Selection of DNA aptamers with two modified bases

Bharat N. Gawande⁎, John C. Rohloff†, Jeffrey D. Carter⁎, Ira von Carlowitz⁎, Chi Zhang⁎, Daniel J. Schneider⁎, and Nebojsa Janičić⁎,†

⁎SomaLogic, Inc., Boulder, CO 80301

Edited by Jack W. Szostak, Massachusetts General Hospital, Boston, MA, and approved February 3, 2017 (received for review September 15, 2016)

The nucleobases comprising DNA and RNA aptamers provide considerably less chemical diversity than protein-based ligands, limiting their versatility. The introduction of novel functional groups at just one of the four bases in modified aptamers has recently led to dramatic improvement in the success rate of identifying nucleic acid ligands to protein targets. Here we explore the benefits of additional enhancement in physicochemical diversity by selecting modified DNA aptamers that contain amino-acid–like modifications on both pyrimidine bases. Using propionate convertase subtilisin/kexin type 9 as a representative protein target, we identify specific pairwise combinations of modifications that result in higher affinity, metabolic stability, and inhibitory potency compared with aptamers with single modifications. Such doubly modified aptamers are also more likely to be encoded in shorter sequences and occupy nonoverlapping epitopes more frequently than aptamers with single modifications. These highly modified DNA aptamers have broad utility in research, diagnostic, and therapeutic applications.

SELEX | modified aptamer | PCSK9 | SOMAmer | PSMA

SELEX is a powerful tool for identifying nucleic-acid-based ligands, or aptamers (1–3), to a wide range of molecular targets. However, the utility of aptamers in research, diagnostic, and therapeutic applications has been constrained by the limited chemical diversity of the nucleic acid libraries. Compared with protein-based ligands, like antibodies, natural nucleic acids are composed of fewer building blocks (four bases vs. 20 amino acids), have a polyanionic backbone with relatively hydrophilic character, and possess a smaller repertoire of functional groups available for target recognition. This limitation is partially compensated for by the sheer size of the random libraries available for screening (>10¹⁵ molecules).

Several attempts have been made to increase the functional diversity of nucleic acids either with modified nucleobases or sugar–phosphate backbone modifications aimed at generating ligands with improved binding properties and metabolic stability. One approach has been to expand the four-letter genetic alphabet to include an additional base pair with a unique hydrogen bonding pattern (4). This six-letter library has been shown to be compatible with the required synthetic, PCR amplification, and sequencing steps of SELEX, resulting in the successful identification of aptamers to cell surface targets (4). Whereas the information density of such libraries is clearly higher, and may result in a larger repertoire of folding motifs, the new bases are structurally and functionally similar to the natural bases, so the expansion of the diversity of such libraries is still limited. In contrast, Kimoto et al. (5) introduced novel base pairs based on hydrophobic interactions and steric complementarity within the double helix that allow the addition of one hydrophobic base in the sense strand. Although the introduction of an unnatural hydrophobic base improved aptamer affinity with select targets, this method is encumbered with the need to introduce the modified base at predetermined positions in the starting libraries, thus limiting sequence diversity (5).

More recently, the application of enzyme engineering or directed evolution approaches has resulted in the identification of novel polymers that are capable of generating synthetic genetic polymers, or xeno nucleic acids (6) with alternate backbone chemistries not found in natural nucleic acids. Such XNA libraries, which replace ribose with alternate backbone chemistries not found in natural nucleic acids, such as locked nucleic acids (LNAs) (8), L-enantiomer RNA (L-RNA) (9), as well as the 2′-amino, 2′-fluoro, and 2′-O-methyl modifications (10, 11) have been of considerable interest and utility because of their enhanced resistance to nuclease degradation. However, none of these alternatives provide fundamentally enhanced functional group diversity in the nucleic acid libraries. Nucleobase functional groups that resemble amino acid side chains or fragments of small molecules have been introduced at either the 5-position of pyrimidines or the 7- or 8-position of purines to enhance catalytic activity or ligand binding properties (10, 12–19).

We recently reported that introducing diversity-enhancing functional groups at one of the four nucleobases in DNA dramatically enhances the efficiency of SELEX (12, 20). Among several types of moieties tested, large, hydrophobic functional groups that resemble similar side chains in proteins have proven to be the most effective modifications for a wide range of protein targets (20). Such modifications, uniformly introduced at the 5-position of deoxyuridine (dU) (12), have allowed us to add a demand for kinetic stability of the complexes during selections and identify slow off-rate modified aptamer (SOMAmer) reagents for over 3,000 proteins. Based on several co-crystal structures of SOMAmer–protein complexes, these side chains participate in shaping the target epitope interaction surfaces and create novel intramolecular motifs that facilitate SOMAmer folding (21–24). This is consistent with the observation that amino acids of similar type are overrepresented in antibody paratopes that recognize protein epitopes (25, 26). For example, phenylalanine, tyrosine, and tryptophan are as much as fivefold more abundant in paratopes compared with the rest of the antibody surface (25). Remarkably, functional antibodies can be obtained from libraries containing only two amino acids in the complementarity determining regions (CDRs), as long as the pair is judiciously chosen, like tyrosine and serine (27). In these antibodies, tyrosine forms the predominance of protein contacts, whereas serine, which is both small and abundant on protein surfaces, mainly creates space for tyrosine side chains to adopt optimal orientation with respect to the protein surface.

Significance

Aptamers are now used ubiquitously as binding agents for a broad range of applications. Natural (unmodified) DNA and RNA aptamers have considerably less chemical diversity than protein-based ligands such as antibodies, limiting their utility. Aptamers possessing a single chemical modification have helped bridge this diversity gap. We report the selection and identification of aptamers with two diversity-enhancing chemical modifications that bind and inhibit proprotein convertase subtilisin/kexin type 9 (PCSK9), a representative human therapeutic protein target. The addition of a second modification, especially in certain pairwise combinations, resulted in significant improvements in affinity, ligand efficiency, epitope coverage, metabolic stability, and inhibitory activity. Extensively chemically functionalized aptamers have the potential to become the next generation of nucleic-acid–based ligands.

Author contributions: B.N.G., I.v.C., C.Z., D.J.S., and N.J. designed research; B.N.G., I.v.C., and C.Z. performed research; B.N.G., J.C.R., and J.D.C. contributed new reagents/analytic tools; B.N.G., D.J.S., and N.J. analyzed data; and B.N.G., D.J.S., and N.J. wrote the paper.

Conflict of interest statement: All authors are employees and/or shareholders of SomaLogic, Inc. SOMAmer reagent is a registered trademark of SomaLogic, Inc.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

⁎To whom correspondence should be addressed. Email: njanic@somalogic.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1615475114/-/DCSupplemental.
Double Mod

Selection of nucleic acid ligands from DNA libraries containing C5-position 2899

groups tested on dC were Nap and Pp (red lines), whereas R = | group and modifications on no. 11

PNAS | vol. 114 | 3) of single-modified and double-modified li-

Top 1.3%), Nap-dC/Moe-dU (40 ± 5.2%), and Pp-dC/ Nap-du (43 ± 2.7%) libraries (Fig. 1

To compare the relative efficiency of SELEX with a representative human protein,

potency of the resulting aptamers, and identify the best combination

of deoxyuridine have been among the most impactful single modifi-

cations, as well as more hydrophilic side chains on dU to identify the

optimal combinations for selecting high-affinity ligands. Although we

have observed in a number of previous SELEX experiments with

protein targets that hydrophilic side chains generally have modest,

if any, impact on the success rate of SELEX, their contribution in

the context of a second modification was difficult to anticipate.

Each of the 18 starting libraries were synthesized using

Thermococcus kodakarensis (KOD) DNA polymerase (exo-)

which accepts a wide variety of 5-position modified dC and dU

triphosphates as substrates (12, 28, 30). Compared with the

unmodified DNA control, we generally obtained lower yields for

libraries containing two modified nucleotides, notably for Nap-dC/ 

Nap-du (28 ± 1.3%), Nap-du/Moe-du (40 ± 5.2%), and Pp-dC/ 

Nap-du (43 ± 2.7%) libraries (Fig. 1C and SI Appendix, Fig. S1). 

However, this did not result in a notable bias in the base composition 

of the starting libraries, with only a modest underrepresentation of 

modified dU in some of the libraries (Fig. 1D, Top). With the notion 

that a higher density of modified bases could support high-affinity 

interactions across shorter SOMAmer lengths, we used 30N 

randomized libraries instead of the typical 40N libraries (20, 31).

Results

SOMAmer Selection. To compare the relative efficiency of SELEX 

with two modified bases, we tested two modifications on dC (Nap and 
Pp, with the unmodified dC as a control) in all possible pairwise 

combinations with five modifications on dU (Nap, Pp, Moe, Tyr and 

Thr, with the unmodified dT as a control), for a total of 18 starting 

libraries (Fig. 1B). The types of modifications tested included hydro-

phobic aromatic side chains on both dC and dU, which in the context 
of deoxuridine have been among the most impactful single modifi-
cations, as well as more hydrophilic side chains on dU to identify the 

optimal combinations for selecting high-affinity ligands. Although we

have observed in a number of previous SELEX experiments with

protein targets that hydrophilic side chains generally have modest,

if any, impact on the success rate of SELEX, their contribution in

the context of a second modification was difficult to anticipate.

Each of the 18 starting libraries were synthesized using

Thermococcus kodakarensis (KOD) DNA polymerase (exo-)

which accepts a wide variety of 5-position modified dC and dU

triphosphates as substrates (12, 28, 30). Compared with the

unmodified DNA control, we generally obtained lower yields for

libraries containing two modified nucleotides, notably for Nap-dC/ 

Nap-du (28 ± 1.3%), Nap-du/Moe-du (40 ± 5.2%), and Pp-dC/ 

Nap-du (43 ± 2.7%) libraries (Fig. 1C and SI Appendix, Fig. S1). 

However, this did not result in a notable bias in the base composition 

of the starting libraries, with only a modest underrepresentation of 

modified dU in some of the libraries (Fig. 1D, Top). With the notion 

that a higher density of modified bases could support high-affinity 

interactions across shorter SOMAmer lengths, we used 30N 

randomized libraries instead of the typical 40N libraries (20, 31).

Results

SOMAmer Selection. To compare the relative efficiency of SELEX 

with two modified bases, we tested two modifications on dC (Nap and 
Pp, with the unmodified dC as a control) in all possible pairwise 

combinations with five modifications on dU (Nap, Pp, Moe, Tyr and 

Thr, with the unmodified dT as a control), for a total of 18 starting 

libraries (Fig. 1B). The types of modifications tested included hydro-

phobic aromatic side chains on both dC and dU, which in the context 
of deoxuridine have been among the most impactful single modifi-
cations, as well as more hydrophilic side chains on dU to identify the 

optimal combinations for selecting high-affinity ligands. Although we

have observed in a number of previous SELEX experiments with

protein targets that hydrophilic side chains generally have modest,

if any, impact on the success rate of SELEX, their contribution in

the context of a second modification was difficult to anticipate.

Each of the 18 starting libraries were synthesized using

Thermococcus kodakarensis (KOD) DNA polymerase (exo-)

which accepts a wide variety of 5-position modified dC and dU

triphosphates as substrates (12, 28, 30). Compared with the

unmodified DNA control, we generally obtained lower yields for

libraries containing two modified nucleotides, notably for Nap-dC/ 

Nap-du (28 ± 1.3%), Nap-du/Moe-du (40 ± 5.2%), and Pp-dC/ 

Nap-du (43 ± 2.7%) libraries (Fig. 1C and SI Appendix, Fig. S1). 

However, this did not result in a notable bias in the base composition 

of the starting libraries, with only a modest underrepresentation of 

modified dU in some of the libraries (Fig. 1D, Top). With the notion 

that a higher density of modified bases could support high-affinity 

interactions across shorter SOMAmer lengths, we used 30N 

randomized libraries instead of the typical 40N libraries (20, 31).

Results

SOMAmer Selection. To compare the relative efficiency of SELEX 

with two modified bases, we tested two modifications on dC (Nap and 
Pp, with the unmodified dC as a control) in all possible pairwise 

combinations with five modifications on dU (Nap, Pp, Moe, Tyr and 

Thr, with the unmodified dT as a control), for a total of 18 starting 

libraries (Fig. 1B). The types of modifications tested included hydro-

phobic aromatic side chains on both dC and dU, which in the context 
of deoxuridine have been among the most impactful single modifi-
cations, as well as more hydrophilic side chains on dU to identify the 

optimal combinations for selecting high-affinity ligands. Although we

have observed in a number of previous SELEX experiments with

protein targets that hydrophilic side chains generally have modest,

if any, impact on the success rate of SELEX, their contribution in

the context of a second modification was difficult to anticipate.

Each of the 18 starting libraries were synthesized using

Thermococcus kodakarensis (KOD) DNA polymerase (exo-)

which accepts a wide variety of 5-position modified dC and dU

triphosphates as substrates (12, 28, 30). Compared with the

unmodified DNA control, we generally obtained lower yields for

libraries containing two modified nucleotides, notably for Nap-dC/ 

Nap-du (28 ± 1.3%), Nap-du/Moe-du (40 ± 5.2%), and Pp-dC/ 

Nap-du (43 ± 2.7%) libraries (Fig. 1C and SI Appendix, Fig. S1). 

However, this did not result in a notable bias in the base composition 

of the starting libraries, with only a modest underrepresentation of 

modified dU in some of the libraries (Fig. 1D, Top). With the notion 

that a higher density of modified bases could support high-affinity 

interactions across shorter SOMAmer lengths, we used 30N 

randomized libraries instead of the typical 40N libraries (20, 31).
We carried out selection experiments as previously reported (20, 31) (SI Appendix, Fig. S2 and Table S1F), initiating selections by mixing each library (≥1,000 pmol or ~10^8 unique sequences) with His-tagged human PCSK9 protein (50 pmol) immobilized on paramagnetic beads. After six rounds of selection, enriched pools were sequenced and analyzed to determine copy numbers of unique and conserved sequence patterns. The base composition of the affinity-enriched libraries indicated a slight-to-moderate shift from the frequencies observed in the starting libraries, generally toward an increase in the fraction of modified bases in libraries with a single modification, and an increase in modified dC in libraries with double modifications (Fig. 1D, Bottom). Several unique sequences from each of the 18 libraries were chosen for further analysis based on their frequency in the affinity-enriched pool and representation in sequence families (SI Appendix, Table S2). All were synthesized as 40-mers composed of the 30-nucleotide random region and 5-nucleotide primer sequence on each end. Alignments of representative sequences from the Pp-dC/Nap-dU, Nap-dC/Tyr-dU, and Nap-dC/Tyr-d pools are shown in SI Appendix, Tables S3–S5.

**SOMAmers with Two Modifications Show Higher Affinity Encoded in Shorter Sequences.** Sequences with high affinity (K_d ≤ 1 nM) were identified in all libraries that contained at least one hydrophobic modification (with the exception of Pp-dC/dT), but not in the control library (dC/dT) or in libraries containing only hydrophilic modifications (dC/Mp-dC, dC/Mp-Tyr-dU; Fig. 2A). The highest-affinity ligands were obtained from libraries containing two modifications, in particular a hydrophobic modification on dC paired with Tyr-dU (Nap-dC/Tyr-dU and Pp-dC/Tyr-dU). This is noteworthy in view of the fact that Tyr alone on dU did not produce ligands with affinities better than 20 nM. Even more remarkably, Pp alone on dC did not produce any high-affinity ligands. Similar results were observed with another target protein, prostate-specific membrane antigen (PSMA) (SI Appendix, Fig. S3). In addition, the high-affinity ligands with two modified nucleotides were highly abundant in their respective affinity-enriched pools, making them easy to identify without deep sequencing (SI Appendix, Fig. S4).

To determine the importance of similar hydrophobic aromatic side chains in antibody paratopes (25, 26), we further truncated all 40-mer SOMAmers with K_d < 1 nM (gray shaded region in Fig. 2A) to 30-mers (containing only the originally random region) and measured their binding affinities to PCSK9 (Fig. 2B). Among ligands with a single modified nucleotide, only 3 out of 14 (21%) were active as 30-mers (SI Appendix, Fig. S5), yellow bar), two (from the dC/dC/Tyr-dU library) and one from the dC/Nap-dU library. In comparison, 38 out of 93 (41%) ligands from libraries containing two modified nucleotides could be truncated to 30-mers, with the highest number arising from libraries containing Pp-dC (27 out of 45, or 60%; SI Appendix, Fig. S5, pink bar). It is possible that the extended methylene linker in the Pp side chain (Fig. 1B) when combined with the second modification is able to access deeper binding pockets than the other modifications we tested, thus requiring fewer nucleotides for high-affinity binding. However, this feature is apparently not sufficient for libraries containing only a single Pp side chain (Fig. 2B).

**Specificity with Species Cross-Reactivity Is More Readily Achieved.** To assess the effect of the second modified nucleotide on specificity, we tested the affinity of the three highest-affinity SOMAmers from each library (11 with single and 22 with double modification, all 40-mers) to other human proprotein convertases (PCs) (32). PCSK9 belongs to a family of nine secretory serine proteases that are related to bacterial subtilisin and yeast kexin (32). No binding activity was observed in human PCSK1, PCSK2, furin, PCSK4, or PCSK7 (13.5–16.2% identity to PCSK9) at concentrations up to 100 nM, demonstrating that specificity was not compromised by the inclusion of the second modification (Fig. 2C). To determine species cross-reactivity, we tested truncated SOMAmers (30-mers) that bind to human PCSK9 with K_d < 1 nM (3 with single and 38 with double modification) for their ability to bind to monkey, rat, and mouse PCSK9, which exhibit 96%, 77%, and 64% identity to PCSK9 from human, monkey, mouse, and rat. Horizontal bars represent median values.

**Fig. 2.** Binding properties of 40-mer and 30-mer ligands from single-modified and double-modified libraries. (A) Binding affinities of partially truncated 40-mer SOMAmers (30 nucleotides from the originally randomized region, plus 5 nucleotides from each primer end) to PCSK9 selected from various libraries. SOMAmers with K_d ≤ 1 nM are highlighted in the gray shaded region and ligands with K_d = 320 nM showed no detectable binding with up to 32 nM PCSK9. Median ligand K_d for each library is indicated with a black horizontal line. (B) SOMAmers with K_d ≤ 1 nM (highlighted in the gray shaded region in A) were truncated to 30 nucleotides (that is, only the originally random region), and binding affinities to PCSK9 were measured. Ligands highlighted in the gray shaded region retained equivalent binding affinity after truncation from 40-mers to 30-mers. Median ligand K_d for each library is indicated with a black horizontal line. (C) Binding characterization of selected PCSK9 SOMAmers for specificity and species cross-reactivity. Target binding specificity of high-affinity PCSK9 SOMAmers from single- and double-modified libraries to other proprotein convertases (PCs). Solution affinity measurements were carried out for high-affinity ligands (40-mers, n = 33) from single-modified and double-modified libraries. The aptamers below the dotted line at 100 nM affinity indicate no detectable binding at a 100 nM concentration of protein. (C. Inset) Species cross-reactivity of SOMAmers. The affinity of single-modified truncated 30-mer SOMAmers (K_d ≤ 1 nM) to PCSK9 from human, monkey, mouse, and rat. Horizontal bars represent median values.
76% amino acid sequence identity with human PCSK9, respectively. Most of the SOMAmers (34 out of 41, or 83%) bound to monkey PCSK9 with comparable affinity (Fig. 2C, Inset and SI Appendix, Table S6). In contrast, only 5 of these SOMAmers bound to rat and 6 bound to mouse (12% and 15%, respectively) with \( K_d \) < 45 nM, all of which were derived from libraries containing two modified nucleotides. Some of these SOMAmers, for example, those derived from Pp-dC/Nap-dU and Pp-dC/Tyr-dU libraries, can bind to both rodent and human/monkey PCSK9 with \( K_d \) < 1 nM (SI Appendix, Table S6), suggesting that augmented side chain diversity allows access to a larger fraction of epitopes on protein surfaces, including those that are shared among proteins from different species. This species cross-reactivity has practical benefit from the standpoint of identifying ligands that can be tested in animal models of efficacy without the need for alternating selection rounds between human and animal targets (33).

**SOMAmers with Two Modifications Show Greater Epitope Coverage.** Enhanced epitope coverage is also useful for the identification of pairs of ligands able to bind simultaneously to the same protein. With conventional aptamers, this has been difficult because of the strong tendency for recognition of electropositive epitopes. Whereas specific selection methods have been used to direct aptamers to different sites on the target protein, such as multivalent aptamer isolation (MAI-SELEX) (34), and array-based discovery platform for multivalent aptamers (AD-MAP) (35), this has been difficult for the same reason, especially when starting libraries with a single modification (but different side chains) are used for selections against the same protein, coupled with a selection strategy in which a primary SOMAmer is used in excess to block the first epitope and force the second SOMAmer to bind a noncompeting second epitope (36). To test whether expanded chemical diversity could facilitate access to additional epitopes, we used a multiplexed assay (Fig. 3) to screen SOMAmers for their ability to form sandwich pairs. Of the 96 unique SOMAmers (40-mers, \( K_d \) < 1 nM) identified from libraries with single and double modifications when tested in all pairwise combinations (9,216 pairs), 70 pairs showed 50-fold or greater target-dependent signal compared with no protein controls (Fig. 3B and SI Appendix, Figs. S6 and S7). The majority of high-signaling pairs had two modifications in both SOMAmers, and nearly all contained at least one SOMAmer with two modifications (Fig. 3B and C). Therefore, the use of libraries with two modified nucleotides clearly facilitates the identification of SOMAmer sandwich pairs, and by inference, allows extensive epitope coverage.

**Facilitating the Development of SOMAmer Sandwich Assays.** Based on dose–response curves with the 70 highest-signaling sandwich pairs, one pair, composed of a primary (capture) SOMAmer with a single modification (SL1061, dC/Pp-dU, \( K_d \) = 175 pM, SI Appendix, Fig. S8A) and a secondary (detection) SOMAmer with two modifications (SL1062, Nap-dC/Nap-dU, \( K_d \) = 531 pM, SI Appendix, Fig. S8B), showed a linear, sensitive, and quantitative signal with both wild-type PCSK9 and gain-of-function mutant D374Y (SI Appendix, Fig. S8D). For this pair, orientation of the capture and detection probes was critical, with signal abolished when the orientation was reversed (SI Appendix, Fig. S8C).

PCSK9 is involved in cholesterol homeostasis and several gain-of-function (GOF) mutations in PCSK9 cause hypercholesterolemia (29). PCSK9 binds to LDL-R and promotes its degradation in lysosomes preventing recycling of LDL-R to the surface of hepatocytes, which results in increased plasma LDL cholesterol (37). The US Food and Drug Administration (FDA) has recently approved two drugs targeting PCSK9 (Alirocumab and Evolocumab) for the treatment of high-risk individuals with hypercholesterolemia (38). In addition, the quantification of PCSK9 in plasma by immunoaffinity (IA) enrichment LC-MS with a SOMAmer containing a single modified nucleotide was recently reported (39). PCSK9 D374Y mutant has higher affinity for LDL-R than the wild-type PCSK9 and is reported to be overexpressed in patients with the severe form of familial hypercholesterolemia (FH) (40). With the goal of developing a SOMAmer sandwich assay to detect circulating concentrations of plasma PCSK9 in human clinical samples, we evaluated the limit of quantification (SI Appendix, Fig. S9A and B and Table S7), precision (SI Appendix, Table S8), accuracy (SI Appendix, Table S9), and plasma dilution linearity (SI Appendix, Fig. S9C) to validate assay performance. We then used this assay to show that a group of subjects on atorvastatin therapy (\( n = 42 \)) has statistically significantly higher plasma levels of PCSK9 (\( P = 0.0044 \)) than a control group (\( n = 42 \)), as expected (41) (Fig. 3D). Such a companion diagnostic assay could be used to identify patients with elevated plasma PCSK9 who might benefit from anti-PCSK9 therapy. We also used this SOMAmer sandwich assay to measure PCSK9 concentrations in cell-free supernatants in PCSK9 overexpressing HepG2 cells, demonstrating the utility of these reagents as research tools (SI Appendix, Fig. S9D).

**Second Modification Enhances Nuclease Resistance.** Nuclease resistance is an important feature for many applications of aptamers, especially for therapeutic uses. We reported previously that SOMAmers with single 5-position modifications on dU impart a considerable degree of nuclease protection (31). To assess the impact of a second modified nucleotide, we tested the stability of the highest-affinity 30-mers from each library in the presence of human serum at 37 °C. SOMAmers with two modified nucleotides showed considerably enhanced stability compared with those containing a single modified nucleotide, and the best stability was observed with combinations of Pp-dC/Pp-dU, Nap-dC/Tyr-dU, Pp-dC/Nap-dU, and Pp-dC/Tyr-dU (<20% degradation over 96 h; Fig. 4A and SI Appendix, Fig. S10).

---

**Fig. 3.** Identification of sandwich pairs and PCSK9 measurements. (A) Screening of sandwich pairs in a multiplex bead-based assay (Luminex). Individual avidin-conjugated bead showing captured biotinylated primary SOMAmer binding to PCSK9 and secondary SOMAmer binding to second noncompeting epitope resulting in luminescence signal through phycoerythrin conjugate of streptavidin (SA-PE) and biotin interaction. (B) Number of sandwich pairs showing high relative signal (≥50) for each library combination. (C) Number of sandwich pairs obtained from combinations of single/single (SS), single/double (S/D), and double/double (D/D) modified libraries. (D) Statistical differentiation of patients on statin therapy vs. no statin therapy using the SOMAmer sandwich assay. PCSK9 levels were measured in 84 plasma samples divided into two groups, a control group (\( n = 42 \)) and a study group in which subjects were on atorvastatin therapy (\( n = 42 \), by self-report). Horizontal bars represent median values.
SOMAmers with Two Modifications Are the Most Potent PCSK9 Inhibitors. It is now well-established that PCSK9 antagonism is an effective therapeutic option for lowering plasma cholesterol levels (38). Aside from antibodies that have received approval (38), there is considerable interest in identifying other types of PCSK9 antagonists as potential therapeutics (42–44). To identify a SOMAmer inhibitor of PCSK9, 41 SOMAmers (30-mers, $K_d < 1 \text{nM}$) were screened in a plate-based sandwich assay where biotinylated PCSK9 was incubated with or without SOMAmer, added to an LDL-R-coated plate, and detected using streptavidin-HRP conjugate in a chemiluminescent readout (SI Appendix, Fig. S11). Over 70% of the SOMAmers showed $>90\%$ inhibition of PCSK9 in this assay, and 41% exhibited IC$_{50}$ values of 0.1–1 nM (Fig. 4B). One representative SOMAmer identified from the Pp-dC/Nap-dU library (SL1063, 30-mer) potently inhibited LDL-R-mediated internalization of fluorescently labeled LDL induced by both the wild-type human PCSK9 (IC$_{50}$ = 2.8 nM) and the D374Y mutant (IC$_{50}$ = 35 pM) (Fig. 4C). The 80-fold higher potency against the D374Y mutant is larger than the 3-fold difference in affinity of SL1063 for the two forms of the protein (wild type, $K_d = 14.7 \text{pM}$; D374Y mutant, $K_d = 5.2 \text{pM}$). The reason for this discrepancy remains to be elucidated; however, it is possible that a combination of steric and/or kinetic effects related to PCSK9-induced LDL cellular uptake result in a more effective functional antagonism of the D374Y mutant than could be explained by the differences in the $K_d$ values. SL1063 also showed high-affinity binding to PCSK9 from Rhesus monkey ($K_d = 11.3 \text{pM}$), mouse ($K_d = 77 \text{pM}$), and rat ($K_d = 165 \text{pM}$) (SI Appendix, Fig. S12), but did not bind other human PCs (Fig. 2C). Furthermore, this SOMAmer neutralized PCSK9 activity and subsequent LDL-R degradation in wild-type HepG2 cells (IC$_{50}$ = 13.5 nM) in a fluorescently labeled LDL uptake reversal assay (SI Appendix, Fig. S13) and increased the LDL-R expression levels in HepG2 cells treated with wild-type PCSK9 in a concentration-dependent manner (IC$_{50}$ = 312 nM, Fig. 4D). The high affinity, species cross-reactivity, inhibitory potency, and metabolic stability (SI Appendix, Fig. S10) of SL1063 highlight the therapeutic potential of SOMAmers with two modified nucleotides.

Discussion

The ability to introduce multiple, protein-like modifications into random nucleic acid libraries on two of four nucleobases reported here greatly expands their chemical diversity and allows the selection of ligands with higher affinity compared with ligands with only a single modification. Because there is not much bias in the base composition of the starting libraries, even though some of the combinations of modifications are clearly less efficiently incorporated by the KOD polymerase, the sequence space coverage is either not substantially compromised or is offset by the gain in chemical diversity. Within the set of modifications tested, the pairing of Tyr-dU or Pp-dC produced the best-performing combinations. The synergistic benefit of combining two modified nucleotides in the same library is apparent in the fact that libraries with either Tyr-dU or Pp-dC alone were not sufficient to generate high-affinity ligands. An assessment of the degree to which these observations are generalizable to a broad range of proteins will be an important goal of future experiments. The higher density of modified nucleotides available in libraries with two modifications also permitted more facile truncation to sequences within the 30N random region, where all modifications were confined by design. Importantly, enhanced affinity and more efficient encoding of binding domains were achieved while maintaining high specificity.

Another advantage of using libraries with two modified nucleotides is enhanced epitope coverage, which is evidenced by the ability to identify ligands that are cross-reactive with PCSK9 from different species, and the fact that sandwich pairs can be identified with both higher frequency and improved performance. One should take care not to misinterpret the species cross-reactivity as a lack of specificity, but instead as an example of the ability to identify ligands that specifically bind regions of high epitope conservation, made possible by the enhanced epitope coverage. It is worth noting, the demand for optimal performance in a sandwich assay where one of the ligands is immobilized on a surface is distinct from that for binding of a single ligand in solution, where only one (presumably optimal) epitope needs to be recognized. The composition of the best performing sandwich pair may well be idiosyncratic depending on the nature of the target. We have shown that such a
SOMAmer-based sandwich assay can be used to measure PCSK9 concentration in human clinical samples with high sensitivity. For development of therapeutics, many of the features described above, such as high affinity, specificity, ligand efficiency, and species cross-reactivity, are highly desirable. We have shown that the 11 most potent inhibitory SOMAmeres identified, 9 are derived from libraries with two modifications. Dramatically enhanced nuclease resistance observed for most SOMAmeres with two modifications, without the need of sugar-phosphate backbone modifications adds an important and exciting feature for nucleic acid ligands intended for therapeutic use. The results obtained with the selection method described here suggest that increased structural and functional group diversity afforded by libraries comprising two modified nucleotides represents the next stage of improvement in SELEX utility.

Materials and Methods
Modified cytidine and uridine triphosphates and phosphoramidites were synthesized in-house (12, 28). Random library preparation and six rounds of selection for binding were performed as described (20, 31) (SI Appendix, Fig. S2). Representative SOMAmeres were chemically synthesized and tested for binding activity to PCSK9 using a filter binding assay. The metabolic stability of SOMAmeres was assessed (31) in 90% human serum at 37 °C and the fraction of full-length SOMAmer was determined by polyacrylamide gel electrophoresis (PAGE). SOMAmer sandwich pairs were screened for binding to noncompeting epitopes on PCSK9 in a multiplexed Luminex assay (36) (Fig. 3A) and plasma measurements of PCSK9 were made in the same assay format. SOMAmer inhibition of the interaction between PCSK9 and LDL-R was measured with a plate-based sandwich assay (45) (SI Appendix, Fig. S11) with 1 nM PCSK9. Uptake of fluorescently labeled LDL in HepG2 cells was used as a measure of cellular LDL-R activity to assess the effect of PCSK9 inhibition on LDL-R expression (46) (SI Appendix, Fig. S12), and LDL-R recovery in PCSK9-treated HepG2 cells was measured by flow cytometry (46, 47). Detailed materials and methods can be found in SI Appendix, SI Materials and Methods.

ACKNOWLEDGMENTS. We thank Cate Flower, Brian Ream, Greg Wardle, Mike Nelson, and Michelle Carlson for help with synthesis of modified dNTPs, phosphoramidites, and oligonucleotides; Amy Gellin, Kirk DeIsle, and Fintan Steele for help with figures; and our SomaLogic colleagues for many insightful discussions.

1. Tuerk C, Gold L (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 249(4968):505–510.
2. Ellington AD, Szostak JW (1990) In vitro selection of RNA molecules that bind specific ligands. Nature 346(6287):868–871.
3. Robertson DL, Joyce GF (1990) Selection in vitro of a DNA enzyme that specifically cleaves single-stranded DNA. Nature 344(6265):467–468.
4. Sealf K, et al. (2014) In vitro selection with artificial expanded genetic information systems. Proc Natl Acad Sci USA 111(44):1449-1454.
5. Kimo M, Yamashige R, Matsuura K, Yokoyama S, Hirota I (2013) Generation of high-affinity DNA aptamers using an expanded genetic alphabet. Nat Biotechnol 31(5):453–457.
6. Pinheiro VB, et al. (2012) Synthetic genetic polymers capable of heredity and evolution. Science 336(6079):341–344.
7. Taylor Al, et al. (2015) Catalysts from synthetic genetic polymers. Nature 518(7539):427–430.
8. Grouiller L, et al. (2012) Efficient reverse transcription using locked nucleic acid nucleotides towards the evolution of nucleic resistant RNA aptamers. PLoS One 7(4):e35990.
9. O’Lea C, et al., Weidmann J, Dawson PE, Joyce GF (2015) An L-RNA aptamer that binds and inhibits influenza. Chem Biol 22(11):1437–1441.
10. Roholff JC, et al. (2014) Nuclear acid ligands with protein-like side chains: Modified aptamers and their use as diagnostic and therapeutic agents. Mol Ther Nucleic Acids 3:e201.
11. Chen T, et al. (2016) Evolution of thermophila DNA polymers for the recognition and amplification of C2-modified DNA. Nat Chem 8(6):556–562.
12. Vaught JD, et al. (2010) Expanding the chemistry of DNA for in vitro selection. J Am Chem Soc 132(12):4141–4151.
13. Perrin DN, Garret-Tweed T, Chebl A (1999) Expanding the catalytic repertoire of nucleic acid catalysts: Simultaneous incorporation of two modified deoxynucleobis(adenosine) triphosphates bearing ammoinium and imidazolyl functional groups. Nucleosides Nucleotides Nucleic Acids 18(3):377–391.
14. Thomas JM, Yoon JK, Perrin DM (2009) Investigation of the catalytic mechanism of a synthetic DNAzyme with protein-like functionality: an RNAsa mimic. J Am Chem Soc 131(15):5648–5658.
15. Latham JA, Johnson R, Toole JJ (1994) The application of a modified nucleotide in aptamer selection: Novel thrombin aptamers containing 5-(1-pentynyl)-2-deoxyuridine. Nucleic Acids Res 22(14):2817–2822.
16. Santoro SW, Joyce GF, Sakhivel V, Gramatikova S, Barbas CF, 3rd (2000) RNA cleavage by a DNA enzyme with enhanced chemical functionality. J Am Chem Soc 122(11):2433–2439.
17. Sidorov AV, Grabsky JA, Williams DM (2004) Sequence-specific cleavage of RNA in the absence of divalent metal ions by a DNAzyme incorporating imidazoyl and amino functional groups. Nucleic Acids Res 32(4):1591–1601.
18. Tarasov TM, Tarasov SL, Eaton BE (1997) RNA-catalysed carbon-carbon bond formation. Nature 386(6646):54–57.
19. Vaish NK, Larralde R, Fraley AW, Szotak JW, McLaughlin LW (2003) A novel modification-dependent ATP-binding aptamer selected from an RNA library incorporating a cationic functionality. Biochemistry 42(29):8842-8851.
20. Gold L, et al. (2010) Aptamer-based multiplexed protein technology for biomarker discovery. PLoS One 5(12):e15004.
21. Davies DR, et al. (2012) Unique motifs and hydrophobic interactions shape the binding of modified DNA ligands to protein targets. Proc Natl Acad Sci USA 109(49):19637–19642.
22. Gelinas AD, et al. (2014) Crystal structure of interleukin-6 in complex with a modified nucleic acid ligand. J Biol Chem 289(2):8720–8734.
23. Gelinas AD, Davies DR, Janicij N (2016) Embracing proteins: Structural themes in aptamer-protein complexes. Curr Opin Struct Biol 36:122–132.