UC Irvine
UC Irvine Electronic Theses and Dissertations

Title
Towards the In Vitro Selection of Therapeutic Aptamers

Permalink
https://escholarship.org/uc/item/1ch1j4tf

Author
McCloskey, Cailen Marie

Publication Date
2021

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,
IRVINE

Towards the In Vitro Selection of Therapeutic Aptamers

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

In Pharmaceutical Sciences

by

Cailen Marie McCloskey

Dissertation Committee:
Professor John C. Chaput, Chair
Professor Robert Spitale
Professor Andrej Lupták

2021

Portions of Chapter 1© 2017 Springer Nature, 2017 and 2010 American Chemical Society, 2016 Elsevier
Portion of Chapter 2© 2020 American Chemical Society
Portion of Chapter 4© 2019 American Chemical Society

DEDICATION

To

My parents, Michael and Diane, my brother Ian,
and my partner Kegan, for their boundless support and love.

TABLE OF CONTENTS

ABBREVIATIONS

°C	degrees Celsius
2-Me-imidazole	2-methylimidazole
3D	Three-dimensional
5-I	Five iodo-
A	Adenine
AB	Aptamer Binding Buffer
ABT	Aptamer Binding Buffer with 0.005% Tween 20
Ac_2O-DMAP	Acetic anhydride-dimethylaminopyridine
AMD	Age-related Macular Degeneration
AMP	Adenosine Monophosphate
Aq.	Aqueous
ATP	Adenosine Triphosphate
Avg	Average
BLI	Biolayer Interferometry
$BnNH_2$	Benzylamine
bp	base pair
BSA	Bovine Serum Albumin
Bst	Bacillus stearothermophilus
C	Cytosine
C5	Carbon in the five position
$CaCl_2$	Calcium chloride
$CaCO_3$	Calcium carbonate
CH_3CN	Acetonitrile
CO	Carbon monoxide
COVID-19	Coronavirus Disease 2019
Cq	quantification cycle
Cryo-EM	Cryogenic Electron Microscopy
CuBr	Copper (I) Bromide
D-	Dextrorotatory
DCM	Dichloromethane
ddPCR	Digital Droplet Polymerase Chain Reaction

DB	Denaturing Buffer
DIBAL-H	Di-isobutyl aluminum hydride
DMA	Dimethylacetamide
DMAP	4-Dimethylaminopyridine
DME	Dimethoxyethane
DMF	Dimethylformamide
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dNTP(s)	Deoxyribonucleic Acid Triphosphate(s)
ds	double stranded
dT	Deoxythymidine
DTT	Dithiothreitol
dU	Deoxyuracil
dUTP	Deoxyuracil triphosphate
EDTA	Ethylenediaminetetraacetic Acid
ELISA	Enzyme-Linked Immunosorbent Assay
EMSA	Electrophoretic Mobility Shift Assay
ESI	Electrospray Ionization
Et_3N	Triethylamine
EtOH	Ethanol
FANA	2’ Fluoro-arabinonucleic Acid
FDA	U.S. Food and Drug Administration
f_{hetero}	Fraction of hetero-duplexed DNA
FLAG	Peptide Sequence DYKDDDDK
Fnorm	Normalized Fluorescence
FT	Flow Through
G	Guanine
xg	Units of times gravity
h	Hour(s)
H_2	Hydrogen Gas
H_2O	Water
H_2O_2	Hydrogen Peroxide

HA	Hemagglutinin
HFNAP(s)	Highly Functionalized Nucleic Acid Polymer(s)
HIV	Human Immunodeficiency Virus
HIV RT	Human Immunodeficiency Virus Reverse Transcriptase
HNA	1,5-Anhydrohexitol Nucleic Acid
IDT	Integrated DNA Technologies
IR680	IRDye 680RD Infrared Dye
IR800	IRDye 800CW Infrared Dye
J	Joules
K^{-1}	per Kelvin
KCl	Potassium Chloride
K_D	Dissociation Constant
kJ	Kilojoules
k_{off}	Off-Rate Constant
k_{on}	On-Rate Constant
L-	Levorotatory
LB-amp	Luria-Bertani Broth with Ampicillin
mAb(s)	Monoclonal Antibody(ies)
MeOH	Methanol
mg/mL	Milligrams per Milliliter
$MgCl_2$	Magnesium Chloride
min	Minute(s)
mL	Milliliter
mM	Millimolar
$MnCl_2$	Manganese Chloride
mol	Mole(s)
Mol^{-1}	per Mole
M	Molar
mRNA	Messenger Ribonucleic Acid
MST	Microscale Thermophoresis
n	Sample Size
N’	85% Parent Nucleotide, 15% Other 3 Nucleotides

N_{40}	40 Random Nucleotides in a Row
NaCl	Sodium Chloride
NaOAc	Sodium Acetate
ng/ml	Nanogram per Milliliter
NGS	Next Generation Sequencing
NH_2OH	Hydroxylamine
NH_4OH	Ammonium Hydroxide
Ni-NTA	Nickel-Nitrilotriacetic Acid Resin
nM	Nanomolar
nm	Nanometer
nmol	Nanomole
nt	Nucleotide(s)
OH	Hydroxyl
P	Phosphorous
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Primer Binding Site
PCR	Polymerase Chain Reaction
$Pd(Ph_3)_4$	N-Phenyl-bis(trifluoromethanesulfonimide)
Pd/C	Palladium on Carbon
PDB	Protein Database
PEG	Polyethylene glycol
pH	Potential of Hydrogen
Phe	Phenylalanine
pM	Picomolar
PPh_3	Triphenylphosphine
qPCR	Quantitative Polymerase Chain Reaction
r.t.	Room Temperature
RBD	Receptor Binding Domain
RNA	Ribonucleic Acid
rpm	Rotations per Minute
RT	Reverse Transcriptase
R	Ideal Gas Constant

T	Temperature
ln	Natural Logarithm
S1	S1 protein domain of SARS-CoV-2
SARS-CoV	Sudden Acute Respiratory Syndrome Coronavirus
sec	Second(s)
SOMAmer	Slow Off-rate Modified Aptamer
ss	Single Stranded
SW	Stringent Wash
SVPE	Snake Venom Phosphodiesterase
T	Thymine
t-BuOH	Tertiary Butanol
T3	Escherichia Virus T3
T4	Escherichia Virus T4
T7	Escherichia Virus T7
tATP	TNA Adenosine Triphosphate
TBDPSCl	Tert-Butyl(chloro)diphenylsilane
TBE	Tris Borate EDTA
TBTA	Tris(benzyltriazolylmethyl)amine
tCTP	TNA Cytidine Triphosphate
tGTP	TNA Guanosine Triphosphate
THF	Tetrahydrofuran
TMSOTf	Trimethylsilyl Triflate
TNA	Threose Nucleic Acid
TNF	Tumor Necrosis Factor
tNTP(s)	Threose Nucleic Acid Triphosphates
Tris-HCl	Tris Hydrochloride
Trp	Tryptophan
TTP	Thymidine Triphosphate
tTTP	TNA Thymidine Triphosphate
tUTP	TNA Uridine Triphosphate
$tUTP^{Phe}$	Phenyl Modified TNA Uridine Triphosphate
$tUTP^{Trp}$	Tryptophan Modified TNA Uridine Triphosphate

μg/ml	Microgram per Milliliter
μL	Microliter
μM	Micromolar
U	Unit(s)
UV	Ultraviolet
V/V	Volume by Volume
VEGF	Vascular Endothelial Growth Factor
W	Watts
w/V	Weight by Volume
XNA(s)	Xenobiotic Nucleic Acid(s)
$ZnCl_2$	Zinc Chloride
ΔG°	Gibbs Free Energy at Standard Conditions

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank my advisor, Professor John Chaput, for taking a chance on a lost transfer student and providing me the wonderful opportunity to perform this research. With your guidance I've grown so much as a scientist and pushed myself further than I thought I could. You've created a near perfect environment for any student to thrive.

I would like to thank my committee members, Professor Andrej Lupták and Professor Robert Spitale for their mentorship, attention, and discourse over the years. I am grateful for your perspective.

I would like to thank all the current and former members of the Chaput Lab for all your advice, discussion, and challenges over the years. I am especially grateful to Professor Yajun Wang, Dr. Nick Chim, Dr. Saikat Bala, Eric Yik and Arlene Ngor for the insightful discussions, indispensable troubleshooting sessions, guidance, and support.

I would like to thank those I've had the privilege to collaborate with over the years, including Dr. Ivan Grubisic, Dr. Val Cavett and Professor Brian Paegel. You've expanded my horizons. I would like to also thank the Genomics High Throughput Facility at UCI, especially Dr. Melanie Oakes and Charlotte Spassoff, who worked with us so closely to get through the stumbling blocks we found along the way. You made this work possible.

I would also like to thank Professor Ryan Hili and the members of the Hili lab at UGA who introduced me to this field in the first place and for all the fun.

I would like to thank my friends and family, for their understanding, love, and support.

Finally, I would like to thank Trevor Slawson, with whom I started this journey, and whom I still miss.

Thank you to those sources of financial support awarded to the Chaput Lab and provided by University of California, Irvine for the partial funding of my doctoral research. DARPA Folded Non-Natural Polymers with Biological Function Fold F(x) Program under award number N66001-16-2-4061, NSF under grant number 1607111 and MCB: 1946312, and our unnamed commercial partner.

VITA

EDUCATION

University of California, Irvine **August 2017-present**

- Ph.D. in Pharmaceutical Sciences, Spring 2021
- Dissertation title: Towards the In Vitro Selection of Therapeutic Aptamers

Barrett, the Honors College at Arizona State University **August 2011-May 2015**

- B.S. in Biochemistry, *cum laude*
- B.S.E. in Chemical Engineering, *cum laude*
- Honors thesis: The Effect of Polymer Film Roughness on the Performance of Polyamide Reverse Osmosis Membranes

RESEARCH EXPERIENCE

Laboratory for Chemical and Synthetic Biology **August 2017-present**

University of California, Irvine in Irvine, CA

Advisor: John C. Chaput

- Optimized a DNA display strategy for the in vitro selection of TNA aptamers.
- Developed streamlined workflows to validate machine learning predicted aptamers, from single sequences to 100K at a time.
- Pioneered the use of TNA as a novel soft material for digital information storage.
- Developed a method to ligate TNA oligomers using commercial enzymes.

Laboratory for Molecular Evolution **November 2015-July 2017**

University of Georgia in Athens, GA

Advisor: Ryan Hili

- Explored fundamentals related to the extension of synthetic polymers in a sequence-defined manner, fabricated small PEG oligomers via DNA backbones.
- Optimized recombinant expression and crystallization conditions for T4 DNA ligase.

Center for Innovations in Medicine **August 2014-May 2015**

Arizona State University in Tempe, AZ

Advisor: Chris Diehnelt

- Characterized binding kinetics of peptide synthetic antibodies (synbodies).

Advanced Membrane Materials Lab **April 2013-May 2015**

Arizona State University in Tempe, AZ

Advisor: Mary Laura Lind

- Synthesized, tested, and characterized thin film nanocomposite membranes and latex thin film composite membranes to improve desalination in a custom reverse osmosis cross-flow system.
- Explored the utility of graphene oxide in thin film nanocomposite membranes.

NanoVoltaics Inc. **February-September 2014**

Phoenix, AZ

Advisors: Chang Gong Wang and Mary Laura Lind

- Collaborated with NanoVoltaics Inc. to compare their energy efficient nanoparticles against commercial Linde Type A zeolites in thin film nanocomposite membranes.

TEACHING EXPERIENCE

Teaching Assistant (University of California, Irvine)

- PHRMSCI 90 – Speaking Science — Spring 2020
- PHRMSCI 177L - Medicinal Chemistry Laboratory — Winter 2018

Teaching Assistant (University of Georgia)

- BCMB 3600 - Genomics and Bioinformatics — Spring 2017
- CHEM 1211L - Freshman Chemistry Laboratory I — Fall 2016
- BCMB 3100 - Introductory Biochemistry and Molecular Biology — Fall 2016

Teaching Assistant (Arizona State University)

- CHM 113R - General Chemistry I Recitation — Fall 2012

AWARDS AND GRANTS

Innovative and Interdisciplinary Research Grant — Summer 2017
Barrett Thesis Funding — Spring 2015
NASA Space Grant — Fall 2014-Spring 2015

PUBLICATIONS

6. **C. McCloskey***, Q. Li*, N. Chim, E. Yik, A. Ngor, E. Medina, J. C. Chaput. Evolution of Functionally Enhanced TNA Aptamers. *In preparation.*
5. K. Yang, **C. McCloskey**, J. C. Chaput. Reading and Writing Digital Information in TNA. *ACS Synth Biol.* September 23, 2020. Articles ASAP.
4. M. R. Dunn, **C. McCloskey**, P. Buckley, K. Rhea, J. C. Chaput. Generating Biologically Stable TNA Aptamers that Function with High Affinity and Thermal Stability. *J Am Chem Soc.* April 16 2020. 142, 17, 7721-7724.
3. **C. McCloskey**, J. Liao, S. Bala, J. C. Chaput. Ligase-Mediated Threose Nucleic Acid Synthesis on DNA Templates. *ACS Synth Biol.* February 15 2019, 8 (2), 282-286.
2. H. Mei, J. Liao, R. M. Jimenez, Y. Wang, S. Bala, **C. McCloskey**, C. Switzer, and J. C. Chaput. Synthesis and Evolution of TNA Libraries Bearing 7-Deaza-7-Substituted Guanosine Analogues. *J Am Chem Soc.* April 18 2018, 140, 17, 5706–5713.
1. P. Cay-Durgun, **C. McCloskey**, J. Konecny, A. Khosravi, M. L. Lind. Evaluation of thin film nanocomposite reverse osmosis membranes for long-term brackish water desalination performance. *Desalination*. 17 February 2017, 404, 304–312.

POSTERS AND PRESENTATIONS

C. McCloskey, J. C. Chaput. Evolution of Functionally Enhanced TNA Aptamers. Paper presented: 2021 February Synthetic and Chemical Biology Club. February 26, 2021. Irvine, CA.

C. McCloskey, J. C. Chaput. Closing the Gap on Therapeutic Aptamers. Paper presented: 2019 Vertex Day. March 7, 2019. Irvine, CA.

C. McCloskey, J. C. Chaput. Ligase-Mediated TNA Synthesis. Paper Presented: 2018 Vertex Day. March 8, 2018. Irvine, CA.

C. McCloskey, Z. A. Tolchin, R. Hili. Development of Sequence-Defined Synthetic Polymers. Poster presented: Southeastern Chemical Biology Symposium. April 14, 2017. Athens, GA.
C. McCloskey, R. Hili. DNA-Templated Synthetic Polymer Production Using T4 DNA Ligase. Poster presented: UGA Chemistry Department Research Retreat. September 3, 2016. Athens, GA.
C. McCloskey, M. L. Lind. The Effect of Polymer Film Roughness on the Performance of Polyamide Reverse Osmosis Membranes. Paper presented: 24th Annual Arizona/NASA Undergraduate Research Symposium. May 2015. Tempe, AZ.
C. McCloskey, P. Cay-Durgun, M. L. Lind. The Effect of Polymer Film Roughness on the Performance of Polyamide Reverse Osmosis Membranes. Poster presented: 2015 ASU/NASA Space Grant Poster Session. Tempe, AZ.

VOLUNTEER INVOLVEMENT

Audiobook Narrator and Proof Listener with Learning Ally	April 2017-present
Women in Science Mentoring Program	August 2016-July 2017
Georgia Science and Engineering Fair Judge	April 2016, 2017
Georgia Junior Science and Humanities Symposium Reader	January 2016, 2017
Volunteer Teacher with Science Detectives	August 2013-May 2015
Hospital Volunteer and Club President with ASU Muralcles	August 2011-May 2015

ABSTRACT OF THE DISSERTATION

Towards the In Vitro Selection of Therapeutic Aptamers

by

Cailen Marie McCloskey

Doctor of Philosophy in Pharmaceutical Sciences

University of California, Irvine, 2021

Professor John C. Chaput, Chair

Aptamers are a class of nucleic acid molecules that mimic antibodies by folding, in a sequence specific manner, into 3D structures that can bind a target. Vast random libraries, routinely around 10^{15} nucleic acid sequences undergo iterative cycles of binding their target, stringent washing to remove weak and non-binders, elution, and amplification to become enriched in sequences with high binding affinity to the target. Aptamers have many benefits over other drug archetypes, such as their ease of discovery, reproducibility due to their being chemically synthesizable, reversible folding and thermal stability, and low cost relative to antibodies. This makes them attractive drug candidates, but aptamers have been largely relegated to diagnostic settings due to the poor biological stability of natural nucleic acids.

The field has long been dominated by RNA and DNA aptamers, which boast an ease of discovery born of their origins in nature. However, recent advancements in polymerase engineering have enabled the discovery of polymerases with the ability to transcribe DNA into xeno nucleic acids (XNAs), synthetic nucleic acids with backbones unrelated to DNA or

RNA. Although these engineered variants function with reduced activity when compared to their natural substrate, the ability to transfer information from one genetic polymer to another has had enabled the selection of XNA aptamers, among other technologies. One such XNA, threose nucleic acid or TNA, is unique in that it can form stable antiparallel duplexes with itself or DNA, while being completely recalcitrant to nuclease digestion. This makes TNA well suited for therapeutic use, as it will be more stable in the body.

In this dissertation, we first developed a method to allow the selection of TNA aptamers covalently linked to their encoding DNA using an engineered TNA polymerase. Next, aptamer chemical diversity was expanded with the addition of base-modified TNA nucleotides. Finally, the scalability of TNA aptamer production was explored through TNA ligation with natural ligases.

Chapter 1 is a review of literature pertaining to natural nucleic acid and xeno nucleic acid aptamers, outlining what has been achieved and what remains to be overcome. Specifically, we focus on TNA aptamer selection, and describe a foundational approach to TNA aptamer selection, DNA display, which is utilized in Chapters 2-3. This new strategy borrows from previous work on mRNA display to overcome one of the biggest stumbling blocks to XNA aptamer selection: library regeneration after selection. The resultant TNA aptamers are characterized to reveal picomolar binding affinity, a magnitude previously not known to be achievable for TNA aptamers.

The next chapter details the discovery of a TNA aptamer to HIV RT with picomolar affinity. While TNA's biological stability was previously known, this TNA aptamer demonstrates high thermal stability of TNA molecules that is expected from DNA and RNA. The use of an engineered TNA polymerase and a display strategy that provides a powerful

genotype–phenotype linkage allowed the discovery of aptamers ranging in affinity from ~0.4-4 nM. The TNA aptamers remain intact and capable of binding in the presence of nucleases. The combined properties of biological stability, high binding affinity, and thermal stability make TNA aptamers a powerful system for the development of diagnostic and therapeutic agents.

Chapter 3 describes work to inject greater chemical diversity into aptamer selections by incorporating amino-acid-like side chains onto standard uracil bases. This strategy parallels work done by Mayer, Liu, and others to expand chemical diversity in aptamers. The base and backbone modified aptamers are referred to as 'threomers', which bind with lower K_Ds and slower off rates than their unmodified TNA counterparts. Kinetic measurements reveal that side chain modifications are critical for generating threomers with slow off-rate binding kinetics. These findings expand the chemical space of evolvable non-natural genetic systems to include functional groups that enhance protein target binding by mimicking the structural properties of traditional antibodies. A detailed investigation of the impact benzyl and indole displaying bases have on selection outcomes are reported here.

Chapter 4 describes initial work to increase the scale of synthesizable TNA through a combination of solid phase TNA synthesis and ligation with natural ligases. A thorough exploration of ligases and reaction conditions with TNA-TNA and DNA-TNA ligations led to the identification of T3 DNA ligase as the most active on TNA substrates. This knowledge could serve as a starting point for future directed evolution of a TNA ligase.

CHAPTER 1

1.1 Abstract

Nucleic acid aptamers rival antibodies in terms of selectivity and specificity, but they have been limited by the use of natural nucleic acids which are not biologically stable. Xenobiotic nucleic acids (XNAs) are inherently biologically stable architectures, several of which are capable of Darwinian evolution. One of these, threose nucleic acid, has notable properties and functions: 1) is recalcitrant to nuclease digestion, 2) base pairs efficiently with DNA, and 3) can be synthesized efficiently by an engineered polymerase. Improved selection strategies, such as the incorporation of chemically diverse base modifications, could help TNA aptamers realize their therapeutic potential.

1.2 Natural Nucleic Acid Aptamers

Aptamers are nucleic acid molecules capable of folding into a three-dimensional conformation to bind a target of interest, such as a small molecule or protein (Figure 1.1). Since their discovery, aptamers have demonstrated the ability to compete directly with proteinaceous antibodies in terms of binding affinity and selectivity (Dunn et al., 2017; Ellington & Szostak, 1990; Tuerk & Gold, 1990). Aptamers even claim a few key benefits over antibodies. Among these, aptamers have greatly improved thermal stability. Antibody drugs require a cold chain, in which the distribution pathway from manufacturer to consumer requires the antibodies must be kept frozen or they face irreversible denaturation. Nucleic acids, however, fold reversibly and therefore do not require the same strictness of supply chain. Another key benefit is that aptamers are chemically synthesized,

which cuts discovery costs and increases the ease by which production can be scaled for widespread use. This also reduces batch-to-batch variability, which is a major issue faced by commercial antibodies in both research and therapeutic applications (Baker, 2015). However, antibodies remain the gold standard of therapeutic care.

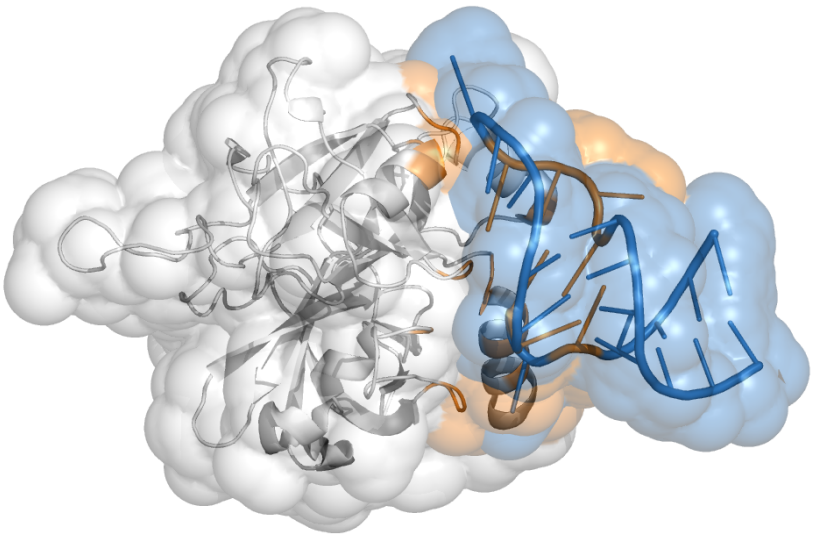


Figure 1.1. RNA aptamer (blue) with interacting residues (orange) in complex with thrombin target (white). Figure adapted from PDB: 3DD2, (Long et al., 2008).

Although aptamers hold great promise as therapeutics, their use in medicine has been largely restricted to diagnostics. Only one aptamer has ever successfully run the gauntlet to garner FDA approval. This RNA aptamer drug, named Macugen, was developed as an antagonist of VEGF for the treatment of age-related macular degeneration (AMD) (Ng et al., 2006). The lack of aptamers in the therapeutic arena relative to their abundance in research is likely due to the poor biological stability of natural nucleic acids, which have half-lives in human serum ranging from seconds for single stranded RNA to about an hour for DNA (Kratschmer & Levy, 2017). Biological half-lives can be extended with the

inclusion of chemical groups that disguise the DNA from nuclease recognition. An example of this would be the incorporation of a phosphorothioate or phosphorodithioate linkage in the backbone. In fact, careful placement of this linkage has an added benefit in that it can increase affinity significantly, as demonstrated by Abeydeera and colleagues in which a single phosphorothioate linkage was able to increase target affinity 1000-fold (Abeydeera et al., 2016). However, this linkage can increase nonspecific binding, and fully modified oligonucleotides have shown toxicity in animal models (Ni et al., 2017).

Another strategy involves electronegative chemical modifications in place of the 2' hydroxyl on ribose, such as 2'-fluoro, or 2'-methoxy groups (Shigdar et al., 2013). This was successful for other nucleic acid therapeutics, such as antisense oligonucleotides. In the case of Macugen, the aptamer was first selected in RNA, then the ribose sugars of the parent sequence were carefully modified, which extended the biological half-life to about 10 days. Macugen also required an inverted dT cap at the 3' end for further exonuclease protection. While this first-in-class drug proves aptamers can make good therapeutics, the extensive process of replacing every purine with the 2'-methoxy variant and every pyrimidine with the 2'-fluoro variant can perturb the carefully selected structural characteristics of an aptamer, and is not a feasible route for all aptamers.

In the late 1990s, a strategy for discovering biologically stable aptamers made from L-ribonucleic acid was invented (Klussmann et al., 1996; Vater & Klussmann, 2015). Aptamer selection is typically performed by challenging a vast pool of random sequences to bind a target, partitioning the binders away from nonbinders, and then amplifying by PCR to regenerate the pool of sequences (Figure 1.2). Over the course of many rounds of selection, the pool becomes enriched in sequences that bind the target, which can then be identified

by Next Generation Sequencing (NGS) or Sanger sequencing. In this way, aptamers can be generated for any conceivable target. Taking advantage of this straightforward discovery pathway, natural D-RNA aptamers were selected against a mirror image version of the target of interest (with D-amino acids). Termed "Spiegelmers", for the German word for mirror, the resultant aptamers were then synthesized from L-RNA and shown to retain affinity to their selected target when presented in its natural L-amino acid form. While Spiegelmers exhibit greatly improved biological stability, the range of pursuable targets is limited to those that can be chemically synthesized as mirror images. Through a combination of solid phase peptide synthesis and native chemical ligation, the longest chemically synthesized protein yet achieved is only 358 amino acids (Jiang et al., 2017). In order to expand the range of druggable targets, it is necessary to develop a biologically stable framework capable of Darwinian evolution to discover therapeutic aptamers.

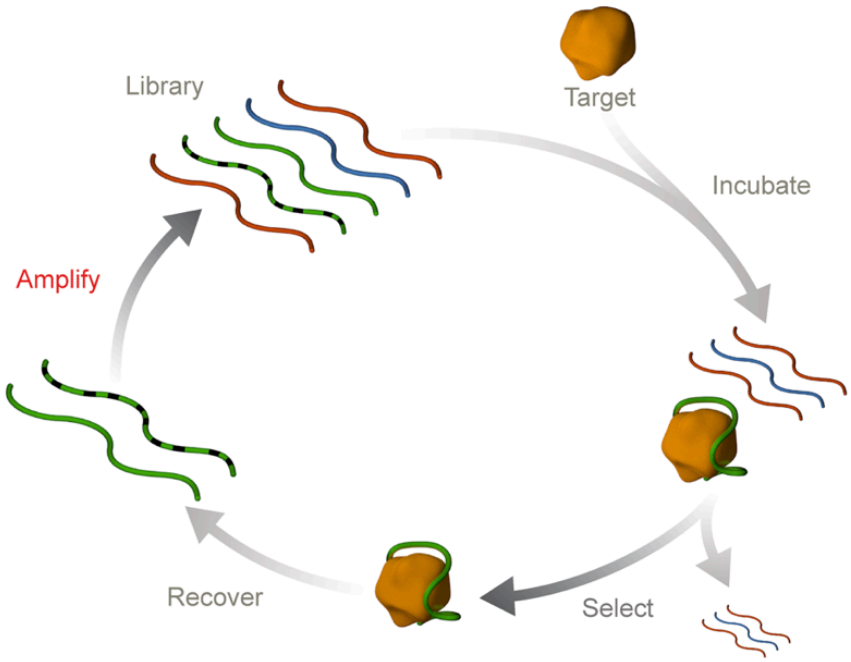


Figure 1.2. General aptamer selection cycle. Figure adapted with permission from *Nat. Rev. Chem.* 2017, 1, 0076.

1.3 Threose Nucleic Acid and other Xenobiotic Nucleic Acids

1.3.1 Xenobiotic Nucleic Acids

In concert with the development of increasingly better aptamer selection methods is the invention of many synthetic nucleic acids, also called xenobiotic nucleic acids (XNAs). XNAs differ from natural or chemically modified DNA or RNA in that they have a unique chemical moiety in place of the ribose or phosphate. This moiety can theoretically be anything, but only a small subset of possible chemical modifications retains the ability to base pair with natural nucleic acid and allow for the transfer of genetic information. Among these, select few genetic polymers have enzymes, natural or engineered, that allow for efficient synthesis. Recognition by natural enzymes can be a double-edged sword, however, where nucleic acids that are more easily incorporated by polymerases often also exhibit greater risk of degradation by nucleases. This tradeoff is illustrated by the comparison of polymerization and nuclease degradation of two common XNAs: 2'-fluoro-arabino nucleic acid (FANA) and threose nucleic acid (TNA) in Figure 1.3. While the first generation engineered TNA polymerase, Kod-RI, and FANA polymerase, Tgo-D4K, show similarly reduced activity on DNA substrates, Tgo-D4K demonstrates considerably greater activity on FANA monomers than Kod-RI does on TNA monomers (Nikoomanzar et al., 2017). Impressively, it was later shown that the rate of wild type Tgo polymerase on FANA triphosphates is 15 nt/min, comparable to the rate of the evolved polymerases on DNA triphosphates (Wang et al., 2018). However, when comparing nuclease resistance, the TNA oligonucleotide showed no sign of degradation in human liver microsomes and human serum (Figure 1.3C). Conversely, in the presence of human liver microsomes, the FANA oligonucleotide degraded completely by three hours of incubation (Culbertson et al., 2016).

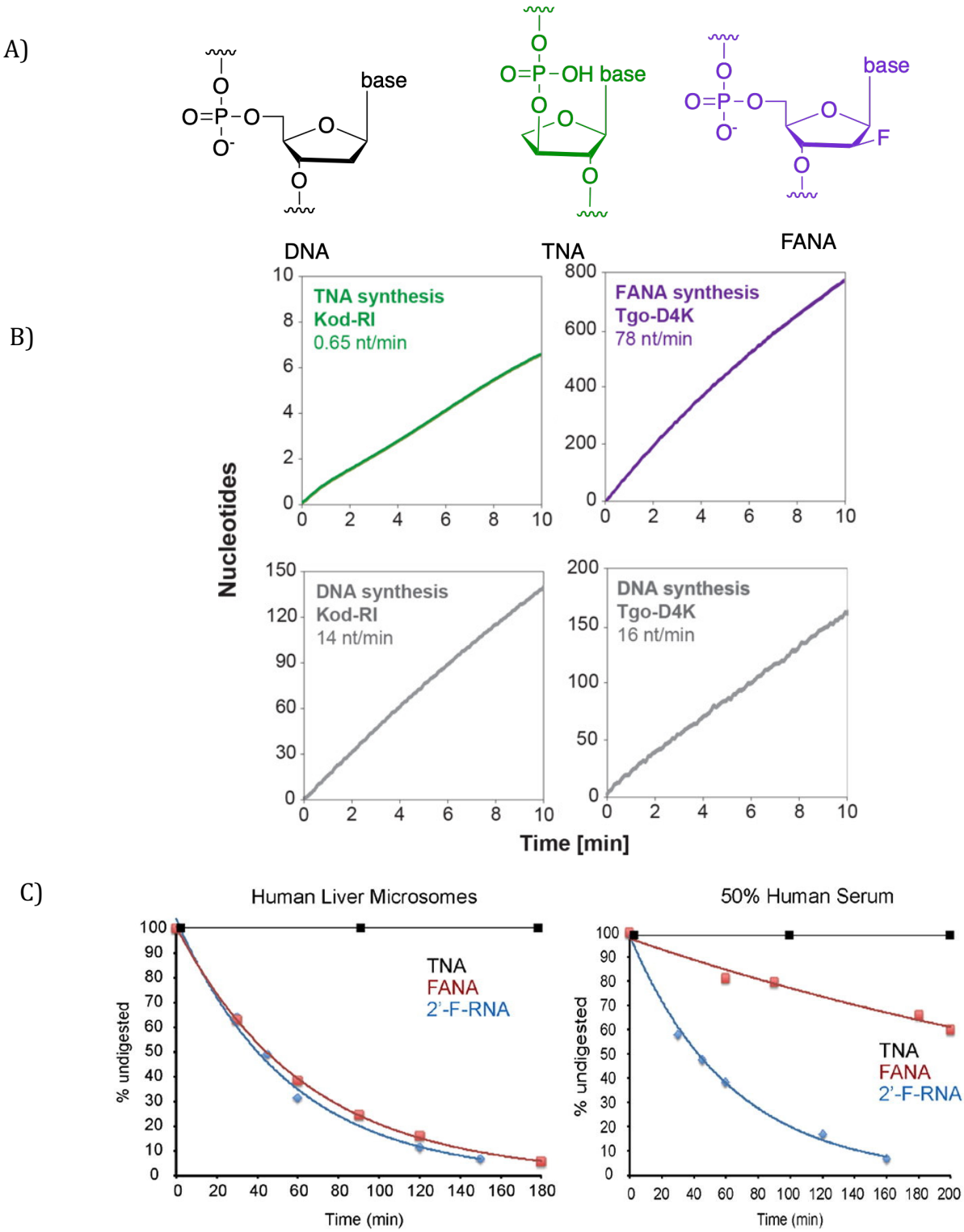


Figure 1.3. A) DNA, FANA, and TNA chemical structures. B) Polymerization of TNA, FANA, and DNA by first generation evolved polymerases, Kod-RI (left) and Tgo-D4K (right). Figure adapted with permission from A*nal. Chem.* 2017, 89, 23, 12622-12625. Copyright 2017

American Chemical Society. C) Nuclease degradation of FANA and TNA. Figure adapted with permission from *Bioorg. Med. Chem. Lett.* 2016, 26, 10, 2418-2421. Copyright 2016 Elsevier.

Due to the challenges of XNA synthesis, XNAs are typically made by DNA-templated primer extension. This has been made possible by continued efforts to improve synthesis of XNA monomers and to engineer polymerases with activity on XNA congeners (Liao et al., 2019; Nikoomanzar et al., 2019; Pinheiro et al., 2012). Despite significant progress, these enzymes continue to suffer from poor fidelity, as compared to natural polymerases on cognate substrates, and they lack the exonuclease activity needed to correct misincorporation. This is magnified by the need for both a polymerase and a reverse transcriptase, as XNAs are typically synthesized on DNA templates and reverse transcribed back into DNA. However, while tedious, the interconversion of information between XNA and DNA allows for the exploitation of powerful tools, such as PCR and NGS, to be applied to XNA systems. As such, several xenobiotic nucleic acid architectures have already been utilized to select aptamers to different targets ranging from small molecules to proteins. These include but are not limited to: a TNA aptamer for ATP (Zhang & Chaput, 2020); FANA and TNA aptamers to HIV RT, a key drug target for HIV and an aptagenic nucleic acid binding protein (Dunn et al., 2020; Ferreira-Bravo et al., 2015; Mei et al., 2018); an 1,5-anhydrohexitol nucleic acid (HNA) aptamer for vascular endothelial growth factor (VEGF), the same protein targeted by Macugen for the treatment of AMD (Eremeeva et al., 2019). These are among the most commonly selected targets for natural nucleic acid aptamers and are well studied.

1.3.2 Threose Nucleic Acid, Benefits and Pitfalls

Among those XNAs capable of Darwinian evolution, TNA stands out due to its recalcitrance to nuclease degradation. This coupled with TNA's ability to efficiently base pair with DNA and itself and the availability of engineered TNA polymerases in recent years makes TNA well suited for therapeutic use. Originally considered as a possible progenitor to RNA in a prebiotic world, TNA consists of a four-carbon sugar oriented in a similar geometry with phosphodiester bonds connecting the 2' and 3' carbons of adjacent nucleosides. The simpler four carbon sugar would be more easily formed in a prebiotic world, however the orientation of the phosphodiester linkage and the lack of a 5' carbon adds steric constraints relative to the ribose sugar backbones. This makes it difficult to chemically synthesize TNA monomers, as the 3' hydroxyl competes for space with the nucleobase and has similar nucleophilicity to the 2' hydroxyl, making regioselectivity challenging (Sau et al., 2016; Schoning et al., 2000).

As TNA is one of the most structurally diverse XNAs, these differences also contribute to the challenge of synthesizing TNA enzymatically. Unlike DNA:DNA homoduplexes, TNA:DNA heteroduplexes are more A-form like than B-form, due to the more rigid, shorter backbone of the TNA strand (Anosova et al., 2016). Hybridization alone is not sufficient for information transfer; polymerases themselves can also have unfavorable contacts with XNA substrates. Crystal structures obtained of the first generation TNA polymerase, Kod-RI, reveal that the incoming TNA nucleobase has a high degree of flexibility in the active site, weakening the Watson-Crick base pairing interaction with the template such that the incoming base and templating base are not in a planar alignment (Chim et al., 2017). This distortion likely accounts for the reduced activity and fidelity of engineered TNA

polymerases, and may be improved in the second and third generation TNA polymerases, Kod-RS and Kod-RSGA (Nikoomanzar et al., 2019, 2020). This is important because the fidelity of the system for the faithful incorporation of substrates in the generation of full length product is a major consideration for Darwinian evolution. An abundance of mutations during aptamer selection would significantly decrease the enrichment of a given sequence. Enrichment scores provide greater context to the functionality of the resultant aptamers in terms of their binding affinity.

Finally, while great strides have been made recently in terms of monomer synthesis and engineered polymerases, solid phase synthesis of TNA remains a significant challenge. A ligase mediated strategy is proposed herein to extend the lengths achievable by solid phase synthesis with ligase activity on TNA strands.

1.3.3 In Vitro Selection of TNA Aptamers

The in vitro selection of functional nucleic acids uses solid phase synthesis to generate large DNA libraries of over 10^{15} sequences, which typically consist of a central random region flanked on both sides by conserved primer regions. Putative aptamers are challenged to bind a target, usually a small molecule or protein. Next, binders are separated from the rest of the pool. A common way to accomplish this is through the use of a target with a label, such as biotin, 6X histidine tags, or FLAG tags. By passing free and complexed target over an affinity resin, binding sequences are retained on the column, while weak and nonbinders are washed through. Remaining sequences are then eluted off the column and amplified by PCR.

In the case of functional nucleic acids, the genotype and phenotype are one and the same. Over the course of 7-20 iterative cycles of binding, eluting, and amplification, the

pool becomes enriched in functional sequences which are identified through NGS. As the pool converges, the abundance of any given sequence typically correlates to its functionality. While aptamer selection is conceptually simple, it can be technically challenging. In order to maximize selection stringency while minimizing the necessary number of rounds performed, a foundational approach to TNA aptamer selection termed DNA display has been used. This system, analogous to mRNA display strategies, covalently links the phenotype-the TNA aptamer-to its encoding DNA genotype. The first TNA aptamer reported was selected against the protein thrombin in this manner, using Therminator DNA polymerase and diaminopurine in place of adenine for improved stability (Yu et al., 2012). With the evolution of new TNA polymerases and discovery of a reverse transcriptase with activity on TNA (Jackson et al., 2019), the selection strategy has evolved to transcribing DNA templates into free TNA, which may be partitioned and reverse transcribed into DNA for PCR amplification. However, reverse transcriptases are known to function with relatively poor activity in the presence of low substrate concentrations, which make it particularly challenging to recover sequences after stringent selections.

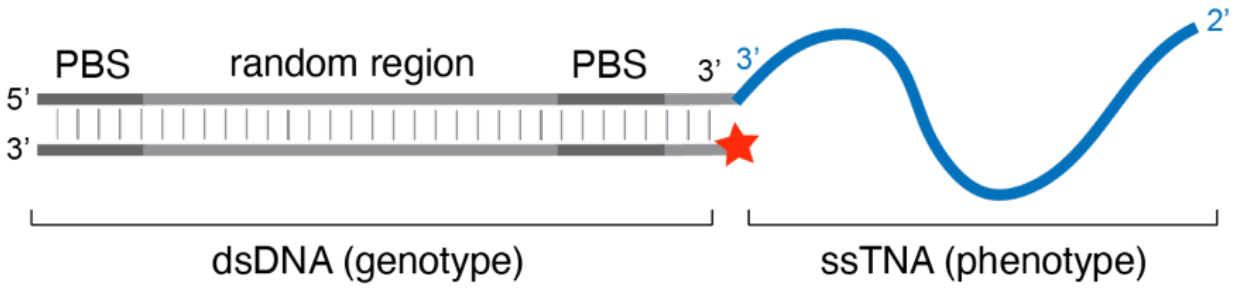


Figure 1.4. DNA display molecule design, with dsDNA (grey) flanked by primer binding sites (PBS), attached to ssTNA (blue). A dye molecule (red star) allows the library to be visualized during selection.

Herein, we return to the original DNA display strategy and build on previous methods of TNA aptamer selection but remove the need for a reverse transcriptase. This method only requires a single engineered polymerase to transcribe DNA genetic information into functional TNA, and the recovered molecules are easily PCR amplified.

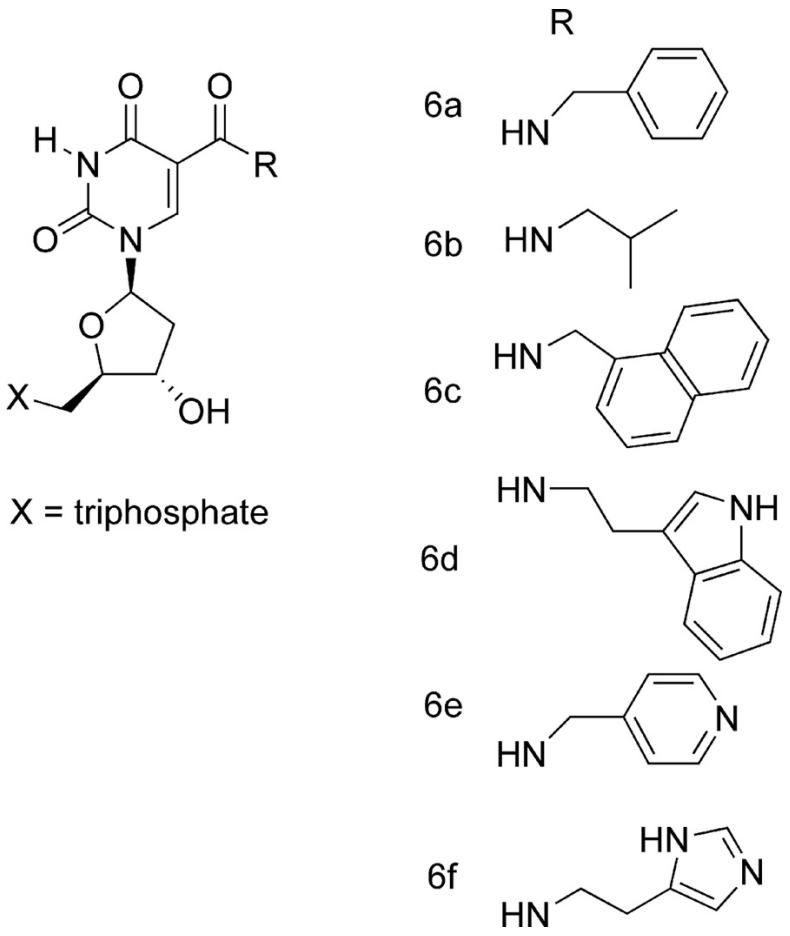


Figure 1.5. Selected SOMAmer dU modifications at C5 with amide linkage. Figure reprinted with permission from *J. Am. Chem. Soc.* 2010, 132, 12, 4141-4151. Copyright 2010 American Chemical Society.

This DNA display strategy was further extended to explore new areas of chemical space by including modified TNA nucleobases. A major thrust in the aptamer field aims to expand

the chemical functionality of aptamers by adding new chemical groups on the nucleobase moiety, typically in the 5 position of pyrimidines or the 7 position of purines, chemical modifications into the major groove, which are less likely to interfere with polymerase activity. This strategy is best exemplified by slow off-rate modified aptamers (SOMAmers). These DNA aptamers contain modified dU residues, where the 5 position of the nucleobase is modified by a chemically diverse moiety, often hydrophobic or amino acid-like in nature, as shown in Figure 1.4 (Vaught et al., 2010). The modification is linked to the nucleobase by an amide bond, which provides some planarity that can help keep the group oriented when interacting with the target. An aptamer identified by Vaught et al. and selected with an indole modification (Figure 1.4, 6d) showed a total loss of affinity when substituted with TTP or other dU analogs, 6a and 6c, indicating that these modifications are specific and necessary for binding.

While functionally enhanced DNAs have previously been demonstrated to improve binding kinetics, extension of these concepts to XNAs has been limited by major synthetic challenges. While 5-iodo-dUTP, a common starting material, is commercially available, the equivalent TNA triphosphate takes more than ten synthetic steps to prepare. An additional hurdle for XNA aptamer development is the paucity of polymerases capable of handling XNA substrates, let alone base-modified XNAs. Interestingly, KOD XL, a variant of Kod polymerase, was shown by Vaught et al. to be one of the most active polymerases on base modified DNA, which bodes well for modified TNA synthesis. The incorporation of deep NGS paired with machine learning algorithms, allows for the generation of functionally enhanced TNA libraries that produce a higher proportion of high affinity binders than standard nucleobases alone. This observation holds across three libraries selected against

two targets. It is likely that antigenic sites targeted by antibodies display a higher degree of hydrophobicity, which is more easily targeted by antibodies than by unmodified aptamers. Aptamers with side chains that emulate amino acids by presenting planar, hydrophobic rings can mimic antibodies in this regard. The results presented herein are consistent with this hypothesis, in which libraries containing phenylalanine and tryptophan derivatives notably produced aptamers with lower K_Ds under the same selection stringency. What is less clear is which moieties, or which combination thereof, are best suited for a particular target-a question which necessitates future work to explore other chemical modifications.

1.4 Conclusion

In summary, while obstacles have prevented the widespread use of natural nucleic acid aptamers in medicine, XNAs are ideally poised to join the ranks of antibodies as therapeutic affinity reagents. TNA in particular is inherently biologically stable, and continuing advances in monomer synthesis and polymerase engineering have allowed studies of TNA molecules at greater depth and breadth. This thesis describes several methods to improve the selection, characterization, and synthesis of these exceptional molecules.

1.5 References

Abeydeera, N. D., Egli, M., Cox, N., Mercier, K., Conde, J. N., Pallan, P. S., Mizurini, D. M., Sierant, M., Hibti, F. E., Hassell, T., Wang, T., Liu, F. W., Liu, H. M., Martinez, C., Sood, A. K., Lybrand, T. P., Frydman, C., Monteiro, R. Q., Gomer, R. H., … Yang, X. (2016). Evoking picomolar binding in RNA by a single phosphorodithioate linkage. *Nucleic Acids Research*, *44*(17), 8052–8064. https://doi.org/10.1093/nar/gkw725

Anosova, I., Kowal, E. A., Sisco, N. J., Sau, S., Liao, J., Bala, S., Rozners, E., Egli, M., Chaput, J. C., & Van Horn, W. D. (2016). Structural Insights into Conformation Differences between DNA/TNA and RNA/TNA Chimeric Duplexes. *ChemBioChem*, *17*(18), 1705–1708. https://doi.org/10.1002/cbic.201600349

Baker, M. (2015). Blame it on the antibodies. *Nature*, *521*(7552), 274–276. https://doi.org/10.1038/521274a

Chim, N., Shi, C., Sau, S. P., Nikoomanzar, A., & Chaput, J. C. (2017). Structural basis for TNA synthesis by an engineered TNA polymerase. *Nature Communications*, *8*(1), 1–11. https://doi.org/10.1038/s41467-017-02014-0

Culbertson, M. C., Temburnikar, K. W., Sau, S. P., Liao, J. Y., Bala, S., & Chaput, J. C. (2016). Evaluating TNA stability under simulated physiological conditions. *Bioorganic and Medicinal Chemistry Letters*, *26*(10). https://doi.org/10.1016/j.bmcl.2016.03.118

Dunn, Matthew R., Jimenez, R. M., & Chaput, J. C. (2017). Analysis of aptamer discovery and technology. *Nature Reviews Chemistry*, *1*(10). https://doi.org/10.1038/s41570-017-0076

Dunn, M.R., Mccloskey, C. M., Buckley, P., Rhea, K., & Chaput, J. C. (2020). Generating Biologically Stable TNA Aptamers that Function with High Affinity and Thermal Stability. *Journal of the American Chemical Society*, *142*(17). https://doi.org/10.1021/jacs.0c00641

Ellington, A. D., & Szostak, J. W. (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature*, *346*(6287). https://doi.org/10.1038/346818a0

Eremeeva, E., Fikatas, A., Margamuljana, L., Abramov, M., Schols, D., Groaz, E., & Herdewijn, P. (2019). Highly stable hexitol based XNA aptamers targeting the vascular endothelial growth factor. *Nucleic Acids Research*, *47*(10), 4927–4939. https://doi.org/10.1093/nar/gkz252

Ferreira-Bravo, I. A., Cozens, C., Holliger, P., & DeStefano, J. J. (2015). Selection of 2′-deoxy-2′-fluoroarabinonucleotide (FANA) aptamers that bind HIV-1 reverse transcriptase with picomolar affinity. *Nucleic Acids Research*, *43*(20), 9587–9599. https://doi.org/10.1093/nar/gkv1057

Jackson, L. N., Chim, N., Shi, C., & Chaput, J. C. (2019). Crystal structures of a natural DNA polymerase that functions as an XNA reverse transcriptase. *Nucleic Acids Research*, *47*(13), 6973–6983. https://doi.org/10.1093/nar/gkz513

Jiang, W., Zhang, B., Fan, C., Wang, M., Wang, J., Deng, Q., Liu, X., Chen, J., Zheng, J., Liu, L., & Zhu, T. F. (2017). Mirror-image polymerase chain reaction. *Cell Discovery*, *3*, 17037. https://doi.org/10.1038/celldisc.2017.37

Klussmann, S., Nolte, A., Bald, R., Erdmann, V. A., & Fürste, J. P. (1996). Mirror-image RNA that binds D-Adenosine. *Nature Biotechnology*, *14*(9), 1112–1115. https://doi.org/10.1038/nbt0996-1112

Kratschmer, C., & Levy, M. (2017). Effect of Chemical Modifications on Aptamer Stability in Serum. *Nucleic Acid Therapeutics*, *27*(6), 335–344. https://doi.org/10.1089/nat.2017.0680

Liao, J. Y., Bala, S., Ngor, A. K., Yik, E. J., & Chaput, J. C. (2019). P(V) Reagents for the Scalable Synthesis of Natural and Modified Nucleoside Triphosphates. *Journal of the American Chemical Society*, *141*(34), 13286–13289. https://doi.org/10.1021/jacs.9b04728

Long, S. B., Long, M. B., White, R. R., & Sullenger, B. A. (2008). Crystal structure of an RNA aptamer bound to thrombin. *RNA*, *14*(12), 2504–2512. https://doi.org/10.1261/rna.1239308

Mei, H., Liao, J.-Y., Jimenez, R. M., Wang, Y., Bala, S., McCloskey, C., Switzer, C., & Chaput, J. C. (2018). Synthesis and Evolution of a Threose Nucleic Acid Aptamer Bearing 7-Deaza-7-Substituted Guanosine Residues. *Journal of the American Chemical Society*, *140*(17). https://doi.org/10.1021/jacs.7b13031

Ng, E. W. M., Shima, D. T., Calias, P., Cunningham, E. T., Guyer, D. R., & Adamis, A. P. (2006). Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nature Reviews Drug Discovery*, *5*(2). https://doi.org/10.1038/nrd1955

Ni, S., Yao, H., Wang, L., Lu, J., Jiang, F., Lu, A., & Zhang, G. (2017). Chemical modifications of nucleic acid aptamers for therapeutic purposes. In *International Journal of Molecular Sciences* (Vol. 18, Issue 8). MDPI AG. https://doi.org/10.3390/ijms18081683

Nikoomanzar, A., Dunn, M. R., & Chaput, J. C. (2017). Evaluating the Rate and Substrate Specificity of Laboratory Evolved XNA Polymerases. *Analytical Chemistry*, *89*(23), 12622–12625. https://doi.org/10.1021/acs.analchem.7b03807

Nikoomanzar, A., Vallejo, D., & Chaput, J. C. (2019). Elucidating the Determinants of Polymerase Specificity by Microfluidic-Based Deep Mutational Scanning. *ACS Synthetic Biology*, *8*(6), 1421–1429. https://doi.org/10.1021/acssynbio.9b00104

Nikoomanzar, A., Vallejo, D., Yik, E. J., & Chaput, J. C. (2020). Programmed Allelic Mutagenesis of a DNA Polymerase with Single Amino Acid Resolution. *ACS Synthetic Biology*, *9*(7), 1873–1881. https://doi.org/10.1021/acssynbio.0c00236

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. (2012). Synthetic genetic polymers capable of heredity and evolution. *Science*, *336*(6079), 341–344. https://doi.org/10.1126/science.1217622

Sau, S. P., Fahmi, N. E., Liao, J. Y., Bala, S., & Chaput, J. C. (2016). A Scalable Synthesis of α-L-Threose Nucleic Acid Monomers. *Journal of Organic Chemistry*, *81*(6), 2302–2307. https://doi.org/10.1021/acs.joc.5b02768

Schoning, K. U., Scholz, P., Guntha, S., Wu, X., Krishnamurthy, R., & Eschenmoser, A. (2000). Chemical etiology of nucleic acid structure: The α-threofuranosyl-(3'→2') oligonucleotide system. *Science*, *290*(5495), 1347–1351. https://doi.org/10.1126/science.290.5495.1347

Shigdar, S., Macdonald, J., O'Connor, M., Wang, T., Xiang, D., Al.Shamaileh, H., Qiao, L., Wei, M., Zhou, S.-F., Zhu, Y., Kong, L., Bhattacharya, S., Li, C., & Duan, W. (2013). Aptamers as Theranostic Agents: Modifications, Serum Stability and Functionalisation. *Sensors*, *13*(10), 13624–13637. https://doi.org/10.3390/s131013624

Tuerk, C., & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, *249*(4968). https://doi.org/10.1126/science.2200121

Vater, A., & Klussmann, S. (2015). Turning mirror-image oligonucleotides into drugs: The evolution of Spiegelmer® therapeutics. In *Drug Discovery Today* (Vol. 20, Issue 1, pp. 147–155). Elsevier Ltd. https://doi.org/10.1016/j.drudis.2014.09.004

Vaught, J. D., Bock, C., Carter, J., Fitzwater, T., Otis, M., Schneider, D., Rolando, J., Waugh, S., Wilcox, S. K., & Eaton, B. E. (2010). Expanding the chemistry of DNA for in vitro selection. *Journal of the American Chemical Society*, *132*(12), 4141–4151. https://doi.org/10.1021/ja908035g

Wang, Y., Ngor, A. K., Nikoomanzar, A., & Chaput, J. C. (2018). Evolution of a General RNA-Cleaving FANA Enzyme. *Nature Communications*, *9*(1), 1–10. https://doi.org/10.1038/s41467-018-07611-1

Yu, H., Zhang, S., & Chaput, J. C. (2012). Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. *Nature Chemistry*, *4*(3), 183–187. https://doi.org/10.1038/nchem.1241

Zhang, L., & Chaput, J. C. (2020). In vitro selection of an ATP-binding TNA aptamer. *Molecules*, *25*(18). https://doi.org/10.3390/molecules25184194

CHAPTER 2

Publication Note

This paper was originally published in the Journal of the American Chemical Society. Adapted with permission from: Dunn, M.R.; McCloskey, C.M.; Buckley, P.; Rhea, K.; Chaput, J.C. Generating Biologically Stable TNA Aptamers that Function with High Affinity and Thermal Stability. J. Am. Chem. Soc. 2020, 142, 17, 7721-7724. Copyright 2020 by American Chemical Society.

2.1 Contribution Statement

M.D., C.M., and J.C. conceived of the project and designed the experiments. M.D., C.M., P.B., and K.R. performed the experiments. C.M. and J.C. wrote the manuscript. All authors reviewed and commented on the manuscript.

2.2 Abstract of Chapter

Aptamers are often prone to nuclease digestion, which limits their utility in many biomedical applications. Here we describe a xeno-nucleic acid system based on α-L-threofuranosyl nucleic acid (TNA) that is completely refractory to nuclease digestion. The use of an engineered TNA polymerase permitted the isolation of functional TNA aptamers that bind to HIV reverse transcriptase (HIV RT) with K_Ds of~0.4–4.0 nM. The aptamers were identified using a display strategy that provides a powerful genotype–phenotype linkage. The TNA aptamers remain active in the presence of nuclease and exhibit markedly higher thermal stability than monoclonal antibodies. The combined properties of biological stability, high binding affinity, and thermal stability make TNA aptamers a powerful system for the development of diagnostic and therapeutic agents.

2.3 Introduction

Aptamers, nucleic acid molecules that mimic antibodies by folding into shapes with ligand binding affinity (Ellington & Szostak, 1990), have enormous potential as diagnostic and therapeutic agents (Zhou & Rossi, 2017). Aptamers are generated by in vitro selection from libraries of combinatorial sequences (Wilson & Szostak, 1999). Similar to natural selection, in vitro selection is a process of selective amplification in which a population of nucleic acid molecules is challenged to bind a target or catalyze a chemical reaction. Molecules having a desired fitness are recovered and amplified to generate a new population that has become enriched in members with a particular activity. Although hundreds of aptamers have been reported in the literature (Dunn et al., 2017), most are unsuitable for in vivo applications because they are susceptible to nuclease digestion (Keefe et al., 2010). One exception is Spiegelmers, mirror-image aptamers, but such reagents are restricted to achiral targets or targets that can be generated by chemical synthesis (Vater & Klussmann, 2015). Modifications made to the 2′ sugar position offer some relief against nucleases but are generally still prone to digestion (Dunn et al., 2017; Keefe et al., 2010).

2.4 Results and Discussion

We describe an approach for evolving biologically stable aptamers based on the framework of α-L-threofuranosyl nucleic acid (TNA, Figure 2.1a), a xeno-nucleic acid (XNA) system in which the ribose sugar in RNA has been replaced with a threose sugar (Schoning et al., 2000). This approach is sufficiently general that it could be applied to any XNA for which a polymerase is available to copy DNA templates into XNA (Pinheiro et al., 2012; Yu et al., 2012). The strategy is analogous to protein display technologies, such as mRNA

display, that provide a covalent link between the encoding mRNA and translated protein (Roberts & Szostak, 1997). Here, freshly synthesized TNA is physically linked to its complementary DNA template, which is present in double stranded (ds) form (Figure 2.1b). In this configuration, TNA molecules are amplified by PCR using dsDNA as the template. This approach simplifies the selection by avoiding the need for a reverse transcriptase.

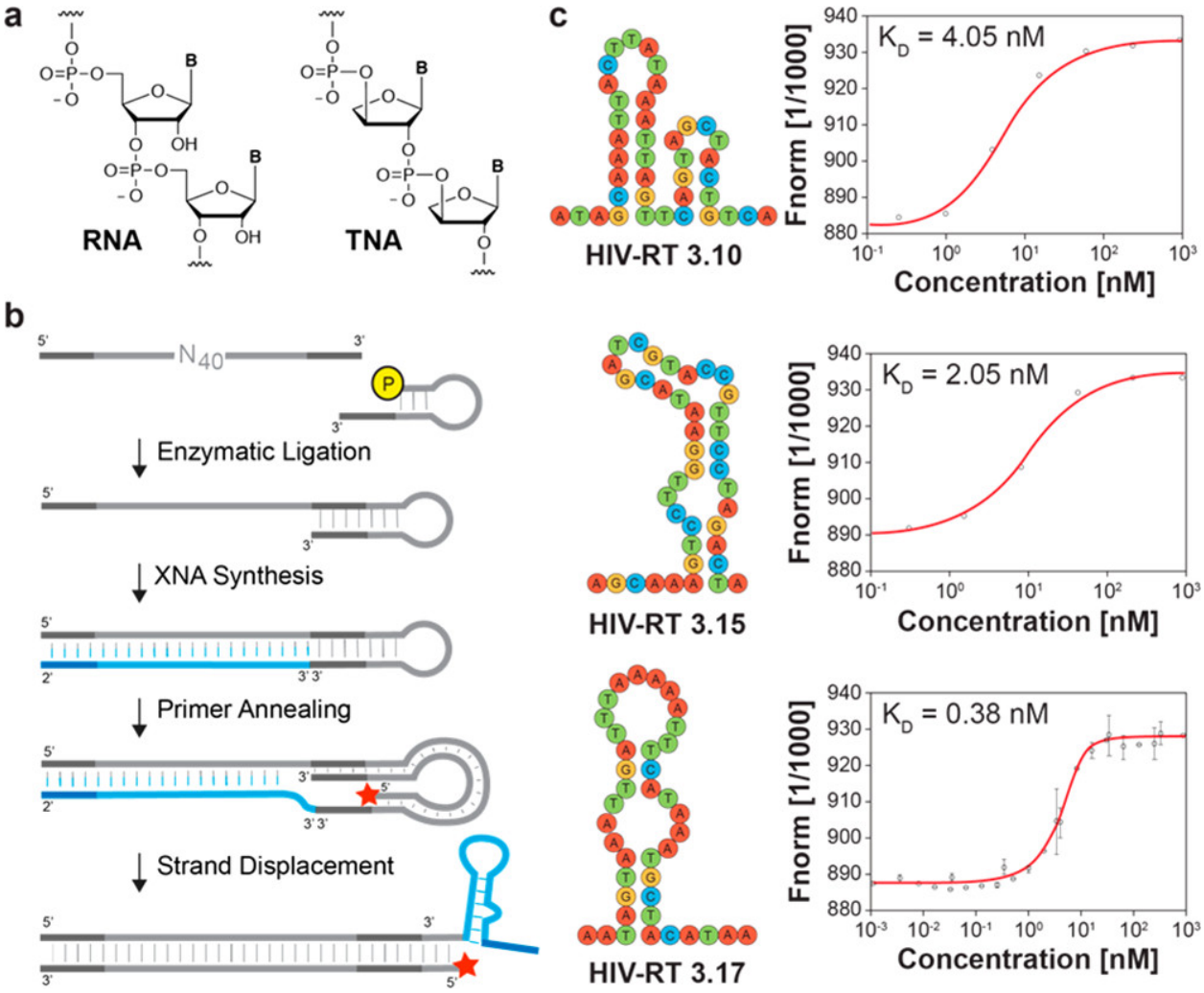


Figure 2.1. Selection of high-affinity TNA aptamers. (a) Chemical structures of RNA and TNA. (b) In vitro selection strategy. (c) MST binding affinity curves for the highest affinity aptamers. HIV-RT 3.17 was evaluated in triplicate. Error bars, standard deviation of each data point.

The selection (Figure 2.1b) was performed by extending a self-priming DNA library with chemically synthesized TNA triphosphates (tNTPs)(Liao et al., 2019) using an engineered TNA polymerase that was developed to synthesize TNA on DNA templates (Nikoomanzar et al., 2019). The product of the primer-extension step is a chimeric TNA-DNA hairpin where a 40 nt random region and downstream primer binding site are copied into TNA. The TNA portion of the duplex was displaced in a separate step by extending a DNA primer annealed to the loop region of the hairpin with DNA, which results in a combinatorial library of single-stranded TNA molecules that are each physically linked to their encoding dsDNA templates. To enrich for TNA molecules with affinity to a desired target, the TNA library was incubated with a protein target, and bound sequences were recovered and amplified by PCR (SI Table 2.1). A second PCR step was performed with a PEG-modified DNA primer, and the single-stranded, PEGylated DNA template was obtained after purification by denaturing polyacrylamide gel electrophoresis (PAGE). The selection was continued by ligating the template to the DNA stem-loop.

We applied this approach to isolate TNA aptamers with affinity to a reverse transcriptase (RT) found in the human immunodeficiency virus (HIV) (Bala et al., 2018). Although previous selections have generated DNA(Lai & DeStefano, 2012) and FANA (2′-fluoroarabino nucleic acid) (Ferreira-Bravo et al., 2015) aptamers to HIV RT, both classes of aptamers are susceptible to nuclease digestion (Culbertson et al., 2016; Noronha et al., 2000) . We therefore sought to establish a biologically stable aptamer that could function with high binding affinity without the need for extensive chemical modifications. A key question that we wished to address was whether TNA could adopt structures with the same level of high affinity commonly observed for the best monoclonal antibodies (mAbs).

Although we previously isolated a TNA aptamer that binds to HIV RT with a K_D of ~5 nM, that example utilized a reverse transcription step and required NGS analysis (Mei et al., 2018). Recognizing that our TNA reverse transcriptase functions with weak activity, we reasoned that the new approach should make it possible to recover higher affinity binders that may have been lost due to the limitations of our TNA reverse transcriptase.

We performed three rounds of selection starting from 10^{13} unique TNA molecules. For each round, the library was incubated in buffer (20 mM Tris-HCl [pH 7.4], 140 mM NaCl, 5 mM KCl, 1 mM $MgCl_2$, and 1 mM $CaCl_2$) with HIV RT immobilized on the surface of an ELISA plate. After a 1 h incubation (25 °C), the well was drained, unbound sequences were removed, and nonspecific binders were disrupted with a high ionic strength buffer. High-affinity aptamers that remained bound to the target were then recovered by denaturing the complex at 70 °C with buffer containing 3.5 M urea. After three rounds of selection, a portion of the pool was cloned and 20 library members were submitted for sequencing. From this set, we obtained 15 high-quality reads with no significant sequence similarity (SI Table 2.2).

We synthesized nine randomly selected clones by primer-extension and determined their K_Ds by microscale thermophoresis (MST). The screen resulted in K_D values ranging from 400 pM to 70 nM with six of the nine sequences having K_D values of <10 nM (SI Table 2.3). The three highest affinity TNA aptamers (HIV-RT 3.10, HIV-RT 3.15, and HIV-RT 3.17) bound to HIV RT with K_Ds of 4 nM, 2 nM, and 380 pM, respectively (Figure 2.1c). The tightest affinity aptamer, HIV-RT 3.17 (K_D 380 ± 115 pM), was evaluated in triplicate. This value compares favorably with the starting library, which shows no significant binding to HIV RT (K_D > 2 μM, SI Figure 2.1).

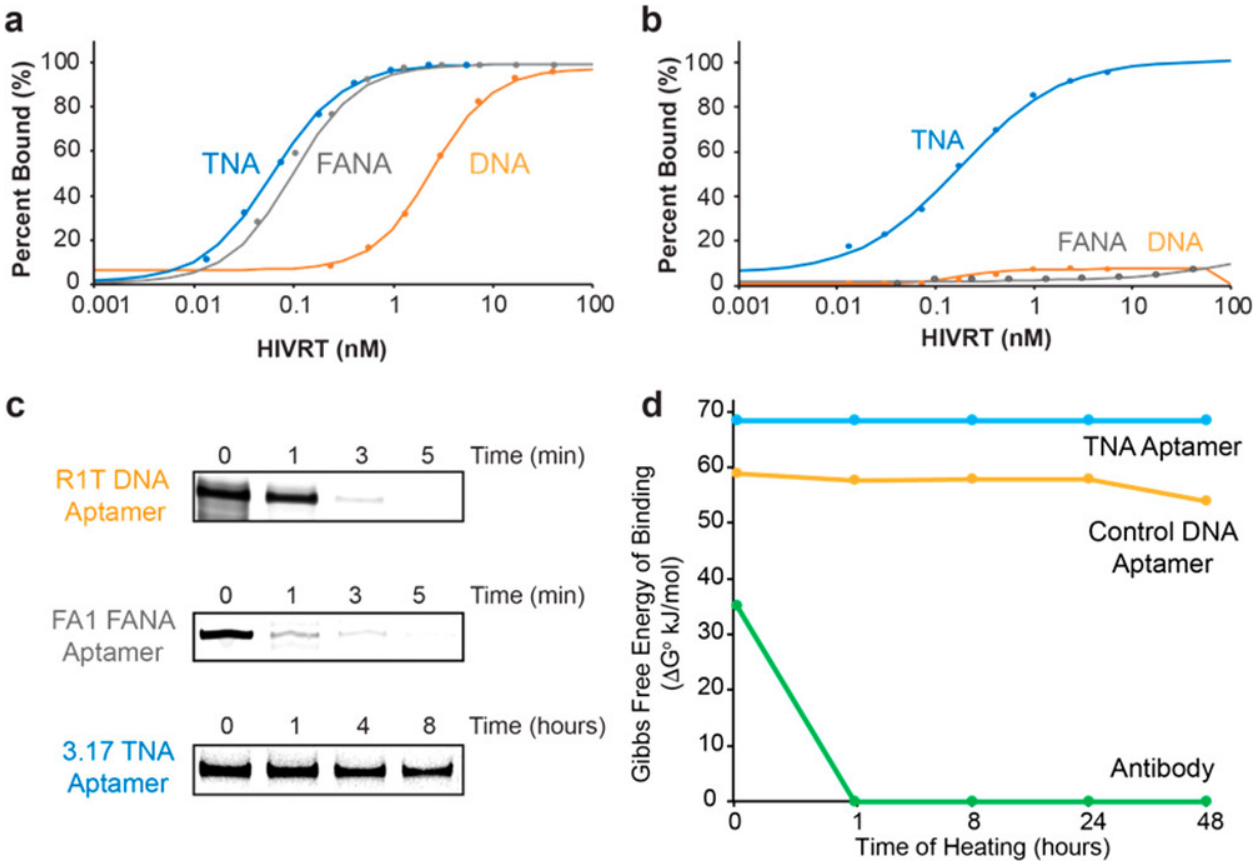


Figure 2.2. Characterization of HIV-RT 3.17. (a) Representative EMSA data in buffer. (b) Representative EMSA data with SVPE. (c) Biostability validation using SVPE. (d) Thermal stability validation after heating at 75 °C for 0–48 h and cooling to 24 °C for analysis.

To validate the affinity of HIV-RT 3.17, a second analytical technique was used to measure the K_D. In this case, we used an electrophoretic mobility shift assay (EMSA) to evaluate the binding interaction by incubating the protein with a low concentration of labeled aptamer and varying the concentration of the protein target (Figure 2.2a, SI Figures 2.2, 2.3). As positive controls, we measured the K_D of R1T and FA1, known DNA (Lai & DeStefano, 2012) and FANA (Ferreira-Bravo et al., 2015) aptamers, respectively, that were previously selected to bind HIV RT. The resulting isotherms reveal that HIV-RT 3.17 binds to HIV RT with a K_D of 196 ± 86 pM, R1T binds with a K_D of 4.5 ± 0.41 nM, and FA1 binds

with a K_D of 324 ± 50 pM. These values closely agree with our MST-derived value of 380 pM for HIV-RT 3.17 and the literature K_D for R1T (1.1 ± 0.2 nM). However, the FA1 aptamer differed considerably from its reported value (4 ± 3 pM), suggesting that the fluorescent dye used in place of the radiolabel may have interfered with binding.

Next, we evaluated the binding of HIV-RT 3.17 under conditions where DNA aptamers and mAbs lose their activity. Since most aptamers are susceptible to nuclease digestion, we performed the HIV RT binding assay in the presence of snake venom phosphodiesterase (SVPE), a highly active 3′ exonuclease. The HIV-RT 3.17, R1T, and FA1 aptamers were incubated with SVPE for 1 h at 37 °C prior to performing the equilibrium binding assay by EMSA. Analysis of the resulting binding curves revealed that HIV-RT 3.17 remains active in the presence of SVPE (K_D of ~500 pM), while DNA and FANA digestion abrogates all protein-binding activity for both the R1T and FA1 aptamers (Figure 2.2b, SI Figure 2.2). Time-dependent analysis of digestion by denaturing PAGE reveals that the DNA and FANA aptamers degrade in minutes, while the TNA aptamer remains undigested after 8 h at 37 °C (Figure 2.2c). Moreover, the TNA aptamer showed no signs of degradation after 24 h at 37 °C with 50% human serum or concentrated human liver microsomes (SI Figure 2.4). The functional activity of HIV-RT 3.17 was further validated using a pull-down assay (SI Figure 2.5).

Recognizing that antibodies unfold irreversibly, we performed a thermal challenge by heating HIV-RT 3.17, R1T, and a commercial mAb in buffer at 75 °C. At designated time points between 0 and 48 h, the reagents were cooled to 24 °C and assayed for binding to HIV RT by biolayer interferometry (BLI). The K_Ds from this data were used to calculate changes in the Gibbs free energy at standard conditions ($\Delta G°$), which were plotted as a

function of temperature to illustrate the effect of heating on the affinity reagents (Figure 2.2d, SI Table 2.4, SI Figure 2.7). Visualizing these data in terms of binding energy as opposed to affinity affords a more intuitive picture of the effect of heating on the function of each binder. Our results indicate that HIV-RT 3.17 and R1T retain full activity after 48 h of heating at 75 °C, while the monoclonal antibody lost activity within the first hour of heating. After 72 h, the activities of both aptamers are reduced to 50%, presumably due to thermal degradation (SI Figure 2.6).

To gain further insight, we performed a secondary structure prediction calculation using mFold (Zuker, 2003). HIV-RT 3.17 was predicted to adopt a stem-loop structure that contains two bulges in the base-pairing region (Figure 2.1c). It also contains a predicted T:G wobble pair and flanking sequences on both the 5′ and 3′ sides of the primary stem-loop structure. Detailed structure–activity relationship studies confirmed the importance of the stem-loop structure and the wobble base pair for binding to HIV RT. In general, deletions made throughout the primary stem-loop structure reduced the binding activity of HIV-RT 3.17 by 2–10-fold. However, more significant drops in activity (17–100-fold) were observed when changes were made to the five base pairs that define the stem near the 5′ and 3′ termini. Even mutation of the T:G wobble pair to a C:G Watson–Crick base resulted in a 10-fold loss in activity. Deletion of either of the two 5′ flanking nucleotides or the five 3′ flanking nucleotides resulted in a ~3-fold reduction in binding affinity (Figure 2.3).

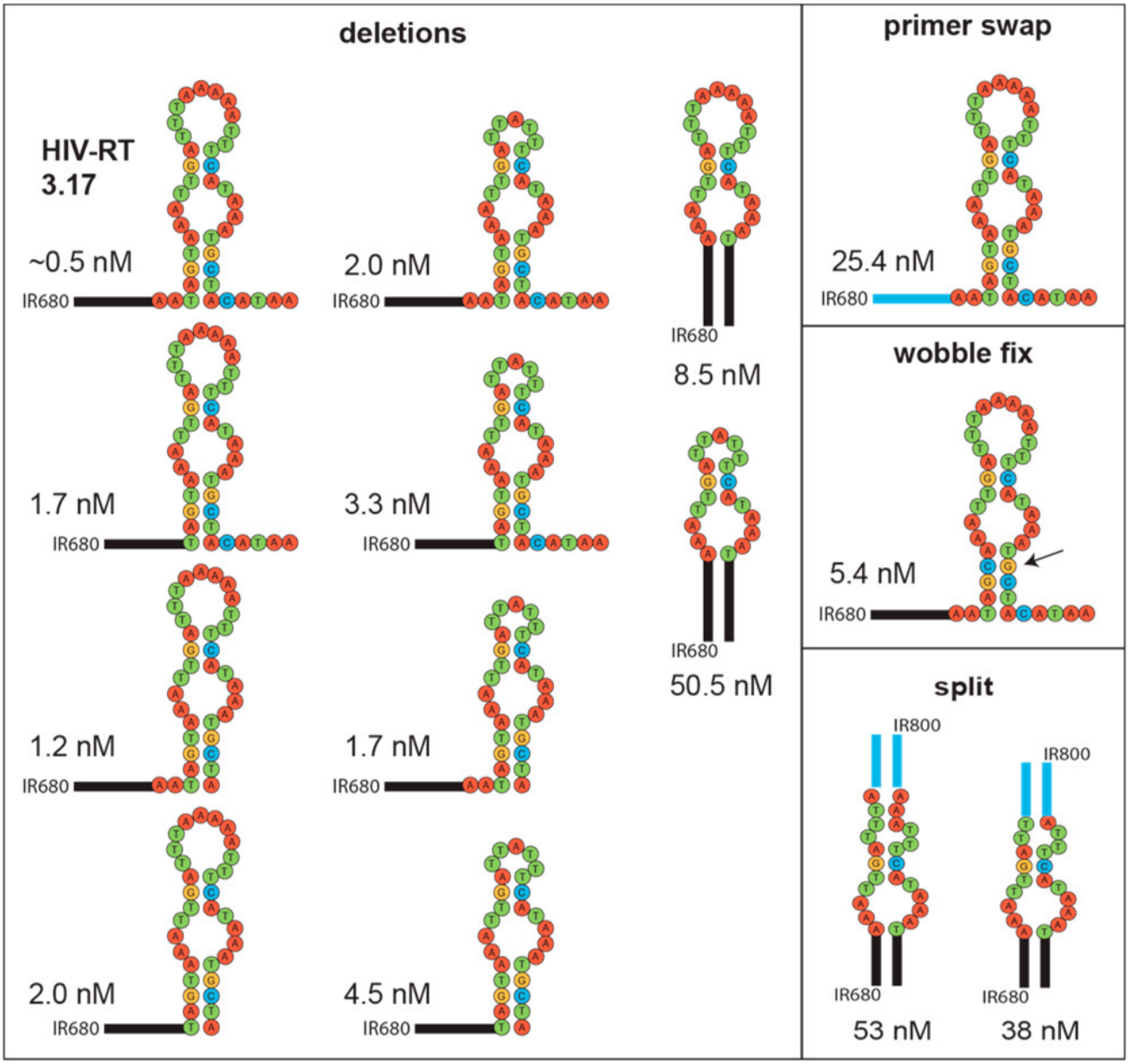


Figure 2.3. Structure– activity relationship of HIV-RT 3.17. MST-derived K_D values were used to evaluate the structure and function of HIV-RT 3.17. Aptamer constructs were made with deletions, mutations, and structural permutations at defined positions.

Confident that HIV-RT 3.17 adopts a stem-loop structure, we designed two split-aptamers that function with K_Ds of 40–50 nM (Figure 2.3). The split-aptamer design is a common approach for biological sensors and has not previously been demonstrated for

XNA polymers (Chen et al., 2016). Interestingly, replacing the five base pair stem in the primary stem-loop region in both the whole and split-aptamers resulted in a similar reduction in binding affinity. This suggests that the precise stem structure selected for HIV-RT 3.17 likely provides more than structural integrity and may manifest in a more complex level of interaction with its protein target. Further engineering of the split-aptamer system may yield a split-aptamer with even greater affinity.

Although antibodies remain the gold standard for affinity reagents, they also represent the greatest source of problems in biomedical research (Baker, 2015). The sensitivity of antibodies to elevated temperatures demonstrates a critical weakness that limits their shelf life, reproducibility, and performance in functional assays. Thermal instability could be linked to problems with batch-to-batch variability that have plagued antibody research (Begley & Ellis, 2012). Given these problems, researchers are calling for better standardization and access to new types of affinity reagents (Bradbury & Plückthun, 2015).

Unlike conventional aptamers, or even certain XNAs, TNA is completely recalcitrant to nuclease digestion and is amenable to in vitro selection against any biological protein of therapeutic or diagnostic interest. The latter provides an important benefit over Spiegelmers, whose chiral targets must be generated by chemical synthesis (Vater & Klussmann, 2015). Relative to antibodies, TNA can achieve similar binding while avoiding unwanted problems associated with thermal denaturation. The ability to fold cooperatively into functional structures, combined with their chemical synthesis, solves the cold-chain problem and should allow for improved performance in functional assays.

In summary, our results establish that TNA aptamers can function with high biological stability, binding affinity, and thermal stability. These data offer a possible solution to the antibody problem and provide strong support for continued development of TNA reagents for biomedical applications.

2.5 Experimental Details

2.5.1 General Information

Taq, Therminator, and Q5 high-fidelity DNA polymerases as well as T4 DNA ligase, 10x T4 DNA ligase buffer, and 10x ThermoPol buffer were purchased from New England Biolabs (Ipswich, MA). Kod-RI TNA polymerase was expressed and purified as described previously1. Noted experiments were performed in DNA LoBind tubes, purchased from Eppendorf. DNA triphosphates were purchased from Sigma-Aldrich (St. Louis, MO). TNA triphosphates bearing natural bases were synthesized as previously described. DNA oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), purified by denaturing polyacrylamide gel electrophoresis (PAGE), electroeluted, buffer exchanged and concentrated using Millipore YM-10 or YM-30 Centricon centrifugal filter units, and quantified by UV absorbance via NanoDrop (Thermo Fisher Scientific, Waltham, MA). Sequences were cloned into TOPO vectors using TOPO-TA Cloning Kit, with pCR2.1 Vector (Thermo Fisher Scientific, US). For the aptamer selection, Nunc Immobilizer Amino plates (Lock Well C8) were used (Thermo Fisher Scientific, US). HIV Reverse Transcriptase was obtained from Worthington Biochemical Corp (Lakewood, NJ). Anti-HIV1 Reverse Transcriptase antibody [39/4.12.2] (ab9066) was purchased from abcam (Cambridge, MA). EvaGreen dye was purchased from Biotium (Fremont, CA).

2.5.2 Library preparation using DNA display method

The library was synthesized such that 35% of bases were A, 35% T, 15% G, 15% C. This design was chosen to improve the TNA transcription efficiency with the engineered polymerase. The initial library containing an internal variable region of 40 nucleotides flanked on both sides by fixed priming sites (PBS7 and PBS8) was synthesized on an ABI3400 DNA synthesizer and purified by 10% denaturing PAGE. The band corresponding to full length library was excised from the gel using a scalpel, electroeluted, concentrated by EtOH precipitation, resuspended in water, and quantified by UV absorbance. The phosphorylated DNA hairpin primer was ligated to the library by combining 10nmolof both hairpin and template in a final volume of 1 mL of 1x T4 DNA ligase buffer. The solution was denatured for 5 min at 95°C and then annealed by incubating for 15 min at room temperature. Once annealed, 20,000 U of T4 DNA ligase was added and the reaction was incubated at 16°C overnight. The following morning, the hairpin library was purified by 10% denaturing PAGE and quantified by UV absorbance. The hairpin library was extended with tNTPs to form a chimeric TNA:DNA hairpin heteroduplex as follows. A 250 µL reaction volume containing 0.5 nmol hairpin library,1x ThermoPol Buffer, and 100 µM each tNTPs was heated for 5 min at 95°C to denature and then annealed by incubating for 15 min at room temperature. TNA polymerization was initiated by adding Kod-RI polymerase (1 µM) and then incubated for 4 h at 55°C. To confirm reaction completion, a small fraction of 80 nt library, 140 nt hairpin library, and 200 nt TNA extension products were run on an 8% native TBE PAGE at 200 V, 4°C for 2 h, stained with ethidium bromide and UV imaged. Full length TNA extended library was purified from incomplete reaction products by 10% denaturing PAGE as described above. Following electroelution, the TNA:DNA hairpin library was concentrated and buffer exchanged into water using a YM-30 Centricon filter

device. Library was quantified by UV absorbance. To form the TNA display product with the single-stranded TNA library (phenotype) physically connected to its encoding double-stranded DNA template (genotype), the TNA strand was displaced by extending a DNA primer annealed to the hairpin loop with dNTPs. Accordingly, the TNA:DNA library (1 μM final) was resuspended in 1x ThermoPol buffer containing 500 μM dNTPs and 2 μM strand displacement primer. After heating for 5 min at 95°C and slowly cooling to room temperature, the displacement reaction was initiated by adding Therminator DNA polymerase to a final concentration of 0.2 U/μL. The reaction was incubated for 30 min at 55°C followed by 90 min at 65°C. Therminator polymerase was extracted with phenol/chloroform. Once treated, the displayed TNA library was purified from residual phenol and reaction components, concentrated, and exchanged in aptamer binding (AB) buffer [20 mM Tris-HCl pH 7.4, 140 mM NaCl, 5 mM KCl, 1 mM MgCl2, and 1 mM CaCl2] using a YM-30 Centricon filter device.

2.5.3 Estimation of non-specific background binding

To estimate the amount of background binding of the library to the Nunc plates, three wells of a plate were passivated by overnight incubation with AB buffer, which was aspirated before library application. The overnight pretreatment of the wells with AB buffer neutralizes unreacted electrophiles displayed on the end-grafted polyglycol surface. Approximately 300 pmol (10^{13} sequences) of starting DNA library in 300 μL of aptamer binding buffer supplemented with 0.005% Tween 20 (ABT) was heated to 50°C and then incubated at room temperature for 15 min to anneal. The library (100 μL each) was then added to the three pretreated wells and incubated for 1 h at 25°C under gentle agitation (500 rpm) in a Thermomixer equipped with a plate adapter. Unbound and weakly bound

members were removed by first washing three times with 300 µL of ABT. A second, more stringent set of three washes with 300 µL stringent wash (SW) buffer (20 mM Tris-HCl pH 7.4, 4 M NaCl, 5 mM KCl, 1 mM MgCl2, 1 mM CaCl2, 0.005% Tween 20) was then applied to remove non-specifically bound members. All remaining adsorbed library members were eluted by three incubations with denaturing buffer (DB) (40 mM Tris-HCl pH 8.0, 3.5 M urea, 10 mM EDTA, 0.005% Tween 20) at 70°C for 10 min. The elutions were combined and desalted using a YM-30 Centricon filter unit. To determine the amount of DNA retained in the wells through non-specific adsorption, the elutions were amplified by RT-qPCR alongside a standard titration series of library. Accordingly, the elutions and titration series were mixed with1x Q5 high-fidelity DNA polymerase master mix and 2x EvaGreen in a final volume of 25 µL. Each reaction was thermocycled as follows in an Applied Biosystems Quantstudio 6 Flex quantitative real-time PCR instrument: 60 sec at 95°C followed by 40 cycles of 15 sec at 95°C, 15 sec at 58°C, and 60 sec at 72°C. Fluorescence was monitored at the end of each extension step and plotted against cycle number. The Cq of each titration series and sample was utilized to determine the concentration of library bound to each well.

2.5.4 Aptamer selection

HIV RT poised at a concentration of 80 nM was immobilized onto the surface of wells of Nunc Amino Immobilizer C8 strips by incubating overnight at 4°C in 100 µL of 100 mM sodium phosphate (pH 7.5). Following target immobilization, wells were washed three times with 300 µL of AB. The wells were then filled once more with 300 µL of AB and left to incubate at room temperature for 1 h under gentle agitation (500 rpm). The incubation buffer was aspirated and three additional washes of 300 µL ABT were performed before

sealing the wells and storing them at 4°C until further use. Approximately 100 pmol (10^{13} sequences) of TNA displayed library was blocked with an equimolar amount of PBS7 DNA primer in 100 µL of ABT. The solution was heated to 50°C and then incubated at room temperature for 15 min to anneal. This step ensures that the fixed primer binding site is double stranded, preventing it from interfering with the proper folding of the library variable region. The folded, primer-blocked library was then added to a functionalized and passivated well and incubated for 1 h at 25°C under gentle agitation (500 rpm) in a Thermomixer equipped with a plate adapter. Unbound and weakly bound members were removed by first washing three times with 300 µL each of ABT. A second more stringent set of three washes with 300 µL of SW buffer was then applied. All remaining library members were eluted by three incubations with DB buffer at 70°C for 10 min, then desalted using a YM-30 Centricon device.

2.5.5 Droplet digital PCR amplification and template regeneration

Desalted aptamer pools were quantified by qPCR in an Applied Biosystems Quantstudio 6 Flex quantitative real-time PCR instrument in order to determine the amount of material that was eluted from each well. The eluted pool was then subjected to droplet digital PCR amplification (ddPCR) (QX-100, Bio-Rad) where the number of wells required was determined by the quantity of library members eluted as to have 50 library members per droplet. Each ddPCR reaction was carried out in a 20 µL volume by partitioning the reaction mix (ddPCR Supermix for Probes without dUTP) containing the entire eluted pool mixed with 500 nM of each DNA amplification primer (PBS7 and PBS8). The mixture was loaded into a well of a DG8 cartridge before addition of fluorinated oil to the corresponding oil wells of the cartridge. Droplets were formed by inserting the cartridge into the droplet

generator instrument (Bio-Rad). The resulting emulsion, containing ~20,000 droplets per well, was carefully transferred into a 96-well PCR plate and subjected to thermocycling as follows: 60 sec at 95°C followed by 25 cycles of 15 sec at 95°C, 15 sec at 58°C, and 60 sec at 72°C. Immediately after amplification, all wells were pooled together, mixed with an equal volume of chloroform, vortexed vigorously to break the droplets, and spun at 13,000xg to separate the amplicon-containing aqueous phase from the organic phase. The recovered aqueous phase was pipetted to a separate tube. The organic phase was washed with an equal amount of water and pooled with the first recovery. Single-stranded DNA templates were regenerated by PCR amplification and gel purification. Each PCR reaction (500 µL total volume) contained 500 pmol of each DNA primer (PBS7 PEG Long and PBS8) and 5 pmol of double stranded DNA template. Reactions were performed with Taq DNA polymerase in a 1x ThermoPol buffer with 400 µM dNTPs. The solution was divided into 50 µL reactions and thermocycled as follows: 60 sec at 95°C followed by 25 cycles of 15 sec at 95°C, 15 sec at 58°C, and 60 sec at 72°C. The reactions were combined and supplemented with 5 mM EDTA and 50% (w/V) urea. The solution was heat denatured for 5 min at 95°C, and the PEG-modified strand was purified by 10% denaturing PAGE. The corresponding band was excised, electroeluted, and concentrated with a YM-10 Centricon centrifugal filter device. The template was recovered in 100 µL of nuclease free water and quantified by UV absorbance. The enriched single stranded DNA library was used to generate a new TNA display library as described above.

2.5.6 Diversity estimation

The sequence diversity of a retained pool of library members was estimated by RT-qPCR. Accordingly, a small aliquot of a retained pool was amplified by qPCR withQ5 high-fidelity

DNA polymerase master mix in a final volume of 25 µL. Each reaction contained a final concentration of 2x EvaGreen. The reactions were cycled as previously described: 60 sec at 95°C followed by 25 cycles of 15 sec at 95°C, 15 sec at 58°C, and 60 sec at 72°C. The dye-bearing amplicons were then melted by heating from 55°C to 95°C in 0.5°C increments in an Applied Biosystems Quantstudio 6 Flex quantitative real-time PCR instrument. The resulting melt curves were analyzed by computing the areas beneath the two normalized transition peaks centered at ~64°C and 76°C, which represent melting of hetero-duplexed and homo-duplexed amplicons, respectively. The two peak areas were used to compute the fraction f_{hetero} of amplicons in the hetero-duplexed state: $f_{hetero}=A_{64°C}/(A_{64°C}+A_{76°C})$. Melt analyses of serial reductions in the sequence diversity of the starting library were used to create a standard curve relating f_{hetero} to sequence diversity. The value of f_{hetero} for the retained pool thereby permitted estimation of sequence diversity.

2.5.7 Sequencing

Library members that remained in the pool after three rounds of selection were cloned using the TOPO-TA cloning kit and transformed into XL1-blue chemically competent *E. coli*. Cells were plated onto LB-amp plates and grown overnight at 37°C. The following morning, isolates were grown at 37°C in 5mL liquid cultures of LB-amp overnight with shaking at 225 rpm. Plasmids were recovered by miniprepping and sequenced. Sequencing traces were analyzed using CLC Main Workbench.

2.5.8 Microscale thermophoresis

Aptamers identified by DNA sequencing were prepared as IR680-labeled TNA molecules using the PBS8 IR680 10mer DNA primer and the corresponding H3.X template (where X represents the isolate number). A 250 µL reaction volume containing 0.5 nmol of both

primer and template,1x ThermoPol Buffer, 1 mM MnCl2, and 100 μM of each tNTP was heated for 5 min at 95°C to anneal. TNA polymerization was initiated by adding Kod-RI polymerase (~1 μM) and then incubated for 4 h at 55°C. Full length IR680-labeled aptamers were purified by 20% denaturing polyacrylamide gel electrophoresis for 3 h at 12 W constant. TNA was recovered from the gel by electroelution and then buffer exchanged into water and concentrated using a YM-10 Centricon centrifugal filter device. TNA concentration was quantified by NanoDrop absorbance. Aptamers were folded in binding buffer by denaturing for 5 min at 90°C and then cooling for 30 min at room temperature. Equilibrium binding reactions were performed in a 500 μL LoBind Eppendorf tube in a final volume of 20 μL. Folded structures at a final concentration of 1nM were incubated with HIV RT poised at concentrations spanning the expected range of K_D (typically 1 pM–500 nM) for 1 hat room temperature. Each titration sample was then loaded into a premium capillary (NanoTemper, Munich, Germany) and analyzed on the Monolith NT.115 Pico (NanoTemper) [20% MST power, 20% laser power]. Binding isotherms were deduced using NanoTemper's NT Analysis software. Aptamer deletion and split aptamer studies were performed as described above using the templates described in the sequence table. To ensure reproducibility, the binding assay for aptamer HIV-RT 3.17 was performed in triplicate.

2.5.9 Validation of binding affinity by native gel shift

Folded IR800-labeled DNA primed TNA H3.17 molecules at a final concentration of 4 nM were incubated for 1 h at room temperature with HIV-RT, serially diluted from 10 nM to 78 pM in 1x AB buffer. Each titration sample was then loaded into a single well of a 6% native TBE PAGE gel and run for 35 min at 12 W constant. Gels and running buffers were cooled to

4°C prior to running to prevent heat denaturing the aptamer-protein complex. Gels were then imaged on the LI-COR Odyssey CLx imager (Lincoln, NE). To ensure reproducibility, binding assays were performed in duplicate. For the K_D measurement at elevated temperature, the samples were incubated at 37°C for 1 h and the gel and running buffer were heated to 37°C prior to running the gel.

2.5.10 Nuclease challenge

Folded IR800-labeled aptamers (all TNA HIV-RT 3.17, FANA FA1, and DNA R1T) at a concentration of 8 nM was incubated for 1 h at 37°C with SVPE in 40 mM Tris pH 8.4, 10 mM MgCl2. The solution was then mixed with HIV RT, serially diluted from 10 nM to 78 pM in 1x AB buffer (TNA and DNA) or 50 mM Tris–HCl pH 8, 80 mM KCl, 2 mM MgCl2 and 1 mM DTT (FANA) in equal volume, so the final concentration of aptamer was 4 nM. Each sample was loaded into a 6% native TBE PAGE gel and run for 35 min at 12 W constant. Gels and running buffers were cooled to 4°C prior to running to prevent heat denaturing the aptamer-protein complex. Gels were then imaged on the LI-COR Odyssey CLx imager. To ensure reproducibility, binding was performed in duplicate. For the biological stability challenge, folded IR labeled all TNA HIV-RT3.17 at a concentration of 8 nM was incubated for 24 h at 37°C with 50% human serum (Sigma-Aldrich) or 2.0 mg/mL human liver microsomes (Sekisui XenoTech, Kansas City, KS) diluted in 1x ThermoPol. Samples were denatured with 95% formamide and 25 mM EDTA and analyzed by 20% denaturing PAGE. Gels were then imaged on the LI-COR Odyssey CLx imager. To ensure reproducibility, this experiment was performed in duplicate.

2.5.11 Temperature challenge

For the temperature challenge, all TNA H3.17 aptamer, DNA R1T aptamer, and Anti-HIV 1 Reverse Transcriptase antibody were tested for binding activity for HIVRT using Biolayer Interferometry (BLI). All BLI experiments were performed on a FortéBio (Menlo Park, CA) Octet RED384 using Streptavidin biosensors. Assays were performed in 384-well microplates at 28°C, with 50 µL volumes. Biotinylated ligands (antibody, DNA aptamer, and TNA aptamer) were each immobilized onto a streptavidin biosensor surface and a baseline was established in 1x AB buffer. These bound sensors were dipped into HIV RT in AB buffer. A basic 60 sec baseline, 300 sec association, and 900 sec dissociation method was used to determine the kinetic constants. Nonspecific binding was measured by screening sensors not exposed to ligand versus analytes under conditions identical to those of the binding assays and with the same analyte samples. Data were analyzed using the Octet System Data Analysis 9.0 Software, and constants were determined by using the global one-state fitting model on the association/dissociation curves. Gibbs free energy changes were calculated from binding affinities (dissociation constants) using the formula:

$$\Delta G^0 = -RT \ln K_D \quad (1)$$

at standard temperature and in units of J mol-1K-1. Experiments were performed to determine the ability of the HIV RT antibody, DNA aptamer and TNA aptamer to retain binding activity after exposure to heat. Each ligand (antibody, DNA aptamer, and TNA aptamer) was diluted to 50 nM in AB buffer. One aliquot of each ligand was kept on ice for the duration of the experiment and marked “time 0”. The remaining four aliquots were heated to 75°C on a calibrated heat block for 1, 8, 24, and 48 h each. After each time point, the corresponding aliquot was removed and placed in an ice bath. After the final time point, all of these samples were then tested for binding activity using BLI. Each ligand was tested

at their ideal binding concentrations of HIV RT, starting at a concentration of 125 μg/mL and ending with 915 ng/mL for the antibody, 62.5 μg/mL and ending with 487.5 ng/mL for DNA binding, and 7.8 μg/mL and ending with 122 ng/mL for TNA binding. For the PAGE analysis of thermal degradation, 50 nM samples of DNA R1T aptamer, DNA primed TNA H3.17 aptamer, and all TNA H3.17 aptamer were incubated in 1X AB buffer at 75°C. Aliquots were taken at 0, 48, and 72 hours, then placed in an ice bath. After the final time point, samples were denatured with 95% formamide and 25 mM EDTA and analyzed by 20% denaturing PAGE. Gels were then imaged on the LI-COR Odyssey CLx imager.

2.5.12 Structure-Activity Relationship Studies

The H3.17 aptamer was mapped for activity by truncating and splitting the aptamer, swapping primer sites, and correcting wobble base pairing (T→C) found in the sequence, based on the prediction of secondary structure given by the Mfold web server under RNA folding constraints, as TNA is an RNA analog (http://unafold.rna.albany.edu/?q=mfold). The predicted secondary structures are a best guess for the TNA aptamers as Mfold is not configured for TNA parameters. Complementary templates were ordered from IDT and transcribed into TNA as described above. IR680 dye-labeled DNA PBS8 primers were used for transcription to enable screening of each new aptamer by MST, as described above.

2.6 References

Baker, M. (2015). Blame it on the antibodies. *Nature*, *521*(7552), 274–276. https://doi.org/10.1038/521274a

Bala, J., Chinnapaiyan, S., Dutta, R. K., & Unwalla, H. (2018). Aptamers in HIV research diagnosis and therapy. In *RNA Biology* (Vol. 15, Issue 3, pp. 327–337). Taylor and Francis Inc. https://doi.org/10.1080/15476286.2017.1414131

Begley, C. G., & Ellis, L. M. (2012). Drug development: Raise standards for preclinical cancer research. *Nature*, *483*(7391), 531–533. https://doi.org/10.1038/483531a

Bradbury, A., & Plückthun, A. (2015). Reproducibility: Standardize antibodies used in research. In *Nature* (Vol. 518, Issue 7537, pp. 27–29). Nature Publishing Group. https://doi.org/10.1038/518027a

Chen, A., Yan, M., & Yang, S. (2016). Split aptamers and their applications in sandwich aptasensors. In *TrAC - Trends in Analytical Chemistry* (Vol. 80, pp. 581–593). Elsevier B.V. https://doi.org/10.1016/j.trac.2016.04.006

Culbertson, M. C., Temburnikar, K. W., Sau, S. P., Liao, J. Y., Bala, S., & Chaput, J. C. (2016). Evaluating TNA stability under simulated physiological conditions. *Bioorganic and Medicinal Chemistry Letters*, *26*(10). https://doi.org/10.1016/j.bmcl.2016.03.118

Dunn, M. R., Jimenez, R. M., & Chaput, J. C. (2017). Analysis of aptamer discovery and technology. *Nature Reviews Chemistry*, *1*(10). https://doi.org/10.1038/s41570-017-0076

Ellington, A. D., & Szostak, J. W. (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature*, *346*(6287). https://doi.org/10.1038/346818a0

Ferreira-Bravo, I. A., Cozens, C., Holliger, P., & DeStefano, J. J. (2015). Selection of 2′-deoxy-2′-fluoroarabinonucleotide (FANA) aptamers that bind HIV-1 reverse transcriptase with picomolar affinity. *Nucleic Acids Research*, *43*(20), 9587–9599. https://doi.org/10.1093/nar/gkv1057

Keefe, A. D., Pai, S., & Ellington, A. (2010). Aptamers as therapeutics. In *Nature Reviews Drug Discovery* (Vol. 9, Issue 7, pp. 537–550). Nature Publishing Group. https://doi.org/10.1038/nrd3141

Lai, Y.-T., & DeStefano, J. J. (2012). DNA Aptamers to Human Immunodeficiency Virus Reverse Transcriptase Selected by a Primer-Free SELEX Method: Characterization and Comparison with Other Aptamers. *Nucleic Acid Therapeutics*, *22*(3), 162–176. https://doi.org/10.1089/nat.2011.0327

Liao, J. Y., Bala, S., Ngor, A. K., Yik, E. J., & Chaput, J. C. (2019). P(V) Reagents for the Scalable Synthesis of Natural and Modified Nucleoside Triphosphates. *Journal of the American Chemical Society*, *141*(34), 13286–13289. https://doi.org/10.1021/jacs.9b04728

Mei, H., Liao, J.-Y., Jimenez, R. M., Wang, Y., Bala, S., McCloskey, C., Switzer, C., & Chaput, J. C. (2018). Synthesis and Evolution of a Threose Nucleic Acid Aptamer Bearing 7-Deaza-7-Substituted Guanosine Residues. *Journal of the American Chemical Society*, *140*(17). https://doi.org/10.1021/jacs.7b13031

Nikoomanzar, A., Vallejo, D., & Chaput, J. C. (2019). Elucidating the Determinants of Polymerase Specificity by Microfluidic-Based Deep Mutational Scanning. *ACS Synthetic Biology*, *8*(6), 1421–1429. https://doi.org/10.1021/acssynbio.9b00104

Noronha, A. M., Wilds, C. J., Lok, C. N., Viazovkina, K., Arion, D., Parniak, M. A., & Damha, M. J. (2000). Synthesis and biophysical properties of arabinonucleic acids (ANA): Circular dichroic spectra, melting temperatures, and ribonuclease H susceptibility of ANA·RNA hybrid duplexes. *Biochemistry*, *39*(24), 7050–7062. https://doi.org/10.1021/bi000280v

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. (2012). Synthetic genetic polymers capable of heredity and evolution. *Science*, *336*(6079), 341–344. https://doi.org/10.1126/science.1217622

Roberts, R. W., & Szostak, J. W. (1997). RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(23), 12297–12302. https://doi.org/10.1073/pnas.94.23.12297

Schoning, K. U., Scholz, P., Guntha, S., Wu, X., Krishnamurthy, R., & Eschenmoser, A. (2000). Chemical etiology of nucleic acid structure: The α-threofuranosyl-(3'→2') oligonucleotide system. *Science*, *290*(5495), 1347–1351. https://doi.org/10.1126/science.290.5495.1347

Vater, A., & Klussmann, S. (2015). Turning mirror-image oligonucleotides into drugs: The evolution of Spiegelmer® therapeutics. In *Drug Discovery Today* (Vol. 20, Issue 1, pp. 147–155). Elsevier Ltd. https://doi.org/10.1016/j.drudis.2014.09.004

Wilson, D. S., & Szostak, J. W. (1999). In Vitro Selection of Functional Nucleic Acids. *Annual Review of Biochemistry*, *68*(1). https://doi.org/10.1146/annurev.biochem.68.1.611

Yu, H., Zhang, S., & Chaput, J. C. (2012). Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. *Nature Chemistry*, *4*(3), 183–187. https://doi.org/10.1038/nchem.1241

Zhou, J., & Rossi, J. (2017). Aptamers as targeted therapeutics: Current potential and challenges. In *Nature Reviews Drug Discovery* (Vol. 16, Issue 3, pp. 181–202). Nature Publishing Group. https://doi.org/10.1038/nrd.2016.199

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, *31*(13), 3406–3415. https://doi.org/10.1093/nar/gkg595

CHAPTER 3

Publication Note

This manuscript is in preparation for submission. McCloskey, C. M.;* Li, Q.;* Yik, E. J.; Chim, N.; Ngor, A.; Medina, E.; Grubisic, I.; Keh, L. C. T; and Chaput, J. C. Evolution of Functionally Enhanced TNA Aptamers.

*Signifies co-first author status.

3.1 Contribution Statement

J.C. and I.G. conceived of the project and designed the experiments. C.M., Q.L., E.Y., N.C., A.N., E.M., and I.G. performed the experiments. J.C. wrote the manuscript with drafts from C.M. and N.C. All authors reviewed and commented on the manuscript.

3.2 Abstract of Chapter

Synthetic genetic polymers (XNAs) have the potential to transition aptamers from laboratory tools to therapeutic agents, but additional functionality is needed to compete with antibodies. Here we describe the evolution of a biologically stable artificial genetic system comprised of α-L-threofuranosyl nucleic acid (TNA) that facilitates the production of backbone- and base-modified aptamers termed 'threomers' that function as high quality protein capture reagents. Threomers were discovered against two prototypical protein targets implicated in human diseases through a combination of in vitro selection and next-generation sequencing using uracil nucleotides that are uniformly equipped with aromatic side chains commonly utilized by antibodies in their recognition of protein antigens. Kinetic measurements reveal that side chain modifications are critical for generating threomers with slow off-rate binding kinetics. These findings expand the chemical space of

evolvable non-natural genetic systems to include functional groups that enhance protein target binding by mimicking the structural properties of traditional antibodies.

3.3 Introduction

The effectiveness of antibodies in diagnostic and therapeutic applications have inspired efforts to explore the chemical space of evolvable non-natural genetic systems in search of sequence-defined macromolecules that can recapitulate antibody binding by folding into shapes that recognize disease-associated proteins with high affinity and high specificity (Brudno & Liu, 2009; Dunn et al. 2017; Keefe et al., 2010; Zhou & Rossi, 2017). Toward this goal, much attention has been given to the establishment of Darwinian evolution systems that allow for the isolation of artificial genetic polymers with nucleobase or sugar-phosphate backbone modifications that function with enhanced target binding affinity or elevated biological stability (Nikoomanzar et al., 2020; Packer & Liu, 2015). However, despite many notable accomplishments, including the evolution of novel synthetic genetic polymers (XNAs) with backbone chemistries not found in nature (Pinheiro et al., 2012; Yu et al., 2012) and the creation of highly functionalized nucleic acid polymers (HFNAPs) (Chen et al., 2018; Kong et al., 2017), the ability to generate affinity reagents that are both recalcitrant to biological nucleases and capable of recognizing their cognate protein target with low dissociation rate constants has proven elusive. Since slow-off rate binding is a hallmark of a high quality antibody, renewed efforts are needed to establish examples of biologically stable affinity reagents that mirror the binding properties of the best monoclonal antibodies.

Previously, we have developed an in vitro selection system that allows for the rapid isolation of high affinity xeno-nucleic acid (XNA) aptamers by covalently linking

synthesized XNA strands to their encoding DNA template (Yu et al., 2012). The DNA display configuration (Figure 3.1a) is conceptually similar to the protein evolution technique of mRNA display, whereby a genotype-phenotype relationship is created by covalently linking freshly translated proteins to their encoding mRNA strands (Roberts & Szostak, 1997). This approach allows for stringent washing conditions to be employed as functional XNA aptamers (phenotype) that remain bound to the protein are directly amplified using the polymerase chain reaction (PCR) to copy the DNA coding region (genotype) of DNA-TNA chimera (Roberts & Szostak, 1997). DNA display simplifies the selection cycle by avoiding the need for an XNA reverse transcriptase. In a recent demonstration, we applied the DNA display strategy to evolve α-L-threofuranosyl nucleic acid (TNA, Figure 3.1b) aptamers against the protein target human HIV reverse transcriptase (Dunn et al., 2020). Because TNA is recalcitrant to nuclease digestion (Culbertson et al., 2016), the evolved aptamers remained active in the presence of biological nucleases (Dunn et al., 2020). Similar results have also been achieved with mirror image aptamers (Spiegelmers), but such reagents are restricted to protein targets that are accessible to solid-phase peptide synthesis (Vater & Klussman, 2015).

We postulated that the site-specific introduction of chemical groups into the nucleobase moiety would allow TNA aptamers to function with enhanced binding affinity relative to equivalent molecules produced from TNA libraries having only standard nucleobases. We focused our attention on planar aromatic side chains (Figure 3.1b), such as phenylalanine (Phe) and tryptophan (Trp) that are overrepresented at the paratope-epitope interface of antibody-antigen co-crystal structures (Kringelum et al., 2013; Ramaraj et al., 2012).

Several studies have reported the use of similar hydrophobic moieties installed on the nucleobases of other DNA-encoded affinity reagents (Cheung et al., 2020; Gawande et al., 2017; Kimoto et al., 2013; Ren et al., 2017; Tolle et al. 2015; Vaught et al., 2010), a subset of which have been solved by X-ray crystallography as binary protein-aptamer complexes (Gelinas et al., 2016). However, the synthetic challenges of generating evolvable systems equipped with diversity-enhancing functional groups has prevented the application of this concept to synthetic genetic polymers with backbone structures that require laboratory evolved polymerases for their synthesis and evolution in a test-tube. Here we demonstrate that chemically modified uracil residues uniformly incorporated into TNA library repertoires enable the evolution of biologically stable TNA aptamers (threomers) that function with enhanced binding affinity. The ability to generate threomers with stable backbone structures and functionally enhanced nucleobases offers hope that aptamers may soon be able to compete with antibodies in the therapeutic arena.

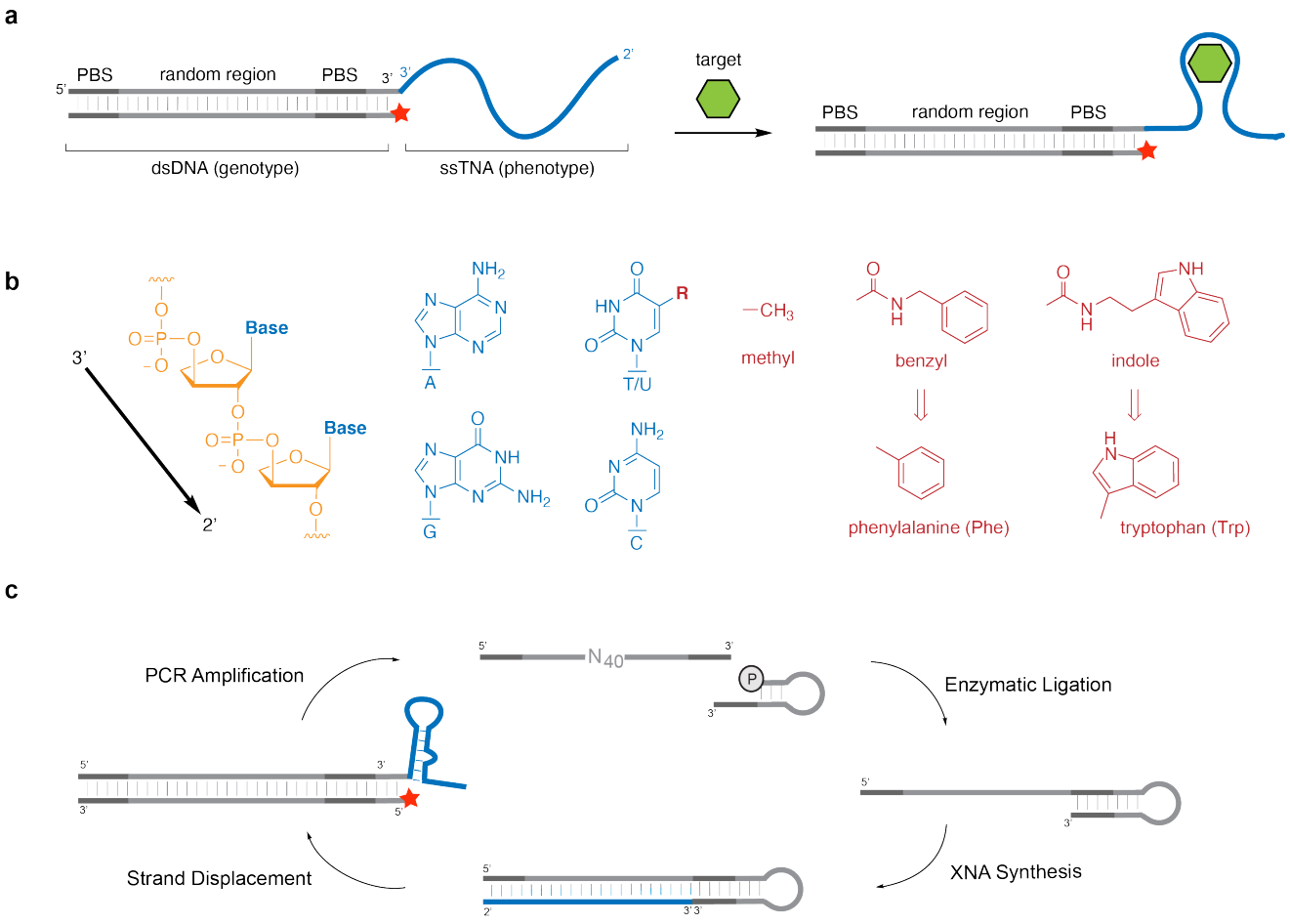


Figure 3.1. DNA display approach to evolving functionally enhanced TNA aptamers. a, DNA display design showing the genotype-phenotype relationship of encoding double-stranded (ds) DNA molecules covalently linked to their single-stranded (ss) TNA. TNA molecules that bind a desired protein target are recovered and PCR amplified using primers that are complementary to the primer binding sites (PBS) in the encoding dsDNA. b, Constitutional structure for the linearized backbone of 3',2'-a-L-threofuranosyl nucleic acid (TNA) with natural and chemically modified bases mimicking the amino acids phenylalanine and tryptophan. c, Replication cycle. DNA molecules (grey) ligated to a self-priming stem-loop structure are extended with TNA (blue), strand displaced by copying a DNA primer annealed to the loop with DNA, and amplified by PCR.

3.4 Results

In an effort to improve the quality of TNA aptamers produced by in vitro selection, we chose to evaluate the potential for aromatic side chains to enhance the functional properties of large random-sequence TNA libraries. Our study focused on tUTP derivatives that are chemically modified at the C5 position with Phe and Trp side chains (Figure 3.1b). Importantly, crystal structures of Kod DNA polymerase bound to modified DNA provided confidence that our derivatized tUTP nucleotides would be viable substrates for TNA synthesis (Hottin & Marx, 2016). The structures reveal a large cavity in the enzyme active site that can accommodate bulky substituents in the major groove of the DNA duplex. A recent crystal structure of Kod-RSGA (unpublished data) indicates that the large active site cavity is maintained in our best laboratory evolved TNA polymerase (Nikoomanzar et al., 2020).

We began by synthesizing the desired C5 modified tUTP substrates (SI Figure 3.1) following a palladium cross-catalyzed carboxyamidation reaction that couples 5-iodo-1-(2'-*O*-benzoyl-α-L-threofuranosyl)-uracil to aromatic amines via in situ formation of an amide linkage (Vaught et al., 2010). Although other conjugation chemistries are available for linking functional groups to uracil nucleosides (Tolle et al., 2015), we viewed this strategy as a highly versatile route to uniformly modified TNA oligonucleotides. In addition to broad chemical reactivity, the resulting amide linkage (Figure 3.1b) benefits from reduced hydrophobicity relative to carbon linkers and limits the number of rotatable bonds between the functional group and nucleobase.

The starting 5-iodo-1-(2'-*O*-benzoyl-α-L-threofuranosyl)-uracil nucleoside was prepared in eight steps from L-ascorbic acid (vitamin C) using a Vorbrüggen reaction to

conjugate 5-iodo-uracil to an orthogonally protected threose sugar (Supplementary Scheme 3.1) (Mei & Chaput, 2018). Subsequent conjugation of the aromatic side chain to the nucleobase was accomplished in separate reactions by heating either the benzyl or tryptophan amine with the TNA nucleoside in the presence of 10 mol% $Pd(Ph_3)_4$, and CO (Supplementary Schemes 3.2-3.3). After purification by silica gel chromatography, the C5 modified uracil nucleosides were converted to their corresponding nucleoside-3'-triphosphates by transforming the 3'-hydroxyl group into an activated nucleoside 3'-monophosphate that was coupled to pyrene pyrophosphate to produce a fully protected TNA nucleoside-3'-triphosphate (Liao et al., 2019). Following purification, the triphosphate precursor was deprotected in concentrated ammonium hydroxide and precipitated to afford the desired α-L-threofuranosyl uracil nucleoside triphosphates carrying the phenylalanine and tryptophan side chains at the C5 position. This study also required the preparation of all four TNA triphosphates (SI Figure 3.1) with the natural bases of adenine (tATP), cytosine (tCTP), thymine (tTTP), and guanine (tGTP), which were each synthesized in 13 steps from vitamin C (Supplementary Scheme 3.4) (Liao et al., 2019; Sau et al., 2016).

We confirmed that base modified TNA triphosphates were viable substrates for TNA synthesis by evaluating their incorporation into TNA oligonucleotides using a standard primer extension assay. Accordingly, an IR-labeled DNA primer annealed to a DNA template was incubated with a mixture of chemically synthesized tNTPs and Kod-RSGA for 2 hours at 55 °C. In these reactions, the tNTP mixture contained either standard bases only or a tNTP solution in which the tTTP substrate was replaced with either $tUTP^{Phe}$ or $tUTP^{Trp}$. Analysis of the primer extension reactions by denaturing polyacrylamide gel electrophoresis (PAGE) indicates that the primer was extended to full-length product in all

cases, as evidenced by the presence of a discrete slower moving band in each lane of the gel (SI Figure 3.2). Definitive evidence of $tUTP^{Phe}$ and $tUTP^{Trp}$ incorporation into the TNA oligonucleotides was obtained by mass spectrometry (SI Figure 3.3).

Encouraged by the synthesis and successful enzymatic incorporation of base-modified tNTPs into TNA, we prepared three chemically distinct libraries containing 10^{14} TNA oligonucleotides, each displayed on their encoding double-stranded (ds) DNA (Figure 3.1a). The libraries were designed at the DNA level to contain a 40 nucleotide (nt) random region that was flanked on the 5' and 3' sides with fixed-sequence primer binding sites for PCR amplification. Adjacent to the random region was a 6 nt bar code identifying the library chemotype (standard, Phe, and Trp). The libraries were designed to randomly sample different regions of TNA polymer space for individual sequences that could fold into structures with high affinity to a desired protein target. In particular, we were interested in evaluating the distribution of threomers that arise when the same in vitro selection protocol is applied to different TNA library chemotypes. We operated under the hypothesis that functionally enhanced libraries should yield a larger abundance of TNA aptamers with higher binding affinity to the desired protein target due to their ability to mimic the protein-protein interface.

As a model system, we chose to enrich for TNA sequences that bound to the spike protein (S1) of SARS-CoV-2 (Figure 3.2a), the causative pathogen of COVID-19 (Wrapp et al., 2020; Yan et al., 2020). Successive rounds of in vitro selection and amplification (Figure 3.1c) were performed in parallel by ligating a DNA stem-loop structure onto a chemically synthesized single-stranded (ss) DNA library. The ligated stem-loop structure serves as the DNA primer for TNA extension across from the DNA template. Kod-RSGA mediated

extension of the primer with tNTPs yields a chimeric TNA-DNA heteroduplex, which undergoes strand displacement in a second primer extension step using Bst DNA polymerase to extend a DNA primer annealed to the loop portion of the hairpin with dNTPs. The product of strand displacement is a library of TNA molecules that are each physically linked to their encoding dsDNA templates.

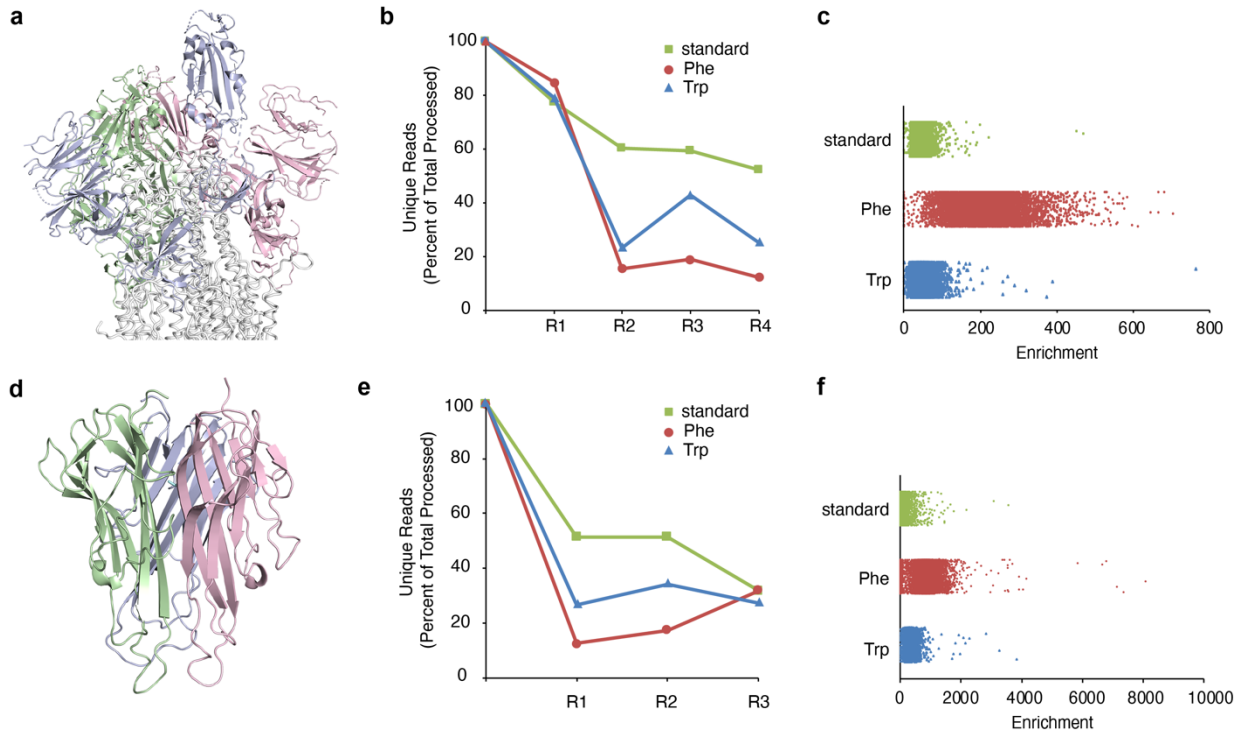


Figure 3.2. Selection performance of standard and functionally enhanced libraries. a, Cryo-EM structure of the SARS-CoV-2 spike protein trimer, with the S1 domains colored in blue, green, and pink (PDB ID: 6VSB). The RBD domain of the blue subunit is in the 'up' position. b, Unique reads obtained by bioinformatic analysis of the NGS data collected over the course of the selection against S1. c, Enrichment values obtained for each library chemotype against S1. d, X-ray crystal structure of the TNFα trimer, subunits colored blue,

green, and pink (PDB ID: 2TNF). e, Unique reads over the course of the selection against TNFα. f, Enrichment values obtained for each library chemotype against TNFα.

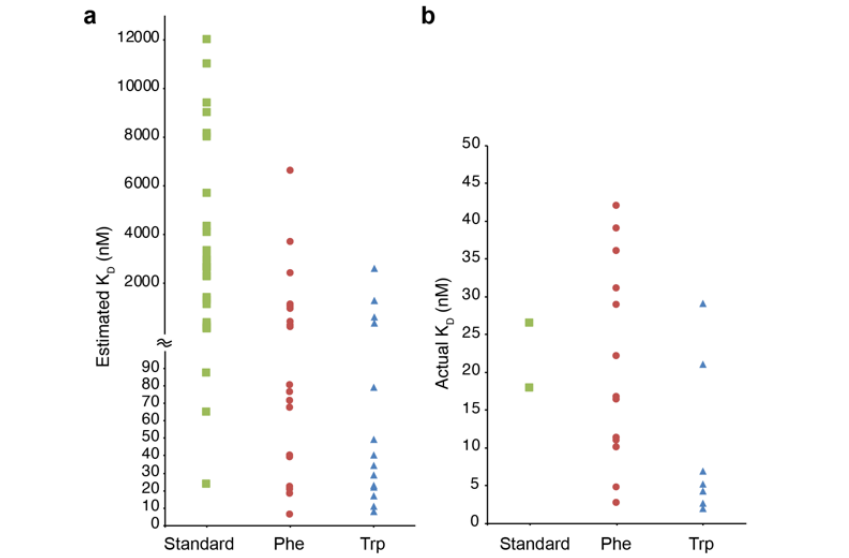
For each round of selection, the TNA libraries were incubated with free Ni-NTA beads to remove any TNA sequences that bound nonspecifically to the solid-support matrix. In round 1, a second negative selection step was performed on the naïve library using His-tagged hemagglutinin (HA), which was chosen as a generic off-target viral-coat protein. The eluted material was incubated with the S1 protein containing a C-terminal His-6 tag. Aptamers that bound specifically to the S1 protein were partitioned away from the unbound material by passing the solution over Ni-NTA beads to capture the aptamer-protein complexes. The beads were washed under high ionic strength conditions to remove weak affinity sequences that bound primarily through electrostatic interactions and the bound material was eluted from the Ni-NTA beads with imidazole. The eluted fractions were exchanged into water and amplified by PCR. Regeneration of the ssDNA library required a second PCR step using a PEG-modified DNA primer that allows for size separation of the PEGylated DNA template by denaturing PAGE. The ssDNA material served as the starting point for the next round of in vitro selection and amplification.

After round 4, the evolving pools of TNA sequences were subjected to high throughput next-generation DNA sequencing (NGS) using an Illumina NovaSeq platform to sequence the elution fractions from each round of selection (Supplementary Table 3.2). The combined power of DNA display and NGS allows us to bypass the need for extensive rounds of selection by identifying sequences with enrichment profiles that are indicative of high affinity binders. Bioinformatic analysis of the data reveals that populations containing the

Phe and Trp side chains converged more rapidly toward a smaller number of unique sequences as compared to the unmodified library composed entirely of standard bases (Figure 3.2b). Enrichment calculations performed on the top 10,000 most abundant sequences observed after four rounds of protein-binding selection indicate a strong enrichment in TNA sequences with Phe side chains. By comparison, TNA sequences equipped with the Trp side chain show only a modest improvement over the unmodified library. However, it's interesting to note that the most enriched sequence derives from the Trp chemotype, suggesting that the Trp population may nevertheless contain several high affinity binders even though the pool exhibited a lower-than-expected enrichment profile.

To better illuminate the relationship between sequence enrichment and function, members of each chemotype family were screened for affinity to the S1 protein. In this experiment, 65 sequences with promising enrichment profiles were selected and individually synthesized by polymerase-mediated primer extension on complementary DNA templates. Estimated binding affinities (K_D) were determined by biolayer interferometry (BLI) using biotinylated aptamers that were immobilized onto the surface of streptavidin-coated biosensors and assayed for binding activity to S1 poised at a protein concentration of 100 nM. The screen revealed striking differences in the number of high-affinity aptamers produced from the modified libraries versus those produced from the standard base library. A plot of estimated K_D value versus library chemotype highlights these differences, which bin according to 3380 nM (n=30), 910 nM (n=20), and 430 nM (n=15) for the average binding affinities observed for the standard, Phe, and Trp libraries, respectively (Figure 3.3a). Analysis of the data indicates that more than half of the

functionally enhanced aptamers are predicted to bind the S1 protein with a K_D of <90 nM versus only 10% for the standard base library. In addition to reduced binding affinity, the



Entry	K_D (nM)	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})
1	17.6	1.6×10^{4}	2.7×10^{-4}
2	27.3	9.0×10^{3}	2.5×10^{-4}
3	2.6	9.0×10^{4}	2.4×10^{-4}
4	5.2	1.4×10^{4}	7.5×10^{-5}
5	10.0	1.8×10^{4}	1.8×10^{-4}
6	11.3	1.5×10^{4}	1.7×10^{-4}
7	12.0	2.5×10^{4}	3.0×10^{-4}
8	16.3	1.2×10^{4}	2.0×10^{-4}
9	16.7	1.2×10^{4}	1.9×10^{-4}
10	21.8	1.2×10^{4}	2.7×10^{-4}
11	22.1	9.5×10^{3}	2.1×10^{-4}
12	28.8	9.1×10^{3}	2.6×10^{-4}
13	31.0	2.0×10^{4}	6.2×10^{-4}
14	36.0	9.4×10^{3}	3.3×10^{-4}
15	39.0	4.3×10^{3}	1.7×10^{-4}
16	42.0	5.2×10^{3}	2.2×10^{-4}
17	2.0	6.7×10^{4}	1.3×10^{-4}
18	2.7	5.2×10^{4}	1.4×10^{-4}
19	4.2	3.3×10^{4}	1.4×10^{-4}
20	5.1	2.1×10^{4}	1.1×10^{-4}
21	6.9	9.7×10^{4}	6.7×10^{-4}
22	21.1	8.7×10^{4}	1.8×10^{-3}
23	29.1	3.9×10^{4}	1.1×10^{-3}
24	76.0	7.0×10^{3}	5.4×10^{-4}

Figure 3.3. Kinetic analysis of aptamer binding to the S1 domain. a, BLI screen of highly enriched sequences at a defined S1 concentration of 100 nM. The Phe and Trp modified libraries (n=20 and 15, respective) exhibit higher affinity binding than the standard base library (n=30). b, Full kinetic analysis of 24 high performing TNA aptamers provided a detailed view of the binding kinetics of modified and unmodified TNA aptamers. c, Binding curves of representative high performance TNA aptamers isolated from the unmodified

(top) and modified (bottom) TNA libraries. Standard-2 and Trp-1 correspond to entries 2 and 17 in the table provided in panel b.

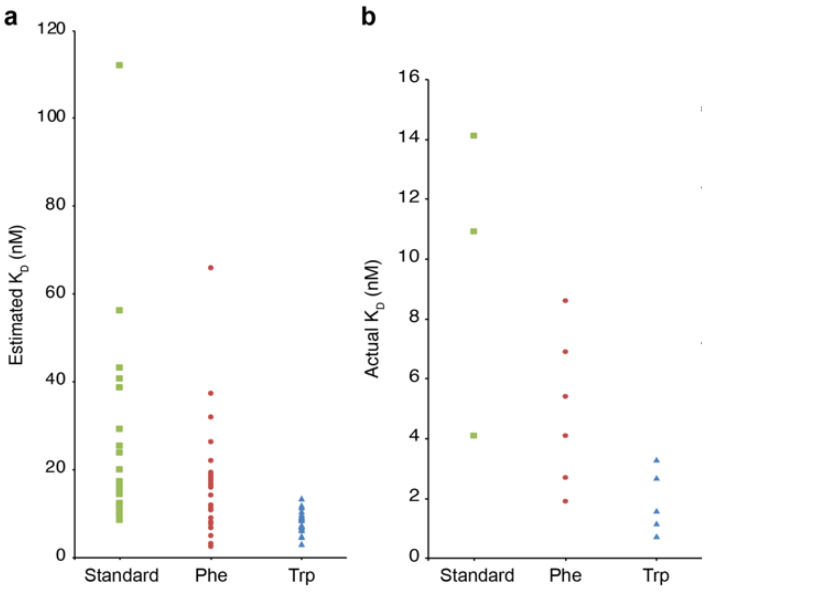
standard library shows a wider range of K_D values, indicating the pool either contains fewer high affinity binders or is less convergent, which is consistent with the higher number of unique reads observed in the NGS data (Figure 3.2b). Together, these data support the hypothesis that expanded chemical functionality is necessary for generating threomers with higher protein binding affinity.

To more precisely evaluate the binding properties of the in vitro selected TNA aptamers isolated from the three chemotype libraries, full kinetic measurements were performed on a subset of the sequences that were predicted to bind S1 with a K_D value of <50 nM. In total, 24 TNA aptamers were selected for further characterization (Figure 3.3b, Supplementary Table 3.3). Curve fitting reveals that all of the data conform to a 1:1 binding model with the top 8 binders having K_D value in the range of 1 – 10 nM, and deriving exclusively from the functionally enhanced libraries. In general, TNA aptamers isolated from functionally enhanced libraries bind S1 with K_D values that are >10-fold higher than aptamers isolated from the standard library. Moreover, modified aptamers exhibit higher binding responses (>2-fold) to the aptamer-coated biosensor than unmodified aptamers, indicating a stronger aptitude for overall binding to the S1 protein.

To confirm the reproducibility of our data, we evaluated a representative high performance S1 aptamer from the standard and modified chemotype libraries in triplicate using independently synthesized material for each kinetic binding assay. The best modified and unmodified S1 aptamers (designated Trp-1 and standard-2, respectively) have average

affinity values of 3.1 ± 1.0 nM and 34 ± 11 nM, respectively, for the S1 protein of SARS-CoV-2 (Figure 3.3c, Supplementary Table 3.5). These values, which closely approximate the K_D values observed in the original kinetic titration assay, provide confidence that the data provided in Figure 3.3b is an accurate reflection of the binding activity of the evolved threomers. Analysis of the BLI sensorgrams indicates the enhanced activity of Trp-1 is due to a combination of faster on-rates, slower-off rates, and higher overall binding to the aptamer-coated biosensor than the unmodified standard-2 aptamer.

To more fully sample the sequence space adjacent to the Trp-1 sequence, a doped DNA library was synthesized in which 85% of the amidite solution used corresponded to the nucleotide present in Trp-1, while the other 15% contained an equal mixture of the other three nucleobases. After a single round of selection, the most abundant sequence remained the seed, Trp-1. Full kinetic characterization was performed on the top ten most abundant sequences, revealing K_Ds from 3-10 nM, similar to the seed sequence.



Entry	K_D (nM)	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})
1	4.1	1.1×10^6	4.4×10^{-3}
2	10.9	1.2×10^6	1.3×10^{-2}
3	14.1	7.5×10^5	1.1×10^{-2}
4	1.9	2.6×10^6	5.0×10^{-3}
5	2.7	2.9×10^6	7.8×10^{-3}
6	4.1	3.0×10^6	1.2×10^{-2}
7	5.4	1.7×10^6	9.1×10^{-3}
8	7.0	2.2×10^6	1.5×10^{-2}
9	8.6	2.3×10^6	2.0×10^{-3}
10	0.7	9.8×10^6	7.1×10^{-3}
11	1.2	4.6×10^6	5.4×10^{-3}
12	1.6	3.3×10^6	5.2×10^{-3}
13	2.7	2.6×10^6	7.1×10^{-3}
14	3.3	4.2×10^6	1.4×10^{-2}

Figure 3.4. Binding kinetic analysis of aptamers to TNFα. a, Single concentration screens of highly enriched sequences show a wider spread of estimated K_Ds for standard bases (green) than for Phe- (blue) and Trp-modified sequences (red). Trp-modified aptamers again show the tightest spread and highest affinities. b, Full kinetic analysis performed on 14 aptamers with high estimated affinity. Functionally enhanced sequences exhibit higher average binding affinity, and in most cases, slower k_{off} values. c, Binding curves of exemplary standard base aptamer (top) and Trp-modified aptamer (bottom) showcase the

difference in affinity as well as significantly slower k_{off} for the functionally enhanced aptamer, Trp-13.

To evaluate the generality of functionally enhanced TNA libraries as a rich source of chemical diversity for producing superior TNA-based affinity reagents, we performed a similar in vitro selection experiment against a different protein target. For this study, we chose tumor necrosis factor-alpha (TNFα, Figure 3.2d), a cytokine that is closely linked to inflammatory diseases (Kalliolias & Ivashkiv, 2016). NGS analysis of the elution fractions obtained after 3 rounds of in vitro selection and amplification for TNA aptamers with affinity to TNFα indicates that TNA libraries carrying the Phe and Trp chemotypes exhibit higher levels of enrichment than the standard base library (Figure 3.2e,f). An activity screen of 66 TNA aptamers reveals an estimated K_D range of 24 – 1123 nM with most of the Phe- and Trp-modified aptamers having K_D values that cluster below 200 nM (Figure 3.4a). A full kinetic analysis of the top performing aptamers from each chemotype library uncovered a strong preference for Trp-modified aptamers (Figure 3.3b). A comparison of representative high performing standard and modified TNA aptamers (Figure 3.3c) reveals that the higher activity of the modified aptamers is primarily due to slower off-rates in the protein-aptamer binding equilibrium. In this case, the best aptamers isolated from the standard and modified chemotype libraries (designated standard-2 and Trp-3, respectively) bind with average K_D values of 82 ± 24 nM and 20 ± 6.1 nM (Figure 3.4c, Supplementary Table 3.5), respectively, when assayed as three independent replicates. As with the S1 aptamers, modified aptamers immobilized on the streptavidin-coated biosensor exhibit higher binding responses than the unmodified aptamers. These data

further support the general conclusion that functionally enhanced TNA libraries improve the quantity and quality of high activity threomers generated by in vitro selection.

Last, we wished to evaluate the contribution of the modified side chains toward the activity and specificity of threomer binding to the S1 and TNFα target proteins. We evaluated the importance of the hydrophobic amino acid side chains by measuring the binding affinity of modified and unmodified versions of the Trp-1 and Trp-3 aptamers selected to bind S1 and TNFα, respectively. In both cases, the resulting BLI sensorgrams clearly show that protein binding affinity is completely abrogated when the in vitro selected TNA sequences are synthesized with the natural base chemotype (Figure 3.5 a). To investigate the potential for protein target binding specificity, we challenged the Trp-1 and Trp-3 aptamers to bind closely related protein targets. The resulting BLI sensorgram reveals that Trp-1 exhibits a strong preference for the S1 protein of SARS-CoV-2 versus SARS-CoV-1 (Figure 3.5b). An even more striking result was observed for the Trp-3 threomer, which shows no detectable binding for TNFβ (Figure 3.5d). Interestingly, our BLI analysis also revealed that Trp-1 binds outside the RBD domain of the S1 protein, suggesting that additional screening or selections could lead to an RBD binder that could be used in a sandwich assay or as a therapeutic candidate.

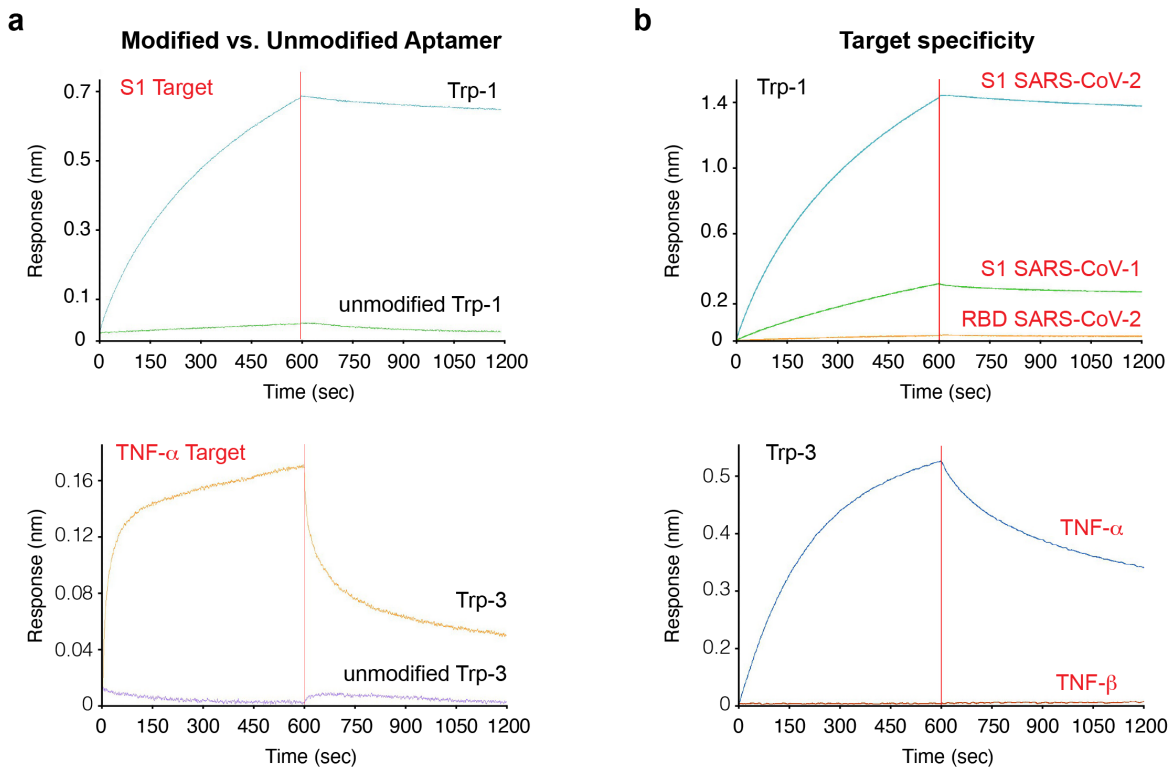


Figure 3.5. Protein binding activity and specificity of the S1 and TNFa threomers. a, BLI sensorgrams observed for the Trp-1 and Trp-3 aptamers containing and lacking the indole chemotype against the S1 protein of SARS-CoV-2 and TNFα, respectively. b, BLI sensorgrams showing the specificity of the Trp-1 S1 SARS-CoV-2 aptamer against the S1 protein of SARS-CoV-1 and the RBD domain of SARS-CoV-2, and separate BLI sensorgrams showing the specificity of the Trp-3 TNFα aptamer against TNFβ.

3.5 Discussion

The data collected in this study provides compelling evidence that functionally enhanced TNA libraries produce higher affinity binders than standard TNA libraries carrying only natural bases. This result was achieved through a combination of DNA display and high throughput NGS sequencing, which allowed for the isolation of base-modified threomers with K_D values in the low nM range after only 3-4 rounds of selective amplification. The

absence of significant numbers of high affinity binders from the unmodified library supports the hypothesis that planar aromatic side chains have the ability to mimic amino acid residues that are overrepresented at the paratope-epitope interface of antibody-antigen complexes. The generality of the conjugation chemistry established to construct uniformly modified TNA libraries is sufficiently versatile that it should be possible to explore new regions of chemical space by synthesizing a broader range of C5 modified tUTP substrates. As these studies continue, it will be interesting to learn which side chains are best suited for a given protein target, as such information could one day be used to create 'smart' libraries that are better able to converge on optimal solutions to a given biochemical problem.

The path to functionally enhanced XNA libraries was limited by a number of synthetic challenges that are less severe for DNA-based affinity reagent technologies. TNA, like most XNAs, is assembled from building blocks that are not commercially available and must be obtained by chemical synthesis. In our case, a complete monomer set of all four TNA nucleoside triphosphates (tNTPs) requires 52 chemical transformation to convert vitamin C into each of the four tNTPs (tATP, tCTP, tTTP, and tGTP) (Liao et al., 2019; Sau et al., 2016). The $tUTP^{Phe}$ and $tUTP^{Trp}$ substrates demanded an additional 30 chemical transformations (15 steps each) with the critical step being the palladium-catalyzed cross-coupling reaction required to functionalize the uracil nucleobase with an aromatic side chain. Once the substrates are prepared, it then becomes necessary to identify a polymerase that can synthesize functionally enhanced TNA oligonucleotides. In the current study, this was accomplished using Kod-RSGA, a recently evolved TNA polymerase developed for enhanced TNA synthesis activity on DNA templates (Nikoomanzar et al., 2020). However, future studies would benefit from a new polymerase variant that can recognize C5 modified tUTP

substrates more efficiently than Kod-RSGA. The modest activity of Kod-RSGA toward modified tNTP substrates may explain the observed discrepancy between enrichment and affinity in the Phe- and Trp-modified libraries in which Trp threomers showed lower enrichment but higher binding affinity in both of the selection experiments.

Although antibodies remain the gold standard as protein affinity reagents, aptamers have a number of advantages that have caused them to grow in popularity (Dunn et al., 2017). In addition to prolonged storage and shipping at ambient temperature, aptamers are produced by cell-free synthesis (chemical or enzymatic), which avoids viral or bacterial contamination problems associated with cellular protein production systems. Aptamers have low immunogenicity, low batch-to-batch variability, and greater ease of chemical modifications than antibodies and other protein-based affinity reagents. However, despite these advantages, aptamers have historically suffered from poor biological stability and fast off-rates. The current study is part of a program of research designed to narrow the gap between antibodies and aptamers. Our focus has been to establish an artificial genetic system that is both biologically stable and amenable to Darwinian evolution, and then to augment this system with the chemical functionality required to achieve high affinity interactions that are driven by slow off-rate binding kinetics. Future studies will explore new technological advances for identifying threomers that function with high affinity and specificity.

In conclusion, we have established threomers as new biologically stable affinity reagents with enhanced functional activity. We expect that further advances in aptamer technology will pave the way towards the discovery of new therapeutic and diagnostic agents that are better equipped to compete with traditional antibody technology.

3.6 Experimental Details

3.6.1 General Information

Taq DNA polymerase, Bst 3.0 DNA polymerase, T4 DNA ligase, 10x T4 DNA ligase buffer, and 10x ThermoPol buffer were purchased from New England Biolabs (Ipswich, MA). PCRBIO HiFi polymerase was purchased from PCRBiosystems (Wayne, PA). Kod-RSGA TNA polymerase was expressed and purified as described previously1. Experiments were performed in DNA LoBind tubes, purchased from Eppendorf (Hamburg, Germany). DNA triphosphates were purchased from Thermo Fisher Scientific (Waltham, MA). TNA triphosphates bearing natural bases were synthesized as previously described. DNA oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), purified by denaturing polyacrylamide gel electrophoresis (PAGE), electroeluted, buffer exchanged and concentrated using Millipore YM-3 or YM-30 Centricon centrifugal filter units, and quantified by UV absorbance via NanoDrop (Thermo Fisher Scientific). For the aptamer selection, Ni-NTA agarose resin was purchased from Qiagen (Hilden, Germany). His-tagged SARS-CoV-2 spike (S1) protein, SARS-CoV-2 receptor binding domain (RBD) of S1, and SARS-CoV-1 S1 protein were purchased from ACROBiosystems (Newark, DE). His-tagged TNFα was expressed from E. coli as previously described. His-tagged TNFα was purchased from Sino Biological (Wayne, PA). Streptavidin and Ni-NTA biosensor were purchased from ForteBio (Fremont, CA).

3.6.2 Library synthesis

Three libraries were synthesized at the DNA level using a custom nucleotide distribution to minimize the occurrence of G-quadruplex structures. Each library contained a unique 6 nt barcode located between the forward primer and random region to signify

the library chemotype (standard, Phe, and Trp). The initial libraries containing an internal variable region of 40 nucleotides (30% A, 30% T, 20% G, 20% C) flanked on both sides by fixed-sequence primer binding sites. The libraries were purchased from the Keck Oligonucleotide Synthesis Facility (Yale University) and purified by 10% denaturing PAGE. The band corresponding to full-length product was excised, electroeluted, desalted, and quantified by UV absorbance. The phosphorylated DNA hairpin primer was ligated to the library by combining 5 nmol DNA library with 6 nmol of DNA hairpin in a final volume of 1 mL of 1x T4 DNA ligase buffer. The solution was denatured for 5 min at 95°C and then annealed by incubating for 30 min at room temperature. Once annealed, 10,000 U of T4 DNA ligase was added and the reaction was incubated overnight at 24°C. The following day, the hairpin library was purified by 10% denaturing PAGE and quantified by UV absorbance.

The hairpin library was extended with tNTPs to form a chimeric TNA:DNA hairpin heteroduplex. In a 1 mL reaction volume containing 1 nmol hairpin library and 1x ThermoPol buffer was heated for 5 min at 95°C to denature and then annealed by incubating for 30 min at 24°C. TNA polymerization was initiated by adding Kod RSGA (1 μM) and 100 μM of each tNTP (tATP:tCTP:tGTP:tTTP or $tUTP^{Phe}$ or $tUTP^{Trp}$) and incubating for 2 h at 55°C. Excess polymerase was extracted with phenol:chloroform:isoamyl alcohol (25:24:1, saturated with 10 mM Tris, pH 8.0, 1 mM EDTA). Following extraction, the TNA:DNA hairpin library was concentrated and the remaining tNTPs were removed by buffer exchange into water using a YM-30 Centricon filter device. Library was quantified by UV absorbance.

Next, the TNA strand was displaced by extending a DNA primer annealed to the hairpin loop with dNTPs. Accordingly, the TNA:DNA library (1 μM final) was resuspended in 1x

ThermoPol buffer containing 500 μM dNTPs and 2 μM strand displacement primer. After heating for 5 min at 95°C and slowly cooling to room temperature, the displacement reaction was initiated by adding Bst 3.0 DNA polymerase to a final concentration of 80 U/mL. The reaction was incubated for 1 h at 50°C. Bst polymerase was extracted with phenol:chloroform as described above. The TNA library was separated from the residual phenol and reaction components in a YM-30 Centricon. Any remaining phenol was removed by adding 900 μL of ethanol to the solution and drying completely. The libraries were resuspended in selection buffer (S1: 150 mM NaCl, 25 mM Tris pH 8.0, TNFα: 20 mM imidazole, 450 mM NaCl, 20 mM Tris pH 8.0).

3.6.3 Aptamer selection

Aptamer selections were performed by passing each library through unmodified Ni-NTA agarose beads for 30 minutes at 24°C with rotation to remove any sequences that bound to the affinity matrix. This step was performed for each round of selection. For the S1 selection, an additional negative selection step was performed with 1 μM hemagglutinin (HA) for 30 minutes at 24°C with rotation, chosen as a generic viral coat protein. For this step, the flow through from the empty beads was passed through Ni-NTA beads containing the HA protein. The negative HA selection step was only performed in the first round of the S1 selection. The material collected from the flow-through of the negative selection was incubated with the appropriate His-tagged protein (either the S1 protein from SARS-CoV-2 or TNFα) poised at a concentration of 1 μM for 15 minutes at 24°C with rotation. The mixture was incubated in a disposable plastic column with Ni-NTA beads for 15 minutes at 24°C with rotation. The column was drained and washed three times with 400 μL of selection buffer to remove unbound and weakly bound molecules. A more stringent wash

step was performed with 1 M NaCl to reduce the occurrence of nonspecific electrostatic interactions (S1: 1 M NaCl, 25 mM Tris pH 8.0, TNFα: 20 mM imidazole, 1 M NaCl, 20 mM Tris pH 8.0), followed by an additional wash with selection buffer to return to baseline salt concentrations before elution. Six 250 μL elutions were performed with elution buffer (S1: 500 mM imidazole, 150 mM NaCl, 25 mM Tris pH 8.0, TNFα: 100 mM imidazole, 450 mM NaCl, 20 mM Tris pH 8.0). Elution fractions were imaged on a LI-COR Odyssey CLx. The two elution fractions with the highest fluorescence were pooled and desalted using a NAP-5 DNA purification column (Cytiva, Marlborough, MA). TNA aptamers were amplified by PCR with library-specific PCR primers and purified using DNA Clean and Concentrator columns (Zymo Research, Irvine, CA). The single-stranded DNA library was regenerated through a second PCR using a PEGylated forward primer and the high fidelity PCRBIO HiFi polymerase, then purifying by 10% denaturing PAGE, cutting the corresponding band with a scalpel and recovering by electroelution.

3.6.4 NGS & Analysis

86 bp library members that remained in the pool after each round of selection were PCR amplified and purified as above. Amplicons were prepared as barcoded libraries and sequenced on a NovaSeq6000 with the S4 flow cell with the paired end setting for 200 cycles for an estimated 2-400 million reads per library, per round.

Sequencing data was parsed by filtering, trimming and aligning reads. Enrichment scores over the course of the selection were calculated for the top 10,000 most abundant sequences in the final round.

3.6.5 Directed evolution of Trp-1

A doped library of Trp-1 was synthesized at the DNA level with 15% doping of the other 3-nts at each position in the aptamer sequence. The DNA displayed TNA library was generated and taken through one round of selection exactly as described above, excluding the HA negative selection, with the exception that the S1 protein was poised at a concentration of 0.5 μM. TNA aptamers collected in the elution fractions were amplified by PCR with library-specific PCR primers, purified using DNA Clean and Concentrator columns (Zymo Research, Irvine, CA), and submitted for NGS analysis.

3.6.6 Aptamer preparation for kinetic analysis

Aptamers identified by sequencing were prepared as biotin-labeled TNA molecules using the PBS8 biotin 20mer DNA primer and the corresponding template. A 500 μL reaction volume containing 0.5 nmol of both primer and template as well as 1x ThermoPol Buffer was heated for 5 min at 95°C to anneal. TNA polymerization was initiated by adding 100 μM of each tNTP and Kod-RSGA polymerase (1 μM for standard bases or Phe, 2 μM for Trp) and then incubated for 2 h at 55°C. Full length biotin-labeled aptamers were purified by 10% denaturing polyacrylamide gel electrophoresis for 1.5 h at 18 W constant. TNA was recovered from the gel by electroelution and then buffer exchanged into water and concentrated using a YM-3 Centricon centrifugal filter device. TNA concentration was quantified by NanoDrop absorbance. Aptamers at a concentration of 100 nM were folded in BLI binding buffer (125 mM NaCl, 20 mM HEPES pH 7.5, 2 mM CaCl2, 0.05% Tween 20) by denaturing for 15 min at 95°C and then cooling for 1 h at room temperature.

3.6.7 BLI screening

Aptamers were screened en masse to identify sequences of interest for further characterization. Prior to testing, all sensors were equilibrated in BLI binding buffer for ≥ 1 h. Aptamers at a concentration of 100 nM were loaded on a single biosensor. A fixed target concentration of 100 nM for S1 and 1000 nM for TNFα was used.

After aptamer folding, the BLI run was performed with the following steps: a buffer only baseline for 60 sec to equilibrate sensors, loading the aptamer for 200 sec, a second buffer only baseline for 200 sec, an association phase with the target protein for 600 sec for S1 or 120 sec for TNFα, and a dissociation phase for 600 sec for S1 or 120 sec for TNFα. The plate and reagents were incubated at 30°C for S1 samples or 24°C for TNFα for the duration of the experiment and for 10 min prior to each run. Data was analyzed using the Octet Data Analysis HT software. For screens, Savitzky-Golay filtering was applied before fitting both association and dissociation curves together to calculate approximate K_D values.

3.6.8 Full kinetic measurements for S1

Aptamer sequences of interest were characterized with four different concentrations of target and one buffer only sensor to determine background. Prior to testing, all sensors were equilibrated in BLI binding buffer (defined above) for ≥ 1 h. After aptamer folding, the BLI run was performed with the following steps: a buffer only baseline for 60 sec to equilibrate sensors, loading the aptamer for 200 sec, a second buffer only baseline for 200 sec, an association phase with the target protein for 600 sec for S1 or 120 sec for TNFα, and a dissociation phase for 600 sec for S1 or 120 sec for TNFα. The plate and reagents were incubated at 30°C for S1 or 24°C for TNFα for the duration of the experiment and for 10

min prior to each run. Data was analyzed using the Octet Data Analysis HT software. For full kinetics measurements, the buffer only baseline sample was used to subtract background from all samples before applying Savitzky-Golay filtering and fitting both the association and dissociation curves together and applying a global fit to determine K_D and other metrics.

3.7 References

Brundo, Y. & Liu, D. R. (2009) Recent progress toward the templated synthesis and directed evolution of sequence-defined synthetic polymers. Chem. Biol. 16, 265-276.

Chen, Z., Lichtor, P. A., Berliner, A. P., Chen, J. C. & Liu, D. R. Evolution of sequence-defined highly functionalized nucleic acid polymers. Nat. Chem., (2018).

Cheung, Y. W., Röthlisberger, P., Mechaly, A. E., Weber, P., Levi-Acobas, F., Lo, Y., Wong, A. W. C., Kinghorn, A. B., Haouz, A., Savage, G. P., Hollenstein, M., & Tanner, J. A. (2020) Evolution of abiotic cubane chemistries in a nucleic acid aptamer allows selective recognition of a malaria biomarker. Proc. Natl. Acad. Sci. U.S.A. 117, 16790-16798

Culbertson, M. C., Temburnikar, K. W., Sau, S. P., Liao, J. Y., Bala, S., & Chaput, J. C. (2016). Evaluating TNA stability under simulated physiological conditions. Bioorganic and Medicinal Chemistry Letters, 26(10).

Dunn, M. R., McCloskey, C. M., Buckley, P., Rhea, K. & Chaput, J. C. (2020) Generating biologically stable TNA aptamers that function with high affinity and thermal stability. J. Am. Chem. Soc. 142, 7721-7724.

Dunn, M. R., Jimenez, R. M. & Chaput, J. C. (2017) Analysis of aptamer discovery and technology. Nat. Rev. Chem. 1, 0076.

Gawande, B. N., Rohloff, J. C., Carter, J. D., von Carlowitz, I., Zhang, C., Schneider, D. J., & Janjic, N. Selection of DNA aptamers with two modified bases. (2017) Proc. Natl. Acad. Sci. U.S.A. 114, 2898-2903.

Gelinas, A. D., Davies, D. R. & Janjic, N. (2016) Embracing proteins: structural themes in aptamer-protein complexes. Curr. Opin. Struct. Biol. 36, 122-132.

Hottin, A. & Marx, A. (2016) Structural Insights into the Processing of Nucleobase-Modified Nucleotides by DNA Polymerases. Acc. Chem. Res. 49, 418-427.

Kalliolias, G. D. & Ivashkiv, L. B. (2016) TNF biology, pathogenic mechanisms and emerging therapeutic strategies. Nat Rev Rheumatol 12, 49-62.

Keefe, A. D., Pai, S. & Ellington, A. D. Aptamers as therapeutics. Nat. Rev. Drug Discov. 9, 537-550, (2010).

Kimoto, M., Yamashige, R., Matsunaga, K.-I., Yokoyama, S. & Hirao, I. (2013) Generation of high affinity DNA aptamers using an expanded genetic alphabet. Nat. Biotechnol. 31, 453-457.

Kong, D., Yeung, W. & Hili, R. (2017) In Vitro Selection of Diversely Functionalized Aptamers. J. Am. Chem. Soc. 139, 13977-13980.

Kringelum, J. V., Nielsen, M., Padkjaer, S. B. & Lund, O. (2013) Structural analysis of B-cell epitopes in antibody:protein complexes. Mol Immunol 53, 24-34.

Liao, J.-Y., Bala, S., Ngor, A. K., Yik, E. J. & Chaput, J. C. (2019) P(V) Reagents for the Scalable Synthesis of Natural and Modified Nucleoside Triphosphates. J. Am. Chem. Soc. 141, 13286-13289.

Mei, H. & Chaput, J. C. (2018) Expanding the chemical diversity of TNA with tUTP derivatives that are substrates for a TNA polymerase. Chem. Commun. 54, 1237-1240.

Nikoomanzar, A., Chim, N., Yik, E. J. & Chaput, J. C. (2020) Engineering polymerases for applications in synthetic biology. Q. Rev. Biophys. 53, e8.

Nikoomanzar, A., Vallejo, D., Yik, E. J. & Chaput, J. C. (2020) Programmed allelic mutagenesis of a DNA polymerase with single amino acid resolution. ACS Synth. Biol. 9, 1873-1881.

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

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. (2012) Synthetic genetic polymers capable of heredity and evolution. Science 336, 341-344.

Ramaraj, T., Angel, T., Dratz, E. A., Jesaitis, A. J. & Mumey, B. (2012) Antigen-antibody interface properties: composition, residue interactions, and features of 53 non-redundant structures. Biochim. Biophys. Acta 1824, 520-532.

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

Roberts, R. W. & Szostak, J. W. (1997) RNA-peptide fusions for the in vitro selection of peptides and proteins. Proc. Natl. Acad. Sci. U.S.A. 94, 12297-12302.

Sau, S. P., Fahmi, N. E., Liao, J.-Y., Bala, S. & Chaput, J. C. (2016) A scalable synthesis of α-L-threose nucleic acid monomers. J. Org. Chem. 81, 2302-2307.

Tolle, F., Brandle, G. M., Natzner, D. & Mayer, G. (2015) A versatile approach towards nucleobase-modified aptamers. Angew. Chem. Int. Ed. 54, 10971-10974.

Vater, A. & Klussmann, S. (2015) Turning mirror-image oligonucleotides into drugs: the evolution of Spiegelmer therapeutics. Drug Discovery Today 20, 147-155.

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

Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh C.-L., Abiona, O., Graham, B. S., & McLellan, J. S. (2020) Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 367, 1260-1263.

Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., & Zhou, Q. (2020) Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science 367, 1444-1448.

Yu, H., Zhang, S. & Chaput, J. C. (2012) Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. Nat. Chem. 4, 183-187.

Zhou, J. & Rossi, J. (2017) Aptamers as targeted therapeutics: current potential and challenges. Nat. Rev. Drug Discov. 16, 181-202.

CHAPTER 4

Publication Note

This article was originally published in ACS Synthetic Biology. Reprinted with permission from: McCloskey, C. M.; Liao, J. Y.; Bala, S.; and Chaput, J. C. Ligase-Mediated Threose Nucleic Acid Synthesis on DNA Templates. *ACS Synth. Biol.* 2019, 8, 2, 282–286. Copyright 2019 by American Chemical Society.

4.1 Contribution Statement

J.C., C.M, and J.Y.L. conceived of the project and designed the experiments. C.M., J.Y.L., and S.B. performed the experiments. C.M. and J.C. wrote the manuscript. All authors reviewed and commented on the manuscript.

4.2 Abstract of the Chapter

Ligases are a class of enzymes that catalyze the formation of phosphodiester bonds between an oligonucleotide donor with a 5′ terminal phosphate and an oligonucleotide acceptor with a 3′ terminal hydroxyl group. Here, we wished to explore the substrate specificity of naturally occurring DNA and RNA ligases to determine whether the molecular recognition of these enzymes is sufficiently general to synthesize alternative genetic polymers with backbone structures that are distinct from those found in nature. We chose threose nucleic acid (TNA) as a model system, as TNA is known to be biologically stable and capable of undergoing Darwinian evolution. Enzyme screening and reaction optimization identified several ligases that can recognize TNA as either the donor or acceptor strand with DNA. Less discrimination occurs on the acceptor strand indicating that the determinants of substrate specificity depend primarily on the composition of the donor

strand. Remarkably, T3 and T7 ligases were able to join TNA homopolymers together, which is surprising given that the TNA backbone is one atom shorter than that of DNA. In this reaction, the base composition of the ligation junction strongly favors the formation of A-T and A-G linkages. We suggest that these results will enable the assembly of TNA oligonucleotides of lengths beyond what is currently possible by solid-phase synthesis and provide a starting point for further optimization by directed evolution.

4.3 Introduction

Ligases are a class of naturally occurring enzymes that catalyze the formation of a phosphodiester bond between a terminal 5′ monophosphate and a 3′ hydroxyl group on adjacent oligonucleotides of a nicked duplex structure (Lehman, 1974). The reaction proceeds by way of a two-step mechanism involving a series of transesterification reactions (Cherepanov & de Vries, 2002; Crut et al., 2008). The first step entails the transfer of AMP to the 5′ phosphate of the donor strand from a covalent enzyme-AMP complex. The second step involves attack of the 3′ hydroxyl group on the AMP-activated phosphate to produce a covalent phosphodiester bond joining the donor and acceptor strands. These enzymes play a critical role in maintaining genome stability in cells and have found broad utility in practical applications that require DNA synthesis, including DNA sequencing and vector assembly (Kosuri & Church, 2014). The reaction has been studied in detail and is known to synthesize DNA and RNA oligonucleotides in high yield (Enland & Uhlenbeck, 1978).

Here we sought to evaluate the substrate specificity of naturally occurring DNA and RNA ligases to determine whether the molecular recognition of these enzymes is sufficiently general to allow for the synthesis of alternative genetic polymers (*i.e.*, XNAs)

with backbone structures that are distinct from those found in nature (Anosova et al, 2016; Chaput et al, 2012). Previous studies have documented the tolerance of ligases toward gaps and base-pair mismatches in the substrate as well as blunt-end cohesion (Goffin et al., 1987; Nilsson & Magnusson, 1982; Wu & Wallace, 1989). Other studies have examined substrates bearing chemical modifications on the base and sugar moieties of nucleotides at or near the ligation site (Hili et al., 2013; Kestemont et al., 2018). However, in all cases investigated thus far, the substrates maintain the same six-atom backbone repeat unit found in DNA and RNA. By comparison, threose nucleic acid (TNA, Figure 4.1), an artificial genetic polymer in which the natural ribose sugar found in RNA has been replaced with an unnatural threose sugar, is a more radical departure from the chemical space of DNA and RNA (Schöning et al., 2000). Despite a backbone repeat unit that is one atom shorter than that of natural genetic polymers, TNA is capable of forming stable antiparallel Watson–Crick duplex structures with itself and with DNA and RNA (Wu et al., 2002; Yang et al., 2007). Due to its chemical simplicity and high biological stability, TNA has received significant attention as a possible RNA progenitor in the evolution of life and as a new analogue for RNA therapeutics (Joyce, 2002; Orgel, 2000).

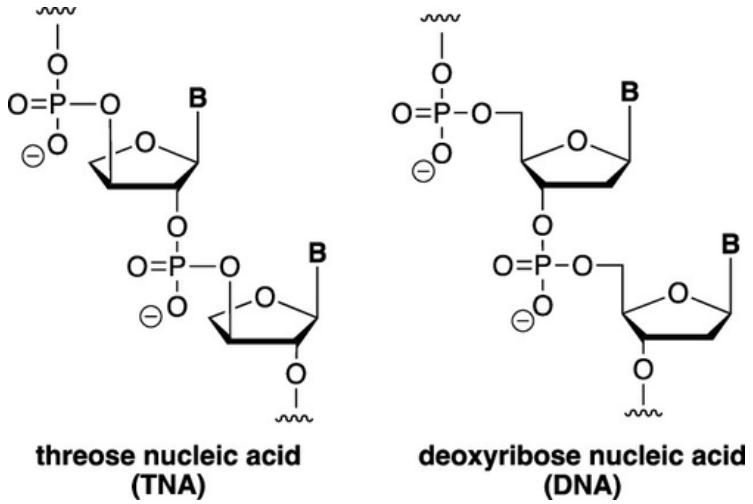


Figure 4.1. Molecular structure of TNA and DNA. Constitutional structure for the linearized backbone of α-l-threofuranosyl-(3′,2’) nucleic acid (left) and DNA (right). TNA is an unnatural genetic polymer composed of repeating α-l-threofuranosyl units that are vicinally connected by 2′,3′-phosphodiester bonds.

4.4 Results and Discussion

We used a nick joining assay to compare the efficiency of TNA-DNA, DNA-TNA, and TNA-TNA ligation on a complementary DNA template (Figure 4.2). The TNA oligonucleotides were obtained by solid-phase synthesis using chemically synthesized TNA phosphoramidites (Sau et al., 2016). In this assay, stoichiometric concentrations of the donor, acceptor, and template strands are combined and annealed by heat denaturing and slow cooling in buffer. The reaction is then initiated with the addition of ATP and enzyme. Following the incubation period (typically 24 h), product formation was assessed by denaturing polyacrylamide gel electrophoresis (PAGE) using an IR-labeled acceptor strand.

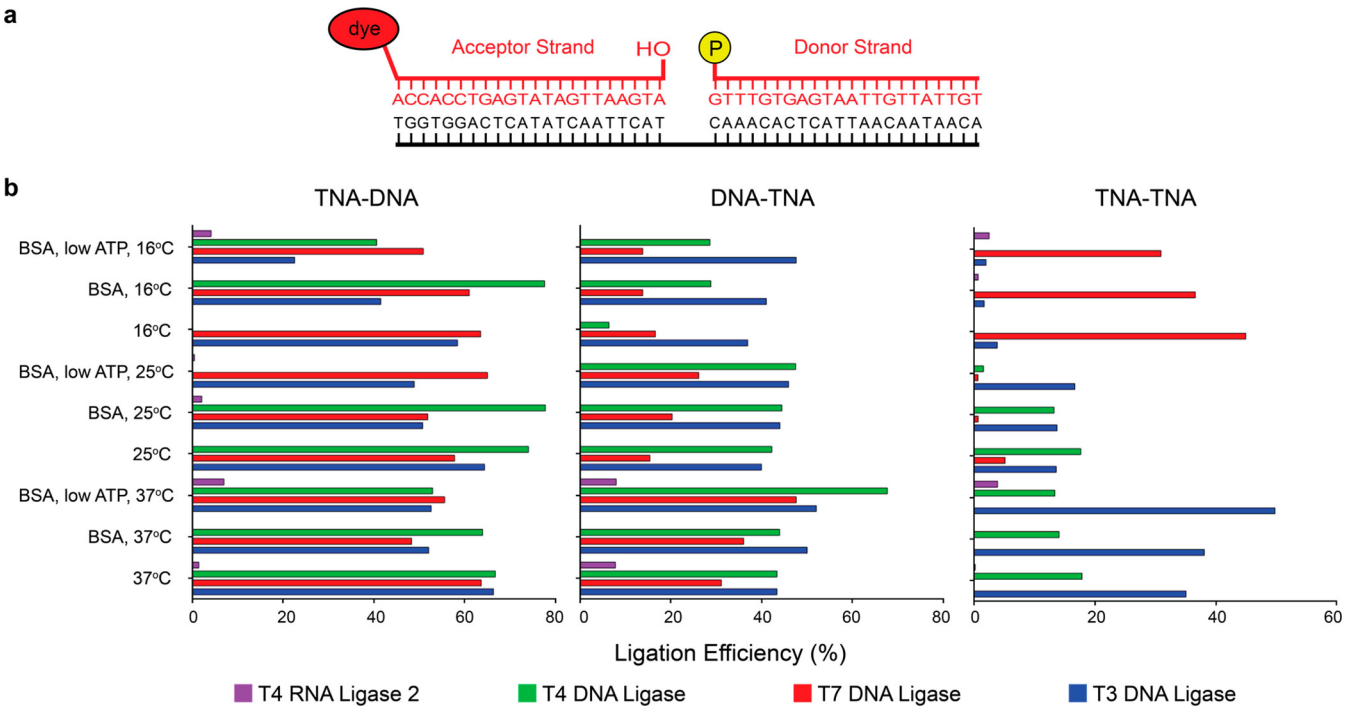


Figure 4.2. Initial screen of ligases and conditions. (a) Sequence and schematic illustration of the nicked duplex substrate. (b) Ligation efficiency determined under different reaction conditions and temperatures. Ligation efficiency was measured by denaturing polyacrylamide gel electrophoresis after 24 h of incubation.

We screened several different types of ligases including bacteriophage T3, T4, and T7 DNA ligases as well as a T4 RNA ligase. The screen was performed across a range of temperatures (16, 25, and 37 °C) and reaction conditions that included standard (1 mM) and low (25 μM) ATP concentrations and 0.1 mg/mL BSA as a crowding reagent (Hili et al., 2013). In all cases, T4 RNA ligase consistently showed little or no activity for the TNA substrates. This result was surprising given that TNA, like RNA, has a strong preference for A-form helical structures (Anosova et al., 2016; Ebert, M.-O. et al., 2008). This may be due to the ability for RNA ligases to discriminate between 3′- and 2′-hydroxyls during nick sealing.

By contrast, all DNA ligases tested in our assay showed at least some activity for the TNA substrates. The level of activity depended on the location of the TNA strand, with the highest levels observed when TNA was present as the acceptor strand with a downstream DNA donor strand. This reaction produced a chimeric TNA-DNA oligonucleotide with a 2′,5′-phosphodiester bond at the ligation junction. Remarkably, conditions were found where 60–80% product formation could be obtained for all three enzymes, suggesting that DNA ligases have less discrimination against the 2′ hydroxyl position on the acceptor strand.

However, the ligation efficiency was noticeably lower when TNA was present as the downstream strand, which produced a chimeric DNA-TNA oligonucleotide with a 3′,3′-phosphodiester bond at the ligation junction. Although one condition was identified where T4 DNA ligase could produce the chimeric product in 70% yield, most of the reaction conditions were substantially less efficient (~10–40% yields), indicating that the determinants of substrate specificity depend on the composition and presumably sequence of the donor strand.

The trend toward reduced activity continued for the all-TNA system, where TNA oligonucleotides were present in both the donor and acceptor strands. In this system, two conditions emerged as promising approaches for TNA synthesis. The best conditions required T3 DNA ligase at 37 °C with BSA and 25 μM ATP or T7 DNA ligase at 16 °C with no added BSA and 1 mM ATP. Under these conditions, TNA strands with 2′,3′-phosphodiester bonds at the ligation junction could be generated in 45–50% yield. This result was surprising given the structural differences between TNA and DNA (Schöning et al., 2000).

A summary of the best results for each nicked duplex structure is shown in Figure 4.3. The global trend of TNA-DNA > DNA-TNA > TNA-TNA is consistent with previous studies documenting the increased intermolecular interaction between the ligase and the downstream sequence (Cherepanov & de Vries, 2002). However, these studies were performed with natural DNA and RNA substrates, which form optimal contacts with the enzyme active site.

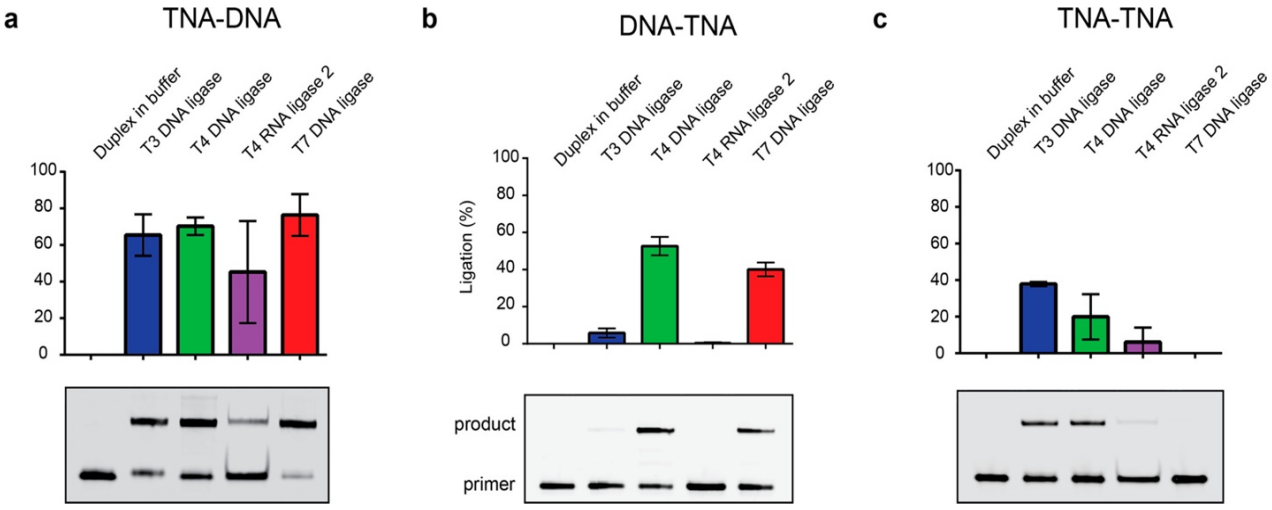


Figure 4.3. Enzymatic ligation under optimized reaction conditions. The ligation efficiency of (a) TNA-DNA, (b) DNA-TNA, and (c) TNA-TNA substrates was measured by denaturing polyacrylamide gel electrophoresis after 24 h of incubation. Reactions were performed under the optimal conditions determined for each enzyme–substrate pair.

To determine the sequence requirements of ligation, we screened all of the 16 possible sequence combinations that define ligation junctions at the terminal nucleotide positions (Figure 4.4a,b, SI Figure 4.1). Using T3 DNA ligase, we compared DNA and TNA synthesis using donor and acceptor strands that generate the all-DNA and all-TNA products. As expected, DNA synthesis proceeds efficiently for all 16 nucleotide junctions,

giving an average yield of 70% product formation. However, TNA synthesis was found to be highly sequence-dependent with a clear-cut bias for A-T and A-G junctions (~50–60% yield). With the exception of G-G and G-T junctions (~25% yield), the remaining sequence combinations showed little or no activity (Figure 4.4b). This result was significant, as ability to synthesize TNA oligonucleotides beyond what is currently possible with conventional solid-phase synthesis (~30-mers) provides a new tool for the synthetic biology arsenal.

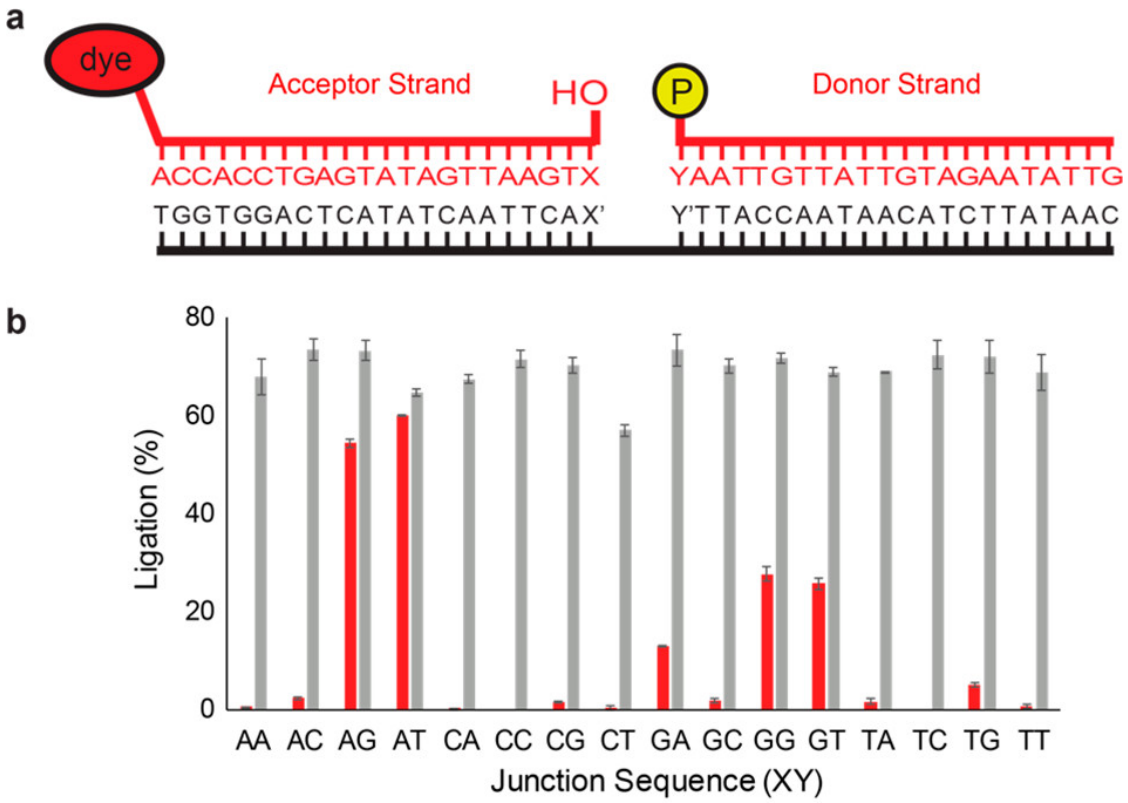


Figure 4.4. TNA and DNA synthesis by T3 DNA ligase. (a) Structure and sequence of the nicked duplex structure. The letters X and Y denote the upstream and downstream terminal nucleotides at the ligation junction. (b) Ligation efficiency profile for the all-DNA (gray) and all-TNA (red) strands.

To investigate the rate of TNA synthesis, we performed a time course analysis using T3 DNA ligase and the TNA strands carrying the A and T terminal residues designed to form the A-T junction. As shown in Figure 4.5, the reaction is complete in ~4 h with a maximum product conversion of 60% full-length TNA. The reaction was scaled-up, and product formation was validated by mass spectrometry (SI Figure 4.2), which produced a mass similar to the calculated theoretical mass (observed: 13 135, calculated: 13 138 Da).

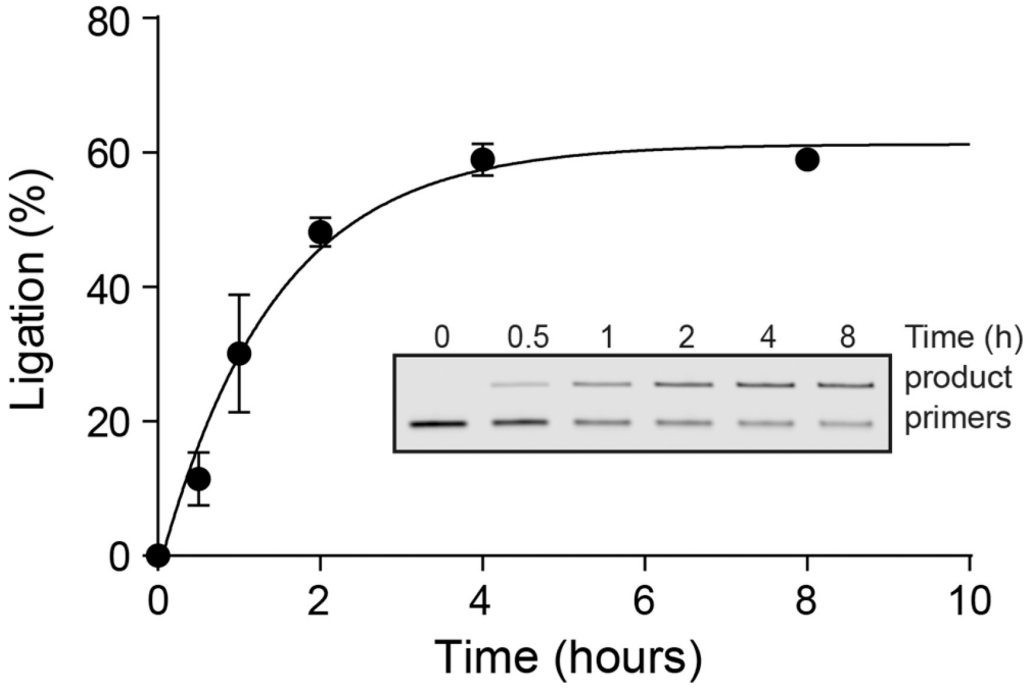


Figure 4.5. TNA-TNA ligation by T3 DNA ligase. Time course analysis of TNA synthesis using the A-T junction. Polyacrylamide gel electrophoresis data provided as an inset. Experiments were performed in duplicate.

Despite a narrow sequence preference, the current results represent a significant improvement over previous nonenzymatic ligation studies by Eschenmoser and colleagues that required a 3′-amino group on the donor strand, which achieved 37% product conversion (Wu et al. 2002). Significantly, the ability to generate longer TNA polymers by

enzymatic ligation avoids the need to chemically synthesize amino-modified TNA monomers, which is a costly and labor-intensive process.

In summary, we have identified bacteriophage T7, T4, and T3 DNA ligases as enzymes capable of synthesizing TNA-DNA, DNA-TNA, and TNA-TNA oligonucleotide products by template-directed ligation, respectively. We suggest that directed evolution of these enzymes, especially T3 DNA ligase, which is responsible for TNA-TNA synthesis, could lead to new enzyme variants with higher efficiency and broader substrate specificity. Such enzymes would provide an important new tool for synthetic genetics where researchers are studying artificial genetic polymers inspired by nature.

4.5 Experimental Details

4.5.1 General Information

Ligases were purchased from New England Biolabs (NEB, Ipswich, MA). DNA oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA), purified by denaturing polyacrylamide gel electrophoresis, electroeluted, desalted using Amicon Ultra 0.5 mL centrifugal filters from Millipore Sigma (Burlington, MA), and UV quantified using a NanoDrop spectrophotometer from ThermoFisher (Waltham, MA).

4.5.2 TNA Oligonucleotide Synthesis and Preparation

TNA phosphoramidites were synthesized as reported previously (Sau et al., 2016). Standard β-cyanoethyl phosphoramidite chemistry and an Applied Biosystems 3400 DNA Synthesizer were used to synthesize TNA oligonucleotides on Universal Support II CPG columns (1 μM scale, Glen Research, Sterling, VA). Standard DNA coupling procedures were modified such that coupling time for TNA amidites is increased to 2000 s and detritylation is performed in two cycles, 60 s each. Upstream oligonucleotides were coupled with 5'-

Hexynyl Phosphoramidite (Glen Research) for later tagging via click chemistry. Downstream oligonucleotides were 3′ phosphorylated with Chemical Phosphorylation Reagent II (Glen Research). Cleavage from the solid support was achieved in NH_4OH (33%) for 18 h at 55 °C. Oligonucleotides were purified and deprotected using a Poly-Pak II Cartridge (Glen Research).

To dye-label alkynylated oligonucleotides, 50 nmol of IR800 azide and freshly prepared solution containing final concentrations 22 mM CuBr and 44 mM TBTA ligand in 3:1 DMSO/*t*-BuOH were added to 25 nmol of TNA. The mixture was held at 37 °C overnight. The reaction was subsequently diluted with 0.3 M NaOAc, and the TNA precipitated using 1 mL cold EtOH. The precipitate was resuspended in 95% formamide stop buffer and purified by denaturing PAGE.

4.5.3 Broad Screen for TNA Ligase Activity

1 μM of dye-labeled acceptor oligonucleotide (DNA P1 or TNA P1), 1 μM of phosphorylated donor oligonucleotide (DNA P2 or TNA P2), 1 μM of DNA splint were added to a total of 10 μL of reaction volume with 1× of the buffers provided by NEB for each ligase. ATP was tested at two different concentrations, 1 mM and 25 μM. For samples containing BSA, BSA (Calbiochem, La Jolla, CA) was added to a final concentration of 0.1 mg/mL. Reactions were initiated with the addition of 50 U of ligase and allowed to incubate for 24 h at 16, 25, or 37°C. The reaction was then diluted (1:4, v/v) in 95% formamide stop buffer and analyzed by denaturing PAGE.

4.5.4 Optimized Condition for TNA-TNA Ligation

1 μM of dye-labeled acceptor oligonucleotide, 1 μM of phosphorylated donor oligonucleotide, and 1 μM of complementary template were added to 1× of T3/T7 buffer.

The mixture was heated to 95 °C for 10 s, cooled to 65 °C for 4 min, then ramped down to 4 °C at 100 s/°C. After annealing, 25 μM ATP, 0.1 mg/mL BSA, and 37.5 U/μL T3 DNA ligase were added to the reaction mixture on ice for a total volume of 10 μL. Reactions were incubated for 24 h at 37 °C. The reaction was then diluted (1:4, v/v) in 95% formamide stop buffer and analyzed by denaturing PAGE.

4.5.5 Ligation Junction Assay

TNA oligonucleotides were constructed by solid-phase synthesis (see above) that contained all the possible sequence combinations to evaluate the terminal nucleotide position on the donor and acceptor strands for TNA-TNA ligation. The ligation reaction was performed using the optimized conditions for TNA-TNA ligation (see above). A parallel set of experiments were performed as controls using DNA oligonucleotides of the same sequence. The DNA ligation reactions were performed in a manner identical to the TNA ligation reactions.

4.5.6 Rate Determination Assay

The rate of TNA synthesis by T3 DNA ligase was determined for the all-TNA (A-T junction) using 1 μM of TNA P1A, 1 μM of TNA P3T, and 1 μM of complementary template in a solution containing 1× of T3/T7 buffer. The mixture was heated to 95 °C for 10 s, cooled to 65 °C for 4 min, then ramped down to 4 °C at 100 s/°C. After annealing, 25 μM ATP, 0.1 mg/mL BSA, and 37.5 U/μL T3 DNA ligase were added to the reaction mixture on ice for a total volume of 60 μL. The reaction was split into 6 tubes, 10 μL each, and incubated at 37 °C for 8 h, during which time points were taken by removing a single tube and immediately stopping the reaction by dilution (1:4, v/v) in 95% formamide stop buffer and analyzed by denaturing PAGE.

4.5.7 Mass Spectrometry Validation

1 μM of TNA P1A, 1 μM of TNA P3T, and 1 μM of complementary template (DNA splint 2) were mixed in a solution containing 1× of T3/T7 buffer. The mixture was heated to 95 °C for 10 s, cooled to 65 °C for 4 min, then ramped down to 4 °C at 100 s/°C. After annealing, 25 μM ATP, 0.1 mg/mL BSA, and 37.5 U/μL T3 DNA ligase were added to the reaction mixture on ice for a total volume of 300 μL. The reaction was incubated at 37 °C for 18 h, then stopped by dilution (1:4, v/v) in 95% formamide stop buffer and analyzed by denaturing PAGE. The band corresponding to product was cut and the product electroeluted and exchanged into water using an Amicon Ultra centrifugal filter. The sample was UV quantified by Nanodrop. For external validation by mass spectrometry, the sample was sent to Novatia (Newtown, PA) for high-resolution electrospray ionization (ESI).

4.6 References

Anosova, I., Kowal, E. A., Dunn, M. R., Chaput, J. C., Van Horn, W. D., and Egli, M. (2016) The Structural Diversity of Artificial Genetic Polymers. Nucleic Acids Res. 44, 1007– 1021.

Anosova, I., Kowal, E. A., Sisco, N. J., Sau, S. P., Liao, J.-Y., Bala, S., Rozners, E., Egli, M., Chaput, J. C., and Van Horn, W. D. (2016) Structural Insights into Conformational Differences between DNA/TNA and RNA/TNA Chimeric Duplexes. ChemBioChem 17, 1705– 1708.

Chaput, J. C., Yu, H., and Zhang, S. (2012) The Emerging World of Synthetic Genetics. Chem. Biol. 19, 1360– 1371.

Cherepanov, A. V. and de Vries, S. (2002) Dynamic mechanism of nick recognition by DNA ligase. Eur. J. Biochem. 269, 5993– 9.

Crut, A., Nair, P. A., Koster, D. A., Shuman, S., and Dekker, N. H. (2008) Dynamics of phosphodiester synthesis by DNA ligase. Proc. Natl. Acad. Sci. U. S. A. 105, 6894– 9.

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

Ebert, M.-O., Mang, C., Krishnamurthy, R., Eschenmoser, A., and Jaun, B. (2008) The Structure of a TNA–TNA Complex in Solution: NMR Study of the Octamer Duplex Derived from α-(l)-Threofuranosyl-(3′-2′)-CGAATTCG. J. Am. Chem. Soc. 130, 15105–15115.

Enland, T. E. and Uhlenbeck, O. C. (1978) Enzymatic Oligoribonucleotide Synthesis with T4 RNA Ligase. Biochemistry 17, 2069– 76.

Goffin, C., Bailly, V., and Verly, W. G. (1987) Nicks 3′ or 5′ to AP sites or to mispaired bases, and one-nucleotide gaps can be sealed by T4 DNA ligase. Nucleic Acids Res. 15, 8755–71.

Hili, R., Niu, J., and Liu, D. R. (2013) DNA Ligase-Mediated Translation of DNA Into Densely Functionalized Nucleic Acid Polymers. J. Am. Chem. Soc. 135, 98– 101.

Joyce, G. F. (2002) The antiquity of RNA-based evolution. Nature 418, 214– 221.

Kestemont, D., Renders, M., Leonczak, P., Abramov, M., Schepers, G., Pinheiro, V. B., Rozenski, J., and Herdewijn, P. (2018) XNA ligation using T4 DNA ligase in crowding conditions. Chem. Commun. 54, 6408– 6411.

Kosuri, S. and Church, G. M. (2014) Large-scale de novo DNA synthesis: technologies and applications. Nat. Methods 11, 499– 507.

Lehman, I. R. (1974) DNA Ligase: Structure, Mechanism, and Function. Science 186, 790–797.

Liu, L. S., Leung, H. M., Tam, D. Y., Lo, T. W., Wong, S. W., and Lo, P. K. (2018) α-l-Threose Nucleic Acids as Biocompatible Antisense Oligonucleotides for Suppressing Gene Expression in Living Cells. ACS Appl. Mater. Interfaces 10, 9736– 9743.

Nilsson, S. V. and Magnusson, G. (1982) Sealing of gaps in duplex DNA by T4 DNA ligase. Nucleic Acids Res. 10, 1425– 37.

Orgel, L. E. (2000) Origin of life. A simpler nucleic acid. Science 290, 1306– 1307.

Sau, S. P., Fahmi, N. E., Liao, J.-Y., Bala, S., and Chaput, J. C. (2016) A Scalable Synthesis of α-L-Threose Nucleic Acid Monomers. J. Org. Chem. 81, 2302– 2307.

Schöning, K. U., Scholz, P., Guntha, S., Wu, X., Krishnamurthy, R., and Eschenmoser, A. (2000) Chemical etiology of nucleic acid structure: the alpha-threofuranosyl-(3′-->2′) oligonucleotide system. Science 290, 1347– 1351.

Wu, D. Y. and Wallace, R. B. (1989) Specificity of the nick-closing activity of bacteriophage T4 DNA ligase. Gene 76, 245– 54.

Wu, X., Delgado, G., Krishnamurthy, R., and Eschenmoser, A. (2002) 2,6-diaminopurine in TNA: effect on duplex stabilities and on the efficiency of template-controlled ligations. Org. Lett. 4, 1283– 1286.

Yang, Y.-W., Zhang, S., McCullum, E. O., and Chaput, J. C. (2007) Experimental Evidence that GNA and TNA Were Not Sequential Polymers in the Prebiotic Evolution of RNA. J. Mol. Evol. 65, 289– 295.

APPENDIX A

SUPPLEMENTARY TABLES

SI Table 2.1 Oligonucleotide sequences used for selection and characterization of HIV-RT 3.17. Modifications are written in IDT nomenclature. Primers are indicated in lowercase, blue letters.

Name	Sequence (5' to 3')
Library	ggatcgtcagtgcattgaga – N_{40} – ggtggtatccccaaggggac
Hairpin	/5Phos/ACTACGTACCCACAACCTCGGCCGTACCACGGTACGTAGTgtccccttggggataccacc
PBS8	gtccccttggggataccacc
PBS7	ggatcgtcagtgcattgaga
PBS7 PEG long	$(AACA)_{10}$ /iSp9/ ggatcgtcagtgcattgaga
PBS7 IR680 10mer	/5IR700/ tgcattgaga
PBS8 IR680 10mer	/5IRD700/ ggataccacc
PBS8 IR800 10mer	/5IRD800/ ggataccacc
PBS8 Biotin 10mer	/5BIOSG/ ggataccacc
Displacement Primer	ggtggtatccccaaggggacACTACGTACCGTGGTACGGCCGAGGTTGTG
30mer template	TCTCTATAGTGAGTCGTATAggtggtatcc
H3.1 template	TATGTAATGTTGACGTAATTTATCTCTGTGAGAATAAGCAggtggtatcc
H3.3 template	AGCATGTCAGTACTAGTAGGATATAAATCATGTGGTATTTggtggtatcc
H3.6 template	CATGGTGTCAATACTATACCTTTTGCCGATGGTAACAGTTggtggtatcc
H3.8 template	TATTGCAAATCCAAAATAAATAGAAAAAAAAAAACTAATAAggtggtatcc
H3.9 template	AGACATAGTACATAACCGATCAATGAACATTTAATTAAATggtggtatcc
H3.10 template	TGACAGTACGTACTGAACTAATTTATAAGTAATTTGCTATggtggtatcc
H3.12 template	GATGAATATCGCAATGAAATTATATATGTATAATTGGAACggtggtatcc
H3.15 template	TAGTCTAGGAACGGTACGATCGTATTCCAAGGACTTTGCTggtggtatcc
H3.16 template	CAATCTGCTAAAACTTAAAGCCTATTAATGAAATTGTTAAggtggtatcc
H3.17 template	TTATGTAGCATTTATGAAATTTTTAAATCAATTTACTATTggtggtatcc
Primer swap template	TTATGTAGCATTTATGAAATTTTTAAATCAATTTACTATTtctcaatgca
Wobble fix template	TTATGTAGCATTTATGAAATTTTTAAATCAATTTGCTATTggtggtatcc
H3.17 5d2 template	TTATGTAGCATTTATGAAATTTTTAAATCAATTTACTAggtggtatcc
H3.17 3d5 template	TAGCATTTATGAAATTTTTAAATCAATTTACTATTggtggtatcc
H3.17 5d2 3d5 template	TAGCATTTATGAAATTTTTAAATCAATTTACTAggtggtatcc
H3.17 dL template	TTATGTAGCATTTATGAAATAATCAATTTACTATTggtggtatcc
H3.17 5d2 dL template	TTATGTAGCATTTATGAAATAATCAATTTACTAggtggtatcc
H3.17 3d5 dL template	TAGCATTTATGAAATAATCAATTTACTATTggtggtatcc
H3.17 5d2 3d5 dL template	TAGCATTTATGAAATAATCAATTTACTAggtggtatcc
H3.17 dS template	ATTTATGAAATTTTTAAATCAATTTggtggtatcc
H3.16 dS dL template	ATTTATGAAATAATCAATTTggtggtatcc
H3.17 split P1 template	cctatggtggTTTAACTAAATagagttacgt
H3.17 split P2 template	acgtaactctTTTAAAGTATTTAccaccatagg
H3.17 split dL P1	cctatggtggTTTAACTAAagagttacgt
H3.17 split dL P2	acgtaactctTAAAGTATTTAccaccatagg
IR R1T DNA aptamer	/5IR700/ CGCCTGATTAGCGATACTCAGGCGTTGGGTGGGTGGGTGGG
Bio R1T DNA aptamer	/5BIOSG/ CGCCTGATTAGCGATACTCAGGCGTTGGGTGGGTGGGTGGG
IR PF1 DNA aptamer	/5IR700/ AGGAAGGCTTTAGGTCTGAGATCTCGGAAT
Bio PF1 DNA aptamer	/5BIOSG/ AGGAAGGCTTTAGGTCTGAGATCTCGGAAT

SI Table 2.2 TNA sequences identified by Sanger Sequencing, written in the 3' to 2' strand direction.

SI Table 2.3 Binding constants for nine chosen TNA aptamers. MST derived K_D values obtained from a single screen of 10 concentrations.

Entry	Aptamer	K_D (nM)
1	HIV-RT 3.1	34.5
2	HIV-RT 3.5	37.2
3	HIV-RT 3.6	6.4
4	HIV-RT 3.9	7.6
5	HIV-RT 3.10	4.6
6	HIV-RT 3.12	69.9
7	HIV-RT 3.15	2.0
8	HIV-RT 3.16	5.6
9	HIV-RT 3.17	0.41

SI Table 2.4 Thermal stability affinity measurements using BLI.

Time at 75°C	Antibody	DNA aptamer	TNA aptamer
0 h	7.20E-07 M	4.75E-11 M	1.00E-12 M
1 h	0	7.93E-11 M	1.00E-12 M
8 h	0	7.47E-11 M	1.00E-12 M
24 h	0	7.13E-11 M	1.00E-12 M
48 h	0	3.62E-10 M	1.00E-12 M

SI Table 3.1 Oligonucleotide sequences used for selection and characterization of threomers. Modifications are written in IDT nomenclature. N represents a 30 A:20 C:20 G:30 T combination. N' represents a combination of 85 parent nucleobase:15 other three nucleobases (5 each).

Name	Sequence (5'→ 3')
PBS7	GGATCGTCAGTGCATTGAGA
PEG PBS7	AACAAACAAACAAACAAACAAACAAACAAACA/iSp9/GGATCGTCAGTGCATTGAGA
PBS8	GTCCCCTTGGGGATACCACC
PBS12	TTCTCATGTTTGACAGCTTA
PEG PBS12	AACAAACAAACAAACAAACAAACAAACAAACA/iSp9/TTCTCATGTTTGACAGCTTA
PBS13	TCATCGATAAGCTTTAATGC
PEG PBS13	AACAAACAAACAAACAAACAAACAAACAAACA/iSp9/TCATCGATAAGCTTTAATGC
PBS15	TTGCTAACGCAGTCAGGCAC
PEG PBS15	AACAAACAAACAAACAAACAAACAAACAAACA/iSp9/TTGCTAACGCAGTCAGGCAC
DNA Display Hairpin	/5Phos/ACTACGTACCCACAACCTCGGCCGTACCACGGTACGTAGTGTCCCCTTGGGGATACCACC
Displacement Primer	/5IRD700/GGTGGTATCCCCAAGGGGACACTACGTACCGTGGTACGGCCGAGGTTGTG
3T Incorporation Template	TCTCTATTGTGAGTCGTCTAGGTGGTATCCTTTTTTTTTTTTTTTTTTT
S1 Std Library	TTCTCATGTTTGACAGCTTAATTCCT-N_{40}-GGTGGTATCCCCAAGGGGAC
S1 Phe Library	TCATCGATAAGCTTTAATGCCACTCA-N_{40}-GGTGGTATCCCCAAGGGGAC
S1 Trp Library	TTGCTAACGCAGTCAGGCACCCAACA-N_{40}-GGTGGTATCCCCAAGGGGAC
TNFα Std Library	GGATCGTCAGTGCATTGAGAATTCCT-N_{40}-GGTGGTATCCCCAAGGGGAC
TNFα Phe Library	GGATCGTCAGTGCATTGAGACACTCA-N_{40}-GGTGGTATCCCCAAGGGGAC
TNFα Trp Library	GGATCGTCAGTGCATTGAGACCAACA-N_{40}-GGTGGTATCCCCAAGGGGAC
Doped Trp-1 Library	TTGCTAACGCAGTCAGGCACTGGCCT-N'_{40}-GGTGGTATCCCCAAGGGGAC

SI Table 3.2 Sequencing reads obtained by NovaSeq.

Protein	Base	Round	Processed Reads	Unique Sequences
S1	Standard	1	26,286,663	20,412,876
		2	102,327,271	61,923,742
		3	71,836,059	42,735,631
		4	67,706,251	35,428,544
		Total	268,156,244	160,500,793
	Phe	1	23,583,405	20,059,840
		2	73,162,007	11,267,290
		3	73,322,141	13,938,779
		4	40,614,749	4,959,383
		Total	210,682,302	50,225,292
	Trp	1	30,106,901	23,793,731
		2	94,452,041	22,026,578
		3	68,928,155	29,680,238
		4	71,299,717	17,979,790
		Total	264,786,814	93,480,337
TNFα	Standard	1	518,460,356	266,703,781
		2	367,601,928	189,040,364
		3	591,683,108	188,726,113
		Total	1,477,745,392	644,470,258
	Phe	1	505,151,129	63,166,002
		2	483,294,442	84,340,855
		3	69,009,793	21,768,993
		Total	1,057,455,364	169,275,850
	Trp	1	369,586,308	98,436,052
		2	129,108,753	44,024,662
		3	217,754,947	59,876,218
		Total	716,450,008	202,336,932

SI Table 3.3 S1 aptamers chosen for full kinetics. Color scheme: standard library (green), Phe library (red), and Trp library (blue).

* Signifies an independent replicate.

Entry	TNA Sequence (3' to 2')	K_D (nM)	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	Fit (R^2)	Variance (Chi^2)
1	TGGCGAGGGGTAGTGAGCTGACGGATAGGGCTAGTGGAAC	17.6	1.6×10^4	2.7×10^{-4}	0.99	0.20
2	CGTGCACGATCTAGCCAGTGCTCGCTTGAAATGGGCGCAG	27.3	9.0×10^3	2.5×10^{-4}	0.99	0.16
3	TATTGGAACTCCCGAGCAGGCTGTCCCATGGAAGACTACA	2.6	9.0×10^4	2.4×10^{-4}	0.99	0.43
4	AGGCTTATGAAATCTTATGTGAACCGGGCATACGGGCGAT	5.2	1.4×10^4	7.5×10^{-5}	0.99	2.69
5*	AGGCTTATGAAATCTTATGTGAACCGGGCATACGGGCGAT	10.0	1.8×10^4	1.8×10^{-4}	0.99	0.41
6	GAGAGTCATAGTAGTTCACGGGTCTGCAGCTTACGCATGT	11.3	1.5×10^4	1.7×10^{-4}	0.99	0.70
7*	GAGAGTCATAGTAGTTCACGGGTCTGCAGCTTACGCATGT	12.0	2.5×10^4	3.0×10^{-4}	0.99	0.61
8*	TATTGGAACTCCCGAGCAGGCTGTCCCATGGAAGACTACA	16.3	1.2×10^4	2.0×10^{-4}	0.99	0.22
9	TGTACGCCATATAGGCAGTCATCGAGCAACGAAGTAATGG	16.7	1.2×10^4	1.9×10^{-4}	0.99	0.29
10*	AGGCTTATGAAATCTTATGTGAACCGGGCATACGGGCGAT	21.8	1.2×10^4	2.7×10^{-4}	0.99	4.30
11	ACGATGGACTTGGGTAGCTAGTCAGGTTAGGCTGTTTGGA	22.1	9.5×10^3	2.1×10^{-4}	0.99	0.73
12	AGAACTTCGTCAGCGTCTTACCGAGTCCAAGGACTGGCTA	28.8	9.1×10^3	2.6×10^{-4}	0.99	0.71
13	ATCGAGTTGTGACGTTAACGATAATAGCCCGCTCCGATTC	31.0	2.0×10^4	6.2×10^{-4}	0.99	0.14
14	TACATTACGCGTTAGCGTTAATGGGTATGTCCGGACAGTG	36.0	9.4×10^3	3.3×10^{-4}	0.99	26.65
15	GGTGTAATAAGGAGAGTACGTCGTTTGGTGGGACCGGGTA	39.0	4.3×10^3	1.7×10^{-4}	0.99	40.40
16*	TACATTACGCGTTAGCGTTAATGGGTATGTCCGGACAGTG	42.0	5.2×10^3	2.2×10^{-4}	0.99	3.30
17	TTGGCGAGATCCCCTATACGTGAGCTTCCTTTTATCCTTT	2.0	6.7×10^4	1.3×10^{-4}	0.99	2.94
18	CGTACAACTCCGGTTTCGGGCGGCCCGTTCTTCAGTTTGA	2.7	5.2×10^4	1.4×10^{-4}	0.99	0.52
19	TAAGGAAAGGGGAGCGGGTAGGAGGTTCGTCGGGTGCGAC	4.2	3.3×10^4	1.4×10^{-4}	0.99	0.73
20	ATTGTGCGAAGTATACATAGCAGATGCGTGACATAACAAG	5.1	2.1×10^4	1.1×10^{-4}	0.99	0.12
21*	CGTACAACTCCGGTTTCGGGCGGCCCGTTCTTCAGTTTGA	6.9	9.7×10^4	6.7×10^{-4}	0.99	2.02
22	AAACTGTTGGGACTTGTAATGTCGTCATCAGCTAAGAGTC	21.1	8.7×10^4	1.8×10^{-3}	0.72	16.72
23	TTTCTATGCATCTCGGGGATGCGGGATATAAAAGAGAGAC	29.1	3.9×10^4	1.1×10^{-3}	0.97	1.85
24*	TTTCTATGCATCTCGGGGATGCGGGATATAAAAGAGAGAC	76.0	7.0×10^3	5.4×10^{-4}	0.98	30.24

SI Table 3.4 TNFa aptamers chosen for full kinetics. Color scheme: standard library (green), Phe library (red), and Trp library (blue).

Entry	TNA Sequence (3' to 2')	K_D (nM)	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	Fit (R^2)	Variance (Chi^2)
1	GCGGGGGGTGTTGTGTAACTGTTTTTTGCCGGAAAGACGC	42	1.1×10^{5}	4.4×10^{-3}	0.88	0.20
2	GACGCTATGACTGGTATGGTGTGTGAGCTTCGAGGTCTGA	109	1.2×10^{5}	1.3×10^{-2}	0.92	0.15
3	CGCGGCGTTGCGTAGCGCCAATCCTGCACAGAGGCGGGCG	141	7.5×10^{4}	1.1×10^{-2}	0.88	0.11
4	TGACACCTGGCACGAATGCGCAGACGACGTTGTTTTGCGT	19	2.6×10^{5}	5.0×10^{-3}	0.89	0.07
5	ACTTTTGCTTGAGATTCGGCGCATTTCGTGCAGTGATGTA	27	2.9×10^{5}	7.8×10^{-3}	0.91	0.15
6	CTTAGGTCGGCCTTTGACCATTCGTGCAAAACTTAGCGAA	41	3.0×10^{5}	1.2×10^{-2}	0.94	0.18
7	GAATGCGCTTGAGCGCGCCCGATGATATAGCGAACCAAGC	54	1.7×10^{5}	9.1×10^{-3}	0.87	0.11
8	TTAATTTTGAGATCTTTGTCGAGTTGCTCGGTGAAAAGTC	70	2.2×10^{5}	1.5×10^{-2}	0.98	0.24
9	AGAGCAAGTCATGATCGCGAGTTTAGCTTTTTCATTTCCC	86	2.3×10^{5}	2.0×10^{-2}	0.97	0.14
10	GGCACTATCGATTATCGTGCGCATTGGTTGTGTCCAGGCT	7.3	9.8×10^{5}	7.1×10^{-3}	0.97	0.07
11	TCACGTCTGTTGTACACTTGTGCAATGCAAGGACGGAAAA	12	4.6×10^{5}	5.4×10^{-3}	0.99	0.17
12	ACGAGTAGGCCGCATTATGATGATTTAAACTGTTTCATGC	16	3.3×10^{5}	5.2×10^{-3}	0.99	0.37

SI Table 3.5 High performing standard and modified S1 and TNFα aptamers chosen for replicative trials. Color scheme: standard library (green) and Trp library (blue).

avg, average; std, standard deviation

Protein Target	Aptamer	TNA Sequence (3'-2')	K_D (nM)	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	Fit (R^2)	Variance (Chi^2)
S1	Standard-2	CGTGCACGATCTAGCCAGTGC TCGCTTGAAATGGGCGCAG	27	9.0×10^3	2.5×10^{-4}	0.99	0.16
			46	5.5×10^3	2.6×10^{-4}	0.99	0.58
			28	9.8×10^3	2.7×10^{-4}	0.99	0.74
		avg ± std	34 ± 11				
S1	Trp-1	TTGGCGAGATCCCCTATACGT GAGCTTCCTTTTATCCTTT	2.0	6.7×10^4	1.3×10^{-4}	0.99	2.94
			3.8	3.6×10^4	1.4×10^{-4}	0.99	0.23
			3.5	3.6×10^4	1.3×10^{-4}	0.99	0.66
		avg ± std	3.1 ± 1.0				
TNFα	Standard-2	GACGCTATGACTGGTATGGTG TGTGAGCTTCGAGGTCTGA	109	1.2×10^5	1.3×10^{-2}	0.92	0.15
			76	7.5×10^4	5.7×10^{-3}	0.94	0.28
			62	9.2×10^4	5.7×10^{-3}	0.95	0.20
		avg ± std	82 ± 24				
TNFα	Trp-3	ACGAGTAGGCCGCATTATGAT GATTTAAACTGTTTCATGC	16	3.3×10^5	5.2×10^{-3}	0.99	0.37
			17	5.8×10^5	9.9×10^{-3}	0.99	0.09
			27	2.6×10^5	7.1×10^{-3}	0.99	0.23
		avg ± std	20 ± 6.1				

SI Table 3.6 DNA S1 and TNFa control aptamers.

Protein	Aptamer	K_D (nM)	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	Fit (R^2)	Variance (Chi^2)
TNFα	S01	25.7	7.21×10^{4}	1.85×10^{-3}	0.93	1.55
TNFα	B01	90.5	3.36×10^{4}	3.04×10^{-3}	0.95	2.49
Spike	CoV2-RBD-1C	0.08	6.40×10^{5}	5.06×10^{-5}	0.88	301
Spike	CoV2-RBD-4C	2.6	2.05×10^{5}	5.23×10^{-4}	0.99	7.92
S1 RBD	CoV2-RBD-1C	5.4	8.30×10^{5}	4.46×10^{-3}	0.97	26.4
S1 RBD	CoV2-RBD-4C	1.6	4.70×10^{5}	7.53×10^{-4}	0.98	15.15

SI Table 4.1 TNA oligonucleotides, written from 3'-2' using IDT notation for modified positions.

Oligo Name	Sequence (3'-2')
TNA P1A	/5IRD800/ACC ACC TGA GTA TAG TTA AGT A
TNA P1C	/5IRD800/ACC ACC TGA GTA TAG TTA AGT C
TNA P1G	/5IRD800/ACC ACC TGA GTA TAG TTA AGT G
TNA P1T	/5IRD800/ACC ACC TGA GTA TAG TTA AGT T
TNA P2	/5Phos/GTT TGT GAG TAA TTG TTA TTG T
TNA P3A	/5Phos/AAA TTG TTA TTG TAG AAT ATT G
TNA P3C	/5Phos/CAA TTG TTA TTG TAG AAT ATT G
TNA P3G	/5Phos/GAA TTG TTA TTG TAG AAT ATT G
TNA P3T	/5Phos/TAA TTG TTA TTG TAG AAT ATT G

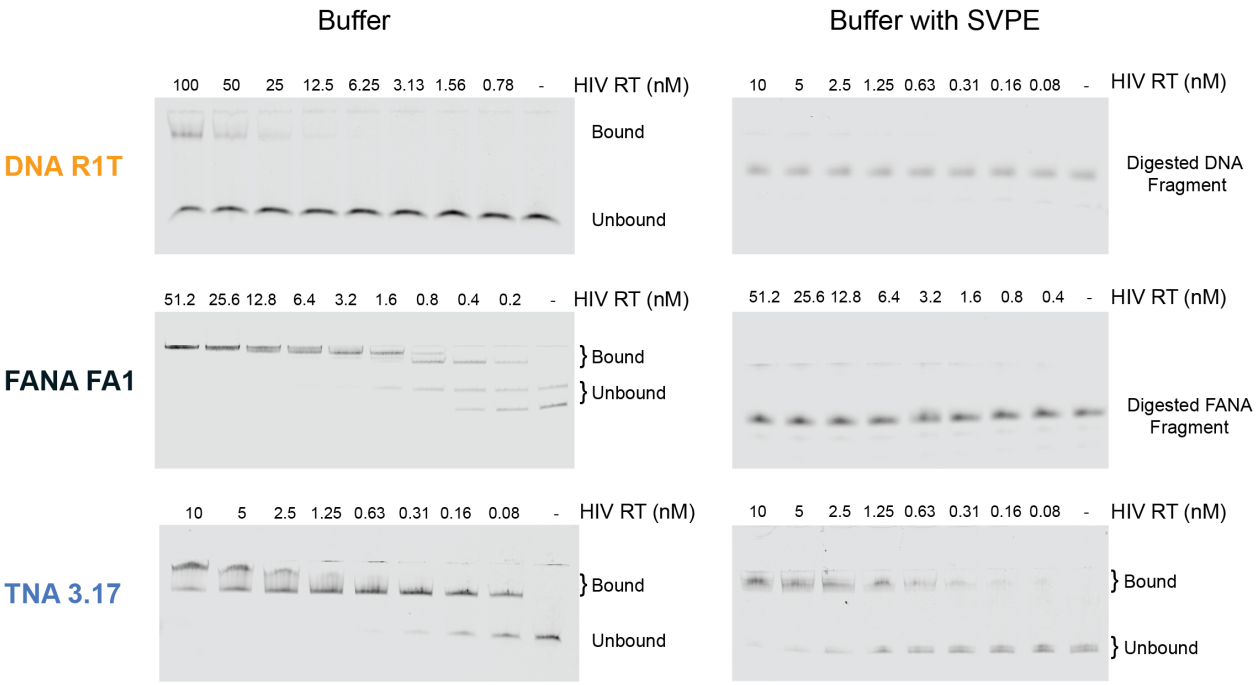
SI Table 4.2 DNA oligonucleotides, written from 5'-3' using IDT notation for modified positions.

Oligo Name	Sequence (5'-3')
DNA splint	ACA ATA ACA ATT ACT CAC AAA CTA CTT AAC TAT ACT CAG GTG GT
DNA splint 2	AAA AAA AAC AAT ATT CTA CAA TAA CAA TTA TAC TTA ACT ATA CTC AGG TGG TAA AAA AAA
DNA P1A	/56-FAM/ACC ACC TGA GTA TAG TTA AGT A
DNA P1C	/56-FAM/ACC ACC TGA GTA TAG TTA AGT C
DNA P1G	/56-FAM/ACC ACC TGA GTA TAG TTA AGT G
DNA P1T	/56-FAM/ACC ACC TGA GTA TAG TTA AGT T
DNA P2	/5Phos/GTT TGT GAG TAA TTG TTA TTG T
DNA P3A	/5Phos/AAA TTG TTA TTG TAG AAT ATT G
DNA P3C	/5Phos/CAA TTG TTA TTG TAG AAT ATT G
DNA P3G	/5Phos/GAA TTG TTA TTG TAG AAT ATT G
DNA P3T	/5Phos/TAA TTG TTA TTG TAG AAT ATT G

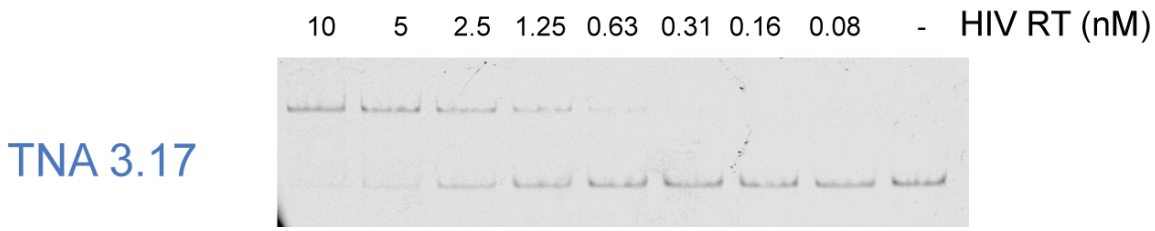
APPENDIX B

SUPPLEMENTARY FIGURES

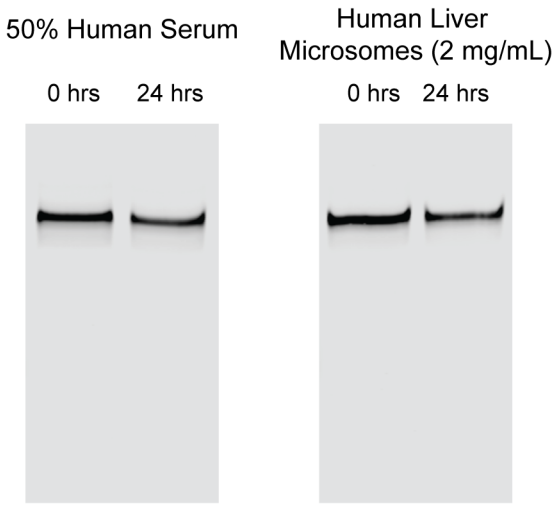
SI Figure 2.1 Representative native PAGE images used to calculate the K_D values for IR680 labeled aptamers in buffer (left) or in buffer after a 1-hour pre-incubation with snake venom phosphodiesterase (SVPE, right).



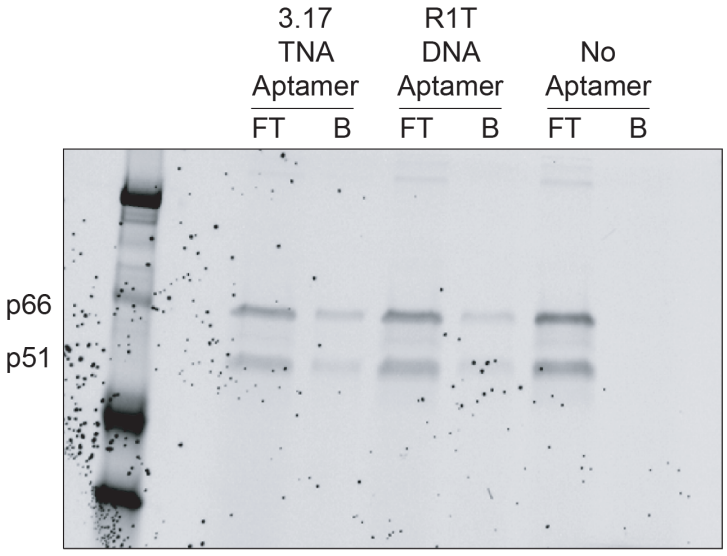
SI Figure 2.2 Characterization of HIV-RT 3.17 at Elevated Temperature, where native PAGE was used to calculate the K_D value for HIV-RT 3.17 at 37°C. The gel shift assay reveals a K_D of 2.5 nM at this temperature**.**



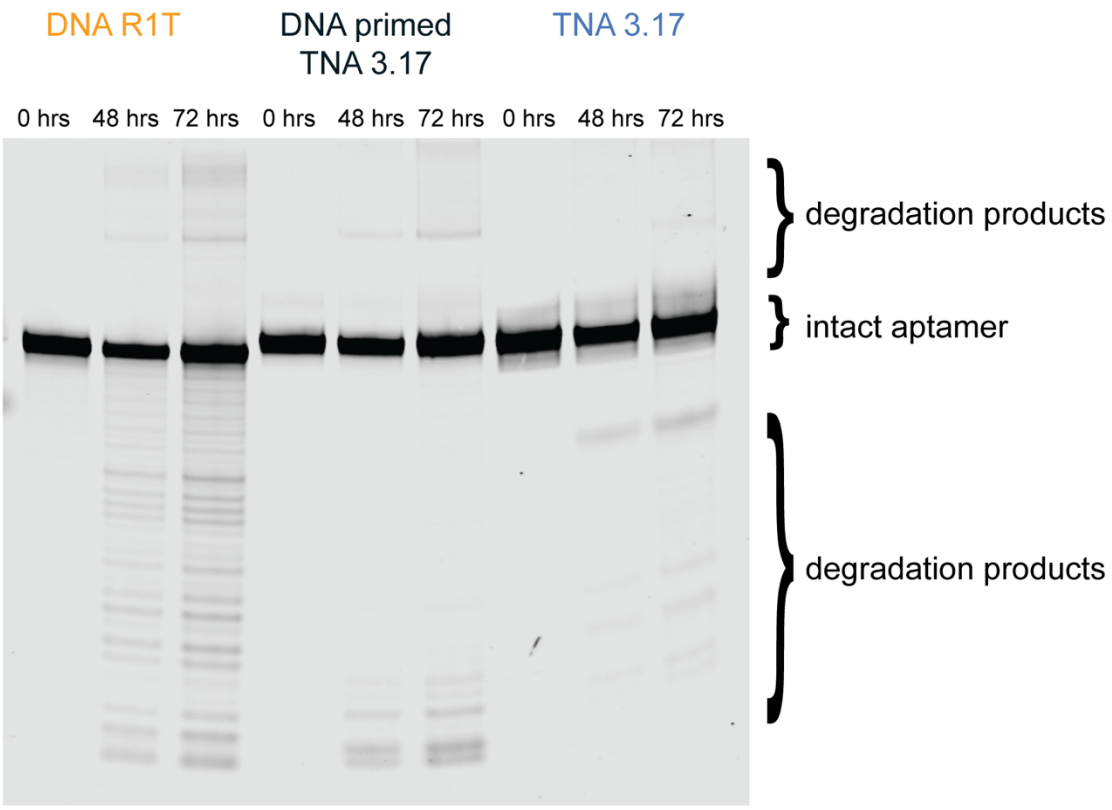
SI Figure 2.3 Biostability validation of all TNA HIV-RT 3.17 using 50% human serum and 2 mg/mL human liver microsomes.



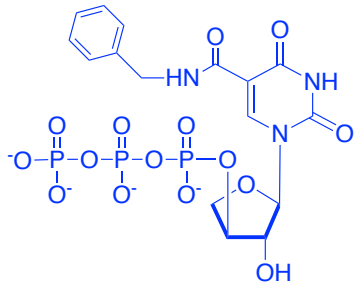
SI Figure 2.4 SDS-PAGE image shows the HIV RT (subunits p66 and p51) in the flow through (FT) or bound to the beads (B), after incubation with biotinylated 3.17, R1T, or buffer alone.



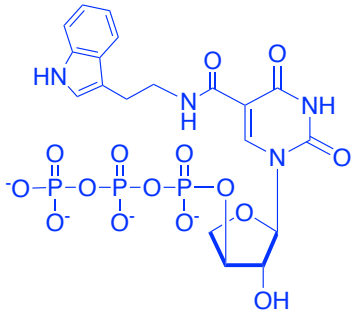
SI Figure 2.5 DNA R1T, DNA primed TNA HIV-RT 3.17, and all TNA HIV-RT 3.17 were heated at 75 °C for 0-72 hours, then analyzed by denaturing PAGE for signs of decomposition. All three constructs show similar thermal stability.



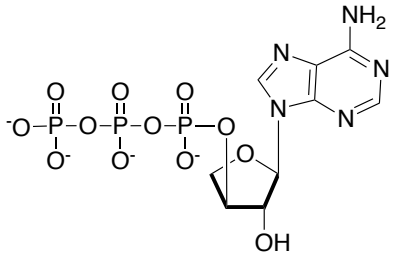
SI Figure 3.1 TNA triphosphate monomers used in this study.



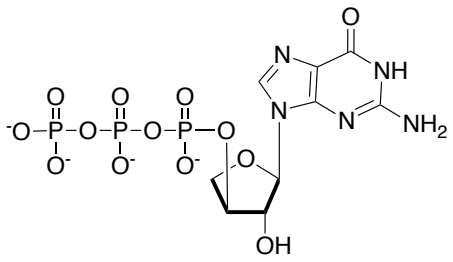
tUTP^{Phe}



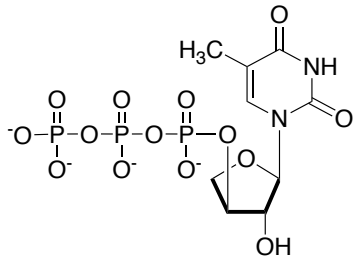
tUTP^{Trp}



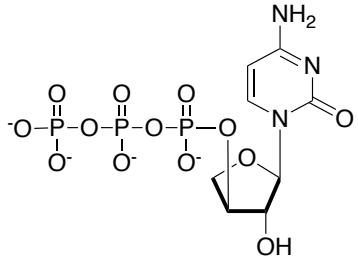
tATP



tGTP

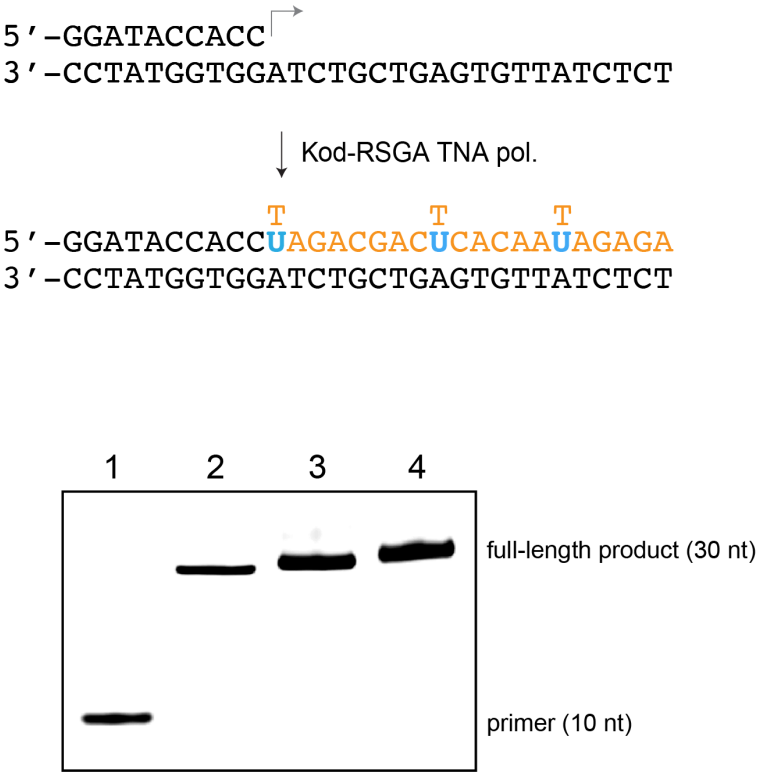


tTTP

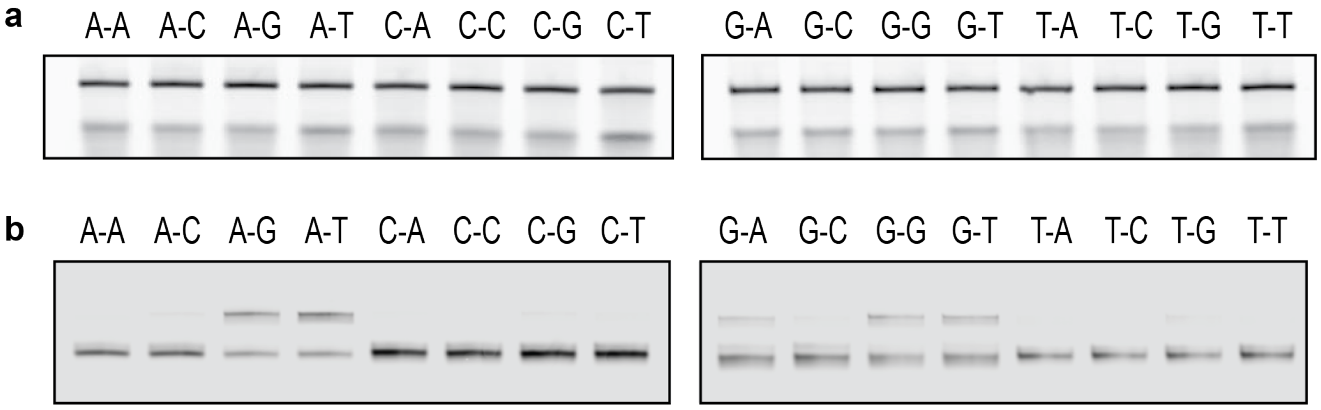


tCTP

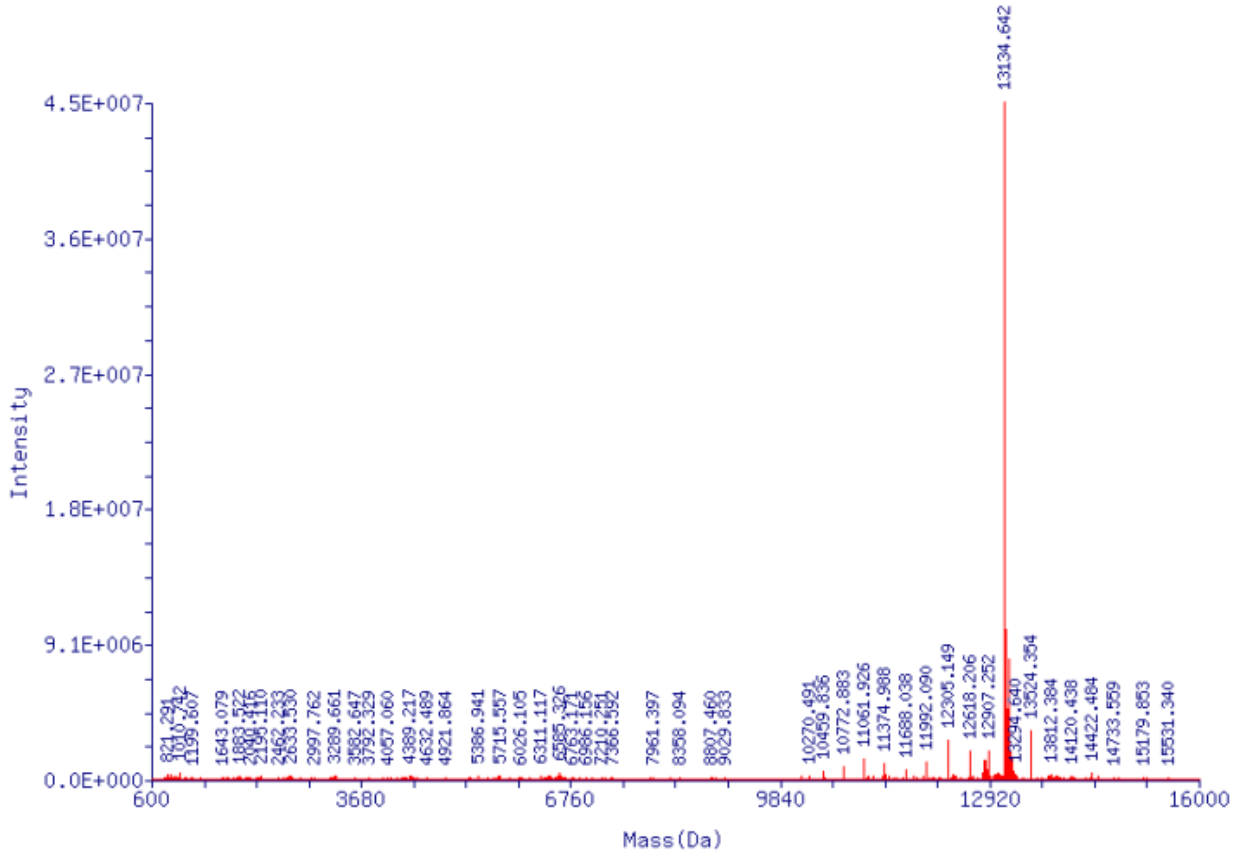
SI Figure 3.2 Primer extension assay. A 5'-labeled DNA primer annealed to a DNA template was extended with TNA using Kod-RSGA. Lane 1: primer only. Lane 2: TNA synthesis with natural bases. Lane 3: TNA synthesis with a tNTP mixture containing $tUTP^{Phe}$ in place of tTTP. Lane 4: TNA synthesis with a tNTP mixture containing $tUTP^{Trp}$ in place of tTTP. Color scheme: DNA (black), TNA with natural bases (orange), and TNA with modified bases (bold blue).



SI Figure 4.1 A representative set of raw data for the ligation junction assay shown in Figure 4.4b. The sequence specificity of T3 DNA ligase was measured for (a) DNA and (b) TNA substrates using oligonucleotides that spanned all of the 16 possible ligation junctions for the terminal nucleotide position on the donor and acceptor strands. Reaction yields were monitored by denaturing polyacrylamide gel electrophoresis.



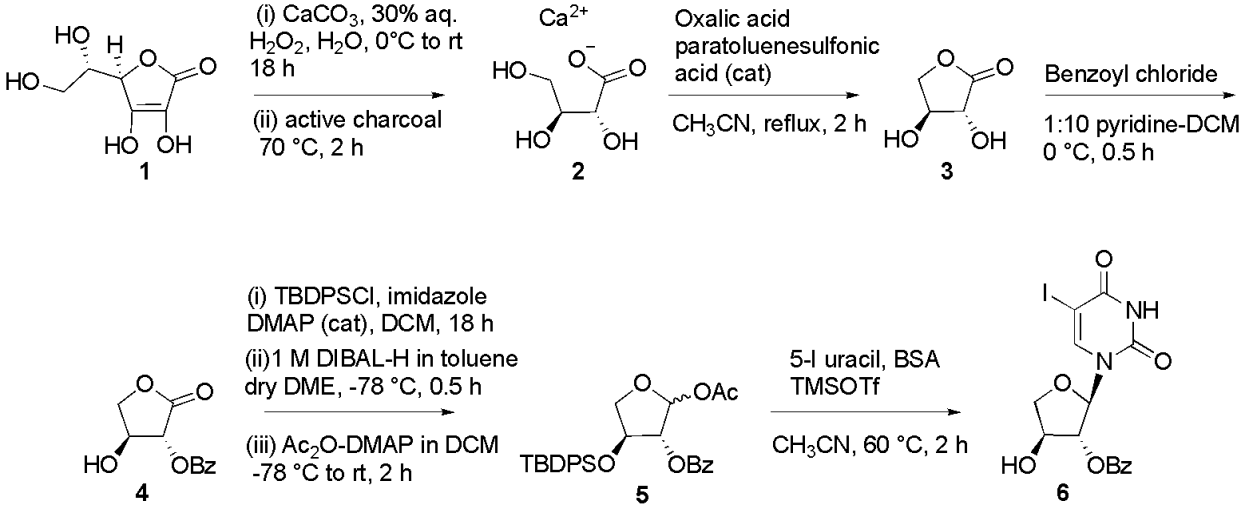
SI Figure 4.2 The mass spectrum for the ligated TNA-TNA product measured by high resolution ESI. Theoretical mass 13,138 Da, observed mass 13,135 Da.



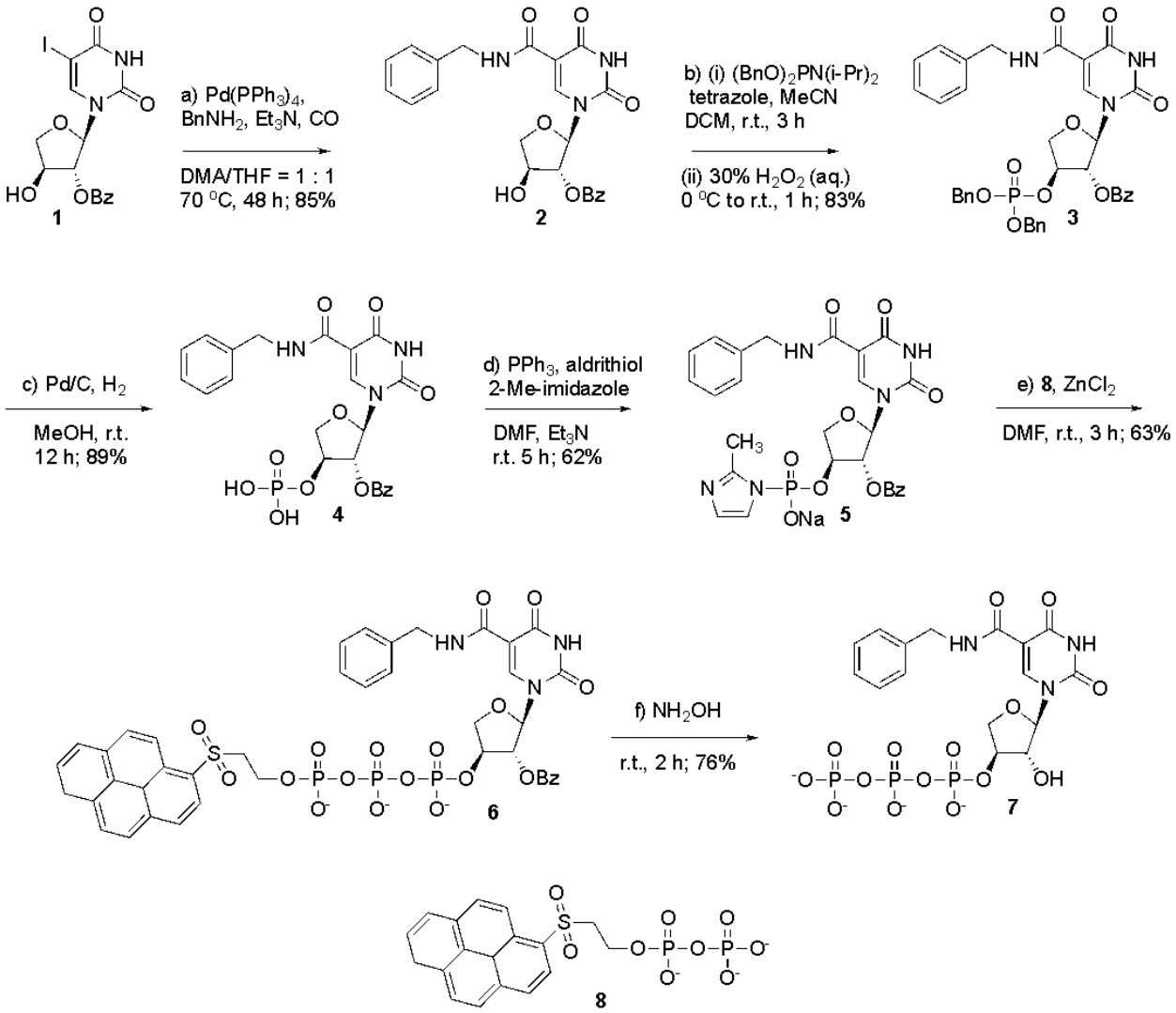
APPENDIX C

SUPPLEMENTARY SCHEMES

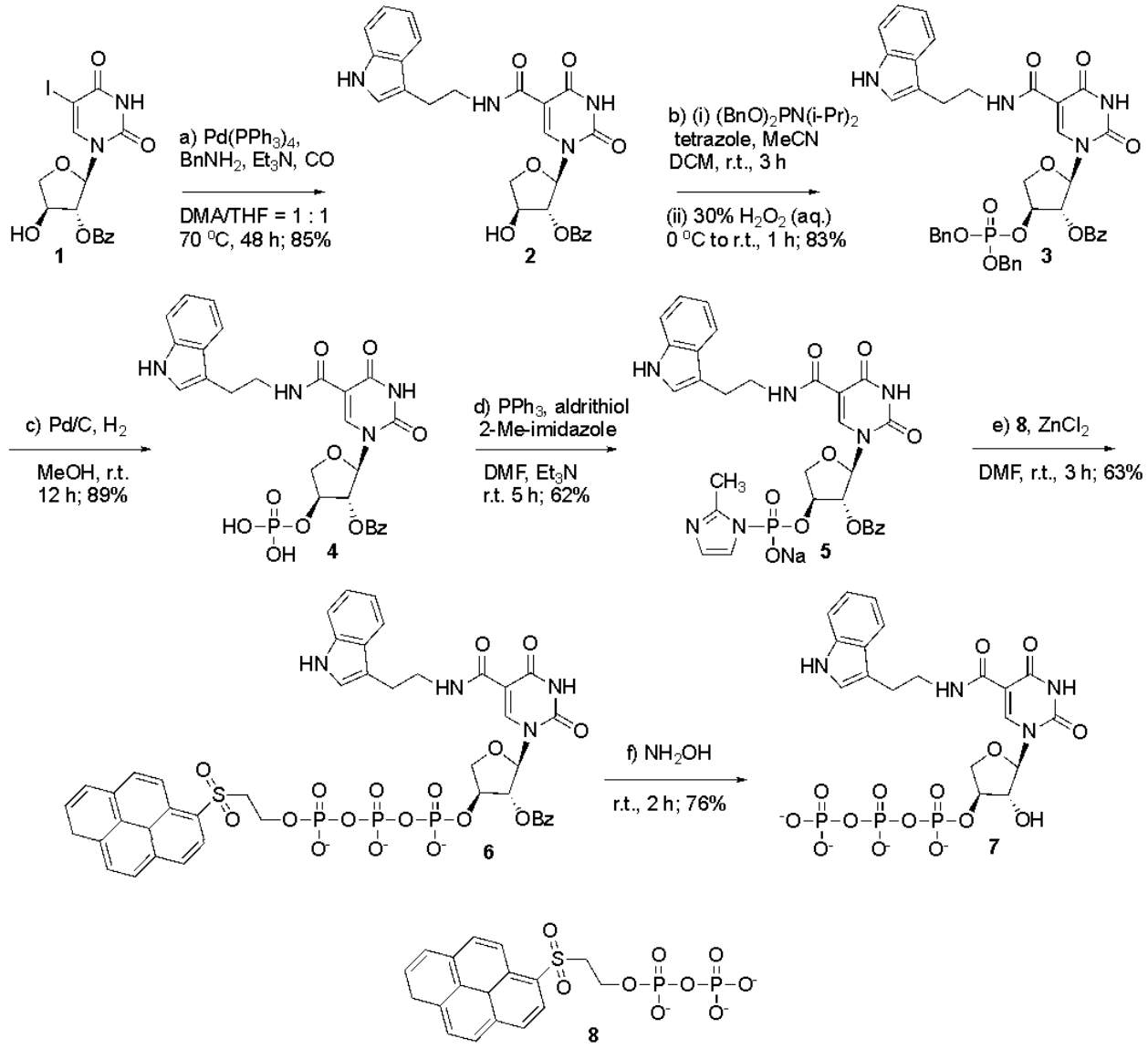
SI Scheme 3.1 Chemical synthesis of 5-iodo-1-(2'-*O*-benzoyl-α-L-threofuranosyl)-uracil.



SI Scheme 3.2 Chemical synthesis of C5-benzyl modified tUTP.



SI Scheme 3.3 Chemical synthesis of C5-tryptophan modified tUTP.



SI Scheme 3.4 Chemical synthesis of tNTP with natural bases.