Enhancing aptamer function and stability via in vitro selection using modified nucleic acids

Kirsten N. Meek,‡ Alexandra E. Rangel,‡ Jennifer M. Heemstra*
Department of Chemistry and the Center for Cell and Genome Science, University of Utah, 315 S 1400 E, Salt Lake City, Utah 84112, United States

Abstract
Nucleic acid aptamers have emerged as a promising alternative to antibodies for use as recognition elements in therapeutics, bioimaging, and analytical applications. A key benefit that aptamers possess relative to antibodies is their ability to be chemically synthesized. This advantage, coupled with the broad range of modified nucleotide building blocks that can be constructed using chemical synthesis, has enabled the discovery and development of modified aptamers having extraordinary affinity, specificity, and biostability. Early efforts to generate modified aptamers focused on selection of a native DNA or RNA aptamer, followed by post-selection trial-and-error testing of modifications. However, recent advances in polymerase engineering and templated nucleic acid synthesis have enabled the direct selection of aptamers having modified backbones and nucleobases. This review will discuss these technological advances and highlight the improvements in aptamer function that have been realized through in vitro selection of non-natural nucleic acids.

Keywords
Aptamer; Nucleic acid; In vitro selection; Affinity reagent; Polymerase engineering; Templated synthesis

1. Introduction
Nucleic acids are best known for their ability to encode, store, and decode genetic information in cells. However, the functional capacity of DNA and RNA stretches far beyond information storage to encompass complex tasks such as regulation and catalysis. These capabilities have emerged as a result of evolution over millions, or even billions, of years. Inspired by the ability of evolution to provide nucleic acids having this large diversity of functions, researchers have sought to replicate the evolutionary process in the laboratory. In 1990, the Szostak, Joyce, and Gold labs individually reported the successful implementation of selection and evolution in vitro,
and this process was termed Systematic Evolution of Ligands by EXponential enrichment (SELEX) [1-3]. Over the past 25 years, SELEX has been most broadly applied for the *in vitro* evolution of aptamers, which are nucleic acid sequences capable of binding to specific small-molecule or protein targets. Using the SELEX process, functional nucleic acid sequences can be enriched and evolved from a starting library on the time scale of days, providing a powerful route to rapidly generate biopolymers having new capabilities [4]. To date, over 1,000 well-characterized aptamers to over 550 target molecules have been reported in the literature, demonstrating the broad interest of the research community in this class of affinity reagents [5].

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

The development of Macugen began with the use of SELEX to generate an RNA aptamer for the vascular endothelial growth factor (VEGF) isoform VEGF_{165}, which regulates blood vessel growth in the eye, and thus is a primary target for the treatment of age-related macular degeneration [9]. After identification and truncation of the initial aptamer sequence, a number of modified sequences having 2’-O-Me, 2’-NH₂, and 2’-F modifications were generated and tested (Figure 1a) [10]. Modifications to the termini of the aptamer were also investigated as a method for increasing stability and circulation lifetime [6]. As a result of these modifications, the affinity, IC₅₀, and biological half-life were each improved by orders of magnitude, and the optimized aptamer structure was carried forward into testing and approved as Macugen (Figure 1b).
Figure 1. Backbone modifications used in post-selection testing of anti-VEGF<sub>165</sub> aptamer. (a) Chemical structure of RNA and modified backbones. (b) Sequence of Macugen. Red letters with circles are 2'-F pyrimidines and blue letters with squares are 2'-O-Me purines.

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

2. SELEX using backbone-modified nucleic acids

Aptamers have found use in a wide variety of applications including biosensing, affinity purification, therapeutics, and biological targeting. However, the low biological stability of DNA and RNA can pose limitations for the use of aptamers in these applications [11]. Because DNA and RNA are naturally occurring biomolecules, a number of nuclease enzymes have co-evolved to catalyze their degradation, which becomes especially problematic for the use of aptamers in vivo. Modification of the nucleic acid backbone can greatly reduce nuclease susceptibility, and this strategy has been successfully implemented to enable the use of nucleic acids in antisense therapy and targeted drug delivery. These backbone-modified nucleic acids are generally referred to as xeno nucleic acids (XNAs), where the “X” can be replaced by a specific letter (or letters) to indicate a particular modification (Figure 2). Also included in Figure 2 is L-DNA, which is the enantiomer of native D-DNA. L-DNA differs from XNAs in that it does not have a chemically modified backbone. But, due to its chiral modification, we still choose to classify L-DNA a modified nucleic acid. For antisense applications, XNA oligonucleotides having the desired activity can be generated with relative ease, as their activity relies upon Watson-Crick base pairing, and thus function is only minimally perturbed by the introduction of modifications [12]. However, as outlined above, post-synthetic modification of aptamers is significantly more challenging and time-intensive. A preferable approach is the direct selection of aptamers from XNA libraries, but this introduces new challenges, as the backbone modifications that prevent recognition by nucleases also make recognition by polymerases less efficient. Excitingly, great strides have been made recently in the field of polymerase engineering, which has enabled the development of polymerases that can both encode (transcribe) DNA into XNA, and decode (reverse transcribe) XNA back into DNA, enabling selection of aptamers from XNA libraries. This section will discuss the implementation of these engineered polymerases for the in vitro selection of XNA aptamers, and will highlight recent exciting developments in the area of non-enzymatic templated synthesis, which holds promise to eliminate the reliance on enzymes for encoding and decoding of sequences during SELEX.
Figure 2. Chemical structure of XNA backbones that have been explored for in vitro selection experiments.

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

2.2 Selection of XNA aptamers using engineered polymerases
As described above, a key challenge for the selection of XNA aptamers is that XNA nucleotides can be poor substrates for naturally occurring polymerases. However, a number of groups have devoted significant effort to screening commercially available polymerases, and at times engineering new polymerases, in order to efficiently encode and decode XNA sequences during the SELEX process.

2.2.1 Selection of XNA aptamers using polymerase-mediated encoding
One particularly promising XNA for use in biological applications is threose nucleic acid (TNA), which has been shown by Chaput and coworkers to possess excellent nuclease resistance [15]. The commercially available Therminator DNA polymerase (NEB) was found to be capable of encoding (transcribing) DNA libraries into TNA, at least in the context of TNA sequences having T, G, and diaminopurine (D, used in place of adenine) nucleotides [16]. Thus, the Chaput Lab designed a “DNA display” selection strategy that obviates the need for a decoding (reverse transcription) step, enabling the implementation of a full SELEX cycle (Figure 3a) [17]. To achieve this, a DNA hairpin was designed that contains both the library template and primer, and thus can be used to synthesize a TNA library via primer extension using Therminator. A DNA primer was then annealed in the loop region of the hairpin, and extension of this primer displaced the TNA, “displaying” it on a DNA duplex where it could then be screened for binding to the target molecule. Using this strategy, the genetic information of each TNA strand is maintained in the covalently attached DNA template, which can be directly PCR amplified after selection, circumventing the need to decode the selected XNA sequences. Using this strategy coupled with capillary electrophoresis to separate target-bound and free sequences, the Chaput Lab successfully selected a thrombin-binding TNA aptamer having a $K_D$ value of 200 nM [16]. They have since shown that incorporation of 7-deazaguanine in the DNA template allows for the transcription of four-letter TNA libraries, which will enable the selection of TNA aptamers that take advantage of the full scope of Watson-Crick base pairing [18]. Additionally, the Chaput Lab have recently discovered that commercially available SuperScript II polymerase is capable of reverse transcribing TNA into DNA, which will enable selections to be carried out in a manner analogous to RNA SELEX [18].
2.2.2 Selection of XNA aptamers using polymerase-mediated encoding and decoding

A key advance in the selection of XNA aptamers was realized with the development of compartmentalized self-tagging (CST) by Holliger and coworkers, which enables the rapid evolution of polymerases for XNA substrates \cite{19}. Using this method, a series of enzymes was evolved that are collectively capable of both transcription and reverse transcription of 1,5-anhydrohexitol nucleic acid (HNA), cyclohexenyl nucleic acid (CeNA), 2’-O,4’-C-methylene-b-D-ribonucleic acid (locked nucleic acid, LNA), arabinonucleic acid (ANA), 2’-fluoro-arabino-nucleic acid (FANA), and TNA (Figure 2) \cite{20}. Because these enzymes are capable of both encoding and decoding XNA libraries, SELEX can be carried out as shown in Figure 3b. Much like RNA SELEX strategies, XNA selections using encoding and decoding polymerases begin with generation of an XNA library via transcription from a DNA library. Once purified away from the DNA template, a selection step can be employed to enrich the XNA library for sequences having the desired function, and these sequences can be reverse-transcribed back into DNA for amplification or sequencing.
Using the encode-select-decode strategy coupled with a standard bead-based enrichment method, Holliger and coworkers used their engineered polymerases to generate HNA aptamers for both HIV TAR RNA and Hen Egg Lysozyme (HEL) [20]. Both aptamers bind to their intended targets with good selectivity and $K_D$ values in the nanomolar range. Similarly, DeStephano and coworkers used CST-derived polymerases to select for a FANA aptamer that binds to HIV-1 reverse transcriptase (RT) with picomolar affinity [21]. This affinity was higher than that achieved by a DNA aptamer for the same target, demonstrating that XNA backbones can also provide increased affinity for the target, potentially through energetically favorable hydrophobic interactions. Interestingly, the enzyme used to transcribe DNA into FANA not only functions as an isothermal primer extension catalyst, but can also be employed in thermal cycling protocols similar to those used in PCR, suggesting that direct amplification of XNA may be possible in the future.

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

2.3 Non-enzymatic templated synthesis of genetically-encoded libraries

The selection methods described up to this point rely on engineered polymerases to encode and decode XNA sequence information. Liu and coworkers envisioned that the information stored in a DNA template could potentially be encoded into XNA, or even non-nucleic acid polymers without the use of enzymes, which would dramatically expand the structural diversity that could be explored in selection experiments. This principle was initially demonstrated for peptide nucleic acid (PNA), which is capable of Watson-Crick base pairing with DNA and RNA [24] but does not possess a phosphodiester backbone, and thus cannot be generated using even highly engineered polymerases. Similar to the encode-only “display” strategy, a self-priming DNA hairpin was used as a template, and PNA building blocks having sequences complementary to this template were annealed and covalently polymerized using reductive amination. To illustrate the potential utility of this method for in vitro selection, a mock selection
was performed in which one PNA building block was synthesized having a biotin, and sequences containing this building block were enriched from a random pool [25].

Taking one step further, Liu and coworkers envisioned that templated synthesis could be used to generate libraries of sequence-defined polymers that have no inherent affinity for the DNA template. This was accomplished using a design inspired by the ability of tRNAs to mediate the templated synthesis of peptides. As shown in Figure 4, PNA building blocks similar to those used in their previous experiments were equipped with a non-nucleic acid polymer segment via two cleavable linkers. In this case, the functional groups required for ligation were placed on the non-nucleic acid polymer, which enabled templated ligation of these segments to create the genetically encoded polymer. Subsequent cleavage of the linkers resulted in removal of the PNA adapter segments, allowing the polymer to be covalently displayed on its encoding DNA sequence [26]. While these templated polymerization strategies have yet to be successfully employed in the selection of aptamers, there is significant promise in their ability to encode genetic information into a structurally diverse range of sequence-defined polymers.

**Figure 4.** Non-enzymatic templated polymerization of non-nucleic acid polymers. Polymer segments are covalently attached to an “adaptor” nucleic acid fragment, which directs templated polymerization, but can be subsequently cleaved to reveal the encoded polymer library. While selection experiments have yet to be reported, regeneration of the enriched templates via PCR would enable iterative rounds of selection and amplification.

### 3. SELEX using modified nucleobases

Proteins have historically dominated the pool of available catalysts and affinity reagents, likely as a result of the diverse array of amino acid side chains that they can employ. In contrast, DNA and RNA are each comprised of only four nucleotide building blocks, and these all possess a relatively similar palette of functional groups. Aiming to generate aptamers that have similar affinity and specificity to protein-based receptors, a number of research groups have explored
the *in vitro* selection of aptamers using monomers having unnatural nucleobases. The earliest example of SELEX using modified nucleotides was presented by Toole and coworkers 1994, where an aptamer for thrombin was generated using a library in which all thymine monomers were replaced by 5-pentynyl-dU [27]. This aptamer was found to have a dissociation constant comparable to that of a previously selected DNA aptamer for thrombin, and showed unique secondary and tertiary structures. While this early example established precedence for the idea of utilizing unnatural nucleobases in SELEX, the true potential of this approach was demonstrated later by SomaLogic, Inc., who carried out selections in which thymine was replaced with a 2'-deoxyuridine (dUTP) having a hydrophobic functional group at the 5-position of the nucleobase [28]. Using libraries containing these monomers, high-affinity “SOMAmers” were generated for a number of challenging targets, and were subsequently incorporated into an aptamer-based biomarker detection platform capable of detection limits in the fM range. The exploration of SELEX using unnatural nucleobases has rapidly expanded in the past few years, and these efforts can be separated into two general approaches: (1) development of unnatural base pair systems that are orthogonal to the Watson-Crick base pairs, and (2) addition of peptide-like functional groups to native nucleobases. As is the case for *in vitro* selection using modified backbones, a key challenge to the use of unnatural nucleobases in SELEX is the requirement that the unnatural nucleotides be encoded and decoded during the selection process, and this often requires compatibility with polymerases. In the case of modified nucleobases, and additional challenge is that the selectivity of base pairing with or between modified bases must occur with high fidelity during the encoding and decoding steps. However, a variety of innovative methods have been recently developed to circumvent these challenges and expand the scope of nucleobase modifications that can be employed for the *in vitro* selection of aptamers.

### 3.1 Development of unnatural base pairs

Expansion of the genetic alphabet to include base pairs beyond naturally occurring A:T/U and G:C was first pioneered by Alex Rich in 1962, with his proposal that an orthogonal base pair could be formed between iso-G and iso-C [29]. The key requirements for the development of these new base pairs was that the two modified nucleobases must pair with one another, but not with the canonical nucleobases. Following up on this initial work, Benner and coworkers proposed three non-canonical donor-acceptor patterns, which could in principle be used in conjunction with the two Watson-Crick base pairs to provide an artificially expanded genetic information system (AEGIS) having five different base pairs [30]. Taking a different approach to
the design of unnatural base pairs, the Kool, Romesberg, and Hirao groups have each reported modified nucleotides that are able to form orthogonal base pairs through hydrophobic interactions rather than hydrogen bonding [31-33]. While these base pairs have been utilized for a number of innovative applications over the past several years, only recently have the technological hurdles required for their use in aptamer selections been overcome.

3.1.1 Aptamer selection using base pairs having alternative hydrogen bonding patterns
In the design of the AEGIS base pairs, the Benner Lab hypothesized that placing electron density on the minor groove face of each nucleobase would enable recognition by polymerases, allowing the enzymatic synthesis and amplification of oligonucleotides incorporating the AEGIS base pairs [34]. Among the proposed AEGIS base pairs, this requirement was most efficiently met by pyrimidine donor-donor-acceptor (pyDDA) and purine acceptor-acceptor-donor (puAAD) motifs. From this insight, the Z:P base pair was reported in 2006 (Figure 5), and was shown to be highly specific, exhibiting orthogonality to the natural nucleobases and having a favorable effect on overall duplex stability. Moreover, the Z-P base pair was recently confirmed to adopt Watson-Crick geometry, supporting the generation of a six nucleotide genetic alphabet [35].

Equipped with an unnatural base pair that was capable of enzymatic amplification, Benner and coworkers turned to exploring the use of Z:P in the selection of aptamers. Using a library containing a 1.5:1.2:1.0:1.0:1.0:0.5 ratio of T:G:A:C:Z:P, SELEX experiments were carried out against the MDA-MB-231 breast cancer cell line [36]. Because the AEGIS base pair is recognized by native polymerases, sequences that bound to the target cells were able to be directly PCR amplified using all six nucleotide triphosphates, and the resulting oligonucleotides carried into the subsequent selection round. After 11 rounds of selection, the resulting library was sequenced by dividing it into two fractions, which were independently barcoded. For each fraction, the Z and P nucleotides were converted into natural nucleotides using a different protocol. In the first fraction, the Z and P nucleotides were converted to C and G, respectively. In the second fraction, the Z nucleotides were converted to a mixture of C/T and the P nucleotides were converted to a mixture of A/G. These two fractions were then combined and subjected to next-generation sequencing, and alignment of the resulting sequences revealed the locations of the unnatural nucleotides in the selected pool. The dominant aptamer was found to have a binding affinity of 30 nM, and replacement of the Z:P base pairs with Watson-Crick base pairs greatly diminished this affinity, highlighting the important role of these modified nucleotides in the target-binding interaction.
Figure 5. Natural and unnatural DNA base pairs used for the *in vitro* selection of aptamers. The Z:P base pair utilizes an orthogonal donor-acceptor hydrogen bonding pattern, whereas the Ds:Px base pair relies on hydrophobic interactions.

### 3.1.2 Aptamer selection using hydrophobic base pairs

In 2006, Hirao and coworkers introduced the Ds:Pa hydrophobic base pair, and showed that the unnatural nucleotides could serve as substrates for natural polymerases [33, 37]. While the Ds-Pa system was successful in both replication and transcription experiments, in part owing to the unique shape complementarity between the hydrophobic bases, undesired Ds-Ds and A-Pa base pairs were observed, leading to reduced fidelity. Subsequently, 2-nitro-4-propynylpyrrole (Px) was synthesized as an optimized pairing partner for Ds, having nitro and propynyl groups to alter electrostatic interactions and increase hydrophobicity, respectively (Figure 5) [38]. The increased hydrophobicity of Px improved incorporation opposite Ds, suppressing Ds-Ds mispairing, and the electrostatic repulsion due to the nitro group discouraged A-Px mispairing.

Using the Ds:Px unnatural base pair system, aptamer selection experiments were carried out against vascular endothelial cell growth factor (VEGF<sub>165</sub>) and interferon-γ (IFN-γ) [7]. The hydrophobicity of Ds was hypothesized to improve interactions with hydrophobic regions of the protein targets, and to ensure access of the unnatural nucleotides to the protein target, Px was intentionally left out of the library, as this prevented the enrichment of sequences in which the Ds was sequestered into base pairs. Unlike the Benner Z:P unnatural base pair, a convenient
method for decoding the Ds:Px pairs after selection was not available. Thus, libraries were synthesized having 1-3 Ds nucleotides at fixed locations in the sequence, and these libraries were barcoded using a unique 2-3 nucleotide tag. The libraries were incubated with each protein target, and bound sequences were amplified in the presence of Ds and Px deoxyribonucleotide triphosphates (dDsTP and dPxTP) to prepare material for the subsequent selection round. After completion of the selection rounds, the libraries were amplified in the absence of dDsTP and dPxTP, which caused reversion of the unnatural base pairs to A:T base pairs, which could then be sequenced. As a control, selections for the two protein targets were also carried out using libraries having entirely native DNA nucleotides (contVG and contIF). The highest affinity aptamers for VEGF\textsubscript{165} and IFN-\gamma from the Ds-modified library had $K_D$ values of 1.69 pM and 0.124 pM, respectively. These affinities were dramatically higher than those obtained for the control aptamers, which had $K_D$ values of 370 pM and 3 nM for binding to VEGF\textsubscript{165} and IFN-\gamma, respectively (Table 1). Compared to the AEGIS approach, the Ds:Px base pair provides hydrophobicity that is generally lacking in the natural nucleobases, and the power of this is reflected in the high affinities obtained by the Hirao Lab. However, the current minor limitation of the Ds:Px base pair relative to AEGIS is that the sites of the Ds modifications must be fixed in the sequence, compared to the Z:P nucleotides, which can be distributed randomly throughout the library and decoded post-selection using sequencing.

Table 1. Comparison of target binding affinity for aptamers having only natural nucleobases and aptamers having modified Ds nucleobases.

<table>
<thead>
<tr>
<th>DNA library</th>
<th>VEG-F 165 ($K_D$)</th>
<th>IFN-\gamma ($K_D$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>natural nucleobases</td>
<td>370 pM</td>
<td>3 nM</td>
</tr>
<tr>
<td>natural + Ds nucleobases</td>
<td>1.69 pM</td>
<td>0.124 pM</td>
</tr>
</tbody>
</table>

### 3.2 Modified nucleotides having peptide-like functional groups

While aptamers have a number of benefits relative to antibodies, including increased thermal stability and their ability to be chemically synthesized, antibodies benefit significantly from the chemical diversity available in the repertoire of naturally occurring amino acid side chains. In the early 2000s, it was envisioned that grafting the functional groups present in amino acid side chains onto DNA nucleobases could lead to the evolution of aptamers or DNA enzymes (DNAzymes) having new or improved function. In 2001, Perrin and coworkers utilized a library
containing nucleobases modified with imidazole and amine side chains to generate a DNAzyme that functioned as an RNaseA mimic, and they later expanded these selections to include nucleobases having guanidine functional groups [39, 40]. Selections using a library containing a combination of imidazole, amine, and guanidine modifications resulted in the generation of a DNAzyme capable of cleaving a RNA at rates much higher than previously reported for unmodified DNAzymes. Moreover, this modified DNAzyme maintained function in the absence of divalent metal cations, which are typically required for catalytic nucleic acids. These experiments demonstrated that the addition of protein-like functional groups could greatly enhance the potential catalytic abilities of nucleic acids. In parallel with the efforts of the Perrin Lab, Famulok and coworkers demonstrated that DNA sequences could be enzymatically polymerized in which all four nucleotides were substituted with an unnatural functional group, paving the way for in vitro selection using libraries having a high density of amino acid-like functional groups.[41] While these specific libraries have yet to be utilized for aptamer selection experiments, the knowledge of which positions on the nucleobases can be modified without disrupting enzymatic polymerization proved to be critical for the selection of nucleobase-modified aptamers such as SOMAmers. Additionally, in recent years, innovative methods have been developed to further expand the diversity of functional groups that can be appended to nucleobases for in vitro selection [42, 43].

3.2.1 Nucleic acids as scaffolds for peptide display

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

Building upon this work, Hili and coworkers have investigated the use of building blocks having polypeptide side chains, as this would provide nucleic acid libraries that mimic the multivalency which characterizes many antibody-antigen interactions. In these experiments, pentanucleotide building blocks were chosen for the T4 DNA ligase-mediated polymerization, with the hypothesis that this length would be amenable to polymerization with longer peptide side chains.
[45]. Using a self-priming DNA template, pentanucleotide building blocks functionalized with either a cationic or anionic octapeptide were efficiently polymerized, suggesting that this scaffold system could display a wide variety of peptide sequences. A mock in vitro selection experiment was performed using seven pentanucleotide building blocks that had been shown to undergo efficient ligation and an eighth building block having a hexahistidine tag. The resulting library was selected for binding to Co$^{2+}$ magnetic beads, and an XbaI cleavage site was included in the template encoding the hexahistidine monomer in order to quantify enrichment (Figure 6). After completing one selection cycle, the hexahistidine-containing sequence was enriched nearly 200-fold, suggesting the potential of the T4 DNA ligase-based polymerization to function for the in vitro selection of aptamers or DNazymes.

Figure 6. Selection of DNA sequences having displayed peptides. Peptide-bearing fragments undergo templated ligation, and selected sequences can be PCR amplified and purified to regenerate single-stranded self-priming library.

3.2.2 Click-SELEX as a versatile method for incorporating functional groups

The Mayer Group has recently presented a new approach termed “click-SELEX” that enables the construction of nucleobase-modified aptamers utilizing the popular copper (I) azide-alkyne cycloaddition reaction [46]. In this method, all thymidine nucleotides within the library are replaced with C5-ethynyl-2′deoxyuridine (EdU), which introduces multiple alkyne functional groups into the sequences. The library is then further functionalized through reaction with an azide, which in the case of this initial demonstration was 3-(2-azidoethyl)indole. The resulting library was then subjected to in vitro selection to enrich aptamers having affinity for the protein cycle 3 GFP (C3-GFP). Sequences surviving each round of selection were amplified using
PCR, with EdU triphosphate used in place of dTTP. This provided the enriched library having alkyne functional groups, which could again be reacted with the azide-containing molecule and utilized in the subsequent selection round (Figure 7). Notably, the bulk of the nucleobase modification is introduced after DNA amplification, which allows access to larger functional groups while avoiding the enzymatic incompatibility that often accompanies such modifications. After 15 rounds of selection, the most abundant aptamer sequence was analyzed and found to have a dissociation constant of 18.4 nM. This aptamer also proved to have excellent specificity, displaying no observable binding to off-target proteins including streptavidin, ERK2, and the closely related mE-GFP. A series of azides were utilized to assess the importance of the specific indole moiety used for the selection process. Aptamer variants functionalized with the alternative azides were not capable of binding to the target protein even when the azide was structurally similar to the indole employed in the selections. This highlights the important role of the appended functional groups in target binding. The versatility with which side chains having a wide variety of sizes and chemical properties can be easily introduced using this method is anticipated to greatly accelerate the discovery of nucleic acid catalysts and affinity reagents having novel functions.

**Figure 7.** Selection of nucleobase-modified aptamers via “click-SELEX.” Post-synthetic modification of the DNA library enables the incorporation of large functional groups. These modifications are reverted to alkynes for PCR amplification. This enables the use of modifications that are not necessarily compatible with existing polymerases.

4. Conclusions and future outlook
As highlighted throughout this review, recent advances in molecular design and polymerase engineering have paved the way for the *in vitro* selection of aptamers having unnatural backbones and nucleotides. Excitingly, this expansion of chemical functionality has already
delivered on its promise to provide affinity reagents having increased biological stability or binding affinity, which will significantly advance the use of aptamers in both therapeutics and diagnostics. Historically, the limited availability of modified triphosphate and phosphoramidite building blocks has hindered the selection and characterization of modified aptamers. However, significant effort has been invested into developing improved synthetic methods, and many modified building blocks are now commercially available. Looking to the future, a number of technologies are in development which will enable the exploration of even greater swaths of chemical structure space. One of these areas is the selection of aptamers having both a modified backbone and modified nucleobases, an initial example of which is provided by the TNA aptamer containing diaminopurine nucleotides in place of adenine. Further expansion in this area to encompass a wide diversity of nucleotide modifications may be possible using the polymerases that are currently available, or is certainly feasible through further polymerase engineering. Additionally, the ability to generate genetically encoded sequence-defined polymers via non-enzymatic templated synthesis has potential to dramatically advance in vitro selection technology. While the implementation of this methodology to generate affinity reagents has yet to be demonstrated, success in this arena will likely be transformative, as it will enable the evolution of completely abiotic polymers that could display entirely new structures and functions.

**Acknowledgements**
This work was supported by the DARPA Folded Non-Natural Polymers with Biological Function (Fold F(x)) Program under award number N66001-14-2-4054. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of DARPA.

**References**


Graphical Abstract

enzyme engineering templated synthesis

in vitro selection

improved affinity increased stability