Non-Nucleosidic Modifications in the Selection of Thrombin-Binding Aptamers

By

Shaun Anderson

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Science

Department of Chemistry

McGill University

Montréal, Québec, Canada

December 2021

Abstract

Aptamers are single-stranded oligonucleotides that bind to targets in a manner analogous to antibodies and have been termed synthetic antibodies. Aptamers offer several key advantages over antibodies – their synthesis is cheaper and more reproducible, their selection is much faster, they are more easily functionalized, and they are more shelf stable. However, aptamers suffer most crucially from deficiencies in biostability, as they are susceptible to nuclease degradation and renal filtration, and are limited in the diversity of their building blocks. There have been many modifications to date that have improved the pharmacokinetic properties of aptamers, such as modified backbones, 2'-modified sugars, and bulky end groups. There still exists a true limitation of aptamers in the chemical space probed by SELEX experiments (systematic evolution of ligands by exponential enrichment), in which libraries of oligonucleotides are selected for binding to a target by iterative rounds of selection and amplification. This thesis will address the current limitations of the field and introduce a modified aptamer library that expands the chemical space accessible in aptamer selections. Using non-nucleosidic phosphoramidites alongside solid-phase synthesis and DNA-encoded libraries, a novel approach to selecting modified aptamers was designed and tested on the thrombin-binding aptamer. Aptamer candidates were identified based on sequencing results and were synthesized and characterized individually for comparison with the unmodified sequence.

Résumé

Les aptamères sont des structures d'oligonucléotides qui s'associent aux cibles de façon similaire aux anticorps et alors se nomment des anticorps synthétiques. Les aptamères offrent certains avantages – leur synthèse est moins chère et plus reproductible, leur sélection est plus vite, elles sont plus facilement fonctionnalisées, et elles ont des durées de conservation plus longues. Cependant, les aptamères souffrent de déficiences en stabilité *in vivo*, puisqu'elles ont tendance à être filtrées par les reins et dégradées par les nucléases, et sont limitées par la diversité de leur composante de base. Il existe plusieurs modifications chimiques qui ont amélioré leurs propriétés pharmacocinétiques, incluant les squelettes modifiés, les sucres modifiés à la position 2', et les volumineux groupes terminaux. Malheureusement, les aptamères sont encore limitées à une mince espace chimique dans les expériences SELEX (évolution systématique de ligands par enrichissements exponentiel), dans lesquelles des populations aléatoires d'acides nucléiques sont sélectionnées pour une affinité à une cible par des tours de sélection et amplification. Cette thèse fera face aux limites actuelles du domaine et introduira une banque d'aptamères modifiées qui améliore l'espace chimique exploré par les sélections d'aptamères. En utilisant des phosphoramidites sans bases d'acides nucléiques avec la synthèse en état solide et les banques encodées par l'ADN, une nouvelle approche pour la sélection d'aptamères modifiées a été conçue et testée avec l'aptamère pour la thrombine. Des candidats ont été identifiés selon les résultats de séquençage et ont été synthétisés et caractérisés individuellement afin de les comparer à la séquence originale.

Acknowledgements

I would like to express my very sincere gratitude to Prof. Hanadi Sleiman for her help, advice, guidance, and compassion during the duration of my degree. I would also like to thank all my fellow lab members for their friendship and support. To those who have been involved with this project, both present and past, I am very appreciative and indebted to you all. This research was funded in part by the Natural Sciences and Engineering Research Council of Canada.

Glossary of important abbreviations

AEGIS – artificially expanded genetic information systems

ASO – antisense oligonucleotides

CEP – β-cyanoethyl-N,N’-diisopropylamino-phosphoramidite

CPG – controlled pore glass

DEL – DNA-encoded library

DMT – 4,4’-dimethoxytrityl

DNA – deoxyribonucleic acid

G-quadruplex – guanine quadruplex

NGS – next-generation sequencing

PAGE – polyacrylamide gel electrophoresis

PCR – polymerase chain replication

PO – phosphodiester

PS – phosphorothioate

RNA – ribonucleic acid

SELEX – systematic evolution of ligands by exponential enrichment

SOMAmers – slow off-rate modified aptamers

SPS – solid-phase synthesis

ssDNA – single-stranded DNA

TBA – thrombin-binding aptamer

XNA – xeno nucleic acid

Table of Contents

1. Introduction

1.1. Aptamers vs antibodies

Antibodies are Y-shaped immunoglobulins produced by white blood cells that bind to epitopes and are generated by the body to fight infections and target foreign bodies. Antibodies have been widely used for the detection of analytes since the 1970s, aided largely by advancements in the culture of monoclonal colonies, which have enabled the large-scale production of a single antibody.[1] Since then, antibodies have been regularly used in diagnostic assays wherein a binding event is required between a target (antigen) and an antibody, such as ELISAs (Enzyme-Linked ImmunoSorbent Assay) and western blots. Antibodies also have therapeutic applications, as they are able to block cellular receptors and growth factors, recruit other functional molecules, and deliver cytotoxic cargo.[2] However, antibody technologies do have some notable drawbacks.[3] The production of antibodies is costly and time-consuming, as they must be obtained from live cell cultures that are prone to batch-to-batch variations and contamination. They also cannot be easily modified and are prone to irreversible denaturation, limiting their shelf stability. In addition, since antibodies are produced as part of an immune response, their targets are typically restricted to molecules that elicit immune responses.[4]

Aptamers are single-stranded oligonucleotides often considered chemical antibodies due to the nature of their synthesis. Much like antibodies, they fold into three-dimensional structures and bind strongly to targets, but they are made synthetically using enzymes or solid-phase synthesis. The field of aptamers is much newer and began in the 1990s with the *in vitro* selection of single-stranded RNA molecules that bind to organic dyes[5] and DNA polymerases[6] and single-stranded DNA molecules that bind to thrombin.[7] Aptamers are selected by a process termed SELEX, standing for the systematic evolution of ligands by exponential enrichment (**Figure 1.4**). In this process, a library of random oligonucleotides with constant primer regions is synthesized via solid-phase synthesis and introduced to an immobilized target of interest, typically a protein. Washing steps remove unbound sequences and elutions recover successful binders. Sequences that are recovered after the selection are amplified by PCR (polymerase chain replication) in

order to generate a new library of lower diversity. This cycle of selection and amplification is repeated anywhere from 5-20 cycles,[8] following which high-throughput sequencing is used to identify the most enriched sequences. Individual sequences can then be synthesized and tested for binding affinities and serum stabilities.

Aptamers present several key advantages over antibodies. As they are typically made by solid-phase synthesis, they can be produced much more quickly and inexpensively, and to a higher degree of purity. Since their selection process does not rely on specific physiological conditions like those in which antibodies must be produced, aptamers can be designed to bind to a variety of targets in a diverse range of environments or applications. Contrary to antibodies, they are not limited to targets that elicit an immune response. They are not constrained by Watson-Crick-Franklin base pair interactions; aptamers can form a range of structures such as guanine-quadruplexes, i-motifs, hairpins, and loops.[9] By virtue of their chemical nature, they are also easily functionalized with reporter molecules such as fluorophores. Aptamers can be used in diagnostic applications for ligand-based assays or biosensors, as ligands for targeted drug delivery, as therapeutics themselves with target-binding disrupting function, and they can even be used in purification systems.[10]

Table 1.1. A comparison of key features of aptamers and antibodies is shown below.

	Aptamers	**Antibodies**
Composition	Single-stranded DNA or RNA (4 nucleotides)	Protein chains (20 amino acids)
Structures	Loops, hairpins, guanine-quadruplexes, i-motifs, pseudoknots	Beta-sheets, alpha-helices, turns, loops
Size	~10-30 kDa	~150-170 kDa
Cost	Low	High
Time	Can be synthesized in days and selected in weeks	Require months to produce
Purity	Easily purified and highly reproducible	Properties vary from batch-to-batch and cells are prone to contamination

Shelf stability	Long-term stability at moderate temperatures, can reversibly denature	Short-term stability at low temperatures, prone to irreversible denaturation
Serum stability	Prone to nuclease degradation	Stable *in vivo*
Binding affinity	High affinity, typically in µM or nM regime, but mainly for positively charged targets	Very high binding affinities, usually in nM or pM range
Specificity	Prone to cross reactivity but selection conditions are easily tuned for increased specificity	Prone to cross reactivity
Targets	Small molecules, metal ions, nucleic acids, proteins, cells	Proteins, nucleic acids, but targets must produce an immune response
Modifications	Can be easily modified at specific sites during chemical synthesis	Can be tagged with signaling or binding molecules
Chemical diversity	Low, as nucleic acids are limited to two similar types of aromatic structures that engage in very similar types of interactions with target molecules (e.g., H-bonding)	High, as amino acids can be positively, negatively, or neutrally charged at physiological pH, they feature a variety of hydrophobic and polar chains, and may contain additional atoms like sulfur

However, they are not perfect. Aptamers suffer most critically from nuclease degradation in physiologically relevant environments, limiting their applicability *in vivo*. They are also smaller and thus more easily filtered by kidneys, resulting in shorter *in vivo* circulation times than antibodies.[11] Fundamentally, aptamers also suffer from a narrower chemical space as a result of its constituent monomers. While oligonucleotides are made of four very similar nucleotides, antibodies are made of twenty amino acids that cover a more extensive range of shapes, chemistry, and polarity. These important drawbacks of aptamers have led to a plethora of chemical modifications that have been incorporated into aptamers over the last thirty years. **Table 1.1** compares some of the major features of aptamers and antibodies.

1.2. The thrombin-binding aptamer

Thrombin is a protein involved in blood coagulation. In one of its important functions as a serine protease, it converts soluble fibrinogen to insoluble fibrin as part of the clotting cascade. Two

distinct features of thrombin are its surface loops responsible for binding to sodium ions and its anion-binding exosites positioned on opposing sides of the substrate-binding site.[12] Exosite I, also known as the fibrinogen-binding exosite, is needed for the enzymatic cleavage of fibrinogen by thrombin and is comprised mainly of arginine, lysine, and tyrosine residues. Exosite II, also known as the heparin-binding exosite, also features basic arginine and lysine residues but it interacts with heparin, a cofactor for antithrombin, which inhibits thrombin activity.

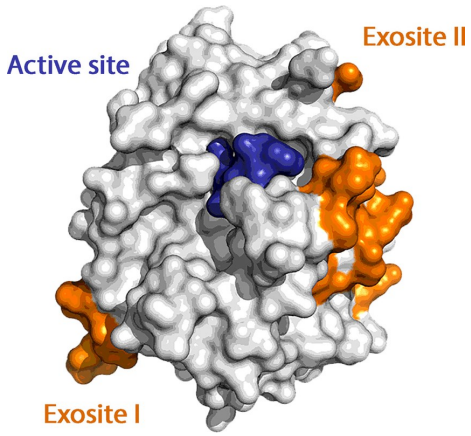


*Figure 1.1. The structure of thrombin is shown above with the relevant regions highlighted.[13] In orange are exosites I (the fibrinogen-binding exosite) and II (the heparin-binding exosite) and in blue is the active site. Reproduced with permission from reference **13.** Copyright 2016 Oxford University Press.*

In 1992, Bock et al. discovered a 15mer ssDNA sequence that folds into a guanine-quadruplex (G-quadruplex) and is able to bind to and inhibit thrombin.[14] Using SELEX, the systematic evolution of ligands by exponential enrichment (see **below**), a library of 60mer sequences were tested for binding and one particular 6mer (5' GGTTGG 3') repeat unit was conserved across most selected aptamers. Two repeat units linked by a TGT loop were found to inhibit thrombin successfully, with the full sequence being 5' GGTTGGTGTGGTTGG 3'. In this 15mer, there are four guanine pairs that form a G-quadruplex based on Hoogsteen hydrogen-bonding, as shown in **Figure 1.2** and **Figure 1.3 (left)**. Telomeres are regions of repetitive non-coding nucleotides at the ends of linear chromosomes that prevent the loss of important genomic DNA during cell division. G-quadruplexes are present in these telomeric repeat units rich in guanines and regulate the extension of telomeres by polymerases.[15] The TBA (thrombin-binding aptamer) consists of a unimolecular, antiparallel G-quadruplex (shown in **Figure 1.3**), with the TT and TGT loops thought to be most involved with binding to thrombin. The aptamer can fold into this chair-like

conformation in the presence of metal cations, such as potassium, or in the presence of its target, thrombin.[16]

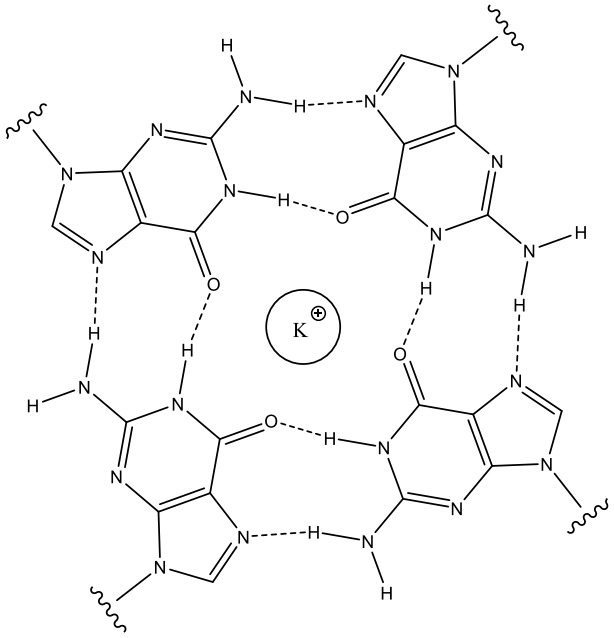


Figure 1.2. The hydrogen-bonding motif involved in guanine-quadruplexes is shown above, with a potassium metal ion coordinated by the electron rich centre between two guanine quartet layers.

As the thrombin-binding aptamer was one of the very first aptamers ever selected, it has since become arguably the most widely studied and best understood aptamer. However, there are still conflicting arguments as to the nature of the interaction between aptamer and protein. XRD (X-ray diffraction) models indicate the TGT loop interacts with the fibrinogen-binding site while NMR (nuclear magnetic resonance) models indicate that T3 and T12, of the two TT loops, are responsible for binding to the fibrinogen site of thrombin. These differences occur as a result of the two models predicting electron deficient regions in different regions of the aptamer, with more electron deficient regions less likely to be involved in the primary interactions with the electropositive fibrinogen-binding site.[17] Both models predict that the loops not involved with exosite I will interact weakly with exosite II of another thrombin molecule. The 29-base aptamer HD22 was discovered not long after and binds to the heparin-binding exosite with a mixed duplex/G-quadruplex structure.[18] There have been several studies that have modified the sequence, the size, or the bases themselves, which have primarily supported the NMR model in which the TT loops interact with the fibrinogen-binding site (see **Figure 1.3**). Generally, larger

groups are less well tolerated in the binding loops with modifications in these regions geared toward improving binding affinity while modifications in the TGT loop tend to focus on improving the stability of the G-quadruplex.[19-24] Specific modifications will be discussed later in the section on modified aptamers.

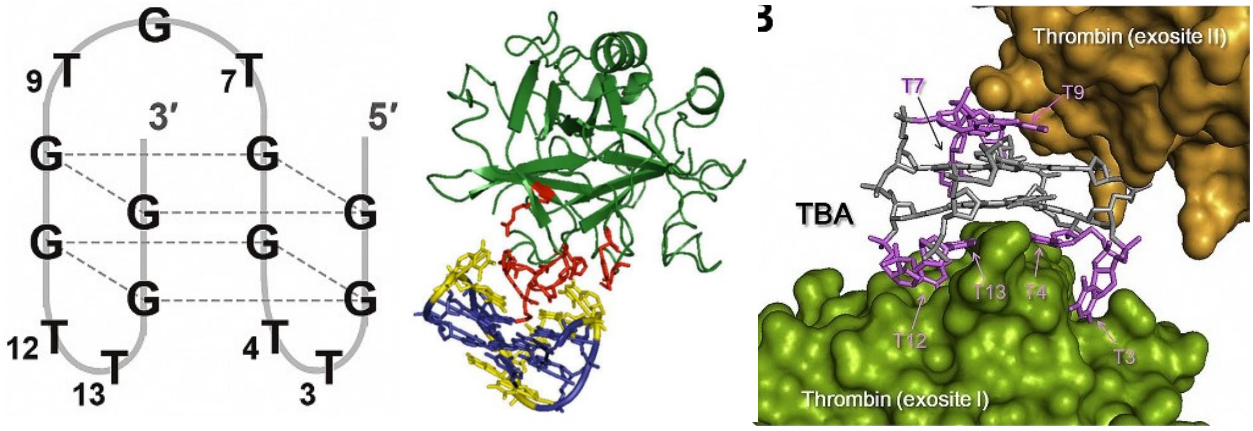


*Figure 1.3. Left – the antiparallel guanine-quadruplex is shown as part of the 15mer TBA (thrombin-binding aptamer).[25] Middle – the TT loops of the TBA (TT and TGT loops shown in yellow, GG pairs in blue) are shown interacting with exosite I of thrombin (thrombin is shown in green with interacting amino acids in red).[17] Right – the TT loops are shown interacting with thrombin's exosite I and the TGT loop with exosite II (TT and TGT loops shown in purple).[25] Adapted with permission from references **25** and **17**. Copyright 2012 Bentham Science Publishers and 2011 Elsevier, respectively.*

1.3. Aptamer selection

1.3.1. The SELEX process

SELEX, the systematic evolution of ligands by exponential enrichment, was developed in the 1990s by two independent research groups to evolve DNA or RNA oligomers capable of folding and binding to a target of interest. By applying selective pressures to a randomized library of potential aptamers, SELEX mimics natural selection by requiring a trait (in this case binding) for oligonucleotides to survive selection, but over a much more practical timeframe as individual cycles of SELEX take just days. Oligonucleotides add another element of convenience compared to other oligomers as the genotype and phenotype are one and the same, as the nucleic acids that form specific three-dimensional shapes can be sequenced and reproduced reliably. SELEX begins with an oligonucleotide library of fixed length and randomized sequence flanked by constant primer regions required for amplification. The entire library is incubated with the target, classically proteins, and potential aptamers compete for binding to the target. Unbound

aptamers can be removed with washing steps and retained aptamers eluted. An enriched library is thus obtained and can be amplified for future cycles of competitive binding. This allows successfully binding sequences initially present at similar levels to non-binding aptamers to become more represented in enriched libraries. Finally, the library can be sequenced to discover which sequences were the most successful at surviving the directed evolution. **Figure 1.4** below provides a visual depiction of the above cycle.

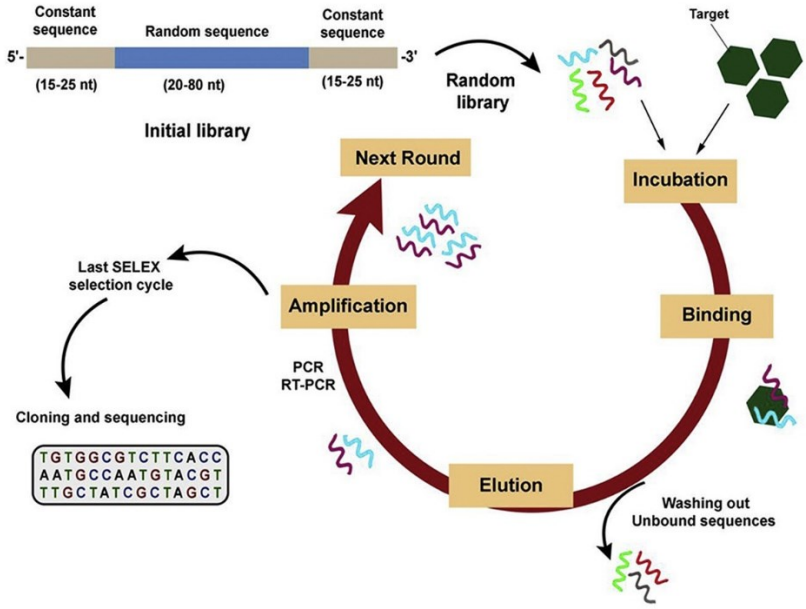


*Figure 1.4. An overview of the SELEX process is presented above.[26] An initial randomized oligonucleotide library flanked by fixed primer regions is incubated with a target and allowed to bind. Unbound sequences are washed away while bound sequences are eluted and amplified for future rounds of selection except in the final cycle, in which they are cloned and sequenced for the identification of enriched aptamers. Reproduced with permission from reference **26**. Copyright 2018 Elsevier.*

1.3.2. Improvements to SELEX

1.3.2.1. Library generation

The first step of SELEX is to generate a library of oligonucleotides for the selection of aptamers. These libraries are typically generated by solid-phase synthesis using phosphoramidite chemistry. With four possible bases at each position of the random sequence, the number of unique sequences created can easily be on the order of 10^{18} for sequences as short as 30 bases, though practical limitations of SELEX restrict this to around 10^{13}-10^{15} for a single selection experiment, which is still orders of magnitude higher than most conventional drug libraries.[27] The sequences of the constant primer regions are important since they should be optimized for PCR

amplification and ideally be designed to avoid hybridization between primers that could interfere with the binding of the aptamer itself during SELEX and could result in primer-dimer by-products during PCR amplification.[28]

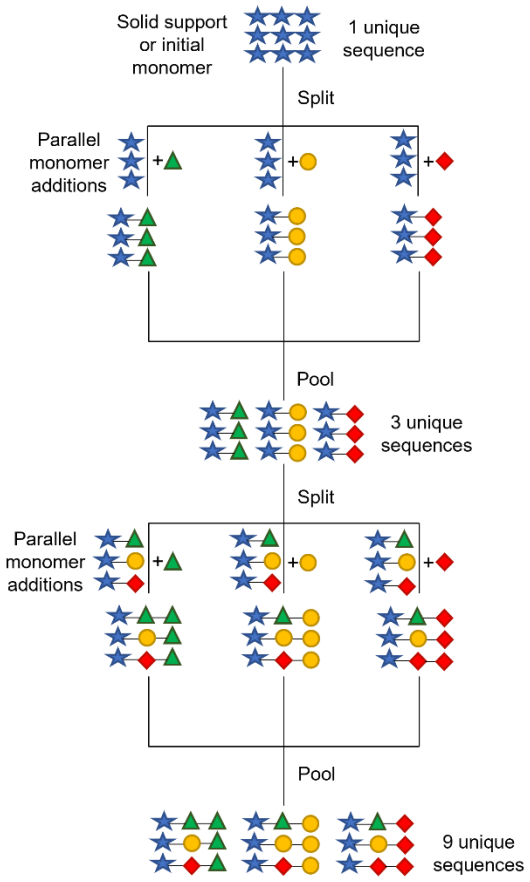


Figure 1.5. A split-and-pool combinatorial library synthesis is schematically shown above. The library is divided into equal amounts for parallel additions of x different monomers and pooled over y steps to give x^y unique sequences. In this case, the blue star can be either a functionalized solid support or an initial monomer that is consistent across all future sequences. The other coloured shapes represent different monomers. The workflow moves from top to bottom.

There are two common approaches to the creation of these libraries – solid-phase synthesis and enzymatic synthesis. Split-and-pool combinatorial chemistry (**Figure 1.5**) relies on splitting a library of growing polymer chains into separate aliquots with each then undergoing the addition of a unique monomer, after which the samples are pooled together and mixed prior to the next split-and-pool cycle. In this way, highly diverse libraries can be generated in relatively few synthetic steps. Aptamer libraries can be generated this way using solid-phase chemistry by splitting the library into four portions, one per nucleotide, and pooling and mixing after each individual coupling. Even simpler, an equimolar mixture of all four phosphoramidite bases can be

used at each coupling step without any need for splitting and pooling.[14] Aptamer libraries have also been generated using enzymatic methods without solid-phase synthesis. Error-prone PCR may be used to introduce random mutations in a parent sequence to create a library of variants that can undergo selection. This method does require an initial sequence from which to work and so is better suited to improving known aptamers rather than discovering new aptamers. Non-homologous random recombination can also be used to generate aptamer libraries, in which entire sections of sequences are shuffled between different oligonucleotides. RNA aptamers for streptavidin were selected from libraries that had underwent error-prone PCR or non-homologous random recombination, starting from parent aptamers derived from standard SELEX experiments.[29] Error-prone PCR resulted in somewhat improved binding affinities compared to parent sequences and non-homologous random recombination resulted in greatly improved binding affinities. Doped-SELEX represents a more recent way to generate a library for SELEX by solid-phase synthesis in which a parental sequence is favoured as in error-prone PCR. Instead of using equimolar mixtures of the four bases during solid-phase synthesis, a biased mixture is used with one of the bases present at a much higher ratio.[30] A doped-SELEX experiment for RNA aptamers for T7 RNA polymerase with only 15% mutagenic variants did not result in any aptamers with improved binding affinity but did indicate which regions were highly conserved and aided in the design of more successful subsequent doped-SELEX experiments.[31] Libraries have also been generated from other sources such as in genomic SELEX, which uses the genomic sequences of an organism to produce sequences for SELEX.[32] In this case, the genomes are subject to digestion, ligation, fragmentation, and PCR amplification in order to get a library of oligonucleotides flanked by primer regions.[33] Aptamer libraries have also been created using particle-display, wherein a library of particles display multiple copies of the same aptamer (**Figure 1.6**). Using emulsion PCR, a solution-phase library of aptamers can be transformed into aptamer particles and subsequently sorted according to binding affinity using fluorescence-activated cell sorting (FACS), allowing for high-throughput screening rather than selection, though the starting aptamer library typically requires a preliminary SELEX experiment to reduce its diversity.[34]

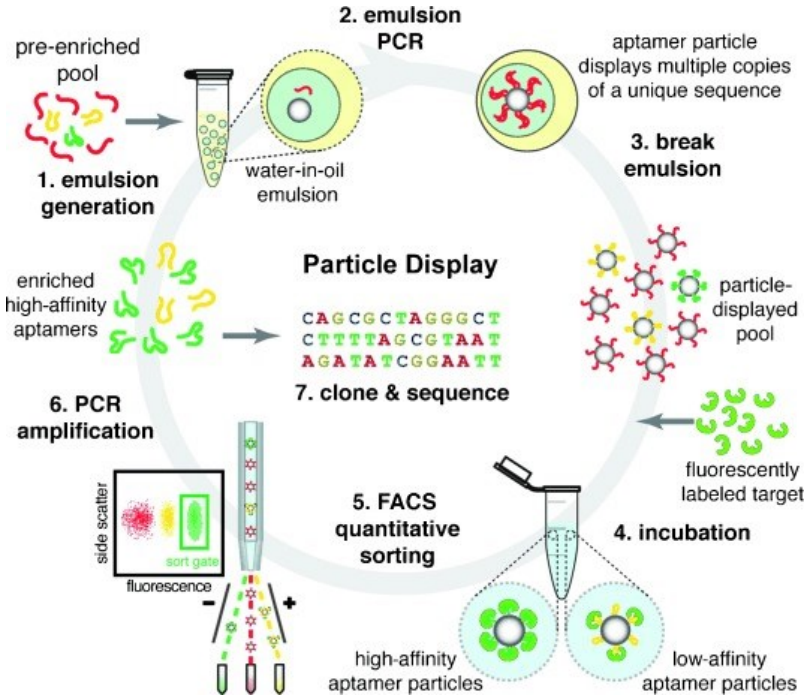


*Figure 1.6. An overview of the particle display screening method is depicted above.[34] A pre-enriched pool, typically from a conventional SELEX experiment, is used to generate a particle-displayed library using emulsion PCR. Fluorescently labeled targets enable quantitative sorting by FACS (fluorescence-activated cell-sorting) based on observed binding affinity. This enriched library is amplified by PCR and is either used to begin another cycle or sequenced for identification of the aptamers. Reproduced with permission from reference **34**. Copyright 2014 John Wiley and Sons.*

An interesting development of SELEX regarding library generation is to shorten or eliminate the primer regions altogether. Most libraries have primer regions that are around 20 bases long each, which can interfere with aptamer binding and reduce the synthetic yields of aptamer libraries. Tailored-SELEX shortens the fixed regions to a total of only ten bases and was first demonstrated in the selection of neuropeptide-binding spiegelmers, which are mirror-image aptamers made of L-DNA.[35] After selection, the library is modified using double-stranded adapters (termed ligates), which enables primer binding sites to be ligated to the short fixed regions that allow subsequent amplification by PCR. The products of PCR, termed amplicons, contain sites that can be cleaved under alkaline conditions, allowing for regeneration of the truncated aptamers prior to future cycles of SELEX.[35] Another variant of SELEX annealed the complementary oligonucleotides to the primer regions along with the single-stranded library prior to selection to avoid interference from primer secondary structures.[36] Primer-free SELEX has also been conducted, with an oligonucleotide library lacking any primer regions undergoing selection. In this case, primers with restriction enzyme sites were ligated to both ends of the library prior to amplification and

enzymatically removed afterward.[37] A more recent variant of library generation is to incorporate conformation-changeable primers as a way to control interactions between targets and primer sequences. Using structures like i-motifs that are sensitive to the ionic strength and pH of the environment, these primers can be reversibly switched between conformations and thus, using a two-step selection process, aptamers can be selected that bind to their target regardless of the structure of the primer sequences.[38] Libraries have also been enhanced with modifications such as those presented **below**.

1.3.2.2. Target and library immobilization

Prior to selection, the target is usually immobilized to allow for unbound and bound sequences to be easily separated from each other during the selection process, though in some cases the library itself is immobilized. The method of immobilization is important as it can induce conformational changes in the target or restrict access to interfaces that could be involved in binding, both of which affect the strength and types of interactions that may occur between aptamer and target.

Targets are typically proteins in aptamer selections and feature several functional residues that can be used to anchor to solid supports. Small molecule targets like L-tyrosinamide have been covalently attached to Sepharose matrices by linking the primary amine group to resin epoxide groups.[39] Protein chimeras or chemically modified proteins can also be used in order to anchor them to surfaces for future separation by affinity chromatography. Protein targets have been immobilized using matrices such as agarose resins or nitrocellulose fibers that adsorb the target and any target-aptamer complexes.[40] More recently, proteins have been non-covalently linked to columns using bound antibodies in a process termed immuno affinity SELEX.[41] These methods require no modification of the target, but with smaller targets this immobilization is no longer effective and other methods must be used. Sol-Gel SELEX has allowed for the non-covalent immobilization of smaller targets like xanthine, with much improved results over SELEX with small targets covalently bound to beads.[42] Sol-gels feature micron-sized channels for aptamer mobility and nanopores for retention of small targets. Many improvements to immobilization methods in SELEX have relied on bead-based systems, which are applicable to a wide range of targets but do limit the interaction surfaces available.[43] With the targets covalently linked to magnetic beads,

the separation of successfully bound aptamers can be much simplified. A process termed FluMag-SELEX immobilized streptavidin on magnetic beads, selected streptavidin-binding DNA aptamers via magnetic separation, and quantified them using fluorescence labeling;[44] neurotoxins were bound to magnetic beads via carbodiimide chemistry for M-SELEX (microfluidic SELEX), allowing DNA aptamers to be separated with a continuous-flow magnetic activated chip-based separation (CMACS) device;[45] and thrombin was covalently linked to magnetic beads via an amide bond for the selection of aptamers in MACE-SELEX (microbead-assisted capillary electrophoresis SELEX), allowing for magnetic and electrophoretic separation.[46]

The aptamers themselves have also been immobilized prior to their selection, though less frequently as this may complicate the other parts of SELEX such as amplification and sequencing. Typically, the 3' end of aptamers is prioritized as the site of attachment as it tends to be more easily digested by exonucleases. Spacers are often introduced to increase rotational freedom, usually with thymidine linkers that are less prone to non-specific interactions than other nucleotides.[47] Thiols, amines, and biotin are commonly added for covalent or non-covalent tethering to membranes and supports. An interesting development is capture-SELEX, which uses DNA hybridization to immobilize a library of aptamers to magnetic beads decorated with partially complementary ssDNA.[48] In this case, the library is composed of two random regions of DNA separated by a 12mer docking sequence, complementary to the capture ssDNA on the beads, and flanked by primers (**Figure 1.7**). Upon binding to the target, the aptamers form three-dimensional structures and dissociate from the capture-DNA, allowing for easy separation using magnets. Self-reporting aptamers (SRAs) have also been immobilized for SELEX experiments (**Figure 1.7**). In this example, the library consisted of two random regions separated by an 18mer DNAzyme sequence and flanked by primers.[49] In the presence of a strand partially complementary to the DNAzyme and in the absence of target binding, the DNAzyme is inactive and produces no fluorescent signal. The library was immobilized onto streptavidin-decorated magnetic beads using the partially complementary and biotinylated DNA, which dissociates from the aptamer and DNAzyme upon successful binding to the target thrombin. In another SELEX experiment, biotin-labeled aptamers for ractopamine were linked to streptavidin-covered magnetic beads and dissociated from the magnetic beads upon binding to the target, allowing

for magnetic separation of unbound aptamers still attached to the beads from aptamers bound to ractopamine.[50] Single-stranded DNA adsorbs non-specifically onto the surface of gold nanoparticles (AuNPs), allowing for the weak immobilization of aptamers in GOLD SELEX, which is sufficient to retain aptamers that do not strongly bind to their target, enabling separation by centrifugation.[51]

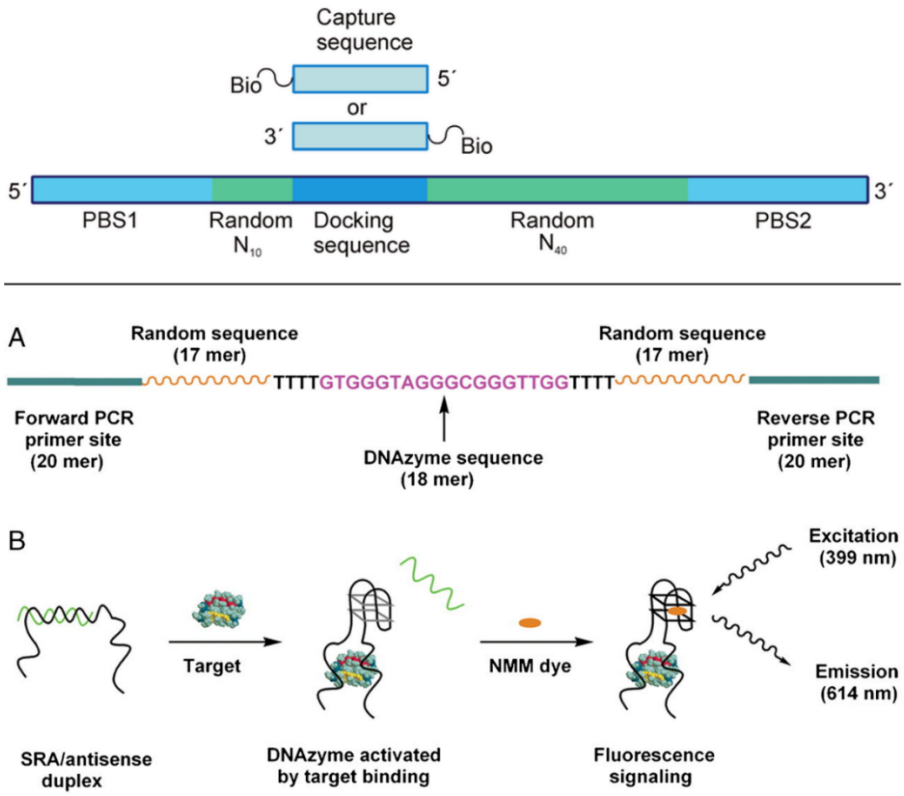


*Figure 1.7. Two ways that structure-switching aptamers have been used in selection experiments are presented above. In the top half, the general design of a capture-SELEX library is depicted.[52] The library is designed with two flanking primer-binding regions, PBS1 and PBS2, with a constant docking sequence placed in the middle of the randomized region. A biotin-tagged complement to the docking sequence, termed the capture sequence, is used to anneal the aptamers to magnetic beads decorated with streptavidin. In the presence of the target, sequences that bind to the target change structure and are displaced from the magnetic beads, allowing facile magnetic separation of unbound and bound aptamers. In the lower half of the figure, a similar approach is shown for self-reporting aptamers (SRAs).[49] In (A), a similarly designed library is shown, with a constant DNAzyme sequence contained within the library's randomized region. The library is initially present with the antisense DNAzyme sequence, resulting in an inactive DNAzyme. For the selection of SRAs, the antisense DNAzyme strand is tagged with biotin to enable conjugation to streptavidin-decorated magnetic beads, as was the case for capture-SELEX. This allows easy separation of sequences that have changed structure and bound their target from those that did not and are still attached to the beads. In (B), the function of SRAs is shown. In the presence of the target, the inactive duplex formed by the SRA and the complementary sequence of the DNAzyme changes shape and the DNAzyme sequence dissociates from its complement. The DNAzyme in its active state produces a fluorescent signal in the presence of its substrate. Adapted with permission from references **49** and **52**. Copyright 2010 National Academy of Sciences and 2018 Multidisciplinary Digital Publishing Institute, respectively.*

Finally, there have been developments in the SELEX field that allow for immobilization-free assays, with aptamers and targets free in solution during the selection step. CE-SELEX (capillary

electrophoresis) allows free targets and aptamers to bind in solution, with separation occurring based on different electrophoretic mobilities upon binding.[53] This immobilization-free method can reduce the number of cycles of SELEX required but is limited to larger targets as it relies on differences in electrophoretic mobility for separation. Other immobilization-free methods rely on the immobilization of the bound or unbound aptamers following selection. Single-stranded DNA adsorbs easily onto graphene oxide by virtue of π-π stacking and electrostatic interactions but dissociates upon hybridization or, in certain cases, as a result of a conformational change induced by target binding.[54] In GO-SELEX (graphene oxide SELEX), the aptamers and target are incubated in solution without any form of immobilization, but are introduced together to graphene oxide, which allows for the separation of unbound DNA from bound DNA via centrifugation.[55] MCP-SELEX (magnetic cross-linking precipitation) also begins with aptamers and targets free in solution but immobilizes the successfully bound sequences for ease of separation.[56] In this case, protein targets are covalently linked to activated carboxylic acid-functionalized beads following incubation with the aptamer library, allowing for magnetic separation.

1.3.2.3. Aptamer selection

The most important part of SELEX is unsurprisingly the aptamer selection. In order to find sequences that bind well to a target, it is important to have a selection step that allows for aptamers that bind successfully to be differentiated from those that do not. Early additions to SELEX focused on improving the methodology to achieve greater specificity and binding in more stringent conditions. Negative-SELEX is an important control in aptamer selections that typically precedes a normal SELEX experiment. In this case, the library is incubated with the matrix or column in the absence of the target in order to remove oligonucleotides that bind non-specifically to the medium itself.[5] Counter-SELEX also introduces an additional selection step but with similar targets in order to remove aptamers that do not bind specifically to the target.[57, 58] Domain targeted SELEX uses the irreversible denaturation of protein domains to select for aptamers that bind only to the native form of a domain of interest via a counter-SELEX step.[59] Other variants of SELEX select for binding to similar targets or epitopes, such as parallel-SELEX, which presents multiple targets at once, and toggle-SELEX, which uses multiple selection rounds with different

targets to isolate aptamers with plural specificity.[60] MAI-SELEX (multivalent aptamer isolation SELEX) selects for aptamers that bind to specific domains of a protein by performing an initial selection with the whole protein and then conducting several individual selection experiments with the enriched library in the presence of different subunits of the protein.[61] ES-SELEX (epitope-specific SELEX) introduces a competitor molecule in later selection steps that is a specific epitope of the protein target in order to elute aptamers that specifically bind to this epitope.[62] Conditional SELEX is a two-step process that selects for aptamers that require either the absence or presence of a regulator molecule in order to bind a protein target.[63]

Aptamer selection has also varied in the nature of the target itself. Though overwhelmingly tested on protein targets, aptamers have been selected for targets ranging from small molecules to cells. Cell-SELEX uses entire cells as the target of a selection experiment rather than individual surface proteins.[64] This has its advantages as surface proteins are present in their natural orientation when bound to the surfaces of cells as opposed to during typical immobilization methods used in the generation of aptamers.[65] Counter-SELEX steps can be implemented with different cell types in order to produce aptamers that specifically target a cell type of interest, which is particularly useful for targeted drug delivery. There have been several notable variants of cell-SELEX. Crossover-SELEX and hybrid-SELEX add the recombinant surface proteins to selection cycles, with crossover-SELEX introducing them alongside the cells within the same selection cycle and hybrid-SELEX introducing them on their own in a toggle-SELEX style with aptamers selected for binding to the cells and to the individual proteins.[66] Cell-internalization SELEX specifically selects for aptamers that bind to a targeted surface receptor and are internalized, selecting against aptamers that are internalized through other pathways and against aptamers that bind to the receptor but are not internalized.[67, 68] FACS SELEX (fluorescence activated cell-sorting SELEX) uses fluorescently labeled aptamers and a cell-sorting device to differentiate and separate cells that are bound or not bound to aptamers.[69] FACS SELEX allows for the simultaneous removal of aptamers bound to unwanted cell types (for counter-SELEX) and unbound aptamers. 3D cell SELEX uses three-dimensional cell cultures as the environments for selection, which are better able to mimic tissue structures and thus are more applicable for biological targeting applications.[70] Similarly, in vivo SELEX generates aptamers that bind better to

their targets in physiologically relevant environments. In vivo SELEX has been used for the selection of tumor-targeting aptamers in mice, with the added benefit of also selecting for nuclease resistance, though the entire SELEX process requires much more time as the library must be introduced intravenously and subsequently harvested from liver tissue during each cycle.[71]

1.3.2.4. Library amplification

The enriched library obtained from selection is amplified by PCR prior to future selection rounds or sequencing. Single-stranded aptamers result in double-stranded amplicons, which requires an additional step to isolate the single-stranded amplified aptamers from their complementary strands. In asymmetric PCR, for instance, the coding strand is preferentially amplified due to unequal primer concentrations. Double-stranded amplicons can be converted to single-stranded oligonucleotides by exonuclease digestion or systematically separated into coding and complementary strands by techniques such as denaturing gel electrophoresis or biotin-labeling coupled with streptavidin-decorated magnetic beads.[72] For each selection experiment, there are certain factors that affect amplification, such as the enzymes used, the annealing temperature, and the number of PCR cycles.[73] With regard to polymerase engineering, there has been plenty of research to tolerate modified aptamers during the amplification step; this will be discussed later. PCR can introduce by-products such as primer-dimers and contaminate samples with artifacts, with PCR bias influencing the observable enrichment of aptamers.[74, 75] Neutral-SELEX aims to provide a baseline for the effects of amplification on the library by running an experiment in parallel to normal SELEX but without the selection step. Hi-Fi SELEX (high fidelity SELEX) uses droplet digital PCR to eliminate amplification artifacts and facilitate the more efficient conversion of double-stranded amplicons into single-stranded aptamers using lambda-exonuclease.[76] Droplet digital PCR divides the library into nanoliter droplets, each containing the necessary reagents for amplification, which significantly reduces the heterogeneity of each droplet and results in less by-product formation than bulk amplification. Complementary ring mediated rolling circle amplification (CRM-RCA) has also been used for the amplification of aptamers during SELEX.[77] The advantage of this form of amplification is that the amplicons are single-stranded and do not require additional steps to convert double-stranded products to a single-stranded library,

though the library must be circularized prior to amplification using restriction enzymes and ligases. During SELEX, PCR is often used as a metric by which to assess the progress of the selection experiment. Quantitative PCR (qPCR) has been used to assess the level of enrichment of cell-bound aptamers in cell-SELEX in real time, providing the absolute number of aptamers, which is relevant to the proper optimization of ensuing PCR amplification cycles.[78] Two-step PCR is commonly applied, with analytical PCR followed by PCR amplification of the library. One recent improvement to this system is to use a quicker one-step specific PCR setup, relying on a dual-microfluidic amplified system (dual-MAS), which monitors the progress of the amplification in real-time to elucidate the ideal number of PCR cycles required.[79] The amplification of the library has also been modified in some SELEX experiments to prevent the interference of primer regions with target binding. In primer-switching SELEX, prior to amplification, entirely new primer regions are ligated to the aptamers after removing the previous primers with restriction enzymes, allowing for the newly amplified library to contain different primer sequences and therefore the selection of successful binding independent of primer sequences.[36]

1.3.2.5. Sequencing

Finally, when the aptamer library has been sufficiently enriched, the cycle ends after amplification and sequencing occurs to reveal the sequences of the most prominent aptamers. Conventional Sanger sequencing uses chain-terminating fluorescent nucleotides and capillary electrophoresis to differentiate and decipher the sequences of aptamers.[80] Next generation sequencing methods have been particularly useful for SELEX experiments with large aptamer libraries as they allow for high-throughput sequencing of samples with 10^8 sequences. Illumina sequencing, for instance, uses massively parallel sequencing by synthesis, with fluorescent nucleotides featuring blocked 3'-OH groups incorporated, imaged, and then fluorophores and protecting groups removed for future nucleotide incorporation. Third-generation sequencing exists as well, requiring no DNA synthesis, such as with nanopores to which voltages are applied that sequence strands based on differences in measured current. However, current approaches with aptamer selections use next-generation sequencing as it is widely available and adept at finding enriched aptamers in SELEX experiments. As computational advances have allowed for improved data analysis, the data obtained from high-throughput sequencing in SELEX

experiments have been used to study aptamer selection experiments beyond the enrichment of individual sequences toward conserved motifs and sequence alignment.

High-throughput SELEX (HT-SELEX) relies on sequencing samples taken after different selection rounds, not only after the final selection, which results in more information about the evolution of enriched sequences and can help discount PCR biases.[81, 82] These samples can be sequenced individually or simultaneously using barcodes introduced prior to sequencing that distinguish samples from different selection rounds.[83] Further work has focused on the development of algorithms to identify and group mutations that arise in later rounds of SELEX during amplification due to polymerase errors. AptaCluster and AptaMut are two computational methods that were used, respectively, to cluster families of mutated aptamers and to identify beneficial mutations.[84] COMPAS (COMmonPAttenS) software also allows for higher-level analysis of the sequencing data of SELEX experiments.[85] COMPAS provides quality control on the initial library by checking the positional distribution of individual monomers, the distribution of fixed length sequence motifs, and the length distribution of the randomized region. COMPAS also monitors the library diversity throughout the entire SELEX process and clusters similar sequences into families, which enables sequences to be identified earlier and sequence motifs essential for binding to be discovered. FASTAptamer is another software that can measure the distribution of sequence populations, cluster sequences into related families, identify sequence motifs, and monitor enrichment during the selection.[86] APTANI focuses on overall sequence frequency and specific secondary structure motifs, such as hairpins and loops, when clustering sequences from HT-SELEX.[87] AptaTRACE identifies sequence-structure motifs by clustering according to both primary and secondary structures.[88] AptaTRACE prioritizes shifts in structure over changes in abundance, allowing less prevalent motifs to be identified earlier on in the selection. Fast string-based clustering (FSBC) has been used for HT-SELEX and estimates target-binding clusters by searching for different lengths of overabundant strings using data from the final selection round.[89] Machine learning has even been used for the classification of aptamers in HT-SELEX via sequential multidimensional analysis algorithm for aptamer discovery (SMART-Aptamer), which takes into account motif enrichment, aptamer families, and secondary structures.[90]

1.3.3. Limits of SELEX

While versatile in its ability to generate aptamers for a variety of targets, SELEX does suffer from some limitations. An important practical constraint of SELEX is the lack of universal protocols, as there are many different variations of SELEX available that each have their own set of strengths and weaknesses and may not be applicable for selecting aptamers for a particular target. Automated SELEX has been performed and could make SELEX more widely applicable in the research community outside of the field of DNA nanotechnology, though it is not usually commercially accessible. For research laboratories without automated DNA synthesizers, the cost of SELEX experiments is much greater as library generation, amplification, and screening all require synthesized DNA.

Although RNA is more expensive and difficult to synthesize, particularly for longer sequences, it is often preferred due to the diversity of three-dimensional shapes it can form compared with DNA. RNA libraries, however, suffer from decreased stability and require additional enzymatic steps for PCR amplification, making their application in SELEX experiments more difficult. The DNA or RNA-based aptamers obtained from SELEX are also easily degraded by nucleases and limited to short circulation times *in vivo*, limiting the value of aptamer selections. However, there has been plenty of research into aptamer modifications to improve their physicochemical properties as well as their binding affinity.

1.4. Modified aptamers

Unmodified single-stranded DNA and RNA feature nitrogenous bases, ribose sugars and phosphate backbones. All three main features of nucleic acids can be modified to change their binding affinity and their physicochemical properties. **Figure 1.8** indicates the addressable functionalities of aptamers and indicates some of the examples of modifications that will be explored in this section.

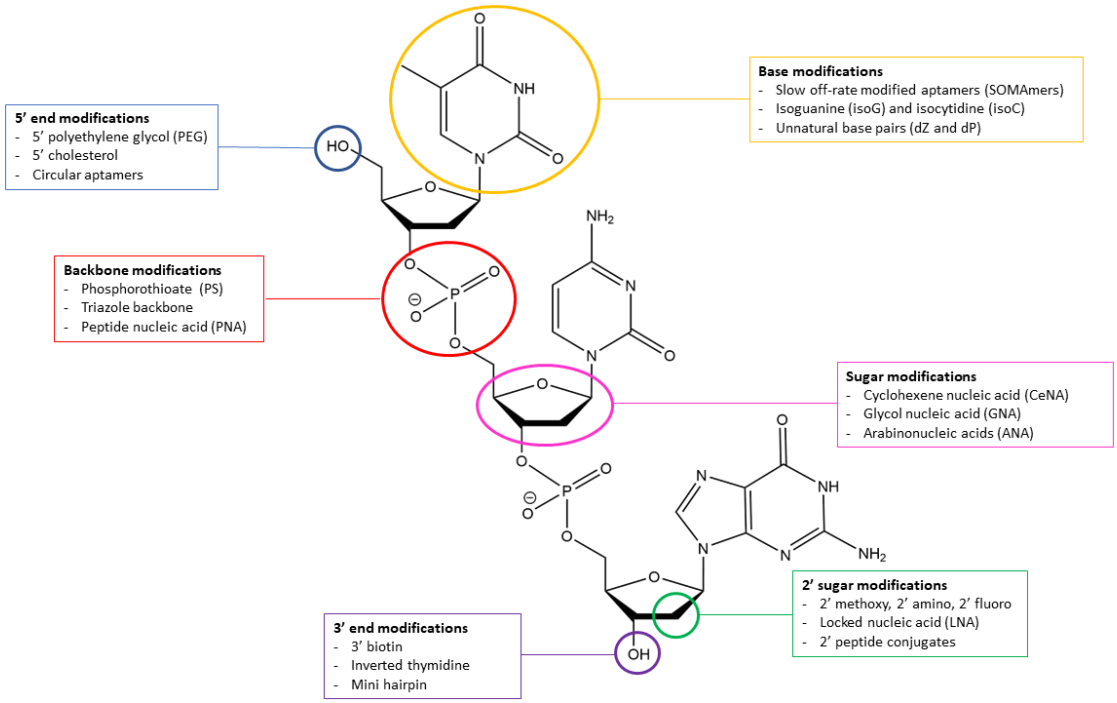


Figure 1.8. An overview of possible sites of modifications for aptamers is depicted above. Examples of specific modifications are given for each region of the aptamer. For simplicity, an unmodified DNA 3mer is shown.

1.4.1. Backbone modifications

The phosphodiester (PO) backbone is commonly modified to increase target binding affinity and improve stability to nucleases. Exonucleases hydrolyse the terminal phosphodiester bonds, irreversibly fragmenting strands of DNA and RNA in physiological conditions. Modifications of the non-bridging oxygens in the phosphodiester linkage have yielded promising results in the field of oligonucleotide therapeutics, which includes aptamers as well as siRNA (small interfering RNA) and ASO (antisense oligonucleotides). Phosphorothioate (PS) and phosphorodithioate (PS2) backbones replace, respectively, one and two non-bridging oxygen atoms with sulfur (**Figure 1.9**). Phosphorothioated oligonucleotides have shown improved nuclease resistance and cellular uptake[91, 92] and, in some cases, improved binding affinities.[93] They are easily synthesized by solid-phase synthesis using sulfurizing agents or thiophosphoramidites and have been incorporated successfully in a variety of aptamers.[94-97] However, it is worth noting that phosphorothioated aptamers introduce additional chiral centers to a normally achiral phosphodiester backbone. In the case of phosphorothioate, the conventional synthetic method does not favour a specific enantiomer and therefore allows for the synthesis of diastereomers when multiple modifications

are present. Diastereomers can be separated by anion exchange column chromatography, though only for short DNA with few modifications.[98] PCR amplification of phosphorothioated aptamers typically favour one particular diastereomer, depending on the polymerase used.[99] For many applications, the racemic mixture of diastereomers does not matter but it is an important consideration for mechanistic studies where the exact aptamer structure is needed. In the case of siRNA, chirality has much less of an effect on binding affinity compared to the strand's sequence but it can affect its resistance to ribonucleases.[100] Alternative synthetic strategies using building blocks with predetermined chiralities allow for the synthesis of specific diastereomers[101] and some post-synthetic separation methods exist using DNAzymes that preferentially cleave specific diastereomers.[102]

Boranophosphates (**Figure 1.9**) have also been introduced into the backbones of aptamers[103] and have shown increased exonuclease resistance while retaining a negatively charged backbone.[104] The backbone is more hydrophobic than PS and PO backbones due to the dispersed negative charge arising from a lack of lone pairs and different electronegativites.[105] As a result, the borane moiety (BH_3) has a higher lipophilicity,[106] which is a property favourable for cell membrane penetration, but is still very water soluble. Boranophosphate backbones have not been extensively incorporated into aptamers and selected aptamers typically do not benefit from substitutions with boranophosphate analogs. However, aptamer selections with boranophosphates still yield strongly binding aptamers in which the boron-modified monomers are important to their function.[103]

A structurally similar but electrostatically different modification is that of the methylphosphonate backbone (**Figure 1.9**). In this case, a non-bridging oxygen is replaced with a neutral methyl group that reduces the polarity of the backbone. This type of modification is important for crossing cell membranes that feature lipid bilayers, as size and polarity are two essential factors that influence the permeability of a membrane toward a molecule.[107] It has been incorporated into the thrombin-binding aptamer (TBA), producing a more biologically stable aptamer despite destabilizing the guanine-quadruplex.[108] The TBA, 5' GGTTGGTGTGGTTGG 3', is an anti-parallel quadruplex with four pairs of guanines forming two tetrads linked by TT and TGT loops. The neutral methylphosphonate backbone is also relevant for aptamers with targets that

are not positively charged. A similar modification is that of ethylphosphonate (**Figure 1.9**), which introduces a slightly longer alkyl chain in the backbone and produces aptamers with similar properties.[109] As was the case for phosphorothiated aptamers, both boranophosphate and methylphosphonate backbones introduce chirality at the linking phosphorus atom, though it is possible to synthesize and purify specific diastereomers with additional steps.[110] N3' → P5' phosphoramidate backbones replace the 3' hydroxyl with 3' amino (**Figure 1.9**) and have been found to improve the thermal stability of duplexes and triplexes.[111] The phosphoramidate backbone was used in the optimization of an RNA aptamer for the TAR (transactivation-responsive element) of HIV-1, yielding a modified aptamer with improved thermal stability but a slightly worse binding affinity.[112]

Chirality is used as a defense against nuclease degradation in the case of spiegelmers (**Figure 1.9**), which are mirror-image aptamers made up of L-DNA. Since endonucleases and exonucleases hydrolyze D-DNA, aptamers have been synthesized with L-DNA to avoid being substrates of nucleases and thus be more physiologically stable.[113] As they differ only in chirality, spiegelmers represent an interesting modification of aptamers in that their molecular composition is no different than DNA aptamers but their ability to be recognized by enzymes for which DNA is a substrate is substantially reduced. Spiegelmers have been successfully used to bind to and block the functions of hepcidin,[114] a regulator protein relevant to anemia, and amylin,[115] a peptide hormone relevant to pancreatic diseases.

There exist even more drastic modifications to the backbones of aptamers that eliminate the phosphate variants altogether. Formacetal backbones (**Figure 1.9**) have been introduced in aptamers to replace the phosphorus atom and its non-bridging oxygens with a conformationally similar methylene group. Formacetal backbones are achiral, uncharged, stable to nuclease degradation, and exhibit higher cellular uptake.[116] Formacetal backbones were introduced at specific positions of the TBA, with slight differences in activity observed at different structural regions of the guanine-quadruplex.[117] Triazole backbones based on click chemistry (**Figure 1.9**) have also been used in order to impart nuclease resistance to aptamers while also adding rigidity and hydrophobicity to the backbone. When incorporated in the thrombin binding aptamer, the triazole modification did not improve its binding affinity but did decrease its susceptibility to

nuclease degradation.[118] Peptide nucleic acids (PNAs) represent another interesting backbone for aptamers, as they are structurally similar to PO backbones but feature no sugar unit and are less rigid and more hydrophobic (**Figure 1.9**). PNA aptamers benefit from the biostability of antibodies as they lack the sugar phosphate backbone susceptible to degradation and benefit from the synthetic programmability of aptamers. The thrombin-binding aptamer (TBA), when replaced completely with a PNA backbone, was found to have improved binding affinity as well as much improved serum stability.[119]

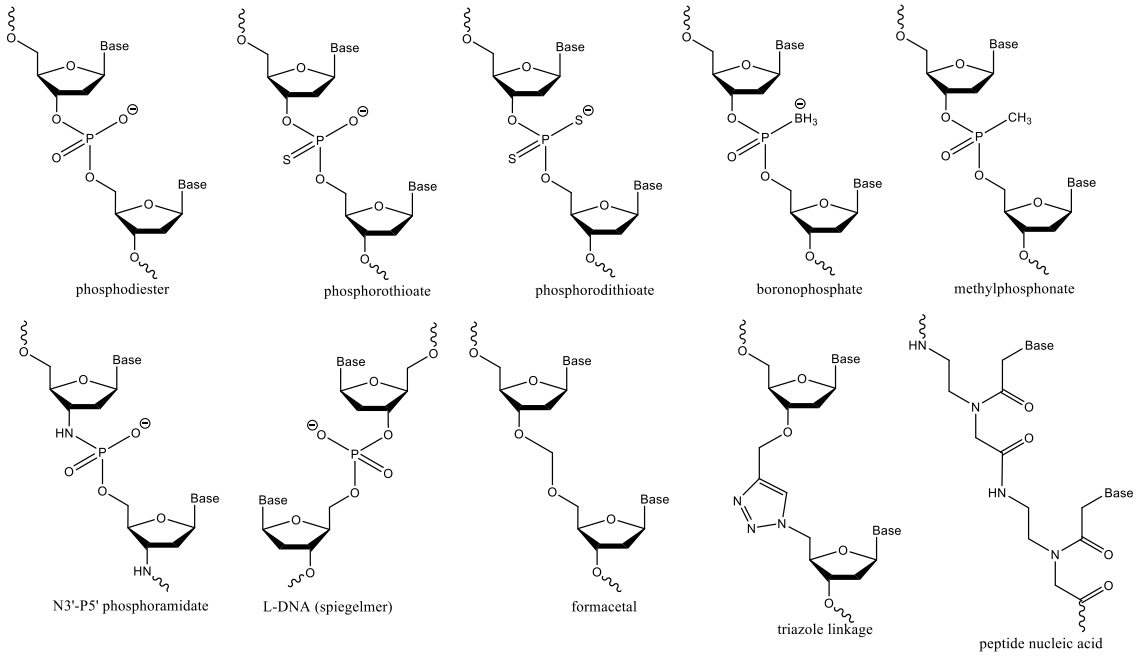


Figure 1.9. Modifications introduced into the backbone of DNA-based aptamers are shown above.

1.4.2. Sugar modifications

Aptamers are commonly modified at the sugar site as it tends not to interfere with base pairing interactions and can lead to improvements in stability and binding affinity. There are two main types of modifications relevant to the ribose and deoxyribose units of aptamers; modifications at the C2' position of the ribose ring (see **Figure 1.11**) and modifications that replace ribose with other cyclic or acyclic units (see **Figure 1.12**).

1.4.2.1. 2' modifications

RNA and DNA differ compositionally by the presence of a 2' hydroxyl group and the substitution of thymine with uracil, which differ only in the presence of a methyl group. The 2' substituent

has a large effect on the structure and stability of the two types of nucleic acids. They both form double helices but adopt different structures; DNA adopts a B-form helix and RNA adopts an A-form helix. This difference arises from two stereoelectronic effects relevant to ribose – the gauche effect and the anomeric effect. The gauche effect provides a basis for preferential conformations of electronegative substituents along a carbon-carbon bond. The presence of the 2'-OH for RNA introduces several additional torsional angles between electronegative atoms on neighbouring carbons of the ribose ring that favour a different pseudorotational orientation than that of DNA.[120] The anomeric effect for nucleic acids relates to the favourable overlap of the filled lone pair orbital of the endocyclic oxygen with the empty anti-bonding orbital of the glycosidic bond.[120] These two effects result in North (C3' endo, A-form) and South (C2' endo, B-form) pucker formations that influence the shape and spacing of the helices formed by RNA and DNA (**Figure 1.10**). In general, groups linked to the 2'-carbon of ribose via electronegative elements like oxygen and fluorine tend to favour the RNA structure and groups linked via less electronegative groups like nitrogen, carbon or sulfur tend to favour the DNA structure, although steric hindrance from large R groups can affect the preferred sugar conformation.[121] RNA is also less stable than DNA and this is related to the 2' OH group which increases the susceptibility to strand cleavage via ribonuclease-catalyzed nucleophilic attack. Ultimately, these effects are important to consider when modifying aptamers at the 2' position, as changes in structure and stability will affect their activity.

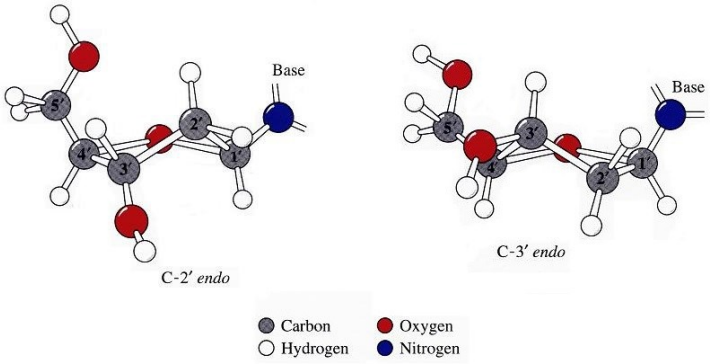


Figure 1.10. The two sugar pucker formations of DNA and RNA are shown above.[122] The left figure, C2'-endo, also known as the North pucker or A-form helix, is the structure adopted by RNA. The right figure, C3'-endo, also known as the South pucker or B-form helix, is the structure adopted by RNA. The two structures form duplexes with different helical spacings as a result of the

*angles of the CO bonds at the 3' and 5' positions. Reproduced with permission from reference **122**. Copyright 1994 International Union of Biochemistry and Molecular Biology, Inc.*

One of the first 2' ribose modifications incorporated in aptamers was the 2'-fluoro modification (**Figure 1.11)**. This modification replaces an electronegative oxygen with an even more electronegative fluorine. 2'-fluoro aptamers favour the A-form helix adopted by RNA.[123] There are many examples of aptamers incorporating this modification to increase nuclease resistance and improve binding affinity. Aptamers with 2'-fluoro modifications have been successfully selected for targets like human neutrophil elastase,[124] prion proteins[125] and viral cellular receptors.[126] 2'-amino groups (**Figure 1.11)** have also been introduced as modifications that improve *in vivo* stability and can improve binding affinity. They have a destabilizing effect on the thermal stability of nucleic acid duplexes, especially for those containing RNA as the modification favours the B-helix of DNA.[127] They have been incorporated in aptamers that target prion proteins[128] and human tumor necrosis factor α[129] and in both cases bore aptamers that were resistant to nuclease degradation and bound to their targets in the low nanomolar (nM) range.

Another widely used modification is the 2'-O-methyl group (**Figure 1.11)**, for which there is a foundation in nature. Methylation of the 2'-OH group is a common post-transcriptional modification of RNA and is highly conserved across different types of RNA such as transfer RNA (tRNA), messenger RNA (mRNA) and ribosomal RNA (rRNA).[130] The increased hydrophobicity and stability of methylated nucleic acids affects their interactions with biomolecules.[131] The methoxy modification has been implemented in several aptamers[132, 133] and has resulted in substantially increased nuclease stability and binding affinity, most notably for an aptamer targeting the vascular endothelial growth factor that led to the development of a clinically approved anti-angiogenic medicine.[134] 2'-O-methoxyethyl modified nucleotides (**Figure 1.11)** have been used extensively in antisense applications for their improvements in nuclease resistance and duplex stability[135, 136] but have also been incorporated in RNA aptamers that target the cytokine interleukin-23.[137] Longer alkyl chains tend to destabilize nucleic acid duplexes as a result of increased steric hindrance, though the introduction of charged moieties can improve duplex thermal stability.[138, 139] Azido groups have also been added to the 2' position of aptamers in order to conjugate bulky hydrophobic groups like 1-pentyne, 1-octyne and 4-phenyl-1-butyne using

click chemistry (**Figure 1.11)**, generating aptamers that bind to human neutrophil elastase with binding affinities below 100 nM.[140]

Finally, there are some more extreme 2' structural modifications that influence the rigidity of the sugar phosphate backbone. Locked nucleic acids (LNAs) introduce a methylene bridge between the 2'-oxygen and the 4'-carbon of RNA (**Figure 1.11)**, resulting in a locked 3'-endo conformation favoured by RNA.[141] LNAs demonstrate high helical thermal stability and are strongly resistant to nuclease degradation.[142] A DNA aptamer was synthesized with LNA modifications that allowed it to bind to a folded RNA target relevant to HIV type 1 with a low nanomolar binding affinity and with robust protection from nuclease degradation.[143] LNA modifications were also used in an aptamer targeting the protein Tenascin-C, with substitutions in the non-binding stem improving the thermal stability of the aptamer and LNA modifications generally improving plasma stability without detrimental effects on binding.[144] On the other end of conformational constraint are unlocked nucleic acids (UNAs), as the bond between the 2'- and 3'-carbons is cleaved, resulting in a highly flexible acyclic sugar unit (**Figure 1.11)**. UNA modifications tend to destabilize hybridized nucleic acids, though they can improve mismatch discrimination and increase thermal stability in some aptamers.[145] The TBA was found to be more thermodynamically stable when modified in the loops flanking the guanine pairs of a guanine-quadruplex and also found to be a better inhibitor of thrombin.[146] The TBA was also modified with a 2'-C-piperazino-UNA monomer (**Figure 1.11)**, a modification that is less destabilizing than regular UNA likely due to the protonation of the piperazino group,[147] which improved the thermal stability of the aptamer while not substantially affecting its binding affinity and activity.[148]

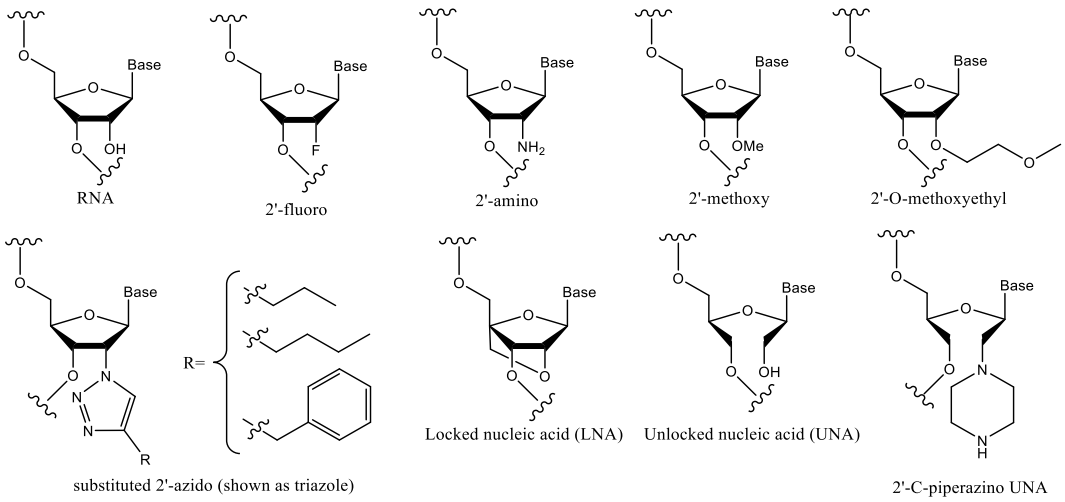


Figure 1.11. 2' sugar modifications present in the literature on modified aptamers.

1.4.2.2. Other sugar modifications

Arabinose is an epimer of ribose, having an opposite configuration at a chiral centre. Specifically, the 2'-OH group is directed above the sugar ring for arabinose and directed below the ring for ribose. Arabinose nucleic acids (ANAs) have been incorporated extensively in antisense oligonucleotides for their added stability (**Figure 1.12**). The change in stereochemistry results in a different favoured pseudorotational conformation, favouring the B-helix of DNA.[149] 2'-deoxy-2'-fluoroarabinonucleic acids (2'-FANA) are particularly prominent, with increased nuclease resistance and thermal stability (**Figure 1.12**).[150] FANA modifications in the TBA substantially improved serum stability, stabilized the quadruplex when modified in specific positions of the aptamer, and resulted in aptamers with improved binding affinity and thermal stability.[151] Another selection experiment with FANA found improved binding in modified TBAs and observed that the FANA modifications helped pre-organize the loops in the aptamer to optimize aptamer-protein interactions.[152] FANA modifications were also incorporated for an aptamer targeting HIV type 1 integrase, with picomolar affinity obtained and improved targeting compared with 2'-fluororibonucleic acid modified analogues.[153]

An alternative modification that improves nuclease resistance in aptamers is α-ribose (**Figure 1.12**), an anomer of β-ribose (present in RNA), which refers in this case to the orientation of the glycosidic bond with respect to the sugar ring.[154] When incorporated in the TBA, the α-2'-deoxyribose modification was beneficial to the thermal stability and inhibitory activity at certain

positions within the loops but detrimental when at core positions of the guanine-quadruplex.[155] α-L-LNA has also been used in aptamers as a stereoisomer of LNA that favours the A-form duplex of DNA but retains the added duplex stability and biostability of LNA (**Figure 1.12**).[156] A DNA-based streptavidin-binding aptamer was modified with both LNA and α-L-LNA, with the aptamer structure containing a stem-bulge-stem-loop motif.[157] LNA modifications improved binding affinity when in the stem regions, particularly in 3' region, but were detrimental when in the bulge or loop region, which are thought to be directly involved in binding. Similarly, α-L-LNA modifications within the terminal stem improved target binding and were detrimental to binding elsewhere.

Conformationally locked aptamers have also been generated using a carbocyclic bicyclo[3.1.0]hexane template, which consists of a cyclopropane unit conjugated to a 5-membered homocyclic ring featuring similarly oriented OH groups as DNA (**Figure 1.12**). The TBA features a guanine-quadruplex with guanine nucleotides oriented in specific conformations to engage in Hoogsteen hydrogen bonding interactions. Guanine nucleotides in the G-quadruplex are present in anti and syn conformations, which refer to the rotation of the glycosidic bond and affect the presentation of hydrogen bonding groups on the bases. Two types of locked methanocarbanucleotides, named North and South for the sugar pucker conformation they respectively favour, were incorporated into the aptamer at positions of the core quadruplex.[158] North puckers (such as RNA) favour anti conformations and South puckers (such as DNA) favour syn conformations, so unsurprisingly substitutions of regular DNA with the locked variants of the corresponding favoured conformation were found to be stabilizing to the aptamer.[158]

The ribose unit can also be modified by replacing the heterocyclic oxygen with sulfur, giving 4'-thio nucleic acids (**Figure 1.12**). Although excessive substitution with 4'-thio nucleic acids in the binding region of RNA aptamers targeting a nuclear factor resulted in a loss of binding affinity, partially substituted sequences had greatly improved serum stability and duplexes formed with RNA or 4'-thioRNA were more thermodynamically stable.[159] An aptamer modified with 4'-thioRNA was also selected for thrombin, with low nanomolar binding affinity achieved.[160] When the DNA-based TBA was replaced with 4'-thioDNA at loop positions, the modified aptamer displayed a twofold increased inhibition of thrombin activity.[161]

An interesting modification of the sugar is threose nucleic acid (TNA), which replaces the five-carbon ribose with a four-carbon threose sugar, resulting in a different 5-membered heterocyclic ring that connects to other monomers via 2' and 3' OH groups (**Figure 1.12**). Despite the shorter backbone repeat unit, TNA is still able to efficiently bind to DNA, RNA, and TNA. Like other modified structures of nucleic acids, TNA greatly improves the nuclease resistance of modified oligonucleotides and as such has been incorporated in a variety of aptamers. TNA aptamers for ATP (adenosine triphosphate)[162] and HIV reverse transcriptase[163] achieved comparable binding affinities to DNA analogues with substantial improvements in serum stability. 4'-methoxymethyl threose nucleic acid was incorporated into the TGT loop of the TBA, improving the activity, thermal stability and biostability of the aptamer.[164] Biostable and strongly binding TNA aptamers to PD-L1 (programmed death-ligand 1) were isolated and tested in mouse models, with results showing targeted inhibition of tumour growth via immunotherapy.[165] An alternative modification that retains the 5-membered ring but introduces a different linkage is that of isomeric DNA (isoDNA), which forms a 5' to 2' sugar phosphate backbone that favours the RNA sugar conformation (**Figure 1.12**).[166] When substituted in the TBA in place of natural DNA, the aptamer retained its ability to inhibit thrombin with improved nuclease resistance.[167]

There are also modifications to the sugar unit of aptamers that introduce 6-membered rings like HNA (1-5,-anhydrohexitol nucleic acid). HNA oligonucleotides (**Figure 1.12**) adopt RNA-like conformations, allowing them to hybridize stably with RNA, and exhibit higher rigidity compared to RNA.[168] An RNA aptamer for TAR RNA was modified with HNA nucleotides and the observed activities of fully modified aptamers were poor, despite improvements in the stabilities of stem structures within the aptamer.[169] However, aptamers with fewer modifications had similar activities to unmodified aptamers, particularly if modifications were not in the loop-closing residues, while also being less susceptible to enzymatic degradation. Fully modified HNA aptamers were selected for VEGF (vascular endothelial growth factor), with the equivalent DNA structures exhibiting binding affinities about five times worse than HNA aptamers and remaining stable in serum for multiple days fewer than HNA aptamers.[170] Another 6-membered ring modification is that of *p*-DNA, which stands for 3'-deoxyribopyranose (4'→2') nucleic acids (**Figure 1.12**). This interesting structure forms stronger duplexes with itself than do RNA and DNA

and features a more selective pairing system, as it participates only in Watson-Crick-Franklin base pairing and not Hoogsteen or reverse-Hoogsteen pairing.[171, 172] This modification was included in the FMN-binding (flavin mononucleotide) aptamer in the stem and stem-loop motifs that flank the bulge region responsible for binding to the target.[173] Modifications in these double-stranded regions stabilized the binding region and resulted in improved binding affinity compared with the unmodified aptamer.

A drastic modification of the sugar, which affects the backbone as well, is that of *S*-glycerol nucleic acids (GNAs) (**Figure 1.12**). Featuring an acyclic backbone of the same length as TNA, but with the conformational freedom of PNA and UNA, GNA introduces flexibility into oligonucleotides while also improving nuclease resistance.[174] When incorporated in a G-quadruplex aptamer targeting HIV integrase, single modifications in some loops proved beneficial to the inhibitory effect of the aptamer.[175] Finally, aptamers can be modified with spacers, which remove the base and sugar units entirely with the goal of increased flexibility or size. A three-carbon spacer (**Figure 1.12**) was added to the TBA in the TGT loop, with little effect on binding affinity but improved thermal stability and anticoagulant activity.[21]

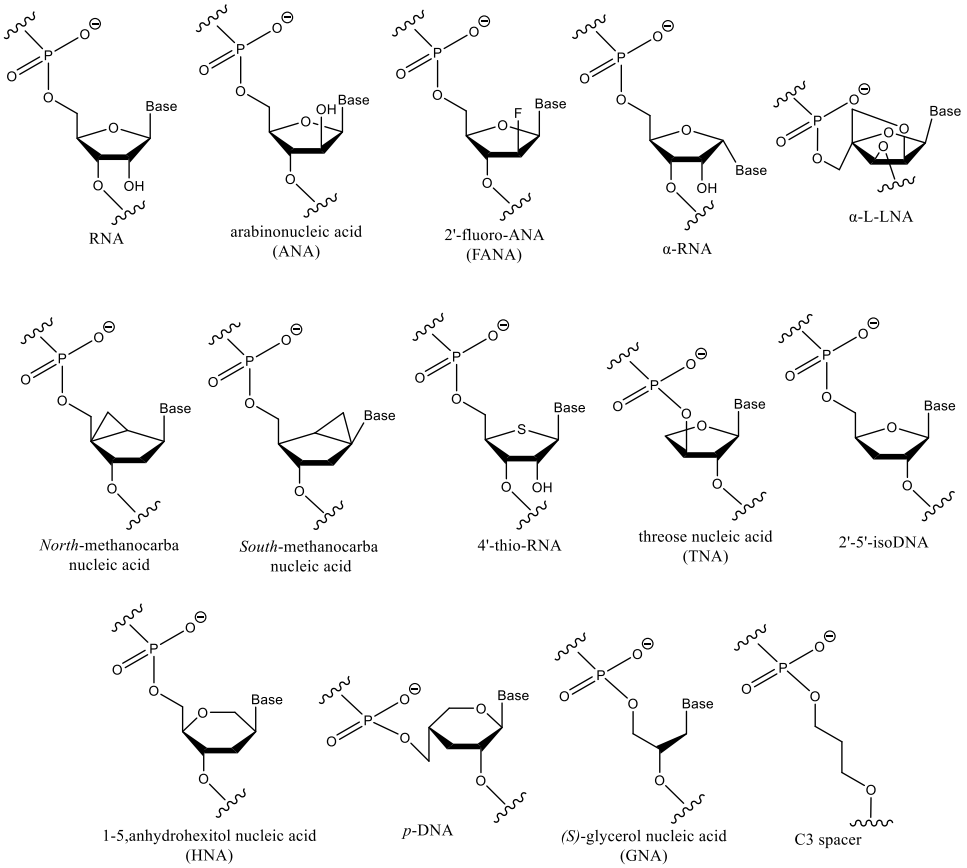


Figure 1.12. Sugar modifications that have been used in aptamers are depicted above.

1.4.3. Base modifications

The final major structural element of nucleotides is the nitrogenous aromatic base. Nucleic acids rely on Watson-Crick-Franklin hydrogen-bonding interactions to form double-stranded helices in a sequence-dependent manner. Generally, these hydrogen-bonding interactions are important to the structure and shape of aptamers, particularly with secondary structures like hairpins and loops. As such, modifications to the bases of aptamers focus on regions outside the Watson-Crick-Franklin hydrogen-bonding motif. These modifications aim to improve binding affinity and specificity with less emphasis on nuclease stability.

1.4.3.1. Pyrimidine modifications

The most common and earliest modifications of nucleobases were introduced at the C5-position of uridine, as it is easily functionalized and not involved in Watson-Crick-Franklin bonding. This position can be modified with groups like halides, alkyl chains, and even amino acids, and has

been prominently used in modified analogues of the TBA. 5-fluoro-2'-deoxyuridine (**Figure 1.13**) has been incorporated in the TBA in place of thymidine in the TT loops, with noted improvements in melting temperature and anticoagulant activity.[176] 5-hydroxymethyl-2'-deoxyuridine (**Figure 1.13**), which differs only slightly from thymidine, was individually introduced into the TBA at all thymidine positions, with the most positive effects seen in the TGT loop that is assumed not to be directly involved in binding.[177] The modification tended to stabilize the quadruplex structure and improved the anticoagulant properties of the aptamer when modified in the TGT loop, despite lower binding affinities for all modified aptamers.[177] 5-iodouridine has been used as a modification in SELEX experiments for its ability to crosslink with aromatic amino acids and cysteine residues under ultraviolet irradiation, allowing for better separation of binding and non-binding aptamers.[178]

3-aminopropyl and 3-amino-1-propenyl (**Figure 1.13**) have also been conjugated to the C5 of 2'-deoxyuridine, which introduce cationic functionalities to the nucleobase. An RNA aptamer binding to adenosine triphosphate was selected for with 5-(3-aminopropyl)uridine, producing an aptamer of similar binding affinity but with a sequence motif distinct from that of the normal RNA aptamer, indicating that the modification introduces different interactions with the target.[179] 5-(3-amino-1-propenyl)-uridine was determined to be paramount to binding to an oncogenic microRNA precursor with high affinity and specificity.[180] 5-(3-amino-1-propenyl)-2'-deoxyuridine, as well as 5-(indolyl-3-acetyl-3-amino-1-propenyl)-2'-deoxyuridine and 5-((4-hydroxybenzoyl)-3-amino-1-propenyl)-2'-deoxyuridine (**Figure 1.13**), were included in a selection for modified aptamers for a thymocyte antigen.[181] The best binding aptamer featured at least one occurrence of each modification and bound the target with a binding affinity of 57 nM.[181] 5-(indolyl-3-acetyl-3-amino-1-propenyl)-2'-deoxyuridine and 5-(methyl-3-acetyl-3-amino-1-propenyl)-2'-deoxyuridine were separately substituted in place of thymidine in the TBA and crystal structures were evaluated to rationalize binding effects.[182] Replacing one or two of the thymidine residues within a TT loop with the indole or methyl modification improved the binding affinity substantially, which was rationalized from the introduction of different stacking and hydrophobic interactions between aptamer and target.[182] Another cationic C5 pyrimidine modification with a longer carbon chain is 5-*N*-(6-aminohexyl)carbamoylmethyl-2'-deoxyuridine

(**Figure 1.13**), which has been used in the selection of aptamers in place of thymidine for targets such as sialyllactosylamine and thalidomide. For siallylactosylamine, the modified bases were beneficial to binding due to the positive charge of the modification and the negatively charged carboxyl residues of the target, particularly for modifications located at three-way junctions presumed to be directly involved in binding.[183] The modification also improved binding affinity and allowed for the enantioselective targeting of thalidomide,[184] and was expected to increase nuclease resistance based on earlier work.[185]

An aptamer to thrombin was selected using 5-(1-pentynyl)-2'-deoxyuridine (**Figure 1.13**) in place of thymidine, which resulted in sequences that when replaced with thymidine no longer bound thrombin, indicating different interactions at play with the hydrophobic substituent, although the overall binding affinities were worse than the normal TBA.[186] A similar 5-propynyl-2'-deoxyuridine modification (**Figure 1.13**) was incorporated in the loops of the TBA, with position-dependent stabilizing effects most influential in the TT loops.[187] A series of oligoethylene glycols were appended to 5-propynyl-2'-deoxyuridine, ranging from four to 16 monomers long within a chain, and tested in the TBA. All substitutions proved beneficial to the thermal stability of the aptamer without significantly affecting the structure of the aptamer, with the added enthalpic contribution likely from CH-π and lone pair-π interactions from the oligoethylene glycol units and some stability from π-π interactions from the alkyne.[187]

Another more recent modification that involves an alkyne is 5-(pyren-1-yl-ethynyl)-2'-deoxyuridine, which improves nuclease resistance in serum for the TBA.[22] Pyrene groups are able to partake in π-π stacking, allowing them to stabilize structures like G-quadruplexes via stacking or intercalation, though in this case the modification only stabilized the TBA when in the 3'-end loop nucleotides.[22] Pyrene-modified aptamers have also been designed using click chemistry by linking pyrene units to the C5 of 2'-deoxyuridine via a tetrazole linkage (**Figure 1.13**).[188] In this case, aptamers for epidermal growth factor receptor (EGFR) were modified with individual pyrene-modified deoxyuridines and resulted in considerably improved binding affinities in the subnanomolar range. Click chemistry has allowed aptamers to be functionalized with a variety of large groups at the C5 of uridine. For instance, a boronic acid moiety (**Figure 1.13**) was included in a selection for aptamers targeting fibrinogen.[189] Boronic acid interacts with hydroxyl groups

and diols, which allowed the moiety to focus binding to the glycosylation site of fibrinogen, though not with any improved binding affinity compared to unmodified aptamers. Cubane-modified aptamers (cubamers) (**Figure 1.13**) have also been designed to target a biomarker of malaria.[190] These unnatural structures can partake in non-classical hydrogen bonds and form hydrophobic cubane clusters, and although the binding affinity of a cubamer was worse than previously reported unmodified aptamers, there was added specificity that allowed the aptamer to target a particular variant of the biomarker.

Slow off-rate modified aptamers (SOMAmers) have been designed with amide linkages to include amino acid-like groups and other hydrophobic and aromatic moieties (**Figure 1.13**).[191] Aptamers with these modifications have been used for targets that are problematic for nucleic acid ligands to bind to, such as kinases and growth factors.[192] Modifications include hydrophobic and aromatic groups like benzene, naphthalene, tryptophan, isobutane, and thiophene.[193] As the chemical space of regular aptamers often limits them to positively charged targets, due in part to their negatively charged backbones, these modifications allow a more diverse set of interactions possible between target and aptamer. Amino acid moieties have been appended to nucleotides in other ways, such as by using a different amide linkage with a hexyl spacer.[194] Aptamers for glutamic acid were selected for using arginine-modified 2'-deoxyuridine and two modified aptamers had good binding affinities and displayed enantioselectivity.[195] A tryptophan-modified aptamer, linked this time to the C5 of uridine via click chemistry, was found to bind VEGF with four times better affinity than a published unmodified aptamer.[196] Nucleobases themselves have even been appended to uridine at C5 and used in the selection of high-affinity aptamers for camptothecin, which featured adenine-modified 2'-deoxyuridine and demonstrated improved binding compared with reference aptamers.[197]

Though less common, pyrimidines have been modified at the N3 position with carbohydrates, aromatic groups and amino acid derivatives. Though these modifications disrupt base-pairing interactions, they can affect aptamer-target interactions, particularly when outside of double-stranded regions of aptamers such as hairpins and stems. The TBA was modified with several N3-modified dT residues using click or amide chemistry.[198] Two specific positions of the aptamer were chosen for modification as they were part of the TT loops believed to be in contact with the

protein. Most modifications improved the thermal stability of the aptamer itself via stacking or hydrophobic interactions, particularly for tryptophan, phenylalanine, and leucine modifications. The modifications had more varied effects on the binding affinity of the aptamer, with serine, leucine, and carbohydrate modifications improving the binding affinity. Most modifications also improved the serum stability, though only the lactose modification improved the anticoagulant properties of the aptamer. 4-thio-2'-deoxyuridine has been incorporated into the TBA as well, which is a modification that increases the hydrophobicity of the aptamer. A modified aptamer with four 4-thio-2'-deoxyuridine residues inhibited thrombin activity twice as well as the unmodified TBA.[161]

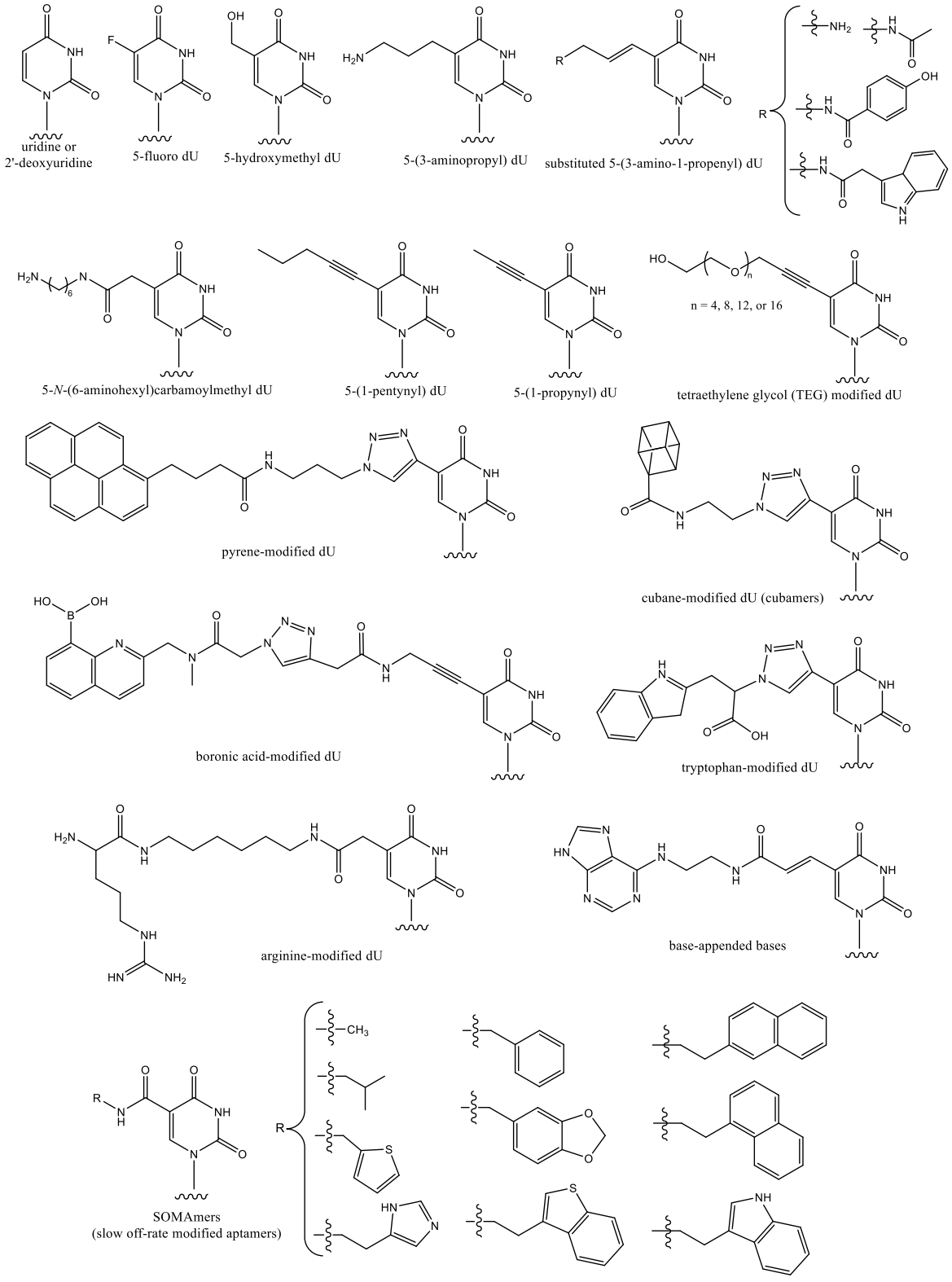


Figure 1.13. Pyrimidine C5 modifications incorporated in aptamers are depicted above. The sugar phosphate backbone is omitted for clarity.

1.4.3.2. Purine modifications

An early purine modification, 7-deaza-2'-deoxyguanosine (**Figure 1.14**), which substitutes the nitrogen at position 7 with carbon, was found to be quite destabilizing to the TBA.[199] Although N7 is outside the Watson-Crick-Franklin hydrogen bonding region, it is involved in hydrogen bonding in guanine quadruplexes, and so its inclusion in the guanine pairs of the TBA lowered the thermal stability of the complex. More recently, 7-deaza-7-phenyl guanine (**Figure 1.14**) was used in a TNA library, with aptamers for HIV reverse transcriptase selected having low nanomolar binding affinity.[163] Inosine (**Figure 1.14**), a compound identical to guanosine except for the lack of an amino group at the C2 position, destabilized the G-quadruplex of the TBA as the amino group was also involved in hydrogen-bonding integral to the structure of the aptamer.[200] 8-methyl-2'-deoxyguanosine (**Figure 1.14**), a variant that favours the *syn* conformation due to steric effects, was incorporated into the TBA as well, with stabilizing effects for *syn* G positions and destabilizing effects for *anti* G positions.[199] Similar results were found for 8-(1-propynyl)-2'-deoxyguanosine, but not for larger substituents like 1-phenylethynyl (**Figure 1.14**), where steric hindrance likely destabilized the modified aptamer.[201] 8-bromo-2'-deoxyguanosine (**Figure 1.14**) is beneficial to the thermal stability of the TBA when incorporated in *syn* residues of guanine tetrads and detrimental when incorporated elsewhere, as it stabilizes the *syn* conformation.[202] However, when both *anti* residues within a guanine tetrad were replaced with 8-bromo-2'-deoxyguanosine, the thermal stability was improved compared with the unmodified aptamer, though the binding affinity decreased.[203] 8-amino-2'-deoxyguanosine (**Figure 1.14**), on the other hand, is destabilizing to the TBA, due to reduced stacking interactions.[204] The amino modification is known to stabilize triplex structures, making it useful for favouring triplex formation in structures capable of both types of complexes.[204] 8-furyl-2'-deoxyguanosine and 8-(4-cyanophenyl)-2'-deoxyguanosine (**Figure 1.14**) were incorporated in the TBA with similar effects, destabilizing *anti* positions within the tetrads of the TBA and stabilizing the *syn* residues.[205] The advantage of these modifications is that they allow for a fluorescent response to be measured upon transition from single-stranded DNA to a guanine-quadruplex, which signals target binding.[206] 6-thioguanine (**Figure 1.14**) was found to destabilize the quadruplex of the TBA, thought to be from a reduced electronegativity that weakens the hydrogen-bonding interactions

of the tetrads.[207] Guanines have also been modified at the N2 amino position involved in Watson-Crick-Franklin hydrogen-bonding (**Figure 1.14**). Benzyl and 1-naphthylmethyl groups increased the activity of the TBA when in two specific *anti* residues on the backside of the aptamer but had little effect elsewhere.[201] In the same positions, 1-adamantylmethyl decreased the activity of the aptamer, with the contrasting effects likely due to different binding interactions with thrombin.[201]

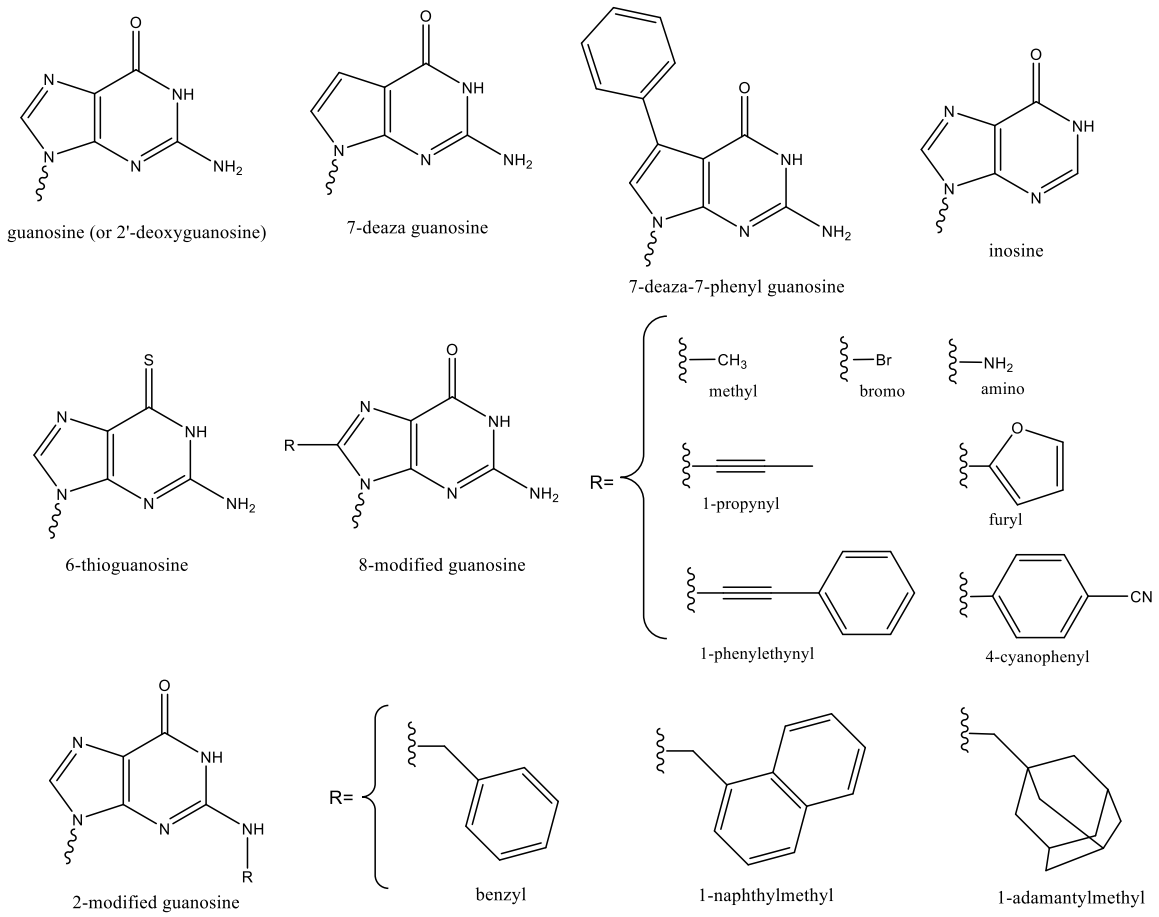


Figure 1.14. Modified purines that have been incorporated into aptamers are depicted above.

1.4.3.3. Artificial bases

Beyond modifications that append functionalities to purines and pyrimidines there are alternative base pair systems that present different binding interfaces. For instance, isoguanine and isocytidine represent an alternative base pair in which the amine and ketone substituents swap places (**Figure 1.15**). Isoguanosine actually occurs naturally and is known to form several high order structures like tetrameric and pentameric complexes.[208] When replacing guanine

residues in the TBA with isoguanine, loop modifications and modifications of *syn* residues in the quartets improved the binding affinity of the aptamer.[20] Artificial base pairs represent orthogonal base pairing systems that can replace or complement that of normal DNA and RNA, forming hydrogen bonds between purines and pyrimidines but with different donor-acceptor patterns. Artificial base pair 2-amino-8*H*-imidazo-[1,2-*a*][1,3,5]triazin-4-one (Z) and 6-amino-5-nitropyridin-2-one (P) have rearranged hydrogen bonding motifs (**Figure 1.15**). These artificial base pairs have been used in the selection of aptamers for the cell surface protein glypican 3.[209] The strongest binding aptamers had binding affinities in the low nanomolar range and featured Z and P bases that when replaced with any of the purines or pyrimidines did not retain binding to the target, indicating the modifications were important to the functionality of the aptamers. Beyond alternative hydrogen bonding motifs are hydrophobic base pairs such as Ds and Px (**Figure 1.15**). 7-(2-thienyl)imidazo[4,5-*b*]pyridine (Ds) and diol-modified 2-nitro-4-propynylpyrrole (Px) form an orthogonal base pair that relies on hydrophobic interactions. Ds was included in the selection of aptamers for VEGF-165 and interferon-γ, but without Px to prioritize interactions between Ds and the targets.[210] In both cases, aptamers were selected with affinities over 100-fold greater than those from a control library with only natural bases.[210] On the other side of unnatural bases is 5-nitroindole (NI) (**Figure 1.15**), which acts as a universal base, meaning it hybridizes with all normal DNA or RNA bases, though it relies on intercalation rather than base pairing.[211] Although it lacks hydrogen bonding groups, 5-nitroindole overlaps well with nearby bases based on stacking interactions and its *anti* conformation.[211] The TBA was modified with NI residues in the TT and TGT loops, with most modifications having little effect. However, replacing the guanine of the TGT loop resulted in much improved binding affinity and anticoagulant properties, supporting the idea that the TGT loop interacts with another site of thrombin.[212]

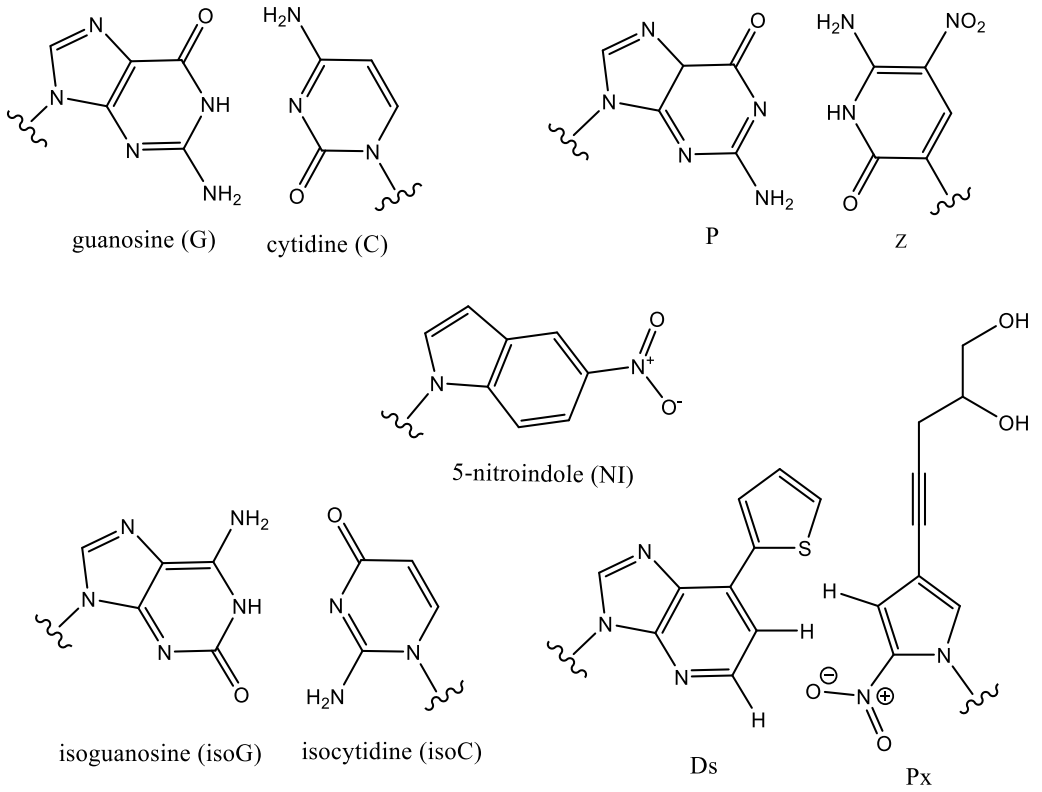


*Figure 1.15. Artificial bases that have been incorporated in aptamers are pictured above. P refers to 6-amino-5-nitropyridin-2-one, Z refers to 2-amino-8*H*-imidazo-[1,2-a][1,3,5]triazin-4-one, Ds refers to 7-(2-thienyl)imidazo[4,5-*b*]pyridine (Ds), and Px refers to diol-modified 2-nitro-4-propynylpyrrole. The sugar phosphate backbone is omitted for clarity.*

1.4.4. End modifications

Finally, aptamers are modified at the ends in order to improve their pharmacokinetic properties. As aptamers suffer in biological environments from nuclease degradation and renal filtration, 3' and 5' end modifications help address these limitations of aptamers. Aptamers are often at least tenfold smaller than antibodies, which helps them bind to a varied size range of targets but also makes them more susceptible to renal filtration. This limits their circulation time *in vivo* and reduces their effectiveness. A common strategy is to append large functional groups to the ends of aptamers to improve their cellular uptake and circulation times. Polyethylene glycol (PEG) has been widely used in drug delivery as it is non-toxic, very water soluble and FDA (Food and Drug Administration) approved.[213] PEGylated aptamers have been designed that show improved circulation times and reduced exonuclease degradation, most famously in the clinically successful pegaptanib, which features 5' PEG groups that increase the aptamer-based drug's half-life.[214] More recently, PEG was appended to the 5' end of aptamers targeting interleukin-17A using a

symmetrical branching phosphoramidite.[215] The PEGylated aptamers displayed improved stability and circulation times *in vivo* and were more biologically active. 3′ PEG was also effective for a spiegelmer targeting MCP-1 (monocyte chemoattractant protein), increasing the bioavailability of the aptamer *in vivo* as well as its activity.[216] Cholesterol (**Figure 1.16**) has also been commonly added to drugs to help with cell internalization and pharmacokinetic properties.[217] Cholesterol was conjugated to the 5′ end of a 2′-F RNA aptamer against an HCV (hepatitis C virus) polymerase.[218] The aptamer was effectively internalized and inhibited the replication of HCV, with the cholesterol modification improving half-life without inducing cellular toxicity. The surfaces of liposomes, commonly used as drug delivery vehicles, have also been decorated with nucleolin aptamers using a 3′-cholesterol group, allowing for better targeting of breast cancer cells compared with undecorated liposomes.[219] 5′ dialkylglycerol (**Figure 1.16**) has also been appended to modified aptamers targeting VEGF to improve circulation time.[220] The lipid group allows the aptamers to be anchored in liposome bilayers, improving their resistance to renal filtration without affecting their ability to bind to their target. Biotin (**Figure 1.16**) and biotin-streptavidin bioconjugates have been appended to the 3′ end of aptamers targeting thrombin.[221] 3′-biotin modified aptamers were more resistant to nuclease degradation but still suffered from short circulation times, while maintaining comparable binding affinities. 3′-biotin-streptavidin modified aptamers were resistant to nucleases as well and had comparable binding affinities but were much longer-lived *in vivo*.

Other modifications at the ends of aptamers focus on structural changes to impart increased nuclease resistance to aptamers. For instance, 5′-5′ and 3′-3′ inversion of polarity sites have been introduced in aptamers to improve their resistance to certain nucleases. A 3′ inverted thymidine residue (**Figure 1.16**) was added to SOMAmers for interleukin-6 receptor to impart 3′-to-5′ exonuclease resistance,[222] based on the observation that ribozymes displayed much higher serum half-lives with a 3′ inverted thymidine cap.[223] Inversion sites were introduced at the ends of the TBA, with effects seen on thermal stability, binding affinity and serum stability. An internal 5′-5′ inversion in the TT loop of TBA did not result in improved serum stability, as the 3′ ends were still easily cleaved by the more active 3′-exonucleases, though it did improve its thermal stability slightly.[224] Inversion sites between guanine pairs of the TBA prevented the formation of guanine-

quadruplexes.[225] 3′ inverted dTs improved the serum and thermal stability of thrombin aptamers, though adding 5′ inverted bases partially counteracted these improvements and on their own did not improve their resistance to nuclease degradation.[226] Another structural end modification that adds stability to aptamers are mini-hairpin DNA. DNA-based aptamers for interferon-γ were modified by adding to the 3′ end of a terminal stem a 9-mer consisting of a stem of three GC pairs and a GAA loop.[227] This end modification improved not only the thermal stability of the aptamer but also its serum stability, as the 3′ end of the terminal aptamer stem is covalently protected while the 5′ end is protected by base stacking.[227] A step further is to circularize aptamers and remove all available ends for exonuclease digestion. Circular DNA aptamers were selected for glutamate dehydrogenase, with two candidates showing high affinity and serum stability, though only one retained function when linearized.[228] Circular aptamers have potential in diagnostic applications as they are compatible with rolling circle amplification, which is useful for signal amplification. However, not all aptamers can be circularized without a loss of function, so this type of modification is not as versatile as other end modifications.

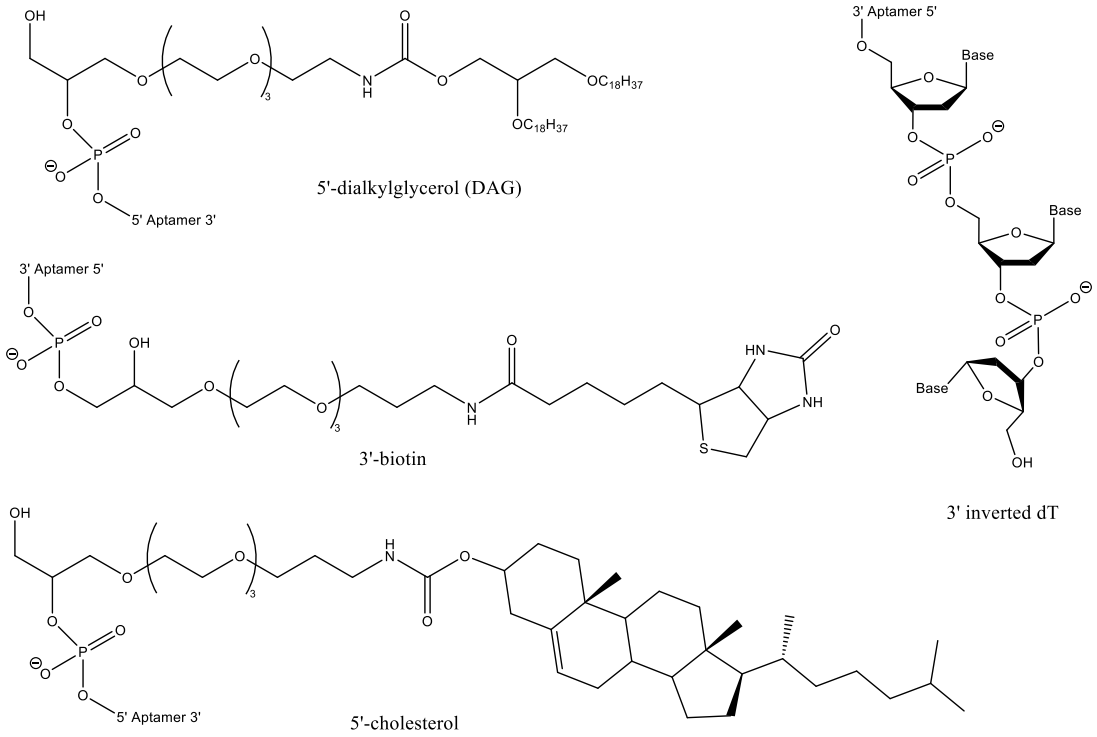


Figure 1.16. Some key end modifications that have been used in aptamers are shown above.

1.5. The selection of modified aptamers

Although modifications to DNA and RNA improve the pharmacokinetic properties of aptamers and may improve binding affinity, particularly with targets considered challenging for normal aptamers, several steps of SELEX do not tolerate modified oligonucleotides. The amplification step relies on polymerases to synthesize copies of the enriched aptamers and modifications that impair nuclease activity often also impair polymerase activity. Two main strategies to circumvent this issue are to engineer polymerases to accept modified nucleotides during amplification or to use DNA-templated synthesis to convert libraries between a modified form used in selection and a corresponding natural form for amplification (and sequencing).

1.5.1. Engineered polymerases

One of the most important benefits of modified aptamers is their superior resistance to nuclease degradation, which results from structural changes that impact how well aptamers behave as substrates for nucleases. These same changes are detrimental to DNA polymerases involved in amplification steps, as they cannot as easily incorporate these modified nucleotides into growing oligonucleotide chains. Initial attempts to amplify modified nucleic acids resulted in lower yields of incorporation for minor 2' modifications, particularly with heavily substituted libraries. RNA aptamer libraries require transcription and reverse transcription to temporarily convert the library to DNA for amplification. DNA polymerases are specific toward DNA over RNA and differentiate the two with the use of a steric gate that disqualifies nucleotides with 2' substituents.[229, 230] Mutations in polymerases that affect this steric gate have relaxed the substrate specificity of DNA polymerases, allowing for minor modifications to be tolerated. T7 RNA polymerase, for instance, has been engineered to accept modifications in nucleotides such as 2'-OMe, 2'-F, and 2'-NH_2.[134, 231, 232] These aptamer libraries require T7 promoter regions in the primers in order to successfully incorporate minor modifications in DNA. Minor backbone modifications are typically well-tolerated, with variants such as phosphorothioate[92] and boranophosphate[105] successfully incorporated by normal DNA polymerases using the appropriate modified nucleotide triphosphates.

Click-SELEX replaces thymidine residues with 5-ethyne-2'-deoxyuridine during library synthesis and amplification, relying on an unmodified Pwo polymerase (*Pyrococcys woesei*).[233] This allows

one functional group to be appended to 2'-deoxyuridine using click chemistry following library generation or amplification. Large C5 groups are not well-tolerated by polymerases but click-SELEX circumvents this limitation by introducing an additional step after amplification. A similar approach is to use 2'-azido modified nucleotides that can be tolerated by an evolved Taq (*Thermus aquaticus*) DNA polymerase and later coupled with hydrophobic groups using click chemistry prior to selection.[140] Click-PD (particle display) has also been used for the introduction of functional groups to aptamers, relying on click chemistry and emulsion PCR to get beads decorated with modified aptamers.[234] Using alkyne-modified dUTP (2'-deoxyuridine triphosphate) and DBCO-modified (dibenzocyclooctyne) dUTP that are tolerated by DNA polymerases, mannose and boronic acid groups, respectively, can be appended to the 2'-deoxyuridine positions of aptamers.

SOMAmers use amino allyl modified uridines in place of thymidine for amplification of the modified library using KOD (*Thermococcus kodakaraensis)* DNA polymerase.[191] During PCR amplification, the antisense strand is biotinylated and made of normal DNA using the modified SOMAmer library as the sense strand. The antisense strands can be easily isolated following amplification using streptavidin coated beads and then used as the template for the synthesis of a single-stranded modified library using KOD DNA polymerase. Spiegelmers, which are made of L-DNA, are typically selected for using mirror images of the protein targets and normal DNA, which indirectly results in sequences that can be used for L-DNA aptamers to target the proteins of interest.[113] More recently, L-DNA has been synthesized enzymatically using the mirror images of polymerases and using L-DNA building blocks.[235] It was also found that the polymerase and its mirror image could operate orthogonally in the presence of both chiral forms of nucleotides.

Xeno nucleic acids (XNAs) feature different carbohydrate ring structures and are not well recognized by normal polymerases, especially those that take on different structures from the canonical nucleic acids. In this case, it is typically necessary to convert the library into DNA prior to PCR amplification using engineered transcriptases and reverse transcriptases. If XNAs can bind specifically to DNA, can be synthesized enzymatically using a DNA template, and can serve as a template for DNA synthesis, they can be used in SELEX experiments. Variants such as LNA, FANA and HNA have all been successfully used in SELEX by relying on engineered XNA polymerases such

as Tgo (*Thermococcus gorgonarius)* polymerase.[236] Compartmentalized self-tagging was used to select for XNA polymerases and reverse transcriptases that displayed activity toward the different types of XNA. In this process, bacterial cells expressing single polymerase clones are isolated using an emulsion and primer extension with modified nucleotide triphosphates is used to identify the polymerases most capable of DNA templated XNA synthesis based on which contain the most stable primer-plasmid complexes.[237]

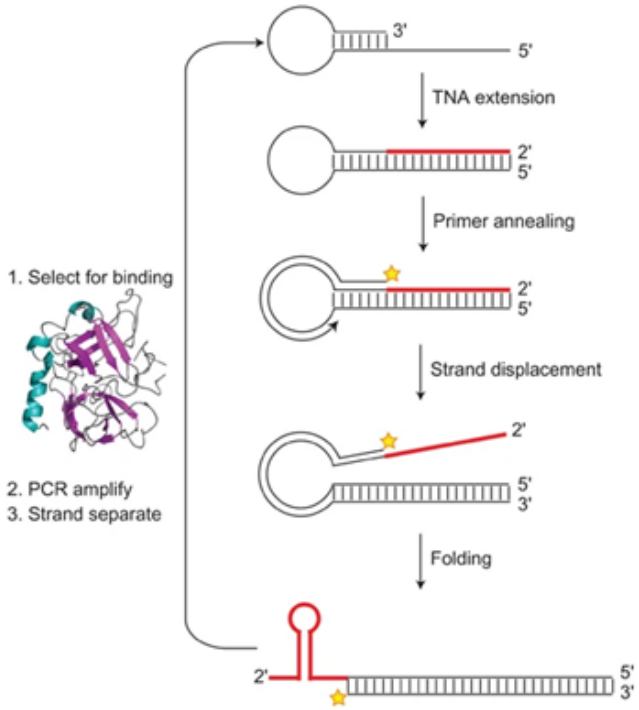


Figure 1.17. Above is shown the methodology of DNA display, in which a TNA aptamer is encoded by a covalently linked DNA duplex.[238] A single-stranded DNA library is designed with a 5' fixed primer-binding site and the 3' end consists of a stem-loop structure that also contains a primer-binding site. This single-stranded library is extended at the 3' end with TNA complementary to the randomized region of the library, forming a chimeric hairpin. A fluorescently tagged primer is used to synthesize a strand complementary to the entire DNA sequence, which displaces the TNA region. The library of DNA-displayed TNA aptamers is incubated with the target, allowed to bind, and separated via capillary electrophoresis. The duplex is amplified by PCR and made single-stranded to begin the cycle anew. Reproduced with permission from reference ***122****. Copyright 2012 Springer Nature.*

TNA has also been used in SELEX experiments using Therminator DNA polymerase to convert a DNA template into TNA, though early issues were encountered with the transcription of diverse libraries compared to individual sequences as a result of incompatibilities of the enzyme with repeat guanine regions.[239] An interesting design of a TNA aptamer library used a DNA template containing a stem-loop structure consisting of the anti-sense DNA template as well as the primer (**Figure 1.17**).[238] This design allows TNA to be extended from the primer region to create a single-

stranded hairpin. Another DNA strand can be extended from a fluorescently labeled primer spanning the constant stem-loop region to create a strand that displaces the TNA from the TNA-DNA double-stranded region, turning it into a single-stranded TNA aptamer covalently linked to a double-stranded DNA template. Denaturing the duplex following selection allows for subsequent amplification of the DNA template by PCR and subsequent regeneration of the enriched TNA library. This approach, termed DNA display, eliminates the need for an enzyme that converts TNA into DNA.

1.5.2. Expanded genetic code

A key limitation of the above methods is that they restrict selection experiments to just four building blocks, as modified bases can only replace their unmodified counterparts. In order to introduce multiple different types of modifications alongside unmodified nucleotides, modifications must be introduced after SELEX via trial-and-error or the library must be setup differently to accommodate more than four building blocks.

Post-SELEX modifications are commonly used to enhance the affinity and physicochemical properties of aptamers discovered through in vitro methods.[240] The key limitation of this method is that the structural changes imparted from modifications may affect binding affinity and specificity in ways that are challenging to predict. Phosphorothioate backbones are often introduced post-SELEX in biologically relevant aptamers to improve stability and cell permeability, though fully phosphorothioated aptamers can have toxic effects on cells. The usefulness of adding modifications after SELEX is that they need not be compatible with PCR amplification or sequencing, so non-nucleosidic monomers may be included in aptamers and it is possible to exceed four building blocks. However, this approach is not easily scalable and is typically only performed with modifications that are limited in number and structural variety. Sequencing results can be used to find which regions of aptamers are more important for binding and which stabilize the aptamer itself, allowing the focus of post-SELEX modifications to specifically target improved binding or improved stability.

Looper SELEX, standing for ligase-catalyzed oligonucleotide polymerization, uses DNA barcoding and DNA templated synthesis to expand the number of modifications that can be simultaneously

included in an aptamer library.[241] Normal DNA-templated synthesis of modified aptamers is restricted to four monomers, as they are each coded for by its unmodified analogue. Looper SELEX uses T4 DNA ligase to connect 5-mer building blocks appended with different groups on adenosine, defined as ANNNN, where N can be A, C, G, or T. In this way, the DNA template of unmodified DNA codes for a specific sequence of building blocks that are then ligated together. This method was used to create 16 different sub-libraries, each 5 bases long with a uniquely modified adenosine, with modifications ranging from aliphatic and aromatic groups to basic and acidic groups.[242] Following selection, the library was amplified using KOD polymerase and the modified aptamers were regenerated using the obtained DNA templates, with separation of template and aptamer via conjugation with biotin and separation with streptavidin. While the method is limited as modifications occur only every five bases and there can be biases toward different modifications based solely on their codon sequence, Looper SELEX represents an exciting development of the SELEX of modified aptamers with expanded genetic codes.

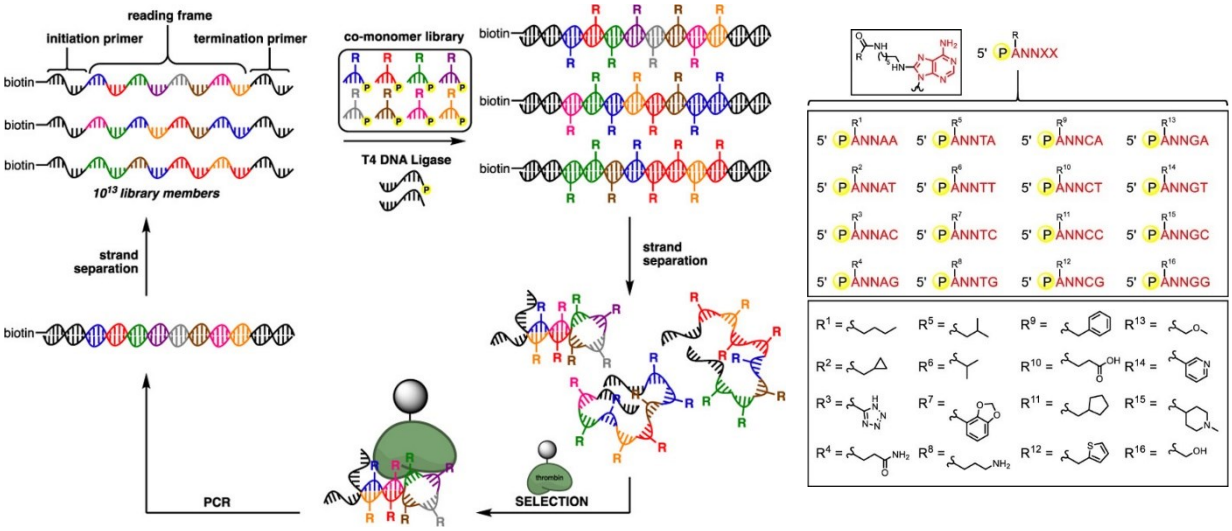


Figure 1.18. Looper SELEX is shown above.[241] In the left figure, the process of the selection is shown. A biotin-labeled single-stranded library of DNA is designed with fixed primers and a randomized region consisting of eight codons with the general sequence 5' NNNNT 3', where N is any of the four standard DNA nucleotides. This set of 256 different codons corresponds to 16 distinct families of anticodons, with each family featuring a uniquely modified adenosine nucleotide, as shown in the right figure. Each anticodon has two unique identifying bases (XX), in the form 5' ANNXX 3'. Using T4 DNA ligase, a modified aptamer library can be grown using the DNA strand as a template, as shown in the top figure. Following strand separation facilitated by the biotin tag, the single-stranded aptamer library containing modified bases can be incubated with the target in order to select for high affinity modified aptamers. The selected aptamers can be amplified by PCR to regenerate the unmodified DNA strands containing

*the codons. This enriched library can be used as the input for another cycle or can be sequenced to identify the sequences of the isolated modified aptamers. Adapted with permission from reference **241**. Copyright 2017 American Chemical Society.*

Codons are one way to increase the genetic diversity of aptamer libraries. Another way is to introduce orthogonal bases so that aptamers can be made enzymatically with more than four bases. AEGIS-SELEX (artificial expanded genetic information systems) introduces base pair Z and P that feature hydrogen-bonding motifs orthogonal to those in AT and CG pairs.[243] These modified base pairs were incorporated into aptamers using Hot Start Taq DNA polymerase during library generation and amplification. Sequencing required more steps as DNA sequencing methods are not equipped to deal with bases outside of the four canonical nucleotides. A conversion strategy was required to elucidate where the artificial bases occurred in the aptamers.[244] The enriched library was barcoded, amplified and then split in two and underwent two different conversion strategies. Barcoded sequences with notable changes between samples indicated the presence of the artificial bases.

The hydrophobic base pair Ds and Px has also been used in SELEX experiments as a third orthogonal base pair.[245] As with Z and P, it can be incorporated enzymatically alongside normal DNA bases using DeepVent DNA polymerase and thus represents a way to increase the chemical space of aptamer libraries for SELEX.[246] The SELEX experiments involving this base pair chose to incorporate only Ds to focus on interactions with targets instead of base pairing. In one setup, Ds bases were placed at up to 3 specific locations in the library and each variant was identified using unique 3-mer codons attached during synthesis. In another setup, the location of Ds bases was randomized and was identified using replacement PCR and a modified form of Sanger sequencing using complementary Px strands.[247]

1.5.3. Beyond SELEX

SELEX as an aptamer discovery platform is versatile and has been used successfully for the selection of aptamers. However, it does have clear limitations when focusing on modified aptamers, particularly for expanded genetic alphabets. There have been several attempts to perform aptamer selections without relying on the multiple rounds of selection and amplification required in SELEX (**Figure 1.19**). Firstly, non-SELEX uses multiple cycles of selection but uses only one round of amplification followed by sequencing. Secondly, single round SELEX uses just one

full cycle of the SELEX process, performing only one selection step followed by amplification and sequencing. Both methods of aptamer selection offer promise for the selection of modified aptamers, though to date they have primarily been used with unmodified aptamers.

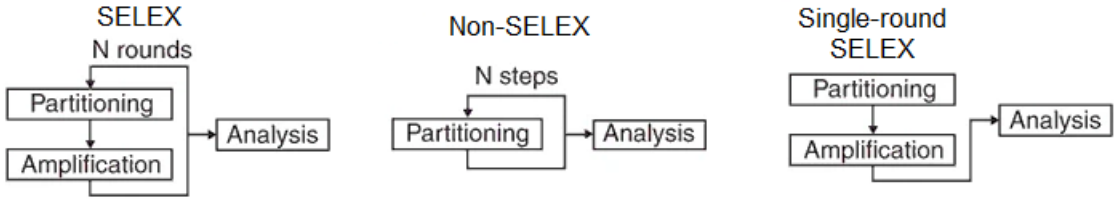


*Figure 1.19. Above is shown a schematic comparison of conventional SELEX with non-SELEX and single-round SELEX.[248] SELEX consists of multiple rounds of alternating partitioning and amplification before analyzing the enriched library. Non-SELEX eliminates the amplification steps of the cycle and performs repeated partitioning steps before analyzing the enriched library. Single-round SELEX uses just one cycle of the normal SELEX process, with no further partitioning steps following amplification. Adapted with permission from reference **248**. Copyright 2006 American Chemical Society.*

1.5.3.1. Non-SELEX

Non-SELEX refers to the selection of aptamers using repetitive partitioning steps without intermediary amplification steps, with amplification only occurring if needed for sequencing. A prominent example of non-SELEX uses the non-equilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) for the important partitioning step, similar to CE-SELEX, but without the use of amplification between partitioning steps.[248, 249] Eliminating repetitive PCR reduces the time required to select aptamers and can help limit biases introduced by PCR inefficiencies. The method is best suited for larger targets like proteins and peptides as separation relies on different migration times for aptamers when bound or not to the target. It is also limited to smaller library sizes given the small injection volumes and fractions collected for efficient separation. However, the method would also be compatible with DNA-encoded libraries, which could lend itself well to modified aptamer libraries. RAPID-SELEX (RNA aptamer isolation via dual-cycles SELEX) uses a combination of non-SELEX and normal SELEX to reduce the time of SELEX experiments by skipping unnecessary PCR amplification steps (**Figure 1.20**).[250] This hybrid SELEX variant is particularly useful for non-DNA libraries, such as RNA libraries, that require additional steps for amplification to occur and can introduce additional biases. SELCOS, the systematic evolution of ligands by competitive selection, also eliminates the amplification steps between partitioning steps and has been used to obtain very selective aptamers that can target specific viral subtypes.[251] Using two different types of targets immobilized on beads, the aptamers that

bind to either of the two targets are retained while non-binding ones are removed. Subsequent cycles with the undesired target allow for the removal of aptamers that are not specific to the intended target and PCR amplification occurs after all partitioning cycles are complete. Finally, an interesting non-SELEX method uses a centrifugation-based partitioning for bacteria-targeting aptamers.[252] Bacteria can be easily grown in suspension and separated from aqueous mixtures by centrifugation, allowing for the removal of unbound aptamers in selection experiments by removing the aqueous layer following centrifugation.[253] This methodology consists of multiple centrifugation-based partitioning cycles but only one round of elution and amplification, which greatly simplifies the process of obtaining aptamers. The main motivation behind non-SELEX is to simplify the selection of aptamers by reducing the number of amplification steps, which can help with PCR biases and, more practically, the cost and time required to perform these selection experiments. As with single round SELEX, it would be feasible to use non-SELEX for the selection of modified aptamers that cannot be amplified between cycles and so also represents a direction for the future of SELEX with modified aptamers.

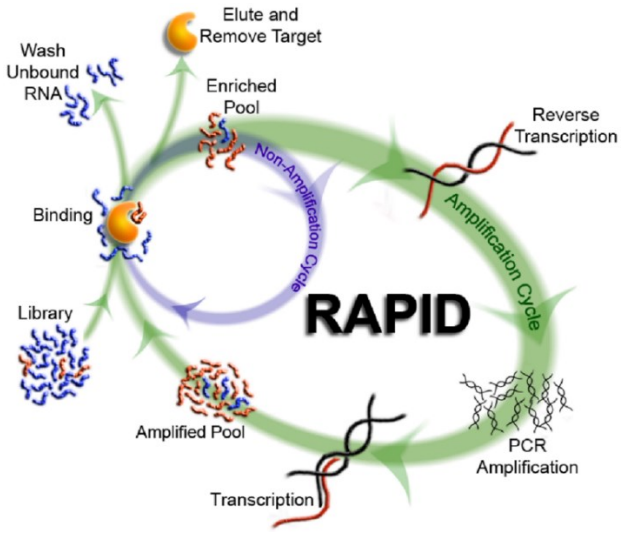


*Figure 1.20. RAPID-SELEX is depicted above for a selection of RNA aptamers.[250] It uses cycles of regular SELEX (green circular arrow) and non-SELEX (blue circular arrow), with the reverse transcription, PCR amplification, and transcription steps not included in non-SELEX steps. Reproduced with permission from reference **250**. Copyright 2013 Szeto et al.*

1.5.3.2. Single round SELEX

In SELEX, multiple rounds of enrichment allow for the discovery of strongly binding aptamers. Typical experiments, if ended after only one round, would not yield sufficiently enriched

aptamers to allow for the identification of key aptamers. However, with more stringent conditions pertaining to library design, selection parameters, or even sequencing, it is possible to limit SELEX to one cycle only. Virus-binding DNA aptamers have been selected for using MonoLEX, in which the physical segmentation of columns is used to enhance the separation of aptamers with different binding properties.[254] This one step protocol begins the same way with a synthetic DNA library that is incubated with the affinity column without the target to eliminate non-specific aptamers, followed by incubation in an affinity column with resin-bound viral particles to eliminate non-binding aptamers. The column was then physically divided into slices and the number of aptamers in each slice was estimated using quantitative real-time PCR, with higher affinity aptamers expected to be nearer to the start of the column. Aptamers were eluted from slices with high aptamer populations and amplified by PCR prior to sequencing.

NanoSelection is another aptamer selection experiment relying on just one round of selection (**Figure 1.21**).[255] It uses a library of bead-bound fluorescently labeled oligonucleotides and performs selection on a glass coverslip functionalized with immobilized targets. Selection occurs by flowing the library of beads over the surface and washing away non-binding aptamers. The locations of fluorescent signals, where strongly binding aptamers are expected to be found, can be observed with a fluorescent microscope and used to direct the tip of an atomic force microscope (AFM) to capture a high-resolution image of the bead and then extract it. Small copy number PCR is used to amplify the individual aptamers obtained by AFM, which are then sequenced prior to characterization. While this method is not well-suited to large and diverse libraries, it can be used to complement other varieties of SELEX or with enriched or focused libraries. It provides additional information about binding sites and can be used to prioritize aptamers that bind to a specific site of a target. PCR bias is also no longer an issue as the samples amplified are homogeneous in nature.

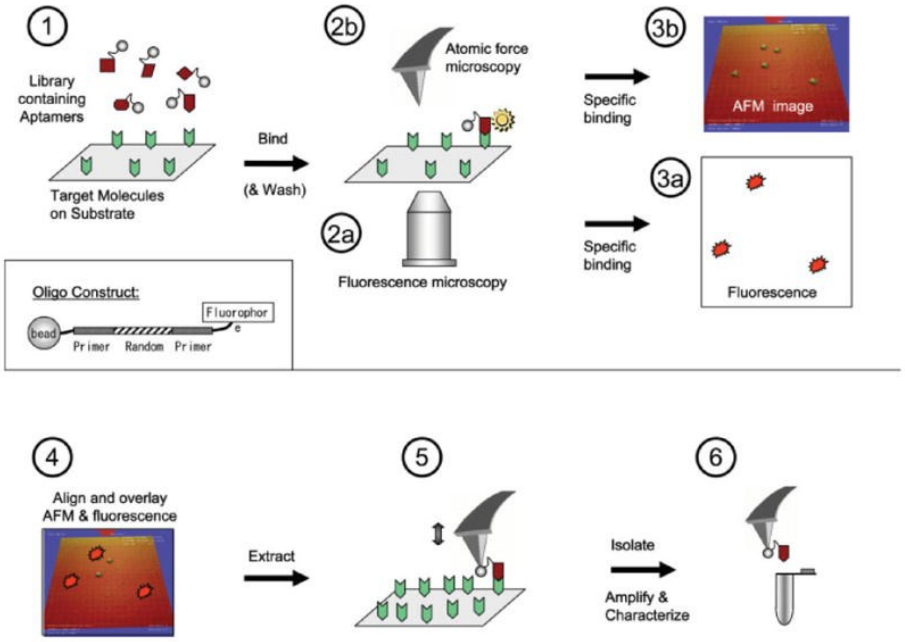


*Figure 1.21. NanoSelection is shown above schematically.[255] A library of oligonucleotides attached to both a bead and a fluorophore is flowed onto a substrate decorated with target molecules and non-binding samples are washed away (step 1). Fluorescence and atomic force microscopy (AFM) are used to detect binding events and capture images (steps 2 and 3). The location of the aptamers is determined (step 4) and used to physically extract them with the AFM tip (step 5), after which amplification and sequencing enable identification of the aptamer (step 6). Reproduced with permission from reference **255**. Copyright 2007 John Wiley and Sons.*

Single round SELEX has also been achieved thanks to improvements in sequencing methods and bioinformatics. An acyclic SELEX process used deep sequencing to analyse overrepresented libraries of thrombin-binding aptamers following only one cycle of selection and amplification.[256] This technique is limited to libraries with fully randomized regions of 25 bases for adequate over-representation, which refers to the average number of copies of unique sequences within a library and relates to the likelihood of all possible sequences being present. The original library used for the selection of TBA[7] was underrepresented as it used a 60-base long randomized region with the same amount of library, which meant that it was statistically unlikely for any unique sequences to have any copies present in the library. The combined use of deep sequencing and library design parameters allowed for more aptamers to be discovered with high signal to noise ratio and in only one cycle. Another single round SELEX method, called aptamer selection by k-mer analysis of sequences (ASKAS), used high throughput sequencing and bioinformatic analysis to discover thrombin-binding aptamers in one round from a library with a 33 base randomized region.[257] The high-throughput sequencing looked at sub-sequences within the 33mer region in

order to find conserved motifs of different lengths. The advantage of this method is that it can focus on different sized motifs, though there is a natural bias toward shorter regions and as the length of the randomized region increases it will likely be more difficult to identify longer aptamers.

Another aptamer selection variant uses a stringent selection protocol to isolate high binding aptamers in just one cycle. This method relies on DNase (deoxyribonuclease) digestion and washing with urea to remove weakly bound sequences from the immobilized targets (**Figure 1.22**).[258] Starting with a highly diverse library of 10^{15} unique sequences, each present around ten times, aptamers are incubated with the membrane-immobilized target and washed with buffer, then with DNase and a gradient urea buffer. DNA can resist enzyme-mediated cleavage when tightly bound to targets,[259] allowing this step to remove any weakly bound aptamers. DNase I digestion, followed by washes with increasing concentrations of urea from 7 to 9 M, was able to reduce the number of molecules left to under 10^3, which greatly simplifies sequencing analysis and the identification of strongly binding aptamers.

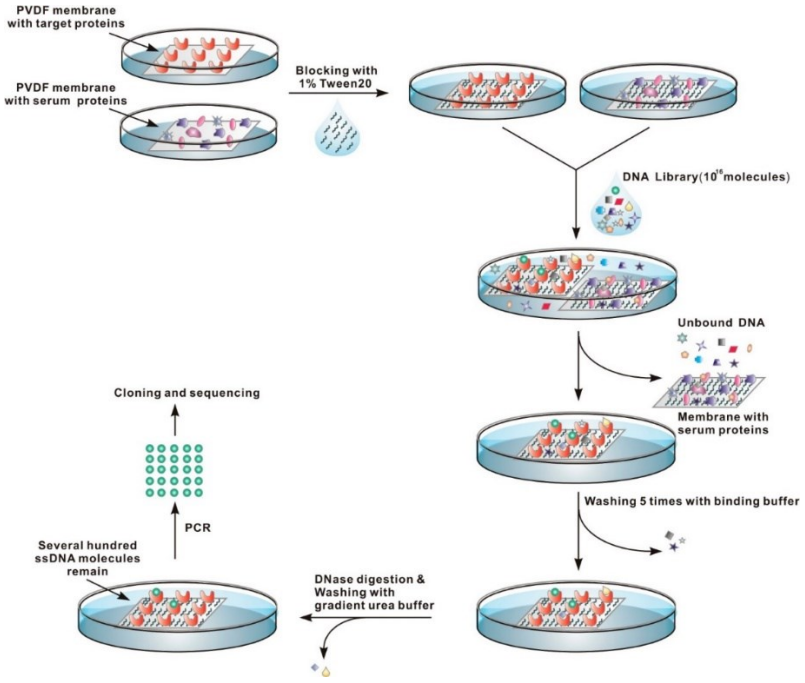


Figure 1.22. Above is a scheme for a single-round selection of aptamers using DNase digestion.[258] The process begins first with a counter selection step to remove aptamers that bind to the serum proteins. Non-binding aptamers are removed by washes with binding buffer, leaving only those bound to the membrane-bound proteins. DNAse I is then used to digest weakly bound aptamers, leaving behind only strongly bound aptamers. Washes with gradient urea buffer are also used to improve the removal of weakly

*bound aptamers. The remaining aptamers are amplified and sequenced. Reproduced with permission from reference **258**. Copyright 2012 American Chemical Society.*

Finally, a single cycle aptamer selection was conducted using the magnetic separation of a bead-based library and protein targets attached to magnetic beads using biotin-streptavidin or protein-antibody interactions.[260] In this case, the library contains modified nucleotides that are either limited to specific bases in pre-determined positions (**Figure 1.23**)[261] or encoded by barcodes contained in the forward primers, the latter necessitating the use of additional primers during amplification. This methodology allows for an expanded genetic code to be used and does not require the regeneration of the functional library for future selection rounds, it requires only that the library be converted to DNA during amplification in such a way that sequencing can reveal the site and identity of modifications.

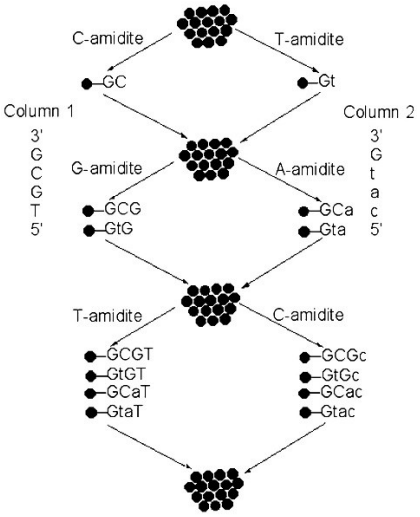


*Figure 1.23. Above is a diagram demonstrating the design of libraries with modified nucleotides in pre-determined positions.[261] In this case, bead-bound aptamer libraries are designed to include both normal nucleotides and nucleotides with phosphorothioated (PS) backbones. In the schematic, two columns are shown for a split-and-pool combinatorial synthesis. In column 1, nucleotides are coupled via a normal phosphodiester (PO) linkage (and are represented as uppercase letters). In column 2, nucleotides are coupled via a PS linkage (and are represented as lowercase letters). Each position of the aptamer has predefined bases unique to each column, such that the sequence can be unambiguously determined after selection. For example, if the sequence GTAT (given as 3' to 5' based on SPS direction of synthesis) were identified in sequencing, the linkages between G and T and between T and A must be PS, as the 1st T and the A must have been made on column 2, and the linkage between A and T must be PO. Reproduced with permission from reference **261**. Copyright 2002 Oxford University Press.*

There are still some important considerations when using single round selections to identify enriched aptamers, as has been recently outlined in great detail by the Krylov research group.[262] The most important part of aptamer (and other ligand) selection experiments is arguably the partitioning of non-binders from binders, yet there are no universal quantitative guidelines for

single-round selection experiments. Mathematically, the output binder to non-binder ratio is dependent upon the input binder to nonbinder ratio, as well as the ratio of the fractions of binders and of non-binders that survive partitioning. While these parameters are hard to measure in randomly generated aptamer libraries, the principles governing their relationships can still be used to perform better single-round selections. They argue that some techniques like deep sequencing or fractionation, in which multiple fractions are taken from the partitioning output, such as in the case of physical column segmentation, do not improve any of these parameters as claimed, though fractionation may be used to identify samples that feature better ratios of binders to non-binders.

In summary, single-round SELEX can be performed by introducing additional stringencies or employing higher level bioinformatic analysis techniques. Though it has not frequently been used for modified aptamers, it does represent a promising avenue to performing selections with libraries that cannot be amplified for future rounds, requiring only an enzyme capable of amplifying modified aptamers into unmodified oligonucleotides. Barcoding represents one possible way to introduce multiple different types of modifications, though its use to date in aptamer selections is very limited.

1.6. DNA-encoded libraries

In the pharmaceutical field, drug discovery has classically occurred via screening methods in which drug candidates are synthesized and characterized individually to determine which structures and sequences are best suited to a therapeutic or diagnostic application. However, this method is time-consuming, costly, and not well-suited to large libraries. It takes multiple synthetic steps to generate leads and optimize a potential drug, even when compounds are rationally designed based on prior knowledge of shape or structure. With polypeptides, it is possible to discover the sequences of isolated molecules using chemical or enzymatic digestion paired with mass spectrometry, allowing for selections with larger libraries, though sequencing peptides is nowhere near as simple as sequencing oligonucleotides. One method that uses DNA sequencing is antibody phage display (see **Figure 1.24**), which links an oligopeptide on the surface of a bacteriophage with the corresponding DNA sequence contained in the phage's genome.[263]

This link between genotype and phenotype allows for the synthesis of a large library of bacteriophages that can be selected for based on the binding affinities of the presented oligomers, though building blocks are limited to natural amino acids as the display technology requires transcription and translation. This means that drug discovery for libraries not composed entirely of nucleotides or of amino acids are limited to screening as there is no way to identify hits following a library selection.

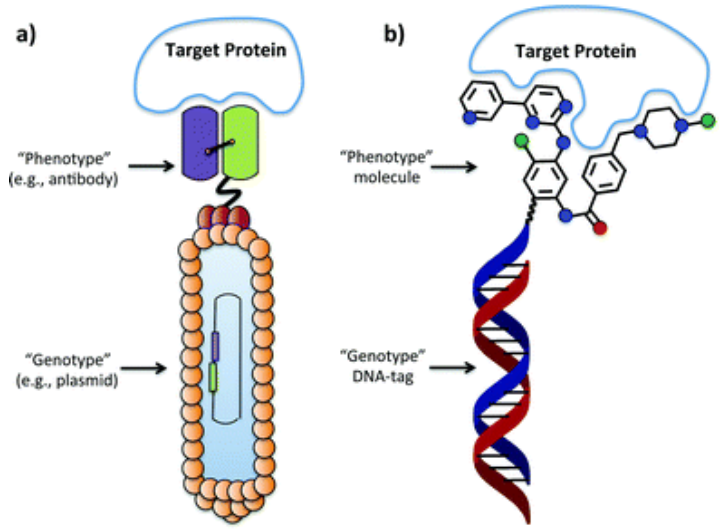


Figure 1.24. Two ways of encoding a phenotype with an oligonucleotide-based genotype are shown above.[264] In a), phage display links a surface-presented peptide to its untranslated sequence contained in the plasmid of a phage. In b), DNA-encoded libraries, a DNA-tag unique and specific to a synthetic polymer is covalently linked to the same polymer. Both methods allow for the easy identification of functional polypeptides in selection experiments using DNA barcodes. Reproduced with permission from reference ***264.*** *Copyright 2011 The Royal Society of Chemistry.*

In 1992, Brenner and Lerner devised an encoded combinatorial library, later referred to as a DNA-encoded library (DEL), that has since revolutionized small molecule drug discovery. They covalently linked a custom DNA tag with a synthetic peptide in order to perform selections of DNA-encoded libraries that were amenable to amplification and sequencing by virtue of a DNA barcode.[265] This approach relied on a split-and-pool combinatorial approach, in which parallel syntheses are used with intervening pooling and splitting steps to create a diverse library in fewer synthetic steps. A schematic for split-and-pool synthesis is shown in **Figure 1.5**, which illustrates its advantages in the synthesis of libraries with multiple positions featuring different possible monomers. The method was termed retrogenetic, as it couples the addition of a monomer with the subsequent addition of a codon unique to that monomer, followed by pooling and splitting

steps. The DNA code was designed with triplet codons chosen carefully to ensure no codon sequences could be misread in the case of reading frame errors. The primers were chosen so that they would not occur within the DNA barcode itself and restriction enzyme sites were incorporated to allow for biotin incorporation and strand separation. The coding region was devised to include at least 15 nucleotides for proper hybridization, which is a condition met for libraries with at least five monomers.

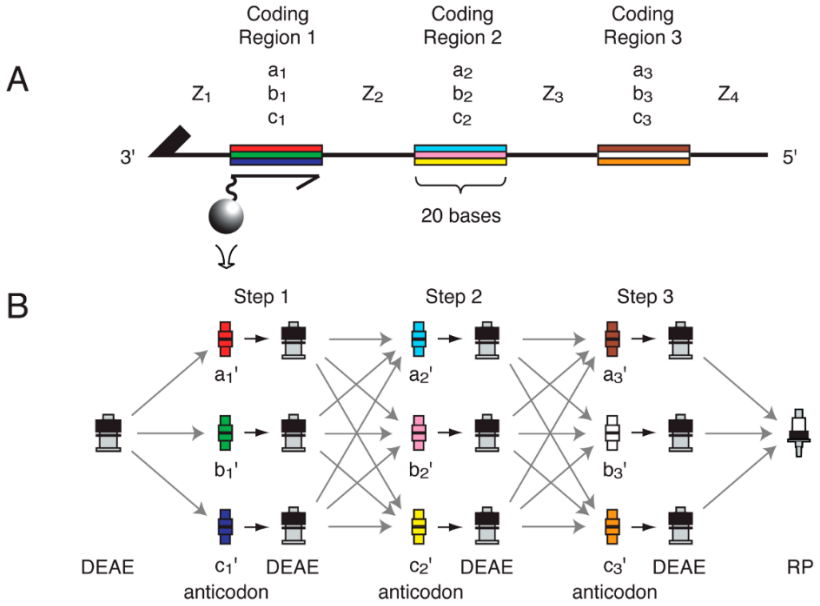


*Figure 1.25. Sequence-encoded routing of DNA is shown above for the synthesis of combinatorial libraries of aptamers.[266] In (A), the design of the library is shown, with alternating 20-base noncoding regions (Z1-Z4) and 20-base coding regions. The coding regions at each position can be one of three unique codons (a,b,c). Using columns with the anticodons coated on resin beads, it is possible to direct the path of an individual oligonucleotide based on its sequence (as shown in (B)), as hybridization between codon and anticodon can effectively partition sequences with different codons at a particular position. This routing approach allows a specific monomer of a growing synthetic polymer to be added to each sequence following partitioning. This eliminates any need to construct the DNA barcode and the synthetic polymer at the same time, as only a library of barcodes needs to be synthesized combinatorially. DEAE stands for diethylaminoethyl, which refers to a type of Sepharose column. RP refers to a reverse phase column. Reproduced with permission from reference **266**. Copyright 2004 Halpin and Harbury.*

Further work in DELs have allowed for the synthesis of a polymer given a fully constructed genetic code. One method uses DNA hybridization to direct the separation of growing polymer strands during split-and-pool combinatorial chemistry (**Figure 1.25**).[266] The programmable nature of DNA allows selective hybridization to partition populations of growing polymers based on their DNA sequence using anticodon capture columns. This routing approach requires individual codons to be long enough for efficient hybridization and separated by non-coding regions, which can result in very long coding sequences, even for polymers with relatively few monomeric units. This

method uses a ssDNA code with an appended amine-functionalized PEG linker and relies on Fmoc (9-fluorenylmethoxycarbonyl)-based peptide synthesis. Proximity-based DNA-templated synthesis has also been used by associating building blocks with unique anticodons so that a complete DNA code can template the assembly of a corresponding synthetic polymer.[267] This method does limit reactions to conditions favourable for DNA hybridization and can also limit the number of units that can be successively encoded.[268] Another DNA-templated method creates a macrocyclic library using amide chemistry, with the DNA code containing three consecutive codons that can bind to anticodons covalently linked to building blocks (**Figure 1.26**).[269] The method is limited to four-membered rings of the same shape and can only include functionalities tolerated by the frameworks of the three types of building blocks. dsDNA has also been used for DEL synthesis without relying on DNA-templated synthesis, with triazine-based libraries created combinatorially using the enzymatic ligation of codons and step-dependent chemical reactions to create libraries with three or four different monomeric units consisting of amino acids, amines, and acids.[270] dsDNA was chosen over ssDNA because of its higher stability and resistance to reagents during the synthesis of the library, as the double helix can protect the nitrogen-containing bases of DNA. Two-base 3' overhangs were used to ligate 7-base double-stranded codons, with overhang sequences unique to each step of the synthesis. This DEL limits the type of building blocks that can be included at each position, but positions of the oligomer feature up to 384 different options, so libraries may contain up to 10^8 different sequences. To date, DELs have been sparingly used, if at all, in the field of aptamer selection and they represent an interesting avenue for the selection of modified aptamers.

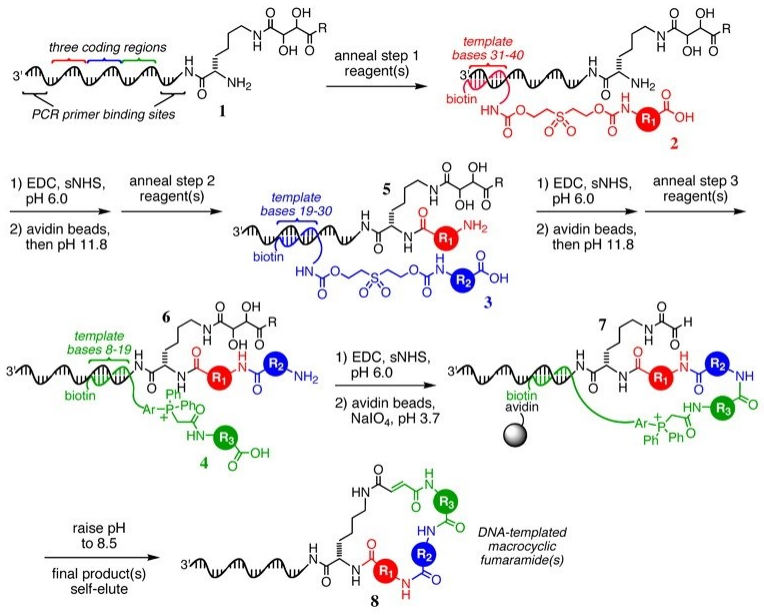


*Figure 1.26. The construction of a DNA-templated macrocyclic library is shown above.[269] In this method, a library of oligonucleotides consists of three consecutive codons flanked by primer regions and features a 5' functionalized end used for the synthesis of a macrocycle. The functional library is synthesized stepwise since the macrocycle is grown in a specific order with different families of building blocks (as shown by the green, red, and blue-coloured monomers). The building blocks, covalently linked to the anticodons and tagged with biotin, are added to the library of oligonucleotides by alternating steps of annealing, coupling, and magnetic purification. The final product is shown in the last step, with three different functional groups included in a macrocyclic framework. Reproduced with permission from reference **269**. Copyright 2004 The American Association for the Advancement of Science.*

1.7. Previous work

In general, nucleotide modifications have been very important to the field of aptamers, ASOs, and siRNA, with noted improvements in stability, circulation times, and binding affinities gained from their incorporation in oligonucleotides. Dr. Donatien de Rochambeau, a previous student in the Sleiman lab, designed several versatile platforms for creating DMT-protected (4,4'-dimethoxytrityl) modified phosphoramidites useful in various DNA nanotechnology applications, particularly for oligonucleotide therapeutics.[271] The synthetic platforms rely on amide or click chemistry to append functional groups containing the appropriate reactive moiety (see **Figure 1.27**). The platforms are also designed to be non-nucleosidic, featuring a nitrogen-containing acyclic backbone and lacking nitrogenous bases. These modifications can be incorporated into nucleic acids via solid-phase synthesis with phosphoramidites but they cannot be used in normal SELEX as they are not compatible with DNA polymerases involved in amplification or sequencing. More importantly, as the monomers cannot base pair with normal DNA, they cannot be

transcribed from DNA or reverse transcribed into DNA using engineered enzymes as in the case of some examples listed earlier. As a result, these modifications can only be used in screening applications or in post-SELEX strategies where neither sequencing nor amplification is required.

However, with the use of DELs, these modifications can be incorporated into aptamer libraries in a sequence-defined manner and used in competitive selections for target binding. A DEL construct was synthesized, with its intended use for the selection of modified aptamers.[272] DELs can be designed in multiple ways; the DNA code and polymer can be grown orthogonally and simultaneously, the polymer can be grown from a fully synthesized code, or compartmentalized polymers can be appended with unique DNA codes. The first method was chosen for the synthesis of DNA-encoded modified aptamer libraries as it was best suited to solid-phase DNA synthesis, which can tolerate non-nucleosidic phosphoramidites (see **Figure 1.29** for a depiction of the aptamer-code construct). Solid-phase synthesis relies on DMT chemistry to grow ssDNA strands in a sequence-defined manner in the 3′ to 5′ direction, as shown in **Figure 1.28**.

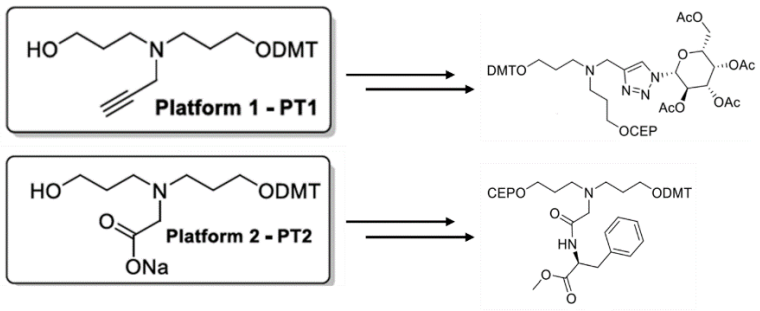


Figure 1.27. The two synthetic platforms on the left can be coupled with moieties containing azides or primary amines.[272] Appropriately protected molecules can be reliably coupled to the platforms and turned into DMT-protected phosphoramidites in relatively few steps. On the right are shown two examples of non-nucleosidic modifications that can be synthesized using these platforms. CEP = β-cyanoethyl-N,N′-diisopropylamino-phosphoramidite, DMT = 4,4′-dimethoxytrityl.

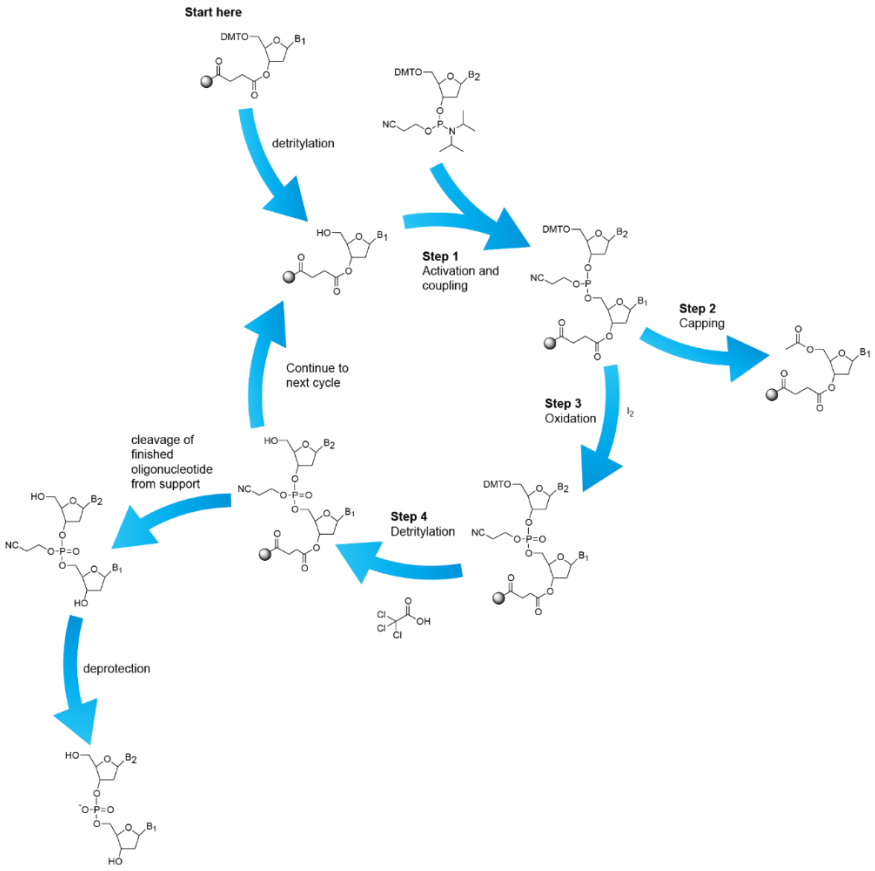


Figure 1.28. Solid-phase synthesis (SPS) is shown for the construction of ssDNA using DMT and phosphoramidite chemistry.[273] *Phosphoramidites with 3'-OH groups protected by DMT are coupled to a growing chain with an available 5'-OH group using a tetrazole-based activator. Any support bound chains with free hydroxyl groups are capped using an acylating agent to prevent the occurrence of deletion mutations and maintain a high degree of purity. Oxidation with iodine and water convert the phosphorus of the backbone from P(III) to P(V). The newly coupled monomer is then deprotected at the 5' end with trichloroacetic acid and cycles can be repeated as the strand is synthesized. Finally, the DNA is cleaved from the solid support and the backbone and bases are deprotected with aqueous ammonia. Reproduced with permission from reference* ***273****. Copyright 2011 ATDBio.*

CPG (controlled pore glass) supports with a cleavable linker and available hydroxyl groups serve as the starting point for the activated coupling of a protected nucleosidic phosphoramidite (**Figure 1.28**, Step 1), followed by the capping of unreacted and unprotected OH groups (Step 2), the oxidation of the newly coupled monomer (Step 3), and the deprotection of the 5' OH for future couplings (Step 4). However, for a DNA code and modified aptamer to be synthesized combinatorially and concurrently using solid-phase synthesis, there are two glaring issues; the chemistries of the coding strand and aptamer strand need to be orthogonal, and the aptamer-code construct needs to be covalently linked using a branching unit.

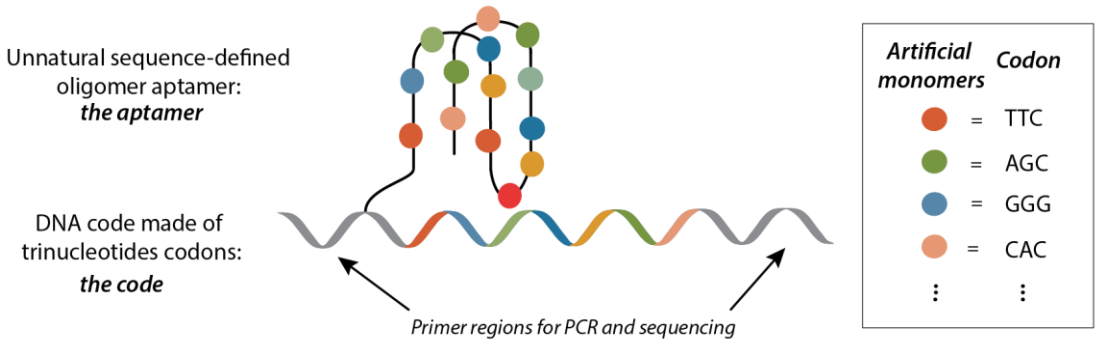


*Figure 1.29. The generalized aptamer-construct is shown above, with a DNA code consisting of codons corresponding to unnatural monomers in the modified aptamer. In grey are the primer regions necessary for amplification and sequencing.[272] The code is covalently linked to the aptamer via a synthetic branching unit. Reproduced with permission from reference **272**. Copyright 2019 de Rochambeau.*

To solve the first issue, levulinyl was used as a protecting group for the DNA code while DMT was used as the protecting group for the modified aptamer. Levulinyl is deprotected by hydrazine treatment and DMT is deprotected by acetic acid treatment. A split-and-pool approach thus allows a monomer to first be added to the growing aptamer using DMT chemistry and then the corresponding DNA codon to be added to the growing DNA barcode using levulinyl chemistry. The second issue arises from the necessity of growing both strands together, which requires two ends to be growing at the same time given one starting point. A branching unit was designed using a synthesized modified uracil phosphoramidite containing a levulinyl-protected 5' OH group and a DMT-protected OH group appended to the C5 of uracil. This modified uracil was placed in the reverse primer, PCR amplification with Taq polymerase was optimized for the aptamer code construct, and sequencing revealed an error rate comparable with that expected in solid-phase synthesis.[272] This set the stage for the selection of DNA-encoded modified thrombin-binding aptamers. An overview of the synthetic steps required for the addition of a monomer to an aptamer and the addition of the corresponding codon to the DNA code is shown in **Figure 1.30**.

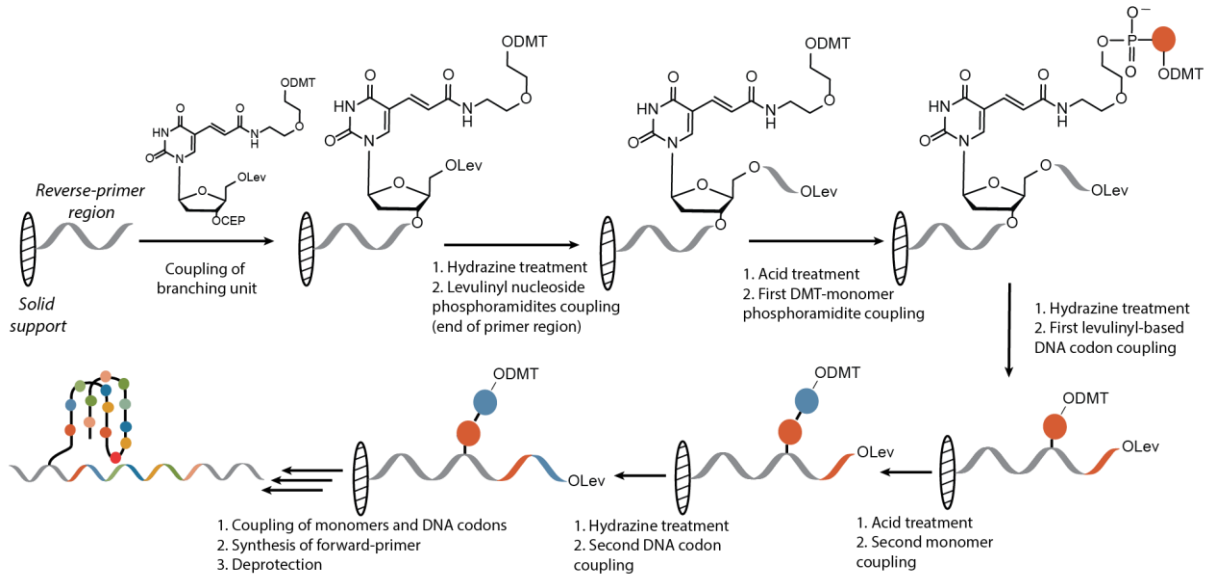


Figure 1.30. The overall scheme for the synthesis of DNA-encoded aptamers is shown above.[272] The branching unit, shown in the very first step, is coupled to and included in the partially complete reverse-primer region grown from a solid support. The branching unit is suitably protected at the hydroxyl positions of C5′ in the 2′-deoxyribose ring, which will be used for the growth of the aptamer using DMT chemistry, and of C5 in the modified uracil, which will be used for the growth of the DNA code using levulinyl chemistry. The reverse primer is completed after the coupling of the branching unit. A monomer is added to the aptamer chain with DMT chemistry and the associated 3-base codon is added to the barcode with levulinyl chemistry. Monomers and their codons are added until the final structure is achieved, with a forward primer region added at the end. Though it is not shown here, in the synthesis of combinatorial libraries, pooling and splitting would occur after each aptamer-codon addition, with several different aptamer-codon additions occurring in parallel. Reproduced with permission from reference ***272****. Copyright 2019 de Rochambeau.*

2. Results and Discussion

2.1. Acknowledgement of authors' contributions

Dr. Donatien de Rochambeau designed the DNA-encoded library as well as the synthetic platforms for modified phosphoramidites. Dr. Hanadi Sleiman, as well as Dr. Maureen McKeague, helped design the project, provided valuable insight, and helped analyze data. Serhii Hirka performed the selection and binding affinity experiments. Daniel Saliba performed the PCR amplification and serum stability experiments. Shaun Anderson analyzed the sequencing data, synthesized aptamers for individual study, and synthesized compounds relevant to the library and its building blocks.

2.2. Introduction

Antibodies have long been the gold standard in targeting and diagnostic applications as they bind to antigens with strong affinity and specificity. However, the selection of antibodies for new targets is limited to those that elicit an immune response, and the production of monoclonal antibodies is costly, time-consuming, and prone to contamination. Aptamers are typically synthesized by solid-phase synthesis and represent a cheaper, faster, and more versatile avenue for discovering ligands. What has most limited their use is not their production but their physicochemical properties. Many modifications to date have helped improve their nuclease stability and *in vivo* circulation times. The chemical space explored by aptamers is currently limited by their method of selection, as SELEX involves the amplification and sequencing of oligonucleotides. Minor modifications have been incorporated into aptamer libraries for selection by SELEX using engineered polymerases or DNA-templated sequencing, though non-nucleosidic modifications are rarely tolerated. Using an approach started in our lab inspired by the DNA-encoded libraries of Stenner and Lerner, these modifications can be included in combinatorial libraries of aptamers using a barcoded approach, wherein a synthetic polymer is distinctly coded for by a series of codons each corresponding to a unique monomer. A DEL of modified thrombin-binding aptamers was rationally designed, the library was incubated in the presence of thrombin to select for binders, the library was amplified and sequenced to discover

sequences and beneficial modifications, and the promising aptamers were individually synthesized and characterized to determine the effect of the non-nucleosidic modifications. The focus of this thesis is on the last stages of sequencing and aptamer characterization, as the initial library design and selection was performed previously by other lab members.

2.2.1. Rational design of modified library

SELEX experiments require several cycles to achieve sufficient enrichment to isolate aptamers with promising sequences or subsequences from huge libraries with as many as 10^{15} unique sequences.[27] However, for multiple cycles of enrichment to occur, SELEX uses PCR amplification to generate new libraries for subsequent selection rounds. In the case of non-nucleosidic modifications, it is very difficult, if not impossible, to amplify and generate a functional library. One notable successful example is that of AEGIS-SELEX, which stands for artificial expanded genetic information systems SELEX.[243] This variant introduces an orthogonal base pair that can be amplified by PCR at the same time as normal DNA bases, allowing six different monomers to be included in libraries compatible with SELEX. Additional synthetic conversion steps are required to elucidate the sequences of the structure as the orthogonal base pair is not compatible with high-throughput sequencing. Including non-nucleosidic modifications without relying on DNA base pairing interactions presents key problems for amplification and sequencing. DNA-encoded libraries represent a promising path for the selection of aptamers with non-nucleosidic modifications. To date, they have not been used in aptamer selections, with the exception of some methods that enzymatically ligate DNA barcodes in non-combinatorial syntheses. Another method combinatorially synthesizes a library containing both normal nucleotides and phosphorothioated nucleotides to achieve a larger genetic code, though the modifications are added in a pre-determined way for a specific nucleotide in a specific position so that the sequence can be read and the location of phsophorothioated nucleotides will be known.[261] This method is limited in its application to non-nucleosidic modifications and induces additional constraints on the diversity of the library. Using our method, DNA-encoded libraries of modified aptamers can be generated with no constraints on the types of modifications or the positions where they can be introduced, so long as they are compatible with the conditions of oligonucleotide solid-phase

synthesis. One key restraint of this method is that the library cannot be regenerated after amplification, and so it cannot be used for multiple cycles of SELEX. However, there have been several examples as listed in the introduction that demonstrate the potential for single-round SELEX or non-SELEX, which perform only one round of amplification but rely on multiple partitioning events or a more stringent selection protocol. As this selection was performed with classic agarose-immobilized targets, a library of modified thrombin-binding aptamers was rationally designed to limit the size of the library to enable sufficient enrichment in one single round of SELEX.

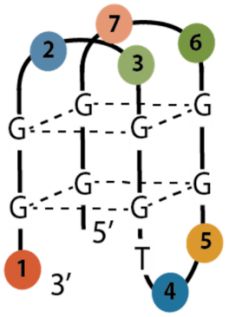


*Figure 2.1. The thrombin-binding aptamer is shown with seven nucleosidic positions highlighted as the focus of modifications for a DNA-encoded combinatorial library of modified TBAs. Reproduced with permission from reference **272**. Copyright 2019 de Rochambeau.*

The thrombin-binding aptamer is the most widely-studied and well-known aptamer, and with its convenient 15-base size it represents a good candidate for modification with non-nucleosidic groups. Using carboxyl, alkyne, and diol platforms, the modified phosphoramidites shown in **Figure 2.2** were synthesized for use in the modified library. Phosphoramidites were designed with hydrophobic groups such as naphthalene (Nap), anthracene (Ant), and alkyl chains in branched (Bal) and linear (C12) forms. Amino acid-based monomers were also incorporated, with phenylalanine (Phe), histidine (His), and tryptophan (Trp) present as functional moieties. Carbohydrate (Sug) and alkyne (Alk) phosphoramidites were also included. The 15mer aptamer was not modified at every position and was instead modified at most positions outside the GG quartets integral to the structure of the G-quadruplex of the TBA (as shown in **Figure 2.1**). The sequence of the unmodified TBA is 5′ GGTTGGTGTGGTTGG 3′, containing two TT loops thought to bind the fibrinogen-binding exosite of thrombin and a TGT loop thought to be involved in secondary interactions with the heparin-binding exosite of thrombin. For improved binding

affinities, the TT loops are of particular interest as they are expected to be involved in the important binding interactions with thrombin. An additional nucleotide was added to the 3' end as there were literature reports of added stability upon extension with a guanine or thymine nucleotide.[19] The guanine quadruplex was left intact as most modifications to this region have been detrimental to the stability and binding affinity of the aptamer, with very few modifications that stabilize the quartet proving to be beneficial, such as those that favor a particular glycosidic bond conformation.

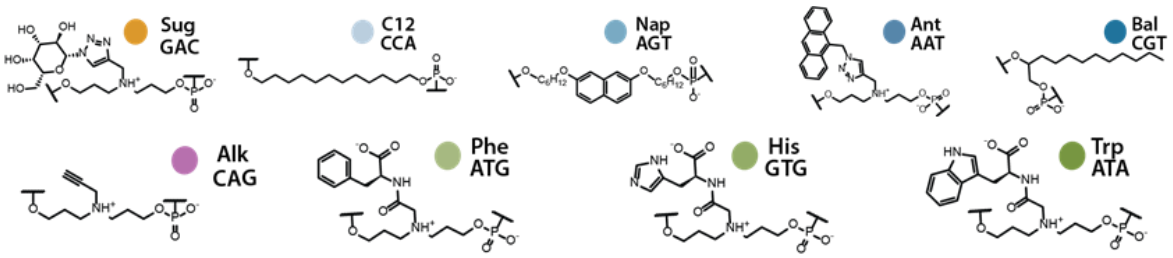


*Figure 2.2. Non-nucleosidic modifications used in the generation of a modified thrombin-binding aptamer library.[272] In this case, unique codons are shown (in the 3' to 5' direction of SPS) that correspond to each monomer. The abbreviations, in order from left to right and then top to bottom, are Sug=Sugar, C12=carbon-12, Nap=naphthalene, Ant=anthracene, Bal=branched alkyl, Alk=alkyne, Phe=phenylalanine, His=histidine, and Trp=tryptophan. Reproduced with permission from reference **272**. Copyright 2019 de Rochambeau.*

To generate the modified aptamer library, Dr. Donatien de Rochambeau performed split-and-pool steps following the addition of a monomer at one of the pre-determined positions, with six possible monomers for each position, giving a maximum library size of 6^7, around 2.8×10^6 unique sequences. The choice of six different monomers per position was practical as the automated DNA synthesizer used to construct the library featured six ports. The monomers used at each position are summarized in **Table 2.1**. Bal (the branched alkyl) was not included in the library over concerns of chirality as it was present as a racemic mixture. Anthracene was limited to two positions of the aptamer to prevent higher order assemblies. Most of the remaining non-nucleosidic modifications present in the library were included in at least four of the seven positions being modified. Monomer additions on the aptamer branch were coupled with 3-base codon additions to the covalently linked DNA code, as shown in **Figure 1.30**. A 7T (thymidine) spacer was inserted between the aptamer and the branching unit to limit steric interference from the DNA barcode during selection experiments.

Table 2.1. The monomers used in the generation of a DNA-encoded modified TBA library are depicted below.[272] The unmodified TBA is contained in the library based on the monomers used in Aliquot B. Sug=Sugar, C12=carbon-12, Nap=naphthalene, Ant=anthracene, Alk=alkyne, Phe=phenylalanine, His=histidine, and Trp=tryptophan.

Position	Aliquot A	Aliquot B	Aliquot C	Aliquot D	Aliquot E	Aliquot F
1	G	T	Alk	C12	Nap	Phe
2	A	T	Alk	C12	Sug	Trp
3	C	T	Ant	Nap	Phe	Trp
4	A	G	Alk	C12	Nap	Sug
5	G	T	C12	Phe	Sug	Trp
6	A	T	Alk	His	Nap	Phe
7	G	T	Ant	C12	Sug	Trp

2.2.2. Selection of modified thrombin-binding aptamers

A negative selection step with agarose devoid of thrombin was used to remove non-specific binders first, and then the library was incubated with thrombin immobilized in agarose. The column was washed multiple times to remove non-binding aptamers prior to the elution of aptamer candidates. Elution methods varied slightly in several different rounds, though distinct samples were always taken after each elution and amplified for sequencing. An analogous fluorescently tagged library was used once for comparison and to facilitate the monitoring of the selection progress. Elutions occurred with competitive agents, such as the unmodified TBA (lacking primers, branching unit, and barcode) and fibrinogen, in order to displace bound aptamers from the immobilized thrombin. Free thrombin solution and hot water were also used to pull down aptamers, to varying levels of success. All selection steps were performed by Serhii Hirka. Sequencing results were then used to determine which sequences were of interest and guide future single-round selections. These results were then used to determine aptamer sequences of interest, which were synthesized and characterized to determine the effect of the non-nucleosidic modifications. **Figure 2.3** depicts the workflow for selecting modified aptamers.

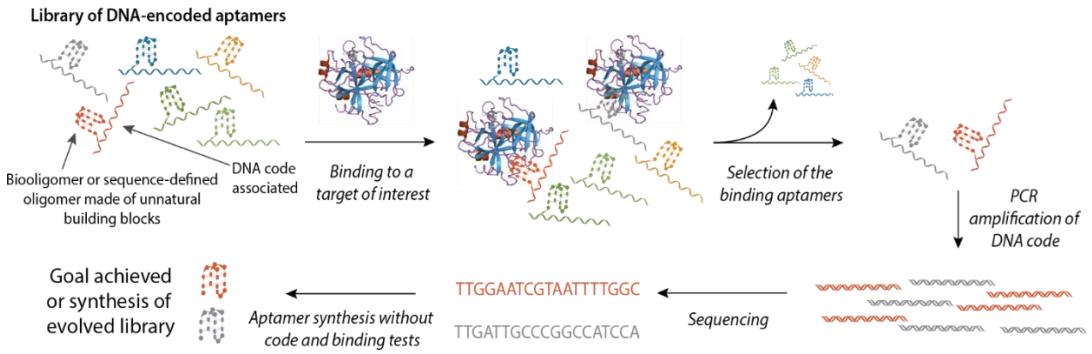


*Figure 2.3. The single-round selection process for DNA-encoded modified aptamers is shown above.[272] The DEL is incubated with the target, allowed to bind, and then binding aptamers are differentiated from non-binding aptamers with washes and elutions. The selected aptamers are then amplified by PCR, reducing the aptamer-code construct to just the code. These unmodified oligonucleotides are sequenced to reveal possible strong-binding aptamers. These candidates are then synthesized independently for characterization. Reproduced with permission from reference **272**. Copyright 2019 de Rochambeau.*

2.3. Sequencing analysis

2.3.1. Interpretation of sequencing data

During the two main selection experiments performed, samples were isolated and amplified for sequencing after each distinct elution. Library samples were also sequenced in order to compare the enrichment of sequences compared to the initial distribution. Samples were sequenced by MiSeq Illumina next-generation sequencing[274] at the Genome Quebec Innovation Centre, in Montreal, Quebec, Canada. An additional step is required prior to sending samples for sequencing, as primers specific to the sequencing methods are appended to the DNA barcodes, which already contain primers from the initial library design. Sequencing results were obtained as fastq files, with two read files per sample corresponding to the forward and reverse strands. Fastq files are text files that list the following information for each uniquely indexed sequence read: a unique identifier, a complete DNA sequence, and quality scores provided for each base of the sequence (**Figure 4.1**). Each file contained anywhere from 30,000 to 250,000 individual reads, depending on the goal of the sequencing. It is possible to use clustering algorithms or high-end analysis techniques, with several commercial products available, though simpler methods were used to identify enriched sequences and beneficial modifications. Cutadapt is a freely available program that can trim adapter sequences from high-throughput sequencing reads and was used to trim the full sequences to just the 21mer DNA barcode (**Figure 4.2**).[275] To determine the most popular sequences, the count function of FASTAptamer, a program used for analyzing

large sequencing files for libraries of oligonucleotides,[86] was used for the trimmed sequences containing the entire DNA barcode (**Figure 4.3**).

2.3.1.1. Validation of library design

The library construct was validated in prior work[272] by counting the number of times the desired sequence was found out of the total number of reads. For these samples, total reads ranged from 30,000 to 60,000 per sample and 93-95 % of the sequences featured the expected primer sequence, which was either the forward primer or the reverse primer, depending on the read file. This translated to an average error rate per base near 0.5 %, which is comparable with the error rate of solid-phase synthesis, in which around 99 % coupling efficiency is expected per base.[276] The code regions were more error-prone, with only around 70 % of the expected sequences found in the samples, which translated to an error rate per base ranging from 0.5 to 1.7 %. The library design was deemed appropriate for amplification and sequencing, as error rates were still low in the case of highly modified aptamers. Hot start Taq polymerase was the enzyme used for amplification and in the future the use of alternative or engineered polymerases could improve the error rate for aptamer-code constructs.

2.3.1.2. Popular sequences

The goal of SELEX and other selection experiments is to find which members of a library are best suited to binding a target under specific conditions. Multiple rounds of selection in SELEX yield enriched aptamers that can be discovered based on the abundance of sequences from elutions. Abundance can be assessed within a sample, with the abundance determined relative to the other sequences in a sample, or it can be determined by comparison with the initial unenriched library to determine what level of enrichment has occurred for a specific sequence over the course of the selection. In our case, we used a single-cycle selection and thus focused on the most abundant aptamers in each elution. The assumption was that later elutions would yield aptamers with higher binding affinities, as aptamers still bound to the target after multiple washes are likely to have lower dissociation constants than those that are no longer bound to the target. The sequences in the fastq files containing the expected primer regions were trimmed with Cutadapt to isolate the DNA barcode. Code regions that were not 21 bases long were discarded, as they did not contain exactly seven 3-base codons. FASTAptamer count was used to

arrange the code sequences by abundance and frequency, with frequency being a metric that could be compared across different samples. The code sequences were deciphered using the established genetic code to determine the sequences of the abundant samples.

2.3.1.3. Positional tendencies

The other type of analysis used for sequencing data was to compare the prevalence of modifications at a given position for a specific eluted sample. For each position of the aptamer that was modified, there were six possible monomers. 16.67 % would be expected in an unbiased library, so positions at which monomers were present at a much higher proportion were noted. This was performed using a simple python program we designed that trimmed the barcodes to determine how often a particular codon appeared in a certain position. For a given position, the sum of the six codons never quite equaled 100 %, with about 3 % of codons at each position not corresponding to any of the possible monomers. Similarly, about 3 % of 21-base sequences featured at least one erroneous codon. This was likely due to errors in synthesis or replication inherent to solid-phase synthesis and PCR amplification, respectively. By the same approximation used in previous work,[272] this translates to an error rate per base near 1 %, which is in agreement with previous results. This type of analysis can provide information on individual modifications that are beneficial in general to the aptamer, while the sequence popularities can identify multiple modifications that are mutually beneficial.

2.3.2. Setup 1

In the first selection setup (**Figure 2.4**), the elution steps in order were as follows: one elution with free thrombin, one elution with the regular TBA, one elution with fibrinogen, and one elution with hot water. The regular TBA is expected to competitively bind to the immobilized thrombin and displace modified aptamers, particularly those that are less strongly bound. Fibrinogen is expected to act in a similar manner and displace aptamers from the fibrinogen-binding site, where the TBA and its modified derivatives are expected to bind. Free thrombin is expected to bind to the aptamers and remove them from the column. Hot water can denature or deform the structures of bound aptamers and dislodge them.

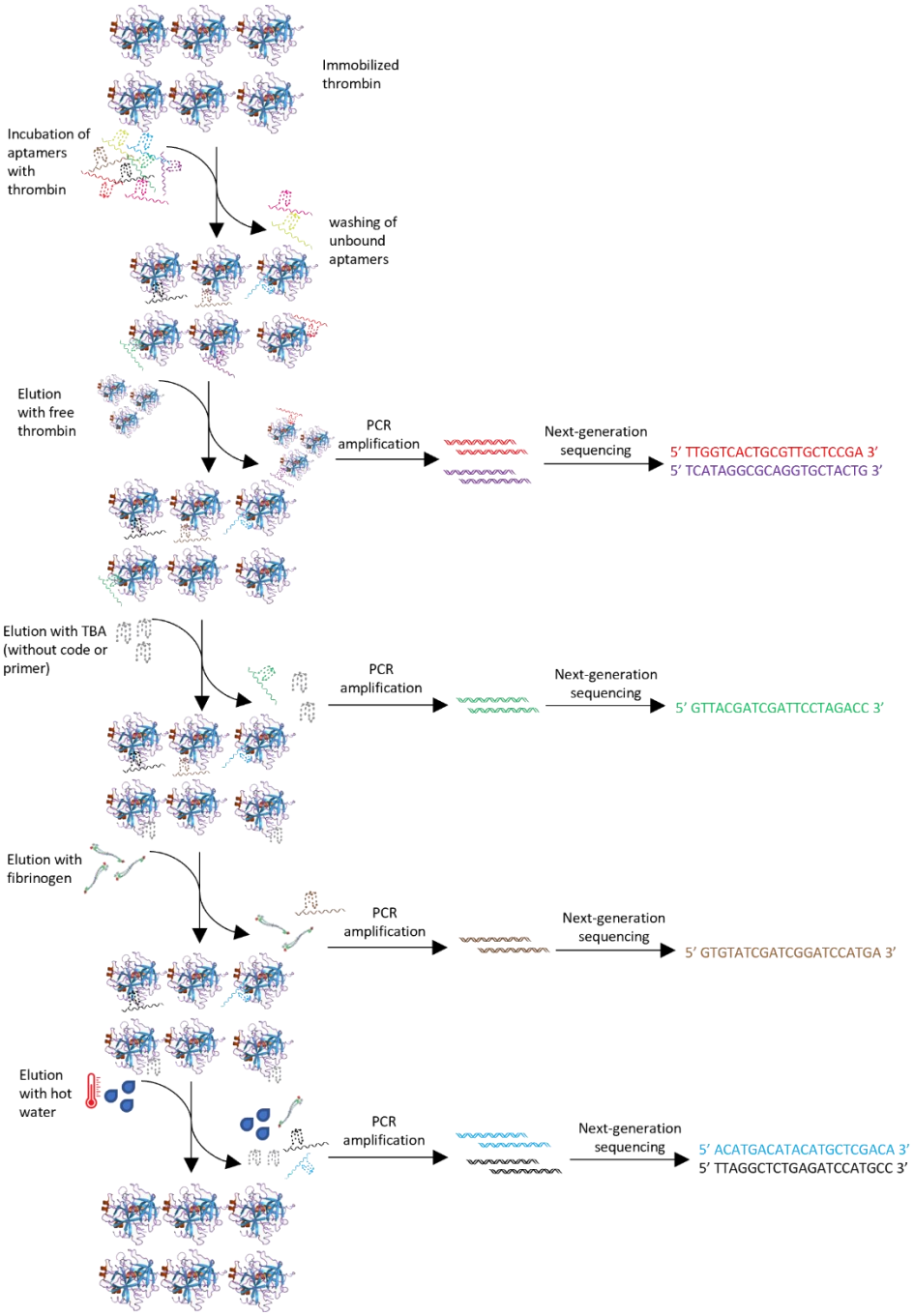


Figure 2.4. An overview of the first selection setup is presented above (the counter selection step is omitted in the schematic). The library of modified DNA-encoded aptamers was incubated with immobilized thrombin and unbound aptamers were removed by washing steps. An elution with free thrombin was performed next, with any eluted aptamers amplified by PCR and sequenced. Next, an elution with the unmodified thrombin-binding aptamer (TBA), lacking any primers or code necessary for amplification, was performed and any recovered sequences that were displaced were amplified and sequenced. Fibrinogen was used in the next elution and any displaced aptamers were amplified and sequenced. Finally, an elution with hot water was performed to recover any sequences still bound to the column, which were amplified and sequenced. In both cases, a sample from the initial library was also amplified and sequenced. A sample from the initial library was also amplified and sequenced.

2.3.2.1. Popular sequences

In the first elution, the unmodified aptamer, which was a sequence present in the library, was the third most popular sequence. In the second elution, the unmodified sequence was the most popular, though other sequences appeared at similar frequencies. In the third elution with fibrinogen, the unmodified TBA was by far the most popular sequence, with other sequences only present at a fraction of that at which the unmodified aptamer was present. As a result, this sample was not very helpful for identifying enriched aptamers. In the fourth elution with hot water, the unmodified aptamer was the tenth most popular sequence, with several other sequences appearing earlier. In the following tables, the condensed sequence is shown in place of the full sequence. The full sequence is 5' GG **76** GG **54**T GG **32** GG **1** 3', which the bolded numbers corresponding to the condensed sequence, 5' **7 6 5 4 3 2 1** 3'.

Table 2.2. The 10 most popular sequences from a sample of the library prior to any selection are presented below. The unmodified TBA has a condensed sequence of 5' T T T G T T T 3', which was present at a frequency of 148 reads per million (not shown in the table).

Rank	**Condensed sequence (5' to 3')**							**Reads per million**
	7	**6**	**5**	**4**	**3**	**2**	**1**	
1	T	T	T	G	Phe	C12	Nap	494
2	Trp	A	T	G	Nap	C12	Phe	432
3	Sug	T	Phe	G	Nap	C12	Nap	420
4	G	T	C12	C12	Phe	C12	G	407
5	C12	T	T	C12	Nap	T	G	395
6	G	A	C12	A	Phe	C12	Nap	383
7	T	A	Phe	G	Nap	C12	T	370
8	C12	A	G	C12	Phe	T	Nap	358
9	Sug	Phe	Phe	G	Nap	C12	G	346
10	Trp	A	Phe	A	Nap	C12	T	346

The most popular sequences from the library sample, which was not incubated with thrombin, are shown in **Table 2.2**. This sample featured approximately 81,000 reads, with the most popular sequence, 5' T T T G Phe C12 Nap 3', appearing 40 times at a frequency of 494 reads per million. A perfectly unbiased library, made up of 6^7 (279,936) different members, would be expected to have all sequences present at the same level. One out of 279,936 equates to 3.57 reads per

million. Clearly, this level of distribution was not achieved by the library. However, a sequence occurring only once in 81,000 reads would translate to 12.3 reads per million, and for this sample, approximately 44,000 unique sequences were read with 30,000 of these sequences occurring only once. While this size of sampling cannot identify all members of the library due to its sheer size, for the purposes of detecting enriched aptamers in future rounds of selection, this method should provide sufficient detail to isolate aptamers that are present at much higher levels than initial levels of the library.

There are several possible reasons for the disparity and bias introduced into the library. The synthesis of the library relies on split-and-pool combinatorial chemistry, which for the case of solid-phase chemistry occurring on CPG (controlled pore glass) supports consisted of physical pooling, mixing, and redistribution of support-bound growing oligonucleotides.[272] This process was not automated and so it is likely that bias was introduced based on unequal splitting or incomplete mixing steps. PCR amplification is well known for its biases, particularly in SELEX, as the most enriched sequences following amplification are not necessarily the most abundant sequences prior to amplification. Aptamer-code constructs with more extensive modifications may also be less well-tolerated as substrates for polymerases necessary to the amplification of the oligonucleotide code.

Biases for certain monomers can be visually observed in **Table 2.2** in position 2, where C12 is present in most of the popular sequences, in position 3, where Nap, Phe, and T are strongly represented, and in position 2, where T and A account for almost all the popular sequences. However, in positions 7, a much greater diversity is observed with nearly all possible monomers present in the most popular sequences. It is difficult to observe any other trends or biases based on the sequences present above, though it does establish an important baseline. In the library, individual sequences were present at frequencies ranging from 12 to 493 reads per million. This means that aptamers that underwent selection will need to be present at much higher levels than 500 reads per million to argue they were enriched by the selection. Samples taken from the elutions with free thrombin, free TBA, and hot water featured many sequences well above this approximate baseline, though the fibrinogen elution featured only the unmodified aptamer at a frequency above 2000 reads per million.

Table 2.3. The 10 most popular sequences from the first elution (with free thrombin) in the first setup are presented below. The unmodified TBA has a condensed sequence of 5' T T T G T T T 3', which is the third most frequent sequence of this sample. An additional column presents the frequency of the sequence in the initial library.

Rank	Condensed sequence (5' to 3')							Reads per million (sample)	Reads per million (library)
	7	6	5	4	3	2	1		
1	T	T	Phe	C12	Phe	T	G	8439	321
2	T	T	Trp	G	Trp	T	G	7971	160
3	T	T	T	G	T	T	T	6661	148
4	Trp	T	Phe	C12	T	T	G	5447	210
5	G	T	Phe	C12	Phe	T	Nap	5240	173
6	T	T	Phe	A	Phe	T	Nap	4344	74
7	G	T	Phe	G	Phe	T	Nap	4151	99
8	G	T	Trp	A	Phe	T	G	4054	160
9	Trp	T	G	Sug	T	T	Nap	3930	74
10	T	T	Trp	A	Trp	T	G	3875	49

The sequences listed in **Table 2.3** are taken from an elution with free thrombin. This sample differs from the library in that it was subjected first to a negative selection to remove non-specific binders, then to a positive selection in which non-binding aptamers were removed by multiple washing steps, and finally to an elution using free thrombin to collect aptamers that were still bound to the immobilized thrombin. A key thing to observe is that the first 10 sequences range from 3875 to 8439 reads per million, which is about 8 to 16 times that of the most common sequences in the initial library. This demonstrates in general the utility of selection experiments, as there are sequences that are now overrepresented with respect to the original library as they fared better than the norm at binding the target. It is also interesting to compare the frequency of a particular sequence within the library sample and within the free thrombin elution sample. The most popular sequence of the elution with free thrombin was the 11^{th} most popular sequence in the library, so although it was the most popular sequence it was not necessarily the most enriched. It is also of note that the unmodified TBA sequence is the 3^{rd} ranked sequence,

with only two other sequences being more popular, though both were also present at higher levels in the initial library.

Another observation is that thymidine is overwhelmingly favoured in positions 6 and 2, which are the monomers of the unmodified TBA. Position 7 favours nucleosidic monomers G and T, though tryptophan monomers are also present. Position 5 has no clear favourite, with all six possible monomers present across the first 25 sequences (though sequences 11-25 are omitted in the table). Position 3 has a slight preference for phenylalanine monomers, though tryptophan and thymidine are also common.

Table 2.4. The 10 most popular sequences from the second elution (with free TBA) in the first setup are presented below. The unmodified TBA has a condensed sequence of 5' T T T G T T T 3', which is the most frequent sequence of this sample. An additional column presents the frequency of the sequence in the initial library.

Rank	Condensed sequence (5' to 3')							Reads per million (sample)	Reads per million (library)
	7	6	5	4	3	2	1		
1	T	T	T	G	T	T	T	28398	148
2	Trp	T	Phe	C12	T	T	G	21005	210
3	G	T	Phe	C12	Phe	T	Nap	17954	173
4	Trp	T	Phe	Sug	T	T	G	15623	99
5	T	T	Trp	G	Trp	T	G	14726	160
6	G	T	Phe	G	Phe	T	Nap	11431	99
7	Trp	T	Phe	G	T	T	G	10996	25
8	G	T	Trp	A	Phe	T	G	9656	160
9	T	T	Phe	C12	Phe	T	G	8934	321
10	Trp	T	C12	C12	T	T	Nap	8681	86

The sequences listed above in T are from an elution with solution containing the thrombin-binding aptamer lacking primers and the branching unit/code construct. Although the aptamer is present in the elution, it cannot be amplified by PCR as it lacks the necessary primer regions and so it does not contribute to the population of the unmodified sequence from the DNA-encoded library. This sample had approximately 230,000 reads, with greater enrichment observed for the most popular sequences, as the range of 4156 to 28398 reads per million is about 8 to 56 times the frequencies of the most popular sequences in the library. Interesting

comparisons can also be made between the two elution samples. The two most popular sequences in the first elution, 5′ T T Phe C12 Phe T G 3′ and 5′ T T Trp G Trp T G 3′, were ranked 5th and 9th overall in the second elution, behind the unmodified sequence. The 4th and 5th ranked sequences of the first elution were found in the same order but now as the 2nd and 3rd most popular sequences of the second elution. Later elutions would be expected to yield stronger binding aptamers, so this indicated some promise for the sequences 5′ Trp T Phe C12 T T G 3′ and 5′ G T Phe C12 Phe T Nap 3′. Again, thymidine was the only monomer present in the 10 most popular sequences in positions 2 and 6. Across the first 25 sequences (only 10 are shown in the table), position 7 displayed a strong but seemingly balanced preference toward tryptophan, thymidine, and guanosine, while position 3 preferred thymidine, phenylalanine, and tryptophan.

Table 2.5. The 10 most popular sequences from the third elution (with free fibrinogen) are presented below. The unmodified TBA has a condensed sequence of 5′ T T T G T T T 3′, which is the most frequent sequence of this sample. An additional column presents the frequency of the sequence in the initial library.

Rank	**Condensed sequence (5′ to 3′)**							**Reads per million (sample)**	**Reads per million (library)**
	7	**6**	**5**	**4**	**3**	**2**	**1**		
1	T	T	T	G	T	T	T	36499	148
2	Sug	T	Phe	G	Nap	C12	Nap	1650	420
3	Trp	T	Trp	Nap	Nap	C12	G	1340	210
4	Ant	Phe	C12	Alk	Nap	C12	Phe	1053	198
5	C12	Alk	C12	C12	T	Alk	Nap	1026	86
6	C12	Nap	Phe	Nap	Nap	Sug	G	984	185
7	C12	Nap	G	G	Nap	Sug	G	953	86
8	Trp	Phe	C12	Alk	Nap	Alk	Nap	912	160
9	T	Alk	Trp	Nap	Nap	C12	T	898	185
10	Trp	Nap	G	G	Nap	T	Nap	889	99

As mentioned earlier, the fibrinogen elution was dominated by the thrombin-binding aptamer in its unmodified form, with 36499 reads per million compared with the next most popular sequence at only 1650 reads per million. These sequences were not very enriched with respect to the original library, and it was difficult to draw conclusions from this sample other than the unmodified sequence being abundant. Approximately 220,000 reads were collected for the

sample, with 54,707 unique sequences, and one sixth of the total reads were for the unmodified thrombin binding aptamer. Some of the prior prominent sequences from the free aptamer and thrombin elutions were present at frequencies ranging from 59 to 392 reads per million, which were comparable with their rates in the initial library. Most of the trends visible in earlier tables were not followed in **Table 2.5**.

Table 2.6. The 10 most popular sequences from the fourth elution (with hot water) are presented below. The unmodified TBA has a condensed sequence of 5' T T T G T T T 3', which is the 10th most frequent sequence of this sample. An additional column presents the frequency of the sequence in the initial library.

Rank	Condensed sequence (5' to 3')							Reads per million (sample)	Reads per million (library)
	7	6	5	4	3	2	1		
1	G	T	Phe	C12	Phe	T	Nap	9204	173
2	T	T	Trp	G	Trp	T	G	6782	160
3	Trp	T	Phe	Sug	T	T	G	6263	99
4	Trp	T	Phe	G	T	T	G	5485	25
5	G	T	Phe	G	Phe	T	Nap	5287	99
6	Trp	T	C12	C12	T	T	Nap	5201	86
7	Trp	T	Phe	C12	T	T	G	5176	210
8	T	T	C12	C12	Trp	T	T	4991	86
9	G	T	Sug	G	Trp	T	Phe	3953	62
10	T	T	T	G	T	T	T	3731	148

The elution with hot water provided more meaningful results than the elution with fibrinogen, as it was not dominated by the unmodified sequence and had frequencies comparable to those of the free thrombin elution. In approximately 81,000 reads there were 32,000 unique sequences, this time with only 400 copies of the unmodified aptamer. The most abundant sequence was the sequence that previously had appeared as the third most common sequence in the aptamer elution, 5' G T Phe C12 Phe T Nap 3'. The sequence that was second most common in the aptamer elution, 5' G T Phe C12 Phe T Nap 3', was identified as the 7th most common sequence in the hot water elution. The sequence that was 2nd and 5th overall in the first and second elutions, 5' T T Trp G Trp T G 3', was 2nd overall in the hot water elution sample. Most of the other highly ranked sequences in the hot water elution could be found in the tables for the

free thrombin and aptamer elutions. The positional trends observed in these elutions were also observed in hot water elution.

2.3.2.2. Positional tendencies

The tables above are useful for finding entire sequences that may present comparable or better binding affinities than the unmodified aptamer. However, it is difficult to make conclusions about the preferences for monomers in specific positions of the aptamer. Another way to look at the sequencing data is to compare how frequently the various monomers for a given position appear with respect to each other. This was accomplished by a simple program we designed that was able to count how many times a codon appeared in a position and normalize that number with respect to the other monomers available at that position.

The below figures visually depict the positional tendencies for each of the five samples. **Table 2.7** summarizes the charts with the most popular modification for a position for a sample.

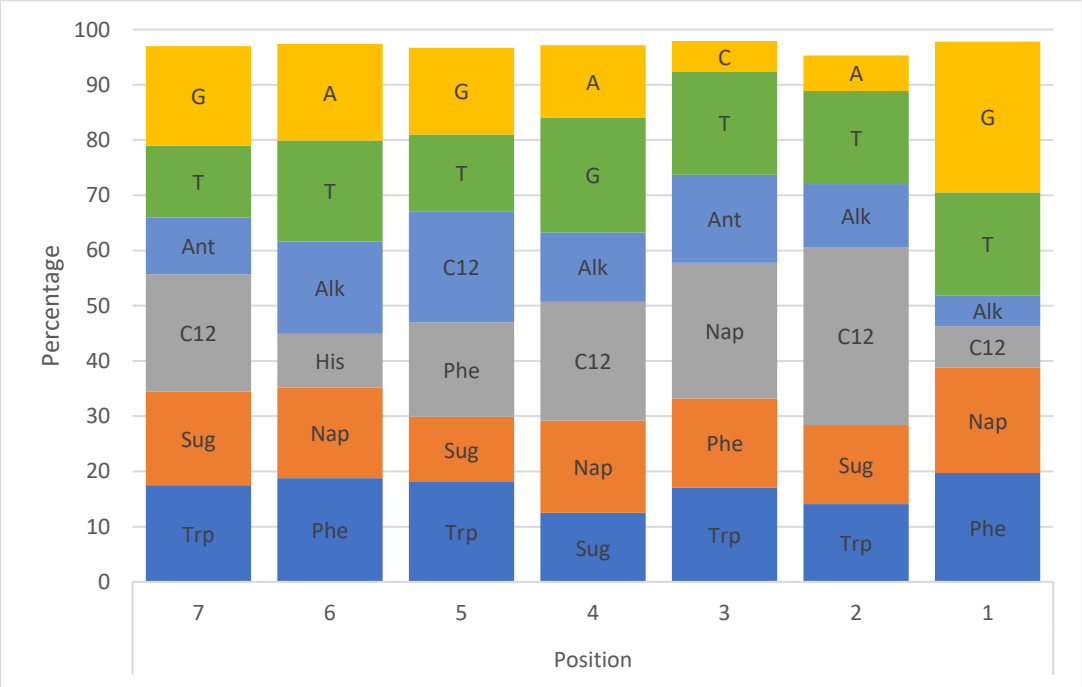


Figure 2.5. The percentages of each possible monomer in a numbered position are displayed above for the library sample in setup 1, which was taken prior to any selection steps.

Bias was visible in the library based on **Figure 2.5**, as several positions are not as well-balanced as they would be in an unbiased library. An ideal library would feature each monomer at 16.7 %; however, in this setup, there are some clear anomalies. For instance, in position 2, C12 is present

in 32.2 % of sequences while A is only present in 6.4 % of sequences. This clear bias will affect the positional tendencies of any downstream samples. This level of disparity is likely due to unequal splitting during the split-and-pool step for position 2, as the same difference between C12 and A is not seen for position 4, in which they are possible monomers. Positions 7 through 4 are reasonably well balanced, ranging from 10.3 - 22.0 % per monomer, with over half of monomers falling within 3 % of the expected 16.7 % value. Positions 3, 2, and 1 all have one monomer present in less than 7 % of sequences and one present in more than 24 % of sequences. This distribution is important to consider when making conclusions based off the data from samples that underwent a form of selection.

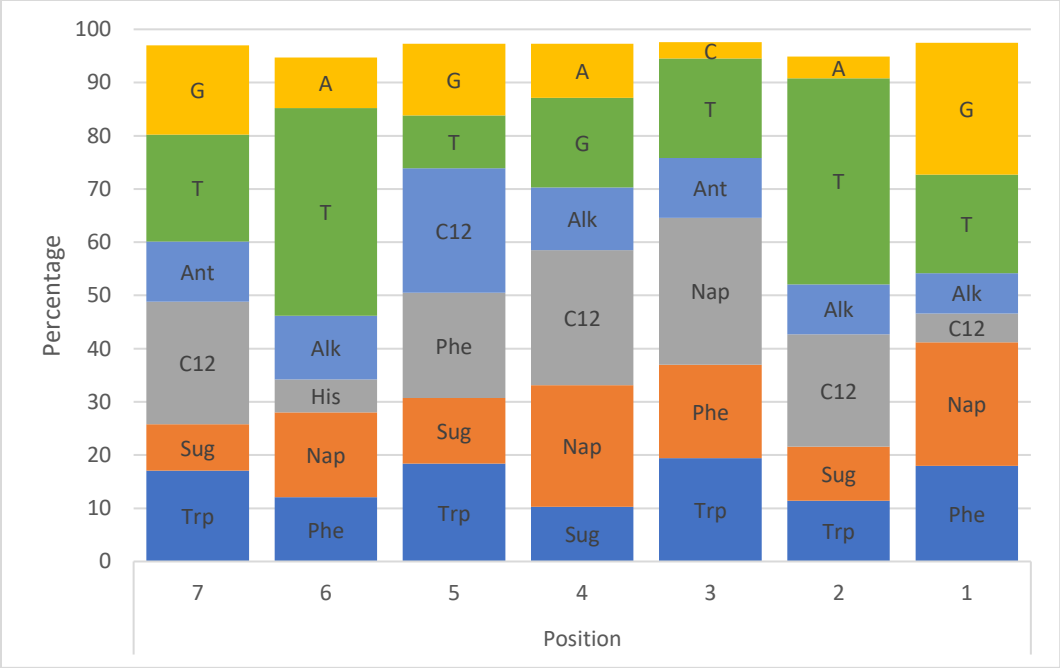


Figure 2.6. The percentages of each possible monomer in a numbered position are shown above for the sample from eluting with free thrombin in setup 1.

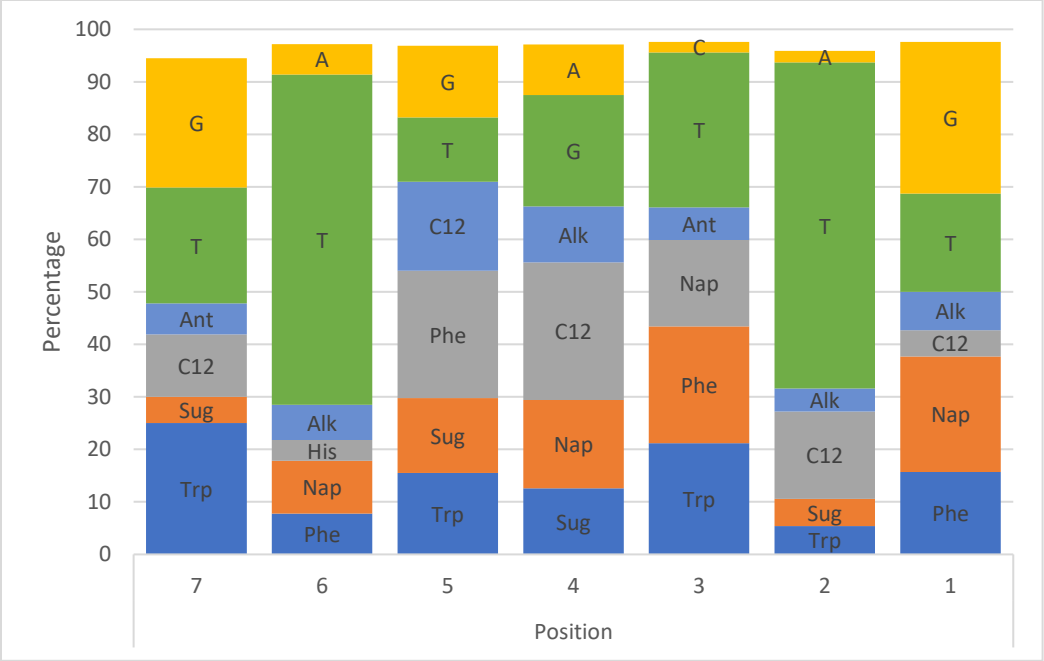


Figure 2.7. The percentages of each possible monomer in a numbered position are shown above for the sample from eluting with free thrombin-binding aptamer in setup 1.

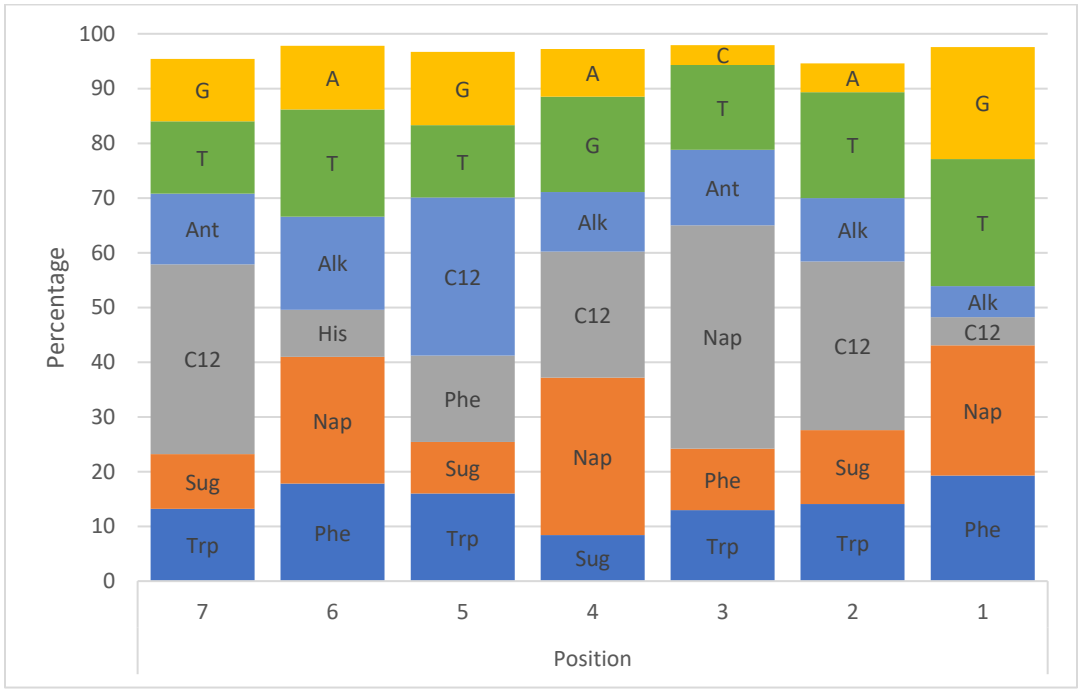


Figure 2.8. The percentages of each possible monomer in a numbered position are shown above for the sample from eluting with free fibrinogen in setup 1.

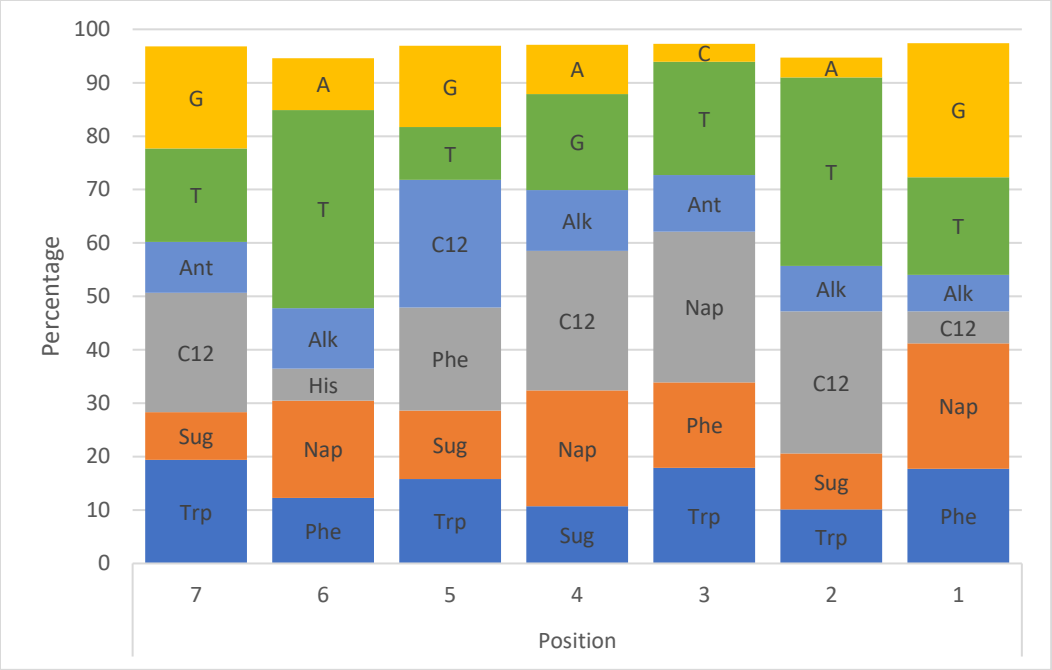


Figure 2.9. The percentages of each possible monomer in a numbered position are shown above for the sample from eluting with hot buffer solution in setup 1.

Of the samples above, the fibrinogen elution (**Figure 2.8**) appears to the be the anomaly as it lacks the typical preference for thymidine in positions 2 and 6. This result is strange, particularly as one sixth of the reads were for the unmodified aptamer that features thymidine in these positions. It is possible that the other aptamers bound to a different region of the immobilized thrombin and were displaced by steric or allosteric effects. The other samples did indicate some trends and preferences for certain monomers in the selected libraries. Thymidine is the most common monomer in positions 2 and 6 for all three types of elutions (see **Figure 2.6** for the free thrombin elution, **Figure 2.7** for the TBA elution, and **Figure 2.9** for the hot water elution). The free thrombin and hot water elutions are very similar. The main differences for the TBA elution were that it exhibited a stronger preference for thymidine in positions 2, 3, and 6, a preference of phenylalanine over C12 in position 5, and a preference for tryptophan and guanosine over C12 and thymidine in position 7.

Table 2.7. The most common monomer per position for each sample for the first selection experiment is presented below. The positions are labeled as before from 7 to 1 (5' to 3'). In brackets are the actual percentages at which the monomers occurred in their sample at the position indicated. Individual entries are bolded in cases where the most common monomer is at least 10% greater (in difference) than the second most common monomer.

Position **Sample**	**7** **Mod (%)**	**6** **Mod (%)**	**5** **Mod (%)**	**4** **Mod (%)**	**3** **Mod (%)**	**2** **Mod (%)**	**1** **Mod (%)**
Library	C12 (21.0)	Phe (18.8)	C12 (20.1)	C12 (22.0)	Nap (25.0)	**C12 (32.0)**	G (27.3)
Thrombin elution	C12 (23.0)	**T (39.0)**	C12 (23.4)	C12 (25.4)	Nap (27.6)	**T (38.7)**	Nap (23.2)
TBA wash	Trp (25.0)	**T (62.9)**	Phe (24.2)	C12 (26.2)	T (29.5)	**T (62.1)**	G (28.9)
Fibrinogen elution	**C12 (34.7)**	Nap (23.2)	**C12 (28.9)**	**Nap (40.8)**	Nap (28.8)	**C12 (30.8)**	Nap (23.8)
Hot water elution	C12 (22.4)	**T (37.1)**	C12 (23.9)	C12 (26.1)	Nap (28.2)	T (35.3)	G (25.1)

More general trends across all samples can be observed in **Table 2.7**, which presents the most common monomer for a specified position and sample. The library favoured C12 in several positions, particularly in position 2 where it was well outside the typical range. C12 is the most common modification in most instances for positions 7, 6, and 4, though it tends to be only slightly more enriched than the library, apart from outliers in the fibrinogen elution. Similar results can be observed with position 1, where there is an initial bias for guanosine that extends to the TBA and hot water elutions. Tryptophan was enriched for the TBA elution in position 7, as it increased from 17.5 to 25.0 %; naphthalene was enriched for all elutions in position 1, increasing from 19.1 to as much as 23.8 %; and phenylalanine was enriched in position 5, increasing from 17.0 to 24.2 %. It is also noteworthy that most of the non-nucleosidic monomers that were enriched to some degree tended to be more hydrophobic monomers, such as naphthalene, C12, and phenylalanine. As the modifications target primarily binding interfaces, these modifications are likely involved in favourable interactions with residues in the fibrinogen-binding site of thrombin, particularly those in positions 7, 6, 3, and 2, which are part of the TT loops believed to interact with the site. Cation–π, π–π and C-H–pi interactions[277] are all possible non-covalent interactions between the basic residues in the fibrinogen-binding sites and the aromatic monomers favoured in the library screening, particularly for phenylalanine and naphthalene. In one example in the literature, a post-SELEX screening of modified TBAs used amino acid derivatized monomers and found that phenylalanine, tryptophan and methionine were all favourable in the T3 and T12 (positions 7 and 3, respectively, in our design).[278] In

particular, T3 phenylalanine modifications resulted in the most enhanced binding affinity, thought to be from favourable interactions with a tyrosine residue on thrombin based on molecular modeling.[278] The carbon chains in the C12 and naphthalene monomers can also engage in non-specific hydrophobic interactions with the aliphatic chains of the lysine and arginine residues present in the binding site. In positions of the aptamer outside of the TT loops, these monomers may also provide added flexibility or rotational freedom for stronger interactions between the other modified monomers and the target; however, in the literature, flexible monomers tend to improve stability and rarely improve binding affinity.[21, 146, 148]

2.3.2.3. Validation of sequencing data

To assess what level of variability could be attributed to the sequencing itself, the library, thrombin, and TBA samples were re-sequenced. Positional tendencies and sequence popularities were compared across the three samples. Samples were re-sequenced with less reads, around 21,000 reads per sample, with similar results obtained for sequence popularities. The order of the first five sequences for both the thrombin and TBA samples was the same, with slight variations in frequencies, and very few of the other top 25 sequences were not found when re-sequenced. The library sample was the least consistent across sequencing, with only 7 of 25 sequences found in both sets of data, though the most popular sequence was the same for both. In terms of positional tendencies, the library also varied slightly in its biases, though the general trends observed earlier still applied to the almost all the data from the re-sequencing. For the library, the re-sequencing presented a less biased sample, with the proportion of sequences having C12 in positions 7 and 2 now closer to the norm. There were no notable differences for the much less diverse thrombin and TBA samples upon re-sequencing, only slight variations in percentages but no changes in order. Overall, this check indicated some variation in the library is to be expected with repetitive sampling, though the order of the most popular sequences and positional preferences remains quite consistent. Greater sources of error are expected from PCR amplification and split-and-pool combinatorial synthesis.

2.3.3. Setup 2

In the second selection setup (**Figure 2.10**), the elution steps in order were as follows: two elutions with regular TBA and one elution with free thrombin. The fibrinogen and hot water elutions were omitted this time as they did not prove as useful as the preceding elutions. The order of the free thrombin and TBA aptamer elutions were switched to determine if the thrombin step interfered with future elutions and to test an alternative selection process. The library used for this setup was from the same source as in setup 1. A fluorescently tagged library was also used in an analogous selection to facilitate elutions and as a source of comparison for the normal library. This library was constructed simply by adding a fluorescein phosphoramidite to the 5' end of the aptamer for a portion of the library that was used for both setups. Sequence popularities and positional tendencies were tabulated and compared as with the previous setup.

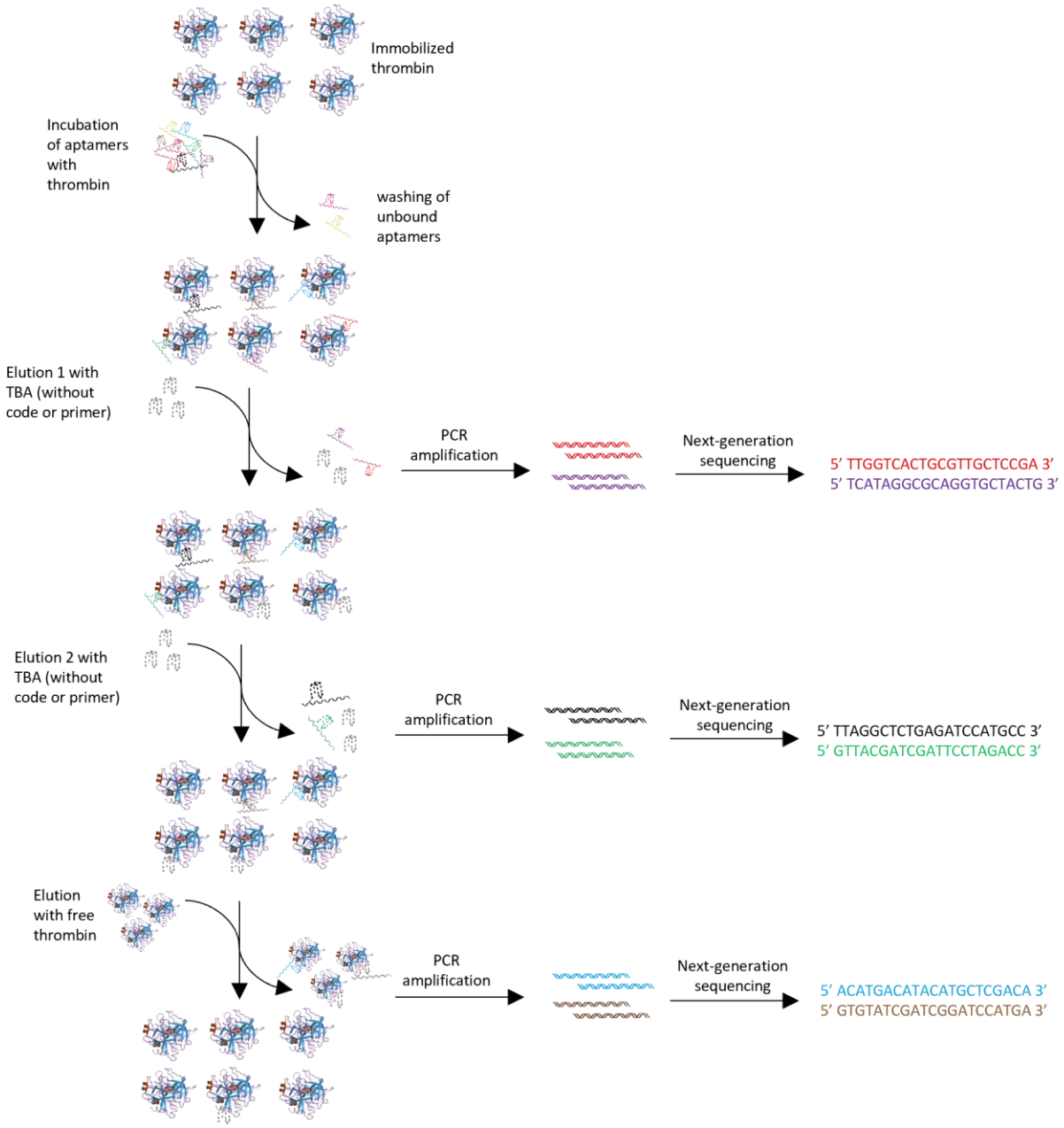


Figure 2.10. An overview of the second selection setup is presented above (counter-selection step not shown). The library of modified DNA-encoded aptamers was incubated with immobilized thrombin and unbound aptamers were removed by washing steps. An elution with the unmodified thrombin-binding aptamer (TBA), lacking any primers or code necessary for amplification, was performed first and any recovered sequences that were displaced were amplified by PCR and sequenced. Another elution with the unmodified TBA was performed and recovered sequences were amplified and sequenced. Finally, an elution with free thrombin was performed to collect aptamers still bound to the immobilized thrombin, which were amplified and sequenced. The same process was also followed for a modified library tagged with a fluorophore. In both cases, a sample from the initial library was also amplified and sequenced.

2.3.3.1. Popular sequences

The library and fluorescently tagged libraries were analyzed and the unique sequences were ranked based on their populations. The library sample featured approximately 33,000 reads with

21,000 unique sequences, of which 15,000 occurred only once. The fluorescent library, referred to in short as library*, with the asterisk denoting a fluorescent tag, featured about 48,000 reads with 29,000 unique sequences, out of which 20,000 occurred without copies. These ratios are similar to those of the library sample sequenced in the first selection setup. The proportion of unique sequences is quite similar for the two libraries in the second setup, though the popular sequences differed quite noticeably. The unmodified TBA sequence was found at 30 reads per million in the library sample and at 21 reads per million in the fluorescent library sample (library*). Most of the top popular sequences in the two library samples were not detected in the other, as shown in the rightmost columns of **Table 2.8** and **Table 2.9**. As with the previous selection setup, it is hard to draw many conclusions from the data presented in the tables. There is a noticeable bias toward C12 in position 2 for both libraries, particularly for the library* sample. This trend was observed in the previous selection. The library* sample also featured primarily G in position 4 of the popular sequences. Once again, in the following tables, the condensed sequence is shown in place of the full sequence. The full sequence is 5' GG **76** GG **54**T GG **32** GG **1** 3', which the bolded numbers corresponding to the condensed sequence, 5' **7 6 5 4 3 2 1** 3'.

Table 2.8. The 10 most popular sequences of the library sample (lib) used in the second selection setup are ranked below. An additional column presents the frequency of the same sequences in the fluorescent library sample.

Rank	Condensed sequence (5' to 3')							Reads per million (library)	Reads per million (library*)
	7	**6**	**5**	**4**	**3**	**2**	**1**		
1	Trp	A	C12	Nap	Nap	T	G	639	0
2	G	T	C12	G	Nap	C12	G	578	126
3	Sug	T	Trp	C12	Nap	C12	Phe	548	105
4	G	T	C12	G	T	Alk	Nap	518	0
5	Sug	Alk	Trp	G	Nap	C12	Nap	518	0
6	T	A	G	A	Phe	C12	Nap	487	63
7	Sug	Alk	T	C12	Nap	C12	G	487	0
8	T	T	C12	A	Ant	T	Nap	426	0
9	Sug	T	C12	C12	Nap	C12	Phe	426	21
10	Trp	A	T	G	T	C12	Nap	396	0

Table 2.9. The 10 most popular sequences of the fluorescently tagged library sample (library) used in the second selection setup are ranked below. An additional column presents the frequencies at which the most popular sequences appeared in the non-fluorescent library.*

Rank	Condensed sequence (5' to 3')							Reads per million (library*)	Reads per million (library)
	7	6	5	4	3	2	1		
1	G	Phe	Trp	Nap	Nap	C12	G	566	0
2	G	A	C12	A	Phe	C12	Nap	503	0
3	T	T	T	G	Phe	C12	Nap	503	0
4	C12	Nap	T	G	Trp	C12	Nap	503	0
5	Sug	T	Phe	G	Nap	C12	Nap	503	0
6	Sug	Nap	Trp	G	T	C12	G	482	30
7	Sug	Phe	Phe	G	Nap	C12	G	461	0
8	Trp	A	T	G	Nap	C12	Phe	461	0
9	Trp	T	T	C12	Trp	C12	G	461	183
10	T	Phe	C12	G	Phe	C12	T	420	0

Two elutions were consecutively performed with the unmodified aptamer lacking primers and the code construct of the DEL. The samples from each wash are termed Apt1 (**Table 2.10**) and Apt2 (**Table 2.11**), with analogous samples using the fluorescent library termed Apt1* (**Table 2.12**) and Apt2* (**Table 2.13**).

Table 2.10. The 10 most popular sequences of the first elution with TBA (Apt1) used in the second selection setup are ranked below. An additional column presents the frequencies of the sequences in the original non-fluorescent library.

Rank	Condensed sequence (5' to 3')							Reads per million (Apt1)	Reads per million (library)
	7	6	5	4	3	2	1		
1	T	T	G	G	Phe	T	Nap	20715	213
2	G	T	Trp	Sug	Phe	T	Nap	15892	213
3	T	T	T	C12	Ant	T	Nap	12029	183
4	Trp	T	C12	C12	T	T	G	11615	274
5	G	T	C12	C12	T	T	Nap	11174	91
6	G	T	T	Nap	T	T	Nap	10578	244
7	Trp	T	Phe	G	T	T	Nap	10085	91
8	G	T	C12	C12	Trp	T	T	9852	183
9	T	T	T	C12	Phe	T	G	9385	152

10	T	T	C12	Alk	Phe	T	G	8451.73	61

Table 2.11. The 10 most popular sequences of the first elution with TBA (Apt1) used in the second selection setup for the fluorescent library are ranked below. An additional column presents the frequencies of the sequences in the original fluorescent library.*

Rank	Condensed sequence (5′ to 3′)							Reads per million (Apt1*)	Reads per million (library*)
	7	6	5	4	3	2	1		
1	T	T	Phe	C12	Phe	T	G	21289	189
2	G	T	Phe	C12	Phe	T	Nap	16314	168
3	T	T	Trp	G	Trp	T	G	15945	84
4	Trp	T	Phe	C12	T	T	G	13338	231
5	Trp	T	C12	C12	T	T	Nap	11535	105
6	Trp	T	G	Sug	T	T	Nap	11014	84
7	G	T	Phe	G	Phe	T	Nap	10775	231
8	T	T	Phe	A	Phe	T	Nap	10101	84
9	G	T	Trp	A	Phe	T	G	8190	189
10	T	T	Sug	C12	Trp	T	T	8190	126

As was observed in the previous selection scheme, thymidine dominates positions 2 and 6. Similar trends are also observed with phenylalanine prominent in position 3 and naphthalene prominent in position 1. The individual sequences vary quite strongly between Apt1 and Apt1*, much like the initial libraries (library and library*). Some of the sequences that stood out in setup 1 also appear at higher frequencies, with 5′ G T Phe C12 Phe T Nap 3′, 5′ T T Trp G Trp T G 3′, and 5′ Trp T Phe C12 T T G 3′ appearing in rows 2 through 4 of

Table 2.11. It is unexpected though that these samples, which occurred previously for the non-fluorescent library, are only found in prominent positions for the selection using the fluorescent library. In Apt1 and Apt1*, the unmodified TBA sequence was found at 648 and 717 reads per million, respectively.

Table 2.12. The 10 most popular sequences of the second elution with TBA (Apt2) used in the second selection setup are ranked below. An additional column presents the frequencies of the sequences in the original non-fluorescent library.

Rank	Condensed sequence (5′ to 3′)							Reads per million (Apt2)	Reads per million (library)
	7	6	5	4	3	2	1		
1	G	T	C12	C12	Trp	T	C12	33036	122
2	Trp	T	C12	C12	T	T	G	24532	274
3	G	T	Trp	Sug	Phe	T	Nap	19841	213
4	G	T	C12	Nap	Phe	T	Nap	13897	91
5	G	T	C12	C12	Phe	T	G	13570	91
6	T	T	G	G	Phe	T	Nap	12216	213
7	Trp	T	Phe	G	T	T	Nap	11238	91
8	G	T	C12	G	Phe	T	Nap	10836	152
9	G	T	C12	C12	Trp	T	T	9080	183
10	T	T	C12	Nap	Trp	T	G	8955	122

Table 2.13. The 10 most popular sequences of the second elution with TBA (Apt2) used in the second selection setup for the fluorescently tagged library are ranked below. An additional column presents the frequencies of the sequences in the original fluorescent library.*

Rank	Condensed sequence (5′ to 3′)							Reads per million (Apt2*)	Reads per million (library*)
	7	6	5	4	3	2	1		
1	G	T	Phe	C12	Phe	T	Nap	25318	168
2	Trp	T	C12	C12	T	T	Nap	16998	105
3	Trp	T	Sug	Nap	T	T	C12	14012	42
4	G	T	Phe	G	Phe	T	Nap	13757	231
5	Trp	T	Phe	C12	T	T	G	13348	231
6	Trp	T	Phe	Nap	T	T	C12	11434	42
7	T	T	Trp	G	Trp	T	G	11204	84
8	Trp	T	G	Sug	T	T	Nap	10388	84
9	G	T	Sug	Nap	Phe	T	Nap	9852	273
10	Trp	T	Phe	Sug	T	T	G	8065	84

In the second elution with aptamer solution, similar frequencies as those of the first elution were observed for the 10 most popular sequences, ranging from about 8,000 to 30,000 reads per million, compared with levels around 600 reads per million for the most popular sequences of the library samples. Most sequences that appeared in Apt1 appeared in Apt2, as was the case for Apt1* and Apt2*, though variations in rankings were observed. In Apt2 and Apt2*, the unmodified TBA sequence appeared at frequencies of 1806 and 1889 reads per million, respectively. This sequence through both elutions has been enriched to very similar levels for both the fluorescent and non-fluorescent library. Similar trends with positions 7, 6, 2, and 1 were observed for both samples.

Table 2.14. The 10 most popular sequences of the third elution with free thrombin (Thr) used in the second selection setup are ranked below. An additional column presents the frequencies of the sequences in the original non-fluorescent library.

Rank	Condensed sequence (5' to 3')							Reads per million (Thr)	Reads per million (library)
	7	6	5	4	3	2	1		
1	T	T	T	G	T	T	T	25877	30
2	G	T	Trp	Sug	Phe	T	Nap	1839	213
3	G	T	C12	C12	Trp	T	T	1467	183
4	Trp	T	Phe	G	T	T	Nap	1442	91
5	Trp	T	C12	C12	T	T	G	1392	274
6	Trp	Sug	Phe	Ant	C12	A	Nap	1367	30
7	T	T	G	G	Phe	T	Nap	1317	213
8	Trp	T	Phe	G	T	T	Phe	1143	61
9	C12	Nap	G	G	T	Sug	G	1119	213
10	G	T	T	CAT	Trp	T	Nap	1044	91

Table 2.15. The 10 most popular sequences of the third elution with free thrombin (Thr) used in the second selection setup for the fluorescently tagged library are ranked below. An additional column presents the frequencies of the sequences in the original fluorescent library.*

Rank	Condensed sequence (5' to 3')							Reads per million (Thr*)	Reads per million (library*)
	7	6	5	4	3	2	1		
1	T	T	T	G	T	T	T	37117	21
2	Sug	T	Phe	G	Nap	C12	Nap	3638	503

3	Trp	Sug	Phe	Ant	C12	A	Nap	1854	0
4	Trp	T	Trp	Nap	Nap	C12	G	1691	378
5	C12	Alk	C12	Nap	T	T	Nap	1367	252
6	T	Alk	Trp	Nap	Nap	C12	T	1344	126
7	Trp	Nap	G	G	Nap	T	Nap	1251	105
8	Ant	A	Trp	Alk	Nap	C12	G	1089	252
9	C12	T	Phe	Nap	Nap	T	Nap	1042	147
10	C12	Alk	C12	C12	T	Alk	Nap	1042	147

The final elutions with free thrombin showed similar results as those seen in setup 1 with the fibrinogen elution, as the unmodified TBA sequence appeared most prominently and dwarfed all other samples by more than a factor of 10. Other sequences appeared only at frequencies below 2000 reads per million, with the small exception of the second ranked sequence in **Table 2.15**, 5' Sug T Phe G Nap C12 Nap 3', which appeared at 3638 reads per million. This sequence also appeared in a similar position in the fibrinogen sample, though the only other samples in which it appeared prominently were in library samples, indicating its prominence in later rounds may have been due to a pre-existing bias. The sequences previously tracked through the fluorescently tagged samples did not appear in the 10 most prominent sequences, though 5' G T Phe C12 Phe T Nap 3' did appear at 857 and 423 reads per million in the samples Thr* and Thr, respectively. The free thrombin samples were also the only ones to counter the pervasive thymidine trend at positions 2 and 6.

2.3.3.2. Positional tendencies

The below figures visually depict the positional tendencies for each of the three samples for both types of libraries. **Table 2.16** summarizes the charts with the most popular modification for a position for a sample.

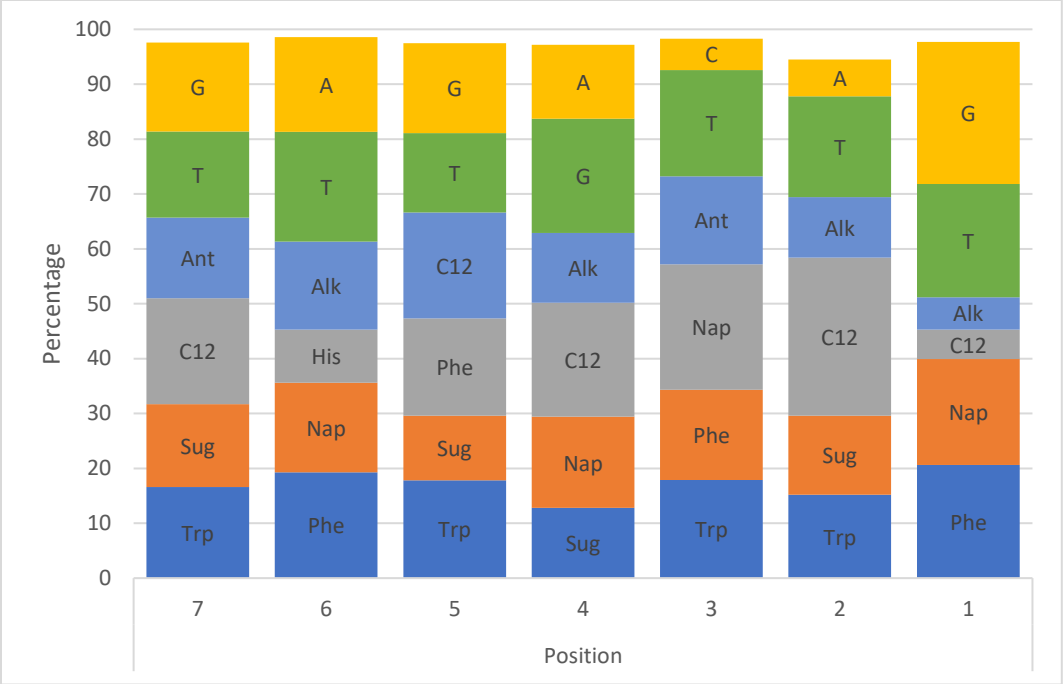


Figure 2.11. The percentages of each possible monomer in a numbered position are shown above for a sample of the library used in setup 2 (library).

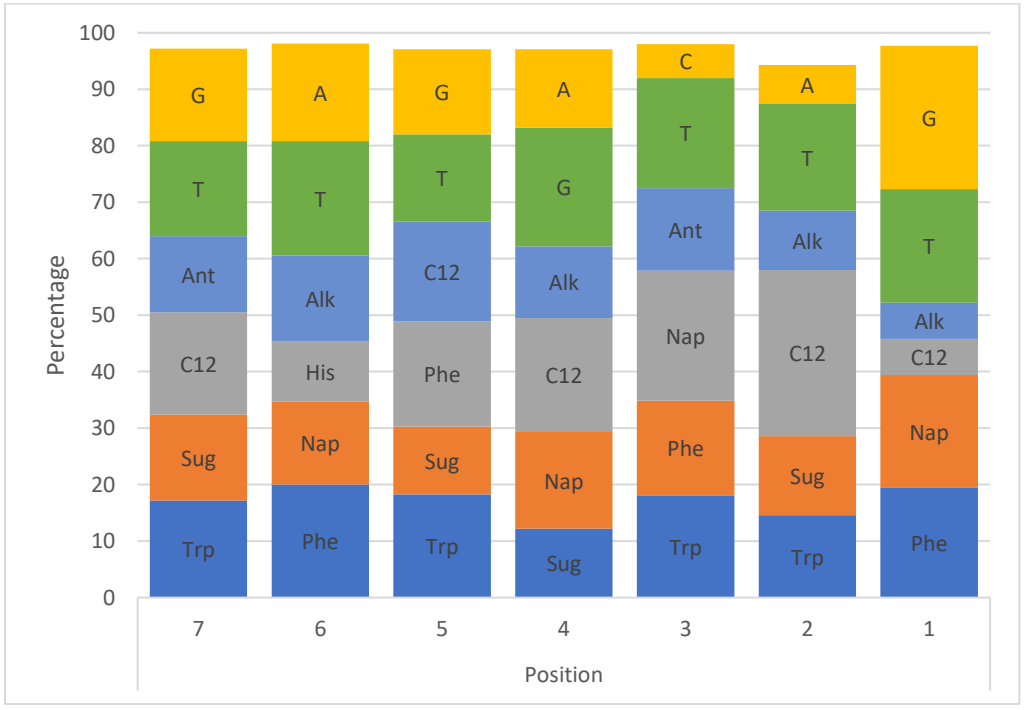


Figure 2.12. The percentages of each possible monomer in a numbered position are shown above for a sample of the fluorescently tagged library (library) used in setup 2.*

The positional distributions were very similar for both library samples, as can be observed visually in **Figure 2.11** (library) and **Figure 2.12** (library*), despite differing results for individual sequence popularities. The library samples revealed some biases primarily in positions 1, 2, and 3. Guanosine was overrepresented in position 1, while C12 and alkyne were underrepresented. In position 2, C12 was overrepresented and adenosine was underrepresented. In position 3, naphthalene was slightly overrepresented while cytosine was underrepresented. Positions 7 through 4 were again more well-balanced, with position 7 showing the narrowest range of percentages at 14.7-19.3 % for the non-fluorescent library and 13.5-18.1 % for the fluorescent library. The remaining distributions are presented in the figures below.

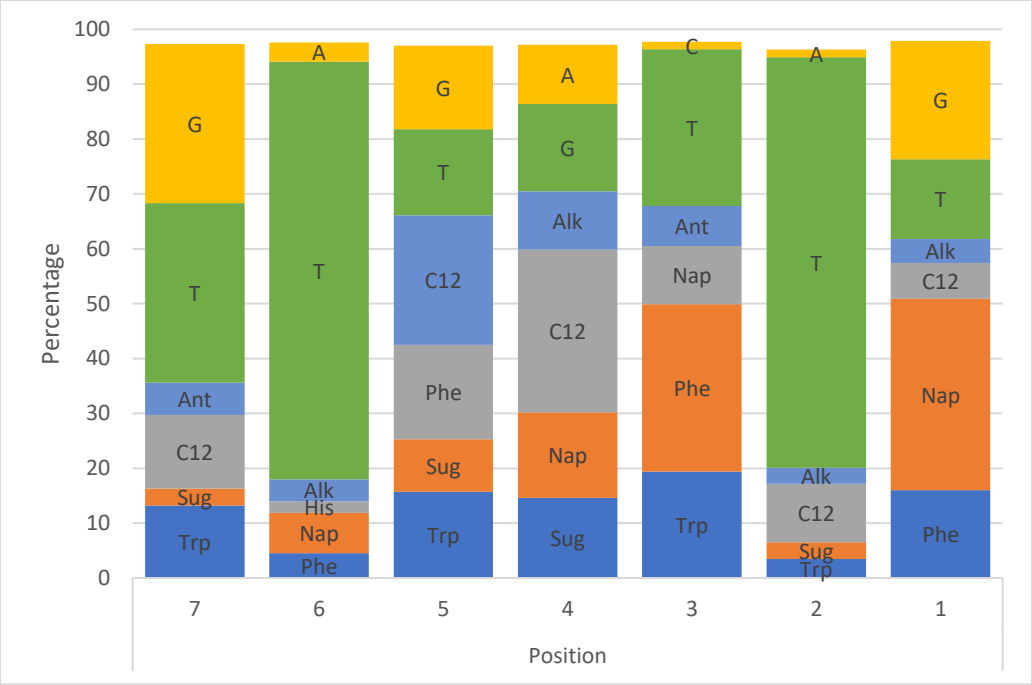


Figure 2.13. The percentages of each possible monomer in a numbered position are shown above for a sample from the first elution with TBA solution (Apt1) for the non-fluorescent library used in setup 2.

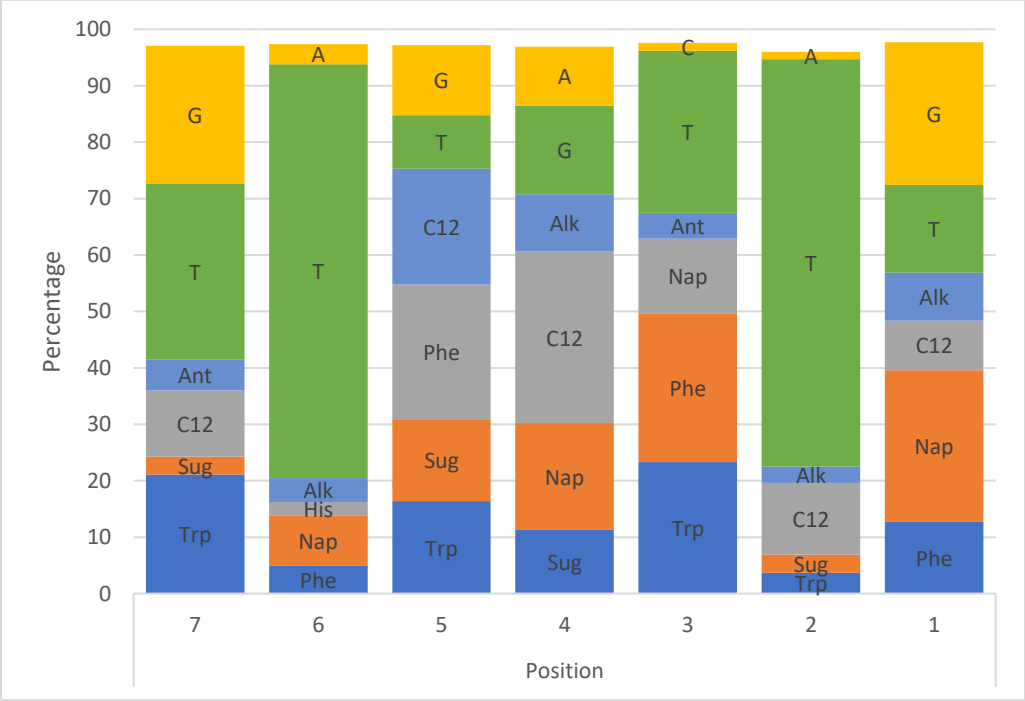


Figure 2.14. The percentages of each possible monomer in a numbered position are shown above for a sample from the first elution with TBA solution (Apt1) for the fluorescently tagged library used in setup 2.*

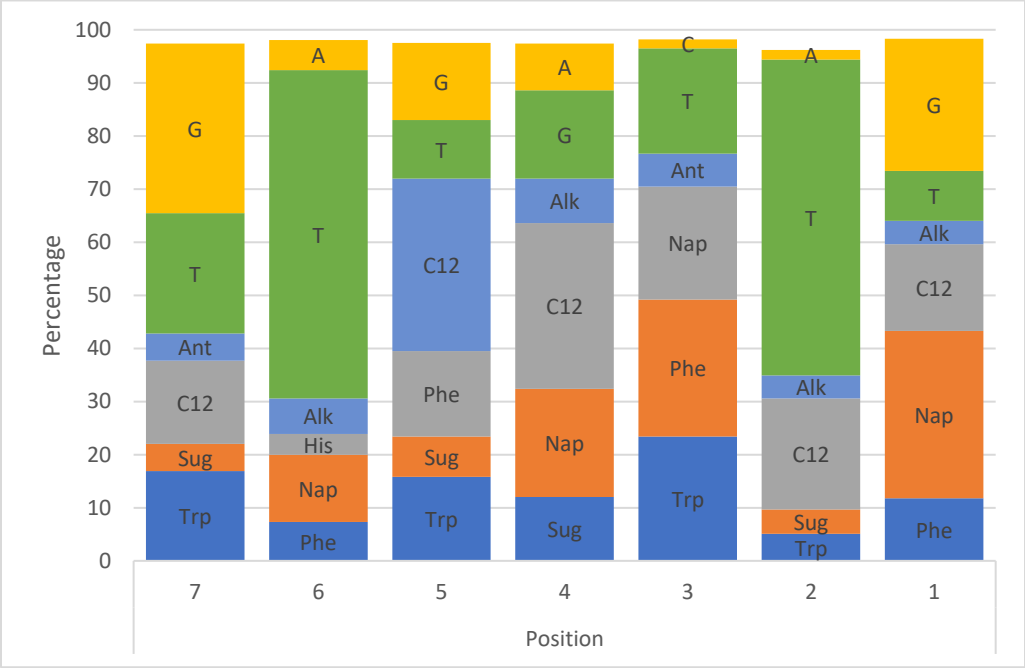


Figure 2.15. The percentages of each possible monomer in a numbered position are shown above for a sample from the second elution with TBA solution (Apt2) for the non-fluorescent library used in setup 2.

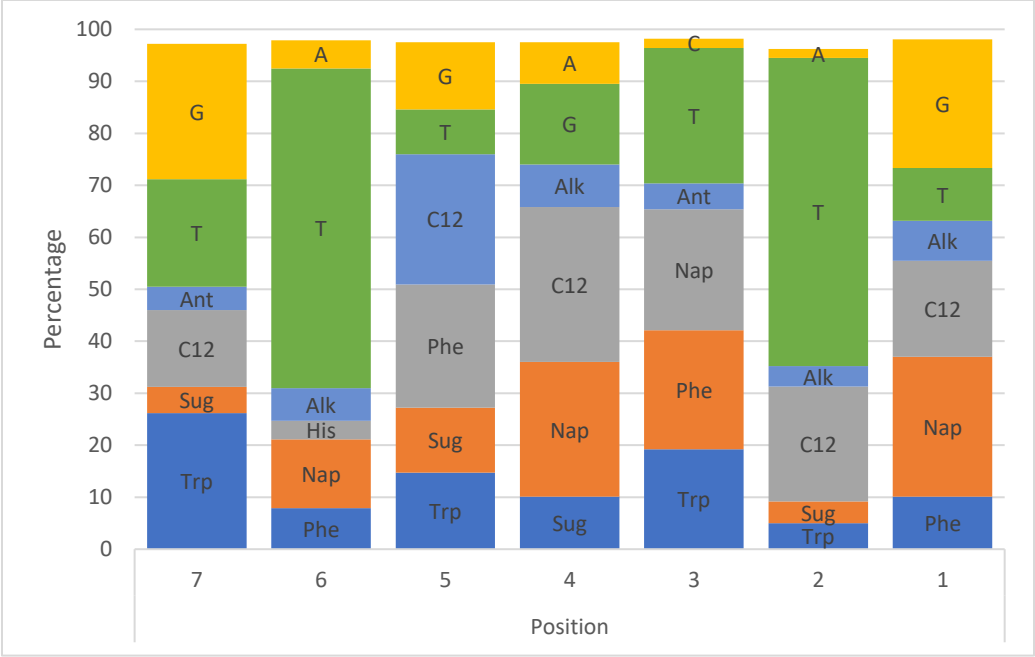


Figure 2.16. The percentages of each possible monomer in a numbered position are shown above for a sample from the second elution with TBA solution (Apt2) for the fluorescently tagged library used in setup 2.*

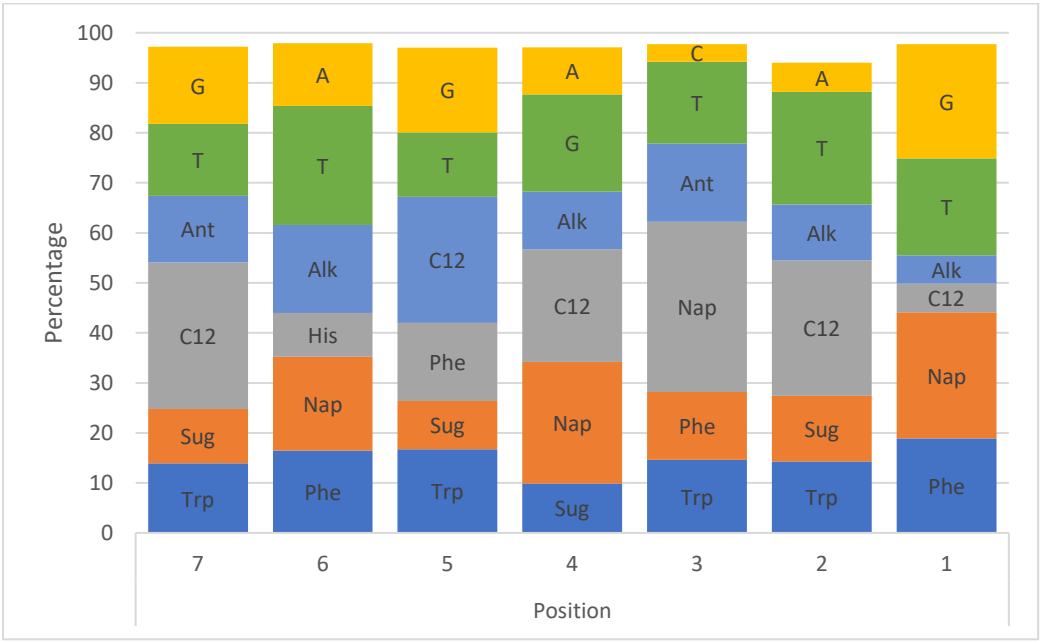


Figure 2.17. The percentages of each possible monomer in a numbered position are shown above for a sample from the third elution with free thrombin solution (Thr) for the non-fluorescent library used in setup 2.

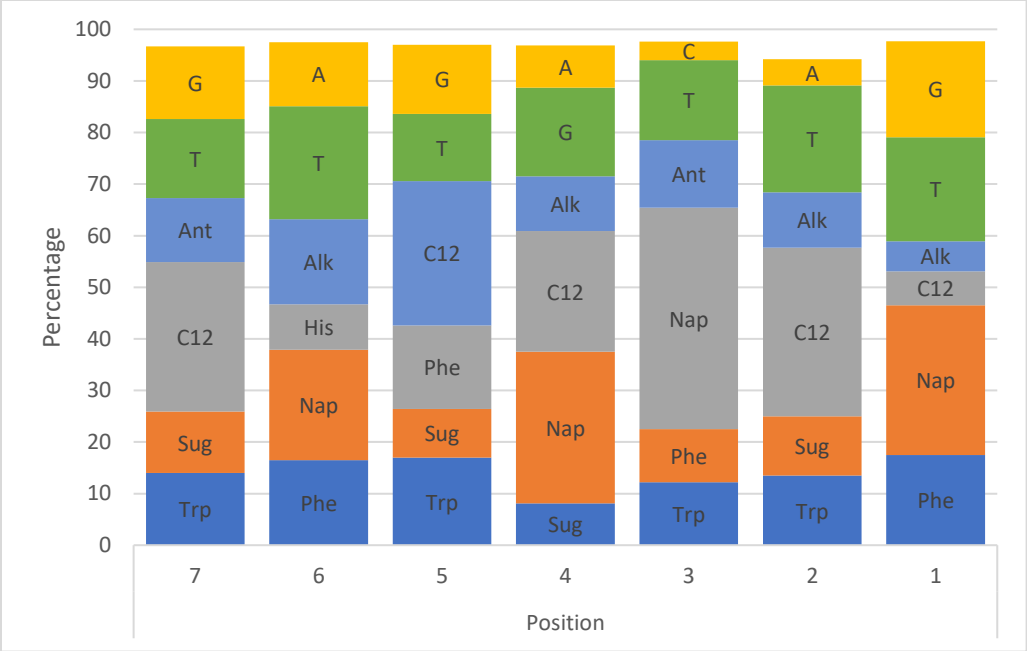


Figure 2.18. The percentages of each possible monomer in a numbered position are shown above for a sample from the third elution with free thrombin solution (Thr) for the fluorescently tagged library used in setup 2.*

Figure 2.13 (Apt1) and **Figure 2.14** (Apt1*) show the positional tendencies for samples taken from the first elution with the unmodified TBA (without code or primers) for the selections originating with libraries lacking and containing fluorescent tags, respectively. These samples display a very strong preference toward thymidine in positions 2 and 6, as this monomer is found in at least 60 % of all sequences for both samples in these positions. Naphthalene is strongly favoured in position 1, phenylalanine and thymidine are both favoured in position 3, C12 is favoured in position 4, C12 and phenylalanine are favoured in position 5, and guanosine and thymidine are favoured in position 7. Compared with the first elution of the first selection setup, there were only some minor differences; in setup 1, guanosine was strongly favoured in position 1 and naphthalene was favoured in position 3. **Figure 2.15** (Apt2) and **Figure 2.16** (Apt2*) show the positional tendencies for the second elutions for the non-fluorescent and fluorescent libraries using unmodified TBA in solution. The results were very similar to the previous elution, with only slight differences noticeable, such as a smaller proportion of thymidine in positions 1, 2, and 6, a larger proportion of C12 in positions 1 and 7, and a larger proportion of tryptophan in position 7. **Figure 2.17** (Thr) and **Figure 2.18** (Thr*) display the proportions of monomers at each position for samples from the final elutions with free thrombin in solution for the non-fluorescent and

fluorescent libraries. The thrombin samples were the most dissimilar, with thymidine not nearly as widespread in positions 2 and 6. Naphthalene and C12 were primarily favoured in these samples, though it is worth noting that the library tended to be biased toward these monomers in certain positions, such as position 2 with C12. **Table 2.16** presents the most common monomer per position per sample for setup 2.

Table 2.16. The most common monomer per position for each sample for the first selection experiment is presented below. The positions are labeled as before from 7 to 1 (5' to 3'). In brackets are the actual percentages at which the monomers occurred in their sample at the position indicated. Individual entries are bolded in cases where the most common monomer is at least 10% greater (in absolute difference, not percent difference) than the second most common monomer. As was the convention earlier, asterisks denoted the original library was fluorescently tagged.

Position / **Sample**	**7** **Mod (%)**	**6** **Mod (%)**	**5** **Mod (%)**	**4** **Mod (%)**	**3** **Mod (%)**	**2** **Mod (%)**	**1** **Mod (%)**
Library	C12 (19.3)	T (20.0)	C12 (19.3)	C12/G (20.8)	Nap (22.9)	**C12 (28.8)**	G (25.9)
Library*	C12 (18.1)	T (20.2)	Phe (18.6)	G (21.0)	Nap (23.0)	**C12 (29.4)**	G (25.4)
Apt1 elution	T (32.7)	**T (76.1)**	C12 (23.6)	**C12 (29.7)**	Phe (30.5)	**T (74.8)**	Nap (34.9)
Apt1* elution	T (31.2)	**T (73.3)**	Phe (23.9)	**C12 (30.4)**	T (28.7)	**T (72.2)**	Nap (26.9)
Apt2 elution	G (31.9)	**T (61.8)**	**C12 (32.5)**	**C12 (31.2)**	Phe (25.8)	**T (59.5)**	**Nap (31.5)**
Apt2* elution	Trp (26.2)	**T (61.5)**	C12 (25.1)	C12 (29.8)	T (26.0)	**T (59.3)**	Nap (26.9)
Thr elution	**C12 (29.3)**	T (23.8)	C12 (25.2)	Nap (24.3)	**Nap (34.0)**	C12 (27.1)	Nap (25.2)
Thr* elution	**C12 (29.0)**	T (21.9)	**C12 (28.0)**	Nap (29.4)	**Nap (42.9)**	**C12 (32.7)**	Nap (29.0)

2.3.4. Summary and considerations

Two different setups were used to select for modified thrombin-binding aptamers enriched from a combinatorially synthesized DNA-encoded library. The first setup used three different types of elutions to collect aptamers bound to a column containing immobilized thrombin. The first elution used free thrombin in solution to collect aptamers from the column, the second elution used the 15mer TBA, but lacking any code or primers necessary for amplification by PCR, to competitively dislodge aptamers from the column, the third elution used free fibrinogen in solution to competitively bind to thrombin and displace bound aptamers, and the fourth elution used hot water to interfere with binding interactions between thrombin and the aptamers. The second setup used only two different types of elutions, with two consecutive elutions using the 15mer TBA and one final elution with free thrombin. A fluorescent library was also tested using

the second setup under the same conditions. In all cases, stronger binding aptamers would be expected to survive until later rounds of selection and be present at higher levels, though bias can be introduced into libraries from different coupling yields during library synthesis or from inefficiencies in amplification. The sequencing data was analyzed by two main approaches – looking at the popularity of individual sequences within a sample, and comparing positional tendencies for certain monomers within a sample.

The library samples, which were taken prior to any form of selection, provided a quasi baseline to assess levels of enrichment in both individual sequences and specific positions. In the first setup, individual sequences were observed at frequencies as high as 500 reads per million, while in the second setup, the maxima were around 600 reads per million. These results, which held across different sample sizes, indicated a level that had to be exceeded in future rounds to isolate enriched aptamers. Some monomers were represented at higher levels than expected in the library, particularly with C12 at position 2. While not ideal, this issue only meant that it was important to consider when determining which monomers were enriched in future samples.

The sequence of the unmodified thrombin-binding aptamer was a member of the library and its progress through the different selections was also monitored and compared with other enriched aptamer sequences. For all three selections, this sequence was enriched with reference to the initial library in all elution samples. Most elution samples provided useful information in terms of popular sequences, except for one sample per selection in which the sequence of the unmodified TBA was the only sequence to be enriched at levels much higher than the established baseline of the library. Although this sequence was often present at levels higher than other high-ranking modified aptamers in samples that underwent selection, this signified that the selection process worked, since the known aptamer was successfully enriched from a library consisting of 6^7 sequences. This validation of the modified aptamer selection method is important for any future uses in which the goal is to discover new aptamers.

Several noteworthy sequences were identified across multiple selection experiments and samples as promising aptamers. 5’ GG **GT** GG **PheC12**T GG **PheT** GG **Nap** 3’ appeared prominently in both selection setups, ranking as the 5^{th}, 3^{rd}, and 1^{st} most popular sequence of the free

thrombin, non-amplifiable TBA, and hot water elutions, respectively, in setup 1. It also was the 2nd, 1st, and 14th most popular individual sequence in the second setup with the fluorescent library for the first TBA elution (Apt1*), the second TBA elution (Apt2*), and the free thrombin elution (Thr*). The sequence 5' GG **TrpT** GG **PheC12**T GG **TT** GG **G** 3' appeared alongside this sequence as well, ranking in 4th, 2nd, and 7th of the first, second, and fourth elutions for the first setup. This second sequence ranked well in second setup as well. 5' GG **TT** GG **PheC12**T GG **PheT** GG **G** 3' also appeared alongside the previous two sequences in the first setup, though it did not feature as prominently in the second setup. The sequence 5' GG **TT** GG **TrpG**T GG **TrpT** GG **G** 3' did rank similarly to the first two, ranking 2nd, 5th, and 2nd overall in the first, second, and fourth elutions of the first setup and 3rd and 7th in the Apt1* and Apt2* samples, though it was not highly ranked in the Thr* sample. 5' GG **SugT** GG **PheG**T GG **NapC12** GG **Nap** 3' 5' Sug T Phe G Nap C12 Nap 3' was another heavily modified sequence that appeared in several samples as a popular sequence. It was ranked just after the unmodified TBA in the fibrinogen and Thr* samples, though these samples were dominated by the unmodified TBA sequence. However, it was also one of the most popular sequences of the fluorescent library and did not appear prominently in any other samples. 5' GG **GT** GG **TrpSug**T GG **PheT** GG **Nap** 3' and 5' GG **TT** GG **GG**T GG **PheT** GG **Nap** 3' were highly positioned in the second selection setup for samples originating from the non-fluorescent library, appearing in the top ten for all three elution samples, though they did not appear in either of the other two selection experiments.

Positional trends were analyzed as well, with the most prominent enrichments observed at positions 2 and 6 of the aptamer for the thymidine monomer. While this result was not necessarily the most exciting, as it did not occur with the unnatural monomers, it did demonstrate the utility of this selection experiment. There were, however, several instances where non-nucleosidic modifications were more common than nucleosidic monomers at specific positions. Naphthalene was particularly overrepresented at positions 1 and 3, C12 was widespread in positions 4 and 5, phenylalanine was strongly represented in positions 3 and 5, and tryptophan was one of the favoured monomers in position 7.

The two approaches for identifying enrichment in aptamer libraries during selection proved complementary. Positional tendencies are useful to determine what types of monomers tend to

be favourable in selection experiments at certain positions, though they do not necessarily reveal entire sequences as monomers at different positions of the aptamer may or may not behave cooperatively in terms of binding to a target. Individual sequence popularities provide much information about the sequences that are favourably selected, though it is difficult to draw more conclusions from the sequences regarding site-specific modifications. Both methods were next used to identify monomers and sequences for individual analysis.

2.4. Aptamer synthesis and characterization

2.4.1. Choice of aptamer candidates

Based on the quantitative and qualitative results from sequencing, the aptamers in **Table 2.17** were selected as potential strong binding candidates and were synthesized without any branching unit, primers or DNA barcodes for binding affinity studies. 5' 6-carboxyfluorescein phosphoramidite was appended to the aptamers to enable fluorescence anisotropy measurements of binding affinity and serum stability measurements. The aptamers were synthesized and purified as detailed in the Experimental section. Non-nucleosidic monomers such as phenylalanine were synthesized as needed, the details of which are also in the Experimental section. For reference, the unmodified thrombin-binding aptamer sequence in the library is 5' GG TT GG TGT GG TT GG T 3', which differs from the classic 15mer TBA sequence by a terminal thymidine. The following different strategies were used to pick five aptamers for future characterization: the popularity of individual sequences, the positional tendencies across entire samples, and the prevalence of a positional monomer in popular sequences.

Table 2.17. The sequences chosen for further study are presented in the below table. Monomers that differ from the unmodified sequence (5' GG TT GG TGT GG TT GG T 3') are bolded. All aptamers were synthesized with a 5' FAM fluorescent group to enable subsequent characterization.

Sequence name	Full sequence (5' to 3')	Rationale for inclusion
Aptamer1	5' FAM GG TT GG **Phe**GT GG TT GG T 3'	Phe prominent in position 5 for many popular sequences
Aptamer2	5' FAM GG **G**T GG **PheC12**T GG **Phe**T GG **Nap** 3'	Prominent sequence in several eluted samples

Aptamer3	5′ FAM GG **Trp**T GG **PheC12**T GG TT GG **G** 3′	Prominent sequence in several eluted samples
Aptamer4	5′ FAM GG TT GG T**C12**T GG TT GG T 3′	C12 prominent in position 4 across many eluted samples
Aptamer5	5′ FAM GG TT GG TGT GG TT GG **Nap** 3′	Nap prominent in position 1 across many eluted samples

Aptamer2 and Aptamer3 were based on the popularity of individual sequences, with both appearing frequently in the samples that had undergone selection. In the first selection setup, Aptamer2 was the 5th most popular sequence of the free thrombin sample (**Table 2.3**), the 3rd most popular in the non-amplifiable TBA sample (**Table 2.4)**, and the most popular in the hot water sample (**Table 2.5**). In the same respective samples, Aptamer3 was the 4th, 2nd, and 7th most popular sequence. In the second selection setup, for the fluorescent library, Aptamer2 was also the 2nd most popular sequence in the first non-amplifiable TBA elution (

Table 2.11), the most popular sequence in the second TBA elution (**Table 2.13**), and the 14th most popular sequence in the final elution with free thrombin (**Table 2.15**). Aptamer3 was the 4th and 5th popular individual sequence for the first and second elutions with TBA in the same setup. Aptamer2 and Aptamer3 were chosen over other prominent sequences such as 5′ Sug T Phe G Nap C12 Nap 3′, which was ruled out as it appeared prominently solely in the library samples and the elutions in which the unmodified aptamer sequence was substantially overrepresented. Other sequences such as 5′ T T Phe C12 Phe T G 3′ and 5′ T T Trp G Trp T G 3′ were prominent in the same samples as Aptamer2 and Aptamer3, though they were ranked lower in the later elutions.

Aptamer1, which features a single phenylalanine modification, was based on the prominence of the phenylalanine monomer in position 5 in several popular sequences. In the first setup, phenylalanine was the most common modification at position 5 in the first ten sequences of the samples from elutions with free thrombin, non-amplifiable TBA, and hot water. It was also the most common modification across all sequences for the TBA elution sample in setup 1 and for the first TBA elution in setup 2 for the fluorescent library. Aptamer4 and Aptamer5 also featured one single site modification and were chosen based only on positional tendencies across entire

samples. Naphthalene was a monomer that for all eluted samples in the second selection setup appeared most frequently in position 1. For the first selection setup, naphthalene was the most popular monomer in two of four elution samples and the second most popular monomer (behind guanosine) in the other two samples. In all three library samples, naphthalene was ranked as the third or fourth most common modification in position 1, behind thymidine, guanosine, and phenylalanine. This indicated that naphthalene in position 1 was beneficial to binding thrombin, as it was enriched with respect to the initial library at a higher rate than the other 5 monomer options at that position (G, T, Alk, C12, and Phe). Aptamer5 was therefore designed with a single naphthalene modification at position 1 and the rest of the sequence unmodified. Similarly, Aptamer4 was designed with a single C12 modification at position 4 based on its presence in eluted samples. C12 and G were the most abundant monomers at position 4 for all library samples, occurring at similar levels with each present in about 20 % of all sequences for the library samples. For the first setup, the proportion of C12 in position 4 across all sequences increased, relative to the library, for all the eluted samples. Contrastingly, guanosine decreased or remained the same in all eluted samples relative to the initial library. Apart from the Thr and Thr* samples, C12 was the most common monomer in position 4 for the eluted samples of the second setup, for both fluorescent and non-fluorescent libraries. The structures of the non-nucleosidic modifications present in the five candidates synthesized individually are shown in **Figure 2.19**.

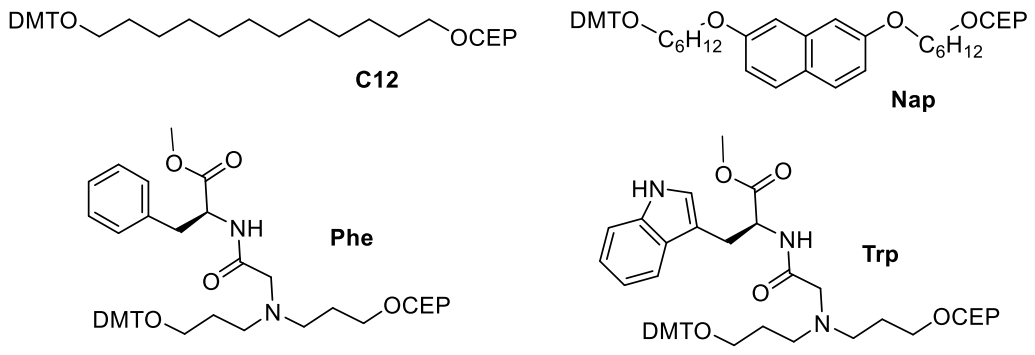


Figure 2.19. The structures of the non-nucleosidic modifications included in sequences selected for further study. C12 stands for a 12-carbon unit, Nap stands for naphthalene, Phe stands for phenylalanine, and Trp stands for tryptophan. DMT refers to 4,4'-dimethoxytrityl and CEP refers to β-cyanoethyl-N,N'-diisopropylamino-phosphoramidite.

2.4.2. Evaluation of candidates

Serum stability and binding affinity were two properties that were assessed for the aptamer candidates. Daniel Saliba measured serum stability in terms of half-life using denaturing gel electrophoresis (**Figure 4.4**). Serhii Hirka measured binding affinity for thrombin using fluorescence polarization. In this measurement technique, fluorophores (or molecules tagged with fluorophores) are irradiated with linearly polarized light and the degree of polarization is measured based on the intensity of fluorescence emission parallel and perpendicular to the light source.[279] The polarization of a fluorophore is very dependent on its environment, so the binding affinity of a fluorescently labeled ligand can be measured based on the associated change in polarization of a fluorophore, which is affected by rotational diffusion (**Figure 2.20**). Fluorescein is widely used as a fluorophore in binding studies as they are spherical in shape, they are easily conjugated to drug candidates, and feature sufficiently long lifetimes for measurements of rotation. Dissociation constants were calculated from binding isotherms, which were obtained by plotting the change in depolarization with respect to changes in thrombin concentration. This method had been previously used to assess what binding affinities were observable for DNA-encoded aptamers. As was expected, a loss in binding affinity was observed for aptamers that were attached to the branch and code construct, as the thrombin-binding aptamer experienced a roughly threefold decrease in its dissociation constant (K_D) when modified at the 3' end with the branch and code.[272]

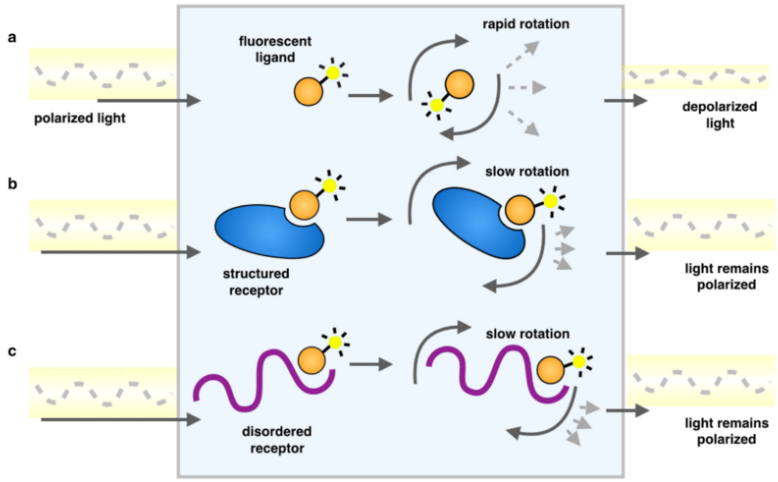


Figure 2.20. Above is a diagram of the general principles of fluorescence polarization assays.[280] The extent of depolarization of plane-polarized light by a fluorophore is very dependent on its environment. When not bound to a target, the fluorophore can rotate rapidly, resulting in mostly depolarized light. When bound to a target, the fluorophore is rotationally restricted as its

*effective size increases and so the light remains mostly polarized. Reproduced with permission from reference **280**. Copyright 2017 Heller et al.*

The binding affinities of the aptamer candidates, along with the serum stability half-lives, are presented in **Table 2.18**. A scrambled sequence was used as a negative control and it did not display any binding activity with up to 1 µM of thrombin added. The unmodified thrombin-binding aptamer sequence was also included as a reference.

Table 2.18. The binding affinities (in units of nanomolar) and serum half lives (in units of hours) of the unmodified TBA and the modified aptamer candidates are presented below. The condensed sequences list only the monomers for positions at which the library was modified.

Compound	Condensed Sequence (5' to 3')							(K_D) Binding affinity (nM)	Serum half life (hr)
TBA	T	T	T	G	T	T	T	11 ± 8	< 0.3
Aptamer1	T	T	Phe	G	T	T	T	26 ± 4	< 0.3
Aptamer2	G	T	Phe	C12	Phe	T	Nap	10 ± 5	5.0
Aptamer3	Trp	T	Phe	C12	T	T	G	6 ± 4	< 0.3
Aptamer4	T	T	T	G	T	T	Nap	37 ± 5	4.5
Aptamer5	T	T	T	C12	T	T	T	24 ± 4	< 0.3

Aptamer2 and Aptamer4 were the only aptamers that displayed an improved serum stability, with half-lives near 5 hours compared to the other sequences that were very quickly broken down by the nucleases present. This suggested that the naphthalene modification at the 3' end provided additional nuclease resistance. 3' nucleases tend to be more active and modifications at the 3' end for aptamers have often been successful at improving nuclease stability. These were the aptamers to feature a modification at the 3' end, so it in the future it would be interesting to see if other non-nucleosidic modifications could also confer added nuclease resistance without losses in binding affinity. Based on the results of Aptamer5 and Aptamer3, no added nuclease resistance was observed with C12 in position 4, phenylalanine in position 5, or tryptophan in position 7.

In terms of binding affinity, two of the aptamer candidates had dissociation constants as low or even lower than the unmodified aptamer, which had an experimental binding affinity of 11 nM. Aptamer2, featuring four non-nucleosidic monomers, had a binding affinity of 10 nM, while Aptamer3, featuring 3 non-nucleosidic monomers, had a binding affinity of 6 nM. These were

promising results, though we cannot conclusively say that Aptamer3 has a better binding affinity than the unmodified sequence since the two results are within each other's range of error. Nevertheless, these results are quite interesting as they provide modified aptamers that can bind at least as well to thrombin as the unmodified aptamer, if not better. The other sequences with just one modification suffered from slightly worse binding affinities of 24, 26, and 37 nM, for the C12, phenylalanine, and naphthalene modified sequences. The single site modifications did result in slightly worse binding affinities, particularly for Aptamer4 with Naphthalene. This suggests that while the modification improved nuclease resistance, it worsened binding affinity, likely due to less favourable interactions with thrombin, steric hindrance, or poorer stacking with the aptamer itself. It is also interesting to note that the phenylalanine in position 5 resulted in a worse binding affinity on its own (in Aptamer1) but it was present at position 5 of the two sequences that bound thrombin at least as well as the unmodified sequence (Aptamer2 and Aptamer3). Similarly, C12 in position 4 was detrimental to the binding affinity of Aptamer5, on its own, but its presence in Aptamer2 and Aptamer3 was beneficial. It may be that the two monomers are needed together to effectively bind thrombin. One possible scenario is that the added flexibility of the C12 monomer is sufficient to enable favourable interactions between phenylalanine and thrombin or even the aptamer itself.

Positions 2, 3, 6, and 7 were the TT loops of the aptamer that have been shown to interact directly with thrombin's fibrinogen-binding site, which contains several positively charged and basic residues such as arginine, lysine, and tyrosine.[12] Arginine and lysine consist of aliphatic chains with basic amine residues, so their interactions with monomers at these positions are likely limited to hydrogen-bonding and ionic interactions, which makes nucleotides such as thymidine effective at these positions, as was seen throughout the various selections performed for the DNA-encoded libraries. Tyrosine features a phenol unit that can form hydrogen bonds and interact with other aromatic molecules via π-π stacking. Tryptophan and phenylalanine were often found in the samples that underwent selection in positions 7 and 3, respectively. Both monomers contain aromatic residues that could form stabilizing interactions with tyrosine residues. Aptamer2 and Aptamer 3 featured these modifications, and both were able to bind to thrombin as well as the unmodified aptamer. These two modifications could also prove beneficial

for membrane penetration as they contain a nitrogen in their backbone that in physiological conditions would have a positive charge, as the negatively charged backbone of normal oligonucleotides hinders their ability to pass through lipid bilayers.

Positions 4 and 5 were part of the TGT loop that is not directly involved in binding thrombin, though it may bind to another site of thrombin, the heparin-binding exosite. The TGT loop is important for the stability of the quadruplex and the aptamer itself. In the literature, modifications such as pyrene[22] and aromatic amino acid-modified nucleotides[198] have engaged in stabilizing stacking interactions with the G-quadruplex of the TBA. It is possible that aromatic non-nucleosidic monomers like phenylalanine may stack well with the nitrogenous bases of the quadruplex. Other modifications in this loop have focused on the flexibility of monomers, such as unlocked nucleic acids (UNAs)[146] or even small spacers,[21] which have increased the thermal stability of the aptamer and improved its anticoagulant activity. C12 features a much less rigid structure than the aromatic rings of phenylalanine and tryptophan, and so it may add flexibility to the TGT loop and enable better formation of the quadruplex necessary to binding thrombin.

Two of the multiply modified aptamers identified from sequencing results were found to bind thrombin better or at least as well as the unmodified aptamer. Singly modified aptamers were less successful, though they still featured binding affinities on the order of the unmodified aptamer. Serum stability experiments revealed additional stability in the case of aptamers containing the naphthalene monomer at the 3’ end. Altogether, these selection experiments produced an aptamer that was more stable than the unmodified sequence in serum and could bind to thrombin with comparable or better affinity.

3. Conclusion

3.1. Conclusion

While aptamers present several key advantages over antibodies, namely their ease and cost of synthesis and selection, they are still plagued by deficiencies in physiological stability and low binding affinities. Their stability toward nuclease degradation has been effectively addressed with modifications that make aptamers worse substrates for endo and exonucleases, such as modifications to the ribose unit and the phosphodiester backbone. Large end groups like polyethylene glycol have also improved the circulation times of aptamers *in vivo*, as unmodified aptamers are susceptible to kidney excretion due to their small size. While aptamers have been used to target a range of molecules, from ions to whole cells, they often suffer from lower binding affinities, particularly with negatively or non-charged targets. This in large part due to the smaller chemical space available to nucleic acids compared with amino acids. The four nucleotides of DNA display much less variety in shape and polarity than the twenty natural amino acids. This thesis aimed to demonstrate this limitation of aptamer libraries, highlight some solutions that have been explored in the literature, such as SOMAmers (slow off-rate modified aptamers), and present a new method for selecting modified aptamers that probe a wider chemical space. A DNA-encoded library of modified aptamers, designed by a former lab member, was used to select for aptamers that bound thrombin. Non-nucleosidic modifications were combinatorially introduced at 7 positions outside of the structural motif of the thrombin-binding aptamer. Multiple single-round selection experiments were conducted, with samples taken from different elutions. Next-generation sequencing was used to identify enriched sequences and positional trends across selected samples. Five prominent aptamer candidates were chosen for individual synthesis and characterization. Two multiply modified aptamers were found to bind thrombin with an equal or better affinity than the unmodified aptamer, highlighting the ability of the single-round selection experiment to identify high affinity modified aptamers. The singly modified aptamers, chosen based on positional trends, had slightly worse binding affinities than the unmodified aptamer. Serum stability was also compared across the aptamer candidates and

revealed that the sequences featuring a 3' naphthalene modification were much more stable to nuclease degradation than the other unmodified aptamers. In particular, the sequence 5' GG GT GG PheC12T GG PheT GG Nap 3' featured a very similar binding affinity to the unmodified aptamer but with much added nuclease stability.

3.2. Relevance of work

This work is particularly relevant to the selection of modified aptamers, as it presents the successful implementation of a new method for identifying aptamers from a library of oligonucleotides with non-nucleosidic modifications in a single round of selection. Most methods that introduce modifications into DNA or RNA for aptamer selections do not expand the chemical space of aptamers, as one of the four natural nucleotides is usually replaced with a modified nucleotide, such as in the case of click-SELEX or SOMAmers. These methods allow libraries to be amplified and used in future SELEX cycles; however, the library is still limited to four building blocks. Other attempts to increase the number of building blocks in a library require a predetermined placement of nucleosidic modifications to be able to identify whether a nucleotide is modified via sequencing, such as in the case of libraries containing both PS and PO linkages. To include non-nucleosidic modifications that cannot base pair to normal DNA, as required for PCR amplification and sequencing, current methodologies are limited to post-SELEX methods or high-throughput screening in which sequences are sequenced and studied individually. There have been successful implementations of SELEX with non-nucleosidic modifications, such as AEGIS-SELEX and Looper-SELEX. AEGIS-SELEX introduces a synthetic base pair orthogonal to the base pairs of DNA that relies on the same type of hydrogen-bonding motifs but can be amplified simultaneously with normal nucleic acids. This modification does not introduce any drastically different structural motifs, besides a nitro group, though it does expand the number of building blocks that can be used concurrently in one library. Similarly, hydrophobic base pairs have been used that are orthogonal to the base pairs of DNA and introduce more structural diversity into aptamers, though sequencing requires additional steps to determine where modifications were present in the aptamers. Looper-SELEX uses a series of 5-base codons, each appended with one unique modification, to generate libraries with as many as 16 different

pentameric building blocks, though the placement of modifications within the library is still quite limited. Our method of DNA-encoded libraries allows for the use of, in theory, any monomer that can be transformed into a phosphoramidite and is compatible with the conditions of solid-phase synthesis. Non-nucleosidic modifications with unique 3mer barcodes can be introduced into libraries combinatorially using a branched aptamer-code construct and two orthogonal chemistries. These synthetic aptamers, that are not amplifiable or sequenceable, can be identified by amplifying and sequencing the distinct DNA barcodes. Next-generation sequencing methods were shown to be sufficient to identify enriched sequences and positional trends, which resulted in the discovery of several high affinity modified aptamers. While it was demonstrated successfully only for the thrombin-binding aptamer, this method could be used to select aptamers for a wide range of targets.

3.3. Next steps

There are several possible next steps to follow the work presented in this thesis. One possible avenue is the design of another library for thrombin using different monomers. With the versatile platforms available for conjugation of functional groups, it would be feasible to include other non-nucleosidic modifications in the combinatorial library, such as modifications used in SOMAmers. Several aromatic groups such as phenylalanine, tryptophan, and naphthalene were favourable in the library, so other modifications that feature aromatic rings but in different structures could be implemented in the library. A benzene monomer is currently in the preliminary stages of synthesis using one of the synthetic platforms for phosphoramidites and its use in the library could be beneficial as an alternative to phenylalanine. In addition, there are many commercially available modified phosphoramidites compatible with SPS and their inclusion in another iteration of a library could prove beneficial. If another library were synthesized for the TBA, its focus could also be narrower as certain positions of the aptamer were shown to overwhelmingly prefer a nucleosidic monomer, so there would be no need to modify these positions. Some bias was also observed at certain positions of the library, so future studies could also standardize the splitting and pooling steps to minimize product loss or unequal divisions of aliquots.

Another element of aptamers to consider is their activity, which is related to both their ability to bind their target and their stability in the environment of the target. The TBA binds thrombin and inhibits its function as a member of the clotting cascade, so a measure of activity can be performed by evaluating how long it takes to form a clot from fibrinogen. This type of analysis would be interesting to look at for the aptamers highlighted in this thesis, as it is possible that one of the modified aptamers may prove more active. Inhibitory activity for aptamers relates to the stability of the aptamer as well as its binding affinity, so it possible that aptamers with poorer binding affinities, such as the singly modified aptamers, may offer additional stability and have comparable activities to the unmodified aptamer. In this study, modified TBAs were tested in the presence of thrombin in thrombin-binding buffer, which does not contain the enzymes present *in vivo* responsible for catalyzing the degradation of oligonucleotides. The inhibitory activity of modified aptamers in the presence of nucleases could also be tested to see what combinations of modifications most effectively bind to thrombin in more biologically relevant environments. The thermodynamic stability of the modified aptamers could also be investigated to determine what effect the modifications have on the structure or stability of the G-quadruplex. Aromatic groups may π-stack with the aromatic bases in the guanine tetrads and stabilize the structure. Guanine quadruplexes have distinct signals in circular dichroism (CD) spectra, so CD spectroscopy could be one way to examine the effect of the modifications on the structure of the G-quadruplexes. The thermal stability of the G-quadruplexes could also be investigated in the presence of thrombin. Computer models are also very adept at investigating thermodynamic properties and could be used for the study of individual modified aptamers following selection, though modeling the thrombin interface with the modified aptamer may prove to be computationally demanding. Sequencing represents another area to consider improvements, as it is possible to observe other trends or patterns via clustering algorithms. While identifying conserved sequence motifs may be less useful in the case of aptamers modified only at certain positions of the sequence, it would be interesting to determine if certain monomers are more likely to occur in tandem with other monomers and act cooperatively.

The methodology used in this thesis for the identification of DNA-encoded modified aptamers to thrombin is versatile and could be used for several other targets. In fact, it may prove more

successful in the case of targets for which aptamers have typically suffered from low binding affinities, as many of the non-nucleosidic modifications introduced could be used to target hydrophobic or uncharged molecules that have proved difficult for conventional aptamers. This was observed in the case of SOMAmers, where several aromatic and aliphatic groups appended to a modified uridine introduced in aptamer libraries proved beneficial for difficult targets, such as platelet-derived growth factor B.[281] Several of the monomers also contain a positively charged amine, which could prove useful for cellular applications involving membrane penetration. The constraint of our approach is that the large libraries necessary to probe a variety of modified monomers in many positions of an aptamer may not be well accommodated by a single round of selection. The library in this thesis consisted of 279,936 unique sequences with six different monomers at seven different positions of a 16mer aptamer. Libraries could easily be orders of magnitude larger for longer aptamers modified at more positions or with more monomers available per position. There have been several improvements to aptamer selection methods that have enabled sufficient enrichment of large libraries using just one round of amplification, such as NECEEM-SELEX, a type of non-SELEX, which uses the non-equilibrium capillary electrophoresis of equilibrium mixtures to repeatedly partition a library of aptamers in the presence of the target. One variant of single-round SELEX has even used nuclease digestion to increase the partitioning efficiency, though this would be problematic for DNA-encoded modified aptamers, or DNA-encoded libraries in general, as the DNA barcode not involved in binding would be susceptible to nuclease degradation regardless of target binding. It would be possible to use a technique like NECEEM-partitioning or bead-based separation in order to achieve better enrichment of modified aptamers, which may be necessary for testing larger libraries of DNA-encoded modified aptamers. Regardless, the results presented here represent exciting progress in the selection of modified aptamers and the methodology presented may allow for the future discovery of diversely modified aptamers, as was demonstrated here for the selection of modified thrombin-binding aptamers.

4. Experimental

4.1. Chemicals

Unless specified, chemicals were used as is without further purification. Solvents such as acetone, acetonitrile, dichloromethane, ethanol, and methanol, were purchased from Fisher Scientific. DMT-protected nucleosidic phosphoramidites (thymidine, and isobutyryl-protected guanosine), dry packs, activator solution, and 1000 Å universal 1 µmol controlled pore glass columns were purchased from Bioautomation. MyTaq HS Red Mix kits were purchased from Froggabio for hot-start PCR and QIAquick PCR purification kits were purchased from Qiagen. β-cyanoethyl N,N'-diisopropylamino phosphonamidic-chloride (CEP-Cl) was purchased from ChemGenes. Sephadex G-25 and 5'-6-carboxyfluorescein phosphoramidite were purchased from Glen Research. Deuterated chloroform was purchased from Cambridge Isotope Laboratories and stored on molecular sieves to prevent possible DMT deprotection. Ammonium hydroxide and acrylamide/bis-acrylamide were purchased from Bioshop Canada Inc. The remaining reagents were purchased from Sigma-Aldrich, including Thrombin from human plasma. TEAA (triethylammonium acetate) buffer was prepared with 50 mM TEA (triethylammonium) adjusted to a pH of 8.0 using glacial acetic acid. TBE (Tris-borate-EDTA) buffer was prepared with 90 mM Tris (tris(hydroxymethylaminomethane)), 90 mM boric acid, and 1.1 mM EDTA (ethylenediaminetetraacetic acid) and adjusted to a pH of 8.0.

4.2. Instrumentation

Automated solid-phase synthesis of DNA was performed on a MerMade MM6 synthesizer from Bioautomation. Oxygen and moisture-sensitive off-column couplings were performed in a Vacuum Atmospheres Co. glove box. DNA quantification by UV absorption was performed using a NanoDrop Lite spectrophotometer from Thermo Scientific. PAGE (polyacrylamide gel electrophoresis) was performed for purification with Hoefer SE 600 electrophoresis systems and for quantitative analysis using Bio-Rad Mini-PROTEAN units. Gel images were captured using a ChemiDoc MP System from Bio-Rad Laboratories. Eppendorf Mastercycler 96-well thermocycler and Bio-Rad T100TM thermal cycler were used for PCR. Fluorescence polarization was performed

with a Synergy H4 Hybrid Multi-Mode Microplate Reader from BioTek. Dry solvents were obtained using an Innovation Technology device. Column chromatography was performed using a CombiFlash Rf system from Teledyne ISCO. NMR spectra were taken with a Bruker 400 or 500 MHz or a Varian 300 or 400 MHz instrument using deuterated chloroform or dimethylsulfoxide as internal lock solvents. Low resolution mass spectrometry (MS) was performed with an electron-spray ionization (ESI) ion trap MS on a Finnigan LCQ Duo device.

4.3. Analysis of sequencing results

Following selection, the eluted samples were individually amplified using hot start Taq DNA polymerase under previously optimized conditions.[272] Indices were appended to the amplicons to enable binding to flow cells for Illumina MiSeq high-throughput sequencing. Sequencing was performed by the Genome Quebec Innovation Centre, in Montreal, Canada. Two read files in fastq format (**Figure 4.1**) were obtained per sample, with each corresponding to one of the two complementary strands. Fastq files are text files that for every unique sequence provide an identifier, a full DNA sequence, and quality scores for each base. The scores are called Phred quality scores and use ASCII (American standard code for information interchange) characters to indicate the probability of a specific base being incorrectly sequenced. The higher the score, the lower the probability of the base being wrong.[282] Sequences that appear in multiple copies prior to the indexing step are each prescribed a unique index, so in fastq files they appear as distinct sequences containing the same 21mer DNA barcode.

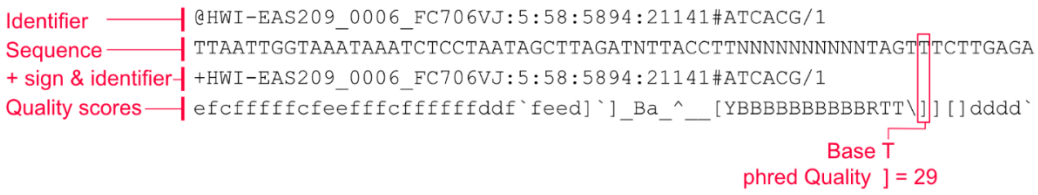


*Figure 4.1. An example of a fastq file is shown above.[282] Fastq files are often the output format for high-throughput sequencing reads. It is a text-based format that contains three types of information per sequence read. Line 1 provides a sequence identifier and often run information, line 2 provides the full sequence, and line 4 provides the quality scores for each base of the sequence. The quality scores are encoded as ASCII characters and are called Phred quality scores. Line 3 is usually simply a plus sign with the sequence identifier repeated. Reproduced with permission from reference **282**. Copyright 2020 Akalin.*

Cutadapt is a freely available program that can trim adapter sequences from high-throughput sequencing reads and was used to trim the full sequences to just the 21mer DNA barcode (**Figure**

4.2).[275] Cutadapt was used to remove all bases prior to and including the forward primer and then was used to remove all bases following and including the reverse primer. Any remaining sequences that were not exactly 21 bases long were discarded.

```
subprocess.call(["cutadapt",
"-g", "CGTCGAGGCCC",
"--discard-untrimmed",
"input.fastq.gz",
"-o", "cut1.fastq"])

subprocess.call(["cutadapt",
"-a", "AGGATACACGTCACGCC",
"--discard-untrimmed",
"cut1.fastq",
"-o", "cut2.fastq"])

subprocess.call(["cutadapt",
"-m", "21",
"-M", "21",
"cut2.fastq",
"-o", "21.fastq"])
```

Figure 4.2. Above is a snapshot of the code we used to trim the sequences in the fastq files. The first paragraph uses cutadapt to search for the sequence specified (the 5' adapter) starting from the 5' end and to keep only the bases to the right of the specified sequence. This creates a new fastq file (cut1.fastq) in which the 5' adapter and any preceding bases have been removed. The second paragraph performs the same function but for the 3' end, removing the 3' adapter and any bases following it (outputting cut2.fastq). These two steps should leave only the 21mer DNA barcode. The final paragraph discards any sequences that are longer or shorter than 21 bases, leaving only sequences with exactly 21 bases in the output file 21.fastq.

Cutadapt was then used indirectly to find out how often each codon occurred in a position for a given sample. The 21mer sequences were trimmed from the 5' end using six different codons and the total number of sequences trimmed was tabulated for each codon. This process was repeated for all seven codons using progressively shorter input sequences. To determine the most popular sequences, the count function of FASTAptamer, a program used for analyzing large sequencing files for libraries of oligonucleotides, was used for the trimmed sequences containing the entire DNA barcode. This function counts the number of a times a unique sequence occurs and outputs a fasta format file listing the sequences in order of their abundance and frequency within the samples (**Figure 4.3**). The most popular sequences were decoded to determine the identity of the monomers.

```
>1-799-20714.51
CTTCTTCGTCGTGTACTTTGA
>2-613-15892.36
CGTCTTACACAGGTACTTTGA
>3-464-12029.45
CTTCTTCTTTGCTAACTTTGA
>4-448-11614.64
ACACTTTGCTGCCTTCTTCGT
>5-431-11173.91
CGTCTTTGCTGCCTTCTTTGA
>6-408-10577.62
CGTCTTCTTTGACTTCTTTGA
>7-389-10085.04
ACACTTGTACGTCTTCTTTGA
>8-380-9851.71
CGTCTTTGCTGCACACTTCTT
>9-362-9385.05
CTTCTTCTTTGCGTACTTCGT
>10-326-8451.73
CTTCTTTGCGACGTACTTCGT
>11-308-7985.07
TGCCTTCAGTGCCTTCTTTGA
>12-306-7933.22
CGTCTTTGCCGTGTACTTTGA
```

Monomer	Codon
Phe	GTA
His	GTG
Trp	ACA
C12	TGC
Nap	TGA
Ant	TAA
Sug	CAG
Alk	GAC
A	CAT
T	CTT
C	CCT
G	CGT

Figure 4.3. Above on the left is an example of the output of FASTAptamer's count function when applied to the 21.fastq file. The format ">#-#-#" refers, from left to right, to the rank, abundance, and frequency of the sequence across the entire sample. On the right is the genetic code, with a unique codon referring to a distinct monomer in the library. For example, the first sequence is T T G G Phe T Nap, which refers to the condensed sequence of modified positions in the thrombin-binding aptamer.

4.4. Synthesis of individual aptamers

Automated solid-phase synthesis was performed on a MerMade MM6 synthesizer from Bioautomation using DMT-protected nucleotide phosphoramidites from ChemGenes and 1 µmol 1000 Å universal synthesis columns from Bioautomation. Off-column couplings were performed manually under argon for any non-nucleosidic monomers, as well as for 5' 6-FAM (6-carboxyfluorescein). 20 µmol of the monomers per coupling were dissolved in 200 µL of DCM and manually mixed with 200 uL of activator solution, with the subsequent oxidation and capping steps performed on the DNA synthesizer and the next deprotection observed for visual assessment of coupling efficiency. For any amino acid-based monomers with carboxylic acid groups protected as methyl esters, an additional deprotection step was required prior to conventional deprotection and cleavage from the solid support using 28% aqueous ammonia for 16-18 hours at 65 °C. This additional step was performed in 1:3 by volume of tertbutylamine in water for 6 hours at 65 °C to deprotect methyl esters into carboxylates. Following removal of the CPG by filtration, the aptamers were purified conventionally by denaturing PAGE using 20 % polyacrylamide gel in TBE with 7M urea. Fluorescent bands were excised and the DNA was

extracted by centrifugation using autoclaved milliQ water. The aptamers were salt purified with sephadex and quantified by UV absorption. Aptamers were stored in autoclaved water at -20 °C under aluminum foil.

4.5. Serum stability of modified aptamers

PAGE-purified aptamers were treated with serum (Dulbecco's modified Eagle's medium with fetal bovine serum), using 10 % by volume DNA and 90 % by volume serum. The mixtures were incubated at 37 °C and equivolume aliquots were taken at different time points and frozen. Once all samples were collected, they were defrosted and treated with proteinase K to denature the proteins in serum. The samples were loaded on a denaturing PAGE gel for quantification of the bands and determination of half-lives using Gel Red and Fluorescein channels. See **Figure 4.4** for the serum stability measurements taken for the aptamer candidates.

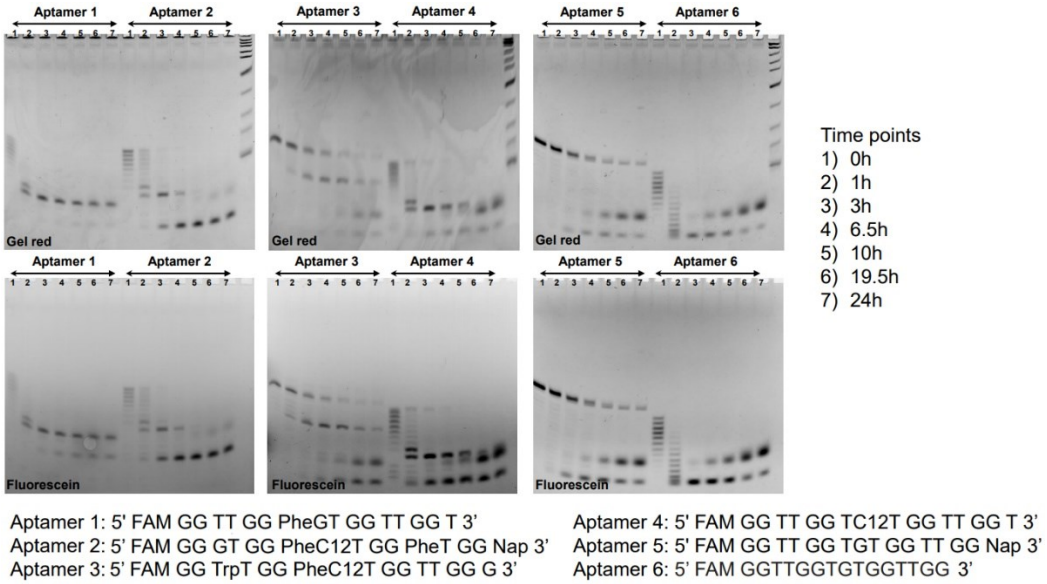


Figure 4.4. Serum stability measurements for the aptamer candidates are shown above. In this case, Aptamer6 is just the unmodified thrombin-binding aptamer. Gel red and fluorescein channels are shown in the top and bottom rows, respectively. Aptamer3 and Aptamer5 are the only sequences that display a half-life greater than an hour.

4.6. Binding affinity of modified aptamers

The fluorescence polarization assay was based on previous work[283, 284] with several modifications. All model compounds were synthesized as described above. Each model compound was dissolved in thrombin binding buffer (20 mM Tris, 140 mM NaCl, 20 mM KCl, 1 mM CaCl2, 1 mM

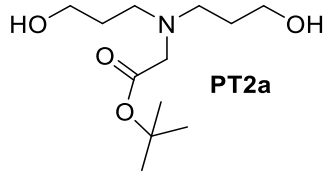
$MgCl_2$, 0.05% Triton-X 100) to a concentration of 10 nM and stored in freezer in the dark until further use. For the assay, the model compounds were thawed, then heated at 95°C for 4 minutes, then cooled at 4°C for 15 minutes. Thrombin (T6684, Sigma Aldrich) was prepared fresh each time by dissolving the entire vial with 200 µL thrombin binding buffer, then measuring the absorbance on the NanoDrop Lite spectrophotometer, and calculating the concentration using E1%/280, 18.3. Dilutions of thrombin from 400 nM to 0.39 nM were made in thrombin binding buffer. Next, 50 µL of each model compound and 50 µL of each thrombin dilution were mixed separately in a 96-well black/conical bottom microplate. The plate was shaken on a Cytation5 plate reader for 20 minutes and incubated at room temperature for 30 minutes. Finally, polarization was measured for each sample using a Xenon Flash light source with green FP filter set, with relevant filter parameters as follows: excitation at 485 nm, emission at 528 nm, and dichroic mirror at 510 nm. The change in polarization, defined as the different between the polarization value measured in the absence of thrombin and the presence, was plotted against the concentration of thrombin added. A binding isotherm was fit using GraphPad Prism 6 using which one site binding (hyperbola) fit. K_D values were reported as the mean and standard deviation of three independent experiments.

4.7. Small molecule synthesis

4.7.1. Platform 2 and phenylalanine phosphoramidite synthesis

Molecules were synthesized as detailed below according to previous work.[272] ^{1}H NMR spectra were taken to confirm the identity of the species synthesized, for which both ^{1}H and ^{13}C NMR spectra exist.

4.7.1.1. PT2a (tert-butyl bis(3-hydroxypropyl)glycinate)



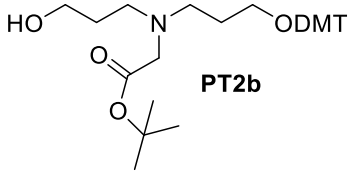
PT2a

To a solution of glycine tert-butyl ester hydrochloride (2.57 g, 15.3 mmol) in acetonitrile were added $KHCO_3$ (4.60 g, 45.9 mmol, 3 equiv.), KI (254 mg, 1.53 mmol, 0.1 equiv.) and 3-bromo-1-propanol (2.77 mL, 30.6 mmol, 2 equiv.) under stirring. Temperature was raised until a reflux of

acetonitrile was reached and left under stirring overnight. The reaction was followed by TLC. After completion, solvent was evaporated under reduced pressure (40 °C). Product was extracted twice with ethyl acetate from 10 % Na_2CO_3 aqueous solution, washed once with a 10 % Na_2CO_3 aqueous solution, dried over $MgSO_4$, filtered, and solvent was evaporated under reduced pressure (40 °C). Compound obtained was a pale yellow oil and was used as is for the next step (3.4 g, 90 %).

^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) 3.74 (m, 4H), 3.16 (s, 2H), 2.62 (m, 4H), 1.73 (m, 4H), 1.44 (s, 9H).

4.7.1.2. PT2b (tert-butyl *N*-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-*N*-(3-hydroxypropyl)glycinate)

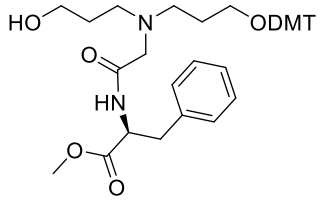


PT2b

13.8 mmol of PT2a and 5.57 mL of TEA (41.3 mmol, 3 equiv.) were dissolved in 120 mL of dry DCM in a 500 mL dry round bottom flask. DMT chloride (4.67 g, 13.8 mmol, 1 equiv.) was dissolved in 80 mL of dry DCM and added dropwise to the reaction mixture at 0 °C under vigorous stirring. After the chloride addition, the reaction was allowed to warm up to room temperature and left under stirring for 2.5 hours. The product was extracted twice with DCM from a 10 % Na_2CO_3 aqueous solution, washed once with a 10 % Na_2CO_3 aqueous solution, dried over $MgSO_4$, filtered, and solvent was evaporated under reduced pressure (40 °C). The crude product was loaded on celite and purified using the CombiFlash system with a 220 g SiO_2 Gold column. Hexanes/TEA (10:1) and ethyl acetate were used in a gradient from 0 to 35 % EtOAc (~25 CV). A clear yellow oil was isolated (0.80 g, 10 %). Platform 2b was more stable than Platform 2, which featured a carboxylate group instead of a t-butyl ester, so it was stored long term as PT2b.

^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) 7.42 (d, *J=8* Hz, 2H), 7.28 (m, 5H), 7.19 (m, 1H), 6.81 (d, *J*=8.5 Hz, 4H), 3.78 (s, 6H), 3.40 (br. s, 1H), 3.19 (s, 2H), 3.13, (m, 1H), 3.06 (m, 1H), 2.69-2.60 (m, 4H), 1.78 (m, 2H), 1.67 (d, *J=7* Hz, 2H), 1.49-1.41 (m, 9H).

4.7.1.3. Methyl *N*-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-*N*-(3-hydroxypropyl)glycyl-L-phenylalaninate



Compound PT2b (80 mg, 0.15 mmol) was dissolved in about 2 mL of methanol. To the mixture was added 5 mL of 0.4M NaOH in MeOH/water 4:1. The reaction mixture was left under stirring for 3h at 65 °C. The reaction was monitored by TLC. When the higher mobility spot disappeared, methanol was evaporated under reduced pressure (60 °C) until a precipitate appeared but a small amount of water remained. DCM was added to the mixture and 2 equivalents of tetrabutylammonium chloride (85 mg, 0.31 mmol) were added. Two extractions with DCM from saturated Na_2CO_3 solution and one washing with 10 % Na_2CO_3 solution were performed. The organic layers were dried over $MgSO_4$, filtered, and the solvent was removed under reduced pressure (40 °C). The salt (PT2) was suspended in DMF (2 mL) and 1-hydroxybenzotriazole (31 mg, 0.20 mmol, 1.3 eq.) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (38 mg, 0.20 mmol, 1.3 eq.) were added in sequence. The solution was stirred for 5 minutes until complete dissolution of EDC-Cl. Phenylalanine methyl ester hydrochloride (35 mg, 0.16 mmol, 1.05 eq.) and triethylamine (0.11 mL, 0.77 mmol, 5 eq.) were added to the flask. The mixture was left under vigorous stirring overnight. The solvent was evaporated under reduced pressure (60 °C) by coevaporation with toluene. The crude product was loaded on celite and purified using the combiFlash system with a SiO_2 "Gold" column pretreated with 0.1 % TEA in DCM. DCM/TEA (100:0.1) and DCM/Methanol/TEA (90:10:0.1) were used in a gradient. A pale yellow solid was isolated but analysis with NMR indicated the product was not pure and further difficulties were encountered with subsequent filtrations. The reaction was performed by Dr. Violeta Toader with greater success and the DMT-protected phenylalaninate was converted into a phosphoramidite. Proton NMR data presented below is from a previous work.[272]

^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) 7.60 (d, J=8Hz, 1H), 7.41 (d, J=7Hz, 2H), 7.31-7.18 (m, 10H), 7.08 (d, J=7Hz, 2H), 6.83 (d, J=9Hz, 4H), 4.85 (q, J=8Hz, 1H), 3.79 (s, 6H), 3.71 (s, 3H), 3.53 (t, J=6Hz, 2H), 3.15-2.94 (m, 6H), 2.61-2.53 (m, 2H), 2.49-2.43 (m, 2H), 1.68-1.61 (m, 2H), 1.60-1.55 (m, 2H).

4.7.2. Branching unit

4.7.2.1. Scheme for total synthesis of branching unit

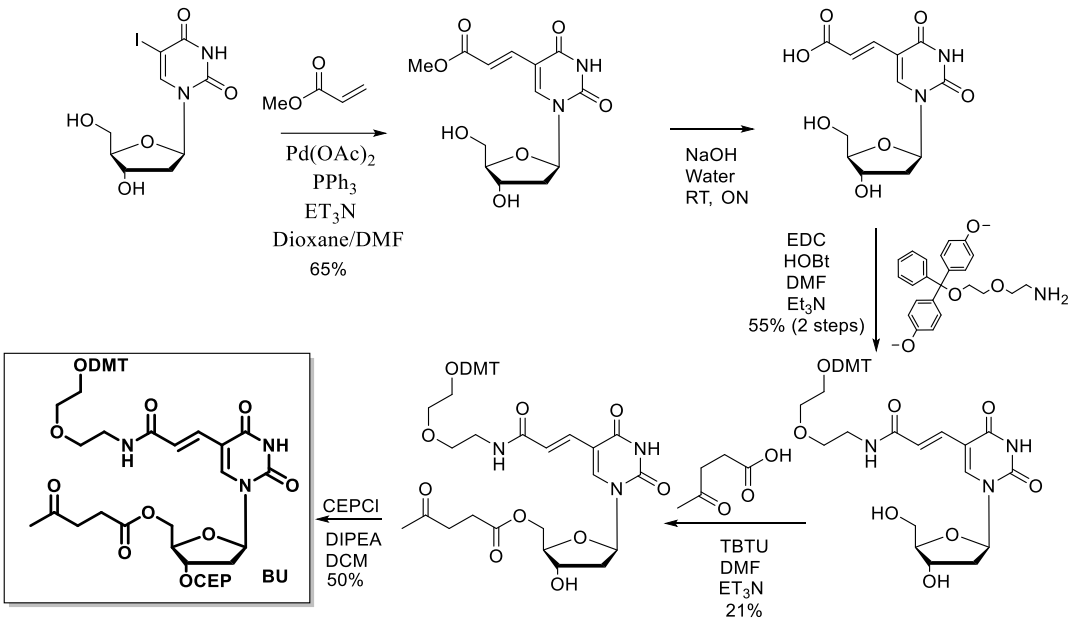
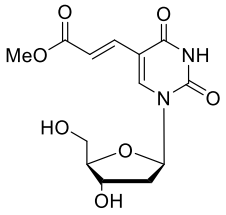


Figure 4.5. The total synthesis is shown above for the branching unit relevant to the construction of the aptamer-code construct. Only the first couple steps were performed and are presented in the following sections.

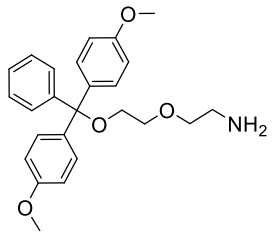
4.7.2.2. Methyl (E)-3-(1-((2R,4S,5R)-4-Hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)acrylate



Protocol inspired by Liu et al.[285] To a solution of 5'-deoxy-5-iodouridine (2.02 g, 5.70 mmol, 1 eq.), methyl acrylate (2.56 mL, 11.4 mmol, 2 eq.), triphenylphosphine (741 mg, 1.14 mmol, 0.2 eq.), and triethylamine (1.97 mL, 11.4 mmol, 2 eq.) in a mixture of dry DMF (25 mL) and dry dioxane (25 mL) was added $Pd(OAc)_2$ (317 mg, 1.41 mmol, 0.1 eq.) at 25 °C under Ar current. The mixture was heated to 90 °C and stirred for 16 h. The reaction mixture was evaporated under reduced pressure to remove the dioxane followed by coevaporation with toluene to dryness. The residue was purified by column chromatography (SiO_2, DCM/10 % methanol in DCM from 10 to 80 % in 7 column volumes) to give the product (1.04 g, 3.33 mmol, 47 % yield) as a white solid.

^{1}H NMR (500 MHz, DMSO): δ (ppm) 11.66 (s, 1H), 8.43 (s, 1H), 7.37 (d, J = 15.8 Hz, 1H), 6.86 (d, J = 15.8 Hz, 1H), 6.13 (t, J = 6.5 Hz, 1H), 5.27 (d, J = 4.3 Hz, 1H), 5.18 (t, J = 5.3 Hz, 1H), 4.26 (dq, J = 6.0, 4.1 Hz, 1H), 3.80 (q, J = 3.7 Hz, 1H), 3.68 (s, 1H), 3.59 (m, 1H), 2.24 – 2.12 (m, 2H).

4.7.2.3. 2-(2-(Bis(4-methoxyphenyl)(phenyl)methoxy)ethoxy)ethan-1-amine



To a solution of 2-(2-aminoethoxy)ethan-1-ol (1.0 mL, 9.97 mmol, 1 equiv.) and triethylamine (4.17 mL, 29.9 mmol, 3 equiv.) in dry DCM (50 mL) was slowly added DMT-Cl (6.75 g, 29.9 mmol, 2 equiv.) on ice. The solution was left under stirring at room temperature under Ar for 2 hours. The solvent was evaporated under reduced pressure, the residue was resuspended in DCM, and ethylthiotetrazole (~0.93 g) was added until a pale pink/orange color appeared. The solution was left under stirring for a few minutes. Product was extracted twice with DCM from 10 % Na_2CO_3 aqueous solution, washed once with a 10 % Na_2CO_3 aqueous solution, dried over $MgSO_4$, filtered, and the solvent was evaporated under reduced pressure (40 °C). The residue was purified by column chromatography (solid loading on celite, SiO_2 pretreated with 0.1 %TEA in DCM, DCM/TEA (100:0.1) and DCM/MeOH/TEA (90:10:0.1)) to give the product (1.52 g, 3.73 mmol, 37 % yield) as a yellow oil. Mass spectrometry was used to confirm the isolation of the pure product.

^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) 7.46 (d, J = 8.6 Hz, 2H), 7.40 – 7.14 (m, 7H), 6.86 – 6.74 (m, 4H), 3.78 (s, 6H), 3.64 (t, J = 5.1 Hz, 2H), 3.62 – 3.47 (m, 2H), 3.27 – 3.15 (m, 2H), 2.91 (s, 2H).

5. References

1. Köhler, G.; Milstein, C., Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **1975,** *256* (5517), 495-497.
2. Trikha, M.; Yan, L.; Nakada, M. T., Monoclonal antibodies as therapeutics in oncology. *Curr. Opin. Biotech.* **2002,** *13* (6), 609-614.
3. Jayasena, S. D., Aptamers: an emerging class of molecules that rival antibodies in diagnostics. *Clin. Chem.* **1999,** *45* (9), 1628-1650.
4. Zhou, J.; Rossi, J., Aptamers as targeted therapeutics: current potential and challenges. *Nat. Rev. Drug Discov.* **2017,** *16* (3), 181-202.
5. Ellington, A. D.; Szostak, J. W., In vitro selection of RNA molecules that bind specific ligands. *Nature* **1990,** *346* (6287), 818-822.
6. Tuerk, C.; Gold, L., Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Sci.* **1990,** *249* (4968), 505-510.
7. Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J., Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature* **1992,** *355* (6360), 564-566.
8. Wu, Y. X.; Kwon, Y. J., Aptamers: The "evolution" of SELEX. *Methods* **2016,** *106*, 21-28.
9. Patel, D. J.; Suri, A. K.; Jiang, F.; Jiang, L.; Fan, P.; Kumar, R. A.; Nonin, S., Structure, recognition and adaptive binding in RNA aptamer complexes. *J. Mol. Biol.* **1997,** *272* (5), 645-664.
10. Germer, K.; Leonard, M.; Zhang, X., RNA aptamers and their therapeutic and diagnostic applications. *Int. J. Biochem. Mol. Biol.* **2013,** *4* (1), 27-40.
11. Keefe, A. D.; Pai, S.; Ellington, A., Aptamers as therapeutics. *Nat. Rev. Drug Discov.* **2010,** *9* (7), 537-550.
12. Davie, E. W.; Kulman, J. D. In *An overview of the structure and function of thrombin*, Seminars in thrombosis and hemostasis, Copyright© 2006 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New …: 2006; pp 003-015.
13. Pica, A.; Russo Krauss, I.; Parente, V.; Tateishi-Karimata, H.; Nagatoishi, S.; Tsumoto, K.; Sugimoto, N.; Sica, F., Through-bond effects in the ternary complexes of thrombin sandwiched by two DNA aptamers. *Nucleic Acids Res.* **2016,** *45*.
14. Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J., Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature* **1992,** *355* (6360), 564.
15. Phan, A. T.; Mergny, J. L., Human telomeric DNA: G-quadruplex, i-motif and Watson–Crick double helix. *Nucleic Acids Res.* **2002,** *30* (21), 4618-4625.
16. Mao, X.-a.; Gmeiner, W. H., NMR study of the folding–unfolding mechanism for the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG). *Biophys. Chem.* **2005,** *113* (2), 155-160.
17. Avino, A.; Fabrega, C.; Tintore, M.; Eritja, R., Thrombin binding aptamer, more than a simple aptamer: chemically modified derivatives and biomedical applications. *Curr. Pharm. Design* **2012,** *18* (14), 2036-2047.
18. Tasset, D. M.; Kubik, M. F.; Steiner, W., Oligonucleotide inhibitors of human thrombin that bind distinct epitopes. *J. Mol. Biol.* **1997,** *272* (5), 688-698.

19. Smirnov, I.; Shafer, R. H., Effect of loop sequence and size on DNA aptamer stability. *Biochem.* **2000,** *39* (6), 1462-1468.
20. Nallagatla, S. R.; Heuberger, B.; Haque, A.; Switzer, C., Combinatorial synthesis of thrombin-binding aptamers containing iso-guanine. *J. Comb. Chem.* **2009,** *11* (3), 364-369.
21. Aaldering, L. J.; Poongavanam, V.; Langkjær, N.; Murugan, N. A.; Jørgensen, P. T.; Wengel, J.; Veedu, R. N., Development of an Efficient G-Quadruplex-Stabilised Thrombin-Binding Aptamer Containing a Three-Carbon Spacer Molecule. *ChemBioChem* **2017,** *18* (8), 755-763.
22. Kovačič, M.; Podbevšek, P.; Tateishi-Karimata, H.; Takahashi, S.; Sugimoto, N.; Plavec, J., Thrombin binding aptamer G-quadruplex stabilized by pyrene-modified nucleotides. *Nucleic Acids Res.* **2020,** *48* (7), 3975-3986.
23. Russo Krauss, I.; Merlino, A.; Giancola, C.; Randazzo, A.; Mazzarella, L.; Sica, F., Thrombin–aptamer recognition: a revealed ambiguity. *Nucleic Acids Res.* **2011,** *39* (17), 7858-7867.
24. Reshetnikov, R.; Golovin, A.; Spiridonova, V.; Kopylov, A.; Šponer, J., Structural Dynamics of Thrombin-Binding DNA Aptamer d(GGTTGGTGTGGTTGG) Quadruplex DNA Studied by Large-Scale Explicit Solvent Simulations. *J. Chem. Theory Comput.* **2010,** *6* (10), 3003-3014.
25. Nagatoishi, S.; Isono, N.; Tsumoto, K.; Sugimoto, N., Loop residues of thrombin-binding DNA aptamer impact G-quadruplex stability and thrombin binding. *Biochimie* **2011,** *93* (8), 1231-1238.
26. Bayat, P.; Nosrati, R.; Alibolandi, M.; Rafatpanah, H.; Abnous, K.; Khedri, M.; Ramezani, M., SELEX methods on the road to protein targeting with nucleic acid aptamers. *Biochimie* **2018,** *154*, 132-155.
27. Vant-Hull, B.; Payano-Baez, A.; Davis, R. H.; Gold, L., The mathematics of SELEX against complex targets. *J. Mol. Biol.* **1998,** *278* (3), 579-597.
28. Sampson, T., Aptamers and SELEX: the technology. *World Pat. Inf.* **2003,** *25* (2), 123-129.
29. Bittker, J. A.; Le, B. V.; Liu, D. R., Nucleic acid evolution and minimization by nonhomologous random recombination. *Nat. Biotech.* **2002,** *20* (10), 1024-1029.
30. Goers, E. S.; Purcell, J.; Voelker, R. B.; Gates, D. P.; Berglund, J. A., MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. *Nucleic Acids Res.* **2010,** *38* (7), 2467-2484.
31. Ohuchi, S.; Mori, Y.; Nakamura, Y., Evolution of an inhibitory RNA aptamer against T7 RNA polymerase. *FEBS Open Bio* **2012,** *2*, 203-207.
32. Singer, B. S.; Shtatland, T.; Brown, D.; Gold, L., Libraries for genomic SELEX. *Nucleic Acids Res.* **1997,** *25* (4), 781-786.
33. Zimmermann, B.; Bilusic, I.; Lorenz, C.; Schroeder, R., Genomic SELEX: A discovery tool for genomic aptamers. *Methods* **2010,** *52* (2), 125-132.
34. Wang, J.; Gong, Q.; Maheshwari, N.; Eisenstein, M.; Arcila, M. L.; Kosik, K. S.; Soh, H. T., Particle Display: A Quantitative Screening Method for Generating High-Affinity Aptamers. *Angew. Chem.* **2014,** *126* (19), 4896-4901.
35. Vater, A.; Jarosch, F.; Buchner, K.; Klussmann, S., Short bioactive Spiegelmers to migraine-associated calcitonin gene-related peptide rapidly identified by a novel approach: Tailored-SELEX. *Nucleic Acids Res.* **2003,** *31* (21), e130-e130.

36. Shtatland, T.; Gill, S. C.; Javornik, B. E.; Johansson, H. E.; Singer, B. S.; Uhlenbeck, O. C.; Zichi, D. A.; Gold, L., Interactions of Escherichia coli RNA with bacteriophage MS2 coat protein: genomic SELEX. *Nucleic Acids Res.* **2000,** *28* (21), e93-e93.
37. Pan, W.; Clawson, G. A., Primer-Free Aptamer Selection Using a Random DNA Library. In *RNA Therapeutics: Function, Design, and Delivery*, Sioud, M., Ed. Humana Press: Totowa, NJ, 2010; pp 367-383.
38. Wang, B., A New Design for the Fixed Primer Regions in an Oligonucleotide Library for SELEX Aptamer Screening. *Front. Chem.* **2020,** *8*, 475-475.
39. Vianini, E.; Palumbo, M.; Gatto, B., In vitro selection of DNA aptamers that bind l-tyrosinamide. *Bioorg. Med. Chem.* **2001,** *9* (10), 2543-2548.
40. Fitzwater, T.; Polisky, B., [17] A SELEX primer. In *Meth. Enzym.*, Academic Press: 1996; Vol. 267, pp 275-301.
41. Setlem, K.; Mondal, B.; Ramlal, S.; Kingston, J., Immuno Affinity SELEX for Simple, Rapid, and Cost-Effective Aptamer Enrichment and Identification against Aflatoxin B1. *Front. Microbiol.* **2016,** *7* (1909).
42. Bae, H.; Ren, S.; Kang, J.; Kim, M.; Jiang, Y.; Jin, M. M.; Min, I. M.; Kim, S., Sol-gel SELEX circumventing chemical conjugation of low molecular weight metabolites discovers aptamers selective to xanthine. *Nucleic Acid Ther.* **2013,** *23* (6), 443-449.
43. Ozer, A.; Pagano, J. M.; Lis, J. T., New Technologies Provide Quantum Changes in the Scale, Speed, and Success of SELEX Methods and Aptamer Characterization. *Mol. Ther. Nucleic Acids* **2014,** *3*, e183.
44. Stoltenburg, R.; Reinemann, C.; Strehlitz, B., FluMag-SELEX as an advantageous method for DNA aptamer selection. *Anal. Bioanal. Chem.* **2005,** *383* (1), 83-91.
45. Lou, X.; Qian, J.; Xiao, Y.; Viel, L.; Gerdon, A. E.; Lagally, E. T.; Atzberger, P.; Tarasow, T. M.; Heeger, A. J.; Soh, H. T., Micromagnetic selection of aptamers in microfluidic channels. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (9), 2989-2994.
46. Wakui, K.; Yoshitomi, T.; Yamaguchi, A.; Tsuchida, M.; Saito, S.; Shibukawa, M.; Furusho, H.; Yoshimoto, K., Rapidly Neutralizable and Highly Anticoagulant Thrombin-Binding DNA Aptamer Discovered by MACE SELEX. *Mol. Ther. Nucleic Acids* **2019,** *16*, 348-359.
47. Dong, Y.; Wang, Z.; Wang, S.; Wu, Y.; Ma, Y.; Liu, J., Introduction of SELEX and Important SELEX Variants. In *Aptamers for Analytical Applications*, 2018; pp 1-25.
48. Stoltenburg, R.; Nikolaus, N.; Strehlitz, B., Capture-SELEX: selection of DNA aptamers for aminoglycoside antibiotics. *J. Anal. Methods Chem.* **2012,** *2012*.
49. Oh, S. S.; Plakos, K.; Lou, X.; Xiao, Y.; Soh, H. T., In vitro selection of structure-switching, self-reporting aptamers. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107* (32), 14053-14058.
50. Duan, N.; Gong, W.; Wu, S.; Wang, Z., An ssDNA library immobilized SELEX technique for selection of an aptamer against ractopamine. *Anal. Chim. Acta* **2017,** *961*, 100-105.
51. Chatterjee, B.; Kalyani, N.; Anand, A.; Khan, E.; Das, S.; Bansal, V.; Kumar, A.; Sharma, T. K., GOLD SELEX: a novel SELEX approach for the development of high-affinity aptamers against small molecules without residual activity. *Microchim. Acta* **2020,** *187* (11), 618.
52. Vorobyeva, M.; Davydova, A.; Vorobjev, P.; Ven'yaminova, A., Key Aspects of Nucleic Acid Library Design for in Vitro Selection. *Int. J. Mol. Sci.* **2018,** *19*, 470.

53. Mendonsa, S. D.; Bowser, M. T., In Vitro Evolution of Functional DNA Using Capillary Electrophoresis. *J. Am. Chem. Soc.* **2004,** *126* (1), 20-21.
54. WooáKim, D.; BockáGu, M., Immobilization-free screening of aptamers assisted by graphene oxide. *Chem. Commun.* **2012,** *48* (15), 2071-2073.
55. Shi, H.; Kou, Q.; Wu, P.; Sun, Q.; Wu, J.; Le, T., Selection and Application of DNA Aptamers Against Sulfaquinoxaline Assisted by Graphene Oxide–Based SELEX. *Food Anal. Method.* **2021,** *14* (2), 250-259.
56. Qiao, N.; Li, J.; Wu, X.; Diao, D.; Zhao, J.; Li, J.; Ren, X.; Ding, X.; Shangguan, D.; Lou, X., Speeding up in Vitro Discovery of Structure-Switching Aptamers via Magnetic Cross-Linking Precipitation. *Anal. Chem.* **2019,** *91* (21), 13383-13389.
57. Jenison, R. D.; Gill, S. C.; Pardi, A.; Polisky, B., High-resolution molecular discrimination by RNA. *Sci.* **1994,** *263* (5152), 1425-1429.
58. Sedighian, H.; Halabian, R.; Amani, J.; Heiat, M.; Amin, M.; Fooladi, A. A. I., Staggered Target SELEX, a novel approach to isolate non-cross-reactive aptamer for detection of SEA by apta-qPCR. *J. Biotech.* **2018,** *286*, 45-55.
59. Waybrant, B.; Pearce, T. R.; Wang, P.; Sreevatsan, S.; Kokkoli, E., Development and characterization of an aptamer binding ligand of fractalkine using domain targeted SELEX. *Chem. Commun.* **2012,** *48* (80), 10043-10045.
60. Song, M. Y.; Nguyen, D.; Hong, S. W.; Kim, B. C., Broadly reactive aptamers targeting bacteria belonging to different genera using a sequential toggle cell-SELEX. *Sci. Rep.* **2017,** *7* (1), 43641.
61. Gong, Q.; Wang, J.; Ahmad, K. M.; Csordas, A. T.; Zhou, J.; Nie, J.; Stewart, R.; Thomson, J. A.; Rossi, J. J.; Soh, H. T., Selection strategy to generate aptamer pairs that bind to distinct sites on protein targets. *Anal. Chem.* **2012,** *84* (12), 5365-5371.
62. Lao, Y.-H.; Chiang, H.-Y.; Yang, D.-K.; Peck, K.; Chen, L.-C., Selection of aptamers targeting the sialic acid receptor of hemagglutinin by epitope-specific SELEX. *Chem. Commun.* **2014,** *50* (63), 8719-8722.
63. Smith, J. D.; Gold, L., Conditional-selex. Google Patents: 2004.
64. Sefah, K.; Shangguan, D.; Xiong, X.; O'donoghue, M. B.; Tan, W., Development of DNA aptamers using Cell-SELEX. *Nat. Protoc.* **2010,** *5* (6), 1169.
65. Fang, X.; Tan, W., Aptamers generated from cell-SELEX for molecular medicine: a chemical biology approach. *Accounts Chem. Res.* **2010,** *43* (1), 48-57.
66. Hicke, B. J.; Marion, C.; Chang, Y.-F.; Gould, T.; Lynott, C. K.; Parma, D.; Schmidt, P. G.; Warren, S., Tenascin-C aptamers are generated using tumor cells and purified protein. *J. Biol. Chem.* **2001,** *276* (52), 48644-48654.
67. Thiel, W. H.; Thiel, K. W.; Flenker, K. S.; Bair, T.; Dupuy, A. J.; McNamara, J. O.; Miller, F. J.; Giangrande, P. H., Cell-internalization SELEX: method for identifying cell-internalizing RNA aptamers for delivering siRNAs to target cells. In *RNA Interference*, Springer: 2015; pp 187-199.
68. Uemachi, H.; Kasahara, Y.; Tanaka, K.; Okuda, T.; Yoneda, Y.; Obika, S., Hybrid-Type SELEX for the Selection of Artificial Nucleic Acid Aptamers Exhibiting Cell Internalization Activity. *Pharmaceutics* **2021,** *13* (6), 888.
69. Mayer, G.; Ahmed, M.-S. L.; Dolf, A.; Endl, E.; Knolle, P. A.; Famulok, M., Fluorescence-activated cell sorting for aptamer SELEX with cell mixtures. *Nat. Protoc.* **2010,** *5* (12), 1993-2004.

70. Souza, A. G.; Marangoni, K.; Fujimura, P. T.; Alves, P. T.; Silva, M. J.; Bastos, V. A. F.; Goulart, L. R.; Goulart, V. A., 3D Cell-SELEX: Development of RNA aptamers as molecular probes for PC-3 tumor cell line. *Exp. Cell. Res.* **2016,** *341* (2), 147-156.
71. Mi, J.; Liu, Y.; Rabbani, Z. N.; Yang, Z.; Urban, J. H.; Sullenger, B. A.; Clary, B. M., In vivo selection of tumor-targeting RNA motifs. *Nat. Chem. Biol.* **2010,** *6* (1), 22-24.
72. Marimuthu, C.; Tang, T.-H.; Tominaga, J.; Tan, S.-C.; Gopinath, S. C., Single-stranded DNA (ssDNA) production in DNA aptamer generation. *Analyst* **2012,** *137* (6), 1307-1315.
73. Bruno, J. G.; Kiel, J. L., In vitro selection of DNA aptamers to anthrax spores with electrochemiluminescence detection. *Biosens. Bioelectron.* **1999,** *14* (5), 457-464.
74. Kanagawa, T., Bias and artifacts in multitemplate polymerase chain reactions (PCR). *J. Biosci. Bioeng.* **2003,** *96* (4), 317-323.
75. Tolle, F.; Wilke, J.; Wengel, J.; Mayer, G., By-product formation in repetitive PCR amplification of DNA libraries during SELEX. *PloS One* **2014,** *9* (12), e114693.
76. Ouellet, E.; Foley, J. H.; Conway, E. M.; Haynes, C., Hi-Fi SELEX: A high-fidelity digital-PCR based therapeutic aptamer discovery platform. *Biotech. Bioeng.* **2015,** *112* (8), 1506-1522.
77. Song, S.; Wang, X.; Xu, K.; Li, Q.; Ning, L.; Yang, X., Selection of highly specific aptamers to Vibrio parahaemolyticus using cell-SELEX powered by functionalized graphene oxide and rolling circle amplification. *Anal. Chim. Acta* **2019,** *1052*, 153-162.
78. Avci-Adali, M.; Wilhelm, N.; Perle, N.; Stoll, H.; Schlensak, C.; Wendel, H. P., Absolute quantification of cell-bound DNA aptamers during SELEX. *Nucleic Acid Ther.* **2013,** *23* (2), 125-130.
79. Chen, J.; Liu, X.; Xu, M.; Li, Z.; Xu, D., Accomplishment of one-step specific PCR and evaluated SELEX process by a dual-microfluidic amplified system. *Biomicrofluidics* **2021,** *15* (2), 024107.
80. Sanger, F.; Nicklen, S.; Coulson, A. R., DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **1977,** *74* (12), 5463-5467.
81. Beier, R.; Boschke, E.; Labudde, D., New strategies for evaluation and analysis of SELEX experiments. *BioMed Res. Int.* **2014,** *2014*.
82. Cho, M.; Xiao, Y.; Nie, J.; Stewart, R.; Csordas, A. T.; Oh, S. S.; Thomson, J. A.; Soh, H. T., Quantitative selection of DNA aptamers through microfluidic selection and high-throughput sequencing. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107* (35), 15373-15378.
83. Schütze, T.; Wilhelm, B.; Greiner, N.; Braun, H.; Peter, F.; Mörl, M.; Erdmann, V. A.; Lehrach, H.; Konthur, Z.; Menger, M., Probing the SELEX process with next-generation sequencing. *PloS One* **2011,** *6* (12), e29604.
84. Hoinka, J.; Berezhnoy, A.; Dao, P.; Sauna, Z. E.; Gilboa, E.; Przytycka, T. M., Large scale analysis of the mutational landscape in HT-SELEX improves aptamer discovery. *Nucleic Acids Res.* **2015,** *43* (12), 5699-5707.
85. Blank, M., Next-Generation Analysis of Deep Sequencing Data: Bringing Light into the Black Box of SELEX Experiments. In *Nucleic Acid Aptamers: Selection, Characterization, and Application*, Mayer, G., Ed. Springer New York: New York, NY, 2016; pp 85-95.
86. Alam, K. K.; Chang, J. L.; Burke, D. H., FASTAptamer: a bioinformatic toolkit for high-throughput sequence analysis of combinatorial selections. *Mol. Ther. Nucleic Acids* **2015,** *4*, e230.

87. Caroli, J.; Taccioli, C.; De La Fuente, A.; Serafini, P.; Bicciato, S., APTANI: a computational tool to select aptamers through sequence-structure motif analysis of HT-SELEX data. *Bioinformatics* **2015,** *32* (2), 161-164.
88. Dao, P.; Hoinka, J.; Takahashi, M.; Zhou, J.; Ho, M.; Wang, Y.; Costa, F.; Rossi, J. J.; Backofen, R.; Burnett, J.; Przytycka, T. M., AptaTRACE Elucidates RNA Sequence-Structure Motifs from Selection Trends in HT-SELEX Experiments. *Cell Syst.* **2016,** *3* (1), 62-70.
89. Kato, S.; Ono, T.; Minagawa, H.; Horii, K.; Shiratori, I.; Waga, I.; Ito, K.; Aoki, T., FSBC: fast string-based clustering for HT-SELEX data. *BMC Bioinform.* **2020,** *21* (1), 263.
90. Song, J.; Zheng, Y.; Huang, M.; Wu, L.; Wang, W.; Zhu, Z.; Song, Y.; Yang, C., A Sequential Multidimensional Analysis Algorithm for Aptamer Identification based on Structure Analysis and Machine Learning. *Anal. Chem.* **2020,** *92* (4), 3307-3314.
91. Bielinska, A.; Shivdasani, R. A.; Zhang, L.; Nabel, G. J., Regulation of gene expression with double-stranded phosphorothioate oligonucleotides. *Science* **1990,** *250* (4983), 997-1000.
92. Eckstein, F., Phosphorothioates, essential components of therapeutic oligonucleotides. *Nucleic Acid Ther.* **2014,** *24* (6), 374-387.
93. Volk, D. E.; Lokesh, G. L., Development of phosphorothioate DNA and DNA thioaptamers. *Biomedicines* **2017,** *5* (3), 41.
94. King, D. J.; Ventura, D. A.; Brasier, A. R.; Gorenstein, D. G., Novel combinatorial selection of phosphorothioate oligonucleotide aptamers. *Biochemistry* **1998,** *37* (47), 16489-16493.
95. Higashimoto, Y.; Matsui, T.; Nishino, Y.; Taira, J.; Inoue, H.; Takeuchi, M.; Yamagishi, S.-i., Blockade by phosphorothioate aptamers of advanced glycation end products-induced damage in cultured pericytes and endothelial cells. *Microvasc. Res.* **2013,** *90*, 64-70.
96. King, D. J.; Bassett, S. E.; Li, X.; Fennewald, S. A.; Herzog, N. K.; Luxon, B. A.; Shope, R.; Gorenstein, D. G., Combinatorial Selection and Binding of Phosphorothioate Aptamers Targeting Human NF-κB RelA(p65) and p50. *Biochemistry* **2002,** *41* (30), 9696-9706.
97. Yang, X.; Fennewald, S.; Luxon, B. A.; Aronson, J.; Herzog, N. K.; Gorenstein, D. G., Aptamers containing thymidine 3'-O-phosphorodithioates: Synthesis and binding to nuclear factor-κB. *Bioorg. Med. Chem. Lett.* **1999,** *9* (23), 3357-3362.
98. Thayer, J. R.; Wu, Y.; Hansen, E.; Angelino, M. D.; Rao, S., Separation of oligonucleotide phosphorothioate diastereoisomers by pellicular anion-exchange chromatography. *J. Chromatogr. A* **2011,** *1218* (6), 802-808.
99. Saran, R.; Huang, Z.; Liu, J., Phosphorothioate nucleic acids for probing metal binding, biosensing and nanotechnology. *Coordin. Chem. Rev.* **2021,** *428*, 213624.
100. Østergaard, M. E.; De Hoyos, C. L.; Wan, W. B.; Shen, W.; Low, A.; Berdeja, A.; Vasquez, G.; Murray, S.; Migawa, M. T.; Liang, X.-h.; Swayze, E. E.; Crooke, S. T.; Seth, P. P., Understanding the effect of controlling phosphorothioate chirality in the DNA gap on the potency and safety of gapmer antisense oligonucleotides. *Nucleic Acids Res.* **2020,** *48* (4), 1691-1700.
101. Nukaga, Y.; Yamada, K.; Ogata, T.; Oka, N.; Wada, T., Stereocontrolled solid-phase synthesis of phosphorothioate oligoribonucleotides using 2'-O-(2-cyanoethoxymethyl)-nucleoside 3'-O-oxazaphospholidine monomers. *J. Org. Chem.* **2012,** *77* (18), 7913-7922.
102. Huang, P. J. J.; Liu, J., Rational evolution of Cd2+-specific DNAzymes with phosphorothioate modified cleavage junction and Cd2+ sensing. *Nucleic Acids Res.* **2015,** *43* (12), 6125-6133.

103. Lato, S. M.; Ozerova, N. D.; He, K.; Sergueeva, Z.; Shaw, B. R.; Burke, D. H., Boron-containing aptamers to ATP. *Nucleic Acids Res.* **2002,** *30* (6), 1401-1407.
104. Li, P.; Sergueeva, Z. A.; Dobrikov, M.; Shaw, B. R., Nucleoside and oligonucleoside boranophosphates: chemistry and properties. *Chem. Rev.* **2007,** *107* (11), 4746-4796.
105. Shaw, B. R.; Dobrikov, M.; Wang, X.; Wan, J.; He, K.; Lin, J.-l.; Li, P.; Rait, V.; Sergueeva, Z. A.; Sergueev, D., Reading, Writing, and Modulating Genetic Information with Boranophosphate Mimics of Nucleotides, DNA, and RNA. *Ann. N. Y. Acad. Sci.* **2003,** *1002* (1), 12-29.
106. Wan, J., Enzymatic synthesis, properties, and functions of boranophosphate RNA. **2005**.
107. Yang, N. J.; Hinner, M. J., Getting across the cell membrane: an overview for small molecules, peptides, and proteins. *Method. Mol. Biol.* **2015,** *1266*, 29-53.
108. Saccà, B.; Lacroix, L.; Mergny, J.-L., The effect of chemical modifications on the thermal stability of different G-quadruplex-forming oligonucleotides. *Nucleic Acids Res.* **2005,** *33* (4), 1182-1192.
109. Arangundy-Franklin, S.; Taylor, A. I.; Porebski, B. T.; Genna, V.; Peak-Chew, S.; Vaisman, A.; Woodgate, R.; Orozco, M.; Holliger, P., A synthetic genetic polymer with an uncharged backbone chemistry based on alkyl phosphonate nucleic acids. *Nat. Chem.* **2019,** *11* (6), 533-542.
110. Thiviyanathan, V.; Vyazovkina, K. V.; Gozansky, E. K.; Bichenchova, E.; Abramova, T. V.; Luxon, B. A.; Lebedev, A. V.; Gorenstein, D. G., Structure of Hybrid Backbone Methylphosphonate DNA Heteroduplexes: Effect of R and S Stereochemistry. *Biochemistry* **2002,** *41* (3), 827-838.
111. Gryaznov, S. M., Oligonucleotide N3'→P5' phosphoramidates as potential therapeutic agents. *Biochim. Biophys. Acta, Gene Struct. Expression* **1999,** *1489* (1), 131-140.
112. Darfeuille, F.; Arzumanov, A.; Gryaznov, S.; Gait, M. J.; Di Primo, C.; Toulmé, J.-J., Loop–loop interaction of HIV-1 TAR RNA with N3'→ P5' deoxyphosphoramidate aptamers inhibits in vitro Tat-mediated transcription. *Proc. Natl. Acad. Sci. U.S.A.* **2002,** *99* (15), 9709-9714.
113. Eulberg, D.; Klussmann, S., Spiegelmers: biostable aptamers. *ChemBioChem* **2003,** *4* (10), 979-983.
114. Schwoebel, F.; van Eijk, L. T.; Zboralski, D.; Sell, S.; Buchner, K.; Maasch, C.; Purschke, W. G.; Humphrey, M.; Zöllner, S.; Eulberg, D.; Morich, F.; Pickkers, P.; Klussmann, S., The effects of the anti-hepcidin Spiegelmer NOX-H94 on inflammation-induced anemia in cynomolgus monkeys. *Blood* **2013,** *121* (12), 2311-2315.
115. Bilik, K. U.; Ergüven, E.; Klussmann, S.; Jarosch, F.; Wielinga, P. Y.; Lutz, T. A.; Riediger, T., In-vitro and in-vivo antagonistic action of an anti-amylin Spiegelmer. *Neuroreport* **2007,** *18* (17), 1855-1859.
116. Gao, X.; Brown, F. K.; Jeffs, P.; Bischofberger, N.; Lin, K. Y.; Pipe, A. J.; Noble, S. A., Probing structural factors stabilizing antisense oligonucleotide duplexes: NMR studies of a DNA. cntdot. DNA duplex containing a formacetal linkage. *Biochemistry* **1992,** *31* (27), 6228-6236.
117. He, G.-X.; Williams, J. P.; Postich, M. J.; Swaminathan, S.; Shea, R. G.; Terhorst, T.; Law, V. S.; Mao, C. T.; Sueoka, C.; Coutré, S., In vitro and in vivo activities of oligodeoxynucleotide-based thrombin inhibitors containing neutral formacetal linkages. *J. Med. Chem.* **1998,** *41* (22), 4224-4231.
118. Varizhuk, A. M.; Tsvetkov, V. B.; Tatarinova, O. N.; Kaluzhny, D. N.; Florentiev, V. L.; Timofeev, E. N.; Shchyolkina, A. K.; Borisova, O. F.; Smirnov, I. P.; Grokhovsky, S. L., Synthesis,

characterization and in vitro activity of thrombin-binding DNA aptamers with triazole internucleotide linkages. *Eur. J. Med. Chem.* **2013,** *67*, 90-97.
119. JeongáLee, E.; KyungáLim, H.; SeuláCho, Y.; SooáHah, S., Peptide nucleic acids are an additional class of aptamers. *RSC Adv.* **2013,** *3* (17), 5828-5831.
120. Plavec, J.; Thibaudeau, C.; Chattopadhyaya, J., How do the energetics of the stereoelectronic gauche and anomeric effects modulate the conformation of nucleos (t) ides? *Pure Appl. Chem.* **1996,** *68* (11), 2137-2144.
121. Freier, S. M.; Altmann, K.-H., The ups and downs of nucleic acid duplex stability: Structure-stability studies on chemically-modified DNA:RNA duplexes. *Nucleic Acids Res.* **1997,** *25* (22), 4429-4443.
122. Moran, L.; Scrimgeour, K.; Horton, H.; Ochs, R.; Rawn, J., Biochemistry; 2" ed.; Moran., LA, ed. Neil Patterson Publishers–Prentice Hall: Englewood Cliffs, NJ: 1994.
123. Kawasaki, A. M.; Casper, M. D.; Freier, S. M.; Lesnik, E. A.; Zounes, M. C.; Cummins, L. L.; Gonzalez, C.; Cook, P. D., Uniformly modified 2'-deoxy-2'-fluoro-phosphorothioate oligonucleotides as nuclease-resistant antisense compounds with high affinity and specificity for RNA targets. *J. Med. Chem.* **1993,** *36* (7), 831-841.
124. Thirunavukarasu, D.; Chen, T.; Liu, Z.; Hongdilokkul, N.; Romesberg, F. E., Selection of 2′-Fluoro-Modified Aptamers with Optimized Properties. *J. Am. Chem. Soc.* **2017,** *139* (8), 2892-2895.
125. Rhie, A.; Kirby, L.; Sayer, N.; Wellesley, R.; Disterer, P.; Sylvester, I.; Gill, A.; Hope, J.; James, W.; Tahiri-Alaoui, A., Characterization of 2′-fluoro-RNA aptamers that bind preferentially to disease-associated conformations of prion protein and inhibit conversion. *J. Biol. Chem.* **2003,** *278* (41), 39697-39705.
126. Hervas-Stubbs, S.; Soldevilla, M. M.; Villanueva, H.; Mancheño, U.; Bendandi, M.; Pastor, F., Identification of TIM3 2′-fluoro oligonucleotide aptamer by HT-SELEX for cancer immunotherapy. *Oncotarget* **2016,** *7* (4), 4522.
127. Aurup, H.; Tuschl, T.; Benseler, F.; Ludwig, J.; Eckstein, F., Oligonucleotide duplexes containing 2′-amino-2′deoxycytidines: thermal stability and chemical reactivity. *Nucleic Acids Res.* **1994,** *22* (4), 701-701.
128. Proske, D.; Gilch, S.; Wopfner, F.; Schätzl, H. M.; Winnacker, E. L.; Famulok, M., Prion-protein-specific aptamer reduces PrPSc formation. *Chembiochem* **2002,** *3* (8), 717-725.
129. Yan, X.; Gao, X.; Zhang, Z., Isolation and characterization of 2′-amino-modified RNA aptamers for human TNFα. *Genomics Proteomics Bioinformatics* **2004,** *2* (1), 32-42.
130. Dimitrova, D. G.; Teysset, L.; Carré, C., RNA 2′-O-methylation (Nm) modification in human diseases. *Genes* **2019,** *10* (2), 117.
131. Sproat, B. S.; Lamond, A. I.; Beijer, B.; Neuner, P.; Ryder, U., Highly efficient chemical synthesis of 2′-O-methyloligoribonucleotides and tetrabiotinylated derivatives; novel probes that are resistant to degradation by RNA or DNA specific nucleases. *Nucleic Acids Res.* **1989,** *17* (9), 3373-3386.
132. Kuwasaki, T.; Hatta, M.; Takeuchi, H.; Takaku, H., Inhibition of human immunodeficiency virus 1 replication in vitro by a self-stabilized oligonucleotide with 2′-O-methyl-guanosine-uridine quadruplex motifs. *J. Antimicrob. Chemother.* **2003,** *51* (4), 813-819.

133. Darfeuille, F.; Arzumanov, A.; Gait, M. J.; Di Primo, C.; Toulmé, J.-J., 2'-O-Methyl-RNA Hairpins Generate Loop−Loop Complexes and Selectively Inhibit HIV-1 Tat-Mediated Transcription. *Biochemistry* **2002,** *41* (40), 12186-12192.
134. Burmeister, P. E.; Lewis, S. D.; Silva, R. F.; Preiss, J. R.; Horwitz, L. R.; Pendergrast, P. S.; McCauley, T. G.; Kurz, J. C.; Epstein, D. M.; Wilson, C., Direct in vitro selection of a 2'-O-methyl aptamer to VEGF. *Chem. Biol.* **2005,** *12* (1), 25-33.
135. Bhat, B.; Esau, C.; Davis, S.; Propp, S.; Kinberger, G.; Gaus, H.; Freier, S.; Swayze, E. E.; Bennett, C. F., 2'-O-Methoxyethyl/2'-Fluoro Modified Oligonucleotides Result in More Potent Inhibition of micro RNA-122 in Vivo: A Target implicatedin HCV Replication. *Nucleic Acids Symp. Ser.* **2008,** *52* (1), 69-69.
136. Geary, R. S.; Watanabe, T. A.; Truong, L.; Freier, S.; Lesnik, E. A.; Sioufi, N. B.; Sasmor, H.; Manoharan, M.; Levin, A. A., Pharmacokinetic properties of 2'-O-(2-methoxyethyl)-modified oligonucleotide analogs in rats. *J. Pharmacol. Exp. Ther.* **2001,** *296* (3), 890-897.
137. Lenn, J. D.; Neil, J.; Donahue, C.; Demock, K.; Tibbetts, C. V.; Cote-Sierra, J.; Smith, S. H.; Rubenstein, D.; Therrien, J.-P.; Pendergrast, P. S., RNA aptamer delivery through intact human skin. *J. Invest. Dermatol.* **2018,** *138* (2), 282-290.
138. Werner, D.; Brunar, H.; Noe, C. R., Investigations on the influence of 2'-O-alkyl modifications on the base pairing properties of oligonucleotides. *Pharm. Acta Helv.* **1998,** *73* (1), 3-10.
139. Noe, C. R.; Winkler, J.; Urban, E.; Gilbert, M.; Haberhauer, G.; Brunar, H., Zwitterionic oligonucleotides: a study on binding properties of 2'-O-aminohexyl modifications. *Nucleosides Nucleotides Nucleic Acids* **2005,** *24* (8), 1167-1185.
140. Shao, Q.; Chen, T.; Sheng, K.; Liu, Z.; Zhang, Z.; Romesberg, F. E., Selection of Aptamers with Large Hydrophobic 2'-Substituents. *J. Am. Chem. Soc.* **2020,** *142* (5), 2125-2128.
141. Kaur, H.; Babu, B. R.; Maiti, S., Perspectives on Chemistry and Therapeutic Applications of Locked Nucleic Acid (LNA). *Chem. Rev.* **2007,** *107* (11), 4672-4697.
142. Wahlestedt, C.; Salmi, P.; Good, L.; Kela, J.; Johnsson, T.; Hökfelt, T.; Broberger, C.; Porreca, F.; Lai, J.; Ren, K., Potent and nontoxic antisense oligonucleotides containing locked nucleic acids. *Proc. Natl. Acad. Sci. U.S.A.* **2000,** *97* (10), 5633-5638.
143. Darfeuille, F.; Hansen, J. B.; Orum, H.; Primo, C. D.; Toulmé, J. J., LNA/DNA chimeric oligomers mimic RNA aptamers targeted to the TAR RNA element of HIV-1. *Nucleic Acids Res.* **2004,** *32* (10), 3101-3107.
144. Schmidt, K. S.; Borkowski, S.; Kurreck, J.; Stephens, A. W.; Bald, R.; Hecht, M.; Friebe, M.; Dinkelborg, L.; Erdmann, V. A., Application of locked nucleic acids to improve aptamer in vivo stability and targeting function. *Nucleic Acids Res.* **2004,** *32* (19), 5757-5765.
145. Langkjær, N.; Pasternak, A.; Wengel, J., UNA (unlocked nucleic acid): a flexible RNA mimic that allows engineering of nucleic acid duplex stability. *Bioorg. Med. Chem.* **2009,** *17* (15), 5420-5425.
146. Pasternak, A.; Hernandez, F. J.; Rasmussen, L. M.; Vester, B.; Wengel, J., Improved thrombin binding aptamer by incorporation of a single unlocked nucleic acid monomer. *Nucleic Acids Res.* **2011,** *39* (3), 1155-1164.
147. Jensen, T. B.; Langkjær, N.; Wengel, J., Unlocked nucleic acid (UNA) and UNA derivatives: Thermal denaturation studies. *Nucleic Acids Symp. Ser.* **2008,** *52* (1), 133-134.

148. Jensen, T. B.; Henriksen, J. R.; Rasmussen, B. E.; Rasmussen, L. M.; Andresen, T. L.; Wengel, J.; Pasternak, A., Thermodynamic and biological evaluation of a thrombin binding aptamer modified with several unlocked nucleic acid (UNA) monomers and a 2′-C-piperazino-UNA monomer. *Bioorg. Med. Chem.* **2011,** *19* (16), 4739-4745.
149. Martín-Pintado, N.; Yahyaee-Anzahaee, M.; Campos-Olivas, R.; Noronha, A. M.; Wilds, C. J.; Damha, M. J.; González, C., The solution structure of double helical arabino nucleic acids (ANA and 2'F-ANA): effect of arabinoses in duplex-hairpin interconversion. *Nucleic Acids Res.* **2012,** *40* (18), 9329-39.
150. Kalota, A.; Karabon, L.; Swider, C.; Viazovkina, E.; Elzagheid, M.; Damha, M.; Gewirtz, A., 2′-deoxy-2′-fluoro-β-D-arabinonucleic acid (2′ F-ANA) modified oligonucleotides (ON) effect highly efficient, and persistent, gene silencing. *Nucleic Acids Res.* **2006,** *34* (2), 451-461.
151. Peng, C. G.; Damha, M. J., G-quadruplex induced stabilization by 2′-deoxy-2′-fluoro-D-arabinonucleic acids (2′ F-ANA). *Nucleic Acids Res.* **2007,** *35* (15), 4977-4988.
152. Lietard, J.; Abou Assi, H.; Gómez-Pinto, I.; González, C.; Somoza, M. M.; Damha, M. J., Mapping the affinity landscape of Thrombin-binding aptamers on 2′F-ANA/DNA chimeric G-Quadruplex microarrays. *Nucleic Acids Res.* **2017,** *45* (4), 1619-1632.
153. Alves Ferreira-Bravo, I.; Cozens, C.; Holliger, P.; DeStefano, J. J., Selection of 2′-deoxy-2′-fluoroarabinonucleotide (FANA) aptamers that bind HIV-1 reverse transcriptase with picomolar affinity. *Nucleic Acids Res.* **2015,** *43* (20), 9587-9599.
154. Morvan, F.; Debart, F.; Vasseur, J. J., From Anionic to Cationic α-Anomeric Oligodeoxynucleotides. *Chemistry & biodiversity* **2010,** *7* (3), 494-535.
155. Kolganova, N. A.; Varizhuk, A. M.; Novikov, R. A.; Florentiev, V. L.; Pozmogova, G. E.; Borisova, O. F.; Shchyolkina, A. K.; Smirnov, I. P.; Kaluzhny, D. N.; Timofeev, E. N., Anomeric DNA quadruplexes: Modified thrombin aptamers. *Artificial Dna: Pna & Xna* **2014,** *5* (2), e28422.
156. Frieden, M.; Christensen, S. M.; Mikkelsen, N. D.; Rosenbohm, C.; Thrue, C. A.; Westergaard, M.; Hansen, H. F.; Ørum, H.; Koch, T., Expanding the design horizon of antisense oligonucleotides with alpha-l-LNA. *Nucleic Acids Res.* **2003,** *31* (21), 6365-6372.
157. Jørgensen, A. S.; Hansen, L. H.; Vester, B.; Wengel, J., Improvement of a streptavidin-binding aptamer by LNA- and α-l-LNA-substitutions. *Bioorg. Med. Chem. Lett.* **2014,** *24* (10), 2273-2277.
158. Saneyoshi, H.; Mazzini, S.; Aviñó, A.; Portella, G.; González, C.; Orozco, M.; Marquez, V. E.; Eritja, R., Conformationally rigid nucleoside probes help understand the role of sugar pucker and nucleobase orientation in the thrombin-binding aptamer. *Nucleic Acids Res.* **2009,** *37* (17), 5589-5601.
159. Hoshika, S.; Minakawa, N.; Matsuda, A., Synthesis and physical and physiological properties of 4′-thioRNA: application to post-modification of RNA aptamer toward NF-κB. *Nucleic Acids Res.* **2004,** *32* (13), 3815-3825.
160. Kato, Y.; Minakawa, N.; Komatsu, Y.; Kamiya, H.; Ogawa, N.; Harashima, H.; Matsuda, A., New NTP analogs: the synthesis of 4′-thioUTP and 4′-thioCTP and their utility for SELEX. *Nucleic Acids Res.* **2005,** *33* (9), 2942-2951.
161. Mendelboum Raviv, S.; Horvath, A.; Aradi, J.; Bagoly, Z.; Fazakas, F.; Batta, Z.; Muszbek, L.; Harsfalvi, J., 4-Thio-deoxyuridylate-modified thrombin aptamer and its inhibitory effect on

fibrin clot formation, platelet aggregation and thrombus growth on subendothelial matrix. *J. Thromb. Haemost.* **2008,** *6* (10), 1764-1771.
162. Zhang, L.; Chaput, J. C., In Vitro Selection of an ATP-Binding TNA Aptamer. *Molecules* **2020,** *25* (18), 4194.
163. Mei, H.; Liao, J.-Y.; Jimenez, R. M.; Wang, Y.; Bala, S.; McCloskey, C.; Switzer, C.; Chaput, J. C., Synthesis and evolution of a threose nucleic acid aptamer bearing 7-deaza-7-substituted guanosine residues. *J. Am. Chem. Soc.* **2018,** *140* (17), 5706-5713.
164. Varada, M.; Aher, M.; Erande, N.; Kumar, V. A.; Fernandes, M., Methoxymethyl Threofuranosyl Thymidine (4′-MOM-TNA-T) at the T7 Position of the Thrombin-Binding Aptamer Boosts Anticoagulation Activity, Thermal Stability, and Nuclease Resistance. *ACS Omega* **2020,** *5* (1), 498-506.
165. Li, X.; Li, Z.; Yu, H., Selection of threose nucleic acid aptamers to block PD-1/PD-L1 interaction for cancer immunotherapy. *Chem. Commun.* **2020,** *56* (93), 14653-14656.
166. Polak, M.; Manoharan, M.; Inamati, G. B.; Plavec, J., Tuning of conformational preorganization in model 2′,5′- and 3′,5′-linked oligonucleotides by 3′- and 2′-O-methoxyethyl modification. *Nucleic Acids Res.* **2003,** *31* (8), 2066-2076.
167. Gunjal, A. D.; Fernandes, M.; Erande, N.; Rajamohanan, P.; Kumar, V. A., Functional isoDNA aptamers: modified thrombin binding aptamers with a 2′-5′-linked sugar-phosphate backbone (isoTBA). *Chem. Commun.* **2014,** *50* (5), 605-607.
168. Herdewijn, P., Nucleic acids with a six-membered ‘carbohydrate’mimic in the backbone. *Chem. Biodivers.* **2010,** *7* (1), 1-59.
169. Kolb, G.; Reigadas, S.; Boiziau, C.; van Aerschot, A.; Arzumanov, A.; Gait, M. J.; Herdewijn, P.; Toulmé, J.-J., Hexitol Nucleic Acid-Containing Aptamers Are Efficient Ligands of HIV-1 TAR RNA. *Biochemistry* **2005,** *44* (8), 2926-2933.
170. Eremeeva, E.; Fikatas, A.; Margamuljana, L.; Abramov, M.; Schols, D.; Groaz, E.; Herdewijn, P., Highly stable hexitol based XNA aptamers targeting the vascular endothelial growth factor. *Nucleic Acids Res.* **2019,** *47* (10), 4927-4939.
171. Ackermann, D.; Pitsch, S., Synthesis and Pairing Properties of 3′-Deoxyribopyranose (4′→ 2′)-Oligonucleotides (‘p-DNA'). *Helv. Chim. Acta* **2002,** *85* (5), 1443-1462.
172. Schlönvogt, I.; Pitsch, S.; Lesueur, C.; Eschenmoser, A.; Jaun, B.; Wolf, R. M., Pyranosyl-RNA (‘p-RNA’): NMR and molecular-dynamics study of the duplex formed by self-pairing of ribopyranosyl-(C-G-A-A-T-T-C-G). *Helv. Chim. Acta* **1996,** *79* (8), 2316-2345.
173. Ackermann, D.; Wu, X.; Pitsch, S., 3′-Deoxyribopyranose (4′→ 2′)-Oligonucleotide Duplexes (= p-DNA Duplexes) as Substitutes of RNA Hairpin Structures. *Helv. Chim. Acta* **2002,** *85* (5), 1463-1478.
174. Horhota, A. T.; Szostak, J. W.; McLaughlin, L. W., Glycerol Nucleoside Triphosphates: Synthesis and Polymerase Substrate Activities. *Org. Lett.* **2006,** *8* (23), 5345-5347.
175. Rivieccio, E.; Tartaglione, L.; Esposito, V.; Dell'Aversano, C.; Koneru, P. C.; Scuotto, M.; Virgilio, A.; Mayol, L.; Kvaratskhelia, M.; Varra, M., Structural studies and biological evaluation of T30695 variants modified with single chiral glycerol-T reveal the importance of LEDGF/p75 for the aptamer anti-HIV-integrase activities. *Biochim. Biophys. Acta Gen. Subj.* **2019,** *1863* (2), 351-361.

176. Virgilio, A.; Petraccone, L.; Vellecco, V.; Bucci, M.; Varra, M.; Irace, C.; Santamaria, R.; Pepe, A.; Mayol, L.; Esposito, V.; Galeone, A., Site-specific replacement of the thymine methyl group by fluorine in thrombin binding aptamer significantly improves structural stability and anticoagulant activity. *Nucleic Acids Res.* **2015,** *43* (22), 10602-10611.
177. Virgilio, A.; Petraccone, L.; Scuotto, M.; Vellecco, V.; Bucci, M.; Mayol, L.; Varra, M.; Esposito, V.; Galeone, A., 5-Hydroxymethyl-2'-Deoxyuridine Residues in the Thrombin Binding Aptamer: Investigating Anticoagulant Activity by Making a Tiny Chemical Modification. *ChemBioChem* **2014,** *15* (16), 2427-2434.
178. Jensen, K. B.; Atkinson, B. L.; Willis, M. C.; Koch, T. H.; Gold, L., Using in vitro selection to direct the covalent attachment of human immunodeficiency virus type 1 Rev protein to high-affinity RNA ligands. *Proc. Natl. Acad. Sci. U.S.A.* **1995,** *92* (26), 12220-12224.
179. Vaish, N. K.; Larralde, R.; Fraley, A. W.; Szostak, J. W.; McLaughlin, L. W., A novel, modification-dependent ATP-binding aptamer selected from an RNA library incorporating a cationic functionality. *Biochemistry* **2003,** *42* (29), 8842-8851.
180. Kabza, A. M.; Sczepanski, J. T., An l-RNA Aptamer with Expanded Chemical Functionality that Inhibits MicroRNA Biogenesis. *ChemBioChem* **2017,** *18* (18), 1824-1827.
181. Wang, H.; Li, X.; Lai, L. A.; Brentnall, T. A.; Dawson, D. W.; Kelly, K. A.; Chen, R.; Pan, S., X-aptamers targeting Thy-1 membrane glycoprotein in pancreatic ductal adenocarcinoma. *Biochimie* **2021,** *181*, 25-33.
182. Dolot, R.; Lam, C. H.; Sierant, M.; Zhao, Q.; Liu, F.-W.; Nawrot, B.; Egli, M.; Yang, X., Crystal structures of thrombin in complex with chemically modified thrombin DNA aptamers reveal the origins of enhanced affinity. *Nucleic Acids Res.* **2018,** *46* (9), 4819-4830.
183. Mehedi Masud, M.; Kuwahara, M.; Ozaki, H.; Sawai, H., Sialyllactose-binding modified DNA aptamer bearing additional functionality by SELEX. *Bioorg. Med. Chem.***2004,** *12* (5), 1111-1120.
184. Shoji, A.; Kuwahara, M.; Ozaki, H.; Sawai, H., Modified DNA aptamer that binds the (R)-isomer of a thalidomide derivative with high enantioselectivity. *J. Am. Chem. Soc.* **2007,** *129* (5), 1456-1464.
185. Ozaki, H.; Nakamura, A.; Arai, M.; Endo, M.; Sawai, H., Novel C5-substituted 2'-deoxyuridine derivatives bearing amino-linker arms: synthesis, incorporation into oligodeoxyribonucleotides, and their hybridization properties. *Bull. Chem. Soc. Jpn.* **1995,** *68* (7), 1981-1987.
186. Latham, J. A.; Johnson, R.; Toole, J. J., The application of a modified nucleotide in aptamer selection: novel thrombin aptamers containing-(1-pentynyl)-2'-deoxyuridine. *Nucleic Acids Res.* **1994,** *22* (14), 2817-2822.
187. Tateishi-Karimata, H.; Ohyama, T.; Muraoka, T.; Podbevsek, P.; Wawro, A. M.; Tanaka, S.; Nakano, S.-i.; Kinbara, K.; Plavec, J.; Sugimoto, N., Newly characterized interaction stabilizes DNA structure: oligoethylene glycols stabilize G-quadruplexes CH–π interactions. *Nucleic Acids Res.* **2017,** *45* (12), 7021-7030.
188. Zavyalova, E.; Turashev, A.; Novoseltseva, A.; Legatova, V.; Antipova, O.; Savchenko, E.; Balk, S.; Golovin, A.; Pavlova, G.; Kopylov, A., Pyrene-Modified DNA Aptamers with High Affinity to Wild-Type EGFR and EGFRvIII. *Nucleic Acid Ther.* **2020,** *30* (3), 175-187.

189. Li, M.; Lin, N.; Huang, Z.; Du, L.; Altier, C.; Fang, H.; Wang, B., Selecting aptamers for a glycoprotein through the incorporation of the boronic acid moiety. *J. Am. Chem. Soc.* **2008,** *130* (38), 12636-12638.
190. Cheung, Y.-W.; Röthlisberger, P.; Mechaly, A. E.; Weber, P.; Levi-Acobas, F.; Lo, Y.; Wong, A. W. C.; Kinghorn, A. B.; Haouz, A.; Savage, G. P.; Hollenstein, M.; Tanner, J. A., Evolution of abiotic cubane chemistries in a nucleic acid aptamer allows selective recognition of a malaria biomarker. *Proc. Natl. Acad. Sci. U.S.A.* **2020,** *117* (29), 16790-16798.
191. Gold, L.; Ayers, D.; Bertino, J.; Bock, C.; Bock, A.; Brody, E.; Carter, J.; Cunningham, V.; Dalby, A.; Eaton, B., Aptamer-based multiplexed proteomic technology for biomarker discovery. *Nat. Preced.* **2010**, 1-1.
192. Mehan, M. R.; Ostroff, R.; Wilcox, S. K.; Steele, F.; Schneider, D.; Jarvis, T. C.; Baird, G. S.; Gold, L.; Janjic, N., Highly Multiplexed Proteomic Platform for Biomarker Discovery, Diagnostics, and Therapeutics. In *Complement Therapeutics*, Lambris, J. D.; Holers, V. M.; Ricklin, D., Eds. Springer US: New York, NY, 2013; pp 283-300.
193. Brody, E.; Gold, L.; Mehan, M.; Ostroff, R.; Rohloff, J.; Walker, J.; Zichi, D., Life's Simple Measures: Unlocking the Proteome. *J. Mol. Biol.* **2012,** *422* (5), 595-606.
194. Kuwahara, M.; Hanawa, K.; Ohsawa, K.; Kitagata, R.; Ozaki, H.; Sawai, H., Direct PCR amplification of various modified DNAs having amino acids: Convenient preparation of DNA libraries with high-potential activities for in vitro selection. *Bioorg. Med. Chem.***2006,** *14* (8), 2518-2526.
195. Ohsawa, K.; Kasamatsu, T.; Nagashima, J.-i.; Hanawa, K.; Kuwahara, M.; Ozaki, H.; Sawai, H., Arginine-modified DNA aptamers that show enantioselective recognition of the dicarboxylic acid moiety of glutamic acid. *Analytical Sciences* **2008,** *24* (1), 167-172.
196. Wu, D.; Feagin, T.; Mage, P.; Rangel, A.; Wan, L.; Li, A.; Coller, J.; Eisenstein, M.; Pitteri, S.; Soh, H. T., Automated platform for high-throughput screening of base-modified aptamers for affinity and specificity. *bioRxiv* **2020**, 2020.04.25.060004.
197. Imaizumi, Y.; Kasahara, Y.; Fujita, H.; Kitadume, S.; Ozaki, H.; Endoh, T.; Kuwahara, M.; Sugimoto, N., Efficacy of base-modification on target binding of small molecule DNA aptamers. *J. Am. Chem. Soc.* **2013,** *135* (25), 9412-9419.
198. Smirnov, I.; Kolganova, N.; Troisi, R.; Sica, F.; Timofeev, E., Expanding the recognition interface of the thrombin-binding aptamer HD1 through modification of residues T3 and T12. *Mol. Ther. Nucleic Acids* **2021,** *23*, 863-871.
199. Krawczyk, S. H.; Bischofberger, N.; Griffin, L. C.; Law, V. S.; Shea, R. G.; Swaminathan, S., Structure-activity study of oligodeoxynucleotides which inhibit thrombin. *Nucleosides Nucleotides Nucleic Acids* **1995,** *14* (3-5), 1109-1116.
200. Wang, K. Y.; Krawczyk, S. H.; Bischofberger, N.; Swaminathan, S.; Bolton, P. H., The tertiary structure of a DNA aptamer which binds to and inhibits thrombin determines activity. *Biochemistry* **1993,** *32* (42), 11285-11292.
201. He, G.-X.; Krawczyk, S. H.; Swaminathan, S.; Shea, R. G.; Dougherty, J. P.; Terhorst, T.; Law, V. S.; Griffin, L. C.; Coutré, S.; Bischofberger, N., N2- and C8-Substituted Oligodeoxynucleotides with Enhanced Thrombin Inhibitory Activity in Vitro and in Vivo. *J. Med. Chem.* **1998,** *41* (13), 2234-2242.

202. Goji, S.; Matsui, J., Direct Detection of Thrombin Binding to 8-Bromodeoxyguanosine-Modified Aptamer: Effects of Modification on Affinity and Kinetics. *Journal of Nucleic Acids* **2011,** *2011*, 316079.
203. Aviñó, A.; Mazzini, S.; Fàbrega, C.; Peñalver, P.; Gargallo, R.; Morales, J. C.; Eritja, R., The effect of l-thymidine, acyclic thymine and 8-bromoguanine on the stability of model G-quadruplex structures. *Biochim. Biophys. Acta Gen. Subj.* **2017,** *1861* (5, Part B), 1205-1212.
204. López de la Osa, J.; González, C.; Gargallo, R.; Rueda, M.; Cubero, E.; Orozco, M.; Aviñó, A.; Eritja, R., Destabilization of Quadruplex DNA by 8-Aminoguanine. *ChemBioChem* **2006,** *7* (1), 46-48.
205. Sproviero, M.; Manderville, R. A., Harnessing G-tetrad scaffolds within G-quadruplex forming aptamers for fluorescence detection strategies. *Chem. Commun.* **2014,** *50* (23), 3097-3099.
206. Blanchard, D. J.; Fadock, K. L.; Sproviero, M.; Deore, P. S.; Cservenyi, T. Z.; Manderville, R. A.; Sharma, P.; Wetmore, S. D., Photophysical properties of push–pull 8-aryl-deoxyguanosine probes within duplex and G-quadruplex structures. *J. Mater. Chem. C* **2016,** *4* (14), 2915-2924.
207. Marathias, V. M.; Sawicki, M. J.; Bolton, P. H., 6-Thioguanine alters the structure and stability of duplex DNA and inhibits quadruplex DNA formation. *Nucleic Acids Res.* **1999,** *27* (14), 2860-2867.
208. Chaput, J. C.; Switzer, C., A DNA pentaplex incorporating nucleobase quintets. *Proc. Natl. Acad. Sci. U.S.A.* **1999,** *96* (19), 10614-10619.
209. Zhang, L.; Yang, Z.; Le Trinh, T.; Teng, I. T.; Wang, S.; Bradley, K. M.; Hoshika, S.; Wu, Q.; Cansiz, S.; Rowold, D. J., Aptamers against cells overexpressing glypican 3 from expanded genetic systems combined with cell engineering and laboratory evolution. *Angew. Chem.* **2016,** *128* (40), 12560-12563.
210. Kimoto, M.; Yamashige, R.; Matsunaga, K.-i.; Yokoyama, S.; Hirao, I., Generation of high-affinity DNA aptamers using an expanded genetic alphabet. *Nat. Biotech.* **2013,** *31* (5), 453-457.
211. Gallego, J.; Loakes, D., Solution structure and dynamics of DNA duplexes containing the universal base analogues 5-nitroindole and 5-nitroindole 3-carboxamide. *Nucleic Acids Res.* **2007,** *35* (9), 2904-2912.
212. Tsvetkov, V. B.; Varizhuk, A. M.; Pozmogova, G. E.; Smirnov, I. P.; Kolganova, N. A.; Timofeev, E. N., A Universal Base in a Specific Role: Tuning up a Thrombin Aptamer with 5-Nitroindole. *Sci. Rep.* **2015,** *5* (1), 16337.
213. Veronese, F. M.; Pasut, G., PEGylation, successful approach to drug delivery. *Drug Discov. Today* **2005,** *10* (21), 1451-1458.
214. Ng, E. W. M.; Shima, D. T.; Calias, P.; Cunningham, E. T.; Guyer, D. R.; Adamis, A. P., Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nat. Rev. Drug Discov.* **2006,** *5* (2), 123-132.
215. Kazuhiko Haruta, N. O., Masakazu Nagamine, Tomoyoshi Kayo, Asako Sasaki, Shinsuke Hiramoto, Masayuki Takahashi, Kuniyoshi Hota, Hideaki Sato, and Hiroaki Yamazaki, A Novel PEGylation Method for Improving the Pharmacokinetic Properties of Anti-Interleukin-17A RNA Aptamers. *Nucleic Acid Ther.* **2017,** *27* (1), 36-44.

216. Kulkarni, O.; Pawar, R. D.; Purschke, W.; Eulberg, D.; Selve, N.; Buchner, K.; Ninichuk, V.; Segerer, S.; Vielhauer, V.; Klussmann, S., Spiegelmer inhibition of CCL2/MCP-1 ameliorates lupus nephritis in MRL-(Fas) lpr mice. *J. Am. Soc. Nephrol.* **2007,** *18* (8), 2350-2358.
217. de Smidt, P. C.; Doan, T. L.; Falco, S. d.; Berkel, T. J. C. v., Association of antisense oligonucleotides with lipoproteins prolongs the plasma half-life and modifies the tissue distribution. *Nucleic Acids Res.* **1991,** *19* (17), 4695-4700.
218. Lee, C. H.; Lee, S.-H.; Kim, J. H.; Noh, Y.-H.; Noh, G.-J.; Lee, S.-W., Pharmacokinetics of a Cholesterol-conjugated Aptamer Against the Hepatitis C Virus (HCV) NS5B Protein. *Mol. Ther. Nucleic Acids* **2015,** *4*, e254.
219. Nsairat, H.; Mahmoud, I. S.; Odeh, F.; Abuarqoub, D.; Al-Azzawi, H.; Zaza, R.; Qadri, M. I.; Ismail, S.; Al Bawab, A.; Awidi, A., Grafting of anti-nucleolin aptamer into preformed and remotely loaded liposomes through aptamer-cholesterol post-insertion. *RSC Adv.* **2020,** *10* (59), 36219-36229.
220. Willis, M. C.; Collins, B.; Zhang, T.; Green, L. S.; Sebesta, D. P.; Bell, C.; Kellogg, E.; Gill, S. C.; Magallanez, A.; Knauer, S., Liposome-anchored vascular endothelial growth factor aptamers. *Bioconj. Chemistry* **1998,** *9* (5), 573-582.
221. Dougan, H.; Lyster, D. M.; Vo, C. V.; Stafford, A.; Weitz, J. I.; Hobbs, J. B., Extending the lifetime of anticoagulant oligodeoxynucleotide aptamers in blood. *Nucl. Med. Biol.* **2000,** *27* (3), 289-297.
222. Gupta, S.; Hirota, M.; Waugh, S. M.; Murakami, I.; Suzuki, T.; Muraguchi, M.; Shibamori, M.; Ishikawa, Y.; Jarvis, T. C.; Carter, J. D., Chemically modified DNA aptamers bind interleukin-6 with high affinity and inhibit signaling by blocking its interaction with interleukin-6 receptor. *J. Biol. Chem.* **2014,** *289* (12), 8706-8719.
223. Beigelman, L.; McSwiggen, J. A.; Draper, K. G.; Gonzalez, C.; Jensen, K.; Karpeisky, A. M.; Modak, A. S.; Matulic-Adamic, J.; DiRenzo, A. B.; Haeberli, P.; Sweedler, D.; Tracz, D.; Grimm, S.; Wincott, F. E.; Thackray, V. G.; Usman, N., Chemical Modification of Hammerhead Ribozymes: Catalytic Activity and Nuclease Resisitance. *J. Biol. Chem.* **1995,** *270* (43), 25702-25708.
224. Martino, L.; Virno, A.; Randazzo, A.; Virgilio, A.; Esposito, V.; Giancola, C.; Bucci, M.; Cirino, G.; Mayol, L., A new modified thrombin binding aptamer containing a 5′–5′ inversion of polarity site. *Nucleic Acids Res.* **2006,** *34* (22), 6653-6662.
225. Esposito, V.; Galeone, A.; Mayol, L.; Randazzo, A.; Virgilio, A.; Virno, A., A mini-library of TBA analogues containing 3′-3′ and 5′-5′ inversion of polarity sites. *Nucleosides Nucleotides Nucleic Acids* **2007,** *26* (8-9), 1145-1149.
226. Esposito, V.; Scuotto, M.; Capuozzo, A.; Santamaria, R.; Varra, M.; Mayol, L.; Virgilio, A.; Galeone, A., A straightforward modification in the thrombin binding aptamer improving the stability, affinity to thrombin and nuclease resistance. *Org. Biomol. Chem.* **2014,** *12* (44), 8840-8843.
227. Matsunaga, K.-i.; Kimoto, M.; Hanson, C.; Sanford, M.; Young, H. A.; Hirao, I., Architecture of high-affinity unnatural-base DNA aptamers toward pharmaceutical applications. *Sci. Rep.* **2015,** *5* (1), 1-7.
228. Liu, M.; Yin, Q.; Chang, Y.; Zhang, Q.; Brennan, J. D.; Li, Y., In vitro selection of circular DNA aptamers for biosensing applications. *Angew. Chem. Int. Ed.* **2019,** *58* (24), 8013-8017.

229. Astatke, M.; Ng, K.; Grindley, N. D.; Joyce, C. M., A single side chain prevents Escherichia coli DNA polymerase I (Klenow fragment) from incorporating ribonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* **1998,** *95* (7), 3402-3407.
230. Patel, P. H.; Loeb, L. A., Multiple amino acid substitutions allow DNA polymerases to synthesize RNA. *J. Biol. Chem.* **2000,** *275* (51), 40266-40272.
231. Mi, J.; Zhang, X.; Giangrande, P. H.; McNamara, J. O.; Nimjee, S. M.; Sarraf-Yazdi, S.; Sullenger, B. A.; Clary, B. M., Targeted inhibition of αvβ3 integrin with an RNA aptamer impairs endothelial cell growth and survival. *Biochem. Biophys. Res. Commun.* **2005,** *338* (2), 956-963.
232. Padilla, R.; Sousa, R., Efficient synthesis of nucleic acids heavily modified with non-canonical ribose 2'-groups using a mutant T7 RNA polymerase (RNAP). *Nucleic Acids Res.* **1999,** *27* (6), 1561-1563.
233. Tolle, F.; Brändle, G. M.; Matzner, D.; Mayer, G., A versatile approach towards nucleobase-modified aptamers. *Angew. Chem. Int. Ed.* **2015,** *54* (37), 10971-10974.
234. Gordon, C. K. L.; Wu, D.; Pusuluri, A.; Feagin, T. A.; Csordas, A. T.; Eisenstein, M. S.; Hawker, C. J.; Niu, J.; Soh, H. T., Click-Particle Display for Base-Modified Aptamer Discovery. *ACS Chem. Biol.* **2019,** *14* (12), 2652-2662.
235. Wang, Z.; Xu, W.; Liu, L.; Zhu, T. F., A synthetic molecular system capable of mirror-image genetic replication and transcription. *Nat. Chem.* **2016,** *8* (7), 698-704.
236. Pinheiro, V. B.; Taylor, A. I.; Cozens, C.; Abramov, M.; Renders, M.; Zhang, S.; Chaput, J. C.; Wengel, J.; Peak-Chew, S.-Y.; McLaughlin, S. H., Synthetic genetic polymers capable of heredity and evolution. *Science* **2012,** *336* (6079), 341-344.
237. Pinheiro, V. B.; Arangundy-Franklin, S.; Holliger, P., Compartmentalized Self-Tagging for In Vitro-Directed Evolution of XNA Polymerases. *Curr. Protoc. Nucleic Acid Chem.* **2014,** *57*, 9.9.1-18.
238. Yu, H.; Zhang, S.; Chaput, J. C., Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. *Nat. Chem.* **2012,** *4* (3), 183-187.
239. Ichida, J. K.; Horhota, A.; Zou, K.; McLaughlin, L. W.; Szostak, J. W., High fidelity TNA synthesis by Therminator polymerase. *Nucleic Acids Res.* **2005,** *33* (16), 5219-5225.
240. Eaton, B. E.; Gold, L.; Hicke, B. J.; Janjié, N.; Jucker, F. M.; Sebesta, D. P.; Tarasow, T. M.; Willis, M. C.; Zichi, D. A., Post-SELEX combinatorial optimization of aptamers. *Bioorg. Med. Chem.***1997,** *5* (6), 1087-1096.
241. Kong, D.; Yeung, W.; Hili, R., In Vitro Selection of Diversely Functionalized Aptamers. *J. Am. Chem. Soc.* **2017,** *139* (40), 13977-13980.
242. Guo, C.; Hili, R., Fidelity of the DNA ligase-catalyzed scaffolding of peptide fragments on nucleic acid polymers. *Bioconj. Chem.* **2017,** *28* (2), 314-318.
243. Sefah, K.; Yang, Z.; Bradley, K. M.; Hoshika, S.; Jiménez, E.; Zhang, L.; Zhu, G.; Shanker, S.; Yu, F.; Turek, D., In vitro selection with artificial expanded genetic information systems. *Proc. Natl. Acad. Sci. U.S.A.* **2014,** *111* (4), 1449-1454.
244. Yang, Z.; Chen, F.; Alvarado, J. B.; Benner, S. A., Amplification, mutation, and sequencing of a six-letter synthetic genetic system. *J. Am. Chem. Soc.* **2011,** *133* (38), 15105-15112.
245. Hirao, I.; Kimoto, M.; Lee, K. H., DNA aptamer generation by ExSELEX using genetic alphabet expansion with a mini-hairpin DNA stabilization method. *Biochimie* **2018,** *145*, 15-21.

246. Yamashige, R.; Kimoto, M.; Takezawa, Y.; Sato, A.; Mitsui, T.; Yokoyama, S.; Hirao, I., Highly specific unnatural base pair systems as a third base pair for PCR amplification. *Nucleic Acids Res.* **2012,** *40* (6), 2793-2806.
247. Matsunaga, K.-i.; Kimoto, M.; Hirao, I., High-affinity DNA aptamer generation targeting von Willebrand factor A1-domain by genetic alphabet expansion for systematic evolution of ligands by exponential enrichment using two types of libraries composed of five different bases. *J. Am. Chem. Soc.* **2017,** *139* (1), 324-334.
248. Berezovski, M.; Musheev, M.; Drabovich, A.; Krylov, S. N., Non-SELEX Selection of Aptamers. *J. Am. Chem. Soc.* **2006,** *128* (5), 1410-1411.
249. Ashley, J.; Ji, K.; Li, S. F., Selection of bovine catalase aptamers using non-SELEX. *Electrophoresis* **2012,** *33* (17), 2783-2789.
250. Szeto, K.; Latulippe, D. R.; Ozer, A.; Pagano, J. M.; White, B. S.; Shalloway, D.; Lis, J. T.; Craighead, H. G., Rapid-SELEX for RNA aptamers. *PloS One* **2013,** *8* (12), e82667.
251. Kushwaha, A.; Takamura, Y.; Nishigaki, K.; Biyani, M., Competitive non-SELEX for the selective and rapid enrichment of DNA aptamers and its use in electrochemical aptasensor. *Sci. Rep.* **2019,** *9* (1), 6642.
252. Kim, H. R.; Song, M. Y.; Chan Kim, B., Rapid isolation of bacteria-specific aptamers with a non-SELEX-based method. *Anal. Biochem.* **2020,** *591*, 113542.
253. Hamula, C. L.; Zhang, H.; Guan, L. L.; Li, X.-F.; Le, X. C., Selection of aptamers against live bacterial cells. *Anal. Chem.* **2008,** *80* (20), 7812-7819.
254. Nitsche, A.; Kurth, A.; Dunkhorst, A.; Pänke, O.; Sielaff, H.; Junge, W.; Muth, D.; Scheller, F.; Stöcklein, W.; Dahmen, C., One-step selection of Vaccinia virus-binding DNA aptamers by MonoLEX. *BMC Biotech.* **2007,** *7* (1), 1-12.
255. Peng, L.; Stephens, B. J.; Bonin, K.; Cubicciotti, R.; Guthold, M., A combined atomic force/fluorescence microscopy technique to select aptamers in a single cycle from a small pool of random oligonucleotides. *Microsc. Res. Tech.* **2007,** *70* (4), 372-381.
256. Kupakuwana, G. V.; Crill II, J. E.; McPike, M. P.; Borer, P. N., Acyclic identification of aptamers for human alpha-thrombin using over-represented libraries and deep sequencing. *PloS one* **2011,** *6* (5), e19395.
257. Hoon, S.; Zhou, B.; Janda, K. D.; Brenner, S.; Scolnick, J., Aptamer selection by high-throughput sequencing and informatic analysis. *BioTechniques* **2011,** *51* (6), 413-416.
258. Liu, Y.; Wang, C.; Li, F.; Shen, S.; Tyrrell, D. L. J.; Le, X. C.; Li, X.-F., DNase-Mediated Single-Cycle Selection of Aptamers for Proteins Blotted on a Membrane. *Anal. Chem.* **2012,** *84* (18), 7603-7606.
259. Galas, D. J.; Schmitz, A., DNAase footprinting a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res.* **1978,** *5* (9), 3157-3170.
260. Lokesh, G. L.; Wang, H.; Lam, C. H.; Thiviyanathan, V.; Ward, N.; Gorenstein, D. G.; Volk, D. E., X-Aptamer Selection and Validation. *Method. Mol. Biol.* **2017,** *1632*, 151-174.
261. Yang, X.; Bassett, S. E.; Li, X.; Luxon, B. A.; Herzog, N. K.; Shope, R. E.; Aronson, J.; Prow, T. W.; Leary, J. F.; Kirby, R.; Ellington, A. D.; Gorenstein, D. G., Construction and selection of bead-bound combinatorial oligonucleoside phosphorothioate and phosphorodithioate aptamer libraries designed for rapid PCR-based sequencing. *Nucleic Acids Res.* **2002,** *30* (23), e132-e132.

262. Le, A. T. H.; Krylova, S. M.; Beloborodov, S. S.; Wang, T. Y.; Hili, R.; Johnson, P. E.; Li, F.; Veedu, R. N.; Belyanskaya, S.; Krylov, S. N., How to Develop and Prove High-Efficiency Selection of Ligands from Oligonucleotide Libraries: A Universal Framework for Aptamers and DNA-Encoded Small-Molecule Ligands. *Anal. Chem.* **2021,** *93* (13), 5343-5354.
263. Smith, G. P.; Petrenko, V. A., Phage Display. *Chem. Rev.* **1997,** *97* (2), 391-410.
264. Mannocci, L.; Leimbacher, M.; Wichert, M.; Scheuermann, J.; Neri, D., 20 years of DNA-encoded chemical libraries. *Chem. Commun.* **2011,** *47* (48), 12747-12753.
265. Brenner, S.; Lerner, R. A., Encoded combinatorial chemistry. *Proc. Natl. Acad. Sci. U.S.A.* **1992,** *89* (12), 5381-5383.
266. Halpin, D. R.; Harbury, P. B.; Joyce, G., DNA display I. Sequence-encoded routing of DNA populations. *PLoS Biol.* **2004,** *2* (7), e173.
267. Gartner, Z. J.; Liu, D. R., The generality of DNA-templated synthesis as a basis for evolving non-natural small molecules. *J. Am. Chem. Soc.* **2001,** *123* (28), 6961-6963.
268. Halpin, D. R.; Harbury, P. B.; Joyce, G., DNA display II. Genetic manipulation of combinatorial chemistry libraries for small-molecule evolution. *PLoS Biol.* **2004,** *2* (7), e174.
269. Gartner, Z. J.; Brian, N. T.; Grubina, R.; Doyon, J. B.; Snyder, T. M.; Liu, D. R., DNA-templated organic synthesis and selection of a library of macrocycles. *Science* **2004,** *305* (5690), 1601-1605.
270. Clark, M. A.; Acharya, R. A.; Arico-Muendel, C. C.; Belyanskaya, S. L.; Benjamin, D. R.; Carlson, N. R.; Centrella, P. A.; Chiu, C. H.; Creaser, S. P.; Cuozzo, J. W., Design, synthesis and selection of DNA-encoded small-molecule libraries. *Nat. Chem. Biol.* **2009,** *5* (9), 647-654.
271. de Rochambeau, D.; Sun, Y.; Barlog, M.; Bazzi, H. S.; Sleiman, H. F., Modular strategy to expand the chemical diversity of DNA and sequence-controlled polymers. *J. Org. Chem.* **2018,** *83* (17), 9774-9786.
272. de Rochambeau, D. Expanding the Scope of Sequence-Defined Oligo(Phosphodiester)s with Novel Building Blocks. McGill University, 2019.
273. Brown, T.; Brown, T. J., Solid-phase oligonucleotide synthesis. In *Nucleic Acids*, ATDBio Ltd.: atdbio.com, 2005.
274. Ravi, R. K.; Walton, K.; Khosroheidari, M., MiSeq: A Next Generation Sequencing Platform for Genomic Analysis. *Method. Mol. Biol.* **2018,** *1706*, 223-232.
275. Martin, M., Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **2011,** *17* (1), 3.
276. Agbavwe, C.; Kim, C.; Hong, D.; Heinrich, K.; Wang, T.; Somoza, M. M., Efficiency, error and yield in light-directed maskless synthesis of DNA microarrays. *J. Nanobiotech.* **2011,** *9* (1), 57.
277. Wheeler, S. E., Understanding substituent effects in noncovalent interactions involving aromatic rings. *Acc. Chem. Res.* **2013,** *46* (4), 1029-1038.
278. Yum, J. H.; Ishizuka, T.; Fukumoto, K.; Hori, D.; Bao, H.-L.; Xu, Y.; Sugiyama, H.; Park, S., Systematic Approach to DNA Aptamer Design Using Amino Acid–Nucleic Acid Hybrids (ANHs) Targeting Thrombin. *ACS Biomater. Sci. Eng.* **2021,** *7* (4), 1338-1343.
279. Zhang, H.; Wu, Q.; Berezin, M. Y., Fluorescence anisotropy (polarization): from drug screening to precision medicine. *Expert Opin. Drug Discov.* **2015,** *10* (11), 1145-1161.
280. Heller, G.; Aprile, F.; Vendruscolo, M., Methods of probing the interactions between small molecules and disordered proteins. *Cell. Mol. Life Sci.* **2017,** *74*.

281. Davies, D. R.; Gelinas, A. D.; Zhang, C.; Rohloff, J. C.; Carter, J. D.; O'Connell, D.; Waugh, S. M.; Wolk, S. K.; Mayfield, W. S.; Burgin, A. B.; Edwards, T. E.; Stewart, L. J.; Gold, L.; Janjic, N.; Jarvis, T. C., Unique motifs and hydrophobic interactions shape the binding of modified DNA ligands to protein targets. *Proc. Natl. Acad. Sci. U.S.A.* **2012,** *109* (49), 19971-19976.
282. Akalin, A., *Computational genomics with R*. CRC Press: 2020.
283. Gokulrangan, G.; Unruh, J. R.; Holub, D. F.; Ingram, B.; Johnson, C. K.; Wilson, G. S., DNA aptamer-based bioanalysis of IgE by fluorescence anisotropy. *Anal. Chem.* **2005,** *77* (7), 1963-1970.
284. Giamberardino, A.; Labib, M.; Hassan, E. M.; Tetro, J. A.; Springthorpe, S.; Sattar, S. A.; Berezovski, M. V.; DeRosa, M. C., Ultrasensitive norovirus detection using DNA aptasensor technology. *PloS one* **2013,** *8* (11).
285. Chen, Z.; Lichtor, P. A.; Berliner, A. P.; Chen, J. C.; Liu, D. R., Evolution of sequence-defined highly functionalized nucleic acid polymers. *Nat. Chem.* **2018,** *10* (4), 420-427.