

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization

International Bureau

(43) International Publication Date
01 November 2018 (01.11.2018)



(10) International Publication Number
WO 2018/200478 A2

(51) International Patent Classification:

A61K 39/395 (2006.01) C07K 16/30 (2006.01)

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(21) International Application Number:

PCT/US2018/029079

Published:

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))
- with sequence listing part of description (Rule 5.2(a))

(22) International Filing Date:

24 April 2018 (24.04.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/489,243 24 April 2017 (24.04.2017) US

(71) Applicants: UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC. [US/US]; Boyd Graduate Studies Research Center, Athens, GA 30602-7411 (US). GLYCOSENSORS AND DIAGNOSTICS, LLC [US/US]; 111 Riverbend Road, Athens, GA 30602 (US).

(72) Inventors: YANG, Loretta; 3937 Nobel Dr. Unit 130, San Diego, CA 92122 (US). SAMLI, Kausar, N.; 14023 95th Avenue NE, Kirkland, WA 98034 (US). WOODS, Robert, J.; 250 Greystone Terrace, Athens, GA 30606 (US). WU, Shengcheng; 112 Buttonwood Loop, Athens, GA 30605-4948 (US). COOPER, John, C.; 63 Cranberry Street Apt. 4, Brooklyn, NY 11201 (US). PAUL, Mal-lory, K.; 1425 Staples St NE Apt. 3, Washington, DC 20002 (US). SAUNDERS, Matthew, J.; 5075 Brighton Ave., San Diego, CA 92107 (US). ELETR, Ziad, M.; 77 Porter Street, Apt. 15, Somerville, MA 02143 (US).

(74) Agent: PROVENCE, David, L.; Mueting, Raasch & Gebhardt, P.A., P.O.Box 581336, Minneapolis, MN 55458-1336 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

(54) Title: SIALIC ACID BINDING POLYPEPTIDE

(57) Abstract: Sialic acid recognizing-affinity reagents engineered from the neuraminidase NanB have lectin-like properties and defined specificities for sialic acid.



WO 2018/200478 A2

SIALIC ACID BINDING POLYPEPTIDE

5

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application Serial No. 62/489,243, filed April 24, 2017, which is incorporated by reference herein.

10

GOVERNMENT FUNDING

This invention was made with government support under Grant No. R41GM113351 awarded by National Institutes of Health. The government has certain rights in the invention.

SEQUENCE LISTING

15

This application contains a Sequence Listing electronically submitted via EFS-Web to the United States Patent and Trademark Office as an ASCII text file entitled "Seq-List-02740201_ST25.txt" having a size of 22 kilobytes and created on April 16, 2018. The information contained in the Sequence Listing is incorporated by reference herein.

20

BACKGROUND

Glycans have several distinct properties that make their development as disease biomarkers appealing. Their location on cell surfaces makes them the first point of contact for cellular interactions and thus they are crucial in the control of normal metabolic processes. They also function as pathogen adhesion receptors. Glycan structures that are absent or are present in low amounts in a normal state on glycoprotein can proliferate or alter their sequence in disease states. A distinguishing feature of many glycans is a terminal sialic acid.

25

SUMMARY OF THE INVENTION

The present invention provides novel sialic acid-recognizing affinity reagents. The affinity reagents recognize, and bind to, sialic acid (also referred to as 5-(acetylamino)-3,5-dideoxy-D-glycero- α -D-galacto-non-2-ulopyranosonic acid, Neu5Ac, *N*-acetyl neuraminic acid, NANA, and

30

Sia) that is present on a glycosylated biomolecule, such as a glycoprotein, glycopeptide, glycolipid, oligosaccharide, or polysaccharide.

The sialic acid-recognizing affinity reagents provided herein are engineered Lectenz[®]-type proteins that have affinity and specificity for sialylated glycans. Lectenz[®] is a registered trademark of Glycosensors and Diagnostics (d/b/a Lectenz[®] Bio). Embodiments of the engineered sialic acid-recognizing affinity reagent of the invention include, without limitation: 1) pan-specific sialic acid-recognizing affinity reagents with broad specificity for sialo-glycans independent of linkage (α 2,3-, α 2,6-, and α 2,8- linkages) (referred to herein as a Sia-PS reagent), with a lead candidate Sia-PS1, described in Example I; and 2) α 2,3 sialic acid-recognizing affinity reagents specific for α 2,3-linked sialo-glycans over α 2,6, and α 2,8 linkages (referred to herein as a Sia-3S reagent) with a lead candidate, Sia-3S1, described in Example II. These reagents have low or no affinity for non-sialylated glycans or peptide backbones.

Advantageously, the sialic acid-recognizing affinity reagents of the invention have enhanced substrate specificity compared to antibodies or lectins that bind to sialic acid. The substrate specificity is tunable, as evidenced by comparing the Sia-PS affinity reagents and the Sia-3S affinity reagents described herein. Substrate specificity need not be context dependent, and the affinity reagents may be evolved to have desirable binding kinetics. Conveniently, the affinity reagents of the invention can be as efficiently produced as monomeric proteins.

In some embodiments, structurally guided genetic manipulations of *S. pneumoniae* NanB carbohydrate processing enzyme were used to identify sites for mutation, and select mutants with enhanced affinity for sialic acid, thereby converting NanB into a high specificity affinity reagent, known as a Lectenz[®]. Design and synthesis methods are described, in general, in US Pat. Pub. US2012/0040474 ("Glycan-Specific Analytical Tools") and WO2015/161201 (U.S. Ser. No. 15/304,725), each of which is explicitly incorporated by reference, and more specifically in Examples I and II below. However, it should be understood that the invention, as it relates to the various and particular sialic acid-recognizing affinity reagents identified and described herein, is not limited in any way by the method of making such reagents.

The sialic acid-recognizing affinity reagent of the invention is useful in both research and clinical settings. For example, it is useful for the detection of disease related sialic acid modifications of glycopeptides and glycoproteins. The affinity reagent may be employed as a capture reagent or recognition element in a variety of applications for the discovery of glycan-based disease markers, as well as in the quality control analysis of recombinantly produced

biopharmaceuticals, many of which are glycoproteins such as antibodies. These reagents can be adapted for such platforms as affinity chromatography enrichment, Western blot, or FACS-based detection.

5

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1 shows a three-dimensional representation of *S. pneumoniae* NanB. FIG. 1B is rotated 90 degrees from the view of FIG. 1A. In these images, NanB is shown in complex with a substrate, 2,7-anhydro-Neu5Ac, shown as a sphere, bound to the NanB active site. Gut et al., FEBS Lett., 2008, 582(23-24):2248-3352.

FIG. 2 shows a flowchart outlining the key steps in Lectenz[®] engineering.

FIG. 3A shows a model of 3'sialyllactosamine (3'SLN, stick) bound to enzyme with residues proximal (≤ 4.5 Å) to sialic acid shown in left panel as shaded residues and those proximal to the Gal/GlcNAc residues of 3'SLN shown in right panel as shaded residues.

FIG. 3B shows a sequence display of wild-type *Streptococcus pneumoniae* NanB (PDB ID: 2vw0). The 697 amino acid sequence (SEQ ID NO:1) is depicted and annotated. The three domains are labeled 2vw0A01, 2vw0A02, 2vw0A03. The image is from the RCSB PDB (www.rcsb.org) of PDB ID 2vw0 (Xu et al., J. Mol. Biol., 2008, 384: 436-449).

FIG. 3C shows an alignment of amino acid sequences of NanA (PDB ID 2VVZ; SEQ ID NO:2), NanB (PDB ID 2VW0; SEQ ID NO:1), and NanC (PDB ID 5F9T; SEQ ID NO:3). NanB and NanC have 50% sequence identity and both share 25% identity with NanA. "*" refers to fully conserved; ":" refers to conservation between groups of amino acids with strongly similar properties; and "." refers to conservation between groups of amino acids with weakly similar properties.

FIG. 3D shows an alignment of wild-type NanB (PDB ID 2VW0; SEQ ID NO:1) with the NanB fragment (30-697; SEQ ID NO:4) that is expressed from a pET28a+ based plasmid ("NanB:AA30-697(pET28a)") in Example I. Residues 1-29 in wild-type NanB represent a signal or leader sequence. "*" refers to fully conserved; ":" refers to conservation between groups of amino acids with strongly similar properties; and "." refers to conservation between groups of amino acids with weakly similar properties.

FIG. 4 shows relative enzymatic activity of NanB and Lectenz[®] clones.

FIG. 5 shows BLI sensograms for the binding of Lectenz[®] candidates to 3'SL immobilized on biosensors.

FIG. 6 shows BLI sensograms for the binding of Sia-PS Lectenz[®] candidates (analytes) to immobilized oligosaccharides. Lsp = lactose (no sialic acid), 3'SL = α 2,3-sialyllactose, 6'SL = α 2,6-sialyllactose, GD3 = GD3 oligosaccharide that contains a terminal α 2,8-linked Neu5Ac residue and an internal α 2,3-linked Neu5Ac residue. For example, Sia-PS1 analyte exhibits a steady-state K_D (3'SL) = 0.59 μ M, K_D (6'SL) = 0.69 μ M, K_D (GD3) = 0.30 μ M.

FIGs. 7A, 6B, and 6C show separation of sialylated and non-sialylated proteins by Sia-PS1 Lectenz[®] affinity chromatography (LAC). A purified Sia-PS1 sample (2 mg) was coupled chemically to a 1-mL sepharose column, followed by loading of an analyte. After competitive elution, the column was regenerated for the binding and elution of the next analyte. Experiments were performed on an AKTA Pure chromatography system. FIG. 7A shows the flow-through of non-sialylated glycoprotein horseradish peroxidase (HRP) and retention of fetuin in overlaid chromatograms. FIG. 7B shows an elution profile showing the separation of a mixture of BSA (non-glycosylated) and fetuin (sialylated glycoprotein). FIG. 7C shows an elution profile showing the separation of a mixture of asialofetuin (de-sialylated fetuin) and fetuin.

FIGs. 8A and 7B show LAC elution profiles for Sia-PS1 and synthetic neoglycoproteins. FIG. 8A shows a Sia-PS1 LAC elution profile of 3'-sialyllactose-BSA. FIG. 8A shows a Sia-PS1 LAC elution profile of 6'-sialyllactose-BSA. Each analyte (200 μ g) was applied onto the LAC column in binding buffer, washed and competitively eluted under gravity flow. The LAC column used was the same as the one described in FIG. 7. Protein concentration of collected fractions was determined by spectrophotometric absorption at 280 nm. For comparison purpose (see FIG. 9), the neoglycoproteins were each loaded at 200 μ g in 1 mL buffer volume to match the recommendation for a typical 1-mL lectin affinity columns. In each experiment, the amount of loaded analyte was nearly fully recovered by competitive elution.

FIGs. 9A and 8B show elution profiles for standard lectin affinity columns and synthetic neoglycoproteins: FIG. 9A shows an elution profile for 3'-sialyllactose-BSA (200 μ g) from a commercially available MAA column. FIG. 9B shows an elution profile for 6'-sialyllactose-BSA (200 μ g) from a commercially available SNA column.

FIGs. 10A, 10B and 10C show Sia-PS1 LAC enrichment of sialo-glycoconjugates from MCF7 cell-free extract (CFE). A sample (approximately 500 μ g) of the CFE, which was prepared according to Lee et al. [36], was loaded on to the Sia-PS1 LAC column (see FIG. 7), washed and

competitive eluted by a 1:1 mixture of 3'SL and 6'SL at 50 mM each in loading buffer. FIG. 10A shows a LAC profile showing the separation of the unbound proteins (flow-through peak; 319 μ g) at the binding step and the bound proteins (elution peak; 30 μ g). FIG. 10B shows SDS-PAGE of CFE (~125 μ g loaded) and LAC elution fraction (~7 μ g loaded). FIG. 10C shows MALDI profiles showing enrichment of small molecules in the LAC elution fraction. Note that the mass spectrometer used for the MALDI experiment is not sensitive enough to resolve molecules larger than 15 kD in size.

FIG. 11 shows glycan array screening results for Sia-PS1 indicating the ability to specifically recognize all sialylated glycans (Neu5Ac-glycans) on the array (I) and many non-human Neu5Gc-Sia glycans (II).

FIG. 12 shows amino acid iceLogo of enriched clone sequences. The wildtype sequence is shown across the bottom. Preferred amino acids at the five randomized positions are shown as a graphical representation. Y-axis indicates percentage of detection (% difference) for each amino acid. The prevalence of wild-type sequence at D237 and S673 is ignored by the iceLogo, hence a blank space at the top of the iceLogo for these positions.

FIG. 13 shows enzymatic activity of wild-type (wt) NanB, the R193A point mutant, and Sia-3S Lectenz[®] candidates.

FIGs. 14A and 13B show BLI sensograms for the binding of Sia-3S1 (analyte) under different buffer conditions to oligosaccharides (immobilized ligands): 3'SL = α 2,3-sialyllactose, 6'SL = α 2,6-sialyllactose. FIG. 14A shows a BLI sensogram for the binding of Sia-3S1 (analyte) to oligosaccharides using a buffer composed of 10 mM EPPS, 10 mM NaCl, pH 7.5. FIG. 13B shows a BLI sensogram for the binding of Sia-3S1 (analyte) to oligosaccharides using a buffer composed of 10 mM EPPS, 25 mM NaCl, pH 7.5.

FIG. 15 shows differential elution of fetuin bound to a Sia-3S1 LAC column. Purified Sia-3S1 (2 mg) was covalently coupled to a 1-mL sepharose column, followed by loading of 100 μ g fetuin. Fetuin molecules bound onto the Sia-3S1 column were competitively eluted with a 50 mM 3'SL or 6'SL solution in loading buffer. (The A280 UV spike at 18 mL retention volume is an instrument pump artifact.)

FIGs. 16A and 16B show Sia-3S1 LAC chromatography demonstrating selective binding to 3'SL-BSA over 6'SL-BSA. FIG. 16A shows results for 100 μ g 3'SL-BSA loaded onto the Sia-3S1 LAC column. FIG. 16B shows results for 100 μ g 6'SL-BSA loaded onto the Sia-3S1 LAC column. Sialylated BSA molecules bound onto the Sia-3S1 column were competitively eluted with 50 mM

3'SL or 6'SL solution before regeneration with 10 mM EPPS (pH 7.5) buffer containing 1 M NaCl ("NaCl regeneration"). Flow through amounts under each condition are approximately 5 µg.

FIG. 17 shows that Sia-3S1 LAC column binds only to 3'SL-BSA using binding buffer with 110 mM NaCl. 50 µg 3'SL-BSA, or 6'SL-BSA was loaded onto the Sia-3S1 LAC column.

5 Sialylated BSA molecules bound onto the Sia-3S1 column were eluted using 10 mM EPPS buffer (pH 7.5) with 1 M NaCl.

FIG. 18 shows that Sia-PS1, Sia-PS4, Sia-PS5, and Sia-3S1 bind to their corresponding glycoconjugates.

10 DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

The sialic acid-recognizing affinity reagent of the invention can be referred to as a Lectenz[®], which term describes a novel class of engineered, lectin-like enzyme-derived reagents with specificity for glycan structures of biological significance. Lectenz[®] reagents are engineered from a parent enzyme, but have little or no enzymatic activity, and have lectin-like carbohydrate-binding properties. In some embodiments, Lectenz[®] reagents retain the specificity of the endogenous parent enzymes from which they are engineered. In other embodiments, the specificity of a Lectenz[®] reagent is tuned to a similar glycan structure or motif. Lectenz[®] reagents therefore have several advantages over either lectins or antibodies, including tunable specificity, tunable affinity, and ease of production.

The sialic acid-recognizing affinity reagents described herein offer numerous advantages over sialic acid-recognizing lectins, because they have been engineered to have high affinity, yet possess substrate specificity based on the carbohydrate processing enzyme, NanB. In some embodiments, substrate specificity the sialic acid-recognizing affinity reagent may be the same as NanB. In some embodiments, the substrate specificity the sialic acid-recognizing affinity reagent may be broader than NanB, i.e., it can be expanded or broadened, as in the pan-specific Sia-PS affinity reagents described herein. In some embodiments, the substrate specificity of the sialic acid-recognizing affinity reagent may be narrower than NanB (enhanced or increased specificity). In designing or assessing sialic acid-recognizing affinity-enhanced modified NanB molecules using structurally guided genetic manipulations as described herein and more generally in US Pat. Pub. 2012/0040474 ("Glycan-Specific Analytical Tools") and WO2015/161201 (U.S. Ser. No.

15/304,725), the binding affinity and specificity of the sialic acid are independently tunable depending on the needs of the researcher or clinician.

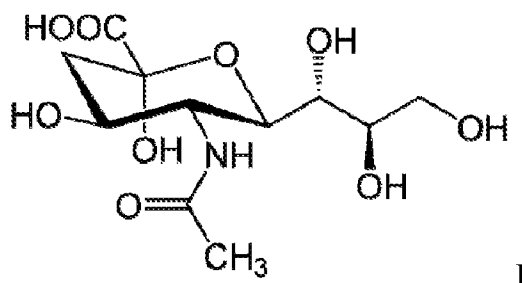
NanB sialidase

5 Sialidases (neuraminidases) remove sialic acid, a 9-carbon carbohydrate, as the terminal sugar of various glycoconjugates. The present invention utilizes *S. pneumoniae* NanB sialidase as the template for development of novel sialic acid-recognizing affinity reagents. A three-dimensional representation of *S. pneumoniae* NanB is shown in Figure 1. Panel B is rotated 90 degrees from the view of panel A. In these images, NanB is shown in complex with a substrate,
10 2,7-anhydro-Neu5Ac, shown as a sphere, bound to the NanB active site. Gut et al., FEBS Lett., 2008, 582(23-24):2248-3352.

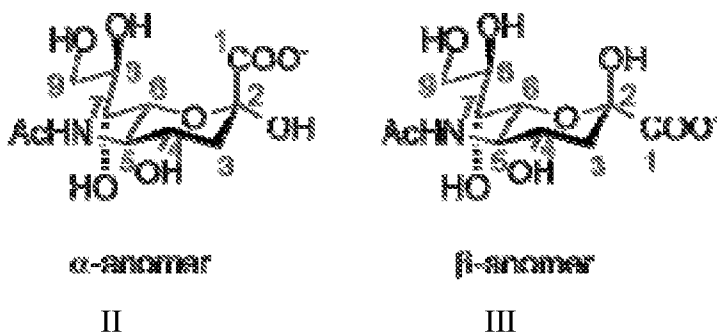
The L-domain (lectin-like), also known as the CBM domain or CBM40 domain, constitutes the carbohydrate-binding module (CBM), which has been classified within the CBM40 family. The catalytic N-domain, also known as the glycosyl hydrolase (GH) domain, of the GH33 domain,
15 which possesses neuraminidase activity, is also shown. A short linker (residues 225–230) connects the L-domain and the N-domain. An irregular domain, the I-domain, is also shown. With reference to SEQ ID NO:1, the approximate locations of the domains of NanB can be characterized as follows: a signal/leader sequence (residues 1-29), a CBM40 domain (residues 30-228), a catalytic domain (residues 229-345 and 459-697), and an irregular domain (residues 346-458). Active site
20 include the classical sialidase arginine triad, Arg245, Arg557, and Arg619, as well as Tyr653, Glu541, and Asp270 (Gut et al., FEBS Lett., 2008, 582(23-24):2248-3352; Xu et al., J. Mol. Biol., 2008, 384: 436-449).

The NanB template can be SEQ ID NO:1, or a truncated version of SEQ ID NO:1. Typically, the signal sequence is not included in the template; thus, a useful NanB template is
25 exemplified by amino acids 30-697 of SEQ ID NO:1.

The term "sialic acid" is often used to refer to a family of acidic sugars with a nine-carbon backbone. The most common sialic acid found in mammalian cells is *N*-acetylneuraminic acid (Neu5Ac, or NANA), which is acetylated at the C-5 position (Formula I).



5 The numbering of the sialic acid structure begins at the carboxylate carbon and continues around the chain. The configuration that places the carboxylate in the axial position is the alpha-anomer (Formula II).



10

In the present invention, "sialic acid" generally refers to Neu5Ac as described in Examples I and II but can, in the appropriate context, include other derivatives of neuraminic acid as well, such as variants that are hydroxylated at the C-5 position (e.g., ketodeoxynonic acid, Kdn), or that have a hydroxylated 5-*N*-acetyl group (e.g., *N*-glycolylneuraminic acid, Neu5Gc) or that are not *N*-acylated (e.g., neuraminic acid, Neu). Varki et al., "Sialic Acids," Ch. 14 in Essentials of Glycobiology, 2nd Ed. (Cummings et al., Ed.), Cold Spring Harbor Press, Cold Spring Harbor, NY (2009). Sialic acids can be present within a glycan at either a terminal position or an internal position. They can be found as a terminal saccharide component of biologically relevant *N*-glycans, *O*-glycans, and glycosphingolipids (gangliosides). They also are known to cap side chains of glycosylphosphatidylinositol (GPI) anchors. They may also be found as oligosialic acids or polysialic acids. Varki et al., "Sialic Acids," Ch. 14 in Essentials of Glycobiology, 2nd Ed. (Cummings et al., Ed.), Cold Spring Harbor Press, Cold Spring Harbor, NY (2009). Sialic acids may be present as a component of oligosaccharide chains of mucins, glycoproteins, glycopeptides,

15

20

and glycolipids, including glycosphingolipids and glycolipopeptides. Typically, sialic acid is linked to another monosaccharide in an α -linkage.

Sialic acid recognition

5 The sialic acid-recognizing affinity reagent of the invention recognizes (i.e., binds to) a sialic acid component an oligosaccharide. A sialic acid can be covalently bound to an adjacent monomer of the oligosaccharide in any of several linkage topographies. While two linkage conformations (α and β) are possible, typically the linkage to the adjacent saccharide is in the α conformation. Linkage can occur at any of a number of positions on the adjacent monomer, such as
10 the 3 position, 6 position, or 8 position of the adjacent monomer, resulting in an α 2,3 linkage, an α 2,6 linkage or an α 2,8 linkage, respectively. The oligosaccharide containing the sialic acid can be a disaccharide (2 monomers) or a higher oligosaccharide (more than 2 monomers). Advantageously, the oligosaccharide can be a part of (i.e., a component of) a glycan structure that is present in a glycosylated biomolecule, such as a glycoprotein, glycopeptide, or glycolipid.

15 The sialic-acid recognizing affinity reagents of the invention are based on, i.e., engineered from, the amino acid sequence of NanB, as exemplified by NanB from *S. pneumoniae* and shown in SEQ ID NO:1 (the full wild-type amino acid sequence, including the leader peptide; Fig. 3C) and SEQ ID NO:4 (the amino acid sequence without the leader peptide; Fig. 3D). Exemplary mutation sites are listed in Table 1, and can include one or more sites designated as important for binding
20 (e.g., R619, W674, R245, R676, Y653, S675, R264, H269, I246, I326, R557, and D327) and/or one or more sites designated as "tepid" or as making weak contributions to binding (e.g., E541, E669, S673, T539, L679, T268, P492, D270, Y671, A538, N353, N316, P660 and N683).

 It should be understood that affinity reagents engineered from NanB, such as Sia-PS1, Sia-PS2, Sia-PS3, Sia-PS4, Sia-PS5, Sia-3S1, Sia-3S2, Sia-3S3, Sia-3S5, Sia-3S6, Sia-3S7, Sia-3S7*,
25 Sia-3S8, Sia-3S9, which are engineered SEQ ID NO:4 and described in Examples I and II can, in some embodiments, include further modifications. For example, they can include truncations at or additions to the ends of SEQ ID NO:4, for example truncations or additions of 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids, and/or conservative mutations in regions not involved in activity or substrate binding, as long as the overall domain structure of the affinity reagent remains intact, and pan-
30 specific binding to sialic acid (in the case of Sia-PS reagents) or specific binding to Sia α 2,3 linkages (in the case of Sia-3S reagents) is preserved. Additionally, some embodiments of the sialic acid-recognizing affinity reagent may contain the carbohydrate binding module (CBM) domain of

NanB (e.g., residues from 30 through about 229 of SEQ ID NO:1), but not the glycosyl hydrolase (GH) domain; other embodiments of the sialic acid-recognizing affinity reagent may contain the glycosyl hydrolase (GH) domain (e.g., residues from about 229 through 697 of SEQ ID NO:1), but not the carbohydrate binding module (CBM) domain of NanB.

5

Pan-specific sialic acid-recognizing affinity reagent

In one embodiment, the sialic acid-recognizing affinity reagent is pan-specific in its recognition of (i.e., binding to) a sialic acid. This pan-specific affinity reagent recognizes a sialic acid (Neu5Ac) independent of the linkage configuration between the Neu5Ac and the adjacent
10 saccharide. The linkage configurations recognized by the pan-specific affinity reagent include Sia α 2,3-, α 2,6-, and α 2,8- linkages. Optionally and advantageously, as in the case of Sia-PS1, the pan-specific sialic acid-recognizing affinity reagent also recognizes common modifications of Neu5Ac, such as 9-*O* acetylation, and variants of the *N*-glycolyl form of sialic acid.

The pan-specific sialic acid-recognizing affinity reagent of the invention is exemplified by a
15 NanB fragment (residues 30-697; SEQ ID NO:4) with a single mutation. Exemplary mutation sites include Y653, D270, and E541. Exemplary pan-specific affinity reagents, each having a single mutation in the NanB GH domain as noted, are shown in Table 2 and include Sia-PS1 (Y653F), Sia-PS2 (D270A), Sia-PS3 (D270N), Sia-PS4 (E541A), and Sia-PS5 (E541Q). Surprisingly, these single site mutations have a dual effect. They reduce (Sia-PS2 and Sia-PS3) or effectively
20 eliminate (Sia-PS1, Sia-PS4, and Sia-PS5) catalytic activity, and they broaden specificity for sialic acid ligand binding. The pan-specific sialic acid-recognizing affinity reagent we chose for further validation study was Sia-PS1 that contains the Y653F at active site residue Tyr653. This one mutation, which involves removal of a single hydroxyl group, not only eliminates enzyme activity, but astonishingly introduces pan-specific binding into the mutated, inactivated enzyme, despite its
25 distance from the carbohydrate binding module (CBM) domain of the NanB.

Linkage-specific sialic acid-recognizing affinity reagent

In another embodiment, the sialic acid-recognizing affinity reagent is specific for a single sialic acid linkage configuration, namely, an α 2,3-linked sialic acid. These affinity reagents
30 recognize (i.e., bind to) a sialic acid linked to the adjacent saccharide in an α 2,3-linkage, but not a sialic acid linked to the adjacent saccharide in an α 2,6-linkage or an α 2,8-linkage.

The linkage-specific sialic acid-recognizing affinity reagent of the invention is exemplified by a NanB fragment (residues 30-697; SEQ ID NO:4) with multiple mutations. Together, the mutations reduce or eliminate catalytic activity, and alter the binding specificity of the inactivated enzyme to make it specific for α 2,3-linked sialic acid (versus α 2,6-linked sialic acid and α 2,8-linked sialic acid). Exemplary mutation sites include R193, D270, A538, E541, P660, S673, D327, and N683. Exemplary sialic acid-recognizing affinity reagents with heightened specificity for an α 2,3-linked sialic acid are shown in Table 4, include the following, with mutations sites identified:

| | | |
|----|----------|--|
| | Sia-3S1 | R193A, D270Q, A538V, E541D |
| 10 | Sia-3S2 | R193A, D270G, A538W, E541D, P660Q, S673A |
| | Sia-3S3 | R193A, D270G, A538V, E541D, S673A |
| | Sia-3S5 | R193A, D327E, A538F, E541A |
| | Sia-3S6 | R193A, D270G, A538H, E541A, N683S |
| | Sia-3S7 | R193A, D270H, D427Y, A538M, E541A, L690F |
| 15 | Sia-3S7* | R193A, D270H, A538M, E541A |
| | Sia-3S8 | R193A, D270I, D327V, A538F, E541I, S673Q |
| | Sia-3S9 | R193A, D270G, A538I, E541S, S673A |

Sia-3S7* is based on Sia-3S7, but without mutations D427Y and L690F. These two missing mutations are not expected to affect binding specificity.

Conjugates

The invention also includes conjugates of the sialic acid-recognizing affinity reagent. A conjugate includes, as a first component, a sialic acid-recognizing affinity reagent, which is covalently linked to at least one second component, which may be a proteinaceous component or a nonproteinaceous component. In some embodiments, a conjugate that includes a proteinaceous component can be synthesized as a fusion protein using well-known recombinant DNA methods. In some embodiments, the conjugate includes a proteinaceous or non-proteinaceous component that is chemically or enzymatically conjugated to the sialic acid-recognizing affinity reagent.

One example of a conjugate of the invention includes a sialic acid-recognizing affinity reagent conjugated to a therapeutic agent, also referred to herein as a drug. This conjugate is analogous to the well-known antibody-drug conjugate (ADC) except that the sialic acid-recognizing

affinity reagent is used in place of the antibody. Drugs that can be conjugated to a sialic acid-recognizing affinity reagent include, without limitation, cytotoxins, anti-metabolites, alkylating agents, antibiotics and anti-mitotic agents.

5 Anti-cancer, anti-inflammatory, pro-inflammatory, and immune-moderating drugs are particularly suitable for conjugation to a sialic acid-recognizing affinity reagent, since cancerous and precancerous conditions, as well as inflammation and immune conditions, are often associated with changes in protein glycosylation patterns. For example, a therapeutic or diagnostic radioactive agent can be coupled to or incorporated into a sialic acid-recognizing affinity reagent to yield a "Lectenz[®]-drug" conjugate that can be targeted to a cancer glycomarker. In one embodiment, the
10 therapeutic or diagnostic agent can be targeted to mucus linings or membranes, such as in the lungs or gut.

Likewise, anti-viral and anti-bacterial drugs are also particularly suitable for incorporation into a "Lectenz[®]-drug" conjugate, as targeting viral or bacterial glycosylated biomolecules has great therapeutic potential.

15 Another example of a conjugate of the invention includes a sialic acid-recognizing affinity reagent conjugated to a diagnostic or detection agent. The diagnostic or detection agent can include a detectable label, including but not limited to a radioactive, fluorescent, phosphorescent, colorimetric, enzymatic, immunological, magnetic, paramagnetic, diamagnetic or electromagnetic label. It should be understood that a sialic acid-recognizing affinity reagent need not be conjugated
20 to function as a diagnostic or detection agent, as the sialic acid-recognizing affinity reagent can be detected directly, e.g., via immunoassay.

Another example of a conjugate of the invention includes a sialic acid-recognizing affinity reagent conjugated to a marker sequence, for example a peptide such as hexa-histidine or hemagglutinin, to facilitate purification. Included in the invention are, for example, fusion proteins
25 that include a sialic acid-recognizing affinity reagent covalently linked to glutathione S-transferase (GST), thioredoxin, bovine serum albumin, bovine pancreatic trypsin inhibitor, or fluorescent proteins such as green fluorescent protein (GFP).

Methods of use

30 The vast number of potential applications for a sialic acid-recognizing affinity reagent, because of its lectin-like properties, will be immediately apparent to persons skilled in the art. The sialic acid-recognizing affinity reagents of the invention have many uses as diagnostic, therapeutic,

bioengineering or research reagents in clinical or research settings. For example, changes in the normal levels of glycan structures such as terminal sialic acid may be markers of disease states. New highly specific reagents are thus desirable in order to overcome current limitations in the discovery and exploitation of disease-related glycans. The sialic acid-recognizing affinity reagent
5 can be used to identify new, or detect previously identified, disease-associated sialic acid modifications of glycopeptides, glycoproteins and glycolipids. For example, the sialic-acid recognizing affinity reagent is useful for the detection of sialic acid and analysis of markers relevant to prostate-specific antigen (PSA) and influenza virus.

As another example, the sialic acid-recognizing affinity reagent can be employed as a
10 research or diagnostic tool in an affinity matrix for sample enrichment and can, in conjunction with existing mass spectrometric based methods, provide linkage information, as further described below.

The sialic acid-recognizing affinity reagent can be utilized in any method can otherwise be performed with an anti-glycan antibody or a lectin. Thus, sialic acid-recognizing affinity reagents
15 of the invention can be advantageously substituted for anti-glycan antibodies in numerous medical and laboratory methods, including diagnostic, analytical and therapeutic methods. Likewise, the sialic acid-recognizing affinity reagents of the invention can be advantageously substituted for lectins in numerous diagnostic and analytical laboratory methods. For example, a sialic acid-recognizing affinity reagent can be employed as a capture reagent, as in affinity chromatography
20 enrichment, or in histological studies, Western blots, FACS-based detection, and the like. As a capture reagent or recognition element, it is useful in a variety of applications for the discovery of glycan-based disease markers, as in glycoprofiling analysis (see, e.g., U.S. Pat Publ. 2014/0005069, published January 2, 2014). For example, sialic acid-recognizing reagent of the invention can be a component of a multicomponent microarray that contains other Lectenz[®] reagents and/or lectins,
25 antibodies, or other carbohydrate-binding molecules, with varying specificities.

Therapeutic uses are also envisioned and are described in more detail below. For example, the sialic acid-recognizing affinity reagents can be used for detection of sialic acid relevant for quality control in the manufacturing of biologics. More specifically, they can be used to detect the presence, absence, or linkage position of sialic acid on synthetic or recombinantly produced
30 glycoprotein-based or glycolipid-based biologics, such as therapeutic antibodies, as step in quality control during synthesis or storage.

Diagnostic and analytical methods

A sialic acid-recognizing affinity reagent or conjugate thereof can be used to detect sialic acid-containing glycans in a biological or synthetic sample. For example, a biological sample, such as a tissue or fluid, can be contacted with the sialic acid-recognizing affinity reagent or conjugate thereof, alone or in conjunction with other analytical reagents, to detect and/or characterize the level or type of glycosylation and/or glycation in the biological sample. Characterization can include of the glycan can include identifying a constituent saccharide of the glycan, determining saccharide composition of the glycan, determining linkage positions within the glycan, or determining stereochemistry of the glycan. As another example, a sialic acid-recognizing affinity reagent or conjugate thereof can be used for quality control in the synthesis of therapeutic biologics, for example in the synthesis of therapeutic antibodies, to monitor the level or type of glycosylation. See PCT patent publication WO2012/118928, published September 7, 2012, and US Pat. Publ. 2014/0005069, published January 2, 2014. A sialic acid-recognizing affinity reagent or conjugate thereof can be utilized as an affinity reagent or as part of an affinity matrix; for example, it can be tethered to a solid support, such as a surface, column, resin, bead, particle or nanoparticle, and used in methods to detect or enrich for sialic acid-containing compounds in or from biological or synthetic samples. Tethered sialic acid-recognizing affinity reagents can also be used to isolate and/or purify synthetic glycosylated compounds.

Diagnostics can be performed on a biological sample obtained from a subject, but can also be performed *in vivo*. In *in vivo* applications, a sialic acid-recognizing affinity reagent or conjugate thereof is administered to a subject, and binding of the sialic acid-recognizing affinity reagent within the subject is detected. Preferably, a conjugate is administered to the subject, wherein the conjugate includes a detectable label so as to facilitate biomedical imaging. Examples of a suitable conjugate include a sialic acid-recognizing affinity reagent conjugated to a radiolabel, a paramagnetic label, or a diamagnetic label.

The sialic acid-recognizing affinity reagent can be used to interrogate biological samples in the search for abnormal glycosylation. Examples of biological samples include, but are not limited to, any biological fluid, tissue, or organ. Examples of the biological fluids include, but are not limited to blood, urine, serum, saliva, cerebra-spinal fluid, and semen. In other embodiments, a sialic acid-recognizing affinity reagent can be used for a detection of the presence or amount of a target analyte in biological fluids and tissues. Examples of targets are exogenously consumed species, such as plant polysaccharides, carbohydrate-based drugs, and pathogens, whose surfaces

are often coated in complex distinct glycans. The sialic acid-recognizing affinity reagent also has application in drug discovery and evaluation of biological activity of new glycan-based compounds.

The sialic acid-recognizing affinity reagent can be used for diagnosing, and/or treating diseases manifested by abnormal glycosylation. It can be used to detect certain tumor antigens comprising glycoproteins, glycolipids, and/or a variety of carbohydrate epitopes. A number of these tumor antigens have been found to be up-regulated in the neoplastic disease state. Examples of tumor antigens that can signal a development and progression of a neoplastic disorder include, but are not limited to, carcinoembryonic antigen (CEA), which is a glycoprotein associated with colorectal, gastric, pancreatic, lung, and breast carcinomas, and the developing fetus; carbohydrate antigen 19-9 (CA 19-9), or sialylated Lewis A antigen, which is present in a glycolipid found in patients with pancreatic cancer; and carbohydrate antigen 15-3 (CA15-3), associated with breast cancer.

The presence of the antigen does not necessarily indicate transformation to a cancerous cell; however, its localization in the cell is indicative, as in the case of CEA. For this reason, there is a need for highly selective and high affinity analytical tools. The diagnostic tests currently rely on antibodies that were often generated against the peptide portions of the glycoprotein or sugar portions of glycolipid, however, the exact epitopes are only now being defined. In the examples in which the glycans have been characterized, multiple glycoforms are often present (CEA, for example). Lacking reagents that are able to discriminate between glycoforms, it is currently impossible to determine the extent to which subtle variations in glycosylation correlate with disease state, cancer type, or tissue localization. At present, these questions can be addressed primarily by MS analyses of isolated glycoproteins, which are examined as mixtures of glycoforms. Typically, the only level of glycoform-focusing that is performed is the enrichment in high-mannose containing glycans using lectin (concanavalin A, (Con A)) affinity chromatography. More efficient laboratory analyses and routine clinical diagnostic techniques remain severely limited by the lack of glycoform-specific reagents.

The sialic acid-recognizing affinity reagent may have utility for quantifying the relative abundances of various glycoforms present for any given glycoprotein in a biological sample. As used herein, the term "glycoform" refers to a protein with a specific glycan attached. A glycoprotein can have multiple glycoforms. More specifically, a glycoform is an isoform of a protein that differs only with respect to the number or type of attached glycan; the amino acid sequence is the same for the various glycoforms. Glycoproteins often comprise a number of

different glycoforms, with alterations in the attached saccharide or oligosaccharide.

Advantageously, a sialic acid-recognizing affinity reagent can be used to enrich the biological sample for sialic acid-containing glycans. It can likewise be used to identify specific glycosylation sites on the protein surface to which the glycans are attached. It can also be used to separate intact glycopeptides from a proteolytic digest of a glycoprotein. Enriching the sample in the analyte of interest is of great assistance in the further characterization of the glycopeptides fractions. In particular, enrichment facilitates the identification of the peptide sequence and the glycan structure, which can enable the identification within the intact protein of the glycosylation sites and the characterization of the particular glycans present at each glycosylation site.

The sialic acid-recognizing affinity reagent can be used in monitoring specific glycan modifications of proteins in biological fluids, tissues, organs, or living cells. Recognition is not expected to depend on the identity of the protein, and the sialic acid-recognizing affinity reagent is expected to be able to recognize any protein sequence that includes a sialic acid (consistent with the particular specificity of the sialic acid-recognizing affinity reagent, i.e., whether it is pan-specific or specific for a particular linkage), and therefore will be very useful for detection of particular glycan modifications.

In yet other embodiments, the sialic acid-recognizing affinity reagent can be used for *in vitro* or *in vivo* staining cells or tissues.

The sialic acid-recognizing affinity reagent can be employed to monitor sialylation of glycans in a mixture, as might arise during the production of recombinant glycoproteins for use in the pharmaceutical or research industries.

In the foregoing embodiments, the sialic acid-recognizing affinity reagent can be tagged with a stain or a dye and applied to a biological sample comprising cells or tissues or glycoproteins or glycopeptides or oligosaccharides or polysaccharides of interest.

Another aspect of the present invention provides methods of using sialic acid-recognizing affinity reagent for analytical applications. The sialic acid-recognizing affinity reagent of the present invention can be used as sialic acid-specific analytical tool. Glycan-specific analytical tools have potential use as a method of detection in many areas, including environmental, fermentation, food and medical areas and could be used for *in vivo* or *in vitro* sensing in humans or animals. For example, the sialic acid-recognizing affinity reagent of the present invention can be used as an affinity reagent or as vehicle for tissue staining. As another example, the sialic acid-recognizing affinity reagent can be used for enriching a biological sample for sialic acid-containing glycans. In

yet other examples, the sialic acid-recognizing affinity reagent can be used to determine specific glycosylation sites on glycoproteins.

In certain embodiments, the sialic acid-recognizing affinity reagent can be used as a reagent for affinity separation, including, for example, affinity chromatography. Affinity chromatography is a method of separating biochemical mixtures, based on a highly specific biological interaction such as that between the binding protein and the glycan. The present invention is not limited to any specific design or chromatographic system. In general, the sialic acid-recognizing affinity reagent will be either covalently attached or otherwise immobilized to the solid support, and will constitute a stationary phase. In certain embodiments, the stationary phase that is derivatized with the sialic acid-recognizing affinity reagent can be used in column chromatography. In these embodiments, the particles of the solid stationary phase will be used to fill the whole inside volume of the tube (packed column). Alternatively, the solid phase particles will be concentrated on or along the inside tube wall leaving an open, unrestricted path for a biological sample (i.e., the mobile phase) in the middle part of the tube (open tubular column). In other embodiments, the derivatized stationary phase can be used for batch chromatography. In these embodiments, the stationary phase can be added to a vessel and mixed with the biological sample. Although the foregoing example generally focused on affinity chromatography, it is understood that these principals are readily applied to other affinity purification protocols.

20 *Therapeutic methods*

In certain embodiments, the sialic acid-recognizing affinity reagent of the invention can be used as a therapeutic agent or modified for delivery of an active therapeutic agent. Since the sialic acid-recognizing affinity reagent of the present invention has a defined glycan specificity, a delivery of the therapeutic agents can be targeted only to those cells, tissues, or organs that display a biomolecule, such as a glycoprotein or glycolipid with the glycan structure recognized by the sialic-acid recognizing affinity reagent.

The potential therefore exists for the sialic acid-recognizing affinity reagent to be used as a therapeutic in many applications such as targeted drug delivery. Changes in the levels and locations of sialic acid-containing glycans have been shown to be associated with many diseases, including cancer. The sialic acid-recognizing affinity reagent is thus expected to have direct applications in the field of cancer research, potentially leading to the development of a product for the detection of certain forms of cancer. It is also expected to have utility as a reagent for use in

glycomics, wherein it may be used to enrich samples containing sialic acid-containing glycans, thereby enabling detection and analysis of these important carbohydrates. A sialic acid-recognizing affinity reagent can be used as a vehicle for targeted delivery of therapeutic agents.

5 A sialic acid-recognizing affinity reagent, or conjugate thereof, can be administered to a subject to treat or prevent an infection, disease, or disorder. The infection can be, for example, viral, bacterial, parasitic, or fungal. The disease or disorder can result from an exogenous agent, or it can be autologous or autoimmune.

10 In one embodiment, a sialic acid-recognizing affinity reagent is administered to a subject so as to bind to a sialic acid-containing glycan present within the subject, so as to achieve a therapeutic or prophylactic effect. The sialic acid-containing glycan can be an endogenous biomolecule produced by the subject, or it can be an exogenous biomolecule produced by a pathogen. In one embodiment, the sialic acid-recognizing affinity reagent binds to an endogenous biomolecule, for example a biomolecule associated with cancer, a precancerous condition, or an immune disorder of the subject. In another embodiment, the sialic acid-recognizing affinity reagent prevents binding of a pathogen to a host cell; in another embodiment, the sialic acid-recognizing affinity reagent prevents internalization of a pathogen into a host cell.

15 In another embodiment, a conjugate of a sialic acid-recognizing affinity reagent is administered to a subject, wherein the conjugate includes a therapeutic agent as exemplified above. The therapeutic agent can be an antibiotic agent, for example an agent that targets a microbial pathogen. The therapeutic agent can be an agent that targets an autologous or autoimmune disease, for example an anti-cancer agent, such as a cytotoxin, or an immunoactive agent. Examples of therapeutic agents that can be used for site-specific delivery include, but are not limited to, various chemotherapeutic, antibiotic, and antiviral agents, toxins, radioisotopes, cytokines, etc.

20 A sialic acid-recognizing affinity reagent or conjugate thereof for therapeutic use can be tested for toxicity in suitable animal model systems, for example in rats, mice, monkeys, or rabbits. The usefulness of a sialic acid-recognizing affinity reagent or conjugate thereof to treat or prevent a viral infection can be assessed by evaluating its ability to inhibit viral replication, inhibit viral transmission or to treat or prevent symptoms associated with viral infection. Likewise the usefulness of a sialic acid-recognizing affinity reagent or conjugate thereof to treat or prevent a bacterial infection can be assessed by evaluating its ability to inhibit the bacterial replication, or to treat or prevent symptoms associated with bacterial infection. Usefulness in treating cancer can be

evaluated by assessing the ability of a sialic acid-recognizing affinity reagent or conjugate thereof to inhibit the growth or metastasis of cancerous cells, to inhibit angiogenesis, or to cause cell death.

Method of making

5 The sialic acid-recognizing affinity reagent of the invention may be expressed in a host cell using genetic engineering techniques. The term "cell" is meant to include any type of biological cell. The host cell can be a eukaryotic cell or a prokaryotic cell. Preferably, the host cell is a prokaryotic cell such as a bacterial cell; however single cell eukaryotes such as protists or yeasts are also useful as host cells. Preferred host cells are microbial cells, preferably the cells of single-
10 celled microbes such as bacterial cells or yeast cells. Notwithstanding the above preferences for bacterial and/or microbial cells, it should be understood that the sialic acid-recognizing affinity reagent can be expressed without limitation in the cell of an animal, plant, insect, yeast, protozoan, bacterium, or archaeobacterium. Examples of microbial cells that can be engineered to express the sialic acid-recognizing affinity reagent of the invention, in addition to *E. coli*, include a wide
15 variety of bacteria and yeast including members of the genera *Escherichia*, *Salmonella*, *Clostridium*, *Zymomonas*, *Pseudomonas*, *Bacillus*, *Rhodococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Lactobacillus*, *Enterococcus*, *Arthrobacter*, *Brevibacterium*, *Corynebacterium*, *Candida*, *Hansenula*, *Pichia* and *Saccharomyces*. Preferred microbial cells include, without limitation, *Escherichia coli*, *Bacillus subtilis*, *Bacillus licheniformis*, *Alcaligenes eutrophus*,
20 *Rhodococcus erythropolis*, *Paenibacillus macerans*, *Pseudomonas putida*, *Enterococcus faecium*, *Saccharomyces cerevisiae*, *Lactobacillus plantarum*, *Enterococcus gallinarum* and *Enterococcus faecalis*.

 A cell that has been genetically engineered to express the sialic acid-recognizing affinity reagent of the invention may be referred to as a "host" cell, a "recombinant" cell, a "genetically
25 engineered" cell or simply an "engineered" cell. These and similar terms are used interchangeably. A genetically engineered cell contains one or more artificial sequences of nucleotides which have been created through standard molecular cloning techniques to bring together genetic material that is not natively found together. DNA sequences used in the construction of recombinant DNA molecules can originate from any species. For example, plant DNA may be joined to bacterial
30 DNA, or human DNA may be joined with fungal DNA. Alternatively, DNA sequences that do not occur anywhere in nature may be created by chemical synthesis of DNA or by directed mutation of DNA, and incorporated into recombinant molecules. Proteins that result from the expression of

recombinant DNA are often termed recombinant proteins. Examples of recombination are described in more detail below and may include inserting foreign polynucleotides (obtained from another species of cell) into a cell, inserting synthetic polynucleotides into a cell, or relocating or rearranging polynucleotides within a cell. Any form of recombination may be considered to be genetic engineering and therefore any recombinant cell may also be considered to be a genetically engineered cell.

As will be appreciated by a person of skill in the art, expression of a protein, such as the sialic acid-recognizing affinity reagent of the invention, can be achieved through a number of molecular biology techniques. For example, expression can be achieved by introducing into the host cell one or more copies of a polynucleotide encoding the desired protein. The polynucleotide encoding the desired protein may be endogenous or heterologous to the host cell. Preferably, the polynucleotide is introduced into the cell using a vector; however, naked DNA may also be used. The polynucleotide may be circular or linear, single-stranded or double stranded, and can be DNA, RNA, or any modification or combination thereof. The vector can be any molecule that may be used as a vehicle to transfer genetic material into a cell. Examples of vectors include plasmids, viral vectors, cosmids, and artificial chromosomes, without limitation. Examples of molecular biology techniques used to transfer nucleotide sequences into a microorganism include, without limitation, transfection, electroporation, transduction, and transformation. These methods are well known in the art. Insertion of a vector into a target cell is usually called transformation for bacterial cells and transfection for eukaryotic cells, however insertion of a viral vector is often called transduction. The terms transformation, transfection, and transduction, for the purpose of the instant invention, are used interchangeably herein. A polynucleotide which has been transferred into a cell via the use of a vector is often referred to as a transgene.

Preferably, the vector is an expression vector. An "expression vector" or "expression construct" is any vector that is used to introduce a specific polynucleotide into a target cell such that once the expression vector is inside the cell, the protein that is encoded by the polynucleotide is produced by the cellular transcription and translation machinery. Typically an expression vector includes regulatory sequences operably linked to the polynucleotide encoding the desired protein. Regulatory sequences are common knowledge to the person of the skill in the art and may include for example, an origin of replication, a promoter sequence, and/or an enhancer sequence. The polynucleotide encoding the desired protein can exist extrachromosomally or can be integrated into the host cell chromosomal DNA.

Extrachromosomal DNA may be contained in cytoplasmic organelles, such as mitochondria (in most eukaryotes), and in chloroplasts and plastids (in plants). More typically, extrachromosomal DNA is maintained within the vector on which it was introduced into the host cell. In many instances, it may be beneficial to select a high copy number vector in order to maximize the expression of the protein. Optionally, the vector may further contain a selectable marker. Certain selectable markers may be used to confirm that the vector is present within the target cell. Other selectable markers may be used to further confirm that the vector and/or transgene has integrated into the host cell chromosomal DNA. The use of selectable markers is common in the art and the skilled person would understand and appreciate the many uses of selectable markers. Optionally, the vector may further contain a reporter gene. Reporter genes may be used to confirm that the vector is expressing within the target cell, and may be further used to monitor the expression from the vector. The use of reporter genes is common in the art and the skilled person would understand and appreciate the many uses of reporter genes.

A sialic acid-recognizing affinity reagent of the invention can be isolated and optionally purified from any genetically engineered cell described herein. It can be isolated directly from the cells, or from the culture medium, for example, during an aerobic or anaerobic fermentation process. Isolation and/or purification can be accomplished using known methods.

Also provided by the invention is a kit that includes a sialic acid-recognizing affinity reagent, conjugate, fusion protein or affinity matrix of any of the preceding claims, and instructions for use.

EXEMPLARY EMBODIMENTS

Embodiment 1. A sialic acid-recognizing affinity reagent comprising a catalytically inactive NanB neuraminidase protein, or fragment thereof, having at least one amino acid mutation compared to a corresponding wild-type NanB neuraminidase protein, wherein the mutation

(a) reduces or eliminates the neuraminidase activity of the NanB protein; and

(b) affects sialic acid binding affinity or binding specificity;

and wherein said affinity reagent binds to a sialic acid component of a glycan.

Embodiment 2. The sialic acid-recognizing affinity reagent of embodiment 1, wherein the NanB fragment comprises at least one of the carbohydrate binding module (CBM) domain of the NanB protein and the glycosyl hydrolase (GH) domain of the NanB protein.

Embodiment 3. The sialic acid-recognizing affinity reagent of embodiment 1 or 2, wherein the affinity reagent is pan-specific for sialic acid.

Embodiment 4. The sialic acid-recognizing affinity reagent of embodiment 3, wherein the affinity reagent binds to Neu5Ac linked to an adjacent saccharide monomer in an α 2,3 linkage, Neu5Ac linked to an adjacent saccharide monomer in an α 2,6 linkage, and Neu5Ac linked to an adjacent saccharide monomer in an α 2,8 linkage.

Embodiment 5. The sialic acid-recognizing affinity reagent of embodiment 3 or 4 which binds to at least one variant of Neu5Ac.

Embodiment 6. The sialic acid-recognizing affinity reagent of any of embodiments 3-5, wherein the mutation is at a site selected from the group consisting of Y653, D270, and E541, as represented in SEQ ID NO:1 or SEQ ID NO:4.

Embodiment 7. The sialic acid-recognizing affinity reagent of embodiment 6, wherein the mutation is in *S. pneumoniae* NanB (SEQ ID NO:1 or SEQ ID NO:4) or fragment thereof, or a corresponding position in a homologous NanB sequence.

Embodiment 8. The sialic acid-recognizing affinity reagent of embodiment 6 or 7, wherein the mutation is selected from the group consisting of Y653F, D270A, D270N, E541A, and E541Q.

Embodiment 9. The sialic acid-recognizing affinity reagent of embodiment 8, wherein the mutation is Y653F.

Embodiment 10. The sialic acid-recognizing affinity reagent of embodiment 8 selected from the group consisting of Sia-PS1, Sia-PS2, Sia-PS3, Sia-PS4, and Sia-PS5.

Embodiment 11. The sialic acid-recognizing affinity reagent of embodiment 10 which is Sia-PS1.

Embodiment 12. The sialic acid-recognizing affinity reagent of embodiment 1 or 2, wherein the affinity reagent has a plurality of amino acid mutations compared to a corresponding wild-type NanB neuraminidase protein.

Embodiment 13. The sialic acid-recognizing affinity reagent of embodiment 12, wherein said plurality of mutations comprises:

(a) at least one first mutation that reduces or eliminates the catalytic activity of the NanB protein; and

(b) at least one second mutation that affects binding affinity or binding specificity.

Embodiment 14. The sialic acid-recognizing affinity reagent of embodiment 12 or 13, wherein the affinity reagent is specific for Neu5Ac that is linked to an adjacent saccharide monomer in an α 2,3 linkage.

Embodiment 15. The sialic acid-recognizing affinity reagent of any of embodiments 12-14, wherein the plurality of mutations are at sites selected from the group consisting of R193, D270, A538, E541, P660, S673, D327, and N683, as represented in SEQ ID NO:1 or SEQ ID NO:4.

Embodiment 16. The sialic acid-recognizing affinity reagent of embodiment 15, wherein the plurality of mutations are in *S. pneumoniae* NanB (SEQ ID NO:1 or SEQ ID NO:4) or fragment thereof, or a corresponding position in a homologous NanB sequence.

Embodiment 17. The sialic acid-recognizing affinity reagent of embodiment 15 or 16, wherein the plurality of mutations are selected from the group consisting of R193A, D270Q, D270G, D270H, D270I, A538V, A538W, A538F, A538H, A538M, A538I, E541D, E541A, E541I, E541S, P660Q, S673A, S673Q, D327E, D327V, and N683S.

Embodiment 18. The sialic acid-recognizing affinity reagent of embodiment 17, wherein the plurality of mutations comprise R193A, D270Q, A538V, and E541D.

Embodiment 19. The sialic acid-recognizing affinity reagent of embodiment 17 selected from the group consisting of Sia-3S1, Sia-3S2, Sia-3S3, Sia-3S5, Sia-3S6, Sia-3S7, Sia-3S7*, Sia-3S8, and Sia-3S9.

Embodiment 20. The sialic acid-recognizing affinity reagent of embodiment 19 which is Sia-PS1.

Embodiment 21. The sialic acid-recognizing affinity reagent of any of the preceding embodiments, wherein the glycan is a constituent of a glycosylated biomolecule.

Embodiment 22. The sialic acid-recognizing affinity reagent of embodiment 21, wherein the glycosylated biomolecule comprises a glycoprotein, a glycopeptide, a glycolipid, a glycolipoprotein, or a glycolipopeptide.

Embodiment 23. A conjugate comprising a first component comprising a sialic acid-recognizing affinity reagent of any of the preceding embodiments, covalently linked to a second component.

Embodiment 24. The conjugate of embodiment 23, wherein the second component is a proteinaceous component.

Embodiment 25. The conjugate of embodiment 23, wherein the second component is a nonproteinaceous component.

Embodiment 26. The conjugate of any of embodiments 23-25 wherein the second component is a therapeutic or diagnostic agent.

Embodiment 27. A fusion protein comprising a sialic acid-recognizing affinity reagent of any of embodiments 1 to 22.

Embodiment 28. An affinity matrix comprising a sialic acid-recognizing affinity reagent, conjugate or fusion protein of any of the preceding embodiments.

Embodiment 29. The affinity matrix of embodiment 28 selected from the group consisting of a solid support, surface, column, resin, bead, microarray, particle and nanoparticle.

Embodiment 30. A kit comprising a sialic acid-recognizing affinity reagent, conjugate, fusion protein or affinity matrix of any of the preceding embodiments, and instructions for use.

Embodiment 31. An isolated polynucleotide encoding a sialic acid-recognizing affinity reagent or conjugate of any of embodiments 1 to 24, or the fusion protein of embodiment 27.

Embodiment 32. A vector comprising a polynucleotide of embodiment 31.

Embodiment 33. The vector of embodiment 32, wherein the vector is an expression vector.

Embodiment 34. A host cell comprising a vector of embodiment 32 or 33.

Embodiment 35. The host cell of embodiment 34, wherein the host cell is a bacterial cell.

Embodiment 36. A method for making the sialic acid-recognizing affinity reagent or conjugate of any of embodiments 1-24, or the fusion protein of embodiment 27, the method comprising expressing the affinity reagent, conjugate or fusion protein in host cell.

Embodiment 37. A method for detecting a sialic acid component of a glycan, the method comprising:

contacting a biological or laboratory sample with a sialic acid-recognizing affinity reagent, conjugate or fusion protein of any of embodiments 1-27 under conditions to allow binding of the affinity reagent, conjugate or fusion protein to the sialic acid; and detecting the sialic acid.

Embodiment 38. The method of embodiment 37, wherein the biological or laboratory sample comprises a recombinant antibody.

Embodiment 39. The method of embodiment 37, wherein the glycan is a biomarker.

Embodiment 40. The method of embodiment 39, wherein the biomarker is a cancer biomarker.

Embodiment 41. A method for enriching, isolating or purifying a sialic acid-containing glycan, the method comprising:

contacting a sialic acid-recognizing affinity reagent, conjugate or fusion protein of any of embodiments 1-27 under conditions to allow binding of the affinity reagent, conjugate or fusion protein to the sialic acid so as to yield an enriched, isolated or purified sialic acid-containing glycan.

Embodiment 42. A diagnostic or therapeutic composition comprising a sialic acid-recognizing affinity reagent, conjugate or fusion protein of any of embodiments 1-27.

Embodiment 43. The diagnostic or therapeutic composition of embodiment 42 wherein the sialic acid-recognizing affinity reagent, conjugate or fusion protein is detectably labeled.

Embodiment 44. The diagnostic or therapeutic composition of embodiment 43 wherein the detectable label comprises a radioactive label, a fluorescent label, a phosphorescent label, a colorimetric label, an enzymatic label, an immunological label, a magnetic label, a paramagnetic label, a diamagnetic label or an electromagnetic label.

Embodiment 45. Use of a sialic acid-recognizing affinity reagent, conjugate or fusion protein of any of embodiments 1-27 as a therapeutic agent.

Embodiment 46. Use of a sialic acid-recognizing affinity reagent, conjugate or fusion protein of any of embodiments 1-27 as a diagnostic agent.

Embodiment 47. Use of a sialic acid-recognizing affinity reagent, conjugate or fusion protein of any of embodiments 1-27 for targeted drug delivery.

Embodiment 48. Use of a sialic acid-recognizing affinity reagent, conjugate or fusion protein of any of embodiments 1-27 for detection of the presence or amount of sialic acid in a biological or laboratory sample.

Embodiment 49. A compound, composition, or method including one or more of the features described herein.

The present invention is illustrated by the following examples. It is to be understood that the particular examples, materials, amounts, and procedures are to be interpreted broadly in accordance with the scope and spirit of the invention as set forth herein.

5

EXAMPLES

Example I. Pan-Specific Sialic Acid Reagents

10 Introduction

At present multiple lectins (typically from *S. Nigra*, *M. amurensis*, and wheat germ) have to be employed in serial lectin affinity chromatography to enrich sialylated glycans [2]. Because these lectins (especially WGA and MAL) show cross-reactivity with non-sialylated glycans [3, 4] their use risks capturing irrelevant components. The Sia-PS1 Lectenz[®] (PS = pan-specific) is under development for the detection and enrichment of general sialo-glycoconjugates regardless of terminal Neu5Ac (Sia α 2,3-, α 2,6-, and α 2,8-) linkages to the adjacent saccharide. To our knowledge, this is the first reagent that is pan-specific for sialic acid (Sia, neuraminic acid, Neu5Ac), and as such, it has the potential to significantly enhance sialo-glycan biomarker detection and glycomics enrichment.

15

20 Methods

The flowchart presented in Figure 2 outlines schematically the methods employed in Example I and II to screen, optimize and select the Lectenz[®] candidates. The specificity and affinity of potential Lectenz[®] candidates is evaluated through structurally-guided site-specific and/or site-saturation mutagenesis assays designed to enable the selection of optimal clones. Bio-layer Interferometry (BLI) is used to determine the binding kinetics and affinities for binding to glycans and glycoproteins. Affinity chromatography is performed by immobilizing the Lectenz[®]

25

candidates in a column format to evaluate their ability to fractionate sialylated- and asialo-glycans and glycoconjugates. All promising Lectenz[®] candidates are screened against the glycan array at the CFG, which currently contains over 600 *O*- and *N*-linked glycans.

5 Results and Conclusions

Computational simulation of enzyme–ligand complexes predicted Sia-PS1 specificity

NanB is a well-characterized enzyme that recognizes sialylated glycan substrates. Its sequence and additional characteristics have been previously described in the literature [9].

10 Computational simulations [10] were employed to identify residues in NanB that are most likely to be suitable to mutagenesis for substrate affinity optimization. The 3D models shown in Figure 3A were built using the online tools at GLYCAM-Web (www.glycam.org).

After ligand alignment, the enzyme-ligand complexes were subjected to fully solvated molecular dynamics (MD) simulation [11-13], followed by calculation of average interaction
15 energies using molecular mechanics–generalized Born surface approximation (MM-GBSA) method [12, 14]. See Table 1 for a list of NanB free energies for selected amino acids bound to 3'-sialyllactosamine (Neu5Ac α 2,3Gal β 1,4Glc β NAc, or 3'SLN). The simulation analysis indicates that Sia-PS Lectenz[®] should display significant affinity for sialylated glycans with no measurable affinity for non-sialylated glycans. The generated Sia-PS Lectenz[®] are listed in Table 2, which also
20 identifies mutagenesis site sequence identities. Table 3 lists the differing physical and chemical properties of these Sia-PS Lectenz[®] reagents.

Table 1. Per-residue binding free energies (kcal/mol) for WT NanB bound to 3'SLN. Residues listed contributed at least 0.5 kcal/mol to either the total molecular mechanical energies via van der Waals forces (ΔE_{VDW}) or electrostatic (ΔE_{ELE}) interactions, or to the total binding free energy ($\Delta G_{Binding}$). Library columns apply to Example II, and indicate residues selected for optimization for knowledge-based library design (X-C-P = any amino acids for mutagenesis, but Cys or Pro; X-C-E-P = any amino acids, except Cys, Glu, or Pro).

| Residues Important for Binding | ΔE_{VDW} | ΔE_{ELE} | ΔG_{GB-SA} | $\Delta G_{Binding}$ | Library 1 | Library 2 | Library 3 |
|---|------------------|------------------|--------------------|----------------------|-----------|-----------|-----------|
| R619 | 0.5 | -47.3 | 40.0 | -6.8 | | | |
| W674 | -4.7 | 0.7 | -0.4 | -4.3 | | | |
| R245 | -0.1 | -41.7 | 37.7 | -4.2 | | | |
| R676 | -2.4 | -23.6 | 22.7 | -3.2 | | | |
| Y653 | -1.9 | 0.4 | -0.6 | -2.2 | | | |
| S675 | -0.1 | -2.3 | 1.0 | -1.5 | | | |
| R264 | -0.2 | -27.7 | 26.6 | -1.3 | | | |
| H269 | -0.9 | -1.3 | 0.9 | -1.2 | | | |
| I246 | -1.0 | -0.8 | 0.8 | -1.0 | | | |
| I326 | -0.9 | -0.6 | 0.6 | -0.9 | | | |
| R557 | -0.4 | -28.3 | 27.9 | -0.7 | | | |
| D327 | -0.5 | 6.1 | -6.3 | -0.7 | X-C-P | X-C-P | X-C-P |
| Tepid residues making only weak contributions | | | | | | | |
| E541 | -0.9 | 21.8 | -21.6 | -0.6 | X-C-E-P | X-C-E-P | X-C-E-P |
| E669 | -0.1 | 23.5 | -23.9 | -0.6 | | | |
| S673 | -1.1 | 2.4 | -2.0 | -0.6 | X-C-P | X-C-P | X-C-P |
| T539 | -0.7 | -1.4 | 1.7 | -0.5 | | | |
| L679 | -0.5 | -0.6 | 0.6 | -0.5 | | | |
| T268 | -0.4 | 1.4 | -1.4 | -0.4 | | | |
| P492 | -0.4 | 0 | 0.1 | -0.4 | | | |
| D270 | -2.3 | 18.5 | -16.5 | -0.4 | X-C-P | X-C-P | X-C-P |
| Y671 | -0.2 | 0.9 | -1.0 | -0.2 | | | |
| A538 | -0.2 | -1.1 | 1.3 | 0 | X-C-P | X-C-P | X-C-P |
| N353 | -0.1 | -0.7 | 0.7 | 0 | | | |
| N316 | -0.3 | 0.4 | -0.1 | 0 | | | |
| P660 | 0.0 | -0.1 | 0.1 | 0 | | | |
| N683 | -4.7 | -0.3 | 0.3 | 0 | | | |

Table 2: Site mutations of selected pan-specific sialic acid-recognizing Lectenz[®] reagents based on NanB. The shaded boxes indicate retention of the wild type amino acid at that site.

| wtNanB | D270 | E541 | Y653 |
|---------|------|------|------|
| Sia-PS1 | D | E | F |
| Sia-PS2 | A | E | Y |
| Sia-PS3 | N | E | Y |
| Sia-PS4 | D | A | Y |
| Sia-PS5 | D | Q | Y |

5 **Table 3:** Physical and chemical properties of NanB and pan-specific sialic acid recognizing Lectenz[®] reagents. ExPASy ProtParam calculated properties based on amino acid sequence are reported. [15] Molecular weight, isoelectric point, and extinction coefficients (ϵ) values are listed.

| Proteins (690 Amino Acids) | Molecular Weight | Isoelectric Point | ϵ (M ⁻¹ cm ⁻¹) | ϵ (L g ⁻¹ cm ⁻¹) | ϵ 1% |
|-------------------------------|---------------------|----------------------|--|--|---------------|
| wtNanB | 76965.9 | 6.53 | 97640 | 1.269 | 12.69 |
| Sia-PS1 | 76949.9 | 6.53 | 96150 | 1.250 | 12.50 |
| Sia-PS2 | 76921.9 | 6.67 | 97640 | 1.269 | 12.69 |
| Sia-PS3 | 76964.9 | 6.67 | 97640 | 1.269 | 12.69 |
| Sia-PS4 | 76907.9 | 6.67 | 97640 | 1.270 | 12.70 |
| Sia-PS5 | 76964.9 | 6.67 | 97640 | 1.269 | 12.69 |

10 **Sia-PS Lectenz[®] have no or attenuated enzymatic activity**

The neuraminidase activity of the wt NanB enzyme and its mutants was measured using, *p*NP- Neu5Ac as a colorimetric substrate. As shown in Figure 4, in comparison to the wild type enzyme, Sia-PS3 was 19% active, Sia-PS2 was 9% active and others were inactive. Thus, Sia- PS1, Sia-PS4 and Sia-PS5 are suitable Lectenz[®] candidates.

15

Sia-PS Lectenz[®] reagents have unique target affinity

Steady-state binding constants between Lectenz[®] candidates and sialic acid ligands were measured by BLI. Figure 5 showed that Sia-PS4 has the highest binding response with a sub-micromolar K_D value to 3'SL while Sia-PS2, a candidate that has residual neuraminidase activity, has the lowest binding response with a higher K_D value. The five candidates' binding to 6'-sialyllactose (Neu5Ac α 2,6Gal β 1,4Glc β , or 6'SL) yielded similar K_D values, albeit at a reduced

binding response (Figure 6). Notably, the affinity of Sia-PS1, Sia-PS4, and Sia-PS5 to 3'SL or 6'SL is comparable or superior to reported values for the lectins MAL ($KD \sim 1-3 \mu\text{M}$ [16, 17]) and SNA ($0.6 - 0.9 \mu\text{M}$ [16]).

Sia-PS1 detects Neu5Ac independent of linkage

Additional BLI measurements confirmed that Sia-PS1, Sia-PS2, Sia-PS3, Sia-PS4, and Sia-PS5 recognizes Neu5Ac α 2,3-, α 2,6-, and α 2,8- linkages present in 3'SL, 6'SL and the oligosaccharide of ganglioside GD3 (Neu5Ac α 2,8Neu5Ac α 2,3Gal β 1,4Glc β), respectively. The
5 binding affinity is micromolar or sub-micromolar for each of the oligosaccharides (Figure 6).

Lectenz[®] affinity chromatography

A typical application of the Lectenz[®] we are developing is using affinity chromatography of complex samples for biomarker enrichment. The low KD value and off rate for Sia-PS1 ($k_{off} = 0.008 \text{ s}^{-1}$) make it ideal as a matrix for affinity chromatography. Indeed, in Lectenz[®] Affinity
10 Chromatography (LAC), Sia-PS1 coupled to a Sepharose column did not retain any non-sialylated glycoprotein horseradish peroxidase (HRP); all of it was in the flow-through. In contrast, sialylated fetuin was retained on the column and eluted (Figure 7A). In 1:1 mixtures, the column captured sialylated fetuin, while non-sialylated BSA or de-sialylated asialofetuin
15 flows through the column at the analyte's loading step (Figures 7B and 7C).

Sia-PS1 specifically binds to both 3'-sialylated and 6'-sialylated glycoproteins

In addition to the BLI data (Figure 6), we conducted LAC using synthetic sialylated glycoproteins (neoglycoproteins), 3'-sialyllactose-BSA (3'SL-BSA) and 6'-sialyllactose-BSA
20 (6'SL-BSA), as the analytes. The Sia-PS1 Lectenz[®] column was able to capture each of the neoglycoproteins that were subsequently eluted by 3'SL and 6'SL, respectively (Figure 8).

Sia-PS1 is superior to lectins as an enrichment reagent for sialo-glycoconjugates

For comparison, commercially available affinity columns (1 mL) of MAA (known as 3'Sia-specific) and SNA (known as 6'Sia-specific) were loaded with 3'SL-BSA and 6'SL-BSA,
25 respectively, at the recommended loading levels of 200 μg analyte in 1 mL loading buffer. As evident in Figure 9, while each column retained some of the neoglycoconjugates, significant amounts were observed in the flow-through fractions. According to the manufacturer's technical data, the binding capacity of the lectin columns far exceeds the 200 μg glycoproteins loaded. In

addition, the elution peaks are far broader than those eluted from Sia-PS1 LAC (Figure 8). Although a number of factors, such as the quality of the lectin columns, or the incompatibility between the immobilized lectins and the neoglycoconjugate analytes, or the manufacturer's recommended eluant might have affected the results, these experiments do indicate the
5 superiority of Sia-PS1 as an enrichment reagent for sialo-glycoconjugates.

Sia-PS1 enriches sialo-glycoconjugates from MCF7 cell-free extract

Application of Sia-PS1 LAC has been applied to enrichment of sialo-glycoconjugates from MCF7 cell-free extracts [18]. As shown in Figure 10, about 10% of the total proteins in the
10 MCF7 extract were affinitive to the Sia-PS1 LAC column and could be eluted by competitive reagent 3'SL & 6'SL. We were able to enhance binding efficiency by circulation of about 1.5 mg of MCF7 proteins through the column at 4°C overnight, which resulted in an eluted peak of 207 µg proteins, or 13.8%. It is expected that comparison of the identity and relative abundance of the Sia-PS1 LAC-enriched proteins to known glycosylation databases will further demonstrate
15 the specificity and effectiveness of Sia-PS1 as a biomarker discovery reagent.

Glycan array screening confirms Sia-PS1 Lectenz[®] specificity for sialo-glycans

Glycan array screening was performed using a biotinylated Sia-PS1 Lectenz[®] sample at a
200 µg/mL concentration on a sialoglycan microarray with approximately 64 glycan structures
20 [19]. Secondary detection was achieved with streptavidin-Cy3 and the experiment was carried out as previously described. The specificity data (Figure 11) confirm binding to all linkage variants of Neu5Ac on the array, including α2,3, α2,6 as well as to common modifications of Neu5Ac, such as 9-*O* acetylation. Further, all variants of the *N*-glycolyl form of sialic acid (Neu5Gc, Gc-Sia) were detected, with the exception of some 9-*O*-acetylated structures. No
25 signal was observed for non-sialylated terminal galactose structures, reaffirming the specificity of Sia-PS1 for sialic acid.

Western Blot using Sia-PS1, Sia-PS4, and Sia-PS5 as probes

A sample (1 µg each) of asialofetuin, fetuin, BSA, 3'SL-BSA or 6'SL-BSA was run on
30 an SDS-PAGE gel. The gel was transferred to a nitrocellulose membrane and blocked. Sia-Lectenz[®] at 1 µM each was incubated with the blot to probe for terminal sialic acid. After

washing, anti-Poly His-HRP conjugate was incubated with the blot to detect the N-terminal hexa-histidine tag of Sia-PS1. The blot was washed and developed with a chemiluminescent substrate. The image (Figure 18) shows that Sia-PS1, Sia-PS4, and Sia-PS5 bind sialylated proteins fetuin, 3' SL-BSA, and 6' SL-BSA, but not non-sialylated proteins asialofetuin and BSA. For comparison, commercial biotinylated lectins MAL II and SNA, known to be specific for Sia α Sia α 2,3- and Sia α 2,6 linkages, respectively, were used as probes. HRP streptavidin was used as the secondary detection reagent. The lectins did not bind their non-cognate BSA neoglycoprotein. This demonstrates that the Sia-PS Lectenz[®] are pan-specific, unlike the commercial lectins.

10

Summary

- 1) Computational modeling provides structural insights relevant for converting the wt NanB parent enzyme to catalytically inactive pan-specific sialic acid Lectenz[®] reagents.
- 2) All pan-specific Lectenz[®] candidates have significantly diminished or un-detected catalytic activity relative to the wt parent enzyme. Notably, the lead candidate Sia-PS1 is catalytically inactive.
- 3) The Sia-PS Lectenz[®] reagents have unique affinity, with Sia-PS1 having the highest affinity.
- 4) The Sia-PS1 Lectenz[®] is pan-specific for Sia α 2,3-, α 2,6-, and α 2,8- linkages present in 3' SL, 6' SL, and GD3 glycan (Neu5Ac α 2,8Neu5Ac α 2,3Gal β 1,4Glc) with sub-micromolar affinity for all three target glycan structures.
- 5) The Sia-PS1 Lectenz[®] is a functional capture reagent for the enrichment of fetuin from 50:50 mixtures of fetuin:asialofetuin and fetuin:BSA.
- 6) The Sia-PS1 Lectenz[®] is a functional capture reagent for the enrichment of both 3'-sialyllactose-BSA and 6'-sialyllactose-BSA glycoconjugates.
- 7) The Sia-PS1 Lectenz[®] is a superior capture reagent for the enrichment of sialoglycoconjugates compared to the lectins MAA and SNA.
- 8) The Sia-PS1 Lectenz[®] is a functional capture reagent for the enrichment of sialoglycoconjugates from a complex MCF7 breast cancer cell extract.
- 9) The Sia-PS1 Lectenz[®] exhibits binding to sialylated glycans and no binding is observed to non-sialylated terminal galactose bearing glycans via glycan array screen.

Example II. Reagents Specific for α 2,3-Linked Sialic Acid

Introduction

5 The lectins *Maackia amurensis* leucoagglutinin (MAL) and hemagglutinin (MAH) are widely believed to be specific for the terminal glycan sequence Neu5Ac α 2,3Gal β . Yet remarkably, commercial *Maackia* lectins are often sold as mixtures, have multiple names (MAA, MAL I, MAL II, etc.), and have a range of reported specificities [5]. To fully define their specificity, these lectins have recently been screened against glycan arrays at the Consortium for
10 Functional Glycomics, which confirmed that although they do bind to the Neu5Ac α 2,3Gal β sequence, they are equally able to bind to glycans terminating in other carbohydrates, including: 3S-Gal β , 3S-GalNAc β , Neu5Gc α 2,3Gal β b, Neu5Ac α 2,8Neu5Gca, and Neu5Gc α 2,8Neu5Gc α (MAL), and 3S-Gal β , KDN α 2,3Gal β , but not Neu5Gc α (MAH). Given that terminal 3S-Gal β [6] and Neu5Gc α [7] modifications are both potential markers for cancers, the ability to
15 distinguish normal Neu5Ac α 2,3Gal β sequences from others is critical. Furthermore, the presence of the normal “ α 2,3” sequence in certain glycoproteins can correlate with disease, as in the case of prostate specific antigen (PSA) [8]. In this report, we illustrate the development of an α 2,3-sialic acid specific Lectenz[®] reagent, Sia-3S1. Furthermore, we demonstrate its potential as a capture reagent for Lectenz[®] affinity chromatography applications.

20

Methods

The methods are as described in in Example I.

Results and Conclusions

25

Library construction

Using computationally-guided directed evolution (Figure 2, Example I), multiple Lectenz[®] candidates with varying specificity for 3'SL over 6'SL have been identified. Three distinct yeast surface display libraries were constructed [1] using the WT NanB enzyme as the
30 scaffold (Table 1). Specifically, the NanB enzyme was expressed as a fusion protein of Aga-2p on the yeast cell surface. These yeast libraries were constructed with site saturation mutagenesis

at positions of interest as indicated in Table 1. Library 1 includes the NanB GH33 domain (residues 230-697). Library 2 includes the NanB CBM and GH33 domains (residues 30-697). Library 3 includes the same coverage as Library 2 with a R193A mutation. Library 2 and 3 were both biopanned against a 3'SL glycan target. Colonies from rounds 3, 6, and 9 were sequenced to identify enriched clones. Enriched yeast clones from library 3 were cloned into *E. coli* expression vectors for protein expression, purification, and characterization. Library 2 clones have also been sequenced and will be analyzed in a similar manner. An amino acid iceLogo representation of the candidates selected from library 3 is illustrated in Figure 12, indicating prevalence of different amino acids at each library mutagenesis residue. These candidates all contain an Arginine 193 residue to Alanine mutation (R193A) in addition to library mutations that are identified in Table 4. This R193A mutation is located in the carbohydrate binding module of NanB and has 88% enzymatic activity of wt NanB (Figure 13). We have previously observed attenuated binding of sialo-glycoconjugates when the R193A mutation is introduced into the Sia-PS1 Lectenz[®]. Taken together, introduction of the R193A mutation to the Sia-3S Lectenz[®] library prior to selection should attenuate pan-specific sialic acid binding, with little effect on candidate enzymatic activity. This will enhance enrichment of candidates that preferentially bind to 3'SL via directed evolution.

These Sia-3S Lectenz[®] candidates are being characterized by BLI and their utility as capture reagents for enrichment of 2,3-linked sialic acid glycans is currently being assessed. A partial list of these Sia-3S Lectenz[®] is included in Table 4, along with their mutations. The amino sequences of these mutants are represented by wild-type sequence SEQ ID NO:4 in Figure 3D modified at the amino acid sites as shown in Table 4. Physical and chemical properties of these candidate reagents are shown in Table 5.

25

Table 4: Site mutations of selected 3'-linkage specific sialic acid-recognizing Lectenz[®] reagents from Library 3, based on NanB (R193A). The shaded boxes indicate retention of the wild type amino acid at that site.

| wtNanB | R193 | D270 | D327 | A538 | E541 | P660 | S673 | N683 |
|---------|------|------|------|------|------|------|------|------|
| Sia-3S1 | A | Q | D | V | D | P | S | N |
| Sia-3S2 | A | G | D | W | D | Q | A | N |
| Sia-3S3 | A | G | D | V | D | P | A | N |
| Sia-3S5 | A | D | E | F | A | P | S | N |
| Sia-3S6 | A | G | D | H | A | P | S | S |
| Sia-3S7 | A | H | D | M | A | P | S | N |
| Sia-3S8 | A | I | V | F | I | P | Q | N |
| Sia-3S9 | A | G | D | I | S | P | A | N |

- 5 Sia-3S7 additionally has mutations D427Y and L690F; however, these were not among the sites selected for mutation (based on Table 1) and, rather, are artifacts of cloning.

10 **Table 5:** Physical and chemical properties of NanB and 3'-linkage specific sialic acid recognizing Lectenz[®] reagents. ExPASy ProtParam calculated properties based on amino acid sequence are reported. [15] Molecular weight, isoelectric point, and extinction coefficients (ϵ) values are listed.

| Proteins (690 Amino Acids) | Molecular Weight | Isoelectric Point | ϵ ($M^{-1} cm^{-1}$) | ϵ ($L g^{-1} cm^{-1}$) | ϵ 1% |
|-------------------------------|---------------------|----------------------|---------------------------------|-----------------------------------|---------------|
| wtNanB | 76965.9 | 6.53 | 97640 | 1.269 | 12.69 |
| Sia-3S1 | 76907.9 | 6.53 | 97640 | 1.270 | 12.70 |
| Sia-3S2 | 76938.9 | 6.53 | 103140 | 1.341 | 13.41 |
| Sia-3S3 | 76820.8 | 6.53 | 97640 | 1.271 | 12.71 |
| Sia-3S5 | 76912.9 | 6.53 | 97640 | 1.269 | 12.69 |
| Sia-3S6 | 76803.8 | 6.70 | 97640 | 1.271 | 12.71 |
| Sia-3S7 | 76904.9 | 6.70 | 97640 | 1.270 | 12.70 |
| Sia-3S8 | 76964.1 | 6.84 | 97640 | 1.269 | 12.69 |
| Sia-3S9 | 76806.8 | 6.67 | 97640 | 1.271 | 12.71 |

Presented below is data for a promising lead Lectenz[®] candidate designated Sia-3S1.

15 **Sia-3S Lectenz[®] have no detectable enzymatic activity**

We screened these candidates against wt NanB and the R193A point mutant to assess whether the identified Sia-3S Lectenz[®] candidates exhibit neuraminidase activity. The colorimetric substrate *p*NP-Neu5Ac was used to measure activity. As shown in Figure 13, in comparison to the wild type enzyme, no activity could be measured above background for the
20 Sia-3S Lectenz[®] candidates. Therefore, all Sia-3S Lectenz[®] reagents are suitable candidates.

Sia-3S1 is specific for 3'SL over 6'SL

Sia-3S1, which is enzymatically inactive, exhibits high 3'SL affinity under well-defined conditions. Figure 14 shows its BLI binding kinetics to 3'SL and 6'SL, respectively, in two different binding buffers.

Sia-3S1 retains fetuin and is competitively eluted using 3'SL

5 Preliminary Sia-3S1 LAC was performed on fetuin, a highly sialylated glycoprotein with both 3'-linked and 6'-linked sialic acid residues, followed by competitive elution with either 3'SL or 6'SL. As shown in Figure 15, of the ~ 85 µg fetuin bound to the Sia-3S1 column, 100% was competitively eluted by 3'SL, while only 50% eluted by 6'SL. These results strongly indicate the 3'-linkage preference of Sia-3S1.

10 3'SL-BSA bound to a Sia-3S1 LAC column is competitively eluted by 3'SL but not 6'SL

A Sia-3S1 LAC column was injected with 100 µg 3'SL-BSA (Figure 16A) or 6'SL-BSA (Figure 16B) before eluting with 50 mM 3'SL or 6'SL. In Figure 16A, 3'SL is able to elute 90 µg of previously bound 3'SL-BSA. Attempting to elute bound 3'SL-BSA using 50 mM 6'SL only results in a small elution peak and the vast majority of 3'SL-BSA is eluted during the NaCl regeneration step. When 100 µg 6'SL-BSA is bound to the Sia-3S1 column (Figure 16B), both 15 50 mM 3'SL and 6'SL elute 45 µg 6'SL-BSA. Taken together, these results demonstrate Sia-3S1 is binding selectively to 3'SL-BSA over 6'SL-BSA. The remaining quantity of 6'SL-BSA is in the flow through as indicated by the shallow and broad peak from 1.5 mL through 10 mL retention volume.

20

Separation of 3'SL-BSA over 6'SL-BSA using Sia-3S1 LAC

As illustrated in Figure 17, Sia-3S1 LAC column was injected with 50 µg 3'SL-BSA or 6'SL-BSA in 10 mM EPPS (pH 7.5) binding buffer containing 110 mM NaCl. Under these conditions, all 50 µg of 6'SL-BSA was detected in the flow through, whereas only 5 µg 3'SL- 25 BSA did not bind to the column. Subsequently, bound 3'SL-BSA was eluted using 10 mM EPPS (pH 7.5) buffer with 1 M NaCl. No 6'SL-BSA elution could be detected using the same EPPS buffer with 1 M NaCl. These results suggest that adjusting the NaCl concentration in binding buffer can result in 3'SL-BSA binding to the Sia-3S1 LAC column, whereas 6'SL-BSA does not.

30

Western Blot using Sia-3S1

A sample (1 μ g each) of asialofetuin, fetuin, BSA, 3'SL-BSA or 6'SL-BSA was run on an SDS-PAGE gel. The gel was transferred to a nitrocellulose membrane and blocked. 1 μ M Sia-3S1 was incubated with the blot to probe for Neu5Ac α 2,3Gal. After washing, anti-Poly His-
5 HRP conjugate was incubated with the blot to detect the N-terminal hexa-histidine tag of Sia-3S1. The blot was washed and developed with a chemiluminescent substrate. The image (Figure 18) shows that Sia-3S1 binds only proteins with α 2,3-linked sialylation: fetuin and 3'SL-BSA. (A faint band for asialofetuin may be due to trace amounts of fetuin in the commercial asialofetuin.) It does not bind BSA and, significantly, does not bind 6'SL-BSA. This is in
10 contrast to the "Sia α 2,3 specific" MAL II lectin, which shows no binding, except for a very faint band for fetuin. A band would be predicted for 3'SL-BSA, but none was observed.

Summary

- 1) Sia-3S1 Lectenz[®] has preferential affinity for 3'SL over 6'SL.
- 15 2) The Sia-3S1 Lectenz[®] is a functional capture reagent for the enrichment of 2,3-linked sialoglycoproteins: more fetuin was eluted when using 3'SL for competitive elution than 6'SL as a competitive elution reagent.
- 3) Any 6'SL binding to Sia-3S1 Lectenz[®] on LAC is eliminated using a buffer with well-defined NaCl concentration.
- 20 4) Western blotting demonstrates that Sia-3S1 is a molecular probe specific to α 2,3-linked sialic acid glycoconjugates that is far superior to commercial MAL II.

REFERENCES CITED

- 25 1. Chao, G., et al., *Isolating and engineering human antibodies using yeast surface display*. Nat. Protocols, 2006. **1**(2): p. 755-768.
2. Zhao, J., et al., *Comparative serum glycoproteomics using lectin selected sialic acid glycoproteins with mass spectrometric analysis: application to pancreatic cancer serum*. J Proteome Res, 2006. **5**(7): p. 1792-802.
- 30 3. Bai, X., et al., *Enhanced 3-O-sulfation of galactose in Asn-linked glycans and Maackia amurensis lectin binding in a new Chinese hamster ovary cell line*. Glycobiology, 2001. **11**(8): p. 621-32.
4. Nicholls, J.M., et al., *Sialic acid receptor detection in the human respiratory tract: evidence for widespread distribution of potential binding sites for human and avian influenza viruses*. Respir Res, 2007. **8**: p. 73.
- 35

5. Geisler, C. and D.L. Jarvis, *Effective glycoanalysis with Maackia amurensis lectins requires a clear understanding of their binding specificities*. Glycobiology, 2011. **21**(8): p. 988-93.
6. Zheng, J., et al., *Serum 3'-sulfo-Lea indication of gastric cancer metastasis*. Clin Chim Acta, 2009. **405**(1-2): p. 119-26.
7. Samraj, A.N., et al., *Involvement of a Non-Human Sialic Acid in Human Cancer*. Front Oncol, 2014. **4**: p. 33.
8. Tajiri, M., C. Ohyama, and Y. Wada, *Oligosaccharide profiles of the prostate specific antigen in free and complexed forms from the prostate cancer patient serum and in seminal plasma: a glycopeptide approach*. Glycobiology, 2008. **18**(1): p. 2-8.
9. Gut, H., S.J. King, and M.A. Walsh, *Structural and functional studies of Streptococcus pneumoniae neuraminidase B: An intramolecular trans-sialidase*. FEBS Lett, 2008. **582**(23-24): p. 3348-52.
10. Ford, M.G., et al., *MD Simulations of Galectin-1 Oligosaccharide Complexes Reveal the Molecular Basis for Ligand Diversity*. Proteins: Struct. Funct. Genet., 2003. **53**: p. 229-240.
11. Martin, J.C., et al., *Defining the Structural Origin of the Substrate Sequence Independence of O-GlcNAcase Using a Combination of Molecular Docking and Dynamics Simulation*. Glycobiology, 2013.
12. Kadirvelraj, R., et al., *Structure and Binding Analysis of Polyporus squamosus Lectin in Complex with the Neu5Aca2-6Gal β 1-4GlcNAc Human-type Influenza Receptor*. Glycobiology, 2011. **21**(7): p. 973-984.
13. Woods, R.J. and M.B. Tessier, *Computational Glycoscience: Characterizing the Spatial and Temporal Properties of Glycans and Glycan-Protein Complexes*. Curr. Opin. Struct. Biol., 2010. **20**: p. 575-583.
14. Hadden, J.A., et al., *Calculating binding free energies for protein-carbohydrate complexes*. Method. Mol. Biol., 2015. **1273**: p. 431-65.
15. Gasteiger E., H.C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A., *Protein Identification and Analysis Tools on the ExPASy Server*, in *The Proteomics Protocols Handbook* J.M. Walker, Editor. 2005, Copyright Humana Press. p. 571-607.
16. Haseley, S.R., et al., *Characterization of the carbohydrate binding specificity and kinetic parameters of lectins by using surface plasmon resonance*. Anal Biochem, 1999. **274**(2): p. 203-10.
17. Yamamoto, K., Y. Konami, and T. Irimura, *Sialic acid-binding motif of Maackia amurensis lectins*. J. Biol. Chem., 1997. **121**(4): p. 756-61.
18. Lee, L.Y., et al., *An optimized approach for enrichment of glycoproteins from cell culture lysates using native multi-lectin affinity chromatography*. J Sep Sci, 2012. **35**(18): p. 2445-52.
19. Padler-Karavani, V., et al., *Cross-comparison of Protein Recognition of Sialic Acid Diversity on Two Novel Sialoglycan Microarrays*. Journal of Biological Chemistry, 2012. **287**(27): p. 22593-22608.

The complete disclosure of all patents, patent applications, and publications, and electronically available material (including, for instance, nucleotide sequence submissions in, e.g., GenBank and RefSeq, and amino acid sequence submissions in, e.g., SwissProt, PIR, PRF,

PDB, and translations from annotated coding regions in GenBank and RefSeq) cited herein are incorporated by reference. In the event that any inconsistency exists between the disclosure of the present application and the disclosure(s) of any document incorporated herein by reference, the disclosure of the present application shall govern. The foregoing detailed description and
5 examples have been given for clarity of understanding only. No unnecessary limitations are to be understood therefrom. The invention is not limited to the exact details shown and described, for variations obvious to one skilled in the art will be included within the invention defined by the claims.

10 All headings are for the convenience of the reader and should not be used to limit the meaning of the text that follows the heading, unless so specified.

WHAT IS CLAIMED IS:

1. A sialic acid-recognizing affinity reagent comprising a catalytically inactive NanB neuraminidase protein, or fragment thereof, having at least one amino acid mutation compared to a corresponding wild-type NanB neuraminidase protein, wherein the mutation
 - (a) reduces or eliminates the neuraminidase activity of the NanB protein; and
 - (b) affects sialic acid binding affinity or binding specificity;and wherein said affinity reagent binds to a sialic acid component of a glycan.
2. The sialic acid-recognizing affinity reagent of claim 1, wherein the NanB fragment comprises at least one of the carbohydrate binding module (CBM) domain of the NanB protein and the glycosyl hydrolyase (GH) domain of the NanB protein.
3. The sialic acid-recognizing affinity reagent of claim 1 or 2, wherein the affinity reagent is pan-specific for sialic acid.
4. The sialic acid-recognizing affinity reagent of claim 3, wherein the affinity reagent binds to Neu5Ac linked to an adjacent saccharide monomer in an α 2,3 linkage, Neu5Ac linked to an adjacent saccharide monomer in an α 2,6 linkage, and Neu5Ac linked to an adjacent saccharide monomer in an α 2,8 linkage.
5. The sialic acid-recognizing affinity reagent of claim 3 which binds to at least one variant of Neu5Ac.
6. The sialic acid-recognizing affinity reagent of claim 3, wherein the mutation is at a site selected from the group consisting of Y653, D270, and E541, as represented in SEQ ID NO:1 or SEQ ID NO:4.
7. The sialic acid-recognizing affinity reagent of claim 6, wherein the mutation is in *S. pneumoniae* NanB (SEQ ID NO:1 or SEQ ID NO:4) or fragment thereof, or a corresponding position in a homologous NanB sequence.

8. The sialic acid-recognizing affinity reagent of claim 6, wherein the mutation is selected from the group consisting of Y653F, D270A, D270N, E541A, and E541Q.
9. The sialic acid-recognizing affinity reagent of claim 8, wherein the mutation is Y653F.
10. The sialic acid-recognizing affinity reagent of claim 8 selected from the group consisting of Sia-PS1, Sia-PS2, Sia-PS3, Sia-PS4, and Sia-PS5.
11. The sialic acid-recognizing affinity reagent of claim 10 which is Sia-PS1.
12. The sialic acid-recognizing affinity reagent of claim 1, wherein the affinity reagent has a plurality of amino acid mutations compared to a corresponding wild-type NanB neuraminidase protein.
13. The sialic acid-recognizing affinity reagent of claim 12, wherein said plurality of mutations comprises:
 - (a) at least one first mutation that reduces or eliminates the catalytic activity of the NanB protein; and
 - (b) at least one second mutation that affects binding affinity or binding specificity.
14. The sialic acid-recognizing affinity reagent of claim 12, wherein the affinity reagent is specific for Neu5Ac that is linked to an adjacent saccharide monomer in an α 2,3 linkage.
15. The sialic acid-recognizing affinity reagent of claim 12, wherein the plurality of mutations are at sites selected from the group consisting of R193, D270, A538, E541, P660, S673, D327, and N683, as represented in SEQ ID NO:1 or SEQ ID NO:4.
16. The sialic acid-recognizing affinity reagent of claim 15, wherein the plurality of mutations are in *S. pneumoniae* NanB (SEQ ID NO:1 or SEQ ID NO:4) or fragment thereof, or a corresponding position in a homologous NanB sequence.

17. The sialic acid-recognizing affinity reagent of claim 15, wherein the plurality of mutations are selected from the group consisting of R193A, D270Q, D270G, D270H, D270I, A538V, A538W, A538F, A538H, A538M, A538I, E541D, E541A, E541I, E541S, P660Q, S673A, S673Q, D327E, D327V, and N683S.
18. The sialic acid-recognizing affinity reagent of claim 17, wherein the plurality of mutations comprise R193A, D270Q, A538V, and E541D.
19. The sialic acid-recognizing affinity reagent of claim 17 selected from the group consisting of Sia-3S1, Sia-3S2, Sia-3S3, Sia-3S5, Sia-3S6, Sia-3S7, Sia-3S7*, Sia-3S8, and Sia-3S9.
20. The sialic acid-recognizing affinity reagent of claim 19 which is Sia-PS1.
21. The sialic acid-recognizing affinity reagent of claim 1, wherein the glycan is a constituent of a glycosylated biomolecule.
22. The sialic acid-recognizing affinity reagent of claim 21, wherein the glycosylated biomolecule comprises a glycoprotein, a glycopeptide, a glycolipid, a glycolipoprotein, or a glycolipopeptide.
23. A conjugate comprising a first component comprising a sialic acid-recognizing affinity reagent of claim 1, covalently linked to a second component.
24. The conjugate of claim 23, wherein the second component is a proteinaceous component.
25. The conjugate of claim 23, wherein the second component is a nonproteinaceous component.
26. The conjugate of claim 23, wherein the second component is a therapeutic or diagnostic agent.

27. A fusion protein comprising a sialic acid-recognizing affinity reagent of claim 1.
28. An affinity matrix comprising a sialic acid-recognizing affinity reagent, conjugate or fusion protein of claim 1, 23, or 27.
29. The affinity matrix of claim 28 selected from the group consisting of a solid support, surface, column, resin, bead, microarray, particle and nanoparticle.
30. A kit comprising a sialic acid-recognizing affinity reagent, conjugate, fusion protein or affinity matrix of any of the preceding claims, and instructions for use.
31. An isolated polynucleotide encoding a sialic acid-recognizing affinity reagent, conjugate or fusion protein of claim 1, 23, or 27.
32. A vector comprising a polynucleotide of claim 31.
33. The vector of claim 32, wherein the vector is an expression vector.
34. A host cell comprising a vector of claim 32 or 33.
35. The host cell of claim 34, wherein the host cell is a bacterial cell.
36. A method for making the sialic acid-recognizing affinity reagent or conjugate of any of claims 1-24, or the fusion protein of claim 27, the method comprising expressing the affinity reagent, conjugate or fusion protein in host cell.
37. A method for detecting a sialic acid component of a glycan, the method comprising:
contacting a biological or laboratory sample with a sialic acid-recognizing affinity reagent, conjugate or fusion protein of claim 1, 23, or 27 under conditions to allow binding of the affinity reagent, conjugate or fusion protein to the sialic acid; and detecting the sialic acid.

38. The method of claim 37, wherein the biological or laboratory sample comprises a recombinant antibody.

39. The method of claim 37, wherein the glycan is a biomarker.

40. The method of claim 39, wherein the biomarker is a cancer biomarker.

41. A method for enriching, isolating or purifying a sialic acid-containing glycan, the method comprising:

contacting a sialic acid-recognizing affinity reagent, conjugate or fusion protein of claim 1, 23, or 27 under conditions to allow binding of the affinity reagent, conjugate or fusion protein to the sialic acid so as to yield an enriched, isolated or purified sialic acid-containing glycan.

42. A diagnostic or therapeutic composition comprising a sialic acid-recognizing affinity reagent, conjugate or fusion protein of claim 1, 23, or 27.

43. The diagnostic or therapeutic composition of claim 42 wherein the sialic acid-recognizing affinity reagent, conjugate or fusion protein is detectably labeled.

44. The diagnostic or therapeutic composition of claim 43 wherein the detectable label comprises a radioactive label, a fluorescent label, a phosphorescent label, a colorimetric label, an enzymatic label, an immunological label, a magnetic label, a paramagnetic label, a diamagnetic label or an electromagnetic label.

45. Use of a sialic acid-recognizing affinity reagent, conjugate or fusion protein of claim 1, 23, or 27 as a therapeutic agent.

46. Use of a sialic acid-recognizing affinity reagent, conjugate or fusion protein of claim 1, 23, or 27 as a diagnostic agent.

47. Use of a sialic acid-recognizing affinity reagent, conjugate or fusion protein of claim 1, 23, or 27 for targeted drug delivery.

48. Use of a sialic acid-recognizing affinity reagent, conjugate or fusion protein of claim 1, 23, or 27 for detection of the presence or amount of sialic acid in a biological or laboratory sample.

49. A compound, composition, or method including one or more of the features described herein.

FIG. 1B

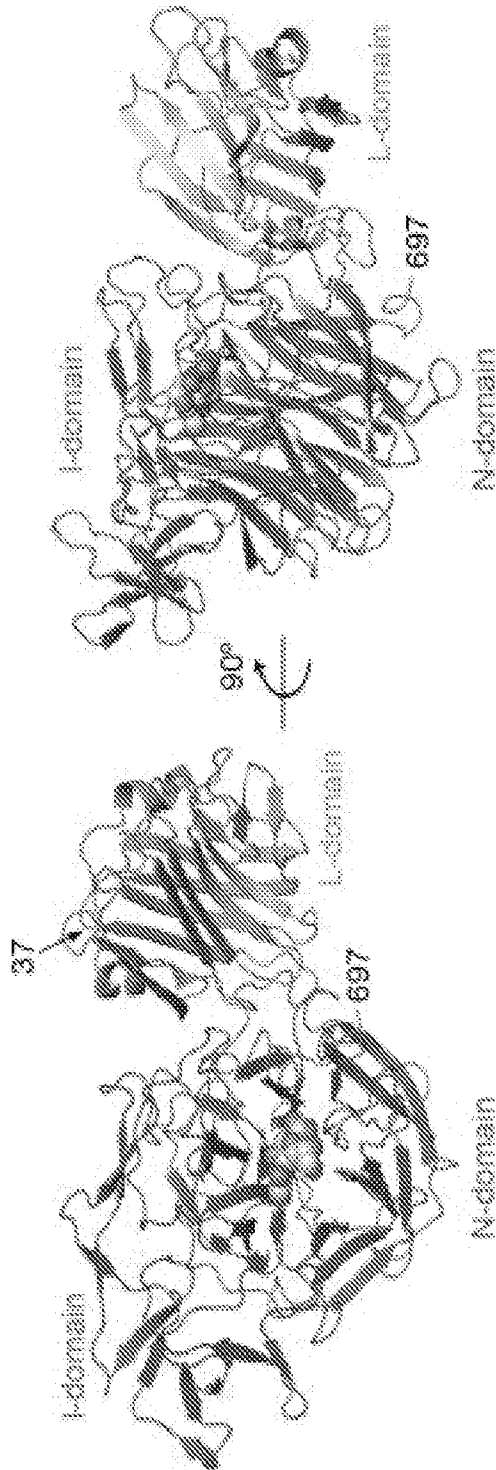


FIG. 1A

FIG. 2

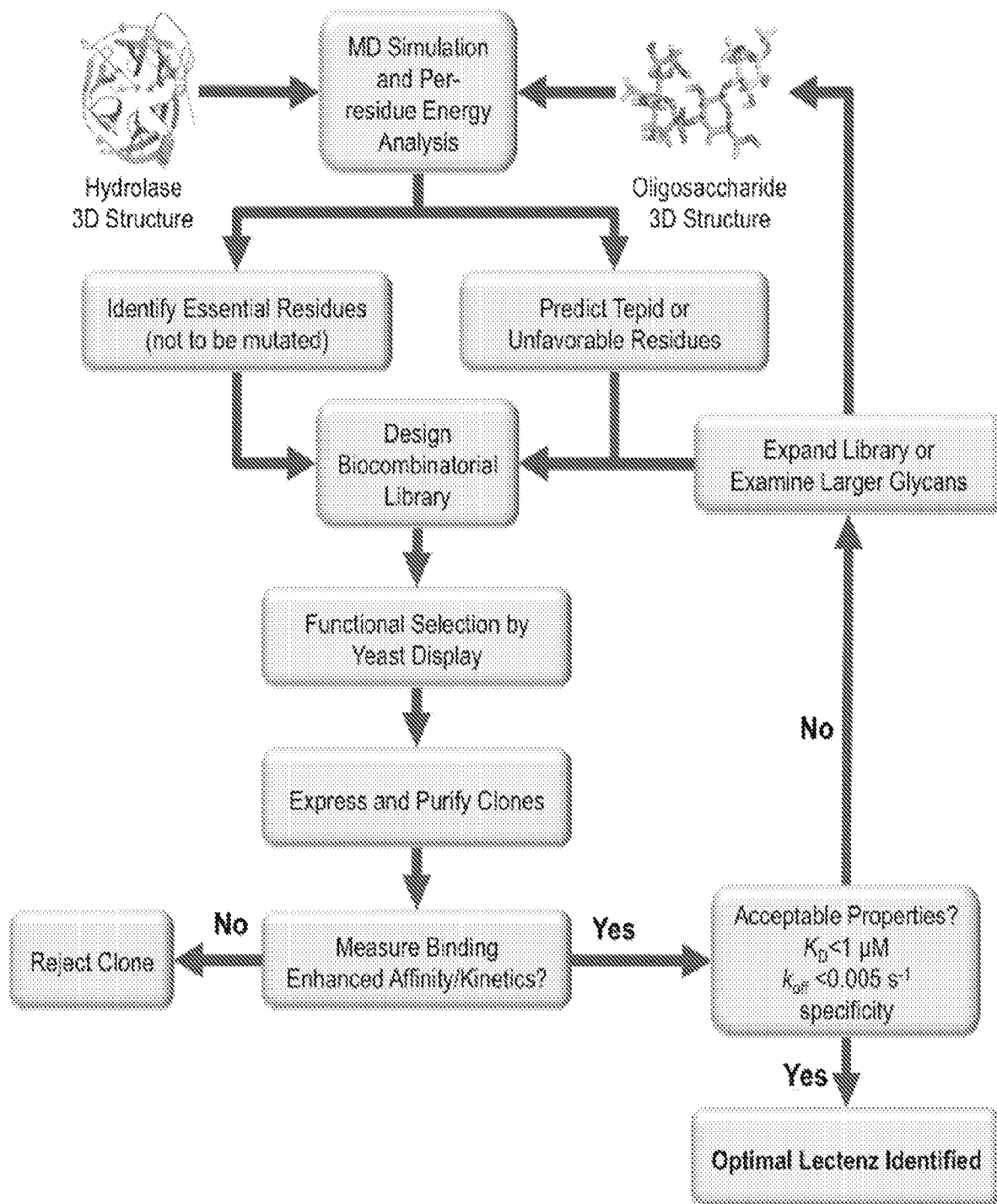


FIG. 3A

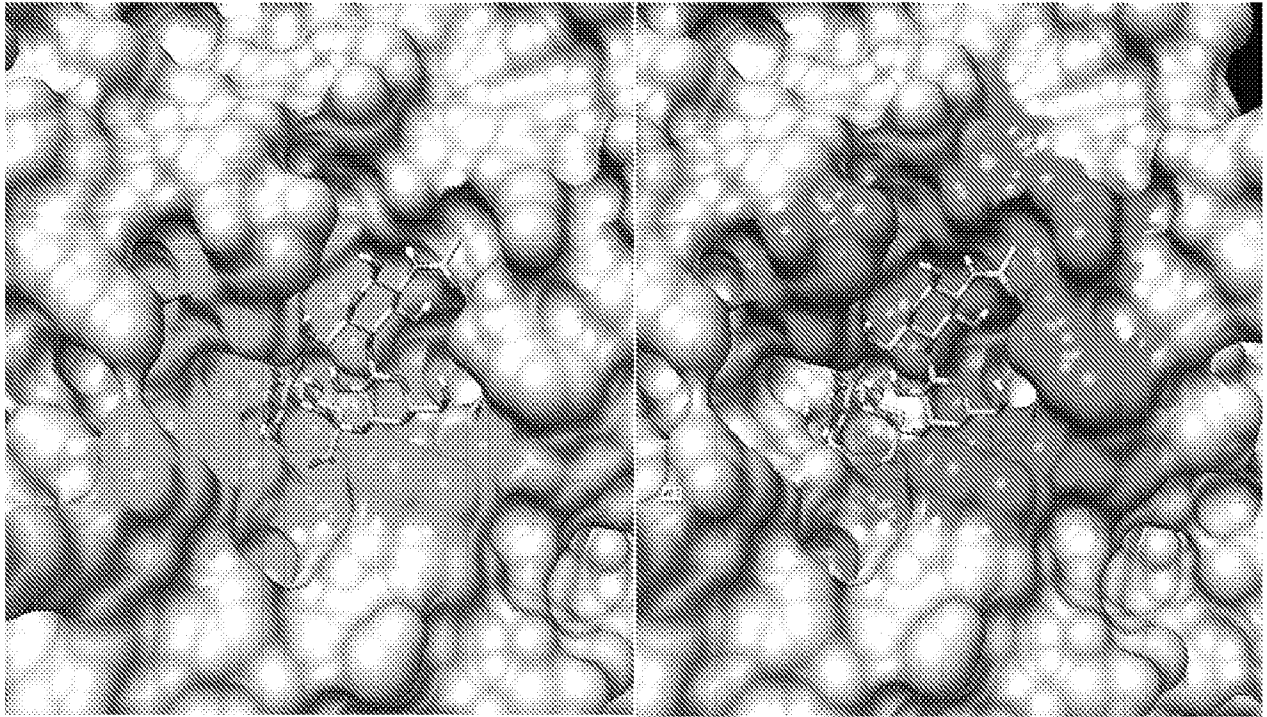


FIG. 3B

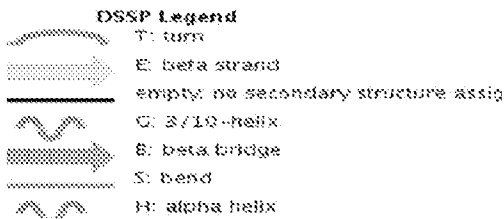
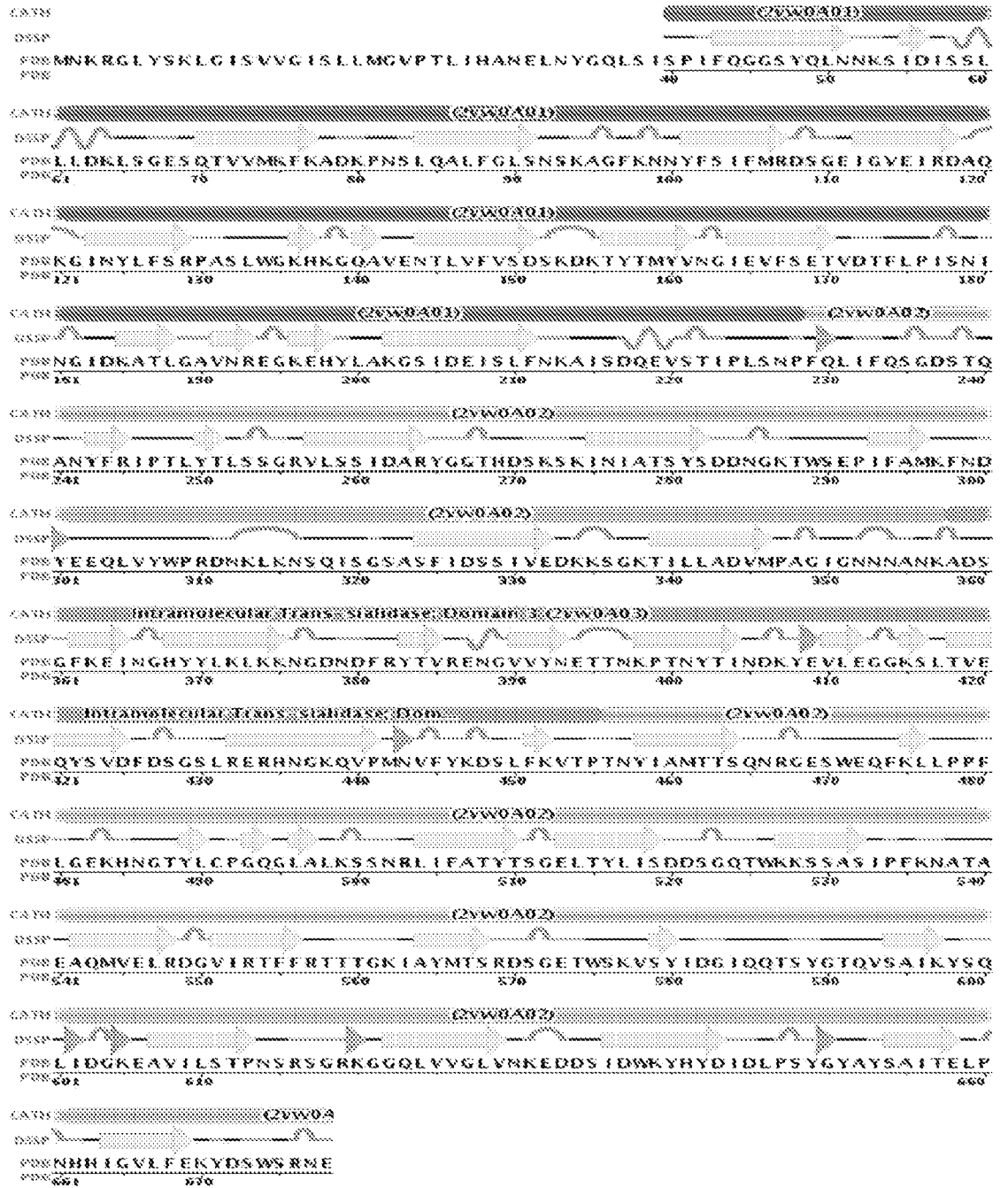


FIG. 3C

```

2VVZ:A|PDBID|CHAIN|SEQUENCE ----- 0
2VW0:A|PDBID|CHAIN|SEQUENCE MNKRGLYSRLGISVVGISLLMGVPTLIHANELNYGQLSISPIFQGGSYQLNKRKSIDISL 60
5F9T:A|PDBID|CHAIN|SEQUENCE -----ETPVLEKNNVTLTGGGENVTKE 22

2VVZ:A|PDBID|CHAIN|SEQUENCE ----- 0
2VW0:A|PDBID|CHAIN|SEQUENCE LLDKLSGESQTVVMKFKADKPNLSLQALFGLSNSKAGFKNNYFSIFMRD3SGEIGVEIRDAQ 120
5F9T:A|PDBID|CHAIN|SEQUENCE LKDKFTSGEDFTVVIKYNQSSSEKGLQALFGISNSKPGQONSIVTVFLRDNNGELGMEARDTS 82

2VVZ:A|PDBID|CHAIN|SEQUENCE ----- 0
2VW0:A|PDBID|CHAIN|SEQUENCE KGINYLFSRPASLWGXKRGQAVENTLVFVSDSKDKTYTMYVNGIEVFSSETVDTLFLPISNI 180
5F9T:A|PDBID|CHAIN|SEQUENCE SNKNNLVSRPASVWGXKYQEAVENTVAVVADSVKKTYSLYANGTKVVEKKVDNPLNIKDI 142

2VVZ:A|PDBID|CHAIN|SEQUENCE -----EGGAL-----TEKTDIFESGRNG 18
2VW0:A|PDBID|CHAIN|SEQUENCE NGIDKATLGAVNREGKEHYLAKSISIDEISLFNKAISDQEVSTIFLS-NPFQLIFQSG----- 236
5F9T:A|PDBID|CHAIN|SEQUENCE KGLDYVYMLGGVVKPAGKTAFGFNGTLENIKFFNSALDEETVKKMTTNAVTFGLIYNTAN---- 199
      . * : * : .

2VVZ:A|PDBID|CHAIN|SEQUENCE KPNKDGIKSYRIPALLKTDKGTLIAGADEPRLHS SDW-GDIGMVIIRS EDNGKRTWGD RVT 77
2VW0:A|PDBID|CHAIN|SEQUENCE --DSTQANYFRIPPLYTLSSGVLSSIDARYGGTHDSKSKINIATSYSDDNGKTFWSEPIF 294
5F9T:A|PDBID|CHAIN|SEQUENCE --DTTGSNYFRIPVLYTFSNGEVFSSIDARYGGTHDFLNKINIATSYSDDNGKTTWKPKL 257
      . : : : * * . . . * : : . * * : * . . * . : * : * * * * .

2VVZ:A|PDBID|CHAIN|SEQUENCE ITNLPD-----NPKASDEISGSPVNIEMVLVQDRETFRIFSIYDMFPEGKGI 126
2VW0:A|PDBID|CHAIN|SEQUENCE AMKFNDEEYEQLVYWPDKLKNSSQISGSASFIDSIVEDKKSCKTILLADVMPAGIGNN-- 353
5F9T:A|PDBID|CHAIN|SEQUENCE TLAFFDFAPVLEWPREVGGRLQISGGATYIDSVIVEKK-NKQVLMFADVMPAGVSR- 315
      : * . * . * * : * : . : : : * : * * .

2VVZ:A|PDBID|CHAIN|SEQUENCE MSSQKEEAYKKIDKRTYQILYREGK-GAYTIRENGTVYTPDGKA-TDYPVVV----- 177
2VW0:A|PDBID|CHAIN|SEQUENCE NANKADSGFRKINGHYLKLKKNNGDNDFRYTVRENGVYVNETNKPTNYTINDHYEVLEG 413
5F9T:A|PDBID|CHAIN|SEQUENCE EATRKDSGNKQIDGNYYLKLKPKQGDYVNYTIRENGTVYDDPTNRPTFFSVKKNFGIKQN 375
      . : : : * : * * : * * : * : . * * : * * * * * : * : :

2VVZ:A|PDBID|CHAIN|SEQUENCE -----DP--VKPAYS DKGDLKGNQLLGNIFYTTNKTSFPFIADKSYLMMSSYSDDDGKT 229
2VW0:A|PDBID|CHAIN|SEQUENCE GKS LTV EQYSVDFDSGSLRERHNGEQVPMNVFY---KDSLFKVTPTNYIAMTTSQNRGES 470
5F9T:A|PDBID|CHAIN|SEQUENCE GNYLTV EQYSVDFENNKKTEYRNGTKVHMNIIFY---KDALPKVPTNYIAYISNDHGES 432
      : * . . . : : * : * * : * : * : * : . * : * : * :

2VVZ:A|PDBID|CHAIN|SEQUENCE WSAFQDITPMKADWMKFLVSGPGTGI VLRNGPHKGRILIPYTTNNVSHLNGSQSSRII 299
2VW0:A|PDBID|CHAIN|SEQUENCE WEQFKLLPPFLGKKN-GTYLCPGGLALK---SNRLIEATYTSGE-----LTYL 517
5F9T:A|PDBID|CHAIN|SEQUENCE WSAPTLLPPTINGLNRN-APYLGPRGRIIES---STGRILIPSYTKE-----SAFI 479
      * . : * : . : * * * : . . : * : * :

2VVZ:A|PDBID|CHAIN|SEQUENCE YSDDRGKTWHAGEAVNDNRQVDGQKIHSSSTMNRRQAQNTSESTVVQLNNGDVKLFMRGLTS 349
2VW0:A|PDBID|CHAIN|SEQUENCE ISDSDGQGTWAKSSASIPF-----KNATAEAQMVLRDGVIRTFPTTTF 561
5F9T:A|PDBID|CHAIN|SEQUENCE YSDENGASWKVVP LPS-----S-WSAEAQFVELSPGVIQAYMRTNNG 522
      * * * * * : * : . . : * : * * * * * : * : : * * .

2VVZ:A|PDBID|CHAIN|SEQUENCE DLQVATSKDGGVTWERDKRYP-QVKDVYVQMSA---IHTMHEGKEYIILSNAGGP-KREN 405
2VW0:A|PDBID|CHAIN|SEQUENCE KIANMTSRDSGETWSKVSYIDGICQTSGYQVSAIKYSQLIDGKEAVILSTPNSRSRGRK 621
5F9T:A|PDBID|CHAIN|SEQUENCE KIA YLT SKDAGTTWSAPEYLKPVSNPSYGTQLSINYSQLIDGKKAVIDLSTPNSRNGRKH 582
      . : * * : * * * . . . . * : * : : * * : * * . . * :

2VVZ:A|PDBID|CHAIN|SEQUENCE GMVHLARV-EENGELTWLKHNP IQG---EFAYNSLQELNGEYGIIVHTEKGN-AYT- 460
2VW0:A|PDBID|CHAIN|SEQUENCE GQLVVLVKNKEDDSIDWKYHYDIDLPSYGYAISAITELPNHHIGVLFEKYDWSRPNELHL 691
5F9T:A|PDBID|CHAIN|SEQUENCE GQIWLGLIN-DNTIDWRVHHVDVDSNYGYSYSTLTELPHHEIGLNFERFDSWSRPNELHM 641
      * : : . : . : * * : : : * : * : * * * : * : * : . .

2VVZ:A|PDBID|CHAIN|SEQUENCE ----LSFPKFNWDFLSNDLISPTEARVHP TREMGKGVIGLEPDSSEVLV 504
2VW0:A|PDBID|CHAIN|SEQUENCE SNVVQYIDLEINDLTK----- 697
5F9T:A|PDBID|CHAIN|SEQUENCE KNVVPYITFRIEDLRKNL----- 659
      : : : : * . *

```

FIG. 3D

```

2VW0:A|PDBID|CHAIN|SEQUENCE      MNKRGLYSKLGLISVVGISLLMGVPTLIHANELNYGQLSISPIFQGGSYQLNPKSIDISSL 60
NanB:AA30-697 (pET28a)          -----NELNYGQLSISPIFQGGSYQLNPKSIDISSL 31
                                   *****

2VW0:A|PDBID|CHAIN|SEQUENCE      LLDHLSGESQTVVMKFKADKPNLSLQALFGLSNKAGFKNNYFSIFMRDSSGEIGVEIRDAQ 120
NanB:AA30-697 (pET28a)          LLDKLSGESQTVVMKFKADKPNLSLQALFGLSNKAGFKNNYFSIFMRDSSGEIGVEIRDAQ 91
                                   *****

2VW0:A|PDBID|CHAIN|SEQUENCE      KGINYLFSRPASLWGRKHKQAVENTLVFVSDSKDKTYTMVNGIEVFSSETVDTFLPI SNI 180
NanB:AA30-697 (pET28a)          KGINYLFSRPASLWGRKHKQAVENTLVFVSDSKDKTYTMVNGIEVFSSETVDTFLPI SNI 151
                                   *****

2VW0:A|PDBID|CHAIN|SEQUENCE      NGIDKATLGAVNREGKEHYLAHGSIDEI SLENKAI SDQEVSTIPLSNFQQLIFQSGDSTQ 240
NanB:AA30-697 (pET28a)          NGIDKATLGAVNREGKEHYLAHGSIDEI SLENKAI SDQEVSTIPLSNFQQLIFQSGDSTQ 211
                                   *****

2VW0:A|PDBID|CHAIN|SEQUENCE      ANYFPIPTLYTLSSGRVLSSIDARYGGTWDKSKKINIATSYSDDNGHTWSEPIFAMKFN 300
NanB:AA30-697 (pET28a)          ANYFRIPTLYTLSSGRVLSSIDARYGGTWDKSKKINIATSYSDDNGHTWSEPIFAMKFN 271
                                   *****

2VW0:A|PDBID|CHAIN|SEQUENCE      YEEQLVYWPRDNKLNKLSQISGSASFIDSSIVEDKKS GKTILLADVMEAGIGNNMANKADS 360
NanB:AA30-697 (pET28a)          YEEQLVYWPRDNKLNKLSQISGSASFIDSSIVEDKKS GKTILLADVMEAGIGNNMANKADS 331
                                   *****

2VW0:A|PDBID|CHAIN|SEQUENCE      GFKEINGHYLLKLNKGDNDFFRYTVRENGVVYNETTNKPTNYTINDKYEVLGGKSLTVE 420
NanB:AA30-697 (pET28a)          GFKEINGRYLLKLNKGDNDFFRYTVRENGVVYNETTNKPTNYTINDKYEVLGGKSLTVE 391
                                   *****

2VW0:A|PDBID|CHAIN|SEQUENCE      QYSVDFDSGSLRERHNGKQVPMNVFYKDSL FKVTFETNYIAMTTSQNRGESWEQFKLLPPF 480
NanB:AA30-697 (pET28a)          QYSVDFDSGSLRERHNGKQVPMNVFYKDSL FKVTFETNYIAMTTSQNRGESWEQFKLLPPF 451
                                   *****

2VW0:A|PDBID|CHAIN|SEQUENCE      LGEKHNGTYLCPGQGLALKSNRLIFATYTS GELTYLISDSSGQTWKSSASIPFKNATA 540
NanB:AA30-697 (pET28a)          LGEKHNGTYLCPGQGLALKSNRLIFATYTS GELTYLISDSSGQTWKSSASIPFKNATA 511
                                   *****

2VW0:A|PDBID|CHAIN|SEQUENCE      EAQMVELRDGVIRTFRTTTTGKIAYMTSKDSGETWSKVS YIDGIIQQTSYGTQVSAIKYSQ 600
NanB:AA30-697 (pET28a)          EAQMVELRDGVIPTRTFRTTTTGKIAYMTSKDSGETWSKVS YIDGIIQQTSYGTQVSAIKYSQ 571
                                   *****

2VW0:A|PDBID|CHAIN|SEQUENCE      LLDGKEAVILSTPNRSRSGRKGQLVVGVLV NKEDDSIDWKYHYDIDLPSYGYAYSAITELP 660
NanB:AA30-697 (pET28a)          LLDGKEAVILSTPNRSRSGRKGQLVVGVLV NKEDDSIDWKYHYDIDLPSYGYAYSAITELP 631
                                   *****

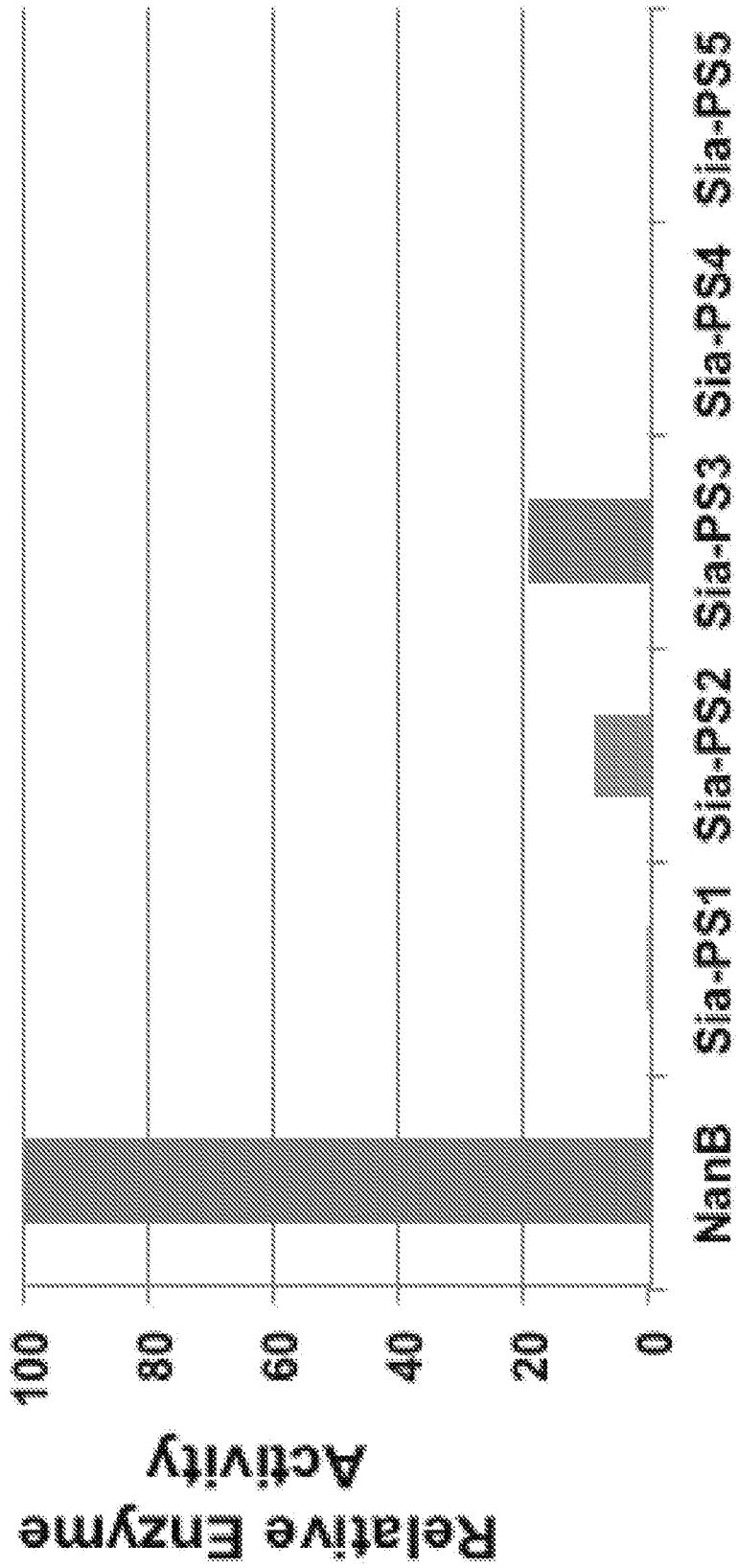
2VW0:A|PDBID|CHAIN|SEQUENCE      NHHIGVLFEKYDSWSRNLHLNSNVVQYI DLEINDLTK          697
NanB:AA30-697 (pET28a)          NHHIGVLFEKYDSWSRNLHLNSNVVQYI DLEINDLTK          668
                                   *****

```

Alignments made using Clustal Omega Multiple Sequence Alignment web tool:
<https://www.ebi.ac.uk/Tools/msa/clustalo/>

PDB = Research Collaboratory for Structural Bioinformatics Protein Database:
<http://www.rcsb.org/pdb/home/home.do>

FIG. 4



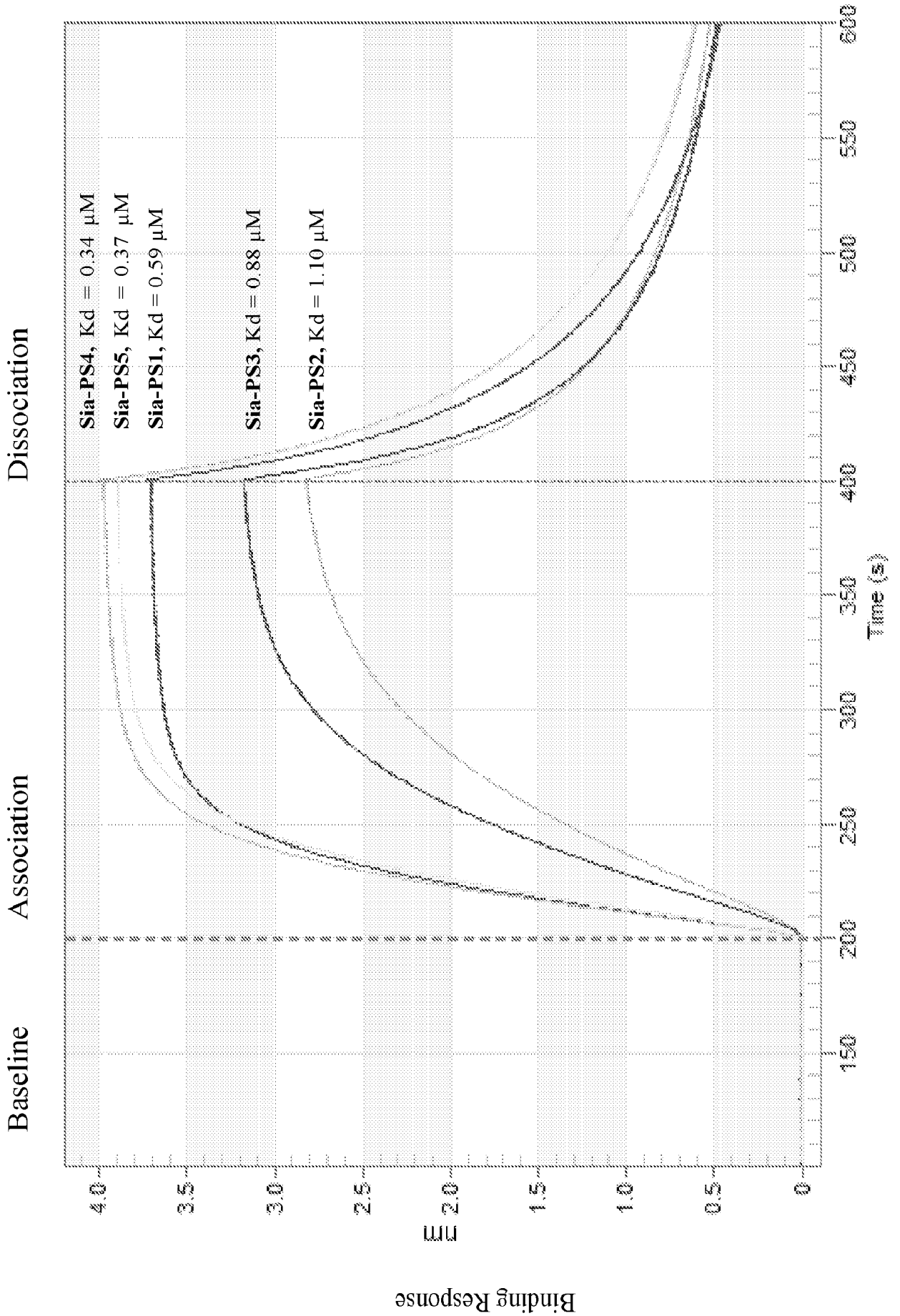
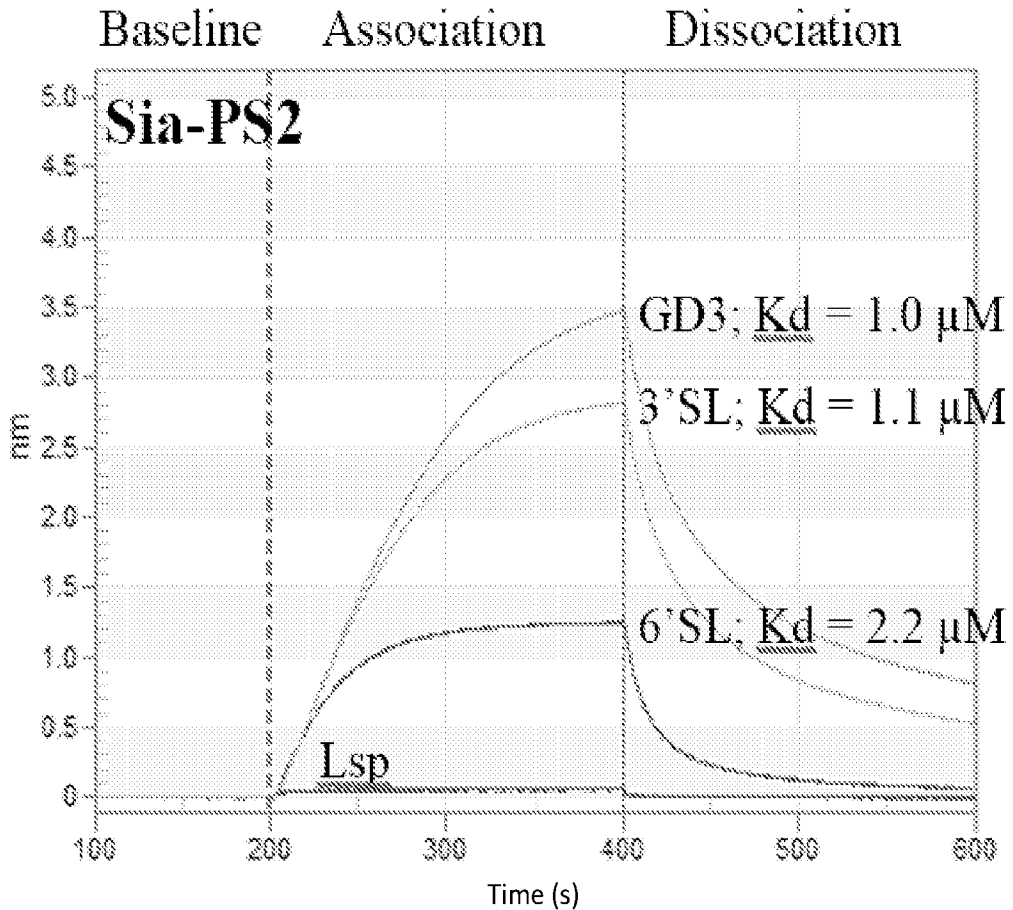
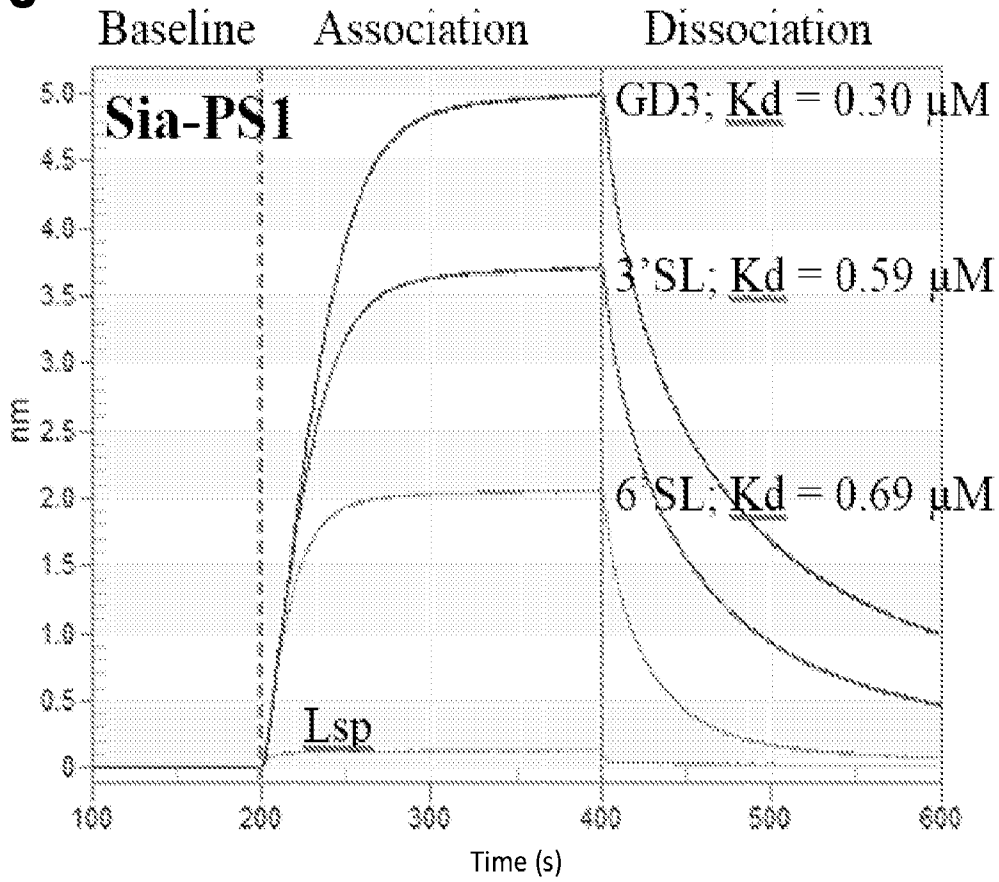


FIG. 5

9/23

FIG. 6



10/23

FIG. 6(Cont.)

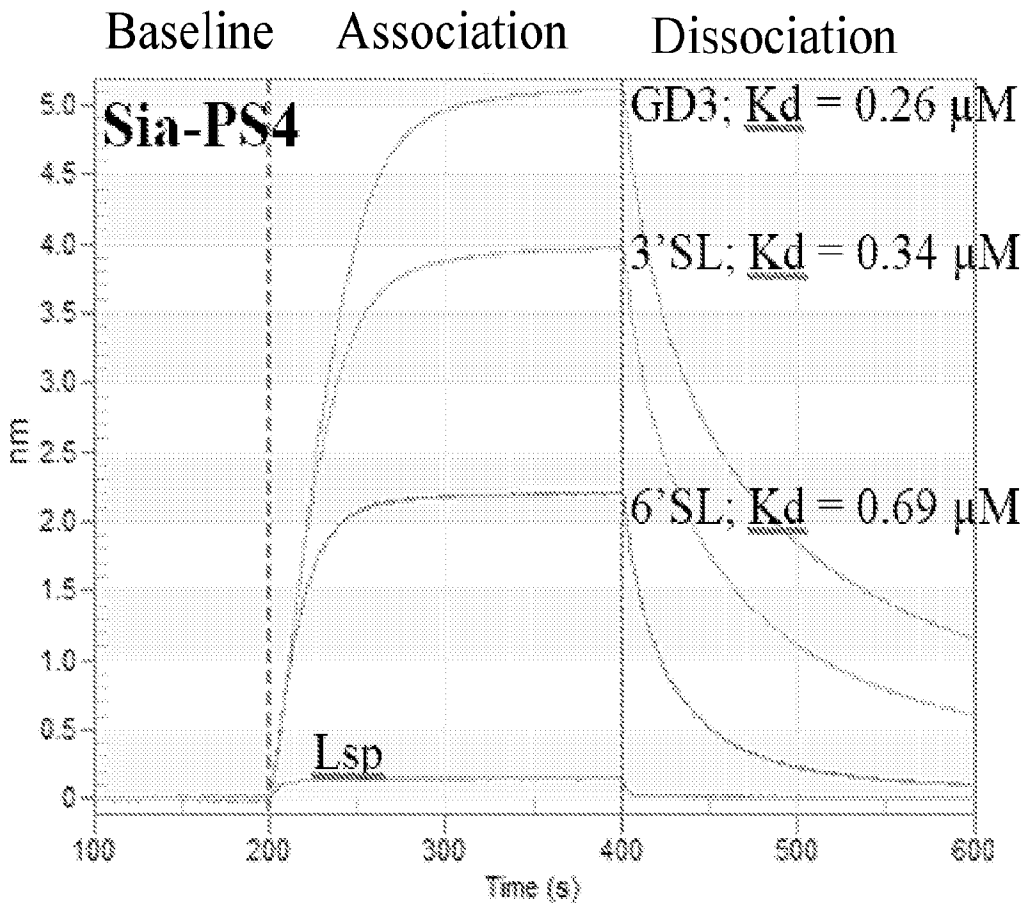
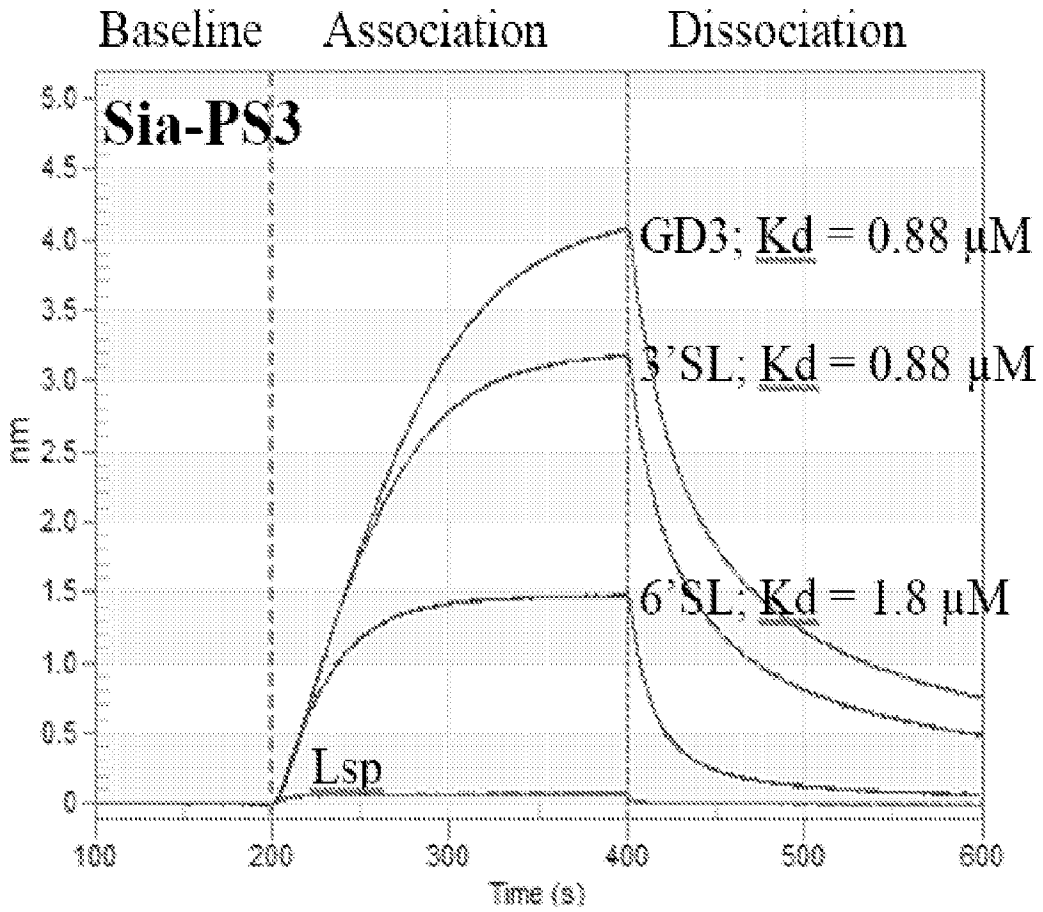


FIG. 6(Cont.)

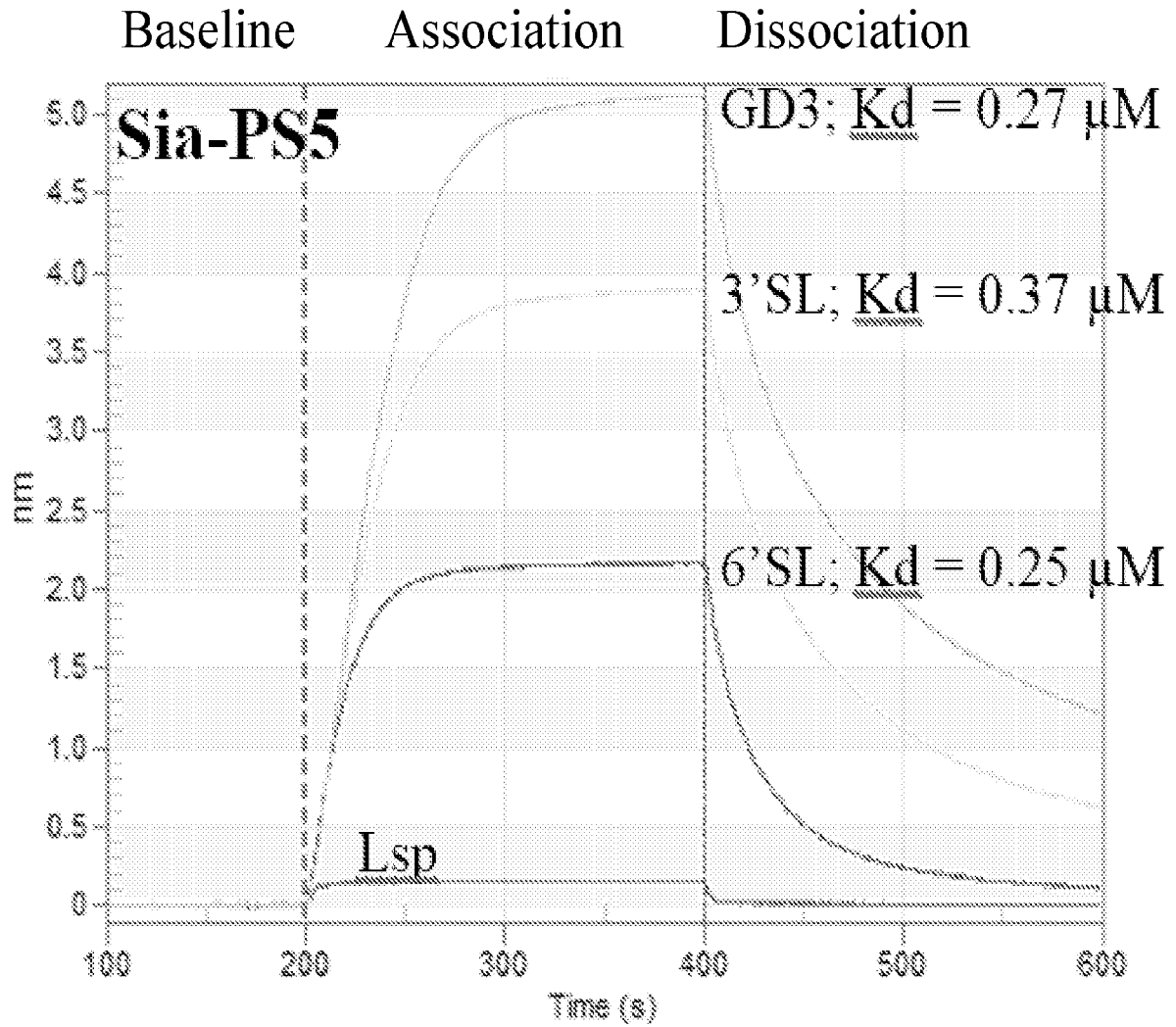


FIG. 7C

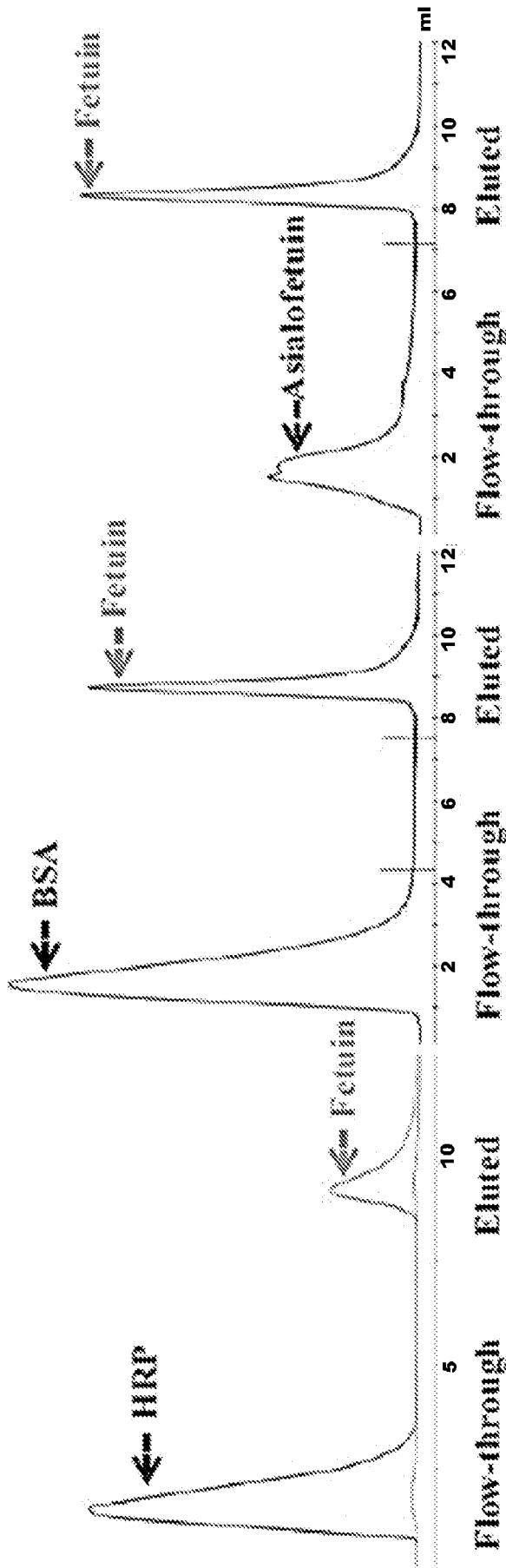


FIG. 7B

FIG. 7A

FIG. 8A

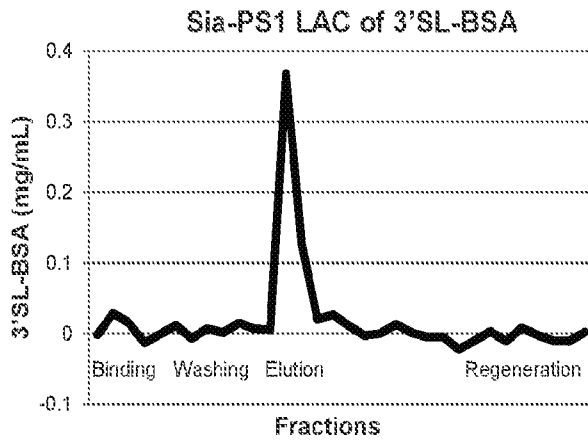


FIG. 8B

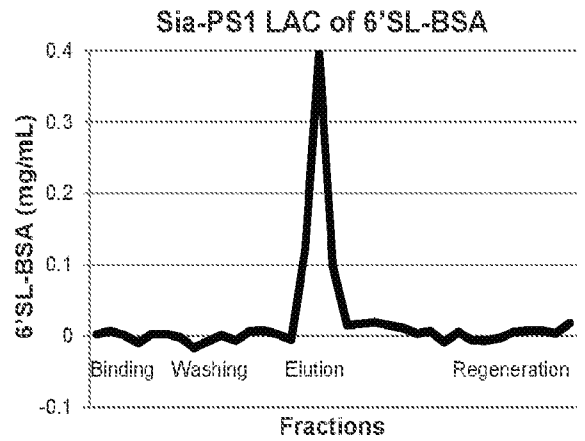


FIG. 9A

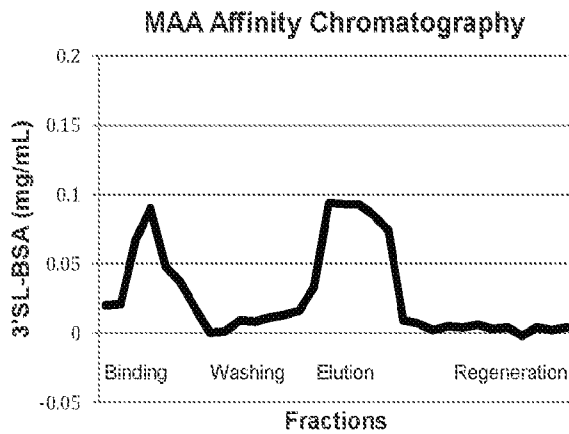


FIG. 9B

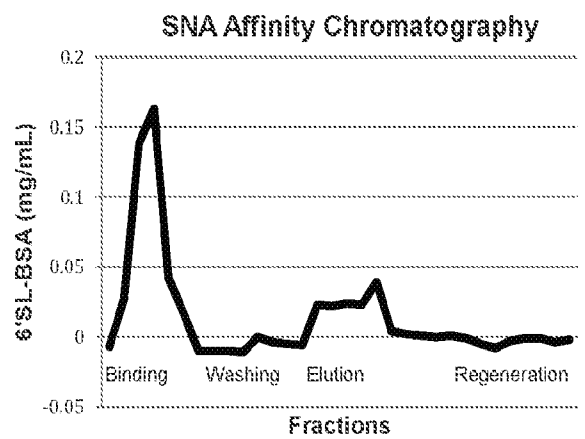


FIG. 10C

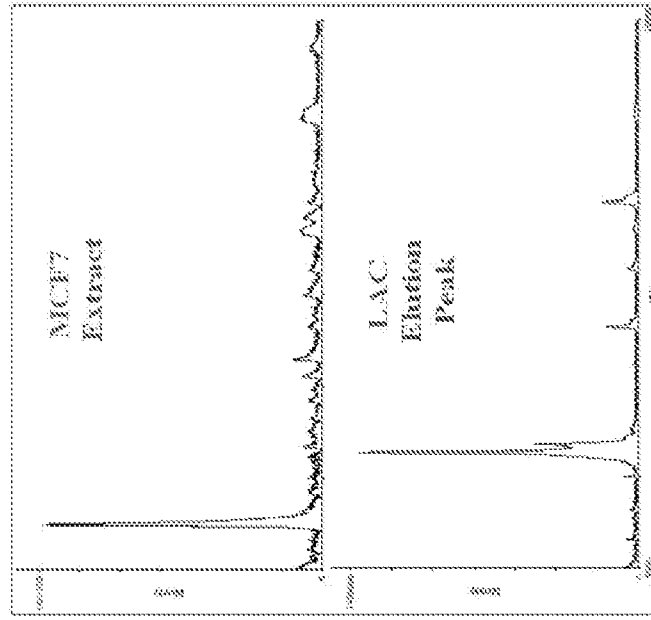


FIG. 10B

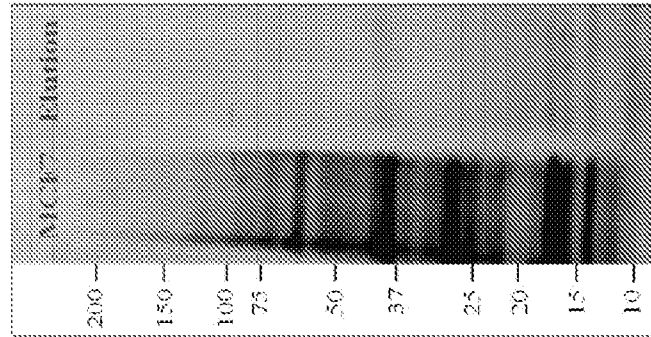


FIG. 10A

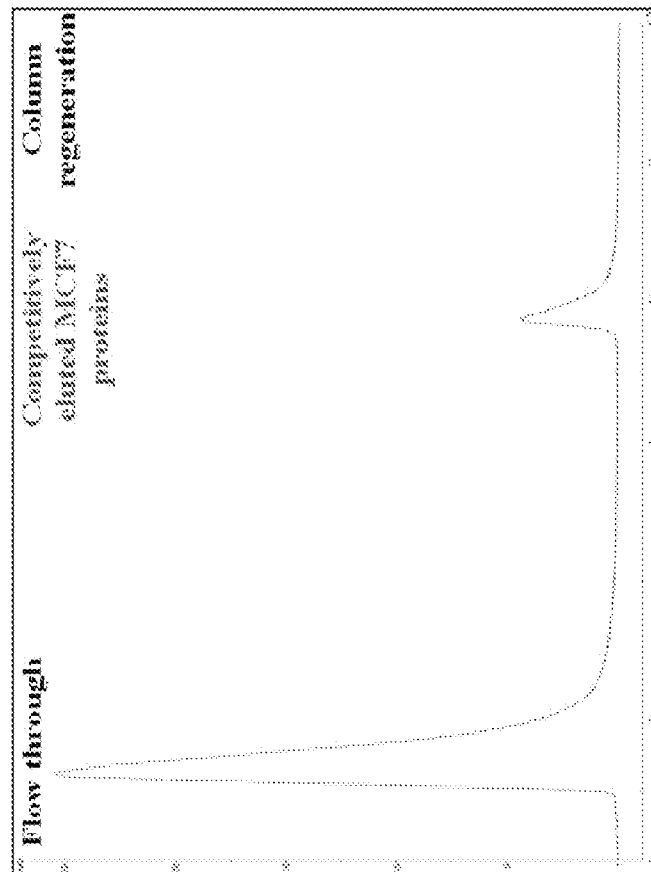


Fig. 11

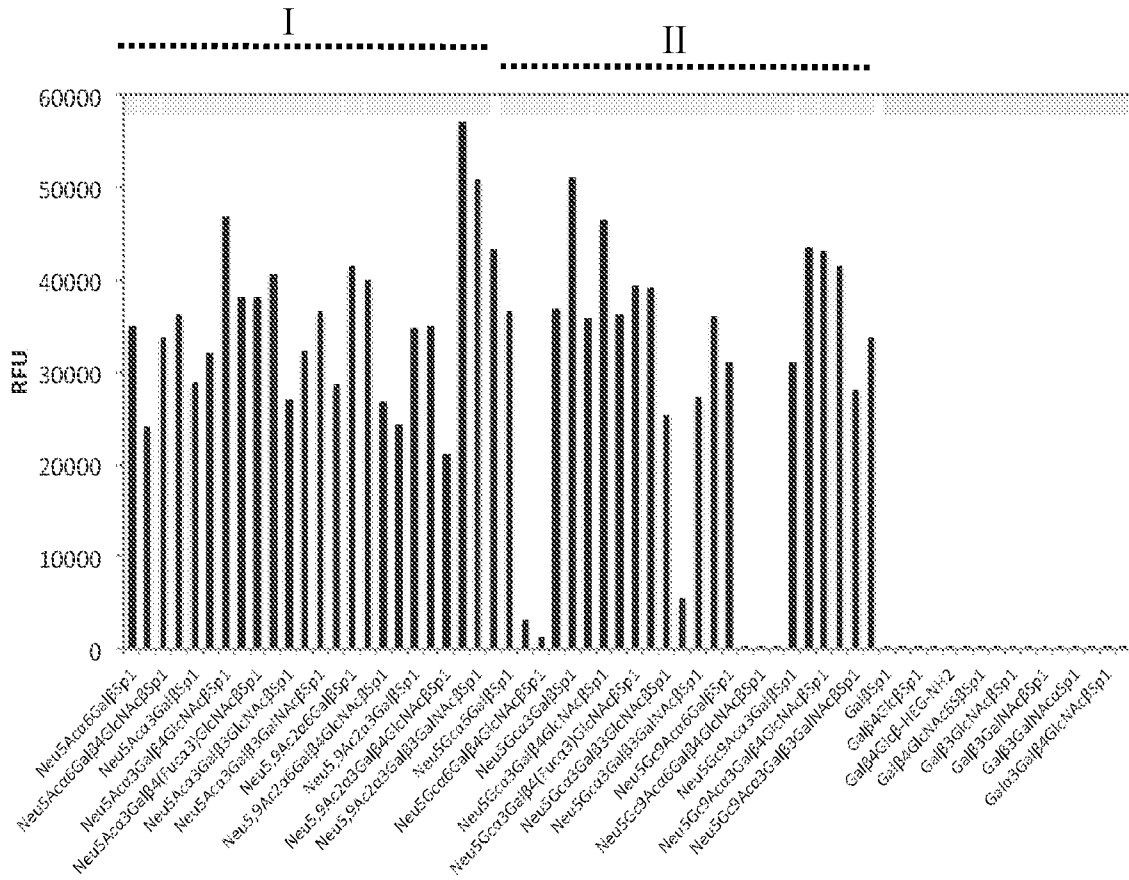


FIG. 12

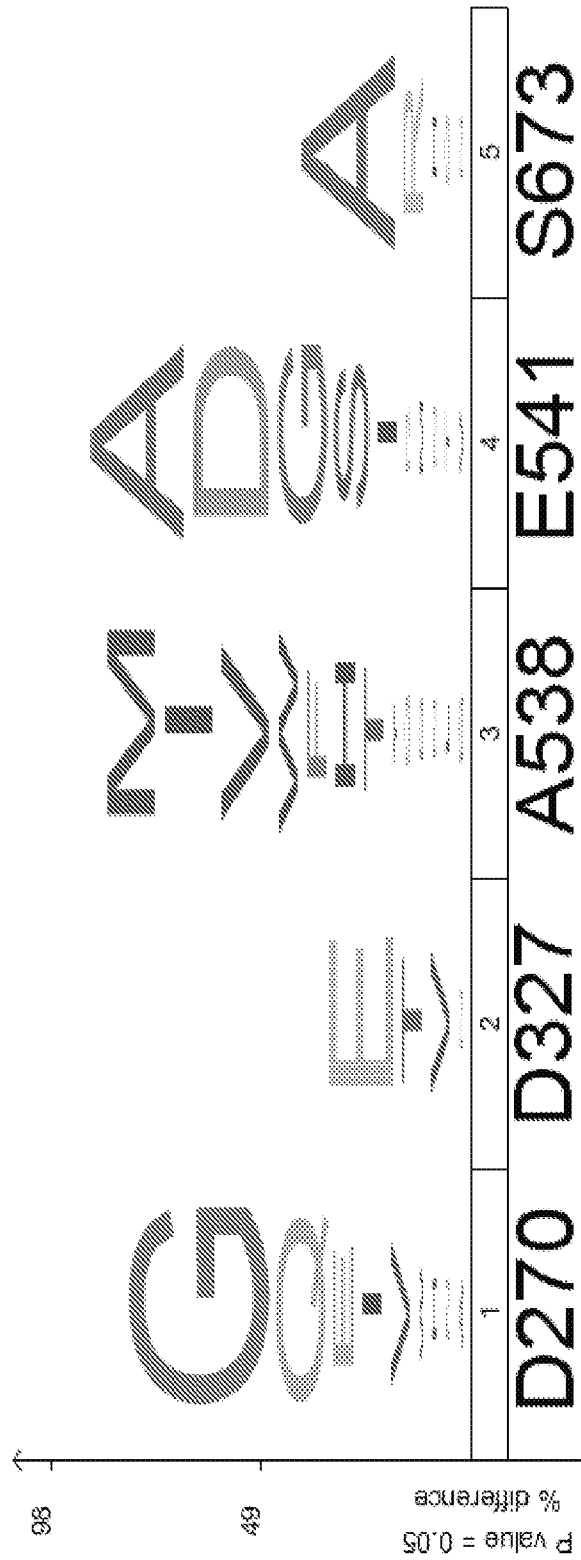


FIG. 13

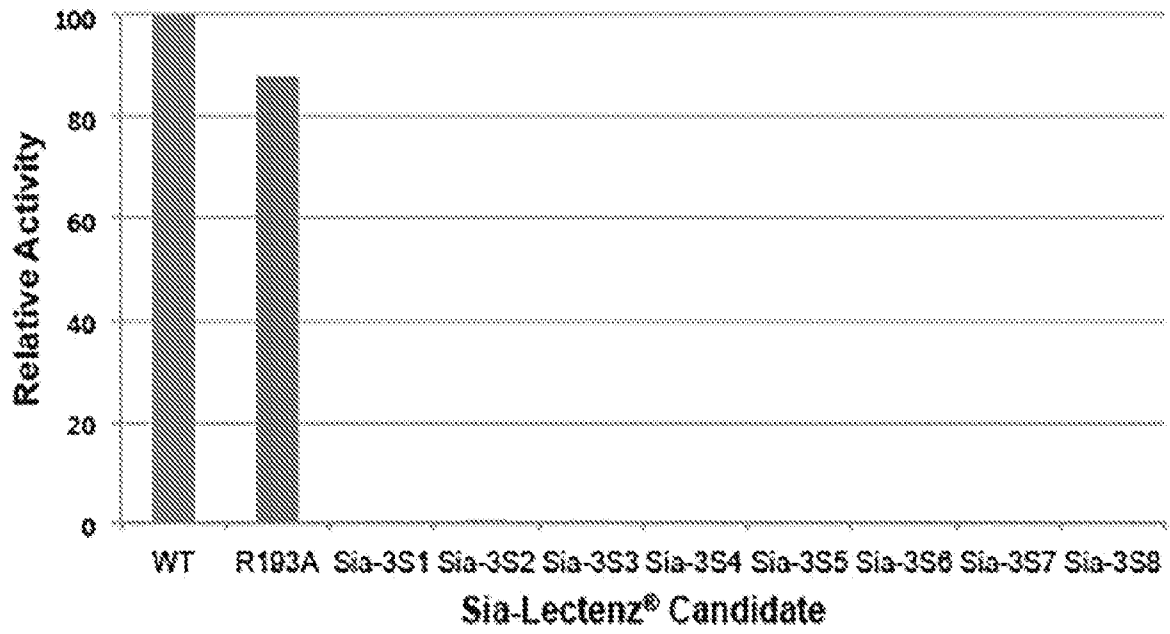


FIG. 14B

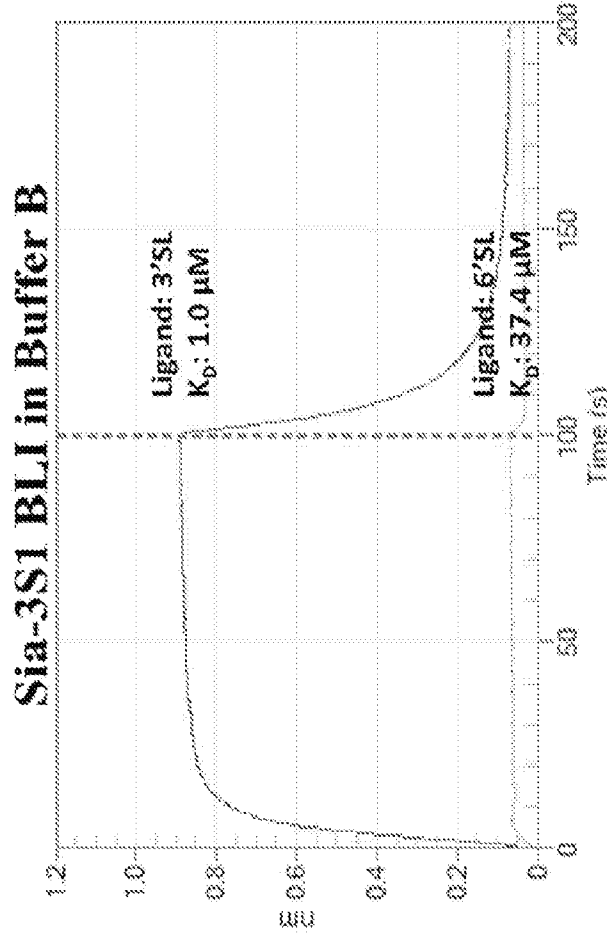
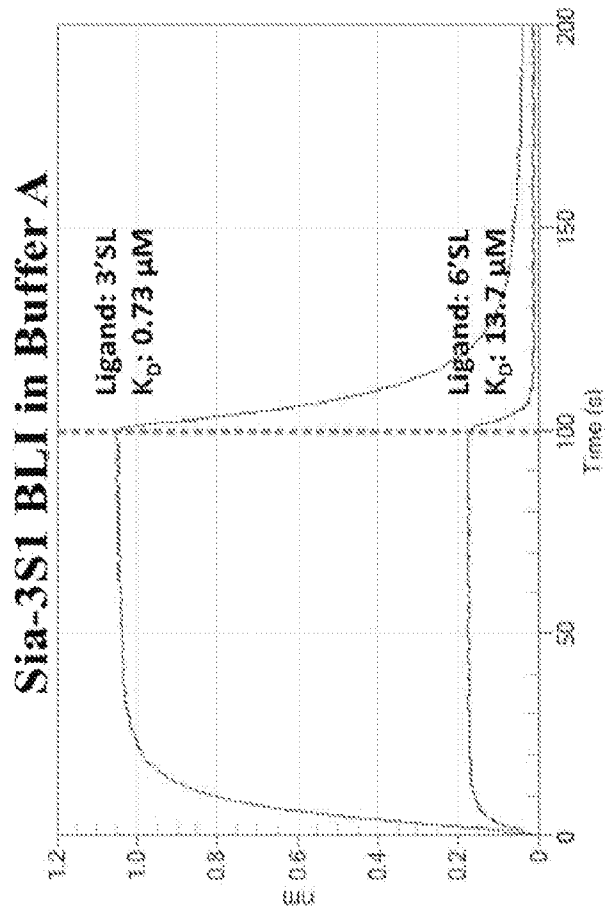


FIG. 14A



19/23

FIG. 15

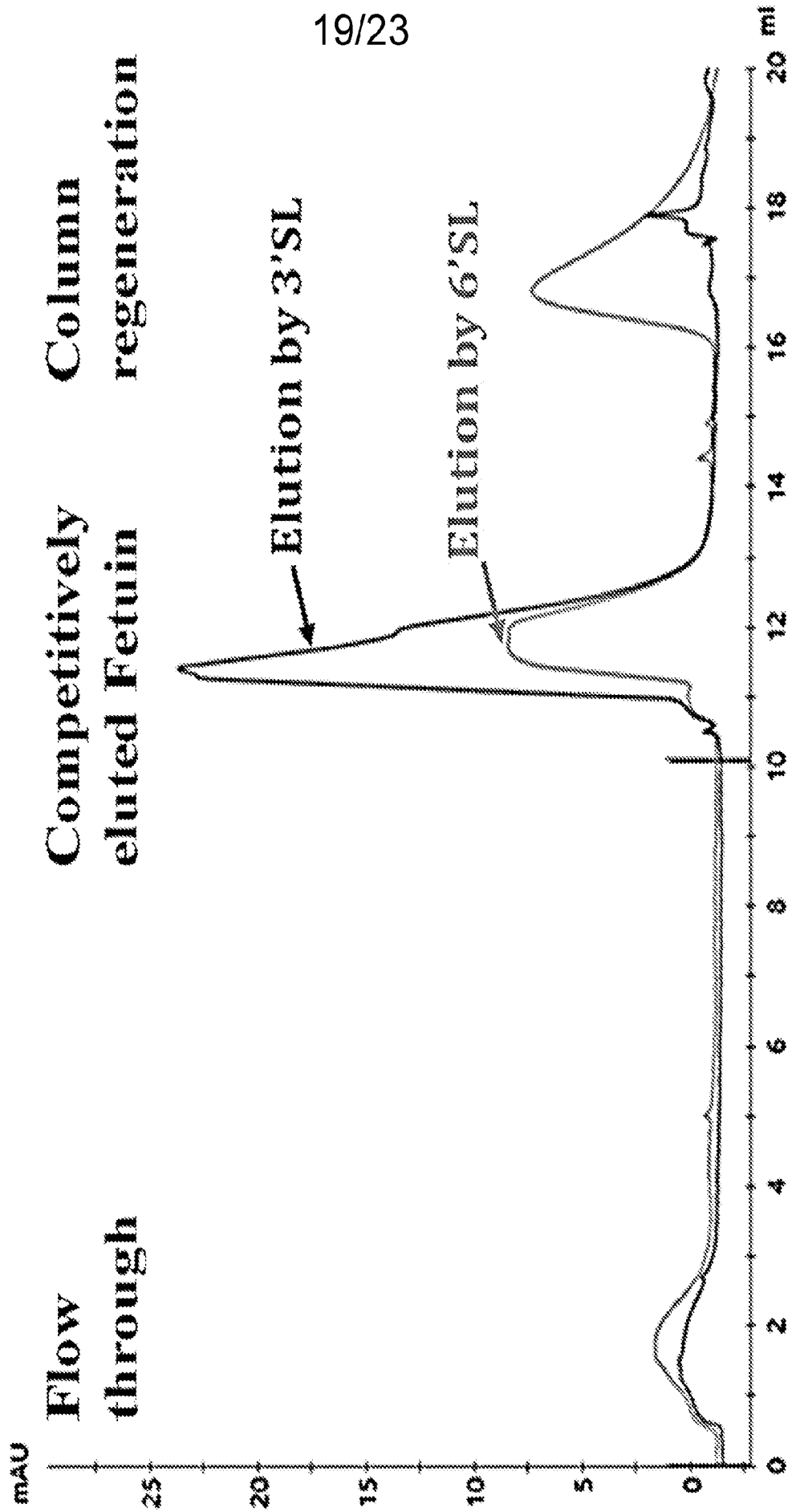


FIG. 16A

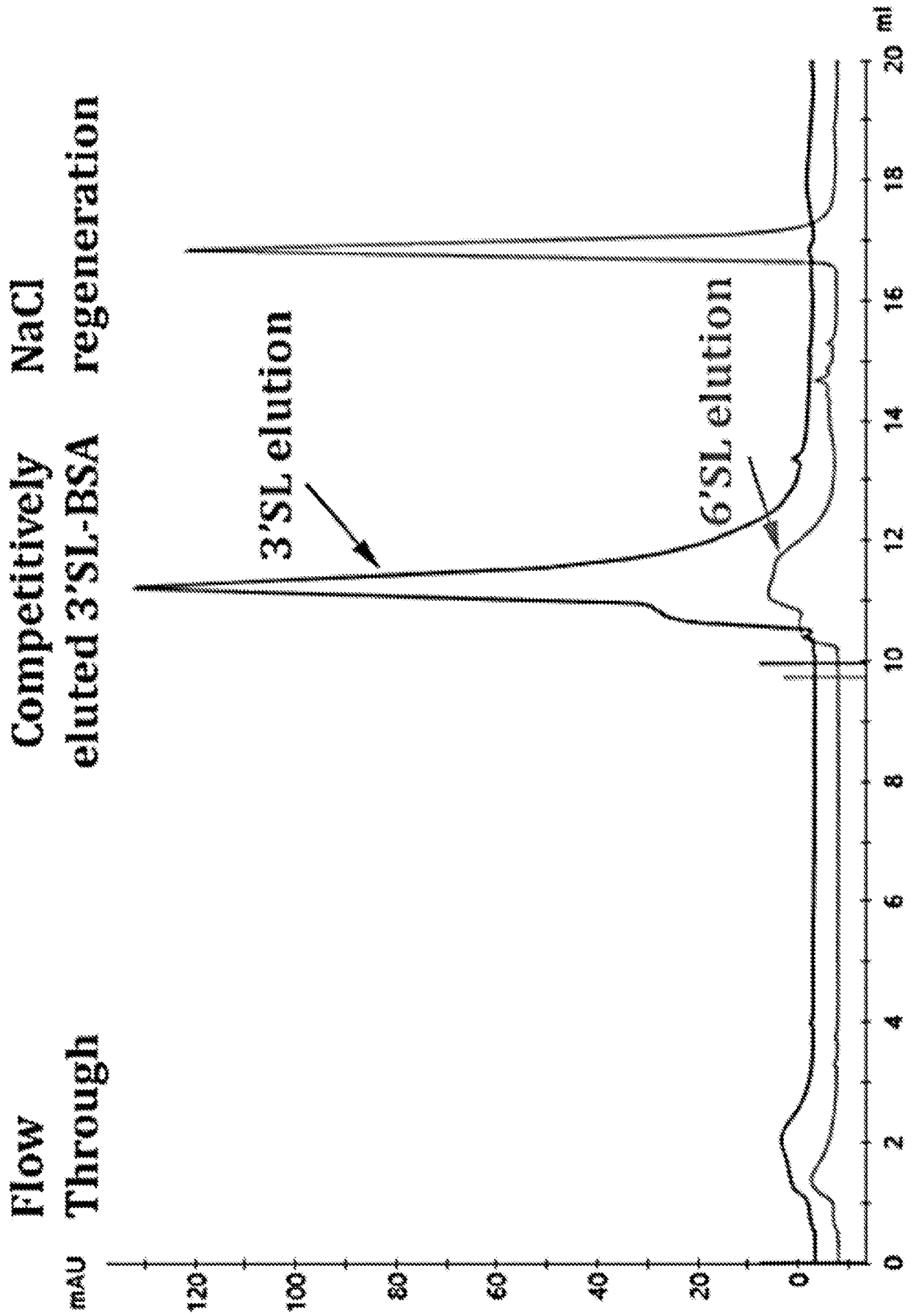


FIG. 16B

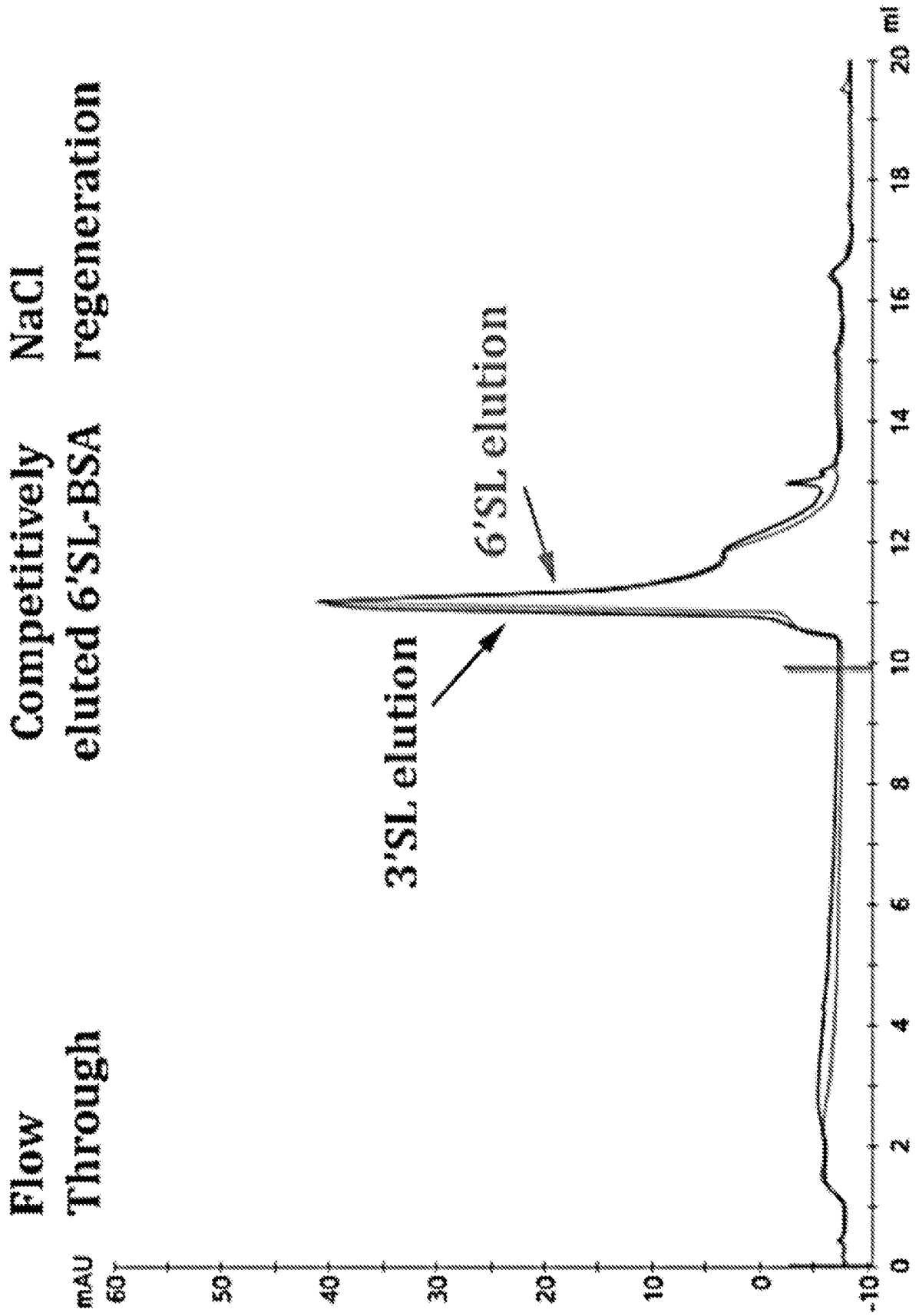


FIG. 17

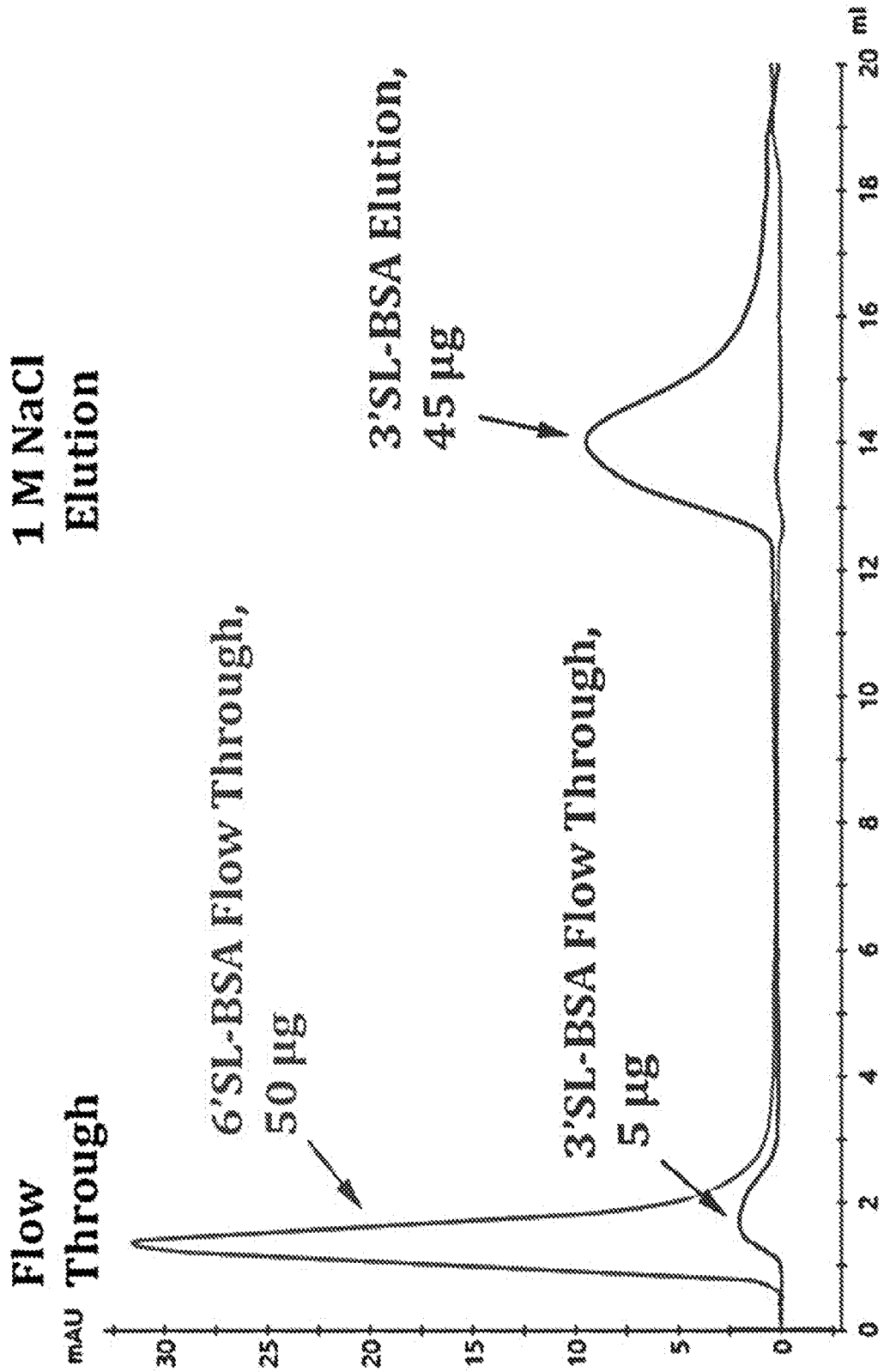


FIG. 18

