DNA ENZYMES FOR TYROSINE PEGYLATION AND AZIDO-ADENYLYLATION OF PEPTIDE AND PROTEIN SUBSTRATES

BY

PUZHOU WANG

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, 2018

Urbana, Illinois

Doctoral Committee:

Professor Scott K. Silverman, Chair
Professor Jeffrey S. Moore
Assistant Professor David Sarlah
Professor Wilfred A. van der Donk

Abstract

Proteins and RNA are used as enzymes in nature, while DNA is used for the storage and transfer of genetic information. Proteins and RNA are biopolymers that can fold into complex secondary and tertiary structures to enable substrate binding and catalysis. Given the structural similarity to RNA, single-stranded DNA should also be able to function as enzymes. DNA enzymes, or deoxyribozymes, have not been found in nature, but in vitro selection has led to the identification of deoxyribozymes for a variety of reactions. De novo enzyme identification favors the use of nucleic acids over proteins for several reasons. First, nucleic acids can be amplified by natural enzymes, whereas proteins cannot be amplified in any way. Second, the number of possible sequences is smaller for nucleic acids (4^n, where n is the length of the biopolymer) than for proteins (20^n). Therefore, selection experiments for identifying nucleic acid enzymes will cover a larger fraction of total sequence space. Furthermore, within the sequence space evaluated, a large portion of nucleic acid sequences will fold into secondary and tertiary structures, whereas most random protein sequences are unlikely to fold into high-order structures. Considering nucleic acid enzymes, DNA has additional advantages over RNA because DNA can be directly amplified by polymerases whereas RNA requires an extra reverse transcription step. DNA is also cheaper and more stable compared to RNA.

Post-translational modifications (PTMs) are essential for protein functions. The ability to site-specifically modify peptides and proteins will enable better understanding and applications of these biomolecules. The idea of using DNA enzymes for peptide and protein modification is very attractive, especially considering that DNA enzymes with site selectivity can be de novo identified without the requirement of a known enzyme as the starting point.

PEGylation is an important artificial PTM for therapeutic peptides and proteins. PEGylation improves the pharmacokinetic properties of biopharmaceuticals by increasing circulation half-time, reducing immunogenicity, increasing solubility, and suppressing aggregation. Peptide and protein PEGylation is most commonly achieved by solely chemical means. However, these chemical strategies generally lack site selectivity among different target

sites in the substrates. Some chemical strategies also suffer from off-target reactivity, i.e., poor chemoselectivity. Enzymatic approaches for PEGylation have also been developed, yet their application is limited by the substrate specificities and the sequence selectivities of the natural enzymes used. In Chapter 2, DNA enzymes were identified for PEGylation of tyrosine in a DNA-tethered peptide substrate using a 5′-phosphorimidazolide-activated oligonucleotide-PEG conjugate (Imp-oligo-PEG_{5k}) as the PEG donor. Two different approaches are described for the identification of DNA enzymes that are functional with untethered peptide substrates. The first approach is to increase the length of the tether between the peptide substrate and the DNA anchor for mimicking a peptide free in solution. The selection experiment using this approach did not lead to deoxyribozymes, and further analysis of other deoxyribozymes with untethered peptide reactivities suggests that the long tethers may interfere with catalysis. Thus, the first approach was discontinued. The second approach is to alternate the position of the tether between the peptide substrate and the DNA anchor. The rationale is that DNA enzymes are expected to perform catalysis without the requirement of any tether if the enzymes are identified from selection experiments with alternating tether positions. Ongoing efforts include selection experiments using the second approach and mixed-sequence peptide substrates to identify deoxyribozymes with untethered peptide reactivity.

In Chapter 3, a two-step strategy is described for DNA-catalyzed peptide modification. In this strategy, a DNA enzyme first catalyzes the transfer of the 2′-azido-2′-deoxyadenosine 5′-monophosphoryl group (2′-Az-dAMP) from the analogous 5′-triphosphate (2′-Az-dATP) onto the tyrosine hydroxyl group (azido-adenylylation). Second, a particular modification of interest is attached to the azido group by copper-catalyzed azide-alkyne cycloaddition (CuAAC) using an alkyne-functionalized reagent. Eleven deoxyribozymes with azido-adenylylation activity are described in Chapter 3. One of the DNA enzymes is selective for the YPR sequence motif and is able to discriminate between tyrosine residues within a single peptide on the basis of sequence context. Another deoxyribozyme is peptide sequence-general, functions with free peptides, and allows their subsequent CuAAC labeling with moieties such as PEG and fluorescein. The use of

azido-adenylylation deoxyribozymes is a versatile method for the synthesis of site-specifically modified peptides and proteins, since the azide group installed by the DNA enzyme can be used for any particular modification as long as the corresponding alkyne derivative is available.

One of our long-term goals is DNA-catalyzed site-specific modification of protein substrates. In Chapter 4, two proteins, human annexin V and human TNF-related apoptosis-inducing ligand (TRAIL) 114–281, and a 36-mer peptide pancreatic polypeptide (PP) with an additional C-terminal cysteine were used as the substrates to evaluate two different approaches for identifying deoxyribozymes. The first approach is to directly use protein substrates during in vitro selection experiments. This approach requires the surviving deoxyribozymes to simultaneously adopt functions of both binding to the protein substrates and catalyzing the modification. However, the selection experiments using this approach did not lead to deoxyribozymes. The second approach is to decouple the binding and catalytic functions required for DNA-catalyzed protein modification. In this modular approach, the binding function is assigned to the predefined aptamer domain, which is placed adjacent to the initially random enzyme domain. The sequence of the enzyme domain will be subsequently identified through in vitro selection in the presence of the aptamer domain. Ongoing efforts are focused on the identification of DNA aptamers that bind to annexin V, TRAIL, and PP, with benzyl, naphthyl, and indolyl modifications. Once the DNA aptamers are identified, they will be used as the binding modules in selection experiments to identify DNA enzymes for protein modification.

"The accomplished scholar is not a utensil."

Confucius, The Analects, Chapter 2

“君子不器”

——《论语•为政》

Acknowledgements

I am grateful for everyone who has helped me and supported me throughout my Ph.D. study. I would like to thank my research advisor, Professor Scott K. Silverman for his guidance and endless support. I cherish every conversation we had, and I am grateful for his patience with me at times. I am also thankful to my thesis committee members, Professor Jeffrey Moore, Professor David Sarlah, and Professor Wilfred van der Donk. Thank you for all the help provided throughout my graduate school. I would also like to thank Professor Ryan Bailey for discussions during my literature seminar, preliminary exam, and original research proposal.

I had a wonderful time working in the Silverman lab and I am grateful for the talented and nice coworkers who have provided daily support and fun. I would like to thank Ben for his instructions when I first joined the lab; Ying and Victor for their guidance; Jag, for all the conversations we had about science, life, and Game of Thrones; Shannon, for being there all the time whenever I had questions about research and English writing; Jimmy, for the discussions about lab work and NBA; Cong, for all the roles he played as a labmate, friend, roommate, and basketball teammate; Josh, Spurti, Peter, Yves, and Shukun, for the scientific assistance. I was also blessed to have the opportunity to mentor Kevin, Ramiz, and James – I appreciate your hard work and I have also learned a lot from each of you.

To all the friends that I met through my years in Champaign-Urbana − without you, the life in Illinois would have been much less colorful.

I am extremely thankful to my parents for their love and understanding. I am grateful for Yuting's love, support, and motivation. Thank you for being by my side, calming me down at key moments, and helping me fight through the tough times. I would not have made it without you.

Table of Contents

Chapter 1: Introduction to Natural and Engineered Enzymes

1.1 Enzymatic Catalysis

Enzymes are macromolecular biological catalysts that adopt three-dimensional structures and accelerate the chemical conversion of substrates to products by enthalpy effects and/or entropy effects.[1] Organisms in our world use enzymes as the solutions to the challenges faced in daily life. Some enzymes have evolved to achieve high rate enhancement. An extreme example is triose phosphate isomerase, which catalyzes the reversible interconversion of dihydroxyacetone phosphate and D-glyceraldehyde 3-phosphate at a rate that is limited merely by substrate diffusion.[2] Another essential property of enzymes is their high specificity, especially when they are compared with small-molecule catalysts. Unlike small-molecule catalysts that often react with any relevant functional group accessible, enzymes can discriminate between different reaction sites on their substrates to enable chemical reactions with high chemoselectivity and site selectivity. One example is TEV protease, which specifically cleaves the amide bond between glutamine and serine residues in the peptide sequence ENLYFQS.[3]

1.2 Natural Enzymes

1.2.1 Proteins

Most enzymes found in nature are protein enzymes. Proteins are polymers of amino acids connected via amide linkages. The natural amino acids contain a variety of chemical functional groups to promote protein structure and enable catalysis (Figure 1.1). Throughout known life, proteins are composed of 22 genetically encoded (proteinogenic) amino acids, 20 in the standard genetic code and additional two incorporated by special translation mechanisms. The diversity of the 22 proteinogenic amino acids is expanded by unnatural amino acids and post-translational modifications. A protein's function depends upon its secondary, tertiary, and quaternary structure.

Local primary amino acid sequences can stabilize secondary structures such as α-helices, β-sheets, and turns. Tertiary structures are formed by the combination of secondary structures, and quaternary structures containing multiple protein subunits can be formed. In the structure of a protein enzyme, a small region is designated as the active site, where the chemical reaction is catalyzed. The amino acid side chains in the active site are responsible for the interactions with the substrate(s) to enable the catalysis.

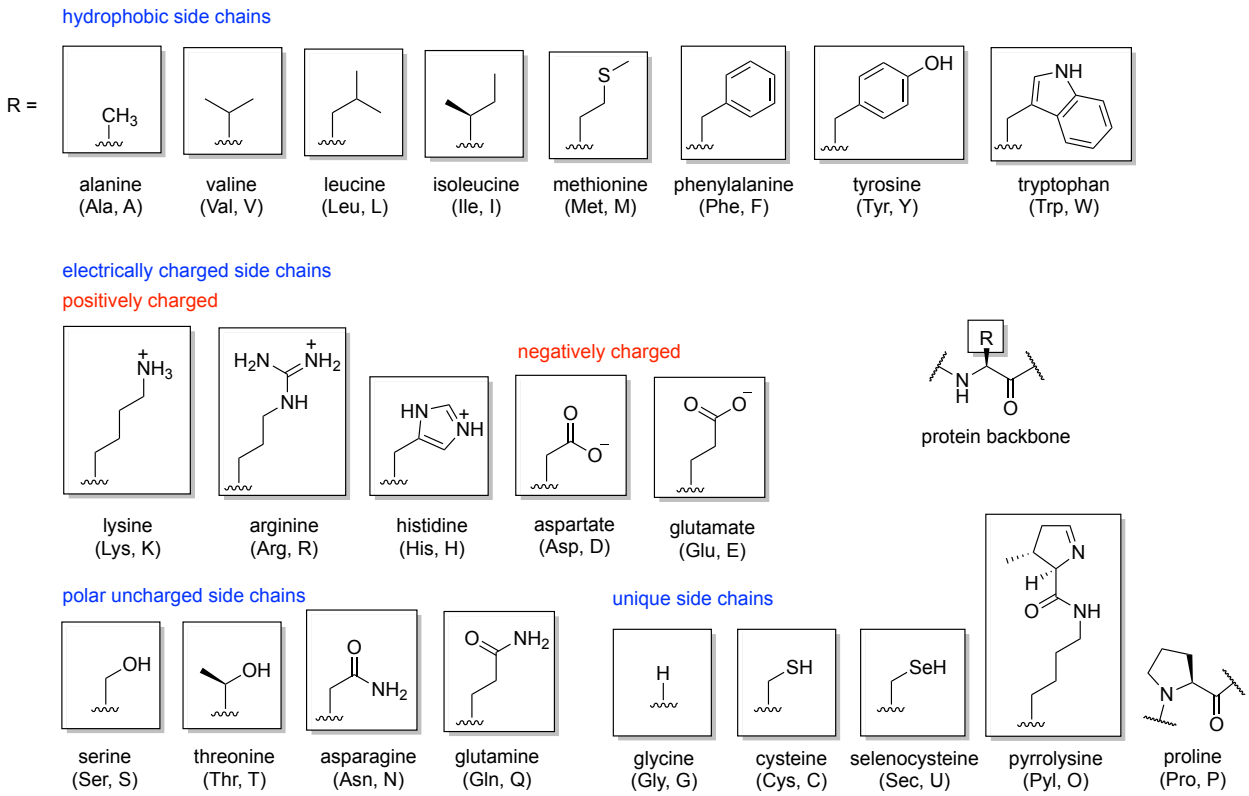


Figure 1.1. Structures of protein backbone and amino acid side chains. Proteins are made of amino acids connected by amide bonds.

Protein enzymes are responsible for many essential biochemical procedures. Four important reactions involved in the central dogma of molecular biology are catalyzed by protein enzymes, namely DNA replication, RNA replication, transcription, and reverse transcription. Problems in the expression, folding, localization, or catalytic efficiency of protein enzymes can

result in diseases. Protein enzymes have also been used in a variety of industries such as pharmaceuticals, biofuels, and consumer products.[4]

1.2.2 Ribozymes

RNA was originally thought to be involved only in the transfer of information between the DNA data storage to the proteins that catalyzed essential cellular functions. However, RNA was discovered to be catalytic in the early 1980s.[5,6] RNA is a polymer of four different ribonucleotides connected via phosphodiester linkages (Figure 1.2A). These four ribonucleotides can be further modified after the transcription. More than 100 different post-transcriptional modifications are known so far,[7] some of which are involved in the catalytic functions of RNA enzymes.[8] Similar to proteins, RNA can also form complex secondary and tertiary structures. Hydrogen bonds between different ribonucleotides lead to secondary structures like hairpin loops, bulges, and internal loops. Tertiary structures are formed by the interactions of secondary structures. In RNA, secondary structures can be stable without being involved in any tertiary structure, whereas secondary structures in proteins generally cannot form without the presence of tertiary structures. The negatively charged RNA backbone enables interaction with metal ions, which can be important for structure or catalysis.[9,10]

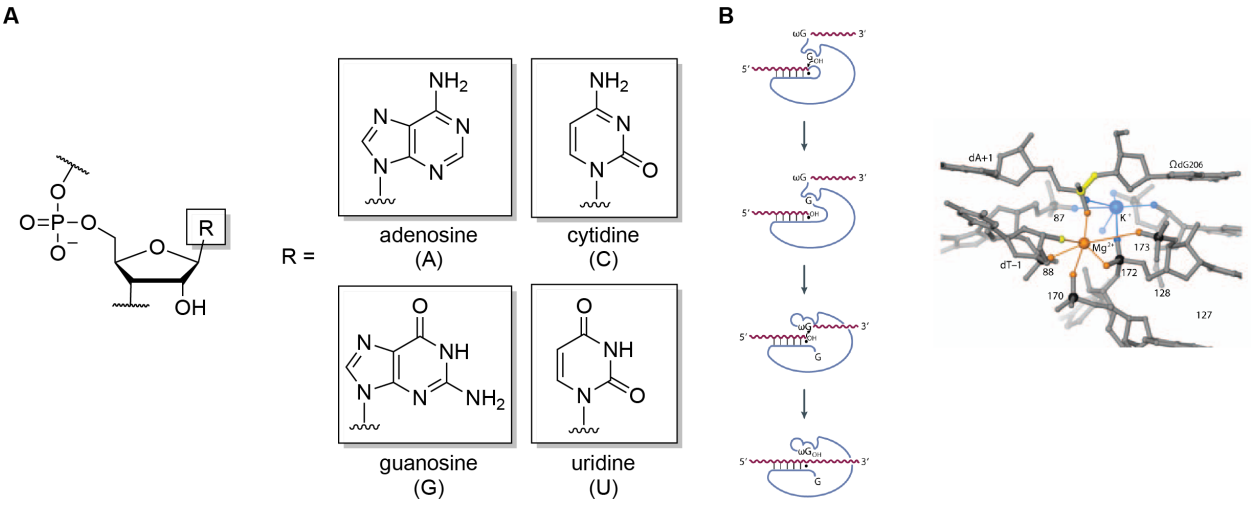


Figure 1.2. Ribozymes. (A) Structures of RNA backbone and nucleobases. RNA is made of ribonucleotides connected by phosphodiester bonds. (B) Group I intron from *Tetrahymena thermophila*. The self-splicing pathway for the group I intron is shown. A guanosine residue binds and attacks the 5′-splice site. A conformation change follows and the 3′-hydroxyl group attacks the 5′-end of the exon, resulting in ligation. The active site is shown in the crystal structure of the group I intron. Panel B adapted with permission from ref. 24.

The first-discovered RNA enzyme was the *Tetrahymena thermophile* group I intron, which catalyzes RNA cleavage as part of intron splicing (Figure 1.2B).[5,11] The RNase P ribozyme was also discovered in the early 1980s to catalyze the cleavage of RNA phosphodiester backbone.[6] RNase P is a ribonucleoprotein complex (RNP), with RNA catalyzing the direct RNA hydrolysis using a water molecule. The RNA component alone is sufficient for catalysis in vitro, but in vivo activity requires the additional protein component.[12] Many other natural ribozymes have since been discovered to catalyze a variety of RNA processing reactions such as self-splicing,[13-15] RNA hydrolysis,[16,17] and self-cleavage.[18,19] Similar to RNase P, some ribozymes function as part of RNP complexes where the RNA performs the catalysis, though both the RNA and protein are required. The most well-known RNP is the ribosome, in which the RNA is responsible for catalyzing the peptidyl transfer reaction during protein translation.[20-23] The abundance of natural catalytic RNAs and their involvement in gene expression and protein synthesis indicate the important roles of catalytic RNAs in biological systems.[24]

1.3 Engineered and Artificial Enzymes

The catalytic efficiencies and substrate specificities of natural protein and RNA enzymes result from billions of years of evolutionary pressure. However, these properties of natural enzymes only occasionally overlap with features sought by humans. Therefore, methods such as directed evolution, rational design, and in vitro selection have been developed both to engineer existing enzymes and to design or identify new enzymes. Natural enzymes modified or artificially developed in the laboratory are referred to as engineered enzymes or artificial enzymes.

1.3.1 Proteins

Two general strategies for protein enzyme engineering are directed evolution and rational design. Directed evolution mimics the process of natural selection to evolve protein enzymes toward desired properties. Rational design of protein enzymes utilizes the understanding of structure-function relationships to predict protein sequences that fold into specific structures for novel functions. Moreover, these two methods are not mutually exclusive; researchers often apply both for protein engineering.

1.3.1.1 Directed Evolution

As an accelerated evolution methodology, directed evolution allows protein enzymes to evolve through laboratory selection pressures for desired catalytic properties on a practical time scale (Figure 1.3).[25] A parent enzyme, which is a known enzyme with catalytic activity similar to that of the desired new enzyme, is required as the starting point. The gene diversification step generates a diverse library of gene mutants. This library is then translated into a corresponding library of protein variants, which are screened or selected for the desired property. The entire process is iterated until protein enzymes with the desired property are identified.

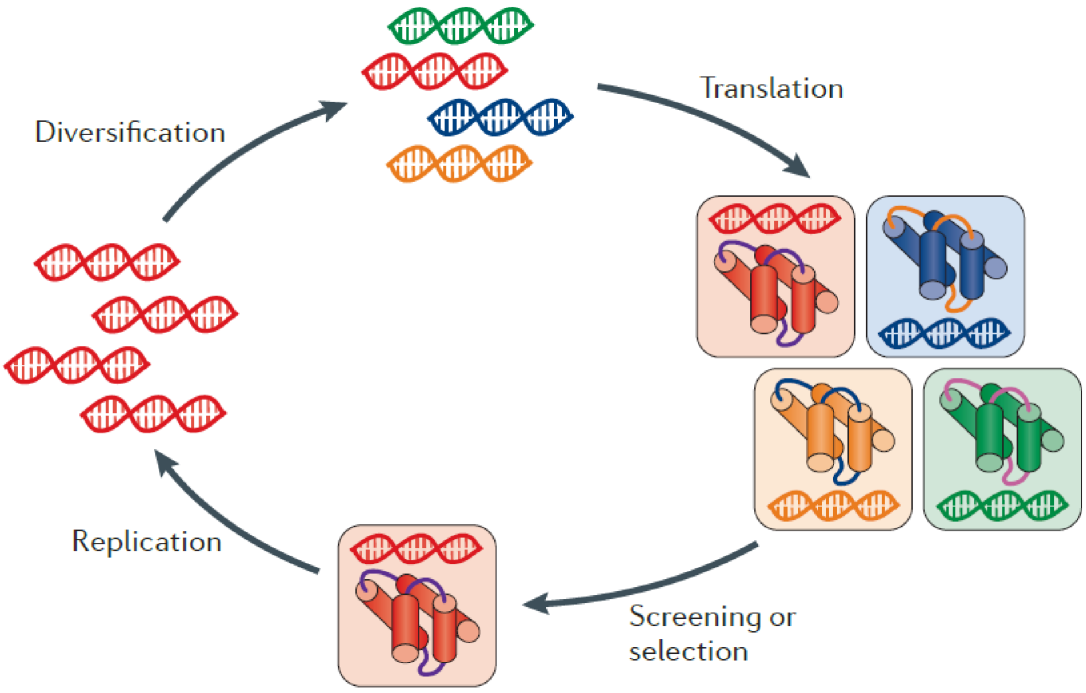


Figure 1.3. Schematic diagram for directed evolution of protein enzymes. A library of genetic variants is prepared by gene diversification. Translation is performed to express the protein mutants, which are screened or selected. The process is iterated, until protein enzymes with the desired property are identified. Figure adapted with permission from ref. 25.

Gene diversification techniques allow the exploration of a protein's sequence space. In the absence of known structure-function relationships, random mutagenesis is usually used for library preparation. Chemical mutagens including ethyl methanesulfonate,[26] nitrous acid,[27] and 2-aminopurine[28] can cause DNA damage and result in errors during replication. However, these approaches are biased in the mutations formed and therefore not commonly used for library generation. Error-prone PCR (epPCR) harnesses the low fidelity of DNA polymerases under certain conditions to enable in vitro random mutagenesis.[29] The DNA polymerases used in epPCR also exhibit mutational biases, but unbalanced deoxyribonucleoside triphosphates (dNTPs) concentrations and proprietary mixtures of polymerases can help to reduce the imbalance in the mutational spectrum.[30,31] Focused mutagenesis is used to maximize the possibility that improved variants are included in the generated library, provided that substantial structural information is available regarding substrate binding or catalysis of the parent enzyme. The most straightforward

approach for focused mutagenesis is to use synthetic DNA oligonucleotides that contain one or more degenerate codons at positions corresponding to targeted amino acid residues. In an ideal case, a degenerated DNA library of 10^{12} members can target about ten amino acid residues in a directed evolution experiment. The mutagenic oligonucleotide is generally incorporated into the gene library as a mutagenic cassette[32] via either restriction enzyme cloning or gene assembly.[33-35] Recombination strategies have also been used to further increase the diversity of the genetic library. Both homologous and non-homologous recombination techniques have been developed. DNA shuffling[36] and the staggered extension process[37] are two examples from the former category, while incremental truncation for the creation of hybrid enzymes (ITCHY)[38] and non-homologous random recombination (NRR)[39] belong to the latter category.

To identify protein enzymes with the desired property, the library of protein variants needs to be screened or selected in a manner that maintains the correspondence between genotype (the gene) and phenotype (the function of the protein enzyme). One approach is in vivo compartmentalization, which utilizes an organism as the individual compartment for both the gene and the translated protein. Yeast display[40] and phage display[41] are two examples for the in vivo compartmentalization approach. In vitro compartmentalization is an alternative format for maintaining the connection between genotype and phenotype. Aqueous droplets in water-oil emulsions are used to compartmentalize individual genes together with gene products generated via in vitro transcription and translation.[42] The linkage between genotype and phenotype can also be achieved by attaching protein variants to their encoding mRNA via either covalent or non-covalent interaction, as shown respectively in mRNA display[43] and ribosome display.[44]

The screening or selection step is used to enrich desired variants and to eliminate undesired ones. Screening is a one-at-a-time process, while selection allows the examine of many variants in parallel. The throughput of screening is determined by the readout techniques. Traditional methods such as chromatography, mass spectrometry, and colorimetric assays have a practical throughput limit of ~10^4 variants.[25] Methods such as fluorescence-activated cell sorting (FACS) can screen larger libraries of ~10^8 members.[40] Unlike screening methods that are typically limited by

measurement throughput, selection methods suffer from a transformation bottleneck which restricts the library size to ~10^9–10^{10} per experiment.[45,46] The recently developed phage-assisted continuous evolution (PACE) system allows directed evolution experiments with libraries of 10^{12} members. [47] In the PACE system, the phage's ability to infect host cells depends on the desired protein function. Because the evolution cycle is based on phage replication, PACE enables several hundred rounds of selection, mutation, and replication to take place per week without manual intervention.

The impact of directed evolution is significant in the field of protein enzyme engineering. Many engineered protein enzymes have been developed for enhanced thermostability,[48] catalytic activity,[49] and new reactions.[50] However, directed evolution suffers from several disadvantages. A parent enzyme is required as the starting point, which may limit the reaction scope of the evolved enzymes. In addition, directed evolution experiments for reprogramming substrate selectivity are often challenging, leading to relaxed selectivity rather than changed selectivity.[51]

1.3.1.2 De Novo Enzyme Design

De novo enzyme design models the active site within the context of a full-length protein based on the understanding of fundamental reaction mechanisms. With the knowledge of individual reaction steps such as bond breaking, bond formation, and charge transfer, a minimalistic three-dimensional model of the active site, labelled theozyme, is assembled (Figure 1.4).[52] This theozyme is docked in silico into structurally characterized scaffolds from the Protein Data Bank (PDB) to choose scaffolds compatible with the presence of the theozyme. Common programs used for the docking step are RosettaMatch,[53] ORBIT,[54] and Scaffold Select.[55] Redesign of the residues in and around this implanted active-site pocket are often needed for optimal packing of the transition state and the catalytic groups.[56,57] The top ranking in silico designs are expressed and assayed for their enzymatic properties. In most cases, directed evolution is also performed to improve the properties of the designed enzymes.[58] Following this general scheme, many protein

enzymes have been generated to catalyze reactions such as retroaldolization,[59] Kemp elimination,[60] Diels-Alder cycloaddition,[61] and ester hydrolysis.[62]

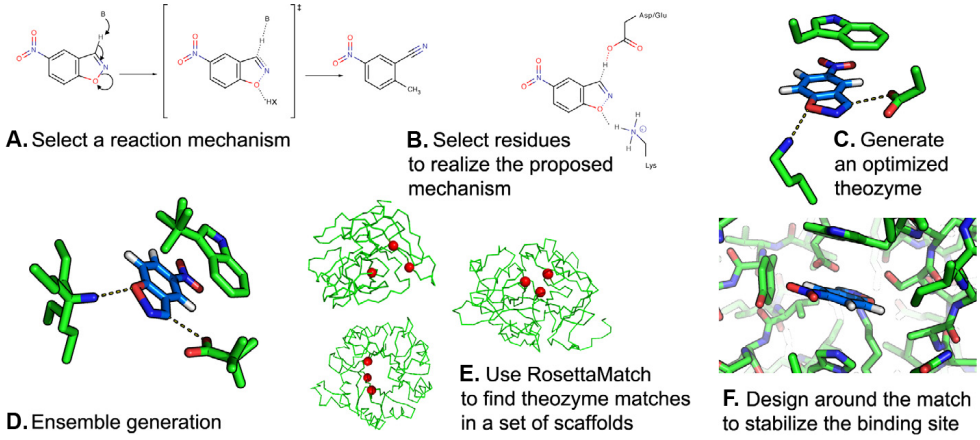


Figure 1.4. Schematic illustration of de novo design of protein enzymes. (A) A mechanism is chosen for the desired reaction, in the example the Kemp elimination. (B) An active-site motif is chosen to realize the stabilizing interactions required by the mechanism. (C) An optimal geometry is defined for the theozyme via quantum chemistry methods. (D) An ensemble of side-chain conformations is generated to stabilize the optimal theozyme. (E) Protein scaffolds from the PDB database are scanned for potential grafting sites for the proposed theozyme using computer programs such as RosettaMatch. (F) The chosen scaffolds are minimized, and the surrounding residues are further designed to stabilize the binding-site geometry. Figure adapted with permission from ref. 52.

Completely de novo protein enzymes have also been designed with both the active site and the scaffold artificially built.[63] The catalytically active Cys-His-Glu triads were installed into heptameric α–helical barrels, which resulted in artificial hydrolases with activities comparable to the most efficient redesigned hydrolases based on natural protein scaffolds. This work demonstrates that completely de novo design of enzymes with control of all residues is possible, although the extent to which a wide variety of reactions are subject to this approach is unclear.

While the power to design novel enzymes is exciting, there are still many challenges in predicting and designing enzymatic protein structures. Besides the difficulty due to technical and theoretical issues, the precise description of structure-function relationships is complicated by the dynamic nature of enzymes.[58] In almost all cases, directed evolution is needed to improve the first generation of newly designed enzymes. However, with our understanding of protein enzyme

structure and function becoming more thorough and our computational algorithm becoming more powerful, computational design of protein enzymes will be more applicable for obtaining useful new enzymes.

1.3.2 Nucleic Acid Enzymes

The discovery of natural ribozymes instigated great interest in the capabilities of nucleic acid enzymes. In the laboratory, both RNA and DNA enzymes have been identified to catalyze various chemical reactions. The technique used for the identification of nucleic acid enzymes is in vitro selection.

1.3.2.1 In Vitro Selection

In vitro selection for nucleic acid enzymes is the experimental process by which random sequences of DNA or RNA are examined in parallel to identify those particular sequences that have a desired function.[64] A general in vitro selection procedure for nucleic acid enzymes starts with a large population of random sequences (~10^{14} molecules; Figure 1.5). Catalytically active sequences are enriched through a selection step followed by a separation method to isolate the nucleic acid molecules with desired enzymatic activity. Some inactive sequences will also survive due to either artifacts of the separation technique or the background of the enzymatic reaction sought in the selection experiment. However, after the selection step, the fraction of catalytically active sequences in the pool increases, which is vital for the success of in vitro selection. The surviving sequences from selection step are amplified to generate more copies of each sequence for the subsequent selection round. This selection-amplification process is iterated until desired catalysis is observed. Finally, individual enzymes from the pool are sequenced and characterized.

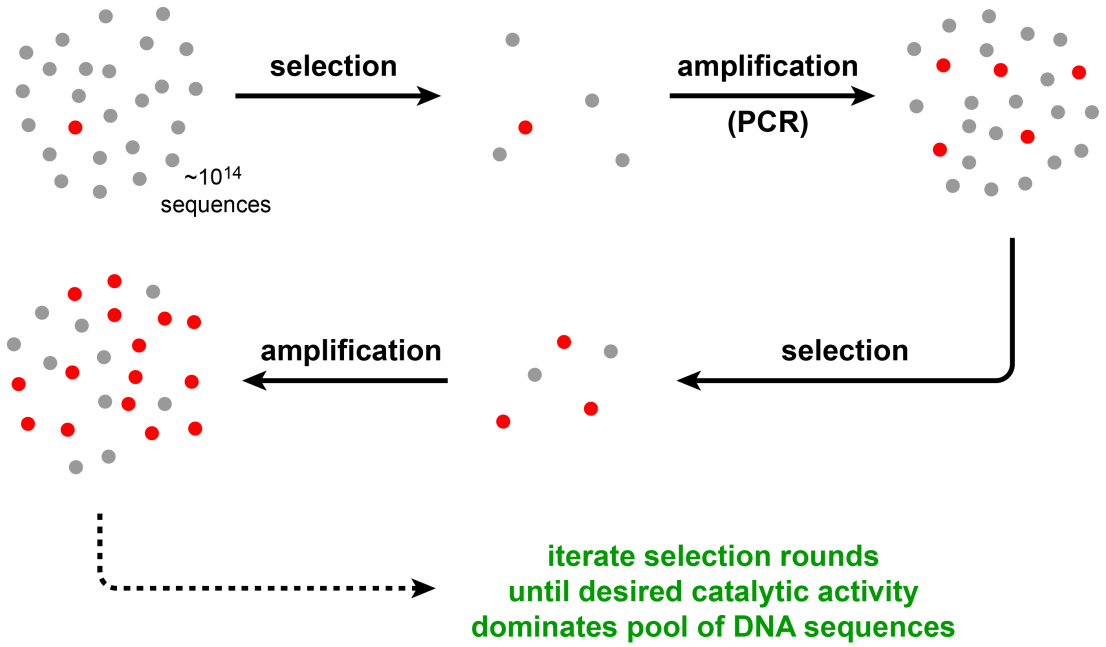


Figure 1.5. Schematic illustration of in vitro selection. The initial pool contains ~10^{14} sequences. The active sequences, represented as red dots, are enriched through a selection step and then amplified by PCR. The selection and amplification steps are iterated until active sequences dominate the selection pool.

Ribozymes and deoxyribozymes are ideal for in vitro selection experiments for several reasons. First and foremost, both RNA and DNA can be amplified, which enables the multiple-round process of in vitro selection. The amplification of RNA molecules involves reverse transcription of RNA to cDNA, PCR of cDNA, and transcription of replicated cDNA. For DNA enzymes, the amplification step is more straightforward, as DNA molecules can be directly replicated via PCR. On the contrary, there is no known technique yet to amplify protein molecules. Second, a much larger fraction of sequence space can be explored in nucleic acid selection experiments than protein selection experiments. Nucleic acids have 4 monomers and a smaller sequence space compared to proteins that have 20 monomers. Within the sequence space evaluated, a large fraction of nucleic acid sequences will fold into complex secondary and tertiary structures, whereas most random protein sequences are unlikely to fold into high-order structures. Therefore, nucleic acid enzymes can be more easily identified via in vitro selection. Considering nucleic acid

enzymes, DNA has several practical advantages compared to RNA in its stability, ease of synthesis, and simplicity in amplification.

1.3.2.2 Ribozymes

Artificial ribozymes have been identified via in vitro selection for RNA modification reactions including RNA cleavage by transesterification,[65] ligation,[66,67] polymerization,[68,69] and phosphorylation,[70,71] providing important contributions for supporting the RNA World hypothesis.[72,73] RNA enzymes have also been found to catalyze other reactions such as acyl transfer,[74,75] biphenyl isomerization,[76] Diels-Alder cycloaddition,[77,78] and aldolization.[79] Despite the success in the identification of novel ribozymes, detailed structural and mechanistic understanding of their functions is limited, even though such understanding is not necessary for the use of in vitro selection. Currently, only three artificial ribozymes, an RNA-cleaving lead-dependent ribozyme,[80,81] a ribozyme that catalyzes the Diels-Alder reaction,[82] and class I ligase ribozymes,[83,84] have crystal structures, providing us detailed knowledge about their structures and functions (Figure 1.6).

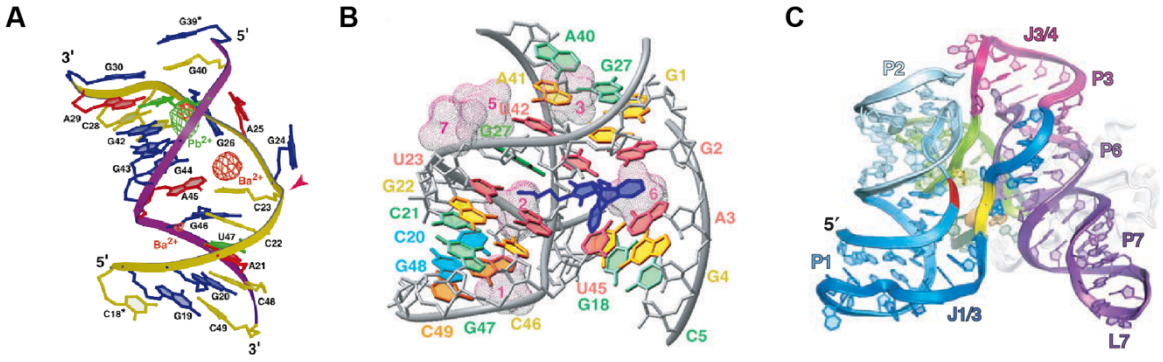


Figure 1.6. Structures of artificial ribozymes. (A) Crystal structure of a lead-dependent RNA-cleaving ribozyme. The ribozyme is shown with the magenta backbone, and the substrate strand is yellow. The red arrowhead indicates the cleavage site. (B) Crystal structure of a ribozyme that catalyzes the Diels-Alder reaction. The structure shows the ribozyme-product complex with the product in blue. The electron density shows the location of bound metal ions. (C) Crystal structure of a class I ligase ribozyme. The structure shows the active site in blue and the ligation junction in red. Panels A, B, and C are reprinted with permission from ref. 80, 82, and 83, respectively.

1.3.2.3 Deoxyribozymes

The chemical structure of DNA is very similar to that of RNA. DNA lacks the 2′-hydroxyl group of RNA, and the nucleobase thymine is present in place of uracil (Figure 1.7A). Natural deoxyribozymes have not been identified. In nature, DNA is primarily double-stranded and used as the long-term repository of genetic information. However, in single-stranded form, DNA can fold into secondary and tertiary structures required for catalysis.

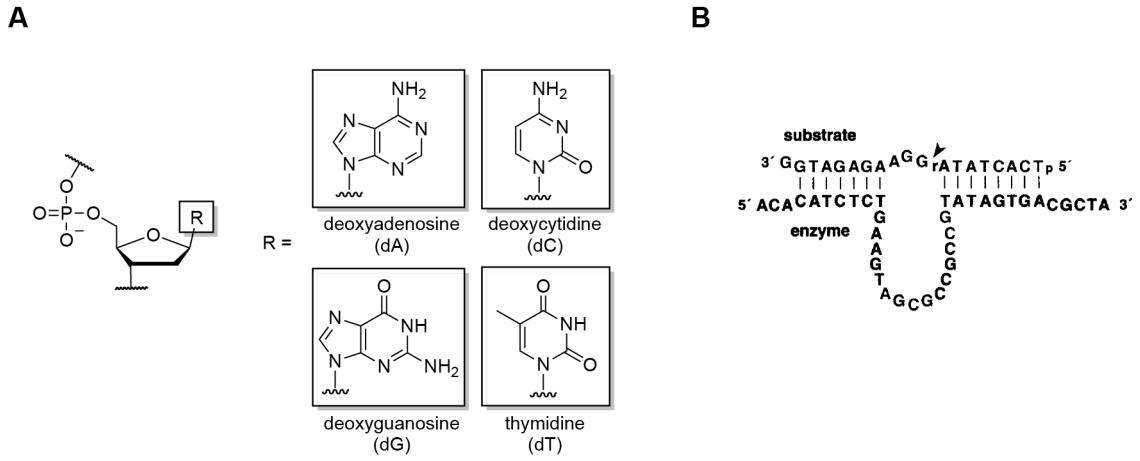


Figure 1.7. Deoxyribozymes. (A) Structure of DNA backbone and nucleobases. DNA is made of 2′-deoxyribonucleotides connected via phosphodiester linkages. (B) The first-discovered deoxyribozyme catalyzes the cleavage of a ribonucleotide embedded within a DNA substrate. The cleavage site is indicated with an arrowhead. Panel B adapted with permission from ref. 85.

DNA enzymes have been identified in a laboratory via in vitro selection. The first deoxyribozyme was discovered in 1994 to catalyze the cleavage of a single ribonucleotide embedded in a DNA substrate (Figure 1.7B).[85] Since then, many DNA enzymes have been identified to catalyze the modification of nucleic acid substrates, as base-pairing interactions can be easily engineered between the substrate and the deoxyribozyme. These modification reactions include RNA cleavage via transesterification[86,87] or hydrolysis;[88] RNA ligation;[89-91] DNA cleavage via hydrolysis,[92,93] deglycosylation,[94] or radical formation;[95] DNA ligation;[96] and DNA phosphorylation.[97,98] The reaction scope of catalysis by DNA is not limited to nucleic acid substrates. In particular, DNA enzymes have been found to catalyze the modification of peptide

substrates by phosphorylation,[99,100] dephosphorylation,[101] nucleopeptide formation,[102-107] and the formation of dehydroalanine from phosphoserine.[108] Some of these deoxyribozymes can catalyze their modification reaction with peptide sequence selectivity.[100,106]

Similar to artificial ribozymes, structural and mechanistic understanding of deoxyribozyme function is limited. Currently, crystal structures are available for two deoxyribozymes, the 9DB1 RNA-ligating deoxyribozyme[109] and the 8-17 RNA-cleaving deoxyribozyme (Figure 1.8).[110] However, the knowledge gained from these two structures is hardly applicable to the identification of novel deoxyribozymes.

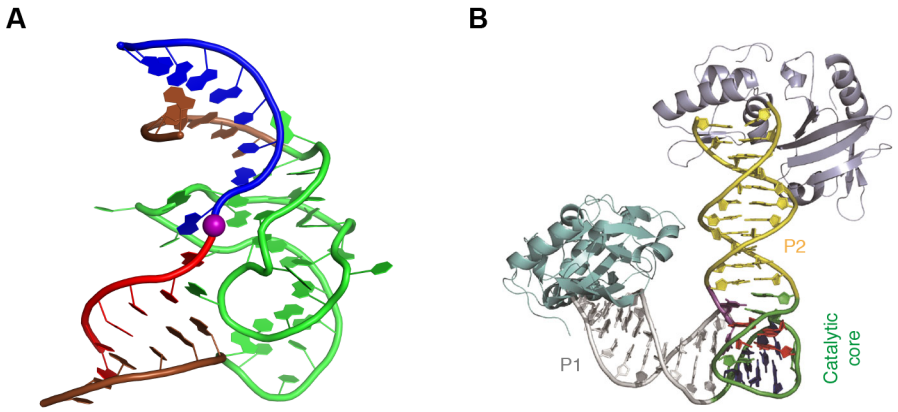


Figure 1.8. Structures of deoxyribozymes. (A) Crystal structure of the 9DB1 deoxyribozyme. The deoxyribozyme is shown in green, and the substrate strands are blue and red. The ligation site is marked with a purple sphere. (B) Crystal structure of the 8-17 deoxyribozyme. *Asfv*PolX proteins, which were utilized to facilitate the crystallization, are colored in light blue or cyan. Panel A and B adapted with permission from ref. 109 and ref. 110, respectively.

1.4 Protein Side Chain Modification

Many natural proteins are modified by enzymes after translation for the real-time dynamics and regulation of their structures and functions. Common natural post-translational modifications (PTMs) include not only additions of small chemical groups such as acetylation, methylation, and phosphorylation, but also attachment of other biopolymers such as glycosylation and

ubiquitination.[111] PTMs are essential for almost all cellular processes including trafficking, differentiation, migration, and signaling (Figure 1.9).[112]

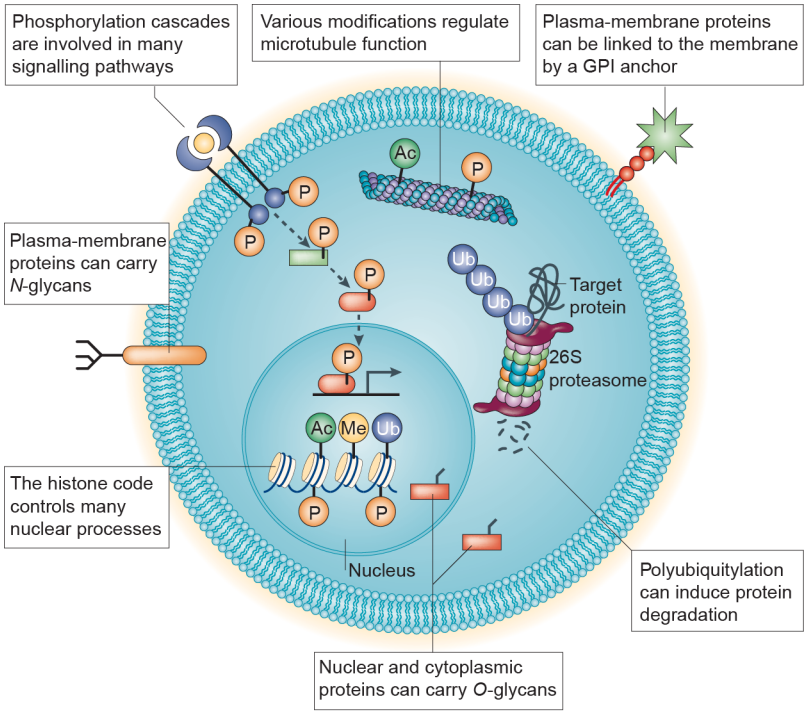


Figure 1.9. Biological roles played by post-translational modifications. This schematic figure shows the locations and functions of some post-translational modifications. Ac, acetyl group; GPI, glycosylphosphatidylinositol; Me, methyl group; P, phosphoryl group; Ub, ubiquitin. Figure reprinted with permission from ref. 112.

Artificial PTMs can also be introduced into proteins. For example, the installment of fluorophores is widely used to study spatial behaviors of target proteins.[113] Another example is the addition of polyethylene glycol molecules (PEGylation), which improves the pharmaceutical properties of protein drugs.[114] Given the significance of PTMs, the production of modified proteins in a highly efficient and site-specific way is of great scientific interest. Many efforts have been focused on the development of novel chemistry for protein modification.[115-117] However, powerful as they are, these chemical approaches lack site selectivity to discriminate two reaction sites in the same protein substrate. In some cases, poor chemoselectivity is also an issue faced by these chemical approaches.

1.5 Thesis Research Focus

The ability to site-specifically modify peptides and proteins will enable better understanding and applications of these biomolecules. Therefore, the research described in this thesis is focused on the identification of deoxyribozymes for covalent modifications at tyrosine residues in peptides and proteins. The work has focused on two specific modification reactions: PEGylation and azido-adenylylation. Chapter 2 reports the identification of deoxyribozymes for tyrosine PEGylation in peptide substrates. These deoxyribozymes use a 5′-phosphorimidazolide-activated oligonucleotide-PEG conjugate as the PEG donor. These PEGylation DNA enzymes catalyze the formation of a phosphodiester bond between the phosphorimidazolide and the tyrosine side chain within a peptide substrate that is covalently attached via a tether to a DNA anchor oligonucleotide. Chapter 3 describes the identification of deoxyribozymes for tyrosine azido-adenylylation in peptide substrates. These deoxyribozymes transfer the 2′-azido-2′-deoxyadenosine 5′-monophosphoryl group (2′-Az-dAMP) from the analogous 5′-triphosphate (2′-Az-dATP) onto the tyrosine hydroxyl group of a peptide, which is either tethered to a DNA anchor or free. One deoxyribozyme is able to differentiate two tyrosine sites within the same peptide substrate based on their sequence contexts. Another deoxyribozyme is able to azido-adenylylate untethered peptides, and the modified peptides are further derivatized via copper-catalyzed azide-alkyne cycloaddition (CuAAC). Chapter 4 discusses the efforts to identify deoxyribozymes for tyrosine azido-adenylylation in protein substrates. The direct approach and the modular approach are the two strategies studied in this chapter. In the direct approach, in vitro selection experiments were performed to identify deoxyribozymes that can simultaneously adopt both functions of binding to the protein substrates and catalyzing the modifications. The modular approach was designed to decouple the binding and catalytic functions, by incorporating a predefined aptamer module into the selection experiments. The modular approach is very different from the studies where independently identified aptamer and enzyme modules are concatenated to form new

catalysts. In the modular approach, the enzyme modules will be identified from initially random sequences with the presence of the aptamer modules.

1.6 References

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

(2) Blacklow, S. C.; Raines, R. T.; Lim, W. A.; Zamore, P. D.; Knowles, J. R. Triosephosphate isomerase catalysis is diffusion controlled. *Biochemistry* **1988**, *27*, 1158-1167.

(3) Carrington, J. C.; Dougherty, W. G. A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 3391-3395.

(4) Singh, R.; Kumar, M.; Mittal, A.; Mehta, P. K. Microbial enzymes: industrial progress in 21st century. *3 Biotech* **2016**, *6*, 174.

(5) Cech, T. R.; Zaug, A. J.; Grabowski, P. J. In vitro splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell* **1981**, *27*, 487-496.

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

(7) Cantara, W. A.; Crain, P. F.; Rozenski, J.; McCloskey, J. A.; Harris, K. A.; Zhang, X. et al. The RNA modification database, RNAMDB: 2011 update. *Nucleic Acids Res.* **2011**, *39*, D195-D201.

(8) King, T. H.; Liu, B.; McCully, R. R.; Fournier, M. J. Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Mol. Cell* **2003**, *11*, 425-435.

(9) Lippert, B. Ligand-pK_a shifts through metals: potential relevance to ribozyme chemistry. *Chem. Biodiversity* **2008**, *5*, 1455-1474.

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

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

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

(13) Peebles, C. L.; Perlman, P. S.; Mecklenburg, K. L.; Petrillo, M. L.; Tabor, J. H.; Jarrell, K. A. et al. A self-splicing RNA excises an intron lariat. *Cell* **1986**, *44*, 213-223.

(14) van der Veen, R.; Arnberg, A. C.; van der Horst, G.; Bonen, L.; Tabak, H. F.; Grivell, L. A. Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro. *Cell* **1986**, *44*, 225-234.

(15) Schmelzer, C.; Schweyen, R. J. Self-splicing of group II introns in vitro: mapping of the branch point and mutational inhibition of lariat formation. *Cell* **1986**, *46*, 557-565.

(16) Decatur, W. A.; Einvik, C.; Johansen, S.; Vogt, V. M. Two group I ribozymes with different functions in a nuclear rDNA intron. *EMBO J.* **1995**, *14*, 4558-4568.

(17) Einvik, C.; Decatur, W. A.; Embley, T. M.; Vogt, V. M.; Johansen, S. *Naegleria* nucleolar introns contain two group I ribozymes with different functions in RNA splicing and processing. *RNA* **1997**, *3*, 710-720.

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

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

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

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

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

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

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

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

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

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

(28) Freese, E. The specific mutagenic effect of base analogues on phage T4. *J. Mol. Biol.* **1959**, *1*, 87-105.

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

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

(31) Vanhercke, T.; Ampe, C.; Tirry, L.; Denolf, P. Reducing mutational bias in random protein libraries. *Anal. Biochem.* **2005**, *339*, 9-14.

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

(33) Gibson, D. G.; Young, L.; Chuang, R. Y.; Venter, J. C.; Hutchison, C. A., III; Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **2009**, *6*, 343-345.

(34) Quan, J.; Tian, J. Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS One* **2009**, *4*, e6441.

(35) Nour-Eldin, H. H.; Geu-Flores, F.; Halkier, B. A. USER cloning and USER fusion: the ideal cloning techniques for small and big laboratories. *Methods Mol. Biol.* **2010**, *643*, 185-200.

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

(37) Zhao, H.; Giver, L.; Shao, Z.; Affholter, J. A.; Arnold, F. H. Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol.* **1998**, *16*, 258-261.

(38) Ostermeier, M.; Shim, J. H.; Benkovic, S. J. A combinatorial approach to hybrid enzymes independent of DNA homology. *Nat. Biotechnol.* **1999**, *17*, 1205-1209.

(39) Bittker, J. A.; Le, B. V.; Liu, J. M.; Liu, D. R. Directed evolution of protein enzymes using nonhomologous random recombination. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 7011-7016.

(40) Boder, E. T.; Wittrup, K. D. Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol.* **1997**, *15*, 553-557.

(41) McCafferty, J.; Griffiths, A. D.; Winter, G.; Chiswell, D. J. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* **1990**, *348*, 552-554.

(42) Tawfik, D. S.; Griffiths, A. D. Man-made cell-like compartments for molecular evolution. *Nat. Biotechnol.* **1998**, *16*, 652-656.

(43) Wilson, D. S.; Keefe, A. D.; Szostak, J. W. The use of mRNA display to select high-affinity protein-binding peptides. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 3750-3755.

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

(45) Dower, W. J.; Miller, J. F.; Ragsdale, C. W. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.* **1988**, *16*, 6127-6145.

(46) Clackson, T.; Hoogenboom, H. R.; Griffiths, A. D.; Winter, G. Making antibody fragments using phage display libraries. *Nature* **1991**, *352*, 624-628.

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

(48) Besenmatter, W.; Kast, P.; Hilvert, D. Relative tolerance of mesostable and thermostable protein homologs to extensive mutation. *Proteins* **2007**, *66*, 500-506.

(49) Choi, Y. H.; Kim, J. H.; Park, J. H.; Lee, N.; Kim, D. H.; Jang, K. S. et al. Protein engineering of α2,3/2,6-sialyltransferase to improve the yield and productivity of in vitro sialyllactose synthesis. *Glycobiology* **2014**, *24*, 159-169.

(50) Hult, K.; Berglund, P. Engineered enzymes for improved organic synthesis. *Curr. Opin. Biotechnol.* **2003**, *14*, 395-400.

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

(52) London, N.; Ambroggio, X. An accurate binding interaction model in de novo computational protein design of interactions: if you build it, they will bind. *J. Struct. Biol.* **2014**, *185*, 136-146.

(53) Zanghellini, A.; Jiang, L.; Wollacott, A. M.; Cheng, G.; Meiler, J.; Althoff, E. A. et al. New algorithms and an in silico benchmark for computational enzyme design. *Protein Sci.* **2006**, *15*, 2785-2794.

(54) Choi, E. J.; Mayo, S. L. Generation and analysis of proline mutants in protein G. *Protein Eng. Des. Sel.* **2006**, *19*, 285-289.

(55) Choi, E. J.; Jacak, R.; Kuhlman, B. A structural bioinformatics approach for identifying proteins predisposed to bind linear epitopes on pre-selected target proteins. *Protein Eng. Des. Sel.* **2013**, *26*, 283-289.

(56) Pearson, A. D.; Mills, J. H.; Song, Y.; Nasertorabi, F.; Han, G. W.; Baker, D. et al. Trapping a transition state in a computationally designed protein bottle. *Science* **2015**, *347*, 863-867.

(57) Siegel, J. B.; Smith, A. L.; Poust, S.; Wargacki, A. J.; Bar-Even, A.; Louw, C. et al. Computational protein design enables a novel one-carbon assimilation pathway. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 3704-3709.

(58) Paladino, A.; Marchetti, F.; Rinaldi, S.; Colombo, G. Protein design: from computer models to artificial intelligence. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, e1318.

(59) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A. et al. De novo computational design of retro-aldol enzymes. *Science* **2008**, *319*, 1387-1391.

(60) Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J. et al. Kemp elimination catalysts by computational enzyme design. *Nature* **2008**, *453*, 190-195.

(61) Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; St.Clair, J. L. et al. Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science* **2010**, *329*, 309-313.

(62) Richter, F.; Blomberg, R.; Khare, S. D.; Kiss, G.; Kuzin, A. P.; Smith, A. J. T. et al. Computational design of catalytic dyads and oxyanion holes for ester hydrolysis. *J. Am. Chem. Soc.* **2012**, *134*, 16197-16206.

(63) Burton, A. J.; Thomson, A. R.; Dawson, W. M.; Brady, R. L.; Woolfson, D. N. Installing hydrolytic activity into a completely de novo protein framework. *Nat. Chem.* **2016**, *8*, 837-844.

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

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

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

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

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

(69) 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.

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

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

(72) Higgs, P. G.; Lehman, N. The RNA World: molecular cooperation at the origins of life. *Nat. Rev. Genet.* **2015**, *16*, 7-17.

(73) Pressman, A.; Blanco, C.; Chen, I. A. The RNA World as a model system to study the origin of life. *Curr. Biol.* **2015**, *25*, R953-R963.

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

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

(76) Prudent, J. R.; Uno, T.; Schultz, P. G. Expanding the scope of RNA catalysis. *Science* **1994**, *264*, 1924-1927.

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

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

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

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

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

(82) Serganov, A.; Keiper, S.; Malinina, L.; Tereshko, V.; Skripkin, E.; Höbartner, C. et al. Structural basis for Diels-Alder ribozyme-catalyzed carbon-carbon bond formation. *Nat. Struct. Mol. Biol.* **2005**, *12*, 218-224.

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

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

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

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

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

(88) 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.

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

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

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

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

(93) 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.

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

(95) Lee, Y.; Klauser, P. C.; Brandsen, B. M.; Zhou, C.; Li, X.; Silverman, S. K. DNA-catalyzed DNA cleavage by a radical pathway with well-defined products. *J. Am. Chem. Soc.* **2017**, *139*, 255-261.

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

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

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

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

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

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

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

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

(104) 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.

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

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

(107) 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.

(108) 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.

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

(110) Liu, H.; Yu, X.; Chen, Y.; Zhang, J.; Wu, B.; Zheng, L. et al. Crystal structure of an RNA-cleaving DNAzyme. *Nat. Commun.* **2017**, *8*, 2006.

(111) Yang, X. J. Multisite protein modification and intramolecular signaling. *Oncogene* **2005**, *24*, 1653-1662.

(112) Jensen, O. N. Interpreting the protein language using proteomics. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 391-403.

(113) Schneider, A. F. L.; Hackenberger, C. P. R. Fluorescent labelling in living cells. *Curr. Opin. Biotechnol.* **2017**, *48*, 61-68.

(114) Dozier, J. K.; Distefano, M. D. Site-specific PEGylation of therapeutic proteins. *Int. J. Mol. Sci.* **2015**, *16*, 25831-25864.

(115) Wright, T. H.; Bower, B. J.; Chalker, J. M.; Bernardes, G. J. L.; Wiewiora, R.; Ng, W. L. et al. Posttranslational mutagenesis: A chemical strategy for exploring protein side-chain diversity. *Science* **2016**, *354*, aag1465.

(116) Yang, A.; Ha, S.; Ahn, J.; Kim, R.; Kim, S.; Lee, Y. et al. A chemical biology route to site-specific authentic protein modifications. *Science* **2016**, *354*, 623-626.

(117) Lin, S.; Yang, X.; Jia, S.; Weeks, A. M.; Hornsby, M.; Lee, P. S. et al. Redox-based reagents for chemoselective methionine bioconjugation. *Science* **2017**, *355*, 597-602.

Chapter 2: DNA Enzymes with Tyrosine PEGylation Activity on Peptide Substrates[†]

2.1 Introduction

2.1.1 Basic Features of PEGylation

Since the approval of recombinant human insulin by the US Food and Drug Administration (FDA) in 1982, the field of peptide and protein therapeutics has grown significantly. Compared to small molecules, peptide and protein drugs offer many advantages, including high specificity, high biological activity, and few adverse effects. In addition, due to their vast diversity, peptides and proteins can be used to treat a number of different diseases including bacterial infection, various cancers, endocrine disorders, and autoimmune diseases. However, such therapy has its own pitfalls. Common issues associated with therapeutic peptides and proteins are rapid body clearance, poor physical stability, enzymatic degradation, and immunogenicity.[1,2] One strategy to overcome the aforementioned limitations is PEGylation, defined as the attachment of poly(ethylene glycol) (PEG) chains. PEGylation can improve pharmacokinetic properties of proteins by increasing their clearance time, suppressing aggregation, and shielding them from proteases and the immune system.[3] Currently ten PEGylated protein biopharmaceuticals are marketed in the US (Table 2.1).[4] Most of these therapeutic proteins are modified by non-specific PEGylation methods. However, these methods inevitably lead to a heterogeneous mixture of PEGylated isomers that have different bio-activity and stability properties among each other. On the other side, site-specific PEGylation methods allow the synthesis of modified proteins with

[†] Former Silverman laboratory graduate student Shannon M. Walsh performed selections for kinase deoxyribozymes with HEG_{10}-tethered peptides and performed experiments to evaluate the tether dependence of the nucleopeptide-forming 8XJ105 deoxyribozyme.[23]

Silverman laboratory undergraduate student James E. Brady performed in vitro selection experiments to identify PEGylation deoxyribozymes with untethered peptide activity using alternating tethers and mixed-sequence peptide substrates.

optimal therapeutic properties, and potentially streamline the production and purification of modified proteins.

Table 2.1: Marketed PEGylated biopharmaceuticals in the United States. The column "site-specific" demonstrates whether the drug is synthesized by site-specific PEGylation. Table adapted with permission from ref. 4.

brand name	protein	indication	year approved	site-specific
Adagen®	adenosine deaminase	severe combined immunodeficiency	1990	No
Oncaspar®	asparaginase	acute lymphoblastic leukemia	1994	No
PegIntron®	interferon α-2b	chronic hepatitis C	2000	No
PEGASYS®	interferon α-2a	chronic hepatitis C	2001	No
Neulasta®	granulocyte colony-stimulating factor (G-CSF)	neutropenia	2002	Yes
SOMAVERT®	human growth hormone (hGH)	acromegaly	2003	No
Mircera®	erythropoietin (EPO)	anemia	2007	No
CIMZIA®	anti-tumor necrosis factor α Fab′	Crohn's disease, rheumatoid arthritis, psoriatic arthritis, and ankylosing spondylitis	2008	Yes
KRYSTEXXA®	urate oxidase	chronic gout	2010	No
PLEGRIDY®	interferon β-1a	relapsing forms of multiple sclerosis	2014	Yes

2.1.2 Non-Enzymatic PEGylation

As an artificial post-translational modification, PEGylation is not endogenously catalyzed by any natural protein enzymes. Therefore, various chemical approaches have been developed to target functional groups in peptide and protein substrates for PEGylation. All of the marketed PEGylated biopharmaceuticals are synthesized via these chemical means.

2.1.2.1 PEGylation Approaches Targeting Lysine Residues

The initial PEGylation work used cyanuric chloride to prepare activated PEG for attachment to protein substrates at lysine residues. The PEG-dichlorotriazine derivative reacts with

a lysine side chain to form a secondary amine linkage (Figure 2.1A).[5,6] However, the activated PEG can also react with other nucleophilic functional groups in a protein such as tyrosine, serine, threonine, cysteine, and histidine. In addition, the remaining chloride is reactive enough to allow crosslinking of protein molecules that contain additional nucleophilic residues.

Another activated PEG species widely used for lysine modification is the *N*-hydroxysuccinimide ester of PEG carboxylic acid (PEG-NHS, Figure 2.1B).[7] Several marketed biopharmaceuticals with PEGylation modification are synthesized with PEG-NHS, including Adagen®, Oncaspar®, PegIntron®, PEGASYS®, SOMAVERT®, and Mircera®.[4] However, all of these protein drugs are PEGylated non-site specifically, leading to heterogeneous mixture of PEGylated material with each PEG conjugate having its own activity and stability properties.

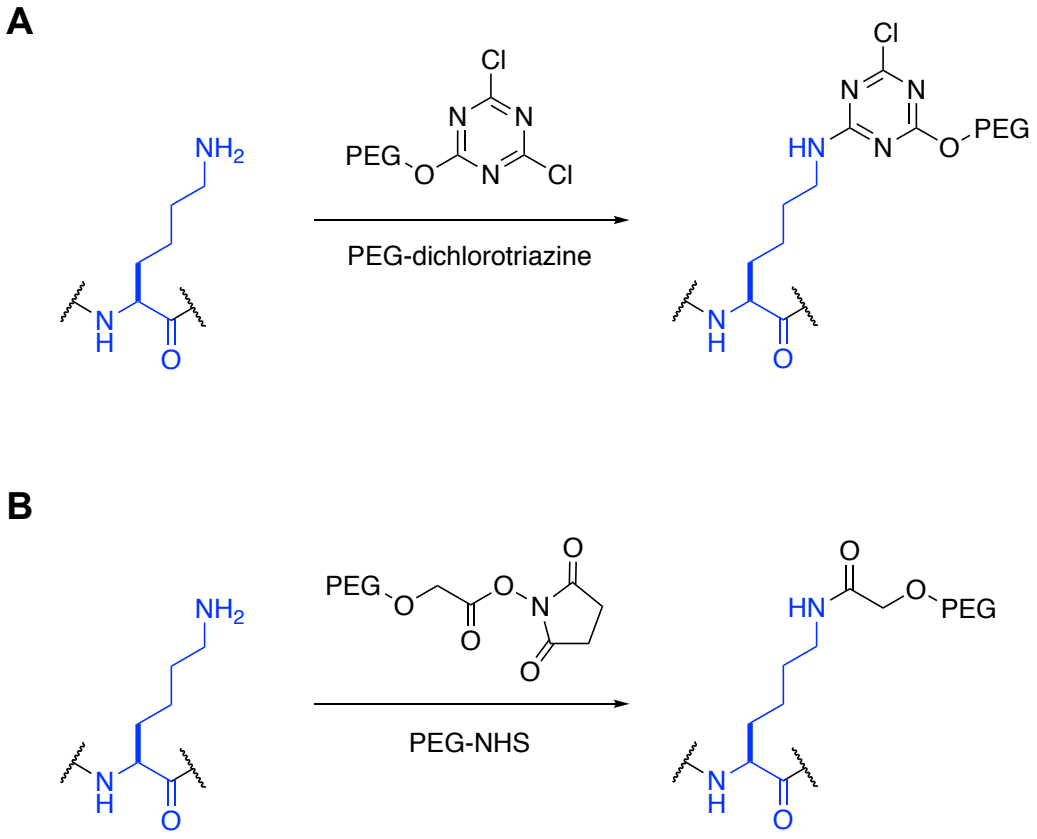


Figure 2.1. PEGylation approaches targeting lysine residues. (A) Reaction with PEG-dichlorotriazine. (B) Reaction with PEG-NHS.

2.1.2.2 PEGylation Approaches Targeting N-Terminal Amines

N-terminal PEGylation takes advantage of the fact that the α-amino group at the protein N-terminus has a lower pK_a (about 7.6–8) than the pK_a of the ε-amino side chain of lysines (about 9.3–9.5).[8] The selective PEGylation at the N-terminus is achieved by performing the reaction in mildly acidic conditions (e.g., pH 6–6.5) and with less reactive PEGylating reagents, such as aldehyde-functionalized PEG (PEG-aldehyde, Figure 2.2).[9] However, in most cases complete selectivity is not observed, even though the extensive heterogeneity frequently observed in lysine chemistry is greatly reduced after careful optimization of reaction conditions.[10] Two site-specific PEGylated protein drugs, Neulasta® and PLEGRIDY®, are synthesized with PEG-aldehyde reagents.[4]

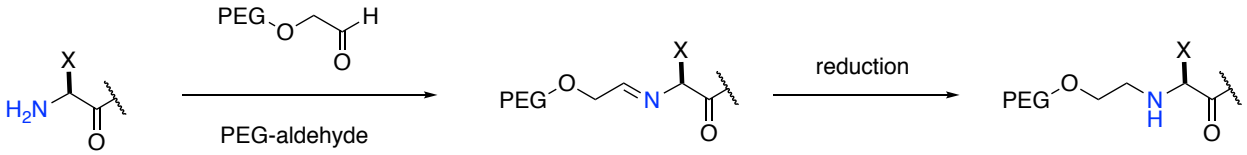


Figure 2.2. PEGylation at the N-terminal amine with PEG-aldehyde. The imine product can be reduced (e.g., with cyanoborohydride) to form a more stable derivative.

2.1.2.3 PEGylation Approaches Targeting Cysteine Residues

Cysteine is another target for site-specific PEGylation, mainly due to its low abundance in natural proteins and the high nucleophilicity of the thiol group.[11] In the absence of a free cysteine in a native protein, one or more free cysteines can be introduced by genetic engineering.[12] However, incorrect disulfide formation and protein dimerization are the issues associated with expressing a protein that contains an odd number of cysteine residues.[13]

Several PEG derivatives, such as PEG-maleimide (Figure 2.3A)[12], PEG-iodoacetamide (Figure 2.3B)[14], and PEG-(2-pyridyl disulfide) (Figure 2.3C)[15], have been developed for cysteine-targeting modification. Among these three different strategies, the PEG-maleimide has been used to synthesize the marketed protein drug, CIMZIA®.[4]

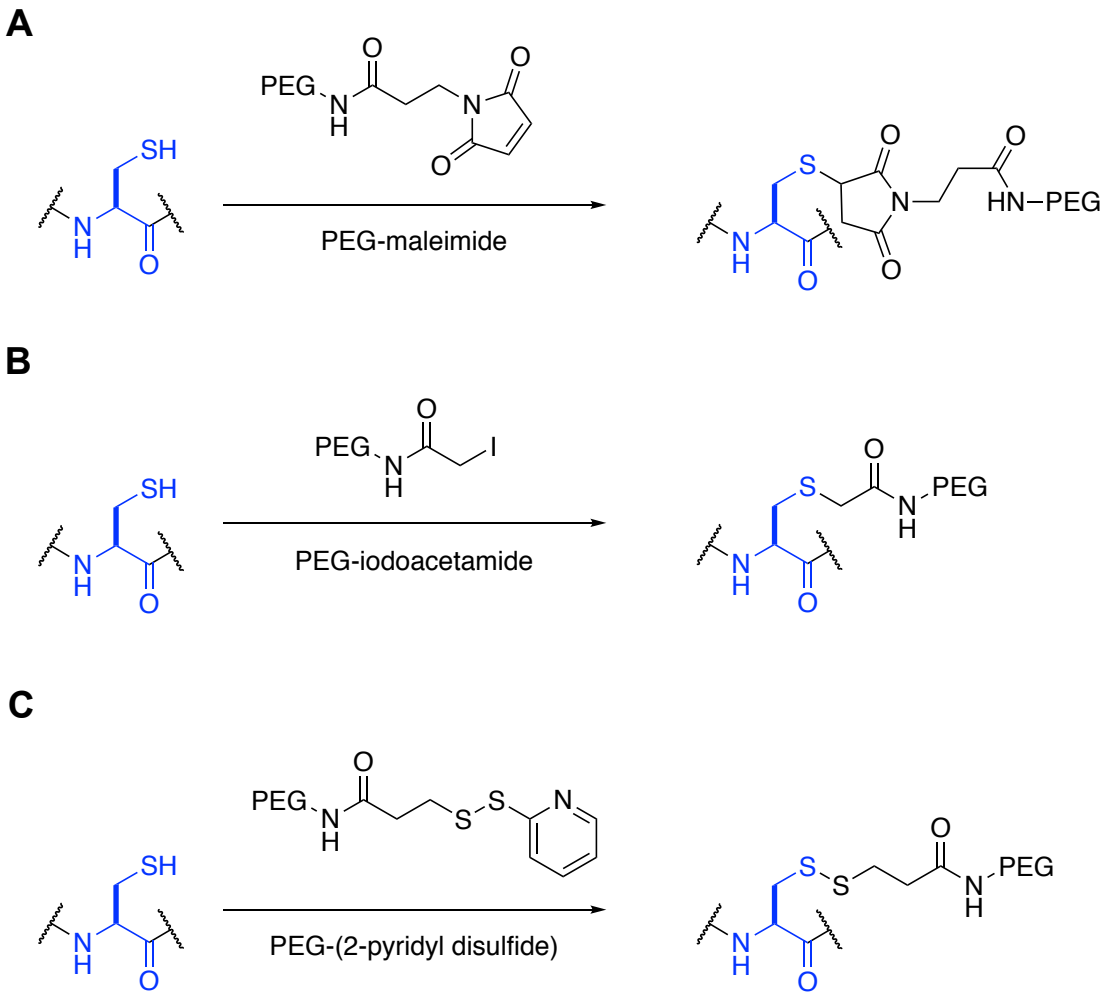


Figure 2.3. PEGylation approaches targeting cysteine residues. (A) Reaction with PEG-maleimide. (B) Reaction with PEG-iodoacetamide. (C) Disulfide exchange with PEG-(2-pyridyl disulfide).

2.1.2.4 PEGylation Approaches Targeting Unnatural Amino Acids

The site-specific incorporation of unnatural amino acids with bioorthogonal reactivity provides another versatile tool for PEGylation of peptides and proteins. Two examples of unnatural amino acids used for PEGylation include *p*-acetylphenylalanine[16] and *p*-azidophenylalanine.[17] The former phenylalanine derivative reacts selectively with PEG-hydrazide or PEG-aminooxy (Figure 2.4A), while the latter derivative can be coupled to PEG-alkyne through copper-catalyzed azide-alkyne cycloaddition (CuAAC) (Figure 2.4B). The main difficulty with unnatural amino acid-targeting PEGylation is the relatively low yield for the expression of protein targets.[18]

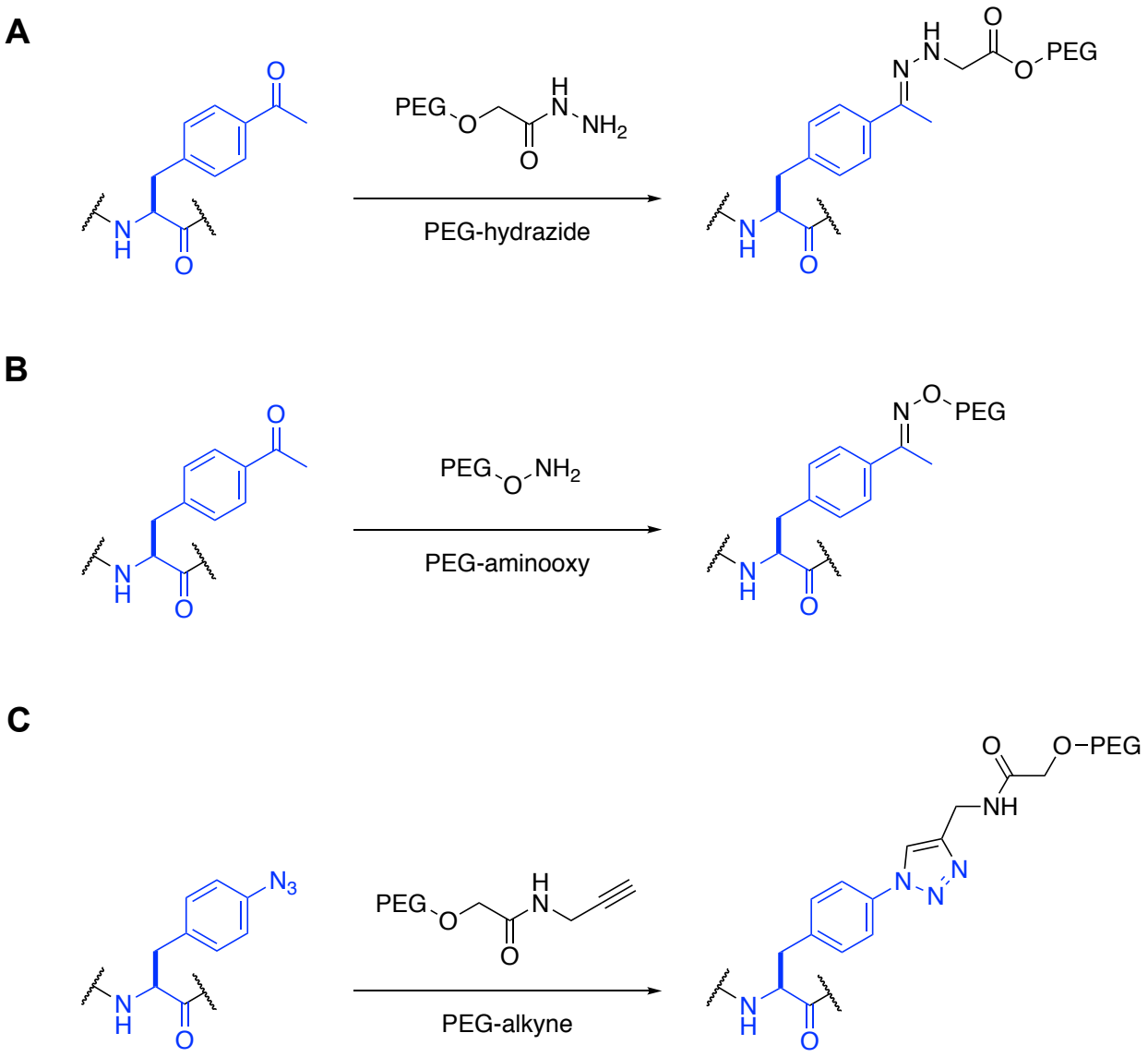


Figure 2.4. PEGylation approaches targeting unnatural amino acids. Reaction of *p*-acetylphenylalanine with PEG-hydrazide (A) or PEG-aminooxy (B). (C) Reaction of *p*-azidophenylalanine with PEG-alkyne via CuAAC.

2.1.3 Enzymatic PEGylation

Natural protein enzymes with promiscuous substrate specificity have also been utilized for PEGylation in a few cases. For example, Sortase A (SrtA) derived from *Staphylococcus aureus* catalyzes a transpeptidase reaction between the N-terminal amino group of glycine and a specific internal peptide sequence, LPETG, on the target protein.[19] When a polyglycine-derived compound

with a pendant PEG group (PEG-polyglycine) is provided, SrtA conjugates the PEG-polyglycine with the LPETG-tagged target protein (Figure 2.5A).

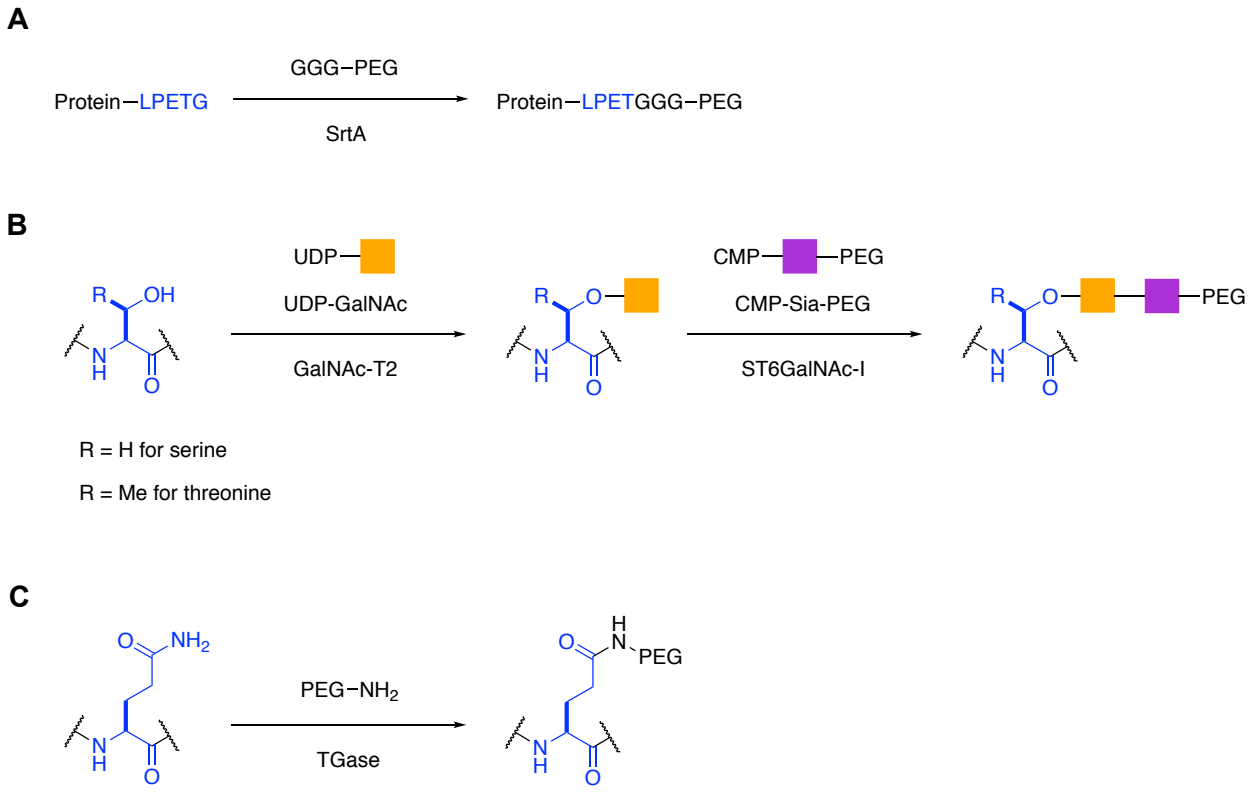


Figure 2.5. Enzymatic PEGylation approaches. (A) PEGylation catalyzed by SrtA. (B) GlycoPEGylation. (C) PEGylation mediated by TGase.

In an approach named GlycoPEGylation, a natural sialylation system is harnessed to introduce PEG molecules into target proteins (Figure 2.5B).[20] A protein is first incubated with a galactosyltransferase and uridine 5′-diphospho-*N*-acetylgalactosamine (UDP-GalNAc). Once the GalNAc is installed, the protein is incubated with a sialyltransferase and a cytidine monophosphate (CMP) activated sialic acid derivative containing a PEG moiety. The sialyltransferase adds the sialyl-PEG group to the GalNac residue and thus PEGylates the protein.

Another class of enzymes for peptide and protein PEGylation are the transglutaminases (TGases, Figure 2.5C).[21] These enzymes create a covalent bond between the primary amine from the PEG molecule and the carboxamide group of glutamine via an acyl transfer reaction. However,

this enzymatic strategy is limited by the requirement that the glutamine residue must be in a flexible loop portion of the target protein for catalysis by the transglutaminase.

These enzymatic PEGylation methods rely solely on the substrate scope of the natural enzymes for site selectivity. However, in some cases, the site selectivity (if any) provided by the natural enzymes is not satisfactory for particular purposes. Therefore, DNA enzymes with PEGylation activity will be valuable alternative tools, especially considering that DNA enzymes with site selectivity can be de novo identified without the requirement of a known enzyme as the starting point.

2.2 Results and Discussion

2.2.1 In Vitro Selection Strategy

To achieve DNA-catalyzed peptide and protein PEGylation, the PEG donors should be recognized by DNA enzymes. Considering the amorphous structure of PEG molecules after their hydration in water, additional recognition moieties in the PEG donors would be desired. The most expedient candidate is an oligonucleotide linker. A 5′-phosphorylated 3′-amino pentamer, pCGTGC-NH_2, was chosen. The pentamer was conjugated to PEG carboxylic acid with average molecular weight of 5,000 (PEG_{5k}-COOH) by the amide coupling reaction (Figure 2.6). Following the 5′-phosphorimidazolide (Imp) activation, this Imp-oligo-PEG_{5k} conjugate served as the PEG donor in the PEGylation selection experiments.

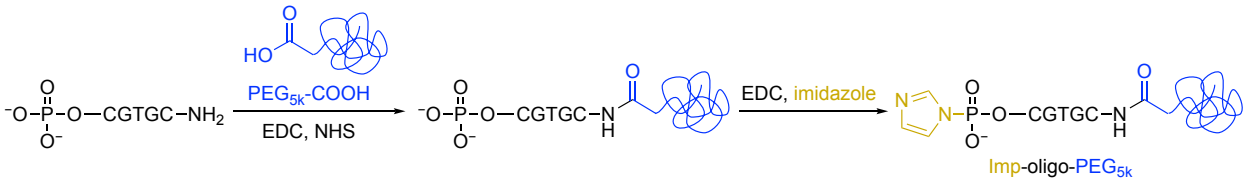


Figure 2.6. Preparation of the PEG donor for in vitro selection experiments. PEG_{5k}-COOH was first activated by *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide (EDC) and *N*-hydroxy-succinimide (NHS), and then coupled with 5′-phosphorylated 3′-amino oligonucleotide at room temperature overnight. The yield of oligo-PEG_{5k} synthesis was 40%. The 5′-phosphoryl group was then transformed into 5′-phosphorimidazolide in quantitative yield by 2 h treatment with EDC and imidazole.

Previous efforts by former Silverman lab graduate student Chih-Chi Chu for DNA-catalyzed nucleopeptide formation showed that the Imp moiety is a reactive electrophile for the phenolic hydroxyl group on tyrosine side chain.[22] Thus, the initial GA1 selection experiment was performed with the tripeptide CYA prepared by solid-phase peptide synthesis as the substrate (Figure 2.7). All in vitro selection experiments are given alphanumeric codes in the following pattern: A-Z, AA-ZZ, AA1-ZZ1, etc. This tripeptide was covalently attached via a disulfide linkage and a hexa(ethylene glycol) [HEG] tether to a DNA anchor oligonucleotide, which was bound by Watson-Crick base pairs to one of the fixed-sequence binding arms of the random N_{40} DNA pool. The DNA-anchored peptide substrate was ligated to the initially random DNA pool using T4 DNA ligase. The initially random region of the DNA sequences was flanked by fixed-sequence regions to enable PCR amplification of the DNA sequences and substrate binding.

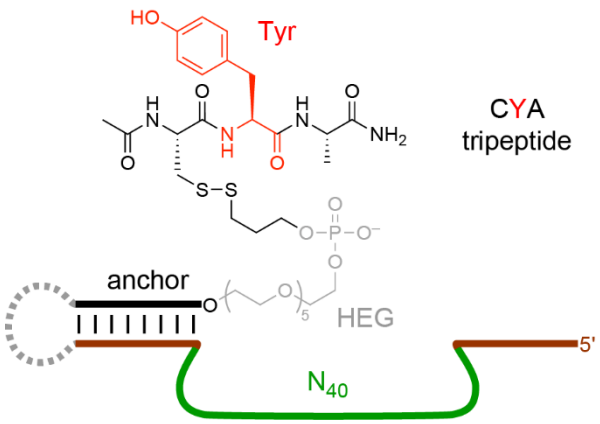


Figure 2.7. Arrangement of initially random DNA pool (N_{40}) and peptide substrate for the GA1 selection experiment of deoxyribozymes for PEGylation. The dashed loop enables selection but is dispensable for catalysis.

The GA1 selection experiment began with ligation of the initially random DNA pool to the DNA-anchored peptide substrate. The ligation product was isolated by polyacrylamide gel electrophoresis (PAGE). Next, the selection step was performed in which the DNA sequences were incubated with Imp-oligo-PEG_{5k} and divalent metal ions in a buffered solution to enable DNA catalysis. After exposure of the substrate-conjugated N_{40} pool to the PEG donor, the catalytically active DNA sequences, which were now attached to an PEG group, were separated from the inactive pool by PAGE. As the standard for PAGE shift of the PEGylated DNA sequences, a 5′-phosphorylated 3′-amino octamer, pTGGGTGCG-NH_2, was conjugated to PEG_{5k}-COOH by the amide coupling reaction and then ligated to the DNA pool (Figure 2.8). PCR was performed to amplify the catalytically active DNA sequences that survived the previous selection step. The entire selection process was iterated, each time enriching the population with catalytically active sequences, until the DNA enzymes dominated the population.

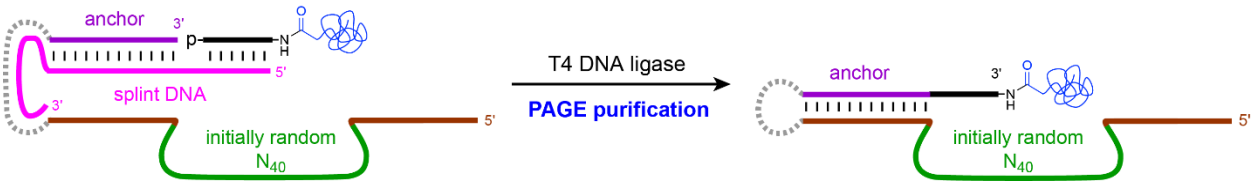


Figure 2.8. Preparation of PAGE shift standard for in vitro selection experiments. A 5′-phosphorylated 3′-amino octamer was conjugated to PEG_{5k}-COOH as described in the synthesis of the oligo-PEG_{5k} conjugate and then ligated to the initially random pool.

2.2.2 Identification and Characterization of Deoxyribozymes from the GA1 Selection Experiment

The GA1 selection experiment was performed. In each selection round, the incubation conditions included 10 μM Imp-oligo-PEG_{5k} in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h. After 7 rounds, the PEGylation activity of the pool reached 32% (Figure 2.9), and individual deoxyribozymes were cloned, which led to the identification of four different enzymes (Figure 2.10). All four sequences have PEGylation activity with 13-36% yield in 24 h (Figure 2.11A), when these enzymes are assayed in cis, i.e., ligating the DNA-anchored peptide substrate to the DNA pool. However, when these active sequences are assayed in trans, i.e., without ligating the DNA-anchored peptide substrate to the pool, only one enzyme, 7GA102, showed 34% PEGylation yield (Figure 2.11B).

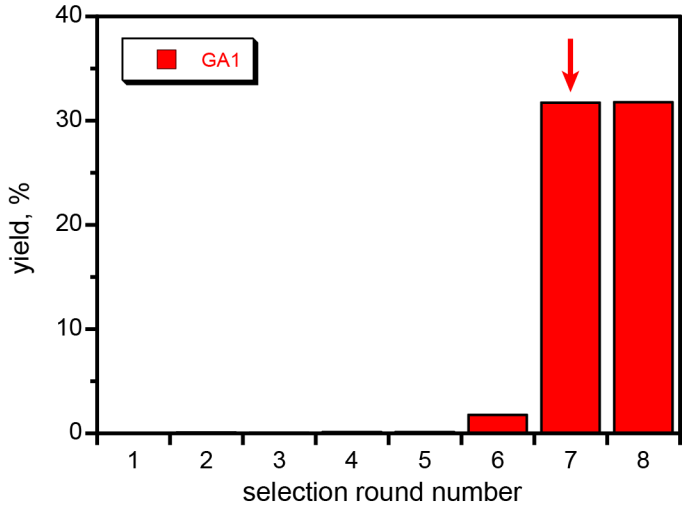


Figure 2.9. Progression of the GA1 selection experiment. Arrow marks the cloned round.

```
                1         10        20        30        40
                |         |         |         |         |
7GA102  ACCTGTTCCG TTTGTATATA GGGACGATAG CGCCGTCTAT 40
7GA106  CTTG.CG..A AG..G.G.GC .AA.GCGCTT ...G.-GGC. 39
7GA107  .GACAAGA.T GCG.GT.GCC CT.CG.GG.T AA.AAGGAGC 40
7GA109  C.GC.ACG.T GCA.A.CG.G T..TTAGGT. A...A.AGGG 40
```

Figure 2.10. Sequences of the initially random region of the PEGylation deoxyribozymes identified from the GA1 selection experiment. A dot denotes conservation with the uppermost sequence, and a dash denotes a gap.

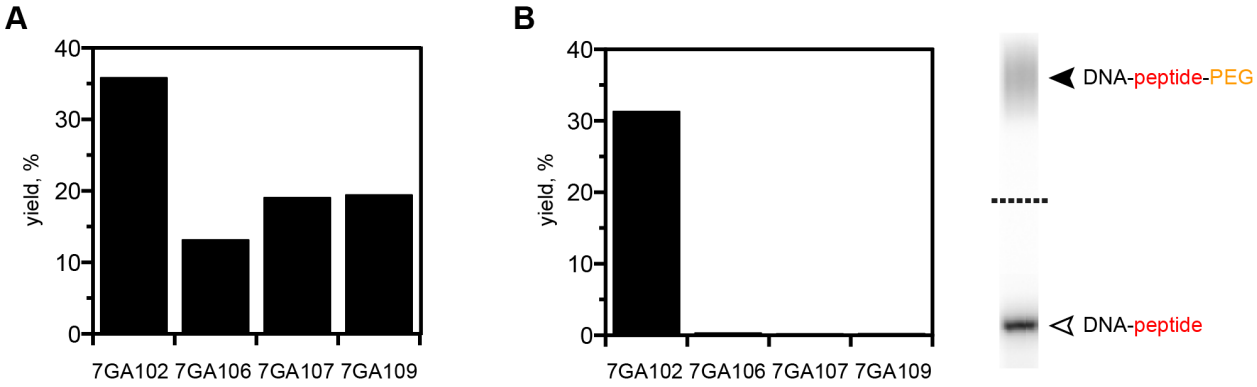


Figure 2.11. Assays of the deoxyribozymes identified from the GA1 selection experiment (yield at 24 h). (A) In cis conditions. (B) In trans conditions. Incubation conditions: 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C with 10 μM Imp-oligo-PEG_{5k}. The PAGE image shows the PEGylation catalyzed by 7GA102 at 24 h.

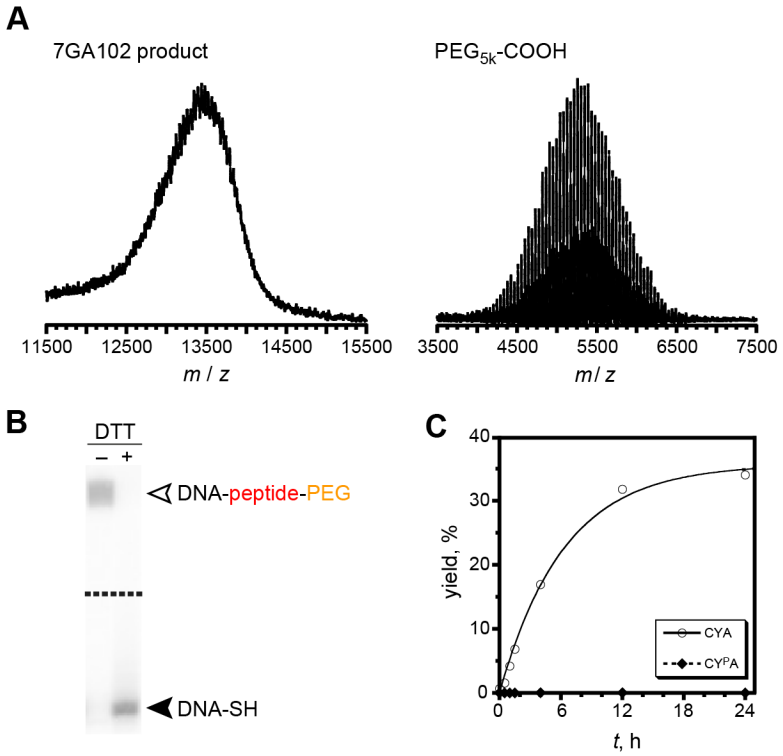
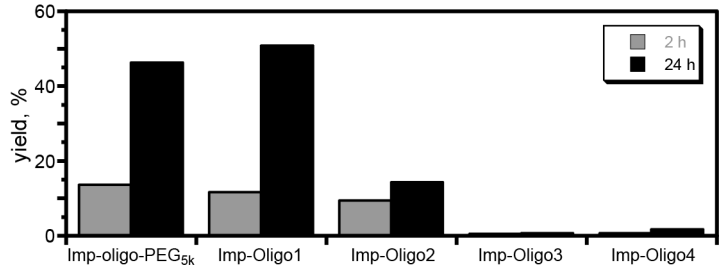


Figure 2.12. Characterization of 7GA102 deoxyribozyme. (A) MALDI mass spectrometry analysis of 7GA102 product and PEG_{5k}-COOH. (B) DTT assay for 7GA102 product. (C) Kinetic comparison of 7GA102 activity with CYA substrate and CY^PA substrate.

Tyrosine PEGylation catalyzed by 7GA102 was verified. MADLI mass spectrometry of the PEGylation product confirms the addition of the PEG moiety to the DNA-anchored peptide substrate (Figure 2.12A). 7GA102 product has a mass distribution with a full width at half maximum (FWHM) of 1000 that is similar to the FWHM for the mass distribution of PEG_{5k}-COOH. Additionally, the average mass of 7GA102 product (~13500) matched with the calculated mass (13500) based on the average mass of PEG_{5k}-COOH. The PEGylation product was reduced with dithiothreitol (DTT) to cleave the disulfide bond between the DNA anchor oligonucleotide and the tripeptide substrate (Figure 2.12B). The large PAGE shift upon DTT reduction suggested that the PEG group is installed on the tripeptide, in comparison with a much smaller PAGE shift induced by the removal of merely the tripeptide if PEGylation occurred at another position (for example, a nucleobase functional group). Moreover, the PEGylation catalyzed by 7GA102 is specific for tyrosine. With the DNA-anchored CYA substrate, 7GA102

deoxyribozyme has 34% PEGylation yield with k_{obs} of 0.16 h^{-1} in 24 h (Figure 2.12C). On the contrary, the DNA-anchored phosphorylated tyrosine substrate (CY^PA) is not a substrate.

The interaction between 7GA102 deoxyribozyme and the Imp substrate was further evaluated. Several 5′-phosphorylated 20-mer oligonucleotides were synthesized. After the Imp activation, these 20-mer oligonucleotides were used in place of Imp-oligo-PEG_{5k} for the in trans assay (Figure 2.13). Catalytic activity with 51% yield in 24 h was observed when all five nucleotides (CGTGC) from the Imp-oligo-PEG_{5k} were conserved at the 5′-end of the 20-mer oligonucleotide (p-CGTGCAACAACAACAACAAC, referred to as Oligo1), establishing that the pentamer segment in the Imp substrate is the recognition motif by 7GA102 enzyme. A lower yield (14% yield in 24 h) was observed for 7GA102 deoxyribozyme when it was incubated with the Imp-activated 20-mer oligonucleotide substrate (p-CGCAACGCAATTAATGTGAG, referred to as Oligo2) in which only the first two nucleotides (CG) were conserved. Catalytic activity of 7GA012 enzyme was abolished (<1% yield in 24 h) when the Imp-activated 20-mer oligonucleotide substrate used (p-CTACCTTTATGCGTATCGAA, referred to as Oligo3) has only a single nucleotide (5′-C) conserved. The abolish of activity with Oligo3 suggested that a single nucleotide is not sufficient for the recognition by 7GA102 deoxyribozyme. The activity of 7GA102 deoxyribozyme was also abolished when the Imp-activated 20-mer oligonucleotide substrate used (p-AACGGCCGCATCAGACAGCG, referred to as Oligo4) shares no conservation with the Imp-oligo-PEG_{5k} substrate. The data above also suggests that 7GA102 enzyme can be used for the introduction of almost any modification X, by using Imp-CGTGC-X in place of Imp-oligo-PEG_{5k} during the catalysis. However, some synthetic effort might be required to prepare the p-CGTGC-X conjugate. For the sake of discussion, deoxyribozymes identified with Imp-oligo-PEG_{5k} substrate will still be referred to as PEGylation deoxyribozymes in this dissertation.



oligonucleotide substrate	sequence
Oligo1	p-CGTGCAACAACAACAACAAC
Oligo2	p-CGCAACGCAATTAATGTGAG
Oligo3	p-CTACCTTTATGCGTATCGAA
Oligo4	p-AACGGCCGCATCAGACAGCG

Figure 2.13. 7GA102 reactivity with different Imp substrates. The table shows the sequences of oligonucleotides used. p represents a phosphoryl group. Underlined are conserved nucleotides from the Imp-oligo-PEG_{5k} substrate used in the selection experiment.

2.2.3 Reselection of 7GA102 Deoxyribozyme

Reselection of 7GA102 deoxyribozyme was performed in the KJ1 selection experiment by generating a partially randomized pool (25% randomization). The fixed-sequence binding arms flanking the initially random region were changed to prevent contamination of the reselection experiment by the parent deoxyribozyme. A hexapeptide substrate CAAYAA was used instead of the tripeptide CYA. Activity was observed in reselection round 8 with 22% PEGylation yield (Figure 2.14A). However, when the active pool was assayed in trans, PEGylation was almost abolished with only 1.2% yield in 24 h (data not shown). This observation that selection pool is only active in cis was unexpected for this reselection experiment, because the parent deoxyribozyme for the reselection experiment, 7GA102, is active both in cis and in trans. One reasonable explanation is that before deoxyribozymes with desired activity were able to emerge, the population of selection pool had already been taken over by sequences that catalyze side reactions. A possible side reaction is the PEGylation of 5′-hydroxyl group, which was supported by the elimination of in cis activity when 5′-hydroxyl group was replaced with 5′-phosphoryl group (Figure 2.14B).

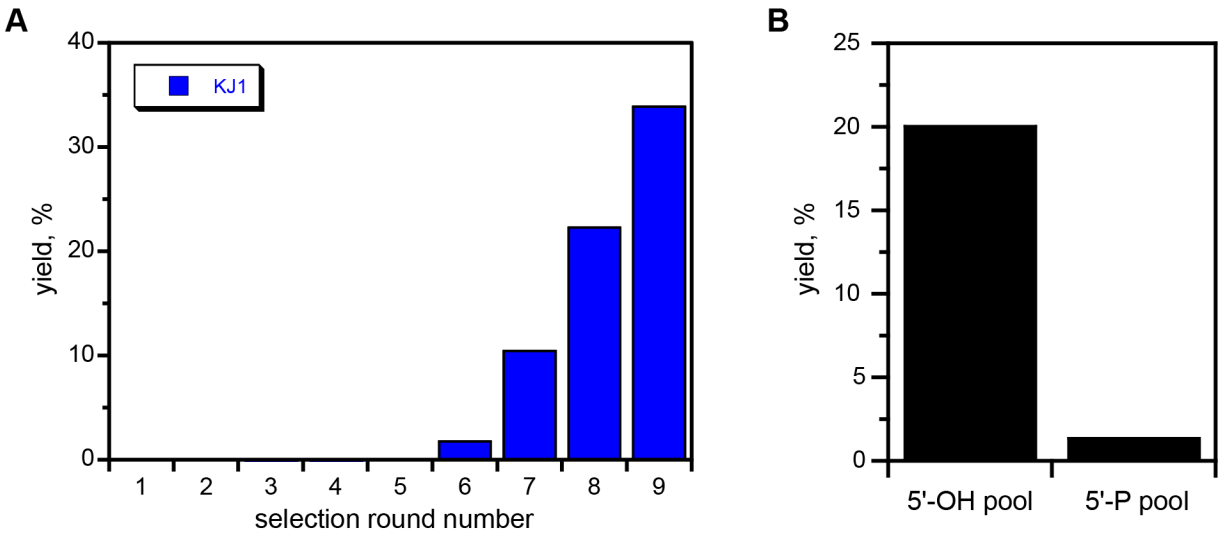


Figure 2.14. Evaluation of the KJ1 reselection experiment. (A) Progression of the selection experiment. (B) In cis assay of KJ1 pool from round 8 with 5′-hydroxyl group and 5′-phosphoryl group (yield at 24 h).

To avoid the emergence of sequences that PEGylate 5′-hydroxyl group, another reselection experiment of 7GA102 deoxyribozyme, NH1, was initiated from DNA pool with an additional 5′-methoxy-T at the 5′-end (Figure 2.15A). Primer oligonucleotide with 5′-methoxy-T at the 5′-end was also used during PCR step. After 7 rounds, the PEGylation activity of the pool reached 38% (Figure 2.15B). The active pool from round 7 showed robust PEGylation activity under both in cis and in trans conditions (Figure 2.15C). Time pressure was applied at round 7 by incubating the selection sample at 37 °C for 2 h. Upon the imposition of time pressure, the activity did not decrease much, indicating that fast deoxyribozymes had already dominated the pool. The active pool from round 7 was cloned, and eight deoxyribozymes were identified (Figure 2.16). The sequence alignment of 7GA102 and its reselected variants revealed that the 3′-half of the initially random region was highly conserved.

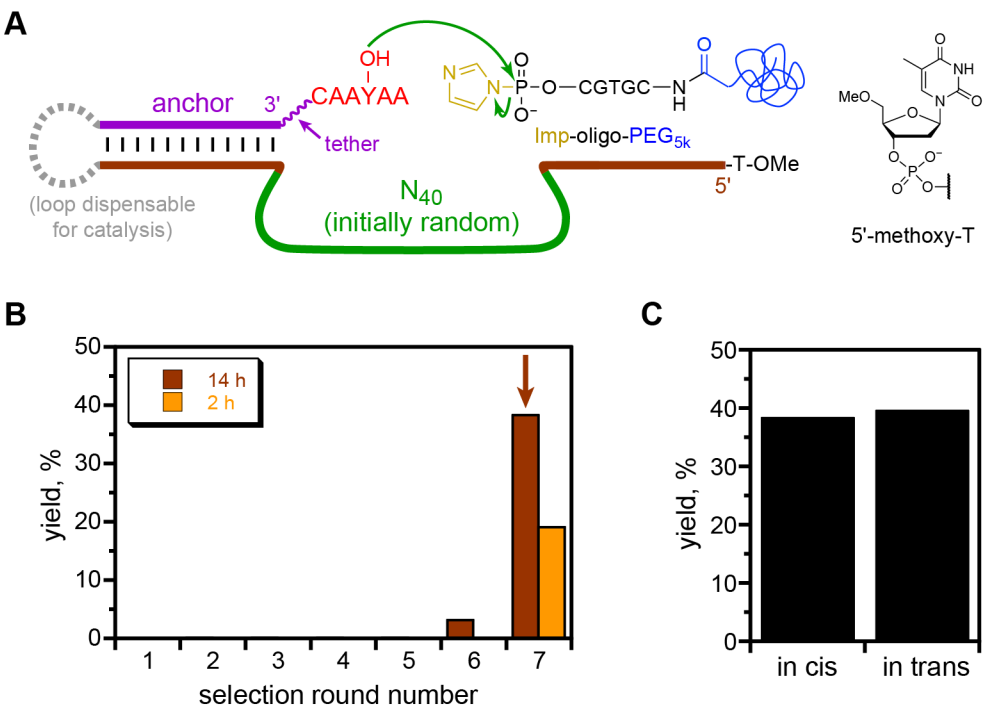


Figure 2.15. Evaluation of the NH1 reselection experiment. (A) Design of the NH1 selection experiment. The 5′-end of the initially random pool was protected by 5′-methoxy-T, the structure of which is shown on the right. (B) Progression of the selection experiment. Arrow marks the cloned round. (C) Assay of the round 7 active pool under in cis and in trans conditions (yield at 24 h).

```
                1                   10                  20                  30                  40
                |                   |                   |                   |                   |
7GA102  ACCTGTTCCG  TTTGTATATA  GGGACGATAG  CGCCGTCTAT 40
7NH101  CGT.CAA.AT  CAG...G...  ....G.T...  ......AG.. 40
7NH102  CGT.AAA..T  AAGTA.G...  ..........  .......AGA 40
7NH103  CGT.AAA..T  AAATA.G...  ..........  .......AGA 40
7NH108  CGT.TAA...  .AA.ATG...  ..........  .......A.A 40
7NH118  CGT.AAA..T  AAATG.G...  ..........  .......AGA 40
7NH122  CGT.AAA..C  AAATA.G...  ..........  .......AGA 40
7NH132  CGT.TAA...  CAA.ATG...  ..........  .......A.A 40
7NH136  CGT.CAA...  ....A.G...  ....A.....  .......A.A 40
```

Figure 2.16. Sequences of the initially random region of the PEGylation deoxyribozymes identified from the NH1 selection experiment. The sequence of 7GA102 deoxyribozyme is also included. A dot denotes conservation with the uppermost sequence. A strongly conserved region is observed at the 3′-end of the sequences, indicated by the grey box.

Two representative deoxyribozymes, 7NH101 and 7NH108, were synthesized and characterized. Both enzymes show robust PEGylation activity (56% and 40%, respectively) on the

tethered CAAYAA substrate after 24 h incubation (Figure 2.17A). 7NH101 and 7NH108 deoxyribozymes can also utilize the 20-mer DNA substrate Oligo1 after the Imp activation (Figure 2.17B).

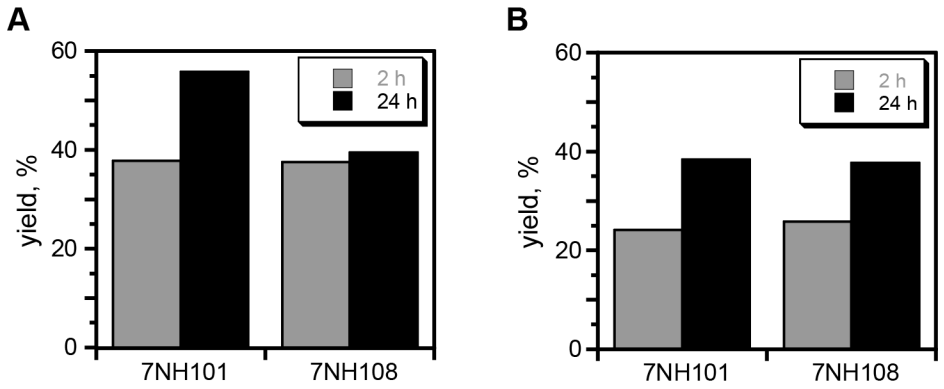


Figure 2.17. Evaluation of the deoxyribozymes identified from the NH1 selection experiment. (A) Single-turnover PEGylation activity of 7NH101 and 7NH108. (B) Single-turnover catalytic activity of 7NH101 and 7NH108 with the activated 20-mer DNA substrate Oligo1.

2.2.4 Selection Efforts towards Untethered Peptide Activity with Long Tether

For the eventual goal of peptide and protein modification, deoxyribozymes such as 7GA102 will be most useful if they can function with free peptide substrates that are not tethered to a DNA anchor oligonucleotide. Unfortunately, neither 7GA102 nor its reselected variants catalyze PEGylation of untethered peptide substrate.

To identify PEGylation deoxyribozymes with untethered peptide activity, the NJ1 selection experiment was performed with a hexapeptide CAAYAA covalently attached via a disulfide linkage and a HEG_{10} tether to the DNA anchor (Figure 2.18A). Activity was observed in round 9 with 6.5% PEGylation yield (Figure 2.18B). However, when the active pool was assayed in trans, catalyzed PEGylation was almost abolished with <0.5% yield in 24 h (data not shown). When in cis selection sample from round 8 was treated with DTT, only a small PAGE shift was observed for the PEGylated pool (Figure 2.18C), suggesting that the PEG moiety was not installed on the peptide but at another position (for example, a nucleobase functional group). This emergence of DNA sequences that catalyze side reactions is similar to the result of the KJ1 selection experiment.

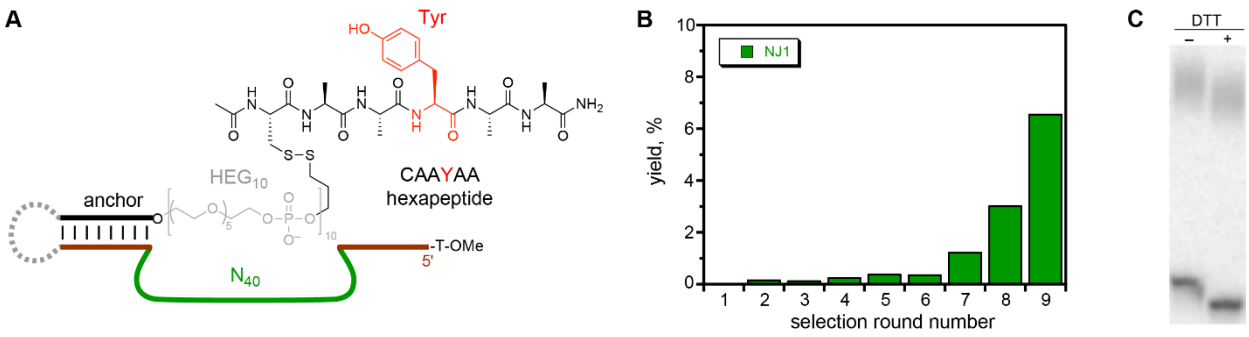


Figure 2.18. NJ1 selection experiment with HEG_{10} tether. (A) Arrangement of initially random DNA pool (N_{40}) and peptide substrate for the NJ1 selection experiment. (B) Progression of the selection experiment. (C) DTT assay for the round 8 active pool.

Extensive study was performed by former Silverman lab graduate student Shannon Walsh using HEG_{10}-tethered peptides during in vitro selection experiments to identify kinase deoxyribozymes with untethered peptide activity.[23] Yet, no activity was observed for any of these selection experiments. Therefore, the tether dependence was evaluated for the nucleopeptide-forming 8XJ105 deoxyribozyme that was identified from a selection experiment using untethered peptide substrates (Figure 2.19). Use of a tethered peptide substrate was expected to result in an increased yield compared to an untethered peptide substrate, and the peptide substrate connected via a longer tether such as the HEG_{10} tether was expected to react similarly to the untethered peptide substrate. As expected with the untethered peptide substrate, 8XJ105 catalyzes nucleopeptide formation in 30% yield. The nucleopeptide formation yield increases to 50% with the DNA-anchored HEG-tethered peptide substrate. Surprisingly, the HEG_6 and HEG_{10}-tethered peptide substrates are modified with only <5% yield. These results suggested that the initial hypothesis was incorrect; and a long tether between the DNA anchor and peptide substrate did not mimic an untethered peptide substrate. The long tether was not inert as expected and instead prevented the 8XJ105 deoxyribozyme from catalyzing the reaction of the peptide substrate. These results are similar to those observed for the 15MZ36 deoxyribozyme.[24] 15MZ36 catalyzes nucleopeptide formation in 75%, 60%, and 10% yield with C_3-tethered, HEG-tethered, and untethered CYA, respectively. However, HEG_{10}-tethered CYA is not modified (<0.5% in 24 h) by 15MZ36. Combined, these results suggest that the use of tethers longer than HEG interfere with

DNA-catalyzed peptide modification. On this basis, the use of tethers longer than HEG during selection experiments was discontinued.

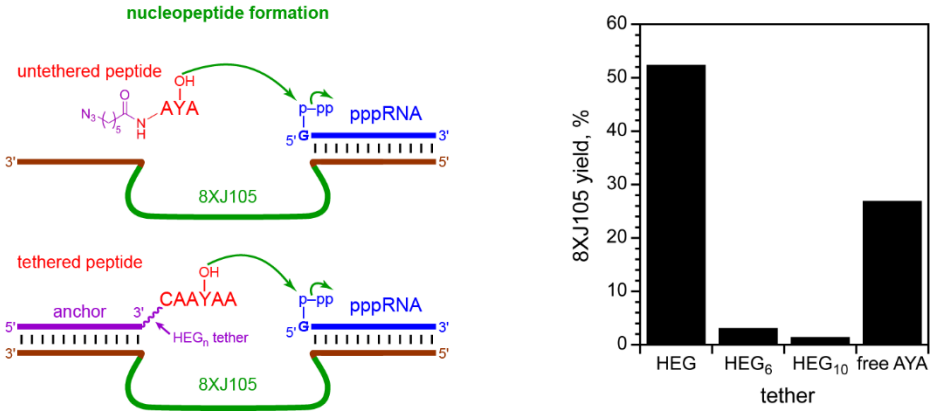


Figure 2.19. The tether dependence of the nucleopeptide-forming 8XJ105 deoxyribozyme.[23] Untethered and tethered peptide substrates were evaluated. 8XJ105 successfully catalyzes the reaction of a free peptide and HEG-tethered peptide substrate. However, peptide substrates connected via a HEG_6 or HEG_{10} tether were not accepted as substrates. Figure adapted with permission from ref. 23.

2.2.5 Selection Efforts towards Untethered Peptide Activity Using Alternating Tethers

Since the long tether between the peptide substrate and DNA anchor oligonucleotide does not mimic untethered peptide substrates, a new method to identify deoxyribozymes with untethered peptide reactivity is needed. In vitro selection experiments with alternating positions of the tethers between the peptide substrate and the DNA anchor are intended to identify deoxyribozymes that perform catalysis without the requirement of any tether (Figure 2.20).

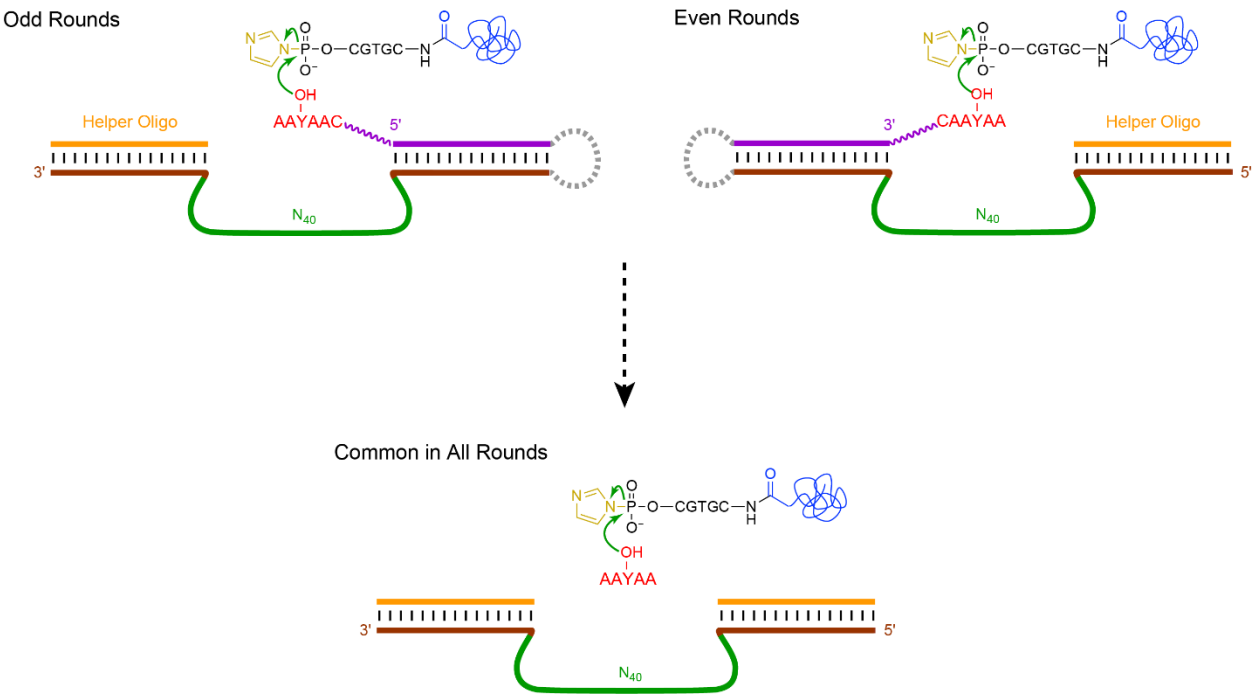


Figure 2.20. The in vitro selection strategy using alternating tethers.

In all previous efforts with PEGylation deoxyribozymes, in vitro selection experiments were performed with L substrates, i.e., peptide substrates tethered to the DNA anchors that were bound to the left-hand binding arms of DNA pools by Watson-Crick base pairs. The PN1 selection experiment was performed to validate that PEGylation deoxyribozymes can also be identified from selection experiments using R substrates, peptide substrates tethered to the DNA anchors base-paired with the right-hand binding arms of DNA pools (Figure 2.21A). Considering the side reactions observed in the KJ1 and NJ1 selection experiments, an additional DTT reduction step was included. After the selection step, the PEGylated pool purified by PAGE was reduced by DTT, and another PAGE separation was performed. If the PEG group is attached not at the peptide substrate but instead at some undesired position within the DNA itself (for example, the 5′-hydroxyl group or a nucleobase functional group), then disulfide cleavage will lead to only a small PAGE shift, in comparison with the much larger PAGE shift upon removal of the entire PEG-modified hexapeptide. Therefore, the difference in PAGE shift enables the separation of DNA enzymes with tyrosine PEGylation activity from sequences that catalyze side reactions. After 7

rounds, the PEGylation activity of the pool reached 16% (Figure 2.21B), and individual deoxyribozymes were cloned, which led to the identification of two different enzymes, 7PN106 and 7PN113 (Figure 2.22). Both enzymes show PEGylation activity (5.6% and 14%, respectively) on the tethered AAYAAC substrate after 24 h incubation (Figure 2.23).

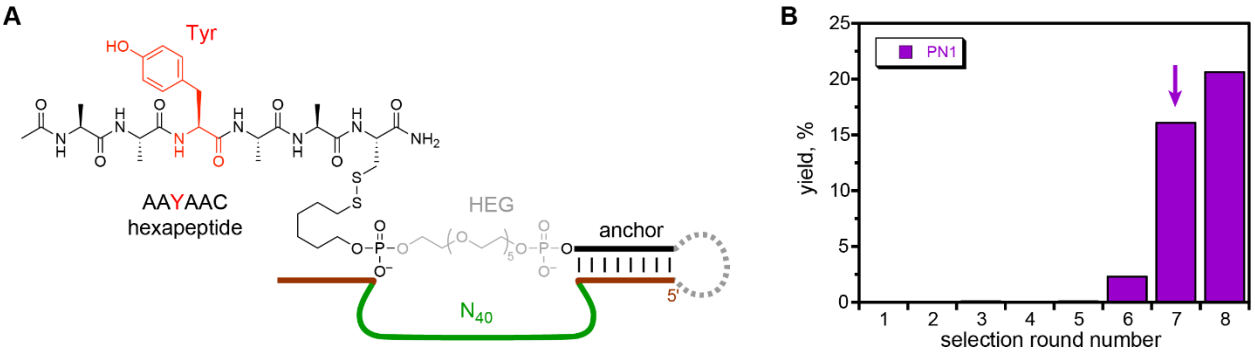


Figure 2.21. The PN1 selection experiment using the R substrate. (A) Arrangement of initially random DNA pool (N_{40}) and peptide substrate. (B) Progression of the selection experiment. Arrow marks the cloned round.

```
          1         10        20        30        40
          |         |         |         |         |
7PN106    CACCTCGCAG TAAAGGATGG AAAGTTCTAT AACCAGCGCT 40
7PN113    AGGGCG.GG. ATTG..GGC. ..GCGCT.TG TGAG.C.C.C 40
```

Figure 2.22. Sequences of the initially random region of the PEGylation deoxyribozymes identified from the PN1 selection experiment. A dot denotes conservation with the uppermost sequence.

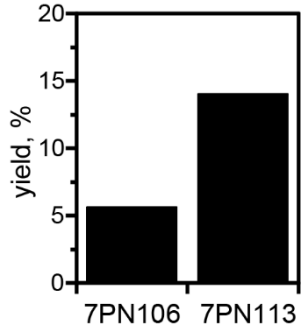


Figure 2.23. Single-turnover PEGylation activity of 7PN106 and 7PN113 (yield at 24 h).

After the success in identifying deoxyribozyme from the PN1 selection experiment, a new in vitro selection experiment, TX1, was performed with alternating tether positions during odd and even rounds of the selection experiment (Figure 2.20). PEGylation was observed in round 7 with 19% activity and in round 8 with 12% activity (Figure 2.24). The active pool from round 7 was cloned, and three deoxyribozymes, 7TX101, 7TX107, and 7TX113, were identified (Figure 2.25). All of these three deoxyribozymes have PEGylation yield of 42-69% in 24 h with the L substrate that was used during odd rounds (Figure 2.26A), while only 7TX101 and 7TX113 show robust PEGylation activity with the R substrate (37% and 16%, respectively). Similar to 7GA102 and its reselected variants, all three enzymes require only the pentamer segment in the Imp-oligo-PEG_{5k} for catalysis (Figure 2.26B). Untethered peptide activity was evaluated by PAGE using 5′-^{32}P-radiolabeled 20-mer DNA oligonucleotide substrate, Oligo1, for the two enzymes, 7TX101 and 7TX113, that are active with both L and R substrates. However, no addition of the untethered peptide to Imp-activated Oligo1 was observed after 24 h incubation with either deoxyribozyme (Figure 2.26C). This result suggests that these two deoxyribozymes require a tethered peptide substrate for catalysis, though they do not require a specific tether position. This result also indicates a lack of interaction between the peptide substrate and the DNA pool during the TX1 selection experiment.

In three previous efforts with tyrosine-modifying deoxyribozymes, deoxyribozymes with better interaction with their peptide substrates have been found from selection experiments in which mixed-sequence peptides are used.[22,25,26] Therefore, two new selection experiments, YR1 and YS1, were designed, using the respective DNA-anchored peptides, TDNYTRL and VKKYLNS (Figure 2.27). Both peptide sequences are derived from human vasoactive intestinal peptide (VIP), which is involved in the regulation of circadian rhythm.[27] The YR1 and YS1 selection experiments are currently being performed by Silverman lab undergraduate student James Brady.

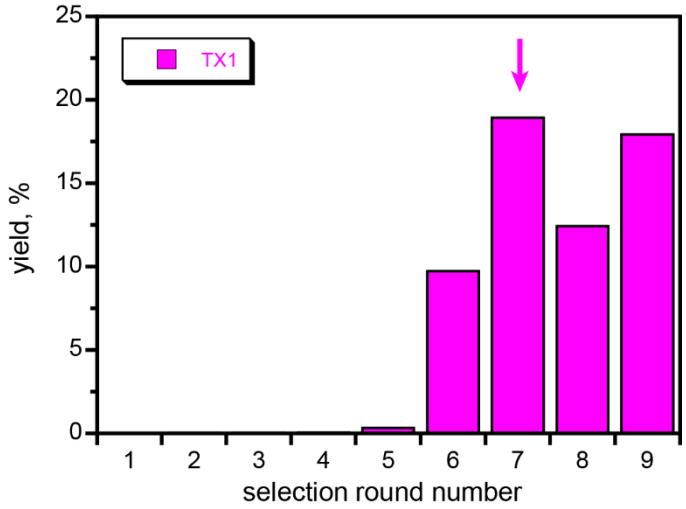


Figure 2.24. Progression of the TX1 selection experiment. Arrow marks the cloned round.

```
                1                  10                  20                  30                  40
                |                  |                   |                   |                   |
7TX101  GCACGGAGAC  ACAGCACATA  GCTTTCTTTC  CGGTAAATCA 40
7TX107  AAG..CG.CG  .ATCAT..GT  AGACAGCAAG  ...AGGTG.G 40
7TX113  .....AGCCT  GAC.TTTTC.  CTCGCAGAGG  GTC.GCGAGC 40
```

Figure 2.25. Sequences of the initially random region of the PEGylation deoxyribozymes identified from the TX1 selection experiment. A dot denotes conservation with the uppermost sequence.

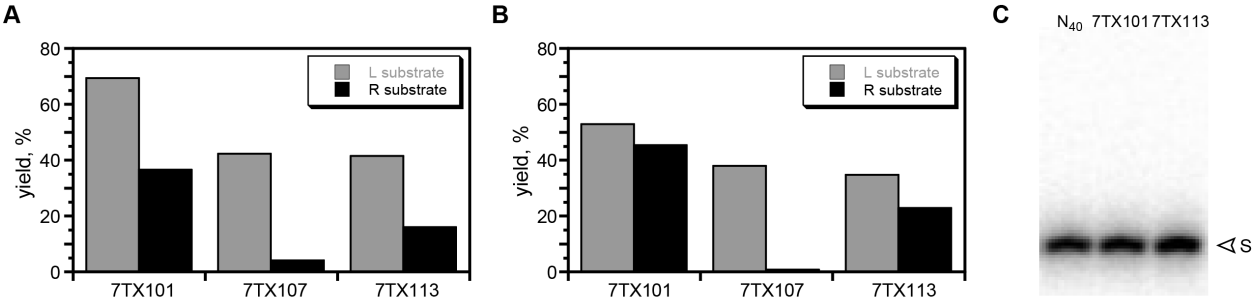


Figure 2.26. Evaluation of deoxyribozymes identified from the TX1 selection experiment (yield at 24 h). (A) Single-turnover PEGylation activity assayed with both L and R substrates. (B) Single-turnover activity with the activated 20-mer DNA substrate Oligo1. (C) Untethered peptide assay with 7TX101 and 7TX113 (S = substrate).

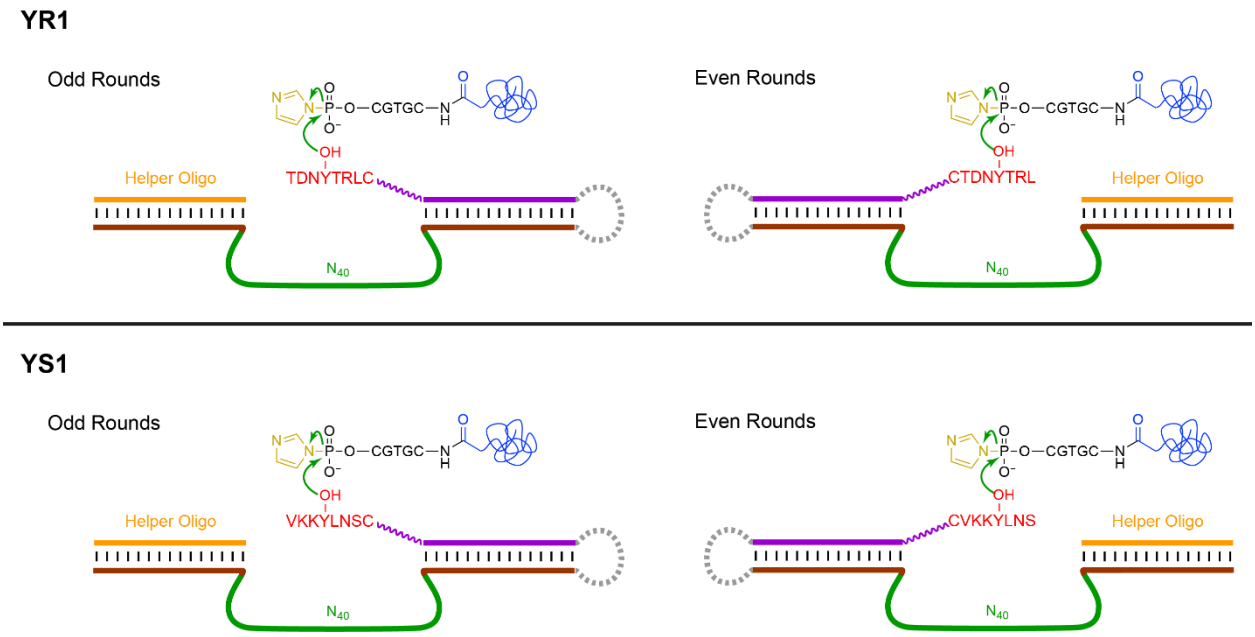


Figure 2.27. Design of in vitro selection experiments YR1 and YS1.

2.3 Summary and Future Directions

PEGylation is an important artificial post-translational modification for therapeutic peptides and proteins. PEGylation improves the pharmacokinetic properties of biopharmaceuticals by increasing circulation half-time, reducing immunogenicity, increasing solubility, and suppressing aggregation. Peptide and protein PEGylation is most commonly achieved by solely chemical means. However, these chemical strategies generally lack site selectivity among different target sites in the substrates. Some chemical strategies also suffer from off-target reactivity, i.e., poor chemo-selectivity. Enzymatic approaches for PEGylation have also been developed, yet their application is limited by substrate specificity and sequence selectivity of the natural enzymes used. These issues of current PEGylation methods can be addressed by de novo identification of deoxyribozymes.

DNA enzymes were identified for PEGylation of tyrosine in a peptide substrate using a 5′-phosphorimidazolide-activated oligonucleotide-PEG conjugate (Imp-oligo-PEG_{5k}) as the PEG donor. The first deoxyribozyme, 7GA102, validated the in vitro selection strategy. Reselection of

7GA102 resulted in the emergence of DNA sequences that PEGylate the 5′-hydroxyl group of the DNA pool. Another reselection experiment performed with an additional 5′-methoxy-T at the 5′-end of the pool provided eight sequence variants of 7GA012. However, none of these deoxyribozymes functions with untethered peptide substrate.

The NJ1 selection experiment was performed using the hexapeptide substrate CAAYAA connected to the DNA anchor with long tether. However, only DNA sequences that catalyze side reactions were enriched in the selection experiment. The study on tether dependence of nucleopeptide-forming deoxyribozymes with untethered peptide activity undermined the original hypothesis that a sufficiently long tether between the peptide and DNA anchor would mimic an untethered peptide substrate. Thus, the use of long tether during selection experiments was discontinued.

A new method to identify deoxyribozymes with untethered peptide activity was developed. In vitro selection experiments with alternating positions of the tether between the peptide substrate and the DNA anchor are expected to identify deoxyribozymes that perform catalysis without the requirement of any tether. Moreover, after the selection step, the PEGylated pool purified by PAGE was reduced by DTT and another PAGE separation was performed. This additional DTT/PAGE operation was designed to avoid emergence of DNA sequences that catalyze undesired reactions. 7PN106 and 7PN113 deoxyribozymes were successfully identified from the PN1 selection experiment using peptide substrate tethered to DNA anchor base-paired with the right binding arm of the DNA pool. With the success of the PN1 selection experiment, the TX1 selection experiment with alternating tether positions was performed, leading to the identification of 7TX101 and 7TX113 deoxyribozymes. However, no untethered peptide activity was observed for either DNA enzyme, even though robust PEGylation was catalyzed by both enzymes with tethered peptide substrates used during their identification. This result suggests that more interaction between the peptide substrate and the DNA pool is needed. Ongoing selection experiments YR1 and YS1 are using mixed-sequence peptide substrates with alternating tether positions for identification of deoxyribozymes with untethered peptide reactivity.

2.4 Materials and Methods

2.4.1 Substrate Preparation Procedures

Oligonucleotides and reagents. 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. 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.[28,29] PEG_{5k}-COOH was obtained from JenKem Technology USA (Plano, TX).

Solid-phase peptide synthesis. All amino acid monomers were obtained from Chem-Impex (Wood Dale, IL). Peptides were prepared by solid-phase synthesis using *N,N,N′,N′*-tetramethyl-*O*-(7-azabenzotriazol-1-yl)uronium hexafluorophosphate (HATU) as coupling agent, and using Fmoc Rink amide MBHA resin from Chem-Impex. Each synthesis was performed at 0.2 mmol scale, initiated using 0.260 g of Rink amide resin with a loading capacity of 0.77 mmol/g. All peptide synthesis operations were performed as previously described.[30]

Synthesis of DNA-anchored peptide conjugates. DNA-anchored peptide conjugates were synthesized by disulfide formation between a DNA HEG-tethered 5′- or 3′-thiol and the N-terminal or C-terminal cysteine side chain of the peptide, as previously reported [HEG = hexa(ethylene glycol)].[31]The 5′-thiol DNA anchor oligonucleotide was 5′-HO-C_6-SS-C_6-p-HEG-X-3′, and the 3′-thiol DNA anchor oligonucleotide was 5′-X-HEG-p-C_3-SS-C_3-OH-3′, where X represents the specific oligonucleotide sequence (Table 2.2). The 5′- or 3′-disulfide linker was introduced via standard solid-phase DNA synthesis and unmasked to a thiol by DTT treatment. The free thiol was activated as the pyridyl disulfide by treatment with 2,2′-dipyridyl disulfide in DMF. The activated DNA substrate was coupled with the free thiol of the peptide N-terminal or C-terminal Cys, forming the DNA-anchored peptide conjugate. All steps were performed as previously described.[31]

Synthesis of oligo-PEG_{5k} conjugate. The oligo-PEG_{5k} conjugate was synthesized by amide formation between a DNA 3′-amino group and the carboxyl group of the PEG_{5k}-COOH. The 3′-amino DNA oligonucleotide was 5′-p-X-p-C_7-NH_2-3′, where X represents the specific

oligonucleotide sequence (Table 2.2). A sample containing 5 nmol of 3′-amino oligonucleotide was brought to 100 μL total volume containing 100 mM MES, pH 6.0, 5 mM PEG_{5k}-COOH, 5 mM *N*-hydroxysuccinimide (NHS), 50 mM *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide (EDC), and 10% (v/v) DMF. The sample was incubated at room temperature for 12 h and purified by 15% PAGE.

Procedures for the preparation of 5′-Imp substrate. A sample containing 4 nmol of 5′-phosphorylated oligonucleotide was brought to 40 μL total volume containing 100 mM EDC and 100 mM imidazole (pH 6.0 with HCl). The sample was incubated at room temperature for 2 h. A Micro Bio-Spin P-6 desalting column (Bio-Rad) was prepared by centrifuging at 1000 *g* for 1 min and rinsing 4× by adding 500 μL of water followed by centrifuging at 1000 *g* for 1 min. The 40 μL sample was applied to the column and eluted by centrifuging at 1000 *g* for 4 min. The eluent was quantified by UV absorbance (A_{260}). The volume of the sample was brought down by SpeedVac so that the concentration of 5′-Imp substrate was 100 μM.

Procedure for 3′-^{32}P-radiolabeling of AAYAAC-HEG-DNA substrate. The AAYAAC-HEG-DNA substrate was 3′-^{32}P-radiolabeled by incubating 20 pmol of the conjugate, 20 μCi of α-^{32}P-dCTP (800 Ci/mmol), and 10 units of terminal deoxytransferase (Fermentas) in 20 μL of 1× TdT reaction buffer (200 mM potassium cacodylate, 25 mM Tris, pH 7.2, 0.01% Triton X-100, and 1 mM $CoCl_2$) at 37 °C for 30 min. The sample was purified by 20% PAGE.

Table 2.2: Oligonucleotide sequences used in the selection experiments. All sequences are written 5′ to 3′. **T** represents 5′-methoxy-T. p represents a phosphoryl group. For all selections, the reverse PCR primer was (AAC)$_4$XCCATCAGGATCAGCT, where X denotes the hexa(ethylene glycol) [HEG] spacer to stop Taq polymerase.

oligonucleotide purpose	oligonucleotide sequence
pentamer DNA for oligo-PEG conjugate	p-CGTGC-C_7-NH_2
Preparation of selection standard	
octamer DNA for oligo-PEG conjugate	p-TGGGTGCG-C_7-NH_2
random pool for standard preparation	CGAAGTCGCCATCTCTTC-N_{40}-ATAGTGAGTCGTATTAAGCTGATCCTGATGGGGATAATACGACTCACTAT
splint for ligation step in standard preparation	CGCACCCAATAGTGAGTCGTATTATCCCCATCAGGATCAGCT

Table 2.2 (cont.)

oligonucleotide purpose	oligonucleotide sequence
Selection with DNA-HEG-CYA (leading to 7GA102)	
DNA-HEG-CYA substrate	GGATAATACGACTCACTAT-HEG-CYA
forward primer for selection	CGAAGTCGCCATCTCTTC
random pool for selection	CGAAGTCGCCATCTCTTC-N_{40}-ATAGTGAGTCGTATTAAGCTGATCCTGATGG
splint for ligation step during selection	ATATGTCTTTCAATAGTCCCCATCAGGATCAGCTCTATTGAAAGACATAT
DNA substrate for evaluation of Imp substrate requirement	
DNA substrate Oligo1	p-CGTGCAACAACAACAACAAC
DNA substrate Oligo2	p-CGCAACGCAATTAATGTGAG
DNA substrate Oligo3	p-CTACCTTTATGCGTATCGAA
DNA substrate Oligo4	p-AACGGCCGCATCAGACAGCG
Reselection of 7G102 deoxyribozyme (KJ1 selection)	
DNA-HEG-CAAYAA substrate	GGAATATCTCGTTTCTTAT-HEG-CAAYAA
forward primer for selection	CGAATTAAGACTGAATTC
partially randomized pool for selection	CGAAGTATAAACCTGTTC-N_{40}-ATAAGAAACGAGATATAGCTGATCCTGATGG
splint for ligation step during selection	ATAAGAAACGAGATATTCCCCATCAGGATCAGCTATATCTCGTTTCTTAT
Reselection of 7GA102 deoxyribozyme (leading to 7NH101 and 7NH108)	
DNA-HEG-CAAYAA substrate	GGATCCTGGATACAAATAT-HEG-CAAYAA
forward primer for selection	TCGAAGTATAAACCTGTTC
random pool for selection	TCGAAGTATAAACCTGTTC-N_{40}-ATATTTGTATCCAGGAAGCTGATCCTGATGG
splint for ligation step during selection	ATATTTGTATCCAGGATCCCCATCAGGATCAGCTTCCTGGATACAAATAT
Selection with DNA-HEG_{10}-CAAYAA (NJ1 selection)	
DNA-HEG_{10}-CAAYAA substrate	GGATAATACGACTCACTAT-HEG_{10}-CAAYAA
forward primer for selection	TCGAAGTCGCCATCTCTTC
random pool for selection	TCGAAGTCGCCATCTCTTC-N_{40}-ATAGTGAGTCGTATTAAGCTGATCCTGATGG
splint for ligation step during selection	ATATGTCTTTCAATAGTCCCCATCAGGATCAGCTCTATTGAAAGACATAT
Selection with AAYAAC-HEG-DNA (leading to 7PN106 and 7PN113)	
AAYAAC-HEG-DNA substrate	AAYAAC-HEG-GAACAGGTTTATACGGA
forward primer for selection	CGAAGTATAAACCTGTTC
random pool for selection	CGAAGTATAAACCTGTTC-N_{40}-ATATTTGTATCCAGGAAGCTGATCCTGATGG
splint for ligation step during selection	GAACAGGTTTATACTTCGTCCGTATAAACCTGTTC

Table 2.2 (cont.)

oligonucleotide purpose	oligonucleotide sequence
Selection with alternating AAYAAC-HEG-DNA and DNA-HEG-CAAYAA (leading to 7TX101, 7TX107, and 7TX113)	
AAYAAC-HEG-DNA substrate (odd rounds)	AAYAAC-HEG-GAAATAGCCATCATGGA
DNA-HEG-CAAYAA substrate (even rounds)	GGACTATTGAAAGACATAT-HEG-CAAYAA
forward primer for selection	p-CGAAATGATGGCTATTTC
random pool for selection	p-CGAAATGATGGCTATTTC-N_{40}-ATATGTCTTTCAATAGAGCTGATCCTGATGG
splint for ligation step in odd round	GAAATAGCCATCATTTCGTCCATGATGGCTATTTC
splint for ligation step in even round	ATATGTCTTTCAATAGTCCCCATCAGGATCAGCTCTATTGAAAGACATAT
helper DNA for selection step in odd round	GGACTATTGAAAGACATAT
helper DNA for selection step in even round	GAAATAGCCATCATGGA
Selection with alternating TDNYTRLC-HEG-DNA and DNA-HEG-CTDNYTRL	
TDNYTRLC-HEG-DNA substrate (odd rounds)	TDNYTRLC-HEG-GAAGGAGGCTTTCGGGA
DNA-HEG-CTDNYTRL substrate (even rounds)	GGACTACCTTTATGCGTAT-HEG-CTDNYTRL
forward primer for selection	p-CGAACGAAAGCCTCCTTC
random pool for selection	p-CGAACGAAAGCCTCCTTC-N_{40}-ATACGCATAAAGGTAGAGCTGATCCTGATGG
splint for ligation step in odd round	GAGGCTTTCGTTCGTCCCGAAAGCCTCCTTC
splint for ligation step in even round	ATACGCATAAAGGTAGTCCCCATCAGGATCAGCTCTACCTTTATGCGTAT
helper DNA for selection step in odd round	GGACTACCTTTATGCGTAT
helper DNA for selection step in even round	GAAGGAGGCTTTCGGGA
Selection with alternating VKKYLNSC-HEG-DNA and DNA-HEG-VKKYLNS	
VKKYLNSC-HEG-DNA substrate (odd rounds)	VKKYLNSC-HEG-GAACTGTTGCAATGGGA
DNA-HEG-CVKKYLNS substrate (even rounds)	GGAGAGCCACTGATGATAT-HEG-CAAYAA
forward primer for selection	p-CGAACATTGCAACAGTTC
random pool for selection	p-CGAACATTGCAACAGTTC-N_{40}-ATATCATCAGTGGCTCAGCTGATCCTGATGG
splint for ligation step in odd round	GAACTGTTGCAATGTTCGTCCCATTGCAACAGTTC
splint for ligation step in even round	ATATCATCAGTGGCTCTCCCCATCAGGATCAGCTGAGCCACTGATGATAT
helper DNA for selection step in odd round	GGAGAGCCACTGATGATAT
helper DNA for selection step in even round	GAACTGTTGCAATGGGA

2.4.2 In Vitro Selection Procedures for GA1, KJ1, NH1, NJ1, and PN1 Selection Experiments

Procedure for ligation step in round 1. A 25 μL sample containing 500 pmol of DNA pool, 630 pmol of DNA splint, and 750 pmol of 5′-phosphorylated DNA-anchored peptide substrate was annealed in 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min

and cooling on ice for 5 min. To this solution was added 3 μL of 10× T4 DNA ligase buffer that lacks DTT (400 mM Tris, pH 7.8, 100 mM $MgCl_2$, and 5 mM ATP) and 2 μL of 5 U/μL T4 DNA ligase (Thermo Fisher). The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

Procedure for ligation step in subsequent rounds. A 17 μL sample containing the PCR-amplified DNA pool (~5–10 pmol), 40 pmol of DNA splint, and 60 pmol of 5′-phosphorylated DNA-anchored peptide substrate was annealed in 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. To this solution was added 2 μL of 10× T4 DNA ligase buffer that lacks DTT (400 mM Tris, pH 7.8, 100 mM $MgCl_2$, and 5 mM ATP) and 1 μL of 1 U/μL T4 DNA ligase (Thermo Fisher). 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 pool. A 16 μL sample containing 200 pmol of 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 40 μL total volume containing 70 mM HEPES, pH 7.5, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$, 150 mM NaCl, and 10 μM Imp-oligo-PEG_{5k}. The Mn^{2+} was added from a 10× stock solution containing 200 mM $MnCl_2$. The Zn^{2+} was added from a 10× stock solution containing 10 mM $ZnCl_2$, 20 mM HNO_3, and 200 mM HEPES at pH 7.5; this stock solution was freshly prepared from a 100× stock of 100 mM $ZnCl_2$ in 200 mM HNO_3. The metal ion stocks were added last to the final sample. The sample was incubated at 37 °C for 14 h and separated by 8% PAGE.

Procedure for selection step in subsequent round. An 8 μL sample containing the 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, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$, 150 mM NaCl, and 10 μM Imp-oligo-PEG_{5k}. The sample was incubated at 37 °C for 14 h and separated by 8% PAGE.

Procedure for DTT reduction in each round for PN1 selection. The sample from the selection step was brought to 20 μL total volume containing 100 mM HEPES, pH 7.5, and 50 mM DTT. The sample was incubated at 37 °C for 2 h and separated by 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 DTT reduction product, 200 pmol of forward primer, 50 pmol of reverse primer, 20 nmol of each dNTP, 10 μL of 10× 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], and Pfu polymerase. This sample was cycled 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. Pfu 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), 5 μL of 10× polymerase buffer, and Pfu polymerase. This sample was cycled according to the following PCR program: 94 °C for 2 min, 30× (94 °C for 30 s, 47 °C for 30 s, 72 °C for 30 s), 72 °C for 5 min. Samples were separated by 8% PAGE.

2.4.3 In Vitro Selection Procedures for TX1, YR1, and YS1 Selection Experiments

Procedures for ligation in even rounds, DTT reduction, and PCR were performed identically as those in GA1, KJ1, NH1, NJ1, and PN1 selection experiments.

Procedure for ligation step in round 1. A 25 μL sample containing 500 pmol of 5′-phosphorylated DNA pool, 630 pmol of DNA splint, and 750 pmol of DNA-anchored peptide substrate was annealed in 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. To this solution was added 3 μL of 10× T4 DNA ligase buffer that lacks DTT (400 mM Tris, pH 7.8, 100 mM $MgCl_2$, and 5 mM ATP) and 2 μL of 5 U/μL T4 DNA ligase (Thermo Fisher). The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

Procedure for ligation step in subsequent odd rounds. A 17 μL sample containing the PCR-amplified 5′-phosphorylated DNA pool (~5–10 pmol), 40 pmol of DNA splint, and 60 pmol of DNA-anchored peptide substrate was annealed in 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. To this solution was added 2 μL of 10× T4 DNA ligase buffer that lacks DTT (400 mM Tris, pH 7.8, 100 mM $MgCl_2$, and 5 mM ATP) and 1 μL of 1 U/μL T4 DNA ligase (Thermo Fisher). 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 pool. A 16 μL sample containing 200 pmol of ligated pool and 250 pmol of helper DNA (included to base pair with one of the fixed-sequence binding arms of the DNA 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, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$, 150 mM NaCl, and 10 μM Imp-oligo-PEG_{5k}. The Mn^{2+} was added from a 10× stock solution containing 200 mM $MnCl_2$. The Zn^{2+} was added from a 10× stock solution containing 10 mM $ZnCl_2$, 20 mM HNO_3, and 200 mM HEPES at pH 7.5; this stock solution was freshly prepared from a 100× stock of 100 mM $ZnCl_2$ in 200 mM HNO_3. The metal ion stocks were added last to the final sample. The sample was incubated at 37 °C for 14 h and separated by 8% PAGE.

Procedure for selection step in subsequent round. An 8 μL sample containing the ligated pool and 50 pmol of helper DNA 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, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$, 150 mM NaCl, and 10 μM Imp-oligo-PEG_{5k}. The sample was incubated at 37 °C for 14 h and separated by 8% PAGE.

2.4.4 Cloning and Screening

Using 1 μL of a 1:1000 dilution of the 10-cycle PCR product from the end of the selection round being cloned, 30-cycle PCR was performed using the above procedure, omitting α-^{32}P-dCTP and using 25 pmol of forward primer and 25 pmol of reverse cloning primer, where the extensions with TAA stop codons in each frame were included to suppress false negatives in blue-white screening.[32] The PCR product was purified on 2% agarose and cloned using a TOPO TA cloning kit (Invitrogen). Miniprep DNA samples derived from individual *E. coli* colonies were assayed by digestion with EcoRI to ascertain the presence of the expected inserts. Using the miniprep DNA samples, PCR (same conditions as 30-cycle PCR during selection, omitting α-^{32}P-dCTP) was performed to obtain the deoxyribozymes. Half of the PCR product was used to perform screening assays. Each screening assay used ~0.1 pmol of 5′-^{32}P-radiolabeled DNA-anchored peptide substrate or 3′-^{32}P-radiolabeled DNA-anchored peptide substrate, 50 pmol of helper DNA (if needed), and ~20 pmol of deoxyribozyme and followed the single-turnover assay procedure described in the subsequent section.

2.4.5 Single-Turnover Deoxyribozyme in cis Assay Procedure

The DNA-anchored peptide substrate was 5′-^{32}P-radiolabeled using γ-^{32}P-ATP and T4 polynucleotide kinase. A 17 μL sample containing 0.4 pmol of 5′-^{32}P-radiolabeled DNA-anchored peptide substrate, 2 pmol of DNA splint, and 4 pmol of deoxyribozyme 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 that lacks DTT (400 mM Tris, pH 7.8, 100 mM $MgCl_2$, and 5 mM ATP) and 1 μL of 1 U/μL T4 DNA ligase (Thermo Fisher). The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

A 4 μL sample containing the DNA-anchored peptide substrate ligated to an individual 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 10 μL total volume containing 70 mM HEPES, pH 7.5, 40 mM $MgCl_2$,

20 mM $MnCl_2$, 1 mM $ZnCl_2$, 150 mM NaCl, and 10 μM freshly activated Imp substrate. The sample was incubated at 37 °C. At appropriate time points, a 2 μL aliquot was quenched with 6 μL of 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 8% PAGE and quantified with a PhosphorImager.

2.4.6 Single-Turnover Deoxyribozyme in trans Assay Procedure

The DNA-anchored peptide substrate was 5′-^{32}P-radiolabeled using γ-^{32}P-ATP and T4 polynucleotide kinase. An 8 μL sample containing 0.1 pmol of 5′-^{32}P-radiolabeled DNA-anchored peptide substrate or 3′-^{32}P-radiolabeled DNA-anchored peptide substrate was annealed with 2 pmol of deoxyribozyme 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 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, 150 mM NaCl, and 10 μM freshly activated Imp substrate. The sample was incubated at 37 °C. At appropriate time points, a 2 μL aliquot was quenched with 6 μL of 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.

2.4.7 Free Peptide Assay Procedure

The 20-mer DNA substrate (Oligo1) was 5′-^{32}P-radiolabeled using γ-^{32}P-ATP and T4 polynucleotide kinase. A sample containing 4 nmol of 5′-phosphorylated Oligo1 and 2 pmol of 5′-^{32}P-radiolabeled Oligo1 was brought to 40 μL total volume containing 100 mM EDC and 100 mM imidazole (pH 6.0 with HCl). The sample was incubated at room temperature for 2 h. A Micro Bio-Spin P-6 desalting column (Bio-Rad) was prepared by centrifuging at 1000 *g* for 1 min and

rinsing 4× by adding 500 μL of water followed by centrifuging at 1000 *g* for 1 min. The 40 μL sample was applied to the column and eluted by centrifuging at 1000 *g* for 4 min. The eluant was dried by SpeedVac and redissolved in 8 μL of water.

A 4 μL sample containing 200 pmol of deoxyribozyme and 220 pmol of both helper DNA oligonucleotides used during the identification of the deoxyribozyme was annealed with 175 pmol of AAYAA pentapeptide 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 10 μL total volume containing 70 mM HEPES, pH 7.5, 1.5 mM $ZnCl_2$ (optimized), 20 mM $MnCl_2$, 40 mM $MgCl_2$, 150 mM NaCl, and 25 μM Imp substrate. The sample was incubated at 37 °C for 24 h. Samples were separated by 20% PAGE and quantified with a PhosphorImager.

2.4.8 Mass Spectrometry

The PEGylation product was prepared from a 10 μL sample containing 1.0 nmol of DNA-anchored HEG-tethered tripeptide substrate and 1.2 nmol of deoxyribozyme, which were annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The DNA-catalyzed PEGylation reaction was initiated by bringing the sample to 50 μL total volume containing 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, 150 mM NaCl, and 150 μM Imp-oligo-PEG_{5k}. The sample was incubated at 37 °C for 24 h and separated by 15% PAGE. The product was quantified after gel extraction by UV absorbance (A_{260}), providing 55% overall yield for the PEGylation of the DNA-anchored peptide substrate. The gel-purified sample was analyzed by mass spectrometry on an Applied Biosystems Voyager DE-STR instrument with matrix 3-hydroxypicolinic acid in linear negative ion mode.

2.5 References

(1) Leader, B.; Baca, Q. J.; Golan, D. E. Protein therapeutics: a summary and pharmacological classification. *Nat. Rev. Drug Discov.* **2008**, *7*, 21-39.

(2) Martínez, J. L.; Liu, L.; Petranovic, D.; Nielsen, J. Pharmaceutical protein production by yeast: towards production of human blood proteins by microbial fermentation. *Curr. Opin. Biotechnol.* **2012**, *23*, 965-971.

(3) Turecek, P. L.; Bossard, M. J.; Schoetens, F.; Ivens, I. A. PEGylation of biopharmaceuticals: a review of chemistry and nonclinical safety information of approved drugs. *J. Pharm. Sci.* **2016**, *105*, 460-475.

(4) Dozier, J. K.; Distefano, M. D. Site-specific PEGylation of therapeutic proteins. *Int. J. Mol. Sci.* **2015**, *16*, 25831-25864.

(5) Abuchowski, A.; McCoy, J. R.; Palczuk, N. C.; van Es, T.; Davis, F. F. Effect of covalent attachment of polyethylene glycol on immunogenicity and circulating life of bovine liver catalase. *J. Biol. Chem.* **1977**, *252*, 3582-3586.

(6) Abuchowski, A.; van Es, T.; Palczuk, N. C.; Davis, F. F. Alteration of immunological properties of bovine serum albumin by covalent attachment of polyethylene glycol. *J. Biol. Chem.* **1977**, *252*, 3578-3581.

(7) Pfister, D.; Morbidelli, M. Process for protein PEGylation. *J. Control. Release* **2014**, *180*, 134-149.

(8) Zhu, L.; Kemple, M. D.; Yuan, P.; Prendergast, F. G. N-terminus and lysine side chain pK_a values of melittin in aqueous solutions and micellar dispersions measured by ^{15}N NMR. *Biochemistry* **1995**, *34*, 13196-13202.

(9) Hu, J.; Sebald, W. N-terminal specificity of PEGylation of human bone morphogenetic protein-2 at acidic pH. *Int. J. Pharm.* **2011**, *413*, 140-146.

(10) Kinstler, O.; Molineux, G.; Treuheit, M.; Ladd, D.; Gegg, C. Mono-N-terminal poly(ethylene glycol)-protein conjugates. *Adv. Drug Deliv. Rev.* **2002**, *54*, 477-485.

(11) Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G. Chemical modification of proteins at cysteine: opportunities in chemistry and biology. *Chem. Asian J.* **2009**, *4*, 630-640.

(12) Goodson, R. J.; Katre, N. V. Site-directed PEGylation of recombinant interleukin-2 at its glycosylation site. *Bio/Technology* **1990**, *8*, 343-346.

(13) Basu, A.; Yang, K.; Wang, M.; Liu, S.; Chintala, R.; Palm, T. et al. Structure-function engineering of interferon-β-1b for improving stability, solubility, potency, immunogenicity, and pharmacokinetic properties by site-selective mono-PEGylation. *Bioconjug. Chem.* **2006**, *17*, 618-630.

(14) Kogan, T. P. The synthesis of substituted methoxy-poly(ethyleneglycol) derivatives suitable for selective protein modification. *Synth. Commun.* **1992**, *22*, 2417-2424.

(15) Woghiren, C.; Sharma, B.; Stein, S. Protected thiol-polyethylene glycol: a new activated polymer for reversible protein modification. *Bioconjug. Chem.* **1993**, *4*, 314-318.

(16) Cho, H.; Daniel, T.; Buechler, Y. J.; Litzinger, D. C.; Maio, Z.; Putnam, A. M. H. et al. Optimized clinical performance of growth hormone with an expanded genetic code. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 9060-9065.

(17) Deiters, A.; Cropp, T. A.; Summerer, D.; Mukherji, M.; Schultz, P. G. Site-specific PEGylation of proteins containing unnatural amino acids. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5743-5745.

(18) Wals, K.; Ovaa, H. Unnatural amino acid incorporation in *E. coli*: current and future applications in the design of therapeutic proteins. *Front. Chem.* **2014**, *2*, 15.

(19) Popp, M. W.; Dougan, S. K.; Chuang, T. Y.; Spooner, E.; Ploegh, H. L. Sortase-catalyzed transformations that improve the properties of cytokines. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 3169-3174.

(20) DeFrees, S.; Wang, Z. G.; Xing, R.; Scott, A. E.; Wang, J.; Zopf, D. et al. GlycoPEGylation of recombinant therapeutic proteins produced in *Escherichia coli*. *Glycobiology* **2006**, *16*, 833-843.

(21) Mero, A.; Schiavon, M.; Veronese, F. M.; Pasut, G. A new method to increase selectivity of transglutaminase mediated PEGylation of salmon calcitonin and human growth hormone. *J. Control. Release* **2011**, *154*, 27-34.

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

(23) Walsh, S. M. Kinase Deoxyribozymes. Ph.D. Thesis, University of Illinois at Urbana-Champaign, **2017**.

(24) 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.

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

(27) Vosko, A. M.; Schroeder, A.; Loh, D. H.; Colwell, C. S. Vasoactive intestinal peptide and the mammalian circadian system. *Gen. Comp. Endocrinol.* **2007**, *152*, 165-175.

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

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

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

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

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

Chapter 3: DNA Enzymes with Tyrosine Azido-Adenylylation Activity on Peptide Substrates[†]

3.1 Introduction

3.1.1 Chemical Methods Targeting Tyrosine for Peptide and Protein Modification

Synthesis of modified peptides and proteins is essential for understanding the functions of natural post-translational modifications, monitoring protein distributions in vivo, modulating therapeutic properties of peptides, proteins, antibodies, and antibody-drug conjugates, and other applications.[1,2] In nature, tyrosine is modified via a wide array of chemical changes.[3] Tyrosine has often been targeted for synthetic modification on the basis of its electron-rich nature and typically low abundance on protein surfaces. These chemical approaches are powerful tools for the synthesis of modified proteins, even though they each rely solely on differential physical accessibilities to discriminate among several tyrosines in the substrate. Some of these chemical methods also suffer from off-target reactivity of other electron-rich residues such as tryptophan.

3.1.1.1 Transition Metal-Free Approaches

Several transition-metal-free strategies have emerged for modifying tyrosine residues in protein substrates. Electron-deficient diazonium salts have been shown to react with tyrosine ortho to the phenol group (Figure 3.1A).[4,5] The introduction of ketones, aldehydes, and more recently alkynes[6] at the para position in the diazonium salts provides an additional biorthogonal handle for a second-step modification, for example, with fluorophore reagents.

[†] This research has been published:

Wang, P.; Silverman, S. K. DNA-Catalyzed Introduction of Azide at Tyrosine for Peptide Modification. *Angew. Chem. Int. Ed.* **2016**, *55*, 10052-10056.

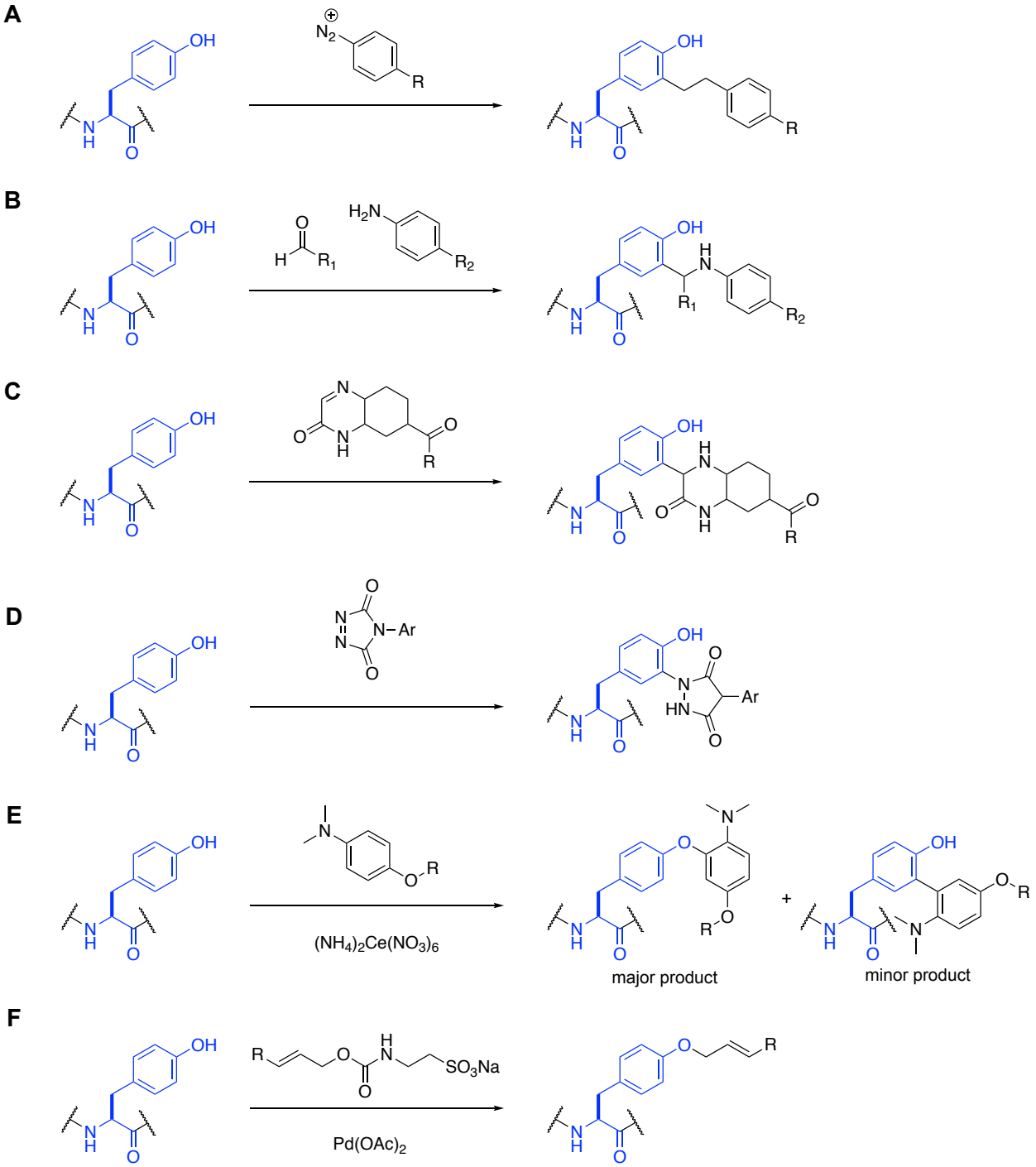


Figure 3.1. Chemical methods targeting tyrosine for peptide and protein modification. (A) Reaction with diazonium salts.[4,5] (B) Three-component Mannich reaction.[7] (C) Reaction with preformed imines.[8] (D) Ene-type reaction with diazodicarboxylate reagents.[9] (E) Ce(IV)-mediated oxidative coupling reaction.[10] (F) Pd-catalyzed allylic *O*-alkylation.[11]

Another approach developed for tyrosine modification is a three-component Mannich-type reaction with aldehydes and anilines (Figure 3.1B).[7] Solvent-exposed tyrosine residues found in lysozyme, RNase A, and chymotrypsinogen A can be successfully modified via this transformation. However, this approach has off-target activity with accessible tryptophan residues and reduced disulfides.

Other methods for tyrosine modification include reaction with preformed imines (Figure 3.1C)[8] and aqueous ene-type reaction with azomaleimides (Figure 3.1D).[9] The latter approach has been utilized for the preparation of well-defined glycoconjugate vaccines[12] and DNA-protein conjugates in model streptavidin and myoglobin proteins.[13]

3.1.1.2 Transition Metal-Mediated Approaches

Oxidative couplings of tyrosine residues mediated by Ce(IV) have been explored (Figure 3.1E).[10] *N,N*-Dialkylated aniline bearing an additional electron-donating group in the para position was reported as the most reactive. With optimizations of incubation period and reagent concentrations, this strategy can be compatible with a variety of substrates.

Tyrosine can also be modified by palladium-catalyzed Tsuji-Trost reaction (Figure 3.1F).[11] Tyrosine *O*-alkylation with a rhodamine dye or lipophilic moieties is achieved by using π-allylpalladium complexes that are generated from allyl acetates/carbamates and $Pd(OAc)_2$/triphenylphosphine tris(sulfonate) (TPPTS) as a catalytic system.

3.1.2 Enzymatic Methods Targeting Tyrosine for Peptide and Protein Modification

Natural protein enzymes have been harnessed for tyrosine modification in a few cases. For example, an enzymatic cascade utilizes the tandem reaction comprising a fungal tyrosinase-catalyzed hydroxylation of tyrosine to L-DOPA residue, followed by *O*-alkylation catalyzed by catechol-*O*-methyltransferase (COMT) (Figure 3.2A).[14] However, this strategy cannot be used to modify tyrosine residues present within α-helices.

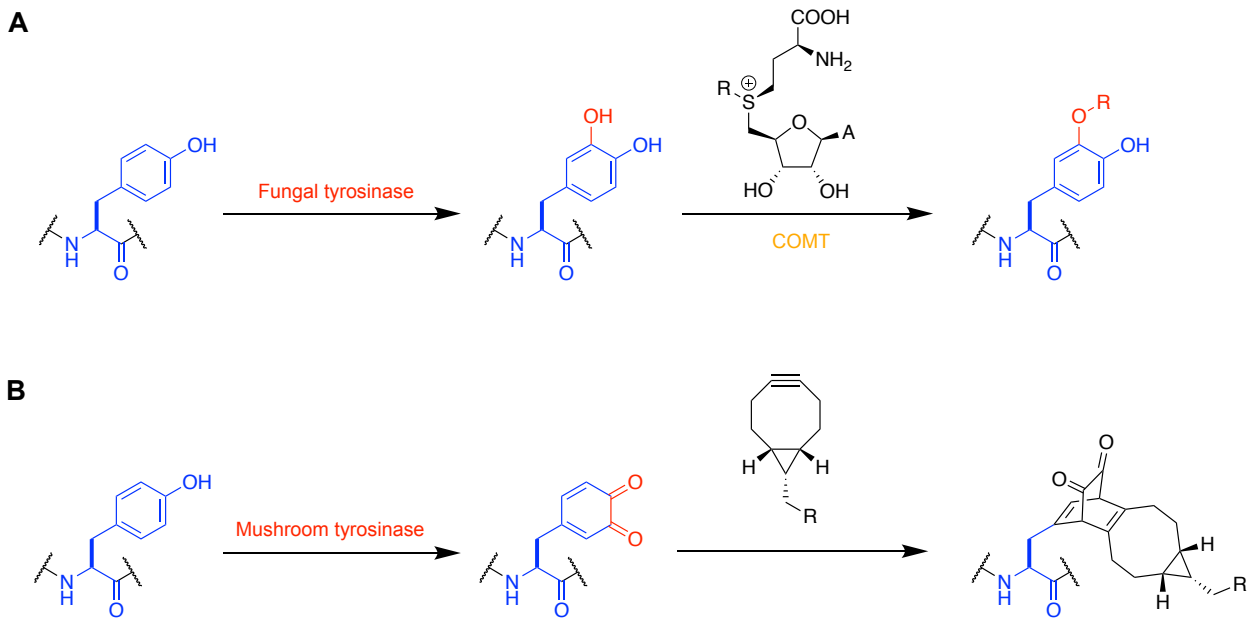


Figure 3.2. Enzymatic methods targeting tyrosine for peptide and protein modification. (A) Tandem enzymatic reaction with a fungal tyrosinase and the mammalian catechol *O*-methyltransferase (COMT).[14] (B) Tyrosine oxidation followed by strain-promoted (4 + 2) cycloaddition.[15]

In another example, a mushroom tyrosinase is used for in situ generation of 1,2-quinone from tyrosine, which is followed by strain-promoted quinone-alkyne cycloaddition (SPOCQ) with bicyclo[6.1.0]nonyne (BCN) (Figure 3.2B).[15] This method was applied for the synthesis of an antibody-drug conjugate where BCN-monomethyl auristatin F conjugate was coupled with AT1002, a potent anti-influenza antibody, at the genetically engineered C-terminal tyrosine residue via SPOCQ.

As discussed above, these enzymatic methods targeting tyrosine for peptide and protein modification rely on the substrate scope of the natural enzymes for site selectivity among different tyrosines. However, in some cases, the site selectivity (if any) provided by the natural enzymes is not satisfactory for particular purposes. Therefore, DNA enzymes with tyrosine modification activity will be valuable alternative tools, especially considering that DNA enzymes with site selectivity can be de novo identified without the requirement of a known enzyme as the starting point.

3.2 Results and Discussion

3.2.1 In Vitro Selection Strategy

To achieve DNA-catalyzed peptide and protein modification, a two-step approach was pursued (Figure 3.3A). In this approach, a deoxyribozyme first catalyzes the transfer of the 2′-azido-2′-deoxyadenosine 5′-monophosphoryl group (2′-Az-dAMP) from the analogous 5′-triphosphate (2′-Az-dATP) onto the tyrosine hydroxyl group (azido-adenylylation). Second, a particular modification of interest is attached to the azido group by copper-catalyzed azide-alkyne cycloaddition (CuAAC)[16,17] using an alkyne-functionalized reagent.

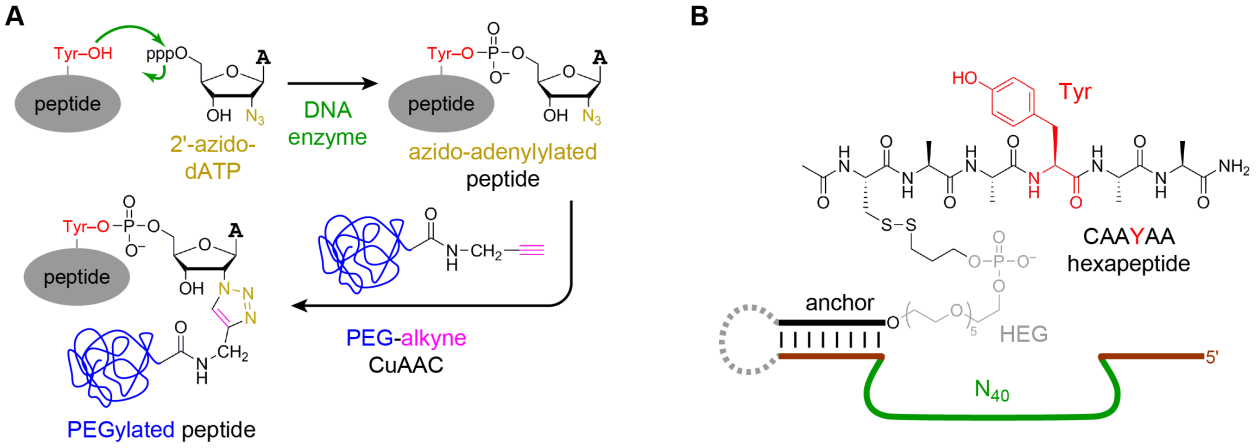


Figure 3.3. Overview of DNA-catalyzed introduction of azide at tyrosine residues. (A) Two-step peptide modification by DNA-catalyzed azido-adenylylation and CuAAC, with PEGylation as the specific example. (B) Arrangement of initially random DNA pool (N_{40}) and peptide substrate for in vitro selection of deoxyribozymes for azido-adenylylation. The dashed loop enables selection but is dispensable for catalysis. Figure adapted with permission from ref. 19.

In vitro selection experiment was initially performed to establish viability of this approach. The hexapeptide CAAYAA prepared by solid-phase peptide synthesis was used as the substrate (Figure 3.3B). This hexapeptide was covalently attached via a disulfide linkage and a hexa(ethylene glycol) [HEG] tether to a DNA anchor oligonucleotide, which was bound by Watson-Crick base pairs to one of the fixed-sequence binding arms of the random N_{40} DNA pool. The DNA-anchored peptide substrate was ligated to the initially random DNA pool using

T4 DNA ligase. The initially random region of the DNA sequences was flanked by fixed-sequence regions to enable PCR amplification of the DNA sequences and substrate binding.

The in vitro selection process began with ligation of the initially random DNA pool to the DNA-anchored peptide substrate (Figure 3.4). The ligation product was isolated by polyacrylamide gel electrophoresis (PAGE). Next, the selection step was performed in which the DNA sequences were incubated with 2′-Az-dATP and divalent metal ions in a buffered solution to enable DNA catalysis. After exposure of the substrate-conjugated N_{40} pool to 2′-Az-dATP, the catalytically active DNA sequences, which were now attached to an azido group, were increased in mass by CuAAC with alkyne-modified poly(ethylene glycol) that has an average molecular weight of 5000 (PEG_{5k}-alkyne). For catalytically active DNA sequences, this "capture step"[18] attachment of PEG induces a sufficiently large PAGE shift that enables the selection process. At the outset of these selection experiments, we did not have any way to prepare an exact standard with an azido-adenylylated peptide moiety. Therefore, as the capture control for both the CuAAC reaction and the PAGE migration, we used an azido-modified substrate that was prepared by conjugating a 3′-C_3-NH_2 oligonucleotide with 6-azidohexanoic acid (Figure 3.5).

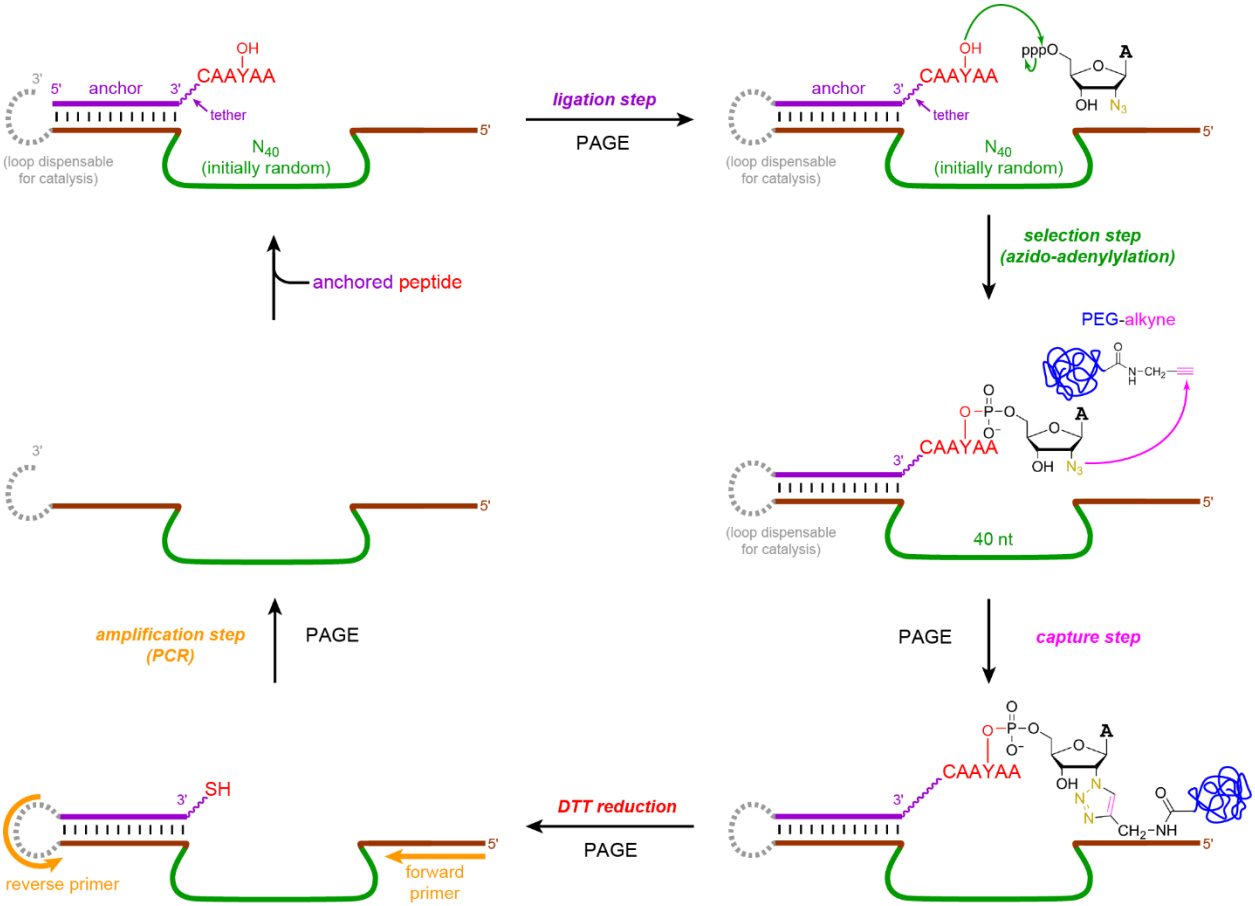


Figure 3.4. Diagram of in vitro selection procedure for deoxyribozymes with tyrosine azido-adenylylation activity.

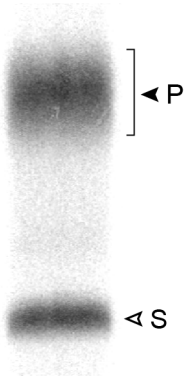


Figure 3.5. PAGE image of capture control reaction using PEG_{5k}-alkyne (S = substrate, P = product). Figure adapted with permission from ref. 19.

After PAGE separation, treatment with dithiothreitol (DTT) cleaved the disulfide bond, and another PAGE separation was performed. As discussed in Chapter 2, this additional

DTT/PAGE operation proved necessary to avoid emergence of DNA sequences that catalyze undesired reactions. If the PEG-modified 2′-Az-dAMP group is attached not at the tyrosine residue but instead at some undesired position within the DNA itself (for example, the 5′-hydroxyl group or a nucleobase functional group), then disulfide cleavage will lead to only a small PAGE shift, in comparison with the much larger PAGE shift upon removal of the entire PEG-modified hexapeptide. Therefore, the difference in PAGE shift enables the separation of DNA enzymes with azido-adenylylation activity from sequences that catalyze side reactions. PCR was performed to amplify the catalytically active DNA sequences that survived the previous selection, capture, and disulfide cleavage steps. The entire selection process was iterated, each time enriching the population with catalytically active sequences, until the DNA enzymes dominated the population.

3.2.2 Identification and Characterization of DzAz1 Deoxyribozyme Family

The initial selection experiment designated PP1 according to our systematic alphanumeric nomenclature was performed. In each selection round, the incubation conditions included 100 μM 2′-Az-dATP in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h. After 8 rounds, the azido-adenylylation activity of the pool reached 16% (Figure 3.6A), and individual deoxyribozymes were cloned.[19] A single sequence, named DzAz1 (deoxyribozyme for azido-adenylylation 1, Figure 3.7), was identified.

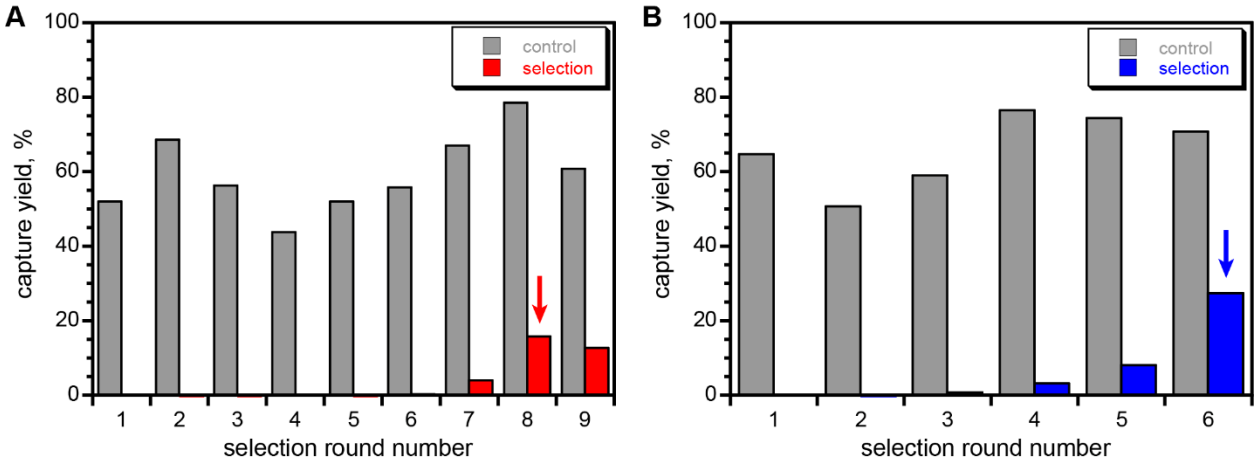


Figure 3.6. Progression of the selection experiments PP1 (A) and RV1 (B). In each round, "control" refers to the yield for the control capture reaction using an azido-modified substrate, and "selection" refers to the yield for the CuAAC capture reaction for the indicated selection experiment. Arrows mark the cloned rounds. Figure adapted with permission from ref. 19.

```
                1         10        20        30        40
                |         |         |         |         |
DzAz1   AACCACCTTT GTATAGTTGG GGGGCGGGCC ACCGTGACAC 40
DzAz1b  .C..TT.ACA .C.A.CAA.. .......... .........C. 40
DzAz1c  ..G....AA. .CGAT.GAC. .......... .......... 40
```

Figure 3.7. Sequences of the initially random region of the DzAz1 deoxyribozyme and reselected variants DzAz1b and DzAz1c. Dots denote the same nucleotide as in the uppermost sequence of the listing. Figure adapted with permission from ref. 19.

Reselection of the DzAz1 deoxyribozyme was performed in the RV1 selection experiment by generating a partially randomized pool (25% randomization). The fixed-sequence binding arms flanking the initially random region were changed to prevent contamination of the reselection experiment by the parent deoxyribozyme (Table 3.2). The reselection was cloned from round 6 with 27% activity (Figure 3.6B). Two reselected variants DzAz1b and DzAz1c were identified (Figure 3.7). The sequence alignment of DzAz1, DzAz1b, and DzAz1c revealed that the 5′-half of the N_{40} sequence was highly conserved.

With the DNA-anchored CAAYAA substrate, the reselected variant DzAz1b has 58% single-turnover azido-adenylylation yield with k_{obs} of 0.38 ± 0.03 h^{-1}, and variant DzAz1c has 56% yield with k_{obs} of 0.19 ± 0.03 h^{-1}; parent DzAz1 has 40% yield with k_{obs} of 0.79 ± 0.07 h^{-1}

(each n = 3, mean ± sd; Figure 3.8). Product identity was validated by MALDI mass spectrometry (Table 3.3). Both Zn^{2+} and one of Mn^{2+} or Mg^{2+} are required for catalysis; Mn^{2+} is much more effective (Figure 3.9). All three deoxyribozymes in the DzAz1 family retain substantial activity when natural ATP is used in place of 2′-Az-dATP (Figure 3.8), suggesting utility of this general approach for studying natural tyrosine adenylylation (AMPylation).[3,20,21]

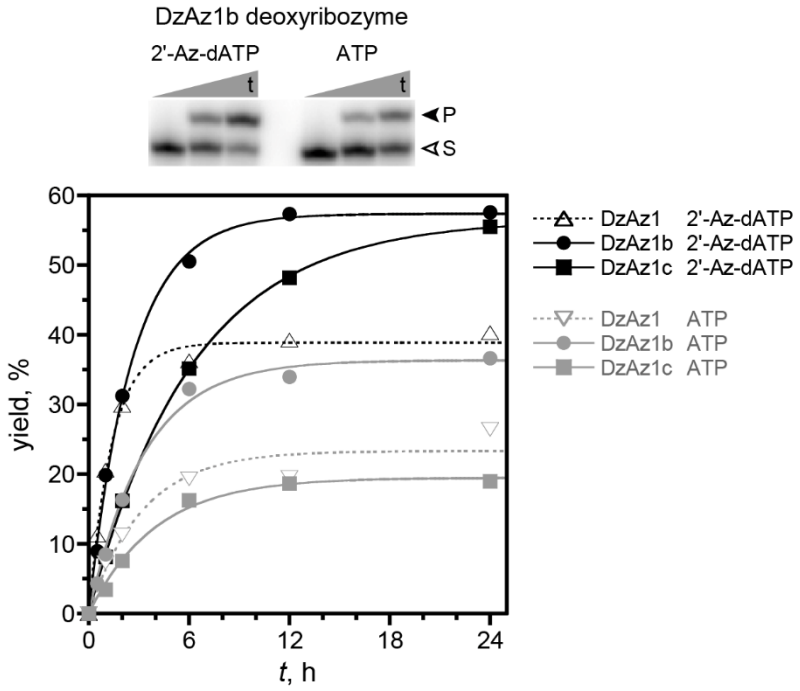


Figure 3.8. Single-turnover kinetic assays of the DzAz1 deoxyribozyme and its variants DzAz1b and DzAz1c. Incubation conditions: 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C with 100 μM 2′-Az-dATP or natural ATP. In the PAGE image, representative time points are shown for DzAz1b at t = 0, 2, and 24 h. S = substrate, P = product. Figure adapted with permission from ref. 19.

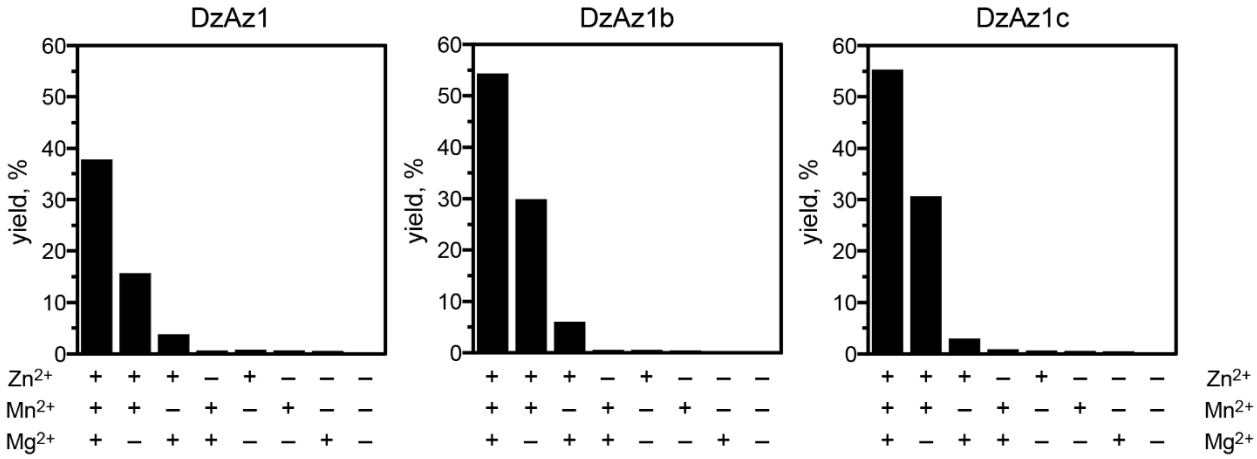


Figure 3.9. Assays of metal ion dependence of the DzAz1 deoxyribozyme and its DzAz1b and DzAz1c sequence variants. Each assay used the DNA-anchored CAAYAA peptide substrate. Incubation conditions: 70 mM HEPES, pH 7.5, combinations of 1 mM $ZnCl_2$, 20 mM $MnCl_2$, and 40 mM $MgCl_2$ as appropriate, and 150 mM NaCl at 37 °C with 100 μM 2′-Az-dATP. The yield at 24 h for each metal ion combination is shown. Figure adapted with permission from ref. 19.

The peptide sequence requirement of DzAz1 family deoxyribozymes was studied with four DNA-anchored octapeptides CLQTYPRT, CQQPYITN, CERSYLMK, and CFQPYMQE (Figure 3.10). The first peptide sequence comprises amino acids 19–25 from the 32-mer salmon calcitonin (sCT),[22] and the remaining three sequences correspond respectively to amino acids 15–21, 54–60 (with C56S), and 78–84 of human interleukin-22 (146-mer, hIL-22, amino acids 34–179 of the genomic sequence);[23] an artificial N-terminal cysteine was appended onto each peptide to allow disulfide linkage to the DNA anchor oligonucleotide. sCT is a hormone prescribed for bone-related disorders such as osteoporosis, Paget's disease, and hypercalcemia; hIL-22 is a cytokine involved in proliferation, host defense, and the inflammatory response. All three deoxyribozyme show partial discrimination among the four peptide sequences.

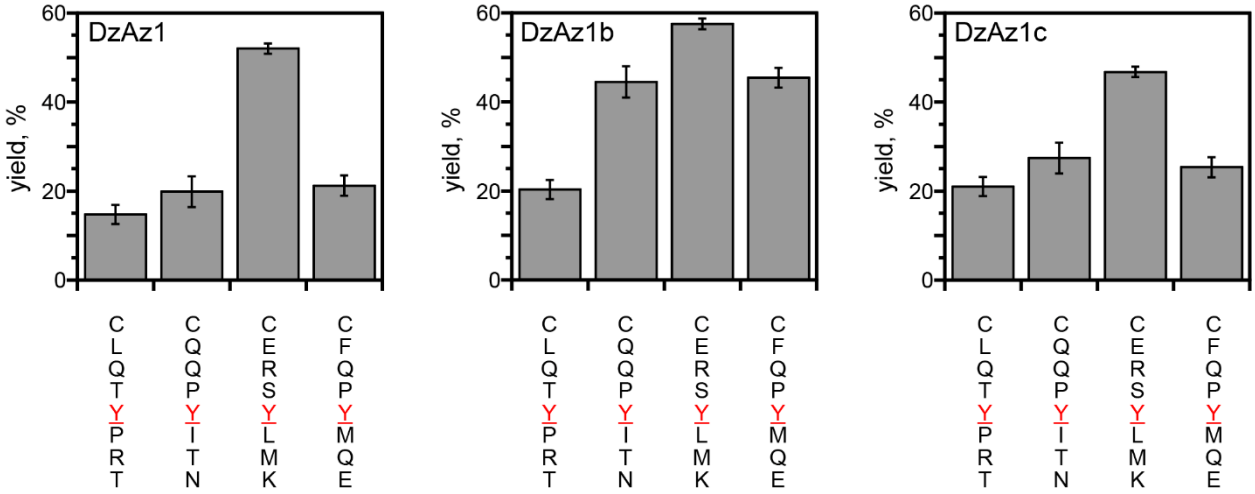


Figure 3.10. Activities of DzAz1 family deoxyribozymes with the four DNA-anchored peptide substrates (yield at 24 h; $n = 3$, mean ± sd). Figure adapted with permission from ref. 19.

3.2.3 Identification and Characterization of Peptide Sequence-Selective Deoxyribozymes

In two previous efforts with tyrosine-modifying deoxyribozymes, peptide sequence-selective catalysis by DNA was found from selection experiments in which mixed-sequence peptides were the substrates.[24,25] Therefore, four new in vitro selection experiments, RW1–RZ1, were performed, each using one of the four respective DNA-anchored peptides CLQTYPRT, CQQPYITN, CERSYLMK, and CFQPYMQE. Each selection experiment was performed using the same incubation conditions as for the PP1 selection experiment. The fixed-sequence binding arms flanking the initially random region were varied between the selection experiments (Table 3.2). After 8–12 rounds, each selection experiment showed 16–44% pool azido-adenylylation activity (Figure 3.11), and individual deoxyribozymes were cloned.

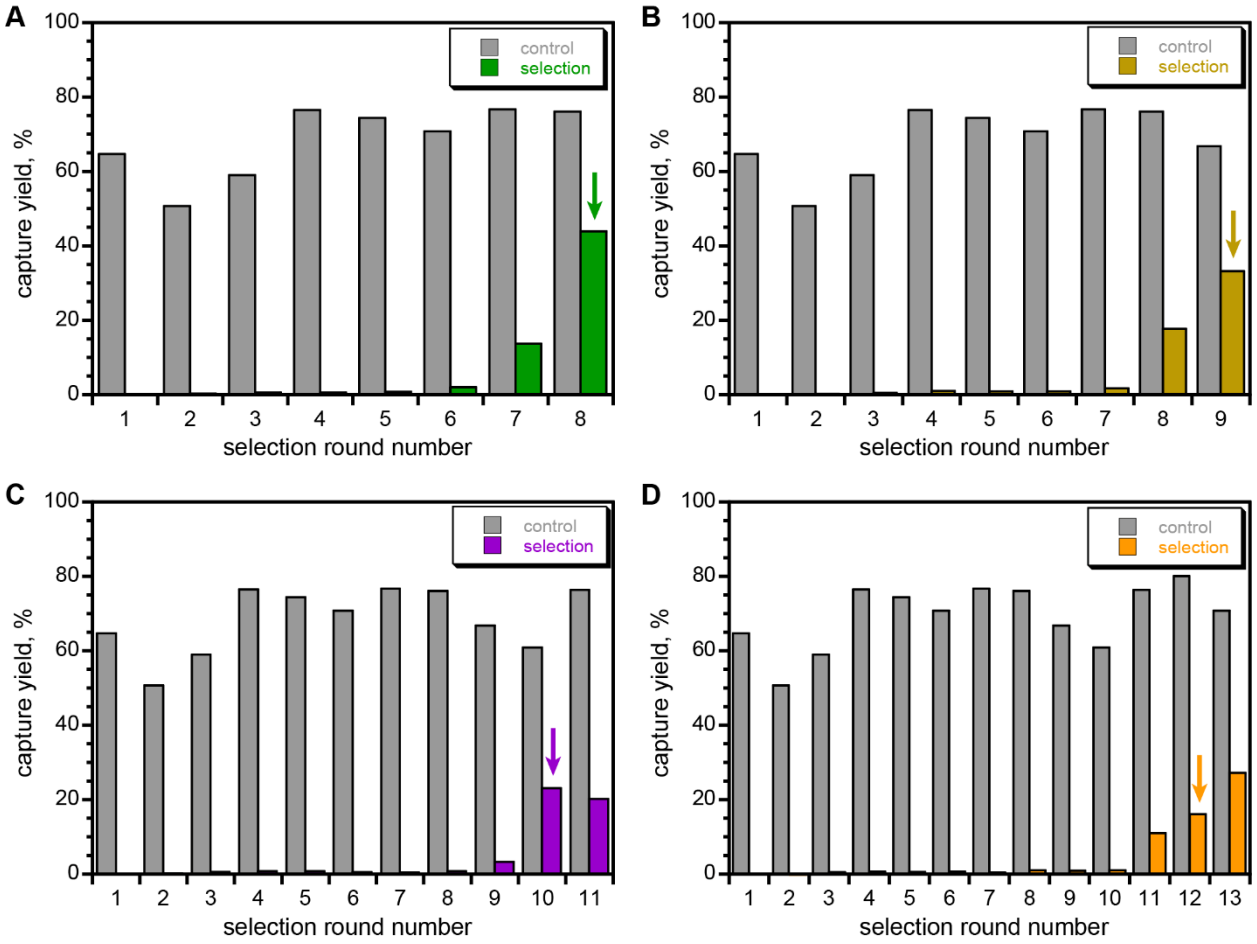


Figure 3.11. Progression of the selection experiments RW1 (A), RX1 (B), RY1 (C), and RZ1 (D). In each round, "control" refers to the yield for the control capture reaction using an azido-modified substrate, and "selection" refers to the yield for the CuAAC capture reaction for the indicated selection experiment. Arrows mark the cloned rounds. Figure adapted with permission from ref. 19.

Five unique deoxyribozymes sequences designated DzAz2 through DzAz6 were identified from the RW1 selection experiment with the sCT-derived octapeptide CLQTYPRT, while the other three selections, RX1, RY1, and RZ1, with the hIL-22 peptides each led to a single DNA sequence, respectively named DzAz7, DzAz8, and DzAz9 (Figure 3.12). These eight new DNA enzymes share no obvious sequence conservation among themselves or with DzAz1 family enzymes.

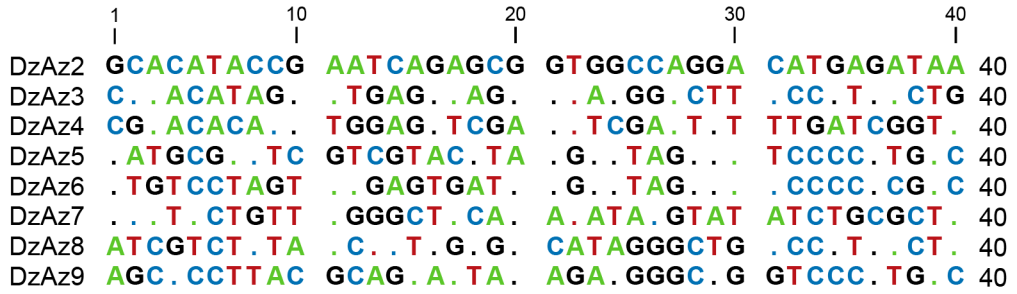


Figure 3.12. Sequences of the initially random region of DzAz2–DzAz9 deoxyribozymes. Dots denote the same nucleotide as in the uppermost sequence of the listing. Figure adapted with permission from ref. 19.

Among these eight deoxyribozymes, DzAz2, DzAz7, and DzAz8 catalyze azido-adenylylation of their corresponding peptide substrates in 61–87% yield in 24 h with k_{obs} of 0.1–0.5 h^{-1} (Figure 3.13A); the other deoxyribozymes have lower 9–27% yields. Natural ATP is generally tolerated well in place of 2′-Az-dATP (Figure 3.13B). Product identity was validated by MALDI mass spectrometry for DzAz2, DzAz7, and DzAz8 (Table 3.3)

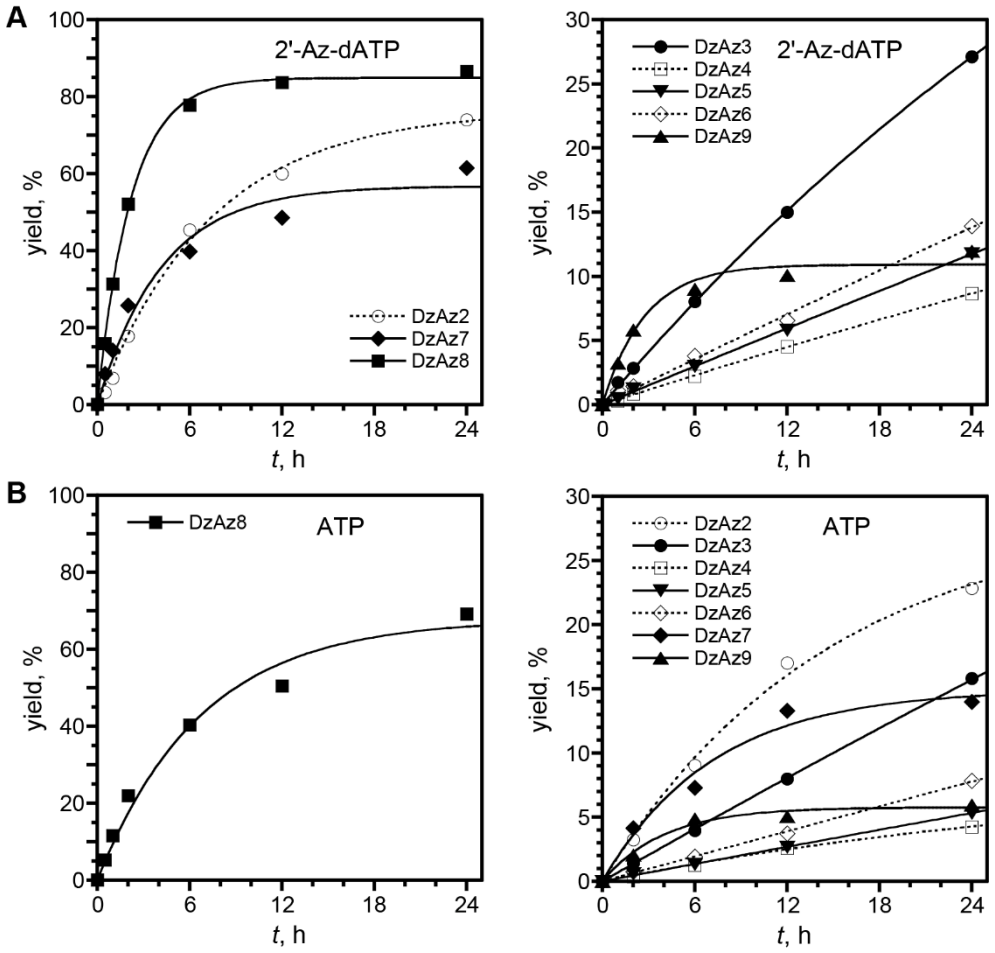


Figure 3.13. Single-turnover kinetic assays of DzAz2–DzAz9 deoxyribozymes. Incubation conditions: 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C with 100 μM 2′-Az-dATP (A) or natural ATP (B). Figure adapted with permission from ref. 19.

All eight deoxyribozymes were evaluated for azido-adenylylation activity with each of the four mixed-sequenced DNA-anchored peptides (Figure 3.14). DzAz8 does not discriminate among any of the peptide sequences. In sharp contrast, DzAz2 functions well only with the CLQTYPRT substrate that was used during its identification. The other six deoxyribozymes each exhibit partial discrimination among the four peptide sequences, which is similar to the behavior observed with DzAz1 family enzymes.

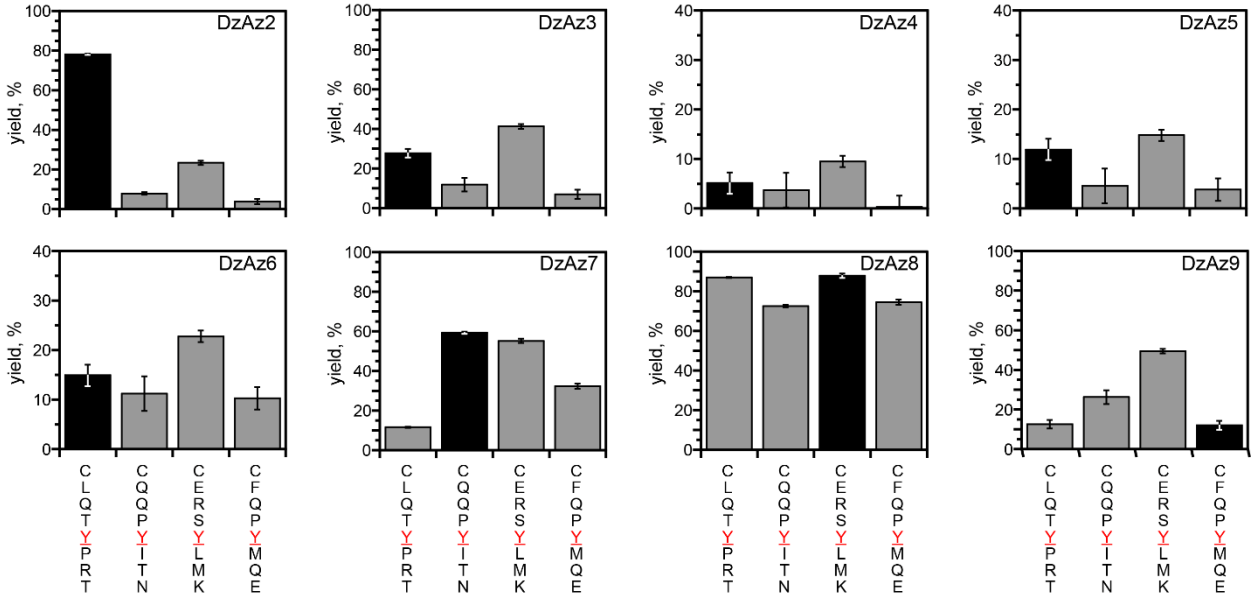


Figure 3.14. Activities of DzAz2–DzAz9 with the four DNA-anchored peptide substrates (yield at 24 h; $n = 3$, mean ± sd). Black bars denote the particular peptide substrate used during identification of that deoxyribozyme; grey bars are for the other peptide. Figure adapted with permission from ref. 19.

The peptide sequence dependence of DzAz2 was analyzed further by testing its azido-adenylylation activity with systematic mutants of the DNA-anchored CLQTYPRT substrate. As expected, mutation of the tyrosine residue to phenylalanine or serine abolished activity (Figure 3.15A). The six amino acids surrounding the tyrosine residue were then individually replaced with alanine, revealing that the required motif is YPR (Figure 3.15B). Introduction of this YPR motif into the other three peptide substrates enabled their azido-adenylylation by DzAz2 (Figure 3.15C), establishing YPR as both necessary and sufficient for DzAz2 reactivity. In contrast, although the activity of the non-selective DzAz8 deoxyribozyme was also abolished upon mutating the tyrosine residue to phenylalanine or serine, all alanine mutations of CLQTYPRT were tolerated well (Figure 3.16).

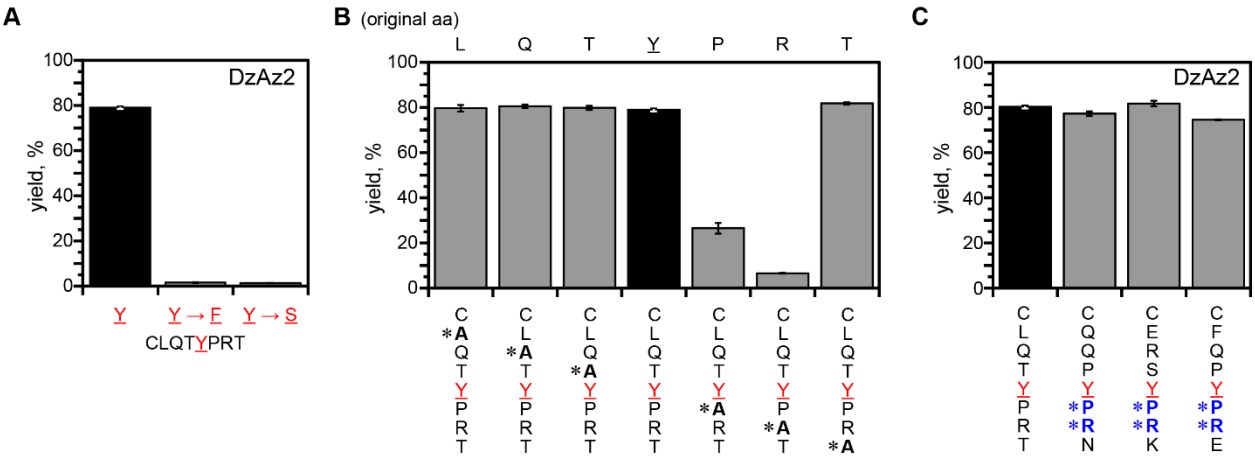


Figure 3.15. Peptide sequence requirement of DzAz2 (yield at 24 h; $n = 3$, mean ± sd). (A) Activity of DzAz2 upon mutation of Tyr to Phe or Ser in the DNA-anchored peptide substrate. Either mutation abolished activity. (B) Azido-adenylylation activity of DzAz2 with a series of DNA-anchored peptide substrates in which a single amino acid was mutated to Ala. (C) DzAz2 activity requires only the YPR motif in the peptide substrate. Each of the three alternative DNA-anchored peptide sequences was mutated to introduce the YPR motif while retaining all other existing amino acids. In all three cases, full activity was observed. Figure adapted with permission from ref. 19.

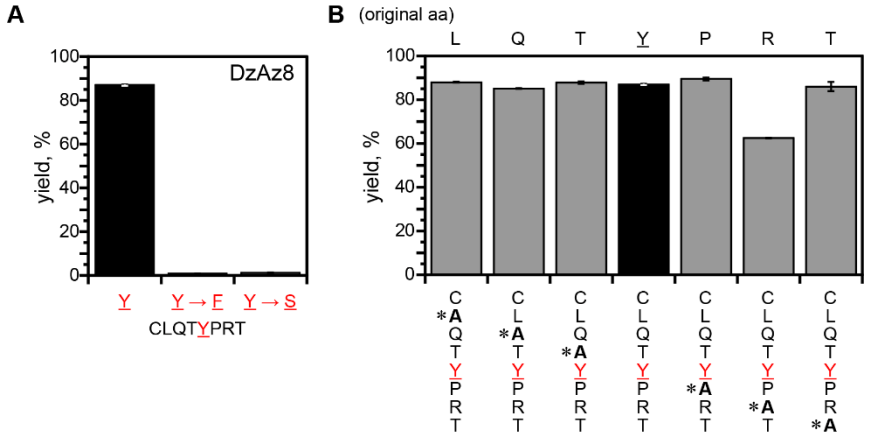


Figure 3.16. Peptide sequence tolerance of DzAz8 (yield at 24 h; $n = 3$, mean ± sd). (A) Activity of DzAz8 upon mutation of Tyr to Phe or Ser in the DNA-anchored peptide substrate. Either mutation abolished activity. (B) Azido-adenylylation activity of DzAz8 with a series of DNA-anchored peptide substrates in which a single amino acid was mutated to Ala. Figure adapted with permission from ref. 19.

The ability of DzAz2 to differentiate between two tyrosine residues within one peptide substrate was also evaluated. The LQTYPRT (site 1) and QQPYITN (site 2) sequence motifs were concatenated in either order within a longer peptide substrate via an arbitrary tripepide

linker (ASK). The longer DNA-anchored peptide substrate was azido-adenylylated by DzAz2, and the product (as a mixture of site 1 and site 2 modifications) was PAGE-separated. The discrimination between the two sites was revealed by MALDI mass spectrometry (Figure 3.17), after Lys-C cleavage to separate the two peptide segments and DTT treatment to remove the DNA anchor oligonucleotide. Peak intensity ratios from MALDI mass spectrometry were used to calculate the discrimination factors, assuming that the response factor for each peptide fragment does not change upon azido-adenylylation. With the peptide substrate CLQTYPRTASKQQPYITN (peptide A), the two discrimination factors were 3.6 and 130; with the peptide substrate CQQPYITNASKLQTYPRT (peptide B), the two discrimination factors were 438 and 109. In each case, DzAz2 preferentially modified LQTYPRT (site 1), which suggested that DzAz2 is able to discriminate between two tyrosine residues within one longer peptide substrate on the basis of their sequence contexts.

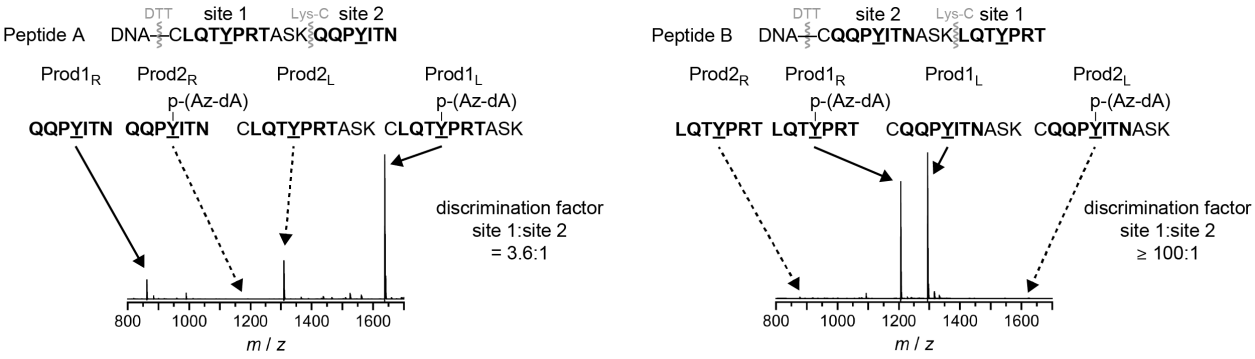


Figure 3.17. DzAz2 differentiates between two Tyr in one peptide substrate. Each peptide product fragment is labelled Prod1 or Prod2 corresponding to site 1 or site 2 modification, and with a subscripted L or R corresponding to the fragment location on the left or right side of the longer peptide before Lys-C cleavage. From each mass spectrum, peak intensity ratios were used to calculate two discrimination factors, one as $Prod1_L/Prod2_L$ and the other as $Prod1_R/Prod2_R$, assuming that the response factor for each peptide fragment does not change upon azido-adenylylation. For peptide A, $Prod1_L/Prod2_L$ was 3.6, and $Prod1_R/Prod2_R$ was 130; the smaller value is reported. For peptide B, $Prod1_L/Prod2_L$ was 438, and $Prod1_R/Prod2_R$ was 109. Figure adapted with permission from ref. 19.

3.2.4 Azido-Adenylylation of Untethered Peptide Substrates

For the eventual goal of peptide and protein modification, deoxyribozymes such as DzAz8 will be most useful if they can function with free peptide substrates that are not tethered to a DNA anchor oligonucleotide. Azido-adenylylation of untethered sCT was achieved by DzAz8 in 14% yield in 24 h, as revealed by HPLC (Figure 3.18A). In contrast, the sCT-specific DzAz2 did not function with untethered sCT. DzAz8 was also assayed with the 28-mer N-terminal fragment of atrial natriuretic peptide (atriopeptin, ANP), which can induce natriuresis (sodium excretion in urine) and vasodilation and is commercialized in Japan for treatment of heart failure.[26] Azido-adenylylation of ANP by DzAz8 in 59% yield in 24 h was observed (Figure 3.18A). Both sCT and ANP were also adenylylated with natural ATP lacking the 2′-azido group, in respective yields of 6.7% and 16% (Figure 3.18B).

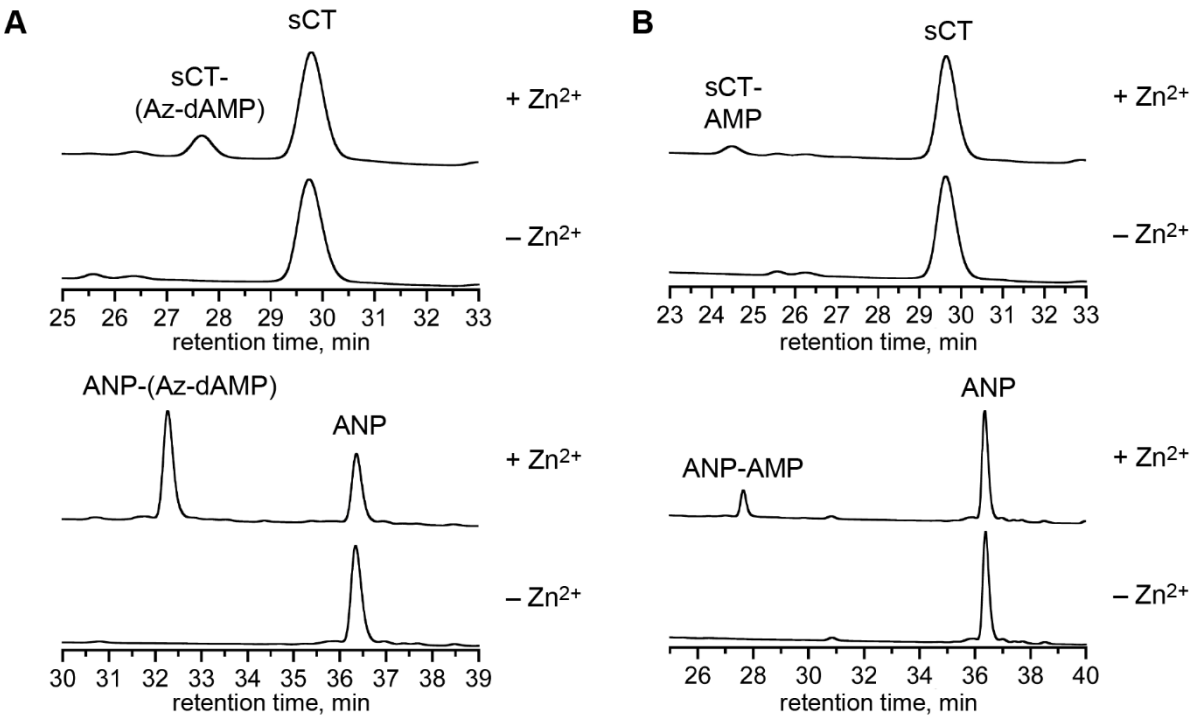


Figure 3.18. Azido-adenylylation (A) and adenylylation (B) by DzAz8 of sCT and ANP, assayed by HPLC (t = 24 h). Figure adapted with permission from ref. 19.

The HPLC-purified azido-adenylylated sCT and ANP were further derivatized quantitatively via CuAAC using either PEG_{5k}-alkyne or fluorescein-alkyne (Figure 3.19). The fluorescein-modified sCT and ANP were also validated by MALDI mass spectrometry (Table 3.3).

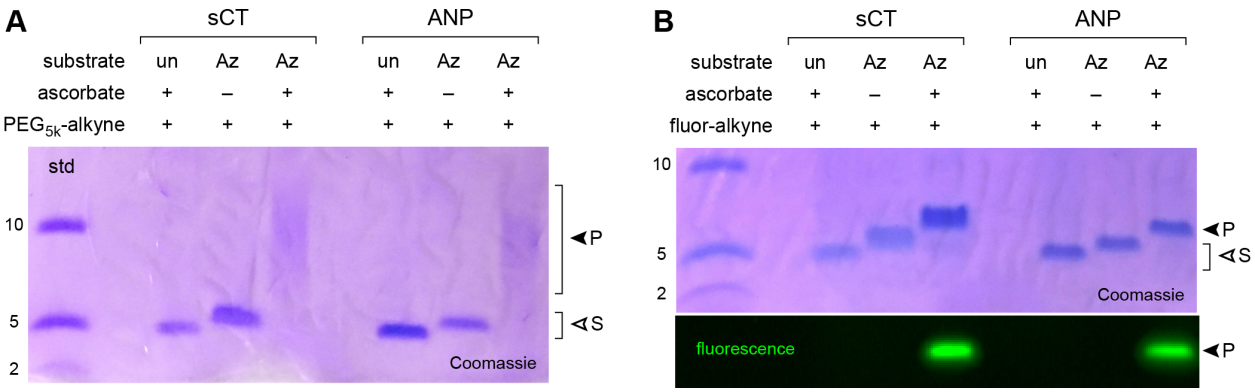


Figure 3.19. Modification via CuAAC of azido-adenylylated sCT and ANP with PEG_{5k}-alkyne (A) and fluorescein-alkyne (B), assayed by SDS-PAGE and imaging by Coomassie stain or fluorescence. un = unmodified substrate; Az = azido-adenylylated substrate. S = substrate; P = product. std = 2, 5, 10 kDa. Ascorbate is the reducing agent required for CuAAC. Figure adapted with permission from ref. 19.

3.2.5 Comparison between PEGylation Deoxyribozymes and Azido-Adenylylation Deoxyribozymes

The properties of PEGylation deoxyribozymes discussed in Chapter 2 and azido-adenylylation deoxyribozymes are compared in Table 3.1. In terms of the substrate scope, the PEGylation deoxyribozymes are only active with tethered peptide substrates. None of the PEGylation DNA enzymes work with untethered peptide substrates, and none were further tested with protein substrates. The azido-adenylylation deoxyribozymes are functional with both tethered and free peptide substrates. With tethered peptide substrates, some of the azido-adenylylation DNA enzymes show peptide sequence selectivity. One such deoxyribozyme, DzAz8, is able to catalyze modification of free peptide substrates. However, no azido-adenylylation product was observed when DzAz8 was tested with free protein substrates (data

not shown). Therefore, one of the future directions is to identify deoxyribozymes that can catalyze azido-adenylylation of protein substrates.

Table 3.1: Comparison between PEGylation deoxyribozymes and azido-adenylylation deoxyribozymes. In the columns of substrate scope, "yes" or "no" represents whether the activity on the corresponding substrate is observed; "n/a" represents that the substrate was not assayed.

deoxyribozyme type	substrate scope			reaction workflow	
	tethered peptide	untethered peptide	protein	step(s) to final product	reagent availability
PEGylation	yes	no	n/a	1	synthesis needed
azido-adenylylation	yes	yes	no	2	commercially available

In terms of the workflow for the preparation of modified peptides and proteins, PEGylation deoxyribozymes provides a one-step synthetic route to final products, whereas azido-adenylylation deoxyribozymes need a subsequent CuAAC step to incorporate the modification of interest. As discussed in Chapter 2, the PEGylation deoxyribozymes can be used to introduce any particular modification of interest by using the corresponding phosphorimidazolide-activated substrate where the PEG moiety used for PEGylation is replaced with the modification of interest, although some synthetic effort might be required to prepare the activated substrate. For the application of azido-adenylylation deoxyribozymes, 2′-Az-dATP can be directly purchased, and various kinds of alkyne-functionalized reagents are commercially available.

3.3 Summary and Future Directions

Tyrosine has been explored as the target for the synthesis of modified peptides and proteins. Both chemical and enzymatic strategies for tyrosine modification have been developed. However, these approaches lack site selectivity among different tyrosines in the substrate. Some

tyrosine derivatization reactions suffer from off-target reactivity of other electron-rich residues such as tryptophan. The issues of poor site selectivity and chemoselectivity can be addressed by de novo identification of deoxyribozymes.

DNA enzymes were identified for azido-adenylylation of tyrosine residues in peptide substrates, which provides the azido group for subsequent derivatization via CuAAC. The first deoxyribozyme, DzAz1, and its reselected mutants validated the in vitro selection strategy. Eight different deoxyribozymes were identified from in vitro selection experiments that used mixed-sequence peptide substrates. One of the DNA enzymes, DzAz2, is selective for the YPR sequence motif and is able to discriminate between tyrosine residues within a single peptide on the basis of sequence context. Another deoxyribozyme, DzAz8, is peptide sequence-general, functions with free peptides, and allows their subsequent CuAAC labeling with moieties such as PEG and fluorescein.

These results reveal DNA enzymes as potential catalysts for the synthesis of site-selectively modified peptides and proteins. Finding peptide sequence selectivity by deoxyribozymes such as DzAz2, together with such observations made for two previously reported DNA-catalyzed reactions,[24,25] establishes that DNA enzymes have the broader ability to interact with side chains of peptide substrates. The identification of DzAz8 also expands the list of DNA enzymes that can catalyze reactions of free peptide substrates. Future efforts are focused on combining the key features of peptide sequence selectivity and reactivity with free peptide substrates, and extending DNA-catalyzed reactivity from peptides to larger protein substrates.

3.4 Materials and Methods

3.4.1 Substrate Preparation Procedures

Oligonucleotides, peptides, and DNA-anchored peptide conjugates were prepared as described in Chapter 2. PEG_{5k}-alkyne was obtained from Laysan Bio (Arab, AL). Fluorescein-alkyne was obtained from Tenova Pharmaceuticals (San Diego, CA). 2′-Az-dATP was obtained

from TriLink BioTechnologies (San Diego, CA). The peptides sCT and ANP were obtained from Anaspec (Fremont, CA).

Preparation of azido-modified oligonucleotide for capture control. A sample containing 5 nmol of 3′-C_3-NH_2 oligonucleotide was brought to 100 μL total volume containing 100 mM MES, pH 6.0, 6 mM 6-azidohexanoic acid (Chem-Impex), 6 mM *N*-hydroxysuccinimide (NHS), 50 mM *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide (EDC), and 10% (v/v) DMF. The sample was incubated at room temperature for 12 h and purified by HPLC [Shimadzu Prominence instrument; Phenomenex Gemini-NX C_{18} column, 5 μm, 10 × 250 mm; gradient of 10% solvent A (20 mM triethylammonium acetate in 50% acetonitrile/50% water, pH 7.0) and 90% solvent B (20 mM triethylammonium acetate in water, pH 7.0) at 0 min to 70% solvent A and 30% solvent B at 40 min with flow rate of 3.5 mL/min].

Table 3.2: Oligonucleotide sequences used in the selection experiments. All sequences are written 5′ to 3′. For all selections, the reverse PCR primer was (AAC)$_4$XCCATCAGGATCAGCT, where X denotes the hexa(ethylene glycol) [HEG] spacer to stop Taq polymerase. Table adapted with permission from ref. 19.

oligonucleotide purpose	oligonucleotide sequence
Selection with DNA-HEG-CAAYAA (leading to DzAz1)	
DNA-HEG-CAAYAA substrate	GGACTATTGAAAGACATAT-HEG-CAAYAA
forward primer for selection	CGAAATGATGGCTATTTC
random pool for selection	CGAAATGATGGCTATTTC-N_{40}-ATATGTCTTTCAATAGAGCTGATCCTGATGG
splint for ligation step during selection	ATATGTCTTTCAATAGTCCCCATCAGGATCAGCTCTATTGAAAGACATAT
Reselection of DzAz1 deoxyribozyme (leading to DzAz1b and DzAz1c)	
DNA-HEG-CAAYAA substrate	GGATCCTGGATACAAATAT-HEG-CAAYAA
forward primer for selection	CGAAGTATAAACCTGTTC
partially randomized pool for selection	CGAAGTATAAACCTGTTC-N_{40}-ATATTTGTATCCAGGAAGCTGATCCTGATGG
splint for ligation step during selection	ATATTTGTATCCAGGATCCCCATCAGGATCAGCTTCCTGGATACAAATAT
Selection with DNA HEG-CLQTYPRT (leading to DzAz2–DzAz6)	
DNA-HEG-CLQTYPRT substrate	GGACTACCTTTATGCGTAT-HEG-CLQTYPRT
forward primer for selection	CGAACGAAAGCCTCCTTC
random pool for selection	CGAACGAAAGCCTCCTTC-N_{40}-ATACGCATAAAGGTAGAGCTGATCCTGATGG
splint for ligation step during selection	ATACGCATAAAGGTAGTCCCCATCAGGATCAGCTCTACCTTTATGCGTAT

Table 3.2 (cont.)

oligonucleotide purpose	oligonucleotide sequence
Selection with DNA HEG-CQQPYITN (leading to DzAz7)	
DNA-HEG-CQQPYITN substrate	GGAATATCTCGTTTCTTAT–HEG-CQQPYITN
forward primer for selection	CGAATTAAGACTGAATTC
random pool for selection	CGAATTAAGACTGAATTC–N_{40}–ATAAGAAACGAGATATAGCTGATCCTGATGG
splint for ligation step during selection	ATAAGAAACGAGATATTCCCCATCAGGATCAGCTATATCTCGTTTCTTAT
Selection with DNA HEG-CERSYLMK (leading to DzAz8)	
DNA-HEG-CERSYLMK substrate	GGAATGGCTTGATTGGTAT–HEG-CERSYLMK
forward primer for selection	CGAATTGAGTAAATATTC
random pool for selection	CGAATTGAGTAAATATTC–N_{40}–ATACCAATCAAGCCATAGCTGATCCTGATGG
splint for ligation step during selection	ATACCAATCAAGCCATTCCCCATCAGGATCAGCTATGGCTTGATTGGTAT
Selection with DNA HEG-CFQPYMQE (leading to DzAz9)	
DNA-HEG-CFQPYMQE substrate	GGATCAGGTTACTAATTAT–HEG-CFQPYMQE
forward primer for selection	CGAAATAGATTATCATTC
random pool for selection	CGAAATAGATTATCATTC–N_{40}–ATAATTAGTAACCTGAAGCTGATCCTGATGG
splint for ligation step during selection	ATAATTAGTAACCTGATCCCCATCAGGATCAGCTTCAGGTTACTAATTAT

3.4.2 In Vitro Selection Procedures

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

Procedure for selection step in round 1. Each selection experiment was initiated with 200 pmol of the ligated pool. A 16 μL sample containing 200 pmol of 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 40 μL total volume containing 70 mM HEPES, pH 7.5, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$, 150 mM NaCl, and 100 μM 2′-Az-dATP. The Mn^{2+} was added from a 10× stock solution containing 200 mM $MnCl_2$. The Zn^{2+} was added from a 10× stock solution containing 10 mM $ZnCl_2$, 20 mM HNO_3, and 200 mM HEPES at pH 7.5; this stock solution was freshly prepared from a 100× stock of 100 mM $ZnCl_2$ in 200 mM HNO_3. The metal ion stocks were added last to the final sample. The sample was incubated at 37 °C for 14 h and precipitated with ethanol.

Procedure for selection step in subsequent round. An 8 μL sample containing the 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, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$, 150 mM NaCl, and 100 μM 2′-Az-dATP. The sample was incubated at 37 °C for 14 h and precipitated with ethanol.

Procedure for capture step in each round. The sample from the selection step was brought to 20 μL total volume containing 100 mM HEPES, pH 7.5, 10 mM PEG_{5k}-alkyne, 40 mM tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), 10 mM sodium ascorbate, and 5 mM $CuSO_4$. The concentrations of all reagents were optimized (not shown). The four reagents were added from 100 mM (10×), 1 M (25×), freshly prepared 100 mM (10×), and 100 mM (20×) stock solutions, respectively; the four reagents were mixed together and then added at once to the other components. The sample was incubated at 4 °C for 1 h and separated by 8% PAGE.

3.4.3 Cloning and Screening

Cloning and screening were performed as described in Chapter 2.

3.4.4 Single-Turnover Deoxyribozyme Assay Procedure

The DNA-anchored peptide substrate was 5′-^{32}P-radiolabeled using γ-^{32}P-ATP and T4 polynucleotide kinase. An 8 μL sample containing 0.1 pmol of 5′-^{32}P-radiolabeled DNA-anchored peptide substrate and 2 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 azido-adenylylation 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 $MgCl_2$, 150 mM NaCl, and 100 μM 2′-Az-dATP or ATP. The sample was incubated at 37 °C. At appropriate time points, a 2 μL aliquot was quenched with 6 μL of 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.

3.4.5 Tyrosine Site Discrimination Assay for DzAz2

The azido-adenylylation product of peptide A or peptide B was prepared from an 8 μL sample containing 600 pmol of DNA-anchored HEG-tethered peptide substrate and 660 pmol of deoxyribozyme, which were annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The DNA-catalyzed azido-adenylylation 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 $MgCl_2$, 150 mM NaCl, and 100 μM 2′-Az-dATP. The sample was incubated at 37 °C for 24 h and separated by 20% PAGE. The product was quantified by absorbance (A_{260}); in each case, ~400 pmol was obtained. From each sample, 50 pmol was dissolved in 5 μL of water and brought to a total volume of 10 μL containing 100 mM Tris, pH 8.0, 40 mM DTT, and 0.6 ng/μL Lys-C (Roche cat. no. 11420429001). The sample was incubated at 37 °C for 12 h, desalted by Millipore C_{18} ZipTip, and analyzed by MALDI mass spectrometry (Bruker UltrafleXtreme, matrix 2,5-dihydroxybenzoic acid).

3.4.6 Free Peptide Assay Procedure

The azido-adenylylation or adenylylation product from each untethered peptide was prepared from a 10 μL sample containing 3.3 nmol of free unmodified DNA anchor oligonucleotide (lacking 3′-phosphate, HEG, and thiol) and 3 nmol of deoxyribozyme, which were annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The DNA-catalyzed azido-adenylylation reaction was

initiated by bringing the sample to 30 μL total volume containing 70 mM HEPES, pH 7.5, 5 mM $ZnCl_2$ (optimized), 20 mM $MnCl_2$, 40 mM $MgCl_2$, 150 mM NaCl, 100 μM untethered peptide and 1 mM 2′-Az-dATP or ATP. The sample was incubated at 37 °C for 24 h and separated by HPLC (Shimadzu Prominence instrument; Phenomenex Jupiter Proteo C_{12} column, 4 μm, 10 × 250 mm). The reaction for azido-adenylylation or adenylylation of sCT was analyzed with a gradient of 28% solvent A (acetonitrile) and 72% solvent B (0.1% TFA in water) at 0 min to 40% solvent A and 60% solvent B at 60 min with flow rate of 2 mL/min. The reaction for azido-adenylylation or adenylylation of ANP was analyzed with a gradient of 17% solvent A and 83% solvent B at 0 min to 29% solvent A and 71% solvent B at 60 min with flow rate of 2 mL/min. The HPLC-purified sample was analyzed by MALDI mass spectrometry (Bruker UltrafleXtreme, matrix 2,5-dihydroxybenzoic acid). The HPLC response factor of azido-adenylylated or adenylylated peptide over unmodified peptide was assigned as 1.2 on the basis of UV absorbance data (A_{205}) for sCT, ANP, and AMP. The yield of each DzAz8-catalyzed reaction was calculated from the peak integrals, accounting for the response factor.

3.4.7 CuAAC Derivatization of Azido-Adenylylated Peptide Substrates

For CuAAC modification of each azido-adenylylated peptide product, 0.4 nmol of HPLC-purified sample was brought to 5 μL total volume containing 100 mM HEPES, pH 7.5, 10 mM PEG_{5k}-alkyne or fluorescein-alkyne, 40 mM tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), 10 mM sodium ascorbate, and 5 mM $CuSO_4$. The final four reagents were added from 100 mM (10×; PEG_{5k}-alkyne in water and fluorescein-alkyne in DMSO), 1 M (25×), freshly prepared 100 mM (10×), and 100 mM (20×) stock solutions, respectively; the four reagents were mixed together and then added at once to the other components. The sample was incubated at 4 °C for 1 h. For SDS-PAGE analysis, the sample was separated by 16.5% Tris-Tricine SDS-PAGE. For mass spectrometry, the sample was analyzed by MALDI mass spectrometry (Bruker UltrafleXtreme, matrix 2,5-dihydroxybenzoic acid).

3.4.8 Mass Spectrometry

The azido-adenylylation product was prepared from an 8 μL sample containing 600 pmol of DNA-anchored HEG-tethered peptide substrate and 660 pmol of deoxyribozyme, which were annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The DNA-catalyzed azido-adenylylation 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 $MgCl_2$, 150 mM NaCl, and 100 μM 2′-Az-dATP. The sample was incubated at 37 °C for 24 h and separated by 20% PAGE. The PAGE-purified sample was desalted by Millipore C_{18} ZipTip and analyzed by MALDI mass spectrometry (Bruker UltrafleXtreme; matrix 3-hydroxypicolinic acid).

Table 3.3: MALDI mass spectrometry data. "DNA" here refers to the DNA anchor oligonucleotide. The azido group in azido-adenylylation product of peptide A or peptide B was reduced to an amino group by the DTT used to cleave the peptide from the DNA anchor oligonucleotide.[27] The calculated mass takes this reaction into account. Table adapted with permission from ref. 19.

deoxyribozyme	product	$[M+H]^+$ calcd.	$[M+H]^+$ found	error, % (found – calcd.)
DzAz1	DNA-HEG-CAAY(Az-dAMP)AA	7317.4	7319.6	+0.03
DzAz2	DNA-HEG-CLQTY(Az-dAMP)PRT	7664.6	7667.6	+0.04
DzAz7	DNA-HEG-CQQPY(Az-dAMP)ITN	7639.6	7642.5	+0.04
DzAz8	DNA-HEG-CERSY(Az-dAMP)LMK	7832.7	7835.8	+0.04
DzAz2	peptide A, Prod 1_R	862.4	862.5	+0.01
DzAz2	peptide A, Prod 2_R	1190.0	1190.6	+0.05
DzAz2	peptide A, Prod 2_L	1309.6	1309.7	+0.008
DzAz2	peptide A, Prod 1_L	1637.7	1637.8	+0.006
DzAz2	peptide B, Prod 2_R	877.5	877.5	0
DzAz2	peptide B, Prod 1_R	1205.0	1205.6	+0.05
DzAz2	peptide B, Prod 1_L	1294.6	1294.7	+0.008
DzAz2	peptide B, Prod 2_L	1622.7	1622.8	+0.006
DzAz8	sCT-(Az-dAMP)	3784.8	3787.6	+0.07
DzAz8	ANP-(Az-dAMP)	3433.5	3434.7	+0.03
DzAz8	sCT-AMP	3759.8	3762.9	+0.08
DzAz8	ANP-AMP	3408.5	3410.3	+0.05
–	sCT-Fluorescein	4197.9	4199.4	+0.04
–	ANP-Fluorescein	3846.6	3844.6	–0.05

3.5 References

(1) Boutureira, O.; Bernardes, G. J. L. Advances in chemical protein modification. *Chem. Rev.* **2015**, *115*, 2174-2195.

(2) Krall, N.; da Cruz, F. P.; Boutureira, O.; Bernardes, G. J. Site-selective protein-modification chemistry for basic biology and drug development. *Nat. Chem.* **2016**, *8*, 103-113.

(3) Jones, L. H.; Narayanan, A.; Hett, E. C. Understanding and applying tyrosine biochemical diversity. *Mol. BioSyst.* **2014**, *10*, 952-969.

(4) Hooker, J. M.; Kovacs, E. W.; Francis, M. B. Interior surface modification of bacteriophage MS2. *J. Am. Chem. Soc.* **2004**, *126*, 3718-3719.

(5) Gavrilyuk, J.; Ban, H.; Nagano, M.; Hakamata, W.; Barbas, C. F., III. Formylbenzene diazonium hexafluorophosphate reagent for tyrosine-selective modification of proteins and the introduction of a bioorthogonal aldehyde. *Bioconjug. Chem.* **2012**, *23*, 2321-2328.

(6) Zhang, J.; Ma, D.; Du, D.; Xi, Z.; Yi, L. An efficient reagent for covalent introduction of alkynes into proteins. *Org. Biomol. Chem.* **2014**, *12*, 9528-9531.

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

(8) Guo, H. M.; Minakawa, M.; Ueno, L.; Tanaka, F. Synthesis and evaluation of a cyclic imine derivative conjugated to a fluorescent molecule for labeling of proteins. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 1210-1213.

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

(10) Seim, K. L.; Obermeyer, A. C.; Francis, M. B. Oxidative modification of native protein residues using cerium(IV) ammonium nitrate. *J. Am. Chem. Soc.* **2011**, *133*, 16970-16976.

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

(12) Nilo, A.; Allan, M.; Brogioni, B.; Proietti, D.; Cattaneo, V.; Crotti, S.; Sokup, S.; Zhai, H.; Margarit, I.; Berti, F.; Hu, Q. Y.; Adamo, R. Tyrosine-directed conjugation of large glycans to proteins via copper-free click chemistry. *Bioconjug. Chem.* **2014**, *25*, 2105-2111.

(13) Bauer, D. M.; Ahmed, I.; Vigovskaya, A.; Fruk, L. Clickable tyrosine binding bifunctional linkers for preparation of DNA-protein conjugates. *Bioconjug. Chem.* **2013**, *24*, 1094-1101.

(14) Struck, A. W.; Bennett, M. R.; Shepherd, S. A.; Law, B. J.; Zhuo, Y.; Wong, L. S.; Micklefield, J. An enzyme cascade for selective modification of tyrosine residues in structurally diverse peptides and proteins. *J. Am. Chem. Soc.* **2016**, *138*, 3038-3045.

(15) Bruins, J. J.; Westphal, A. H.; Albada, B.; Wagner, K.; Bartels, L.; Spits, H.; van Berkel, W. J. H.; van Delft, F. L. Inducible, site-specific protein labeling by tyrosine oxidation-strain-promoted (4 + 2) cycloaddition. *Bioconjug. Chem.* **2017**, *28*, 1189-1193.

(16) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. Bioconjugation by copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition. *J. Am. Chem. Soc.* **2003**, *125*, 3192-3193.

(17) Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. G. Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angew. Chem. Int. Ed.* **2009**, *48*, 9879-9883.

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

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

(20) Itzen, A.; Blankenfeldt, W.; Goody, R. S. Adenylylation: renaissance of a forgotten post-translational modification. *Trends Biochem. Sci.* **2011**, *36*, 221-228.

(21) Hedberg, C.; Itzen, A. Molecular perspectives on protein adenylylation. *ACS Chem. Biol.* **2015**, *10*, 12-21.

(22) Jones, M. W.; Mantovani, G.; Blindauer, C. A.; Ryan, S. M.; Wang, X.; Brayden, D. J.; Haddleton, D. M. Direct peptide bioconjugation/PEGylation at tyrosine with linear and branched polymeric diazonium salts. *J. Am. Chem. Soc.* **2012**, *134*, 7406-7413.

(23) Dudakov, J. A.; Hanash, A. M.; van den Brink, M. R. M. Interleukin-22: immunobiology and pathology. *Annu. Rev. Immunol.* **2015**, *33*, 747-785.

(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) Vesely, D. L. Which of the cardiac natriuretic peptides is most effective for the treatment of congestive heart failure, renal failure and cancer? *Clin. Exp. Pharmacol. Physiol.* **2006**, *33*, 169-176.

(27) Abe, H.; Wang, J.; Furukawa, K.; Oki, K.; Uda, M.; Tsuneda, S.; Ito, Y. A reduction-triggered fluorescence probe for sensing nucleic acids. *Bioconjug. Chem.* **2008**, *19*, 1219-1226.

Chapter 4: Efforts towards DNA-Catalyzed Tyrosine Azido-Adenylylation of Protein Substrates[†]

4.1 Introduction

4.1.1 DNA-Catalyzed Modification of Protein Substrates

Deoxyribozymes with the ability to site-specifically modify protein substrates can be powerful tools for biochemical studies. However, the phosphatase 14WM9 deoxyribozyme is the only DNA enzyme known to function with a protein substrate.[1] The 14WM9 deoxyribozyme can catalyze the dephosphorylation reaction in a 91-mer protein that has the phosphotyrosine-containing peptide segment, CAAY^{P}AA, at the C-terminus (Figure 4.1). Although this study showed that the 14WM9 DNA enzyme was capable of recognizing the Y^{P} residue in the context of a protein substrate, the reaction site was located at the unstructured end of the protein. Therefore, this result is still very different from modifying an amino acid side chain in a protein region with a well-defined structure.

[†] Former Silverman laboratory graduate student Jagadeeswaran Chandrasekar identified the phosphatase deoxyribozymes and evaluated the 14WM9 activity with the protein substrate.[1]

Former Silverman laboratory graduate student Chih-Chi Chu evaluated the recruiting effects of His_6 tag and thrombin-binding aptamers on deoxyribozymes.[2, 3]

Former Silverman laboratory staff Kevin M. Duffin performed in vitro selection experiments with unmodified DNA pools to identify aptamers that bind human annexin V and human TNF-related apoptosis-inducing ligand (TRAIL) 114–281.

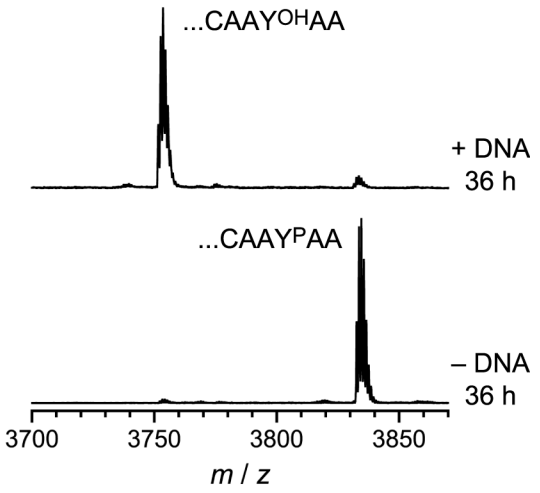


Figure 4.1. Dephosphorylation of a larger protein substrate by the 14WM9 deoxyribozyme. The 91-mer protein derived from prochlorosin ProcA2.8 and ending with ...CAAYPAA at its C-terminus was dephosphorylated by the 14WM9 deoxyribozyme, digested in-gel by Lys-C, and analyzed by MALDI mass spectrometry. Figure adapted with permission from ref. 1.

One of the challenges associated with DNA-catalyzed modification of protein substrates is the accessibility by a deoxyribozyme to the target residue on a structured protein, where both the DNA enzyme and the protein substrate are large biomolecules. Another challenge is the sufficient interaction between the deoxyribozyme and the target peptide segment to enable catalysis within the context of the protein structure. Both challenges might be overcome by identifying DNA enzymes from selection experiments that directly use protein substrates. If a deoxyribozyme has the ability to access and interact with the target residue for catalysis, then another major practical challenge is to achieve catalysis at low protein concentrations. In the previous chapter, azido-adenylylation of untethered peptides was performed with 100 μM of the substrates. Such concentration might be impractically high for protein substrates. The concentration issue can be addressed if deoxyribozymes are recruited to their target proteins.

4.1.2 Recruiting DNA Enzymes to Their Substrates by Hexahistidine (His_6) Tags

The ability of hexahistidine (His_6) tags on peptide and protein substrates to recruit deoxyribozymes for modifying those substrates has been evaluated by former Silverman lab graduate student Chih-Chi Chu.[2] In this study, the metal-mediated interaction between a His_6 tag

and a nitrilotriacetic acid (NTA) group was used (Figure 4.2A). The tris(NTA)-modified DNA anchor oligonucleotide forms Watson-Crick base pairs with the fixed-sequence binding arm of the deoxyribozyme, which allows recruiting of the DNA enzyme to the His_6-tagged peptide or protein substrate.

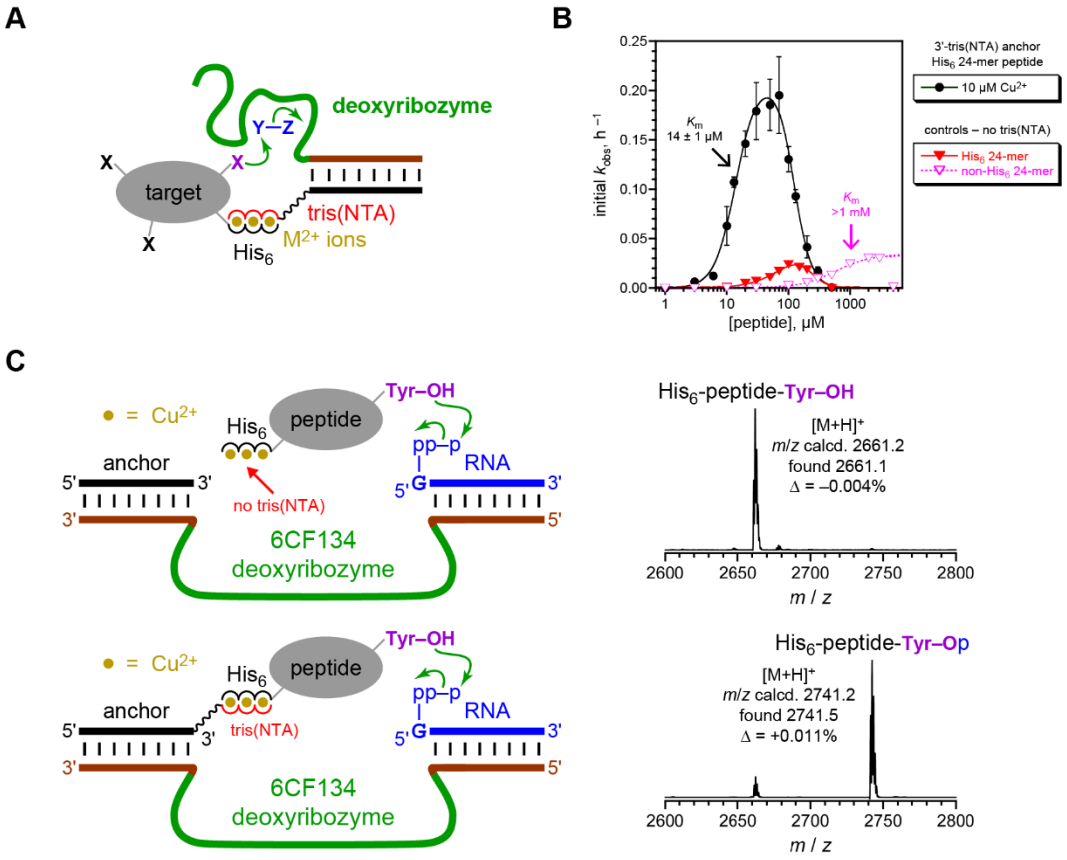


Figure 4.2. Recruiting DNA enzymes to their substrates by hexahistidine (His_6) tags. (A) The strategy of recruiting a DNA enzyme to its peptide or protein substrate for DNA-catalyzed modification of peptide or protein side chains. X = the amino acid side chain; Y–Z = the reaction partner for the peptide or protein. (B) Histidine tag recruiting improves both reaction rate and K_m for the nucleopeptide-forming 8XJ105 deoxyribozyme using Cu^{2+} as the recruiting metal ion. (C) MALDI mass spectrometry analysis of the histidine recruiting effect for 6CF134. A negative control omitting the 3′-tris(NTA) is shown with 0.9% conversion of the product. Figure modified with permission from ref. 2.

For the nucleopeptide-forming 8XJ105 deoxyribozyme, this His_6 tag recruiting strategy provided substantial improvements in yield, k_{obs}, and K_m for peptide substrates (Figure 4.2B). The His_6/Cu^{2+} recruiting effect also allowed, for the first time, successful DNA-catalyzed tyrosine

phosphorylation of a discrete, untethered peptide substrate by the tyrosine kinase 6CF134 deoxyribozyme (Figure 4.2C). However, where tested, DNA-catalyzed protein modification using this recruiting strategy was not observed. This unsuccessful outcome with protein substrates may be due to the inaccessibility of the targeting amino acid side chain or the difficulty for the deoxyribozymes to have enough interactions with the peptide segment for catalysis within the context of the protein structure.

4.1.3 Recruiting DNA Enzymes to Their Substrates by Aptamer Modules

The use of aptamers that bind to target proteins is another promising strategy for recruiting DNA enzymes to their protein substrates. Former Silverman lab graduate student Chih-Chi Chu also investigated the recruiting effect of thrombin-binding aptamer on the nucleopeptide-forming 8XJ105 deoxyribozyme.[3] In this study, the previously identified thrombin-binding 29-mer DNA aptamer HD22-29 (Apt29) was connected to the 8XJ105 deoxyribozyme at either the 3′- or the 5′-end of the DNA enzyme by an oligonucleotide linker (A_{15}). 8XJ105 is a peptide-sequence-general deoxyribozyme which shows robust catalytic yields with four thrombin-related tyrosine-containing peptides (Figure 4.3A).

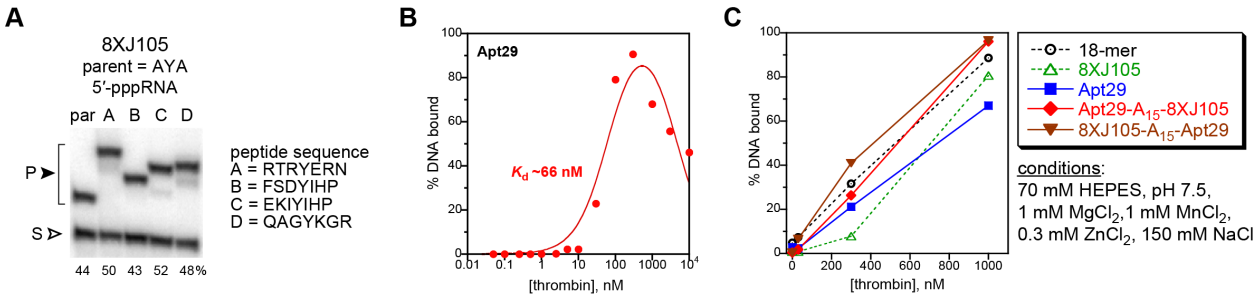


Figure 4.3. Recruiting DNA enzymes to their substrates by aptamer modules, exemplified by the 8XJ105 deoxyribozyme. (A) Peptide sequence tolerance of the 8XJ105 deoxyribozyme. The four peptides used are thrombin-related peptide segments located on the protein surface. (B) Determination of K_d value of Apt29 by filter-binding assay. (C) Thrombin binding of Apt29 and Apt29-8XJ105 conjugates in conditions where the Apt29-8XJ105 conjugates are catalytically active. Nonspecific binding was also observed for 8XJ105 and the arbitrary 18-mer DNA oligonucleotide $(AAC)_6$. Figure adapted with permission from ref. 3.

Substantial binding of thrombin was observed for Apt29 at protein concentration of 100 nM, when the aptamer was not connected to the 8XJ105 deoxyribozyme and tested under aptamer binding conditions (Figure 4.3B). However, under DNA-catalyzed reaction conditions, severe nonspecific binding was observed in negative controls with only 8XJ105 or the arbitrary 18-mer DNA oligonucleotide $(AAC)_6$, even after extensive optimization of the conditions (Figure 4.3C). The binding of the Apt29-8XJ105 conjugates to thrombin was not verified due to nonspecific binding. Furthermore, no DNA-catalyzed reaction product was observed in SDS-PAGE for assays of Apt29-8XJ105 conjugates with thrombin (data not shown).

In conclusion, for the case evaluated (one deoxyribozyme with one protein substrate), joining a DNA aptamer unit to a deoxyribozyme did not enable DNA-catalyzed modification of a protein substrate. However, the major issue observed in this study was the incompatibility between the aptamer binding conditions and the DNA-catalyzed reaction conditions. We anticipate that this issue can be resolved by de novo identification of suitable DNA aptamers for the DNA enzyme selection conditions.

4.1.4 Modified Nucleotides Incorporated into Aptamers

DNA aptamers have been identified to bind to many protein substrates via SELEX (Systematic Evolution of Ligands by EXponential enrichment). To increase the binding affinity between the aptamers and protein substrates, DNA aptamers with hydrophobic modifications have been identified (Figure 4.4A).[4-6] In many cases, unmodified aptamers have weak binding affinities (K_d > 100 nM). Upon incorporation of hydrophobic modifications, selection experiments result in the identification of modified aptamers that have better binding affinities with dissociation constants in the low nanomolar range. The use of hydrophobic modifications during selection experiments is different from efforts in which modifications are incorporated into aptamers to improve their binding affinities after their identification from unmodified DNA pools.

A

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

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

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

B

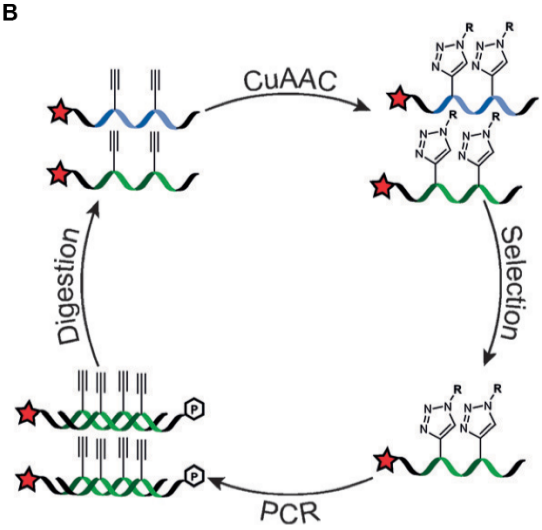


Figure 4.4. Modified nucleotides incorporated into aptamers. (A) Binding affinities of SELEX libraries selected for binding to the 14 different proteins indicated with either unmodified, Bn-modified, iBu-modified, or Trp-modified DNA. Binding affinities are shown (K_d, nM). Figure adapted with permission from ref. 5. (B) Schematic representation of the click-SELEX process. Figure adapted with permission from ref. 7.

The introduction of hydrophobic modifications can be achieved by either using modified nucleotide triphosphates in PCR reactions or using modified phosphoramidites during solid-phase DNA synthesis. Considerable synthetic efforts are usually required for the preparation of modified nucleotide triphosphates or phosphoramidites. A different approach, known as click-SELEX, employs copper-catalyzed azide-alkyne cycloaddition (CuAAC) to generate modified DNA libraries for aptamer selection experiments (Figure 4.4B).[7] First, an alkyne-modified DNA library is prepared by replacing thymidine with C5-ethynyl-2′-deoxyuridine (^{E}dU). Second, this library is further modified through click reaction with an azide-containing compound, e.g., 3-(2-azidoethyl)indole as shown in Figure 4.4B. This library can be then applied for in vitro selection experiments. As a proof of concept, the indole-modified DNA library was used in a selection experiment for binding to an improved GFP derivative, cycle 3 GFP (C3-GFP). An aptamer C12 was identified from the enriched pool of round 15 and showed strong binding to C3-GFP with both high affinity (K_d = 18.4 nM) and high specificity. Therefore, this click-SELEX strategy is likely to be a versatile approach towards identification of nucleobase-modified

aptamers. In this chapter, we used the click-SELEX strategy to perform selection experiments for the identification of modified aptamers that bind protein substrates.

4.2 Results and Discussion

4.2.1 Development of Capture Method with Alkyne-Modified Resin

To identify deoxyribozymes with azido-adenylylation activity on proteins, selection experiments were designed to directly use target proteins as the substrates. As discussed in Chapter 3, the capture step is essential to the identification of deoxyribozymes from in vitro selection experiments. The previously used CuAAC capture reaction with alkyne-modified poly(ethylene glycol) that has an average molecular weight of 5000 (PEG_{5k}-alkyne) enables the isolation of active DNA sequences based on their migration in polyacrylamide gel electrophoresis (PAGE). Even though there is no literature report about the recovery efficiency of DNA-protein conjugates from PAGE gels, the recovery of proteins by gel extraction is generally low in yield.[8] Therefore, this gel-based capture approach was not used initially in the selection experiments with protein substrates.

A resin-based capture method was developed as the alternative (Figure 4.5A). After the selection step, the sample was precipitated and incubated with alkyne-modified controlled pore glass (CPG-alkyne) together with CuAAC reagents. After the connection of active DNA sequences to the CPG, the resin was washed, removing the unbound and inactive DNA sequences. The bound deoxyribozymes were eluted from the resin by treatment with dithiothreitol (DTT) that cleaved the disulfide bond between the modified protein substrate and the active DNA sequences. Scintillation counting of the wash and elution fractions was used to quantify activity.

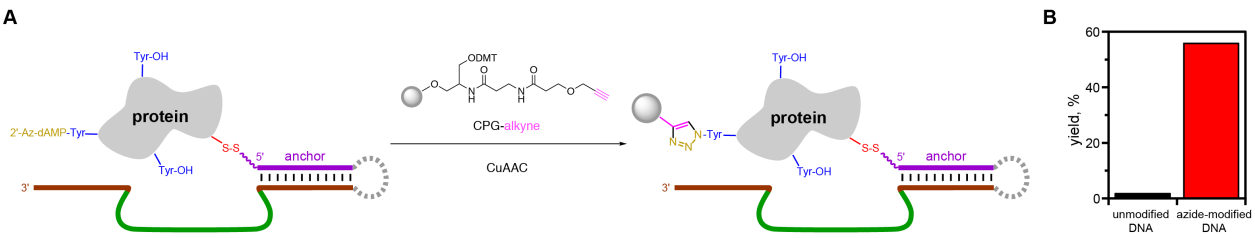


Figure 4.5. Capture method using CPG-alkyne. (A) Schematic illustration of the CuAAC capture reaction between azide-modified product and CPG-alkyne. (B) Capture assay with unmodified DNA and azide-modified DNA standard.

The CPG-alkyne capture method was optimized for use during in vitro selection of azido-adenylylation deoxyribozymes. To prevent the DNA pool from adhering to the reaction tube, low-retention tubes were coated with a 5% PEG solution, incubated at room temperature overnight, and rinsed prior to use. To the sample was also added 0.5 µg of herring sperm DNA as a sacrificial oligonucleotide. The CPG-alkyne capture method was efficient for an azide-modified DNA oligonucleotide (prepared by conjugating a 3′-C_3-NH_2 oligonucleotide with 6-azidohexanoic acid, as described in Chapter 3) with 56% capture yield, whereas <3% of the unmodified DNA was captured (Figure 4.5B).

One major concern with the resin-based capture method is the unintended enrichment of aptamer sequences that bind to the resin via non-covalent interactions. A common strategy used to avoid the emergence of DNA aptamers for the resin is the inclusion of a pre-selection step between the ligation and selection steps. In this pre-selection step, the DNA sequences ligated to the substrate are subjected to CuAAC capture reaction with CPG-alkyne. The DNA sequences that do not bind to the CPG-alkyne are recovered in the wash and taken on to the selection step. However, this pre-selection step is not applicable in the selection experiments with protein substrates, because the reaction conditions used for CPG-alkyne capture may denature the protein substrates. Therefore, a pre-selection step could not be included in the selection experiments, even though the outcome of the selection experiments can be the emergence of aptamer sequences that bind to the CPG.

4.2.2 Selection Experiments with Protein Substrates Using Resin-Based Capture Method

In vitro selection experiments were performed to identify azido-adenylylation deoxyribozymes with protein substrates. Two proteins, human annexin V and human TNF-related apoptosis-inducing ligand (TRAIL) 114–281, and a 36-mer peptide pancreatic polypeptide (PP) with an additional C-terminal cysteine were used as the substrates for the selection experiments (Figure 4.6). Annexin V is commonly used to detect apoptotic cells by its ability to bind phosphatidylserine that is a marker of apoptosis.[9-11] TRAIL binds to the death receptors DR4 and DR5 to induce caspase-8-dependent apoptosis.[12] Full-length TRAIL is composed of an N-terminal cytoplasmic domain, a transmembrane domain, and a C-terminal extracellular domain. The protein substrate TRAIL 114–281 is a soluble portion of the extracellular domain with similar function to the full-length protein and is henceforth abbreviated simply as TRAIL. PP is a polypeptide secreted in the pancreas to self-regulate pancreatic secretion activities.[13] Both annexin V and TRAIL have only one cysteine residue in their amino acid sequences, which allows site-specific connections to the DNA anchor oligonucleotides. An artificial N-terminal cysteine was appended to PP to enable disulfide linkage to the DNA anchor oligonucleotide. The two protein substrates have very distinctive structures, with annexin V rich in α-helices and TRAIL rich in β-sheets. Therefore, the result from the selection experiments using these proteins may provide information about the preference for protein structure by deoxyribozymes. Each protein or peptide substrate was covalently attached at the cysteine residue via a disulfide linkage and a hexa(ethylene glycol) [HEG] tether to a DNA anchor oligonucleotide, which was bound by Watson-Crick base pairs to one of the fixed-sequence binding arms of the random DNA pool. The DNA anchor oligonucleotide contained three ribonucleotides (gga) at the 3′-end that enables the ligation to the initially random DNA pool via T4 RNA ligase. The initially random region of the DNA sequences was flanked by fixed-sequence regions to enable PCR amplification of the DNA sequences and substrate binding.

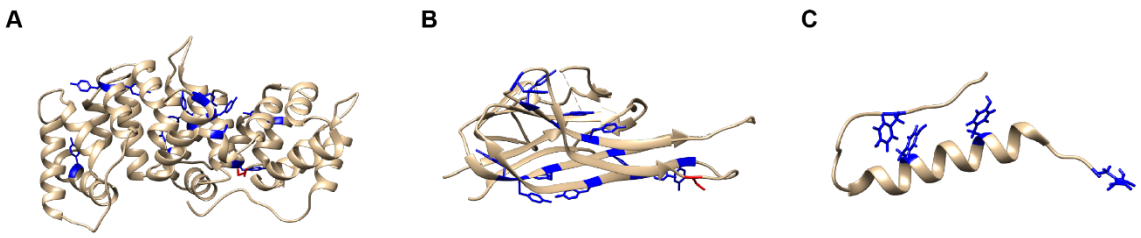


Figure 4.6. Structures of the protein and peptide substrates used in the selection experiments. (A) Annexin V (PDB: 1HVG). (B) TRAIL (PDB: 1DG6). (C) PP (PDB: 1BBA). The amino acid sequences of these substrates can be found via the PDB entry. The amino acid sequence of PP is also listed in Table 4.5.

The in vitro selection process began with ligation of the initially random DNA pool to the DNA-anchored substrate (Figure 4.7). The ligation product was desalted by a Micro Bio-Spin P-6 column (Bio-Rad). Next, the selection step was performed in which the DNA sequences were incubated with 2′-Az-dATP and divalent metal ions in a buffered solution to enable DNA catalysis. After exposure of the substrate-conjugated pool to 2′-Az-dATP, the catalytically active DNA sequences, which were now attached to an azido group, were precipitated and captured with the CPG-alkyne via the CuAAC reaction. After the CPG was washed extensively, the active DNA sequences were eluted from the resin by treatment with dithiothreitol (DTT), which cleaved the disulfide bond between the protein or peptide substrate and the DNA anchor. PCR was performed to amplify the catalytically active DNA sequences that survived the previous selection, capture, and disulfide cleavage steps. The entire selection process was iterated, each time enriching the population with catalytically active sequences, until the DNA enzymes dominated the population.

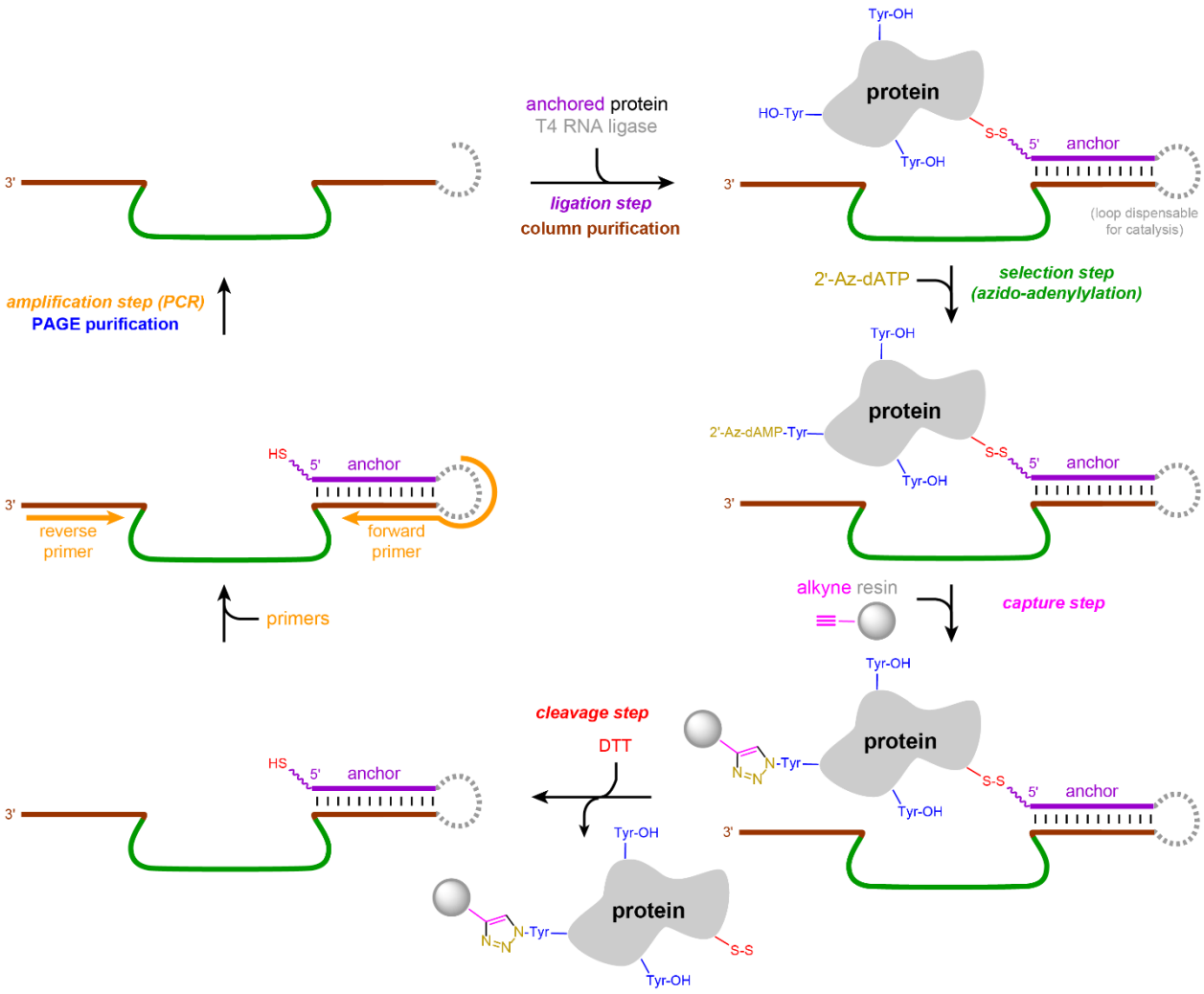


Figure 4.7. Diagram of in vitro selection procedure for deoxyribozymes with tyrosine azido-adenylylation activity on protein substrates using resin-based capture method.

The selection experiments designated as AK2–AP2 according to our systematic alphanumeric nomenclature were performed to identify azido-adenylylation deoxyribozymes that are functional with protein and long peptide substrates (Figure 4.8). The selection experiments with protein substrates were performed with an N_{40} or N_{80} initially random region. The length of the initially random region is an important variable in deoxyribozyme selection experiments.[14] Larger random regions may enable more complex catalytic DNA structures, which may lead to more interactions between the DNA enzyme and the protein substrate. However, all selection experiments evaluate 10^{14} DNA sequences, which is a much smaller fraction of the 10^{48} possible

N_{80} sequences compared to the fraction of the 10^{24} possible N_{40} sequences. Therefore, for any specific random region length used for a selection experiment, a compromise is made between the structural complexity and the sequence-space coverage.

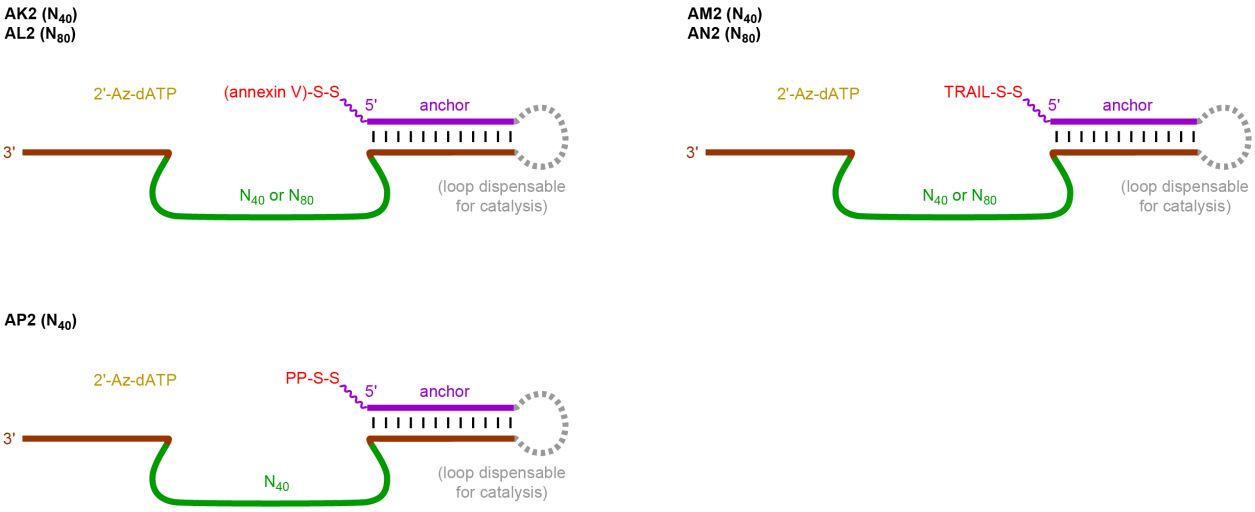


Figure 4.8. Design of the AK2–AP2 selection experiments to identify azido-adenylylation deoxyribozymes with protein and long peptide substrates. Initially random region lengths are indicated.

All selections were performed with 100 μM 2′-Az-dATP in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h. After 4 rounds of selection, the capture yield of the AP2 pool reached 34% (Figure 4.9A). However, when the 4AP2 pool was not incubated with 2′-Az-dATP, a capture yield of 32% was observed (Figure 4.9B). This result indicated that the 4AP2 pool was dominated by aptamer sequences that bind to the CPG-alkyne, and the observed capture yield was due to non-catalytic binding rather than azido-adenylylation. This negative outcome of the AP2 selection experiment reemphasized that the lack of a pre-selection step during selection experiments using resin-based capture methods makes these selection experiments prone to the emergence of aptamer sequences. Therefore, all five of the AK2–AP2 selection experiments were discontinued. The capture approach based on CPG-alkyne was also discontinued.

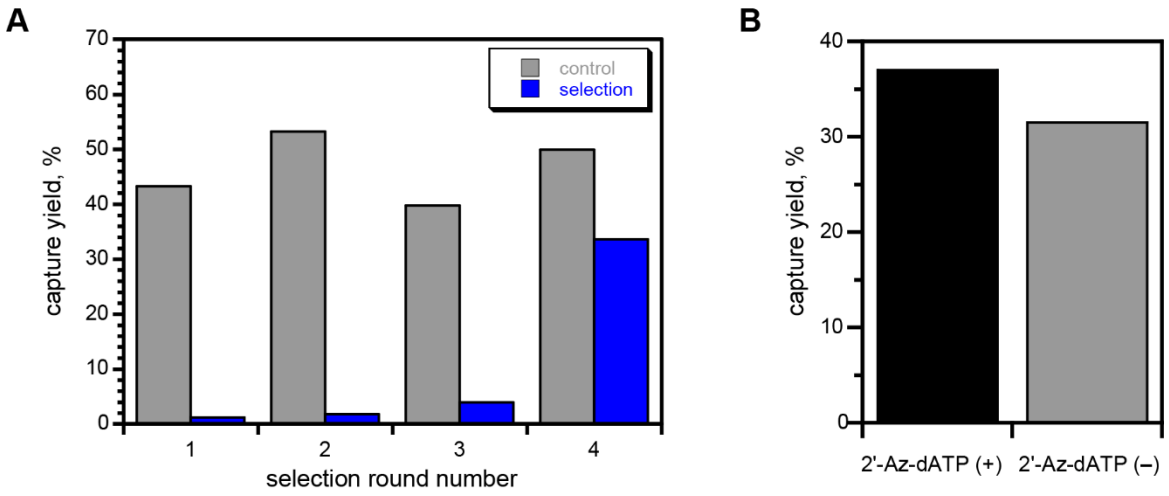


Figure 4.9. The AP2 selection experiment. (A) Progression of the selection experiment. In each round, "control" refers to the yield for the control capture reaction using an azido-modified substrate, and "selection" refers to the yield for the CuAAC capture. (B) Capture of the 4AP2 pool after the selection step where the pool was incubated with or without 2′-Az-dATP.

4.2.3 Selection Experiments with Protein Substrates Using Gel-Based Capture Method

An alternative capture method was required to enable selection experiments with protein substrates. The gel-based CuAAC capture reaction with PEG-alkyne was originally considered unsuitable, because the recovery of proteins from PAGE gels is generally low in yield. However, if the protein moiety is removed by reducing the disulfide bond via DTT treatment, the DNA moiety in the capture product should be extracted from the PAGE gel with high efficiency. To validate this process, the ligation products of both proteins to the N_{80} DNA pool were treated with *N*-hydroxysuccinimide carboxyl PEG that has an average molecular weight of 10000 (PEG_{10k}-NHS). Lysine residues on the proteins were randomly modified by PEG_{10k}-NHS, which provided model products of the CuAAC capture reactions. These model products were separated by 8% SDS-urea PAGE, in which the PAGE gel included each of 0.1% (w/v) SDS and 7 M urea. PEG molecules with larger molecular weight were used to induce a sufficiently large PAGE shift that enables the separation. When the excised gel pieces were incubated in the standard TEN buffer (10 mM Tris, pH 8.0, 300 mM NaCl, 1 mM EDTA) supplemented with 50 mM DTT, efficient recovery yields of the DNA moieties (>85%) were observed (Figure 4.10). When DTT was omitted

during the extraction, only ~30% of PEGylated DNA-protein conjugates were recovered from the gel pieces. This result showed that a gel-based CuAAC capture reaction with PEG_{10k}-alkyne will be viable for selection experiments with protein substrates when the capture product is extracted using the TEN buffer supplemented with DTT.

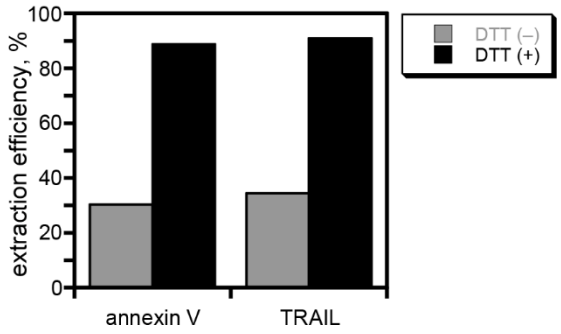


Figure 4.10. SDS-PAGE gel extraction efficiency. Gel pieces were incubated with 10 mM Tris, pH 8.0, 300 mM NaCl, 1 mM EDTA, with or without 50 mM DTT at 37 °C for 2 h. Scintillation counting of the extractant and the gel pieces was used to quantify extraction yield.

The CP2–CT2 selection experiments were performed similarly to the previous set of selection experiments AK2–AP2, except using a different capture step (Figure 4.11). After 14 rounds, no activity was observed in any of these selection experiments, which were discontinued. The negative outcome of these selection experiments suggested the difficulty for DNA pools to simultaneously adopt functions of both binding to the protein substrates and catalyzing their modification.

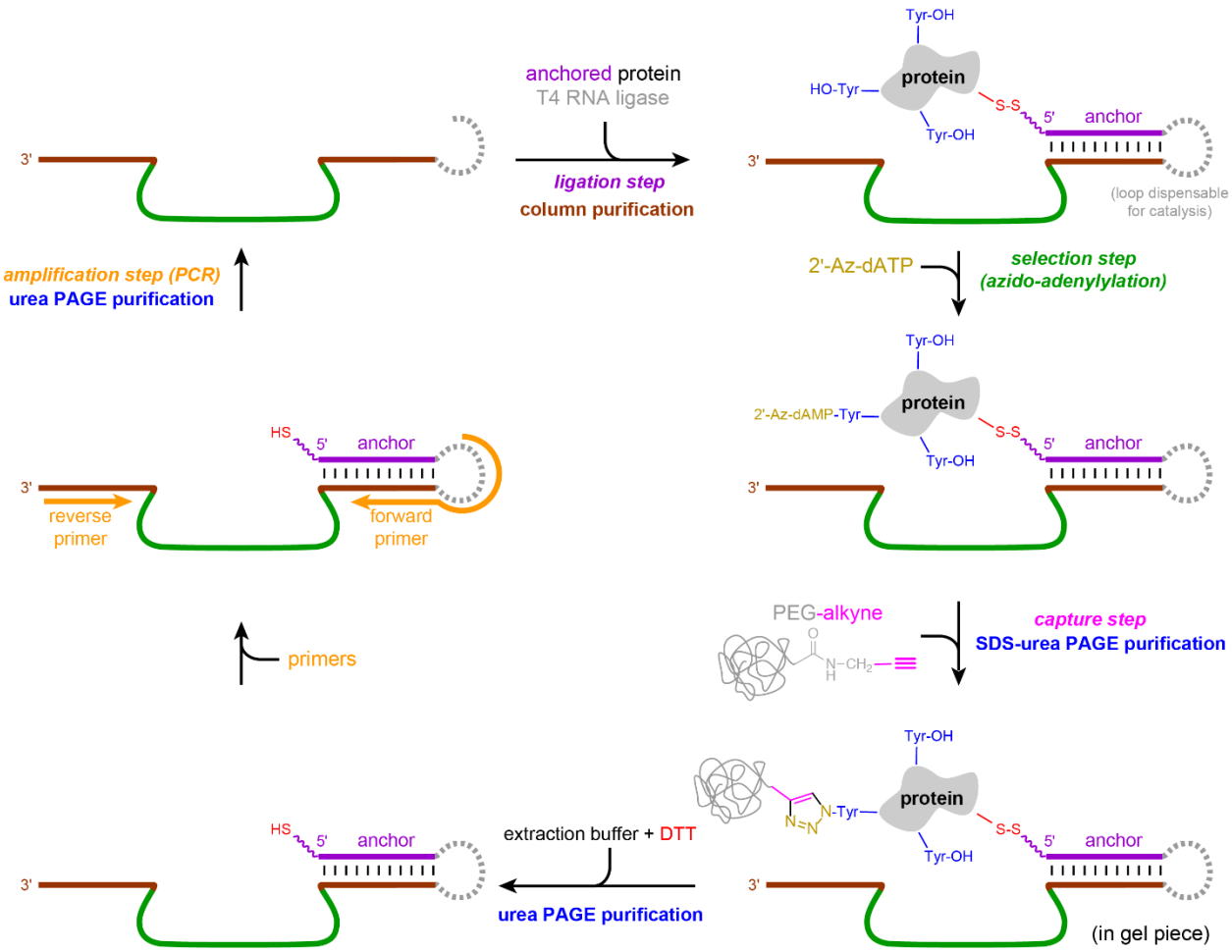


Figure 4.11. Diagram of in vitro selection procedure for deoxyribozymes with tyrosine azido-adenylylation activity on protein substrates using gel-based capture method.

We considered that a modular approach might facilitate the identification of deoxyribozymes with activity on protein substrates, by decoupling the binding and catalytic functions (Figure 4.12). In this modular approach, a predetermined aptamer domain engages in noncovalent interaction with the protein substrate, allowing a distinct enzyme domain, subsequently identified through in vitro selection in the presence of the aptamer domain, to be devoted to catalysis.

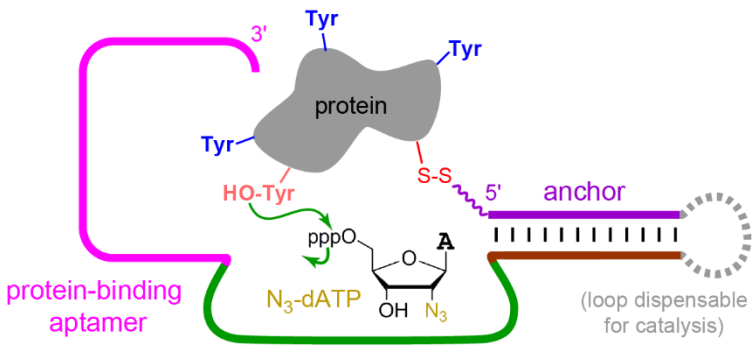


Figure 4.12. Schematic illustration of the modular approach to identify deoxyribozymes with tyrosine azido-adenylylation activity on protein substrates.

4.2.4 Selection Experiments with Unmodified DNA Pools for Aptamers that Bind to Annexin V and TRAIL

To perform selection experiments with predetermined aptamer domains, DNA aptamers that bind to the protein substrates must be first identified. Both annexin V and TRAIL were immobilized with Ni-NTA agarose via their His_6 tag to enable the aptamer selection experiments. The selection process began with a pre-selection step in which the DNA pools were incubated with unmodified Ni-NTA agarose which did not have any protein immobilized (Figure 4.13). The flow-through from the pre-selection step was incubated with the protein-bound Ni-NTA agarose. After incubation, the protein-bound resin was washed with the aptamer binding buffer to remove the unbound and inactive DNA sequences. Scintillation counting of the wash fractions and the protein-bound Ni-NTA agarose was used to quantify binding activity. The active DNA sequences that were bound to the protein was recovered by phenol/chloroform extraction followed by ethanol precipitation. PCR was performed to amplify the active DNA sequences that survived the previous selection step. The entire selection process was iterated until the aptamer sequences dominated the population.

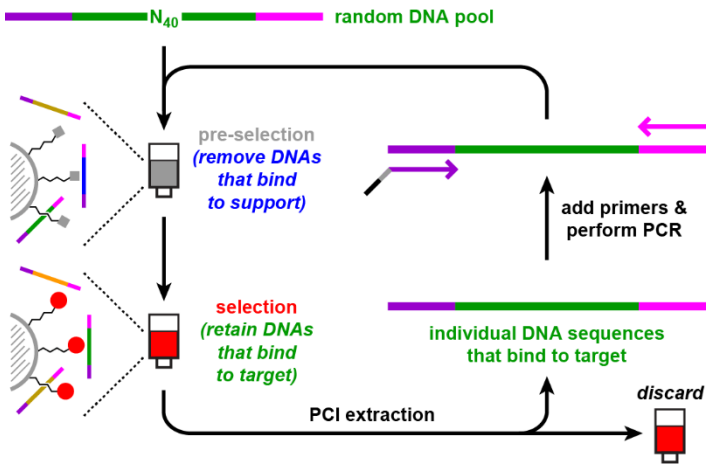


Figure 4.13. Schematic illustration of the in vitro selection strategy to identify DNA aptamers that bind to protein substrates.

The AQ2–AY2 selection experiments were performed by former Silverman lab staff Kevin Duffin to identify DNA aptamers that bind to annexin V and TRAIL. Different conditions were used for these selection experiments, exploring key variables like temperature, Mg^{2+} concentration, and pool/protein ratio (Table 4.1). However, after 16 rounds of selection, no binding activity above background level (~5%) was observed in any of the selection experiments, which were discontinued.

Table 4.1: Design of the AQ2–AY2 selection experiments.

selection experiment	target protein	temperature / °C	$[Mg^{2+}]$ / mM	pool/protein ratio	
				round 1–5	round 6–16
AQ2	annexin V	37	40	10:1	1:1
AR2	annexin V	25	40	10:1	1:1
AS2	annexin V	37	5	10:1	1:1
AT2	annexin V	37	40	1:1	1:4
AV2	TRAIL	37	40	10:1	1:1
AW2	TRAIL	25	40	10:1	1:1
AX2	TRAIL	37	5	10:1	1:1
AY2	TRAIL	37	40	1:1	1:4

4.2.5 Selection Experiments with Known Aptamers that Bind to Egg White Lysozyme

Given the negative outcome from the AQ2–AY2 selection experiments, the design of aptamer selection experiments must be validated with positive controls. Two lysozyme aptamers, Apta1 and Apta8, were used for these control selection experiments.[15] Egg white lysozyme was conjugated with the C-terminal carboxylic acid of the octapeptide HHHHHHAA (referred to as His_6Ala_2 peptide) via the amide coupling reaction to prepare the His_6-lysozyme. Apta1 and Apta8 each bound to His_6-lysozyme that was immobilized on Ni-NTA agarose with 23% and 29% yields, respectively (Figure 4.14). As the negative control, random DNA pool bound to His_6-lysozyme with ~5% background. Even though the binding yields of Apta1 and Apta8 to His_6-lysozyme were not 100%, they were still 3- to 5-fold higher than the background binding of random DNA pool, which was sufficient for the control selection experiments.

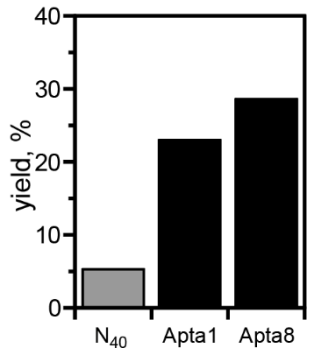


Figure 4.14. Binding of Apta1 and Apta8 aptamers to His_6-lysozyme immobilized on Ni-NTA agarose. N_{40} denotes random DNA pool that serves as the negative control. Incubation conditions: 25 mM Tris, pH 8.3, 5 mM K_2HPO_4, and 192 mM glycine at room temperature for 30 min.

Nine selection experiments, DE2–DM2, were performed with DNA pools that included Apta1 or Apta8 at different percentages (Table 4.2). The process for the DE2–DM2 selection experiments was similar to that for the AQ2–AY2 selection experiments, except that the reported aptamer binding conditions were used as the selection conditions and the pool/protein ratio was kept as 1:1. Binding activity with ~20% yield was observed in round 2 for the selection experiments initiated with DNA pools that contained 10% (DE2 and DJ2) and 1% (DF2 and DK2)

known aptamers (Figure 4.15). The binding activity of the DNA pools from these four selections in round 2 was comparable to that observed for Apta1 and Apta8, suggesting that the sequence of Apta1 or Apta8 dominated each selection pool. Thus, the enrichment factors for these four selection experiments were estimated as 10–100, based on the calculation that 1–10% active sequences were enriched to 100% after one round of enrichment. For the other selection experiments, 25–30% binding yields were observed in round 3. The enrichment factors were estimated as 32–100 (i.e., the square root of 1000–10000), assuming that the enrichment factor for each selection does not change among different rounds. The outcome of the DE2–DM2 selection experiments validated that the design of aptamer selection experiments can enrich DNA sequences binding to target proteins.

Table 4.2: Design of the DE2–DM2 selection experiments.

selection experiment	aptamer used	percentage of known aptamer in the initial pool
DE2	Apta1	10
DF2	Apta1	1
DG2	Apta1	0.1
DH2	Apta1	0.01
DJ2	Apta8	10
DK2	Apta8	1
DL2	Apta8	0.1
DM2	Apta8	0.01

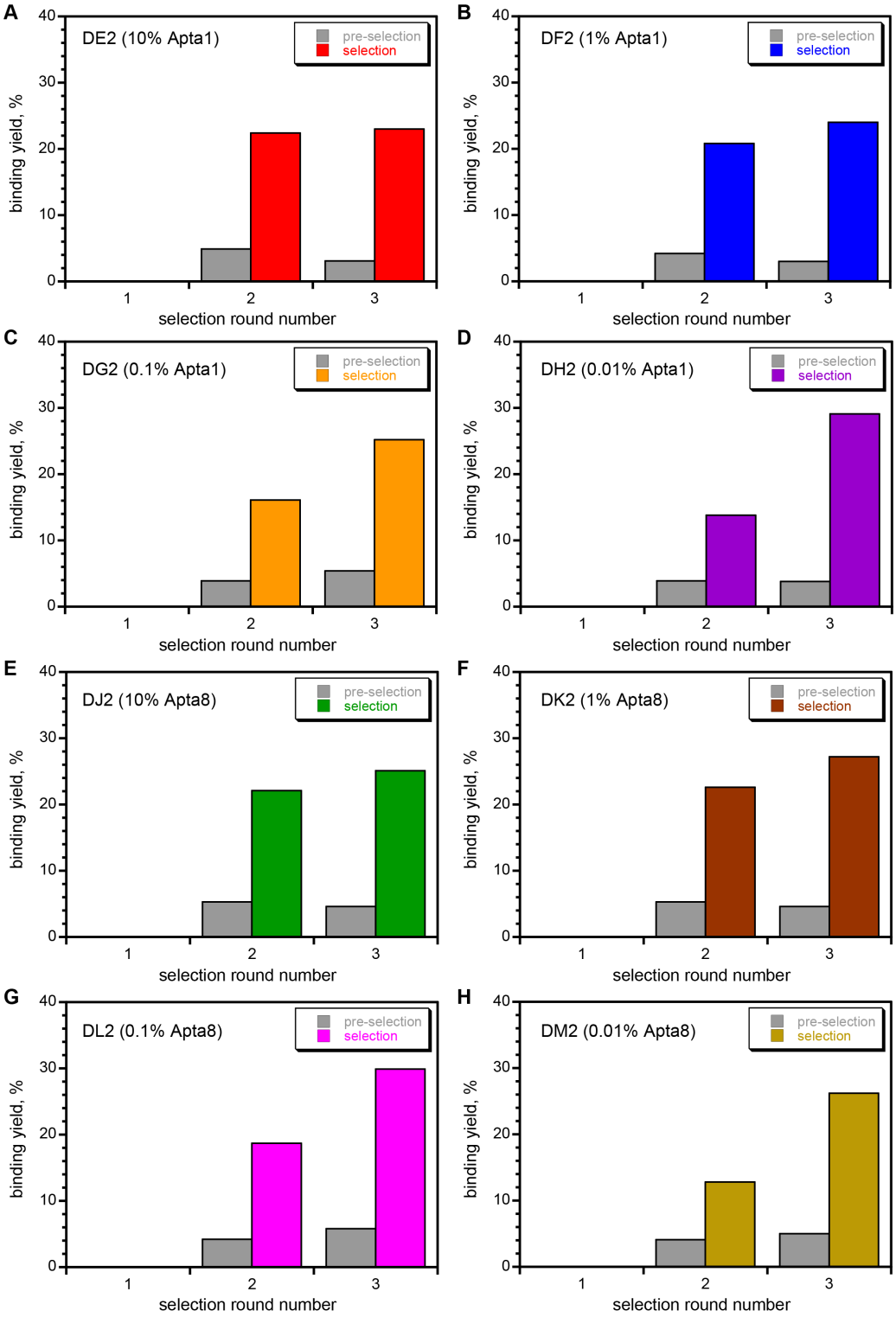


Figure 4.15. Progression of the in vitro selection experiments DE2–DM2 (panels A–H). In each round, "pre-selection" refers to the background binding observed in the pre-selection step using unmodified Ni-NTA agarose, and "selection" refers to the yield observed in the selection step using the Ni-NTA agarose-bound lysozyme for the indicated selection experiment. The indicated selection experiment and percentage of known aptamer in the initial pool was labeled in each panel.

4.2.6 Selection Experiments with Modified DNA Pools for Aptamers that Bind to Annexin V, TRAIL, and PP

Inspired by the success in identifying aptamers from DNA pools modified with hydrophobic moieties,[4-6] the EA2–EJ2 selection experiments were designed to incorporate benzyl, naphthyl, and indolyl groups via the click-SELEX strategy (Table 4.3).[7] Benzyl azide was purchased from the commercial vendor. The other two azido-containing compounds were synthesized by reacting the corresponding bromide compounds with sodium azide (Figure 4.16).[16] These selection experiments were performed similarly as the AQ2–AY2 selection experiments, except that additional primer extension and click steps were incorporated (Figure 4.17). The alkyne-modified DNA pool was generated by primer extension of the reverse template with dATP, dGTP, dCTP, and ^{E}dUTP. After purified by PAGE, the primer extension product was further modified with one of the three azide compounds by the CuAAC reaction. After the CuAAC reaction, the product was precipitated with ethanol, and a second click reaction was performed to ensure that all the alkyne groups in the DNA pool were modified.

Table 4.3: Design of the EA2–EJ2 selection experiments.

selection experiment	target protein	hydrophobic modification
EA2	annexin V	benzyl
EB2	annexin V	naphthyl
EC2	annexin V	indolyl
ED2	TRAIL	benzyl
EE2	TRAIL	naphthyl
EF2	TRAIL	indolyl
EG2	PP	benzyl
EH2	PP	naphthyl
EJ2	PP	indolyl

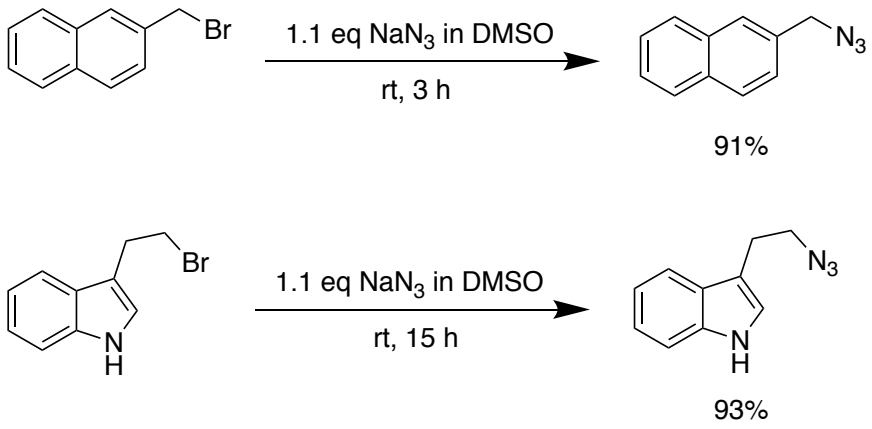


Figure 4.16. Scheme for the synthesis of 2-(azidomethyl)naphthalene and 3-(2-azidoethyl)indole.

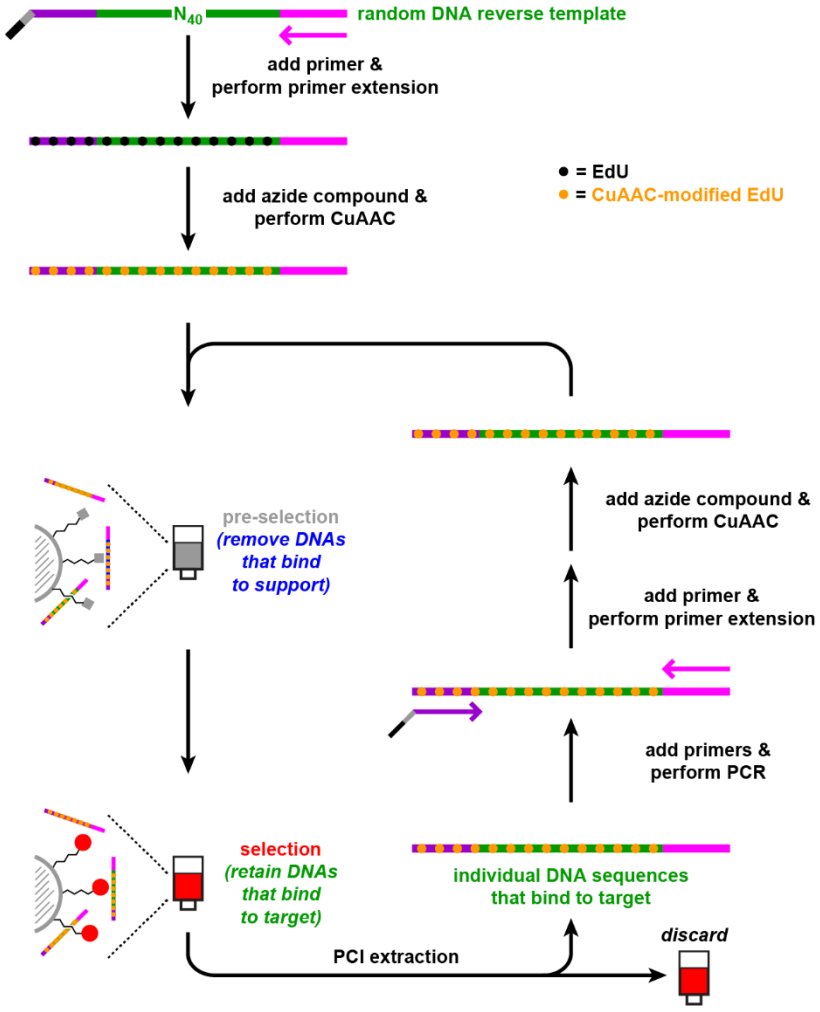


Figure 4.17. Schematic illustration of the in vitro selection strategy to identify hydrophobically modified DNA aptamers that bind to protein substrates.

The incubation conditions for the EA2–EJ2 selection experiments were 70 mM HEPES, pH 7.5, 5 mM $MgCl_2$, 150 mM NaCl, and 1 mM 2-mercaptoethanol. The pool/protein ratio was kept as 1:1 for all nine selection experiments. After 3 selection rounds, 31–71% background binding to unmodified Ni-NTA agarose was observed in the pre-selection step for the selection experiments with naphthyl and indolyl modifications (Figure 4.18). For the selection experiments with benzyl modification, 14–18% background binding was observed. The increase in background binding suggested the emergence of aptamer sequences that bind to unmodified Ni-NTA agarose, even though a pre-selection step was included during the selection experiments to remove those sequences.

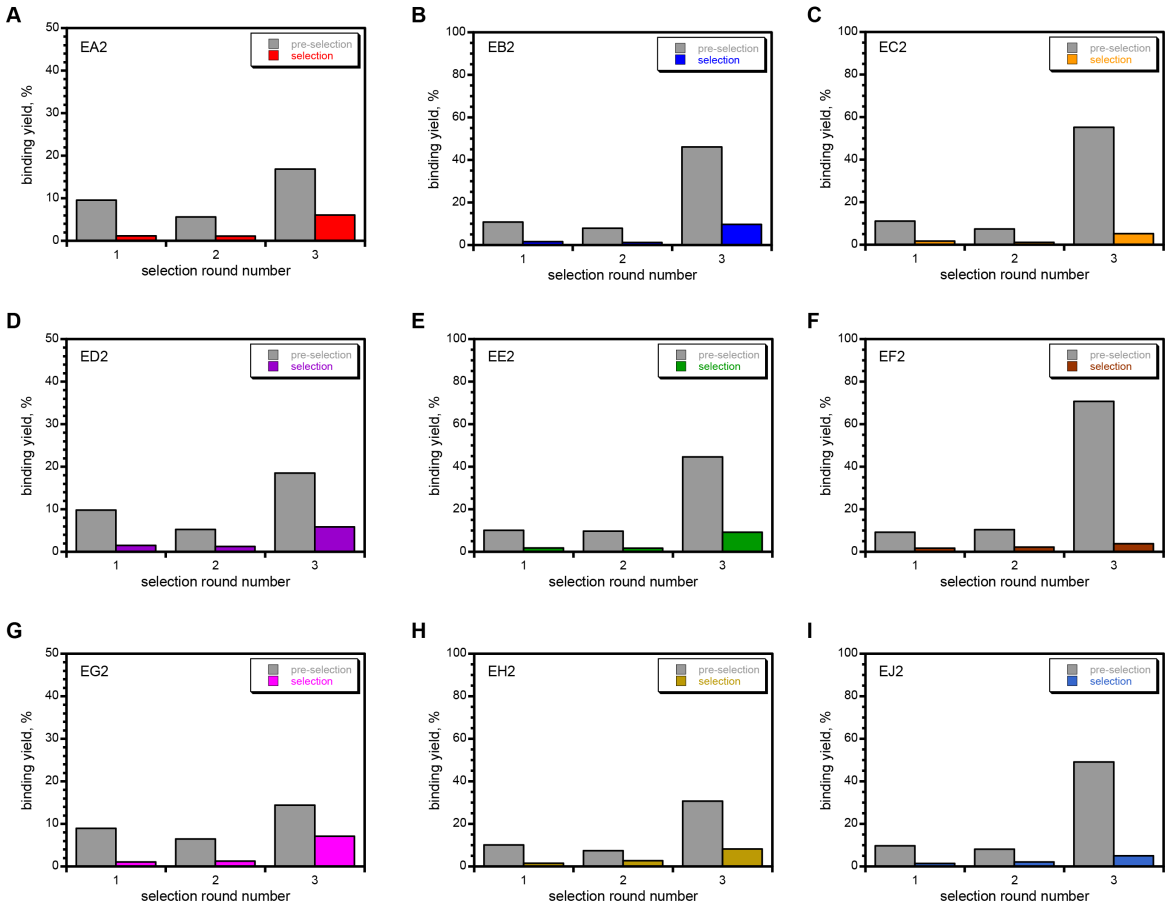


Figure 4.18. Progression of the in vitro selection experiments EA2–EJ2 (panels A–I). In each round, "pre-selection" refers to the background binding observed in the pre-selection step using unmodified Ni-NTA agarose, and "selection" refers to the yield observed in the selection step using Ni-NTA agarose-bound protein for the indicated selection experiment.

DNA aptamers modified with these hydrophobic groups have been identified from selection experiments where His_6-tagged target proteins were immobilized on magnetic beads.[4,6] Therefore, the FA2–FJ2 selection experiments were designed to use magnetic beads in place of Ni-NTA agarose (Table 4.4). The FA2–FJ2 selection experiments are currently ongoing.

Table 4.4: Design of the FA2–FJ2 selection experiments.

selection experiment	target protein	hydrophobic modification
FA2	annexin V	benzyl
FB2	annexin V	naphthyl
FC2	annexin V	indolyl
FD2	TRAIL	benzyl
FE2	TRAIL	naphthyl
FF2	TRAIL	indolyl
FG2	PP	benzyl
FH2	PP	naphthyl
FJ2	PP	indolyl

4.3 Summary and Future Directions

One of our long-term goals is DNA-catalyzed site-specific modification of protein substrates. To achieve such goal with known deoxyribozymes, two different recruiting strategies via hexahistidine (His_6) tags and DNA aptamers have been pursued, yet neither recruiting approach provided DNA-catalyzed protein modification.[2,3] Therefore, methods to identify de novo deoxyribozymes with azido-adenylylation activity on protein substrates are investigated in this chapter.

Two proteins, human annexin V and human TNF-related apoptosis-inducing ligand (TRAIL) 114–281, and a 36-mer peptide pancreatic polypeptide (PP) with an additional C-terminal cysteine were used as the substrates to evaluate various strategies for identifying deoxyribozymes. Both resin-based and gel-based capture methods were developed for selection experiments using protein substrates. For the selection experiments using resin-based capture methods, a

pre-selection step was intentionally omitted to avoid potential denaturation of the protein substrates, and aptamer sequences that bind to the resin used in the capture step quickly dominated the DNA pools. Thus, the selections using resin-based capture methods were discontinued. On the other hand, the gel-based capture method avoided the emergence of aptamer sequences. Therefore, selection experiments with protein substrates were performed using the gel-based capture method. Unfortunately, no activity was observed from these selection experiments after 14 rounds.

To achieve the identification of deoxyribozymes that are functional with proteins, a modular approach was designed to decouple the binding and catalytic functions required for DNA-catalyzed protein modification. The binding function is assigned to the predetermined aptamer domain, which is placed adjacent to the initially random enzyme domain. The sequence of the enzyme domain will be subsequently identified through in vitro selection in the presence of the aptamer domain. The aptamer domain is required to bind the protein substrate under conditions compatible with DNA catalysis. Therefore, aptamer selection experiments were performed to identify DNA sequences that bind to target proteins. No binding activity was observed in any of the selection experiments with unmodified DNA pools. Inspired by previous reports in which DNA aptamers with hydrophobic modifications bind to their target proteins with strong affinity and specificity,[4-6] ongoing selection experiments are seeking DNA aptamers that bind to annexin V, TRAIL, and PP, with benzyl, naphthyl, and indolyl modifications introduced via the click-SELEX strategy.[7]

4.4 Materials and Methods

4.4.1 Substrate Preparation Procedures

Oligonucleotides, peptides, and DNA-anchored peptide conjugates were prepared as described in Chapter 2. ^{E}dUTP was obtained from baseclick GmbH (Munich, Germany). Benzyl azide was obtained from Alfa Aesar (Tewksbury, MA). 2-(Bromomethyl)naphthalene, 3-(2-bromoethyl)indole, and lysozyme from chicken egg white were obtained from Sigma-Aldrich (St.

Louis, MO). Ni-NTA agarose and Dynabeads® His-Tag Isolation & Pulldown (referred to as magnetic beads) was obtained from Thermo Fisher (Waltham, MA).

Annexin V expression and purification. Human annexin V was expressed and purified by the previously reported procedures with minor modifications.[17] The bacterial expression plasmid pProEx.Htb.annexin V that encodes annexin V was provided by Prof. S.J. Martin (Trinity College, Dublin, Ireland). Using 5 μL of the plasmid, 50 μL of competent DH5α *E. coli* bacteria was transformed by the heat shock method. The seed culture for protein expression was prepared by inoculation of an individual *E. coli* colony in 3 mL of LB containing 100 μg/mL of ampicillin and incubation at 37 °C with shaking at 220 rpm overnight. The seed culture was diluted into 300 mL of LB containing 100 μg/mL of ampicillin. After the cells were grown at 37 °C with shaking at 220 rpm to an OD_{600} of 0.4–0.6, protein expression was induced by the addition of isopropyl-β-D-1-thiogalactoside (IPTG) to a final concentration of 500 μM. After another 3 h incubation at 37 °C with shaking at 220 rpm, cells were harvested by centrifugation at 7500×*g* and 4 °C. The bacterial pellet was resuspended in 5 mL of lysis buffer [50 mM Tris, pH 8.5, 10 mM 2-mercaptoethanol, 1 mM phenylmethanesulfonyl fluoride (PMSF), 2 μg/mL aprotinin, and 10 μg/mL leupeptin] and disrupted on ice for 10 min (alternating between 1 s pulses and 3 s pauses) using a Branson sonicator at 15% amplitude with a ⅛-inch tip. The supernatant from the bacterial lysate was recovered by centrifugation for 15 min at 15000×*g* and 4 °C and incubated with 250 μL of Ni-NTA agarose by nutating at 4 °C overnight. Unbound protein was removed from the Ni-NTA agarose by centrifuging at 1000×*g* and 4 °C for 1 min. The Ni-NTA agarose was washed with 3× 500 μL of ice-cold wash buffer [20 mM Tris, pH 8.5, 100 mM KCl, 10% (v/v) glycerol, 10 mM 2-mercaptoethanol, and 15 mM imidazole]. Annexin V was eluted from the Ni-NTA agarose with 3× 500 μL of ice-cold elution buffer [20 mM Tris, pH 7.5, 100 mM KCl, 10% (v/v) glycerol, 10 mM 2-mercaptoethanol, 1 mM PMSF, 2 μg/mL aprotinin, 10 μg/mL leupeptin, and 150 mM imidazole]. Protein concentration was quantified by Bradford protein assay.[18] From 300 mL of cell culture, 9.0 mg of annexin V was obtained. Therefore, the overall protein production yield was 30 mg/L.

TRAIL expression and purification. Human TNF-related apoptosis-inducing ligand (TRAIL) 114–281 was expressed and purified by the previously reported procedures with minor modifications.[19] The bacterial expression plasmid pETdwHisTRAIL that encodes soluble TRAIL (amino acids 114–281 of the genomic sequence) was provided by Dr. D.W. Seol (Chung-Aung University, Seoul, South Korea). Using 5 µL of the plasmid, 50 µL of competent Rosetta (DE3) *E. coli* bacteria was transformed by the heat shock method. The seed culture for protein expression was prepared by inoculation of an individual *E. coli* colony in 3 mL of LB containing 100 µg/mL of ampicillin and incubation at 37 °C with shaking at 220 rpm overnight. The seed culture was diluted into 300 mL of LB containing 100 µg/mL of ampicillin. After the cells were grown at 37 °C with shaking at 220 rpm to an OD_{600} of 0.4–0.6, protein expression was induced by the addition of IPTG to a final concentration of 500 µM. After an additional overnight incubation at 30 °C with shaking at 220 rpm, cells were harvested by centrifugation at 7500×*g* and 4 °C. The bacterial pellet was resuspended in 5 mL of lysis buffer (50 mM sodium phosphate, pH 8.5, 10 mM 2-mercaptoethanol, 1 mM PMSF, 2 µg/mL aprotinin, and 10 µg/mL leupeptin) and disrupted on ice for 10 min (alternating between 1 s pulses and 3 s pauses) using a Branson sonicator at 15% amplitude with a ⅛-inch tip. The supernatant from the bacterial lysate was recovered by centrifugation for 15 min at 15000×*g* and 4 °C and incubated with 250 µL of Ni-NTA agarose by nutating at 4 °C overnight. Unbound protein was removed from the Ni-NTA agarose by centrifuging at 1000×*g* and 4 °C for 1 min. The Ni-NTA agarose was washed with 3× 500 µL of ice-cold wash buffer [20 mM sodium phosphate, pH 8.5, 100 mM KCl, 10% (v/v) glycerol, 10 mM 2-mercaptoethanol, and 30 mM imidazole]. TRAIL protein was eluted from the Ni-NTA agarose with 3× 500 µL of ice-cold elution buffer [20 mM sodium phosphate, pH 7.5, 100 mM KCl, 10% (v/v) glycerol, 10 mM 2-mercaptoethanol, 1 mM PMSF, 2 µg/mL aprotinin, 10 µg/mL leupeptin, and 500 mM imidazole]. Protein concentration was quantified by Bradford protein assay. From 300 mL of cell culture, 3.6 mg of TRAIL was obtained. Therefore, the overall protein production yield was 12 mg/L.

The differences between the preparation procedures for annexin V and TRAIL were as follows. Annexin V was expressed in competent DH5α *E. coli* bacteria. TRAIL was expressed in competent Rosetta (DE3) *E. coli* bacteria. After the induction of protein expression by IPTG, the cell culture for annexin V expression was shaken at 220 rpm and 37 °C overnight. After the induction of protein expression by IPTG, the cell culture for TRAIL expression was shaken at 220 rpm and 30 °C overnight. Tris was used in all the buffers for annexin V purification. Sodium phosphate was used in all the buffers for TRAIL purification. For annexin V purification, the wash and elution buffers contained, respectively, 15 mM and 150 mM of imidazole. For TRAIL purification, the wash and elution buffers contained, respectively, 30 mM and 500 mM of imidazole.

Synthesis of DNA-anchored protein conjugates. DNA-anchored protein conjugates were synthesized by disulfide formation between a DNA HEG-tethered 3′-thiol [HEG = hexa(ethylene glycol)] and the cysteine side chain of the protein (Cys316 for annexin V or Cys230 for TRAIL), similar to the synthetic procedure for DNA-anchored peptide conjugates. The 5′-thiol DNA anchor oligonucleotide was 5′-HO-C_6-SS-C_6-p-HEG-X-3′, where X represents the specific oligonucleotide sequence (Table 4.5). The 5′-disulfide linker was introduced via standard solid-phase DNA synthesis and unmasked to a thiol by DTT treatment. A 50 μL sample containing 5 nmol of DNA anchor oligonucleotide in 100 mM HEPES, pH 7.5, and 50 mM DTT was incubated at 37 °C for 2 h. The reduced product was precipitated to remove DTT by addition of 50 μL of water, 10 μL of 3 M NaCl, and 300 μL of ethanol. The precipitated product (HS-C_6-p-HEG-DNA) was dissolved in 35 μL of water and 10 μL of 100 mM triethylammonium acetate, pH 7.0. Activation as the pyridyl disulfide was achieved by adding 5 μL of 100 mM 2,2′-dipyridyl disulfide in DMF and incubating at 37 °C for 2 h. The product (PySS-C_6-p-HEG-DNA) was precipitated by addition of 50 μL of water, 10 μL of 3 M NaCl and 300 μL of ethanol and dissolved in 35 μL of water. Meanwhile, 6 nmol of protein that was expressed and purified with Ni-NTA agarose was buffer-exchanged into 10 mM NaCl using a Micro Bio-Spin P-6 column (Bio-Rad). Conjugation to the protein was performed by combining the activated DNA substrate and the protein into 40

μL total volume containing 20 mM triethylammonium acetate, pH 7.0. The sample was incubated at 37 °C for 12 h, and the DNA-anchored protein was purified by immobilized metal ion affinity chromatography (IMAC) using Ni-NTA agarose as described above in the protein purification section. The elution was buffer-exchanged into 1× annealing buffer (5 mM Tris, pH 7.5, and 15 mM NaCl) using a Micro Bio-Spin P-6 column (Bio-Rad) and sample concentration was quantified by UV absorbance (A_{260}). The overall yield is 10–20%, providing 0.5–1 nmol of the DNA-protein conjugates.

Immobilization of His_6-tagged substrates on Ni-NTA agarose. Low-retention microcentrifuge tubes were treated by allowing to stand overnight with 2 mL of 5% (w/v) PEG_{3350} (Sigma-Aldrich, cat. no. P3640) and rinsing with 3× 1 mL of water. A 100 μL sample containing 30 nmol of His_6-tagged peptide or protein substrate and 50 μL of Ni-NTA agarose was incubated in 20 mM sodium phosphate, pH 8.5, and 100 mM KCl by shaking on a vortexer at the lowest setting overnight at 4 °C. The Ni-NTA agarose was washed with 3× 50 μL of aptamer binding buffer (70 mM HEPES, pH 7.5, 150 mM NaCl, and 1 mM 2-mercaptoethanol), and stored in 50 μL of aptamer binding buffer. Immobilization efficiency was quantified by Bradford protein assay as ~60%.

Synthesis of His_6-lysozyme. A sample containing 10 nmol of chicken egg white lysozyme was brought to 50 μL total volume containing 100 mM MES, pH 6.0, 1 mM His_6Ala_2 peptide, 1 mM *N*-hydroxysuccinimide (NHS), 100 mM *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide (EDC), and 10% (v/v) DMF. The sample was incubated at room temperature for 12 h and buffer-exchanged into 10 mM NaCl using a Micro Bio-Spin P-6 column.

Synthesis of 2-(azidomethyl)naphthalene. The synthesis of 2-(azidomethyl)naphthalene was performed by the previously reported procedures with minor modifications.[16] To a 0.5 M solution of NaN_3 (1.1 equiv) in 1 mL of DMSO was added 111 mg of 2-(bromomethyl)naphthalene (0.5 mmol, 1.0 equiv), and the clear colorless solution was stirred at room temperature for 3 h. The mixture was quenched with 4 mL of water and extracted with 3× 5 mL of ethyl acetate. The organic layer was separated, washed with 3× 10 mL of brine, and dried over Na_2SO_4. After filtration and

concentration in vacuo, the product was obtained as a white solid (83 mg, 91%). ^{1}H NMR ($CDCl_3$, 500 MHz, δ; ppm): 7.88–7.80 (3H, m), 7.76 (1H, s), 7.52–7.48 (2H, m), 7.42 (1H, dd), 4.49 (2H, s). EI-MS: *m*/z calcd. for $C_{11}H_9N_3$ $[M]^+$ 183.1; found 183.0.

Synthesis of 3-(2-azidoethyl)indole. The synthesis of 3-(2-azidoethyl)indole was performed by the previously reported procedures with minor modifications.[16] To a 0.5 M solution of NaN_3 (1.1 equiv) in 1 mL of DMSO was added 112 mg of 3-(2-bromoethyl)indole (0.5 mmol, 1.0 equiv), and the clear yellow solution was stirred at room temperature for 15 h. The mixture was then quenched with 4 mL of water and extracted with 3× 5 mL of ethyl acetate. The organic layer was separated, washed with 3× 10 mL of brine, and dried over Na_2SO_4. After filtration and concentration in vacuo, the product was obtained as yellow oil (87 mg, 93%). ^{1}H NMR ($CDCl_3$, 500 MHz, δ; ppm): 7.95 (1H, s), 7.53 (1H, d), 7.31 (1H, d), 7.22 (1H, m), 7.08 (1H, m), 7.01 (1H, s), 3.51 (2H, t), 3.01 (2H, d). EI-MS: *m*/z calcd. for $C_{10}H_{10}N_4$ $[M]^+$ 186.1; found 186.1.

Immobilization of His_6-tagged substrates on the magnetic beads. The immobilization of His_6-tagged substrates with the magnetic beads was performed by the previously reported procedures.[4] Unless indicated, all washes were performed by resuspending the magnetic beads in the appropriate solution, mixing for 30 s, separating the beads with a magnet, and removing the supernatant. In a microcentrifuge tube, 125 μL of the magnetic beads was washed with 3× 500 μL of BW buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, and 0.01% Tween-20) and mixed with 0.1 mg of His_6-tagged substrate in 1 mL of BW buffer. The mixture was nutated at room temperature for 30 min and stored at 4 °C until use.

Table 4.5: Oligonucleotide sequences used in the selection experiments. All sequences are written 5′ to 3′. Lowercase letters represent ribonucleotides. p represents a phosphoryl group. For all selections, the reverse PCR primer was (AAC)$_4$XCCATCAGGATCAGCT, where X denotes the hexa(ethylene glycol) [HEG] spacer to stop Taq polymerase.

oligonucleotide purpose	oligonucleotide sequence
	AK2 and CP2 selection experiments
(annexin V)-HEG-DNA substrate	(annexin V)–HEG–GAAGAGATGGCGACgga
forward primer for selection	p-CGAAGTCGCCATCTCTTC
random pool for selection	p-CGAAGTCGCCATCTCTTC-N_{40}-ATAGTGAGTCGTATTAAGCTGATCCTGATGG
	AL2 and CQ2 selection experiments
(annexin V)-HEG-DNA substrate	(annexin V)–HEG–GAAATAGCCATCATgga
forward primer for selection	p-CGAAATGATGGCTATTTC
random pool for selection	p-CGAAATGATGGCTATTTC-N_{80}-ATATGTCTTTCAATAGAGCTGATCCTGATGG
	AM2 and CR2 selection experiments
TRAIL-HEG-DNA substrate	TRAIL–HEG–GAAATAGCCATCATgga
forward primer for selection	p-CGAAATGATGGCTATTTC
random pool for selection	p-CGAAATGATGGCTATTTC-N_{40}-ATATGTCTTTCAATAGAGCTGATCCTGATGG
	AN2 and CS2 selection experiments
TRAIL-HEG-DNA substrate	TRAIL–HEG–GAAGAGATGGCGACgga
forward primer for selection	p-CGAAGTCGCCATCTCTTC
random pool for selection	p-CGAAGTCGCCATCTCTTC-N_{80}-ATAGTGAGTCGTATTAAGCTGATCCTGATGG
	AP2 and CT2 selection experiments
PP-HEG-DNA substrate	APLEPVYPGDNATPEQMAQYAADLRRYINMLTRPRYC–HEG–GAATTCAGTCTTAAgga
forward primer for selection	p-CGAATTAAGACTGAATTC
random pool for selection	p-CGAAGTATAAACCTGTTC-N_{40}-ATAAGAAACGAGATATAGCTGATCCTGATGG
	AQ2 and AV2 selection experiments
forward primer for selection	CGAAGTCGCCATCTCTTC
random pool for selection	CGAAGTCGCCATCTCTTC-N_{30}-ATAGTGAGTCGTATTAAGCTGATCCTGATGG
	AR2 and AW2 selection experiments
forward primer for selection	CGAACGAAAGCCTCCTTC
random pool for selection	CGAACGAAAGCCTCCTTC-N_{30}-ATACGCATAAAGGTAGAGCTGATCCTGATGG
	AS2 and AX2 selection experiments
forward primer for selection	CGAAATGATGGCTATTTC
random pool for selection	CGAAATGATGGCTATTTC-N_{30}-ATATGTCTTTCAATAGAGCTGATCCTGATGG

Table 4.5 (cont.)

oligonucleotide purpose	oligonucleotide sequence
	AT2 and AY2 selection experiments
forward primer for selection	CGAAATAGATTATCATTC
random pool for selection	CGAAATAGATTATCATTC-N_{30}-ATAATTAGTAACCTGAAGCTGATCCTGATGG
	DE2–DM2 selection experiments
forward primer for selection	CGAACAACGCTAAAATTC
random pool for selection	CGAACAACGCTAAAATTC-N_{30}-ATAGGAGACGGGCAACAGCTGATCCTGATGG
lysozyme aptamer Apta1	CGAACAACGCTAAAATTCGCAGCTAAGCAGGCGGCTCACAAAACCATTCGCATGCGGC-ATAGGAGACGGGCAACAGCTGATCCTGATGG
lysozyme aptamer Apta8	CGAACAACGCTAAAATTCGCACCTTGATGACATGATAGTCGTTGTGTATGCAGTTGGC-ATAGGAGACGGGCAACAGCTGATCCTGATGG
	EA2–EJ2 and FA2–FJ2 selection experiments
forward primer for selection	CGAATTAAGACTGAATTC
random pool for selection	CGAAGTATAAACCTGTTC-N_{40}-ATAAGAAACGAGATATAGCTGATCCTGATGG

4.4.2 DNA Enzyme In Vitro Selection Procedures

4.4.2.1 In Vitro Selection Procedures for the AK2–AP2 Selection Experiments

Procedure for PCR was performed as described in Chapter 2.

Procedure for ligation step in round 1. A 21 μL sample containing 200 pmol of 5′-phosphorylated DNA pool and 220 pmol of DNA-anchored protein or peptide substrate was cooled on ice for 5 min in 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA. To this solution was added 6 μL of 5× T4 RNA ligase buffer that lacks DTT (250 mM Tris, pH 7.5, 25 mM $MgCl_2$, and 0.25 mM ATP) and 3 μL of 10 U/μL T4 RNA ligase (Thermo Fisher). The sample was incubated at 37 °C for 12 h and buffer-exchanged into 1× annealing buffer (5 mM Tris, pH 7.5, and 15 mM NaCl) using a Micro Bio-Spin P-6 column (Bio-Rad).

Procedure for ligation step in subsequent rounds. A 15 μL sample containing the PCR-amplified 5′-phosphorylated DNA pool (~5–10 pmol) and 50 pmol of DNA-anchored protein or peptide substrate was cooled on ice for 5 min in 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA. To this solution was added 4 μL of 5× T4 RNA ligase buffer that lacks DTT (250 mM

Tris, pH 7.5, 25 mM $MgCl_2$, and 0.25 mM ATP) and 1 μL of 10 U/μL T4 RNA ligase (Thermo Fisher). The sample was incubated at 37 °C for 2 h and buffer-exchanged into 1× annealing buffer (5 mM Tris, pH 7.5, and 15 mM NaCl) using a Micro Bio-Spin P-6 column (Bio-Rad).

Procedure for selection step in round 1. A 16 μL sample containing ligated pool was cooled on ice for 5 min in 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA. The selection reaction was initiated by bringing the sample to 40 μL total volume containing 70 mM HEPES, pH 7.5, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$, 150 mM NaCl, and 100 μM 2′-Az-dATP. The Mn^{2+} was added from a 10× stock solution containing 200 mM $MnCl_2$. The Zn^{2+} was added from a 10× stock solution containing 10 mM $ZnCl_2$, 20 mM HNO_3, and 200 mM HEPES at pH 7.5; this stock solution was freshly prepared from a 100× stock of 100 mM $ZnCl_2$ in 200 mM HNO_3. The metal ion stocks were added last to the final sample. The sample was incubated at 37 °C for 14 h and precipitated with ethanol.

Procedure for selection step in subsequent round. An 8 μL sample containing the ligated pool was cooled on ice for 5 min in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA. The selection reaction was initiated by bringing the sample to 20 μL total volume containing 70 mM HEPES, pH 7.5, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$, 150 mM NaCl, and 100 μM 2′-Az-dATP. The sample was incubated at 37 °C for 14 h and precipitated with ethanol.

Procedure for capture step in each round. Low-retention microcentrifuge tubes were treated by allowing to stand overnight with 2 mL of 5% (w/v) PEG_{3350} (Sigma-Aldrich, cat. no. P3640) and rinsing with 3 × 1 mL of water. All centrifugation steps were performed at 1000×*g* for 1 min. The sample from the selection step was dissolved in a PEG-treated tube to 10 μL total volume containing 100 mM HEPES, pH 7.5, 2 mg 3′-alkyne-modifier serinol CPG (Glen Research, 42 μmol/g), 40 mM tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), 10 mM sodium ascorbate, 5 mM $CuSO_4$, and 0.5 μg herring sperm DNA (Sigma-Aldrich). The concentrations of THPTA, sodium ascorbate, and $CuSO_4$ were optimized (not shown). The three reagents were added from 1 M (25×), freshly prepared 100 mM (10×), and 100 mM (20×) stock solutions, respectively; the three reagents were mixed together and then added at once to the other components. The sample

was shaken on a vortexer at the lowest setting for 1 h at 4 °C, transferred to a spin column, centrifuged, washed with 3× 100 µL of 50 mM Na_3EDTA, pH 8.0 and 8 M urea, and further washed with 3× 100 µL of water. The washes were combined into a single wash fraction. The bound sequences were eluted with 20 µL of 100 mM HEPES, pH 7.5, and 50 mM DTT at 37 °C for 1 h. The supernatant in the sample was transferred to a scintillation vial (elution fraction). Scintillation counting of the wash and elution fractions was used to quantify selection activity. The elution fraction was precipitated with ethanol and taken onward to the PCR step of the in vitro selection procedure.

4.4.2.2 In Vitro Selection Procedures for the CP2–CT2 Selection Experiments

Procedure for PCR was performed as described in Chapter 2.

Procedures for ligation and selection were identical to those in the selections with protein substrates using the resin-based capture method.

Procedure for capture step in each round. The sample from the selection step was brought to 10 µL total volume containing 100 mM HEPES, pH 7.5, 10 mM PEG_{10k}-alkyne, 40 mM tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), 10 mM sodium ascorbate, and 5 mM $CuSO_4$. The concentrations of PEG_{10k}-alkyne, THPTA, sodium ascorbate, and $CuSO_4$ were optimized (not shown). The four reagents were added from 100 mM (10×), 1 M (25×), freshly prepared 100 mM (10×), and 100 mM (20×) stock solutions, respectively; the three reagents were mixed together and then added at once to the other components. The sample was incubated at 4 °C for 1 h and separated by 8% SDS-urea PAGE, in which the PAGE gel included each of 0.1% (w/v) SDS and 7 M urea. The product was extracted from the gel in 10 mM Tris, pH 8.0, 300 mM NaCl, 1 mM EDTA, and 50 mM DTT. The extracted product was precipitated with ethanol and purified again by 8% PAGE.

4.4.3 DNA Aptamer In Vitro Selection Procedures

4.4.3.1 In Vitro Selection Procedures for the AQ2–AY2 Selection Experiments

Procedure for PCR was performed as described in Chapter 2.

Procedure for selection step in round 1. Each selection experiment was initiated with 200 pmol of the DNA pool. Low-retention microcentrifuge tubes were treated by allowing to stand overnight with 2 mL of 5% (w/v) PEG_{3350} (Sigma-Aldrich, cat. no. P3640) and rinsing with 3× 1 mL of water. All centrifugation steps were performed at 1000×*g* for 1 min. A 10 μL sample containing 200 pmol of pool in a PEG-treated tube 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. A pre-selection step was performed by bringing the sample to 50 μL total volume containing 70 mM HEPES, pH 7.5, 5 or 40 mM $MgCl_2$, 150 mM NaCl, 1 mM 2-mercaptoethanol, 1 μg herring sperm DNA, and 10 μL of unmodified Ni-NTA agarose. The sample was shaken on a vortexer at the lowest setting for 1 h at 37 °C or room temperature, transferred to a spin column, and centrifuged. The spin column used in the pre-selection step was transferred to a scintillation vial (sample A). The retained flow-through from the pre-selection step was brought to 100 μL total volume containing 70 mM HEPES, pH 7.5, 5 or 40 mM $MgCl_2$, 150 mM NaCl, 1 mM 2-mercaptoethanol, 2 μg herring sperm DNA, and 20 or 200 pmol of target protein immobilized on Ni-NTA agarose. The selection incubation was performed under identical conditions as those of the pre-selection step. After incubation, the sample was transferred to a spin column, centrifuged, washed with 3× 100 μL of aptamer binding buffer (70 mM HEPES, pH 7.5, 150 mM NaCl, and 1 mM 2-mercaptoethanol). The washes were combined into a single wash fraction (sample B). The spin column used in the selection step was transferred to a scintillation vial (sample C). Scintillation counting of the three samples A–C was used to quantify background binding and selection activity. After counting, the filter of the spin column in sample C was transferred to a low-retention microcentrifuge tube that contained 100 μL of water. DNA sequences that bound to target peptide or protein were recovered from the filter by phenol/chloroform extraction followed by ethanol precipitation.

Procedure for selection step in rounds 2–5. The procedure was identical to the selection step in round 1, except that 0.5 μL of unmodified Ni-NTA agarose was used in the pre-selection

step and 1 or 10 pmol of target protein immobilized on Ni-NTA agarose was used in the selection step.

Procedure for selection step in rounds 6–16. The procedure was identical to the selection step in round 1, except that 2 μL of unmodified Ni-NTA agarose was used in the pre-selection step and 10 or 40 pmol of target protein immobilized on Ni-NTA agarose was used in the selection step.

4.4.3.2 In Vitro Selection Procedures for the DE2–DM2 Selection Experiments

Procedure for PCR was performed as described in Chapter 2.

Procedure for selection step in round 1. Each selection experiment was initiated with 200 pmol of DNA sequence population that is composed of random pool and reported lysozyme aptamer at the intended ratio (Table 4.2). Low-retention microcentrifuge tubes were treated by allowing to stand overnight with 2 mL of 5% (w/v) PEG_{3350} (Sigma-Aldrich, cat. no. P3640) and rinsing with 3× 1 mL of water. All centrifugation steps were performed at 1000×*g* for 1 min. A 10 μL sample containing the mixture of random DNA pool and Apta1 or Apta8 aptamer sequence in a PEG-treated tube was annealed in 25 mM Tris, pH 8.3, 5 mM K_2HPO_4, and 192 mM glycine by heating at 95 °C for 3 min and cooling on ice for 5 min. A pre-selection step was performed by bringing the sample to 50 μL total volume containing 25 mM Tris, pH 8.3, 5 mM K_2HPO_4, 192 mM glycine, 1 μg herring sperm DNA, and 10 μL of unmodified Ni-NTA agarose. The sample was shaken on a vortexer at the lowest setting for 30 min at room temperature, transferred to a spin column, and centrifuged. The spin column used in the pre-selection step was transferred to a scintillation vial (sample A). The retained flow-through from pre-selection step was brought to 100 μL total volume containing 25 mM Tris, pH 8.3, 5 mM K_2HPO_4, 192 mM glycine, 2 μg herring sperm DNA, and with 200 pmol of His_6-lysozyme immobilized on Ni-NTA agarose. The selection incubation was performed under identical conditions as those of the pre-selection step. After incubation, the sample was transferred to a spin column, centrifuged, washed with 3× 100 μL of washing solution that contained 25 mM Tris, pH 8.3, 5 mM K_2HPO_4, 192 mM glycine. The washes

were combined into a single wash fraction (sample B). The spin column used in the selection step was transferred to a scintillation vial (sample C). Scintillation counting of the three samples A–C was used to quantify background binding and selection activity. After counting, the filter of the spin column in sample C was transferred to a low-retention microcentrifuge tube that contained 100 μL of water. DNA sequences that bound to lysozyme were recovered by phenol/chloroform extraction followed by ethanol precipitation.

Procedure for selection step in subsequent rounds. The procedure was identical to selection step in round 1, except that 0.5 μL of unmodified Ni-NTA agarose was used in the pre-selection step and 10 pmol of His_6-lysozyme immobilized on Ni-NTA agarose was used in the selection step.

4.4.3.3 In Vitro Selection Procedures for the EA2–EJ2 Selection Experiments

Procedure for selection was performed as described above for the AQ2–AY2 selection experiments, except that the ratio of immobilized target protein or peptide to DNA pool was kept as 1:1 throughout the selection rounds.

Primer extension to generate ^{E}dU-modified pools for round 1. To generate the initial 25% ^{E}dU-modified pools by primer extension, a reverse complement of the pool containing 25% A was synthesized with a 3′-tail to enable separation from the product after primer extension. A 25 μL sample was prepared containing 250 pmol of reverse-complement template, 300 pmol of forward primer, 7.5 nmol each of dATP, dGTP, dCTP, and ^{E}dUTP, 2.5 μL of 10× KOD XL polymerase buffer [1.2 M Tris-HCl, pH 8.0, 60 mM $(NH_4)_2SO_4$, 100 mM KCl, 1% Triton X-100, 0.01% BSA], and 2 μL of KOD XL polymerase. To generate the initial pool, 8 × 25 μL reactions were performed. Primer extension was performed in a PCR thermocycler according to the following program: 94 °C for 2 min, 47 °C for 2 min, 72 °C for 1 h. The 8 × 25 μL samples were combined, precipitated with ethanol, and separated by 8% PAGE.

Primer extension to generate ^{E}dU-modified pools for subsequent rounds. A 25 μL sample was prepared containing the reverse-complement single strand, 50 pmol of forward primer, 7.5

nmol each of dATP, dGTP, dCTP, and ^{E}dUTP, 10 μCi of α-^{32}P-dCTP (3000 Ci/mmol), 2.5 μL of 10× KOD XL polymerase buffer [1.2 M Tris-HCl, pH 8.0, 60 mM $(NH_4)_2SO_4$, 100 mM KCl, 1% Triton X-100, 0.01% BSA], and 0.5 μL of KOD XL polymerase. Primer extension was performed in a PCR thermocycler according to the following program: 94 °C for 2 min, 47 °C for 2 min, 72 °C for 1 h. The sample was separated by 8% PAGE.

Procedure for click reaction of ^{E}dU-modified pools. The sample from the primer extension step was brought to 20 μL total volume containing 100 mM HEPES, pH 7.5, 10 mM azide compound, 40 mM tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), 10 mM sodium ascorbate, 5 mM $CuSO_4$, and 60% (v/v) DMSO. The concentrations of azide compound, THPTA, sodium ascorbate, and $CuSO_4$ were optimized (not shown). The four reagents were added from 100 mM (10×), 1 M (25×), freshly prepared 100 mM (10×), and 100 mM (20×) stock solutions, respectively; the four reagents were mixed together and then added at once to the other components. The sample was incubated at 37 °C for 1 h and precipitated with ethanol. A second click reaction was performed with the precipitated sample. The product was purified by 8% PAGE.

Procedure for PCR of modified pools. In each selection round, two PCR reactions were performed, a 10-cycle PCR followed by a 30-cycle PCR. First, a 100 μL sample was prepared containing the capture product, 50 pmol of forward primer, 200 pmol of reverse primer, 20 nmol of each dNTP, and 10 μL of 10× KOD XL polymerase buffer [1.2 M Tris-HCl, pH 8.0, 60 mM $(NH_4)_2SO_4$, 100 mM KCl, 1% Triton X-100, 0.01% BSA], and 1 μL of KOD XL polymerase. This sample was primer-extended and then cycled according to the following PCR program: 94 °C for 2 min, 47 °C for 2 min, 72 °C for 30 min; then 94 °C for 2 min, 10× (94 °C for 1 min, 47 °C for 1 min, 72 °C for 1 min), 72 °C for 5 min. KOD XL polymerase was removed by phenol/chloroform extraction. Second, a 50 μL sample was prepared containing 1 μL of the 100 μL 10-cycle PCR product, 25 pmol of forward primer, 100 pmol of reverse primer, 10 nmol of each dNTP, 5 μCi of α-^{32}P-dCTP (3000 Ci/mmol), 5 μL of 10× Taq polymerase buffer, and 0.5 μL of Taq polymerase. This sample was cycled 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 reverse-complement single strand was separated by 8% PAGE for the subsequent primer extension.

4.4.3.4 In Vitro Selection Procedures for the FA2–FJ2 Selection Experiments

Procedures for primer extension, click reaction, and PCR were performed identically as those in EA2–EJ2 selection experiments.

Procedure for buffer-exchange and resuspension of the target protein or peptide immobilized on the magnetic beads. In a microcentrifuge tube, 50 μL of the target protein or peptide immobilized on the magnetic beads was washed with 3× 50 μL of BW buffer then 1× 100 μL of SB buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 1 mM 2-mercaptoethanol, and 0.05% Tween-20). The washed beads were resuspended in 50 μL of SB buffer.

Procedure for buffer-exchange and resuspension of the His_6Ala_2 peptide immobilized on the magnetic beads. In a microcentrifuge tube, 50 μL of the His_6Ala_2 peptide immobilized on the magnetic beads was washed with 3× 50 μL of BW buffer then 1× 100 μL of SB buffer. The washed beads were resuspended in 25 μL of SB buffer.

Procedure for selection step. Each selection experiment was initiated with 200 pmol of DNA pool. A pre-selection step was performed by incubating DNA pool with 25 μL of the His_6Ala_2 peptide immobilized on the magnetic beads and 5 μL of SB buffer at 37 °C for 15 min. The magnetic beads used in the pre-selection step were separated with a magnet and transferred to a scintillation vial (sample A). The supernatant from the pre-selection step was incubated with 50 μL of the target protein or peptide immobilized on the magnetic beads, 20 μL of SB buffer, 10 μg of herring sperm DNA, 0.1 mg of BSA (Sigma-Aldrich, cat. no. A9418), and 100 nmol of casein (USB, cat. no. 12840) at 37 °C for 15 min. After incubation, the supernatant was separated with a magnet, and the magnetic beads used in the selection step were washed with 3× 100 μL of 37 °C SB buffer that contained 10 μg of herring sperm DNA. The washes were combined with the supernatant from the selection step into a single fraction (sample B). DNA sequences bound to the target protein or peptide immobilized on the magnetic beads were eluted with 100 μL of SB buffer

that contained 2 M of guanidine-HCl via incubation at 37 °C for 5 min. The elution fraction was separated with a magnet and transferred to a scintillation vial (sample C). The magnetic beads after the elution step were transferred to another scintillation vial (sample D). Scintillation counting of the samples A–D was used to quantify background binding and selection activity. After scintillation counting, DNA sequences in sample C were precipitated with ethanol and taken onward to the PCR step.

4.5 References

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

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

(3) Chu, C. DNA Enzymes for Peptide-Nucleic Acid Conjugation and for Lysine Methylation. Ph.D. Thesis, University of Illinois at Urbana-Champaign, **2017**.

(4) Gold, L.; Ayers, D.; Bertino, J.; Bock, C.; Bock, A.; Brody, E. N. et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One* **2010**, *5*, e15004.

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

(6) Ren, X.; Gelinas, A. D.; von Carlowitz, I.; Janjic, N.; Pyle, A. M. Structural basis for IL-1α recognition by a modified DNA aptamer that specifically inhibits IL-1α signaling. *Nat. Commun.* **2017**, *8*, 810.

(7) Tolle, F.; Brändle, G. M.; Matzner, D.; Mayer, G. A versatile approach towards nucleobase-modified aptamers. *Angew. Chem. Int. Ed.* **2015**, *54*, 10971-10974.

(8) Hjertén, S.; Liu, Z. Q.; Zhao, S. L. Polyacrylamide gel electrophoresis: recovery of non-stained and stained proteins from gel slices. *J. Biochem. Biophys. Methods* **1983**, *7*, 101-113.

(9) Meers, P.; Mealy, T. Phospholipid determinants for annexin V binding sites and the role of tryptophan 187. *Biochemistry* **1994**, *33*, 5829-5837.

(10) Koopman, G.; Reutelingsperger, C. P.; Kuijten, G. A.; Keehnen, R. M.; Pals, S. T.; van Oers, M. H. J. Annexin V for flow cytometric detection of phosphatidylserine expression on B cells undergoing apoptosis. *Blood* **1994**, *84*, 1415-1420.

(11) Vermes, I.; Haanen, C.; Steffens-Nakken, H.; Reutelingsperger, C. A novel assay for apoptosis. Flow cytometric detection of phosphatidylserine expression on early apoptotic cells using fluorescein labelled Annexin V. *J. Immunol. Methods* **1995**, *184*, 39-51.

(12) Song, J. J.; Lee, Y. J. Differential cleavage of Mst1 by caspase-7/-3 is responsible for TRAIL-induced activation of the MAPK superfamily. *Cell. Signalling* **2008**, *20*, 892-906.

(13) Lonovics, J.; Devitt, P.; Watson, L. C.; Rayford, P. L.; Thompson, J. C. Pancreatic polypeptide. *Arch. Surg.* **1981**, *116*, 1256-1264.

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

(15) Tran, D. T.; Janssen, K. P. F.; Pollet, J.; Lammertyn, E.; Anné, J.; Van Schepdael, A.; Lammertyn, J. Selection and characterization of DNA aptamers for egg white lysozyme. *Molecules* **2010**, *15*, 1127-1140.

(16) Suzuki, T.; Ota, Y.; Ri, M.; Bando, M.; Gotoh, A.; Itoh, Y.; Tsumoto, H.; Tatum, P. R.; Mizukami, T.; Nakagawa, H.; Iida, S.; Ueda, R.; Shirahige, K.; Miyata, N. Rapid discovery of highly potent and selective inhibitors of histone deacetylase 8 using click chemistry to generate candidate libraries. *J. Med. Chem.* **2012**, *55*, 9562-9575.

(17) Logue, S. E.; Elgendy, M.; Martin, S. J. Expression, purification and use of recombinant annexin V for the detection of apoptotic cells. *Nat. Protoc.* **2009**, *4*, 1383-1395.

(18) Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248-254.

(19) Lee, H. W.; Lee, S. H.; Lee, H. W.; Ryu, Y. W.; Kwon, M. H.; Kim, Y. S. Homomeric and heteromeric interactions of the extracellular domains of death receptors and death decoy receptors. *Biochem. Biophys. Res. Commun.* **2005**, *330*, 1205-1212.