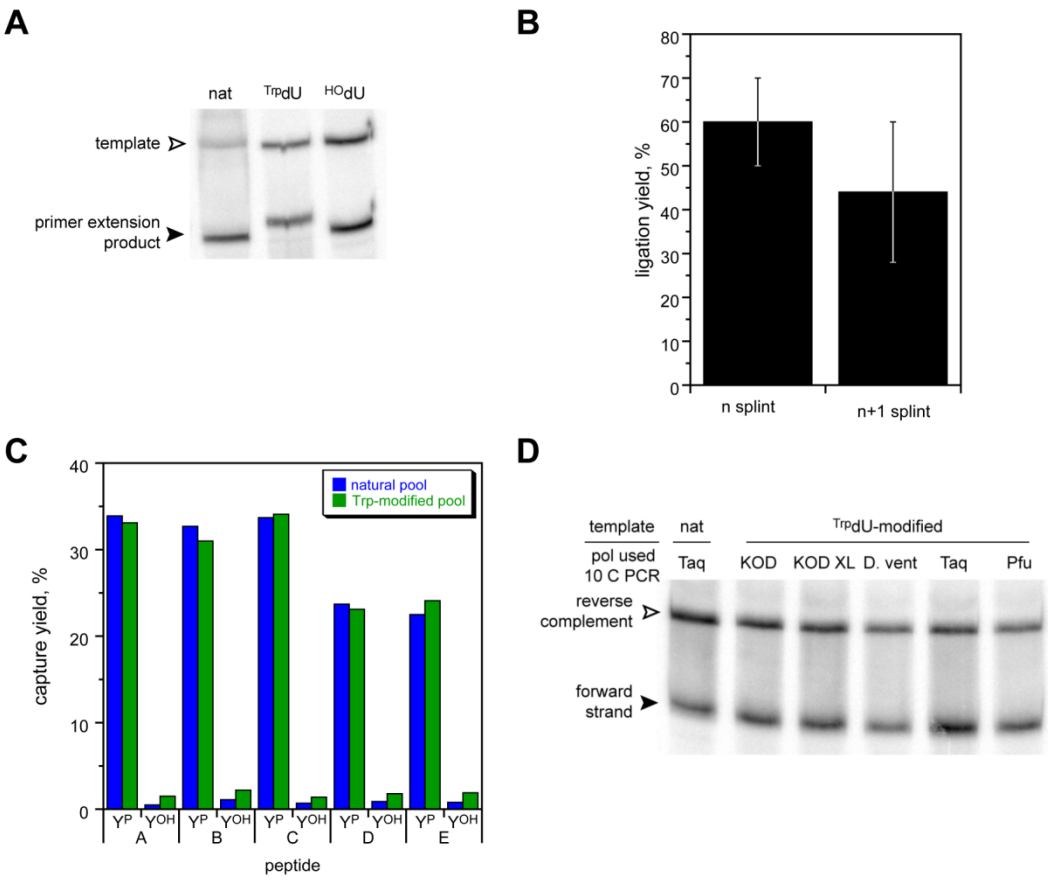


Figure 4.16. Evaluation of the use of Trp-modified pools during in vitro selection experiments. (A) Primer extension to incorporate ^{Trp}dU. The PAGE image shows the results of primer extension from a reverse-complement pool template incorporating natural (nat) nucleotides, ^{Trp}dU, or ^{HO}dU using KOD XL polymerase. (B) Ligation of the Trp-modified pool to the DNA-anchored peptide substrate. The plot shows the yield of ligation between the DNA-anchored peptide substrate and Trp-modified pool using either an *n* or *n* + 1 splint (mean ± sd, *n* = 3). (C) Phosphotyrosine immunoprecipitation capture yield of phosphopeptide products connected to either a natural DNA pool or a Trp-modified DNA pool. Peptide substrates with either pTyr (Y^P) or Tyr (Y^{OH}) were as follows: A = CMTG<u>Y</u>VAT; B = CKVI<u>Y</u>DFI; C = CEDI<u>Y</u>TPG; D = CARHTDDEMTG<u>Y</u>VATRWAYRAPE; E = CNPDWEIGEDI<u>Y</u>TPGKAGDALR. (D) PCR amplification of the Trp-modified DNA pool. The PAGE image shows the results of the subsequent round 30-cycle PCR with natural nucleotides using Pfu polymerase. The polymerase used during the previous 10-cycle PCR is indicated. The two DNA strands were separable because the reverse primer contained a nonamplifiable spacer.

The presence of Trp-modified DNA during the pTyr immunoprecipitation capture step was evaluated (Figure 4.16C). Incorporation of hydrophobic modifications into DNA can improve aptamer binding to proteins, and the capture step includes the use of two proteins, a pTyr antibody and protein A-coated resin. The capture yield of the phosphotyrosine-containing

peptide products was not affected by the presence of ^{Trp}dU in the connected DNA pool. A small increase in background capture of the tyrosine-containing peptide substrate was observed when the peptide was connected to a Trp-modified DNA pool (~2%) compared to a natural DNA pool (~1%). A small 2% background should not interfere with the selection experiments.

After the capture step, 10-cycle PCR was performed incorporating natural nucleotides based on the Trp-modified DNA template. Different polymerases were evaluated for use in 10-cycle PCR to determine their ability to copy the Trp-modified DNA template. Using the 10-cycle PCR product, a 30-cycle PCR with natural nucleotides and Pfu polymerase was performed to amplify the catalytic DNA sequences (Figure 4.16D). All polymerases used during 10-cycle PCR led to expected 30-cycle PCR product yields. For simplicity, KOD XL was used during 10-cycle PCR and Pfu polymerase was used for 30-cycle PCR during the selection experiments. Therefore, ^{Trp}dU was successfully incorporated, and all steps of the selection process were validated.

Selection experiments with ^{Trp}dU were designed with various peptide substrates (Figure 4.17). All selections used an N_{40} initially random region. The selection experiments used peptide substrates derived from natural proteins with an artificial cysteine residue at their N-terminus to enable disulfide conjugation to the DNA anchor oligonucleotide. The WP1 selection used the CMTG<u>Y</u>VAT peptide substrate corresponding to residues 179-185 of MAPK14/p38a (mitogen-activated protein kinase 14) containing the naturally phosphorylated Tyr182 residue.[27,28] The TyrKinC1 deoxyribozyme was previously identified to phosphorylate this peptide substrate in a partially selective manner, but TyrKinC1 was not highly selective and did not catalyze the phosphorylation of untethered peptide substrates (Chapter 3).[29] When new Trp-modified deoxyribozymes are identified, their sequence selectivity and untethered peptide reactivity will be compared to that of TyrKinC1. The WQ1 selection used the CKVI<u>Y</u>DFI peptide substrate corresponding to residues 253-259 of N-WASP (neuronal Wiskott-Aldrich Syndrome protein) where Tyr256 is naturally phosphorylated.[30] The previous selection experiment using natural nucleotides with this peptide substrate did not result in activity (Chapter 3). Therefore, the ^{Trp}dU

may enable DNA-catalyzed phosphorylation of this peptide sequence. Two selection experiments WR1-WS1, with either natural or Trp-modified nucleotides, used the new peptide substrate CEDIYTPG corresponding to residues 440-446 of CNP (cereus neutral protease).[31,32]

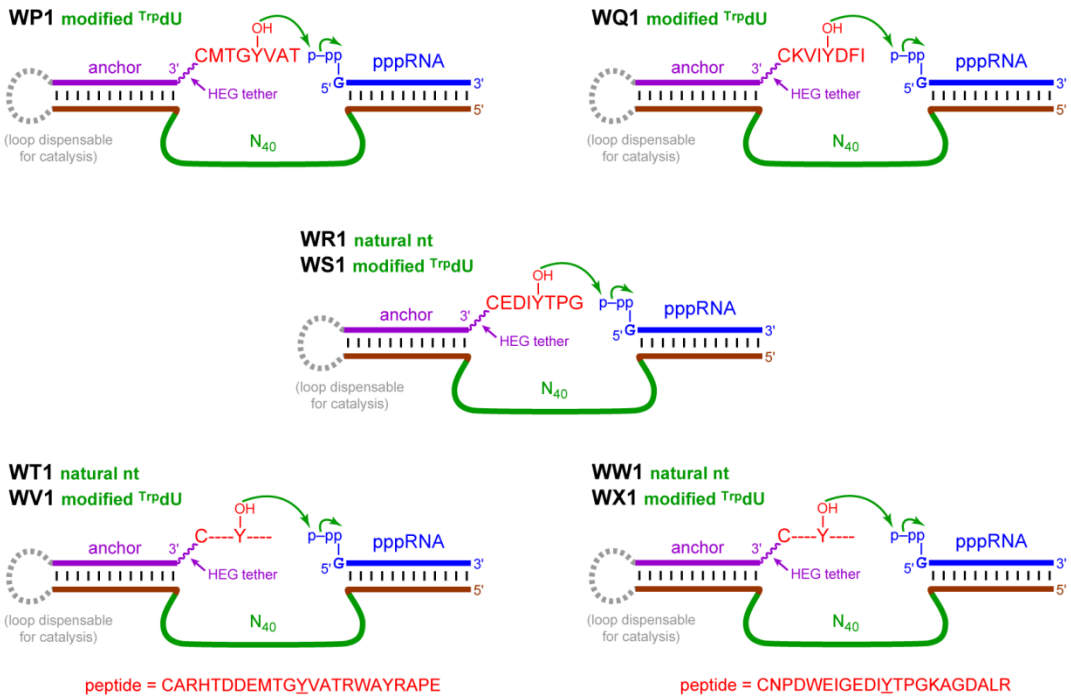


Figure 4.17. Design for the WP1-WX1 in vitro selection experiments to identify tyrosine kinase deoxyribozymes containing hydrophobic modifications. In all selections, the peptide substrate was connected to a DNA anchor oligonucleotide via a disulfide bond and a HEG tether, pppRNA was used as the phosphoryl donor, and the initially random region was N_{40}. Selections used only natural nucleotides or incorporated ^{Trp}dU as indicated. The peptide substrate for each selection is shown. The selection step was performed in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h.

The SOMAmers SL5 with 8 modified nucleotides and SL1025 with 10 modified nucleotides have hydrophobic interactions with 24 and 11 amino acids in their respective protein substrates.[18,20] Based on these results, the increased binding affinity of the SOMAmers may be in part due to the large number of interactions between the hydrophobically modified aptamers and the amino acid residues in the protein target. Therefore, four selection experiments, WT1-WX1, used longer peptide substrates to enable more interactions between the deoxyribozyme and peptide substrate. The long peptide substrates with 22 residues were expanded versions of the

short peptide substrates based on the natural protein sequences. The peptides CARHTDDEMTGYVATRWYRAPE and CNPDWEIGEDIYTPGKAGDALR, respectively, correspond to the residues flanking Tyr182 of MAPK14/p38a (172-192) and Tyr443 of CNP (433-453). Each of these peptide substrates was used in selection experiments with either natural or Trp-modified DNA pools. The selection experiments were performed with the pTyr immunoprecipitation capture method, because the long, 22 residue peptide substrates were capturable by the pTyr immunoprecipitation method and not capturable by the 8VP1 deoxyribozyme.

The WP1-WX1 selections were initiated and all five selection experiments with Trp-modified nucleotides resulted in no or very little 30-cycle PCR product to begin round 2, while all three selection experiments with natural nucleotides were successful (Figure 4.18). When the WP1, WR1, and WS1 selection experiments were repeated with the 8VP1 capture method, similar results were observed. The other selection experiments could not be performed with the 8VP1 capture method because the long peptide substrates could not be captured or an 8VP1 capture deoxyribozyme with binding arms complementary to the pool was not available. The problem with PCR amplification was unexpected based on the success of the previous control experiments. However, in all of the control experiments the amount of DNA sequences expected to survive the first round was estimated to be 0.01 pmol, but this is still 10^{10} DNA sequences. The actual number of DNA sequences that catalyze the desired reaction and survive the selection pressure may be orders of magnitude lower than 10^{10}. Therefore, the control experiments may not have accurately represented the population during the actual selection experiments.

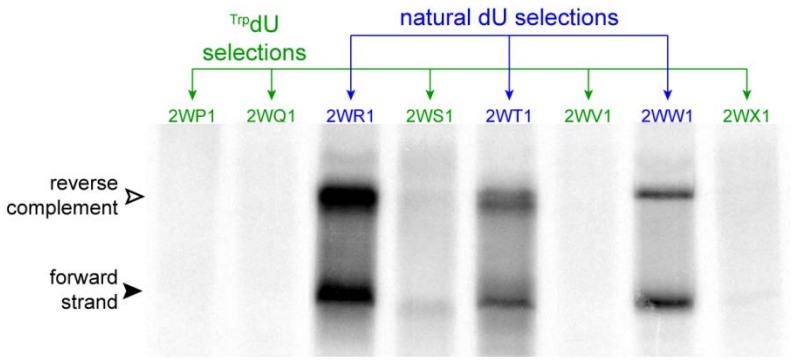
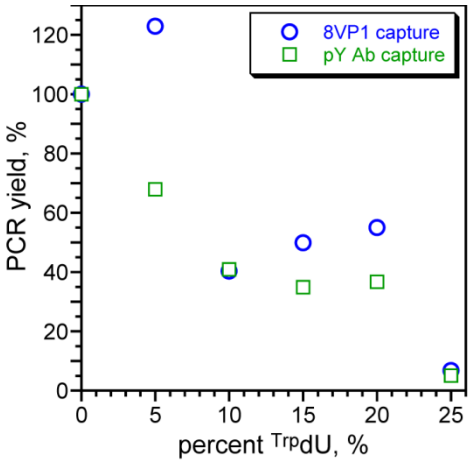


Figure 4.18. PCR gel for round 2 of the WP1-WX1 selection experiments. The two DNA strands were separable because the reverse primer contained a nonamplifiable spacer. Round 1 was performed using the pTyr immunoprecipitation capture method. The PAGE image shows the results of 30-cycle PCR to begin round 2 for the WP1-WX1 selections. All three selection experiments with natural nucleotides led to PCR product. However, no or very little PCR product was observed for the five selection experiments that incorporated modified ^{Trp}dU. To continue the selection round, the reverse-complement strand is needed as the template for primer extension.

The only reported Trp-modified aptamer sequence was identified from an N_{40} random pool and contained 9 ^{Trp}dU residues.[16] Surprisingly, none of the modified residues in the aptamer are adjacent. Therefore, it was hypothesized that in the WP1-WX1 selection experiments the initially random region with 25% ^{Trp}dU contained too many modifications to survive the amplification step or to be able to catalyze the phosphorylation of the peptide substrate. To test the effect the percentage of ^{Trp}dU has on the ability to execute selection experiments, the WP1 selection was performed with 0%, 5%, 10%, 15%, 20%, and 25% ^{Trp}dU in the initially random region (Figure 4.19). The remaining three natural nucleotides equally composed the remainder of the initially random region for each pool. The pools with different percentages of ^{Trp}dU residues were used in a round 1 selection experiment with the 8VP1 capture method and, separately, the pTyr immunoprecipitation capture method. The 30-cycle PCR product for round 2 was quantified based on the intensity of the desired product band (the reverse-complement) in each experiment relative to the intensity of the product band in the selection experiment with 0% ^{Trp}dU (all natural nucleotides). As previously observed, 25% ^{Trp}dU led to very little PCR product. However, reducing the average amount of ^{Trp}dU in the initially random region led to an increase

in the amount of round 2 PCR product. Pools containing 10-20% ^{Trp}dU resulted in 40% PCR yield compared to the unmodified control, which was acceptable for selection experiments with modified nucleotides. Based on these results, all of the Trp-modified selection experiments were restarted using an initially random pool with 15% ^{Trp}dU modification, providing a compromise between reliable PCR product yields and a large number of modified nucleotides in the initially random region.



percent ^{Trp}dU	average ^{Trp}dU residues
0	0
5	2
10	4
15	6
20	8
25	10

Figure 4.19. Evaluation of pools containing varying percentage of ^{Trp}dU in the N_{40} initially random region for use during in vitro selection experiments. The plot shows the round 2 WP1 30-cycle PCR yields for selection experiments with initially random pools containing the indicated average percentage of ^{Trp}dU residues. PCR yields were normalized to the yield of the PCR product with 0% ^{Trp}dU. Each pool was used in a selection experiment with the 8VP1 capture method and the pTyr immunoprecipitation capture method. The table shows the average number of ^{Trp}dU residues in the initially random region of an N_{40} pool at the various percentages used.

No activity was observed in any of the WP1-WX1 selection experiments, and they were discontinued after 12 rounds. It is not clear why the selection experiments with natural nucleotides were not successful. The challenges incorporating the ^{Trp}dU modified nucleotides into the in vitro selection experiments may contribute to the lack of success in the selections with Trp-modified pools. Despite the encouraging literature results showing that the incorporation of ^{Trp}dU into DNA aptamers led to the identification of aptamers with low nanomolar K_d values for all 14 of the proteins tested,[17] very few Trp-modified aptamers have been reported. Most of the

SOMAmers reported have been identified from aptamer selection experiments using benzyl-modified residues. Based on the lack of reported Trp-modified aptamers and the experimental difficulties of incorporating ^{Trp}dU into the in vitro selection experiments to identify kinase deoxyribozymes, the large, hydrophobic indole modification of ^{Trp}dU may be challenging to use in selection experiments. The use of the smaller benzyl hydrophobic modification may still provide improved binding interactions between the peptide substrate and the deoxyribozyme. Additionally, ^{Bn}dU has been used successfully during multiple reported aptamer selection experiments, and therefore may be incorporated more easily into kinase deoxyribozyme selection experiments.

4.2.3 DNA-Catalyzed Phosphorylation of Untethered Peptides

In separate efforts, Chih-Chi Chu evaluated the use of hexahistidine (His_6) tags on peptide substrates to recruit deoxyribozymes for peptide modification.[33] A similar recruiting strategy has been used for DNA-templated protein modification.[34] This general strategy used a DNA anchor oligonucleotide with a 3′-tris(NTA) moiety (NTA = nitrilotriacetic acid). The NTA moiety binds to divalent metal ions and interacts with the His_6 tag on the peptide substrate. The DNA anchor base-pairs to the binding arm of the deoxyribozyme. Therefore, the deoxyribozyme is recruited to the peptide substrate in solution and catalyzes the modification of the peptide. This recruiting strategy improved the k_{obs} and K_m of the nucleopeptide-forming 8XJ105 deoxyribozyme that was identified from selection experiments in which untethered peptide substrates were used.

The His_6 recruiting strategy was then evaluated with tyrosine kinase deoxyribozymes that were previously unable to phosphorylate untethered peptide substrates (Figure 4.20).[33] The 6CF134 tyrosine kinase deoxyribozyme phosphorylates a peptide substrate only when the peptide is covalently tethered to a DNA anchor oligonucleotide (Chapter 2).[5] 6CF134 successfully phosphorylated a His_6-tagged peptide substrate when the 3′-tris(NTA) oligonucleotide and 100 μM Cu^{2+} was present, as determined by matrix-assisted laser desorption

ionization (MALDI) mass spectrometry. This is the first example of DNA-catalyzed phosphorylation of a non-covalently tethered peptide substrate. The interaction between the His_6 tag and the tris(NTA) was sufficient to replace the previously used covalent tether and enabled 6CF134-catalyzed phosphorylation of an untethered peptide substrate.

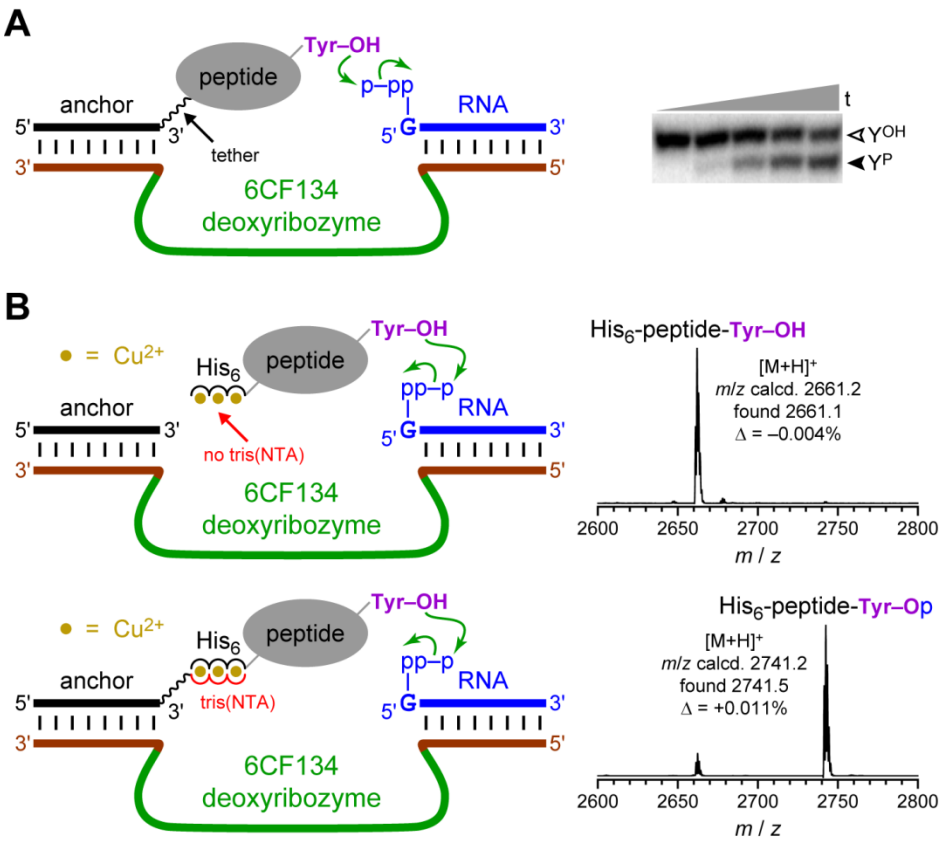


Figure 4.20. Evaluation of the 6CF134 tyrosine kinase deoxyribozyme with recruiting of His_6-tagged peptide substrate. (A) Original arrangement of the 6CF134 deoxyribozyme with a disulfide-tethered peptide substrate required for 6CF134-catalyzed phosphorylation. The PAGE image shows the phosphorylation of the DNA-anchored peptide substrate with ~50% phosphorylation yield; t = 30 s, 30 min, 2 h, 5 h, and 16 h. Y^{OH} = substrate; Y^P = product. Incubation conditions were 70 mM HEPES, pH 7.5, 0.5 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at room temperature. (B) Recruiting of a His_6-tagged peptide substrate by an NTA-modified DNA anchor oligonucleotide resulted in 88% phosphorylation yield by 6CF134 as determined by MALDI mass spectrometry. Peptide phosphorylation was only observed when the DNA anchor was 3′-tris(NTA) modified. The peptide substrate was H_6SAGERASAEDMARAAYAA. Incubation conditions were 70 mM HEPES, pH 7.5, 2 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, 150 mM NaCl, and 100 µM $Cu(NO_3)_2$ at room temperature for 20 h. Figure reprinted with permission from Ref. 33.

4.3 Summary and Future Directions

DNA enzymes that catalyze the modification untethered peptide and protein substrates are desired. The development of a method to reliably identify deoxyribozymes with untethered peptide and protein reactivity is valuable in the pursuit of catalysts to enable studies of biological protein modifications. The use of DNA-anchored HEG-tethered peptide substrates during in vitro selection has not reliably led to the identification of deoxyribozymes with untethered peptide reactivity. This includes all current tyrosine kinase deoxyribozymes, which require DNA-anchored HEG-tethered peptide substrates. The identification of kinase deoxyribozymes requires that the phosphopeptide product is connected to the DNA sequence that catalyzed its phosphorylation. Therefore, a tethered peptide substrate is required during in vitro selection.

The original hypothesis was that a sufficiently long tether between the peptide and DNA anchor would mimic an untethered peptide substrate. Unfortunately, selection experiments to identify capture deoxyribozymes or tyrosine kinase deoxyribozymes using HEG_6 or HEG_{10}-tethered peptide substrates were unsuccessful. The 8XJ105 nucleopeptide-forming deoxyribozyme with untethered peptide reactivity was evaluated with tethered peptide substrates, and the results suggest that the longer tethers do not mimic an untethered peptide substrate. Therefore, the longer tethers were not inert as expected and prevented DNA enzymes from binding or reacting with the peptide substrate. The combined results of many unsuccessful selection experiments and the inability of 8XJ105 and 15MZ36 to catalyze the nucleopeptide formation of HEG_6 and HEG_{10}-tethered peptide substrates suggest that the hypothesis that long tethers are similar to the absence of a tether was incorrect. Thus, HEG_6 and HEG_{10}-tethered peptide substrates will no longer be used during in vitro selection experiments.

The identification of deoxyribozymes and aptamers with modified nucleotides has expanded the catalytic scope of deoxyribozymes and the binding affinities of aptamers. The incorporation of hydrophobic modifications into deoxyribozymes may increase the binding affinity between the peptide substrate and the DNA enzyme, thus leading to untethered peptide

reactivity. The modified ^{Trp}dUTP was successfully synthesized and incorporated into a DNA pool. Unexpected challenges in PCR amplification of active DNA sequences during in vitro selection experiments were encountered and attributed to the ^{Trp}dU modification. Decrease in the percentage of ^{Trp}dU residues in the initially random region of the pool enabled selection experiments. The data shows that the use of ^{Trp}dU resulted in the identification of tight-binding aptamers for all proteins evaluated, whereas the use of ^{Bn}dU lead to tight-binding aptamers for only 60% of proteins evaluated. Despite these results, the majority of SOMAmers were identified using ^{Bn}dU for an unknown reason. Future efforts to identify deoxyribozymes with modified nucleotides will try the hydrophobic benzyl modification. The magnitude of the binding interaction required between the deoxyribozyme and its peptide substrate to enable untethered peptide reactivity is not known, and ^{Bn}dU may be sufficient to enable untethered peptide reactivity. The ^{Bn}dUTP can be synthesized similarly to the ^{Trp}dUTP by replacing tryptamine with benzylamine.

The aptamers with hydrophobic modifications bind tightly to their protein substrates. Therefore, the in vitro selection strategy may want to avoid a bead-based immunoprecipitation method in which tight-binding aptamers for the resin or immobilized proteins may prevent the successful enrichment of catalytic DNA sequences. The previously used, PAGE-based 8VP1 capture method was not amenable to the longer peptide substrates. A new PAGE-based capture method using Phos-tag acrylamide is promising to enable a sufficient PAGE separation of active DNA sequences connected to their phosphotyrosine-containing peptide products from the inactive DNA sequences.[35-39] The Phos-tag PAGE capture method should prevent the enrichment of DNA aptamers for the proteins or resin used in the immunoprecipitation capture method and still allow for the use of long peptide substrates.

DNA enzymes containing modified nucleotides with protein-like functional groups have expanded the scope of DNA-catalyzed reactions to include amide bond cleavage.[15] The active sites of known protein kinases contain aspartate and lysine residues.[40] Therefore, incorporation of amino or carboxyl modifications into kinase deoxyribozymes may be beneficial for catalysis

and provide improved rate, yield, or substrate binding (either peptide or NTP substrate). Ongoing selection experiments VJ1-VN1, initiated by Stephanie Konecki and continued by Tiyaporn Tangpradabkul, are seeking tyrosine kinase deoxyribozymes with ^{Am}dU or ^{COOH}dU modifications in the initially random region. The DNA-anchored HEG-tethered peptide substrate is CMTGYVAT, and the phosphoryl donor is either pppRNA or 100 μM ATP. Selection experiments use the 8VP1 capture deoxyribozyme method.

While the DNA-catalyzed phosphorylation of untethered peptide substrates and a reliable method to identify these deoxyribozymes are still desired, the ability to recruit peptide substrates to the DNA enzyme does enable phosphorylation of peptide substrates. The interaction between a His_6 tag and a tris(NTA)-oligonucleotide is sufficient to replace the previously used covalent tether and enables 6CF134-catalyzed phosphorylation of an untethered peptide substrate.

4.4 Materials and Methods

4.4.1 Substrate Preparation Procedures

Oligonucleotides, peptides, and DNA-anchored peptide conjugates were prepared as described in Chapter 2.

4.4.2 In Vitro Selection Procedures

Procedures for ligation, selection, 8VP1 PAGE-based capture, and PCR were performed as described in Chapter 2. Phosphotyrosine antibody bead-based capture and preclear was performed as described in Chapter 3. Any changes to incorporate modified nucleotides are described below.

Primer extension to generate modified pools for round 1. To generate the initial 15% Trp-modified pools by primer extension, a reverse-complement of the pool containing 15% A was synthesized and ligated to a 3′-tail to enable separation from the product after primer extension. A 25 μL sample containing 250 pmol of reverse-complement template, 300 pmol of forward

primer, 7.5 nmol each of dATP, dGTP, dCTP, and ^{Trp}dUTP, 2.5 μL of 10× KOD XL polymerase buffer (1.2 M Tris-HCl, pH 8.0, 60 mM $(NH_4)_2SO_4$, 100 mM KCl, 1% Triton X-100, 0.01% BSA), and 2 μL of KOD XL polymerase. To generate the initial pool, 8 × 25 μL reactions were performed. Primer extension was performed in a PCR thermocycler according to the following program: 94 °C for 2 min, 47 °C for 2 min, 72 °C for 1 h. The 8 × 25 μL samples were combined, precipitated with ethanol, and separated by 8% PAGE.

Primer extension to generate modified pools for subsequent rounds. A 25 μL sample containing the reverse-complement single strand, 50 pmol of forward primer, 7.5 nmol each of dATP, dGTP, dCTP, and ^{Trp}dUTP, 20 μCi of α-^{32}P-dCTP (800 Ci/mmol), 2.5 μL of 10× KOD XL polymerase buffer (1.2 M Tris-HCl, pH 8.0, 60 mM $(NH_4)_2SO_4$, 100 mM KCl, 1% Triton X-100, 0.01% BSA), and 0.5 μL of KOD XL polymerase. Primer extension was performed in a PCR thermocycler according to the following program: 94 °C for 2 min, 47 °C for 2 min, 72 °C for 1 h. The sample was separated by 8% PAGE, providing the modified DNA pool for the next selection round.

Procedure for ligation of modified pools in round 1. A 30 μL sample containing ~200 pmol of modified DNA pool, 250 pmol of DNA splint, and 300 pmol of 5′-phosphorylated DNA-anchored peptide substrate was annealed in 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. To this solution were added 4 μL of 10× T4 DNA ligase buffer (400 mM Tris, pH 7.8, 100 mM $MgCl_2$, and 5 mM ATP) and 1 μL of 5 U/μL T4 DNA ligase (Fermentas). The sample was incubated at 37 °C for 14 h and separated by 8% PAGE.

Procedure for PCR of modified pools. In each selection round, two PCR reactions were performed, a 10-cycle PCR followed by a 30-cycle PCR. First, a 100 μL sample was prepared containing the capture product, 50 pmol of forward primer, 200 pmol of reverse primer, 20 nmol of each dNTP, and 10 μL of 10× KOD XL polymerase buffer (1.2 M Tris-HCl, pH 8.0, 60 mM $(NH_4)_2SO_4$, 100 mM KCl, 1% Triton X-100, 0.01% BSA), and 1 μL KOD XL polymerase. This sample was primer-extended and then cycled 10 times according to the following PCR program:

94 °C for 2 min, 47 °C for 2 min, 72 °C for 30 min; then 94 °C for 2 min, 10× (94 °C for 1 min, 47 °C for 1 min, 72 °C for 1 min), 72 °C for 5 min. KOD XL polymerase was removed by phenol/chloroform extraction. Second, a 50 µL sample was prepared containing 1 µL of the 100 µL 10-cycle PCR product, 25 pmol of forward primer, 100 pmol of reverse primer, 10 nmol of each dNTP, 5 µCi of α-^{32}P-dCTP (800 Ci/mmol), 5 µL of 10× Taq polymerase buffer, and Taq polymerase. This sample was cycled 30 times according to the following PCR program: 94 °C for 2 min, 30× (94 °C for 30 s, 47 °C for 30 s, 72 °C for 30 s), 72 °C for 5 min. Samples were separated by 8% PAGE, and the reverse-complement single strand was separated for the subsequent primer extension.

4.4.3 Synthesis of ^{Trp}dUTP

Reagents were commercial grade and used without purification. Thin-layer chromatography (TLC) was performed on silica gel plates pre-coated with fluorescent indicator, with visualization by UV light (254 nm). Flash column chromatography was performed with silica gel (230-400 mesh). NMR spectra were recorded on a Varian Unity instrument. Mass spectrometry data were obtained at the UIUC School of Chemical Sciences Mass Spectrometry Laboratory using a Waters Quattro II instrument (LR-ESI).

Synthesis of ^{Trp}dU. The ^{Trp}dU was synthesized from (*E*)-5-(2-carbomethoxyvinyl)-2′-deoxyuridine *N*-hydroxysuccinimide ester (Berry & Associates, cat. no. PY7190). The 2′-deoxyuridine (118 mg, 0.30 mmol) was dissolved in 2 mL of dimethylformamide (DMF). To this solution was added tryptamine (48 mg, 0.30 mmol) and *N*,*N*-diisopropylethylamine (105 µL, 0.60 mmol), and the solution was stirred overnight at room temperature. The solution was dried on a rotary evaporator and the oily residue was purified by silica gel column chromatography, eluting with 10% CH_3OH in CH_2Cl_2 to yield the desired product as a colorless oil (39 mg, 30%).

TLC: R_f = 0.55 [10% CH_3OH in CH_2Cl_2].

ESI-MS: *m*/*z* calculated for $C_{22}H_{24}N_4O_6$ $[M–H]^-$ 439.2; found 439.1.

Synthesis of ^{Trp}dU 5′-triphosphate (^{Trp}dUTP). The triphosphorylation procedure was modified from the procedure reported by Huang and co-workers.[41] 2-Chloro-1,3,2-benzodioxaphosphorin-4-one (TCI, cat. no. C1210) was portioned into a dried vial in a glove box. ^{Trp}dU and tributylammonium pyrophosphate (Sigma, cat. no. P83533) were dried under vacuum. Tributylamine was dried over 4 Å molecular sieves. Tributylamine (0.35 mL, 1.47 mmol) was added to a solution of tributylammonium pyrophosphate (99 mg, 0.18 mmol) dissolved in 1 mL of anhydrous DMF under argon. This solution was added into a stirred solution of 2-chloro-1,3,2-benxodioaphosphorin-4-one (26 mg, 0.13 mmol) in 0.55 mL of anhydrous DMF under argon. The resulting solution was stirred at room temperature for 1 h and then was added dropwise over 5 min to a solution of ^{Trp}dU (38 mg, 0.09 mmol) in 1 mL of anhydrous DMF under argon, cooled in an ice-salt bath at 0 to –10 °C. The solution was stirred for 5 h and raised to room temperature and incubated for 30 min. A solution of 0.02 M iodine in THF/pyridine/water (Glen Research, 8 mL) was added until a permanent brown color was maintained. The brown solution was stirred at room temperature for 30 min. Two sample volumes of water were added. The two-layer mixture was stirred vigorously at room temperature for 1.5 h, followed by addition of 0.1 volumes of 3 M NaCl and 3 volumes of ethanol. The thoroughly mixed sample was precipitated at –80 °C overnight and centrifuged for 1 h at 10,000 × *g* and 4 °C. The precipitated sample was dried, dissolved in water, and purified by HPLC [Shimadzu Prominence instrument; Phenomenex Gemini-NX C_{18} column, 5 µm, 10 × 250 mm; linear gradient of 0.5% solvent A (20 mM triethylammonium acetate in 50% acetonitrile/50% water, pH 7.0) and 99.5% solvent B (20 mM triethylammonium acetate in water, pH 7.0) at 0 min to 25% solvent A and 75% solvent B at 45 min with a flow rate of 3.5 mL/min]. The ^{Trp}dUTP was purified as its triethylammonium salt [41 mg, 20% from (*E*)-5-(2-carbomethoxyvinyl)-2′-deoxyuridine *N*-hydroxysuccinimide ester].

ESI-MS: *m*/*z* calculated for $C_{22}H_{23}N_4O_{15}P_3$ $[M–H]^-$ 679.1; found 679.2.

4.4.4 Procedures for Expression of KOD Polymerase

Construction of the KOD and KOD exo⁻ plasmids. The introns from the full-length KOD polymerase gene (GenBank D29671.2) were removed, and the gene was codon-optimized for expression in *E. coli* using the Integrated DNA Technologies (IDT) Codon Optimization Tool.

Table 4.1. Oligonucleotide sequences of the KOD polymerase gBlocks and related primers.

oligonucleotide purpose	oligonucleotide sequence
forward primer	GTCCACATATGATTTTAGACACCG
reverse primer	TGTCAGGATCCTCGAGTG
forward mutagenic primer	AGAAAGATCCAGACGTCCTGATTACCTAT***GAT***GGCGACAACTTTGATTTTGCGTACCTG
reverse mutagenic primer	CAGGTACGCAAAATCAAAGTTGTCGCC***ATC***ATAGGTAATCAGGACGTCTGGATCTTTCT
sequencing primer 1	GCGCACGCGTGATTACAT
sequencing primer 2	TGGCTCCCAACAAGCCAG
sequencing primer 3	TCAATGCGAAACTCCCAG

gBlock1[a]

GTCCACATATGATTTTAGACACCGACTATATCACGGAAGATGGCAAGCCAGTGATCCGTATCTTCAAAAAGGAGAATGGCGAGTTCAAAATCGAGTA
TGACCGTACGTTTGAACCGTACTTTTACGCTCTGCTGAAAGATGACTCCGCCATCGAGGAGGTAAAGAAGATCACCGCGGAACGGCATGGCACGGTG
GTTACGGTTAAACGCGTGGAAAAGGTTCAAAAGAAATTCCTGGGCCGCCCGGTAGAAGTCTGGAAACTTTATTTCACCCACCCTCAGGACGTGCCGG
CGATCCGGGACAAGATTCGCGAACATCCCGCTGTAATTGACATCTATGAATATGATATCCCCTTCGCGAAACGTTATCTGATCGATAAAGGTTTAGT
CCCCATGGAGGGCGATGAGGAACTCAAAATGCTGGCGTTCGACATTGAGACCTTATATCATGAAGGCGAGGAATTTGCCGAAGGTCCGATTTTAATG
ATTAGTTACGCGGATGAAGAAGGCGCACGCGTGATTACATGGAAGAACGTTGACTTGCCTTATGTGGACGTTGTGTCCACAGAGCGCGAAATGATTA
AACGTTTCCTTCGTGTGGTGAAGGAGAAAGATCCAGACGTCCTGATTACCTAT***AAT***GGCGACAACTTTGATTTTGCGTACCTGAAAAAACGTTGCGA
GAAACTGGGAATCAATTTTGCGCTGGGTCGCGACGGTTCTGAGCCAAAAATTCAGCGTATGGGAGATCGCTTCGCCGTTGAAGTTAAAGGCCGTATT
CATTTCGATCTGTATCCAGTCATCCGGCGCACTATCAACCTTCCGACTTATACTCTTGAAGCCGTATATGAAGCTGTGTTTGGACAGCCTAAAGAAA
AAGTGTATGCGGAAGAAATTACTACCGCTTGGGAAACAGGCGAGAACTTAGAACGCGTGGCGCGTTATAGCATGGAAGATGCAAAAGTTACCTATGA
ATTAGGTAAAGAGTTTTTGCCTATGGAAGCACAGCTGTCGCGTCTGATCGGGCAGAGTCTTTGGGATGTGAGCCGTAGCTCAACAGGCAACCTGGTG
GAATGGTTCCTGCTCCGTAAAGCTTATGAACGCAACGAGCTGGCTCCCAACAAGCCAGATGAAAAAGAACTGGCGCGGCGCCGCCAGTCGTATGAAG
GGGGTTACGTAAAAGAGCCGGAACGCGGTCTGTG

gBlock2[b]

TTACGTAAAAGAGCCGGAACGCGGTCTGTGGGAGAACATCGTCTATTTGGACTTTCGTAGCCTGTATCCTTCTATCATTATTACTCACAACGTCAGC
CCTGACACGCTGAATCGCGAAGGGTGCAAGGAGTACGACGTTGCGCCGCAAGTCGGGCACCGCTTTTGTAAAGATTTCCCGGGATTTATTCCATCTC
TGCTGGGCGATCTGTTGGAAGAACGTCAAAAGATTAAGAAAAAAATGAAAGCAACCATTGATCCGATCGAACGTAAGCTCCTGGACTATCGCCAGCG
CGCAATCAAAATCCTGGCAAACTCGTATTACGGGTATTACGGCTACGCGCGTGCCCGTTGGTATTGCAAGGAGTGTGCCGAGTCCGTAACCGCATGG
GGCCGCGAATATATTACTATGACCATTAAAGAGATTGAAGAGAAGTATGGTTTCAAGGTCATCTACTCGGATACAGATGGGTTCTTTGCCACTATCC
CTGGCGCGGACGCGGAGACTGTGAAGAAAAAGGCAATGGAGTTTTTAAAATACATCAATGCGAAACTCCCAGGTGCACTGGAACTGGAATATGAAGG
GTTTTATAAACGCGGCTTTTTCGTGACGAAAAAGAAGTATGCGGTAATCGATGAAGAGGGGAAGATTACTACCCGCGGCCTGGAAATCGTTCGGCGT
GATTGGTCGGAAATTGCAAAAGAAACTCAAGCGCGTGTGCTGGAAGCCCTCCTGAAGGACGGGGACGTGGAGAAAGCAGTCCGCATCGTTAAGGAGG
TGACAGAGAAATTGTCGAAGTACGAGGTGCCGCCTGAAAAGCTGGTAATTCATGAACAGATCACGCGCGATCTGAAGGATTACAAAGCGACCGGACC
GCACGTGGCGGTGGCAAAACGGTTGGCCGCGCGTGGTGTTAAAATCCGTCCGGGGACCGTGATCTCCTATATCGTCCTGAAAGGAAGTGGCCGGATT
GGCGATCGGGCTATTCCGTTTGACGAATTCGACCCGACCAAACATAAATATGATGCAGAATACTATATCGAAAATCAGGTTCTGCCGGCGGTAGAAC
GCATTTTGCGGGCTTTTGGCTATCGCAAAGAAGATCTGCGCTATCAGAAGACGCGTCAGGTGGGTCTGAGCGCGTGGCTGAAGCCGAAAGGCACC**TA**
AAAAAGCGATAGATATCACTCGAGGATCCTGACA

[a]The 8 underlined bp at the 5′-end are removed upon cleavage by NdeI, and the red nucleotides are the NdeI recognition site. The bold nucleotides represent the codon for N210 and the mutagenic primers change the first A to G. The 30 underlined bp at the 3′-end base-pair with the 5′-end of gBlock 2.

[b]The 30 underlined bp at the 5′-end base-pair with the 3′-end of gBlock 1. The underlined bp at the 3′-end are removed upon cleavage by Bam HI, and the red nucleotides are the BamHI recognition site. The bold nucleotides are the stop codon.

PCR amplification of the full-length KOD and KOD exo⁻ polymerase genes. A 50 µL sample was prepared containing 10 ng of each gBlock gene fragment, 25 pmol of each primer, 10 nmol of each dNTP, 5 µL of 10× Pfu polymerase buffer (200 mM Tris-HCl, pH 8.8, 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 20 mM $MgSO_4$, 1% Triton X-100), and 0.5 uL of Pfu polymerase. This sample was cycled 30 times according to the following PCR program: 94 °C for 2 min, 30× (94 °C for 30 s, 47 °C for 30 s, 72 °C for 5 min), 72 °C for 5 min. The sample were purified on 1% agarose gel. Site-directed mutagenesis to generate the KOD exo⁻ gene was performed using mutagenic primers and 10 ng of the full-length KOD polymerase gene. PCR was performed as above with each of the mutagenic primers, and the full-length KOD exo⁻ gene was prepared by PCR and purified on a 1% agarose gel.

Restriction digestion of the KOD genes and vector. Each of the full-length KOD polymerase genes and the pET-16b vector containing the Pfu gene were digested with BamHI and NdeI. A 30 µL sample was prepared containing 500 ng of vector or 300 ng of the KOD gene and 3 µL of 10× NEB buffer 3 (500 mM Tris-HCl, pH 7.9, 1 M NaCl, 100 mM $MgCl_2$, 10 mM DTT), 100 µg/mL BSA, 1 µL of 20,000 U/mL of BamHI (New England Biolabs), and 1 µL of 20,000 U/mL of NdeI (New England Biolabs). Samples were incubated at 37 °C for 14 h. The digested vector was dephosphorylate by adding 4 µL of 10× calf intestine alkaline phosphatase (CIAP) buffer (500 mM Tris-HCl, pH 8.5, 1 mM EDTA) and 1 µL of 20 U/µL CIAP (Fermentas) to the digested product for a final volume of 40 µL. The samples were incubated at 37 °C for 2 h. The digested, dephosphorylated vector and digested genes were purified on a 1% agarose gel.

Ligation and transformation of the KOD and KOD exo⁻ polymerase plasmids. A 20 µL sample containing 50 ng of digested and dephosphorylated pET-16b vector, 100 ng of digested gene, 2 µL of 10× T4 DNA ligase buffer (400 mM Tris-HCl, pH 7.8, 100 mM $MgCl_2$, 100 mM DTT, 5 mM ATP), and 1 µL of 1 U/µL of T4 DNA Ligase (Fermentas) was incubated at 16 °C for 16 h. To 50 µL of DH5α *E. coli* cells was added 4 µL of the ligation product, and the sample was incubated on ice for 30 min. The sample was heat-shocked at 42 °C for 45 s, followed by

incubation on ice for 2 min. To the sample was added 250 µL of room-temperature SOC medium. The sample was shaken at 220 rpm and 37 °C for 1 h, and plated on LB containing 100 µg/mL ampicillin and incubated at 37 °C for 16 h. Individual colonies were used to inoculate 3 mL of LB containing 100 µg/mL ampicillin, and the culture was shaken at 37 °C for 16 h. Plasmids were isolated by miniprep and sequenced.

Expression and purification of KOD and KOD exo$^-$ polymerase. To 50 µL of BL21 (DE3) pLysS *E. coli* was added 1 µL of a 1:100 dilution of the miniprep purified plasmid. The sample was mixed gently and incubated on ice for 30 min. The cells were heat-shocked at 42 °C for 1 min, followed by incubation on ice for 2 min. To the sample was added 500 µL of room-temperature SOC medium. The sample was shaken at 220 rpm and 37 °C for 1 h, and then plated on LB plates containing 100 µg/mL ampicillin and incubated at 37 °C for 16 h. Individual colonies were used to inoculate 9 mL starter cultures of LB with 100 µg/mL ampicillin and 34 µg/mL chloramphenicol. Cultures were grown by shaking at 220 rpm and 37 °C for 16 h. The starter culture with the highest OD_{600} was used to inoculate the induction culture. To 750 mL of LB containing 100 µg/mL ampicillin and 34 µg/mL chloramphenicol, 7.5 mL of starter culture was added. The culture was grown by shaking at 220 rpm at 37 °C until OD_{600} = 0.3. The culture was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM and was shaken at 220 rpm at 37 °C for an additional 4 h. The cells were pelleted by centrifugation at 9,000 ×*g* and 4 °C for 10 min.

The cell pellet was resuspended in 22 mL of Buffer A (50 mM Tris-HCl, pH 7.9, 50 mM glucose, 1 mM EDTA, 5 mg/mL lysozyme) and incubated at room temperature for 30 min. The sample was sonicated for 90 s at 35% amplitude with 4 s pulse-on and 9.9 s pulse-off. To the cell lysate, 22 mL of Buffer B (10 mM Tris-HCl, pH 7.9, 50 mM KCl, 1 mM EDTA, 0.5% Tween-20, 0.5% Nonidet P40) was added, and the sample was stirred at 75 °C for 30 min. Cell debris was removed by centrifugation at 12,000 ×*g* and 4 °C for 20 min. An equal volume of 50% glycerol storage buffer (50 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 1 mM DTT, 0.1% Tween-20, 0.1 % Nonidet P40) was added slowly. Next, an equal volume of 75% glycerol storage buffer

was added slowly, resulting in a final solution of 50% glycerol. The polymerase was stored at –20 °C. PCR. The 10× KOD XL polymerase buffer contains 1.2 M Tris-HCl, pH 8.0, 60 mM $(NH_4)_2SO_4$, 100 mM KCl, 1% Triton X-100, and 0.01% BSA.

4.5 References

(1) Chandrasekar, J.; Silverman, S. K. Catalytic DNA with phosphatase activity. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5315-5320.

(2) 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.

(3) Chandrasekar, J.; Wylder, A. C.; Silverman, S. K. Phosphoserine lyase deoxyribozymes: DNA-catalyzed formation of dehydroalanine residues in peptides. *J. Am. Chem. Soc.* **2015**, *137*, 9575-9578.

(4) Wang, P.; Silverman, S. K. DNA-catalyzed introduction of azide at tyrosine for peptide modification. *Angew. Chem. Int. Ed.* **2016**, *55*, 10052-10056.

(5) Walsh, S. M.; Sachdeva, A.; Silverman, S. K. DNA catalysts with tyrosine kinase activity. *J. Am. Chem. Soc.* **2013**, *135*, 14928-14931.

(6) Brandsen, B. M.; Velez, T. E.; Sachdeva, A.; Ibrahim, N. A.; Silverman, S. K. DNA-catalyzed lysine side chain modification. *Angew. Chem. Int. Ed.* **2014**, *53*, 9045-9050.

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

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

(9) Chu, C.; Wong, O.; Silverman, S. K. A generalizable DNA-catalyzed approach to peptide-nucleic acid conjugation. *ChemBioChem* **2014**, *15*, 1905-1910.

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

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

(12) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. A DNAzyme with three protein-like functional groups: Enhancing catalytic efficiency of M^{2+}-independent RNA cleavage. *ChemBioChem* **2009**, *10*, 1988-1992.

(13) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. A self-cleaving DNA enzyme modified with amines, guanidines and imidazoles operates independently of divalent metal cations (M^{2+}). *Nucleic Acids Res.* **2009**, *37*, 1638-1649.

(14) Hollenstein, M. Expanding the catalytic repertoire of DNAzymes by modified nucleosides. *Chimia* **2011**, *65*, 770-775.

(15) Zhou, C.; Avins, J. L.; Klauser, P. C.; Brandsen, B. M.; Lee, Y.; Silverman, S. K. DNA-catalyzed amide hydrolysis. *J. Am. Chem. Soc.* **2016**, *138*, 2106-2109.

(16) Vaught, J. D.; Bock, C.; Carter, J.; Fitzwater, T.; Otis, M.; Schneider, D.; Rolando, J.; Waugh, S.; Wilcox, S. K.; Eaton, B. E. Expanding the chemistry of DNA for in vitro selection. *J. Am. Chem. Soc.* **2010**, *132*, 4141-4151.

(17) Rohloff, J. C.; Gelinas, A. D.; Jarvis, T. C.; Ochsner, U. A.; Schneider, D. J.; Gold, L.; Janjic, N. Nucleic acid ligands with protein-like side chains: Modified aptamers and their use as diagnostic and therapeutic agents. *Mol. Ther. Nucleic Acids* **2014**, *3*, 201.

(18) Davies, D. R.; Gelinas, A. D.; Zhang, C.; Rohloff, J. C.; Carter, J. D.; O'Connell, D.; Waugh, S. M.; Wolk, S. K.; Mayfield, W. S.; Burgin, A. B.; Edwards, T. E.; Stewart, L. J.; Gold, L.; Janjic, N.; Jarvis, T. C. Unique motifs and hydrophobic interactions shape the binding of modified DNA ligands to protein targets. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 19971-19976.

(19) Gupta, S.; Hirota, M.; Waugh, S. M.; Murakami, I.; Suzuki, T.; Muraguchi, M.; Shibamori, M.; Ishikawa, Y.; Jarvis, T. C.; Carter, J. D.; Zhang, C.; Gawande, B.; Vrkljan, M.; Janjic, N.; Schneider, D. J. Chemically modified DNA aptamers bind interleukin-6 with high affinity and inhibit signaling by blocking its interaction with interleukin-6 receptor. *J. Biol. Chem.* **2014**, *289*, 8706-8719.

(20) Gelinas, A. D.; Davies, D. R.; Edwards, T. E.; Rohloff, J. C.; Carter, J. D.; Zhang, C.; Gupta, S.; Ishikawa, Y.; Hirota, M.; Nakaishi, Y.; Jarvis, T. C.; Janjic, N. Crystal structure of interleukin-6 in complex with a modified nucleic acid ligand. *J. Biol. Chem.* **2014**, *289*, 8720-8734.

(21) Sachdeva, A.; Chandra, M.; Chandrasekar, J.; Silverman, S. K. Covalent tagging of phosphorylated peptides by phosphate-specific deoxyribozymes. *ChemBioChem* **2012**, *13*, 654-657.

(22) Dokukin, V.; Silverman, S. K. A modular tyrosine kinase deoxyribozyme with discrete aptamer and catalyst domains. *Chem. Commun.* **2014**, *50*, 9317-9320.

(23) Takagi, M.; Nishioka, M.; Kakihara, H.; Kitabayashi, M.; Inoue, H.; Kawakami, B.; Oka, M.; Imanaka, T. Characterization of DNA polymerase from *Pyrococcus* sp. strain KOD1 and its application to PCR. *Appl. Environ. Microbiol.* **1997**, *63*, 4504-4510.

(24) Nishioka, M.; Mizuguchi, H.; Fujiwara, S.; Komatsubara, S.; Kitabayashi, M.; Uemura, H.; Takagi, M.; Imanaka, T. Long and accurate PCR with a mixture of KOD DNA polymerase and its exonuclease deficient mutant enzyme. *J. Biotechnol.* **2001**, *88*, 141-149.

(25) Grosjean, H.; Fiers, W. Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. *Gene* **1982**, *18*, 199-209.

(26) Sun, Z.; Cai, J. Purification of recombinant Pfu DNA polymerase using a new JK110 resin. *Korean J. Chem. Eng.* **2006**, *23*, 607-609.

(27) Raingeaud, J.; Gupta, S.; Rogers, J. S.; Dickens, M.; Han, J.; Ulevitch, R. J.; Davis, R. J. Pro-inflammatory cytokines and environmental stress cause p38 mitogen-activated protein

kinase activation by dual phosphorylation on tyrosine and threonine. *J. Biol. Chem.* **1995**, *270*, 7420-7426.

(28) Salvador, J. M.; Mittelstadt, P. R.; Guszczynski, T.; Copeland, T. D.; Yamaguchi, H.; Appella, E.; Fornace, A. J., Jr.; Ashwell, J. D. Alternative p38 activation pathway mediated by T cell receptor-proximal tyrosine kinases. *Nat. Immunol.* **2005**, *6*, 390-395.

(29) Walsh, S. M.; Konecki, S. N.; Silverman, S. K. Identification of sequence-selective tyrosine kinase deoxyribozymes. *J. Mol. Evol.* **2015**, *81*, 218-224.

(30) Yokoyama, N.; Lougheed, J.; Miller, W. T. Phosphorylation of WASP by the Cdc42-associated kinase ACK1: dual hydroxyamino acid specificity in a tyrosine kinase. *J. Biol. Chem.* **2005**, *280*, 42219-42226.

(31) Feder, J.; Keay, L.; Garrett, L. R.; Cirulis, N.; Moseley, M. H.; Wildi, B. S. Bacillus cereus neutral protease. *Biochim. Biophys. Acta* **1971**, *251*, 74-78.

(32) Stark, W.; Pauptit, R. A.; Wilson, K. S.; Jansonius, J. N. The structure of neutral protease from Bacillus cereus at 0.2-nm resolution. *Eur. J. Biochem.* **1992**, *207*, 781-791.

(33) Chu, C. C.; Silverman, S. K. Assessing histidine tags for recruiting deoxyribozymes to catalyze peptide and protein modification reactions. *Org. Biomol. Chem.* **2016**, *14*, 4697-4703.

(34) Rosen, C. B.; Kodal, A. L.; Nielsen, J. S.; Schaffert, D. H.; Scavenius, C.; Okholm, A. H.; Voigt, N. V.; Enghild, J. J.; Kjems, J.; Tørring, T.; Gothelf, K. V. Template-directed covalent conjugation of DNA to native antibodies, transferrin and other metal-binding proteins. *Nat. Chem.* **2014**, *6*, 804-809.

(35) Kinoshita, E.; Takahashi, M.; Takeda, H.; Shiro, M.; Koike, T. Recognition of phosphate monoester dianion by an alkoxide-bridged dinuclear zinc(II) complex. *Dalton Trans.* **2004**, 1189-1193.

(36) Kinoshita, E.; Yamada, A.; Takeda, H.; Kinoshita-Kikuta, E.; Koike, T. Novel immobilized zinc(II) affinity chromatography for phosphopeptides and phosphorylated proteins. *J. Sep. Sci.* **2005**, *28*, 155-162.

(37) Kinoshita, E.; Kinoshita-Kikuta, E.; Takiyama, K.; Koike, T. Phosphate-binding tag, a new tool to visualize phosphorylated proteins. *Mol. Cell. Proteomics* **2006**, *5*, 749-757.

(38) Kinoshita, E.; Kinoshita-Kikuta, E.; Koike, T. A single nucleotide polymorphism genotyping method using phosphate-affinity polyacrylamide gel electrophoresis. *Anal. Biochem.* **2007**, *361*, 294-298.

(39) Tilmans, N. P.; Krusemark, C. J.; Harbury, P. A. Expedient synthesis of a modular phosphate affinity reagent. *Bioconjug. Chem.* **2010**, *21*, 1010-1013.

(40) Adams, J. A. Kinetic and catalytic mechanisms of protein kinases. *Chem. Rev.* **2001**, *101*, 2271-2290.

(41) Caton-Williams, J.; Hoxhaj, R.; Fiaz, B.; Huang, Z. Use of a novel 5'-regioselective phosphitylating reagent for one-pot synthesis of nucleoside 5'-triphosphates from unprotected nucleosides. *Curr. Protoc. Nucleic Acid Chem.* **2013**, *Chapter 1*, Unit 1.30.

Chapter 5: Identification of Serine Kinase Deoxyribozymes[†]

5.1 Introduction

Serine and threonine phosphorylation is an abundant regulatory post-translational modification in biology. The hydroxyl group of the amino acid side chain attacks the γ-phosphate of a phosphoryl donor, resulting in transfer of a phosphoryl group. The aromatic hydroxyl group of tyrosine is more nucleophilic than the aliphatic hydroxyl group of serine or threonine side chains, making phosphorylation of serine and threonine side chains more challenging. However, the ability to site-specifically phosphorylate serine or threonine residues within peptide and protein substrates will enable the study of phosphoserine modifications of biological relevance.

5.1.1 Serine-Modifying Deoxyribozymes

While many peptide-modifying deoxyribozymes have focused on tyrosine because it is a good nucleophile, efforts have expanded to include DNA-catalyzed serine modification. The first DNA-catalyzed serine modification was nucleopeptide formation.[1] The serine-containing

[†] University of Illinois graduate student Amit Sachdeva performed kinase deoxyribozyme selection experiments using γ-thiophosphoryl donors for the phosphorylation of 2′-hydroxyl of RNA and the 3′-hydroxyl of DNA.[12]

University of Illinois undergraduate student Alison J. Camden identified and characterized kinase deoxyribozymes that catalyze phosphorylation of the 3′-hydroxyl of DNA.[13]

University of Illinois undergraduate student Nickolaus C. Lammer performed selection experiments for serine kinase deoxyribozymes using complex peptide sequences and investigated the results.

University of Illinois undergraduate student Stephanie N. Konecki performed the initial selections for serine kinase deoxyribozymes using ATP as the phosphoryl donor. S. Konecki also began selections for serine kinase deoxyribozymes using modified nucleotides within the initially random region.

University of Illinois undergraduate student Tiyaporn Tangpradabkul performed selections for serine kinase deoxyribozymes using modified nucleotides within the initially random region.

tripeptide was incorporated between two DNA binding arms, which base-paired to the deoxyribozyme in a highly organized three-helix-junction configuration. Deoxyribozymes were identified to catalyze the reaction of the hydroxyl of the serine side chain with the α-phosphate of a 5′-triphosphorylated RNA oligonucleotide (pppRNA), resulting in nucleopeptide formation. Two of the deoxyribozymes identified, SerB1 and SerB2, have 85% yield in 20 h. Both deoxyribozymes also catalyzes the modification of tyrosine with yields of 10% and 65% for SerB1 and SerB2, respectively.

Subsequent efforts for DNA-catalyzed nucleopeptide formation of serine residues used a DNA-anchored tethered peptide substrate with an open architecture in place of the three-helix-junction architecture.[2] Selection experiments were performed with a tethered aliphatic hydroxyl group, and when activity was observed during the selection the substrate was changed to a tethered CSA tripeptide substrate. The 15MZ36 deoxyribozyme identified from this selection experiment catalyzed nucleopeptide formation of a C_3-tethered CSA substrate with 60% yield. However, only 5% activity was observed with a HEG-tethered CSA substrate, and no activity was observed with a free CSA peptide. 15MZ36 also has 80%, 80%, and 50% yield with C_3-tethered, HEG-tethered, and untethered CYA peptide substrates, respectively. Therefore, 15MZ36 favors tyrosine-containing peptide substrates.

5.1.2 Nucleic Acid-Catalyzed Phosphorylation

Artificial ribozymes have been identified to catalyze self-phosphorylation at their 5′-terminus or at an internal 2′-hydroxyl group.[3-7] These catalysts used ATPγS as the phosphoryl donor. Separate efforts identified DNA enzymes to catalyze self-phosphorylation at their 5′-terminus as well.[8-11] In the Silverman Lab, early efforts by Amit Sachdeva to identify kinase deoxyribozymes using γ-thiophosphoryl donors resulted in activity for selection experiments seeking phosphorylation of 2′-hydroxyl of RNA and the 3′-hydroxyl of DNA.[12] Additional efforts by Alison Camden identified a deoxyribozyme for phosphorylation of the 3′-hydroxyl of DNA.[13] Oligonucleotide phosphorylation is distinct from peptide phosphorylation. The 2′-

hydroxyl of RNA, the 3′-hydroxyl of DNA, and the hydroxyl of threonine are secondary aliphatic hydroxyl groups. The 5′-hydroxyl of DNA and the hydroxyl group of serine are primary hydroxyl group. The ability of DNA to catalyze the phosphorylation of RNA and DNA suggests that serine and threonine phosphorylation may also be possible.

All of the previously identified tyrosine kinase deoxyribozymes were identified from in vitro selection experiments using tyrosine-containing peptide substrates (Chapter 2, 3).[14-16] When each of these deoxyribozymes were evaluated with analogous serine-containing peptide substrates, no phosphorylation activity was observed. Therefore, all of these deoxyribozymes are specifically tyrosine kinase deoxyribozymes. All five of the deoxyribozymes identified from the selection experiments using the CADPYDQS peptide substrates were peptide sequence-specific, and all of the deoxyribozymes phosphorylated the tyrosine residue despite the presence of a serine residue within the same peptide substrate (Chapter 3).[16] Therefore, separate selection experiments with serine-containing peptides may enable the identification of serine kinase deoxyribozymes.

5.2 Results and Discussion

5.2.1 Capture of Phosphoserine-Containing Peptides

Similarly to selection for tyrosine kinase deoxyribozymes, in vitro selection for serine kinase deoxyribozymes also requires a method to separate the catalytically active DNA sequences from the inactive DNA sequences. The nucleopeptide-forming deoxyribozyme 8VP1 was identified from selection experiments performed with a HEG-tethered phosphotyrosine-containing peptide substrate.[17] The 8VP1 capture deoxyribozyme has been used successfully in selection experiments to identify tyrosine kinase deoxyribozymes (Chapter 2).[14] Upon evaluation, 8VP1 was shown to also catalyze reaction of tethered phosphoserine-containing peptide substrates. The tether dependence of the 8VP1-catalyzed phosphoserine reactivity was evaluated (Figure 5.1). 8VP1-catalyzed capture of the CAAS^{P}AA peptide substrate connected to

a DNA anchor via a C_3, HEG, HEG_3, or HEG_6 tether was 29%, 40%, 41%, and 24%, respectively. No activity was observed with the CAASAA peptide substrate. Therefore, the 8VP1 capture method could also be applied to selection experiments for the identification of serine kinase deoxyribozymes.

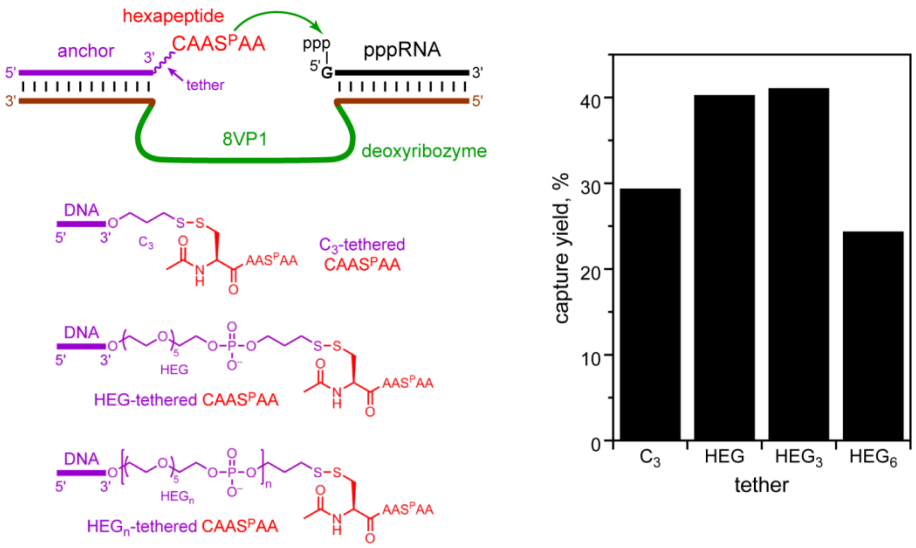


Figure 5.1. DNA-catalyzed covalent modification of phosphoserine substrates. (A) The 8VP1 deoxyribozyme catalyzes attachment of a 5′-triphosphorylated RNA oligonucleotide to the phosphorylated amino acid residue within a DNA-anchored hexapeptide substrate. The tethers between the DNA anchor and the $CAAS^PAA$ peptide substrate are shown. The plot shows 8VP1-catalyzed capture yields of phosphoserine-containing peptide substrates connected via various tether lengths. No activity was observed with serine-containing peptide substrates. Incubation conditions were 50 mM HEPES, pH 7.5, 20 mM $MnCl_2$, and 150 mM NaCl at 37 °C for 16 h.

5.2.2 Identification of Serine Kinase Deoxyribozymes

In vitro selection experiments AC1-AD1, BC1-BD1, and ED1 were performed to identify deoxyribozymes to catalyze serine phosphorylation. The peptide substrate was HEG-tethered CAASAA, and Mn^{2+} and Mg^{2+} were used as divalent metal ion cofactors during the selection step. No activity was observed in any of these selection experiments, and they were discontinued after 12 rounds. However, similar selection experiments to identify tyrosine kinase deoxyribozymes using Zn^{2+}, Mn^{2+}, and Mg^{2+} during the selection step were successful (Chapter 2, 3).[14-16] The tyrosine kinase deoxyribozymes required Zn^{2+} for catalysis. Based on these results,

all three divalent metal ions were used in new selection experiments to identify serine kinase deoxyribozymes (Figure 5.2). In addition, the tether connecting the DNA anchor to the peptide substrate was decreased in length to a C_3 tether and a 5′-triphosphorylated RNA oligonucleotide (pppRNA) was used as the phosphoryl donor. Three selection experiments were performed using N_{30}, N_{40}, and N_{50} random region lengths.

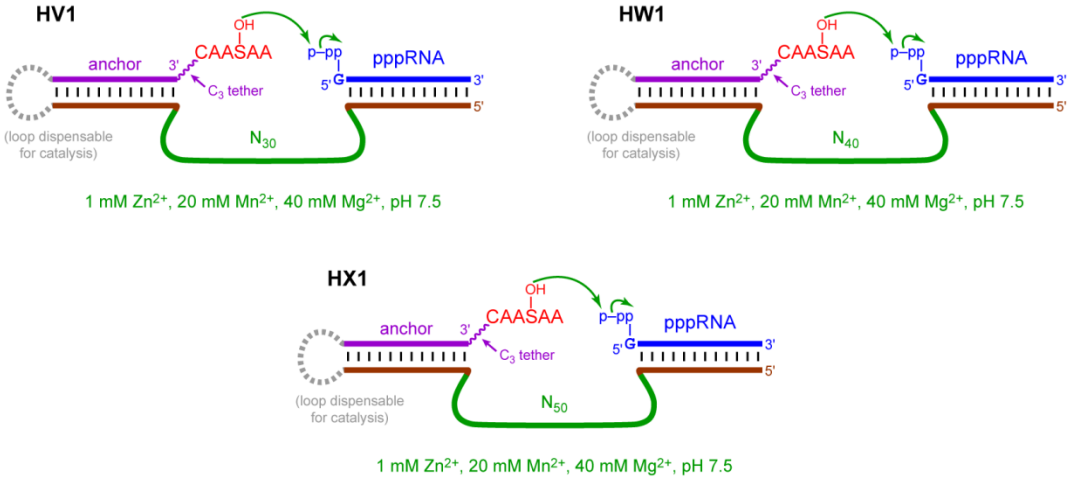


Figure 5.2. Design for the HV1-HX1 in vitro selections to identify serine kinase deoxyribozymes. In all selections, the peptide substrate CAASAA was connected to a DNA anchor via a disulfide bond and a C_3 tether, and the phosphoryl donor was pppRNA. The HV1, HW1, and HX1 selections had N_{30}, N_{40}, and N_{50} initially random regions, respectively. The selection step was performed in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h.

Each of these three selection experiments led to activity (Figure 5.3). Selection pressure for the ability to phosphorylate a more loosely HEG-tethered peptide substrate was applied after activity was observed. Selection pressure began at rounds 8, 7, and 10 for HV1, HW1, and HX1, respectively. After 7 rounds with the selection pressure, activity did not increase for either the HV1 or HX1 selection experiments. Therefore, the HV1 and HX1 selections using the short C_3 tether were cloned from rounds 7 and 10, respectively. Upon imposition of selection pressure, an increase in activity was observed in the HW1 selection. Therefore, the HW1 selection was cloned at round 6 from the selection using the short C_3 tether and round 11 from the selection pressure using the longer HEG tether.

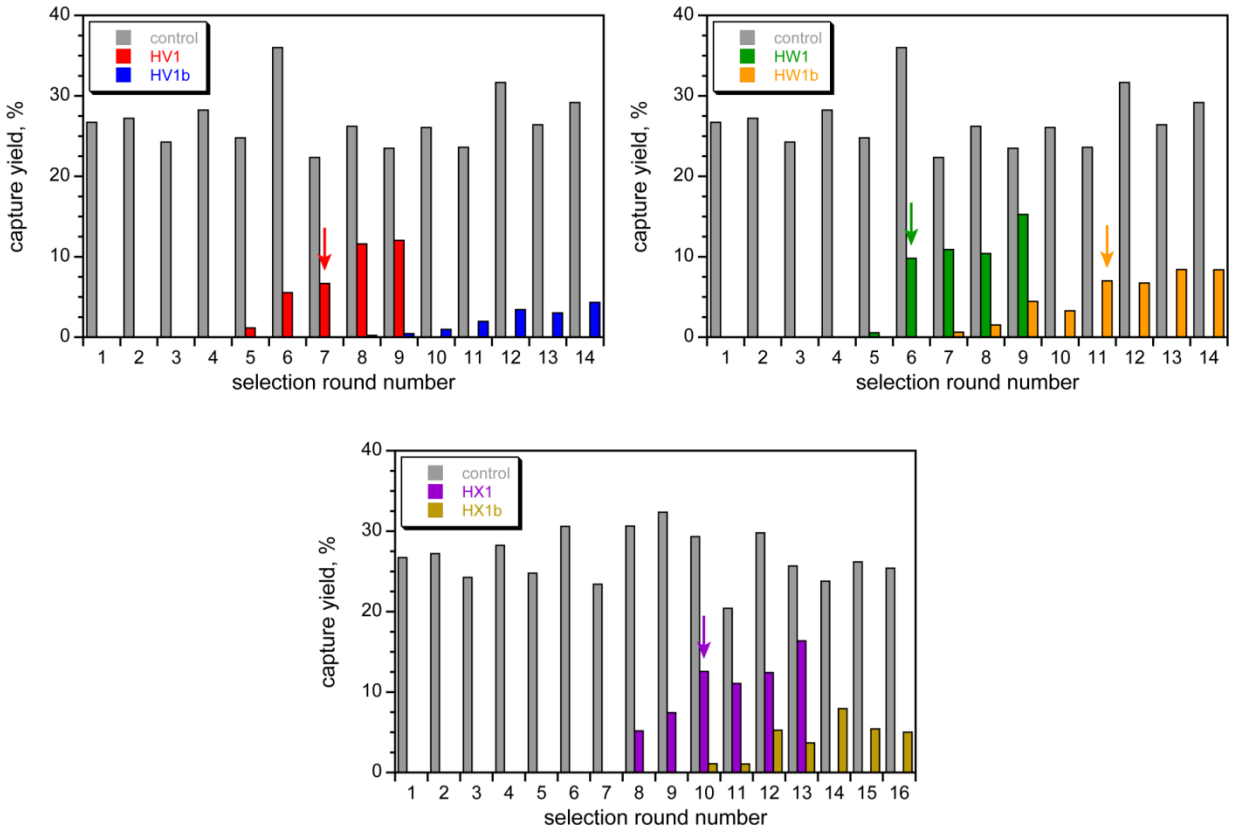


Figure 5.3. Progression of the in vitro selection experiments for serine kinase deoxyribozymes. In each round, "control" refers to the yield for the 8VP1-catalyzed capture reaction using a CAAS^{P}AA substrate, and the 8VP1-catalyzed capture reaction for the indicated selection experiment is shown. Arrows mark the cloned rounds. The HV1, HW1, and HX1 selections have N_{30}, N_{40}, and N_{50} random regions, respectively. The HV1, HW1, and HX1 selections have C_3-tethered peptide substrates, and the HV1b, HW1b, and HX1b selections have HEG-tethered peptide substrates.

Four, two, and two sequences were identified for each of HV1 (N_{30}), HW1 (N_{40}), and HX1 (N_{50}), respectively (Fig 5.4). All of the deoxyribozymes were assayed for serine phosphorylation activity in trans (without the substrate ligated to the pool). The seven deoxyribozymes identified from selection with the C_3-tethered peptide substrate all show robust phosphorylation yields ranging from 47 to 81% in 16 h (Fig 5.5A). None of these deoxyribozymes catalyze the phosphorylation of HEG-tethered CAASAA, and thus require the short C_3 tether. The 11HW104 deoxyribozyme identified from selection pressure with a longer HEG-tethered CAASAA peptide substrate has modest phosphorylation yields of 15% and 9% with the C_3 and HEG-tethered peptide substrate, respectively (Fig 5.5B). None of the

deoxyribozymes catalyze phosphorylation of a C_3-tethered CAAYAA substrate. Therefore, they are all specifically serine kinase deoxyribozymes.

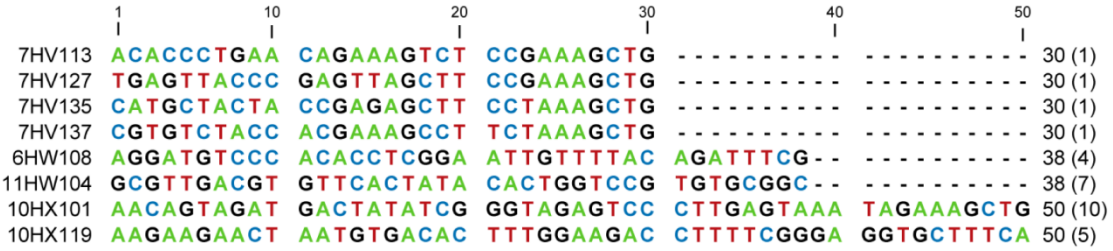


Figure 5.4. Sequences of the initially random region of the serine kinase deoxyribozymes. Only the initially random region is shown. Next to the sequence length on the right, in parentheses, is the number of times each sequence was found.

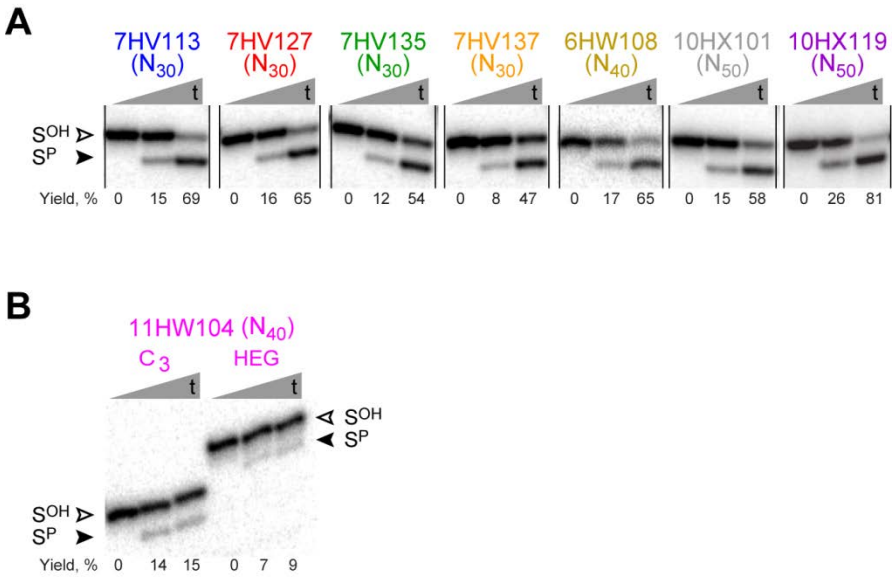


Figure 5.5. Activity of the serine kinase deoxyribozymes. All assays were performed in trans. Incubation conditions were 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C. S^{OH} = substrate; S^P = product. (A) Deoxyribozymes identified from selection with the C_3-tethered CAASAA peptide substrate. The PAGE image shows representative time points (t = 30 s, 2 h, 16 h) for seven of the deoxyribozymes assayed with the C_3-tethered peptide substrate. (B) The 11HW104 deoxyribozyme identified from selection with the HEG-tethered CAASAA peptide substrate. The PAGE image shows representative time points (t = 30 s, 2 h, 16 h) for 11HW104 with both the C_3 and HEG-tethered peptide substrate.

The identification of these serine kinase deoxyribozymes establishes that DNA can catalyze the phosphorylation of serine. However, limitations remain in the utility of these deoxyribozymes. Improvements necessary include the ability to phosphorylate more loosely tethered and untethered peptide substrates, to use ATP as the phosphoryl donor, and to phosphorylate serine residues in a sequence-selective manner.

Subsequently, the KW1-KZ1 selection experiments were performed using a HEG or HEG_6-tethered CAASAA peptide substrate and pppRNA as the phosphoryl donor (Figure 5.6A). Surprisingly, none of these efforts were successful, despite the previous success of HEG-tethered tyrosine-containing peptide substrates. Separately, LF1-LJ1 selection experiments were performed by Stephanie Konecki with tethered CAASAA peptide substrates and 100 μM ATP as the phosphoryl donor (Figure 5.6B). Previous efforts with DNA-catalyzed tyrosine phosphorylation were successful using GTP or ATP as the phosphoryl donor (Chapter 2),[14,15] suggesting that ATP could also be used in serine phosphorylation. However, no activity was observed for selection experiments with 100 μM ATP using N_{30}, N_{40}, or N_{50} initially random region lengths with C_3-tethered CAASAA or N_{40} with HEG-tethered CAASAA, and they were discontinued after 14 rounds.

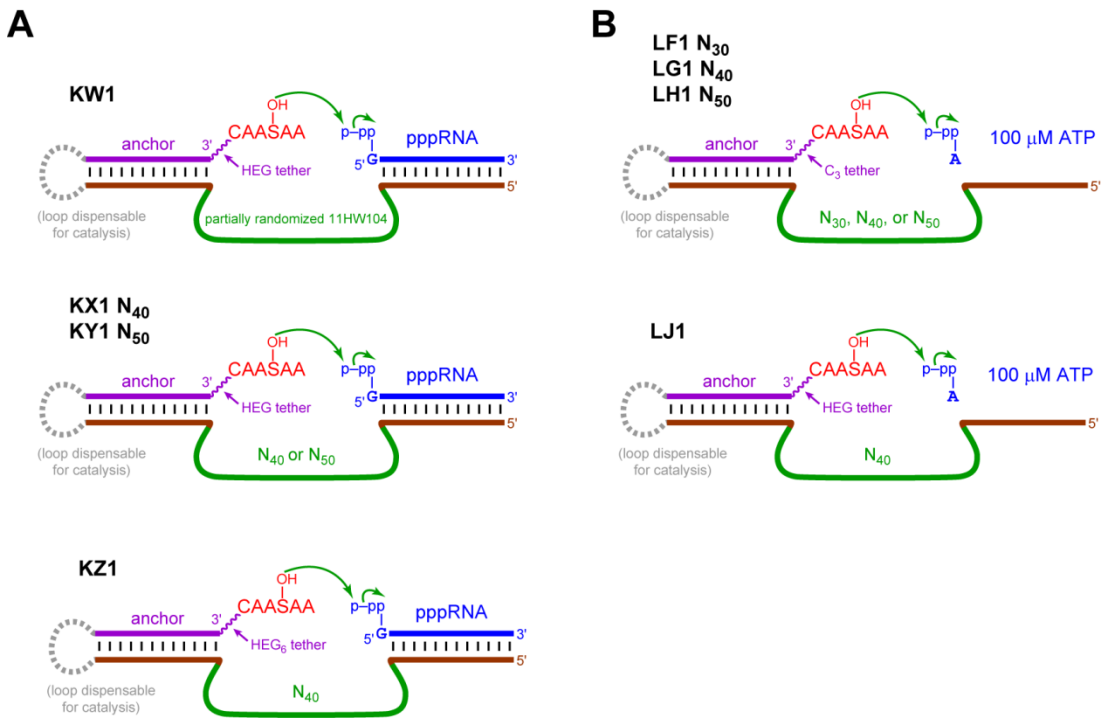


Figure 5.6. Design for the KW1-KZ1 and LF1-LJ1 in vitro selection experiments to identify serine kinase deoxyribozymes. In all selections, the peptide substrate CAASAA was connected to a DNA anchor via a disulfide bond and tether. Selection conditions were 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h. (A) The KW1-KZ1 selection experiments. The phosphoryl donor was pppRNA. The KW1-KY1 selections used a HEG tether, and the KZ1 selection used a HEG_6 tether. The KW1 selection was a reselection of 11HW104 from a partially randomized pool. The KX1 and KZ1 selection had an N_{40} initially random region, and the KY1 selection had an N_{50} initially random region. (B) The LF1-LJ1 selection experiments. The phosphoryl donor was 100 μM ATP. The LF1-LH1 selections used a C_3-tethered peptide substrate, and the LJ1 selections used a HEG-tethered peptide substrate. The length of the initially random region is indicated.

5.2.3 Use of Complex Serine-Containing Peptide Sequences

Tyrosine's aromatic hydroxyl group is more nucleophilic than serine's aliphatic hydroxyl group, and the tyrosine side chain is larger than serine's. Increasing the ability of the DNA catalysts to interact with the serine-containing peptide substrate may enable reactivity of more loosely tethered peptide substrates. To increase the ability of the DNA catalyst to interact with the serine-containing peptide substrate, peptides with non-alanine residues flanking the serine residue were used during in vitro selection experiments.

Separate selection experiments were performed by Nickolaus Lammer with three peptide substrates: CAASFA, CQKESTLH, and CNRTSLPW (Figure 5.7). The CAASFA peptide substrate was designed to contain serine and a single phenyl side chain via the phenylalanine residue. This aromatic group is most similar to a tyrosine residue, and DNA-catalyzed phosphorylation of peptide substrates containing a single tyrosine residue flanked by alanine residues is successful. The two other complex peptide sequences are derived from the amino acid sequence flanking the naturally phosphorylated serine residues. The CQKESTLH peptide corresponds to residues 62-68 of ubiquitin.[18-20] The CNRTSLPW peptide corresponds to residues 215-221 of casein kinase 1α (CK1α).[21] The N-terminal cysteine residue is incorporated to enable conjugation of each peptide substrate to the DNA-anchor.

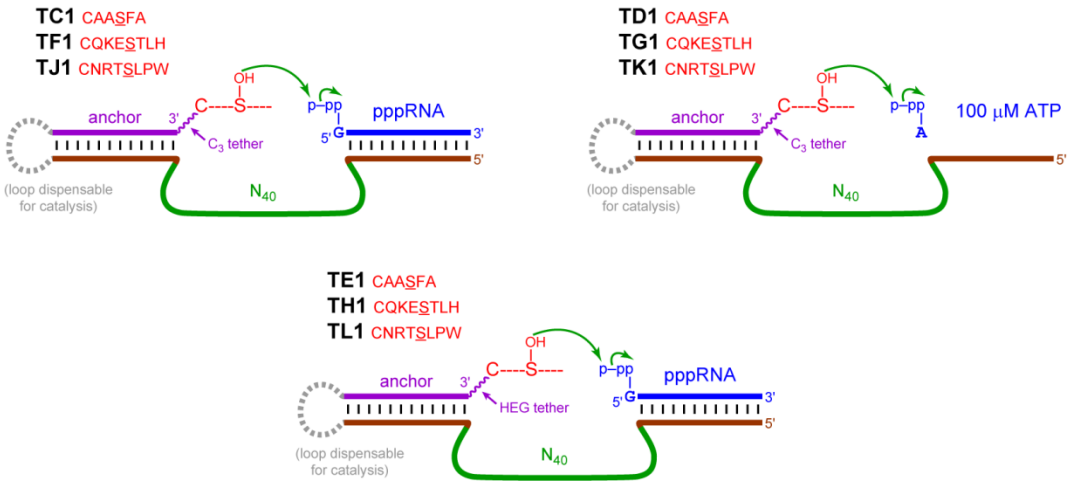


Figure 5.7. Design for the TC1-TL1 in vitro selection experiments to identify serine kinase deoxyribozymes. In all selections, the initially random region was N_{40}. In the TC1, TF1, and TJ1 selections, the peptide substrate was connected to a DNA anchor via a C_3 tether, and the phosphoryl donor was pppRNA. In the TD1, TG1, and TK1 selections, the peptide substrate was connected to a DNA anchor via a C_3 tether, and the phosphoryl donor was 100 μM ATP. In the TE1, TH1, and TL1 selections, the peptide substrate was connected to a DNA anchor via a HEG tether, and the phosphoryl donor was pppRNA. The peptide substrates were as indicated. All selections were performed in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C.

For each peptide substrate, three selection experiments were performed. A C_3-tethered peptide acceptor was used with either pppRNA or 100 μM ATP as the phosphoryl donor, and a

HEG-tethered peptide acceptor was used with a pppRNA phosphoryl donor. After 13 rounds of selection, activity was only observed in the TD1 selection experiment. Upon evaluation of the round 13 and 14 pools in trans, no phosphorylation activity was observed. Similar to previous kinase deoxyribozyme selections (Chapter 2, 4), the 8VP1 capture method was used, and the length and sequence of the capture oligonucleotide was alternated. The short capture oligonucleotide was a 17 nt pppRNA. The long capture oligonucleotide is a 54 nt pppRNA-DNA chimera made by ligating a 17 nt pppRNA oligonucleotide to a 37 nt DNA oligonucleotide using a DNA splint and T4 DNA ligase. Continuation of selection rounds showed that activity was observed only when the short capture oligonucleotide was used, suggesting that the activity was not due to the desired phosphorylated serine product (Figure 5.8).

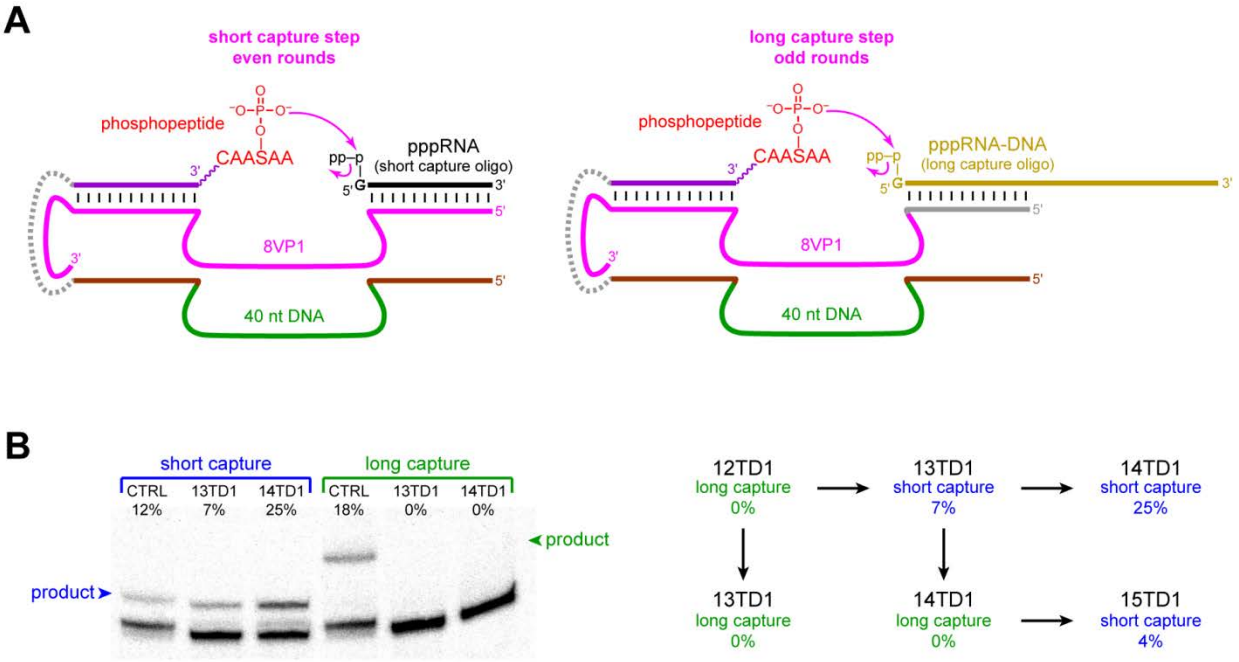


Figure 5.8. Results of the TD1 selection experiment. (A) The diagram of the 8VP1-catalyzed capture reaction shows the difference between the short and long capture oligonucleotide in alternating selection rounds. The long capture oligonucleotide was a RNA-DNA hybrid and has a different sequence than the short capture oligonucleotide. (B) The PAGE image shows the in cis capture reaction for rounds 13 and 14 using both the short and long capture oligonucleotide, and the diagram shows the capture yields of the selection rounds with the capture oligonucleotide indicated. No activity is observed when the long capture oligonucleotide was used.

Further evaluation showed that the pppRNA capture oligonucleotide was attached to the DNA pool and not the DNA-anchored peptide substrate. However, only the short capture oligonucleotide reacted with the pool, which is consistent with the in cis selection results. Both the short and long capture oligonucleotides share the same four nucleotides at the 5′-end (pppGGAA...), which may explain the tolerance of the long capture oligonucleotide despite the lack of observable yield. When the 5′-hydroxymethyl group of the DNA pool was replaced with only a methyl group, the pool was still modified. The location of the modification within the pool was not investigated further. The product resulting from attachment of the capture oligonucleotide to the DNA pool caused a PAGE shift in the same region as the desired capture product, and these DNA sequences were enriched. An alternate capture method for in vitro selection of serine kinase deoxyribozymes is needed to avoid this issue.

5.3 Summary and Future Directions

Serine phosphorylation is a common biological post-translational modification of proteins. Based on success with DNA-catalyzed tyrosine phosphorylation, deoxyribozymes that catalyze serine phosphorylation were sought. Previous DNA-catalyzed serine modification suggested that DNA-catalyzed serine phosphorylation was possible. The previously identified tyrosine kinase deoxyribozymes were identified from selection experiments using tyrosine-containing peptide substrates and are all specifically tyrosine kinase deoxyribozymes without serine reactivity. Therefore, selection experiments for serine kinase deoxyribozymes were performed with serine-containing peptide substrates. Similarly to the tyrosine kinase deoxyribozyme selection experiments, the 8VP1 capture deoxyribozyme was used in serine kinase selection experiments to isolate the active DNA sequences.

Serine kinase deoxyribozymes were identified to phosphorylate C_3-tethered CAASAA from these selection efforts. All of the DNA enzymes are specific for serine and do not catalyze tyrosine phosphorylation. This specificity is unlike the previously identified nucleopeptide-

forming deoxyribozymes that catalyze the modification of serine but also tyrosine. However, of the current serine kinase deoxyribozymes, only one can catalyze the phosphorylation of HEG-tethered peptide substrates. These results establish that DNA can catalyze serine phosphorylation, but some limitations remain. Selection experiments with HEG-tethered serine-containing peptide substrates or ATP as the phosphoryl donor were unsuccessful, and the use of small-molecule NTP phosphoryl donors remains a goal of future efforts. The characteristics of the serine kinase deoxyribozymes are similar to those of nucleopeptide-forming deoxyribozymes with serine substrates.[2] Robust activity is observed with C_3-tethered peptides, but little activity is observed with more loosely HEG-tethered substrates.

The use of more complex serine-containing peptide sequences was thought to enable more interactions between the peptide substrate and the deoxyribozyme. These interactions may enable phosphorylation of peptides connected via longer HEG tethers or untethered peptide substrates and may enable peptide-sequence selectivity. Unexpectedly, these experiments led to DNA-catalyzed reaction of the capture oligonucleotide with the DNA pool instead of phosphorylation of the peptide substrate. To prevent this undesired reaction from dominating the pool during selection experiments, a new capture method is needed. The phosphotyrosine immunoprecipitation capture method works well for tyrosine kinase deoxyribozyme selections, and a parallel method for phosphoserine could be developed. However, there are only a few commercially available phosphoserine antibodies, and they are cost-prohibitive for use during in vitro selection experiments. Therefore, a Phos-tag PAGE selection method may be useful for these efforts.[22-25] Phos-tag will bind to the phosphoserine products connected to the active DNA sequences and cause these DNA sequences to migrate slower on a PAGE gel, therefore separating the active sequences from the inactive ones. Preliminary tests indicate that this method may provide a large enough PAGE shift for use during in vitro selection experiments. Currently, efforts are underway to synthesize Phos-tag acrylamide.[26]

The use of modified nucleotides has been successful at enabling catalysis of reactions previously unable to be catalyzed by deoxyribozymes with only natural nucleotides.[27] While

serine phosphorylation can be catalyzed by natural DNA, incorporation of modified nucleotides may enable catalysis to have more desired features such as increased rate and yield, reactivity with untethered peptides, and peptide-sequence specificity. The mechanism of known protein kinases involves interactions between lysine, aspartate, glutamate, and asparagine residues with the phosphoryl donor and serine hydroxyl substrate.[28] Therefore, nucleotides with amino modifications (^{Am}dU, mimicking lysine) or carboxyl modifications (^{COOH}dU, mimicking aspartate and glutamate) were incorporated into the initially random region of the DNA catalysts.

The VP1-VS1 selection experiments were initiated by Stephanie Konecki and continued by Tiyaporn Tangpradabkul. The four selection experiments were performed with the CQKESTLH peptide substrate. The longer HEG tether was used with pppRNA as the phosphoryl donor. Separately, the short C_3 tether was used with 100 μM ATP as the phosphoryl donor. For each, either ^{Am}dU or ^{COOH}dU was incorporated. These efforts are ongoing.

5.4 Materials and Methods

5.4.1 Substrate Preparation Procedures

Oligonucleotides, peptides, and DNA-anchored peptide conjugates were prepared as described in Chapter 2.

Table 5.1: Oligonucleotide sequences of the pools, substrates, and primers used in the selection experiments.

oligonucleotide purpose	oligonucleotide sequence
HV1 selection with N_{30} random region	
DNA-C_3-CAASAA substrate	GGACTACCTTTATGCGTAT-C_3-CAASAA
pppRNA phosphoryl donor	pppGGAAGGAGGCUUUCGGG
forward primer for selection	CGAACGAAAGCCTCCTTC
reverse primer for selection	$(AAC)_4$-HEG-CCATCAGGATCAGCT
reverse primer for cloning	TAATTAATTAATTACCCATCAGGATCAGCT
random pool for selection	CGAACGAAAGCCTCCTTC-N_{30}-ATACGCATAAAGGTAGAGCTGATCCTGATGG
splint for ligation step during selection	ATACGCATAAAGGTAGTCCTCCATCAGGATCAGCTCTACCTTTATGCGTAT

Table 5.1 (cont.)

oligonucleotide purpose	oligonucleotide sequence
HW1 selection with N_{40} random region	
DNA-C_3-CAASAA substrate	GGAATATCTCGTTTCTTAT-C_3-CAASAA
pppRNA phosphoryl donor	pppGGAAUUCAGUCUUAAGG
forward primer for selection	CGAATTAAGACTGAATTC
reverse primer for selection	$(AAC)_4$-HEG-CCATCAGGATCAGCT
reverse primer for cloning	TAATTAATTAATTACCCATCAGGATCAGCT
partially randomized pool for selection	CGAATTAAGACTGAATTC-N_{40}-ATAAGAAACGAGATATAGCTGATCCTGATGG
splint for ligation step during selection	ATAAGAAACGAGATATTCCTCCATCAGGATCAGCTATATCTCGTTTCTTAT
HX1 selection with N_{50} random region	
DNA-C_3-CAASAA substrate	GGAATGGCTTGATTGGTAT-C_3-CAASAA
pppRNA phosphoryl donor	pppGGAAUAUUUACUCAAGG
forward primer for selection	CGAATTGAGTAAATATTC
reverse primer for selection	$(AAC)_4$-HEG-CCATCAGGATCAGCT
reverse primer for cloning	TAATTAATTAATTACCCATCAGGATCAGCT
random pool for selection	CGAATTGAGTAAATATTC-N_{50}-ATACCAATCAAGCCATAGCTGATCCTGATGG
splint for ligation step during selection	ATACCAATCAAGCCATTCCTCCATCAGGATCAGCTATGGCTTGATTGGTAT

5.4.2 In Vitro Selection Procedures

Procedures for ligation, selection, 8VP1 capture, and PCR) were performed as described in Chapter 2.

5.4.3 Cloning and Screening

Procedures for cloning and screening were performed as described in Chapter 2.

5.4.4 Deoxyribozyme Activity Assay Procedure

The DNA-anchored hexapeptide substrate was 5′-^{32}P-radiolabeled using γ-^{32}P-ATP and T4 polynucleotide kinase (Fermentas), using 10× buffer that lacks DTT (500 mM Tris, pH 7.6, 100 mM $MgCl_2$, and 1 mM spermidine). A 10 µL sample containing 0.25 pmol of 5′-^{32}P-radiolabeled DNA-anchored peptide substrate, 10 pmol of deoxyribozyme, and 20 pmol of pppRNA phosphoryl donor was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The DNA-catalyzed

phosphorylation reaction was initiated by bringing the sample to 20 µL total volume containing 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl (or other ion concentrations as appropriate). The Mn^{2+} was added from a 10× stock solution containing 200 mM $MnCl_2$. The Zn^{2+} was added from a 10× stock solution containing 5 mM $ZnCl_2$, 10 mM HNO_3, and 100 mM HEPES at pH 7.5; this stock solution was freshly prepared from a 200× stock of 100 mM $ZnCl_2$ in 200 mM HNO_3. The metal ion stocks were added last to the final sample, which was divided into 2-µL aliquots that were all incubated at 37 °C. At appropriate time points, 2 µL aliquots were quenched with 5 µL stop solution (80% formamide, 1× TBE [89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3], 50 mM EDTA, 0.025% bromophenol blue, 0.025% xylene cyanol). Samples were separated by 20% PAGE and quantified with a PhosphorImager. Values of k_{obs} were obtained by fitting the yield versus time data directly to first-order kinetics; i.e., yield $Y = Y_{max} \bullet (1 - e^{-kt})$, where $k = k_{obs}$ and Y_{max} is the final yield. Errors in k_{obs} values were calculated as the standard deviation from the indicated number of independent determinations.

5.5 References

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

(2) 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.

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

(4) Lorsch, J. R.; Szostak, J. W. Kinetic and thermodynamic characterization of the reaction catalyzed by a polynucleotide kinase ribozyme. *Biochemistry* **1995**, *34*, 15315-15327.

(5) Burke, D. H.; Rhee, S. S. Assembly and activation of a kinase ribozyme. *RNA* **2010**, *16*, 2349-2359.

(6) Biondi, E.; Maxwell, A. W.; Burke, D. H. A small ribozyme with dual-site kinase activity. *Nucleic Acids Res.* **2012**, *40*, 7528-7540.

(7) Curtis, E. A.; Bartel, D. P. Synthetic shuffling and in vitro selection reveal the rugged adaptive fitness landscape of a kinase ribozyme. *RNA* **2013**, *19*, 1116-1128.

(8) Li, Y.; Breaker, R. R. Phosphorylating DNA with DNA. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 2746-2751.

(9) Li, Y.; Breaker, R. R. In vitro selection of kinase and ligase deoxyribozymes. *Methods* **2001**, *23*, 179-190.

(10) 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.

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

(12) Sachdeva, A. "Deoxyribozymes for Peptide Substrates: Exploring the Landscape of Nucleophiles and Electrophiles". Ph.D. Thesis, University of Illinois at Urbana-Champaign, **2011**.

(13) Camden, A. J.; Walsh, S. M.; Suk, S. H.; Silverman, S. K. DNA oligonucleotide 3'-phosphorylation by a DNA enzyme. *Biochemistry* **2016**, *55*, 2671-2676.

(14) Walsh, S. M.; Sachdeva, A.; Silverman, S. K. DNA catalysts with tyrosine kinase activity. *J. Am. Chem. Soc.* **2013**, *135*, 14928-14931.

(15) Dokukin, V.; Silverman, S. K. A modular tyrosine kinase deoxyribozyme with discrete aptamer and catalyst domains. *Chem. Commun.* **2014**, *50*, 9317-9320.

(16) Walsh, S. M.; Konecki, S. N.; Silverman, S. K. Identification of sequence-selective tyrosine kinase deoxyribozymes. *J. Mol. Evol.* **2015**, *81*, 218-224.

(17) Sachdeva, A.; Chandra, M.; Chandrasekar, J.; Silverman, S. K. Covalent tagging of phosphorylated peptides by phosphate-specific deoxyribozymes. *ChemBioChem* **2012**, *13*, 654-657.

(18) Koyano, F.; Okatsu, K.; Kosako, H.; Tamura, Y.; Go, E.; Kimura, M.; Kimura, Y.; Tsuchiya, H.; Yoshihara, H.; Hirokawa, T.; Endo, T.; Fon, E. A.; Trempe, J. F.; Saeki, Y.; Tanaka, K.; Matsuda, N. Ubiquitin is phosphorylated by PINK1 to activate parkin. *Nature* **2014**, *510*, 162-166.

(19) Kane, L. A.; Lazarou, M.; Fogel, A. I.; Li, Y.; Yamano, K.; Sarraf, S. A.; Banerjee, S.; Youle, R. J. PINK1 phosphorylates ubiquitin to activate Parkin E3 ubiquitin ligase activity. *J. Cell Biol.* **2014**, *205*, 143-153.

(20) Kazlauskaite, A.; Kondapalli, C.; Gourlay, R.; Campbell, D. G.; Ritorto, M. S.; Hofmann, K.; Alessi, D. R.; Knebel, A.; Trost, M.; Muqit, M. M. Parkin is activated by PINK1-dependent phosphorylation of ubiquitin at Ser65. *Biochem. J.* **2014**, *460*, 127-139.

(21) Clokie, S.; Falconer, H.; Mackie, S.; Dubois, T.; Aitken, A. The interaction between casein kinase Ialpha and 14-3-3 is phosphorylation dependent. *FEBS J.* **2009**, *276*, 6971-6984.

(22) Kinoshita, E.; Takahashi, M.; Takeda, H.; Shiro, M.; Koike, T. Recognition of phosphate monoester dianion by an alkoxide-bridged dinuclear zinc(II) complex. *Dalton Trans.* **2004**, 1189-1193.

(23) Kinoshita, E.; Yamada, A.; Takeda, H.; Kinoshita-Kikuta, E.; Koike, T. Novel immobilized zinc(II) affinity chromatography for phosphopeptides and phosphorylated proteins. *J. Sep. Sci.* **2005**, *28*, 155-162.

(24) Kinoshita, E.; Kinoshita-Kikuta, E.; Takiyama, K.; Koike, T. Phosphate-binding tag, a new tool to visualize phosphorylated proteins. *Mol. Cell. Proteomics* **2006**, *5*, 749-757.

(25) Kinoshita, E.; Kinoshita-Kikuta, E.; Koike, T. A single nucleotide polymorphism genotyping method using phosphate-affinity polyacrylamide gel electrophoresis. *Anal. Biochem.* **2007**, *361*, 294-298.

(26) Tilmans, N. P.; Krusemark, C. J.; Harbury, P. A. Expedient synthesis of a modular phosphate affinity reagent. *Bioconjug. Chem.* **2010**, *21*, 1010-1013.

(27) Zhou, C.; Avins, J. L.; Klauser, P. C.; Brandsen, B. M.; Lee, Y.; Silverman, S. K. DNA-catalyzed amide hydrolysis. *J. Am. Chem. Soc.* **2016**, *138*, 2106-2109.

(28) Adams, J. A. Kinetic and catalytic mechanisms of protein kinases. *Chem. Rev.* **2001**, *101*, 2271-2290.