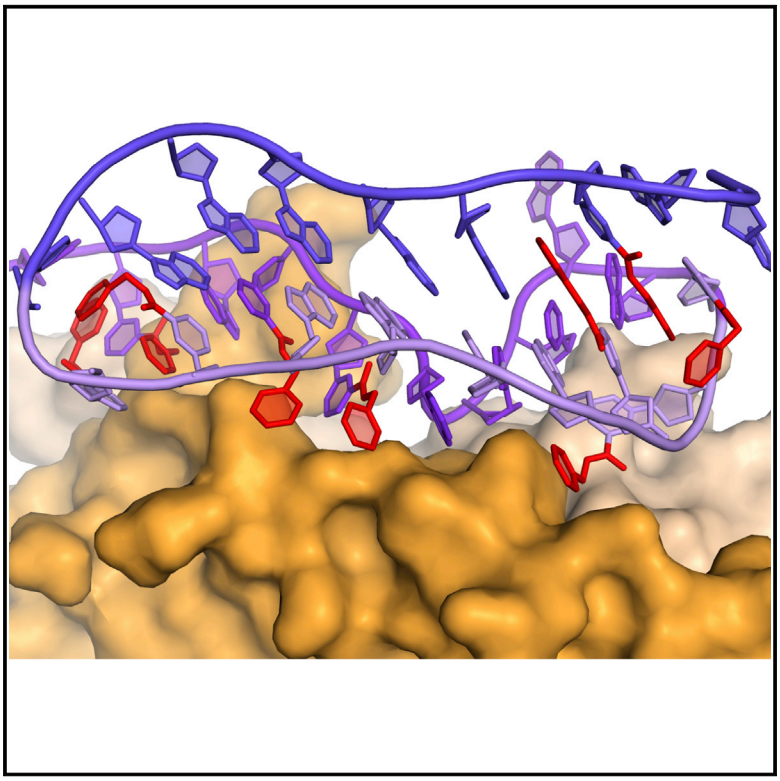
Article

Structure

Non-helical DNA Triplex Forms a Unique Aptamer Scaffold for High Affinity Recognition of Nerve Growth Factor

Graphical Abstract



Authors

Thale C. Jarvis, Douglas R. Davies, Akihiko Hisaminato, ..., Amy D. Gelinas, Daniel J. Schneider, Nebojsa Janjic

Correspondence

njanjic@somalogic.com

In Brief

Jarvis et al. report the structure of a modified DNA aptamer bound to beta nerve growth factor. The aptamer folds into an unusual configuration that lacks any known nucleic acid structural motifs, while hydrophobic modifications at the 5-position of deoxyuridine provide enhanced chemical diversity for unique intra- and intermolecular interactions.

Highlights

- A high affinity modified DNA aptamer binds and neutralizes nerve growth factor
- The co-crystal structure reveals an unprecedented triangular prism DNA structure
- DNA structure is completely non-helical with unusual base pairings and bond angles
- Primarily hydrophobic contacts with few H bonds at complex interface

Accession Numbers

4ZBN

Jarvis et al., 2015, Structure *23*, 1293–1304
July 7, 2015
http://dx.doi.org/10.1016/j.str.2015.03.027

Structure
Article

Non-helical DNA Triplex Forms a Unique Aptamer Scaffold for High Affinity Recognition of Nerve Growth Factor

Thale C. Jarvis,[1,7] Douglas R. Davies,[2,7] Akihiko Hisaminato,[3,7] Daniel I. Resnicow,[1] Shashi Gupta,[1] Sheela M. Waugh,[1] Akira Nagabukuro,[3] Takashi Wadatsu,[3] Haretsugu Hishigaki,[3] Bharat Gawande,[1] Chi Zhang,[1] Steven K. Wolk,[1] Wesley S. Mayfield,[1] Yuichiro Nakaishi,[4] Alex B. Burgin,[5] Lance J. Stewart,[6] Thomas E. Edwards,[2] Amy D. Gelinas,[1] Daniel J. Schneider,[1] and Nebojsa Janjic[1,*]

[1]SomaLogic, Inc., 2945 Wilderness Place, Boulder, CO 80301, USA
[2]Beryllium, 7869 Northeast Day Road West, Bainbridge Island, WA 98110, USA
[3]Otsuka Pharmaceutical Co., Ltd., Institute of Biomedical Innovation, 463-10 Kagasuno, Kawauchi-cho, Tokushima 771-0192, Japan
[4]Otsuka Pharmaceutical Co., Ltd., Medicinal Chemistry Research Institute, 463-10 Kagasuno, Kawauchi-cho, Tokushima 771-0192, Japan
[5]The Broad Institute, 7 Cambridge Center, Cambridge, MA 02142, USA
[6]The Institute for Protein Design, University of Washington, 3946 West Stevens Way Northeast, Seattle, WA 98195-1655, USA
[7]Co-first author
*Correspondence: njanjic@somalogic.com
http://dx.doi.org/10.1016/j.str.2015.03.027

SUMMARY

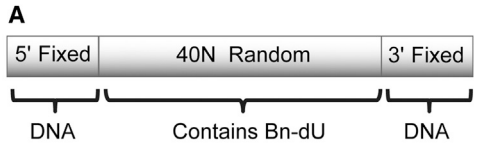
Discerning the structural building blocks of macromolecules is essential for understanding their folding and function. For a new generation of modified nucleic acid ligands (called slow off-rate modified aptamers or SOMAmers), we previously observed essential functions of hydrophobic aromatic side chains in the context of well-known nucleic acid motifs. Here we report a 2.45-Å resolution crystal structure of a SOMAmer complexed with nerve growth factor that lacks any known nucleic acid motifs, instead adopting a configuration akin to a triangular prism. The SOMAmer utilizes extensive hydrophobic stacking interactions, non-canonical base pairing and irregular purine glycosidic bond angles to adopt a completely non-helical, compact S-shaped structure. Aromatic side chains contribute to folding by creating an unprecedented intercalating zipper-like motif and a prominent hydrophobic core. The structure provides compelling rationale for potent inhibitory activity of the SOMAmer and adds entirely novel motifs to the repertoire of structural elements uniquely available to SOMAmers.

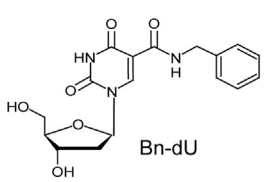
INTRODUCTION

The in vitro process of SELEX (systematic evolution of ligands by exponential enrichment) has been used extensively to identify nucleic acid ligands to biological targets of interest. However, given the inherent limitations of natural nucleotides, certain targets remained refractory to traditional SELEX methods. The use of modified nucleotides with enhanced chemical diversity has dramatically increased the success rate of SELEX (Gold et al., 2010; Vaught et al., 2010). Such chemically modified aptamers are particularly well-suited to selection schemes that favor long-lived complexes; hence, we refer to this class of aptamers as SOMAmers (slow off-rate modified aptamers). SOMAmers have transformed the landscape of proteomics, leading to a sensitive and highly multiplexed platform with applications in biomarker discovery and early-stage disease detection (De Groote et al., 2013; Gold et al., 2010, 2012; Loffredo et al., 2013; Lollo et al., 2014; Mehan et al., 2012, 2013; Ostroff et al., 2010, 2012), as well as sandwich-based single analyte detection assays (Ochsner et al., 2013, 2014), rapid histochemistry tools (Gupta et al., 2011), and therapeutic applications (Gelinas et al., 2014; Gupta et al., 2014).

Herein we describe a SOMAmer that binds and neutralizes beta nerve growth factor (NGF), a member of the neurotrophin family of secreted proteins critical for development and maintenance of the nervous system in mammals. NGF functions as an important mediator of itching, pain, and inflammation, and inhibition of NGF-mediated signaling has been explored as a treatment for a variety of disorders including neuropathic pain, nociceptive pain, osteoarthritis pain, and fibromyalgia (Ghilardi et al., 2011; Hefti et al., 2006; McKelvey et al., 2013). NGF binds two cell surface receptors, tropomyosin receptor kinase (TrkA) and p75 neurotrophin receptor ($p75^{NTR}$), a member of the tumor necrosis family of receptors. NGF binding to TrkA stimulates autophosphorylation and consequent signal transduction via the mitogen-activated protein kinase and phosphatidylinositol 3-kinase pathways, promoting neuronal survival and/or differentiation. NGF binding to $p75^{NTR}$ potentiates TrkA signaling (Ceni et al., 2010; Eibl et al., 2012). Several structures of NGF, apo and receptor bound, have been published (He and Garcia, 2004; McDonald et al., 1991; Wehrman et al., 2007; Wiesmann et al., 1999). The active form of NGF is a non-covalent homodimer with tightly associated monomers that fold into a pair of β-pleated strands, with four intervening loop regions (Bothwell and Shooter, 1977).

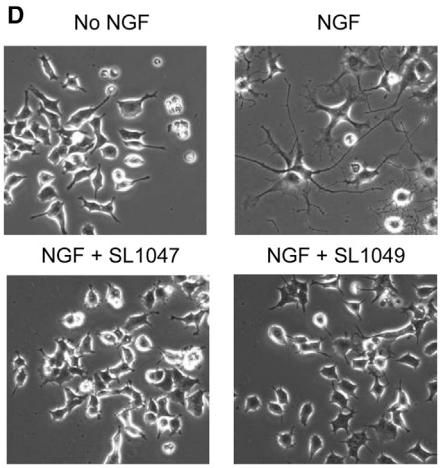
NGF has been a SELEX target in the past, with the aptamers identified as having either relatively weak binding affinity (K_d ~200 nM) (Binkley et al., 1995) or requiring long sequences

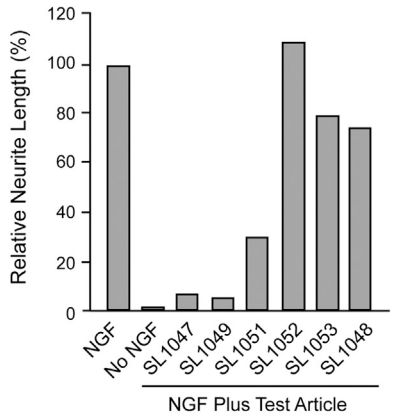


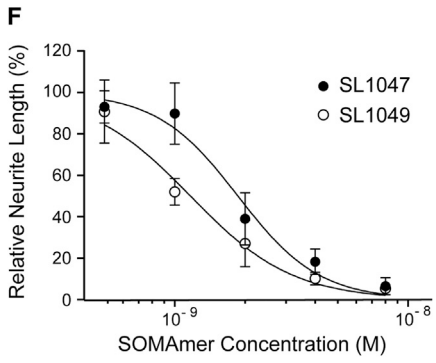


C

SOMAmer	Length	Sequence	K_d (nM)
SL1047	76	...CAGCGGGACACAZGAGGACZZGGGGZZZAGCCGZZGZGG...	0.14 ± 0.04
SL1048	50	...CAGCGGGACACAZGAGGACZZGGGGZZZAGCCGZ	>100
SL1049	28	CAZGAGGACZZGGGGZZZAGCCGZZGZG	0.21 ± 0.08
SL1051	27	AZGAGGACZZGGGGZZZAGCCGZZGZG	0.09 ± 0.03
SL1052	26	ZGAGGACZZGGGGZZZAGCCGZZGZG	0.30 ± 0.04
SL1053	25	GAGGACZZGGGGZZZAGCCGZZGZG	1.47 ± 0.65







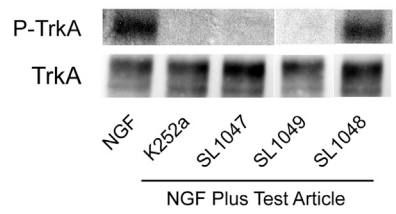


Figure 1. Short NGF SOMAmers Potently Inhibit NGF Activity In Vitro

(A) Key elements of the random DNA SELEX library, including the random region (comprising dA, dC, dG, and Bn-dU nucleotides) and the 5′ and 3′ fixed regions for primer binding.

(B) Structure of 5-(*N*-benzylcarboxamide)-2′-deoxyuridine (Bn-dU).

(C) Sequences and binding affinities of SL1047 and its truncated variants. 5′ and 3′ fixed region sequences are represented by " … " and K_d values are reported as the mean ± SD of three replicate measurements.

(D) Effect of NGF SOMAmer on neurite outgrowth of PC12 cells stimulated by NGF. Cultured PC12 cells were treated for 5 days with or without NGF (0.38 nM) or NGF pre-equilibrated with test article (10 nM). Cells were photographed and neurite length was measured using the NeuronJ plugin for ImageJ.

(E) Neurite length plotted relative to that observed in the absence of added SOMAmer.

(F) Dose-response of NGF SOMAmers SL1047 and SL1049 on neurite outgrowth. Each data point represents the mean ± SD of three replicate measurements at each SOMAmer concentration.

(G) Inhibition of NGF-induced TrkA phosphorylation by NGF SOMAmers. Cultured PC12 cells were treated for 10 min with NGF (0.38 nM) or NGF pre-equilibrated with test article (0.2 μM K252a or 10 nM oligonucleotide). Cell lysates were immunoprecipitated with an anti-Trk antibody C-14 and resolved with SDS-PAGE. Phosphorylated and total TrkA were examined by anti-phospho-tyrosine antibody 4G10 and anti-Trk antibody C-14, respectively. See also Figure S1.

(>45 nucleotides) for high affinity binding (Nakamura et al., 2013). Here, we describe the identification of a considerably shorter SOMAmer that exhibits high affinity binding to NGF and potently inhibits NGF-mediated TrkA phosphorylation and neurite outgrowth in vitro.

To understand its mode of action, we solved the crystal structure of this SOMAmer bound to recombinant human NGF to 2.45-Å resolution. We have previously reported the co-crystal structures of two SOMAmer-target complexes (Davies et al., 2012; Gelinas et al., 2014). These structures revealed extensive utilization of modified nucleotide side chains at the protein binding interface, serving to mimic native receptor interactions. The resulting SOMAmer-protein binding surfaces exhibit more hydrophobic character and considerably fewer polar interactions compared with traditional aptamers (Davies et al., 2012). Both structures utilized scaffolds composed of familiar nucleic acid structural motifs such as pseudoknots, G-quartets, and hairpins, which were augmented by modified nucleotides. In the present work, we report the structure of a third SOMAmer-target complex that reveals a marked departure from previous observations and highlights the extraordinary structural plasticity afforded by selective chemical modifications of natural nucleotides.

RESULTS

Identification of an NGF SOMAmer by SELEX

To identify a SOMAmer to NGF, we performed SELEX with a modified DNA library containing the hydrophobic nucleotide 5-(*N*-benzylcarboxamide)-2′-deoxyuridine (Bn-dU) replacing dT (Figures 1A and 1B). Following affinity screening, we chose SL1047 for further analysis, one of the highest affinity SOMAmers that belongs to a family of related sequences comprising about 9% of the affinity-enriched pool (Figure S1A).

Table 1. Crystallographic Data Collection and Refinement Statistics

Data Collection	Native	NaI Soak
Space group	C2	C2
Unit-cell parameters: *a, b, c* (Å), β (°)	109.20, 60.71, 68.59, 95.57	107.46, 60.60, 67.44, 95.78
Wavelength (Å)	0.97950	1.54
Resolution range (Å)	50.00–2.45 (2.00–2.45)	50.00–3.15 (3.23–3.15)
Unique reflections	15,532	14,556
Completeness (%)	98.2 (84.1)	99.3 (100)
R_{merge}	0.038 (0.303)	0.101 (0.372)
Mean I/σ (I)	23.3 (2.03)	10.92 (2.97)
Refinement	**Native**	**N/A**
Resolution range (Å)	50.00–2.45 (2.51–2.45)	
R_{cryst}	0.202 (0.380)	
R_{free}	0.257 (0.342)	
RMSD bonds (Å)	0.0134	
RMSD angles (°)	2.112	
Total no. of atoms	2,883	
Wilson *B* factor (Å^2)	55.28	
Mean *B* factor (Å^2), all atoms	55.61	
Mean *B* factor (Å^2), protein atoms	53.92	
Mean *B* factor (Å^2), nucleic acid atoms	59.68	
Error in coordinates by Luzzati plot (Å)	0.405	
Disordered side chains	11	
Residues in favored regions (%)	175 (96.15)	
Residues in disallowed regions (%)	0 (0.00)	
MolProbity score (percentile)	2.12 (91st)	
PDB code	PDB: 4ZBN	

Values in parentheses indicate the values for the highest of 20 resolution shells.

$R_{merge} = \Sigma_h\Sigma_i|I_i(h) - \langle I(h)\rangle|/\Sigma_h\Sigma_i I_i(h)$.

$R_{free} = \Sigma_h||F_{obs}| - |F_{calc}||/\Sigma_h|F_{obs}|$.

The free R factor was calculated using 5% of the reflections omitted from the refinement (Collaborative Computational Project, Number 4, 1994).

We determined the minimal sequence for SL1047 by comparing the affinities of synthetic variants truncated systematically from the 5′ and 3′ ends. Several truncated variants comprised of 25–28 nucleotides exhibited equal or better affinity compared with the full-length sequence (Figure 1C). The full-length SOMAmer (SL1047) and the 28-mer truncated variant (SL1049) inhibited NGF-mediated neurite outgrowth in PC12 cells in vitro (Figures 1D and 1E) in a concentration-dependent manner with half maximal inhibitory concentration values of 2 nM and 1 nM, respectively (Figure 1F). Shorter binding-competent variants were less active (27-mer SL1051) or inactive (26-mer SL1052 and 25-mer SL1053) (Figures 1C and 1E). The reason for the lack of activity of these shorter variants is not clear (even with the benefit of the structure; see below); we have speculated these results could be due to structural rearrangement of the SOMAmer, decreased stability of the SOMAmer structure (or the complex), or less efficient binding in the cell medium and assay conditions. The control sequence, SL1048, with deletion of the 3′ region essential for high affinity binding was also inactive (Figures 1C and 1E). SL1047 and SL1049 at 10 nM also inhibited NGF-induced phosphorylation of TrkA to the same degree as the small molecule tyrosine kinase inhibitor K252a (Tapley et al., 1992) at 200 nM, while SL1048 showed minimal inhibition (Figure 1G). SL1049 exhibited a relatively rapid association rate ($k_{on} = 8.0 \times 10^5\ M^{-1}\ s^{-1}$) and slow dissociation rate ($k_{off} = 2.5 \times 10^{-4}\ s^{-1}$; $t_{1/2} = 46$ min) (Figures S1B–S1D), with a calculated K_d value of 0.3 nM, in close agreement with equilibrium measurements (Figure 1C). SL1049 was highly selective for NGF, with more than 100-fold lower affinity to related neurotrophins NT3, NT4, and brain-derived neurotrophic factor (Figure S1E), and showed similar high affinity to human, mouse, and rat NGF (Figure S1F).

Crystallization of SOMAmer-NGF Complexes

To gain insight into the mode of binding and inhibition by the SOMAmer, we determined the crystal structure of SL1049 in a complex with NGF. Size exclusion chromatography indicated that SL1049 formed a complex with NGF in a 1:1 stoichiometry (i.e., two SOMAmer molecules per NGF dimer; data not shown). Initial sparse matrix screens for crystallization conditions produced crystal hits with the human NGF ortholog but not with the mouse or rat orthologs. Native X-ray data were collected at 2.45-Å resolution and the structure was solved through a combination of molecular replacement and heavy atom phasing techniques (Table 1; Figure S3).

Structure Overview

Two SOMAmers (chains D and E) bind to symmetrical sites at the NGF dimer interface (chains A and B) (Figure 2A). The structure of NGF protein within the co-crystal is globally similar to previously published NGF structures (He and Garcia, 2004; Holland et al., 1994; Wehrman et al., 2007; Wiesmann et al., 1999), although the conformations of some loops and side chains are perturbed by interactions with the SOMAmer. A few short loops are disordered in the SOMAmer-bound structure, including Asn45-Asn46 and Pro61-Asp65, and Asp93-Gln96 (only disordered in chain B). The only significant asymmetry in the dimeric crystal structure is evident at the N terminus of NGF chain B and the C terminus of chain A. Residues His4-Glu10 and Arg114 are ordered only in chains B and A, respectively. This helix interacts with the 3′ end of SOMAmer chain E. The asymmetric model likely arises due to a crystal contact between His4 on the NGF B chain and dC1 on the SOMAmer E chain of a symmetry mate. The SOMAmer chain D is virtually identical to chain E (root-mean-square [RMS] = 0.315, 521 atoms), with the exception of the crystal contact mentioned above and the extruded base, dG20. We focus on the SOMAmer chain E in subsequent figures as it illuminates more protein contacts with the N-terminal amino acids of NGF than chain D.

The SOMAmer adopts a compact S-shaped fold with a striking absence of helicity (Figures 2B and 2C). The structure can be

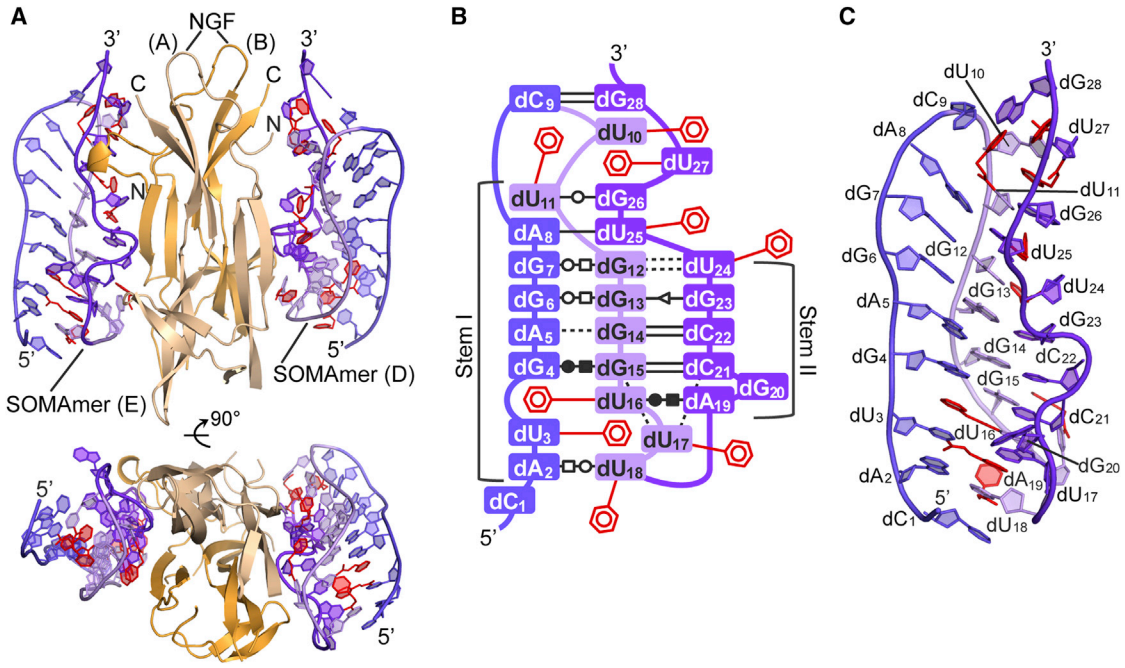


Figure 2. SOMAmer Forms a Compact S-Shape and Binds NGF at the Dimer Interface

The NGF homodimer is represented as wheat (chain A) and gold (chain B) colored ribbons and the SOMAmer (chains D and E) is shown as a backbone trace with the bases colored by strand. The strands are numbered from 5′ to 3′: strand 1 (slate blue; nucleotides 1–9), strand 2 (light purple, nucleotides 10–18), and strand 3 (purple, nucleotides 19–28). Modified nucleotides are colored red.

(A) The upper view shows the complex perpendicular to the non-crystallographic 2-fold axis of symmetry; the lower view is parallel to this 2-fold axis of symmetry, viewed from the 5′ end.

(B) Cartoon rendition of the SOMAmer, showing base-pairing and stacking interactions, with base pairs denoted according to the nomenclature of Leontis and Westhof (2001).

(C) SOMAmer chain E, colored by strand with modified nucleotides colored red. See also Figure S2.

divided into three strands delineated by tight backbone turns with strands 1–3 (Str1–3) comprising nucleotides 1–9, 10–18, and 19–28, respectively (Figures 2B and 2C). The SOMAmer contains several secondary structure features that are common in RNA but less common in previously observed DNA structures, such as Hoogsteen base pairing, triple strand, and purine-phosphate interactions. Remarkably, the tightly folded structure makes extensive use of non-canonical base pairs, with only four Watson-Crick (WC) base pairs out of 17 total hydrogen bond-mediated base-pairing interactions (Figures 2B and S2C). Although the SOMAmer exhibits features suggestive of an extreme H-type pseudoknot, it defies categorization into classical pseudoknot sub-elements (Figure 2C) and the complete lack of helical twist renders the structure more reminiscent of a triangular prism. Three major structural influences contribute to the unconventional conformation adopted by the NGF SOMAmer: (1) intercalation and edge-to-face π stacking of modified nucleotides, (2) glycosidic bond conformations including both *syn* and *anti* purines, and (3) non-canonical base pairing, triple and quadruple bases, and base-to-backbone hydrogen bonding.

Structural Motifs: Intercalating Zipper and Benzyl Cluster

In a notable departure from previous SOMAmer structures, most of the benzyl side chains of modified nucleotides in the NGF SOMAmer (six of nine) do not contact the protein (Figures 3A–3F). Instead, these residues play a purely structural role, supporting the unusual conformation of the nucleic acid scaffold. At the 5′ end, an intriguing motif is the intercalating zipper formed by Bn-dU3 and Bn-dU16. The two Bn-dU nucleotides exhibit an offset dyad symmetry, allowing the teeth of the zipper to align based on reciprocal π-stacking of the benzyl moiety from one side with the uridine from the opposite side (Figures 3B, 3C, and S3A). Each benzyl ring (Bn) thereby effectively acts as a pseudo base, covalently linked to uridine, taking the place of a conventional hydrogen-bonded base pair. The zipper is supported by an edge-to-face π-stacking interaction with Bn18.

A second zipper-like motif is present in the benzyl cluster comprising Bn-dU10, Bn-dU11 and Bn-dU27, where Bn-dU10 and Bn-dU27 actually mimic the Bn-dU3:Bn-dU16 intercalating zipper, exhibiting the characteristic reciprocal stacking between the benzyl and uridine moieties from opposing strands (Figures 3E, 3F, and S3B). However, rotational freedom in the linkers allows both of the benzyl groups to twist, in apparent synchrony, out of the plane of the base to allow face-to-face stacking with the opposite uridine (Figures 3E and 3F). The non-planar nature of these modified nucleotides creates three distinct π-stacking systems, interconnected through Bn27. First, Bn11 stacks with dC9 and interacts edge-to-face with Bn27 (Figures 3E and 3F). Bn27 in turn intercalates between dU10 and dU11 (dU specifies

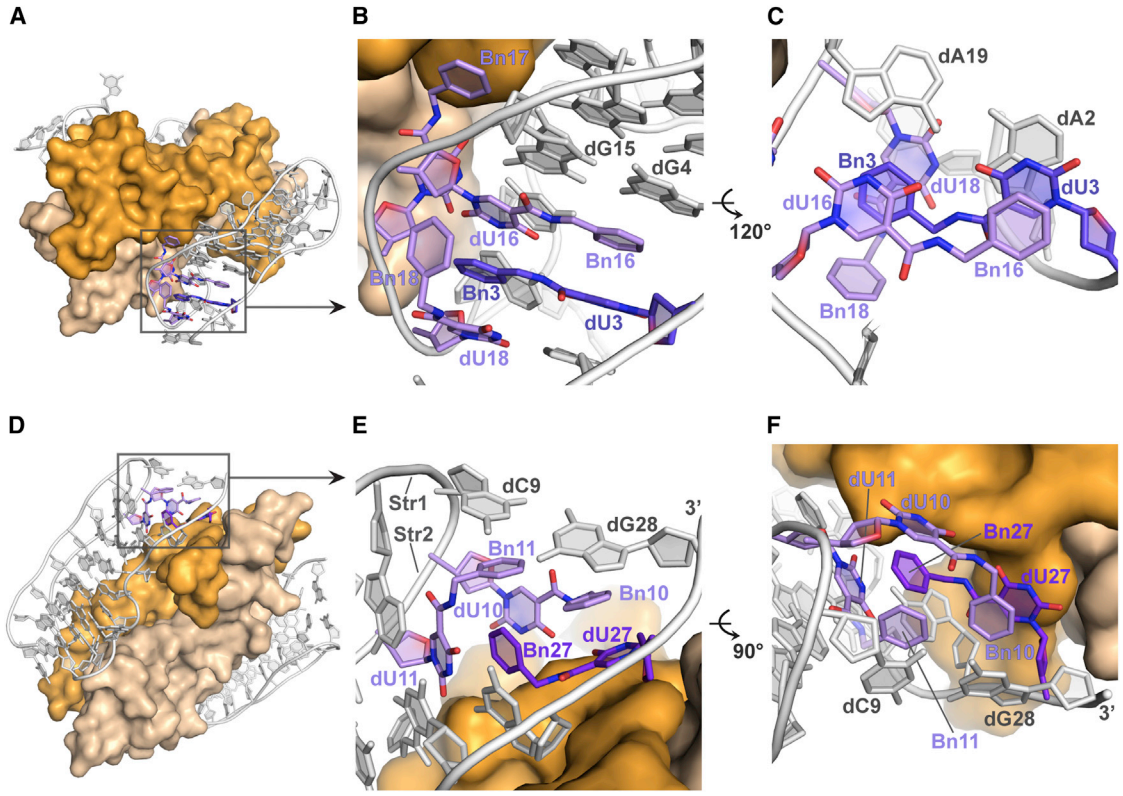


Figure 3. Modified Nucleotides Are Integral to Both the Scaffold and the Protein Binding Interface

(A) The benzyl zipper is located near the 5′ end of the SOMAmer and the Str2/Str3 turn.

(B) Zoomed in view of the Bn-dU3:Bn-dU16 zipper shows reciprocal π-stacking between uridines and benzyl moieties with additional aromatic interactions from neighboring Bn-dU18. Bn-dU17 is extruded by the zipper, allowing it to make significant contact with NGF.

(C) Rotated view of the benzyl zipper highlights the continuous stacking that results from the intercalated Bn-dU nucleotides.

(D) The benzyl cluster is located near the 3′ end of the SOMAmer and the Str1/Str2 turn.

(E) Zoomed in view of (D) shows the three π-stacking systems of the benzyl cluster, interconnected through Bn27. The dC9:dG28 base pair caps the benzyl cluster and connects Str1/Str2 of the SOMAmer to the 3′ end.

(F) Rotated view of the benzyl cluster. See also Figures S3 and S4.

just the uridine base of a modified nucleotide) creating the second aromatic stacking system (Figure 3E). The third system contains a quadruple stacking arrangement including both SOMAmer (dG28, Bn10, dU27) and NGF (Phe12) aromatic groups (Figures 3F and 5D); Bn10 also makes an edge-to-edge contact with Bn27 and an edge-to-face interaction with the sugar ring of dG28. The dC9:dG28 base pair serves as a cap for the hydrophobic benzyl cluster, protecting it from solvent while the extensive network of aromatic interactions connects the 3′ end of the SOMAmer to the backbone turn between Str1 and Str2 (Figures 3D and 3E). Although the benzyl cluster is primarily an internal structural motif where the benzyl side chains do not interact with the protein, all three nucleotides do make contact with NGF (Table S1).

Given the dual utilization of Bn-dU nucleotides at the protein binding interface and as internal structural elements, we wanted to determine the essential role of the benzyl groups for the structure and function of the NGF SOMAmer. To that end, we sequentially replaced each Bn-dU nucleotide with dT and assessed the effect on binding affinity. We found that only Bn-dU18 was dispensable within the SOMAmer, with all other substitutions resulting in significantly reduced binding affinities (from 6- to >100-fold) (Figure S2D). These results confirmed the critical function of the Bn-dU nucleotides in forming the unusual conformation that provides high affinity binding to NGF.

An Unconventional Triple-Stranded Core

The central region of the SOMAmer consists of a triple strand comprising dG4-dG7 (Str1), dG12-dG15 (Str2), and dC21-dU24 (Str3) (Figure 4A). The triple-stranded region of the NGF SOMAmer includes one base triple and three base quadruples (Figures 4 and S5). In the dG4-dG15-dC21-dU17 base quadruple (Figures 4B and S5A), dC21 forms five H bonds, thereby exhibiting complete utilization of every available H-bond donor and acceptor. The next base triple, dA5-dG14-dC22 (Figure 4C) involves a WC base pair between dG14 and dC22, and an unusual interaction between dG14 and dA5 involving at least one H bond between dG14 O6 and N6 of dA5. A second H bond between dG14 and dA5 could potentially occur if the adenosine is protonated at N1 (because its pKa is raised by the local environment) or assumes the imino

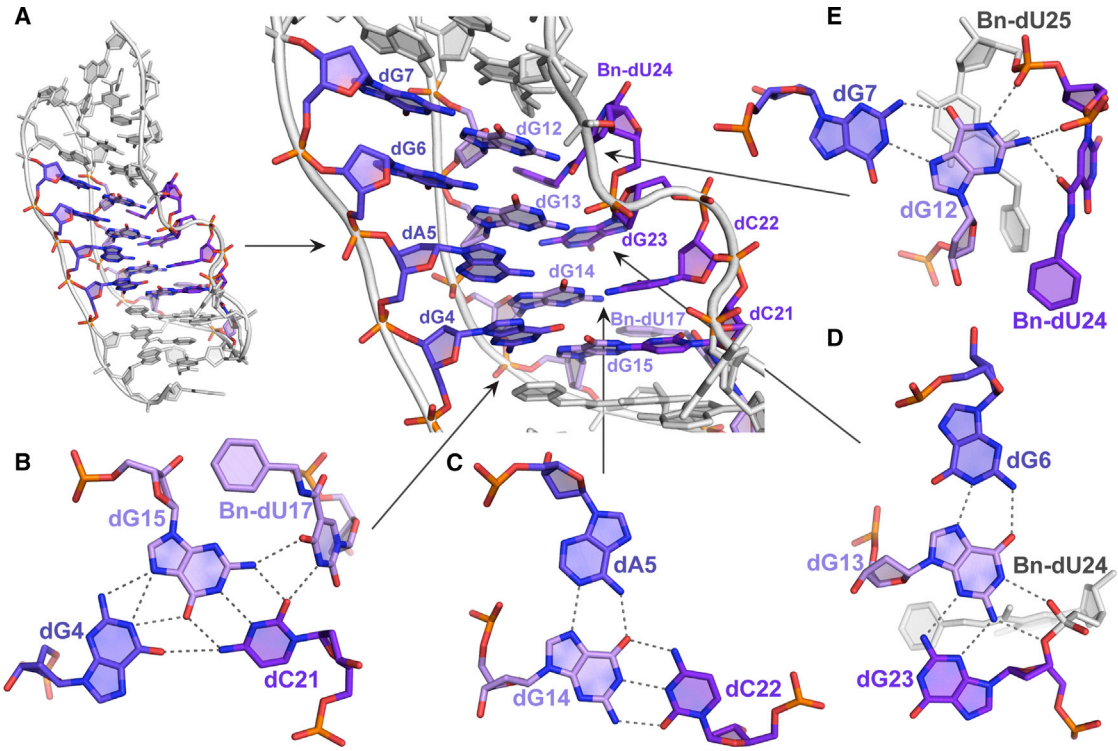


Figure 4. Core Region Composed of Base Triples and Base Quadruples

(A) One base triple and three base quadruples are located in the central region of the SOMAmer.

(B) A base quadruple consisting of dG4-dG15-Bn-dU17-dC21 has eight H bonds with dC21 utilizing every available H-bond donor and acceptor.

(C) The base triple of dA5-dG14 and dC22. The interaction between dA5 and dG14 includes one H bond between dG14 O6 and N6 of dA5 and a second potential H bond if dA5 adopts the imino tautomeric form with protonation of N1. The same type of H bond could be formed with dA5 protonated at N1.

(D) The dG6-dG13-dG23-dU24 base quadruple centered on dG13 exhibits nearly complete engagement of every available H-bond donor and acceptor. This quadruple includes two base-to-backbone H bonds between the WC face of dG13 and a non-bridging oxygen at U24, and to a bridging oxygen at dG23.

(E) The base quadruple consisting of dG7-dG12-dU24-dU25 includes two H bonds between the base of dG12 and the backbone at dU24 and dU25.

The arrow in (A) indicates the zoomed in region of the SOMAmer. All additional arrows point to the region of the SOMAmer where the corresponding bases can be seen. See also Figure S5.

tautomeric form with a proton at N1 (Figures 4C and S5B). dG6-dG13-dG23-dU24 form a base quadruple (Figures 4D and S5C) centered on dG13, which exhibits nearly complete engagement of every available H-bond donor and acceptor via WC, Hoogsteen and sugar edge base pairing and base-to-backbone interactions. In the final base quadruple, dG7-dG12-dU24-dU25 (Figures 4E and S5D), dG12 makes a Hoogsteen base pair with dG7 and three atypical H bonds on its WC face, one to the Bn-dU24 linker and the remaining two from base to backbone. Overall, this core region comprises a dense network of 24 H bonds. With additional stabilization contributed by an extensive system of π-π stacking interactions (Figures 2B and 3A), the central core clearly represents an important structural element in the folded organization of the SOMAmer. Consistent with these observations, SL1049 exhibits a thermal melting transition with a T_m value of 50°C, approximately 20°C higher than the unmodified DNA with the same sequence (Figures S4A and S4B). As observed previously with other SOMAmers (Gupta et al., 2014), SL1049 is also substantially more resistant to serum nucleases compared with its all-DNA equivalent (Figures S4C and S4D), which is likely due, in part, to its unusual structure.

Protein Recognition Domain Supported by a Flat Duplex Scaffold

The overall structure of the SOMAmer can be viewed as a triangular prism with the backbones of each strand defining the three edges. Two sides (Str1 and Str2) serve as a scaffold, while the third side (Str3) serves as the recognition domain for binding NGF. A 180° backbone turn between Str1 and Str2 forms a hairpin composed of almost entirely non-canonical base pairs that create an extraordinarily flat duplex (stem I; Figure 2B). Structurally, stem I benefits from contiguous stacking of multiple elements, beginning with a Hoogsteen dA2-dU18 base pair, which stacks with the intercalating zipper (Bn-dU3:Bn-dU16). The zipper stacks with four purine-purine base pairs in the triple-stranded core. Stem I is completed by two cross-over base pairs between Str1-Str3 (dA8-dU25) and Str2-Str3 (dU11-dG26), respectively. Stem I forms a planar duplex characterized by an almost non-existent minor groove and an extremely wide major groove. This is achieved through a combination of non-canonical base pairing, intercalation of the modified nucleotides, and unusual glycosidic bond conformations; more than half the purines (6 of 11 in stem I, and 8 of 15 in the SOMAmer altogether) adopt a *syn* conformation (Table S1). Between Str2 and Str3, the

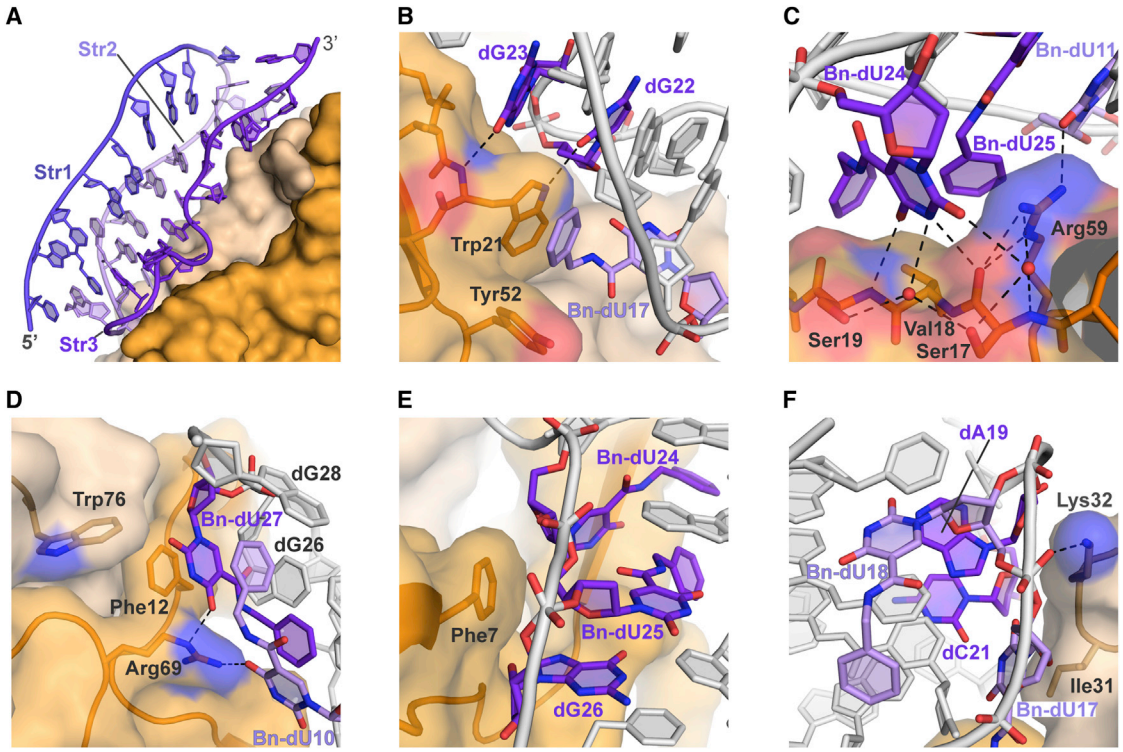


Figure 5. NGF-SOMAmer Binding Interface

(A) Str3 of the SOMAmer conforms to the NGF surface through atypical curvature of the backbone, contrary to Str1 and Str2, which form a relatively flat structure distal to the binding interface.

(B) Bn17 sits in a hydrophobic pocket on the NGF surface, stacking face-to-face with Trp21 and edge-to-face with Tyr52. dG23 makes an H bond to the protein backbone at Trp21, while dG22 makes an H bond to the side chain.

(C) The uridine base of Bn-dU24 makes five H bonds with NGF, two mediated by presumed water molecules at the interface. dU11 makes an H bond to Arg59 while Bn-dU25 stacks with the methylene side chain.

(D) Bn-dU27 forms a zipper-like motif with Bn-dU10 which stabilizes the 3′ end of the SOMAmer. dU27 stacks with Phe12 on NGF and it also makes an H bond to the Arg69 side chain and contacts Trp76. Arg69 makes a second H bond to dU10.

(E) The deoxyribose rings of dU24, dU25, and dG26 make non-polar sugar contacts with Phe7.

(F) A salt bridge is present between the SOMAmer backbone at dU18 and Lys32. Lys32 is sterically constrained by the sugar rings of dU17, dA19, and dC21, consequently positioning the side chain to form the salt bridge. See also Figure S6 and Table S1.

SOMAmer backbone makes a second sharp reversal, allowing Str3 to form stem II, a triple-stranded interaction with stem I. The Str3 backbone makes irregular twists and turns, conforming remarkably well to the NGF surface and forming a topographically complex recognition domain that sharply contrasts with the flat plane exhibited by the stem I scaffold (Figure 5A). Str3 residues exhibit the lowest B factors in the structure (not shown), consistent with close packing of these residues between stem I and NGF.

NGF Binding Interface

The SOMAmer inserts into a trough at the NGF dimer interface with Str3 at the apex of the wedge. Binding of each SOMAmer subunit buries 1,072 Å^2 of solvent-accessible surface area (SAS) (calculated as $((SAS_{NGF} + SAS_{SOMAmer}) - SAS_{Complex})/2$). Bn-dU residues play a central role in the interface, accounting for hydrophobic interactions with 15 amino acids (Table S1). Several of the natural bases in Str3 also contact NGF, resulting in almost continuous hydrophobic contact between Str3 and the protein. Of note, Bn17 sits in a hydrophobic furrow on the protein surface, stacking face-to-face with Trp21 and edge-to-face with Tyr52 (Figure 5B). The Bn-dU17 linker contacts Phe101, dU17 contacts Ile31, and the deoxyribose ring contacts the methylene side chain of Lys32 (Figure S6B). Although technically deriving from Str2, Bn-dU17 does not stack with Str2 bases in stem I; instead, it projects away from stem I, placing it in close proximity to the protein (Figures 2B and 5B). The extrusion of Bn-dU17 from stem I is undoubtedly facilitated by intercalation of Bn3 between dU16 and dU18, as part of the Bn-dU intercalating zipper motif (Figure 3B).

Three other modified nucleotides contribute to the NGF binding interface. Bn24 and Bn25, which contact each other edgewise, form a hydrophobic face that complements the hydrophobic patch on the protein created by Val18, Val20, and the methylene side chains of Ser19 and Arg59 (Figure 5C). The deoxyribose rings of dU24 and dU25 make non-polar sugar contacts with Phe7 (Figure 5E). Bn-dU27 exhibits dual functions in the SOMAmer with the benzyl moiety stabilizing the 3′ end as part of the hydrophobic cluster and the uridine residue contributing to the protein binding interface; to that end, dU27 stacks

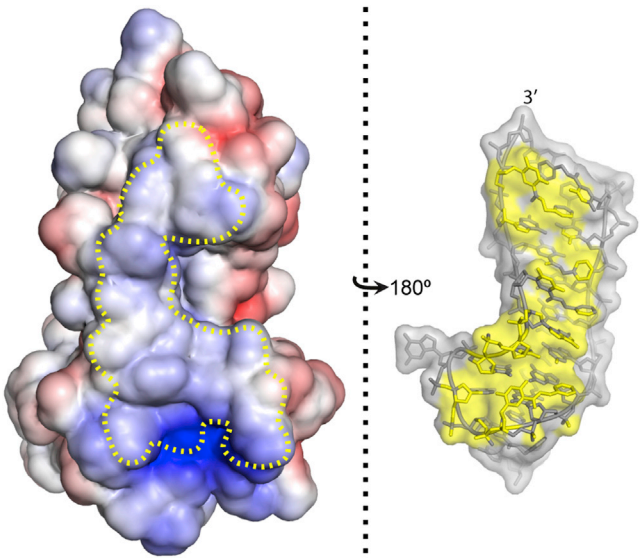


Figure 6. Open-Book View of the NGF-SOMAmer Binding Interface
An electrostatic surface potential rendering of NGF (±5 kT/e) shows the slightly basic composition of the surface. The SOMAmer interface area is outlined in yellow on NGF while atoms on the SOMAmer within 4 Å of the protein are colored yellow. The electrostatic surface rendering was prepared in PyMOL (Delano, 2002) using the APBS plugin (Baker et al., 2001).

with Phe12 and contacts Trp76 and the methylene sidechain of Arg69 (Figure 5D), while the deoxyribose sugar of dU27 contacts Gly10 and the backbone at Arg114 (Figure S6C). dG26 makes non-polar contact with Phe7 and Phe12 (Figures 5D and 5E).

A single salt bridge is observed between the SOMAmer backbone (at dU18) and Lys32 (Figure 5F). The orientation of Lys32 is constrained by the sugar rings of dU17, dA19, and dC21, which together form a hydrophobic channel that cradles the methylene side chain of the lysine, assisting to position the primary amine to form the salt bridge with the SOMAmer backbone. Other polar interactions are limited, with only seven observed H bonds between the SOMAmer and NGF (Table S1; Figure 5C). Although the NGF binding surface has a slightly basic composition, the few H bonds appear to be strategically located to support major hydrophobic interactions mediated by the aromatic moieties of the Bn-dU nucleotides (Figure 6). Previously identified RNA aptamers (Binkley et al., 1995; Nakamura et al., 2013) compete with SL1049 for binding NGF, indicating, at a minimum, partial overlap in binding sites (data not shown). This result is not surprising given the protein and ligands are of a similar size, and no general conclusions can be made regarding any similarities at the nucleotide-residue interface level.

SOMAmer Binding Domain Overlaps Receptor Binding Epitopes

Binding of NGF to the NGF-specific TrkA receptor was structurally characterized first with TrkA-d5 (Wiesmann et al., 1999) and subsequently confirmed with a co-crystal structure of the entire TrkA extracellular domain bound to NGF (Wehrman et al., 2007). TrkA binds in the saddle formed along the twisted β-sheet core of the NGF dimer. Binding of each TrkA molecule buries 1,120 Å^2 of SAS, which is remarkably similar to SOMAmer binding (Figures 7B and 7C). The NGF residues involved in TrkA binding (Figure 7A) include a 10 amino acid TrkA specificity region near the N terminus and a 23 amino acid neurotrophin common patch. The overall conformation of NGF is similar between the SOMAmer-bound and TrkA-bound structures (RMS = 0.709, 166 atoms), with the exception of the N-terminal α-helix, which interacts with both the SOMAmer and TrkA but orients in opposite directions in the structures (Figure S7A). Binding of the SOMAmer directly overlaps 48% of the TrkA binding residues; SOMAmer nucleotides Bn-dU17 to Bn-dU25 extensively occlude the N-terminal portion of the neurotrophin common binding patch, and the SOMAmer region encompassing nucleotides Bn-dU24 to Bn-dU27 occludes parts of the specificity patch. Binding of the SOMAmer therefore sterically hinders TrkA binding (Figures 7B and 7C), consistent with the observation that the SOMAmer blocks NGF-induced TrkA autophosphorylation and neurofilament outgrowth. The aromatic rings of the Bn-dU-modified nucleotides sometimes appear to mirror specific receptor-ligand interactions. For example, TrkA Pro382 stacks face-to-face with NGF Trp21 and edge-to-face with NGF Tyr52 in much the same way that the SOMAmer Bn17 interacts with those residues (Figure 7E).

The co-crystal structure of NGF bound to p75^{NTR} reveals that the receptor binds at the NGF dimer interface, burying 2,300 Å^2 of solvent-accessible surface and contacting NGF primarily in two hydrophobic regions designated site I and site II (He and Garcia, 2004). The global conformation of NGF is highly similar in the SOMAmer-bound and p75^{NTR}-bound structures (RMS = 0.536, 145 atoms) (Figure S7B). SOMAmer binding occludes half (13/26) of the NGF residues involved in p75^{NTR} binding (Figure 7A). Given the binding footprints (Figures 7B and 7D) and affinity of the SOMAmer, it is highly likely that SOMAmer binding would block p75^{NTR} binding to NGF. As seen with TrkA, we observe cases of receptor mimicry by the SOMAmer, where Bn25 occupies the same position as Tyr83 from p75^{NTR} (Figure 7F).

DISCUSSION

As one of only three structurally elucidated SOMAmers, the NGF SOMAmer has greatly advanced our understanding of the structural vocabulary of these novel ligands. The NGF SOMAmer shares a number of similarities with the platelet-derived growth factor (PDGF)-B (Davies et al., 2012) and interleukin-6 (IL-6) SOMAmers (Gelinas et al., 2014). All three are similar in length (24–32 nucleotides), the number of modified nucleotides (8–10), and the size of the protein binding interface (1,097–1,248 Å^2), which may reflect underlying similarities in starting libraries (all Bn-dU modified DNA libraries with 40N random region) and selection conditions. Compared with traditional aptamers, which rely heavily on polar interactions, these SOMAmers all utilize remarkably few H bonds and charge-charge interactions (Davies et al., 2012), instead achieving high target binding affinity via a primarily hydrophobic protein binding interface. All three SOMAmers exhibit extensive receptor mimicry, resulting in potent inhibition of cytokine-stimulated receptor activation. Despite these global similarities, the NGF SOMAmer exhibits several unique features. For example, both the PDGF and IL-6 SOMAmers exhibit two distinct domains composed of

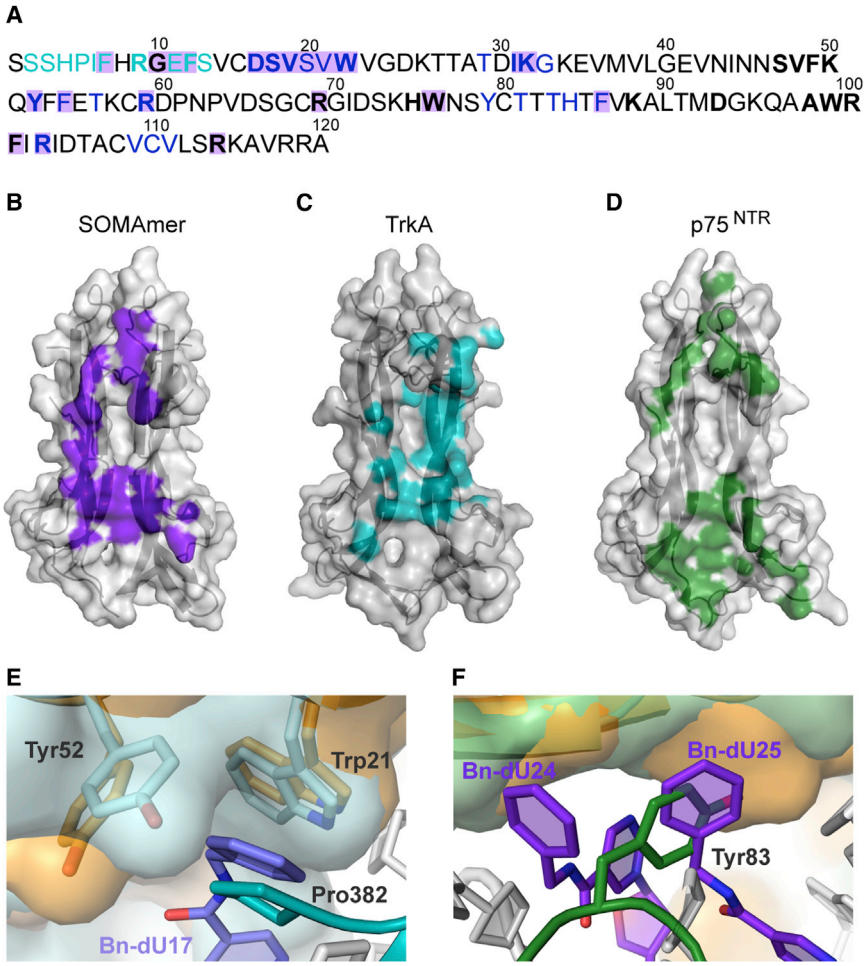


Figure 7. SOMAmer Binding Occludes TrkA and p75^{NTR} Binding Domains on NGF

(A–D) The sequence of the mature form of human NGF (UniProt: P01138) is shown (A), with residues involved in receptor interactions indicated: TrkA specificity patch (teal text), TrkA common patch (blue text), and p75^{NTR} (bold text) (He and Garcia, 2004; Wehrman et al., 2007; Wiesmann et al., 1999). Residues involved in SOMAmer binding are shaded (purple). Comparison of residues on NGF within 4 Å of the SOMAmer (B), purple, TrkA (PDB: 1WWW) (Wiesmann et al., 1999) (C), teal, or p75 (PDB: 1SG1) (He and Garcia, 2004) (D), green.

(E) Superposition of the NGF from the SOMAmer structure (gold) and PDB: 1WWW (light teal) shows TrkA residue Pro382 (teal) and SOMAmer Bn-dU17 (light purple) occupying the same position and interacting with NGF residues Trp21 and Tyr52.

(F) Superposition of NGF from the SOMAmer structure (gold) and PDB: 1SG1 (light green) shows p75^{NTR} residue Tyr83 (green) and SOMAmer-modified nucleotides Bn-dU24 and Bn-dU25 (purple) occupying similar positions at the binding interface. See also Figure S7.

known nucleic acid structural elements such as stem loops, pseudoknots, and G-quartets. In contrast, the NGF SOMAmer assumes a single compact S-shaped structure containing an intercalating zipper at one end and a Bn-dU cluster at the other end, flanking a non-helical triple-stranded core. Whereas the PDGF and IL-6 SOMAmers rely primarily on canonical base pairing, the NGF SOMAmer makes extensive use of Hoogsteen and sugar edge interactions, engaging virtually every available H-bond donor and acceptor in a network of base triples and quadruples. In the PDGF and IL-6 SOMAmers, the modified nucleotide side chains reside almost exclusively at the protein binding interface; conversely, the modified nucleotides play a more prominent role in the internal structure of the NGF SOMAmer with fewer side chains contributing directly to the protein binding interface. The unusual conformation of the NGF SOMAmer positions the extra aromatic rings of the modified nucleotides to maximize intramolecular π-stacking, thus stabilizing an entirely novel configuration of non-canonical base pairs and uncommon glycosidic bond angles.

We have identified two unique intramolecular structural motifs formed by Bn-dU residues: the intercalating zipper and the cluster. Upon closer inspection, the zipper and the cluster can be viewed as variants of the same structural motif, each consisting of three Bn-dU residues that exhibit certain hallmark characteristics. These include (1) reciprocal stacking of benzyl and uridyl rings from two Bn-dU residues arising from opposite strands, forming the teeth of the zipper, and (2) an adjacent Bn-dU adopting an L-shape that facilitates face-to-face stacking of its uridine with the teeth of the zipper and edge-to-face stacking of its benzyl group with one of the benzyl teeth. Despite the underlying similarity between the two motifs, the differences between the motifs are equally notable. Whereas the zipper represents an entirely internal structural element that stabilizes the nucleic acid scaffold, the cluster contributes significantly (albeit indirectly) to the protein interface in addition to its structural role. The aromatic rings in the cluster participate in a more extensive network of intermolecular and intramolecular π-stacking interactions than seen in the zipper motif. These differences exemplify the exceptional range of conformations that are accessible as a result of the rotational freedom afforded by the Bn-dU linker. In view of these observations, it is not surprising that all of the modified nucleotides, with the exception of Bn-dU18, contribute substantially to NGF binding affinity; single substitution of deoxythymidine (dT) for Bn-dU at any of these positions results in marked loss of affinity.

The triple-stranded SOMAmer core bears little resemblance to any of the three major subtypes of DNA triple helices that have been described. In typical DNA triple-helical structures, a WC duplex containing an oligopurine strand (R strand) and an oligopyrimidine strand (Y strand) pairs with a third strand via the Hoogsteen face of the R strand in the major groove of the duplex (Duca et al., 2008). Only one base triple in the SOMAmer fits the pattern of an antiparallel GT triplex motif: WC base pairing between dG15 and dC21 and a reverse Hoogsteen base pair between dG15 and dG4. Beyond that, any similarity to known

triplex motifs breaks down since the SOMAmer lacks an extended WC duplex or any semblance of helicity. In traditional DNA triple helices, the backbones of the R and third strands are in close proximity, bringing negatively charged atoms less than 6 Å apart. Unfavorable charge repulsion thus destabilizes the triplex structure and supra-physiological levels of divalent or polycationic agents are necessary to stabilize the triplex (Duca et al., 2008). Although the SOMAmer backbone comes in close proximity at one point in stem I, with ~6 Å between the non-bridging oxygens of dA5 and dG14, in general the non-helical triangular prism configuration of the SOMAmer maintains a greater separation of charge than typically seen in triple-helical DNA. Thus, the unusual base triples at the central core are likely a contributing factor to the observed absence of helicity in this structure. This feature is also consistent with the observation that the SOMAmer does not require divalent cations for NGF binding. These unique motifs almost certainly contribute thermodynamic stability to the structure, driving the SOMAmer to adopt unconventional base conformations and non-canonical base pairing in order to fold into a maximally stacked and highly compact structure.

Pi stacking interactions reminiscent of the SOMAmer zipper motif are also observed in traditional DNA/RNA aptamer protein structures (Cheung et al., 2013; Convery et al., 1998; Huang et al., 2003, 2009; Nomura et al., 2010; Russo Krauss et al., 2011), where an extruded base interacts with a side chain on the protein. Unlike SOMAmers, the expelled nucleotides are disengaged from the primary DNA/RNA structure, exhibiting few if any intramolecular interactions, but like the SOMAmers, the extrusion appears to be necessary, not only for interactions with the target but to accommodate tight turns or other aspects of the folded structure. In addition to π-stacking, the extruded nucleotides of the traditional aptamers are often positioned to participate in a number of H bonds with the protein, which are noticeably sparse in SOMAmer structures (Davies et al., 2012).

SOMAmers embody the intersection between evolutionary biology and synthetic chemistry. Combining the evolutionary power of SELEX with the enhanced shape and functional group diversity afforded by chemically modified nucleotide libraries has opened the door to a whole new class of molecules. SOMAmer-target co-crystal structures provide a startling glimpse into the rich and varied structures that emerge from chemically enhanced SELEX libraries. The insights afforded by the NGF structure provide inspiration for further chemical modifications aimed at refining the interface area. We have shown that many different C5-linked modifications of deoxyuridine are compatible with SELEX (Davies et al., 2012; Gelinas et al., 2014; Vaught et al., 2010) and each has the potential to impart unique characteristics. Different chemical modifications lead to intrinsic variations in SOMAmer specificity and affinity, as well as nuclease resistance (Gupta et al., 2014). The ability to fine-tune the physico-chemical properties of the SOMAmer can facilitate the selection of molecules with optimal characteristics for diagnostic and therapeutic applications.

EXPERIMENTAL PROCEDURES

See Supplemental Experimental Procedures for additional experimental details.

SOMAmer Discovery and Synthesis

NGF SOMAmers were discovered by using biotinylated recombinant human NGF in the SELEX with a kinetic challenge step, as described previously (Gold et al., 2010). The DNA library comprised 40 random positions with Bn-dU replacing dT in the random region. Selections were performed in SB17T buffer (40 mM HEPES [pH 7.5], 102 mM NaCl, 5 mM KCl, 5 mM $MgCl_2$, 1 mM EDTA, 0.05% TWEEN 20). Bn-dU SOMAmers were synthesized by solid phase synthesis using the phosphoramidite method (Beaucage and Caruthers, 1981), with modified deoxyuridine-5-carboxamide amidite reagents, as described previously (Davies et al., 2012; Gupta et al., 2014).

Measurement of Equilibrium Binding Constant

Equilibrium binding constants (K_d) of SOMAmers were measured in SB18T buffer (SB17 buffer with no EDTA, 0.01% TWEEN 20) at 37°C as described previously Gold et al., 2010). Briefly, heat-cooled radiolabeled SOMAmer was equilibrated for 17 hr (to ensure equilibrium was reached) with different concentrations of NGF and NGF-SOMAmer complexes, captured with ZORBAX PSM-300 resin, and quantified with a phosphorimager. The fraction of captured SOMAmer was plotted as a function of NGF concentration and the data were fit to a three-parameter sigmoid dose-response model to determine a K_d value.

Measurement of the Association Rate Constant (k_{on})

NGF at various concentrations (0.32–32 nM) was incubated with heat-cooled radiolabeled SOMAmer at 37°C. Aliquots were removed at various times between 0 and 1,050 s and complexes were isolated and quantified as described above. For each data point, the fraction of SOMAmer bound to NGF was normalized to the value at 1,050 s with 32 nM NGF. Relative fraction bound values were plotted as a function of time and the data were fit to a one-phase association model to determine the apparent association rate constant (k_{on}^{app}) for each NGF concentration. A plot of (k_{on}^{app}) as a function of NGF concentration showed a linear relationship with a slope equal to the association rate constant (k_{on}).

Measurement of the Dissociation Rate Constant (k_{off})

NGF (1×10^{-8} M) was incubated with heat-cooled radiolabeled SOMAmer at 37°C for 30 min. After adding unlabeled SOMAmer at the final concentration of 1×10^{-7} M, aliquots were removed at various times between 0 and 360 min. Complexes were isolated, quantified, and normalized to the value at t = 0 as described above. Relative fraction bound values were plotted as a function of time and the data were fit to a one-phase dissociation model to determine the k_{off} value.

Serum Stability Assay

Heat-cooled SOMAmer (500 nM) was incubated with 90% fresh human serum at 37°C. Aliquots were removed at various time points from 0 to 72 hr, extracted with phenol and chloroform, and analyzed by denaturing PAGE, as described previously (Gupta et al., 2014). DNA stained with SYBR gold was imaged with a fluorescent image analyzer (FUJI FLA-3000) and quantified using the ImageGauge software package to determine the fraction of intact SOMAmer at each time point.

Thermal Stability Analysis

Thermal melt studies were performed in SB18T buffer as described previously (Davies et al., 2012). Because baselines were not easily determined, T_m values were estimated using the maximum of first derivative method (Mergny and Lacroix, 2003). Derivatives were calculated as five-point tangential slopes at each temperature.

TrkA Phosphorylation Assay

PC12 cells (rat pheochromocytoma, ATCC) were seeded on 100-mm collagen IV coated plates (BD Bioscience). After overnight attachment, low serum medium (LSM) was added (0.2% FBS, 0.1% horse serum, 10 U/ml penicillin, 10 μg/ml streptomycin/RPMI-1640). The cells were left in LSM overnight and were then treated for 10 min with NGF alone (10 ng/ml, or 0.38 nM), NGF with TrkA phosphorylation inhibitor K252a (0.2 μM), and NGF pre-equilibrated with 10 nM SOMAmer. Cells were lysed in RIPA lysis buffer supplemented with PMSF, protease inhibitor cocktail, and sodium orthovanadate (Santa Cruz Biotechnology). The protein concentration in the lysates was measured by

DC protein assay (Bio-Rad). Aliquots of the lysates with equal amounts of protein were immuno-precipitated with an anti-Trk antibody C-14 (Santa Cruz Biotechnology) for 1 hr at 4°C. The precipitated proteins were purified using Protein A Sepharose Fast Flow (GE Healthcare). The immuno-precipitate was run on an SDS-PAGE gel, transferred to a polyvinylidene fluoride membrane, and probed with anti-phospho-tyrosine antibody 4G10 (Millipore) to quantify the amount of phosphorylated TrkA. The blot was stripped and probed with anti-Trk antibody C-14 to quantify the amount of total TrkA. Immuno-reactive bands were detected using an ECL Plus Western Blotting Detection System (GE Healthcare). Percent TrkA phosphorylation (ratio of phosphorylated TrkA/total TrkA) was normalized to a control obtained in the absence of added oligonucleotides.

Neurite Outgrowth Inhibition Assay

PC12 cells were seeded on 60-mm dishes at a density of 1.5×10^5 cells per dish. After 24 hr, the medium was replaced with LSM to induce differentiation. NGF (100 ng/ml or 3.8 nM) was mixed with SOMAmer (100 nM) for 1 hr, then added to the plates to final concentrations of 10 ng/ml NGF (0.38 nM) and 10 nM SOMAmer. After 3 days, fresh medium with NGF-SOMAmer complexes was added. On day 5, images of the cells were captured with a phase-contrast microscope and neurite length was measured using the NeuronJ plugin for ImageJ (NIH program). Neurite length/cell was calculated and normalized to a value of 100 observed in the absence of added oligonucleotides.

Crystallization and Heavy Atom Soaking

The complex of SL1049 and recombinant human NGF (Creative Biomart, CHO cell expression) was crystallized at 16°C by combining 0.4 μl of protein-SOMAmer complex (0.33 mM in PBS buffer) with 0.4 μl of crystallization solution (14–16 % PEG 8000, 100 mM HEPES [pH 6.0], 100 mM magnesium acetate, and 20% glycerol) in a sitting drop vapor diffusion tray (Compact 300 96-well crystallization plates, Rigaku Reagents). Sodium iodide was used for the heavy atom soak for phasing.

Data Collection and Structure Determination

A 2.45-Å native dataset was collected at the Canadian Light Source synchrotron. The NGF structure could be solved by molecular replacement but the solution did not provide enough phasing power to build the SOMAmer in the resulting electron density. Therefore, the technique of single anomalous dispersion (SAD) was used from a crystal soaked in an iodide salt (Abendroth et al., 2011; Dauter and Dauter, 2007; Dauter et al., 2000). This resulting low-resolution dataset (3.15 Å) was solved using the NGF portion of PDB: 1WWW. Four iodide sites were found and used for SAD phasing and combined with phases from the molecular replacement solution using the program PHASER in the CCP4 software suite in "experimental phasing" mode (McCoy et al., 2007; Winn et al., 2011).

The electron density from the phased data was modified using the program Parrot. The resulting maps showed features consistent with nucleic acid but identifying sequence position or even chain polarity was difficult. As an intermediate step, 16 phosphate ions were modeled into the strongest peaks of the putative DNA density. The expanded partial model was then used for another round of SAD with model phasing. The map quality improved sufficiently to allow modeling of four nucleotide residues in each subunit. Another round of phasing resulted in incrementally improved maps again. The expanded model (NGF plus eight nucleotide residues plus dummy phosphates) was then carried into the 2.45-Å native dataset. The iodide atoms were removed from the model and restrained refinement using Refmac (Vagin et al., 2004) was carried out. The resulting data were subjected to density modification using Parrot. The modified maps were a significant improvement over the low-resolution maps and model building was completed by bootstrapping using maps generated from the 2.45-Å data. All nucleotide residues in the final model were manually placed in the electron density maps and the final maps were of sufficient quality to allow unambiguous placement of the nucleoside bases and modifications.

ACCESSION NUMBERS

The accession number for the NGF-SOMAmer co-crystal structure reported in this paper is PDB: 4ZBN.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.str.2015.03.027.

AUTHOR CONTRIBUTIONS

T.C.J., D.R.D, D.J.S., A.H., D.I.R., S.G., S.M.W., H. H., S.K.W., A.B.B., L.J.S., D.J.S. and N.J. designed the research; T.C.J., D.R.D., A.H., D.I.R., S.M.W., A.N., T.W., T.O., B.G., C.Z. W.S.M., and T.E.E. the performed research; T.C.J., D.R.D., A.H., D.I.R., S.G., S.M.W., H.H., Y.N., A.D.G., D.J.S., and N.J. analyzed the data; T.C.J, D.R.D, A.D.G, A.H., S.G., D.J.S., and N.J. wrote the paper.

ACKNOWLEDGMENTS

We thank Jan Abendroth of Beryllium for assistance with data collection and software advice for phasing. We also thank Drs. Toshihide Ono and Jun Shimada of Otsuka Pharmaceutical Co., for helpful discussion. SOMAmer is a registered trademark of SomaLogic, Inc. T.C.J., D.I.R., S.G., S.M.W., B.G., S.K.W., W.S.M., A.D.G, D.J.S., and N.J. are employees and/or shareholders of SomaLogic, Inc. D.R.D., T.E.E., A.B.B., and L.J.S. are employees and/or shareholders of Beryllium. A.H., A.N., T.W., T.O., H.H., and Y.N. are employees and/or shareholders of Otsuka Pharmaceutical Co., Ltd.

Received: December 23, 2014
Revised: March 24, 2015
Accepted: March 25, 2015
Published: May 28, 2015

REFERENCES

Abendroth, J., Gardberg, A.S., Robinson, J.I., Christensen, J.S., Staker, B.L., Myler, P.J., Stewart, L.J., and Edwards, T.E. (2011). SAD phasing using iodide ions in a high-throughput structural genomics environment. J. Struct. Funct. Genomics *12*, 83–95.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. USA *98*, 10037–10041.

Beaucage, S., and Caruthers, M. (1981). Deoxynucleoside phosphoramidites - a new class of key intermediates for deoxypolynucleotide synthesis. Tetrahedron Lett. *22*, 1859–1862.

Binkley, J., Allen, P., Brown, D.M., Green, L., Tuerk, C., and Gold, L. (1995). RNA ligands to human nerve growth factor. Nucleic Acids Res. *23*, 3198–3205.

Bothwell, M.A., and Shooter, E.M. (1977). Dissociation equilibrium constant of beta nerve growth factor. J. Biol. Chem. *252*, 8532–8536.

Ceni, C., Kommaddi, R.P., Thomas, R., Vereker, E., Liu, X., McPherson, P.S., Ritter, B., and Barker, P.A. (2010). The p75NTR intracellular domain generated by neurotrophin-induced receptor cleavage potentiates Trk signaling. J. Cell Sci. *123*, 2299–2307.

Cheung, Y.W., Kwok, J., Law, A.W., Watt, R.M., Kotaka, M., and Tanner, J.A. (2013). Structural basis for discriminatory recognition of *Plasmodium* lactate dehydrogenase by a DNA aptamer. Proc. Natl. Acad. Sci. USA *110*, 15967–15972.

Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Convery, M.A., Rowsell, S., Stonehouse, N.J., Ellington, A.D., Hirao, I., Murray, J.B., Peabody, D.S., Phillips, S.E., and Stockley, P.G. (1998). Crystal structure of an RNA aptamer-protein complex at 2.8 A resolution. Nat. Struct. Biol. *5*, 133–139.

Dauter, M., and Dauter, Z. (2007). Phase determination using halide ions. Methods Mol. Biol. *364*, 149–158.

Dauter, Z., Dauter, M., and Rajashankar, K.R. (2000). Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. Acta Crystallogr. D Biol. Crystallogr. *56*, 232–237.

Davies, D.R., Gelinas, A.D., Zhang, C., Rohloff, J.C., Carter, J.D., O'Connell, D., Waugh, S.M., Wolk, S.K., Mayfield, W.S., Burgin, A.B., et al. (2012). Unique motifs and hydrophobic interactions shape the binding of modified DNA ligands to protein targets. Proc. Natl. Acad. Sci. USA *109*, 19971–19976.

De Groote, M.A., Nahid, P., Jarlsberg, L., Johnson, J.L., Weiner, M., Muzanyi, G., Janjic, N., Sterling, D.G., and Ochsner, U.A. (2013). Elucidating novel serum biomarkers associated with pulmonary tuberculosis treatment. PLoS One *8*, e61002.

Delano, W.L. (2002). The PyMOL Molecular Graphics System (Delano Scientific, LLC).

Duca, M., Vekhoff, P., Oussedik, K., Halby, L., and Arimondo, P.B. (2008). The triple helix: 50 years later, the outcome. Nucleic Acids Res. *36*, 5123–5138.

Eibl, J.K., Strasser, B.C., and Ross, G.M. (2012). Structural, biological, and pharmacological strategies for the inhibition of nerve growth factor. Neurochem. Int. *61*, 1266–1275.

Gelinas, A.D., Davies, D.R., Edwards, T.E., Rohloff, J.C., Carter, J.D., Zhang, C., Gupta, S., Ishikawa, Y., Hirota, M., Nakaishi, Y., et al. (2014). Crystal structure of interleukin-6 in complex with a modified nucleic acid ligand. J. Biol. Chem. *289*, 8720–8734.

Ghilardi, J.R., Freeman, K.T., Jimenez-Andrade, J.M., Mantyh, W.G., Bloom, A.P., Bouhana, K.S., Trollinger, D., Winkler, J., Lee, P., Andrews, S.W., et al. (2011). Sustained blockade of neurotrophin receptors TrkA, TrkB and TrkC reduces non-malignant skeletal pain but not the maintenance of sensory and sympathetic nerve fibers. Bone *48*, 389–398.

Gold, L., Ayers, D., Bertino, J., Bock, C., Bock, A., Brody, E.N., Carter, J., Dalby, A.B., Eaton, B.E., Fitzwater, T., et al. (2010). Aptamer-based multiplexed proteomic technology for biomarker discovery. PLoS One *5*, e15004.

Gold, L., Walker, J.J., Wilcox, S.K., and Williams, S. (2012). Advances in human proteomics at high scale with the SOMAscan proteomics platform. N. Biotechnol. *29*, 543–549.

Gupta, S., Thirstrup, D., Jarvis, T.C., Schneider, D.J., Wilcox, S.K., Carter, J., Zhang, C., Gelinas, A., Weiss, A., Janjic, N., et al. (2011). Rapid histochemistry using slow off-rate modified aptamers with anionic competition. Appl. Immunohistochem. Mol. Morphol. *19*, 273–278.

Gupta, S., Hirota, M., Waugh, S.M., Murakami, I., Suzuki, T., Muraguchi, M., Shibamori, M., Ishikawa, Y., Jarvis, T.C., Carter, J.D., et al. (2014). Chemically modified DNA aptamers bind interleukin-6 with high affinity and inhibit signaling by blocking its interaction with interleukin-6 receptor. J. Biol. Chem. *289*, 8706–8719.

He, X.L., and Garcia, K.C. (2004). Structure of nerve growth factor complexed with the shared neurotrophin receptor p75. Science *304*, 870–875.

Hefti, F.F., Rosenthal, A., Walicke, P.A., Wyatt, S., Vergara, G., Shelton, D.L., and Davies, A.M. (2006). Novel class of pain drugs based on antagonism of NGF. Trends Pharmacol. Sci. *27*, 85–91.

Holland, D.R., Cousens, L.S., Meng, W., and Matthews, B.W. (1994). Nerve growth factor in different crystal forms displays structural flexibility and reveals zinc binding sites. J. Mol. Biol. *239*, 385–400.

Huang, D.B., Vu, D., Cassiday, L.A., Zimmerman, J.M., Maher, L.J., 3rd, and Ghosh, G. (2003). Crystal structure of NF-kappaB (p50)2 complexed to a high-affinity RNA aptamer. Proc. Natl. Acad. Sci. USA *100*, 9268–9273.

Huang, R.H., Fremont, D.H., Diener, J.L., Schaub, R.G., and Sadler, J.E. (2009). A structural explanation for the antithrombotic activity of ARC1172, a DNA aptamer that binds von Willebrand factor domain A1. Structure *17*, 1476–1484.

Leontis, N.B., and Westhof, E. (2001). Geometric nomenclature and classification of RNA base pairs. RNA *7*, 499–512.

Loffredo, F.S., Steinhauser, M.L., Jay, S.M., Gannon, J., Pancoast, J.R., Yalamanchi, P., Sinha, M., Dall'Osso, C., Khong, D., Shadrach, J.L., et al. (2013). Growth differentiation factor 11 is a circulating factor that reverses age-related cardiac hypertrophy. Cell *153*, 828–839.

Lollo, B., Steele, F., and Gold, L. (2014). Beyond antibodies: new affinity reagents to unlock the proteome. Proteomics *14*, 638–644.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

McDonald, N.Q., Lapatto, R., Murray-Rust, J., Gunning, J., Wlodawer, A., and Blundell, T.L. (1991). New protein fold revealed by a 2.3-A resolution crystal structure of nerve growth factor. Nature *354*, 411–414.

McKelvey, L., Shorten, G.D., and O'Keeffe, G.W. (2013). Nerve growth factor-mediated regulation of pain signalling and proposed new intervention strategies in clinical pain management. J. Neurochem. *124*, 276–289.

Mehan, M.R., Ayers, D., Thirstrup, D., Xiong, W., Ostroff, R.M., Brody, E.N., Walker, J.J., Gold, L., Jarvis, T.C., Janjic, N., et al. (2012). Protein signature of lung cancer tissues. PLoS One *7*, e35157.

Mehan, M.R., Ostroff, R., Wilcox, S.K., Steele, F., Schneider, D., Jarvis, T.C., Baird, G.S., Gold, L., and Janjic, N. (2013). Highly multiplexed proteomic platform for biomarker discovery, diagnostics, and therapeutics. Adv. Exp. Med. Biol. *735*, 283–300.

Mergny, J.L., and Lacroix, L. (2003). Analysis of thermal melting curves. Oligonucleotides *13*, 515–537.

Nakamura, Y., Jin, L., and Hiramatsu, H. (2013). Aptamer for NGF and Use Thereof (Shionogi & Co., Ltd., Ribomic Inc.).

Nomura, Y., Sugiyama, S., Sakamoto, T., Miyakawa, S., Adachi, H., Takano, K., Murakami, S., Inoue, T., Mori, Y., Nakamura, Y., et al. (2010). Conformational plasticity of RNA for target recognition as revealed by the 2.15 A crystal structure of a human IgG-aptamer complex. Nucleic Acids Res. *38*, 7822–7829.

Ochsner, U.A., Katilius, E., and Janjic, N. (2013). Detection of *Clostridium difficile* toxins A, B and binary toxin with slow off-rate modified aptamers. Diagn. Microbiol. Infect. Dis. *76*, 278–285.

Ochsner, U.A., Green, L.S., Gold, L., and Janjic, N. (2014). Systematic selection of modified aptamer pairs for diagnostic sandwich assays. BioTechniques *56*, 125–128, 130, 132–123.

Ostroff, R.M., Bigbee, W.L., Franklin, W., Gold, L., Mehan, M., Miller, Y.E., Pass, H.I., Rom, W.N., Siegfried, J.M., Stewart, A., et al. (2010). Unlocking biomarker discovery: large scale application of aptamer proteomic technology for early detection of lung cancer. PLoS One *5*, e15003.

Ostroff, R.M., Mehan, M.R., Stewart, A., Ayers, D., Brody, E.N., Williams, S.A., Levin, S., Black, B., Harbut, M., Carbone, M., et al. (2012). Early detection of malignant pleural mesothelioma in asbestos-exposed individuals with a noninvasive proteomics-based surveillance tool. PLoS One *7*, e46091.

Russo Krauss, I., Merlino, A., Giancola, C., Randazzo, A., Mazzarella, L., and Sica, F. (2011). Thrombin-aptamer recognition: a revealed ambiguity. Nucleic Acids Res. *39*, 7858–7867.

Tapley, P., Lamballe, F., and Barbacid, M. (1992). K252a is a selective inhibitor of the tyrosine protein kinase activity of the trk family of oncogenes and neurotrophin receptors. Oncogene *7*, 371–381.

Vagin, A.A., Steiner, R.A., Lebedev, A.A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G.N. (2004). REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. Acta Crystallogr. D Biol. Crystallogr. *60*, 2184–2195.

Vaught, J.D., Bock, C., Carter, J., Fitzwater, T., Otis, M., Schneider, D., Rolando, J., Waugh, S., Wilcox, S.K., and Eaton, B.E. (2010). Expanding the chemistry of DNA for in vitro selection. J. Am. Chem. Soc. *132*, 4141–4151.

Wehrman, T., He, X., Raab, B., Dukipatti, A., Blau, H., and Garcia, K.C. (2007). Structural and mechanistic insights into nerve growth factor interactions with the TrkA and p75 receptors. Neuron *53*, 25–38.

Wiesmann, C., Ultsch, M.H., Bass, S.H., and de Vos, A.M. (1999). Crystal structure of nerve growth factor in complex with the ligand-binding domain of the TrkA receptor. Nature *401*, 184–188.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. *67*, 235–242.

