EXPANDING SELEX TECHNOLOGY TO ENGINEER FUNCTIONAL NUCLEIC ACIDS OPTIMIZED FOR THE DEVELOPMENT OF SMALL-MOLECULE BIOSENSORS

by

Alexandra Eusebia Rangel

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

The University of Utah

May 2018

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Alexandra Eusebia Rangel**

has been approved by the following supervisory committee members:

Jennifer Heemstra	, Chair	**Nov 30, 2017** Date Approved
Joel Harris	, Member	**Nov 30, 2017** Date Approved
Shelley Minteer	, Member	**Nov 30, 2017** Date Approved
Ilya Zharov	, Member	**Nov 30, 2017** Date Approved
Julie Hollien	, Member	**Nov 30, 2017** Date Approved

and by **Cynthia Burrows**, Chair/Dean of the Department/College/School of **Chemistry**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

The development of systematic evolution of ligands by exponential enrichment, or SELEX, in 1990 propelled interest in nucleic acid engineering. The ability of nucleic acids to undergo directed evolution has resulted in the widespread interest in the development of aptamers, or single stranded nucleic acids that possess affinity to a specific target molecule. Aptamers exist to diverse classes of molecules, and because of their versatility are being regarded as robust recognition elements for a variety of biotechnology applications. Further, the small-molecule harnessing potential of aptamers has shown great promise for the development of biosensors. Importantly, traditional selection methods can be expanded upon to enable the generation of functional nucleic acids which are optimized for biosensing platforms.

Novel SELEX methods are a powerful tool to develop innovative technologies for detection and imaging using aptamer-based biosensors. Our lab has contributed to a variety of areas within this field including the use of ribozymes as a tag for intracellular imaging of RNA. We have designed a novel IP-SELEX method which allows for the generation of ribozymes that are capable of performing a reaction to covalently attach a small-molecule fluorophore to itself (Chapter 2). We show that this ribozyme can label in cellular conditions and hold great promise as a genetically encodable tag for live cell labeling of mRNA.

We have also developed a new class of small-molecule sensors through the evolution of nucleic acids with noncanonical backbone structures, or xeno nucleic acids (XNA). XNA molecules have many advantageous properties because of their noncanonical structures, notably nuclease resistance. Using our unique method, optimized by the use of an XNA primer, we were able generate the first artificial genetic polymers capable of small-molecule recognition (Chapter 3). We discovered threose nucleic acid (TNA) aptamers with affinity for a small-molecule mycotoxin (OTA). We show that these aptamers have outstanding biostability in the presence of nucleases and retain the ability to bind the target in these environments.

We acknowledge the potential for structure-switching (SS) aptamer biosensors as a privileged architecture for small-molecule detection and the development of fluorescence assays. We propose an innovative technique for the direct selection of SS biosensors, in which we take advantage of both restriction digests and polymerase chain reaction (PCR) as key steps to eliminate nonfunctional sequences. Importantly, because amplification simultaneously serves as the selection step there is no requirement for a solid support. We anticipate this will overcome the limitations of bead-based selection methods and enable for efficient and effective generation of SS biosensors.

For my grandma who always believed in me
and to my family (and cats) who have
supported me through this journey

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

λ_{em}	emission wavelength
λ_{ex}	excitation wavelength
AcCoA	acetyl coA
ADP	adenosine diphosphate
Apt	aptamer
DNA	2'-deoxyribonucleic acid
DS	displacement strand
cAMP	cyclic adenosine monophosphate
cDNA	complementary DNA
CoA	coenzyme A
CNG5	cyclic nucleotide gated subunit 5
CS	complementary strand
ETBR	ethidium bromide
FAM	fluorescein
FIA	fluorescein iodoacetamide
FP	fluorescence polarization
GPCR	G protein-coupled receptor
HMT	histone methyltransferase
HPLC	high pressure liquid chromatography

HTS high-throughput screening

IP immunoprecipitation

K_D dissociation constant

mRNA messenger ribonucleic acid

MS mass spectrometry

MST microscale thermophoresis

Nt nucleotides

OTA ochratoxin A

PAGE polyacrylamide gel electrophoresis

PBS primer binding site

PCR polymerase chain reaction

PEG polyethylene glycol

RNA ribonucleic acid

SELEX systematic evolution of ligands by exponential enrichment

SS structure-switching

TMRIA tetramethyl rhodamine iodoacetamide

TNA threose nucleic acid

tNTP threose nucleic acid triphosphate

TRIP transcript-specific immunoprecipitation

UTR untranslated region

XNA xeno nucleic acid

ACKNOWLEDGEMENTS

I would like to first thank my research advisor Prof. Jennifer Heemstra for being an exceptional mentor throughout my graduate school career. I am extremely grateful for the advice and insight she has shared with me over the years, which has been invaluable to my personal and scientific growth.

I would like to thank my committee members Professors Joel Harris, Julie Hollien, Shelley Minteer, and Ilya Zharov for their time, guidance, and support. I would also like to thank the Department of Chemistry and the University of Utah for providing me with financial support. I also extend my gratitude to the faculty and staff of the department.

I am also grateful for my group members who created a unique atmosphere of collegiality and comradery: Ashwani Sharma, Amberlyn Peterson, Tilani de Costa, Brad Green, Trevor Feagin, Kirsten Meek, Petr Simon, Zhe Chen, Satish Ellipili, Nicholas Spiropulos, David Olsen, Zhesen Tan, Tewoderos Ayele, Collin Swenson, Erin Price, Hershel Lackey, Misael Romero, Riley Giesler, Steve Knutson, April Anamisis, Alexandra Kent, and Evelyn Kimbrough.

Most of all I would like to thank my family for their love and support through this process, especially my sister whom I can always lean on. They helped me through the toughest times and I would have not succeeded without them. Last but not least, thank you Nick for all the adventures and laughs that helped me through.

CHAPTER 1

APTAMERS WITH UNIQUE PROPERTIES AND FUNCTION

Introduction*

Nucleic acids are best known for their ability to encode, store, and decode genetic information in cells. However, the functional capacity of DNA and RNA stretches far beyond information storage to encompass complex tasks such as regulation and catalysis. These capabilities have emerged as a result of evolution over millions, or even billions, of years. Inspired by the ability of evolution to provide nucleic acids having this large diversity of functions, researchers have sought to replicate the evolutionary process in the laboratory. In 1990, the Szostak, Joyce, and Gold labs individually reported the successful implementation of selection and evolution *in vitro*, and this process was termed Systematic Evolution of Ligands by EXponential enrichment (SELEX).[1-3] Over the past 25 years, SELEX has been most broadly applied for the *in vitro* evolution of aptamers, which are nucleic acid sequences capable of binding to specific small-molecule or protein targets. Using the SELEX process, functional nucleic acid sequences can be enriched and

*Adapted from Methods, Vol. 106, Meek, K. N.; Rangel, A. E.; Heemstra, J. M., Enhancing aptamer function and stability via in vitro selection using modified nucleic acids, Pages 29-36, Copyright 2016, with permission from Elsevier.

evolved from a starting library on the time scale of days, providing a powerful route to rapidly generate biopolymers having new capabilities.[4] To date, over 1000 well-characterized aptamers to over 550 target molecules have been reported in the literature, demonstrating the broad interest of the research community in this class of reagents.[5]

In addition to providing functional sequences on a dramatically accelerated timescale, *in vitro* selection also benefits from the fact that the building blocks for evolution and the resulting biopolymers can be chemically synthesized, increasing the diversity of structural space that can be explored. Specifically, the sugar- phosphate backbone can be modified to render aptamers resistant to nuclease cleavage, and in some cases, increase affinity for the specified target.[6] Alternatively, the nucleobases can be modified to expand the functional scope of the aptamers, providing similar affinity to antibodies, which benefit from the functional diversity of the naturally occurring amino acids.[7] Early examples of nonnative nucleic acid aptamers were primarily generated via postselection modification. In this process, the aptamer was first selected using native DNA or RNA, then modifications were systematically introduced and their impact on affinity, stability, and function quantified. A classic example of this approach is the drug Macugen (pegaptanib), which is FDA approved for the treatment of age-related macular degeneration, and represents the first aptamer-based therapeutic to be successfully commercialized.[8]

The development of Macugen began with the use of SELEX to generate an RNA aptamer for the vascular endothelial, which regulates blood vessel growth in the eye, and thus is a primary target for the treatment of age-related macular dege-

neration.[9] After identification and truncation of the initial aptamer sequence, a number of modified sequences having 2'-O-Me, 2'-NH_2, and 2'-F modifications were generated and tested (Figure 1.1a).[10] Modifications to the termini of the aptamer were also investigated as a method for increasing stability and circulation lifetime.[6] As a result of these modifications, the affinity, IC_{50}, and biological half-life were each improved by orders of magnitude, and the optimized aptamer structure was carried forward into testing and approved as Macugen (Figure. 1.1b).

While the story of Macugen highlights the potential gains that can be realized through postselection modification of aptamers, an even more powerful approach is the direct selection of aptamers from libraries bearing modified backbones or nucleobases. In this approach, a large number of modified sequences

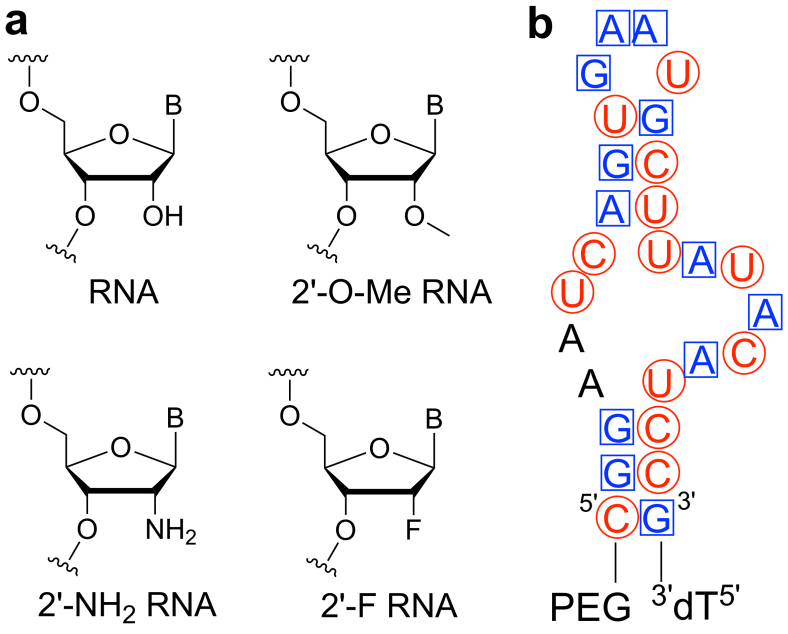


Figure 1.1. Backbone modifications used in postselection testing of anti-$VEGF_{165}$ aptamer. (a) Chemical structure of RNA and modified backbones. (b) Sequence of Macugen. Red Letters with circles are 2'-F pyrimidines and blue letters with squares are 2'-O-Me purines.

can be investigated in parallel, and the most effective modifications are directly enriched in the selection process, in contrast to the trial-and-error process of post-selection modification. As a result, aptamers having unique and improved function can be generated in shorter amounts of time. For several years, a key challenge preventing realization of this goal was the incompatibility of modified building blocks with the polymerases used to transcribe, reverse transcribe, or amplify nucleic acids during SELEX. However, over the last decade, exciting progress has been made in two key areas: (1) the development of engineered polymerases capable of utilizing nonnatural building blocks as substrates, and (2) the invention of clever strategies for the nonenzymatic templated synthesis of modified nucleic acids, which eliminates the need for polymerases to accept the non-native nucleic acids as substrates. This review will highlight many of these recent developments, and explore their application for in vitro selection of modified nucleic acid aptamers. The review is organized into two main sections, which separately address the selection of aptamers having modified backbones and modified nucleobases. These two different types of modifications create different challenges for the implementation of SELEX, and thus have required slightly different forms of technology development. Excitingly, the progress that has been made in each of these areas has provided a key foundation for the development of *in vitro* selection methods that combine both backbone and nucleobase modifications, which will enable the future exploration of even greater swaths of structural and functional space.

SELEX using backbone-modified nucleic acids

Aptamers have found use in a wide variety of applications including biosensing, affinity purification, therapeutics, and biological targeting. However, the low biological stability of DNA and RNA can pose limitations for the use of aptamers in these applications.[11] Because DNA and RNA are naturally occurring biomolecules, a number of nuclease enzymes have co-evolved to catalyze their degradation, which becomes especially problematic for the use of aptamers *in vivo*. Modification of the nucleic acid backbone can greatly reduce nuclease susceptibility, and this strategy has been successfully implemented to enable the use of nucleic acids in antisense therapy and targeted drug delivery. These backbone-modified nucleic acids are generally referred to as xeno nucleic acids (XNAs), where the ‘‘X” can be replaced by a specific letter (or letters) to indicate a particular modification (Figure 1.2). Also included in Figure 1.2 is L-DNA, which is the enantiomer of native D-DNA. L-DNA differs from XNAs in that it does not have a chemically modified backbone. But, due to its chiral modification, we still choose to classify L-DNA as a modified nucleic acid. For antisense applications, XNA oligonucleotides having the desired activity can be generated with relative ease, as their activity relies upon Watson–Crick base pairing, and thus function is only minimally perturbed by the introduction of modifications.[12] However, as outlined above, postsynthetic modification of aptamers is significantly more challenging and time-intensive. A preferable approach is the direct selection of aptamers from XNA libraries, but this introduces new challenges, as the backbone modifications that prevent recognition by nucleases also make recognition by polymerases less efficient. Excitingly,

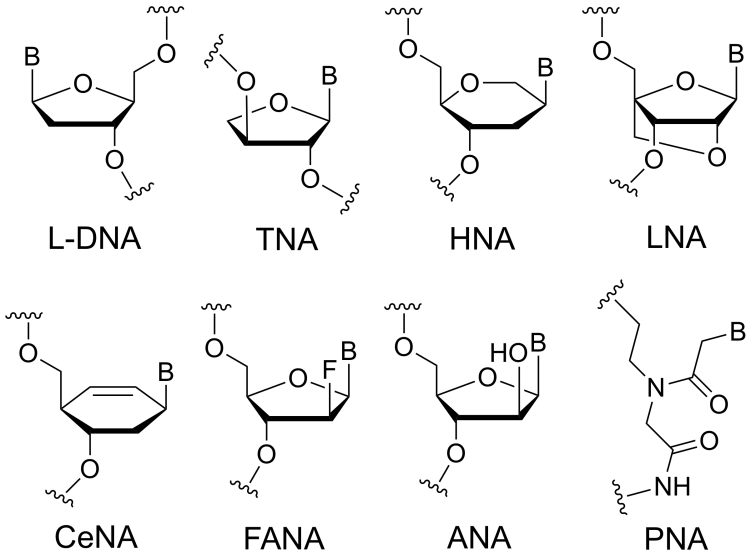


Figure 1.2. Chemical structure of XNA backbones that have been explored for *in vitro* selection experiments.

great strides have been made recently in the field of polymerase engineering, which has enabled the development of polymerases that can both encode (transcribe) DNA into XNA, and decode (reverse transcribe) XNA back into DNA, enabling selection of aptamers from XNA libraries. This section will discuss the implementation of these engineered polymerases for the *in vitro* selection of XNA aptamers, and will highlight recent exciting developments in the area of nonenzymatic templated synthesis, which holds promise to eliminate the reliance on enzymes for encoding and decoding of sequences during SELEX.

Mirror-image selection of L-DNA aptamers

While postselection incorporation of chemically modified nucleotides is challenging, one unique class of XNA for which postselection modification is relatively trivial is L-DNA. Because it is the mirror image of naturally occurring D-DNA, these aptamers can be initially selected using standard DNA SELEX against the

enantiomer of the desired target. Then, the resulting sequences can be chemically synthesized from enantiomeric L-DNA, and due to the principal of reciprocal chiral substrate selectivity, they will bind to the desired target with identical affinity and specificity as the initially selected aptamer. The Joyce Lab has been a key pioneer in the selection of L-DNA aptamers, and recently reported an L-DNA sequence that utilizes non-Watson-Crick interactions to bind to the HIV-1 trans-activation responsive (TAR) RNA.[13] The Klussman Lab have also been key pioneers of mirror-image selection, and have recently implemented this process to generate a D-RNA sequence capable of binding to the enantiomer of complement factor C5a, a pro-inflammatory mediator involved in the pathogenesis of many inflammatory diseases. Interestingly, after minimization of the D-RNA sequence, the aptamer was then synthesized using a mixture of L-DNA and L-RNA monomers and was found to bind the target with high efficiency.[14] While this approach significantly benefits from the fact that selections are carried out using native DNA or RNA, a minor drawback is the need to generate the enantiomer of the target for use in the selection experiments. In the case of small molecules and peptides, this is typically not problematic. However, generating an enantiomeric target can become challenging when undertaking selections for large proteins that are not amenable to chemical synthesis.

Selection of XNA aptamers using engineered polymerases

As described above, a key challenge for the selection of XNA aptamers is that XNA nucleotides can be poor substrates for naturally occurring polymerases. However, a number of groups have devoted significant effort to screening comer-

cially available polymerases, and at times engineering new polymerases, in order to efficiently encode and decode XNA sequences during the SELEX process.

Selection of XNA aptamers using polymerase-mediated encoding

One particularly promising XNA for use in biological applications is threose nucleic acid (TNA), which has been shown by Chaput and coworkers to possess excellent nuclease resistance.[15] The commercially available Therminator DNA polymerase (NEB) was found to be capable of encoding (transcribing) DNA libraries into TNA, at least in the context of TNA sequences having T, G, and diaminopurine (D, used in place of adenine) nucleotides.[16] Thus, the Chaput Lab designed a ‘‘DNA display” selection strategy that obviates the need for a decoding (reverse transcription) step, enabling the implementation of a full SELEX cycle (Figure 1.3a). To achieve this, a DNA hairpin was designed that contains both the library template and primer, and thus can be used to synthesize a TNA library via primer extension using Therminator. A DNA primer was then annealed in the loop region of the hairpin, and extension of this primer displaced the TNA, ‘‘displaying” it on a DNA duplex where it could then be screened for binding to the target molecule. Using this strategy, the genetic information of each TNA strand is maintained in the covalently attached DNA template, which can be directly PCR amplified after selection, circumventing the need to decode the selected XNA sequences. Using this strategy coupled with capillary electrophoresis to separate target-bound and free sequences, the Chaput Lab successfully selected a thrombin-binding TNA aptamer having a K_D value of 200 nM. They have since shown that incorporation of 7-

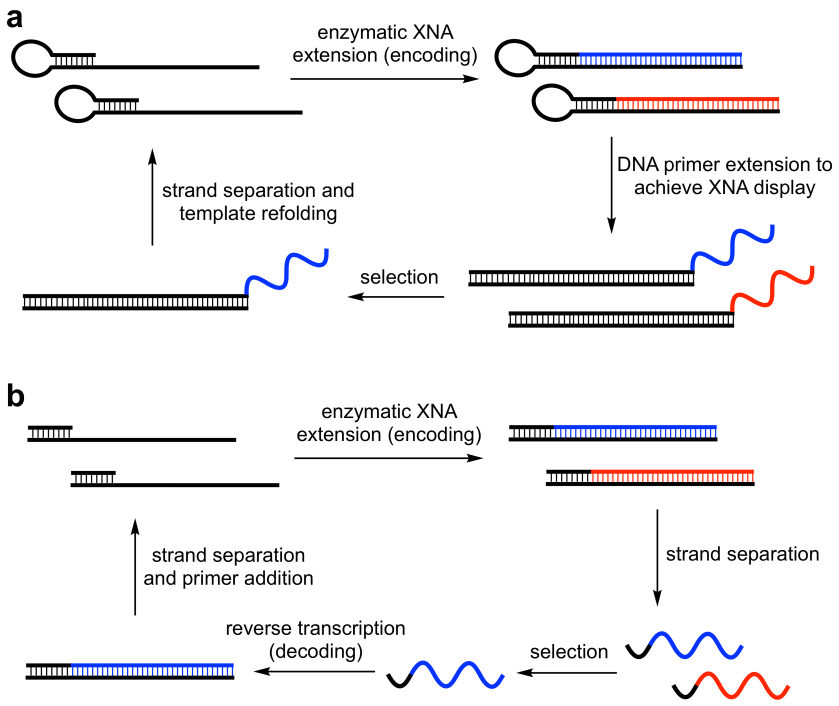


Figure 1.3. Selection of XNA aptamers using engineered polymerases. (a) "DNA display" strategy only requires a polymerase for encoding the DNA library into XNA. (b) Two-enzyme selection system enables both encoding and decoding of XNA, so selection step can be carried out using single-stranded XNA library.

deazaguanine in the DNA template allows for the transcription of four-letter TNA libraries, which will enable the selection of TNA aptamers that take advantage of the full scope of Watson-Crick base pairing.[18] Additionally, the Chaput Lab have recently discovered that commercially available SuperScript II polymerase is capable of reverse transcribing TNA into DNA, which will enable selections to be carried out in a manner analogous to RNA SELEX.[18]

Selection of XNA aptamers using polymerase-mediated encoding and decoding

A key advance in the selection of XNA aptamers was realized with the development of compartmentalized self-tagging (CST) by Holliger and coworkers,

which enables the rapid evolution of polymerases for XNA substrates.[19] Using this method, a series of enzymes was evolve that are collectively capable of both transcription and reverse transcription of 1,5-anhydrohexitol nucleic acid (HNA), cyclohexitol nucleic acid (CeNA), 2'-O,4'-C-methylene-b-D-ribonucleic acid (locked nucleic acid, LNA), aribinonucleic acid (ANA), 2'-fluoro-arabino-nucleic acid (FANA), and TNA (Figure 1.2). Because these enzymes are capable of both encoding and decoding XNA libraries, SELEX can be can carried out as shown in Figure 1.3b. Much like RNA SELEX, XNA selections using encoding and decoding polymerases begin with generation of an XNA library via transcription from a DNA library. Once purified away from the DNA template, a selection step can be employed to enrich the XNA library for sequences having the desired function, and these sequences can be reverse-transcribed back into DNA for amplification or sequencing.

Using the encode-select-decode strategy coupled with a standard bead-based enrichment method, Holliger and coworkers used their engineered polymerases to generate HNA aptamers for both HIV TAR RNA and Hen Egg Lysozyme (HEL).[20] Both aptamers bind to their intended targets with good selectivity and K_D values in the nanomolar range. Similarly, DeStephano and coworkers used CST-derived polymerases to select for a FANA aptamer that binds to HIV-1 reverse transcriptase (RT) with picomolar affinity.[21] This affinity was higher than that achieved by a DNA aptamer for the same target, demonstrating that XNA backbones can also provide increased affinity for the target, potentially through energetically favorable hydrophobic interactions. Interestingly, the enzyme used to transcribe DNA into FANA not only functions as an isothermal primer extension cata-

lyst, but can also be employed in thermal cycling protocols similar to those used in PCR, suggesting that direct amplification of XNA may be possible in the future.

Taking a slightly different approach, Kuwahara and coworkers utilized an engineered polymerase for LNA, but the primer used for the encoding step was comprised of DNA having LNA inserted at every third nucleotide. Similarly, the libraries that were generated were primarily composed of DNA, but had all thymines replaced with locked thymine analogues. Using this library of chimeric DNA:LNA sequences, four thrombin-binding aptamers were selected via capillary electrophoresis, and were found to have K_D values of 18, 20, 23, and 27 nM.[22,23]

Nonenzymatic templated synthesis of genetically encoded libraries

The selection methods described up to this point rely on engineered polymerases to encode and decode XNA sequence information. Liu and coworkers envisioned that the information stored in a DNA template could potentially be encoded into XNA, or even nonnucleic acid polymers without the use of enzymes, which would dramatically expand the structural diversity that could be explored in selection experiments. This principle was initially demonstrated for peptide nucleic acid (PNA), which is capable of Watson-Crick base pairing with DNA and RNA[24] but does not possess a phosphodiester backbone, and thus cannot be generated using even highly engineered polymerases. Similar to the encode-only ‘‘display” strategy, a self-priming DNA hairpin was used as a template, and PNA building blocks having sequences complementary to this template were annealed and covalently polymerized using reductive amination. To illustrate the potential utility of

this method for *in vitro* selection, a mock selection was performed in which one PNA building block was synthesized having a biotin, and sequences containing this building block were enriched from a random pool.[25]

Taking one step further, Liu and coworkers envisioned that templated synthesis could be used to generate libraries of sequence-defined polymers that have no inherent affinity for the DNA template. This was accomplished using a design inspired by the ability of tRNAs to mediate the templated synthesis of peptides. As shown in Figure 1.4, PNA building blocks similar to those used in their previous experiments were equipped with a nonnucleic acid polymer segment via two cleavable linkers. In this case, the functional groups required for ligation were placed on the nonnucleic acid polymer, which enabled templated ligation of these segments to create the genetically encoded polymer. Subsequent cleavage of the linkers res-

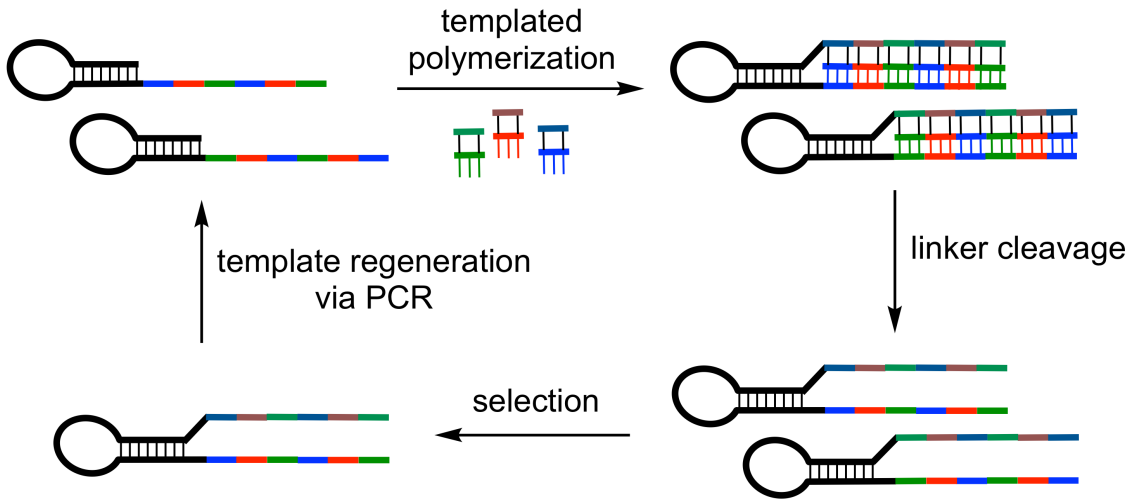


Figure 1.4. Nonenzymatic templated polymerization of nonnucleic acid polymers. Polymer segments are covalently attached to an "adaptor" nucleic acid fragment, which directs templated polymerization, but can be subsequently cleaved to reveal the encoded polymer library.

ulted in removal of the PNA adapter segments, allowing the polymer to be covalently displayed on its encoding DNA sequence.[26] While these templated polymerization strategies have yet to be successfully employed in the selection of aptamers, there is significant promise in their ability to encode genetic information into a structurally diverse range of sequence-defined polymers.

SELEX using modified nucleobases

Proteins have historically dominated the pool of available catalysts and affinity reagents, likely as a result of the diverse array of amino acid side chains that they can employ. In contrast, DNA and RNA are each comprised of only four nucleotide building blocks, and these all possess a relatively similar palette of functional groups. Aiming to generate aptamers that have similar affinity and specificity to protein-based receptors, a number of research groups have explored the *in vitro* selection of aptamers using monomers having unnatural nucleobases. The earliest example of SELEX using modified nucleotides was presented by Toole and coworkers in 1994, where an aptamer for thrombin was generated using a library in which all thymine monomers were replaced by 5-pentynyl-dU.[27] This aptamer was found to have a dissociation constant comparable to that of a previously selected DNA aptamer for thrombin, and showed unique secondary and tertiary structures. While this early example established precedence for the idea of utilizing unnatural nucleobases in SELEX, the true potential of this approach was demonstrated later by SomaLogic, Inc., who carried out selections in which thymine was replaced with a 2'-deoxyuridine (dUTP) having a hydrophobic functional group at the 5' position of the nucleobase.[28] Using libraries containing these monomers,

high-affinity ‘‘SOMAmers” were generated for a number of challenging targets, and were subsequently incorporated into an aptamer-based biomarker detection platform capable of detection limits in the fM range. The exploration of SELEX using unnatural nucleobases has rapidly expanded in the past few years, and these efforts can be separated into two general approaches: (1) development of unnatural base pair systems that are orthogonal to the Watson-Crick base pairs, and (2) addition of peptide-like functional groups to native nucleobases. As is the case for *in vitro* selection using modified backbones, a key challenge to the use of unnatural nucleobases in SELEX is the requirement that the unnatural nucleotides be encoded and decoded during the selection process, and this often requires compatibility with polymerases. In the case of modified nucleobases, and additional challenge is that the selectivity of base pairing with or between modified bases must occur with high fidelity during the encoding and decoding steps. However, a variety of innovative methods have been recently developed to circumvent these challenges and expand the scope of nucleobase modifications that can be employed for the *in vitro* selection of aptamers.

Development of unnatural base pairs

Expansion of the genetic alphabet to include base pairs beyond naturally occurring A:T/U and G:C was first pioneered by Alex Rich in 1962, with his proposal that an orthogonal base pair could be formed between iso-G and iso-C.[29] The key requirements for the development of these new base pairs was that the two modified nucleobases must pair with one another, but not with the canonical nucleobases. Following up on this initial work, Benner and coworkers proposed

three non-canonical donor–acceptor patterns, which could in principle be used in conjunction with the two Watson-Crick base pairs to provide an artificially expanded genetic information system (AEGIS) having five different base pairs.[30] Taking a different approach to the design of unnatural base pairs, the Kool, Romesberg, and Hirao groups have each reported modified nucleotides that are able to form orthogonal base pairs through hydrophobic interactions rather than hydrogen bonding.[31-33] While these base pairs have been utilized for a number of innovative applications over the past several years, only recently have the technological hurdles required for their use in aptamer selections been overcome.

Aptamer selection using base pairs having alternative hydrogen bonding patterns

In the design of the AEGIS base pairs, the Benner Lab hypothesized that placing electron density on the minor groove face of each nucleobase would enable recognition by polymerases, allowing the enzymatic synthesis and amplification of oligonucleotides incorporating the AEGIS base pairs.[34] Among the proposed AEGIS base pairs, this requirement was most efficiently met by pyrimidine donor-donor-acceptor (pyDDA) and purine acceptor-acceptor- donor (puAAD) motifs. From this insight, the Z:P base pair was reported in 2006 (Figure 1.5), and was shown to be highly specific, exhibiting orthogonality to the natural nucleobases and having a favorable effect on overall duplex stability. Moreover, the Z-P base pair was recently confirmed to adopt Watson-Crick geometry, supporting the generation of a six nucleotide genetic alphabet.[35]

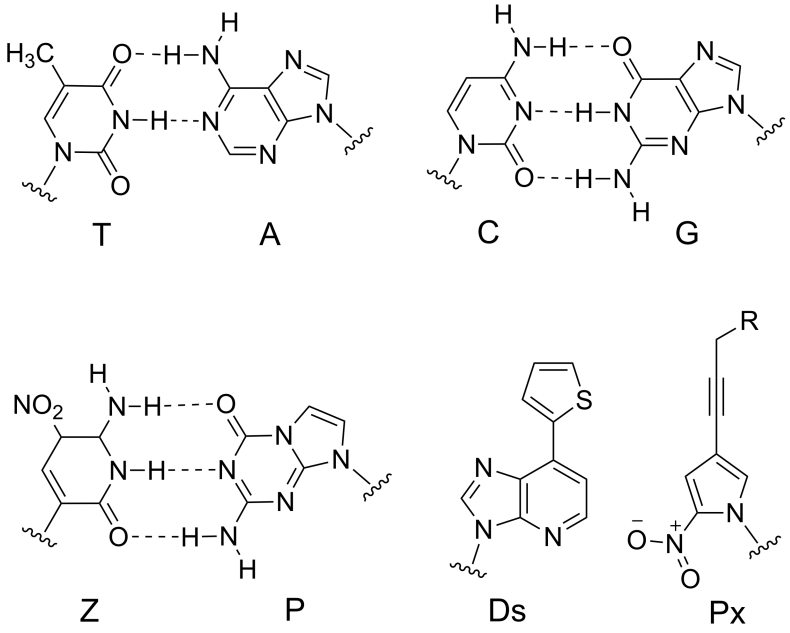


Figure 1.5. Natural and unnatural DNA base pairs used for the *in vitro* selection of aptamers. The Z:P base pair utilizes an orthogonal donor-acceptor hydrogen bonding pattern, whereas the Ds:Px base pair relies on hydrophobic interactions.

Equipped with an unnatural base pair that was capable of enzymatic amplification, Benner and coworkers turned to exploring the use of Z:P in the selection of aptamers. Using a library containing a 1.5:1.2:1.0:1.0:1.0:0.5 ratio of T:G:A:C:Z:P, SELEX experiments were carried out against the MDA-MB-231 breast cancer cell line.[36] Because the AEGIS base pair is recognized by native polymerases, sequences that bound to the target cells were able to be directly PCR amplified using all six nucleotide triphosphates, and the resulting oligonucleotides carried into the subsequent selection round. After 11 rounds of selection, the resulting library was sequenced by dividing it into two fractions, which were independently barcoded. For each fraction, the Z and P nucleotides were converted into natural nucleotides using a different protocol. In the first fraction, the Z and P nucleotides were converted to C and G, respectively. In the second fraction,

the Z nucleotides were converted to a mixture of C/T and the P nucleotides were converted to a mixture of A/G. These two fractions were then combined and subjected to next-generation sequencing, and alignment of the resulting sequences revealed the locations of the unnatural nucleotides in the selected pool. The dominant aptamer was found to have a binding affinity of 30 nM, and replacement of the Z:P base pairs with Watson-Crick base pairs greatly diminished this affinity, highlighting the important role of these modified nucleotides in the target-binding interaction.

Aptamer selection using hydrophobic base pairs

In 2006, Hirao and coworkers introduced the Ds:Pa hydrophobic base pair, and showed that the unnatural nucleotides could serve as substrates for natural polymerases.[33,37] While the Ds-Pa system was successful in both replication and transcription experiments, in part owing to the unique shape complementarity between the hydrophobic bases, undesired Ds-Ds and A-Pa base pairs were observed, leading to reduced fidelity. Subsequently, 2-nitro-4-propynylpyrrole (Px) was synthesized as an optimized pairing partner for Ds, having nitro and propynyl groups to alter electrostatic interactions and increase hydrophobicity, respectively (Figure 1.5).[38] The increased hydrophobicity of Px improved incorporation opposite Ds, suppressing Ds-Ds mispairing, and the electrostatic repulsion due to the nitro group discouraged A-Px mispairing.

Using the Ds:Px unnatural base pair system, aptamer selection experiments were carried out against vascular endothelial cell growth factor($VEGF_{165}$) and interferon-γ (IFN-γ).[7] The hydrophobicity of Ds was hypothesized to improve

interactions with hydrophobic regions of the protein targets, and to ensure access of the unnatural nucleotides to the protein target, Px was intentionally left out of the library, as this prevented the enrichment of meaningful sequences in which the Ds was sequestered into base pairs. Unlike the Benner Z:P unnatural base pair, a convenient method for decoding the Ds:Px pairs after selection was not available. Thus, libraries were synthesized having 1–3 Ds nucleotides at fixed locations in the sequence, and these libraries were barcoded using a unique 2–3 nucleotide tag. The libraries were incubated with each protein target, and bound sequences were amplified in the presence of Ds and Px deoxyribonucleotide triphosphates (dDsTP and dPxTP) to prepare material for the subsequent selection round. After completion of the selection rounds, the libraries were amplified in the absence of dDsTP and dPxTP, which caused reversion of the unnatural base pairs to A:T base pairs, which could then be sequenced. As a control, selections for the two protein targets were also carried out using libraries having entirely native DNA nucleotides (contVG and contIF). The highest affinity aptamers for $VEGF_{165}$ and IFN-γ from the Ds –modified library had K_D values of 1.69 pM and 0.124 pM respectively. These affinities were dramatically higher than those obtained for control aptamers contVG (370 pM) and contIF (3 nM) for binding to $VEGF_{165}$ and IFN-γ respectively. Compared to the AEGIS approach, the Ds:Px base pair provides hydrophobicity that is generally lacking in the natural nucleobases, and the power of this is reflected in the high affinities obtained by the Hirao Lab. However, the current minor limitation of the Ds:Px base pair relative to AEGIS is that the sites of the Ds modifications must be fixed in the sequence, compared to the Z:P nucleotides, which

can be distributed randomly throughout the library and decoded post-selection using sequencing.

Modified nucleotides having peptide-like functional groups

While aptamers have a number of benefits relative to antibodies, including increased thermal stability and their ability to be chemically synthesized, antibodies benefit significantly from the chemical diversity available in the repertoire of naturally occurring amino acid side chains. In the early 2000s, it was envisioned that grafting the functional groups present in amino acid side chains onto DNA nucleobases could lead to the evolution of aptamers or DNA enzymes (DNAzymes) having new or improved function. In 2001, Perrin and coworkers utilized a library containing nucleobases modified with imidazole and amine side chains to generate a DNAzyme that functioned as an RNaseA mimic, and they later expanded these selections to include nucleobases having guanidine functional groups.[39,40] Selections using a library containing a combination of imidazole, amine, and guanidine modifications resulted in the generation of a DNAzyme capable of cleaving RNA at rates much higher than previously reported for unmodified DNAzymes. Moreover, this modified DNAzyme maintained function in the absence of divalent metal cations, which are typically required for catalytic nucleic acids. These experiments demonstrated that the addition of protein-like functional groups could greatly enhance the potential catalytic abilities of nucleic acids. In parallel with the efforts of the Perrin Lab, Famulok and coworkers demonstrated that DNA sequences could be enzymatically polymerized in which all four nucleotides were substituted with an unnatural functional group, paving the way for in vitro selection using libraries

having a high density of amino acid-like functional groups.[41] While these specific libraries have yet to be utilized for aptamer selection experiments, the knowledge of which positions on the nucleobases can be modified without disrupting enzymatic polymerization proved to be critical for the selection of nucleobase-modified aptamers such as SOMAmers. Additionally, in recent years, innovative methods have been developed to further expand the diversity of functional groups that can be appended to nucleobases for *in vitro* selection.[42,43]

Nucleic acids as scaffolds for peptide display

While polymerase-mediated incorporation of modified nucleotides is a powerful approach for the selection of functionalized aptamers, Liu and coworkers recognized that a greater diversity of side chains could be accessed via enzymatic templated synthesis of modified oligonucleotides. Specifically, they constructed trinucleotide building blocks in which one nucleobase was functionalized with a side chain, and demonstrated that these could be efficiently polymerized on a DNA template using T4 DNA ligase.[44] To demonstrate the potential of this technology for *in vitro* selection, a mock selection was carried out in which one building block was functionalized with a known ligand for carbonic anhydrase II.

Building upon this work, Hili and coworkers have investigated the use of building blocks having polypeptide side chains, as this would provide nucleic acid libraries that mimic the multivalency which characterizes many antibody-antigen interactions. In these experiments, pentanucleotide building blocks were chosen for the T4 DNA ligase-mediated polymerization, with the hypothesis that this length would be amenable to polymerization with longer peptide side chains.[45] Using a

self-priming DNA template, pentanucleotide building blocks functionalized with either a cationic or anionic octapeptide were efficiently polymerized, suggesting that this scaffold system could display a wide variety of peptide sequences. A mock *in vitro* selection experiment was performed using seven pentanucleotide building blocks that had been shown to undergo efficient ligation and an eighth building block having a hexahistidine tag. The resulting library was selected for binding to Co^{2+} magnetic beads, and an XbaI cleavage site was included in the template encoding the hexahistidine monomer in order to quantify enrichment (Figure 1.6). After completing one selection cycle, the hexahistidine-containing sequence was enriched nearly 200-fold, suggesting the potential of the T4 DNA-ligase-based polymerization to function for the *in vitro* selection of aptamers or DNAzymes.

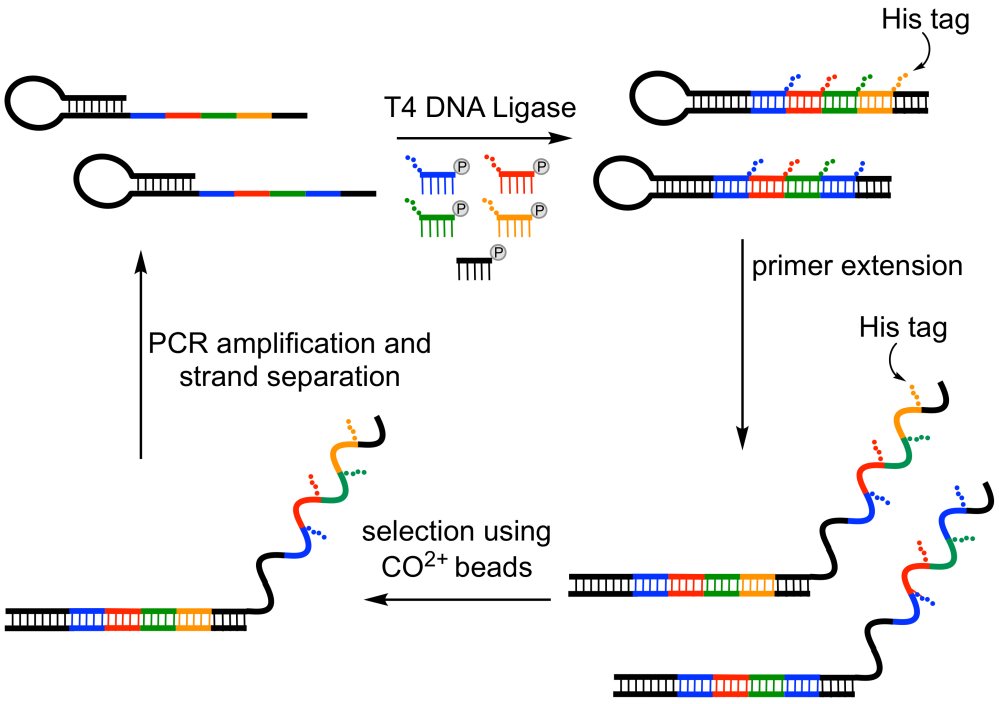


Figure 1.6. Selection of DNA sequences having displayed peptides. Peptide-bearing fragments undergo templated ligation, and selected sequences can be PCR amplified and purified to regenerate single-stranded self-priming library.

Click-SELEX as a versatile method for incorporating functional groups

The Mayer Group has recently presented a new approach termed ‘‘click-SELEX” that enables the construction of nucleobase-modified aptamers utilizing the popular copper (I) azide-alkyne cycloaddition reaction.[46] In this method, all thymidine nucleotides within the library are replaced with C5-ethynyl-2’-deoxyuridine (EdU), which introduces multiple alkyne functional groups into the sequences. The library is then further functionalized through reaction with an azide, which in the case of this initial demonstration was 3-(2-azidoethyl)indole. The resulting library was then subjected to *in vitro* selection to enrich aptamers having affinity for the protein cycle 3 GFP (C3-GFP). Sequences surviving each round of selection were amplified using PCR, with EdU triphosphate used in place of dTTP. This provided the enriched library having alkyne functional groups, which could again be reacted with the azide-containing molecule and utilized in the subsequent selection round (Figure 1.7). Notably, the bulk of the nucleobase modification is introduced after DNA amplification, which allows access to larger functional groups while avoiding the enzymatic incompatibility that often accompanies such modifications. After 15 rounds of selection, the most abundant aptamer sequence was analyzed and found to have a dissociation constant of 18.4 nM. This aptamer also proved to have excellent specificity, displaying no observable binding to off-target proteins including streptavidin, ERK2, and the closely related mE-GFP. A series of azides were utilized to assess the importance of the specific indole moiety used for the selection process. Aptamer variants functionalized with the alternative azides were not capa-

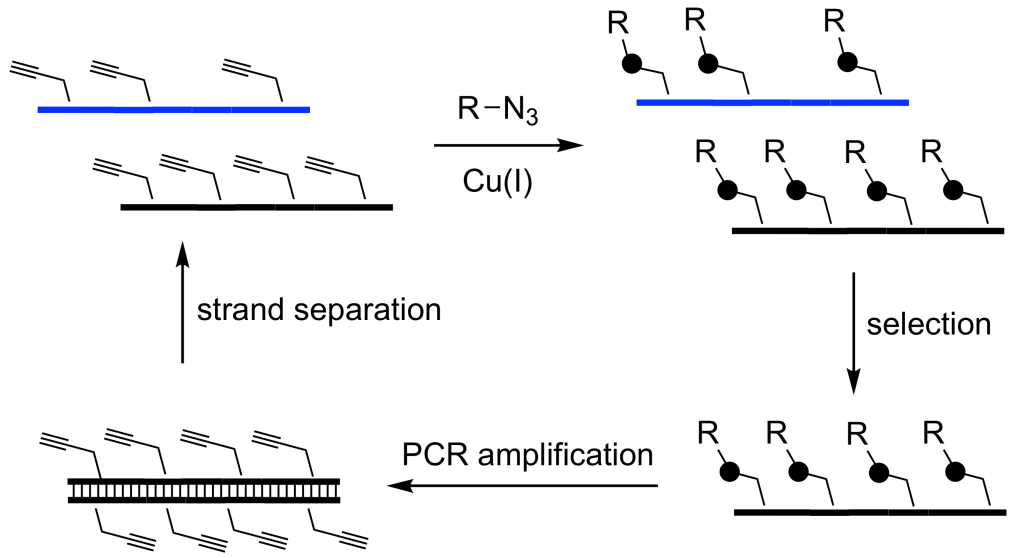


Figure 1.7. Selection of nucleobase-modified aptamers via click-SELEX. Post-synthetic modification of the DNA library enables the incorporation of large functional groups. These modifications are reverted to alkynes for PCR amplification. This enables the use of modifications that are not necessarily compatible with existing polymerases.

ble of binding to the target protein even when the azide was structurally similar to the indole employed in the selections. This highlights the important role of the appended functional groups in target binding. The versatility with which side chains having a wide variety of sizes and chemical properties can be easily introduced using this method is anticipated to greatly accelerate the discovery of nucleic acid catalysts and affinity reagents having novel functions.

Aptamer-based small-molecule biosensors

This increased availability of aptamers to a variety of targets has empowered research effort investigating their potential as alternatives to antibodies. Aptamers provide many advantages such as increased thermal stability, time and cost efficient production, little batch-to-batch variation, and ease of modification to

suit a variety of applications.[47] Because of these benefits, aptamers have been integrated into various applications such as environmental screening, therapeutics, drug delivery, and sensing.[47] In particular, there has been growing interest in the utilization of aptamers as recognition elements in novel biosensor platforms.[48] Biosensors utilize a biological recognition event that can be translated into a measurable signal (i.e., optical, electronic, electrochemical, fluorescent) using a transducer element.[49] Fluorescence-based platforms have become widely popular, due to the wide variety of fluorophores available, ease of synthesis, and sensitivity.[50] Aptamers have been developed against molecules ranging from metal ions to whole cells,[51] and thus together aptamer-based fluorescence assays hold great the potential to detect even the most difficult targets with high sensitivity.

Small-molecules (< 900 g/mol) comprise varied classes of molecules including toxins, biomarkers, drugs, and metabolites.[52] Consequently, biosensors which respond to small molecules have many applications including triggered therapeutic response, environmental toxin detection, RNA imaging, medical diagnostics, and can provide insight into cellular pathways.[53] There is great interest in utilizing aptamers as recognition elements for small-molecule detection, however, there remain challenges in generating aptamers to small-molecule targets.[54] Less than 25% of aptamers selected to date are to small-molecules likely due to the lack of functional groups present in comparison to protein targets.[52] Therefore, small-molecules give lower probability of the pi-pi stacking, hydrogen bonding, and electrostatic interactions that govern aptamer-ligand binding.[55] Additionally, there are technical challenges within the selection and characterization of small-molecule

aptamers.[54] The small size of the target does not lead to a considerable mass change upon target binding, which complicates the partitioning of bound and unbound sequences. This difficulty in distinguishing between the unbound and bound complex not only affects enrichment of high affinity aptamers but also postselection characterization. These obstacles are evident by the observed trend showing that ligand molecular weight is proportional to aptamer affinity (i.e., higher affinity for larger targets). Fortunately, *in vitro* evolution provides a customizable platform to strategically alter traditional selection methods in order to generate high affinity aptamers with enhanced properties.

Favorable architectures for small-molecule detection

While traditional selection methods are well-established for the generation of aptamers, modifications are needed to function as signaling probes for biosensing. Various aptamer sensor architectures have been developed to facilitate their integration as signaling molecules for fluorescence detection. The earliest example of this is molecular beacons, in which an aptamer is functionalized on both termini with either a fluorophore and quencher or a fluorescent donor-acceptor pair.[56] Depending on which moieties are used, fluorescence signal can be turned on or off in the presence of target. While this method is simplistic, it requires a very specific secondary structure that is also sensitive to changes in environment which will alter spacing of the fluorophore and quencher.[48]

Split-aptamers are another type of sensor architecture uniquely suited for small-molecule detection. Split-aptamers consist of two strands that selectively as-

semble in the presence of their small-molecule target.[57] Our lab has recently described Split Aptamer Proximity Ligation (StAPL) showing that the termini of the split-aptamer fragments can be functionalized with reaction groups to promote DNA ligation (Figure 1.8a).[58] Further research in our lab expanded this StAPL technology to enable successful implementation of a split-aptamer–based sandwich assay for cocaine detection (Figure 1.8b). Using the known split-aptamer for cocaine, the capture strand was functionalized with an azide and the detection strand was modified with a cyclooctyne and biotin on opposite ends.[59] The presence of the cocaine drives assembly of the split-aptamer which will place the reactive groups in close proximity, promoting cycloaddition and resulting in covalent ligation of the fragments. A streptavidin-reporter enzyme conjugate can then bind the biotin on the detection strand resulting in measurable colorimetric signal corresponding to analyte concentration. Despite this promise for biosensing assays, few aptamers have been successfully engineered into split-aptamers, likely due to the difficulty in dividing most aptamer structures without disrupting the binding site.

Aptamers are well-suited as recognition elements because the conformational change they undergo upon target binding can be exploited to generate signal output. Recent biosensor development has focused on the exploitation of this target-induced conformational changes in the form of structure-switching (SS) aptamer sensors. This architecture is defined as a DNA/DNA duplex consisting of a fluorescently labeled aptamer and quencher modified, short complementary strand (CS) (Figure 1.9a). In the absence of target the duplex remains hybridized, however, in the presence of target the short CS will be displaced and fluorescence

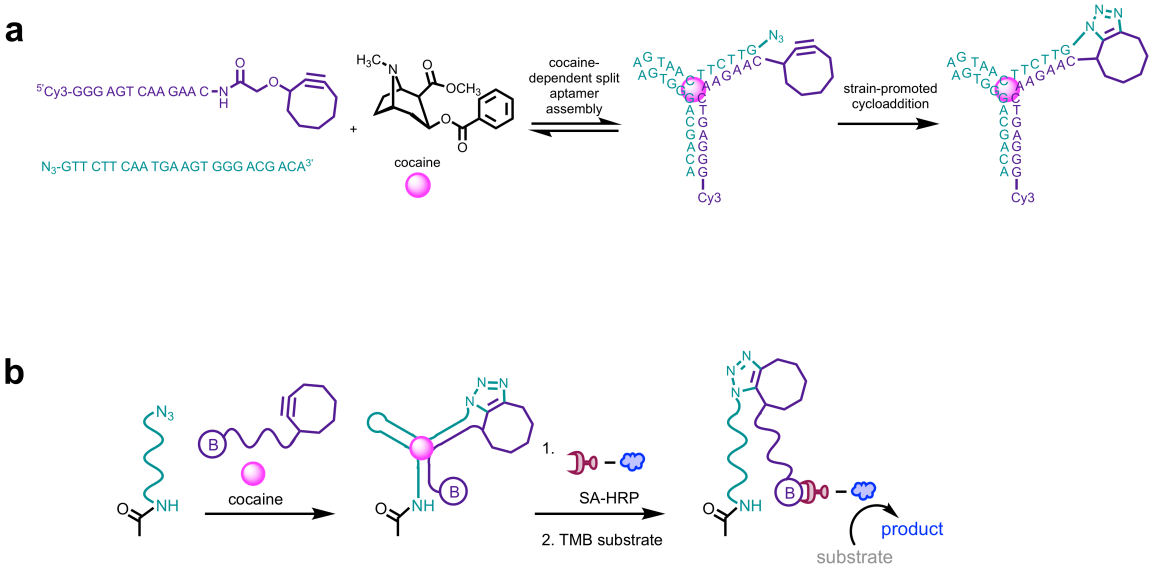


Figure 1.8. Split aptamer technology. (a) Split aptamer ligation proximity ligation (StAPL) (b) Enzyme-linked small-molecule detection using StAPL.

signal will be recovered. This architecture was first realized by Li and coworkers who created structure-switching sensors from existing aptamers for ATP and thrombin.[60] However, because postselection engineering of aptamers into SS sensors without insight into three-dimensional structure has proven difficult, there have been proposed methods for the direct selection of this architecture. This was first exhibited by Nutiu and Li using a DNA library comprised of fixed motif flanked by two short random regions.[61] As shown in Figure 1.9b, this fixed sequence at the center was designed to be complementary to the short CS which was immobilized on magnetic beads. After hybridization of the library to the CS-functionalized beads, sequences that successfully undergo a target-dependent structural change will become dehybridized and released into solution and enriched. The resultant aptamers to ATP and GTP could be directly converted into signaling aptamers, however, they produced only modest fluorescence enhancement in the presence of

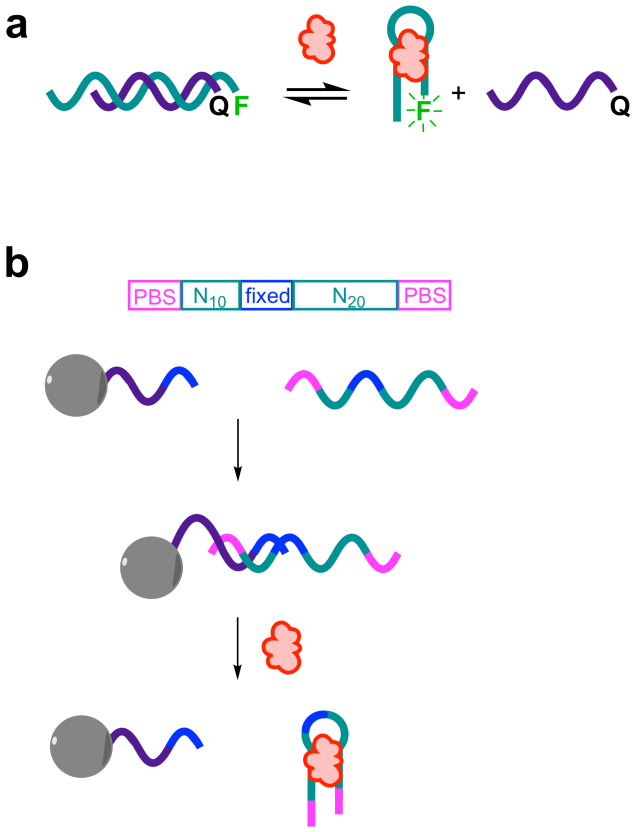


Figure 1.9 SS aptamer-based biosensor design and selection**.** (a) SS-biosensor for small-molecule detection (b) Typical bead-based selection for generating SS-biosensors as reported by Li and coworkers.

target.[61] This is likely due to the fact that small-molecule targets are challenging for the direct selection of SS biosensors as they lead to much smaller changes in aptamer structure in comparison to their protein counterparts. While this elegant approach provided a great advance in the direct selection of privileged sensor formats, it highlights the limitations of bead-based selection methods. The rate of enrichment during SELEX is dictated by the resolution of separating binding from nonbinding sequences. Because affinity based methods (i.e., magnetic beads) result in the nonspecific retention of sequences, partitioning is low which hinders enrichment.[62] The inability of traditional selection methods to guarantee the gener-

ation of aptamers possessing the desired SS architecture that produce sufficient fluorescence signal poses great technical challenges. Consequently, there remain great opportunities in the innovation of selection methods that can produce sensors that are sensitive to small-molecule binding without laborious post-selection optimization.

RNA aptamers to small-molecule fluorophores

RNA is an incredibly fascinating and complex biopolymer that performs various cellular duties. The discovery of novel functions such as riboswitches, RNA-protein complexes, and noncoding RNAs[63] has shattered the preconceived idea that it operates solely as an intermediate in the transfer of genetic information. The ability to image RNA in cells has the potential to provide insight into their functions and elucidate the mechanisms of different pathologies. Nucleic acids have found utility as probes for the imaging of native mRNAs due to the high fidelity of Watson-Crick base pairing. In the 1980s, fluorescence *in situ* hybridization (FISH) was introduced, a method in which a fluorescently labeled nucleic acid complement is delivered into fixed (dead) cells to hybridize to the target RNA.[64] While this method is able to achieve site-specific labeling, the use of fixed cells provides a stationary image of RNA. Therefore, methods to image RNA in live cells are in high demand as they allow the ability to resolve spatial and temporal information regarding localization patterns. The introduction of hybridization probes (FRET, molecular beacon, quenched autoligation) benefited from reduced background fluorescence, allowing for visualization of endogenous transcripts in living cells (Figure 1.10 a-c).[65] Despite their simplicity, hybridization probes have not been widely implemented

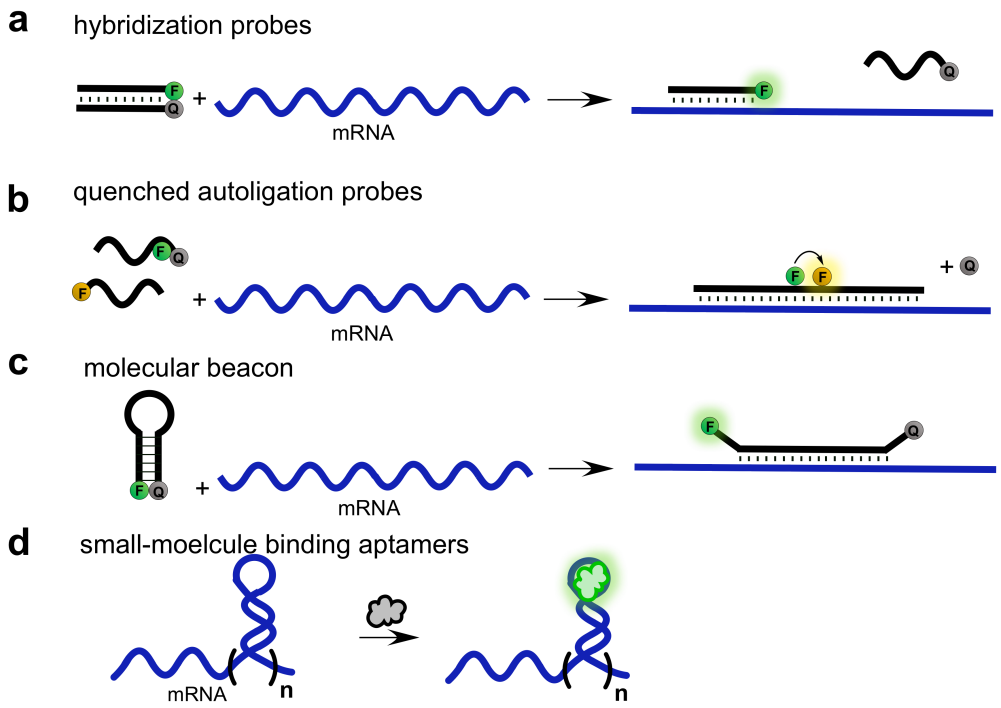


Figure 1.10. Methods for mRNA imaging using nucleic acid probes: (a) hybridization probes, (b) quenched autoligation probes, (c) molecular beacon, and (d) small-molecule binding RNA aptamers.

because delivery, stability, and subcellular distribution remain difficult obstacles to overcome.

In the search for novel RNA visualization techniques, researchers have regarded the molecular recognition capabilities of aptamers as a promising alternative. Several RNA aptamers have been reported that bind to small-molecule fluorophores which benefit from high quantum yield and small size that is unlikely to perturb RNA behavior. Tsien and coworkers were the first to report an RNA aptamer that could bind to a fluorogenic small-molecule, that is only fluorescent in the bound complex, in order to image β-lactamase expression in biological samples.[66] Aptamer-based strategies were adapted to live-cell imaging with the development of an RNA aptamer that bound a mimic of the GFP fluorophore, 3,5-

difluoro-4-hydroxybenzylidine imidazolinone (DFHBI) (Figure 1.10d).[1] Importantly, the use of an RNA aptamer allows the tag to be genetically encoded, obviating the need to exogenously introduce the probe. This aptamer called “spinach” due to the bright green fluorescence of the DFHBI upon binding, was successfully used to visualize ribosomal RNA in live mammalian cells.[67] The spinach aptamer provides compelling evidence that nucleic acids are of great utility as labeling tags for RNA imaging, however, there are still drawbacks such as imaging sensitivity due to the noncovalent nature of the binding interaction.

Dissertation overview

Functional nucleic acids hold great promise for a variety of biosensing applications, and *in vitro* selection provides a powerful method to optimize nucleic acids for this purpose. This dissertation will explore novel SELEX methods that enable the generation of aptamers with enhanced functionality for small-molecule detection and imaging. We anticipate the aptamers generated by the techniques discussed herein will be optimized for the development of fluorescence assays. Collectively, these results suggest that the ability to push the boundaries of current selection technologies will be advantageous for the advancement of aptamer-based biosensors.

The discovery of naturally occurring ribozymes has shown that nucleic acids can function in more than base-pair and ligand-binding interactions. Their ability to catalyze a variety of chemical reactions has widened the scope of potential applications for catalytic RNA. In Chapter 2, we propose the use of ribozymes as novel tags for mRNA labeling and imaging. Using a novel IP-SELEX method, we were

able to generate self-alkylating ribozymes that are capable of catalyzing the covalent attachment of a small-molecule fluorophore to itself. The use of a fluorophore binding antibody in the selection step provides a convenient method to recover sequences that successfully conjugate the label. We anticipate the covalent nature of this method will be advantageous over previous small-molecule binding aptamers to allow for reduced background and prevent dissociation of the label and facilitate sensitive imaging of target RNA. We also explored the specificity and substrate scope of these ribozymes. Furthermore, we established that the ribozyme can function when fused to an mRNA sequence, validating its promise for labeling in live cells.

The evolution of enzymes that can transcribe and reverse transcribe XNA, has provided the ability to generate aptamers with novel attributes such as high nuclease stability. While such polymers have been evolved having affinity to protein targets, and even catalytic function, there exists no small-molecule binding XNA aptamer. We have developed a novel selection method to enable the creation of the first XNA aptamer capable of small-molecule recognition. We report the first aptamer selection using a fully XNA library, which obviates the need for post selection alteration which may affect the integrity of the binding site. Using this method, we were able to generate TNA aptamers with high affinity to ochratoxin A, a food-borne mycotoxin. Further, we show that our best sequences can be structurally minimized to enable chemical synthesis of our aptamers. We exhibit that these TNA aptamers are stable in the presence of nucleases and retain binding in conditions where DNA aptamers for the same target are completely degraded.

We recognize the promise of structure-switching biosensors in the detection of small molecules, but acknowledge the limitations of current bead-based technologies for acquiring the correct architecture. In Chapter 4, we introduce a novel method for the selection of SS biosensors to overcome the limitations of current SELEX capabilities. We propose a homogeneous PCR-based technique in which the library is designed to form a restriction enzyme cut site when hybridize to the CS to form the biosensor complex. Functional SS biosensors will undergo a conformation change upon target binding and displace the CS and nonfunctional sequences will remain hybridized. Upon introduction of the restriction enzyme, the non-functional sequences will be digested and only the functional SS-biosensors will be amplified to form full-length product. The use of the restriction enzyme provides a clever method of selectively amplifying functional sequences, as the non-functional sequences will have their primer binding site removed during digest ion and will not form full-length product. Additionally, because the amplification step also serves simultaneously as the method of enriching functional sequences, this proposed selection is completely homogenous. This obviates the need for a solid support and anticipate this will overcomes the nonspecific elution associated with bead selections that negatively affects enrichment. We apply this method for the development of SS biosensors to coenzyme A to be used in the development of FP assays for histone acetyltransferases.

References

(1) Ellington, A. D.; Szostak, J. W. In Vitro Selection of RNA Molecules That Bind Specific Ligands. *Nature* **1990,** *346*, 818-822.

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

(3) Tuerk, C.; Gold, L. Systematic Evolution of Ligands By Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase. *Science* **1990,** *249*, 505-510.

(4) Stoltenburg, R.; Reinemann, C.; Strehlitz, B. SELEX—A (R)evolutionary Method to Generate High-Affinity Nucleic Acid Ligands. *Biomol. Eng.* **2007,** *24*, 381-403.

(5) Cruz-Toledo, J.; McKeague, M.; Zhang, X.; Giamberadino, A., McConnell, E.; Francis, T.; DeRosa, M. C.; Dumontier, M. Aptamer Base: A Collaborative Knowledge Base to Describe Aptamers and SELEX Experiments. *Database* **2012,** *2012*, bas006.

(6) Green, L. S.; Jellinek, D.; Bell, C.; Beebe, L. A. Feistner, B. D.; Gill, S. C.; Jucker, F. M.; Janjić, N. Nuclease-Resistant Nucleic Acid Ligands to Vascular Permeability Factor/Vascular Endothelial Growth Factor. *Chem. Biol.* **1994,** *2*, 683-695.

(7) Kimoto, M.; Yamashige, R.; Matsunaga, K-I.; Yokoyama, S., Hirao, I. Generation of High-Affinity DNA Aptamers Using an Expanded Genetic Alphabet. *Nat. Biotech.* **2013,** *31*, 453-457.

(8) Ng, E. W. M.; Shima, D. T.; Calias, P.; Cunningham Jr., E. T., Guyer, D. R.; Adamis, A. P. Pegaptanib, A Targeted Anti-Vegf Aptamer For Ocular Vascular Disease. *Nat. Rev. Drug Discov.* **2006,** *5*, 123-132.

(9) Jellinek, D.; Green, L. S.; Bell, C.; Janjic, N. Inhibition of Receptor Binding by High-Affinity RNA Ligands to Vascular Endothelial Growth Factor. *Biochemistry* **1994,** *33*, 10450-10456.

(10) Ruckman, J.; Green, L. S.; Beeson, J.; Waugh, S.; Gillette, W. L.; Henninger, D. D.; Claesson-Welsh, L.; Janjić, N. 2′-Fluoropyrimidine RNA-Based Aptamers to the 165-Amino Acid Form of Vascular Endothelial Growth Factor (VEGF165): Inhibition of Receptor Binding and Vegf-Induced Vascular Permeability Through Interactions Requiring the Exon 7-Encoded Domain. *J. Biol. Chem.* **1998,** *273*, 20556-20567.

(11) Griffin, L. C.; Tidmarsh, G. F.; Bock, L. C.; Toole, J. J.; Leung, L. L. In Vivo Anticoagulant Properties of a Novel Nucleotide-Based Thrombin Inhibi-

tor 27 and Demonstration of Regional Anticoagulation in Extracorporeal Circuits. *Blood* **1993,** *81*, 3271-3276.

(12) Kurreck, J. Antisense Technologies. *Eur. J. Biochem.* **2003,** *270*, 1628-1644.

(13) Sczepanski, J. T.; Joyce, G. F., Binding of a Structured D-RNA Molecule by an L-RNA Aptamer. *J. Am. Chem. Soc.* **2013,** *135*, 13290-13293.

(14) Hoehlig, K.; Maasch, C.; Shushakova, N.; Buchner, K.; Huber-Lang, M.; Purschke, W. G.; Vater, A.; Klussmann, S. A Novel C5a-Neutralizing Mirror-Image (L-)Aptamer Prevents Organ Failure and Improves Survival in Experimental Sepsis. *Mol. Ther.* **2013,** *21*, 2236-2246.

(15) Yu, H.; Zhang, S.; Dunn, M. R.; Chaput, J. C. An Efficient and Faithful In Vitro Replication System for Threose Nucleic Acid. *J. Am. Chem. Soc.* **2013,** *135*, 3583-3591.

(16) Yu, H.; Zhang, S.; Chaput, J. C. Darwinian Evolution of An Alternative Genetic System Provides Support for TNA as an RNA Progenitor. *Nat. Chem.* **2012,** *4*, 183-7.

(17) Dunn, M. R.; Chaput, J. C. An In Vitro Selection Protocol for Threose Nucleic Acid (TNA) Using DNA Display. *Current Protocols in Nucleic Acid Chemistry / Edited by Serge L. Beaucage ... [et al.]* **2014,** *57*, 9.8.1-9.8.19.

(18) Dunn, M. R.; Larsen, A. C.; Zahurancik, W. J.; Fahmi, N. E.; Meyers, M. Suo, Z.; Chaput, J. C. DNA Polymerase-Mediated Synthesis of Unbiased Threose Nucleic Acid (TNA) Polymers Requires 7-Deazaguanine to Suppress G:G Mispairing During TNA Transcription. *J. Am. Chem. Soc.* **2015,** *137*, 4014-4017.

(19) Pinheiro, V. B.; Arangundy-Franklin, S.; Holliger, P. Compartmentalized Self-Tagging For In Vitro-Directed Evolution of XNA Polymerases. *Current Protocols in Nucleic Acid Chemistry / Edited by Serge L. Beaucage ... [et al.]* **2014,** *57*, 9.9.1-9.9.18.

(20) Pinheiro, V. B.; Taylor, A. I.; Cozens, C.; Abramov, M.; Abramov, M.; Renders, M.; Zhang, S.; Chaput, J. C.; Wengel, J.; Peak-Chew, S-Y.; Mclaughlin, S. H.; Herdewijn, P.; Holliger, P. Synthetic Genetic Polymers Capable of Heredity and Evolution. *Science* **2012,** *336*, 341-344.

(21) Alves Ferreira-Bravo, I.; Cozens, C.; Holliger, P.; Destefano, J. J. Selection Of 2'-Deoxy-2'-Fluoroarabinonucleotide (FANA) Aptamers That Bind HIV-1 Reverse Transcriptase with Picomolar Affinity. *Nucleic Acids Res.* **2015,** *43*, 9587-9599.

(22) Kasahara, Y.; Irisawa, Y.; Fujita, H.; Yahara, A.; Ozaki, H.; Obika, S.; Kuwahara, M. Capillary Electrophoresis–Systematic Evolution of Ligands by Exponential Enrichment Selection of Base- and Sugar-Modified DNA Aptamers: Target Binding Dominated by 2′-O,4′-C-Methylene-Bridged/Locked Nucleic Acid Primer. *Anal. Chem.* **2013,** *85*, 4961-4967.

(23) Kasahara, Y.; Irisawa, Y.; Ozaki, H.; Obika, S., Et Al., 2′,4′-BNA/LNA Aptamers: CE-SELEX Using A DNA-Based Library of Full-Length 2′-O,4′-C-Methylene-Bridged/Linked Bicyclic Ribonucleotides. *Bioorg. Med. Chem. Lett.* **2013,** *23*, 1288-1292.

(24) Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. PNA Hybridizes To Complementary Oligonucleotides Obeying the Watson-Crick Hydrogen-Bonding Rules. *Nature* **1993,** *365*, 566-568.

(25) Brudno, Y.; Birnbaum, M. E.; Kleiner, R. E.; Liu, D. R., An In Vitro Translation, Selection and Amplification System for Peptide Nucleic Acids. *Nat. Chem. Biol.* **2010,** *6*, 148-155.

(26) Niu, J.; Hili, R.; Liu, D. R. Enzyme-Free Translation Of DNA Into Sequence-Defined Synthetic Polymers Structurally Unrelated to Nucleic Acids. *Nat. Chem.* **2013,** *5*, 282-292.

(27) Latham, J. A.; Johnson, R.; Toole, J. J. The Application of A Modified Nucleotide In Aptamer Selection: Novel Thrombin Aptamers Containing 5-(1-Pentynyl)-2'-Deoxyuridine. *Nucleic Acids Res.* **1994,** *22*, 2817-2822.

(28) Gold, L.; Ayers, D.; Bertino, J.; Bock, C., Et Al. Aptamer-Based Multiplexed Proteomic Technology for Biomarker Discovery. *Plos ONE* **2010,** *5*, E15004.

(29) Rich, A., *Horizons In Biochemistry*. Academic Press: New York, 1962; pp 103-126.

(30) Sismour, A. M.; Benner, S. A., The Use of Thymidine Analogs To Improve The Replication of An Extra DNA Base Pair: A Synthetic Biological System. *Nucleic Acids Res.* **2005,** *33*, 5640-5646.

(31) Morales, J. C.; Kool, E. T. Efficient Replication Between Non-Hydrogen Bonded Nucleoside Shape Analogs. *Nat Struct. Mol. Biol.* **1998,** *5*, 950-954.

(32) Mcminn, D. L.; Ogawa, A. K.; Wu, Y.; Liu, J.; Shultz, P. G.; Romesberg, F. E. Efforts Toward Expansion of the Genetic Alphabet: DNA Polymerase Recognition of A Highly Stable, Self-Pairing Hydrophobic Base. *J. Am. Chem. Soc.* **1999,** *121*, 11585-11586.

(33) Hirao, I.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Kawai, R.; Satp, A.; Harada, Y.; Yokoyama, S. An Unnatural Hydrophobic Base Pair System: Site-Specific Incorporation of Nucleotide Analogs Into DNA and RNA. *Nat. Meth.* **2006,** *3*, 729-735.

(34) Yang, Z.; Hutter, D.; Sheng, P.; Sismour, A. M., Benner, S. A. Artificially Expanded Genetic Information System: A New Base Pair with an Alternative Hydrogen Bonding Pattern. *Nucleic Acids Res.* **2006,** *34*, 6095-6101.

(35) Georgiadis, M. M.; Singh, I.; Kellett, W. F.; Hoshika, S.; Benner, S. A.; Richards, N. G. J. Structural Basis for a Six Nucleotide Genetic Alphabet. *J. Am. Chem. Soc.* **2015,** *137*, 6947-6955.

(36) Sefah, K.; Yang, Z.; Bradley, K. M.; Hoshika, S., et al. In Vitro Selection With Artificial Expanded Genetic Information Systems. *Proc. Natl. Acad. Sci.* **2014,** *111*, 1449-1454.

(37) Hirao, I.; Mitsui, T.; Kimoto, M.; Yokoyama, S. An Efficient Unnatural Base Pair for PCR Amplification. *J. Am. Chem. Soc.* **2007,** *129*, 15549-15555.

(38) Kimoto, M.; Kawai, R.; Mitsui, T.; Yokoyama, S.; Hirao, I. An Unnatural Base Pair System for Efficient PCR Amplification and Functionalization of DNA Molecules. *Nucleic Acids Res.* **2009,** *37*, E14-E14.

(39) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. A Self-Cleaving DNA Enzyme Modified With Amines, Guanidines and Imidazoles Operates Independently of Divalent Metal Cations (M(2+)). *Nucleic Acids Res.* **2009,** *37*, 1638-1649.

(40) Perrin, D. M.; Garestier, T.; Hélène, C. Bridging The Gap Between Proteins and Nucleic Acids: A Metal-Independent RNase A Mimic With Two Protein-Like Functionalities. *J. Am. Chem. Soc.* **2001,** *123*, 1556-1563.

(41) Thum, O.; Jäger, S.; Famulok, M. Functionalized DNA: A New Replicable Biopolymer. *Angew. Chem. Int. Ed.* **2001,** *40*, 3990-3993.

(42) Vaught, J. D.; Bock, C.; Carter, J.; Fitzwater, T., et al. Expanding the Chemistry of DNA for in Vitro Selection. *J. Am. Chem. Soc.* **2010,** *132*, 4141-4151.

(43) Renders, M.; Miller, E.; Hollenstein, M.; Perrin, D. A Method For Selecting Modified DNAzymes Without the use of Modified DNA as a Template in PCR. *Chem. Comm.* **2015,** *51*, 1360-1362.

(44) Hili, R.; Niu, J.; Liu, D. R. DNA Ligase-Mediated Translation of DNA into Densely Functionalized Nucleic Acid Polymers. *J. Am. Chem. Soc.* **2013,** *135*(1), 98-101.

(45) Guo, C.; Watkins, C. P.; Hili, R. Sequence-Defined Scaffolding of Peptides On Nucleic Acid Polymers. *J. Am. Chem. Soc.* **2015,** *137* (34), 11191-11196.

(46) Tolle, F.; Brandle, G. M.; Matzner, D.; Mayer, G. A Versatile Approach Towards Nucleobase-Modified Aptamers. *Angew. Chem. Int. Ed.* **2015,** *54*, 10971-4.

(47) Dunn, M. R.; Jimenez, R. M.; Chaput, J. C. Analysis of Aptamer Discovery and Technology. *Nat. Rev. Chem.* **2017,***1*, 0076.

(48) Cho, E. U.; Lee, J-W.; Ellington, A. D. Applications Of Aptamers as Sensors. *Annu. Rev. Anal. Chem*. **2009**, *2*, 241-264.

(49) Navani, N. K.; Li, Y. Nucleic Acid Aptamer and Enzymes as Sensors. *Curr. Opin. Chem.* Biol. **2006,** *10*, 272-281.

(50) Liu, J.; Cao, Z.; Lu,Y. Functional Nucleic Acid Sensors. *Chem. Rev*. **2009,** *109*, 1948-1998.

(51) Mckeague, M.; Mcconnell, E. M.; Cruz-Toledo, J.; Bernard, E.; Foster, A.; Mastronardi, E.; Zhang, X.; Beking, M.; Francis, T.; Giamberardino, A.; Cabecinha, A.; Ruscito, A.; Aranda-Rodriguez, R.; Dumontier, M.; Derosa, M.C. Analysis of In Vitro Aptamer Selection Parameters. *J. Mol. Evol*. **2015**, *81*, 150-161.

(52) Ruscito, A.; Derosa, M. C. Small-Molecule Binding Aptamers: Selection Strategies, Characterization, and Applications. *Front. Chem.* **2016**, *4,* 14.

(53) Pfeiffer, F.; Mayer, G. Selection And Biosensor Application of Aptamers for Small Molecules. *Front. Chem.* **2016,** *4,* 25.

(54) Mckeague, M.; Derosa, M.C. Challenges and Opportunities for Small Molecule Aptamer Development. *J. Nucleic Acids*, **2012**, 1–20.

(55) Li, Y.; Lu, Y. *Functional Nucleic Acids for Analytical Applications*. Springer: New York, 2010; pp 47-108.

(56) Hamaguchi, N.; Ellington, A.; Stanton, M. Aptamer Beacons for the Direct Detection of Proteins. *Anal. Biochem*. **2001**, *294*, 126-131.

(57) Kent, A. D.; Spiropulos, N. G; Heemstra, J. M. General Approach for Engineering Small-Molecule-Binding DNA Split Aptamers. *Anal. Chem.* **2013,** *85*, 9916–9923.

(58) Sharma, A. K.; Heemstra, J. M. Small-Molecule Dependent Split-Aptamer Ligation. *J. Am. Chem. Soc.* **2011**, *133,* 12426-12429.

(59) Sharma, A. K.; Kent, A. D.; Heemstra, J. M. Enzyme-Linked Small-Molecule

Detection Using Split Aptamer Ligation. *Anal. Chem.* **2012,** *84*, 6104–6109.

(60) Nutiu, R.; Li, Y. Structure-Switching Signaling Aptamers. *J. Am. Chem. Soc*. **2003,** *125*, 4771-4778.

(61) Nutiu, R.; Li, Y. In Vitro Selection of Structure-Switching Signaling Aptamers. *Agnew. Chem. Int. Ed.* **2005**, *125*, 4771-4778.

(62) Mendonsa, S. D.; Bowser, M. T. In Vitro Selection of High Affinity DNA Ligands for Human Ige Using Capillary Electrophoresis. *Anal. Chem.* **2004,** *76,* 5387-5392.

(63) Armitage, B. Imaging of RNA in Live Cells. *Curr. Opin. Chem. Biol*. **2011**, *15*, 806-812.

(64) Singer, R. H.; Ward, D. C. Actin Gene Expression Visualized in Chicken Muscle Tissue Culture By Using In Situ Hybridization with a Biotinated Nucleotide Analog. *Proc. Natl. Acad. Sci. U. S. A.* **1982,** *79,* 7331-7335.

(65) Tyagi, S. Imaging Intracellular RNA Distribution and Dynamics in Living Cells. *Nat. Methods* **2009**, *6*, 331-338.

(66) Gao, W.; Xing, B.; Tsien, R. Y.; Rao, J. Novel Fluorogenic Substrates for Imaging B-Lactamase Gene Expression. *J. Am. Chem. Soc.* **2003,***125*, 11146–11147.

(67) Paige, J. S.; Wu, K. Y.; Jaffrey, S. RNA Mimics of Green Fluorescent Protein. *Science* **2011**, *333*, 642–646.

CHAPTER 2

DEVELOPING AND EXPLORING THE SUBSTRATE SCOPE OF A SELF-ALKYLATING RIBOZYME FOR MRNA LABELING

Introduction[†]

Localization and transport patterns for mRNA are key to many biological processes including cell growth and differentiation.[1] Consequently, atypical RNA localization has been associated with the progression of several detrimental diseases.[2-4] Fragile X syndrome, which causes mental retardation and autism in males, is characterized by a mutation in the FMR1 gene. This gene codes for the RNA-binding protein FMRP, which is known to bind to ~4% of mRNA in the brain and is responsible for localization and translation of these RNAs in dendrites.[5,6] These processes are critical for proper synapse development,[7] and thus dysregulation of those mRNA can be deleterious. Therefore, there is a growing demand for techniques to visualize mRNA dynamics which could help elucidate these cellular mechanisms and provide insight into the etiology and potential therapeutics

[†]Adapted with permission from ACS Chemical Biology, Ashwani K. Sharma, Joshua J. Plant, Alexandra E. Rangel, Kirsten N. Meek, April J. Anamisis, Julie Hollien, and Jennifer M. Heemstra. Fluorescent RNA Labeling Using Self-Alkylating Ribozymes. Vol 9., Pages 1680-1684. Copyright 2014 American Chemical Society.

for such diseases. There are currently a number of sophisticated methods to label RNA, however, these methods suffer from limitations making them less suitable for cellular studies. The predominant method, the MS2-GFP system, utilizes the MS2 bacteriophage coat protein (MCP) binding to a specific hairpin sequence (Figure 2.1a).[8] Multiple copies of this sequence are encoded onto an mRNA of interest, which is co-expressed with a fusion of the MCP to a fluorescent protein such as GFP. The MCP-GFP fusion binds to the hairpin sequence with extremely high affinity, allowing visualization of the target mRNA.[9] While effective, this method is characterized by drawbacks such as high background fluorescence and a bulky tag which is likely to perturb native RNA structure.[10]

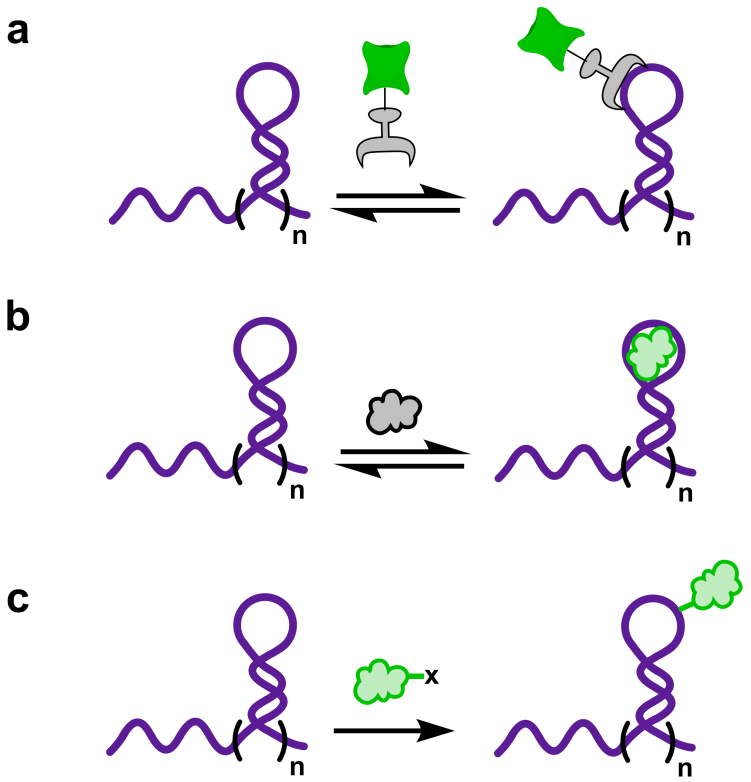


Figure 2.1. Methods for fluorescently labeling RNA using (a) RNA-binding proteins, (b) aptamer and fluorogenic small molecule, and (c) fluorescently self-alkylating ribozyme.

Recently, the small-molecule harnessing potential of aptamers, or short sequences of nucleic acids, have been seen as promising tools for improvements over previous labeling methods.[11] Small molecule fluorophores have many benefits for labeling including high quantum yield and small size less likely to perturb native RNA. Jaffrey and coworkers revolutionized the field in 2011 when they presented an aptamer for 4-hydroxybenzlidene imidazolinone, a GFP analogue, which could be attached to ribosomal RNA to enable visualization in live cells (Figure 2.1b).[12] Perhaps the most novel aspect of this system is the use of a fluorogenic small molecule. Steric constraints in the aptamer bound state turns on fluorescence, meaning that signal is only exhibited when the fluorophore is bound to the aptamer sequence.[11] This work provided many advantages over the MCP-GFP system such as small cargo size and elimination of background fluorescence. However, the noncovalent interaction between aptamer and small molecule prevents imaging sensitivity due to the possibility of the fluorophore dissociating from the aptamer. Nonetheless, this work provided huge improvements over previous labeling methods and established aptamers as a powerful tool to be implemented for such applications.

The goal of the present research is to circumvent the limitations of previous RNA labeling methods through the development of a self-alkylating ribozyme that acts as the recognition element and provides a covalent linkage to the fluorophore (Figure 2.1c). Ribozymes, or catalytic RNAs, exist for a variety of functions and have shown promise for RNA labeling but have not been adapted for live cell applications.[13,14] The covalent nature attachment not only offers additional stability of

the label, but also allows for excess fluorophore to be washed out of cells, consequently reducing signal loss and background fluorescence. Importantly, our fluorescently self-labeling ribozyme can be fused to an mRNA of interest to enable monitoring of localization and dynamics in living cells. Additionally, this covalent linkage also allows for critical advancement over previous methods in the way of potential applications.

One such application is the ability to perform pulse-chase labeling experiments. With this method, the target is labeled for a set amount of time and then excess fluorophore is removed which allows for time-resolved tracking of target biomolecules.[15] We also hope to have the ability to identify proteins that bind to specific mRNA sequences using a transcript-specific RNA immunoprecipitiation (TRIP) process. Traditional IP methods rely on isolation using an antibody for the target protein and subsequent RT-PCR to identify bound RNA sequences.[16] The lack of affinity reagents for RNA that function under IP conditions has made the identification of proteins that bind to specific RNA sequences a difficult task, and thus only limited examples of such experiments can be found in the literature.[17] Further, many of these methods rely on the affinity tag being attached to *in vitro* transcribed RNA with subsequent incubation in a cell lysate, which may alter RNA-protein interactions.[17] Our ribozyme will ensure that the fluorescein will bind to a specific RNA sequence *in vivo*, allowing for immunoprecipitiation of endogenous RNA-protein complexes. In conjunction, these experiments will provide vital information regarding transport mechanisms and localization patterns in cellular context. Our self-alkylating ribozyme is the first RNA labeling probe to have all of the

characteristics necessary to successfully perform immunoprecipitiation and pulse-chase labeling of specific RNA transcripts.

In the present work we generate self-alkylating ribozymes capable of covalently attaching fluorescein iodoacetamide (FIA) to itself using a novel systematic evolution of ligands by exponential enrichment (SELEX) method.[18-21] Our motivation to utilize this chemistry originated from Wilson and Szostak's report of an RNA capable of self-alkylation with biotin iodoacetamide (Figure 2.2a).[22] Biotin is readily available in cells and has no potential for direct optical detection; hence we were motivated to use fluorescein because of its superior brightness (Figure 2.2b). We aimed to fully characterize our ribozymes by mapping the self-alkylation reaction site and exploring substrate specificity and scope with other fluorophores. We also determined the ability of our ribozyme to label when fused to an mRNA sequence-

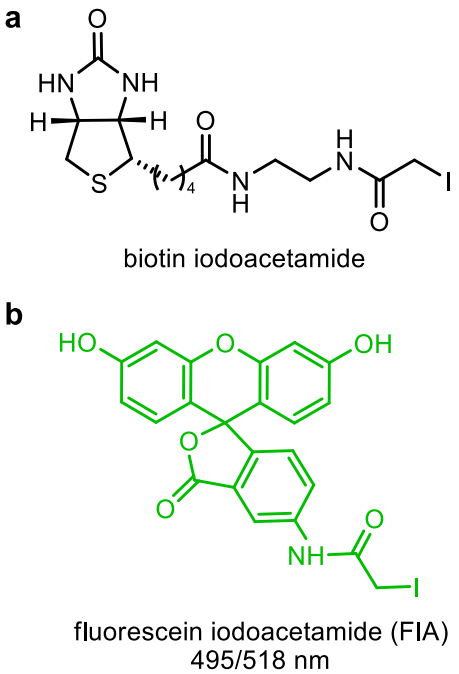


Figure 2.2. Chemical structure of (a) biotin iodoacetamide and (b) FIA.

a critical indicator of our system's potential success for cellular labeling experiments. Further, we show preliminary data supporting our ribozyme's compatibility with conditions necessary for TRIP experiments. Altogether, this study demonstrates the potential of our ribozymes as a promising tool for RNA imaging applications.

Results and discussion

IP-SELEX to generate self-alkylating ribozymes

In the process of selecting for our fluorescently self-alkylating ribozymes, we implemented an immunoprecipitation SELEX (IP-SELEX) process, in which "hit" sequences are pulled down using magnetic beads functionalized with an antibody for the fluorescein label (Figure 2.3a). This provides a convenient and highly selective method for separating ribozymes from the over-whelming majority of inactive sequences.[23] Additionally, reverse transcription of the ribozymes is carried out on the beads, obviating the need for an elution step. As shown in Figure 2.3a, a starting DNA library was synthesized having a 116 nt random region (N116) flanked by two primer binding sites. We amplified this DNA by PCR using a forward primer containing the T7 promoter sequence and converted the products into the corresponding RNA library via T7 in vitro transcription. We then incubated the RNA library with 200 μM FIA (initial incubation time was 8 h but was reduced to 15 min over iterative SELEX cycles to drive the selection toward ribozymes having fast reaction rates). The labeling reaction was quenched with 2-mercaptoethanol, and the RNA was purified by ethanol precipitation. To isolate RNA sequences that had been successfully labeled, we incubated the total RNA pool with magnetic beads

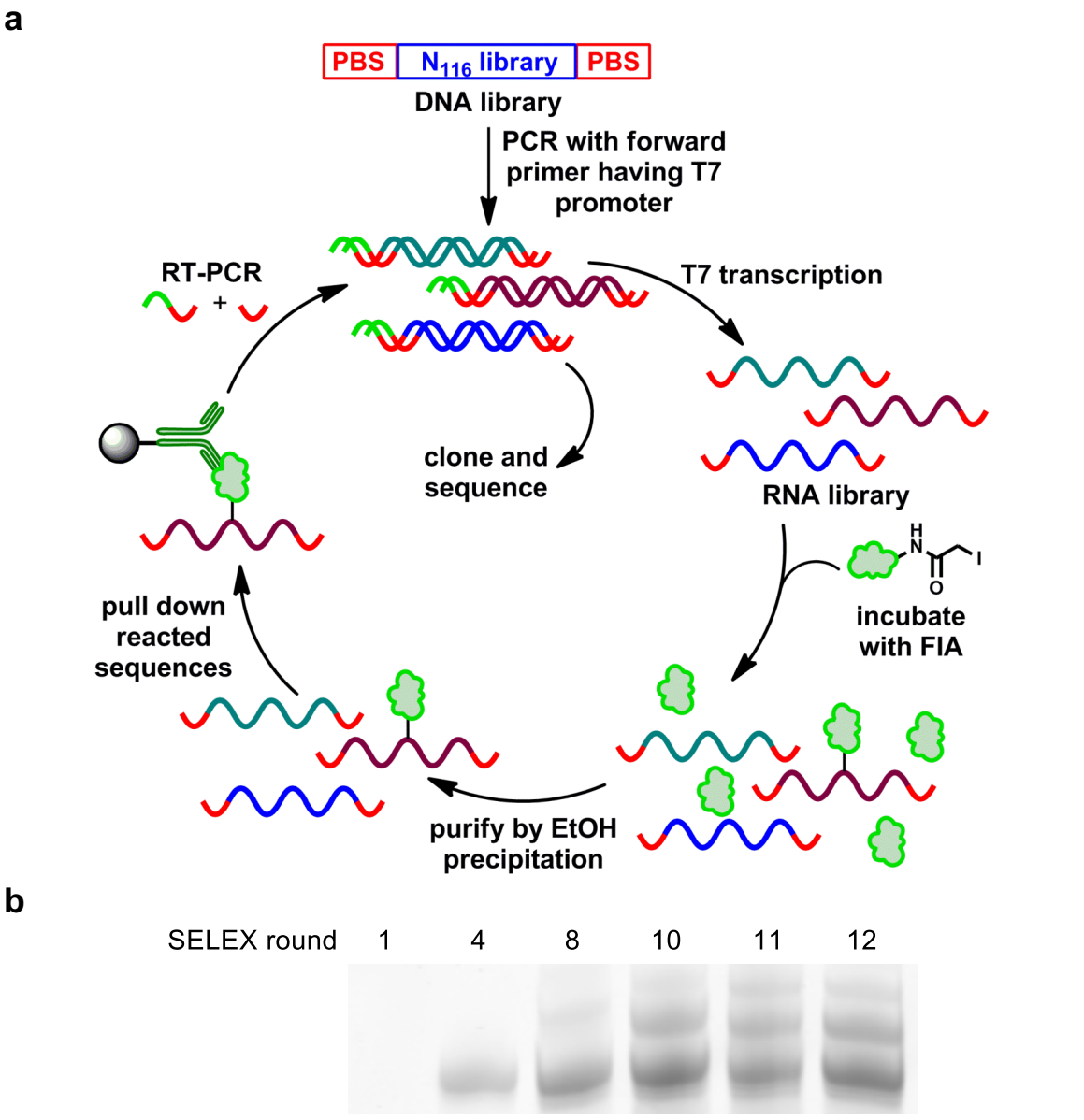
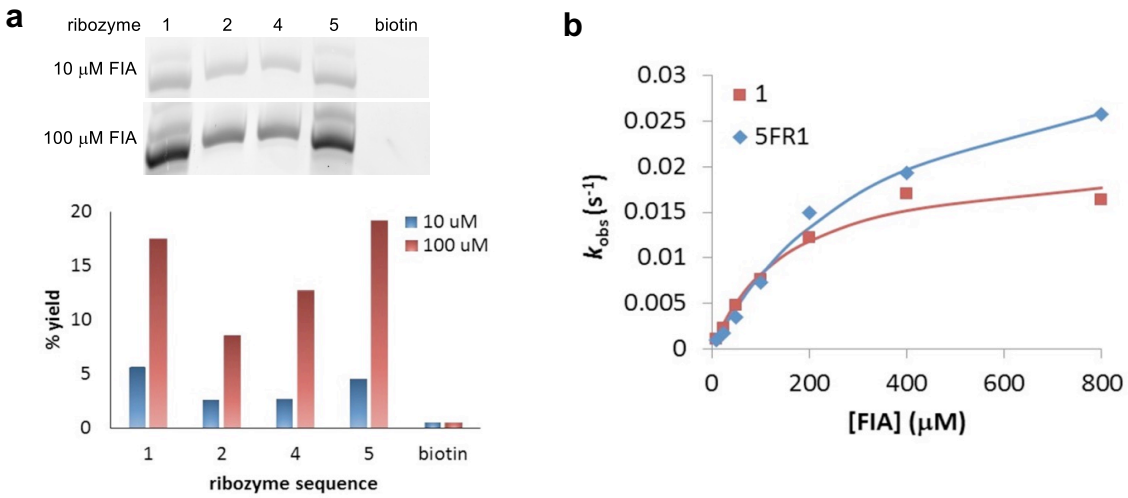


Figure 2.3. IP-SELEX to generate self-alkylating ribozymes. (a) SELEX Scheme-a DNA library is PCR amplified and then transcribed into an RNA library. The RNA library is allowed to react with FIA, and then active sequences are pulled down using an antifluorescein antibody. (b) Denaturing PAGE showing fluorescein labeling of RNA from SELEX rounds.

prefunctionalized with an antifluorescein antibody. Unbound RNA was removed by washing, and bound RNAs were reverse transcribed and PCR amplified, resulting in a double-stranded DNA library that could be subjected to a subsequent round of selection.

We evaluated the overall reactivity of the RNA library after rounds 4, 7, 10, 11, and 12. In these experiments, we incubated the library with 100 μM FIA for 15 min, ethanol precipitated the RNA to remove excess FIA, and examined the products using denaturing PAGE followed by fluorescence imaging. The same quantity of RNA was loaded into each lane of the gel, such that the intensity of the fluorescein band is representative of reaction efficiency. Importantly, the denaturing conditions of the gel disrupt any noncovalent RNA-fluorophore interactions, ensuring that fluorescence is observed only when the FIA is covalently tethered to the RNA sequence. Using this analysis, labeling efficiency appeared to plateau at round 12 (Figure 2.3b). Thus, selection cycles were stopped, and the RNA library was amplified by reverse transcription-PCR, inserted into a TOPO cloning vector, and amplified in E. coli.

We sequenced 20 individual clones from the round 12 RNA library. This revealed four different types of ribozyme sequences, each represented by several highly related or identical clones. We then used PCR and in vitro transcription to generate representative ribozymes from each of the 4 groups and compared them using the denaturing PAGE assay described above (Figure 2.4a, lanes 1−4). All four sequences showed self-labeling activity, but sequences 1 and 5 appeared to be the most reactive. Importantly, control RNA sequences including the biotin self-



c

Ribozyme 1 (157 nt):
GAGGCACACCACGGCTGGATCCGAGTGCGGCGTTAATCGCAATCTTACGGAAGCGGTAGTCGTGCGT
GCTAATCAAGGGTGTCCCACGAACATCAAGTGTCAGAGGACATAGGCGGGAAAACCGCGGTGTCGGG
TAGGCCCTTGGTCATTAGGATCG

Ribozyme 5FR1 (137 nt):
GAGGCACCACGGCTGGATCCCCACGAGCTGAGTTGAACACTTTCGCGTCTCTGCACGATCCGAATCAG
CACACGCGAGATCGGATCGGCGACTAGTGGCAATTCGGCCGGGACGACGGCCTAGGATATCCGCAGG
TG

Figure 2.4. IP-SELEX results. (a) Denaturing PAGE showing fluorescein labeling of ribozyme sequences from SELEX using 10 μM and 100 μM FIA. Labeling yields determined by HPLC. (b) Pseudo-first-order rate constant vs. [FIA] for ribozymes 1 and 5FR1. (c) Sequences for most reactive ribozymes. Sequences are displayed 5' to 3' and primer binding sites are underlined.

alkylating ribozyme and total cellular RNA do not show observable reaction with FIA (Figure 2.4a, lane 5), indicating that FIA reacts only with the selected ribozyme sequences and does not label RNA molecules nonspecifically.

The PAGE assay described above is useful for determining whether RNA sequences show self-labeling activity with FIA and can provide insight into the reactive reactivity of these ribozyme sequences. However, this assay does not provide quantitative information regarding reaction yields. Thus, to quantify reaction yields, we turned to HPLC (Figure 2.4a). Both reacted and unreacted RNA elute together as a single peak, but the relative peak areas (A) for absorbance at 260 nm (RNA) and 495 nm (fluorescein) can be used in conjunction with the molar absorptivity (ε) values for the RNA and fluorescein to calculate reaction yield according to the equation:

$$\%\, yield = \left[\frac{A_{495} * \varepsilon_{260}}{A_{260} * \varepsilon_{495}}\right] * 100 \quad (1)$$

Figure 2.4b shows the ligation yield for each of the ribozyme sequences after reaction with either 10 μM or 100 μM FIA for 30 min. These data verified that sequences 1 and 5 show the fastest reaction kinetics, and thus we selected these two sequences for further analysis.

Ribozyme minimization

To evaluate ribozymes 1 and 5, we first sought to determine whether the sequences could be truncated without impacting reactivity. Using denaturing PAGE and HPLC, we evaluated sequences having 18 nt removed from either the

5′ or 3′ terminus For ribozyme 1, shortening the sequence from either terminus to give 1F1R or 1FR1 resulted in a dramatic decrease in reactivity. For ribozyme 5, truncation at the 5′ terminus to give 5F1R also resulted in a dramatic decrease in reactivity. However, we were able to remove 18 nt from the 3′ terminus of ribozyme 5 to give 5FR1 without impacting reactivity. Further truncation at the 3′ terminus to give 5FR2 resulted in a decrease in reactivity, and thus sequence 5FR1 was selected for further study (sequences shown in Figure 2.4).

Kinetics of self-labeling reaction

To characterize the kinetics of the self-labeling reaction, we carried out labeling reactions under pseudo-first-order conditions using a large excess of FIA. Ribozymes 1 and 5FR1 were each reacted with FIA, then quenched, purified, and analyzed by HPLC as described above. For each concentration of FIA, we fit data from a series of time points to the integrated rate law $[A] = [A]_0\ e^{-kt}$ to give the pseudo-first-order rate constant (k_{obs}) for that reaction (Figure 2.4b). We then calculated the second order rate constant (k_2) for each reaction was using the equation $k_2 = k_{obs}/[FIA]$. The values of k_2 was averaged over the region that k_{obs} vs [FIA] was linear (10−50 μM for 1 and 10−200 μM for 5FR1), providing k_2 values of 100 ± 8 and 77 ± 12 $M^{-1}s^{-1}$ ribozymes 1 and 5FR1, respectively. Encouragingly, these rate constants are on par with those of other reactions that have been used successfully for intracellular biomolecule labeling.[24,25]

We hypothesize that ribozymes 1 and 5FR1 bind to FIA, increasing effective molarity and thus promoting the self- labeling reaction. The corresponding reaction mechanism is represented by the equation below in which R is the ribozyme, F is

FIA, P is labeled between the unimolecular rate constant for conversion of ribozyme-FIA complex into labeled product:[26]

$$R + F \overset{K_a}{\rightleftharpoons} R{\cdot}F \overset{k_u}{\rightarrow} P \qquad (2)$$

Ribozyme specificity

We first sought to evaluate the substrate specificity of our ribozyme sequences against other fluorophore iodoacetamides. We chose BODIPY FL iodoacetamide (BFIA) for these studies, as the core structure of the molecule is much different and would allow us to determine if the ribozymes were specific for the xanthene core of the FIA (Figure 2.5a). Importantly, these molecules share similar spectral properties and can be compared directly on the same gel. We labeled sequences 1 and 5FR1 with 100 μM FIA/BFIA and then analyzed our results with a denaturing PAGE assay (Figure 2.5b). We include the ethidium bromide staining in our analysis to quantify all RNA present in the lane for reference. Approximate selectivity values were determined by comparing ratios of fluorescein/ethidium bromide intensities for each sample. We were pleased to observe that our ribozymes are selective for FIA over BFIA with approximate selectivity of 160:1 and 280:1 for sequences 1 and 5FR1 respectively.

Exploring the substrate scope

For further exploration, we reacted both ribozyme sequences with tetramethylrhodamine iodoacetamide (TMRIA) and analyzed the results using PAGE analysis. Interestingly, we observed that not only do both ribozymes react with

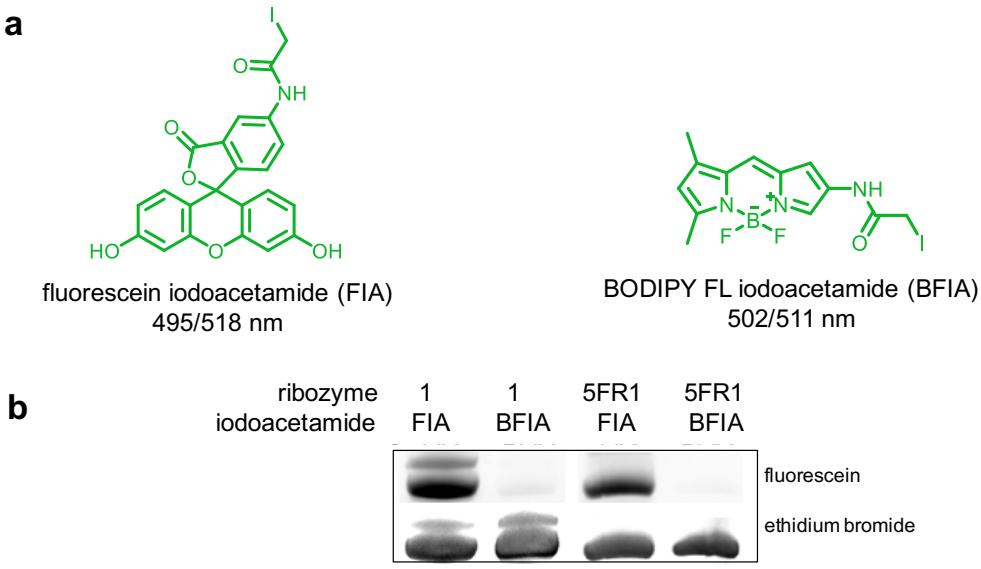


Figure 2.5. Ribozyme specificity experiments. (a) Chemical Structures of FIA and BFIA (b) Denaturing PAGE gel shows labeling of ribozymes with 100 μM FIA and BFIA to explore substrate specificity. Ethidium bromide image of same gel after staining included to depict relative RNA loading.

TMRIA but do so with higher efficiency than FIA. This is not surprising considering only positive selections were carried out in SELEX and TMRIA and FIA both have a similar xanthene based structures. With these results, we were curious to see if our ribozymes would show selectivity for the TMRIA over BODIPY 507 (Figure 2.6a), similar to the experiment discussed above with FIA/BFIA. We determined that our ribozymes are selective for TMRIA over B507 with approximate selectivity of 500:1 and 102:1 for sequences 1 and 5FR1 respectively (Figure 2.6b). This result is encouraging as it shows that our ribozymes are specific for xanthene based fluorophores.

We were then interested in determining how the kinetics of the ribozyme labeling reactions differ between FIA and TMRIA. The kinetics of the self-alkylation reaction for both ribozymes with TMRIA were characterized by HPLC analysis as

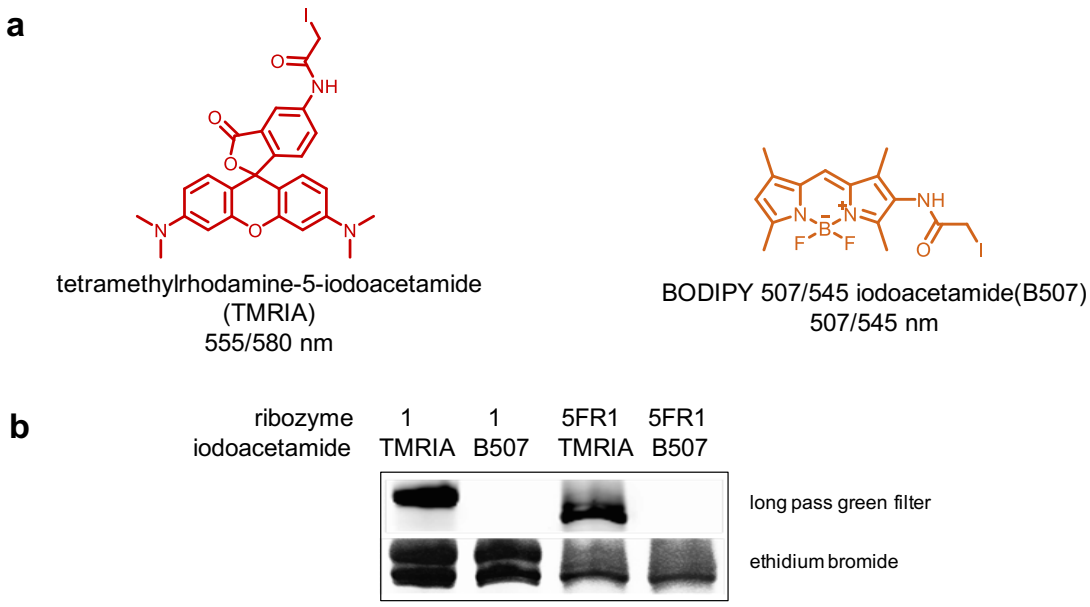
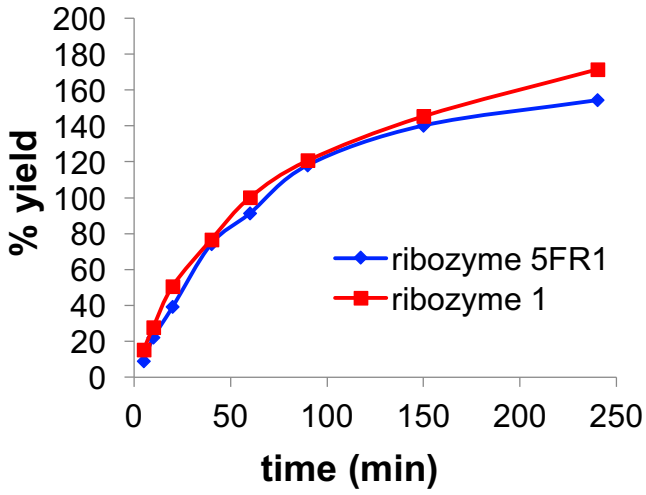


Figure 2.6. Ribozyme labeling with other xanthene iodoacetamides. (a) Chemical Structures of TMRIA and B507IA (b) Denaturing PAGE gel showing labeling of ribozymes with 100 μM TMRIA and B507IA to explore substrate scope. Ethidium bromide image of same gel after staining included to depict relative RNA loading.

discussed for the initial characterization of the ribozyme reactivity with FIA. Ribozymes 1 and 5FR1 were reacted with 100 μM TMRIA, and the reaction was quenched at different time points and analyzed by HPLC (Figure 2.7a). The relative peak areas at 260 nm (RNA) and 555 nm (TMRIA) were used to compute the reaction yield according to the equation used in the quantification of FIA labeling yields mentioned previously. Since one concentration of TMRIA was used for this experiment, only the k_2 could be determined, however, this still allows us to relatively compare rates for both ribozymes with both iodoacetamides. With 100 μM TMRIA, ribozymes 1 and 5FR1 react 1.9 and 2.1-times faster compared to the FIA labeling reaction (Figure 2.7b). We hypothesize that promiscuity of our ribozyme will make our system easily integrated into a variety of imaging applications where orthogonal fluorophores are advantageous. In particular, these results suggest that

a



b

	k_2 ($M^{-1}s^{-1}$)	
	1	**5FR1**
TMRIA	146	154
FIA	100±8	77±12

Figure 2.7. Ribozyme Labeling with xanthene-based fluorophores. (a) Labeling yields determined by HPLC for the reaction of ribozymes 1 and 5FR1 with 100 μM TMRIA. (b) Comparison of second order rate constants for labeling of ribozymes with FIA and TMRIA.

our ribozymes will allow for color flexibility for time-resolved imaging, which requires that each fluorophore be imaged individually (and emission spectra must not overlap).

Mapping the reaction site

Importantly, we aimed to locate the alkylation site for FIA on the ribozyme. We chose to use sequence 5FR1 for these experiments since it exhibits better reaction efficiency with the FIA. Ribozyme previously labeled with FIA was then

biotinylated on the 3’ terminus using a commercially available kit. Biotinylated ribozyme was then subjected to hydrolysis in alkaline buffer according to the scheme in Figure 2.8. The alkaline hydrolysis fragments were then immobilized onto M-270 streptavidin magnetic beads and nonbound fragments were separated using a magnetic separation stand and discarded. The size of hydrolysis fragments eluted from the beads were analyzed using a denaturing PAGE and compared to the decade marker ladder to determine size. The gel was scanned for fluorescence and stained with SYBR Green II, which we chose because of its superior detection of RNA. The gel shows labeled 5FR1 sequences having 3’ terminus intact after hydrolysis. Only the full sequence shows fluorescein signal, and the next longest sequence observed using SYBR Green II has an approximate length of 100 nt (Figure 2.9). This indicates that the fluorescein reaction site is within the 40 nt closest to the 5’ terminus of the ribozyme. We hypothesize that our inability to resolve the specific reaction site is due to secondary structure present in the ribozyme that prevents complete hydrolysis. We postulate that our ribozymes bind FIA, increasing effective molarity, and thus promotes the alkylation reaction. We propose that our ribozymes either bind FIA in a conformation that places the iodoacetamide in close proximity to a nucleophilic group, or specific nucleotides in the ribozyme activate the electrophile for nucleophilic attack.

Cellular labeling compatibility

While the aforementioned experiment provides very useful information about the way our ribozymes function, we ultimately want to determine their compatibility with mRNA labeling experiments. To use our ribozymes for such studies,

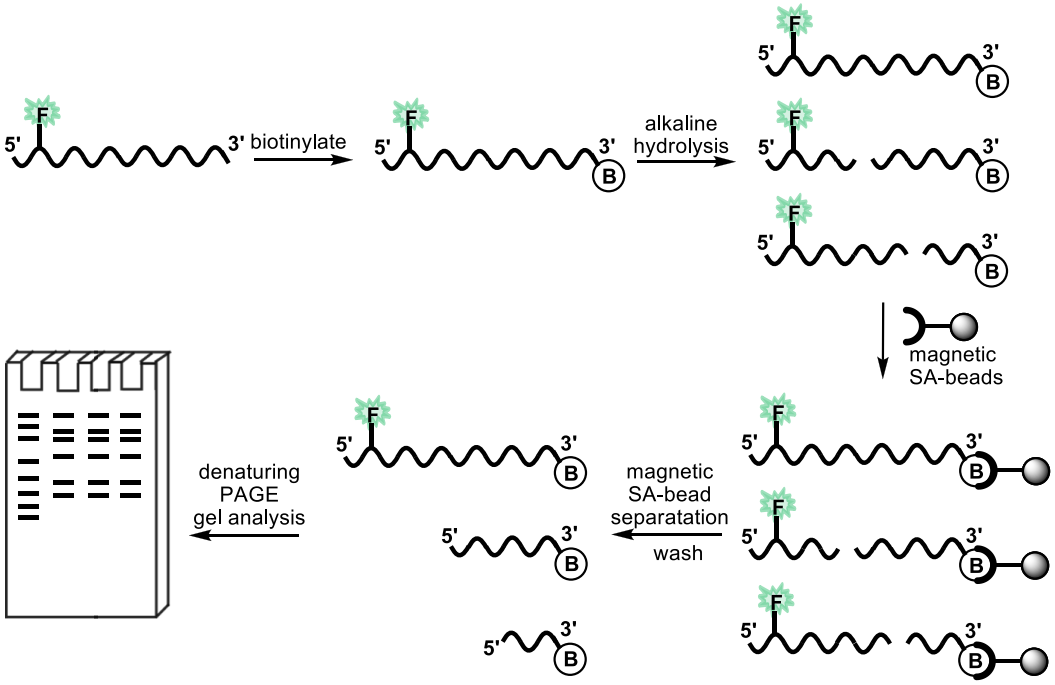


Figure 2.8. Scheme of mapping protocol to determine reaction site for ribozyme 5FR1.

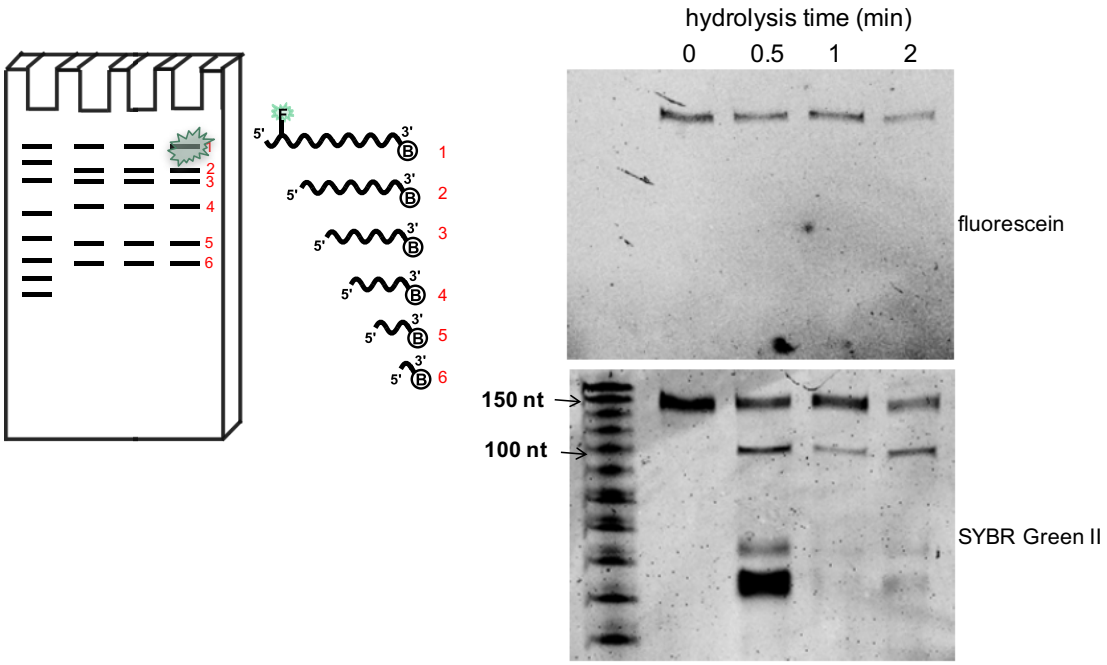


Figure 2.9. Denaturing PAGE showing fragments having 3' terminus intact after hydrolysis. SYBR green II staining shown below to depict all RNA present in sample.

it must be genetically encoded as a fusion to an mRNA of interest. In order for this to be plausible, we desired to verify that attaching the ribozyme to a much longer mRNA sequence would not impact reactivity. Ribozyme sequences were inserted into the 3' untranslated region (UTR) of an RNA containing 432 nucleotides of the sequence coding for the mCherry fluorescent protein. This slightly truncated version of the full mCherry sequence was used for better compatibility with PAGE analysis. Both 1 and 5FR1 were fused to the mRNA sequence to give fusions with single and multiple copies of the ribozymes, and were reacted with FIA and compared to reactivity of an mCherry control sequence with no ribozyme attached. Unfortunately, our initial experiments showed no labeling of the fusions, even at very high concentrations of FIA and long incubation times. Our next steps were to determine what factors were impeding labeling and what necessary changes should be made to obtain fusion labeling. We hypothesized that the mRNA, being a much larger sequence, might place itself in a conformation that prevents the self-alkylating reaction when fused to the ribozyme. We reacted our ribozymes with FIA in the presence of increasing amounts of total cellular RNA (by weight) to investigate if large amounts of RNA (i.e. the mCherry sequence) inhibit labeling (Figure 2.10a). A progressive decrease in fluorescence intensity was observed with increasing amount of total RNA, however, at 8X total RNA by weight labeling is still observed. We were encouraged that in the presence of such a large amount of RNA our ribozymes were still capable of reacting with FIA, however, labeling efficiency is significantly diminished. We hypothesized that using a high concentration of ribozyme for our labeling reactions (4 uM) causes the local concentration of RNA

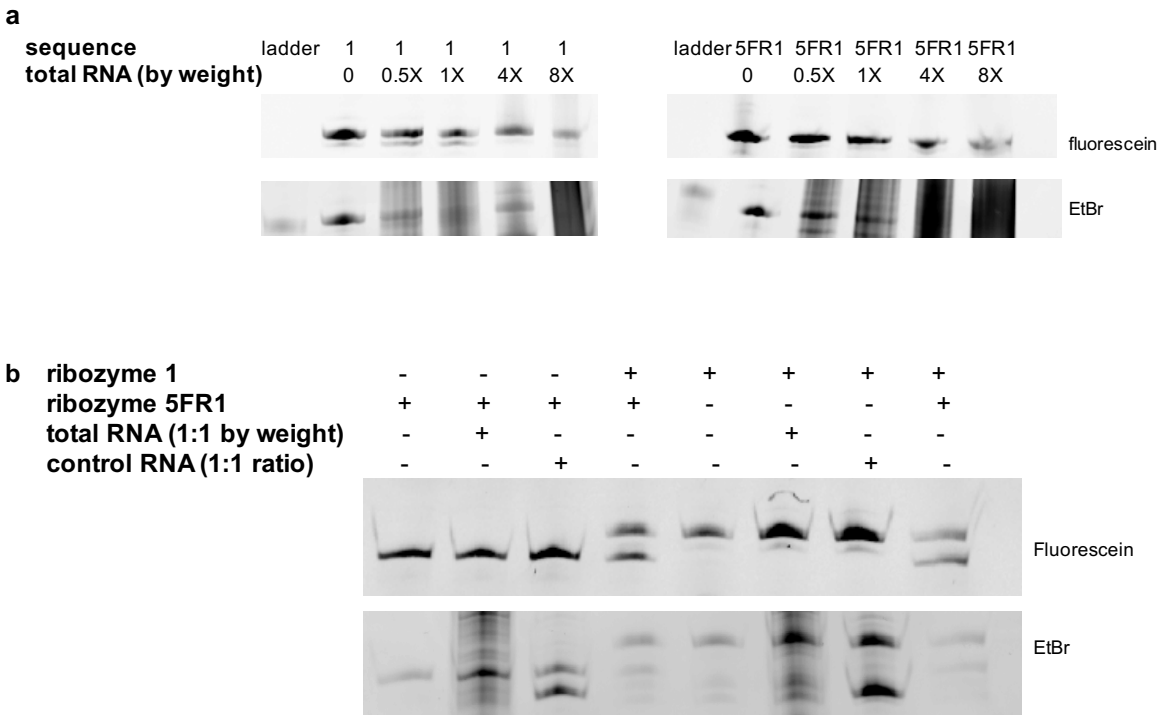


Figure 2.10. Ribozyme labeling in the presence of other RNAs**.** (a) Labeling of ribozymes in the presence of increasing total cellular RNA. (b) Effect of lowering RNA concentration (from 4 μM to 400 nM) on the ability of ribozymes to be labeled in the presence of other RNAs.

to be so high that it cannot be effectively labeled. Consequently, we performed an experiment in which we lowered the concentration of ribozyme 10-fold to 400 nM and attempted labeling in the presence of other RNAs (ribozymes, mCherry control sequence, and total cellular RNA). From these experiments, we observed no obvious difference in fluorescence intensity of both ribozymes in the presence of other RNAs as seen in Figure 2.10b. We concluded that lowering our ribozyme concentration by a factor of ten would be more favorable for our fusion labeling experiments. In addition, we hypothesized that our SELEX buffer, containing a high concentration of lithium ions, was causing the mRNA to fold around the ribozyme, preventing effective labeling. As discussed previously, one application for our labe-

ling system is to be able to identify proteins that bind to mRNA sequences using our TRIP methodology. For our experiments, we will crosslink the RNA and proteins in live cells, and fluorescein labeling will occur after cell lysis. As a preliminary experiment, we wanted to determine if the conditions used to crosslink mRNAs to their protein binding partners would impact the reactivity of the ribozyme (Figure 2.11). We incubated ribozyme 5FR1 with 1 and 0.1% formaldehyde, quenched with glycine and then attempted ribozyme labeling with FIA. We observed that the labeling of ribozyme incubated in buffer alone. These results suggest that our system will be compatible with TRIP experiments and allow for successful identification of protein which bind to specific RNA sequences.

Labeling of mRNA fusions

To test the ability of the ribozymes to function when fused to an RNA of interest, we inserted each ribozyme into the 3′ untranslated region (UTR) of an RNA containing 432 nucleotides of the coding sequence for the fluorescent protein exposed to formaldehyde was comparable to the labeling efficiency of ribozyme

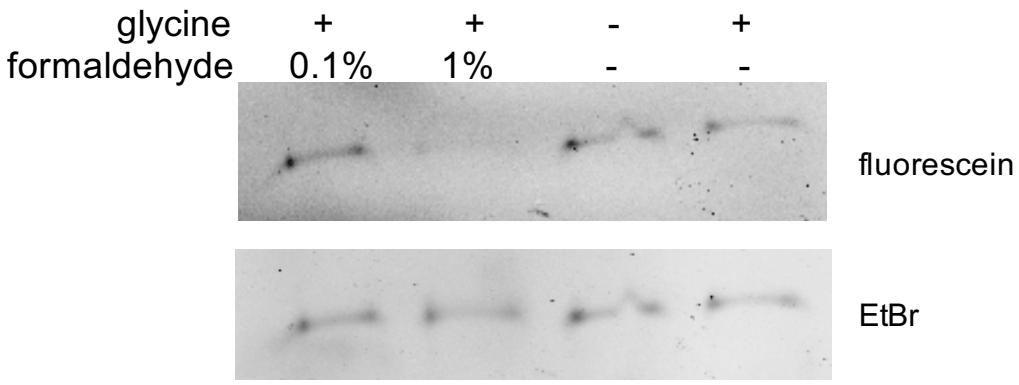


Figure 2.11. Labeling of ribozyme 5FR1 in crosslinking conditions

mCherry. As observed in Figure 2.12, we observe labeling for both of these ribozyme-RNA fusions compared with the mCherry RNA alone as a control. We also sought to evaluate the ability of the ribozymes to undergo labeling in a complex biological sample, as the thiols present in proteins and glutathione may outcompete the ribozyme for labeling with FIA. We were pleased to observe that labeling of the ribozymes proceeded in cell lysate with efficiencies similar to those observed in buffer. Including multiple copies of the ribozymes does not seem to have any benefit on the labeling yield.

Conclusions

We demonstrated that generation of self-alkylating ribozymes for small molecule targets is possible through a novel IP-SELEX method. Further, we show that our ribozyme is capable of fluorescently labeling with FIA through a covalent attachment and are selective for FIA over other fluorophore iodoacetamides. Interestingly, our ribozymes show faster reaction kinetics with TMRIA, which shares the same xanthene core structure as FIA. This promiscuity of substrate binding, while

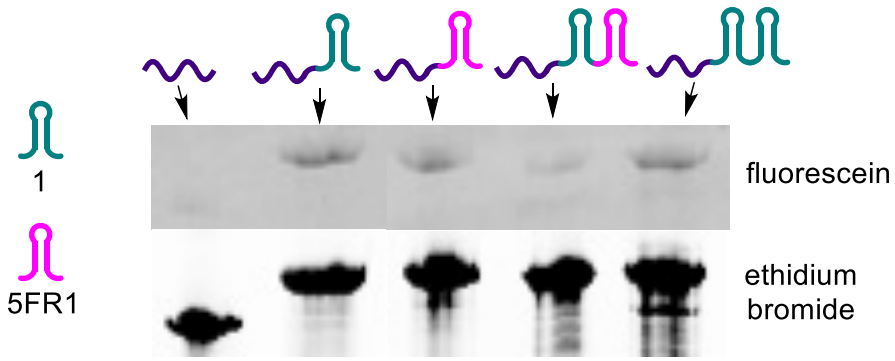


Figure 2.12. Labeling of ribozymes 1 and 5FR1 when inserted into the 3' UTR of an RNA encoding a portion of the mCherry protein. Constructs included single and multiple copies of the ribozyme sequences.

surprising, suggests our ribozyme has multicolor imaging capabilities lending it as a versatile tool for potential imaging applications. Additional experiments were performed to characterize the approximate location of the alkylation site and we determined that the reactive site is within the first 40 nucleotides near the 5' end of the ribozyme. Importantly, we show that both sequences show significant labeling when fused to an mRNA of interest, which is encouraging for success of our ribozyme *in vivo*. Last, we show that labeling is still observed after RNA is subjected to crosslinking conditions, supporting the compatibility of our system for cellular labeling and applications such as TRIP.

Materials and methods

All DNA sequences were ordered from the University of Utah's DNA and Peptide synthesis core facility. Enzymes were purchased from New England Biolabs.

Attachment of antifluorescein antibody to magnetic beads

100 µL of M-280 streptavidin beads (Life Technologies) were washed three times with 200 µL of PBS. After each wash, the beads were separated from the supernatant using a magnetic separation stand (DynaMag-Spin, Life Technologies). The beads were suspended in 180 µL of PBS and 20 µL (20 µg) of biotinylated anti-fluorescein antibody (Abcam) was added. The suspension was agitated on a nutating mixer for 3 h at 37 °C, followed by overnight shaking at 4 °C to further ensure the maximum loading of biotinylated antibody onto the streptavidin beads.

The beads were washed three times with 200 µL of PBS, resuspended in 100 µL PBS and stored at 4 °C until further use.

IP-SELEX

A dsDNA library (0.4−1 µg) was transcribed in a 20 µL reaction for 5 h at 37 °C using an AmpliScribe T7-Flash transcription kit (Epicentre Biotechnologies), followed by DNase treatment for 1 h at 37 °C. The resulting RNA was precipitated using sodium acetate/ethanol. The precipitated RNA library was dissolved in 500 µL of selection buffer (100 mM LiCl, 10 mM Na-HEPES, 5 mM $MgCl_2$, pH 7.4) and coincubated with 200 µM 5-(iodoacetamido)fluorescein (Aldrich) at 37°C. The incubation time was 8 h for the initial round of SELEX and was gradually shortened to 15 min over subsequent rounds of SELEX to increase the stringency of selection. The reaction was quenched by the addition of 100 mM β-mercaptoethanol, 5 mM EDTA, 0.3 M NaCl, and 50 µg of yeast tRNA (Ambion). After 15 min of quenching, the reaction was precipitated using sodium acetate/ethanol.

Twenty microliters of antibody-functionalized magnetic beads (suspended in PBS) was placed on a magnetic separation stand to enable removal of the supernatant. The precipitated RNA was resuspended in 100 µL of PBS and added to the antibody functionalized magnetic beads. The mixture was incubated at RT for 3 h with mild shaking. The unbound RNA was removed by washing the beads three times with 100 µL of wash buffer (1 M NaCl, 10 mM Na-HEPES, 5 mM EDTA, pH 7.4) and then three times with 100 µL of water. The RNA-bound beads were subjected to reverse transcription in a 25 µL reaction using 1 µL (200 units) of M-MLV reverse transcriptase (Life Technologies). The reaction was incubated at 37 °C for

1 h, then diluted into ThermoPol buffer (New England Biolabs), and subjected to 34 rounds of PCR in a 500 µL reaction using the thermocycling protocol and primers described in the Supporting Information. The resulting dsDNA was precipitated using sodium acetate/ethanol. The dsDNA was separated on a 2% agarose gel at 70 V for 45 min, and the desired band was cut from the gel and purified using a Zymoclean Gel DNA Recovery Kit (Zymo Research). The resulting dsDNA was used as a template for in vitro transcription for the next round of SELEX. After 12 rounds of SELEX, the dsDNA was cloned into a TOPO vector using a TOPO TA cloning kit (Life Technologies) and amplified in E. coli. Plasmid DNA from 20 individual colonies was recovered using a MiniPrep kit (Qiagen) and was sequenced at the University of Utah DNA sequencing core facility.

Ribozyme labeling with fluorophore iodoacetamides

Reactions were typically carried out in a volume of 250 µL, in a solution containing 4 µM (unless specified otherwise) transcribed RNA and the specified concentration of IA (FIA, FBA, TMRIA, B507) in selection buffer or Dulbecco's PBS. The reactions were incubated at 37°C for the specified time, then quenched by the addition of 100 mM β-mercaptoethanol, 5 mM EDTA, 0.3 M NaCl. After 15 min of shaking at room temperature, the reactions were precipitated using sodium acetate/ethanol, stored at -80 °C for 12 h, then centrifuged at 17,000 × g for 15 min at room temperature. The resulting RNA pellet was washed with 2 x 200 µL of 70% ethanol, air-dried, and resuspended in 30 µl of RNAse free water.

PAGE analysis of labeling reactions

The RNA purified from each reaction was loaded onto a 6% polyacrylamide gel containing 8 M Urea. For reactions using the fusion constructs, samples were run on 3.5% polyacrylamide gels in order to resolve the larger sequences. The gels were run at 150 Volts for 1 h, and were analyzed on a Typhoon 9500 fluorescence laser scanner (Amersham Biosciences) using a 473 nm excitation laser/long pass blue (LPB) emission filter for fluorescein images and 532 nm excitation laser/long pass green (LPG) emission filter for rhodamine images. If required, the gels were stained with ethidium bromide or SYBR Green II solution for 15 min, then visualized on a UV transilluminator or the Typhoon gel scanner. Some gels show slight laddering of the RNA which likely results from transcription errors.

HPLC analysis of labeling reactions

Analytical high performance liquid chromatography (HPLC) was performed on an Agilent Technologies 1200 series HPLC system equipped with anion-exchange DNAPac PA 100 column (4 x 250 mm, ThermoScientific). The HPLC solvents used were: 10% CH3CN in H2O (buffer A) and 10% CH3CN in 0.25 M Tris buffer, 0.375 M NaClO4, pH 8 (buffer B). The percentage of buffer B in buffer A was ramped from 20–40% over 20 min with a flow rate of 1.0 mL min-1. Elution from the HPLC column was monitored at both 260 nm and 495/555 nm (fluorescein/TMRIA). Peak areas for the RNA at each of these wavelengths were obtained by manual integration.

Hydrolysis to determine reaction site

Ribozyme 5FR1 was reacted with FIA in a volume of 250 µL, in a solution containing 4 µM transcribed RNA and 100 µM FIA in PBS (Aldrich). The reaction was incubated at 37 °C for 30 min, then quenched by the addition of 100 mM β-mercaptoethanol, 5 mM ETDA, 0.3 M NaCl. After 15 min of shaking at room temperature, the reaction was precipitated using sodium acetate/ethanol and stored at -80 °C overnight. This was then centrifuged at 17,000 x g for 30 min at 4 °C and the pellet was washed with 200 µL of 70% ethanol, air dried, and resuspended in 20 µL of RNAse free water. Approximately 1-50 pmol of this RNA was end labeled on the 3' terminus with biotin using the RNA 3' Biotinylation Kit (ThermoScientific).

The biotinylated RNA (0.1-3 µg) was mixed with 3 µg of yeast tRNA in a volume to not exceed 5 µL. 1X Alkaline Hydrolysis Buffer (50 mM Sodium Carbonate, 1mM EDTA, pH 9.2) was added to give a final volume of 15 µL. This mixture was divided into three tubes labeled 1-3, and heated to 95 °C to hydrolyze RNA. The tubes were moved to ice after 30 s, 1 min, and 2 min, and then precipitated with sodium acetate/ethanol for 1 hr at -80 °C. Tubes were centrifuged at 17,000 x g for 15 min at 4 °C and the pellet was washed with 200 µL of 70% ethanol, air dried, and resuspended in 40 µL of 2X Binding and washing (B&W) buffer (10 mM Tris, 1 mM EDTA, 2M NaCl). As a control, 40 µL B&W buffer was added to an equivalent amount (0.1-3 µg) biotinylated ribozyme (no hydrolysis) to serve as an additional size marker and to verify the success of biotin pull down on the magnetic beads. M-270 streptavidin beads (400 µL, Life Technologies) were washed three times with 400 µL of B&W buffer. After each wash, the beads were separated from

the supernatant using a magnetic separation stand (DynaMag-Spin, Life Technologies). The beads were aliquoted into four tubes and resuspended in 40 µL of the solution containing resuspended fragments from the different hydrolysis times or the control biotinylated ribozyme in B&W buffer. The suspension was gently agitated on a nutating mixer for 20 min at room temperature. The beads were washed three times with 40 µL of B&W buffer, and then resuspended in 40 µL of Elution Buffer (40 mM Tris, 10 mM EDTA, 3.5 M urea, 0.02% tween, pH 7.2). The solutions were heated to 90 °C for 5 min to break the biotin streptavidin bond. The beads were separated and the supernatant was removed immediately and placed at -20 °C for further analysis. The RNA eluted from the beads was loaded onto an 8% (Tris/borate/EDTA) polyacrylamide gel with 8 M urea along with the Decade Marker (Life Technologies) ladder to determine the size of hydrolysis fragments. The gels were run at 200 V for 1 hr and imaged on a Typhoon 9500 fluorescence laser scanner (GE healthcare) using a 473 nm laser and long pass blue (LPB) emission filter. The gels were then stained with a 1X SYBR Green II (Life Technologies) solution for 20 min and then imaged using the 473 nm laser and a LPB emission filter.

Ribozyme function in crosslinking conditions

A total amount of 4 µM RNA was incubated with 1% or 0.1% formaldehyde for 10 min in PBS at room temp and quenched with 250 mM glycine. 400 nM RNA was subsequently labeled with 100 µM FIA in PBS, 30 min, 37°C.

Construction of ribozyme fusions to mRNA

To assess the labeling ability of ribozymes in the context of a longer RNA sequence, we constructed vectors containing the coding sequence of the mCherry fluorescent protein with and without the R1 and 5FR1 ribozymes inserted into the vector 3'UTR. The parent plasmid was a vector designed to express green fluorescent protein (GFP) in Drosophila cells (1), in which we replaced the GFP coding sequence with that of mCherry, using the Kpn1 and Xho1 restriction sites. We then further inserted the R1 ribozyme sequence between the Xho1 and BstB1 restrictions sites or the 5FR1 sequence between the Mlu1 and Sac1 sites; all of these sites are contained within the vector SV40 3'UTR. To synthesize RNA from these vectors for labeling tests, we used primers (forward: TTAATACGACTCACTATAGGGAGACCCCGAGGGCTTCAAGTGGGAGC, which also contains the T7 RNA polymerase binding site, and reverse: ACTGCATTCTAGTTGTGGT) to amplify regions containing 432 nucleotides of the mCherry coding sequence followed by the 3'UTR with and without the ribozyme sequences (total RNA lengths were therefore 659 nucleotides for the control, 800 nucleotides for the R1-containing RNA and 724 nucleotides for the 5FR1-containing RNA). These PCR products were subjected to in vitro transcription and labeling as for the short ribozymes, with exceptions noted in the text.

References

(1) Bashirullah, A.; Cooperstock, R. L.; Lipshitz, H. D. RNA Localization in Development. *Annu. Rev. Biochem.* **1998**, *67,* 335-394.

(2) Tolino, M.; Köhrmann, M.; Kiebler, M. A. RNA-Binding Proteins Involved in RNA Localization and Their Implications in Neuronal Diseases. *Eur. J. Neurosci.* **2012,** *35*, 1818-1836.

(3) 6.Cooper, T. A.; Wan, L.; Dreyfuss, G. RNA and Disease. *Cell* **2009,** *136*, 777-793.

(4) Liu-Yesucevitz, L.; Bassell, G. J.; Gitler, A. D.; Hart, A. C.; Klann, E.; Richter, J. D.; Warren, S. T.; Wolozin, B. Local RNA Translation at the Synapse and in Disease. *J. Neurosci.* **2011,** *31*, 16086-16093.

(5) Ashley, C. T.; Wilkinson, K. D.; Reines, D.; Warren, S. T. FMR1 Protein: Conserved RNP Family Domains and Selective RNA Binding. *Science* **1993,** *262*, 563-566.

(6) Bassell, G. J.; Warren, S. T. Fragile X Syndrome: Loss of Local Mrna Regulation Alters Synaptic Development and Function. *Cell* **2008,** *60*, 201-214.

(7) Bassell, G. J.; Kelic, S. Binding Proteins for mRNALocalization and Local Translation, and Their Dysfunction in Genetic Neurological Disease. *Curr. Opin. Neurobiol.* **2004,** *14*, 574-581.

(8) Lecuyer, K. A.; Behlen, L. S.; Uhlenbeck, O. C. Mutagenesis of a Stacking Contact in the MS2 Coat Protein-RNA Complex. *EMBO J.* **1996,** *15*, 6847-6853

(9) Armitage, B. A. Imaging of RNA in Live Cells. *Curr. Opin. Chem. Bio.* **2011**, *15,* 806-812.

(10) Tyagi, S. Imaging Intracellular RNA Distribution and Dynamics in Living Cells. *Nat. Methods.* **2009**, *6,* 331-338.

(11) Babendure, J. R.; Adams, S. R.; Tsien, R. Y. Aptamers Switch on Fluorescence of Triphenylmethane Dyes. *J. Am. Chem. Soc.* **2003,** *125*, 14716-14717.

(12) Paige, J. S.; Wu, K. Y.; Jaffrey, S. R. RNA Mimics of Green Fluoresencent Protein. *Science.* **2011,** *333,* 642-646.

(13) Jäschke, A. Artificial Ribozymes and Deoxyribozymes. *Curr. Opin. Struct. Bio.* **2001,** *11,* 321-326.

(14) Mcdonald, R. I.; Guilinger, J. P.; Mukherji, S.; Curtis, E. A.; Lee, W. I.; Liu, D. R. Electrophilic Activity-Based RNA Probes Reveal a Naturally Occurring Self-Alkylating Ribozyme That Can Be Used for Selective RNA Labeling. *Nat. Chem. Biol.* **2014,** *10*, 1049-1054.

(15) Gaietta, G.; Deernick, T. J.; Adams, S. R.; Bouwer, J.; Tour, O.; Laird, D. W.; Sosinksy, G. E.; Tsien, R. Y.; Ellisman, M. H. Multicolor and Electron Microscopic Imaging of Connexin Trafficking. *Science.* **2002,** *296,* 503-507.

(16) Niranjanakumari, S.; Lasda, E.; Brazas, R.; Garcia-Blanco, M. A. Reversible Cross-Linking Combined With Immunoprecipitiation to Study RNA-Protein Interactions *in Vivo. Methods.* **2002,** *26,*

(17) Slobodin, B.; Gerst, J. E. A Novel Mrna Affinity Purification Techniques for the Identification Of Interacting Proteins and Transcripts In Ribonucleoprotein Complexes. *RNA.* **2010**, *16*, 2277-2290.

(18) Sharma, A. K.; Plant, J. J.; Rangel, A. E.; Khoe, K. N.; Anamisis, A. J.; Hollien, J.; Heemstra, J. M. Fluorescent RNA Labeling Using Self-Alkylating Ribozymes. *ACS Chem. Biol.* **2014,** *9*, 1680-1684.

(19) Tuerk, C.; Gold, L., Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands To Bacteriophage T4 DNA Polymerase. *Science (New York, N.Y.)* **1990,** *249* (4968), 505-10.

(20) Ellington, A. D.; Szostak, J. W., In Vitro Selection of RNA Molecules That Bind Specific Ligands. *Nature* **1990,** *346* (6287), 818-22.

(21) Robertson, D. L.; Joyce, G. F., Selection In Vitro Of An RNA Enzyme That Specifically Cleaves Single-Stranded DNA. *Nature* **1990,** *344* (6265), 467-8.

(22) Wilson, C.; Szostak, J. W. In Vitro Evolution of A Self-Alkylating Ribozyme. *Nature.* **1995,** *374,* 777-781.

(23) Debets, M. F.; Van Berkel, S. S.; Schoffelen, S.; Rutjes, F. P. J. T.; Van Hest, J. C. M.; Van Delft, F. L. Aza-Dibenzocyclooctynes for Fast and Efficient Enzyme Pegylation Via Copper-Free (3+2) Cycloaddition. *Chem. Commun.* **2010**, 97-99.

(24) Silverman, S. K. Catalytic DNA (Deoxyribozymes) For Synthetic Applications-Current Abililites and Future Prospects. *Chem. Commun*. **2008**, 3467-3485.

(25) Yao, J. Z.; Uttamapinant, C.; Poloukhtine, A. A.; Baskin, J. M.; Codelli, J. A.; Sletten, E. M.; Bertozzi, C. R.; Popik, V. V.; Ting, A. Y. Fluorophore Targeting to Cellular Proteins Via Enzyme-Mediated Azide Ligation and Strain-Promoted Cycloaddition. *J. Am. Chem. Soc.* **2012,** *134*, 3720-

3728.

(26) Heemstra, J. M.; Moore, J. S. Enhanced Methylation Rate Within a Foldable Molecular Receptor. *J. Org. Chem*. **2004** 69, 9234-9237.

CHAPTER 3

EVOLUTION OF ARTIFICIAL GENETIC POLYMERS CAPABLE OF SMALL-MOLECULE RECOGNITION

Introduction

Life is controlled on a molecular level by the unparalleled ability of nucleic acids to store and process genetic information. This has inspired significant research effort focused on understanding how nucleic acids evolve and function, and how these capabilities can be harnessed for applications in biotechnology. *In vitro* evolution has proven to be a powerful tool for exploring these questions, and simultaneously provides a method for generating nucleic acid polymers having novel molecular recognition functionality. These polymers, known as aptamers, are single stranded oligonucleotides that can be evolved through a process called SELEX (Systematic Evolution of Ligands by EXponential enrichment) to bind to a wide variety of target molecules with high affinity and specificity.[1-3] Since their discovery, aptamers have shown promise in a variety of applications including therapeutics, affinity purification, and biosensing, where antibodies have long been considered the gold standard. Aptamers provide numerous advantages over antibodies including increased thermal stability and tolerance to surfactants. Additionally, aptamers

can be chemically synthesized, enabling more cost-effective production and allowing for the incorporation of a wide range of labels and nonnatural functional groups.[4] However, aptamers comprised of native DNA and RNA retain some limitations due to their instability in biological fluids and cellular environments.[5]

Recently, there has been a growing interest in backbone-modified nucleic acids, or xeno nucleic acids (XNA), as these polymers are not readily recognized and degraded by nucleases, and thus are well-suited for *in vivo* applications.[6,7] Arguably, the most famous XNA aptamer to date is Macugen, a backbone-modified aptamer that binds vascular endothelial growth factor isoform $VEGF_{165}$ to prevent the overgrowth of blood vessels associated with macular degeneration.[8] In the development of Macugen, aptamer sequences were evolved using native RNA, then the native nucleotides were systematically replaced with 2'-OMe or 2'-F nucleotides to generate a biostable XNA aptamer that retained the necessary binding affinity to $VEGF_{165}$. While this example supports the utility of XNA aptamers, the postselection modification strategy is expensive and time-consuming. In contrast, the direct selection of aptamers from XNA libraries offers a more robust approach. Fortuitously, many XNAs are capable of Watson-Crick base pairing with DNA and RNA, which has provided encouragement for their use as alternative genetic systems.[9] Recent efforts in this field have generated replication technologies that permit information transfer between DNA and different XNAs, establishing the capacity for *in vitro* evolution using these artificial genetic systems.[10] While the direct selection of XNA aptamers holds significant promise for generating biopolymers having novel functions, these artificial genetic systems can present challenges for

in vitro selection experiments. An initial challenge is the limited availability of the building blocks and polymerases needed to create XNA polymers. For many nonnatural nucleic acids, the phosphoramidites and triphosphates needed for solid phase and enzymatic synthesis, respectively, are neither commercially available nor easily synthesized. Additionally, while recent advances in enzyme engineering have provided polymerases capable of XNA transcription and reverse transcription, they are often characterized by lower fidelity and processivity.[11] Consequently, transcription efficiency is often reduced, which can lead to undersampling of library diversity and introduce sequence bias that negatively impacts the efficacy of the selection.[12] Additionally, the difficulty of generating and utilizing XNA primers has limited transcription to techniques where the polymer is extended from a DNA primer to yield a DNA-XNA chimera. While aptamer selections using these chimeric libraries have been successful, the DNA portion is often involved in target binding and cannot be completely removed or converted to XNA postselection. Under biological conditions, this vital region of the aptamer may then be degraded, compromising the binding affinity of the aptamer. To date, direct selection using chimeric DNA-XNA libraries has provided hexitol nucleic acid (HNA) aptamers to hen-egg lysozyme and HIV trans-activating response RNA element,[11] a (3',2')-α-ı-threose nucleic acid (TNA) aptamer to human thrombin,[13] a 2'-deoxy-2'-fluoroarabinonucleotide (FANA) aptamer to HIV-1 reverse transcriptase,[14] and a 2'-fluoro-modified aptamer for human neutrophil elastase (HNE).[15]

These examples demonstrate that despite the challenges associated with *in vitro* selection using XNA, aptamers comprising a variety of backbone structures

can be generated for protein targets. However, successful selection of an XNA aptamer to a small-molecule target has yet to be reported. It is widely accepted that small molecules are more difficult targets for aptamer selection compared to proteins. In general, small molecules present significantly less surface area and fewer functional groups for the aptamer to interact with to gain favorable binding interactions.[16] These factors also complicate the partitioning of bound from unbound sequences, as selection methods typically rely on immobilization of the target molecule. This inherently compromises one functional group on the target, which carries a significantly greater risk in the case of small-molecule targets.[17] Furthermore, the relatively small change in size between free XNAs and those bound to the small-molecule target presents complications in the screening and characterization of putative aptamer sequences. Given these challenges, it is not surprising that only ~25% of all reported selection experiments are for small-molecule targets.[17] However, many small molecules including toxins and biomarkers play key roles in biology and the environment, therefore developing receptors that can sensitively and selectively detect these targets is of high utility.

Given the exciting capabilities provided by XNA in the context of protein recognition, we sought to extend these by selecting the first XNA aptamer to a small-molecule target. As our target molecule, we chose ochratoxin A (OTA), a food-contaminating mycotoxin which causes nephrotoxic effects when consumed (Figure 3.1a).[18] OTA has a negative immunogenic response, which makes it an interesting target for aptamer development, as antibodies are difficult to generate. In addition, several DNA aptamers have been evolved for OTA, enabling comparis-

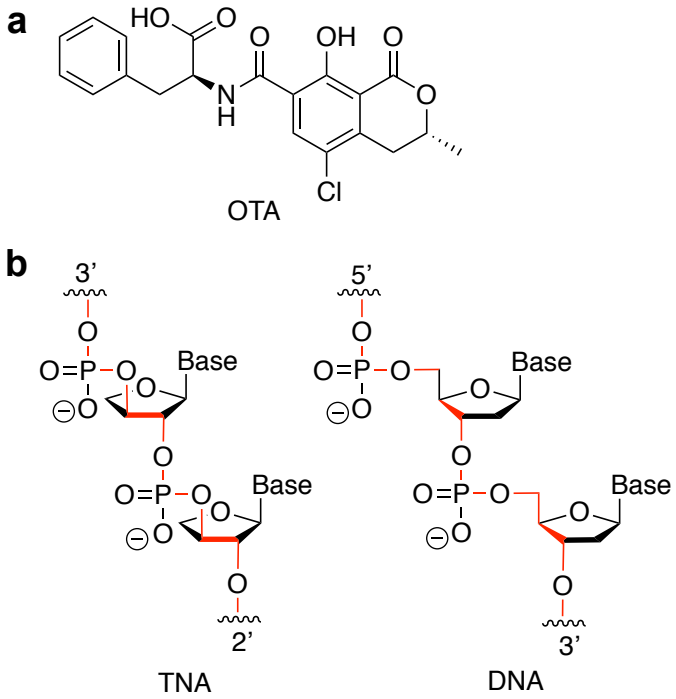


Figure 3.1. Relevant chemical structures of (a) ochratoxin A (OTA) and (b) TNA and DNA polymers. TNA has a five-atom backbone perodicity and phosphodiester linkages at the 3' and 2' positions of the furanose ring

on and functional validation of our newly generated XNA aptamers. We anticipated that TNA would be well-suited for *in vitro* evolution because of the availability of enzymes that support efficient *in vitro* transcription and reverse transcription, as well as its superb stability in the presence of nucleases (Figure 3.1b).[19,20] Given the increased difficulty of generating aptamers to small-molecule targets, we introduced new capabilities in XNA SELEX aimed at increasing enrichment efficiency and ensuring binding of the free target molecule. Central among these was the use of a TNA primer to give a fully nonnative library (Figure 3.2). Additionally, the TNA primer allows us to digest DNA contaminants that may hinder the enrichment of high affinity TNA aptamers. Small-molecule selections typically take 10 ± 6 round to achieve sufficient enrichment,[21] and we anticipate the ability to perform the

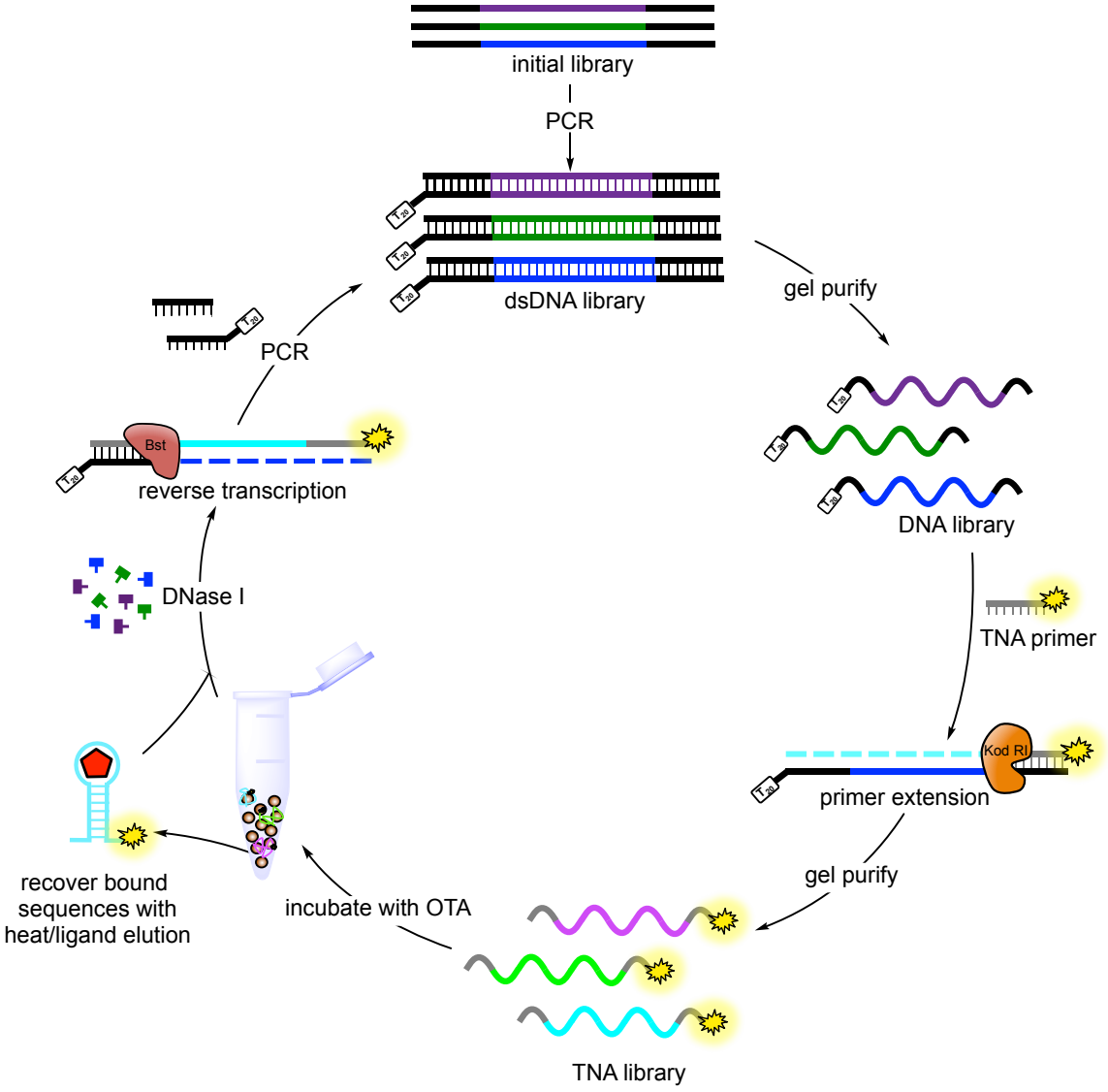


Figure 3.2. TNA SELEX to generate OTA-binding aptamers. The initial ssDNA library is amplified using a forward primer modified with a PEG spacer and polyT tail to enable separation and recovery by denaturing PAGE. The PEGylated DNA template is then annealed to the FAM-labeled TNA primer and extended using KOD RI polymerase to generate the TNA library for each selection round. The TNA library is incubated with OTA-functionalized magnetic beads, and bound sequences recovered by either heat (rounds 1-4) or ligand elution (rounds 5-9). These sequences are then treated with DNase I to digest any remaining DNA template. The TNA is then reverse transcribed back into DNA using Bst DNA polymerase and PCR amplified for the next round of selection.

DNase digestion will improve the efficiency of our selection. Using our improved SELEX protocol, we generated seven sequences having affinity for OTA. Analysis using microscale thermophoresis (MST) revealed that five of these sequences have K_D values in the nM range, with the best being 92 ± 2 nM. Importantly, the use of MST to quantify the binding affinity of our aptamer sequences confirms that despite the use of immobilized target during the selection experiments, our aptamers are capable of binding to the free target in solution. Using secondary structure predictions, we were able to minimize two of our aptamer sequences to a functional core <35 nt, a length that enables chemical synthesis using phosphoramidite monomers available in our lab. Aptamer A04T.2, our shortest sequence at 31 nt, was found to have a K_D of 71 ± 8 nM. This represents tighter binding than even the best DNA aptamer to OTA, and is the first example of a TNA aptamer to be validated by chemical synthesis. Using this truncated aptamer sequence, we demonstrate the impressive resistance of TNA to degradation by DNase I, human liver microsomes, and human blood serum. Furthermore, we highlight the selectivity of our TNA aptamer in that it retains binding affinity for OTA in the presence of 50% human serum. Together, these experiments demonstrate that XNA aptamers having high affinity and selectivity can be generated for small-molecule targets, and provide a toxin-binding aptamer for potential use in diagnostics and therapeutics.

Materials and methods

Oligonucleotides synthesis

TNA triphosphates used for primer extension reactions were synthesized as described previously.[22,23] TNA phosphoramidites were synthesized by unpublished methods developed in our lab for solid phase synthesis of TNA polymers. All DNA and chemically synthesizable TNA sequences were purchased from the University of Utah DNA/Peptide Synthesis Core Facility. All oligonucleotides were purified by denaturing PAGE and the desired band was excised and incubated in crush and soak buffer (300 mM ammonium acetate, 1 mM EDTA, pH 8.0) at 40°C overnight. Samples were separated from gel pieces and buffer exchanged into water using Amicon Ultra-0.5 Centrifugal Unit with Ultracel 10 membrane (EMD Millipore).

Attachment of ochratoxin A to magnetic beads

300 µg of OTA (Enzo Life Sciences) was suspended in 500 µl of 0.1 M MES buffer, pH 5. After equilibration to room temperature, 0.4 mg EDC and 1.1 mg sulfo-NHS were added to the OTAA solution and allowed to react for 15 min. 1.2 µl of 2-mercaptoethanol was added to the reaction to quench the remaining EDC. 500 µl of M-270 amine beads (Life Technologies) were washed three times with 1 mL 1x PBS. After each wash, the beads were separated from the supernatant using a magnetic separation stand (DynaMag-Spin, Life Technologies). The beads were re-suspended in 500 µl 1x PBS and added to the solution of OTA. After allowing the reaction to proceed for 1 h, hydroxylamine was added to a final concentration of 10 mM to quench the remaining NHS ester. The functionalized beads were washed three times with 1 mL binding buffer and re-suspended in 500 µL binding

buffer for storage at 4°C. Successful attachment of OTA was validated by monitoring fluorescence intensity (λ_{ex} = 380 nm, λ_{em} = 444 nm).

PCR amplification of ssDNA library

A single stranded DNA library having the sequence 5'-GCGCAT-ACCAGCTTATTCAATT-N_{50}-AGATAGTAAGTGCAATCTCGGC-3' (20 pmol) was amplified in 50 µL PCR reactions containing 0.2 µM template, 0.2 µM each primers, 200 µM dNTPs, and 2.5 U Taq DNA polymerase in 1x Thermopol buffer (20 mM Tris-HCl, 10 mM $(NH_4)_2SO_4$, 10 mM KCl. 2 mM $MgSO_4$, 0.1% Triton X-100, pH 8.8, New England Biolabs). The forward primer had the sequence 5'-TTTTTTTTTTTTTTTTTTTT/Sp9/GCGCATACCAGCTTATTCAATT-3' and the reverse primer had the sequence 5′-/FAM/GCCGAGATTGCACTTACTATCT-3′. PCR was carried using a program with an initial denaturation at 95°C for 3 min, 20 cycles of (95°C for 30 s, 54 °C for 30 s, and 72 ^{o}C for 20 s), and a final extension 72 ^{o}C for 2 min. The amplified double stranded DNA was purified using a PCR cleanup column (Qiagen) and the PEG-functionalized strand was separated from the FAM labeled strand on a denaturing 10% polyacrylamide gel. The gel was stained with ethidium bromide and the desired band was excised and incubated in crush and soak at 40°C overnight. Samples were recovered using size-exclusion centrifugal units as described previously.

Transcription of TNA library

The TNA library was generated by primer extension using 75 pmol of purified PEGylated ssDNA library. Primer extension reactions were performed in 10 µL

volumes containing 100 μM tNTPs with 500 nM primer-template complex. The FAM labeled TNA primer having sequence 3′-/FAM/GCCGAGATTGCACTTACTA-TCT-2′ was annealed to the PEGylated ssDNA in 1x Thermopol buffer by heating to 95 °C for 5 min, cooling to 4°C for 10 min, and incubating at 50°C until addition of the enzyme. KOD RI TNA polymerase was pretreated with $MnCl_2$ (1 mM) in a 1:1 v:v ratio and added to the mixture. tNTPs were added last at a final concentration of 100 μM, and the reactions were incubated at 50 °C for 10 h. Primer extension reactions were quenched using TNA stop buffer (50% (w/v) urea, 10 mM EDTA) and denatured at 95°C for 5 min. Reactions were pooled and purified on a denaturing 10% polyacrylamide gel as described above.

TNA SELEX

The purified TNA library (~10^{13} sequences) was folded in 1x selection buffer (100 mM NaCl, 2 mM $MgCl_2$, 5 mM KCl, 1 mM $CaCl_2$, 0.02% Tween 20, 20 mM Tris-HCl, pH 7.6) in a total volume of 500 μL by heating to 95 °C for 5 min, cooling at 4 °C for 10 min, and a final incubation at 25 °C for 10 min. 70 μL (1.8 x10^8) of OTA functionalized beads were washed three times with selection buffer and then re-suspended in the 500 μL of folded TNA library. After incubation for 1 h at 25 °C, the unbound sequences were removed by washing three times with 100 μL binding buffer, using the magnetic stand to separate the supernatant from the beads. The beads were re-suspended in 100 μL binding buffer and the bound sequences were recovered by elution with heat at 95 °C for 5 min. The heat elution step was repeated twice. Starting at round 5, bound sequences were recovered by target elution using excess OTA: after the three washes, the beads were resuspended in

500 µL of OTA (100 µM) for 1 h. After incubation, the supernatant containing the bound sequences was separated from the beads and concentrated using a 10 kDa spin filter. The amount of recovered TNA was quantified by comparison to a calibration curve created using a FAM-labeled TNA standard of the same size. The standard and SELEX output were analyzed for fluorescence intensity using a Bio-Tek Synergy 2 multimode plate reader.

DNase digestion to remove DNA contaminants

The recovered TNA was divided into 10 µL portions and 1.5 U DNase I (New England Biolabs) was added to each aliquot. The solutions were incubated for 45 min at 37°C. The DNase was denatured for 10 min at 75°C and the TNA solutions left on ice until further use.

TNA Reverse transcription and regeneration of ssDNA library

Reverse transcription reactions were performed in a 20 µL reaction volume. 7.4 µL of the TNA/DNase mixture was combined with dNTPs (500 µM) and $MgCl_2$ (3 mM) in 1x Thermopol buffer. The recovered TNA was annealed with 5 pmol of unmodified forward primer in 1x Thermopol buffer as previously described. 1.6 U of Bst DNA polymerase I large fragment (New England Biolabs) was added to the primer-template complex and the reaction incubated at 55 °C for 3.5 h. The RT reactions were divided into 5 µL aliquots and PCR amplified as previously described for the DNA library using unmodified versions of the forward and reverse PCR primers and an annealing temperature of 50°C. This material was purified

using MinElute PCR clean up columns and reamplified using the PEG-forward and FAM-reverse PCR primers to enable strand separation. The recovered PEG-functionalized DNA template was then used to initiate transcription of the TNA library for the subsequent round of selection.

Initial screening of putative aptamer sequences

After 9 rounds of selection, cDNA from rounds 7, 8, and 9 were amplified in separate PCR reactions using unmodified primers and cloned into a pCR™4-TOPO® TA vector using a TOPO TA cloning kit for sequencing (Life Technologies). The vector was amplified in *E. coli* and plasmid DNA was recovered from 25 colonies per selection round using a Miniprep Kit (Quiagen). These DNA plasmids were sequenced at the University of Utah DNA sequencing core facility. Sequences were aligned using Multalin and secondary structure analysis was performed using the DNA folding form in Mfold (Figure 3.3).[24] DNA templates corresponding to the 10 potential TNA aptamer sequences were ordered having a PEG-T_{20} modification, enabling them to be directly used for TNA synthesis. The PEGylated templates were annealed to the FAM-labeled TNA reverse primer as previously described to generate TNA for each potential aptamer.

MST analysis of binding interactions with OTA

MST experiments were performed by 2bind GmbH. Briefly, a serial dilution of OTA in the appropriate binding buffer was carried out to provide solutions having a range of concentrations between 1.832 nM and 60 µM. 5 µL of each solution was mixed with 5 µL of the FAM-labeled aptamer, which was held at a constant concent-

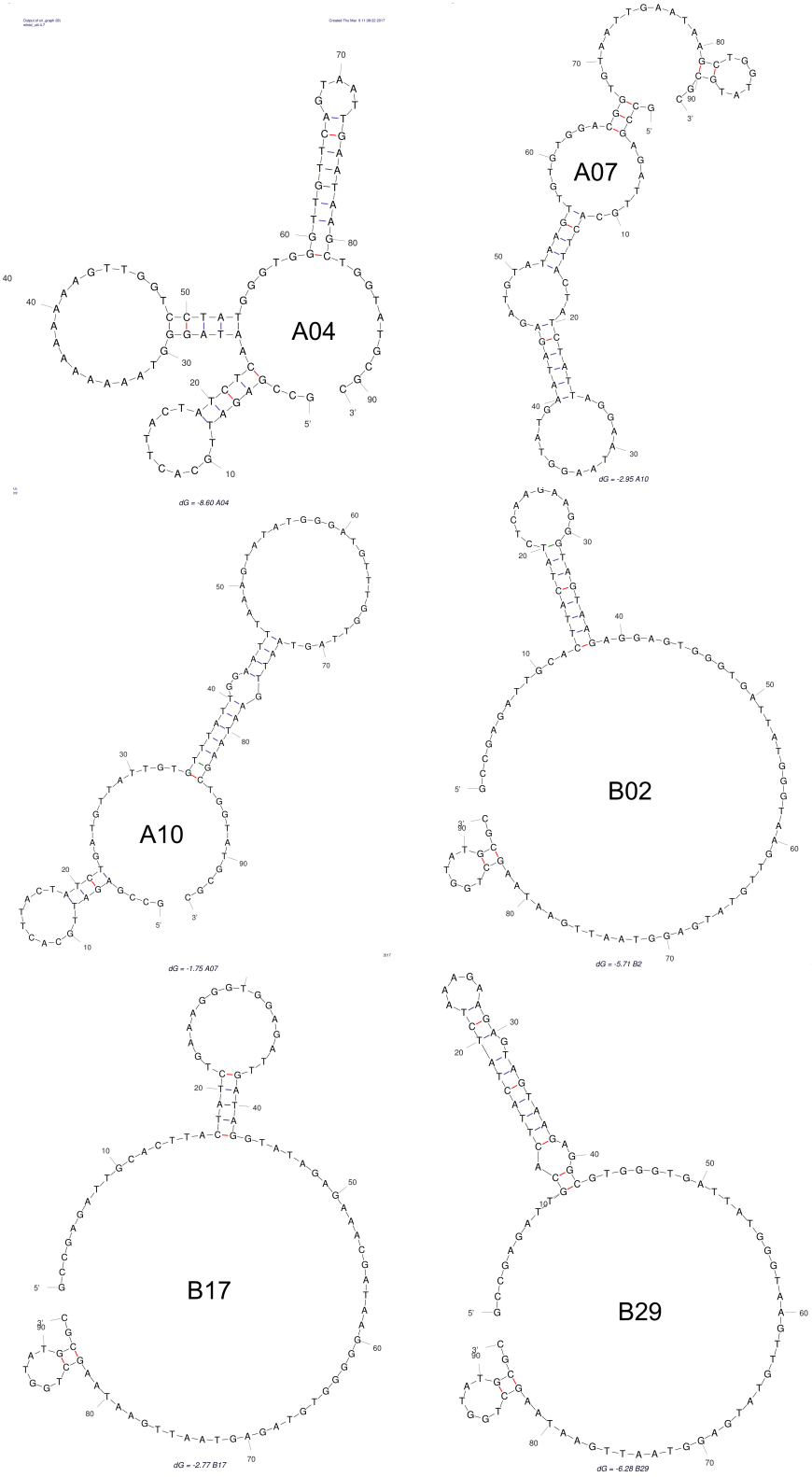


Figure 3.3. Secondary structure analysis of putative aptamer sequences using Mfold. For these predictions, the folding temperature was set to 25 °C and the ionic conditions were $[Na^{+}]$= 100 mM and $[Mg^{2+}]$= 2 mM.

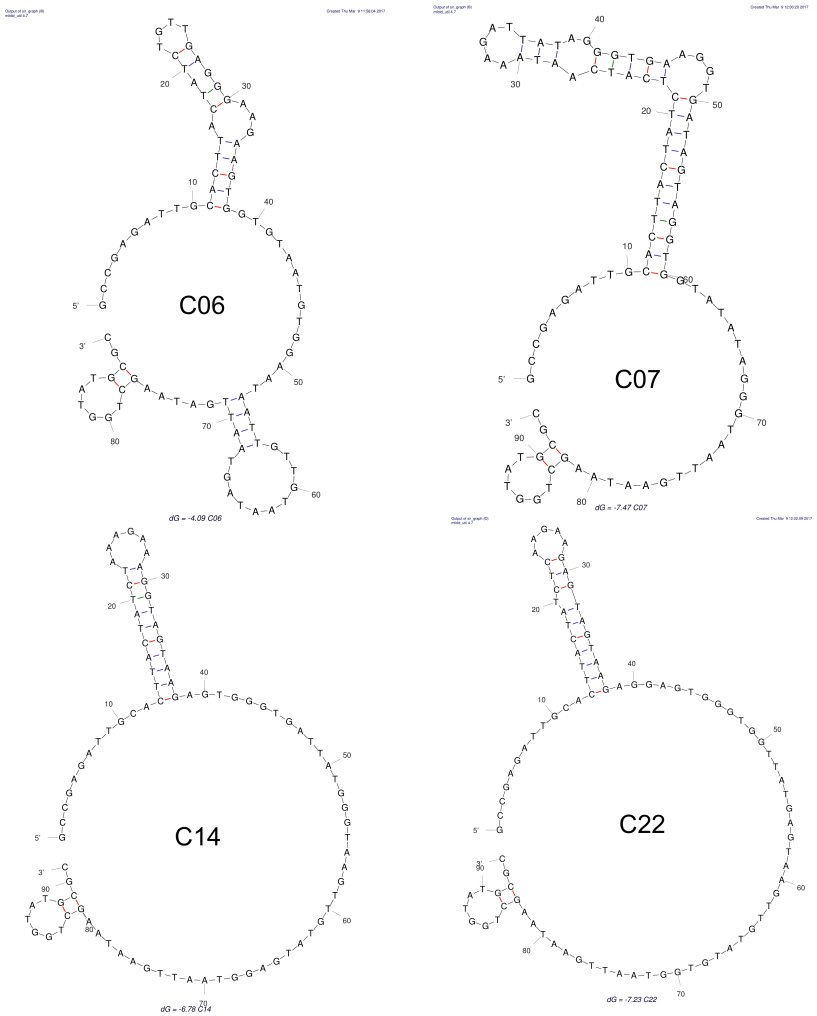


Figure 3.3. continued.

ration of 10 nM. The final concentrations of OTA in each capillary ranged from 916 pM to 30 μM. Each sample was analyzed on a Monolith NT.115 Pico at 25 °C, with 40% LED power and 80% laser power. Data were fitted using Origin Lab analysis software to determine the aptamer K_D values. To calculate the fraction bound (FB), raw fluorescence data acquired for each concentration point during the MST experiments were averaged (avg) for technical repeats and plotted versus concentration. A dose response fit was performed for each data set and the maximum(complexed) and minimum (free) values for each curve were recorded. These values were then used to determine the normalized fraction bound for each aptamer using the equation FB=(avg–free)/(complexed–free). A Hill fit can then be performed to determine K_D.

Sequence minimization and evaluation

To qualitatively assess binding of enzymatically-synthesized truncated sequences A04T and B29T, OTA functionalized magnetic beads were washed three times with binding buffer and separated into aliquots of 20 μL (2×10^7) beads. 4 pmol of each aptamer sequence was folded in binding buffer and then added to an aliquot of beads and incubated for 1 h at 25 °C. Unbound sequences were removed using three washes with binding buffer and then the bound fraction was eluted into 100 μL buffer by heating to 95 °C for 10 min. Fluorescence was analyzed using a plate reader as described above to determine the percentage of aptamer bound to the beads. Binding affinities for truncated aptamer sequences were determined by MST as described above.

TNA stability assay

FAM-labelled DNA aptamer A08 and TNA aptamer A04T.2 were prepared in three nuclease challenge conditions: 1.5 U DNase I (RNase-free, New England Biolabs) in 1x DNase I reaction buffer (10 mM Tris-HCl, 2.5 mM $MgCl_2$, 0.5 mM $CaCl_2$, pH 7.6), 50% v/v human blood serum (normal pool, Fisher BioReagents) in 1x Dulbecco's PBS, 0.5 mg/mL pooled human liver microsomes (XenoTech) in 1x selection buffer. TNA and DNA aptamer (4 pmol) were each folded in the respective buffer for each condition and DNase I, human blood serum, and human liver microsomes were added to the desired concentration. These samples were then incubated at 37 °C for three days, separated on a denaturing 10% polyacrylamide gel, and analyzed using a GE Typhoon laser gel scanner. TNA and DNA aptamers folded in selection buffer were used for comparison in these experiments. The samples were also assessed for their ability to retain OTA binding using a bead-binding assay. Samples were prepared as described above and incubated with an aliquot of 20 µL (2×10^7) OTA-functionalized beads for 30 min at 25 °C. Bound sequences were recovered after three washes using heat elution and quantified by fluorescence. For the MST experiments, Cy5-labeled DNA aptamer A08 and TNA aptamer A04T.2 were folded in 1x Dulbecco's PBS at a final concentration of 100 nM. An equal volume of human blood serum was added and the mixture was incubated at 37 °C for either 3 or 7 d. Binding affinities were measured for the aptamers in serum using the MST method described above.

Results and discussion

Transcription of a fully XNA library using a TNA primer

Typical XNA SELEX experiments utilize a DNA primer, upon which an XNA library is synthesized using enzyme-mediated primer extension. This can create challenges postselection, as removing the primer sequence or converting it to XNA may lead to a decrease in binding affinity of the aptamer. Additionally, because the engineered polymerases used to transcribe and reverse transcribe XNA still have a strong bias toward DNA, small amounts of DNA contamination can undermine the enrichment of aptamer sequences. We anticipated that this would be especially problematic in the selection of aptamers for small-molecule targets, but recognized that if we could use a TNA primer to provide fully TNA library, then DNase digestion could be employed to eliminate unwanted DNA contamination during the selection rounds. The gel image in Figure 3.4 shows that in the absence of DNase, residual DNA template amplifies efficiently, evident by the comparable fluorescent band in the positive and negative reactions. We know that any product being amplified in the negative reaction is due to DNA contamination as TNA cannot be direct amplified by PCR. With the use of DNase we see selective amplification of TNA that has been reverse transcribed which is supported by the higher fluorescence signal in comparison to the control. We observe that excess DNase seems to slightly inhibit PCR and concluded we would use1.5 U DNase as it provided the largest selectivity between the positive and negative reactions

Thus, before beginning our aptamer selection, we performed initial experments to determine whether TNA transcription would be possible using a TNA pri

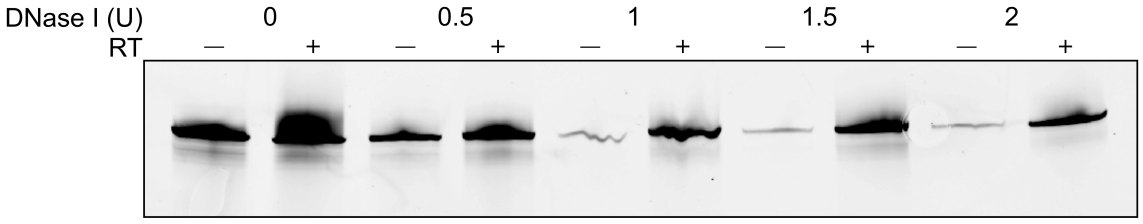


Figure 3.4. Optimization of DNase I digestion. 200 pmol of TNA library was incubated with designated amount of DNase I for 45 min at 37 °C and heat inactivated at 75 °C for 10 min. Each sample was then divided in half and used in a positive (with Bst DNA polymerase) or negative (no polymerase) reverse transcription reaction).

mer. Given the significant structural differences between TNA and DNA/RNA, a key challenge in TNA SELEX is the availability of polymerases capable of transcribing DNA into TNA and reverse transcribing TNA into DNA. The Chaput lab has recently made a number of exciting advances in this field, and while completely unbiased libraries still remain somewhat problematic due to G:G mispairing during TNA transcription, libraries having all four letters can now be transcribed efficiently.[25-27] In exploring transcription from a TNA primer, we anticipated that a library having an N_{50} random region composed of 38.5% each of A and T and 11.5% each of C and G would allow for sufficient diversity and secondary structural elements, while also permitting efficient and faithful transcription. To evaluate the fidelity of transcribing this four-letter library, we first needed to optimize conditions for the primer extension reaction (Figure 3.5). An initial transcription reaction using our SELEX library and TNA primer revealed that the original incubation temperature for the primer extension reaction (55 °C) yielded no full-length product. We performed a temperature screen, and found that an incubation temperature of 50 °C and time of 10 h was optimal for reactions using the TNA primer. In Figure 3.6

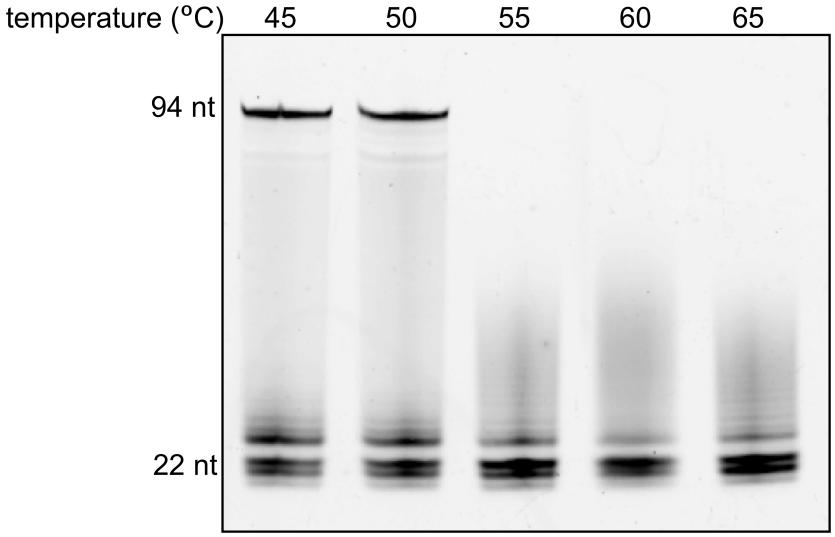


Figure 3.5 Temperature screen of TNA transcription using a TNA primer. The TNA primer was annealed to the SELEX library and incubated in a primer extension reaction for 10 h at the designated temperatures. No product was observed at temperatures above 55°C.

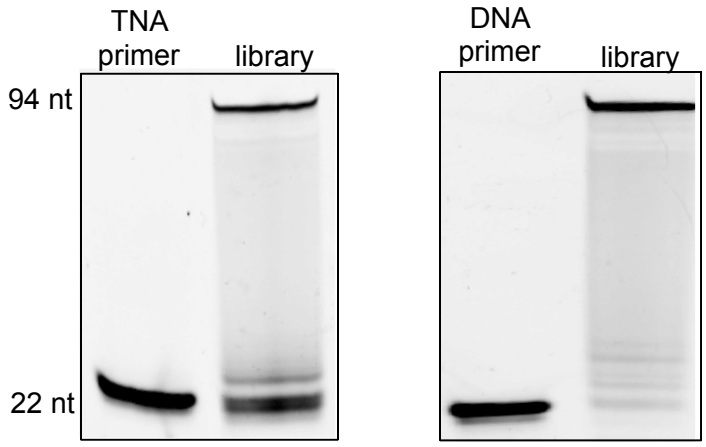


Figure 3.6. Comparison of TNA transcription efficiency of our SELEX library using a DNA versus TNA primer. 500 nM primer and 11.5% G,C and 38.5% A,T library were annealed and TNA was transcribed by primer extension at 50 °C for 10 h. The transcription efficiency using the TNA primer (19%) was approximately half of that achieved using the DNA primer (39%).

we observe the yield of this reaction is slightly less than what we observe using a DNA primer, but we hypothesized that this would not be problematic if the sequences that were being transcribed had the expected nucleotide content, indicating that transcription from the TNA primer does not introduce unwanted sequence bias. To determine nucleotide distribution and evaluate the fidelity of the transcription of our library, TNA from a scaled up primer extension reaction was recovered and reverse transcribed into the corresponding DNA library. The DNA library was amplified using the appropriate primers and inserted into a TOPO vector, cloned into *E. coli*, and DNA from 20 clones was submitted for Sanger sequencing. Analysis of the library sequences revealed that the average amount of each nucleotide in the random region was 32.3% A, 10.3% G, 39.8% T, and 17.2% C (Figure 3.7). While these percentages differ slightly from the starting composition of our library, the retention of all four nucleobases suggests that any sequence bias is limited, and given the transcription efficiency achieved, we concluded that primer extension using the TNA primer under our optimized conditions would be suitable for use in our selection experiments.

TNA SELEX for small-molecule targets

To initiate selection experiments, our ssDNA library was amplified using modified PCR primers (Figure 3.2), which allow strand separation by gel electrophoresis. The PEGylated ssDNA library was recovered and annealed to a FAM-labeled TNA primer, then extended using KOD RI polymerase and tNTPs to generate the TNA library. We began our selections with approximately 15 pmol (~10^{13} sequences) of TNA, which we incubated with OTA-functionalized magnetic beads.

Clone 1 (31,6,54,10)
ATTTAATGAAACTAATCTTTAACTTATGCTAATTGTATTCATTTTTATTT
Clone 2 (25,12,52,12)
AATCATTTTTAGAAATTCTTTTATTTCTATGGATTGACCTTTGTAACGTT
Clone 3 (25,0,35,40)
ACCTCATACAACTTACCCATAATCACCCACTCCTCTTACTACCCTTCTTT
Clone 4 (33,0,38,29)
ATTAATACACTAACACACCTTATAAACCTCATTCTCTTTCCTCTTACAT
Clone 5 (36,12,42,10)
ATATAACTATATTAGATAACTGCTTATATAGGCCTATTGATAATATTTTG
Clone 6 (28,10,44,18)
AGATATTACTCTTCATCAGGCTTCCCTATTATTATTATTCGTAAGATATT
Clone 7 (42,0,48,10)
ATATTAACAATCTTTATTTAAATTATTAATTTTTAAAATACCAATTTACT
Clone 8 (32,12,44,12)
TTACTTCTTGTACCATTATAACCGGTAATGGTTTTAGATATATATTAATA
Clone 9 (38,10,40,12)
AGACGATTACTAAAAATTACATATATCTTATAAGCTTGGATATTTTTACT
Clone 10 (33,0,35,31)
ATTAATACACTAACACACCTTATAAACCTCATTCTCTTCCTCTTACAT
Clone 11 (35,8,35,22)
CCTATATGCCTTGAATAATCATTTATTAAATTCGAGCCTCAATTACCAA
Clone 12 (34,18,42,6)
ACTGATTATATATAATTAATTGTCTGGTGAGATTGAATATCAGATTGATT
Clone 13 (32,18,36,14)
AGCAGATTTTAAAAACTTCGTACTATGTATTGACCGTTAATTTAATGGGC
Clone 14 (35,14,41,10)
ATAGATTTTTGGAATGCTATTCATGTAATAATCAACAGTCTAATAGTTT
Clone 15 (22,12,56,10)
AATTCGTTTTTAGTATTTTTACTTAATGGCTTCTGGTAATTTTTACTATT
Clone 16 (38,8,40,14)
ATTTTATCCAAAAATATTTTTCACTGTACTATATATACAAAAGGATGTCT
Clone 17 (48,12,30,10)
ACTAGAAACTAAAGGATCATACTAAAATATTTTATTAGAAACGTATATAG
Clone 18 (21,0,40,38)
ACCCTCCACCTATCAACCACAACTCATTCTCCCTTCTTTTTATTTTCTAC
Clone 19 (31,8,43,18)
ATCTCATTAACTTCTTATAATTATTTCGAGTCATACTACTATTTACAGG
Clone 20 (28,10,44,18)
TTTTTTAAGAAGGGATATATTCATCACCCTCAATACTCAATTGTTTCTTT

Figure 3.7. Sequence analysis of SELEX library transcription. TNA generated from a scaled up primer extension of the 11.5% G,C and 38.5% A,T DNA library was reverse transcribed and the cDNA was inserted into a vector using a TOPO TA cloning kit for sequencing. Sanger sequencing was used to determine the nucleotide composition of 20 unique clones. Numbers in parentheses indicate the percentage of A,G,T,C nucleotides, respectively, in the N_{50} random region.

In rounds 1-4, bound sequences were recovered by incubation at 95 °C. While this strategy maximizes recovery in early rounds, a key risk is that the aptamers generated will only bind to the immobilized version of the target molecule. Thus, for rounds 5-9, we partitioned bound sequences from the beads by elution with a solution containing an excess of OTA, as this ensures that all sequences recovered are capable of binding to the unmodified target molecule. Additionally, since the attachment of the OTA to amine beads renders the phenylalanine region of the molecule less accessible, we expected that elution with free target molecule would allow for more interaction with this region of the molecule, and therefore enrich tighter binders. The recovered sequences from each round were incubated with DNase I prior to reverse transcription to eliminate any DNA template that was not effectively removed during the gel purification step. If this step is not performed, even a small amount of residual DNA will efficiently amplify during RT-PCR, compromising the ability to enrich the desired target-binding sequences. Importantly, this step was only possible because of our choice to utilize a TNA primer to generate a fully TNA library. Additionally, the use of a FAM-labeled TNA primer to generate our library enabled us to analyze the fraction of library recovered in each round using fluorescence spectroscopy. From these data, we monitored enrichment throughout the selection experiment (Figure 3.8). We observed a gradual increase in enrichment until round 5, at which point the slight decrease in percentage of library recovered was expected due to initiation of the ligand elution. This presumably results because there is a fraction of the sequences that have been enriched to bind tightly to the immobilized target, but they do not bind to free OTA in solution.

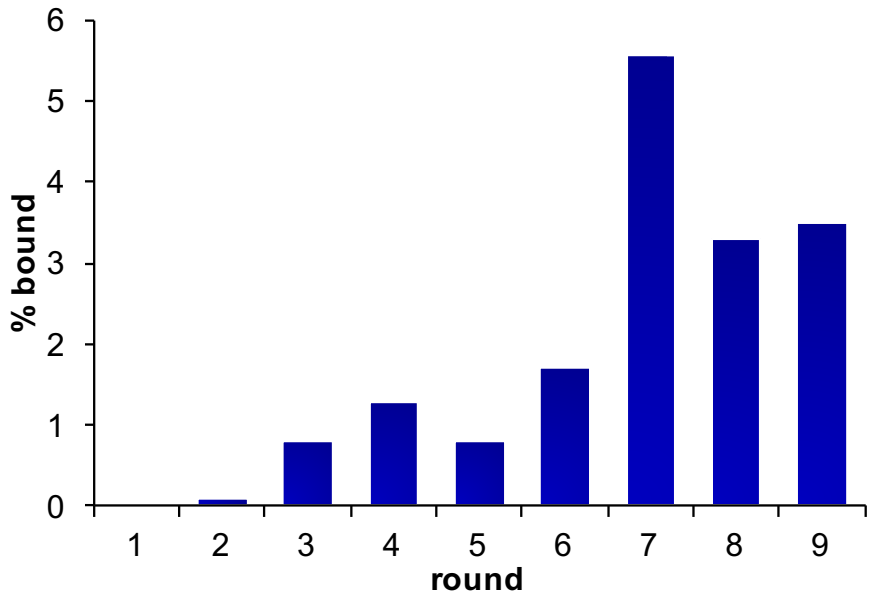


Figure 3.8. Enrichment of OTA aptamers during TNA SELEX.

Sequence recovery and screening of initial aptamer candidates

We observed a considerable increase in library recovery at round 7, and carried out two additional selection rounds to determine whether these would result in further enrichment. However, recovery appeared to show a slight decrease and then plateau with rounds 8 and 9. Despite the lower recovery in the last two rounds, we hypothesized that it would be valuable to compare the sequencing results across rounds 7, 8, and 9, as sequences that persisted through multiple rounds would have a higher probability of being genuine hits. Using Sanger sequencing, 25 clones from each round were analyzed, with the goal of identifying consensus within and between the final rounds of selection (Figure 3.9). We identified 10 sequences for which there were multiple copies in a single round, or which were conserved over multiple rounds of selection using Multalin to generate sequence

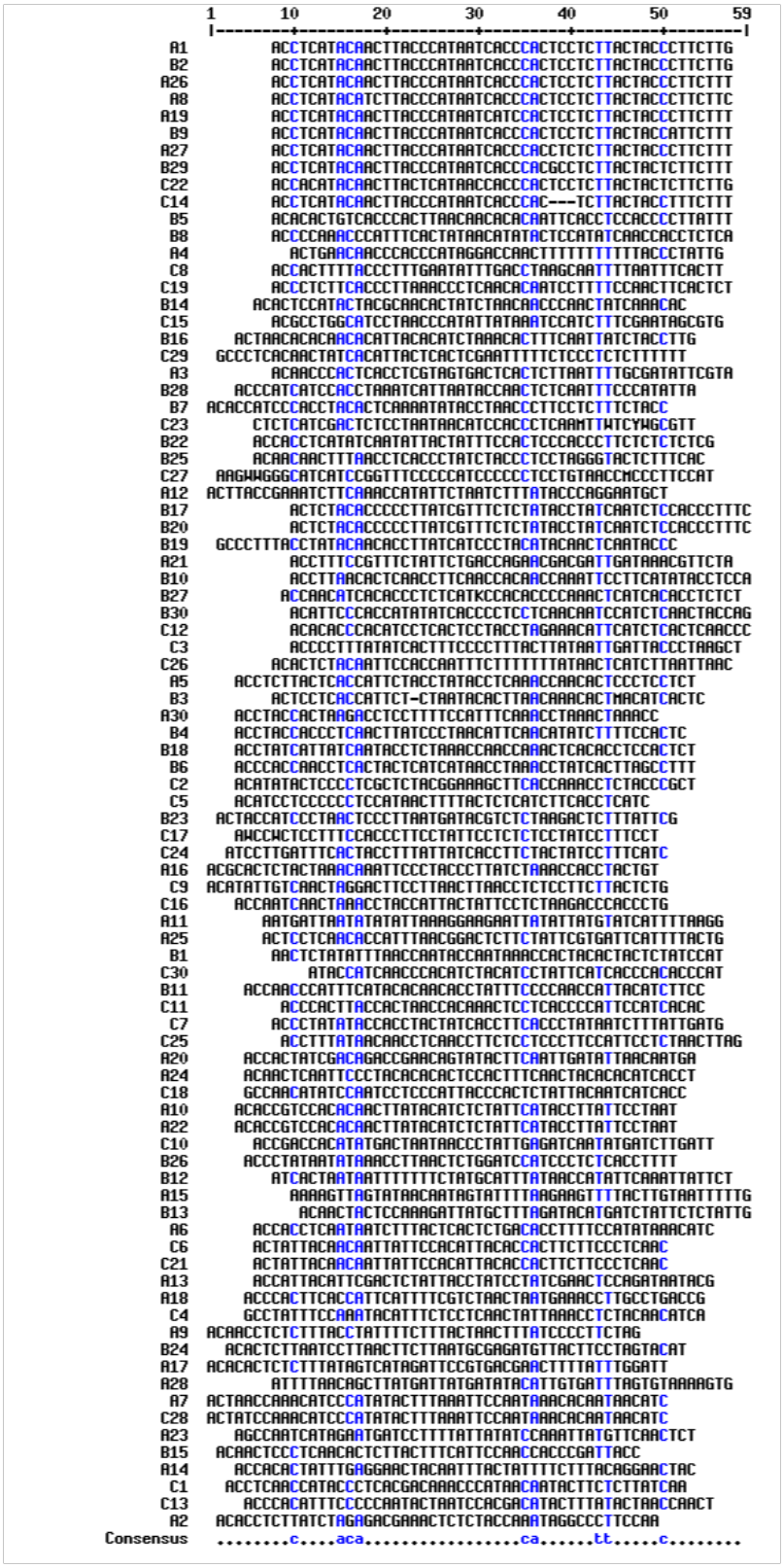


Figure 3.9 Sequences obtained from selection experiments. Letters A, B, and C correspond to clones from rounds 7, 8, and 9, respectively. Alignment was generated using Multalin.

alignment.[28] We purchased PEGylated DNA templates corresponding to each of these sequences, enabling their synthesis using enzymatic primer extension. With these sequences in hand, we next turned to characterizing their binding affinity with the OTA target. In choosing analysis method, a primary consideration was to avoid immobilization of the target molecule. As described earlier, immobilization of small-molecule targets can result in dramatic changes to chemical structure, and we were especially wary of carrying out of analysis using the same immobilization chemistry that had been employed in the selection rounds, as this could lead to false positive results. MST appeared to be a promising approach, as it relies on the change in thermophoretic movement of a molecule in response to the addition of a binding partner, and does not require immobilization of either the receptor or target.[29] Additionally, MST has sufficient sensitivity to detect even minor perturbations, such as those caused by binding of a small-molecule target to a receptor, enabling us to place the fluorescent tag on the TNA strand and avoid modification of the OTA target.[30] The MST data revealed that seven of the 10 TNA sequences showed binding to OTA, with five of these sequences having binding affinities in the nM range (Table 3.1 and Figure 3.10). Impressively, the two best TNA aptamers A04 and A07, were found to have K_D values of 92 ± 2 and 96 ± 4 nM, respectively, representing higher affinity for OTA than the best full length DNA aptamer for this target.[31]

Table 3.1. TNA sequences of potential aptamers and their corresponding K_D values determined by MST. K_D value for best reported DNA aptamer (A08) included for comparison.

	TNA **Sequence (3'→2')**	K_D **(nM)**
A04	GCCGAGATTGCACTTACTATCTCAATAGGGTAAAAAAAAAAAGTTGGTCC-TATGGGTGGGTTGTTCAGTAATTGAATAAGCTGGTATGCGC	92 ± 2
A07	GCCGAGATTGCACTTACTATCTGATGTTATTGTGTTTATTGGAATTTAAAGTATATGG-GATGTTTGGTTAGTAATTGAATAAGCTGGTATGCGC	96 ± 4
A10	GCCGAGATTGCACTTACTATCTATTAGGAATAAGGTATGAATAGAGATGTATAAGTT-GTGTGGACGGTGTAATTGAATAAGCTGGTATGCGC	175 ± 29
B02	GCCGAGATTGCACTTACTATCTCAAGAAGGGTAGTAAGAGGAGTGGGTGATTATGGG-TAAGTTGTATGAGGTAATTGAATAAGCTGGTATGCGC	1968 ± 139
B17	GCCGAGATTGCACTTACTATCTGAAAGGGTGGAGATTGATAGGTATAGAGAAACGA-TAAGGGGGTGTAGAGTAATTGAATAAGCTGGTATGCGC	--
B29	GCCGAGATTGCACTTACTATCTAAAGAAGAGTAGTAAGAGGCGTGGGTGATTATGGG-TAAGTTGTATGAGGTAATTGAATAAGCTGGTATGCGC	184 ± 7
C06	GCCGAGATTGCACTTACTATCTGTTGAGGGAAGAAGTGGTGTAATGTGGAATAATT-GTTGTAATAGTAATTGATAAGCTGGTATGCGC	1414 ± 30
C07	GCCGAGATTGCACTTACTATCTCATCAATAAAGATTATAGGGTGAAGGTGA-TAGTAGGTGGTATATAGGGTAATTGAATAAGCTGGTATGCGC	--
C14	GCCGAGATTGCACTTACTATCTAAAGAAAGGTAGTAAGAGTGGGTGATTATGGG-TAAGTTGTATGAGGTAATTGAATAAGCTGGTATGCGC	--
C22	GCCGAGATTGCACTTACTATCTCAAGAAGAGTAGTAAGAGGAGTGGGTGGTTATGAG-TAAGTTGTATGTGGTAATTGAATAAGCTGGTATGCGC	406 ± 79
	DNA Sequence (5'→3')	
DNA aptamer	GGCAGTGTGGGCGAATCTATGCGTACCGTTCGATATCGTG	148 ± 3

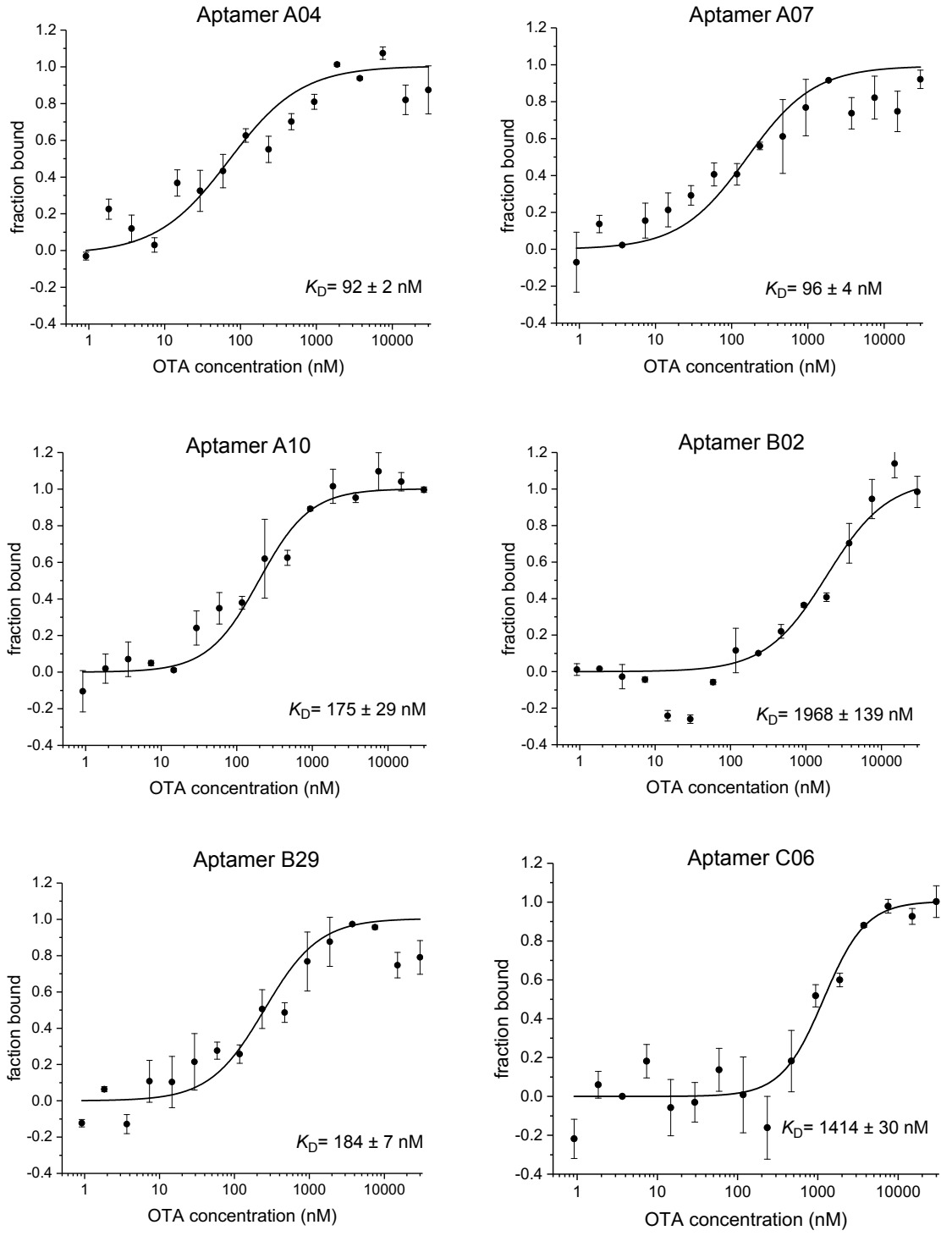


Figure 3.10. Binding isotherms for aptamer sequences determined by MST. Graphs are displayed as fraction bound versus OTA concentration. Each K_D value and error is the average and standard deviation, respectively, of two trials.

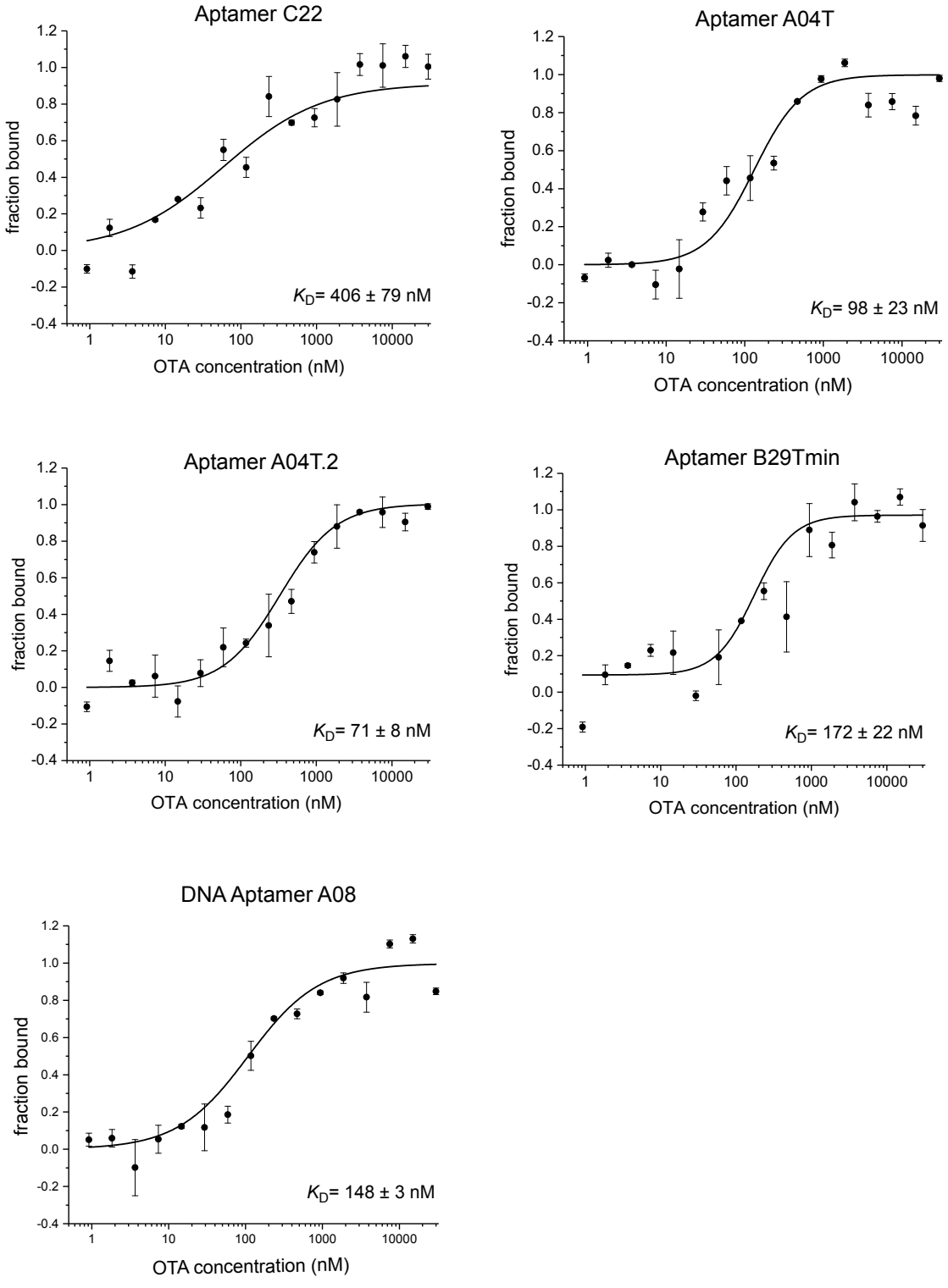


Figure 3.10 continued.

Determination of the minimal binding region for TNA aptamers

We next sought to determine the minimal binding region of the aptamers having the highest affinities. A key goal in these experiments was to identify sequences that could be generated using chemical oligonucleotide synthesis, as this method is significantly more scalable than enzymatic synthesis, and thus is far preferable for future practical applications of our TNA aptamers. In contrast to DNA synthesis, TNA synthesis presents a unique challenge in that both the phosphoramidite and alcohol coupling partners are attached to secondary carbon atoms. As a result, steric hindrance significantly decreases the coupling yield for each step, limiting the length of oligonucleotides that can be practically synthesized. In the case of DNA oligonucleotides, sequences of up to 200 nt can be routinely synthesized, but chemical synthesis of TNA is currently limited to ~40 nt. Thus, of the five aptamers having nM binding affinities, aptamers A04 and B29 were chosen for our minimization studies, as secondary structure prediction suggested putative binding regions of <40 nt (Figure 3.11). In both of these sequences, the main structural elements are near the 3' terminus (which is analogous to the 5' terminus of DNA), allowing us to design initial truncations that could be enzymatically synthesized using the same TNA primer that was employed in the SELEX experiments. For sequence A04, 37 nt were removed from the 2' terminus to give A04T (54 nt), and for sequence B29, 51 nucleotides were removed from the 2' terminus give B29T (43 nt). Initial screening using OTA-functionalized beads showed that both truncated sequences were capable of binding to the target. This provided the

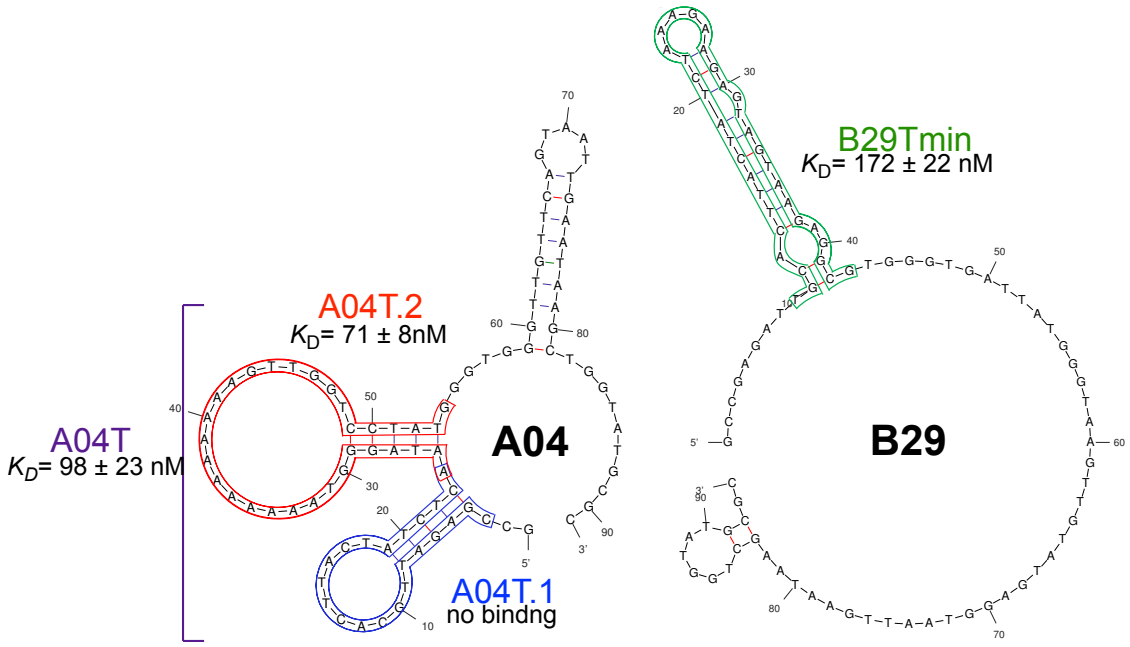


Figure 3.11 Rational truncation of aptamers A04 and B29 based on secondary structure predictions generated using Mfold. Outlined stem loop regions show minimized structures that were chemically synthesized using phosphoramidites.

necessary encouragement to pursue chemical synthesis of sequences having further truncations, and we synthesized the stem loop of B29 (B29Tmin, 35 nt) and loops 1 (A04T.1, 22 nt) and 2 (A04T.2, 31 nt) of A04, each having one additional flanking nucleotide on either side of the predicted structural element (Figure 3.11). Using MST analysis, we measured a K_D value of 172 ± 22 nM for the 35 nt truncation B29Tmin, which is similar to that of the full length B29 aptamer, confirming that only the main stem loop is required for binding to OTA. In the case of the A04-derived sequences, we found that A04T.1 showed no binding to OTA, but the 31 nt sequence A04T.2 bound OTA with a K_D value 71 ± 8 nM. This establishes that loop 2 of the sequence is the binding site for OTA, and it is interesting and fortuitous that the truncated sequence shows higher affinity than either the parent A04 apta-

mer or the A04T truncation containing both stem loops. Excitingly, the ability to generate the highest affinity aptamer using solid-phase synthesis provides a scalable route to acquiring material for use in diagnostic or therapeutic applications. Additionally, the ability to replicate affinity data between enzymatically and chemically synthesized oligonucleotides validates the faithfulness of information transfer between DNA and TNA in the selection and sequencing experiments. To our knowledge, this is the first example demonstrating this transition from enzymatic to chemical synthesis for a fully XNA aptamer.

Biostability comparison of TNA and DNA aptamers with affinity for OTA

In addition to quantifying binding affinity, we also sought to assess the stability of our aptamer in a variety of nuclease rich environments. The Chaput lab has shown that TNA has superior stability compared to RNA and other XNAs such as FANA when challenged with nucleases in biologically relevant matrices.[32] However, there have been no reports that evaluate the ability of TNA aptamers to bind to their target molecule in these complex, nuclease rich environments. For these experiments, we challenged the highest affinity DNA aptamer and our best TNA aptamer (A04T.2) to increasingly harsh conditions including DNase I, 50% human serum in phosphate buffered saline (PBS), and 0.5 mg/mL human liver microsomes. Of these conditions, microsomes provide the most stringent test of nucleic acid stability due to the abundance of nucleases having differing activities and sequence preferences.[32] Figure 3.12a shows that after incubation at 37°C for 3 d in these harsh conditions, the DNA aptamer is fully degraded, while the TNA remains

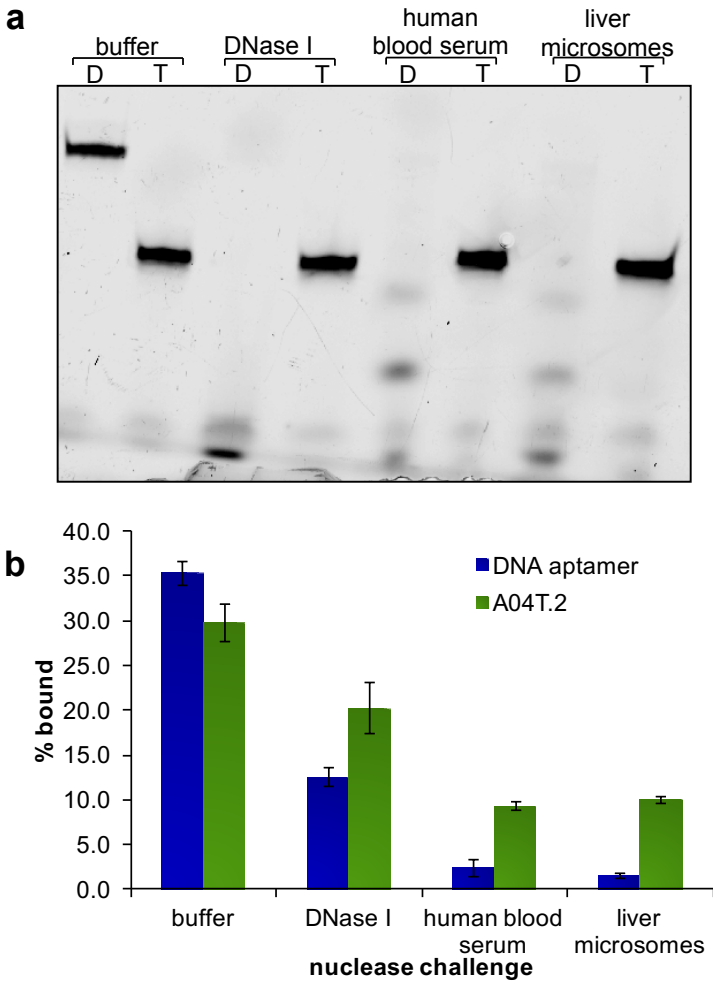


Figure 3.12 Comparison of the biostability of FAM-labeled TNA aptamer A04T.2 and DNA aptamer A08. (a) Denaturing PAGE analysis of the TNA (T) and DNA (D) aptamers after incubation in conditions of increasing nuclease stringency: selection buffer (control), 1.5 U DNase I, 50% human blood serum in PBS, and 0.5 mg/mL human liver microsomes. Samples were incubated under these conditions for 3 days at 37 °C. (b) Bead-binding assay to determine retention of aptamer binding in the presence of nucleases. Each column and error bar represents the average and standard deviation of two trials.

intact with no degradation products detectable by denaturing PAGE. After the nuclease treatment, each sample was incubated with OTA-functionalized beads to qualitatively assess whether the target-binding capability of the TNA and DNA aptamers was retained. The binding data in Figure 3.12b support the PAGE results showing degradation of the DNA aptamer, as we observe significantly less binding after incubation in all three nuclease-rich conditions in comparison to the percent bound by a control DNA sample incubated only in buffer. In comparison, A04T.2 shows a smaller change in binding capacity after incubation in DNase I, human serum, or liver microsomes. Given the high biological stability of TNA, we hypothesize that the decreased binding that we do observe may be attributable to interference from the proteins and other molecules present in the serum and microsome matrices.

Retention of binding in the presence of nucleases

To test this hypothesis, we returned to MST analysis to quantify binding affinity of the aptamers under biological conditions. Human serum is currently regarded as the standard for evaluating oligonucleotide stability and aptamer selectivity, and therefore we preformed all experiments using 50% human serum in PBS. To avoid spectral overlap with the background fluorescence of the serum, Cy5 was employed in place of FAM for labeling the DNA and TNA aptamers in these experiments. Both aptamers were incubated in 50% human serum at 37°C for 3 d or 7d, then MST measurements were acquired for each sample and compared to results obtained earlier for each aptamer in buffer alone. After incubation in serum for 3 d, our TNA aptamer A04T.2 shows a 3.5-fold reduction in binding

affinity (K_D = 245 ± 35 nM), while the DNA aptamer shows a 16-fold reduction in affinity (K_D = 2396 ± 539 nM) (Figure 3.13). We were surprised to not observe a greater reduction in binding affinity of the DNA aptamer after 3 days in serum, as our PAGE and bead-based experiments suggested near complete degradation of the DNA. However, there is evidence that terminal functionalization of nucleic acids can protect against degradation in serum, and given the sensitivity of MST, a binding isotherm can still be acquired even if only a small amount of intact aptamer remains.[33] This notion is supported by the data acquired after seven days of incubation in serum, as the binding affinity of aptamer A04T.2 remains unchanged (K_D = 238 ± 29 nM), but the DNA aptamer shows no detectable binding to OTA, suggesting complete degradation. The ability of our TNA aptamer to maintain target binding capability under such harsh conditions highlights the exceptional nuclease resistance of TNA polymers. Furthermore, the fact that the TNA aptamer binds to OTA with nM affinity despite the large excess of competing molecules in human serum suggests a high level of target binding selectivity.

Conclusions

In summary, we have demonstrated *in vitro* evolution of the first XNA aptamers capable of binding to a small molecule target. Our best TNA aptamers have low nM affinity for the OTA target, surpassing the affinity of the best reported DNA aptamer for the same target. We were able to successfully minimize two of the aptamer sequences to lengths of $\leq$35 nt, enabling them to be generated using solid-phase oligonucleotide synthesis. The 31 nt truncated sequence A04T.2 exhibited the highest affinity of any sequence tested, and validates the integrity of

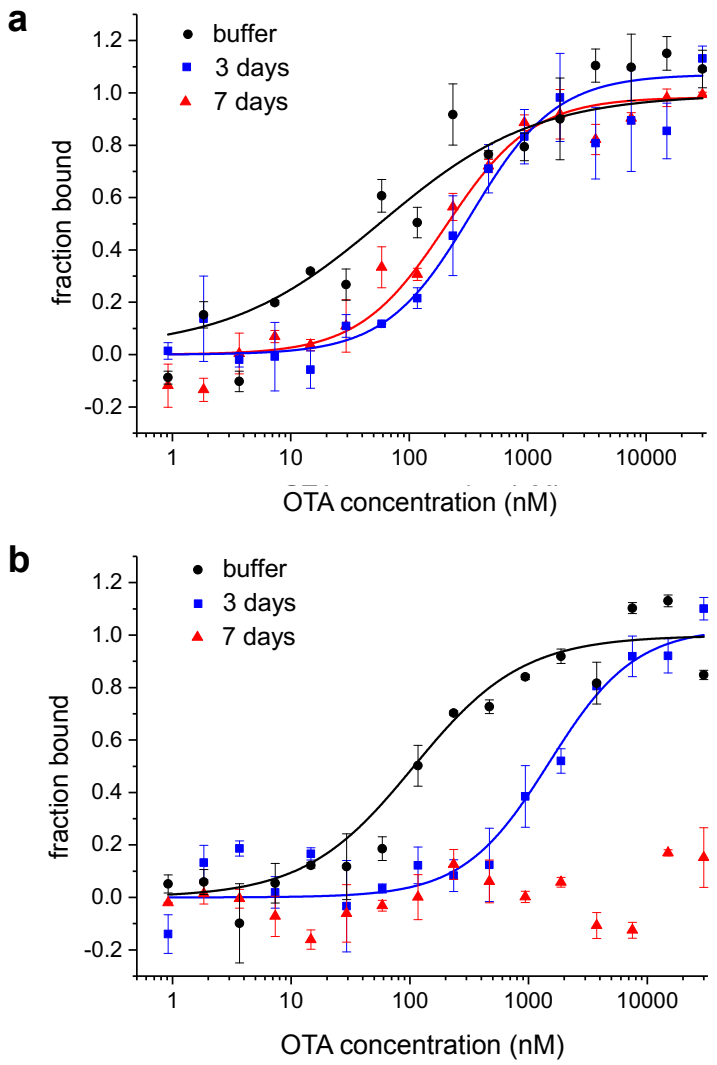


Figure 3.13 MST analysis of (a) TNA and (b) DNA aptamers binding to OTA after incubation in 50% human blood serum in PBS at 37 °C for 3 or 7 d. For both aptamers, the binding affinity measured in selection buffer was used as the control. Error bars represent the standard deviation of two independent trials.

information transfer between enzymatically and chemically synthesized XNA. We also show that the TNA aptamer is capable of binding to OTA after a 7-day incubation in human serum, whereas the best DNA aptamer no longer functions after exposure to these conditions. Achieving success in these experiments required the expansion of XNA selection technology to enhance enrichment by suppressing the background arising from DNA contamination. We envision that this selection approach will be amenable to many interesting small-molecule targets, greatly expanding the reach of XNA applications. Additionally, the demonstration that a TNA aptamer can be minimized to a length that is compatible with chemical synthesis provides access to significantly greater quantities of material, enabling new applications that were previously not feasible due to the limited practical scale of enzymatic synthesis. For example, efforts are currently underway in our laboratory to engineer modified versions of our TNA aptamers into biosensors for use in toxin detection. We envision that these sensors will find high utility in field-based testing, where accurate quantitation of a desired target is needed with minimal sample preparation and using assays capable of enduring rigorous shipment and storage conditions. Additionally, the ability of XNA aptamers to simultaneously provide high affinity, selectivity, and biostability suggests a bright future for these biopolymers in therapeutic applications.

References

(1) Ellington, A. D.; Szostak, J. W. *In Vitro* Selection Of RNA Molecules That Bind Specific Ligands. *Nature* **1990,** *346*, 818-822.

(2) Roberston, D. L.; Joyce, G. F. Selection *In Vitro* Of an RNA Enzyme That Specifically Cleaves Single-Stranded DNA. *Nature* **1990**, *344*, 467-468.

(3) Tuerk, C.; Gold, L. Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase. *Science* **1990***, 249,* 505-510.

(4) Peterson, A. M.; Jahnke, F. M.; Heemstra, J. M. Modulating the Substrate Selectivity of DNA Aptamers Using Surfactants. *Langmuir* **2015***, 31*, 11769-11773.

(5) Taylor, A. I.; Arrangundy-Franklin, S.; Holliger, P. Towards Applications of Synthetic Genetic Polymers In Diagnosis and Therapy. *Curr. Opin. Chem. Biol*. **2014**, *22*, 79-84.

(6) Herdewijn, P.; Marliere, P. Toward Safe Genetically Modified Organisms Through the Chemical Diversification of Nucleic Acids. *Chem Biodivers*., **2009,** *6,* 791-808.

(7) Steele F. R.; Gold, L. The Sweet Allure of XNA. *Nat Biotech*. **2012**, *30*, 624-625.

(8) Ruckman, J.; Green, L. S.; Beeson, J.; Waugh, S.; Gillette, W. L.; Heninger, D. D.; Claesson-Welsh, L.; Janjić, J. 2'-Fluoropyrimidine RNA-based Aptamers to the 165-Amino Acid Form of Vascular Endothelial Growth Factor (VEGF165): Inhibition of Receptor Binding and VEGF-Induced Vascular Permeability Through Interactions Requiring the Exon 7-Encoded Domain. *J. Biol. Chem*. **1998**, *273*, 20556-20567.

(9) Pinherio, V. B.; Holliger, P. The XNA World: Progress Towards Replication and Evolution of Synthetic Genetic Polymers. *Curr. Opin. Chem. Biol.* **2012,** *16,* 245-252.

(10) Meek, K. N.; Rangel, A. E.; Heemstra, J. M. Enhancing Aptamer Function and Stability Via *In Vitro* Selection Using Modified Nucleic Acids. *Methods* **2016***, 106*, 29-36.

(11) Pinheiro, V. B.; Taylor, A. I.; Cozens, C.; Abramov, M.; Renders, M.; Zhang, S.; Chaput, J. C.; Wengel, J.; Peak-Chew, S-Y.; McLaughlin, S. H.; Herdewijn, P.; Holliger, P. Synthetic Genetic Polymers Capable of Heredity and Evolution. *Science* **2012**, *336*, 341-344.

(12) Taylor, A. I.; Holliger, P. Directed Evolution of Artificial Enzymes

(XNAzymes) From Diverse Repertoires of Synthetic Genetic Polymers*. Nat. Protoc.* **2015,** *10*, 1625-1642.

(13) Yu, H.; Zhang, S.; Chaput, J. C. Darwinian Evolution of an Alternative Genetic System Provides Support for TNA as an RNA Progenitor. *Nat. Chem*. **2012**, *4*, 183-187.

(14) Alves Ferreira-Bravo, I. A.; Cozens, C.; Holliger, P.; Destefano, J. J. Selection of 2'-deoxy-2'-fluoroarabinonucleotide (FANA) Aptamers That Bind HIV-1 Reverse Transcriptase with Picomolar Affinity. *Nucleic Acids Res.* **2015**,*15*, 9587-9599.

(15) Thirunavukarasu, D.; Chen, T.; Hongdilokkul, N.; Romesburg, F. E. Selection of 2'-Fluoro-Modified Aptamers with Optimized Properties. *J. Am. Chem. Soc.* **2017,** *139*, 2892-2895.

(16) Mckeague, M.; DeRosa, M. C. Challenges and Opportunities for Small Molecule Aptamer Development. *J. Nucleic Acids. **2012**,* 748913.

(17) Ruscito, A.; DeRosa, M. C. Small-Molecule Binding Aptamers: Selection Strategies, Characterization, and Applications. *Front. Chem.* **2016**, *4*.

(18) Cruz-Aguado, J. A.; Penner, G. Determination of Ochratoxin A with a DNA Aptamer. *J. Agric. Food Chem.* **2008**, *56*, 10456-10461.

(19) Ichida, J. K.; Zou, K.; Horhota, A.; Yu, B.; McLaughlin, L. W.; Szostak, J. W. An in Vitro Selection System for TNA. *J. Am. Chem. Soc.* **2005**, *127*, 2802–2803.

(20) Yu, H.; Zhang, S.; Dunn, M. R.; Chaput, J. C. An Efficient and Faithful In Vitro Replication System for Threose Nucleic Acid. *J. Am. Chem. Soc*. **2013**, *135*, 3583-3591.

(21) Mckeague, M.; McConnell, E. M.; Cruz-Toledo, J.; Bernard, E. D.; Pach, A. Mastronardi, E.; Zhang, X.; Beking, M.; Francis, T.; Giamberardino, A.; Cabecinha, A.; Ruscito, A. Aranda-Rodriguex, R.; Dumontier, M.; DeRosa, M. C. Analysis of In Vitro Selection Parameters. *J. Mol. Evol*. **2015**, *81*, 150-161.

(22) Sau, S. P.; Fahmi, N. E.; Liao, J-Y.; Bala, S.; Chaput, J. C*.* A Scalable Synthesis of **a**-L-Threose Nucleic Acid Monomers**.** *J. Org. Chem.* **2016**, *8*, 2302-2307.

(23) Sau, S. P.; Chaput, J. C**.** A One-Pot Synthesis of a-L Threofuranosyl Nucleoside Triphosphates (tNTPs)*. Bioorg. Med. Chem. Lett.* **2016***, 26*, 3271-3273.

(24) X. Zuker, M. Mfold Web Server for Nucleic Acid Folding and Hybridization

Prediction. *Nucleic Acids Res.* **2003***, 31,* 3406-3415*.*

(25) Dunn, M. R.; Larsen, A. C.; Zahurancik, W. J.; Fahmi, N. E.; Meyers, M.; Suo, Z.; Chaput, J. C. DNA Polymerase-Mediated Synthesis of Unbiased Threose Nucleic Acid (TNA) Polymers Requires 7-Deazaguanine To Suppress G:G Mispairing during TNA Transcription. *J. Am. Chem. Soc.,* **2015**, *137*, 4014-4017.

(26) Larsen, A. C.; Dunn, M. R.; Hatch, A.; Sau, S. P.; Youngbull, C.; Chaput, J. C. A general strategy for expanding polymerase function by droplet microfluidics. *Nat. Commun.* **2016***, 7*, 11235.

(27) Dunn, M. R.; Chaput, J. C. Reverse Transcription of Threose Nucleic Acid by a Naturally Occurring DNA Polymerase. *Chem. Bio. Chem.* **2016**, *17*, 1804-1808.

(28) Corpet, F. Multiple Sequence Alignment with Hierarchical Clustering. *Nucleic Acids Res.* **1988**, *16*, 10881-10890.

(29) Duhr, S.; Braun, D. Thermophoretic Depletion Follows Boltzmann Distribution*. Phys. Rev. Lett.,* **2006,** *96*, 168301.

(30) Entzian, C.; Schubert, T. Studying Small Molecule-Aptamer Interactions Using MicroScale Thermophoresis (MST). *Methods* **2016***, 97*, 27-34.

(31) McKeague, M.; Velu, R.; Hill, K.; Bardóczy, V.; Mészáros, T.; De Rosa, M. C. Selection and Characterization of a Novel DNA Aptamer for Label-Free Fluorescence Biosensing of Ochratoxin A. *Toxins **2014**, 6*, 2435-2452.

(32) Culbertson, M. C.; Temburnikar, K. W.; Sau, S. P.; Liao, J. Y.; Bala, S.; Chaput, J. C. Evaluating TNA Stability Under Stimulated Physiological Conditions. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 2418-2421.

(33) Shaw, J. P.; Kent, K.; Bird, J.; Froehler, B. Modified Deoxynucleotides Stable to Exonuclease Degradation in Serum. *Nucleic Acids Res.* **1991***, 19*, 747-750.

CHAPTER 4

PCR-BASED SELEX FOR THE GENERATION OF STRUCTURE-SWITCHING BIOSENSORS

Introduction

High-throughput screening (HTS) provides a powerful platform for drug discovery through the *in vitro* quantification of small-molecule cofactors and metabolites. For example, screening of GPCR targets and kinases is primarily accomplished through the detection of cAMP in cell extracts and ADP respectively.[1,2] Despite the promise of the HTS platform, there remain challenges for the development of detection reagents, primarily in achieving specific molecular recognition of the target and the transduction of the binding event into an appreciable signal. Because only few reagents can fulfill both of these requirements, multiple steps are necessary in the assay design. Bellbrook's EPIGEN assay for HMTs utilizes two enzymes to convert S-adenosylhomocystine (SAH) to AMP, which can then be subsequently detected using a competitive immunoassay where binding to the antibody displaces a fluorescent tracer which is measured as a change in fluorescence polarization (FP) signal (Figure 4.1).[3] This enzyme coupling is necessary due to the difficulties in developing an antibody that can specifically recognize the

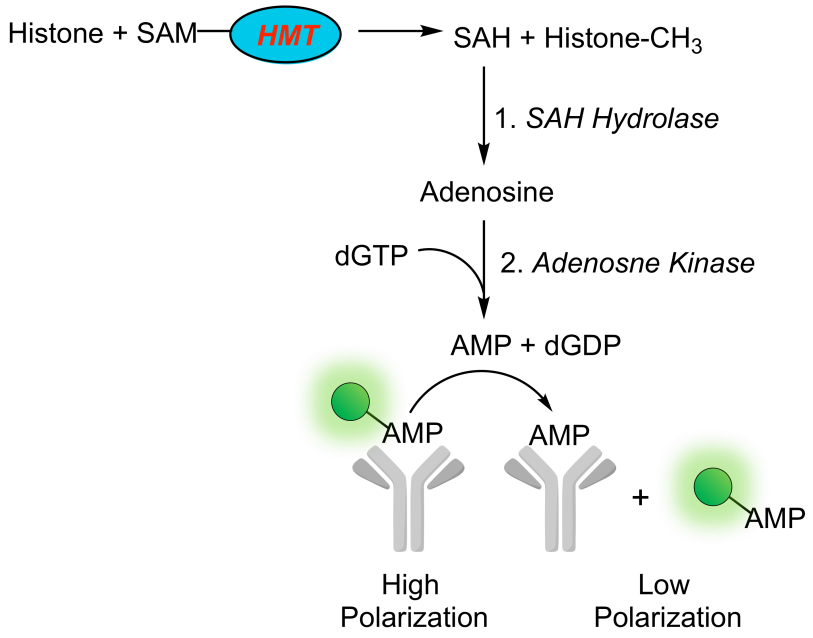


Figure 4.1. Coupled assay for detection of SAH produced by HMT enzymes.

SAH product in the presence of an excess of the donor S-adenosylmethionine. Designing multi-step assays suitable for HTS often requires the identification and synthesis of multiple reagents in addition to the optimization of the integrated output, making the process time and cost inefficient.

Although antibodies are the gold standard for HTS assays, they are not well-suited for the quantification of small-molecule analytes due to their inability to discriminate between structurally similar molecules. Functionalization of small molecules is difficult to accomplish without modifying key recognition elements, which in part leads to poor selectivity. Further, conjugation to a carrier protein for antibody generation is challenging and may reduce the affinity and specificity of the antibodies that are produced. Many assay formats (including FP) require conjugation of the small molecule to a reporter to serve as a competitor in the assay, resulting in similar synthetic challenges.

One target of interest in the development of HTS assays is CoA which can be used for the measurement of histone acetyltransferase (HAT) activity (Figure 4.2b). Because dysfunctional epigenetic regulation contributes to a variety of diseases (neurological disorders, cancers, cardiovascular disease), HATs are being regarded as potential therapeutic targets.[4-7] There is great interest in methods to detect CoA, a product of all HAT reactions, that will be advantageous over other techniques that detect specific acetylation events. There are 18 known human HATs with diverse substrate requirement, thus a universal detection platform for the detection of CoA would be of high utility. Importantly, measurement of HAT enzymatic activity requires specific recognition of CoA in the presence of excess acetyl CoA (AcCoA). Current commercially available assays rely on the reaction of a fluorescent probes with the free thiol in CoA which results in reduced fluorescence, however, this method suffers from frequent false positives and often few inhibitors are identified (i.e. 3 validated out of 1587 primary hits).[8,,9] Other methods utilize coupled enzyme approaches with colorimetric output, however, these assays are insensitive, require great amounts of HATs, and are subject to interference from the coupling enzymes.[10]

Nucleic acid aptamers are a promising alternative to antibodies as affinity reagents in HTS assays.[11-14] Because SELEX doesn't require conjugation of the target molecule to a carrier protein,[15,16] aptamers are more effective than antibodies at discriminating minute structural differences between similar small-molecules.[17] Additionally, the *in vitro* method by which aptamers are generated is less expensive and results in less batch-to-batch variation. Importantly, aptamers can

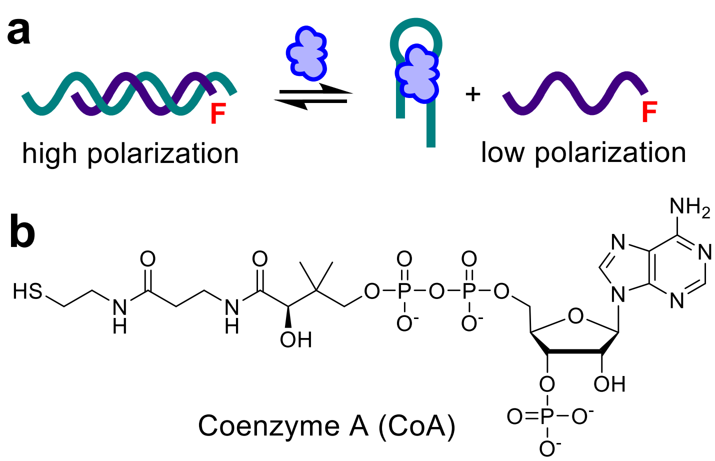


Figure 4.2. Structure-switching biosensors for small-molecule detection**.** (a) SS biosensors provide a dose-dependent FP signal without labeling of the target. (b) Chemical structure of Coenzyme A (CoA).

provide a direct fluorescence signal in a single step which obviates the need to functionalize the small-molecule target. One approach to aptamer-based detection is the structure-switching (SS) biosensor format, in which the aptamer is hybridized to a short complementary sequence to form a duplex in the absence of analyte.[18,19] Upon binding of the target, the complementary strand is displaced in a dose-dependent manner. If the complementary stand is functionalized with a fluorophore, the displacement of the labeled strand causes a change in FP signal (Figure 4.2a).[20]

Despite their potential for various applications, SS biosensors are not often utilized because many aptamers cannot be engineered into this configuration post-selection. This is because the intention of the SELEX process is to identify sequences that bind a certain target, but success is not contingent on a target-dependent conformational change. Consequently, only a few small-molecule binding aptamers (e.g. OTA, cocaine, ATP) are used for analytical applications because it

was discovered that they coincidentally undergo a conformation change upon binding.[21] Recently, research effort has been focused on the development of SELEX methods that directly select for aptamers that exhibit this structure-switching behavior in the presence of analyte.[22,23] All reported SS-SELEX methods are bead-based and consequently suffer limitations related to interference due to the immobilization of the library on a solid support.[24] In bead-based SS-SELEX methods, the library is bound to a complementary strand that is attached to the solid support, and addition of the target results in the selective release of aptamers that function as structure-switching biosensors. However, a significant fraction of nonfunctional sequences are simultaneously eluted because hybridization to the complementary strand is in dynamic equilibrium. This phenomenon lowers the efficiency of the selection process, requiring more rounds (>10) of SELEX in order to observe significant enrichment which is time-consuming and does not always successfully generate the SS biosensor construct.[25] This obstacle is supported by the fact that in some cases traditional SELEX can produce aptamers, but current SS-SELEX methods fail to produce biosensors. The development of a more generalizable selection method for the rapid and reliable generation of SS biosensors would be of great utility in the production of novel small-molecule detection platforms, and to the aptamers field in general.

We propose an innovative new method for the direct selection of SS biosensors using a homogeneous PCR-based SELEX process (Figure 4.3). To carry out the proposed SS-SELEX method, we will use a DNA library with an N_{40} random region which will be flanked by two primer binding sites, which will be hybridized

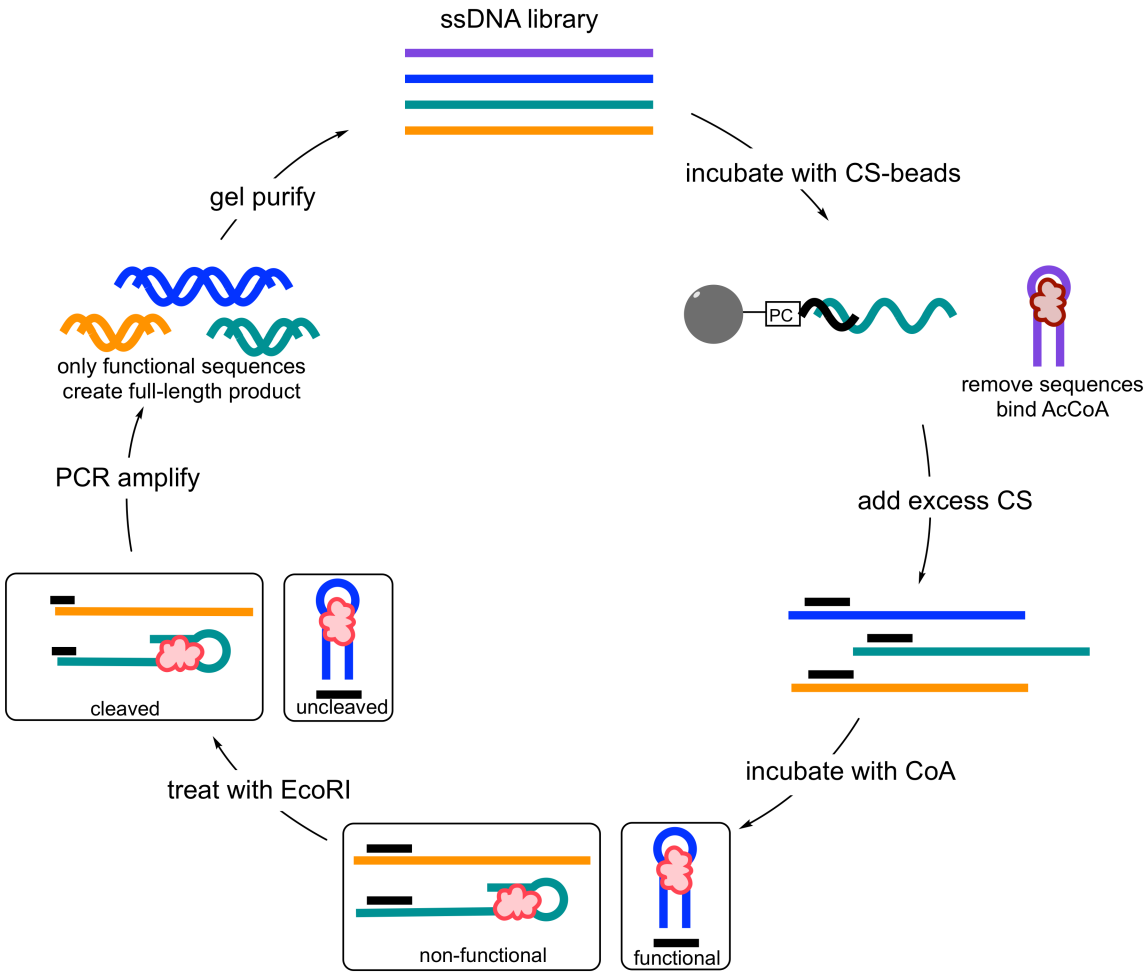


Figure 4.3. Homogenous PCR-based SS-SELEX protocol.

to a short CS near the start of the random region on the 5′ terminus. These sequences will be designed so that when the two strands are hybridized they form the double stranded palindromic recognition site for the EcoRI restriction enzyme. We chose this enzyme because it can provide essentially quantitative cleavage in a variety of buffer conditions. An important benefit over traditional bead-based methods is that the homogenous nature of this design allows the use of excess CS to prevent nonspecific dissociation of library members. Upon addition of the CoA, library members that function as SS biosensors will become dehybridized from the CS (functional), and those that do not bind the target or are capable of binding without releasing the CS will remain in a duplex (nonfunctional). In the presence of EcoRI restriction enzyme, nonfunctional sequences will be cleaved therefore removing one of the primer binding sites necessary for PCR amplification. Conversely, functional sequences that have successfully displaced the CS as a result of target-binding will not be digested and will remain intact. This allows us to selectively amplify those sequences that function as SS biosensors which we can then purify to obtain the ssDNA for the next round of selection. To increase the selectivity of the SS biosensors we will perform a negative selection using acetyl CoA (AcCoA). We anticipate that our proposed SS-SELEX method will enhance enrichment of sequences that function as SS biosensors, requiring fewer rounds to obtain the desired SS architecture. We anticipate these SS biosensors will be optimized for the development fluorescence-based assay for histone acetyl HATs, as the proposed biosensor will based on a ligand-dependent FP signal. Further, we anticipate that SS-biosensors will be selective for CoA which will accommodate

HATs with high affinities for AcCoA. Together we predict these SS biosensors will be instrumental in the development of robust HAT enzyme assays.

Results and discussion

Screening possible restriction enzymes

In the initial optimization of our library design we needed to identify which restriction endonuclease we would utilize for the cleavage step in our selection protocol. In choosing and enzyme, we narrowed our search down to those that had 100% activity in a variety of buffer conditions and proceeded with BamHI and EcoRI.[26] We designed two starting libraries having BamHI and EcoRI cut sites to assess and compare the efficiencies of both enzymes. We synthesized five short complementary sequences for each library ranging in length from 8-12 nt, which included the cut site for the restriction enzymes plus 2-6 additional bases on the 3' end. Previous reports support that this DS length is sufficient for good hybridization but short enough to displaced upon target binding.[27,28] In our first experiments, we hybridized both libraries to CS 8-12 and incubated these samples with their respective restriction enzymes. We analyzed the efficiency of the digestion by denaturing PAGE to monitor percent cleaved. Fluorescence volumes were calculated for each band and percent cleaved was determined by the ratio of truncated product to total band volume in a lane. Neither of these initial libraries showed any cleavage even after an overnight (~20 h) incubation.

Library design and optimization

Upon further research, we discovered information suggesting that restriction endonucleases require flanking regions on their recognition sequences near the end of a DNA fragment for effective cleavage.[29] Without a sufficient number of bases, the enzyme cannot properly dock onto the strand resulting in reduced efficiency. With this knowledge, we designed new CS sequences for two new libraries (EcoRI and BamHI libraries) which incorporated 1-3 flanking regions on either side of recognition site resulting in three strands that were 8, 10, and 12 nt in length (Figure 4.4a). We included a 14 nt CS for the BamHI library because the enzyme is slightly less efficient than EcoRI and longer sequences are needed for complete digestion. Excitingly, we observed cleavage of both libraries by their respective enzymes with efficiencies of 78% with the 8 nt CS and 85-90% for the longer CS sequences (Figure 4.4b). This difference in yield is expected as longer sequences are expected to improve efficiency. Although the ability of the enzymes to cleave their respective libraries were comparable, the efficiencies were slightly higher for EcoRI at all CS lengths. Therefore, we decided to move forward with EcoRI as our restriction endonuclease for optimization and ECS 12 which had the highest efficiency (90%).

Maximizing cleavage efficiency

While we were excited to observe such high percentages of cleaved library, our selections require that this efficiency be as close to 100% as possible. During our selection we will monitor enrichment by quantifying the percent of library that remains uncleaved since the functional sequences will displace the CS and will not

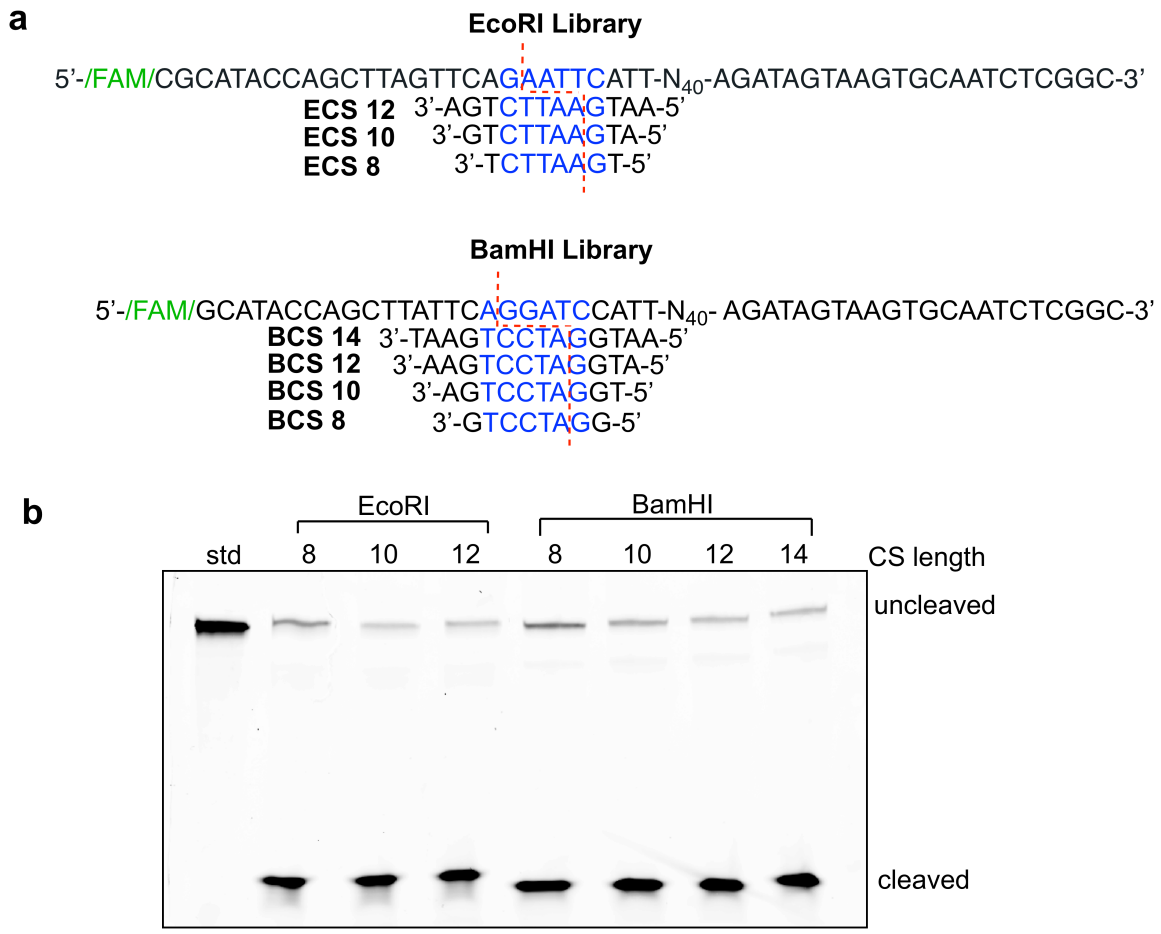


Figure 4.4. Optimizing library and digest conditions. (a) Library and CS sequences for both BamHI and EcoRI with recognition sequence in blue. CS sequences listed under library to illustrate cut site. (b) Digestion of libraries using CS with flanking regions. Library/DS complexes were incubated with enzyme for 2 hrs and analyzed by PAGE to quantify % cleaved.

be degraded. Any nonfunctional library members that remain uncleaved as a result of incomplete digestion have the potential to dominate the sequence pool and make it more difficult to identify functional sequences. As such, the digestion of our library using EcoRI needs to be optimized prior to beginning selection. We hypothesized that the reduced efficiency of the EcoRI digestion could be attributed to library members which possess inherent secondary structure that prevents the hybridization of the CS. Although these sequences don't undergo a conformational

change upon target binding, they will be propagated to the next selection making it difficult to enrich sequences that function as SS biosensors. To test this theory, we ordered a sequence with minimal predicted secondary structure which we call “EcoRI test sequence” (Figure 4.5a). We hypothesized that if it was inherent secondary structure that was inhibiting hybridization to the displacement strand, the use of this test sequence would allow for efficient hybridization, resulting increased cleavage efficiency.

We tested the ability of EcoRI to cleave this test sequence using a 1:1 and 1:2 ratio of the library to ECS 8, 10, and 12. As evident in Figure 4.5b, we saw no improvement upon the 90% cleavage efficiency we observed using the library even after a 20 h incubation period. Despite this result, we remained convinced that it

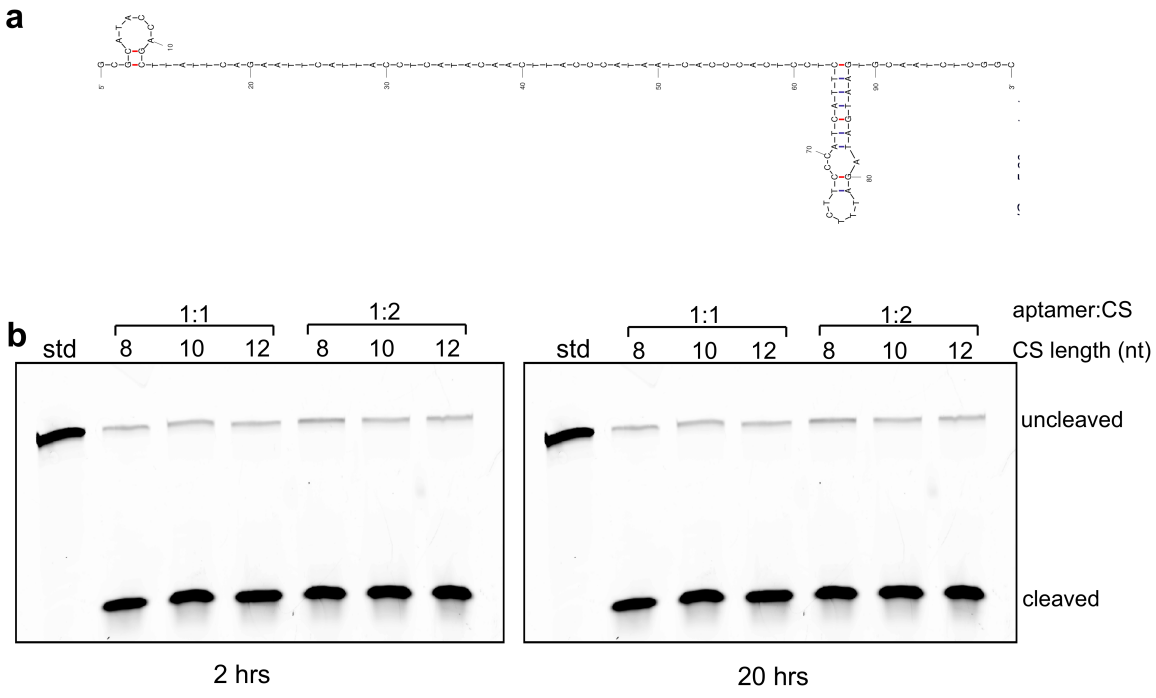


Figure 4.5. Cleavage of EcoRI control sequence. (a) Secondary structure prediction for EcoRI test sequence. (b) Digestion of test sequence using different CS lengths, ratios, and incubation times.

was in fact prohibitive secondary structure preventing complete digestion of our library. We proposed another idea to validate this claim using magnetic beads functionalized with the CS. We hypothesized this would be a clever way to separate sequences that can effectively hybridize the CS. We modified the 12 nt CS with an amine on the 3′ terminus to enable attachment to carboxylic acid functionalized magnetic beads (Figure 4.6a). Since the CS is complementary to the part of the library closest to the random region, we also added two PEG spacers in between the CS and terminal amine to ensure that there would be no steric hindrance preventing the library from hybridizing to the beads. Additionally, we included a photocleavable (PC) spacer in front of the DS to allow recovery of the library/CS

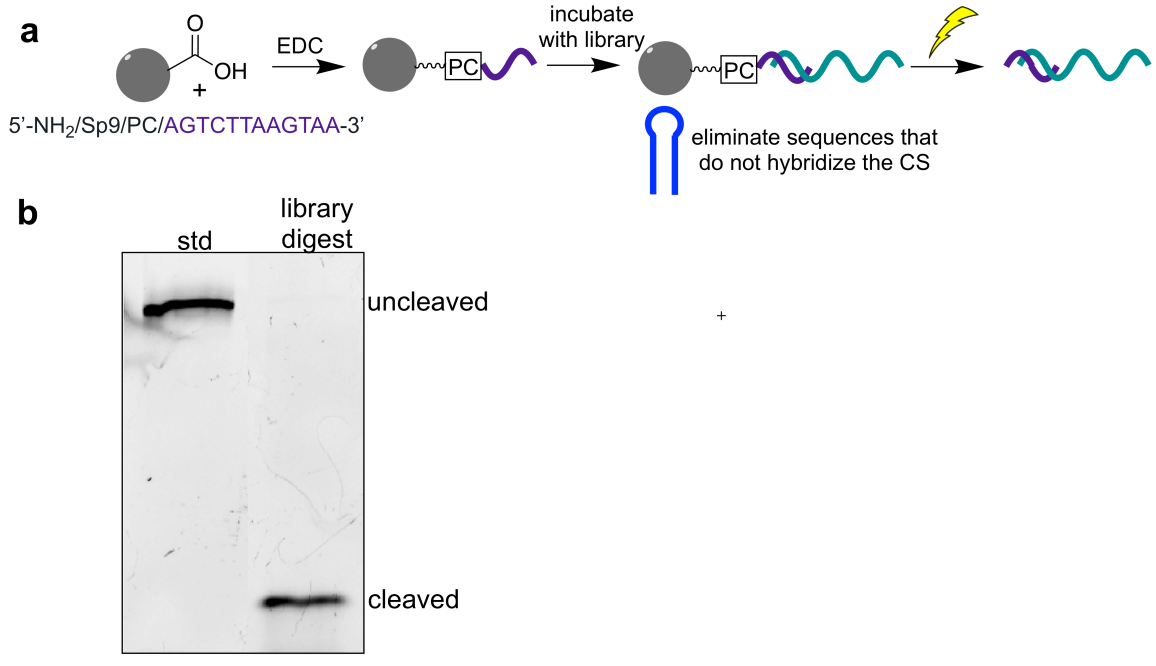


Figure 4.6. Elimination of sequences with prohibitive secondary structure. (a) Conjugation of photocleavable CS to magnetic beads and hybridization of library to form biosensor complex attached to beads. (b) Digestion of library/CS complex by EcoRI after recovery from beads by UV irradiation.

complex from magnetic beads. We incubated our library with the CS-beads, and once hybridized we performed a series of washes in order to remove nonbound sequences. We then used UV light to remove the biosensor complex which was subsequently used in an EcoRI digestion reaction. We were pleased to observe that 99% of our library was cleaved using this bead step (Figure 4.6b). We determined that this new level of efficiency we achieved using this step would be suitable for our selections. This confirmed that the certain sequences are problematic for the generation of biosensors using our proposed method and can be effectively removed using the aforementioned bead step. Further, we anticipate that the inclusion of this step in the first round of our SELEX method will aid in reducing any bias that may affect the efficacy of the selection.

Reducing bias during selection

One important question we wanted to answer is what happens to the <1% of library that doesn't get cleaved over the SELEX process, as we believe that it has the potential to bias the enrichment of nonfunctional sequences. We hypothesized that we could gain valuable information if we performed the bead step we developed, PCR amplified the uncleaved portion of the library, and attempted to re-digest this material (without prior bead step). This would allow us to verify that we don't reintroduce bias after amplification and also to estimate whether we need to perform the bead step every round. After completing the bead method to hybridize and recover the library/CS complex, the remaining ~1% of uncleaved library was amplified and purified. This purified library was then hybridized to ECS 12 at a 1:2 ratio in solution and incubated with EcoRI. We were pleased to observe <0.1%

of the library remained uncleaved which further validates the benefit of including the bead step prior to digestion. This result suggests that the beads will only be required in the first round as we are able to retain highly efficient digestion of the library after amplification and re-digestion. However, we concluded that it would be beneficial to include this step every three rounds to provide a convenient method to eliminate any bias that may develop. Importantly, because selection step will occur after the biosensor complex is recovered from the magnetic beads, we are able to maintain the homogenous nature of our proposed SS-SELEX method.

PCR optimization

In preparation for selection, it was important to carefully consider the conditions for PCR amplification. The library in this study present a unique challenge due to the fact that there is the restriction site within the 5' primer binding site of the library which will also be present in the forward primer for PCR. Because the recognition sequence is a palindrome, the forward primers are likely to form dimers which can be problematic during PCR amplification. While some off-target product during PCR will not be detrimental since we will perform a gel purification after amplification, we wanted to minimize the propensity for primer dimers to ensure our potential aptamers can be effectively enriched during PCR. Using too much primer in PCR reactions increases the probability that they will bind to eachother to form dimers and also to nonspecific sites on the template.[30] In addition, increasing the annealing temperature can increase the specificity of primer binding in order to mitigate the formation of primer dimers. Therefore, we began the optimization of our PCR protocol by using a range of primer concentrations and annealing

temperatures and assessing the fidelity of library amplification (Figure 4.7a). We used three annealing temperatures (55°C, 60°C, 65°C) and three primer concentrations (50 nM, 100 nM, and 200 nM) in various combinations and analyzed the results by denaturing PAGE. We observe that at higher concentrations of primer and lower annealing temperatures we see multiple bands likely resulting from primer dimers and nonspecific hybridization to the library. We obtained relatively clean product at 50 nM primer and an annealing temperature of 60°C and decided to use those conditions during selection. While we observe reduced yield with these conditions we expect that this can be compensated for by increasing the number of PCR reactions we perform. There have been reports that organic additives can improve PCR amplification by enhancing specificity to increase the yield of full length product.[31,32] We tried various combinations of betaine, DMSO, and 1,2-propanediol which are all additives that have been validated to improve PCR fidelity (Figure 4.7b). We do observe the reduction of nonspecific products with the combination of 1.25 M betaine and 5% DMSO (lane 5), when compared to the control (lane 1). However, this increased specificity occurs at the cost of yield as the bands appear proportionally faint. We concluded there would be no incremental benefit to including additives, and that the modifications to the annealing temperature and primer concentration provide sufficient specificity for selection.

SELEX method and progress

In addition to eliminating some nonfunctional sequences, we hypothesized that the CS-bead step would provide a convenient way to perform a negative selection. Every three rounds a solution of AcCoA will be added to the library hybridi-

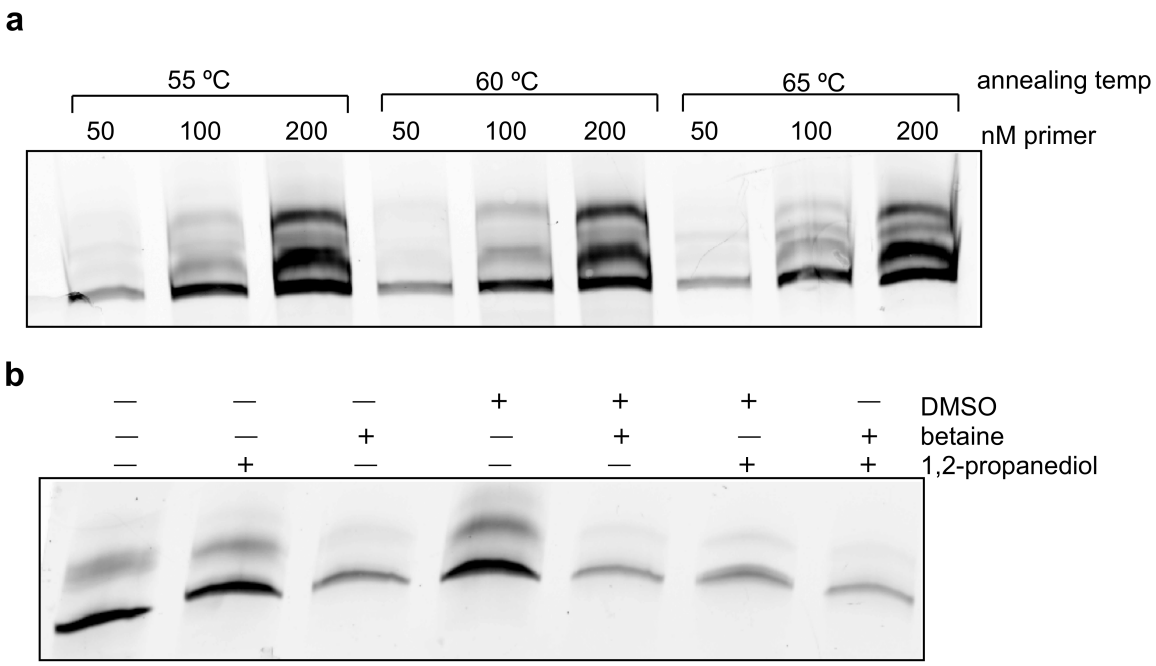


Figure 4.7. PCR optimization using EcoRI library. (a) Optimization of annealing temperature and primer concentration for PCR of EcoRI library. (b) Exploring the effect of PCR additives (0.8 M 1,2-propanediol, 1.25 M betaine, and 5% DMSO) on yield of full-length product.

zed to the CS-beads for an hour incubation. Any sequences that undergo a conformational change due to binding of AcCoA, will be separated from the remaining library members using a magnetic stand. This allows us to remove any sequences that bind AcCoA, so that we enrich those sequences that function selectively as SS biosensors for CoA. We began our positive selection at 1 mM CoA in the first round and plan to reduce the concentration once we observe enrichment. Reducing the concentration of the CoA will enable us to select for aptamers with better affinity for the target. In round 3, we introduced the negative selection to the bead bound library/DS complex using 100 μM AcCoA. We increased this amount to 250 μM AcCoA in round 6 to increase the stringency of our negative selection and we plan to gradually increase the AcCoA amount to 1 mM over subsequent rounds.

We expect that challenging the library to an equivalent concentration of off-target as where we initiated our positive selection will allow us to select for biosensors that have superior selectivity. We have completed six rounds of selection to date, and while we have not observed significant enrichment, we believe these results look promising (Figure 4.8). There is a decrease in percent uncleaved library in round 3 which we would expect due to the introduction of the negative selection. We are encouraged that the subsequent rounds of selection all exhibit higher percent uncleaved than round 1, supporting that we are successfully enriching our library. Further, it is promising that we observe higher enrichment in round 6 compared to round 3, even with the increased stringency of the negative selection.

Conclusions

We propose a novel PCR-based SELEX method for the generation of SS-biosensors. In this protocol, functional sequences be digested by a restriction enzyme and will be selectively amplified by PCR. We anticipate that this method will

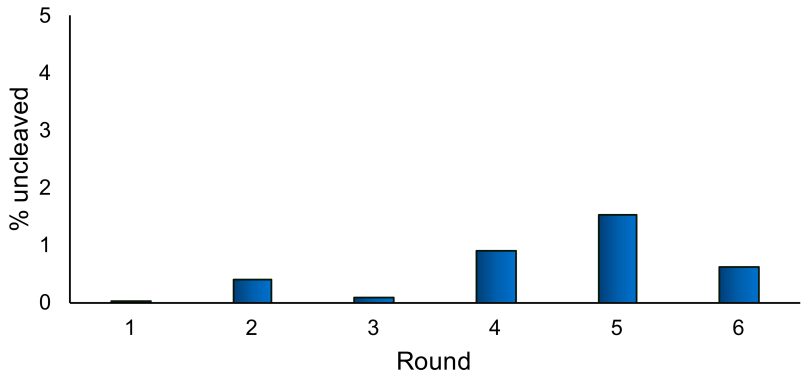


Figure 4.8. Progress for the selection of SS biosensors for CoA. Enrichment is monitored by the amount of uncleaved library reamaining after EcoRI digestion.

be advantageous as the selection step does not require the use of a solid support. This will overcome the limitations of traditional bead-based SS-SELEX methods such as nonspecific elution and retention of nonfunctional sequences, and allow for high partitioning and enrichment. We have carefully designed our SELEX library to enable digestion by restriction enzyme when hybridized to a short CS. We have chosen EcoRI for our method, and have optimized conditions for maximal digest efficiency and PCR amplification of our library. We anticipate that together these optimized conditions will lead to reduced bias during our selection and effective enrichment of sequences that function as SS biosensors. We are encouraged by our SELEX progress thus far, and are confident this method holds promise for the generation of selective, high affinity biosensors for the detection of CoA. We are optimistic that SS biosensors for CoA generated by our innovative method will be of high utility in the development FP-based assays for HAT detection. Further, we anticipate that successful implementation of this novel SELEX method will provide a generalizable evolution strategy to develop SS biosensors to a variety of small-molecule targets, and streamline the development of aptamer-based biosensors for diverse applications.

Materials and methods

Oligonucleotides synthesis

All DNA sequences were purchased from the University of Utah DNA/Peptide Synthesis Core Facility. All oligonucleotides were purified by denaturing PAGE and the desired band was excised and incubated in crush and soak buffer (300 mM ammonium acetate, 1 mM EDTA, pH 8.0) at 40°C overnight. Samples were

separated from gel pieces and buffer exchanged into water using Amicon Ultra-0.5 Centrifugal Unit with Ultracel 10 membrane (EMD Millipore).

Functionalization of photocleavable CS to magnetic beads

400 µl of M-270 carboxylic acid beads (Life Technologies) were washed three times with 1 mL 25 mM MES, pH5. After each wash, the beads were separated from the supernatant using a magnetic separation stand (DynaMag-Spin, Life Technologies). 8400 pmol of PC-DS12 (AATGAATTCTGA/PC spacer/Sp18/amine) was suspended in 240 µl of 25 mM MES buffer, pH 5 and mixed with the washed beads and incubated for 30 min at room temperature with slow tilt rotation. After equilibration to room temperature, EDC was dissolved in cold 100 mM MES, pH 5 to a final concentration of 100 mg/mL. 120 µl of the EDC solution was added to bead/ligand mixture and mixed well. The beads were brought to a final volume of 400 µl with 25 mM MES, pH 5 and incubated at 4°C overnight with slow tilt rotation. The functionalized beads were washed three times with 1 mL 1X EcoRI buffer (100 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, 0.025% Triton X-100, pH 7.5) and re-suspended in 500 µL of this buffer for storage at 4°C. Successful attachment of PC-DS12 was validated by monitoring absorbance at 260 nm.

Hybridization of library to CS-beads and recovery of biosensor complex

5 pmol of EcoRI.3 library were added to 10 µL (2×10^7) of /PC/ECS 12 functionalized beads to a final volume of 100 µL in 1X EcoRI buffer. The library was

hybridized to the DS-beads by heating to 95 °C for 5 min, cooling at 4 °C for 10 min, and a final incubation at 25 °C overnight, the beads were then washed three times with 200 µL EcoRI buffer and resuspended 100 µL (having an extra equivalence of ECS12. The biosensor complex was cleaved from the beads by exposure to UV light (365 nm) for 45 min. This step was performed in a 4°C refrigerator to minimize dissociation of the library and CS.

EcoRI digest of biosensor complex

The library/DS complex was divided into 10 µL portions and 1.5 U EcoRI (New England Biolabs) was added to each aliquot. The solutions were incubated for 2 hrs at 37°C. The EcoRI was denatured for 20 min at 65°C and the library was recovered using a minelute PCR cleanup column (Qiagen). Samples were analyzed by denaturing PAGE on a 10% polyacrylamide gel to monitor digestion of the library. The gels were imaged using a blue laser (473 nm)/long pass blue filter and fluorescence volumes were calculated for all lanes. Digest efficiency was determined by the percent cleaved (fluorescence band volume for the cleaved product/ total lane volume).

PCR amplification of library after digest

A single stranded DNA library having the sequence 5'- CGCAT-ACCAGCTTAGTTCAGAATTCATT-N_{40}-AGATAGTAAGTGCAATCTCGGC -3' (20 pmol) was amplified in 50 µL PCR reactions containing 0.2 µM template, 0.05 µM each primers, 200 µM dNTPs, and 2.5 U Taq DNA polymerase in 1x Thermopol buffer (20 mM Tris-HCl, 10 mM $(NH_4)_2SO_4$, 10 mM KCl. 2 mM $MgSO_4$, 0.1% Triton

X-100, pH 8.8, New England Biolabs). The forward primer had the sequence 5'-/FAM/CGCATACCAGCTTAGTTCAGAATTCATT-3' and the reverse primer had the sequence 5' TTTTTTTTTTTTTTTTTTTT/Sp9/GCCGAGATTGCACTTACTA-TCT -3'. PCR was carried using a program with an initial denaturation at 95°C for 3 min, 12 cycles of (95°C for 30 s, 60 °C for 30 s, and 72 ^{o}C for 20 s), and a final extension 72 ^{o}C for 2 min. The amplified double stranded DNA was purified using a PCR cleanup column (Qiagen) and the PEG-functionalized strand was separated from the FAM labeled strand on a denaturing 10% polyacrylamide gel. FAM-labeled ssDNA library was recovered by elution in crush and soak buffer as mentioned previously. Samples were recovered using size-exclusion centrifugal units as described previously.

PCR-based SS-SELEX

The purified DNA library (~10^{14} sequences) was folded in 1X EcoRI buffer in a total volume of 100 µL by heating to 95 °C for 5 min, cooling at 4 °C for 10 min, and a final incubation at 25 °C for 10 min. CoA was added to the solution at a final concentration of 1 mM and incubated for 1 h at 25 °C. For the negative selection every three rounds, the library was hybridized to the CS-beads and incubated with the desired concentration AcCoA (start at 100 µM). The off-target binding sequences were removed by washing three times with 100 µL 1X EcoRI buffer, using the magnetic stand to separate the supernatant from the beads. The beads were resuspended in 100 µL binding buffer and the bound sequences were recovered by exposure to UV light as previously described. 1.5 U EcoRI was added to the mixture and incubated for 2 h at 37 °C. Enrichment was monitored as the amount

of uncleaved material that remained postdigestion. The restriction enzyme was purified using a PCR cleanup column and the recovered DNA was amplified using the method described previously.

References

(1) Norskov-Lauritsen, L.; Thomsen, A. R.; Brauner-Osborne, H. G. Protein Coupled Receptor Signaling Analysis Using Homogenous Time-Resolved Forster Resonance Energy Transfer (HTRF(R)) Technology. *Int. J. Mol. Sci.* **2014,** *15*, 2554-72.

(2) Huss, K. L.; Blonigen, P. E.; Campbell, R. M. Development of a Transcreener Kinase Assay for Protein Kinase A and Demonstration of Concordance of Data With a Filter-Binding Assay Format. *J. Biomol. Screen.* **2007,** *12*, 578-84.

(3) Klink, T. A.; Staeben, M.; Twesten, K.; Kopp, A. L.; Kumar, M.; Dunn, R. S.; Pinchard, C. A.; Kleman-Leyer, K. M.; Klumpp, M.; Lowery, R. G. Development and Validation of a Generic Fluorescent Methyltransferase Activity Assay Based on the Transcreener AMP/GMP Assay. *J. Biomol. Screen.* **2012,** *17*, 59-70.

(4) Gajer, J. M.; Furdas, S. D.; Grunder, A.; Gothwal, M.; Heinicke, U.; Keller, K.; Colland, F.; Fulda, S.; Pahl, H. L.; Fichtner, I.; Sippl, W.; Jung, M. Histone Acetyltransferase Inhibitors Block Neuroblastoma Cell Growth In Vivo. *Oncogenesis* **2015,** *4*, e137.

(5) Judes, G.; Rifai, K.; Ngollo, M.; Daures, M.; Bignon, Y. J.; Penault-Llorca, F.; Bernard-Gallon, D. A Bivalent Role of TIP60 Histone Acetyl Transferase in Human Cancer. *Epigenomics* **2015,** *7*, 1351-63.

(6) Schneider, A.; Chatterjee, S.; Bousiges, O.; Selvi, B. R.; Swaminathan, A.; Cassel, R.; Blanc, F.; Kundu, T. K.; Boutillier, A. L. Acetyltransferases (HATs) as Targets for Neurological Therapeutics. *Neurotherapeutics* **2013,** *10*, 568-88.

(7) Wang, Y.; Miao, X.; Liu, Y.; Li, F.; Liu, Q.; Sun, J.; Cai, L. Dysregulation of Histone Acetyltransferases and Deacetylases in Cardiovascular Diseases. *Oxid. Med. Cell Longev.* **2014,** *2014*, 641979.

(8) Gao, T.; Yang, C.; Zheng, Y. G. The fluorescence-based acetylation assay using thiol-sensitive probes. *Methods Mol. Biol.* **2013,** *981*, 229-38.

(9) Dahlin, J. L.; Sinville, R.; Solberg, J.; Zhou, H.; Han, J.; Francis, S.; Strasser, J. M.; John, K.; Hook, D. J.; Walters, M. A.; Zhang, Z. A Cell-Free Fluorometric High-Throughput Screen for Inhibitors of Rtt109-Catalyzed Histone Acetylation. *PLoS One* **2013,** *8*, e78877.

(10) Berndsen, C. E.; Denu, J. M. Assays for Mechanistic Investigations of Protein/Histone Acetyltransferases. *Methods* **2005,** *36*, 321-31.

(11) Liu, J.; Cao, Z.; Lu, Y. Functional Nucleic Acid Sensors. *Chem. Rev*. **2009**, *109*, 1948-1998.

(12) O'Sullivan, C. K. Aptasensors--The Future of Biosensing? *Anal. Bioanal. Chem*. **2002**, *372*, 44-48.

(13) Cho, E. J.; Lee, J.-W.; Ellington, A. D. Applications of Aptamers as Sensors. *Annu. Rev. Anal. Chem*. **2009**, *2,* 241-264.

(14) Famulok, M.; Hartig, J. S.; Mayer, G. Functional Aptamers and Aptazymes In Biotechnology, Diagnostics, And Therapy. *Chem. Rev*. **2007**, *107*, 3715-3743.

(15) Ellington, A. D.; Szostak, J. W. In Vitro Selection of RNA Molecules That Bind Specific Ligands. *Nature* **1990**, *346,* 818-822.

(16) Tuerk, C.; Gold, L. Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands To Bacteriophage T4 DNA Polymerase. *Science* **1990**, *249*, 505-510.

(17) McKeague, M.; DeRosa, M. C. Challenges and Opportunities for Small Molecule Aptamer Development. *J. Nucleic Acids* **2012**, *2012*, 748913.

(18) Nutiu, R.; Li, Y. Structure-Switching Signaling Aptamers. *J. Am. Chem. Soc*. **2003**, 125, 4771-4778.

(19) Jhaveri, S. D.; Kirby, R.; Conrad, R.; Maglott, E. J.; Bowser, M.; Kennedy, R. T.; Glick, G.; Ellington, A. D. Designed Signaling Aptamers That Transduce Molecular Recognition to Changes in Fluorescence Intensity. *J. Am. Chem. Soc.* **2000,** *122*, 2469-2473.

(20) Zhu, Z.; Schmidt, T.; Mahrous, M.; Guieu, V.; Perrier, S.; Ravelet, C.; Peyrin, E. Optimization of the Structure-Switching Aptamer-Based Fluorescence Polarization Assay for the Sensitive Tyrosinamide Sensing. *Anal. Chim. Acta* **2011**, *707*, 191-196.

(21) Uphoff, K. W.; Bell, S. D.; Ellington, A. D. In Vitro Selection of Aptamers: The Dearth of Pure Reason. *Curr. Opin. Struct. Biol.* **1996,** *6*, 281-288.

(22) Nutiu, R.; Li, Y. In Vitro Selection of Structure-Switching Signaling Aptamers. *Angew. Chem. Int. Ed.* **2005,** *44*, 1061-1065.

(23) Stoltenburg, R.; Nikolaus, N.; Strehlitz, B. Capture-SELEX: Selection of DNA Aptamers for Aminoglycoside Antibiotics. *J. Anal. Methods Chem.* **2012**, 415679.

(24) Mosing, R. K.; Bowser, M. T. Isolating Aptamers Using Capillary Electrophoresis-SELEX (CE-SELEX). *Methods Mol. Biol.* **2009,** *535*, 33-43.

(25) Feagin, T. A.; Heemstra, J. M. *unpublished results*.

(26) New England Biolabs. NEBuffer Activity/Performance Chart with Restriction Enzymes. https://www.neb.com/tools-and-resources/usageguidelines/nebuffer-performance-chart-with-restriction-enzymes.html (accessed May 12, 2017).

(27) Zhu, Z.; Shmidt, T.; Mahrous, M.; Guieu, V.; Perrier, S.; Ravelet, C.; Peyrin, E. Optmization of the Structure-Switching Aptamer-Based Fluoresence Polarization Assay for the Sensitive Tyrosynamide Tensing. *Anal. Chim. Acta***. 2011**, *707*, 1-2.

(28) Chen, J.; Fang, Z.; JieLiu, J.; Zeng, L. A Simple and Rapid Biosensor for Ochratoxin A Based on a Structure-Switching Signaling Aptamer. *Food Control* **2012**, *25*, 555-560.

(29) New England Biolabs. Cleavage Close to the End of DNA Fragments (oligonucleotides). https://www.neb.com/-/media/nebus/files/chart image/cleavage_olignucleotides_old.pdf?la=enhttp://sis.nlm.nih.gov/enviro.html (accessed Jul 23, 2017).

(30) Roux, K. H. Optimization and Troubleshooting in PCR. *Cold Spring Harb. Protoc.* **2009**, *4*, 1-6.

(31) Kang, J.; Lee, M-S.; Gorenstein, D. G. The Enhancement of PCR Amplification of a Random Sequence DNA Library by DMSO and Betaine: Application to In Vitro Combinatorial Selection of Aptamers. *J Biochem. Biophys. Methods* **2005***, 64,* 147-151*.*

(32) Mousavian, Z.; Sadeghi, H. M.; Sabzghabaee, A. M.; Moazen, F. Polymerase Chain Reaction Amplification of a GC Rich Region by Adding 1,2 Propanediol*. Adv. Biomed. Res.* **2014,** *3*, 65.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

Conclusions

We have described novel selection methods to generate enhanced functional nucleic acids for the development of aptamer based-biosensors. We have implemented a IP-SELEX method to generate self-alkylating ribozymes for fluorescent mRNA labeling. These ribozymes are capable of catalyzing the covalent attachment of a small-molecule fluorophore to itself. Further we show that the ribozyme can be fused onto a longer mRNA sequence and successfully react with the label in simulated physiological conditions. We anticipate these ribozymes will prove advantageous over previous noncovalent small-molecule binding aptamers for live cell labeling. We have designed a novel selection method for the generation of TNA aptamers to OTA. Our ability to utilize a TNA primer allowed for the elimination of contaminating DNA which aided in the efficient enrichment of TNA aptamers. This method enabled us to generate the first XNA aptamer of any kind to a small-molecule target. Further, we exhibit the impressive stability in the presence of nucleases and their ability to retain ligand-binding in these conditions where comparable DNA aptamers fail. Last, we introduce an innovative PCR-based SELEX method for the generation of SS biosensors to CoA. In this protocol, PCR

amplification serves as the step to selectively enrich functional sequences, obviating the need for a solid support for selection. We have optimized library design to include an EcoRI cut site to allow the efficient digestion of non-functional sequences. We anticipate this technique will be generalizable for any small-molecule target, and successfully generate the SS architecture that can be easily developed into a fluorescence assay for detection. These data support *in vitro* evolution as a powerful technique to generate molecular recognition elements, which can be expanded upon to attain aptamers with unique function and attributes. Through the methods discussed we have generated nucleic acids polymers with enhanced functionality such as nuclease resistance, covalent attachment of a label, and privileged architectures that hold great promise as biosensors. We are encouraged by the results we have attained thus far, and are excited to move forward into further optimization of these systems. Upon further studies, we are enthusiastic about the potential biosensing applications using the functional nucleic acids we have developed.

Future directions

Self-alkylating ribozymes

While we are excited by the results we have obtained with our current ribozymes we have developed, we have several modifications to the SELEX protocol for the generation of our second-generation ribozymes. There are some aspects of our original selection method that may have biased generation of ribozymes that have low affinity for the FIA. Small-molecule binding RNA aptamers typically have high nM to low μM range[1,2], and we hypothesize that ribozymes with high affinity

were lost in the selection process. We plan on reducing the random region form N_{116} to N_{50}, as this length has been shown to be successful in the generation of ribozymes.[3,4] This will result in shorter ribozymes that will be more functional for imaging as they will be less likely to perturb the target mRNA sequence. Additionally, we will decrease the concentration of fluorophore form 200 μM to 1 μM over the course of the selection to bias towards ribozymes that have low micromolar affinities. The IP step will be carried out in 0.1X PBS buffer and 2M urea to disrupt RNA-fluorophore binding and permit more efficient pull-down of sequences that bind tightly to the fluorophore.

In addition to modification to the SELEX protocol, we will generate second-generation ribozymes using alternative fluorophore, electrophiles, and linkers shown in Figure 5.1. We will use both fluorescein and malachite green (MG) as they have advantages for different applications. Fluorescein, which is constantly fluorescent, is ideal for pulse-chase experiments as it is necessary to verify fluorophore washout from cells. MG is fluorogenic[5] and will provide reduced background as signal will only be generated when bound to the target transcript. Wagoner, Bruchez, and coworkers have demonstrated that MG analogues with short side chains are cell permeable, do not bind nonspecifically in cells, and have low cytotoxicity.[6,7] Additionally, the emission wavelength of MG bound to activating RNA is in the range of 650-680 which will minimize signal overlap from cell autofluorescence.[8] We realize that the use of FIA was not well-suited for living cells as iodoacetamides are strong electrophiles that can react with molecules readily present in cells (i.e., thiols from proteins and glutathione,) and increase background

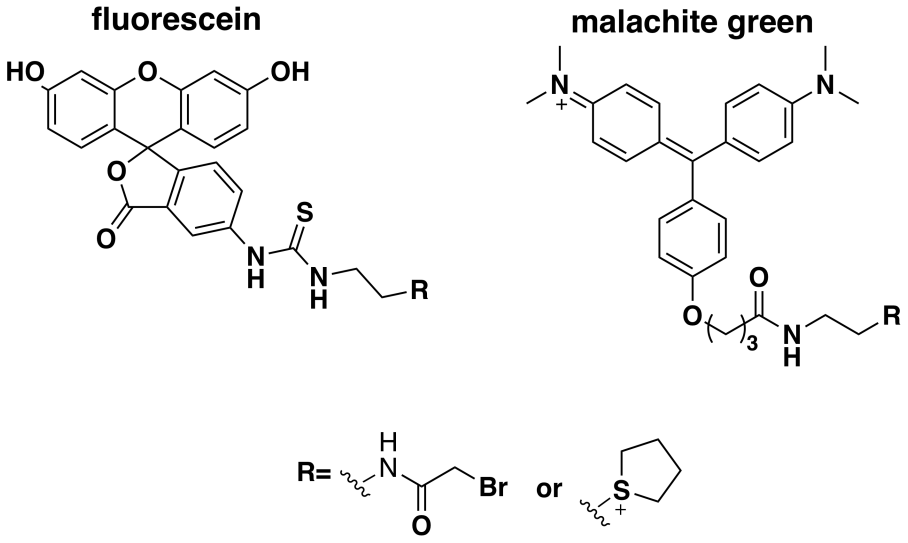


Figure 5.1. Structures of alternative fluorophores and electrophiles to be used for second-generation ribozyme selection

fluorescence. Thus, we will use bromoacetamide (BA) and sulfonium (SF) as alternative electrophiles. We were motivated to use BA as it is a weaker electrophile leading to attenuated reactivity and is anticipated to show selectivity for nitrogen nucleophiles.[9] The use of SF was bio-inspired by the role S-adenosylmethionine (SAM) plays in the alkylation of DNA and RNA,[10,11] however, only has minimal reactivity with cellular nucleophiles when unassisted by an enzyme. Further, we will increase the length of the linker between the fluorophore and electrophile in order to improve access to nucleophiles on the RNA.

We have previously discussed that our ribozymes generated by IP-SELEX are uniquely suited to enable identification transcript-specific RNA-binding proteins using our proposed TRIP methodology (Figure 5.2). Because our self-alkylating ribozymes covalently attach fluorescein to the target RNA transcript and can be efficiently captured by antibodies, they are ideal pull-down assays. New RNA reg-

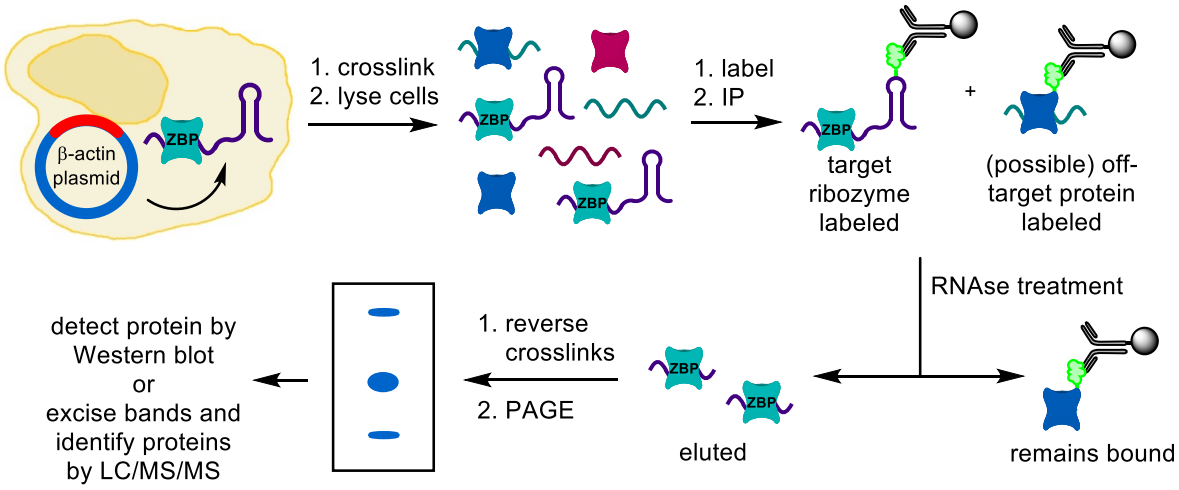


Figure 5.2. Transcript-specific immunoprecipitation (TRIP) to identify ZBP-1 bound to β-actin mRNA

ulatory mechanisms are often discovered by expression studies where the proteins responsible for the phenotype are unknown, and thus our proposed TRIP method holds great promise. Further, we have shown that ribozyme 5FR1 is able to self-label in the crosslinking conditions necessary for IP protocols. We aim to demonstrate pull-down of the well-studied zipcode-binding protein (ZBP-1), which is known to bind to the 54 nt "zipcode" sequence in the the 3′ UTR of β-actin mRNA.[12] Once ZBP-1 is recruited by β-actin mRNA, the transcript is localized to the leading edge of fibroblast cells.[12] As shown in Figure 5.2, we will fuse the ribozyme to β-actin mRNA and subsequently express this fusion in cells. These cells will then be incubated with formaldehyde to crosslink RNAs with their bound proteins, lysed, and then FIA will be added to the lysate. Antibody-functionalized magnetic beads will then be used to pull-down RNAs that were successfully labeled. RNA-protein complexes will be removed from the beads using RNAse and the crosslinks will be

reversed to dissociate the protein. The proteins will be analyzed by gel electrophoresis and Western blotting will be used to detect ZBP-1 or LC/MS/MS when the protein is unknown.

Nuclease-resistant TNA aptamers

In the short-term future of this project, we have ongoing research in the lab towards the development of a SS biosensor using our TNA aptamer A04T.2 for the detection of OTA. Analysis of OTA is currently accomplished by chromatographic methods such as thin-layer chromatography (TLC), HPLC, and GC-MS.[13] These techniques require training and specialized equipment, and extensive sample preparation in order to attain sensitive detection. There have been several reports of aptamer-based sensors for OTA, however, these require extensive modifications to generate signal and do not function in complex samples.[14,15] In the search for a fast a reliable alternative, Zeng and coworkers developed fluorescence-based assay for the detection of OTA using a SS-signaling DNA aptamer.[16] Using this method, they were able to detect OTA down to 0.8 ng/mL and in addition could recover toxin in corn samples. Despite this promise for detection in food samples, indication of prolonged exposure to OTA is measured in human blood serum/plasma where DNA will likely degrade.[17] We anticipate that a SS biosensor composed completely of TNA will be advantageous as the high stability will expand the potential matrices the biosensor will function in. We plan to initially construct CS strands of different lengths and optimize the site of hybridization on the aptamer in order to maximize fluorescence enhancement. While there is a possibility aptamer A04T.2 will not function as a biosensor as we did not directly select for

this architecture, we are encouraged that the DNA aptamer for OTA was engineered successfully post-selection for SS activity.

We hypothesize that the impressive stability of our TNA aptamers will be advantageous for real-world applications where natural nucleic acids are limited. As previously discussed, small-molecule toxins can pose a serious threat, however, simply detecting the concentration of toxin is insufficient. Because the biological response to toxins can occur quickly[18], prophylactic therapy which that could be given in advance to remedy potential exposure would be extremely beneficial. Macromolecular affinity reagents provide a promising approach as they can adsorb toxin that enters the bloodstream. Traditional receptors (antibodies, DNA/RNA) are problematic for this purpose as they are readily degraded by nucleases and proteases, therefore our TNA aptamers are a promising alternative due to their remarkable stability and low immunogenicity. We have described that these aptamers are capable of binding OTA with high affinity, however, they are not optimized for treatment alone due to limitations such as circulation lifetime and persistence of binding. We propose that the attachment of TNA aptamers to the interior and exterior of virus-like particles (VLPs) will be a powerful tool for toxin sequestration (Figure 5.3). VLPs have shown great promise over delivery vehicles as they can be easily functionalized and protect their cargo from degradation.[19] We anticipate the benefits of using VLPs will be two-fold: to enhance circulation lifetime and also allow multivalent display of the aptamer for more effective retention of sequestered toxin. We expect that the exterior aptamers will provide fast capture, and if the toxin becomes unbound over time it is likely to diffuse through the pores into the interior of the

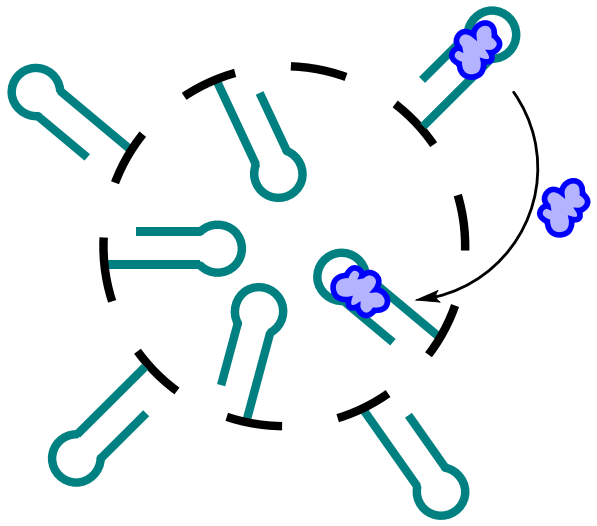


Figure 5.3. Rapid capture by surface aptamers on VLPs followed by diffusion to interior, which provides longer term sequestration.

VLP. The effective concentration of aptamers will be higher in the interior to enable long-term capture and sequestration. We will express the monomeric subunit of the cowpea chlorotic mottle virus (CCMV) having an azide-functionalized amino acid on the interior or exterior surface of each monomer (Figure 5.4). Alkyne-functionalized TNA aptamer will be conjugated to the monomers using copper catalyzed azide-alkyne cycloaddition (CuAAC). The aptamer modified monomers will be combined with unmodified monomers in varying ratios to assemble the VLPs *in vitro.* We will measure the binding affinity and kinetics of the aptamer-functionalized VLPs to gain insight into how aptamer behavior is altered once attached. We will also perform dialysis experiments to determine the rate of sequestration and persistence of binding of the VLPs in solutions of OTA. We expect that we will gain fundamental knowledge regarding binding capacity, sequestration rate, and mechanism of the aptamer-functionalized VLPs. We envision this will be a powerful method for advanced treatment of toxin exposure.

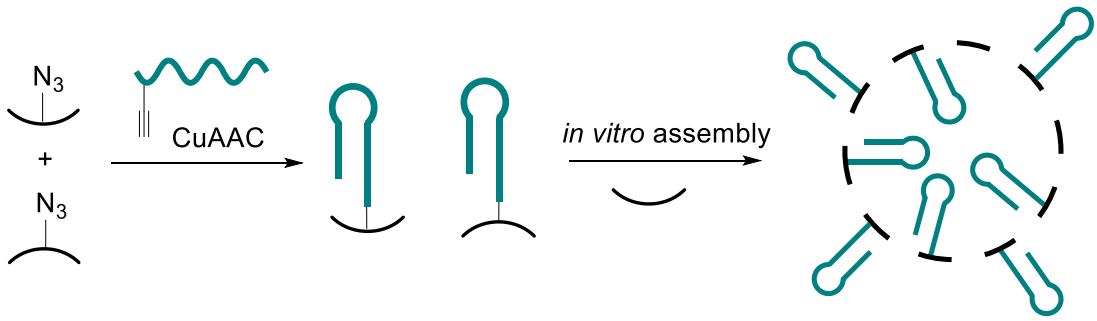


Figure 5.4. Synthesis and assembly of aptamer-VLP conjugates.

Structure-switching biosensors

Once we generate SS biosensors for CoA, we will integrate these biosensors into a HAT enzymatic assay with an FP readout (Figure 5.5). FP assays are the simplest fluorescence molecular binding assay because they only require labeling of one component and signal is not dependent on proximity between fluorophores. FP is traditionally used in a competitive format where the ligand displaces a small, fluorescent tracer from the receptor molecule, which causes a decrease in polarization. Using our SS biosensors, a fluorescently labeled CS will dissociate from the aptamer and the fluorophore will undergo a decrease in FP. We anticipate that optimizing the fluorophore attachment and relative sizes of the aptamer and CS will enable an appreciable FP signal in response to low levels of CoA and in the presence of excess AcCoA. FP has been accomplished using the direct measurement of fluorophore labeled aptamer, however, this method is limited to interactions where the target is much larger.[21] Because the rate of rotation of a molecule is highly dependent on molar mass of the molecule being measured, FP assays are more challenging for small molecule targets.[22] SS biosensors are advantageous for FP assays as the additive effects of changes to the fluorophore micro-

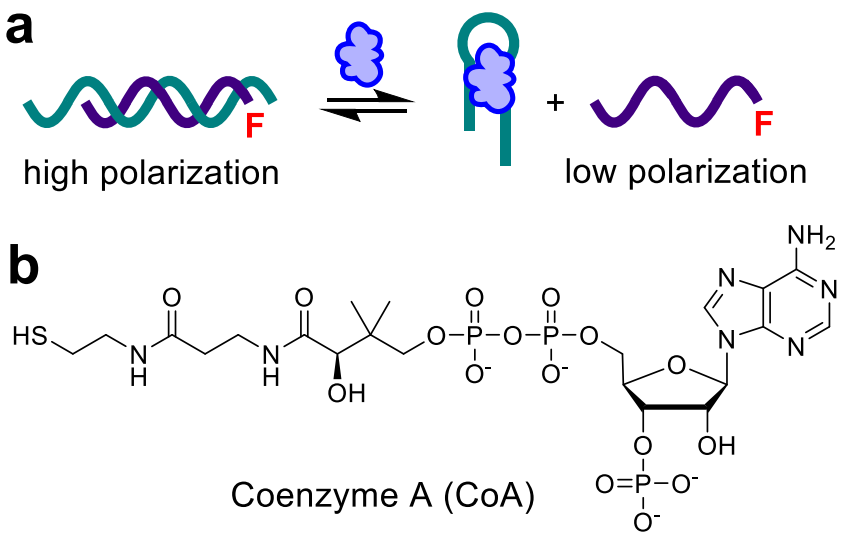


Figure 5.5. SS-biosensors for small-molecule detection. (a) Structure-switching biosensors provide a dose-dependent FP signal without labeling of target. (b) Chemical structure of Coenzyme A.

environment and size resulting from the binding-induced dissociation will enhance the polarization signal.[14] After identifying the best FP constructs using the CoA SS biosensors we selected using our PCR-based SELEX method, we will test these for the detection of purified HATs. There are a two HATs, including PCAF and CGN5, which have been validated in coupled enzyme assays for CoA detection, and will serve as a validation of the of our biosensors. These samples will be analyzed using a fluorescence plate reader, and standard curves will be constructed to convert FP signals to the amount of CoA formed.

Importantly, we anticipate that this method will be generalizable to enable the generation of SS biosensors to a variety of targets. Hence In the future we anticipate our SELEX method can be used to develop SS biosensors for small-molecule enantiopurity measurement to accelerate biocatalyst discovery. Stereoselective catalysis is vital to synthetic organic chemistry,[22] and while naturally

occurring enzymes remain the most efficient catalysts, they are not optimized for specific chemical transformations.[23] Highly selective enzymes can be evolved screening a large library of mutants to screen for those that provide the highest stereoselctivity.[24] This is usually analyzed by chiral chromatography or mass spectrometry, which are both limited in their sample throughput. Our group has recently shown that SS-biosensors composed of naturally occurring D-DNA and enantiomeric L-DNA can be used simultaneously to enable enantiopurity measurement in solution.[25] The principle of reciprocal chiral substrate selectivity states that an aptamer synthesized from opposite enantiomers of DNA will form mirror image complexes that bind to opposite enantiomers of a target with the same affinity and selectivity (Figure 5.6 a,b).[26] Our lab has recently developed enantiomeric sensors using the previously reported DNA SS biosensor for L-tyrosynamide (L-tym). As shown in Figure 5.6c, these sensors could be modified with spectrally orthogonal fluorophores to enable simultaneous measurement of L-tym and D-tym concentrations in the same solution.[24] These measurements can be performed in a microwell plate which has the potential to dramatically increase the throughput of analysis. Looking ahead to enzyme evolution, we were intrigued by 2-chloromandelic acid (2-CMA) as an interesting target since it is a key intermediate in the synthesis of the drug Plavix (Figure 5.7a).[26] Arnold and coworkers have demonstrated that P450 enzymes can be used to generate the S-enantiomer of mandelic acid esters with 57-94% enatiomeric excess (ee) (Figure 5.7c).[27] Our group has regarded 2-CMA as a compelling target to verify the power of our screening methods, as enzyme variants could potentially improve the enatioselectivity to >99 %ee. Past

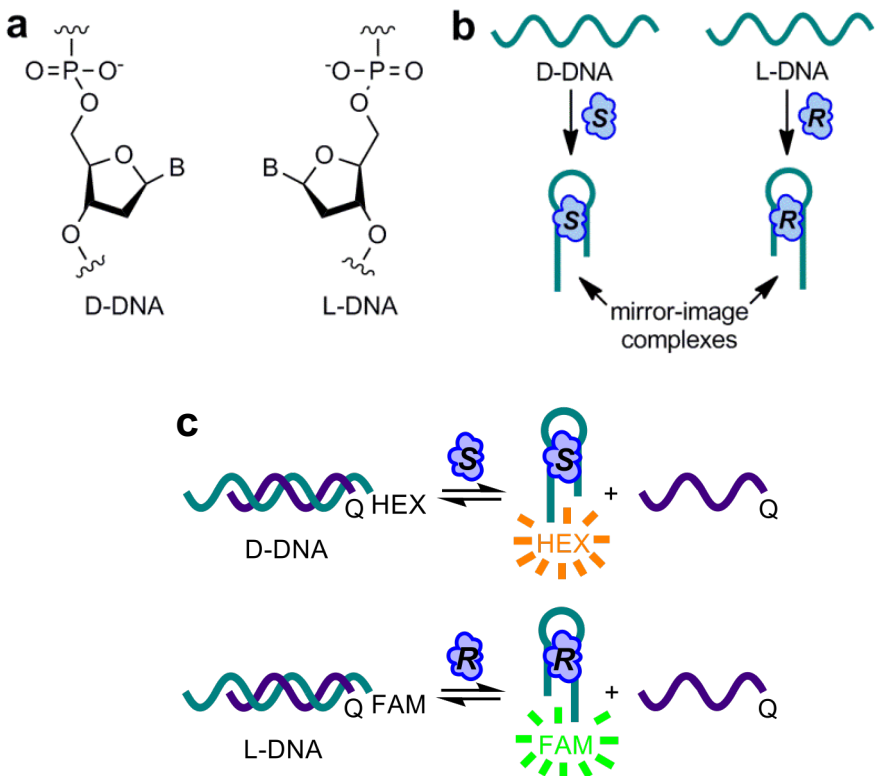


Figure 5.6. Enantiomeric DNA biosensors. (a) Chemical structure of *D*- and *L*-DNA. (b) *D*- and *L*-DNA aptamers each bind to a different enantiomer of the small-molecule target to produce mirror-image complexes. (c) Enantiomeric SS biosensors modified with orthogonal fluorophores enables simultaneous quantification of both enantiomers of a target.

members have generated aptamers that bind S-2-CMA over R-2-CMA with greater than 1000:1 selectivity using traditional bead-based selections.[28] Despite these encouraging results, none of these aptamers were able to produce a structure-switching response in the presence of target. We anticipate that our newly developed PCR-based SS-SELEX method will overcome this obstacle to successfully and reliably generate SS biosensors. The ability to develop biosensors with the desired architecture to a variety of targets has the potential to greatly accelerate enzyme evolution and the discovery of highly efficient catalysts via our robust enantiomeric aptamer assay.

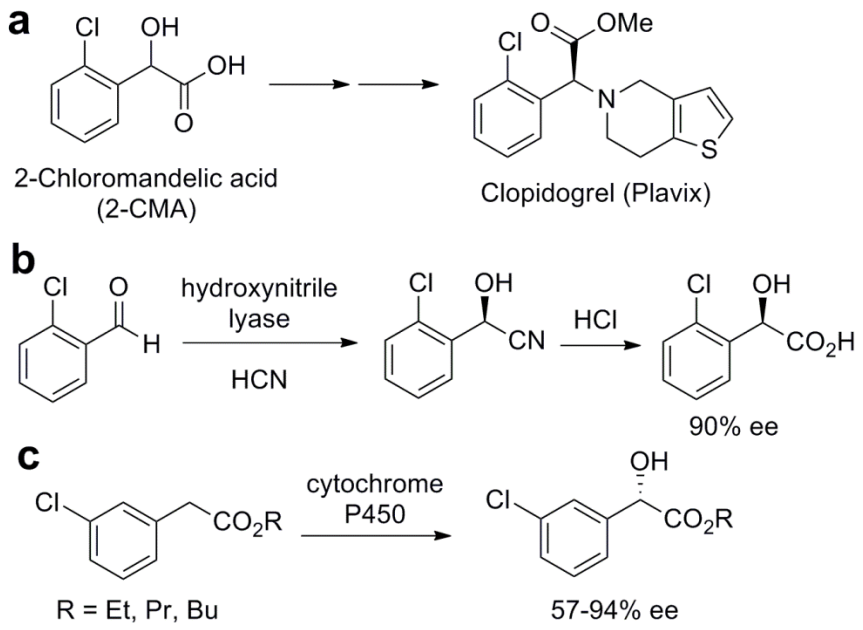


Figure 5.7. Synthetic schemes involving 2-chloromandelic acid. (a) 2-CMA is a key intermediate in the synthesis of clopidogrel. (b) Enzymatic synthesis of *R*-2-CMA using hydroxynitrile lyase. (c) Enzymatic synthesis of *S*-3CME using cytochrome *P450*.

References

(1) Famulok, M. Oligonucleotide Aptamers That Recognize Small Molecules. *Curr. Opin. Struct. Biol*. **1999**, *9*, 324-329.

(2) Osborne, S. E.; Ellington, A. D. Nucleic Acid Selection and the Challenge of Combinatorial Chemistry. *Chem. Rev*. **1997**, *97*, 349-370.

(3) Robertson, M. P.; Ellington, A. D. In Vitro Selection of Nucleoprotein Enzymes. *Nat. Biotechnol*. **2001**, *19*, 650-655.

(4) Jaeger, L. The New World of Ribozymes. *Curr. Opin. Struct. Biol*. **1997**, *7*, 324-335.

(5) Babendure, J. R.; Adams, S. R.; Tsien, R. Y. Aptamers Switch on Fluorescence of Triphenylmethane Dyes. *J. Am. Chem. Soc.* **2003,** *125*, 14716-14717.

(6) Fitzpatrick, J. A. J.; Yan, Q.; Sieber, J. J.; Dyba, M.; Schwarz, U.; Szent-Gyorgyi, C.; Woolford, C. A.; Berget, P. B.; Waggoner, A. S.; Bruchez, M. P. STED Nanoscopy in Living Cells Using Fluorogen Activating Proteins. *Bioconjugate Chem.* **2009,** *20*, 1843-1847.

(7) Szent-Gyorgyi, C.; Schmidt, B. F.; Creeger, Y.; Fisher, G. W.; Zakel, K. L.; Adler, S.; Fitzpatrick, J. A. J.; Woolford, C. A.; Yan, Q.; Vasilev, K. V.; Berget, P. B.; Bruchez, M. P.; Jarvik, J. W.; Waggoner, A. Fluorogen-Activating Single-Chain Antibodies for Imaging Cell Surface Proteins. *Nat. Biotechnol.* **2008,** *26*, 235-240.

(8) Saunders, M. J.; Szent-Gyorgyi, C.; Fisher, G. W.; Jarvik, J. W.; Bruchez, M. P.; Waggoner, A. S. Fluorogen Activating Proteins In Flow Cytometry for the Study of Surface Molecules and Receptors. *Methods* **2012,** *57*, 308-317.

(9) Carey, F. A.; Sundberg, R. J., *Advanced Organic Chemistry; Part A: Structure and Mechanisms*. 4th ed.; Kluwer Academic: New York, NY, 2000; pp 253-388.

(10) Jia, G.; Fu, Y.; He, C. Reversible RNA Adenosine Methylation In Biological Regulation. *Trends Genetics* **2013,** *29*, 108-115.

(11) Fontecave, M.; Atta, M.; Mulliez, E. S-Adenosylmethionine: Nothing Goes to Waste. *Trends Biochem. Sci.* **2004,** *29*, 243-249.

(12) Ross, A. F.; Oleynikov, Y.; Kislauskis, E. H.; Taneja, K. L.; Singer, R. H. Characterization of a B-Actin mRNA Zipcode-Binding Protein. *Mol. Cell. Biol.* **1997,** *17*, 2158-2165.

(13) Songsermsakul, P.; Razzazi-Fazeli, E. A Review of Recent Trends in Applications of Liquid Chromatography-Mass Spectrometry for Determination of Mycotoxins. *J. Liq. Chromatogr. Relat. Technol.* **2008,** *31*, 1641–1686

(14) Cruz-Aguado, J. A.; Penner, G. Fluorescence Polarization Based Displacement Assay for the Determination of Small Molecules with Aptamers. *Anal. Chem.* 2008, *80*, 8853–8855.

(15) Bonel, L.; Vidal, J. C.; Duato, P.; Castillo, J. R. An Electrochemical Competitive Biosensor for Ochratoxin A Based on a DNA Biotinylated Aptamer. *Biosens. Bioelectron.* **2011**, *26*, 3254–3259.

(16) Chen, J.; Fang, Z.; Liu, J.; Zeng, L. A Simple and Rapid Biosensor for Ochratoxin A Based on a Structure-Switching Signaling Aptamer. *Food Control*. **2012**, *25*, 555–560.

(17) Dohnal, V.; Dvořák, V.; Malíř, F.; Ostrý, V.; Roubal, T. A Comparison of ELISA And HPLC Methods for Determination of Ochratoxin A in Human Blood Serum in the Czech Republic. *Food Chem. Toxicol.* **2013,** *62*, 427–431.

(18) Holstege, C. P.; Bechtel, L. K.; Reilly, T. H.; Wispelwey, B. P.; Dobmeier, S. G. Unusual but Potential Agents of Terrorists. *Emerg. Med. Clin. North Am.* **2007,** *25*, 549–566.

(19) Rohovie, M. J.; Nagasawa, M.; Swartz, J. R. Virus-Like Particles: Next-Generation Nanoparticles for Targeted Therapeutic Delivery. *Bioeng. Trans. Med.* **2017,** *2*, 43–57.

(20) Zhao, Q.; Lv, Q.; Wang, H. Aptamer Fluorescence Anisotropy Sensors for Adenosine Triphosphate by Comprehensive Screening Tetramethylrhodamine Labeled Nucleotides. *Biosens. Bioelectron.* **2015,** *70*, 188-93.

(21) Knowles, W. S.; Sabacky, M. J. Catalytic Asymmetric Hydrogenation Employing a Soluble, Optically Active, Rhodium Complex. *Chem. Commun.* **1968**, 1445-1446.

(22) Reetz, M. T. Laboratory Evolution of Stereoselective Enzymes: A Prolific Source of Catalysts for Asymmetric Reactions. *Angew. Chem., Int. Ed.* **2011,** *50*, 138-174.

(23) Jayasena, S. D. Aptamers: An Emerging Class of Molecules That Rival Antibodies in Diagnostics. *Clin. Chem*. **1999**, *45*, 1628–1650.

(24) Reetz, M. T., *Evolutionary Methods in Biotechnology*. Brakmann, S.; Schwienhorst, A., Eds. Wiley-VCH: Weinheim, 2004.

(25) Feagin, T. A.; Olsen, D. P. V.; Headman, Z. C.; Heemstra, J. M. High-Throughput Enantiopurity Analysis Using Enantiomeric DNA-Based Sensors. *J. Am. Chem. Soc.* **2015,** *137*, 4198-4206.

(26) Klussmann, S.; Nolte, A.; Bald, R.; Erdmann, V. A.; Furste, J. P. Mirror-Image RNA That Binds D-adenosine. *Nat. Biotechnol.* **1996,** *14*, 1112-1115.

(27) Breuer, M.; Ditrich, K.; Habicher, T.; Hauer, B.; Kebeler, M.; Sturmer, R.; Zelinski, T. Industrial Methods for the Production of Optically Active Intermediates. *Angew. Chem., Int. Ed.* **2004,** *43*, 788-824.

(28) Landwehr, M.; Hochrein, L.; Otey, C. R.; Kasrayan, A.; Backvall, J.- E.; Arnold, F. H. Enantioselective r-Hydroxylation of 2-Arylacetic Acid Derivatives and Buspirone Catalyzed by Engineered Cytochrome P450 BM-3. *J. Am. Chem. Soc.* **2006,** *128*, 6058-6059.

(29) Feagin, T. A.; Heemstra, J. M. unpublished results.