Supporting Information for:

Elucidating Glycosaminoglycan-Protein-Protein Interactions using Carbohydrate Microarray and Computational Approaches

Claude J. Rogers, Peter M. Clark, Sarah E. Tully, Ravinder Abrol, K. Christopher Garcia, William A. Goddard III, and Linda C. Hsieh-Wilson

This file includes:

Supplementary Methods and References

Supplementary Figure Legends

Supplementary Figures

Supplementary Tables

SI Methods

Microarray analysis. Tetrasaccharide and polysaccharide microarrays were prepared as described previously (1, 2). Microarray experiments with the individual NTs were performed using methods described previously (3). To measure glycosaminoglycan-protein-protein interactions using the tetrasaccharide microarrays, a perimeter was drawn around 3 microarrays with a hydrophobic marker (PapPen) for each protein-receptor pair. The microarrays were treated with $NaBH_4$ (5 min, 66 mM in PBS) and washed five times with PBS. The first microarray was incubated with the protein ligand of interest (TNF-α, NGF, BDNF, NT-3, or NT-4/5; R&D Systems, 1 μM in 0.1% Triton X-100 in PBS). The second microarray was incubated with the receptor of interest (TNFR1-Fc, TrkA-Fc, TrkB-Fc, TrkC-Fc, or TrkB-Fc, respectively; R&D Systems, 1 μM in 0.1% Triton X-100 in PBS). The third microarray was incubated with a 1:1 mixture of both the protein ligand and receptor (1 μM each in 0.1% Triton X-100 in PBS). After 3 h, the microarrays were washed (5 × PBS) and treated with a rabbit primary antibody against the protein ligand (R&D Systems, 1:1000 in 0.1% Triton X-100 in PBS). After 1 h, the microarray was washed (5 × PBS) and incubated with Cy3-conjugated anti-human Fc IgG and Cy5-conjugated anti-rabbit IgG antibodies (Invitrogen, 1:5000 in 0.1% Triton X-100 in PBS) for 1 h. The microarrays were then washed (3 × PBS, 2 × H_2O) and dried under a stream of air. A similar procedure was employed for the polysaccharide microarrays. After drawing perimeters around the array regions, the microarrays were incubated in 10% fetal bovine serum (FBS) in PBS for 1 h at 37 ºC. The remainder of the procedure was identical, except that 1% FBS in PBS was used as the incubating buffer instead of 0.1% Triton X-100 in PBS. All microarrays were analyzed using a GenePix

5000a scanner, and fluorescence quantification was performed using GenePix 6.1 software with correction for local background. Each protein was analyzed in triplicate, and the data represent an average of 8–10 spots for a given carbohydrate concentration. As controls, the microarrays were treated with the protein ligand (NT, TNFα or FGF2), followed by the anti-ligand antibody and Cy3 anti-human Fc IgG. Similarly, the microarrays were incubated with the receptor (Trk-Fc, TNFR1-Fc or FGFR1-Fc), followed by the corresponding anti-ligand antibody (rabbit species) and Cy5 anti-rabbit IgG. In all cases, no signal was observed, confirming that the antibodies showed no cross-reactivity (Fig. S11).

Computational Methods. *Structures and Homology Models*. For the computational studies, the protein coordinates were obtained from solution or crystal structures when available. Homology models were generated for protein complexes where such data was absent. In all cases, crystal structures of highly related protein complexes were available for construction of the homology models. The following PDB files were used: DBL3x (3BQK), DBL6 (2WAU), FGF2 solution structure (1BLA), FGF2-heparin co-crystal structure (1BFB), FGFR1 (1FQ9, chain C), FGF2-FGFR1 quaternary complex (1FQ9, chains A, B, C, D), FGF2-FGFR1-heparin co-crystal structure (1FQ9), TNF-α trimer (1TNF), TNF-β-TNFR1 (1TNR), NT-4/5 dimer (1B98 and 1HCF, chains A, B), human NGF dimer (2IFG, chains E, F), mouse NGF dimer (1BET), NT-3 dimer (1NT3), BDNF monomer (1BND, chain A), TrkA ligand-binding domain (1WWW chain X), TrkB ligand-binding domain (1HCF, chain X), NGF-TrkA complex (1WWW).

The TNF-α-TNFR1 homology model was determined by aligning chain B of TNF-α in the TNF-α trimer crystal structure with TNF-β in the TNF-β-TNFR1 crystal

structure (TNF-α, TNF-β RMSD = 1.08 Å) and replacing TNF-β with the TNF-α trimer structure. The BDNF dimer homology model was constructed by replacing the NT-3 monomer with the BDNF monomer in the BDNF-NT-3 dimer structure from the PDB file 1BND (BDNF, NT-3 RMSD = 0.967 Å). The NT-4/5-TrkA homology model was created by replacing TrkB in the PDB file 1HCF with TrkA (TrkA, TrkB RMSD = 0.804 Å). The BDNF-TrkB homology model was created by replacing the NT-4/5 dimer in the PDB file 1HCF with the BDNF dimer homology model (NT-4/5 dimer, BDNF dimer RMSD = 0.757 Å). The NT-3-TrkA homology model was created by replacing the NT-4/5 dimer and TrkB in the PDB file 1HCF with the NT-3 dimer and TrkA, respectively (NT-4/5 dimer, NT-3 dimer RMSD = 1.042 Å).

The DREIDING FF (4) was used throughout the modeling. Protein files were downloaded from the RCSB Protein Data Bank (www.pdb.org) and loaded into the Swiss PDB Viewer (5) to fix incomplete side chains. The WhatIF program (6) was used to add hydrogen atoms. CHARMM22 (7) charges were added, and the protein was fully minimized in the presence of sodium and chloride ions under conditions of Surface Generalized Born (SGB) continuum solvation (8).

The structures of a heparin tetrasaccharide and octasaccharide were extracted from the PDB files 1BFB and 1FQ9, respectively (heparin 'A'). Hydrogen atoms were added and charges were assigned to each atom using the charge equilibration (QEq) method (9). The ligands were then fully minimized under conditions of Surface Generalized Born (SGB) continuum solvation (8). The solution structures of the CS-A, CS-C, and CS-E tetrasaccharides were determined using molecular dynamics as reported (3).

Putative Binding Site Determination. Coarse binding sites were first determined by rigid body docking of one oligosaccharide conformation as described previously (10), except with the following modifications. The parameters *radmax* = 5.0 and *dotlim* = –0.5 were used for the autoMS program in Dock4.0 when creating the molecular surface. To determine the potential binding site, the twenty-five lowest energy docked structures and corresponding binding sites were tabulated and ranked by energy. Next, the sum of the inverse energy ranks for each binding site was determined. Any binding site with a value of 1 or greater was considered a potential glycosaminoglycan binding site.

Final Binding Site Determination. The potential glycosaminoglycan binding sites were inputted into GenMSCDock (11), and rigid body docking of one oligosaccaride conformation was performed with standard input parameters. Briefly, up to 120 different docked orientations were obtained in the first step, which represented the 40 lowest energy orientations each, as measured by Coulombic interaction energy, van der Waals interaction energy, and total interaction energy. Next, for each of these structures, residues within 4 Å of any of the bound oligosaccharides were rotated, the complexes were briefly minimized, and the energy was calculated. Finally, the universal cavity energy, which consists of the energy of the oligosaccharide and those protein residues within 5 Å of the oligosaccharide, was determined. Residues within 5 Å of the oligosaccharide in more than one of the five minimum energy structures were considered part of the glycosaminoglycan binding site. The average binding energies of the CS-A, CS-C, and CS-E tetrasaccharides were determined by averaging the differences between the energy of the docked CS-protein structure and the combined energy of the individual CS and protein structures for the five minimum energy structures.

Images were created using PyMOL (12), and the electrostatic maps were derived using Adaptive Poisson-Boltzmann Solver (APBS) software (13).

Cellular Assays. PC12 cells were propagated on collagen-coated 10-cm dishes in DMEM (Gibco) supplemented with 6.5% FBS, 6.5% horse serum, 2 mM L-glutamine, 100 $U \cdot ml^{-1}$ penicillin, and 100 $U \cdot ml^{-1}$ streptomycin. Cells were grown at 37 ºC in a humidified atmosphere enriched with 5% CO_2 and sub-cultured at a 1:3 ratio every 5–6 days. For the assays, PC12 cells were cultured on poly-DL-lysine-coated 60-mm dishes and grown as described above. After 4–5 days (70–80% confluence), the media was replaced with a minimal media composed of DMEM supplemented with 0.5% FBS, 1% horse serum, 2 mM L-glutamine, 100 $U \cdot ml^{-1}$ penicillin, and 100 $U \cdot ml^{-1}$ streptomycin. Cells were incubated for 12 h before use in the following experiments.

For chondroitinase experiments, 1 $U \cdot ml^{-1}$ chondroitinase ABC (Seikagaku) was added to the media. After 2 h, the cells were gently washed with fresh media three times before treatment. For experiments with adsorbed CS-E on the dish, cells were split 1:1 and plated on poly-DL-lysine dishes that had been incubated with a solution of CS-E-enriched polysaccharides (500 $ng \cdot ml^{-1}$in PBS; Seikagaku) for 12 h at 37 ºC and then washed (3 x PBS). The cells were allowed to adhere to the dish for 2–3 h prior to treatment. For experiments with exogenous CS-E in solution, cells were exposed to fresh media containing CS-E-enriched polysaccharides (500 $ng \cdot ml^{-1}$ or 5000 $ng \cdot ml^{-1}$) for 2 h prior to treatment. In all cases, cells treated with NTs were exposed to NGF (5 $ng \cdot ml^{-1}$) for 5 min, NT-4/5 (0.5 $\mu g \cdot ml^{-1}$) for 15 min, or a vehicle control. The cells were subsequently washed (3 × PBS), lysed, and analyzed as previously described (14). Western blots were analyzed by blotting for total TrkA (763; Santa Cruz Biotechnology)

and p-TrkA (Y490; Cell Signaling). Relative TrkA activation was calculated by dividing p-TrkA signal by total TrkA signal for each condition and normalizing with respect to the control (NT treatment of WT cells).

Immunohistochemistry. PC12 cells were grown on poly-DL-lysine-coated coverslips and treated with chondroitinase as described above. For the co-localization studies, cells were treated with NGF for 60 min. Following each treatment, cells were fixed with 4% paraformaldehyde in PBS for 20 min, washed (2 × PBS), and permeabilized with 0.5% Triton-X 100 for 5 min. After washing (2 × PBS), cells were blocked with 10% FBS in PBS for 1 h and then incubated with an anti-CS-E mouse antibody (3, 15) (1:500) and an anti-tubulin rat antibody (Sigma, 1:1000) in 10% FBS in PBS for 3 h. The cells were washed (5 × PBS), treated with anti-mouse-AlexaFluor488 (Invitrogen, 1:1000) and anti-rat-AlexaFluor546 (Invitrogen, 1:1000) secondary antibodies for 1 h, washed again (5 × PBS), mounted onto slides, and then imaged by confocal fluorescence microscopy. Images were analyzed using ImageJ (http://rsbweb.nih.gov/ij/) using the RG2B Co-localization plugin, which returns a pixel in the blue channel for each pixel with signal from both the red and green channel. The total number of pixels in the blue channel were counted for each cell and normalized with respect to area.

ELISA. TrkA (3.3 $\mu g \cdot ml^{-1}$ in 50 mM Na_2CO_3, pH 9.6; 25 µl per well) was added to a 384-well plate (Maxisorp) and incubated in a humidified chamber at 4 ºC for 12 h. The wells were blocked with 1% BSA in PBS at 25 °C for 2 h. Varying concentrations of NGF, NT-3, or NT-4/5 (0.45 – 230 nM in 0.1% BSA in PBS) were added to each well. After 2 h, the wells were incubated with a rabbit antibody against the NT of interest

(1:1000 in 0.1% BSA in PBS) at 25 °C for 1 h, followed by an anti-rabbit antibody conjugated to horseradish peroxidase (Invitrogen, 1:10,000 in 0.1% BSA in PBS) at 25 °C for 1 h. Between each incubation, the wells were washed three to five times in 0.05% Tween-20 in PBS. NT binding was detected using a TMB Substrate Kit (Pierce) according to the manufacturer's protocol.

References

1. Tully SE, Rawat M, & Hsieh-Wilson LC (2006) Discovery of a TNF-alpha antagonist using chondroitin sulfate microarrays. *J. Am. Chem. Soc.* **128,** 7740-7741.
2. Shipp EL & Hsieh-Wilson LC (2007) Profiling the sulfation specificities of glycosaminoglycan interactions with growth factors and chemotactic proteins using microarrays. *Chem. Biol.* **14,** 195-208.
3. Gama CI*, et al.* (2006) Sulfation patterns of glycosaminoglycans encode molecular recognition and activity. *Nat. Chem. Biol.* **2,** 467-473.
4. Mayo SL, Olafson BD, & Goddard WA, 3rd (1990) DREIDING: a generic force field for molecular simulations. *J. Phys. Chem.* **94,** 8897-8909.
5. Guex N & Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18,** 2714-2723.
6. Vriend G (1990) WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* **8,** 52-56, 29.
7. MacKerell Jr AD*, et al.* (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **102,** 3586-3616.
8. Ghosh A, Rapp CS, & Friesner RA (1998) Generalized Born model based on a surface integral formulation. *J. Phys. Chem. B* **102,** 10983-10990.
9. Rappe AK & Goddard WA, 3rd (1991) Charge equilibration for molecular dynamics simulations. *J. Phys. Chem.* **95,** 3358-3363.
10. Vaidehi N*, et al.* (2002) Prediction of structure and function of G protein-coupled receptors. *Proc. Natl. Acad. Sci. USA* **99,** 12622-12627.

11. Kim SK, Li Y, Park C, Abrol R, & Goddard WA, 3rd (2010) Prediction of the three-dimensional structure for the rat urotensin II receptor, and comparison of the antagonist binding sites and binding selectivity between human and rat receptors from atomistic simulations. *ChemMedChem* **5,** 1594-1608.

12. Schrodinger, LLC (2010).

13. Baker NA, Sept D, Joseph S, Holst MJ, & McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* **98,** 10037-10041.

14. Colangelo AM*, et al.* (2005) Recombinant human nerve growth factor with a marked activity *in vitro* and *in vivo*. *Proc. Natl. Acad. Sci. USA* **102,** 18658-18663.

15. Tully SE*, et al.* (2004) A chondroitin sulfate small molecule that stimulates neuronal growth. *J. Am. Chem. Soc.* **126,** 7736-7737.

Supporting Figure Legends

Figure S1. Structures of representative glycosaminoglycan family members. For heparin, heparan sulfate, dermatan sulfate, and keratan sulfate: R = H or SO_3^-, R' = H, Ac, or SO_3^-. n = 10-200.

Figure S2. Co-localization of FGF2 (red) and FGFR1 (green) on the polysaccharide microarray

Figure S3. (A) Relative binding of TNF-α to the indicated tetrasaccharide of defined sulfation sequence in the presence (red) or absence (black) of TNFR1. (B) Relative binding of TNFR1 to the indicated tetrasaccharide of defined sulfation sequence in the presence (red) or absence (black) of TNF-α. Binding relative to the maximum signal for each plot is shown. Each protein was analyzed in triplicate, and the data represent an average of 8–10 spots for a given carbohydrate concentration.

Figure S4. (A) Relative binding of TrkC to the indicated tetrasaccharides in the presence (red) or absence (black) of NT-3. (B) Relative binding of TrkB to the indicated tetrasaccharides in the presence (red) or absence (black) of NT-4/5. Binding relative to the maximum signal for each plot is shown. Each protein was analyzed in triplicate, and the data represent an average of 8–10 spots for a given carbohydrate concentration.

Figure S5. Binding of Trk receptors to the indicated polysaccharides at 0.5, 5, and 10 μM concentration in the presence (red) or absence (black) of the indicated NT.

Figure S6. Binding of the indicated NTs to immobilized TrkA, as determined by ELISA. NGF has the strongest affinity for TrkA. NT-3 and NT-4/5 also bind, but with >100-fold weaker affinity.

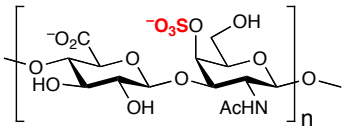
Figure S7. Comparison of predicted CS-E binding sites (yellow and green) on the BDNF, NGF, NT-4/5, and NT-3 dimer. Residues highlighted in green correspond to monomer A of the dimer; residues highlighted in yellow correspond to monomer B. NT-3 does not contain β-strands 4 and 5.

Figure S8. Predicted CS-E binding sites in the (A) BDNF, NGF, NT-4/5, and NT-3 dimers and (B) the NT binding domain (domain 5) of TrkA and TrkB. Top: Ribbon representation of each protein. Middle: Connolly surface with CS-E binding sites depicted in blue (non-basic residues) and yellow (basic residues). Bottom: Electrostatic maps generated by Adaptive Poisson-Boltzmann Solver software. All images were created in PyMOL.

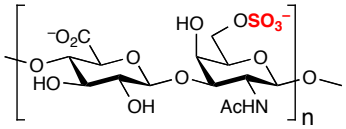
Figure S9. Predicted CS-E binding sites in the NGF-TrkA, BDNF-TrkB, NT-3-TrkA, and NT-4/5-TrkA complexes. Top: Ribbon representation of each protein. Middle: Connolly surface with CS-E binding sites depicted in blue (non-basic residues) and yellow (basic residues). Bottom: Electrostatic maps generated by Adaptive Poisson-Boltzmann Solver software. All images were created in PyMOL.

Figure S10. Treatment of PC12 cells with ChABC results in loss of CS-E-positive immunostaining.

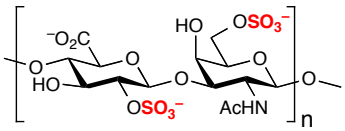
Figure S11. Microarray controls to confirm lack of antibody cross-reactivity. (A) NGF binding to the microarray was detected using an anti-NGF rabbit primary antibody and anti-rabbit IgG secondary antibody conjugated to Cy5. (B) No signal was observed with the anti-Fc antibody conjugated to Cy3 used to detect Trk-Fc receptors. (C) TrkA-Fc binding to the microarray was detected using an anti-Fc antibody conjugated to Cy3. (D) No signal was observed with the anti-rabbit antibody conjugated to Cy5 used to detect NGF.



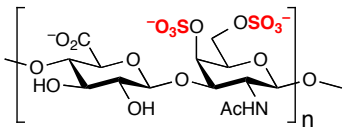
Chondroitin Sulfate-A



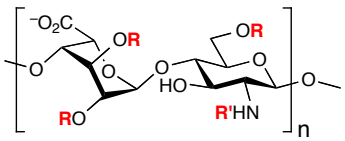
Chondroitin Sulfate-C



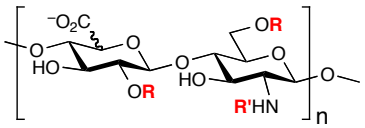
Chondroitin Sulfate-D



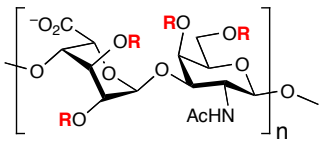
Chondroitin Sulfate-E



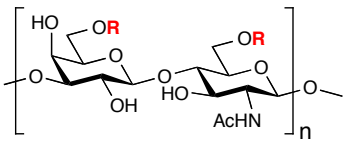
Heparin



Heparan Sulfate



Dermatan Sulfate



Keratan Sulfate

Figure S1

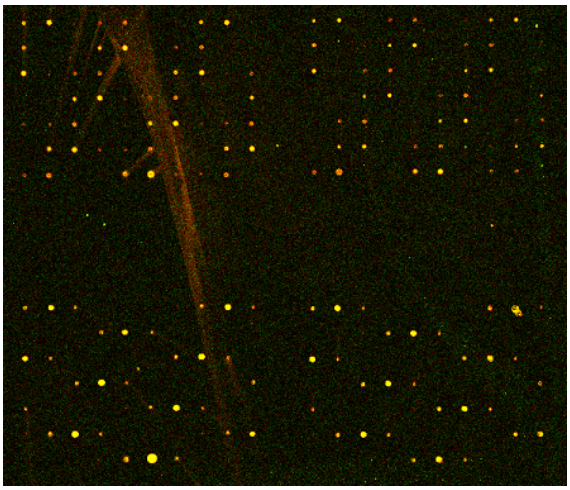


Figure S2

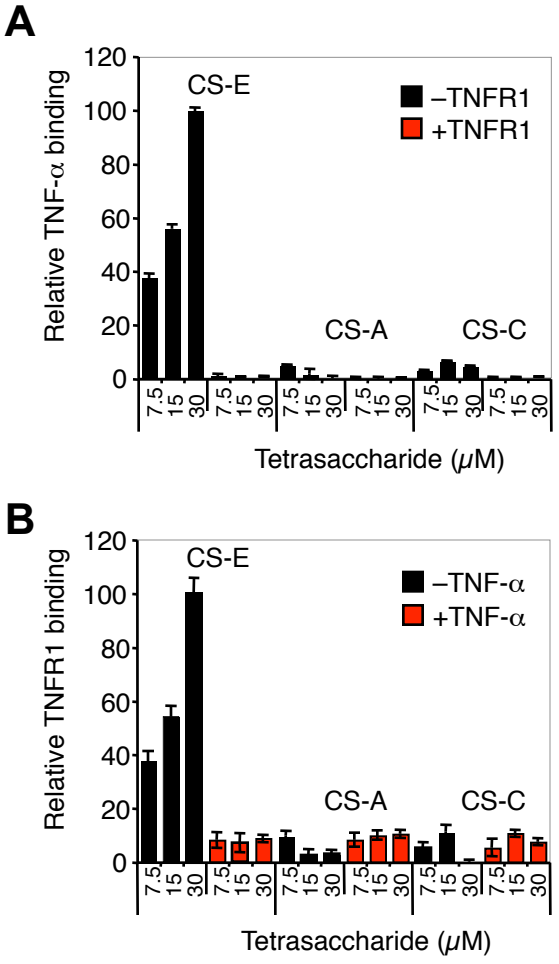

A
Relative TNF-α binding
120
100
80
60
40
20
0
CS-E
CS-A
CS-C
–TNFR1
+TNFR1
7.5
15
30
Tetrasaccharide (μM)
B
Relative TNFR1 binding
CS-E
CS-A
CS-C
–TNF-α
+TNF-α
Tetrasaccharide (μM)

Figure S3

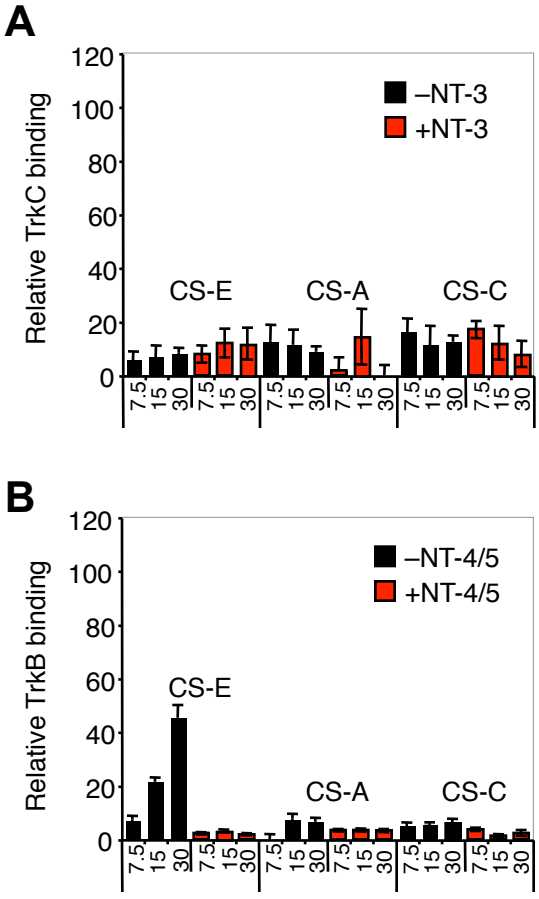
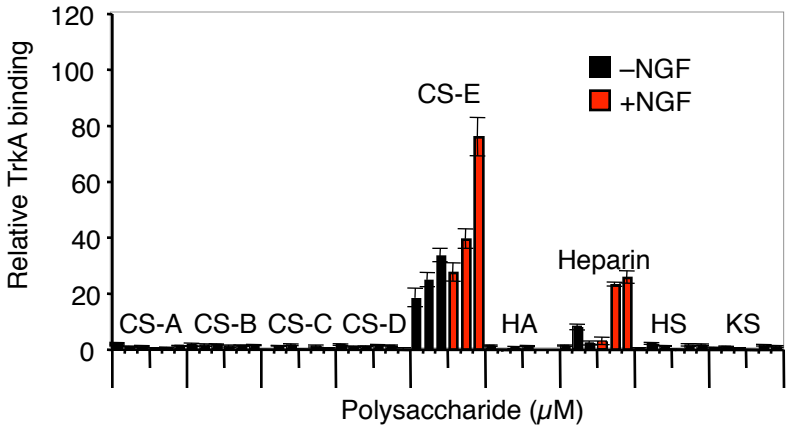
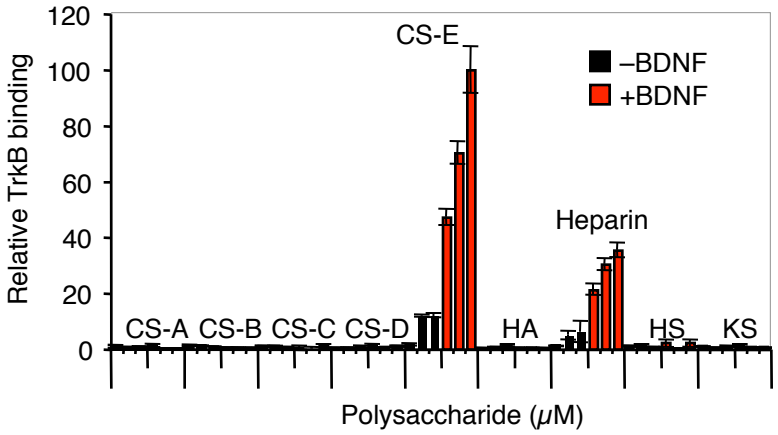
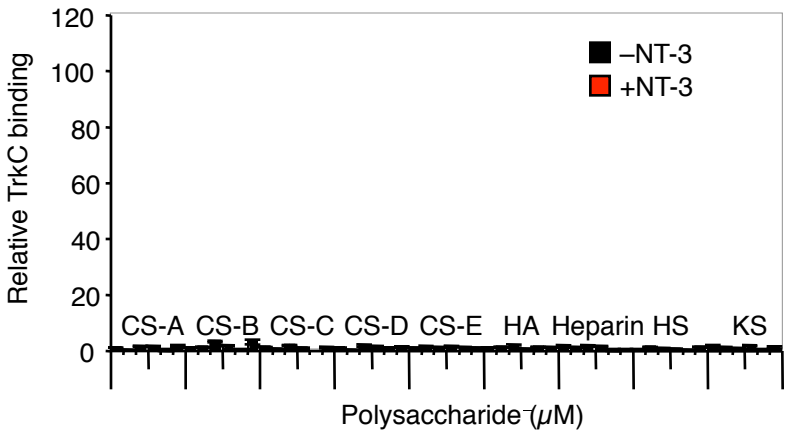

A
Relative TrkC binding
120
100
80
60
40
20
0
–NT-3
+NT-3
CS-E
CS-A
CS-C
7.5
15
30
B
Relative TrkB binding
–NT-4/5
+NT-4/5

Figure S4


Relative TrkA binding
120
100
80
60
40
20
0
–NGF
+NGF
CS-E
Heparin
CS-A CS-B CS-C CS-D HA HS KS
Polysaccharide (μM)


Relative TrkB binding
120
100
80
60
40
20
0
–BDNF
+BDNF
CS-E
Heparin
CS-A CS-B CS-C CS-D HA HS KS
Polysaccharide (μM)


Relative TrkC binding
120
100
80
60
40
20
0
–NT-3
+NT-3
CS-A CS-B CS-C CS-D CS-E HA Heparin HS KS
Polysaccharide (μM)

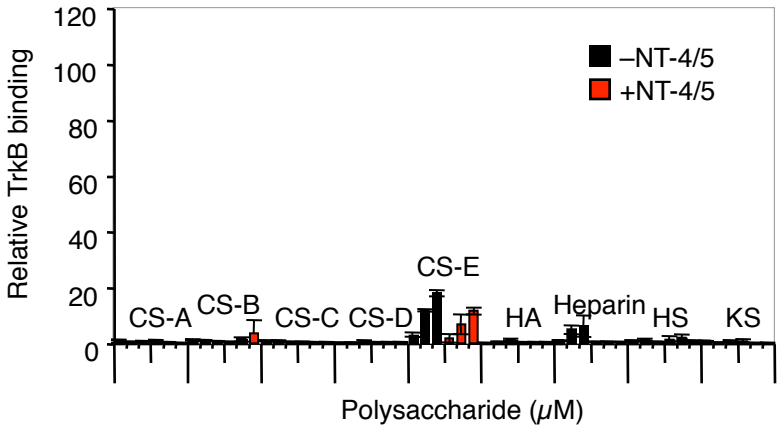

Relative TrkB binding
120
100
80
60
40
20
0
–NT-4/5
+NT-4/5
CS-E
CS-A CS-B CS-C CS-D HA Heparin HS KS
Polysaccharide (μM)

Figure S5

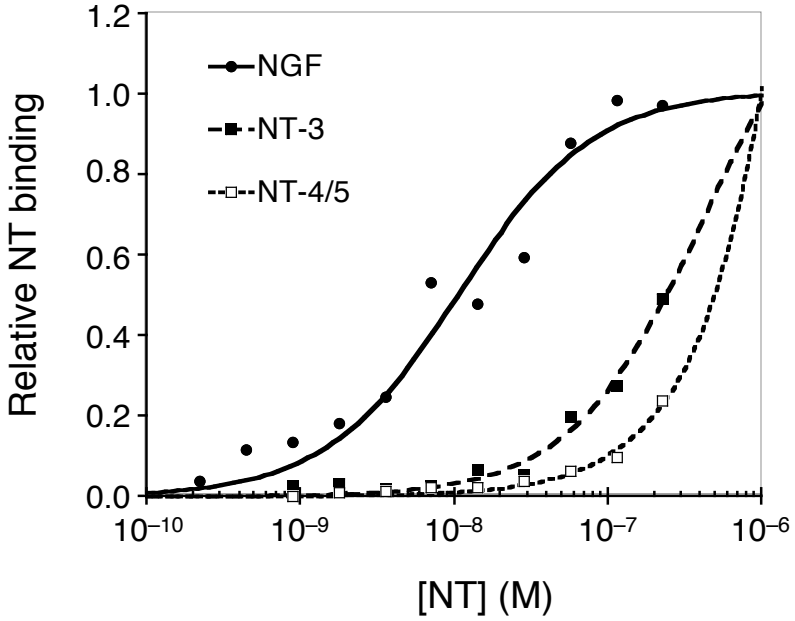

1.2
1.0
0.8
0.6
0.4
0.2
0.0
Relative NT binding
NGF
NT-3
NT-4/5
10–10
10–9
10–8
10–7
10–6
[NT] (M)

Figure S6

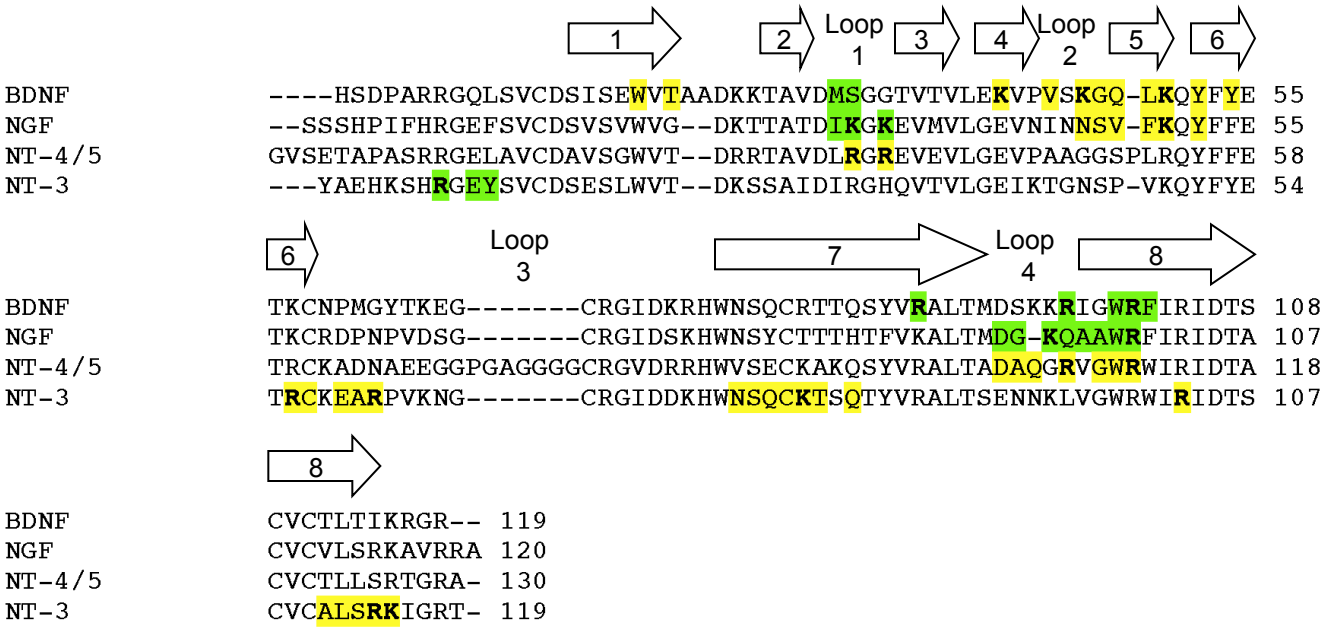

1
2
Loop 1
3
4
Loop 2
5
6
BDNF ----HSDPARRGQLSVCDSISEWVTAADKKTAVDMSGGTVTVLEKVPVSKGQ-LKQYFYE 55
NGF --SSSHPIFHRGEFSVCDSVSVWVG--DKTTATDIKGKEVMVLGEVNINNSV-FKQYFFE 55
NT-4/5 GVSETAPASRRGELAVCDAVSGWVT--DRRTAVDLRGREVEVLGEVPAAGGSPLRQYFFE 58
NT-3 ---YAEHKSHRGEYSVCDSESLWVT--DKSSAIDIRGHQVTVLGEIKTGNSP-VKQYFYE 54
6
Loop 3
7
Loop 4
8
BDNF TKCNPMGYTKEG-------CRGIDKRHWNSQCRTTQSYVRALTMDSKKRIGWRFIRIDTS 108
NGF TKCRDPNPVDSG-------CRGIDSKHWNSYCTTTHTFVKALTMDG-KQAAWRFIRIDTA 107
NT-4/5 TRCKADNAEEGGPGAGGGGCRGVDRRHWVSECKAKQSYVRALTADAQGRVGWRWIRIDTA 118
NT-3 TRCKEARPVKNG-------CRGIDDKHWNSQCKTSQTYVRALTSENNKLVGWRWIRIDTS 107
8
BDNF CVCTLTIKRGR-- 119
NGF CVCVLSRKAVRRA 120
NT-4/5 CVCTLLSRTGRA- 130
NT-3 CVCALSRKIGRT- 119

Figure S7

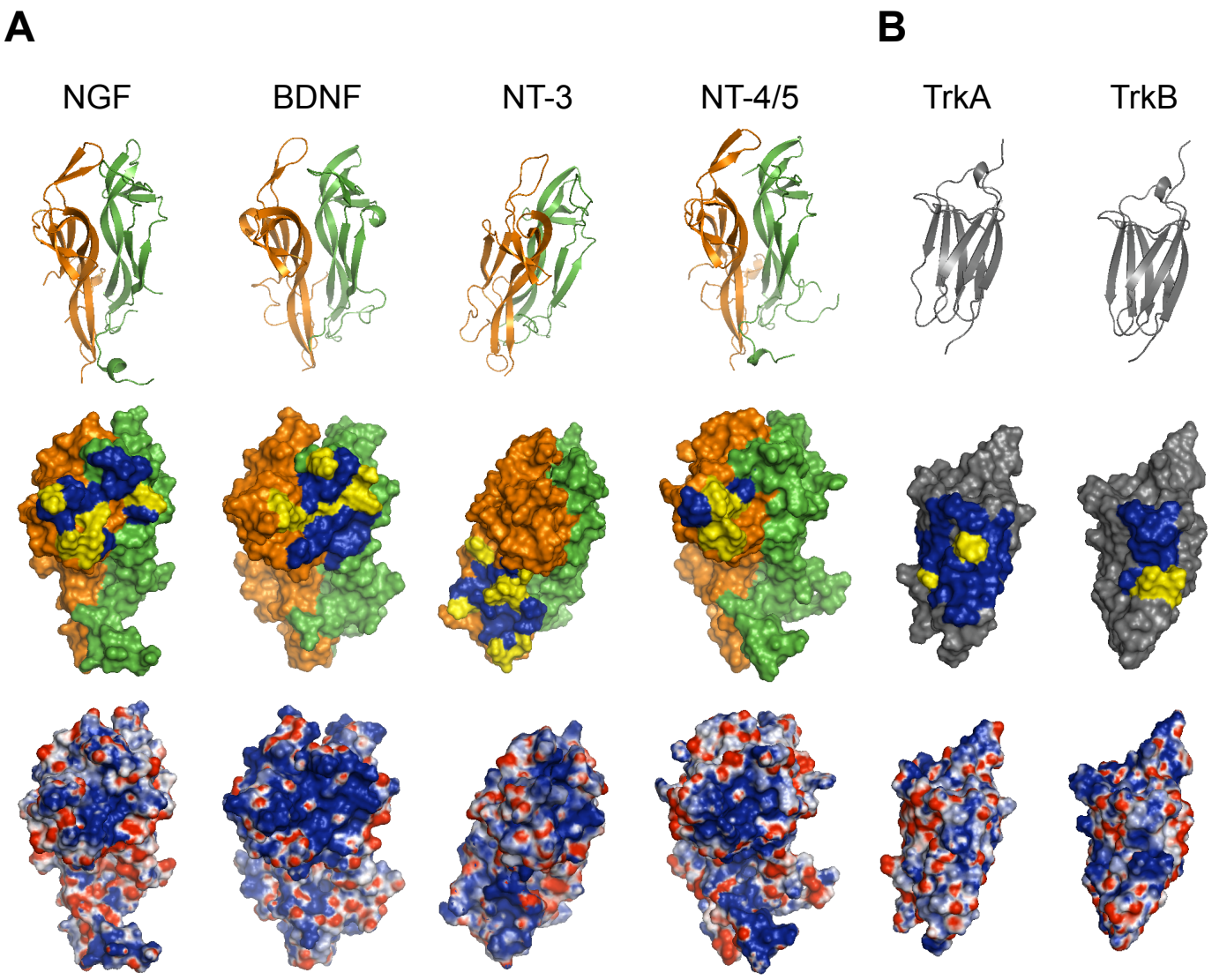


Figure S8

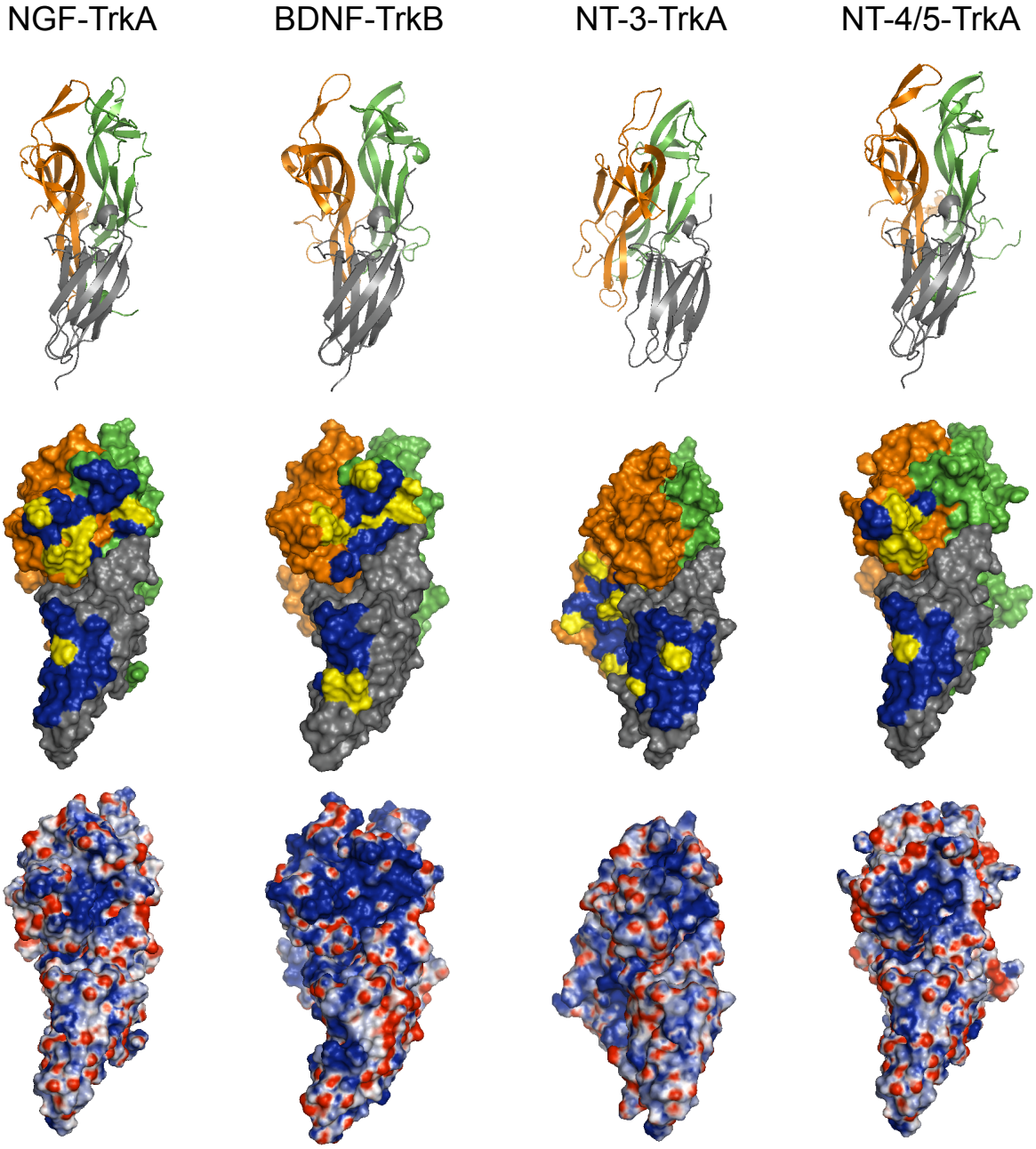


Figure S9

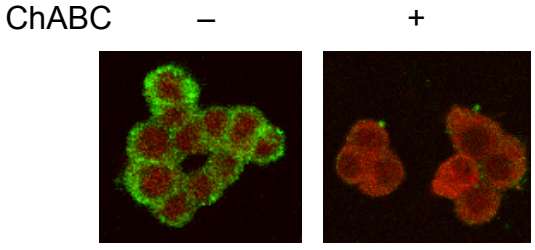


Figure S10

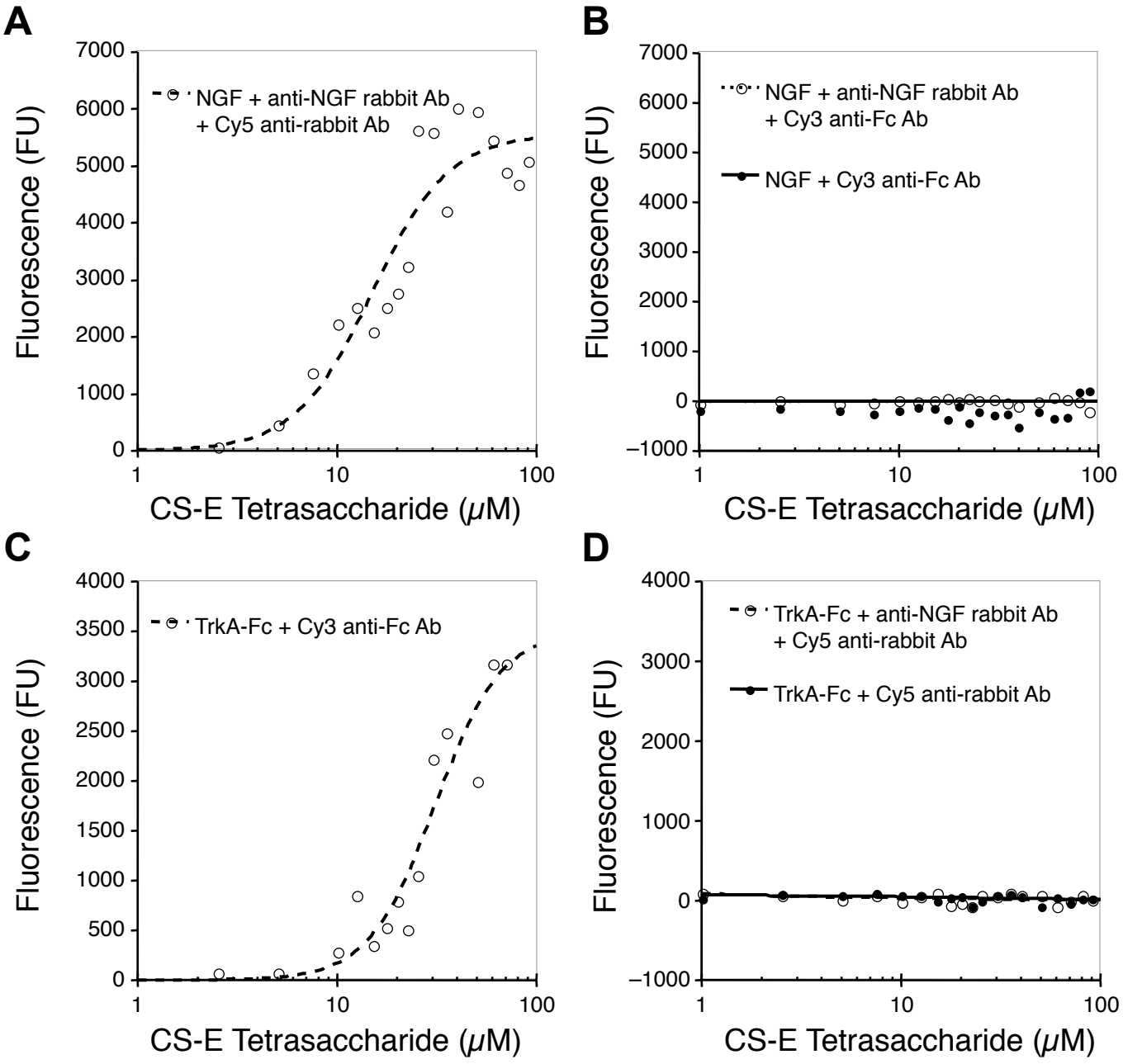

A
B
C
D
Fluorescence (FU)
CS-E Tetrasaccharide (μM)
NGF + anti-NGF rabbit Ab
+ Cy5 anti-rabbit Ab
NGF + anti-NGF rabbit Ab
+ Cy3 anti-Fc Ab
NGF + Cy3 anti-Fc Ab
TrkA-Fc + Cy3 anti-Fc Ab
TrkA-Fc + anti-NGF rabbit Ab
+ Cy5 anti-rabbit Ab
TrkA-Fc + Cy5 anti-rabbit Ab

Figure S11

Table S1. Predicted CS-A binding site on the DBL6 domain of VAR2CSA and associated K_D values for the interaction of DBL6 mutants with purified CS-A

Predicted CS-A Binding Site	K_D values of DBL6 mutants[a]
Lys2388	
Arg2389	
Asp2390	
Pro2391	
Lys2392	**Lys2932Ala**: ND[b]
Phe2394	
Lys2395	**Lys2395Ala**: ND[b]
Ile2452	
Leu2453	
Gly2454	
Lys2462	
Trp2466	
Met2469	
Asn2470	

[a] Data are from reference 28. For comparison, WT DBL6 had a K_D value of 80 μM. Blank entries correspond to residues whose contribution to the binding site was not tested

[b] K_D value too weak to be determined

Table S2. Predicted CS-A binding site on the DBL3x domain of VAR2CSA and associated K_D values for the interaction of DBL3x mutants with purified CS-A

Predicted CS-A Binding Site	K_D values of DBL3x mutants[a]
Asp1236	
Gly1237	
Lys1238	
Ile1239	
Phe1240	
Gly1242	
Lys1243	**Lys1243Ala**: 367 μM
Gly1244	
Gly1245	
Glu1246	
Gly1318	
Thr1319	
Ile1321	
Lys1324	**Lys1324Ala**: 122 μM
Asn1325	
Lys1328	**Lys1328Ala**: 89 μM
Gly1329	
Gln1330	
Arg1467	**Arg1467Ala**: 122 μM
Arg1503	
	Lys1504Ala: 172 μM
Lys1507	
Lys1510	**Lys1510Ala**: 193 μM
	Lys1515Ala: 488 μM[b]

[a] Data are from reference 28. For comparison, WT DBL3x had a K_D value of 33 μM. Blank entries correspond to residues whose contribution to the binding site was not tested

[b] Corresponds to a buried lysine residue that may not directly interact with CS-A

Table S3. Comparison of the heparin binding site on FGF2 as determined using computational and crystallographic methods[a]

Computational Prediction	Crystal Structure
Lys26[b]	
Asn27	Asn27
Gly28	
Leu118	Leu118
Lys119	**Lys119**
Arg120	**Arg120**
Thr121	Thr121
Lys125	**Lys125**
Lys129	**Lys129**
Gly133	Gly133
Gln134	Gln134
Lys135	**Lys135**
Ala136	Ala136
	Ile137

[a] Residues are numbered according to FGF2 structure 1FQ9

[b] Found in the heparin binding site of the structure 1BFC

Table S4. Comparison of the heparin binding site in the FGF2-FGFR1 complex determined using computational or crystallographic methods

Computational Prediction	Crystal Structure[a]	Computational Prediction	Crystal Structure[a]
FGF2 A, **Lys26**	FGF2 A, **Lys26**	FGFR1 A, **Lys160**	FGFR1 A, **Lys160**
FGF2 A, Asn27	FGF2 A, Asn27	FGFR1 A**, Lys163**	FGFR1 A, **Lys163**
FGF2 A, Gly28	FGF2 A, Gly28	FGFR1 A, His166	FGFR1 A, His166
FGF2 A, Leu118	FGF2 A, Leu118	FGFR1 A, Val168	
FGF2 A, **Lys119**	FGF2 A, **Lys119**	FGFR1 A, **Lys172**	FGFR1 A, **Lys172**
FGF2 A, **Arg120**	FGF2 A, **Arg120**	FGFR1 A, Thr173	FGFR1 A, Thr173
FGF2 A, Thr121	FGF2 A, Thr121	FGFR1 A, Val174	FGFR1 A, Val174
FGF2 A, **Lys125**	FGF2 A, **Lys125**	FGFR1 A, **Lys175**	FGFR1 A, **Lys175**
	FGF2 A, Leu126	FGFR1 A, Phe176	
	FGF2 A, Ser128	FGFR1 A, **Lys177**	FGFR1 A, **Lys177**
FGF2 A, **Lys129**	FGF2 A, **Lys129**		FGFR1 A, Tyr206
FGF2 A, Gly133	FGF2 A, Gly133	FGFR1 A, **Lys207**	FGFR1 A, **Lys207**
FGF2 A, Gln134	FGF2 A, Gln134		FGFR1 A, Val208
FGF2 A**, Lys135**	FGF2 A, **Lys135**	FGFR1 A, **Arg209**	FGFR1 A**, Arg209**
FGF2 A, Ala136	FGF2 A, Ala136	FGFR1 A, Thr212	
	FGF2 A, Ile137	FGFR1 A, Ser214	
	FGF2 B, Tyr24	FGFR1 A, Ile216	FGFR1 A, Ile216
FGF2 B, **Lys26**	FGF2 B, **Lys26**	FGFR1 A, Asp218	FGFR1 A, Asp218
FGF2 B, Asn27	FGF2 B, Asn27	FGFR1 B, **Lys160**	FGFR1 B, **Lys160**
FGF2 B, Gly28	FGF2 B, Gly28	FGFR1 B, **Lys163**	FGFR1 B, **Lys163**
	FGF2 B, Gly29	FGFR1 B, His166	FGFR1 B, His166
	FGF2 B, Ala117	FGFR1 B, Val168	FGFR1 B, Val168
FGF2 B, Leu118	FGF2 B, Leu118	FGFR1 B, **Lys172**	
FGF2 B, **Lys119**	FGF2 B, **Lys119**	FGFR1 B, Thr173	FGFR1 B, Thr173
FGF2 B, **Arg120**	FGF2 B, **Arg120**	FGFR1 B, Val174	
FGF2 B, Thr121		FGFR1 B, **Lys175**	FGFR1 B, **Lys175**
FGF2 B, **Lys125**	FGF2 B, **Lys125**	FGFR1 B, Phe176	
	FGF2 B, Leu126	FGFR1 B, **Lys177**	
	FGF2 B, Ser128		FGFR1 B, Tyr206
FGF2 B, **Lys129**	FGF2 B, **Lys129**	FGFR1 B, **Lys207**	FGFR1 B, **Lys207**
	FGF2 B, Thr130		FGFR1 B, Val208
	FGF2 B, Gly131	FGFR1 B, **Arg209**	FGFR1 B, **Arg209**
FGF2 B, Gly133	FGF2 B, Gly133	FGFR1 B, Thr212	
FGF2 B, Gln134	FGF2 B, Gln134	FGFR1 B, Ser214	
FGF2 B, **Lys135**	FGF2 B, **Lys135**	FGFR1 B, Ile216	FGFR1 B, Ile216
FGF2 B, Ala136	FGF2 B, Ala136	FGFR1 B, Asp218	FGFR1 B, Asp218
	FGF2 B, Ile137		

[a] Residues found within 5 Å of heparin in the heparin-FGF2-FGFR1 structure 1FQ9

Table S5. Predicted CS-E binding site on TNF-α

Monomer A	Monomer B
Ser71	**Lys65**
Thr72	Gln67
His73	Gly108
Val74	Ala109
Leu75	Glu110
Arg103	Ala111
Thr105	**Lys112**
Arg138	Pro113
	Tyr115
	Leu143

Table S6. Average binding energies (kcal/mol) of CS-A, CS-C, and CS-E for TNF-α and the NTs

	TNF	BDNF	NGF	NT-4/5	NT-3
CS-A	-116.80 ± 12.57	-560.93 ± 7.29	-265.65 ± 8.47	-250.11 ± 15.71	-309.20 ± 15.89
CS-C	-114.23 ± 11.29	-538.91 ± 20.16	-220.21 ± 10.45	-240.06 ± 10.30	-315.32 ± 11.42
CS-E	-176.12 ± 10.22	-824.94 ± 9.25	-446.0 ± 8.69	-407.77 ± 25.07	-569.74 ± 19.10

Table S7. Predicted CS-E binding sites on the NTs and Trks

BDNF	NGF	NT-4/5	NT-3	TrkA	TrkB
Monomer A, Met31	Monomer A, Ile31	Monomer A, **Arg34**	Monomer A, **Arg56**	Ala310	Ala314
Monomer A, Ser32	Monomer A, **Lys32**	Monomer A, **Arg36**	Monomer A, Cys57	Pro311	Gln316
Monomer A, **Arg88**	Monomer A, **Lys34**	Monomer A, Asp103	Monomer A, Glu59	Ser312	Phe318
Monomer A, **Arg97**	Monomer A, Asp93	Monomer A, Ala104	Monomer A, Ala60	Leu313	Ala322
Monomer A, Trp100	Monomer A, Gly94	Monomer A, Gln105	Monomer A, **Arg61**	**Arg314**	Ile323
Monomer A, **Arg101**	Monomer A, **Lys95**	Monomer A, **Arg107**	Monomer A, Asn76	Leu316	Leu324
Monomer A, Phe102	Monomer A, Gln96	Monomer A, Gly109	Monomer A, Ser77	Gly319	Asn325
Monomer B, Trp19	Monomer A, Ala97	Monomer A, Trp110	Monomer A, Gln78	Val321	Ile362
Monomer B, Thr21	Monomer A, Ala98	Monomer A, **Arg111**	Monomer A, Cys79	Asn323	**Lys364**
Monomer B, **Lys41**	Monomer A, Trp99		Monomer A, **Lys80**	Glu324	**Lys369**
Monomer B, Val44	Monomer A, **Arg100**		Monomer A, Thr81	Thr330	
Monomer B, **Lys46**	Monomer B, Asn46		Monomer A, Gln83	Glu331	
Monomer B, Gly47	Monomer B, Ser47		Monomer A, **Arg103**	Phe332	
Monomer B, Gln48	Monomer B, Val48		Monomer A, Ala111	**Arg342**	
Monomer B, Leu49	Monomer B, Phe49		Monomer A, Leu112	Thr360	
Monomer B, **Lys50**	Monomer B, **Lys50**		Monomer A, Ser113	Leu362	
Monomer B, Tyr52	Monomer B, Tyr52		Monomer A, **Arg114**	Ala364	
Monomer B, Tyr54			Monomer A, **Lys115**	Asn365	
			Monomer B, **Arg8**	Pro366	
			Monomer B, Glu10	Gly368	
			Monomer B, Tyr11	Gln369	
				Ser371	