NUCLEOBASE-MODIFIED DEOXYRIBOZYMES FOR AMIDE AND PEPTIDE HYDROLYSIS

BY

CONG ZHOU

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 Paul J. Hergenrother
Assistant Professor Kami L. Hull
Associate Professor Douglas A. Mitchell

Abstract

Nature has evolved proteins and RNA as enzymes. These biopolymers can fold into complex structures to enable catalysis. DNA is primarily double-stranded and is non-catalytic in nature. However, considering the structural similarity to RNA, single-stranded DNA should also be able to form complex structures and perform catalysis. In fact, artificial DNA enzymes (deoxyribozymes) have been identified in laboratories by in vitro selection. The identification of new enzymes 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 total 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, with a practical amount of material, selection experiments for nucleic acid enzymes can cover a larger fraction of total sequence space. Furthermore, nucleic acid sequences can easily fold into secondary and tertiary structures, whereas most random sequences of proteins will misfold and aggregate. Between the two types of nucleic acid, DNA has several practical advantages over RNA. DNA can be directly amplified by polymerases whereas RNA requires an extra reverse transcription step. DNA is also easier to synthesize and more stable compared to RNA.

Proteases, which catalyze cleavage of proteins, are essential enzymes in nature. Natural and engineered proteases are tremendously useful for various academic, therapeutic, and industrial applications. Engineering of natural proteases for novel cleavage sites is an exciting prospect, but this process usually leads to a relaxed, rather than truly altered, substrate specificity. Because deoxyribozymes are identified from pools of random DNA sequences, no inherent peptide sequence biases must be overcome during the selection process, and thus the prospect of truly selective artificial proteases is reasonable. Previous efforts seeking DNA-catalyzed peptide cleavage resulted in the identification of deoxyribozymes that cleave a DNA phosphodiester bond. Subsequent selection experiments for DNA-catalyzed ester and amide bond hydrolysis employed an additional capture step to avoid identifying deoxyribozymes for phosphodiester cleavage,

leading to the identification of deoxyribozymes for ester hydrolysis and aromatic amide hydrolysis. However, no deoxyribozymes for aliphatic amide hydrolysis were identified. Chapter 2 describes the identification of nucleobase-modified deoxyribozymes for amide bond hydrolysis. By introducing protein-like functional groups in the nucleobases of deoxyuridines in DNA, amino, hydroxyl, and carboxyl modified deoxyribozymes were identified to catalyze the hydrolysis of a simple amide bond embedded between two DNA anchors.

Several efforts to identify modified deoxyribozymes for peptide hydrolysis are described in Chapter 3. In the initial effort, no deoxyribozymes were identified using the same strategy as described in Chapter 2. Subsequent selection experiments were performed to include two types of functional group, among amino, imidazolyl, hydroxyl, carboxyl, benzyl, and thiol, with the anticipation that some combination would function in synergy to enable the catalysis, similar to what natural proteases do in the active sites. Unfortunately, no peptide hydrolysis activity was observed in these selections. One hypothesis is that DNA-catalyzed peptide hydrolysis might require strong substrate binding by the DNA enzyme. Therefore, additional selection experiments were designed, focusing on exploring hydrophobic modifications along with protein-like functional groups. This strategy is based on considerable evidence in DNA aptamer studies that incorporating hydrophobic modifications into DNA can increase the interaction between DNA and protein. Such strong interactions between DNA and the peptide substrate may enable the DNA-catalyzed peptide hydrolysis. These selections are still in progress.

In the selections for identifying deoxyribozymes that catalyze amide and peptide hydrolysis, DNA-catalyzed radical-based oxidative DNA cleavage was identified. Surprisingly, no redox active metal ions were required for catalysis. The discovery of the DNA-catalyzed oxidative DNA cleavage indicated the general ability of DNA to catalyze redox reactions.

Acknowledgements

I am grateful for everyone who has helped me and supported me throughout my Ph.D. studies. I would like to thank my research advisor, Professor Scott Silverman for his outstanding guidance and support. His passion for science always inspires me.

I am thankful to my thesis committee members, Professor Paul Hergenrother, Professor Kami Hull, and Professor Douglas Mitchell. Thank you for all the helpful discussions and suggestions during my literature seminar, preliminary exam, original research proposal, and final defense.

I am also grateful for the talented and wonderful coworkers who have provided tremendous support every day. I would like to thank Ying, Victor, and Josh for their great guidance; Ben, Jag, Shannon, and Jimmy for all the scientific and nonscientific discussions and advice; Puzhou for all the happy time and tough time in research and life; Peter, Yves, and Shukun for the scientific assistance; Nick for the fun talks about anime and games; To the undergrads who worked with me, Paul, Yujeong, and Sheila – I appreciate your hard work and I also learned a lot from each of you.

I'm thankful to all my friends, especially Zhe, Kaimin, Ruibo, and Xinyu, for all the fun experiences and delicious food we had together. You made me feel that I was not far away from home.

I am extremely thankful to Qinyu for her love and support. Thank you for being by my side and helping me fight through the tough times in this long journey. I could not have made it without you.

Finally, I would like to thank my family, especially my parents and my grandparents, for their unconditional love, encouragement, and support throughout my life.

Table of Contents

Chapter 1: Introduction to Natural and Artificial Enzymes

1.1 Enzymatic Catalysis

Enzymes are macromolecules that catalyze biochemical reactions by adopting three-dimensional structures to bind specifically to substrates and accelerate their transformations to the corresponding products. The high reaction rate enhancement in enzymatic catalysis is the key feature of enzymes. With the help of enzymes, biochemical reactions can reach high rates, which are important for biological functions.[1] During enzymatic catalysis, the interactions between enzymes and substrates can destabilize substrate ground states and stabilize transition states, thereby reducing activation energies and enhancing reaction rates.[2] In addition, enzymes are highly specific for their substrates, which is essential for many biological processes. Enzymes adopt complex tertiary structures and interact with their substrates at multiple binding sites, thereby enabling specific substrate binding.[3] Because of the high reaction rate and substrate specificity, enzymes are essential both in nature and in many academic, therapeutic, and industrial applications.[4–6]

1.2 Natural Enzymes

Natural enzymes are biopolymers composed of small-molecule building blocks such as amino acids and ribonucleotides. Based on the different building blocks, natural enzymes can be categorized into protein enzymes and ribozymes.

1.2.1 Protein Enzymes

Protein enzymes are the predominant enzymes in nature, catalyzing most of the biochemical reactions in biological systems. Proteins are polymers of amino acids connected by peptide bonds (Figure 1.1). The chemical diversity of the functional groups on amino acid side

chains contributes to the wide variety of protein structures and functions. In addition to the 20 natural amino acids, unnatural amino acids and post-translational modifications can expand the diversity of proteins. Specific amino acid sequences can form secondary structures such as α-helices and β-sheets. These secondary structures cooperatively form tertiary structures, which can further form quaternary structures by assembly of multiple protein subunits into large protein complexes.

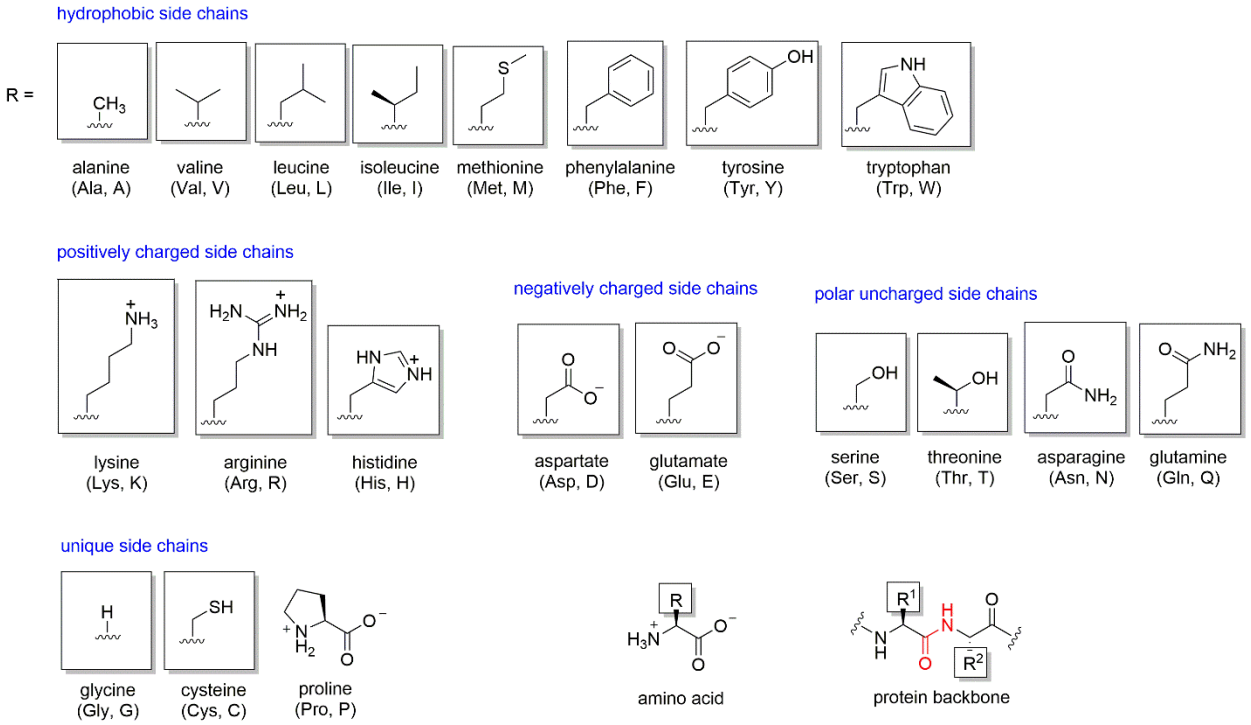

Figure 1.1. Structures of the amino acid side chains and protein backbone. Proteins are made of amino acids connected by peptide bonds (shown in red).

For a protein enzyme, the overall structure serves as a scaffold to support and position the active site, a critical region of an enzyme where the substrate molecules bind and undergo the enzymatic reaction. The active site of a protein enzyme usually consists of several amino acids with functional groups that are required for catalysis. These functional groups on amino acid side chains can engage substrates and facilitate catalysis for many different chemical reactions and through many different mechanisms. For example, in ribonucleases, imidazolyl groups can

function as proton donors or acceptors to perform acid-base catalysis.[7] In serine or cysteine proteases, hydroxyl or thiol groups can perform nucleophilic attack to form covalent intermediate with the substrates.[8] For metalloenzymes, functional groups such as imidazolyl and carboxyl groups can act as metal-binding ligands to coordinate appropriate metal ions for stabilizing transition states or reducing the pK_a of hydroxyl groups and water molecules.[9]

1.2.2 Ribozymes

Ribozymes are ribonucleic acid (RNA) molecules that can function as enzymes. RNA is a polymer of ribonucleotides connected by phosphodiester bonds (Figure 1.2). RNA was originally thought to be involved only in transferring the genetic information from the information storage molecule, DNA, to the functional molecule, protein. However, many natural RNA molecules have been identified as catalytic RNA (ribozymes) since the early 1980s, when the enzymatic abilities of natural RNA were discovered.[10] Even though the functional groups on the ribonucleotides in RNA are much less diverse compared to those on amino acids in proteins, ribozymes can still fold into complex tertiary structures and catalyze reactions that are involved in essential biological processes, namely RNA cleavage, RNA ligation, and peptide bond formation.[11]

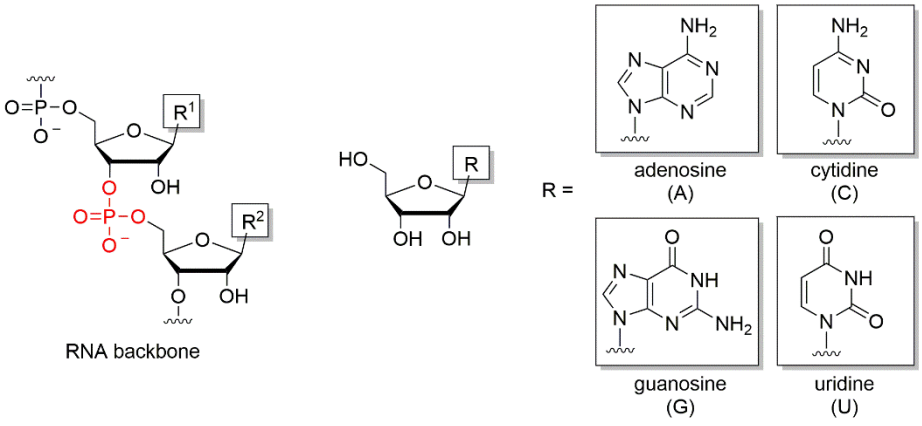

Figure 1.2. Structures of the RNA backbone and nucleobases. RNA is made of ribonucleotides connected by phosphodiester bonds (shown in red).

The first natural ribozyme discovered in the early 1980s was the group I intron from *Tetrahymena thermophila*.[10] The group I intron contains a roughly 200-nucleotide long catalytic core and catalyzes self-cleavage during RNA splicing using the 3′-hydroxyl group of an exogenous guanosine cofactor as the nucleophile. Another example, RNase P, was also identified independently in the early 1980s.[12] RNase P is a RNA-protein complex where the RNA component (containing ~200 to ~500 nucleotides) directly catalyzes the site-specific hydrolysis of precursor RNA substrates such as tRNA.[13] Rather than using the 3′-hydroxyl group of a nucleotide, RNase P catalyzes the direct nucleophilic attack of a water molecule to cleave the RNA phosphodiester backbone. Similar to RNase P, other RNA-protein complexes with catalytic RNA components were also identified, such as the ribosome, in which the rRNA component catalyzes the peptidyl transfer reaction during protein translation.[14–16]

In addition to the group I intron, RNase P, and the catalytic rRNA in the ribosome, many smaller natural ribozymes have also been identified, such as the hammerhead, hairpin, and hepatitis delta virus (HDV) ribozymes.[17] All three of these ribozymes catalyze the site-specific RNA self-cleavage via transesterification, resulting in a 5′-phosphate and 2′,3′-cyclic phosphate.[18–20] The hammerhead and hairpin ribozymes have also been shown to catalyze the reverse RNA ligation reaction.[19,21]

Given the negatively charged and hydrophilic nature of the RNA backbone and the limited diversity of functional groups in the ribonucleotide building blocks compared to amino acids, metal ions are very important for ribozymes. During RNA catalysis, metal ions can stabilize the ribozyme structures, and in most cases divalent metal ions are required to directly participate in catalysis. For example, studies of the group I intron ribozyme demonstrated that divalent metal ions are required to deprotonate the 3′-hydroxyl group, position the two substrates, and stabilize the negative charge on the leaving oxygen atom in the ribozyme active site (Figure 1.3A).[22,23] However, not all ribozymes require metal ions in the active site for catalysis. Other than metal ions, nucleobases can also participate in RNA catalysis. For example, the HDV ribozyme uses a protonated cytosine nucleobase as a general acid, donating a proton to the leaving 5′-oxygen atom

during the transesterification reaction (Figure 1.3B).[24,25] Studies also suggested a metal ion-independent mechanism for the hairpin ribozyme, likely using the nucleobases as general acid and base in catalysis.[19,26]

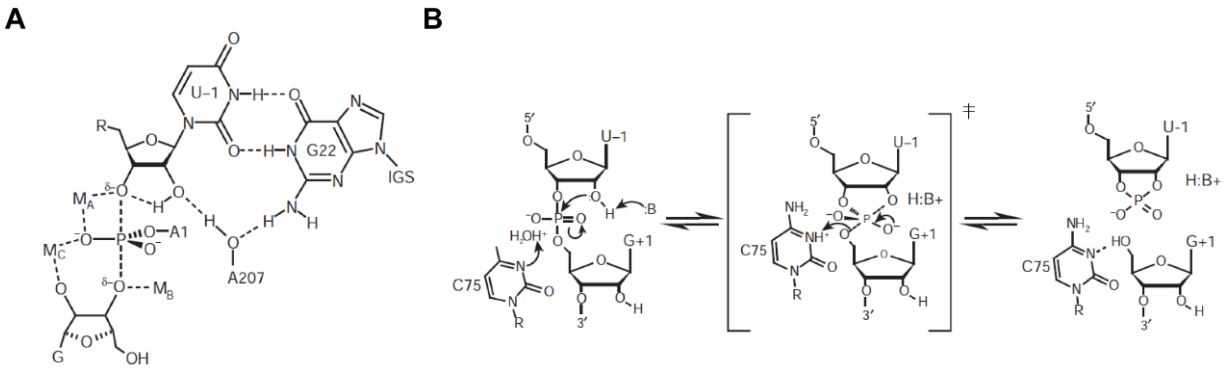

Figure 1.3. Ribozyme reaction mechanisms. (A) Mechanism of RNA transesterification by the group I intron. The image depicts the RNA cleavage by the 3′-hydroxyl group of an exogenous guanosine (G). The M_A metal ion coordinates the 3′-oxygen atom of U−1, stabilizing the developing negative charge on the leaving 3′-oxygen atom; the M_B metal ion coordinates the 3′-oxygen atom of the exogenous G; the M_C metal ion assists in positioning the substrates and stabilizes the reaction transition state. (B) Mechanism of RNA self-cleavage by the HDV ribozyme. The protonated N3 atom of the C75 nucleobase acts as a general acid, donating a proton to the 5′-oxygen of the leaving ribonucleotide. Figure reprinted with permission from ref. 11.

Ribozymes are involved in many important biological processes, such as gene expression and protein synthesis. The catalytic abilities demonstrated by these ribozymes suggest the potential value of nucleic acids as promising enzymes in both biochemical studies and industrial applications.

1.3 Artificial Enzymes

Natural enzymes have evolved for billions of years under the evolutionary pressure in nature. Their catalytic capabilities and substrate specificities are developed and fine-tuned by nature, specifically for the corresponding biochemical reactions that are essential for biological systems.[27] Inspired by the natural enzymes, engineered or artificial enzymes have been developed in the laboratory for various purposes, aiming for modified or new substrate specificities or

catalytic capabilities. Strategies such as directed evolution and rational design have been developed for engineered protein enzymes, and in vitro selection has been developed for artificial nucleic acid enzymes.

1.3.1 Development of Engineered Protein Enzymes

1.3.1.1 Directed Evolution

Directed evolution has been the most common strategy for protein engineering to improve or alter the activity of existing protein enzymes.[28–32] The highly accelerated evolution cycle allows the enzymes to be evolved to enzyme variants with the desired new functions on a practical time scale in the laboratory. An existing protein enzyme with a similar function to that of the desired enzyme variant is required as a starting point. In the directed evolution cycle (Figure 1.4), gene diversification techniques are used to generate libraries of gene variants to accelerate the exploration of protein sequence space. Screening or selection methods are used to isolate the library members with desired functions in ways that maintain the association between genotype (genes) and phenotype (functional variants).

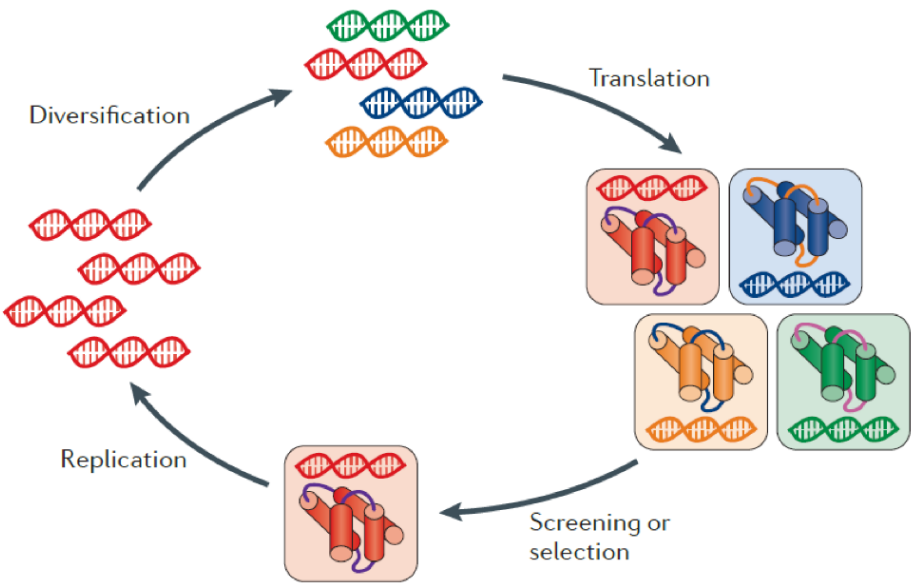

Figure 1.4. Key steps in the cycle of directed evolution for protein enzymes. The diversification step generates a library of gene variants from a parent gene. The translation step expresses the protein variants. The screening or selection step isolates the functional variants in a manner that maintains the genotype-phenotype association. The replication step amplifies the gene variants that are correlated with the isolated functional variants. The process is iterated until proteins with the desired activity are identified. Figure reprinted with permission from ref. 32.

Gene diversification strategies such as random mutagenesis, focused mutagenesis, and gene recombination are used in directed evolution to promote gene mutations. For random mutagenesis, random mutations are introduced into genes by either chemical or non-chemical methods. The chemical methods, which utilize chemical reagents such as alkylating[33] and deaminating[34] compounds to induce mutagenesis, are not commonly used for gene diversification in directed evolution due to the biases towards certain types of transitions and transversions. Error-prone PCR as a non-chemical method is widely preferred in gene diversification due to its high mutation rates and relatively low mutation biases.[35] This method is performed under conditions that reduce the base-pairing fidelity of DNA polymerases, thereby increasing the point mutation rates during PCR amplification of a gene of interest.[36,37] For focused mutagenesis, substantial genetic and structural information for the parent enzyme is required. Mutations are introduced only in specific protein sequence regions that can likely influence the functions of the protein enzymes, such as the active site.[38,39] Because the variants in the generated library contain focused mutations

in small key regions, a more comprehensive exploration of the possible sequence space around the parent enzyme can be achieved. The challenge of focused mutagenesis is that the understanding of the structure-activity relationship of the parent protein enzyme is usually limited, and potentially key mutations outside of the active site are excluded. Gene recombination methods such as homologous recombination is another strategy for gene diversification to access beneficial combinations of mutations. The original DNA shuffling method includes gene fragmentation using DNase followed by random reassembly using PCR to generate gene variants.[40] Other homologous recombination methods have been developed using different fragmentation and reassembly approaches.[41,42] These recombination methods rely on gene libraries of closely related proteins with high sequence homology to preserve gene structure among recombinants.

Screening or selection strategies are required to identify enzyme variants with desired catalytic properties.[32] Traditional screening methods such as chromatography, mass spectrometry, and colorimetric assays have limited throughput of 10^2–10^4 in a practical time. High-throughput flow cytometry screening methods such as fluorescence-activated cell sorting (FACS) can screen up to 10^8 library members in less than 24 hours.[43,44] In contrast to screening, selection strategies do not require examining each library member separately. For example, in some selection strategies, the desired protein function can be associated with the survival of a host organism such as *E. coli*.[45] The survival and replication of the active host cells leads to the enrichment of the desired gene variants within the cells. These strategies can examine libraries containing up to 10^8–10^{10} variants, usually limited by the transformation efficiency.[46] In addition, phage-assisted continuous evolution (PACE) was developed as a powerful evolution technique that allows multiple rounds of evolution in a short time period without manual intervention.[47] The desired protein function is associated with the ability of phage to infect host cells. The library size for PACE can be 10^8–10^{12}, limited by the phage titre.[32]

Directed evolution has made significant contributions to protein engineering for improved catalytic efficiency, altered substrate specificity, and new catalytic capability of protein enzymes. Despite the improvements of various evolution methods, several drawbacks remain for directed

evolution. The requirement of an existing parent enzyme as the starting point may limit the activity scope of the evolved protein enzymes. In addition, directed evolution for challenging substrate specificity usually leads to a relaxed, rather than truly altered, specificity.[48]

1.3.1.2 Rational Design

Protein rational design uses knowledge of the structures and functions of proteins to modify existing enzymes or develop new enzymes for desired functions with the assistance of computational power. In contrast to directed evolution, de novo design of proteins can develop new enzymes without an existing enzyme as the starting point and therefore has the potential to achieve completely novel functions (Figure 1.5).[49] The process is usually initiated by determining a feasible mechanism for the desired reaction, followed by the computational construction of an active site with amino acid residues required for the designed mechanism. This three-dimensional active-site model, named theozyme, is assembled into a set of protein scaffolds and optimized in silico. The resulting designed proteins are expressed and experimentally evaluated for the desired enzymatic activity. Because de novo enzyme design is very challenging, the identified enzymes are usually suboptimal and require directed evolution to achieve high activity. Although several successful examples of designed enzymes have been reported for reactions such as Kemp elimination,[50] retro-aldol reaction,[51] Diels-Alder reaction,[52] and ester hydrolysis,[53] in general enzyme design from scratch is still very difficult. In a study of designing enzymes for the retro-aldol reaction, the originally designed reactive lysine residue in the active site was abandoned by the enzyme after optimization via directed evolution. Instead, a new catalytic lysine residue at a different position was evolved, together with an alternative substrate-binding pocket.[54] These results further highlight the challenge of de novo enzyme design.

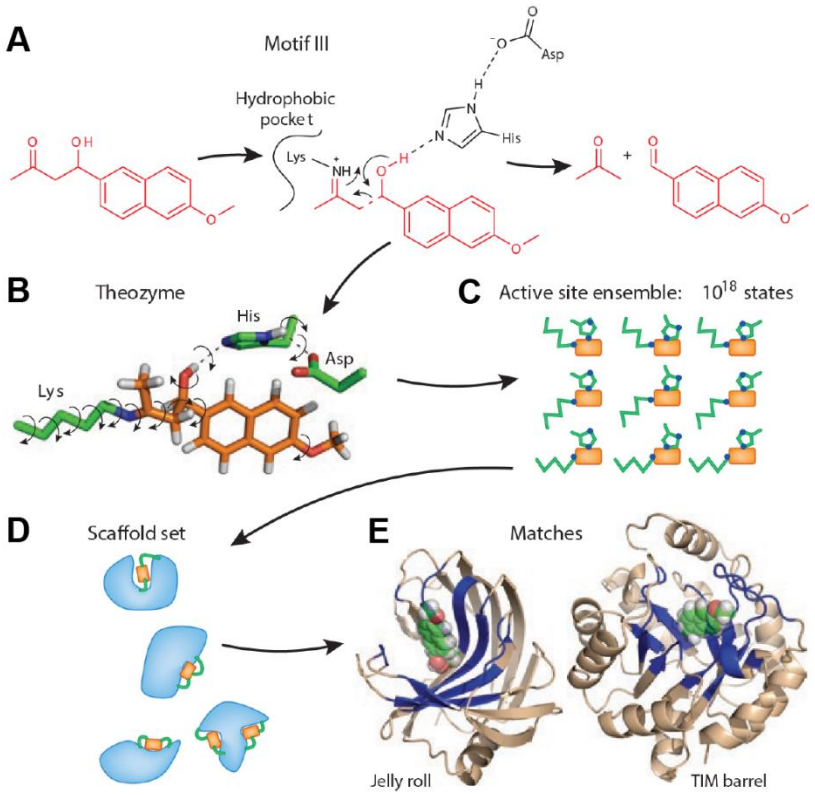

Figure 1.5. The key steps in de novo design of protein enzymes. (A) The reaction mechanism and the required amino acid side chain residues are determined. (B) The theozyme is computed by quantum mechanics optimization of the substrate transition state and catalytic residues. (C) Active-site configurations are determined. (D) The active-site configurations are mapped on a set of scaffolds to determine compatible sites. (E) Matches are further optimized by redesigning amino acids adjacent to the active site. Figure reprinted with permission from ref. 49.

1.3.2 Development of Artificial Nucleic Acid Enzymes

The discovery and mechanism studies of natural ribozymes demonstrated the potential catalytic capabilities of nucleic acids, leading to an increasing interest of identifying artificial nucleic acid enzymes for different chemical reactions.[55] Although DNA is non-catalytic and primarily double-stranded in nature, single-stranded DNA shares very similar chemical structures with RNA, and therefore DNA is considered to have similar catalytic capabilities to RNA. Many artificial RNA enzymes (ribozymes) and DNA enzymes (deoxyribozymes) have been identified from completely random sequence pools by in vitro selection.[56–58]

The general process of in vitro selection starts with a large pool (~10^{14} molecules) of random sequences (Figure 1.6). Sequences with desired activity are enriched though a carefully

designed selection step. The survival of some inactive sequences is inevitable due to the background reaction (uncatalyzed desired reaction) and the artifacts of the separation method. The selected population enriched with catalytically active sequences is subjected to an amplification step, which is critical for in vitro selection to generate a sufficient total number of sequences for the next selection round. For nucleic acid enzyme selections, RNA sequences need to be reverse-transcribed into cDNA that can be amplified by PCR and then transcribed back to RNA, whereas DNA sequences can be directly amplified by PCR. The selection rounds are iterated until the desired catalytic activity is observed, at which point the active pool is cloned and sequenced to identify individual nucleic acid enzymes.

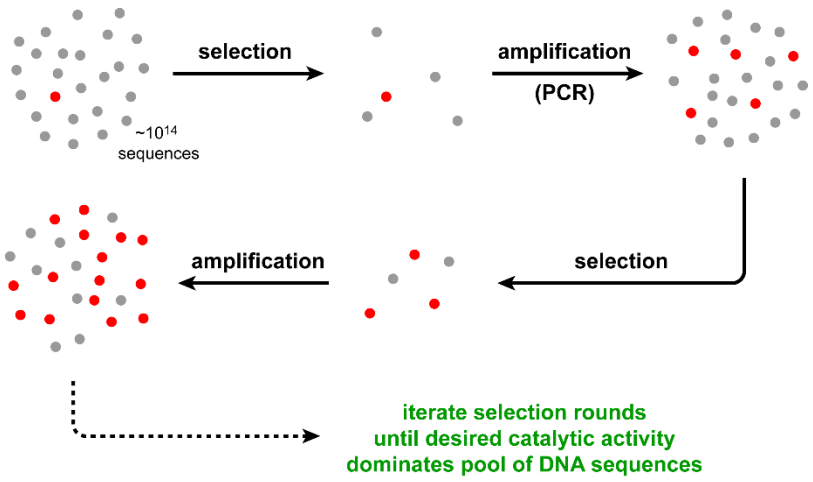

Figure 1.6. The general process of in vitro selection. The initial pool contains ~10^{14} sequences. The catalytically active sequences (red dots) are enriched in the selection step. All the survived catalytic sequences (red dots) and noncatalytic sequences (gray dots) are amplified in the amplification step.

Nucleic acids have many inherent advantages relative to proteins in the context of identifying artificial enzymes using selection or evolution-based methods.[58] In vitro selection can be performed by starting with completely random nucleic acid sequences, which allows the identification of entirely new nucleic acid enzymes without any known enzyme sequences, structures, or catalytic mechanisms as the starting point. By contrast, in vitro selection of proteins is infeasible because protein sequences cannot be amplified by any known method. In addition,

with a practically limited amount of material, nucleic acid selections can cover a much larger fraction of sequence space compared to protein selections, considering that nucleic acids only have four building blocks and therefore have a smaller total sequence space (4^n, where n is the random sequence length) relative to proteins (20^n, where 20 comes from the number of amino acid building blocks in proteins). Moreover, random nucleic acid sequences can readily fold into secondary and tertiary structures, whereas most random protein sequences suffer from misfolding and aggregation, which further limits the actual coverage of protein sequence space. Among the nucleic acids, practical considerations favor DNA over RNA in terms of stability, cost, ease of synthesis, and simplicity of amplification.

1.3.2.1 Artificial Ribozymes

Artificial ribozymes have been identified for various reactions using RNA as substrates, such as RNA cleavage,[59,60] ligation,[61,62] polymerization,[63,64] and phosphorylation.[65,66] In addition, artificial ribozymes have also been identified to catalyze chemical reactions such as Diels-Alder reaction,[67] aldol reaction,[68], alcohol oxidation,[69] and acyl transfer reaction.[70] The wide reaction scope of artificial ribozymes demonstrated the catalytic potential of RNA as an enzyme. However, detailed structural and mechanistic understanding of artificial ribozymes is still limited due to the small number of crystal structures available for artificial ribozymes.[71–73]

1.3.2.2 Deoxyribozymes

DNA has a similar chemical structure compared to RNA, only lacking the 2′-hydroxyl group and having the nucleobase thymine in place of uracil (Figure 1.7). Similar to RNA, single-stranded DNA can fold into complex secondary and tertiary structures and function as enzymes. The first deoxyribozyme was reported in 1994 and catalyzes the cleavage at a single ribonucleotide embedded in a DNA sequence.[74]

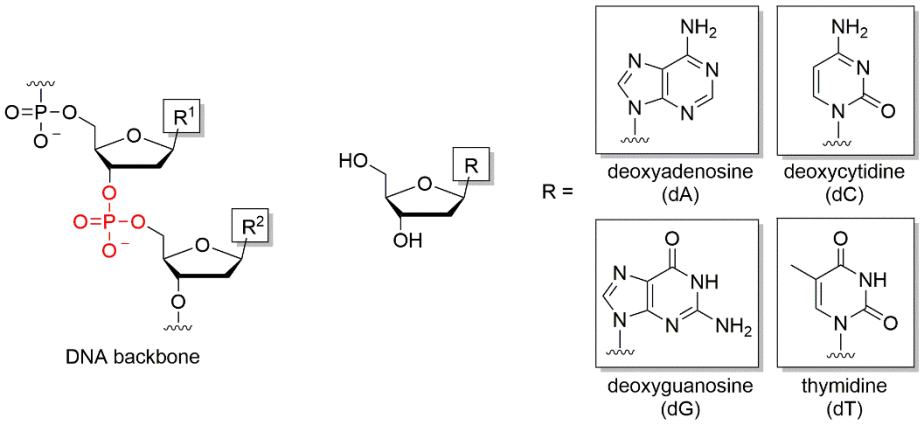

Figure 1.7. Structures of the DNA backbone and nucleobases. DNA is made of 2′-deoxyribonucleotides connected by phosphodiester bonds (shown in red).

Deoxyribozymes have been identified by in vitro selection for a variety of chemical reactions using oligonucleotides as substrates. These reactions include RNA cleavage by transesterification[75,76] or hydrolysis,[77] RNA ligation,[78–80] DNA cleavage by hydrolysis[81,82] or deglycosylation,[83] DNA ligation,[84] and DNA phosphorylation.[85,86]

The reaction scope of deoxyribozymes has been expanded to reactions using non-oligonucleotide substrates, such as modifications of peptide substrates (Figure 1.8).[87] These reactions include conjugation of oligonucleotides to tyrosine,[88–90] tyrosine phosphorylation,[91,92] phosphotyrosine (pTyr) and phosphoserine (pSer) dephosphorylation,[93] formation of dehydroalanine by elimination of phosphate from pSer,[94] tyrosine modification via azido-adenylylation,[95] and lysine modification by phosphorimidazolidated oligonucleotides.[96]

Modified nucleotides have been included in deoxyribozymes for various reactions.[58] This topic will be further discussed in Chapter 2 and Chapter 3.

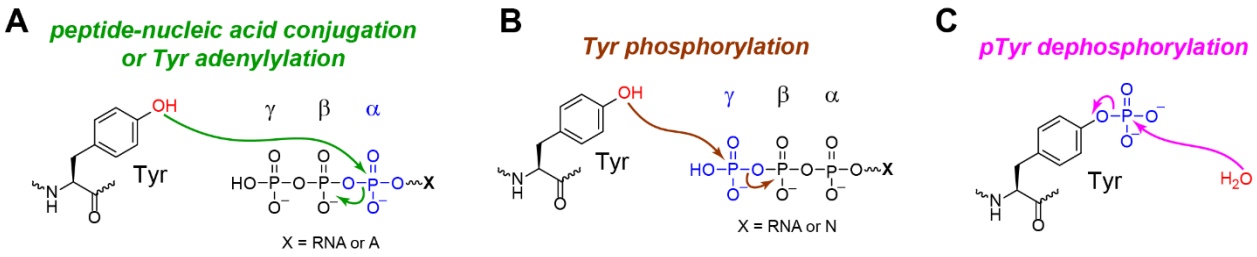

Figure 1.8. Examples of DNA-catalyzed peptide modifications. (A) Peptide-nucleic acid conjugation (X = RNA) and Tyr adenylylation (X = adenosine). (B) Tyr phosphorylation (X = RNA or nucleoside). (C) pTyr dephosphorylation.

Structural and mechanistic understanding of deoxyribozymes is very limited, because obtaining crystal structures of DNA enzymes is extremely difficult.[97] The first high-resolution crystal structure of a DNA enzyme was reported in 2016.[98] The crystal structure of the RNA-ligating 9DB1 deoxyribozyme provided important insights into DNA catalysis (Figure 1.9). For example, the crystal structure revealed no metal ions at the active site, but instead, an internucleotide phosphodiester oxygen atom was shown to have an important interaction with the newly formed phosphodiester linkage. Although 9DB1 requires metal ions for activity, the structure of the active site supports the conclusion that DNA enzymes are not necessarily metalloenzymes in terms of mechanism. More recently, the crystal structure of the RNA-cleaving 8-17 deoxyribozyme was reported, revealing that a guanine nucleobase acts as a general base and a metal-bound water acts as a general acid in the catalysis of RNA cleavage by transesterification.[99]

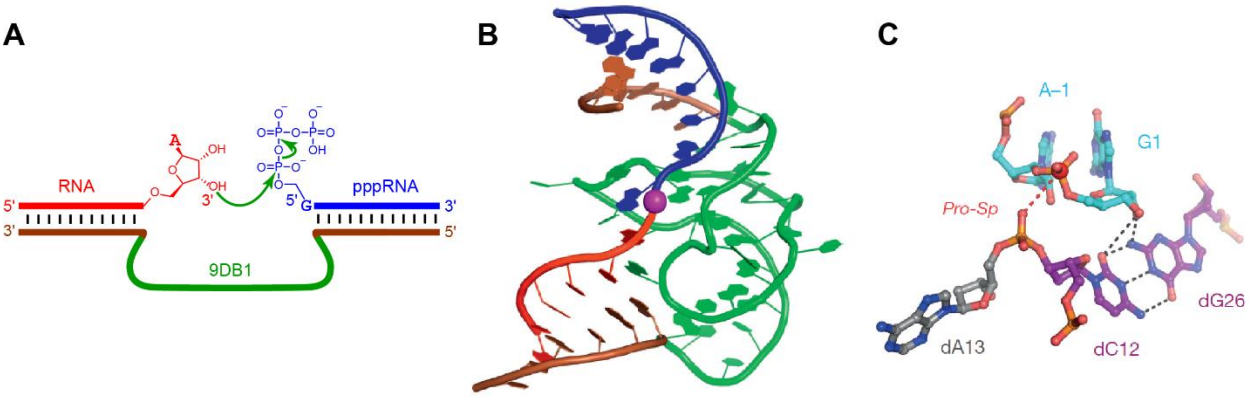

Figure 1.9. 9DB1 deoxyribozyme. (A) The RNA ligation reaction catalyzed by 9DB1. (B) The post-ligation crystal structure of 9DB1. The color codes match the image in panel A. The ligation site is marked with a purple sphere. (C) The structure of the active site. The internucleotide phosphodiester oxygen between dC12 and dA13 interacts with the phosphorus of the newly formed phosphodiester bond. Panels B and C adapted with permission from ref. 58 and ref. 98, respectively.

1.4 Proteases

1.4.1 Natural Proteases

Proteases are enzymes that catalyze protein and peptide cleavage by peptide bond hydrolysis, which is a critical biochemical reaction involved in many essential biological processes such as activation of precursor proteins, initiation of enzymatic cascades, and regulation of protein levels by degradation.[100–102] Natural proteases can be categorized into several mechanistic classes (Figure 1.10), including serine proteases, cysteine proteases, aspartyl proteases, and metalloproteases, among others.[8,103] Serine and cysteine proteases catalyze peptide hydrolysis by nucleophilic attack of a serine or cysteine residue to form a covalent intermediate followed by hydrolysis, whereas aspartyl and metalloproteases directly use a water molecule as the nucleophile.

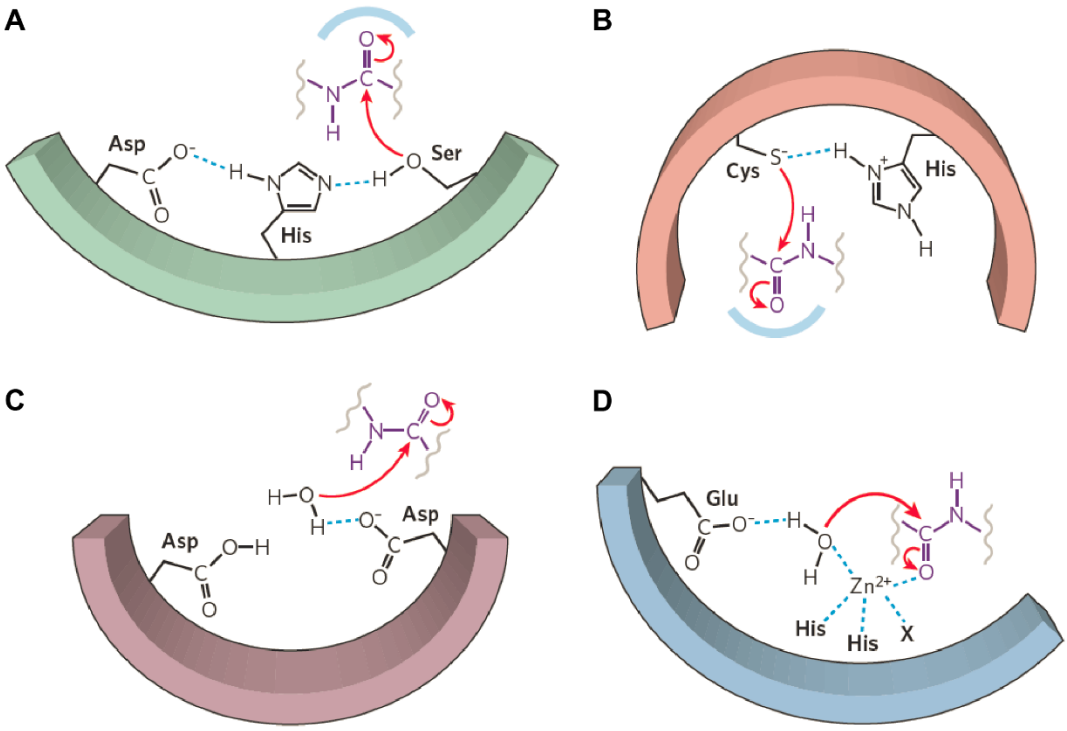

Figure 1.10. Mechanisms of peptide bond hydrolysis by natural proteases. (A) Serine proteases. (B) Cysteine proteases. (C) Aspartyl proteases. (D) Metalloproteases. Figure reprinted with permission from ref. 8.

Serine proteases use a catalytic triad consisting of an aspartate, a histidine, and a serine amino acid residue in the active site (Figure 1.10A).[104] The histidine, assisted by the aspartate, acts as a general base to deprotonate the serine hydroxyl group, enabling its nucleophilic attack on the carbonyl group of the substrate peptide bond. The acyl enzyme intermediate is hydrolyzed to regenerate the active enzyme and complete the cleavage process. During the catalysis, an oxyanion hole is usually formed by protein backbone NH groups, stabilizing the tetrahedral intermediate.

Similar to serine proteases, cysteine proteases also form a covalent intermediate and undergo subsequent hydrolysis to cleave the peptide bond, but the nucleophile is the sulfhydryl group of a cysteine residue (Figure 1.10B).[105] A histidine residue is present in the active site to function as the general base to activate the nucleophile, and an oxyanion hole is formed to stabilize the reaction transition state.

The aspartyl proteases use two aspartate residues to perform general acid-base catalysis (Figure 1.10C).[106] A water molecule coordinated in the active site is activated by hydrogen bonding with a aspartate residue. The water molecule attacks the carbonyl group of the peptide bond, leading to the bond cleavage.

The metalloproteases coordinate a metal ion such as zinc in the active site using histidine residues and sometimes also a negatively charged residue (Figure 1.10D).[107] A water molecule is bound and activated by the zinc and a glutamate residue, performing the direct nucleophilic attack to cleave the peptide bond. The zinc can also stabilize the oxyanion formed during the reaction.

Proteases have been widely used for various academic, therapeutic, and industrial applications.[108] For example, trypsin digestion for protein mass spectrometry is an important step in protein analysis and proteomics.[109] Proteases with high sequence specificity, such as thrombin and TEV protease, have been used to cleave fusion proteins or remove affinity tags.[110] Several recombinant proteases have been approved for therapeutic regulation of thrombosis and hemostasis.[111] These applications highlight the value of proteases with desired substrate specificity as biotechnological tools.

1.4.2 Artificial Proteases

The promiscuity of most natural proteases generally limits their application scope. Site-specific protein cleavage by proteases is a valuable biotechnological tool. The application scope of proteases would be significantly expanded if new proteases with unique and high substrate specificities can be achieved. For example, in proteomics studies, the protein samples must be digested by proteases before the mass spectrometry analysis,[112] because the whole proteins are often too large to be sufficiently ionized, detected, and assigned. However, for methods using the currently available proteases, the digestion-generated peptide fragments are often too small (< 2 kDa), leading to problems such as sample complexity, difficulties in aligning peptide fragments into proteins, and loss of post-translational modification information.[113] To solve these problems, middle-down proteomics was developed, ideally using peptide fragments with sizes of 2–20 kDa

in the mass spectrometry analysis.[114] However, the currently available proteases or fragmentation methods to achieve these fragment sizes are far from ideal.[113] Therefore, new artificial proteases specific for certain sequences will be valuable.

Another example for the potential applications of artificial proteases is their use in the reconstitution of biosynthesis. The biosynthesis of ribosomally synthesized and post-translationally modified peptides (RiPPs) requires a leader peptide for recognition by post-translational tailoring enzymes. After tailoring, the leader peptide is removed by the endogenous protease. Unfortunately, these endogenous proteases are often membrane-bound, and reconstituting them in vitro is challenging or impossible.[115] The use of commercial proteases, combined with the introduction of the corresponding cleavage site into the leader peptide, also has limitations, such as presence of the cleavage site in the core peptide or interference with the tailoring enzyme activity.[116] Using artificial proteases with desired sequence specificity would be a valuable method to obviate the requirement of reconstituting endogenous proteases or designing specific strategies for cleaving leader peptides in the reconstitution of RiPP biosynthesis.

Engineering proteases to achieve new substrate specificities has been proved difficult.[48] Rational design using structure-guided mutagenesis for desired new substrate specificity usually leads to a loss of or decrease in catalytic activity, because mutations at one site often disrupt structures of neighboring sites or residues important for catalysis. Directed evolution methods generally lead to a relaxed, rather than truly altered, substrate specificity,[117,118] and the selection techniques to harness proteins in directed evolution methods are limited to a few specific cases for proteases.[119] These results suggested that mutating existing proteases to achieve new substrate specificities might not be a broadly applicable way to identify new proteases. Developing deoxyribozymes as artificial proteases by in vitro selection could potentially solve these problems. Deoxyribozymes are identified from completely random sequences without any active enzyme as the starting point. Therefore, deoxyribozymes with DNA sequences that are ideal for the new substrate specificities can be directly identified.

1.5 Thesis Research Focus

The research described in this thesis is focused on the identification of nucleobase-modified deoxyribozymes for amide and peptide hydrolysis. Chapter 2 describes an adjusted selection strategy that enables the inclusion of modified nucleotides with protein-like functionality for the identification of amide-hydrolyzing deoxyribozymes. Several deoxyribozymes were identified to hydrolyze a simple amide bond embedded between two DNA anchors in a model substrate. Chapter 3 discusses the efforts toward modified deoxyribozymes for peptide hydrolysis. Peptide substrates with one or two DNA anchors were used. Selections were performed using either one or two types of modified nucleotides with protein-like functionalities. In addition, a new strategy was used in selections to incorporate multi-functional hydrophobic modifications into deoxyribozymes by click chemistry. Chapter 4 briefly describes the interesting discovery of deoxyribozymes with radical-based DNA cleavage activity.

1.6 References

(1) Wolfenden, R.; Snider, M. J. *Acc. Chem. Res.* **2001**, *34*, 938.

(2) Gao, J.; Ma, S.; Major, D. T.; Nam, K.; Pu, J.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 3188.

(3) Srere, P. A. *Trends Biochem. Sci.* **1984**, *9*, 387.

(4) Rittié, L.; Perbal, B. *J. Cell Commun. Signal.* **2008**, *2*, 25.

(5) Vellard, M. *Curr. Opin. Biotechnol.* **2003**, *14*, 444.

(6) Kirk, O.; Borchert, T. V.; Fuglsang, C. C. *Curr. Opin. Biotechnol.* **2002**, *13*, 345.

(7) Raines, R. T. *Chem. Rev.* **1998**, *98*, 1045.

(8) Erez, E.; Fass, D.; Bibi, E. *Nature* **2009**, *459*, 371.

(9) Vallee, B. L.; Williams, R. J. *Proc. Natl. Acad. Sci. USA* **1968**, *59*, 498.

(10) Cech, T. R.; Zaug, A. J.; Grabowski, P. J. *Cell* **1981**, *27*, 487.

(11) Doudna, J. A.; Cech, T. R. *Nature* **2002**, *418*, 222.

(12) Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. *Cell* **1983**, *35*, 849.

(13) Evans, D.; Marquez, S. M.; Pace, N. R. *Trends Biochem. Sci.* **2006**, *31*, 333.

(14) Nissen, P.; Hansen, J.; Ban, N.; Moore, P. B.; Steitz, T. A. *Science* **2000**, *289*, 920.

(15) Steitz, T. A. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 242.

(16) Hiller, D. A.; Singh, V.; Zhong, M.; Strobel, S. A. *Nature* **2011**, *476*, 236.

(17) Doherty, E. A.; Doudna, J. A. *Annu. Rev. Biochem.* **2000**, *69*, 597.

(18) Pley, H. W.; Flaherty, K. M.; McKay, D. B. *Nature* **1994**, *372*, 68.

(19) Nesbitt, S.; Hegg, L. A.; Fedor, M. J. *Chem. Biol.* **1997**, *4*, 619.

(20) Perrotta, A. T.; Been, M. D. *Nature* **1991**, *350*, 434.

(21) Canny, M. D.; Jucker, F. M.; Pardi, A. *Biochemistry* **2007**, *46*, 3826.

(22) Shan, S.; Kravchuk, A. V.; Piccirilli, J. A.; Herschlag, D. *Biochemistry* **2001**, *40*, 5161.

(23) Stahley, M. R.; Strobel, S. A. *Curr. Opin. Struct. Biol.* **2006**, *16*, 319.

(24) Ferré-D'Amaré, A. R.; Zhou, K.; Doudna, J. A. *Nature* **1998**, *395*, 567.

(25) Chen, J.; Yajima, R.; Chadalavada, D. M.; Chase, E.; Bevilacqua, P. C.; Golden, B. L. *Biochemistry* **2010**, *49*, 6508.

(26) Rupert, P. B.; Ferré-D'Amaré, A. R. *Nature* **2001**, *410*, 780.

(27) Bar-Even, A.; Noor, E.; Savir, Y.; Liebermeister, W.; Davidi, D.; Tawfik, D. S.; Milo, R. *Biochemistry* **2011**, *50*, 4402.

(28) Arnold, F. H.; Volkov, A. A. *Curr. Opin. Chem. Biol.* **1999**, *3*, 54.

(29) Bloom, J. D.; Arnold, F. H. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9995.

(30) Renata, H.; Wang, Z. J.; Arnold, F. H. *Angew. Chem. Int. Ed.* **2015**, *54*, 3351.

(31) Currin, A.; Swainston, N.; Day, P. J.; Kell, D. B. *Chem. Soc. Rev.* **2015**, *44*, 1172.

(32) Packer, M. S.; Liu, D. R. *Nat. Rev. Genet.* **2015**, *16*, 379.

(33) Lai, Y.; Huang, J.; Wang, L.; Li, J.; Wu, Z. *Biotechnol. Bioeng.* **2004**, *86*, 622.

(34) Myers, R. M.; Lerman, L. S.; Maniatis, T. *Science* **1985**, *229*, 242.

(35) McCullum, E. O.; Williams, B. A. R.; Zhang, J.; Chaput, J. C. *Methods Mol. Biol.* **2010**, *634*, 103.

(36) Cadwell, R. C.; Joyce, G. F. *PCR Methods Appl.* **1992**, *2*, 28.

(37) Zaccolo, M.; Williams, D. M.; Brown, D. M.; Gherardi, E. *J. Mol. Biol.* **1996**, *255*, 589.

(38) Wells, J. A.; Vasser, M.; Powers, D. B. *Gene* **1985**, *34*, 315.

(39) Cherny, I.; Greisen, P.; Ashani, Y.; Khare, S. D.; Oberdorfer, G.; Leader, H.; Baker, D.; Tawfik, D. S. *ACS Chem. Biol.* **2013**, *8*, 2394.

(40) Stemmer, W. P. *Nature* **1994**, *370*, 389.

(41) Zhao, H.; Giver, L.; Shao, Z.; Affholter, J. A.; Arnold, F. H. *Nat. Biotechnol.* **1998**, *16*, 258.

(42) Ness, J. E.; Kim, S.; Gottman, A.; Pak, R.; Krebber, A.; Borchert, T. V.; Govindarajan, S.; Mundorff, E. C.; Minshull, J. *Nat. Biotechnol.* **2002**, *20*, 1251.

(43) Fulwyler, M. J. *Science* **1965**, *150*, 910.

(44) Boder, E. T.; Wittrup, K. D. *Nat. Biotechnol.* **1997**, *15*, 553.

(45) Santoro, S. W.; Schultz, P. G. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 4185.

(46) Dower, W. J.; Miller, J. F.; Ragsdale, C. W. *Nucleic Acids Res.* **1988**, *16*, 6127.

(47) Esvelt, K. M.; Carlson, J. C.; Liu, D. R. *Nature* **2011**, *472*, 499.

(48) Khersonsky, O.; Roodveldt, C.; Tawfik, D. S. *Curr. Opin. Chem. Biol.* **2006**, *10*, 498.

(49) Nanda, V. *Nat. Chem. Biol.* **2008**, *4*, 273.

(50) Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. *Nature* **2008**, *453*, 190.

(51) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F., III; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. *Science* **2008**, *319*, 1387.

(52) Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; St Clair, J. L.; Gallaher, J. L.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L.; Houk, K. N.; Michael, F. E.; Baker, D. *Science* **2010**, *329*, 309.

(53) Richter, F.; Blomberg, R.; Khare, S. D.; Kiss, G.; Kuzin, A. P.; Smith, A. J. T.; Gallaher, J.; Pianowski, Z.; Helgeson, R. C.; Grjasnow, A.; Xiao, R.; Seetharaman, J.; Su, M.; Vorobiev, S.; Lew, S.; Forouhar, F.; Kornhaber, G. J.; Hunt, J. F.; Montelione, G. T.; Tong, L.; Houk, K. N.; Hilvert, D.; Baker, D. *J. Am. Chem. Soc.* **2012**, *134*, 16197.

(54) Giger, L.; Caner, S.; Obexer, R.; Kast, P.; Baker, D.; Ban, N.; Hilvert, D. *Nat. Chem. Biol.* **2013**, *9*, 494.

(55) Jaeger, L. *Curr. Opin. Struct. Biol.* **1997**, *7*, 324.

(56) Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1999**, *68*, 611.

(57) Breaker, R. R. *Chem. Rev.* **1997**, *97*, 371.

(58) Silverman, S. K. *Trends Biochem. Sci.* **2016**, *41*, 595.

(59) Pan, T.; Uhlenbeck, O. C. *Nature* **1992**, *358*, 560.

(60) Jayasena, V. K.; Gold, L. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 10612.

(61) Ekland, E. H.; Szostak, J. W.; Bartel, D. P. *Science* **1995**, *269*, 364.

(62) Hager, A. J.; Szostak, J. W. *Chem. Biol.* **1997**, *4*, 607.

(63) Ekland, E. H.; Bartel, D. P. *Nature* **1996**, *382*, 373.

(64) Johnston, W. K.; Unrau, P. J.; Lawrence, M. S.; Glasner, M. E.; Bartel, D. P. *Science* **2001**, *292*, 1319.

(65) Lorsch, J. R.; Szostak, J. W. *Nature* **1994**, *371*, 31.

(66) Curtis, E. A.; Bartel, D. P. *RNA* **2013**, *19*, 1116.

(67) Seelig, B.; Jäschke, A. *Chem. Biol.* **1999**, *6*, 167.

(68) Fusz, S.; Eisenführ, A.; Srivatsan, S. G.; Heckel, A.; Famulok, M. *Chem. Biol.* **2005**, *12*, 941.

(69) Tsukiji, S.; Pattnaik, S. B.; Suga, H. *Nat. Struct. Biol.* **2003**, *10*, 713.

(70) Lohse, P. A.; Szostak, J. W. *Nature* **1996**, *381*, 442.

(71) Wedekind, J. E.; McKay, D. B. *Biochemistry* **2003**, *42*, 9554.

(72) Serganov, A.; Keiper, S.; Malinina, L.; Tereshko, V.; Skripkin, E.; Höbartner, C.; Polonskaia, A.; Phan, A. T.; Wombacher, R.; Micura, R.; Dauter, Z.; Jäschke, A.; Patel, D. J. *Nat. Struct. Mol. Biol.* **2005**, *12*, 218.

(73) Shechner, D. M.; Bartel, D. P. *Nat. Struct. Mol. Biol.* **2011**, *18*, 1036.

(74) Breaker, R. R.; Joyce, G. F. *Chem. Biol.* **1994**, *1*, 223.

(75) Breaker, R. R.; Joyce, G. F. *Chem. Biol.* **1995**, *2*, 655.

(76) Santoro, S. W.; Joyce, G. F. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 4262.

(77) Parker, D. J.; Xiao, Y.; Aguilar, J. M.; Silverman, S. K. *J. Am. Chem. Soc.* **2013**, *135*, 8472.

(78) Flynn-Charlebois, A.; Wang, Y.; Prior, T. K.; Rashid, I.; Hoadley, K. A.; Coppins, R. L.; Wolf, A. C.; Silverman, S. K. *J. Am. Chem. Soc.* **2003**, *125*, 2444.

(79) Coppins, R. L.; Silverman, S. K. *Nat. Struct. Mol. Biol.* **2004**, *11*, 270.

(80) Purtha, W. E.; Coppins, R. L.; Smalley, M. K.; Silverman, S. K. *J. Am. Chem. Soc.* **2005**, *127*, 13124.

(81) Chandra, M.; Sachdeva, A.; Silverman, S. K. *Nat. Chem. Biol.* **2009**, *5*, 718.

(82) Xiao, Y.; Wehrmann, R. J.; Ibrahim, N. A.; Silverman, S. K. *Nucleic Acids Res.* **2012**, *40*, 1778.

(83) Dokukin, V.; Silverman, S. K. *Chem. Sci.* **2012**, *3*, 1707.

(84) Cuenoud, B.; Szostak, J. W. *Nature* **1995**, *375*, 611.

(85) Li, Y.; Breaker, R. R. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 2746.

(86) Camden, A. J.; Walsh, S. M.; Suk, S. H.; Silverman, S. K. *Biochemistry* **2016**, *55*, 2671.

(87) Silverman, S. K. *Acc. Chem. Res.* **2015**, *48*, 1369.

(88) Pradeepkumar, P. I.; Höbartner, C.; Baum, D. A.; Silverman, S. K. *Angew. Chem. Int. Ed.* **2008**, *47*, 1753.

(89) Wong, O. Y.; Pradeepkumar, P. I.; Silverman, S. K. *Biochemistry* **2011**, *50*, 4741.

(90) Chu, C.; Wong, O. Y.; Silverman, S. K. *ChemBioChem* **2014**, *15*, 1905.

(91) Walsh, S. M.; Sachdeva, A.; Silverman, S. K. *J. Am. Chem. Soc.* **2013**, *135*, 14928.

(92) Walsh, S. M.; Konecki, S. N.; Silverman, S. K. *J. Mol. Evol.* **2015**, *81*, 218.

(93) Chandrasekar, J.; Silverman, S. K. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5315.

(94) Chandrasekar, J.; Wylder, A. C.; Silverman, S. K. *J. Am. Chem. Soc.* **2015**, *137*, 9575.

(95) Wang, P.; Silverman, S. K. *Angew. Chem. Int. Ed.* **2016**, *55*, 10052.

(96) Brandsen, B. M.; Velez, T. E.; Sachdeva, A.; Ibrahim, N. A.; Silverman, S. K. *Angew. Chem. Int. Ed.* **2014**, *53*, 9045.

(97) Nowakowski, J.; Shim, P. J.; Prasad, G. S.; Stout, C. D.; Joyce, G. F. *Nat. Struct. Biol.* **1999**, *6*, 151.

(98) Ponce-Salvatierra, A.; Wawrzyniak-Turek, K.; Steuerwald, U.; Höbartner, C.; Pena, V. *Nature* **2016**, *529*, 231.

(99) Liu, H.; Yu, X.; Chen, Y.; Zhang, J.; Wu, B.; Zheng, L.; Haruehanroengra, P.; Wang, R.; Li, S.; Lin, J.; Li, J.; Sheng, J.; Huang, Z.; Ma, J.; Gan, J. *Nat. Commun.* **2017**, *8*, 2006.

(100) Neurath, H.; Walsh, K. A. *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 3825.

(101) Puente, X. S.; Sánchez, L. M.; Overall, C. M.; López-Otín, C. *Nat. Rev. Genet.* **2003**, *4*, 544.

(102) Overall, C. M.; Blobel, C. P. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 245.

(103) Deu, E.; Verdoes, M.; Bogyo, M. *Nat. Struct. Mol. Biol.* **2012**, *19*, 9.

(104) Hedstrom, L. *Chem. Rev.* **2002**, *102*, 4501.

(105) Verma, S.; Dixit, R.; Pandey, K. C. *Front. Pharmacol.* **2016**, *7*, 107.

(106) Northrop, D. B. *Acc. Chem. Res.* **2001**, *34*, 790.

(107) Hernick, M.; Fierke, C. A. *Arch. Biochem. Biophys.* **2005**, *433*, 71.

(108) Li, Q.; Yi, L.; Marek, P.; Iverson, B. L. *FEBS Lett.* **2013**, *587*, 1155.

(109) Vandermarliere, E.; Mueller, M.; Martens, L. *Mass Spectrom. Rev.* **2013**, *32*, 453.

(110) Waugh, D. S. *Protein Expr. Purif.* **2011**, *80*, 283.

(111) Craik, C. S.; Page, M. J.; Madison, E. L. *Biochem. J.* **2011**, *435*, 1.

(112) Switzar, L.; Giera, M.; Niessen, W. M. A. *J. Proteome Res.* **2013**, *12*, 1067.

(113) Cristobal, A.; Marino, F.; Post, H.; van den Toorn, H. W. P.; Mohammed, S.; Heck, A. J. R. *Anal. Chem.* **2017**, *89*, 3318.

(114) Wu, C.; Tran, J. C.; Zamdborg, L.; Durbin, K. R.; Li, M.; Ahlf, D. R.; Early, B. P.; Thomas, P. M.; Sweedler, J. V.; Kelleher, N. L. *Nat. Methods* **2012**, *9*, 822.

(115) Dunbar, K. L.; Mitchell, D. A. *ACS Chem. Biol.* **2013**, *8*, 473.

(116) Bindman, N. A.; Bobeica, S. C.; Liu, W. R.; van der Donk, W. A. *J. Am. Chem. Soc.* **2015**, *137*, 6975.

(117) Yoo, T. H.; Pogson, M.; Iverson, B. L.; Georgiou, G. *ChemBioChem* **2012**, *13*, 649.

(118) Yi, L.; Gebhard, M. C.; Li, Q.; Taft, J. M.; Georgiou, G.; Iverson, B. L. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 7229.

(119) Varadarajan, N.; Rodriguez, S.; Hwang, B.; Georgiou, G.; Iverson, B. L. *Nat. Chem. Biol.* **2008**, *4*, 290.

Chapter 2: Modified Deoxyribozymes for Amide Bond Hydrolysis*

2.1 Introduction

2.1.1 Previous Efforts toward DNA-Catalyzed Amide Bond Hydrolysis

2.1.1.1 Seeking DNA-Catalyzed Amide Bond Hydrolysis Using Tripeptide Substrate with Two DNA Anchors

In the initial attempt for DNA-catalyzed amide bond hydrolysis, in vitro selection was performed using a DNA-Ala-Ser-Ala-DNA tripeptide substrate in which the Ala-Ser-Ala tripeptide is embedded between two DNA anchors (Figure 2.1A).[1,2] The tripeptide was presented to a DNA pool with 40-nucleotide random region (N_{40}) by Watson-Crick interactions between two DNA anchors (fixed DNA sequences) in the substrate and complementary DNA binding arms in the pool. In the selection experiments, the tripeptide substrate was ligated to the N_{40} pool and incubated in selection condition with divalent metal ions. DNA sequences that cleave the tripeptide substrate can be separated by PAGE gel shift caused by losing the left DNA anchor upon cleavage. Surprisingly, instead of amide-hydrolyzing deoxyribozymes, DNA-hydrolyzing deoxyribozymes were discovered in this selection. Each of these deoxyribozymes catalyzes the hydrolysis of a specific DNA phosphodiester linkage in one of the DNA anchors (Figure 2.1B). This finding was unexpected considering the uncatalyzed half-life of DNA phosphodiester hydrolysis is about 10^5-fold greater than that of peptide hydrolysis.[3]

---

* This research has been published:
Zhou, C.; Avins, J. L.; Klauser, P. C.; Brandsen, B. M.; Lee, Y.; Silverman, S. K. *J. Am. Chem. Soc.* **2016**, *138*, 2106.

University of Illinois postdoc Joshua Avins performed early stage material preparation.
University of Illinois graduate student Benjamin Brandsen assisted with the substrate preparation and cloning.
University of Illinois undergraduate student Paul Klauser and Yujeon Lee performed cloning of ^{HO}dU and ^{COOH}dU selection.

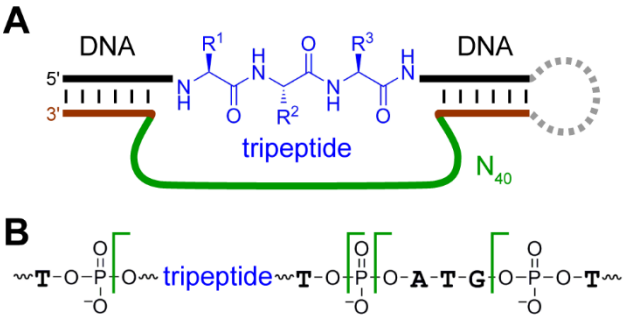

Figure 2.1. Initial attempt toward DNA-catalyzed peptide bond cleavage and unexpected discovery of DNA-hydrolyzing deoxyribozymes. (A) In vitro selection design for DNA-catalyzed peptide hydrolysis using tripeptide substrate. Catalytically active sequences were separated by a downward PAGE shift. (B) Cleavage sites of DNA-hydrolyzing deoxyribozymes. Each deoxyribozyme cleaves at only one particular site. Figure adapted with permission from ref. 2.

2.1.1.2 The Use of Capture Step and Identification of Deoxyribozymes with Ester and Aromatic Amide Hydrolysis Activity

The initial effort demonstrated that it is necessary to use a strategy that can differentiate between desired and undesired reaction in selection process. Therefore, in the subsequent effort, a "capture" step was included in the selection process to specifically capture the DNA sequences that catalyze only the desired reaction.[4] For DNA-catalyzed amide bond hydrolysis, the desired reaction will generate a carboxylic acid as product, whereas the undesired DNA phosphodiester hydrolysis will give a different product, phosphomonoester. To differentiate these two products and specifically capture the desired one, a coupling reaction was optimized and used in the capture step. This coupling reaction, which uses the compound DMT-MM as an activating agent, can selectively join the amide hydrolysis product carboxylic acid to a 5′-amino "capture oligonucleotide", causing an upward PAGE gel shift to separate the DNA sequences with amide hydrolysis activity (Figure 2.2A). Using this capture step in each selection round, the in vitro selection in this study led to several deoxyribozymes for ester hydrolysis and several for aromatic amide hydrolysis, but none for aliphatic amide hydrolysis (Figure 2.2B).

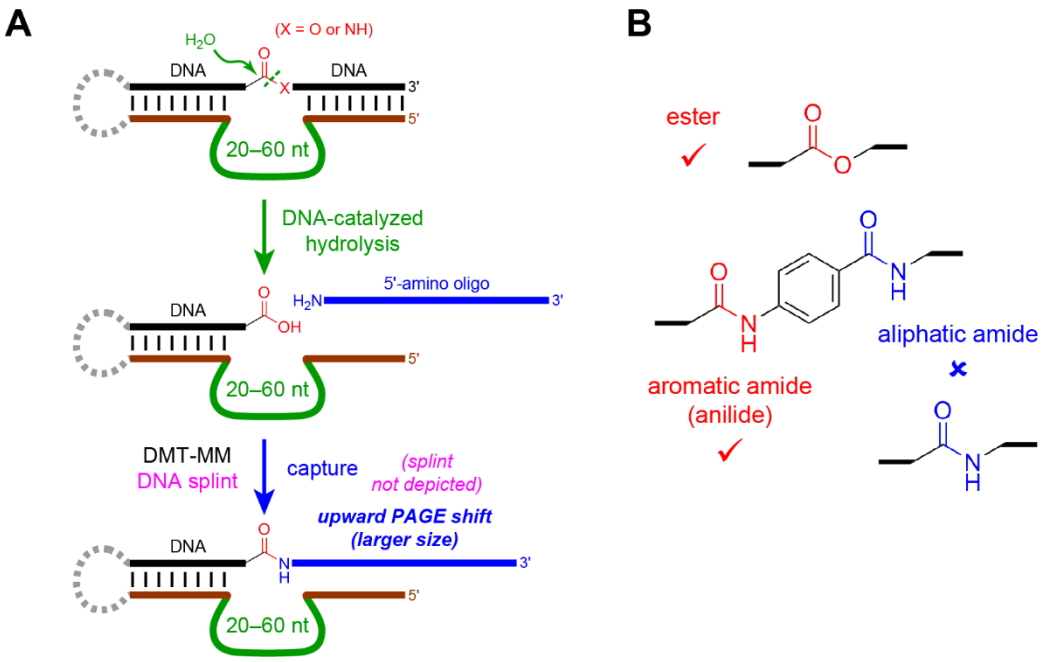

Figure 2.2. DNA-catalyzed ester and aromatic amide hydrolysis. (A) Capture step for the carboxylic acid product of ester or amide hydrolysis. (B) Selection outcome. DNA-catalyzed hydrolysis was observed for ester and aromatic amide (anilide) substrates, but not for aliphatic amides. Figure adapted with permission from ref. 4.

2.1.1.3 Other Attempts for DNA-Catalyzed Amide Bond Cleavage

Several additional efforts using fundamentally different strategies were performed for DNA-catalyzed amide and peptide bond cleavage. The previous results suggested that water might not be nucleophilic enough, or the aliphatic amide carbonyl not electrophilic enough, for deoxyribozymes to catalyze the amide bond cleavage. One direction is to explore stronger nucleophiles such as aliphatic amine, hydrazide and thiol. The other direction is to perform selections using a series of amide with different electrophilicity, for example acetylglycine or fluoro-, difluoro- and trifluoroacetylglycine. Selections using these strategies were performed but no deoxyribozymes were identified. Another strategy is to promote intramolecular peptide self-cleavage by including an amino acid with a nucleophilic side chain in the peptide substrate, such as cysteine, serine, or homoserine. However, no deoxyribozymes were identified in the selection.

2.1.2 The Use of Modified Nucleotides in Deoxyribozymes

Despite the success of deoxyribozymes that catalyze different chemical reactions, one distinct disadvantage of DNA is its limited chemical functionality, especially when compared to protein, the most common enzyme in nature. The four nucleobases (A, G, T, C) in DNA are all structurally similar aromatic heterocycles, which could potentially limit the scope of activities for DNA as catalyst. Therefore, for certain activities that are difficult to achieve using standard DNA, incorporating suitable modified nucleotides can be a viable way to enable the intended activity by expanding the chemical functionality of DNA. Similarly, for deoxyribozymes with poor activity, incorporating modifications could be beneficial for catalysis in terms of rate, yield, substrate scope, or cofactor dependence.[5]

In the context of modified deoxyribozymes, nucleotide triphosphates with modifications on the nucleobases are often used to incorporate functional groups in DNA via suitable polymerases.[6,7] Many laboratories have used such methods in the identification of modified RNA-cleaving deoxyribozymes, aiming to reduce or obviate the divalent metal ion (M^{2+}) requirment.[8–10] In one representative study from the Perrin laboratory, deoxyribozymes with three different types of modification (Figure 2.3) are able to catalyze RNA cleavage in the absence of M^{2+}.[11–13]

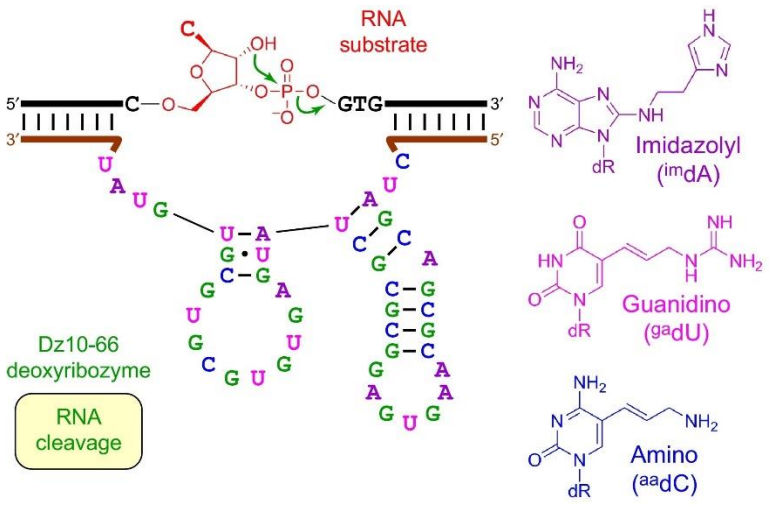

Figure 2.3. M^{2+}-independent RNA-cleaving Dz10-66 deoxyribozyme with three types of modified nucleotide. The substrate has a single RNA nucleotide embedded within a DNA sequence. Figure reprinted with permission from ref. 5.

Not only could the chemical functionality of DNA be expanded by incorporating modifications on nucleobases, but also the structure diversity of DNA could be explored by using a class of synthetic genetic polymer called xeno nucleic acid (XNA). Instead of ribose (in RNA) and deoxyribose (in DNA), XNAs have unnatural backbone structures such as those of arabino nucleic acid (ANA), 2′-fluoroarabino nucleic acid (FANA), hexitol nucleic acid (HNA), or cyclohexene nucleic acid (CeNA). A study of Holliger and coworkers demonstrated the capability of XNA as catalyst,[14] indicating potential value of XNA in de novo enzyme development.

2.1.3 Using Modified Nucleotides in Deoxyribozymes for Amide Hydrolysis

For DNA-catalyzed amide hydrolysis, the previous extensive efforts using unmodified DNA did not lead to any deoxyribozymes. Considering the limited chemical functionality of unmodified DNA, one way to improve the catalytic capability of DNA is to include modified nucleotides with functional groups that are not found in natural DNA. Inspired by the functionality diversity and success of protein enzymes in nature, we decided to experimentally evaluate the

contributions of protein-like functional groups in the context of identifying deoxyribozymes for amide bond hydrolysis.

2.2 Results and Discussion

2.2.1 In Vitro Selection Strategy

We designed four types of modified nucleotide with protein-like functional groups of primary amino, primary hydroxyl, carboxyl, or imidazolyl attached to the 5-position of thymidine, thereby forming 5-substituted 2′-deoxyuridine derivatives ^{Am}dU, ^{HO}dU, ^{COOH}dU, or ^{Im}dU (Figure 2.4). These functional groups could potentially function as nucleophiles, general acids/bases, or metal-binding ligands in plausible mechanisms of amide hydrolysis.

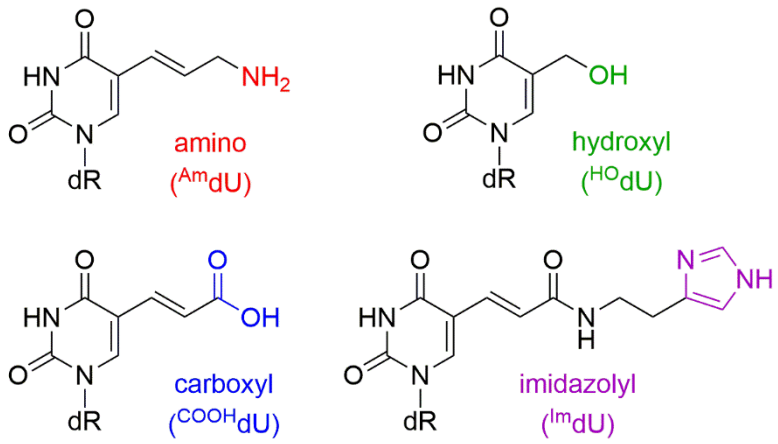

Figure 2.4. Structures of four modified dU with protein-like functionality (dR = deoxyribose).

The amide substrate used in selection has a simple aliphatic amide bond embedded between two DNA anchors (Figure 2.5). Note that this arrangement is structurally similar to the previous selection (Figure 2.2).[4]

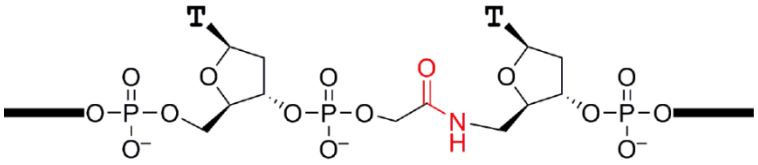

Figure 2.5. Amide substrate with detailed structure near the amide bond cleavage site. This amide substrate was prepared by a splinted coupling reaction using a 3′-AHA (α-hydroxyacetic acid) oligonucleotide and a 5′-amino-5′-deoxythymidine oligonucleotide. Figure reprinted with permission from ref. 15.

To prepare the N_{40} pools for selections using ^{Am}dU, ^{HO}dU, and ^{COOH}dU, the corresponding modified nucleoside phosphoramidites were used in solid-phase oligonucleotide synthesis. For selection using ^{Im}dU, the pool was prepared by primer extension on a reverse complement template N_{40} pool using ^{Im}dUTP and KOD XL DNA polymerase. ^{Am}dU, ^{HO}dU, and ^{COOH}dU phosphoramidite were purchased, and ^{Im}dUTP was synthesized in our lab.

The in vitro selection process was adjusted as required for incorporating the modified nucleotides (Figure 2.6).[15] Selection experiments started with the ligation of the amide substrate to the modified N_{40} pool. The subsequent selection step for amide hydrolysis was performed in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, 150 mM NaCl at 37 °C for 14 h, where each of Zn^{2+}, Mn^{2+}, and Mg^{2+} have been useful cofactors for other deoxyribozymes. An optimized capture step, using compound EDC as an activating agent and compound HOBt to form stable activated ester, was included to specifically capture the carboxylic acid product of amide hydrolysis. At the end of each selection round, the sequence population was amplified by PCR, and the reverse complement DNA strand was isolated. The new modified DNA pool with enriched catalytically active sequences was generated by primer extension from the reverse complement strand population using the appropriate modified dUTP and KOD XL DNA polymerase.

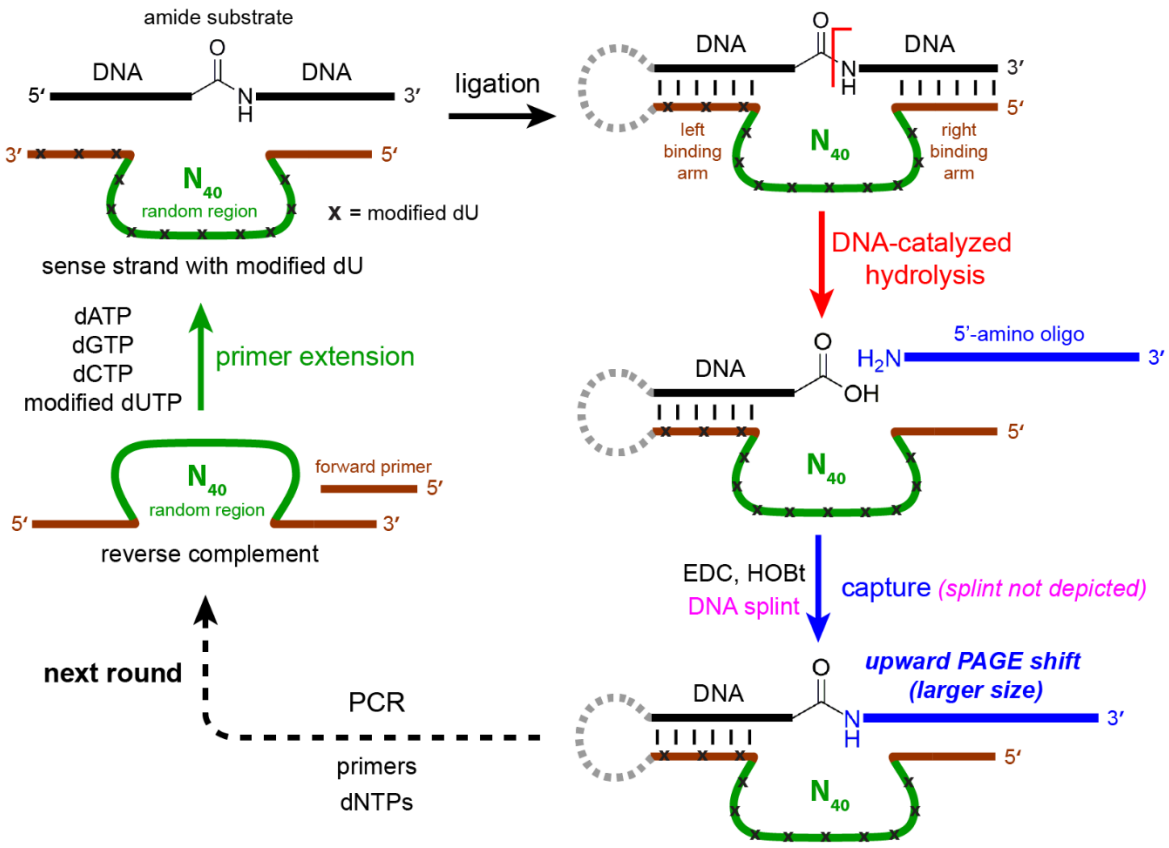

Figure 2.6. In vitro selection using modified DNA to achieve amide bond hydrolysis.

2.2.2 Identification and Characterization of Deoxyribozymes for Amide Hydrolysis

In vitro selection experiments were performed for DNA-catalyzed amide hydrolysis using ^{Am}dU (JV1 selection, JV1 is the systematic alphabetical designation for the particular selection), ^{HO}dU (JY1 selection), ^{COOH}dU (JX1 selection) and ^{Im}dU (JW1 selection). Deoxyribozymes with amide hydrolysis activity were identified for the ^{Am}dU, ^{HO}dU and ^{COOH}dU selections. However, no detectable activity (< 0.5%) was observed for the ^{Im}dU selection after 14 rounds, and the selection was discontinued.

2.2.2.1 ^{Am}dU-Modified Deoxyribozymes for Amide Hydrolysis

For the JV1 selection using ^{Am}dU, 16% capture yield of the DNA pool was observed after 8 rounds (Figure 2.7). To identify individual deoxyribozyme sequences, the active pool was cloned from round 8. Eleven different catalytically active sequences were identified, and they all belong

to one single sequence family (Figure 2.8), with a predominant sequence designated as AmideAm1 and ten variants with only one or two nucleotide differences. For AmideAm1, in the initially random region containing 40 nucleotides, only six are ^{Am}dU. Interestingly, in the sequence alignment, four of these six positions (^{Am}dU in AmideAm1) are mutated to dC or dA in other catalytically active variants, suggesting that these four ^{Am}dU in AmideAm1 are not required for catalysis. However, the remaining two positions of ^{Am}dU (positions 34 and 38 of the initially random region of 40 nucleotides) are completely conserved across all variants, indicating that these two ^{Am}dU are catalytically relevant.

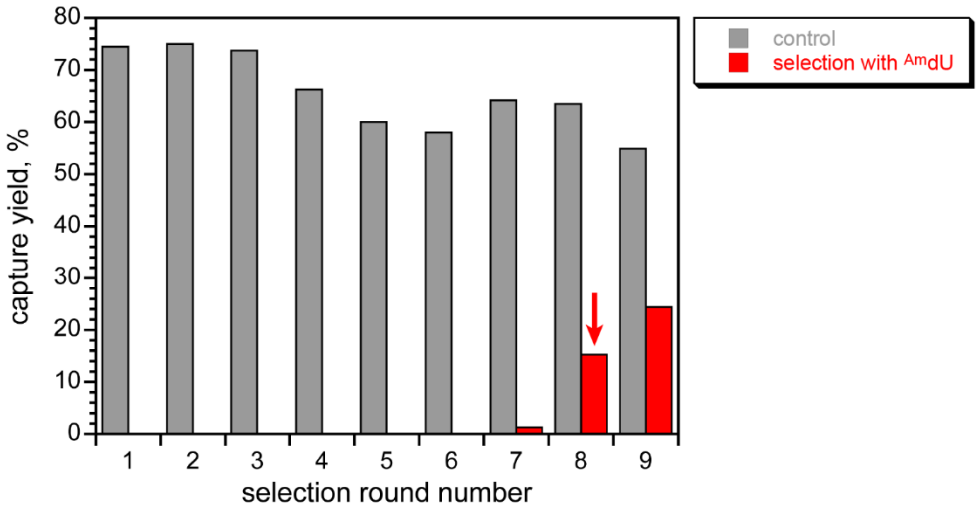

Figure 2.7. Progression of the JV1 selection with ^{Am}dU. In each round, "control" refers to the EDC-promoted capture yield for the product that has a 3′-COOH group, and "selection" refers to the EDC-promoted capture yield using the deoxyribozyme pool for that round. Arrow marks the cloned round.

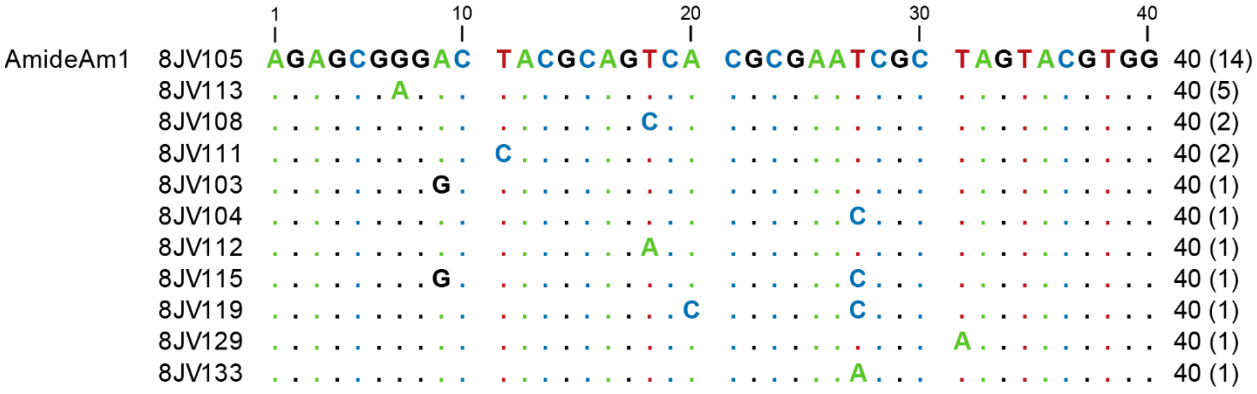

Figure 2.8. Sequences of the initially random region of the ^{Am}dU-modified deoxyribozymes for amide hydrolysis. All deoxyribozymes were used as 5′-CGAAGTCGCCATCTCTTC-N_{40}-ATAGTGAGTCGTATTA-3′, where N_{40} represents the specific 40 nucleotides of the initially random region as determined through the selection process. All alignments show only the initially random region. A dot denotes conservation, i.e., the same nucleotide as in the uppermost sequence. Next to each sequence length in nucleotides on the far right is shown (in parentheses) the number of times that sequence was found during cloning. Each deoxyribozyme is listed with its internal laboratory designation such as 8JV105, where 8 is the round number, JV1 is the systematic alphabetical designation for the particular selection, and 05 is the clone number.

A new variant, the AmideAm1 quadruple mutant, was generated by simultaneously mutating those four dispensable ^{Am}dU in AmideAm1 into dC or dA, with only positions 34 and 38 retained as ^{Am}dU (Figure 2.9A). Both AmideAm1 (parent) and its quadruple mutant were characterized for cleavage of amide substrate in the same incubation condition as in the selection process (Figure 2.9B). The parent AmideAm1 catalyzed amide hydrolysis with k_{obs} of 0.11 h^{-1} and 64% yield at 48 h, while the quadruple mutant had k_{obs} 0.06 h^{-1} and slightly lower 49% yield at 48 h (Figure 2.9C). The amide hydrolysis activity was confirmed by MALDI mass spectrometry of the cleavage products (Figure 2.9D). The scope of alternative amide substrates for AmideAm1 was briefly evaluated by using analogous amide substrate derivatives with GHB (γ-hydroxybutyric acid) amide or Ala-Phe-Ala tripeptide embedded between two DNA anchors, but no cleavage activity was observed. The cleavage of the analogous all-DNA substrate lacking the amide bond was also evaluated, and no cleavage activity was detected.

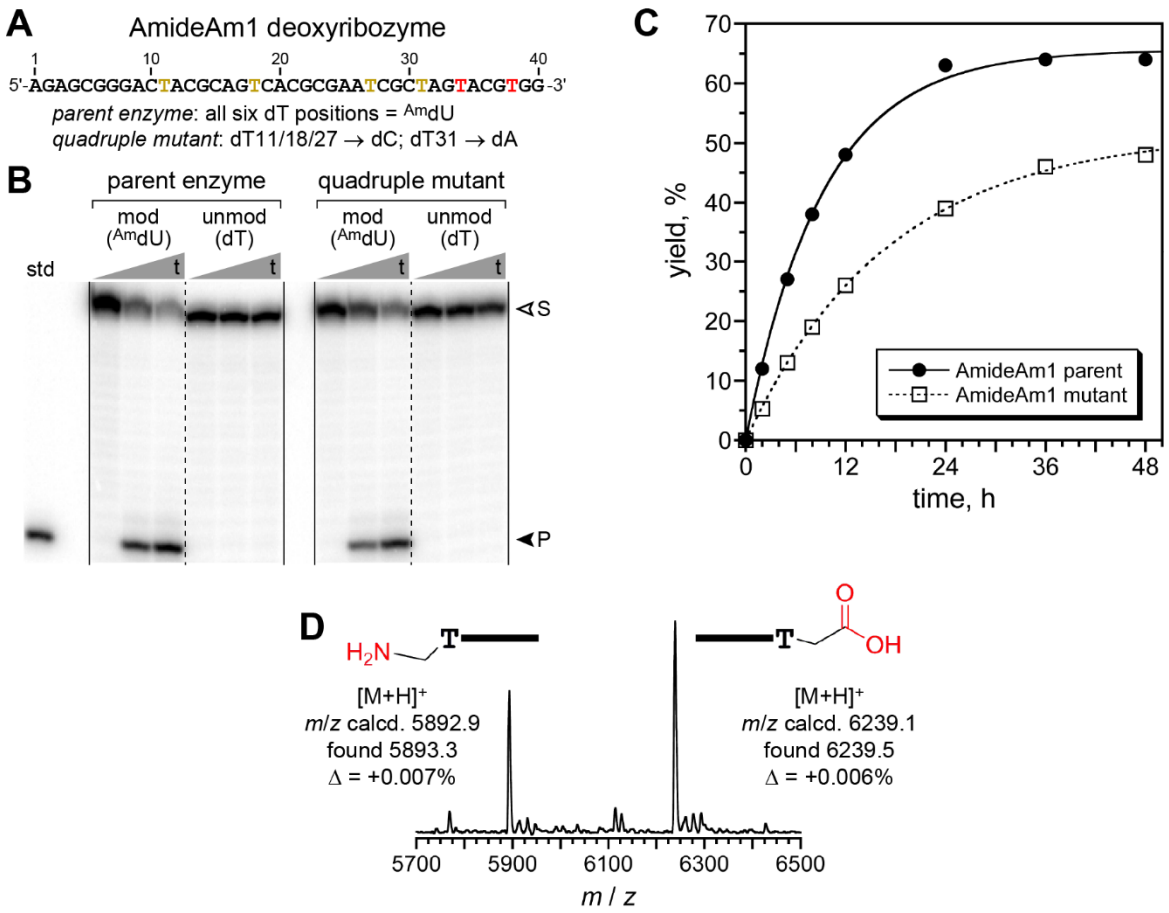

Figure 2.9. ^{Am}dU-modified AmideAm1 deoxyribozyme for amide hydrolysis. (A) Sequence of the initially random (N_{40}) region of AmideAm1. The four yellow T positions are ^{Am}dU in the parent sequence but mutatable as indicated. (B) PAGE assay (t = 0, 12, 48 h). (C) Kinetic plots. k_{obs}, h^{-1} (n = 3): parent 0.11 ± 0.01 and quadruple mutant 0.057 ± 0.007. Because k_{bkgd} $<10^{-4}$ h^{-1} (i.e., <0.5% in 48 h), the rate enhancement k_{obs}/k_{bkgd} is $>10^3$. (D) Mass spectrometry analysis of AmideAm1 product.

In the experiments of Figure 2.9, each ^{Am}dU-modified deoxyribozyme was prepared by primer extension from the corresponding reverse complement template, as was done in every selection round for generating the modified pool. For all the ^{Am}dU-modified deoxyribozymes generated using this method, ^{Am}dU were not only incorporated in the initially random region of 40 nucleotides, but also included in the fixed-sequence 3′-binding arm. To assess the functional contributions of ^{Am}dU in the 3′-binding arm of AmideAm1, a version of AmideAm1 without ^{Am}dU modifications in the 3′-binding arm was prepared using splint ligation of a truncated primer extension product and an unmodified 3′-binding arm. No obvious difference was observed

between AmideAm1 with and without 3′-binding arm modifications in terms of the amide hydrolysis rate and yield (Figure 2.10), suggesting that ^{Am}dU modifications in the 3′-binding arm are not catalytically relevant.

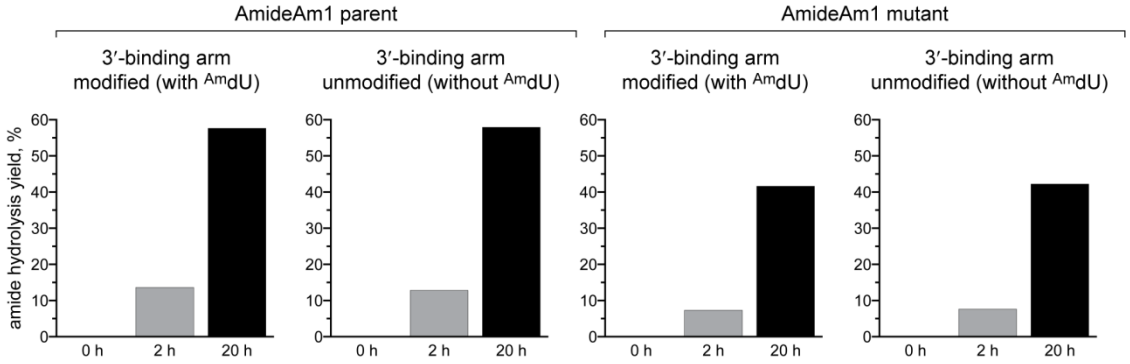

Figure 2.10. Assays of ^{Am}dU requirement in 3′-binding arm for AmideAm1 parent and mutant deoxyribozymes. Both the parent version of AmideAm1 (with all six ^{Am}dU nucleotides marked in Figure 2.9A) and the mutant version of AmideAm1 (with only positions 34 and 38 of the initially random N_{40} region as ^{Am}dU) were assayed when ^{Am}dU was either included or not included within the 3′-binding arm.

To understand the ^{Am}dU requirements in AmideAm1 and dissect the catalytic roles of the two critical ^{Am}dU at positions T34 and T38, several versions of AmideAm1 (Figure 2.11A) were prepared by solid-phase DNA synthesis using the ^{Am}dU phosphoramidite specifically for one or both of positions T34 and T38. All other nucleotides that were originally ^{Am}dU in AmideAm1 sequence were replaced by unmodified dT. These AmideAm1 versions with specific ^{Am}dU and unmodified dT positions cannot be prepared by primer extension, because it cannot differentiate ^{Am}dU and dT in a site-specific manner. Assays using AmideAm1 variants **3**, **4** and **5** (Figure 2.11B) revealed that ^{Am}dU at both positions T34 and T38 (in variant **3**) are required for high catalytic activity, although reduced activity (~10% yield at 48 h; k_{obs} decrease by ~3 fold) was still observed for single-^{Am}dU variants **4** and **5**. Moreover, comparison of the results with variants **4** and **5** to those with variants **7** and **8** indicated that the T nucleobases at positions 34 and 38 are also important for catalysis. In summary, AmideAm1 only requires as few as one ^{Am}dU in the initially

random region of 40 nucleotides to achieve amide hydrolysis, whereas the remaining 39 of 40 can be unmodified.

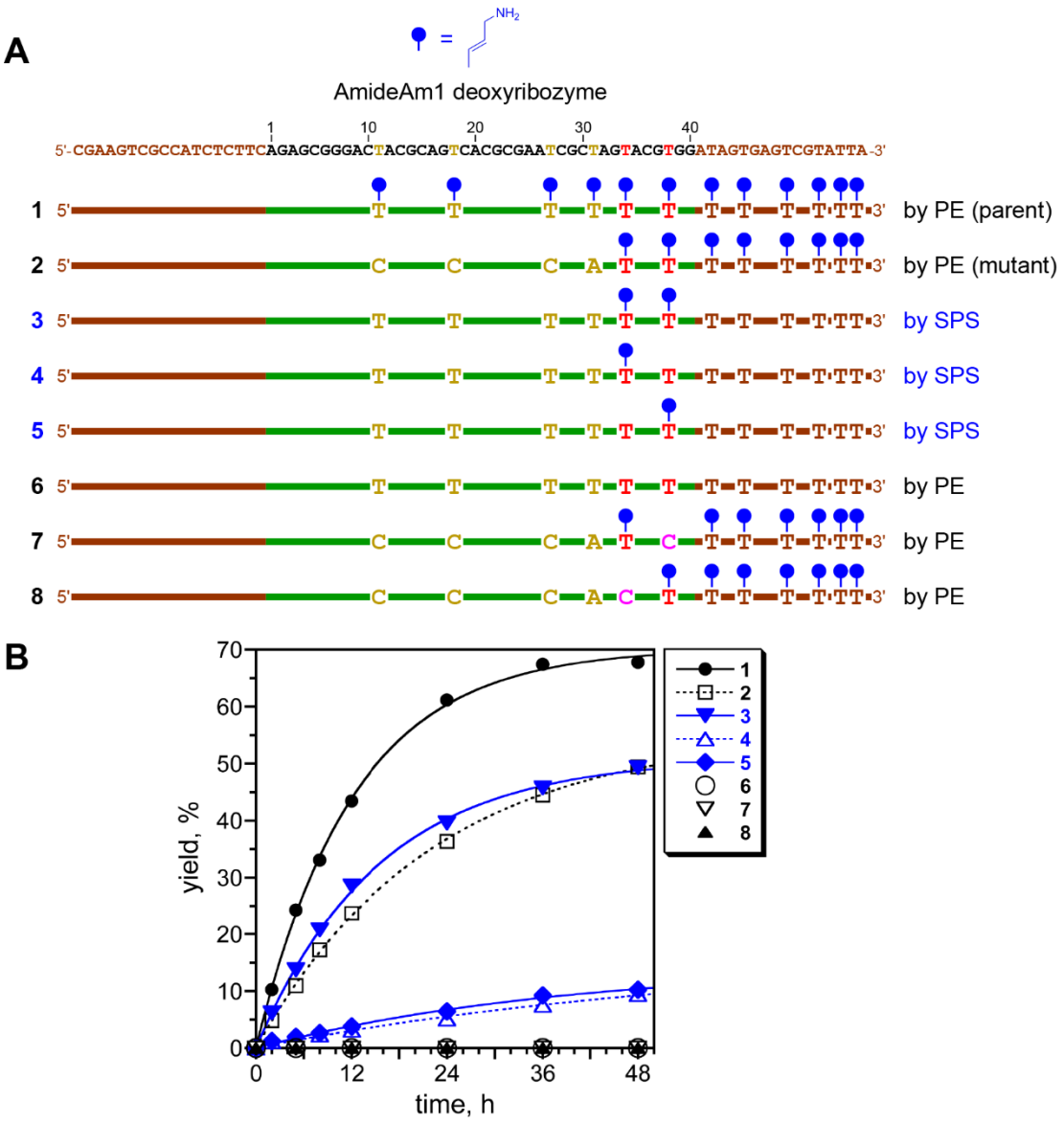

Figure 2.11. Assays of ^{Am}dU requirements at T34 and T38 positions for AmideAm1. Shown are assays for eight variants of AmideAm1. Variants **1** and **2** are the parent and mutant deoxyribozymes from Figure 2.9, respectively, each prepared by primer extension ("PE"). Therefore, all dT positions in the initially random N_{40} region and 3′-binding arm are ^{Am}dU, and removing the four N_{40} primary amino groups at positions T11, T18, T27, and T31 required mutation to dC/dA (with choice of nucleotide dictated by the phylogeny in Figure 2.8), because primer extension does not allow site-specific replacements. Variants **3**, **4**, and **5** were each prepared by solid-phase synthesis ("SPS"), which allows individual replacement of ^{Am}dU with unmodified dT. Variant **3** includes ^{Am}dU at only positions 34 and 38. Variants **4** and **5** include ^{Am}dU at only position 34 or 38, respectively (k_{obs} ~0.02 h^{-1}). Variants **6**, **7**, and **8** were each prepared by primer extension (although variant **6** could equally well be prepared by solid-phase synthesis), and each had no detectable activity.

2.2.2.2 ^{HO}dU-Modified Deoxyribozymes for Amide Hydrolysis

For the JY1 selection using ^{HO}dU, the capture yield was 16% at round 14 (Figure 2.12), and the active pool was cloned from that round. Three unique sequences were identified and designated as AmideHy1, AmideHy2 and AmideHy3 (Figure 2.13). In these three deoxyribozymes, there are 7, 9 and 6 ^{HO}dU modifications, respectively, in the initially random (N_{40}) region (Figure 2.14A). AmideHy1 and AmideHy2 had k_{obs} 0.08–0.09 h^{-1} and ~50% yield at 48 h, and AmideHy3 had k_{obs} 0.17 h^{-1} and 25% yield at 48 h (Figure 2.14B,C). The amide hydrolysis activity was confirmed by MALDI mass spectrometry (Figure 2.14D). As for the ^{Am}dU-modified deoxyribozymes, assays were performed to assess the modification requirements in the 3′-binding arm for these ^{HO}dU-modified deoxyribozymes (Figure 2.15). The results suggested that ^{HO}dU modifications in 3′-binding arm are not required for catalysis.

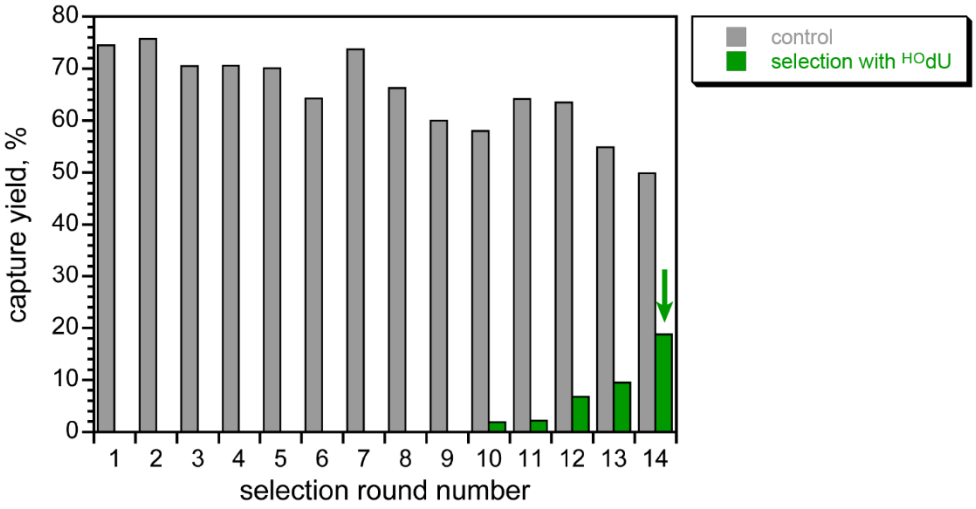

Figure 2.12. Progression of the JY1 selection with ^{HO}dU. In each round, "control" refers to the EDC-promoted capture yield for the product that has a 3′-COOH group, and "selection" refers to the EDC-promoted capture yield using the deoxyribozyme pool for that round. Arrow marks the cloned round.

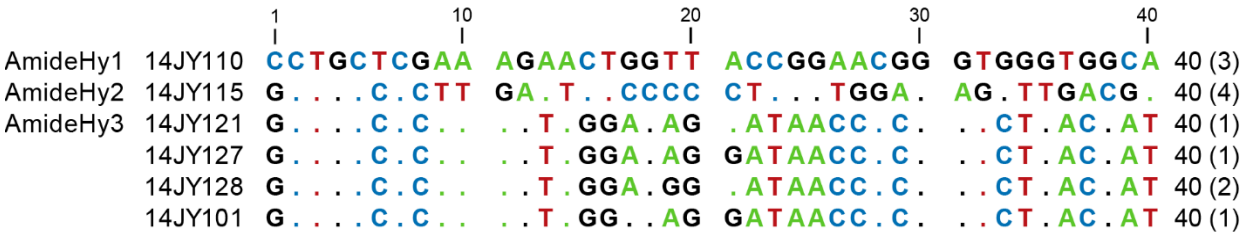

Figure 2.13. Sequences of the initially random region of the ^{HO}dU-modified deoxyribozymes for amide hydrolysis. All deoxyribozymes were used as 5′-CGAAGTCGCCATCTCTTC-N_{40}-ATAGTGAGTCGTATTA-3′, where N_{40} represents the specific 40 nucleotides of the initially random region as determined through the selection process. All alignments show only the initially random region. A dot denotes conservation, i.e., the same nucleotide as in the uppermost sequence. Next to each sequence length in nucleotides on the far right is shown (in parentheses) the number of times that sequence was found during cloning.

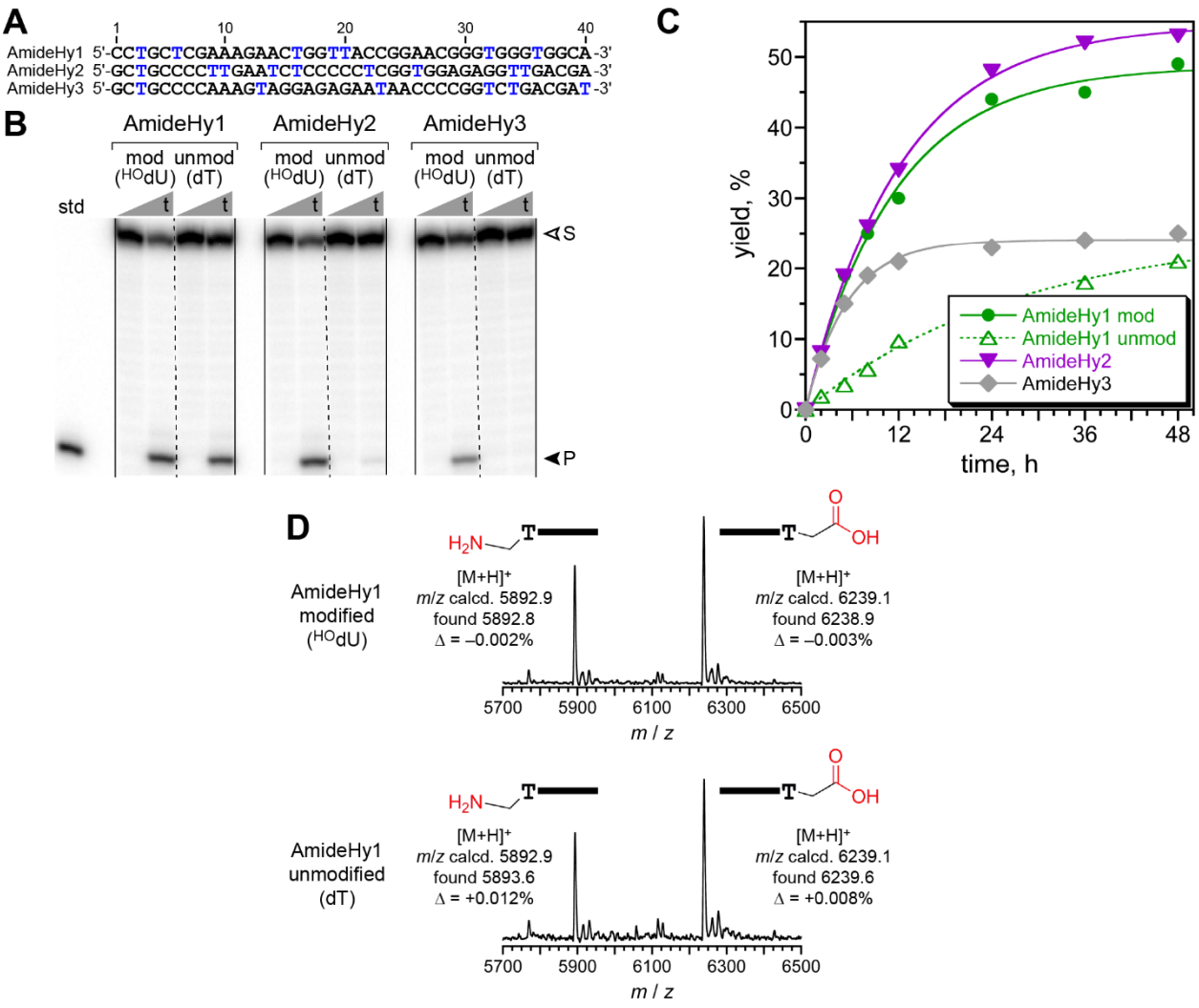

Figure 2.14. ^{HO}dU-modified deoxyribozymes for amide hydrolysis. (A) Sequence of the initially random (N_{40}) regions of AmideHy1, 2, and 3. Each blue T is ^{HO}dU in the deoxyribozyme. (B) PAGE assay (t = 0, 48 h). (C) Kinetic plots. k_{obs}, h^{-1} (n = 3): AmideHy1 mod 0.089 ± 0.011 and unmod 0.034 ± 0.003; AmideHy2 0.084 ± 0.004; AmideHy3 0.17 ± 0.02. (D) Mass spectrometry analysis of products (similar for AmideHy2,3; not shown).

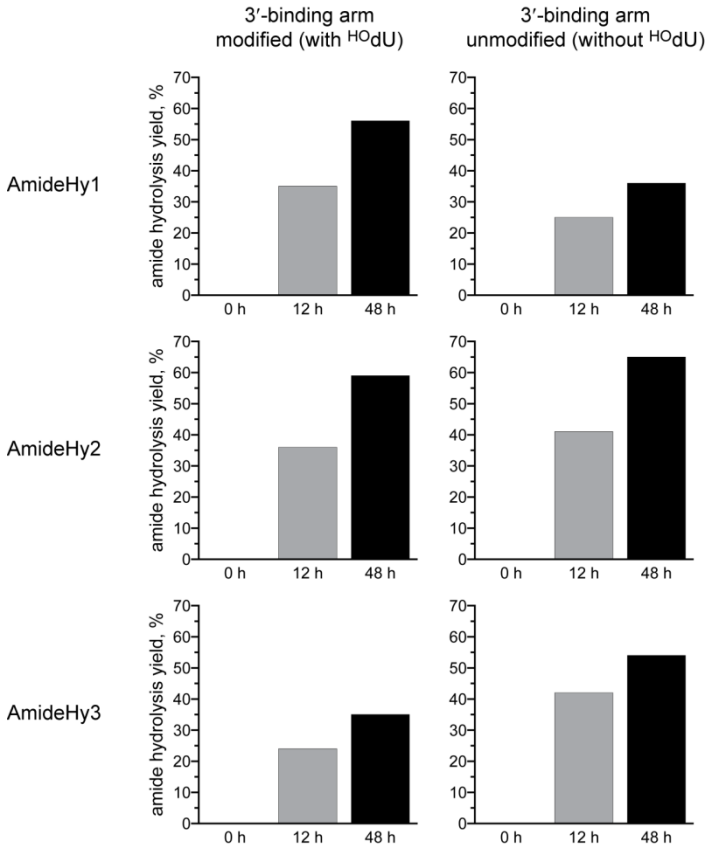

Figure 2.15. Assays of ^{HO}dU requirement in 3′-binding arm for ^{HO}dU-modified deoxyribozymes. The results reveal that ^{HO}dU modifications are not required in the 3′-binding arm for any of these three deoxyribozymes. Indeed, for AmideHy3 the yield was higher without 3′-binding arm modifications.

When prepared with unmodified dT instead of modified ^{HO}dU nucleotides, this unmodified version of AmideHy1 (without any primary hydroxyl modifications) was unexpectedly found to retain substantial catalytic activity (20% yield in 48 h; Figure 2.14B,C), and the amide hydrolysis products were confirmed by MALDI mass spectrometry (Figure 2.14D). Similarly, the unmodified version of AmideHy2 also showed a trace of amide hydrolysis activity (1.2% yield in 48 h; Figure 2.14B). These findings are surprising in the context of our prior efforts, because unmodified DNA was never observed to catalyze amide bond hydrolysis.[1,4] One hypothesis, not yet tested experimentally, is that when changing ^{HO}dU to dT, the 5-hydroxymethyl group in ^{HO}dU can be replaced by the 5-methyl group of dT and a water molecule. In some cases, the combination of the water molecule and dT could retain similar function as the catalytically relevant ^{HO}dU in the

deoxyribozyme. This would be analogous to abasic and other rescue experiments for ribozymes[16–18] and similar rescue studies for protein mutants.[19,20] This "rescue effect" of water could be relatively efficient for AmideHy1 and much less efficient for AmideHy2, thus explaining the amide hydrolysis activities for the unmodified version of AmideHy1 and AmideHy2.

2.2.2.3 ^{COOH}dU-Modified Deoxyribozymes for Amide Hydrolysis

For JX1 selection using ^{COOH}dU, the capture yield was 7% at round 11 (Figure 2.16), and the subsequent cloning process identified two unique sequences, designated as AmideCa1 and AmideCa2 (Figure 2.17). These two deoxyribozymes included 7 and 3 ^{COOH}dU modifications, respectively, in the initially random region (Figure 2.18A). AmideCa1 and AmideCa2 had k_{obs} 0.3–0.4 h^{-1} and 10–17% yield at 48 h for amide hydrolysis (Figure 2.18B,C). Assays using variants prepared by splint ligation revealed that ^{COOH}dU modifications in the 3′-binding arm are not needed for amide hydrolysis; in fact, yields were higher (16–24%) when ^{COOH}dU modifications were omitted (Figure 2.18B,C). No amide hydrolysis activity was observed for unmodified AmideCa1 and AmideCa2(Figure 2.18B).

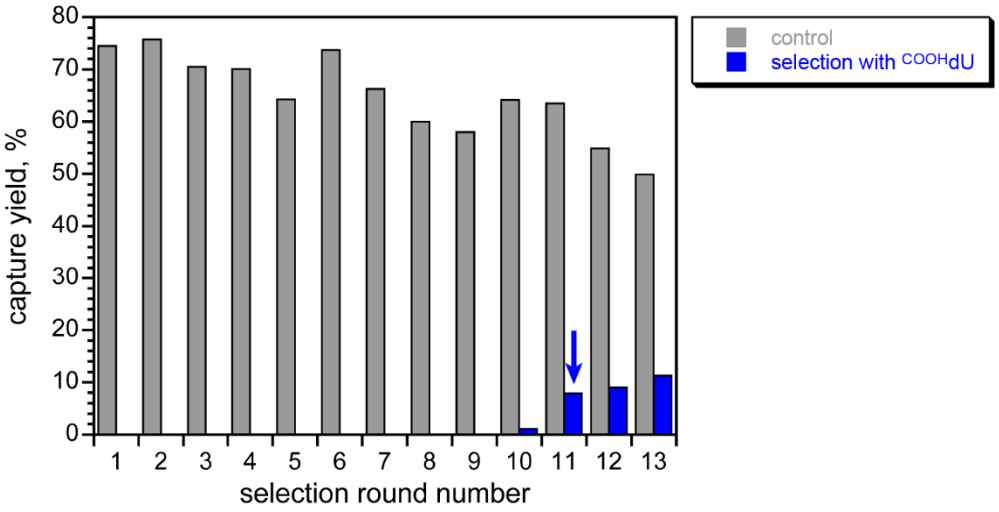

Figure 2.16. Progression of the JX1 selection with ^{COOH}dU. In each round, "control" refers to the EDC-promoted capture yield for the product that has a 3′-COOH group, and "selection" refers to the EDC-promoted capture yield using the deoxyribozyme pool for that round. Arrow marks the cloned round.

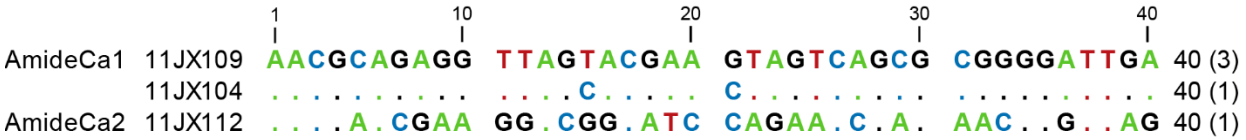

Figure 2.17. Sequences of the initially random region of the ^{COOH}dU-modified deoxyribozymes for amide hydrolysis. All deoxyribozymes were used as 5′-CGAAGTCGCCATCTCTTC-N_{40}-ATAGTGAGTCGTATTA-3′, where N_{40} represents the specific 40 nucleotides of the initially random region as determined through the selection process. All alignments show only the initially random region. A dot denotes conservation, i.e., the same nucleotide as in the uppermost sequence. Next to each sequence length in nucleotides on the far right is shown (in parentheses) the number of times that sequence was found during cloning.

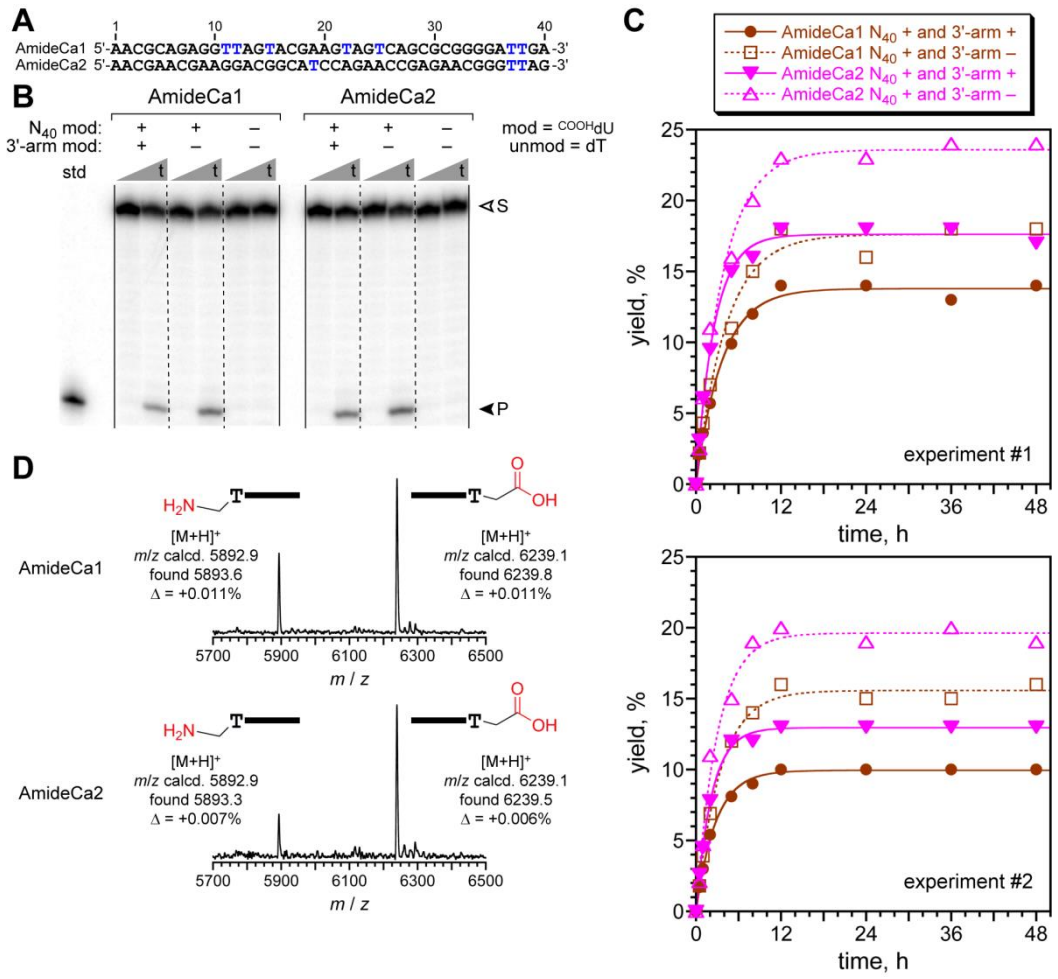

Figure 2.18. ^{COOH}dU-modified deoxyribozymes for amide hydrolysis. (A) Sequence of the initially random (N_{40}) regions of AmideCa1 and AmideCa2. Each blue T is ^{COOH}dU in the deoxyribozyme. (B) PAGE assay (t = 0, 48 h). The ^{COOH}dU modifications were either present (+) or absent (–) in the initially random region and 3′-binding arm as marked. (C) Kinetic plots. Shown are results from duplicate experiments performed on different days. k_{obs}, h^{-1}: AmideCa1 N_{40} + and 3′-arm + 0.27, 0.36; AmideCa1 N_{40} + and 3′-arm – 0.24, 0.29; AmideCa2 N_{40} + and 3′-arm + 0.39, 0.45; AmideCa2 N_{40} + and 3′-arm – 0.26, 0.34. (D) Mass spectrometry analysis of products.

2.2.2.4 Metal Ion and pH Dependence of Deoxyribozymes for Amide Hydrolysis

Modified RNA-cleaving deoxyribozymes can have reduced divalent metal ion dependence.[8–10] Although such a feature was not our goal, we were still interested to see whether our deoxyribozymes could work when omitting some or even all divalent metal ions. Therefore, metal ion dependence was investigated for these amide-hydrolyzing deoxyribozymes (Figure 2.19).

All of the deoxyribozymes require Zn^{2+} for full activity, in most cases additionally requiring Mn^{2+} or Mg^{2+}. For AmideHy2, only Zn^{2+} was required for high activity.

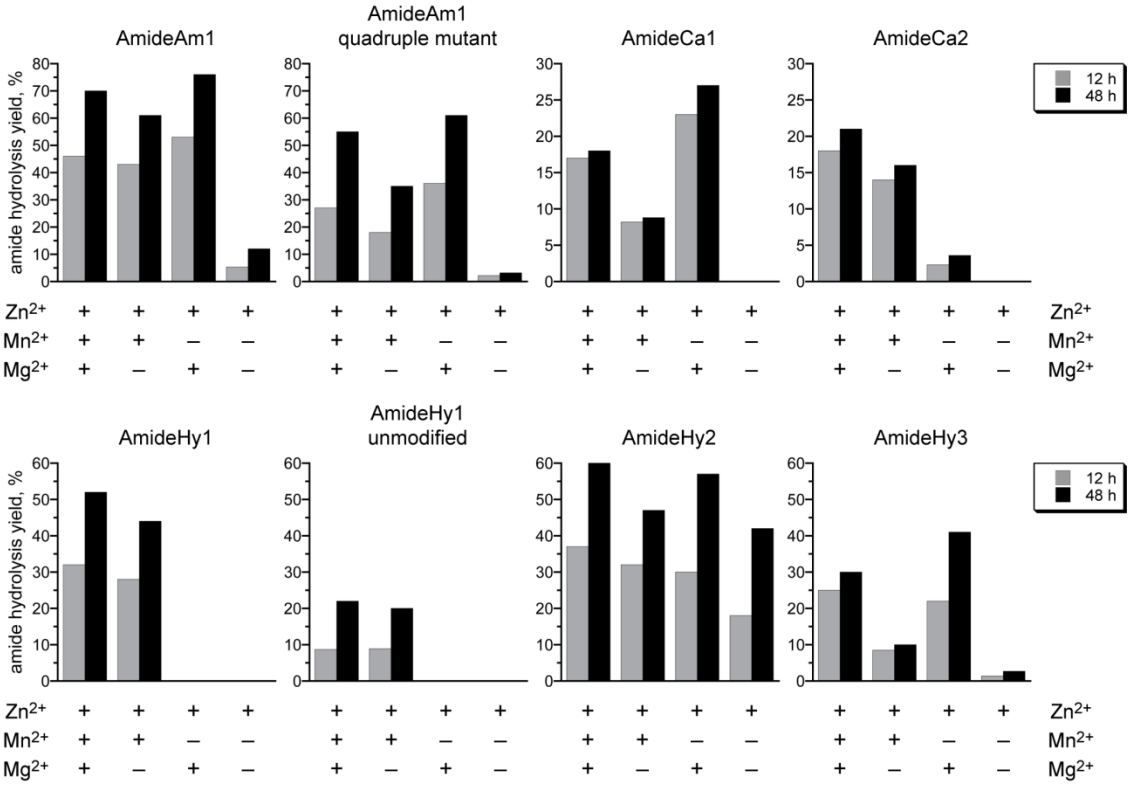

Figure 2.19. Divalent metal ion dependence of the amide-hydrolyzing deoxyribozymes. Incubation conditions were 70 mM HEPES, pH 7.5, combinations of 1 mM $ZnCl_2$, 20 mM $MnCl_2$, and 40 mM $MgCl_2$ as indicated, and 150 mM NaCl at 37 °C. All deoxyribozymes were prepared by primer extension with modifications throughout the sequence, including the 3′-binding arm. For each of the eight deoxyribozymes, omitting all three metal ions led to no observed amide hydrolysis (<0.5%).

For these deoxyribozymes, another possible key factor is pH, especially considering the different pK_a of primary amino, primary hydroxyl, and carboxylic acid modifications. Assays showed that all of the deoxyribozymes have reduced activity below pH 7.5, with little or no activity at pH 7.0. We did not test pH above 7.5, because Zn^{2+} precipitates and Mn^{2+} oxidizes at higher pH.

2.3 Summary and Future Directions

Although DNA-catalyzed ester and aromatic amide hydrolysis was achieved in previous efforts, the aliphatic amide bond hydrolysis was still elusive despite extensive attempts using unmodified DNA. Here, an in vitro selection strategy using modified DNA was developed for DNA-catalyzed aliphatic amide hydrolysis. Intriguingly, for the first time, deoxyribozymes were identified to catalyze the aliphatic amide bond hydrolysis, showing that including modified nucleotides with protein-like functional groups enables the identification of amide-hydrolyzing deoxyribozymes.

DNA with modified nucleotides is still DNA from the practical viewpoint of in vitro selection. Because modified nucleotides can be tolerated by appropriate DNA polymerases, the in vitro process is still fundamentally possible. Therefore, DNA that contains modified nucleotides with protein-like functional groups could have expanded chemical functionality like protein, while still maintaining the advantages of DNA for in vitro selection.

Besides the identification of modified deoxyribozymes for amide hydrolysis, the finding that unmodified AmideHy1 can have substantial amide hydrolysis activity is also exciting. However, performing selections using modified DNA is not a reliable way of finding unmodified deoxyribozymes, and this special case for amide hydrolysis is difficult to expand to other studies.

In the present study, the amide substrate was just a model substrate with a simple amide bond embedded between two DNA anchors. The subsequent efforts (described in Chapter 3) are focused on expanding this DNA-catalyzed amide hydrolysis activity to peptide substrates. The long-term goal is to achieve DNA-catalyzed amide hydrolysis for peptide and protein substrates with sequence selectivity and free peptide/protein activity,[21,22] which will have valuable applications in molecular biology, chemical biology, and proteomics.[23,24]

2.4 Materials and Methods

2.4.1 Preparation of Oligonucleotides and Amide Substrate

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), extracted from the polyacrylamide with TEN buffer (10 mM Tris, pH 8.0, 1 mM EDTA, 300 mM NaCl), and precipitated with ethanol as described previously.[25] The amide substrate was synthesized using procedures similar to those used to prepare the ester-linked substrate in our previous report.[26]

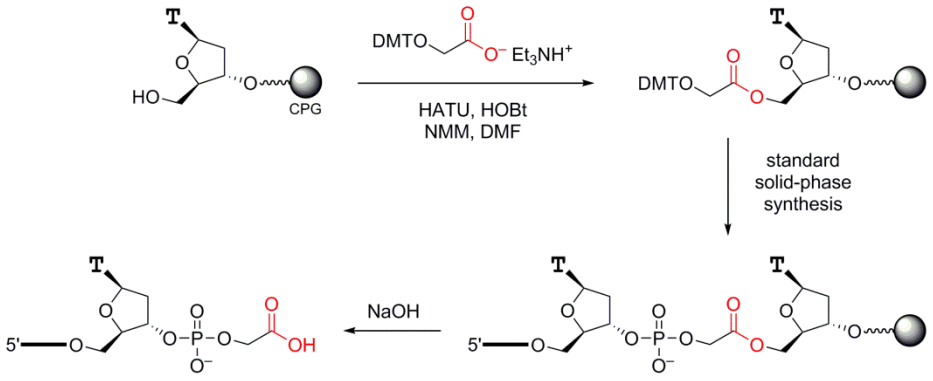

Figure 2.20. Solid-phase synthesis of 3′-CO_2H oligonucleotide.

Synthesis of 3′-CO_2H oligonucleotide. 5′-DMT-thymidine-derivatized CPG solid support (40 µmol/g; 5 mg, 200 nmol) was deprotected with 3% (v/v) TCA/CH_2Cl_2 (3 × 1 mL) and washed with CH_2Cl_2 (3 × 5 mL). The support was rinsed with 2% (v/v) NMM/CH_2Cl_2 and CH_2Cl_2 (3 × 5 mL) to improve the efficiency of the next coupling step and dried under a stream of nitrogen. *O*-(4,4′-Dimethoxytrityl)glycolate triethylammonium salt (48 mg, 100 µmol; prepared as described in section 2.4.8), HATU (38 mg, 100 µmol), and HOBt (14 mg, 100 µmol) were dissolved in 500 µL of dry DMF. NMM (26 µL, 230 µmol) was added to this solution, which was introduced into the DNA synthesis column via two syringes. The coupling step was performed for 16 h with mixing every 5 min for the first 1 h and no mixing for the remaining 15 h. The column was washed

with DMF (3 × 5 mL) and CH_2Cl_2 (3 × 5 mL) and dried under a stream of nitrogen. The remainder of the oligonucleotide was synthesized on the ABI 394 instrument using standard procedures. Cleavage of the oligonucleotide from the CPG and nucleobase deprotection were performed using 200 µL of 400 mM NaOH in 4:1 MeOH:H_2O for 24 h at room temperature. The sample was diluted to 600 µL with water and precipitated with ethanol. The deprotection was completed by treating the sample with concentrated ammonium hydroxide for 16 h at 55 °C. The sample was dried under vacuum and purified by 20% PAGE. 50 nmol of the product was obtained. MALDI-MS $[M+H]^+$ calcd. 6239.1, found 6237.7.

Synthesis of 5′-amino-5′-deoxythymidine oligonucleotide. Oligonucleotide synthesis was performed on the ABI 394 instrument using standard procedures and MMT-protected 5′-amino-5′-deoxythymidine phosphoramidite (Glen Research), providing an oligonucleotide terminating in a 5′-amino-5′-deoxythymidine residue.

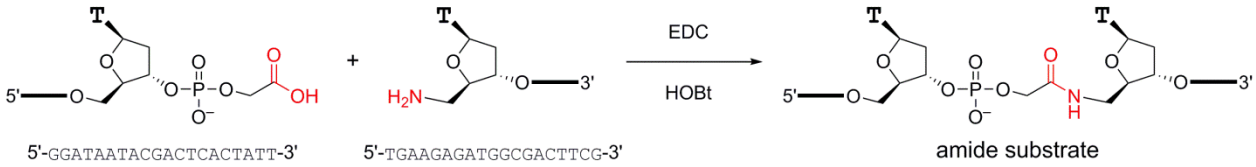

Figure 2.21. Synthesis of amide substrate by splinted coupling of 3′-CO_2H oligonucleotide and 5′-amino-5′-deoxythymidine oligonucleotide.

Solution-phase coupling of 3′-CO_2H oligonucleotide and 5′-amino-5′-deoxythymidine oligonucleotide. A 45 µL sample containing 4.0 nmol of 3′-CO_2H oligonucleotide, 6.0 nmol of 5′-amino-5′-deoxythymidine oligonucleotide, and 5.0 nmol of DNA splint was annealed in 110 mM MOPS, pH 7.0, and 1.1 M NaCl by heating at 95 °C for 3 min and cooling on ice for 5 min. The sequence of the DNA splint was 5′-CGAAGTCGCCATCTCTTCA**T**AATAGTGAGTCGTATTA-TCCAACAACAACAACAACAAC-3′, where the boldface **T** is unpaired, and the underlined region is included to shift the splint away from the desired product on PAGE. The reaction was initiated by bringing the sample to 50 µL volume containing 100 mM MOPS, pH 7.0, 1 M NaCl, 50 mM EDC, 50 mM HOBt, and 5% (v/v) DMF, which was from the HOBt stock solution and was required to

dissolve the HOBt. The sample was incubated at room temperature for 24 h, precipitated with ethanol, and purified by 20% PAGE. 2.8 nmol of the product was obtained. MALDI-MS $[M+H]^+$ calcd. 12113.0, found 12113.8.

2.4.2 2′-Deoxynucleotide 5′-Triphosphates and 2′-Deoxynucleoside 3′-Phosphoramidites

The ^{Am}dU 5′-triphosphate (^{Am}dUTP) was from TriLink BioTechnologies (cat. no. N-0249). The ^{HO}dU 5′-triphosphate (^{HO}dUTP) was from TriLink BioTechnologies (cat. no. N-0259). The ^{Am}dU phosphoramidite was from Berry & Associates (cat. no. BA0311). The ^{COOH}dU phosphoramidite was from Glen Research (cat. no. 10-1035). The ^{HO}dU phosphoramidite was from Glen Research (cat. no. 10-1093). The ^{COOH}dU 5′-triphosphate (^{COOH}dUTP) and ^{Im}dU 5′-triphosphate (^{Im}dUTP) were synthesized as described in section 2.4.8.

2.4.3 In Vitro Selection Procedure

The selection procedure, cloning, and initial screening of individual clones were performed essentially as described previously,[27] with additional steps to enable incorporation of the modified nucleotides. An overview of the key selection and capture steps of each round is shown in Figure 2.6. A depiction of these steps with nucleotide details is shown in Figure 2.22. The random deoxyribozyme pool was 5′-CGAAGTCGCCATCTCTTC-N_{40}-ATAGTGAGTCGTATTAAGCTGAT-CCTGATGG-3′, where the initially random N_{40} region includes modified nucleotides. Each of the ^{Am}dU, ^{HO}dU and ^{COOH}dU random pools was prepared by solid-phase synthesis, replacing all dT in the N_{40} region with one of ^{Am}dU, ^{HO}dU, or ^{COOH}dU using the corresponding 2′-deoxynucleoside 3′-phosphoramidite. The ^{Im}dU random pool was prepared by primer extension from a reverse complement template pool using synthesized ^{Im}dUTP and KOD XL polymerase. The reverse complement template pool was 5′-CCATCAGGATCAGCTTAATACGACTCACTAT-N_{40}-GAAGAGATGGCGACTTCGAGATCACGTCGATAACAACAACAACAACAACAACAACAAC-3′, where the underlined nucleotides are arbitrary sequence included to allow PAGE separation of the primer extension product from the template. Primers were 5′-CGAAGTCGCCATCTCTTC-3′ (forward

primer) and 5′-(AAC)$_{18}$XCCATCAGGATCAGCT-3′, where X is the HEG spacer to stop Taq polymerase (reverse primer). The (AAC)$_{18}$ tail was especially long to assure PAGE separation of template and ^{Am}dU/^{Im}dU-modified product strands. For convenience, the same tail length was used with the ^{HO}dU and ^{COOH}dU modifications, although the more standard (AAC)$_4$ tail can be used instead. In each round, the ligation step to attach the deoxyribozyme pool at its 3′-end with the 5′-end of the amide substrate was performed using a DNA splint and T4 DNA ligase. The ligation splint sequence was 5′-ATAGTGAGTCGTATTATCCTCCATCAGGATCAGCTTAA–TACGACTCACTAT-3′, where the underlined T is included to account for the untemplated A nucleotide that is added at the 3′-end of each PCR product by KOD XL polymerase. This T nucleotide was omitted from the splint used for ligation of the initially random N_{40} pool, which was prepared by solid-phase synthesis without the untemplated A. Nucleotide sequences of the amide substrate, the deoxyribozyme binding arms, the 5′-amino capture oligonucleotide, and the capture splint are shown in Figure 2.22. T4 DNA ligase was from Thermo Fisher. KOD XL polymerase was from Novagen.

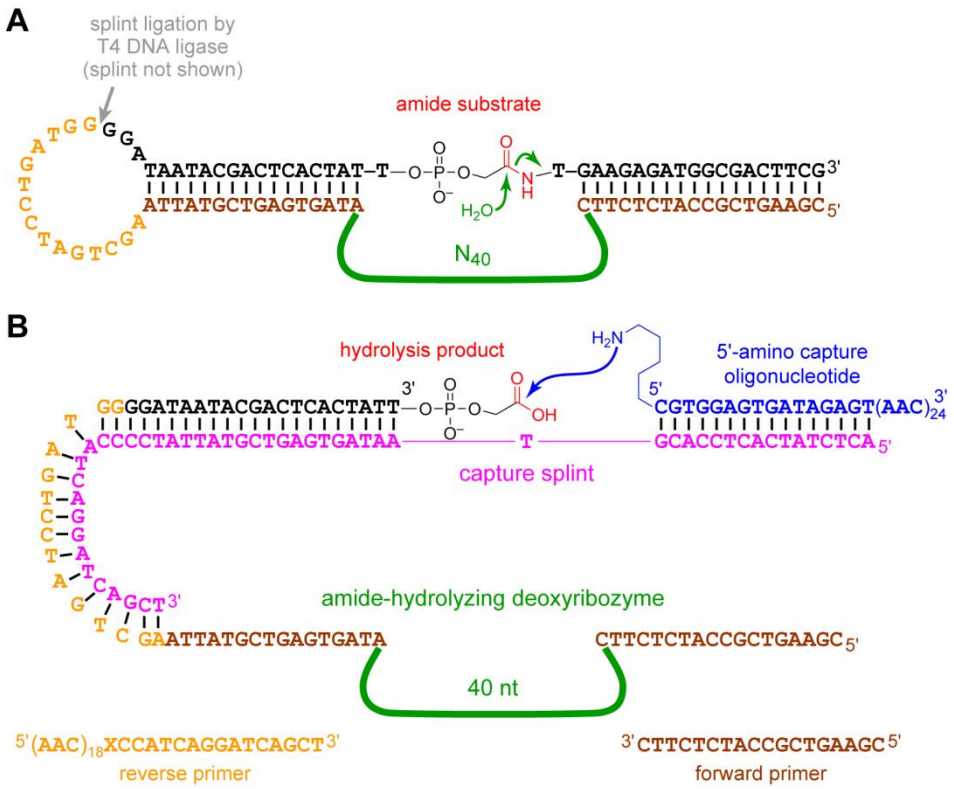

Figure 2.22. Nucleotide details of the selection and capture steps of in vitro selection. (A) Selection step. (B) Capture step. See text for details of PCR and primer extension using the forward and reverse primers. In the reverse primer, X is the HEG spacer (Glen Spacer 18) that stops extension by DNA polymerase and leads to a size difference between the two PCR product strands. The capture splint enforces proximity of the 5′-amino group of the capture oligonucleotide and the carboxyl group of the hydrolysis product, such that amino or carboxyl groups within the initially random region or 3′-binding arm do not react appreciably during the capture step incubation time.

Procedure for primer extension to prepare the ^{Im}*dU random pool.* A 50 µL sample was prepared containing 250 pmol of the reverse complement template pool (which contains a 3′-segment of arbitrary sequence to allow PAGE separation of the primer extension product from the template), 500 pmol of forward primer, 15 nmol each of dATP, dGTP, dCTP, and ^{Im}dUTP, 5 µL of 10× KOD XL polymerase buffer, and 1 µL of 2.5 U/µL KOD XL polymerase. Primer extension was performed by cycling 4 times in a PCR thermocycler according to the following program: 94 °C for 2 min, 4× (94 °C for 2 min, 47 °C for 2 min, 72 °C for 30 min), 72 °C for 30 min. The sample was separated by 8% PAGE to provide the ^{Im}dU random pool for the selection.

Procedure for ligation step in round 1. A 25 µL sample containing 600 pmol of ^{X}dU-modified DNA pool (X = Am, HO, COOH, or Im), 700 pmol of DNA splint, and 800 pmol of amide 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 and 2 µL of 5 U/µL T4 DNA ligase. The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

Procedure for ligation step in subsequent rounds. A 16 µL sample containing the modified DNA pool synthesized by primer extension (~5–10 pmol), 25 pmol of DNA splint, and 50 pmol of amide substrate was annealed in 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. To this solution was added 2 µL of 10× T4 DNA ligase buffer and 2 µL of 1 U/µL T4 DNA ligase. The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

Procedure for selection step in round 1. The selection experiment was initiated with 200 pmol of the ligated N_{40} pool. A 20 µL sample containing 200 pmol of ligated N_{40} pool was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The selection reaction was initiated by bringing the sample to 40 µL total volume containing 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl. The 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 rounds. A 10 µL sample containing ligated pool was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The selection reaction was initiated by bringing the sample to 20 µL total volume containing 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl. The sample was incubated at 37 °C for 14 h and precipitated with ethanol.

Procedure for alkaline treatment in round 1 of ^{HO}dU selection. For the selection experiment that included ^{HO}dU, a brief alkaline treatment was performed between the selection and capture steps. This treatment accommodated the possibility that a hydroxyl group from a ^{HO}dU nucleotide could serve in place of water as the nucleophile, thereby leading to a macrolactone intermediate that would require subsequent hydrolysis to complete the cleavage process. The precipitated selection sample was dissolved in 45 µL of water, to which was added 5 µL of 500 mM NaOH (50 mM final NaOH concentration). The sample was incubated at room temperature for 30 min and precipitated with ethanol.

Procedure for alkaline treatment in subsequent rounds of ^{HO}dU selection. The precipitated selection sample was dissolved in 9 µL of water, to which was added 1 µL of 500 mM NaOH (50 mM final NaOH concentration). The sample was incubated at room temperature for 30 min and precipitated with ethanol.

Procedure for capture step in round 1. A 90 µL sample containing the selection product, 300 pmol of capture splint, and 500 pmol of 5′-amino capture oligonucleotide was annealed in 100 mM MOPS, pH 7.0, and 1 M NaCl by heating at 95 °C for 3 min and cooling on ice for 5 min. The capture reaction was initiated by bringing the sample to 100 µL total volume containing 100 mM MOPS, pH 7.0, 1 M NaCl, 100 mM EDC, 100 mM HOBt, and 10% (v/v) DMF, which was from the HOBt stock solution and was required to dissolve the HOBt. The sample was incubated at room temperature for 15 h, precipitated with ethanol, and separated by 8% PAGE.

Procedure for capture step in subsequent rounds. A 22.5 µL sample containing the selection product, 30 pmol of capture splint, and 50 pmol of 5′-amino capture oligonucleotide was annealed in 100 mM MOPS, pH 7.0, and 1 M NaCl by heating at 95 °C for 3 min and cooling on ice for 5 min. The capture reaction was initiated by bringing the sample to 25 µL total volume containing 100 mM MOPS, pH 7.0, 1 M NaCl, 50 mM EDC, 50 mM HOBt, and 5% (v/v) DMF, which was from the HOBt stock solution and was required to dissolve the HOBt. The sample was incubated at room temperature for 15 h and loaded directly on 8% PAGE. In parallel, during each round a standard capture reaction was performed using the N_{40} pool ligated to the hydrolysis

product. The yield is shown as "control" in Figure 2.7, Figure 2.12 and Figure 2.16. The PAGE position of this standard capture product was used as a reference to excise the capture product for the selection sample.

Procedure for PCR. In each selection round, two PCR reactions were performed, 10-cycle PCR followed by 30-cycle PCR. First, a 100 µL sample was prepared containing the PAGE-purified capture product, 50 pmol of forward primer, 200 pmol of reverse primer, 20 nmol of each dNTP, 10 µL of 10× KOD XL polymerase buffer, and 2 µL of 2.5 U/µL KOD XL polymerase. This sample was primer-extended and then cycled 10 times according to the following PCR program: 94 °C for 2 min, 47 °C for 2 min, 72 °C for 30 min; then 94 °C for 2 min, 10× (94 °C for 1 min, 47 °C for 1 min, 72 °C for 1 min), 72 °C for 5 min. KOD XL polymerase was removed by phenol/chloroform extraction. Second, a 50 µL sample was prepared containing 1 µL of the 10-cycle PCR product, 25 pmol of forward primer, 100 pmol of reverse primer, 10 nmol of each dNTP, 20 µCi of α-^{32}P-dCTP (800 Ci/mmol), 5 µL of 10× Taq polymerase buffer, and 0.5 µL of Taq polymerase (no modified nucleotides are being incorporated). This sample was cycled 30 times according to the following PCR program: 94 °C for 2 min, 30× (94 °C for 30 s, 47 °C for 30 s, 72 °C for 30 s), 72 °C for 5 min. The sample was separated by 8% PAGE, and the reverse-complement single strand was isolated for subsequent primer extension. The forward and reverse single-stranded PCR products were separable because formation of the reverse product was initiated using the reverse primer, which contains a nonamplifiable spacer that stops Taq polymerase.

Procedure for primer extension to generate the modified DNA pool in subsequent rounds. A 25 µL sample was prepared containing the reverse-complement single strand (which contains a nonamplifiable spacer that stops KOD XL polymerase), 50 pmol of forward primer, 7.5 nmol each of dATP, dGTP, dCTP, and modified ^{X}dUTP (X = Am, HO, COOH, or Im), 20 µCi of α-^{32}P-dCTP (800 Ci/mmol), 2.5 µL of 10× KOD XL polymerase buffer, and 0.5 µL of 2.5 U/µL 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 to provide the modified DNA pool for the next selection round.

Procedure for cloning and initial screening. Using 1 μL of a 1:1000 dilution of the 10-cycle PCR product from the relevant selection round, 30-cycle PCR was performed using the same procedure as described, omitting α-^{32}P-dCTP and using primers 5′-CGAAGTCGCCATCTCTTC-3′ (forward primer, 25 pmol) and 5′-TAATTAATTAATTACCCATCAGGATCAGCT-3′ (reverse primer, 25 pmol), where the extensions with TAA stop codons in each frame were included to suppress false negatives in blue-white screening.[3] 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) was performed to synthesize the reverse-complement templates for the subsequent primer extension reactions. For primer extension, 25 μL samples were prepared, each containing a reverse-complement template, 100 pmol of forward primer, 5.0 nmol each of dATP, dGTP, dCTP, and modified ^{X}dUTP (X = Am, HO, COOH, Im), 2 μCi of α-^{32}P-dCTP (800 Ci/mmol), 2.5 μL of 10× KOD XL polymerase buffer, and 0.5 μL of 2.5 U/μL KOD XL polymerase. Primer extension was performed in a PCR thermocycler according to the following program: 94 °C for 2 min, 4× (94 °C for 2 min, 47 °C for 2 min, 72 °C for 1 h), 72 °C for 1 h. The samples were separated by 8% PAGE to provide the modified deoxyribozymes. The concentration of each PAGE-purified deoxyribozyme strand was estimated from the radioactive intensity relative to suitable standards. Each screening assay used ~0.1 pmol of 5′-^{32}P-radiolabeled amide substrate and at least 10 pmol of deoxyribozyme in the single-turnover assay procedure described in section 2.4.6.

2.4.4 Procedure for Preparing Modified Deoxyribozymes by Primer Extension

A 50 μL sample was prepared containing 400 pmol of the reverse-complement template 5′-TAATACGACTCACTAT-$\mathbf{N}_{40}$-GAAGAGATGGCGACTTCGGAAAGAGCAATAGTAA-3′ (the boldface $\mathbf{N}_{40}$ is complementary to the 40 nt initially random region in the deoxyribozyme; the

underlined nucleotides are arbitrary sequence included to allow PAGE separation of the primer extension product from the template), 800 pmol of forward primer, 15 nmol each of dATP, dGTP, dCTP, and modified ^{X}dUTP (X = Am, HO, COOH), 5.0 μL of 10× KOD XL polymerase buffer, and 1.0 μL of 2.5 U/μL KOD XL polymerase. Primer extension was performed in a PCR thermocycler according to the following program: 94 °C for 2 min, 4× (94 °C for 2 min, 47 °C for 2 min, 72 °C for 1 h), 72 °C for 1 h. The sample was separated by 12% PAGE to provide the modified deoxyribozyme.

2.4.5 Procedure for Preparing AmideAm1 without 3′-Binding Arm Modifications

To assess functional contributions of ^{Am}dU nucleotides within the 3′-binding arm of AmideAm1 (Figure 2.10), we used splint ligation to prepare a version of this deoxyribozyme without 3′-binding arm modifications. The 5′-segment (i.e., 5′-binding arm + initially random 40 nucleotides) was prepared by primer extension using the forward primer from Figure 2.22 and reverse complement template 5′-**CCACGTACTAGCGATTCGCGTGACTGCGTAGTCCCGCTCT**–GAAGAGATGGCGACTTCGGAAAGAGCAATAGTAA-3′, where the boldface nucleotides are complementary to the initially random 40 nucleotides, the standard nucleotides are complementary to the 5′-binding arm (binding site for forward primer), and the underlined nucleotides are arbitrary sequence included to allow PAGE separation of the primer extension product from the template. The 3′-segment 5′-AGAGTGAGTCGTATTA-3′ was prepared by solid-phase synthesis. The splint sequence was 5′-TAATACGACTCACTAT**CCACGTACTAGCGATTCGCG**-3′. To prepare the deoxyribozyme without 3′-binding arm modifications, a 26 μL sample containing 200 pmol of the 5′-segment, 300 pmol of DNA splint, and 400 pmol of 3′-segment 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 and 1 μL of 5 U/μL T4 DNA ligase. The sample was incubated at 37 °C for 12 h and separated by 12% PAGE.

2.4.6 Single-Turnover Deoxyribozyme Assay Procedure

The amide substrate was 5′-^{32}P-radiolabeled using γ-^{32}P-ATP and T4 polynucleotide kinase. A 10 μL sample containing 0.2 pmol of 5′-^{32}P-radiolabeled amide substrate and 20 pmol of deoxyribozyme was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The DNA-catalyzed amide hydrolysis reaction was initiated by bringing the sample to 20 μL total volume containing 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl. The sample was incubated at 37 °C. At appropriate time points, 2 μL aliquots were quenched with 5 μL stop solution (80% formamide, 1× TBE [89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3], 50 mM EDTA, 0.025% bromophenol blue, 0.025% xylene cyanol). Samples were separated by 20% PAGE and quantified with a PhosphorImager. Values of k_{obs} were obtained by fitting the yield versus time data directly to first-order kinetics; i.e., yield = Y•$(1 - e^{-kt})$, where $k = k_{obs}$ and Y is the final yield. Each k_{obs} value is reported with error calculated as the standard deviation from the indicated number of independent determinations.

2.4.7 Mass Spectrometry Procedure

To prepare the cleavage products using a deoxyribozyme, a 25 μL sample containing 100 pmol of amide substrate and 120 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 amide hydrolysis reaction was initiated by bringing the sample to 50 μL total volume containing 70 mM HEPES, pH 7.5, 1.0 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl. The sample was incubated at 37 °C for 48 h, precipitated with ethanol, desalted by Millipore C_{18} ZipTip, and analyzed by MALDI mass spectrometry. Data were acquired on a Bruker UltrafleXtreme MALDI-TOF mass spectrometer with matrix 3-hydroxypicolinic acid in positive ion mode at the UIUC School of Chemical Sciences Mass Spectrometry Laboratory.

2.4.8 Organic Synthesis Procedures

Reagents were commercial grade and used without purification unless otherwise indicated. Dry pyridine was obtained from Acros Acroseal bottles. Thin-layer chromatography (TLC) was performed on silica gel plates pre-coated with fluorescent indicator, with visualization by UV light (254 nm). Flash column chromatography was performed with silica gel (230-400 mesh). NMR spectra were recorded on a Varian Unity instrument. The chemical shifts in parts per million (δ) are reported downfield from TMS or DSS (0 ppm) and referenced to the residual proton signal for $CDCl_3$ (7.26 ppm) or DSS (0 ppm). Apparent multiplicities of ^1H NMR peaks are reported as s (singlet), d (doublet), t (triplet), q (quartet), or m (multiplet and overlapping spin systems), along with values for apparent coupling constants (*J*, Hz). Mass spectrometry data were obtained at the UIUC School of Chemical Sciences Mass Spectrometry Laboratory using a Waters Quattro II instrument (LR-ESI).

Synthesis of O-*(4,4′-dimethoxytrityl)glycolate triethylammonium salt.* Glycolic acid (228 mg, 3.0 mmol; Alfa Aesar) was dissolved in dry pyridine (25 mL) under argon. 4,4′-Dimethoxytrityl chloride (1.21 g, 3.6 mmol; Sigma) and a catalytic amount of DMAP (35 mg, 0.3 mmol; Sigma) were added to the solution, which was stirred overnight at room temperature. A portion of CH_3OH (1.5 mL, 36 mmol) was added to quench any unreacted 4,4′-dimethoxytrityl chloride. The solution was diluted with ethyl acetate (100 mL) and washed with 80% (w/v) aqueous $NaHCO_3$ (30 mL) and saturated NaCl (30 mL). The organic layer was dried over anhydrous $MgSO_4$ and concentrated on a rotary evaporator. The oily residue was purified by silica gel column chromatography, eluting with 0–3% CH_3OH in CH_2Cl_2 containing 2% (v/v) Et_3N to yield the desired compound as a colorless oil (906 mg, 63%).

TLC: R_f = 0.19 [1% CH_3OH in CH_2Cl_2 with 2% (v/v) Et_3N].

^1H NMR: (500 MHz, $CDCl_3$) δ 7.54 (dd, *J* = 8.3, 1.3 Hz, 2H), 7.42 (d, *J* = 8.8 Hz, 4H), 7.24 (t, *J* = 7.9 Hz, 2H), 7.15 (tt, *J* = 7.3, <2 Hz, 1H), 6.78 (d, *J* = 8.8 Hz, 4H), 3.76 (s, 6H), 3.58 (s, 2H), 2.98 (q, *J* = 7.3 Hz, 6H), 1.22 (t, *J* = 7.3 Hz, 9H) ppm.

ESI-MS: *m/z* for carboxylic acid calcd. for $C_{23}H_{21}O_5$ $[M–H]^-$ 377.4; found 377.4.

Synthesis of ^{COOH}dU 5′-triphosphate (^{COOH}dUTP). The ^{COOH}dUTP was synthesized from (*E*)-5-(2-carbomethoxyvinyl)-2′-deoxyuridine (Berry & Associates, cat. no. PY7170). The triphosphorylation procedure was modified from the procedure reported by Huang and coworkers.[28] 2-Chloro-1,3,2-benzodioxaphosphorin-4-one (TCI, cat. no. C1210) was portioned into a dried vial in a glove box. 2′-Deoxyuridine and tributylammonium pyrophosphate (Sigma, cat. no. P8533) were dried under vacuum. Tributylamine was dried over 4 Å molecular sieves. Tributylamine (0.65 mL, 2.74 mmol) was added to a solution of tributylammonium pyrophosphate (188 mg, 0.34 mmol) dissolved in 1 mL of anhydrous DMF under argon. This solution was added into a stirred solution of 2-chloro-1,3,2-benzodioxaphosphorin-4-one (50 mg, 0.25 mmol) in 1 mL of anhydrous DMF under argon. The resulting solution was stirred at room temperature for 1 h and then added dropwise over 5 min to a solution of 2′-deoxyuridine (50 mg, 0.16 mmol) in 1 mL of anhydrous DMF under argon, cooled in an ice-salt bath at 0 to –10 °C. The solution was stirred for 5 h and raised to room temperature for 30 min. A solution of 0.02 M iodine in THF/pyridine/water (Glen Research, 8 mL) was added until a permanent brown color was maintained. The brown solution was stirred at room temperature for 30 min. Two sample volumes of water were added. The two-layer mixture was stirred vigorously at room temperature for 1.5 h, followed by addition of 0.1 volumes of 3 M NaCl and 3 volumes of ethanol. The thoroughly mixed sample was precipitated at –80 °C overnight and centrifuged for 30 min at 10,000 × *g* and 4 °C. The precipitated sample of ^{COOMe}dUTP was dried, dissolved in water, and purified by HPLC [Shimdazu Prominence instrument; Phenomenex Gemini-NX C_{18} column, 5 µm, 10 × 250 mm; gradient of 0.5% solvent A (20 mM triethylammonium acetate in 50% acetonitrile/50% water, pH 7.0) and 99.5% solvent B (20 mM triethylammonium acetate in water, pH 7.0) at 0 min to 25% solvent A and 75% solvent B at 45 min with flow rate of 3.5 mL/min]. To a solution of the ^{COOMe}dUTP in 50 µL of water was added 50 µL of 0.2 M NaOH, forming ^{COOH}dUTP. The sample was incubated at room temperature for 2 h, neutralized with 10 µL of 2.5 M NaOAc buffer, pH 5.2, and purified by HPLC (same gradient as before). The ^{COOH}dUTP was isolated as its triethylammonium salt [12 mg, 8.0% from (*E*)-5-(2-carbomethoxyvinyl)-2′-deoxyuridine].

^{1}H NMR: (500 MHz, D_2O with 0.01 mg/mL DSS) δ 8.05 (s, 1H), 7.17 (d, *J* = 16.0 Hz, 1H), 6.81 (d, *J* = 16.0 Hz, 1H), 6.29 (dd, *J* = 7.6, 6.3 Hz, 1H), 4.81 (D_2O), 4.64 (m, 1H), 4.21 (m, 3H), 3.19 (q, *J* = 7.3 Hz, 24H), 2.46 (m, 1H), 2.39 (ddd, *J* = 14.2, 6.4, 3.4 Hz, 1H), 1.27 (t, *J* = 7.3 Hz, 36H) ppm. From the integrations, 4 triethylammonium ions were present for each ^{COOH}dUTP molecule.
ESI-MS: *m/z* calcd. for $C_{12}H_{17}N_2O_{16}P_3$ $[M–H]^-$ 537.0; found 537.1.

Synthesis of ^{Im}dU. The (*E*)-5-(2-carbomethoxyvinyl)-2′-deoxyuridine *N*-hydroxysuccinimide ester (158 mg, 0.4 mmol) was dissolved in 2 mL of DMF. To this solution was added histamine dihydrochloride (147 mg, 0.8 mmol) and *N*,*N*-diisopropylethylamine (DIPEA; 278 μL, 1.6 mmol), and the solution was stirred overnight at room temperature. The solution was dried on a rotary evaporator. The residue was dissolved in 1 mL of acetone and dry-loaded on a silica gel column. The column was eluted with 5−20% CH_3OH in CH_2Cl_2 to yield the desired product as a white solid (88 mg, 56%).
ESI-MS: *m/z* calcd. for $C_{17}H_{21}N_5O_6$ $[M+H]^+$ 392.2; found 392.2.

Synthesis of ^{Im}dU 5′-triphosphate (^{Im}dUTP). The ^{Im}dUTP was synthesized from ^{Im}dU. The procedure was similar to the synthesis of ^{COOH}dUTP. The triphosphorylation procedure was modified from the procedure reported by Huang and coworkers.[28] 2-Chloro-1,3,2-benzodioxaphosphorin-4-one was portioned into a dried vial in a glove box. ^{Im}dU and tributylammonium pyrophosphate were dried under vacuum. Tributylamine was dried over 4 Å molecular sieves. Tributylamine (0.65 mL, 2.74 mmol) was added to a solution of tributylammonium pyrophosphate (188 mg, 0.34 mmol) dissolved in 1 mL of anhydrous DMF under argon. This solution was added into a stirred solution of 2-chloro-1,3,2-benzodioxaphosphorin-4-one (50 mg, 0.25 mmol) in 1 mL of anhydrous DMF under argon. The resulting solution was stirred at room temperature for 1 h and then added dropwise over 5 min to a solution of ^{Im}dU (67 mg, 0.17 mmol) in 1 mL of anhydrous DMF under argon, cooled in an ice-salt bath at 0 to –10 °C. The solution was stirred for 5 h and raised to room temperature for 30 min. A solution of 0.02 M iodine in THF/pyridine/water (Glen Research, 8 mL) was added until a permanent brown color was maintained. The brown solution was stirred at room temperature for

30 min. Two sample volumes of water were added. The two-layer mixture was stirred vigorously at room temperature for 1.5 h, followed by addition of 0.1 volumes of 3 M NaCl and 3 volumes of ethanol. The thoroughly mixed sample was precipitated at –80 °C overnight and centrifuged for 30 min at 10,000 × *g* and 4 °C. The precipitated sample of ^{Im}dUTP was dried, dissolved in water, and purified by HPLC [Shimdazu 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 40% solvent A and 60% solvent B at 45 min with flow rate of 3.5 mL/min]. The ^{Im}dUTP was isolated as its triethylammonium salt (14 mg, 8.8% from ^{Im}dU).

ESI-MS: *m/z* calcd. for $C_{17}H_{24}N_5O_{15}P_3$ $[M–H]^-$ 630.1; found 630.2.

2.5 References

(1) Chandra, M.; Sachdeva, A.; Silverman, S. K. *Nat. Chem. Biol.* **2009**, *5*, 718.

(2) Silverman, S. K. *Acc. Chem. Res.* **2015**, *48*, 1369.

(3) Schroeder, G. K.; Lad, C.; Wyman, P.; Williams, N. H.; Wolfenden, R. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4052.

(4) Brandsen, B. M.; Hesser, A. R.; Castner, M. A.; Chandra, M.; Silverman, S. K. *J. Am. Chem. Soc.* **2013**, *135*, 16014.

(5) Silverman, S. K. *Trends Biochem. Sci.* **2016**, *41*, 595.

(6) Jäger, S.; Rasched, G.; Kornreich-Leshem, H.; Engeser, M.; Thum, O.; Famulok, M. *J. Am. Chem. Soc.* **2005**, *127*, 15071.

(7) Vaught, J. D.; Bock, C.; Carter, J.; Fitzwater, T.; Otis, M.; Schneider, D.; Rolando, J.; Waugh, S.; Wilcox, S. K.; Eaton, B. E. *J. Am. Chem. Soc.* **2010**, *132*, 4141.

(8) Santoro, S. W.; Joyce, G. F.; Sakthivel, K.; Gramatikova, S.; Barbas, C. F., III. *J. Am. Chem. Soc.* **2000**, *122*, 2433.

(9) Lermer, L.; Roupioz, Y.; Ting, R.; Perrin, D. M. *J. Am. Chem. Soc.* **2002**, *124*, 9960.

(10) Sidorov, A. V.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2004**, *32*, 1591.

(11) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. *Nucleic Acids Res.* **2009**, *37*, 1638.

(12) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. *ChemBioChem* **2009**, *10*, 1988.

(13) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. *ACS Comb. Sci.* **2013**, *15*, 174.

(14) Taylor, A. I.; Pinheiro, V. B.; Smola, M. J.; Morgunov, A. S.; Peak-Chew, S.; Cozens, C.; Weeks, K. M.; Herdewijn, P.; Holliger, P. *Nature* **2015**, *518*, 427.

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

(16) Peracchi, A.; Beigelman, L.; Usman, N.; Herschlag, D. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 11522.

(17) Perrotta, A. T.; Shih, I.; Been, M. D. *Science* **1999**, *286*, 123.

(18) Kuzmin, Y. I.; Da Costa, C. P.; Fedor, M. J. *J. Mol. Biol.* **2004**, *340*, 233.

(19) Toney, M. D.; Kirsch, J. F. *Science* **1989**, *243*, 1485.

(20) Tu, C. K.; Silverman, D. N.; Forsman, C.; Jonsson, B. H.; Lindskog, S. *Biochemistry* **1989**, *28*, 7913.

(21) Walsh, S. M.; Konecki, S. N.; Silverman, S. K. *J. Mol. Evol.* **2015**, *81*, 218.

(22) Chu, C.-C.; Wong, O. Y.; Silverman, S. K. *ChemBioChem* **2014**, *15*, 1905.

(23) Yoo, T. H.; Pogson, M.; Iverson, B. L.; Georgiou, G. *ChemBioChem* **2012**, *13*, 649.

(24) Wu, C.; Tran, J. C.; Zamdborg, L.; Durbin, K. R.; Li, M.; Ahlf, D. R.; Early, B. P.; Thomas, P. M.; Sweedler, J. V.; Kelleher, N. L. *Nat. Methods* **2012**, *9*, 822.

(25) Flynn-Charlebois, A.; Wang, Y.; Prior, T. K.; Rashid, I.; Hoadley, K. A.; Coppins, R. L.; Wolf, A. C.; Silverman, S. K. *J. Am. Chem. Soc.* **2003**, *125*, 2444.

(26) Wang, Y.; Silverman, S. K. *J. Am. Chem. Soc.* **2003**, *125*, 6880.

(27) Langner, J.; Klussmann, S. *BioTechniques* **2003**, *34*, 950.

(28) Caton-Williams, J.; Hoxhaj, R.; Fiaz, B. *Curr. Protoc. Nucleic Acid Chem.* **2013**, *52*, 1.30.1.

Chapter 3: Efforts toward Modified Deoxyribozymes for Peptide Hydrolysis

3.1 Introduction

3.1.1 Identifying Deoxyribozymes for Peptide Hydrolysis

As described in Chapter 2, the incorporation of protein-like functional groups enabled the identification of amide-hydrolyzing deoxyribozymes that can cleave a model amide bond embedded in the DNA strand.[1] However, these deoxyribozymes are unable to catalyze peptide cleavage, which has higher practical value in chemical biology, molecular biology, and proteomics studies. Therefore, selection strategies to identify modified deoxyribozymes for peptide hydrolysis are needed.

3.1.2 Incorporating Multiple Types of Modification for Improved Catalytic Activity

The strategy of incorporating multiple types of modification has been developed for RNA-cleaving deoxyribozymes, to reduce or obviate the divalent metal ion (M^{2+}) requirement.[2–4] The modifications were inspired by protein enzymes like RNase A (Figure 3.1A),[3,5,6] many of which can effectively catalyze RNA cleavage using protein side chain functional groups without the assistance of M^{2+}. Several modified RNA-cleaving deoxyribozymes with reduced or no M^{2+} requirement were identified by incorporating protein-like functional groups.[2–4] In later studies,[7–9] up to three types of modification (primary amino, imidazolyl, and guanidino) were incorporated to achieve more efficient DNA-catalyzed RNA-cleavage in the complete absence of M^{2+} (Figure 3.1B), compared to the previous studies that used only one or two modifications. Although not many mechanistic data were provided, incorporating more than one modification was demonstrated to improve the catalytic activity of deoxyribozymes, at least in the practical viewpoint for this one particular case of a bond cleavage reaction. For a much more difficult bond cleavage reaction like peptide hydrolysis, a similar strategy of incorporating multiple types of protein-like functionality can also be evaluated.

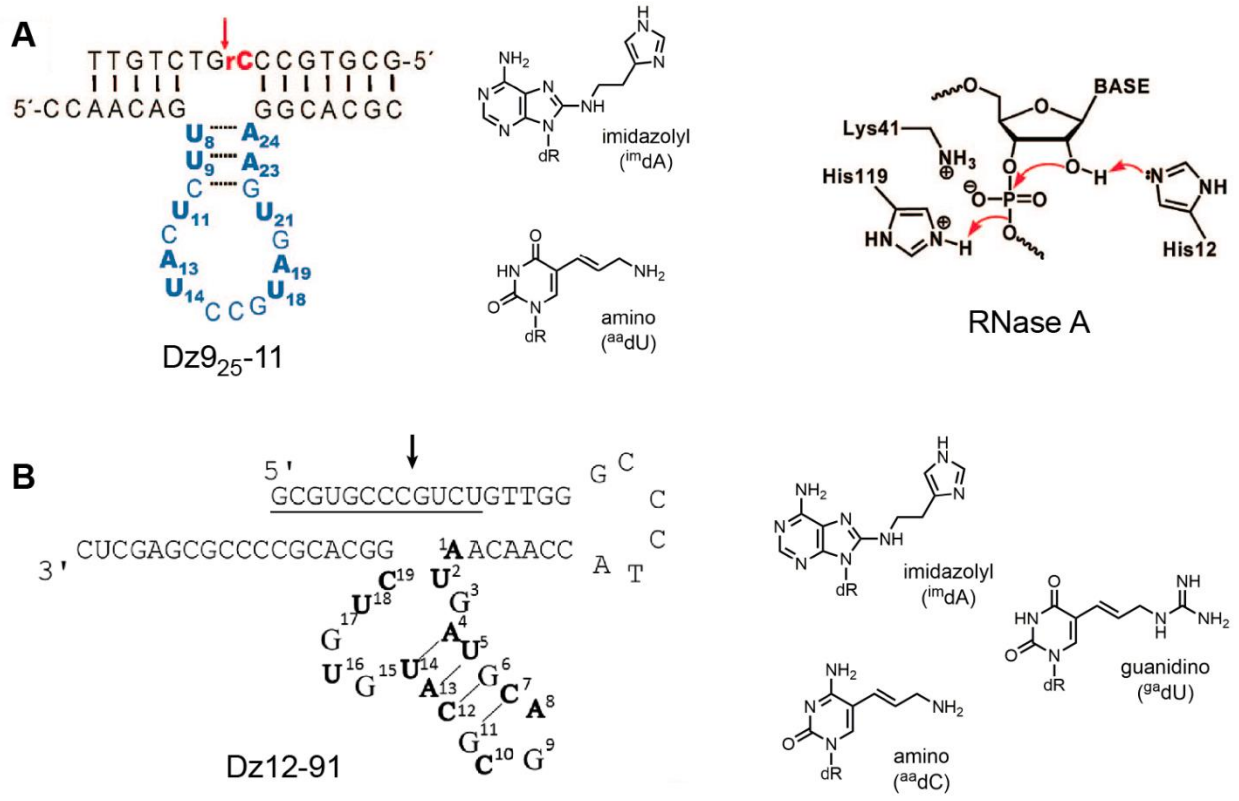

Figure 3.1. Modified nucleotides in M^{2+}-independent RNA-cleaving deoxyribozymes. (A) Dz9$_{25}$-11, with two types of modification (imidazolyl and amino) that were inspired by the RNase A active site. The substrate has a single RNA nucleotide embedded within a DNA sequence. (B) Dz12-91, with three types of modification (imidazolyl, amino, and guanidino). The underscored sequence indicates an all-RNA substrate. Figure adapted with permission from refs. 6 and 9.

3.1.3 Incorporating Hydrophobic Modifications for Improved Binding Affinity

To increase the binding affinity between DNA aptamers and their protein targets, the strategy of incorporating hydrophobic modifications into DNA was developed.[10–14] Many types of hydrophobic modification (such as benzyl, indolyl, and naphthyl) have been shown to increase the binding affinity of DNA aptamers for proteins.[11,14,15] These modified aptamers with low dissociation constants are called slow off-rate modified aptamers (SOMAmers). Crystal structures of several SOMAmer-protein binding complexes revealed that the hydrophobically modified residues within the aptamers engage with the protein targets via hydrophobic interactions.[11,13,15] Similarly, strong interactions are also needed for enzymes to bind their substrates. This

hydrophobic modification strategy could be applied to identifying deoxyribozymes that require strong binding to peptide or protein substrates.

3.1.4 Exploring More Modifications Using Click Chemistry

Rapid and versatile approaches to access the nucleic acid libraries that bear novel functionality are always desired for in vitro selections. A modular functionalization method using click chemistry was developed for SELEX experiments to identify modified DNA aptamers.[16] During this click-SELEX process (Figure 3.2), an alkyne-modified DNA library was prepared by incorporating 5-ethynyl-2′-deoxyuridine (^{E}dU) in place of canonical thymidine (dT) in the DNA. The functionalization step was completed by copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) using an azide-containing compound that bears the desired modification. This functionalization method is readily integrated into the in vitro selection process, facilitating broad access to many chemical modifications for deoxyribozymes.

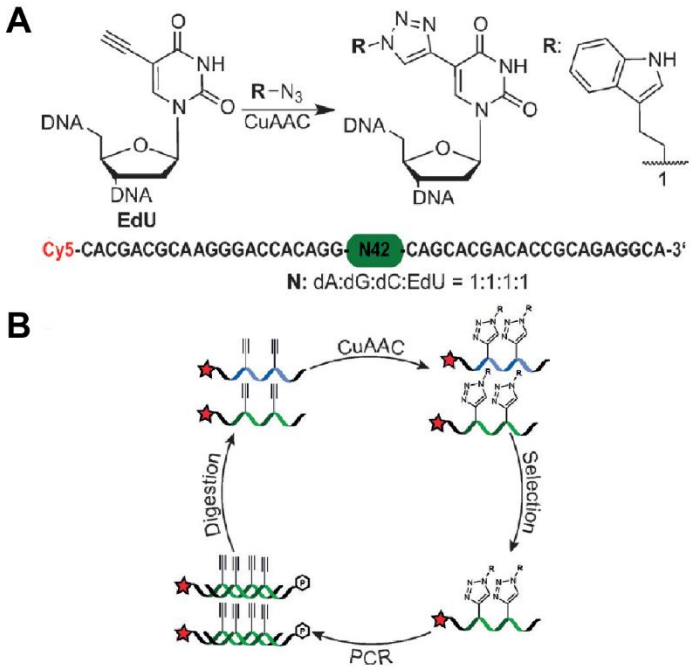

Figure 3.2. Design of click-SELEX for modified-DNA aptamer. (A) CuAAC functionalization of ^{E}dU-containing DNA with an azide compound that contains an indolyl group. (B) Schematic representation of the click-SELEX process. Figure adapted with permission from ref. 16.

3.2 Results and Discussion

3.2.1 In Vitro Selection for Peptide-Hydrolyzing Deoxyribozymes with Modified dU

The identification of modified deoxyribozymes for amide hydrolysis (described in Chapter 2) suggested that the strategy of using modified dU in selections could also be beneficial for identifying peptide-hydrolyzing deoxyribozymes. Therefore, the SR1-SZ1 selection experiments (Figure 3.3) were performed using the same modified nucleotide selection strategy as described in Chapter 2, with the capture step adjusted for the peptide cleavage products using DMT-MM, instead of EDC and HOBt, as the activating compound.

The same four types of modified ^{X}dU (X = Am, Im, HO, or COOH) were used (Figure 3.3A) as in the JV1-JY1 selections described in Chapter 2. The N_{40} pools for the ^{Am}dU (SR1 and SV1), ^{HO}dU (ST1 and SX1), and ^{COOH}dU (SY1 and SZ1) selections were prepared by solid-phase synthesis using the corresponding modified nucleoside phosphoramidites. The N_{40} pools for the ^{Im}dU (SS1 and SW1) selections were prepared by primer extension on a reverse complement template N_{40} pool using ^{Im}dUTP and KOD XL DNA polymerase.

Two different peptide-containing substrates were used in the new selections (Figure 3.3B), AFA tripeptide substrate for the SR1, SS1, ST1, and SY1 selections and AFASWR peptide substrate for the SV1, SW1, SX1, and SZ1 selections. The AFA (Ala-Phe-Ala) tripeptide substrate has two DNA anchors, which bring the tripeptide close to the potential deoxyribozyme via Waston-Crick base pairing. This arrangement is designed structurally similar to the JV1-JY1 selections in Chapter 2, in anticipation that the previously successful selection strategy would also work for this structurally similar peptide substrate. Because the peptide is already embedded in the DNA strand and held tightly in the enzyme-substrate complex, few interactions are required for the deoxyribozyme to engage with the peptide. In contrast, the AFASWR (Ala-Phe-Ala-Ser-Trp-Arg) peptide substrate has only one DNA anchor. The arbitrarily chosen peptide sequence is loosely attached to the DNA anchor via a succinic acid linker and a long HEG tether. For this substrate,

the deoxyribozyme must has a sufficient level of interaction with the peptide to perform the catalysis.

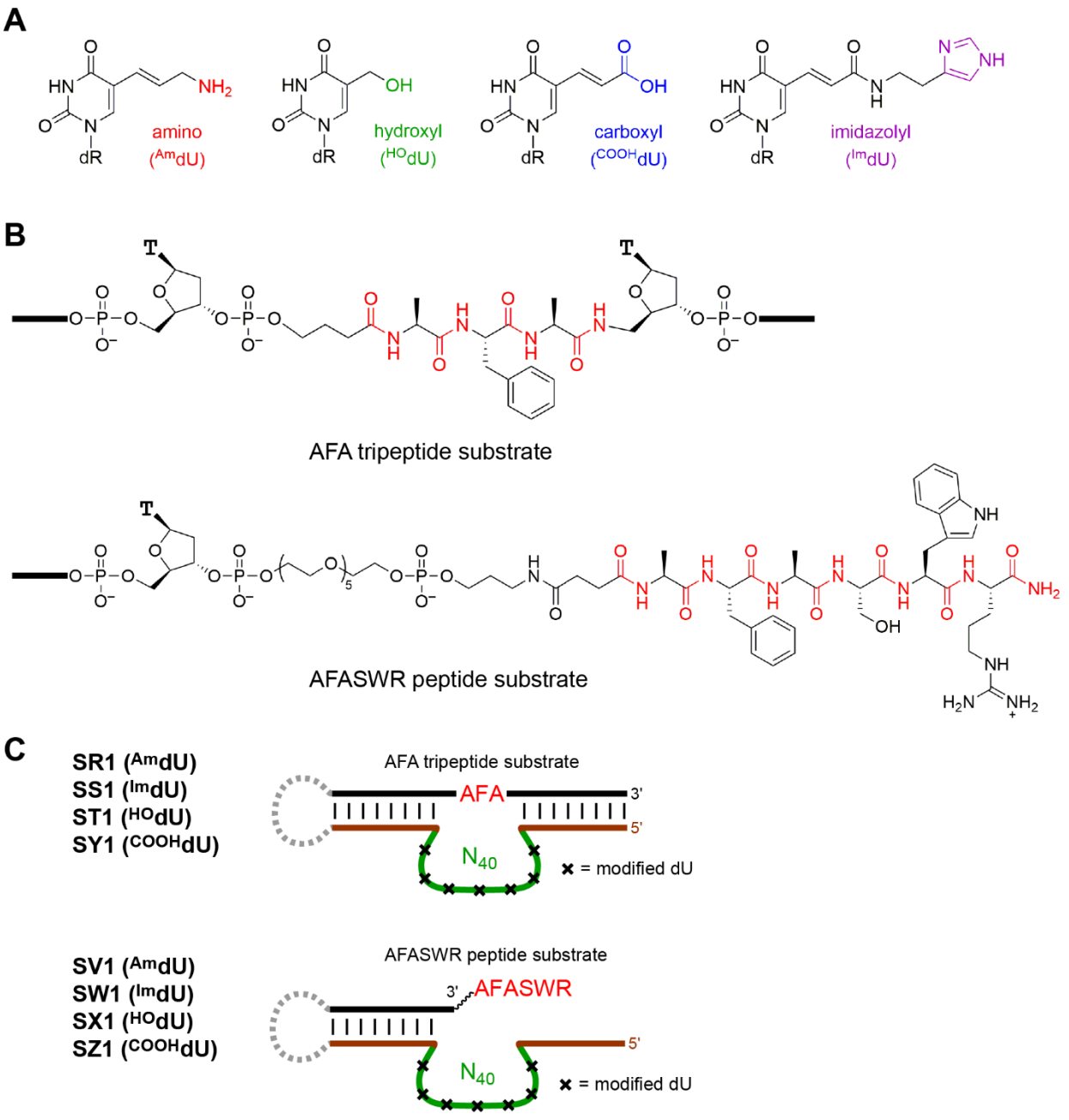

Figure 3.3. SR1-SZ1 selections to identify peptide-hydrolyzing deoxyribozymes with modified dU. (A) Structures of ^{Am}dU, ^{HO}dU, ^{COOH}dU, and ^{Im}dU (dR = deoxyribose). (B) Structures of the peptide-containing substrates for SR1-SZ1 selections. The AFA tripeptide substrate has a GHB (γ-hydroxybutyric acid) linker and an Ala-Phe-Ala tripeptide embedded between two DNA anchors. The AFASWR peptide substrate has an Ala-Phe-Ala-Ser-Trp-Arg peptide attached to one DNA anchor via a succinic acid linker and a HEG tether. The desired peptide cleavage sites are shown in red. The thick black lines indicate the DNA anchors. (C) The selection design for SR1-SZ1. The selection step was performed in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h.

The selection conditions were 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h. In each round, a standard capture reaction was performed using the peptide cleavage product with 3′-GHB-COOH (for SR1, SS1, ST1, and SY1; capture yield 50-60%) or 3′-HEG-succinic-AFA-COOH (for SV1, SW1, SX1, and SZ1; capture yield 20-40%). In round 7 for the ST1 selection, 5% activity was observed. However, when assayed in trans without the AFA tripeptide substrate ligated to the PCR product, the cleavage site was clearly not within the tripeptide but instead within the right-hand DNA anchor. Round 11 of the SR1 selection also showed a similar result. Because our capture reaction excludes the capture of any DNA hydrolysis product, this cleavage reaction cannot be DNA hydrolysis. A radical-based DNA cleavage is more likely, which will be described further in Chapter 4. For the other selections, no detectable activity (<0.5%) was observed after 14-16 rounds, and the selections were discontinued.

These findings suggested that using the doubly anchored peptide substrate could likely lead to the cleavage within the right-hand DNA anchor instead of the peptide. Because the singly anchored peptide substrate is more like a discrete peptide and therefore more relevant to practical applications, only the singly anchored peptide substrate was used in subsequent selections. Moreover, the result of the SR1-SZ1 selections also indicated that the peptide hydrolysis is a more difficult reaction than the simple amide bond hydrolysis in the practical context of the deoxyribozyme identification. Therefore, a strategy that could further boost the catalytic capability of DNA was evaluated.

3.2.2 In Vitro Selection for Peptide-Hydrolyzing Deoxyribozymes with Two Modifications

One way to increase the catalytic capability of DNA is by using a combination of multiple different modifications. For a difficult reaction like peptide hydrolysis, two or more types of modification may be required to function in synergy, just as in the catalytic mechanism of some natural proteases.[17] To evaluate the strategy of using two types of protein-like functionality for

DNA-catalyzed peptide hydrolysis, ZA1-ZG1, ZY1, and VZ1 selections were designed (Figure 3.4).

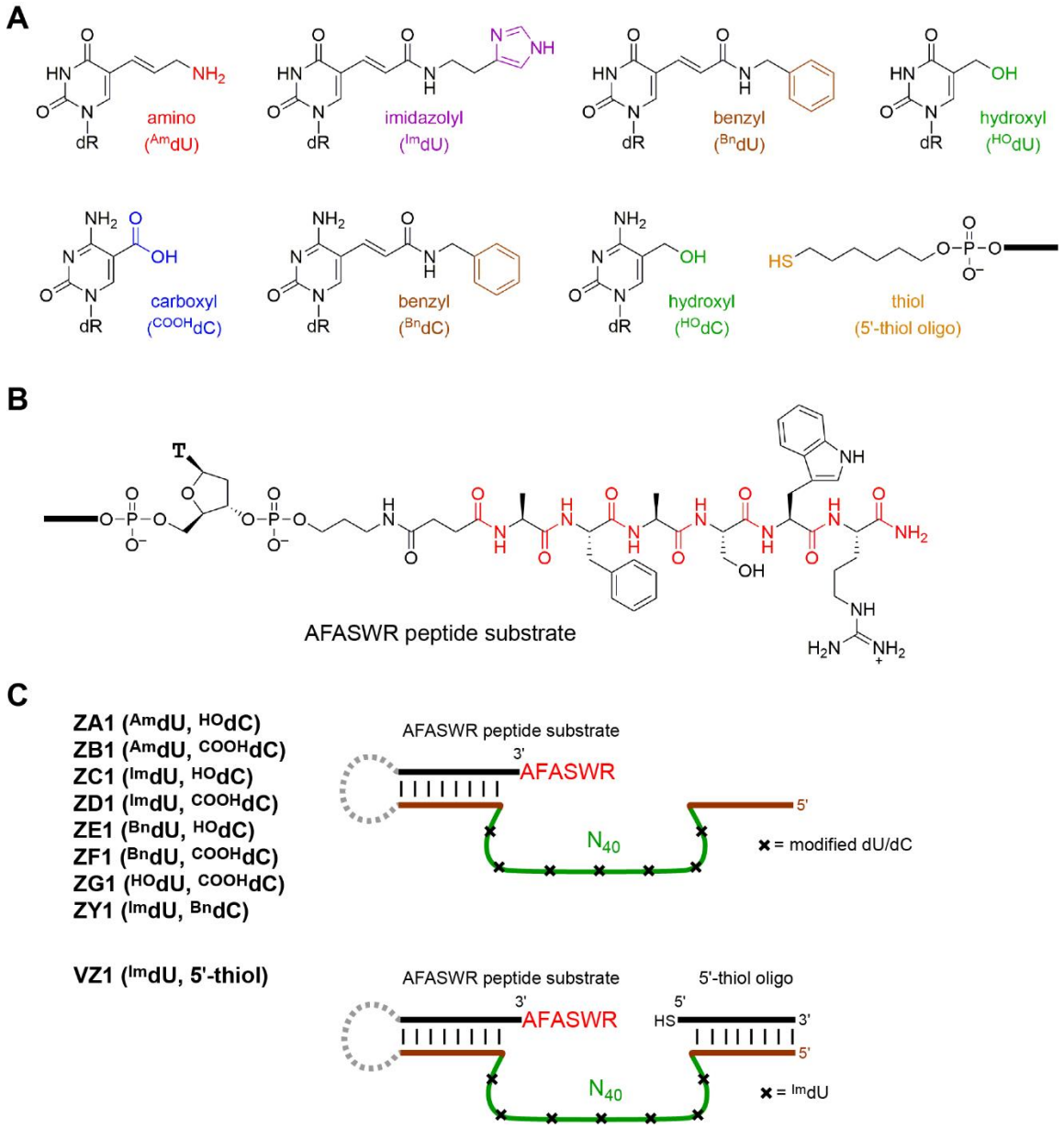

Figure 3.4. The ZA1-ZG1, ZY1, and VZ1 selections to identify peptide-hydrolyzing deoxyribozymes with two types of modification. (A) Structures of modified dU, modified dC, and 5′-thiol oligo (dR = deoxyribose; the thick black line indicates the DNA anchor). (B) Structures of the AFASWR peptide substrate for the ZA1-ZG1, ZY1, and VZ1 selections. The AFASWR peptide substrate has an Ala-Phe-Ala-Ser-Trp-Arg peptide attached to a DNA anchor via a succinic acid linker and a C_3 tether. The desired peptide cleavage sites are shown in red. (C) The design for the ZA1-ZG1, ZY1, and VZ1 selections. The selection step was performed in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h.

To incorporate two types of modification into one DNA strand, both modified dU and dC were used (Figure 3.4A). A benzyl group (in ^{Bn}dU and ^{Bn}dC) was added to the modification repertoire due to its promising property of increasing binding affinity between the modified DNA and the peptide via hydrophobic interactions.[11,12] A thiol group (in 5′-thiol oligo) was also used as a protein-like functionality, inspired by the mechanism of cysteine proteases.[17]

The AFASWR peptide substrate (Figure 3.4B) contains the same Ala-Phe-Ala-Ser-Trp-Arg peptide sequence and succinic acid linker as in the SV1-SZ1 selections, but with a shorter C_3 tether rather than a HEG tether. This short-tethered substrate is considered an easier substrate for catalysis, because the closely attached peptide has a higher chance to interact with the deoxyribozyme.

The N_{40} pools for the ZA1 and ZB1 selections were prepared by solid-phase synthesis, each using the relevant modified nucleoside phosphoramidites. The N_{40} pools for the other selections were prepared by primer extension on a reverse complement template using the relevant modified nucleoside triphosphates and KOD XL DNA polymerase. ^{Im}dUTP and ^{Bn}dUTP were synthesized from NHS-activated ^{COOH}dU [(*E*)-5-(2-carbomethoxyvinyl)-2′-deoxyuridine *N*-hydroxysuccinimide ester]. ^{Bn}dCTP was synthesized from 5-iodo-2′-deoxycytidine (Figure 3.5). The phosphoramidites and other triphosphates were purchased.

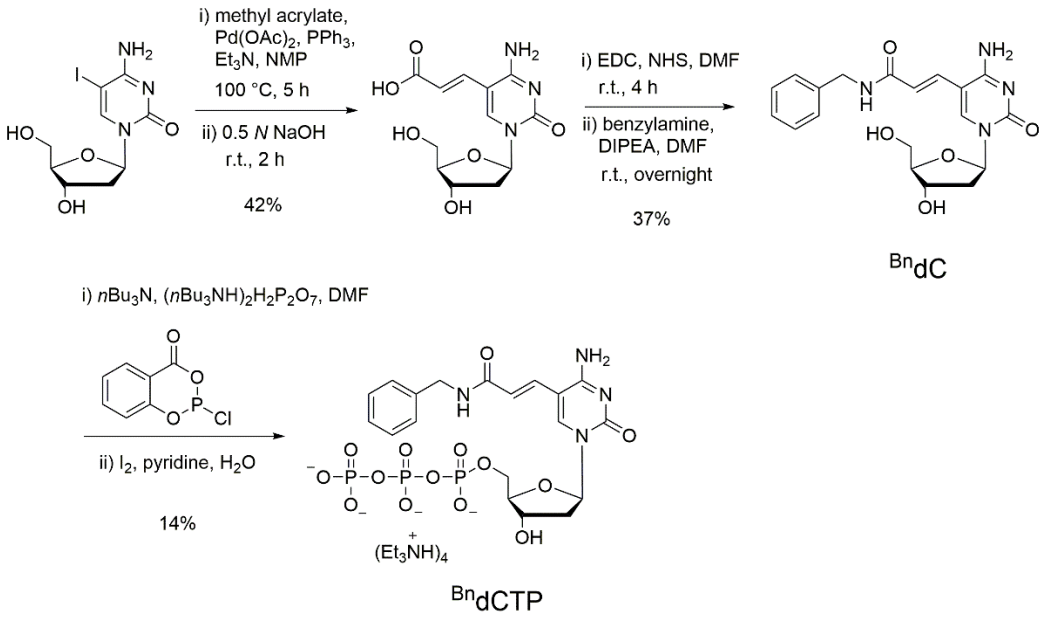

Figure 3.5. Synthesis of ^{Bn}dCTP. For the first step, (*E*)-5-(2-carboxyvinyl)-2′-deoxycytidine was prepared from 5-iodo-2′-deoxycytidine by a Heck reaction and a subsequent alkaline hydrolysis. For the second step, ^{Bn}dC was synthesized by coupling benzylamine with (*E*)-5-(2-carboxyvinyl)-2′-deoxycytidine using EDC and NHS. For the third step, ^{Bn}dCTP was synthesized by the triphosphorylation of ^{Bn}dC.

The ZA1-ZG1, ZY1, and VZ1 selections were performed using strategies slightly adjusted from those used in the SR1-SZ1 selections described in Section 3.2.1. For the ZA1-ZG1 and ZY1 selections, the primer extension step was adjusted by using both modified dUTP and modified dCTP as required for incorporating two types of modified nucleotides. For the VZ1 selection, a 5′-thiol oligo was included in the selection step.

The selection conditions were 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h. In each round, a separate standard capture reaction was performed using the peptide cleavage product with 3′-C_3-succinic-AFA-COOH (capture yield 30-50%). Unfortunately, no activity was observed for these selections after 14 rounds. One possible explanation is that the fundamental catalytic capability of hydrolyzing a peptide bond is not the major challenge for deoxyribozymes, especially considering that similar functional groups enabled the identification of deoxyribozymes that hydrolyze an amide bond as described in Chapter 2. Instead, the main reason could be that the DNA sequences were unable to

interact sufficiently with the peptide substrate to perform the catalysis, therefore could not be identified by these selections.

3.2.3 In Vitro Selection for Peptide-Hydrolyzing Deoxyribozymes with Hydrophobic Modifications Incorporated by Click Chemistry

Driven by the lessons learned from the previous successful and unsuccessful selections, we sought to develop a strategy that could improve the interaction between the deoxyribozyme and the peptide substrate. Others have reported that hydrophobic modifications on nucleobases can improve the binding affinity of DNA aptamers to protein targets by increasing the number of hydrophobic interactions.[10,14,16] Based on these studies, the same hydrophobic effects may also help to identify deoxyribozymes that interact more strongly with the peptide substrates.

To explore the contributions of hydrophobic modifications comprehensively while obviating extensive synthetic efforts, the click chemistry functionalization approach related to an aptamer study[16] was integrated into our previous modified nucleotide selection strategy for identifying deoxyribozymes with peptide hydrolysis activity (Figure 3.6). In the primer extension step, 5-ethynyl-2′-deoxyuridine (^{E}dU) was incorporated into the deoxyribozymes using ^{E}dUTP and KOD XL DNA polymerase. The subsequent click step functionalized the DNA sequences by copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) reaction using the azide compounds that contain the desired modifications.

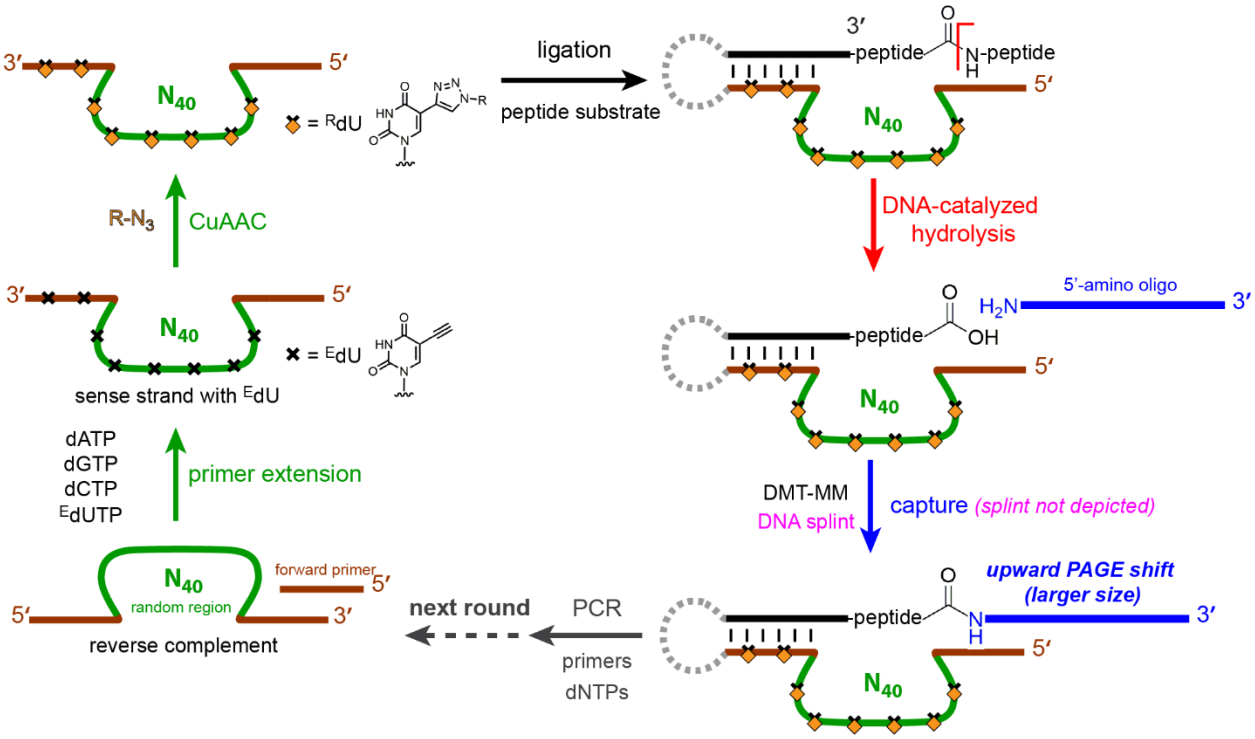

Figure 3.6. In vitro selection strategy for peptide-hydrolyzing deoxyribozymes functionalized by click chemistry. The ^{E}dU N_{40} pool was prepared by primer extension from a reverse complement template. Modifications (depicted as R group) were incorporated by CuAAC reaction using the corresponding azide compounds (R-N_3). A DNA-anchored peptide substrate was ligated to the modified pool. The selection step for DNA-catalyzed peptide hydrolysis was performed in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, 150 mM NaCl at 37 °C for 14 h. In the capture step, the catalytically active DNA sequences were captured by DMT-MM coupling reaction using a 5′-amino oligo. The new pool enriched in active DNA sequences was amplified by PCR, and the reverse complement pool was isolated for the primer extension of the next round.

This functionalization approach facilitated the exploration of diversified modifications. A variety of modifications containing different types of hydrophobic moiety were designed (Figure 3.7A), in consideration of the ease of synthesis as well as the functional group diversity. Many of these modifications contain not only a hydrophobic moiety but also some combination of amino, hydroxyl, and carboxyl groups, which were demonstrated in Chapter 2 to be critical for identifying DNA-catalyzed amide bond hydrolysis.

These modifications could be incorporated into deoxyribozymes by CuAAC using the corresponding azide compounds (Figure 3.7B). Az-Naph was synthesized by Silverman lab member Puzhou Wang. Az-Phe was purchased. I synthesized the other azide compounds. Az-

Naph-NH_2 was prepared by azide substitution of 2,6-bis(bromomethyl)naphthalene, followed by partial reduction using triphenylphosphine. Each of Az-HO-NH-Bn/Naph/Trp/Im was synthesized by nucleophilic attack of the corresponding amine into (*S*)-epichlorohydrin, followed by azide substitution. The imidazole nitrogen of histamine was protected with a trityl group due to its substantial nucleophilicity that could cause side reactions.

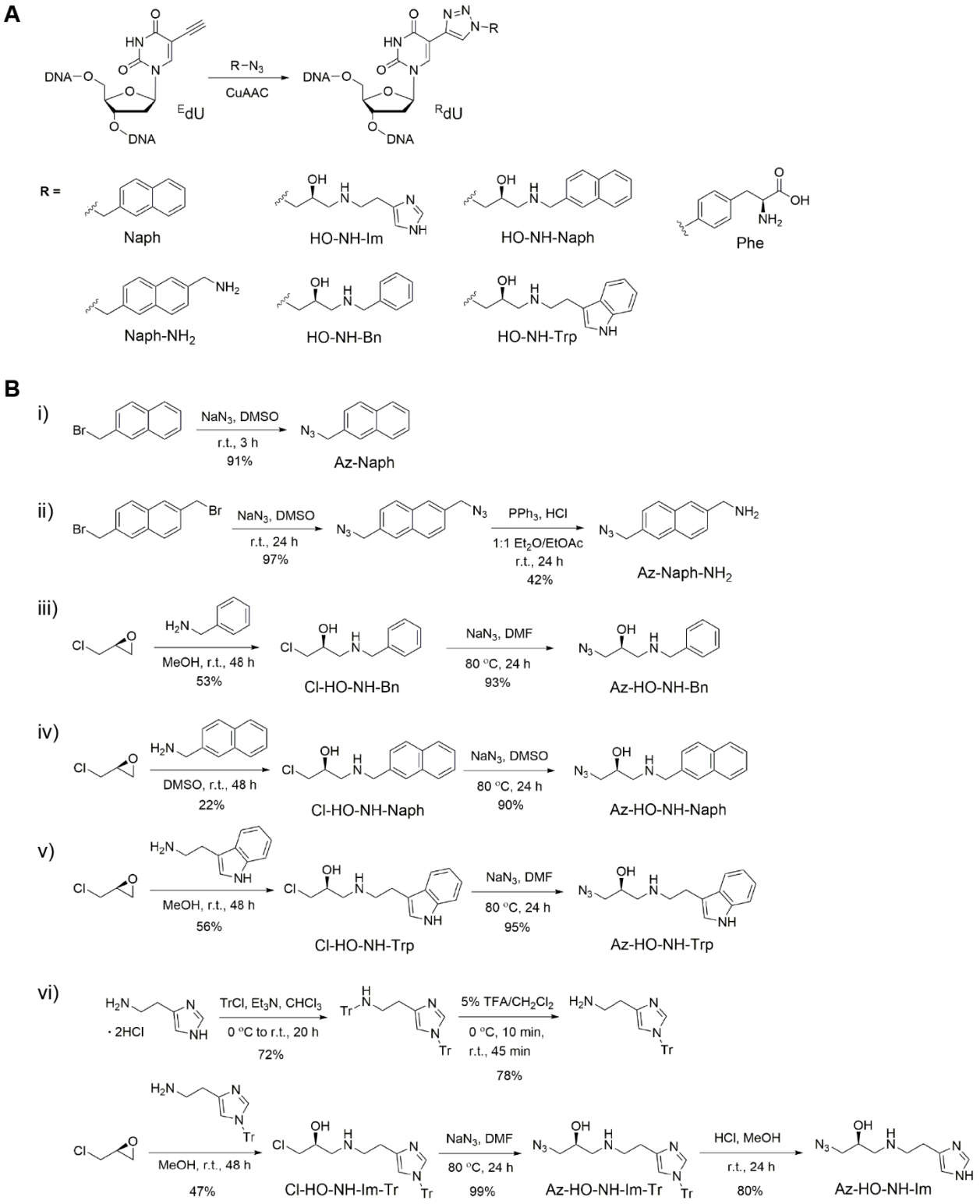

Figure 3.7. Design and synthesis of the azide compounds for modification incorporation. (A) Hydrophobic modifications. (B) Synthesis of (i) Az-Naph, (ii) Az-Naph-NH_2, (iii) Az-HO-NH-Bn, (iv) Az-HO-NH-Naph, (v) Az-HO-NH-Trp, and (vi) Az-HO-NH-Im.

The DP2-DX2 selections (Figure 3.8) were performed using the same AFASWR peptide substrate as described in Section 3.2.2. The DP2-DW2 selections used the hydrophobic modifications shown in Figure 3.7. DX2 used the Naph-NH_2 modification, as well as a 5′-thiol oligo that was present in the selection step. These selections are still in progress.

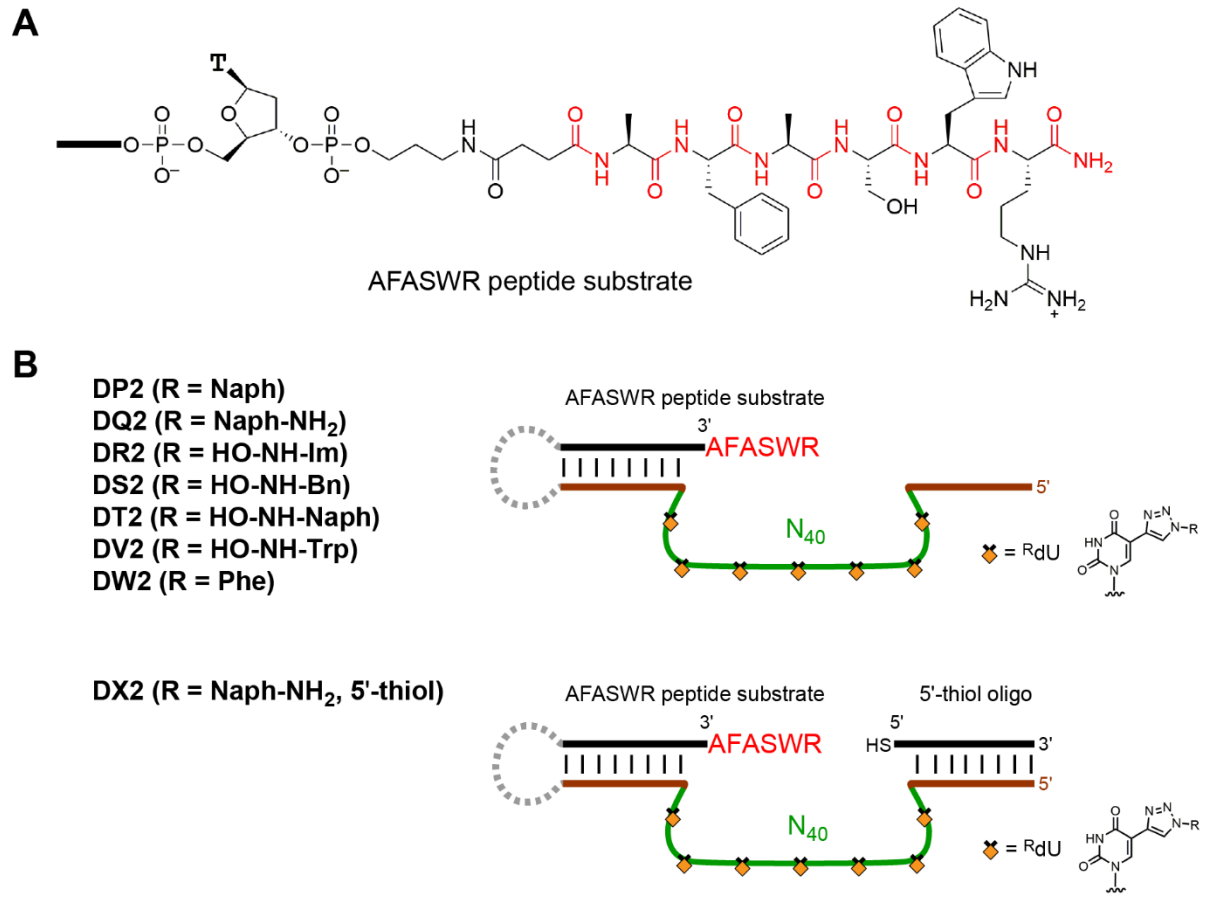

Figure 3.8. DP2-DX2 selections to identify peptide-hydrolyzing deoxyribozymes with hydrophobic modifications. (A) Structure of the AFASWR peptide substrate. The Ala-Phe-Ala-Ser-Trp-Arg peptide is attached to a DNA anchor via a succinic acid linker and a C_3 tether. The desired peptide cleavage sites are shown in red. (B) The design for the DP2-DX2 selections. The selection step was performed in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C for 14 h.

3.3 Summary and Future Directions

Although DNA-catalyzed amide hydrolysis was achieved by incorporating protein-like functional groups (primary amino, hydroxyl, or carboxyl) into the deoxyribozymes (described in Chapter 2), the same selection strategy was difficult to expand from the amide model substrate to peptide substrates. No deoxyribozymes were identified in the SR1-SZ1 selections for cleaving either an AFA tripeptide substrate with a two-anchor structural arrangement or an AFASWR peptide substrate with just one DNA anchor. These results suggested that DNA-catalyzed peptide hydrolysis is a more difficult reaction than plain amide bond hydrolysis, at least from the practical viewpoint of deoxyribozyme identification. The DNA cleavage side reaction found in the ST1 and SR1 selections suggested not to use doubly anchored peptide substrates in subsequent peptide hydrolysis selections because of the unavoidable side reaction. In addition, singly anchored peptide substrates require the deoxyribozymes to interact more strongly with the peptides, therefore are more relevant to the potential applications of the deoxyribozymes, which is cleaving untethered peptides.

Two approaches were used to boost the catalytic capability of DNA for the difficult peptide hydrolysis. The first approach (the ZA1-ZG1, ZY1 and VZ1 selections) was focused on using two different types of protein-like functional groups, with the anticipation that some combination would function in synergy to enable the catalysis, similar to what natural proteases do in the active sites. However, no activity was observed in these selections, which led us to the second approach focusing on exploring hydrophobic modifications and protein-like functionality at the same time. The hypothesis is that if the deoxyribozymes could not sufficiently interact with the peptide substrate, then the hydrolysis would not happen even if the deoxyribozymes had the necessary functional groups to participate in the catalysis. Others established that hydrophobic modifications can improve the binding affinity of DNA aptamers to the protein targets. Therefore, the incorporation of hydrophobic modifications may also increase the interactions between

deoxyribozymes and their peptide substrates. To evaluate this strategy, the DP2-DX2 selections are currently in progress.

In the long term, DNA-catalyzed peptide hydrolysis with sequence selectivity and free peptide activity is desired. To achieve peptide sequence selectivity, deoxyribozymes must recognize the specific structures and side chain residues of the peptide substrates through certain enzyme-substrate interactions. To achieve free peptide activity, deoxyribozymes must bind the untethered peptide substrates tightly in solution through strong enzyme-substrate interactions. Therefore, finding the appropriate modifications that can introduce specific and strong enzyme-substrate interactions is a promising approach to achieve sequence selectivity and free peptide activity for DNA-catalyzed peptide hydrolysis. Results from the DP2-DX2 selections can be used to guide the selection design for other peptide/protein substrates. For future applications, the most immediate selection plan is to identify nucleobase-modified deoxyribozymes for leader peptide cleavage in the context of peptide natural product biosynthesis.[18,19] For a certain natural product candidate with a leader peptide to be cleaved, useful deoxyribozymes need only moderate sequence selectivity (no miscleavage within the peptide sequence) and moderate binding affinity (K_m in low mM range), which are achievable according to our previous research for other deoxyribozymes that use peptide substrates.[20–22]

3.4 Materials and Methods

3.4.1 Preparation of Oligonucleotides and Peptide Substrates

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), extracted from the polyacrylamide with TEN buffer (10 mM Tris, pH 8.0, 1 mM EDTA, 300 mM NaCl), and precipitated with ethanol as described previously.[23,24]

Peptides were prepared by solid-phase synthesis as described,[25] with a slight adjustment to incorporate the succinic acid linker at the N-terminus of the peptide. After the last amino acid was installed and deprotected on the N-terminus, succinic anhydride (10 equivalents, 200 mg, 2 mmol) instead of acetic anhydride was used in the final capping step. Other steps were the same as previously described.[25]

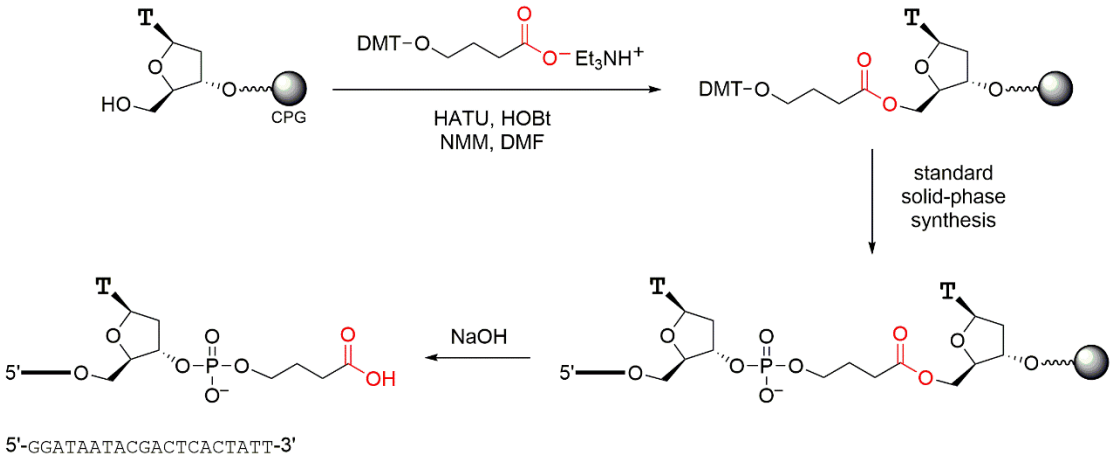

Figure 3.9. Solid-phase synthesis of 3′-GHB-COOH oligonucleotide.

Synthesis of 3′-GHB-COOH oligonucleotide. 5′-DMT-thymidine-derivatized CPG solid support (40 µmol/g; 5 mg, 200 nmol) was deprotected with 3% (v/v) TCA/CH_2Cl_2 (3 × 1 mL) and washed with CH_2Cl_2 (3 × 5 mL). The support was rinsed with 2% (v/v) NMM/CH_2Cl_2 and CH_2Cl_2 (3 × 5 mL) to improve the efficiency of the next coupling step and dried under a stream of nitrogen. *O*-(4,4′-Dimethoxytrityl)-4-hydroxybutanoate triethylammonium salt (50 mg, 100 µmol; prepared by Silverman lab member Benjamin Brandsen), HATU (38 mg, 100 µmol), and HOBt (14 mg, 100 µmol) were dissolved in 500 µL of dry DMF. NMM (26 µL, 230 µmol) was added to this solution, which was introduced into the DNA synthesis column via two syringes. The coupling step was performed for 16 h with mixing every 5 min for the first 1 h and no mixing for the remaining 15 h. The column was washed with DMF (3 × 5 mL) and CH_2Cl_2 (3 × 5 mL) and dried under a stream of nitrogen. The remainder of the oligonucleotide was synthesized on the ABI 394

instrument using standard procedures. Cleavage of the oligonucleotide from the CPG and nucleobase deprotection were performed using 200 µL of 400 mM NaOH in 4:1 MeOH:H_2O for 24 h at room temperature. The sample was diluted to 600 µL with water and precipitated with ethanol. The deprotection was completed by treating the sample with concentrated ammonium hydroxide for 16 h at 55 °C. The sample was dried under vacuum and purified by 20% PAGE. 70 nmol of the product was obtained. MALDI-MS $[M+H]^+$ calcd. 6267.1, found 6263.9.

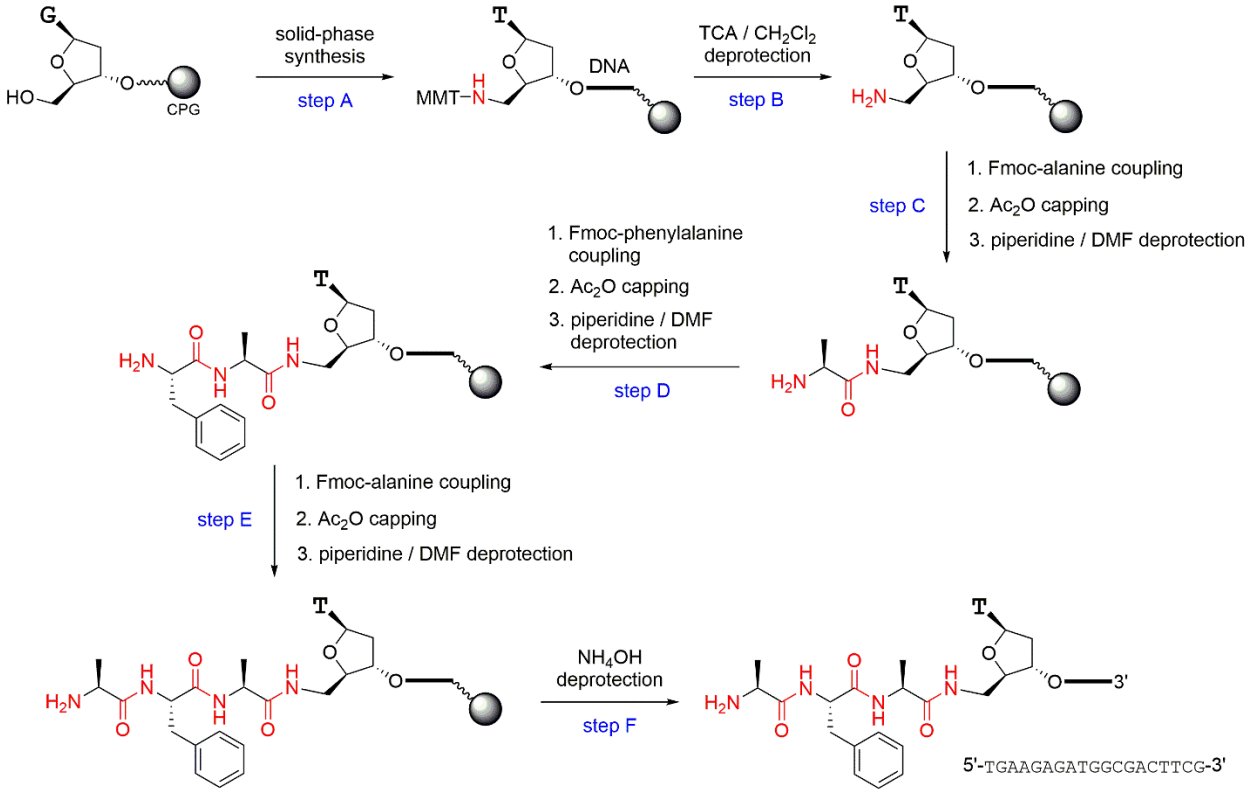

Figure 3.10. Solid-phase synthesis of 5′-amino-AFA oligonucleotide.

Synthesis of 5′-amino-AFA oligonucleotide.

Synthesis and deprotection of 5′-amino-5′-deoxythymidine oligonucleotide (steps A and B). Oligonucleotide synthesis was performed on the ABI 394 instrument using standard procedures, providing a CPG-supported oligonucleotide terminating in a *N*-MMT-5′-amino-5′-deoxythymidine residue. The MMT protecting group was removed with 3% (v/v) TCA/CH_2Cl_2

(3 × 1 mL), and the support was rinsed with CH_2Cl_2 (3 × 5 mL). The support was rinsed with 2% (v/v) NMM/CH_2Cl_2 and CH_2Cl_2 (3 × 5 mL) to improve the efficiency of the next coupling step and dried under a stream of nitrogen.

Coupling of Fmoc-alanine (step C1). Fmoc-alanine (31 mg, 100 µmol), HATU (38 mg, 100 µmol), and HOBt (14 mg, 100 µmol) were dissolved in 500 µL of dry DMF. NMM (26 µL, 230 µmol) was added to this solution, which was introduced into the DNA synthesis column via a syringe under an argon balloon. The coupling step was performed for 2 h with mixing every 15 min. The support was washed with DMF (3 × 5 mL) and CH_2Cl_2 (3 × 5 mL) and dried under a stream of nitrogen.

Capping of unreacted 5′-amino group (step C2). Capping of unreacted amino group was performed using 500 µL of Glen Research Cap Mix A (10% Ac_2O/10% pyridine in THF) and 500 µL of Glen Research Cap Mix B (10% 1-methylimidazole in THF). Cap Mix A and Cap Mix B were combined and introduced into the DNA synthesis column via a syringe under an argon balloon. The capping step was performed for 30 min with mixing every 5 min. The support was washed with CH_2Cl_2 (3 × 5 mL) and dried under a stream of nitrogen.

Deprotection of Fmoc group (step C3). The Fmoc protecting group was removed with 20% (v/v) piperidine/DMF (4 × 1 mL). The support was washed with DMF (3 × 5 mL) and CH_2Cl_2 (3 × 5 mL) and dried under a stream of nitrogen.

Coupling of Fmoc-phenylalanine (step D1). Fmoc-phenylalanine was coupled to the N-terminal amino group via the coupling step described above (step C1), using Fmoc-phenylalanine (39 mg, 100 µmol) in place of Fmoc-alanine and extending the coupling reaction time from 2 h to 18 h.

Capping of unreacted N-terminus of the alanine (step D2). Capping with acetic anhydride was performed as described above (step C2).

Deprotection of Fmoc group (step D3). The Fmoc protecting group was removed as described above (step C3).

Coupling of Fmoc-alanine (step E1). Fmoc-alanine was coupled to the terminal amino group via the coupling step described above (step C1).

Capping of unreacted N-terminus of the phenylalanine (step E2). Capping with acetic anhydride was performed as described above (step C2).

Deprotection of Fmoc group (step E3). The Fmoc protecting group was removed as described above (step C3).

Completion of synthesis (step F). The product was cleaved from the solid support and deprotected by treating the sample with concentrated ammonium hydroxide for 16 h at 55 °C. The sample was dried under vacuum and purified by 20% PAGE. 73 nmol of the product was obtained. MALDI-MS $[M+H]^+$ calcd. 6182.2, found 6186.0.

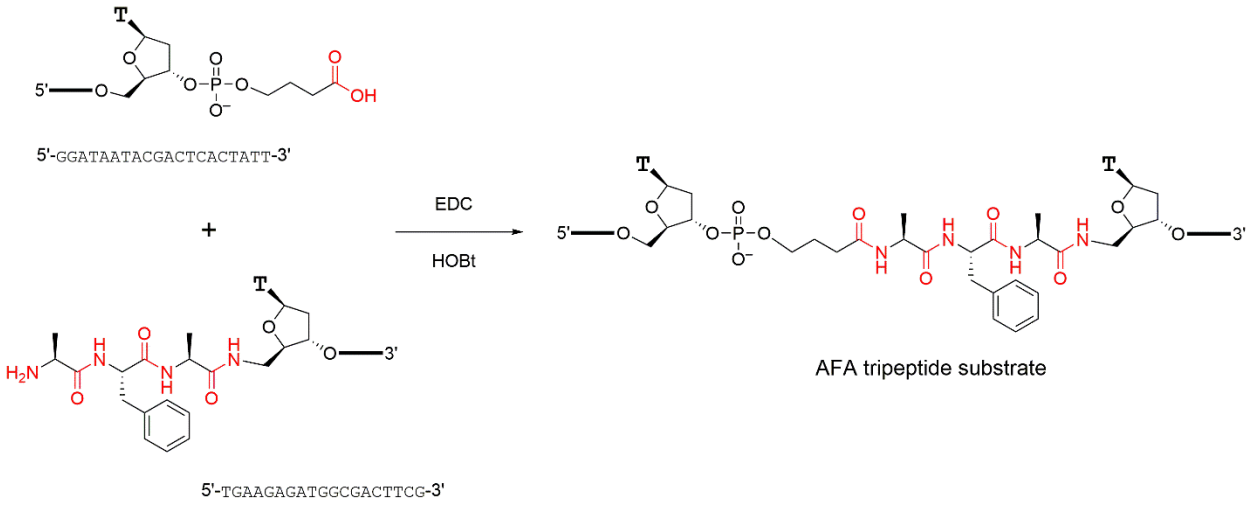

Figure 3.11. Synthesis of AFA tripeptide substrate by splinted coupling of 3′-GHB-COOH oligonucleotide and 5′-amino-AFA oligonucleotide.

Synthesis of AFA tripeptide substrate by solution-phase coupling of 3′-GHB-COOH oligonucleotide and 5′-amino-AFA oligonucleotide. A 45 µL sample containing 4.0 nmol of 3′-GHB-COOH oligonucleotide, 5.0 nmol of DNA splint, and 6.0 nmol of 5′-amino-AFA oligonucleotide was annealed in 110 mM MOPS, pH 7.0, and 1.1 M NaCl by heating at 95 °C for 3 min and cooling on ice for 5 min. The sequence of the DNA splint was 5′-

CGAAGTCGCCATCTCTTCA**T**AATAGTGAGTCGTATTATCCAACAACAACAACAACAAC-3′, where the boldface **T** is unpaired, and the underlined region is included to shift the splint away from the desired product on PAGE. The reaction was initiated by bringing the sample to 50 µL volume containing 100 mM MOPS, pH 7.0, 1 M NaCl, 50 mM EDC, 50 mM HOBt, and 5% (v/v) DMF, the latter of which was from the HOBt stock solution and was required to dissolve the HOBt. The sample was incubated at room temperature for 24 h, precipitated with ethanol, and purified by 20% PAGE. 3.2 nmol of the product was obtained. MALDI-MS $[M+H]^+$ calcd. 12430.3, found 12431.6.

Synthesis of AFASWR peptide substrate with HEG or C_3 tether. DNA-anchored AFASWR peptide conjugates were synthesized by a DMT-MM coupling reaction between a DNA tethered 3′-amine and the succinic acid linker at the N-terminus of the HOOC-succinic-Ala-Phe-Ala-Ser-Trp-Arg peptide. The 3′-amino-modified DNA oligonucleotides were synthesized as 5′-DNA-HEG-p-C_3-NH_2-3′ (for HEG-tethered peptide substrate) or 5′-DNA-p-C_3-NH_2-3′ (for C_3-tethered peptide substrate). The DNA sequence was 5'-GGATAATACGACTCACTATT-3'. To a 5 µL sample containing 5.0 nmol of the 3′-amino-modified DNA oligonucleotide was added 4 µL of 1 M MOPS, pH 7.0, 8 µL of DMF, 2 µL of 300 mM peptide in DMF, and 2 mg of DMT-MM. The sample was incubated at room temperature for 15 h, desalted by 3K Amicon Ultra-0.5 centrifugal filter device (Millipore), and purified by 20% PAGE. 3.3 nmol of the HEG-tethered peptide substrate was obtained. MALDI-MS $[M+H]^+$ calcd. 7096.1, found 7096.5. 3.7 nmol of the C_3-tethered peptide substrate was obtained. MALDI-MS $[M+H]^+$ calcd. 6751.8, found 6751.0.

3.4.2 2′-Deoxynucleotide 5′-Triphosphates and 2′-Deoxynucleoside 3′-Phosphoramidites

The ^{Am}dU 5′-triphosphate (^{Am}dUTP) was from TriLink BioTechnologies (cat. no. N-0249). The ^{HO}dU 5′-triphosphate (^{HO}dUTP) was from TriLink BioTechnologies (cat. no. N-0259). The ^{HO}dC 5′-triphosphate (^{HO}dCTP) was from TriLink BioTechnologies (cat. no. N-0260). The ^{COOH}dC 5′-triphosphate (^{COOH}dCTP) was from TriLink BioTechnologies (cat. no. N-0263). The ^{E}dU 5′-triphosphate (^{E}dUTP) was from Baseclick GmbH (cat. no. BCT-08-L).

The ^{Am}dU phosphoramidite was from Berry & Associates (cat. no. BA0311). The ^{HO}dU phosphoramidite was from Glen Research (cat. no. 10-1093). The ^{COOH}dU phosphoramidite was from Glen Research (cat. no. 10-1035). The ^{HO}dC phosphoramidite was from Glen Research (cat. no. 10-1062). The ^{COOH}dC phosphoramidite was from Glen Research (cat. no. 10-1066).

The ^{COOH}dU 5′-triphosphate (^{COOH}dUTP) and ^{Im}dU 5′-triphosphate (^{Im}dUTP) were synthesized as described in Chapter 2. The ^{Bn}dU 5′-triphosphate (^{Bn}dUTP) and ^{Bn}dC 5′-triphosphate (^{Bn}dCTP) were synthesized as described in Section 3.4.4.1.

3.4.3 In Vitro Selection Procedures

3.4.3.1 SR1-SZ1 Selections

Procedures for ligation, selection, PCR, and primer extension were performed as described in Chapter 2. The capture step was adjusted for peptide hydrolysis products by using DMT-MM as the activating compound.

Procedure for capture step in round 1. A 90 µL sample containing the selection product, 300 pmol of capture splint, and 500 pmol of 5′-amino capture oligonucleotide was annealed in 100 mM MOPS, pH 7.0, and 1 M NaCl by heating at 95 °C for 3 min and cooling on ice for 5 min. The capture reaction was initiated by bringing the sample to 100 µL total volume containing 100 mM MOPS, pH 7.0, 1 M NaCl, and 100 mM DMT-MM. The sample was incubated at room temperature for 15 h, precipitated with ethanol, and separated by 8% PAGE.

Procedure for capture step in subsequent rounds. A 22.5 µL sample containing the selection product, 30 pmol of capture splint, and 50 pmol of 5′-amino capture oligonucleotide was annealed in 100 mM MOPS, pH 7.0, and 1 M NaCl by heating at 95 °C for 3 min and cooling on ice for 5 min. The capture reaction was initiated by bringing the sample to 25 µL total volume containing 100 mM MOPS, pH 7.0, 1 M NaCl, and 50 mM DMT-MM. The sample was incubated at room temperature for 15 h and loaded directly on 8% PAGE. In parallel, during each round a standard capture reaction was performed using the N_{40} pool ligated to the hydrolysis product. The

PAGE position of this standard capture product was used as a reference to excise the capture product for the selection sample.

3.4.3.2 ZA1-ZG1, ZY1, and VZ1 Selections

For the ZA1-ZG1 and ZY1 selections, procedures for ligation, selection, capture, and PCR were performed as described in Section 3.4.3.1. The primer extension step was adjusted to incorporate both modified dU and modified dC. For the VZ1 selection, procedures for ligation, capture, PCR, and primer extension were performed as described in Section 3.4.3.1. The selection step was adjusted to include the 5′-thiol oligonucleotide that bound to the right binding arm of the pool.

Procedure for primer extension to prepare the random pool for ZA1-ZG1 and ZY1. A 50 µL sample was prepared containing 250 pmol of the reverse complement template pool (which contains a 3′-segment of arbitrary sequence to allow PAGE separation of the primer extension product from the template), 500 pmol of forward primer, 15 nmol each of dATP, dGTP, modified ^{X}dUTP (X = Am, Im, Bn, or HO), and modified ^{X}dCTP (X = HO, COOH, or Bn), 5 µL of 10× KOD XL polymerase buffer, and 1 µL of KOD XL polymerase (prepared by Silverman lab member Shannon Walsh). Primer extension was performed by cycling 4 times in a PCR thermocycler according to the following program: 94 °C for 2 min, 4× (94 °C for 2 min, 47 °C for 2 min, 72 °C for 30 min), 72 °C for 30 min. The sample was separated by 8% PAGE to provide the random pool for the corresponding selection.

Procedure for primer extension to generate the modified DNA pool in subsequent rounds for ZA1-ZG1 and ZY1. A 25 µL sample was prepared containing the reverse-complement single strand (which contains a nonamplifiable spacer that stops KOD XL polymerase), 50 pmol of forward primer, 7.5 nmol each of dATP, dGTP, modified ^{X}dUTP (X = Am, Im, Bn, or HO), and modified ^{X}dCTP (X = HO, COOH, or Bn), 20 µCi of α-^{32}P-dCTP (800 Ci/mmol), 2.5 µL of 10× KOD XL polymerase buffer, 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 to provide the modified DNA pools for the next selection round.

Procedure for VZ1 selection step in round 1. The 5′-thiol-modified DNA oligonucleotide (5'-GAAGAGATGGCGACTTCG-3') was synthesized as 5′-HO-C_6-S-S-C_6-p-DNA-3′. A 20 µL sample containing 500 pmol of 5′-thiol-modified DNA oligonucleotide in 100 mM HEPES, pH 7.5, and 50 mM DTT was incubated at 37 °C for 2 h. The deprotected product was precipitated to remove DTT by addition of 80 µL of water, 10 µL of 3 M NaCl, and 300 µL of ethanol. The selection experiment was initiated with 200 pmol of the ligated N_{40} pool. A 20 µL sample containing 200 pmol of ligated N_{40} pool and 500 pmol of deprotected 5′-thiol-modified DNA oligonucleotide was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The selection reaction was initiated by bringing the sample to 40 µL total volume containing 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl. The 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 VZ1 selection step in subsequent rounds. A 10 µL sample containing 100 pmol of 5′-thiol-modified DNA oligonucleotide in 100 mM HEPES, pH 7.5, and 50 mM DTT was incubated at 37 °C for 2 h. The deprotected product was precipitated to remove DTT by addition of 90 µL of water, 10 µL of 3 M NaCl, and 300 µL of ethanol. A 10 µL sample containing ligated pool and 100 pmol of deprotected 5′-thiol-modified oligonucleotide was annealed in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The selection reaction was initiated by bringing the sample to 20 µL total volume containing 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl. The sample was incubated at 37 °C for 14 h and precipitated with ethanol.

3.4.3.3 DP2-DX2 Selections

Procedures for ligation, selection, capture, and PCR were the same as described in Section 3.4.3.2. The primer extension step was adjusted to incorporate ^{E}dU, followed by an additional CuAAC step to incorporate hydrophobic modifications using the corresponding azide compound synthesized as described in Section 3.4.4.2.

Procedure for primer extension to prepare the ^{E}dU random pool. A 50 µL sample was prepared containing 250 pmol of the reverse complement template pool (which contains a 3′-segment of arbitrary sequence to allow PAGE separation of the primer extension product from the template), 500 pmol of forward primer, 15 nmol each of dATP, dGTP, dCTP, and ^{E}dUTP, 5 µL of 10× KOD XL polymerase buffer, and 1 µL of KOD XL polymerase. Primer extension was performed by cycling 4 times in a PCR thermocycler according to the following program: 94 °C for 2 min, 4× (94 °C for 2 min, 47 °C for 2 min, 72 °C for 30 min), 72 °C for 30 min. The sample was separated by 8% PAGE.

Procedure for DP2-DV2 and DX2 CuAAC step to incorporate hydrophobic modifications in round 1. The sample was brought to 10 µL total volume containing 200 pmol of the ^{E}dU random pool, 100 mM HEPES, pH 7.5, 10 mM corresponding azide compound, 40 mM tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), 10 mM sodium ascorbate, 5 mM $CuSO_4$, and 60% (v/v) DMSO (6 µL). The azide compound was added from 100 mM (10×) stock solution in DMSO (1 µL). The other 5 µL of DMSO was required to dissolve the azide compound. The 400 mM (10×) THPTA, 200 mM (20×) sodium ascorbate (freshly prepared), and 100 mM (20×) $CuSO_4$ stock solutions in water were premixed and added at last. The sample was incubated at 37 °C for 1 h and precipitated with ethanol.

Procedure for DW2 CuAAC step to incorporate hydrophobic modifications in round 1. The sample was brought to 10 µL total volume containing 200 pmol of the ^{E}dU random pool, 100 mM HEPES, pH 7.5, 10 mM 4-azido phenylalanine, 40 mM tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), 10 mM sodium ascorbate, and 5 mM $CuSO_4$. The 4-azido phenylalanine was added from a 25 mM (10×) stock aqueous solution in 20 mM NaOH.

The NaOH was required to dissolve 4-azido phenylalanine in stock solution. The 400 mM (10×) THPTA, 200 mM (20×) sodium ascorbate (freshly prepared), and 100 mM (20×) $CuSO_4$ stock solutions in water were premixed and added at last. The sample was incubated at 37 °C for 1 h and precipitated with ethanol.

Procedure for primer extension to generate the ^{E}dU-modified DNA pool in subsequent rounds. A 25 µL sample was prepared containing the reverse-complement single strand (which contains a nonamplifiable spacer that stops KOD XL polymerase), 50 pmol of forward primer, 7.5 nmol each of dATP, dGTP, dCTP, and ^{E}dUTP, 20 µCi of α-^{32}P-dCTP (800 Ci/mmol), 2.5 µL of 10× KOD XL polymerase buffer, 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 to provide the ^{E}dU-modified DNA pool for the next selection round.

Procedure for DP2-DV2 and DX2 CuAAC step to incorporate hydrophobic modifications in subsequent rounds. The sample was brought to 10 µL total volume containing the ^{E}dU random pool, 100 mM HEPES, pH 7.5, 10 mM corresponding azide compound, 40 mM tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), 10 mM sodium ascorbate, 5 mM $CuSO_4$, and 60% (v/v) DMSO (6 µL). The azide compound was added from a 100 mM (10×) stock solution in DMSO (1 µL). The other 5 µL of DMSO was required to dissolve the azide compound. The 400 mM (10×) THPTA, 200 mM (20×) sodium ascorbate (freshly prepared), and 100 mM (20×) $CuSO_4$ stock solutions in water were premixed and added at last. The sample was incubated at 37 °C for 1 h and precipitated with ethanol.

Procedure for DW2 CuAAC step to incorporate hydrophobic modifications in subsequent rounds. The sample was brought to 10 µL total volume containing the ^{E}dU random pool, 100 mM HEPES, pH 7.5, 10 mM 4-azido phenylalanine, 40 mM tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), 10 mM sodium ascorbate, and 5 mM $CuSO_4$. The 4-azido phenylalanine was added from 25 mM (10×) stock aqueous solution in 20 mM NaOH. The NaOH was required to dissolve 4-azido phenylalanine in stock solution. The 400 mM (10×)

THPTA, 200 mM (20×) sodium ascorbate (freshly prepared), and 100 mM (20×) $CuSO_4$ stock solutions in water were premixed and added at last. The sample was incubated at 37 °C for 1 h and precipitated with ethanol.

3.4.4 Organic Synthesis Procedures

Reagents were commercial grade and used without purification unless otherwise indicated. Dry pyridine was obtained from Acros Acroseal bottles. Thin-layer chromatography (TLC) was performed on silica gel plates pre-coated with fluorescent indicator, with visualization by UV light (254 nm). Flash column chromatography was performed with silica gel (230-400 mesh). NMR spectra were recorded on a Varian Unity instrument. The chemical shifts in parts per million (δ) are reported downfield from TMS (0 ppm) and referenced to the residual proton signal for $CDCl_3$ (7.26 ppm), CD_3OD (3.31 ppm), or DMSO-d_6 (2.50 ppm). Apparent multiplicities of ^{1}H NMR peaks are reported as s (singlet), d (doublet), t (triplet), q (quartet), or m (multiplet and overlapping spin systems), along with values for apparent coupling constants (*J*, Hz). Mass spectrometry data were obtained at the UIUC School of Chemical Sciences Mass Spectrometry Laboratory using a Waters Quattro II instrument (LR-ESI).

3.4.4.1 Synthesis Procedures for Modified Nucleotides and Modified Nucleoside Triphosphates

Synthesis of ^{Bn}dU. The ^{Bn}dU was synthesized from (*E*)-5-(2-carbomethoxyvinyl)-2′-deoxyuridine *N*-hydroxysuccinimide ester (Berry & Associates, cat. no. PY7190). The 2′-deoxyuridine (158 mg, 0.4 mmol) was dissolved in 2 mL of DMF. To this solution was added benzylamine (87.4 μL, 0.8 mmol) and *N*,*N*-diisopropylethylamine (DIPEA; 139 μL, 0.8 mmol), and the solution was stirred overnight at room temperature. The solution was dried on a rotary evaporator. The residue was dissolved in 1 mL of acetone and dry-loaded on a silica gel column. The column was eluted with 5−20% CH_3OH in CH_2Cl_2 to yield the desired product as a white solid (80 mg, 52%).

^{1}H NMR: (500 MHz, DMSO-d_6) δ 11.55 (s, 1H), 8.59 (t, J = 6.0 Hz, 1H), 8.29 (s, 1H), 7.27 (m, 5H), 7.16 (d, J = 15.5 Hz, 1H), 7.08 (d, J = 15.5 Hz, 1H), 6.15 (t, J = 6.5 Hz, 1H), 5.26 (d, J = 4.3 Hz, 1H), 5.15 (t, J = 5.1 Hz, 1H), 4.35 (d, J = 5.1 Hz, 2H), 4.26 (m, 1H), 3.80 (dd, J = 6.7, 3.3 Hz, 1H), 3.65 (m, 1H), 3.57 (m, 1H), 2.15 (m, 2H) ppm.

ESI-MS: *m/z* calcd. for $C_{19}H_{21}N_3O_6$ $[M+H]^+$ 388.1; found 388.1.

Synthesis of ^{Bn}dU 5′-triphosphate (^{Bn}dUTP). The ^{Bn}dUTP was synthesized from ^{Bn}dU. The procedure was similar to the synthesis of ^{COOH}dUTP described in Chapter 2. The triphosphorylation procedure was modified from the procedure reported by Huang and coworkers.[26] 2-Chloro-1,3,2-benzodioxaphosphorin-4-one (TCI, cat. no. C1210) was portioned into a dried vial in a glove box. ^{Bn}dU and tributylammonium pyrophosphate (Sigma, cat. no. P8533) were dried under vacuum. Tributylamine was dried over 4 Å molecular sieves. Tributylamine (0.65 mL, 2.74 mmol) was added to a solution of tributylammonium pyrophosphate (188 mg, 0.34 mmol) dissolved in 1 mL of anhydrous DMF under argon. This solution was added into a stirred solution of 2-chloro-1,3,2-benzodioxaphosphorin-4-one (50 mg, 0.25 mmol) in 1 mL of anhydrous DMF under argon. The resulting solution was stirred at room temperature for 1 h and then added dropwise over 5 min to a solution of ^{Bn}dU (62 mg, 0.16 mmol) in 1 mL of anhydrous DMF under argon, cooled in an ice-salt bath at 0 to –10 °C. The solution was stirred for 5 h and raised to room temperature for 30 min. A solution of 0.02 M iodine in THF/pyridine/water (Glen Research, 8 mL) was added until a permanent brown color was maintained. The brown solution was stirred at room temperature for 30 min. Two sample volumes of water were added. The two-layer mixture was stirred vigorously at room temperature for 1.5 h, followed by addition of 0.1 volumes of 3 M NaCl and 3 volumes of ethanol. The thoroughly mixed sample was precipitated at –80 °C overnight and centrifuged for 30 min at 10,000 × g and 4 °C. The precipitated sample of ^{Bn}dUTP was dried, dissolved in water, and purified by HPLC [Shimdazu 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 40% solvent A and 60% solvent B at 45

min with flow rate of 3.5 mL/min]. The ^{Bn}dUTP was isolated as its triethylammonium salt (30 mg, 14.0% from ^{Bn}dU).

ESI-MS: *m/z* calcd. for $C_{19}H_{24}N_3O_{15}P_3$ $[M–H]^-$ 626.0; found 626.4.

Synthesis of (E)-5-(2-carboxyvinyl)-2′-deoxycytidine. This synthesis followed the procedure reported by Okamoto and co-workers.[27] 5-Iodo-2′-deoxycytidine (Ark Pharm, cat. no. AK-54708; 1.41 g, 4.0 mmol), palladium(II) acetate (90 mg, 0.4 mmol), and triphenylphosphine (315 mg, 1.2 mmol) were suspended in *N*-methyl-2-pyrrolidone (NMP; 5 mL). Methyl acrylate (1.4 mL, 16 mmol) and triethylamine (1.1 mL, 8.0 mmol) were added to the suspension, which was stirred at 100 °C for 5 h under argon. The reaction mixture was filtered and concentrated under reduced pressure. The residue was dissolved in 0.5 M aqueous NaOH (15 mL). The solution was stirred at room temperature for 4 h. After addition of acetic acid (0.87 mL, 15 mmol) to neutralize the reaction mixture, the solvent was evaporated. The product was purified by silica gel column chromatography, eluting with 20−50% CH_3OH and 1% acetic acid in CH_2Cl_2. After evaporation, the residue was dissolved in a small amount of methanol (~1 mL), and acetonitrile (~10 mL) was added to the solution. The precipitate was filtered, washed with acetonitrile, and dried under reduced pressure. The product was obtained as a white solid (497 mg, 42%).

^{1}H NMR: (500 MHz, CD_3OD) δ 8.61 (s, 1H), 7.26 (d, *J* = 15.8 Hz, 1H), 6.28 (d, *J* = 15.7 Hz, 1H), 6.25 (t, *J* = 6.2 Hz, 1H), 4.42 (dt, *J* = 6.2 ,4.3 Hz, 1H), 3.97 (dd, *J* = 6.9 ,3.2 Hz, 1H), 3.90 (dd, *J* = 11.8 ,2.9 Hz, 1H), 3.78 (dd, *J* = 11.8 ,3.2 Hz, 1H), 2.40 (ddd, *J* = 13.6, 6.3, 4.5 Hz, 1H), 2.21 (dt, *J* = 13.6, 6.2 Hz, 1H) ppm.

ESI-MS: *m/z* calcd. for $C_{12}H_{15}N_3O_6$ $[M–H]^-$ 296.1; found 296.3.

Synthesis of ^{Bn}dC. The (*E*)-5-(2-carboxyvinyl)-2′-deoxycytidine (475 mg, 1.6 mmol) was dissolved in 5 mL of DMF. To this solution was added *N*-hydroxysuccinimide (NHS; 395 mg, 3.4 mmol), and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC; 660 mg, 3.4 mmol). The mixture was stirred at room temperature for 4 h. Benzylamine (375 μL, 3.4 mmol) and DIPEA (598 μL, 3.4 mmol) were added, and the solution was stirred overnight at room temperature. The solution was dried on a rotary evaporator. The residue was dissolved in 5 mL of

acetone and dry-loaded on a silica gel column. The column was eluted with 5−20% CH_3OH in CH_2Cl_2 to yield the desired product as a white solid (230 mg, 37%).

^{1}H NMR: (500 MHz, DMSO-d_6) δ 8.77 (t, *J* = 6.1 Hz, 1H), 8.53 (s, 1H), 7.38 (d, *J* = 15.6 Hz, 1H), 7.30 (m, 5H), 6.41 (d, *J* = 15.6 Hz, 1H), 6.13 (t, *J* = 6.2 Hz, 1H), 4.36 (d, *J* = 6.1 Hz, 2H), 4.26 (m, 1H), 3.83 (dd, *J* = 6.6, 3.2 Hz, 1H), 3.70 (dd, *J* = 11.6, 3.2 Hz, 1H), 3.61 (dd, *J* = 11.7, 3.1 Hz, 1H), 2.19 (m, 1H), 2.03 (m, 1H) ppm.

ESI-MS: *m/z* calcd. for $C_{19}H_{22}N_4O_5$ $[M+H]^+$ 387.2; found 387.1.

Synthesis of ^{Bn}dC 5′-triphosphate (^{Bn}dCTP). The ^{Bn}dCTP was synthesized from ^{Bn}dC. The procedure was same as for the synthesis of ^{Bn}dUTP, only with ^{Bn}dC (63 mg, 0.16 mmol) in place of ^{Bn}dU. The ^{Bn}dCTP was obtained as its triethylammonium salt (23 mg, 13.5% from ^{Bn}dC).

ESI-MS: *m/z* calcd. for $C_{19}H_{25}N_4O_{14}P_3$ $[M–H]^–$ 625.1; found 625.3.

3.4.4.2 Synthesis Procedures for Azide Compounds

Synthesis of 2,6-bis(azidomethyl)naphthalene. To a solution of 2,6-bis(bromomethyl)naphthalene (Ark Pharm, cat. no. AK133130; 314 mg, 1.0 mmol) in 5 mL of DMSO was added sodium azide (163 mg, 2.5 mmol), and the solution was stirred for 24 h at room temperature. Water (30 mL) was added, and the product was extracted with ethyl acetate (EtOAc) (3 × 10 mL). The organic layer was washed with water (3 × 10 mL), and the solvent was evaporated to yield the desired product as a white solid (231 mg, 97%).

^{1}H NMR: (500 MHz, $CDCl_3$) δ 7.88 (d, *J* = 8.3 Hz, 2H), 7.79 (s, 2H), 7.46 (dd, *J* = 8.4, 1.3 Hz, 2H), 4.52 (s, 4H) ppm.

ESI-MS: *m/z* calcd. for $C_{12}H_{10}N_6$ $[M+H]^+$ 239.1; found 239.1.

Synthesis of Az-Naph-NH_2. To a mixture of 2,6-bis(azidomethyl)naphthalene (100 mg, 0.42 mmol) in 4 mL of 1:1 (v/v) Et_2O/EtOAc and 3 mL of 1 M aqueous HCl, triphenylphosphine (132 mg, 0.50 mmol) was added, and the heterogenous mixture was stirred vigorously for 24 h at room temperature. The organic layer was separated, and the aqueous layer was extracted with EtOAc (3 × 10 mL) to remove triphenylphosphine oxide and remaining starting materials. The pH of the

aqueous layer was adjusted to 10 by adding 2 M aqueous NaOH (~3 mL) and extracted with CH_2Cl_2 (3 × 10 mL). The combined extracts were dried with anhydrous Na_2SO_4, filtered, and evaporated. The residue was purified by silica gel column chromatography, eluting with 0−50% CH_3OH in CH_2Cl_2 to yield the desired product as a white solid (60 mg, 67%).

^{1}H NMR: (500 MHz, DMSO-d_6) δ 7.91 (m, 4H), 7.60 (dd, *J* = 8.3, 1.7 Hz, 1H), 7.49 (dd, *J* = 8.4, 1.8 Hz, 1H), 4.62 (s, 2H), 3.99 (s, 2H) ppm.

ESI-MS: *m/z* calcd. for $C_{12}H_{12}N_4$ $[M+H]^+$ 213.1; found 213.1.

Synthesis of Cl-HO-NH-Bn. To a solution of benzylamine (546 μL, 536 mg, 5 mmol) in CH_3OH (10 mL) was added (*S*)-epichlorohydrin (Synthonix, cat. no. E10841; 392 μL, 463 mg, 5 mmol), and the solution was stirred for 48 h at room temperature. After evaporation under reduced pressure, the residue was purified by silica gel column chromatography, eluting with 0−5% CH_3OH in CH_2Cl_2 to yield the product as a white solid (530 mg, 53%).

^{1}H NMR: (500 MHz, CD_3OD) δ 7.32 (m, 5H), 3.90 (m, 1H), 3.81 (d, *J* = 13.1 Hz, 1H), 3.76 (d, *J* = 13.1 Hz, 1H), 3.56 (dd, *J* = 11.2, 5.2 Hz, 1H), 3.52 (dd, *J* = 11.2, 5.8 Hz, 1H), 2.77 (dd, *J* = 12.2, 4.1 Hz, 1H), 2.63 (dd, *J* = 12.2, 8.1 Hz, 1H) ppm.

ESI-MS: *m/z* calcd. for $C_{10}H_{14}ClNO$ $[M+H]^+$ 200.1; found 200.0.

Synthesis of Az-HO-NH-Bn. To a solution of Cl-HO-NH-Bn (130 mg, 0.65 mmol) in 3 mL of DMF was added sodium azide (63.4 mg, 0.98 mmol), and the solution was stirred at 80 °C for 24 h. Water (3 mL) and saturated aqueous NaCl (5 mL) were added, and the product was extracted with EtOAc (3 × 5 mL). The organic layers were combined, washed with saturated aqueous NaCl (3 × 5 mL), dried with anhydrous Na_2SO_4, and filtered. The solvent was evaporated to yield the desired product as a white solid (125 mg, 93%).

^{1}H NMR: (500 MHz, $CDCl_3$) δ 7.34 (m, 5H), 3.94 (m, 1H), 3.88 (m, 2H), 3.33 (m, 1H), 3.25 (m, 1H), 2.77 (dd, *J* = 12.3, 3.5 Hz, 1H), 2.70 (dd, *J* = 12.3, 8.8 Hz, 1H) ppm.

ESI-MS: *m/z* calcd. for $C_{10}H_{14}N_4O$ $[M+H]^+$ 207.1; found 207.0.

Synthesis of Cl-HO-NH-Naph. To a solution of 1-(2-naphthyl)methanamine (Ark Pharm, cat. no. AK-33985; 314 mg, 2 mmol) in DMSO (10 mL) was added (*S*)-epichlorohydrin (157 μL,

185 mg, 2 mmol), and the solution was stirred for 48 h at room temperature. After evaporation under reduced pressure, the oily residue was purified by silica gel column chromatography, eluting with 0−20% CH_3OH in CH_2Cl_2 to yield the product as a white solid (110 mg, 22%).

^{1}H NMR: (500 MHz, CD_3OD) δ 7.85 (m, 4H), 7.47(m, 3H), 3.97 (m, 3H), 3.58 (dd, *J* = 11.2, 5.2 Hz, 1H), 3.53 (dd, *J* = 11.2, 5.8 Hz, 1H), 2.85 (dd, *J* = 12.3, 4.0 Hz, 1H), 2.71 (dd, *J* = 12.3, 8.1 Hz, 1H) ppm.

ESI-MS: *m/z* calcd. for $C_{14}H_{16}ClNO$ $[M+H]^+$ 250.1; found 250.0.

Synthesis of Az-HO-NH-Naph. To a solution of Cl-HO-NH-Naph (60 mg, 0.24 mmol) in 3 mL of DMSO was added sodium azide (23.4 mg, 0.36 mmol), and the solution was stirred at 80 °C for 24 h. Water (3 mL) and saturated aqueous NaCl (5 mL) were added, and the product was extracted with EtOAc (3 × 5 mL). The organic layers were combined, washed with saturated aqueous NaCl (3 × 5 mL), dried with anhydrous Na_2SO_4, and filtered. The solvent was evaporated to yield the desired product as a white solid (56 mg, 91%).

^{1}H NMR: (500 MHz, CD_3OD) δ 7.84 (m, 4H), 7.47(m, 3H), 4.00 (m, 2H), 3.91 (m, 1H), 3.29 (m, 2H), 2.74 (dd, *J* = 12.3, 4.5 Hz, 1H), 2.70 (dd, *J* = 12.3, 7.9 Hz, 1H) ppm.

ESI-MS: *m/z* calcd. for $C_{14}H_{16}N_4O$ $[M+H]^+$ 257.1; found 257.1.

Synthesis of Cl-HO-NH-Trp. To a solution of tryptamine (801 mg, 5 mmol) in CH_3OH (10 mL) was added (*S*)-epichlorohydrin (392 μL, 463 mg, 5 mmol), and the solution was stirred for 48 h at room temperature. After evaporation under reduced pressure, the residue was purified by silica gel column chromatography, eluting with 0−20% CH_3OH in CH_2Cl_2 to yield the product as a brown oil (703 mg, 56%).

^{1}H NMR: (500 MHz, CD_3OD) δ 7.56 (d, *J* = 7.9 Hz, 1H), 7.34 (d, *J* = 8.1 Hz, 1H), 7.06 (m, 3H), 3.94 (m, 1H), 3.52 (m, 2H), 3.07 (m, 2H), 3.04 (m, 2H), 2.95 (dd, *J* = 12.4, 3.5 Hz, 1H), 2.77 (dd, *J* = 12.4, 8.9 Hz, 1H) ppm.

ESI-MS: *m/z* calcd. for $C_{13}H_{17}ClN_2O$ $[M+H]^+$ 253.1; found 253.0.

Synthesis of Az-HO-NH-Trp. To a solution of Cl-HO-NH-Trp (220 mg, 0.87 mmol) in 3 mL of DMF was added sodium azide (84.9 mg, 1.31 mmol), and the solution was stirred at 80 °C

for 24 h. Water (3 mL) and saturated aqueous NaCl (5 mL) were added, and the product was extracted with EtOAc (3 × 5 mL). The organic layers were combined, washed with saturated aqueous NaCl (3 × 5 mL), dried with anhydrous Na_2SO_4, and filtered. The solvent was evaporated to yield the desired product as a brown oil (214 mg, 95%).

^{1}H NMR: (500 MHz, CD_3OD) δ 7.56 (d, *J* = 7.9 Hz, 1H), 7.35 (d, *J* = 8.1 Hz, 1H), 7.06 (m, 3H), 3.90 (m, 1H), 3.27 (m, 2H), 3.07 (m, 2H), 3.03 (m, 2H), 2.83 (dd, *J* = 12.4, 3.8 Hz, 1H), 2.76 (dd, *J* = 12.4, 8.7 Hz, 1H) ppm.

ESI-MS: *m/z* calcd. for $C_{13}H_{17}N_5O$ $[M+H]^+$ 260.1; found 260.1.

Synthesis of N-*Trityl-2-(1-trityl-1*H-*imidazol-4-yl)ethanamine.* To a solution of histamine dihydrochloride (920 mg, 5 mmol) and triethylamine (2.8 mL, 2.02 g, 20 mmol) in $CHCl_3$ (12.5 mL) was added dropwise a solution of trityl chloride (3.48 g, 12.5 mmol) in $CHCl_3$ (12.5 mL) at 0 °C. The solution was stirred for 20 h at room temperature. After evaporation under reduced pressure, the yellow oily residue was suspended in 25 mL of water. After stirring for 1 h, the mixture was extracted with $CHCl_3$ (2 × 25 mL). The organic layers were combined, dried with anhydrous Na_2SO_4, filtered, and evaporated under reduced pressure to yield the product as a yellow oil (2.15 g, 72%).

ESI-MS: *m/z* calcd. for $C_{43}H_{37}N_3$ $[M+H]^+$ 596.3; found 596.3.

*Synthesis of 2-(1-trityl-1*H-*imidazol-4-yl)ethanamine.* To a solution of *N*-trityl-2-(1-trityl-1*H*-imidazol-4-yl)ethanamine (2.15 g, 3.6 mmol) in CH_2Cl_2 (24 mL) at 0 °C was added dropwise 1.25 mL of TFA. After stirring for 10 min at 0 °C, the solution was stirred at room temperature for 45 min. After evaporation under reduced pressure, the yellow oily residue was neutralized with 25 mL of saturated aqueous $NaHCO_3$ and extracted with $CHCl_3$ (4 × 25 mL). The organic layers were combined, washed with water (25 mL) and saturated aqueous NaCl (25 mL), dried with anhydrous Na_2SO_4, and filtered. After evaporation under reduced pressure, the yellow oily residue was purified by silica gel column chromatography, eluting with 0−5% CH_3OH and 1% triethylamine in $CHCl_3$ to yield the product as a yellow oil (996 mg, 78%).

^{1}H NMR: (500 MHz, $CDCl_3$) δ 8.24 (s, 1H), 7.37 (m, 9H), 7.18 (m, 6H), 7.07 (s, 1H), 3.47 (t, *J* = 6.8 Hz, 2H), 3.34 (t, *J* = 6.8 Hz, 2H) ppm.

ESI-MS: *m/z* calcd. for $C_{24}H_{23}N_3$ $[M+H]^+$ 354.2; found 354.2.

Synthesis of Cl-HO-NH-Im-Tr. To a solution of 2-(1-trityl-1*H*-imidazol-4-yl)ethanamine (707 mg, 2 mmol) in CH_3OH (10 mL) was added (*S*)-epichlorohydrin (157 μL, 185 mg, 2 mmol), and the solution was stirred for 48 h at room temperature. After evaporation under reduced pressure, the yellow oily residue was purified by silica gel column chromatography, eluting with 0−10% CH_3OH and 1% triethylamine in CH_2Cl_2 to yield the product as a white solid (417 mg, 47%).

^{1}H NMR: (500 MHz, CD_3OD) δ 7.39 (s, 1H), 7.37 (m, 9H), 7.15 (m, 6H), 6.74 (s, 1H), 3.86 (m, 1H), 3.51 (m, 2H), 2.87 (m, 2H), 2.75 (m, 2H), 2.68 (m, 2H) ppm.

ESI-MS: *m/z* calcd. for $C_{27}H_{28}ClN_3O$ $[M+H]^+$ 446.2; found 446.2.

Synthesis of Az-HO-NH-Im-Tr. To a solution of Cl-HO-NH-Im-Tr (400 mg, 0.90 mmol) in 3 mL of DMF was added sodium azide (87.8 mg, 1.35 mmol), and the solution was stirred at 80 °C for 24 h. Water (3 mL) and saturated aqueous NaCl (5 mL) were added, and the solution was extracted with EtOAc (3 × 5 mL). The organic layers were combined, washed with saturated aqueous NaCl (3 × 5 mL), dried with anhydrous Na_2SO_4, and filtered. The solvent was evaporated to yield the desired product as a yellow oil (404 mg, 99%).

^{1}H NMR: (500 MHz, CD_3OD) δ 7.98 (s, 1H), 7.38 (m, 9H), 7.15(m, 6H), 6.74 (s, 1H), 3.82 (m, 1H), 3.25 (m, 2H), 2.85 (m, 2H), 2.72 (m, 2H), 2.62 (m, 2H) ppm.

ESI-MS: *m/z* calcd. for $C_{27}H_{28}N_6O$ $[M+H]^+$ 453.2; found 453.2.

Synthesis of Az-HO-NH-Im. To a solution of Az-HO-NH-Im-Tr (400 mg, 0.88 mmol) in 15 mL of CH_3OH was added dropwise 5 mL of concentrated HCl, and the solution was stirred at room temperature for 24 h. The solvents were evaporated, and 20 mL of water was added. The mixture was washed with CH_2Cl_2 (3 × 10 mL) to remove the triphenylmethanol. The aqueous phase was evaporated, and the residue was coevaporated with ethanol (2 × 10 mL). The white solid product was obtained as the dihydrochloride salt (199 mg, 80%).

^{1}H NMR: (500 MHz, CD_3OD) δ 8.90 (s, 1H), 7.52 (s, 1H), 4.13 (m, 1H), 3.46 (m, 1H), 3.42 (m, 2H), 3.35 (m, 1H), 3.24 (m, 2H), 3.21 (m, 1H), 3.13 (m, 1H) ppm.

ESI-MS: *m/z* calcd. for $C_8H_{14}N_6O$ $[M+H]^+$ 211.1; found 211.1.

3.5 References

(1) Zhou, C.; Avins, J. L.; Klauser, P. C.; Brandsen, B. M.; Lee, Y.; Silverman, S. K. *J. Am. Chem. Soc.* **2016**, *138*, 2106.

(2) Santoro, S. W.; Joyce, G. F.; Sakthivel, K.; Gramatikova, S.; Barbas, C. F., III. *J. Am. Chem. Soc.* **2000**, *122*, 2433.

(3) Lermer, L.; Roupioz, Y.; Ting, R.; Perrin, D. M. *J. Am. Chem. Soc.* **2002**, *124*, 9960.

(4) Sidorov, A. V.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2004**, *32*, 1591.

(5) Raines, R. T. *Chem. Rev.* **1998**, *98*, 1045.

(6) Thomas, J. M.; Yoon, J.; Perrin, D. M. *J. Am. Chem. Soc.* **2009**, *131*, 5648.

(7) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. *Nucleic Acids Res.* **2009**, *37*, 1638.

(8) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. *ChemBioChem* **2009**, *10*, 1988.

(9) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. *ACS Comb. Sci.* **2013**, *15*, 174.

(10) Vaught, J. D.; Bock, C.; Carter, J.; Fitzwater, T.; Otis, M.; Schneider, D.; Rolando, J.; Waugh, S.; Wilcox, S. K.; Eaton, B. E. *J. Am. Chem. Soc.* **2010**, *132*, 4141.

(11) Davies, D. R.; Gelinas, A. D.; Zhang, C.; Rohloff, J. C.; Carter, J. D.; O'Connell, D.; Waugh, S. M.; Wolk, S. K.; Mayfield, W. S.; Burgin, A. B.; Edwards, T. E.; Stewart, L. J.; Gold, L.; Janjic, N.; Jarvis, T. C. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 19971.

(12) Gupta, S.; Hirota, M.; Waugh, S. M.; Murakami, I.; Suzuki, T.; Muraguchi, M.; Shibamori, M.; Ishikawa, Y.; Jarvis, T. C.; Carter, J. D.; Zhang, C.; Gawande, B.; Vrkljan, M.; Janjic, N.; Schneider, D. J. *J. Biol. Chem.* **2014**, *289*, 8706.

(13) Gelinas, A. D.; Davies, D. R.; Edwards, T. E.; Rohloff, J. C.; Carter, J. D.; Zhang, C.; Gupta, S.; Ishikawa, Y.; Hirota, M.; Nakaishi, Y.; Jarvis, T. C.; Janjic, N. *J. Biol. Chem.* **2014**, *289*, 8720.

(14) Rohloff, J. C.; Gelinas, A. D.; Jarvis, T. C.; Ochsner, U. A.; Schneider, D. J.; Gold, L.; Janjic, N. *Mol. Ther. Nucleic Acids* **2014**, *3*, 201.

(15) Ren, X.; Gelinas, A. D.; von Carlowitz, I.; Janjic, N.; Pyle, A. M. *Nat. Commun.* **2017**, *8*, 810.

(16) Tolle, F.; Brändle, G. M.; Matzner, D.; Mayer, G. *Angew. Chem. Int. Ed.* **2015**, *54*, 10971.

(17) Erez, E.; Fass, D.; Bibi, E. *Nature* **2009**, *459*, 371.

(18) Dunbar, K. L.; Mitchell, D. A. *ACS Chem. Biol.* **2013**, *8*, 473.

(19) Bindman, N. A.; Bobeica, S. C.; Liu, W. R.; van der Donk, W. A. *J. Am. Chem. Soc.* **2015**, *137*, 6975.

(20) Chu, C.-C.; Wong, O. Y.; Silverman, S. K. *ChemBioChem* **2014**, *15*, 1905.

(21) Walsh, S. M.; Konecki, S. N.; Silverman, S. K. *J. Mol. Evol.* **2015**, *81*, 218.

(22) Wang, P.; Silverman, S. K. *Angew. Chem. Int. Ed.* **2016**, *55*, 10052.

(23) Flynn-Charlebois, A.; Wang, Y.; Prior, T. K.; Rashid, I.; Hoadley, K. A.; Coppins, R. L.; Wolf, A. C.; Silverman, S. K. *J. Am. Chem. Soc.* **2003**, *125*, 2444.

(24) Wang, Y.; Silverman, S. K. *J. Am. Chem. Soc.* **2003**, *125*, 6880.

(25) Chandrasekar, J.; Silverman, S. K. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5315.

(26) Caton-Williams, J.; Hoxhaj, R.; Fiaz, B. *Curr. Protoc. Nucleic Acid Chem.* **2013**, *52*, 1.30.1.

(27) Ikeda, S.; Yuki, M.; Yanagisawa, H.; Okamoto, A. *Tetrahedron Lett.* **2009**, *50*, 7191.

Chapter 4: DNA-Catalyzed Radical-Based Oxidative DNA Cleavage*

4.1 Introduction

DNA cleavage is common in nature and has been extensively studied due to its importance in biological processes and therapeutics.[1–4] In general, two different types of DNA cleavage can be distinguished, namely hydrolytic DNA cleavage and oxidative DNA cleavage. Hydrolytic DNA cleavage occurs either by hydrolysis of the phosphodiester linkages or by deglycosylation with subsequent elimination reactions that lead to strand scission.[1,5] Oxidative DNA cleavage usually occurs by formation of reactive radical species on 2′-deoxyriboses or nucleobases that undergo autoxidation reactions leading to strand scission.[6,7]

The oxidation of 2′-deoxyribose in DNA plays an important role in DNA damage under oxidative stress in biological systems.[8] The oxidation can be initiated by hydrogen atom abstraction at one of the five positions in 2′-deoxyribose to form a carbon radical. The subsequent radical-based reactions can lead to various DNA damage products, some of which can result in strand scission of DNA.[6]

This 2′-deoxyribose oxidation mechanism is also utilized by anticancer drugs like bleomycin to induce DNA damage that triggers cell apoptosis.[4] Bleomycin, first isolated from the bacterium *Streptomyces verticillus*,[9] is a glycopeptide natural product that can cleave both single-stranded and double-stranded DNA in the presence of a metal ion such as Fe^{2+}.[10] The oxidative

* This research has been published:
Lee, Y.; Klauser, P. C.; Brandsen, B. M.; Zhou, C.; Li, X.; Silverman, S. K. *J. Am. Chem. Soc.* **2017**, *139*, 255.

The majority of this research was performed by University of Illinois undergraduate students Yujeong Lee and Paul Klauser. I performed the in vitro selection originally intended for DNA-catalyzed amide hydrolysis (described in Chapter 2) and the initial MALDI mass spectrometry analysis. I also assisted with the hydrogen peroxide assay.

DNA cleavage, starting with formation of the C4′ radical intermediate by bleomycin, can lead to different cleavage products in aerobic and anaerobic conditions (Figure 4.1).

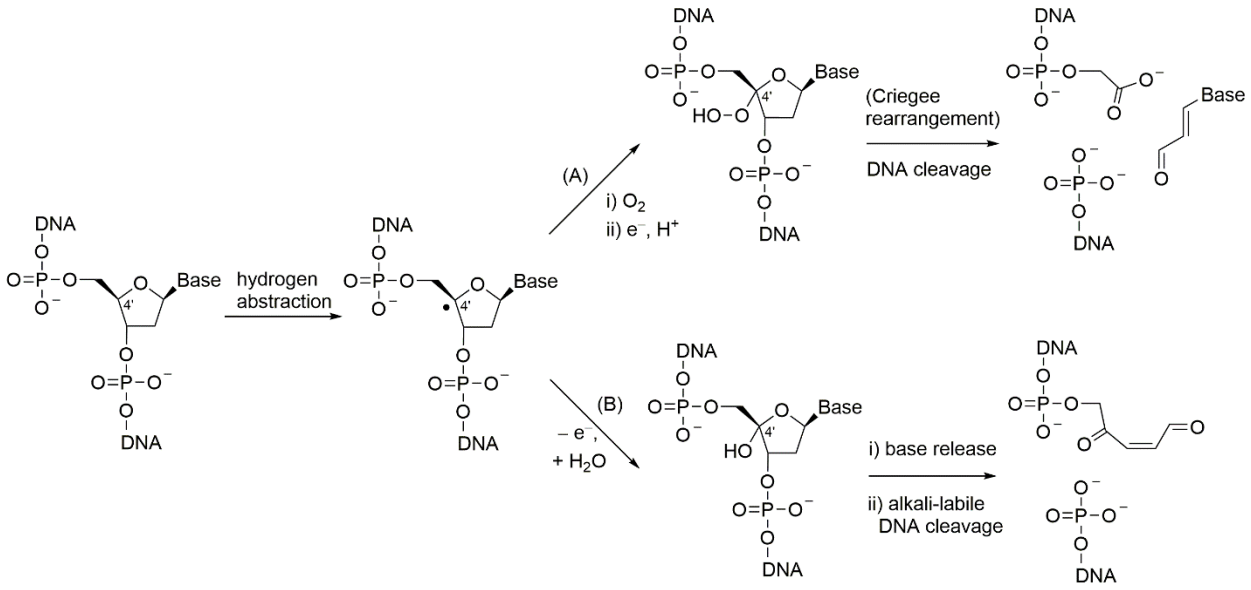

Figure 4.1. Bleomycin-induced oxidative DNA cleavage. The radical pathway is initiated by hydrogen atom abstraction to form a radical at the C4′ position of a 2′-deoxyribose in the DNA substrate. (A) Under aerobic conditions, the C4′ radical intermediate reacts with molecular oxygen. With subsequent reduction, a Criegee rearrangement leads to the formation of the 5′-phosphate, 3′-phosphoglycolate, and base propenal products. (B) Under anaerobic conditions, the C4′ radical intermediate is further oxidized and reacts with water, forming the released base, 5′-phosphate product, and an aldehyde product. Figure adapted with permission from ref. 10.

In the selection for DNA-catalyzed amide bond hydrolysis described in Chapter 2, the RadDz3 deoxyribozyme with oxidative DNA cleavage activity was also identified. The cleavage process occurs by a superoxide/H_2O_2-dependent radical-based pathway and leads to well-defined reaction products that are the same as those formed by the bleomycin-induced DNA cleavage in aerobic conditions. Unlike the previously reported examples of DNA-catalyzed oxidative DNA cleavage,[11,12] RadDz3 does not require redox-active metal ions.

4.2 Results and Discussion

In the selection experiments of Chapter 2 for DNA-catalyzed amide hydrolysis using modified nucleotides,[13] one additional deoxyribozyme named RadDz3 was identified from the ^{HO}dU selection designated JY1. RadDz3 cleaves an all-DNA substrate (Figure 4.2A) that lacks any amide bond, and all of the ^{HO}dU modifications originally in RadDz3 are dispensable for catalysis. All of the characterization experiments were performed using the unmodified version of RadDz3.

The initial MALDI mass spectrometry analysis revealed unusual DNA cleavage products that did not correspond to any possible DNA hydrolysis pathway (Figure 4.2B). Moreover, the total mass of the two cleavage products shown in MALDI mass spectrometry did not add up to the mass of the original DNA substrate, indicating an additional small-molecule fragment released in the cleavage reaction. These findings led us to consider the possibility of an oxidative DNA cleavage pathway, which could result in unusual excisions or modifications of the deoxyriboses or nucleobases.[6,7]

A radical pathway was established by Yujeong Lee and Paul Klauser. The DNA cleavage by RadDz3 was fully inhibited by adding glutathione or catalase, where glutathione quenches radical intermediates and catalase destroys H_2O_2, suggesting the involvement of one or more radical intermediates and H_2O_2. Adding H_2O_2 or potassium superoxide (KO_2) significantly enhanced the catalysis.

Based on the MALDI mass spectrometry analysis and the observations in the characterization assays described above, a bleomycin-like radical pathway (Figure 4.2C) was considered. In the bleomycin radical pathway, a 3′-phosphoglycolate product is formed by the abstraction of a hydrogen atom from the C4′ position of a nucleoside followed by the addition of an O_2 molecule and Criegee rearrangement. This 3′-phosphoglycolate product is a carboxylic acid, which can be captured during the in vitro selection process described in Chapter 2, explaining the survival of RadDz3. Of the various 2′-deoxyribose radical intermediates that can be formed by

hydrogen atom abstraction, only the C4′ radical intermediate as formed in the bleomycin radical pathway leads to a carboxylic acid.[6,8]

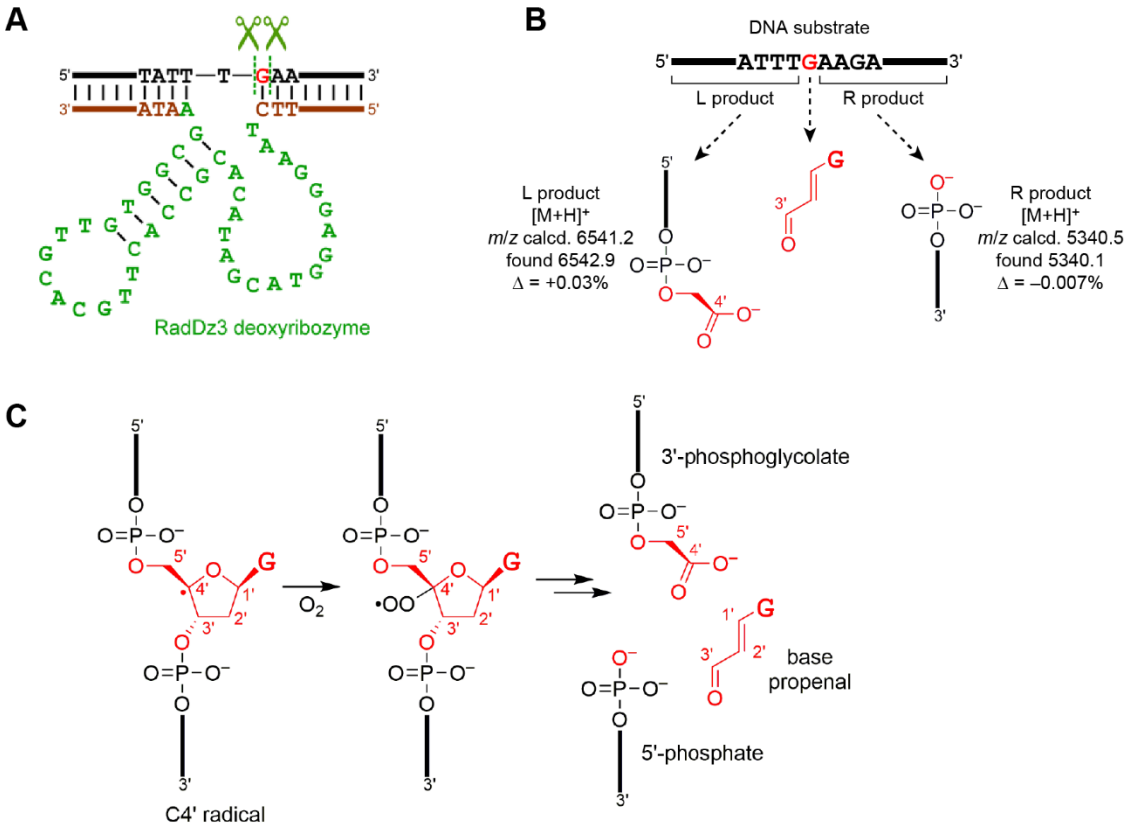

Figure 4.2. The RadDz3 deoxyribozyme for oxidative DNA cleavage by a radical pathway. (A) Mfold-predicted[14] secondary structure of RadDz3 and the cleavage sites within the DNA substrate. (B) Structures and MALDI mass spectrometry analysis of the RadDz3 cleavage products. (C) The bleomycin radical pathway for oxidative DNA cleavage. Figure adapted with permission from ref. 15.

The RadDz3 cleavage products analyzed by MALDI mass spectrometry revealed formation of the 3′-phosphoglycolate and 5′-phosphate products by excision of a particular guanosine nucleoside of the DNA substrate(Figure 4.2B). The small-molecule base propenal fragment formed from the excised mononucleoside was also detected via a colorimetric assay[16] performed by Paul Klauser and Xinyi Li, thereby accounting for all DNA substrate atoms.

In the metal ion dependence assay performed by Yujeong Lee and Paul Klauser, Zn^{2+} and Mg^{2+} together were found to be sufficient for observable RadDz3 activity, demonstrating that RadDz3 does not require a redox-active metal ion cofactor.

4.3 Summary and Future Directions

A new DNA-catalyzed oxidative DNA cleavage reaction was identified, with a radical-based pathway and well-defined cleavage products in the presence of redox-inactive metal ion cofactors. Additional efforts are needed to understand the detailed mechanism of the selective C4′ radical formation, the role of superoxide and H_2O_2 in the reaction pathway, and the role of the metal ion cofactors in catalysis.

The similar DNA-catalyzed oxidative DNA cleavage was identified in the SR1 and ST1 selections intended for DNA-catalyzed peptide hydrolysis in Chapter 3. This reaction can readily interfere with the selection experiments involving capture of carboxylic acid products and therefore must be considered in the selection designs. One example is to use substrates that have only the left-hand DNA anchor to avoid such carboxylic acid products being captured, as shown in Section 3.2.2 and Section 3.2.3.

The DNA-catalyzed oxidative DNA cleavage revealed the general ability of DNA enzymes to catalyze redox reactions. Others have reported that artificial ribozymes can catalyze redox reactions such as alcohol oxidation.[17] Selections using DNA could be performed to identify deoxyribozymes that catalyze such alcohol oxidation and likely other types of redox reactions.

4.4 Materials and Methods

Oligonucleotide preparation, in vitro selection, and MALDI mass spectrometry assay were performed as described in Chapter 2.

4.5 References

(1) Lindahl, T. *Nature* **1993**, *362*, 709.

(2) Bashkin, J. K. *Chem. Rev.* **1998**, *98*, 937.

(3) Kim, H.; Kim, J. *Nat. Rev. Genet.* **2014**, *15*, 321.

(4) Chen, J.; Stubbe, J. *Nat. Rev. Cancer* **2005**, *5*, 102.

(5) Molina, R.; Stella, S.; Redondo, P.; Gomez, H.; Marcaida, M. J.; Orozco, M.; Prieto, J.; Montoya, G. *Nat. Struct. Mol. Biol.* **2015**, *22*, 65.

(6) Pogozelski, W. K.; Tullius, T. D. *Chem. Rev.* **1998**, *98*, 1089.

(7) Burrows, C. J.; Muller, J. G. *Chem. Rev.* **1998**, *98*, 1109.

(8) Dedon, P. C. *Chem. Res. Toxicol.* **2008**, *21*, 206.

(9) Umezawa, H.; Maeda, K.; Takeuchi, T.; Okami, Y. *J Antibiot* **1966**, *19*, 200.

(10) Burger, R. M. *Chem. Rev.* **1998**, *98*, 1153.

(11) Carmi, N.; Breaker, R. R. *Bioorg. Med. Chem.* **2001**, *9*, 2589.

(12) Wang, M.; Dong, J.; Zhang, H.; Tang, Z. *Org. Biomol. Chem.* **2016**, *14*, 2347.

(13) Zhou, C.; Avins, J. L.; Klauser, P. C.; Brandsen, B. M.; Lee, Y.; Silverman, S. K. *J. Am. Chem. Soc.* **2016**, *138*, 2106.

(14) Zuker, M. *Nucleic Acids Res.* **2003**, *31*, 3406.

(15) Lee, Y.; Klauser, P. C.; Brandsen, B. M.; Zhou, C.; Li, X.; Silverman, S. K. *J. Am. Chem. Soc.* **2017**, *139*, 255.

(16) Burger, R. M.; Berkowitz, A. R.; Peisach, J.; Horwitz, S. B. *J. Biol. Chem.* **1980**, *255*, 11832.

(17) Tsukiji, S.; Pattnaik, S. B.; Suga, H. *Nat. Struct. Biol.* **2003**, *10*, 713.