ACS Combinatorial Science

Research Article

pubs.acs.org/acscombsci

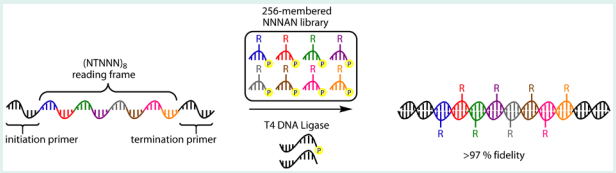
# A High-Fidelity Codon Set for the T4 DNA Ligase-Catalyzed Polymerization of Modified Oligonucleotides

Yi Lei, Dehui Kong, and Ryan Hili*

Department of Chemistry, University of Georgia, 140 Cedar Street, Athens, Georgia 30602, United States

Supporting Information

**ABSTRACT:** In vitro selection of nucleic acid polymers can readily deliver highly specific receptors and catalysts for a variety of applications; however, it is suspected that the functional group deficit of nucleic acids has limited their potential with respect to proteinogenic polymers. This has stimulated research toward expanding their chemical diversity to bridge the functional gap between nucleic acids and proteins to develop a superior biopolymer. In this study, we investigate the effect of codon library size and composition on the sequence specificity of T4 DNA ligase in the DNA-templated polymerization of both unmodified and modified oligonucleotides. Using high-throughput DNA sequencing of duplex pairs, we have uncovered a 256-membered codon set that yields sequence-defined modified ssDNA polymers in high yield and with high fidelity.




**KEYWORDS:** *aptamers, T4 DNA ligase, DNA-templated, SELEX, DNA sequencing*

In the rapidly growing fields of proteomics and glycomics, in-depth analysis of protein expression, modification, and interaction on a genomic scale requires the development of specific and high-affinity reagents.[1−3] Traditionally, proteinogenic polymers, especially antibodies, have served in this role; however, they suffer from a number of shortcomings, ranging from poor stability and variable production quality to limited target availability and issues with specificity profiles.[4−6] Despite over 500 000 commercially available antibodies,[6] researchers continue to express concerns over accessibility to high-quality antibodies for biomedical research.[7−9] These issues have stimulated the development of alternative technologies to generate high-affinity reagents for proteins that rival the performance of traditional antibodies.[10−14]

To this end, the use of nucleic acid aptamers[15] as high-affinity reagents has come to the fore, as they provide numerous advantages over traditional antibodies, such as (i) their generation through in vitro selection enables ready tuning of binding and specificity properties; (ii) their active structure can be reversibly formed by thermal denaturation and cooling; (iii) they exhibit excellent chemical stability and shelf life; and (iv) their chemical synthesis is predictable and scalable.[15] Unfortunately, the functional group deficit of nucleic acids limits their potential to match the performance of proteinogenic affinity reagents. It is anticipated that expanding the functional group repertoire of DNA will increase the heteromultivalent interactions with their molecular targets and better resemble the binding events that take place at protein−protein interfaces.[16] Thus, methods to increase the chemical functionality present in nucleic acid polymers are of significant interest.[17]

The sequence-specific incorporation of chemical functionality throughout a nucleic acid polymer has traditionally relied on polymerase-catalyzed DNA-templated primer extension using nucleobase-modified dNTPs.[18] When using a four-base genetic code, this approach enables up to four different functional groups to be incorporated throughout ssDNA and also provides flexibility over the sugar backbone structure.[18,19] While this has produced nucleic acid polymers with superior function when compared with their unmodified counterparts,[16,20] the ability to incorporate greater functionality remained elusive. Recently, the T4 DNA ligase-catalyzed DNA-templated polymerization of modified 5′-phosphorylated oligonucleotides has expanded the number of different modifications on a ssDNA to eight and relaxed limitations on the size of the modification (Figure 1).[21] As the method relies on codons, rather than single nucleotides, the theoretical number of unique modifications that can be incorporated increases with increasing codon length. Thus, a trinucleotide

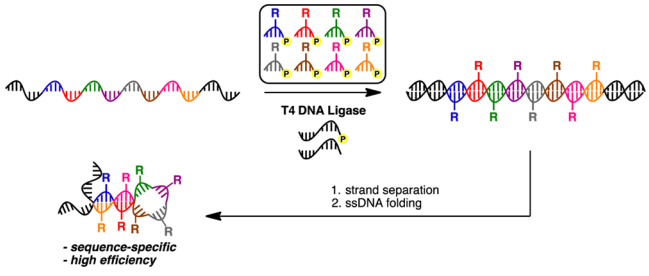



**Figure 1.** Synthesis of modified ssDNA polymers using T4 DNA ligase-catalyzed polymerization of modified DNA oligonucleotides.

**Received:** July 20, 2015
**Revised:** September 30, 2015
**Published:** October 29, 2015





DOI: 10.1021/acscombsci.5b00119
*ACS Comb. Sci.* 2015, 17, 716−721

codon set enables the incorporation of up to 64 unique modifications.

While the specificity of the reported ligase-catalyzed polymerization using an eight-membered trinucleotide codon library was high, there is little known about the specificity of larger libraries that can access greater sequence space. One early report documents the filling of a 12 nt gapped DNA duplex with either a library of 64 trinucleotides or library of 4096 hexanucleotides using *Escherichia coli* ligase.[22] The yield of the full-length product was very low; however, the fidelity of the system at the single nucleotide level was high, with hexanucleotides providing the best fidelity of 0.8 single-nucleotide errors per 12 nt sequence. Unfortunately, when the error rate is considered at the level of hexanucleotide incorporation, this corresponds to 60% fidelity. For longer sequences, it is not clear if such a low level of fidelity could support iterative rounds of in vitro selection, and the ability for *E. coli* ligase to accept chemical modifications was not studied.

We sought to identify a codon set for the T4 DNA ligase-catalyzed DNA-templated polymerization of modified 5′-phosphorylated oligonucleotides that satisfies the following requirements: (i) highly efficient polymerization; (ii) high fidelity DNA-templated polymerization; (iii) broad coverage of sequence space; (iv) tolerant of small modifications on polymerized oligonucleotides; and (v) readily accessible with standard phosphoramidite mixtures. Herein, we report a codon set for the ligase-catalyzed polymerization of modified oligonucleotides that satisfies all of these requirements and should enable the in vitro selection of modified nucleic acids as receptors and catalysts with larger repertoires of chemical functionality.

Building from the initial report on the T4 DNA ligase-catalyzed polymerization of modified trinucleotides using an eight-membered codon set,[21] we pursued a tetranucleotide codon set with the aim of expanding the number of possible modifications and increasing the coverage of sequence space. Unfortunately, the polymerization efficiency of modified tetranucelotides dropped dramatically when expanding the codon set to 64 members (Figure S1). As our objective was to identify a codon set capable of exploring a greater portion of nucleic acid sequence space, we next examined the efficiency of T4 DNA ligase to polymerize a library of 5′-phosphorylated pentanucleotides along a library of DNA templates.[23] The pentanucleotide library consisted of all 256 possible sequences derived from 5′P-ANNNN, where A was either unmodified or modified as an N8-hexylamine derivative (Figure 2c). The corresponding template library comprised two primer sites that flanked eight repeats of the codon NNNNT. The primers used during the polymerization were fluorescently labeled with CY5 and 6FAM fluorophores, which enabled identification of full-length products by 2-channel fluorescent PAGE imaging (Figure 2). In contrast with the trinucleotides and tetranucleotide systems, efficient polymerization of both the modified and unmodified pentanucleotide libraries was observed as judged by the presence of a heavy dual fluorescently labeled product band.

Encouraged by the efficiency of polymerization, we next examined the fidelity of the process. We first sought to determine if T4 DNA ligase was able to efficiently incorporate a pentanucleotide containing a single-nucleotide mismatch and extend from the misincorporation site (i.e., read through). Using a set of 5′-phosphorylated hairpin templates, we assessed the ability of T4 DNA ligase to incorporate a pentanucleotide (5′P-AGAGA) across from a codon with a single-nucleotide

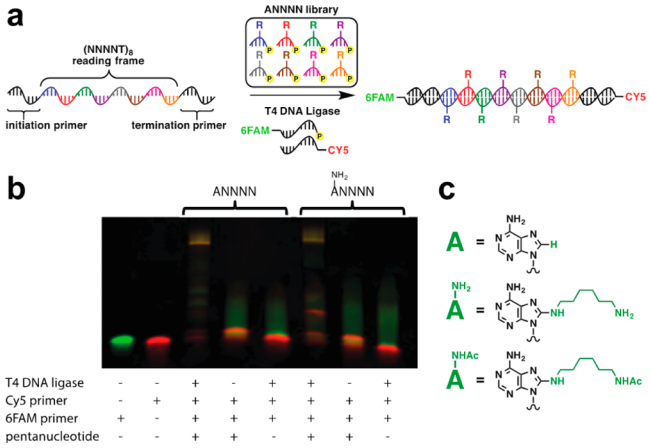



**Figure 2.** (a) Dual fluorphore approach for evaluating polymerization efficiency. (b) Dual-channel fluorescent image of a denaturing PAGE analyzing polymerization efficiency. (c) Unmodified and modified adenosines used in this study.

error (Figure 3). The templates contained eight codons, where the first codon was either the reverse complement or a codon

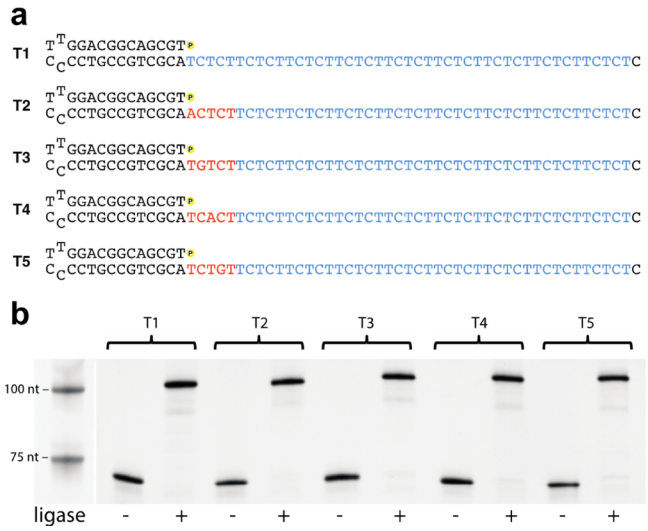



**Figure 3.** (a) Templates used to examine the incorporation of pentanucleotides containing single-nucleotide errors. (b) Denaturing PAGE analysis of T4 DNA ligase-catalyzed polymerization showing efficient incorporation of pentanucleotides with single-nucleotide errors.

containing a single-nucleotide error; a 3′-cytdine was added to the template to preclude blunt-end ligation of the full-length products. Following analysis by denaturing PAGE, full-length products were observed for all cases under the examined polymerization conditions. These experiments demonstrate that misincorporation can readily occur when the matched pentanucleotide sequence is not present in solution and that such misincorporations can be extended to full-length products (Figure 3b). While potentially serving as an issue with respect to fidelity, this feature could serve an important role during in vitro evolution of modified nucleic acids, as mutations will not terminate the nucleic acid polymer synthesis, thus enabling novel fit phenotypes to survive and be replicated to generate novel genotypes for subsequent rounds of selection.



DOI: 10.1021/acscombsci.5b00119
*ACS Comb. Sci.* 2015, 17, 716−721

We next designed a simple chain-termination[21,23,24] competition experiment to provide an estimation of single-nucleotide discrimination during pentanucleotide incorporation, as we anticipated single-nucleotide mismatches to be the greatest threat to fidelity. Thus, a 5′-phosphorylated hairpin template was synthesized containing a TCTCT codon followed by seven repeats of AGAGT and capped with a 3′-cytidine (Figure 4). The template was used to direct the polymerization

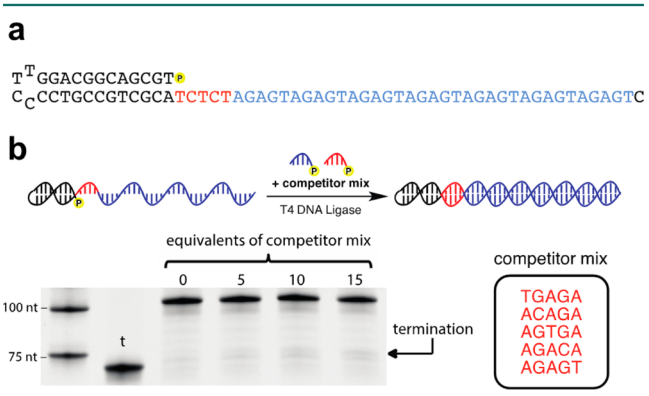



**Figure 4.** (a) 5′-phosphorylated DNA hairpin template used to examine the incorporation of errant pentanucleotides in competition with correct pentanucleotides. (b) Denaturing PAGE analysis of single-nucleotide mismatch competition experiments.

of 5′P-AGAGA and 5′P-ACTCT in the presence of increasing amounts of a mixture of single-nucleotide mismatch terminators of 5′P-AGAGA (mixture contained equal parts non-phosphorylated pentanucleotides 5′-TGAGA, 5′-ACAGA, 5′-AGTGA, 5′-AGACA, and 5′-AGAGT). Since T4 DNA ligase requires 5′-phosphates to continue the templated polymerization, misincorporation of a nonphosphorylated single-nucleotide mismatch pentanucleotide will terminate polymerization. As the terminator mixture was increased from 0 to 15-fold with respect to 5′P-AGAGA, chain termination at codon position one of the template increased from 0 to 5%, suggesting that high single-nucleotide discrimination occurred during the templated polymerization.

Following a series of preliminary sequence specificity experiments involving the analysis of fidelity by restriction enzyme digestion of polymerized products, and analysis of polymerized products by Sanger sequencing (see Figures S1 and S2), we observed sufficiently high fidelity of polymerization to warrant a high-throughput evaluation of the system. The measure of fidelity in XNA systems,[19] which typically comprise a four-nucleotide code, relies on the sequencing analysis of the product of polymerization from a single template. Because of the large number of codons used during ligase-catalyzed pentanucleotide polymerization, a high throughput approach was needed to provide adequate sampling of each codon to determine the fidelity of the process. Since the large size of the pentanucleotide codon set prohibits the ready analysis of sequence specificity by a conventional chain-termination approach, we looked to high-throughput DNA sequencing. Inspired by the "Duplex Sequencing" approach,[25] which was recently developed to increase accuracy in high-throughput DNA sequencing, we reasoned that the barcoding of duplex pairs would enable postsequencing association of template and polymerized strands (Figure 5). Thus, the sequences of the template and polymerized strands could be directly compared

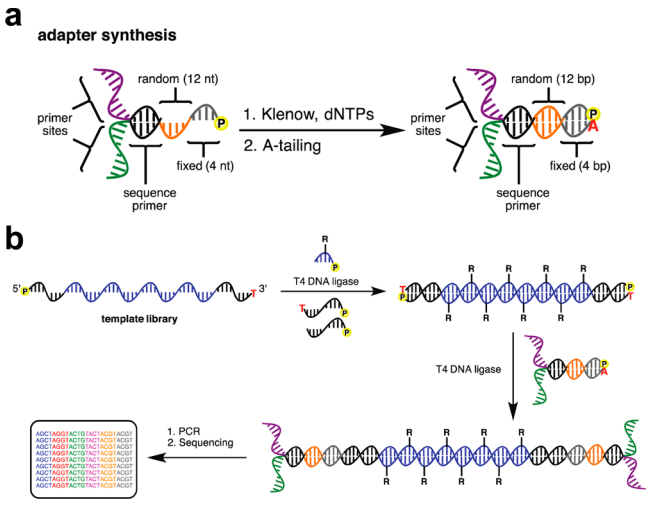



**Figure 5.** (a) Synthesis of DNA duplex adapter. (b) Workflow of duplex DNA sequencing to analyze the fidelity of T4 DNA ligase-catalyzed polymerization of oligonucleotides.

to reveal the fidelity of the ligase system for various codon sets in a high-throughput manner.

DNA template libraries were designed with two primer binding sites flanking a reading frame comprising eight consecutive repeats of various codon sets (Table 1). The

**Table 1. fidelity of unmodified pentanucleotide polymerization**

| entry | pentanucleotide[a] | codon set | library size | reads[b] | fidelity (%)[c] |
|---|---|---|---|---|---|
| 1 | 5′P-NNNNN | NNNNN | 1024 | 3376 | 81.1 |
| 2 | 5′P-ANWNW | WNWNT | 64 | 1800 | 91.6 |
| 3 | 5′P-ANNNN | NNNNT | 256 | 1536 | 86.7 |
| 4 | 5′P-NANNN | NNNTN | 256 | 3968 | 87.5 |
| 5 | 5′P-NNANN | NNTNN | 256 | 288 | 88.5 |
| 6 | 5′P-NNNAN | NTNNN | 256 | 176 | 85.2 |
| 7 | 5′P-NNNAN | NTNNN | 256 | 2.1 × 10[6][d] | 83.7 |
| 8 | 5′P-NNNNA | TNNNN | 256 | 288 | 84.7 |

[a]Unmodified pentanucleotides. [b]Number of codon reads from sequencing data. [c]Fidelity was calculated for pentanucleotide incorporation; the fidelity is considerably higher if evaluated at the single nucleotide level. [d]35 amols instead of 35 fmols were used for PCR prior to sequencing.

DNA templates were modified with a 5′-phosphate and a 3′-T to facilitate downstream ligations of sequencing adapters. Two primers were then annealed to the template library; the initiation primer was modified with a 5′-phosphate and a 3′-T overhang, and the termination primer was modified with a 5′-phosphate. Following T4 DNA ligase-catalyzed polymerization of the corresponding pentanucleotide library, the duplex barcoding tag, which contained standard Illumina sequencing adapters and a 12 bp randomized region, was ligated directly to both ends of the duplex product and subsequently purified by PAGE. The adapter ligated products were PCR amplified with flow-cell adapter sequences and subjected to paired-end DNA sequencing.

After DNA sequencing and processing, duplex pairs were grouped by their matching 24-nucleotide barcode tags and the template and polymerized strands were identified. The reading frames of the polymerized strands were then parsed into


DOI: 10.1021/acscombsci.5b00119
ACS Comb. Sci. 2015, 17, 716−721

pentanucleotide sequences and compared against their template codons to identify errors. The frequency of each 5′-phosphorylated pentanucleotide sequence was determined, and their misincorporation rate was calculated. Readouts of errant codon sites for each pentanucleotide were generated to determine trends in fidelity. Overall fidelity of the codon set was calculated by the aggregated error frequency of all codons. Since PCR amplification (20−30 cycles) was performed after duplex tag ligation, the fidelity is a combined assessment of ligase-catalyzed polymerization and PCR amplification of the modified dsDNA, which is required for iterative rounds of in vitro selection.

Sequencing results for the polymerization of unmodified pentanucleotide libraries are summarized in Table 1. Surprisingly, polymerization with a 1024-membered NNNNN pentanucleotide library resulted in sequence-specific incorporation with 81.1% fidelity (entry 1, Table 1). Considering the complexity of such a large codon library, and the presence of single-nucleotide mismatch competitors, this level of fidelity is remarkable. As expected, with decreasing library complexity, the fidelity of the polymerization increases; when using a 64-membered codon set, 91.6% fidelity was observed (Table 1, entry 2). We also surveyed the fidelity of a 256-membered codon set, where adenosine was held constant at one position along the pentanucleotide (Table 1, entries 3−8). The fidelities for these codon sets were also high, ranging from 83.7 to 88.5%.

We next evaluated the sequence-specificity for ligase-catalyzed polymerization of pentanucleotides modified with chemical functionality. Using the 256-membered codon system, pentanucleotides containing a hexylamine group on the adenine nucleobase (Figure 1c) were polymerized along their corresponding templates and the fidelity of the process was assessed by duplex DNA sequencing (Table 2, entries 1−5, 7).

**Table 2. Fidelity of Modified Pentanucleotide Polymerization**

| entry | pentanucleotide[a] | codon set | reads[b] | yield (%)[c] | fidelity (%)[d] |
|---|---|---|---|---|---|
| 1 | 5′P-ANNNN | NNNNT | 31936 | 75 | 95.1 |
| 2 | 5′P-NANNN | NNNTN | 2912 | 60 | 97.8 |
| 3 | 5′P-NNANN | NNTNN | 5232 | 50 | 98.0 |
| 4 | 5′P-NNNAN | NTNNN | 3032 | 30 | 98.4 |
| 5 | 5′P-NNNAN | NTNNN | $2.4 \times 10^5$[e] | 30 | 98.1 |
| 6 | 5′P-NNNAN[f] | NTNNN | $2.1 \times 10^5$[e] | 30 | 97.6 |
| 7 | 5′P-NNNNA | TNNNN | 600 | 10 | 98.0 |

[a]Modification is a hexylamine unless otherwise noted. [b]Number of codon reads from sequencing data. [c]Yield of polymerization was calculated by gel electrophoresis (see Figure S4). [d]Fidelity was calculated for pentanucleotide incorporation; the fidelity is considerably higher if evaluated at the single nucleotide level. [e]35 amols instead of 35 fmols were used for PCR prior to sequencing. [f]Modification is a *N*-hexylacetamide.

Remarkably, the addition of a modification on the pentanucleotide decreased error rates >5-fold, resulting in fidelities of >95%. Error rates were the highest when the modification was at the 5′-end, but this level of fidelity should not preclude in vitro selection, as 67% of octacodon templates were generated faithfully without error. Notwithstanding, we focused on the NTNNN codon set, which provided 98.4% fidelity. This level of fidelity was confirmed with a second higher throughout sequencing analysis, which used a lower amount of DNA template (35 amols) for PCR prior to DNA sequencing (Table 2, entry 5). We were concerned that the amine functionality, which would be protonated under the ligation conditions (pH 7.6), might have enabled the observed higher sequence-specificity. This would pose problems when pentanucleotides containing various uncharged functional groups are used. To investigate this possible undesired effect, we synthesized a pentanucleotide library that was modified with an uncharged *N*-hexylacetamide group. Gratifyingly, the level of fidelity remained high (97.6%) suggesting that various small functional groups on the adenosine nucleobase could be accommodated by this polymerization method and 256-membered NTNNN codon set.

Although the reasons for the striking increase in fidelity that arise upon nucleobase modification are not entirely understood, one possibility is that they result from the perturbation in melting temperature upon nucleobase modification. It is known that modifications at C-8 of adenine results in the adenosine adopting a *syn* conformation about the N-glycoside bond in solution, which inhibits the kinetics of annealing and decreases the thermal stability of the DNA duplex.[26,27] The resulting drop in thermal duplex stability and slower annealing kinetics should give rise to a more discriminating ligation process. However, sequencing data indicate that GC-content, and thus the $T_m$ of the pentanucleotide has little influence on fidelity within a codon set. For example, when polymerizing the NNNAN pentanucleotide library modified with *N*-hexylacetamide, codons grouped by %GC-content showed little variability with respect to error rate (80%GC = 2.4%; 60%GC = 2.4%; 40%GC = 2.4%; 20%GC = 2.6%; 0%GC = 2.7%). Notwithstanding, the decrease in thermal stability caused by the C-8 modification could have a similar destabilizing effect on each member within the codon set.

We next analyzed the codon frequency of polymerized products to determine the presence of codon sequence bias during pentanucleotide incorporation. We limited our analysis to sequencing data from the NTNNN codon set, which resulted in greater than 100 instances of each codon within the library. For unmodified 5′P-NNNAN pentanucleotides, we observed a 69% increase in codon frequency with decreasing the GC-content (see Figure S5). This trend was repeated, albeit with lesser effect, for amino modified 5′P-NNNAN polymerization, which resulted in a 12% increase in codon frequency. Surprisingly, the *N*-hexylacetamide variant had a reversed trend with a 41% increase in codon frequency upon increasing the GC-content. The frequency range for the NTNNN codon set was also determined for unmodified and modified pentanucleotide polymerizations. While a 6-fold range in codon frequencies was observed for the unmodified pentanucleotides, this frequency range increased significantly with pentanucleotide modification to approximately 20-fold, with 94% of codons being within a 10-fold range. These data should be considered when synthesizing differentially modified pentanucleotide libraries for in vitro selections, as this bias will be reflected in the sequencing results. It is also anticipated that family B polymerases[28] will be required to amplify DNA with larger modifications, which could also affect not only the aggregated fidelity of the process, but also codon bias during selections.

The observed high level of fidelity should allow for in vitro selection by DNA display methods[29] and in vitro evolution since the error rate is low enough to avoid error catastrophe during iterated cycles of selection for lengths typically implemented in SELEX experiments.[30] To highlight the application of this polymerization to in vitro selection, we


DOI: 10.1021/acscombsci.5b00119
ACS Comb. Sci. 2015, 17, 716−721

performed oligonucleotide polymerization, selection, and amplification on a model DNA library. We prepared a DNA library, LIB1, which contained two primer-binding sites flanking eight consecutive repeats of the NTNNN codon set. We also prepared a biotinylated DNA template, POS1, which contained a *BcoDI* restriction site (Figure 6a). We diluted the POS1

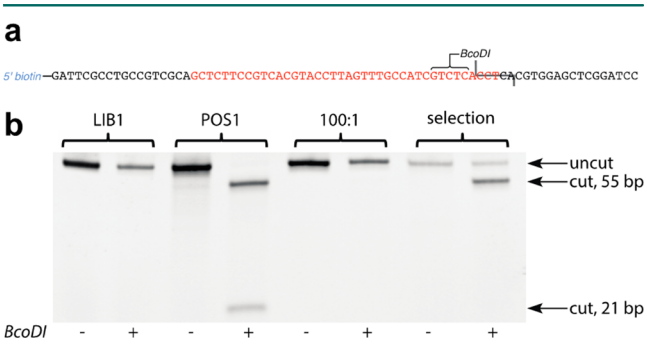



**Figure 6.** (a) POS1 sequence used during in vitro selection. (b) Enrichments after one round of in vitro selection for streptavidin binding.

template 100-fold into LIB1 and performed a single round of in vitro selection. The round involved polymerization with 5′P-NNNAN library that was modified with a *N*-hexylacetamide group; a selection pressure, which involved binding to streptavidin-coated magnetic particles, followed by thorough washing; elution of the modified DNA strand upon incubation in 100 mM NaOH; and then amplification of the eluted strand by PCR. To evaluate enrichment over the single round, the PCR product was digested with *BcoDI* and the ratio was compared against digestion of the amplified products of LIB1 and POS1 (Figure 6b). This single round of mock selection resulted in approximately 350-fold enrichment, suggesting that the codon set could work effectively in an in vitro selection against molecular targets.

In summary, we have described a strategy to determine the fidelity of the T4 DNA ligase-catalyzed DNA-templated polymerization of modified oligonucleotides using high-throughput DNA duplex sequencing. The method allowed the evaluation of fidelity for each pentanucleotide within a large codon set. We applied this strategy to discover a 256-member codon set that enables the incorporation of small molecule functionality throughout a DNA polymer at >97% fidelity. This codon set was implemented in a mock selection for streptavidin binding, exhibiting 350-fold enrichment over one round. This high-fidelity polymerization strategy should find immediate application in DNA nanotechnology, DNA computing, and in vitro selection of functional nucleic acids.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscombsci.5b00119.

Molecular characterization, processed sequencing data, methods, and supplementary data (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

*E-mail: rhili@uga.edu.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by the NSF (DMR 1506667) and the Office for the Vice President of Research, University of Georgia. We would like to thank Dr. Saravanaraj Ayyampalayam for help with the analysis of DNA sequencing data and the PAMS core facility at the University of Georgia for their help in the characterization of oligonucleotides.

## ■ REFERENCES

(1) Bradbury, A. R.; Velappan, N.; Verzillo, V.; Ovecka, M.; Marzari, R.; Sblattero, D.; Chasteen, L.; Siegel, R.; Pavlik, P. Antibodies in proteomics. *Methods Mol. Biol.* **2004**, *248*, 519−546.

(2) Blow, N. Antibodies: The generation game. *Nature* **2007**, *447*, 741−744.

(3) Cummings, R. D.; Pierce, J. M. The challenge and promise of glycomics. *Chem. Biol.* **2014**, *21*, 1−15.

(4) Berglund, L.; Bjoerling, E.; Oksvold, P.; Fagerberg, L.; Asplund, A.; Szigyarto, C. A. K.; Persson, A.; Ottosson, J.; Wernerus, H.; Nilsson, P.; Lundberg, E.; Sivertsson, A.; Navani, S.; Wester, K.; Kampf, C.; Hober, S.; Ponten, F.; Uhlen, M. A genecentric human protein atlas for expression profiles based on antibodies. *Mol. Cell. Proteomics* **2008**, *7*, 2019−2027.

(5) Nollau, P. Generating high-quality protein binders: a large screening effort pays off. *Nat. Methods* **2011**, *8*, 545−546.

(6) Colwill, K.; et al. A roadmap to generate renewable protein binders to the human proteome. *Nat. Methods* **2011**, *8*, 551−558.

(7) Taussig, M. J.; Stoevesandt, O.; Borrebaeck, C. A. K.; Bradbury, A. R.; Cahill, D.; Cambillau, C.; de Daruvar, A.; Dubel, S.; Eichler, J.; Frank, R.; Gibson, T. J.; Gloriam, D.; Gold, L.; Herberg, F. W.; Hermjakob, H.; Hoheisel, J. D.; Joos, T. O.; Kallioniemi, O.; Koegl, M.; Konthur, Z.; Korn, B.; Kremmer, E.; Krobitsch, S.; Landegren, U.; van der Maarel, S.; McCafferty, J.; Muyldermans, S.; Nygren, P. A.; Palcy, S.; Pluckthun, A.; Polic, B.; Przybylski, M.; Saviranta, P.; Sawyer, A.; Sherman, D. J.; Skerra, A.; Templin, M.; Ueffing, M.; Uhlen, M. ProteomeBinders: planning a European resource of affinity reagents for analysis of the human proteome. *Nat. Methods* **2007**, *4*, 13−17.

(8) Bordeaux, J.; Welsh, A. W.; Agarwal, S.; Killiam, E.; Baquero, M. T.; Hanna, J. A.; Anagnostou, V. K.; Rimm, D. L. Antibody validation. *BioTechniques* **2010**, *48*, 197−209.

(9) Uhlen, M.; Bjorling, E.; Agaton, C.; Szigyarto, C. A.; Amini, B.; Andersen, E.; Andersson, A. C.; Angelidou, P.; Asplund, A.; Asplund, C.; Berglund, L.; Bergstrom, K.; Brumer, H.; Cerjan, D.; Ekstrom, M.; Elobeid, A.; Eriksson, C.; Fagerberg, L.; Falk, R.; Fall, J.; Forsberg, M.; Bjorklund, M. G.; Gumbel, K.; Halimi, A.; Hallin, I.; Hamsten, C.; Hansson, M.; Hedhammar, M.; Hercules, G.; Kampf, C.; Larsson, K.; Linskog, M.; Lodewyckx, W.; Lund, J.; Lundeberg, J.; Magnusson, K.; Malm, E.; Nilsson, P.; Odling, J.; Oksvold, P.; Olsson, I.; Oster, E.; Ottosson, J.; Paavilainen, L.; Persson, A.; Rimini, R.; Rockberg, J.; Runeson, M.; Sivertsson, A.; Skollermo, A.; Steen, J.; Stenvall, M.; Sterky, F.; Stromberg, S.; Sundberg, M.; Tegel, H.; Tourle, S.; Wahlund, E.; Walden, A.; Wan, J. H.; Wernerus, H.; Westberg, J.; Wester, K.; Wrethagen, U.; Xu, L. L.; Hober, S.; Ponten, F. A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol. Cell. Proteomics* **2005**, *4*, 1920−1932.

(10) Mahon, C. S.; Fulton, D. A. Mimicking nature with synthetic macromolecules capable of recognition. *Nat. Chem.* **2014**, *6*, 665−672.

(11) Gebauer, M.; Skerra, A. Engineered protein scaffolds as next-generation antibody therapeutics. *Curr. Opin. Chem. Biol.* **2009**, *13*, 245−255.

(12) Mayer, G. The chemical biology of aptamers. *Angew. Chem., Int. Ed.* **2009**, *48*, 2672−2689.

(13) Lollo, B.; Steele, F.; Gold, L. Beyond antibodies: New affinity reagents to unlock the proteome. *Proteomics* **2014**, *14*, 638−644.

(14) Marx, V. Calling the next generation of affinity reagents. *Nat. Methods* **2013**, *10*, 829−833.



DOI: 10.1021/acscombsci.5b00119
*ACS Comb. Sci.* 2015, 17, 716−721

(15) Wilson, D. S.; Szostak, J. W. In vitro selection of functional nucleic acids. *Annu. Rev. Biochem.* **1999**, *68*, 611−647.
(16) Rohloff, J. C.; Gelinas, A. D.; Jarvis, T. C.; Ochsner, U. A.; Schneider, D. J.; Gold, L.; Janjic, N. Nucleic acid ligands with protein-like side chains: modified aptamers and their use as diagnostic and therapeutic agents. *Mol. Ther.–Nucleic Acids* **2014**, *3*, e201.
(17) Taylor, A. I.; Arangundy-Franklin, S.; Holliger, P. Towards applications of synthetic genetic polymers in diagnosis and therapy. *Curr. Opin. Chem. Biol.* **2014**, *22*, 79−84.
(18) Chaput, J. C.; Yu, H. Y.; Zhang, S. The emerging world of synthetic genetics. *Chem. Biol.* **2012**, *19*, 1360−1371.
(19) Pinheiro, V. B.; Holliger, P. The XNA world: progress towards replication and evolution of synthetic genetic polymers. *Curr. Opin. Chem. Biol.* **2012**, *16*, 245−252.
(20) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. Toward the Combinatorial Selection of Chemically Modified DNAzyme RNase A Mimics Active Against all-RNA Substrates. *ACS Comb. Sci.* **2013**, *15*, 174−182.
(21) Hili, R.; Niu, J.; Liu, D. R. DNA Ligase-Mediated Translation of DNA Into Densely Functionalized Nucleic Acid Polymers. *J. Am. Chem. Soc.* **2013**, *135*, 98−101.
(22) James, K. D.; Boles, A. R.; Henckel, D.; Ellington, A. D. The fidelity of template-directed oligonucleotide ligation and its relevance to DNA computation. *Nucleic Acids Res.* **1998**, *26*, 5203−5211.
(23) Guo, C.; Watkins, C. P.; Hili, R. Sequence-Defined Scaffolding of Peptides on Nucleic Acid Polymers. *J. Am. Chem. Soc.* **2015**, *137*, 11191−11196.
(24) Niu, J.; Hili, R.; Liu, D. R. Enzyme-free translation of DNA into sequence-defined synthetic polymers structurally unrelated to nucleic acids. *Nat. Chem.* **2013**, *5*, 282−292.
(25) Schmitt, M. W.; Kennedy, S. R.; Salk, J. J.; Fox, E. J.; Hiatt, J. B.; Loeb, L. A. Detection of ultra-rare mutations by next-generation sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 14508−14513.
(26) Dierckx, A.; Diner, P.; El-Sagheer, A. H.; Kumar, J. D.; Brown, T.; Grotli, M.; Wilhelmsson, L. M. Characterization of photophysical and base-mimicking properties of a novel fluorescent adenine analogue in DNA. *Nucleic Acids Res.* **2011**, *39*, 4513−4524.
(27) Luyten, I.; Herdewijn, P. Hybridization properties of base-modified oligonucleotides within the double and triple helix motif. *Eur. J. Med. Chem.* **1998**, *33*, 515−576.
(28) Jäger, S.; Rasched, G.; Kornreich-Leshman, H.; Engeser, M.; Thum, O.; Famulok, M. A Versatile Toolbox for Variable DNA Functionalization at High Density. *J. Am. Chem. Soc.* **2005**, *127*, 15071−15082.
(29) Li, G.; Zheng, W.; Liu, Y.; Li, X. Novel encoding methods for DNA-templated chemical libraries. *Curr. Opin. Chem. Biol.* **2015**, *26*, 25−33.
(30) Rajamani, S.; Ichida, J. K.; Antal, T.; Treco, D. A.; Leu, K.; Nowak, M. A.; Szostak, J. W.; Chen, I. A. Effect of Stalling after Mismatches on the Error Catastrophe in Nonenzymatic Nucleic Acid Replication. *J. Am. Chem. Soc.* **2010**, *132*, 5880−5885.



DOI: 10.1021/acscombsci.5b00119
*ACS Comb. Sci.* 2015, 17, 716−721