Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Hector S. Argueta-Gonzalez Date

Exploration of Bilingual Peptide Nucleic Acids

By

Hector S. Argueta-Gonzalez
Doctor of Philosophy

Chemistry

Dr. Jennifer Heemstra
Advisor

Dr. Monika Raj
Committee Member

Dr. Vincent Conticello
Committee Member

Accepted:

Kimberly Jacob Arriola, Ph.D.

Dean of the James T. Laney School of Graduate Studies

Date

Exploration of Bilingual Peptide Nucleic Acids

By

Hector S. Argueta-Gonzalez

B.S. Georgia State University, 2017

Advisor: Jennifer M. Heemstra, PhD

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Chemistry
2022

Abstract

Exploration of Bilingual Peptide Nucleic Acids

By Hector S. Argueta-Gonzalez

Peptide nucleic acid (PNA) is an unnatural nucleic acid mimetic that contains a neutral pseudopeptide backbone, replacing the negatively charged phosphate backbone found in DNA and RNA. The change in backbone imparts changes such as having higher affinity and specificity for nucleic acid hybridization and increases stability. Due to these properties, PNA has a great potential for use in various applications such as nucleic acid detection and antisense interactions. As well, due to its modifiable backbone, amino acid residues could be installed at the γ-position to direct for assembly. Previous work has been done in our laboratory to develop a bilingual PNA biopolymer that could incorporate both nucleic acid and amino acid coding to develop a self-assembling oligomer that could undergo stimuli-responsive disassembly upon the recognition of a target sequence. However, this system has yet to be further explored to understand the coding interactions and explore its potential applications. This thesis aims to highlight the elucidation and potential applications of bilingual PNA systems. In chapter 1, I describe PNA as a nucleic acid, its previous applications, and potential for target recognition systems. In chapter 2, I describe a toehold-mediated displacement system development to induce stimuli-responsive assembly in bilingual PNA systems upon the addition of a releasing strand. In chapter 3, I elucidate the nucleic acid and amino acid influences on the bilingual PNA systems to further its potential in nanotechnology and biomedical applications. In chapter 4, I explore the use of triplex forced intercalation PNA probes for the detection of Adenosine to Inosine (A-to-I) edits. In chapter 5, I describe the application of a PNA as a capture strand for the development of a PNA based biosensor. In chapter 6, I discuss the testing and optimization of a coenzyme A (CoA) biosensor. In chapter 7, I describe the synthesis of a photocaged aspartic acid for use in bilingual PNA biopolymers to introduce spatial/temporal control. In chapter 8, I sought to synthesize a novel Nile Red (NR) PNA monomer to further expand PNA detection applications. Finally, I summarize how the presented studies could be utilized for future directions and applications.

Exploration of Bilingual Peptide Nucleic Acids

By

Hector S. Argueta-Gonzalez

B.S. Georgia State University, 2017

Advisor: Jennifer M. Heemstra, PhD

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Chemistry
2022

Acknowledgments

My time at Emory University came with more challenges than I expected to experience. My journey started with having to balance my family responsibilities with my graduate school ones. As I learned, there was no perfect way to handle these two major areas of my life and decided to focus on school because it would only be a temporary shift in focus. I also know that it came at the cost of missing out on a lot of events and experiences. However, as my time in graduate school has come to an end, I want to express gratitude to everyone involved and that understood what this experience and time would mean to me.

The person I want to thank the most is Dr. Kesiya Cherian, PharmD, my partner of almost nine years. You kept being patient with me and supporting me, especially with how much I dedicated to my time to work. I am more than ecstatic to finally have a wrap on a lot of our lives this year, with both of us finishing graduate school, moving onto our careers, and most importantly, finally getting married. I can't wait to start calling you my wife. I'll continue to be happy with you besides me.

I would also like to thank my family. My parents for always inviting me to go everywhere with them, taking me away from work and stress, and always give me advice, even if it was unsolicited at times. My brothers for always understanding my time in school and leaving you guys at home. While I was only really forty minutes away, it always seemed impossible to make time to see each other. I hope to be around more but I also know that work will likely take me away and you guys are more grown now with your own plans. As well, I would like to thank the Cherians for accepting me into their lives and supporting me through school as one of their own. I would also like to thank all my extended family and friends as you all have been a part of my journey as well. Even the little things went a long way. Ate lunch with me? Yup. Talked to me on

discord for one night? Yes. Being a part of and/or helping with my wedding? Definitely. I sincerely appreciate every one of these moments.

Finally, I would like to thank the Heemstra lab. For had they not come to Emory when they did, I may not have gotten this far. Both in terms of accepting me into the lab but also helping me learn and grown. To allow me to extend myself even beyond just the lab. I hope to leave having made the lab and chemistry department a better place than even when I joined.

List of Tables and Figures

Chapter 1: Introduction

1.1 Nucleic Acid Base Biosensors

Nucleic acids are a highly effective molecular platform for storing and distributing information. In nature, the two naturally occurring nucleic acids are DNA and RNA, which can undergo specific Watson-Crick base pairing to form duplexes.

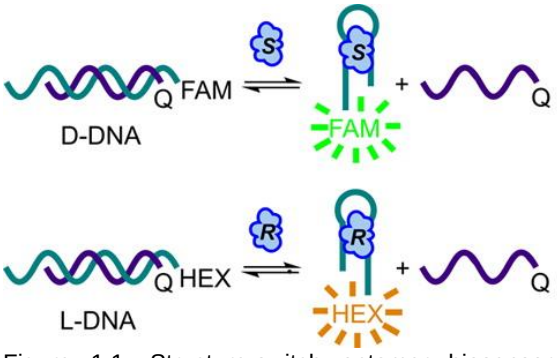


Figure 1.1: Structure-switch aptamer biosensor. Figure reproduced [8]

Nucleic acid based biosensors (Fig. 1) have been developed and function by having a designed oligomeric sequence bound to a complementary sequence with specified targets or targeted sequences. One of the more prominent developments in this form of biosensing comes from the use of an aptamer bound to a complementary sequence. The aptamer is modified to contain a fluorescent dye, while the quencher-labeled complementary strand “quenches” the fluorescent signal, indicating duplex formation. Once the target molecule is introduced into the system, the aptamer-based biosensor can undergo structure switching target recognition, releasing the complementary sequence. This removes the quencher from the fluorophore’s proximity, regenerating the fluorescent signal to quantitatively detect targets of interest through fluorescence spectroscopy [1-3].

While these structure-switching aptamer biosensors have been successfully applied to a variety of small molecule and protein targets, there are still limitations to this technology. A major hindrance to these systems is the stability of natural nucleic acids. In an aptamer based biosensor, hybridization is at equilibrium when the complementary sequences are duplexed [2, 4]. Thus, it is highly unlikely to be able to have full fluorescence quenching due to some fluorescently-labeled oligomers being unbound [3]. This results in background fluorescence and having to optimize for the signal-to-noise ratio generated, reducing the limits of detections for the target of interest. In

biological settings, nucleases recognize and hydrolyze both DNA and RNA. Once broken down, the sequence is no longer retained and its function is lost, severely limiting the practical application capabilities of these DNA-based biosensors. To overcome the biostability limitation, self-assembling nucleic acid polymers were created.

1.2 DNA-polymer conjugate assemblies

Oligomers have also been conjugated to organic polymers and lipids for use in the design and synthesis of structures known as DNA-Polymer Conjugates (DPCs) that can undergo self-assembly. The DPCs are normally dissolved in solvent and not able to form into a structure until the critical micelle concentration (CMC) is reached. The CMC is dependent on chemical properties of the DPC and the environmental factors like pH and ionic strength[5]. Based on the design of the hydrophobic region of the conjugate, the amphiphilic structure can undergo structural reorganization under the specified conditions to form micelles, vesicles, and tubes (Fig) [6]. Due to structural differences, vesicles have amphiphilic guest binding while both micelle and tubes are limited to hydrophobic guest binding[7]. An example of a stimuli-responsive DPCs comes from the work of Chien et al. as they developed an amphiphilic copolymer that reverted between spherical and cylindrical upon the introduction of sequence of interest. Upon binding the oligomer, the structure would change to a spherical shape due to electrostatic repulsion and rigidity. After removal of the sequence through enzymatic cleavage, the structures reverted to a more linear structure. However, this example is limited by DNA truncation that is necessary to alter its structure and cannot reversibly bind.[8].

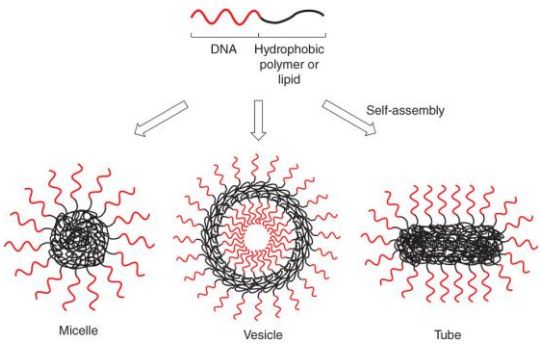


Figure 2: DPC conjugate monomer structure and possible reorganization structures. Figure reproduced [16]

As newer iterations and more sophisticated DPCs were synthesized and tested, the beneficial properties were investigated. The nucleic acid sequences retained their ability to hybridize,

became more sensitive to mismatched sequences, decreased nuclease degradation, and added the attribute of cell permeability with minimized cytotoxicity. These allowed DPCs to be applied to overcome previous hurdles. PEG linkers have been added on to the polymer to increase circulatory lifetime, folate has been used to target cancer cells, and DPCs can act as antisense therapeutic [9-11]. As this new field of research opened up, the methods became optimized but there still remained some drawbacks. A primary drawback stems from DNA only able to act as the hydrophilic region and the polymer act as the hydrophobic region. Also, a another major drawback comes from the modified structures that can hinder their intended function [7].

The design of these self-assembling structures has made improvements to the field but would benefit overall from the application of synthetic nucleic acids that can be highly modifiable. One such synthetic nucleic acid is peptide nucleic acid.

1.3 Peptide Nucleic Acids

Peptide nucleic acids (PNA) have emerged as a promising alternative to DNA and RNA over other synthetic nucleic acids. PNA is comprised of a repeating polyamide rather than phosphodiester backbone (Fig. 2). This charge-neutral backbone displays exceptional Watson-Crick hybridization and sequence selectivity with complementary RNA and DNA sequences and has strand invasion capabilities [12]. This property results in melting temperature differences of 15 ℃ per mismatch in PNA:DNA duplexes versus 11 ℃ in DNA:DNA sequences [13]. The polyamide backbone also adds a benefit to the analog as it decreases recognition by nucleases and proteases, promoting its biostability. Other properties found in PNA are that it can undergo Hoogsteen base pairing, allowing for the formation of triplexes [14]. Finally, PNA also confers biostability to the natural nucleic acids as the

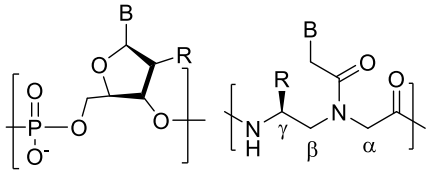


Figure 1.3: Chemical structures of DNA, RNA and PNA

duplex formed has greater thermodynamic stability and increased nuclease resistance [12, 13, 15, 16]. One limitation of PNA is its reduced solubility in solution, however this can be largely avoided through rational sequence design and solubility promoting modifications. PNA has been employed in several biological applications.

PNA has been employed in several biological applications including as an antisense therapeutic and biosensor. For example, the targeting of LacI mRNA to inhibit *E. coli* gene expression, targeting Ha-*ras* mRNA to terminate polypeptide elongation, and, a peptide-PNA conjugation using the CPP (KFF_3)K was used to inhibit bacterial growth [17-19]. As a method for detection, PNA has been used as a graphene-oxide (GO) based biosensor in which the PNA sequences were labeled with a fluorescent probe and the unbound fluorescent sequences were quenched by the GO while the duplexed PNA:DNA fluorescence could be measured [20, 21]. PNA has also been conjugated to polymers in a similar fashion as the aforementioned DPCs. These structures were found to retain the higher binding affinity and sequence selectively properties while undergoing self-assembly into micelle structures [22-24].

1.4 Modified Peptide Nucleic Acid

The applications of PNA have been enhanced further upon the development of modified sequences. Synthetic routes discovered could produce α-, β-, and γ-modified monomers from inexpensive starting materials. These monomers could then be inserted into the sequences, impacting the structural and chemical properties of the strand. The enhanced properties such as affinity and specificity would increase the capabilities of modified PNAs and thus increase their applications for use such as in biosensors and capture probes. Through the introduction of a D-lysine at the α-position, an beneficial increase in PNA:DNA duplex formation was pronounce.[25] However, this effect was noted to be dependent on the charge, size, and steric components of

the amino acid residue incorporated, as the L-lysine was found to reduce stability and glycine residues had no noticeable effect.[25, 26] Furthermore, the inclusion of multiple modified PNA monomers increased the capability of the sequence to distinguish between parallel and antiparallel complementary strands.[27] By taking advantage of the modified PNA backbone properties, a lysine-rich PNA sequence was capable of detecting for DNA mutation R553X found in cystic fibrosis through the use of 3 lysine residues in capillary electrophoresis experiments. [28] However, the introduction of D-arginine produced differing results, as consecutive D-monomers destabilized the duplex and semi-consecutive positioning returned the increased affinity.[29] The arginine modified sequences were also found to have an increase in cellular uptake in mammalian cells and had reduced toxicity, demonstrating the various properties the amino acid repertoire had on the binding capabilities of PNA oligomers.[29, 30]

While single substitution had a notable synthetic pathway for the α-modified PNA monomers, the β-modified PNA monomers typically contained remained limited. As such, a single methyl group was able to be installed at the β-position using both the S- and R- monomers, allowing for comparison. Testing showed that the S- monomer containing PNA sequences were not capable of forming a duplex, likely due to with DNA while the R- monomer sequence had similar properties to unmodified PNA.[31] Other β-modified PNA monomers have been synthesized that contain cyclic structures along the backbone, typically incorporating the γ-position.[13] While these monomers have been synthesized, many of these modifications lead to similar properties that disrupt binding to RNA and DNA sequences. A notable exception was the introduction of cyclopentanes such as in the form of *t*cypPNA.[32-35]

1.5 γ-Modified Peptide Nucleic Acid

The third modification of PNA at the γ-position has been found to most significantly impact both sequence affinity and physiochemical properties. The L- monomer modification coordinates the oligomer into a right-handed helix that results in an increase of 4 ℃ per monomer. An additional benefit is that the modification can be altered to a charged substituent such as arginine known as GPNAs to increase solubility and cell permeability, overcoming insolubility, giving these constructs strong potential for application in a biological setting as previously demonstrated in α-modified PNA sequences [13, 36]. As well, the introduction of L-amino acids were found to increase single base mismatch discrimintation.[37, 38] Another key factor in the development of γ-modified PNA monomers was the low-cost and ease of synthesis, using Fmoc-protected natural L-amino acids, as the high-cost of synthesis for previously mentioned monomers discouraged clinical trials.[39]

Several studies have expanded on the potential applications of γ-modified PNA. For example, Bichismita et al. developed a miniPEG γ-modified monomer that increased both hybridization and water solubility properties [40]. GPNAs were used by Katarzyna Gorska *et al.* for developing a nucleic acid templated reaction, using an mRNA target as a template to facilitate the increase of effective concentration and coordinate the reduction of azidorhodamine [41]. However, the reaction was sequence dependent to increase the effective concentration, mismatches would significantly stop the reaction from occurring.

1.6 Developing a Dual Encoded Biopolymer

Through the incorporation of amino acid residues along the backbone of the PNA monomers, the potential to develop a dual encoded mimetic was presented. Previous works that combined the two codes into one such as through the DPCs and PNA-peptide conjugates still suffered from various issues such as the two sequences acting independently.[22, 23] However, the γ-modified

PNA monomers open the potential to develop a novel construct that could amalgamate both properties. Previous works such as Marchelli and coworkers used a “peptide” sequence along the PNA backbone to enhance the sequence’s activity.[42]

1.7 Summary and Objectives of this Dissertation

Chapter 2: Upon the development of a self-assembling bilingual PNA system that could undergo stimuli-responsive disassembly, our lab focused on the combining toehold-mediated displacement technology for application in bilingual PNA systems. Thus, an initial masking sequence (MS) was designed to be fully complementary to the bilingual PNA sequence, undergoing sequence recognition and inducing disassembly. Upon the introduction of a releasing sequence (RS), the amphiphilic PNA oligomer would be released, allowing for the biopolymer to undergo assembly. In chapter 2, we present the design, synthesis, and characterization of a novel bilingual PNA biopolymer containing a 5-Fluorescein (FAM) at the N-terminal along with 2 alanine and 2 lysine residues at the γ-positions to act as the hydrophobic and hydrophilic sources, respectively. We demonstrated the capability of the biopolymer to undergo fluorescence quenching when hybridized to a quencher-labeled masking sequencing. Further testing demonstrated a dose-dependent increase in fluorescence upon the introduction of a releasing sequence. Imaging techniques demonstrated the sequence was capable of undergoing self-assembly, stimuli-responsive disassembly, and stimuli-responsive assembly. This study demonstrates a novel approach to introduce control over the assembly and stimuli-responsive capabilities of the bilingual PNA biopolymer.

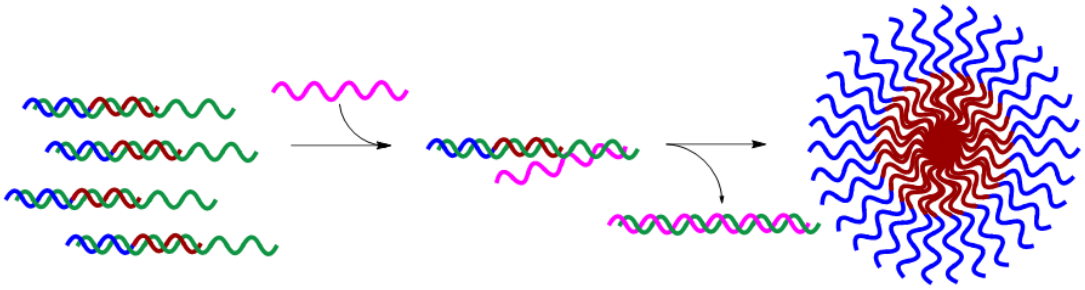


Figure 1.4: Toehold-mediated displacement to induce stimuli-responsive assembly.

Chapter 3: Upon the characterization of the initial PNA biopolymer, the capacity to introduce novel properties became evident. Thus, our lab hypothesized that through the introduction of variations such as differing steric bulk, charge, and sequence length, we would be able to learn the design rules of the amphiphilic oligomers. In Chapter 3, we describe the design, synthesis, and characterization of a small library of bilingual PNA biopolymers with changes in amino acid residues at the γ-positions and alterations in sequence length. We demonstrated that all sequences were capable of undergoing self-assembly into micelle-like structures and sequence recognition upon the introduction of a complementary sequence. Further evaluation elucidated that the length of the sequence determined the size of the structures formed. This systematic exploration of the dual encoded biopolymer introduces tunable properties that allow for an increased utility in biotechnology and nanotechnology applications.

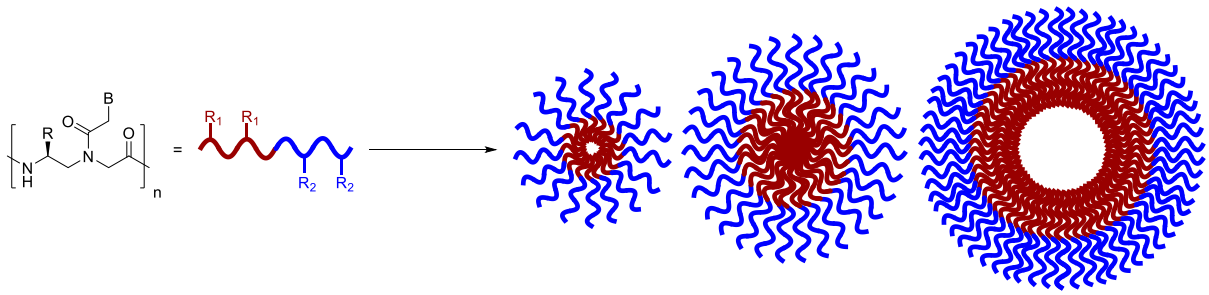


Figure 1.5: Elucidation of dual coding in the bilingual PNA biopolymer

Chapter 4: Previous work in our lab focused on the ability to detect and monitor for conversion of Adenosine to Inosine (A-to-I) through the ADAR enzyme. The dysregulation of the A-to-I edits has been shown to lead to diseased states such as Alzheimer's and various cancers. Thus, our lab attempted to develop novel fluorogenic probes in form of Forced Intercalation (FIT) PNA probes. While Thiazole Orange FIT PNA probes were successful in identifying A-to-I edits in linear RNA sequences, they were not capable of detecting in the known RNA hairpin HER1. Work in other laboratories demonstrated the possibility to develop a triplex FIT (tFIT) PNA probe that could bind via Hoogsteen base pairing interactions. Thus, in Chapter 4, we describe the design, synthesis, and characterization of tFIT TO PNA probes for the detection of A-to-I edits. Initial hybridization studies demonstrated an increase in fluorescence upon the introduction of the RNA target sequences. However, further studies displayed no differences in fluorescence, indicating that the tFIT PNA probes could not distinguish A and I-containing RNA sequences via Hoogsteen base pairing.

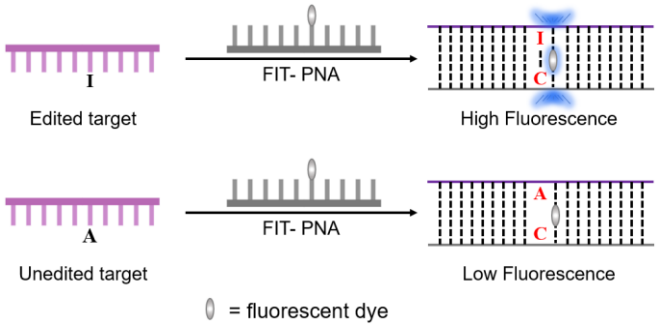


Figure 1.6: A-to-I FIT PNA probe undergoing an increase in fluorescence upon the recognition of target I sequence

Chapter 5: Structure-switching biosensors are highly prominent for the detection and quantification of small molecule targets. These biosensors employ an aptamer, a single-stranded nucleic acid sequence, and a hybridized complementary sequence, that, upon the introduction of the target of interest, induces a conformational change that can be converted into a signal output. While these systems are promising, they suffer from a lack of biostability and background fluorescence. In Chapter 5, we describe the use of PNA to overcome this limitation by designing, synthesizing, and characterizing short PNA complementary sequences to bind to the aptamer and release upon the introduction of the molecule of interest. Evaluation of PNA:DNA system

demonstrated its capability to develop a novel biosensors. With a preliminary PNA:DNA biosensor, we attempted to introduce γ-modified PNA monomers to develop a novel bilingual PNA:DNA biosensor system that would take advantage of the release of PNA to induce stimuli-responsive assembly. However, exploration of the γ-PNA:DNA system did not prove to be optimal for inducing assembly.

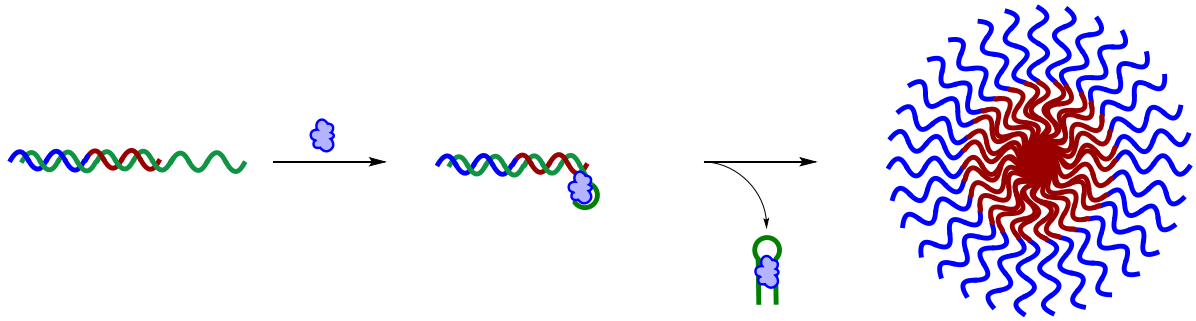


Figure 1.7: Structure-switching Biosensor releasing complementary PNA sequence to bind target of interest, allowing assembly to occur.

Chapter 6: While structure-switching biosensors are highly valued due to their ability to be adapted to a functional output, current Systematic Evolution of Ligands by Exponential Enrichment (SELEX) due not select for structure-switching (SS) aptamers. Thus, our lab previously recognized that a need to develop a novel SELEX method that could select for SS aptamers would expand SS biosensor repertoire of targets. Previous work in our lab designed and developed a homogeneous, PCR-based SELEX method to select for aptamers capable of acting as SS biosensors. In chapter 6, we describe the identification and characterization of six candidate sequences for CoA. Upon evaluation, the sequences were not found to be viable

Figure 8

candidates for SS biosensors that could bind and recognize CoA with high specificity and affinity. The basis of this work was used to further expand and design the RE-SELEX method that was capable of selecting for a SS biosensor for Kanamycin A.

References

1. Chen, J., et al., *A simple and rapid biosensor for ochratoxin A based on a structure-switching signaling aptamer.* 2012. **25**(2): p. 555-560.
2. Tan, Z., T.A. Feagin, and J.M.J.J.o.t.A.C.S. Heemstra, *Temporal control of aptamer biosensors using covalent self-caging to shift equilibrium.* 2016. **138**(20): p. 6328-6331.
3. Feagin, T.A., et al., *High-throughput enantiopurity analysis using enantiomeric DNA-based sensors.* 2015. **137**(12): p. 4198-4206.
4. Nutiu, R. and Y.J.J.o.t.A.C.S. Li, *Structure-switching signaling aptamers.* 2003. **125**(16): p. 4771-4778.
5. Peterson, A.M., et al., *3, 3'-Dioctadecyloxacarbocyanine perchlorate (DiO) as a fluorogenic probe for measurement of critical micelle concentration.* 2015. **7**(16): p. 6877-6882.
6. Edwardson, T.G., et al., *Site-specific positioning of dendritic alkyl chains on DNA cages enables their geometry-dependent self-assembly.* 2013. **5**(10): p. 868.
7. Peterson, A.M., J.M.J.W.I.R.N. Heemstra, and Nanobiotechnology, *Controlling self-assembly of DNA-polymer conjugates for applications in imaging and drug delivery.* 2015. **7**(3): p. 282-297.
8. Chien, M.P., et al., *Programmable shape-shifting micelles.* 2010. **49**(30): p. 5076-5080.
9. Jeong, J.H. and T.G.J.B.c. Park, *Novel polymer− DNA hybrid polymeric micelles composed of hydrophobic poly (D, L-lactic-co-glycolic acid) and hydrophilic oligonucleotides.* 2001. **12**(6): p. 917-923.

10. Alemdaroglu, F.E., et al., *DNA block copolymer micelles–a combinatorial tool for cancer nanotechnology.* 2008. **20**(5): p. 899-902.

11. Magnusson, J.P., et al., *Programmed assembly of polymer–DNA conjugate nanoparticles with optical readout and sequence-specific activation of biorecognition.* 2014. **6**(4): p. 2368-2374.

12. Nielsen, P.E. and M.J.C.I.M.B. Egholm, *An introduction to peptide nucleic acid.* 1999. **1**(1-2): p. 89-104.

13. Moccia, M., et al., *Insights on chiral, backbone modified peptide nucleic acids: Properties and biological activity.* 2014. **5**(3): p. e1107176.

14. Betts, L., et al., *A nucleic acid triple helix formed by a peptide nucleic acid-DNA complex.* 1995. **270**(5243): p. 1838-1841.

15. Karkare, S., D.J.A.m. Bhatnagar, and biotechnology, *Promising nucleic acid analogs and mimics: characteristic features and applications of PNA, LNA, and morpholino.* 2006. **71**(5): p. 575-586.

16. Uhlmann, E.J.B.c., *Peptide nucleic acids (PNA) and PNA-DNA chimeras: from high binding affinity towards biological function.* 1998. **379**: p. 1045-1052.

17. Dias, N., et al., *Antisense PNA tridecamers targeted to the coding region of ha-ras mRNA arrest polypeptide chain elongation1.* 1999. **294**(2): p. 403-416.

18. Good, L., et al., *Bactericidal antisense effects of peptide–PNA conjugates.* 2001. **19**(4): p. 360.

19. Good, L., et al., *Antisense PNA effects in Escherichia coli are limited by the outer-membrane LPS layer.* 2000. **146**(10): p. 2665-2670.

20. Ryoo, S.-R., et al., *Quantitative and multiplexed microRNA sensing in living cells based on peptide nucleic acid and nano graphene oxide (PANGO).* 2013. **7**(7): p. 5882-5891.

21. Ryoo, S.-R., et al., *High-throughput chemical screening to discover new modulators of microRNA expression in living cells by using graphene-based biosensor.* 2018. **8**(1): p. 11413.

22. Lau, C., et al., *Morphological characterization of self-assembled peptide nucleic acid amphiphiles.* 2006. **110**(18): p. 9027-9033.

23. Marques, B.F. and J.W.J.L. Schneider, *Sequence-specific binding of DNA to liposomes containing di-alkyl peptide nucleic acid (PNA) amphiphiles.* 2005. **21**(6): p. 2488-2494.

24. Vernille, J.P., L.C. Kovell, and J.W.J.B.c. Schneider, *Peptide nucleic acid (PNA) amphiphiles: synthesis, self-assembly, and duplex stability.* 2004. **15**(6): p. 1314-1321.

25. Katritzky, A.R. and T. Narindoshvili, *Chiral peptide nucleic acid monomers (PNAM) with modified backbones.* Organic & biomolecular chemistry, 2008. **6**(17): p. 3171-3176.

26. Dueholm, K.L., et al., *Peptide nucleic acid (PNA) with a chiral backbone based on alanine.* Bioorganic & medicinal chemistry letters, 1994. **4**(8): p. 1077-1080.

27. Menchise, V., et al., *Insights into peptide nucleic acid (PNA) structural features: The crystal structure of a d-lysine-based chiral PNA–DNA duplex.* Proceedings of the National Academy of Sciences, 2003. **100**(21): p. 12021-12026.

28. Tedeschi, T., et al., *Detection of the R553X DNA single point mutation related to cystic fibrosis by a "chiral box" d-lysine-peptide nucleic acid probe by capillary electrophoresis.* Electrophoresis, 2005. **26**(22): p. 4310-4316.

29. Zhou, P., et al., *Synthesis of cell-permeable peptide nucleic acids and characterization of their hybridization and uptake properties.* Bioorganic & medicinal chemistry letters, 2006. **16**(18): p. 4931-4935.

30. Dragulescu-Andrasi, A., et al., *Cell-permeable peptide nucleic acid designed to bind to the 5 '-untranslated region of E-cadherin transcript induces potent and sequence-specific antisense effects.* Journal of the American Chemical Society, 2006. **128**(50): p. 16104-16112.

31. Sugiyama, T., et al., *β-PNA: Peptide nucleic acid (PNA) with a chiral center at the β-position of the PNA backbone.* Bioorganic & medicinal chemistry letters, 2011. **21**(24): p. 7317-7320.

32. Myers, M.C., et al., *A cyclopentane conformational restraint for a peptide nucleic acid: Design, asymmetric synthesis, and improved binding affinity to DNA and RNA.* Organic letters, 2003. **5**(15): p. 2695-2698.

33. Pokorski, J.K., et al., *(S, S)-trans-Cyclopentane-constrained peptide nucleic acids. A general backbone modification that improves binding affinity and sequence specificity.* Journal of the American Chemical Society, 2004. **126**(46): p. 15067-15073.

34. Govindaraju, T., V.A. Kumar, and K.N. Ganesh, *cis-Cyclopentyl PNA (cp PNA) as constrained chiral PNA analogues: stereochemical dependence of DNA/RNA hybridization.* Chemical communications, 2004(7): p. 860-861.

35. Govindaraju, T., V.A. Kumar, and K.N. Ganesh, *(1 S, 2 R/1 R, 2 S)-cis-Cyclopentyl PNAs (cp PNAs) as Constrained PNA Analogues: Synthesis and Evaluation of a eg-cp PNA Chimera and Stereopreferences in Hybridization with DNA/RNA.* The Journal of organic chemistry, 2004. **69**(17): p. 5725-5734.

36. Sahu, B., et al., *Synthesis of conformationally preorganized and cell-permeable guanidine-based γ-peptide nucleic acids (γGPNAs).* 2009. **74**(4): p. 1509-1516.

37. Englund, E.A. and D.H. Appella, *γ-Substituted Peptide Nucleic Acids Constructed from l-Lysine are a Versatile Scaffold for Multifunctional Display.* Angewandte Chemie, 2007. **119**(9): p. 1436-1440.

38. Englund, E.A. and D.H. Appella, *Synthesis of γ-substituted peptide nucleic acids: a new place to attach fluorophores without affecting DNA binding.* Organic Letters, 2005. **7**(16): p. 3465-3467.

39. Sahu, B., et al., *Synthesis of conformationally preorganized and cell-permeable guanidine-based γ-peptide nucleic acids (γGPNAs).* The Journal of organic chemistry, 2009. **74**(4): p. 1509-1516.

40. Bahal, R., et al., *Sequence-Unrestricted, Watson–Crick Recognition of Double Helical B-DNA by (R)-MiniPEG-γPNAs.* 2012. **13**(1): p. 56-60.

41. Pianowski, Z., et al., *Imaging of mRNA in live cells using nucleic acid-templated reduction of azidorhodamine probes.* 2009. **131**(18): p. 6492-6497.

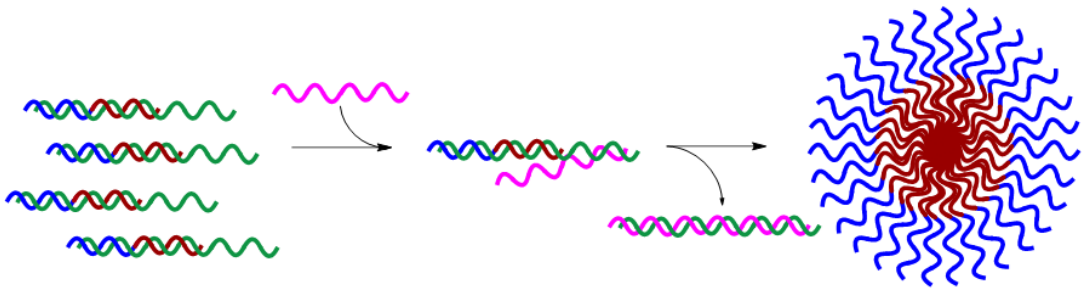
42. Sforza, S., et al., *A peptide nucleic acid embedding a pseudopeptide nuclear localization sequence in the backbone behaves as a peptide mimic*. 2010, Wiley Online Library.

Chapter 2: Stimuli-Responsive Assembly of Bilingual Peptide Nucleic Acids

Hector Argueta-Gonzalez, Colin S. Swenson, George Song, Jennifer M. Heemstra

Abstract:

Peptide nucleic acids (PNAs) are high-affinity synthetic nucleic acid analogs capable of hybridization with native nucleic acids. PNAs synthesized having amino acid sidechains installed at the γ-position along the backbone provide a template for a single biopolymer to simultaneously encode nucleic acid and amino acid sequences. Previously, we reported the development of "bilingual" PNAs through the synthesis of an amphiphilic sequence featuring separate blocks of hydrophobic and hydrophilic amino acid functional groups. These PNAs combined the sequence-specific binding activity of nucleic acids with the structural organization properties of peptides. Like other amphiphilic compounds, these γ-PNAs were observed to assemble spontaneously into micelle-like nanostructures in aqueous solutions and disassembly was induced through



hybridization to a complementary sequence. Here, we explore whether assembly of these bilingual PNAs is possible by harnessing the nucleic acid code. Specifically, we designed an amphiphile-masking duplex system in which spontaneous amphiphile assembly is prevented through hybridization to a nucleic acid masking sequence. We show that the amphiphile is displaced upon introduction of a releasing sequence complementary to the masking sequence through toehold mediated displacement. Upon release, we observe that the amphiphile proceeds to assemble in a fashion consistent with our previously reported structures. Our approach

represents a novel method for controlled stimuli-responsive assembly of PNA-based nanostructures.

Introduction

Nature has evolved two universal “languages” to encode information, function, and structure in biopolymers. The information stored in and communicated through nucleic acid and protein sequences are essential to functional biological systems. Nucleic acid sequences feature the high specificity necessary for accurate transmission of information but are limited by their relatively low diversity of chemical functionality. Conversely, amino acids contain high chemical diversity for structure and function but lack the ability to pass along genetic information.[1-3] We hypothesized that integration of these codes into a single biopolymer could overcome the inherent limitations they possess individually, as the resulting “bilingual” polymers would be able to encode information in both amino acid and nucleobase languages, thus generating materials with novel properties for biomedical and nanotechnology applications.

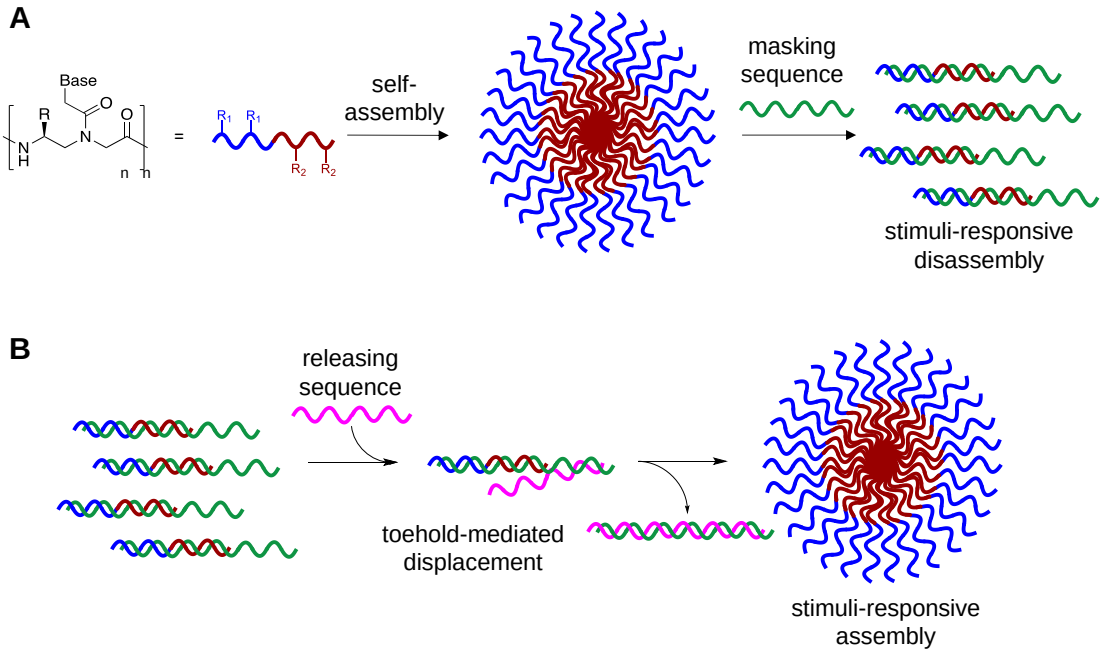


Fig. 1 (A) Bilingual PNAs having amphiphilic side chains at the γ-position to direct self-assembly and a nucleotide sequence to direct disassembly through recognition of a complementary masking sequence. The hydrophobic group (R_1) drives self-assembly by inducing aggregation, while the hydrophilic group (R_2) increases solubility. (B) Stimuli-responsive assembly through toehold-mediated displacement of the masking strand.

Our lab recognized that peptide nucleic acid (PNA) was uniquely suited to generate these bilingual biopolymers. Unlike DNA-peptide conjugates and other hetero-copolymers that contain separated domains, PNA allows for both nucleobases and amino acids to be arrayed along a pseudopeptide backbone, thus integrating nucleic acid and peptide functionality together.[3-5] This unnatural structure also provides key functional properties such as strong resistance to hydrolase degradation while retaining Watson-Crick-Franklin base pairing.[6, 7] Additionally, solid-phase synthesis allows precise placement of amino acid side chains along the backbone, and modifications at the g-position can be conveniently incorporated using amino acid starting materials.[5, 8] Taking these advantages into account, previous efforts in our lab led to the development of a bilingual g-PNA system that combined the structure and sequence specific activities of both classes of biomolecules into a single biopolymer. Encoding amphiphilicity through the amino acid sequence led to spontaneous assembly into micellar-like nanostructures, while the nucleic acid sequence directed stimuli-responsive disassembly through sequence-

specific hybridization to a complementary DNA or RNA strand.[9] While controlling disassembly offers utility for delivery applications, we recognized that potential applications could be broadened by introducing control for assembly as well. [10]

We recognized that this could be possible by masking the hydrophobic portion of the amphiphile with a nucleic acid strand to prevent assembly, and that if this masking strand was designed to have an unpaired toehold region, then addition of a fully complementary DNA or RNA sequence would lead to release of the amphiphilic PNA and subsequent assembly (Fig. 1).[11] Using fluorophore labeled PNA and quencher labeled masking sequence, we were able to monitor hybridization using fluorescence spectroscopy and demonstrate near quantitative release of masking sequences. We also demonstrate the specificity of this system using a scrambled release sequence, which results in minimal displacement. Visualization by transmission electron microscopy (TEM) validated the stimuli-responsive assembly of micelle-sized structures, consistent with our previously reported spontaneously assembling system. This novel system demonstrates the use of toehold mediated displacement for the development of a stimuli-responsive bilingual biopolymer assembly, greatly expanding the potential for use in biomedical and nanotechnology applications.

Results and Discussion

Design and synthesis of amphiphilic PNA. We previously designed and characterized a dodecameric PNA amphiphile (PNA-A1) complementary to a portion of the oncogenic miRNA-21.[12] The sequence was modified at the C-terminus to include 4-dimethylamino-naphthalimide (4-DMN), a fluorophore which displays a solvatochromic shift when in a hydrophobic environment, such as the hydrophobic region generated upon micellar assembly.[9, 13] PNA-A1 displayed spontaneous self-assembly, and the introduction of a complementary DNA or RNA was associated with a disappearance of assemblies attributable to sequence specific hybridization. Based on these results, we recognized the potential to modify the complementary DNA or RNA sequence to include a toehold for subsequent displacement and stimuli responsive PNA assembly. Toehold-mediated displacement systems function by hybridizing an oligonucleotide to a masking sequence having an unpaired toehold region. A fully complementary releasing sequence can then bind to the toehold and displace the target oligonucleotide.[14, 15] In this context, we recognized the potential to temporarily mask the amphiphilicity of the PNA sequence, and restore this upon displacement, thus driving assembly.

Table 1 Oligonucleotide sequences used for stimuli-responsive assembly system

Strand	Sequence
MS-R	$^{5'}$UAGCUUAUCAGACUGAUGUUGA/BHQ1$^{3'}$
MS-D	$^{5'}$TAGCTTATCAGACTGATGTTGA/BHQ1$^{3'}$
MS-D-m1	$^{5'}$TAGCTTATCAGACTGATGTTGT/BHQ1$^{3'}$
PNA-C1-FAM	^{C}CTGACTACAACT /FAM^{N}
PNA-A1-FAM	^{C}CT_AGAC_ATACA_KACT_K/FAM^{N}
RS-R	$^{3'}$AUCAAAUAGUCUGACUACAACU$^{5'}$
RS-D	$^{3'}$ATCAAATAGTCTGACTACAACT$^{5'}$
RS-D-m1	$^{3'}$ATCAAATAGTCTGACTACAACA$^{5'}$
RS-D-sc	$^{3'}$AATGAATTCTGAACTAAGCTGGTATGCG$^{5'}$

"FAM" denotes 5-carboxyfluorescein. "BHQ-1" denotes Black Hole Quencher 1. Subscripts denote the amino acid residues incorporated at the γ-position.

In this design, we envisioned a fluorophore-labeled sequence hybridized to a quencher-labeled complementary masking sequence to maximize signal gain upon displacement. Introduction of

the releasing sequence would then cause a change in hybridization state, separating the fluorophore-quencher pair and increasing fluorescence signal.[15, 16] This change in fluorescence signal provides a convenient output to monitor the fraction of PNA amphiphile remaining hybridized to the masking sequence.

In our previous PNA-A1 sequence, the fluorophore was located on the C-terminus such that it would be embedded in the hydrophobic micelle core upon assembly.[9] However, for the current system, moving the fluorophore to the N-terminus of the sequence would enable more efficient quenching in the toehold construct. Moreover, because the solvatochromic behavior of the dye was no longer needed, the 4-DMN was replaced with 5-carboxyfluoroscein (FAM), which is commonly used in fluorophore-quencher systems.[16, 17] In addition to the desired PNA-A1-FAM sequence, we also synthesized an unmodified control sequence (PNA-C1-FAM) in order to perform initial optimization of the toehold displacement construct (Table 1). Both PNA sequences were synthesized and characterized as previously reported.

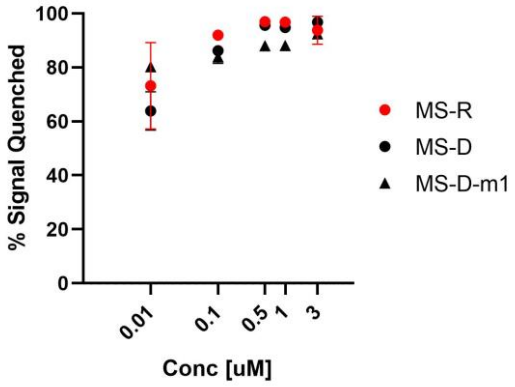


Fig. 2 Hybridization of MS-D, MS-R, and MS-D-1m to PNA-C1-FAM was monitored using fluorescence quenching. PNA and DNA sequences were at a ratio of 1:1. Error bars represent standard error.

Effect of masking sequence on quenching efficiency. We were interested to explore both DNA and RNA as potential masking sequences to disrupt PNA micelle assembly. To this end, we designed complementary sequences having a 10 nt toehold region and modified each with Black

Hole Quencher 1 (BHQ) on the 3' terminus. We investigated the quenching efficiency of DNA and RNA masking sequences MS-D and MS-R, respectively, on the fluorescent PNA-C1-FAM to determine an effective concentration for hybridization and quenching.[16, 18] We also aimed to explore the effect of hybridization region on toehold-mediated displacement through incorporation of a single mismatch at the 3' end of the DNA masking sequence (MS-D-m1).[15] By introducing a mismatch in the hybridization region, we anticipated that the thermodynamic preference for displacement by a fully complementary releasing strand would be increased, leading to more efficient triggering of assembly.[19] We found that the quenching efficiency of all three sequences was high and reached greater than 90% at concentrations above 500 nM (Fig. 2). Moving forward, we chose 3 μM as the working concentration for our PNA amphiphile and masking strands to remain an order of magnitude above the previously reported critical micelle concentration (CMC) of 317 nM.[9, 20, 21] Having optimized the concentration for duplex formation using PNA-C1-FAM, we then moved forward and replicated the hybridization of the amphiphilic sequence PNA-A1-FAM with MS-D and observed similar concentration dependence (Fig. S12).

Effect of releasing sequence on displacement efficiency. We envisioned that the release of the PNA strand from the masking sequence would be largely influenced by the strength of the duplex formed between the masking and releasing sequences.[15, 19] Thus, we tested displacement using both DNA and RNA releasing strands, RS-D and RS-R, respectfully, at several ratios relative to the PNA-C1-FAM:MS-D/R duplex system (Fig. 3A). Displacement was quantified by restoration of fluorescence upon addition of releasing sequence and calculated using Eq.2. In the PNA-C1-FAM:MS-D system, the addition of RS-D up to a 1:1 stoichiometric equivalence resulted in linear displacement, immediately restoring up to 85% of the fluorescence signal (Fig. 3B). We continued to monitor fluorescence for 1 hour after addition of the releasing sequence, and observed no significant increase in fluorescence, indicating that equilibrium for displacement is reached very rapidly. While the response to RS-R was also observed to be linear, we only

observed 20% displacement with one equivalent of releasing sequence, and additional incubation did not lead to an increase in fluorescence signal.

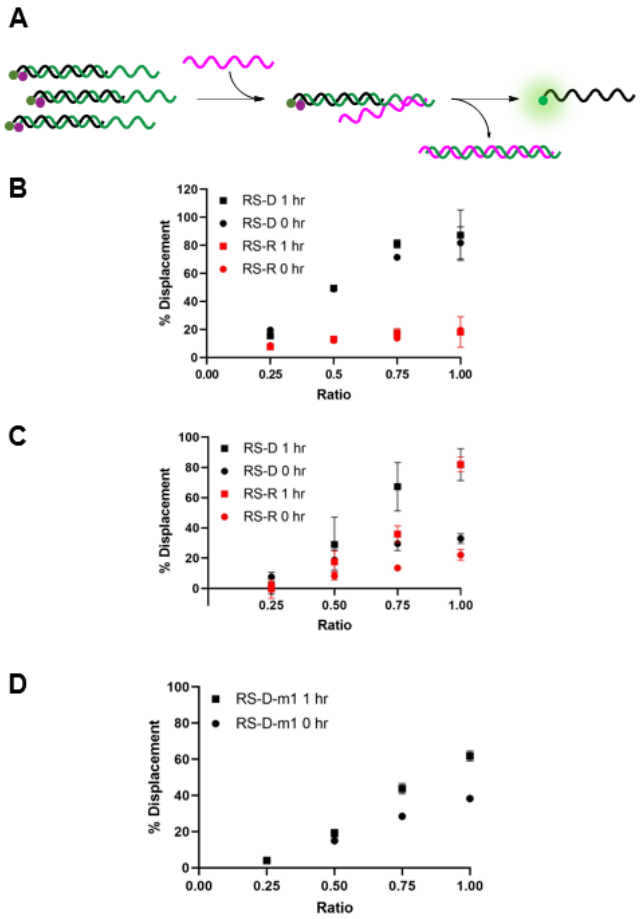


Fig. 3 (A) Toehold mediated displacement system using PNA-C1-FAM. When hybridized to the masking sequence, fluorescence is quenched. When a releasing sequence is introduced, the PNA is released, and fluorescence intensity increases. (B-D) % Displacement from PNA-C1-FAM:MS using RS-D, RS-R, and RS-D-m1, respectively, was evaluated in relation to the equivalents of RS added. (B) MS-D. (C) MS-R. (D) MS-D-m1. PNA:MS duplex was used at 3 µM in 1x PBS. Error bars represent standard error (n=3).

Changing the masking strand to RNA to give a PNA-C1-FAM:MS-R duplex resulted in similar thermodynamic response upon addition of the releasing sequences, but with slower kinetics (Fig. 3C). Specifically, upon the introduction of RS-D and RS-R, similar levels of fluorescence recovery to the PNA-C1-FAM:MS-D system were achieved, however, an incubation period of 1 hour was required to obtain maximum displacement in both systems. Previous reports using fluorophore-

quenching and melting temperature studies show the thermodynamics of duplex stability are: RNA:RNA > DNA:DNA > hybrid duplexes, which explains the superior displacement capability of RS-R with MS-R compared to MS-D.[22, 23] Additionally, predictive models and other kinetics studies show a difference in the rate of displacement depending on nucleic acid backbone, especially in heterochiral strand displacement and duplex formation, with a report demonstrating that RNA releasing strands provide faster rates of displacement.[24, 25] Interestingly, our studies show the opposite trend for RNA releasing sequences, but this may be due to differences in experimental parameters such as concentration, stoichiometry, and buffer conditions.

We hypothesized that the thermodynamics for displacement of the PNA-C1-FAM from the masking sequence would be driven by the relative stability of the PNA-MS and RS-MS duplexes. Seeking to harness this principle to obtain higher levels of displacement, we introduced a single-base mismatch at the 3' terminus of the masking sequence (MS-D-m1) and the complementary releasing sequence (RS-D-m1) such that they retain full complementarity with each other, but hybridization between the PNA and the masking sequence is reduced (Fig. 3D). We chose to locate the mismatch at a terminal site rather than an internal site for two reasons. First, internal mismatches in PNA can have a significant impact on duplex stability and we recognized that this would likely prevent efficient hybridization at room temperature. Second, the primary role of the masking sequence is to shield the hydrophobic block of the PNA sequence, and thus we sought to locate the mismatch as far away as possible from this region of the duplex. Unexpectedly, we observed a decrease in displacement yield and kinetics for the mismatched sequences compared to the fully complementary sequences, indicating that the mechanism for PNA displacement may not be completely driven by relative duplex stability.[15] While somewhat unexpected, similar results have been observed by others when investigating toehold systems.[19, 25] Moreover, our melting temperature studies showed that the presence of the mismatch surprisingly increased the duplex stability with PNA-C1-FAM (Table S5). Given these results, we decided to focus the

following work with PNA-A1-FAM on the MS-D system because it yielded the fastest and most complete displacement of the systems tested.

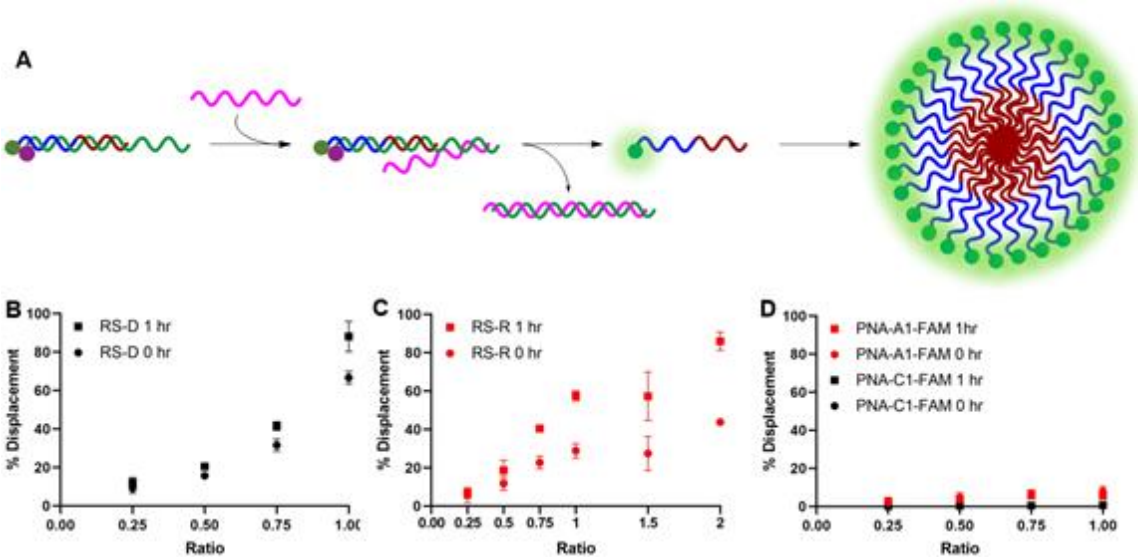


Fig. 4 (A) Toehold mediated displacement system using PNA-A1-FAM. When hybridized to the masking sequence, fluorescence is quenched. When a releasing sequence is introduced, the PNA is released, and fluorescence intensity increases. Upon release, the amphiphile can then self-assemble. (B) % Displacement from PNA-A1-FAM:MS-D using RS-D was evaluated in relation to the equivalents of RS-D added. (C) % Displacement from PNA-A1-FAM:MS-D using RS-R was evaluated in relation to the equivalents of RS-R added. (D) % Displacement from PNA-C1-FAM:MS-D and PNA-A1-FAM:MS-D using RS-D-sc was evaluated in relation to the equivalents of RS-D-sc added. PNA:MS-D duplex was used at 3 μM in 1x PBS. Error bars represent standard error (n=3).

After confirming and optimizing displacement using the PNA-C1-FAM duplex systems, we next tested the response of PNA-A1-FAM:MS-D using RS-R and RS-D (Fig. 4A). Ratiometric addition of RS-D resulted in a similar concentration-dependent response to that observed with the control PNA sequence, with approximately 60% immediate displacement of the amphiphile at a 1:1 ratio, which increased to 90% following a one-hour incubation (Fig. 4B). Addition of RS-R to the amphiphile masking complex resulted in much lower displacements of 30% at a 1:1 ratio immediately after addition, and this increased to 60% following incubation. While lower than the displacement observed using MS-D, this did compare favorably with the activity of the PNA-C1-FAM control with MS-R. Encouraged by the increased displacement, we tested higher ratios of RS-R up to 2:1 and observed 45% displacement immediately after addition of releasing sequence, and an increase to 90% after 1 hour of incubation (Fig. 4C). These results demonstrate that the

presence of γ-modifications or the amphiphilic nature of the PNA can alter both the thermodynamics and kinetics of toehold-mediated displacement. Moreover, our results demonstrate that both DNA and RNA can serve as effective stimuli to trigger displacement and induce assembly.

Finally, both control and amphiphilic systems were tested for sequence specificity through the introduction of a scrambled DNA releasing sequence, RS-D-sc. This sequence retains partial sequence complementarity but introduces several mismatches. Introduction of RS-D-sc under various conditions resulted in minimal (<10%) displacement even at a 1:1 ratio, indicating that the masked PNA is specific for its target sequence (Fig. 4D).

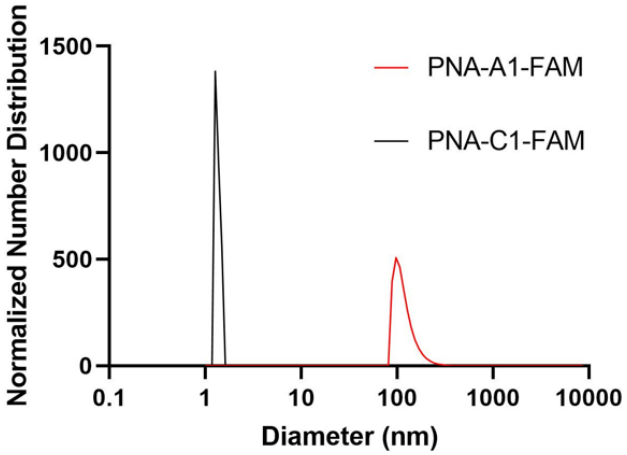


Fig. 5 Normalized size distribution of PNA assemblies. Samples tested at 200 μM in 1x PBS. Average diameter of particles of PNA-C1-FAM = 1.5 ± 0.3 nm; PNA-A1-FAM = 119.1 ± 34.0 nm. Peaks having greater than 1% are reported.

Characterization of PNA assembly. Upon successful toehold-mediated displacement of both control and amphiphilic PNA sequences, we sought to evaluate their assembly and hybridization properties using dynamic light scattering (DLS), circular dichroism (CD), and transmission electron microscopy (TEM). DLS analysis showed that PNA-C1-FAM and PNA-A1-FAM assemblies were similar in size to the previously published PNA-C and PNA-A structures.[9] The control sequence showed a size that would be expected for an unassembled monomer and the

amphiphilic sequence formed structures of 119.1 ± 34.0 nm, comparable to the size of 110.2 ± 31.9 nm of PNA-A (Fig. 5).[26, 27] This indicates that neither the fluorophore nor its location affected the formation and size of assemblies.

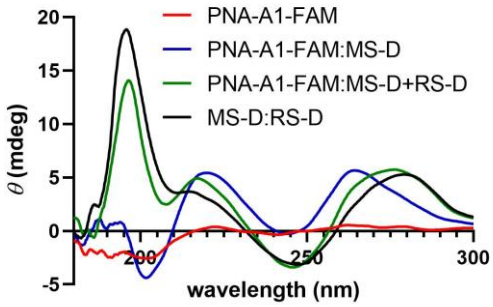


Fig. 6 CD spectroscopy demonstrating the change in maxima and minima upon the addition of RS-D. All samples were prepared using 100 µM PNA and DNA in 1x PBS.

In an attempt to validate hybridization of the masking sequence and subsequent disassembly, PNA-A1-FAM was annealed to MS-D at 200 µM under previously described conditions and the complex analyzed by DLS. While we theoretically anticipated observing a shift in size back to that of PNA-C1-FAM, we instead observed the formation of larger aggregates (Fig. S13). This was not entirely unexpected, however, as our previous studies as well as the TEM data shown below for this work demonstrate that large amorphous aggregates tend to form when DNA or RNA are added to PNA in the presence of buffer salts.[9] As anticipated, these aggregates persisted upon addition of RS-D to the PNA-A1-FAM:MS-D complex (Fig. S13). In DLS, the presence of large aggregates can overwhelm the signal and prevent detection of smaller assemblies, and thus we viewed the DLS data as being inconclusive as to whether we were achieving our goal of hybridization and stimuli-responsive assembly.

We next turned to CD to provide validation for the changes in hybridization upon addition of the masking and releasing sequences. Despite being chiral, PNA-A1-FAM displayed only a weak CD signal, in line with previous reports (Fig. 6).[9] Upon hybridization with MS-D, we observed a strong signal with maxima at 219 nm and 262 nm and with minima at 244 nm and 201 nm, in

similar agreement with previously reported right-handed PNA:DNA duplex structure.[5, 9, 28, 29] Importantly, this signal is distinct from MS-D alone (Fig. S11). After introduction of the RS-D strand to the PNA-A1-FAM:MS-D duplex, we observed a shift in signal, indicating that the PNA amphiphile was released. Additionally, the signal observed after addition of the releasing sequence is very similar to a combination of signals from PNA alone and the MS-D:RS-D duplex, providing additional evidence for successful toehold-mediated displacement.

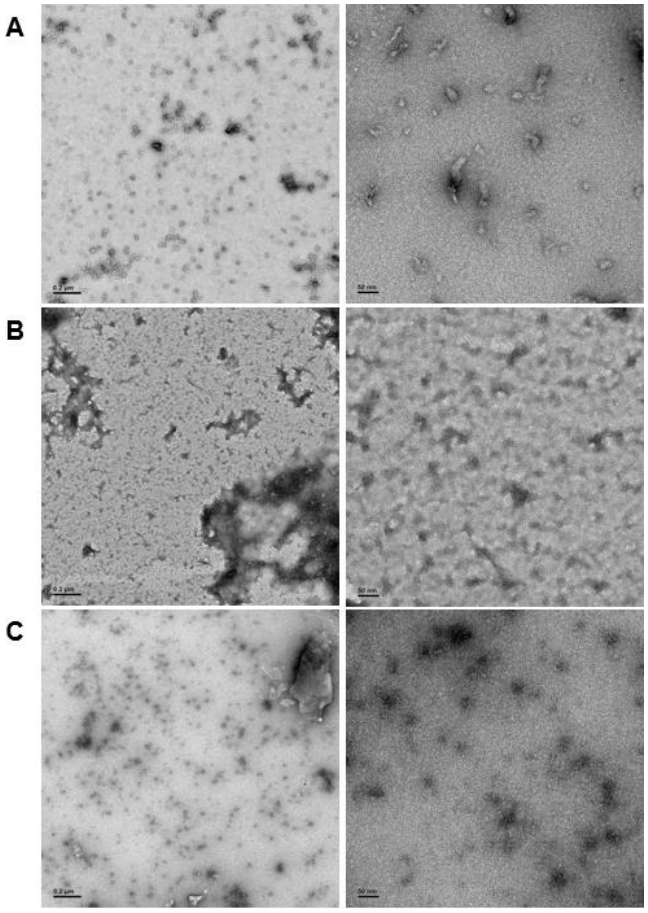


Fig. 7 TEM images of (A) PNA-A1-FAM assembled at 10 μM, (B) PNA-A1-FAM disassembled via MS-D hybridization at 10 μM, and (C) PNA-A1-FAM reassembled via RS-D toehold release at 10 μM. Left column scale bar = 2000 nm. Right column scale bar = 50 nm.

In order to demonstrate DNA-mediated disassembly and stimuli-responsive reassembly, we turned to transmission electron microscopy (TEM). In comparison to DLS, where non-specific aggregates can block detection of smaller assemblies, TEM enables us to visualize all of these

objects, and thus observe PNA assembly even in the presence of other aggregates. Assembly of PNA-A1-FAM was tested by preparing a 10 μM or 100 μM solution, spotting onto the grid, and staining with uranyl acetate for imaging. We observed relatively uniform, micelle-like structures having similar size and morphology compared to previously reported PNA-A (Fig. 7A and S8). This was to be expected, given that both of these concentrations are above the CMC value for the PNA amphiphile. Moreover, the ability of PNA-A1-FAM to self-assemble similarly to PNA-A demonstrates that the amino acid sequence directs assembly into micelle-like structures regardless of the fluorophore and its location.

We then tested disassembly through nucleic acid recognition of the masking sequence. The PNA-A1-FAM:MS-D duplex system was initially prepared at 100 μM, annealed as described, and spotted for imaging (Fig. S9). However, in line with our previous experience and our DLS data, we observed a very high level of non-specific aggregation, with the dark nature of these aggregates attributable to interactions between the DNA phosphate backbone and the uranyl acetate stain. Given our goal of eventually observing micelle re-assembly, we sought to minimize the presence of these non-specific aggregates and diluted the solutions to 10 μM. This reduced the quantity and size of dark aggregates, likely due to the decrease in DNA concentration. Importantly, both the high and low concentration samples show the disappearance of the ordered PNA assemblies, matching our previous report that sequence-specific recognition can induce disassembly (Fig. 7B).[9]

Finally, to validate stimuli-responsive assembly, we combined PNA-A1-FAM:MS-D and RS-D, incubated for 1 hour, spotted the samples onto a TEM grid, stained, and imaged. Using a concentration of 100 μM again resulted in dark amorphous aggregates which we attribute to precipitation (Fig. S10). However, this was reduced by lowering the concentrations to 10 μM, similar to the experiments described above. While we do still observe some non-specific aggregates, we were excited to also observe a large number of ordered assemblies having an

average size of 16.8 ± 5.7 nm (Fig. 7C). We do note that the structures formed upon reassembly are slightly smaller than those formed in the initial PNA assembly. This may be due to the presence of DNA in the solutions or as a result of a different assembly mechanism when the PNA is being released from a DNA complement over time. However, we were excited to observe that our bilingual PNA is able to undergo stimuli-responsive assembly as anticipated.

Conclusion

Biological information codes are extremely powerful for information storage and function, and we have shown that both peptide and nucleic acid codes can be combined into a single bilingual biopolymer using a PNA scaffold. Our initial work focused on harnessing the peptide code for assembly and the nucleic acid code for stimuli-responsive disassembly. Here we explore the extension of the nucleic acid code to also enable stimuli-responsive assembly. This is achieved through the design of a toehold-mediated displacement system in which a masking sequence hybridizes to the PNA to induce disassembly while providing a single-stranded toehold. Subsequent addition of a complementary nucleic acid results in displacement to release the PNA and trigger reassembly. Using fluorescence, we were able to monitor hybridization and subsequent displacement. We observed varying thermodynamics and kinetics for displacement when using different combinations of DNA and RNA sequences, and observed that using DNA as both the masking and releasing sequence provided highly efficient and rapid displacement. The assembled, disassembled, and reassembled structures were characterized by DLS and TEM. The presence of non-specific aggregates induced by the presence of DNA or RNA prevented direct observation of the reassembly process using DLS, however, this challenge was overcome using TEM and we were excited to observe the stimuli-responsive formation of micelle-sized assemblies. Together, this work provides a new approach to harnessing biological information codes to control assembly, providing stimuli-responsive materials that can advance applications in biotechnology and biomedicine.

Experimental Procedures:

Abbreviations: Fmoc, fluorenylmethyloxycarbonyl; FAM, fluorescein; Acetonitrile, ACN; Water, H_2O

PNA monomer synthesis. Unmodified PNA monomers were purchased from PolyOrg, Inc. Thymine acetic acid was purchased from Sigma-Aldrich. The protected nucleobases adenine and cytosine and Fmoc protected γ-methyl and γ-lysyl modified PNA backbones were synthesized following previously published procedures.[9, 30] The Fmoc-protected L-amino acids and nucleobase starting materials were purchased from Chem-Impex Inc. All other reagents were purchased from Fisher Scientific, Chem-Impex, and Sigma-Aldrich and used without additional purification unless otherwise stated. Merck silica gel 60 F254 was used to monitor small molecule synthesis through thin layer chromatography (TLC) with UV light for visualization. Synthesized compounds were purified using flash chromatography with SiliaFlash F60 grade silica purchased from SiliCycle Inc. Characterization of compounds was completed using proton (^{1}H) and carbon (^{13}C) NMR using a Varian Inova 400 MHz spectrometer. The spectra were analyzed using MestReNova Software. The mass of the compounds was confirmed using an Agilent 6230 electrospray ionization time-of-flight (ESI-TOF) mass spectrometer. PNA monomer synthesis was completed as described previously.[9, 31]

PNA oligomer synthesis. PNA oligomers were synthesized as previously described using a Biotage SP Wave semiautomatic peptide synthesizer.[9] Monomer coupling efficiencies were monitored using a Nanodrop 2000 spectrophotometer to measure dibenzofulvene-piperidine adduct absorbance at 301 nm. Upon cleavage, sequences were purified through reverse-phase HPLC using an Agilent Eclipse XDB-C18 5 µm, 9.4 x 250 mm column at 60 ºC with a flow rate of 2 mL/min, monitoring 260 and 495 nm using a linear gradient (10-40%) of 0.1% TFA/CAN in 0.1% TFA/H_2O. The sequences were confirmed by ESI-TOF mass spectrometry.

Melting temperature analysis of PNA strands. All DNA and RNA sequences were purchased from the University of Utah DNA/Peptide Synthesis Core Facility or Integrated DNA Technologies. Samples containing 5 µM of complementary nucleic acids were prepared using stock solutions in 1x PBS. Samples were then added to an 8-cell quartz microcuvette with a 1 cm path length. A Shimadzu UV-1800 spectrophotometer equipped with a temperature controller and Julabo Corio CD water circulator were used for measurement of absorbance at 260 nm while heating from 25 to 95 ºC at a rate of 0.5 ºC/min. The melting temperatures were calculated using a first derivative method and experiments were carried out in triplicate.

Fluorescence monitoring of hybridization and displacement. The fluorescence quenching of the FAM-labeled PNA sequences in the PNA:MS-R or MS-D duplexes was measured by fluorescence emission of the FAM fluorophore. Samples of 100 µL containing PNA alone, as a fluorescent control sample, and PNA:MS-R or MS-D over a range of concentrations (0.01 µM to 3 µM), at a 1:1 ratio were prepared using stock solutions in 1x PBS. The samples were annealed by heating to 95 ºC for 5 mins and cooling 1 ºC every 30 seconds until 25 ºC was achieved. The samples were then transferred into a 384-well flat-bottom plate for analysis on a Biotek Cytation5 plate reader. Fluorescence was then measured using an excitation at 495 nm and emission at 520 nm. The fluorescence measurement was used to plot the % Quenched using Eq. 1 using GraphPad Prism software.

$$\% \text{ Quenched} = \frac{(F_m - F_0)}{(F_m)} \times 100 \qquad (1)$$

To measure the % Displacement via fluorescence intensity, 50 µL of samples containing 6 µM of PNA and MS-D or MS-R each in 1x PBS and 50 µL of RS-R or RS-D samples at 2X final reported quantities (0.25 to 2 equivalents) in 1x PBS were annealed as previously described. The 50 µL RS-R or RS-D samples were then transferred into the PNA:MS-D or MS-R samples to achieve the final reported concentrations. Samples containing PNA alone and PNA:MS-D or MS-R were used and had 1x PBS blank solutions added as controls. The samples were then transferred to a

384-well flat-bottom plate for analysis using the previously described protocol. The samples were measured at reported incubation intervals. The fluorescence was then used to calculate the % Displacement using Eq. 2 as described and plotted using GraphPad Prism software.

$$\% \text{ Displacement} = \frac{(F - F_0)}{(F_m - F_0)} \times 100 \qquad (2)$$

F_m is the fluorescently labeled PNA sequence alone, F_0 is the fluorescence when PNA is hybridized to a quencher-labeled masking sequence, and F is the fluorescence measured when inducing a stimuli-response.

Preparation of PNA amphiphile assemblies. To prepare amphiphile assembly, samples of PNA, PNA:MS-D, and RS-D were dissolved in water or 1x PBS and were annealed as previously described. All samples were allowed to incubate at room temperature for 1 hour. Samples containing RS-D were then transferred to PNA:MS-D samples and allowed to incubate at room temperature for an additional hour.

Determination of size using dynamic light scattering. Samples of PNA, PNA:MS-D, and PNA:MS-D were prepared in 1x PBS at 200 µM from stock solutions and were annealed through heating and slow cooling as described above. Samples were then transferred to a quartz cuvette and analyzed using a Particular Systems NanoPlus DLS nano particle analyzer and data analyzed using NanoPlus Software.

Confirmation of MS-R hybridizing to PNA-A by CD. Circular Dichroism was used to confirm the hybridization of PNA-A:MS-R. Individual samples of PNA-A and MS-R were prepared at 200 µM in 1x PBS. 30 µL of PNA-A was combined with 30 µL of MS-R to form the final 100 µM in 1x PBS solutions. Samples were incubated for 1 hour, after which, they were analyzed using a JASCO J-1500 circular dichroism spectrometer. Data points from 190 to 300 nm were collected at 1 nm intervals using a continuous scanning mode of 200 nm/min at 23.5 ºC. A sample with PNA-A alone was used as a control and had the signal overlapped using the JASCO software.

Confirmation of RS-D hybridizing to MS-D in the PNA-A:MS-D duplex state by CD. Samples containing PNA-A:MS-R and MS-D were prepared at 200 µM in 1x PBS and were annealed through heating and slow cooling as described above. 30 µL of PNA-A:MS-R was combined with 30 µL of RS-D to form the final 100 µM in 1x PBS solutions and incubated for 1 hour at room temperature before being analyzed. Samples containing MS-R:RS-D and RS-D alone were used as controls.

Characterization of PNA-A assemblies by TEM. Samples of PNA-A1-FAM were prepared at 100 µM in water from stock solutions. Using a 200-mesh Formvar/carbon-coated copper grid, 3.5 µL of sample was spotted for 2 minutes and was wicked away with filter paper. Then, 3.5 µL of 1% uranyl acetate stain solution was spotted for 30 seconds before wicking. Grids were dried at room temperature for 30 minutes prior to imaging using a JEOL JEM-1400 transmission electron microscope.

Characterization of PNA-A disassembled using masking sequence DNA by TEM. Samples of PNA-A1-FAM:MS-D at 100 µM were prepared in 1x PBS and annealed through heating and slow cooling as described above. The duplex sample was then spotted and prepared for TEM imaging as described above.

Characterization of PNA-A re-assembled using release sequence DNA by TEM. Samples containing PNA-A1-FAM:MS-D and RS-D at 200 µM were prepared in 1x PBS and annealed through heating and slow cooling as described above. After, 20 µL of RS-D was transferred to 20 µL of PNA-A1-FAM:MS-D and allowed to incubate at room temperature for 1 hour. The sample was then prepared for TEM imaging as described above.

Supporting Information:

Characterization of PNA sequences, T_m data, TEM images, and additional displacement studies

Conflicts of Interest:

The authors report no conflicting interests.

Acknowledgement:

This work was supported by the National Science Foundation (DMR 2003987 to J.M.H.). The authors also acknowledge the Robert P. Apkarian Integrated Electron Microscopy Core and NMR Research Center at Emory University for instrument access and technical assistance. They would also like to thank Mike Hanson and the oligonucleotide and peptide synthesis facility and the University of Utah for oligonucleotide materials.

References:

1. Tan, X., et al., *Nucleic acid-based drug delivery strategies.* Journal of Controlled Release, 2020. **323**: p. 240-252.
2. Wang, L., et al., *Controlling the self-assembly of biomolecules into functional nanomaterials through internal interactions and external stimulations: A review.* Nanomaterials, 2019. **9**(2): p. 285.
3. Peterson, A.M. and J.M. Heemstra, *Controlling self-assembly of DNA-polymer conjugates for applications in imaging and drug delivery.* Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology, 2015. **7**(3): p. 282-297.
4. Egholm, M., et al., *Peptide nucleic acids (PNA). Oligonucleotide analogs with an achiral peptide backbone.* Journal of the American Chemical Society, 1992. **114**(5): p. 1895-1897.
5. Moccia, M., M.F. Adamo, and M. Saviano, *Insights on chiral, backbone modified peptide nucleic acids: properties and biological activity.* Artificial DNA: PNA & XNA, 2014. **5**(3): p. e1107176.
6. Shakeel, S., S. Karim, and A. Ali, *Peptide nucleic acid (PNA)—a review.* Journal of Chemical Technology & Biotechnology: International Research in Process, Environmental & Clean Technology, 2006. **81**(6): p. 892-899.

7. Egholm, M., et al., *PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules.* Nature, 1993. **365**(6446): p. 566-568.

8. Sahu, B., et al., *Synthesis of conformationally preorganized and cell-permeable guanidine-based γ-peptide nucleic acids (γGPNAs).* The Journal of organic chemistry, 2009. **74**(4): p. 1509-1516.

9. Swenson, C.S., et al., *Bilingual peptide nucleic acids: encoding the languages of nucleic acids and proteins in a single self-assembling biopolymer.* Journal of the American Chemical Society, 2019. **141**(48): p. 19038-19047.

10. Peng, L., et al., *Engineering and applications of DNA-grafted polymer materials.* Chemical science, 2013. **4**(5): p. 1928-1938.

11. Xu, H., et al., *Hydrophobic-region-induced transitions in self-assembled peptide nanostructures.* Langmuir, 2009. **25**(7): p. 4115-4123.

12. Kumarswamy, R., I. Volkmann, and T. Thum, *Regulation and function of miRNA-21 in health and disease.* RNA biology, 2011. **8**(5): p. 706-713.

13. Ikeda, H., Y. Nakamura, and I. Saito, *Synthesis and characterization of naphthalimide-containing peptide nucleic acid.* Tetrahedron letters, 2002. **43**(32): p. 5525-5528.

14. Guo, Y., et al., *Recent advances in molecular machines based on toehold-mediated strand displacement reaction.* Quantitative Biology, 2017. **5**(1): p. 25-41.

15. Machinek, R.R., et al., *Programmable energy landscapes for kinetic control of DNA strand displacement.* Nature communications, 2014. **5**(1): p. 1-9.

16. Tan, Z., T.A. Feagin, and J.M. Heemstra, *Temporal control of aptamer biosensors using covalent self-caging to shift equilibrium.* Journal of the American Chemical Society, 2016. **138**(20): p. 6328-6331.

17. Chen, J., et al., *A simple and rapid biosensor for ochratoxin A based on a structure-switching signaling aptamer.* Food Control, 2012. **25**(2): p. 555-560.

18. Sanford, A.A., et al., *RE-SELEX: Restriction Enzyme-Based Evolution of Structure-Switching Aptamer Biosensors.* 2021.

19. Haley, N.E., et al., *Design of hidden thermodynamic driving for non-equilibrium systems via mismatch elimination during DNA strand displacement.* Nature communications, 2020. **11**(1): p. 1-11.

20. Vernille, J.P., L.C. Kovell, and J.W. Schneider, *Peptide nucleic acid (PNA) amphiphiles: synthesis, self-assembly, and duplex stability.* Bioconjugate chemistry, 2004. **15**(6): p. 1314-1321.

21. Lau, C., et al., *Morphological characterization of self-assembled peptide nucleic acid amphiphiles.* The Journal of Physical Chemistry B, 2006. **110**(18): p. 9027-9033.

22. Li, M.-X., et al., *Exploration of the kinetics of toehold-mediated strand displacement via plasmon rulers.* ACS nano, 2018. **12**(4): p. 3341-3350.

23. Lesnik, E.A. and S.M. Freier, *Relative thermodynamic stability of DNA, RNA, and DNA: RNA hybrid duplexes: relationship with base composition and structure.* Biochemistry, 1995. **34**(34): p. 10807-10815.

24. Srinivas, N., et al., *On the biophysics and kinetics of toehold-mediated DNA strand displacement.* Nucleic acids research, 2013. **41**(22): p. 10641-10658.

25. Kundu, N., B.E. Young, and J.T. Sczepanski, *Kinetics of heterochiral strand displacement from PNA–DNA heteroduplexes.* Nucleic Acids Research, 2021. **49**(11): p. 6114-6127.

26. Souza, T.G., V.S. Ciminelli, and N.D.S. Mohallem. *A comparison of TEM and DLS methods to characterize size distribution of ceramic nanoparticles*. in *Journal of physics: conference series*. 2016. IOP Publishing.

27. Domingos, R.F., et al., *Characterizing manufactured nanoparticles in the environment: multimethod determination of particle sizes.* Environmental science & technology, 2009. **43**(19): p. 7277-7284.

28. Dragulescu-Andrasi, A., et al., *A simple γ-backbone modification preorganizes peptide nucleic acid into a helical structure.* Journal of the American Chemical Society, 2006. **128**(31): p. 10258-10267.

29. Faccini, A., et al., *Circular dichroism study of DNA binding by a potential anticancer peptide nucleic acid targeted against the MYCN oncogene.* Chirality: The Pharmacological, Biological, and Chemical Consequences of Molecular Asymmetry, 2008. **20**(3-4): p. 494-500.

30. De Costa, N.T.S. and J.M. Heemstra, *Evaluating the effect of ionic strength on duplex stability for PNA having negatively or positively charged side chains.* PloS one, 2013. **8**(3): p. e58670.

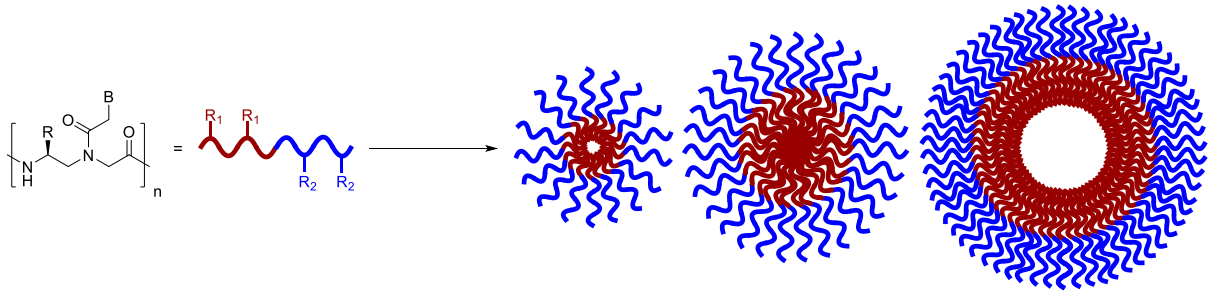
31. Porcheddu, A., et al., *A practical and efficient approach to PNA monomers compatible with Fmoc-mediated solid-phase synthesis protocols*. 2008, WILEY-VCH Verlag Weinheim.

Chapter 3: Elucidating Structure-Assembly Relationships for Bilingual PNA Biopolymers

Hector Argueta-Gonzalez, Colin S. Swenson, Jennifer M. Heemstra

Abstract:

Nucleic acids and proteins each possess an encoded "language" that can be used for information storage or to direct function. However, each biopolymer is limited to encoding its respective “language.” Using a peptide nucleic acid (PNA) scaffold, both nucleobase and amino acid residues can be installed on the backbone, enabling a single polymer to encode for both languages. Our lab previously reported the development of a “bilingual” PNA biopolymer that incorporated both hydrophobic and hydrophilic amino acid residues in the form of alanine and lysine side chains along with a sequence-specific nucleic acid code. We found that the amphiphilic residues directed the biopolymer to undergo self-assembly into micelle-like structures, while nucleic acid recognition could be harnessed for disassembly. Herein, we synthesize and study a series of bilingual PNA sequences having varying side chain charge, hydrophobicity, and spacing in order to elucidate the effect of these parameters on micelle assembly and nucleic acid recognition. By exchanging positive charge for negative in the hydrophilic block or increasing the bulkiness of the hydrophobic side chains, we observed assembly into similarly sized micelles, with the negative charge leading to an increase in CMC. Upon nucleic acid sequence truncation to decrease spacing between side chains, the biopolymers remained capable of self-assembling but formed smaller structures. Characterization of disassembly revealed that each variant retained sequence recognition capabilities and was able to undergo stimuli-responsive disassembly. Together, these data show that both the amino acid and nucleic acid sequences of bilingual biopolymers can be customized to meet the needs of different applications or modulate the properties of assembly.



Introduction

Nature primarily uses two biopolymer structures – nucleic acids and proteins – to encode information and transmit that information into structure and function. Nucleic acids are adept at storing genetic information and performing sequence recognition to regulate biological function. While this system is privileged by predictable molecular recognition properties in the form of Watson-Crick-Franklin base pairing, the relative lack of chemical diversity in backbone and nucleobase components limits its potential functions.[1-3] In contrast, peptides and proteins are decorated with amino acid side chains having a diverse range of chemical structures and functionalities, which leads to a greater diversity of functions, but makes folding and molecular recognition less predictable. Because of their uniquely adaptable capabilities, both biopolymers show great promise in applications outside their canonical biological roles.[4-6] We hypothesized that a “bilingual” biopolymer capable of integrating both the sequence-recognition properties of nucleic acids and the structure and function-defining properties of peptides in a cohesive manner would open new avenues for creating programmable materials.

Researchers have described chimeric polymers containing both nucleic acid and peptide blocks, but the separation of these domains limits the overall function of the polymers.[3, 7, 8] We envisioned that peptide nucleic acid (PNA) would provide an ideal scaffold for simultaneously encoding both nucleic acid and protein sequences.[9, 10] The modifiable pseudopeptide backbone of PNA allows for the precise incorporation of both nucleobase and amino acid

residues, and PNA confers additional advantages such as an increase in hydrolase resistance and increased duplex stability.[3, 11-14] Through the incorporation of amino acid residues at the γ-position along the PNA backbone, we designed and synthesized of a “bilingual” PNA biopolymer.[9] Specifically, we incorporated hydrophilic and hydrophobic residues in the form of lysine and alanine to generate an amphiphilic sequence capable of self-assembly into spherical micelle-like structures. Upon the introduction of a complementary nucleic acid sequence, the sequence recognition event leads to a change in the electrostatic properties and induces a stimuli-responsive disassembly. In this way, we are able to utilize the peptide code for assembly and the nucleic acid code for disassembly. Further exploration of this system enabled us to design a masking system that renders the PNA amphiphile in an initially disassembled state, enabling stimuli-responsive assembly in the presence of a target nucleic acid.[10] Recognizing the potential use of these amphiphilic bilingual biopolymers for applications such as drug delivery and biosensing, we were curious to explore the design rules for assembly and function.[15]

Herein, we describe the design, synthesis, and characterization of a series of amphiphilic PNA sequences that vary systematically in amino acid side chain identity and placement, enabling us to explore the parameters of charge, hydrophobicity, and sequence length (Figure 1A). We demonstrate that altering the specific structure of either the hydrophilic or hydrophobic residues significantly alters the micellar architecture as observed by TEM and DLS (Figure 1B). Variation in sequence length plays a large role in determining the size of the assemblies, as shorter sequences produced smaller assemblies (Figure 1C). Finally, our studies confirmed that all of the bilingual biopolymer sequences remained capable of undergoing stimuli-responsive disassembly upon the introduction of a complementary nucleic acid sequence. These results demonstrate that the amino acid code can be manipulated in order to generate assemblies having diverse physiochemical properties while retaining the capacity to utilize peptide and nucleic acid codes for self-assembly and stimuli-responsive disassembly, respectively. Moreover, we demonstrate

that materials properties such as assembly size can be programmed by altering the structure of the bilingual biopolymer. Together, these results advance our understanding of this new class of materials and provide key information for advancing the use of PNA amphiphiles in biological and nanotechnology applications.

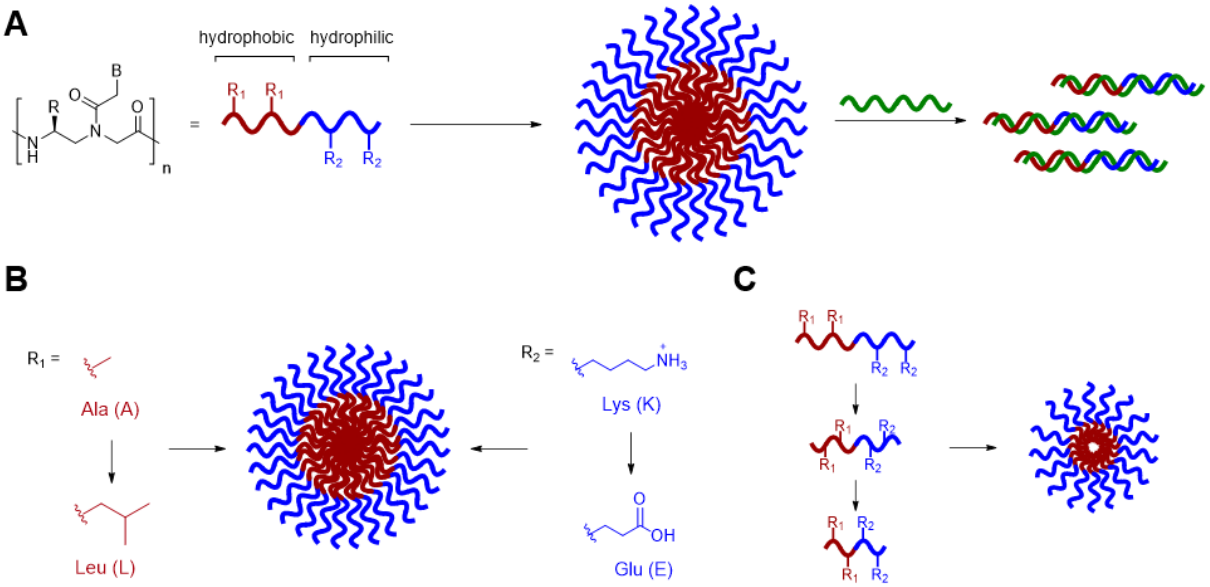


Figure 1. (A) PNA bilingual biopolymers contain hydrophobic and hydrophilic amino acid residues at the γ-position to direct self-assembly. Upon the introduction of a complementary sequence, nucleobase recognition induces stimuli-responsive disassembly (B) Changes to the amino acid code (R_1, R_2) do not significantly impact assembly. (C) Nucleic acid sequence truncation allows for control over assembly size.

Results and Discussion

Design and synthesis of amphiphilic PNA sequences. Previously, our lab designed and characterized PNA-A, a dodecameric PNA sequence that contains two alanine and two lysine residues that create the hydrophobic and hydrophilic blocks, respectively (Table 1).[9, 10] We recognized that the application of amphiphilic bilingual biopolymers would be advanced by elucidating the impact of the amino acid side chain identity and spacing on assembly and disassembly properties. To test the influence of a larger hydrophobic side chain, PNA-B-A was synthesized by replacing alanine with leucine.[16-18] To determine the influence of charge in the hydrophilic block, PNA-N was synthesized having glutamic acid incorporated to switch from

positive to negative charge while maintaining similar sidechain length to lysine.[19, 20] Finally, to test the influence of the side chain spacing and nucleic acid length on the biopolymer characteristics, decamer and octamer sequences PNA-10-A and PNA-8-A where synthesized having the same residue content as PNA-A but in a shorter polymer (Figure 1B).[21, 22] The 4-DMN PNA monomer was included in all sequences to evaluate aggregation and determine critical micelle concentrations (CMC) as previously reported.[9]

To generate the small library of sequences, we first synthesized the γ-modified PNA monomers as previously reported.[23, 24] Using Fmoc-protected L-amino acids, the PNA backbones containing lysine, alanine, leucine, and glutamic acid, respectively, were produced followed by coupling to protected nucleobase acetic acids using HATU in the presence of *N*,*N*-diisopropylethylamine (DIPEA) to produce the desired monomers. The synthesis of the 4-DMN PNA monomer was completed as previously reported.[9] Following the monomer synthesis, the amphiphilic and control PNA sequences were synthesized, purified, and characterized via mass spectrometry and HPLC as previously reported. The C-to-N terminal peptide synthesis produces

Table 1. Oligonucleotide sequences used to explore the impact of structure on self-assembly and stimuli-responsive disassembly.

Strand	Sequence
DNA	$^{5'}$T A G C T T A T C A G A C T G A T G T T G A$^{3'}$
CS-DNA	$^{3'}$A T C A A A T A G T C T G A C T A C A A C T$^{5'}$
PNA-C	^{C}Lys^{+} D C T G A C T A C A A C T^{N}
PNA-A	^{C}D C T$_A$ G A C$_A$ T A C A$_K$ A C T$_K$^{N}
PNA-B-A	^{C}D C T$_L$ G A C$_L$ T A C A$_K$ A C T$_K$^{N}
PNA-N-C	^{C}Gly C T G A C T A C A A C T D Glu^{-N*}
PNA-N-A	^{C}Gly C T$_E$ G A C$_E$ T A C A$_A$ A C T$_A$ D^{N*}
PNA-10-C	^{C}Lys^{+} D G A C T A C A A C T^{N}
PNA-10-A	^{C}D G A$_A$ C T$_A$ A C A$_K$ A C T$_K$^{N}
PNA-8-C	^{C}Lys^{+} D C T A C A A C T^{N}
PNA-8-A	^{C}D C T$_A$ A C$_A$ A A$_K$ C T$_K$^{N}

"D" denotes the 4-DMN dye monomer. * denotes acetyl N terminal. Subscript denotes the amino acid residue single letter code.

oligomers having an N-terminal positive charge, which is ideal for amphiphiles having a positively charged hydrophilic block but could interfere with our design of amphiphiles having a negatively charged hydrophilic block. Thus, in the case of PNA-N-A, we chose to reverse the ordering of the hydrophobic and hydrophilic residues, with the 4-DMN being placed at the N-terminus and an acetyl group used to cap the free amine.[25]

After obtaining the PNA sequences, UV melting temperature studies were performed to analyze hybridization to complementary DNA target sequences. To measure melting temperature, PNA and DNA were at combined in 1x phosphate buffered saline (1x PBS) at final concentrations of 3 μM each, and a heat-cool cycle performed to anneal the two strands. We previously observed that amphiphilic PNA sequences have higher melting temperatures than their control strand counterparts, as the introduction of γ-modifications is known to increase hybridization stability.[26, 27] Of note, in the case of PNA-N-A we observed a decrease in melting temperature from 58.1 to 55.7 °C as a result of adding the amino acid side chains. This may be due in part to the swapping of the hydrophilic and hydrophobic blocks in the sequence, but more likely results from the presence of the additional glutamic acid at the N-terminus and addition of a glycine at the C-terminus, which may reduce preorganization.[27] While electrostatic repulsion between the negative side chains and the negative DNA backbone may also play a role, previous studies have

found that at physiological ionic strength, the polarity of the side chains has little impact on hybridization strength with DNA and RNA.[24]

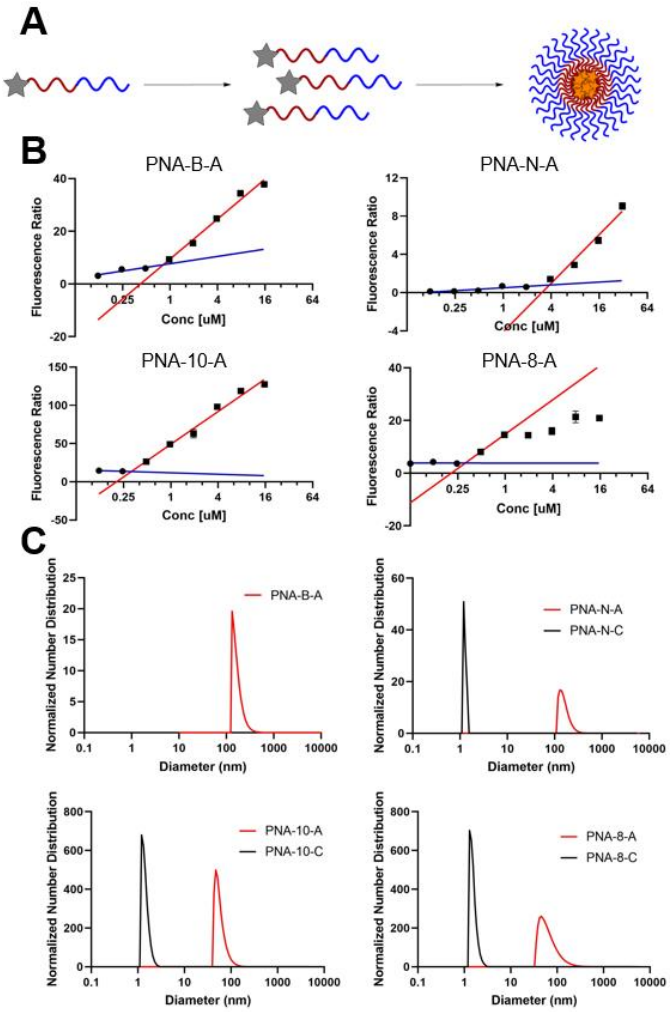


Figure 2. (A) Activation of solvatochromic 4-DMN dye upon PNA assembly. (B) CMC measurement using fluorescence ratio. Samples tested in water. (C) Normalized size distribution of PNA assemblies measured by DLS. Samples tested at 200 μM in 1x PBS.

Characterization of PNA assembly. In parallel with studying hybridization, we wanted to determine the effect of each structural change on the self-assembly properties of the bilingual biopolymers. In our previous study, we had strategically placed the 4-DMN dye such that it would become sequestered in a hydrophobic environment upon assembly, leading to increased fluorescence and enabling us to measure critical micelle concentration (CMC).[9, 28] By measuring the relative fluorescence of each amphiphilic sequence and unbound dye monomer in solution, we were able

to determine the CMC value for each PNA amphiphile (Figure 2A). Similar CMC values were

Table 2. Tabulated CMC values

Strand	CMC
PNA-A	317 nM
PNA-B-A	860 nM
PNA-N-A	3.58 μM
PNA-10-A	548 nM
PNA-8-A	523 nM

found for PNA-B-A, PNA-10-A, and PNA-8-A as compared to the previously reported PNA-A (Table 2).[9] This is likely due to the similarity in peptide sequence that drives assembly, having two hydrophobic and two positively charged residues. In the case of PNA-B-A, this was somewhat surprising, as we anticipated that the increased hydrophobicity of the leucine side chain would increase the driving force for assembly. However, any gain from increased hydrophobicity could be offset by greater steric hindrance to close packing of PNA strands, or it is possible that the positively charged block plays a more significant role in determining the thermodynamics of assembly. In comparison, PNA-N-A containing glutamic acid was found to have a significantly higher CMC than the other sequences, indicating that it has higher solubility and reduced driving

Table 3. Average Size of PNA Assembly Particles using DLS

Strand	Diameter
PNA-B-A	107.8 ± 47.0 nm
PNA-N-C	1.3 ± 0.1 nm
PNA-N-A	161.7 ± 44.4 nm
PNA-10-C	1.6 ± 0.3 nm
PNA-10-A	58.9 ± 17.8 nm
PNA-8-C	1.5 ± 0.3 nm
PNA-8-A	70.0 ± 43.1 nm

Sizes reported are greater than 1% in population

force for assembly. Importantly, these experiments demonstrate our ability to tune the assembly strength of the micelles by altering the side chain identity, enabling these materials to be tuned for varying biological applications. Experiments using the control sequences resulted in lower and delayed increases in fluorescence, similar to the previously reported PNA-C,[9] indicating that the amino acid residues are driving self-assembly.

After evaluating the CMC values of the amphiphiles, we next carried out dynamic light scattering (DLS) experiments to evaluate the size of the assemblies in solution. Samples were prepared using 200 μM of each PNA and then size distribution was measured (Figure 2C). Each amphiphile sequence produced assemblies having a measurable hydrodynamic diameter and PNA-B-A and PNA-N-A were found to be of similar size, 107.8 ± 47.0 nm and 161.7 ± 44.4 nm, respectively. The truncated sequences PNA-10-A and PNA-8-A were found to form smaller assemblies, measuring 58.9 ± 17.8 nm and 70 ± 43.1 nm, respectively (Table 3). This suggests that while the peptide sequence determines assembly strength, the length of the oligomer determines the size. DLS analysis of the control sequences revealed only objects below 5 nm in size, which

corresponds approximately to the size of single-stranded PNA and thus indicates the lack of measurable aggregation.

Finally, we used transmission electron microscopy (TEM) to visualize and characterize the assemblies formed. All sequences were prepared in ultrapure water and diluted to 100 µM. The

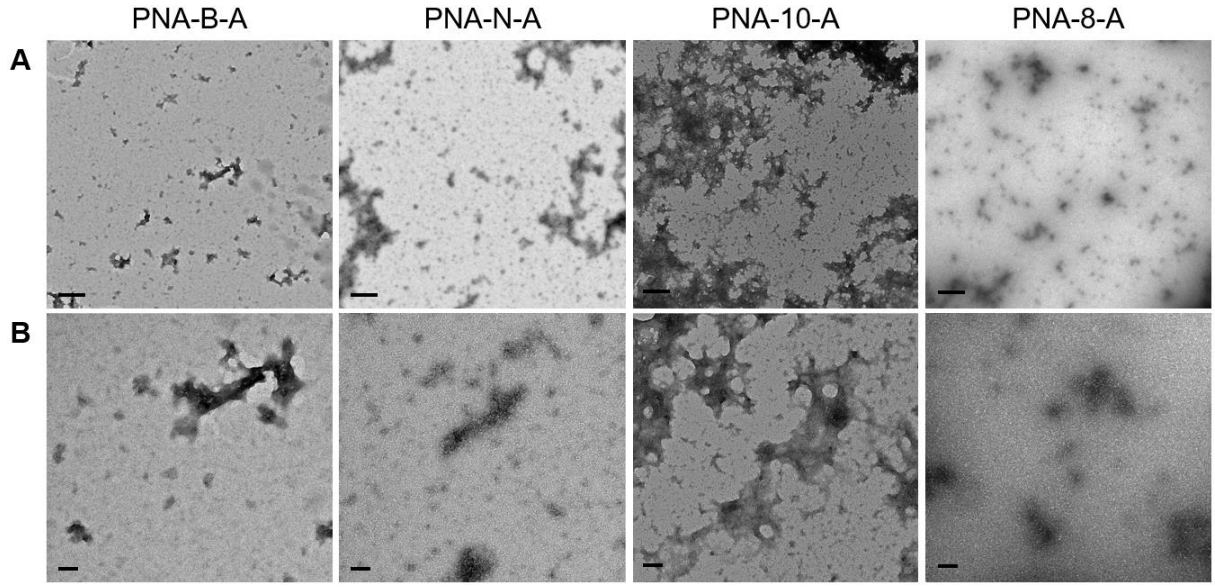


Figure 3. TEM images of PNA-B-A, PNA-N-A, PNA-10-A, and PNA-8-A assembled in water at 100 µM. A) Scale bar = 200 nm. B) Scale bar = 50 nm.

assemblies were then spotted onto the grid and stained with uranyl acetate to image. We observed that all of the amphiphilic sequences were capable of forming uniform, micellar-like assemblies. The dodecameric sequences PNA-B-A and PNA-N-A were found to be similar in size to the parent sequence at 30.3 ± 7.6 and 29.1 ± 14.3 nm, respectively. These assembly sizes were expected, as the DLS measurements of PNA-N-A and PNA-B-A similarly matched those of PNA-A.[9] The shorter sequences PNA-10-A and PNA-8-A were found to be smaller in size, measuring 17.7 ± 4.12 and 18.6 ± 3.1 nm, respectively, again in alignment with the DLS data (Figure 3, Table 4). The control PNA sequences were also analyzed and we did not observe the same spherical assemblies as seen for the amphiphiles (Figure S2). Rather, the only observable objects were amorphous in nature and likely due to uncontrolled PNA aggregation from the high concentrations. Overall, these results demonstrate that the PNA bilingual biopolymers are able to

Table 4. Average size of PNA assemblies as measured by TEM.

Strand	Diameter
PNA-B-A	30.3 ± 7.6 nm
PNA-N-A	29.1 ± 14.3 nm
PNA-10-A	17.7 ± 4.12 nm
PNA-8-A	18.6 ± 3.1 nm

undergo self-assembly, retain uniform shape, and have their size determined by the length of the oligomer.

Characterization of stimuli-responsive PNA disassembly. While the amino acid code of our bilingual biopolymers drives assembly, the nucleic acid code is responsible for stimuli-responsive disassembly, as hybridization to a highly anionic DNA or RNA strand diminishes the amphiphilicity. Thus, we were curious to also study the effect of changing sequence length and amino acid residues on stimuli-responsive disassembly. To test this, we used a 22 base pair DNA

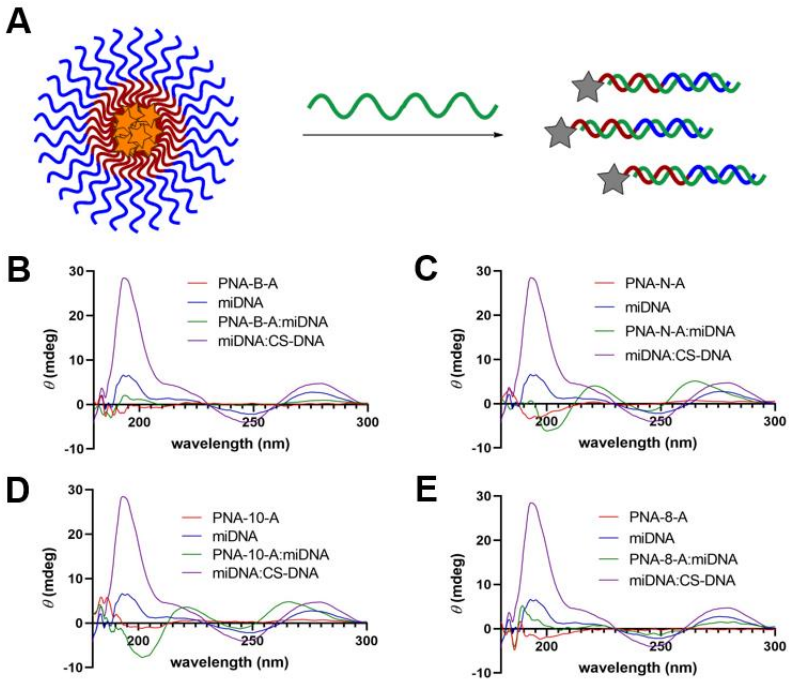


Figure 4. A) PNA assemblies undergoing stimuli-responsive disassembly in the presence of a complementary sequence. CD spectra of B) PNA-B-A C) PNA-N-A D) PNA-10-A, and E) PNA-8-A in 1x PBS at 100 μM to demonstrate the change in maxima and minima up on the addition of DNA.

sequence that would be fully complementary to all of the PNA sequences. Using circular dichroism (CD), we collected spectra for each PNA and PNA:DNA duplex, as well as a DNA:DNA control (Figure 4).[29] All solutions were prepared in 1x PBS at 100 μM final concentration of each oligonucleotide. As previously reported, the unmodified PNA sequences were found to impart no chirality, while the γ-modified sequences were found to produce a small chiral signal, as the modifications orient the sequence into a helical structure.[27, 30] We observed that all of the sequences were capable of binding to the DNA complement, with all but the PNA-B-A and PNA-8-A showing a shift in CD signal to maxima of 220 and 263 nm and minima of 200 and 244 nm, indicative of a PNA:DNA duplex.[12, 27, 29] We also tested the complementary DNA alone and DNA hybridized to a fully complementary DNA sequence (CS-DNA), which enabled us to distinguish between the ssDNA, PNA:DNA, and DNA:DNA CD signals. While PNA-B-A:DNA and PNA-8-A:DNA did not display the expected changes in CD signal for PNA-DNA duplex formation, the decrease in CD signal intensity is likely indicative of a change in hybridization state. For PNA-B-A, this unexpected result could be due to the change in the amino acid bulk, as this may result in a tighter helix confirmation, thus reducing the signal.[16, 31, 32] For PNA-8-A, this result is likely due to the change in the proximal nucleobases adjacent to the γ-modified PNA monomers, as previously reported.[27] The control sequences were also tested under similar conditions and were found to be capable of binding to DNA, inducing a similar change in CD signal.

Finally, to confirm that the bilingual biopolymers can undergo stimuli-responsive disassembly, we used TEM to image the changes in aggregation patterns upon addition of the complementary DNA sequence. Each sample was prepared having a final concentration of 100 μM each of PNA and DNA in 1x PBS. As previously reported, we observe the appearance of dark precipitates in

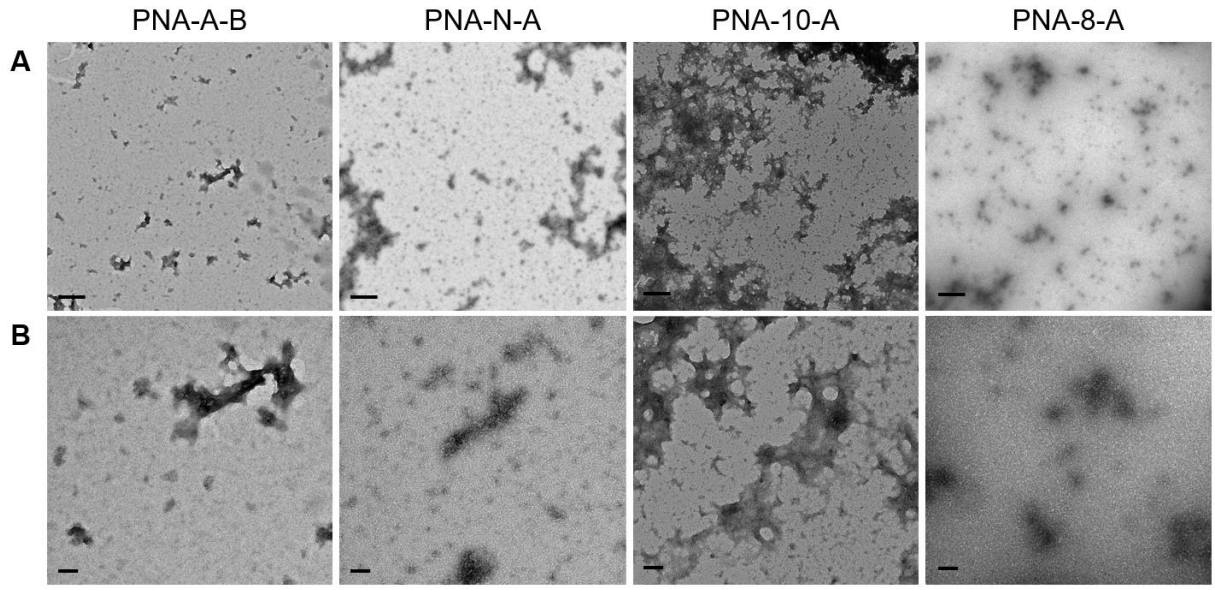


Figure 5. TEM images of PNA-B-A, PNA-N-A, PNA-10-A, and PNA-8-A disassembled in through hybridization to DNA in 1x PBS at 100 μM. A) Scale bar = 200 nm. B) Scale bar = 50 nm.

all samples containing DNA, likely due to the negative stain undergoing a nonspecific interaction with the phosphate backbone in DNA (Figure 5).[9, 10] As expected, we noted disappearance of the previously observed spherical assemblies, indicating disassembly through hybridization. Thus, through CD and TEM, we were able to establish the binding of complementary DNA sequences to the PNA amphiphiles and corresponding disassembly of the micellar assemblies, demonstrating all of the bilingual biopolymers remain capable of nucleic acid molecular recognition, irrespective of amino acid identity and PNA sequence length.

Conclusion

By synthesizing and characterizing a series of novel bilingual PNA biopolymers having variations in both the amino acid and nucleic acid codes, we were able to understand the impact of these parameters on assembly and stimuli-responsive disassembly. Our studies demonstrate that varying the identity and spacing of the amphiphilic side chains can modulate assembly strength but does not significantly alter overall morphology. Somewhat predictably, we found that assembly size can be tuned by altering sequence length. Finally, all the sequences tested were capable of undergoing stimuli-responsive disassembly upon hybridization to a complementary

DNA sequence. This further suggests that the amino acid residues direct assembly but do not interfere with the nucleic acid sequence recognition.

In Nature, the languages of nucleic acids and proteins encode information and impart function. Our previous work brought these two separate information systems together to produce a bilingual biopolymer that incorporates both amino acid and nucleobase sequences to direct assembly and stimuli-responsive disassembly. In our initial studies, we focused on a single sequence, and here we systematically explore the impact of amino acid side chains and nucleic acid sequence length. The ability to modulate assembly and disassembly properties by altering structure enables the generation of materials having tunable properties, providing insight into the assembly process and in turn increasing utility for biological and nanotechnology applications. We envision future studies involving the incorporation of unnatural amino acids and variations in amino acid patterning to increase the structural and functional repertoire of this class of biomaterials.

Experimental Procedures:

Abbreviations: Fmoc, fluorenylmethyloxycarbonyl; FAM, fluorescein; Acetonitrile, ACN; Water, H_2O; DIPEA, diisopropylethylamine; HATU, 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate; TFA, trifluoroacetic acid; DCM, dichloromethane; Boc, tert-butyloxycarbonyl; NMP, N-methyl-2-pyrrolidone.

PNA monomer synthesis. Unmodified PNA monomers were purchased from PolyOrg, Inc. Thymine acetic acid was purchased from Sigma-Aldrich. The protected nucleobases adenine and cytosine and Fmoc protected γ-methyl, γ-lysyl, γ-glutaryl, and γ-leucyl modified PNA backbones were synthesized following previously published procedures. The Fmoc-protected L-amino acids and nucleobase starting materials were purchased from Chem-Impex Inc. All other reagents were purchased from Fisher Scientific, Chem-Impex, and Sigma-Aldrich and used without additional

purification unless otherwise stated. Merck silica gel 60 F254 was used to monitor small molecule synthesis through thin layer chromatography (TLC) with UV light for visualization. Synthesized compounds were purified using flash chromatography with SiliaFlash F60 grade silica purchased from SiliCycle Inc. Characterization of compounds was completed using proton (^{1}H) and carbon (^{13}C) NMR using a Varian Inova 400 MHz spectrometer. The spectra were analyzed using MestReNova Software. The mass of the compounds was confirmed using an Agilent 6230 electrospray ionization time-of-flight (ESI-TOF) mass spectrometer. PNA monomer synthesis was completed as described previously.

PNA oligomer synthesis. PNA oligomers were synthesized as previously described using a Biotage SP Wave semiautomatic peptide synthesizer. Monomer coupling efficiencies were monitored using a Nanodrop 2000 spectrophotometer to measure dibenzofulvene-piperidine adduct absorbance at 301 nm. Upon cleavage, sequences were purified through reverse-phase HPLC using an Agilent Eclipse XDB-C18 5 µm, 9.4 x 250 mm column at 60 ºC with a flow rate of 2 mL/min, monitoring 260 and 495 nm using a linear gradient (10-40%) of 0.1% TFA/CAN in 0.1% TFA/H2O. The sequences were confirmed by ESI-TOF mass spectrometry.

Melting temperature analysis of PNA strands. All DNA sequences were purchased from the University of Utah DNA/Peptide Synthesis Core Facility or Integrated DNA Technologies. Samples containing 3 µM of complementary nucleic acids were prepared using stock solutions in 1x PBS. Samples were then added to an 8-cell quartz microcuvette with a 1 cm path length. A Shimadzu UV-1800 spectrophotometer equipped with a temperature controller and Julabo Corio CD water circulator were used for measurement of absorbance at 260 nm while heating from 25 to 95 ºC at a rate of 0.5 ºC/min. The melting temperatures were calculated using a first derivative method and experiments were carried out in triplicate.

Preparation of PNA amphiphile assemblies. To prepare amphiphile assembly, samples of PNA and PNA:DNA were dissolved in water or 1x PBS and were annealed as previously described. All samples were allowed to incubate at room temperature for 1 hour.

Determination of Critical Micelle Concentration using fluorescence. The critical micelle concentration (CMC) of each sequence was determined using the 4-DMN monomer fluorescence response. PNA samples and fully deprotected 4-DMN ranging from 32 μM to 0.061 μM in 1% DMSO/H_2O were prepared in stock solutions. 4-DMN was initially dissolved in DMSO, followed by dilution with water to a final solution containing 1% DMSO/H_2O. Each sample was prepared using the amphiphile assembly protocol as previously described. Samples were transferred to a 384-well clear-bottom plate for analysis on a Biotek Cytation5 plate reader (higher concentrations) or to a 50 μL quartz cuvette for analysis by a Horiba Scientific Dual-FL fluorometer (lower concentrations). Fluorescence measurements at 550 nm were collected using an excitation of 458 nm. Absorbance was also measured at 458 nm to correct for any changes in concentration between PNA samples and the monomer. A plot of the fluorescence enhancement of the PNA samples over the monomer as a function of concentration was created using GraphPad Prism software. A semilogarithmic fit to the data provided equations for the upper and lower portions of the curve, which were then compared to determine a CMC value.

Determination of size using dynamic light scattering. Samples of PNA and PNA:DNA, were prepared in 1x PBS at 200 μM from stock solutions and were annealed through heating and slow cooling as described above. Samples were then transferred to a quartz cuvette and analyzed using a Particular Systems NanoPlus DLS nano particle analyzer and data analyzed using NanoPlus Software.

Confirmation of DNA hybridizing to PNA by CD. Circular dichroism was used to confirm the hybridization of PNA:DNA. Individual samples of PNA and DNA were prepared at 200 μM in 1x PBS. 30 μL of PNA was combined with 30 μL of DNA to form the final 100 μM in 1x PBS solutions.

Samples were incubated for 1 hour, after which they were analyzed using a JASCO J-1500 circular dichroism spectrometer. Data points from 190 to 300 nm were collected at 1 nm intervals using a continuous scanning mode of 200 nm/min at 23.5 ºC. A sample with PNA alone was used as a control and had the signal overlapped using the JASCO software.

Characterization of PNA assemblies by TEM. Samples of PNA were prepared at 100 µM in water from stock solutions. Using a 200-mesh Formvar/carbon- coated copper grid, 3.5 µL of sample was spotted for 2 minutes and was wicked away with filter paper. Then, 3.5 µL of 1% uranyl acetate stain solution was spotted for 30 seconds before wicking. Grids were dried at room temperature for 30 minutes prior to imaging using a JEOL JEM-1400 transmission electron microscope.

Characterization of PNA-A disassembled using masking sequence DNA by TEM. Samples of PNA:DNA at 100 µM were prepared in 1x PBS and annealed through heating and slow cooling as described above. The duplex sample was then spotted and prepared for TEM imaging as described above.

Supporting Information

Characterization of PNA sequences, T_m data, TEM images, and additional displacement studies

Conflicts of Interest

The authors report no conflicting interests.

Acknowledgement

This work was supported by the National Science Foundation (DMR 2003987 to J.M.H.). The authors also acknowledge the Robert P. Apkarian Integrated Electron Microscopy Core and NMR Research Center at Emory University for instrument access and technical assistance. They would

also like to thank Mike Hanson and the oligonucleotide and peptide synthesis facility and the University of Utah for oligonucleotide materials.

References

1. Tan, X., et al., *Nucleic acid-based drug delivery strategies.* Journal of Controlled Release, 2020. **323**: p. 240-252.
2. Wang, L., et al., *Controlling the self-assembly of biomolecules into functional nanomaterials through internal interactions and external stimulations: A review.* Nanomaterials, 2019. **9**(2): p. 285.
3. Peterson, A.M. and J.M. Heemstra, *Controlling self-assembly of DNA-polymer conjugates for applications in imaging and drug delivery.* Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology, 2015. **7**(3): p. 282-297.
4. Seeman, N.C. and H.F. Sleiman, *DNA nanotechnology.* Nature Reviews Materials, 2017. **3**(1): p. 1-23.
5. Sato, K., et al., *Peptide supramolecular materials for therapeutics.* Chemical Society Reviews, 2018. **47**(20): p. 7539-7551.
6. Habibi, N., et al., *Self-assembled peptide-based nanostructures: Smart nanomaterials toward targeted drug delivery.* Nano Today, 2016. **11**(1): p. 41-60.
7. Kwak, M. and A. Herrmann, *Nucleic acid amphiphiles: synthesis and self-assembled nanostructures.* Chemical Society Reviews, 2011. **40**(12): p. 5745-5755.
8. Zhang, F., et al., *Structural DNA nanotechnology: state of the art and future perspective.* Journal of the American Chemical Society, 2014. **136**(32): p. 11198-11211.
9. Swenson, C.S., et al., *Bilingual peptide nucleic acids: encoding the languages of nucleic acids and proteins in a single self-assembling biopolymer.* Journal of the American Chemical Society, 2019. **141**(48): p. 19038-19047.

10. Argueta-Gonzalez, H.S., et al., *Stimuli-responsive assembly of bilingual peptide nucleic acids.* RSC Chemical Biology, 2022.

11. Egholm, M., et al., *Peptide nucleic acids (PNA). Oligonucleotide analogs with an achiral peptide backbone.* Journal of the American Chemical Society, 1992. **114**(5): p. 1895-1897.

12. Moccia, M., M.F. Adamo, and M. Saviano, *Insights on chiral, backbone modified peptide nucleic acids: properties and biological activity.* Artificial DNA: PNA & XNA, 2014. **5**(3): p. e1107176.

13. Shakeel, S., S. Karim, and A. Ali, *Peptide nucleic acid (PNA)—a review.* Journal of Chemical Technology & Biotechnology: International Research in Process, Environmental & Clean Technology, 2006. **81**(6): p. 892-899.

14. Egholm, M., et al., *PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules.* Nature, 1993. **365**(6446): p. 566-568.

15. Peng, L., et al., *Engineering and applications of DNA-grafted polymer materials.* Chemical Science, 2013. **4**(5): p. 1928-1938.

16. Gurezka, R., et al., *A heptad motif of leucine residues found in membrane proteins can drive self-assembly of artificial transmembrane segments.* Journal of Biological Chemistry, 1999. **274**(14): p. 9265-9270.

17. Wang, K., J.D. Keasling, and S.J. Muller, *Effects of the sequence and size of non-polar residues on self-assembly of amphiphilic peptides.* International Journal of Biological Macromolecules, 2005. **36**(4): p. 232-240.

18. Klein, T., et al., *Impact of amino acids on the aqueous self-assembly of benzenetrispeptides into supramolecular polymer bottlebrushes.* Polymer Chemistry, 2020. **11**(42): p. 6763-6771.

19. Koshti, B., et al., *Unusual Aggregates Formed by the Self-Assembly of Proline, Hydroxyproline, and Lysine.* ACS Chemical Neuroscience, 2021. **12**(17): p. 3237-3249.

20. Dzwolak, W. and P.E. Marszalek, *Zipper-like properties of [poly (L-lysine)+ poly (L-glutamic acid)] β-pleated molecular self-assembly.* Chemical communications, 2005(44): p. 5557-5559.

21. Mendes, A.C., et al., *Self-assembly in nature: using the principles of nature to create complex nanobiomaterials.* Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology, 2013. **5**(6): p. 582-612.

22. Edwards-Gayle, C.J. and I.W. Hamley, *Self-assembly of bioactive peptides, peptide conjugates, and peptide mimetic materials.* Organic & biomolecular chemistry, 2017. **15**(28): p. 5867-5876.

23. Nielsen, P.E., et al., *Peptide nucleic acids (PNAs) containing thymine monomers derived from chiral amino acids: hybridization and solubility properties of D-lysine PNA.* Angewandte Chemie International Edition in English, 1996. **35**(17): p. 1939-1942.

24. De Costa, N.T.S. and J.M. Heemstra, *Evaluating the effect of ionic strength on duplex stability for PNA having negatively or positively charged side chains.* PloS one, 2013. **8**(3): p. e58670.

25. Cavallaro, V., P. Thompson, and M. Hearn, *Solid-phase synthesis of a dendritic peptide related to a retinoblastoma protein fragment utilizing a combined boc-and fmoc-chemistry approach.* Journal of Peptide Science: An Official Publication of the European Peptide Society, 2001. **7**(5): p. 262-269.

26. Sforza, S., et al., *Induction of helical handedness and DNA binding properties of peptide nucleic acids (PNAs) with two stereogenic centres*. 2007, Wiley Online Library.

27. Dragulescu-Andrasi, A., et al., *A simple γ-backbone modification preorganizes peptide nucleic acid into a helical structure.* Journal of the American Chemical Society, 2006. **128**(31): p. 10258-10267.

28. Loving, G. and B. Imperiali, *A versatile amino acid analogue of the solvatochromic fluorophore 4-N, N-dimethylamino-1, 8-naphthalimide: a powerful tool for the study of*

dynamic protein interactions. Journal of the American Chemical Society, 2008. **130**(41): p. 13630-13638.

29. Faccini, A., et al., *Circular dichroism study of DNA binding by a potential anticancer peptide nucleic acid targeted against the MYCN oncogene.* Chirality: The Pharmacological, Biological, and Chemical Consequences of Molecular Asymmetry, 2008. **20**(3-4): p. 494-500.
30. Rapireddy, S., et al., *Strand invasion of mixed-sequence B-DNA by acridine-linked, γ-peptide nucleic acid (γ-PNA).* Journal of the American Chemical Society, 2007. **129**(50): p. 15596-15600.
31. Xu, H., et al., *Hydrophobic-region-induced transitions in self-assembled peptide nanostructures.* Langmuir, 2009. **25**(7): p. 4115-4123.
32. Xu, C., et al., *Rational design of helical nanotubes from self-assembly of coiled-coil lock washers.* Journal of the American Chemical Society, 2013. **135**(41): p. 15565-15578.

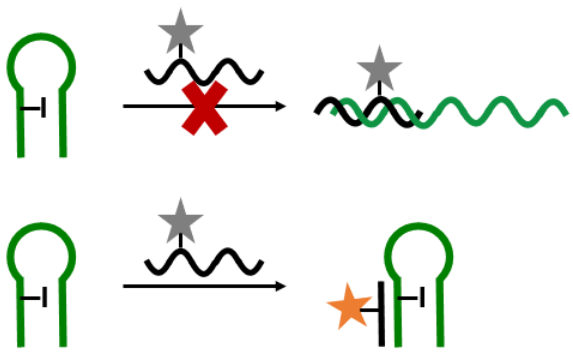
Chapter 4: Synthesis and development of a Triplex Forced Intercalation Peptide Nucleic Acid Probe for the Detection of Adenosine-to-Inosine Modification in Hairpin RNA*

Argueta-Gonzalez, H. S.; Swenson, C. S.; Sterling, S. A.; Robichaux, R.; Knutson, S. D.; Heemstra, J. M.

*Adaptation from 'Forced Intercalation Peptide Nucleic Acid Probes for the Detection of an Adenosine-to-Inosine Modification'

Abstract

One important nucleic acid modification found is the deamination of adenosine to inosine (A-to-I) edit that occurs as a posttranscriptional modification of RNA by ADAR (adenonsine deaminases acting on RNA) enzymes. Previously developed techniques can be used to monitor for edited sites, but they are not capable of real-time RNA analysis. By using forced intercalation (FIT) oligonucleotides, DNA and RNA targets can be monitored, however, they require high affinity and specificity, properties imparted onto PNA due to its backbone. By replacing the canonical nucleobase with a fluorogenic dye, a fluorescence response can be used to detect A-to-I edits. Previous work developed and tested a Malachite Green (MG), 4-DMN, and Thiazole Orange (TO) FIT PNA probes for use in monitoring A-to-I edits. However, the FIT PNA probes were found to be incapable of disrupting RNA hairpins. Herein, we report the synthesis of a triplex-FIT (tFIT) TO PNA probe with testing showed no inosine specific fluorescence change.



Introduction

Nature has evolved the conversion of DNA into protein through the use of RNA. Using RNA as the intermediary, there have been posttranscriptional steps that have evolved as an important mechanism of gene regulation. One specific posttranscriptional RNA edit involves the deamination of adenosine to inosine (A-to-I) through ADAR (adenosine deaminases acting on RNA) enzymes.[1] The modification to inosine can also be read in DNA as damage due to oxidation, affecting both transcription and translation. This deamination alters the base recognition, as inosine base pairs with cytosine, acting as a guanosine. The dysregulation of ADAR enzymes has been associated with various types of cancer and neurological disorders such as epilepsy and amyotrophic lateral sclerosis (ALS).[2-7]

While the A-to-I edit is important, the detection of single nucleotide polymorphisms in targeted sequences remains a key challenge. Previous work in our lab lead to the development of a high-throughput RNA sequencing method using an inosine enrichment technique.[8] However, we recognize that this technique suffered from incompatibility with real-time analysis of RNA editing from multi-step preparation. Another method to overcome this challenge was through the introduction of a reporter molecule covalently labeling edited transcripts.[9, 10] However, selective reactivity with inosine over other modified nucleobases hindered its live-cell and real-time applications. Thus, a need to develop a method for the selective detection of inosine-containing transcripts with high affinity remained.

One method to detect for A-to-I edits is through the use of Forced InTeracalation (FIT) probes, as it has proven to be useful for the detection and imaging of *in vitro* and *in cellulo* DNA and RNA sequences.[11] These FIT probes function through the introduction of a fluorogenic dye such as Thiazole Orange in place of a nucleobase, retaining Watson-Crick-Franklin hybridization.[12, 13]

Upon duplex formation, a change in fluorescent signal from the dye coincides with a site-specific nucleotide change.[14] Due to the high specificity and affinity requirements need for FIT probes application, peptide nucleic acid (PNA) has been demonstrated to be great scaffold.

PNA is an unnatural nucleic acid mimetic comprised of a pseudopeptide backbone.[15] The altered backbone has been shown to increase the affinity, specificity, and biostability, thus acting as a potent candidate to act as a FIT probe. Previous work has demonstrated FIT-PNA probes to be capable of nucleic acid detection in both *in vitro* and live cells.[15-19] Other labs have used thiazole orange (TO) to detect for single nucleotide polymorphisms, such as cytodine-to-uridine (C-to-U) editing.[20-23] However, no A-to-I FIT-PNA probe for the detection of ADAR activity has been reported on to the best of our knowledge.

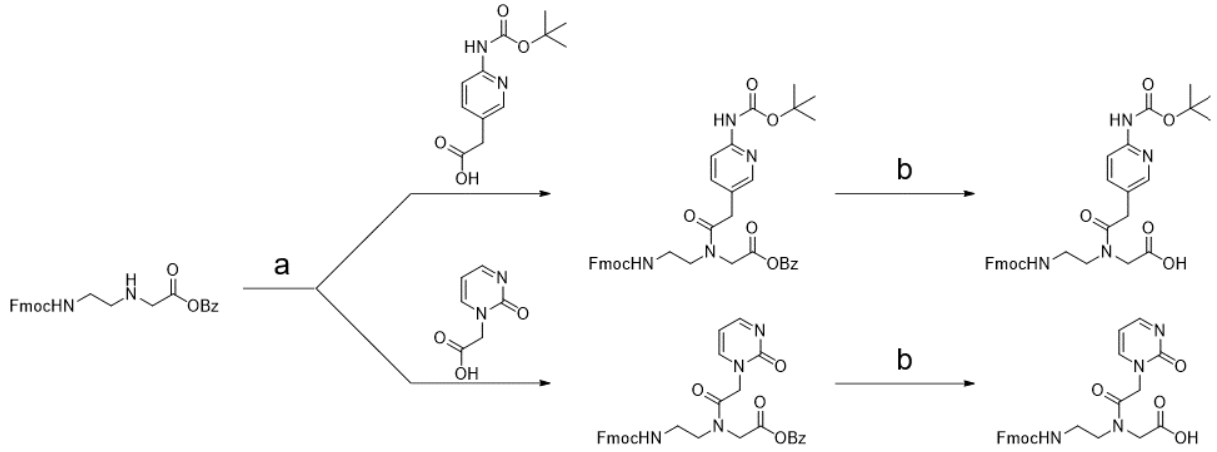
The library of fluorogenic dyes has continued to expand since the initial conception of FIT-PNA probes.[24-29] Our group previously synthesized 4-dimethylamino-naphthalimide (4-DMN) and malachite green (MG) as fluorogenic reporters.[30, 31] Thus, the application of these new fluorogenic dyes was thought possibly expand upon the TO FIT PNA fluorescence capabilities. Thus, previous work in our lab demonstrated that the TO and 4-DMN FIT-PNA probes were capable of detecting for A-to-I edits, however, MG did not have a sufficient fluorescence response. Furthermore, the FIT-PNA probes reported enhanced signal in the presence of linear RNA sequences. Finally, upon the introduction of HER1 RNA, a known hairpin, the probes were not found to undergo duplex invasion. Nishizawa et. al. demonstrated the use of cyanine dyes in triplex forming FIT PNA (tFIT-PNA) probe, capable of undergoing Hoogsteen base pairing and increasing fluorescent signal.[32-35] Thus, we hypothesized that through the introduction of triplex-forming monomers, a tFIT-PNA probe could bind and detect for A-to-I edits in duplex sequences. Herein, we report on the use of M and P PNA monomers for the synthesis and characterization of tFIT-PNA probes in the presence of HER1 RNA hairpin sequences.[36, 37]

The tFIT-PNA probe was found to bind in a triplex fashion, however, it did not demonstrate sufficient fluorescence enhancement for A-to-I edits.

Results and Discussion

Design and synthesis of triplex FIT PNA monomers and sequence

Scheme 1 Synthesis of M and P Monomers[a]



[a] Nucleobase acetic acid, HATU, NMM, DMF and (b) H_2, Pd/C, MeOH.

To create a triplex FIT PNA sequence, unnatural nucleobases that could undergo Hoogsteen base pairing were needed to be synthesized. Previous reports by the Rozner lab entailed the synthesis of both 2-aminopyridine (M) and 2-pyrimidinone (P) monomers that could recognize G and C, respectively, and bind in a triplex formation. To synthesize the Fmoc-protected M and P monomers, we first synthesized the Fmoc-AEG-OBn PNA backbone as previously reported by Faegin et al. After, the PNA backbone was coupled to 2-(2-Aminopyridin-3-yl)acetic acid and 2-(2-oxopyrimidin-1-yl)acetic acid, respectively, using HATU and N-methylmorpholine (NMM) to synthesize the protected monomers as following . The monomers were then subjected to palladium catalyzed hydrogenation and isolated to obtain the Fmoc-protected M and P monomers for use in developing tFIT PNA probes.

For the design of the tFIT PNA probe, we initially chose to bind to the duplex region of the HER1 RNA hairpin region of the sequence. Four HER1 RNA sequences were chosen, HER1-A, HER 1-I, HER 1-A2 and HER 1-I2, to identify the tFIT probe's capability in detecting for the A-to-I edit. The design of the probe was complementary to the 5'-HER1 region. The T monomer has also been reported to be capable of forming triplexes by base pairing with A and the Thiazole Orange (TO) monomer was placed adjacent to the A:C mismatch or I:C match. The probe was synthesized through solid-phase peptide synthesis, HPLC purified, and confirmed by ESI-TOF mass spectrometry.

Characterization of the tFIT PNA probes binding to HER1 targets

Initial fluorescence measurements of the PNA probe were observed in the presence of RNA 1-A, RNA 1-I, and no target. The fluorescence of the probe increased 1.79 and 1.68-fold, respectively, showing relatively similar increases and no specificity for either target as the previous TO PNA probe. We then measured the fluorescence over a range of temperatures to see if specificity via fluorescence could be elucidated. While fluorescence enhancement occurred,

the fluorescence ratio of RNA 1-A and RNA 1-I remained stagnant at a ratio of 1.3-fold increase of A to I. Thus, no fluorescence-based specificity was displayed in the HER1-1 system.

Table 1 List of RNA sequences and tFIT PNA probe

Strand	Sequence
HER 1-A	C G A G U U I G C G G G C 3' / C C C A A C C G C C C 5'
HER 1-I	C G A G U U I G C G G G C 3' / C C C A A C C G C C C 5'
HER 1-A2	C G A G U U G G C G G G C 3' / C C C A A C C A C C C 5'
HER 1-I2	C G A G U U G G C G G G C 3' / C C C A A C C I C C C 5'
TO tFIT PNA	N P T T P TO M P P P C

Underline region indicates tFIT PNA probe hybridization region via Hoogsteen base pairing

To further explore the tFIT PNA probes capabilities, we proceeded to use RNA 2, an unnatural variant of the HER1 RNA sequence with an A-to-I edit closer to the 5' end of the sequence. This change places the edit on the binding side of the tFIT PNA probes, increasing the likelihood that a fluorescence change would occur. A preliminary test measuring the fluorescence of the probe binding to RNA 2-A and RNA 2-I over a range of temperatures was completed. Unfortunately, the fluorescence ratio remained the same as RNA 1, indicating a lack of specificity. Thus, the tFIT TO PNA Probe was not capable of showing fluorescence-based specificity through

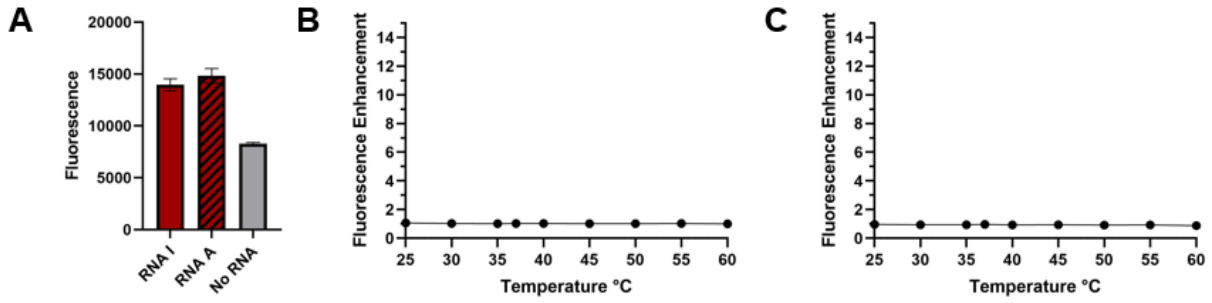


Figure 1 Fluorescence measurements of tFIT PNA probe. A) Fluorescence measurement of tFIT PNA probe with HER1-I, HER1-A, and no RNA at 25 ºC. B) Fluorescence enhancement of HER1-A:HER1-I using the tFIT PNA probe over the range of 25-60 ºC. C) Fluorescence enhancement of HER1-A2:HER1-I2 using the tFIT PNA probe over the range of 25-60 ºC.

a Hoogsteen base pairing interaction. Other works used the Thiazole Orange fluorophore were capable of measuring fluorescence-based specificities through the introductions of mismatches, demonstrating the TO undergoing base-stacking interactions when bound to the RNA sequence. Due to our previous TO FIT PNA probe showing specificity through Watson-Crick-Franklin base pairing, another method to overcome this limitation is through the incorporation of γ-modified PNA monomers. Specifically, through the introduction of guanidinium and acridine modifications, it has been shown to increase affinity and specificity and allows for PNA sequences to undergo both duplex and double duplex invasion using decamer sequences with a semi-consecutive positioning.[38, 39] This would allow for the previous FIT TO PNA probe to undergo duplex invasion of the hairpin HER1 sequence and demonstrate a fluorescence specificity.

Conclusion

Deamination of adenosine to inosine is a important nucleic acid modification that impacts cellular functions. Because of its impact, it is crucial to develop an optimal method for detection and quantification in real time. Previous work in our laboratory explored the use of FIT PNA probes to use a Thiazole Orange dye as a base surrogate and act as a reporter. While it was capable of detecting A-to-I edits in linear RNA sequences, it was not able to invade to invade the HER1 RNA hairpin. Our current work explored the use of unnatural nucleobases to synthesize of M and P monomers that could undergo Hoogsteen base pairing of G and C, respectively. Using these monomers, a triplex FIT TO PNA probe was synthesized and tested on two variants of HER1 RNA hairpins that contained an A-to-I edit. Unfortunately, the tFIT PNA probe was not capable of distinguishing adenosine from inosine. Future work in detecting A-to-I edits will be further explored through the novel incorporation of dyes and duplex invasion properties to enhance target selectivity for use in the detection and quantification of inosine modifications.

Experimental Procedures

Abbreviations: Fmoc, fluorenylmethyloxycarbonyl; Acetonitrile, ACN; Water, H_2O; DMF, dimethylformamide; HATU, 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate; NMM, N-methylmorpholine; DIPEA, diisopropylethylamine; TFA, trifluoroacetic acid; DCM, dichloromethane; Boc, tert-butyloxycarbonyl; NMP, N-methyl-2-pyrrolidone; PyBOP, benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate; FCC, flash column chromatography.

PNA monomer synthesis. Unmodified PNA monomers were purchased from PolyOrg, Inc. 2-(2-Aminopyridin-3-yl)acetic acid and 2-(2-oxopyrimidin-1-yl)acetic acid were purchased from Sigma-Aldrich and 1-Click Chemistry, respectively. All synthetic steps were carried out under an inert atmosphere unless stated otherwise. Fmoc-AEG-OBn backbone was synthesized as previously reported. All other reagents were purchased from Fisher Scientific, Chem-Impex, and Sigma-Aldrich and used without additional purification unless otherwise stated. Merck silica gel 60 F254 was used to monitor small molecule synthesis through thin layer chromatography (TLC) with UV light for visualization. Synthesized compounds were purified using flash chromatography with SiliaFlash F60 grade silica purchased from SiliCycle Inc. Characterization of compounds was completed using proton (^{1}H) and carbon (^{13}C) NMR using a Varian Inova 400 MHz spectrometer. The spectra were analyzed using MestReNova Software. The mass of the compounds was confirmed using an Agilent 6230 electrospray ionization time-of-flight (ESI-TOF) mass spectrometer.

PNA oligomer synthesis. PNA oligomers were synthesized as previously described using a Biotage SP Wave semiautomatic peptide synthesizer. For the thiazole orange PNA monomer, PyBOP was used in place of HATU for activation and coupling steps to improve solubility and efficiency. Monomer coupling efficiencies were monitored using a Nanodrop 2000 spectrophotometer to measure dibenzofulvene-piperidine adduct absorbance at 301 nm. Upon

cleavage, sequences were purified through reverse-phase HPLC using an Agilent Eclipse XDB-C18 5 µm, 9.4 x 250 mm column at 60 ºC with a flow rate of 2 mL/min, monitoring 260 and 495 nm using a linear gradient (10-40%) of 0.1% TFA/CAN in 0.1% TFA/H_2O. The sequences were confirmed by ESI-TOF mass spectrometry.

tFIT PNA fluorescence enhancement studies. PNA and RNA samples were prepared from stock solutions in 1xPBS with samples consisting of 2.5 µM of PNA and target RNA, PNA alone, RNA alone, and 1xPBS blank. The samples were annealed by heating to 95 ºC for 5 mins and cooling 1 ºC every 30 seconds until 25 ºC was achieved. The samples were then transferred into a 384-well flat-bottom plate for analysis on a Biotek Cytation5 plate reader. Fluorescence was then measured using an excitation at 500 nm and emission at 541 nm. The A-to-I fluorescence enhancement was calculated by dividing the adenosine-containing sequence fluorescence by the inosine containing sequence. General fluorescence enhancement used the PNA:RNA fluorescence and divided by PNA alone fluorescence. The fluorescence enhancement was plotted using GraphPad Prism software.

Thermal evaluation of tFIT PNA probes fluorescence enhancement. PNA and RNA samples were prepared as previously described. Fluorescence was measured every 5 ºC from 25-60 ºC, with 37 ºC as well. Fluorescence enhancement was calculated as previously described.

Supporting Information

Characterization of PNA sequences

Conflicts of Interest

The authors report no conflicting interests.

Acknowledgement

This work was supported by the National Science Foundation (DMR 2003987 to J.M.H.). The authors also acknowledge the NMR Research Center at Emory University for instrument access and technical assistance. They would also like to thank Mike Hanson and the oligonucleotide and peptide synthesis facility and the University of Utah for oligonucleotide materials.

References

1. Meier, J.C., et al., *RNA editing—systemic relevance and clue to disease mechanisms?* Frontiers in molecular neuroscience, 2016. **9**: p. 124.
2. Brusa, R., et al., *Early-onset epilepsy and postnatal lethality associated with an editing-deficient GluR-B allele in mice.* Science, 1995. **270**(5242): p. 1677-1680.
3. Chen, L., et al., *Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma.* Nature medicine, 2013. **19**(2): p. 209-216.
4. Kawahara, Y., et al., *RNA editing and death of motor neurons.* Nature, 2004. **427**(6977): p. 801-801.
5. Krestel, H.E., et al., *A genetic switch for epilepsy in adult mice.* Journal of Neuroscience, 2004. **24**(46): p. 10568-10578.
6. Kwak, S. and Y. Kawahara, *Deficient RNA editing of GluR2 and neuronal death in amyotropic lateral sclerosis.* Journal of molecular medicine, 2005. **83**(2): p. 110-120.
7. Maas, S., et al., *Underediting of glutamate receptor GluR-B mRNA in malignant gliomas.* Proceedings of the National Academy of Sciences, 2001. **98**(25): p. 14687-14692.
8. Knutson, S.D., et al., *Selective enrichment of A-to-I edited transcripts from cellular RNA using endonuclease V.* Journal of the American Chemical Society, 2020. **142**(11): p. 5241-5251.
9. Knutson, S.D., T.M. Ayele, and J.M. Heemstra, *Chemical labeling and affinity capture of inosine-containing RNAs using acrylamidofluorescein.* Bioconjugate chemistry, 2018. **29**(9): p. 2899-2903.

10. Knutson, S.D., et al., *Chemical Profiling of A-to-I RNA Editing Using a Click-Compatible Phenylacrylamide.* Chemistry–A European Journal, 2020. **26**(44): p. 9874-9878.
11. Tomoike, F. and H. Abe, *RNA imaging by chemical probes.* Advanced Drug Delivery Reviews, 2019. **147**: p. 44-58.
12. Seitz, O., F. Bergmann, and D. Heindl, *A Convergent Strategy for the Modification of Peptide Nucleic Acids: Novel Mismatch-Specific PNA-Hybridization Probes.* Angewandte Chemie International Edition, 1999. **38**(15): p. 2203-2206.
13. Köhler, O. and O. Seitz, *Thiazole orange as fluorescent universal base in peptide nucleic acids.* Chemical communications, 2003(23): p. 2938-2939.
14. Köhler, O., D.V. Jarikote, and O. Seitz, *Forced intercalation probes (FIT Probes): thiazole orange as a fluorescent base in peptide nucleic acids for homogeneous single-nucleotide-polymorphism detection.* ChemBioChem, 2005. **6**(1): p. 69-77.
15. Nielsen, P.E., et al., *Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide.* Science, 1991. **254**(5037): p. 1497-1500.
16. Demidov, V.V., et al., *Stability of peptide nucleic acids in human serum and cellular extracts.* Biochemical pharmacology, 1994. **48**(6): p. 1310-1313.
17. Egholm, M., et al., *PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules.* Nature, 1993. **365**(6446): p. 566-568.
18. Uhlmann, E., et al., *PNA: synthetic polyamide nucleic acids with unusual binding properties.* Angewandte Chemie International Edition, 1998. **37**(20): p. 2796-2823.
19. Nielsen, P.E., M. Egholm, and O. Buchardt, *Peptide nucleic acid (PNA). A DNA mimic with a peptide backbone.* Bioconjugate chemistry, 1994. **5**(1): p. 3-7.
20. Socher, E., et al., *FIT probes: peptide nucleic acid probes with a fluorescent base surrogate enable real-time DNA quantification and single nucleotide polymorphism discovery.* Analytical biochemistry, 2008. **375**(2): p. 318-330.

21. Dilley, R.J. and W.A. Morrison, *Vascularisation to improve translational potential of tissue engineering systems for cardiac repair.* The international journal of biochemistry & cell biology, 2014. **56**: p. 38-46.

22. Kam, Y., et al., *Detection of endogenous K-ras mRNA in living cells at a single base resolution by a PNA molecular beacon.* Molecular pharmaceutics, 2012. **9**(3): p. 685-693.

23. Fang, G.-m., et al., *A bright FIT-PNA hybridization probe for the hybridization state specific analysis of a C→ U RNA edit via FRET in a binary system.* Chemical science, 2018. **9**(21): p. 4794-4800.

24. Socher, E., A. Knoll, and O. Seitz, *Dual fluorophore PNA FIT-probes− extremely responsive and bright hybridization probes for the sensitive detection of DNA and RNA.* Organic & biomolecular chemistry, 2012. **10**(36): p. 7363-7371.

25. Bethge, L., D.V. Jarikote, and O. Seitz, *New cyanine dyes as base surrogates in PNA: Forced intercalation probes (FIT-probes) for homogeneous SNP detection.* Bioorganic & medicinal chemistry, 2008. **16**(1): p. 114-125.

26. Kolevzon, N., et al., *Single point mutation detection in living cancer cells by far-red emitting PNA–FIT probes.* Chemical Communications, 2016. **52**(11): p. 2405-2407.

27. Tedeschi, T., et al., *A pyrenyl-PNA probe for DNA and RNA recognition: Fluorescence and UV absorption studies.* Artificial DNA: PNA & XNA, 2010. **1**(2): p. 83-89.

28. Mansawat, W., et al., *Clicked polycyclic aromatic hydrocarbon as a hybridization-responsive fluorescent artificial nucleobase in pyrrolidinyl peptide nucleic acids.* Tetrahedron, 2012. **68**(21): p. 3988-3995.

29. Okamoto, A., K. Tanabe, and I. Saito, *Synthesis and properties of peptide nucleic acids containing a psoralen unit.* Organic letters, 2001. **3**(6): p. 925-927.

30. Swenson, C.S., et al., *Bilingual peptide nucleic acids: encoding the languages of nucleic acids and proteins in a single self-assembling biopolymer.* Journal of the American Chemical Society, 2019. **141**(48): p. 19038-19047.

31. Ayele, T.M., et al., *Fluorogenic photoaffinity labeling of proteins in living cells.* Bioconjugate chemistry, 2019. **30**(5): p. 1309-1313.

32. Zengeya, T., P. Gupta, and E. Rozners, *Triple‐helical recognition of RNA using 2‐aminopyridine‐modified PNA at physiologically relevant conditions.* Angewandte Chemie, 2012. **124**(50): p. 12761-12764.

33. Annoni, C., et al., *Triplex-forming peptide nucleic acid modified with 2-aminopyridine as a new tool for detection of A-to-I editing.* Chemical Communications, 2016. **52**(51): p. 7935-7938.

34. Sato, T., Y. Sato, and S. Nishizawa, *Triplex-forming peptide nucleic acid probe having thiazole orange as a base surrogate for fluorescence sensing of double-stranded RNA.* Journal of the American Chemical Society, 2016. **138**(30): p. 9397-9400.

35. Chiba, T., et al., *Red-emissive triplex-forming PNA probes carrying cyanine base surrogates for fluorescence sensing of double-stranded RNA.* Organic & Biomolecular Chemistry, 2017. **15**(37): p. 7765-7769.

36. Hnedzko, D., S.K. Cheruiyot, and E. Rozners, *Using Triple‐Helix‐Forming Peptide Nucleic Acids for Sequence‐Selective Recognition of Double‐Stranded RNA.* Current Protocols in Nucleic Acid Chemistry, 2014. **58**(1): p. 4.60. 1-4.60. 23.

37. Kumpina, I., et al., *Synthesis and RNA-binding properties of extended nucleobases for triplex-forming peptide nucleic acids.* The Journal of Organic Chemistry, 2019. **84**(21): p. 13276-13298.

38. Rapireddy, S., et al., *Strand invasion of mixed-sequence B-DNA by acridine-linked, γ-peptide nucleic acid (γ-PNA).* Journal of the American Chemical Society, 2007. **129**(50): p. 15596-15600.

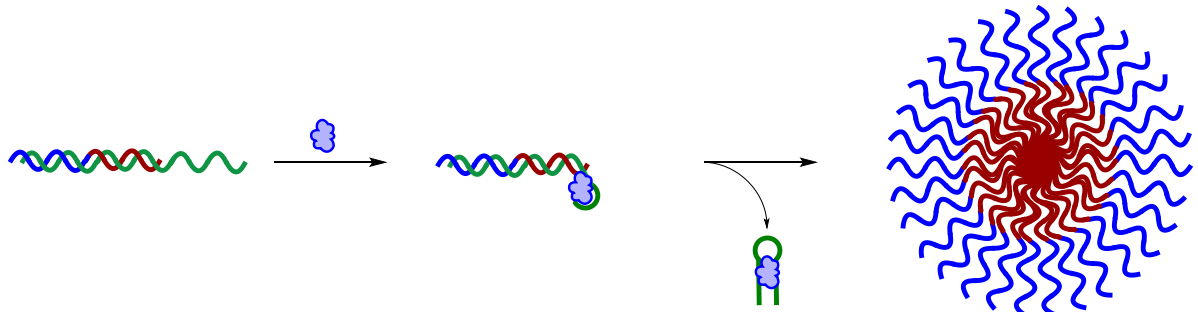
39. Ishizuka, T., et al., *Chiral introduction of positive charges to PNA for double-duplex invasion to versatile sequences.* Nucleic acids research, 2008. **36**(5): p. 1464-1471.

Chapter 5: Development of a Structure-switching DNA:PNA Biosensor for Ochratoxin A

Hector Argueta-Gonzalez, Colin S. Swenson, Jennifer M. Heemstra

Abstract

Structure-switching biosensors are powerful aptamer-based detection methods for targets of interest such as the Ochratoxin A (OTA) biosensor. While these methods are promising, there are limitations due to biostability and background fluorescence. Through the use of bilingual peptide nucleic acid (PNA) biopolymers, these limitations can be overcome by increasing biostability and fluorescent response with the incorporation of a stimuli-responsive assembly that can be induced upon the introduction of the target of interest. Here, we explore the development of a PNA:DNA structure-switching biosensor using the known OTA aptamer as a proof of principle. Upon the introduction of OTA, the structure-switching aptamer would bind the small molecule, releasing the PNA sequence. Initial unmodified PNA oligomers have been evaluated for application as a biosensor, however, optimization of the PNA:DNA based biosensor proved to not be optimal for the development of a stimuli-responsive assembly.



Introduction

Biosensors are powerful tools that can be used for the detection of a specific target such as peptides or small molecules. Particularly, nucleic acid-based biosensors are a promising method for the detection of targets of interest.[1-3] These systems contain single-stranded sequences that bind to targets of interest with high specificity, known as aptamers, which can be generated for specific targets through the *in vitro* process called systematic evolution of ligands

via exponential enrichment (SELEX).[2, 4] These sequences are readily available, stable under long-term storage conditions, and inexpensive to produce with little variability. Another benefit of aptamers is that chemical modifications can be introduced for analytical purposes, creating different types of biosensors such as the use of a fluorescently labeled aptamer for use in the detection of a target of interest using fluorescence polarization studies.

One of the more prominent forms is the structure-switching (SS) biosensor that takes advantage of the ability of some aptamers that can undergo a structural change upon binding to the target of interest.[2, 3] Specifically, a fluorescence-intensity based SS biosensor can be created by fluorescently labeling an aptamer and pairing it with a separate complementary sequence (CS) labeled with a quencher, effectively reducing fluorescence. Once the target is introduced, the aptamer would release the complementary sequence, or displacement strand, and bind the target. This results in a decrease in fluorophore-quencher proximity, thus returning fluorescent signal in relation to the concentration of the target.

While SS biosensors have been developed for a number of targets, one major limitation to this technology is the recognition property. Aptamers can bind to their targets of interest through interactions between the secondary structure formed from the nucleobase sequence and the target, and through molecular interactions. This specificity can be represented through the Ochratoxin A (OTA) aptamer from Xu, et al. that can distinguish between OTA and Ochratoxin B, structural analogues that only differ through the chlorine or hydrogen atom on the isocoumarin ring. Specifically, G5 forms three halogen bonds to the chlorine while three GCC triplets form the binding pocket for the interaction to occur.[4] Thus, if the G5 nucleobase or the triplets were not accessible, target recognition may be reduced or abolished. Another example comes from fluorogenic RNA aptamers Broccoli, Orange Broccoli, and Red Broccoli.[5] These aptamers each bind to 3,5-difluoro-4-hydroxybenzylidene imidzaolinone-2-oxime (DFHO) and each aptamer-target complex has distinct excitation and emission wavelengths. This difference is caused by the

change of a single nucleotide that interacts with DFHO, indicating that molecular interactions are vital in how the aptamer binds and recognizes the target. Thus, the SS biosensors must be tested with the fluorophore at different positions to determine if the location undergoes a significant conformational change and differing CS strands must be tested through changes in hybridization location and number of base pairs.[6, 7] This inherit property may be limiting the number of SS aptamers available for use in this technology as different sequences may have important structures and interactions that are perturbed through hybridization to the complementary sequence. Thus, by shortening the CS, fewer nucleobases from the aptamer are occupied in hybridization, increasing the likelihood that an aptamer that previously did not show SS properties can be used in SS biosensors. One possible avenue to explore shorter complementary sequence is through the incorporation of peptide nucleic acids (PNA).

PNA is a pseudopeptide analog to native nucleic acids that retains the Watson-Crick-Franklin base pairing ability.[8-10] The backbone imparts novel properties to the PNA sequences such as hydrolase resistance while increasing binding affinity and specificity. An additional benefit to PNA is that amino acid functional groups can be introduced into the backbone at the γ-position in a controllable manner to enhance the affinity and specificity further.[11-13] Thus, a shorter complementary sequence can be used in SS biosensor technology, allowing for a larger portion of the aptamer to remain unbound to form key secondary structures and to detect and bind to the target of interest. This would increase system sensitivity and the possibility that an aptamer can display structure-switching properties it previously could not due to length and location of CS hybridization.

Thus, we hypothesized through the combination of a quencher labeled PNA (QPNA) sequence complementary and fluorescently labeled aptamer, a novel biosensor will be developed that can function with previously nonfunctional SS aptamers. The QPNA/DNA complex will induce fluorescence quenching upon hybridization and due to the higher affinity, a shorter sequence can

be used. By introducing γ-modifications, the QPNA will have an additional increase in affinity, allowing for further sequence truncation. Upon the addition of the target of interest into the system, the release of the QPNA will induce the increase in fluorescence higher than native CS due to the increased free aptamer sequence portion for target recognition. Thus, the combination of a γ-modified QPNA with fluorescently labeled aptamer will create a more robust SS biosensor that has an increase in sensitivity and stability of the system with an optimized location to increase the likelihood of an aptamer displaying structure-switching properties.

Herein, we report the design and synthesis of a γ-modified PNA:DNA SS biosensor using the fluorescein-Ochratoxin A (OTA) aptamer from Xu et al. to create a biosensor as a model system. Initial testing revealed the necessary PNA capture strand (CS) length necessary to develop a PNA:DNA biosensor. After, we synthesized γ-modified quencher-labeled PNA (QPNA) CS sequences and evaluated the γ-PNA:DNA biosensor to act as a SS biosensor. We show that the γ-modified QPNA necessary for the biosensor is shorter in length. We also display a comparison in the differences in biosensor sensitivity and displacement based on hybridization length and location. Finally, we adapt a PNA sequence to hybridize to known key residues of the

Table 1 List of PNA and DNA sequences

Strand	Sequence
BHQ-1 DNA	3' BHQ1/A C A G G C T A C G A G G G A A A T G C G G T G G G T G T G G G C T A G 5'
FAM DNA	3' FAM/A C A G G C T A C G A G G G A A A T G C G G T G G G T G T G G G C T A G 5'
CS-DNA	5' BHQ1/T G T C C G A T 3'
PNA-1	N FAM/T G T C C C
PNA-2	N FAM/T G T C C G C
PNA-3	N FAM/T G T C C G A C
PNA-CM-1	N BHQ1/K A T C A T T G T_A C C G A_A K C
PNA-CM-2	N BHQ1/K T T C A T C C G A_A T G C_A K C
PNA-CM-3	N BHQ1/K T T C A A C C G A_AT G C_A K C

Subscript denotes amino acid residue using single letter code

OTA aptamer, thus losing the SS capability of the biosensor. This displays the importance of the aptamer's secondary structure and molecular recognition properties in structure switching capabilities. To the best of our knowledge, this is the first reported fluorescence-intensity based SS biosensor that uses PNA as the complementary sequence.

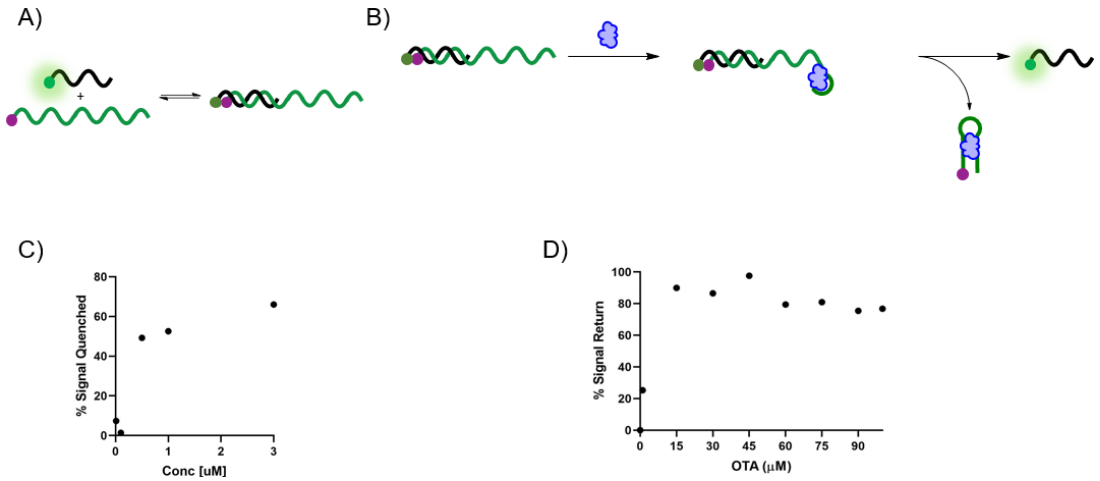


Figure 1 A) Hybridization of unmodified PNA to OTA aptamer sequence. Hybridization induces fluorescence quenching. B) Biosensor displacement upon the introduction of the OTA, the target molecule, measured through a fluorescence increase. C) % Signal Quenched from PNA-3:DNA system from 0.01-3 µM in OTA buffer. C) % Displacement from PNA-3:DNA system in relation to the concentration of OTA introduced.

Results and Discussion

Design of PNA Sequences

To create a γ-modified PNA amphiphile for stimuli-responsive assembly, the DNA-based OTA biosensor was chosen as model system. Short unmodified PNA sequences were synthesized and tested prior to the introduction of γ-modifications to optimize hybridization capabilities. The nucleobase sequence was made complementary to the 3' end of the OTA aptamer evolved by Cruz-Aguado to create an unmodified PNA:DNA biosensor, hybridizing in a similar location to the reported biosensor [14]. PNA oligomers of different lengths were designed shorter than the DNA oligomers designed by Chen *et al.* based on the differences between duplex stability and binding affinity differences between DNA and PNA.[1] Because γ-modified PNA has

an increased melting temperature and binding affinity compared to the same unmodified PNA sequence, it would not be ideal to initially test for greater than 90% signal quenching as this may be too stable to allow for the structure-switching aptamer to function.[15] We also took into account critical micelle concentration CMC of the previously reported bilingual PNA biopolymer, thus a low micromolar concentration was ideal for stimuli-responsive assembly.[11, 16]

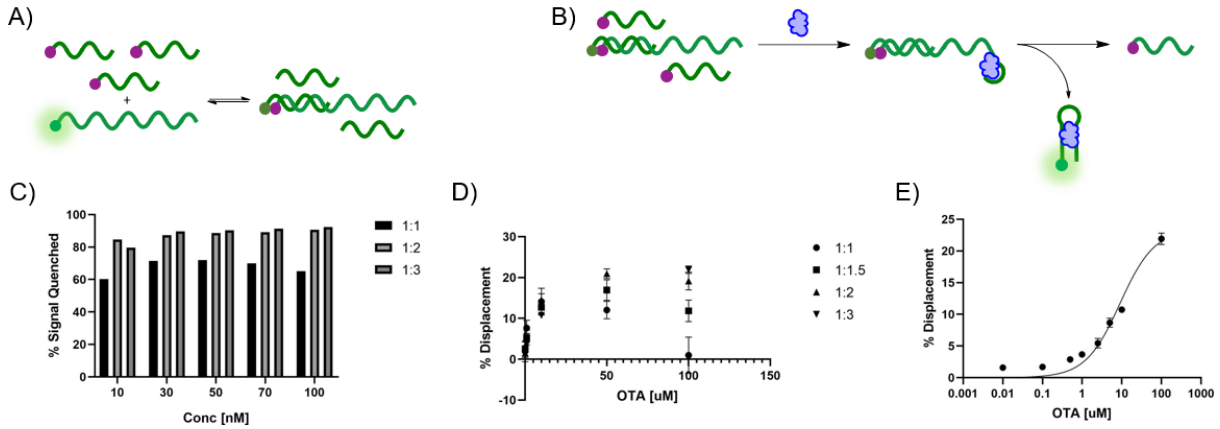


Figure 2 A) Hybridization of CS-DNA to OTA aptamer sequence. Changes in CS-DNA concentration increases fluorescence quenching. B) Biosensor displacement upon the introduction of the OTA, the target molecule, in the presence of higher concentration of CS-DNA. Displacement was measured through fluorescence increase. C) % Signal Quenched from CS-DNA:DNA system from 10-100 nM in OTA buffer. D) % Displacement from CS-DNA:DNA systems with various CS-DNA concentrations in relation to the concentration of OTA introduced. E) K_{sens} of the OTA biosensor system was evaluated at 1:3 Aptamer:CS ratio in the presence of OTA in OTA buffer.

To begin signal quenching studies, a PNA-1, PNA-2, and PNA-3, unmodified PNA sequences labeled with Fluorescein (FAM), were synthesized and characterized as previously reported. The OTA Aptamer was ordered and synthesized to contain black hole quencher-1 (BHQ-1) at the 3' terminal end. The sequences were then tested for signal quenching, annealing and testing for signal quenching at a 1:1 ratio at 3 µM in an OTA-binding buffer comprised of 10 mM Tris-HCl, 120 mM NaCl, 5 mM KCl, 1 mM $CaCl_2$ at pH 8.5. The 5-mer and 6-mer sequences were found to have minimal quenching, below 25% signal quenched, indicating the duplex stability was not sufficient enough to achieve the desired quenching in the concentration range. The 7-mer sequence was capable of 70% signal quenching, indicating sufficient stability for signal quenching. Further testing showed a concentration-dependent signal quenching, suggesting that the quenching is due to a hybridization event. A line of best fit was used to determine 1.46 µM as

a working concentration for 55% signal quenched. As a result, 1.5 μM of PNA-3 was used as the working concentration. Finally, preliminary testing of the biosensor from 0-100 μM OTA to induce displacement was measured and found to have greater than 80% displacement by 15 μM OTA. Thus, we were able to create a PNA:DNA based OTA biosensor.

To further understand the PNA:DNA biosensor and its properties, we further explored the DNA:DNA biosensor. Thus, a FAM-labeled OTA aptamer and BHQ-1 capture strand DNA (CS-DNA) were ordered and tested. We then tested % signal quenching with 1:1, 1:2, and 1:3 ratios of DNA:CS-DNA at various concentrations to determine the effect on % signal quenched. We found that having CS-DNA at 2 equivalents did increase signal quenching and that 3 equivalents lead to similar quenching. Afterwards, a 0-100 μM range of OTA was tested with 1:1, 1:1.5, 1:2, and 1:3 ratios at using 45 nM DNA:DNA biosensor concentration, as previously reported by Zhesen et al., to determine differences in the biosensor response. It was found that having an increase in CS-DNA lead to a higher % displacement at higher OTA concentrations along with a decrease in signal stability. Thus, it was found to be necessary to have a higher ratio of quencher capture strand.

After generation of the PNA:DNA biosensor and replicating the previously reported DNA:DNA biosensor, modified PNA sequences were designed. To match the DNA:DNA biosensor, having the capture strand contain the quencher, and to match our previously reported bilingual PNA biopolymer, 12-mer PNA sequences with the hybridization region to the OTA aptamer were designed. Two alanine residues were introduced at the γ-positions in the binding region to “mask” the hydrophobic effect through sequence recognition, as previously reported. A terminal lysine residue was introduced at both C- and N-terminals of the PNA sequence to assist in solubility and prevent self-assembly from occurring. PNA control modification 1 (PNA-CM-1) was designed to have the 7-nucleotide hybridization region fully complementary. PNA-CM-2 shifted over one nucleobase and contained a mismatch at the N-terminal as well to reduce binding

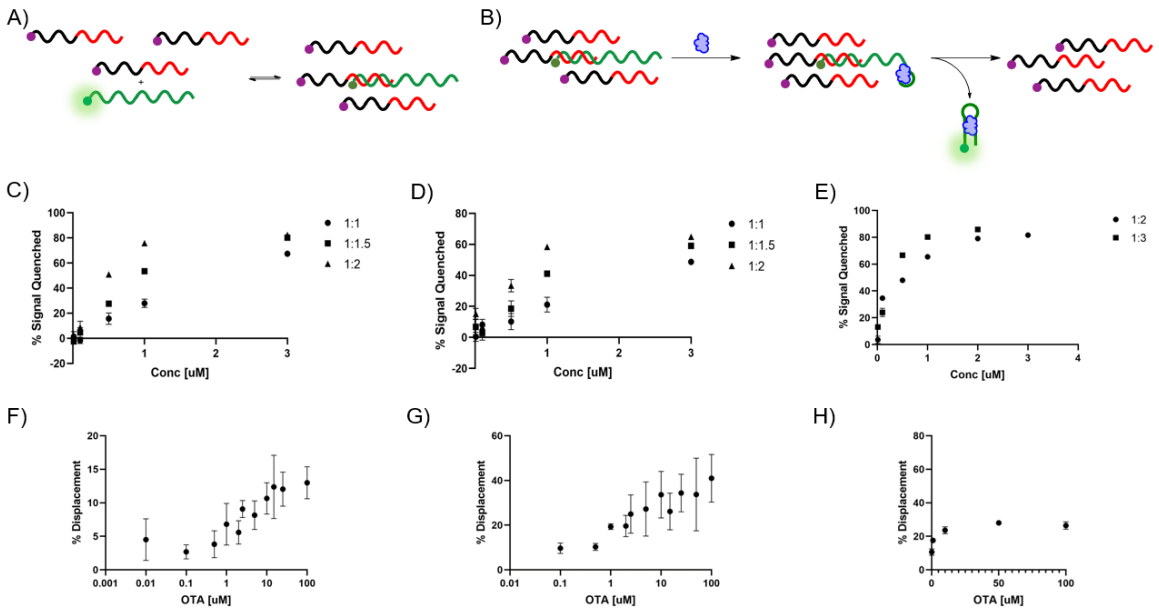


Figure 3 A) Hybridization of PNA-CM to OTA aptamer sequence at increasing CS ratios. B) PNA-CM displacement upon the introduction of OTA. Displacement was measured through fluorescence increase. C-E) % Signal Quenched from PNA-CM:DNA system from 0.1-3 µM in OTA buffer. B) PNA-CM-1. C) PNA-CM-2. D) PNA-CM-3. F-H) % Displacement from PNA-CM:DNA systems in relation to the concentration of OTA introduced. F) PNA-CM-1. G) PNA-CM-2. H) PNA-CM-3. The OTA biosensor systems were evaluated at 1:2 Aptamer:PNA ratio in the presence of OTA in OTA buffer.

affinity. Finally, PNA-CM-3 introduced a single mismatch, reducing it to a 6-nucleotide binding region. All 3 were capable of undergoing signal quenching, with PNA-CM-1 and PNA-CM-2 being tested at various concentrations and ratios. PNA-CM-3 was tested at 1:1, with a 1:2 signal quenching study completed at 1 µM. Matching the DNA:DNA biosensor, higher signal quenching was achieved with higher ratios.

The γ-PNA:DNA biosensors were then evaluated and the system conditions were optimized. Before testing, we recognized that, while higher concentrations and CS ratios for the biosensor systems were more likely to give better signal output, the critical micelle concentration (CMC), the concentration self-assembled structures begin to spontaneously form, of the previously reported PNA-A was 317 nM. Thus, unbound PNA was a concern for stimuli-responsive assembly. To test the PNA-CM-1 and PNA-CM-2 biosensors, 1 µM and 750 nM concentrations were used with 1:2 and 1:1.5 ratios, respectively. Upon the introduction of OTA, both systems showed a dose-dependent response. However, PNA-CM-1 showed a similar

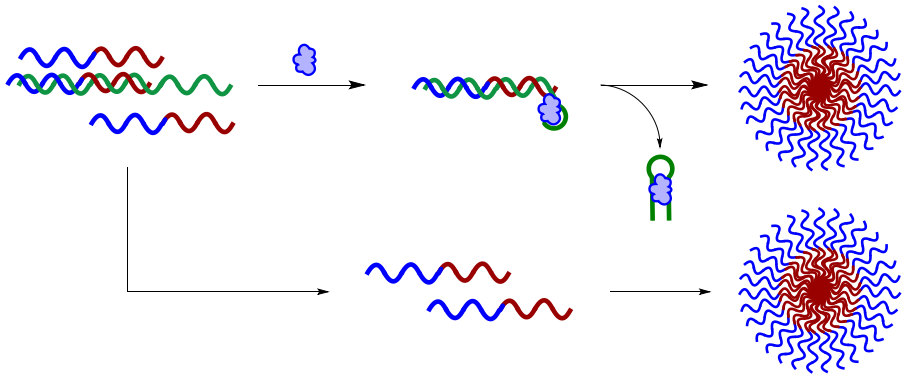


Figure 4 Development of a bilingual PNA biopolymer that would act as a capture strand in the OTA biosensor system. The unbound amphiphilic PNA would begin to assemble prior to the addition of the target molecule, thus only increasing the assembly population.

response to the DNA biosensor while PNA-CM-2 showed a greater response with both systems showing similar decreases in signal-to-noise. Finally, PNA-CM-3 was tested at 1 µM with a 1:2 ratio and showed an increase in both displacement and signal-to-noise ratio. Thus, a preliminary γ-PNA:DNA biosensor was developed. However, all three systems were not candidates for a stimuli-responsive assembly system due to the amount of unbound PNA. Due to hybridization occurring in a 1:1 ratio, the PNA-CM-3 system contained 1 µM of aptamer to 2 µM of PNA-CM-3. A 1 µM concentration of PNA-CM-3 remained unbound under SS biosensor conditions, approximately 3-fold higher than the CMC of PNA-A. Thus, assemblies would be formed prior to the introduction of the target molecule. One possible method to overcome this limitation is through the use of a longer hybridization region. However, it has been shown through computational modeling that structure-switching aptamers like the OTA aptamer require specific nucleobases for target recognition. By increasing the length of the binding region, it increases the likelihood that a nucleobase necessary for OTA recognition would be masked, decreasing the structure-switching biosensor efficacy. A second method to overcome this limitation is through bilingual PNA optimization, as previous work has developed a negatively-charged PNA-N-A had a CMC of 3.5 µM, thus the unbound PNA would not cross the assembly-formation threshold.

Conclusion

Structure-switching biosensors are powerful tools that can be used for target recognition with high specificity. However, due to their sensitivity, structure-switching aptamers are prone to a loss of target recognition upon the introduction of a capture strand sequence and native nucleic acids suffer from a lack of biostability. Herein, we explored the use of γ-PNA to develop a novel biosensor that can induce stimuli-responsive assembly upon the introduction of Ochratoxin A. FAM-labeled, unmodified PNA sequences were synthesized and tested for signal quenching capabilities and found a 7-mer sequence was capable of sufficient hybridization. Through testing of the DNA:DNA biosensor, the Aptamer:CS-DNA ratio was found to be necessary in developing a successful biosensor. After, quencher-labeled, γ-modified PNA sequences for γ-PNA:DNA biosensor optimization were synthesized and tested. All three control modified sequences were capable of ratiometric signal quenching and were able to undergo dose-dependent displacement. PNA-CM-3 was found to achieve the highest displacement and signal-to-noise ratio of the three. However, the current biosensor system was found to not be capable for use in developing a stimuli-responsive assembly. Future bilingual PNA biopolymer work may lead to CMC optimization that would make the SS biosensor system capable.

Experimental Procedures

Abbreviations: Fmoc, fluorenylmethyloxycarbonyl; FAM, fluorescein; Acetonitrile, ACN; Water, H_2O; BHQ-1, Black Hole Quencher-1

PNA monomer synthesis. Unmodified PNA monomers were purchased from PolyOrg, Inc. Thymine acetic acid was purchased from Sigma-Aldrich. The protected nucleobases adenine and cytosine and Fmoc protected γ-methyl and γ-lysyl modified PNA backbones were synthesized

following previously published procedures.[15, 16] The Fmoc-protected L-amino acids and nucleobase starting materials were purchased from Chem-Impex Inc. All other reagents were purchased from Fisher Scientific, Chem-Impex, and Sigma-Aldrich and used without additional purification unless otherwise stated. Merck silica gel 60 F254 was used to monitor small molecule synthesis through thin layer chromatography (TLC) with UV light for visualization. Synthesized compounds were purified using flash chromatography with SiliaFlash F60 grade silica purchased from SiliCycle Inc. Characterization of compounds was completed using proton (^{1}H) and carbon (^{13}C) NMR using a Varian Inova 400 MHz spectrometer. The spectra were analyzed using MestReNova Software. The mass of the compounds was confirmed using an Agilent 6230 electrospray ionization time-of-flight (ESI-TOF) mass spectrometer. PNA monomer synthesis was completed as described previously.[16, 17]

PNA oligomer synthesis. PNA oligomers were synthesized as previously described using a Biotage SP Wave semiautomatic peptide synthesizer.[16] Monomer coupling efficiencies were monitored using a Nanodrop 2000 spectrophotometer to measure dibenzofulvene-piperidine adduct absorbance at 301 nm. Upon cleavage, sequences were purified through reverse-phase HPLC using an Agilent Eclipse XDB-C18 5 µm, 9.4 x 250 mm column at 60 ºC with a flow rate of 2 mL/min, monitoring 260 and 495 nm using a linear gradient (10-40%) of 0.1% TFA/CAN in 0.1% TFA/H_2O. The sequences were confirmed by ESI-TOF mass spectrometry.

Fluorescence monitoring of hybridization and displacement. The fluorescence quenching of the FAM-labeled PNA and DNA sequences in the PNA:DNA or DNA:DNA duplexes were measured by fluorescence emission of the FAM fluorophore. Samples of 100 µL containing PNA alone, as a fluorescent control sample, and PNA:MS-R or MS-D over a range of concentrations (0.01 µM to 3 µM), at a 1:1-3 ratio were prepared using stock solutions in OTA buffer, comprised of 10 mM Tris-HCl, 120 mM NaCl, 5 mM KCl, 1 mM CaCl2 at pH 8.5. The samples were annealed by heating to 95 ºC for 5 mins and cooling 1 ºC every 30 seconds until 25 ºC was achieved. The

samples were then transferred into a 384-well flat-bottom plate for analysis on a Biotek Cytation5 plate reader. Fluorescence was then measured using an excitation at 495 nm and emission at 520 nm. The fluorescence measurement was used to plot the % Quenched using Eq. 1 using GraphPad Prism software.

$$\% \text{ Quenched} = \frac{(F_m - F_0)}{(F_m)} \times 100 \qquad (1)$$

To measure the % Displacement via fluorescence intensity, 50 µL of samples containing 2X the final concentration of PNA or CS-DNA and DNA, each in OTA buffer. 50 µL of OTA at 2X final reported quantities (0.1 to 100 µM) in 2% DMSO were prepared. After, the OTA samples transferred into the PNA:DNA or DNA:DNA samples to achieve the final reported concentrations. Samples containing PNA/CS-DNA alone and PNA:DNA/DNA:DNA were used and had 1x OTA buffer blank solutions added as controls. The samples were then transferred to a 384-well flat-bottom plate for analysis using the previously described protocol. The samples were measured at 0 and 1 hour incubation intervals. The fluorescence was then used to calculate the % Displacement using Eq. 2 as described and plotted using GraphPad Prism software.

$$\% \text{ Displacement} = \frac{(F - F_0)}{(F_m - F_0)} \times 100 \qquad (2)$$

F_m is the fluorescently labeled PNA sequence alone, F_0 is the fluorescence when PNA is hybridized to a quencher-labeled masking sequence, and F is the fluorescence measured when inducing a stimuli-response.

Supporting Information

Characterization of PNA sequences, and additional quenching and displacement studies

Conflicts of Interest

The authors report no conflicting interests.

Acknowledgement

This work was supported by the National Science Foundation (DMR 2003987 to J.M.H.). The authors also acknowledge the Robert P. Apkarian Integrated Electron Microscopy Core and NMR Research Center at Emory University for instrument access and technical assistance. They would also like to thank Mike Hanson and the oligonucleotide and peptide synthesis facility and the University of Utah for oligonucleotide materials.

References

1. Chen, J., et al., *A simple and rapid biosensor for ochratoxin A based on a structure-switching signaling aptamer.* Food Control, 2012. **25**(2): p. 555-560.
2. Sanford, A.A., et al., *RE-SELEX: Restriction Enzyme-Based Evolution of Structure-Switching Aptamer Biosensors.* 2021.
3. Tan, Z., T.A. Feagin, and J.M. Heemstra, *Temporal control of aptamer biosensors using covalent self-caging to shift equilibrium.* Journal of the American Chemical Society, 2016. **138**(20): p. 6328-6331.
4. Xu, G., et al., *Structure-guided post-SELEX optimization of an ochratoxin A aptamer.* Nucleic acids research, 2019. **47**(11): p. 5963-5972.
5. Filonov, G.S., W. Song, and S.R. Jaffrey, *Spectral tuning by a single nucleotide controls the fluorescence properties of a fluorogenic aptamer.* Biochemistry, 2019. **58**(12): p. 1560-1564.
6. Nutiu, R. and Y. Li, *Structure-switching signaling aptamers.* Journal of the American Chemical Society, 2003. **125**(16): p. 4771-4778.
7. Nutiu, R. and Y. Li, *Aptamers with fluorescence-signaling properties.* Methods, 2005. **37**(1): p. 16-25.
8. Egholm, M., et al., *PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules.* Nature, 1993. **365**(6446): p. 566-568.

9. Egholm, M., et al., *Peptide nucleic acids (PNA). Oligonucleotide analogs with an achiral peptide backbone.* Journal of the American Chemical Society, 1992. **114**(5): p. 1895-1897.

10. Nielsen, P.E., et al., *Peptide nucleic acids (PNAs) containing thymine monomers derived from chiral amino acids: hybridization and solubility properties of D-lysine PNA.* Angewandte Chemie International Edition in English, 1996. **35**(17): p. 1939-1942.

11. Argueta-Gonzalez, H.S., et al., *Stimuli-responsive assembly of bilingual peptide nucleic acids.* RSC Chemical Biology, 2022.

12. Moccia, M., M.F. Adamo, and M. Saviano, *Insights on chiral, backbone modified peptide nucleic acids: properties and biological activity.* Artificial DNA: PNA & XNA, 2014. **5**(3): p. e1107176.

13. Rapireddy, S., et al., *Strand invasion of mixed-sequence B-DNA by acridine-linked, γ-peptide nucleic acid (γ-PNA).* Journal of the American Chemical Society, 2007. **129**(50): p. 15596-15600.

14. Cruz-Aguado, J.A., G.J.J.o.a. Penner, and f. chemistry, *Determination of ochratoxin A with a DNA aptamer.* 2008. **56**(22): p. 10456-10461.

15. De Costa, N.T.S. and J.M. Heemstra, *Evaluating the effect of ionic strength on duplex stability for PNA having negatively or positively charged side chains.* PloS one, 2013. **8**(3): p. e58670.

16. Swenson, C.S., et al., *Bilingual peptide nucleic acids: encoding the languages of nucleic acids and proteins in a single self-assembling biopolymer.* Journal of the American Chemical Society, 2019. **141**(48): p. 19038-19047.

17. Porcheddu, A., et al., *A practical and efficient approach to PNA monomers compatible with Fmoc-mediated solid-phase synthesis protocols*. 2008, WILEY-VCH Verlag Weinheim.

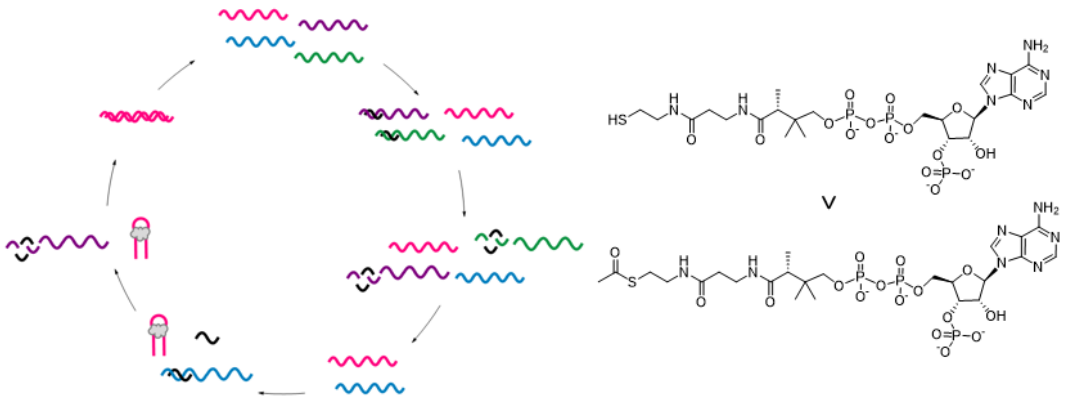
Chapter 6: Optimization of a CoA Structure-Switching Biosensor*

Hector Argueta-Gonzalez, Aimee A. Sanford, Alexandra E. Rangel, Trevor A. Faegin, Robert G. Lowery, Jennifer M. Heemstra

*Adaptation from 'RE-SELEX: restriction enzyme-based evolution of structure-switching aptamer biosensors'

Abstract

Structure-switching aptamers are promising for the development of high-throughput screening methods as the target-induced structural change can be adapted to a functional output. Current methods do not select for the elucidation of aptamers for use in structure-switching biosensors. A novel SELEX selection method to select for a structure-switching biosensor was previously described and identified six candidate sequences for optimization. We optimized the sequences for use in biosensor systems and evaluate target-binding capabilities, showing no specificity for the target molecule, Acetyl Coenzyme A.



Introduction

The detection and quantification of specific targets such as small molecule and metabolites are important as they are vital in human health.[1] One example is the detection of cAMP in cell extracts and ADP to screen for GPCR targets and kinases, respectively.[2, 3] Current

methodologies are limited to high-cost equipment or expertise. Recent efforts have been focused on the development of methodologies that can are highly sensitive and that combine cost-effectiveness and are user friendly.[1, 4] However, due to their small size and fewer binding regions, small molecules remain an elusive target.

Antibodies are generally regarded as the gold standard for HTS assays. However, they are unable to distinguish between similarly structured small molecules. Small molecule functionalization for SELEX and other assay developments further limits antibody capabilities due to the modification of key recognition elements. Additionally, carrier protein conjugation may reduce affinity and specificity towards the target of interest. Many assays such as fluorescence polarization require small molecule conjugation to a reporter to act as a competitor to in the assay.

Our lab recognized the need to develop HTS assays for the detection of coenzyme A (CoA) to measure histone acetyltransferase (HAT) activity, as dysregulation can lead to various diseases such as cancers, neurological disorders, and cardiovascular diseases.[5-8] The 18 known HAT enzymes and their subsequent reactions produce CoA, thus making CoA a target of interest for the detection of universal and specific acetylation events. However, specificity for CoA would need to occur in the presence of acetyl CoA (AcCoA). Current methods of detection use fluorescent probes to react with the free thiol of CoA, reducing fluorescence emission upon reacting. However, this leads to false positives and false identifications.[9, 10] Other methods take advantage of colorimetry using coupled enzyme approaches, however, these methods suffer from insensitivity and can undergo interference.[11]

Nucleic acids are a promising method for the detection of biomolecules through non-covalent interactions, an alternative to antibodies.[12-15] These single-stranded nucleic acid sequences can be evolved *in vitro* through Systemic Evolution of Ligands by Exponential Enrichment (SELEX) without requiring conjugation to a carrier protein and can discriminate between small structural differences with high affinity.[16-18] One approach to detect CoA is through the

development of an aptamer-based sensor, as they have shown great promise in being low-cost, user-friendly, stability, and ease of synthesis.[1, 4] Specifically, the structure-switching (SS) biosensor format is the most promising, as it contains a hybridized short complementary sequence to form a duplex. Upon the addition of the target molecule, the recognition event would induce displacement, with this event being capable of inducing a readable output in a dose-dependent manner.[19, 20]

While SS biosensors have a high potential, these methods are limited due to the lack of post-selection modification process. Thus, the SS biosensors previously reported on contain aptamers with inherit SS capabilities. This limitation stems from the SELEX process, as it is engineered to identify target binding and not on a conformational change.[21-23] Effort to develop a structure-switching selection using SELEX had been underway.[24-26] Previous efforts in our laboratory were towards the development a novel structure-switching biosensor for the detection of Coenzyme A (CoA) using a homogenous PCR-based SELEX process. This methodology used an N_{40} DNA library that contained a capture strand (CS) nearing the random region. This binding region would also be flanked by a double-stranded palindromic recognition EcoRI restriction enzyme binding site. Thus, non-binding sequences would remain hybridized, allowing for EocRI recognition and cleavage and disruption of PCR amplification. Binding sequences would displace the CS and remain intact and bound to the target of interest, allowing the selection process to identify sequences capable of acting as structure-switching biosensors.

Herein, we report on the previously identified coenzyme A SELEX candidates from the novel SS SELEX process. Six sequences were tested for signal quenching capabilities and optimized for biosensor testing. Upon the introduction of CoA, significant fluorescence quenching occurred and was identified to be caused by the target molecule. Aptamer candidates were then tested through microscale thermophoresis (MST) and fluorescence polarization to evaluate binding capabilities. Fluorescence polarization showed binding potential, however, MST failed to show binding

specificity. This project served as a framework for the RE-SELEX methodology to develop a SS biosensor for Kanamycin A.

Table 1 List of DNA sequences

Candidate Sequence #	Sequence 5'-3'
1	CGCATACCAGCTTAGTTCAGAATTCATTTTGAAATAATAAGCTGGTGTCAGACATTC ATTTCAGAATTAGATAGTAAGTGCAATCTCGGC
2	CGCATACCAGCTTAGTTCAGAATTCATTAAAGCTGGTGTGCATCAGGATTCATTTG GTAAAGACTATGAGATAGTAAGTGCAATCTCGGC
3	CGCATACCAGCTTAGTTCAGAATTCATTCTACGGCTGGTTATCGGCAATTCATTTCA GAAGTTCATTGAGATAGTAAGTGCAATCTCGGC
4	CGCATACCAGCTTAGTTCAGAATTCATTAAGCTGGGATGATCAGGATTCATTCAAAA TTGCGCTGTGCAGATAGTAAGTGCAATCTCGGC
5	CGCATACCAGCTTAGTTCAGAATTCATTACGAACTATGTGGTATCGTCAGTTCATTT CAGCAATTCATAGATAGTAAGTGCAATCTCGGC
6	CGCATACCAGCTTAGTTCAGAATTCATTCTGAATAGGTATGCAAACTCAGAAATCAT TTCAGAACTTCAGATAGTAAGTGCAATCTCGGC
Capture Strand (CS)	AATGAATTCTGA

Candidate sequences were labeled at the 5'positon with FAM. Capture strand was labeled at the 3'position with Black Hole Quencher-1 (BHQ-1)

Results and Discussion

DNA Biosensor Optimization

Previous reports on the development of biosensors determined that concentrations nearing total aptamer hybridization to capture strand perform optimally, likely due to enthalpic gain from target binding shifting the equilibrium to aptamer binding, thus capture strand release. Thus, % signal quenched was measured from 10 to 500 nM at a 1:1 ratio to evaluate sequence hybridization. All sequence candidates were labeled with a 5'-FAM and hybridized to a black hole quencher-1 (BHQ-1) labeled capture strand. In testing, most sequences performed similarly, showing a dose-dependent increase in quenching, indicative of hybridization. However, while

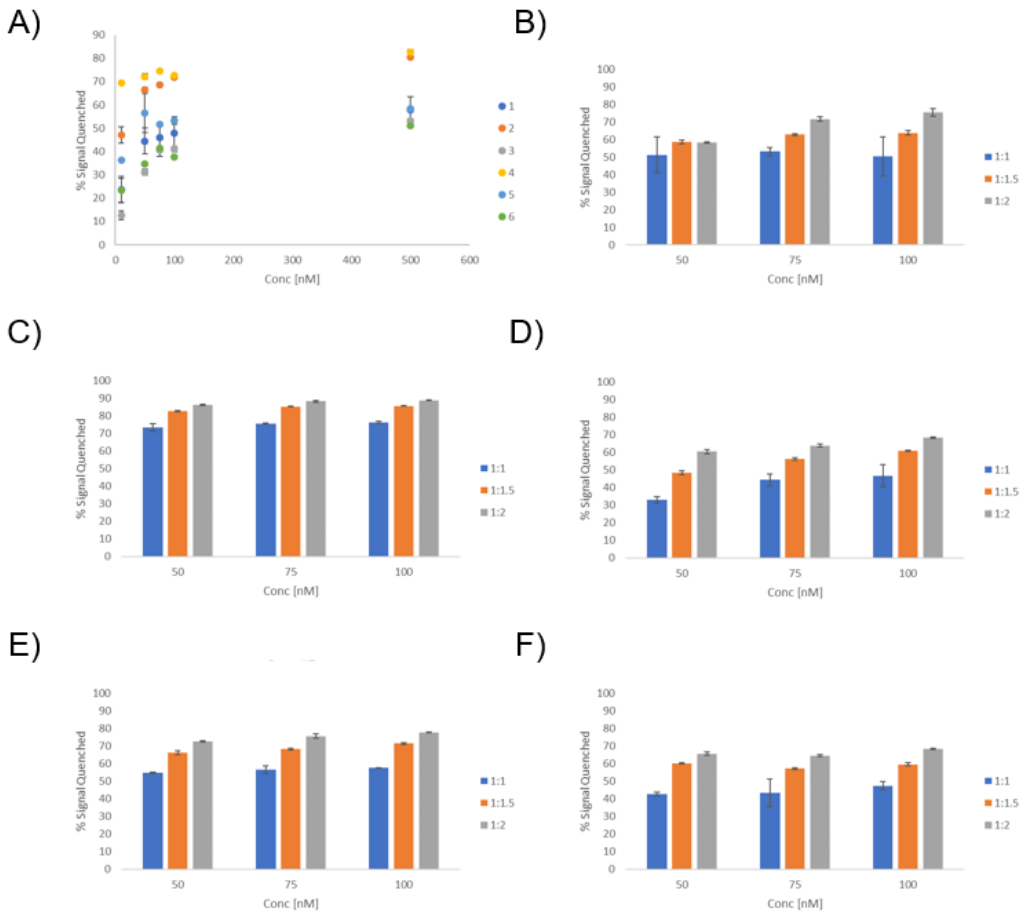


Fig 1 Candidate Sequence Biosensor optimization. A) Candidate sequences were hybridized to capture strand at a 1:1 ratio over a range of concentration to measure signal quenching capabilities. B-F) Candidate sequences were optimized to have varying ratio of capture strand, using 1, 1.5, and 2 equivalents of CS to Aptamer. B) Candidate sequence #1. C) Candidate sequence #2. D) Candidate sequence #3. E) Candidate sequence #5. F) Candidate sequence #6.

sequence #4 demonstrated signal quenching >70% at all concentrations, no dose-dependent increase was displayed. This suggests that the hybridization state does not change and that the equilibrium of bound and unbound sequences remain constant. Thus, Sequence #4 had low likelihood of acting as a biosensor and was removed from optimization. After, the candidate biosensor systems were tested at an aptamer:CS ratio of 1:1, 1:1.5, and 1:2 over a range of concentrations to evaluate the ratiometric signal quenching, as the SELEX method used a 1:2 ratio in sequence evolution and previous systems have been reported. All sequences were found to have increases in signal quenched, as expected.

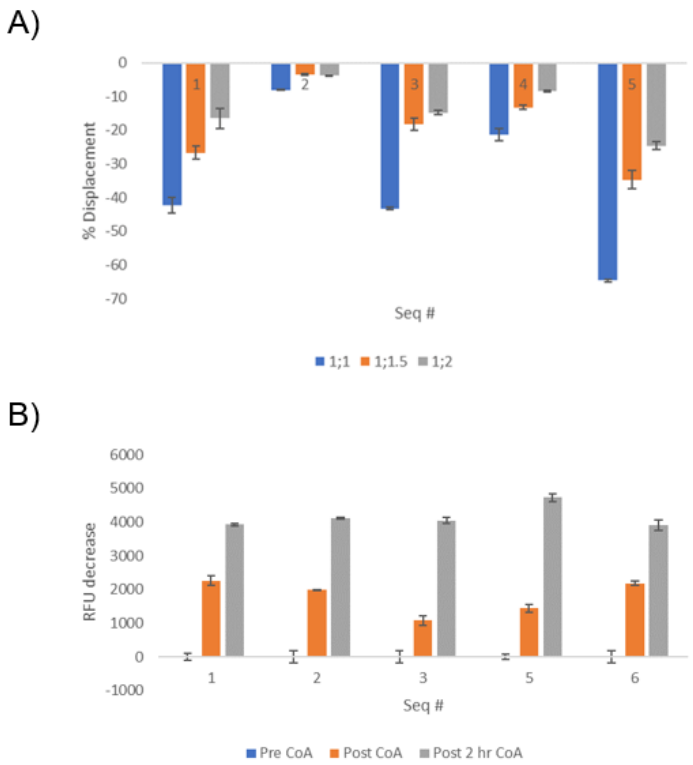


Fig 2 Candidate Biosensor target displacement. A) Candidate biosensor systems tested for displacement using CoA and measure an increase in fluorescence. B) Evaluation of CoA signal quenching using FAM-labeled candidate sequences alone, demonstrating that the signal was reduced upon the introduction of the target molecule due to fluorophore quenching.

After signal quenching was evaluated, the biosensor systems were then evaluated through the introduction of Coenzyme A. The five systems were tested at various concentrations and ratios for an increase in fluorescence to indicate a change in hybridization state occurring due to the release of CS for the target molecule. Upon the introduction of CoA, all systems were found to have a significant decrease in fluorescence. Further testing with the aptamer candidates alone showed similar fluorescence decreases that indicated the target molecule was in contact with the fluorophore and was increasingly causing contact quenching over time. Thus, testing for biosensor activity through a fluorophore-quencher biosensor system was not ideal.

To evaluate the aptamer candidates' binding capabilities alone, fluorescence polarization (FP) and MST were used. For both FP and MST, candidate #6 was used to as an initial candidate. Through FP, 50 nM of the aptamer was tested by titrating 500 nM of CoA and comparing the mP to a control sample with no molecule, showing a difference and indicating a binding event

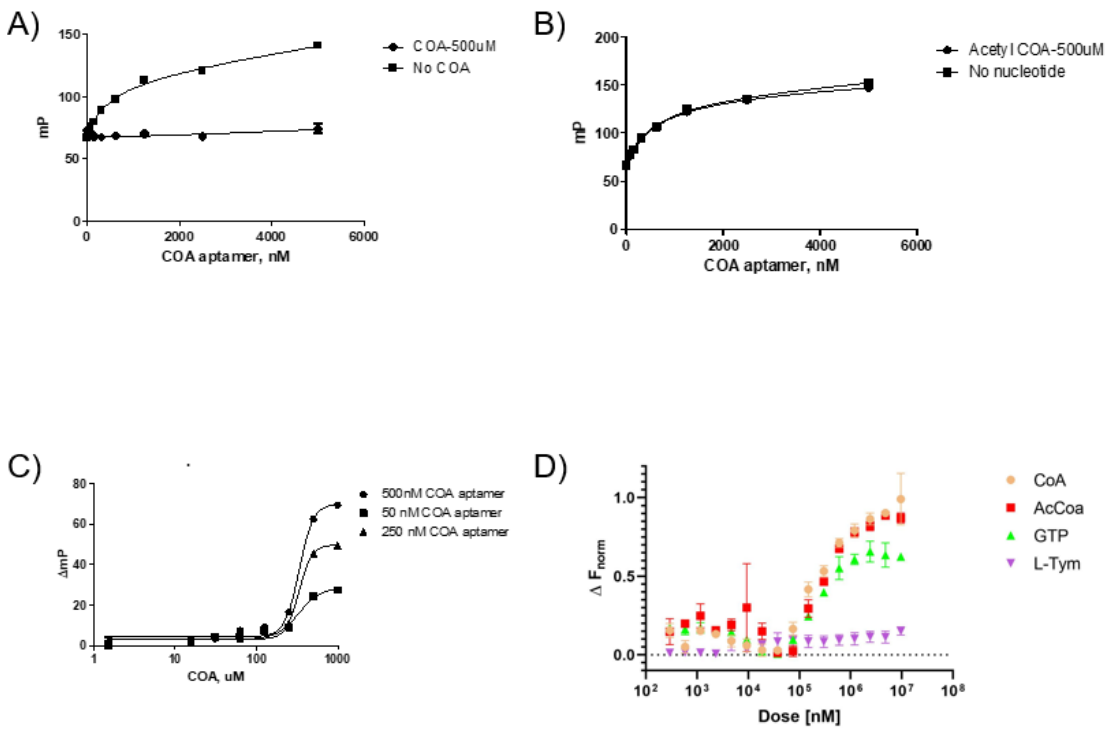


Fig 3 FAM-labeled candidate Sequence #6 binding studies. A) Fluorescence Polarization study titrating 500 µM CoA or blank into the system to measure binding. B) Fluorescence Polarization study titrating 500 µM Acetyl CoA or blank into the system to measure binding. C) B) Fluorescence Polarization study titrating 50, 250, and 500 µM CoA to measure sensitivity of the system. D) MST study using CoA, Acetyl CoA, GTP, and L-tyrosinamide to evaluate sequence's binding capabilities

occurring. Further testing used the negative selection molecule Acetly CoA and demonstrated a comparable signal to a negative blank control, supporting the candidate aptamer's specificity. Finally, testing the 50 nM candidate with three titrations of 50, 250, and 500 nM CoA showed similar EC50 values. This indicates that the concentration of the aptamer determines the binding capability and does not affect the window sensitivity. To further confirm binding specificity, candidate sequence #6 was also tested through MST, using CoA, Acetly CoA, GTP, and L-tyrosinamde. However, this data contradicts the MST experiment, showing no specificity. Candidate sequence #6 was shown to bind to CoA, Acetyl CoA, and GTP with similar affinity and not binding to L-tyrosinamide. This indicates that the sequence is recognizing and binding to the deoxyribose backbone or to the diphosphate and not the terminal thiol region. Together, this data suggests that candidate sequence #6 is capable of binding with no specificity. Because of this lack of specificity within the top candidate sequences, it was also deemed that other sequences

would have similar properties. This project served to act as a basis for SELEX protocol that would gone on to be optimize and better select for more selective sequences.

Conclusion

The high-throughput screening capabilities of a structure-switching biosensor are important for the monitoring of small molecule targets. Previous work had been completed to create a novel SS biosensor through the use of a novel PCR-based SELEX method. Upon the identification of six candidate sequences, they were developed to be tested as a fluorescence-based biosensor. While the sequences were optimized for biosensor conditions, the target molecule induced fluorescence quenching and could not be used further testing. Candidate sequence #6 was then selected for fluorescence polarization and microscale thermophoresis testing. While the aptamer was found to be capable of binding to CoA in both, non-specific binding interactions were also identified. Due to a top candidate sequence not being highly specific, no further tested were conducted. This work served to further optimize the PCR-based SELEX. The following work has since been completed and was submitted and published.

Experimental Procedures

Abbreviations: aCSF, artificial Cerebral Spinal Fluid; FAM, fluorescein; CoA, Conezyme A;

Oligonucleotide synthesis. All DNA sequences were purchased from the University of Utah DNA Synthesis Core Facility. All oligonucleotides were purified by denaturing PAGE and the desired band was excised and incubated in crush and soak buffer (300 mM ammonium acetate, 1 mM EDTA, pH 8.0) at 40°C overnight. Samples were separated from gel pieces and buffer exchanged into water using Amicon Ultra 0.5 Centrifugal Unit with Ultracel 10 membrane from EMD Millipore.

Fluorescence monitoring of hybridization and displacement. The fluorescence quenching of the FAM-labeled candidate DNA sequences in duplexes were measured by fluorescence

emission of the FAM fluorophore. Samples of 100 µL containing candidate sequences alone, as a fluorescent control sample, and duplexes over a range of concentrations (10 nM to 100 nM), at a 1:1-2 ratio were prepared using stock solutions in 50/50 aCSF/EcoRI buffer. The samples were annealed by heating to 95 ºC for 5 mins and cooling 1 ºC every 30 seconds until 25 ºC was achieved. The samples were then transferred into a 384-well flat-bottom plate for analysis on a Biotek Cytation5 plate reader. Fluorescence was then measured using an excitation at 495 nm and emission at 520 nm. The fluorescence measurement was used to plot the % Quenched using Eq. 1 using GraphPad Prism software.

$$\% \text{ Quenched} = \frac{(F_m - F_0)}{(F_m)} \times 100 \qquad (1)$$

To measure the % Displacement via fluorescence intensity, 50 µL of samples containing 2X the final concentration of candidate sequence or duplexed DNA, each in 50/50 aCSF/EcoRI buffer. 50 µL of CoA at 2X final reported quantities (0.1 to 100 µM) in 2% DMSO were prepared. After, the CoA samples transferred into the DNA or DNA:DNA samples to achieve the final reported concentrations. Samples containing DNA/CS-DNA alone and DNA:DNA were used and had 1x OTA buffer blank solutions added as controls. The samples were then transferred to a 384-well flat-bottom plate for analysis using the previously described protocol. The samples were measured at 0 and 1 hour incubation intervals. The fluorescence was then used to calculate the % Displacement using Eq. 2 as described and plotted using GraphPad Prism software.

$$\% \text{ Displacement} = \frac{(F - F_0)}{(F_m - F_0)} \times 100 \qquad (2)$$

F_m is the fluorescently labeled DNA sequence alone, F_0 is the fluorescence when candidate sequence is hybridized to a quencher-labeled capture strand, and F is the fluorescence measured when inducing a structure-switching response.

Microscale Thermophoresis binding. The

Fluorescence polarization binding. Candidate sequence #6 was tested for target binding through fluorescence polarization by our collaborators. Buffer conditions used were 100 mM NaCl, 10 mM $MgCl_2$, and 0.01% Brij-L23.

Supporting Information

Purification and characterization of DNA sequences, and additional quenching and displacement studies

Conflicts of Interest

The authors report no conflicting interests.

Acknowledgement

This work was supported by the National Science Foundation (DMR 2003987 to J.M.H.). The authors also acknowledge the Robert P. Apkarian Integrated Electron Microscopy Core and NMR Research Center at Emory University for instrument access and technical assistance. They would also like to thank Mike Hanson and the oligonucleotide and peptide synthesis facility and the University of Utah for oligonucleotide materials.

References

1. Nguyen, V.-T., Y.S. Kwon, and M.B. Gu, *Aptamer-based environmental biosensors for small molecule contaminants.* Current opinion in biotechnology, 2017. **45**: p. 15-23.
2. Huss, K.L., P.E. Blonigen, and R.M. Campbell, *Development of a Transcreener™ kinase assay for protein kinase A and demonstration of concordance of data with a filter-binding assay format.* Journal of biomolecular screening, 2007. **12**(4): p. 578-584.
3. Norskov-Lauritsen, L., A.R.B. Thomsen, and H. Brauner-Osborne, *G protein-coupled receptor signaling analysis using homogenous time-resolved forster resonance energy transfer (HTRF) technology.* International journal of molecular sciences, 2014. **15**(2): p. 2554-2572.

4. McConnell, E.M., J. Nguyen, and Y. Li, *Aptamer-based biosensors for environmental monitoring.* Frontiers in chemistry, 2020. **8**: p. 434.

5. Gajer, J., et al., *Histone acetyltransferase inhibitors block neuroblastoma cell growth in vivo.* Oncogenesis, 2015. **4**(2): p. e137-e137.

6. Judes, G., et al., *A bivalent role of TIP60 histone acetyl transferase in human cancer.* Epigenomics, 2015. **7**(8): p. 1351-1363.

7. Schneider, A., et al., *Acetyltransferases (HATs) as targets for neurological therapeutics.* Neurotherapeutics, 2013. **10**(4): p. 568-588.

8. Wang, Y., et al., *Dysregulation of histone acetyltransferases and deacetylases in cardiovascular diseases.* Oxidative medicine and cellular longevity, 2014. **2014**.

9. Gao, T., C. Yang, and Y.G. Zheng, *The fluorescence-based acetylation assay using thiol-sensitive probes*, in *Protein Acetylation*. 2013, Springer. p. 229-238.

10. Dahlin, J.L., et al., *A cell-free fluorometric high-throughput screen for inhibitors of Rtt109-catalyzed histone acetylation.* PLoS One, 2013. **8**(11): p. e78877.

11. Berndsen, C.E. and J.M. Denu, *Assays for mechanistic investigations of protein/histone acetyltransferases.* Methods, 2005. **36**(4): p. 321-331.

12. Liu, J., Z. Cao, and Y. Lu, *Functional nucleic acid sensors.* Chemical reviews, 2009. **109**(5): p. 1948-1998.

13. O'Sullivan, C.K., *Aptasensors–the future of biosensing?* Analytical and bioanalytical chemistry, 2002. **372**(1): p. 44-48.

14. Cho, E.J., J.-W. Lee, and A.D. Ellington, *Applications of aptamers as sensors.* Annual review of analytical chemistry, 2009. **2**: p. 241-264.

15. Famulok, M., J.S. Hartig, and G. Mayer, *Functional aptamers and aptazymes in biotechnology, diagnostics, and therapy.* Chemical reviews, 2007. **107**(9): p. 3715-3743.

16. Ellington, A.D. and J.W. Szostak, *In vitro selection of RNA molecules that bind specific ligands.* nature, 1990. **346**(6287): p. 818-822.

17. Tuerk, C. and L. Gold, *Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase.* science, 1990. **249**(4968): p. 505-510.

18. McKeague, M. and M.C. DeRosa, *Challenges and opportunities for small molecule aptamer development.* Journal of nucleic acids, 2012. **2012**.

19. Nutiu, R. and Y. Li, *Structure-switching signaling aptamers.* Journal of the American Chemical Society, 2003. **125**(16): p. 4771-4778.

20. Jhaveri, S.D., et al., *Designed signaling aptamers that transduce molecular recognition to changes in fluorescence intensity.* Journal of the American Chemical Society, 2000. **122**(11): p. 2469-2473.

21. Chen, J., et al., *A simple and rapid biosensor for ochratoxin A based on a structure-switching signaling aptamer.* Food Control, 2012. **25**(2): p. 555-560.

22. Liu, F., et al., *A label-free aptasensor for ochratoxin A detection based on the structure switch of aptamer.* Sensors, 2018. **18**(6): p. 1769.

23. Feagin, T.A., et al., *High-throughput enantiopurity analysis using enantiomeric DNA-based sensors.* Journal of the American Chemical Society, 2015. **137**(12): p. 4198-4206.

24. Nutiu, R. and Y. Li, *In vitro selection of structure-switching signaling aptamers.* Angewandte Chemie International Edition, 2005. **44**(7): p. 1061-1065.

25. Stoltenburg, R., N. Nikolaus, and B. Strehlitz, *Capture-SELEX: selection of DNA aptamers for aminoglycoside antibiotics.* Journal of analytical methods in chemistry, 2012. **2012**.

26. Sanford, A.A., et al., *RE-SELEX: Restriction Enzyme-Based Evolution of Structure-Switching Aptamer Biosensors.* 2021.

Chapter 7: Conclusions and Future Perspectives

Nature has evolved two primary languages in the form of nucleic acids and amino acids. Scientists have introduced novel variants and modifications to these systems, further expanding upon them, with one prominent unnatural nucleic acid being PNA. This pseudopeptide compound would further be explored for its sequence recognition properties through applications such as antisense interactions, sequence detection assays, and drug delivery methodology. However, further exploration of PNA properties revealed that through synthetic design, modifications could be introduced to alter their capabilities. One of the more prominent modifications that could be introduced at the α-, β-, γ-position. This created a single monomer unit that contained could further be modified.

In chapter 1, we described nucleic acid functionality and the introduction of PNA to function in similar capacities. As the properties of the modifications were elucidated, it was determined that modifications at the γ-positon not only retained PNA functionality, but improved on it, and an amino acid residue could be chosen. This would allow for the development of a bilingual biopolymer that could employ both sequences. Our laboratory explored this initial concept and explored its unique properties for characterization. We found that it was capable of utilizing the amino acid code to drive for assembly while also being capable of target sequence recognition, inducing disassembly. With this discovery, the Heemstra laboratory is interested in expanding on the novel biopolymer's capabilities.

7.1 Nucleic Acids and Bilingual Biopolymers

Previous work using nucleic acids have elucidated their properties for various applications. Some of these works include the development of aptamers, single-stranded nucleic acid sequences that can undergo a target recognition and binding event, and of biosensors that can utilize a readable output to use for the detection and quantification of a target of interest. Following, the introduction and framework of PNA was shown as a lead up to the development of the novel bilingual PNA

biopolymer that our research discovered. By understanding the previously established nucleic acid systems, we used these techniques to assess the biopolymer's range of applications such as being adapted for a fluorescence intensity-based biosensor.

7.2 Stimuli-Responsive Assembly of Bilingual Peptide Nucleic Acids.7.3 Nucleic Acids and Bilingual Biopolymers

Through this work, we expanded the capabilities of the PNA bilingual biopolymer to introduce a stimuli-responsive mechanism and to act as a reporter for a target sequence. PWe envisioned the use of the biopolymer to be used in drug delivery systems. As such, to further develop this system and the biopolymer's applications, a two PNA amphiphile system could be developed to work in tandem. Both sequences would be developed to respond to the target of interest and be capable of self-assembling individually and jointly. However, the hydrophobic terminal end of Sequence #1 would contain a reactive compound and the same position on Sequence #2 would be an inactive moiety, likely a Staudinger ligation reaction. With both sequences in the same system, upon the introduction of the external stimulus, both sequences would be released and being assembling. The terminal ends would increase the effective concentration of the two molecules, inducing release of the active molecule as a response. This would further allow for the use of the biopolymer's assembly capabilities to be harnessed for effective stimuli response for drug release and biosensor readouts.

7.3 Elucidating the Structural Influence of Bilingual PNA Biopolymers7.5 Nucleic Acids and Bilingual Biopolymers

Upon the initial design and characterization of the PNA-A bilingual biopolymer, the Heemstra laboratory was interested in further elucidating the roles of the two sequences. After testing, we found that the nucleic acid sequence had a larger role in affecting the size of the assemblies formed that expected. Similar retention of most other properties as the previous amphiphile,

regardless of modification. Further studies that could be explored is the introduction of other natural residues that could alter functionality such as cysteine, as it has been used for the formation of disulfide bonds. Using this, the disassembly of the biopolymer could be prevented, thus allowing for the retention of assemblies to be used as a marker for a nucleic acid target of interest. Arginine would also be a residue of interest as previous research in our lab has been for development of glyoxal caging. Using this methodology, the guanidinium moiety would be caged, acting as a hydrophobic residue, and driving for assembly. Upon decaging, the released residue would revert to a hydrophilic state, altering the structure. Of note, only one variation of the hydrophobic/hydrophilic balance was tested, involving two hydrophobic and two hydrophilic residues on terminal ends. Thus, by tipping the balance, it could change formation of the architecture to broaden the drug delivery capabilities.

7.4 Synthesis and development of a Triplex Forced Intercalation Peptide Nucleic Acid Probe for the Detection of Adenosine-to-Inosine Modification in Hairpin RNA*

Previous A-to-I FIT PNA probe work lead to the fluorescence-based detection of Inosine in linear RNA sequences using Thiazole Orange and attempts to use this in hairpin HER1 RNA proved to be difficult. It was also proposed that through the use of triplex forming monomers, a triplex FIT probe could be developed that could bind using Hoogsteen base pairing and detect for the change. Our results concluded that it was not capable of detecting for I over A. Future endeavors for this work would be through the application of γ-modified PNA monomers, specifically using guanidiniums moieties through the use of GPNAs. These monomers in particular have been found to be capable of DNA duplex invasion. Thus, the previously tested and confirmed Watson-Crick-Franklin recognition would be retained and I edits would likely be detected in hairpin sequences.

7.5 Development of a Structure-switching DNA:PNA Biosensor for Ochratoxin A

Our work for the development of a mixed nucleic acid Biosensor of Ochratoxin A was met with mixed results. While a PNA:DNA system was capable of achieving a dose-dependent response in the presence of OTA through the use of γ-modified PNA monomers, it would not a suitable system for stimuli-responsive assembly. The primary reason for this was derived from the biosensor signal being instable in the presence of a lower ratio of PNA than DNA. Thus, assembly formation would likely occur prior to the introduction of OTA at its tested concentration. One possible method to overcome this would be through the exploration of novel bilingual biopolymers that contain high critical micelle concentration (CMC) values, preventing assembly prior to the additional release of amphiphile. As well, a different biosensor could be adapted that has an initially lower aptamer:complementary sequence ratio due to the OTA biosensor being lowered from a 1:3 to a 1:2 ratio through the use of PNA. If a different system has a lower ratio, the introduction of PNA would likely see a similar reduction in ratio, decreasing the presence of unbound PNA and preventing non-stimulated assembly formation.

7.6 Optimization of a CoA Structure-Switching Biosensor*

Previous reports have demonstrated that biosensors are vital for the detection and quantification of various targets of interest. However, SS biosensors are difficult to develop as previous SELEX methods did not select for this property. Thus, the Heemstra laboratory worked to develop a novel SELEX method to select for structure-switching properties. Upon the identification of sequence candidates, biosensor and binding capabilities were evaluated to determine their efficacies. Unfortunately, target-specific binding was not capable of being reproduced in various systems and biosensor results were not promising. Since the testing of the sequences, a novel RE-SELEX method was developed that was capable of developing novel SS biosensors, using kanamycin A as a target. Thus, to target CoA, the RE-SELEX method could be employed to select for candidate sequences and undergo biosensor development.

Appendix A: Omitted Data from Chapter 2

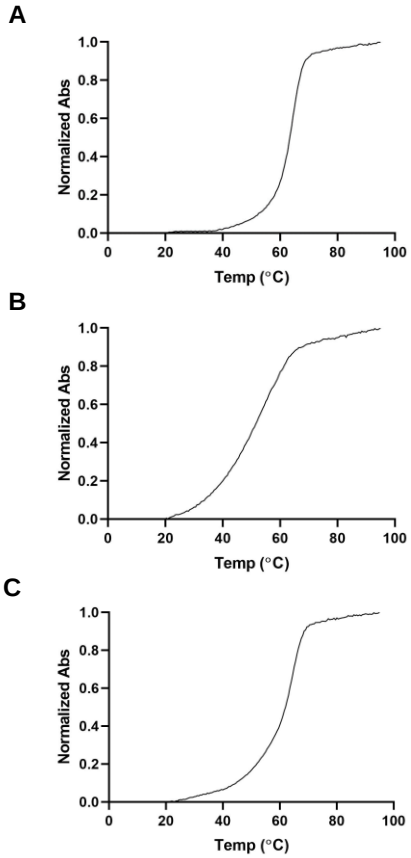


Fig S1 UV-Vis melting curves using normalized absorbance. Samples were prepared at 3μM in 1X PBS. (A) MS-D:RS-D. (B) MS-D:RS-R. (C) MS-D-m1:RS-D-m1.

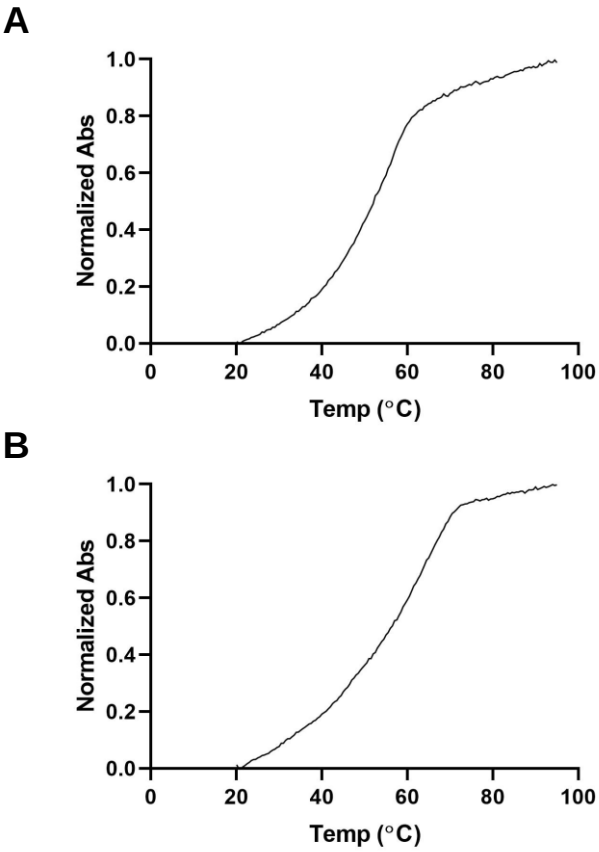


Fig S2 UV-Vis melting curves using normalized absorbance. Sample were prepared at 3μM in 1X PBS. (A) MS-R:RS-D. (B) MS-R:RS-R.

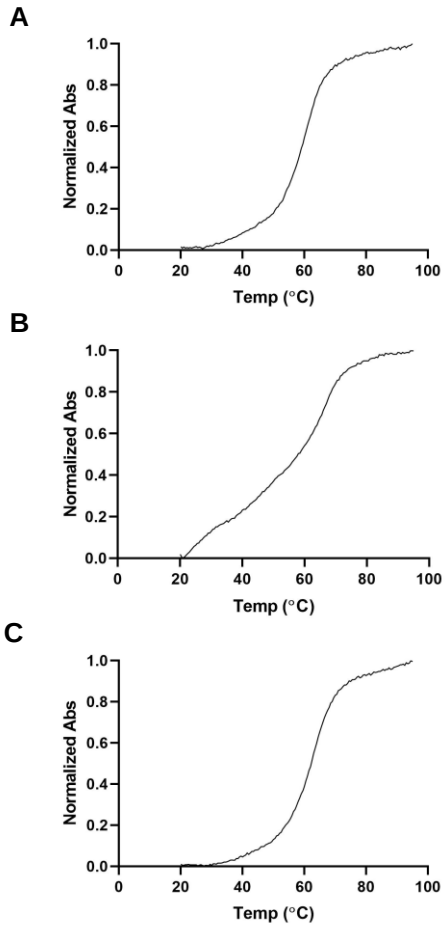


Fig S3 UV-Vis melting curves using normalized absorbance. Sample were prepared at 3μM in 1X PBS. (A) PNA-C1-FAM:MS-D. (B) PNA-C1-FAM:MS-R. (C) PNA-C1-FAM:MS-D-m1.

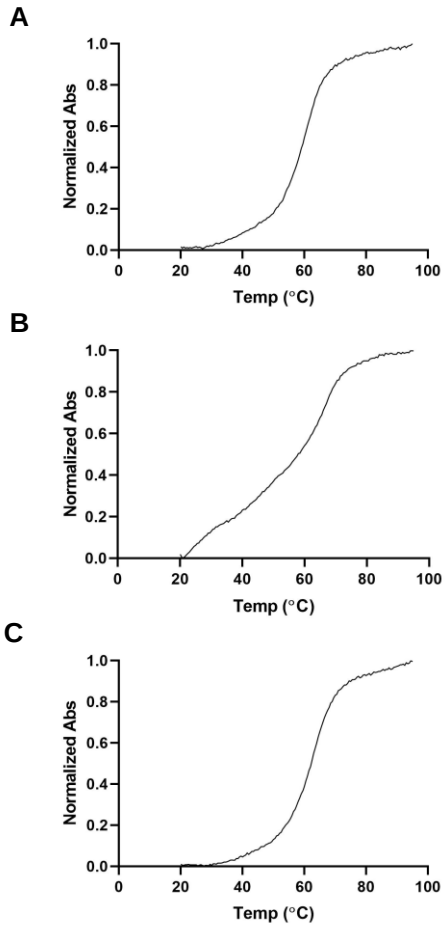


Fig S4 UV-Vis melting curves using normalized absorbance. Sample were prepared at 3μM in 1X PBS. (A) PNA-A1-FAM:MS-D. (B) PNA-A1-FAM:MS-R. (C) PNA-A1-FAM:MS-D-m1

Table S1 Melting temperature measurements using normalized absorbance

Sequences	Melt Temp (°C)
MS-D:RS-D	62.9 ± 0.2
MS-D:RS-R	50.8 ± 0.2
MS-D-m1:RS-D-m1	61.4 ± 1.3
MS-R:RS-D	51.3 ± 0.5
MS-R:RS-R	56.1 ± 1.2
PNA-C1-FAM:MS-D	58.9 ± 0.1
PNA-C1-FAM:MS-R	58.3 ± 0.6
PNA-C1-FAM:MS-D-m1	61.8 ± 0.2
PNA-A1-FAM:MS-D	56.6 ± 1.1
PNA-A1-FAM:MS-R	54.9 ± 1.1
PNA-A1-FAM:MS-D-m1	53.4 ± 0.8

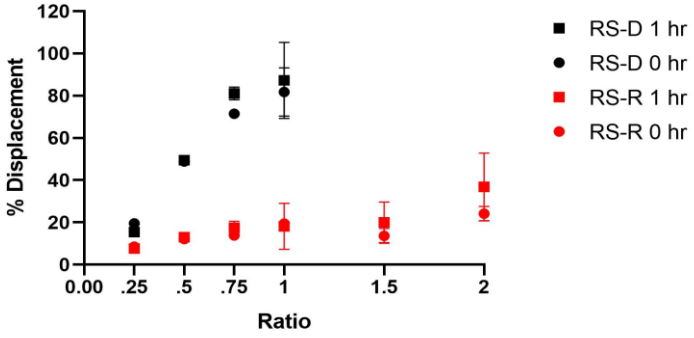


Fig S5 % Displacement from PNA-C1-FAM:MS-D using RS-D and RS-R, respectively, was evaluated in relation to the stoichiometry of the system. PNA:MS-D system was used at 3 µM in 1xPBS. RS-R was tested up to a stoichiometry of 1:2 duplex:RS-R. Error bars represent standard error (n=3).

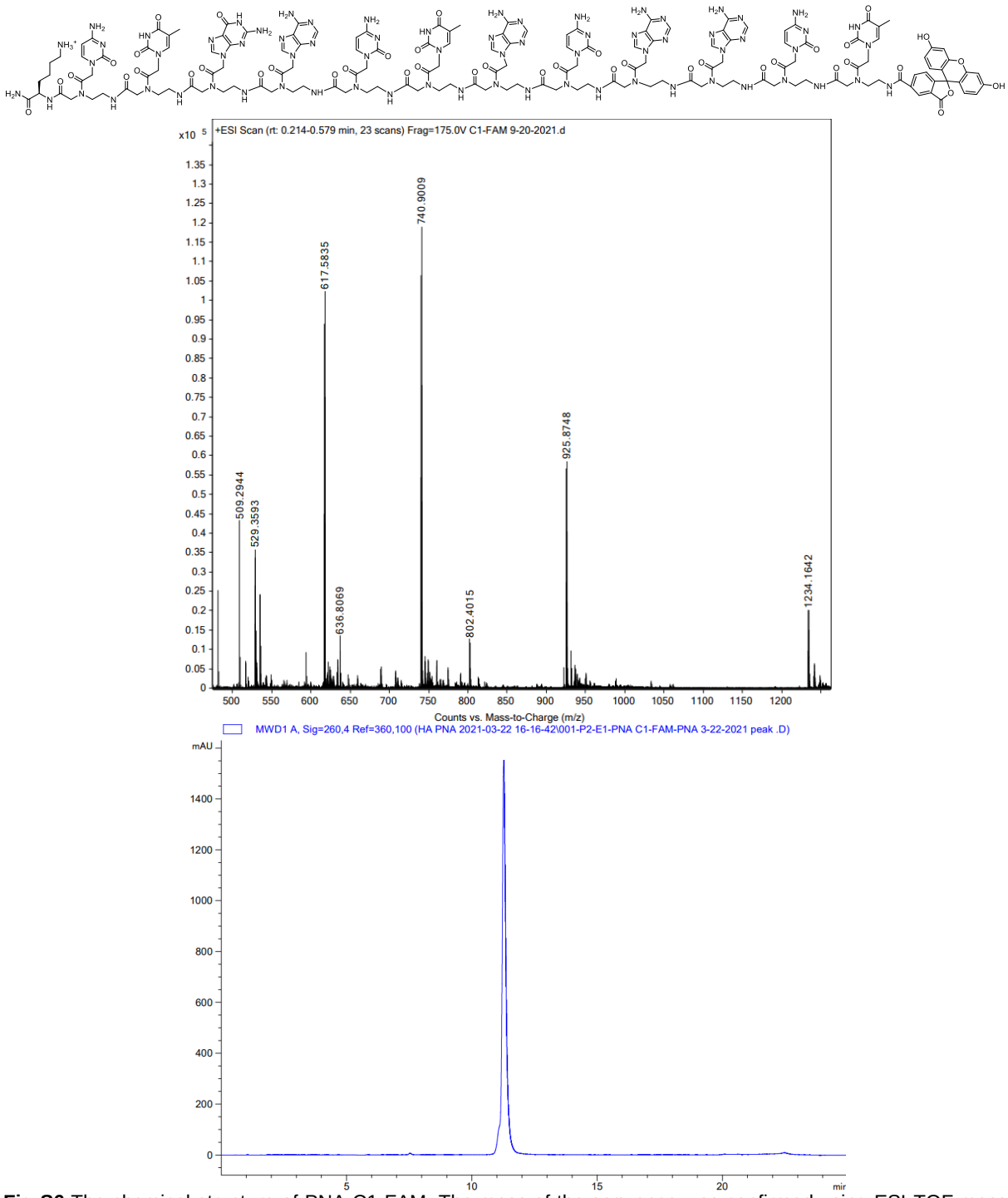


Fig S6 The chemical structure of PNA-C1-FAM. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

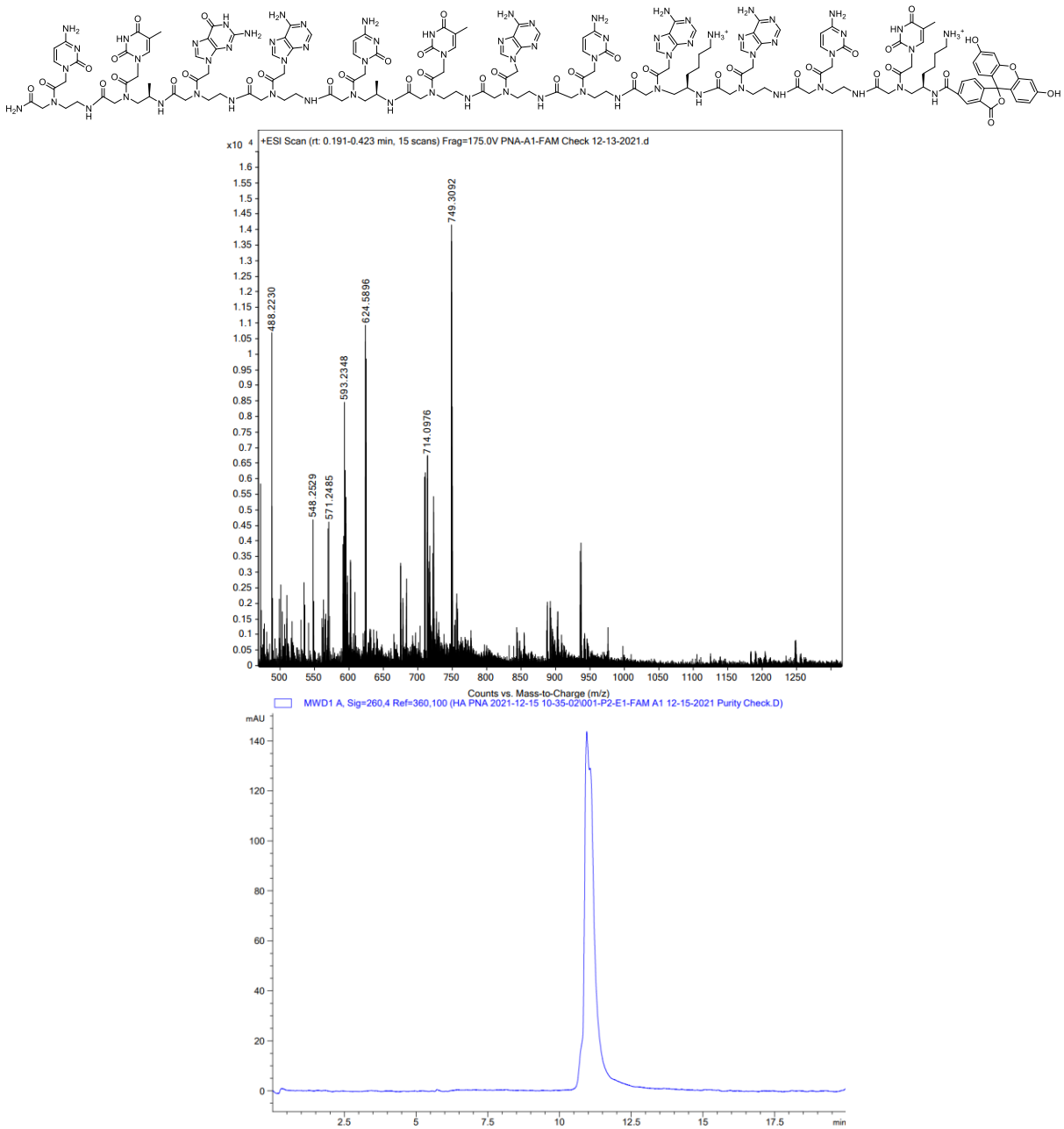


Fig S7 The chemical structure of PNA-A1-FAM. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

Table S2 PNA Sequence, expected mass, and found mass

Strand	Sequence	Expected Mass $(M+5)^{5+}$	Found Mass $(M+5)^{5+}$
PNA-C1-FAM	C-C T G A C T A C A A C T FAM-N	740.3720	740.9009
PNA-A1-FAM	C-C T_A G A C_A T A C A_K A C T_K FAM-N	748.9794	749.3092

"FAM" denotes 5-carboxyfluorescein. Subscripts denote the amino acid residues incorporated at the γ-position. Masses were confirmed by ESI-TOF mass spectrometry.

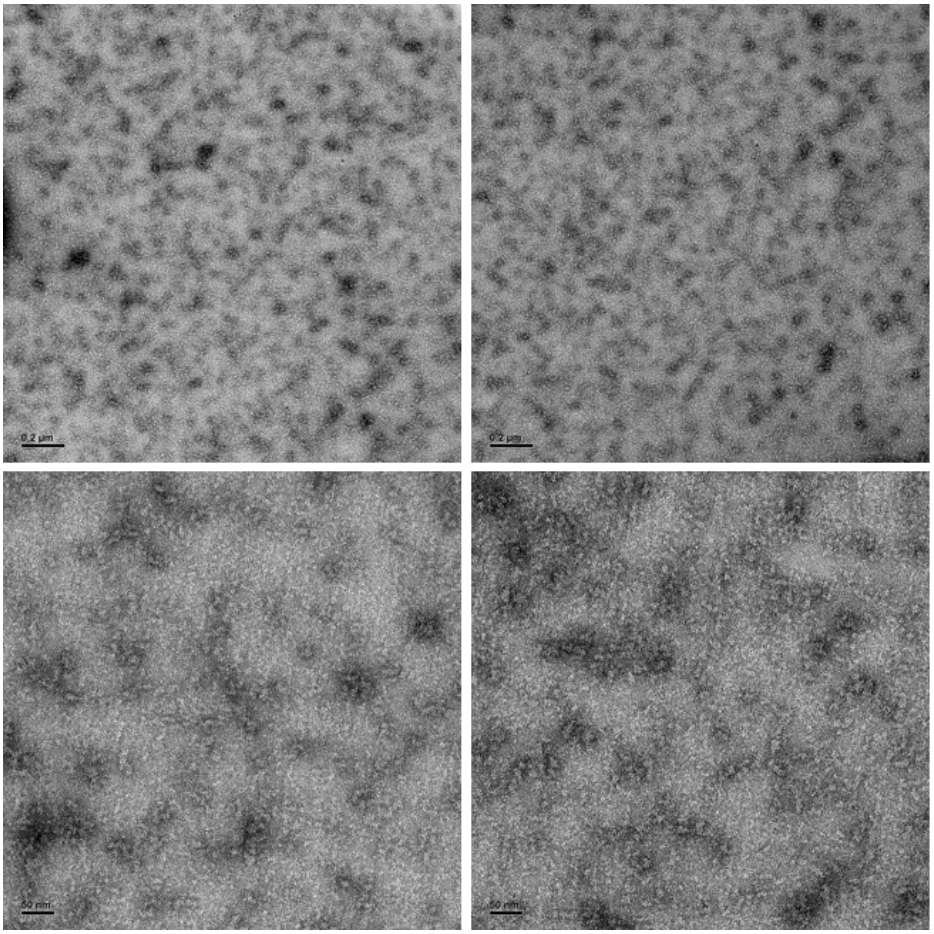


Fig S8 TEM images of PNA-A1-FAM 100 µM in 1X PBS. Top row scale bar = 2000 nm. Bottom row scale bar = 50 nm.

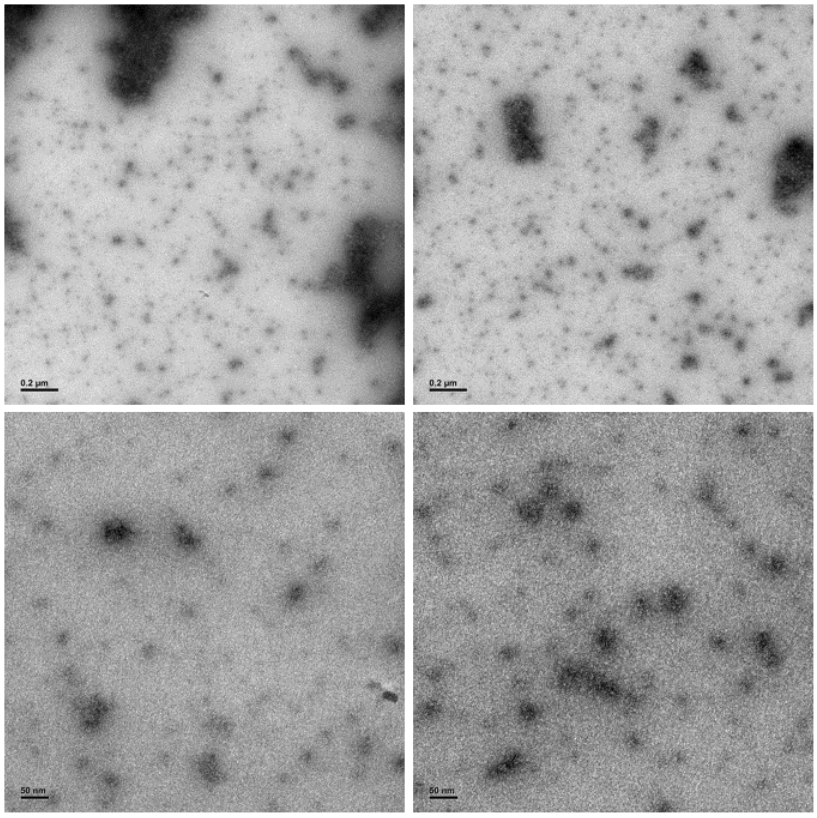


Fig S9 TEM images of PNA-A1-FAM + MS-D 100 µM in 1X PBS. Top row scale bar = 2000 nm. Bottom row scale bar = 50 nm.

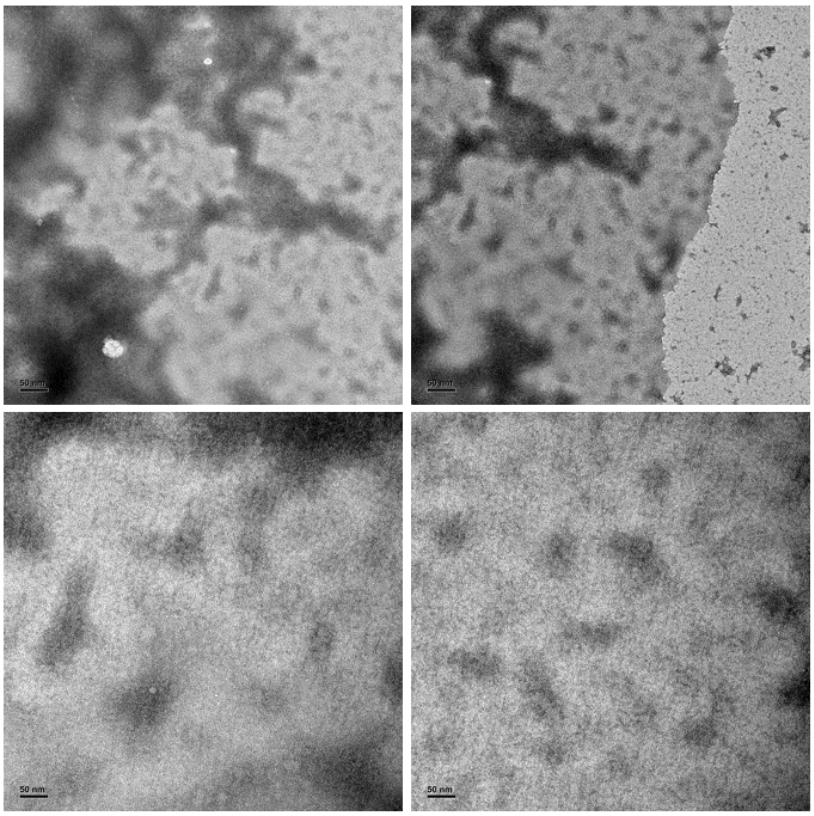


Fig S10 TEM images of PNA-A1-FAM + MS-D + RS-D 100 µM in 1X PBS. Top row scale bar = 2000 nm. Bottom row scale bar = 50 nm.

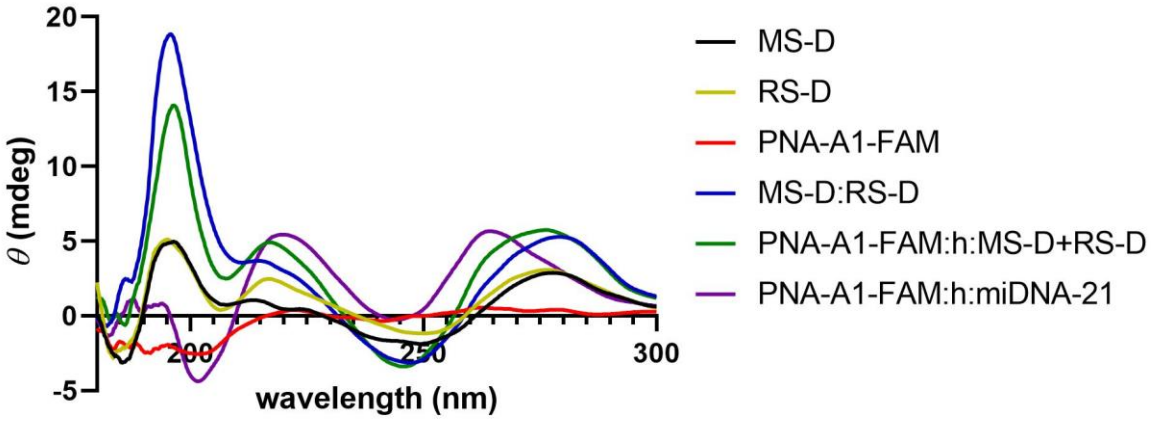


Fig S11 CD spectroscopy demonstrating the change in maxima and minima upon the addition of RS-D-0. All samples were prepared using 100 µM PNA and DNA in 1x PBS.

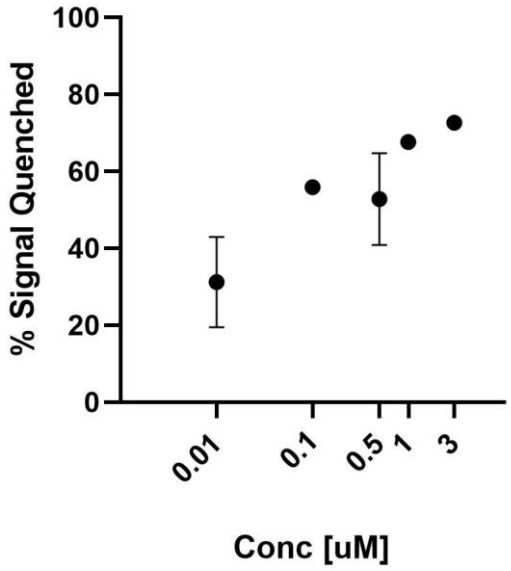


Fig S12 Hybridization of MS-D to PNA-A1-FAM was monitored using fluorescence quenching. Error bars represent standard error (n=3).

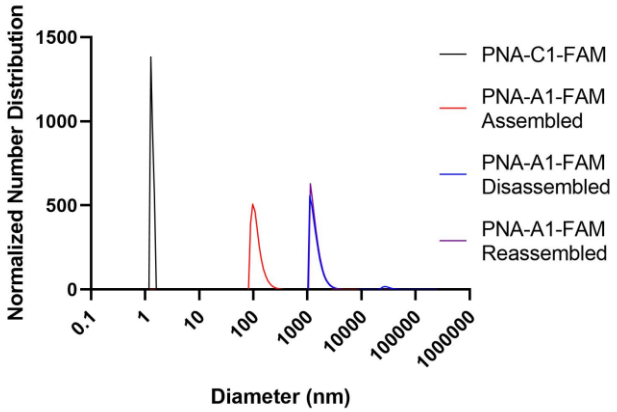


Fig S13 Normalized size distribution of PNA assemblies. Samples tested at 200 μM in 1x PBS. Average diameter of particles of PNA-C1-FAM = 1.5 ± 0.3 nm and 80.4 ± 22.8 nm; PNA-A1-FAM (assembled) = 119.1 ± 34.0 nm. PNA-A1-FAM (disassembled) = 1438.9 ± 400.2 nm. PNA-A1-FAM (reassembled) = 1470.5 ± 417.2 nm.

Appendix B: Omitted Data from Chapter 3

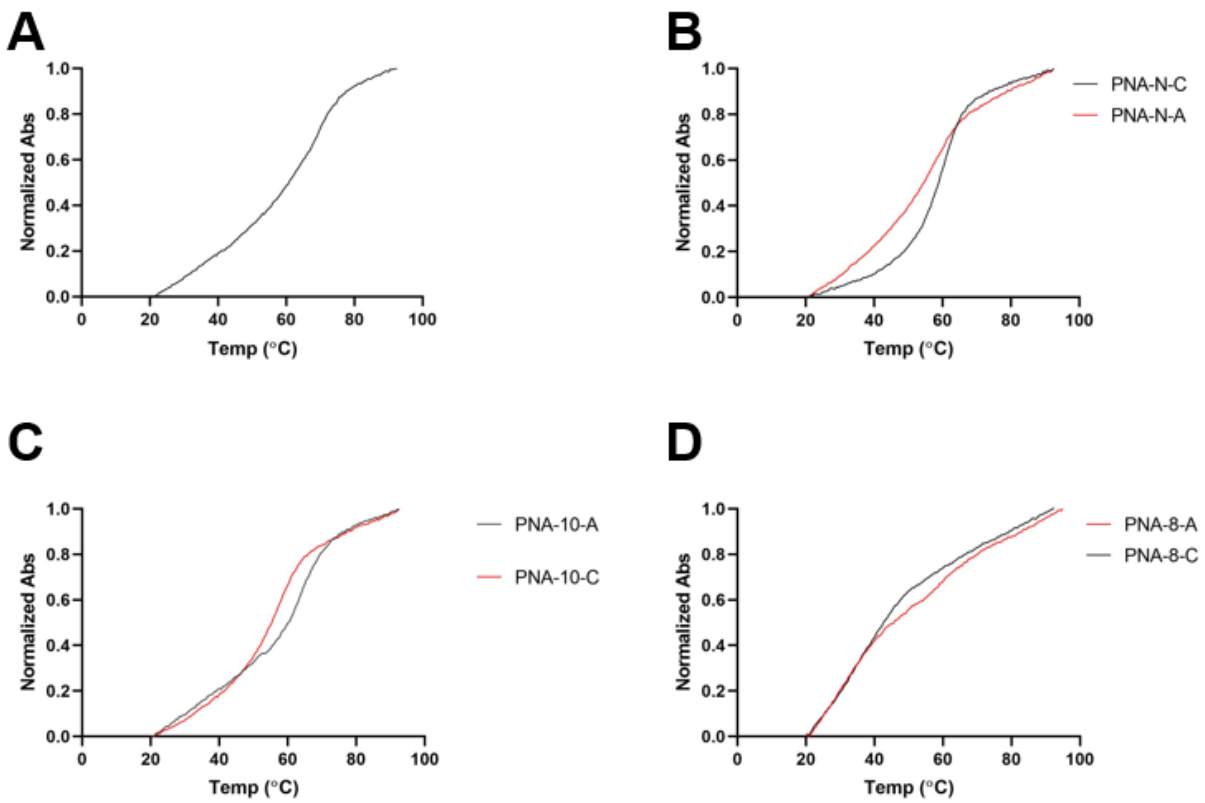


Fig S1 UV-Vis melting curves using normalized absorbance. PNA samples were prepared at 3 µM in 1X PBS with 3 µM of DNA, final concentrations. (A) PNA-B-A. (B) PNA-N-A and PNA-N-C. (C) PNA-10-A and PNA-10-C. (D) PNA-8-A and PNA-8-C.

Table S1 Melting temperature measurements using normalized absorbance using DNA

Strand	Melt Temp (°C)
PNA-B-A	61.96 ± 0.92
PNA-N-C	58.11 ± 0.09
PNA-N-A	55.74 ± 0.38
PNA-10-C	54.32 ± 0.33
PNA-10-A	60.1 ± 0.73
PNA-8-C	45.97 ± 1.4
PNA-8-A	48.08 ± 1.11

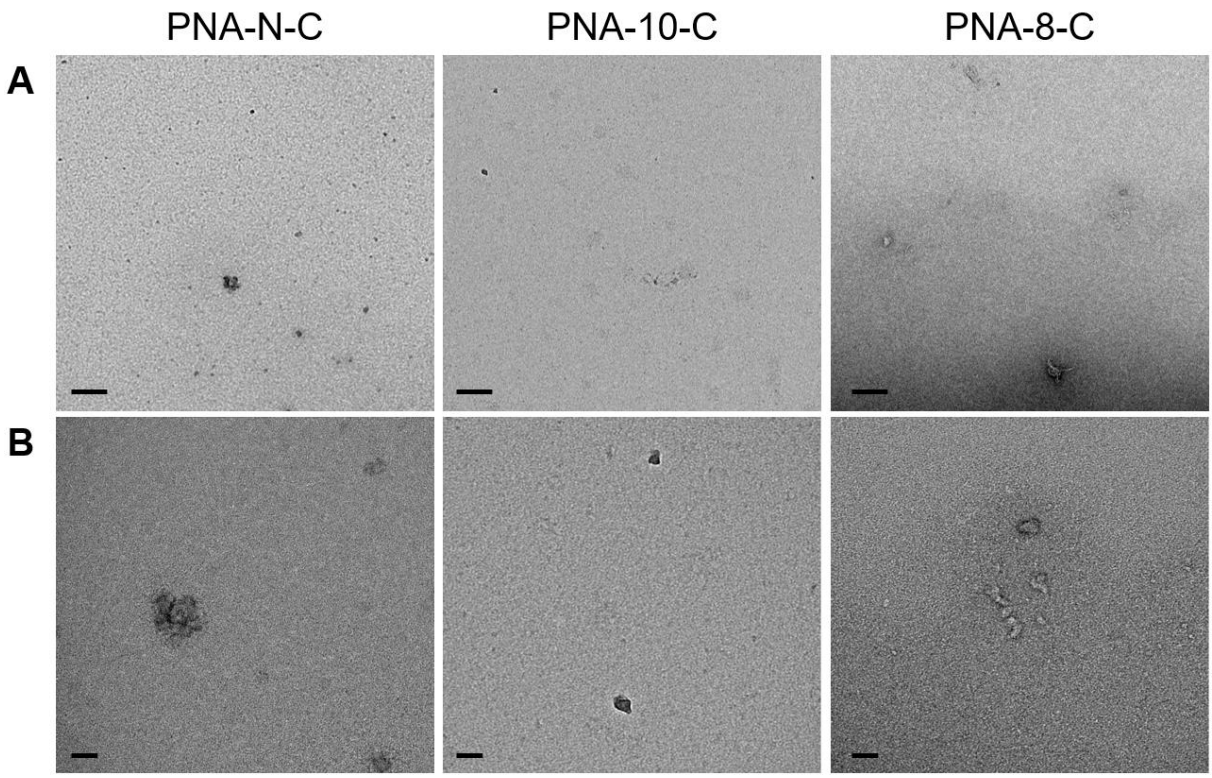


Fig S2. TEM images of PNA-N-C, PNA-10-C, and PNA-8-C assembled in water at 100 µM. A) Scale bar = 200 nm. B) Scale bar = 50 nm.

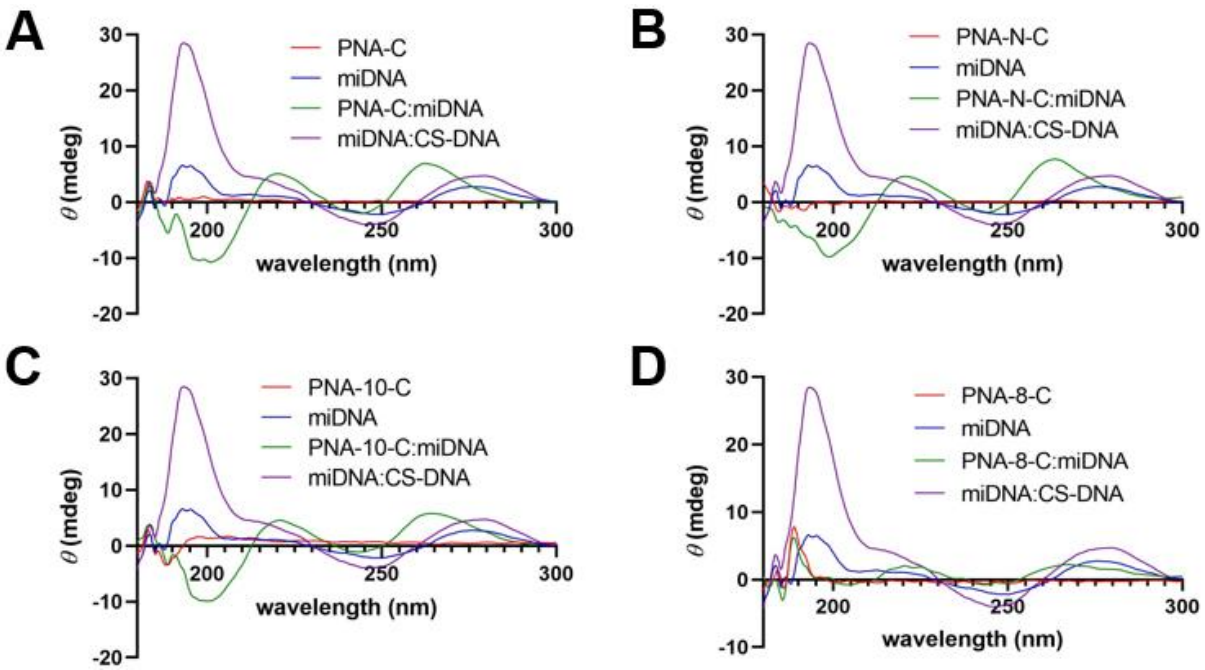


Fig S4 CD Spectroscopy of A) PNA-C B) PNA-N-C C) PNA-10-C, and D) PNA-8-C in 1x PBS at 100 µM to demonstrate the change in maxima and minima up on the addition of DNA.

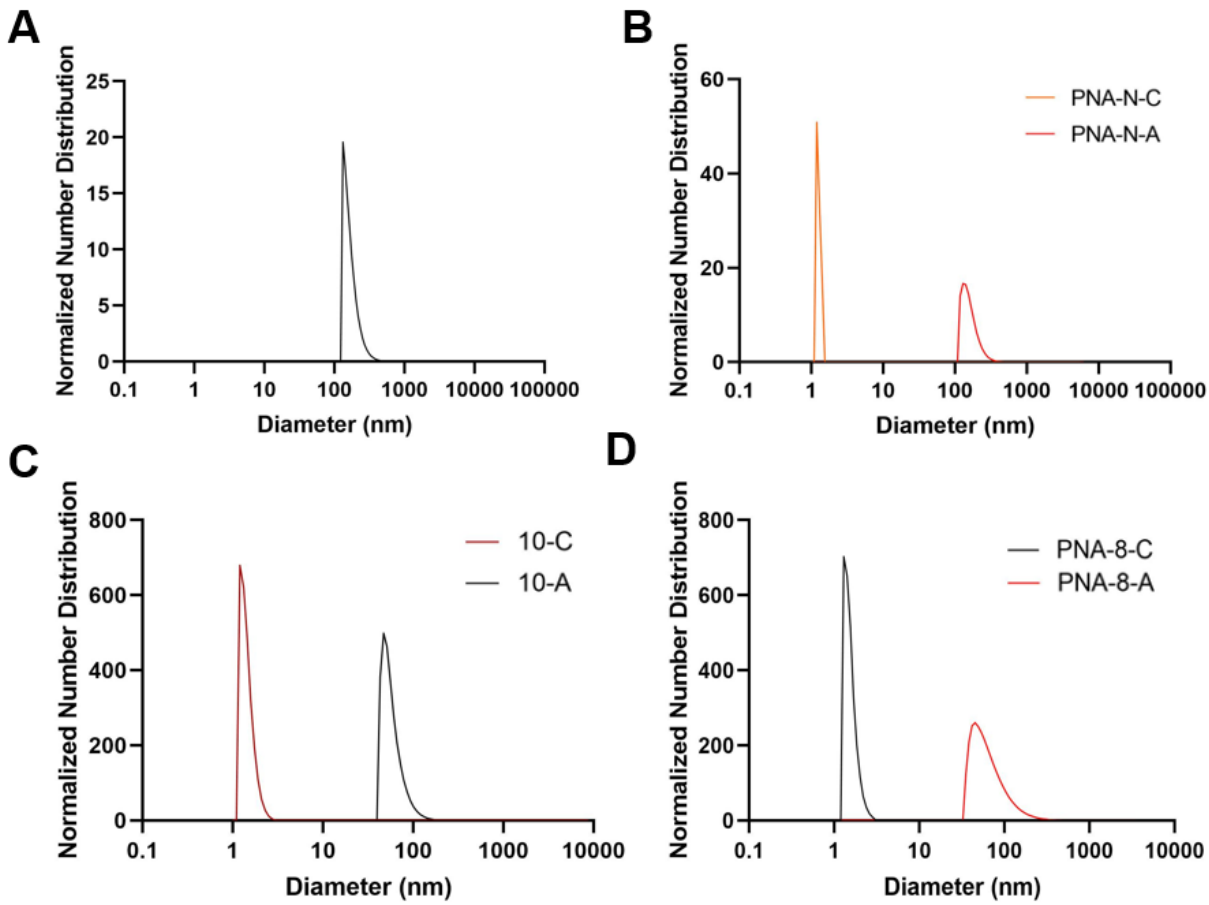


Fig S5 Normalized size distribution of PNA assemblies. Samples tested at 200 µM in 1x PBS.

Table S2 Average Size of PNA Assembly Particles using DLS

Strand	Diameter
PNA-B-A	107.8 ± 47.0 nm
PNA-N-C	1.3 ± 0.1 nm 127.1 ± 36.3 nm
PNA-N-A	161.7 ± 44.4 nm
PNA-10-C	1.6 ± 0.3 nm 191.9 ± 79.1 nm
PNA-10-A	58.9 ± 17.8 nm
PNA-8-C	1.5 ± 0.3 nm 80.4 ± 22.8 nm
PNA-8-A	70.0 ± 43.1 nm

All sizes reported

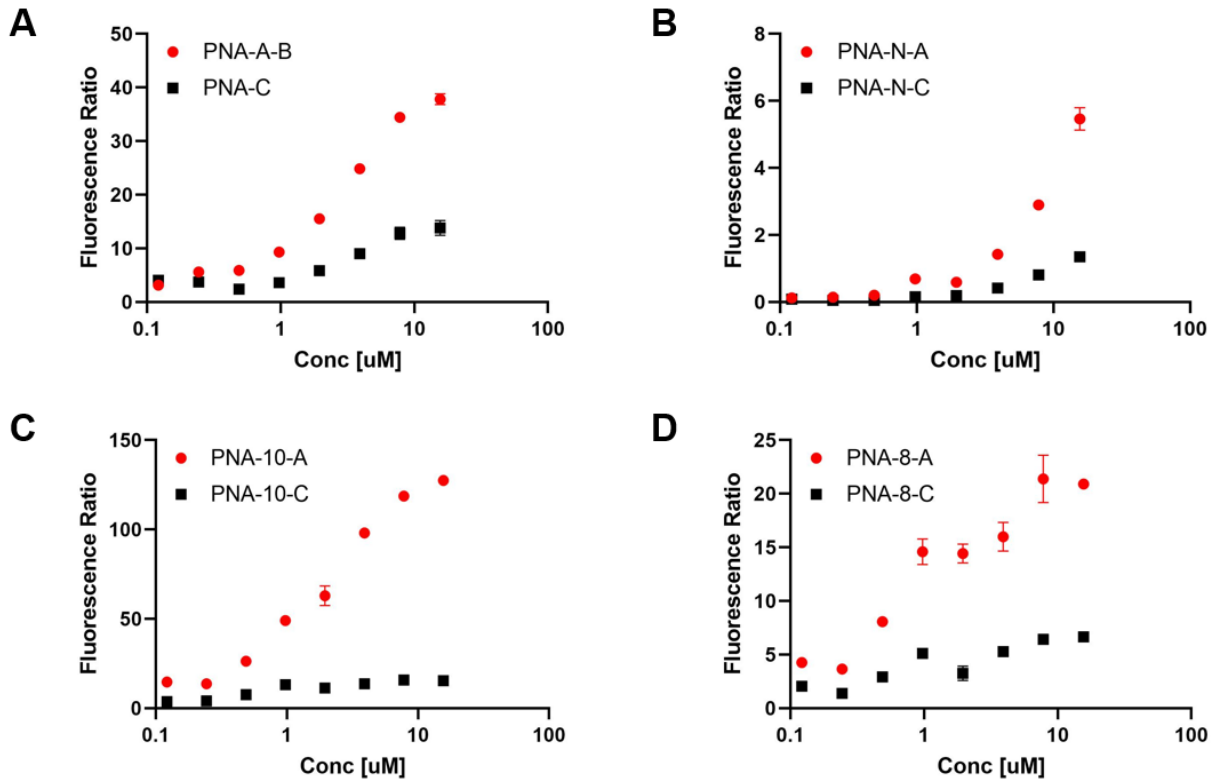


Fig S6 Fluorescence Ratio studies as a function of concentration, testing amphiphilic and control PNA sequences in. Top left: PNA-A-B and PNA-C; Top right: PNA-N-A and PNA-N-C; Bottom left: PNA-10-A and PNA-10-C; Bottom right: PNA-8-A and PNA-8-C. Samples tested in 1%DMSO/water.

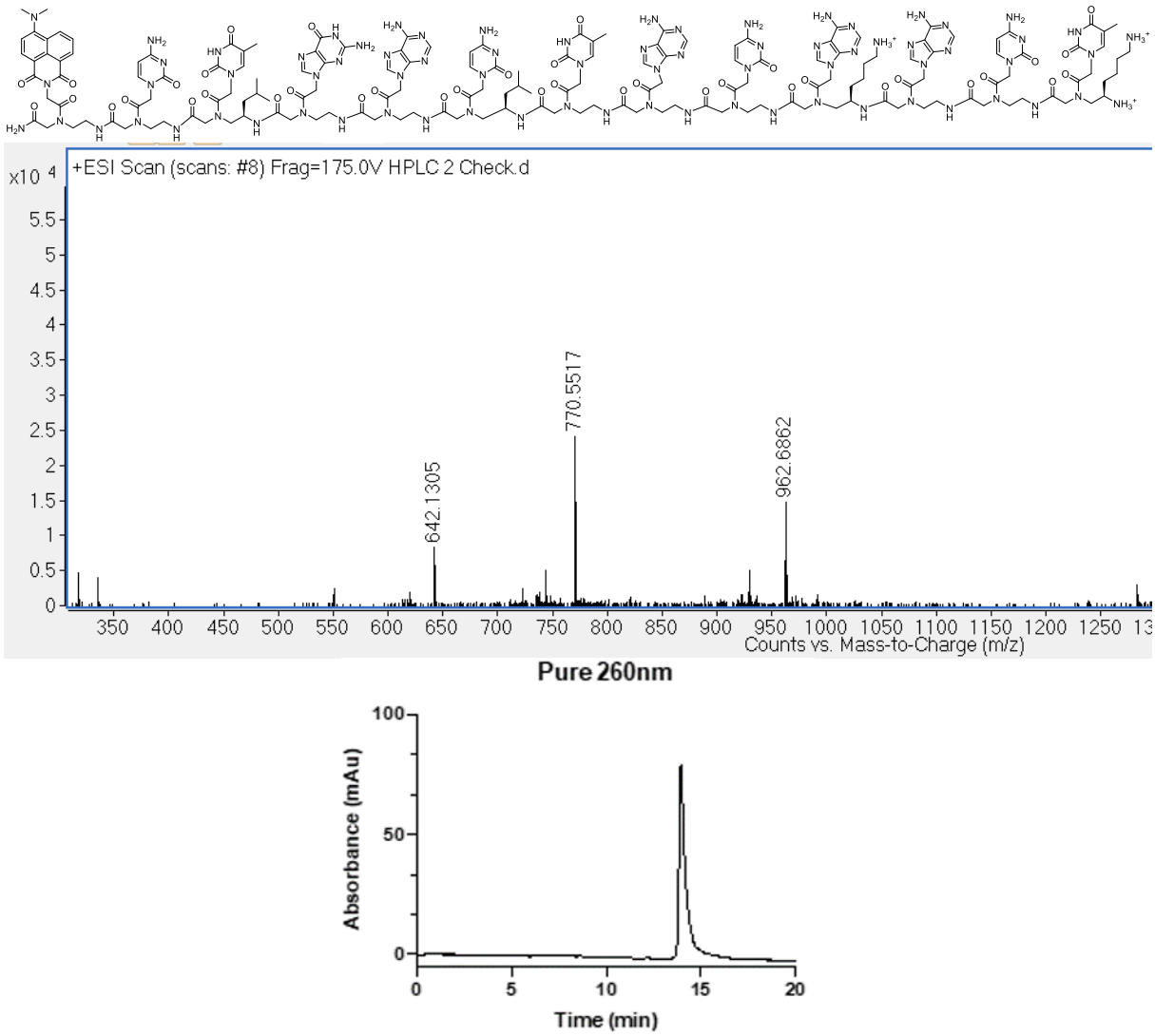


Fig S7 The chemical structure of PNA-B-A. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

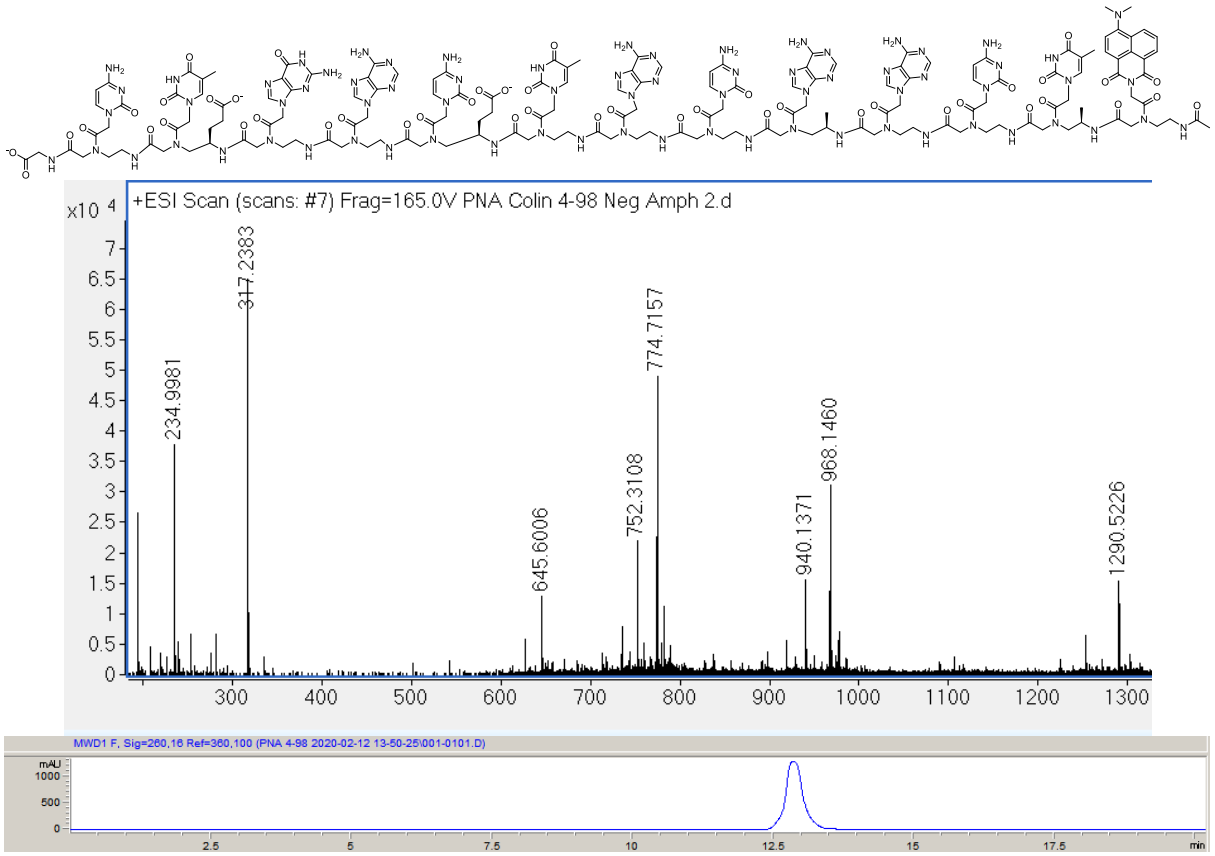


Fig S8 The chemical structure of PNA-N-A. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

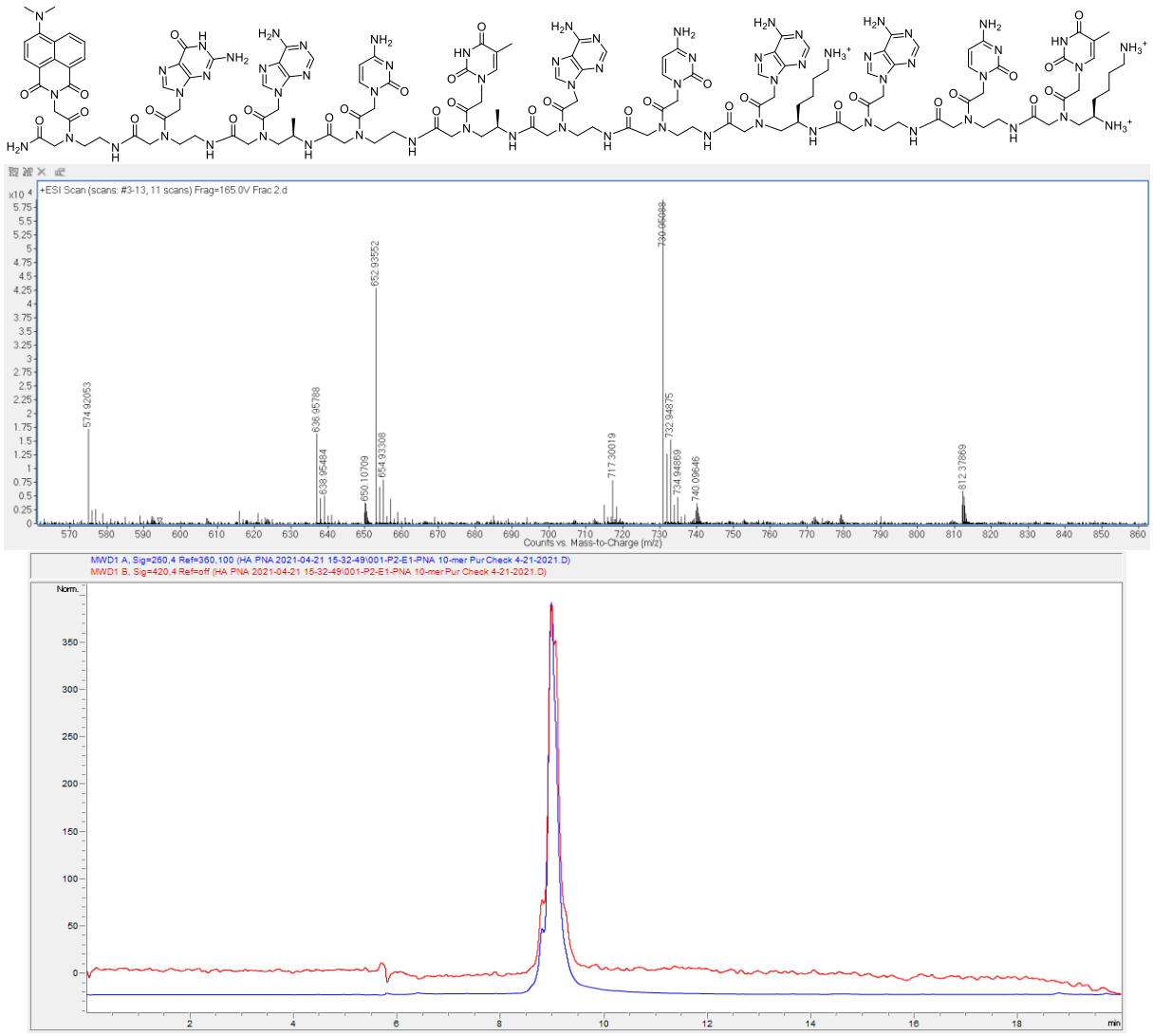


Fig S9 The chemical structure of PNA-10-A. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

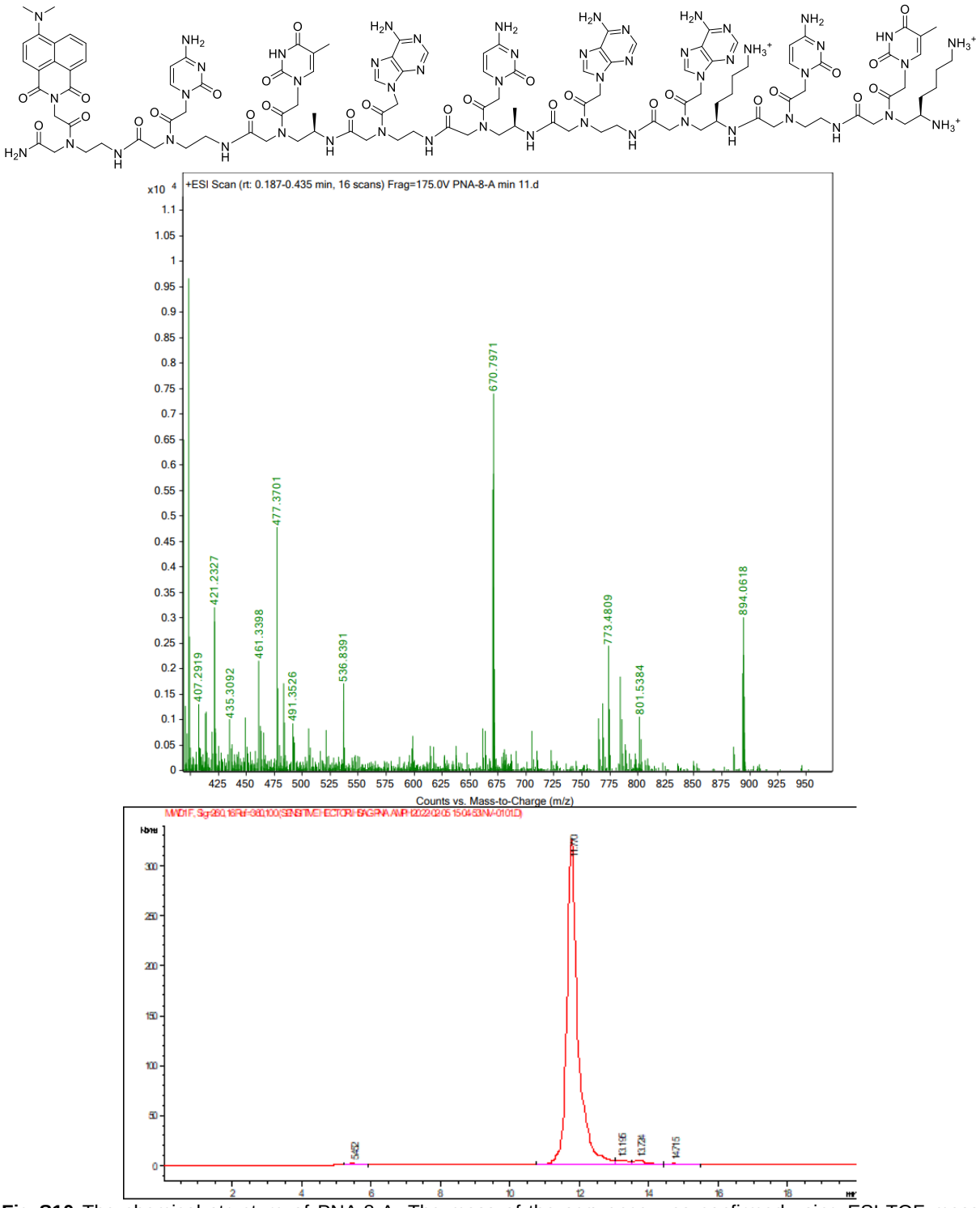


Fig S10 The chemical structure of PNA-8-A. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

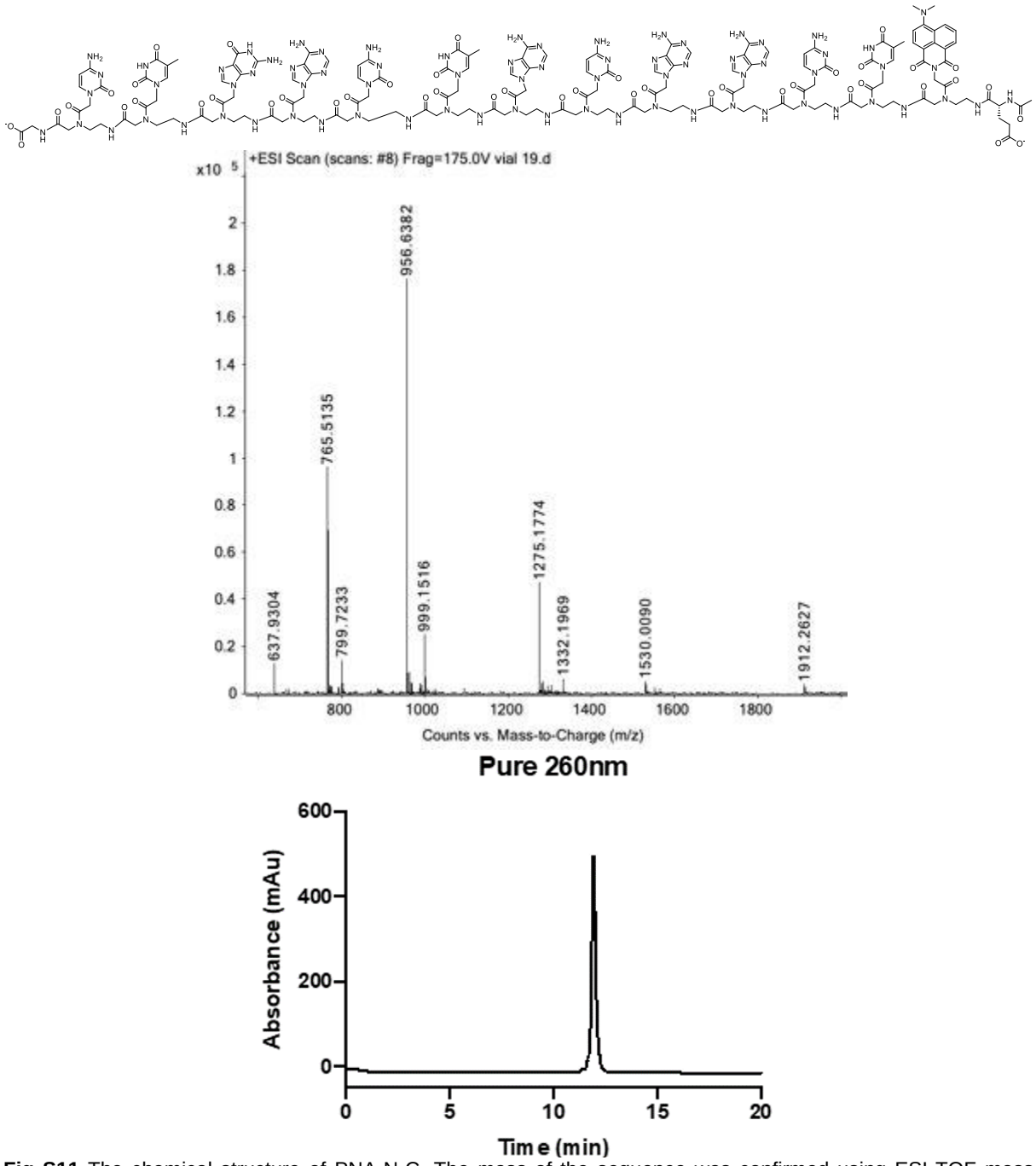


Fig S11 The chemical structure of PNA-N-C. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

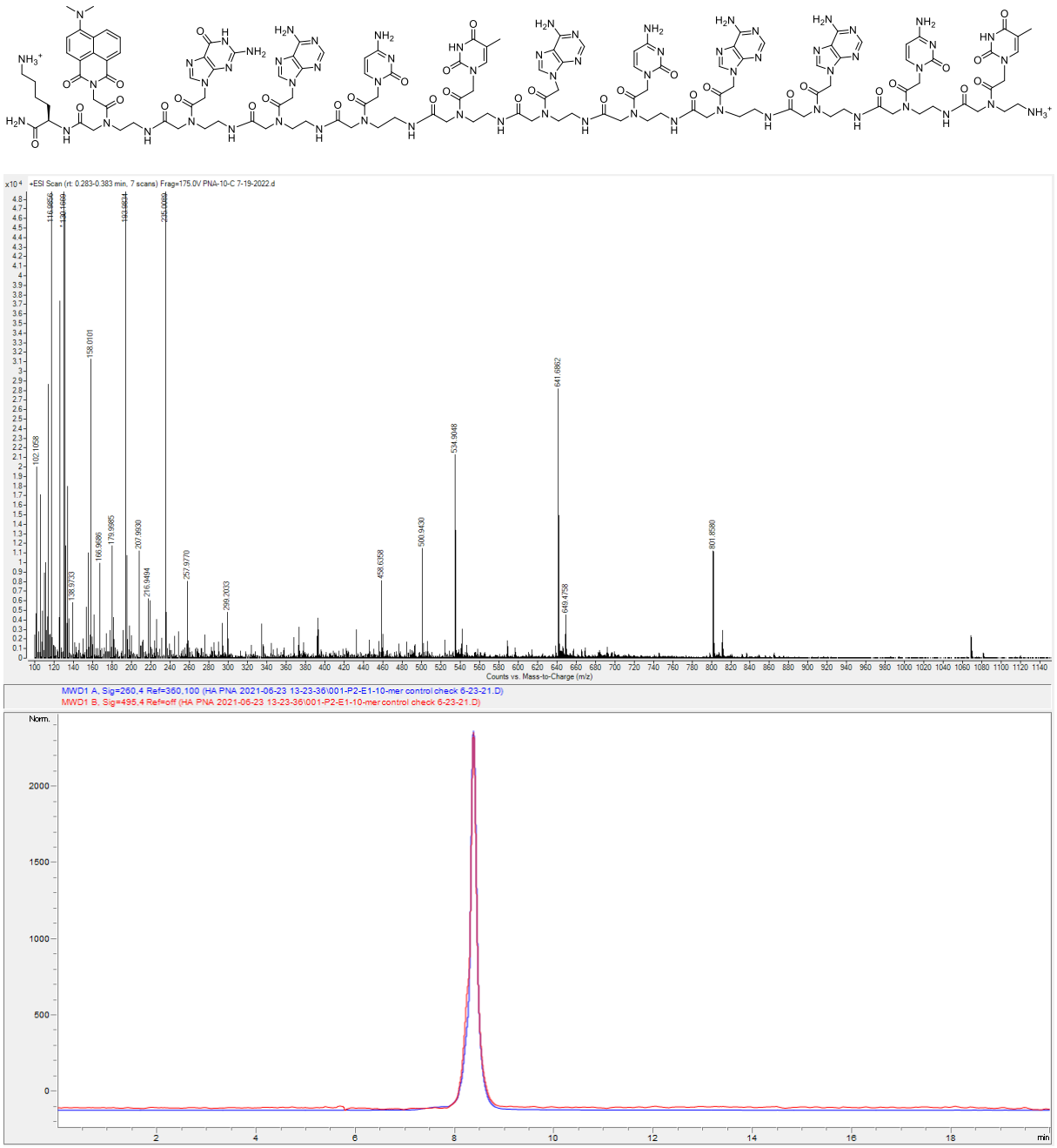


Fig S12 The chemical structure of PNA-10-C. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

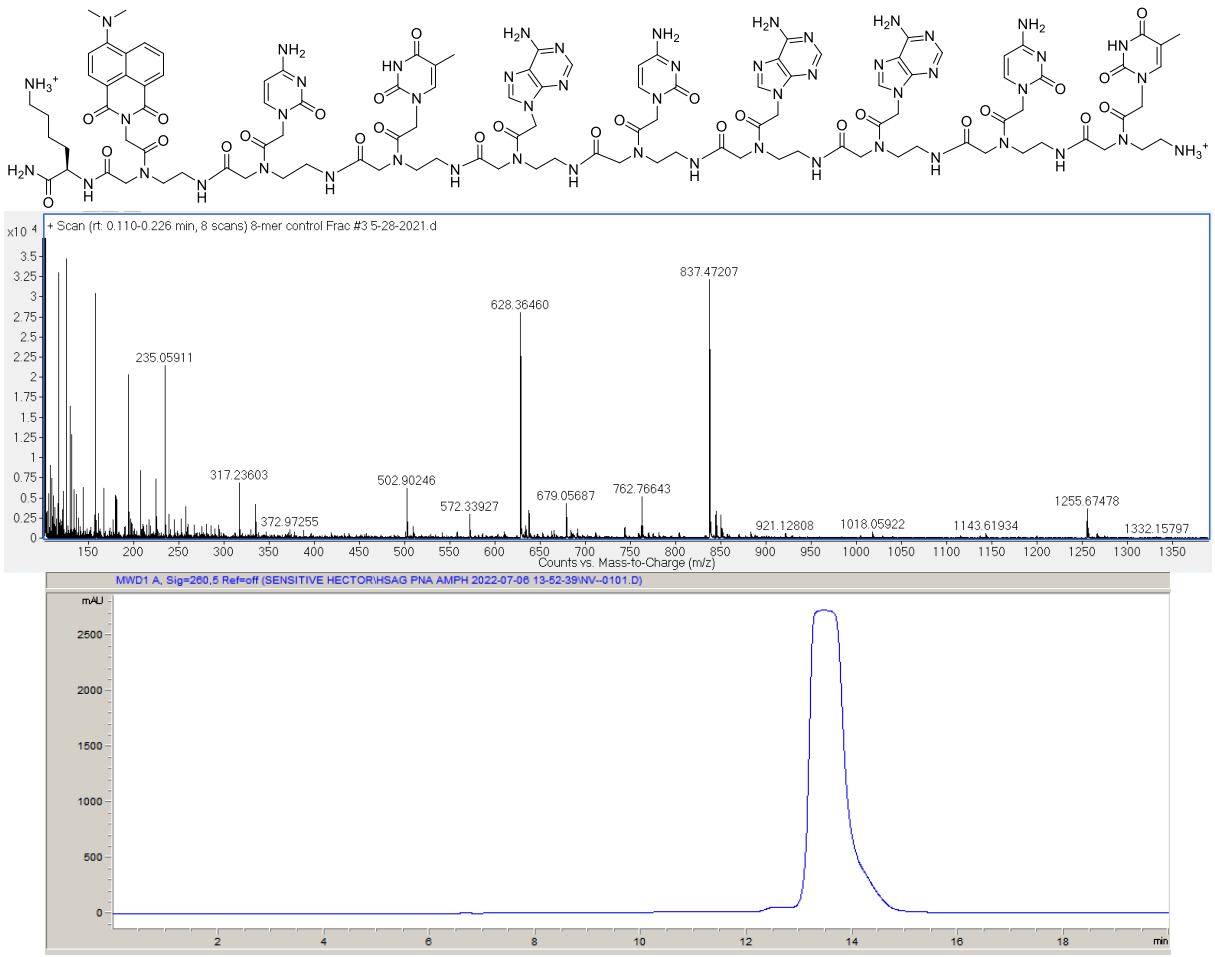


Fig S13 The chemical structure of PNA-8-C. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

Table S3 PNA sequence, expected mass, and found mass

Strand	Expected Mass	Found Mass
PNA-B-A	962.212	962.686
PNA-N-C	955.943	956.638
PNA-N-A	1288.572	1290.523
PNA-10-C	801.123	801.858
PNA-10-A	811.882	812.379
PNA-8-C	627.777	628.365
PNA-8-A	670.321	670.979

Appendix C: Omitted Data from Chapter 4

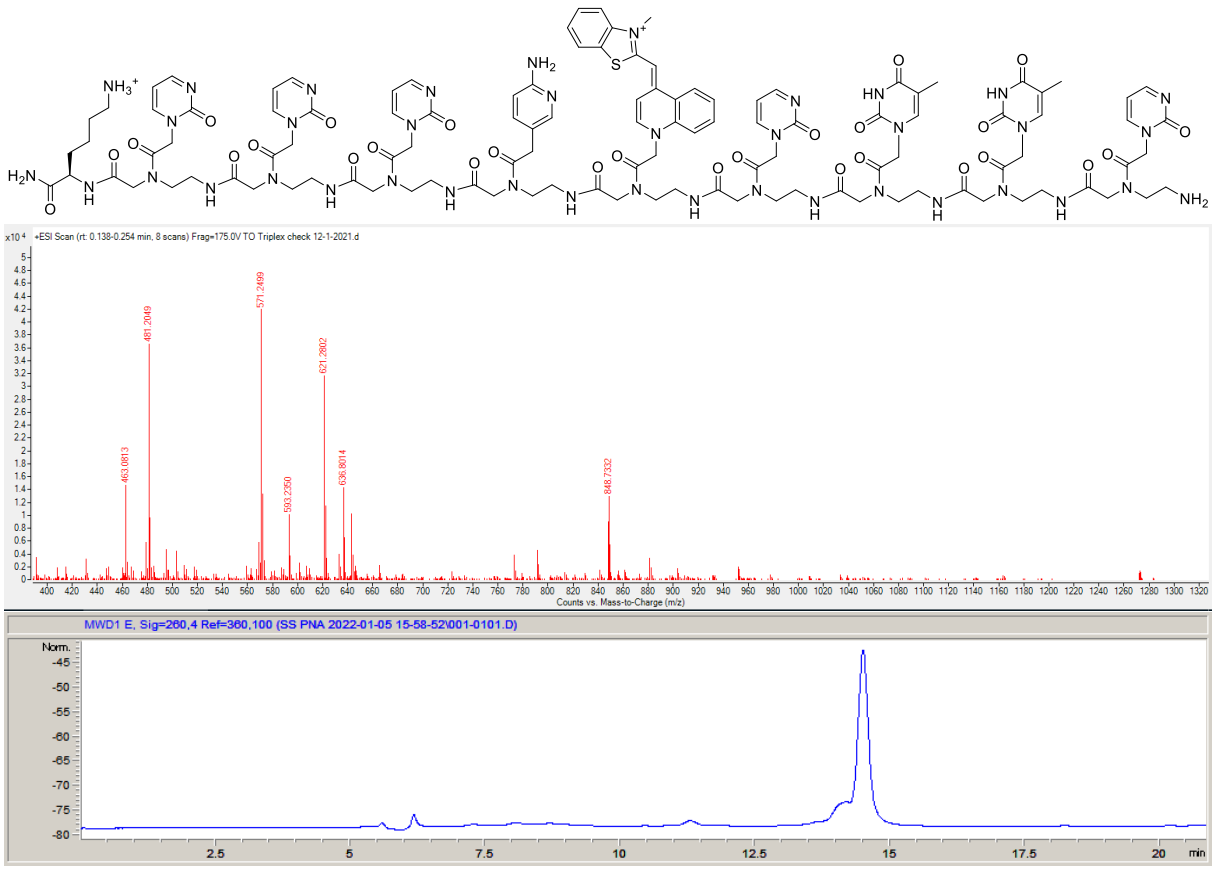


Fig S1 The chemical structure of TO tFIT PNA Probe. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

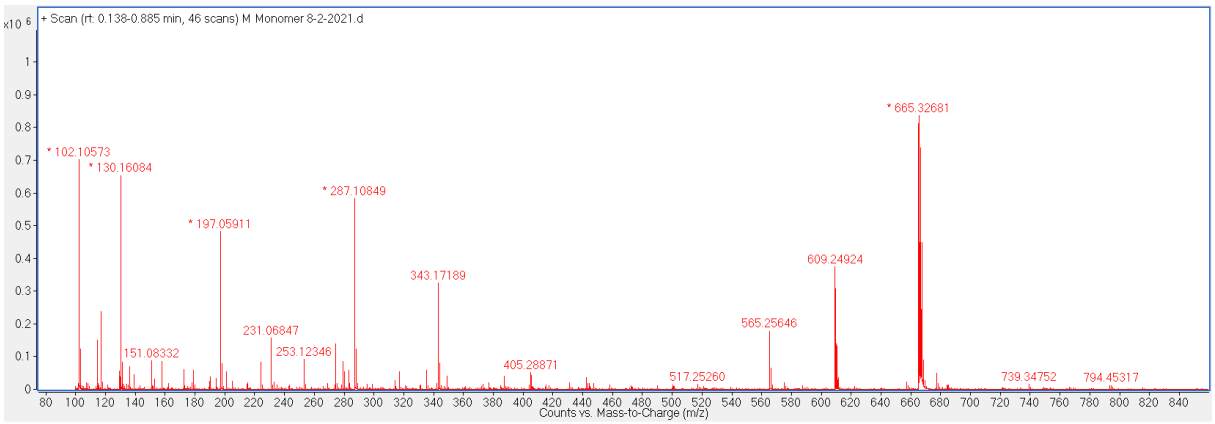


Fig S2 The mass of the Fmoc-M-Bz was confirmed using ESI-TOF mass spectrometry

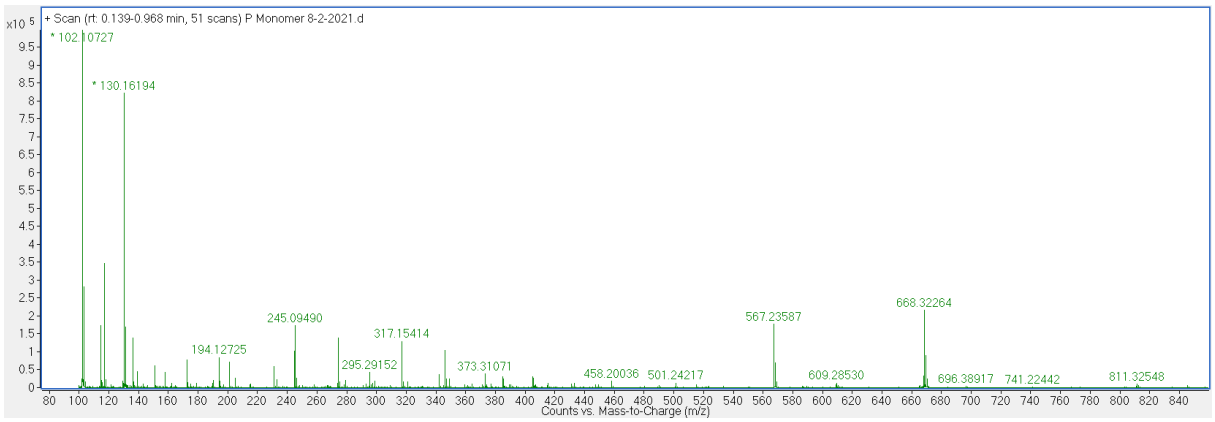


Fig S3 The mass of the Fmoc-P-Bz was confirmed using ESI-TOF mass spectrometry

Appendix D: Omitted Data from Chapter 5

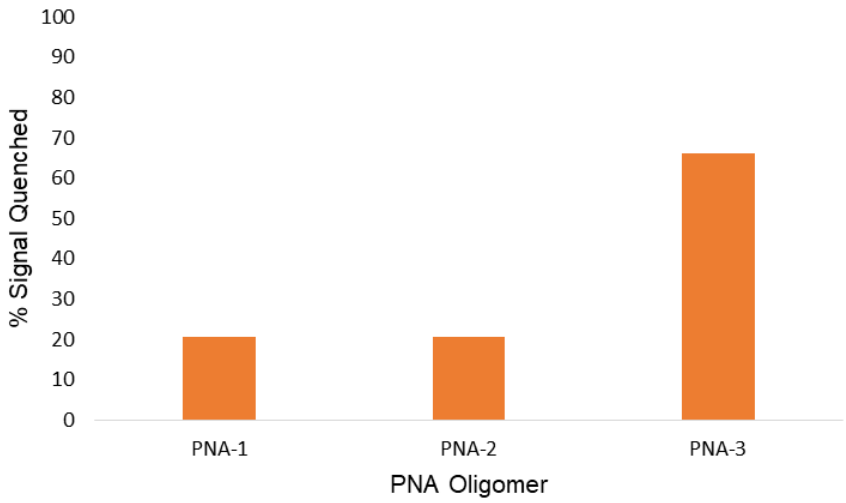


Fig S1 PNA-1, PNA-2, and PNA-3 % signal quenching in the presence of BHQ-1 OTA aptamer in OTA buffer at 3 uM.

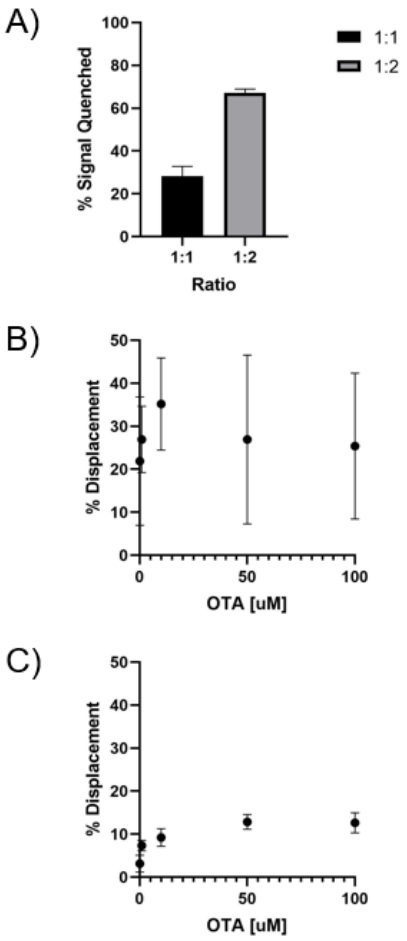


Fig S2 PNA-CM-3 biosensor testing. A) % signal quenching of FAM-labeled OTA aptamer was evaluated in the presence of 1X and 2X PNA-CM-3 in OTA buffer. B) % Displacement of PNA-CM-3:DNA biosensor in the presence of OTA without incubation.

Table S1 FAM-labeled PNA-CM sequences and BHQ-1-labeled OTA aptamer sequence

Strand	Sequence
BHQ-1 DNA	$^{3'}$BHQ1/A C A G G C T A C G A G G G A A A T G C G G T G G G T G T G G G C T A G$^{5'}$
PNA-CM-4	^{N}FAM/ K A T C A T T G T_A C C G A_A K^{C}
PNA-CM-5	^{N}FAM/ K T A T C A T T_A G T C C_A G K^{C}
PNA-CM-6	^{N}FAM/ K A T C A T C G T_A C C G A_A K^{C}

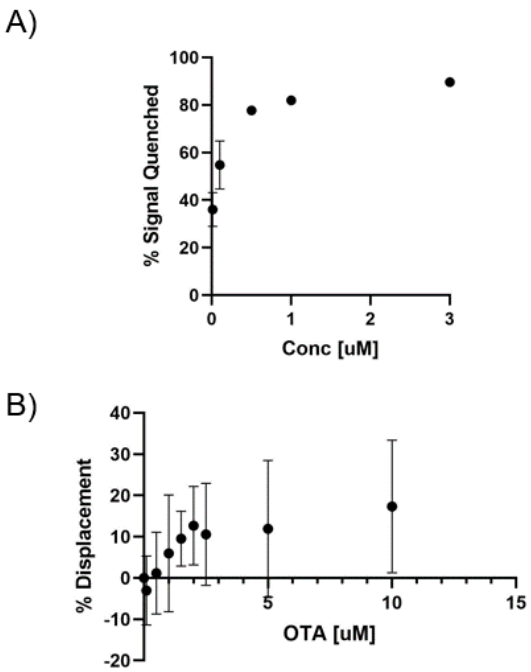


Fig S3 A) % signal quenching using PNA-CM-4, testing at 1:1 Aptamer:PNA-CM-1 ratio. B) % displacement of the PNA-CM-4:DNA biosensor in the presence of OTA in OTA buffer.

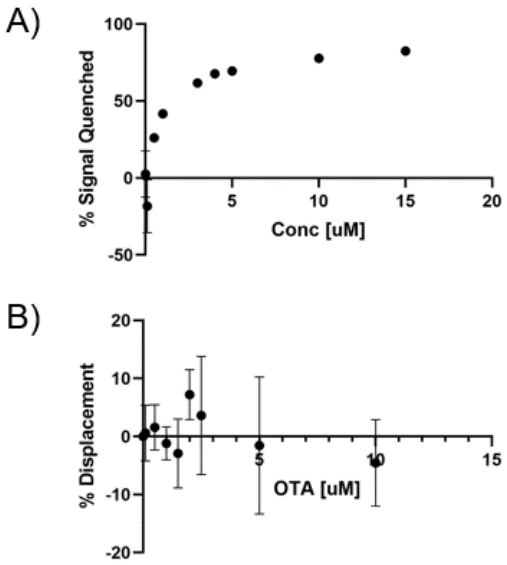


Fig S4 A) % signal quenching using PNA-CM-5, testing at 1:1 Aptamer:PNA-CM-1 ratio. B) % displacement of the PNA-CM-5:DNA biosensor in the presence of OTA in OTA buffer.

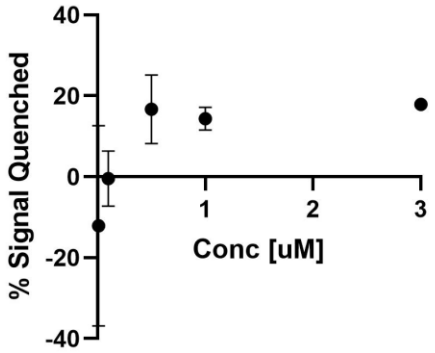


Fig S5 % signal quenching using PNA-CM-6, testing at 1:1 Aptamer:PNA-CM-1 ratio in OTA buffer.

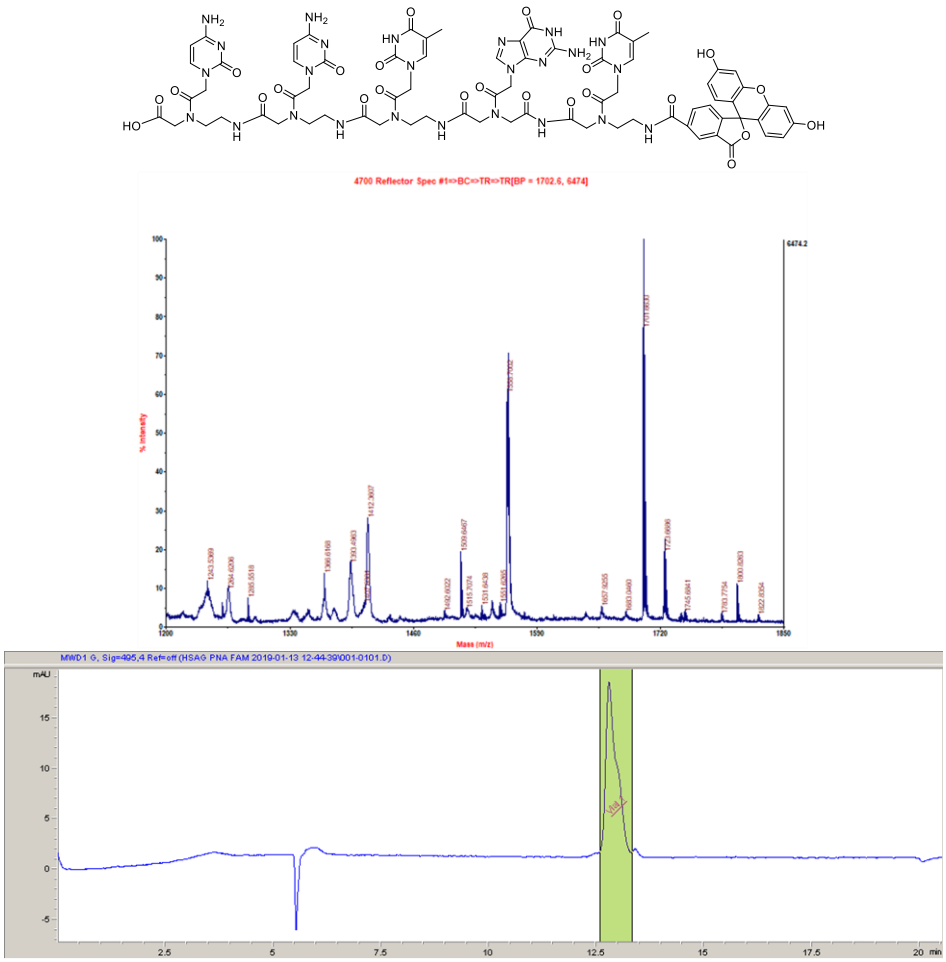


Fig S6 The chemical structure of PNA-1. The mass of the sequence was confirmed using MALDI and purified using RP-HPLC.

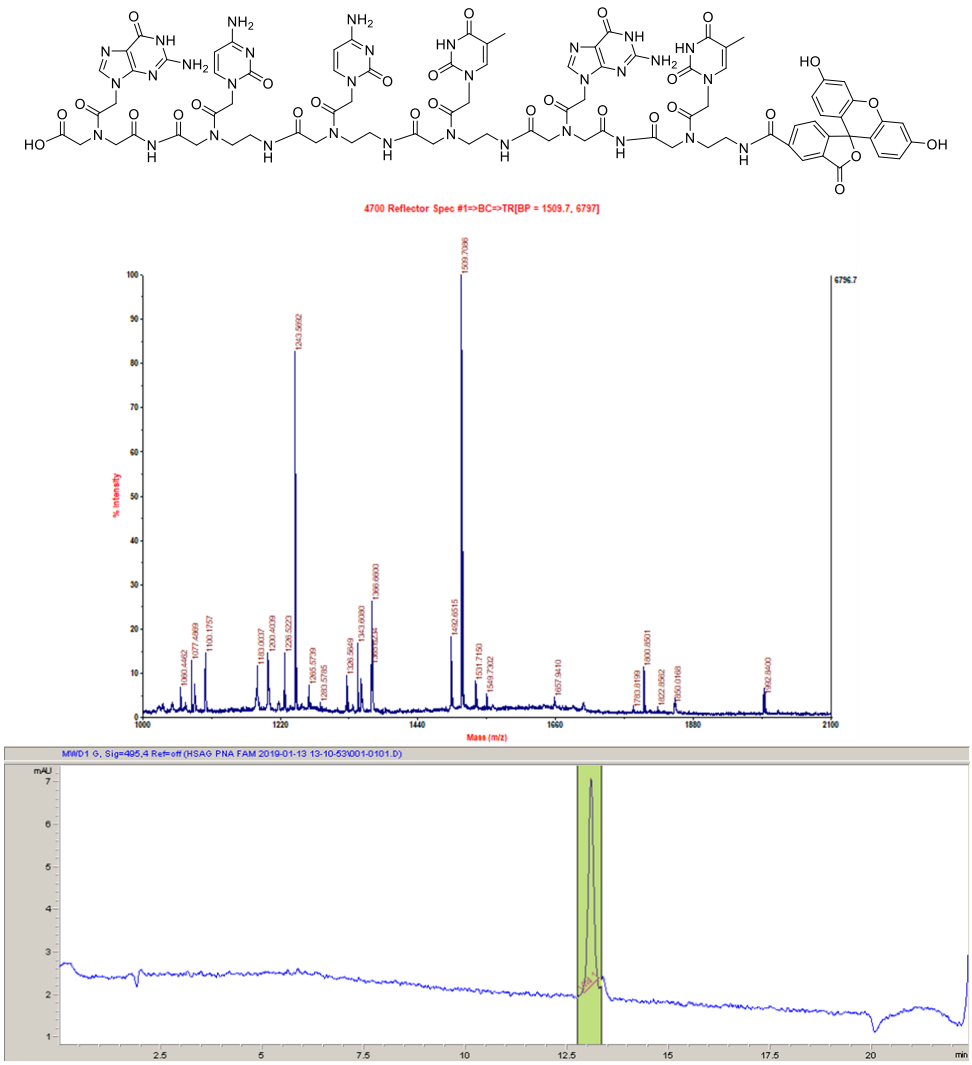


Fig S7 The chemical structure of PNA-2. The mass of the sequence was confirmed using MALDI and purified using RP-HPLC.

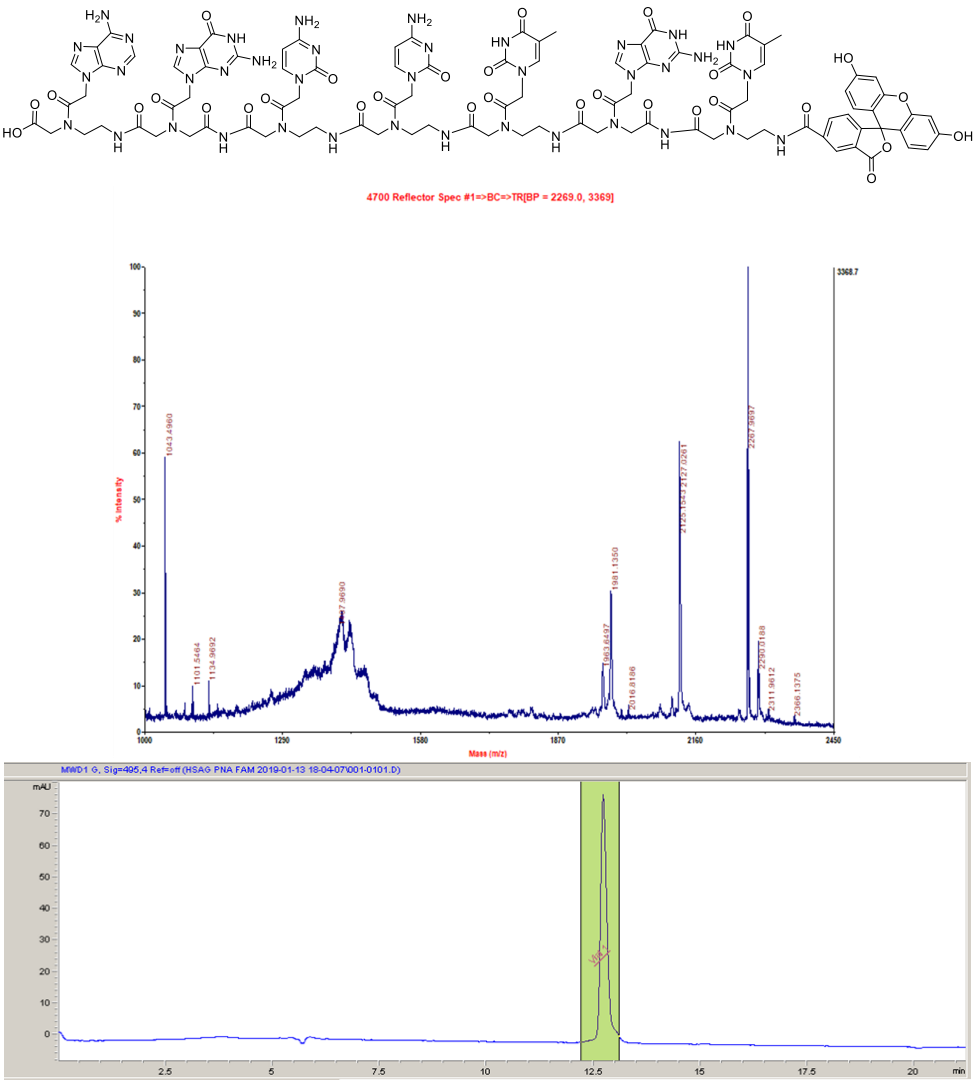


Fig S8 The chemical structure of PNA-3. The mass of the sequence was confirmed using MALDI and purified using RP-HPLC.

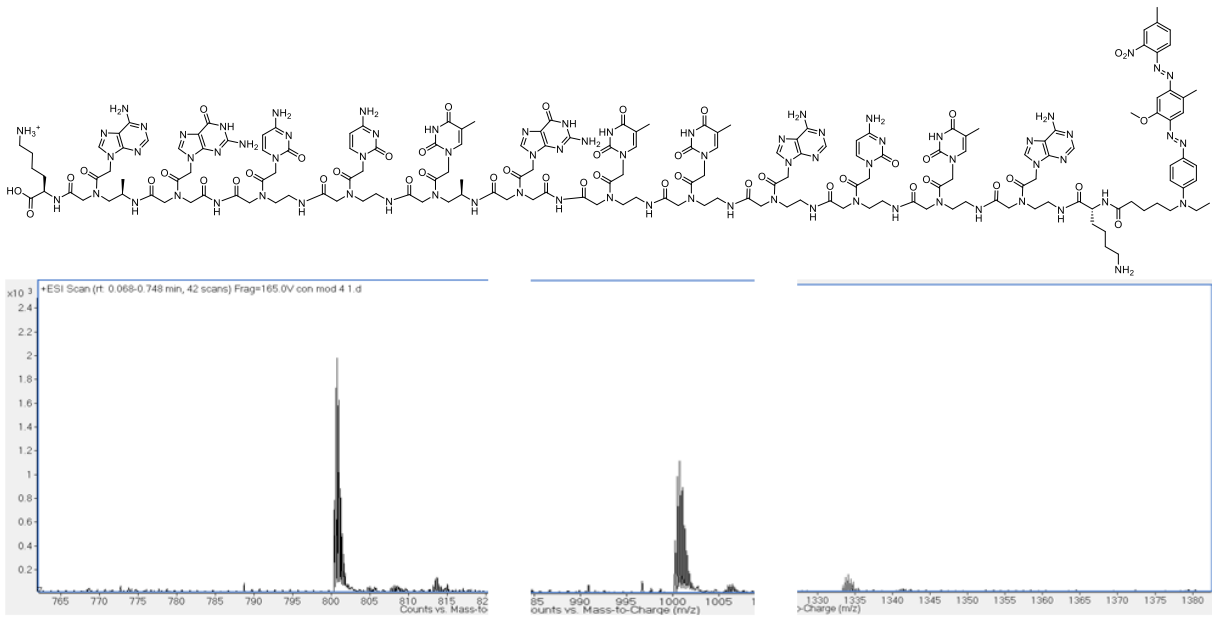


Fig S9 The chemical structure of PNA-CM-1. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

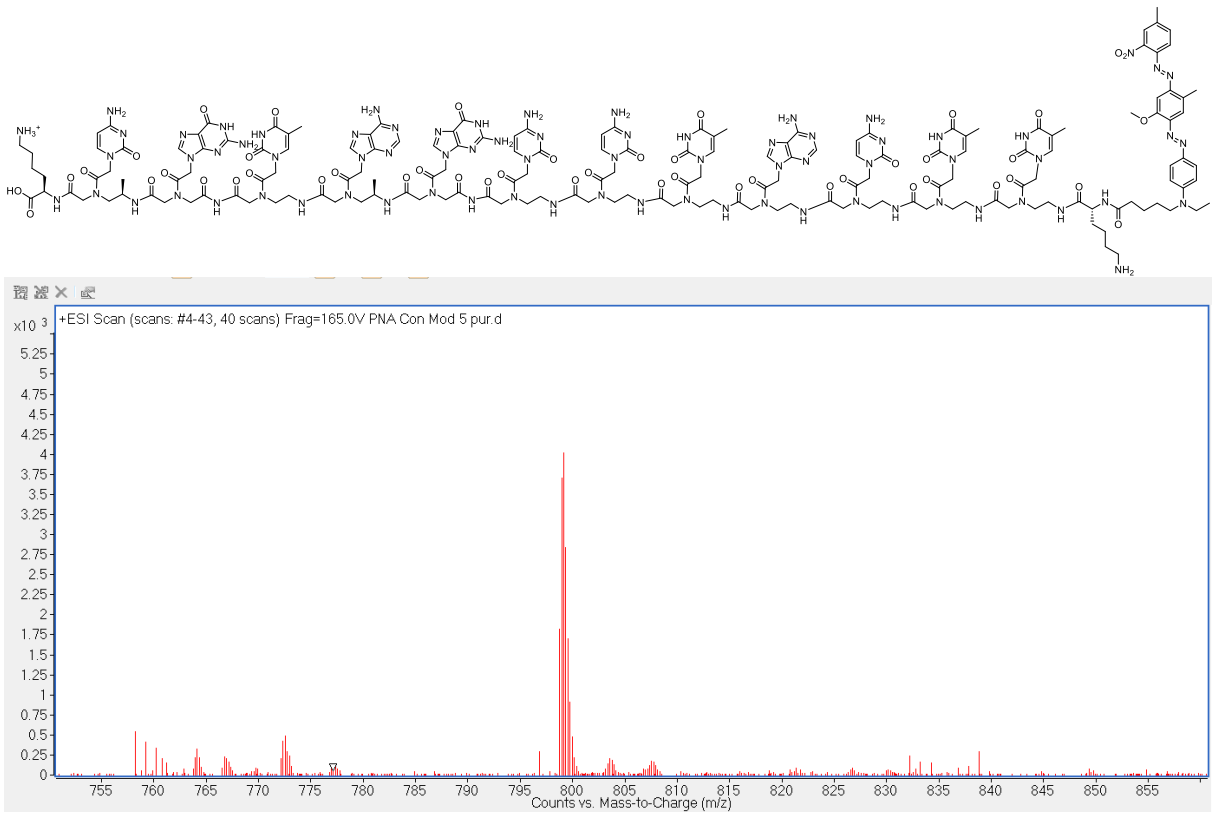


Fig S10 The chemical structure of PNA-CM-2. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

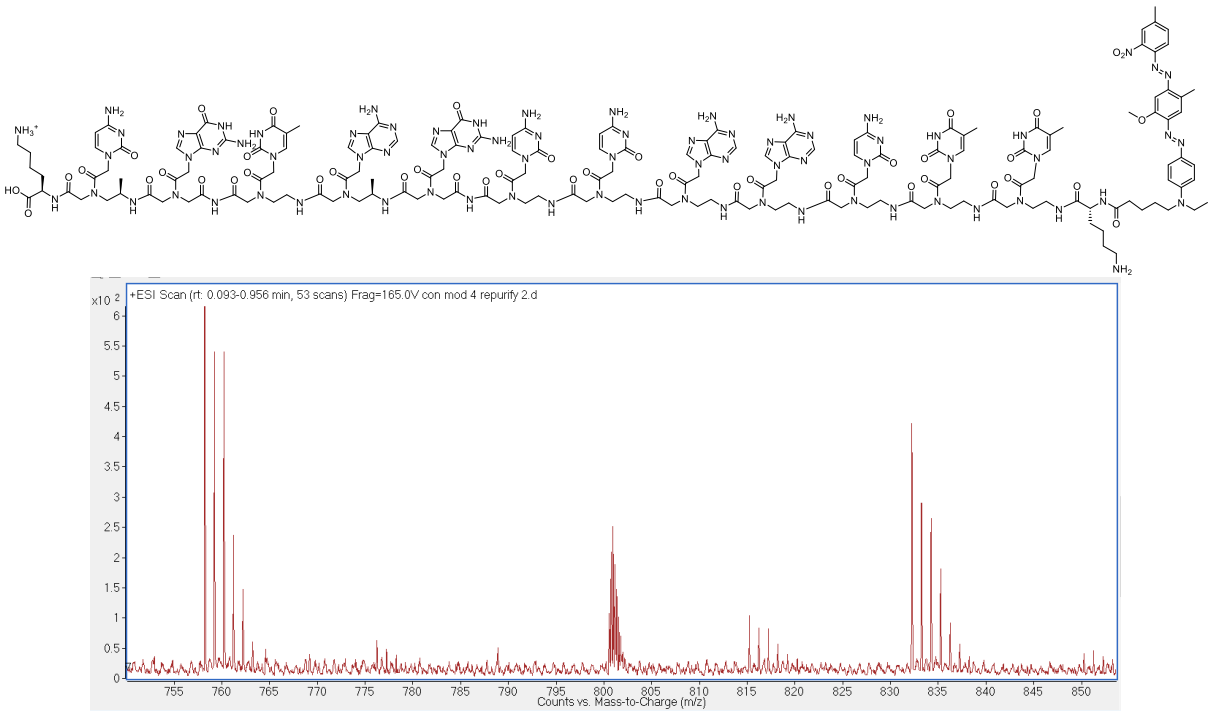


Fig S11 The chemical structure of PNA-CM-3. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

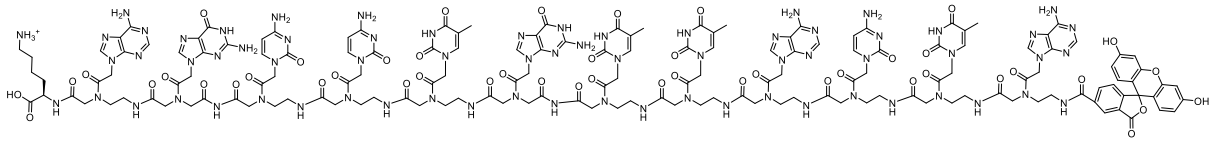


Fig S12 The chemical structure of PNA-CM-4. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

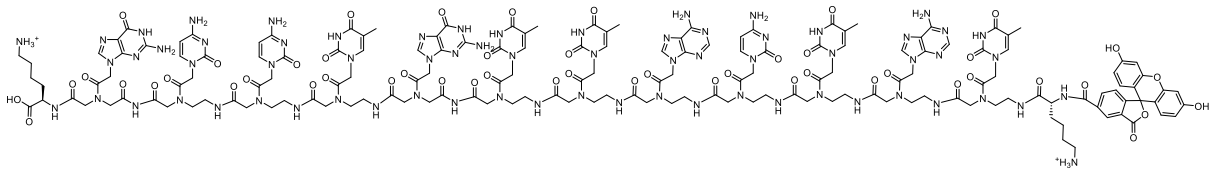


Fig S13 The chemical structure of PNA-CM-5. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

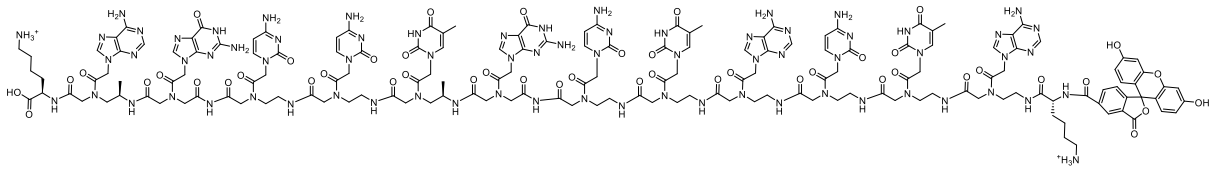


Fig S14 The chemical structure of PNA-CM-6. The mass of the sequence was confirmed using ESI-TOF mass spectrometry and purified using RP-HPLC.

Appendix E: Omitted Data from Chapter 6

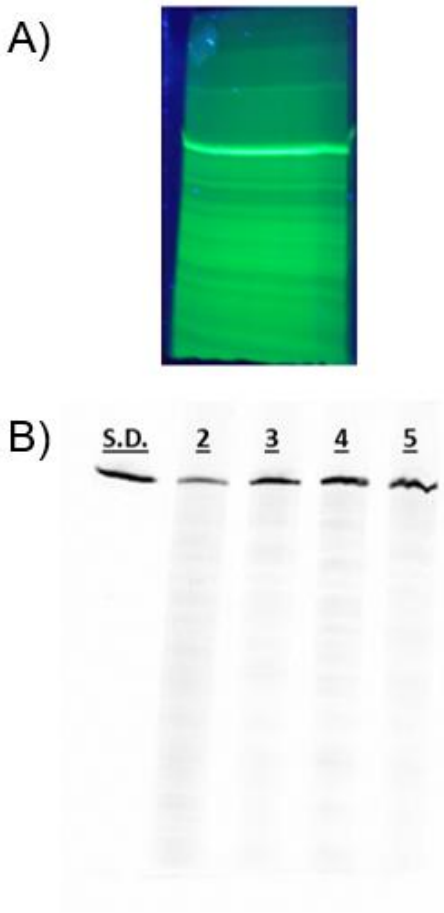


Fig S1 Gel purification of candidate sequences. A) Purification of candidate sequences via 10% PAGE gel, recovered via crush and soak. B) 10% PAGE gel purity confirmation using 90 bp standard (S.D.).

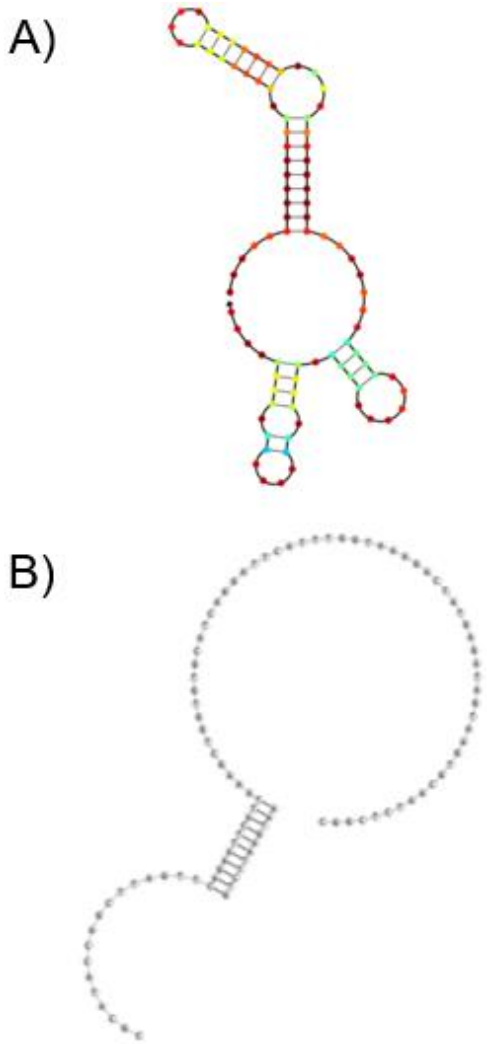


Fig S2 Sequence secondary structures via Nupack. A) Single-stranded. B) Double-stranded using capture strand.