# DNA CATALYSTS AS PHOSPHATASES AND AS PHOSPHOSERINE LYASES

BY

JAGADEESWARAN CHANDRASEKAR

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Scott K. Silverman, Chair
Professor Ryan C. Bailey
Professor Paul J. Hergenrother
Professor Wilfred A. van der Donk

## ABSTRACT

Proteins and RNA are the only known biopolymers that have catalytic roles in nature, whereas DNA is primarily considered to store and transfer genetic information. However, artificial single-stranded DNA has been identified by in vitro selection to catalyze several chemical reactions and several of those are of biological relevance. For in vitro selection or directed evolution of proteins, direct amplification is not possible, and it essential to attach the genotype to the phenotype. For nucleic acids however, the functional biopolymer can be readily amplified. DNA has the advantage of being directly amplified by polymerases, whereas RNA requires an additional reverse transcription step. Moreover, DNA catalysts identified by in vitro selection processes have shown similar catalytic proficiency as RNA. DNA has added advantages of low cost of chemical synthesis and higher stability. Considering these factors combined, identification of artificial DNA catalysts (deoxyribozymes) for chemical reactions is a valuable endeavor with long-term implications.

Protein post-translational modifications (PTMs) are highly important in biological processes involving cellular regulation. Additionally, PTMs serve as important intermediates or key motifs on natural products and bioactive peptides. The natural protein enzymes carrying out the essential modifications may have several shortcomings for biotechnological use. Identification of artificial DNA catalysts with ability to perform chemoselective post-translation chemical reaction would be highly useful in studying biological regulatory processes, performing synthesis and late-stage diversification of post-translationally modified peptides, as well as carrying out other important functions that natural proteins may not readily solve.

My first effort was to identify DNA enzymes with peptide/protein phosphatase activity, more specifically dephosphorylation of peptide side chains. Without a catalyst, phosphomonoester hydrolysis reactions have exceedingly low spontaneous reaction rates. Nature utilizes proficient protein enzymes to perform this challenging reaction with great efficiency. Using a known DNA catalyst for the in vitro selection process, new DNA catalysts were identified with phosphatase activity. The phosphatase DNA catalysts exhibited multiple-turnover activity with phosphotyrosine-containing free peptides and were active even in the presence of externally added cell lysate or bovine serum albumin (BSA). Furthermore, the best DNA phosphatase functioned with a larger protein substrate. This established the fundamental ability of DNA to catalyze dephosphorylation of amino acid side chain residues. The study also suggested that phosphatase DNA catalysts could perform intracellular phosphatase activity. Hence, these deoxyribozymes were functionalized on gold nanoparticles and delivered inside live mammalian cells to investigate if they behave as functional protein analogues (or mimics) of recombinantly expressed Protein Tyrosine Phosphatase (PTP1B). Separately, efforts were directed towards the important goal of identifying sequence-selective phosphatase deoxyribozymes. Although three separate efforts were directed towards identifying sequence-selective phosphatases deoxyribozymes, we were unsuccessful in accomplishing this specific goal of selectivity in the context of peptide sequence discrimination.

Dehydroalanine (Dha) is a non-proteinogenic electrophilic amino acid that serves as a synthetic intermediate or product in the biosynthesis of several bioactive cyclic peptides such as lantibiotics, thiopeptides and microcystins. DNA enzymes were

identified to establish the fundamental catalytic ability to eliminate phosphate from phosphoserine (pSer) to form Dha, namely phosphoserine lyase activity. Furthermore, DhaDz1 was utilized to achieve chemo-enzymatic synthesis of a cyclic cystathionine-containing peptide. Based on this initial success, future efforts will be directed to achieve sequence-general phosphoserine and phosphotyrosine lyase activity. Separately, application of sequence-general lyases in the synthesis of complex lanthipeptides and enrichment of phosphopeptides/proteins in phosphoproteomics will be explored.

## ACKNOWLEDGMENTS

I am extremely grateful to everyone who participated in my graduate journey here at UIUC. First and foremost, I owe my sincere gratitude to Professor Scott Silverman for his patience, guidance, and phenomenal mentorship. I am thankful for each of those one-on-one discussions we have shared that helped me grow both as a scientific researcher and as a person. I am confident that the lessons I have acquired under his guidance will continue to serve me well in my future endeavors.

I am also grateful to my thesis committee members, Professor Wilfred van der Donk, Professor Ryan Bailey, and Professor Paul Hergenrother for their thoughtful suggestions and advices. Interacting with them has always been a great learning experience.

I had the privilege of working with the fantastic Silverman crew. I am thankful to Amit, Adrienne, Madhav and Josh who initially trained and mentored me. I am grateful to Ying, Victor, Ben and Spurti their friendship. I also had the pleasure of working with amazing friends like Shannon, Jimmy, Yves, Puzhou, Peter and Cong. I had an opportunity to mentor wonderful students like Natalia, Adam, Natalie and Miriam who made my graduate experience extremely cheerful.

I want to thank our remarkable chemistry staff members, facility heads, and my other chemistry compatriots. Jessie, Subha, Manuel, Joe, Jonathan, and Soumitra provided me with a lot of help and assistance with unfamiliar research fields. Most importantly, I am grateful to my parents and my sister for their continual support and love. I am truly indebted to them.

# Table of Contents

Chapter 1: Introduction to Enzymatic Catalysis and Applications ..... 1
Chapter 2: Catalytic DNA with Phosphatase Activity ..... 23
Chapter 3: Efforts to Identify Sequence-Selective Phosphatase Deoxyribozymes ..... 65
Chapter 4: Intracellular Application of Phosphatase DNA Catalysts ..... 105
Chapter 5: Catalytic DNA with Phosphoserine and Phosphothreonine Lyase Activity ..... 134

# Chapter 1: Introduction to Enzymatic Catalysis and Applications

## 1.1 Enzyme Rate Enhancements

Biologically relevant chemical reactions have extremely slow spontaneous reaction rates, including exceedingly high half-lives. The sluggishness typically arises because chemical reactions necessitate formation of unfavorable transition states with unusual buildup of positive or negative charges. Reactions also require atoms and molecules to be brought together in a specific orientation that are often entropically unfavorable. However, enzymes achieve tremendous rate enhancements by performing these challenging reactions with high rates, chemoselectivity, and stereospecificity. This section briefly reviews how nature achieves high rate enhancements or accelerations.

### 1.1.1 Proximity and Orientation

Entropy reduction is a key feature of enzymes. The enzymes bring their substrates in a proper spatial orientation for catalysis by minimizing the relative translational and rotational motions. More specifically in cases involving ligation-like reactions, enzymes provide high effective molarity. Jencks initially proposed that entropy is a major contributing factor for high rate accelerations in enzymatic reactions.[1] However, Bruice later proposed that improvement in rate enhancements arise mainly from enthalpic ($\Delta H$) rather than entropic ($T\Delta S$) contributions, with experimental evidence supporting the Nearest Active Conformation theory (NAC).[2] Conformational restriction (or NAC) is primarily determined by enthalpy, and the contribution of entropy to catalysis was originally overestimated.[3] Enzymes can present different activated precursors to close proximity to improve the reaction rates, but true catalysis is attributed to enthalpy.

### 1.1.2 Electrostatics

Desolvation occurs when the substrate binds to the enzyme. This may lead to change in the local dielectric constant, which in turn increases the electrostatic interactions in the active site. Furthermore, certain enzymes may create a protective pocket to shield the reactive groups from water to prevent formation of unwanted side products. If water itself is a secondary substrate, as in the case of hydrolysis reactions, a specific water molecule has to be held in the active site. In all such cases, enzyme has to provide substantial electrostatic stabilization for stripping and shielding the substrates/TS or the intermediates from the solvent. Electrostatic effects are established to be one of the largest contributors of catalysis.[4] Charged transition states are stabilized by active-site residues in enzymes forming ionic bonds. The stabilization usually come from acidic or basic side chains found on the amino acids (lysine, glutamic acid, aspartic acid, arginine) or from the metal cofactors such as magnesium or zinc.[5]

### 1.1.3 Catalysis by Preferential Transition State Binding

Enzymes may have a preferential affinity for the TS over the substrates or the products. Earlier known as “strain and distortion”, preferential binding the TS leads to fast reaction rates. The enzyme leads to reduction in the free energy of the TS and catalysis is achieved. The concept for TS stabilization was initially called strain because it was theorized that enzymes strongly restrain the substrate in a specific transition state conformation to enable catalysis. However, theoretical studies predict that enzymes are too flexible to force conformational strains on the substrates.[6] Binding of substrates to the enzymes by itself does not lead to better catalysis. The more likely scenario is for the enzyme to achieve an induced fit to enable catalysis.[7] This concept of catalysis by

preferential transition-state binding has been utilized in two major fields; namely transition-state inhibitors and catalytic antibodies.

Transition-state analogues are designed on the premise that they bind stronger to enzymes than the actual substrates themselves. Pharmaceutical and medicinal chemists have capitalized on this concept by studying transition states using kinetic isotope effects (KIE) and computational simulations. Some successful examples involve design of methylthioadenosine nucleosidase,[8] thermolysin,[9] and arginase transition-state analogues.[10] (Figure 1.1)

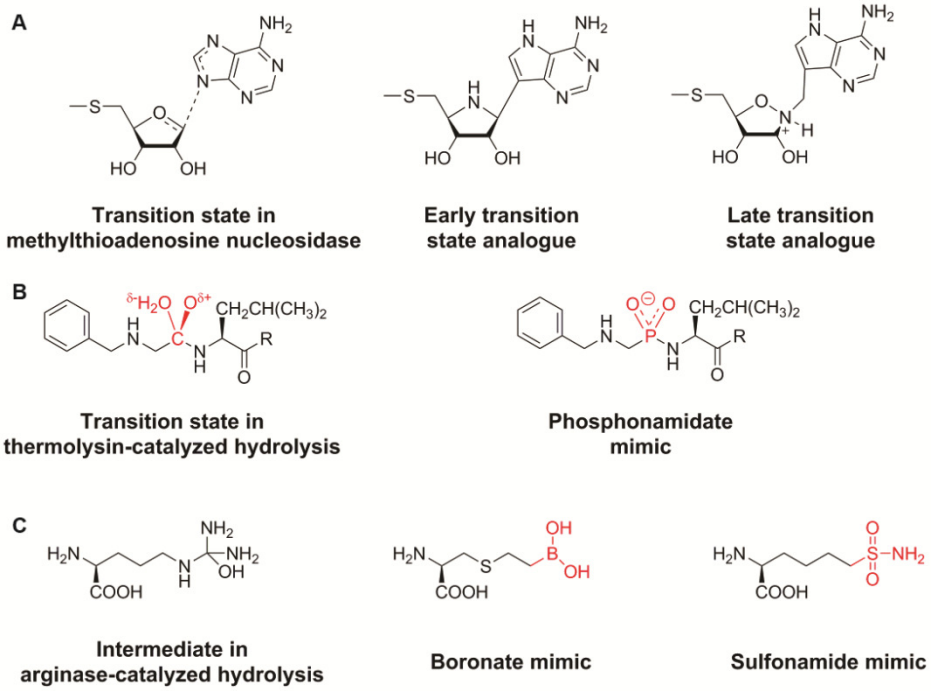


**Figure 1.1.** Transition-state mimics or analogues are successfully used to inhibit enzymes, suggesting that enzymes often have preferential binding affinity for transition states over the substrates or the products. Examples of inhibitors designed as analogues of transition states or intermediates for a few enzymes are listed here. A) Methylthioadenosine nucleosidase inhibitor.[8] B) Thermolysin inhibitor.[9] C) Arginase inhibitor.[10]

Separately, Schultz developed catalytic antibodies and abzymes that strongly rely on transition-state analogues. Antigens are designed to resemble transition states of known chemical reactions. These antigens are used as haptens to develop antibodies that

tightly bind to the transition states. The resulting antibodies are then used as catalysts for chemical reactions.[11] Catalytic antibodies have been explored for therapeutic potential in combating weight gain and drug addiction, treating microbial infections, and curing genetic diseases. Separately, abzymes have also been utilized for non-therapeutic applications such as Diels-Alder reaction and oxidation of water.[12-13]

### 1.1.4 Acid-Base Catalysis

In several enzymatic reactions such as hydrolysis of peptides/esters, nucleophilic addition on carbonyls and phosphoryl transfer reactions, general acid-base catalysis are involved. Formation of charged intermediate in chemical reactions is unfavorable and this leads to high free energies of the TS. Enzymes can appropriately donate or accept protons to enable catalysis.[14] Proteins have several amino acid side chains that function as general acid/base catalysts such as imidazoles, amines, carboxyls, thiols, aliphatic/aromatic hydroxyls, and guanidines ( from arginine) (Figure 1.2). pH can affect the protonation state of the enzyme functional groups. pH of the solution also strongly affects the catalytic proficiency of the enzymes that function with acid-base catalytic mechanism. When both protonation and deprotonation occur synchronously, then the process is called concerted acid-base catalysis.

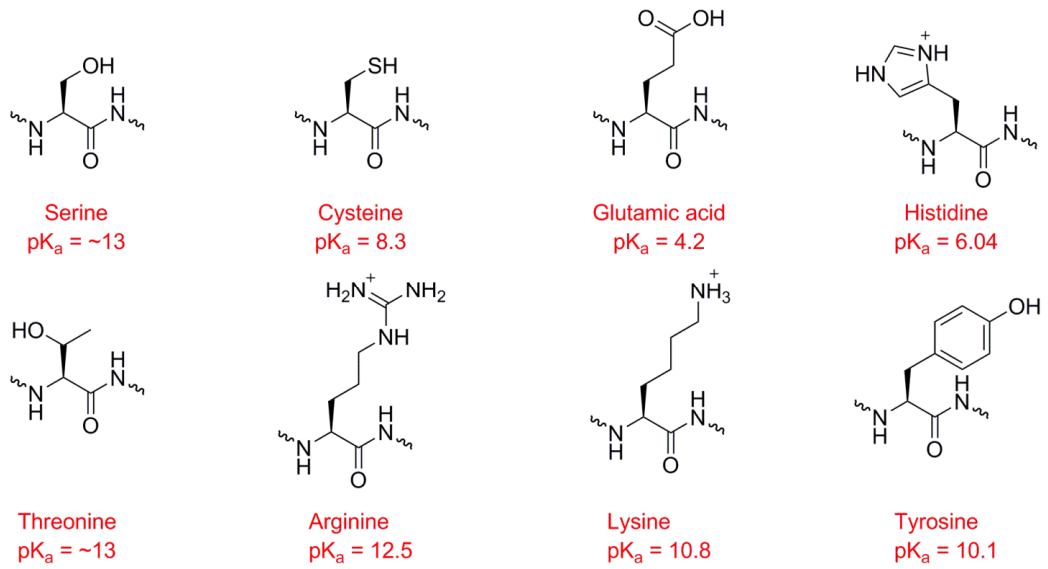


**Figure 1.2.** Protein side chains have functional groups that are capable of general acid-base catalysis. The amino acids shown here are typically found at protein enzyme active sites. Their $pK_a$'s of these amino acids are listed as well.

### 1.1.5 Metal Ion Catalysis

Metal ions are often engaged in enzymatic catalysis. They can be used for binding and orienting the substrates (such as metal-bound water), assisting with oxidation-reduction, or electrostatically shielding negative charges. Metalloenzymes usually contain tightly bound metal ions at the active site to achieve catalysis (strong $K_d$).[15] In cases where metal ions are loosely bound, the enzymes are called metal-activated enzymes. Metal-organic compounds like heme may merely function as prosthetic or signaling groups.[16] Enzymes such as phosphatases utilize metal-bound hydroxide for attacking the phosphomonoester, which lowers the $pK_a$ of water.[17] Metalloproteases use metals to stabilize the negative charge formed in the tetrahedral oxyanion intermediates.[18]

## 1.2 Nucleic Acid Catalysts

Until the 1980s, it was thought that nature exclusively used proteins to perform catalysis. However, in the 1980s Thomas Cech and Sidney Altman independently

discovered that RNA (polymer of ribonucleotides) are enzymatic.[19] Following this discovery, Breaker and Joyce identified a single-stranded DNA catalyst to cleave RNA.[20]

### 1.2.1 Ribozymes

RNA molecules that catalyze specific biochemical reactions are called ribozymes (**ribo**nucleic acid en**zyme**s). Ribozymes are found in nature or can be artificially developed by in vitro selection. Some of the chemical reactions catalyzed by ribozymes are DNA/RNA cleavage, DNA/RNA ligation and peptide bond formation.[5] In ribosomal protein synthesis, ribosomal RNA is a major catalytic subunit that is used to link the specific amino acid residues. Ribozymes also perform a variety of important biological functions such as RNA splicing, viral replication and tRNA biosynthesis.

RNA has received a lot of focus in the field of "origin of life". Ribozymes have been discovered in laboratory that can catalyze their own synthesis (self-replication) from activated monomers using RNA as the template.[21] Further improvements have been made to this ribozyme to add up to 20 nucleotides to a primer template.[22] Holliger and co-workers identified ribozyme-catalyzed transcription of an active ribozyme.[23]

Several natural ribozymes have been established. RNase P was first identified by Sidney Altman. In an independent effort, Thomas Cech identified group I intron. Both Altman and Cech were awarded Nobel Prize for their discovery. Since then, other ribozymes have been identified that include group II self-splicing intron, hairpin ribozyme, hammerhead ribozyme, GIR1 branching ribozyme, *glmS* ribozyme, HDV ribozyme, twister family ribozyme, and VS ribozyme.[19]

Artificial ribozymes have received substantial attention as potential therapeutic agents. RNA sequences have been targeted for sequence-specific cleavage.[24] Another

popular application includes biosensing.[25] RNA-based enzymes however have low stability (easily degraded both *in vitro* and *in vivo*), which is a major challenge in ribozyme therapy. Ribozymes have been explored as oligonucleotide therapeutics to target HIV RNA gene,[26] and hepatitis C viral RNA.[27]

### 1.2.2 Deoxyribozymes

DNA molecules that catalyze specific biochemical reactions are called deoxyribozymes (**deoxyribo**nucleic **acid enzymes)**. As of 2016, no natural DNA enzymes have been found in nature. DNA is widely regarded to store and carry genetic information. However, Breaker and Joyce identified the first artificial DNA catalyst to cleave a rA substrate embedded in a DNA.[20] Some of the DNA-catalyzed reactions include transesterification of RNA, hydrolysis of DNA and RNA, DNA phosphorylation, DNA adenylation, and DNA deglycosylation.[28] More recently, DNA catalysts have been identified to modify peptides and proteins.[29]

### 1.2.3 Functional Groups in Nucleic Acids for Catalysis

Proteins are armed with a diverse set of functional groups that enable general acid-base catalysis. As mentioned earlier, proteins also have functional groups like amine, hydroxyls, thiols, and carboxylates that can have hard-hard or soft-soft interactions with metal cations resulting in either metal-assisted protein catalysis or metalloenzyme catalysis.

Nucleic acids however have four nucleobases with primary or secondary heterocyclic (aryl) amines that have $pK_a$ values over 9, whereas the $pK_a$ of protonated N7 position of adenine/guanine and N3 position of cytosine is less than 4.[30] The RNA 2′-OH is predicted to have $pK_a$ of the range 12.5-13 whereas DNA lacks any hydroxyl groups.[31]

The phosphodiester $pK_a$ value lies close to 2 (Figure 1.3A). Additionally the phosphodiester is also known to have strong affinity with divalent metal cations like $Mg^{2+}$.

Although nucleic acids do not possess the necessary functional groups to match the catalytic repertoire of proteins, they do have substantial catalytic prowess. Initially, ribozymes were predicted to be obligatory metalloenzymes.[32] However, biochemical and structural studies indicate that both ribozymes and deoxyribozymes are capable of performing general acid-base catalysis as well even with functional groups having $pK_a$'s well outside the physiological pH range.[33-35] (Figure 1.3B)

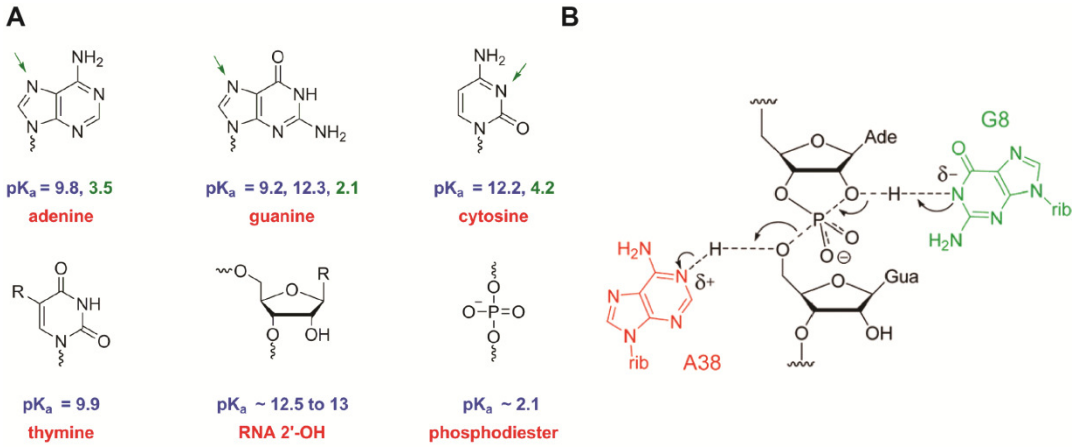


**Figure 1.3.** A) $pK_a$ of functional groups found in nucleic acids are listed here. The green values represent protonated form of the nitrogen. B) Mechanistic evidence for general acid-base catalysis mediated by nucleobases in the hairpin ribozyme (figure from reference 35).

## 1.3 DNA Catalysts

DNA encompasses several features that make it an exceptional choice of biopolymer for performing in vitro selection to identify and discover new biocatalysts. DNA is readily synthesized by solid-phase synthesis. Caruthers and co-workers developed phosphoramidite chemistry that is highly efficient in comparison to the iterative amide coupling chemistry. Because DNA lacks the 2′-OH group, additional

protection/deprotection strategies are not required unlike the case of RNA where silyl ether protection strategy is typically utilized for the 2′-OH group protection. Furthermore, DNA is less prone to degradation whereas RNA is readily broken down by RNases or acidic/basic conditions.

A typical procedure for the identification of DNA catalysts begins with the synthesis of the DNA pool. This is achieved by solid-phase synthesis where nucleoside phosphoramidites (dA, dG, dC, dT) are mixed in a specific ratio to allow each nucleotide an equal coupling opportunity at every iterative synthetic step. DNA sequences with fixed length are synthesized. Fixed sequence portions flank the random nucleotide region at both the 5′ and 3′ ends. These fixed regions enable downstream PCR amplification. Substrates are covalently attached to the DNA sequences in the pool such that every sequence now has a covalently attached substrate. Metal cation cofactors are provided to enable the catalysis. Catalytically active DNA sequences perform chemical reaction on the covalently linked substrate. A chemoselective technique is required to separate and isolate the covalently attached product from the initial substrate. This chemoselective separation of the product in turn isolates the covalently attached catalytically active DNA sequence. This is typically done by a size-based separation using denaturing polyacrylamide gel electrophoresis (PAGE). PCR is utilized to amplify the catalytically active DNA. After iteratively performing this entire process over several rounds, active DNA sequences are enriched, screened, sequenced and characterized for catalysis.

### 1.3.1 DNA-Templated Synthesis vs. DNA Catalysis

DNA-templated synthesis (DTS) or nucleic acid-templated chemistry is a chemical technology developed by David Lynn and co-workers (further popularized by

David Liu) to achieved controlled synthesis of chemical compounds.[36] Oligonucleotides are connected through chemical linkages to synthetic small-molecule precursors. These synthetic precursors are brought to close spatial proximity after the oligonucleotides are hybridized. Because of high effective molarity, the synthetic precursors react to form the product. The generality of DTS has been used as a basis for evolving non-natural small molecules,[37] as well discovering an additional repertoire of chemical reactions among a wide range of substrates.[36] In addition, DTS has also been used as a display technique to identify enzyme inhibitors.[38,39] Szostak utilized templated-directed nucleic acid copying to provide hypotheses about origin of life.[40]

It is important to note the important difference between true enzymatic rate enhancement and the rate increase achieved by high effective molarity. Classic experiments by Bruice suggest that true catalysis primarily have enthalpic than entropic contributions. NACs are not determined by brining synthetic precursors to proximity. This is more evident in cleavage reactions like phosphodiester, phosphomonoester and amide bond hydrolyses where water is the secondary substrate. In spite of water being in high abundance, 55 M, reaction does not occur without aid from an external catalyst or an enzyme. Even for reactions involving phosphorylation and DNA/RNA polymerization, saturating the reaction solution with high amount of secondary substrates such as NTPs or dNTPs need not necessarily accomplish a templated chemical reaction. Enzymes need to achieve improved reaction rates by providing high enthalpic contributions that DTS does not solve. Additionally, enzymes often perform unfavorable chemistry in the presence of other kinetically or thermodynamically favorable competing pathways. Although DTS increases reaction rate by bringing synthetic precursors to proximity, it is

important to draw the crucial distinction between true enzymatic catalysis by deoxyribozymes and DTS.

### 1.3.2 Nucleic Acid vs. Non-nucleic Acid Substrates

Previous selection processes to identify DNA (or RNA) catalysts were mainly directed towards chemical reactions encompassing nucleic acid substrates. Using such substrates for in vitro selection is a rational and conservative effort. Nucleic acid substrates have the propensity to engage in canonical and non-canonical base pairing with DNA/RNA oligonucleotides. Such interactions include Watson-Crick and wobble base pairs. Non-canonical interaction such as edge-to-edge base pairs are also predicted to be important for structural organization of the RNA/DNA molecules.[41]

The first deoxyribozyme identified in 1994 cleaved a RNA linkage in a full DNA substrate. Following that, the Silverman laboratory has been one of the pioneers in developing DNA catalysts for different chemical reactions.[28,29,42] One of such efforts was directed towards RNA ligation. DNA-catalyzed syntheses of linear and branched RNA have been successfully achieved.[42] These DNA catalysts have been used in a few studies as important chemical biology tools.[43-46] Additionally, DNA catalysts have also been identified for phosphodiester hydrolysis in both DNA and RNA.[47,48] Even though phosphodiester hydrolysis has a half-life ($t_{1/2}$) of approximately 30 million years, [49] and kinetically favorable competing reactions such as RNA transesterification has a half-life of 10 years,[50] DNA is capable of achieving site-specific phosphodiester hydrolysis (Figure 1.4).

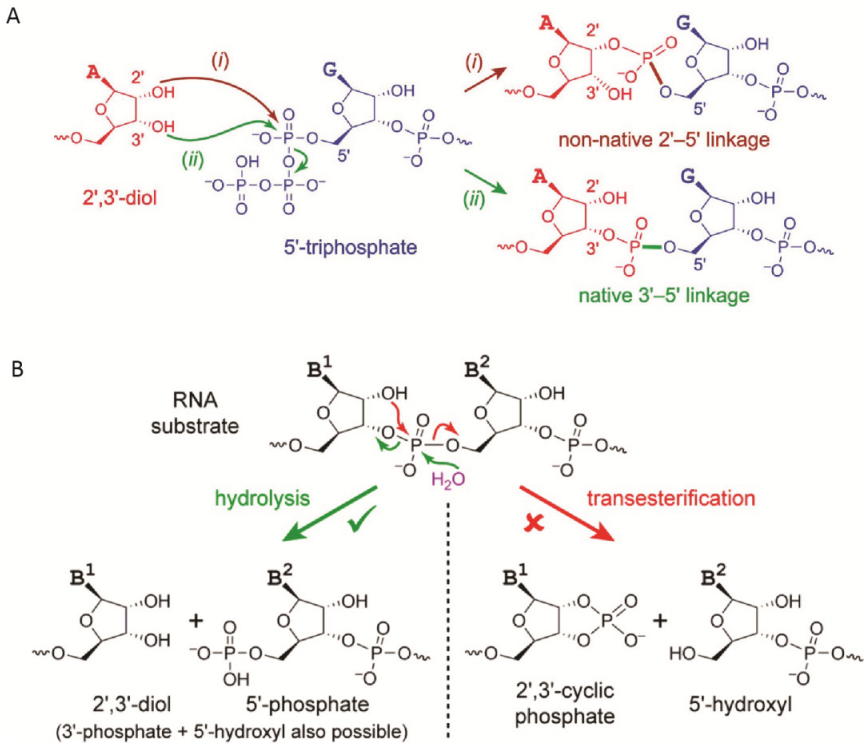


**Figure 1.4.** Reactions catalyzed by DNA involving nucleic acid substrates. A) RNA ligation (figure from reference 29). B) Phosphodiester hydrolysis in RNA (figure from reference 42).

More recently, efforts have been directed to identify DNA catalysts that modify peptides and small molecules (with non-nucleotide character). Extending DNA catalysis to small-molecule substrates with non-nucleotide character can be challenging because of the lack of Watson-Crick base pairing or zero potential to form canonical or non-canonical interactions. However, DNA catalysis has been achieved with Diels-Alder reaction for efficient carbon-carbon bond formation.[51] DNA also has the ability to catalyze covalent modification of protein side chains.[52,53] Chemoselective modifications of peptide or protein side chains have tremendous biological implications that conventional chemical methods do not entirely solve. Fundamental ability of DNA catalysts to modify peptide side chains like tyrosine and serine has been established.[53,54] Hence efforts can be directed to seek DNA catalysts for other protein modifications.

Identifying DNA catalysts that perform cleavage reactions like protease activity, and post-translation modifications (PTMs) such as phosphatase, and kinase activities would be a huge advancement to the chemical and biological research. There is a major impetus to seek deoxyribozymes that function with non-nucleic acid substrates in addition to the conventional nucleic acid substrates.

### 1.3.3 Phosphoryl Transfer Reactions

Phosphoryl groups are constructs made of phosphoric acid. The conjugate base of phosphoric acid is an inorganic phosphate. Phosphoric acid has three acidic protons with $pK_a$ values of 2.1, 7.2 and 12.3. Organophosphates (esters of phosphoric acid) are highly prevalent in biological systems. Nucleic acids (DNA and RNA) are polymers of nucleotides connected by phosphodiester linkages. Phosphates are also prevalent in nature as organic monophosphates, glucose-6-phosphate, isopentenyl diphosphate, and nucleoside triphosphates. Phosphatases are also found on amino acid side chains as post-translational modifications. In several chemical reactions involving phosphoryl transfer, metal cations like magnesium ($Mg^{2+}$) are present to stabilize the negative charge on oxygen and make the phosphorous more electrophilic.[55]

DNA catalysts have shown tremendous promise with phosphoryl transfer reactions. For instance, nucleoside triphosphate, triphosphorylated RNA or nucleoside cyclic phosphate have been used as electrophilic donors for ligation chemistry.[42] Substantial success has been achieved in phosphodiester hydrolysis and transesterification reactions.[29] Furthermore, non-nucleotide substrates such as phosphopeptides can be tagged to RNA oligonucleotides using 5′-triphosphorylated RNA transcript as the electrophilic donors.[56] DNA has had substantial success in catalyzing

reactions involving phosphoryl substrates. This earlier success is suggestive that the chances for identifying DNA catalysts with the ability to perform phosphoryl transfer or phosphorylation-dependent reactions are high, even for non-nucleotide substrates.

## 1.4 Peptide/Protein Post-Translational Modifications

Protein post-translational modifications (PTMs) are covalent modifications on amino acid side chains installed by enzymes following the synthesis of protein by the ribosome. Formally, any of the standard 20 amino acids in a protein that get modified following the ribosomal synthesis are considered to by post-translationally modified amino acids. Some of the common modifications include addition of phosphate, acetate, methyl group, lipid group, monosaccharide and oligosaccharides. PTMs play an important role in biology by modulating the function of several proteins. Some of such important functions include protein folding, stabilization, enzyme-cascade regulation, cellular localization and protein degradation.[57]

### 1.4.1 Kinases and Phosphatases

Protein phosphorylation is one of the most abundant PTMs. Phosphoserine, phosphothreonine and phosphotyrosine are some of the most extensively observed PTMs.[57] Addition or removal of phosphates on peptide side chains are known to regulate cellular pathways, mainly those involved in signal transduction. Most kinases phosphorylate both serine and threonine whereas tyrosine-specific kinases function specifically on tyrosine. Dual-specificity kinases act on all three amino acids.[58]

Conversely, phosphatases remove the phosphate group from the amino acid residue. Protein phosphatases are subdivided into groups based on their substrate specificity or requirement. They can be classified as tyrosine-specific phosphatases,[59]

serine/threonine-specific phosphatases,[60] dual-specificity phosphatases,[61] and histidine phosphatases.[62] Chemically, phosphatase activity is a phosphoryl transfer reaction involving transfer of the phosphate from a peptidic side chain to an external water molecule.

### 1.4.2 Phosphoserine Lyases

Phosphorylated peptidic side chains such as phosphoserine (pSer) or phosphothreonine (pThr) can undergo a β-elimination of the phosphate, leading to formation of dehydroalanine (Dha) and dehydrobutyrine (Dhb) respectively. Dha and Dhb are rare electrophilic sites on peptide/protein side chains. In three specific classes of lantibiotic synthetases, serine/threonine are phosphorylated followed by a lyase activity that involves β-elimination of the phosphorylated residues.[63] Separately, there are examples of pThr lyases that irreversibly inactivate the dual-phosphorylated host MAPKs through β-elimination leading to pathogenic activity.[64] Although pSer/pThr lyase activity is a rare PTM, they are highly important both chemically and biologically.

## 1.5 Thesis Research Focus

This thesis focuses on my efforts to identify DNA catalysts for two important chemical reactions, phosphatase activity and phosphoserine lyase activity. Chapter 2 focuses on catalytic DNA with phosphatase activity. A robust in vitro selection procedure was established to identify phosphatase deoxyribozymes. The starting material is a phosphorylated peptide and the product is the dephosphorylated hydroxyl group. A chemoselective method is required to separate the hydroxyl product, which is covalently linked to the active deoxyribozyme, from the inactive sequences connected to the phosphorylated starting material. This method of separating the active sequences is

formally called a capture method. A know deoxyribozyme forming nucleopeptide linkage with the dephosphorylated product was used as the chemoselective capture method. This capture method allowed for a size-based gel separation. Phosphatase deoxyribozymes were successfully identified. The DNA catalysts increase the reaction rate of phosphomonoester hydrolysis substantially, functioned with free phosphopeptides with multiple turnovers, were robust in the presence of cell lysate and externally added proteins, and functioned with larger protein substrates.

Chapter 3 focuses on three different efforts to identify deoxyribozymes with sequence-selective phosphatase activity. The first effort involved using phosphopeptides with "extreme" (highly functionalized) biophysical/biochemical properties to encourage selectivity. In the second effort, competitive ligands were supplemented in buffer as an external selection process to inspire sequence selectivity. The third approach involved the use of reverse complement oligonucleotides to inhibit the sequence-general DNA catalysts. However, all three efforts failed to provide DNA enzymes with sequence selectivity. Chapter 4 focuses on efforts to deliver DNA enzymes inside live mammalian cells and establish intracellular phosphatase activity. LNCaP cells are known to differentiate to neuroendocrine cells when a known protein-tyrosine phosphatase, PTP1B, is overexpressed.[65] We hope that phosphatase DNA catalysts can serve as a PTP1B mimic.

Chapter 5 focuses on DNA catalysts with pSer lyase activity. In this effort, the substrate was a phosphoserine-containing hexapeptide. The final product Dha, a Michael-like acceptor, was separated by addition of a thiol oligonucleotide to enable a size-based separation on denaturing polyacrylamide gel electrophoresis (PAGE), which was

subsequently enriched by PCR. DNA catalysts were successfully identified and applied in the chemoenzymatic synthesis of a cystathionine-containing cyclic peptide. The chapter also discusses additional efforts to identify sequence-general pSer/pThr lyase deoxyribozymes. Additionally, this chapter focuses on long-term applications of such deoxyribozymes in the synthesis of complex lanthipeptides and utility in phosphoproteomics.

## 1.6 References


(1) Page, M. I.; Jencks, W. P. Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 1678-1683.

(2) Bruice, T. C.; Lightstone, F. C. Ground state and transition state contributions to the rates of intramolecular and enzymatic reactions. *Acc. Chem. Res.* **1999**, *32*, 127-136.

(3) Warshel, A.; Parson, W. W. Dynamics of biochemical and biophysical reactions: insight from computer simulations. *Quart. Rev. Biophys.* **2001**, *34*, 563-679.

(4) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. Electrostatic basis for enzyme catalysis. *Chem. Rev.* **2006**, *106*, 3210-3235.

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

(6) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227-249.

(7) Seeliger, D.; de Groot, B. L. Conformational Transitions upon Ligand Binding: Holo-Structure Prediction from Apo Conformations. *PLoS Comput. Biol.* **2010**, *6*.

(8) Gutierrez, J. A.; Luo, M.; Singh, V.; Li, L.; Brown, R. L.; Norris, G. E.; Evans, G. B.; Furneaux, R. H.; Tyler, P. C.; Painter, G. F.; Lenz, D. H.; Schramm, V. L. Picomolar inhibitors as transition-state probes of 5'-methylthioadenosine nucleosidases. *ACS Chem. Biol.* **2007**, *2*, 725-734.

(9) Holden, H. M.; Tronrud, D. E.; Monzingo, A. F.; Weaver, L. H.; Matthews, B. W. Slow- and fast-binding inhibitors of thermolysin display different modes of binding: crystallographic analysis of extended phosphonamidate transition-state analogues. *Biochemistry* **1987**, *26*, 8542-8553.

(10) Cama, E.; Shin, H.; Christianson, D. W. Design of amino acid sulfonamides as transition-state analogue inhibitors of arginase. *J. Am. Chem. Soc.* **2003**, *125*, 13052-13057.

(11) Wentworth, P.; Janda, K. D. Catalytic antibodies - Structure and function. *Cell Biochem. Biophys.* **2001**, *35*, 63-87.

(12) Gouverneur, V. E.; Houk, K. N.; de Pascual-Teresa, B.; Beno, B.; Janda, K. D.; Lerner, R. A. Control of the exo and endo pathways of the Diels-Alder reaction by antibody catalysis. *Science* **1993**, *262*, 204-208.

(13) Wentworth, P., Jr.; Jones, L. H.; Wentworth, A. D.; Zhu, X.; Larsen, N. A.; Wilson, I. A.; Xu, X.; Goddard, W. A., 3rd; Janda, K. D.; Eschenmoser, A.; Lerner, R. A. Antibody catalysis of the oxidation of water. *Science* **2001**, *293*, 1806-1811.

(14) Jencks, W. P. General acid-base catalysis of complex reactions in water. *Chem. Rev.* **1972**, *72*, 705-718.

(15) Vallee, B. L.; Williams, R. Metalloenzymes: the entatic nature of their active sites. *Proc. Natl. Acad. Sci.* **1968**, *59*, 498-505.

(16) Mense, S. M.; Zhang, L. Heme: a versatile signaling molecule controlling the activities of diverse regulators ranging from transcription factors to MAP kinases. *Cell Res.* **2006**, *16*, 681-692.

(17) Schenk, G.; Elliott, T. W.; Leung, E.; Carrington, L. E.; Mitic, N.; Gahan, L. R.; Guddat, L. W. Crystal structures of a purple acid phosphatase, representing different steps of this enzyme's catalytic cycle. *BMC Struct. Biol.* **2008**, *8*, 6.

(18) Seals, D. F.; Courtneidge, S. A. The ADAMs family of metalloproteases: multidomain proteins with multiple functions. *Genes Dev.* **2003**, *17*, 7-30.

(19) Serganov, A.; Patel, D. J. Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat. Rev. Genet.* **2007**, *8*, 776-790.

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

(21) 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.

(22) Zaher, H. S.; Unrau, P. J. Selection of an improved RNA polymerase ribozyme with superior extension and fidelity. *RNA* **2007**, *13*, 1017-1026.

(23) Wochner, A.; Attwater, J.; Coulson, A.; Holliger, P. Ribozyme-catalyzed transcription of an active ribozyme. *Science* **2011**, *332*, 209-212.

(24) Phylactou, L. A.; Kilpatrick, M. W.; Wood, M. J. Ribozymes as therapeutic tools for genetic disease. *Hum. Mol. Genet.* **1998**, *7*, 1649-1653.

(25) Breaker, R. R. Engineered allosteric ribozymes as biosensor components. *Curr. Opin. Biotechnol.* **2002**, *13*, 31-39.

(26) Khan, A. U. Ribozyme: a clinical tool. *Clin. Chim. Acta* **2006**, *367*, 20-27.

(27) Lieber, A.; He, C. Y.; Polyak, S. J.; Gretch, D. R.; Barr, D.; Kay, M. A. Elimination of hepatitis C virus RNA in infected human hepatocytes by adenovirus-mediated expression of ribozymes. *J. Virol.* **1996**, *70*, 8782-8791.

(28) Silverman, S. K. DNA as a Versatile Chemical Component for Catalysis, Encoding, and Stereocontrol. *Angew. Chem. Int. Ed.* **2010**, *49*, 7180-7201.

(29) 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.

(30) Krishnamurthy, R. Role of pK(a) of Nucleobases in the Origins of Chemical Evolution. *Acc. Chem. Res.* **2012**, *45*, 2035-2044.

(31) Velikyan, I.; Acharya, S.; Trifonova, A.; Foldesi, A.; Chattopadhyaya, J. The pK(a)'s of 2'-hydroxyl group in nucleosides and nucleotides. *J. Am. Chem. Soc.* **2001**, *123*, 2893-2894.

(32) Pyle, A. M. Ribozymes: a distinct class of metalloenzymes. *Science* **1993**, *261*, 709-714.

(33) Suslov, N. B.; DasGupta, S.; Huang, H.; Fuller, J. R.; Lilley, D. M.; Rice, P. A.; Piccirilli, J. A. Crystal structure of the Varkud satellite ribozyme. *Nat. Chem. Biol.* **2015**, *11*, 840-846.

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

(35) Kath-Schorr, S.; Wilson, T. J.; Li, N. S.; Lu, J.; Piccirilli, J. A.; Lilley, D. M. General acid-base catalysis mediated by nucleobases in the hairpin ribozyme. *J. Am. Chem. Soc.* **2012**, *134*, 16717-16724.

(36) Kanan, M. W.; Rozenman, M. M.; Sakurai, K.; Snyder, T. M.; Liu, D. R. Reaction discovery enabled by DNA-templated synthesis and in vitro selection. *Nature* **2004**, *431*, 545-549.

(37) Gartner, Z. J.; Liu, D. R. The generality of DNA-templated synthesis as a basis for evolving non-natural small molecules. *J. Am. Chem. Soc.* **2001**, *123*, 6961-6963.

(38) Kleiner, R. E.; Dumelin, C. E.; Tiu, G. C.; Sakurai, K.; Liu, D. R. In vitro selection of a DNA-templated small-molecule library reveals a class of macrocyclic kinase inhibitors. *J. Am. Chem. Soc.* **2010**, *132*, 11779-11791.

(39) Georghiou, G.; Kleiner, R. E.; Pulkoski-Gross, M.; Liu, D. R.; Seeliger, M. A. Highly specific, bisubstrate-competitive Src inhibitors from DNA-templated macrocycles. *Nat. Chem. Biol.* **2012**, *8*, 366-374.

(40) Schrum, J. P.; Ricardo, A.; Krishnamurthy, M.; Blain, J. C.; Szostak, J. W. Efficient and Rapid Template-Directed Nucleic Acid Copying Using 2 '-Amino-2 ',3 '-dideoxyribonucleoside-5 '-Phosphorimidazolide Monomers. *J. Am. Chem. Soc.* **2009**, *131*, 14560-14570.

(41) Halder, S.; Bhattacharyya, D. RNA structure and dynamics: a base pairing perspective. *Prog. Biophys. Mol. Biol.* **2013**, *113*, 264-283.

(42) Silverman, S. K. Deoxyribozymes: selection design and serendipity in the development of DNA catalysts. *Acc. Chem. Res.* **2009**, *42*, 1521-1531.

(43) Furukawa, K.; Abe, H.; Tamura, Y.; Yoshimoto, R.; Yoshida, M.; Tsuneda, S.; Ito, Y. Fluorescence detection of intron lariat RNA with reduction-triggered fluorescent probes. *Angew. Chem. Int. Ed.* **2011**, 12020-12023.

(44) Buttner, L.; Seikowski, J.; Wawrzyniak, K.; Ochmann, A.; Hobartner, C. Synthesis of spin-labeled riboswitch RNAs using convertible nucleosides and DNA-catalyzed RNA ligation. *Bioorg. Med. Chem.* **2013**, *21*, 6171-6180.

(45) Buttner, L.; Javadi-Zarnaghi, F.; Hobartner, C. Site-Specific Labeling of RNA at Internal Ribose Hydroxyl Groups: Terbium-Assisted Deoxyribozymes at Work. *J. Am. Chem. Soc.* **2014**, *136*, 8131-8137.

(46) Wawrzyniak-Turek, K.; Hobartner, C. Deoxyribozyme-mediated ligation for incorporating EPR spin labels and reporter groups into RNA. *Methods Enzymol.* **2014**, *549*, 85-104.

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

(48) 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.

(49) Schroeder, G. K.; Lad, C.; Wyman, P.; Williams, N. H.; Wolfenden, R. The time required for water attack at the phosphorus atom of simple phosphodiesters and of DNA. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4052-4055.

(50) Li, Y. F.; Breaker, R. R. Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2 '-hydroxyl group. *J. Am. Chem. Soc.* **1999**, *121*, 5364-5372.

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

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

(53) 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.

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

(55) Matte, A.; Tari, L. W.; Delbaere, L. T. J. How do kinases transfer phosphoryl groups? *Struct. Fold. Des.* **1998**, *6*, 413-419.

(56) Sachdeva, A.; Chandra, M.; Chandrasekar, J.; Silverman, S. K. Covalent Tagging of Phosphorylated Peptides by Phosphate-Specific Deoxyribozymes. *ChemBioChem* **2012**, *13*, 654-657.

(57) Prabakaran, S.; Lippens, G.; Steen, H.; Gunawardena, J. Post-translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2012**, *4*, 565-583.

(58) Dhanasekaran, N.; Premkumar Reddy, E. Signaling by dual specificity kinases. *Oncogene* **1998**, *17*, 1447-1455.

(59) Zhang, Z. Y. Protein tyrosine phosphatases: structure and function, substrate specificity, and inhibitor development. *Annu. Rev. Pharmacol. Toxicol.* **2002**, *42*, 209-234.

(60) Mumby, M. C.; Walter, G. Protein serine/threonine phosphatases: structure, regulation, and functions in cell growth. *Physiol. Rev.* **1993**, *73*, 673-699.

(61) Camps, M.; Nichols, A.; Arkinstall, S. Dual specificity phosphatases: a gene family for control of MAP kinase function. *FASEB J.* **2000**, *14*, 6-16.

(62) Baumer, N.; Maurer, A.; Krieglstein, J.; Klumpp, S. Expression of protein histidine phosphatase in *Escherichia coli*, purification, and determination of enzyme activity. *Methods Mol. Biol.* **2007**, *365*, 247-260.

(63) Chatterjee, C.; Miller, L. M.; Leung, Y. L.; Xie, L.; Yi, M.; Kelleher, N. L.; van der Donk, W. A. Lacticin 481 synthetase phosphorylates its substrate during lantibiotic production. *J. Am. Chem. Soc.* **2005**, *127*, 15332-15333.

(64) Zhu, Y. Q.; Li, H. T.; Long, C. Z.; Hu, L. Y.; Xu, H.; Liu, L. P.; Chen, S.; Wang, D. C.; Shao, F. Structural insights into the enzymatic mechanism of the pathogenic MAPK phosphothreonine lyase. *Mol. Cell* **2007**, *28*, 899-913.

(65) Wu, C.; Zhang, L.; Bourne, P. A.; Reeder, J. E.; di Sant'Agnese, P. A.; Yao, J. L.; Na, Y.; Huang, J. Protein tyrosine phosphatase PTP1B is involved in neuroendocrine differentiation of prostate cancer. *Prostate* **2006**, *66*, 1125-1135.

# Chapter 2: Catalytic DNA with Phosphatase Activity[1]

## 2.1 Introduction

In this chapter, catalytic DNA with phosphatase activity is discussed. Phosphatase reaction is hydrolysis of the phosphomonoester bond by an external water molecule to release the hydroxyl group and an inorganic phosphate. Phosphatase activity is discussed in the context of two phosphorylated amino acid side chains, phosphotyrosine ($Y^P$) and phosphoserine ($S^P$). Development of catalysts is a major impetus for much of modern chemical research. Nature's biomolecular protein and RNA catalysts are responsible for a wide range of chemical reactions, and protein enzymes in particular can achieve large rate enhancements.[1] Although DNA catalysts are unknown in nature, in vitro selection (first pioneered for RNA)[2] is readily applied to identify catalytically active artificial DNA sequences.[3-4] Importantly, DNA (and RNA) catalysts can be identified by starting with entirely random sequence pools, whereas directed evolution of proteins typically requires a known, catalytically active starting point.[5-6] A growing range of chemical reactions has been shown to be catalyzed by DNA.[7] For DNA phosphodiester hydrolysis, the uncatalyzed (spontaneous) half-life for P–O bond cleavage of ~30 million years is reduced to as little as 0.5 min by a DNA catalyst.[8,9] However, in this reaction the DNA catalyst interacts with its DNA substrate by extensive Watson-Crick base pairing, and such an approach cannot be generalized to non-oligonucleotide substrates such as peptides and proteins. With the exception of the ribosome, the natural ribozymes catalyze

[1] This research has been published: Chandrasekar, J.; Silverman, S. K. Catalytic DNA with phosphatase activity. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5315-5320.

RNA cleavage and ligation, and they generally have more modest rate enhancements,[10-12] limited in part by the relatively high uncatalyzed half-lives of these reactions.

### 2.1.1 Mechanism and Role of Protein Phosphatases

Natural protein phosphatases[13-14] function either by a double-displacement mechanism (cysteine-dependent phosphatases[15-16] and haloacid dehalogenase (HAD) superfamily aspartate-dependent phosphatases[17-19]) or via direct metal-bound water attack using cofactors such as $Mn^{2+}$ and $Fe^{2+}$ (Figure 2.1).[20] Both phosphomonoester dianion[21] and monoanion[22-23] substrates have been implicated.[24]

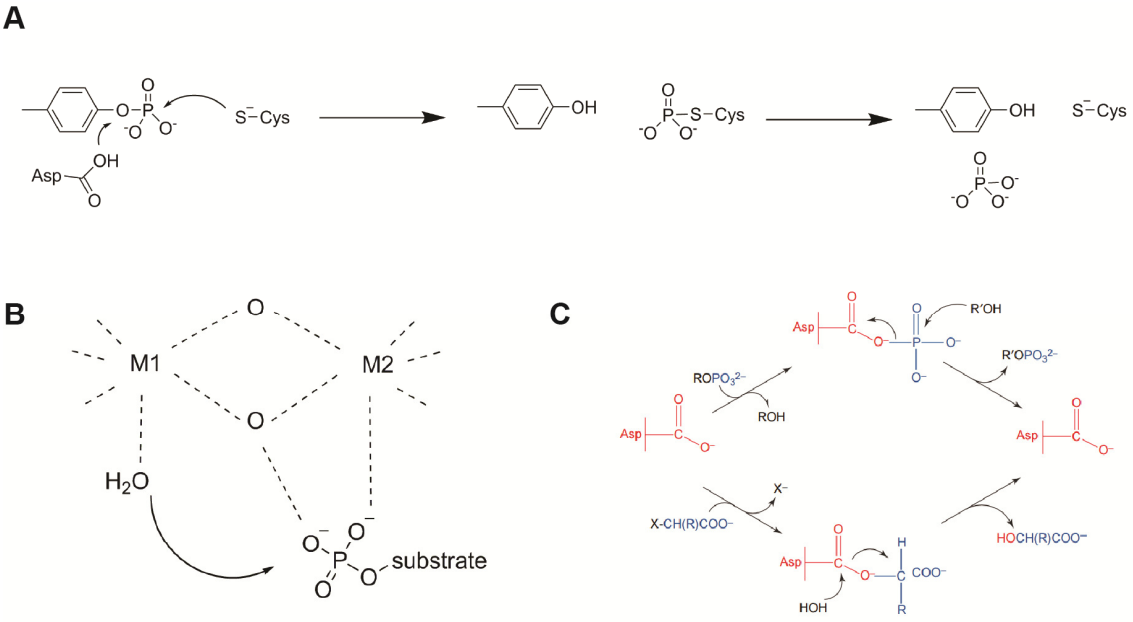


**Figure 2.1.** Catalytic mechanism of natural protein phosphatases. A) Protein-tyrosine phosphatase PTP1B. B) Direct metal-bound water attack using cofactors such as $Mn^{2+}$ and $Fe^{2+}$. C) Haloacid dehalogenase (HAD) superfamily aspartate-dependent phosphatases (figure from reference 17).

### 2.1.2 Capture Strategy to Identify DNA Catalysts

To identify DNA catalysts with (peptide/protein) phosphatase activity, we need to achieve chemoselective isolation of the product (dephosphorylated tyrosine or serine) from the initially phosphorylated starting material. As discussed in chapter 1, this separation of product (attached to active DNA) from the starting material (attached to the

inactive DNA pool) is formally called a capture reaction. An ideal capture reaction would involve a clean size-based separation of the desirable components by denaturing polyacrylamide gel electrophoresis (PAGE).

The capture strategy has to be amenable to a typical in vitro selection pipeline. To enable size-based PAGE separation of the product, it is necessary to attach a secondary tag on tyrosine-OH or serine-OH product, but not $Y^P$ or $S^P$. Few chemical or enzymatic options exist for such chemoselective separation of Y from $Y^P$ (Figure 2.2). A) Tyrosine alkylation using π-allylpalladium complexes.[25] B) Tyrosine conjugation through an aqueous ene-type reaction.[26] C) Deoxyribozyme that forms a nucleopeptide linkage of tyrosine with a 5′-triphosphorylated RNA.[3] The enzymatic technique using a DNA catalyst forming nucleopeptide linkage with tyrosine-OH group is the most promising approach amenable to a typical in vitro selection procedure.

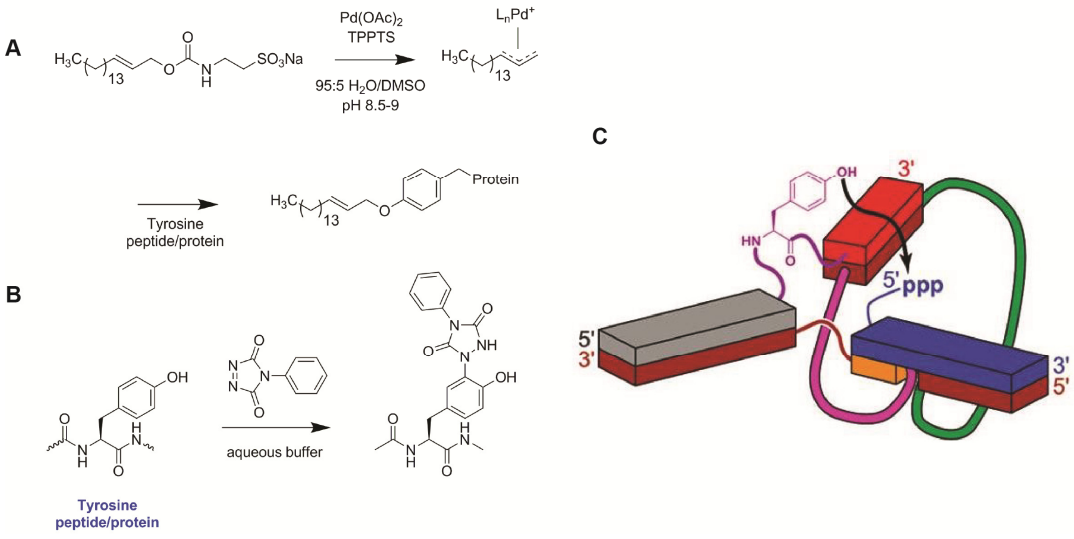


**Figure 2.2.** Chemical and enzymatic techniques to attach chemoselective tag to tyrosine ($Y^{OH}$) and not phosphotyrosine ($Y^{P}$). A) Pd-based technique developed by Francis and co-workers.[25] B) Aqueous-ene method developed by Barbas and co-workers.[26] C) DNA catalyst utilized for tyrosine nucleopeptide conjugation developed by the Silverman laboratory (figure from reference 3).

### 2.1.3 Capture Strategy for Phosphoserine Modification

Analogous chemoselective conjugation techniques are required to isolate Serine from $S^P$. Chemical oxidation of N-terminal serine to an aldehyde using periodate oxidation followed by capture of the resulting electrophilic aldehyde is a well-known bio-conjugation technique.[27] DNA catalysts have also been identified to perform nucleopeptide linkage with an internal serine-OH embedded in a tripeptide (Figure 2.3).[7] However, it should be noted that phosphatase DNA catalysts identified by N-terminal serine oxidation may have limited utility. Phosphatase deoxyribozymes identified from such selection may be programmed to function merely with N-terminal phosphoserine instead of an internal phosphoserine. DNA catalysts with nucleopeptide-linkage forming ability with an internal serine-OH group would be the ideal choice for serine capture.

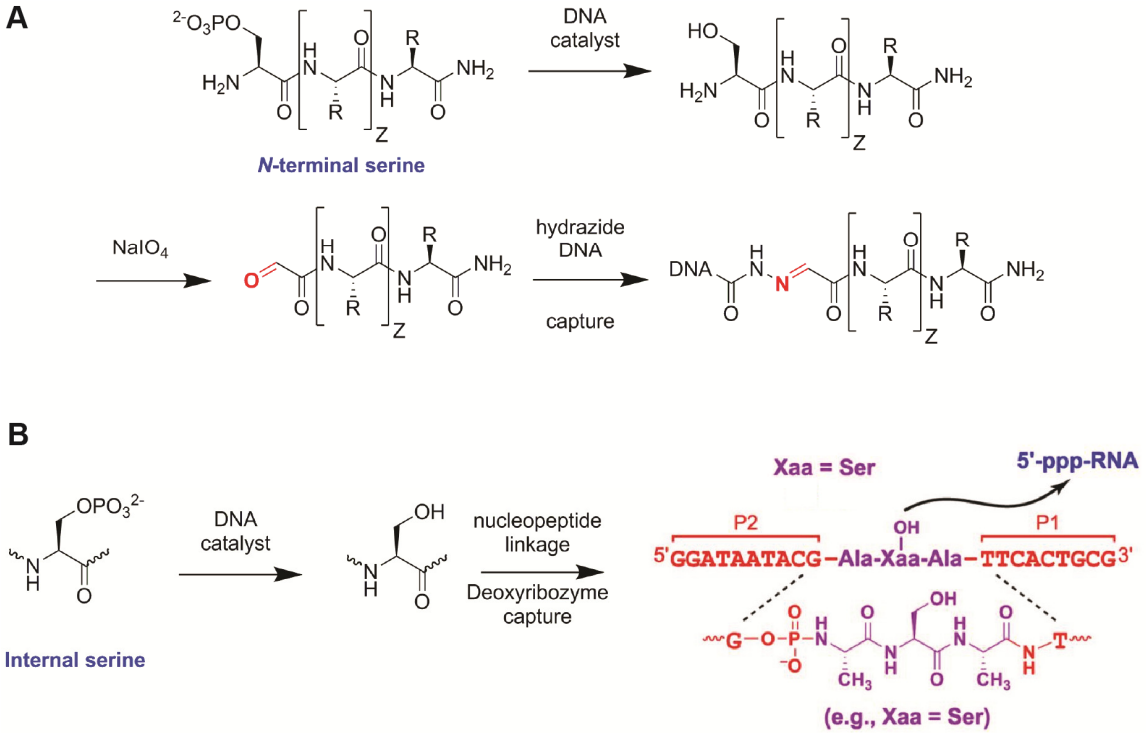


**Figure 2.3.** Techniques to capture dephosphorylated serine $S^{OH}$. A) Capture step involves oxidation of the N-terminal serine using sodium periodate. The resulting aldehyde is capture as a hydrazone using a hydrazide DNA. B) DNA catalyst is used to perform nucleopeptide linkage on the dephosphorylated serine.

## 2.2 Results

DNA catalysts that covalently modify peptide and protein substrates are fundamentally interesting and have likely practical value, especially for biologically relevant chemical modifications. We have initiated studies into DNA-catalyzed modifications of amino acid side chains of peptide substrates, such as nucleopeptide linkage formation involving tyrosine and serine. [7, 28-29] A major challenge in such studies is to achieve catalysis even though the DNA catalyst cannot engage in any preprogrammable Watson-Crick binding interactions with the substrate. In this chapter, it is shown that DNA can catalyze $Zn^{2+}$-dependent hydrolysis of tyrosine and serine phosphomonoesters of peptide substrates, i.e., have phosphatase activity. The half-life for spontaneous phosphoserine hydrolysis at ~37 °C is approximately $4 \times 10^{10}$ years as estimated on the basis of the half-life for methyl phosphate dianion,[13] which emphasizes the difficulty of catalyzing this reaction.

### 2.2.1 Design for Serine and Tyrosine Capture

The selection strategy requires a capture enzyme to separate the active DNA sequences from the inactive ones. Several DNA enzymes are capable of forming nucleopeptide linkage with tyrosine or serine and 5′-triphosphorylated RNA.[29] These deoxyribozymes were tested for the "capture" purpose. Hexapeptide substrates $AAAY^PAA$ and $AAAS^PAA$ that incorporate internal phosphoserine ($S^P$) or phosphotyrosine ($Y^P$) were appended to a DNA anchor oligonucleotide (as discussed in Materials and Methods).

Among deoxyribozymes 15MZ30, 15MZ36, 15MZ49, 9NG14, 10KC3, and 11MN5, 15MZ36 was determined to have the best capture yield with the DNA-anchored

hexapeptide AAAY$^{P}$AA (~65-70%). However, for the serine-containing peptide AAAS$^{P}$AA, the capture yield was relatively weak (~10%). The yield was substantially low to be useful as a capture method for serine phosphatase selections. Hence, reselections were performed to identify capture deoxyribozymes for AAAS$^{P}$AA.

For selections involving AAAY$^{P}$AA, we aimed to utilize $Mg^{2+}$, $Mn^{2+}$, and $Zn^{2+}$ (40 mM, 20 mM and 1 mM respectively) as the metal cation co-factors. Previous efforts with DNA hydrolysis selections strongly suggested that $Zn^{2+}$ plays an important role in phosphodiester hydrolysis.[8,9] We hypothesized that $Zn^{2+}$ would most likely be an important co-factor in phosphomonoester hydrolysis (phosphatase activity). When the capture reaction with 15MZ36 was tested in the presence of $Zn^{2+}$, the nucleopeptide linkage yield dropped from ~70% to 35%. $Zn^{2+}$ was determined to be a poisoning cofactor for the capture enzyme 15MZ36. Using 15MZ36 directly after the selection step (containing 1 mM $Zn^{2+}$) would lead to a sub-optimal capture yield (even if phosphatase activity were to require $Zn^{2+}$). To address this situation, all divalent metal cations (Mg/Mn/Zn) were removed by ethanol precipitation before the capture step. This was immediately followed by addition of 15MZ36, 5′-triphosphorylated RNA, and fresh $Mg^{2+}$/$Mn^{2+}$ co-factors for the capture reaction. However, this procedure is time consuming with the possibility of precipitation errors. Noting that $Zn^{2+}$ has the highest affinity to EDTA in comparison to either $Mg^{2+}$ or $Mn^{2+}$, 2 equivalents of EDTA were added to selectively chelate the 1mM $Zn^{2+}$. Without any precipitation, the 15MZ36 enzyme and the 5′-triphosphorylated RNA were added to the reaction mixture (see Materials and Methods for capture details). The capture yield for this chelation procedure was similar to the precipitation technique (Figure 2.4). Because the selective chelation

was a more time efficient procedure, we used this modified capture method for the in vitro selection procedure.

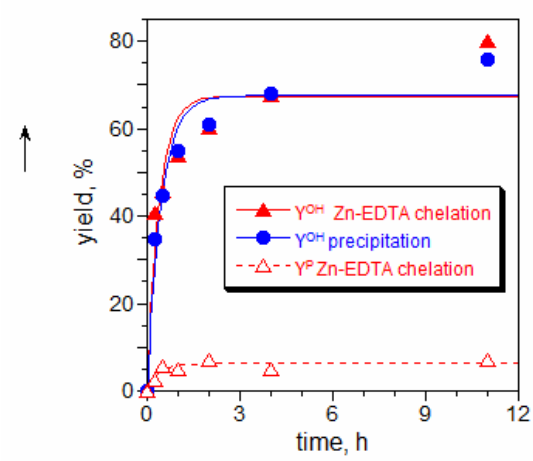


**Figure 2.4.** Capture yield of nucleopeptide linkage with a 5′-triphosphorylated RNA using 15MZ36. A) Blue filled circle, AAAY$^{OH}$AA capture after precipitation and adding fresh Mg/Mn. B) Red filled triangle, capture after selective chelation of $Zn^{2+}$ using 2 eq. of EDTA. C) Red open triangle, AAAY$^{P}$AA capture after selective chelation with 15MZ36.

### 2.2.2 In Vitro Selection Process for Phosphotyrosine

Our strategy to identify phosphatase deoxyribozymes by in vitro selection has key steps depicted in Figure 2.5. A hexapeptide substrate AAAY$^{P}$AA that incorporates an internal phosphotyrosine (Y$^{P}$) residue was covalently appended to a DNA anchor oligonucleotide (Figure 2.6), which was itself ligated to the deoxyribozyme pool. After each key selection step that used incubation conditions of 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, active DNA catalyst sequences were "captured" with the aid of the 15MZ36 deoxyribozyme that selectively uses dephosphorylated tyrosine (Y$^{OH}$) as the nucleophile to attack 5′-triphosphorylated RNA, resulting in a PAGE shift for the corresponding deoxyribozyme sequence (Figure 2.7). [30] The active DNA sequences were amplified by PCR and joined by T4 DNA ligase to the DNA-anchored hexapeptide substrate for entering the next

selection round. After 14 rounds of selection, the phosphatase activity of the DNA pool reached 35% (Figure 2.8), and individual deoxyribozymes were cloned and characterized.

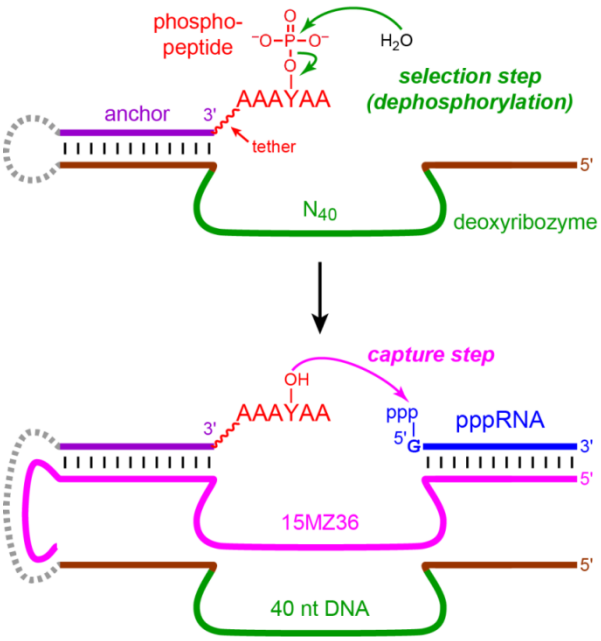


**Figure 2.5.** In vitro selection strategy to identify DNA catalysts with phosphatase activity. Illustrated are the key selection and capture steps of each round. Each final deoxyribozyme has 40 particular nucleotides in its originally random N40 region. The covalent loop on the left side (dashed) was required for the selection process but was dispensable for catalysis by individual deoxyribozymes. The loop was absent for all assays of individual deoxyribozymes.

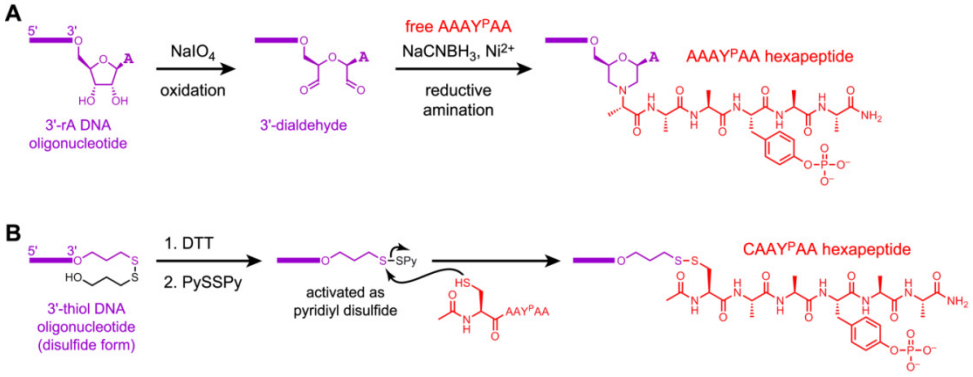


**Figure 2.6.** Synthesis of DNA-anchored hexapeptides. A) Synthesis by attachment of the DNA anchor to the N-terminal α-amino group, with the linkage created by reductive amination. B) Synthesis by attachment of the DNA anchor to the N-terminal cysteine side chain, with the linkage created by disulfide formation.

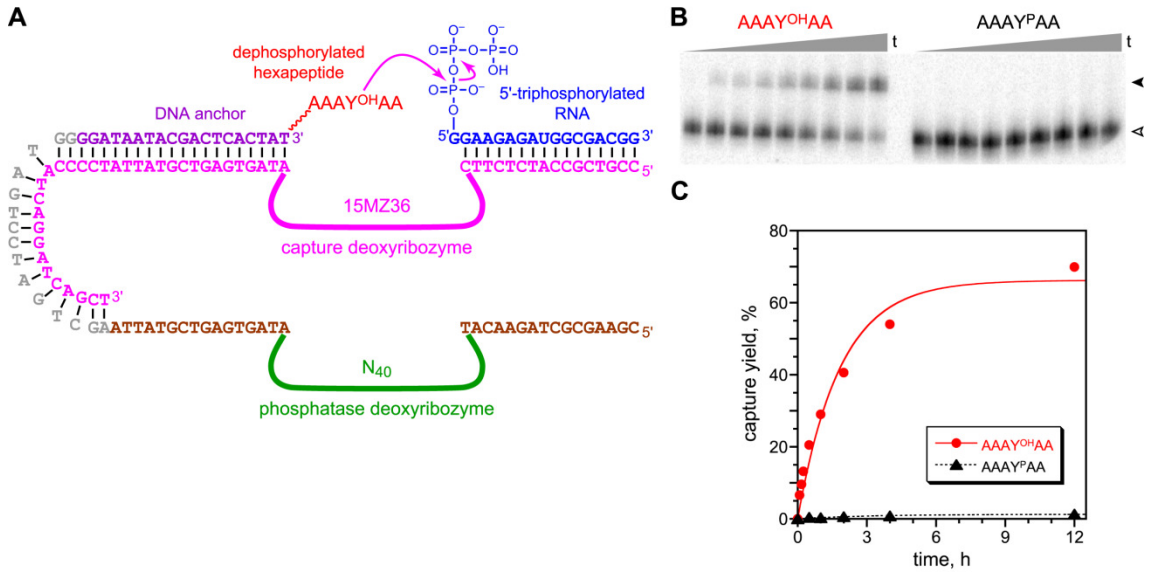


**Figure 2.7.** Validating use of the 15MZ36 deoxyribozyme for capture of catalytically active DNA sequences that have dephosphorylated the $Y^P$ residue of their covalently attached $AAAY^PAA$ hexapeptide substrate. A) Illustration of the 15MZ36-catalyzed capture reaction to attach the 5′-triphosphorylated RNA to the dephosphorylated $AAAY^{OH}AA$ substrate. Note that the substrate molecule is physically connected to the deoxyribozyme sequence that catalyzed its dephosphorylation (denoted as $N_{40}$). The 15MZ36 catalytic region (41 nt) is 5′-CAAGGAGAGTTGTACAAGCTCGGGTCGTGTTCAAAGGGATC-3′.[29] B) PAGE image, illustrating selectivity for $AAAY^{OH}AA$ over $AAAY^PAA$ of the 15MZ36-catalyzed capture reaction. $t$ = 30 s, 5 min, 10 min, 15 min, 30 min, 1 h, 2 h, 4 h, 12 h. The open arrowhead marks the reactant, and the filled arrowhead marks the product. Incubation conditions: 50 mM HEPES, pH 7.5, 20 mM $MnCl_2$, 40 mM $MgCl_2$, 100 mM NaCl, 37 °C. C) Kinetic plot to quantify the PAGE image.

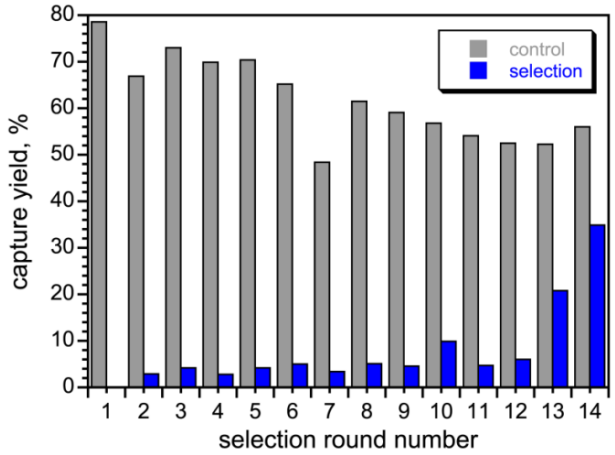


**Figure 2.8.** Progression of the in vitro selection experiment. In each round, "control" refers to the yield for the 15MZ36-catalyzed capture reaction using the DNA-anchored dephosphorylated $AAAY^{OH}AA$ substrate, and "selection" refers to the yield for the 15MZ36-catalyzed capture reaction using the deoxyribozyme pool for that round. Round 1 pool yield was not measured because the pool was not radiolabeled.

Five distinct deoxyribozyme sequences were identified (Figure 2.9). Of these five DNA catalysts, one — named 14WM9 — had substantial phosphatase activity not only

with the initially used $Y^P$ hexapeptide that is connected rather closely to the DNA anchor oligonucleotide, but also with a $Y^P$ hexapeptide that is attached to the DNA anchor through a PEG-type tether that includes 60 ethylene glycol units. The distinct 14WM27 deoxyribozyme also showed a small amount of activity with the PEG-linked $Y^P$ hexapeptide. We focused our attention primarily on 14WM9, because its substantial activity with the PEG-tethered hexapeptide suggested the possibility of catalysis with entirely free peptides as well.

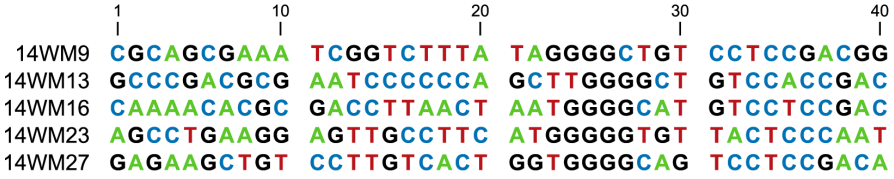


**Figure 2.9.** Sequences of the new phosphatase deoxyribozymes. Only the enzyme region ($N_{40}$) sequences are shown. For single-turnover intramolecular (in *cis*) reactions, each enzyme region was surrounded by additional DNA nucleotides as shown in Figure 2.7A. For multiple-turnover intermolecular (in *trans*) reactions, each enzyme region was surrounded by the two single-stranded binding arms (each 16 nt) as shown in Figure 2.7A.

### 2.2.3 Single-Turnover Dephosphorylation by the 14WM9 Deoxyribozyme

We found that 14WM9 catalyzes $Zn^{2+}$-dependent single-turnover dephosphorylation of both phosphotyrosine ($Y^P$) and phosphoserine ($S^P$) in DNA-anchored hexapeptide substrates (Figure 2.10). The substrates were the same as those used during the selection process, but now the DNA-anchored hexapeptide was not covalently attached to the deoxyribozyme (the dashed loop on the far left side of Figure 2.5 was absent). The identities of the phosphorylated and dephosphorylated DNA-anchored hexapeptides (i.e., substrates and products) were verified by MALDI mass spectrometry analysis directly from the DNA-catalyzed reaction samples (all $[M–H]^-$: $Y^P$ substrate *m*/*z* calcd. 6705.6, found 6708.3, $\Delta = +0.04\%$; $Y^{OH}$ product *m*/*z* calcd. 6625.6, found 6628.0, $\Delta = +0.04\%$; $S^P$ substrate *m*/*z* calcd. 6629.5, found 6633.6, $\Delta = +0.06\%$;

$S^{OH}$ product *m*/*z* calcd. 6549.5, found 6554.3, Δ = +0.07%). Mass spectrometry was also used to verify the identities of separately synthesized standards (Table 2.1).

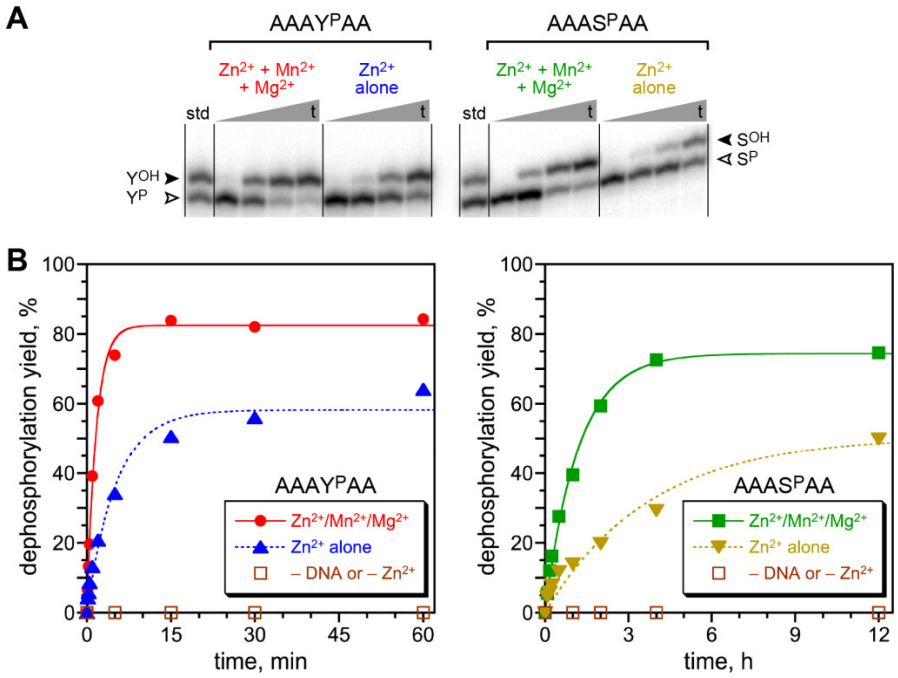


**Figure 2.10.** Single-turnover assays of the 14WM9 deoxyribozyme that dephosphorylates phosphotyrosine and phosphoserine side chains. A) PAGE assays of the 14WM9-catalyzed reactions, using DNA-anchored hexapeptide substrates and showing representative timepoints. The open arrowhead marks the phosphopeptide substrate; the filled arrowhead marks the dephosphorylated product. Each hexapeptide substrate was joined via its N-terminal Ala residue to the 3′-end of the DNA anchor by reductive amination. Assay conditions: 1 μM 14WM9, 12.5 nM 5′-$^{32}$P-radiolabeled DNA-anchored hexapeptide, 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. In the assays with $Zn^{2+}$ alone, the $Mn^{2+}$ and $Mg^{2+}$ were omitted. Timepoints for $Y^P$ substrate: 10 s, 1 min, 5 min, 1 h. Timepoints for $S^P$ substrate: 30 s, 15 min, 2 h, 12 h. B) Kinetic plots for the assays. The negative controls without deoxyribozyme or without $Zn^{2+}$ all showed no detectable dephosphorylation activity (<0.5%).

**Table 2.1.** MALDI mass spectrometry data for DNA-anchored peptide standards. The DNA was either 5′-GGATAATACGACTCACTAT−rA−3′ (for peptides with N-terminal Ala anchored by reductive amination; the rA was oxidized by $NaIO_4$) or 5′-GGATAATACGACTCACTAT-$C_3$-SS-$C_3$-OH-3′ (for peptides with N-terminal Cys anchored by disulfide bond formation; the disulfide linker was reduced by DTT). The HEG linker was hexa(ethylene glycol) as illustrated in Figure 4. The PEG linker was 10 consecutive HEG linkers.

| DNA-anchored peptide | mass calcd. | mass found | error, % (found – calcd.) |
|---|---|---|---|
| DNA-AAAY$^{P}$AA | 6705.6 | 6703.7 [a] | –0.03 |
| DNA-AAAY$^{OH}$AA | 6625.6 | 6625.3 [a] | –0.01 |
| DNA-AAAS$^{P}$AA | 6629.5 | 6627.2 [a] | –0.03 |
| DNA-AAAS$^{OH}$AA | 6549.5 | 6547.6 [a] | –0.03 |
| DNA-$C_3$-CAAY$^{P}$AA | 6636.7 | 6637.2 | +0.01 |
| DNA-HEG-CAAY$^{P}$AA | 6980.9 | 6977.9 | –0.04 |
| DNA-HEG-CAAY$^{OH}$AA | 6901.0 | 6902.5 | +0.02 |
| DNA-HEG-CAAS$^{P}$AA | 6904.8 | 6897.5 | –0.11 |
| DNA-HEG-CAAS$^{OH}$AA | 6824.9 | 6825.5 | +0.01 |
| DNA-HEG-CY$^{P}$A | 6767.7 | 6771.3 | +0.05 |
| DNA-HEG-CS$^{P}$A | 6691.6 | 6693.8 | +0.03 |
| DNA-HEG-CY$^{P}$ | 6696.6 | 6697.5 | +0.01 |
| DNA-HEG-CS$^{P}$ | 6620.5 | 6618.9 | –0.02 |
| DNA-HEG-CAFY$^{P}$AA | 7057.0 | 7059.9 | +0.04 |
| DNA-HEG-CAEY$^{P}$AA | 7039.0 | 7040.6 | +0.02 |
| DNA-HEG-CAKY$^{P}$AA | 7038.0 | 7039.2 | +0.02 |
| DNA-HEG-CAAY$^{P}$FA | 7057.0 | 7057.7 | +0.01 |
| DNA-HEG-CAAY$^{P}$EA | 7039.0 | 7039.8 | +0.01 |
| DNA-HEG-CAAY$^{P}$KA | 7038.0 | 7038.5 | +0.01 |
| DNA-PEG-CAAY$^{P}$AA | 10079.8 | 10080.4 | +0.01 |
| DNA-$C_3$-S$^{P}$AAAAC | 6518.6 | 6512.3 | -0.10 |

[a] The found values for these DNA-anchored peptides were observed by MALDI MS analysis directly from the DNA-catalyzed reaction samples, rather than by using separately synthesized standard samples as tabulated here.

Substantial catalytic activity of 14WM9 with the $Y^P$ substrate was observed not only using the initial selection conditions that included each of 1 mM $Zn^{2+}$, 20 mM $Mn^{2+}$, and 40 mM $Mg^{2+}$ at pH 7.5 and 37 °C (single-turnover $k_{obs} = 0.65 \pm 0.05$ min$^{-1}$; $n = 3$), but also with 1 mM $Zn^{2+}$ alone, albeit with 3-fold reduction in rate constant ($k_{obs} = 0.19 \pm 0.03$ min$^{-1}$; $n = 3$). Activity with the $S^P$ substrate was observed with $Zn^{2+}$ in the presence or absence of $Mn^{2+}$ and $Mg^{2+}$, again with about 3-fold difference in rate constant ($k_{obs} = 0.79 \pm 0.15$ h$^{-1}$ and $0.31 \pm 0.06$ h$^{-1}$, respectively; $n = 3$). For both $Y^P$ and

$S^P$ substrates, no activity was observed in the absence of either the deoxyribozyme or $Zn^{2+}$, and 1 mM $Zn^{2+}$ was optimal; either decreasing or increasing the $Zn^{2+}$ concentration reduced the activity. Using the $Y^P$ substrate ($S^P$ was not tested in this regard), truncating the constant 5′-portion of 14WM9 (16 nt 5′-segment in Figure 2.5) did not substantially disrupt catalytic function. The rate constant did not decrease ($k_{rel} = 1.6$ for 5′-truncated versus full-length deoxyribozyme), and the final yield was essentially unchanged. However, using a random DNA sequence in place of the 40 nt catalytic region — functionally equivalent to performing the first selection round with the $N_{40}$ pool — abolished all catalytic activity, demonstrating that the mere presence of nonspecific DNA and $Zn^{2+}$ is insufficient to support dephosphorylation.

### 2.2.4 Multiple-Turnover Dephosphorylation by the 14WM9 Deoxyribozyme

Although our in vitro selection strategy inherently requires a covalent tether between the peptide and the DNA anchor oligonucleotide (Figure 2.5), we found that 14WM9 also catalyzes efficient multiple-turnover $Zn^{2+}$-dependent dephosphorylation of the entirely free (untethered) $AAAY^PAA$ peptide substrate, as assayed both by HPLC of the peptide mixture (Figure 2.11A) and by colorimetric detection of the released inorganic phosphate, $P_i$, using malachite green dye (Figure 2.11B).

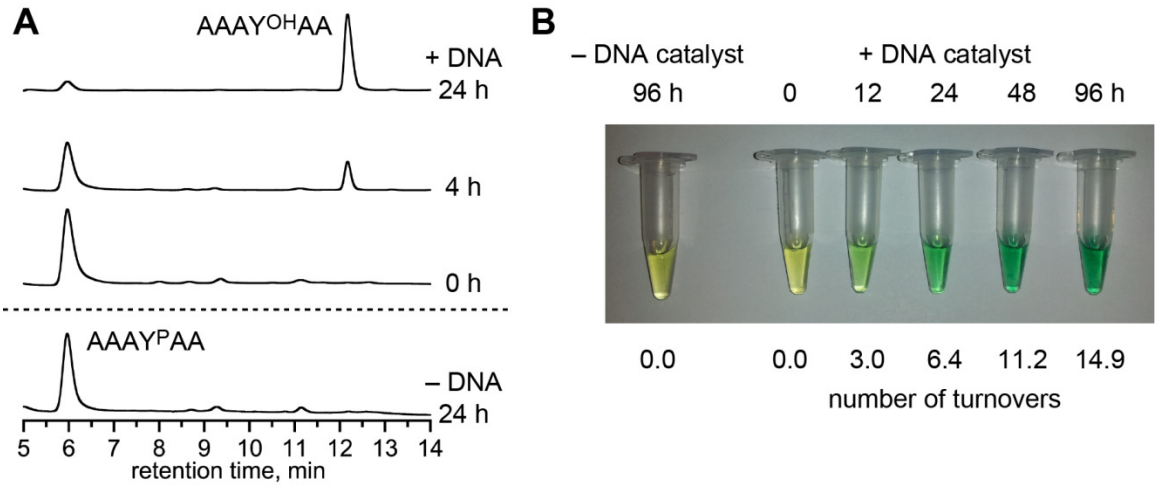


**Figure 2.11.** Multiple-turnover assays of the 14WM9 deoxyribozyme with free phosphopeptide substrate. A) HPLC assay. Conditions: 100 μM 14WM9, 125 μM DNA anchor oligonucleotide, 500 μM free AAAY$^{P}$AA, 70 mM HEPES, pH 7.5, 6 mM $ZnCl_2$, and 150 mM NaCl at 37 °C. ~5 turnovers were observed in 24 h. B) Malachite green dye assay, in which $P_i$ is observed colorimetrically upon binding to ammonium molybdate and malachite green dye. Conditions: 10 μM 14WM9, 12.5 μM DNA anchor oligonucleotide, 500 μM free AAAY$^{P}$AA, 70 mM HEPES, pH 7.5, 2 mM $ZnCl_2$, and 150 mM NaCl at 37 °C. As calculated from $A_{620}$ values, 6 turnovers in 24 h and 15 turnovers in 96 h were observed. In both the HPLC and malachite green assays, omission of $Zn^{2+}$ rather than 14WM9 also led to no activity in 24 h.

LC-ESI mass spectrometry of the reaction sample confirmed the identities of the phosphorylated substrate and dephosphorylated product hexapeptides (all $[M+H]^+$: Y$^{P}$ substrate *m*/*z* calcd. 616.2, found 616.2; Y$^{OH}$ product *m*/*z* calcd. 536.3, found 536.2). These assays revealed 6 turnovers in 24 h and 15 turnovers in 96 h by 14WM9. The DNA anchor oligonucleotide, now unconnected to the free peptide substrate, was required for efficient reaction, presumably to sequester the corresponding deoxyribozyme binding arm (truncation of this binding arm led to 12-fold lower yield at 24 h). In all of these assays, the higher deoxyribozyme concentrations (100 μM or 10 μM, rather than 1 μM as in the single-turnover assays) required elevated $Zn^{2+}$ concentration (6 mM or 2 mM, rather than 1 mM) for maximal activity, apparently to compensate for nonspecific chelation of $Zn^{2+}$ by the increased amount of DNA. The $K_m$ for AAAY$^{P}$AA was ≥1 mM (Figure 2.12; the assays of Figure 2.11 were performed at 0.5 mM peptide).

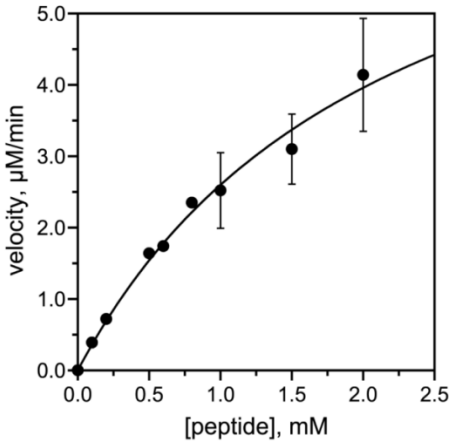


**Figure 2.12.** Determination of $K_m$ for the 14WM9 deoxyribozyme. Data points are $n = 1$ for peptide concentrations below 1.0 mM, $n = 3$ for 1.0 mM, $n = 5$ for 1.5 mM, and $n = 6$ for 2.0 mM (error bars for $n > 1$ are standard deviation). Each velocity determination was derived from a linear fit of product formation (μM) versus time (min) by HPLC. The fit value to the standard equation $V = V_{max} \bullet (K_m/([peptide] + K_m)$ gave $V_{max} = 8.3 \pm 1.1$ μM/min and $K_m = 2.2 \pm 0.5$ mM. However, because the data points do not turn over at high peptide concentration (and >2 mM peptide was not practical), the safest assessment is that $K_m \geq 1$ mM.

The multiple-turnover activity of 14WM9 was not substantially inhibited by an excess of the dephosphorylated AAAY$^{OH}$AA peptide product, but considerable inhibition by $P_i$ was observed when over 2 equivalents was included (Figure 2.13). The sequence-unrelated 14WM27 deoxyribozyme also exhibited multiple-turnover phosphatase activity with the free AAAY$^{P}$AA substrate and $Zn^{2+}$ (Figure 2.14), demonstrating that DNA-catalyzed dephosphorylation of a free peptide substrate is not limited to a single deoxyribozyme.

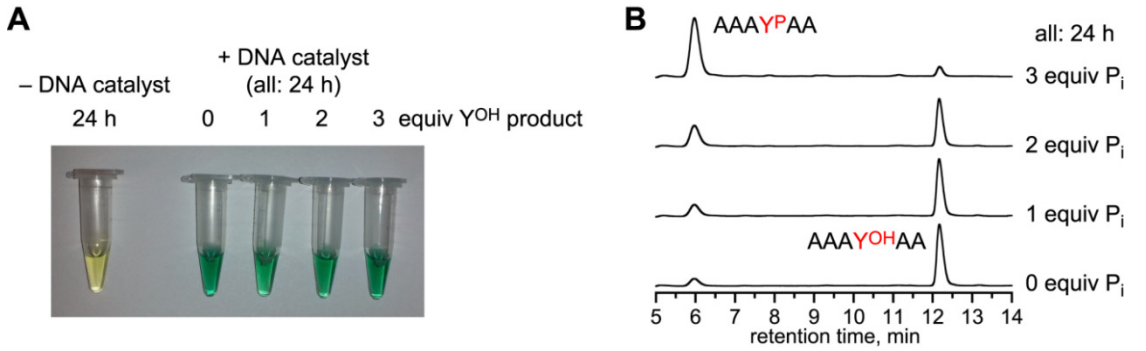


**Figure 2.13.** Evaluating inhibition of the 14WM9 deoxyribozyme by the dephosphorylated AAAY$^{OH}$AA peptide product and by inorganic phosphate ($P_i$). A) Evaluating inhibition by AAAY$^{OH}$AA, as determined by the malachite green dye assay for detection of released $P_i$. Each assay included 100 µM 14WM9, 125 µM DNA anchor oligonucleotide, 500 µM AAAY$^{P}$AA substrate, and up to 1.5 mM AAAY$^{OH}$AA product in 70 mM HEPES, pH 7.5, 6 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C. No inhibition was observed, even with 3 equivalents of product relative to substrate. B) Evaluating inhibition by $P_i$, as determined by HPLC assay. Each assay included 100 µM 14WM9, 125 µM DNA anchor oligonucleotide, 500 µM AAAY$^{P}$AA substrate, and up to 1.5 mM $P_i$ in 70 mM HEPES, pH 7.5, 6 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C. Substantial inhibition was reproducibly observed, especially above 2 equivalents of $P_i$ relative to substrate.

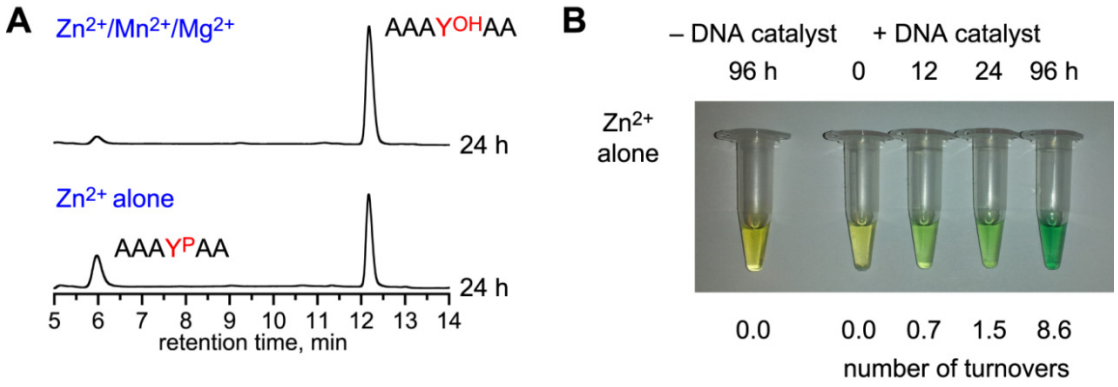


**Figure 2.14.** Activity of the 14WM27 deoxyribozyme with the free AAAY$^{P}$AA substrate. A) HPLC assay, under conditions of Figure 2.11A (either 6 mM $Zn^{2+}$ alone, or 6 mM $Zn^{2+}$ with 20 mM $Mn^{2+}$ and 40 mM $Mg^{2+}$). B) Malachite green assay for released $P_i$, under conditions of Figure 2.11B.

### 2.2.5 Substrate Dependence of 14WM9

To gain insight into the interactions between the DNA catalyst and the substrate, we examined the dependence of 14WM9's single-turnover catalytic activity on the chemical structure of its phosphorylated substrate. When the DNA-anchored hexapeptide was shortened to as little as a dipeptide (CY$^{P}$ or CS$^{P}$, where the cysteine sulfhydryl was

used to attach the dipeptide to the DNA anchor), 14WM9 retained substantial activity (Figure 2.15).

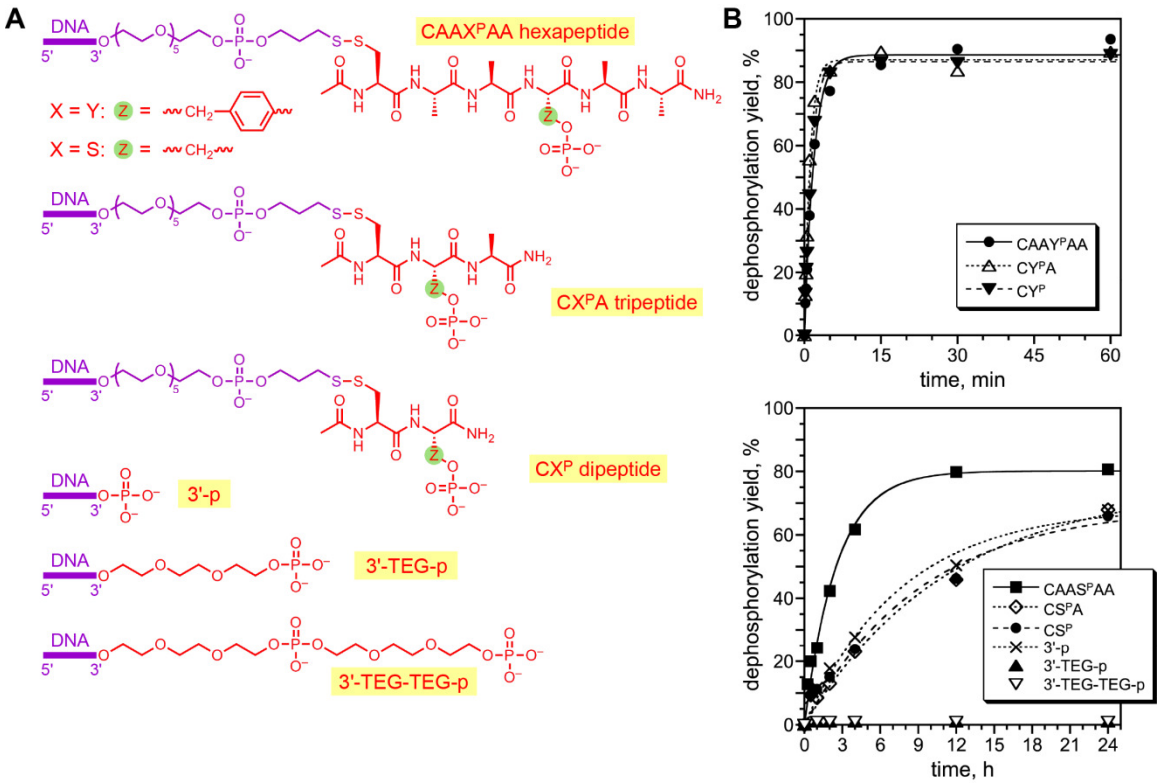


**Figure 2.15.** Dependence of 14WM9 catalytic activity on the chemical structure of the phosphorylated substrate. A) Structures of the evaluated substrates. Note that for these experiments, the hexapeptide was connected via an N-terminal Cys, rather than an N-terminal Ala as in Figure 2.10. B) Single-turnover assays, under conditions of Figure 2.10 ($Zn^{2+}/Mn^{2+}/Mg^{2+}$). $k_{obs}$ values from PAGE assays, top to bottom as listed in figure legends: 0.54, 0.93, 0.75 $min^{-1}$; 0.39, 0.087, 0.11, 0.13 $h^{-1}$ [$6.5 \times 10^{-3}$, $1.5 \times 10^{-3}$, $1.8 \times 10^{-3}$, $2.2 \times 10^{-3}$ $min^{-1}$].

Consistent with this finding, substrates with sequence CA<u>X</u>Y$^{P}$AA and CAAY$^{P}$<u>X</u>A (where <u>X</u> = F, E, or K) were dephosphorylated to similar extents by 14WM9 (Figure 2.16). In contrast, when the substrate was non-peptidic and the phosphate group was instead connected to the DNA anchor through one or two triethylene glycol (TEG) units, 14WM9 had no detectable dephosphorylation activity (<0.5%; Figure 2.15). When the phosphate group was attached directly to the 3′-hydroxyl of the DNA anchor, activity was comparable to that observed for the shorter CS$^{P}$ substrate, indicating that the

floppiness introduced by the TEG linkers contributes to the lack of activity with those substrates (Figure 2.15). From these data, we conclude that 14WM9 substantially prefers peptidic substrates, although only a relatively short peptide segment is required for optimal activity.

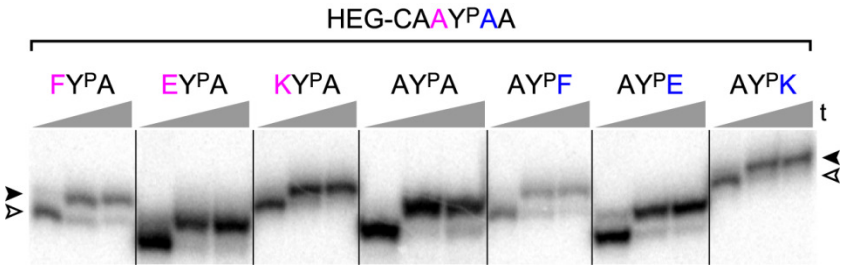


**Figure 2.16.** Assays of 14WM9 with various hexapeptide substrate sequences in which one amino acid adjacent to the $Y^P$ was altered, under conditions of Figure 2.10 ($Zn^{2+}$/$Mn^{2+}$/$Mg^{2+}$). *t* = 0, 2, 12 h. Data were obtained in parallel with Figure 2.15.

### 2.2.6 14WM9 Catalysis Under More Biologically Relevant Conditions

To investigate the ability of 14WM9 to function under more biologically representative conditions, its activity was assayed in the presence of human cell lysate or bovine serum albumin (BSA), using the free $AAAY^PAA$ substrate. When up to 160 μg/mL of H1299 human non-small-cell lung carcinoma lysate protein was supplemented into the same buffer and metal ion components as used for the previous experiments, substantial catalytic activity by 14WM9 was retained (Figure 2.17A). Separately, when 14WM9 was assayed in the presence of BSA, substantial activity was maintained up to 20 mg/mL BSA (Figure 2.17B and Figure 2.18). Multiple-turnover behavior for the 14WM27 deoxyribozyme was also observed in the presence of lysate or BSA (Figure 2.19). Collectively, these findings establish that robust DNA-catalyzed phosphatase activity is compatible with high concentrations of nonspecific proteins and other native cellular components.

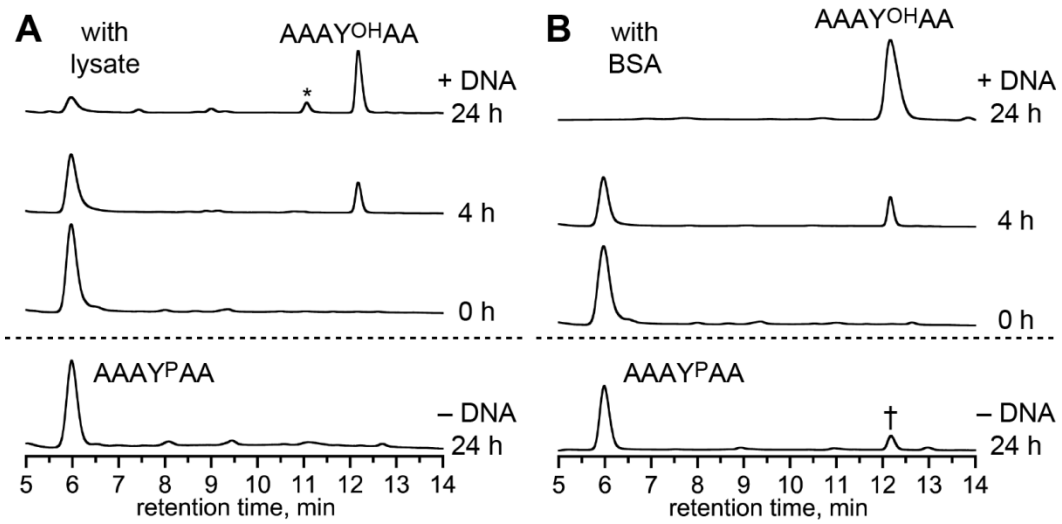


**Figure 2.17.** Activity of the 14WM9 deoxyribozyme in presence of human cell lysate and BSA. A) HPLC multiple-turnover assay with free AAAY$^{P}$AA peptide in the presence of human cell lysate (160 µg/mL of lysate protein), under conditions of Figure 2.11A. The asterisk denotes the pentapeptide AAY$^{OH}$AA formed by nonspecific peptidase activity in the lysate, as validated by ESI-MS ([M+H]$^{+}$ *m*/*z* calcd. 465.2, found 465.2). B) HPLC multiple-turnover assay in the presence of 20 mg/mL BSA, under conditions of Figure 2.11A except 10 mM $Zn^{2+}$. Substantial activity was also observed at 6 and 8 mM $Zn^{2+}$, but reduced presumably due to nonspecific chelation of $Zn^{2+}$ by BSA (Figure 2.18). The dagger denotes a small amount (~14%) of dephosphorylation product formed in 24 h in the absence of 14WM9. In both panels, omission of $Zn^{2+}$ rather than 14WM9 led to no activity in 24 h. This finding suggests for panel B that $Zn^{2+}$ activates nonspecific phosphatases present in the BSA, and inclusion of the DNA suppresses this phosphatase activity.

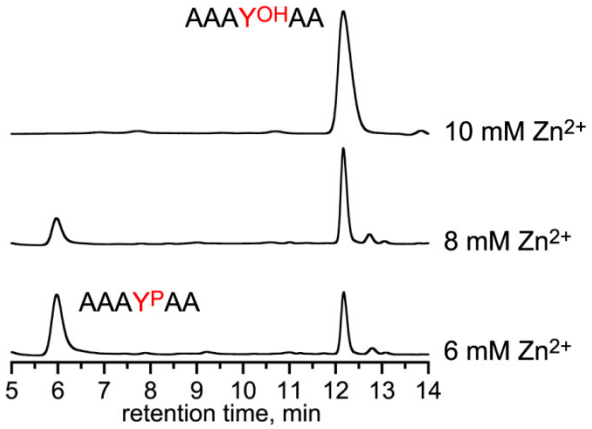


**Figure 2.18.** 14WM9 multiple-turnover activity with free peptide substrate assayed in the presence of 20 mg/mL BSA and different $Zn^{2+}$ concentrations. Data acquired as in Figure 2.17B ($t$ = 24 h).

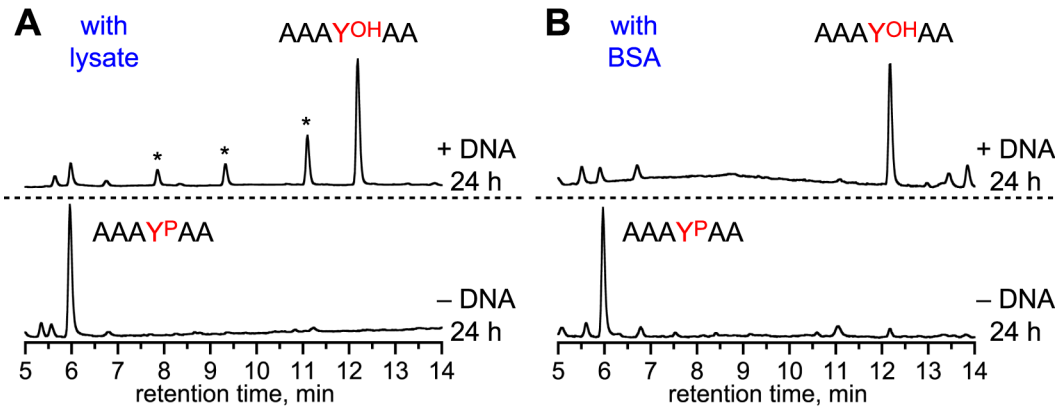


**Figure 2.19.** 14WM27 multiple-turnover activity assayed in the presence of human cell lysate and BSA. A) HPLC assay in the presence of human cell lysate (160 µg/mL lysate protein). Conditions: 100 µM 14WM9, 125 µM DNA anchor oligonucleotide, 500 µM free AAAY$^{P}$AA, 70 mM HEPES, pH 7.5, 6 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C. The asterisks denote shorter peptide fragments formed by nonspecific peptidase activity in the lysate. B) HPLC assay in the presence of 20 mg/mL BSA. Conditions: 10 µM 14WM9, 12.5 µM DNA anchor oligonucleotide, 500 µM free AAAY$^{P}$AA, 70 mM HEPES, pH 7.5, 10 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C.

Finally, we assessed the ability of 14WM9 to dephosphorylate a Y$^{P}$ residue when presented as part of a large 91-mer protein substrate rather than merely on a short hexapeptide. For this purpose, we used as substrate a 66-mer fragment of prochlorosin ProcA2.8, [31] fused to an N-terminal $His_6$ tag and joined at its C-terminus by native chemical ligation to the CAAY$^{P}$AA hexapeptide. The resulting 91-mer protein was incubated with 14WM9 and analyzed by MALDI mass spectrometry, revealing substantial dephosphorylation (94% in 36 h; Figure 2.20).

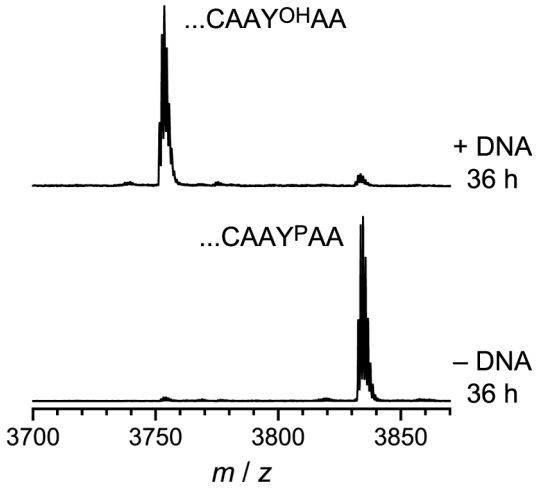


**Figure 2.20.** Dephosphorylation of a larger protein substrate by 14WM9. The 91-mer protein derived from prochlorosin ProcA2.8 and ending with ...CAAY$^{P}$AA at its C-terminus was dephosphorylated by 14WM9, digested in-gel by Lys-C, and analyzed by MALDI mass spectrometry (see Materials and Methods for details). Dephosphorylation conditions: 100 μM 14WM9, 125 μM DNA anchor oligonucleotide, 20 μM protein, 70 mM HEPES, pH 7.5, 6 mM $ZnCl_2$, 40 mM $MgCl_2$, 20 mM $MnCl_2$, and 150 mM NaCl at 37 °C for 36 h. Y$^{OH}$ product Lys-C fragment $[M+H]^+$ *m/z* calcd. 3751.725, found 3751.699, $\Delta = -6.9$ ppm. Difference between product and substrate due to dephosphorylation $\Delta(m/z)$ calcd. 79.966, found 79.993, $\Delta = +0.033\%$.

### 2.2.7 In Vitro Selection Process for N-terminal Phosphoserine and Capture Deoxyribozymes for Serine

A strategy to identify phosphatase deoxyribozymes for N-terminal phosphoserine was devised (Figure 2.21A). The hexapeptide substrate, S$^{P}$AAAAC, was covalently appended to a DNA-anchored oligonucleotide (by disulfide linkage), which was then ligated to the deoxyribozyme pool. After each key selection step that used incubation conditions of 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, active DNA sequences were precipitated and "captured" in a two-step chemical technique. The first step involved oxidation of the N-terminal serine by sodium periodate for 30 min to reveal an aldehyde. After desalting and removing the excess periodate, a secondary capture step was performed with a 5′-hydrazide DNA to form the hydrazone resulting in an upward PAGE shift (Figure 2.21A) (see Materials and Methods for capture conditions). The active DNA sequences were

amplified by PCR and joined by T4 DNA ligase to the DNA-anchored hexapeptide substrate for entering the next selection round. After 7 rounds of selection, positive control yield gradually decreased from 25% to ~10% (Figure 2.21B). Because the positive control itself was giving unreliably low yields, these selections were discontinued.

Separately, for serine-OH capture deoxyribozymes, reselections from 15MZ36 (which exhibits trace activity with AAASAA), and redirection from 11MY (continuation from an earlier selection round that captures non-peptidic aliphatic OH) were initiated. The DNA-anchored AAASAA substrate was prepared by reductive amination (see Materials and Methods). Each selection step used incubation conditions of 70 mM HEPES, pH 7.5, 100 pmol of 5′-triphosphorylated RNA, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h Activity was observed in round 6 for the reselection and round 17 for redirection. Individual sequences were cloned, and sequenced. The best deoxyribozyme from both selections, 17WK3, showed 37% serine capture activity and selective for serine-OH over phosphoserine (Figure 2.21C).

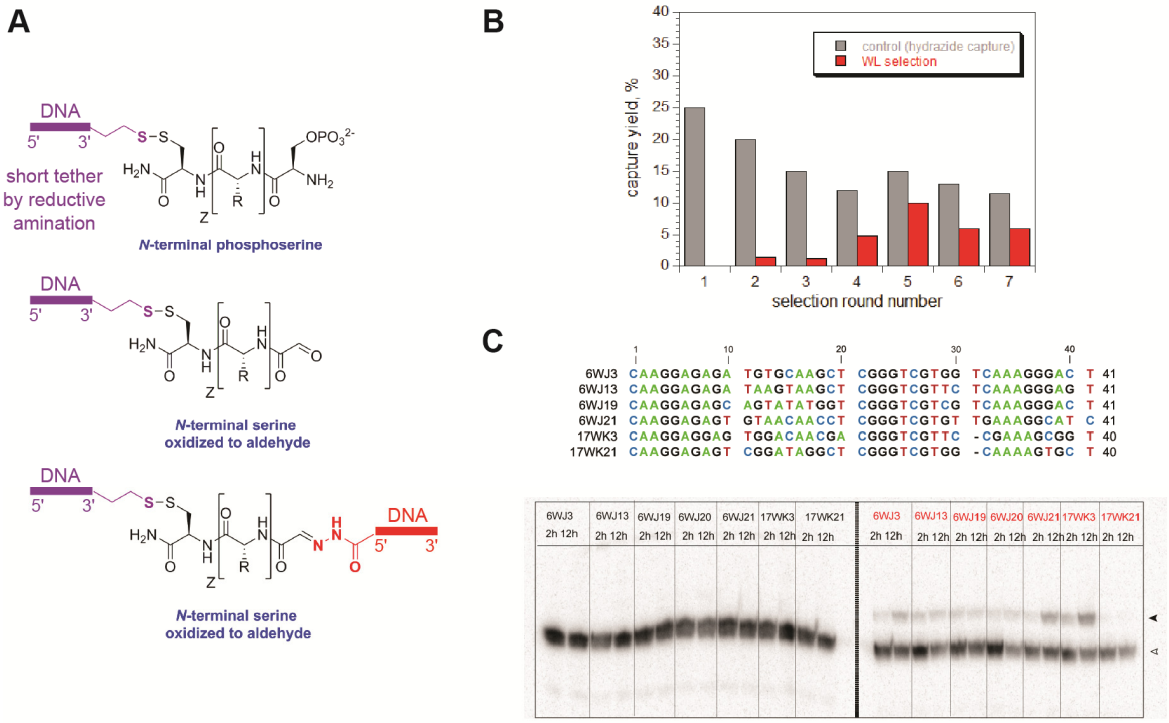


**Figure 2.21.** A) Selection scheme for N-terminal phosphoserine phosphatase activity. After the selection step, dephosphorylated serine is oxidized to reveal an aldehyde, which is captured by the hydrazide oligonucleotide (red). B) Selections were performed for 7 rounds and discontinued. C) The catalytic region for capture deoxyribozyme (40 or 41 nt) and PAGE image illustrating selectivity for $AAAS^{OH}AA$ (red) over $AAAS^{P}AA$ (black) of the 6WJ- and 17WK-catalyzed capture reactions. The open arrowhead marks the reactant, and the filled arrowhead marks the product.

## 2.3 Summary

In this study, we have established the fundamental ability of DNA to catalyze the challenging reaction of phosphomonoester hydrolysis in the context of peptide substrates. The best DNA catalysts function with entirely free peptides. We also observed multiple-turnover behavior even in the presence of human cell lysate or high concentrations of BSA. In addition, substantial phosphatase activity with a larger protein substrate was also observed.

We previously demonstrated DNA-catalyzed DNA phosphodiester hydrolysis, i.e., nuclease activity.[8-9] Both reactions involve "phosphoester" substrates, but a phosphomonoester as used here is substantially more difficult to hydrolyze than is a

phosphodiester, considering the >3 orders of magnitude difference in uncatalyzed half-lives ($4 \times 10^{10}$ years for an aliphatic phosphomonoester dianion, [32] versus $3 \times 10^{7}$ years for a phosphodiester).[33] Regardless of the uncatalyzed reaction rate constants, the peptide substrate for phosphomonoester dephosphorylation cannot be bound by the DNA catalyst using any Watson-Crick base pairs, whereas the DNA substrate for phosphodiester hydrolysis inherently interacts with the deoxyribozyme by extensive preprogrammed base pairing. This means that a phosphatase deoxyribozyme has the much greater challenge to interact well with its substrate. The new DNA-catalyzed phosphatase activity has rate constant as fast as 0.7 $min^{-1}$ for $Y^P$ dephosphorylation, a value commensurate with many other in vitro selected deoxyribozymes for various catalytic activities.[3, 7, 34]

From a fundamental perspective, we wish to understand the mechanism of DNA-catalyzed dephosphorylation, in the context of broader studies of biological phosphoryl transfer.[35] Such mechanistic studies will benefit greatly from future high-resolution structural information that is currently unavailable except for one deoxyribozyme.[36,37] Although we can speculate that 14WM9, 14WM27, and the other newly identified phosphatase deoxyribozymes activate $Zn^{2+}$ for direct hydrolysis of the phosphorylated side chains, determining their detailed mechanisms requires considerable further experiments, especially without a sound structural basis for the catalysis.

The marked preference of the 14WM9 deoxyribozyme for a peptidic substrate suggests direct interaction of this deoxyribozyme with some portion of the substrate's polyamide backbone, although the efficient dephosphorylation of a dipeptide substrate means that any such contact is not very extensive.

## 2.4 Materials and Methods

### 2.4.1 Oligonucleotides

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′-Triphosphorylated RNA oligonucleotides were prepared by in vitro transcription using synthetic DNA templates and T7 RNA polymerase.[38] 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 previously.[39-40]

### 2.4.2 Synthesis of Peptides and DNA-Anchored Peptides

Peptides were prepared by solid-phase synthesis on Fmoc Rink amide MBHA resin (AAPPTec). Each peptide was coupled to the DNA anchor oligonucleotide via either the N-terminal α-amino group (linkage created by reductive amination) or the N-terminal cysteine side chain (linkage created by disulfide formation).

**2.4.2.1 Synthesis of peptides.** Peptides were prepared by solid-phase synthesis using Fmoc Rink amide MBHA resin, with N,N,N′,N′-tetramethyl-O-(7-azabenzotriazol-1-yl)uronium hexafluorophosphate (HATU) as coupling agent. Each synthesis was performed at 0.2 mmol scale, initiated using 308 mg of Rink amide resin with a loading capacity of 0.65 mmol/g (AAPPTec). All steps were monitored by ninhydrin test. For each coupling, 5 equivalents (1.0 mmol) of Fmoc-amino acid, 4.9 equivalents (373 mg, 0.98 mmol) of HATU, and 10 equivalents (350 μL, 2.0 mmol) of N,N-diisopropylethylamine (DIPEA) were mixed in 10 mL of anhydrous DMF. The coupling reaction was initiated by adding this mixture to the resin and agitating by bubbling with

nitrogen for 45 min, followed by washing with DMF (3 □ 10 mL). The N terminus of the newly installed amino acid was deprotected by agitating the resin in 20% piperidine in DMF (5 mL) under nitrogen for 5 min a total of three times, each time washing with DMF (3 □ 5 mL). Peptides with N-terminal cysteine were capped via acetylation of their N-terminus by agitating the resin in 50 equivalents (924 μL, 10 mmol) of acetic anhydride and 25 equivalents of DIPEA (870 □L, 5 mmol) in 5 mL of DMF under nitrogen for 30 min. The peptide was cleaved from the solid support by stirring the resin in a separate vial with a solution containing 5 mL of trifluoroacetic acid (TFA), 125 μL of water, 125 μL of ethanedithiol, and 50 □L of triisopropylsilane for 90 min. The liquid solution was separated from the resin by filtration. This solution was dried on a rotary evaporator, providing a solid (for hexapeptides) or an oily solid (for dipeptides). To this material 20 mL of cold diethyl ether was added, and the peptide was obtained as a white solid that was filtered and purified by HPLC.

**2.4.2.2 Synthesis of DNA-anchored peptides.** Each peptide was coupled to the DNA anchor oligonucleotide via either the N-terminal α-amino group (linkage created by reductive amination) or the N-terminal cysteine side chain (linkage created by disulfide formation).

*Procedure for coupling to N-terminal α-amino group (Figure 2.6A):* The DNA anchor oligonucleotide was 5′-GGATAATACGACTCACTAT-rA-3′, where the 3′-terminal ribonucleotide was oxidized by $NaIO_4$. A 100 μL sample containing 1.0 nmol of DNA anchor oligonucleotide in 100 mM HEPES, pH 7.5, and 10 mM $NaIO_4$ was incubated at room temperature for 1 h. The oxidized product was precipitated to remove excess $NaIO_4$ by addition of 10 μL of 3 M NaCl and 300 μL of ethanol. The precipitated product was

dissolved in 65 μL of water and used directly in the next step. A 100 μL sample containing the $NaIO_4$-oxidized DNA anchor oligonucleotide and 100 nmol (100 equivalents) of peptide in 100 mM NaOAc, pH 5.2, 50 mM $NiCl_2$, and 10 mM $NaCNBH_3$ was incubated at 37 °C for 14 h. The DNA-anchored hexapeptide was precipitated by addition of 10 μL of 3 M NaCl and 300 μL of ethanol and purified by 20% PAGE. A typical yield was 200–400 pmol. See Table 2.1 for MALDI mass spectrometry values.

*Procedure for coupling to N-terminal cysteine side chain (Figure 2.6B):* The DNA anchor oligonucleotide was 5′-GGATAATACGACTCACTAT-$C_3$-SS-$C_3$-OH-3′, where the 3′-disulfide linker was introduced via standard solid-phase DNA synthesis and unmasked to a 3′-thiol by DTT treatment. In some cases, one or ten hexa(ethylene glycol), or HEG, spacer units were included immediately to the 5′-side of the disulfide linker; see structures in Figure 2.11. A 50 μL sample containing 2 nmol of DNA anchor oligonucleotide in 50 mM HEPES, pH 7.5, and 50 mM DTT was incubated at 37 °C for 2 h. The reduced product was precipitated to remove excess DTT by addition of 50 μL of water, 10 μL of 3 M NaCl, and 300 μL of ethanol. The precipitated product (DNA-$C_3$-SH) was dissolved in 45 μL of water. 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-$C_3$-SSPy) was precipitated by addition of 50 μL of water, 10 μL of 3 M NaCl and 300 μL of ethanol and dissolved in 25 μL of water. Conjugation to the peptide was performed by adding 20 μL of 50 mM triethylammonium acetate, pH 7.0, and 5 μL of 20 mM peptide (100 nmol, 100 equivalents). The sample was incubated at 37 °C for 2 h, and the DNA-anchored peptide was purified by 20% PAGE. A typical yield was 1.4–1.8 nmol. See Table 2.1 for MALDI mass spectrometry values.

### 2.4.3 Capture Strategy for Phosphotyrosine Capture using 15MZ36

The DNA-anchored AAAYAA substrate ligated to the DNA pool was 5′-$^{32}$P-radiolabeled using γ-$^{32}$P-ATP and PNK (Fermentas).

*Procedure for capture step using selective chelation of zinc with EDTA.* A 10 μL sample containing 0.1 pmol radiolabeled ligated pool (with DNA-anchored AAAYAA or AAAY$^{P}$AA) 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. To the 20 μL selection sample containing was added 0.8 μL of 50 mM EDTA, pH 8.0 (40 nmol). The sample was then brought to 25 μL total volume containing 40 pmol of 15MZ36 capture deoxyribozyme, 80 pmol of 5′-triphosphorylated RNA substrate, and 20 mM $MnCl_2$, 40 mM $MgCl_2$, and incubated at 37 °C for 14 h.

*Procedure for capture step using ethanol precipitation.* The 20 μL selection sample containing 0.1 pmol radiolabeled ligated pool (with DNA-anchored AAAYAA or AAAY$^{P}$AA) in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl was precipitated in 75% ethanol after adding 2.5 μL of 500 mM EDTA, pH 8.0 (1.25 mmol) to remove all divalent metal cations. The sample was then brought to a 20 μL total volume containing 40 pmol of 15MZ36 capture deoxyribozyme, 80 pmol of 5′-triphosphorylated RNA substrate, and 20 mM $MnCl_2$, 40 mM $MgCl_2$, and incubated at 37 °C for 14 h.

For both cases at the 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.

### 2.4.4 In Vitro Selection Procedure for Tyrosine Phosphatase Activity

The selection procedure, cloning, and initial analysis of individual clones were performed essentially as described previously,[39, 41] but with a different ligation step as recently reported. [30] An overview of the key selection and capture steps of each round is shown in Figure 2.5.

The selection procedure, cloning, and initial analysis of individual clones were performed essentially as described previously [39, 41], but with a different ligation step as recently reported [30]. An overview of the key selection and capture steps of each round is shown in Figure 2.5. The random deoxyribozyme pool was 5′-CGAAGTCCGCCATCTCTTC–$N_{40}$–ATAGTGAGTCGTATTAAGCTGATCCTGATGG-3′. PCR primers were 5′-CGAAGTCGCCATCTCTTC-3′ (forward primer) and 5′-$(AAC)_4$XCCATCAGATCAGCT-3′, where X is the HEG spacer to stop Taq polymerase (reverse primer). In each round, the ligation step to attach the deoxyribozyme pool at its 3′-end with the 5′-end of the DNA-anchored hexapeptide substrate was performed using a DNA splint and T4 DNA ligase. The splint sequence was 5′-ATAGTGAGTCGTATTATCCTCCATCAGGATCAGCTTAATACGACTCACTAT-3′, where the underlined T is included to account for the untemplated A nucleotide that is added at the 3′-end of each PCR product by Taq polymerase. This T nucleotide was omitted from the splint used for ligation of the initially random $N_{40}$ pool, which was prepared by solid-phase synthesis without the untemplated A. Nucleotide sequences of the DNA anchor oligonucleotide, the deoxyribozyme binding arms, the 5′-

triphosphorylated RNA substrate for the capture reaction, and the full 15MZ36 capture deoxyribozyme are shown in Figure 2.7.

*Procedure for ligation step in round 1.* A 34 μL sample containing 1 nmol of DNA pool, 850 pmol of DNA splint, and 750 pmol of 5′-phosphorylated DNA-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 (Fermentas) and 2 μL of 5 U/μL T4 DNA ligase (Fermentas). The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

*Procedure for ligation step in subsequent rounds.* A 16 μL sample containing the PCR-amplified DNA pool (~5–10 pmol), 20 pmol of DNA splint, and 30 pmol of 5′-phosphorylated DNA-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 (Fermentas) and 2 μL of 1 U/μL T4 DNA ligase (Fermentas). The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

*Procedure for selection step in round 1.* Each selection experiment was initiated with 200 pmol of the ligated $N_{40}$ pool. A 15 μL sample containing 200 pmol of ligated $N_{40}$ pool 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 30 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.

*Procedure for selection step in subsequent rounds.* A 10 μL sample containing ligated pool 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. The sample was incubated at 37 °C for 14 h.

*Procedure for capture step in round 1.* To the 30 μL selection sample was added 1.2 μL of 50 mM EDTA, pH 8.0 (60 nmol). The sample was then brought to 45 μL total volume containing 300 pmol of 15MZ36 capture deoxyribozyme, 400 pmol of 5′-triphosphorylated RNA substrate, 20 mM $MnCl_2$, and 40 mM $MgCl_2$ and incubated at 37 °C for 14 h.

*Procedure for capture step in subsequent rounds.* To the 20 μL selection sample was added 0.8 μL of 50 mM EDTA, pH 8.0 (40 nmol). The sample was then brought to 25 μL total volume containing 40 pmol of 15MZ36 capture deoxyribozyme, 80 pmol of 5′-triphosphorylated RNA substrate, and 20 mM $MnCl_2$, 40 mM $MgCl_2$, and incubated at 37 °C for 14 h.

*Procedure for PCR.* 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-purified selection product, 200 pmol of forward primer, 50 pmol of reverse primer, 20 nmol of each dNTP, and 10 μL of 10× Taq polymerase buffer (1× = 20 mM Tris-HCl, pH 8.8, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, and 0.1% Triton X-100). This sample 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 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), and 5 μL of 10× Taq polymerase buffer. 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.

### 2.4.5 Single-Turnover Assay Procedure Using PAGE

The DNA-anchored phosphopeptide substrate was 5′-$^{32}$P-radiolabeled using γ-$^{32}$P-ATP and Optikinase (USB), which lacks the 3′-phosphatase activity that we have found also dephosphorylates tyrosine and serine side chains. A 10 μL sample containing 0.25 pmol of 5′-$^{32}$P-radiolabeled DNA-anchored phosphopeptide substrate and 20 pmol of 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 DNA-catalyzed dephosphorylation 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 ($Mn^{2+}$ and $Mg^{2+}$ were omitted in assays that used $Zn^{2+}$ alone). The sample was 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 = $Y \bullet (1 - e^{-kt})$, where $k = k_{obs}$ and $Y$ is the final yield. Each $k_{obs}$ value is reported with

error calculated as the standard deviation from the indicated number of independent determinations.

### 2.4.6 Multiple-Turnover Assay using HPLC

A 20 μL sample containing 2 nmol (100 μM) of deoxyribozyme, 2.5 nmol (125 μM) of free DNA anchor oligonucleotide, and 10 nmol (500 μM) of free AAAY$^{P}$AA hexapeptide in 70 mM HEPES, pH 7.5, 6 mM $ZnCl_2$, and 150 mM NaCl was incubated at 37 °C. At an appropriate time point (0–24 h), the sample was quenched with 5 μL of 50 mM EDTA, pH 8.0 (250 nmol, versus 120 nmol $Zn^{2+}$ present) and frozen at –80 °C until further processing. To remove the deoxyribozyme, a filtration step was performed (this step was confirmed not to distort substantially the ratio of phosphorylated to dephosphorylated peptide). An Amicon Ultra-0.5 mL 3 kDa centrifugal filter was washed with 300 μL of water by centrifugation at 14000 *g* for 10 min. The sample was diluted with water to 300 μL and transferred to the filter, which was centrifuged at 14000 *g* for 15 min, and the filtrate was collected. Another 300 μL of water was passed through the filter. The combined filtrates were evaporated to dryness, redissolved in 25 μL of water, and analyzed by HPLC. Analysis was performed on a Beckman System Gold instrument with a Beckman Ultrasphere $C_{18}$ column (5 μm, 2 × 150 mm). Samples were analyzed using a gradient of 1% solvent A (acetonitrile) and 99% solvent B (0.1% trifluoroacetic acid in water) at 0 min to 21:79 A:B at 25 min at flow rate 0.5 mL/min.

For assays in the presence of H1299 human non-small-cell lung carcinoma lysate, the lysate (40 mg/mL of lysate protein) was a gift prepared by B. Parkinson in the P. Hergenrother laboratory [lysis in RIPA lysis buffer containing protease inhibitor, with cell debris removed by centrifugation at 16000 *g* for 5 min [42] and quantification by

Bradford and bicinchoninic acid (BCA) assays]. The 20 μL sample additionally contained 1/250 by volume of lysate (160 μg/mL of lysate protein). For assays in the presence of BSA, the 20 μL sample additionally contained 20 mg/mL of BSA (Sigma A4503), and the $Zn^{2+}$ concentration was 10 mM.

### 2.4.7 Multiple-turnover Assay Procedure Using Malachite Green Dye

A 20 μL sample containing 200 pmol (10 μM) of deoxyribozyme, 250 pmol (12.5 μM) of free DNA anchor oligonucleotide, and 10 nmol (500 μM) of free $AAAY^{P}AA$ hexapeptide in 70 mM HEPES, pH 7.5, 2 mM $ZnCl_2$, and 150 mM NaCl was incubated at 37 °C. At an appropriate time point (0–96 h), the sample was quenched with 5 μL of 50 mM EDTA, pH 8.0 (250 nmol) and frozen at –80 °C until further processing. To the sample was added 75 μL of water and 20 μL of malachite green assay solution (BioAssay Systems POMG-25H). After incubation at room temperature for 30 min, the sample was diluted with 180 μL of water, and the absorbance at 620 nm was measured (NanoDrop 2000c, Thermo Scientific). Standard curves for inorganic phosphate ($P_i$) were obtained using solutions containing known amounts of $P_i$ and all other assay components except the free hexapeptide. Numbers of turnovers were calculated directly from the ratio of mole amounts of released $P_i$ and deoxyribozyme.

### 2.4.8 Mass Spectrometry of DNA-Anchored Peptides

The DNA-anchored $AAAY^{OH}AA$ dephosphorylation product was prepared from a 10 μL sample containing 50 pmol of DNA-anchored $AAAY^{P}AA$ substrate and 80 pmol of 14WM9 deoxyribozyme, which 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 dephosphorylation 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 5 min, quenched with 5 μL of 50 mM EDTA, pH 8.0 (250 nmol), desalted by Millipore $C_{18}$ ZipTip, and analyzed by MALDI mass spectrometry (matrix 3-hydroxypicolinic acid).

### 2.4.9 Mass Spectrometry of Free Hexapeptides

The free AAAY$^{OH}$AA dephosphorylation product was prepared from 10 nmol of the free AAAY$^{P}$AA substrate using 2 nmol of the 14WM9 deoxyribozyme and the multiple-turnover procedure described above (24 h incubation). Analysis was performed on a Waters 2795 LC-MS instrument with a Beckman Ultrasphere $C_{18}$ column (5 μm, 2 × 150 mm) and detection by ESI mass spectrometry (Q-Tof Ultima API in positive ion scan mode using the manufacturer's suggested parameters). Samples were analyzed using a gradient of 1% solvent A (0.1% formic acid in acetonitrile) and 99% solvent B (0.1% formic acid in water) at 0 min to 21:79 A:B at 25 min.

### 2.4.10 Large Protein Preparation and Analysis

The 85-mer $His_6$-ProcA2.8(1-66) MESNA thioester was overexpressed and partially purified. The procedure was followed essentially as reported, [43] with minor modifications. *E. coli* BL21(DE3) cells harboring the pTXB1 plasmid that encodes the $His_6$-ProcA2.8(1-66)-intein-CDB fusion protein were grown in 2 L LB with 100 μg/mL ampicillin at 37 °C until an $OD_{600}$ of 0.6-0.7 was reached. Protein expression was induced by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and the cells were grown for an additional 12 h at 18 °C. Cells were harvested by centrifugation at 5500 *g* for 20 min at 4 °C, and the cell pellet was resuspended in 30 mL of column buffer containing 100 mM sodium phosphate, pH 7.0, 500 mM NaCl, 1 mM

EDTA, and 1 mM TCEP. Cells were lysed by sonication on ice at an amplitude of 65% (3 s on, 9 s off) for 16 min. The lysate was centrifuged at 20000 *g* for 20 min, and the supernatant containing the fusion protein was loaded onto a chitin column (20 mL) that was pre-equilibrated with column buffer. The fusion protein was bound to the chitin column by shaking gently at 4 °C for 2 h. The column was mounted upright and washed with 200 mL (10 column volumes) of the column buffer. The column was washed with 40 mL of buffer containing 100 mM sodium phosphate, pH 7.4, 500 mM NaCl, 1 mM EDTA, and 1 mM TCEP. To the column was added 20 mL of cleavage buffer containing 100 mM sodium phosphate, pH 7.4, 500 mM NaCl, 1 mM EDTA, 1 mM TCEP, and 100 mM MESNA. Intein-mediated cleavage of the 85-mer $His_6$-ProcA2.8(1-66) thioester was performed by gentle shaking of the chitin resin for 12 h at 4 °C. The $His_6$-ProcA2.8(1-66) thioester was eluted, concentrated using a 15 mL Amicon 3 kDa filter (Millipore), and lyophilized to dryness.

A 100 µL sample containing 6 mg crude protein thioester (~0.6 mM thioester; most of the mass is salts) and 5 mM $CAAY^PAA$ hexapeptide in 100 mM HEPES, pH 7.5, 1 mM TCEP, and 10 mM MesNa was incubated at 4 °C for 16 h. To alkylate the cysteine side chain, 50 µL of 1 M iodoacetamide was added, and the sample was incubated at 25 °C for 45 min in the dark. The reaction mixture was acidified by adding 3 µL of 5% trifluoroacetic to a final concentration of 0.1% trifluoroacetic acid. The ligated protein was purified by HPLC on a Beckman System Gold instrument with a Grace Vydac $C_4$ column (5 µm, 4.6 × 250 mm) using a gradient of 2% solvent A (80% acetonitrile and 0.1% TFA in water) and 98% solvent B (0.1% trifluoroacetic acid in water) at 0 min to 100% solvent B at 45 min with flow rate of 1 mL/min. The ligated

protein eluted at 29 min and was collected, lyophilized, and redissolved in water. MALDI mass spectrometry $[M+H]^+$ calcd. 9681.5, found 9681.7, $\Delta = +0.002\%$ (matrix 2,5-dihydroxybenzoic acid). The final 91-mer protein sequence was GSSHHHHHHSSGLVPRGSHMSEEQLKAFLTKVQADTSLQEQLKIEGADVVAIA-K^AAGFSITTEDLNSHRQNLSDDELEGVAGGAACAAY$^P$AA, where the caret marks the Lys-C cleavage site closest to the $Y^P$ position, and the single cysteine has been alkylated by iodoacetamide.

To assay 14WM9-catalyzed dephosphorylation of the 91-mer protein, a 10 μL sample containing 1 nmol (100 μM) of deoxyribozyme, 1.25 nmol (125 μM) of free DNA anchor oligonucleotide, and 200 pmol (20 μM) of protein in 70 mM HEPES, pH 7.5, 6 mM $ZnCl_2$, 40 mM $MgCl_2$, 20 mM $MnCl_2$, and 150 mM NaCl was incubated at 37 °C. The deoxyribozyme was omitted in the control experiment. After 36 h, 5 μL of the sample was added to 10 μL of solution containing 60 mM Tris, pH 6.8, 2% SDS, 0.1% bromophenol blue, and 10% glycerol. The sample was separated by SDS-PAGE (1 M Tris, pH 8.45, 0.1% SDS, 5% acrylamide stacking, 15% acrylamide + 10% glycerol resolving), run at 120 V for 75 min. The gel was stained with Imperial Protein Stain (Pierce 24615), and the band corresponding to the protein was excised. In-gel Lys-C digestion was performed as reported,[44] except omitting treatment of the gel slice with DTT and iodoacetamide. For the Lys-C digestion step, the gel slice was treated with 40 μL of a solution containing 100 mM Tris, pH 8.0, 40 mM DTT, and 0.6 ng/μL Lys-C (Roche 11420429001). The acidic extraction step used 5% formic acid in acetonitrile. The sample was concentrated by SpeedVac to 10 μL, desalted by Millipore $C_{18}$ ZipTip,

and analyzed by MALDI mass spectrometry (matrix 2,5-dihydroxybenzoic acid). Monoisotopic *m*/*z* values were calculated using the PeptideMass calculator at ExPASy.

### 2.4.11 Selection Procedure for N-terminal Phosphoserine Phosphatase Activity

The in vitro selection process for N-terminal phosphoserine phosphatase activity is similar to the selection procedure described earlier with the following modifications in the capture steps.

*Procedure for capture step in round 1.* To the 40 μL selection sample was added 10 μL of 500 mM EDTA, pH 8.0 (5 mmol). The sample was precipitated by adding 250 μL of water and 900 μL of ethanol. The sample was brought back to 150 μL total volume containing 200 mM HEPES, pH 7.5, 10 mM $NaIO_4$ and incubated at 25 °C for 30 min to oxidize the N-terminal serine. Amicon Ultra-0.5 mL 3 kDa was used to desalt and remove the excess periodate by addition of 400 μL water followed by centrifugation at 14000 × g for 20 min (twice). The desalted sample was dried and brought up to a 40 μL total volume containing 100 mM NaOAc, pH 4.6, 1 nmol of hydrazide capture oligonucleotide, 50 mM $NiCl_2$ and incubated at 37 °C for 14 h. The sequence of hydrazide oligonucleotide was 5′- XGGAAGAGATGGCGACGGAACAACAACAACAACAAC-3′, where X = 5′-carboxy-modifier C10 + hydrazine.

*Procedure for capture step in subsequent rounds.* To the 20 μL selection sample was added 5 μL of 500 mM EDTA, pH 8.0 (2.5 mmol). The sample was precipitated by adding 275 μL of water and 900 μL ethanol. The sample was brought to 150 μL total volume containing 200 mM HEPES, pH 7.5, 10 mM $NaIO_4$ and incubated at 25 °C for 30 min. Excess periodate was removed using Amicon Ultra-0.5 mL 3 kDa (see above for procedure). The sample was dried and brought up to 20 μL total volume containing 100

mM NaOAc, pH 4.6, 400 pmol of hydrazide capture oligonucleotide, 50 mM $NiCl_2$ and incubated at 37 °C for 14 h.

## 2.5 References


(1) Radzicka, A.; Wolfenden, R. A proficient enzyme. *Science* **1995**, *267*, 90-93.

(2) Joyce, G. F. Forty Years of In Vitro Evolution. *Angew. Chem. Int. Ed.* **2007**, *46*, 6420-6436.

(3) Silverman, S. K. Deoxyribozymes: selection design and serendipity in the development of DNA catalysts. *Acc. Chem. Res.* **2009**, *42*, 1521-1531.

(4) Schlosser, K.; Li, Y. Biologically inspired synthetic enzymes made from DNA. *Chem. Biol.* **2009**, *16*, 311-322.

(5) Dalby, P. A. Strategy and success for the directed evolution of enzymes. *Curr. Opin. Struct. Biol.* **2011**, *21*, 473-480.

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

(7) Silverman, S. K. DNA as a Versatile Chemical Component for Catalysis, Encoding, and Stereocontrol. *Angew. Chem. Int. Ed.* **2010**, *49*, 7180-7201.

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

(9) 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.

(10) Doudna, J. A.; Cech, T. R. The chemical repertoire of natural ribozymes. *Nature* **2002**, 222-228.

(11) Emilsson, G. M.; Nakamura, S.; Roth, A.; Breaker, R. R. Ribozyme speed limits. *RNA* **2003**, *9*, 907-918.

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

(13) Lad, C.; Williams, N. H.; Wolfenden, R. The rate of hydrolysis of phosphomonoester dianions and the exceptional catalytic proficiencies of protein and inositol phosphatases. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5607-5610.

(14) Barford, D.; Das, A. K.; Egloff, M. P. The structure and mechanism of protein phosphatases: insights into catalysis and regulation. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 133-164.

(15) Denu, J. M.; Dixon, J. E. A catalytic mechanism for the dual-specific phosphatases. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 5910-5914.

(16) Tonks, N. K. Protein tyrosine phosphatases: from genes, to function, to disease. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 833-846.

(17) Allen, K. N.; Dunaway-Mariano, D. Phosphoryl group transfer: evolution of a catalytic scaffold. *Trends Biochem. Sci.* **2004**, *29*, 495-503.

(18) Allen, K. N.; Dunaway-Mariano, D. Markers of fitness in a successful enzyme superfamily. *Curr. Opin. Struct. Biol.* **2009**, *19*, 658-665.

(19) Seifried, A.; Schultz, J.; Gohla, A. Human HAD phosphatases: structure, mechanism, and roles in health and disease. *FEBS J.* **2013**, *280*, 549-571.

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

(21) Nikolic-Hughes, I.; Rees, D. C.; Herschlag, D. Do electrostatic interactions with positively charged active site groups tighten the transition state for enzymatic phosphoryl transfer? *J. Am. Chem. Soc.* **2004**, *126*, 11814-11819.

(22) Bone, R.; Frank, L.; Springer, J. P.; Atack, J. R. Structural studies of metal binding by inositol monophosphatase: evidence for two-metal ion catalysis. *Biochemistry* **1994**, *33*, 9468-9476.

(23) Choe, J.-Y.; Fromm, H. J.; Honzatko, R. B. Crystal structures of fructose 1,6-bisphosphatase: mechanism of catalysis and allosteric inhibition revealed in product complexes. *Biochemistry* **2000**, *39*, 8565-8574.

(24) Kamerlin, S. C. L.; Wilkie, J. The role of metal ions in phosphate ester hydrolysis. *Org. Biomol. Chem.* **2007**, *5*, 2098-2108.

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

(26) Ban, H.; Gavrilyuk, J.; Barbas, C. F. Tyrosine Bioconjugation through Aqueous Ene-Type Reactions: A Click-Like Reaction for Tyrosine. *J. Am. Chem. Soc.* **2010**, *132*, 1523-1525.

(27) Geoghegan, K. F.; Stroh, J. G. Site-Directed Conjugation of Nonpeptide Groups to Peptides and Proteins Via Periodate-Oxidation of a 2-Amino Alcohol - Application to Modification at N-Terminal Serine. *Bioconjugate Chem.* **1992**, *3*, 138-146.

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

(29) 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.

(30) Sachdeva, A.; Chandra, M.; Chandrasekar, J.; Silverman, S. K. Covalent Tagging of Phosphorylated Peptides by Phosphate-Specific Deoxyribozymes. *ChemBioChem* **2012**, *13*, 654-657.

(31) Li, B.; Sher, D.; Kelly, L.; Shi, Y.; Huang, K.; Knerr, P. J.; Joewono, I.; Rusch, D.; Chisholm, S. W.; van der Donk, W. A. Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 10430-10435.

(32) Lad, C.; Williams, N. H.; Wolfenden, R. The rate of hydrolysis of phosphomonoester dianions and the exceptional catalytic proficiencies of protein and inositol phosphatases. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5607-5610.

(33) Schroeder, G. K.; Lad, C.; Wyman, P.; Williams, N. H.; Wolfenden, R. The time required for water attack at the phosphorus atom of simple phosphodiesters and of DNA. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4052-4055.

(34) Schlosser, K.; Li, Y. Biologically inspired synthetic enzymes made from DNA. *Chem. Biol.* **2009**, *16*, 311-322.

(35) Lassila, J. K.; Zalatan, J. G.; Herschlag, D. Biological phosphoryl-transfer reactions: understanding mechanism and catalysis. *Annu. Rev. Biochem.* **2011**, *80*, 669-702.

(36) Nowakowski, J.; Shim, P. J.; Prasad, G. S.; Stout, C. D.; Joyce, G. F. Crystal structure of an 82-nucleotide RNA-DNA complex formed by the 10-23 DNA enzyme. *Nat. Struct. Biol.* **1999**, *6*, 151-156.

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

(38) 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.

(39) 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.

(40) Wang, Y.; Silverman, S. K. Characterization of Deoxyribozymes That Synthesize Branched RNA. *Biochemistry* **2003**, *42*, 15252-15263.

(41) Kost, D. M.; Gerdt, J. P.; Pradeepkumar, P. I.; Silverman, S. K. Controlling regioselectivity and site-selectivity in RNA ligation by $Zn^{2+}$-dependent deoxyribozymes that use 2',3'-cyclic phosphate RNA substrates. *Org. Biomol. Chem.* **2008**, *6*, 4391-4398.

(42) Bair, J. S.; Palchaudhuri, R.; Hergenrother, P. J. Chemistry and biology of deoxynyboquinone, a potent inducer of cancer cell death. *J. Am. Chem. Soc.* **2010**, *132*, 5469-5478.

(43) Levengood, M. R.; Kerwood, C. C.; Chatterjee, C.; van der Donk, W. A. Investigation of the substrate specificity of lacticin 481 synthetase by using nonproteinogenic amino acids. *ChemBioChem* **2009**, *10*, 911-919.

(44) Shevchenko, A.; Tomas, H.; Havlis, J.; Olsen, J. V.; Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **2006**, *1*, 2856-2860.

# Chapter 3: Efforts to Identify Sequence-Selective Phosphatase Deoxyribozymes

## 3.1 Introduction

Sequence-selectivity is a key feature and a major goal of biocatalysis. Protein engineering is used to achieve this important objective of specificity/selectivity when nature does not have a preexisting protein enzyme as the starting point.[1] Engineering protein to accomplish selectivity involves either a rational protein design or a directed evolution approach.[2] Prior knowledge of the protein structure, including protein sequence homology, usually serves as an excellent starting point for protein engineering. For rational design, structural information of protein is exploited to make desirable mutations and changes. Site-directed mutagenesis is the most important tool used in rational protein design.[3] A primary disadvantage of rational design is that unavailability of structural data makes progress with valuable mutations almost impossible. In contrast, directed evolution obviates the need for detailed understanding of the protein structure. Several random mutations are introduced and the proteins with the desirable traits are selected, isolated, and investigated. DNA shuffling is a commonly used technique to introduce semi-random mutations in the protein.[4] Shuffling is widely considered as a human mimicry of the natural protein evolution process. This method involves preparation of recombinant variants of related homologous families. Other techniques such as staggered extension process (StEP), assembly PCR or synthetic shuffling, incremental truncation for the creation of hybrid enzymes (ITCHY), and non-homologous random recombination (NRR) exist for genetic recombination in protein evolution.[5] Moreover, better techniques have been developed for introducing degenerate codons in the directed

evolution processes.[6] With continuous advancements, protein engineering continues to remain a powerful field for the development of sequence-selective catalysts.

However, several drawbacks exist with directed evolution (protein engineering). Firstly, proteins have long sequences. Only a limited set of variants can be created from a pre-existing starting point while maintaining the fundamental function of the initial enzyme. Secondly, selecting or screening for winners with desirable traits can be fairly tedious. Certain activities may require sophisticated robotics to screen through a huge number of variants. Not every activity can have a straightforward high-throughput screening readout. Separately, time intensive computational or bioinformatics analysis may be required to rationalize success of winners for subsequent evolution and mutations. Although protein engineering has achieved substantial success, practical considerations still limit the success of protein engineering to achieve sequence selectivity.

In this chapter, efforts to identify DNA catalysts with sequence selectivity are elaborated. With phosphatase activity, we wish to expand our examination of peptide substrate sequences, seeking catalysts that have the selectivity to discriminate among varying phosphopeptides. Classical protein-tyrosine phosphatases have similar sequence amino acid preference for dephosphorylation of $Y^P$ (acid over basic peptides).[7] Protein enzymes like PTP1B are also highly proficient and can dephosphorylate majority of the $Y^P$ substrates in the cytosol.[7] Recombinant expression of protein-tyrosine phosphatases cannot achieve specific targeted catalysis. Tools like phosphatase deoxyribozymes can potentially address this problem. Three independent approaches to identify sequence-selective catalysis in the context of phosphotyrosine phosphatase activity are discussed

here. A robust in vitro selection method to directly identify functional sequence-selective DNA catalysts starting from a library of $10^{14}$ molecules would be a milestone achievement in the biocatalytic endeavor, which will have the potential to strongly complement the protein engineering approach.

### 3.1.1 Utility of Sequence Selective Enzymes in Biotechnology

Engineering specificity and selectivity in proteases is of high importance in biotechnology (See Table 3.1). Proteases are often used in analytical and therapeutic applications.[8] Directed evolution is used to allow identification of protease variants from randomized libraries. For instance, trypsin cleaves arginine and lysine residues at the C-terminal portion.[9] Using protein engineering trypsin was converted to chymotrypsin to cleave hydrophobic sites instead of the positively charged side chains.[10-11] Native chymotrypsin found in pancreatic juices performs amide bond cleavage to the C-terminal portion of tyrosine, tryptophan and phenylalanine.[12] Lysyl endopeptidase isolated from *Achromobacter lyticus* has a preference for lysine (positively charged residue).[13] Glutamyl endopeptidase catalyzes protease activity at the C-terminus of glutamate or aspartate (negatively charged amino acid).[14] Although, few successes have been reported in obtaining specificity or selectivity from an initially different starting point, it is difficult to extensively fine-tune the activity without compromising the overall catalytic activity of the protease.[15]

**Table 3.1.** Proteases are listed with their sequence selectivity. Except Asp-N, all proteases shown here cleave at the C-terminal end of the specific amino acid residues.

| Protease | Residues | Notes |
|---|---|---|
| Trypsin | Lysine, Arginine | Positively charged |
| Chymotrypsin | Tyrosine, Tryptophan, Phenyalanine | Aromatic |
| Lys-C | Lysine | Postively charged |
| Glu-C | Glutamate, Aspartate | Negatively charged |
| Asp-N | Aspartate, Cysteine | N-terminal cleavege |

Protein engineering has also been used for applications in organic synthesis as monooxygenases, ketoreductases, lipases or aldolases providing improved enantioselectivity and specificity.[16] A good (catalytic active) starting point needs to be met with directed evolution, including negative selections for improved selectivity without losing initial activity. In vitro selection processes that directly help achieve identification of both activity and selectivity from an unknown starting point would be an important advancement in biotechnological research.

### 3.1.2 Current Approaches for Achieving Selectivity in Nucleic Acid Catalysts

Several research efforts have been devoted to establish selectivity with artificial nucleic acid catalysts. Negative selection is a process that is performed to remove DNA sequences that do not meet a specified selection criterion. Negative selections have been successfully used to identify selective catalytic DNA biosensors.[17-18,19] Some examples include selective $Co^{2+}$-, $Na^{+}$-, $Pb^{2+}$-, and uranyl-dependent DNA catalysts developed by Yi Lu and coworkers. Engineered molecular (ribo)switches have also been identified by negative selections.[20] In spite of these documented successes, negative selection is not likely to be effective. Low or weak negative selection process allows several sequences with undesirable properties to survive, which eventually get amplified over and over.[21]

Developing a stringent negative selection procedure against phosphopeptide sequences (for selective phosphatase activity) is a not promising method to achieve sequence selectivity.

## 3.2 Results and Discussion

Three independent approaches were tested for identification of DNA catalysts with sequence-selective phosphatase activity. A) Phosphopeptides with highly functionalized physical/chemical properties as the selection substrates. B) Apply selection pressure by spiking the reaction solution with free phosphopeptides (unrelated physical/chemical properties) as competitive ligands. C) Reverse complement inhibition to remove the undesirable sequence-general deoxyribozymes.

### 3.2.1 Highly Functionalized Substrates with Varying Physical/Chemical Properties

#### 3.2.1.1 In vitro selection experiments

New in vitro selection projects were initiated to identify sequence-selective DNA catalysts with phosphatase activity. The rationale was to use highly charged (positive and negative) and aromatic phosphopeptides as substrates for these new selections. Similar to natural and engineered proteases that function specifically with certain amino acid side chain residues, we hypothesized that phosphatase deoxyribozymes could be acquired with similar discriminatory properties. If highly functionalized peptidic substrates with varying chemical and physical attributes are provided, the odds of encouraging sequence selectivity increases. For positively charged peptide, we decided to use CAKKY$^{P}$KKA, where Y$^{P}$ is a phosphotyrosine. Similarly for negatively charged and aromatic phosphopeptides, we used CAEEY$^{P}$EEA and CAFFY$^{P}$FFA respectively.

We also wished to explore phosphopeptide substrates of biological relevance. For that purpose, we utilized doubly phosphorylated fragment (8-18 residue) from Cyclic-

Dependent Kinase 2 (CDK2).[22] The kinase activity of CDK2 is dependent on the phosphorylation state of this fragment. Two positions, T14 and Y15, need to be in the dephosphorylated state for CDK2 to be enzymatically active. Even mutation of T14 and Y16 leads to increase in kinase activity of the CDK2.[22] Hence the fragment containing T14 and Y15, CEKIGEGT$^{P}$Y$^{P}$GVV with an extra cysteine handle at the N terminus, was used in our selection efforts. In addition, two phosphotyrosine-containing sequences from Insulin Receptor 1 and 2 (IRS1 and IRS2, see uniprot.org taxonomy ID 9606) were also utilized. The sequences were CIEEY$^{P}$TEM (558-604 from IRS1) and CEGWY$^{P}$RAL (133-139 from IRS2), both with an extra cysteine handle at the N terminus for conjugation to the DNA anchor. Each peptide was anchored as a disulfide through its N-terminal Cys residue via a hexa(ethylene glycol) (HEG) tether to a DNA anchor oligonucleotide, which interacted with a fixed-sequence DNA segment by Watson–Crick base pairs.

Contamination with generic enzymes is a major concern for selections efforts involving sequence selectivity. For instance phosphatases discovered in chapter 2 (14WM9, 14M27) functioned promiscuously with a variety of phosphopeptides (Figure 3.1).[23]

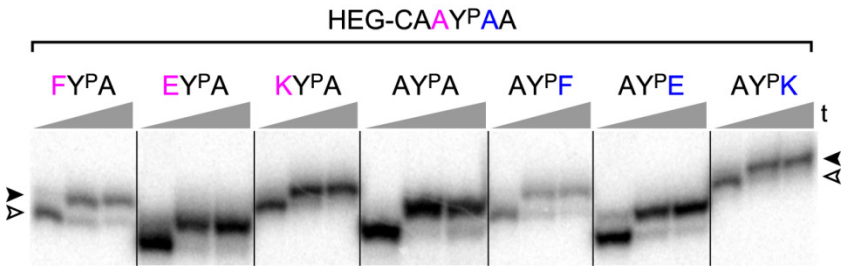


**Figure 3.1.** Assays of 14WM9 with various hexapeptide substrate sequences in which one amino acid adjacent to the Y$^{P}$ was altered, under selection conditions (figure from reference 23).

If independent selections were performed with different peptidic substrates, crossover of generic enzymes among every of these independent selections is a possibility. To prevent this contamination, new pools were designed and synthesized with different fixed constant binding arms flanking the random nucleotide region. This necessitates the use of different primers for every selection during polymerase chain reaction (PCR) amplification. However, crossover or contaminations among these selections are prevented (see Materials and Methods for details on pool, primers, splints and DNA anchors).

The strategy to identify phosphatase deoxyribozymes was similar to the previous efforts described in chapter 2 (Figure 3.2). Phosphotyrosine-containing peptide substrate was covalently appended to a DNA anchor oligonucleotide (Figure 3.3), which was itself ligated to the deoxyribozyme pool.

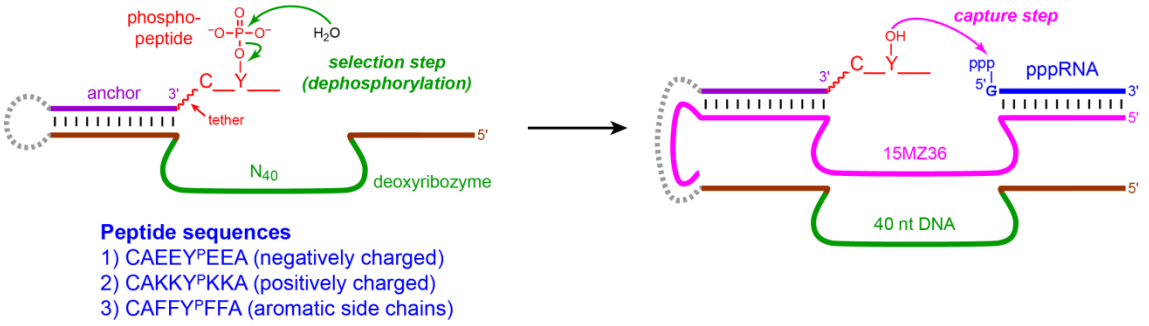


**Figure 3.2.** In vitro selection strategy is similar to that described in chapter 2 with a few key modifications. Each final deoxyribozyme has 40 particular nucleotides in its originally random $N_{40}$ region. For each selection, different peptide was used as the substrate. The peptides are highly functionalized to have negative, positive or aromatic properties.

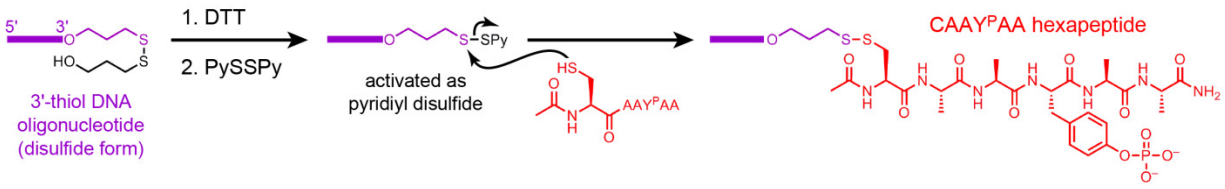


**Figure 3.3.** Conjugation of DNA anchor to the peptide. N-terminal cysteine of the peptide was conjugated to 3'-thiol oligonucleotide by dipyridyl disulfide activation. A model peptide ($CAAY^{P}AA$) is shown as an example. Peptide sequences were varied according to the selection

When the capture enzyme 15MZ36 was tested for all different peptides with tyrosine-OH, only slight variability in capture yield was observed. However, for peptide CAFFYFFA, the capture yield was insufficient to be useful for in vitro selection (<15%). All other peptides were efficiently captured by 15MZ36 (Figure 3.4). Selections were named FT1, FV1, FW1, FX1 and FY1 for E-rich, K-rich, CDK2, IRS1, and IRS2 phosphotyrosine-containing peptidic fragments.

After each key selection step that used incubation conditions of 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, active DNA catalyst sequences were "captured" with the aid of the 15MZ36 deoxyribozyme that selectively uses dephosphorylated tyrosine ($Y^{OH}$) as the nucleophile to attack 5′-triphosphorylated RNA, resulting in a PAGE shift for the corresponding deoxyribozyme sequence.[23] The active DNA sequences were amplified by PCR and joined by T4 DNA ligase to the DNA-anchored peptide substrate for entering the next selection round. When the DNA pool reached considerable activity, individual deoxyribozymes were cloned and characterized.

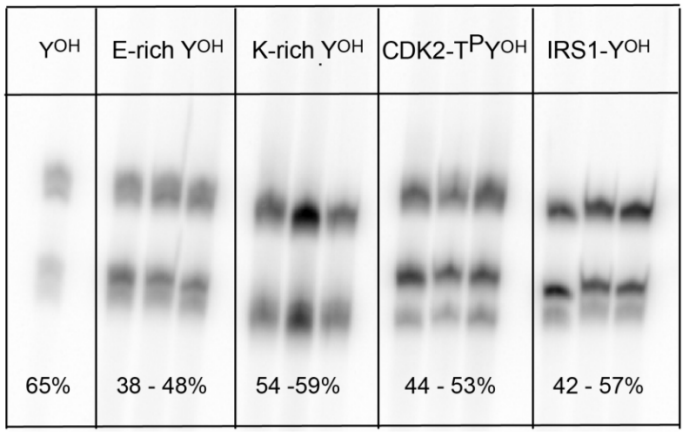


**Figure 3.4.** Capture test with 15MZ36. Substantial capture yield was observed for all peptides. The capture method involved selective chelation of zinc (see chapter 2 and material and methods for more information). For peptide fragment from IRS2, 30% yield was observed (data not shown). PAGE image illustrating 15MZ36-catalyzed capture for DNA-anchored CAAY$^{OH}$AA, CAEEY$^{OH}$EEA, CAKKY$^{OH}$KKA, CEKIGETPY$^{OH}$GVV and CIEEY$^{OH}$TEM (left to right). Incubation conditions: 70 mM HEPES, pH 7.5, 20 (40 and 60) mM $MnCl_2$, 40 mM $MgCl_2$, 100 mM NaCl, 37 °C. The capture yield is depicted at the bottom.

### 3.2.1.2 Cloning and sequences of the selection

After several rounds of selection, activity was observed for at round 8 for FT1, round 10 FV1, and round 9 for FX1 (Figure 3.5). Individual sequences ere cloned and characterized. For FW1 and FY1, no activity was observed at round 14 and hence they were discontinued.

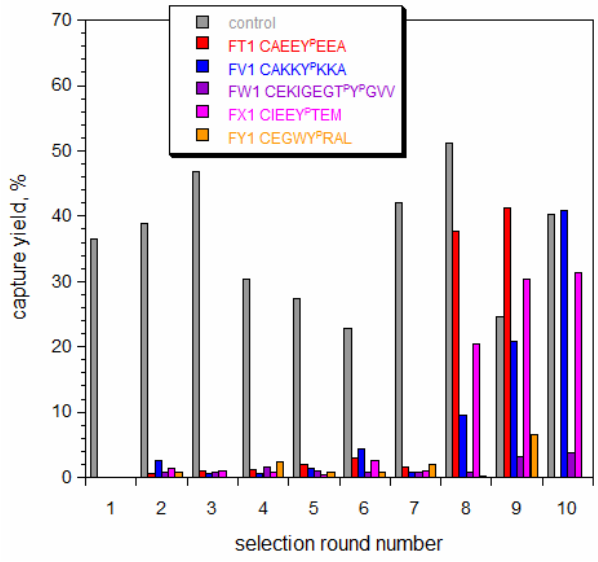


**Figure 3.5.** Progression of the in vitro selection experiment. In each round, "control" refers to the yield for the 15MZ36-catalyzed capture reaction using the DNA-anchored dephosphorylated CEGWY$^{OH}$RAL substrate. Selection FT1, FV1 and FX1 were cloned at rounds 8, 10 and 9 respectively.

Active sequences from FT1, FV1 and FX1 were aligned with phosphatases from earlier selections (14WM9 and 14WM27). If contamination were present, a strong similarity in sequence alignment would be observed in the $N_{40}$ nucleotide region. Although a small portion of the 3'-end was conserved for sequences from both FT1 and FV1 selections with the 14WM9 and 14WM27 selections, no strong homology was otherwise observed. This was strongly indicative that none of the sequences were contamination products or relatives of the previous WM selections. FX1 selections shared no conserved region with WM selections at all. When all the sequences from FT1, FV1 and FX1 were aligned together, no major homology was observed except for some conserved portion in the 3'-end among FT1 and FV1 (Figure 3.6).

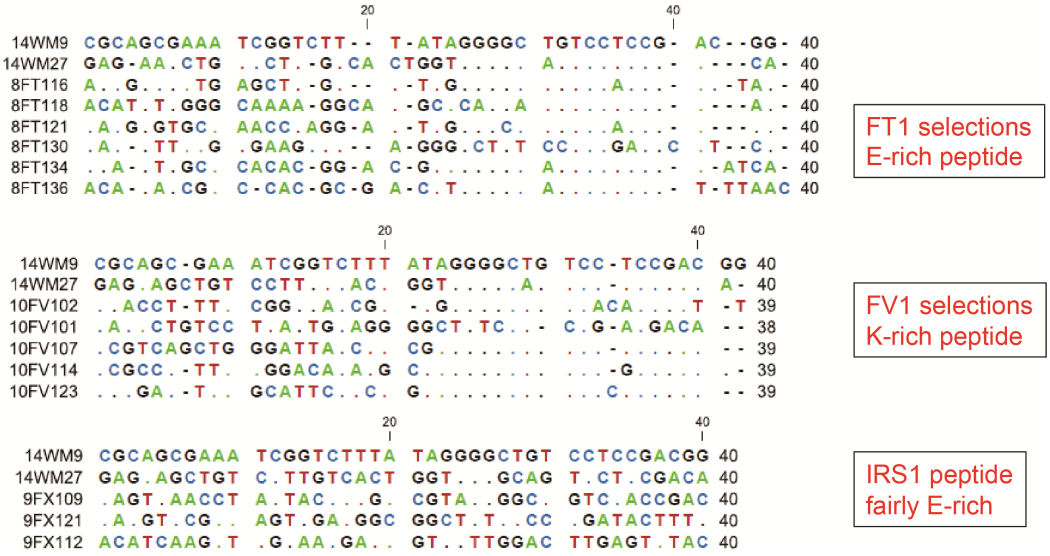


**Figure 3.6.** Sequences of the new phosphatase deoxyribozymes. Only the enzyme region ($N_{40}$) sequences are shown. For single-turnover intramolecular reactions, each enzyme region was surrounded by additional DNA nucleotides (refer to Chapter 2 and Table 3.4). A) FT1 sequences with E-rich phosphopeptide were aligned with WM sequences show some conservation at the 3' end. B) FV1 sequences with E-rich peptide show conservation at the 3' end as well C) FX1 selections with the peptide CIEEY$^{P}$TEM shows no homology.

### 3.2.1.3 Single-Turnover dephosphorylation assay for selectivity

FT1, FV1 and FX1 enzymes were tested for single-turnover dephosphorylation in the presence of $Mg^{2+}$, $Zn^{2+}$ and $Mn^{2+}$ with the respective DNA-anchored phosphopeptide substrates for sequence selectivity. First, the FT1, FV1 and FX1 deoxyribozymes were tested for catalytic activity with DNA-anchored CAKKY$^{P}$KKA. Even though selection FT1 and FX1 were performed with E-rich phosphopeptides, substantial dephosphorylation activity was observed with all enzymes for the K-rich phosphotyrosine substrate suggesting no sequence-selectivity with a K-rich substrate (Figure 3.7A). Similarly when all the FT1, FV1 and FX1 enzymes when tested with the DNA-anchored CAEEY$^{P}$EEA substrate (negatively charged phosphopeptide), no selectivity was observed (Figure 3.7B).

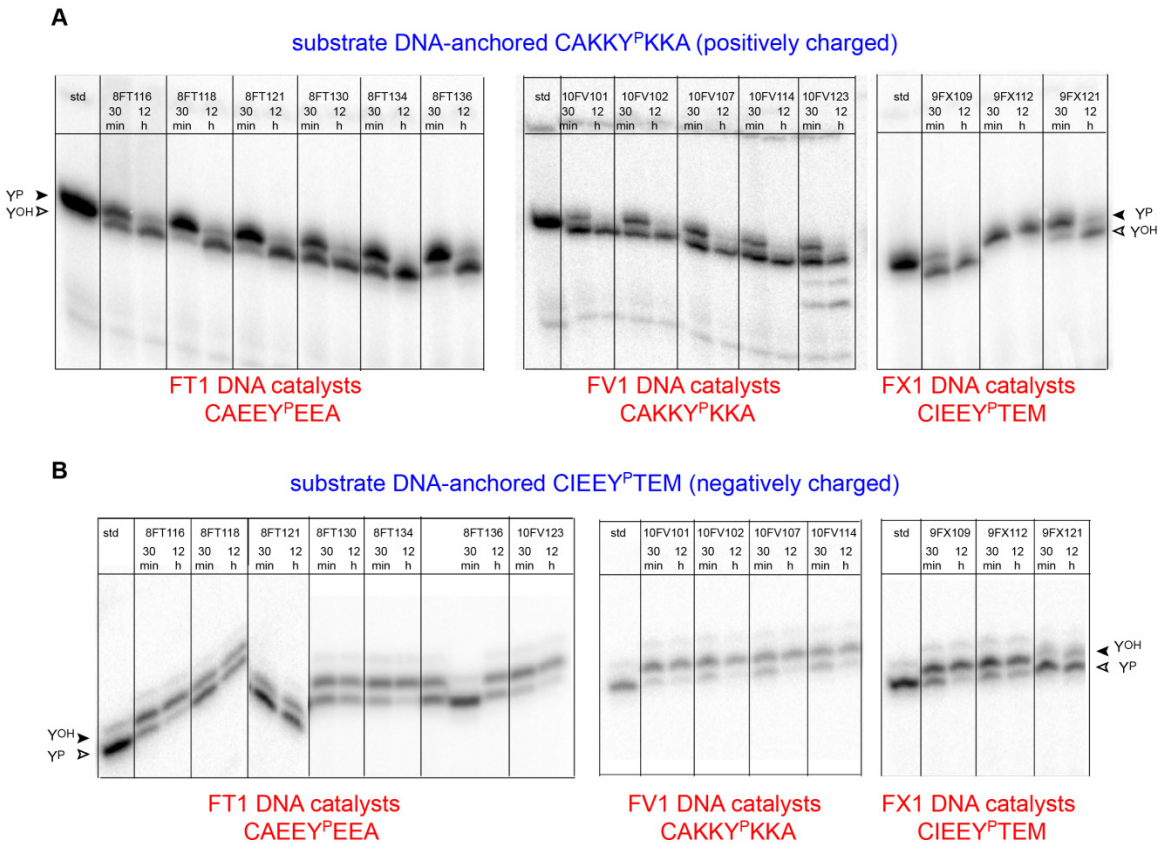


**Figure 3.7.** DNA catalysts from selections FT1, FV1 and FX1 were tested for sequence selectivity with DNA-anchored K-rich phosphopeptide and E-rich phosphopeptide. A) Top panel shows the single-turnover test with DNA-anchored CAKKY$^{P}$KKA. B) Lower panel shows the single-turnover assay with DNA-anchored CIEEY$^{P}$TEM. Note, same results were observed with CAEEY$^{P}$EEA (data not shown). FT1, FV1 and FX1 enzymes has different DNA anchor sequences (see Materials and Methods).

This was rather surprising because we expected some degree of sequence selectivity. DNA is a negatively charged biopolymer. It is reasonable to expect some preference for the FV1 enzymes to have preference for the positively charged K-rich peptide over the E-rich counterpart. Nonetheless, empirical data suggests that no sequence selectivity was achieved.

### 3.2.2 Competitive Peptides for Sequence Selectivity[2]

Although we had reasonable expectations for sequence-selectivity from the first approach, empirical data suggested that the new phosphatase deoxyribozymes were sequence general. Even so, selection pressure was not actively applied in the earlier effort to encourage sequence selectivity. There was no major incentive for DNA catalysts to recognize any specific sequence. To address this shortcoming in the previous approach, new selection efforts were initiated.

In this approach, DNA-anchored phosphopeptide that is ligated to the pool is considered the actual substrate. In the selection buffer, free phosphotyrosine-containing peptides were unrelated physical and chemical properties were supplemented. These free phosphopeptides are expected to compete with the actual DNA-anchored phosphopeptide substrate. Only DNA catalysts that ignore the presence of free phosphopeptides and recognize the covalently tethered phosphopeptide for catalysis are expected to survive the selections (Figure 3.8).

[2] Prof. Silverman and I conceived and planned the project. University of Illinois undergraduate student Natalia Recko performed the in vitro selections with competitive peptides under my supervision.

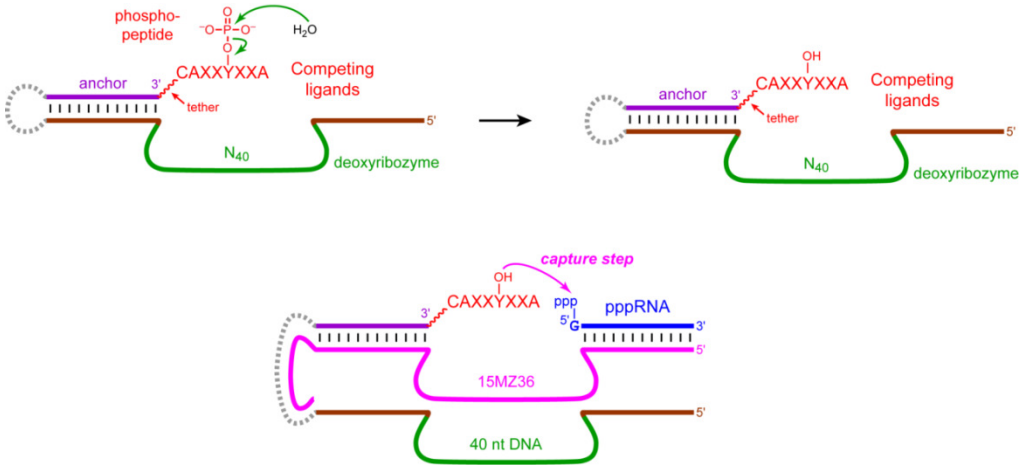


**Figure 3.8.** In vitro selection in the presence of competitive ligands (or peptide) as an active selection pressure to encourage selectivity. DNA-anchored phosphopeptide is the actual substrate. Competitive phosphopeptides of unrelated properties are supplemented in the reaction solution. Only sequence-selective deoxyribozymes that are selective for the actual substrate and are not outcompeted by unrelated peptides will survive the selection.

### 3.2.2.1 Testing selection conditions for competitive ligands

To apply appreciable selection pressure, free peptide concentration needs to be substantially high to compete with the actual DNA-anchored phosphopeptide substrate. However, if the free peptide concentration is exceedingly high, non-specific inhibitory (or aggregation) effects will be observed leading to the failure of the in vitro selection efforts. To empirically determine a reasonable compromise in the concentration, 14WM9 phosphatase deoxyribozyme was initially tested for catalytic activity in the presence of competitive phosphopeptides.

DNA-anchored phosphopeptide (CAAY$^{P}$AA), was tested with 14WM9 in 70 mM HEPES, pH 7.5, X mM $ZnCl_2$ (where X is 1 mM, 1.5 mM, 2 mM and 3 mM), 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h in the presence of 2 mM free peptide peptides. The free peptides were systematically changed from A-rich, E-rich, K-rich and F-rich phosphopeptides; i.e. AAAY$^{P}$AAA, AEEY$^{P}$EEA, AKKY$^{P}$KKA and AFFY$^{P}$FFA respectively (Figure 3.9). The F-rich peptide had the highest inhibitory effect

on DNA catalysis. This can be attributed to the hydrophobic property of AFFY$^{P}$FFA. To minimize any such non-specific inhibition, the concentration of AFFY$^{P}$FFA peptide was reduced from 2 mM to 0.5 mM.

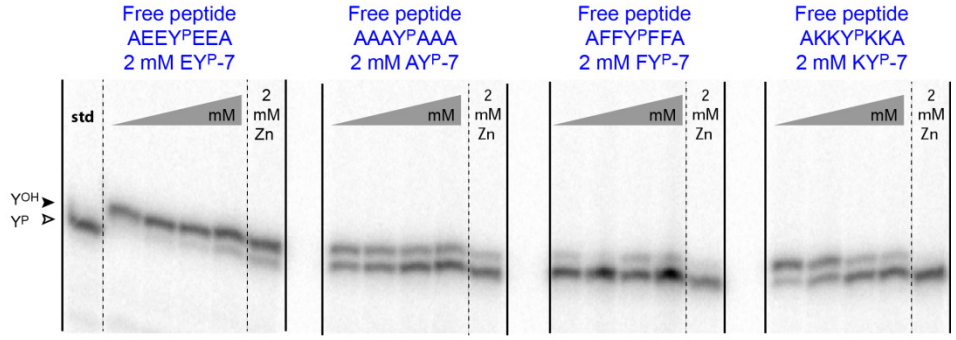


**Figure 3.9.** Single-turnover assays of the 14WM9 deoxyribozyme that dephosphorylates phosphotyrosine and phosphoserine side chains in the presence of competitive peptides. Assay conditions: 1 µM 14WM9, 12.5 nM 5′-$^{32}$P-radiolabeled DNA-anchored hexapeptide, 70 mM HEPES, pH 7.5, X mM $ZnCl_2$(where X is 1, 1.5, 2, 3), 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C. In the assays with $Zn^{2+}$ alone, the $Mn^{2+}$ and $Mg^{2+}$ were omitted. Timepoints for CAAY$^{P}$AA substrate is 12 h. Free peptides were supplemented in the selection buffer at the respective concentration.

### 3.2.2.2 In vitro selections with competitive peptides

Selections HM1, HN1, HP1 and HQ1 were initiated as depicted in Table 3.2. DNA-anchored peptides were prepared as mentioned in the earlier section via an N terminus cysteine handle. Peptides that lacking cysteines AXXY$^{P}$XXA, where X = F, E, A, and K, were synthesized and purified by HPLC. These HPLC-purified peptides were used as competitive ligands. Based on the earlier results (Figure 3.9), 1mM of E-rich, K-rich or A-rich phosphopeptides was a reasonable compromise between active selection pressure and disruptive non-specific inhibition. For F-rich peptide, a lower concentration of 0.5 mM was used because of the higher hydrophobic inhibition. The selection conditions were 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$ (2 mM $ZnCl_2$ for F-rich peptide), 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h supplemented with

the respective HPLC-purified peptide at the appropriate concentration (0.5 or 1 mM). After 15 rounds of selection, no activity was observed and these selections were discontinued.

**Table 3.2** Selection HN1, HP1, HQ1, HM1 were designed with the following DNA-anchored phosphopeptides and HPLC-purified competitive peptides supplemented at the appropriate concentration. All peptides were spiked upto 1 mM in the selection step (except F-rich peptide, 0.5 mM)

| Selection | DNA-anchored Substrate | Competitive Peptide |
|---|---|---|
| HN1 | CAKKY$^{P}$KKA | AAAY$^{P}$AAA (1mM) |
| HP1 | CAKKY$^{P}$KKA | AEEY$^{P}$EEA ( 1 mM) |
| HQ1 | CAKKY$^{P}$KKA | AFFY$^{P}$FFA (0.5 mM) |
| HM1 | CAAAY$^{P}$AAA | AKKY$^{P}$KKA (1 mM) |

### 3.2.3 Reverse Complement Inhibition to Remove Undesirable Sequences[3]

Because the two earlier efforts to identify sequence-selective deoxyribozymes were unsuccessful, a third approach was designed (Figure 3.10).

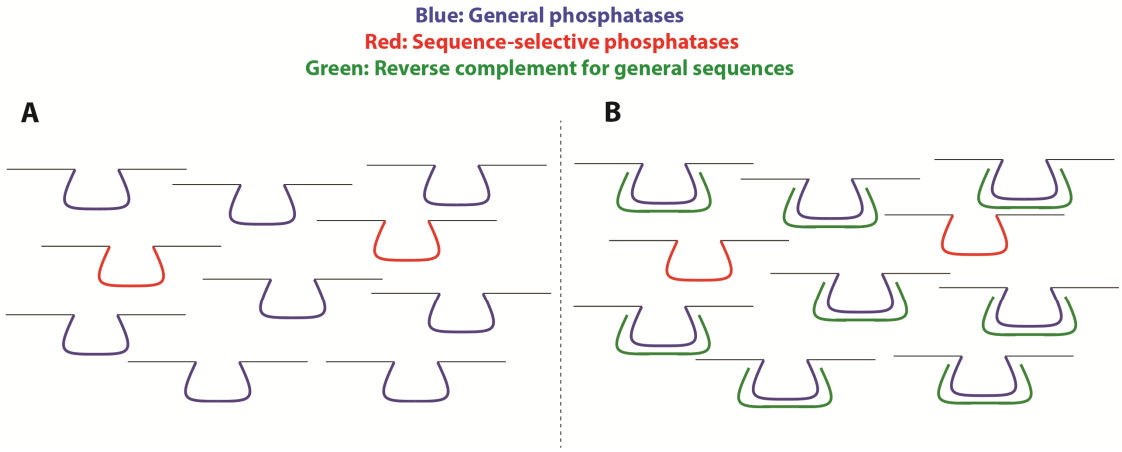


**Figure 3.10.** Reverse complement selection involves inhibition of undesirable sequences A) Initial phosphatase pool contains sequence-general enzymes (blue) and sequence-selective enzymes (red). B) Sequence-general enzymes are inhibited as a double-stranded DNA by Watson-Crick base pairing.

This approach relies on inactivation of undesirable sequence-general single-stranded deoxyribozymes using the complementary reverse strand. Inactivation of

[3] Prof. Silverman and I conceived and planned the project for reverse complement inhibition.

undesirable sequences by Watson-Crick base pairing is expected to be a powerful alternative to performing a negative selection.

#### 3.2.3.1 In vitro selection strategy with reverse complement method

A procedure was carefully devised to identify sequence-selective phosphatases with the reverse complement technique. Selections would be performed with the desirable substrate for a certain amount of rounds. A late-stage selection pressure would be applied when the pool is already enriched with phosphatase sequences, including both sequence-general and sequence-selective DNA catalysts. At this point, the selection would be split into two routes (Figure 3.11). The first portion of the pool would be passed through a selection process involving the undesirable substrate (red pathway). After the selection and capture, PCR amplification would be performed. However, instead of isolating the sense strand, the complementary (anti-sense) strand would be purified. The second portion of the pool will be passed through a selection process involving the desirable substrate (shown in green). During the selection step with the desirable substrate, the purified complementary (anti-sense) strands will be pooled together. This leads to specific inactivation of the single-stranded DNA that are functional with the undesirable substrate from the red pathway. Importantly, sequences that work selectively with the desirable substrate will be unaffected and survive the selection round. In summary, the overall process will lead to suppression of the sequence-general catalyst enabling the enrichment of the sequence-selective DNA catalysts. These selective DNA catalysts are expected to function specifically with the desirable used in the green pathway.

This selection process, if successful, gives a great degree of control on the desired selectivity. The desired and the undesired substrates can be fined tuned for a specific downstream biological application and is not limited to just phosphatase activity.

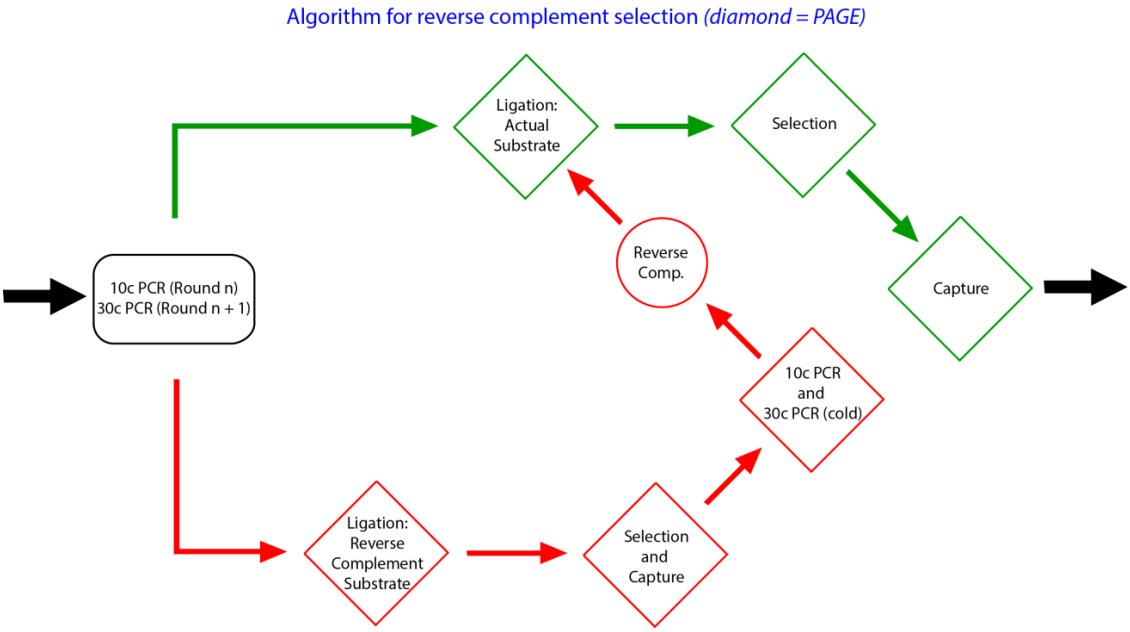


**Figure 3.11.** Reverse complement inhibition is a selection pressure projected to strongly suppress the undesirable sequences. A) Green curve involves the steps with the desirable substrate. B) Red pathway involves a selection step with an undesirable substrate. The reverse complement is isolated after the PCR enrichment and used to inhibit the general sequences during the selection step with the actual substrate.

### 3.2.3.2 Validation of the strategy

It is important to validate the reverse complement inhibition strategy before initiating the new selection process. For that purpose, positive and negative controls were designed and performed with a catalytically active phosphatase pool (round 14WM selections mentioned in Chapter 2).

Three test cases were investigated. A) No reverse complement, B) Random reverse complement, C) Exact reverse complement to the pool (see Materials and Methods for details). Selection step was performed with 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, in the

presence of respective reverse complement (anti-sense) sequences. The selection sample was then purified by an 8% denaturing PAGE to remove these additional reverse complement components. The solution was brought back to 20 μL with 50 pmol of 15MZ36, 100 pmol of 5'-triphosphorylated RNA in 70 mM HEPES, pH 7.0, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h. Only catalytically active DNA sequences should be captured. Results indicated that catalytic activity of the active pool was unaffected in the presence of random reverse complement. However, when the exact reverse complement was utilized, catalytic activity was suppressed because of the formation of a double-stranded DNA indicating that catalytic suppression was highly specific (Figure 3.12). These controls were strongly indicative that this reverse complement inhibition strategy is a viable pathway to allow enrichment of sequence-selective DNA catalysts.

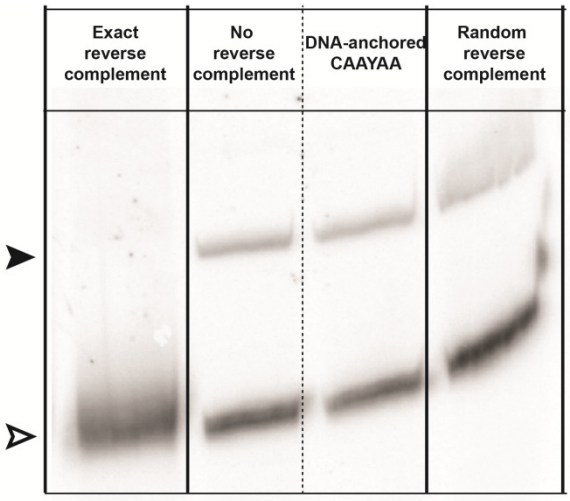


**Figure 3.12.** Assay for validation of the reverse complement technique with 14WM phosphatase pool. This assay ensures that reverse complement inhibition is specific. A) Lane 1 shows that exact reverse complement leads to complete phosphatase inhibition. B) No reverse complement was added to the phosphatase pool. C) Capture activity with $CAAY^{OH}AA$. D) Random reverse complement sequences led to negligible catalytic suppression (see Materials and Methods for details).

### 3.2.3.3 In vitro selection with the reverse complement strategy

Because we are concerned about non-specific inhibition of DNA sequence in the initial rounds by random reverse complement sequences, we theorized that the best strategy is to apply a late-stage selection pressure when phosphatase deoxyribozymes already dominate the pool. If the sequence-selective DNA sequences are in lower copy number compared to the general sequences, we still expect the general sequences to get inhibited or suppressed by the reverse complement approach. The suppression of general phosphatases eventually enables the desirable sequence-selective deoxyribozymes to get enriched. This is a strong feedback approach for removing unwanted sequences. New selections involving the reverse complement approach were redirected from the later rounds of WM, FT1 and FV1. The plans for reverse complement selections are depicted in Table 3.3.

Table 3.3. Selection MK1 to MS1. These selections were intiated from WM, FT1 and FV1 selections. The desirable, undesirable substrates, including the expected outcome are shown for each selection.

| Selection | Desirable substrate (green pathway) | Undesirabale substrate (red pathway) | Expected outcome |
|---|---|---|---|
| 13WM → 14MK1 | **CAAAY$^{P}$AAA** | **None** | **Sequence general** |
| 13WM → 14ML1 | **CAAAY$^{P}$AAA** | **Random** | **Sequence general** |
| 13WM → 14MM1 | **CAAAY$^{P}$AAA** | **Exact** | **No activity** |
| 13WM → 14MN1 | **CAAAY$^{P}$AAA** | **CAFFY$^{P}$FFA** | **A-rich vs. F-rich** |
| 7FT1 → 8MP1 | **CAEEY$^{P}$EEA** | **CAKKY$^{P}$KKA** | **E-rich vs. K-rich** |
| 7FT1 → 8MQ11 | **CAEEY$^{P}$EEA** | **CAFFY$^{P}$FFA** | **E-rich vs F-rich** |
| 7FV1 → 8MR1 | **CAKKY$^{P}$KKA** | **CAEEY$^{P}$EEA** | **K-rich vs. E-rich** |
| 7FV1 → 8MS1 | **CAKKY$^{P}$KKA** | **CAFFY$^{P}$FFA** | **K-rich vs. F-rich** |

These selections had properly designed controls MK1, ML1, MM1. The selection MK1 was continued from 13WM (active pool) without addition of any reverse complement. Selection ML1 had random reverse complements supplemented in the

selection process. Because random sequences are unrelated to the pool, no inhibitory activity by Watson-Crick base pairing is expected. For MM1 selections, the reverse complement from the same selection round was used for inhibition. Completely suppression by Watson-Crick base pairing is expected here between the sense and anti-sense strand. Other selections were designed for specifically controlled desirable selectivity (see Table 3.3). For instance, MN1 was designed with the peptide CAAAY$^{P}$AAA (in the green pathway) as the desirable substrate and CAFFY$^{P}$FFA (red pathway) as the undesirable substrate. The likely outcome would be DNA catalysts with preferential activity for A-rich peptide over the F-rich peptide. Similarly, MP1, MQ1, MR1 and MS1 were designed for varied selectivity outcomes (Table 3.3).

15MZ36 couldn't function properly with CAFFYFFA as a "capture" enzyme. However another nucleopeptide-linkage forming deoxyribozymes 8HJ102,[24] functioned with the F-rich peptide forming nucleopeptide linkage with Y$^{OH}$ in a F-rich peptide as the nucleophile (see material and methods). Capture enzyme 8HJ102 was used in pathways involving F-rich peptides.

The selections were pursued for 4 additional rounds. No activity was observed for the selections MN1 to MS1 even though the controls selections functioned normally as expected (Figure 3.13). Hence these selections were discontinued.

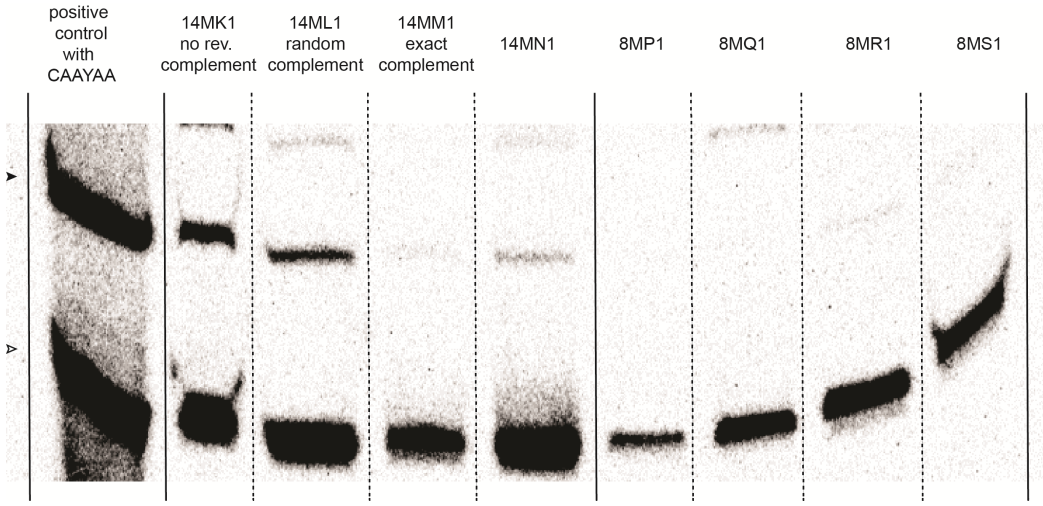


**Figure 3.13.** Capture step for selection with the reverse complement inhibition approach. Rounds 14MK1-14MN1 and 8MP1-8MS1 are depicted here. The left lane denotes the capture control with CAAY$^{OH}$AA. 14MK1 (no reverse complement) and 14ML1 (random reverse complement) control selections show similar activity. For 14ML1 with the exact reverse complement, catalytic suppression of activity is observed. All the other selections showed background capture with little or no activity. The lower band is the DNA-anchored peptide ligated to the pool. The top band denotes capture with 15MZ36 and a 5'-triphosphorylayted RNA/DNA chimera.

## 3.3 Summary

In summary, three different approaches were applied to enable identification of sequence-selective deoxyribozymes with phosphatase activity. The first approach utilized highly functionalized peptides (extreme physical or chemical properties) as the DNA-anchored substrates (FT1-FY1). They were highly charged (positive or negative) or aromatic. The hypothesis was that, DNA catalysts arising from such selections would have a natural propensity to function superficially with closely related phosphotyrosine-containing sequence. Unfortunately, the DNA catalysts identified from such selections were sequence general and functioned with any phosphotyrosine-containing sequence showing no noticeable discriminatory properties. In these efforts, DNA pools were synthesized with different primer-binding regions flanking the random nucleotide region. This was a pre-emptive measure to disallow general deoxyribozymes from contaminating

or crossing over to unrelated selections. However, no active pressure for selectivity was applied in this first effort to encourage selectivity.

In the second effort, competitive peptides were utilized as an active selection pressure to enable selectivity. DNA-anchored phosphopeptide was formally the actual substrate for phosphatase activity. The selection buffer was supplemented with millimolar quantity of free phosphopeptide. The free peptides $Y^P$-containing peptides had physical and chemical properties unrelated to the DNA-anchored substrate (selections HM1-HQ1). Only selective deoxyribozymes that are capable of ignoring the presence of free phosphopeptides, but function specifically with the DNA-anchored phosphopeptides should survive the selections. Concentration of the free peptide was carefully optimized to overcome the effective molarity of the DNA-anchored peptide, but to not have non-specific inhibition or aggregation on catalysis. However, these efforts failed to provide activity. We theorized that adding millimolar amounts of free phosphopeptides in solution are harsh conditions to enable DNA catalysis from the initially random pool, even though our preliminary design establishes that DNA phosphatases to be active under such conditions millimolar quantities of free peptide.

In the third effort, a rational reverse complement approach was utilized to identify sequence-selective DNA catalysts. This approach involved splitting the selection process into two pathways. One pathway involved the undesirable substrate while the other involved the desirable one. The pool ligated to undesirable substrate was passed through a selection step followed by the enrichment and capture of the active sequences by PCR amplification. The reverse complement (or the anti-sense) sequences are isolated. These sequences are added in the selection step of the desirable pathway (where the correct

substrate is ligated to the pool). Specific inactivation or suppression by Watson-Crick base pairing is expected for the DNA catalysts in the pool that function with undesirable substrate. Only selective sequences that merely function with the desirable substrate on the other hand are expected to survive the process and get enriched. Although controls were establish, validated and carried in parallel with the actual selections (MK1-MM1), no activity was observed for actual selections.

It is rather surprising that we have not been able to identify sequence-selective phosphatase deoxyribozymes. We expect DNA to show some preference for the adjacent peptide side chain residues. For instance, a DNA catalyst functioning with super positively charged phosphopeptide should get repelled by a super negatively charged peptidic substrate. Or DNA being a negatively charged molecule itself, should have preferential binding and catalysis for a positive charged (K-rich) phosphopeptide substrate. Likewise, we would expect a negatively charged biopolymer like DNA to have bias for positively charged phosphopeptides over the negatively charged peptides. In spite of having rational and well-controlled methods to improve enrichment of sequence-selective deoxyribozymes, we haven't been successful in such endeavors. On a promising note, DNA has shown its ability to be sequence-selective for other activities.[25] This is suggestive that although sequence selectivity has not been achieved with phosphatase activity, DNA is capable of such sequence-selective catalysis in other processes.

## 3.4 Materials and Methods

### 3.4.1 Oligonucleotides and In vitro Selection

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′-Triphosphorylated RNA oligonucleotides were prepared by in vitro transcription using synthetic DNA templates and T7 RNA polymerase.[26] 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 previously. The 5′-Triphosphorylated RNA/DNA chimera was synthesized as described in Walsh et al.[27] A different 5′-Triphosphorylated RNA was used for capture reaction with the enzyme 8HJ102. The synthesis and sequence of the 5′-triphosphorylated RNA used for 8HJ102 capture is described in Chu et al.[24]

### 3.4.2 Synthesis of Peptides and DNA-Anchored Peptides

Peptides were prepared by solid-phase synthesis on Fmoc Rink amide MBHA resin (AAPPTec). Each peptide was coupled to the DNA anchor oligonucleotide via the N-terminal cysteine side chain (linkage created by disulfide formation).

**3.4.2.1 Synthesis of peptides.** Peptides were prepared by solid-phase synthesis using Fmoc Rink amide MBHA resin, with *N,N,N′,N′*-tetramethyl-*O*-(7-azabenzotriazol-1-yl)uronium hexafluorophosphate (HATU) as coupling agent. Each synthesis was performed at 0.2 mmol scale, initiated using 308 mg of Rink amide resin with a loading capacity of 0.65 mmol/g (AAPPTec). All steps were monitored by ninhydrin test. For each coupling, 5 equivalents (1.0 mmol) of Fmoc-amino acid, 4.9 equivalents (373 mg, 0.98 mmol) of HATU, and 10 equivalents (350 μL, 2.0 mmol) of *N*,*N*-diisopropylethylamine (DIPEA) were mixed in 5 mL of anhydrous DMF. The coupling reaction was initiated by adding this mixture to the resin and agitating by bubbling with nitrogen for 60 min, followed by washing with DMF (3 × 10 mL). The N-terminus of the newly installed amino acid was deprotected by agitating the resin in 20% piperidine in

DMF (5 mL) under nitrogen for 5 min a total of three times, each time washing with DMF (3 × 5 mL). Peptides with N-terminal cysteine were capped via acetylation of their N-terminus by agitating the resin in 50 equivalents (924 μL, 10 mmol) of acetic anhydride and 25 equivalents of DIPEA (870 μL, 5 mmol) in 5 mL of DMF under nitrogen for 30 min. The peptide was cleaved from the solid support by stirring the resin in a separate vial with a solution containing 5 mL of trifluoroacetic acid (TFA), 125 μL of water, 125 μL of ethanedithiol, and 50 μL of triisopropylsilane for 90 min. The liquid solution was separated from the resin by filtration. This solution was dried on a rotary evaporator, providing a solid. To this material 15 mL of cold diethyl ether was added, and the peptide was obtained as a white solid that was filtered. Peptides with cysteine at the N terminus were used without HPLC purification (as crude substrates) for conjugating with the DNA anchor. All other peptides were purified by HPLC.

**3.4.2.2 Synthesis of DNA-anchored peptides.** Each peptide was coupled to the DNA anchor oligonucleotide via the N-terminal cysteine side chain (linkage created by disulfide formation). The DNA anchor oligonucleotide (see Table 3.4 for sequence information for each selection) ended with a HEG-$C_3$-SS-$C_3$-OH-3′, where HEG stands for hexaethylene glycol linker.[23] The HEG-3′-disulfide linker was introduced via standard solid-phase DNA synthesis and unmasked to a 3′-thiol by DTT treatment. A 50 μL sample containing 2-5 nmol of DNA anchor oligonucleotide in 50 mM HEPES, pH 7.5, and 50 mM DTT was incubated at 37 °C for 2 h. The reduced product was precipitated to remove excess DTT by addition of 50 μL of water, 10 μL of 3 M NaCl, and 300 μL of ethanol. The precipitated product (DNA-HEG-$C_3$-SH) was dissolved in 45 μL of water. 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-$C_3$-SSPy) was precipitated by addition of 50 µL of water, 10 µL of 3 M NaCl and 300 µL of ethanol and dissolved in 25 µL of water. Conjugation to the peptide was performed by adding 20 µL of 50 mM triethylammonium acetate, pH 7.0, and 5 µL of 40 mM peptide (200 nmol). The sample was incubated at 37 °C for 2 h, and the DNA-anchored peptide was purified by 20% PAGE. A typical yield was 2-3 nmol. See Table 3.5 for MALDI mass spectrometry values.

**Table 3.4**. Different pools were used for each selection. Sequences are written 5′ to 3′. X denotes a HEG spacer to stop Taq polymerase from extending further. This enables separation of the single-stranded DNA. Sequence Y stands for the 40-nucleotide portion of 15MZ36, CAAGGAGAGTTGTACAAGCTCGGGTCGTGTTCAAAGG-GATC

| oligonucleotide purpose | oligonucleotide sequence |
|---|---|
| *Selection FT1 details (with CAKKY$^{P}$KKA)* | |
| DNA-HEG- substrate | GGACTACCTTTATGCGTAT-HEG-$C_3$-SH |
| forward primer for selection | CGAACGAAAGCCTCCTTC |
| reverse primer for selection | $(AAC)_4$XCCATCAGGATCAGCT |
| random pool for selection | CGAACGAAAGCCTCCTTC-$N_{40}$-ATACGCATAAAGGTAGAGCTGATCCTGATGG |
| 15MZ36 capture enzyme | CCGTCGCCATCTCTTC-Y-ATACGCATAAAGGTAGTCCCCATCAGGATCAGC |
| splint for ligation step during selection [a] | ATAAGAAACGAGATATTCCTCCATCAGGATCAGCTATATCTCGTTTCTTAT |
| *Selection FV1 details (with CAEEY$^{P}$EEA)* | |
| DNA-HEG- substrate | GGAATATCTCGTTTCTTAT-HEG-$C_3$-SH |
| forward primer for selection | CGAATTAAGACTGAATTC |
| reverse primer for selection | $(AAC)_4$XCCATCAGGATCAGCT |
| random pool for selection | CGAATTAAGACTGAATTC-$N_{40}$-ATAAGAAACGAGATATAGCTGATCCTGATGG |
| 15MZ36 capture enzyme | CCGTCGCCATCTCTTC-Y-ATAAGAAACGAGATATTCCCCATCAGGATCAGC |
| splint for ligation step during selection [a] | ATACCAATCAAGCCATTCCTCCATCAGGATCAGCTATGGCTTGATTGGTAT |
| *Selection FW1 details with (CEKIGEGT$^{P}$Y$^{P}$GVV)* | |
| DNA-HEG- substrate | GGAATGGCTTGATTGGTAT-HEG-$C_3$-SH |
| forward primer for selection | CGAATTGAGTAAATATTC |
| reverse primer for selection | $(AAC)_4$XCCATCAGGATCAGCT |
| random pool for selection | CGAATTGAGTAAATATTC-$N_{40}$-ATACCAATCAAGCCATAGCTGATCCTGATGG |
| 15MZ36 capture enzyme | CCGTCGCCATCTCTTC-Y-ATACCAATCAAGCCATTCCCCATCAGGATCAGC |
| splint for ligation step during selection [a] | ATATGTCTTTCAATAGTCCTCCATCAGGATCAGCTCTATTGAAAGACATAT |
| *Selection FX1 details (with CIEEY$^{P}$TEM)* | |
| DNA-HEG- substrate | GGACTATTGAAAGACATAT-HEG-$C_3$-SH |
| forward primer for selection | CGAAATGATGGCTATTTC |
| reverse primer for selection | $(AAC)_4$XCCATCAGGATCAGCT |
| random pool for selection | CGAAATGATGGCTATTTC-$N_{40}$-ATATGTCTTTCAATAGAGCTGATCCTGATGG |
| 15MZ36 capture enzyme | CCGTCGCCATCTCTTC-Y-ATATGTCTTTCAATAGTCCCCATCAGGATCAGC |
| splint for ligation step during selection [a] | ATAATTAGTAACCTGATCCTCCATCAGGATCAGCTTCAGGTTACTAATTAT |
| *Selection FY1 details (with CEGWY$^{P}$RAL)* | |
| DNA-HEG- substrate | GGATCAGGTTACTAATTAT-HEG-$C_3$-SH |
| forward primer for selection | CGAAATAGATTATCATTC |
| reverse primer for selection | $(AAC)_4$XCCATCAGGATCAGCT |
| random pool for selection | CGAAATAGATTATCATTC-$N_{40}$-ATAATTAGTAACCTGAAGCTGATCCTGATGG |
| 15MZ36 capture enzyme | CCGTCGCCATCTCTTC-Y-ATAATTAGTAACCTGATCCCCATCAGGATCAGC |
| splint for ligation step during selection [a] | ATATCATCAGTGGCTCTCCTCCATCAGGATCAGCTGAGCCACTGATGATAT |
| *Selection HQ1 details* | |
| DNA-HEG- substrate | GGATCCTGGATACAAATAT-HEG-$C_3$-SH |
| forward primer for selection | CGAAGTATAAACCTGTTC |
| reverse primer for selection | $(AAC)_4$XCCATCAGGATCAGCT |
| random pool for selection | CGAAGTATAAACCTGTTC-$N_{40}$-ATATTTGTATCCAGGAAGCTGATCCTGATGG |
| 15MZ36 capture enzyme | CCGTCGCCATCTCTTC-Y-CCGTCGCCATCTCTTCCCATCAGGATCAGC |
| splint for ligation step during selection [a] | ATATTTGTATCCAGGATCCTCCATCAGGATCAGCTTCCTGGATACAAATAT |

[a] The underlined T in the splint was absent in the round 1 ligation step because the DNA pool in this step was prepared by solid-phase synthesis and therefore did not have an untemplated A nucleotide at its 3′-end. The underlined T was included in all other selection rounds to account for the untemplated A nucleotide that is added at the 3′-end of each PCR product by Taq polymerase.

Table 3.5. MALDI-MS for the DNA-anchored peptide substrates are depicted in this table. [a] indicates that mass spec was not specifically obtained for this substrate, but we have high confidence because the MS of the oligonucleotide and peptide were independtly determined and simultaenous conjugations were performed in parallel with MS of the other conjugates validated. All masses are [M+H] in positive mode. Refer to Table 3.4 for sequence information about the DNA anchor.

| DNA-anchored peptide | mass calcd. | mass found | error, % (found – calcd.) |
|---|---|---|---|
| **For selections WM, MK1, and ML1** | | | |
| DNA-HEG-CAAAY$^{P}$AAA | 7125.1 | 7124.8 | –0.004 |
| DNA-HEG-CAFFY$^{P}$FFA | 7429.4 | 7429.1 | –0.004 |
| **For selections FT1, MP1, and MQ1** | | | |
| DNA-HEG-CAEEY$^{P}$EEA | 7355.2 | 7355.3 | +0.001 |
| DNA-HEG-CAKKY$^{P}$KKA | 7351.4 | 7349.4 | –0.027 |
| DNA-HEG-CAFFY$^{P}$FFA | 7427.4 | 7428.1 | +0.009 |
| **For selections FV1, MR1, and MS1** | | | |
| DNA-HEG-CAEEY$^{P}$EEA | 7345.2 | [a] | Na |
| DNA-HEG-CAKKY$^{P}$KKA | 7341.4 | 7340.9 | –0.007 |
| DNA-HEG-CAFFY$^{P}$FFA | 7417.4 | 7418.6 | +0.016 |
| **For selection FW1 and HM1** | | | |
| DNA-HEG-CEKIGEGT$^{P}$Y$^{P}$GVV | 7866.7 | 7868.3 | +0.020 |
| DNA-HEG-CAAAY$^{P}$AAA | 7243.1 | [a] | Na |
| **For selection FX1 and HN1** | | | |
| DNA-HEG-CIEEY$^{P}$TEM | 7495.4 | 7494.0 | –0.019 |
| DNA-HEG-CAKKY$^{P}$KKA | 7417.5 | [a] | Na |
| **For selection FY1 and HP1** | | | |
| DNA-HEG-CEGWY$^{P}$RAL | 7457.4 | 7459.1 | +0.023 |
| DNA-HEG-CAKKY$^{P}$KKA | 7399.5 | 7398.2 | –0.018 |
| **For selection HQ1** | | | |
| DNA-HEG-CAKKY$^{P}$KKA | 7434.5 | 7434.1 | –0.005 |

### 3.4.3 In Vitro Selection Procedure for FT1-FY1 Selections

The selection procedure, cloning, and initial analysis of individual clones were performed essentially as described previously.[23,28] The random deoxyribozyme pool, PCR primers (both forward and reverse) are shown in the Table 3.4. In each round, ligation step was used to attach the deoxyribozyme pool at its 3′-end to the 5′-end of the DNA-anchored phosphopeptide substrate using a DNA splint and T4 DNA ligase. The

splint sequence is shown in Table 3.4 as well. A nucleotide is added at the 3′-end of each PCR product by Taq polymerase. This T nucleotide was omitted from the splint used for ligation of the initially random $N_{40}$ pool, which was prepared by solid-phase synthesis. Nucleotide sequences of the DNA anchor oligonucleotide, the deoxyribozyme binding arms and the full 15MZ36 capture deoxyribozymes are depicted in Table 3.4. 5'-Triphosphorylated RNA sequence is the same as used previously in chapter 2.

*Procedure for ligation step in round 1.* A 34 μL sample containing 1 nmol of DNA pool, 850 pmol of DNA splint, and 750 pmol of 5′-phosphorylated DNA-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 was added 4 μL of 10× T4 DNA ligase buffer (Fermentas) and 2 μL of 5 U/μL T4 DNA ligase (Fermentas). The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

*Procedure for ligation step in subsequent rounds.* A 16 μL sample containing the PCR-amplified DNA pool (~5–10 pmol), 20 pmol of DNA splint, and 30 pmol of 5′-phosphorylated DNA-phosphopeptide 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 (Fermentas) and 2 μL of 1 U/μL T4 DNA ligase (Fermentas). The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

*Procedure for selection step in round 1.* Each selection experiment was initiated with 200 pmol of the ligated $N_{40}$ pool. A 15 μL sample containing 200 pmol of ligated $N_{40}$ pool 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 30 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.

*Procedure for selection step in subsequent rounds.* A 10 μL sample containing ligated pool 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. The sample was incubated at 37 °C for 14 h.

*Procedure for capture step in round 1.* To the 30 μL selection sample was added 1.2 μL of 50 mM EDTA, pH 8.0 (60 nmol). The sample was then brought to 45 μL total volume containing 300 pmol of 15MZ36 capture deoxyribozyme, 400 pmol of 5′-triphosphorylated RNA substrate, 20 mM $MnCl_2$, and 40 mM $MgCl_2$ and incubated at 37 °C for 14 h.

*Procedure for capture step in subsequent rounds.* To the 20 μL selection sample was added 0.8 μL of 50 mM EDTA, pH 8.0 (40 nmol). The sample was then brought to 25 μL total volume containing 40 pmol of 15MZ36 capture deoxyribozyme, 80 pmol of 5′-triphosphorylated RNA substrate, and 20 mM $MnCl_2$, 40 mM $MgCl_2$, and incubated at 37 °C for 14 h.

*Procedure for PCR.* 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-purified selection product, 200 pmol of forward primer, 50 pmol of reverse primer, 20 nmol of each dNTP, and 10 μL of 10× Taq polymerase buffer (1× = 20 mM

Tris-HCl, pH 8.8, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, and 0.1% Triton X-100). This sample 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 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), and 5 μL of 10× Taq polymerase buffer. 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.

### 3.4.4 Testing for Sequence Selectivity (FT1, FV1, and FX1)

The DNA-anchored phosphopeptide substrate was 5′-$^{32}$P-radiolabeled using γ-$^{32}$P-ATP and Optikinase (USB), which lacks the 3′-phosphatase activity that we have found also dephosphorylates tyrosine and serine side chains. A 10 μL sample containing 0.25 pmol of 5′-$^{32}$P-radiolabeled DNA-anchored phosphopeptide substrate and 20 pmol of 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 DNA-catalyzed dephosphorylation 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. At appropriate time points, 5 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). Correctly matched DNA-anchored peptides were tested with the appropriate enzymes. For instance, when testing FT1 deoxyribozymes, the DNA anchor

used for FT1 selection (shown in Table 3.4) was used to attach and test the different peptidic substrates. The peptide sequences were varied from CAKKY$^{P}$KKA, CAEEY$^{P}$EEA and CIEEY$^{P}$TEM. Using a 20% PAGE, single-turnover desphosphorylation was difficult to investigate for DNA-anchored CAEEY$^{P}$EEA, because the substrate and the product migrated closely (data not shown). Hence peptide CIEEY$^{P}$TEM was formally used as a surrogate for E-rich phosphopeptide in the selectivity assay although CAEEY$^{P}$EEA was confirmed for dephosphorylation as well.

### 3.4.5 In Vitro Selection Procedure for HM1 to HQ1

For selections involving competitive ligands, the procedure was similar to that described in section 3.4.3 with the following modification; free peptides (with properties different to the actual DNA-anchored substrate) were supplemented in the buffer as an active selection pressure to encourage sequence selectivity. Details for the pool, primers, ligation splint, 15MZ36 capture enzyme, and DNA anchor can be found in Table 3.4. HM1 selection was similar to FW1 selection with DNA-anchored CAAAY$^{P}$AAA as the actual substrate and 1 mM AKKY$^{P}$KKA (free peptide) supplemented in the selection buffer as a selection pressure. HN1 mimicked FX1 selection with CAKKY$^{P}$KKA as DNA-anchored peptide and 1 mM AAAY$^{P}$AAA supplemented in the selection buffer. HN1 mimicked the FX1 selection with CAKKY$^{P}$KKA as DNA-anchored peptide and 1 mM AAAY$^{P}$AAA supplemented in the selection buffer.

For HQ1 selection, information for the sequences is provided in Table 3.4. The actual substrate was CAKKY$^{P}$KKA and the competitive substrate was AFFY$^{P}$FFA (0.5 mM). The selection condition used 2 mM of $ZnCl_2$ instead of the conventional 1 mM $Zn^{2+}$ during the capture step. For the 15MZ36 capture step, the amount of EDTA was

increased twice to selectively quench the excess zinc followed by addition of the 15MZ36 capture enzyme and the 5'-triphosphorylated DNA.

### 3.4.6 In Vitro Selection Procedure for MK1-MN1

Selections 14MK1, 14ML1, 14MM1 and 14MN1 were continued from 13WM (WM selections discussed in chapter 2). For a general overview of the reverse complement inhibition selections, please refer to Figure 3.11. Information about primers, pool, ligation splint, and capture deoxyribozyme 15MZ36 is provided in chapter 2. The pool was divided into two portions, one for the undesirable substrate and the other for the desirable substrate. Details about the desirable and undesirable substrates are discussed in Table 3.3.

#### 3.4.6.1 Procedure for the undesirable substrate (Red pathway)

For MK1 and ML1 were control selections with no undesirable substrates. For MM1, DNA-anchored CAAAY$^{P}$AAA was the substrate. For MN1, the substrate was DNA-anchored CAFFY$^{P}$FFA. Ligation, selection and capture steps were performed with 15MZ36 for MM1. However, the capture enzyme for MN1 (F-rich peptide) was 8HJ102 with the following sequence, 5'-CCCGAAAGCCTCCTTCGGGAGATGTCTCTCAGACGGAAACTTTC-AGTACGGAATGGATAGTGAGTCGTATTATCCCCATCAGGATCAGC-3′, and a different 5'-triphosphorylated RNA sequence specific to 8HJ102.[24] Addition of EDTA for selection chelation in the capture step was omitted because 8HJ102 is a zinc-dependent enzyme that functions optimally with 40 mM $Mg^{2+}$, 20 mM $Mn^{2+}$ and 1 mM $Zn^{2+}$.

After the capture step, 10-cycle PCR was performed. Second, a 50 μL sample was prepared containing 1 μL of the 10-cycle PCR product, 25 pmol of forward primer, 100 pmol of modified long reverse primer (5'-$(AAC)_{12}$XCCATCAGGATCAGCT-3'), 10 nmol of

each dNTP, and 5 μL of 10× Taq polymerase buffer. This sample (scaled up to 6 times) 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. The PCR sample was combined to 300 μL and precipitated by adding 30 μL of 3M NaCl, and 990 μL of ethanol. The anti-sense reverse complement strand was separated by 8% PAGE and isolated.

**3.4.6.2 Procedure for the desirable substrate**

The desirable substrates for MK1, ML1, MM1 and MN1 were CAAAY$^{P}$AAA. In the selection step, 15 pmol of the quantified reverse complement (anti-sense) oligonucleotide was added. In addition, 30 pmol of binding arm primers were provided to prevent the constant binding arm portion of the reverse complement from interacting non-specifically with the deoxyribozyme pool (Figure 3.14). Selection step was performed with these added components and incubated at 37 °C for 14 h. After this step, the pool was separated by 8% PAGE to remove the additional components, including the reverse complement. For MK1, no components were added in the selection step. For ML1, reverse complement prepared from a random oligonucleotide pool was supplement in the selection, including the binding arms described above.

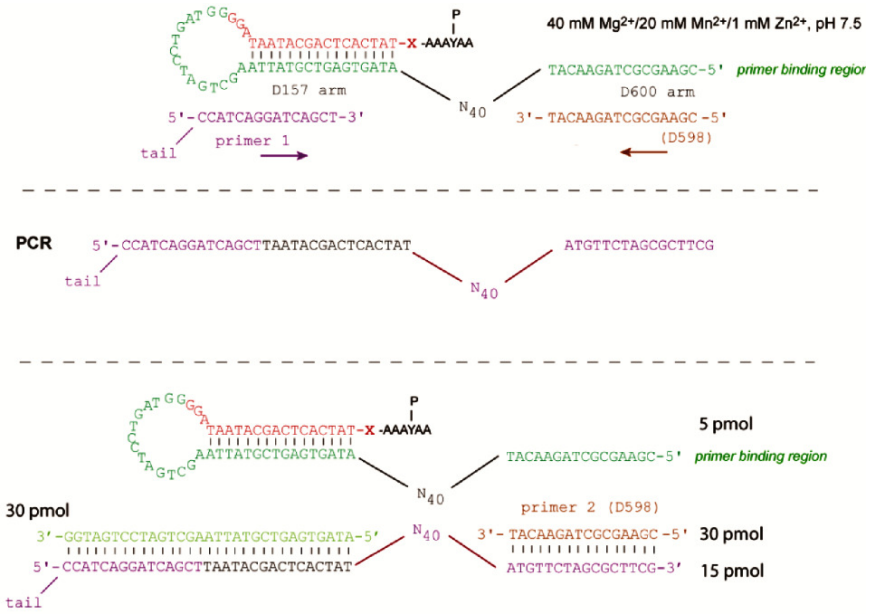


**Figure 3.14.** Once the reverse PCR primer with a long 5' tail is isolated (shown in purple), 15 pmol of this component was added to the selection step in the desirable pathway. In addition, two binding arms for the pool shown in light green and brown were added in (30 pmol) to prevent the reverse strand (purple) from engaging with non-specific interactions with the pool. Watson-Crick base pairing is exclusively favored at the N40 region for the sense and anti-sense strands.

Capture step is performed by bringing the reaction solution to 20 µL total volume containing 50 pmol of 15MZ36 (sequence provided in chapter 2), 100 pmol of 5'-triphosphorylated RNA/DNA chimera (check Walsh et al.),[27] 20 mM $MnCl_2$, 40 mM $MgCl_2$, and incubated at 37 °C for 14 h.

### 3.4.7 In Vitro Selection Procedure for MP1-MS1

Selections 8MP1, 8MQ1 were initiated from 7FT1 and 8MR1, 8MS1 were initiated from 7FV1. The information about primers, pool, splint for ligation, capture deoxyribozyme are also provided in Table 3.4. MP1/MQ1 correspond to FT1 and MR1/MS1 correspond to FV1. The pool was divided into two portions, one for the undesirable substrate and the other for the desirable substrate. Details about the desirable and undesirable substrates, including the desired outcome are discussed in Table 3.4

### 3.4.7.1 Procedure for the undesirable substrate:

For MP1 and MQ1, the desired substrate was E-rich peptide and selectivity was pursued against K-rich peptide and F-rich peptides. Similarly for MR1 and MS1, K-rich peptide was the actual substrate with selectivity pressure against E-rich and F-rich peptides. 8HJ102 sequences used were 5'-CCCGAAAGCCTCCTTCGGGAGATGTCTCTCAGACGG-AAACTTTCAGTACGGAATGGATACGCATAAAGGTAGTCCCCATCAGGATCAGC-3' for MQ1 and 5'-CCCGAAAGCCTCCTTCGGGAGATGTCTCTCAGACGGAAACTTTCAGTACGGAATGGATAAGA-AACGAGATATTCCCCATCAGGATCAGC-3' for MS1 Addition of EDTA in the capture step with 8HJ102 deoxyribozyme was omitted as discussed earlier. After the capture step, PCR was performed to isolate the corresponding anti-sense reverse complement strands.

### 3.4.7.2 Procedure for the desirable substrate

The desirable substrates for MP1/MQ1 were CAEEY$^{P}$EEA and CAKKY$^{P}$KKA for MR1/MS1. In the selection step, 15 pmol of the quantified reverse complement (anti-sense) oligonucleotide was supplemented to inhibit the sequence-general deoxyribozymes. In addition, 30 pmol of binding arm primers (see Figure 3.14 and Table 3.4 for the sequence information) were provided to sequester the constant binding region of the reverse complement. Selection step was performed with these added components and incubated at 37 °C for 14 h. After this step, the pool was separated by 8% PAGE to remove the reverse complement and other additional components.

Capture step was performed by bringing the reaction solution to 20 μL total volume containing 50 pmol of 15MZ36 (sequence provided in chapter 2), 100 pmol of 5'-triphosphorylated RNA/DNA chimera (check Walsh et al.),[27] 20 mM $MnCl_2$, 40 mM $MgCl_2$, and incubated at 37 °C for 14 h.

### 3.4.8 Mass Spectrometry Experiment Procedure

DNA-anchored peptides were desalted by Millipore $C_{18}$ ZipTip, and analyzed by MALDI mass spectrometry. Data were acquired on a Bruker UltrafleXtreme MALDI-TOF mass spectrometer with matrix 3-hydroxypicolinic acid in positive ion mode at the UIUC School of Chemical Sciences Mass Spectrometry Laboratory (instrument purchased with support from NIH S10RR027109A).

## 3.5 References


(1) Brannigan, J. A.; Wilkinson, A. J. Protein engineering 20 years on. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 964-970.

(2) Bornscheuer, U. T.; Pohl, M. Improved biocatalysts by directed evolution and rational protein design. *Curr. Opin. Chem. Biol.* **2001**, *5*, 137-143.

(3) Laible, M.; Boonrod, K. Homemade site directed mutagenesis of whole plasmids. *J. Vis. Exp.* **2009**.

(4) Joern, J. M. DNA shuffling. *Methods Mol. Biol.* **2003**, *231*, 85-89.

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

(6) Reetz, M. T.; Kahakeaw, D.; Lohmer, R. Addressing the numbers problem in directed evolution. *ChemBioChem* **2008**, *9*, 1797-1804.

(7) Selner, N. G.; Luechapanichkul, R.; Chen, X.; Neel, B. G.; Zhang, Z. Y.; Knapp, S.; Bell, C. E.; Pei, D. Diverse levels of sequence selectivity and catalytic efficiency of protein-tyrosine phosphatases. *Biochemistry* **2014**, *53*, 397-412.

(8) Li, Q.; Yi, L.; Marek, P.; Iverson, B. L. Commercial proteases: present and future. *FEBS Lett.* **2013**, *587*, 1155-1163.

(9) Olsen, J. V.; Ong, S. E.; Mann, M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell Proteomics* **2004**, *3*, 608-614.

(10) Hedstrom, L.; Szilagyi, L.; Rutter, W. J. Converting trypsin to chymotrypsin: the role of surface loops. *Science* **1992**, *255*, 1249-1253.

(11) Hedstrom, L.; Perona, J. J.; Rutter, W. J. Converting Trypsin to Chymotrypsin - Residue-172 Is a Substrate-Specificity Determinant. *Biochemistry* **1994**, *33*, 8757-8763.

(12) Appel, W. Chymotrypsin: molecular and catalytic properties. *Clin. Biochem.* **1986**, *19*, 317-322.

(13) Masaki, T.; Tanabe, M.; Nakamura, K.; Soejima, M. Studies on a New Proteolytic-Enzyme from *Achromobacter-Lyticus* M497-1 .1. Purification and Some Enzymatic-Properties. *Biochim. Biophys. Acta* **1981**, *660*, 44-50.

(14) Nemoto, T. K.; Ohara-Nemoto, Y.; Ono, T.; Kobayakawa, T.; Shimoyama, Y.; Kimura, S.; Takagi, T. Characterization of the glutamyl endopeptidase from *Staphylococcus aureus* expressed in *Escherichia coli*. *FEBS J.* **2008**, *275*, 573-587.

(15) Pogson, M.; Georgiou, G.; Iverson, B. L. Engineering next generation proteases. *Curr. Opin. Biotechnol.* **2009**, *20*, 390-397.

(16) Bottcher, D.; Bornscheuer, U. T. Protein engineering of microbial enzymes. *Curr. Opin. Microbiol.* **2010**, *13*, 274-282.

(17) Lu, Y.; Liu, J.; Li, J.; Bruesehoff, P. J.; Pavot, C. M.; Brown, A. K. New highly sensitive and selective catalytic DNA biosensors for metal ions. *Biosens. Bioelectron.* **2003**, 529-540.

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

(19) Bruesehoff, P. J.; Li, J.; Augustine, A. J., 3rd; Lu, Y. Improving metal ion specificity during in vitro selection of catalytic DNA. *Comb. Chem. High Throughput Screening* **2002**, *5*, 327-335.

(20) Breaker, R. R. Engineered allosteric ribozymes as biosensor components. *Curr. Opin. Biotechnol.* **2002**, *13*, 31-39.

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

(22) Gu, Y.; Rosenblatt, J.; Morgan, D. O. Cell cycle regulation of CDK2 activity by phosphorylation of Thr160 and Tyr15. *EMBO J.* **1992**, *11*, 3995-4005.

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

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

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

(26) 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.

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

(28) 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.

# Chapter 4: Intracellular Application of Phosphatase DNA Catalysts[4]

## 4.1 Introduction

In this chapter, strategy to achieve phosphatase activity with deoxyribozymes inside live mammalian is discussed. Some of the earliest *in vivo* efforts with DNA catalysts involved the use of 10-23 deoxyribozyme to cleave and silence mRNA.[1,2] DNA enzymes with RNA cleaving activity have been explored in a few therapeutic avenues.[2] Additionally, DNA catalysts have been used for sensing and determining the concentration of specific metal cation concentrations inside live cells.[3,4] However, deoxyribozymes, so far, have never been utilized to serve as a protein (enzyme) substitutes inside live cells. Establishing the fundamental ability of single-stranded DNA to replace a functional protein inside cells would be an important accomplishment in the field of DNA catalysis.

### 4.1.1 Techniques used in Gene Therapy

Several techniques exist for delivering DNA into cells. A physical approach involves weakening of the cell membrane to enable penetration of external DNA through the cell membrane. Examples of such physical manipulation include electroporation, ultrasound/hydrodynamic treatment, and needle injection. Alternatively, chemical techniques such as viral and non-viral vectors can be used for nucleic acid delivery. Polycationic derivatives and cationic lipids can be used to coat the DNA and shield the negative charge. This helps the DNA pass through the electrostatically repulsive phospholipid cell membrane. Covalent conjugation to polycationic polymers such as

[4] Prof. Silverman and I conceived and planned the project for intracellular phosphatase activity.

polyarginine also serves as excellent delivery agents.[5] These delivery agents can be designed to be biodegradable to enable controlled spatial and temporal release of nucleic acids inside cells.[6] Viral vectors have received substantial success in the field of gene therapy. The most popular viral vectors used for gene therapy are adenovirus, retrovirus, and vaccinia virus.[7] Several therapeutic targets currently being addressed by gene therapy are under clinical trials (Figure 4.1).[8]

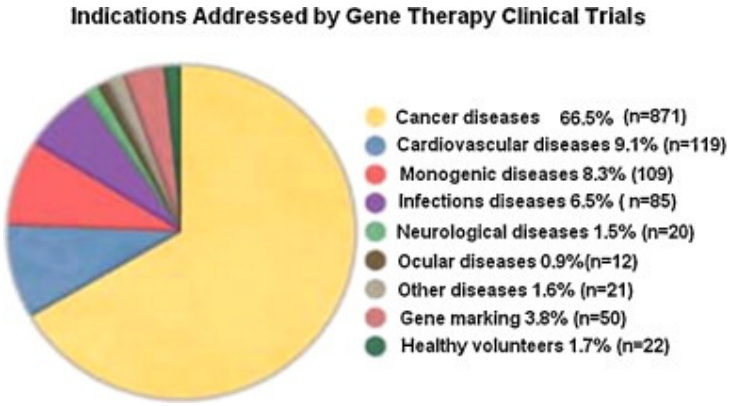


**Figure 4.1.** Clinical targets currently addressed by gene therapy in clinical trials (figure from reference 8).

Advances in gene therapy continue to offer better alternatives for delivery of deoxyribozymes inside cells. One of the earliest efforts to introduce DNA enzymes involved direct *in vivo* expression of single-stranded DNA in mammalian cells.[9] More recently, gold nanoparticles[3,4] and lipidoid-coated iron oxide nanoparticles[10-11] have successfully introduced short DNA inside cell. Some of the promising features of nanomaterials include the ease of synthesis, efficient functionalization with nucleic acids, low toxic effects, and high delivery payload. Classical lipid transfection agents such as Lipofectamine also function as reliable delivery agents, albeit with lower payload efficiency in comparison to nano-particle based methods.[11] Although Lipofectamine is

suitable for larger DNA constructs such as plasmids, better methods exist for delivery of short DNA strands.

### 4.1.2 Choice of Cell System for DNA

This chapter focuses on efforts to achieve dephosphorylation with phosphatase deoxyribozymes on full-length proteins inside live mammalian cells. Phosphatase deoxyribozymes identified so far (namely 14WM9 or 14M27), are zinc-dependent phosphomonoesterases.[12] These enzymes function specifically with phosphotyrosine-containing peptides and proteins. No free peptide activity is observed with phosphoserine.[12] However, it is important to note that these DNA catalysts require up to millimolar ($\geq$ 2 mM) concentration of $Zn^{2+}$ for optimal enzymatic activity.[12] This high dependence on zinc is a major concern for intracellular dephosphorylation. Cells lack millimolar quantity of zinc in cytosol, either in bound or unbound form.

To address this concern, we first need to consider specific cells cell lines that can be maintained or cultured at high concentration of zinc. Secondly, these cells should have the ability to intake high amount of zinc in the cytosol when maintained in zinc-rich conditions. LNCaP cells (malignant human prostatic adenocarcinoma, androgen-dependent) are suitable candidates for such purposes. These cells are known to accumulate the highest concentration of mobile zinc of all soft tissues in our body.[13] Additional of exogenous hormones such as testosterone or prolactin leads to overexpression of zinc-specific transporters on the cell membrane, which in turn increases the concentration of intracellular zinc.[13]

Although the total amount of zinc in cells is easy to estimate, concentration of free (mobile or unbound) zinc is challenging to measure. Zinc can be in either free form

(available for catalysis) or strongly bound to proteins changing their effective free concentration. DNA enzymes need free (mobile) zinc for functioning. If the zinc is tightly bound to cellular components, deoxyribozymes may not function optimally. Lippard and co-workers calculated intracellular concentration of mobile (free) zinc using a zinc fluorescent biosensor, ZPP1.[14-15] Based on their assessment, normal peripheral prostate cancer cell line, RWPE1, cultured in a zinc rich media contains free intracellular zinc of 12–16 fmol/cell.[15] Assuming an average cell volume of 4.5 pL (adapted from Min6)[14], the concentration of free zinc within a normal peripheral prostate cell is approximately 2.5 mM–3.5 mM. Analysis of LNCaP and DU145 (malignant human prostatic adenocarcinoma, androgen-dependent and androgen-independent) reveal negligible free zinc based on ZPPI fluorescence.[14] However, Costello et al. suggest that treatment with hormones (testosterone or prolactin) leads to intake of zinc in the cytosol changing the free zinc concentration to greater than 2 fmol/cell.[13] This implies that treatment with hormones can lead to a final concentration of $\geq$ 0.4 mM loosely-bound intracellular zinc in LNCaP cytosol. Because LNCaP is well studied and established in numerous prostate cancer studies, we pursued LNCaP as the test case.

### 4.1.3 Downstream Investigation of Intracellular Activity

To analyze the effect of phosphatase deoxyribozymes inside LNCaP cells, we need an appropriate positive control. Incidentally, PTP1B (Protein Tyrosine Phosphatase) was overexpressed in LNCaP cells in an earlier report, which led to cell differentiation of LNCaP to neuroendocrine (NE) cells.[16] If phosphatase DNA catalysts function similar to PTP1B inside LNCaP cells, similar phenotypic outcome would be expected.

Protein tyrosine phosphatase PTP1B is a highly proficient enzyme with a $k_{cat}/K_m$ of ~$10^5$ to $10^7 M^{-1} s^{-1}$.[17] PTP1B is predicted to dephosphorylate 50% of phosphotyrosine-containing proteins in the cytosol within 5 s.[18] The deoxyribozyme 14WM9 has $k_{obs}$ of 0.65 $min^{-1}$ and a $K_m$ greater than 1 mM, which equates to (at best) a $k_{cat}/K_m$ of 100 $M^{-1} s^{-1}$. Under molecular crowding conditions of the cytosol (200 mg/mL), catalytic proficiency of 14WM9 can have a huge negative impact in the rate. Moreover, 14WM9 requires optimal zinc concentration for catalytic activity. For 14WM9 activity on larger protein substrates, $k_{obs}$ can be further reduced. Hence it is essential to introduce high amounts of DNA catalysts inside each cell to compensate for the high catalytic proficiency of PTP1B. This requires a highly proficient intracellular delivery of deoxyribozymes.

When PTP1B was overexpressed in LNCaP cells by conventional lipid-based transfection, neuroendocrine phenotype were observed.[16] The neuroendocrine-differentiated cells are characterized by specific NE marker proteins. These proteins, neuron-specific enolase (NSE) and chromogranin A (CGA), can be conveniently tested by qRT-PCR or Western blot based on a previous report.[16]

In conclusion, three important criteria need to demonstrate the proof-of-concept for intracellular dephosphorylation by phosphatase deoxyribozymes (Figure 4.2).

- Cell line needs to intake high amounts of zinc in the cytosol. For that purpose we chose LNCaP cells as the most viable candidate
- We should have an appropriate control where PTP1B has been previously over-expressed in LNCaP. Fortunately previous studies have investigated this.

- We should have a proper downstream readout to verify the effectivity of phosphatase activity. PTP1B overexpression leads to neuroendocrine differentiation of LNCaP cells. Biomarkers such as NSE and CGA can be quantified by qRT-PCR:

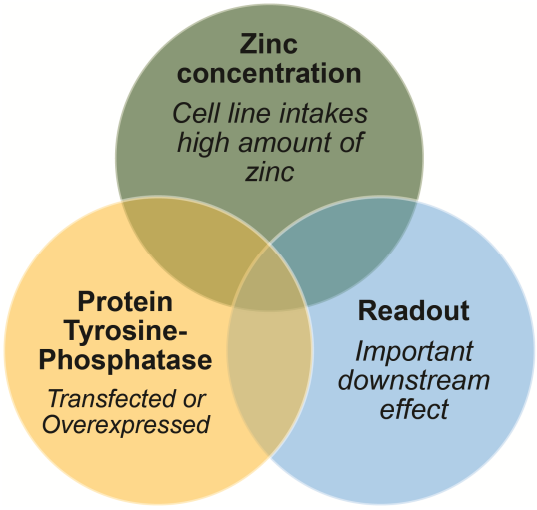


**Figure 4.2.** Three important criteria need to be simultaneously satisfied to establish the ability of DNA catalysts to function as phosphatases inside cells. A) The cell line should be able to accommodate high concentration of zinc in the cytosol. B) Protein tyrosine phosphatase should already be overexpressed in the specific cell line. C) A convenient readout should exist for the downstream effect.

## 4.2 Results and Discussion

### 4.2.1 In Vitro Selection for More Proficient Phosphatase Deoxyribozymes

#### 4.2.1.1 Reselections FA1, FB1 and FC1

In an effort to identify more proficient phosphatase deoxyribozymes, reselections were performed for 14WM9 and 14WM27 (Figure 4.3). Two partially randomized DNA pools were synthesized (each for 14WM9 and 14WM27) maintaining the same constant binding region, but randomizing the 40-nucleotide region by 25% at each position. These selections were name FA1 and FB1 respectively. An additional reselection with 14WM9 called FC1 was initiated, which involved a DNA-anchored phosphopeptide CAKKY$^{p}$KKA (K-rich peptide conjugated by disulfide linkage). Each pool was then

ligated to the DNA-anchored phosphopeptide. The selection strategy followed the same procedure discussed in chapter 2. Each selection step used incubation conditions of 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, active DNA catalyst sequences were "captured" with the aid of the 15MZ36 deoxyribozyme that selectively uses dephosphorylated tyrosine ($Y^{OH}$) as the nucleophile to attack 5′-triphosphorylated RNA, resulting in a PAGE shift for the corresponding deoxyribozyme sequence.

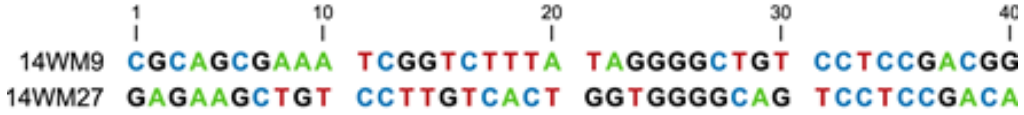


**Figure 4.3.** The 40-nucleotide region of 14WM9 and 14WM27 from WM selection is depicted here. Each position was partially randomized 25% for reselections FA1 and FC1 with 14WM9, and FB1 reselection with 14WM27.

Active DNA sequences were amplified by PCR and joined by T4 DNA ligase to the DNA-anchored hexapeptide substrate for entering the next selection round. After 5 rounds of selection, activity was observed for all selections (Figure 4.4). A two hour time pressure was applied at round 7 for FA1, FB1 and FC1 to enable enrichment of faster deoxyribozymes from the pool. However the activity at 2 h selection pressure was greater than $1/6^{th}$ the activity at 12 h time point (Figure 4.4). This indicated that activity at 12 h was already dominated by fast deoxyribozymes and applying time pressure is redundant. Individual deoxyribozymes were cloned and characterized from round 6 for all selections.

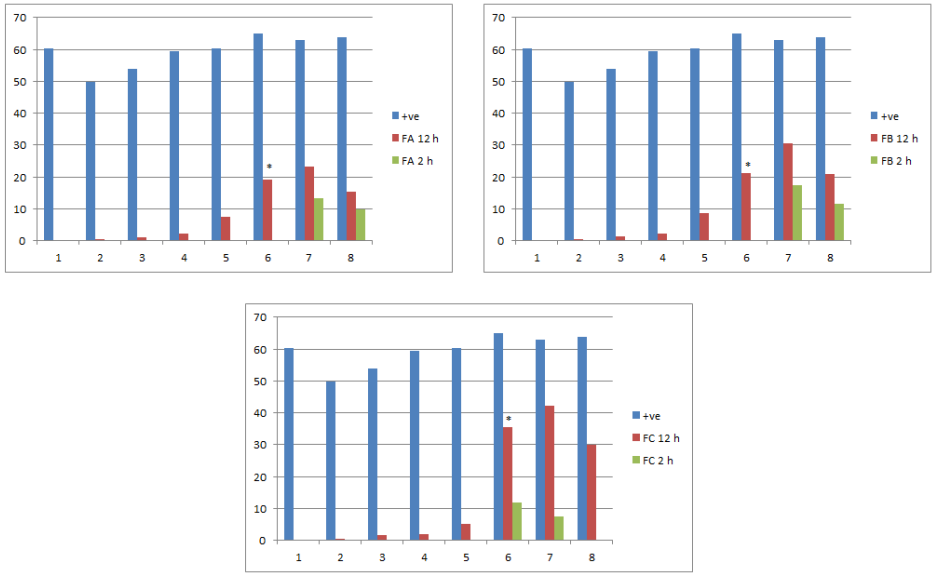


**Figure 4.4.** Initial activity was observed for FA1, FB1 and FC1 at round 5. The blue plot indicates positive control with 15MZ36 capture yield. Cloning was performed at round 6 for FA1, FB1, and FC1. The green plot denotes a selection pressure of 2 h applied for enriching faster enzymes. However, fast sequences already dominated the pool and hence selections with 2 h time point were not pursued for cloning.

Distinct deoxyribozyme sequences were identified for FA1-FC1 (Figure 4.5). Activities of these distinct catalysts were tested with a $Y^P$ hexapeptide that is attached to the DNA anchor through a PEG-type tether that includes 60 ethylene glycol units (Figure 4.6). As discussed in chapter 2, activity with PEG-tethered hexapeptide suggests the possibility of catalysis with entirely free peptides as well. Based on investigation of these newly identified enzymes and their activity with PEG-tethered substrate, it was evident that 14WM9 was still amongst the most proficient phosphatase deoxyribozymes enzymes.

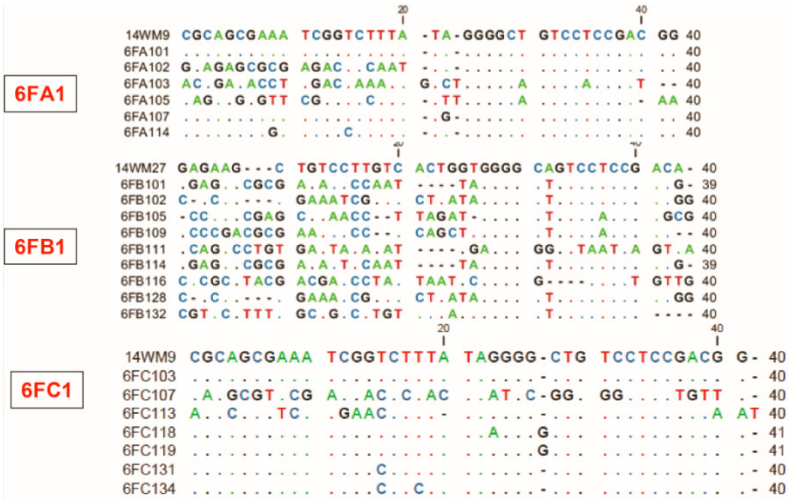


**Figure 4.5.** 40 nucleotide portion of the active phosphatase sequences from 6FA1, 6FB1, and 6FC1.

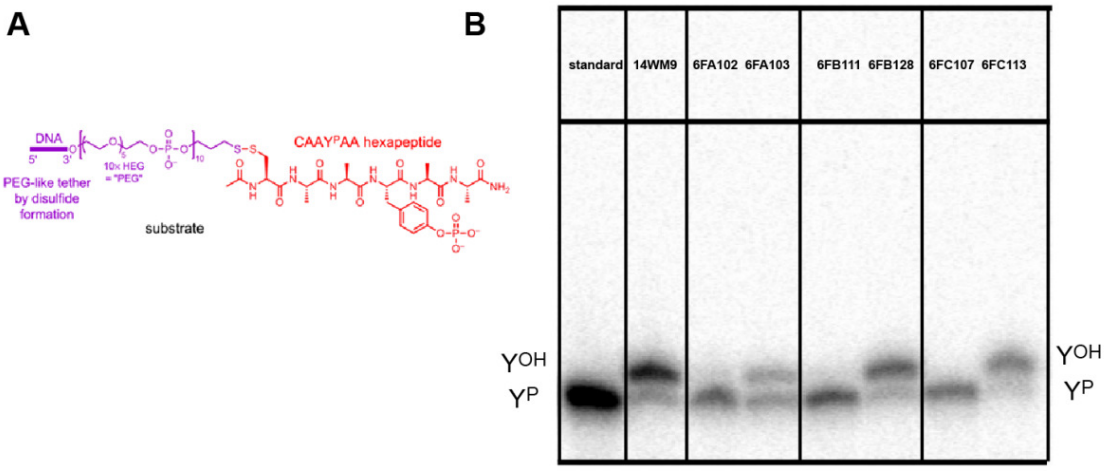


**Figure 4.6.** DNA-anchored CAAY$^{P}$AA was tested with a PEG-type tether linkage (22 h time point). For more details on PEG-type tether, see chapter 2 (Materials and Methods). A) Hexapeptide CAAYPAA is conjugated to the DNA anchor via a PEG-type tether. B) Substantial activity was observed for 14WM9, 6FA103, 6FB128 and 6FC113. However, 14WM9 was still amongst the fastest phosphatase enzymes.

### 4.2.1.2 In vitro selection with different random region length[5]

In vitro selection with different random region lengths were explored for improved phosphatase activity. HJ1 selections had 30 random nucleotides whereas HL1 had 50. DNA-anchored CAAY$^{P}$AA was used as the substrate. Each DNA pool was

[5] University of Illinois undergraduate Natalia Recko performed the HJ1 and HL1 selections under my supervision.

ligated to the DNA-anchored phosphopeptide. Selection procedures are mentioned in (Materials and Methods). No activity was observed with HL1 selections with 50 random nucleotides. However, HJ1 (30-nucleotide region) gave activity at round 7. Round 8 was used for cloning.

When assayed for free peptide activity with AAAY$^{P}$AA substrate using malachite green inorganic phosphate assay, 8HJ102 deoxyribozyme was determined to be more proficient that 14WM9. At 24 h, 8HJ102 gave 22 turnovers (Figure 4.7) compared to 6 turnovers by 14WM9 (see Materials and Methods for details).[12]

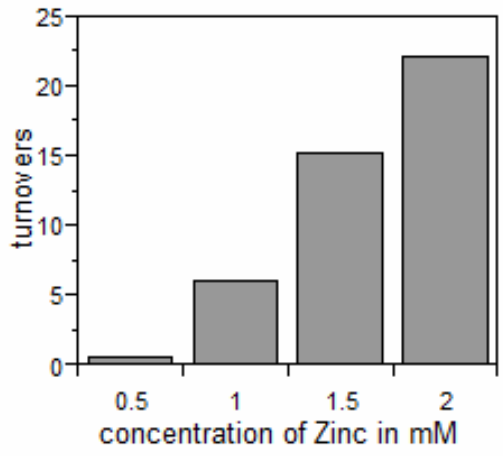


**Figure 4.7.** 8HJ102 is a zinc-dependent phosphatase deoxyribozyme. Up to 22 turnovers are observed in 24 h (2 mM $Zn^{2+}$) as calculated by malachite green assay (see Materials and Methods for experimental details).

### 4.2.2 Preparation of PTP1B Plasmid and the D181A Mutant

Human PTP1B (PTPN1) encoded in pcDNA3.1 vector (Invitrogen) was a generous gift from the Steidl lab.[19] Using the mutagenic primers 5′-ATACCACATGGCCT-GCCTTTGGAGTCCCTGA-3′ and 5′-TCAGGGACTCCAAAG**G**CAGGCCATGTGGTAT-3′, the D181 mutant was prepared (see Materials and Methods). T7 promoter sequence was used as a forward sequence primer to verify the sequences.

### 4.2.3 Cell Culture and Validation of the Positive Control

LNCaP cells were maintained in RPMI 1640 media that was supplemented with 10% FBS. 50,000 cells were transferred in a 24-well plate with 0.5 mL of media per well. jetPRIME (Polyplus) was mixed with the following DNA,

a) pcDNA3.1-PTP1B (500 ng)

b) pcDNA3.1-PTP1B-D181A (500 ng)

c) 14WM9 with Cy3 at the 3′ end (500 ng) and 14WM9 helper oligonucleotide (750 ng)

d) 8HJ102 with Cy3 at the 3′ end (500 ng) and 8HJ102 helper oligonucleotide (750 ng)

Although the deoxyribozymes (14WM9, 8HJ102) were successfully transfected in LNCaP cells (Figure 4.8), the payload of deoxyribozymes is not projected to be as high as gold nanoparticle-based delivery. Following transfection, cells were grown for 0, 3 and 6 days followed by detachment and isolation of the RNA. Investigation of NE genes NSE and CGA can be performed by qRT-PCR.

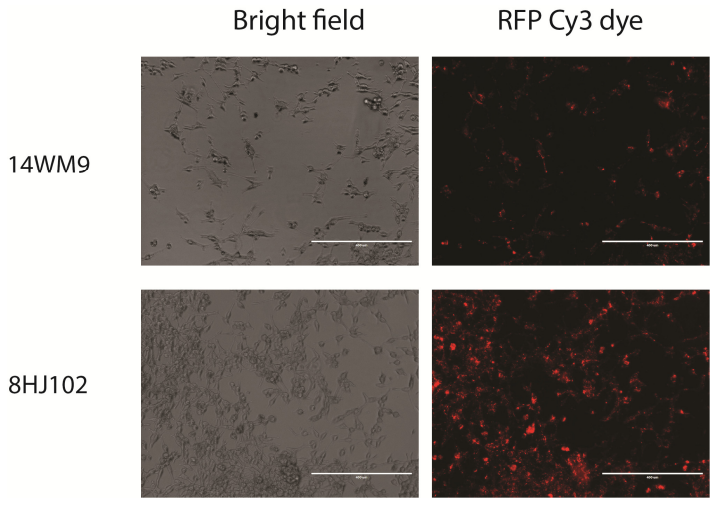


**Figure 4.8.** 14WM9 and 8HJ102 with Cy3 at the 3′ end were transfected in LNCaP cells using jetPRIME transfection reagent. Left panel shows the bright field and the right panel shows the red fluorescence (RFP) window.

### 4.2.4 Synthesis of Gold Nanoparticles and their Conjugation to Deoxyribozymes

Gold nanoparticles have been used to deliver high amounts of DNA inside cells. 3′-thiol DNA can be linked to gold nanoparticles of diameter 13 nm.[3] Each nanoparticle can carry up to 70 DNA enzymes based on previous reports.[3] Up to $10^6$ can be delivered inside cells, which accounts to 10-100 μM of DNA in a 2 pL cell.[3] Both 14WM9 and 8HJ102, with $(HEG)_3$-$C_3$-SH at the 3′ end were functionalized on gold nanoparticles. Appropriate helper (DNA anchor lacking HEG linker) oligonucleotides were hybridized (see material and methods). These gold-nanoparticle functionalized deoxyribozymes will be introduced to LNCaP cells and NSE and CGA levels will be investigated by qRT-PCR under different concentrations of zinc and testosterone.

### 4.2.5 qRT-PCR Assay

In the report by Wu et al.,[16] overexpression of PTP1B in LNCaP cells leads to neuroendocrine differentiation of the cells. Biomarkers for neuroendocrine cells are NSE and CGA, These can be investigated using qRT-PCR. Primers and procedure for qRT-PCR are already reported. [16] Preliminary test with qRT-PCR for a control GADPH gene indicates that the RNA integrity is maintained and qRT-PCR can be successfully performed (Figure 4.9).

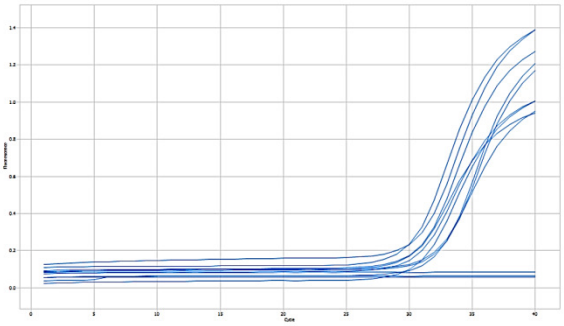

**Figure 4.9.** Initial qRT-PCR for GADPH gene was performed with PTP1B transfected sample (day 6), PTP1B-D181A sample (day 6), and no transfection control (day 6). The Ct values were 28.32 (n = 3), 29.74 (n = 3) and 27.95 (n = 3). No Ct value was obtained for a non-templated control (baseline).

## 4.3 Summary and Future Directions

The main goal of this chapter is to establish the fundamental ability of DNA catalysts to function as a full-length protein enzyme (PTP1B) inside cells (Figure 4.10). Tyrosine dephosphorylation plays a key role in cell signaling and signal transduction pathways.

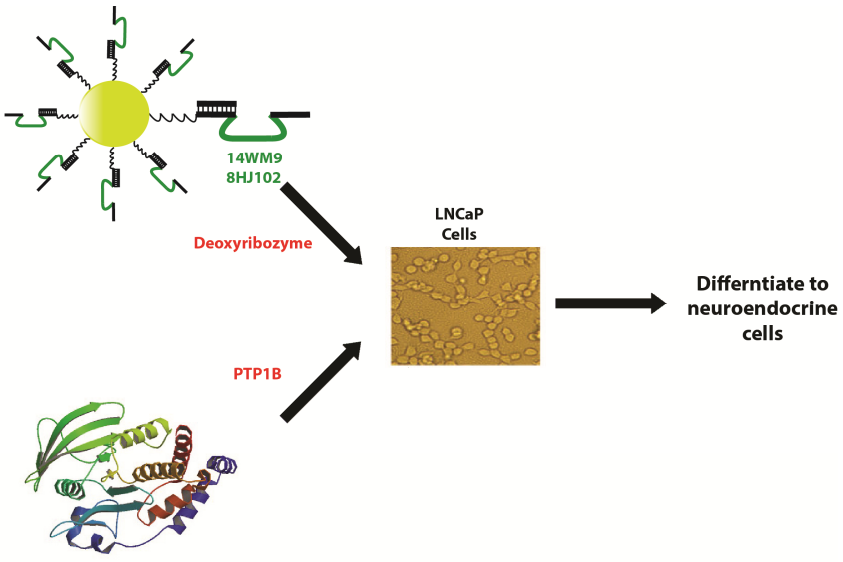


**Figure 4.10.** The main goal (vision) of this study is presented. PTP1B overexpression is known to impart neuroendocrine phenotype (or differentiation) to LNCaP cells. This can be verified by overexpression of the biomarkers CGA and NSE. We aim to achieve the NE phenotype using a phosphotyrosine phosphatase deoxyribozyme (14WM9 or 8HJ102) to establish the fundamental ability of deoxyribozymes to function as overexpressed full-length PTP1B enzyme inside cells.

Anomalous overexpression of protein tyrosine phosphatase (PTP1B) can lead to differentiation of cell types. In some cases, the overexpression might be beneficial, and conversely in detrimental in others. Gold nanoparticle-mediated delivery of phosphatase deoxyribozymes is a promising route for directly introducing enzymatic activity inside cells. The transfection efficiency is expected to be consistently high and reproducible when mediated by a gold nanoparticle-based delivery.

Gold nanoparticles were successfully synthesized and functionalized with deoxyribozymes containing 3′-thiol. LNCaP cells have been determined to be alive and viable in up to 50 μM of external zinc and 10 nM of testosterone, which most likely enables uptake of high amount of zinc in the cytosol. Preliminary qRT-PCR with GADPH has established the integrity or RNA and successful cDNA synthesis. We hope to establish and the replicate overexpression of neuroendocrine biomarkers such as NSE and CGA using qRT-PCR with known primers for pcDNA3.1-PTP1B plasmid, and 14WM9 and 8HJ102 deoxyribozymes.

A major concern with future *in vivo* application is the high metal cation dependence of deoxyribozymes. Proficiency of deoxyribozymes will also greatly impact their future use *in vivo*. Deoxyribozymes are rarely known to function on full-length proteins and their efficiency can be substantially decreased under molecular crowding conditions. Modified nucleotides with protein-inspired functional groups can not only reduce the metal dependence, but also strongly complement the already high catalytic prowess of DNA with phosphomonoester hydrolysis.

## 4.4 Materials and Methods

### 4.4.1 Oligonucleotides

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′-Triphosphorylated RNA oligonucleotides were prepared by in vitro transcription using synthetic DNA templates and T7 RNA polymerase.[20] 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 previously.[12]

### 4.4.2 Synthesis of Peptides and DNA-Anchored Peptides

Peptides were prepared by solid-phase synthesis on Fmoc Rink amide MBHA resin (AAPPTec). Each peptide was coupled to the DNA anchor oligonucleotide via N-terminal cysteine side chain (linkage created by disulfide formation) as discussed in chapter 2.

#### 4.4.2.1 Synthesis of peptides.

Peptides were prepared by solid-phase synthesis using Fmoc Rink amide MBHA resin, with *N,N,N′,N′*-tetramethyl-*O*-(7-azabenzotriazol-1-yl)uronium hexa-fluorophosphate (HATU) as coupling agent. Each synthesis was performed at 0.2 mmol scale. The sequences of peptides were CAAY$^{P}$AA and CAKKY$^{P}$KKA.

#### 4.4.2.2 Synthesis of DNA-anchored peptides.

Each peptide was coupled to the DNA anchor oligonucleotide via either the N-terminal cysteine side chain (linkage created by disulfide formation). The DNA anchor oligonucleotide (see Table 4.1) ending with HEG-$C_3$-SS-$C_3$-OH-3′, where the 3′-disulfide linker was introduced via standard solid-phase DNA synthesis and unmasked to a 3′-thiol

by DTT treatment. A 50 μL sample containing 2-5 nmol of DNA anchor oligonucleotide in 50 mM HEPES, pH 7.5, and 50 mM DTT was incubated at 37 °C for 2 h. The reduced product was precipitated to remove excess DTT by addition of 50 μL of water, 10 μL of 3 M NaCl, and 300 μL of ethanol. The precipitated product (DNA-HEG-$C_3$-SH) was dissolved in 45 μL of water. 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-$C_3$-SSPy) was precipitated by addition of 50 μL of water, 10 μL of 3 M NaCl and 300 μL of ethanol and dissolved in 25 μL of water. Conjugation to the peptide was performed by adding 20 μL of 50 mM triethylammonium acetate, pH 7.0, and 5 μL of 40 mM peptide (200 nmol). The sample was incubated at 37 °C for 2 h, and the DNA-anchored peptide was purified by 20% PAGE. A typical yield was 2-3 nmol.

**Table 4.1**. Different pools were used for each selection. Sequences are written 5′ to 3′. X denotes a HEG spacer to stop Taq polymerase from extending further. This enables separation of the single-stranded DNA. Sequence Y stands for CAAGGAGAGTTGTACAAGCTCGGGTCGTGTTCAAAGGGATC

| oligonucleotide purpose | oligonucleotide sequence |
|---|---|
| *Selection FA1 and FB1 (with CAAY$^P$AA) and FC1 with CAKKY$^P$KKA* | |
| DNA-HEG- substrate | GGATAATACGACTCACTAT-HEG-$C_3$-SH |
| forward primer for selection | CGAAGCGCTAGAACAT |
| reverse primer for selection | $(AAC)_4$XCCATCAGGATCAGCT |
| Partially randomized pool for selection | CGAAGCGCTAGAACAT-$N_{40}$-ATAGTGAGTCGTATTAAGCTGATCCTGATGG |
| 15MZ36 capture enzyme | CCGTCGCCATCTCTTC-Z-ATAGTGAGTCGTATTATCCCCATCAGGATCAGCT |
| splint for ligation step during selection [a] | ATAGTGAGTCGTATTATCCTCCATCAGGATCAGCTTAATACGACTCACTAT |
| *Selection HJ1 (with CAAY$^P$AA)* | |
| DNA-HEG- substrate | GGACTACCTTTATGCGTAT-HEG-$C_3$-SH |
| forward primer for selection | CGAACGAAAGCCTCCTTC |
| reverse primer for selection | $(AAC)_4$XCCATCAGGATCAGCT |
| random pool for selection | CGAACGAAAGCCTCCTTC –$N_{30}$-ATACGCATAAAGGTAGAGCTGATCCTGATGG |
| 15MZ36 capture enzyme | CCGTCGCCATCTCTTC-Z-ATACGCATAAAGGTAGTCCCCATCAGGATCAGC |
| splint for ligation step during selection [a] | ATAAGAAACGAGATATTCCTCCATCAGGATCAGCTATATCTCGTTTCTTAT |
| *Selection HL1 (with CAAY$^P$AA)* | |
| DNA-HEG- substrate | GGAATATCTCGTTTCTTAT-HEG-$C_3$-SH |
| forward primer for selection | CGAATTAAGACTGAATTC |
| reverse primer for selection | $(AAC)_4$XCCATCAGGATCAGCT |
| random pool for selection | CGAATTAAGACTGAATTC -$N_{50}$-ATAAGAAACGAGATATAGCTGATCCTGATGG |
| 15MZ36 capture enzyme | CCGTCGCCATCTCTTC-Z-ATAAGAAACGAGATATTCCCCATCAGGATCAGC |
| splint for ligation step during selection [a] | ATACCAATCAAGCCATTCCTCCATCAGGATCAGCTATGGCTTGATTGGTAT |

[a] The underlined T in the splint was absent in the round 1 ligation step because the DNA pool in this step was prepared by solid-phase synthesis and therefore did not have an untemplated A nucleotide at its 3′-end. The underlined T was included in all other selection rounds to account for the untemplated A nucleotide that is added at the 3′-end of each PCR product by Taq polymerase.

### 4.4.3 In Vitro Selection Procedure for FA1-FC1 and HJ1 and HL1

The selection procedure, cloning, and initial analysis of individual clones were performed essentially as described previously in chapter 2 and 3. The reselections FA1/FC1 and FB1 were partially randomized from the initial sequences 14WM9 and 14WM27 respectively. The pools were 5′-CGAAGTCCGCCATCTCTTC-$N_{40}$-ATAGTGAGTCGTATTAAGCTGATCCTGATGG-3′ with 40-nucleotide region partially randomized 25% at each position (see Figure 4.3 for 14WM9 and 14WM27). PCR primers are shown

in Table 4.1. In each round, the ligation step attached the deoxyribozyme pool at its 3′-end with the 5′-end of the DNA-anchored phosphopeptide substrate using a DNA splint and T4 DNA ligase. Capture step was performed using 15MZ36 and 5′-triphosphorylated RNA. See information on selection and sequences in Table 4.1. The peptidic substrates for FA1 and FB1 were CAAY$^{P}$AA. The substrate for FC1 was CAKKY$^{P}$KKA. For different N30 and N50 random nucleotide region selections HJ1 and HL1, hexapeptide substrate CAAY$^{P}$AA was used.

*Procedure for ligation step in round 1.* A 34 μL sample containing 1 nmol of DNA pool, 850 pmol of DNA splint, and 750 pmol of 5′-phosphorylated DNA-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 (Fermentas) and 2 μL of 5 U/μL T4 DNA ligase (Fermentas). The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

*Procedure for ligation step in subsequent rounds.* A 16 μL sample containing the PCR-amplified DNA pool (~5–10 pmol), 20 pmol of DNA splint, and 30 pmol of 5′-phosphorylated DNA-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 (Fermentas) and 2 μL of 1 U/μL T4 DNA ligase (Fermentas). The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

*Procedure for selection step in round 1.* Each selection experiment was initiated with 200 pmol of the ligated $N_{40}$ pool. A 15 μL sample containing 200 pmol of ligated $N_{40}$ pool 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 30 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.

*Procedure for selection step in subsequent rounds.* A 10 μL sample containing ligated pool 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. The sample was incubated at 37 °C for 14 h.

*Procedure for capture step in round 1.* To the 30 μL selection sample was added 1.2 μL of 50 mM EDTA, pH 8.0 (60 nmol). The sample was then brought to 45 μL total volume containing 300 pmol of 15MZ36 capture deoxyribozyme, 400 pmol of 5′-triphosphorylated RNA substrate, 20 mM $MnCl_2$, and 40 mM $MgCl_2$ and incubated at 37 °C for 14 h.

*Procedure for capture step in subsequent rounds.* To the 20 μL selection sample was added 0.8 μL of 50 mM EDTA, pH 8.0 (40 nmol). The sample was then brought to 25 μL total volume containing 40 pmol of 15MZ36 capture deoxyribozyme, 80 pmol of 5′-triphosphorylated RNA substrate, and 20 mM $MnCl_2$, 40 mM $MgCl_2$, and incubated at 37 °C for 14 h.

*Procedure for PCR.* 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-purified selection product, 200 pmol of forward primer, 50 pmol of reverse

primer, 20 nmol of each dNTP, and 10 μL of 10× Taq polymerase buffer (1× = 20 mM Tris-HCl, pH 8.8, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, and 0.1% Triton X-100). This sample 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 10-cycle PCR product, 100 pmol of forward primer, 25 pmol of reverse primer, 10 nmol of each dNTP, 20 μCi of $\alpha$-$^{32}$P-dCTP (800 Ci/mmol), and 5 μL of 10× Taq polymerase buffer. 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.

### 4.4.4 Multiple-Turnover Assay using Malachite Green Dye

A 20 μL sample containing 200 pmol (10 μM) of deoxyribozyme 8HJ102, 250 pmol (12.5 μM) of free DNA anchor oligonucleotide lacking the HEG linker, and 10 nmol (500 μM) of free AAAY$^{P}$AA hexapeptide in 70 mM HEPES, pH 7.5, X mM $ZnCl_2$, and 150 mM NaCl was incubated at 37 °C. X was varied from 0.5 mM, 1 mM, 1.5 mM and 2 mM. At the 24 h time-point, 75 μL of water and 20 μL of malachite green assay solution (BioAssay Systems POMG-25H) was added to the solution. After incubation at room temperature for 30 min, the sample was diluted with 180 μL of water, and the absorbance at 620 nm was measured (NanoDrop 2000c, Thermo Scientific). Standard curves for inorganic phosphate ($P_i$) were obtained using solutions containing known amounts of $P_i$ and all other assay components except the free hexapeptide. Numbers of turnovers were calculated directly from the ratio of mole amounts of released Pi and deoxyribozyme.

### 4.4.5 Cell Culture LNCaP

LNCaP cells were grown and maintained in RPMI 1640, supplemented with 10% FBS. Cells were maintained in T75 Cell Culture flasks (Corning Plastic). For detachment from the flask, 5 mL of sterile Trypsin-EDTA (0.25% trypsin, 1mM EDTA) was treated with the cells at 37 °C for 5 min. 5 mL of cell media was added to stop the trypsin cleavage process. The cells were centrifuged at 250 g for 3 minutes to spin down the cells. The supernatant was aspirated and then 5 mL of fresh media was added to the centrifuged cells. Cells were resuspended in the media by a gentle vortex process. Cells were counted using a hemocytometer and an optical microscope.

In a 24-well plate (Corning® Costar®), 50000 cells were transferred and additional media was supplemented to maintain a total volume of 0.5 mL per well. After 24 h, appropriate plasmids were transfected in each well. jetPRIME (Polyplus) was used as the *in vitro* transfection agent. 50 µL of jetPRIME buffer was mixed with 0.5 µg of DNA. 0.5 µL of jetPRIME was then added to the buffered mixture. After vortexing for 10 s, the sample was spun down and incubated at room temperature for 10 min. This was added dropwise to a single cell. Plasmids pcDNA3.1-PTP1B, pCDNA3.1-PTP1B-D181A and pEGFP were transfected to the LNCaP cells. Plasmid pEGFP was a generous gift from Jessie Peh from the Hergenrother laboratory. Transfection efficiency was investigated by verifying GFP expression using fluorescence microscopy (Figure 4.11).

For downstream investigation with qRT-PCR, the cells were detached from the 24-well using the following procedure. First, the media (after an appropriate time) were aspirated carefully with a pipette. Then 200 µL of trypsin was added to the well to detach the cells from the well. After 3 min, 200 µL of RPMI 1640 medium is added to arrest the

trypsinization. The detached cells were centrifuged at 1000 g for 3 min. The supernatant was aspirated. Cells will be investigated for gene expression using qRT-PCR.

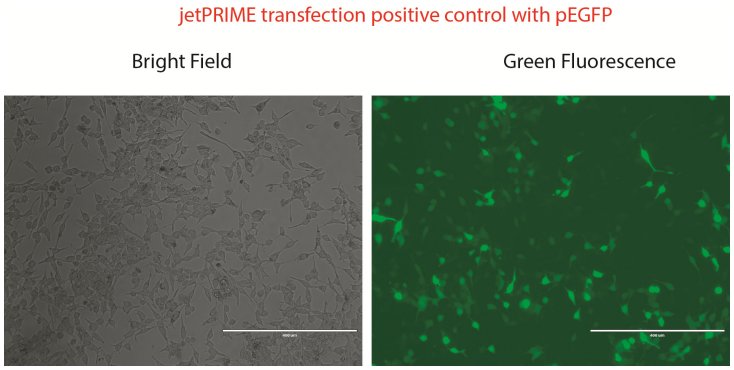


**Figure 4.11.** pEGFP transfection was successful using jetPRIME (Polyplus) transfection reagent. Left panel shows the bright field while the right panel indicates fluorescence field.

### 4.4.6 Site-Directed Mutagenesis of PTP1B D181A

QuikChange protocol was partially followed to perform site-directed mutagenesis. The PCR reactions were performed using the mutagenic primers 5′- ATACCACATGGCCT-GCCTTTGGAGTCCCTGA-3′ (forward) and 5′-TCAGGGACTCCAAAGGCAGGCCATGTGGTAT-3′ (reverse). PCR was performed in 50 µL of 1× Pfu polymerase buffer with 100 pmol of each primer, 25 ng of pcDNA3.1-PTP1B plasmid, 0.5 µL of dNTPs, and 2.5 units of Pfu polymerase. The PCR was performed following the procedure: 95 °C for 30 s, 55 °C for 1 min, 68 °C for 10 min for 30 cycles. The sample was then purified by phenol/chloroform extraction. The purified PCR sample was then digested with DpnI in a 45 µL reaction with 1×NEB buffer 4 and 10 units of DpnI (New England Biolab). The DpnI digestion product was purified on 1% agarose gel.

500 ng of of pcDNA3.1-PTP1B plasmid and the PCR sample were separately double digested in 20 µL reaction with 1×NEB buffer 2 and 10 units of EcoRI and 10 units of BamHI (New England Biolab). Only the pCNDA3.1-PTP1B sample was further

treated with Calf Intestinal Alkaline Phosphatase (CIP). The reaction was brought to a volume of 30 μL with 1×CIP buffer and 10 units of CIP and incubated at 37 °C for 1 h. Both samples were purified by 1% agarose gel. The empty vector was isolated from the PCDNA3.1-PTP1B sample while the PTP1B-D181A fragment was purified from the other sample. The isolated fragments were mixed in 1:10 (vector to insert) ratio. Sample was ligated in a 30 μL containing 5 ng of vector, 50 ng of insert (PTP1B-D181A), 1× T4 DNA ligase buffer (Fermentas) and 10 units of T4 DNA ligase (Fermentas) at 25 °C for 16 h.

DH5α cells were transformed with the ligated product. 5 μL of digestion reaction was mixed with 50 μL of chemically competent cells, and incubated on ice for 30 min. For transformation, the mixture was heat shocked at 42 °C for 1 min, and shaken horizontally in 500 μL of SOC medium at 220 rpm 37 °C for 1 h. The cells (200 μL) were spread on LB + ampicillin agar plates, and incubated at 37 °C overnight. Individual colonies were picked and cultured in 5 mL of LB medium containing 50 μg/mL of ampicillin overnight. Plasmids were harvested via mini-prep as described earlier in Chapter 2. Using T7 promoter the sequences mutant pcDNA3.1-PTP1B-D181A was confirmed.

### 4.4.7 Synthesis of Gold Nanoparticles and Functionalization with Deoxyribozymes

A 200 mL two-neck round-bottom flask was first cleaned with aqua regia for 30 min. The flask was then washed multiple times with Millipore water. 98 mL of Millipore water and 2 mL of 50 mM of $HAuCl_4$ solution were mixed in the clean flask. The flask was heated to reflux with a reflux condenser and a stopper. On refluxing, 10 mL of 38.8 mM sodium citrate as quickly added to the reaction mixture. The color of the solution

changed to deep red within a minute. The reflux was continued for another 20 minutes. The solution was allowed to cool to room temperature while stirring. The size and shape of the resulting nanoparticles were characterized scanning transmission electron microscopy (TEM). 200 kV (JEOL 2100 Cryo TEM). (Figure 4.12) The diameter of the nanoparticles was consistent (approximately 13 nm). The concentration of the gold nanoparticle was 1.9 nM as determined by UV-Vis absorbance. Extinction coefficient of gold nanoparticles (AuNP) is $1.85 \times 10^8\ M^{-1}cm^{-1}$.

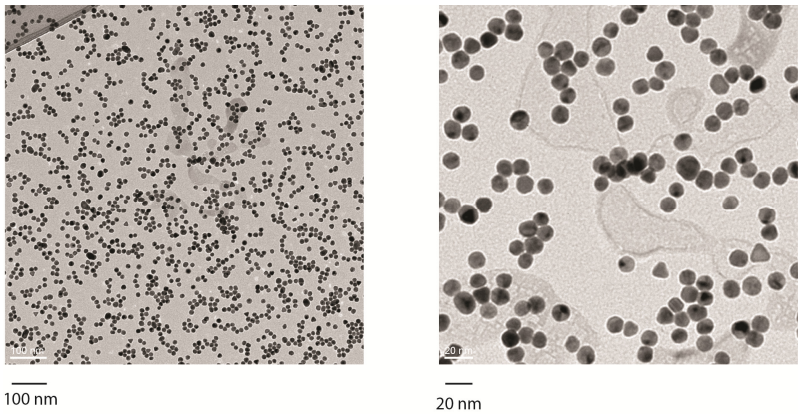


**Figure 4.12.** TEM image of gold nanoparticles. The particles were of consistent size 13 nm.

14WM9 and 8HJ102 were prepared with two phosphorothioate modifications at both the 5′ and 3′ end to prevent exonuclease digestion in cytosol.[21] 14WM9 had the sequence: C*G*AAGCGCTAGAACATCGCAGCGAAATCGGTCTTTATAGGGGCTGTCCTCCGACGG-ATAGTGAGTCGTAT*T*A-HEG$_3$-C$_3$-S-S-C$_3$-OH and 8HJ102 had the sequence C*G*AACGAA-AGCCTCCTTCCGCAGACCATGCATAGGGGCTGTCCCTCCGACATACGCATAAAGGTAGT*C*C-HEG$_3$-C$_3$-S-S-C$_3$-OH, where * indicates inter-nucleotide phosphorothioate modification.

To reduce the disulfide linkage in the initial disulfide linkage 9 µL of 1 mM deoxyribozyme (9 nmol) was mixed with 1.5 µL of 10 mM TCEP (tris(2-carboxyethyl)phosphine) in 500 mM sodium acetate buffer (pH 5.2). After 1 h, the

reduced deoxyribozyme was slowly added to 3 mL of gold nanoparticle solution in a clean vial while gently shaking. The glass vial was treated with 10% NaOH followed by several washing with Millipore water prior to adding the gold nanoparticle solution. The vial was stored in dark for 24 h. 30 μL of 500 mM Tris acetate (pH 8.2) and 300 μL of 1M NaCl were added dropwise sequentially with gentle shaking. The vial was again stored in the dark for 24 h before the next step.

Two buffer solutions were prepared; 25 mM Tris acetate (pH 8.2) with 100 mM NaCl called buffer A, and 20 mM MOPS (pH 7.0) with 100 mM NaCl called buffer B. Deoxyribozymes conjugated to gold nanoparticles (Drz-AuNP) were transferred to a low-bind tube. 500 μL of Drz-AuNP was centrifuged at 16,100 g for 20 min. The supernatant containing unbound deoxyribozyme was removed. Particles were redispersed in 200 μL of buffer A and spun down twice (at 16,100 g) to remove the excess unbound DNA. The Drz-AuNP were then dispersed in 200 μL of buffer B. The supernatants containing the unbound DNA were combined and quantified. The amount of deoxyribozyme attached per gold nanoparticle was estimated. Gold nanoparticle concentration dispersed in solution B can be quantified by utilizing the extinction coefficient at 520 nm.

Anchor (helper) oligonucleotides for 14WM9 [5′-G*G*ATAATACGACTCACT*A*T-3′] and 8HJ102, [5′-G*G*ACTACCTTTATGCGT*A*T-3′] were synthesized with internucleotide phosphorothioate modifications. .After estimating the amount of bound deoxyribozyme, helper oligonucleotides (desalted by PD-10 desalting column GE Healthcare) were added in 1:2 ratio excess. The mixture was heated at 65 °C for 5 min. The sample was allowed to cool down to room temperature for an hour. Immediately, the solution was stored at 4 °C overnight to allow full hybridization (assumed to be 100% efficient). After the

overnight incubation, excess helper oligonucleotide was removed by centrifuging the sample with buffer B thrice with 200 μL buffer B at 16,100 g. Finally the sample was dispersed in 200 μL of 20 mM HEPES (pH 7.5) with 100 mM NaCl.

### 4.4.8 qRT-PCR Plans

GeneJET RNA Purification Kit (K0731) was used to isolate RNA from the cells Integrity of RNA was verified by a bleach agarose gel (Figure 4.13).

RNA was reverse transcribed using the cDNA Syntehsis Kit (ThermoFisher Scientific K1671). Following primers will be used for qRT-PCR assay based on the report by Wu et al.[16], all 5′ to 3′.

PTP1B: GGAGATGGAAAAGGAGTTC (f), TGCTTTTCTGCTCCCACAC (r)

NSE: AGCTGCCCCTGCCTTAC (f), AGCTGCCCCTGCCTTAC (r)

CGA: GCGGTGGAAGAGCCATCAT (f), TCTGTGGCTTCACCACTTTTCTC (r)

B-Actin: GCGGGAAATCGTGCGTGACATT (f), GATGGAGTTGAAGGTAGTTTCGTG (r)

GADPH: ACCACAGTCCATGCCATCAC (f), TCCACCACCCTGTTGCTGT (r)

For the preliminary qRT-PCR test, a 20 μL reaction solution containing the 10 ng cDNA and PerfeCTa SYBR Green FastMix (Quanta Biosciences) was first subjected to UDG incubation at 50 °C for 2 min. Hot start PCR was initiated by heating the reaction solution at 95 °C for 10 min. A cycle containing 95 °C for 10s, 60 °C for 30s was performed for a total of a total 45 rounds, followed by a melt curve cycle of 95 °C for 15 s, 55 °C for 15 s, and 95 °C for 15 s. Eco Real-Time PCR System (Illumina) was used for sample analysis.

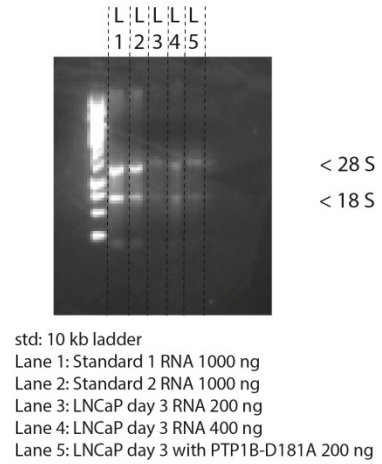


**Figure 4.13.** The RNA integrity was checked on a 1% agarose gel with 1% v/v bleach. A) RNA standard-1 1000 ng, B) RNA standard-2 1000 ng, C) LNCaP RNA 200 ng, D) LNCaP 400 ng, E) LNCaP with PTP1B-D181A 200ng.

## 4.5 References

(1) Joyce, G. F. RNA cleavage by the 10-23 DNA enzyme. *Methods Enzymol.* **2001**, *341*, 503-517.

(2) Khachigian, L. M. DNAzymes: cutting a path to a new class of therapeutics. *Curr. Opin. Mol. Ther.* **2002**, *4*, 119-121.

(3) Wu, P. W.; Hwang, K. V.; Lan, T.; Lu, Y. A DNAzyme-Gold Nanoparticle Probe for Uranyl Ion in Living Cells. *J. Am. Chem. Soc.* **2013**, *135*, 5254-5257.

(4) Torabi, S. F.; Wu, P.; McGhee, C. E.; Chen, L.; Hwang, K.; Zheng, N.; Cheng, J.; Lu, Y. In vitro selection of a sodium-specific DNAzyme and its application in intracellular sensing. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 5903-5908.

(5) Jin, L.; Zeng, X.; Liu, M.; Deng, Y.; He, N. Current progress in gene delivery technology based on chemical methods and nano-carriers. *Theranostics* **2014**, *4*, 240-255.

(6) Jewell, C. M.; Zhang, J.; Fredin, N. J.; Wolff, M. R.; Hacker, T. A.; Lynn, D. M. Release of plasmid DNA from intravascular stents coated with ultrathin multilayered polyelectrolyte films. *Biomacromolecules* **2006**, *7*, 2483-2491.

(7) Nayerossadat, N.; Maedeh, T.; Ali, P. A. Viral and nonviral delivery systems for gene delivery. *Adv. Biomed. Res.* **2012**, *1*, 27.

(8) Ibraheem, D.; Elaissari, A.; Fessi, H. Gene therapy and DNA delivery systems. *Int. J. Pharm.* **2014**, *459*, 70-83.

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

(10) Jiang, S.; Eltoukhy, A. A.; Love, K. T.; Langer, R.; Anderson, D. G. Lipidoid-coated iron oxide nanoparticles for efficient DNA and siRNA delivery. *Nano Lett.* **2013**, *13*, 1059-1064.

(11) Wang, X. Y.; Feng, M. L.; Xiao, L.; Tong, A. J.; Xiang, Y. Postsynthetic Modification of DNA Phosphodiester Backbone for Photocaged DNAzyme. *ACS Chem. Biol.* **2016**, *11*, 444-451.

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

(13) Costello, L. C.; Liu, Y. Y.; Zou, J.; Franklin, R. B. Evidence for a zinc uptake transporter in human prostate cancer cells which is regulated by prolactin and testosterone. *J. Biol. Chem.* **1999**, *274*, 17499-17504.

(14) Zhang, X. A.; Hayes, D.; Smith, S. J.; Friedle, S.; Lippard, S. J. New strategy for quantifying biological zinc by a modified zinpyr fluorescence sensor. *J. Am. Chem. Soc.* **2008**, *130*, 15788-15789.

(15) Ghosh, S. K.; Kim, P.; Zhang, X. A.; Yun, S. H.; Moore, A.; Lippard, S. J.; Medarova, Z. A novel imaging approach for early detection of prostate cancer based on endogenous zinc sensing. *Cancer Res.* **2010**, *70*, 6119-6127.

(16) Wu, C. Y.; Zhang, L.; Bourne, P. A.; Reeder, J. E.; di Sant'Agnese, P. A.; Yao, J. L.; Na, Y. Q.; Huang, J. T. Protein tyrosine phosphatase PTPIB is involved in neuroendocrine differentiation of prostate cancer. *Prostate* **2006**, *66*, 1125-1135.

(17) Pedersen, A. K.; Guo, X. L.; Moller, K. B.; Peters, G. H.; Andersen, H. S.; Kastrup, J. S.; Mortensen, S. B.; Iversen, L. F.; Zhang, Z. Y.; Moller, N. P. Residue 182 influences the second step of protein-tyrosine phosphatase-mediated catalysis. *Biochem J.* **2004**, *378*, 421-433.

(18) Selner, N. G.; Luechapanichkul, R.; Chen, X. W.; Neel, B. G.; Zhang, Z. Y.; Knapp, S.; Bell, C. E.; Pei, D. H. Diverse Levels of Sequence Selectivity and Catalytic Efficiency of Protein-Tyrosine Phosphatases. *Biochemistry* **2014**, *53*, 397-412.

(19) Gunawardana, J.; Chan, F. C.; Telenius, A.; Woolcock, B.; Kridel, R.; Tan, K. L.; Ben-Neriah, S.; Mottok, A.; Lim, R. S.; Boyle, M.; Rogic, S.; Rimsza, L. M.; Guiter, C.; Leroy, K.; Gaulard, P.; Haioun, C.; Marra, M. A.; Savage, K. J.; Connors, J. M.; Shah, S. P.; Gascoyne, R. D.; Steidl, C. Recurrent somatic mutations of PTPN1 in primary mediastinal B cell lymphoma and Hodgkin lymphoma. *Nat Genet* **2014**, *46*, 329-335.

(20) 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.

(21) Wu, P.; Hwang, K.; Lan, T.; Lu, Y. A DNAzyme-gold nanoparticle probe for uranyl ion in living cells. *J. Am. Chem. Soc.* **2013**, *135*, 5254-5257.

# Chapter 5: Catalytic DNA with Phosphoserine and Phosphothreonine Lyase Activity[6]

## 5.1 Introduction

Peptides or proteins rarely form natural electrophilic sites.[1-3] An important exception to this observation is dehydroalanine (Dha) and dehydrobutyrine (Dhb).[1] Dehydroalanine is formally created from serine (Ser) by a net elimination of water.[4-5] Dha and related functional groups are intermediates or products in the biosynthesis of a wide range of cyclic peptides, including lantibiotics,[6-7] thiopeptides,[8-10] microcystins,[11-12] and other natural products.[11-12] When introduced synthetically into a peptide or protein, Dha enables preparation of many valuable natural and artificial modifications, usually via thiol addition.[13-14] One biosynthetic pathway for generating Dha or Dhb involves activation of Serine or Threonine with a phosphate, followed by a secondary β-elimination step to reveal Dha and Dhb. This secondary β-elimination step is also known as phosphoserine (pSer) and phosphothreonine (pThr) lyase activity.

### 5.1.1 Comparison with Natural Phosphoserine/Phosphothreonine Lyases

Proteins with pSer/pThr lyase activity in nature have important biological roles. For instance, the OspF family of pThr lyase (e.g. SpvC from Salmonella) inactivates the phosphorylated host MAPKs irreversibly through a β-elimination reaction.[15] Families of lantibiotic synthetases (class III and IV) have specific lyase domains that are highly homologous to the OspF family pThr lyase.[16] A genome-wide search in human proteins

---

[6] This research has been published: 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.

for similar homology led to annotation of β-elimination function for the BLES03 protein whose function was initially unknown.[17] Comparison with the crystal structure of the OsPF family lyase with lantibiotic VenL lyase is consistent with a mechanism that involves histidine stabilizing the phosphomonoester *O*-bond while lysine residue functions as the general base to abstract the proton from the peptide backbone to help perform the β-elimination (Figure 5.1).[15,18,19]

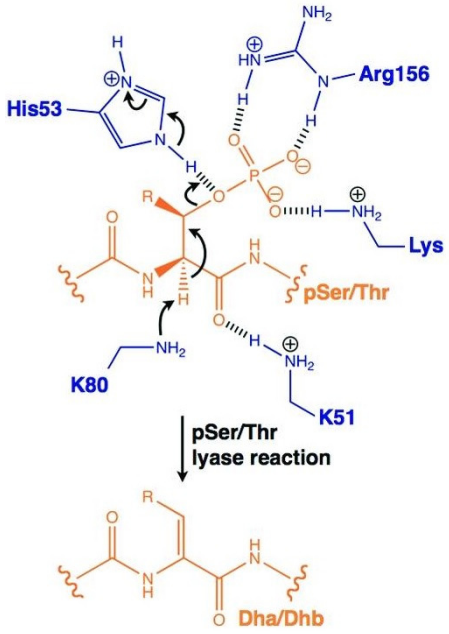


**Figure 5.1.** Histidine, arginine and lysine stabilize the phosphate leaving group. Lysine (K80) serves as the general base to abstract the proton from the peptide backbone to enable β elimination (figure from reference 19).

### 5.1.2 Bioactive Natural Products and Phosphoproteomics

The β-elimination products of pSer/pThr are also known as α,β-dehydroamino acids. To date, about 37 different structural units are known of which Dha/Dhb is the main class.[1] Typically isolated from bacteria, and less often from fungi, marine organisms or plants, these compounds often exhibit important biological activities

(Figure 5.2A). pSer/pThr lyase activity can also be achieved chemically via base-catalyzed elimination of phosphate. However, such chemical methods are harsh with potential for substantial side reactions to be useful in a broader context. Nonetheless, base-treated chemical "pSer/pThr Lyase" reactions have received consideration as phosphopeptide enrichment tools in phosphoproteomics (Figure 5.2B).

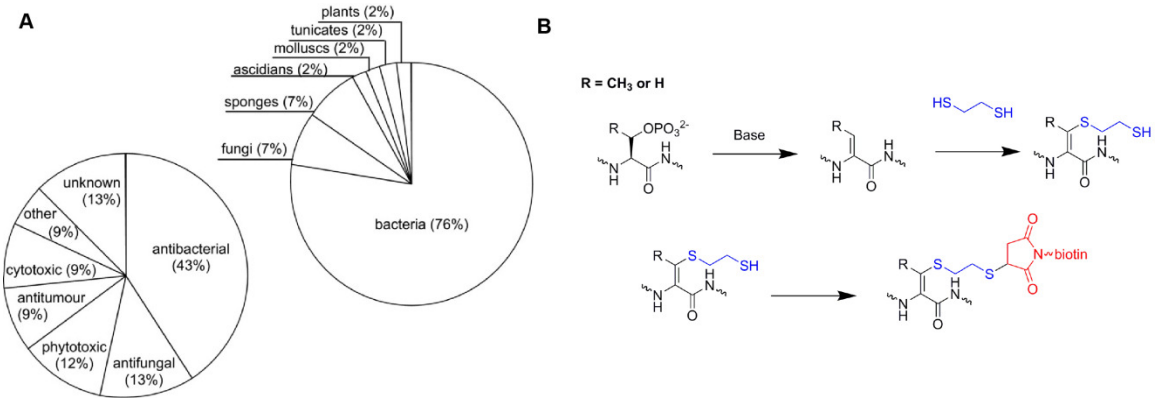


**Figure 5.2.** A) Pie charts illustrates the percentage of natural product containing α,β-dehydroamino acid, including their typical biological uses (figure from reference 1). B) Chemical β-elimination of phosphate from pSer and pThr is used as an enrichment tool in phosphoproteomics. The Dha or Dhb can be treated with ethanedithiol, followed by a reaction with biotin maleimide. Biotinylated peptides are captured using streptavidin beads.

### 5.1.3 Limitations to Current Synthesis of Dha and Dhb

Dha-containing peptides and proteins can be formed via base-catalyzed elimination of phosphate from pSer-containing substrates (Figure 5.2B); e.g., by treatment with strong base (~1 M hydroxide) for phosphoproteomics analysis.[20-22] However, such relatively harsh conditions lead to numerous side reactions and are not typically used to prepare Dha-containing products. Davis and coworkers have surveyed the use of various more sophisticated reagents for conversion of cysteine (Cys) to Dha, concluding that a bis-alkylation-elimination strategy is optimal.[13-14, 23] However, any exposed Cys will react with such reagents, which therefore cannot readily be used when

the newly created Dha is intended for subsequent cyclization with a second Cys residue. Without a gentle and general route available, one must resort to difficult or lengthy synthetic approaches to install Dha or to obviate its need.[13-14, 24] This is especially true considering that the Dha monomer for solid-phase peptide synthesis (SPPS) is—after nitrogen deprotection—a simple enamine that presents substantial difficulties for traditional SPPS, necessitating use of alternative Dha-related monomers (Figure 5.3).[25-27] Hence, this chapter focuses on a chemoenzymatic method to achieve synthesis of Dha from pSer utilizing deoxyribozymes.

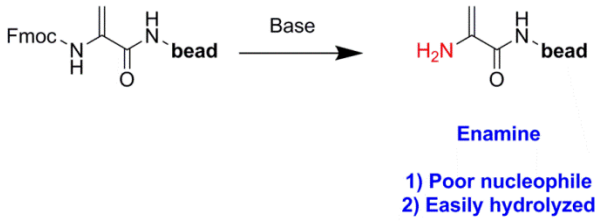


**Figure 5.3.** The *Fmoc* protecting group has to be deprotected using the base (20% piperidine). The next amino acid coupling is inefficient because of the low coupling efficiency of the unmasked enamine. Enamines are also easily hydrolyzed to aldehydes leading to peptide truncations.

## 5.2 Results and Discussion

### 5.2.1 In Vitro Selection Strategy for pSer Lyase Activity

To perform in vitro selection for pSer lyase deoxyribozymes, we needed to develop a reliable "capture" reaction,[28] by which specific DNA sequences that convert pSer to Dha are physically separable from the vast majority of catalytically inactive sequences. This capture was achieved with a simple 5′-thiol oligonucleotide that reacts via Michael-type addition with Dha, increasing the mass of only the catalytic DNA sequences and enabling their separation (Figure 5.4). The capture yield was ~40%, which was readily sufficient to enable the selection process. Using this strategy, in vitro selection was implemented using a DNA-anchored AAAS$^{P}$AA model peptide (S$^{P}$ = pSer)

that was attached via the DNA anchor's remote 5′-end to the deoxyribozyme pool, which included an $N_{40}$ region (40 random nucleotides). During each round's key selection step, 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 for 14 h. In many of our previous selection experiments, the ions $Zn^{2+}$, $Mn^{2+}$, and $Mg^{2+}$ were each effective deoxyribozyme cofactors, sometimes alone and sometimes in combination. Following the selection and capture steps, PAGE-separated sequences were amplified by PCR and ligated to the DNA-anchored pSer peptide substrate for the next selection round.

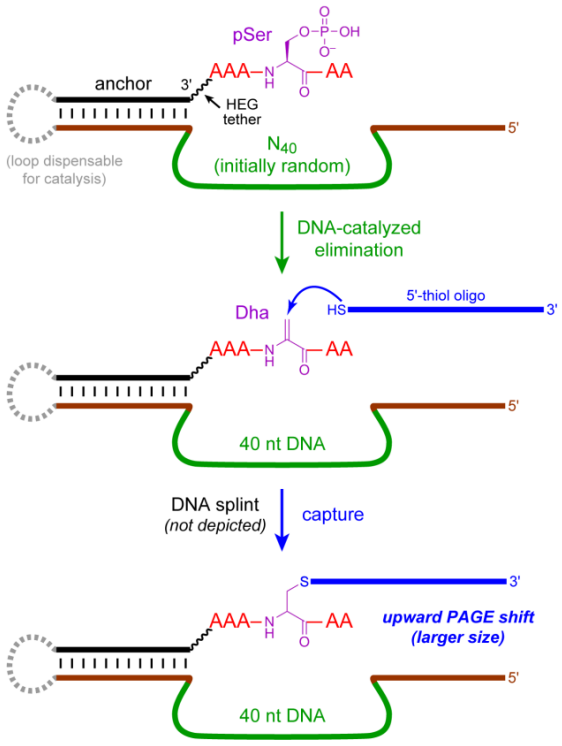


**Figure 5.4.** In vitro selection of pSer lyase deoxyribozymes. DNA-catalyzed elimination of phosphate from the pSer residue in the DNA-anchored peptide, forming Dha, is followed by DNA-splinted capture of Dha using a 5′-thiol oligonucleotide, which results in an upward PAGE shift for only the active DNA sequences. See Figure 5.12 for details, including depiction of the hexa(ethylene glycol) [HEG] tether connecting the DNA anchor to the peptide and validation of the capture reaction. The dashed loop on the left enables the selection process but is dispensable for catalysis.

DNA-catalyzed pSer lyase activity of the pool was 15% at round 9, which is relative to the maximum (control) capture yield of ~40% (Figure 5.5) and individual deoxyribozymes were cloned. In parallel, selection in 50 mM CHES, pH 9.0, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C showed no activity through round 10, and this second selection experiment was discontinued. This experiment was performed because the higher pH could have fostered more efficient phosphate elimination, but catalysis was not observed (note that $Zn^{2+}$ and $Mn^{2+}$ cannot be used at the higher pH because of precipitation and elimination, respectively). From the round 9 pool, two unrelated 76 nt deoxyribozymes were identified (see Materials and Methods for sequences) and named DhaDz1 and DhaDz2, for Dha-forming Deoxyribozymes 1 and 2.

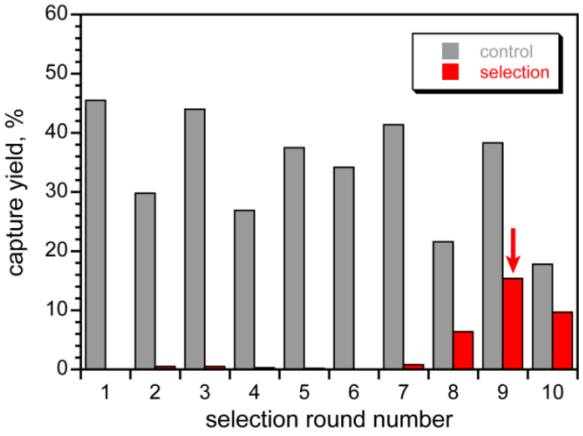


**Figure 5.5.** Progression of the in vitro selection experiment. In each round, "control" refers to the yield for the splint capture reaction using the DNA-anchored AAA(Dha)AA substrate, and "selection" refers to the yield for the splint capture reaction using the deoxyribozyme pool for that round. The arrow marks the cloned round. Round 1 pool yield was not measured because the pool was not radiolabeled.

### 5.2.2 Single-Turnover Assay with DhaDz1 and DhaDz2

DhaDz1 catalyzed single-turnover $Zn^{2+}/Mn^{2+}$-dependent elimination of phosphate from the DNA-anchored $AAAS^{P}AA$ substrate in high yield when this substrate was not attached to the deoxyribozyme (Figure 5.6). Although catalysis strictly required only

$Zn^{2+}$ and $Mn^{2+}$, optimal activity was observed only in the additional presence of $Mg^{2+}$. DhaDz2 catalyzed the same elimination reaction, strictly requiring only $Zn^{2+}$ but with greater activity in the additional presence of $Mn^{2+}$, $Mg^{2+}$, or both of these ions. For DhaDz1 in the presence of each of 1.5 mM $Zn^{2+}$, 20 mM $Mn^{2+}$, and 40 mM $Mg^{2+}$, $k_{obs} = 0.28 \pm 0.01\ h^{-1}$ (n = 3) with 80% yield, For DhaDz2 with 2.0 mM $Zn^{2+}$, 20 mM $Mn^{2+}$, and 40 mM $Mg^{2+}$, $k_{obs} = 0.43 \pm 0.07\ h^{-1}$ (n = 3) with 68% yield.

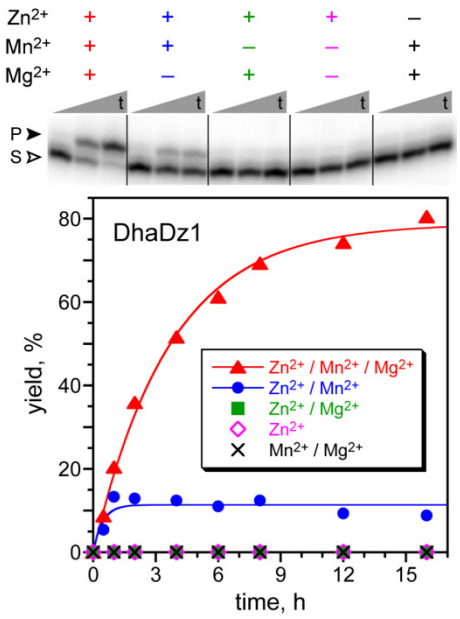


**Figure 5.6.** Single-turnover pSer lyase activity of the DhaDz1 deoxyribozyme with DNA-anchored model peptide AAAS$^{P}$AA. S = pSer substrate; P = Dha product. Incubation conditions: 70 mM HEPES, pH 7.5, combinations of 1.5 mM $ZnCl_2$, 20 mM $MnCl_2$, and 40 mM $MgCl_2$ as appropriate, and 150 mM NaCl at 37 °C (10 nM 5′-$^{32}$P-radiolabeled DNA-anchored peptide, 1 μM DhaDz1). The PAGE image shows timepoints at $t$ = 30 s, 4 h, and 16 h. The $Zn^{2+}$ concentration of 1.5 mM was optimal at 20 mM $Mn^{2+}$ and 40 mM $Mg^{2+}$. At 1.5 mM $Zn^{2+}$ in the absence of $Mg^{2+}$, a higher yield of 55% was observed at 10 mM $Mn^{2+}$ (data not shown). Data for the DhaDz2 deoxyribozyme is in Figure 5.8A.

MALDI mass spectrometry validated the pSer lyase activities of DhaDz1 and DhaDz2. For both DhaDz1 and DhaDz2, omitting the deoxyribozyme or using a random-sequence oligonucleotide in place of the catalytic sequences led to no activity. Removal of the 18 nt fixed-sequence 5′-segment of DhaDz1 or DhaDz2 led to only ~10% yield in

16 h in both cases. This finding suggests that during the selection process, some portion of the fixed-sequence 5′-segment was recruited for folding or catalysis by the initially random $N_{40}$ region.

### 5.2.3 Multiple-Turnover Assay with DhaDz1 and DhaDz2

DhaDz1 and DhaDz2 were assessed for multiple-turnover activity with the free (discrete; not DNA-anchored) model AAAS$^{P}$AA peptide substrate. Each deoxyribozyme was assayed both by colorimetric detection of the released inorganic phosphate using malachite green dye (Figure 5.7A) and by LC-MS of the peptide mixture (Figure 5.7B). DhaDz1 achieved up to 6–7 turnovers in 96 h as determined by both assays.

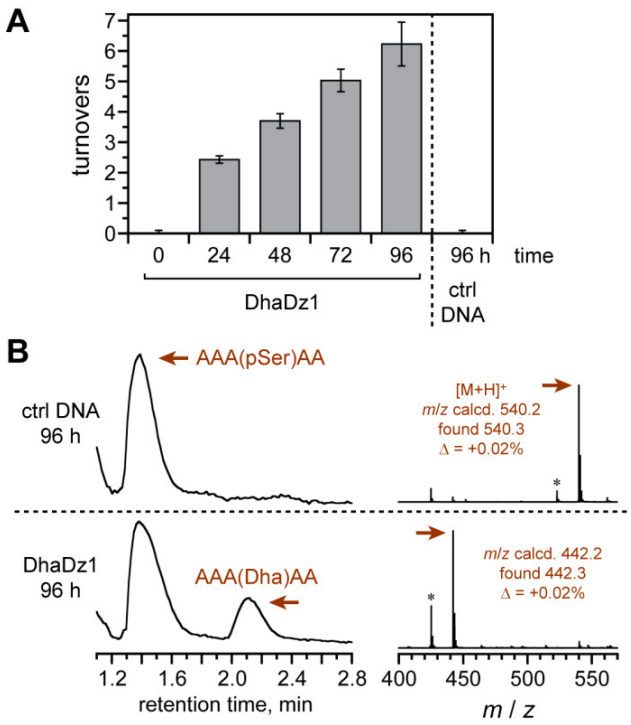


**Figure 5.7.** Multiple-turnover pSer lyase activity of the DhaDz1 deoxyribozyme with free model peptide AAAS$^{P}$AA. The incubation conditions were as in Figure 5.6, except with 6 mM $Zn^{2+}$ as optimized to accommodate the high DNA and peptide concentrations (100 μM DhaDz1, 4 mM peptide). A) Malachite green assay for released inorganic phosphate. At each timepoint, $n = 3$ (± sd); "ctrl DNA" refers to an inactive negative control of arbitrary sequence. B) LC-MS assay (TIC; ESI-MS of indicated peaks). The observed 18% yield is equivalent to 7.2 turnovers. The asterisks in the MS denote loss of ammonia from the peptide N-terminus.[29]

pSer lyase activity was observed only when the 19 nt fixed-sequence 3′-segment of DhaDz1 was part of a Watson-Crick duplex with the separate anchor oligonucleotide; either deletion of the 3′-segment or omission of the anchor led to no reaction (dye assay, 40:1 peptide:DhaDz1, <0.1 turnovers). The single-turnover activity of DhaDz2 was also determined (Figure 5.8A). DhaDz2 also showed multiple-turnover activity (4 turnovers in 96 h; Figure 5.8B).

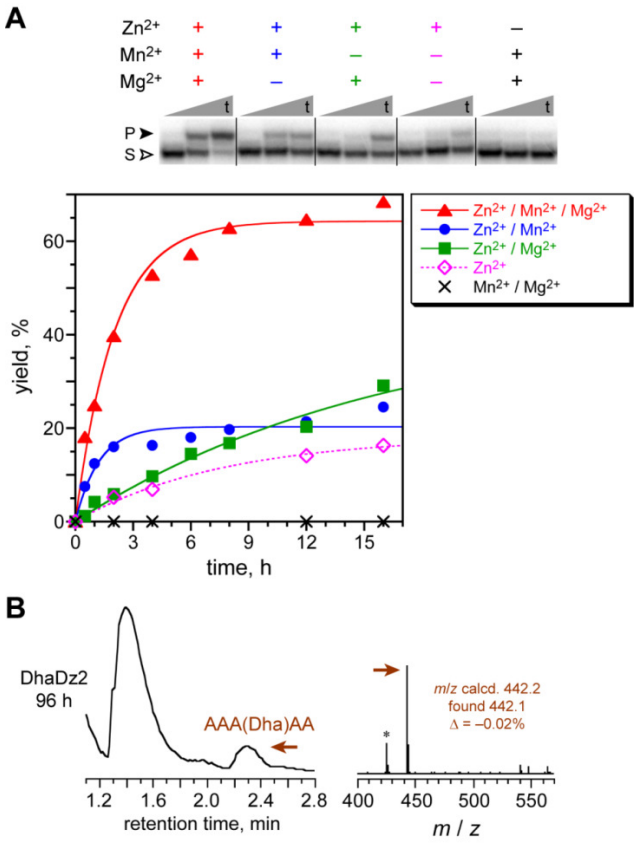


**Figure 5.8.** Single-turnover and multiple-turnover data for the DhaDz2 deoxyribozyme. A) Single-turnover pSer lyase activity of the DhaDz2 deoxyribozyme with DNA-anchored peptide AAASPAA. S = pSer substrate; P = Dha product. Incubation conditions: 70 mM HEPES, pH 7.5, combinations of 2.0 mM $ZnCl_2$, 20 mM $MnCl_2$, and 40 mM $MgCl_2$ as appropriate, and 150 mM NaCl at 37 °C (10 nM 5′-$^{32}$P-radiolabeled DNA-anchored peptide, 1 µM DhaDz2). The PAGE image shows timepoints at $t$ = 30 s, 4 h, and 16 h. The $Zn^{2+}$ concentration of 2.0 mM was optimal at 20 mM $Mn^{2+}$ and 40 mM $Mg^{2+}$. B) Multiple-turnover pSer lyase activity of the DhaDz2 deoxyribozyme with free peptide AAAS$^{P}$AA. LC-MS experiment as in Figure 5.7B, but with 8 mM $Zn^{2+}$. The observed 10% yield is equivalent to 4.0 turnovers.

### 5.2.4 Synthesis of Cystathionine-containing Compstatin Peptide using DhaDz1

For practical utility, a pSer lyase deoxyribozyme must function with a realistic peptide substrate sequence, rather than the model peptide used during selection. Compstatin is a 13 amino acid cyclic peptide, identified by phage display, which inhibits the proteolytic activation of human complement component C3 and may have therapeutic value.[30-33] Disulfide-containing cyclic peptides such as compstatin are stabilized by replacing the disulfide bond with a cystathionine bridge, i.e., –S–$CH_2$– in place of –S–S–. This replacement is typically accompanied by minimal loss of bioactivity,[34-35] including for compstatin.[32] Here we used DhaDz1 to enable synthesis of a more stable cystathionine analogue of compstatin (Figure 5.9A). DhaDz1 was found to catalyze preparative-scale (nanomole) single-turnover formation of Dha within a compstatin-derived linear peptide precursor that includes pSer and homocysteine (Hcy) in place of the two Cys in the native sequence (Figure 5.9B for LC-MS analysis). The linear Dha/Hcy intermediate was then efficiently cyclized by simple base treatment to the compstatin analogue in 70% two-step yield. Analysis of HPLC-purified intermediate and cyclic product after iodoacetamide treatment (to alkylate free cysteine thiol groups) by MALDI mass spectrometry confirmed their structures (Figure 5.9C). Because the cyclization step is not enzyme-catalyzed, both (*R*,*S*) and (*S*,*S*) diastereomers (i.e., epimers at the initially pSer residue) of the compstatin analogue are presumably formed. Formation of diastereomers upon thiol addition into Dha is well-recognized and not an inherent impediment to a wide range of experiments that exploit the products of such a reaction.[13-14, 36-37]

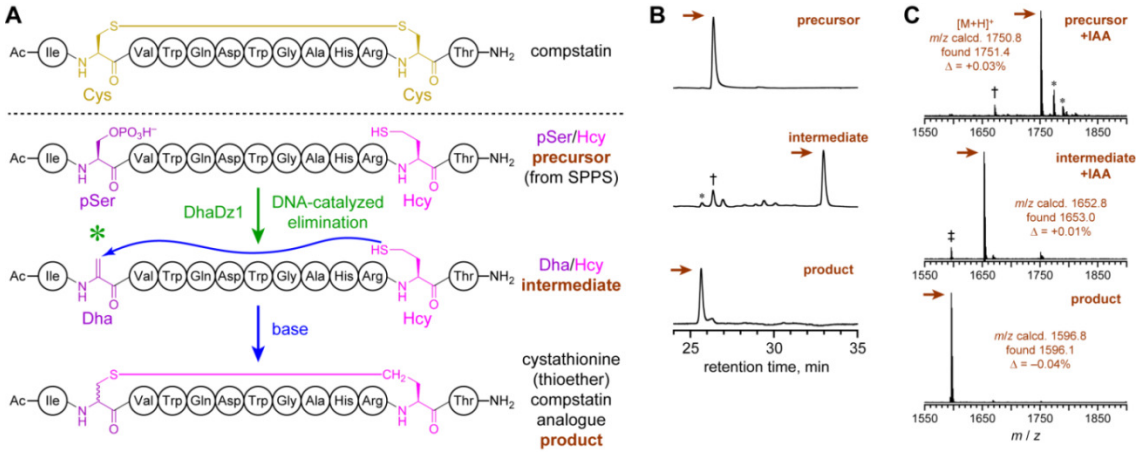


**Figure 5.9.** Application of the DhaDz1 pSer lyase deoxyribozyme for synthesis of a more stable compstatin analogue. A) Compstatin, and its stable cystathionine (thioether) analogue as prepared using DhaDz1. One of the two possible pSer/Hcy combinations is shown. B) HPLC analysis of the reaction process. The three traces show (top to bottom) the HPLC-purified linear precursor pSer/Hcy peptide, the crude linear intermediate Dha/Hcy peptide after treatment of the precursor with DhaDz1, and the crude cyclic product (compstatin thioether analogue) after subsequent base treatment of the HPLC-purified intermediate Dha/Hcy peptide. Single-turnover conditions for Dha formation: 100 μM DhaDz1, 125 μM untethered DNA anchor oligo, 50 μM precursor peptide, 70 mM HEPES, pH 7.5, 4 mM $ZnCl_2$ (optimized), 20 mM $MnCl_2$, 40 mM $MgCl_2$, 150 mM NaCl, 37 °C, 96 h. Conditions for cyclization: 100 mM Tris, pH 8.0, 25 °C, 4 h (shorter times not attempted). The two-step yield was 70%. For the intermediate, the asterisk denotes autocyclization product, the dagger denotes unreacted precursor, and unmarked peaks are side products not specifically identified. C) MALDI mass spectrometry of HPLC-purified precursor, intermediate, and product, each after iodoacetamide (IAA) treatment to modify any free thiol groups (+57; one free thiol in each of precursor and intermediate; no free thiol in product).[38] For the precursor, the asterisks denote $[M+Na]^+$ and $[M+K]^+$, and the dagger denotes hydrolysis of pSer. For the intermediate, the double dagger denotes autocyclization product.

### 5.2.5 Testing for Sequence Generality of DhaDz1

To explore further the substrate generality of DhaDz1, deoxyribozyme conversion of pSer into Dha within four unrelated pSer-containing heptapeptides was tested. Specifically, heptapeptides derived from lantibiotic component cytolysin $CylL_L''$ (Val-pSer-Leu),[39-40] inhibitor of HIV replication T22 (Lys-pSer-Tyr),[41] the α7 nAChR antagonist conotoxin ImI (Arg-pSer-Ala),[34] and the lantibiotic lacticin 481 (Asn-pSer-Trp) were used.[3a] For the first three peptides, single-turnover yields of approximately 95%, 50%, and 30%, respectively, in 48 h were obtained on nanomole scale, after optimizing the $Zn^{2+}$ concentration in the 2 to 5 mM range (MALDI-MS; Figure 5.10). For

the lacticin peptide, no yield (<5%) was observed, perhaps due to the Trp adjacent to pSer; further investigation is required. Separately, the other cystathionine variant of Compstatin was tested, but no activity was observed with this substrate (see Materials and Methods). It is however important to note that these yields are determined based on MALDI-MS peak intensities. These MS signals can significantly differ from the actual ratio of substrate and product in the sample. Therefore, we should be cautious about the accuracy of yield determined on this basis.

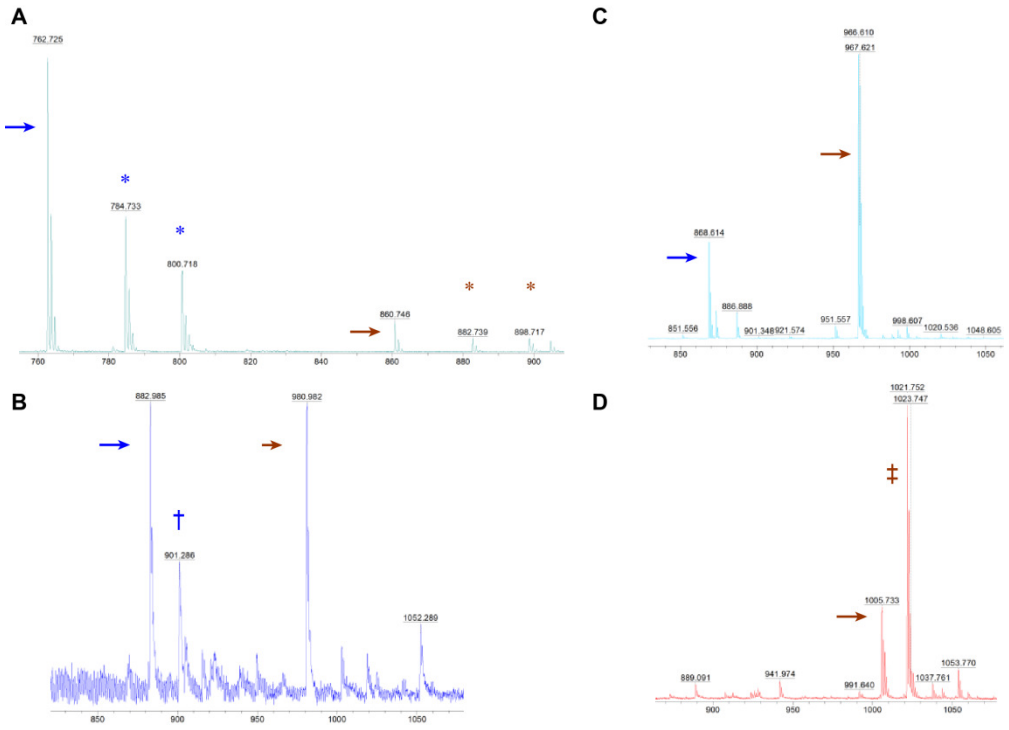


**Figure 5.10.** MALDI-MS for, A) Cytolysin (Val-Val-pSer-Leu-Lys-His). B) T22 (Tyr-Arg-Lys-pSer-Tyr-Lys-Gly). C) Conopeptide (Asp-Pro-Arg-pSer-Ala-Trp-Arg). D) Lacticin (Asn-Met-Asn-pSer-Trp-Gln-Phe). Conditions for Dha formation: 100 µM DhaDz1, 125 µM untethered DNA anchor oligo, 50 µM precursor peptide, 70 mM HEPES, pH 7.5, X mM $ZnCl_2$ (optimized), 20 mM $MnCl_2$, 40 mM $MgCl_2$, 150 mM NaCl, 37 °C, 96 h. Yield was found to be optimal with 5 mM $Zn^{2+}$ (Val-pSer-Leu), 5 mM $Zn^{2+}$ (Lys-pSer-Tyr) and 3 mM $Zn^{2+}$ (Arg-pSer-Ala). No activity was observed with the Asn-pSer-Trp peptide. For the precursor (brown), the asterisks denote $[M+Na]^+$ and $[M+K]^+$, and the double dagger denotes oxidation of methionine. For the product, the blue dagger denotes the hydrolysis peak (spontaneous MS product). See table 5.1 for the mass spectrometry values.

### 5.2.6 New Selections for Phosphothreonine Lyase Deoxyriboymes

Although we have confined our attention here to conversion of pSer to Dha, analogous DNA-catalyzed reaction of pThr to form Dhb should be feasible. The new DhaDz1 and DhaDz2 deoxyribozymes do not function with a pThr substrate; the additional methyl group of pThr suppresses activity. This is unsurprising because the selection experiment that led to DhaDz1 and DhaDz2 was performed with pSer rather than pThr. Selections that directly use pThr-containing phosphopeptides are used to identify pThr lyase deoxyribozymes as well. We have validated that the capture approach of Figure 5.4 is successful with Dhb in place of Dha. (See Materials and Methods for synthesis of the Dhb standard)

### 5.2.7 New Strategy for Long 5′-thiol Capture

Higher capture yield was observed with the short 5′-thiol oligonucleotide ($n$ = 0 in Figure 5.12B; 18-mer) as compared with the long 5′-thiol oligonucleotide ($n$ = 6 in Figure 5.12B; 54-mer). This outcome was attributed to some of the long 5′-thiol oligonucleotide lacking the intended 5′-thiol group, likely due to relatively inefficient coupling of the thiol phosphoramidite (due to partial hydrolysis) along with poor 12% PAGE resolution for such a lengthy oligonucleotide. After this selection experiment was concluded and individual deoxyribozymes were already identified, we prepared a sequence variant of the long 5′-thiol oligonucleotide by splint ligation starting with the 20% PAGE-purified short 5′-thiol oligonucleotide. This variant of the long 5′-thiol oligonucleotide had capture yield equivalent to that of the short 5′-thiol oligonucleotide (ca. 40%).

### 5.2.8 Selection Strategy for Sequence General pSer/pThr Lyases[7]

DhaDz1 couldn't function promiscuously with each pSer-containing peptide we tested for generality (Section 5.2.6). This result is however not too surprising. Pressure was not applied during the selection process to encourage sequence-tolerant pSer lyase activity. Merely performing selections with DNA-anchored AAAS$^P$AA (alanine-rich peptide) is insufficient to achieve promiscuity. Applying a systematic strategic selection pressure may be necessary to obtain sequence generality.

To address this situation, new selection efforts have been designed. Peptides AKKXKKA, AEEXEEA, AFFXFFA were synthesized (where X = S$^P$ or T$^P$) and conjugated to the DNA anchor as by reductive amination (check Materials and Methods for lysine). In the new efforts to identify general DNA catalysts, we are imposing selection pressure by maintaining the pSer or pThr epitope while systematically changing the peptide identity every round of selection; also termed as the "alternation strategy" (Figure 5.11).

When capture with thiol-containing oligonucleotide was tested with E-rich Dha/Dhb-containing peptide, suboptimal capture yield was observed (≤ 10%). This could be attributed to negative charges from the glutamate (E) repelling the "capture" thiolate and preventing Michael-type addition. With K-rich peptide however, high background activity was observed with even pSer or pThr (>15%, which would be a big headwind for selection processes). This can be hypothesized to be an $SN_1$ displacement of pSer/pThr with a thiol oligonucleotide. For these reasons, selections utilizing E-rich and K-rich

[7] University of Illinois undergraduate Adam Wyler initiated the selection process for sequence generality under my supervision.

peptides have placed been on hold. However, alternation strategy is still viable with A-rich and F-rich pSer/pThr-containing peptides. Separately, a biologically useful fragment from cytolysin, GIT$^{P}$PVAF,[42] was also tested as a selection substrate. However, efforts to capture the Dhb standard of cytolysin fragment were unsuccessful. We hypothesize that the proline adjacent to Dhb may provide steric or directional inhibition to attack by a 5′-thiol capture oligonucleotide (proline conformation).

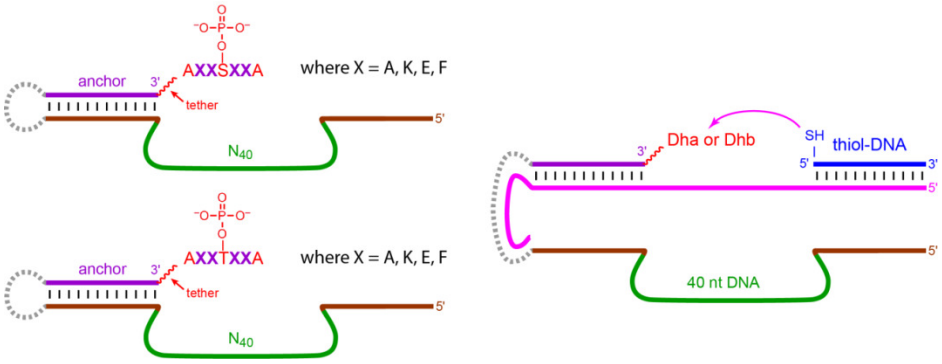


**Figure 5.11.** The strategy for identifying sequence-general deoxyribozymes involves changing the peptitic portion in the DNA-anchored peptide each selection round. The internal pSer/pThr amino acid is maintained constant every selection round, but the identity of the surrounding residues (X) is changed to F: aromatic, E: negatively charged, K: positively charged.

### 5.2.9 Phosphoproteomics

Although DhaDz1 and DhaDz2 exhibit much sequence tolerance, we can envision the use of such sequence-general pSer/pThr lyases in phosphoproteomics as enrichment tools. However, additional selections would be necessary to obtain powerful deoxyribozymes for such purpose. DNA has the following advantages, A) Deoxyribozymes can be used in high stoichiometric ratio (excess to the substrate). B) A mixture of DNA catalysts (cocktail) can be used to achieve a sequence-general even if a single deoxyribozyme doesn't function as a universally general pSer/pThr lyase enzyme. C) Choice of metal cation cofactors is not a limitation for *in vitro* use. D) Acidic peptides

such as glutamates and aspartates will not be enriched, which is a disadvantage of IMAC (immobilized metal affinity chromatography) technology. DNA enzymes are expected to be chemoselective.

DNA will have to overcome three fundamental challenges to be useful in phosphoproteomics. DNA catalysts are required to be sequence general. In addition, deoxyribozymes will require excellent catalytic proficiency, i.e. good values of $k_{cat}$ and $K_m$. The more important factor would be $K_m$ because phosphorylated peptides are expected to be in lower abundance (pico-, femto-, atto-moles). $k_{cat}$ to some extent can be addressed by increasing the enzymatic incubation time.

To demonstrate the proof-of-concept and applicability of deoxyribozymes, two proteins were chosen as test cases, chicken ovalbumin and β-casein. Both proteins serve as excellent source of pSer-containing proteins. Chicken ovalbumin has two pSer (mostly phosphorylated with 2 equivalents of phosphate) whereas β-casein can have about 5 phosphorylated sites with substantial phosphorylation variability (0-5 phosphate equivalents relative to the protein). Bottom-up proteomics involves analysis of typtic digested fragments from protein. On the other hand, top-down approach involves analysis of intact proteins followed by mass-spectrometry induced fragmentation. Both proteins were separately digested with trypsin, desalted on a $C_{18}$ column (solid-phase extraction). Known amounts of digested and full-length proteins were treated with alkaline phosphatase and analyzed by malachite green phosphate assay to quantify the amount of inorganic phosphate released. This would reveal the exact amount of phosphorylation in the protein or protein digest (assuming complete dephosphorylation of the protein by the alkaline phosphatase). Although β-elimination would lead to formation of Dha or Dhb,

inorganic phosphate can be quantified by the same technique. Relative quantification of inorganic phosphate will provide the yield of pSer/pThr Lyase for both the digested and full-length protein (see Materials and Methods). With DhaDz1 negligible activity was observed with either the full-length protein or the digested fragment.

## 5.3 Summary and Future Directions

In this chapter, we utilized enzymatic catalysis rather than reagent-based synthetic approaches to form Dha from a readily introduced precursor. The natural protein enzymes that create Dha residues proceed via an activated serine intermediate, such as phosphoserine (pSer)[4-5] or glutamylated serine.[43] Biosynthesis of lanthipeptides involves enzyme-catalyzed elimination reactions.[44] However, these enzymes often require subsequently removed leader peptides on their substrates, and in general lanthipeptide dehydratases are not preparatively useful beyond their narrow natural sequence contexts.[45] We therefore targeted *de novo* identification of entirely new enzymes for formation of Dha residues.

Using in vitro selection, we identified DhaDz1 and DhaDz2, pSer lyase deoxyribozymes that eliminate phosphate from pSer in a peptide substrate to form Dha. This elimination is a fundamentally new DNA-catalyzed reaction, which expands the repertoire of reactions known to be catalyzed by DNA. DhaDz1 was used to synthesize a Dha-containing linear peptide intermediate to a cystathionine analogue of compstatin, a cyclic inhibitor of complement activation. With further development of pSer and pThr lyase deoxyribozymes to improve the elimination rate and yield and to establish peptide sequence generality beyond compstatin, we anticipate that such DNA catalysts will be useful for enzymatically converting pSer/pThr residues in peptide substrates into the

corresponding Dha/Dhb residues. This ability will enable a broad range of experiments in biological chemistry, both via study of natural products and their analogues as well as by site-specific labeling of peptides and proteins at readily and specifically engineered electrophilic sites. pSer/pThr lyase deoxyribozymes may also find utility in phosphoproteomics experiments.[20-22]

For DNA to function with sequence generality, selections will involve an alternation strategy involving different DNA-anchored phosphopeptides in each selection round. To improve the catalytic proficiency of the pSer/pThr Lyase deoxyribozymes, modified nucleotides with protein-inspired functional groups can be used to augment the catalytic prowess of DNA. In protein OspF pThr lyases, lysine acts as a general base in proton abstraction. Histidine assists with β-elimination of the phosphate by making it a better leaving group. dUTPs with amino allyl and imidazole groups at the C5-heterocyclic position can be can be used in future selection endeavors.[46] Metal cation cofactors that bind tightly with phosphate, such as lanthanides, may also permit efficient DNA-catalyzed β-elimination of phosphate. For applications in phosphoproteomics, DNA catalysts will be required to have excellent $K_m$ values while turnover is not prohibitive for such applications. Ability of DNA to efficiently recognize the minimal epitope of pSer and pThr in any peptidic context/environment would be highly desirable. This should be further complemented by an efficient a proton abstraction from the peptidic backbone resulting in β-elimination. Two model proteins, chicken ovalbumin and β-casein, were arranged for preliminary phosphoproteomics analyses with DhaDz1 with negligible activity. Newly identified pSer/pThr lyase deoxyribozymes in the future

will be directly tested on both full-length protein and digested tryptic fragments from these protein samples.

## 5.4 Materials and Methods

### 5.4.1 Oligonucleotide and Oligonucleotide Conjugates

DNA oligonucleotides and oligonucleotide conjugates were obtained from Integrated DNA Technologies (Coralville, IA) or prepared by solid-phase synthesis on an ABI 394 instrument using reagents from Glen Research. All oligonucleotides and conjugates 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), extracted from the polyacrylamide with TEN buffer (10 mM Tris, pH 8.0, 1 mM EDTA, 300 mM NaCl), and precipitated with ethanol as described previously.[47-48]

Deoxyribozyme sequences were as follows (initially random $N_{40}$ regions underlined):

DhaDz1 5′-CGAATTAAGACTGAATTCACAGGTGGCGGATGTAGTTTACCCGTTTTCTGTAGAGCCATAA-GAAACGAGATATTCC-3′

DhaDz2 5′-CGAATTAAGACTGAATTCACAGGATTCAAAGTGGCTGTCAGAAGTTGGTGAGGGAAGATAA-GAAACGAGATATTCC-3′

### 5.4.2 Synthesis of Peptides and DNA-Anchored Peptides

All amino acid monomers including those for phosphoserine (pSer, $S^P$), homocysteine (Hcy), dehydrobutyrine (Dhb), trifluoroacetyl-lysine ($K^{Tfa}$) and *S*-methylcysteine ($Cys^{Me}$, $C^{Me}$) were obtained from Chem-Impex (Wood Dale, IL). Peptides were prepared by solid-phase synthesis using Fmoc Rink amide MBHA resin from Chem-Impex, with *N,N,N′,N′*-tetramethyl-*O*-(7-azabenzotriazol-1-yl)uronium hexa-fluorophosphate **(**HATU) as coupling agent.[49]

Synthesis of the peptide AAAS$^{P}$AA was performed at 0.2 mmol scale, initiated using 260 mg of Rink amide resin with a loading capacity of 0.77 mmol/g. All steps were monitored by ninhydrin test. For each coupling, 5 equivalents (1.0 mmol) of Fmoc-amino acid, 4.9 equivalents (373 mg, 0.98 mmol) of HATU, and 10 equivalents (350 μL, 2.0 mmol) of *N*,*N*-diisopropylethylamine (DIPEA) were mixed in 5 mL of anhydrous DMF. The coupling reaction was initiated by adding this mixture to the resin and agitating by bubbling with nitrogen for 60 min, followed by washing with DMF (3 × 10 mL). The N-terminus of the newly installed amino acid was deprotected by agitating the resin in 20% piperidine in DMF (5 mL) under nitrogen for 5 min a total of three times, each time washing with DMF (3 × 5 mL). The peptide was cleaved from the solid support by stirring the resin in a separate vial with a solution containing 5 mL of trifluoroacetic acid (TFA), 125 μL of water, and 50 μL of triisopropylsilane for 90 min. The liquid solution was separated from the resin by filtration. This solution was dried on a rotary evaporator, providing a solid. To this material 20 mL of cold diethyl ether was added, and the peptide was obtained as a white solid that was filtered and purified by HPLC.

Synthesis of each of the peptides AAAC$^{Me}$AA, AAA(Dhb)AA, Ac-IS$^{P}$VWQDWGAHR(Hcy)T (compstatin precursor 1), Ac-I(Hcy)VWQDWGAHRS$^{P}$T (compstatin precursor 2) was performed as described above, with the following procedural modifications. Synthesis was performed at 0.1 mmol scale, initiated using 303 mg of 0.33 mmol/g (Hcy) or 130 mg of 0.77 mmol/g (Dhb or C$^{Me}$) Rink amide resin. For each coupling of a standard amino acid, 5 equivalents (0.5 mmol) of Fmoc-amino acid, 4.9 equivalents (187 mg, 0.49 mmol) of HATU, and 10 equivalents (175 μL, 1.0 mmol) of DIPEA were mixed in 3 mL of anhydrous DMF and agitated with the resin for 60 min.

For Fmoc-*S*-trityl-L-homocysteine (Hcy), Fmoc-α,β-dehydro-2-aminobutyric acid or Fmoc-*S*-methyl-L-cysteine ($C^{Me}$), 2 equivalents (0.2 mmol) of Fmoc-amino acid, 1.96 equivalents (75 mg, 0.20 mmol) of HATU, and 4 equivalents (70 μL, 0.4 mmol) of DIPEA were mixed in 3 mL of anhydrous DMF and agitated with the resin for 120 min. For the step involving deprotection of Fmoc-α,β-dehydro-2-aminobutyric acid, extra care was taken to utilize dry solvent to avoid hydrolysis of the enamine. The compstatin peptide was capped via acetylation of its N-terminus by agitating the resin in 50 equivalents (462 μL, 5 mmol) of acetic anhydride and 25 equivalents of DIPEA (435 μL, 2.5 mmol) in 5 mL of DMF under nitrogen for 30 min. The peptide was cleaved from the solid support by stirring the resin in a separate vial with a solution containing 3 mL of TFA, 100 μL of water, 100 μL of ethanedithiol (for compstatin), and 40 μL of triisopropylsilane for 90 min.

The $AAAS^{P}AA$ peptide was coupled to the DNA anchor oligonucleotide via the N-terminal α-amino group with a linkage created by reductive amination (Figure 5.12A). The DNA anchor oligonucleotide was 5′-GGAATATCTCGTTTCTTAT-HEG-rA-3′. This sequence includes a hexa(ethylene glycol) [HEG] spacer, which provides structural flexibility and may induce the emergent deoxyribozymes to function well with free (unanchored) peptide substrates, and a 3′-terminal ribonucleotide, which enables the reductive amination process. Coupling of peptide to DNA anchor was performed using the previously described procedure.[49] pSer/pThr or Dhb-containing peptides were similarly coupled to a DNA anchor using reductive amination. For peptides containing lysines, the primary amines were protected as Lys-Tfa. After the *N*-terminal amine was

conjugated by reductive amination and purified on gel, Tfa was removed using ammonia at 55 °C for 4 h and evaporated to dryness.

To prepare the DNA-anchored AAA(Dha)AA peptide, a precursor peptide incorporating *S*-methylcysteine, $Cys^{Me}$ ($C^{Me}$), was synthesized. The $AAAC^{Me}AA$ peptide was conjugated to the 3′-terminal ribonucleotide of the DNA anchor oligonucleotide as described above. Then, a 50 μL sample containing 500 pmol of DNA-anchored $AAAC^{Me}AA$, 100 mM HEPES, pH 7.5, and 10 mM $NaIO_4$ was incubated at room temperature for 1 h to oxidize the thioether to the sulfone. The sulfone product was precipitated by addition of 220 μL of water, 30 μL of 3 M NaCl, and 900 μL of ethanol. The sulfone product was dissolved in 20 μL of 10 mM NaOH, pH 12.0, and incubated at 37 °C for 4 h to form the final elimination product, DNA-anchored AAA(Dha)AA. This final product was precipitated by addition of 250 μL of water, 30 μL of 3 M NaCl, and 900 μL of ethanol. MALDI mass spectrometry for DNA-anchored $AAAC^{Me}AA$ peptide *m*/*z* calcd. 6913.9, found 6915.5, $\Delta$ = +0.02%. DNA-anchored AAA(Dha)AA elimination product *m*/*z* calcd. 6865.8, found 6864.8, $\Delta$ = –0.01%.

*Procedure for validation of Dha capture reaction during in vitro selection process.* The 5′-terminus of the DNA-anchored $AAAC^{Me}AA$ precursor peptide was ligated to the 3′-terminus of the random deoxyribozyme pool using a DNA splint and T4 DNA ligase (see *Procedure for ligation step*, below). The ligated product was oxidized and eliminated as described above, forming DNA-anchored AAA(Dha)AA linked to the deoxyribozyme pool. A 10 μL sample containing 0.1 pmol of DNA-anchored AAA(Dha)AA linked to the pool, 20 pmol of capture splint, and 100 pmol of the short 5′-thiol capture oligonucleotide

($n = 0$ in Figure 5.12B) in 100 mM $Na_2CO_3$, pH 9.2, was incubated at 37 °C for 12 h, with result shown in Figure 5.12C.

### 5.4.3 In Vitro Selection Procedure

The selection procedure, cloning, and initial screening of individual clones were performed essentially as described previously.[49] An overview of the key selection and capture steps of each round is shown in Figure 5.4. A depiction of the capture step with nucleotide details and validation data is shown in Figure 5.12. The random deoxyribozyme pool was 5′-CGAATTAAGACTGAATTC-$N_{40}$-ATAAGAAACGAGATATAGCTGATCC-TGATGG-3′. PCR primers were 5′-CGAATTAAGACTGAATTC-3′ (forward primer) and 5′-$(AAC)_4$XCCATCAGATCAGCT-3′, where X is the HEG spacer to stop Taq polymerase (reverse primer).

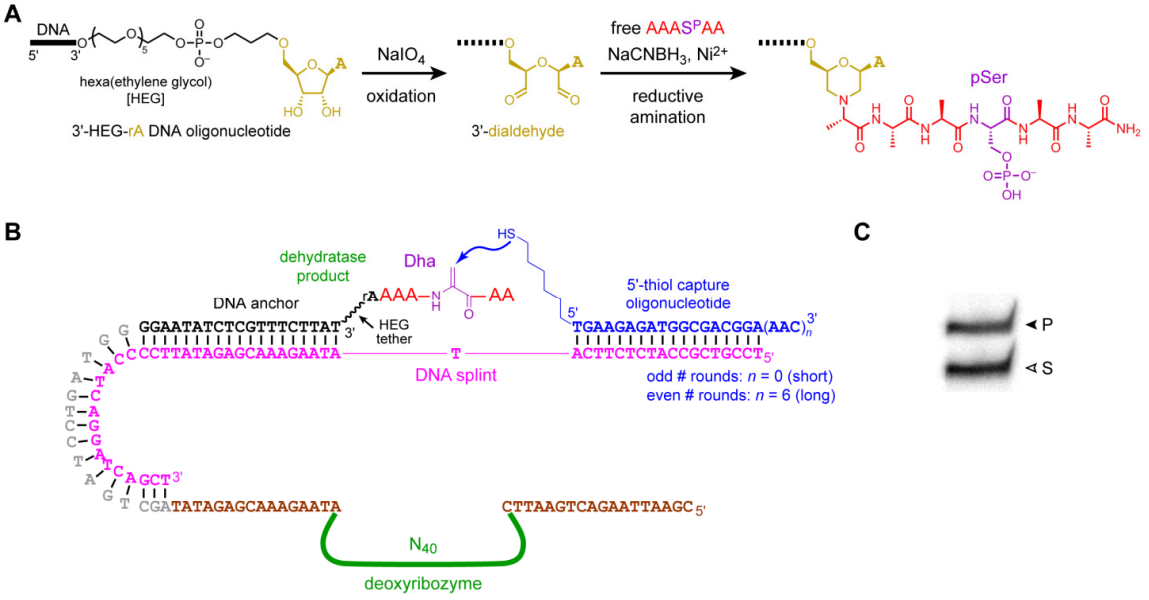


**Figure 5.12.** A) Synthesis of DNA-anchored AAAS$^P$AA peptide by attachment of the DNA anchor to the N-terminal α-amino group, with the linkage created by reductive amination. B) Capture step during in vitro selection. C) Validation of the capture reaction using the short 5′-thiol capture oligonucleotide (S = substrate, P = product).

In each round, the ligation step to attach the deoxyribozyme pool at its 3′-end with the 5′-end of the DNA-anchored hexapeptide substrate was performed using a DNA splint and T4 DNA ligase. The ligation splint sequence was 5′-ATAAGAAACGAGATATTCCTCCATCAGGATCAGCTATATCTCGTTTCTTAT-3′, where the underlined T is included to account for the untemplated A nucleotide that is added at the 3′-end of each PCR product by Taq polymerase. This T nucleotide was omitted from the splint used for ligation of the initially random $N_{40}$ pool, which was prepared by solid-phase synthesis without the untemplated A. The length of the capture oligonucleotide was alternated between shorter (odd-numbered rounds) and longer (even-numbered rounds), to avoid selection of particular DNA sequences that migrate aberrantly at a fixed position on the gel. Nucleotide sequences of the DNA anchor oligonucleotide, the deoxyribozyme binding arms, the splint oligonucleotide, and the two 5′-thiol capture oligonucleotides are shown in Figure 5.12B.

*Procedure for ligation step in round 1.* A 34 μL sample containing 1 nmol of DNA pool, 850 pmol of DNA splint, and 750 pmol of 5′-phosphorylated DNA-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 (Fermentas) and 2 μL of 5 U/μL T4 DNA ligase (Fermentas). The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

*Procedure for ligation step in subsequent rounds.* A 16 μL sample containing the PCR-amplified DNA pool (~5–10 pmol), 25 pmol of DNA splint, and 50 pmol of 5′-phosphorylated DNA-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 (Fermentas) and 2 μL of 1 U/μL T4 DNA ligase (Fermentas). The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

*Procedure for selection step in round 1.* The selection experiment was initiated with 200 pmol of the ligated $N_{40}$ pool. A 20 μL sample containing 200 pmol of ligated $N_{40}$ pool 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 sample was incubated at 37 °C for 14 h.

*Procedure for selection step in subsequent rounds.* A 10 μL sample containing ligated pool 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. The sample was incubated at 37 °C for 14 h.

*Procedure for preparation of 5′-thiol capture oligonucleotide.* The 5′-thiol capture oligonucleotide (short or long version as appropriate; Figure 5.12B) was prepared by DTT reduction of the precursor oligonucleotide (prepared using the Glen Research thiol modifier C6 S-S) immediately prior to each use. A 50 μL sample containing 2 to 5 nmol of precursor oligonucleotide in 100 mM HEPES, pH 7.5, and 25 mM DTT was incubated

at 37 °C for 2 h. To remove excess DTT, the sample was precipitated by adding 220 μL of water, 30 μL of 3 M NaCl, and 900 μL of ethanol.

*Procedure for capture step in round 1.* To the 40 μL selection sample was added 6 μL of 0.5 M EDTA, pH 8.0 (3 μmol). The DNA was precipitated by adding 224 μL of water, 30 μL of 3 M NaCl, and 900 μL of ethanol. The sample was then brought to 40 μL total volume containing 300 pmol of capture splint, and 500 pmol of short 5′-thiol capture oligonucleotide in 100 mM $Na_2CO_3$, pH 9.2 and incubated at 37 °C for 12 h. The sample was loaded directly on 8% PAGE.

*Procedure for capture step in subsequent rounds.* To the 20 μL selection sample was added 3 μL of 0.5 M EDTA, pH 8.0 (1.5 μmol). The DNA was precipitated by adding 247 μL of water, 30 μL of 3 M NaCl, and 900 μL of ethanol. The sample was then brought to 10 μL total volume containing 20 pmol of capture splint, and 100 pmol of the appropriate 5′-thiol capture oligonucleotide (long for even-numbered rounds, short for odd-numbered rounds) in 100 mM $Na_2CO_3$, pH 9.2 and incubated at 37 °C for 12–14 h. The sample was loaded directly on 8% PAGE.

*Procedure for PCR.* 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-purified selection product, 200 pmol of forward primer, 50 pmol of reverse primer, 20 nmol of each dNTP, and 10 μL of 10× Taq polymerase buffer (1× = 20 mM Tris-HCl, pH 8.8, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, and 0.1% Triton X-100). This sample 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 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), and 5 μL of 10× Taq polymerase buffer. 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. The forward and reverse single-stranded PCR products were separable because formation of the reverse product was initiated using the reverse primer, which contains a nonamplifiable spacer that stops Taq polymerase.

*Procedure for cloning and initial screening.* Using 1 μL of a 1:1000 dilution of the 10-cycle PCR product from selection round 9, 25-cycle PCR was performed using the analogous procedure as described above for 30-cycle PCR, omitting α-$^{32}$P-dCTP and using primers 5′-CGAATTAAGACTGAATTC-3′ (forward primer, 25 pmol) and 5′-TAATTAATTAATTACCCATCAGGATCAGCT-3′ (reverse primer, 25 pmol), where the extensions with TAA stop codons in each frame were included to suppress false negatives in blue-white screening.[50] The PCR product was purified on 2% agarose and cloned using a TOPO TA cloning kit (Invitrogen). Initial screening of individual deoxyribozyme clones was performed using DNA strands prepared by PCR from miniprep DNA derived from individual *E. coli* colonies. The miniprep DNA samples were first assayed by digestion with EcoRI to ascertain the presence of the expected insert. The concentration of each PAGE-purified deoxyribozyme strand was estimated from the UV shadowing intensity relative to suitable standards. Each screening assay used ~0.1 pmol of 5′-$^{32}$P-radiolabeled DNA-anchored AAAS$^{P}$AA substrate and at least 20 pmol of deoxyribozyme in the

single-turnover assay procedure described in a subsequent section of this document (here using 1.0 mM $ZnCl_2$).

### 5.4.4 Selection Progression

In vitro selection was also performed with 30-nucleotide and 50-nucleotide random regions (JG1 and JK1 respectively). The DNA anchored oligonucleotide lacked the HEG-type tether. Activity was observed in round 8 for JG1. However, the sequences were inactive when individual clones were screened suggesting that the activity depends on the loop formed by the T4 DNA ligase. No activity was obtained for JK1 and both these selections were not pursued any further.

### 5.4.5 Single-Turnover Procedure Assay

The 5′-end of DNA-anchored $AAAS^{P}AA$ was radiolabeled using γ-$^{32}$P-ATP and Optikinase (USB), which lacks the 3′-phosphatase activity that we observed also dephosphorylates the pSer side chain. A 10 μL sample containing 0.1 pmol of 5′-$^{32}$P-radiolabeled DNA-anchored $AAAS^{P}AA$ and 20 pmol of 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 DNA-catalyzed reaction was initiated by bringing the sample to 20 μL total volume containing 70 mM HEPES, pH 7.5, 1.5 or 2.0 mM $ZnCl_2$ if included, 20 mM $MnCl_2$ if included, 40 mM $MgCl_2$ if included, and 150 mM NaCl, and the sample was incubated at 37 °C. Note that the $Zn^{2+}$ concentration was carefully optimized for each deoxyribozyme. 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 = $Y$•$(1 - e^{-kt})$, where $k = k_{obs}$ and $Y$ is the final yield. Each $k_{obs}$ value is reported with error calculated as the standard deviation from the indicated number of independent determinations.

### 5.4.6 Multiple-Turnover Assay Procedure with Malachite Green Dye

A 20 μL sample containing 2 nmol of deoxyribozyme, 2.5 nmol of free DNA anchor oligonucleotide, and 80 nmol of AAAS$^{P}$AA hexapeptide in 70 mM HEPES, pH 7.5, 6 or 8 mM $ZnCl_2$ as appropriate, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl was incubated at 37 °C. Note that the $Zn^{2+}$ concentration was carefully optimized for each deoxyribozyme. At each time point (0–96 h), 2 μL of the sample was frozen on dry ice and stored at –80 °C until further processing. To the sample was added 98 μL of water and 20 μL of malachite green assay solution (BioAssay Systems POMG-25H). After incubation at room temperature for 30 min, the sample was diluted with 180 μL of water, and the absorbance at 620 nm was measured (NanoDrop 2000c, Thermo Scientific). Standard curves for inorganic phosphate ($P_i$) were obtained using solutions containing known amounts of $P_i$ and all other assay components with 2 nmol of an unrelated, catalytically inactive 74 nt DNA oligonucleotide instead of the deoxyribozyme. Numbers of turnovers were calculated directly from the ratio of mole amounts of released $P_i$ and deoxyribozyme.

### 5.4.7 Multiple-Turnover Assay Procedure with LC-MS

A 20 μL sample containing 2 nmol of deoxyribozyme, 2.5 nmol of free DNA anchor oligonucleotide, and 80 nmol of peptide in 70 mM HEPES, pH 7.5, 6 or 8 mM $ZnCl_2$ (optimized), 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl was incubated at

37 °C for 96 h. To remove the deoxyribozyme before LC-MS analysis, a filtration step was performed. An Amicon Ultra-0.5 mL 3 kDa centrifugal filter (EMD Millipore) was washed with 400 μL of water by centrifugation at 14000 × *g* for 10 min. To the sample was added 400 μL water, and the combined solution was transferred to the filter, which was centrifuged at 14000 × *g* for 20 min. Another 400 μL of water was passed through the filter at 14000 × *g* for 20 min. Finally, 400 μL of 5% aqueous acetonitrile was passed through the filter at 14000 × *g* for 20 min. The combined filtrates were evaporated to dryness, and redissolved in 40 μL of water. The sample was analyzed by LC-MS on an Acquity UPLC instrument with a Phenomenex Jupiter $C_{18}$ column (5 μm, 0.1 × 150 mm) attached to an electrospray ionization mass spectrometer (Q-TOF Synapt-G1 Waters in positive ion scan mode using the manufacturer's suggested parameters) and an isocratic gradient of 1% solvent A (0.1% formic acid in acetonitrile) and 99% solvent B (0.1% formic acid in water) at 0.18 mL/min for 10 min.

### 5.4.8 Procedure for Compstatin Peptide

A 20 μL sample containing 2 nmol of deoxyribozyme, 2.5 nmol of free DNA anchor oligonucleotide, and 1 nmol of peptide in 70 mM HEPES, pH 7.5, 4 mM $ZnCl_2$ (optimized), 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl was incubated at 37 °C for 96 h. The sample was quenched with 3 μL of 0.5 M EDTA, pH 8.0. To reverse formation of potential peptide disulfide dimers, 4 μL of 1 mM tris(2-carboxyethylphosphine) (TCEP; 4 nmol) was added, and the sample was incubated at 25 °C for 15 min. Because compstatin does not pass well through the Amicon filter, the deoxyribozyme was not removed by filtration. The sample was analyzed by HPLC on a Beckman System Gold instrument with a Phenomenex Kinetex $C_{18}$ column (2.6 μm,

4.6 × 150 mm) and a gradient of 15% solvent A (acetonitrile) and 85% solvent B (0.1% trifluoroacetic acid in water) at 0 min to 30% solvent A and 70% solvent B at 60 min with flow rate of 0.5 mL/min.

The HPLC-purified intermediate was cyclized by base treatment. The intermediate was evaporated to dryness, redissolved in 40 μL of 100 mM Tris, pH 8.0, and incubated at 25 °C for 4 h (shorter times not attempted). The sample was analyzed by HPLC as above.

### 5.4.9 Mass Spectrometry Procedure

To prepare the DNA-anchored AAA(Dha)AA using a deoxyribozyme, a 10 μL sample containing 50 pmol of DNA-anchored AAAS$^{P}$AA substrate and 100 pmol of DhaDz1 (or DhaDz2) 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 elimination reaction was initiated by bringing the sample to 20 μL total volume containing 70 mM HEPES, pH 7.5, 1.5 mM $ZnCl_2$ (or 2.0 mM $ZnCl_2$), 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl. The sample was incubated at 37 °C for 16 h and quenched with 5 μL of 0.5 M EDTA, pH 8.0 (2.5 μmol), desalted by Millipore $C_{18}$ ZipTip, and analyzed by MALDI mass spectrometry. Data were acquired on a Bruker UltrafleXtreme MALDI-TOF mass spectrometer with matrix 3-hydroxypicolinic acid in positive ion mode at the UIUC School of Chemical Sciences Mass Spectrometry Laboratory (instrument purchased with support from NIH S10RR027109A).

For the MALDI mass spectrometry analyses of Figure 5.9C, the HPLC-purified precursor, intermediate, or product was evaporated to dryness, redissolved in a total volume of 40 μL containing 100 mM Tris, pH 8.0, 25 mM IAA, and 100 μM TCEP, and

incubated at 25 °C for 2 h in the dark. The sample was desalted and analyzed by MALDI mass spectrometry as described above, except with 2,5-dihydroxybenzoic acid as matrix.

### 5.4.10 Addressing Long 5′-thiol Capture Yield

A 34 μL sample containing 2 nmol of 5′-thiol capture oligonucleotide (short), 2.5 nmol of the splint oligonucleotide 5′-TATCAGTATATGCGATGCTTCCGTCGCCATCTCTTCC-3′, and 3 nmol of 5′-phosphorylated 5′-AGCATCGCATATACTGATAAACAACAACAACAACAAC-3′ 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 (Fermentas) and 2 μL of 5 U/μL T4 DNA ligase (Fermentas). The sample was incubated at 37 °C for 12 h and purified by 12% PAGE to obtain the long 5′-thiol capture oligonucleotide. The sample was treated with DTT (as mentioned earlier) prior to capture use.

### 5.4.11 Procedure for Analysis of Heptapeptides by MALDI for Sequence Generality

Heptapeptides used for studying sequence generality had the following sequences. Lantibiotic component cytolysin $CylL_L''$ was Val-Val-Val-pSer-Leu-Lys-His, inhibitor of HIV replication T22 was Lys-Tyr-Arg--pSer-Tyr-Lys-Gly, the α7 nAChR antagonist conotoxin ImI was Asp-Pro-Arg-pSer-Ala-Trp-Arg and the lantibiotic lacticin 481 was Asn-Met-Asn-pSer-Trp-Gln-Phe. These sequences were purified by HPLC.

A 20 μL sample containing 2 nmol of deoxyribozyme, 2.5 nmol of free DNA anchor oligonucleotide, and 1 nmol of peptide in 70 mM HEPES, pH 7.5, **X** mM $ZnCl_2$ (optimized), 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl was incubated at 37 °C for 96 h. An Amicon Ultra-0.5 mL 3 kDa centrifugal filter (EMD Millipore) was washed with 400 μL of water by centrifugation at 14000 × *g* for 10 min. To the sample was added

400 μL water, and the combined solution was transferred to the filter, which was centrifuged at 14000 × *g* for 20 min. Another 400 μL of water was passed through the filter at 14000 × *g* for 20 min. Finally, 400 μL of 5% aqueous acetonitrile was passed through the filter at 14000 × *g* for 20 min. The combined filtrates were evaporated to dryness. The samples were desalted by desalted by Millipore $C_{18}$ ZipTip, and analyzed by MALDI mass spectrometry.

### 5.4.12 Proof-of-concept for Phosphoproteomics

*Procedure for trypsin digestion.* 5 mg of Chicken Ovalbumin (Sigma A5503) was re-suspended in 500 μL solution containing 100 mM Tris, pH 8.0, 6 M urea, 5 M DTT and incubated at 37 °C for 60 min. Similar procedure was carried separately for β-casein (Sigma C6905). The sequence identities of these proteins are well established (Figure 5.13). After 60 minutes, a 3 mL solution containing 100 mM Tris, pH 8.0 was added to bring the final concentration of urea to 1 M. To this sample was added 2.5 μL of 100 ng/ μL of Trypsin (Promega V5111). The final protein:trypsin ratio has to be 20:1 (wt/wt) to for optimal digestion. The sample was incubated at 37 °C for 24 h. 200 μL of formic acid was added to bring the final pH ≤ 4.

For trypsin-digested sample, Sep-Pak C18 (WAT023501) was used to remove the excess urea and DTT. The column was initially equilibrated with 1 mL of 50% acetonitrile (ACN). Then 5 mL of 1% ACN with 0.1% formic acid was passed through the column. Digested sample was introduced in the column and washed with additional 5 mL of 1% ACN with 0.1% formic acid to enable complete removal of urea and DTT. Digested peptide sample was then eluted with 1 mL of 50% ACN in 0.1% TFA and 1mL of 90% ACN in 0.1% TFA. The elution sample was evaporated to dryness and

resuspended in 500 μ of water to prepare a solution of digested product containing 1 mg/mL sample.

*Procedure for treatment with alkaline phosphatase and malachite green assay.* 1 nmol of the full-length protein or the digested fragments was evaporated to dryness and incubated with 1 U of CIAP (Fermentas) in 10 μL of 1× CIAP buffer (10 mM Tris, pH 7.5, 10 mM MgCl2) at 37 °C for 12 h. The sample was heated at 95 °C for 5 min and cooled on ice for 10 min to inactivate the CIAP. The sample was investigated with malachite green assay (see malachite green assay procedure in section 5.4.6) to determine the amount of phosphorylation in nanomoles.

A 20 μL sample containing 2 nmol of deoxyribozyme, 2.5 nmol of free DNA anchor oligonucleotide, and 1 nmol of tryptic peptides or the full-length protein (based on the amount of phosphorylation) in 70 mM HEPES, pH 7.5, **X** mM $ZnCl_2$ (2, 3, 4, 5, 6, 7 and 8 mM), 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl was incubated at 37 °C for 96 h. Negligible activity was observed with DhaDz1 by the malachite green assay.


**Figure 5.13.** Sequences of full-length proteins are depicted in this figure. Yellow region indicates phosphorylation sites. Blue highlight indicates glycosylation and green indicates N-terminal acetylation. A) Chicken ovalbumin. B) β-casein.

**Table 5.1.** Mass spectrometry of the peptides used in this study. A) MALDI data in black (monoisotopic). B) ESI-MS data in red (average). All the masses indicated as $[M+H]^+$.

| Peptide | mass calcd. | mass found |
|---|---|---|
| Cytolysin (pSer) | 860.6 | 860.5 |
| Cytolysin (Dha) | 762.7 | 762.6 |
| T22 (pSer) | 981.0 | 981.0 |
| T22 (Dha) | 893.0 | 893 |
| Conopeptide (pSer) | 966.6 | 966.5 |
| Conopeptide (Dha) | 868.6 | 868.5 |
| Lacticin (pSer) | 1005.7 | 1005.4 |
| Lacticin (Dha) | 987.7 | na |
| AAAS$^{P}$AA | 540.5 | 540.5 |
| AAAT$^{P}$AA | 554.5 | 555.2 |
| AKKS$^{P}$KKA (Tfa-protected) | 1224.0 | 1225.0 |
| AKKT$^{P}$KKA (Tfa-protected) | 1238 | 1237.4 |
| AEES$^{P}$EEA | 843.7 | 844.3 |
| AEET$^{P}$EEA | 857.7 | 857.3 |
| AFFS$^{P}$FFA | 915.9 | 915.4 |
| AFFT$^{P}$FFA | 929.9 | 929.4 |
| GIT$^{P}$PVAF | 783.8 | 783.4 |

## 5.5 References


(1) Siodłak, D. α,β-Dehydroamino acids in naturally occurring peptides. *Amino Acids* **2015**, *47*, 1-17.

(2) Gersch, M.; Kreuzer, J.; Sieber, S. A. Electrophilic natural products and their biological targets. *Nat. Prod. Rep.* **2012**, *29*, 659-682.

(3) Wang, L.; Zhang, Z.; Brock, A.; Schultz, P. G. Addition of the keto functional group to the genetic code of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 56-61.

(4) Chatterjee, C.; Miller, L. M.; Leung, Y. L.; Xie, L.; Yi, M.; Kelleher, N. L.; van der Donk, W. A. Lacticin 481 synthetase phosphorylates its substrate during lantibiotic production. *J. Am. Chem. Soc.* **2005**, *127*, 15332-15333.

(5) Knerr, P. J.; van der Donk, W. A. Discovery, biosynthesis, and engineering of lantipeptides. *Annu. Rev. Biochem.* **2012**, *81*, 479-505.

(6) Willey, J. M.; van der Donk, W. A. Lantibiotics: peptides of diverse structure and function. *Annu. Rev. Microbiol.* **2007**, *61*, 477-501.

(7) Ross, A. C.; Vederas, J. C. Fundamental functionality: recent developments in understanding the structure-activity relationships of lantibiotic peptides. *J. Antibiot.* **2011**, *64*, 27-34.

(8) Zhang, Q.; Liu, W. Biosynthesis of thiopeptide antibiotics and their pathway engineering. *Nat. Prod. Rep.* **2013**, *30*, 218-226.

(9) Just-Baringo, X.; Albericio, F.; Álvarez, M. Thiopeptide antibiotics: retrospective and recent advances. *Mar. Drugs* **2014**, *12*, 317-351.

(10) Wang, S.; Zhou, S.; Liu, W. Opportunities and challenges from current investigations into the biosynthetic logic of nosiheptide-represented thiopeptide antibiotics. *Curr. Opin. Chem. Biol.* **2013**, *17*, 626-634.

(11) Drahl, C.; Cravatt, B. F.; Sorensen, E. J. Protein-reactive natural products. *Angew. Chem. Int. Ed.* **2005**, *44*, 5788-5809.

(12) Dittmann, E.; Fewer, D. P.; Neilan, B. A. Cyanobacterial toxins: biosynthetic routes and evolutionary roots. *FEMS Microbiol. Rev.* **2013**, *37*, 23-43.

(13) Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernández-González, M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. Methods for converting cysteine to dehydroalanine on peptides and proteins. *Chem. Sci.* **2011**, *2*, 1666-1676.

(14) Spicer, C. D.; Davis, B. G. Selective chemical protein modification. *Nat. Commun.* **2014**, *5*, 4740.

(15) Zhu, Y. Q.; Li, H. T.; Long, C. Z.; Hu, L. Y.; Xu, H.; Liu, L. P.; Chen, S.; Wang, D. C.; Shao, F. Structural insights into the enzymatic mechanism of the pathogenic MAPK phosphothreonine lyase. *Mol. Cell* **2007**, *28*, 899-913.

(16) Goto, Y.; Li, B.; Claesen, J.; Shi, Y. X.; Bibb, M. J.; van der Donk, W. A. Discovery of Unique Lanthionine Synthetases Reveals New Mechanistic and Evolutionary Insights. *PLoS Biol.* **2010**, *8*.

(17) Khater, S.; Mohanty, D. Genome-wide search for eliminylating domains reveals novel function for BLES03-like proteins. *Genome Biol. Evol.* **2014**, *6*, 2017-2033.

(18) Chen, L.; Wang, H.; Zhang, J.; Gu, L.; Huang, N.; Zhou, J. M.; Chai, J. Structural basis for the catalytic mechanism of phosphothreonine lyase. *Nat. Struct. Mol. Biol.* **2008**, *15*, 101-102.

(19) Goto, Y.; Okesli, A.; van der Donk, W. A. Mechanistic studies of Ser/Thr dehydration catalyzed by a member of the LanL lanthionine synthetase family. *Biochemistry* **2011**, *50*, 891-898.

(20) Oda, Y.; Nagasu, T.; Chait, B. T. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat. Biotechnol.* **2001**, *19*, 379-382.

(21) Rusnak, F.; Zhou, J.; Hathaway, G. M. Identification of phosphorylated and glycosylated sites in peptides by chemically targeted proteolysis. *J. Biomol. Tech.* **2002**, *13*, 228-237.

(22) Knight, Z. A.; Schilling, B.; Row, R. H.; Kenski, D. M.; Gibson, B. W.; Shokat, K. M. Phosphospecific proteolysis for mapping sites of protein phosphorylation. *Nat. Biotechnol.* **2003**, *21*, 1047-1054.

(23) Chalker, J. M.; Lercher, L.; Rose, N. R.; Schofield, C. J.; Davis, B. G. Conversion of cysteine into dehydroalanine enables access to synthetic histones bearing diverse post-translational modifications. *Angew. Chem. Int. Ed.* **2012**, *51*, 1835-1839.

(24) Tabor, A. B. The challenge of the lantibiotics: synthetic approaches to thioether-bridged peptides. *Org. Biomol. Chem.* **2011**, *9*, 7606-7628.

(25) Burrage, S. A.; Rayhham, T.; Bradley, M. A highly efficient route to dehydroalanine containing peptides. *Tetrahedron Lett.* **1998**, *39*, 2831-2834.

(26) Okeley, N. M.; Zhu, Y.; van Der Donk, W. A. Facile chemoselective synthesis of dehydroalanine-containing peptides. *Org. Lett.* **2000**, *2*, 3603-3606.

(27) Pattabiraman, V. R.; Stymiest, J. L.; Derksen, D. J.; Martin, N. I.; Vederas, J. C. Multiple on-resin olefin metathesis to form ring-expanded analogues of the lantibiotic peptide, lacticin 3147 A2. *Org. Lett.* **2007**, *9*, 699-702.

(28) Silverman, S. K. Pursuing DNA Catalysts for Protein Modification. *Acc. Chem. Res.* **2015**, *48*, 1369-1379.

(29) Wysocki, V. H.; Resing, K. A.; Zhang, Q.; Cheng, G. Mass spectrometry of peptides and proteins. *Methods* **2005**, *35*, 211-222.

(30) Sahu, A.; Kay, B. K.; Lambris, J. D. Inhibition of human complement by a C3-binding peptide isolated from a phage-displayed random peptide library. *J. Immunol.* **1996**, *157*, 884-891.

(31) Janssen, B. J.; Halff, E. F.; Lambris, J. D.; Gros, P. Structure of compstatin in complex with complement component C3c reveals a new mechanism of complement inhibition. *J. Biol. Chem.* **2007**, *282*, 29241-29247.

(32) Knerr, P. J.; Tzekou, A.; Ricklin, D.; Qu, H.; Chen, H.; van der Donk, W. A.; Lambris, J. D. Synthesis and activity of thioether-containing analogues of the complement inhibitor compstatin. *ACS Chem. Biol.* **2011**, *6*, 753-760.

(33) Mastellos, D. C.; Yancopoulou, D.; Kokkinos, P.; Huber-Lang, M.; Hajishengallis, G.; Biglarnia, A. R.; Lupu, F.; Nilsson, B.; Risitano, A. M.; Ricklin, D.; Lambris, J. D. Compstatin: a C3-targeted complement inhibitor reaching its prime for bedside intervention. *Eur. J. Clin. Invest.* **2015**, *45*, 423-440.

(34) Dekan, Z.; Vetter, I.; Daly, N. L.; Craik, D. J.; Lewis, R. J.; Alewood, P. F. α-Conotoxin ImI incorporating stable cystathionine bridges maintains full potency and identical three-dimensional structure. *J. Am. Chem. Soc.* **2011**, *133*, 15866-15869.

(35) Muttenthaler, M.; Andersson, A.; de Araujo, A. D.; Dekan, Z.; Lewis, R. J.; Alewood, P. F. Modulating oxytocin activity and plasma stability by disulfide bond engineering. *J. Med. Chem.* **2010**, *53*, 8585-8596.

(36) Hofmann, F. T.; Szostak, J. W.; Seebeck, F. P. In vitro selection of functional lantipeptides. *J. Am. Chem. Soc.* **2012**, *134*, 8038-8041.

(37) Tang, W.; Jimenez-Oses, G.; Houk, K. N.; van der Donk, W. A. Substrate control in stereoselective lanthionine biosynthesis. *Nat. Chem.* **2015**, *7*, 57-64.

(38) Mukherjee, S.; van der Donk, W. A. Mechanistic studies on the substrate-tolerant lanthipeptide synthetase ProcM. *J. Am. Chem. Soc.* **2014**, *136*, 10450-10459.

(39) Gilmore, M. S.; Segarra, R. A.; Booth, M. C.; Bogie, C. P.; Hall, L. R.; Clewell, D. B. Genetic structure of the Enterococcus faecalis plasmid pAD1-encoded cytolytic toxin system and its relationship to lantibiotic determinants. *J. Bacteriol.* **1994**, *176*, 7335-7344.

(40) Tang, W.; van der Donk, W. A. The sequence of the enterococcal cytolysin imparts unusual lanthionine stereochemistry. *Nat. Chem. Biol.* **2013**, *9*, 157-159.

(41) Masuda, M.; Nakashima, H.; Ueda, T.; Naba, H.; Ikoma, R.; Otaka, A.; Terakawa, Y.; Tamamura, H.; Ibuka, T.; Murakami, T.; et al. A novel anti-HIV synthetic peptide, T-22 ([Tyr5,12,Lys7]-polyphemusin II). *Biochem. Biophys. Res. Commun.* **1992**, *189*, 845-850.

(42) Tang, W.; van der Donk, W. A. The sequence of the enterococcal cytolysin imparts unusual lanthionine stereochemistry. *Nat. Chem. Biol.* **2013**, *9*, 157-159.

(43) Ortega, M. A.; Hao, Y.; Zhang, Q.; Walker, M. C.; van der Donk, W. A.; Nair, S. K. Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB. *Nature* **2015**, *517*, 509-512.

(44) Yu, Y.; Zhang, Q.; van der Donk, W. A. Insights into the evolution of lanthipeptide biosynthesis. *Protein Sci.* **2013**, *22*, 1478-1489.

(45) Yang, X.; van der Donk, W. A. Ribosomally synthesized and post-translationally modified peptide natural products: new insights into the role of leader and core peptides during biosynthesis. *Chem. Eur. J.* **2013**, *19*, 7662-7677.

(46) 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.

(47) 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.

(48) Wang, Y.; Silverman, S. K. Characterization of Deoxyribozymes That Synthesize Branched RNA. *Biochemistry* **2003**, *42*, 15252-15263.

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

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