

(51) International Patent Classification:
C12N 9/10 (2006.01)

8837 Chapel Square Drive, Suite C, Cincinnati, OH 45249 (US).

(21) International Application Number:
PCT/US2021/041537

(74) Agent: NESBITT, Daniel, F.; Hasse & Nesbitt LLC, 8837 Chapel Square Drive, Suite C, Cincinnati, OH 45249 (US).

(22) International Filing Date:
14 July 2021 (14.07.2021)

(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, IT, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(25) Filing Language: English

(26) Publication Language: English

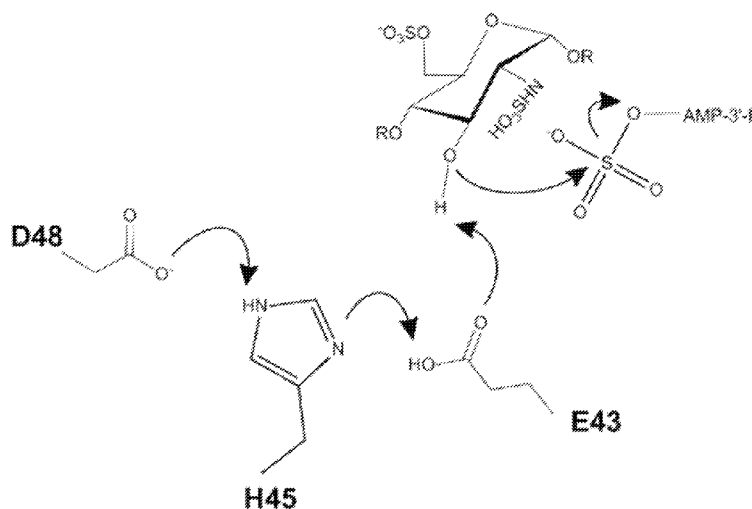
(30) Priority Data:
63/051,764 14 July 2020 (14.07.2020) US

(71) Applicant: OPTIMVIA, LLC [US/US]; 458 W. Main Street, Batavia, OH 45103 (US).

(72) Inventors: GESTEIRA, Tarsis, Ferreira; 458 W. Main Street, Batavia, OH 45103 (US). LAJINESS, Daniel, H.;

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH,

(54) Title: METHODS FOR SYNTHESIZING NON-ANTICOAGULANT HEPARAN SULFATE



Reaction

Figure 1A

(57) **Abstract:** Non-anti coagulant heparan sulfate polysaccharides, and methods for preparing thereof using non-naturally occurring, engineered sulfotransferase enzymes that are designed to react with aryl sulfate compounds instead of the natural substrate, PAPS, to facilitate sulfo group transfer to polysaccharide sulfo group acceptors. Suitable aryl sulfate compounds include, but are not limited to, p-nitrophenyl sulfate or 4-nitrocatechol sulfate. One class of heparan sulfate polysaccharides produced by methods of the present invention comprise N- and 6-O-sulfated glucosamine residues, but are neither 2-O sulfated or 3-O sulfated. Such A Γ -, 6-O sulfated heparan sulfate polysaccharides can have comparable pharmacological activity compared to (9-desulfated heparan (ODSH), without possessing any of ODSH's low-level anticoagulant activity.



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

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*
- *of inventorship (Rule 4.17(iv))*

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
- *with sequence listing part of description (Rule 5.2(a))*

METHODS FOR SYNTHESIZING NON-ANTICOAGULANT HEPARAN SULFATE

CROSS-REFERENCE TO RELATED APPLICATIONS

[0000] The instant application claims the benefit of U.S. Provisional Application No. 63/051,764, filed on July 14, 2020, the disclosure of which is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

[0001] The present invention relates to methods for synthesizing anticoagulant polysaccharides using engineered, non-natural sulfotransferase enzymes that are designed to react with aryl sulfate compounds as sulfo group donors.

REFERENCE TO SEQUENCE LISTING

[0002] The present application is being filed along with a sequence listing in electronic format. The sequence listing is provided as a file entitled "OPT-008P_sequence_disclosure.txt" created on June 30, 2020 and which is 125,627 bytes in size. The information in electronic format of the sequence listing is incorporated by reference in its entirety.

BACKGROUND OF THE INVENTION

[0003] Although heparin and low-molecular weight heparin (LMWH) are commonly known and prescribed as anticoagulants to reduce or prevent blood clotting, they can also be useful during the treatment of conditions such as, for example: cancers; inflammation; thrombocytopenia; neutropenia; apoptosis; asthma; emphysema; bronchitis; adult respiratory distress syndrome; cystic fibrosis; and ischemia-reperfusion related conditions. However, for many patients, diminished blood coagulation is an unwanted side effect that in some instances can cause more problems than the treatment itself is designed to remedy.

[0004] One solution to this problem has been to develop 2-*O*, 3-*O*-desulfated heparin (ODSH) derivatives, which have a greatly diminished anticoagulant activity relative to heparin or LMWH. Generally, ODSH compositions are prepared by reacting unfractionated heparin or LMWH with a strong base, often at cold temperatures, to remove 2-*O* and 3-*O* sulfate groups from the heparan sulfate polysaccharide backbone. Methods of preparing ODSH and their use in the treatment of patients is described in U.S. Patents 10,052,346, 9,271,999, 7,468,358, 6,489,311, 5,990,097, 5,912,237, 5,808,021, 5,668,118, and 5,296,471, the disclosures of which are incorporated by reference in their entireties (*see also* Mousavi, S. et al., (2015) *Advances in Pharmacological Sciences*, Article ID 507151, available at <http://dx.doi.org/10.1155/2015/S07151>, incorporated by reference in its entirety). However, although ODSH has been produced that has a reduction of up to

99% of anticoagulant activity relative to heparin, the complete removal of all anticoagulant activity, while maintaining the pharmacological benefits of ODSH, has not been reported.

[0005] Further, because ODSH is prepared from unfractionated heparin, which is isolated and purified from the internal organs of animals, such as pigs and cows, they are susceptible to disruptions in the worldwide supply due to potential contamination of heparin (over 200 people died as a result of contaminated compounds in 2007 in the United States alone), cross-species transmission of the flu and/or other animal viruses into humans, or geopolitical tensions with global suppliers, particularly China. As a result, there has been a recent push to try to synthesize heparin *in vitro*.

[0006] Generally, sulfated polysaccharides, including heparin, are synthesized by the catalytic transfer of sulfate functional groups, also called “sulfo groups”, from a sulfo group donor to a polysaccharide, which acts as a sulfo group acceptor. Each sulfo group transfer is catalyzed by a sulfotransferase enzyme, and there are often multiple sulfotransfer reactions catalyzed by multiple sulfotransferase enzymes to ultimately arrive at each sulfated polysaccharide product. Sulfotransferases are nearly ubiquitous in nature, and they exist in nearly all types of organisms, including bacteria, yeast, and animals, including humans. Similarly, sulfotransferase enzymes play an integral role in the sulfation of a wide array of sulfo group acceptors, including many types of steroids, polysaccharides, proteins, xenobiotics, and other molecules.

[0007] With respect to polysaccharides in particular, there are several polysaccharides that can be utilized as sulfo group acceptors, including, for example, dermatan, keratan, heparosan, and chondroitin. In particular, heparin is formed from heparosan, which comprises repeating dimers of 1→4 glycosidically-linked glucuronic acid and *N*-acetylated glucosamine residues. In nature, heparin is formed upon the removal of *N*-acetyl groups, inversion of stereochemistry of glucuronic acid residues, and reaction with four different sulfotransferase enzymes that transfer sulfo groups to multiple positions within the polysaccharide.

[0008] As wide-ranging and voluminous as the set of sulfo group acceptors can be, the number of molecules used as sulfo group donors for reactions catalyzed by sulfotransferase enzymes is relatively small. Most typically, 3'-phosphoadenosine 5'-phosphosulfate (PAPS) is utilized as the sulfo group donor, and in reactions in which a polysaccharide is the sulfo group acceptor, PAPS is the only known sulfo group donor. However, PAPS is often unsuitable for use as a sulfo group donor to catalyze enzymatic syntheses of sulfated polysaccharides *in vitro*, particularly in large scale syntheses, because it has an extremely short shelf life and can readily decompose into adenosine 3',5'-diphosphate, which actively inhibits the sulfotransferases' biological activity. In contrast, *in*

in vivo systems have evolved to exclusively and efficiently react with PAPS because adenosine 3',5'-diphosphate can either readily be converted back into PAPS or be broken down into one or more compounds that do not inhibit sulfotransferase activity. As a result, the natural activity of sulfotransferase enzymes to exclusively utilize PAPS as a sulfo group donor presents a steep barrier to the *in vitro* synthesis of heparin, from which ODSH is prepared.

[0009] On the other hand, aryl sulfate compounds, such as *p*-nitrophenyl sulfate (PNS) and 4-methylumbelliferyl sulfate (MUS) have been identified as cheap, widely-available compounds that can be useful in limited situations as sulfo donors with sulfotransferases to synthesize certain small molecule products (see Malojcic, G., et al. (2008) *Proc. Nat. Acad. Sci.* **105** (49):19217-19222 and Kaysser, L., et al., (2010) *J. Biol. Chem.* **285** (17):12684-12694, the disclosures of which are incorporated by reference in their entirety). As described by Malojcic, these aryl sulfate sulfotransferases undergo a two-step mechanism, where the enzyme first removes the sulfo group from the aryl sulfate compound and forms a sulfohistidine intermediate in which the sulfo group is covalently bonded with an amino acid side chain, typically a histidine residue, within the active site. The sulfate-bound form of the enzyme can then recognize and bind with a sulfo group acceptor to complete the sulfo group transfer

[0010] Yet, only a small number of bacterial sulfotransferases have been shown to react with aryl sulfate compounds as sulfo group donors, and eukaryotic sulfotransferases that react with polysaccharides as sulfo group acceptors demonstrate no biological activity when aryl sulfate compounds are used as sulfo group donors. Instead, such sulfotransferases exclusively react with PAPS as the sulfo donor, as described above. As a result, when sulfotransferases are used in *in vitro* syntheses of sulfated polysaccharides, the sulfotransferases cannot catalyze transfer of the sulfo group from aryl sulfate compounds to the polysaccharides directly. Instead, aryl sulfate compounds can only be used indirectly to repopulate the system with PAPS (see U.S. Pat. No. 6,255,088, the disclosure of which is incorporated by reference in its entirety).

[0011] Consequently, there is a need to develop facile methods of synthesizing heparan sulfate polysaccharides *in vitro*, particularly *N*-sulfated, 6-*O* sulfated heparan sulfate polysaccharides, which can be utilized as analogs to ODSH while not possessing any of its low-level anticoagulant activity.

SUMMARY OF THE INVENTION

[0012] The present invention provides methods for producing heparan sulfate *in vitro* using non-naturally occurring sulfotransferase enzymes that have been engineered to catalyze the transfer of sulfo groups from aryl sulfate compounds as sulfo group donors to react with polysaccharides as

sulfo group acceptors. Non-anticoagulant *N*-, 6-*O* sulfated heparan sulfate (NS6S/HS) products synthesized by such methods can be formed as structural analogs of ODSH, which retain some glucosamine 3-*O* sulfation and accordingly, low-level anticoagulant activity. On the other hand, NS6S/HS products can contain no 3-*O* sulfated glucosamine residues, and optionally, no 2-*O* sulfated hexuronic acid residues. Such fully non-anticoagulant NS6S/HS compositions can be utilized in the treatment of several medical conditions, including, as non-limiting examples, cancers; inflammation; thrombocytopenia; neutropenia; apoptosis; asthma; emphysema; bronchitis; adult respiratory distress syndrome; cystic fibrosis; and ischemia-reperfusion related conditions. Such treatments are described in further detail, below.

[0013] In an aspect of the present invention, heparan sulfate can be synthesized enzymatically by combining a heparosan-based polysaccharide, an aryl sulfate compound, and an engineered sulfotransferase enzyme having a biological activity comprising the transfer of a sulfo group from the aryl sulfate compound to the polysaccharide. Heparosan-based polysaccharides are derived from heparosan $[\beta(1,4)\text{GlcA}-\alpha(1,4)\text{GlcNAc}]_n$, and comprise repeating dimers of 1→4 glycosidically-linked hexuronic acid and glucosamine residues, wherein each hexuronic acid is either glucuronic acid (GlcA, above) or iduronic acid (IdoA). Generally, the amine group in each glucosamine residue can either be *N*-acetylated, *N*-sulfated, or *N*-unsubstituted. When at least one of the glucosamine residues within heparosan is *N*-unsubstituted, the polysaccharide can also be called *N*-deacetylated heparosan. *N*-sulfated glucosamine residues can also be 3-*O* and 6-*O* sulfated, while any of the GlcA or IdoA residues can be sulfated at the 2-*O* position. Heparosan-based polysaccharides that contain at least one sulfate group in any of the above positions within a hexuronic acid or glucosamine residue can also be called heparan sulfate (HS).

[0014] In various embodiments, a sulfated polysaccharide product formed in a first sulfotransfer reaction can be utilized as a sulfo group acceptor in a subsequent reaction with another sulfotransferase enzyme, which can either be performed in the same reaction mixture as the first sulfotransfer reaction, or in a separate reaction mixture after isolating the sulfated polysaccharide product and combining it with a sulfo group donor and a sulfotransferase enzyme. In various embodiments, a plurality of sulfotransfer reactions can be carried out, either sequentially or simultaneously, on a single heparosan-based polysaccharide, including at least two, at least three, or at least four sulfotransfer reactions. Each of the plurality of sulfotransfer reactions on a heparosan-based polysaccharide can be catalyzed by at least two, at least three, or up to four sulfotransferase enzymes. In various embodiments, at least one, and preferably all, of the sulfotransfer reactions are catalyzed by an engineered sulfotransferase enzyme which recognizes, binds, and reacts with the

aryl sulfate compound as a sulfo group donor. In further embodiments, at least one, and preferably all, of the sulfotransfer reactions are carried out in reaction mixtures that contain only an aryl sulfate compound and do not contain PAPS.

[0015] In another aspect of the invention, each engineered sulfotransferase enzyme comprises several amino acid mutations made within the active site of a corresponding natural sulfotransferase enzyme, in order to shift the enzyme's biological activity from reacting with PAPS as the sulfo group donor to reacting with an aryl sulfate compound as a sulfo group donor. However, in various embodiments, each engineered sulfotransferase enzyme retains the natural enzyme's biological activity with its particular sulfo acceptor polysaccharide. As a non-limiting example, the *N*-sulfotransferase domain of a natural *N*-deacetylase/*N*-sulfotransferase (NDST) enzyme, which has a biological activity in which the enzyme reacts with PAPS as a sulfo group donor and *N*-deacetylated heparosan as a sulfo group acceptor, can be mutated in multiple amino acid positions to generate an engineered glucosaminyl *N*-sulfotransferase enzyme that recognizes, binds, and reacts with an aryl sulfate compound as a sulfo group donor, but that still reacts with *N*-deacetylated heparosan as a sulfo group acceptor. Such engineered glucosaminyl *N*-sulfotransferase (NST) enzymes, and others, are described in further detail below.

[0016] In another aspect of the invention, NS6S/HS can be synthesized using a method comprising the following steps: (a) providing a starting polysaccharide composition comprising *N*-deacetylated heparosan; (b) reacting the starting polysaccharide composition within a reaction mixture comprising an *N*-sulfation agent, to form an *N*-sulfated heparan sulfate (NS/HS) product; and (c) reacting the NS/HS product within a reaction mixture comprising an aryl sulfate compound and an engineered glucosaminyl 6-*O* sulfotransferase (6OST) enzyme, thereby forming the NS6S/HS product; wherein the biological activity of the engineered 6OST enzyme comprises the transfer of a sulfo group from an aryl sulfate compound to a heparosan-based polysaccharide. In various embodiments, the biological activity of the engineered 6OST enzyme consists of the transfer of a sulfo group from an aryl sulfate compound to a heparosan-based polysaccharide, preferably in the absence of PAPS.

[0017] In various embodiments, an engineered 6OST enzyme utilized in any of the methods described herein can comprise any amino acid sequence so long as the enzyme catalyzes the transfer of a sulfo group from an aryl sulfate compound to the 6-*O* position of a glucosamine residue within a heparosan-based polysaccharide. Generally, the heparosan-based polysaccharides that can be utilized as sulfo group acceptors with an engineered 6OST enzyme are identical to that of a natural 6OST enzyme, in which the amine group of the glucosamine residue receiving the sulfo group at the 6-*O* position can be *N*-sulfated, *N*-unsubstituted, or *N*-acetylated. In various embodiments,

glucosamine residues within the heparosan-based polysaccharide that do not receive the sulfo group can be optionally be *N*-, 3-*O*, and/or 6-*O* sulfated, *N*-acetylated, or *N*-unsubstituted, and any of the hexuronic acid residues within the heparosan-based polysaccharide can either be iduronic acid or glucuronic acid, and can optionally be 2-*O* sulfated, prior to reacting with the 6OST enzyme. In various embodiments, an engineered 6OST can react with *N*-sulfated heparosan that is neither 2-*O*, 6-*O*, nor 3-*O* sulfated, prior to reacting with the 6OST.

[0018] In various embodiments, the engineered 6OST enzymes can be mutants of natural sulfotransferases that have heparan sulfate 6OST activity, which are members of enzyme class EC 2.8.2.-. According to the present invention, an engineered 6OST enzyme can comprise several amino acid mutations relative to one or more of the natural EC 2.8.2.- enzymes with 6OST activity, in order to reconfigure the active site to bind and react with an aryl sulfate compound as a sulfo group donor instead of PAPS. Non-limiting examples of engineered 6OST enzymes that can be utilized in accordance with any of the methods described herein can comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 59, SEQ ID NO: 60, and SEQ ID NO: 61, each of which contains several amino acid mutations made relative to highly conserved regions within natural 6OST enzymes. In various embodiments, the engineered 6OST comprises the amino acid sequence SEQ ID NO: 18. In various embodiments, engineered 6OST enzymes utilized in accordance with any of the methods described herein can also comprise an amino acid sequence having one or more amino residue differences or mutations from, and/or is a biological functional equivalent of, an amino acid sequence selected from the group consisting of SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 59, SEQ ID NO: 60, and SEQ ID NO: 61. Non-limiting examples of such residue differences include amino acid insertions, deletions, substitutions, or any combination of such changes.

[0019] In various embodiments, any natural 6OST enzyme, or a biologically-active fragment thereof, can be utilized to catalyze 6-*O* sulfation during the synthesis of HS products, particularly NS6S/HS products, in which an engineered sulfotransferase enzyme is utilized in at least one other sulfation step, for example in the enzymatic *N*-sulfation of *N*-deacetylated heparosan. Reaction mixtures comprising a natural 6OST enzyme also comprise PAPS as a sulfo group donor.

[0020] In various embodiments, NS/HS polysaccharides can be isolated and purified prior to reacting with the 6OST in a separate reaction mixture. In other embodiments, 6-*O* sulfation of glucosamine residues can take place in the same reaction mixture as the *N*-sulfation of *N*-deacetylated heparosan.

[0021] In various embodiments, the step of providing the starting polysaccharide reaction mixture can comprise the chemical synthesis of *N*-sulfated heparosan, comprising the following sub-steps: (i) providing a precursor polysaccharide composition comprising heparosan; (ii) combining the precursor polysaccharide composition with a reaction mixture comprising a base, preferably lithium hydroxide or sodium hydroxide, for a time sufficient to *N*-deacetylate at least one of the *N*-acetylated glucosamine residues within the heparosan to form the starting polysaccharide composition.

[0022] In various embodiments, the step of providing the precursor polysaccharide composition comprising heparosan can further comprise the sub-step of isolating heparosan from a bacterial or eukaryotic cell culture, preferably a bacterial cell culture, and more preferably a bacterial cell culture comprising bacteria selected from the group consisting of the K5 strain of *Escherichia coli* (*E. coli*) and the BL21 strain of *E. coli*. Heparosan can be isolated from *E. coli* as a polydisperse mixture of polysaccharides having a weight-average molecular weight of at least 10,000 Da, and up to at least 1,000,000 Da. In various embodiments, at least 90% of the glucosamine residues within the heparosan are *N*-acetylated.

[0023] Treating heparosan with a base, such as lithium hydroxide or sodium hydroxide, removes acetyl groups from *N*-acetyl glucosamine residues, forming *N*-unsubstituted glucosamine residues that can subsequently be *N*-sulfated by an *N*-sulfation agent. In various embodiments, precursor polysaccharides can be treated with a base for a time sufficient to reduce the relative number of *N*-acetylated glucosamine residues to a desired level. The reaction time can be dependent on factors such as the average molecular weight of the heparosan within the precursor polysaccharide composition, the *N*-acetyl glucosamine content of the heparosan prior to reacting with the base, the desired *N*-acetyl content within the *N*-deacetylated heparosan composition, and the concentration and identity of the base itself. In various embodiments, the time sufficient to *N*-deacetylate the heparosan within the precursor polysaccharide composition can be the time sufficient to form an *N*-deacetylated heparosan composition in which less than 60%, down to less than 5%, preferably in the range of 12% to 18%, and more preferably 15%, of the glucosamine residues remain *N*-acetylated.

[0024] Additionally, treating the precursor polysaccharide composition with a base to reduce the number of *N*-acetylated glucosamine residues can also have the effect of depolymerizing the

heparosan, causing the *N*-deacetylated heparosan composition to have a lower average molecular weight relative to the precursor polysaccharide composition. Accordingly, in various embodiments, the precursor polysaccharide composition can be treated with a base for a time sufficient to form an *N*-deacetylated heparosan composition having a desired average molecular weight. As with above, the reaction time can depend on several factors, including the average molecular weight of the heparosan within the precursor polysaccharide composition, and the desired average molecular weight of the polysaccharides within the *N*-deacetylated heparosan composition itself. In various embodiments, the time sufficient to *N*-deacetylate the heparosan within the precursor polysaccharide composition can be the time sufficient to form an *N*-deacetylated polysaccharide composition having a weight-average molecular weight in a range from 1,500 Da to 100,000 Da, for example, from at least 9,000 Da, and up to 12,500 Da.

[0025] In various embodiments, once the *N*-deacetylated heparosan is formed, the resulting *N*-unsubstituted glucosaminyl residues can then receive a sulfo group upon reacting within a reaction mixture comprising an *N*-sulfation agent, to form NS/HS. In various embodiments, one or more of the *N*-unsubstituted glucosamine residues within *N*-deacetylated heparosan can be chemically *N*-sulfated. A non-limiting example of a chemical *N*-sulfation agent can comprise a reaction mixture comprising a sulfur trioxide-containing compound or adduct, particularly a sulfur trioxide-trimethylamine adduct.

[0026] In various embodiments, the *N*-sulfation agent is an engineered NST or natural NDST enzyme. In various embodiments, enzymatic *N*-sulfation can either supplement or replace chemical *N*-sulfation of *N*-deacetylated heparosan. In various embodiments, a natural NDST enzyme can be utilized when at least one additional enzymatic sulfation step, particularly 6-*O* sulfation, is catalyzed by an engineered aryl sulfate-dependent sulfotransferase. Similar to the 6OSTs above, reaction mixtures containing a natural NDST enzyme also comprise PAPS as a sulfo group donor.

[0027] In various embodiments, an engineered NST enzyme utilized in any of the methods described herein can comprise any amino acid sequence so long as the enzyme catalyzes the transfer of a sulfo group from an aryl sulfate compound to the amine functional group of an *N*-unsubstituted glucosamine residue of a heparosan-based polysaccharide, preferably *N*-deacetylated heparosan. In various embodiments, the engineered NST enzyme can be a mutant of the *N*-sulfotransferase domain of a natural NDST enzyme, which is a member of the enzyme class, EC 2.8.2.8. In various embodiments, an engineered NST enzyme can comprise several amino acid mutations relative to the *N*-sulfotransferase domain of one or more natural NDST enzymes, in order to engineer the active site to bind and react with an aryl sulfate compound as a sulfo group donor instead of PAPS.

[0028] Engineered NST enzymes utilized in accordance with any of the methods described herein can comprise any amino acid sequence so long as the enzyme has activity with an aryl sulfate compound as a sulfo group donor. In a non-limiting example, and in some embodiments, the amino acid sequence of the engineered NST enzyme can be selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40, each of which contains several amino acid mutations made relative to highly conserved regions within the *N*-sulfotransferase domain of natural NDST enzymes within EC 2.8.2.8. In various embodiments, engineered NST enzymes utilized in accordance with any of the methods described herein can also comprise an amino acid sequence having one or more amino residue differences or mutations from, and/or is a biological functional equivalent of, an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40. Non-limiting examples of such residue differences include amino acid insertions, deletions, substitutions, or any combination of such changes.

[0029] In various embodiments, any of the engineered glucosaminyl *N*-sulfotransferase enzymes described above can further include an *N*-deacetylase domain that is either identical or mutated relative to the *N*-deacetylase domain that is present in any native NDST enzyme. In various embodiments, any of the engineered NST enzymes can further include other domains or fusions with other proteins to facilitate solubility or secondary biochemical reactions.

[0030] Glucosamine residues within the heparosan-based polysaccharide that do not receive the sulfo group can be *N*-, 3-*O*-, and/or 6-*O* sulfated, *N*-acetylated, or *N*-unsubstituted, and hexuronic acid residues can include GlcA or IdoA, either of which can be sulfated at the 2-*O* position. Preferably, the heparosan-based polysaccharide is *N*-deacetylated heparosan, and all other positions within the polysaccharide are unsulfated. In other embodiments, the 6-*O* group of an *N*-unsubstituted glucosamine residue can already be sulfated prior to the *N*-sulfation reaction.

[0031] In various embodiments, by either chemical and/or enzymatic *N*-sulfation, at least about 10%, and up to at least about 95%, of the glucosaminyl residues within *N*-deacetylated heparosan are *N*-sulfated, prior to subsequently being sulfated at any of the 2-*O*, 3-*O*, or 6-*O* positions.

[0032] In another aspect of the invention, a non-anticoagulant *N*-sulfated, 2-*O* sulfated, 6-*O* sulfated heparan sulfate polysaccharide (NS2S6S/HS) product can be synthesized by a method comprising the following steps: (a) providing a starting polysaccharide reaction mixture comprising *N*-

deacetylated heparosan; (b) reacting the starting polysaccharide composition within a reaction mixture comprising an *N*-sulfation agent, to form an NS/HS product; (c) combining the NS/HS product with a reaction mixture comprising a sulfo group donor and an intermediate sulfotransferase enzyme selected from the group consisting of a hexuronyl 2-*O* sulfotransferase (2OST) enzyme and a 6OST enzyme, to form an intermediate HS product; (d) combining the intermediate HS product with a reaction mixture comprising a finishing sulfotransferase enzyme, wherein the finishing sulfotransferase enzyme is the enzyme that was not selected in step (c), to form the NS2S6S/HS product; wherein (i) at least one of the sulfotransferase enzymes is an engineered sulfotransferase enzyme that is dependent on reacting with an aryl sulfate compound as a sulfo group donor to catalyze a sulfotransfer reaction, and (ii) in a reaction mixture comprising an engineered sulfotransferase enzyme, the reaction mixture consists of an aryl sulfate compound as a sulfo group donor. In various embodiments, the intermediate sulfotransferase enzyme is a 2OST enzyme and the finishing sulfotransferase enzyme is a 6OST enzyme. In various embodiments, all of the sulfotransferase enzymes are engineered aryl sulfate-dependent sulfotransferase enzymes, and each of the sulfotransfer reactions are performed in the absence of PAPS.

[0033] In various embodiments, an engineered 2OST enzyme utilized in any of the methods described herein can comprise any amino acid sequence so long as the enzyme catalyzes the transfer of a sulfo group from an aryl sulfate compound to the 2-*O* position of a hexuronic acid residue within a heparosan-based polysaccharide, particularly NS/HS. In various embodiments, the engineered 2OST enzymes can be mutants of natural sulfotransferases that have 2OST activity, which are members of enzyme class EC 2.8.2.-. In various embodiments, an engineered 2OST enzyme can comprise several amino acid mutations relative to one or more of the natural EC 2.8.2.- enzymes with 2OST activity, in order to engineer the active site to bind and react with an aryl sulfate compound as a sulfo group donor instead of PAPS.

[0034] Engineered 2OST enzymes utilized in accordance with any of the methods described herein can comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 41, and SEQ ID NO: 42, each of which contains several amino acid mutations made relative to highly conserved regions within natural 2OST enzymes. In various embodiments, engineered 2OST enzymes utilized in accordance with any of the methods described herein can also comprise an amino acid sequence having one or more amino residue differences or mutations from, and/or is a biological functional equivalent of, an amino acid sequence selected from the group consisting of SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 41, and SEQ ID NO: 42.

Non-limiting examples of such residue differences include amino acid insertions, deletions, substitutions, or any combination of such changes.

[0035] In various embodiments, any natural 2OST enzyme within EC 2.8.2.-, or a biologically-active fragment thereof, can be utilized to catalyze 2-*O* sulfation during the synthesis of HS products, particularly NS2S6S/HS products, in which engineered sulfotransferase enzymes are utilized to catalyze at least one other sulfation step, for example, the *N*-sulfation step and/or the 6-*O* sulfation step. Reaction mixtures comprising a natural 2OST enzyme also comprise PAPS as a sulfo group donor.

[0036] In various embodiments, NS/HS polysaccharides can be isolated and purified prior to reacting with the 2OST in a separate reaction mixture. In other embodiments, 2-*O* sulfation of hexuronic acid residues can take place in the same reaction mixture as the *N*-sulfation of *N*-deacetylated heparosan.

[0037] In various embodiments, a hexuronic acid residue that can receive a sulfo group from the 2OST enzyme can be either glucuronic acid or iduronic acid, and preferably iduronic acid, while other hexuronic acid residues within the polysaccharide can be glucuronic acid or iduronic acid, either of which can be 2-*O* sulfated. Both glucosamine residues adjacent to the hexuronic acid residue receiving the sulfo group can be, and preferably are, *N*-sulfated prior to reacting with the engineered or natural 2OST. In various embodiments, the sulfo acceptor polysaccharide reacting with a 2OST enzyme is NS/HS. In various embodiments, glucosamine residues that are not adjacent to the hexuronic acid residue receiving the sulfo group can optionally be *N*-, 3-*O*-, and/or 6-*O* sulfated, *N*-acetylated, or *N*-unsubstituted.

[0038] In various embodiments, the non-anticoagulant NS2S6S/HS product can be utilized directly in the treatment of a subject having a medical condition, including, as non-limiting examples, cancers; inflammation; thrombocytopenia; neutropenia; apoptosis; asthma; emphysema; bronchitis; adult respiratory distress syndrome; cystic fibrosis; and ischemia-reperfusion related conditions. In other embodiments, the non-anticoagulant NS2S6S/HS product can be modified by cold alkaline hydrolysis for a time sufficient to remove at least a portion of, and in some embodiments substantially all, of the 2-*O* sulfate groups from the non-anticoagulant NS2S6S/HS product, according to the methods described in Fryer, A. et al., 1997, *J. Pharmacol. Exp. Ther.* **282**: 208-219, and U.S. Patents 10,052,346 and 9,271,999.

[0039] In another aspect of the invention, an ODSH polysaccharide composition can be formed from an *N*-, 2-*O*-, 3-*O*-, 6-*O*-sulfated heparan sulfate (NS2S6S3S-HS) product that is synthesized by a method that utilizes at least one engineered, aryl sulfate-dependent sulfotransferase, the method

comprising the following steps: (a) providing a starting polysaccharide reaction mixture comprising *N*-deacetylated heparosan; (b) reacting the starting polysaccharide composition within a reaction mixture comprising an *N*-sulfation agent, to form an NS/HS product; (c) combining the NS/HS product with a reaction mixture comprising a sulfo group donor and a first intermediate sulfotransferase enzyme selected from the group consisting of a 2OST enzyme and a 6OST enzyme, to form a first intermediate HS product; (d) combining the first intermediate HS product with a reaction mixture comprising a second intermediate sulfotransferase enzyme, wherein the second intermediate sulfotransferase enzyme is the enzyme that was not selected in step (c), to form a second intermediate HS product; and (e) combining the second intermediate HS product with a reaction mixture comprising a sulfo group donor and a glucosaminyl 3-*O* sulfotransferase (3OST) enzyme, to form the NS2S6S3S-HS product; wherein (i) at least one of the sulfotransferase enzymes is an engineered sulfotransferase enzyme that is dependent on reacting with an aryl sulfate compound as a sulfo group donor to catalyze a sulfotransfer reaction, and (ii) in a reaction mixture comprising an engineered sulfotransferase enzyme, the reaction mixture consists of an aryl sulfate compound as a sulfo group donor. In various embodiments, the first intermediate sulfotransferase enzyme is a 2OST enzyme and the second intermediate sulfotransferase enzyme is a 6OST enzyme. In various embodiments, all of the sulfotransferase enzymes are engineered aryl sulfate-dependent sulfotransferase enzymes, and each of the sulfotransfer reactions are performed in the absence of PAPS. In various embodiments, the NS2S6S3S-HS product has anticoagulant activity. In various embodiments, the NS2S6S3S-HS product is substantially equivalent in molecular weight, purity, and anticoagulant activity to any of the heparin compounds described by CAS NO: 9005-49-6 or CAS NO: 9041-08-1. In various embodiments, once the NS2S6S3S-HS product is formed, it can be modified by cold alkaline hydrolysis for a time sufficient to remove at least a portion of, and in some embodiments substantially all, of the 2-*O* and 3-*O* sulfate groups to form an ODSH polysaccharide composition. In various embodiments, the NS2S6S3S-HS product can be modified by any method known in the art for forming an ODSH polysaccharide composition and/or until the synthesized ODSH polysaccharide composition is substantially equivalent to any ODSH polysaccharide composition described in the art. Such prior art compositions and methods for forming them are described above.

[0040] In various embodiments, an engineered 3OST enzyme utilized in any of the methods described herein can comprise any amino acid sequence so long as the enzyme catalyzes the transfer of a sulfo group from an aryl sulfate compound to the 3-*O* position of a glucosamine residue within a heparosan-based polysaccharide, particularly NS2S6S/HS polysaccharides. In various embodiments,

engineered 3OST enzymes can be mutants of natural sulfotransferases that have 3OST activity, which are members of enzyme class EC 2.8.2.23. In various embodiments, an engineered 3OST enzyme can comprise several amino acid mutations relative to one or more of the natural 3OST enzymes, in order to engineer the active site to bind and react with an aryl sulfate compound as a sulfo group donor instead of PAPS.

[0041] Engineered 3OST enzymes utilized in accordance with any of the methods described herein can comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58, each of which contains several amino acid mutations made relative to highly conserved regions within natural 3OST enzymes within EC 2.8.2.23. In various embodiments, engineered 3OST enzymes utilized in accordance with any of the methods described herein can also comprise an amino acid sequence having one or more amino residue differences or mutations from, and/or is a biological functional equivalent of, an amino acid sequence selected from the group consisting of SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58. Non-limiting examples of such residue differences include amino acid insertions, deletions, substitutions, or any combination of such changes. In various embodiments, the engineered 3OST enzyme comprises the amino acid sequence of SEQ ID NO: 28.

[0042] In various embodiments, any natural 3OST enzymes within EC 2.8.2.23, or a biologically-active fragment thereof, can be utilized to catalyze 3-*O* sulfation during the synthesis of NS2S6S3S-HS products, in which engineered sulfotransferase enzymes are utilized to catalyze the *N*-, 2-*O*, and/or 6-*O* sulfation of the polysaccharide. In various embodiments, reaction mixtures comprising a natural 3OST enzyme also comprise PAPS as a sulfo group donor. In various embodiments, an engineered 3OST enzyme is utilized to catalyze 3-*O* sulfation of an HS polysaccharide even if a natural HS sulfotransferase is utilized in one or more of the *N*-, 2-*O*, or 6-*O* sulfation steps to form the NS2S6S3S-HS product.

[0043] In various embodiments, NS2S6S/HS polysaccharides can be isolated and purified prior to reacting with the 3OST in a separate reaction mixture. In other embodiments, 3-*O* sulfation can take place in the same reaction mixture as the 6-*O* sulfation of NS2S/HS.

[0044] In various embodiments, glucosamine residues within the HS polysaccharide that can receive a sulfo group at the 3-*O* position are *N*-sulfated, and can optionally comprise a 6-*O* sulfo group as well. Any other glucosamine residue within the sulfo acceptor polysaccharide can be optionally be

N-, 3-*O*, and/or 6-*O* sulfated, *N*-acetylated, or *N*-unsubstituted. In various embodiments, one or more of the glucosamine residues within the HS polysaccharide, including the glucosamine residue being 3-*O* sulfated, can be both *N*-sulfated and 6-*O* sulfated. According to the present invention, the glucosamine residue being 3-*O* sulfated is adjacent to an unsulfated glucuronic acid residue at the non-reducing end and an iduronic acid residue, which can optionally be 2-*O* sulfated, at the reducing end. Any of the other hexuronic acid residues within the polysaccharide can optionally be iduronic acid or glucuronic acid, and can optionally be 2-*O* sulfated.

[0045] In various embodiments, aryl sulfate compounds that can be utilized as sulfo donors with any of the engineered sulfotransferase enzymes are organosulfates that comprise a sulfo group covalently bound to an aromatic moiety, bound together by a sulfate ester linkage comprising a C-O bond. Non-limiting examples of aryl sulfate compounds that are suitable substrates with the engineered enzymes of the present invention include *p*-nitrophenyl sulfate (PNS), 4-methylumbelliferyl sulfate (MUS), 7-hydroxycoumarin sulfate, phenyl sulfate, 4-acetylphenyl sulfate, indoxyl sulfate, 1-naphthyl sulfate, 2-naphthyl sulfate (2NapS), and 4-nitrocatechol sulfate (NCS). In various embodiments, engineered enzymes utilized in accordance with any of the methods of the present invention can recognize, bind, and react with PNS. In some embodiments, PNS can be used as the aryl sulfate compound in every sulfotransfer reaction during the synthesis of the HS polysaccharide product. In various embodiments, engineered enzymes utilized in accordance with any of the methods of the present invention can recognize, bind, and react with NCS. In some embodiments, NCS can be used as the aryl sulfate compound in every sulfotransfer reaction during the synthesis of the HS polysaccharide product. According to the present invention, a single engineered enzyme utilized in accordance with any of the methods of the present invention can recognize, bind, and react with multiple aryl sulfate compounds.

[0046] In various embodiments, each of the engineered sulfotransferase enzymes utilized in any of the methods described herein can be selected to react with the same aryl sulfate compound as a sulfo group donor. In other embodiments, one or more of the engineered sulfotransferase enzymes can have a biological activity with different aryl sulfate compounds than other enzymes utilized in the same synthesis. As a non-limiting example, in syntheses in which multiple sulfotransfer reactions occur in a single reaction mixture, both PNS and NCS can be included within the reaction mixture.

[0047] In various embodiments, a glucuronyl C5-epimerase enzyme can be added into the reaction mixture for any of the sulfation steps within any of the methods described herein. In a non-limiting example, and in some embodiments, the glucuronyl C5-epimerase enzyme can comprise the amino acid sequence of SEQ ID NO: 29, preferably residues 34-617 of SEQ ID NO: 29. In a non-limiting

example, a glucuronyl C5-epimerase enzyme comprising either the amino acid sequence of SEQ ID NO: 29 or residues 34-617 of SEQ ID NO: 29 can be included within a reaction mixture comprising *N*-sulfated heparosan and an engineered or natural 6OST, to form NS6S/HS polysaccharides comprising one or more IdoA residues. In another non-limiting example, a glucuronyl C5-epimerase enzyme comprising either the amino acid sequence of SEQ ID NO: 29 or residues 34-617 of SEQ ID NO: 29 can be included within a reaction mixture comprising an engineered or natural 2OST, during the formation of an NS2S6S/HS or NS2S6S3S-HS product.

[0048] In various embodiments, within any reaction mixture or composition comprising heparosan-based polysaccharides, whether used as starting materials or formed as products, the reaction mixture or compositions can be a polydisperse mixture of heparosan-based polysaccharides having variable chain lengths, molecular weights, *N*-acetylation, and/or *N*-, 2-*O*, 6-*O*, or 3-*O* sulfation. Alternatively, any of the polysaccharides described above can be present or provided as a homogeneous composition comprised of polysaccharides having identical chain lengths, molecular weights, *N*-acetylation, and/or *N*-, 2-*O*, 6-*O*, or 3-*O* sulfation.

[0049] In various embodiments, heparosan-based polysaccharides that can be used as sulfo group acceptors in any of the sulfotransfer reactions described herein can generally be any molecular weight greater than 1,000 Da, including greater than 1,000,000 Da. In various embodiments, compositions or mixtures comprising *N*-deacetylated heparosan polysaccharides can preferably have a weight-average molecular weight in the range of at least 9,000 Da, and up to 12,500 Da. In various embodiments, sulfated polysaccharide products of any of the reactions described herein any of the methods described above can comprise molecular weights associated with the addition of a single sulfo group (about 80 Da), and up to the addition of sulfo groups to all available *N*, 2-*O*, 3-*O*, and/or 6-*O* positions, based on the molecular weight of the polysaccharide used as the sulfo group acceptor.

[0050] In various embodiments, in any of the methods described herein, any reaction mixture comprising an engineered sulfotransferase enzyme and an aryl sulfate compound can further comprise one or more components for repopulating the aryl sulfate compound. In various embodiments, the one or more components for repopulating the aryl sulfate compound can comprise an aryl sulfate sulfotransferase (ASST) enzyme and a secondary aryl sulfate compound. According to the present invention, the engineered sulfotransferase enzyme has minimal or no activity with the secondary aryl sulfate compound as a sulfo group donor. The ASST enzyme from any bacteria can be utilized, and can either be isolated from the bacteria directly or generated recombinantly from an

expression host *in vitro*. In various embodiments, the ASST enzyme can be a recombinant ASST from *E. coli* strain CFT073, comprising the amino acid sequence of SEQ ID NO: 55.

[0051] In one non-limiting example, a reaction mixture comprising an NS/HS product, NCS, and an engineered 6OST enzyme comprising the amino acid sequence SEQ ID NO: 18 can further comprise an ASST enzyme and PNS, with which the engineered enzyme comprising the amino acid sequence SEQ ID NO: 18 is not active. Upon being formed as a product of the sulfotransfer reaction, 4-nitrocatechol can then act as a sulfo group acceptor for a reaction between PNS and the ASST enzyme, thereby reforming NCS for subsequent reactions with the engineered enzyme comprising the amino acid sequence SEQ ID NO: 18. Alternatively, the NCS utilized for the sulfotransfer reaction to form an NS6S/HS product can be generated *in situ* by forming a reaction mixture comprising the engineered 6OST enzyme comprising the amino acid sequence SEQ ID NO: 18, 4-nitrocatechol, PNS, and an ASST enzyme.

[0052] In various embodiments, the anticoagulant effect of antithrombin activation can be quantified, particularly as a function of its subsequent effect on the activity of Factor IIa and Factor Xa, in terms of International Units of activity per milligram (IU mg⁻¹). In various embodiments, HS polysaccharides made by methods of the present invention can have an anticoagulant activity of less than about 9 IU mg⁻¹. In various embodiments, HS polysaccharides made by methods of the present invention can have an anti-Factor Xa (anti-Xa) activity of less than 5 IU mg⁻¹. In various embodiments, HS polysaccharides made by methods of the present invention can have an anti-Factor IIa (anti-IIa) activity of less than about 2 IU mg⁻¹. In various embodiments, HS polysaccharides made by methods of the present invention, particularly methods to synthesize NS/HS, NS2S/HS, NS6S/HS, and NS2S6S/HS in the absence of a 3OST enzyme, have no anti-IIa or anti-Xa activity.

[0053] In various embodiments, any of the HS product mixtures produced by any of the methods above can have an average molecular weight of at least about 1,500 Da, depending on the weight average molecular weight of polysaccharides utilized as sulfo group acceptors, as described above. In various embodiments, non-anticoagulant HS products, particularly NS6S/HS products, can have a weight-average molecular weight in the range of about 2,000 Da to about 15,000 Da. In embodiments which an ODSH product is generated by modifying an NS2S6S3S-HS product using cold alkaline hydrolysis, the NS2S6S3S-HS product can have a molecular weight profile such that: (a) the weight-average molecular weight of the NS2S6S3S-HS product is at least 15,000 Da, and up to 19,000 Da; (b) less than or equal to 20% of the polysaccharides within the NS2S6S3S-HS product have a molecular weight greater than 24,000 Da; and (c) the number of polysaccharide chains within the NS2S6S3S-HS product having a molecular weight between 8,000 Da and 16,000 Da is greater

than the number of polysaccharide chains having a molecular weight between 16,000 Da and 24,000 Da.

[0054] In various embodiments, engineered sulfotransferase enzymes having biological activity with aryl sulfate compounds as sulfo group donors can be expressed from a nucleic acid comprising a nucleotide sequence that encodes for any of the amino acid sequences described above. Non-limiting examples of such nucleotide sequences include SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, and SEQ ID NO: 27. Persons skilled in the art can determine appropriate nucleotide sequences that encode for polypeptides having the amino acid sequence of SEQ ID NOs: 33-54 and 56-61, based on the nucleotide sequences above.

[0055] In another aspect of the invention, any of the HS products produced by any of the methods described above can be further modified by one or more subsequent processes to depolymerize and/or modify the HS product to form a low molecular weight (LMW)-HS product. In various embodiments, a non-anticoagulant LMW-HS composition can be synthesized from compositions comprising non-anticoagulant NS/HS, NS2S/HS, NS6S/HS, NS2S6S/HS, or NS2S6S3S-HS. In various embodiments, the LMW-HS composition is an LMW-NS6S/HS composition. In various embodiments, an anticoagulant LMW-HS composition can be synthesized from anticoagulant NS2S6S3S-HS and subsequently subjected to cold alkaline hydrolysis to form LMW-ODSH. In other embodiments, an anticoagulant NS2S6S3S-HS composition can be modified using cold alkaline hydrolysis to form an ODSH composition, which can subsequently depolymerized to form LMW-ODSH.

[0056] In various embodiments, an HS product produced by any method described above can be referred to as an “unfractionated” HS product, relative to an LMW-HS product or LMW-ODSH product. Unfractionated HS products can include one or more non-anticoagulant NS/HS, NS2S/HS, NS6S/HS, NS2S6S/HS, or NS2S6S3S-HS products, and/or anticoagulant NS2S6S3S-HS.

[0057] Generally, methods of the present invention for synthesizing an LMW-HS or LMW-ODSH product can comprise the following steps: (a) synthesizing an unfractionated HS product according to any of the above methods; (b) providing one or more depolymerization agents; and (c) treating the unfractionated HS product with the one or more depolymerization agents for a time sufficient to depolymerize at least a portion of the polysaccharides within the unfractionated HS product, thereby forming the LMW-HS or LMW-ODSH product. In various embodiments, the weight-average

molecular weight of the LMW-HS or LMW-ODSH product is at least 2,000 Da, and up to 12,000 Da, and preferably at least 3,000 Da, and up to 8,000 Da.

[0058] In various embodiments, the one or more depolymerization agents can be formed by, and/or be comprised of, one or more reaction components within one or more reaction mixtures, that can be combined with an unfractionated HS product to chemically and/or enzymatically depolymerize the unfractionated HS product and form an LMW-HS or LMW-ODSH product. In various embodiments, the selection of the depolymerization agent can determine which chemical or enzymatic depolymerization process occurs, as well as the chemical structure and/or anticoagulant activity of the depolymerized product. Such depolymerization processes can include, but are not limited to: chemical and/or enzymatic β -elimination reactions; deamination reactions; and oxidation reactions, including combinations thereof. In various embodiments, an unfractionated HS product can be treated with any combination of depolymerization agents in order to form an LMW-HS or LMW-ODSH product.

[0059] In various embodiments, the amount of time that an unfractionated HS product is treated with the one or more depolymerization agents can be controlled to form an LMW-HS or LMW-ODSH product with a desired molecular weight and/or chemical structure. According to the present invention, with respect to the same depolymerization agent, the amount of time that an unfractionated HS product is treated with the depolymerization agent can be varied to form products with similar chemical structures, but different molecular weights relative to each other.

[0060] In one-non-limiting example, an unfractionated HS product can be depolymerized by an enzymatic β -elimination reaction to form an LMW-HS or LMW-ODSH product. In various embodiments, the depolymerization agent can comprise a heparinase reaction mixture comprising at least one heparinase enzyme, preferably at least one heparinase enzyme comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 30, SEQ ID NO: 31, and SEQ ID NO: 32. In various embodiments, the unfractionated HS product can be treated with the heparinase reaction mixture for a time sufficient to catalyze β -eliminative cleavage of the unfractionated HS product and form an enzymatically-depolymerized LMW-HS or LMW-ODSH product. In various embodiments, the weight-average molecular weight of the enzymatically-depolymerized LMW-HS or LMW-ODSH product can be in the range of 2,000 Da to 10,000 Da, preferably 5,500 Da to 7,500 Da, and more preferably 6,500 Da. In various embodiments, the enzymatically-depolymerized LMW-HS or LMW-ODSH product can comprise polysaccharides having a 4,5-unsaturated uronic acid residue at the non-reducing end. In various embodiments, prior to 2-O and 3-O desulfation to form LMW-ODSH, an anticoagulant enzymatically-depolymerized

LMW-HS product can comprise a substantially equivalent chemical structure, weight-average molecular weight, and/or anticoagulant activity relative to tinzaparin.

[0061] In another non-limiting example, an unfractionated HS product can be depolymerized by a chemical β -elimination reaction. In various embodiments, the depolymerization agent for a chemical β -elimination reaction can comprise a base, preferably a base selected from the group consisting of sodium hydroxide, a quaternary ammonium hydroxide, and a phosphazene base, including any combination thereof, and the unfractionated HS product can be treated with the base for a time sufficient to cause β -eliminative cleavage of the unfractionated HS product and form a chemically β -eliminative, LMW-HS or LMW-ODSH product.

[0062] In various embodiments, a benzethonium salt can be formed prior to reacting the unfractionated HS product with the base. Accordingly, in various embodiments, the step of treating the unfractionated HS product with the depolymerization agent can comprise the following sub-steps: (i) reacting the unfractionated HS product with a benzethonium salt, preferably benzethonium chloride, to form a benzethonium HS salt; and (ii) combining the benzethonium HS salt with a reaction mixture comprising the base for a time sufficient to form the chemically β -eliminative, LMW-HS or LMW-ODSH product. In various embodiments, the weight-average molecular weight of the chemically β -eliminative, LMW-HS or LMW-ODSH product can be at least 2,000 Da, up to 10,000 Da, and preferably in the range of 2,000 Da to 6,000 Da. In various embodiments, the chemically β -eliminative, LMW-HS or LMW-ODSH product can comprise polysaccharides having a 4,5-unsaturated uronic acid residue at the non-reducing end.

[0063] In various embodiments, once the benzethonium HS salt is formed, it can be subsequently treated with a base for a time sufficient to form the chemically β -eliminative, LMW-HS or LMW-ODSH product. In various embodiments, the base can be a quaternary ammonium hydroxide, preferably benzyl trimethylammonium hydroxide (Triton[®] B). In various embodiments, the weight-average molecular weight of the chemically β -eliminative, LMW-HS or LMW-ODSH product can be in the range of 3,000 Da to 4,200 Da, and preferably 3,600 Da. In various embodiments, prior to 2-*O* and 3-*O* desulfation, an anticoagulant enzymatically-depolymerized LMW-HS product can comprise a substantially equivalent chemical structure, weight-average molecular weight, and/or anticoagulant activity relative to bemiparin.

[0064] In various embodiments, the benzethonium HS salt can be further modified prior to reacting with the base. In one non-limiting example, the benzethonium HS salt can be converted to a benzyl ester form of HS upon reacting with a benzyl halide, particularly benzyl chloride. In various

embodiments, the conversion to the benzyl ester can take place within a chlorinated solvent, including but not limited to methylene chloride and chloroform.

[0065] In various embodiments, once the benzyl ester HS is formed, it can be subsequently reacted with a base to initiate depolymerization. In various embodiments, the base can be sodium hydroxide. In various embodiments, the chemically β -eliminative, LMW-HS or LMW-ODSH product can comprise polysaccharides having a 1,6-anhydromannose or 1,6-anhydroglucosamine residue at the reducing end in addition to the 4,5-unsaturated uronic acid residue at the non-reducing end. In various embodiments, the weight-average molecular weight of the chemically β -eliminative, LMW-HS or LMW-ODSH product can be in the range of 3,800 Da to 5,000 Da, preferably 4,500 Da. In various embodiments, prior to 2-*O* and 3-*O* desulfation to form LMW-ODSH, an anticoagulant enzymatically-depolymerized LMW-HS product can comprise a substantially equivalent chemical structure, weight-average molecular weight, and/or anticoagulant activity relative to enoxaparin.

[0066] In various embodiments, the benzyl ester HS can instead be transalified in the presence of a benzethonium salt, preferably benzethonium chloride, in order to form a benzethonium benzyl ester HS, which can then be subsequently depolymerized using a base. In various embodiments, the base is a phosphazene base, preferably 2-*tert*-butylimino-2-diethylamino-1,3-dimethylperhydro-1,2,3-diaza-phosphorine (BEMP). After depolymerization is complete, the remaining benzyl esters within the chemically β -eliminative, LMW-HS or LMW-ODSH product can be saponified and removed. In various embodiments, the weight-average molecular weight of the chemically β -eliminative, LMW-HS or LMW-ODSH product can be in the range of 2,000 Da to 3,000 Da, and is preferably 2,400 Da. In various embodiments, prior to 2-*O* and 3-*O* desulfation to form LMW-ODSH, an anticoagulant enzymatically-depolymerized LMW-HS product can comprise a substantially equivalent chemical structure, weight-average molecular weight, and/or anticoagulant activity relative to semuloparin.

[0067] In various embodiments, unfractionated HS products can optionally be depolymerized by both an enzymatic and a chemical β -elimination reaction. For example, an enzymatically-depolymerized LMW-HS or LMW-ODSH product can subsequently be subjected to a chemical β -elimination reaction by reacting with a base. In another example, a chemically β -eliminative, LMW-HS product can subsequently be subjected to an enzymatic β -elimination reaction by reacting one or more heparinase enzymes.

[0068] In another non-limiting example, an unfractionated HS product can be depolymerized by a deamination reaction. In various embodiments, the depolymerization agent can comprise a deamination reaction mixture comprising a deamination agent, preferably a deamination agent selected from the group consisting of isoamyl nitrate and nitrous acid, for a time sufficient to cause

deaminative cleavage of the unfractionated HS product, thereby forming a deaminated LMW-HS or LMW-ODSH product.

[0069] In various embodiments, the deamination agent can be nitrous acid. In various embodiments, the deamination reaction mixture can comprise stoichiometric quantities of an acid, preferably acetic acid or hydrochloric acid, and an alkali or alkaline earth metal nitrite salt, preferably sodium nitrite, to form nitrous acid *in situ*. In various embodiments, the deaminated LMW-HS or LMW-ODSH product can comprise polysaccharides having a 2,5-anhydro-D-mannose residue at the reducing end. In various embodiments, the weight-average molecular weight of the deaminated LMW-HS or LMW-ODSH product can be in the range of 2,000 Da to 10,000 Da, preferably in the range of 4,000 Da to 6,000 Da.

[0070] In one non-limiting example, the weight-average molecular weight of the deaminated LMW-HS or LMW-ODSH product can be in the range of 3,600 Da to 5,500 Da, preferably 4,300 Da. In various embodiments, prior to 2-*O* and 3-*O* desulfation to form LMW-ODSH, an anticoagulant deaminated LMW-HS product can comprise a substantially equivalent chemical structure, weight-average molecular weight, and/or anticoagulant activity relative to nadroparin.

[0071] In another non-limiting example, the weight-average molecular weight of the deaminated LMW-HS or LMW-ODSH product can be in the range of 5,600 Da to 6,400 Da, preferably 6,000 Da. In various embodiments, prior to 2-*O* and 3-*O* desulfation to form LMW-ODSH, an anticoagulant deaminated LMW-HS product can comprise a substantially equivalent chemical structure, weight-average molecular weight, and/or anticoagulant activity relative to dalteparin.

[0072] In another non-limiting example, the weight-average molecular weight of the deaminated LMW-HS or LMW-ODSH product can be in the range of 4,200 Da to 4,600 Da, preferably 4,400 Da. In various embodiments, prior to 2-*O* and 3-*O* desulfation to form LMW-ODSH, an anticoagulant deaminated LMW-HS product can comprise a substantially equivalent chemical structure, weight-average molecular weight, and/or anticoagulant activity relative to reviparin.

[0073] In another non-limiting example, the deamination agent is isoamyl nitrate, and the weight-average molecular weight of the deaminated LMW-HS or LMW-ODSH product can be in the range of 5,000 Da to 5,600 Da, preferably 5,400 Da. In various embodiments, prior to 2-*O* and 3-*O* desulfation to form LMW-ODSH, an anticoagulant deaminated LMW-HS product can comprise a substantially equivalent chemical structure, weight-average molecular weight, and/or anticoagulant activity relative to certoparin.

[0074] In another non-limiting example, an unfractionated HS product can be depolymerized by an oxidation reaction. In various embodiments, the depolymerization agent can comprise an oxidation

agent, preferably an oxidation agent selected from the group consisting of a peroxide or a superoxide, and more preferably hydrogen peroxide to form an oxidized LMW-HS or LMW-ODSH product. In various embodiments, the step of treating an unfractionated HS product with the oxidation agent can comprise the following sub-steps: (i) acidifying the unfractionated HS product to form an acidified HS product; (ii) combining the acidified HS product with the oxidation reaction mixture; and (iii) incubating the acidified HS product within the oxidation reaction mixture at a temperature of at least 50 °C for a time sufficient to form the oxidized LMW-HS or LMW-ODSH product.

[0075] In various embodiments, the sub-step of acidifying the unfractionated HS product can comprise the addition of a reaction mixture comprising an acid, preferably ascorbic acid, to the HS product to form the acidified HS product. Alternatively, the sub-step of acidifying the unfractionated HS product can further comprise the sub-steps of: loading the unfractionated HS product into a cation exchange resin, preferably a cation exchange resin suspended within a chromatography column; and eluting the unfractionated HS product from the cation exchange resin, forming the acidified HS product. In various embodiments, the pH of the acidified HS product can be at least 3.0, and up to 5.0, and preferably in a range of 3.0 to 3.5.

[0076] In various embodiments, the weight-average molecular weight of the oxidized LMW-HS or LMW-ODSH product can be in the range of 2,000 Da to 12,000 Da, preferably in the range of 4,000 Da to 6,000 Da.

[0077] In one non-limiting example, the weight-average molecular weight of the oxidized LMW-HS or LMW-ODSH product can be in the range of at least 4,000 Da up to 6,000 Da, and is preferably 5,000 Da. In various embodiments, prior to 2-*O* and 3-*O* desulfation to form LMW-ODSH, an anticoagulant oxidized LMW-HS product can comprise a substantially equivalent chemical structure, weight-average molecular weight, and/or anticoagulant activity relative to parnaparin.

[0078] In another non-limiting example, the weight-average molecular weight of the oxidized LMW-HS product can be in the range of 5,500 Da to 6,500 Da, preferably 6,000 Da. In various embodiments, prior to 2-*O* and 3-*O* desulfation to form LMW-ODSH, an anticoagulant oxidized LMW-HS product can comprise a substantially equivalent chemical structure, weight-average molecular weight, and/or anticoagulant activity relative to ardeparin.

[0079] According to the present invention, and useful in combination with any one or more of the above aspects and embodiments, any of the non-anticoagulant or anticoagulant HS products and/or LMW-HS or LMW-ODSH products prepared according to any of the methods described above can be prepared as pharmaceutically-acceptable salts, particularly alkali or alkali earth salts including, but not limited to, sodium, lithium, or calcium salts.

[0080] These and other embodiments of the present invention will be apparent to one of ordinary skill in the art from the following detailed description.

BRIEF DESCRIPTION OF THE FIGURES

[0081] Figures 1A-1C show an example reaction mechanism between the human 3OST enzyme, PAPS, and an *N*-sulfated, 6-*O* sulfated glucosamine residue within heparan sulfate.

[0082] Figure 2 shows a non-limiting example of an *N*-deacetylated heparosan polysaccharide capable of reacting as a sulfo group acceptor for both natural NDST enzymes and engineered NST enzymes that can be used in accordance with methods of the present invention.

[0083] Figures 3A-3C show a multiple sequence alignment for the *N*-sulfotransferase domains of fifteen natural NDST enzymes within enzyme class EC 2.8.2.8, illustrating conserved amino acid sequence motifs that are present regardless of overall sequence identity.

[0084] Figures 4A-4C show a reaction mechanism between conserved residues within the *N*-sulfotransferase domain of a natural NDST enzyme, PAPS, and *N*-deacetylated heparosan.

[0085] Figure 5 shows a three-dimensional model of an aryl sulfate compound bound within the active site of a first group of engineered NST enzymes, superimposed over the crystal structure of PAPS bound within the *N*-sulfotransferase domain of a natural human NDST enzyme.

[0086] Figure 6 shows an alternate view of the modelled active site of the engineered NST enzyme shown in Figure 5, illustrating amino acid mutations present within the active site.

[0087] Figure 7 shows a three-dimensional model of an aryl sulfate compound bound within the active site of a second group of engineered NST enzymes, superimposed over the crystal structure of the *N*-sulfotransferase domain of a natural human NDST enzyme.

[0088] Figure 8 shows an alternate view of the modelled active site of the engineered NST enzyme shown in Figure 7, illustrating amino acid mutations present within the active site.

[0089] Figure 9 shows a sequence alignment of polypeptides comprising the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, respectively, depicting the position and identity of amino acid residues differences between each of the illustrated sequences and relative to the human NDST1 enzyme.

[0090] Figure 10 shows the 2-*O* sulfation of one non-limiting example of an *N*-sulfated heparosan polysaccharide, catalyzed by either a natural or engineered 2OST enzyme in accordance with methods of the present invention, wherein the polysaccharide comprises *N*-sulfated, *N*-acetylated, and unsubstituted glucosaminy residues.

[0091] Figure 11 shows the 2-*O* sulfation of a glucuronic acid residue within another non-limiting example of an *N*-sulfated heparosan polysaccharide, catalyzed by either a natural or engineered 2OST enzyme in accordance with methods of the present invention.

[0092] Figure 12 shows the 2-*O* sulfation of an iduronic acid residue within the *N*-sulfated heparosan polysaccharide shown in Figure 11, catalyzed by either a natural or engineered 2OST enzyme in accordance with methods of the present invention.

[0093] Figure 13 shows the 2-*O* sulfation of a glucuronic acid residue and an iduronic acid residue within the *N*-sulfated heparosan polysaccharide shown in Figure 11, catalyzed by either a natural or engineered 2OST enzyme in accordance with methods of the present invention.

[0094] Figures 14A-14D show a multiple sequence alignment for twelve natural 2OST enzymes within EC 2.8.2.-, illustrating conserved amino acid sequence motifs that are present regardless of overall sequence identity.

[0095] Figures 15A-15C show a reaction mechanism between conserved residues within a natural 2OST enzyme, PAPS, and a hexuronic acid residue within *N*-sulfated heparosan.

[0096] Figure 16 shows a three-dimensional model of an aryl sulfate compound bound within the active site of an engineered 2OST enzyme, superimposed over the crystal structure of PAPS bound within the active site of the chicken 2OST enzyme.

[0097] Figure 17 shows the 6-*O* sulfation of one non-limiting example of an *N*-sulfated, 2-*O* sulfated heparan sulfate polysaccharide, catalyzed by either a natural or engineered 6OST enzyme in accordance with methods of the present invention, wherein multiple glucosamine residues within the polysaccharide are capable of receiving a sulfate group.

[0098] Figures 18A-18C show a multiple sequence alignment for fifteen natural 6OST enzymes within EC 2.8.2.-, illustrating conserved amino acid sequence motifs that are present regardless of overall sequence identity.

[0099] Figures 19A-19C show a reaction mechanism between conserved residues within a natural 6OST enzyme, PAPS, and an *N*-sulfated glucosamine residue within heparan sulfate.

[0100] Figure 20 shows a three-dimensional model of an aryl sulfate compound bound within the active site of an engineered 6OST enzyme, superimposed over the crystal structure of PAPS bound within the zebrafish 6OST3 enzyme.

[0101] Figure 21 shows a sequence alignment of polypeptides comprising the amino acid sequences of SEQ ID NO: 18, SEQ ID NO: 20, and SEQ ID NO: 22, respectively, depicting the position and identity of amino acid residues differences between each of the illustrated sequences and relative to the mouse 6OST1 enzyme.

[0102] Figure 22 shows the 3-*O* sulfation of one non-limiting example of an *N*-sulfated, 2-*O* sulfated, 6-*O* sulfated heparan sulfate polysaccharide, catalyzed by either a natural or engineered 3OST enzyme in accordance with methods of the present invention.

[0103] Figures 23A-23C show a multiple sequence alignment for fifteen natural 3OST enzymes within EC 2.8.2.23, illustrating conserved amino acid sequence motifs that are present regardless of overall sequence identity.

[0104] Figure 24 shows a three-dimensional model of an aryl sulfate compound bound within the active sites of three superimposed engineered 3OST enzymes, which themselves are superimposed over the crystal structure of PAPS bound within the mouse 3OST1 enzyme.

[0105] Figure 25 shows a sequence alignment of polypeptides comprising the amino acid sequences of SEQ ID NO: 24, SEQ ID NO: 26, and SEQ ID NO: 28, respectively, depicting the position and identity of amino acid residues differences between each of the illustrated sequences and relative to the mouse 3OST1 enzyme.

[0106] Figure 26 shows a series of overlaid SAX-HPLC chromatograms of digested *N*-sulfated heparosan products synthesized using an engineered NST enzyme, compared to commercial standards.

[0107] Figures 27A-27B show a series of LCMS chromatograms of digested *N*-, 2-*O*-sulfated polysaccharide products synthesized using an engineered 2OST having the amino acid sequence SEQ ID NO: 14 or SEQ ID NO: 16, respectively.

[0108] Figures 28A-28C shows an LCMS chromatogram of digested *N*-, 2-*O*-, 6-*O*-sulfated polysaccharide products synthesized using an engineered 6OST having the amino acid sequence SEQ ID NO: 18, SEQ ID NO: 20, or SEQ ID NO: 22, respectively.

[0109] Figures 29A-29B show a series of overlaid LCMS chromatograms of digested *N*-, 2-*O*-, 6-*O*-, 3-*O*-sulfated polysaccharide products synthesized using engineered 3OST enzymes having the amino acid sequence SEQ ID NO: 24, SEQ ID NO: 26, or SEQ ID NO: 28, compared to a series of disaccharide and polysaccharide standards.

[0110] Figure 30 shows the reaction scheme for deuterium labeling of protons of interest for nuclear magnetic resonance (NMR) studies.

[0111] Figure 31 shows an expanded view of ¹H-NMR spectra for engineered 3OST enzymes having the amino acid sequence SEQ ID NO: 24, SEQ ID NO: 26, or SEQ ID NO: 28, either with PNS or NCS.

[0112] Figure 32 shows a magnified view of the 3.5ppm to 4.5ppm region of the ¹H-NMR spectra illustrated in Figure 31.

[0113] Figure 33 shows a SAX-HPLC chromatogram of a chemically *N*-sulfated polysaccharide product, compared to a commercial standard.

[0114] Figure 34 shows a SAX-HPLC chromatogram of an enzymatically 2-*O* sulfated polysaccharide product prepared using the chemically *N*-sulfated polysaccharide product of Example 7 as the sulfo acceptor polysaccharide, compared to a commercial standard.

[0115] Figure 35 shows a SAX-HPLC chromatogram of an enzymatically 2-*O* sulfated polysaccharide product prepared using the chemically *N*-sulfated polysaccharide product of Example 7 as the sulfo acceptor polysaccharide and with a C5-hexuronyl epimerase included in the reaction mixture, compared to a commercial standard.

[0116] Figure 36 shows a SAX-HPLC chromatogram of an enzymatically 6-*O* sulfated polysaccharide product prepared using the sulfated polysaccharide product of Example 8 as the sulfo group acceptor, compared to a commercial standard.

[0117] Figure 37 shows a SAX-HPLC chromatogram of an enzymatically 6-*O* sulfated polysaccharide product prepared using the chemically *N*-sulfated polysaccharide product of Example 7 as the sulfo group acceptor, compared to a commercial standard.

DEFINITIONS

[0118] The term, “active site,” refers to sites in catalytic proteins, in which catalysis occurs, and can include one or more substrate binding sites. Active sites are of significant utility in the identification of compounds that specifically interact with, and modulate the activity of, a particular polypeptide. The association of natural ligands or substrates with the active sites of their corresponding receptors or enzymes is the basis of many biological mechanisms of action. Similarly, many compounds exert their biological effects through association with the active sites of receptors and enzymes. Such associations may occur with all or any parts of the active site. An understanding of such associations helps lead to the design of engineered active sites within sulfotransferases that are capable of binding to and reacting with aryl sulfate compounds instead of PAPS.

[0119] The term, “amino acid,” refers to a molecule having the structure wherein a central carbon atom (the alpha-carbon atom) is linked to a hydrogen atom, a carboxylic acid group (the carbon atom of which is referred to herein as a “carboxyl carbon atom”), an amino group (the nitrogen atom of which is referred to herein as an “amino nitrogen atom”), and a side chain group, R. When incorporated into a peptide, polypeptide, or protein, an amino acid loses one or more atoms of its amino and carboxylic groups in the dehydration reaction that links one amino acid to another. As a result, when incorporated into a protein, an amino acid is referred to as an “amino acid residue.” In

the case of naturally occurring proteins, an amino acid residue's R group differentiates the 20 amino acids from which proteins are synthesized, although one or more amino acid residues in a protein may be derivatized or modified following incorporation into protein in biological systems (e.g., by glycosylation and/or by the formation of cysteine through the oxidation of the thiol side chains of two non-adjacent cysteine amino acid residues, resulting in a disulfide covalent bond that frequently plays an important role in stabilizing the folded conformation of a protein, etc.). Additionally, when an alpha-carbon atom has four different groups (as is the case with the 20 amino acids used by biological systems to synthesize proteins, except for glycine, which has two hydrogen atoms bonded to the carbon atom), two different enantiomeric forms of each amino acid exist, designated D and L. In mammals, only L-amino acids are incorporated into naturally occurring polypeptides. Engineered sulfotransferase enzymes utilized in accordance with methods of the present invention can incorporate one or more D- and L-amino acids, or can be comprised solely of D- or L-amino acid residues.

[0120] Non-naturally occurring amino acids can also be incorporated into any of the sulfotransferase enzymes utilized in accordance with the methods of the present invention, particularly engineered sulfotransferase enzymes having aryl sulfate-dependent activity. Examples of such amino acids include, without limitation, alpha-amino isobutyric acid, 4-amino butyric acid, L-amino butyric acid, 6-amino hexanoic acid, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, cysteic acid, t-butyl glycine, t-butyl alanine, phenylglycine, cyclohexyl alanine, beta-alanine, fluoro-amino acids, designer amino acids (e.g., beta-methyl amino acids, alpha-methyl amino acids, alpha-methyl amino acids) and amino acid analogs in general.

[0121] The term, “and/or,” when used in the context of a listing of entities, refers to the entities being present singly or in combination. For example, the phrase “A, B, C, and/or D” includes A, B, C, and D individually, but also includes any and all combinations and sub-combinations of A, B, C, and D.

[0122] The term, “API heparin,” refers to the form of heparin that is regulated for administering to patients, and which conforms to the United States Pharmacopeia (USP) reference standard with respect to identity, strength, quality, purity, and potency. Properties defined by the USP monograph for heparin sodium include: a characteristic ¹H-NMR spectrum; chromatographic purity, particularly with respect to dermatan sulfate and oversulfated chondroitin sulfate; anti-Factor Xa activity; anti-Factor IIa activity; the ratio of anti-factor Xa activity relative to anti-factor IIa activity; the presence or absence of inorganic and inorganic impurities; and a characteristic molecular weight distribution

or profile. In particular, the USP Heparin Sodium standard has an anti-Factor Xa activity of not less than 180 IU mg⁻¹; an anti-factor IIa activity of not less than 180 IU mg⁻¹; a ratio of anti-Factor Xa activity to anti-Factor IIa activity of 0.9-1.1; the amount of polysaccharide chains greater than 24,000 Da is less than 20% of a heparin sample; the amount of polysaccharide chains between 8,000 Da and 16,000 Da being greater than the amount of polysaccharide chains between 16,000 Da and 24,000 Da within the heparin sample; and a weight average molecular weight of the heparin sample in the range of at least 15,000 Da and up to 19,000 Da

[0123] The terms, “aryl sulfate” or “aryl sulfate compound,” refer to any compound, functional group, or substituent derived from an aromatic ring in which one or more of the hydrogen atoms directly bonded to the aromatic ring is replaced by a sulfate functional group. Typically, the sulfate functional group is covalently bound to the aromatic moiety of an aryl sulfate compound through a sulfate ester linkage. Exemplary aryl sulfate compounds that can donate a sulfo group to a polysaccharide, particularly a heparosan-based polysaccharide, using any of the engineered sulfotransferases include, but are not limited to, *p*-nitrophenyl sulfate (PNS), 4-methylumbelliferyl sulfate (MUS), 7-hydroxycoumarin sulfate, phenyl sulfate, 4-acetylphenyl sulfate, indoxyl sulfate, 1-naphthyl sulfate, 2-naphthyl sulfate, and 4-nitrocatechol sulfate (NCS).

[0124] The term, “aryl sulfate-dependent sulfotransferase,” refers to the collective group of engineered sulfotransferases that possess biological or catalytic activity with aryl sulfate compounds as sulfo donors. Non-limiting examples of aryl sulfate compounds upon which the biological activity of the sulfotransferase can be dependent include PNS and NCS. As described herein, engineered sulfotransferases having biological activity with aryl sulfate compounds as sulfo group donors can possess biological activity with polysaccharides, particularly heparosan-based polysaccharides, as sulfo group acceptors. “Aryl sulfate-dependent sulfotransferase” also includes both nucleic acids and polypeptides encoding for any aryl sulfate-dependent sulfotransferase, including mutants derived from the sequences disclosed herein.

[0125] The term, “average molecular weight,” with respect to any of the polysaccharide starting materials, intermediates, and/or products used or generated according to any of the methods of the present invention, and unless otherwise indicated, can refer to any accepted measure of determining the molar mass distribution or molar mass average of a mixture of polymers having varying degrees of polymerization, functionalization, and molar mass, including but not limited to “number-average molecular weight,” “mass-average molecular weight,” “weight-average molecular weight,” “Z (centrifugation) average molar mass,” or “viscosity average molar mass.”

[0126] The term, “weight-average molecular weight,” refers to a method of reporting the average molecular weight of polysaccharides in a mixture, calculated using the mole fraction distribution of the polysaccharides within the sample, using the equation $\bar{M}_w = \frac{\sum_i N_i M_i^2}{\sum_i N_i M_i}$, wherein N_i is the number of polysaccharides of molecular mass M_i .

[0127] The term, “number-average molecular weight,” refers to a method of reporting the average molecular weight of polysaccharides in a mixture, calculated by dividing the total weight of all of the polysaccharides in the sample divided by the number of polysaccharides in a sample, using the equation, $\bar{M}_n = \frac{\sum_i N_i M_i}{\sum_i N_i}$, wherein N_i is the number of polysaccharides of molecular mass M_i .

Accordingly, the weight-average molecular weight, $\bar{M}_w$, is necessarily skewed toward higher values corresponding to polysaccharides within the sample that are larger than other polysaccharides within the same mixture, and will always be larger than the number-average molecular weight, $\bar{M}_n$, except when the sample is monodisperse, and $\bar{M}_w$ equals $\bar{M}_n$. If a particular sample of polysaccharides within the sample has a large dispersion of actual weights, then $\bar{M}_w$ will be much larger than $\bar{M}_n$. Conversely, as the weight dispersion of polysaccharides in a sample narrows, $\bar{M}_w$ approaches $\bar{M}_n$.

[0128] The terms, “relative molecular weight” or “relative molar mass” (M_r), refers to another method of reporting the average molecular weight of polysaccharides in a mixture as a unitless quantity, most broadly determined by dividing the average mass of the molecule by an atomic mass constant, such as 1 atomic mass unit (amu) or 1 Dalton (Da). With respect to polysaccharides, M_r does not take into account the different chain-lengths, functionalization, and/or weight distribution of the polysaccharides in the sample, and instead simply represents the true average mass of the polysaccharides in the sample in a manner similar to small molecules.

[0129] The terms, “biological activity” or “catalytic activity,” refer to the ability of an enzyme to catalyze a particular chemical reaction by specific recognition of a particular substrate or substrates to generate a particular product or products. In some embodiments, the engineered enzymes of the present invention possess a biological or catalytic activity that is dependent on binding and reacting with aryl sulfate compounds, particularly PNS, as substrates. Additionally, some engineered enzymes are capable of having promiscuous catalytic activity with one or more alternate aryl sulfate compounds in addition to PNS, including but not limited to MUS, 7-hydroxycoumarin sulfate, phenyl sulfate, 4-acetylphenyl sulfate, indoxyl sulfate, 1-naphthyl sulfate, 2-naphthyl sulfate, and NCS.

[0130] The term, “coding sequence,” refers to that portion of a nucleic acid, for example, a gene, that encodes an amino acid sequence of a protein.

[0131] The term, “codon-optimized” refers to changes in the codons of the polynucleotide encoding a protein to those preferentially used in a particular organism such that the encoded protein is efficiently expressed in the organism of interest. Although the genetic code is degenerate in that most amino acids are represented by several codons, it is well known that codon usage by particular organisms is non-random and biased toward particular codon triplets. In some embodiments of the invention, the polynucleotide encoding for an engineered enzyme may be codon optimized for optimal production from the host organism selected for expression.

[0132] The terms, “corresponding to,” “reference to,” or “relative to,” when used in the context of the numbering of a given amino acid or polynucleotide sequence, refers to the numbering of the residues of a specified reference sequence when the given amino acid or polynucleotide sequence is compared to the reference sequence. In other words, the residue number or residue position of a given polymer is designated with respect to the reference sequence rather than by the actual numerical position of the residue within the given amino acid or polynucleotide sequence.

[0133] The term, “deletion,” refers to modification of a polypeptide by removal of one or more amino acids from the reference polypeptide. Deletions can comprise removal of 1 or more amino acids, the net result of which is retaining the catalytic activity of the reference polypeptide. Deletions can be directed to the internal portions and/or terminal portions of a polypeptide. Additionally, deletions can comprise continuous segments or they can be discontinuous.

[0134] The term, “disaccharide unit,” refers to the smallest repeating backbone unit within many polysaccharides, including linear polysaccharides, in which the smallest repeating unit consists of two sugar residues. With respect to a heparosan-based polysaccharide, the disaccharide unit consists of a hexuronic acid residue and a glucosamine residue, either of which can be functionalized and in which the hexuronic acid residue can either be glucuronic acid or iduronic acid. Each disaccharide unit within the heparosan-based polysaccharide can be described by its backbone structure and by the number and position of sulfo groups that are present. Further, the relative abundance of disaccharide units having the same structure within the same polysaccharide, and/or within the same sample of polysaccharides, can be characterized to determine the amount of sulfation at a particular position as a result of reacting with any of the sulfotransferases described herein.

[0135] The terms, “fragment” or “segment,” refer to a polypeptide that has an amino- or carboxy-terminal deletion, but where the remaining amino acid sequence is identical to the corresponding positions in a reference sequence. Fragments can be at least 50 amino acids or longer, and up to 70%, 80%, 90%, 95%, 98%, and 99% of a full-length aryl sulfate-dependent or natural sulfotransferase enzyme.

[0136] The terms, “functional site” or “functional domain,” generally refer to any site in a protein that confers a function on the protein. Representative examples include active sites (i.e., those sites in catalytic proteins where catalysis occurs) and ligand binding sites. Ligand binding sites include, but are not limited to, metal binding sites, co-factor binding sites, antigen binding sites, substrate channels and tunnels, and substrate binding domains. In an enzyme, a ligand binding site that is a substrate binding domain may also be an active site. Functional sites may also be composites of multiple functional sites, wherein the absence of one or more sites comprising the composite results in a loss of function. As a non-limiting example, the active site of a particular sulfotransferase enzyme may include multiple binding sites or clefts, including one site for the sulfo donor and one site for the sulfo acceptor.

[0137] The terms, “gene,” “gene sequence,” and “gene segment,” refer to a functional unit of nucleic acid unit encoding for a functional protein, polypeptide, or peptide. As would be understood by those skilled in the art, this functional term includes both genomic sequences and cDNA sequences. The terms, “gene,” “gene sequence,” and “gene segment,” additionally refer to any DNA sequence that is substantially identical to a polynucleotide sequence disclosed herein encoding for engineered enzyme gene product, protein, or polysaccharide, and can comprise any combination of associated control sequence. The terms also refer to RNA, or antisense sequences, complementary to such DNA sequences. As used herein, the term “DNA segment” includes isolated DNA molecules that have been isolated free of recombinant vectors, including but not limited to plasmids, cosmids, phages, and viruses.

[0138] The term, “glycosaminoglycan,” refers to long, linear polysaccharides consisting of repeating disaccharide units. Examples of glycosaminoglycans (GAGs) include chondroitin, dermatan, heparosan, hyaluronic acid, and keratan. GAGs are generally heterogeneous with respect to mass, length, disaccharide unit structure and functionalization, degree of sulfation.

[0139] The term, “heparosan,” refers to a particular GAG having repeating $[\beta(1,4)\text{GlcA}-\alpha(1,4)\text{GlcNAc}]_n$ disaccharide units, in which GlcA is glucuronic acid and GlcNAc is *N*-acetyl glucosamine.

[0140] The term, “heparosan-based polysaccharide,” refers to polysaccharides having the same backbone structure as heparosan, in which the disaccharide unit contains 1→4 glycosidically-linked hexuronic acid and glucosamine residues. The hexuronic acid residue can either be GlcA, as in heparosan, or iduronic acid (IdoA), and can optionally have a sulfo group at the 2-*O* position. The glucosamine residue can either be *N*-acetylated, as in heparosan, *N*-sulfated, or *N*-unsubstituted, and can optionally be sulfated at the *N*-, 3-*O*, or 6-*O* position. As used herein, the term “*N*-unsubstituted,”

with respect to a glucosamine residue, is equivalent to an “N-deacetylated” glucosamine residue, and refers to an amine functional group that is capable of receiving a sulfo group either chemically, or enzymatically using an NST enzyme. According to the present invention, heparosan-based polysaccharides can be utilized as starting materials, formed as intermediates, acting as sulfo group acceptors and/or synthesized as products according to any of the methods described herein.

[0141] The term, “insertion,” refers to modifications to the polypeptide by addition of one or more amino acids to the reference polypeptide. Insertions can be in the internal portions of the polypeptide, or to the C- or N-termini of the polypeptide. Insertions can include fusion proteins as is known in the art and described below. The insertions can comprise a continuous segment of amino acids or multiple insertions separated by one or more of the amino acids in the reference polypeptide.

[0142] The term, “isolated nucleic acid” as used herein with respect to nucleic acids derived from naturally-occurring sequences, means a ribonucleic or deoxyribonucleic acid which comprises a naturally-occurring nucleotide sequence and which can be manipulated by standard recombinant DNA techniques, but which is not covalently joined to the nucleotide sequences that are immediately contiguous on its 5’ and 3’ ends in the naturally-occurring genome of the organism from which it is derived. As used herein with respect to synthetic nucleic acids, the term “isolated nucleic acid” means a ribonucleic or deoxyribonucleic acid which comprises a nucleotide sequence which does not occur in nature and which can be manipulated by standard recombinant DNA techniques. An isolated nucleic acid can be manipulated by standard recombinant DNA techniques when it may be used in, for example, amplification by polymerase chain reaction (PCR), *in vitro* translation, ligation to other nucleic acids (e.g., cloning or expression vectors), restriction from other nucleic acids (e.g., cloning or expression vectors), transformation of cells, hybridization screening assays, or the like.

[0143] The terms, “naturally occurring” or “natural,” refer to forms of an enzyme found in nature. For example, a naturally occurring or natural polypeptide or polynucleotide sequence is a sequence present in an organism that can be isolated from a source in nature and which has not been intentionally modified by human manipulation. A natural polypeptide or polynucleotide sequence can also refer to recombinant proteins or nucleic acids that can be synthesized, amplified, and/or expressed *in vitro*, and which have the same sequence and biological activity as an enzyme produced *in vivo*. In contrast to naturally occurring or natural sulfotransferase enzymes, the engineered aryl sulfate-dependent sulfotransferase enzymes utilized in accordance with methods of the present invention have different amino acid and nucleic acid sequences, biological activity with aryl sulfate compounds instead of PAPS as sulfo group donors, and cannot be found in nature.

[0144] The term, “oligosaccharide,” refers to saccharide polymers containing a small number, typically three to nine, sugar residues within each molecule.

[0145] The term, “percent identity,” refers to a quantitative measurement of the similarity between two or more nucleic acid or amino acid sequences. As a non-limiting example, the percent identity can be assessed between two or more engineered enzymes of the present invention, two or more naturally occurring enzymes, or between one or more engineered enzymes and one or more naturally occurring enzymes. Percent identity can be assessed relative to two or more full-length sequences, two or more truncated sequences, or a combination of full-length sequences and truncated sequences.

[0146] The term, “polysaccharide,” refers to polymeric carbohydrate structures formed of repeating units, typically monosaccharide or disaccharide units, joined together by glycosidic bonds, and which can range in structure from a linear chain to a highly-branched three-dimensional structure. Although the term “polysaccharide,” as used in the art, can refer to saccharide polymers having more than ten sugar residues per molecule, “polysaccharide” is used within this application to describe saccharide polymers having more than one sugar residue, including saccharide polymers that have three to nine sugar residues that may be defined in the art as an “oligosaccharide.” According to the present invention, the term “polysaccharide,” is also used to generally describe GAGs and GAG-based compounds, including chondroitin, dermatan, heparosan, hyaluronic acid, and keratan compounds.

[0147] The terms, “protein,” “gene product,” “polypeptide,” and “peptide” can be used interchangeably to describe a biomolecule consisting of one or more chains of amino acid residues. In addition, proteins comprising multiple polypeptide subunits (e.g., dimers, trimers or tetramers), as well as other non-proteinaceous catalytic molecules will also be understood to be included within the meaning of “protein” as used herein. Similarly, “protein fragments,” i.e., stretches of amino acid residues that comprise fewer than all of the amino acid residues of a protein, are also within the scope of the invention and may be referred to herein as “proteins.” Additionally, “protein domains” are also included within the term “protein.” A “protein domain” represents a portion of a protein comprised of its own semi-independent folded region having its own characteristic spherical geometry with hydrophobic core and polar exterior.

[0148] The term, “recombinant,” when used with reference to, for example, a cell, nucleic acid, or polypeptide, refers to a material that has been modified in a manner that would not otherwise exist in nature. Non-limiting examples include, among others, recombinant cells expressing genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise expressed at a different level.

[0149] The term, “reference sequence,” refers to a disclosed or defined sequence used as a basis for sequence comparison. A reference sequence may be a subset of a larger sequence, for example, a segment of a full-length gene or polypeptide sequence. Generally, a reference sequence refers to at least a portion of a full-length sequence, typically at least 20 amino acids, or the full-length sequence of the nucleic acid or polypeptide.

[0150] The term, “saccharide,” refers to a carbohydrate, also known as a sugar, which is a broad term for a chemical compound comprised of carbon, hydrogen, and oxygen, wherein the number of hydrogen atoms is essentially twice that of the number of oxygen atoms. Often, the number of repeating units may vary in a saccharide. Thus, disaccharides, oligosaccharides, and polysaccharides are all examples of chains composed of saccharide units that are recognized by the engineered sulfotransferase enzymes of the present invention as sulfo group acceptors.

[0151] The term, “substantially equivalent,” with respect to polysaccharides utilized as starting materials, formed as intermediates, acting as sulfo group acceptors, and/or synthesized as products according to any of the methods described herein, refers to one or more properties of a polysaccharide sample that are identical to those found in a polysaccharide sample characterized in the prior art. Such properties may include, but are not limited to, chemical structure, sulfation frequency and location, disaccharide unit composition, molecular weight profile, and/or anticoagulant activity. Even if the two polysaccharide samples have additional properties that may be different, such differences do not significantly affect their substantial equivalence. In a non-limiting example, anticoagulant NS2S6S3S-HS products synthesized according to methods of the present invention can be substantially equivalent to the United States Pharmacopeia (USP) reference standard (CAS No: 9041-08-1) with respect to chemical structure, molecular weight profile, and/or anticoagulant activity, but can be produced at a different purity than the USP reference standard, which is isolated from natural sources and can contain non-trace amounts of other GAGs in the same sample.

[0152] The term, “substantially pure,” with respect to protein preparations, refers to a preparation which contains at least 60% (by dry weight) the protein of interest, exclusive of the weight of other intentionally included compounds. Particularly the preparation is at least 75%, more particularly at least 90%, and most particularly at least 99%, by dry weight the protein of interest, exclusive of the weight of other intentionally included compounds. Purity can be measured by any appropriate method, e.g., column chromatography, gel electrophoresis, or high-performance liquid chromatography (HPLC) analysis. If a preparation intentionally includes two or more different proteins of the invention, a “substantially pure” preparation means a preparation in which the total

dry weight of the proteins of the invention is at least 60% of the total dry weight, exclusive of the weight of other intentionally included compounds. Particularly, for such preparations containing two or more proteins of the invention, the total weight of the proteins of the invention can be at least 75%, more particularly at least 90%, and most particularly at least 99%, of the total dry weight of the preparation, exclusive of the weight of other intentionally included compounds.

[0153] The terms, “sulfo” or “sulfuryl” refer to a functional group, substituent, or moiety having the chemical formula SO_3H^- that can be removed from an aryl sulfate compound and/or be transferred from a donor compound to an acceptor compound. In some embodiments, the engineered sulfotransferases of the present invention catalyze the transfer of sulfo groups from aryl sulfate compounds to a polysaccharide, particularly heparosan and/or heparosan-based polysaccharides.

[0154] The term, “sulfotransferase,” refers to any enzyme in an *in vivo* or *in vitro* process that is used to catalyze the transfer of a sulfo group from a sulfo donor compound to a sulfo acceptor compound. “Sulfotransferase” can be used interchangeably to describe enzymes that catalyze sulfotransfer reactions *in vivo* or to describe engineered enzymes of the present invention that catalyze sulfotransfer reactions *in vitro*.

[0155] The term, “transformation,” refers to any method of introducing exogenous a nucleic acid into a cell including, but not limited to, transformation, transfection, electroporation, microinjection, direct injection of naked nucleic acid, particle-mediated delivery, viral-mediated transduction or any other means of delivering a nucleic acid into a host cell which results in transient or stable expression of said nucleic acid or integration of said nucleic acid into the genome of said host cell or descendant thereof.

[0156] The term, “unfractionated heparin,” refers to any synthesized or isolated heparin that has not been modified and/or partially depolymerized to form low molecular weight heparin. With respect to naturally-obtained heparin, the term “unfractionated heparin” generally represents the form of the heparin isolated from the animal, typically from porcine or bovine sources, prior to purification to meet USP reference standards. With respect to products synthesized by methods of the present invention, the term “unfractionated heparin” can refer to the N,2,3,6-HS product having polysaccharides comprising the pentasaccharide sequence of Formula I, prior to purification to form API heparin or low-molecular-weight heparin.

DETAILED DESCRIPTION OF THE INVENTION

[0157] Heparin has been commonly described in a variety of medical treatments, most notably as a blood thinner as a result of its anticoagulant activity. However, in many treatments, that same

anticoagulant activity is not desired, and in many cases, can potentially cause dangerous side effects, including heparin-induced thrombocytopenia and increased risk of uncontrolled bleeding. To mitigate that risk, polysaccharide compositions that can still interact with various targets within the body while having reduced anticoagulant activity are prepared from heparin. Examples of such targets are described in further detail below. Heparin derivatives (also known as “heparinoids”) are typically prepared by *O*-desulfation of heparin, to form *O*-desulfated heparin (ODSH). The generated ODSH heparinoids are substantially desulfated at the 2-*O* position of hexuronic acid residues and/or the 3-*O* position of glucosamine residues within each polysaccharide, while retaining the *N*- and 6-*O* glucosamine sulfation commonly found in heparin. Methods of preparing and controlling the desulfation of heparin to form ODSH are described in U.S. Patents 5,990,097, 5,912,237, 5,808,021, 5,668,118, and 5,296,471.

[0158] Although ODSH heparinoids prescribed in medical treatments are often substantially 2-*O* and 3-*O* desulfated, from at least 85% and up to at least 99% 2-*O* and 3-*O* desulfation, the ODSH products nonetheless retain some of the anticoagulant activity from heparin, indicating that not all of the 2-*O* and 3-*O* positions are desulfated. For example, U.S. Patents 5,296,471 and 5,808,021 both describe the production of 2-*O* and 3-*O* desulfated ODSH compositions having between 1.2 and 10% of the anticoagulant activity of heparin using an activated partial thromboplastin time (aPTT) assay. Both patents also describe references which are referred to as “non-anticoagulant” depolymerized heparins, although these similarly only describe low molecular weight heparins (LMWH) having reduced anticoagulant potency under United States Pharmacopeia (USP) assay reaction conditions (*see, e.g.* Jaseja, M., et al., *Can J. Chem* (1989) 67: 1449-1456 (< 5 IU/mg) and U.S. Patent 6,150,342 (APTT: 54-102 IU/mg; Anti-Factor Xa: 3-8 IU/mg), the disclosures of which are incorporated by reference in their entireties). Similarly, U.S. Patents 5,668,118, 5,912,237, and 5,990,097 describe the production of 2-*O* desulfated heparin with “much reduced anti-coagulant activity” when compared to heparin, while U.S. Patents 6,489,311, 7,468,358, 9,271,999, and 10,052,346 describe the use of substantially 2-*O* and 3-*O* desulfated ODSH compositions having from 6-10 IU/mg of USP activity, 1.9-10 IU/mg of Anti-Factor Xa activity, and 2 IU/mg of Anti-Factor IIa activity. As a result, there is still potential for severe health risks associated with anticoagulant heparin when using ODSH heparinoids derived from anticoagulant heparin.

[0159] In contrast, heparan sulfate compounds that have no aPTT, USP, anti-Factor Xa (anti-Xa), and/or anti-Factor IIa (anti-IIa) anticoagulant activity can be synthesized by constructing such compounds *in vitro*, rather than depolymerizing heparin isolated from natural sources. Heparin synthesized *in vitro* has generally been performed by utilizing recombinant heparan sulfate

sulfotransferase enzymes and PAPS to selectively and sequentially add sulfate groups to *N*-deacetylated heparosan (see e.g. U.S. Pat. No. 8,771,995 and 9,951,149, the disclosures of which are incorporated by reference in its entirety), similar to processes performed *in vivo*. In nature, heparosan is synthesized in the Golgi apparatus as co-polymers of glucuronic acid and *N*-acetylated glucosamine, before being modified by one or more sulfotransferases to form heparan sulfate (HS) products. Such modifications include *N*-deacetylation and *N*-sulfation of glucosamine, C₅ epimerization of glucuronic acid to form iduronic acid residues, 2-*O*-sulfation of iduronic and/or glucuronic acid, as well as 6-*O*-sulfation and 3-*O*-sulfation of glucosamine residues. The natural sulfotransferases that catalyze *N*-sulfation, 2-*O*-sulfation, 6-*O*-sulfation and 3-*O*-sulfation of heparosan and HS polysaccharides *in vivo* exclusively recognize and bind with PAPS, a nearly ubiquitous sulfo group donor recognized by nearly all sulfotransferases, particularly in eukaryotes. An example of a sulfotransfer reaction mechanism between the human glucosaminyl 3-*O* sulfotransferase (3OST) enzyme, PAPS, and heparan sulfate is illustrated in Figures 1A-1C. In particular, the glutamic acid residue at position 43 abstracts the proton from the 3-*O* position of the *N*-sulfoglucosamine residue within the polysaccharide, enabling the nucleophilic attack and removal of the sulfo group from PAPS, whereas His-45 and Asp-48 coordinate to stabilize the transition state of the enzyme before the sulfated polysaccharide product is released from the active site.

[0160] However, although PAPS is the exclusive sulfo donor in eukaryotes, it has a short half-life and can readily decompose into adenosine 3',5'-diphosphate, which acts as a competitive inhibitor during sulfotransfer reactions. Animals can efficiently utilize PAPS because they can metabolize adenosine 3',5'-diphosphate to prevent competitive inhibition and also replenish PAPS for each sulfotransfer reaction, as needed. On the other hand, aryl sulfate compounds, which can be utilized as sulfo donors in a limited number of bacterial systems (see Malojcic, G., et al., above), cannot react with any of the known native sulfotransferase enzymes in eukaryotes, including those that are involved in synthesizing HS polysaccharides *in vivo*. Without being limited by a particular theory, it is believed that the binding pockets for PAPS within the active sites of eukaryotic sulfotransferases either do not have a high enough affinity for aryl sulfate compounds to facilitate binding, and/or that the aryl sulfate compounds are sterically hindered from entering the active site at all.

[0161] The present disclosure includes methods for synthesizing HS polysaccharides using sulfotransferase enzymes that are engineered to recognize and bind with aryl sulfate compounds as sulfo group donors. Particularly, the engineered sulfotransferase enzymes are designed to transfer sulfo groups from aryl sulfate compounds to *N*-deacetylated heparosan or heparan sulfate. The structure and activity (anticoagulant vs. non-anticoagulant) of the resulting HS polysaccharide

products can be controlled, in part, by the enzymes selected for the product synthesis. Depending on their role, HS polysaccharides can contain one or more unique patterns or motifs recognized by specific protein(s) involved in the particular biological process. In an embodiment of the invention, the HS polysaccharide produced by any of the methods described herein can have anticoagulant activity. In another embodiment, HS polysaccharides produced by any of the methods described herein can have zero anticoagulant activity.

[0162] It should be understood that while reference is made to exemplary embodiments and specific language is used to describe them, no limitation of the scope of the invention is intended. Further modifications of the methods described herein, as well as additional applications of the principles of those inventions as described, which would occur to one skilled in the relevant art and having possession of this disclosure, are to be considered within the scope of this invention. Furthermore, unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which embodiments of this particular invention pertain. The terminology used is for the purpose of describing those embodiments only, and is not intended to be limiting unless specified as such. Headings are provided for convenience only and are not to be construed to limit the invention in any way. Additionally, throughout the specification and claims, a given chemical formula or name shall encompass all optical isomers and stereoisomers, as well as racemic mixtures where such isomers and mixtures exist.

In vitro synthesis of heparan sulfate polysaccharides

[0163] In an embodiment of the invention, the synthesis of HS polysaccharides can be accomplished by treating a heparosan-based polysaccharide with an aryl sulfate compound and a sulfotransferase enzyme that has been engineered to recognize, bind, and react with aryl sulfate compounds as sulfo group donors. Each of the engineered sulfotransferase enzymes, including their sequences, structures, and biological activities, are described in further detail below. Without being limited by a particular theory, it is believed that sulfotransferase enzymes that recognize polysaccharides as sulfo group acceptors, but also bind and react with aryl sulfate compounds as sulfo donors, have neither been observed in nature nor described previously.

[0164] Those skilled in the art will appreciate that the engineered sulfotransferase enzymes utilized in the methods of the present invention have several advantages over *in vitro* and *in vivo* reaction mechanisms that are unable to bind and react with aryl sulfate compounds in order to catalyze sulfo transfer. Presently, obtaining large-scale quantities of sulfated polysaccharides, including ODSH, requires isolating heparin produced *in vivo* from animal sources, such as pigs and cattle (*see* Xu, Y.,

et al., (2011) *Science* **334** (6055): 498-501). However, a worldwide contamination crisis of heparin shone a spotlight on the fragility of solely relying on obtaining them from animal sources. Consequently, in recent years, there has been a push to develop synthetic routes to synthesizing heparin and other HS polysaccharides in large enough quantities to compliment or replace animal-sourced products.

[0165] In order to synthesize sulfated polysaccharides *in vitro*, there have historically been two reaction schemes: total chemical synthesis and chemoenzymatic synthesis. While both types of reaction schemes have led to purified products that in some instances are homogeneous, synthetic routes as a whole have been inadequate to produce sulfated polysaccharides, particularly heparin, on an industrial scale. Indeed, the production of such polysaccharides using total chemical synthesis has historically required as many as 60 steps and resulted in very low yields (*see* Balagurunathan, K., et al., (eds.) (2015) *Glycosaminoglycans: Chemistry and Biology*, Methods in Molecular Biology, vol. 1229, DOI 10.1007/978-1-4939-1714-3_2, © Springer Science + Business Media New York).

[0166] Chemoenzymatic synthesis routes, on the other hand, generally utilize far fewer steps and increase the scale of the generated anticoagulant products into multi-milligram amounts (*See* U.S. Pat. No. 8,771,995 and 9,951,149, the disclosures of which are incorporated by reference in its entirety). The improvements in the quantity of obtainable product can be attributed to the ability to combine recombinant natural sulfotransferases and PAPS in a reaction vessel in order to catalyze sulfo group transfer. Yet, chemoenzymatic methods to this point are inadequate for forming heparin on a large scale, because of the natural sulfotransferases' requirement to react with PAPS. PAPS is a highly expensive and unstable molecule that has been an obstacle to the large-scale production of enzymatically sulfated products, including heparin, because the half-life of PAPS at pH 8.0 is only about 20 hours.

[0167] Furthermore, product inhibition by adenosine 3',5'-diphosphate has also been a limiting factor to large-scale synthesis of sulfated products. The highly negative impact of the product inhibition by adenosine 3',5'-diphosphate can be somewhat reduced by employing a PAPS regeneration system (*see* U.S. Pat. No. 6,255,088, above, and Burkhardt, et al. (2000) *J. Org. Chem.* **65**: 5565-5574) that converts adenosine 3',5'-diphosphate into PAPS. Despite the PAPS regeneration system, however, the absolute necessity to supply PAPS to initiate the chemical reaction with native sulfotransferases nonetheless creates an insurmountably high-cost barrier to synthesize sulfated products, including heparin, on an industrial, production-grade scale.

[0168] In contrast to prior chemoenzymatic syntheses of sulfated polysaccharides that require PAPS as a sulfo donor in order to drive activity, the methods of the present invention obviate the need to

use PAPS altogether, because each of the sulfotransferases have been engineered to recognize, bind, and react with aryl sulfate compounds as sulfo group donors. As described above, some aryl sulfate compounds, such as PNS, NCS, or MUS, are cheap, widely-available, and have been shown to react with some bacterial sulfotransferases as sulfo donors (see *Malojcic, G.*, et al., above). However, bacterial sulfotransferases are unsuitable to synthesize sulfated polysaccharides, particularly heparin or ODSH, because the bacterial sulfotransferases only recognize other aromatic compounds as sulfo group acceptors, and cannot bind or react with polysaccharides. Consequently, and without being limited by a particular theory, it is believed that the engineered sulfotransferases utilized in methods of the present invention are the only known sulfotransferases that are capable of catalyzing sulfo group transfer from an aryl sulfate compound to a polysaccharide, particularly heparosan-based polysaccharides. Generally, any of the methods described herein for synthesizing sulfated products such as heparin and ODSH can be performed using one or more engineered sulfotransferases, and such engineered sulfotransferases can comprise any amino acid sequence so long as its biological activity is dependent on transferring a sulfo group from an aryl sulfate compound to heparosan-based polysaccharide. Non-limiting examples of engineered enzymes, aryl sulfate compounds, and heparosan-based polysaccharides are described in further detail, below.

[0169] In nature, heparan sulfate can be sulfated at the 2-*O* position of any hexuronic acid residue and the *N*-, 3-*O*, 6-*O* position of any glucosamine residue within the polysaccharide. Further, several of the hexuronic acid or glucosamine residues within the same polysaccharide chain can be sulfated at any of the above positions, and can form a characteristic sulfation pattern that can be recognized by one or more enzymes or co-factors within the body. As a non-limiting example, heparin contains polysaccharides having a characteristic pentasaccharide sequence with a specific sulfation pattern that is recognized by antithrombin.

[0170] In an embodiment of the invention, methods are provided for chemoenzymatically synthesizing heparan sulfate products, including ODSH and heparin. One or more, and preferably all, of the *N*-, 2-*O*-, 3-*O*-, and 6-*O* sulfation steps can be catalyzed using sulfotransferase enzymes that are engineered to react with aryl sulfate compounds in the absence of PAPS. Each of these enzymes are described in further detail below. Such product compositions that can be synthesized comprise, as non-limiting examples, *N*-, 2-*O* sulfated heparan sulfate (NS2S/HS), *N*-, 6-*O* sulfated heparan sulfate (NS6S/HS), *N*-, 2-*O*, 6-*O* sulfated heparan sulfate (NS2S6S/HS), and *N*-, 3-*O*, 6-*O* sulfated heparan sulfate (NS3S6S/HS). In various embodiments, the NS2S/HS, NS6S/HS, NS2S6S/HS, and NS3S6S/HS product composition(s) synthesized by any of the methods described herein can have substantially the same structure(s) and pharmaceutical activities as any ODSH

composition in the art produced by *O*-desulfating heparin. In various embodiments, NS2S/HS, NS6S/HS, NS2S6S/HS, and NS3S6S/HS products, particularly NS6S/HS products, can have zero USP, aPTT, Anti-Xa, and/or Anti-IIa anticoagulant activity, distinguishing them from ODSH heparinoid compositions, which retain some anticoagulant activity from the heparin or low molecular weight heparin (LMWH) compositions from which they are derived.

[0171] In an embodiment of the invention, methods are provided for chemoenzymatically synthesizing *N*-, 2-*O*-, 3-*O*-, 6-*O*-sulfated-HS (N,2,3,6-HS) products, particularly heparin. One or more, and preferably all, of the *N*-, 2-*O*-, 3-*O*-, and 6-*O* sulfation steps can be catalyzed using sulfotransferase enzymes that are engineered to react with aryl sulfate compounds in the absence of PAPS. Each of these enzymes are described in further detail below. By controlling the molecular weight and *N*-acetyl glucosamine content of heparosan-based polysaccharides utilized as starting materials, an N,2,3,6-HS product composition can be formed that has a comparable molecular weight, sulfation, and anticoagulant activity to the United States Pharmacopeia (USP) reference standard (CAS No: 9041-08-1) for API heparin. In various embodiments, once the NS2S6S3S-HS product is formed according to any of the methods of the present invention, it can subsequently be *O*-desulfated according to any method known in the art to form an ODSH composition *in vitro*. Such ODSH compositions can be completely free of dermatan sulfate and chondroitin sulfate contaminants that can be found in ODSHs produced from animal-sourced heparin.

[0172] Heparin produced *in vitro* and *in vivo* contains heparan sulfate polysaccharides having a consensus pentasaccharide motif, which can only be formed when sulfated in a specific order. Thus, in methods of the present invention in which a heparin product is synthesized, the order of sulfation within the pentasaccharide motif is typically: (1) *N*-sulfation; (2) 2-*O*-sulfation; (3) 6-*O*-sulfation; and (4) 3-*O* sulfation. However, other portions of the polysaccharide can be sulfated in any order, and other HS products, including NS2S/HS, NS6S/HS, NS2S6S/HS, NS3S6S/HS, and non-heparin NS2S6S3S-HS products, can be synthesized by sulfating heparosan-based polysaccharides in any order. Each of the reaction steps utilized to synthesize any HS product can optionally be performed in a single pot, or performed in one or more separate steps in which the products are isolated and purified prior to performing the next sulfation step.

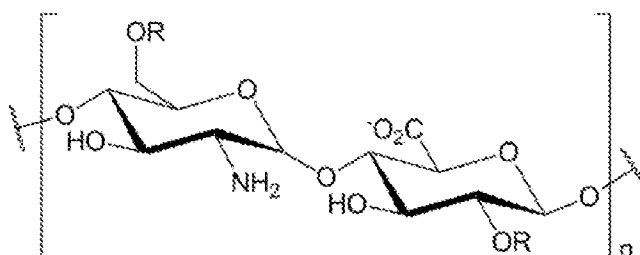
[0173] In general, and as described above, a vast majority of natural sulfotransferases, including all sulfotransferases known to react with polysaccharides, react with PAPS as a sulfo donor. Consequently, each sulfotransferase enzyme is generally classified by the chemical reaction it catalyzes, particularly the sulfo group acceptor and the subsequently-formed product. With respect to sulfotransferases that react with heparosan-based polysaccharides, the enzymes must further

recognize specific structural motifs and sulfation patterns within the polysaccharide chain in order to bind and react. Each of the engineered, aryl sulfate-dependent sulfotransferases, and the sulfo acceptor polysaccharides that they recognize, bind, and react with, are described in further detail below.

Glucosaminyl N-sulfotransferases

[0174] In nature, *N*-sulfation is typically carried out by *N*-deacetylase/*N*-sulfotransferase (NDST) enzymes have dual activity, in which the same enzyme can catalyze the *N*-deacetylation of *N*-acetyl glucosamine residues and the *N*-sulfation of unsubstituted glucosamine residues within heparosan. In particular, *N*-sulfation is accomplished by the enzymatic transfer of a sulfo group from PAPS to the glucosamine residue. The dual *N*-deacetylase and *N*-sulfotransferase activity of NDST is achieved via two separate structural domains—an *N*-deacetylase domain and an *N*-sulfotransferase domain. However, the activity of one of the domains is not a pre-requisite for the activity of the other domain, and recombinant single domain proteins comprising either *N*-deacetylase or *N*-sulfotransferase activity can be expressed and purified. Thus, in *in vitro* syntheses of heparan sulfate products, a single-domain, recombinant *N*-sulfotransferase enzyme is often utilized to carry out the *N*-sulfation step. Similarly, and in an embodiment of the invention, engineered aryl sulfate-dependent NST enzymes can be expressed and purified to comprise a single, *N*-sulfotransferase domain, in order to catalyze the *N*-sulfation of *N*-deacetylated heparosan in the absence of PAPS.

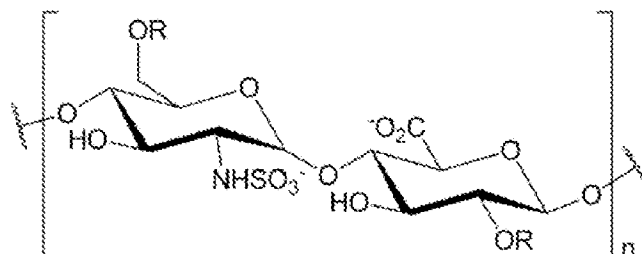
[0175] Naturally-occurring NDST enzymes, which react with PAPS as a sulfo group donor, are members of the EC 2.8.2.8 enzyme class. *N*-deacetylated portions of heparosan that can react with natural NDST enzymes, recombinant *N*-sulfotransferase domains of natural NDST enzymes, and the engineered aryl-sulfate dependent NST enzymes described herein can comprise one or more disaccharide units comprising the structure of Formula II, below:



wherein *n* is an integer and *R* is selected from the group consisting of a hydrogen atom or a sulfo group. Although the portion of the polysaccharide that reacts with the enzyme comprises the structure of Formula II, other portions of the polysaccharide can be *N*- or *O*- substituted. Typically, *N*-deacetylated heparosan comprising the structure of Formula II can comprise at least four

disaccharide units, or eight sugar residues total. Sulfotransfer reactions in which *N*-deacetylated heparosan is utilized as the sulfo group acceptor are discussed in Sheng, J., et al., (2011) *J. Biol. Chem.* **286** (22):19768-76, as well as Gesteira, T.F., et al., (2013) *PLoS One* **8** (8):e70880, the disclosures of which are incorporated by reference in their entireties.

[0176] Upon successfully binding PAPS and *N*-deacetylated heparosan, NDST enzymes can catalyze transfer of the sulfo group to an unsubstituted glucosamine, forming an *N*-sulfated heparosan product comprising the structure of Formula III, below:



wherein *n* is an integer and *R* is selected from the group consisting of a hydrogen atom or a sulfo group. Similarly, when an engineered aryl sulfate-dependent NST enzyme successfully binds with an aryl sulfate compound and *N*-deacetylated heparosan, *N*-sulfation is catalyzed to form an *N*-sulfated heparosan product comprising the structure of Formula III.

[0177] In another embodiment, each of the repeating disaccharide units within the *N*-deacetylated heparosan that reacts with any of the natural NDST enzymes or any of the engineered aryl sulfate-dependent NST enzymes comprises the structure of Formula II. In further embodiments, both of the *R* groups at the 6-*O* position of the glucosaminyl residues and the 2-*O* position of the glucuronic acid residues are hydrogen atoms, in all of the disaccharide units. In other embodiments, in some locations within the polysaccharide, at least a portion of the glucosamine residues are still *N*-acetylated, as shown in Figure 2, although glucosaminyl residues that are *N*-acetylated cannot directly participate as sulfo group acceptors. However, the presence of *N*-acetylated residues within the polysaccharide does not affect the sulfotransferases' binding affinity for non-acetylated residues within the same polysaccharide. In another embodiment, regardless of the structure of the heparosan-based polysaccharide adjacent to portion comprising the structure of Formula II, the *N*-sulfated polysaccharide product generated by reacting with an engineered NST or natural NDST (or the recombinant *N*-sulfotransferase domain of NDST) comprises the structure of Formula III.

[0178] In another embodiment, when there are multiple dimers comprising the structure of Formula II within the polysaccharide, any unsubstituted glucosamine residue can be *N*-sulfated. Similarly, the same polysaccharide can be *N*-sulfated multiple times, including and up to all available unsubstituted glucosaminyl residues that are present within the chain.

[0179] In another embodiment, heparosan-based polysaccharides comprising the structure of Formula II can be provided as a homogenous composition. In still other embodiments, sulfo acceptor polysaccharides comprising the structure of Formula II can be comprised within a composition comprising a polydisperse mixture of polysaccharides having variable chain lengths, molecular weights, and monosaccharide composition and functionalization.

[0180] In another embodiment, heparosan-based polysaccharides comprising the structure of Formula II and utilized in accordance with methods of the present invention can be obtained and/or modified from commercial sources. In other embodiments, heparosan can be isolated from bacterial or eukaryotic sources and subsequently chemically treated in order to produce an *N*-deacetylated polysaccharide that comprises the structure of Formula II. Such processes are discussed in detail in the description and examples, below.

[0181] The *N*-sulfotransferase domains of natural NDST enzymes within EC 2.8.2.8 typically comprise approximately 300 to 350 amino acid residues that can vary greatly in their sequence, yet ultimately have the exact same function, namely, to catalyze the *N*-sulfation of unsubstituted glucosamine residues within *N*-deacetylated heparosan. Without being limited by a particular theory, it is believed that each of the natural *N*-sulfotransferase domains can catalyze the same chemical reaction because there are multiple amino acid sequence motifs and secondary structures that are either identical or highly conserved across all species.

[0182] Further, it is believed that several of the conserved amino acid sequence motifs within NDST are directly involved in binding of either PAPS and/or the polysaccharide, or participate in the chemical reaction itself. The identity of conserved amino acid sequence motifs between the NDST enzymes can be demonstrated by comparing the amino acid sequence of the *N*-sulfotransferase domain of the human NDST1 enzyme, which has a solved crystal structure (PDB code: 1NST) in which amino acid residues within the active site have been identified, with the amino acid sequences of the *N*-sulfotransferase domains of other natural NDSTs. A multiple sequence alignment of the *N*-sulfotransferase domains of fifteen enzymes within EC 2.8.2.8, including several eukaryotic organisms and several isoforms of the human NDST, is shown in Figures 3A-3C, along with their percent identity relative to the human NDST1 (UniProtKB Accession No. P52848). As illustrated in Figures 3A-3C, sequences range from having 98.4% sequence identity with the P52848 reference sequence (entry sp|Q02353|NDST1_RAT) for the rat *N*-sulfotransferase domain down to 55.6% sequence identity (entry sp|Q9V3L1|NDST_DROME) for the fruit fly *N*-sulfotransferase domain. Those skilled in the art would appreciate that the multiple sequence alignment was limited to fifteen sequences for clarity, and that there are hundreds of amino acid sequences encoding for the

N-sulfotransferase domains of other natural NDST enzymes that have been identified and that have highly conserved active site and/or binding regions as well.

[0183] Within Figures 3A-3C, amino acids that are depicted in white with a black background at a particular position, are 100% identical across all sequences. Amino acids that are highly conserved, meaning that the amino acids are either identical or chemically or structurally similar, at a particular position are enclosed with a black outline. Within highly conserved regions, consensus amino acids that are present in a majority of the sequences, are in bold. Amino acids at a particular position that are not identical or highly conserved are typically variable. A period within a sequence indicates a gap that has been inserted into the sequence in order to facilitate the sequence alignment with other sequence(s) that have additional residues between highly conserved or identical region. Finally, above each block of sequences are a series of arrows and coils that indicate secondary structure that is conserved across all sequences, based on the identity of the amino acids within the alignment and using the structure of the natural human *N*-sulfotransferase enzyme as a reference. The β symbol adjacent to an arrow refers to a β -sheet, whereas a coil adjacent to an α symbol or a η symbol refers to a helix secondary structure.

[0184] Within the fifteen aligned sequences in Figures 3A-3C, there are several conserved amino acid motifs that include one or more amino acids that comprise the active site, based on the crystal structure of the *N*-sulfotransferase domain of human NDST1. These conserved amino acid sequence motifs, based on the numbering of the amino acid residues within Figures 3A-3C include residues 40-46 (Q-K-T-G-T-T-A); residues 66-69 (T-F-E-E); residues 101-105 (F-E-K-S-A); residues 139-143 (S-W-Y-Q-H); and residues 255-262 (C-L-G-K/R-S-K-G-R). In further embodiments, some isoforms of the natural sulfotransferase enzymes within EC 2.8.2.8 that comprise the conserved amino acid sequence motif Q-K-T-G-T-T-A further comprise the expanded conserved amino acid sequence motif, Q-K-T-G-T-T-A-L-Y-L, from residues 40-49.

[0185] Without being limited by a particular theory, it is believed that these residues either facilitate or participate in the chemical reaction, or enable binding of PAPS or the polysaccharide within the active site. In particular and as illustrated in Figures 4A-4C, the histidine residue at position 143 (corresponding to position 716 in the amino acid sequence of the full-length natural sulfotransferase enzyme that also includes an *N*-deacetylase domain) is in position to abstract one of the two protons within the amine functional group of the unsubstituted glucosaminyl residue within the polysaccharide, enabling the nitrogen atom to initiate the nucleophilic attack of PAPS and remove the sulfonyl group. Additionally, lysine residues at position 41 and 260 are also universally conserved, and are thought to coordinate with the sulfonyl moiety, driving binding of PAPS within

the active site as well as stabilizing the transition state during the course of the reaction (*see* Gesteira, T.F., et al., above, as well as Sueyoshi, T., et al., (1998) *FEBS Letters* **433**:211-214, the disclosure of which is incorporated by reference in its entirety).

[0186] However, as described above, the natural NDST enzymes are unable to catalyze the transfer of the sulfate group from an aryl sulfate compound to the polysaccharide, because without being limited by a particular theory, it is believed that the binding pocket for PAPS either does not have a high enough affinity for aryl sulfate compounds to facilitate binding and/or that the aryl sulfate compounds are sterically hindered from entering the active site. Consequently, and in another embodiment, the *N*-sulfotransferase domain of a natural NDST enzyme can be mutated in several locations within its amino acid sequence to enable binding of the aryl sulfate compound within the active site and/or to optimally position the aryl sulfate compound so transfer of the sulfate group to the polysaccharide can occur.

[0187] Accordingly, and in another embodiment, engineered NST enzymes that can be utilized in accordance with methods of the present invention can comprise a single *N*-sulfotransferase domain that is mutated relative to the *N*-sulfotransferase domain of any NDST enzyme, including enzymes having the amino acid sequences illustrated in Figures 3A-3C. In other embodiments, engineered NST enzymes that can be utilized in accordance with methods of the present invention can further comprise an *N*-deacetylase domain that has an identical or mutated amino acid sequence of the *N*-deacetylase domain of any natural NDST enzyme.

[0188] In another embodiment, mutations engineered into the amino acid sequences of the engineered enzymes facilitate a biological activity in which aryl sulfate compounds can both bind and react with the engineered NST enzymes as sulfo group donors. In further embodiments, the engineered NST enzyme can bind and react with an aryl sulfate compound as a sulfo group donor, while retaining the corresponding natural sulfotransferases' biological activity with heparosan and/or *N*-deacetylated heparosan as a sulfo group acceptor. Without being limited by a particular theory, it is believed that because of the mutations inserted into the amino acid sequences of the engineered NST enzymes, their sulfotransferase activity may comprise the direct transfer of a sulfuryl group from an aryl sulfate compound to the sulfo acceptor polysaccharide, using a similar mechanism as described in Figures 4A-4C above, except that the PAPS is substituted with the aryl sulfate compound. Otherwise, it is believed that the mutations may cause the sulfotransferase activity to comprise a two-step process including the hydrolysis of an aryl sulfate compound and formation of a sulfohistidine intermediate, followed by the nucleophilic attack of the sulfohistidine intermediate by an *N*-unsubstituted glucosamine within *N*-deacetylated heparosan to form the *N*-sulfated product.

By either mechanism, engineered NST enzymes have been shown to achieve sulfo transfer from an aryl sulfate compound to a polysaccharide, as described in the examples, below.

[0189] In another embodiment, an engineered NST enzyme can comprise one or more mutated amino acid sequence motifs relative to the conserved amino acid sequence motifs described above that are found in the *N*-sulfotransferase domains of natural NDST enzymes within EC 2.8.2.8, as described above and indicated in the multiple sequence alignment in Figure 3. In another embodiment, each mutated amino acid sequence motif that is present in the amino acid sequence of the engineered NST enzyme comprises at least one amino acid mutation relative to the corresponding conserved amino acid sequence motif within the *N*-sulfotransferase domain of a natural NDST. In another embodiment, an engineered NST enzyme comprises one mutated amino acid sequence motif. In another embodiment, an engineered NST enzyme comprises two mutated amino acid sequence motifs. In another embodiment, an engineered NST enzyme comprises three mutated amino acid sequence motifs. In another embodiment, an engineered NST enzyme comprises four mutated amino acid sequence motifs. In another embodiment, an engineered NST enzyme comprises five mutated amino acid sequence motifs. In another embodiment, an engineered NST enzyme that includes at least one mutated amino acid sequence motif relative to an *N*-sulfotransferase domain of any of the natural NDST enzymes within EC 2.8.2.8 can have an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40.

[0190] In another embodiment, upon viewing the crystal structure of the *N*-sulfotransferase domain of the human NDST1 (PDB code: 1NST) within a 3D molecular visualization system (including, as a non-limiting example, the open-source software, PyMOL), the structure of related sequences, such as those of engineered NST enzymes that contain one or more mutated amino acid sequence motifs relative to the human *N*-sulfotransferase domain, can be modeled for comparison as illustrated in Figures 5-8. In one non-limiting example, Figure 5 shows a magnified view of the active site of the human *N*-sulfotransferase domain that is overlaid with an engineered NST enzyme, comprising the amino acid sequence of SEQ ID NO: 10, in which the structure of the engineered enzyme is modelled upon making mutations relative to the human *N*-sulfotransferase domain amino acid sequence. Adenosine 3',5'-diphosphate, which is the product of a sulfotransfer reaction in which PAPS is the sulfo donor, and which was co-crystallized with the human *N*-sulfotransferase domain, is also illustrated within the active site. PNS is also modeled into the engineered enzyme active site, using the consensus solutions of molecular dynamics (MD) simulations that designed to calculate the

optimized position and orientation of a ligand within an enzyme active site adjacent to the polysaccharide binding site (not shown), if such solutions are possible.

[0191] As illustrated in Figure 5, although there are several mutations within SEQ ID NO: 10, relative to sequence of the human *N*-sulfotransferase domain (UniProtKB Accession No. P52848) indicated in Figure 3, the respective protein backbones are in a nearly identical location to one another, enabling a one-to-one comparison of the active sites. Within the structure of the engineered enzyme comprising the sequence of SEQ ID NO: 10, the consensus solutions from MD simulations indicate that the sulfate moiety within PNS is favored to bind adjacent to a histidine residue, His-45, that has been mutated relative to the natural amino acid residue, threonine, which is also universally conserved within EC 2.8.2.8. On the other hand, within the human *N*-sulfotransferase domain, the adenosine 3',5'-diphosphate is located near to the conserved His-143, described above. Although the sulfo group that would be comprised within the PAPS substrate is not shown, those skilled in the art would appreciate that if PAPS were present, the sulfate group would be oriented in a position immediately adjacent to His-143 and partially overlapping with the sulfate group within PNS. Without being limited by a particular theory, it is believed that the nearly overlapping location of the sulfate groups accounts for the engineered enzyme's ability to facilitate sulfo group transfer by using His-143 as a base to remove the proton from the glucosaminyl residue within the polysaccharide.

[0192] However, even though the sulfate groups can bind in a nearly identical location within the active site, aryl sulfate compounds cannot be utilized with EC 2.8.2.8 enzymes to facilitate sulfo group transfer to a polysaccharide. As described above, the amino acid residues within the active site of the natural enzymes are evolved to have strong binding affinity for PAPS, and likely do not have enough affinity for aryl sulfate compounds to drive binding and subsequently, reactivity. Consequently, other mutations must be present within the engineered enzymes to drive binding of aryl sulfate compounds within the active site. Figure 6 illustrates other mutations that surround PNS within the engineered enzyme comprising the amino acid sequence of SEQ ID NO: 10, including Trp-106, His-69, and His-40. Trp-106 and His-69 are positioned to provide π - π stacking binding contacts with aromatic moiety within PNS. Additionally, the ϵ 2 nitrogen atoms within His-69 and His-40 coordinate with the sulfuryl group directly. Lysine residues retained from the natural enzyme sequence, Lys-41 (not shown, for clarity) and Lys-103 are in position to coordinate with the sulfate group during transfer in order to stabilize the transition state. Of note, the natural amino acid residue, Lys-260, which also coordinates with the sulfate group in PAPS, is mutated to a valine residue within the engineered enzyme sequence. Without being limited by a particular theory, it is believed that His-45, which is necessary for the reaction with PNS, would exhibit charge repulsion with a

lysine residue at position 260, and that the mutation to a valine residue retains some steric bulk within the binding site while eliminating the charge repulsion. Lys-103 is nonetheless positioned to coordinate with the sulfonyl group, particularly when the sulfonyl group is associated or bound to His-45, as shown in Figure 6.

[0193] In another non-limiting example, Figure 7 shows a magnified view of the active site of the human *N*-sulfotransferase domain (UniProtKB Accession No. P52848) that is overlaid with a different engineered NST enzyme, comprising the amino acid sequence of SEQ ID NO: 2. PNS is modeled into the engineered enzyme active site, as described above. As with the engineered NST having the amino acid sequence SEQ ID NO: 10, the protein backbone of the enzyme having the amino acid sequence SEQ ID NO: 2 also has a nearly identical structure to the *N*-sulfotransferase domain of the human enzyme. However, the consensus solutions from MD simulations indicate that the sulfate moiety within PNS is favored to bind adjacent to a different histidine mutation (His-49), which is mutated from a natural leucine residue that is conserved in the active site of the *N*-sulfotransferase domain of several of the natural NDST enzymes. Consequently, mutations within SEQ ID NO: 10 that formed binding contacts with PNS are not necessarily present in SEQ ID NO: 2. As illustrated in Figure 8 and similar to SEQ ID NO: 10, there are two mutations present within SEQ ID NO: 2 that appear to form π - π stacking binding contacts surrounding the aromatic moiety of PNS, Trp-45 and His-67. Other mutations that comprise side chains that coordinate with PNS include Ser-69 (coordinating with the nitro functional group of PNS) and His-260 (coordinating with the sulfate moiety). Similar to SEQ ID NO: 10, because the natural lysine residue at position 260 is mutated, the natural Lys-103 residue is utilized within SEQ ID NO: 2 to coordinate with the sulfate moiety within PNS.

[0194] Those skilled in the art would appreciate that engineered NST enzymes of any other amino acid sequence, including, but not limited to, those described by SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 12, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40, would likely exhibit a similar structure to the human *N*-sulfotransferase domain and engineered NST enzymes having the amino acid sequence of SEQ ID NO: 2 and SEQ ID NO: 10. Without being limited by a particular theory, it is also believed that NCS would bind in a similar position as PNS within the active site of any of the engineered NST enzymes, since the structures of the two aryl sulfate compounds are very similar, except that the sulfate group is located *ortho* on the aromatic ring relative to the nitro group, rather than *para* to the nitro group.

[0195] Further, engineered NST enzymes utilized in accordance with methods of the present invention can include mutated amino acid sequence motifs that include the above-described mutations as well as other mutations that facilitate binding of substrates, the sulfotransfer reaction, or the stability of the enzyme during protein expression. In another embodiment, an engineered NST enzyme can include the mutated amino acid sequence motif, X_1 -K-T-G-A-W/F-A/L-L- X_2 -H, mutated from the conserved amino acid sequence Q-K-T-G-T-T-A-L-Y-L within EC 2.8.2.8, wherein X_1 is selected from the group consisting of glutamine, serine, and alanine; and X_2 is selected from the group consisting of tyrosine, threonine, and histidine. Engineered NST enzymes that include the mutated amino acid sequence motif X_1 -K-T-G-A-W/F-A/L-L- X_2 -H include, but are not limited to SEQ ID NO: 2 (described above), as well as SEQ ID NO: 4, SEQ ID NO: 12; SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 36, and SEQ ID NO: 40. In further embodiments, engineered NST enzymes can further include the mutated amino acid sequence motif, T- X_3 - X_4 -S, mutated from the conserved amino acid sequence T-F-E-E, wherein X_3 is a mutation relative to the natural sulfotransferase enzymes within EC 2.8.2.8, selected from the group consisting of histidine and glycine; X_4 is a mutation relative to the natural sulfotransferase enzymes within EC 2.8.2.8, selected from the group consisting of glycine, histidine, and serine; and wherein at least one of X_3 and X_4 is a histidine residue. In some even further embodiments, X_1 is glutamine, X_2 is tyrosine, X_3 is histidine, X_4 is glycine, and the engineered NST enzyme further comprises the mutated amino acid sequence motif, C-L-G-K/R-S-H-G-R. In other even further embodiments, X_1 is serine, X_2 is threonine, X_3 is glycine, X_4 is histidine, and the engineered NST enzyme further comprises the mutated amino acid sequence motif, C-H-G-K/R-R-W-G-R. In still other even further embodiments, X_1 is alanine, X_2 is histidine, X_3 is histidine, X_4 is serine, and the engineered NST enzyme further comprises the mutated amino acid sequence motif, C-A-H-K/R-G-L-G-R.

[0196] In another embodiment, engineered NST enzymes can include the mutated amino acid sequence motif, H- X_5 -T-G- X_6 -H-A, mutated from the conserved amino acid sequence Q-K-T-G-T-T-A, wherein X_5 is selected from the group consisting of lysine and glycine; and X_6 is a mutation relative to the natural sulfotransferase enzymes within EC 2.8.2.8, selected from the group consisting of glycine and valine. Engineered NST enzymes that include the mutated amino acid sequence motif H- X_5 -T-G- X_6 -H-A include, but are not limited to SEQ ID NO: 10 (described above), as well as SEQ ID NO: 6, SEQ ID NO: 8; SEQ ID NO: 34, SEQ ID NO: 37, SEQ ID NO: 38, and SEQ ID NO: 39. In further embodiments, X_5 is glycine and X_6 is glycine. In some even further embodiments, the engineered NST enzyme further comprises the mutated amino acid sequence motif,

C-G-G-K/R-H-L-G-R. In other even further embodiments, the engineered NST enzyme further comprises the mutated amino acid sequence motif, F-E-H-S-G.

[0197] In another embodiment, within any of the engineered NST enzymes that include the mutated amino acid sequence motif, H-X₅-T-G-X₆-H-A, X₅ is selected from the group consisting of lysine and glycine; and X₆ is a mutation relative to the natural sulfotransferase enzymes within EC 2.8.2.8, selected from the group consisting of glycine and valine. In further embodiments, X₅ is selected to be lysine, X₆ is selected to be valine, and the engineered NST enzyme further comprises the mutated amino acid sequence motif, T-G-N-H.

[0198] Furthermore, the amino acid sequences (SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12) of six engineered NST enzymes, which have been experimentally determined to be active with aryl sulfate compounds as sulfo group donors (see Example 2 below) can be compared with the amino acid sequence of the *N*-sulfotransferase domain of the human NDST1 (entry sp|P52848|NDST1_HUMAN) in a multiple sequence alignment to determine if there are relationships between mutations among each of the enzymes. A period within the amino acid sequence of an engineered enzyme indicates identity at a particular position with the human *N*-sulfotransferase domain. As shown in Figure 9, the sequence alignment demonstrates that while over 90% of the amino acid residues within the six sulfotransferase sequences are identical, there are several positions in which multiple amino acids can be chosen. Without being limited by a particular theory, it is believed that these enzymes have a similar relationship with each other as the *N*-sulfotransferase domains of the natural NDST enzymes that comprise EC 2.8.2.8. As a result, and in another embodiment, engineered NST enzymes comprising an amino acid sequence in which multiple amino acids can be chosen at defined positions are disclosed as SEQ ID NO: 33 and SEQ ID NO: 34. Positions at which the identity of an amino acid can be chosen from a selection of possible residues are denoted in terms “Xaa,” “Xn,” or “position n,” where n refers to the residue position.

[0199] In another embodiment, within an engineered NST enzyme comprising the amino acid sequence of SEQ ID NO: 33 or SEQ ID NO: 34, the amino acid residue at position 41 is lysine, the amino acid residue at position 44 is alanine, the amino acid residue at position 45 is an aromatic amino acid residue, preferably tyrosine or phenylalanine, and the amino acid residue at position 49 is histidine. In another embodiment, when the engineered NST enzyme comprises the above residues from positions 41-49, the amino acid residue at position 67 is glycine or histidine, the amino acid residue at position 68 is selected from the group consisting of glycine, histidine, and serine, and the amino acid residue at position 69 is serine.

[0200] In another embodiment, within an engineered NST enzyme comprising the amino acid sequence of SEQ ID NO: 33 or SEQ ID NO: 34, the amino acid residue at position 40 is histidine and the amino acid residue at position 45 is histidine. In further embodiments, the amino acid residue at position 41 is glycine and the amino acid residue at position 44 is glycine. In other further embodiments, the amino acid residue at position 41 is lysine and the amino acid residue at position 44 is valine. In even further embodiments, the amino acid residue at position 67 is glycine and the amino acid residue at position 69 is histidine. In still further embodiments, the amino acid residue at position 106 is tryptophan. In even still further embodiments, the amino acid residue at position 260 is valine.

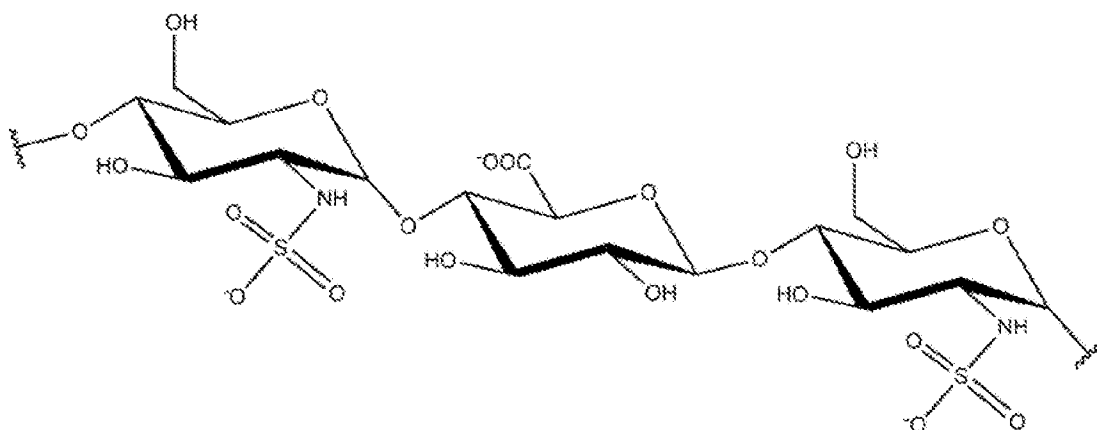
[0201] In another embodiment, within an engineered NST enzyme comprising the amino acid sequence of SEQ ID NO: 33 or SEQ ID NO: 34, the amino acid sequence can optionally include one or more mutations at residue positions not specified by an "Xn" or "Xaa," so long as any such mutations do not eliminate the NST and/or aryl sulfate-dependent activity of the enzyme. In another embodiment, such mutations not eliminating aryl sulfate-dependent activity at positions not specified by an "Xn" or "Xaa" can include substitutions, deletions, and/or additions.

[0202] Accordingly, in another embodiment, an engineered NST enzyme utilized in accordance with any of the methods of the present invention can comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40. In another embodiment, any of the above enzymes react with an aryl sulfate compound, instead of PAPS, as a sulfo group donor. In further embodiments, the aryl sulfate compound is selected from the group consisting of PNS, MUS, 7-hydroxycoumarin sulfate, phenyl sulfate, 4-acetylphenyl sulfate, indoxyl sulfate, 1-naphthyl sulfate, 2-naphthyl sulfate, and NCS. In some even further embodiments, the aryl sulfate compound is PNS. In other even further embodiments, the aryl sulfate compound is NCS.

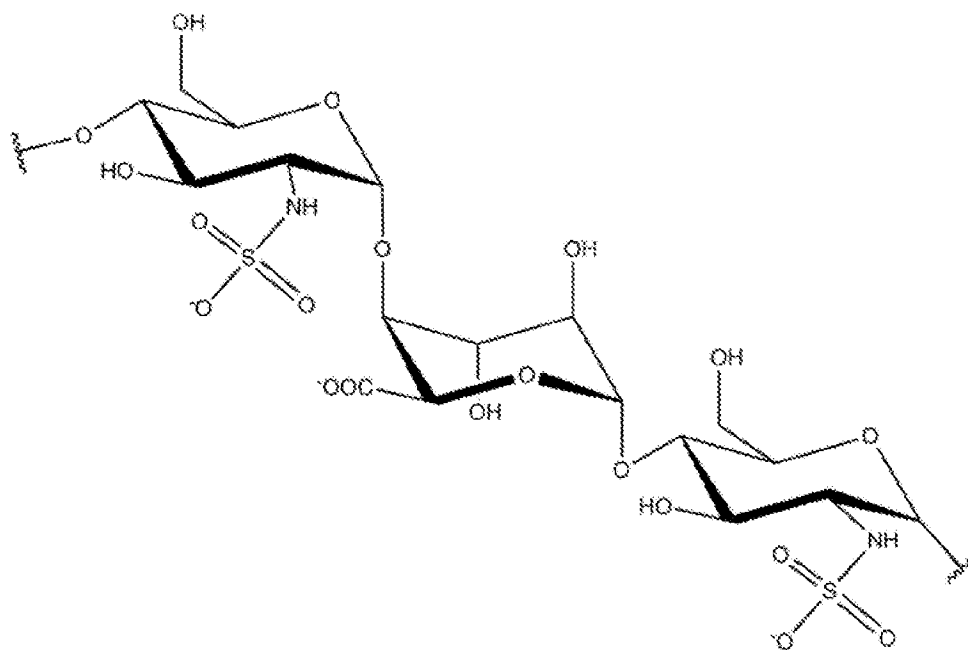
Hexuronyl 2-O sulfotransferases

[0203] In nature, HS hexuronyl 2-O sulfotransferase (2OST) enzymes recognize, bind, and react with *N*-sulfated heparosan-based polysaccharides as sulfo group acceptors. As with the natural NDSTs described above, natural 2OSTs transfer the sulfo group to the polysaccharide upon reacting with PAPS as a sulfo group donor. However, natural 2OSTs are members of the EC 2.8.2.- enzyme class. Generally, a majority of the glucosaminyl residues within the heparosan-based polysaccharide are *N*-sulfated, and the sulfo group is transferred to the 2-*O* position of a hexuronic acid residue,

generally either glucuronic acid or iduronic acid. A first non-limiting example of an *N*-sulfated heparosan that can bind and react with a natural or engineered 2OST is illustrated by the structure of Formula IV, below:



In another non-limiting example, an 2OST enzyme can recognize, bind, and react with heparosan-based polysaccharides having the structure of Formula V, below:

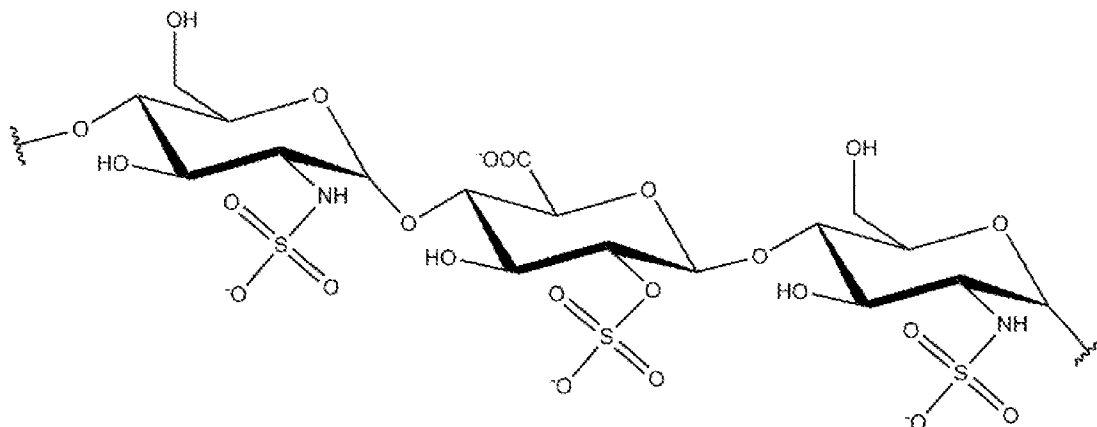


In both instances, the hexuronic acid residue (glucuronic acid in Formula IV, iduronic acid in Formula V) is flanked on either side by *N*-sulfated glucosamine residues that are otherwise unsubstituted at the 3-*O* and 6-*O* positions. Natural 2OST enzymes, and their biological activity with *N*-sulfated heparosan polysaccharides comprising the structures of Formula IV or Formula V, have been described by Rong, J., et al., (2001) *Biochemistry* **40** (18):5548-5555, the disclosure of which is incorporated by reference in its entirety.

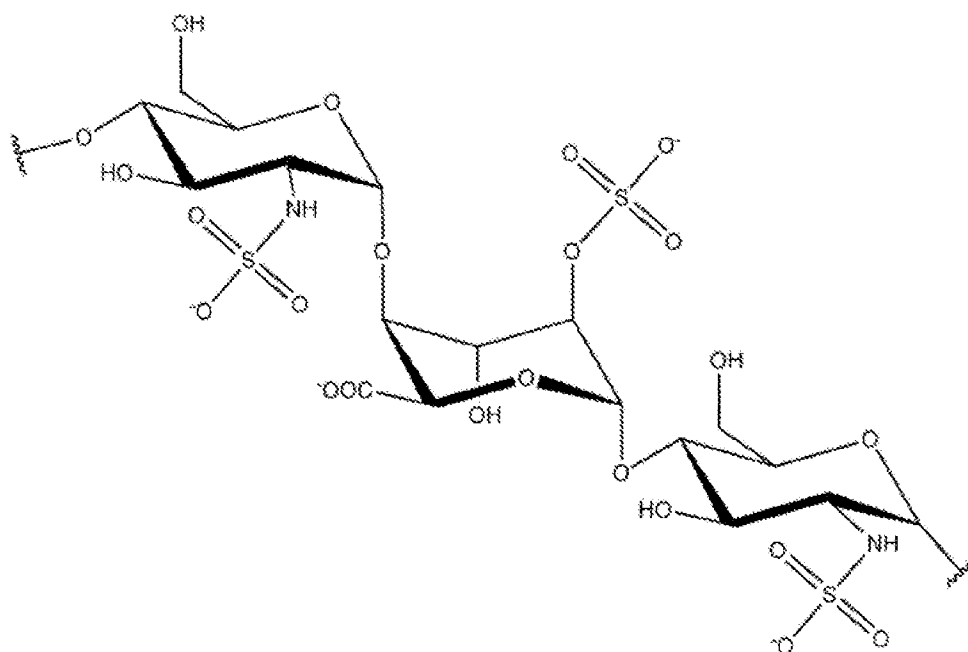
[0204] As described above, although the portion of the polysaccharide that reacts with the enzyme comprises the structure of Formula IV or Formula V, other portions can be *N*- or *O*- substituted. Similarly, the heparosan-based polysaccharides can comprise both the structure of Formula IV and the structure of Formula V within the same polysaccharide, and either or both of the hexuronyl residues within the structure of Formula IV and Formula V polysaccharide can be sulfated by the same enzyme molecule. Typically, *N*-sulfated HS polysaccharides comprising the structure of Formula IV and/or Formula V can comprise at least eight monosaccharide residues. In some embodiments, the heparosan-based polysaccharide is only *N*-sulfated or *N*-acetylated, and is not 3-*O* or 6-*O* sulfated prior to reacting with the 2OST. In another embodiment, engineered 2OSTs that can be utilized in accordance with methods of the present invention have the same biological activity as natural 2OSTs with heparosan-based polysaccharides, particularly those comprising the structure of Formula IV and Formula V, as sulfo acceptors.

[0205] The identity of the hexuronic acid residue in *N*-sulfated heparosan comprising the structure of Formula IV or Formula V can be controlled by the presence of a hexuronyl C₅-epimerase, which reversibly inverts the stereochemistry of the C₅-carbon. However, once the hexuronyl residue within a polysaccharide comprising the structure of Formula IV or Formula V is 2-*O* sulfated, epimerization can no longer occur. In eukaryotic systems, the *N*-sulfated heparosan products of NDST are almost exclusively formed as disaccharide units of *N*-sulfoglucosamine and glucuronic acid. Consequently, the glucuronic acid residue must be epimerized to an iduronic acid residue to form the structures of Formula V prior to reacting with the 2OST enzyme. However, and without being limited by a particular theory, it is believed that natural 2OST enzymes generally have preference for binding and reacting with heparosan-based polysaccharides comprising the structure of Formula V, and that most *N*-, 2-*O* sulfated HS (N,2-HS) polysaccharides produced *in vivo* generally comprise 2-*O* sulfated iduronic acid.

[0206] Upon successfully binding PAPS and *N*-sulfated heparosan comprising the structure of Formula IV, natural 2OST enzymes can catalyze transfer of the sulfo group to the 2-*O* position of the glucuronic acid residue, forming an N,2-HS product comprising the structure of Formula VI, below:



[0207] Similarly, engineered 2OST enzymes that successfully bind and react with an aryl sulfate compound and an *N*-sulfated heparosan comprising the structure of Formula IV can also form an N,2-HS product comprising the structure of Formula VI. Upon successfully binding PAPS and *N*-sulfated heparosan comprising the structure of Formula V, natural 2OST enzymes can catalyze transfer of the sulfo group to the 2-*O* position of the iduronic acid residue, forming an N,2-HS product comprising the structure of Formula VII, below:



Similarly, engineered 2OST enzymes that successfully bind and react with an aryl sulfate compound and an *N*-sulfated heparosan comprising the structure of Formula V can also form an N,2-HS product comprising the structure of Formula VII.

[0208] In another embodiment, in other locations within the *N*-sulfated sulfo acceptor polysaccharide, some of the glucosaminyl residues can be *N*-substituted with a sulfo group, an acetyl group, or a hydrogen, although hexuronyl residues within the polymer must reside between two *N*-

sulfoglucosamine residues, as described above, in order to receive a sulfo group. A non-limiting example of one such polysaccharide is illustrated in Figure 10. In Figure 10, hexuronyl residues 10 within polysaccharide 40 are flanked by glucosaminyl residues 20, 21, and 22, that are either *N*-sulfated, *N*-acetylated, or unsubstituted, respectively. Upon reacting the polysaccharide with either a natural or engineered 2OST, only the hexuronyl residue 10 flanked by two *N*-sulfoglucosaminyl residues 20 is sulfated, ultimately forming a sulfated hexuronyl residue 110 within the product polysaccharide 41.

[0209] In another non-limiting example, sulfo acceptor polysaccharides comprising the structures of Formula IV and Formula V are illustrated by polysaccharide 50 in Figure 11, Figure 12, and Figure 13. Additional monosaccharide residues required for catalysis are omitted for clarity. In Figure 11, Figure 12, and Figure 13, a hexuronyl residue 10 and an epimerized hexuronyl residue 30 reside between the three *N*-sulfoglucosaminyl residues 20 within polysaccharide 50. Although hexuronyl residues 10 and 30 are represented in a chair conformation, those skilled in the art can appreciate that such monosaccharide residues within a longer oligo- or polysaccharide chain can adopt several different conformations, including chair, half-chair, boat, skew, and skew boat conformations, and that those additional conformations are omitted for clarity.

[0210] Upon reacting polysaccharide 50 with any of the engineered aryl sulfate-dependent 2OST enzymes that can be utilized with methods of the present invention, the enzyme can catalyze sulfo group transfer to hexuronyl residue 10 to form a sulfated hexuronyl residue 110 within product polysaccharide 51 (Figure 11), to epimerized hexuronyl residue 30 to form a sulfated epimerized hexuronyl residue 130 within product polysaccharide 52 (Figure 12), or to both hexuronyl residue 10 and epimerized hexuronyl residue 30 to form a sulfated hexuronyl residue 110 and a sulfated epimerized hexuronyl residue 130, respectively, within product polysaccharide 53 (Figure 13).

[0211] In another embodiment, polysaccharides comprising the structure of Formula IV and/or Formula V can be provided as a homogenous composition. In still other embodiments, polysaccharides comprising the structure of Formula IV and/or Formula V can be comprised within a composition comprising a polydisperse mixture of polysaccharides having variable chain lengths, molecular weights, relative abundance of Formula IV and/or Formula V, and overall monosaccharide composition and functionalization.

[0212] In some embodiments, polysaccharides comprising the structure of Formula IV and/or Formula V and utilized in accordance with methods of the present invention can be obtained and/or modified from commercial sources. In other embodiments, polysaccharides comprising the structure of Formula IV and/or Formula V can be obtained by enzymatically or chemically *N*-sulfating

polysaccharides isolated and modified from bacterial or eukaryotic sources. In still other embodiments, polysaccharides comprising the structure of Formula IV and/or Formula V can be obtained by isolating and purifying the sulfated polysaccharide products of any of the other engineered aryl sulfate-dependent sulfotransferases utilized in conjunction with methods of the present invention. Each of these processes are discussed in detail in the description and examples, below.

[0213] Natural 2OSTs within the EC 2.8.2.- enzyme class generally comprise approximately 325-375 amino acid residues that in some cases vary greatly in their sequence, yet ultimately have the exact same function, namely, to catalyze the transfer of a sulfo group from PAPS to the 2-*O* position of hexuronyl residues within heparosan-based polysaccharides, particularly those comprising the structure of Formula IV and/or Formula V. Without being limited by a particular theory, it is believed that each of the natural 2OSTs can catalyze the same chemical reaction because there are multiple amino acid sequence motifs and secondary structures that are either identical or highly conserved across all species.

[0214] Further, it is believed that several of the conserved amino acid sequence motifs are directly involved in binding of either PAPS and/or the polysaccharide, or participate in the chemical reaction itself. The identity between the natural 2OST enzymes can be demonstrated by comparing the amino acid sequence of enzymes with a known crystal structure (e.g. chicken 2-*O* sulfotransferase, PDB codes: 3F5F and 4NDZ), in which amino acid residues within the active site have been identified, with the amino acid sequences of other 2OSTs within the EC 2.8.2.- enzyme class. A multiple sequence alignment of twelve enzymes, including the chicken, human, and other 2OST enzymes, is shown in Figures 14A-14D, along with percent identity relative to the chicken 2OST reference sequence (UniProtKB Accession No. Q76KB1). As illustrated in Figures 14A-14D, sequences range from having 94.9% sequence identity with the Q76KB1 reference sequence (entry tr|T1DMV2|T1DMV2_CROHD) for the timber rattlesnake 2OST, down to 56.3% sequence identity (entry tr|A0A131Z2T4| A0A131Z2T4_RHIAP) for the brown ear tick 2OST. The human enzyme (entry sp|Q7LGA3|HS2ST_HUMAN) has 94.1% sequence identity with the Q76KB1 reference sequence. Those skilled in the art would appreciate that the multiple sequence alignment was limited to twelve sequences for clarity, and that there are hundreds of amino acid sequences encoding for natural 2OST enzymes that have been identified and that have highly conserved active site and/or binding regions as well.

[0215] Within Figures 14A-14D, amino acids that are depicted in white with a black background at a particular position, are 100% identical across all sequences. Amino acids that are highly conserved,

meaning that the amino acids are either identical, or chemically or structurally similar, at a particular position are enclosed with a black outline. Within highly conserved regions, consensus amino acids that are present in a majority of the sequences are in bold. Amino acids at a particular position that are not identical or highly conserved are typically variable. A period within a sequence indicates a gap that has been inserted into the sequence in order to facilitate the sequence alignment with other sequence(s) that have additional residues between highly conserved or identical region. Finally, above each block of sequences are a series of arrows and coils that indicate secondary structure that is conserved across all sequences, based on the identity of the amino acids within the alignment and using the structure of the natural chicken 2OST enzyme as a reference. The β symbol adjacent to an arrow refers to a β -sheet, whereas a coil adjacent to an α symbol or a η symbol refers to a helix secondary structure.

[0216] Within the twelve aligned sequences in Figures 14A-14D, there are several conserved amino acid motifs that include one or more amino acids that comprise the active site, based on the crystal structures of the chicken 2OST enzyme described above. Based on the numbering of the amino acid residues within Figures 14A-14D, these motifs include residues 12-19 (R-V-P-K-T-A/G-S-T), residues 40-44 (N-T-S/T-K-N), residues 71-74 (Y-H-G-H), residues 108-115 (F-L-R-F/H-G-D-D/N-F/Y), residues 121-125 (R-R-K/R-Q-G), and residues 217-222 (S-H-L-R-K/R-T). Without being limited by a particular theory, it is believed that these residues either facilitate or participate in the chemical reaction, or enable binding of PAPS or the polysaccharide within the active site. In particular and as illustrated in Figures 15A-15C, the histidine residue at position 74 abstracts the proton from the 2-*O* position of the iduronic acid residue within the polysaccharide, enabling nucleophilic attack and removal of the sulfo group from PAPS, whereas the lysine residue at position 15 coordinates with the phosphate moiety of PAPS to stabilize the transition state of the enzyme before the N,2-HS product is released from the active site.

[0217] However, as described above, the natural 2OST enzymes are unable to catalyze the transfer of the sulfate group from an aryl sulfate compound to the polysaccharide. As with the NDSTs, it is believed that the binding pocket for PAPS within the active site of the natural sulfotransferase either does not have a high enough affinity for aryl sulfate compounds to facilitate binding and/or that the aryl sulfate compounds are sterically hindered from entering the active site. Consequently, and in another embodiment, a natural 2OST enzyme can be mutated in several locations within its amino acid sequence to enable binding of the aryl sulfate compound within the active site and/or to optimally position the aryl sulfate compound so transfer of the sulfate group to the polysaccharide can occur.

[0218] Accordingly, and in another embodiment, engineered 2OST enzymes that can be utilized with methods of the present invention can be mutants of natural 2OST enzymes within EC 2.8.2.-, including enzymes having the amino acid sequences illustrated in Figures 14A-14D. In another embodiment, the aryl sulfate-dependent, 2OSTs have been engineered to recognize, bind, and react with aryl sulfate compounds as sulfo group donors, while retaining the natural enzymes' ability to recognize, bind, and react with *N*-sulfated, heparosan-based polysaccharides, particularly those comprising the structure of Formula IV and/or Formula V, as sulfo group acceptors. Without being limited by a particular theory, it is believed that because of the mutations inserted into the amino acid sequences of the engineered 2OST enzymes, their sulfotransferase activity may comprise the direct transfer of a sulfuryl group from an aryl sulfate compound to the sulfo acceptor polysaccharide, using a similar mechanism as described in Figures 15A-15C above, except that the PAPS is substituted with the aryl sulfate compound. Otherwise, it is believed that the mutations may cause the sulfotransferase activity to comprise a two-step process including the hydrolysis of an aryl sulfate compound and formation of a sulfohistidine intermediate, followed by the nucleophilic attack of the sulfohistidine intermediate by the oxygen atom at the 2-*O* position of a hexuronic acid residue, to form the N,2-HS product. By either mechanism, engineered 2OST enzymes achieve sulfo transfer from an aryl sulfate compound to a polysaccharide, as described in the examples, below.

[0219] In another embodiment, an engineered 2OST enzyme can comprise one or more mutated amino acid sequence motifs relative to the conserved amino acid sequence motifs described above that are found in the natural 2OST enzymes within EC 2.8.2.-, as described above and indicated in the multiple sequence alignment in Figures 14A-14D. In another embodiment, each mutated amino acid sequence motif that is present in the amino acid sequence of the engineered enzyme comprises at least one amino acid mutation relative to the corresponding conserved amino acid sequence motif within the natural 2OST enzymes. In another embodiment, an engineered 2OST enzyme can comprise one mutated amino acid sequence motif. In another embodiment, an engineered 2OST enzyme can comprise two mutated amino acid sequence motifs. In another embodiment, an engineered 2OST enzyme can comprise three mutated amino acid sequence motifs. In another embodiment, an engineered 2OST enzyme can comprise four mutated amino acid sequence motifs. In another embodiment, an engineered 2OST enzyme can comprise five mutated amino acid sequence motifs. In another embodiment, an engineered 2OST enzyme can comprise six mutated amino acid sequence motifs. In another embodiment, an engineered 2OST enzyme that includes at least one mutated amino acid sequence motif relative to any of the natural 2OST enzymes within EC

2.8.2.- can have an amino acid sequence selected from the group consisting of SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 41, and SEQ ID NO: 42.

[0220] In another embodiment, upon viewing the crystal structure of the chicken 2OST (PDB code: 3F5F) within a 3D molecular visualization system (including, as a non-limiting example, the open-source software, PyMOL), the structure of related sequences, such as those of engineered 2OST enzymes that contain one or more mutated amino acid sequence motifs relative to the chicken sulfotransferase structure, can be modeled for comparison as illustrated in Figure 16. Figure 16 shows a magnified view of the active site of the chicken 2OST enzyme overlaid with two engineered 2OST enzymes, comprising the amino acid sequences of SEQ ID NO: 14 and SEQ ID NO: 16, in which the structure of the engineered enzyme is calculated upon making mutations relative to the chicken 2OST amino acid sequence. Adenosine 3',5'-diphosphate, which is the product of a sulfotransfer reaction in which PAPS is the sulfo donor, and which was co-crystallized with the chicken 2OST, is also illustrated within the active site. The sulfate group that would be present in the natural substrate, PAPS, is modeled onto the 5'-phosphate functional group to illustrate its approximate position within the active site prior to initiating the reaction. NCS is also modeled into the active site of the engineered enzymes, using the consensus solutions of molecular dynamics (MD) simulations that designed to calculate the optimized position and orientation of a ligand within an enzyme active site adjacent to the polysaccharide binding site (not shown), if such solutions are possible. Hydrogen atoms are not shown.

[0221] As illustrated in Figure 16, although there are several mutations made to SEQ ID NO: 14 and SEQ ID NO: 16, relative to the chicken 2OST, the respective protein backbones are in a nearly identical location to one another, enabling a one-to-one comparison of the active sites. When comparing the two active sites, the PAPS is located in the background and adjacent to a lysine residue (position 15 of the Q76KB1 sequence in Figures 14A-14D), whereas the convergent solutions from the above MD simulations indicate that NCS binding within the engineered enzymes is favored on the opposite side of the active site. However, binding of NCS would be sterically hindered in the natural enzyme in part by the lysine residue as well as the phenylalanine residue located on the nearby α -helix (position 108 of the Q76KB1 sequence in Figures 14A-14D). Without being limited by a particular theory, it is believed that binding of NCS in the active site of the engineered enzyme comprising the amino acid sequence of SEQ ID NO: 14 is facilitated by the mutation of the lysine residue to a histidine residue, which creates additional space within the active site and provides a π - π stacking partner for the aromatic ring within NCS. Also without being limited by a particular theory, it is believed that binding of NCS in the active site of the engineered

enzyme comprising the amino acid sequence of SEQ ID NO: 16 is facilitated by the mutation of the lysine to an arginine residue in concert with the adjacent mutation of the proline residue (position 14 of the Q76KB1 sequence in Figures 14A-14D) to a histidine residue. The increased number of conformational degrees of freedom of the arginine side chain facilitate entry of the NCS while still being in a position to provide a polar contact to stabilize the transition state during the transfer reaction, whereas the adjacent histidine provides other binding contacts for NCS.

[0222] Another mutation of note includes the mutation from an arginine residue (position 220 of the Q76KB1 sequence in Figures 14A-14D) to a histidine residue, a mutation that is found at position 221 in both SEQ ID NO: 14 and SEQ ID NO: 16. Without being limited by a particular theory, the mutated histidine residue is in a favorable position to facilitate removal of the sulfate group from NCS. Other illustrated mutations from the chicken 2OST enzyme, particularly mutations present in SEQ ID NO: 16 (His-20, Ser-114, Lys-116, Met-122) may similarly drive binding of NCS within the active site, either by providing a direct binding contact with the sulfate moiety within NCS (His-20), coordinating with other mutated residues (Ser-114 coordinating with His-221), or by increasing the hydrophobic environment near NCS (Met-122).

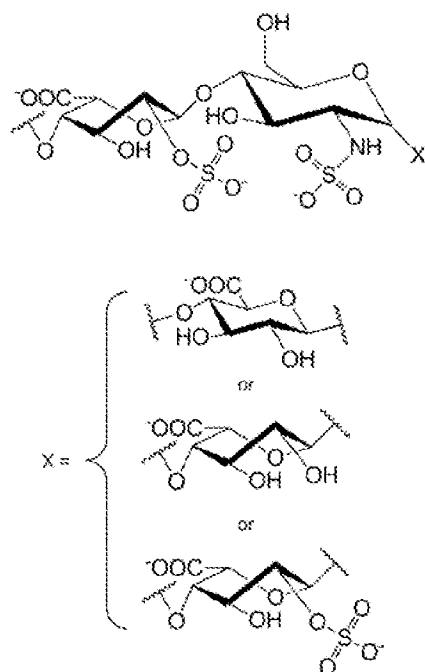
[0223] Those skilled in the art would appreciate that engineered 2OST enzymes of any other amino acid sequence, including, but not limited to, those disclosed by SEQ ID NO: 41 and SEQ ID NO: 42, would likely exhibit a similar structure to the chicken 2OST, as well as engineered 2OSTs having the amino acid sequence of SEQ ID NO: 14 and SEQ ID NO: 16. Without being limited by a particular theory, it is believed that PNS would bind in a similar position as NCS within the active site of any of the engineered 2OST enzymes, since the structures of the two aryl sulfate compounds are very similar, except that the sulfate group is located *ortho* on the aromatic ring relative to the nitro group in NCS, rather than *para* to the nitro group in PNS.

[0224] Accordingly, in another embodiment, an engineered 2OST enzyme utilized in accordance with any of the methods of the present invention can comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 41, or SEQ ID NO: 42. In another embodiment, any of the above 2OST enzymes react with an aryl sulfate compound, instead of PAPS, as a sulfo group donor. In further embodiments, the aryl sulfate compound is selected from the group consisting of PNS, MUS, 7-hydroxycoumarin sulfate, phenyl sulfate, 4-acetylphenyl sulfate, indoxyl sulfate, 1-naphthyl sulfate, 2-naphthyl sulfate, and NCS. In some even further embodiments, the aryl sulfate compound is PNS. In other even further embodiments, the aryl sulfate compound is NCS.

[0225] In another embodiment, within reaction mixtures that comprise any natural or engineered 2OST enzyme, particularly an engineered 2OST enzyme comprising the amino acid sequence of SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 41, or SEQ ID NO: 42, the reaction mixture can further comprise an hexuronyl C₅-epimerase to catalyze formation of an N,2-HS product. In some embodiments, the N,2-HS product can comprise the structure of Formula VI. In other embodiments, the N,2-HS product can comprise the structure of Formula VII. In another embodiment, any isolated or recombinant hexuronyl C₅-epimerase can be used. In another embodiment, the hexuronyl C₅-epimerase can comprise the amino acid sequence of SEQ ID NO: 29. In another embodiment, the hexuronyl C₅-epimerase can comprise residues 34-617 of SEQ ID NO: 29.

Glucosaminyl 6-O sulfotransferases

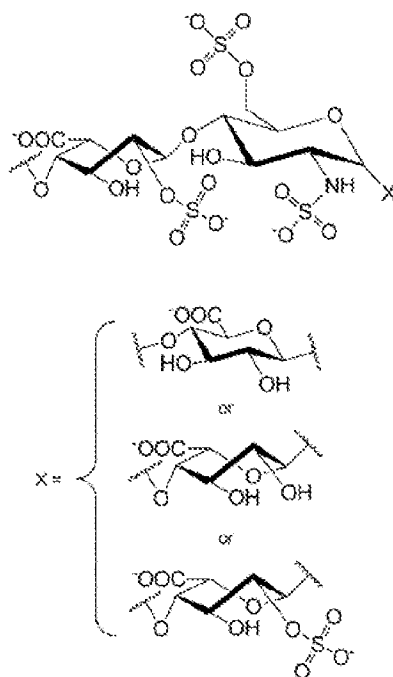
[0226] In nature, 6OSTs recognize, bind, and react with heparosan-based polysaccharides as sulfo group acceptors. Generally, a majority of the glucosaminyl residues are *N*-sulfated, but the enzymes can still transfer sulfo groups to the 6-*O* position of glucosaminyl residues that are *N*-acetylated. Additionally, either adjacent hexuronic acid residue can be either glucuronic acid or iduronic acid, and can optionally be 2-*O* sulfated. Generally, the hexuronic acid at the non-reducing end of the glucosamine residue receiving the 6-*O* sulfo group is 2-*O* sulfated iduronic acid, and in many instances, the glucosamine residue itself is also *N*-sulfated. Similar to the NSTs and 2OSTs, naturally-occurring 6OST enzymes transfer the sulfo group to the polysaccharide upon reacting with PAPS as a sulfo group donor. As with natural 2OSTs, natural 6OST enzymes are also members of the EC 2.8.2.- enzyme class. In a non-limiting example, either natural or engineered 6OST enzymes can recognize, bind, and react with heparosan-based polysaccharides comprising the structure of Formula VIII, below:



wherein the glucosamine residue receiving the 6-*O* sulfo group is *N*-sulfated and is adjacent to a 2-*O* sulfated iduronic acid residue at its non-reducing end, and X comprises any of the hexuronyl residues depicted in Formula VIII, above. 6OST enzymes within EC 2.8.2.- having biological activity with polysaccharides comprising the structure of Formula VIII have been described by Xu, Y., et al., (2017) *ACS Chem. Biol.* **12** (1):73-82 and Holmborn, K., et al., (2004) *J. Biol. Chem.* **279**, (41):42355-42358, the disclosures of which are incorporated by reference in their entireties.

[0227] As described above, although the portion of the heparosan-based polysaccharide that reacts with the 6OST enzyme can comprise the structure of Formula VIII, other portions of the polysaccharide can be *N*- or *O*- substituted, and can comprise other structural motifs that can also react with the enzyme. Similar to the other enzymes above, 6OST enzymes can transfer a sulfo group to multiple positions within the same polysaccharide molecule, and multiple positions within the same polysaccharide molecule can be 6-*O* sulfated by the same enzyme molecule. Typically, heparosan-based polysaccharides that can react with 6OST enzymes, including those comprising the structure of Formula VIII, can comprise at least three monosaccharide residues.

[0228] Upon successfully binding PAPS and a heparosan-based polysaccharide comprising the structure of Formula VIII, natural 6OST enzymes can catalyze transfer of the sulfo group to the 6-*O* position of the glucosamine residue, forming an N,2,6-HS product comprising the structure of Formula IX, below:



wherein X comprises any of the hexuronyl residues depicted in Formula IX, above. Similarly, an engineered 6OST enzyme that binds and reacts with an aryl sulfate compound and a heparosan-based polysaccharide comprising the structure of Formula VIII can form an N,2,6-HS product comprising the structure of Formula IX.

[0229] A non-limiting example of one such polysaccharide sulfo acceptor that can react with an 6OST enzyme is illustrated in Figure 17. Figure 17 shows a heparosan-based polysaccharide 240 that includes three *N*-substituted glucosamine residues 210 that can be *N*-substituted with either an acetyl group 211 or a sulfate group 212. Within the polysaccharide 240, *N*-substituted glucosamine residues 210 that are capable of acting as a sulfo acceptor are flanked by two hexuronyl residues. Hexuronyl residues can include any residue represented by the functional group “X” in Formula VIII, particularly glucuronyl residue 220 and iduronyl residue 230. Either the glucuronyl residue 220 or iduronyl residue 230 can further be substituted by a sulfate group 231 at the 2-*O* position. Upon reacting the polysaccharide 240 with an 6OST enzyme and a sulfo group donor, the 6-*O* position 213 of any of the glucosamine residues 210 can be sulfated, ultimately forming 6-*O* sulfated glucosamine residues 310 within the product polysaccharide 241. In another embodiment, the 6OST enzyme can be an engineered aryl sulfate-dependent enzyme, and the sulfo group donor is an aryl sulfate compound.

[0230] In another embodiment, engineered 6OSTs that can be utilized in accordance with methods of the present invention can have the same biological activity with heparosan-based sulfo acceptor polysaccharides as natural 6OSTs, particularly heparosan-based polysaccharides comprising the

structure of Formula VIII. In another embodiment, when there are multiple portions of the polysaccharide comprising the structure of Formula VIII within the sulfo acceptor polysaccharide, any glucosamine residue can be sulfated by the engineered 6OST enzyme. Similarly, the same polysaccharide can be sulfated multiple times by the engineered 6OST, including and up to all of the glucosamine residues that are present within the polysaccharide.

[0231] In another embodiment, sulfo acceptor polysaccharides that can react with an engineered or natural 6OST, including but not limited to those comprising the structure of Formula VIII, can be provided as a homogenous composition. In still other embodiments, sulfo acceptor polysaccharides that can react with an engineered or natural 6OST can be comprised within a composition comprising a polydisperse mixture of polysaccharides having variable chain lengths, molecular weights, relative abundance of Formula VIII, and overall monosaccharide composition and functionalization.

[0232] In another embodiment, N,2-HS polysaccharides, including but not limited to those comprising the structure of Formula VIII, and utilized in accordance with methods of the present invention with either an engineered or natural 6OST enzyme can be obtained and/or modified from commercial sources. In another embodiment, either an engineered or natural 6OST can be utilized in accordance with methods of the present invention can react with *N*-sulfated heparosan products produced by an NST enzyme in one or more previous steps. In another embodiment, either an engineered or natural 6OST that can be utilized in accordance with methods of the present invention can react with N,2-HS products produced by an NST and/or a 2OST in one or more previous steps. In another embodiment, one or more of the sulfation steps to produce the N,2-HS product was catalyzed by an engineered, aryl sulfate-dependent sulfotransferase. Each of these processes are discussed in detail in the description and examples, below.

[0233] Natural 6OST enzymes within the EC 2.8.2.- enzyme class generally comprise between 300 and 700 amino acid residues that can in some cases vary greatly in their sequence, yet ultimately have the exact same function, namely, to catalyze the transfer of a sulfuryl group from PAPS to the 6-*O* position of glucosamine residues within heparosan-based polysaccharides, particularly those comprising the structure of Formula VIII. Without being limited by a particular theory, it is believed that each of the natural 6OSTs can catalyze the same chemical reaction because there are multiple amino acid sequence motifs and secondary structures that are either identical or highly conserved across all species.

[0234] Further, it is believed that several of the conserved amino acid sequence motifs are directly involved in binding of either PAPS and/or the polysaccharide, or participate in the chemical reaction

itself. The identity between the natural 6OST enzymes can be demonstrated by comparing the amino acid sequence of an enzyme with a known crystal structure (zebrafish 6OST isoform 3-B, PDB codes 5T03, 5T05 and 5T0A), in which amino acid residues within the active site have been identified, with the amino acid sequences of other natural 6OSTs. A multiple sequence alignment of fifteen enzymes is shown in Figures 18A-18C, along with the percent identity of each sequence relative to the mouse 6OST (isoform 1) reference sequence (UniProtKB Accession No. Q9QYK5). As illustrated in Figures 18A-18C, sequences range from having 97.3% identity with the Q9QYK5 reference sequence (entry O60243|H6ST1_HUMAN) down to 53.7% identity (entry A0A3P8W3M9|A0A3P8W3M9_CYSNE). For comparison, the zebrafish 6OST isoform 3-B enzyme (entry A0MGZ7|H6S3B_DANRE) has 60.4% sequence identity with the Q9QYK5 reference sequence. Those skilled in the art would appreciate that the multiple sequence alignment was limited to fifteen sequences for clarity, and that there are hundreds of amino acid sequences encoding for natural 6OST enzymes that have been identified and that have highly conserved active site and/or binding regions as well.

[0235] Within Figures 18A-18C, amino acids that are depicted in white with a black background at a particular position, are 100% identical across all sequences. Amino acids that are highly conserved, meaning that the amino acids are either identical or chemically or structurally similar, at a particular position are enclosed with a black outline. Within highly conserved regions, consensus amino acids that are present in a majority of the sequences, are in bold. Amino acids at a particular position that are not identical or highly conserved are typically variable. A period within a sequence indicates a gap that has been inserted into the sequence in order to facilitate the sequence alignment with other sequence(s) that have additional residues between highly conserved or identical region. Finally, above each block of sequences are a series of arrows and coils that indicate secondary structure that is conserved across all sequences, based on the identity of the amino acids within the alignment and using the structure of the natural mouse 6OST enzymes enzyme as a reference. The β symbol adjacent to an arrow refers to a β -sheet, whereas a coil adjacent to an α symbol refers to a helix secondary structure. Each of the fifteen aligned sequences in illustrated Figures 18A-18C have been truncated relative to their natural full-length sequences to coincide with the engineered enzymes of the present invention, particularly SEQ ID NO: 18, SEQ ID NO: 20, and SEQ ID NO: 22. In particular, the residues illustrated in Figures 18A-18C are aligned with residues 67-377 of the Q9QYK5 reference sequence for the mouse 6OST.

[0236] Within the fifteen aligned sequences in Figures 18A-18C, there are several conserved amino acid sequence motifs that include one or more amino acids that comprise the active site, based on the

crystal structure of the zebrafish 6OST enzyme (entry A0MGZ7|H6S3B__DANRE) described above. Based on the numbering of the amino acid residues within Figures 18A-18C, these conserved amino acid sequence motifs include amino acid residues 29 through 34 (Q-K-T-G-G-T); 81 through 86 (C-G-L-H-A-D); 127 through 139 (S-E-W-R/K-H-V-Q-R-G-A-T-W-K); 178 through 184 (N-L-A-N-N-R-Q); and 227 through 231 (L-T-E-F/Y-Q). In particular, and as illustrated in the reaction mechanism in Figures 19A-19C, the histidine residue within the C-G-L-H-A-D conserved amino acid sequence motif is in position to abstract the hydrogen atom from the 6' hydroxyl group of an *N*-sulfoglucosamine residue, enabling the negatively-charged oxygen atom to then initiate the nucleophilic attack of PAPS and remove the sulfate group. Additionally, the universally conserved lysine residue within the Q-K-T-G-G-T conserved amino acid sequence motif coordinates with the 5'-phosphate in PAPS, while the universally conserved histidine and tryptophan residues at positions 131 and 138 coordinate with the *N*-sulfoglucosamine residue (*see* Xu, Y., et al., above).

[0237] However, as described above, natural 6OST enzymes are unable to catalyze the transfer of the sulfate group from an aryl sulfate compound to a polysaccharide. Without being limited by a particular theory, and as with the NSTs and 2OSTs described above, it is believed that the binding pocket for PAPS within the active site of the natural 6OST either does not have a high enough affinity for aryl sulfate compounds to facilitate binding and/or that the aryl sulfate compounds are sterically hindered from entering the active site. Consequently, and in another embodiment, a natural 6OST enzyme can be mutated in several locations within its amino acid sequence to enable binding of the aryl sulfate compound within the active site and/or to optimally position the aryl sulfate compound so transfer of the sulfate group to the polysaccharide can occur.

[0238] Accordingly, and in another embodiment, engineered 6OST enzymes that can be utilized with methods of the present invention can be mutants of natural 6OST enzymes within EC 2.8.2.-, including enzymes having the amino acid sequences illustrated in Figures 18A-18C. In another embodiment, the engineered 6OST enzymes have been engineered to recognize, bind, and react with aryl sulfate compounds as sulfo group donors, while retaining the natural enzymes' ability to recognize, bind, and react with any of the HS polysaccharides described above, including but not limited to those comprising the structure of Formula VIII, as sulfo group acceptors. Without being limited by a particular theory, it is believed that because of the mutations inserted into the amino acid sequences of the engineered 6OST enzymes, their sulfotransferase activity may comprise the direct transfer of a sulfonyl group from an aryl sulfate compound to the sulfo acceptor polysaccharide, using a similar mechanism as described in Figures 19A-19C, above, except that the PAPS is substituted with the aryl sulfate compound. Otherwise, it is believed that the mutations may

cause the sulfotransferase activity to comprise a two-step process including the hydrolysis of an aryl sulfate compound and formation of a sulfohistidine intermediate, followed by the nucleophilic attack of the sulfohistidine intermediate by the oxygen atom at the 6-*O* position of a glucosamine residue, to form a 6-*O* sulfated HS product. In another embodiment, the 6-*O* sulfated HS product of either sulfotransfer mechanism is an N,2,6-HS product.

[0239] In another embodiment, an engineered 6OST enzyme can comprise one or more mutated amino acid sequence motifs relative to the conserved amino acid sequence motifs found in natural 6OST enzymes within EC 2.8.2.-, as described above and indicated in the multiple sequence alignment in Figures 18A-18C. In another embodiment, each mutated amino acid sequence motif that is present in the amino acid sequence of the engineered enzyme comprises at least one amino acid mutation relative to the corresponding conserved amino acid sequence motif within the natural 6OST enzymes. In another embodiment, an engineered 6OST enzyme can comprise one mutated amino acid sequence motif. In another embodiment, an engineered 6OST enzyme can comprise two mutated amino acid sequence motifs. In another embodiment, an engineered 6OST enzyme can comprise three mutated amino acid sequence motifs. In another embodiment, an engineered 6OST enzyme can comprise four mutated amino acid sequence motifs. In another embodiment, an engineered 6OST enzyme can comprise five mutated amino acid sequence motifs. In another embodiment, an engineered 6OST enzyme that includes at least one mutated amino acid sequence motif relative to any of the natural 6OST enzymes within EC 2.8.2.- can have an amino acid sequence selected from the group consisting of SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 59, SEQ ID NO: 60, and SEQ ID NO: 61.

[0240] In another embodiment, upon viewing any of the crystal structures of the zebrafish 6OST (UniProtKB Accession No. A0MGZ7) within a 3D molecular visualization system (including, as a non-limiting example, the open-source software, PyMOL), the structure of related sequences, such as those of engineered 6OST enzymes that contain one or more mutated amino acid sequence motifs relative to any of the zebrafish 6OST structures, can be modeled for comparison as illustrated in Figure 20. Figure 20 shows a magnified view of the active site of the zebrafish 6OST enzyme (PDB code: 5T03) with one of the engineered enzymes of the present invention, comprising the amino acid sequence of SEQ ID NO: 22, in which the structure of the engineered 6OST enzyme is calculated upon making mutations relative to the zebrafish 6OST amino acid sequence. Adenosine 3',5'-diphosphate, which is the product of a sulfotransfer reaction in which PAPS is the sulfo donor,

and which was co-crystallized with the zebrafish 6OST, is also illustrated within the active site. PNS is also modeled into the active site of the engineered enzymes, using the consensus solutions of molecular dynamics (MD) simulations that designed to calculate the optimized position and orientation of a ligand within an enzyme active site adjacent to the polysaccharide binding site (not shown), if such solutions are possible. Hydrogen atoms are not shown for clarity.

[0241] As illustrated in Figure 20, although there are several mutations made SEQ ID NO: 22, relative to the zebrafish 6OST enzyme, the respective protein backbones are in a nearly identical location to one another, enabling a one-to-one comparison of the active sites. However, when comparing the two active sites, the adenosine 3',5'-diphosphate product is located on the opposite side of the central α -helix as the PNS molecule, as determined by the convergent solutions from the above MD simulations. Without being limited by a particular theory, it is believed that the convergent MD simulation solutions place PNS on the opposite side of the α -helix because there is not enough of an affinity toward PNS in the same or similar position as PAPS within the zebrafish enzyme. As described by Xu, Y., et al., above, the conserved histidine at position 158 of the full-length amino acid sequence is the catalytic histidine that abstracts the proton from the 6' hydroxyl group of *N*-sulfo glucosamine, which is then subsequently able to react with PAPS to initiate sulfo group transfer. Yet, despite the apparent differences in the binding pocket for PAPS and PNS, engineered 6OST enzymes comprising the amino acid sequences of SEQ ID NO: 18, SEQ ID NO: 20, and SEQ ID NO: 22 all achieved sulfo transfer from an aryl sulfate compound to the glucosaminyl 6-*O* position within a heparosan-based polysaccharide, as described in the examples below.

[0242] As a result, and without being limited by a particular theory, one or more mutations present within the active site of engineered 6OST enzymes may assist binding of the sulfate moiety of the aryl sulfate compound in a position in which it can be transferred to the sulfo acceptor HS polysaccharide. As illustrated in Figure 20, the engineered enzyme has the amino acid sequence SEQ ID NO: 22, and the aryl sulfate compound is PNS. However, a sulfo acceptor HS polysaccharide is not illustrated. In a non-limiting example, the histidine residue engineered into position 31 of SEQ ID NO: 22 may be in position to facilitate removal of the sulfate group from PNS using a ping-pong mechanism, as described in Malojcic, et al, above. Additionally, the histidine residue engineered into position 133 of SEQ ID NO: 22 may further coordinate with the sulfate moiety along with the conserved histidine at position 132 of SEQ ID NO: 22 (corresponding to positions 131-132 in each of the sequences in Figures 18A-18C). Mutation to G-A-N at positions 137-139 of SEQ ID NO: 22 (corresponding to the conserved A-T-W motif at positions 136-138 of

the sequences in Figures 18A-18C) removes steric bulk that may prevent binding of PNS in a position where the sulfate can be abstracted by the engineered histidine at position 31 of SEQ ID NO: 22. The mutations to G-A-N within the loop containing A-T-W also appears to cause the loop to move away from PNS, which may further assist PNS to reach its binding pocket. Finally, a serine residue engineered into position 84 of SEQ ID NO: 22, immediately adjacent to a native histidine corresponding to His-158 in the full-length zebrafish 6OST, described above, may create an additional hydrogen-binding contact to assist the engineered enzyme in retaining the zebrafish enzyme's natural activity with the sulfo acceptor polysaccharide.

[0243] Those skilled in the art would appreciate that engineered 6OST enzymes of any other amino acid sequence, including, but not limited to, those disclosed by SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 59, SEQ ID NO: 60, and SEQ ID NO: 61, would exhibit similar structural motifs, particularly within the active site. Without being limited by a particular theory, it is believed that NCS would bind in a similar position as PNS within the active site of any of the engineered enzymes, since the structures of the two aryl sulfate compounds are very similar, except that the sulfate group is located *ortho* on the aromatic ring relative to the nitro group, rather than *para* to the nitro group.

[0244] In another embodiment, engineered 6OST enzymes that can be utilized in accordance with methods of the present invention can comprise one or more mutated amino acid sequence motifs, which can be determined in-part by comparing conserved amino acid sequence motifs indicated in the multiple sequence alignment of Figures 18A-18C with the known structure(s) of natural enzymes and/or modeled engineered enzymes, including but not limited to, as a non-limiting example, enzymes illustrated in Figure 20. In another embodiment, mutated amino acid sequence motifs that can be comprised within an engineered 6OST enzyme can be selected from the group consisting of (a) G-H-T-G-G-T; (b) C-G-X₁-X₂-A-D, wherein X₁ is selected from the group consisting of threonine and serine, and X₂ is selected from the group consisting of asparagine, arginine, and histidine; (c) X₃-X₄-W-R-H-X₅-Q-R-G-G-X₆-N-K, wherein X₃ is selected from the group consisting of serine and glycine, X₄ is selected from the group consisting of glycine and histidine, X₅ is selected from the group consisting of histidine and threonine, and X₆ is selected from the group consisting of alanine and threonine; and (d) N-L-X₇-N-N-R-Q, wherein X₇ is selected from the group consisting of alanine and glycine; including any combination thereof. Each of the mutated amino acid sequence motifs corresponds with a conserved amino acid motif indicated in Figures 18A-18C above: sequence motif (a) corresponds to the conserved amino acid sequence motif, Q-K-T-G-G-T;

mutated amino acid sequence motif (b) corresponds to the conserved amino acid sequence motif, C-G-L-H-A-D; mutated amino acid sequence motif (c) corresponds to the conserved amino acid sequence motif, S-E-W-(R/K)-H-V-Q-R-G-A-T-W-K; and mutated amino acid sequence motif (d) corresponds to the conserved amino acid sequence motif, N-L-A-N-N-R-Q. In another embodiment, engineered 6OST enzymes comprising at least one mutated amino acid sequence motif described above can be selected from the group consisting of: SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 59, SEQ ID NO: 60, and SEQ ID NO: 61.

[0245] In another embodiment and in one non-limiting example, engineered 6OST enzymes can comprise the mutated amino acid sequence motifs (b) and (c) within the same amino acid sequence. Engineered enzymes comprising the mutated amino acid sequence motifs (b) and (c) include, but are not limited to, enzymes comprising the amino acid sequences of SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, and SEQ ID NO: 50. In another embodiment, each of the engineered 6OST enzymes comprising the mutated amino acid sequence motifs (b) and (c) have a similar active site as SEQ ID NO: 22, as illustrated in Figure 20. Without being limited to another theory, it is believed that several of the mutations comprised within mutated amino acid sequence motifs (b) and (c) have one or more functions during sulfotransferase activity, including not limited to: increasing the affinity of aryl sulfate compounds to the active site by reducing the size of the binding pocket, increasing the hydrophobicity of the pocket, removing or creating polar or hydrogen bonding contacts, and/or creating π - π interactions with the aromatic moieties of the aryl sulfate compounds; stabilizing the transition state of the enzyme during the chemical reaction; and/or participating in the chemical reaction itself.

[0246] In another embodiment, within engineered 6OST enzymes that comprise the mutated amino acid sequence motifs (b) and (c), X₄ is glycine and X₅ is histidine. In other embodiments, X₄ is histidine and X₅ is threonine.

[0247] In another embodiment, within engineered 6OST enzymes comprising the mutated amino acid sequence motifs (b) and (c), X₃ is serine, X₆ is alanine, and X₇ is glycine. In other embodiments, X₃ is glycine, X₆ is threonine, and X₇ is alanine.

[0248] Furthermore, the amino acid sequences (SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22) of three engineered 6OST enzymes, which have been experimentally determined to be active with aryl sulfate compounds as sulfo group donors (see Example 4 below) can be compared with the

amino acid sequence of the mouse 6OST enzyme (entry Q9QYK5[H6ST1_MOUSE]) in a multiple sequence alignment to determine if there are relationships between mutations among each of the enzymes. A period within the amino acid sequence of an engineered enzyme indicates identity at a particular position with the mouse 6OST enzyme. As shown in Figure 21, the sequence alignment demonstrates that while over 90% of the amino acid residues within the three sulfotransferase sequences are identical, there are several positions in which multiple amino acids can be chosen. Without being limited by a particular theory, these enzymes have a similar relationship with each other as the 6OST enzymes that comprise EC 2.8.2.-. As a result, and in another embodiment, engineered 6OST enzymes comprising an amino acid sequence in which multiple amino acids can be chosen at defined positions are disclosed as SEQ ID NO: 43 and SEQ ID NO: 44. Positions at which the identity of an amino acid can be chosen from a selection of possible residues are denoted in terms "Xaa," "Xn," or "position n," where n refers to the residue position.

[0249] In another embodiment, within SEQ ID NO: 43, residues having the designation, "Xaa," illustrate known instances in which there is a lack of identity at a particular position within the amino acid sequences of SEQ ID NO: 18, SEQ ID NO: 20, and SEQ ID NO: 22. In another embodiment, the amino acid sequence, SEQ ID NO: 44, also illustrates known instances in which there is a lack of identity at a particular position within the amino acid sequences of SEQ ID NO: 18, SEQ ID NO: 20, and SEQ ID NO: 22, but SEQ ID NO: 44 further comprises N-terminal residues 1-66, and C-terminal residues 378-411, of several full-length 6OST enzymes within EC 2.8.2.-, including, as non-limiting examples, the mouse, human, and pig 6OST enzymes. In contrast, amino acid residues in SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, and SEQ ID NO: 43 correspond with residues 67-377 of several full-length 6OST enzymes within EC 2.8.2.-, including, as non-limiting examples, the mouse, human, and pig 6OST enzymes. To facilitate protein expression, an N-terminal methionine residue was added to each SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, and SEQ ID NO: 43 amino acid sequence, relative to residues 67-377 of the mouse, human, and pig 6OST enzymes.

[0250] In another embodiment, any selection can be made for an Xaa residue, defined by the amino acid sequence SEQ ID NO: 43 or SEQ ID NO: 44, so long as the resulting enzyme maintains its 6OST activity upon reacting with an aryl sulfate compound as a sulfo group donor.

[0251] In another embodiment, within an engineered 6OST enzyme comprising the amino acid sequence of SEQ ID NO: 43, the amino acid residue at position 129 is glycine and the amino acid residue at position 133 is histidine. In another embodiment, within an engineered 6OST enzyme comprising the amino acid sequence of SEQ ID NO: 43, the amino acid residue at position 129 is

histidine and the amino acid residue at position 133 is threonine. In another embodiment, within an engineered 6OST enzyme comprising the amino acid sequence of SEQ ID NO: 44, the amino acid residue at position 194 is glycine and the amino acid residue at position 198 is histidine. In another embodiment, within an engineered 6OST enzyme comprising the amino acid sequence of SEQ ID NO: 44, the amino acid residue at position 194 is histidine and the amino acid residue at position 198 is threonine.

[0252] In another embodiment, within an engineered 6OST enzyme comprising the amino acid sequence of SEQ ID NO: 43, the amino acid residue at position 128 is serine, the amino acid residue at position 138 is alanine, and the amino acid residue at position 181 is glycine. In another embodiment, within an engineered 6OST enzyme comprising the amino acid sequence of SEQ ID NO: 43, the amino acid residue at position 128 is glycine, the amino acid residue at position 138 is threonine, and the amino acid residue at position 181 is alanine. In another embodiment, within an engineered 6OST enzyme comprising the amino acid sequence of SEQ ID NO: 44, the amino acid residue at position 193 is serine, the amino acid residue at position 203 is alanine, and the amino acid residue at position 246 is glycine. In another embodiment, within an engineered 6OST enzyme comprising the amino acid sequence of SEQ ID NO: 44, the amino acid residue at position 193 is glycine, the amino acid residue at position 203 is threonine, and the amino acid residue at position 246 is alanine.

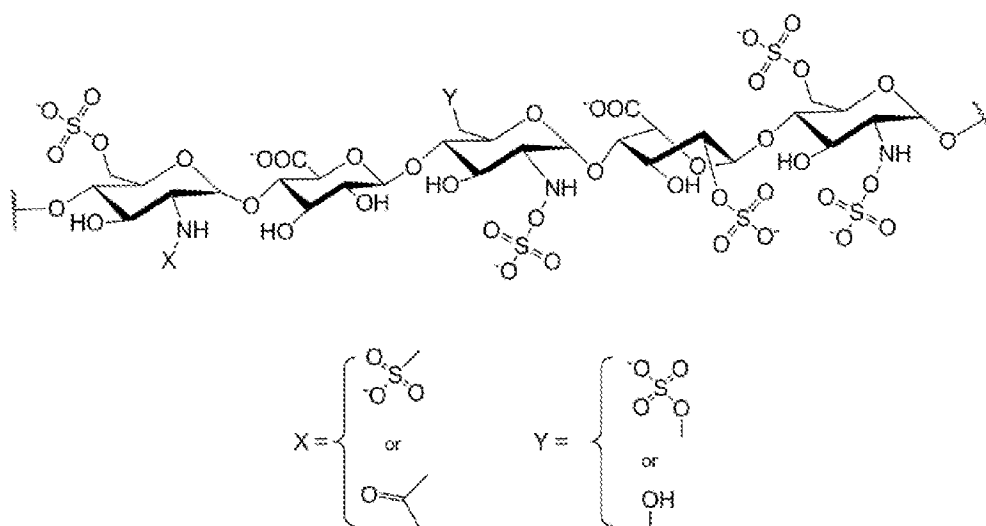
[0253] In another embodiment, within an engineered 6OST enzyme comprising the amino acid sequence of SEQ ID NO: 43 or SEQ ID NO: 44, the amino acid sequence can optionally include one or more mutations at residue positions not specified by an “Xn” or “Xaa,” so long as any such mutations do not eliminate the 6OST and/or aryl sulfate-dependent activity of the enzyme. In another embodiment, such mutations not eliminating aryl sulfate-dependent activity at positions not specified by an “Xn” or “Xaa” can include substitutions, deletions, and/or additions.

[0254] Accordingly, in another embodiment, an engineered 6OST enzyme utilized in accordance with any of the methods of the present invention can comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 59, SEQ ID NO: 60, and SEQ ID NO: 61. In another embodiment, any of the above engineered 6OST enzymes react with an aryl sulfate compound, instead of PAPS, as a sulfo group donor. In further embodiments, the aryl sulfate compound is selected from the group consisting of PNS, MUS, 7-hydroxycoumarin sulfate, phenyl sulfate, 4-acetylphenyl sulfate, indoxyl sulfate, 1-naphthyl sulfate, 2-naphthyl sulfate, and NCS. In some even further embodiments, the

aryl sulfate compound is PNS. In other even further embodiments, the aryl sulfate compound is NCS.

Glucosaminyl 3-O sulfotransferases

[0255] In nature, 3OSTs generally recognize, bind, and react with N,2-HS polysaccharides and N,2,6-HS polysaccharides as sulfo group acceptors. Generally, the glucosamine residue that receives the sulfo group at the 3-*O* position is *N*-sulfated, and is optionally also 6-*O* sulfated. Additionally, either adjacent hexuronic acid residue can be either glucuronic acid or iduronic acid, and can optionally be 2-*O* sulfated. In some embodiments, the hexuronic acid residue on the non-reducing end of the glucosamine residue is unsulfated glucuronic acid, while the hexuronic acid residue on the reducing end of the glucosamine residue is 2-*O* sulfated iduronic acid. Similar to each of the natural sulfotransferases described above, naturally-occurring 3OSTs transfer the sulfo group to the polysaccharide upon reacting with PAPS as a sulfo group donor. Natural 3OST enzymes that utilize PAPS as the sulfo group donor are members of the EC 2.8.2.23 enzyme class. In a non-limiting example, both natural 3OST enzymes and engineered aryl sulfate-dependent 3OST enzymes can recognize, bind, and react with N,2,6-HS polysaccharides comprising the structure of Formula X, below:

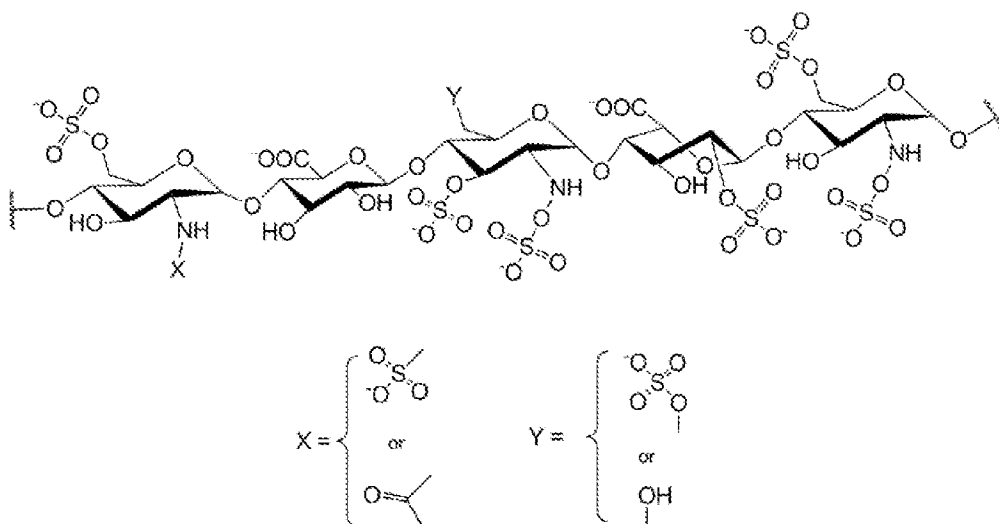


wherein the central glucosamine residue is *N*-sulfated and is adjacent to an unsubstituted glucuronic acid residue at its non-reducing end and a 2-*O* sulfated iduronic acid residue at its reducing end, X can optionally be a sulfate group or an acetyl group, and Y can optionally be a sulfate group or a hydroxyl group.

[0256] As described above, although the portion of the polysaccharide that reacts with the enzyme comprises the structure of Formula X, other portions of the polysaccharide can be *N*- or

O-substituted, and can comprise other structural motifs that can also react with the enzyme. Similar to the other enzymes above, 3OST enzymes can transfer a sulfo group to multiple positions within the same polysaccharide molecule, and multiple positions within the same polysaccharide molecule can be 3-*O* sulfated by the same enzyme molecule. Typically, HS polysaccharides that can react with 3OSTs as sulfo group acceptors typically comprise at least five monosaccharide residues, as shown in Formula X. In another embodiment, polysaccharides comprising the structure of Formula X and can react with 3OSTs as sulfo group acceptors can comprise at least 32 monosaccharide residues.

[0257] Upon successfully binding PAPS and an N,2,6-HS polysaccharide comprising the structure of Formula X, natural 3OST enzymes can catalyze transfer of the sulfo group to the 3-*O* position of the central glucosamine residue, forming an N,2,3,6-HS product comprising the structure of Formula I, below:



wherein X is either a sulfo group or an acetate group and Y is either a sulfo group or a hydroxyl group. Similarly, engineered 3OST enzymes that react with an aryl sulfate compound and an N,2,6-HS polysaccharide comprising the structure of Formula X can also form an N,2,3,6-HS product comprising the structure of Formula I. In further embodiments, the functional group X in the N,2,3,6-HS product is a sulfate group. In other further embodiments, the functional group Y in the N,2,3,6-HS product is a sulfate group. In another embodiment, in some locations within the polymer, at least a portion of the glucosamine residues are *N*-acetylated. Natural 3OST enzymes within EC 2.8.2.23, which have biological activity with N,2,6-HS polysaccharides comprising the structure of Formula X as sulfo group acceptors and form N,2,3,6-HS products comprising the structure of Formula I, have been described by Xu, D., et al., (2008) *Nat. Chem. Biol.* 4(3): 200-202 and

Edavettal, S.C., et al., (2004) *J. Biol. Chem.* **24**(11): 25789-25797, the disclosures of which are incorporated by reference in their entireties.

[0258] A non-limiting example of one such N,2,6-HS sulfo group acceptor for 3OST enzymes is illustrated in Figure 22. Figure 22 shows a polysaccharide 440 that includes three glucosamine residues 410 comprising an *N*-sulfo group 411 at each *N*-position and an *O*-sulfo group 412 at each 6-*O* position. Within the polysaccharide 440, glucosamine residues 410 that are capable of acting as a sulfo acceptor must be flanked by two hexuronic acid residues. Hexuronic acid residues can include any residue represented by the functional group “X” in Formula X, and are shown in Figure 22 as glucuronic acid residue 420 and iduronic acid residue 430. Either hexuronic acid residue can further be substituted by a sulfo group 431 at the 2-*O* position. Upon reacting the polysaccharide 440 with an 3OST enzyme and a sulfo group donor, the 3-*O* position 413 of any of the glucosaminyl residues 410 can be sulfated. As shown in Figure 22, the central glucosamine residue 410 receives a sulfo group, ultimately forming a 3-*O* sulfated glucosaminyl residue 510 within the sulfated product polysaccharide 441. Also as shown, sulfated product polysaccharide 441 comprises the structure of Formula I.

[0259] In another embodiment, engineered 3OSTs that can be utilized in accordance with methods of the present invention can have the same biological activity with heparosan-based sulfo acceptor polysaccharides as natural 3OSTs, particularly heparosan-based polysaccharides comprising the structure of Formula X. In another embodiment, when there are multiple portions of the polysaccharide comprising the structure of Formula X within the sulfo acceptor polysaccharide, any *N*-sulfated glucosamine residue can be 3-*O* sulfated by the engineered 3OST enzyme. Similarly, the same polysaccharide can be sulfated multiple times by the engineered 3OST, including and up to all of the *N*-sulfated glucosamine residues that are present within the polysaccharide. In another embodiment, a heparin mixture, either isolated from an animal source or synthesized according to any of the methods described herein, can also be utilized as a sulfo group acceptor and further 3-*O* sulfated upon reacting with an engineered 3OST enzyme and an aryl sulfate compound, to form an “over-sulfated” heparin mixture.

[0260] In another embodiment, sulfo acceptor polysaccharides that can react with an engineered or natural 3OST, including but not limited to those comprising the structure of Formula X, can be provided as a homogenous composition. In still other embodiments, sulfo acceptor polysaccharides that can react with an engineered or natural 3OST can be comprised within a composition comprising a polydisperse mixture of polysaccharides having variable chain lengths, molecular

weights, relative abundance of Formula X, and overall monosaccharide composition and functionalization.

[0261] In another embodiment, N,2-HS and N,2,6-HS polysaccharides, including but not limited to those comprising the structure of Formula X, and utilized in accordance with methods of the present invention with either an engineered or natural 6OST enzyme, can be obtained and/or modified from commercial sources. In another embodiment, either an engineered or natural 6OST can be utilized in accordance with methods of the present invention can react with N,2-HS products produced by an NST and/or a 2OST in one or more previous steps. In another embodiment, either an engineered or natural 6OST can be utilized in accordance with methods of the present invention can react with N,2,6-HS products produced by an NST, a 2OST, and/or a 6OST in one or more previous steps. In another embodiment, one or more of the sulfation steps to produce the N,2-HS or N,2,6-HS product was catalyzed by an engineered, aryl sulfate-dependent sulfotransferase. In another embodiment, all of the sulfation steps to produce the N,2-HS or N,2,6-HS product was catalyzed by an engineered, aryl sulfate-dependent sulfotransferase. Each of these processes are discussed in detail in the description and examples, below.

[0262] Natural 3OST enzymes within the EC 2.8.2.23 enzyme class generally comprise approximately 300 to 325 amino acid residues that can in some cases vary greatly in their sequence, yet ultimately have the exact same function, namely, to catalyze the transfer of a sulfonyl group from PAPS to the 3-*O* position of *N*-sulfo-glucosamine residues within N,2-HS or N,2,6-HS polysaccharides, particularly those comprising the structure of Formula X. Without being limited by a particular theory, it is believed that each of the natural 3OSTs can catalyze the same chemical reaction because there are multiple amino acid sequence motifs and secondary structures that are either identical or highly conserved across all species.

[0263] Further, it is believed that several of the conserved amino acid sequence motifs are directly involved in binding of either PAPS and/or the polysaccharide, or participate in the chemical reaction itself. The identity between the natural 3OST enzymes can be demonstrated by comparing the amino acid sequence of a particular enzyme with 3OST enzymes that have known crystal structures in which amino acid residues within the active site have been identified, including the mouse (PDB code: 3UAN) and human (PDB code: 1ZRH) 3OST1 enzymes, which have nearly identical active sites and overall structures even though they have only an 83% sequence identity with one another. A multiple sequence alignment of fifteen enzymes within EC 2.8.2.23, including the mouse and human enzymes, is shown in Figures 23A-23C, along with the percent identity of each sequence relative to a human 3OST reference sequence (UniProtKB Accession No. O14792). As illustrated in

Figures 23A-23C, sequences range from having 98% identity with the O14792 reference sequence (entry tr|H9ZG39|H9ZG39_MACMU) for the rhesus monkey 3OST, down to 53% identity (entry sp|Q8IZT8|HS3S5_HUMAN) for human 3OST5. Those skilled in the art would appreciate that the multiple sequence alignment was limited to fifteen sequences for clarity, and that there are hundreds of amino acid sequences encoding for natural 3OST enzymes that have been identified and that have highly conserved active site and/or binding regions as well.

[0264] Within Figures 23A-23C, amino acids that are depicted in white with a black background at a particular position, are 100% identical across all sequences. Amino acids that are highly conserved, meaning that the amino acids are either identical or chemically or structurally similar, at a particular position are enclosed with a black outline. Within highly conserved regions, consensus amino acids that are present in a majority of the sequences, are in bold. Amino acids at a particular position that are not identical or highly conserved are typically variable. A period within a sequence indicates a gap that has been inserted into the sequence in order to facilitate the sequence alignment with other sequence(s) that have additional residues between highly conserved or identical region. Finally, above each block of sequences are a series of arrows and coils that indicate secondary structure that is conserved across all sequences, based on the identity of the amino acids within the alignment and using the structure of the natural human sulfotransferase enzyme as a reference. The β symbol adjacent to an arrow refers to a β -sheet, whereas a coil adjacent to an α symbol or a η symbol refers to a helix secondary structure.

[0265] Within the fifteen aligned sequences in Figures 23A-23C, there are several conserved amino acid sequence motifs that include one or more amino acids that comprise the active site, based on the crystal structures of the mouse (entry sp|O35310|HS3S1_MOUSE) and human 3OST1 (entry sp|O14792|HS3S1_HUMAN) enzymes described above. Based on the numbering of the amino acid residues within Figures 23A-23C, these motifs include residues 16-27 (including G-V-R-K-G-G from residues 18-23), residues 43-48 (E-V/I-H-F-F-D), residues 78-81 (P-A/G-Y-F), residues 112-117 (including S-D-Y-T-Q-V), and residues 145-147 (Y-K-A). It is believed that these residues either facilitate or participate in the chemical reaction, or enable binding of PAPS or the polysaccharide within the active site. In particular, within residues 43-48, as described above and as illustrated in Figure 1, the glutamic acid residue at position 43 abstracts the proton from the 3-O position of the *N*-sulfoglucosamine residue within the polysaccharide, enabling the nucleophilic attack and removal of the sulfo group from PAPS, whereas His-45 and Asp-48 coordinate to stabilize the transition state of the enzyme before the sulfurylated polysaccharide product is released from the active site.

[0266] However, as described above, the natural 3OST enzymes are unable to catalyze the transfer of the sulfate group from an aryl sulfate compound to a polysaccharide. Without being limited by a particular theory, and as with the NSTs, 2OSTs, and the 6OSTs described above, it is believed that the binding pocket for PAPS within the active site of the natural sulfotransferase either does not have a high enough affinity for aryl sulfate compounds to facilitate binding and/or that the aryl sulfate compounds are sterically hindered from entering the active site. Consequently, and in another embodiment, a natural 3OST enzyme can be mutated in several locations within its amino acid sequence to enable binding of the aryl sulfate compound within the active site and/or to optimally position the aryl sulfate compound so transfer of the sulfate group to the polysaccharide can occur.

[0267] Accordingly, and in another embodiment, engineered 3OST enzymes that can be utilized with methods of the present invention can be mutants of natural 3OST enzymes within EC 2.8.2.23, including enzymes having the amino acid sequences illustrated in Figures 23A-23C. In another embodiment, the engineered 3OST enzymes have been engineered to recognize, bind, and react with aryl sulfate compounds as sulfo group donors, while retaining the natural enzymes' ability to recognize, bind, and react with any of the HS polysaccharides described above, including but not limited to those comprising the structure of Formula X, as sulfo group acceptors. Without being limited by a particular theory, it is believed that because of the mutations inserted into the amino acid sequences of the engineered 3OST enzymes, their sulfotransferase activity may comprise the direct transfer of a sulfuryl group from an aryl sulfate compound to the sulfo acceptor polysaccharide, using a similar mechanism as described in Figure 1, above, except that the PAPS is substituted with the aryl sulfate compound. Otherwise, it is believed that the mutations may cause the sulfotransferase activity to comprise a two-step process including the hydrolysis of an aryl sulfate compound and formation of a sulfohistidine intermediate, followed by the nucleophilic attack of the sulfohistidine intermediate by the oxygen atom at the 3-*O* position of a glucosamine residue, to form a 3-*O* sulfated product. In another embodiment, the 3-*O* sulfated HS product is an N,2,3,6-HS product.

[0268] In another embodiment, an engineered 3OST enzyme can comprise one or more mutated amino acid sequence motifs relative to the conserved amino acid sequence motifs found in natural 3OST enzymes within EC 2.8.2.23, as described above and indicated in the multiple sequence alignment in Figures 23A-23C. In another embodiment, each mutated amino acid sequence motif that is present in the amino acid sequence of the engineered enzyme comprises at least one amino acid mutation relative to the corresponding conserved amino acid sequence motif within the natural 3OST enzymes. In another embodiment, an engineered 3OST enzyme can comprise one mutated

amino acid sequence motif. In another embodiment, an engineered 3OST enzyme can comprise two mutated amino acid sequence motifs. In another embodiment, an engineered 3OST enzyme can comprise three mutated amino acid sequence motifs. In another embodiment, an engineered 3OST enzyme can comprise four mutated amino acid sequence motifs. In another embodiment, an engineered 3OST enzyme can comprise five mutated amino acid sequence motifs. In another embodiment, an engineered 3OST enzyme that includes at least one mutated amino acid sequence motif relative to any of the natural 3OST enzymes within EC 2.8.2.23 can have an amino acid sequence selected from the group consisting of SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58.

[0269] In another embodiment, upon viewing the crystal structure of the mouse 3OST1 enzyme within a 3D molecular visualization system (including, as a non-limiting example, the open-source software, PyMOL), the structure of related sequences, such as those of engineered 3OST enzymes that contain one or more mutated amino acid sequence motifs relative to the mouse 3OST1 (UniProtKB Accession No. O35310) structure, can be modeled for comparison as illustrated in Figure 24. Figure 24 shows a magnified view of the active site of the mouse 3OST enzyme (PDB code: 3UAN) with three engineered 3OST enzymes, comprising the amino acid sequences of SEQ ID NO: 24, SEQ ID NO: 26, and SEQ ID NO: 28. Adenosine 3',5'-diphosphate, which is the product of a sulfotransfer reaction in which PAPS is the sulfo donor, and which was co-crystallized with the mouse 3OST, is also illustrated within the active site. PNS is also modeled into the active site of the engineered enzymes, using the consensus solutions of molecular dynamics (MD) simulations that designed to calculate the optimized position and orientation of a ligand within an enzyme active site adjacent to the polysaccharide binding site (not shown), if such solutions are possible. Hydrogen atoms are not shown for clarity.

[0270] As illustrated in Figure 24, although there are several mutations made to SEQ ID NO: 24, SEQ ID NO: 26, and SEQ ID NO: 28, relative to the natural mouse sulfotransferase, the respective protein backbones are in a nearly identical location to one another, enabling a one-to-one comparison of the active sites. However, when comparing the two active sites, the adenosine 3',5'-diphosphate product from the natural sulfotransfer reaction is adjacent to the lysine residue, whereas the convergent solutions from the above MD simulations indicate that PNS binding within the engineered enzymes is favored on the opposite side of the active site. Without being limited by a particular theory, it is believed that the convergent MD simulation solutions place PNS on the opposite side of the active site because there is not enough of an affinity toward PNS in the same or

similar position as PAPS. Yet, despite the apparent differences in the binding pocket for PAPS and PNS, engineered 3OST enzymes comprising the amino acid sequences of SEQ ID NO: 24, SEQ ID NO: 26, and SEQ ID NO: 28 all achieved sulfo transfer from an aryl sulfate compound to the glucosaminyl 3-*O* position within an N,2,6-HS, as described in the examples below.

[0271] Further, the arginine residue corresponding to position 20 of the mouse 3OST1, and conserved in all of the other 3OST enzymes illustrated in Figures 23A-23C, if present in an engineered 3OST enzyme, would block PNS from binding in the position indicated in Figure 24. Accordingly, and in another embodiment, engineered 3OST enzymes that bind PNS can comprise a mutation of the active site arginine residue to a glycine residue, which removes all steric hindrance for PNS to bind within the binding pocket. As indicated in the amino acid sequences for SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 53, and SEQ ID NO: 54, the arginine to glycine mutation is at position 21. As indicated in the amino acid sequences for SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58, the arginine to glycine mutation is at position 99.

[0272] Similarly, the next amino acid residue in each of the engineered enzymes, corresponding to position 22 in the amino acid sequences SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, is mutated to a histidine residue. Without being limited by a particular theory, it is believed that the mutation to a histidine residue from the conserved lysine residue (corresponding to position 21 in each of the amino acid sequences in Figures 23A-23C) facilitates removal of the sulfate group from PNS, using a similar mechanism described by Malojcic, et al., above. As indicated in the amino acid sequences for SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58, the lysine to histidine residue is at position 100.

[0273] Those skilled in the art would appreciate that engineered 3OST enzymes of any other amino acid sequence, including, but not limited to, those disclosed by SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58, would exhibit a similar structure would exhibit similar structural motifs as engineered enzymes having the amino acid sequences of SEQ ID NO: 24, SEQ ID NO: 26, and SEQ ID NO: 28, particularly within the active site. Without being limited by a particular theory, it is also believed that NCS would bind in a similar position as PNS within the active site of any of the engineered enzymes, since the structures of the two aryl sulfate compounds are very similar, except that the sulfate group is located *ortho* on the aromatic ring relative to the nitro group, rather than *para* to the nitro group.

[0274] In another embodiment, engineered 3OST enzymes that can be utilized in accordance with methods of the present invention can comprise one or more mutated amino acid sequence motifs,

which can be determined in-part by comparing conserved amino acid sequence motifs indicated in the multiple sequence alignment of Figures 23A-23C with the known structure(s) of natural enzymes and/or modeled engineered enzymes, including but not limited to, as a non-limiting example, enzymes illustrated in Figure 24. In another embodiment, mutated amino acid sequence motifs that can be comprised within an engineered 3OST enzyme can be selected from the group consisting of (a) G-V-G-H-G-G; (b) H-S-Y-F; (c) S-X₁-X₂-T-H-X₃, wherein X₁ is selected from the group consisting of alanine and leucine; X₂ is selected from the group consisting of tyrosine and glycine, and X₃ is selected from the group consisting of methionine and leucine; and (d) Y-X₄-G, wherein X₄ is selected from the group consisting of valine and threonine; including any combination thereof. Each of the mutated amino acid sequence motifs corresponds with a conserved amino acid motif indicated in Figures 23A-23C above: the mutated amino acid sequence motif G-V-G-H-G-G corresponds to the conserved amino acid sequence motif G-V-R-K-G-G; the mutated amino acid sequence motif H-S-Y-F corresponds to the conserved amino acid sequence motif P-A/G-Y-F; the mutated amino acid sequence motif S-X₁-X₂-T-H-X₃ corresponds to the conserved amino acid sequence motif S-D-Y-T-Q-V; and the mutated amino acid sequence motif Y-X₄-G corresponds to the conserved amino acid sequence motif Y-K-A. In another embodiment, an engineered 3OST enzyme comprising each of the mutated amino acid sequence motifs above can be selected from the group consisting of: SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58.

[0275] In another embodiment, each of the mutated amino acid sequence motifs can comprise at least one mutation that is made relative to the conserved amino acids found in the natural 3OST enzymes within EC 2.8.2.23. In another embodiment, mutated amino acid sequence motif (a) contains an R-K to G-H mutation, relative to the conserved amino acid sequence motif, G-V-R-K-G-G. In another embodiment, mutated amino acid sequence motif (b) contains a P-A/G to an H-S mutation relative to the conserved amino acid sequence motif, P-A/G-Y-F. In another embodiment, in addition to potential mutations made at the X₁, X₂, and X₃ positions, mutated amino acid sequence motif (c) comprises a Q to H mutation, relative to the conserved amino acid sequence motif, S-D-Y-T-Q-V. In another embodiment, in addition to a mutation at the X₄ position, mutated amino acid sequence motif (d) comprises an A to G mutation, relative to the conserved amino acid sequence motif, Y-K-A.

[0276] In another embodiment, X₁ is alanine, X₂ is tyrosine; X₃ is methionine, and X₄ is valine or threonine. In other embodiments, X₁ is leucine, X₂ is glycine, X₃ is leucine, and X₄ is threonine.

Without being limited to another theory, it is believed that one or more of the mutations comprised within mutated amino acid sequence motifs (b), (c), and (d) play a role in stabilizing the transition state of the enzyme during the chemical reaction, or in increasing the affinity of aryl sulfate compounds to the active site, including by reducing the size of the binding pocket, increasing the hydrophobicity of the pocket, and/or creating π - π interactions with the aromatic moieties of the aryl sulfate compounds.

[0277] Furthermore, the amino acid sequences (SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28) of three engineered 3OST enzymes, which have been experimentally determined to be active with aryl sulfate compounds as sulfo group donors (see Example 5 below) can be compared with the amino acid sequence of the first isoform of the human 3OST enzyme (entry sp|O14792|HS3S1_HUMAN) in a multiple sequence alignment to determine if there are relationships between mutations among each of the enzymes. A period within the amino acid sequence of an engineered enzyme indicates identity at a particular position with the human 3OST enzyme. As shown in Figure 25, the sequence alignment demonstrates that while over 90% of the amino acid residues within the three sulfotransferase sequences are identical, there are several positions in which multiple amino acids can be chosen. Without being limited by a particular theory, these enzymes have a similar relationship with each other as the 3OST enzymes that comprise EC 2.8.2.23. As a result, and in another embodiment, an engineered 3OST enzyme comprising an amino acid sequence in which multiple amino acids can be chosen at defined positions is disclosed as SEQ ID NO: 51. Positions at which the identity of an amino acid can be chosen from a selection of possible residues are denoted in terms “Xaa,” “Xn,” or “position n,” where n refers to the residue position.

[0278] In another embodiment, within an engineered 3OST enzyme comprising the amino acid sequence of SEQ ID NO: 51, the amino acid residue at position 114 is alanine and the amino acid residue at position 118 is methionine. In further embodiments, the amino acid residue at position 147 is selected from the group consisting of valine and threonine.

[0279] In another embodiment, within an engineered 3OST enzyme comprising the amino acid sequence of SEQ ID NO: 51, the amino acid residue at position 114 is leucine, the amino acid residue at position 118 is leucine, and the amino acid residue at position 121 is valine. In further embodiments, the amino acid residue at position 115 is glycine. In even further embodiments, the amino acid residue at position 147 is threonine.

[0280] In another embodiment, within an engineered 3OST enzyme comprising the amino acid sequence of SEQ ID NO: 51, the amino acid sequence can optionally include one or more mutations

at residue positions not specified by an “Xn” or “Xaa,” so long as any such mutations do not eliminate the 3OST and/or aryl sulfate-dependent activity of the enzyme. In another embodiment, such mutations not eliminating aryl sulfate-dependent activity at positions not specified by an “Xn” or “Xaa” can include substitutions, deletions, and/or additions.

[0281] Accordingly, in another embodiment, an engineered 3OST enzyme utilized in accordance with any of the methods of the present invention can comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58. In another embodiment, any of the above engineered 6OST enzymes react with an aryl sulfate compound, instead of PAPS, as a sulfo group donor. In further embodiments, the aryl sulfate compound is selected from the group consisting of PNS, MUS, 7-hydroxycoumarin sulfate, phenyl sulfate, 4-acetylphenyl sulfate, indoxyl sulfate, 1-naphthyl sulfate, 2-naphthyl sulfate, and NCS. In some even further embodiments, the aryl sulfate compound is PNS. In other even further embodiments, the aryl sulfate compound is NCS.

In vitro synthesis of sulfated polysaccharides

[0282] As described above, natural sulfotransferases that recognize, bind, and react with heparosan-based polysaccharides as sulfo group acceptors have the ability to produce a wide range of sulfated polysaccharide products *in vivo*, including heparin (*see* Desai, U.R., et al., (1998) *J. Biol. Chem.* **273** (13):7478-7487). The medical use of heparin has been well documented for decades including, but are not limited to, inactivation of Factor IIa (thrombin) and/or Factor Xa, two proteins that are vital in the blood-clotting cascade. In particular, when heparin binds to antithrombin (AT), it causes a conformational change in the enzyme that enables the formation of a ternary complex between the polysaccharide, AT, and either thrombin or Factor Xa (*see* Li, W., et al., (2004) *Nat. Struct. Mol. Biol.* **11** (9):857-862, the disclosure of which is incorporated by reference in its entirety). In order to bind with AT and induce its conformational change, polysaccharides within the heparin composition must have a specific five-residue AT-recognition sequence, which is identical to the structure of Formula I, described above. As a result, ODSH compositions that have even a small amount of anticoagulant activity often contain some polysaccharides that have an AT-recognition sequence. In some embodiments, HS compositions can be synthesized by one or more methods of the present invention, in which none of the polysaccharides contain an AT-recognition sequence. In further embodiments, an 3OST can be omitted from the synthesis of an HS polysaccharide to ensure that no AT-recognition sequences will be present in the product.

[0283] As described above, the hallmark of nearly all sulfotransferases, whether they are utilized in either *in vitro* or an *in vivo* sulfotransfer reaction, is that they universally and exclusively recognize PAPS as the sulfo group donor, as described in U.S. Pat. Nos. 5,541,095, 5,817,487, 5,834,282, 6,861,254, 8,771,995, 9,951,149, and U.S. Pat. Pubs. 2009/0035787, 2013/0296540, and 2016/0122446, the disclosures of which are incorporated by reference in their entireties. These include sulfotransferases in which a polysaccharide is a sulfo group acceptor, particularly HS sulfotransferases that take part in the production of anticoagulant and non-anticoagulant N,2,3,6-HS products. Currently, because PAPS is expensive and unstable in solution, the most convenient and economically feasible method to obtain anticoagulant N,2,3,6-HS polysaccharides in large quantities is to isolate them from animal sources, particularly pigs and cattle, rather than to synthesize them *in vitro*, even when a coupled, enzymatic PAPS regeneration system (see U.S. Pat. No. 6,255,088, above) is employed. Without being limited by a particular theory, utilizing any of the engineered aryl sulfate-dependent sulfotransferases described above to catalyze one or more of the sulfotransfer reactions in the production of N,2,3,6-HS polysaccharides can reduce the industry's reliance on using PAPS as a sulfo group donor, and if an engineered aryl sulfate-dependent sulfotransferase is utilized in all of the enzymatic sulfotransfer steps, the need to use PAPS can be obviated entirely.

[0284] Accordingly, methods for synthesizing an HS compound can comprise any combination of natural or engineered sulfotransferase enzymes, so long as at least one of the reactions comprises an engineered aryl sulfate-dependent sulfotransferase enzyme and an aryl sulfate compound. In some embodiments, NS6S/HS can be synthesized using a method comprising the following steps: (a) providing a starting polysaccharide composition comprising *N*-deacetylated heparosan; (b) reacting the starting polysaccharide composition within a reaction mixture comprising an *N*-sulfation agent, to form an *N*-sulfated heparan sulfate (NS/HS) product; and (c) reacting the NS/HS product within a reaction mixture comprising an aryl sulfate compound and an engineered 6OST enzyme, thereby forming the NS6S/HS product; wherein the biological activity of the engineered 6OST enzyme comprises the transfer of a sulfo group from an aryl sulfate compound to a heparosan-based polysaccharide. In various embodiments, the biological activity of the engineered 6OST enzyme consists of the transfer of a sulfo group from an aryl sulfate compound to a heparosan-based polysaccharide, preferably in the absence of PAPS. In further embodiments, the method can further comprise the step of reacting either NS/HS or NS6S/HS with a 2OST and a sulfo group donor, to form either an NS2S/HS or NS2S6S/HS product, respectively. In some preferred embodiments, the 2OST is an engineered sulfotransferase, and the sulfo group donor consists of an aryl sulfate

compound. In another embodiment, the reaction mixture that comprises the 2OST enzyme further comprises a glucuronyl C5-epimerase enzyme.

[0285] In other embodiments, an ODSH compound can be formed from an NS2S6S3S-HS product synthesized by any combination of natural or engineered sulfotransferase enzymes, so long as at least one of the reactions comprises an engineered aryl sulfate-dependent sulfotransferase enzyme and an aryl sulfate compound. In some embodiments, methods for synthesizing an NS2S6S3S-HS product can comprise the following steps: (a) providing a starting polysaccharide reaction mixture comprising *N*-deacetylated heparosan; (b) reacting the starting polysaccharide composition within a reaction mixture comprising an *N*-sulfation agent, to form an NS/HS product; (c) combining the NS/HS product with a reaction mixture comprising a sulfo group donor and a first intermediate sulfotransferase enzyme selected from the group consisting of a 2OST enzyme and a 6OST enzyme, to form a first intermediate HS product; (d) combining the first intermediate HS product with a reaction mixture comprising a second intermediate sulfotransferase enzyme, wherein the second intermediate sulfotransferase enzyme is the enzyme that was not selected in step (c), to form a second intermediate HS product; and (e) combining the second intermediate HS product with a reaction mixture comprising a sulfo group donor and a 3OST enzyme, to form the NS2S6S3S-HS product. Once the NS2S6S3S-HS product is formed, an ODSH can be formed according to any of the desulfation methods described in above in U.S. Patents 5,990,097, 5,912,237, 5,808,021, 5,668,118, and 5,296,471, and in further detail below. Reaction mixtures that do not comprise an engineered sulfotransferase enzyme can comprise PAPS and a natural HS sulfotransferase enzyme that possesses biological activity with PAPS as the sulfo group donor. In various embodiments, all of the enzymes are engineered aryl sulfate-dependent sulfotransferases.

[0286] In another embodiment, the *N*-sulfation agent can comprise any of the natural NDST or engineered NST enzymes described above. In another embodiment, when the *N*-sulfation agent is a natural NDST enzyme, the reaction mixture can also comprise PAPS as a sulfo group donor. In another embodiment, when the *N*-sulfation agent is an engineered NST, the reaction mixture can also comprise an aryl sulfate compound, preferably PNS or NCS, as a sulfo group donor. In another embodiment, when the *N*-sulfation agent is an engineered NST, the enzyme can comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40.

[0287] In another embodiment, *N*-deacetylated heparosan can be chemically *N*-sulfated, rather than being enzymatically *N*-sulfated. In another embodiment, the *N*-sulfation agent is a chemical agent,

preferably sulfur trioxide and/or one or more sulfur-trioxide containing compounds or adducts. Chemical *N*-sulfation of glucosamine residues within polysaccharides using sulfur trioxide is commonly known in the art (*see* Lloyd, A.G., et al., (1971) *Biochem. Pharmacol.* **20** (3):637-648; Nadkarni, V.D., et al., (1996) *Carbohydrate Research* **290**:87-96; Kuberan, B., et al., (2003) *J. Biol. Chem.* **278** (52):52613-52621; Zhang, Z., et al., (2008) *J. Am. Chem. Soc.* **130** (39):12998-13007; and Wang, et al., (2011), above; *see also* U.S. Pat. No. 6,991,183 and U.S. Pat. Pub. 2008/020789, the disclosures of which are incorporated by reference in their entireties). Sulfur trioxide complexes are generally mild enough bases to enable the selected *N*-sulfation of polysaccharides without causing depolymerization, unlike sodium hydroxide (*see* Gilbert, E.E., (1962) *Chem. Rev.* **62** (6):549-589). Non-limiting examples of sulfur trioxide-containing complexes include sulfur dioxide-pyridine, sulfur dioxide-dioxane, sulfur dioxide-trimethylamine, sulfur dioxide-triethylamine, sulfur dioxide-dimethylaniline, sulfur dioxide-thioxane, sulfur dioxide-Bis(2-chloroethyl) ether, sulfur dioxide-2-methylpyridine, sulfur dioxide-quinoline, or sulfur dioxide-dimethylformamide. In another embodiment, the *N*-sulfation agent comprises a sulfur trioxide-containing adduct selected from the group consisting of a sulfur trioxide-trimethylamine adduct and a sulfur trioxide-pyridine adduct. In another embodiment, the *N*-sulfation agent comprises a sulfur trioxide-trimethylamine adduct.

[0288] In another embodiment, *N*-sulfation, particularly chemical *N*-sulfation, can comprise the first sulfation step, with respect to *N*-deacetylated heparosan. Subsequently, after the *N*-deacetylated heparosan is either enzymatically or chemically *N*-sulfated, the *N*-sulfated heparosan can then be further sulfated using a 2OST, 6OST, and/or 3OST. In embodiments in which an NS2S6S3S-HS product is formed, enzymatic sulfation steps can occur in the order of 2-*O*, 6-*O*, and 3-*O* sulfation. As described above, the 3OST enzyme, and preferably all of the sulfotransferase enzymes, are engineered aryl-sulfate dependent sulfotransferase enzymes, and the reactions are performed in the absence of PAPS.

[0289] In another embodiment, reaction mixtures comprising a 2OST enzyme further comprise a glucuronyl C5-epimerase enzyme, preferably a glucuronyl C5-epimerase enzyme comprising the amino acid sequence of SEQ ID NO: 29, and more preferably a glucuronyl C5-epimerase enzyme comprising the amino acid sequence of residues 34-617 of SEQ ID NO: 29. In another embodiment, when the 2OST enzyme is an engineered enzyme, the enzyme can comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 41, and SEQ ID NO: 42.

[0290] In another embodiment, when the 6OST enzyme is an engineered enzyme, the enzyme can comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 59, SEQ ID NO: 60, and SEQ ID NO: 61.

[0291] In another embodiment, when the 3OST enzyme is an engineered enzyme, the enzyme can comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58.

[0292] In another embodiment, aryl sulfate compounds used as sulfo group donors can be selected from the group consisting of PNS, MUS, 7-hydroxycoumarin sulfate, phenyl sulfate, 4-acetylphenyl sulfate, indoxyl sulfate, 1-naphthyl sulfate, 2-naphthyl sulfate, and NCS. In even further embodiments, the aryl sulfate compound is PNS. In other even further embodiments, the aryl sulfate compound is NCS.

[0293] In another embodiment, in some instances in which the methods of the present invention are used to synthesize an NS2S6S3S-HS product, the NS2S6S3S-HS product comprises an AT-recognition sequence and has anticoagulant activity, which can be characterized by the degree of inhibitory activity that they have against Factor Xa and thrombin, termed “anti-Xa” activity and “anti-IIa” activity, respectively. The amount of inhibition induced by anticoagulant polysaccharides is often measured in International Units per milligram (IU mg⁻¹) and less often as International Units per milliliter (IU mL⁻¹). In either case, an International Unit is an amount approximately equivalent to the quantity required to keep 1-mL of cat’s blood fluid for 24 hours at 0 °C. In another embodiment, anticoagulant NS2S6S3S-HS polysaccharides produced by methods of the present invention can have an anti-Xa activity of at least about 1 IU mg⁻¹, including at least about 50 IU mg⁻¹, at least 75 IU mg⁻¹, 100 IU mg⁻¹, 150 IU mg⁻¹, 200 IU mg⁻¹, or 500 IU mg⁻¹, up to at least about 1,000 IU mg⁻¹. In another embodiment, anticoagulant NS2S6S3S-HS polysaccharides produced by methods of the present invention can have an anti-IIa activity of at least about 1 IU mg⁻¹, including at least about 10 IU mg⁻¹, 25 IU mg⁻¹, 50 IU mg⁻¹, 100 IU mg⁻¹, 150 IU mg⁻¹, or 180 IU mg⁻¹, up to at least about 200 IU mg⁻¹. In another embodiment, the ratio of anti-Xa activity to anti-IIa activity of the NS2S6S3S-HS product is at least 0.5:1, including at least 0.75:1, 0.9:1, 1:1, 1.1:1, 1.3:1, 1.5:1, 2.0:1, 3.0:1, 4.0:1, 5.0:1, 6.0:1, 7.0:1, 8.0:1, 9.0:1, 10.0:1, 20:1, 40:1, 60:1, or 80:1, up to at least 100:1. In another embodiment, the ratio of anti-Xa activity to anti-IIa activity of the NS2S6S3S-HS product is less than 100:1, including less than 80:1, 60:1, 40:1, 20:1, 10.0:1, 9.0:1, 8.0:1, 7.0:1, 6.0:1,

5.0:1, 4.0:1, 3.0:1, 2.0:1, 1.5:1, 1.3:1, 1.1:1, 0.9:1, or 0.75:1, down to less than 0.5:1. In another embodiment, the ratio of anti-Xa activity to anti-IIa activity of the NS2S6S3S-HS product is in a range from about 0.9 to about 1.1. In another embodiment, the ratio of anti-Xa activity to anti-IIa activity of an NS2S6S3S-HS product is in a range from 0.5:1 up to 0.75:1, or 0.9:1, or 1:1, or 1.1:1, or 1.3:1, or 1.5:1, or 2.0:1, or 3.0:1, or 4.0:1, or 5.0:1, or 6.0:1, or 7.0:1, or 8.0:1, or 9.0:1, or 10.0:1.

[0294] Similarly, all polysaccharide mixtures, including NS2S/HS, NS6S/HS, NS2S6S/HS, and NS2S6S3S-HS product mixtures, can be characterized by their weight-average molecular weight ($\bar{M}_w$). Because substantially all of the heparins either isolated from animal sources or synthesized *in vitro* are obtained as a polydisperse mixture of polysaccharides with different chain lengths and degrees of sulfation, expressing the average molecular weight as a weight average, rather than a number average (i.e. a true arithmetic mean ($\bar{M}_n$), is often the most advantageous because it accounts for the effect larger molecules have on anticoagulation. The $\bar{M}_w$ of a polysaccharide mixture can be measured experimentally using light scattering methods or analytical ultracentrifugation (see Mulloy, B., et al., (2014) *Anal. Bioanal. Chem.* **406**:4815-4823, the disclosure of which is incorporated by reference in its entirety). However, determining the $\bar{M}_n$, typically by size exclusion chromatography, can still be useful because the ratio between $\bar{M}_w$ and $\bar{M}_n$ can provide valuable insight into the amount of polydispersity in a particular polysaccharide sample.

[0295] For example, heparins are generally divided into multiple classes based on their average molecular weights, particularly their $\bar{M}_w$. Samples of low-molecular weight heparin (LMWH) typically have an $\bar{M}_w$ of less than 8,000 Da, in which more than 60% of all of the polysaccharide molecules within the sample have an actual molecular weight of less than 8,000 Da (see Linhardt, R.J. and Gunay, N.S., (1999) *Seminars in Thrombosis and Hemostasis* **25** (Suppl. 3):5-16, the disclosure of which is incorporated by reference in its entirety). LMWH is typically prepared by chemically or enzymatically modifying animal-sourced unfractionated heparin or API heparin. Unfractionated heparin typically has an $\bar{M}_w$ of greater than 8,000 Da. To be approved for use in medical treatments, API heparin has strict molecular weight guidelines that must be met, namely: (1) the proportion of polysaccharides within the composition having a molecular weight over 24,000 Da is not more than 20%; (2) the $\bar{M}_w$ of the composition itself is between 15,000 Da and 19,000 Da; and (3) the ratio of the number of polysaccharides within the composition having a molecular weight between 8,000 Da and 16,000 Da relative to the number of polysaccharides within the composition having a molecular weight between 16,000 Da and 24,000 Da is not less than 1.0:1 (see Mulloy, B., et al., above).

[0296] Thus, in another embodiment, ODSH can be prepared from an NS2S6S3S-HS product synthesized according to methods of the present invention, in which the NS2S6S3S-HS product has an $\bar{M}_w$ of at least 1,000 Da, including at least 2,000 Da, 3,000 Da, 4,000 Da, 5,000 Da, 6,000 Da, 7,000 Da, 8,000 Da, 9,000 Da, 10,000 Da, 11,000 Da, 12,000 Da, 13,000 Da, 14,000 Da, 15,000 Da, 16,000 Da, 17,000 Da, 18,000 Da, 19,000 Da, 20,000 Da, 21,000 Da, 22,000 Da, 23,000 Da, or 24,000 Da, up to at least 50,000 Da. In another embodiment, the NS2S6S3S-HS product has an $\bar{M}_w$ of less than 50,000 Da, including less than 24,000 Da, 23,000 Da, 22,000 Da, 21,000 Da, 20,000 Da, 19,000 Da, 18,000 Da, 17,000 Da, 16,000 Da, 15,000 Da, 14,000 Da, 13,000 Da, 12,000 Da, 11,000 Da, 10,000 Da, 9,000 Da, 8,000 Da, 7,000 Da, 6,000 Da, 5,000 Da, 4,000 Da, or 3,000 Da, down to less than 2,000 Da. In another embodiment, the NS2S6S3S-HS product has an $\bar{M}_w$ in a range from 1,000 up to 2,000 Da, or 3,000 Da, or 4,000 Da, or 5,000 Da, or 6,000 Da, or 7,000 Da, or 8,000 Da, or 9,000 Da, or 10,000 Da, or 11,000 Da, or 12,000 Da, or 13,000 Da, or 14,000 Da, or 15,000 Da, or 16,000 Da, or 17,000 Da, or 18,000 Da, or 19,000 Da, or 20,000 Da, or 21,000 Da, 22,000 Da, or 23,000 Da, or 24,000 Da. In another embodiment, the unfractionated anticoagulant NS2S6S3S-HS product has an $\bar{M}_w$ in a range from 8,000 Da up to 9,000 Da, or 10,000 Da, or 11,000 Da, or 12,000 Da, or 13,000 Da, or 14,000 Da, or 15,000 Da, or 16,000 Da, or 17,000 Da, or 18,000 Da, or 19,000 Da, or 20,000 Da, or 21,000 Da, 22,000 Da, or 23,000 Da, or 24,000 Da. In another embodiment, the anticoagulant NS2S6S3S-HS product can have an $\bar{M}_w$ in any range listed above between and inclusive of 1,000 Da and 24,000 Da, and preferably in any range listed above between and inclusive of 15,000 Da and about 19,000 Da.

[0297] In another embodiment, N,2,3,6-HS products prepared by any of the methods of the present invention can satisfy any of the benchmark requirements determined by the USP for API heparin, including but not limited to composition, purity, activity, and/or molecular weight. In another embodiment, the anticoagulant N,2,3,6-HS product can possess any of the properties selected from the group consisting of: an anti-IIa activity of not less than 180 IU mg⁻¹; an anti-Xa activity of not less than 180 IU mg⁻¹; a ratio of anti-Xa to anti-IIa activity in a range of 0.9:1 up to 1.1:1, preferably 1:1; an $\bar{M}_w$ of in a range of 15,000 Da up to 19,000 Da; not more than 20% of the polysaccharides having a molecular weight greater than 24,000 Da; and the ratio of polysaccharides within the composition having a molecular weight between 8,000 Da and 16,000 Da relative to the number of polysaccharides within the composition having a molecular weight between 16,000 Da and 24,000 Da is not less than 1.0:1, including any combination thereof. In another embodiment, anticoagulant N,2,3,6-HS products prepared by any of the methods of the present invention can possess all of the following anticoagulant activity and molecular weight properties: an anti-IIa

activity of not less than 180 IU mg⁻¹; an anti-Xa activity of not less than 180 IU mg⁻¹; a ratio of anti-Xa to anti-IIa activity in a range of 0.9:1 up to 1.1:1, preferably 1:1; an $\bar{M}_w$ of in a range of 15,000 Da up to 19,000 Da; not more than 20% of the polysaccharides having a molecular weight greater than 24,000 Da; and the ratio of polysaccharides within the composition having a molecular weight between 8,000 Da and 16,000 Da relative to the number of polysaccharides within the composition having a molecular weight between 16,000 Da and 24,000 Da is not less than 1.0:1. In another embodiment, anticoagulant N,2,3,6-HS products prepared by any of the methods of the present invention have a substantially equivalent anticoagulant activity and molecular weight properties relative to API heparin (CAS No: 9041-08-1), which is widely commercially-available.

[0298] In another embodiment, anticoagulant N,2,3,6-HS products can satisfy benchmark requirements determined by the USP for API heparin with regard to product purity, particularly purity from other sulfated polysaccharides, including but not limited to chondroitin sulfate. In particular, over-sulfated chondroitin sulfate (OSCS) was determined to be the source of contamination within pharmaceutical heparin compositions that caused hundreds of deaths worldwide in 2007 and 2008. In another embodiment, and without being limited by a particular theory, preparations of the N,2,3,6-HS product formed by any of the methods of the present invention can be prepared substantially or completely free from chondroitin sulfate, particularly OSCS, because it is believed that the *N*-deacetylated heparosan starting material, which can either be obtained commercially or after modifying heparosan isolated from bacteria (described in further detail below), itself is free of chondroitin sulfate. In another embodiment, any of the NS/HS, NS6S/HS, NS2S/HS, and NS2S6S/HS products formed by any of the methods of the present invention can be prepared substantially or completely free from chondroitin sulfate.

[0299] In another embodiment, ODSH can be prepared from a low molecular weight HS (LMW-HS) product, which itself is synthesized from an NS2S6S3S-HS product and described in further detail below. In another embodiment, the LMW-HS product has an $\bar{M}_w$ in a range from 2,000 Da up to 3,000 Da, or 4,000 Da, or 5,000 Da, or 6,000 Da, or 7,000 Da, or 8,000 Da. In another embodiment, the LMW-HS product has an $\bar{M}_w$ in any range listed above between and inclusive of 2,000 Da and about 8,000 Da.

[0300] Similarly, and in another embodiment, NS/HS, NS6S/HS, NS2S/HS, and NS2S6S/HS products can be synthesized to have an $\bar{M}_w$ that is substantially equivalent to the $\bar{M}_w$ of any of the NS2S6S3S-HS products above, including at least 1,000 Da, including at least 2,000 Da, 3,000 Da, 4,000 Da, 5,000 Da, 6,000 Da, 7,000 Da, 8,000 Da, 9,000 Da, 10,000 Da, 11,000 Da, 12,000 Da, 13,000 Da, 14,000 Da, 15,000 Da, 16,000 Da, 17,000 Da, 18,000 Da, 19,000 Da, 20,000 Da, 21,000

Da, 22,000 Da, 23,000 Da, or 24,000 Da, up to at least 50,000 Da. In another embodiment, an NS/HS, NS6S/HS, NS2S/HS, or NS2S6S/HS product can have an $\bar{M}_w$ of less than 50,000 Da, including less than 24,000 Da, 23,000 Da, 22,000 Da, 21,000 Da, 20,000 Da, 19,000 Da, 18,000 Da, 17,000 Da, 16,000 Da, 15,000 Da, 14,000 Da, 13,000 Da, 12,000 Da, 11,000 Da, 10,000 Da, 9,000 Da, 8,000 Da, 7,000 Da, 6,000 Da, 5,000 Da, 4,000 Da, or 3,000 Da, down to less than 2,000 Da. In another embodiment, an NS/HS, NS6S/HS, NS2S/HS, or NS2S6S/HS product can have an $\bar{M}_w$ in a range from 1,000 up to 2,000 Da, or 3,000 Da, or 4,000 Da, or 5,000 Da, or 6,000 Da, or 7,000 Da, or 8,000 Da, or 9,000 Da, or 10,000 Da, or 11,000 Da, or 12,000 Da, or 13,000 Da, or 14,000 Da, or 15,000 Da, or 16,000 Da, or 17,000 Da, or 18,000 Da, or 19,000 Da, or 20,000 Da, or 21,000 Da, 22,000 Da, or 23,000 Da, or 24,000 Da. In another embodiment, an NS/HS, NS6S/HS, NS2S/HS, or NS2S6S/HS product can have an $\bar{M}_w$ in a range from 8,000 Da up to 9,000 Da, or 10,000 Da, or 11,000 Da, or 12,000 Da, or 13,000 Da, or 14,000 Da, or 15,000 Da, or 16,000 Da, or 17,000 Da, or 18,000 Da, or 19,000 Da, or 20,000 Da, or 21,000 Da, 22,000 Da, or 23,000 Da, or 24,000 Da. In another embodiment, an NS/HS, NS6S/HS, NS2S/HS, or NS2S6S/HS product can have an $\bar{M}_w$ in any range listed above between and inclusive of 1,000 Da and 24,000 Da.

[0301] In another embodiment, an NS/HS, NS6S/HS, NS2S/HS, or NS2S6S/HS product can be prepared to have an $\bar{M}_w$ that is larger than 2-*O* and 3-*O*-desulfated ODSH compositions prepared from heparin or LMWH. Generally, 2-*O* and 3-*O*-desulfated ODSH compositions prepared from USP-compliant heparin have a reduced $\bar{M}_w$, typically in a range from about 8,000 Da to about 15,000 Da. Without being limited by a particular theory, it is believed that such ODSH compositions have a comparatively reduced $\bar{M}_w$ because the strongly basic conditions required for desulfation can also have the effect of depolymerizing some of the polysaccharides within the heparin composition. In contrast, NS/HS, NS6S/HS, NS2S/HS, or NS2S6S/HS products made by methods of the present invention can have an $\bar{M}_w$ that is greater than 15,000 Da, and particularly in a range between about 15,000 Da and about 19,000 Da.

[0302] In another embodiment, in order to arrive at an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product having a desired $\bar{M}_w$, the molecular weight of any of the polysaccharides utilized as sulfo group acceptors can be controlled. In a non-limiting example, and in another embodiment, the molecular weight properties of the heparosan-based polysaccharides used as starting materials can be controlled by chemically modifying heparosan until a target set of molecular weight properties is reached. As described below, heparosan and other heparosan-based polysaccharides can be obtained from commercial sources or isolated from bacterial or eukaryotic sources.

[0303] In particular, heparosan and heparosan-based polysaccharides can also be found within bacteria as a capsule that regulates cell entry by metabolites and other exogenous materials. Such bacteria, include, but are not limited to *Pasteurella multocida* and *Escherichia coli* (*E. coli*). In some embodiments, heparosan can be extracted and purified from *E. coli*, particularly K5 strain of *E. coli*, as a polydisperse mixture of polysaccharide molecules having varying molecular weights. Procedures for isolating heparosan from the K5 strain of *E. coli* are discussed and provided in Wang, Z., et al., (2010) *Biotechnol. Bioeng.* **107** (6):964-973, the disclosure of which is incorporated by reference in its entirety; see also DeAngelis, P.L. (2015) *Expert Opinion on Drug Delivery* **12** (3):349-352; Ly, M., et al., (2010) *Anal. Bioanal. Chem.* **399**:737-745; and Zhang, C., et al., (2012) *Metabolic Engineering* **14**:521-527, the disclosures of which are also incorporated in their entireties. However, because substantially all of the heparosan isolated from bacteria, including *E. coli*, is *N*-acetylated, it cannot be used directly as a sulfo acceptor for any of the sulfotransferases described herein and utilized in accordance with the methods of the present invention. As a result, heparosan must be at least partially *N*-deacetylated before it can be utilized as a sulfo group acceptor.

[0304] As a result, and in another embodiment, heparosan can be at least partially *N*-deacetylated by treating it with a base, particularly lithium hydroxide or sodium hydroxide (see Wang, Z., et al., (2011) *Appl. Microbiol. Biotechnol.* **91** (1):91-99, the disclosure of which is incorporated by reference in its entirety; see also PCT publication PCT/US2012/026081, the disclosure of which is incorporated by reference in its entirety). In another embodiment, the base is sodium hydroxide. Depending on the degree of *N*-deacetylation desired, the concentration of the heparosan, and the concentration of the base, one skilled in the art can determine how long to incubate heparosan with the base according to the procedures described in Wang, et al., (2011), above.

[0305] In another embodiment, heparosan can be incubated with a base, preferably sodium hydroxide, until a desired amount of *N*-acetylated glucosamine residues remains within the *N*-deacetylated product. In another embodiment, *N*-acetyl glucosamine residues can comprise less than 60%, including less than 30%, 20%, 18%, 16%, 14%, 12%, or 10%, down to less than 5%, and preferably in a range from 12% and up to 18%, of the glucosamine residues within the *N*-deacetylated heparosan product. In another embodiment, the *N*-acetyl glucosamine can comprise about 15% of the glucosamine residues within the *N*-deacetylated heparosan product.

[0306] Additionally, and without being limited by a particular theory, it is believed that in addition to *N*-deacetylating glucosamine residues, the reaction between heparosan and a base can simultaneously depolymerize the heparosan polysaccharides and reduce their molecular weight, which can in turn reduce the $\bar{M}_w$ of the *N*-deacetylated heparosan composition. Typically, heparosan

polysaccharides isolated from bacteria, including but not limited to *E. coli*, have a molecular weight ranging from about 3,000 Da to about 150,000 Da, and compositions of isolated heparosan can have a $\bar{M}_w$ in the range of about 25,000 Da up to about 50,000 Da (see Ly, M., et al. and Wang, et al., (2011), above). In another embodiment, and independent from its starting $\bar{M}_w$ and overall molecular weight properties, a heparosan composition either obtained from commercial sources or isolated from bacteria, including but not limited to *E. coli*, can be treated with a base, preferably sodium hydroxide, for a time sufficient to reduce the $\bar{M}_w$ of the *N*-deacetylated heparosan product to a target or desired level. In another embodiment, the depolymerized, *N*-deacetylated heparosan product has an $\bar{M}_w$ of at least 1,000 Da, including at least 2,000 Da, 4,000 Da, 6,000 Da, 7,000 Da, 8,000 Da, 8,500 Da, 9,000 Da, 9,500 Da, 10,000 Da, 10,500 Da, 11,000 Da, 11,500 Da, 12,000 Da, 12,500 Da, 13,000 Da, 13,500 Da, 14,000 Da, 15,000 Da, 16,000 Da, or 18,000 Da, up to at least 20,000 Da. In another embodiment, the depolymerized, *N*-deacetylated heparosan product has an $\bar{M}_w$ of less than 20,000 Da, including less than 18,000 Da, 16,000 Da, 15,000 Da, 14,000 Da, 13,500 Da, 13,000 Da, 12,500 Da, 12,000 Da, 11,500 Da, 11,000 Da, 10,500 Da, 10,000 Da, 9,500 Da, 9,000 Da, 8,500 Da, 8,000 Da, 7,000 Da, 6,000 Da, or 4,000 Da, down to less than 2,000 Da. In another embodiment, the depolymerized, *N*-deacetylated heparosan product has an $\bar{M}_w$ in a range from 1,000 up to 2,000 Da, or 4,000 Da, or 6,000 Da, or 7,000 Da, or 8,000 Da, or 8,500 Da, or 9,000 Da, or 9,500 Da, or 10,000 Da, or 10,500 Da, or 11,000 Da, or 11,500 Da, or 12,000 Da, or 12,500 Da, or 13,000 Da, or 13,500 Da, or 14,000 Da, or 15,000 Da, or 16,000 Da, or 18,000 Da, or 20,000 Da. In another embodiment, the anticoagulant NS2S6S3S-HS product has an $\bar{M}_w$ in a range from 7,000 Da up to 8,000 Da, or 8,500 Da, or 9,000 Da, or 9,500 Da, or 10,000 Da, or 10,500 Da, or 11,000 Da, or 11,500 Da, or 12,000 Da, or 12,500 Da, or 13,000 Da, or 13,500 Da, or 14,000 Da, or 15,000 Da. In another embodiment, the depolymerized, *N*-deacetylated heparosan product has an $\bar{M}_w$ in a range from 9,000 Da up to 9,500 Da, or 10,000 Da, or 10,500 Da, or 11,000 Da, or 11,500 Da, or 12,000 Da, or 12,500 Da. In another embodiment, the depolymerized, *N*-deacetylated heparosan product can have an $\bar{M}_w$ in any range listed above between and inclusive of 1,000 Da and 20,000 Da, and preferably in any range listed above between and inclusive of 9,000 Da and 12,500 Da.

[0307] In another embodiment, a heparosan composition can be treated with a base, preferably sodium hydroxide, for a time sufficient to both reduce the $\bar{M}_w$ of the *N*-deacetylated heparosan product to a target or desired level, and to attain a desired amount of glucosamine residues that remain *N*-acetylated within the *N*-deacetylated heparosan product. Methods for providing a starting polysaccharide reaction mixture comprising *N*-deacetylated heparosan comprise the following sub-steps: (a) providing a precursor polysaccharide composition comprising heparosan; and (b)

combining the precursor polysaccharide composition with a reaction mixture comprising a base, preferably lithium hydroxide or sodium hydroxide, for a time sufficient to *N*-deacetylate at least one of the *N*-acetylated glucosamine residues within the heparosan, forming the *N*-deacetylated heparosan composition. In another embodiment, the *N*-deacetylated heparosan product can have an $\bar{M}_w$ in any range listed above between and inclusive of 1,000 Da and 20,000 Da, simultaneously with having less than 60% of the glucosamine residues within the *N*-deacetylated heparosan product present as *N*-acetylglucosamine residues. In another embodiment, the *N*-deacetylated heparosan product can have an $\bar{M}_w$ in any range listed above between and inclusive of 9,000 Da and 12,500 Da, in which from 12% and up to 18% of the glucosamine residues within the *N*-deacetylated heparosan product are *N*-acetylated. The preparation of *N*-deacetylated heparosan having such molecular weight properties and *N*-acetyl content is described in detail in Wang, et al., (2011), above. In another embodiment, the time sufficient to react a heparosan with a base, preferably sodium hydroxide, to form an *N*-deacetylated heparosan product having an $\bar{M}_w$ in a range between 9,000 Da and 12,500 Da, as well as an *N*-acetyl glucosamine content in a range from 12% and up to 18%, can be at least 1 hour, including at least 2, 4, 6, 8, 10, 12, or 18 hours, and up to at least 24 hours, depending on the molecular weight properties and concentration of the heparosan starting material, and the identity and concentration of the base used to carry out the reaction.

[0308] In another embodiment, within any of the methods for forming an NS2S6S3S-HS product described above, any of the reaction mixtures comprising an engineered sulfotransferase and an aryl sulfate compound as a sulfo group donor can further comprise one or more reaction components for repopulating the aryl sulfate compound. In another embodiment, the one or more reaction components comprise an aryl sulfotransferase (ASST) enzyme and a secondary aryl sulfate compound. In nature, aryl sulfotransferase enzymes can catalyze the sulfation of aromatic compounds to form an aryl sulfate compound. Typically, the sulfo donor itself is an aryl sulfate compound. The reactivity of ASST enzymes is generally described, for example, in U.S. Pat Nos. 6,225,088 and 8,771,995, as well as Malojcic, et al., above, the disclosures of which are incorporated by reference in their entireties. Without being limited by a particular theory, it is believed that further including an ASST and a secondary aryl sulfate compound within a reaction mixture comprising an engineered sulfotransferase can have the advantage of reducing potential competitive inhibition of the engineered sulfotransferase by the desulfated aromatic product, as well as repopulating the reaction mixture with the sulfo group donor.

[0309] In another embodiment, the secondary aryl sulfate compound can be any aryl sulfate compound, including those described above. In another embodiment, the secondary aryl sulfate

compound is the same aryl sulfate compound used as the sulfo group donor for the engineered sulfotransferase enzyme. In another embodiment, the secondary aryl sulfate compound is a different aryl sulfate compound than the one used as the sulfo group donor for the engineered sulfotransferase enzyme. As a non-limiting example, and in another embodiment, if the engineered sulfotransferase has biological activity with NCS as a sulfo group donor, then the secondary aryl sulfate compound is PNS. In another non-limiting example, and in another embodiment, if the engineered sulfotransferase has biological activity with PNS as a sulfo group donor, then the secondary aryl sulfate compound is NCS.

[0310] In another embodiment, the ASST enzyme utilized in conjunction with any of the above methods to repopulate the sulfo donor aryl sulfate compound can be any bacterial enzyme, either isolated from *in vivo* sources or generated recombinantly *in vitro*, which transfers a sulfo group from an aryl sulfate compound to an aromatic compound. In another embodiment, and in one non-limiting example, the ASST is a recombinant ASST from *E. coli*, preferably from the *E. coli* strain CFT073 and having the amino acid sequence of SEQ ID NO: 55. In another embodiment, an ASST enzyme, preferably an ASST enzyme comprising the amino acid sequence of SEQ ID NO: 55, when coupled to any of the engineered sulfotransferases described above, can transfer a sulfate group from the secondary aryl sulfate compound to the desulfated aromatic compound formed by the engineered sulfotransferase. Without being limited by a particular theory, it is believed that utilizing the ASST can reduce potential product inhibition by the desulfated aromatic compound, while also regenerating the sulfo group donor for subsequent sulfotransfer reactions to an HS or heparosan-based polysaccharide.

[0311] In another embodiment, and also without being limited by a particular theory, it is believed that coupling the engineered sulfotransferase-catalyzed reaction with ASST can provide a further advantage of generating the aryl sulfate sulfo donor directly from a non-sulfated aromatic compound. The reaction mixture for a particular reaction catalyzed by an engineered sulfotransferase can be formulated to combine a non-sulfated aromatic compound with ASST and a secondary aryl sulfate compound either prior to or simultaneously with addition of the engineered sulfotransferase to the reaction mixture. In a non-limiting example, and in another embodiment, a sulfotransfer reaction catalyzed by an engineered sulfotransferase enzyme can be initiated by combining a non-sulfated aromatic compound, an aryl sulfate compound, and an ASST in the same reaction mixture as the engineered sulfotransferase and the polysaccharide sulfo group acceptor. The reaction between the ASST, the aryl sulfate compound, and the non-sulfated aromatic compound can generate the sulfo donor aryl sulfate compound, which can then react with the engineered sulfotransferase enzyme to

transfer the sulfate group to the polysaccharide. In another embodiment, the aryl sulfate compound produced by the reaction with the ASST enzyme is a different compound than the aryl sulfate compound that reacts with ASST itself. In a non-limiting example, the non-sulfated aromatic compound is NCS, and the aryl sulfate compound that reacts with the ASST is PNS. As NCS is formed by the reaction between PNS and ASST, the sulfo group can then be transferred from the NCS to the polysaccharide, using the engineered sulfotransferase.

Post-Synthesis Processing of HS Products

[0312] As described above, heparins that are prescribed to patients generally adhere to a tightly-regulated set of purity, molecular weight and activity requirements, whereas LMWH compositions typically have an average molecular weight of less than 8,000 Da, in which more than 60% of all of the polysaccharide molecules within the sample have an actual molecular weight of less than 8,000 Da (*see* Linhardt, R.J. and Gunay, N.S., above). Furthermore, pharmaceutical LMWH compositions have their own regulated set of molecular weight and activity requirements in their own right, and are generally prepared from heparin. Accordingly, and in another embodiment, NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS products produced by any of the methods described above can be utilized to produce LMW-HS products, using any well-known means in the art. In another embodiment, an NS2S6S3S-HS product synthesized by a method described above can be utilized to produce an LMW-HS product, which can then subsequently be *O*-desulfated to form an *O*-desulfated LMW-HS product. In a further embodiment, an NS2S6S3S-HS product synthesized by a method described above can have a purity, molecular weight, and/or anticoagulant activity equivalent to USP heparin, and the formed LMW-HS product can have a purity, molecular weight, and/or anticoagulant activity equivalent to a USP LMWH composition. In another embodiment, an NS2S6S3S-HS product synthesized by a method described above can first be *O*-desulfated, and then modified to form an *O*-desulfated LMW-HS product. In another embodiment, NS/HS, NS6S/HS, NS2S/HS, or NS2S6S/HS products can subsequently be reacted to form an LMH-NS/HS, -NS6S/HS, -NS2S/HS, or -NS2S6S/HS product. Non-limiting exemplary methods for synthesizing LMW-HS products from NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS products are described in further detail below.

[0313] In one non-limiting example, and in another embodiment, polysaccharides within an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product mixture that have a low molecular weight, particularly a molecular weight less than 15,000 Da, including less than 14,000 Da, 13,000 Da, 12,000 Da, 11,000 Da, 10,000 Da, 9,000 Da, 8,000 Da, 7,000 Da, 6,000 Da, 5,000 Da, 4,000 Da,

or 3,000 Da, down to less than 2,000 can be separated from other polysaccharides within the same mixture, such as by electrophoretic mobility using gel electrophoresis, size exclusion chromatography, and/or precipitation with salts of a divalent cation and a weak anion, including but not limited to barium, calcium, magnesium, strontium, copper, nickel, cadmium, zinc, mercury, beryllium, palladium, platinum, iron, and tin salts. In another embodiment, the polysaccharides can be separated from higher molecular-weight polysaccharides in bulk, by separating all such polysaccharides under 15,000 Da from those above 15,000 Da, as a non-limiting example. In another embodiment, the polysaccharides can be separated into one or more fractions, such as 10,000 Da to 15,000 Da, 5,000 Da to 10,000 Da, and all polysaccharides under 5,000 Da, as another non-limiting example.

[0314] In another embodiment, NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS polysaccharide product mixtures having an average molecular weight less than 8,000 Da can be utilized as LMW products directly. In other embodiments, NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS polysaccharide product mixtures having an average molecular weight less than 8,000 Da can be combined with other glycosaminoglycans (GAGs) to form HS-GAG mixtures. Although an advantage of several of the methods above, particularly methods in which the heparosan starting material is isolated and purified from *E. coli*, includes the ability to synthesize HS products that are free from chondroitin sulfate, dermatan sulfate, and other sulfated GAGs, some highly-purified HS-GAG mixtures that comprise chondroitin sulfate and/or dermatan sulfate have been successfully prescribed to patients in the past because they have beneficial pharmacological properties relative to UF-HS, even if they don't possess as much anticoagulant activity as UF-HS. In another embodiment, once an HS-GAG mixture comprising an NS2S6S3S-HS product is formed, polysaccharides within the HS-GAG mixture can be *O*-desulfated. In another embodiment, an NS2S6S3S-HS product can be *O*-desulfated prior to forming the HS-GAG mixture.

[0315] Non-limiting examples of HS-GAG mixtures that have been prescribed medically include sulodexide (CAS No: 57821-29-1) and danaparoid (CAS No: 308068-55-5). Historically, sulodexide has been extracted from pig intestinal mucosa (*see* U.S. Pat. No. 3,936,351, herein incorporated by reference in its entirety), but sulodexide can also be prepared by combining dermatan sulfate (CAS No: 24967-94-0) with the "fast-moving" heparin fractions (FM-HS) that can be separated from heparin using salt precipitation (*see* Volpi, N., (1993) *Carbohydr. Res.* **247**:263-278), particularly with barium salts. FM-HS fractions are deemed "fast-moving" based on their electrophoretic mobility relative to heavier, "slow-moving" heparin (SM-HS) that are also formed upon salt precipitation of heparin, and can be purified away from SM-HS, using ultracentrifugation,

as a non-limiting example. Additionally, FM-HS fractions have reduced anticoagulant activity and overall sulfation relative to heparin, and a relative molecular mass, M_r , as determined by high performance size exclusion chromatography (HPSEC) of about 8,000 (*see* Volpi, N., above). However, the mean molecular weight of the FM-HS fraction itself is about 7,000 Da (*see* Coccheri, S. and Mannello, F., (2014) *Drug Design, Development, and Therapy* 8:49-65).

[0316] Thus, in another embodiment, FM-HS fractions can be prepared from NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS products synthesized using engineered aryl sulfate-dependent sulfotransferase enzymes. In another embodiment, the sulfated HS product can be precipitated with divalent-cationic salt, particularly a barium or calcium salt, using a similar procedure described by Volpi, above. Methods for performing a salt precipitation of heparin to form and subsequently purify FM-HS are also described in U.S. Patents 7,687,479 and 8,609,632, the disclosures of which are herein incorporated by reference in their entireties. In another embodiment, once the resulting FM-HS fraction is purified, it can be combined with dermatan sulfate to form an HS-GAG mixture. In another embodiment, methods of the present invention can be utilized to synthesize FM-HS directly, which can then be combined with dermatan sulfate to form an HS-GAG mixture. In another embodiment, the HS-GAG mixture prepared by either method can comprise one or more properties that are identical to sulodexide, including but not limited to a composition comprising 80% of the FM-HS fraction and 20% of dermatan sulfate (*see* Lauver, D.A. Lucchesi, B.R., *Cardio. Drug Rev.* 24 (3-4):214-216), an average molecular weight of 7,000 Da, an M_r of about 8,000, and/or a sulfate to carboxyl group ratio in the range of 2.0:1 to 2.2:1.

[0317] In contrast to sulodexide, the HS-GAG mixture, danaparoid, has been historically prepared from natural HS isolated from porcine sources, rather than from unfractionated heparin (*see* U.S. Patent No. 5,164,377, herein incorporated by reference in its entirety; *see also* “Danaparoid Sodium” (2010) *European Pharmacopoeia* 7.0, 1789-1792). Such HS polysaccharide compositions, as opposed to heparin, contain some polysaccharides having disaccharide units that are generally either unsulfated or are *N*-, 2-*O*, and/or 6-*O* sulfated, resulting in a dramatically reduced anticoagulant activity relative to unfractionated heparin.

[0318] Additionally, upon purifying danaparoid according to the procedures in U.S. Patent No. 5,164,377, the resulting product contains not only HS, but also chondroitin sulfate and dermatan sulfate, that have reduced molecular weights as a result of the addition of a base during the extraction process, similar to the effect of reacting a base with heparosan to reduce the molecular weight. According to the *European Pharmacopoeia*, the weight-average molecular weight ($\bar{M}_w$) of all of the GAGs within a danaparoid HS-GAG composition suitable to be prescribed to patients is in

a range of at least 4,000 Da, up to 7,000 Da, and comprise the following size distribution limits: (a) polysaccharide chains comprising an M_r of less than 2,000 comprise a maximum of 13% (w/w) of the danaparoid mixture; (b) polysaccharide chains comprising an M_r of less than 4,000 comprise a maximum of 39% (w/w) of the danaparoid mixture; (c) polysaccharide chains comprising an M_r between 4,000 and 8,000 comprise a minimum of 50% (w/w) of the danaparoid mixture; (d) polysaccharide chains comprising an M_r of higher than 8,000 comprise a maximum of 19% (w/w) of the danaparoid mixture; and (e) polysaccharide chains comprising an M_r of less than 10,000 comprise a maximum of 11% (w/w) of the danaparoid mixture. With regard to particular composition limits for danaparoid determined by the *European Pharmacopoeia*, chondroitin sulfate can comprise a maximum of 8.5% (w/w) of the danaparoid mixture, and dermatan sulfate can comprise a range from at least 8.0% (w/w) up to 16.0% (w/w) of the danaparoid mixture. As a non-limiting example, the danaparoid composition Orgaran® comprises about 84% (w/w) HS, about 12% (w/w) dermatan sulfate, and about 4% chondroitin sulfate.

[0319] In another embodiment, an HS-GAG mixture comprising an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product produced by any of the methods of the present invention using engineered aryl sulfate-dependent sulfotransferase enzymes, dermatan sulfate, and chondroitin sulfate can be formed that has similar properties to danaparoid (CAS No: 308068-55-5). In another embodiment, the HS product is an NS2S6S/HS product. In another embodiment, the HS product is an NS2S6S3S-HS product. In another embodiment, the HS product synthesized directly from the reaction has an $\bar{M}_w$ in a range from at least 4,000 Da, and up to 8,000 Da, preferably in a range from at least 4,000 Da, up to 7,000 Da. In another embodiment, the HS product has an $\bar{M}_w$ larger than 8,000 Da, and is prepared for inclusion in a danaparoid-like HS-GAG mixture by subsequently reacting it with a base, similar to methods described above for reducing its molecular weight. In another embodiment, chondroitin sulfate and dermatan sulfate are also reacted with a base to reduce their molecular weight. In another embodiment, a composition comprising an HS product produced by any of the methods of the present invention, chondroitin sulfate, and dermatan sulfate can be filtered using a filtration device. Such filtration devices can include, but are not limited to, centrifugal filter units such as an Amicon® Ultra unit (EMD Millipore), or dialysis membranes, either of which have a desired molecular weight cut-off (MWCO). In another embodiment, the MWCO for either a centrifugal filter unit or dialysis membrane is 5,500 Da. In another embodiment, the $\bar{M}_w$ for all of the GAGs in the danaparoid HS-GAG mixture is in a range from at least 4,000 Da, and up to 8,000 Da, preferably in a range from at least 4,000 Da, and up to 7,000 Da, and more preferably in a range from at least 5,000 Da, and up to 6,000 Da. In another embodiment, GAGs

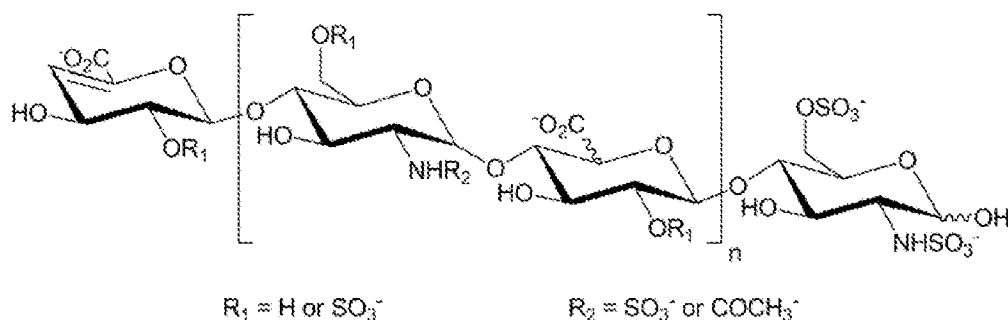
within the danaparoid HS-GAG mixture comprise the following size distribution limits: (a) polysaccharide chains comprising an M_r of less than 2,000 comprise a maximum of 13% (w/w) of the danaparoid HS-GAG mixture; (b) polysaccharide chains comprising an M_r of less than 4,000 comprise a maximum of 39% (w/w) of the danaparoid HS-GAG mixture; (c) polysaccharide chains comprising an M_r between 4,000 and 8,000 comprise a minimum of 50% (w/w) of the danaparoid HS-GAG mixture; (d) polysaccharide chains comprising an M_r of higher than 8,000 comprise a maximum of 19% (w/w) of the danaparoid HS-GAG mixture; and (e) polysaccharide chains comprising an M_r of less than 10,000 comprise a maximum of 11% (w/w) of the danaparoid HS-GAG mixture.

[0320] In another embodiment, the danaparoid HS-GAG mixture can comprise a GAG composition that is either similar or identical to danaparoid (CAS No: 308068-55-5). In another embodiment, the composition of the GAGs within the danaparoid HS-GAG mixture comprises at least 8 % (w/w), up to 16% (w/w), and preferably 12% (w/w) of dermatan sulfate, and less than 8% (w/w), preferably in a range of at least 3 % (w/w), up to 5% (w/w), and more preferably 4% (w/w) of chondroitin sulfate. In another embodiment, the danaparoid HS-GAG mixture can comprise either a similar or identical anticoagulant activity to danaparoid, prior to being *O*-desulfated.

[0321] In another embodiment, NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product mixtures synthesized according to any of the methods of the present invention can be further modified by one or more subsequent processes to depolymerize and/or modify the HS product to form an LMW-HS product, as described above. In further embodiments, an NS6S/HS product mixture is depolymerized and/or modified to form an LMW-NS6S/HS product. Generally, and in another embodiment, the process for forming an LMW-HS product from an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product mixture comprises the following steps: (a) synthesizing an HS product according to any of the above methods; (b) providing one or more depolymerization agents; and (c) treating the HS product with the one or more depolymerization agents for a time sufficient to depolymerize at least a portion of the polysaccharides within the HS product, thereby forming the LMW-HS product. Without being limited by a particular theory, it is believed that the choice in the depolymerization agent can determine the chemical mechanism for forming the LMW-HS product, as well as the product(s) structure, anticoagulant activity (if prepared from an NS2S6S3S-HS product), and pharmacological properties. Known chemical mechanisms for forming an LMW-HS product from an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product include, but are not limited to: chemical and/or enzymatic β -elimination reactions; deamination reactions; and oxidation reactions, including combinations thereof.

[0322] In another embodiment, an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product, synthesized according to any of the methods of the present invention, can be modified by an enzymatic β -elimination reaction to form an enzymatically-depolymerized LMW-HS product. Historically, enzymatically-depolymerized LMWH products have been prepared by incubating USP heparin with one or more heparinase enzymes until the LMW-HS product comprises a desired chemical structure, average molecular weight, anticoagulant activity, and degree of sulfation. (see “Tinzaparin Sodium” (2010) *European Pharmacopoeia* 7.0, 3098; see also Linhardt, R.J. and Gunay, N.S., above).

[0323] As an example, upon reacting an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product with the one or more heparinases, one or more of the polysaccharides within the NS2S6S3S-HS product both depolymerizes and develops a characteristic chemical structure, illustrated by Formula XI, below.



[0324] As illustrated above in Formula XI, n can be any integer from 1-25. Instead of a glucuronic acid or uronic acid residue, the sugar residue at the non-reducing end of one or more of the enzymatically-depolymerized LMW-HS polysaccharides within the product is a 2-*O*-sulfo-4-enepyranosulfonic acid. Additionally, each glucosamine residue at the reducing end is sulfated at the *N*- and 6-*O* positions. When the starting material is NS2S6S3S-HS, the 3-*O* position of a glucosamine residue within one or more of disaccharide units can also be 3-*O* sulfated, and the enzymatically-depolymerized LMW-HS product has anticoagulant activity. When the starting material is NS6S/HS or NS2S6S/HS, the resulting enzymatically-depolymerized LMW-HS product can also comprise the structure of Formula XI, but without having any 3-*O* sulfated glucosamine residues or anticoagulant activity.

[0325] Further, much like USP heparin, enzymatically-depolymerized LMWH products prescribed as anticoagulants must satisfy strict purity and property standards. In particular, one such enzymatically-depolymerized LMW-HS product, tinzaparin (CAS No: 9041-08-1; ATC code: B01AB10), has a particular set of molecular weight, anticoagulant activity, and sulfation content properties in addition to the chemical structure of Formula XI above, including: an $\bar{M}_w$ in a range

from at least 5,500 Da, and up to 7,500 Da, and characteristically 6,500 Da; at least 1.8 and up to 2.5 sulfate groups per disaccharide unit; and an anti-Xa activity of at least 70 IU mg⁻¹ and up to 120 IU mg⁻¹, and/or a ratio of anti-Xa activity to anti-IIa activity of at least 1.5:1, and up to 2.5:1. In another embodiment, an enzymatically-depolymerized LMW-HS product formed from NS2S6S3S-HS comprises one or more properties that are identical to tinzaparin, including but not limited to chemical structure, molecular weight, size distribution, anticoagulant activity, and/or sulfation content. In another embodiment, the enzymatically-depolymerized LMW-HS product formed from NS2S6S3S-HS is substantially identical to tinzaparin. In another embodiment, an enzymatically-depolymerized LMW-HS product, for example, products having one or more properties that are identical to tinzaparin or products that are substantially equivalent to tinzaparin, is subsequently O-desulfated, according to any of the procedures described above, to form an enzymatically-depolymerized ODSH product.

[0326] In another embodiment, the at least one heparinase can be a heparinase from any species, so long as the enzyme catalyzes β -eliminative cleavage of HS polysaccharides. In another embodiment, the at least one heparinase can be selected from the group consisting of the heparinases from *Bacteroides eggerthii* comprising the amino acid sequences of SEQ ID NO: 30, SEQ ID NO: 31, and SEQ ID NO: 32. In another embodiment, the at least one heparinase can comprise one, two, or all three of the enzymes having the amino acid sequences of SEQ ID NO: 30, SEQ ID NO: 31, and SEQ ID NO: 32, respectively.

[0327] In another embodiment, the time sufficient to form the enzymatically-depolymerized LMW-HS product is the time sufficient to cause the product to have a desired average molecular weight. In another embodiment, the $\bar{M}_w$ of the enzymatically-depolymerized LMW-HS product can be in the range of 2,000 Da to 10,000 Da, and when the starting material is an NS2S6S3S-HS product, preferably 5,500 Da to 7,500 Da, and more preferably 6,500 Da. In another embodiment, the enzymatically-depolymerized LMW-HS product can have anticoagulant activity, preferably an anti-Xa activity of at least 70 IU mg⁻¹ and up to 120 IU mg⁻¹, and/or a ratio of anti-Xa activity to anti-IIa activity of at least 1.5:1, and up to 2.5:1.

[0328] In another embodiment, an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product, synthesized according to any of the methods of the present invention, can be modified by a chemical β -elimination reaction to form a chemically β -eliminative, LMW-HS product. Historically, chemically β -eliminative LMWH products have been prepared by treating USP heparin or its quaternary ammonium salt with a base. Under these conditions, chemical β -elimination takes place, forming a chemically β -eliminative LMW-HS product having polysaccharides containing a 4,5-

unsaturated uronic acid residue at the non-reducing end, a feature observed in enzymatically-depolymerized LMW-HS products, described above (*see* Linhardt, R.J. and Gunay, N.S., above).

[0329] Control of the reaction conditions has led to the production of chemically β -eliminative LMW-HS compositions that have either been approved for clinical use or been administered during clinical trials, and are described in more detail below. When the starting material is NS2S6S3S-HS, the chemically β -eliminative LMW-HS product can contain polysaccharides comprising the structure of Formula XI, in which the 3-*O* position of a glucosamine residue within one or more of disaccharide units can also be 3-*O* sulfated. Several of these chemically β -eliminative LMW-HS compositions have anticoagulant activity. When the starting material is NS6S/HS or NS2S6S/HS, the resulting chemically β -eliminative LMW-HS product can also comprise the structure of Formula XI, but without having any 3-*O* sulfated glucosamine residues or anticoagulant activity.

[0330] In a first non-limiting example, an anticoagulant chemically β -eliminative LMW-HS composition that has been prescribed for clinical use is bemiparin (CAS No: 91449-79-5; ATC code: B01AB12) (*see e.g.* Chapman, T.M. and Goa, K.L., (2003) *Drugs* **63** (21):2357-2377; Sanchez-Ferrer, C.F. (2010) *Drugs* **70** Suppl. 2:19-23; Ciccone, M.M., et al., (2014) *Vascular Pharmacology* **62**:32-37). Bemiparin is prepared by alkaline depolymerization of USP heparin, particularly by reacting the benzethonium salt of USP heparin with a quaternary ammonium hydroxide, such as Triton[®] B (benzyl trimethylammonium hydroxide), in the presence of methanol (*see* U.S. Pat. No. 4,981,955 and European Patent EP 0293539, the disclosures of which are incorporated by reference in their entireties). Upon subsequent purification and precipitation, the resulting bemiparin composition comprising the structure of Formula XI has an $\bar{M}_w$ in a range of at least 3,000 Da, up to 4,200 Da, and typically 3,600 Da, and a size distribution such that: less than 35% of the polysaccharide chains have an M_r less than 2,000; a range of at least 50% and up to 75% of the polysaccharide chains have an M_r in a range of at least 2,000 and up to 6,000; and less than 15% of the polysaccharide chains have an M_r greater than 6,000. Additionally, bemiparin compositions can comprise an anti-Xa activity of at least 80 IU mg⁻¹ and up to 120 IU mg⁻¹, an anti-IIa activity of at least 5 IU mg⁻¹ and up to 20 IU mg⁻¹, and/or a ratio of anti-Xa activity to anti-IIa activity of at least 8.0:1, and up to 10:1 (*see* Sanchez-Ferrer, C.F., above).

[0331] Accordingly, a chemically β -eliminative LMW-HS composition can be prepared from an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product by the following steps: (i) reacting the NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product with a benzethonium salt, preferably benzethonium chloride, to form a benzethonium HS salt; and (ii) combining the benzethonium HS salt with a reaction mixture comprising Triton[®] B and methanol for

a time sufficient to form the chemically β -eliminative LMW-HS product. In another embodiment, the time sufficient to depolymerize the benzethonium HS salt is the time sufficient to form a chemically β -eliminative LMW-HS product to having an $\bar{M}_w$ in a range of at least 3,000 Da, up to 4,200 Da, and preferably 3,600 Da, and having a size distribution such that: less than 35% of the polysaccharide chains have an M_r less than 2,000; a range of at least 50% and up to 75% of the polysaccharide chains have an M_r in a range of at least 2,000 and up to 6,000; and less than 15% of the polysaccharide chains have an M_r greater than 6,000. In another embodiment, the step of preparing the chemically β -eliminative LMW-HS product from the benzethonium HS salt comprises the procedure reported in any of the examples in U.S. Pat No. 4,981,955, preferably Example 3. In another embodiment, a chemically β -eliminative LMW-HS product formed from NS2S6S3S-HS comprises one or more properties that are identical to bemiparin, including but not limited to chemical structure, molecular weight, size distribution, anticoagulant activity, and/or sulfation content. In another embodiment, the chemically β -eliminative LMW-HS product formed from NS2S6S3S-HS is substantially identical to bemiparin. In another embodiment, a chemically β -eliminative LMW-HS product formed from NS2S6S3S-HS, for example, products having one or more properties that are identical to bemiparin or products that are substantially equivalent to bemiparin, is subsequently *O*-desulfated, according to any of the procedures described above, to form a chemically β -eliminative ODSH product.

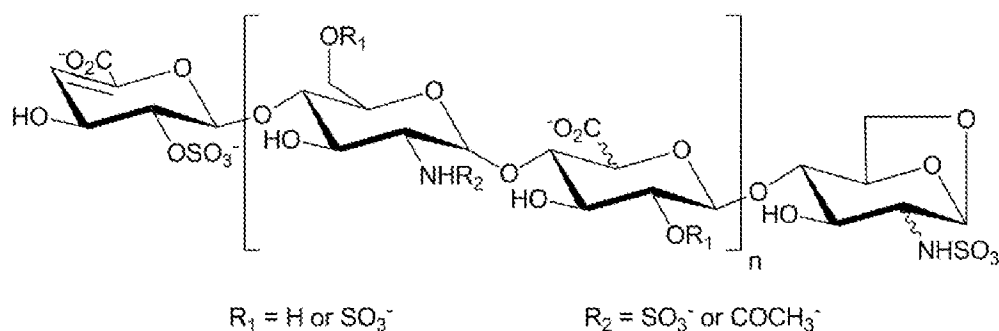
[0332] In another non-limiting example, a chemically β -eliminative LMW-HS composition that has been administered during clinical trials is semuloparin (CAS No: 9041-08-1). Semuloparin is prepared by reacting the benzyl ester of a heparin benzethonium salt with the strong phosphazene base, BEMP (2-tert-butylimino-2-diethylamino-1,3-dimethylperhydro-1,2,3-diaza-phosphorine), with subsequent saponification of the benzyl esters and purification (*see* Viskov, C., et al., (2009) *J. Thromb. Haemost.* 7:1143-1551). Phosphazene bases are among the strongest-known organic bases, by are highly-sterically hindered and non-nucleophilic. As a result, phosphazene bases target the least sterically hindered regions of polysaccharides within USP heparin for β -elimination, and avoid the AT-recognition sequence that comprises the 3-*O* sulfated glucosamine residue. The resulting semuloparin product having the structure of Formula XI has an $\bar{M}_w$ in a range of at least 2,000 Da, up to 3,000 Da, and typically 2,400 Da, and the anticoagulant activity of the semuloparin product comprises an anti-Xa activity of about 160 IU mg⁻¹, an anti-IIa activity of about 2 IU mg⁻¹, and a ratio of anti-Xa activity to anti-IIa activity of about 80:1 (*see* Viskov, C., above).

[0333] Accordingly, a chemically β -eliminative LMW-HS composition can be prepared from an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product by the following steps: (i)

reacting the NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product with a benzethonium salt, preferably benzethonium chloride, to form a benzethonium HS salt; (ii) esterification of the benzethonium HS salt using benzyl chloride to form a benzyl ester HS; (iii) transalification of the benzyl ester HS with a benzethonium salt, preferably benzethonium chloride, to form a benzethonium benzyl ester HS; (iv) depolymerization of the benzethonium benzyl ester HS with BEMP to form a benzyl ester chemically β -eliminative LMW-HS product; and (v) saponification of the benzyl ester chemically β -eliminative LMW-HS product to form the chemically β -eliminative LMW-HS product, as reported in Viskov, C., et al., above. In another embodiment, the time sufficient to depolymerize the benzethonium benzyl ester HS with BEMP is the time sufficient to form a benzyl ester chemically β -eliminative LMW-HS product such that upon saponification of the benzyl esters, the resulting chemically β -eliminative LMW-HS product has an $\bar{M}_w$ in a range of at least 2,000 Da, up to 3,000 Da, and preferably about 2,400 Da. In another embodiment, a chemically β -eliminative LMW-HS product formed from NS2S6S3S-HS comprises one or more properties that are identical to semuloparin, including but not limited to chemical structure, molecular weight, size distribution, anticoagulant activity, and/or sulfation content. In another embodiment, the chemically β -eliminative LMW-HS product formed from NS2S6S3S-HS is substantially identical to semuloparin. In another embodiment, a chemically β -eliminative LMW-HS product formed from NS2S6S3S-HS, for example, products having one or more properties that are identical to semuloparin or products that are substantially equivalent to semuloparin, is subsequently *O*-desulfated, according to any of the procedures described above, to form a chemically β -eliminative ODSH product.

[0334] In another non-limiting example, a chemically β -eliminative LMW-HS composition that has been prescribed for clinical use is enoxaparin (CAS No: 679809-58-6; ATC code: B01AB05) (*see e.g.* Linhardt, R.J. and Gunay, N.S., above). Enoxaparin is prepared similarly to semuloparin in that a benzyl ester form of USP heparin is prepared, before being reacted with a base. The benzyl ester is formed in a chlorinated organic solvent, such as chloroform or methylene chloride, in the presence of a chlorine derivative, such as benzyl chloride, which controls the amount of esterification in the resulting heparin benzyl ester, with about 9-14% efficiency. Once the benzyl ester is formed, it is subsequently treated with a strong, non-sterically hindered base, such as sodium hydroxide, at high temperature (*see* U.S. Patent No. 5,389,618 and U.S. Reissue Patent RE38,743, the disclosures of which are incorporated by reference in their entireties. However, some (about 15% to 25%) polysaccharides within enoxaparin can additionally comprise a terminal 1,6-anhydro sugar residue (either 1,6-anhydromannose or 1,6-anhydroglucosamine) at the reducing end, in addition to the

characteristic 4,5-unsaturated uronic acid at the non-reducing end (see Guerrini, M., (2010) *J. Med. Chem.* **53**:8030-8040). As a result, enoxaparin typically comprises polysaccharides having the characteristic structure illustrated in Formula XII, below, in addition to polysaccharides comprising the structure of Formula XI.



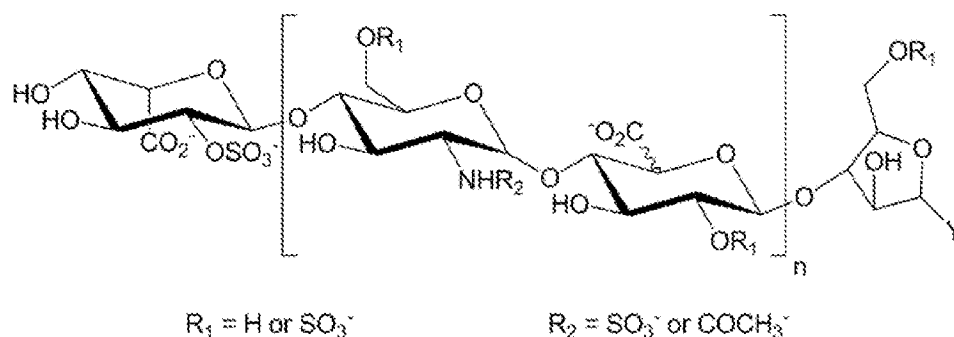
[0335] As illustrated above in Formula XII, n can be any integer from 1-21. Instead of a glucuronic acid or uronic acid residue, the sugar residue at the non-reducing end of enoxaparin polysaccharides can be 2-*O*-sulfo-4-enepyransulfonic acid. Additionally, the glucosamine residue at the reducing end can comprise a 1,6-anhydro moiety, and the stereochemistry around the C2 carbon determines whether the residue is a 1,6-anhydromannose or 1,6-anhydroglucosamine residue. Optionally, the 3-*O* position of a glucosamine residue within one or more of disaccharide units can also be 3-*O* sulfated. Without being limited by a particular theory, it is believed that at least some of the polysaccharides within enoxaparin comprises 3-*O* sulfated glucosamine residues, which ultimately leads to its anticoagulant activity.

[0336] As a commonly prescribed LMWH drug, compositions of enoxaparin that are administered to patients must satisfy a series of stringent size, activity, and purity requirements established by both the European Pharmacopoeia and the USP. (see "Enoxaparin Sodium" (2010) *European Pharmacopoeia* 7.0, 1920-1921). In addition to comprising the structure of Formula XII above, properties that must be present in order to satisfy the requirements include: an $\bar{M}_w$ in a range from at least 3,800 Da, and up to 5,000 Da, and characteristically 4,500 Da; not less than 1.8 sulfate groups per disaccharide unit; and an anti-Xa activity of at least 90 IU mg^{-1} and up to 125 IU mg^{-1} , an anti-IIa activity of at least 20 IU mg^{-1} and up to 35 IU mg^{-1} ; and/or a ratio of anti-Xa activity to anti-IIa activity of at least 3.3:1, and up to 5.3:1. Further, enoxaparin compositions suitable to be administered to patients comprise a size distribution such that: at least 12.0%, up to 20.0% percent, and characteristically about 16%, of the polysaccharide chains have an M_r less than 2,000; a range of at least 68.0%, up to 82.0%, and characteristically about 74%, of the polysaccharide chains have an M_r in a range of at least 2,000 and up to 8,000; and not more than 18.0% of the polysaccharide chains have an M_r greater than 8,000.

[0337] Accordingly, a chemically β -eliminative LMW-HS composition can be prepared from an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product by the following steps: (i) reacting the unfractionated NS2S6S3S-HS product with a benzethonium salt, preferably benzethonium chloride, to form a benzethonium HS salt; (ii) esterification of the benzethonium HS salt using benzyl chloride in the presence of a chlorinated solvent, preferably methylene chloride or chloroform, to form a benzyl ester HS; and (iii) combining the benzyl ester HS with a reaction mixture comprising sodium hydroxide to form the chemically β -eliminative LMW-HS product. In another embodiment, the benzyl ester HS has a degree of esterification of at least 9%, and up to about 14%. In another embodiment, the reaction between the benzyl ester HS and sodium hydroxide is performed at a temperature selected within the range of at least 50 °C, up to 70 °C, and preferably within the range of at least 55 °C, and up to 65 °C. In another embodiment, the benzyl ester HS and chemically β -eliminative LMW-HS product are prepared according to the procedure of Example 3 within US RE38,743. In another embodiment, the time sufficient to depolymerize the benzyl ester HS is the time sufficient to form a chemically β -eliminative LMW-HS product to having an $\bar{M}_w$ in a range of at least 3,800 Da, up to 5,000 Da, and preferably 4,500 Da. In another embodiment, the chemically β -eliminative LMW-HS product comprises a size distribution such that: at least 12.0%, up to 20.0% percent, and preferably about 16%, of the polysaccharide chains have an M_r less than 2,000; a range of at least 68.0%, up to 82.0%, and preferably about 74%, of the polysaccharide chains have an M_r in a range of at least 2,000 and up to 8,000; and not more than 18.0% of the polysaccharide chains have an M_r greater than 8,000. In another embodiment, the chemically β -eliminative LMW-HS product comprises not less than 1.8 sulfate groups per disaccharide unit. In another embodiment, a chemically β -eliminative LMW-HS product formed from NS2S6S3S-HS comprises one or more properties that are identical to enoxaparin, including but not limited to chemical structure, molecular weight, size distribution, anticoagulant activity, and/or sulfation content. In another embodiment, the chemically β -eliminative LMW-HS product formed from NS2S6S3S-HS is substantially identical to enoxaparin. In another embodiment, a chemically β -eliminative LMW-HS product formed from NS2S6S3S-HS, for example, products having one or more properties that are identical to enoxaparin or products that are substantially equivalent to enoxaparin, is subsequently *O*-desulfated, according to any of the procedures described above, to form a chemically β -eliminative ODSH product.

[0338] In another embodiment, an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product, synthesized according to any of the methods of the present invention, can be modified by a deamination reaction to form a deaminated LMW-HS product. Historically, deaminated LMWH

products have been prepared by treating USP heparin with nitrous acid. Under these conditions, a deaminated LMW-HS product is formed that contains polysaccharides having a 2-*O*-sulfo- α -L-idopyranosuronic acid residue at the non-reducing end, and a 6-*O*-sulfo-2,5-anhydro-D-mannitol residue at the reducing end (*see* Linhardt, R.J. and Gunay, N.S., above). Deaminated LMW-HS products comprising 2-*O*-sulfo- α -L-idopyranosuronic acid residues at the non-reducing end and 6-*O*-sulfo-2,5-anhydro-D-mannitol residues at the reducing end generally comprise the structure of Formula XIII, below:



[0339] As illustrated above in Formula XIII, *n* can be any integer from 3-20, and *Y* can be an aldehyde, hydroxyl, or carboxylic acid functional group. In another embodiment, *Y* is a hydroxyl group. Optionally, the 3-*O* position of a glucosamine residue within one or more of disaccharide units can also be 3-*O* sulfated. Without being limited by a particular theory, it is believed that at least some of the polysaccharides within the deaminated LMW-HS product comprises 3-*O* sulfated glucosamine residues, which ultimately leads to its anticoagulant activity.

[0340] When the starting material is NS2S6S3S-HS, the deaminated LMW-HS product can contain polysaccharides comprising the structure of Formula XIII, in which the 3-*O* position of a glucosamine residue within one or more of disaccharide units can also be 3-*O* sulfated. Several of these deaminated LMW-HS compositions have anticoagulant activity. On the other hand, when the starting material is NS6S/HS or NS2S6S/HS, the resulting chemically β -eliminative LMW-HS product can also comprise the structure of Formula XI, but without having any 3-*O* sulfated glucosamine residues or anticoagulant activity. Non-limiting examples of deaminated LMW-HS compositions prepared from USP heparin that have been prescribed for clinical use include dalteparin (CAS No: 9041-08-1; ATC code: B01AB04), nadroparin (CAS No: 9005-49-6; ATC code: B01AB06), reviparin (CAS No: 9005-49-6; ATC code: B01AB08) and certoparin (CAS No: 9005-49-6). Generally, each of dalteparin, nadroparin, and reviparin are prepared by depolymerization using nitrous acid, either added directly or formed *in situ* by the addition of sodium nitrite to an acidic composition. Certoparin is prepared similarly, using a nitrous acid

derivative such as isoamyl nitrite (*see* Linhardt, R.J. and Gunay, N.S., above). Control of the reaction conditions has led to the production of deaminated LMWH compositions that have slightly different anticoagulant activities and molecular weight properties relative to each other, and described, for example, in U.S. Pat Nos. 4,303,651, 4,351,938, 4,438,261, 4,500,519, 4,686,388, 5,019,649, and 5,599,801, the disclosures of which are incorporated by reference in their entireties.

[0341] In a first non-limiting example, an anticoagulant deaminated LMW-HS composition that has been prescribed for clinical use is dalteparin (*see e.g.* Jacobsen, A.F., et al., (2003) *Br J Obstet Gynaecol* **110**:139-144; and Guerrini, M., et al., (2007) *Seminars in Thrombosis and Hemostasis* **33** (5):478-487). Dalteparin is typically prepared as a sodium salt by an acid depolymerization of USP heparin, particularly by reacting USP heparin with nitrous acid (*see e.g.* U.S. Pat. No 5,019,649). Upon subsequent purification and precipitation, the resulting dalteparin composition has an $\bar{M}_w$ in a range of at least 5,600 Da, up to 6,400 Da, and typically 6,000 Da, and a size distribution such that the proportion of polysaccharide chains having an M_r less than 3,000 is not more than 13.0%; and at least 15.0% and up to 25.0% of the chains have an M_r of at least 8,000. Additionally, dalteparin compositions can comprise an anti-Xa activity of at least 110 IU mg⁻¹ and not more than 210 IU mg⁻¹, an anti-IIa activity of at least 35 IU mg⁻¹ and not more than 100 IU mg⁻¹, and/or a ratio of anti-Xa activity to anti-IIa activity of at least 1.9:1, and up to 3.2:1 (*see* “Dalteparin Sodium” (2010) *European Pharmacopoeia* 7.0, 1788-1789).

[0342] In another non-limiting example, an anticoagulant deaminated LMW-HS composition that has been prescribed for clinical use is nadroparin. Nadroparin is commonly prepared as a sodium or calcium salt by an acid depolymerization of USP heparin, using sodium nitrite in the presence of hydrochloric acid to maintain a pH of about 2.5 (*see e.g.* U.S. Pat Nos. 4,686,388 and 5,599,801) the disclosures of which are incorporated by reference in their entireties). Upon subsequent purification and precipitation, the resulting nadroparin composition has an $\bar{M}_w$ in a range of at least 3,600 Da, up to 5,000 Da, and typically 4,300 Da, and a size distribution such that the proportion of chains having an M_r less than 2,000 is not more than 15%; and at least 75% and up to 95% of the chains have an M_r in a range of at least 2,000 and up to 8,000, with at least 35% and up to 55% of the chains having an M_r of at least 2,000 and up to 4,000. Additionally, nadroparin compositions can comprise an anti-Xa activity of not less than 95 IU mg⁻¹ and not more than 130 IU mg⁻¹, and/or a ratio of anti-Xa activity to anti-IIa activity of at least 2.5:1, and up to 4.0:1 (*see* “Nadroparin Sodium” (2010) *European Pharmacopoeia* 7.0, 1788-1789).

[0343] Other non-limiting examples of deaminated LMW-HS compositions that have been prescribed for clinical use is reviparin and certoparin. Reviparin is prepared similarly to dalteparin

and nadroparin, by introducing nitrous acid or forming nitrous acid *in situ* (see Linhardt, R.J. and Gunay, N.S., above), and the resulting reviparin composition comprising the structure of Formula XIII has an $\bar{M}_w$ in a range of at least 4,200 Da, up to 4,600 Da, and typically 4,400 Da, and a ratio of anti-Xa activity to anti-IIa activity of at least 4.0:1, up to 4.5:1, and typically 4.2:1 (see Grey, et al, above). Certoparin is prepared by reacting heparin with isoamyl nitrite in the presence of acetic or hydrochloric acid (see Ahsan, A., et al., (2000) *Clin. Appl. Thrombosis/Hemostasis* 6 (3):169-174). The resulting certoparin composition comprising the structure of Formula XIII has an $\bar{M}_w$ in a range of at least 5,000 Da, up to 5,600 Da, and typically 5,400 Da, and a ratio of anti-Xa activity to anti-IIa activity of at least 2.0:1, up to 2.5:1, and preferably 2.4:1 (see Grey, et al, above).

[0344] Accordingly, a deaminated LMW-HS composition can be prepared from an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product by the following steps: (a) synthesizing an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product according to any of the above methods; (b) providing a deamination reaction mixture comprising a deamination agent, preferably a deamination agent selected from the group consisting of isoamyl nitrate and nitrous acid; and (c) treating the NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product with the deamination reaction mixture for a time sufficient to depolymerize at least a portion of the NS2S6S3S-HS product, thereby forming the deaminated LMW-HS product. In another embodiment, the deamination agent is nitrous acid, the deamination reaction mixture can comprise stoichiometric quantities of an acid, preferably acetic acid or hydrochloric acid, and an alkali or alkaline earth metal nitrite salt, preferably sodium nitrite, wherein the nitrous acid is formed within the deamination reaction mixture *in situ*. In another embodiment, the deamination agent is isoamyl nitrite. In another embodiment, the time sufficient to form the deaminated LMW-HS product is the time sufficient to cause the product to have a desired average molecular weight. In another embodiment, the $\bar{M}_w$ of the deaminated LMW-HS product is in the range of 2,000 Da to 10,000 Da, preferably in the range of 4,000 Da to 6,000 Da. In another embodiment, the $\bar{M}_w$ of the deaminated LMW-HS product is in the range 4,000 Da to 4,500 Da, preferably 4,300 Da. In another embodiment, the $\bar{M}_w$ of the deaminated LMW-HS product is in the range 4,200 Da to 4,600 Da, preferably 4,400 Da. In another embodiment, the $\bar{M}_w$ of the deaminated LMW-HS product is in the range 5,000 Da to 5,600 Da, preferably 5,400 Da. In another embodiment, the $\bar{M}_w$ of the deaminated LMW-HS product is in the range 5,700 Da to 6,300 Da, preferably 6,000 Da.

[0345] In another embodiment, a deaminated LMW-HS product formed from NS2S6S3S-HS comprises one or more properties that are identical to dalteparin, including but not limited to chemical structure, molecular weight, size distribution, anticoagulant activity, and/or sulfation

content. In another embodiment, the deaminated LMW-HS product formed from NS2S6S3S-HS is substantially identical to dalteparin. In another embodiment, a deaminated LMW-HS product formed from NS2S6S3S-HS, for example, products having one or more properties that are identical to dalteparin or products that are substantially equivalent to dalteparin, is subsequently *O*-desulfated, according to any of the procedures described above, to form a deaminated ODSH product.

[0346] In another embodiment, a deaminated LMW-HS product formed from NS2S6S3S-HS comprises one or more properties that are identical to nadroparin, including but not limited to chemical structure, molecular weight, size distribution, anticoagulant activity, and/or sulfation content. In another embodiment, the deaminated LMW-HS product formed from NS2S6S3S-HS is substantially identical to nadroparin. In another embodiment, a deaminated LMW-HS product formed from NS2S6S3S-HS, for example, products having one or more properties that are identical to nadroparin or products that are substantially equivalent to nadroparin, is subsequently *O*-desulfated, according to any of the procedures described above, to form a deaminated ODSH product.

[0347] In another embodiment, a deaminated LMW-HS product formed from NS2S6S3S-HS comprises one or more properties that are identical to certoparin, including but not limited to chemical structure, molecular weight, size distribution, anticoagulant activity, and/or sulfation content. In another embodiment, the deaminated LMW-HS product formed from NS2S6S3S-HS is substantially identical to certoparin. In another embodiment, a deaminated LMW-HS product formed from NS2S6S3S-HS, for example, products having one or more properties that are identical to certoparin or products that are substantially equivalent to certoparin, is subsequently *O*-desulfated, according to any of the procedures described above, to form a deaminated ODSH product.

[0348] In another embodiment, a deaminated LMW-HS product formed from NS2S6S3S-HS comprises one or more properties that are identical to reviparin, including but not limited to chemical structure, molecular weight, size distribution, anticoagulant activity, and/or sulfation content. In another embodiment, the deaminated LMW-HS product formed from NS2S6S3S-HS is substantially identical to reviparin. In another embodiment, a deaminated LMW-HS product formed from NS2S6S3S-HS, for example, products having one or more properties that are identical to reviparin or products that are substantially equivalent to reviparin, is subsequently *O*-desulfated, according to any of the procedures described above, to form a deaminated ODSH product.

[0349] In another embodiment, an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product, synthesized according to any of the methods of the present invention, can be modified by an oxidation reaction to form an oxidized LMW-HS product. Historically, oxidized LMWH products

have been prepared by treating USP heparin with an acid, and then reacting the acidified heparin with an oxidizing agent, particularly a peroxide or a superoxide compound such as hydrogen peroxide, at an elevated temperature. Under these conditions, an oxidized LMW-HS product can be formed that retains the structure of USP heparin, particularly comprising the structure of Formula I, but is in the same approximate molecular weight and anticoagulant activity ranges as other LMWH compounds.

[0350] Control of the reaction conditions has led to the production of oxidized LMW-HS compositions that have different anticoagulant activities and molecular weight properties relative to each other, and described, for example, in U.S. Pat Nos. 4,281,108, 4,629,699, and 4,791,195, as well as European Patent EP0101141, the disclosures of which are incorporated by reference in their entireties. In particular, the acidified heparin has been formed by reacting the USP heparin with a strong acid, such as hydrochloric acid, or a weak acid, such as ascorbic acid. Acidified heparin has also been formed by binding USP heparin to a strong cationic exchange resin. Similarly, the depolymerization conditions can be controlled with respect to the pH and temperature at which the depolymerization takes place, and the oxidizing agent itself.

[0351] When the starting material is NS2S6S3S-HS, the oxidized LMW-HS product can contain polysaccharides having 3-*O* sulfated glucosamine residues and comprising the structure of Formula I. Accordingly, several of these oxidized LMW-HS product compositions have anticoagulant activity. On the other hand, when the starting material is NS6S/HS or NS2S6S/HS, the resulting oxidized LMW-HS product does not have any 3-*O* sulfated glucosamine residues or anticoagulant activity.

[0352] Non-limiting examples of oxidized LMW-HS compositions that have been prescribed for clinical use include parnaparin (CAS No: 91449-79-5; ATC code: B01AB05) and ardeparin (CAS No: 9005-49-6). In particular, Parnaparin has been used in the prevention of venous thromboembolism, in the treatment of chronic venous disorders, and in the treatment of venous and arterial thrombosis (*see e.g.* Camporese, G., et al., (2009) *Vascular Health and Risk Management* 5:819-831). Without being limited by a particular theory, it is believed that parnaparin is produced by forming the acidified heparin using ascorbic acid, and subsequently depolymerizing the acidified heparin under slightly basic conditions in the presence of cupric acetate monohydrate and hydrogen peroxide with incubation at 50 °C (*see* U.S. Pat. No. 4,791,195, Example 1). Parnaparin that has been administered to patients has an $\bar{M}_w$ in a range of at least 4,000 Da, up to 6,000 Da, and typically 5,000 Da, and a size distribution such that the proportion of polysaccharides having an M_r less than 3,000 is not more than 30% of the composition, and the proportion of polysaccharides having an M_r in a range of at least 3,000 and up to 8,000 is between 50% and 60% of the

composition. Additionally, parnaparin compositions can comprise an anti-Xa activity of at least 75 IU mg⁻¹ and not more than 110 IU mg⁻¹, and/or a ratio of anti-Xa activity to anti-IIa activity of at least 1.5:1, and up to 3.0:1 (see “Parnaparin Sodium” (2010) *European Pharmacopoeia* 7.0, 2672). On the other hand, ardeparin compositions that have been prescribed to patients have generally had an $\bar{M}_w$ in a range of at least 5,500 Da, up to 6,500 Da, and typically 6,000 Da, an anti-Xa activity of 120 +/- 25 IU mg⁻¹, and a ratio of anti-Xa activity to anti-IIa activity of at least 2.0:1, up to 2.5:1, and characteristically 2.3:1.

[0353] Accordingly, an oxidized LMW-HS composition can be prepared from an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product by the following steps: (a) synthesizing an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product according to any of the above methods; (b) providing an oxidation reaction mixture comprising an oxidation agent, preferably hydrogen peroxide; and (c) treating the NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product with the oxidation reaction mixture for a time sufficient to depolymerize at least a portion of the NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product, thereby forming the oxidized LMW-HS product. In another embodiment, the step of treating the NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product with the oxidation reaction mixture can comprise the following sub-steps: (i) acidifying the NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product to form an acidified HS product; (ii) combining the acidified HS product with the oxidation reaction mixture; and (c) incubating the acidified HS product within the oxidation reaction mixture at a temperature of at least 50 °C until an oxidized LMW-HS product is formed. In another embodiment, the step of treating the NS2S6S3S-HS product with the oxidation reaction mixture can comprise the procedure of Example 1 of U.S. Patent No. 4,791,195.

[0354] In another embodiment, the time sufficient to form the oxidized LMW-HS product is the time sufficient to cause the product to have a desired average molecular weight. In another embodiment, the $\bar{M}_w$ of the oxidized LMW-HS product is in the range of 2,000 Da to 12,000 Da, preferably in the range of 4,000 Da to 6,500 Da. In another embodiment, the $\bar{M}_w$ of the oxidized LMW-HS product is in the range 4,000 Da to 6,000 Da, preferably 5,000 Da. In a further embodiment, the oxidized LMW-HS product comprises a size distribution such that the proportion of polysaccharides having an M_r less than 3,000 is not more than 30% of the composition, and the proportion of polysaccharides having an M_r in a range of at least 3,000 and up to 8,000 is between 50% and 60% of the composition. In another embodiment, the $\bar{M}_w$ of the oxidized LMW-HS product is in the range 5,500 Da to 6,500 Da, preferably 6,000 Da.

[0355] In another embodiment, an oxidized LMW-HS product formed from NS2S6S3S-HS comprises one or more properties that are identical to parnaparin, including but not limited to chemical structure, molecular weight, size distribution, anticoagulant activity, and/or sulfation content. In another embodiment, the oxidized LMW-HS product formed from NS2S6S3S-HS is substantially identical to parnaparin. In another embodiment, an oxidized LMW-HS product formed from NS2S6S3S-HS, for example, products having one or more properties that are identical to parnaparin or products that are substantially equivalent to parnaparin, is subsequently *O*-desulfated, according to any of the procedures described above, to form an oxidized ODSH product.

[0356] In another embodiment, an oxidized LMW-HS product formed from NS2S6S3S-HS comprises one or more properties that are identical to ardeparin, including but not limited to chemical structure, molecular weight, size distribution, anticoagulant activity, and/or sulfation content. In another embodiment, the oxidized LMW-HS product formed from NS2S6S3S-HS is substantially identical to ardeparin. In another embodiment, an oxidized LMW-HS product formed from NS2S6S3S-HS, for example, products having one or more properties that are identical to ardeparin or products that are substantially equivalent to ardeparin, is subsequently *O*-desulfated, according to any of the procedures described above, to form an oxidized ODSH product.

[0357] Those skilled in the art would appreciate that the examples described above of LMW-HS compositions, and methods for forming them from an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product synthesized using one or more engineered aryl sulfate-dependent sulfotransferase enzymes, are non-exhaustive, and that such other examples are excluded for clarity and brevity. Once an NS/HS, NS6S/HS, NS2S/HS, NS2S6S/HS, or NS2S6S3S-HS product is formed according to any of the methods described above, it can be modified and/or depolymerized by any known process to form a secondary product, particularly an LMW-HS product. Such processes include, but are not limited to: fractionation using solvents (French Patent No. 2,440,376, U.S. Pat. No. 4,692,435); fractionation using an anionic resin (French Patent No. 2,453,875); gel filtration; affinity chromatography (U.S. Pat. No. 4,401,758); controlled depolymerization by means of a chemical agent including, but not limited to, nitrous acid (European Patent EP 0014184, European Patent EP 0037319, European Patent EP 0076279, European Patent EP 0623629, French Patent No. 2,503,714, U.S. Pat. No. 4,804,652 and PCT Publication No. WO 81/03276), β -elimination from a heparin ester (European Patent EP 0040144, U.S. Pat. No. 5,389,618), periodate (European Patent EP 0287477), sodium borohydride (European Patent EP 0347588, European Patent EP 0380943), ascorbic acid (U.S. Pat. No. 4,533,549), hydrogen peroxide (U.S. Pat. No. 4,629,699, U.S. Pat. No. 4,791,195), quaternary ammonium hydroxide from a quaternary ammonium salt of

heparin (U.S. Pat. No. 4,981,955), alkali metal hydroxide (European Patent EP 0380943, European Patent EP 0347588), using heparinase enzymes (European Patent EP 0064452, U.S. Pat. No. 4,396,762, European Patent EP 0244235, European Patent EP 0244236; U.S. Pat. No. 4,826,827; U.S. Pat. No. 3,766,167), by means of irradiation (European Patent EP 0269981), purification and modification of fast-moving HS fractions (US. Pat. No. 7,687,479, U.S. Pat. No. 8,609,632), and other methods or combinations of methods such as those described in U.S. Pat. No. 4,303,651, U.S. Pat. No. 4,757,057, U.S. Publication No. 2007/287683, PCT Publication No. WO 2009/059284 and PCT Publication No. WO 2009/059283, the disclosures of which are incorporated by reference in their entireties. Any of the LMW-HS products formed from NS2S6S3S-HS by any of the above process can subsequently *O*-desulfated, to form an LMW ODSH product.

Preparation of Engineered Aryl Sulfate-Dependent Sulfotransferase Enzymes

[0358] In general, the engineered sulfotransferases encoded by the disclosed nucleic acid and amino acid sequences can be expressed and purified using any microbiological technique known in the art, including as described below. The aryl sulfate-dependent sulfotransferase activity of each purified enzyme can be determined spectrophotometrically or fluorescently and/or using mass spectrometry (MS) or nuclear magnetic resonance (NMR) spectroscopy to characterize the starting materials and/or sulfated polysaccharide products. Methods for isolating, purifying, and assessing the activity of the engineered sulfotransferases described herein are described below in the Examples section, as well as in PCT Publication Nos. WO 2020/150350 and WO 2021/007429 (International Application Nos. PCT/US2020/013677 and PCT/US2020/041404, respectively), the descriptions of which are hereby incorporated by reference in their entireties.

[0359] The engineered gene products, proteins and polypeptides utilized in accordance with methods of the present invention can also include analogs that contain insertions, deletions, or mutations relative to the disclosed DNA or peptide sequences, and that also encode for enzymes that catalyze reactions in which aryl sulfate compounds are substrates. In another embodiment, each analog similarly catalyzes sulfotransfer reactions in which aryl sulfate compounds are utilized as sulfo donors. Analogs can be derived from nucleotide or amino acid sequences as disclosed herein, or they can be designed synthetically *in silico* or *de novo* using computer modeling techniques. Those skilled in the art will appreciate that other analogs, as yet undisclosed or undiscovered, can be used to design and/or construct different sulfate-dependent sulfotransferase enzymes capable of being utilized in accordance with methods of the present invention. There is no need for a gene product, protein, or polypeptide to comprise all or substantially all of a nucleic acid or amino acid sequence

of an engineered sulfotransferase as disclosed herein. Such sequences are herein referred to as “segments.” Further, the gene products, proteins, and polypeptides discussed and disclosed herein can also include fusion or recombinant aryl sulfate-dependent sulfotransferases comprising full-length sequences or biologically functional segments of sequences disclosed in the present invention. Methods of preparing such proteins are known in the art.

[0360] In addition to the nucleic acid and amino acid sequences disclosed herein, methods of the present invention can be practiced by aryl sulfate-dependent sulfotransferases comprising amino acid sequences that are substantially identical to any of the disclosed amino acid sequences above, or expressed from nucleic acids comprising a nucleotide sequence that is substantially identical to a disclosed nucleotide sequence (SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, or SEQ ID NO: 27). Those skilled in the art can determine appropriate nucleotide sequences that encode for polypeptides having the amino acid sequence of SEQ ID NOs: 33-54 and 56-61, based on the nucleotide sequences above. “Substantially identical” sequences, as used in the art, refer to sequences which differ from a particular reference sequence by one or more deletions, substitutions, or additions, the net effect of which is to retain at least some of the biological activity of the engineered polypeptide encoded by the reference sequence. Namely, the biological activity of the engineered aryl sulfate-dependent sulfotransferases comprises the transfer of a sulfo group from a sulfo donor aryl sulfate compound to a polysaccharide acting as a sulfo group acceptor. In another embodiment, the polysaccharide is a heparosan-based and/or HS polysaccharide. Accordingly, as used to describe the aryl sulfate-dependent enzymes of the present invention, “substantial identity” can refer either to identity with a particular gene product, polypeptide or amino acid sequence of an aryl sulfate-dependent enzyme, or a gene or nucleic acid sequence encoding for an aryl sulfate-dependent enzyme. Such sequences can include mutations of the disclosed sequences or a sequence in which the biological activity is altered, enhanced, or diminished to some degree but retains at least some of the original biological activity of a disclosed reference amino acid sequence or polypeptide encoded by a disclosed reference nucleic acid sequence.

[0361] Alternatively, DNA analog sequences are substantially identical to the specific DNA sequences disclosed herein if: (a) the DNA analog sequence is derived from coding regions of the any of the disclosed nucleic acid sequences; or (b) the DNA analog sequence is capable of hybridization of DNA sequences of (a) under stringent conditions and which encode biologically active aryl sulfate-dependent sulfotransferase gene product; or (c) the DNA sequences are

degenerate as a result of alternative genetic code to the DNA analog sequences defined in (a) and/or (b). Substantially identical analog proteins will be greater than about 60% identical to the corresponding sequence of the native protein. Sequences having lesser degrees of identity but comparable biological activity, namely, transferring a sulfo group from an aryl sulfate compound to polysaccharides, particularly heparosan-based or HS polysaccharides, are also considered to be substantially identical. In determining the substantial identity of nucleic acid sequences, all subject nucleic acid sequences capable of encoding substantially identical amino acid sequences are considered to be substantially identical to a reference nucleic acid sequence, regardless of differences in codon sequences or amino acid substitutions to create biologically functional equivalents.

[0362] At a biological level, identity is just that, i.e. the same amino acid at the same relative position in a given family member of a gene family. Homology and similarity are generally viewed as broader terms. For example, biochemically similar amino acids, for example leucine and isoleucine or glutamic acid/aspartic acid, can be alternatively present at the same position—these are not identical per se, but are biochemically “similar.” As disclosed herein, these are referred to as conservative differences or conservative substitutions. This differs from a conservative mutation at the DNA level, which changes the nucleotide sequence without making a change in the encoded amino acid, e.g., TCC to TCA, both of which encode serine.

[0363] In some embodiments, the genes and gene products include within their respective sequences a sequence “essentially as that” of a gene encoding for an aryl sulfate-dependent sulfotransferase or its corresponding protein. A sequence essentially as that of a gene encoding for an aryl sulfate-dependent sulfotransferase refers to sequences that are substantially identical or substantially similar to a portion of a disclosed nucleic acid sequence and contains a minority of bases or amino acids (whether DNA or protein) that are not identical to those of a disclosed protein or a gene, or which are not a biologically functional equivalent. Biological functional equivalence is well understood in the art and is further discussed in detail below. Nucleotide sequences are “essentially the same” where they have between about 75% and about 85%, or particularly, between about 86% and about 90%, or more particularly greater than 90%, or even more particularly between about 91% and about 95%, or still more particularly, between about 96% and about 99%, of nucleic acid residues which are identical to the nucleotide sequence of a disclosed gene. Similarly, peptide sequences which have about 80%, or 90%, or particularly from 90-95%, or more particularly greater than 96%, or even more particularly 95-98%, or still more particularly 99% or greater amino acids which are identical

or functionally equivalent or biologically functionally equivalent to the amino acids of a disclosed polypeptide sequence will be sequences which are “essentially the same.”

[0364] Additionally, alternate nucleic acid sequences that include functionally equivalent codons are also encompassed by this invention. Functionally equivalent codons refer to codons that encode the same amino acid, such as the ACG and AGU codons for serine. Thus, substitution of a functionally equivalent codon into any of the nucleotide sequences above encode for biologically functionally equivalent sulfotransferases. Thus, the present invention includes amino acid and nucleic acid sequences comprising such substitutions but which are not set forth herein in their entirety for convenience.

[0365] Those skilled in the art would recognize that amino acid and nucleic acid sequences can include additional residues, such as additional N- or C-terminal amino acids or 5' or 3' nucleic acid sequences, and yet still be essentially as set forth in one of the sequences disclosed herein, so long as the sequence retains its biological activity with respect to binding and reacting with aryl sulfate compounds as sulfo donors. The addition of terminal sequences particularly applies to nucleic acid sequences which can, for example, include various non-coding sequences flanking either of the 5' or 3' portions of the coding region or can include various internal sequences, or introns, which are known to occur within genes.

[0366] As discussed above, modifications and changes can be made in the sequence of any of the disclosed aryl sulfate-dependent sulfotransferases, including conservative and non-conserved mutations, deletions, and additions while still constituting a molecule having like or otherwise desirable characteristics. For example, certain amino acids can be substituted for other amino acids in a protein structure without appreciable loss of interactive capacity with particular structures or compounds, particularly aryl sulfate compounds and/or sulfo acceptor polysaccharides. This can occur because the ability of a protein to recognize, bind, and react with other structures or compounds within its environment defines that protein's biological functional activity, not the sequence itself. Consequently, certain amino acid sequence substitutions can be made in that protein's sequence to obtain a protein with the equal, enhanced, or diminished properties. One non-limiting example of such amino acid substitutions that can occur without an appreciable loss of interactive activity include substitutions in external domains or surfaces of the protein that do not affect the folding and solubility of the protein. Similarly, amino acids can potentially be added to either terminus of the protein so long as the ability of the protein to fold or to recognize and bind its substrates is not deleteriously affected. One skilled in the art can appreciate that several other methods and/or strategies can be utilized to alter an enzyme's sequence without affecting its activity.

[0367] Consequently, mutations, deletions, additions, or other alterations to a parent enzyme's structure or sequence in which the modified enzyme retains the parent enzyme's biological activity can be defined to be biologically functionally equivalent to the parent enzyme. Thus, biologically functional equivalent enzymes, with respect to the engineered aryl sulfate-dependent sulfotransferases, can include any substitution or modification of any of the amino acid sequences disclosed herein, so long as the resultant modified enzyme is dependent on interacting with aryl sulfate compounds, particularly PNS or NCS, to catalyze sulfo transfer to polysaccharides, particularly heparosan-based and/or HS polysaccharides. In particular, such substitutions or modifications can result from conservative mutations in the amino acid sequence in any portion of the protein, as described below, although non-conservative mutations in non-catalytically active regions of the enzyme are also contemplated. Consequently, engineered aryl sulfate-dependent sulfotransferases suitable to practice the methods of the present invention can be expressed from any nucleic acid having a nucleotide sequence that encodes for a biologically functional equivalent enzyme, although such nucleotide sequences are not set forth herein in their entirety for convenience.

[0368] Alternatively, recombinant DNA technology can be used to create biologically functionally equivalent proteins or peptides in which changes in the protein structure can be engineered, based on considerations of the properties of the amino acids being exchanged. Rationally-designed changes can be introduced through the application of site-directed mutagenesis techniques, for example, to test whether certain mutations affect positively or negatively affect the enzyme's aryl sulfate-dependent catalytic activity or binding of sulfo donors or acceptors within the enzyme's active site.

[0369] Amino acid substitutions, such as those which might be employed in modifying any of the aryl sulfate-dependent sulfotransferases described herein, are generally based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Those skilled in the art are familiar with the similarities between certain amino acids, such as the size, shape and type of the amino acid side-chain substituents. Non-limiting examples include relationships such as that arginine, lysine and histidine are all positively charged residues; that alanine, glycine and serine are all of similar size; and that phenylalanine, tryptophan and tyrosine all have a generally similar shape. Consequently, the amino acids that comprise the following groups—arginine, lysine and histidine; alanine, glycine and serine; and phenylalanine, tryptophan and tyrosine—are defined herein as biologically functional equivalents to the other amino acids in the same group. Other biologically functionally equivalent changes will be appreciated by those of skill in the art.

[0370] In another embodiment, the present invention provides isolated nucleic acids encoding functional fragments of the engineered enzymes of the present invention, or mutants thereof, in which conservative substitutions have been made for particular residues within the amino acid sequence of any of the engineered sulfotransferase enzymes described herein.

[0371] Additionally, isolated nucleic acids used to express aryl sulfate-dependent sulfotransferases capable of practicing the methods of the present invention may be joined to other nucleic acid sequences for use in various applications. Thus, for example, the isolated nucleic acids may be ligated into cloning or expression vectors, as are commonly known in the art and as described in the examples below. Additionally, nucleic acids may be joined in-frame to sequences encoding another polypeptide so as to form a fusion protein, as is commonly known in the art. Fusion proteins can comprise a coding region for the aryl sulfate-dependent sulfotransferase that is aligned within the same expression unit with other proteins or peptides having desired functions, such as for solubility, purification, or immunodetection. Thus, in another embodiment, cloning, expression and fusion vectors comprising any of the above-described nucleic acids, that encode for an aryl sulfate-dependent sulfotransferase that can be utilized in with methods of the present invention are also provided.

[0372] Furthermore, nucleic acid segments of the present invention, regardless of the length of the coding sequence itself, can be combined with other DNA sequences, such as promoters, enhancers, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length can vary considerably. Those skilled in the art would recognize that a nucleic acid fragment of almost any length can be employed, with the total length typically being limited by the ease of preparation and use in the intended recombinant DNA protocol.

[0373] In particular, recombinant vectors in which the coding portion of the gene or DNA segment is positioned under the control of a promoter are especially useful. In some embodiments, the coding DNA segment can be associated with promoters isolated from bacterial, viral, eukaryotic, or mammalian cells. Promoters specific to the cell type chosen for expression are often the most effective. The use of promoter and cell type combinations for protein expression is generally known to those of skill in the art of molecular biology (See, e.g., Sambrook et al. (2012) *Molecular Cloning: A Laboratory Manual*, Fourth Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., incorporated by reference in its entirety). The promoters employed can be constitutive or inducible and can be used under the appropriate conditions to direct high-level expression of the introduced DNA segment, such as is advantageous in the large-scale production of recombinant

proteins or peptides. Appropriate promoter systems that are often effective for high-level expression include, but are not limited to, the vaccinia virus promoter, the baculovirus promoter, and the Ptac promoter.

[0374] Thus, in some embodiments, an expression vector can be utilized that comprises a nucleotide sequence encoding for a biologically-active, aryl sulfate-dependent sulfotransferase suitable for use with methods of the present invention. In one example, an expression vector can comprise any nucleotide sequence that encodes for an aryl sulfate-dependent sulfotransferase gene product. In further embodiments, an expression vector comprises a nucleic acid comprising any of the nucleotide sequences described above, or any nucleotide sequence that encodes for a polypeptide comprising the amino acid sequence of any of the engineered sulfotransferase enzymes described above. In even further embodiments, any nucleic acid sequence encoding for an engineered aryl sulfate-dependent sulfotransferase enzyme of the present invention can be codon-optimized based on the expression host used to produce the enzyme. The preparation of recombinant vectors and codon optimization are well known to those of skill in the art and described in many references, such as, for example, Sambrook et al. (2012) *Molecular Cloning: A Laboratory Manual*, Fourth Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

[0375] Those skilled in the art would recognize that the DNA coding sequences to be expressed, in this case those encoding the aryl sulfate-dependent sulfotransferase gene products, are positioned in a vector adjacent to and under the control of a promoter. As is known in the art, a promoter is a region of a DNA molecule typically within about 100 nucleotide pairs upstream of (i.e., 5' to) the point at which transcription begins (i.e., a transcription start site). That region typically contains several types of DNA sequence elements that are located in similar relative positions in different genes. It is understood in the art that to bring a coding sequence under the control of such a promoter, one generally positions the 5' end of the transcription initiation site of the transcriptional reading frame of the gene product to be expressed between about 1 and about 50 nucleotides "downstream" of (i.e., 3' of) the chosen promoter.

[0376] One can also desire to incorporate into the transcriptional unit of the vector an appropriate polyadenylation site (e.g., 5'-AATAAA-3'), if one was not contained within the original inserted DNA. Typically, poly-A addition sites are placed about 30 to 2000 nucleotides "downstream" of the coding sequence at a position prior to transcription termination.

[0377] Another type of discrete transcription regulatory sequence element is an enhancer. An enhancer imposes specificity of time, location and expression level on a particular coding region or gene. A major function of an enhancer is to increase the level of transcription of a coding sequence

in a cell that contains one or more transcription factors that bind to that enhancer. An enhancer can function when located at variable distances from transcription start sites so long as a promoter is present.

[0378] Optionally, an expression vector of the invention comprises a polynucleotide operatively linked to an enhancer-promoter. As used herein, the phrase “enhancer-promoter” means a composite unit that contains both enhancer and promoter elements. For example, an expression vector can comprise a polynucleotide operatively linked to an enhancer-promoter that is a eukaryotic promoter and the expression vector further comprises a polyadenylation signal that is positioned 3’ of the carboxy-terminal amino acid and within a transcriptional unit of the encoded polypeptide. As used herein, the phrase “operatively linked” means that an enhancer-promoter is connected to a coding sequence in such a way that the transcription of that coding sequence is controlled and regulated by that enhancer-promoter. Techniques for operatively linking an enhancer-promoter to a coding sequence are well known in the art; the precise orientation and location relative to a coding sequence of interest is dependent, *inter alia*, upon the specific nature of the enhancer-promoter.

[0379] An enhancer-promoter used in a vector construct of the present invention can be any enhancer-promoter that drives expression in a cell to be transfected. By employing an enhancer-promoter with well-known properties, the level and pattern of gene product expression can be optimized.

[0380] Sulfotransferase enzymes suitable to practice the methods of the present invention can be expressed within cells or cell lines, either prokaryotic or eukaryotic, into which have been introduced the nucleic acids of the present invention so as to cause clonal propagation of those nucleic acids and/or expression of the proteins or peptides encoded thereby. Such cells or cell lines are useful for propagating and producing nucleic acids, as well as for producing the aryl sulfate-dependent sulfotransferases themselves. As used herein, the term “transformed cell” is intended to embrace any cell, or the descendant of any cell, into which has been introduced any of the nucleic acids of the invention, whether by transformation, transfection, transduction, infection, or other means. Methods of producing appropriate vectors, transforming cells with those vectors, and identifying transformants are well known in the art. (See, e.g., Sambrook et al. (2012) *Molecular Cloning: A Laboratory Manual*, Fourth Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.)

[0381] Prokaryotic cells useful for producing transformed cells include members of the bacterial genera *Escherichia* (e.g., *E. coli*), *Pseudomonas* (e.g., *P. aeruginosa*), and *Bacillus* (e.g., *B. subtilis*, *B. stearothermophilus*), as well as many others well known and frequently used in the art.

Prokaryotic cells are particularly useful for the production of large quantities of the proteins or peptides (e.g., aryl sulfate-dependent enzymes, fragments of those sequences thereof, or fusion proteins including those sequences). Bacterial cells (e.g., *E. coli*) may be used with a variety of expression vector systems including, for example, plasmids with the T7 RNA polymerase/promoter system, bacteriophage λ regulatory sequences, or M13 Phage regulatory elements. Bacterial hosts may also be transformed with fusion protein vectors that create, for example, Protein A, lacZ, trpE, maltose-binding protein (MBP), small ubiquitin-related modifier (SUMO), poly-His tag, or glutathione-S-transferase (GST) fusion proteins. All of these, as well as many other prokaryotic expression systems, are well known in the art and widely available commercially (e.g., pGEX-27 (Amrad, USA) for GST fusions).

[0382] In some embodiments of the invention, expression vectors comprising any of the nucleotide sequences described above can also comprise genes or nucleic acid sequences encoding for fusion proteins with any aryl sulfate-dependent sulfotransferase. In further embodiments, expression vectors can additionally include the *malE* gene, which encodes for the maltose binding protein. Upon inducing protein expression from such expression vectors, the expressed gene product comprises a fusion protein that includes maltose binding protein and any of the aryl sulfate-dependent sulfotransferase enzymes described above. In other further embodiments, an expression vector that includes any of the above nucleic acids that encode for any of the above aryl sulfate-dependent sulfotransferase enzymes can additionally include a gene encoding for a SUMO modifier, such as, in a non-limiting example, SUMO-1.

[0383] In other embodiments, expression vectors according to the present invention can additionally include a nucleic acid sequence encoding for a poly-His tag. Upon inducing protein expression from such expression vectors, the expressed gene product comprises a fusion protein that includes the poly-His tag and any of the aryl sulfate-dependent sulfotransferase enzymes described above. In a further embodiment, expression vectors can include both a nucleic acid sequence encoding for a poly-His tag and the *malE* gene or a SUMO gene, from which a fusion protein can be expressed that includes a poly-His tag, MBP, or SUMO, along with any aryl sulfate-dependent sulfotransferase enzyme.

[0384] Eukaryotic cells and cell lines useful for producing transformed cells include mammalian cells (e.g., endothelial cells, mast cells, COS cells, CHO cells, fibroblasts, hybridomas, oocytes, embryonic stem cells), insect cells lines (e.g., *Drosophila* Schneider cells), yeast, and fungi. Non-limiting examples of such cells include, but are not limited to, COS-7 cells, CHO cells, murine primary cardiac microvascular endothelial cells (CME), murine mast cell line C57.1, human primary

endothelial cells of umbilical vein (HUVEC), F9 embryonal carcinoma cells, rat fat pad endothelial cells (RFPEC), and L cells (e.g., murine LTA tk⁻ cells).

[0385] Vectors may be introduced into the recipient or “host” cells by various methods well known in the art including, but not limited to, calcium phosphate transfection, strontium phosphate transfection, DEAE dextran transfection, electroporation, lipofection, microinjection, ballistic insertion on micro-beads, protoplast fusion or, for viral or phage vectors, by infection with the recombinant virus or phage.

[0386] In another embodiment, the present invention provides aryl sulfate-dependent sulfotransferase variants in which conservative or non-conservative substitutions have been made for certain residues within any of the engineered sulfotransferase amino acid sequences disclosed above. Conservative or non-conservative substitutions can be made at any point in the amino acid sequence, including residues that surround the active site or are involved in catalysis, provided that the enzyme retains measurable catalytic activity; namely, the transfer of a sulfo group from an aryl sulfate compound to a polysaccharide, particularly a heparosan-based and/or HS polysaccharide. In other embodiments, the aryl sulfate compound is PNS. In still other embodiments, the aryl sulfate compound is NCS.

[0387] In another embodiment, the aryl sulfate-dependent sulfotransferase enzymes have at least 50%, including at least 60%, 70%, 80%, 85%, 90% or 95% up to at least 99% amino acid sequence identity to the amino acid sequence of any of the engineered sulfotransferase enzymes disclosed above, including disclosed as SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, and SEQ ID NOs: 33-54 and 56-61, while retaining its catalytic activity of transfer of a sulfo group from an aryl sulfate compound to a polysaccharide, particularly a heparosan-based and/or HS polysaccharide. Such sequences may be routinely produced by those of ordinary skill in the art, and sulfotransferase activity may be tested by routine methods such as those disclosed herein.

[0388] Further, and in another embodiment, the amino acid sequence(s) of any of the engineered aryl sulfate-dependent sulfotransferases utilized in accordance with any of the methods described herein can be characterized as a percent identity relative to a natural sulfotransferase that catalyzes the same reaction using PAPS as the sulfo donor, so long as the sulfotransferase has aryl sulfate-dependent activity. For example, and in another embodiment, an engineered aryl sulfate-dependent glucosaminyl *N*-sulfotransferase that can be utilized in accordance with any of the methods of the present invention can comprise an amino acid sequence that has at least 50%, including at least 55%,

60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, up to at least 97% sequence identity with the amino acid sequence of the *N*-sulfotransferase domain of any of the natural enzymes within the EC 2.8.2.8 enzyme class, including biological functional fragments thereof. In a further embodiment, the engineered aryl sulfate-dependent glucosaminyl *N*-sulfotransferase can comprise at least 50%, including at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, up to at least 97% sequence identity with the amino acid sequence of the *N*-sulfotransferase domain of the natural human glucosaminyl *NDST* enzyme (entry sp|P52848|NDST_1_HUMAN, in Figure 3, above).

[0389] In another embodiment, an engineered aryl sulfate-dependent 2OST that can be utilized in accordance with any of the methods of the present invention can comprise an amino acid sequence that has at least 50%, including at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, up to at least 97% sequence identity with the amino acid sequence of any of the natural 2OST enzymes within the EC 2.8.2.- enzyme class, including biological functional fragments thereof. In a further embodiment, the engineered aryl sulfate-dependent 2OST can comprise at least 50%, including at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, up to at least 97% sequence identity with the amino acid sequence of the natural chicken 2OST enzyme (entry sp|Q76KB1|HS2ST_CHICK, in Figure 14, above).

[0390] In another embodiment, an engineered aryl sulfate-dependent 6OST that can be utilized in accordance with any of the methods of the present invention can comprise an amino acid sequence that has at least 50%, including at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, up to at least 97% sequence identity with the amino acid sequence of any of the natural 6OST enzymes within the EC 2.8.2.- enzyme class, including biological functional fragments thereof. In a further embodiment, the engineered aryl sulfate-dependent 6OST can comprise at least 50%, including at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, up to at least 97% sequence identity with the amino acid sequence of the first isoform of the mouse 6OST (UniProtKB Accession No. Q9QYK5). In a further embodiment, the engineered aryl sulfate-dependent 6OST can comprise at least 50%, including at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, up to at least 97% sequence identity with residues 67-377 of the amino acid sequence of the first isoform of the mouse 6OST (entry Q9QYK5|H6ST1_MOUSE, in Figure 18, above).

[0391] In another embodiment, an engineered aryl sulfate-dependent 3OST that can be utilized in accordance with any of the methods of the present invention can comprise an amino acid sequence that has at least 50%, including at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, up to at least 97% sequence identity with the amino acid sequence of any of the natural enzymes within the EC 2.8.2.23 enzyme class, including biological functional fragments thereof. In a further

embodiment, the engineered aryl sulfate-dependent 3OST can comprise at least 50%, including at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, up to at least 97% sequence identity with residues 48-311 of the amino acid sequence of the first isoform of the natural human 3OST (UniProtKB Accession No. O14792).

[0392] Substantially pure aryl sulfate-dependent sulfotransferases may be joined to other polypeptide sequences for use in various applications. Thus, for example, engineered sulfotransferases may be joined to one or more additional polypeptides so as to form a fusion protein, as is commonly known in the art. The additional polypeptides may be joined to the N-terminus, C-terminus or both termini of the aryl sulfate-dependent sulfotransferase enzyme. Such fusion proteins may be particularly useful if the additional polypeptide sequences are easily identified (e.g., by providing an antigenic determinant), are easily purified (e.g., by providing a ligand for affinity purification), or enhance the solubility of the aryl sulfate-dependent sulfotransferase enzyme in solution.

[0393] In another embodiment, substantially pure proteins may comprise only a portion or fragment of the amino acid sequence of a complete aryl sulfate-dependent sulfotransferase. In some instances, it may be preferable to employ a minimal fragment retaining aryl sulfate-dependent sulfotransferase activity, particularly if the minimal fragment enhances the solubility or reactivity of the enzyme. Thus, in some embodiments, methods of the present invention can be practiced using substantially pure aryl sulfate-dependent sulfotransferases of any length, including full-length forms, or minimal functional fragments thereof. Additionally, these proteins may also comprise conservative or non-conservative substitution variants as described above.

[0394] In some embodiments, the present invention provides substantially pure preparations of aryl sulfate-dependent sulfotransferases, including those comprising any of the amino acid sequences disclosed above. The engineered sulfotransferases may be substantially purified by any of a variety of methods selected on the basis of the properties revealed by their protein sequences. Typically, the aryl sulfate-dependent sulfotransferases, fusion proteins, or fragments thereof, can be purified from cells transformed or transfected with expression vectors, as described above. Insect, yeast, eukaryotic, or prokaryotic expression systems can be used, and are well known in the art. In the event that the protein or fragment localizes within microsomes derived from the Golgi apparatus, endoplasmic reticulum, or other membrane-containing structures of such cells, the protein may be purified from the appropriate cell fraction. Alternatively, if the protein does not localize within these structures, or aggregates in inclusion bodies within the recombinant cells (e.g., prokaryotic cells), the

protein may be purified from whole lysed cells or from solubilized inclusion bodies by standard means.

[0395] Purification can be achieved using standard protein purification procedures including, but not limited to, affinity chromatography, gel-filtration chromatography, ion-exchange chromatography, high-performance liquid chromatography (RP-HPLC, ion-exchange HPLC, size-exclusion HPLC), high-performance chromatofocusing chromatography, hydrophobic interaction chromatography, immunoprecipitation, or immunoaffinity purification. Gel electrophoresis (e.g., PAGE, SDS-PAGE) can also be used to isolate a protein or peptide based on its molecular weight, charge properties and hydrophobicity.

[0396] An aryl sulfate-dependent sulfotransferase, or a fragment thereof, may also be conveniently purified by creating a fusion protein including the desired sequence fused to another peptide such as an antigenic determinant, a poly-histidine tag (e.g., QIAexpress vectors, QIAGEN Corp., Chatsworth, CA), or a larger protein (e.g., GST using the pGEX-27 vector (Amrad, USA), green fluorescent protein using the Green Lantern vector (GIBCO/BRL, Gaithersburg, MD), maltose binding protein using the pMAL vector (New England Biolabs, Ipswich, MA), or a SUMO protein. The fusion protein may be expressed and recovered from prokaryotic or eukaryotic cells and purified by any standard method based upon the fusion vector sequence. For example, the fusion protein may be purified by immunoaffinity or immunoprecipitation with an antibody to the non-aryl sulfate-dependent sulfotransferase portion of the fusion or, in the case of a poly-His tag, by affinity binding to a nickel column. The desired aryl sulfate-dependent sulfotransferase protein or fragment can then be further purified from the fusion protein by enzymatic cleavage of the fusion protein. Methods for preparing and using such fusion constructs for the purification of proteins are well known in the art and numerous kits are now commercially available for this purpose.

[0397] Furthermore, in some embodiments, isolated nucleic acids encoding for any aryl sulfate-dependent sulfotransferase may be used to transform host cells. The resulting proteins may then be substantially purified by well-known methods including, but not limited to, those described in the examples below. Alternatively, isolated nucleic acids may be utilized in cell-free in vitro translation systems. Such systems are also well known in the art.

[0398] While particular embodiments of the invention have been described, the invention can be further modified within the spirit and scope of this disclosure. Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, numerous equivalents to the specific procedures, embodiments, claims, and examples described herein. As such, such equivalents are considered to be within the scope of the invention, and this application is therefore

intended to cover any variations, uses or adaptations of the invention using its general principles. Further, the invention is intended to cover such departures from the present disclosure as come within known or customary practice in the art to which this invention pertains and which fall within the appended claims.

[0399] It is appreciated that certain features of the invention, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. Conversely, various features of the invention, which are, for brevity, described in the context of a single embodiment, may also be provided separately or in any suitable sub-combination or as suitable in any other described embodiment of the invention. Certain features described in the context of various embodiments are not to be considered essential features of those embodiments, unless the embodiment is inoperative without those elements.

[0400] The contents of all references, patents, and patent applications mentioned in this specification are hereby incorporated by reference, and shall not be construed as an admission that such reference is available as prior art to the present invention. All of the incorporated publications and patent applications in this specification are indicative of the level of ordinary skill in the art to which this invention pertains, and are incorporated to the same extent as if each individual publication or patent application was specifically indicated and individually indicated by reference.

[0401] The invention is further illustrated by the following working and prophetic examples, neither of which should be construed as limiting the invention. Additionally, to the extent that section headings are used, they should not be construed as necessarily limiting. Any use of the past tense to describe an example otherwise indicated as constructive or prophetic is not intended to reflect that the constructive or prophetic example has actually been carried out.

EXAMPLES

[0402] The following working and prophetic examples illustrate the embodiments of the invention that are presently best known. However, it is to be understood that the following are only exemplary or illustrative of the application of the principles of the present invention. Numerous modifications and alternative compositions, methods, and systems may be devised by those skilled in the art without departing from the spirit and scope of the present invention. Thus, while the present invention has been described above with particularity, the following examples provide further detail in connection with what are presently deemed to be the most practical and preferred embodiments of the invention.

Example 1: Cloning, Expression, and Purification of the Engineered Aryl Sulfate-Dependent Sulfotransferases

[0403] A study was conducted in accordance with embodiments of the present disclosure to determine whether genes according to the present invention could be transformed into host cells capable of overexpressing engineered aryl sulfate-dependent sulfotransferases. After expression, each aryl sulfate-dependent enzyme was isolated and purified from the host cell.

[0404] Generally, DNA coding for genes of any sequence can be synthesized *de novo* by methods commonly known in the art, including but not limited to oligonucleotide synthesis and annealing. Alternatively, DNA can be synthesized commercially and purchased from any one of several laboratories that regularly synthesize genes of a given sequence, including but not limited to ThermoFisher Scientific, GenScript, DNA 2.0, or OriGene. Persons skilled in the art would appreciate that there are several companies that provide the same services, and that the list provided above is merely a small sample of them. Genes of interest can be synthesized independently and subsequently inserted into a bacterial or other expression vector using conventional molecular biology techniques, or the genes can be synthesized concurrently with the DNA comprising the expression vector itself. Similar to genes of interest, suitable expression vectors can also be synthesized or obtained commercially. Often, bacterial expression vectors include genes that confer selective antibiotic resistance to the host cell, as well as genes that permit the cell to overproduce the protein of interest in response to the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG). Bacterial production of proteins of interest using IPTG to induce protein expression is widely known in the art.

[0405] As described above, expression vectors can also include genes that enable production of fusion proteins that include the desired protein that is co-expressed with an additional, known protein to aid in protein folding and solubility. Non-limiting examples of fusion proteins that are commonly produced and are well-known in the art include fusions with MBP, SUMO, or green fluorescent protein. In particular, MBP fusion proteins facilitate easier purification because MBP possesses high affinity for amylose-based resins used in some affinity chromatography columns, while SUMO fusion proteins can include a poly-histidine tag that enables affinity purification on columns with Ni^{2+} -based resins as a stationary phase. Often, fusion proteins between the protein of interest and MBP and/or SUMO can optionally include an amino acid linking sequence that connects the two proteins. Non-limiting examples of commercial expression vectors that can be purchased to produce MBP fusion proteins include the pMAL-c5ETM and pMAL-c5XTM vectors, which can be

obtained from New England Biolabs. Similarly, and in another non-limiting example, commercial expression vectors can also be purchased to produce SUMO fusion proteins, such as the pE-SUMOpro AMP vector, available from LifeSensors, Inc. Once the fusion proteins are produced and isolated, proteases can be utilized to cleave the fused protein and any associated linker sequences from the sulfotransferase, if cleavage is necessary for activity.

[0406] Additionally, expression vectors can also include DNA coding for a poly-histidine tag that can be synthesized at either the N- or C-terminus of the protein of interest. As with MBP fusions, proteins that include a poly-histidine tag simplify the enzyme purification because the tag has a high affinity for Ni²⁺ resins that are utilized in many purification columns. Additionally, poly-histidine tags can optionally be cleaved after purification if it is necessary for optimal activity of the enzyme. A non-limiting example of an expression vector encoding for a C-terminal poly-histidine tag is the pET21b vector, available from Novagen. Another non-limiting example of an expression vector encoding for a poly-histidine tag is the pE-SUMO vector, which encodes for a poly-histidine tag at the N-terminus of the SUMO protein.

[0407] In the present example, double-stranded DNA fragments comprising the nucleotide sequences of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, or SEQ ID NO: 27, encoding for engineered aryl sulfate-dependent sulfotransferases comprising the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26 or SEQ ID NO: 28, respectively, were synthesized using Integrated DNA Technologies' (IDT) gBlocks® Gene Fragments synthesis service. Polymerase chain reactions (PCR) were initiated to generate copies of each double-stranded DNA fragment, using forward and reverse primers comprising appropriate restriction enzyme recognition sequences to facilitate insertion into an expression vector. Genes encoding for the engineered glucosaminyl *N*-sulfotransferase enzymes (SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11) and 3OST enzymes (SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27) contained NdeI and BamHI restriction enzyme recognition sequences, and were ligated into the pMAL-c5x expression vector using quick ligation kits provided by NEB. Expression vectors were then transformed into competent DH5- α *E. coli* cells. Single clones were incubated in LB medium with 100 μ L/mL ampicillin. Nucleotide sequences of each gene and expression vector within the transformed host cells were confirmed by commercial DNA sequencing (GeneWiz).

[0408] Protein expression of the glucosaminyl *N*- and 3-*O* sulfotransferase enzymes was achieved by first transforming confirmed DNA constructs into competent SHuffle® T7 Express lysY *E. coli* cells. Protein expression of the glucosaminyl *N*- and 3-*O* sulfotransferase enzymes has also been achieved by transforming confirmed DNA constructs into competent BL21 (DE3) *E. coli* cells. From either construct, resultant colonies were used to inoculate 250 mL cultures in LB medium, which were allowed to shake and incubate at 32 °C until an optical density at 600 nm (OD 600) of approximately 0.4 to 0.6 was observed. Expression was induced by the addition of 100 µM IPTG to each culture at 18 °C.

[0409] Upon incubation at 18 °C overnight, expressed cells were harvested by centrifuging at 3,620 g and resuspending the pellet in 10 mL of resuspension buffer (25 mM Tris-HCl, pH 7.5; 0.15 M NaCl; 0.2 mg/mL lysozyme; 10 µg/mL DNase I; 5 mM MgCl₂; and 0.1% (w/v) Triton-X 100). Resuspended cells were lysed upon sonication on ice for three pulses of 10 seconds each, and subsequently passed through a 0.45-µm syringe filter. The resulting supernatant was loaded into a 5-mL spin column (G-biosciences) comprising Dextrin Sepharose® resin (GE Biosciences) suspended in a binding buffer comprising 25 mM Tris-HCl, pH 7.5 and 0.15 M NaCl. Enzymes of interest were eluted from the column upon adding an elution buffer comprising 25 mM Tris-HCl, pH 7.5; 0.15 M NaCl; and 40 mM maltose.

[0410] On the other hand, genes encoding for the engineered 2OST (SEQ ID NO: 13, SEQ ID NO: 15) and 6OST enzymes (SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21) contained BsaI and XbaI restriction enzyme recognition sequences, and were ligated into the pE-SUMO vector (LifeSensors, Inc.). Expression vectors were then transformed into competent BL21-DE3 *E. coli* cells. Single clones were incubated in Terrific Broth with 100 µg/mL ampicillin. Nucleotide sequences of each gene and expression vector within the transformed host cells were confirmed by commercial DNA sequencing (GeneWiz).

[0411] Protein expression of the engineered 2OSTs and 6OSTs was achieved by inoculating 500 mL cultures in Terrific Broth with ampicillin and allowing the cultures to incubate with shaking at 35 °C until an OD 600 of approximately 0.6-0.8 was reached. Protein expression was induced by the addition of 0.2 mM IPTG at 18 °C. Cultures were then allowed to incubate at 18 °C overnight, and were subsequently lysed and filtered using an identical procedure to the glucosaminyl *N*- and 3-*O* sulfotransferase enzymes above. The 2OST and 6OST enzymes were subsequently purified in a 5-mL spin column (G-biosciences) comprising HisPur Ni-NTA resin (ThermoFisher) suspended in a binding buffer comprising 25 mM Tris-HCl, pH 7.5, 0.15 M NaCl, 5 mM MgCl₂, and 30 mM

imidazole. Enzymes of interest were eluted from the column upon adding an elution buffer comprising 25 mM Tris-HCl, pH 7.5, 0.15 M NaCl, 5 mM MgCl₂, and 300 mM imidazole.

Example 2: Mass Spectrometric Characterization of the *N*-Sulfated Polysaccharide Products of Engineered Aryl Sulfate-Dependent Glucosaminyl *N*-Sulfotransferase Enzymes

[0412] A study was conducted in accordance with embodiments of the present disclosure to confirm glucosaminyl *N*-sulfotransferase activity of enzymes comprising the amino acid sequence of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, or SEQ ID NO: 12 by detecting the presence of *N*-sulfated polysaccharide products formed as a result of their sulfotransfer reaction, using mass spectrometry (MS). Each engineered enzyme was purified according to the procedure of Example 1. Sulfotransferase activity was monitored in 100 μ L reactions containing 50 μ M of enzyme. To each purified protein solution, 20 mg of an aryl sulfate compound (either PNS or NCS) was dissolved in 2 mL of reaction buffer (50 mM MES pH 7.0, 2 mM CaCl₂), added to the protein solution, and incubated at 37°C for 10 min. 2.5 mL of 2 mg/mL solution of *N*-deacetylated heparosan was added to protein/donor solution and incubated overnight at 37°C. The *N*-deacetylated heparosan was synthesized according to the protocol described in Balagurunathan, K. et al (eds.) (2015), *Glycosaminoglycans: Chemistry and Biology*, Methods in Molecular Biology, vol. 1229, DOI 10.1007/978-1-4939-1714-3_2, ©Springer Science+Business Media, New York, pp. 11-19 (section 3.1). To purify the *N*-sulfated product, the incubated reaction mixture was centrifuged the following day at 5,000 x g for 10 min. The filter was washed once with 2 mL water, and centrifuged again. The filtrate was added to a 1K MWCO Dialysis membrane, dialyzed for 2 days in Milli-Q water, with water changes at 1 h, 2 h, 8 h, 16 h, 32 h, and then lyophilized.

[0413] The lyophilized *N*-sulfated products from each reaction were subsequently digested with a mixture of three heparinases comprising the amino acid sequences of SEQ ID NO: 30, SEQ ID NO: 31, and SEQ ID NO: 32, which catalyze the β -eliminative cleavage of heparosan-based polysaccharides. Such lyases are available from New England Biolabs, among other chemical and biological commercial entities. 1 μ L of each lyase was incubated with 50 μ g of the lyophilized sulfated polysaccharide product and the provided digestion buffer, and incubated over 24 hours according to the packaged instructions provided by New England Biolabs with each lyase. After digestion, the lyase enzymes were inactivated by heating to 100°C for 5 minutes. Samples were centrifuged at 14,000 rpm for 30 minutes before introduction to a strong anion exchange, high performance liquid chromatography (SAX) analysis. SAX analysis was performed on a Dionex Ultimate 3000 LC system interface. Separation was carried out on a 4.6 $\times$ 250 mm Waters Spherisorb

analytical column with 5.0 μm particle size at 45 °C. Mobile phase solution A was 2.5 mM sodium phosphate, pH 3.5, while mobile phase solution B was 2.5 mM sodium phosphate, pH 3.5, and 1.2 M Sodium perchlorate. After each sample was loaded onto the column, mobile phase solutions were applied to the column at a ratio of 98% mobile phase solution A and 2% mobile phase solution B for five minutes at a flow rate of 1.4 mL/min. After five minutes, a linear gradient of increasing mobile phase solution B was applied until the ratio of mobile phase solution A to mobile phase solution B was 50:50.

[0414] Using the SAX analysis, it was determined that all six of the tested enzymes were active as sulfotransferases. However, each of the sulfotransferases were not necessarily active with both PNS and NCS. Enzymes having the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, and SEQ ID NO: 10 had activity with NCS only, and the enzyme having the amino acid sequence of SEQ ID NO: 12 had activity with PNS only. Enzymes having the amino acid sequences of SEQ ID NO: 6 and SEQ ID NO: 8 had activity with both aryl sulfate compounds.

[0415] Representative chromatograms from SAX analysis illustrating the presence of *N*-sulfated products produced as a result of the reaction are shown in Figure 26. The black chromatogram represents the *N*-deacetylated heparosan starting material and the lavender chromatogram represents the *N*-sulfated product produced by SEQ ID NO: 10. Both the starting material and product were digested with the lyases having the amino acid sequence of SEQ ID NO: 30, SEQ ID NO: 31, and SEQ ID NO: 32 according the digestion procedure described above. The green and orange chromatograms represent two disaccharide standards (HD005 and HD013) that are commercially available from Iduron, Ltd. The HD013 disaccharide comprises an unsubstituted glucosamine residue and a reduced hexuronic acid. The HD005 disaccharide is the same as HD013 except that the glucosamine residue is *N*-sulfated. All of the overlaid chromatograms are normalized so the most prominent peak in each chromatogram is assigned a normalized relative fluorescence value of 1.0.

[0416] As shown in Figure 26, the most prominent peak for HD013 disaccharide (illustrated with a * symbol) elutes almost immediately, whereas the most prominent peak for the HD005 disaccharide (illustrated with a ** symbol) elutes after approximately 17 minutes. This is expected under SAX conditions because positively-charged species (like HD013) typically do not bind to the column, whereas negatively-charged species (like HD005) do bind to the column. The *N*-deacetylated heparosan, which is similarly non-sulfated, most prominently elutes at a nearly identical time as HD013. Similarly, the lyophilized sample produced during the reaction shows a peak at a nearly identical time as HD005, indicating that the sample likely contains an *N*-sulfated product. Other

peaks within each of the chromatograms, particularly within the synthesized starting materials and products, indicate a lack of sample purity based on the use of spin-filtration columns as the sole basis of purifying the polysaccharides in each instance. Those skilled in the art would appreciate that there are several other separations techniques that can be utilized if a more purified product is desired. Additionally, the drifting upward of the baseline of the fluorescent signal in the chromatograms is a known phenomenon when increasing amounts of salt are introduced onto the column via the mobile phase.

Example 3: Mass Spectrometric Characterization of the 2-*O* Sulfated Polysaccharide Products of Engineered Aryl Sulfate-Dependent 2OST Enzymes

[0417] A study was conducted in accordance with embodiments of the present disclosure to confirm 2OST activity of enzymes comprising the amino acid sequence of SEQ ID NO: 14 or SEQ ID NO: 16 by detecting the presence of 2-*O* sulfated polysaccharide products formed as a result of their sulfotransfer reaction, using a similar procedure as in Example 2, except that the sulfo acceptor polysaccharide was commercial UF-HS in which the 2-*O* sulfate groups had been selectively removed by chemical means (product DSH001/2, available from Galen Laboratory Supplies) and analysis of each of the digested samples containing sulfated products was conducted using mass spectrometry, coupled with SAX-based high performance liquid chromatography (LCMS).

[0418] Disaccharides obtained by digesting the 2-*O* sulfated products using the heparinases having the amino acid sequence of SEQ ID NO: 30, SEQ ID NO: 31, and SEQ ID NO: 32 and according to the procedure described above in Example 2 were quantified on a Shimadzu LCMS-8050 Triple Quadrupole Liquid Chromatograph Mass Spectrometer. 100 ng of each of the digested samples, diluted in 10 mM ammonium bicarbonate (pH 10). The disaccharides were separated on a Thermo Hypercarb HPLC column (100x2.1 mm, 5 μ m). The mobile phase consisted of 10 mM ammonium bicarbonate (pH 10), and the disaccharides were eluted with an acetonitrile gradient of 0% to 20% for 2.5 min, held at 20% for the next 2.5 min, with 2 min of equilibration at 0% before the next injection; the flow rate was 0.2 mL/min, and the total run time was 7.1 min.

[0419] The extracted ion chromatograms from the LCMS are shown in Figure 27, corresponding to 2-*O* sulfated products obtained from reactions with engineered enzymes having the amino acid sequences of SEQ ID NO: 14 or SEQ ID NO: 16. Peaks were compared with chromatograms of a series of eight disaccharide standards, as well as a chromatogram from 100 ng of a commercial UF-HS polysaccharide (CAS code: 9041-08-1, available from Millipore Sigma), which was also digested using the lyase mixture. The eight reference disaccharide standards (D0A0, D0S0, D0A6,

D2A0, D0S6, D2S0, D2A6, D2S6) represent disaccharides that are variably sulfated at the *N*-, 2-*O* and 6-*O* positions. In particular, the disaccharide D2S0 represents a disaccharide having a hexuronyl residue sulfated at the 2-*O* position and an *N*-sulfated glucosamine residue. The retention time and peak areas from the spectra from all of the disaccharide standards (not shown), the digested commercial sulfated polysaccharide (not shown), and the sulfated polysaccharide products of the engineered enzymes having the amino acid sequence of SEQ ID NO: 14 or SEQ ID NO: 16 are collected in Table 1, below. Since the ionization of each individual disaccharide is different, the present percent in EIC chromatograms may not represent their actual abundance. However, the ionization efficiency is identical for each disaccharide from sample to sample. Therefore, it is believed that comparing the peak area percent of the same saccharides from sample to sample can still be achieved.

Table 1

Peak No.	Disaccharides	Peak Area %		
		Commercial standard	SEQ ID NO: 14	SEQ ID NO: 16
1	D0A0	3.9	5.9	9.1
2	D0S0	3.9	87.1	85.5
3	D0A6	3.4	ND	ND
4	D2A0	1.8	ND	ND
5	D0S6	11.8	4.1	3.1
6	D2S0	6.6	2.9	2.3
7	D2A6	1.6	ND	ND
8	D2S6	67.0	ND	ND

[0420] Sulfotransferase activity of the engineered enzymes was confirmed by the re-sulfation at the 2-*O* position of hexuronic acid residues within the sulfo acceptor polysaccharide that had previously been desulfated prior to the reaction. This is illustrated by the presence of D2S0 disaccharides within the products isolated from reactions of both engineered enzymes and NCS. Without being limited by a particular theory, it is also believed that the activity of the engineered enzyme is dependent on reacting with a portion of the polysaccharide in which the hexuronic acid residue is adjacent to a glucosamine residue that is *N*-sulfated, but not 6-*O* sulfated. This is illustrated by the lack of D2S6 (2-*O* sulfated hexuronic acid residue and an *N*,6-sulfated glucosamine residue) and D2A6 (2-*O* sulfated hexuronic acid residue and a 6-*O* sulfated *N*-acetyl glucosamine residue) disaccharides detected within the isolated sulfated polysaccharide product. This is a similar

reactivity to wild type 2OSTs within EC 2.8.2.-, which are believed to react with *N*-sulfated heparosan comprising either the structure of Formula IV or Formula V.

Example 4: Mass Spectrometric Characterization of the 6-*O* Sulfated Polysaccharide Products of Engineered Aryl Sulfate-Dependent 6OST Enzymes

[0421] A study was conducted in accordance with embodiments of the present disclosure to confirm 6OST activity of enzymes comprising the amino acid sequence of SEQ ID NO: 18, SEQ ID NO: 20, or SEQ ID NO: 22 by detecting the presence of 6-*O* sulfated polysaccharide products as a result of their sulfotransfer reaction, using a similar LCMS procedure as in Example 3, except that the sulfo acceptor polysaccharide was prepared by chemically 6-*O* desulfating commercially available UF-HS (CAS code: 9041-08-1, available from Millipore Sigma), according to the procedure provided by Kariya, Y., et al., (2000) *J. Biol. Chem.* **275** (34):25949-25958).

[0422] The extracted ion chromatograms corresponding to 6-*O* sulfated products obtained from reactions with engineered enzymes having the amino acid sequences of SEQ ID NO: 18, SEQ ID NO: 20, or SEQ ID NO: 22 are shown in Figure 28A, Figure 28B, and Figure 28C, respectively. Enzymes having the sequence of SEQ ID NO: 18 and SEQ ID NO: 20 were active when NCS was the sulfo group donor, while the enzyme having the sequence of SEQ ID NO: 22 was active when PNS was the sulfo group donor. Assigned peaks were based on the determined retention times of eight reference disaccharide standards. The eight reference disaccharide standards (D0A0, D0S0, D0A6, D2A0, D0S6, D2S0, D2A6, and D2S6) represent disaccharides that are variably sulfated at the *N*-, 2-*O*, and 6-*O* positions. D0A6, D0S6, D2A6, and D2S6 comprise 6-*O* sulfated glucosamine residues. S6 indicates an *N*,6-sulfated glucosamine residue, while A6 indicates a 6-*O* sulfated *N*-acetyl glucosamine residue. Each chromatogram indicates two integrable peaks, D0S6 and D2S6, correlating to the synthesis of *N*,6-sulfated glucosamine residues, adjacent to a hexuronic acid residue that is either non sulfated or sulfated at the 2-*O* position, respectively. The peak area % of all the labelled disaccharides is in Table 2, below. Since the ionization of each individual disaccharide is different, especially for D0A0 and D2S6, the present percent in EIC chromatograms may not represent their actual abundance. However, the ionization efficiency is identical for each disaccharide from sample to sample. Therefore, it is believed that comparing the peak area percent of the same saccharides from sample to sample can still be achieved.

Table 2

Peak No.	Disaccharides	RT (min)	Peak Area %		
			SEQ ID NO: 18	SEQ ID NO: 20	SEQ ID NO: 22
1	D0A0	7.7	4.6	6.0	5.4
2	D0S0	16.4	14.2	18.4	13.0
3	D0A6	ND	ND	ND	ND
4	D2A0	20.0	1.1	1.8	1.3
5	D0S6	23.7	4.0	3.7	5.6
6	D2S0	25.6	73.5	68.4	72.4
7	D2A6	ND	ND	ND	ND
8	D2S6	32.7	2.5	1.7	2.3

[0423] Sulfotransferase activity of the engineered enzymes was confirmed by the re-sulfation at the 6-*O* position of glucosamine residues that had been desulfated by the procedure according to Kariya, Y., et al, above. This is illustrated by the presence of D0S6 and D2S6 disaccharides within the products isolated from the reactions with each enzyme. Among each of the engineered enzymes, it appears that the 6OST having the amino acid sequence of SEQ ID NO: 22 was the most active, based on comparing the peak area percentages of the D0S6 and D2S6 disaccharides. However, while D0A6 and D2A6 polysaccharides were not observed in any of the 6-*O* sulfated products produced by the engineered enzymes, without being limited by any particular theory, it is believed that these enzymes may nonetheless be able to transfer a sulfo group to *N*-acetyl glucosamine residues in different reaction conditions, particularly by increasing the concentration of the enzyme and/or polysaccharide where the presence of *N*-acetyl glucosamine residues is confirmed prior to the reaction, based on the reactivity of natural natural 6OSTs within EC 2.8.2.-.

Example 5: Mass Spectrometric Characterization of the 3-*O* Sulfated Polysaccharide Products of Engineered Aryl Sulfate-Dependent 3OST Enzymes

[0424] A study was conducted in accordance with embodiments of the present disclosure to confirm 3OST activity of enzymes comprising the amino acid sequence of SEQ ID NO: 24, SEQ ID NO: 26, or SEQ ID NO: 28 by detecting the presence of 3-*O* sulfated polysaccharide products as a result of their sulfotransfer reaction, using a reaction, using a similar LCMS procedure as in Example 3, except that the sulfo acceptor polysaccharide was commercially-available UF-HS (CAS code: 9041-08-1, available from Millipore Sigma). Even though the unmodified UF-HS contains ~3.5% (w/w) of 3-*O* sulfated glucosamine residues, about ~60% of the glucosamine residues are *N*,6-sulfated and

are adjacent to a 2-*O* sulfated hexuronic acid residue, as in Formula X. Consequently, these *N*,6-sulfated glucosamine residues can still be 3-*O* sulfated.

[0425] The extracted ion chromatograms are shown in Figure 29, along with chromatograms of a series of ten reference standards and 100 ng of the commercial polysaccharide, which was also digested using the lyase mixture. The ten reference standards (D0A0, D0S0, D0A6, D2A0, D0S6, D2S0, D2A6, D2S6, D0A6G0S3, and D0A6G0S9) represent di- or tetrasaccharides that are variably sulfated at the *N*-, 2-*O*, 3-*O*, and 6-*O* positions (black spectrum). For clarity, reference peaks that include 3-*O* sulfated glucosamine residues (D0A6G0S3) and (D0A6G0S9) are indicated in the digested commercial polysaccharide spectrum, shown in red. Four mass spectra representing the digested sulfated polysaccharide products from reactions with enzymes comprising the amino acid sequence of SEQ ID NO: 24 (PNS, yellow spectrum), SEQ ID NO: 26 (PNS, purple spectrum) (NCS, green spectrum), and SEQ ID NO: 28 (NCS, blue spectrum) are shown below the digested commercial polysaccharide spectrum. The peak area % of all the labelled disaccharides and tetrasaccharides is in Table 3, below. Since the ionization of each individual disaccharide is different, especially for D0A0 and D2S6, the present percent in EIC chromatograms may not represent their actual abundance. However, the ionization efficiency is identical for each disaccharide or tetrasaccharide from sample to sample. Therefore, it is believed that comparing the peak area percent of the same saccharides from sample to sample can still be achieved.

Table 3

peak No.	Disaccharides	RT (min)	Peak Area %				
			Commercial standard	SEQ ID NO: 24	SEQ ID NO: 26 (NCS)	SEQ ID NO: 28	SEQ ID NO: 26 (PNS)
1	D0A0	4.5	1.9	0.6	0.8	1.4	N.D.
2	D0S0	22.5	3.7	1.4	1.7	2.3	N.D.
3	D0A6	24.6	4.2	2.8	3.1	4.5	N.D.
4	D2A0	26.2	2.2	0.5	0.8	0.5	N.D.
5	D0S6	37.5	16.0	10.9	10.6	13.1	N.D.
6	D2S0	38.5	6.5	4.9	5.4	5.4	N.D.
7	D2A6	40.3	1.6	0.8	0.8	0.9	N.D.
8	D2S6	48.4	60.3	73.4	71.6	64.0	100.0
9	D0A6G0S3	52.9	0.6	0.8	0.9	1.4	N.D.
10	D0A6G0S9	58.2	3.0	4.0	4.1	6.5	N.D.

[0426] Sulfotransferase activity of each of the engineered enzymes was confirmed by the increase in the abundance of the D0A6G0S3 (hexuronic acid-6-*O*-sulfated *N*-acetyl glucosamine-glucuronic acid-*N*,3,6-sulfated glucosamine) and D0A6G0S9 (hexuronic acid-6-*O*-sulfated *N*-acetyl glucosamine-glucuronic acid-*N*,3-sulfated glucosamine) tetrasaccharides relative to the commercial UF-HS sample. However, the total abundance of disaccharides in the SEQ ID NO: 26 PNS sample was much lower than other samples. Subsequent trials included re-running the experiment with 10 times more injection volume, and a re-digestion of the sample with the lyase mixture. Nonetheless, only the D2S6 disaccharide could ever be found, indicating that the abundance of the SEQ ID NO: 26 PNS sulfated polysaccharide sample isolated initially was extremely low, and/or that the polysaccharide resists lyase digestion, causing the product to potentially elute from the column with a retention time longer than one hour.

[0427] Nonetheless, NMR studies (indicated below in Example 6) indicated 3-*O* sulfotransferase activity with the enzyme comprising the amino acid sequence SEQ ID NO: 26 when PNS is the aryl sulfate compound. Further, the enzyme having the amino acid sequence of SEQ ID NO: 26 was determined to be active as a sulfotransferase when NCS is the aryl sulfate compound. Therefore, it is believed that the observed results for the SEQ ID NO: 26 PNS sulfated polysaccharide sample during the LCMS experiment result from the sample produced for the purpose of the experiment, and not the activity of the enzyme itself. Otherwise, a higher abundance of 3-*O* sulfation was found in all of the other sulfated polysaccharide products from SEQ ID NO: 24, SEQ ID NO: 26, and SEQ ID NO: 28, relative to the commercial UF-HS standard.

Example 6: Confirmation of Sulfotransferase Activity of the Engineered 3OSTs Using Nuclear Magnetic Resonance

[0428] A study was conducted in accordance with embodiments of the present disclosure to confirm the 3-*O* sulfotransferase activity of the engineered enzymes having the amino acid sequence of SEQ ID NO: 24, SEQ ID NO: 26, and SEQ ID NO: 28, particularly the activity of the enzyme having the amino acid sequence SEQ ID NO: 26 with PNS as the sulfo group donor. Each enzyme was purified according to the procedure of Example 1. To each purified protein solution, 20 mg of an aryl sulfate compound (PNS or NCS) dissolved in 2 mL of reaction buffer (50 mM MES pH 7.0, 2 mM CaCl₂) was added to the protein solution and incubated at 37°C for 10 min. 2.5 mL of 2 mg/mL solution of the commercial UF-HS polysaccharide utilized in Example 5 was added to protein/donor solution and incubated overnight at 37°C.

[0429] Each reaction was centrifuged at 5,000 x g for 10 min, applied to a pre-wetted 30K MWCO Amicon-15 filter and centrifuged at 5,000 x g for 10 min. The filter was washed once with 2 mL water, and centrifuged again. The filtrate was added to a 1K MWCO Dialysis membrane, dialyzed for 2 days in Milli-Q water, with water changes at 1 h, 2 h, 8 h, 16 h, 32 h, and then lyophilized. The dry, white powder was resuspended in 400 μ L D₂O, lyophilized to remove exchangeable protons, then resuspended in 600 μ L D₂O and transferred to NMR tubes (Wilmad, 0.38 mm x 7"). To determine if sulfotransfer took place, ¹H NMR spectra were obtained on a Bruker 600 MHz NMR, 32 scans, with water suppression. The overall reaction scheme is shown in Figure 30. Within Figure 30, positions within the polysaccharide capable of accepting a sulfo group are shown in red, and exchangeable protons having the ability to exhibit resonance upon deuterium exchange are shown in blue. Crude mixture peaks were integrated to literature-referenced spectra for the sulfo acceptor polysaccharide and associated 3-*O* sulfated product.

[0430] As shown in the overlain spectra in Figure 31, a sharp peak at 5.15 ppm that correlates to the proton at the C2 carbon of the 2-*O* sulfated iduronic acid present in the commercial UF-HS (shown in red) disappears upon reacting with enzymes comprising the amino acid sequence of SEQ ID NO: 24, SEQ ID NO: 26, and SEQ ID NO: 28. The proton of interest is circled in green at the polysaccharide shown above the spectra. The ¹H NMR spectrum for the product in the presence of an enzyme comprising the amino acid sequence of SEQ ID NO: 24 reacting with PNS is shown in yellow-green; the ¹H NMR spectrum for the product in the presence of an enzyme comprising the amino acid sequence of SEQ ID NO: 26 reacting with PNS is shown in blue; the ¹H NMR spectrum for the product in the presence of an enzyme comprising the amino acid sequence of SEQ ID NO: 26 reacting with NCS is shown in green; and the ¹H NMR spectrum for the product in the presence of an enzyme comprising the amino acid sequence of SEQ ID NO: 28 reacting with NCS is shown in purple. In each of the product spectra, the IdoA_{2s} peak shifts to between approximately 5.0 and 5.05 ppm. A similar transition is shown when incubating the natural human sulfotransferase enzyme with the same polysaccharide substrate and PAPS (data not shown).

[0431] As shown in Figure 32, the region between 4.5 and 3.5 shows several peaks that similarly shift in response to the addition of the sulfate group to the 3-*O* position of a glucosamine residue, all of which correlate to the same shifts observed upon incubating the natural human 3OST enzyme with the same commercial UF-HS substrate and PAPS. Peaks that shift are indicated in curved arrows, and positions of the peaks from 3-*O* sulfated polysaccharides produced by enzymes having the amino acid sequence of SEQ ID NO: 24, SEQ ID NO: 26, or SEQ ID NO: 28, are shown with straight arrows. The largest shift occurs for H3 of GlcNS_{3S6S}, from 3.7 ppm to 4.2 ppm. This results

from being closest to the newly added 3-*O* sulfate group. Additionally, the H3 proton of Ido_{2S} and H5 of Glc_{NS3S6S} both converge toward a peak at 4.07 ppm, which shows two overlapping peaks. H4 of Glc_{NS3S6S} shifts moderately downfield from the 3.7 ppm region to the 3.8 ppm region, and according to references, many peaks such as H3 & H4 from Glc_{NS6S} and H3, H4, and H5 from GlcA shift from the 3.7 ppm region to the 3.6 ppm region.

Example 7: Chemical Synthesis of *N*-Sulfated Heparosan toward the production of NS_{2S6S3S}-HS Products

[0432] A study was conducted in accordance with embodiments of the present disclosure to chemically synthesize *N*-sulfated heparosan that can be utilized as sulfo acceptor polysaccharides with any of the engineered aryl sulfate-dependent enzymes, particularly either of the engineered 2OST enzymes. *N*-deacetylated heparosan was prepared according to the protocol described in Balagurunathan, K. et al., above. Particularly, the heparosan that eluted from the DEAE resin was precipitated overnight in ethanol saturated with sodium acetate, at -30 °C, before being resuspended in water and dialyzed within a cellulose dialysis membrane having a 1,000 Da molecular weight cut-off (MWCO).

[0433] To *N*-deacetylate the heparosan, enough sodium hydroxide pellets (~4.0 g) were dissolved to make a 2.5 M solution in a 40 mL aliquot of the dialyzed heparosan in water. The solution was incubated at 55 °C for 16 hours, with shaking at 100 rpm. The sodium hydroxide within the sample was then neutralized with acetic acid until the solution reached a pH of ~7.0, and then dialyzed in water overnight within a 1,000 MWCO dialysis membrane.

[0434] Subsequent *N*-sulfation of the *N*-deacetylated heparosan was accomplished by adding 100 mg of sodium carbonate and 100 mg of sulfur trioxide-triethylamine complex, and allowing the composition to incubate at 48 °C until all of the solid was dissolved. The pH of the solution was then readjusted to ~9.5, using acetic acid. After incubation at 48 °C overnight with shaking at 100 rpm, an additional 100 mg of sodium carbonate and 100 mg of sulfur trioxide-triethylamine complex was added, before subsequent readjustment of the pH to ~9.5 using acetic acid. The solution was incubated at 48 °C for an additional 24 hours. The sulfated polysaccharide solution was neutralized with acetic acid to a pH of ~7.0, and dialyzed in water overnight within a 1,000 MWCO dialysis membrane. The dialyzed *N*-sulfated heparosan was then lyophilized prior to further use. The *N*-sulfated heparosan was then further purified by loading it onto a Zenix SEC-100 column and eluting it isocratically with 0.1 M ammonium acetate, pH 9.0.

[0435] The functionalization of the purified heparosan-based polysaccharide was characterized by digesting it with a mixture of three heparinases comprising the amino acid sequences of SEQ ID NO: 30, SEQ ID NO: 31, and SEQ ID NO: 32, and analyzing the digested samples using SAX, using a similar procedure described above. As a positive control, the commercial HD005 disaccharide of Example 2, containing *N*-sulfated glucosamine residues, was also analyzed. Representative chromatograms of both samples are shown in Figure 33, in which the commercial HD005 disaccharide is illustrated in purple, and the synthesized and digested *N*-sulfated sample is shown in red. In both chromatograms, a strong peak is present at about 16.5 minutes, indicating that the synthesized sample contains *N*-sulfated glucosamine residues.

Example 8: Preparation of an NS2S/HS Polysaccharide Product

[0436] A study was conducted in accordance with embodiments of the present disclosure to synthesize an NS2S/HS polysaccharide product using an engineered 2OST, using the *N*-sulfated heparosan synthesized in Example 7 as the sulfo acceptor. In a conical-bottom centrifuge tube, 80 mM aliquots of NCS were dissolved in 50 mM MES pH 7.0, 2 mM CaCl₂. To each solution, 2 mg of the enzyme having the sequence of SEQ ID NO: 14, based on the absorbance of the enzyme sample at 280 nm, was added (about 4 mL). 5 mg of the lyophilized *N*-sulfated heparosan synthesized in Example 7 was resuspended in 1 mL of water and added to the reaction mixture containing the enzyme and NCS. The entire reaction mixture was then incubated at 34 °C with shaking at 30 rpm, for 48 hours. A second set of reactions were prepared using the same procedure, except that 2 mg of a C₅-hexuronyl epimerase comprising the amino acid sequence of SEQ ID NO: 29 was also added to the reaction mixture, prior to incubation.

[0437] The sulfated polysaccharide products from both sets of reactions were purified by first precipitating out the proteins from the reaction mixtures by placing the reaction vessels in boiling water for 10 minutes and centrifuging at high speed to form a pellet. The supernatant containing the polysaccharide products was decanted from the pellet and dialyzed in water overnight within a 1,000 MWCO dialysis membrane. The dialyzed products were then lyophilized for future use.

[0438] To characterize the polysaccharide products, lyophilized samples were resuspended in 400 µL of water, and purified using a Q-Sepharose Fast Flow Column (GE Biosciences). Samples were eluted from the column using a gradient ranging from 0 to 2M NaCl, in 20 mM sodium acetate buffer, pH 5.0. Purified polysaccharides were then digested and analyzed by SAX according to the procedures in Example 2 above, along with a commercial polysaccharide, HD002 (Iduron), which contains disaccharides of 2-*O* sulfated uronic acid and *N*-sulfated glucosamine. Representative

chromatograms of reactions either without or including the epimerase enzyme are shown in Figure 34 and Figure 35, respectively. In Figure 34, the commercial HD002 disaccharide is illustrated in blue, and the NS2S/HS product is shown in purple. The chromatogram for the HD002 disaccharide has a single, sharp peak at about 21.1 minutes, which correlates to a sharp peak at a nearly identical time in the reaction product, indicating the time that an NS2S/HS was formed as a result of the reaction. In Figure 35, the HD002 disaccharide was provided within a mixture containing other disaccharide standards, and the chromatogram is illustrated in black. The disaccharide corresponding to HD002 (circled in red) eluted at 20.5 minutes, while the reaction product has a sharp peak that eluted from the column at a nearly identical time, indicating that an NS2S/HS product was formed as a result of the reaction.

Example 9: Preparation of an NS2S6S/HS Product

[0439] A study was conducted in accordance with embodiments of the present disclosure to synthesize an NS2S6S/HS product using the procedure of Example 8, except that the NS2S/HS product of Example 8 was used as the sulfo acceptor polysaccharide, and the engineered 6OST having the amino acid sequence of SEQ ID NO: 18 was used as the enzyme.

[0440] Representative chromatograms of the sulfated polysaccharide product and a mixture of commercial disaccharides are shown in Figure 36. The commercial mixture is shown in blue, and the synthesized polysaccharide product is shown in red. The chromatogram of the commercial mixture exhibits a peak at about 23.7 minutes, that correlates to HD001 (Iduron), which consists of disaccharides of 2-*O* sulfated uronic acid and *N*-, 6-*O* sulfated glucosamine, while the reaction product exhibits a similar peak at 23.4 minutes, indicating that an NS2S6S/HS was formed as a result of the reaction. Other peaks present within the NS2S6S/HS product include undigested polysaccharide (2.5 min), unsubstituted uronic acid and *N*-acetylated glucosamine (5.5 min), and unsubstituted uronic acid and *N*-, 6-*O* sulfated glucosamine.

Example 10: Preparation of an NS6S/HS Product

[0441] A study was conducted in accordance with embodiments of the present disclosure to synthesize an NS6S/HS product using the procedure of Example 9, except that the *N*-sulfated heparosan synthesized in Example 7 was used as the sulfo acceptor, and the *N*-sulfated heparosan was not 2-*O* sulfated prior to 6-*O* sulfating using the engineered sulfotransferase having the amino acid sequence of SEQ ID NO: 18.

[0442] Representative chromatograms of the sulfated polysaccharide product and a mixture of commercial disaccharides are shown in Figure 37. The commercial mixture is shown in blue

(bottom chromatogram), and the synthesized NS6S/HS product is shown in black (top chromatogram). The chromatogram of the commercial mixture exhibits a peak at about 17.8 minutes, that correlates to disaccharides having an unsulfated uronic acid and an *N*-, 6-*O* sulfated glucosamine (Δ UA-GlcNS,6S), as well as a prominent peak at about 13.8 minutes that correlates to disaccharides having an unsulfated uronic acid and an *N*-sulfated glucosamine (Δ UA-GlcNS).

Example 11: Preparation of an NS2S6S3S-HS Product

[0443] A study was conducted in accordance with embodiments of the present disclosure to synthesize a sulfated polysaccharide product comprising *N*-, 6-*O*, 3-*O* sulfated glucosamine and 2-*O* sulfated hexuronic acid residues, using the procedure of Example 8, except that the chemically synthesized *N*-, 2-*O*, 6-*O* sulfated polysaccharide of Example 9 is used as the sulfo acceptor polysaccharide, and an engineered 3-*O* sulfotransferase enzyme having the amino acid sequence of SEQ ID NO: 28 is used as the sulfotransferase enzyme.

[0444] Sulfated polysaccharide products were digested and analyzed using LCMS to confirm the production of an NS2S6S3S-HS product. To facilitate study using LCMS, sulfated polysaccharide products of the SEQ ID NO: 28 sulfotransferase enzyme were isolated and derivatized with aniline tags, according to the procedures described in Lawrence, R., et al., (2008) *J. Biol. Chem.* **283** (48):33674-33684, the disclosure of which is incorporated by reference in its entirety. Briefly, some GAGs, including commercial UF-HS and other NS2S6S3S-HS polysaccharides, can be quantified and compared ratiometrically using LCMS by chemically modifying the sulfated product. Lawrence, R., et al., describes the tagging of the reducing end of lyase-generated disaccharides and tetrasaccharides with [$^{12}\text{C}_6$]- and [$^{13}\text{C}_6$]-aniline and propionylation of *N*-unsubstituted glucosamine residues. Isotopic tagging of the disaccharides and tetrasaccharides has no effect on the chromatographic retention times, but can be discriminated using mass spectroscopy.

[0445] Sulfated disaccharide and tetrasaccharide products were prepared by anion exchange chromatography, as described in Example 8, and digestion with a mixture of three heparinases comprising the amino acid sequences of SEQ ID NO: 30, SEQ ID NO: 31, and SEQ ID NO: 32, as described above in Example 7. 1 pmol to 10 nmol of the digested samples were transferred to 1.5-ml microcentrifuge tubes and dried down in a centrifugal evaporator. [$^{12}\text{C}_6$]-aniline or [$^{13}\text{C}_6$]-aniline (15 μl , 165 μmol) and 15 μl of 1 M NaCNBH₃ freshly prepared in dimethyl sulfoxide:acetic acid (7:3, v/v) were added to each sample. Reactions were carried out at 65 °C for 4 h, or alternatively at 37 °C for 16 h, and then dried in a centrifugal evaporator.

[0446] Unsubstituted amines were reacted with propionic anhydride. Dried samples were reconstituted in 20 μ l of 50% methanol, and 3 μ l of propionic anhydride (23.3 μ mol) was added. Reactions were carried out at room temperature for 2 h. Acylated samples were subsequently aniline-tagged as described above.

[0447] A quadrupole ion trap Liquid Chromatograph Mass Spectrometer with an electrospray ionization source, similar to the Shimadzu LCMS-8050 mass spectrometer described in Example 3, was used for disaccharide analysis. Derivatized and non-derivatized disaccharide residues were separated on a C18 reversed-phase column with the ion pairing agent dibutylamine (DBA). The isocratic steps were: 100% buffer A (8 mM acetic acid, 5 mM DBA) for 10 min, 17% buffer B (70% methanol, 8 mM acetic acid, 5 mM DBA) for 15 min; 32% buffer B for 15 min, 40% buffer B for 15 min, 60% buffer B for 15 min; 100% buffer B for 10 min; and 100% buffer A for 10 min. Generally, mass spectra for samples containing 3-*O* sulfated product are expected to generate *m/z* peaks corresponding to tetrasaccharides that are resistant to digestion by the heparinases, as described above in Example 5. Tetrasaccharides that can be produced include, but are not limited to: 4,5-unsaturated uronic acid – *N*-acetylated, 6-*O* sulfated glucosamine – glucuronic acid – *N*-sulfated, 3-*O* sulfated glucosamine (Δ U-ANAc6S-G-ANS3S); 4,5-unsaturated uronic acid – *N*-acetylated, 6-*O* sulfated glucosamine – glucuronic acid – *N*-sulfated, 3-*O* sulfated, 6-*O* sulfated glucosamine (Δ U-ANAc6S-G-ANS3S6S); 4,5-unsaturated uronic acid – *N*-sulfated, 6-*O* sulfated glucosamine – glucuronic acid – *N*-sulfated, 3-*O* sulfated, 6-*O* sulfated glucosamine (Δ U-ANS6S-G-ANS3S6S); 4,5-unsaturated, 2-*O* sulfated uronic acid – *N*-sulfoglucosamine – glucuronic acid – *N*-sulfated, 3-*O* sulfated, 6-*O* sulfated glucosamine (Δ U2S-ANS-G-ANS3S6S); and 4,5-unsaturated, 2-*O* sulfated uronic acid – *N*-sulfated, 6-*O* sulfated glucosamine – glucuronic acid – *N*-sulfated, 3-*O* sulfated, 6-*O* sulfated glucosamine (Δ U2S-ANS6S-G-ANS3S6S). In particular, LCMS of the digested polysaccharide samples collected from the reaction with the SEQ ID NO: 28 sulfotransferase enzyme generated mass spectra (not shown) with *m/z* peaks corresponding to the Δ U-ANAc6S-G-ANS3S6S (*m/z* = 1036), Δ U-ANS6S-G-ANS3S6S (*m/z* = 1074), and Δ U2S-ANS6S-G-ANS3S6S (*m/z* = 1154) tetrasaccharides, indicating that the NS2S6S3S-HS product was produced by the reaction with the SEQ ID NO: 28 engineered sulfotransferase enzyme.

Example 12: Confirmation of Anticoagulant Activity of the NS2S6S3S-HS Product

[0448] A study is conducted in accordance with embodiments of the present disclosure to determine whether 3-*O* sulfated polysaccharide products produced in Example 11 have a binding affinity to antithrombin using a procedure similar to Meneghetti, G., et al. (2017) *Org. Biomol. Chem.* **15**:6792-

6799). It is expected that melting curves of antithrombin in the presence of the 3-*O* sulfated polysaccharide products produced in Example 11 will demonstrate a higher melting temperature than antithrombin alone, indicating that the 3-*O* sulfated polysaccharide product produced in Example 11 comprises the structure of Formula I.

Example 13: Determination of Engineered Aryl Sulfate-Dependent Mutants of Other EC

2.8.2.8 Enzymes

[0449] A study is conducted in accordance with embodiments of the present disclosure to engineer additional aryl sulfate-dependent glucosaminyl *N*-sulfotransferase enzymes. As described above, the aryl sulfate-dependent glucosaminyl *N*-sulfotransferase enzymes having the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, and SEQ ID NO: 12 have been engineered to be mutants of the *N*-sulfotransferase domain of the human glucosaminyl *NDST* enzyme (see entry sp|P52848|NDST_1_HUMAN, in Figure 3 above), which is a member of enzyme class EC 2.8.2.8. By generating and analyzing a multiple sequence alignment that includes both the amino acid sequences of the *N*-sulfotransferase domain of one or more of the other glucosaminyl *N*-deacetylase/ *N*-sulfotransferase enzymes within EC 2.8.2.8 as well as the amino acid sequences of aryl sulfate-dependent glucosaminyl *N*-sulfotransferase enzymes having the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, and/or SEQ ID NO: 12, mutations in the amino acid sequences in the engineered glucosaminyl *N*-sulfotransferase enzymes can be observed relative to the amino acid sequences of the natural EC 2.8.2.8 enzymes within the same alignment. Upon selecting the amino acid sequence of the *N*-sulfotransferase domain of a natural 2.8.2.8 enzyme that is not the human glucosaminyl *NDST* enzyme, mutations that are present within the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, and/or SEQ ID NO: 12 can be engineered into the natural sequence in order to form additional mutants that can have aryl sulfate-dependent sulfotransferase activity.

[0450] As a non-limiting example, the amino acid sequence encoding for the *N*-sulfotransferase domain of the pig glucosaminyl *NDST* enzyme (entry tr|M3V841|M3V841_PIG, as illustrated in the sequence alignment in Figure 3), is aligned with the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, and SEQ ID NO: 12. Amino acid mutations that are present in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, and SEQ ID NO: 12 are engineered into their equivalent positions within the amino acid sequence of the *N*-sulfotransferase domain of the pig *NDST* enzyme, in order to generate the

mutant amino acid sequences SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40, respectively. Enzymes comprising the amino acid sequences of SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40, respectively, will be utilized in Example 14 and Example 15, below. However, a person skilled in the art would appreciate that the same procedure can be applied to generate mutants of the *N*-sulfotransferase domain, or the entire enzyme, with respect to any of the other glucosaminyl natural *N*-deacetylase/ *N*-sulfotransferase enzymes within the EC 2.8.2.8 enzyme class, and that those are omitted for clarity.

Example 14: Expression and Purification of Engineered Aryl Sulfate-Dependent EC 2.8.2.8 Mutants

[0451] A study is conducted in accordance with embodiments of the present disclosure to determine whether genes encoding for engineered glucosaminyl *N*-sulfotransferase enzymes having the amino acid sequences SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40, respectively, can be transformed into host cells, and that enzymes comprising each of those amino acid sequences can be subsequently expressed, isolated, and purified according to the procedure of Example 1, above. Codon-optimized nucleotide sequences are determined that encode for enzymes having the amino acid sequences of SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40, respectively, based on the desired expression host. Upon synthesizing or inserting those genes within a suitable expression vector, it is expected that genes encoding for each of the amino acid sequences SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40, respectively, will be transformed into host cells, and that enzymes containing those sequences will be subsequently expressed, isolated, and purified in a sufficient quantity and purity to determine aryl sulfate-dependent glucosaminyl *N*-sulfotransferase activity.

Example 15: Sulfotransferase Activity of EC 2.8.2.8 Mutants

[0452] A study is conducted in accordance with embodiments of the present disclosure to determine whether mutant enzymes comprising the sequences of SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40, respectively, are active sulfotransferases, using the procedures of Example 2. It is expected that SAX studies will confirm the presence of *N*-sulfated polysaccharide products formed as a result of reacting *N*-deacetylated heparosan and an aryl sulfate compound with each of the engineered enzymes comprising the

sequences of SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, or SEQ ID NO: 40, respectively.

Example 16: Determination of Engineered Aryl Sulfate-Dependent Mutants of Other 2OST Enzymes within EC 2.8.2.-

[0453] A study is conducted in accordance with embodiments of the present disclosure to engineer additional aryl sulfate-dependent 2OST enzymes. As described above, the aryl sulfate-dependent 2OST enzymes having the amino acid sequences of SEQ ID NO: 14 and SEQ ID NO: 16 have been engineered to be mutants of the chicken 2OST enzyme (see entry sp|Q76KB1|HS2ST_CHICK, in Figure 14, above), which is a member of enzyme class EC 2.8.2.-. By generating and analyzing a multiple sequence alignment that includes both the amino acid sequences of one or more of the other 2OST enzymes within EC 2.8.2.-, as well as the amino acid sequences of aryl sulfate-dependent 2OST enzymes having the amino acid sequences of SEQ ID NO: 14 and/or SEQ ID NO: 16, mutations in the amino acid sequences in the engineered 2OST enzymes can be observed relative to the amino acid sequences of the natural 2OST enzymes within the same alignment. Upon selecting the amino acid sequence of a natural 2OST enzyme that is not the chicken 2OST enzyme, mutations that are present within the amino acid sequences of SEQ ID NO: 14 and/or SEQ ID NO: 16 can be engineered into the natural sequence in order to form additional mutants that can have aryl sulfate-dependent sulfotransferase activity.

[0454] As a non-limiting example, the amino acid sequence encoding for the human 2OST enzyme (entry sp|Q7LGA3|HS2ST_HUMAN, as illustrated in the sequence alignment in Figure 14), is aligned with the amino acid sequences of SEQ ID NO: 14 and SEQ ID NO: 16. Amino acid mutations that are present in SEQ ID NO 14 and SEQ ID NO: 16 are engineered into their equivalent positions within the amino acid sequence of the human 2OST enzyme, in order to generate the mutant amino acid sequences SEQ ID NO: 41 and SEQ ID NO: 42, respectively. Enzymes comprising the amino acid sequences of SEQ ID NO: 41 and SEQ ID NO: 42, respectively, will be utilized in Example 17 and Example 18, below. However, a person skilled in the art would appreciate that the same procedure can be applied to generate aryl sulfate-dependent mutants with respect to any of the other 2OST enzymes within the EC 2.8.2.- enzyme class, and that those are omitted for clarity.

Example 17: Expression and Purification of EC 2.8.2.- Mutants Having 2OST Activity

[0455] A study is conducted in accordance with embodiments of the present disclosure to determine whether genes encoding for engineered 2OST enzymes having the amino acid sequences

SEQ ID NO: 41 and SEQ ID NO: 42, respectively, can be transformed into host cells, and that enzymes comprising each of those amino acid sequences can be subsequently expressed, isolated, and purified according to the procedure of Example 1, above. Codon-optimized nucleotide sequences are determined that encode for enzymes having the amino acid sequences of SEQ ID NO: 41 and SEQ ID NO: 42, respectively, based on the desired expression host. Upon synthesizing or inserting those genes within a suitable expression vector, it is expected that genes encoding for each of the amino acid sequences SEQ ID NO: 41 and SEQ ID NO: 42, respectively, will be transformed into host cells, and that enzymes containing those sequences will be subsequently expressed, isolated, and purified in a sufficient quantity and purity to determine aryl sulfate-dependent 2OST activity.

Example 18: 2OST Activity of EC 2.8.2.- Mutants

[0456] A study is conducted in accordance with embodiments of the present disclosure to determine whether mutant enzymes comprising the sequences of SEQ ID NO: 41 and SEQ ID NO: 42, respectively, are active sulfotransferases, using the procedures of Example 3. It is expected that MS studies will confirm the presence of NS2S/HS products formed as a result of reacting an *N*-sulfated heparosan-based polysaccharide and an aryl sulfate compound with each of the engineered enzymes comprising the sequences of SEQ ID NO: 41 and SEQ ID NO: 42, respectively. It is also expected that both enzymes will be active with heparosan-based polysaccharides comprising either or both of Formula IV or Formula V.

Example 19: Determination of Engineered Aryl Sulfate-Dependent Mutants of Other 6OST Enzymes within EC 2.8.2.-

[0457] A study is conducted in accordance with embodiments of the present disclosure to engineer additional aryl sulfate-dependent 6OST enzymes. As described above, the aryl sulfate-dependent 6OST enzymes having the amino acid sequences of SEQ ID NO: 18, SEQ ID NO: 20, and SEQ ID NO: 22 have been engineered to be mutants of isoform 1 of the mouse 6OST enzyme (see entry Q9QYK5|H6ST1_MOUSE, in Figure 18, above), which is a member of enzyme class EC 2.8.2.-. By generating and analyzing a multiple sequence alignment that includes both the amino acid sequences of one or more of the other 6OST enzymes within EC 2.8.2.-, as well as the amino acid sequences of aryl sulfate-dependent 6OST enzymes having the amino acid sequences of SEQ ID NO: 18, SEQ ID NO: 20, and/or SEQ ID NO: 22, mutations in the amino acid sequences in the engineered 6OST enzymes can be observed relative to the amino acid sequences of the natural 6OST enzymes within the same alignment. Upon selecting the amino acid sequence of a natural 6OST enzyme that is not the mouse 6OST enzyme, mutations that are present within the amino acid

sequences of SEQ ID NO: 18, SEQ ID NO: 20, and/or SEQ ID NO: 22 can be engineered into the natural sequence in order to form additional mutants that can have aryl sulfate-dependent sulfotransferase activity.

[0458] As a non-limiting example, the amino acid sequence encoding for the pig 6OST enzyme (entry I3LAM6[I3LAM6_PIG, as illustrated in the sequence alignment in Figure 18), is aligned with the amino acid sequences of SEQ ID NO: 18, SEQ ID NO: 20, and SEQ ID NO: 22. Amino acid mutations that are present in SEQ ID NO: 18, SEQ ID NO: 20, and SEQ ID NO: 22 are engineered into their equivalent positions within the amino acid sequence of the pig 6OST enzyme, in order to generate mutant amino acid sequences. Generated mutant amino acid sequences corresponding to residues 67-377 of the pig 6OST enzyme, as illustrated in Figure 18, are disclosed as SEQ ID NO: 45, SEQ ID NO: 46, and SEQ ID NO: 47, respectively. Generated mutant amino acid sequences corresponding to the full-length amino acid sequence for the pig 6OST enzyme (not shown in Figure 18) are disclosed as SEQ ID NO: 48, SEQ ID NO: 49, and SEQ ID NO: 50, respectively.

[0459] In another non-limiting example, the full-length amino acid sequence encoding for the encoding for isoform 3 of the mouse 6OST enzyme (entry Q9QYK4[H6HS3_MOUSE, a truncated sequence for which is illustrated in the sequence alignment in Figure 18) is aligned with the amino acid sequences of SEQ ID NO: 18, SEQ ID NO: 20, and SEQ ID NO: 22. Amino acid mutations that are present in SEQ ID NO: 18, SEQ ID NO: 20, and SEQ ID NO: 22 are engineered into their equivalent positions within the amino acid sequence of isoform 3 of the mouse 6OST enzyme, in order to generate mutant amino acid sequences. The generated full-length amino acid sequences are disclosed as SEQ ID NO: 59, SEQ ID NO: 60, and SEQ ID NO: 61, respectively. Enzymes comprising the amino acid sequences of SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 59, SEQ ID NO: 60, and SEQ ID NO: 61, respectively, will be utilized in Example 20 and Example 21, below. However, a person skilled in the art would appreciate that the same procedure can be applied to generate aryl sulfate-dependent mutants with respect to any of the other 6OST enzymes within the EC 2.8.2.-enzyme class, and that those are omitted for clarity.

Example 20: Expression and Purification of EC 2.8.2.- Mutants Having 6OST Activity

[0460] A study is conducted in accordance with embodiments of the present disclosure to determine whether genes encoding for engineered 6OST enzymes having the amino acid sequences SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ

ID NO: 59, SEQ ID NO: 60, and SEQ ID NO: 61, respectively, can be transformed into host cells, and that enzymes comprising each of those amino acid sequences can be subsequently expressed, isolated, and purified according to the procedure of Example 1, above. Codon-optimized nucleotide sequences are determined that encode for enzymes having the amino acid sequences of SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 59, SEQ ID NO: 60, and SEQ ID NO: 61, respectively, based on the desired expression host. Upon synthesizing or inserting those genes within a suitable expression vector, it is expected that genes encoding for each of the amino acid sequences SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 59, SEQ ID NO: 60, and SEQ ID NO: 61, respectively, will be transformed into host cells, and that enzymes containing those sequences will be subsequently expressed, isolated, and purified in a sufficient quantity and purity to determine aryl sulfate-dependent 6OST activity.

Example 21: 6OST Activity of EC 2.8.2.- Mutants

[0461] A study is conducted in accordance with embodiments of the present disclosure to determine whether mutant enzymes comprising the sequences of SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 59, SEQ ID NO: 60, and SEQ ID NO: 61, respectively, are active sulfotransferases, using the procedures of Example 4. It is expected that MS studies will confirm the presence of NS2S6S/HS products formed as a result of reacting an NS2S/HS polysaccharide and an aryl sulfate compound with each of the engineered enzymes comprising the sequences of SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 59, SEQ ID NO: 60, and SEQ ID NO: 61, respectively.

Example 22: Determination of Engineered Aryl Sulfate-Dependent Mutants of Other 3OST Enzymes within EC 2.8.2.23

[0462] A study is conducted in accordance with embodiments of the present disclosure to engineer additional aryl sulfate-dependent 3OST enzymes. As described above, the aryl sulfate-dependent 3OST enzymes having the amino acid sequences of SEQ ID NO: 24, SEQ ID NO: 26, and SEQ ID NO: 28 have been engineered to be mutants of isoform 1 of the human 3OST enzyme (see entry sp|O14792|HS3S1_HUMAN, in Figure 23, above), which is a member of enzyme class EC 2.8.2.23. By generating and analyzing a multiple sequence alignment that includes both the amino acid sequences of one or more of the other 3OST enzymes within EC 2.8.2.23, as well as the amino acid sequences of aryl sulfate-dependent 3OST enzymes having the amino acid sequences of SEQ

ID NO: 24, SEQ ID NO: 26, and/or SEQ ID NO: 28, mutations in the amino acid sequences in the engineered 3OST enzymes can be observed relative to the amino acid sequences of the natural 3OST enzymes within the same alignment. Upon selecting the amino acid sequence of a natural 3OST enzyme that is not the human 3OST enzyme, mutations that are present within the amino acid sequences of SEQ ID NO: 24, SEQ ID NO: 26, and/or SEQ ID NO: 28 can be engineered into the natural sequence in order to form additional mutants that can have aryl sulfate-dependent sulfotransferase activity.

[0463] As a non-limiting example, the amino acid sequence encoding for isoform 1 of the pig 3OST enzyme (entry tr|I3LHH5| I3LHH5_PIG, as illustrated in the sequence alignment in Figure 23), is aligned with the amino acid sequences of SEQ ID NO: 24, SEQ ID NO: 26, and SEQ ID NO: 28. Amino acid mutations that are present in SEQ ID NO: 24, SEQ ID NO: 26, and SEQ ID NO: 28 are engineered into their equivalent positions within the amino acid sequence of the pig 3OST enzyme, in order to generate mutant amino acid sequences SEQ ID NO: 52, SEQ ID NO: 53, and SEQ ID NO: 54, respectively.

[0464] In another non-limiting example, the full-length amino acid sequence encoding for the encoding for isoform 5 of the mouse 3OST enzyme (not shown in Figure 18) is aligned with the amino acid sequences of SEQ ID NO: 24, SEQ ID NO: 26, and SEQ ID NO: 28. Amino acid mutations that are present in SEQ ID NO: 24, SEQ ID NO: 26, and SEQ ID NO: 28 are engineered into their equivalent positions within the amino acid sequence of isoform 5 of the mouse 3OST enzyme, in order to generate mutant amino acid sequences. The generated full-length amino acid sequences are disclosed as SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58, respectively.

[0465] Enzymes comprising the amino acid sequences of SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58 respectively, will be utilized in Example 23 and Example 24, below. However, a person skilled in the art would appreciate that the same procedure can be applied to generate aryl sulfate-dependent mutants with respect to any of the other 3OST enzymes within the EC 2.8.2.23 enzyme class, and that those are omitted for clarity.

Example 23: Expression and Purification of EC 2.8.2.23 Mutants Having 3OST Activity

[0466] A study is conducted in accordance with embodiments of the present disclosure to determine whether genes encoding for engineered 3OST enzymes having the amino acid sequences SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58, respectively, can be transformed into host cells, and that enzymes comprising each of those amino acid sequences can be subsequently expressed, isolated, and purified according to the

procedure of Example 1, above. Codon-optimized nucleotide sequences are determined that encode for enzymes having the amino acid sequences of SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58, respectively, based on the desired expression host. Upon synthesizing or inserting those genes within a suitable expression vector, it is expected that genes encoding for each of the amino acid sequences SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58, respectively, will be transformed into host cells, and that enzymes containing those sequences will be subsequently expressed, isolated, and purified in a sufficient quantity and purity to determine aryl sulfate-dependent 3OST activity.

Example 24: 3OST Activity of EC 2.8.2.23 Mutants

[0467] A study is conducted in accordance with embodiments of the present disclosure to determine whether mutant enzymes comprising the sequences of SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58, respectively, are active sulfotransferases, using the procedures of Example 5 and/or Example 6. It is expected that MS and/or NMR studies will confirm the presence of NS2S6S3S-HS products formed as a result of reacting an NS2S6S/HS polysaccharide and an aryl sulfate compound with each of the engineered enzymes comprising the sequences of SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 57, and SEQ ID NO: 58, respectively.

We claim:

1. A pharmaceutical composition comprising *N*-sulfated, 6-*O*-sulfated heparan sulfate (NS6S/HS) polysaccharides that are structural analogs of *O*-desulfated heparin (ODSH), wherein each of the hexuronic acid and glucosamine residues within the NS6S/HS polysaccharides are fully unsulfated at their 2-*O* and 3-*O* positions, respectively.
2. The composition according to Claim 1, wherein the composition has 0 international units per milligram (IU/mg) of Anti-Factor Xa or Anti-Factor IIa activity.
3. The composition according to either Claim 1 or Claim 2, wherein the composition comprises no dermatan sulfate or chondroitin sulfate.
4. The composition according to any of Claims 1-3, wherein the weight-average molecular weight of the composition is in a range from at least 2,400 Da, up to 15,000 Da.
5. The composition according to any of Claims 1-4, wherein the composition comprises NS6S/HS polysaccharides having a 4,5-unsaturated uronic acid residue at the non-reducing end.
6. The composition according to Claim 5, wherein the composition comprises NS6S/HS polysaccharides having either a 1,6-anhydromannose residue or a 1,6-anhydroglucosamine residue at the reducing end.
7. The composition according to any of Claims 1-4, wherein the composition comprises NS6S/HS polysaccharides having a 2,5-anhydro-D-mannose residue at the reducing end.
8. The composition according to any of Claims 1-7, wherein the composition further comprises a hypomethylating agent, and the NS6S/HS is present in an amount effective to enhance the effectiveness of the hypomethylating agent in a subject with a myelodysplastic syndrome.
9. The composition according to Claim 8, wherein the hypomethylating agent is azacytidine, and the NS6S/HS has a weight average molecular weight in a range from at least 8,000 Da, up to 15,000 Da.
10. The composition according to any of Claims 1-7, wherein the NS6S/HS is present in an amount effective to promote thrombopoiesis in a subject with thrombocytopenia.
11. The composition according to any of Claims 1-7, wherein the NS6S/HS is present in an amount effective to promote neutrophil production in a subject with neutropenia.

12. The composition according to any of Claims 1-7, wherein the NS6S/HS is present in an amount effective to reduce apoptosis in a subject with a myocardial infarction, stroke, congestive heart failure, or cardiomyopathy.
13. The composition according to any of Claims 1-7, wherein the NS6S/HS is present in an amount effective to reduce the asthmatic response in a subject with asthma.
14. The composition according to any of Claims 1-7, wherein the NS6S/HS is present in an amount effective to reduce or inhibit the activity of neutrophil elastase or cathepsin G in a mammal.
15. The composition according to any of Claims 1-7, wherein the NS6S/HS is present in an amount effective to reduce or inhibit tumor growth in a subject with adenocarcinoma.
16. The composition according to any of Claims 1-7, wherein the NS6S/HS is present in an amount effective to reduce or prevent inflammation in a subject.
17. A method of enzymatically synthesizing an NS6S/HS product comprising NS6S/HS polysaccharides that are structural analogs of *O*-desulfated heparin (ODSH), wherein each of the hexuronic acid and glucosamine residues within the NS6S/HS polysaccharides are fully unsulfated at their 2-*O* and 3-*O* positions, respectively, and the method comprising the steps of:
- (a) providing a starting polysaccharide composition comprising *N*-deacetylated heparosan;
 - (b) reacting the starting polysaccharide composition within a reaction mixture comprising an *N*-sulfation agent, to form an *N*-sulfated heparan sulfate (NS/HS) product;
 - (c) reacting the NS/HS product within a reaction mixture comprising an aryl sulfate compound and an engineered 6OST enzyme, thereby forming the NS6S/HS product; and
- wherein the biological activity of the engineered 6OST enzyme consists of catalyzing the transfer of a sulfo group from an aryl sulfate compound to a heparosan-based polysaccharide in the absence of PAPS.
18. The method according to Claim 17, wherein the engineered 6OST enzyme comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 22; SEQ ID NO: 24; SEQ ID NO: 26; SEQ ID NO: 27; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 30; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 33; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 36; and SEQ ID NO: 37.

19 The method according to either of Claims 17 or 18, wherein the step of providing the starting polysaccharide composition comprises the following sub-steps:

- (i) providing a precursor polysaccharide composition comprising heparosan; and
- (ii) combining the precursor polysaccharide composition with a reaction mixture comprising a base, preferably lithium hydroxide or sodium hydroxide, for a time sufficient to *N*-deacetylate at least one of the *N*-acetylated glucosamine residues within the heparosan to form the starting polysaccharide composition, preferably for a time sufficient to form *N*-deacetylated heparosan polysaccharides in which at least 5% and up to 60%, more preferably at least 12% and up to 18%, and even more preferably 15%, of the glucosamine residues remain *N*-acetylated.

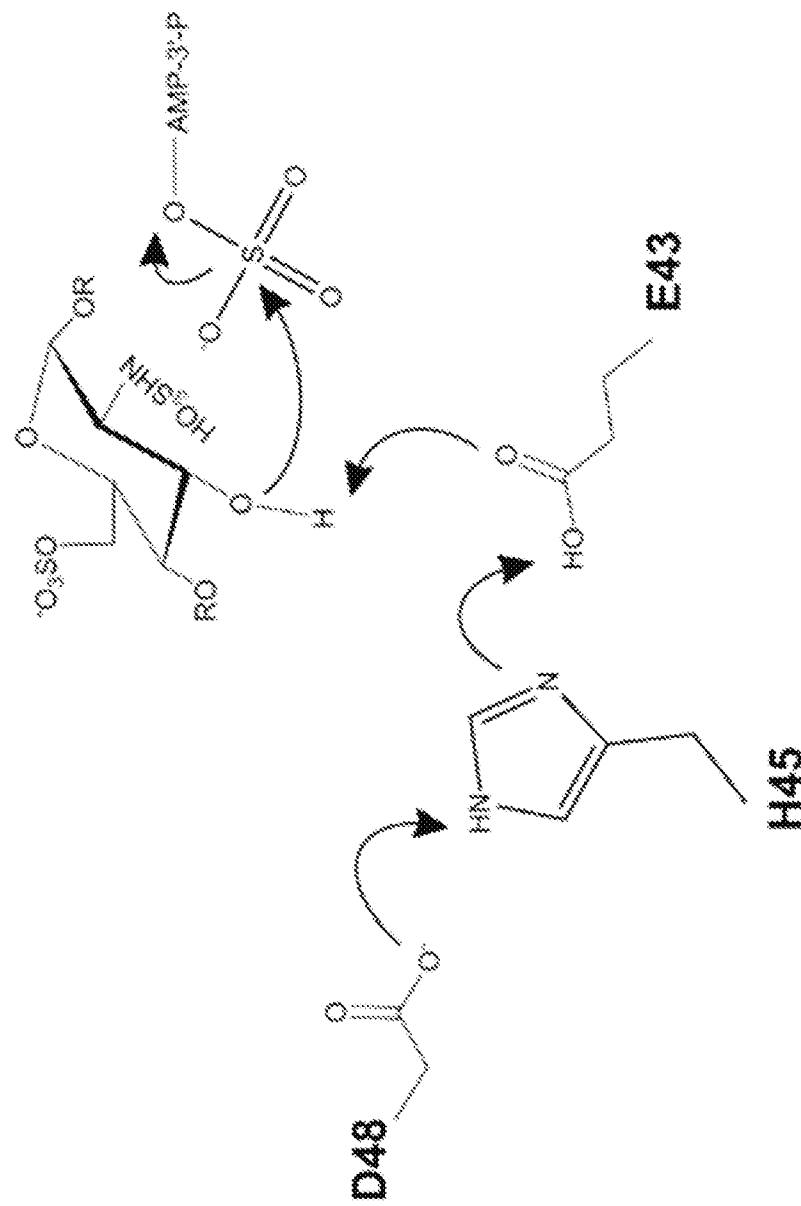
20. The method according to Claim 19, wherein the step of providing the precursor polysaccharide composition comprising heparosan further comprises the sub-step of isolating heparosan from a bacterial or eukaryotic cell culture, preferably a bacterial cell culture, and more preferably a bacterial cell culture comprising bacteria selected from the group consisting of the K5 strain of *Escherichia coli* and the BL21 strain of *Escherichia coli*.

21. The method according to any of Claims 1-5, wherein a sulfur trioxide-trimethylamine adduct is comprised within the *N*-sulfation agent reaction mixture.

22. The method according to any of Claims 1-5, wherein the *N*-sulfation agent is an engineered glucosaminyl *N*-sulfotransferase (NST) enzyme, wherein the biological activity of the engineered NST enzyme consists of the transfer of a sulfo group from an aryl sulfate compound to *N*-deacetylated heparosan, in the absence of PAPS.

23. The method according to Claim 22, wherein the engineered glucosaminyl *N*-sulfotransferase enzyme comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, and SEQ ID NO: 20.

24. The method according to any of Claims 17-23, wherein the aryl sulfate compound is selected from the group consisting of *p*-nitrophenyl sulfate and 4-nitrocatechol sulfate.
25. The method according to any of Claims 17-24, wherein the method further comprises providing a glucuronyl C5-epimerase enzyme and reacting the glucuronyl C5-epimerase enzyme with at least one of: the starting polysaccharide composition; the NS/HS product; and the NS6S/HS product.
26. The method according to any of Claims 17-25, wherein method steps (b) and (c) are performed in a single reaction vessel.
27. The method according to any of Claims 17-26, wherein the NS6S/HS product has no anticoagulant activity.
28. The method according to any of Claims 17-27, wherein the NS6S/HS product has a weight-average molecular weight in a range from about 2,400 Da to about 15,000 Da.
29. The use of the NS6S/HS product of any of Claims 17-28 in the treatment of a subject having a medical condition selected from the group consisting of: cancers, including adenocarcinoma and myelodysplastic syndrome; inflammation; heparin-induced thrombocytopenia; non-heparin-induced thrombocytopenia; neutropenia; apoptosis; asthma; emphysema; bronchitis; adult respiratory distress syndrome; cystic fibrosis; and ischemia-reperfusion related conditions, including myocardial infarctions and strokes.



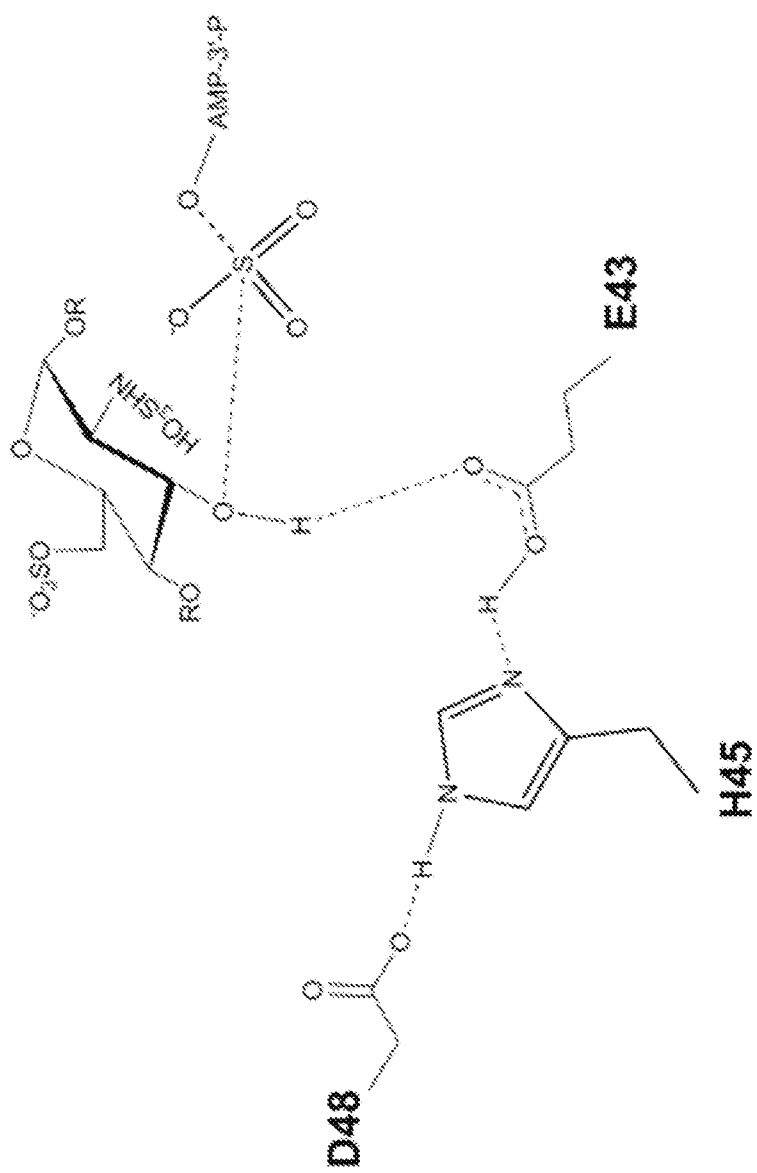
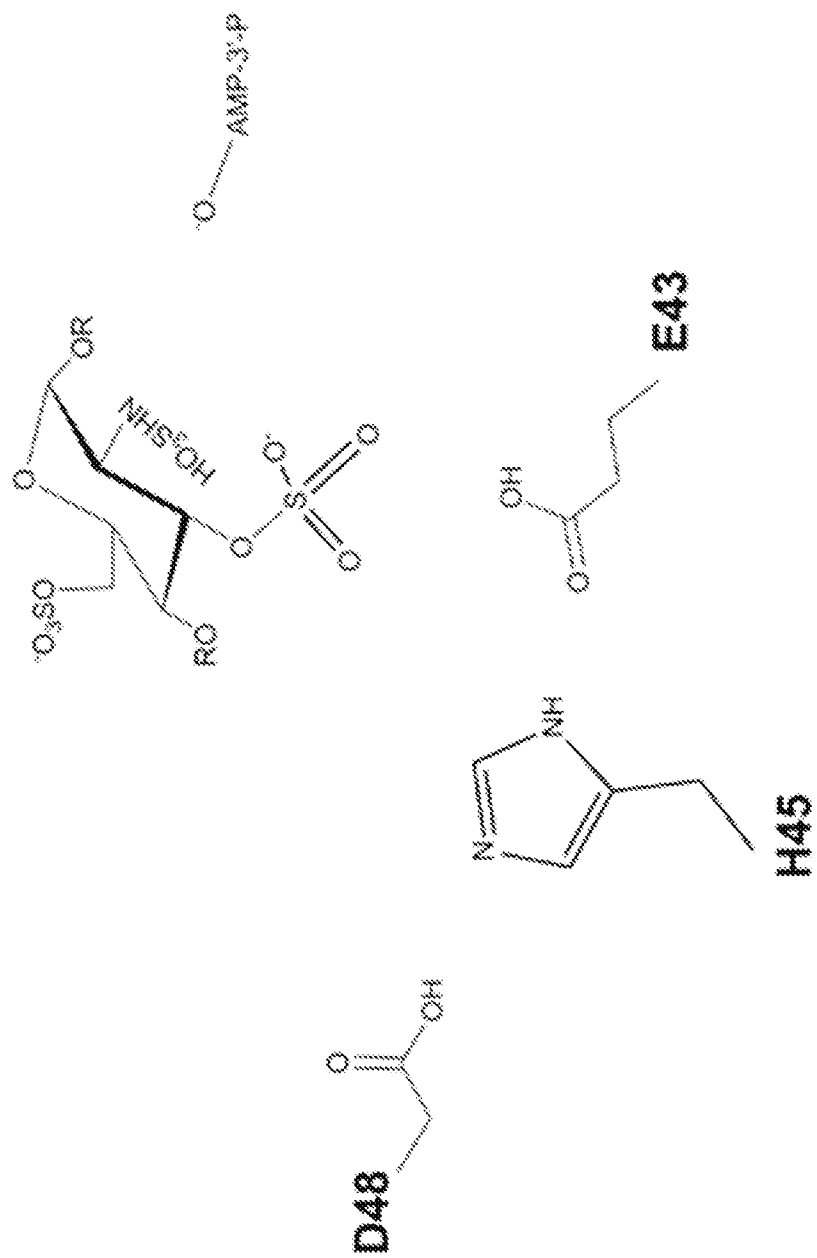


Figure 1B



Products

Figure 1C

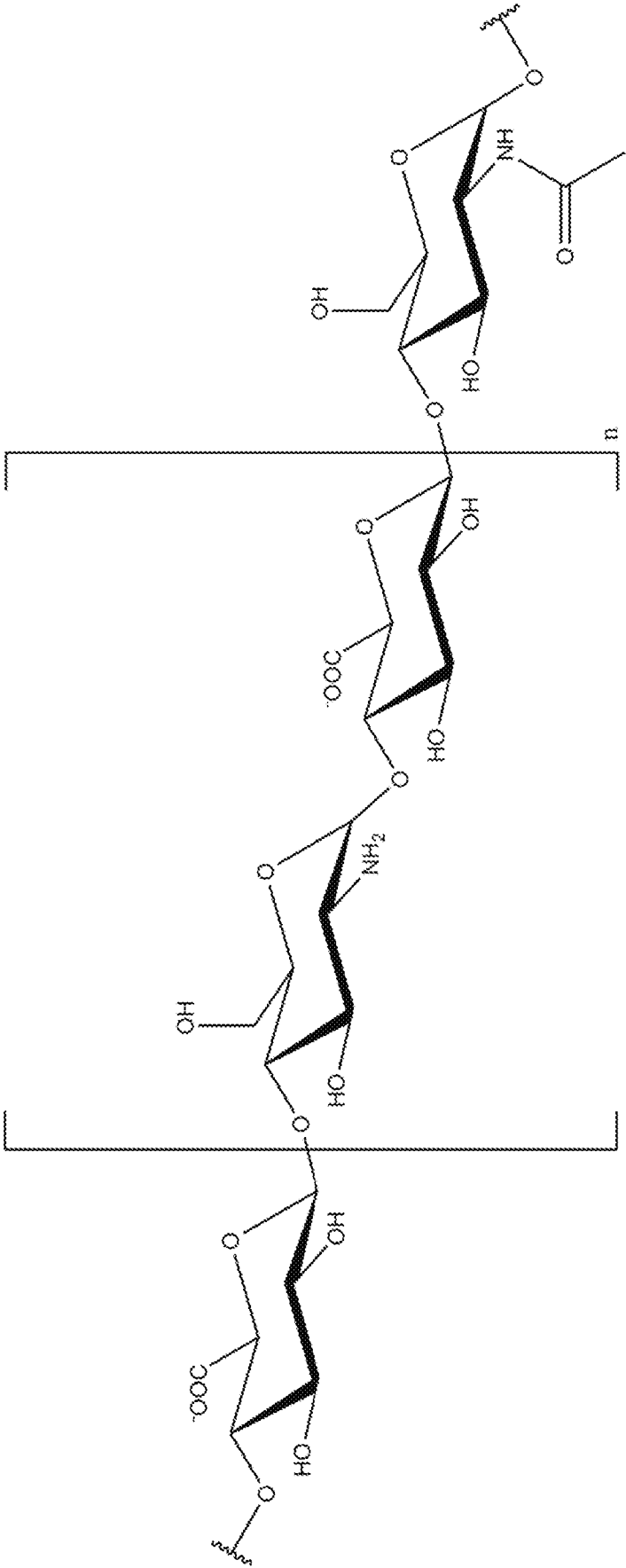


Figure 2

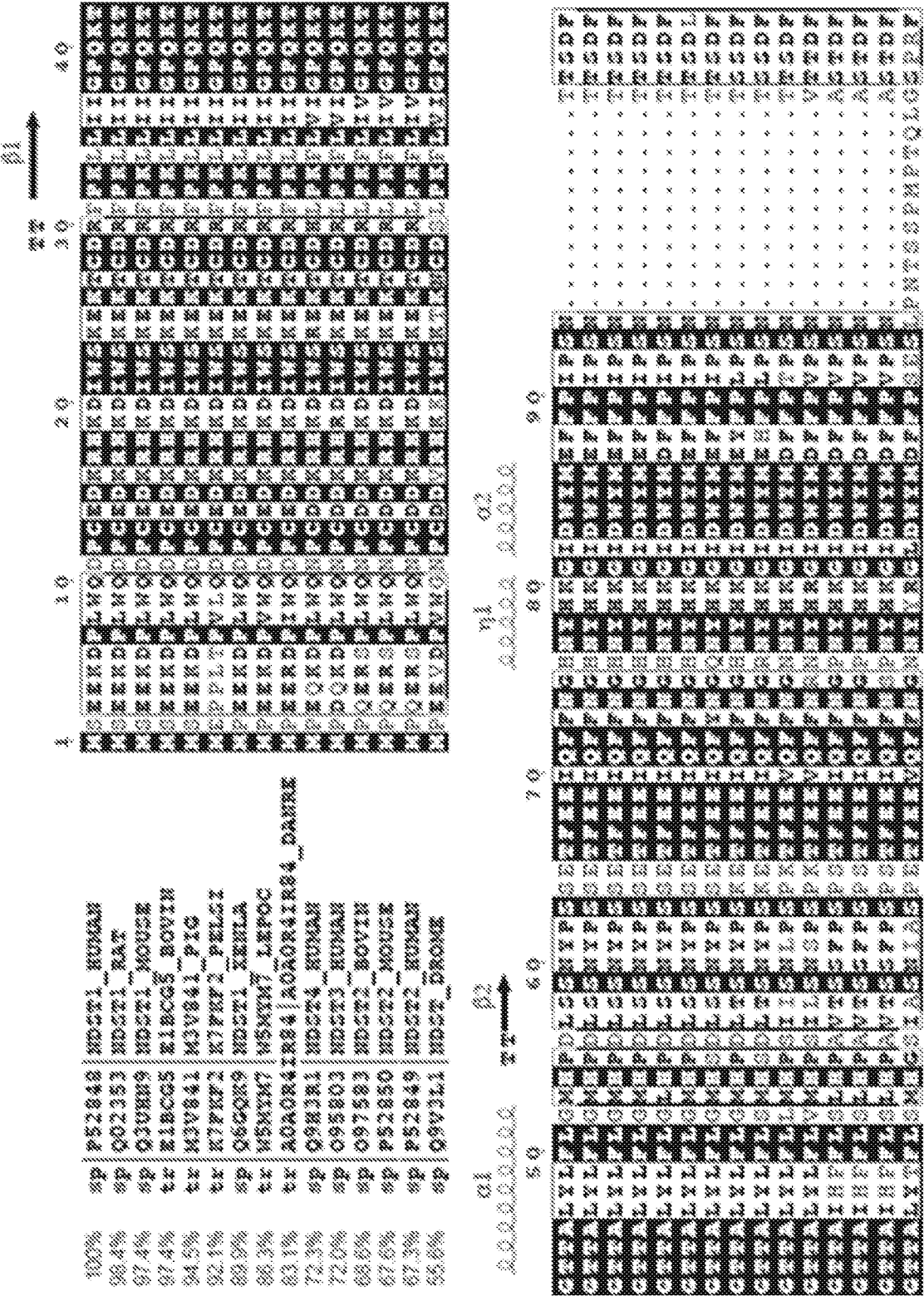


Figure 3A

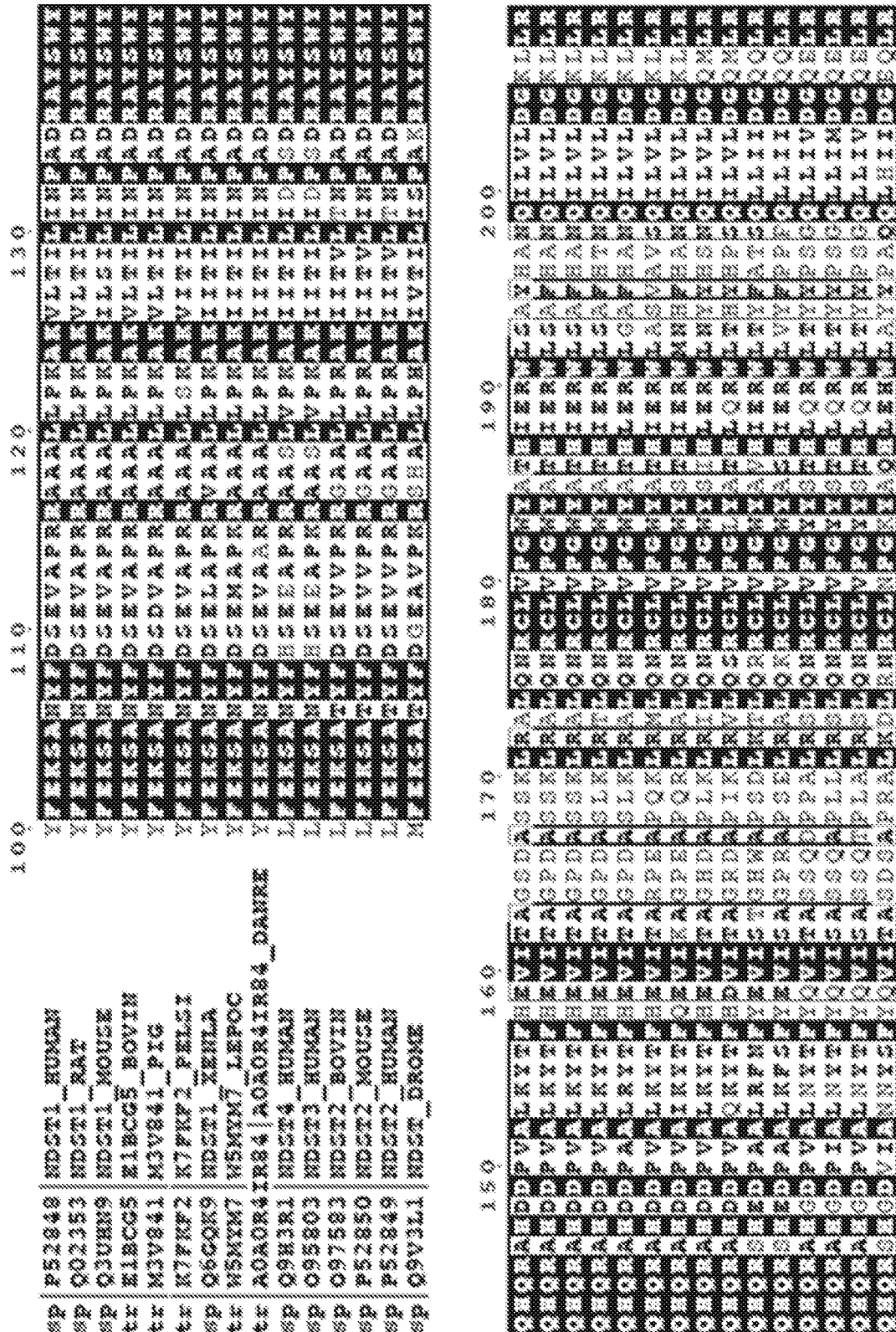


Figure 3B

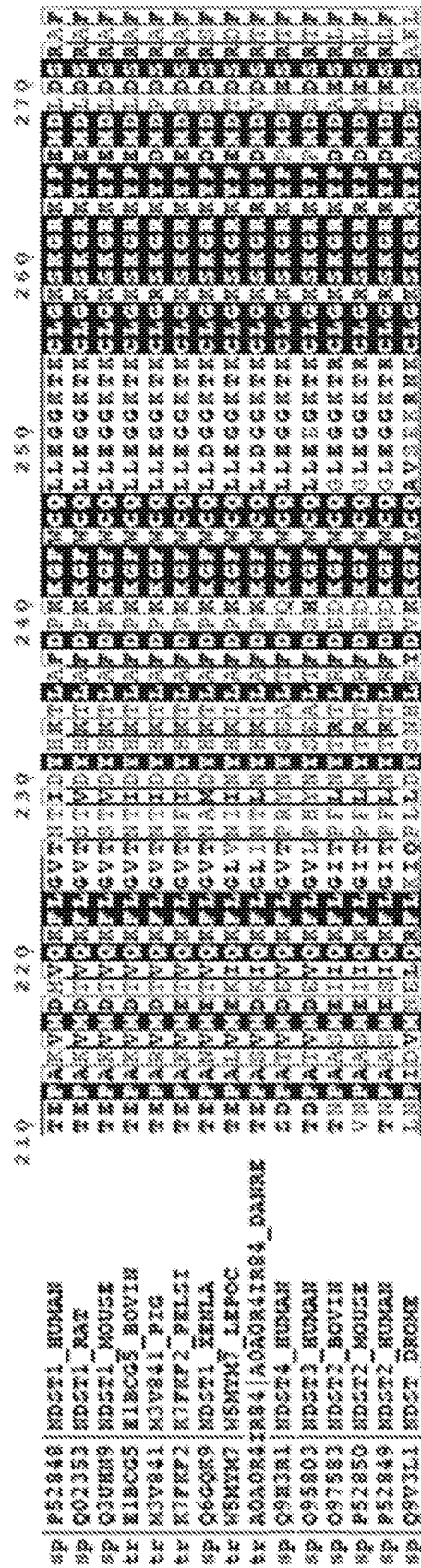
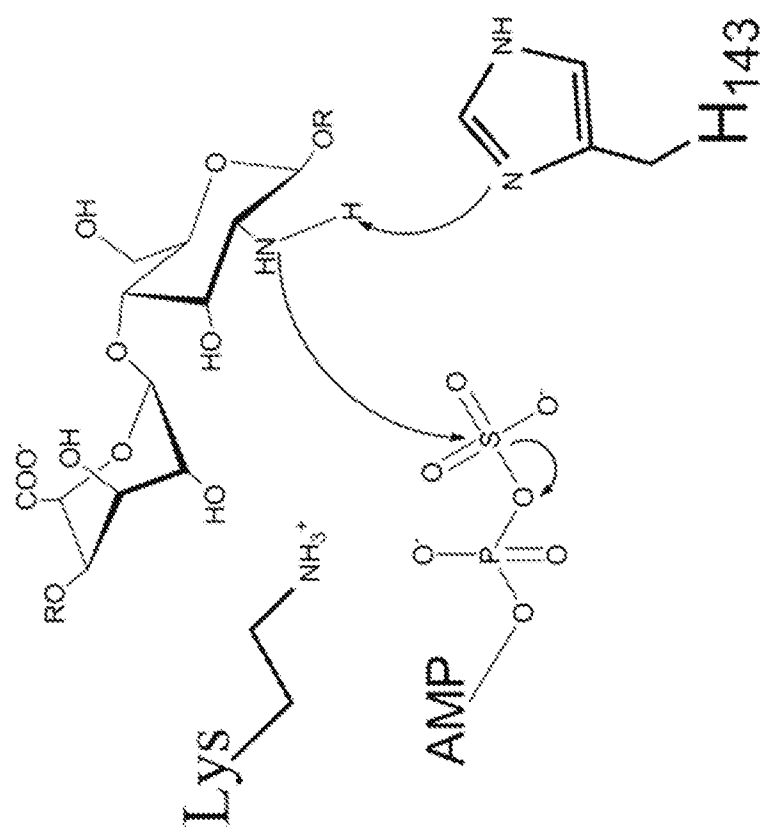
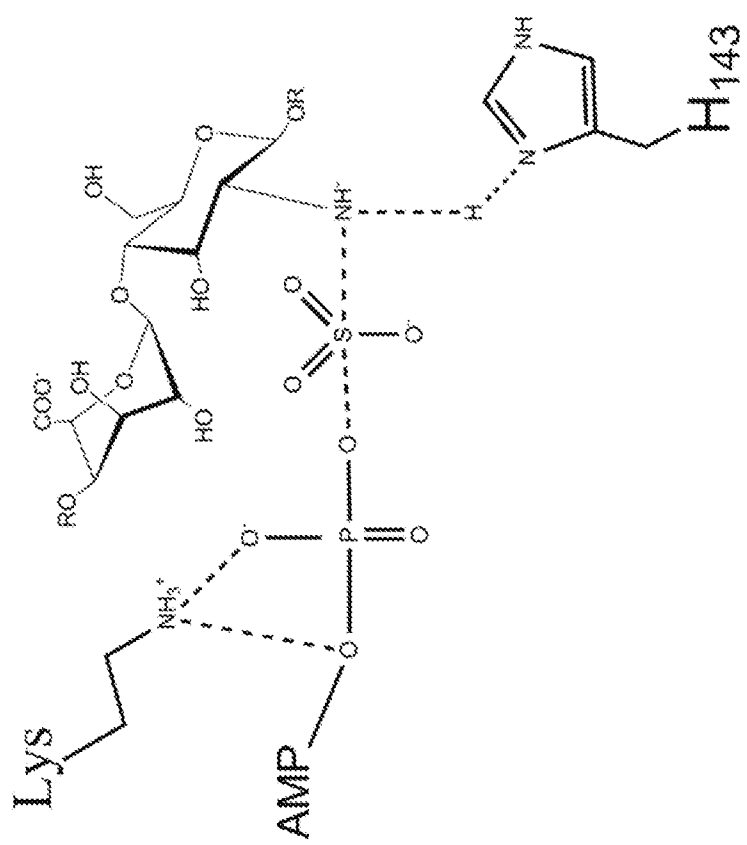


Figure 3C



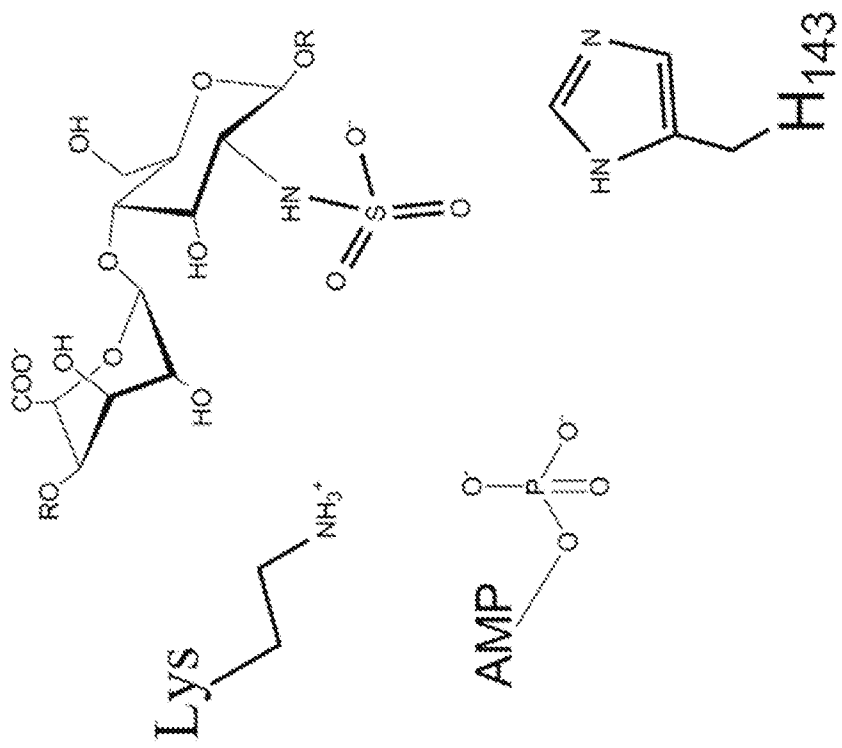
Reaction

Figure 4A



Transition State

Figure 4B



Products

Figure 4C

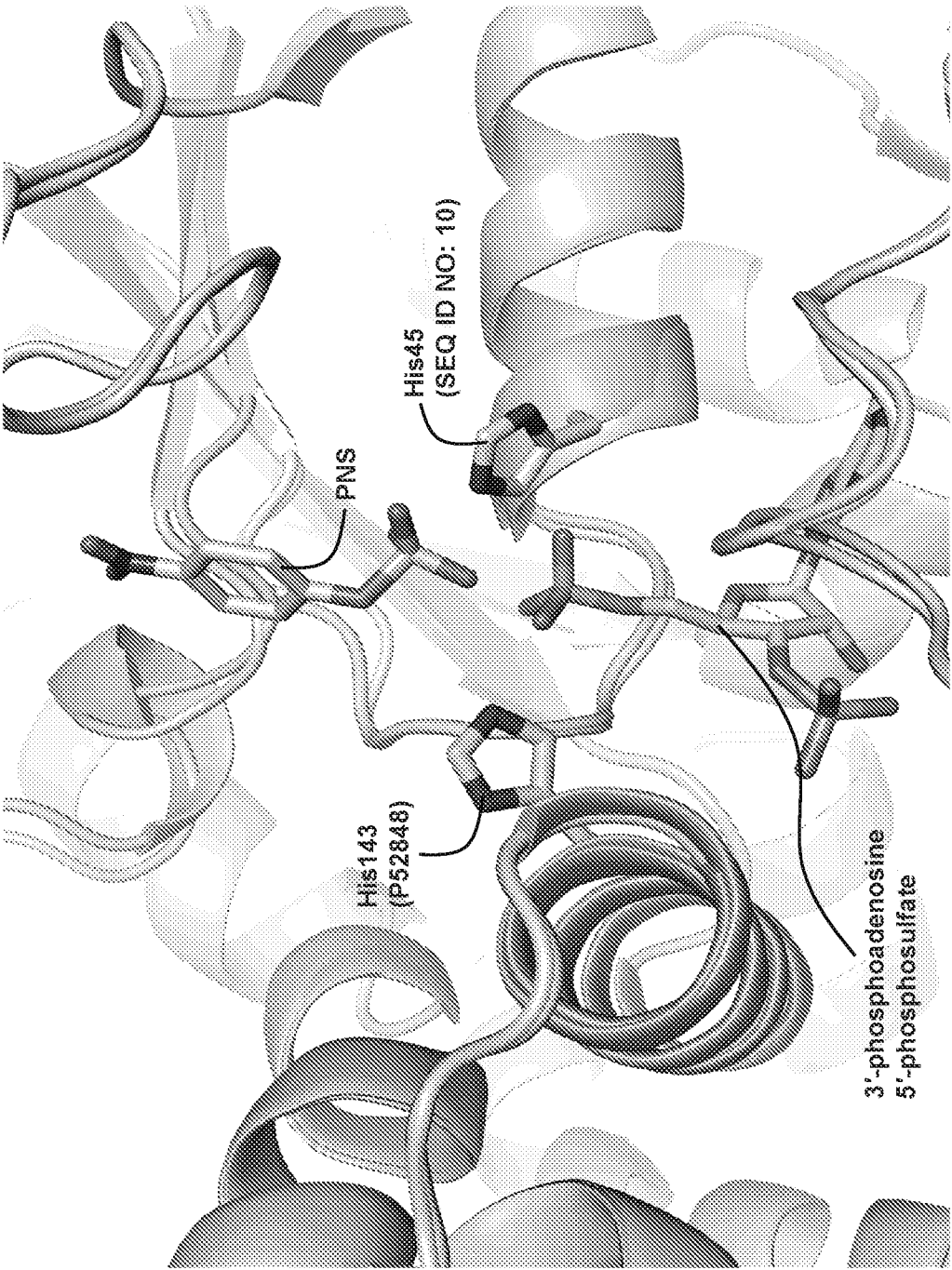


Figure 5

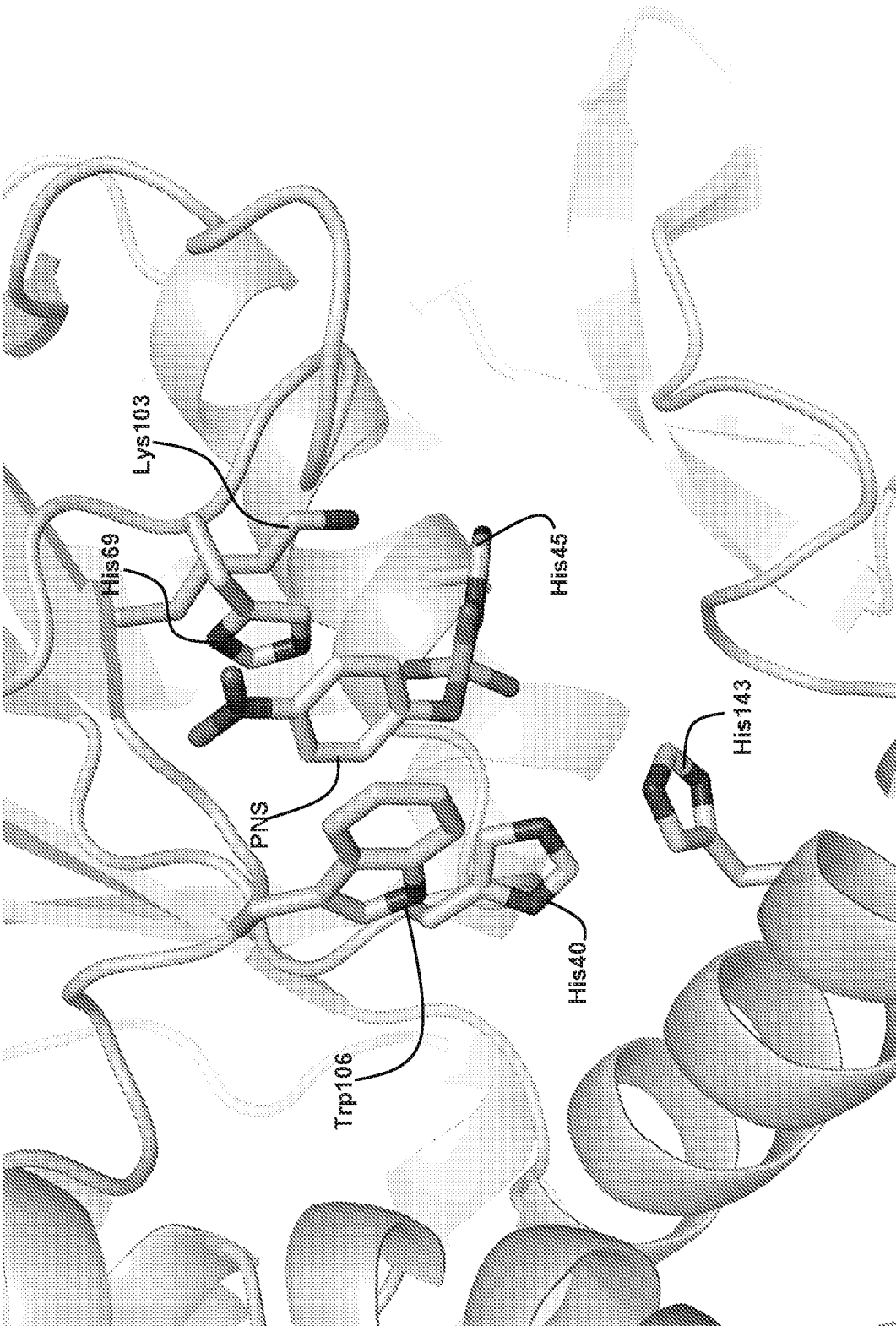


Figure 6

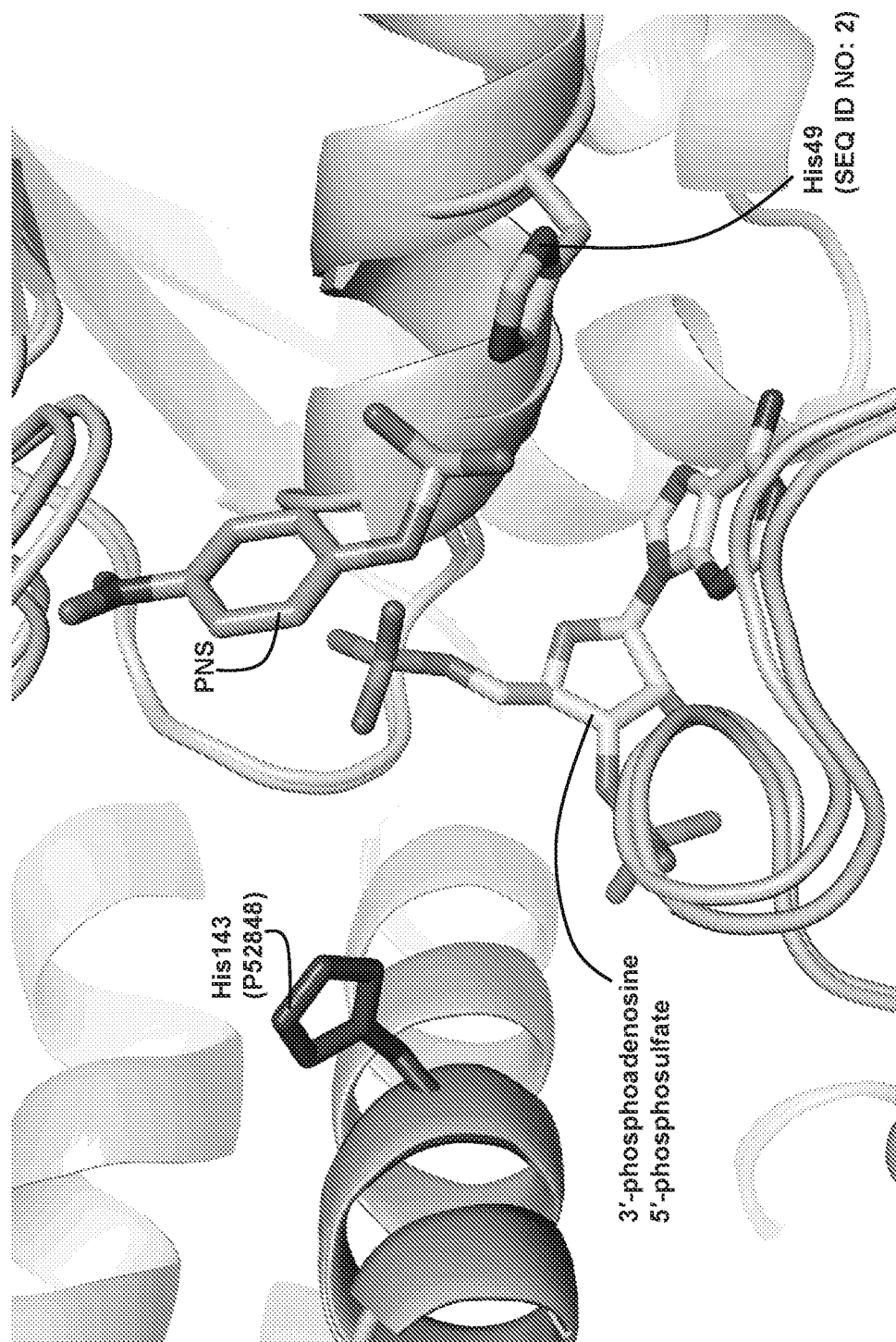


Figure 7

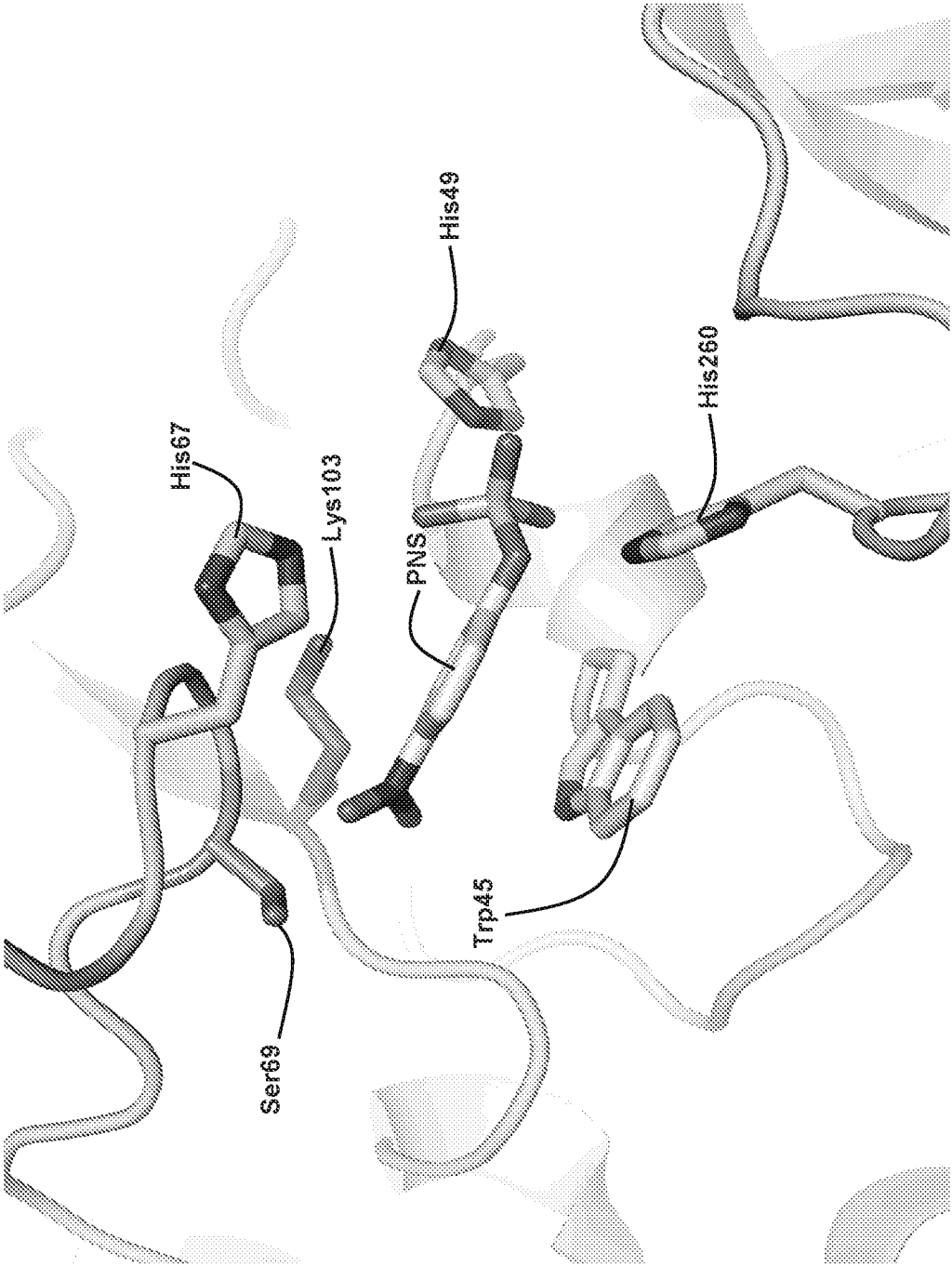


Figure 8

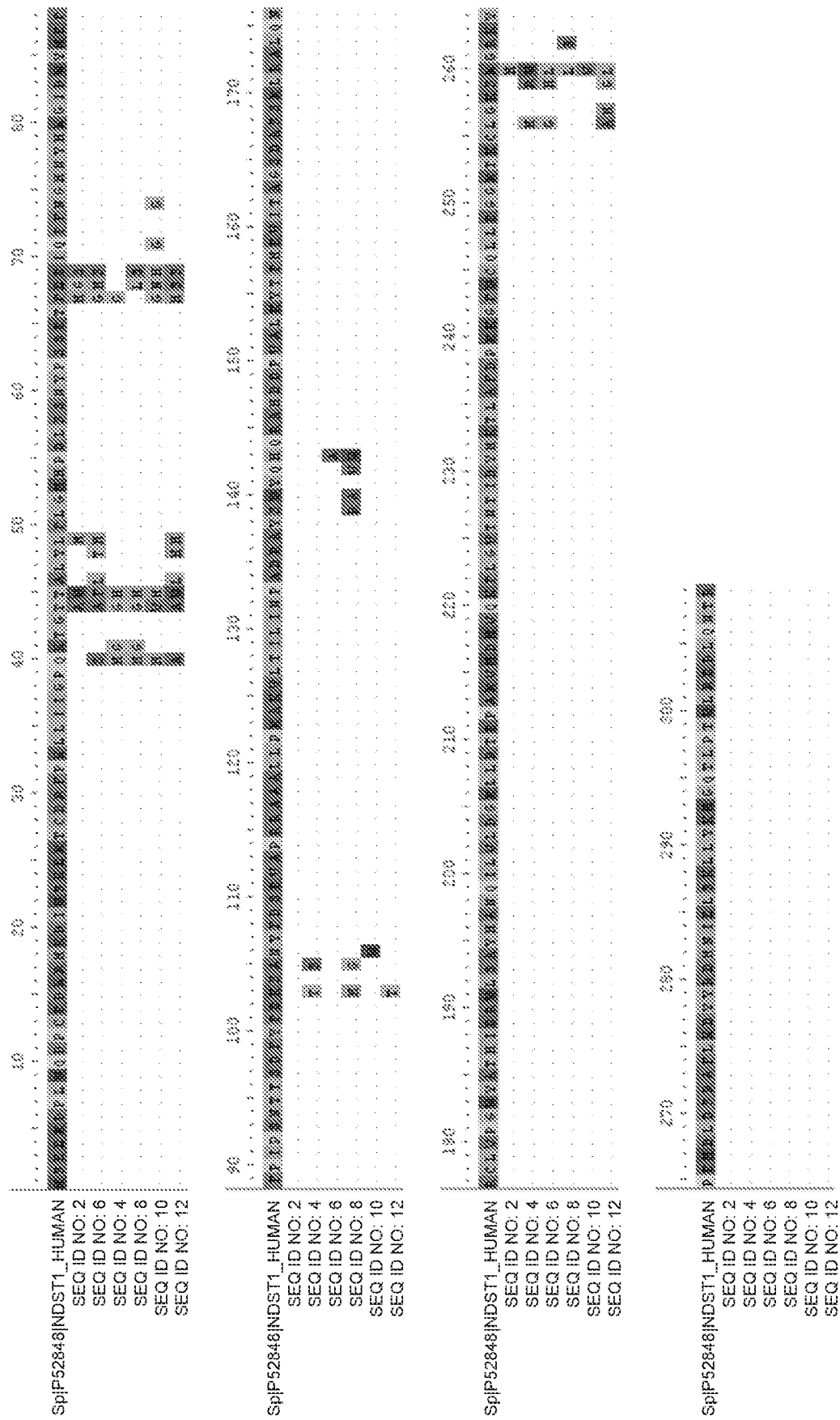


Figure 9

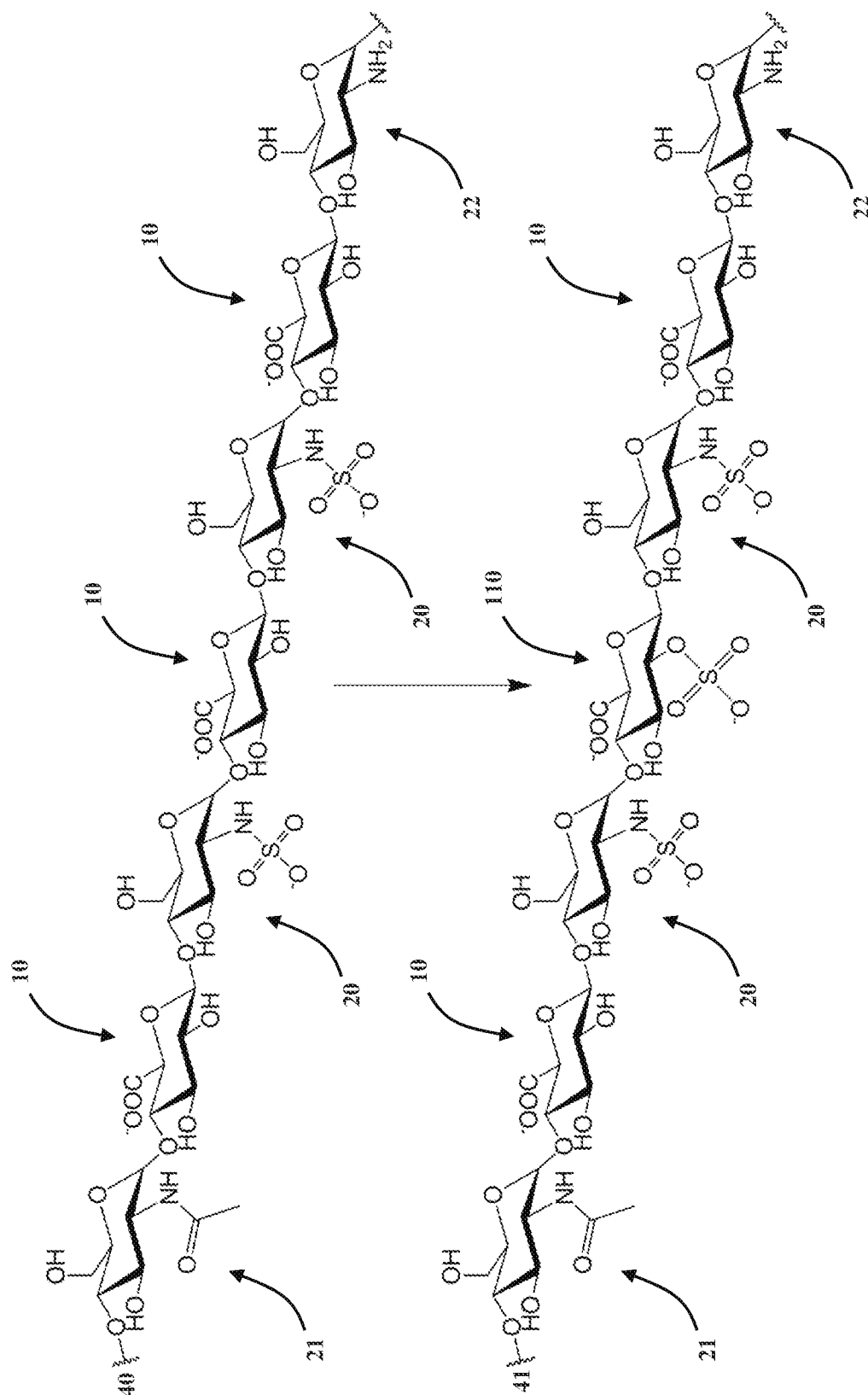


Figure 10

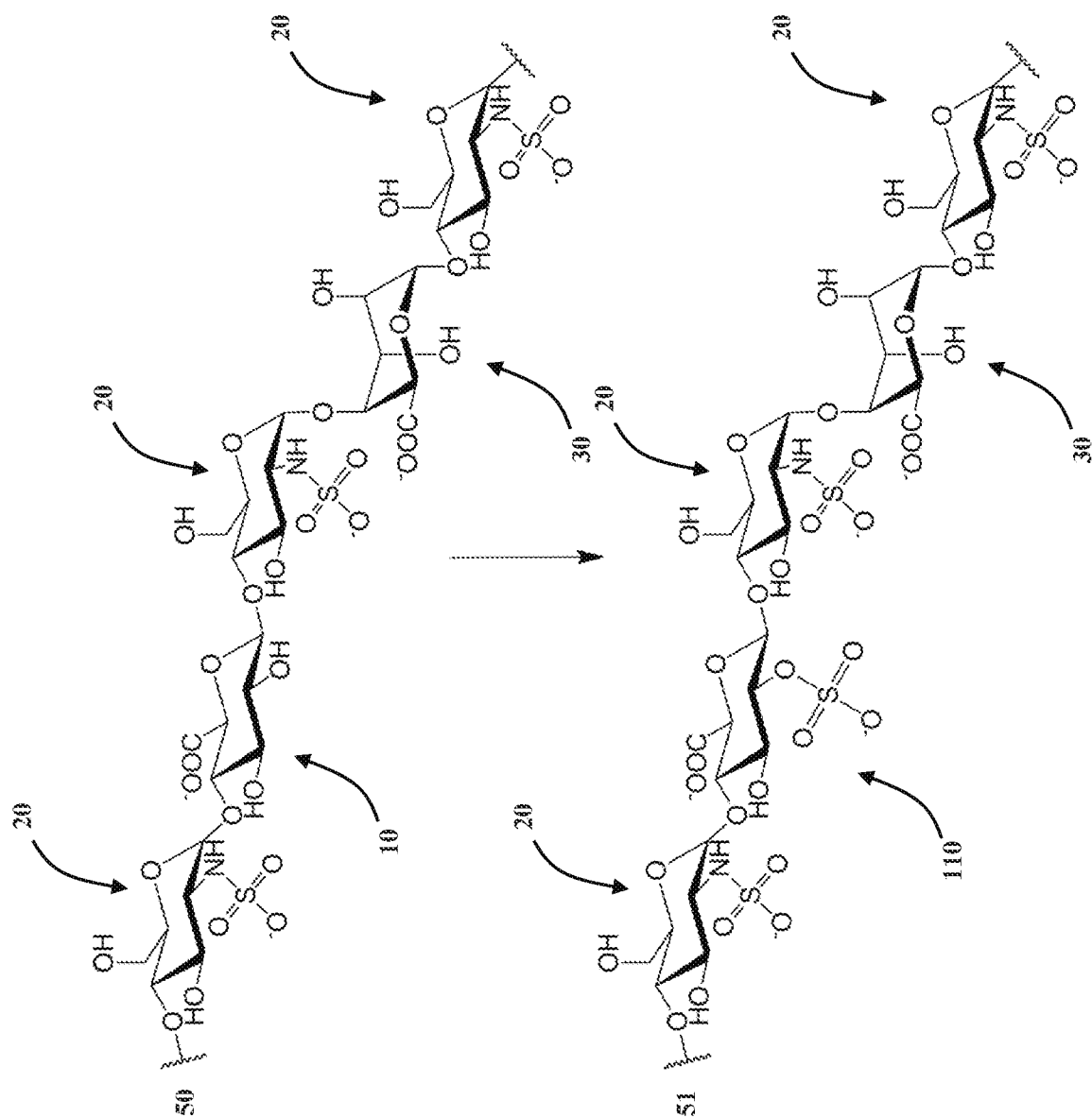


Figure 11

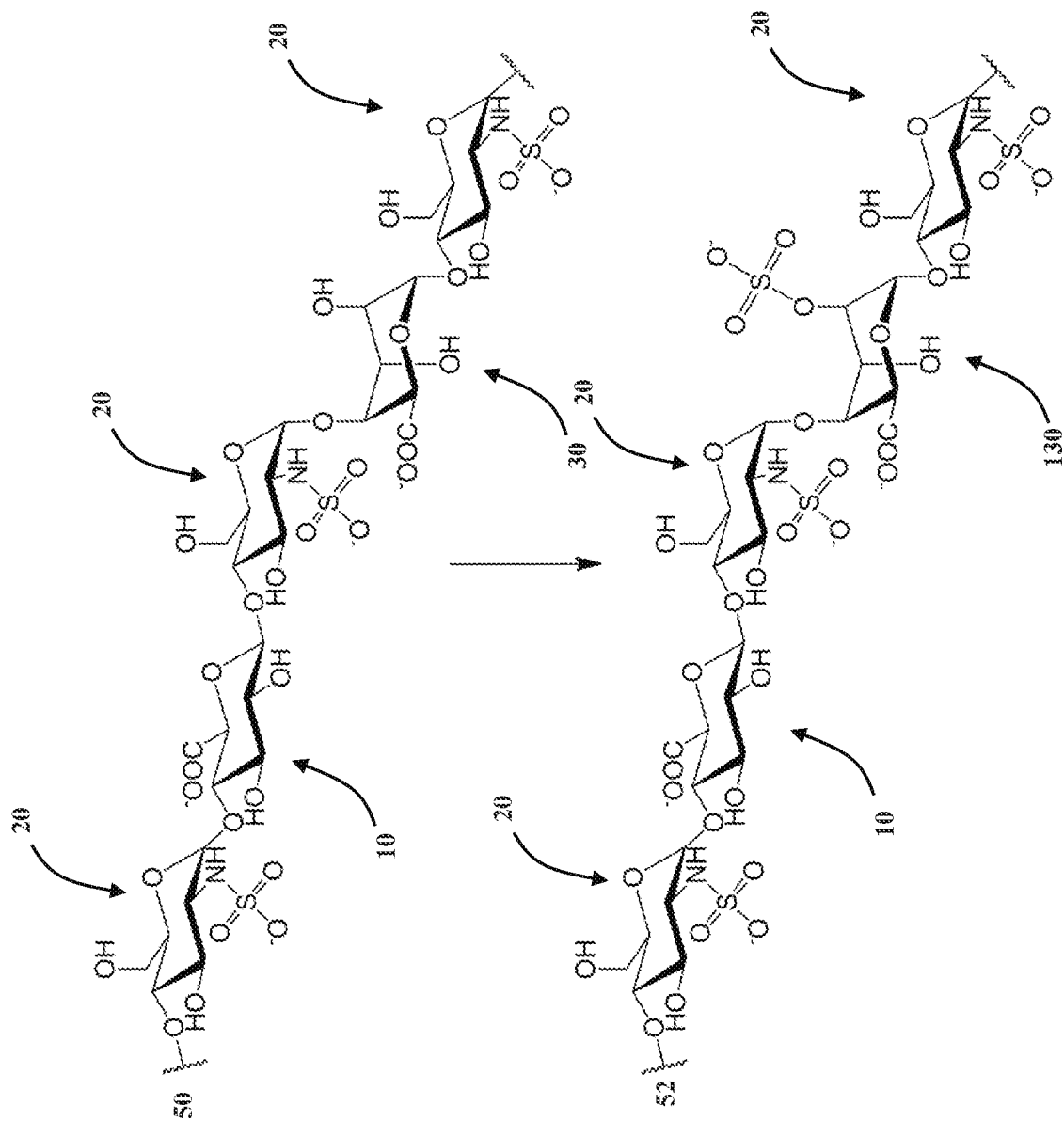


Figure 12

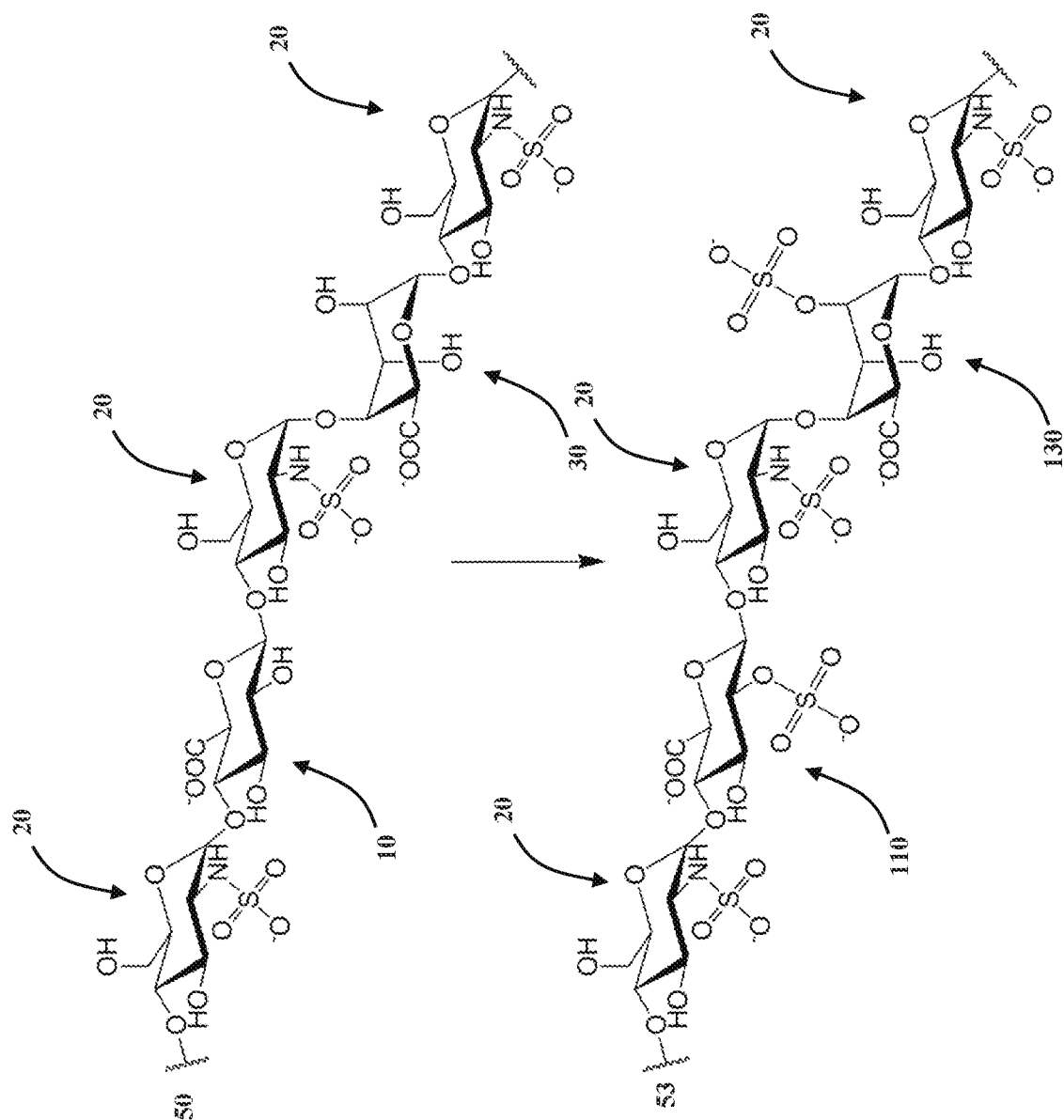


Figure 13

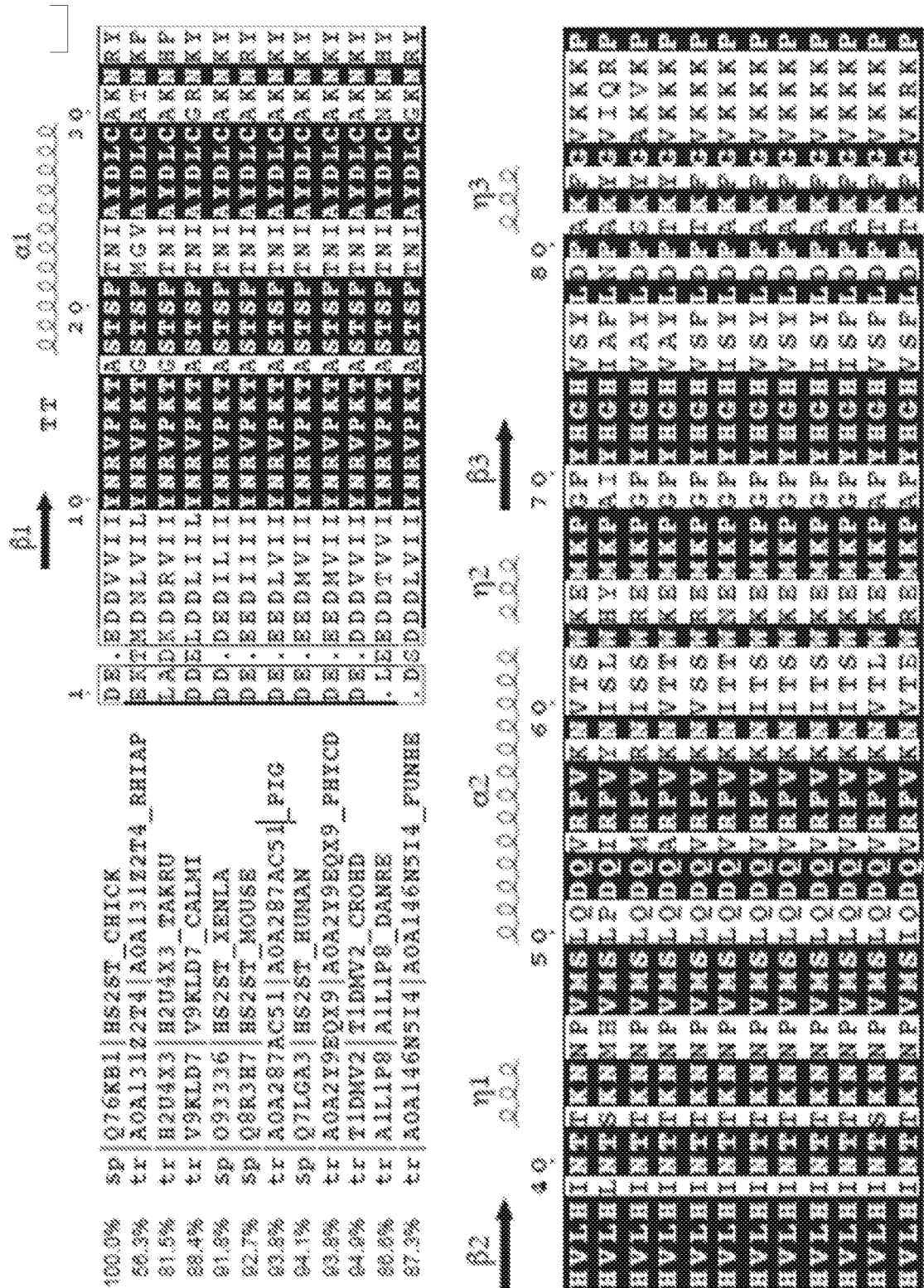


Figure 14A

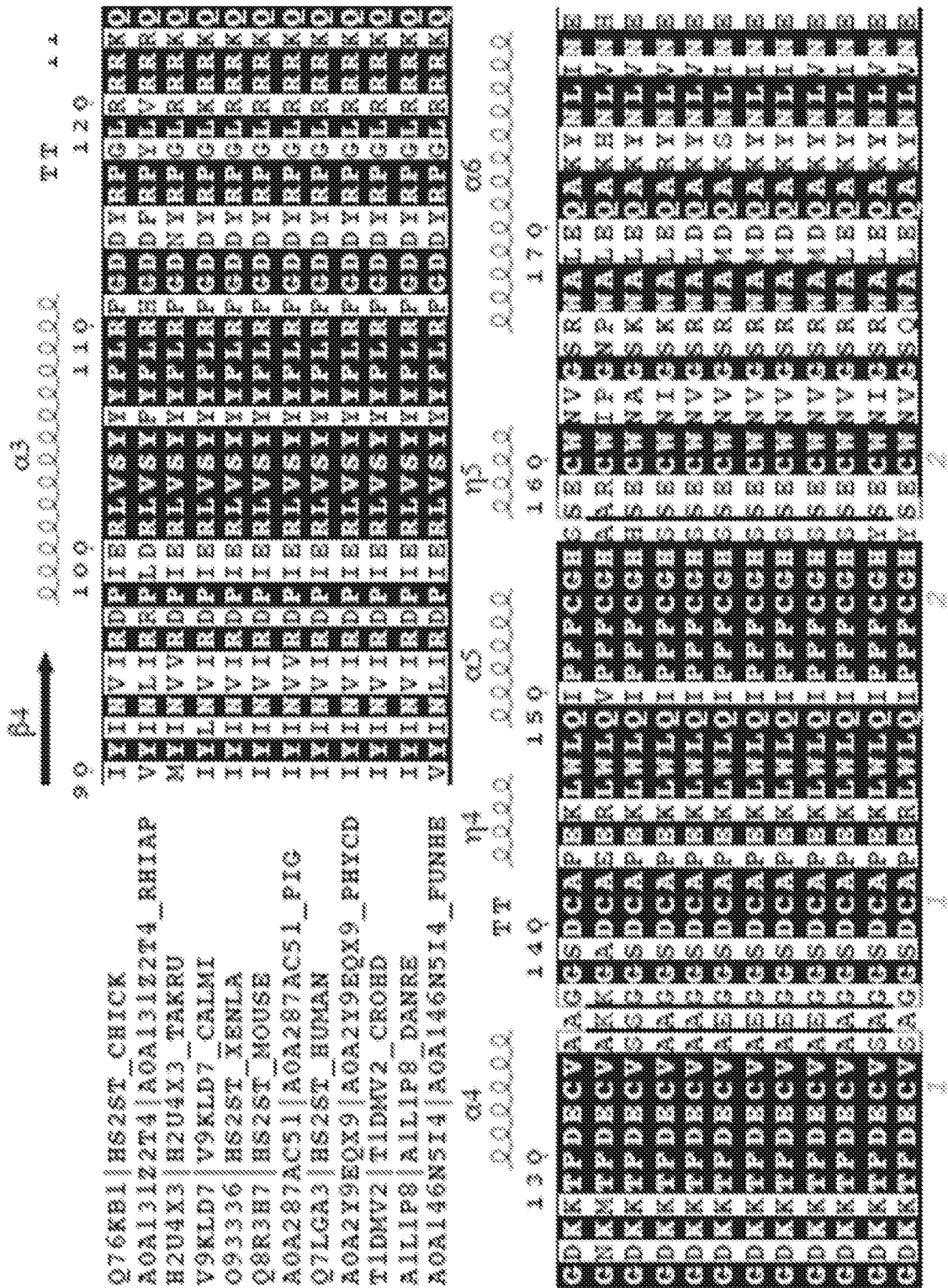
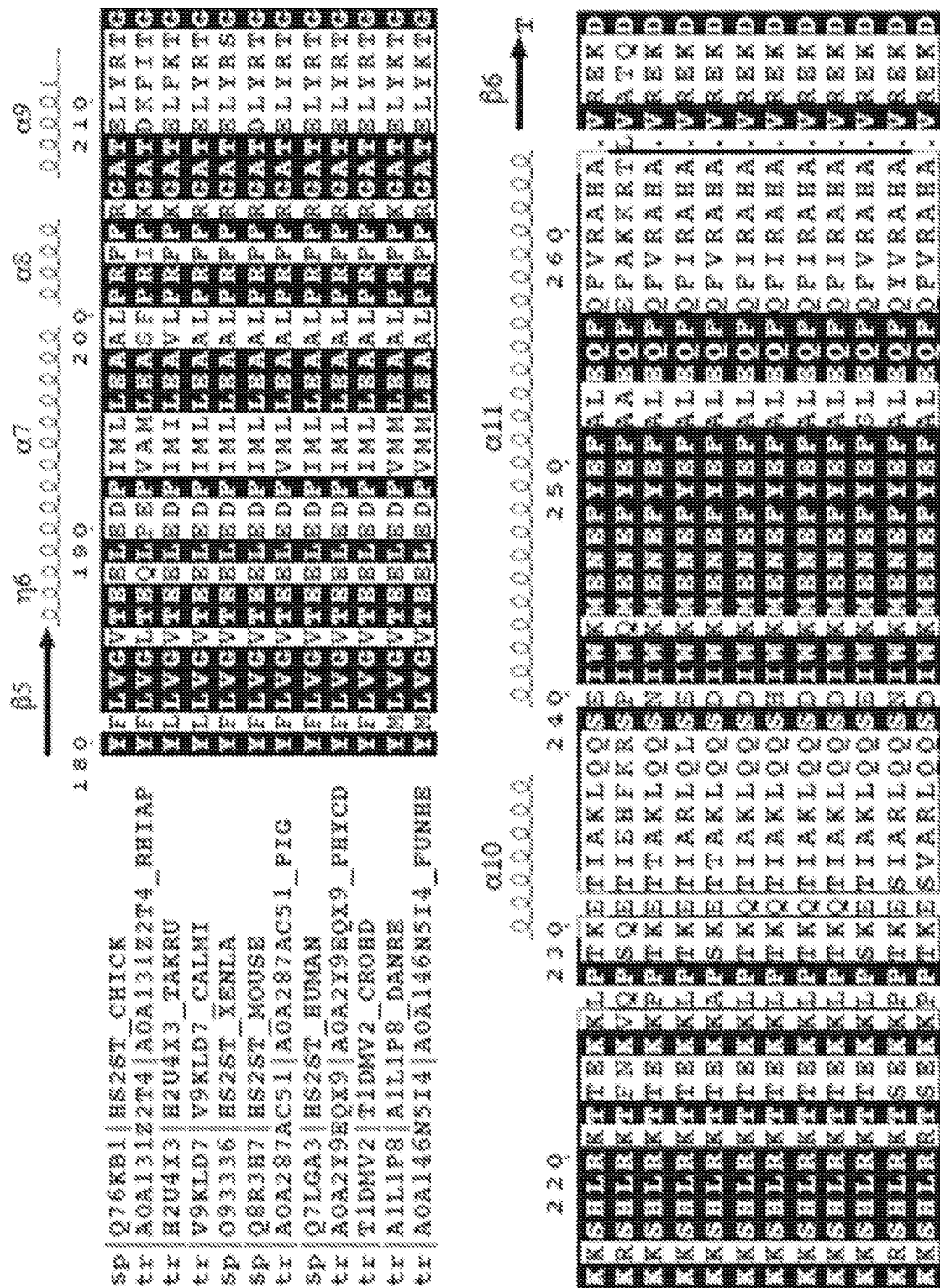


Figure 14B



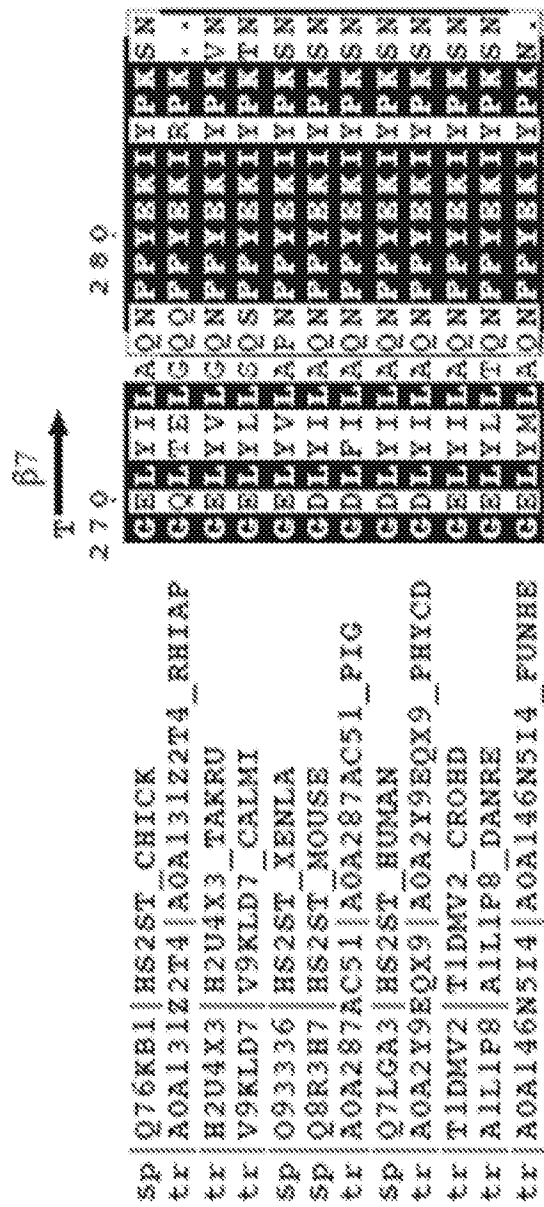
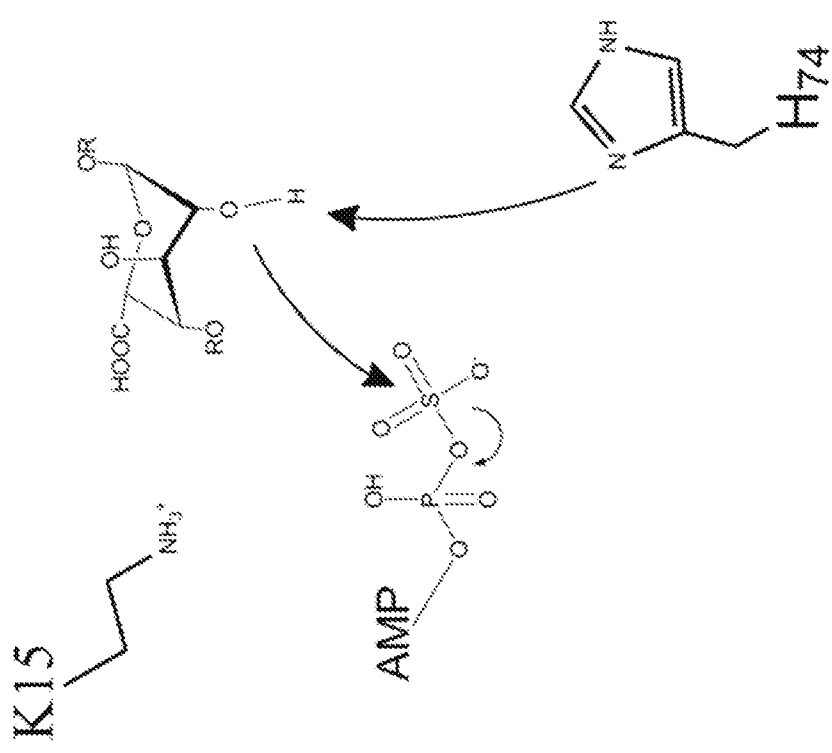
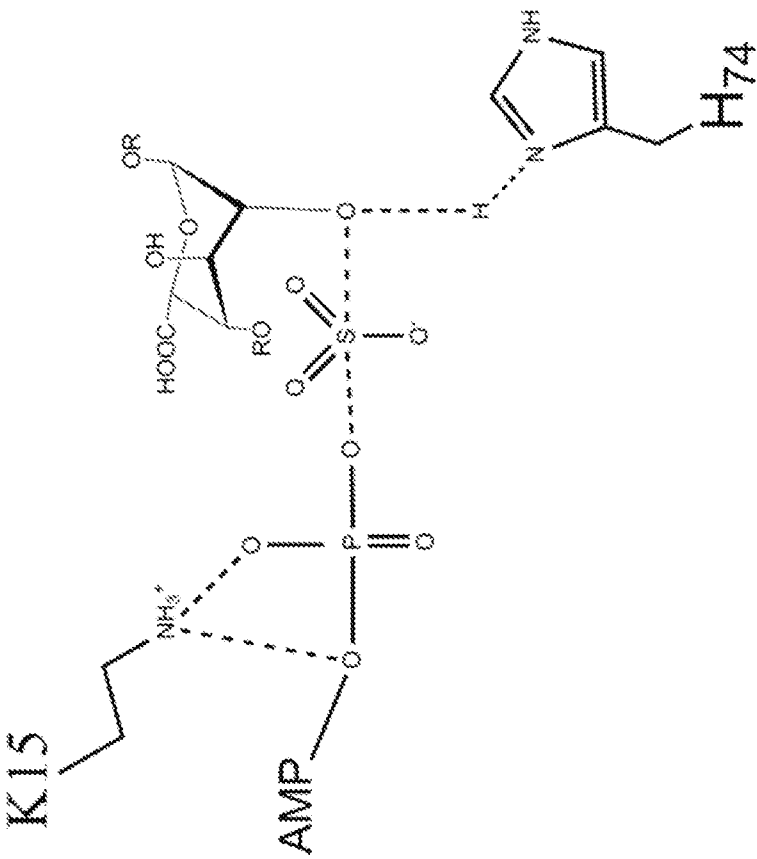


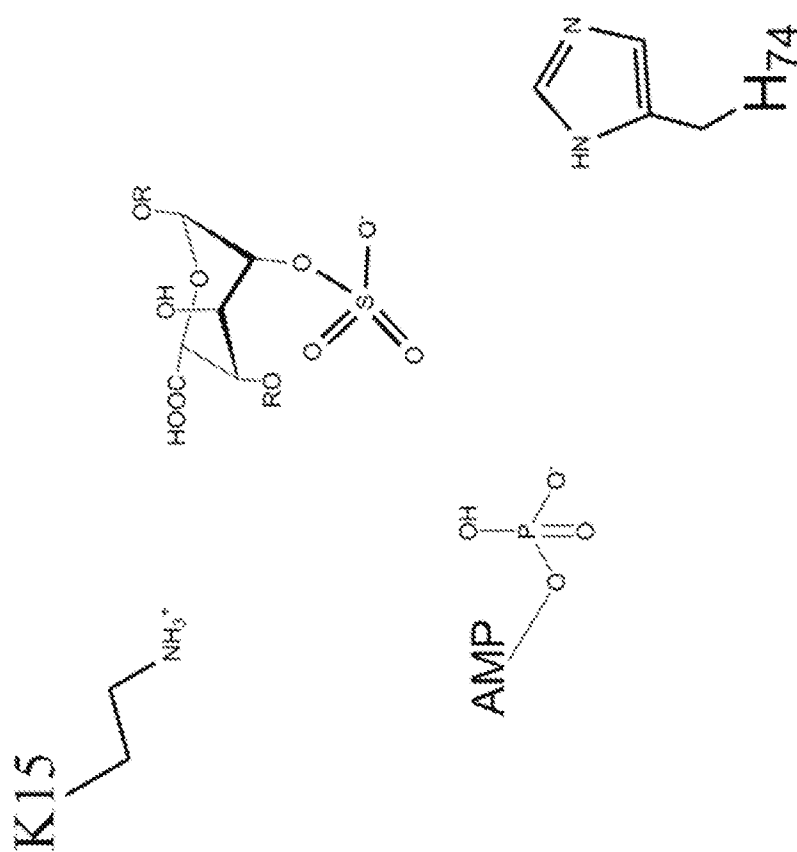
Figure 14D





Transition State

Figure 15B



Products

Figure 15C

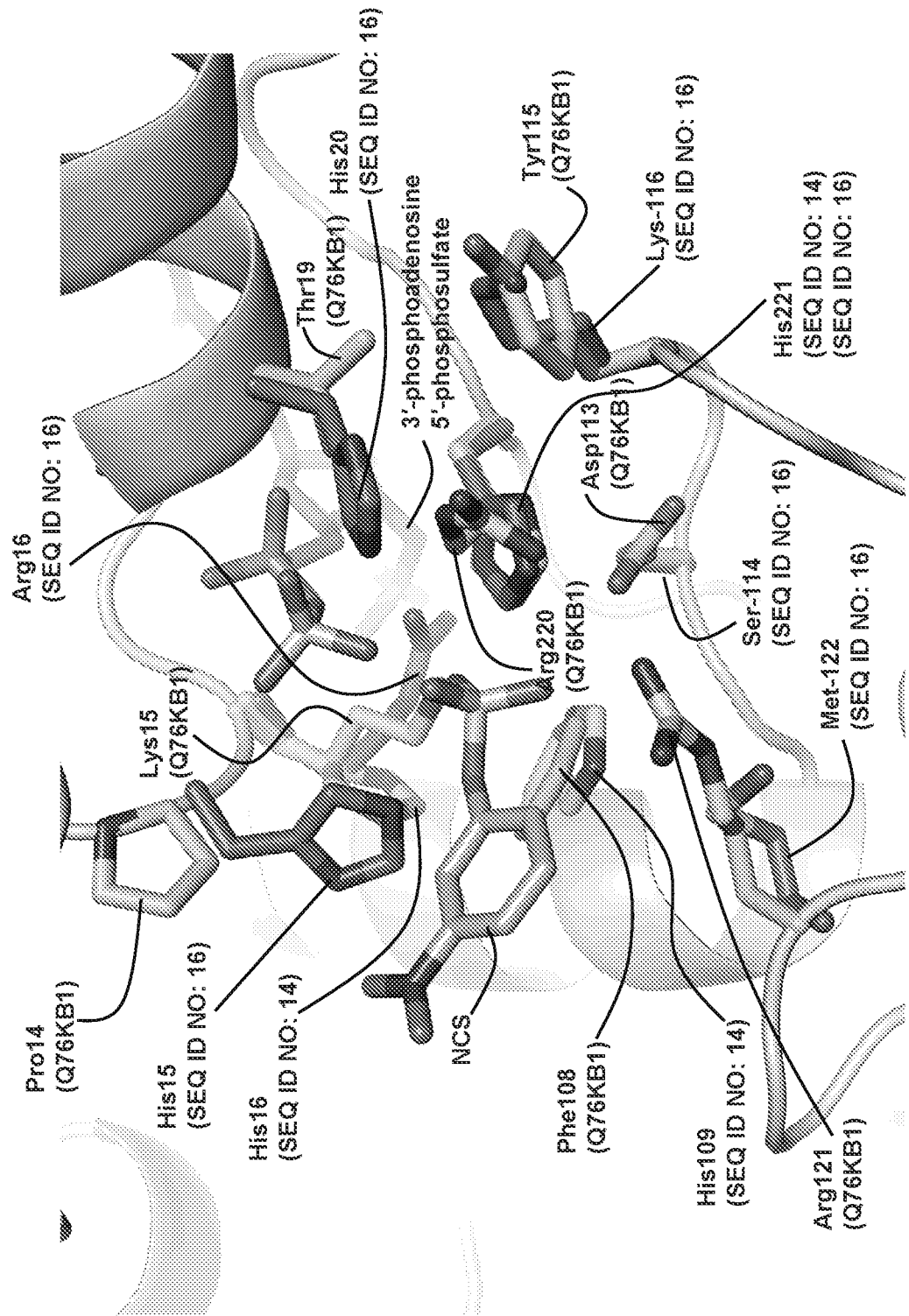


Figure 16

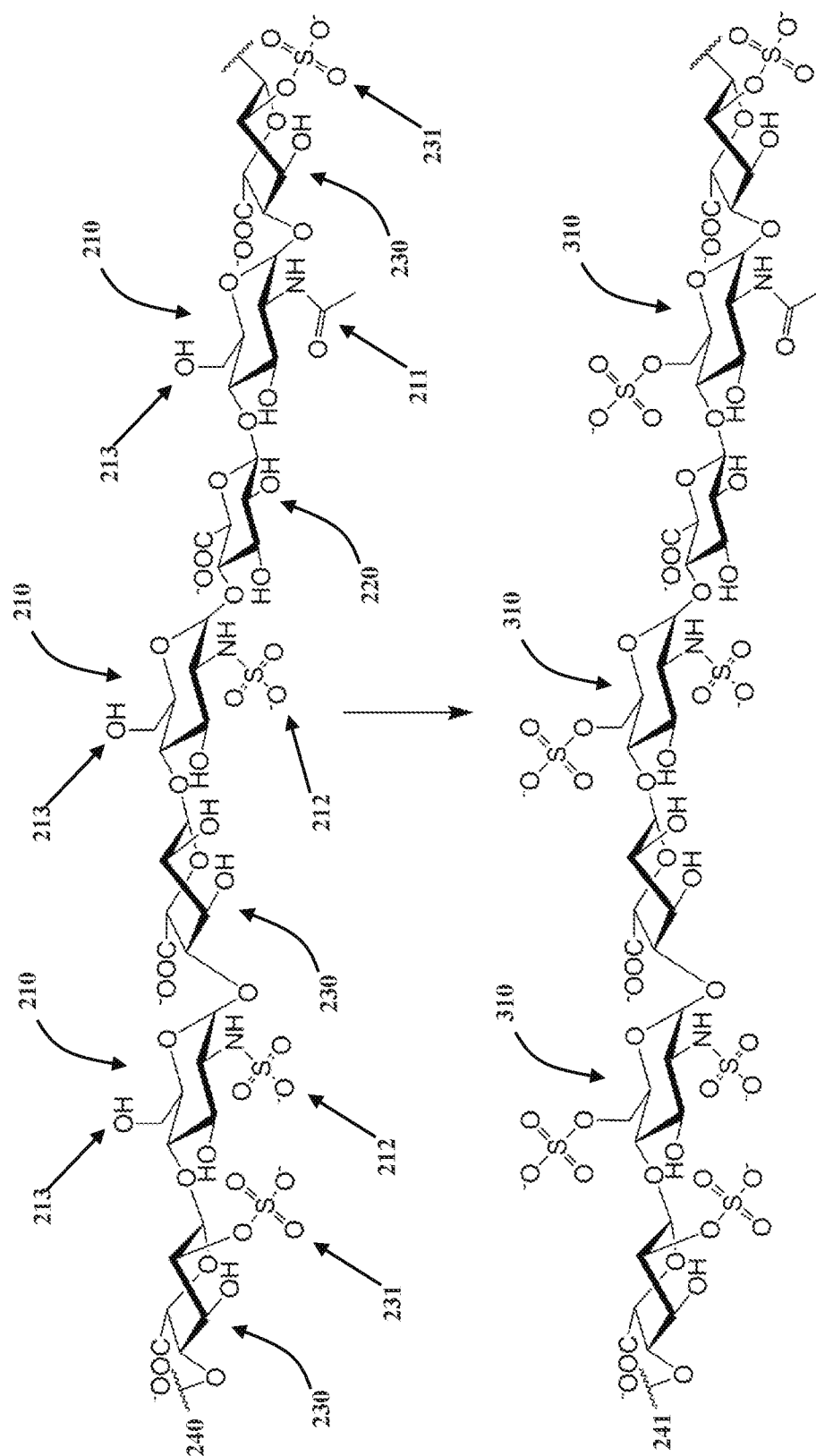
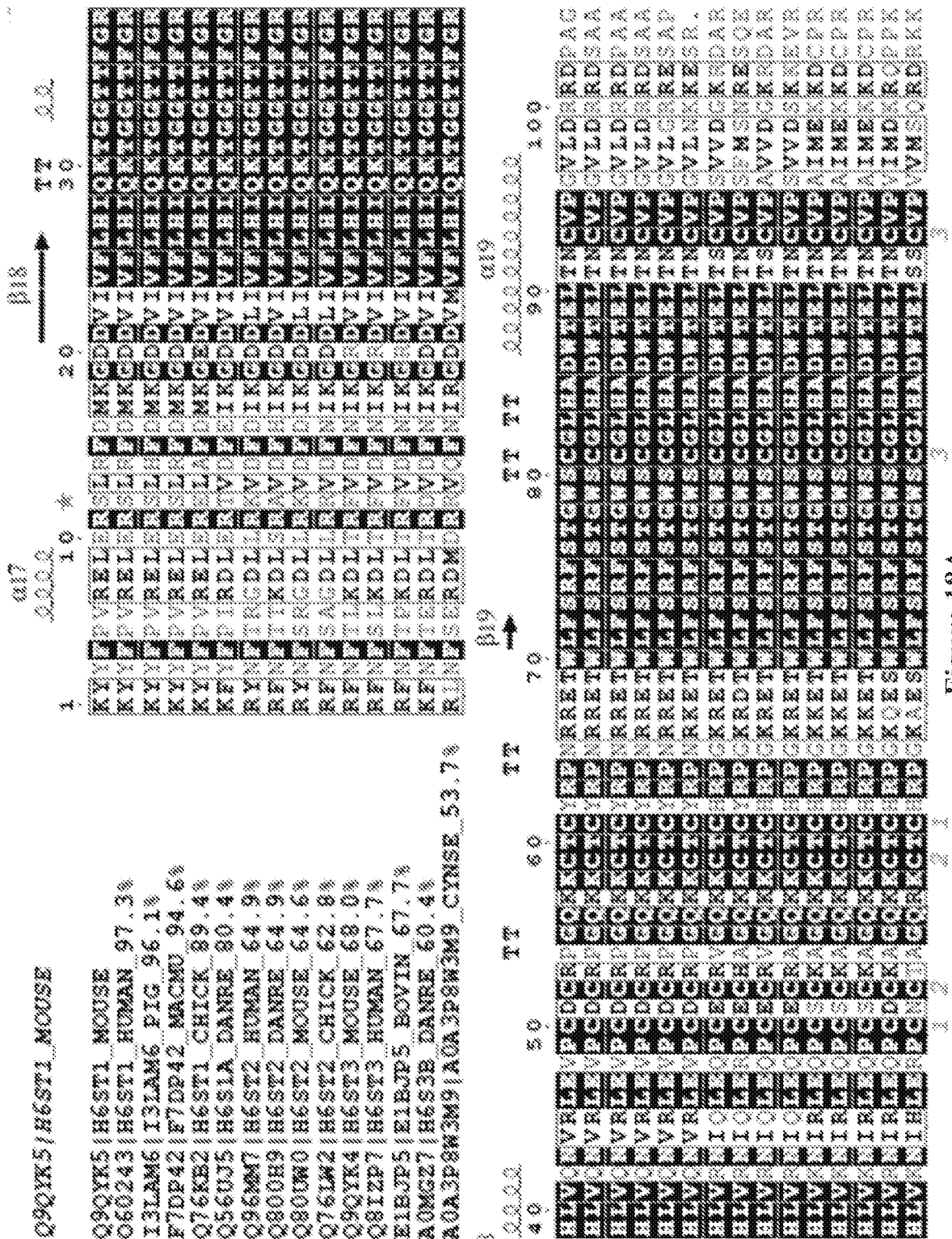


Figure 17



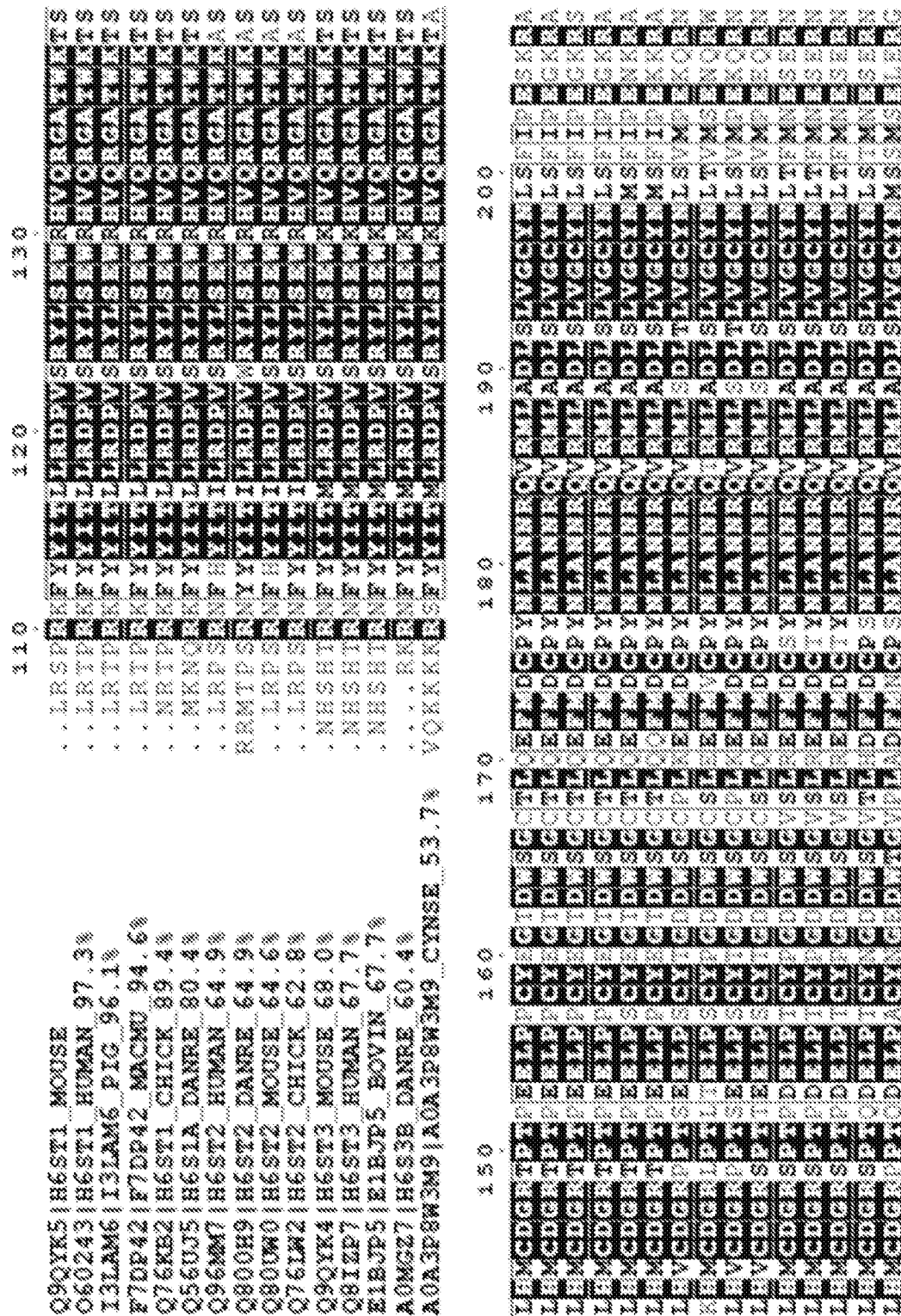


Figure 18B

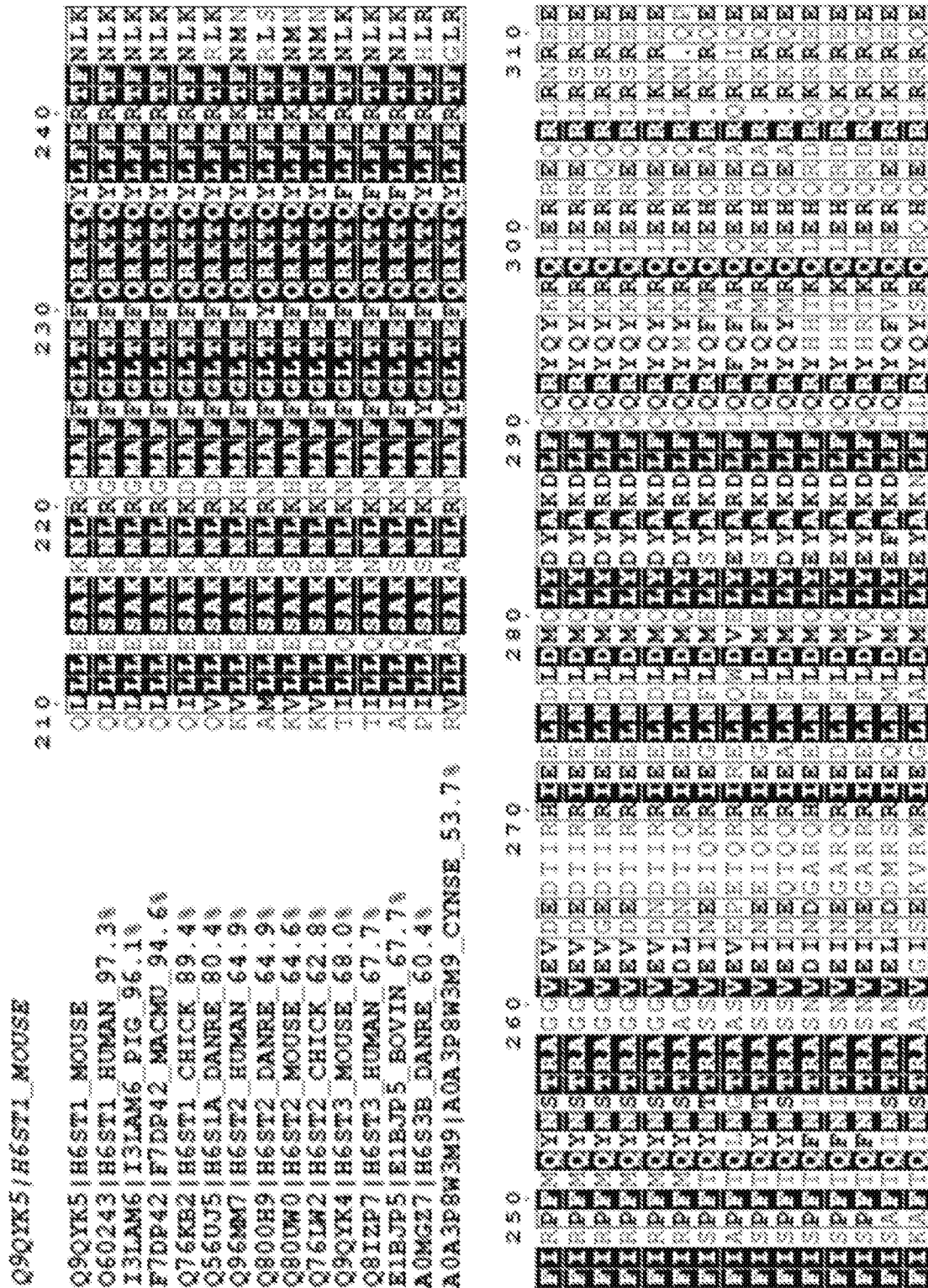
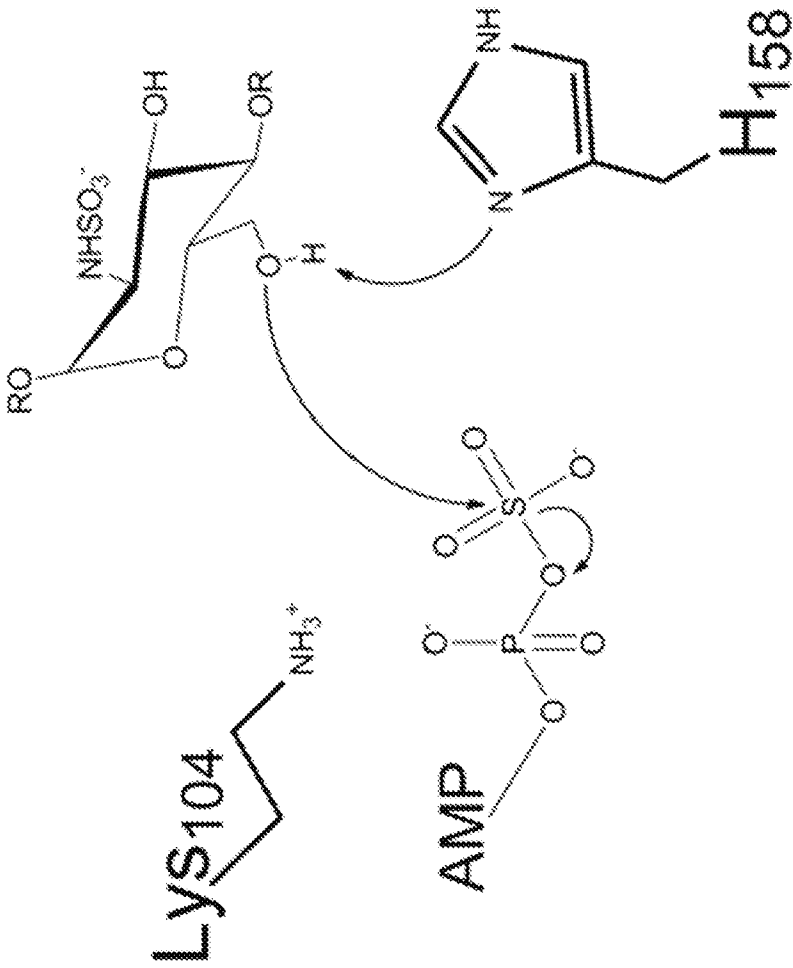
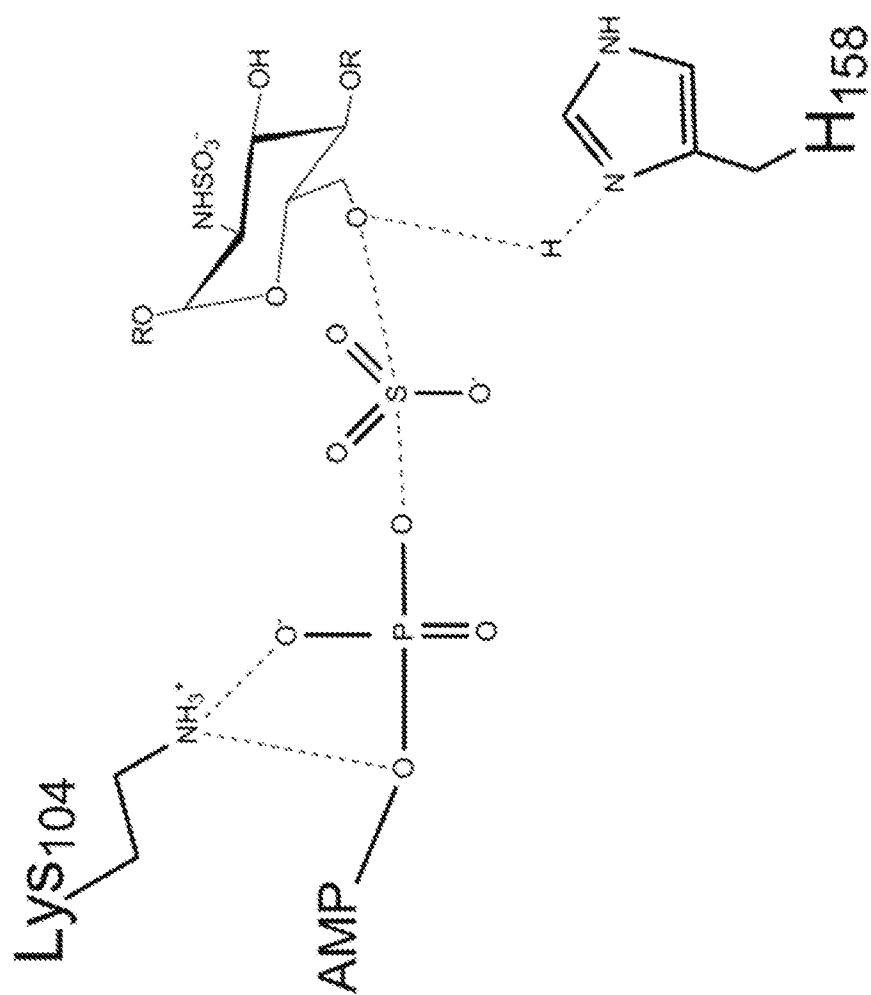


Figure 18C



Reaction

Figure 19A



Transition State

Figure 13

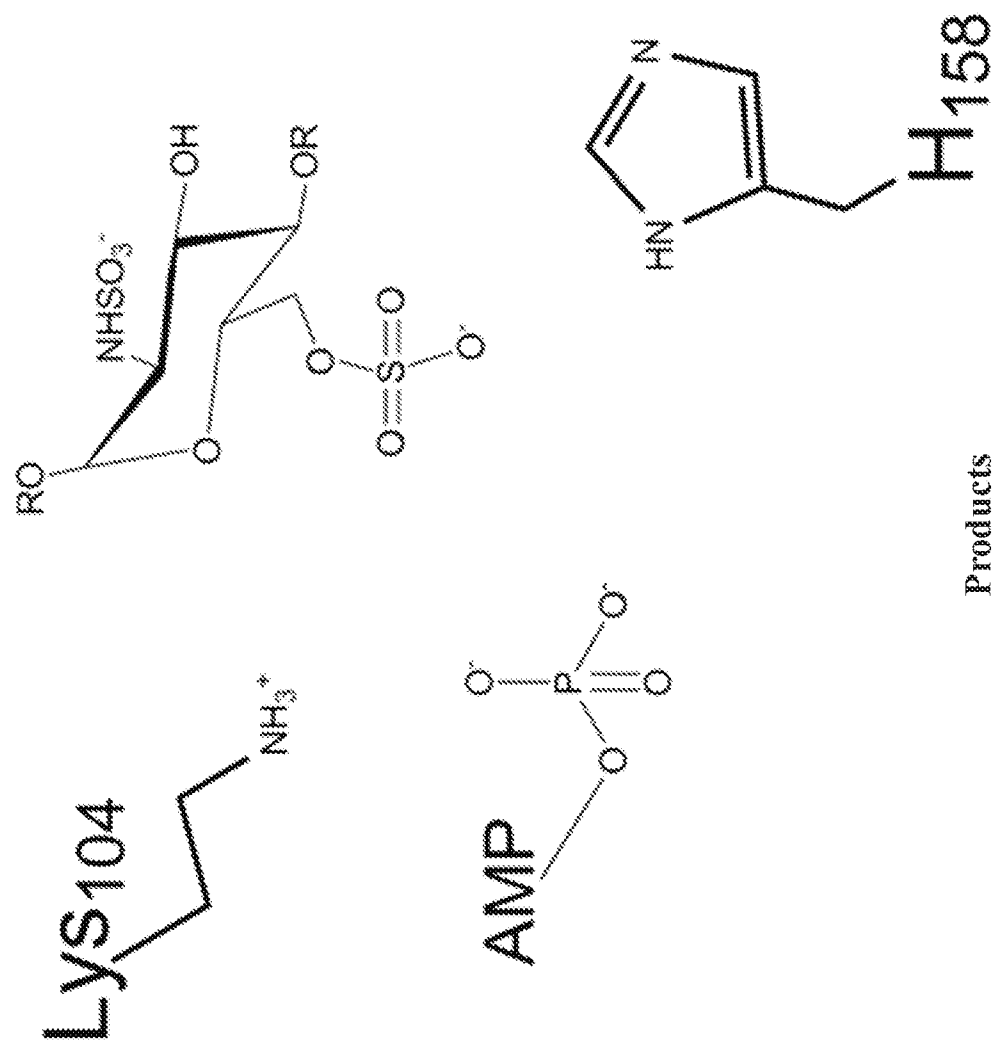


Figure 19C

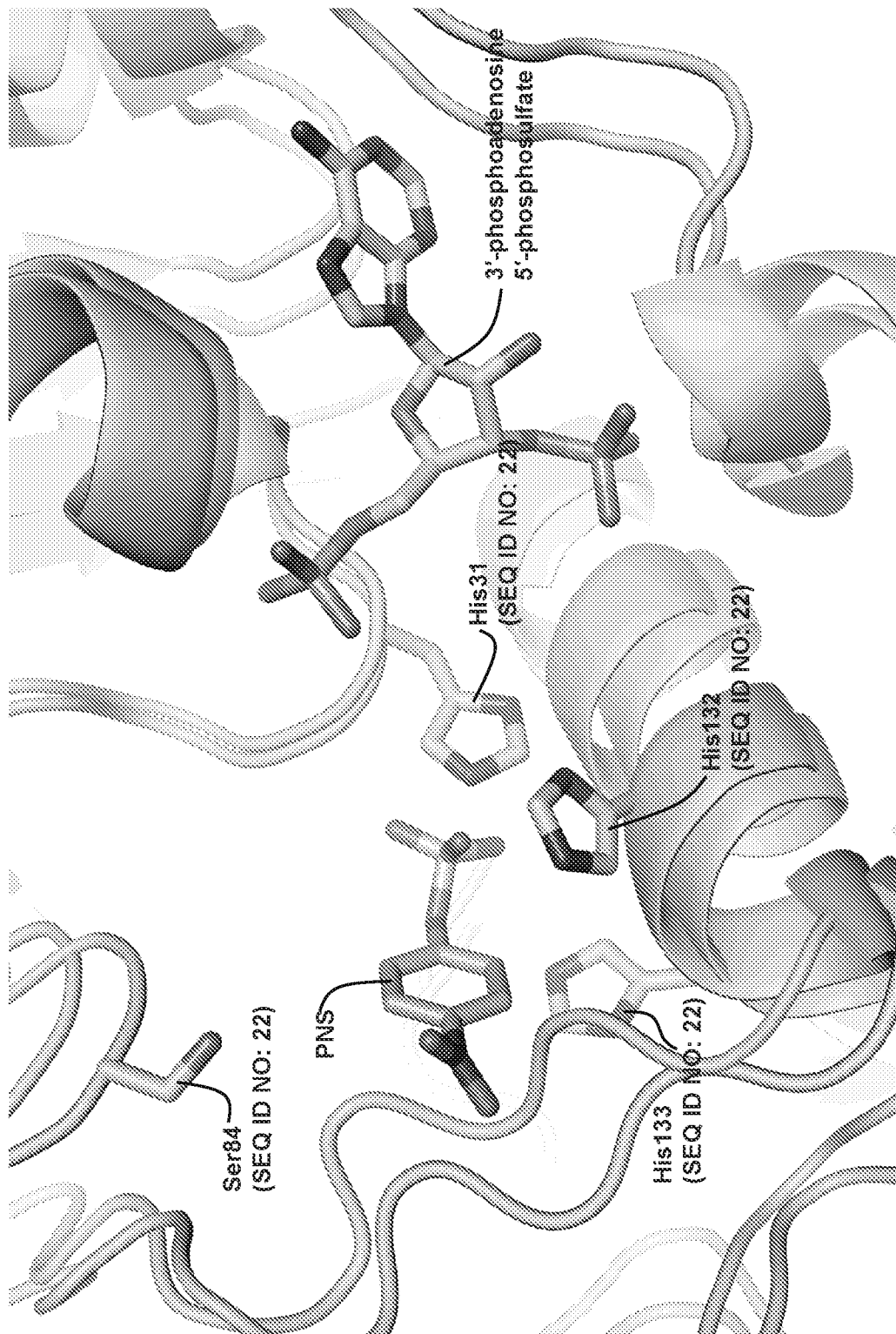


Figure 20

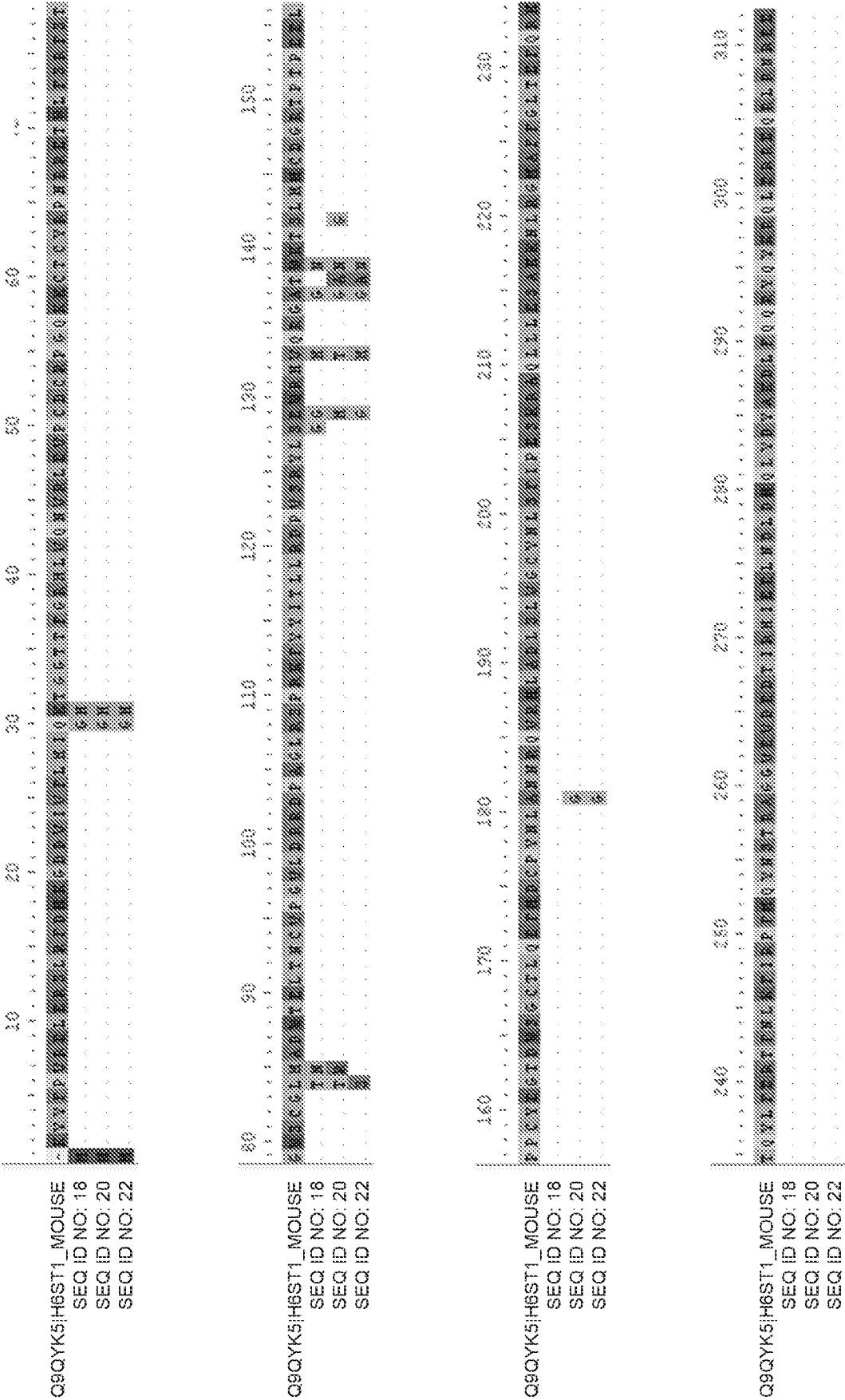


Figure 21

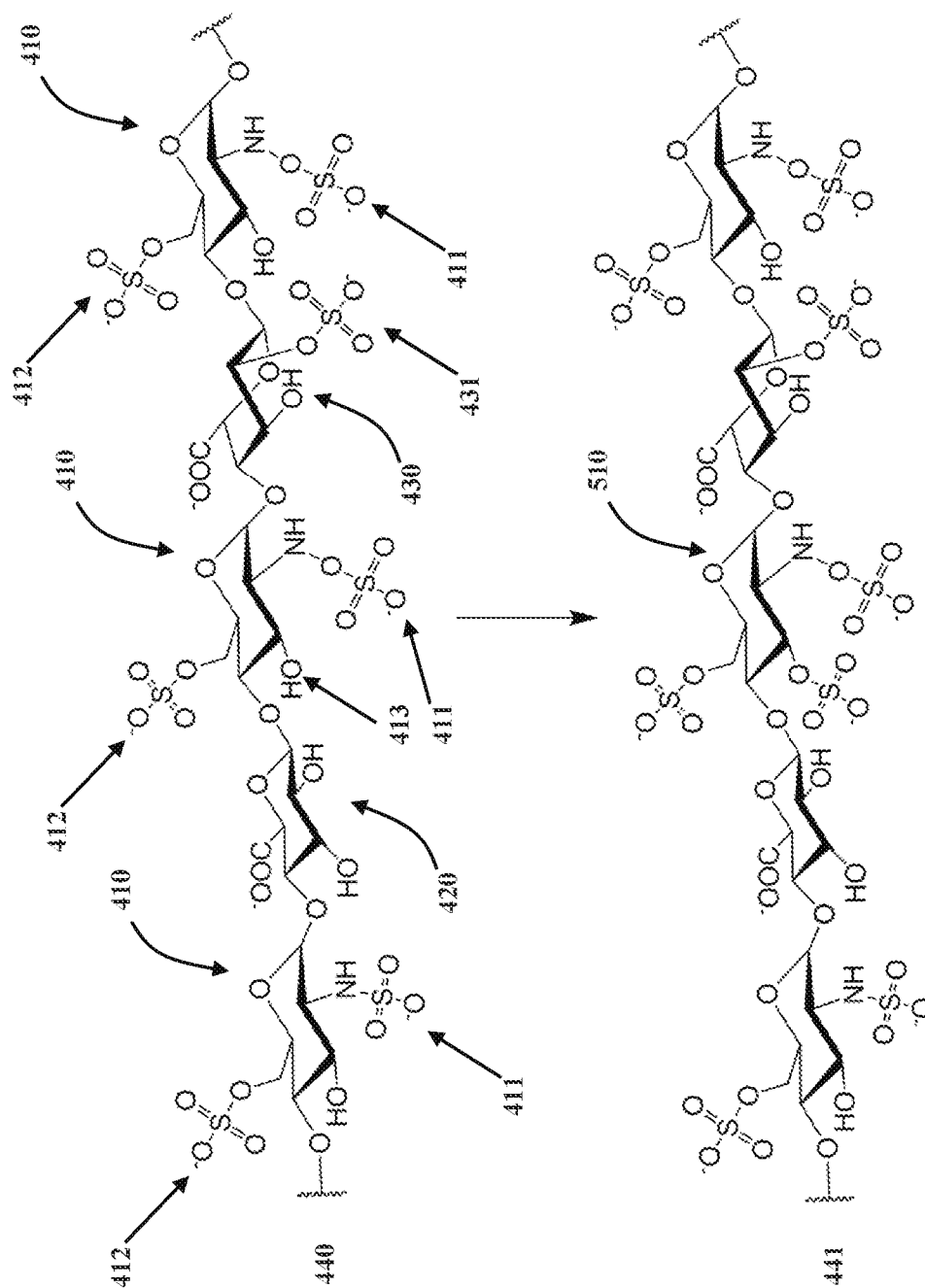
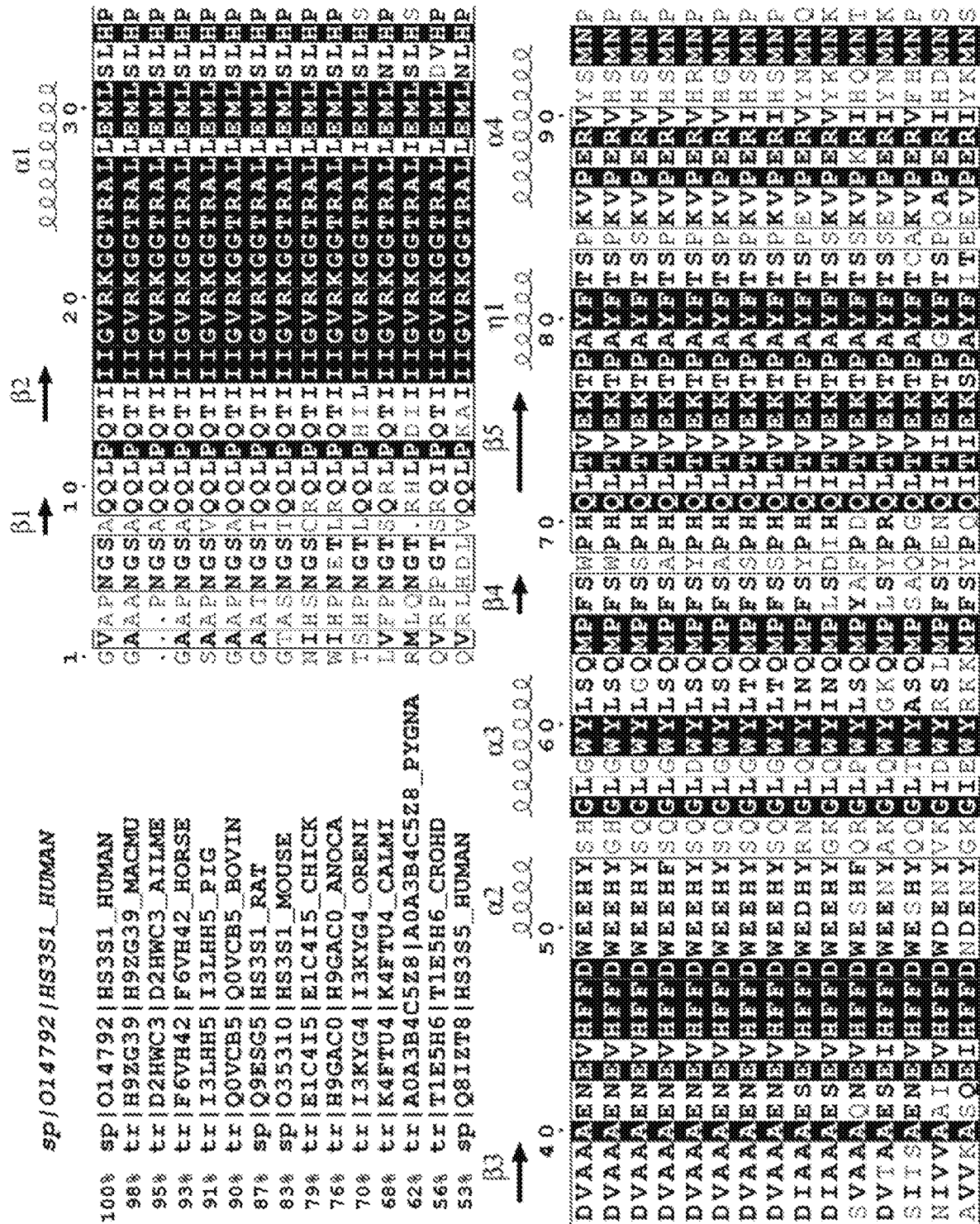
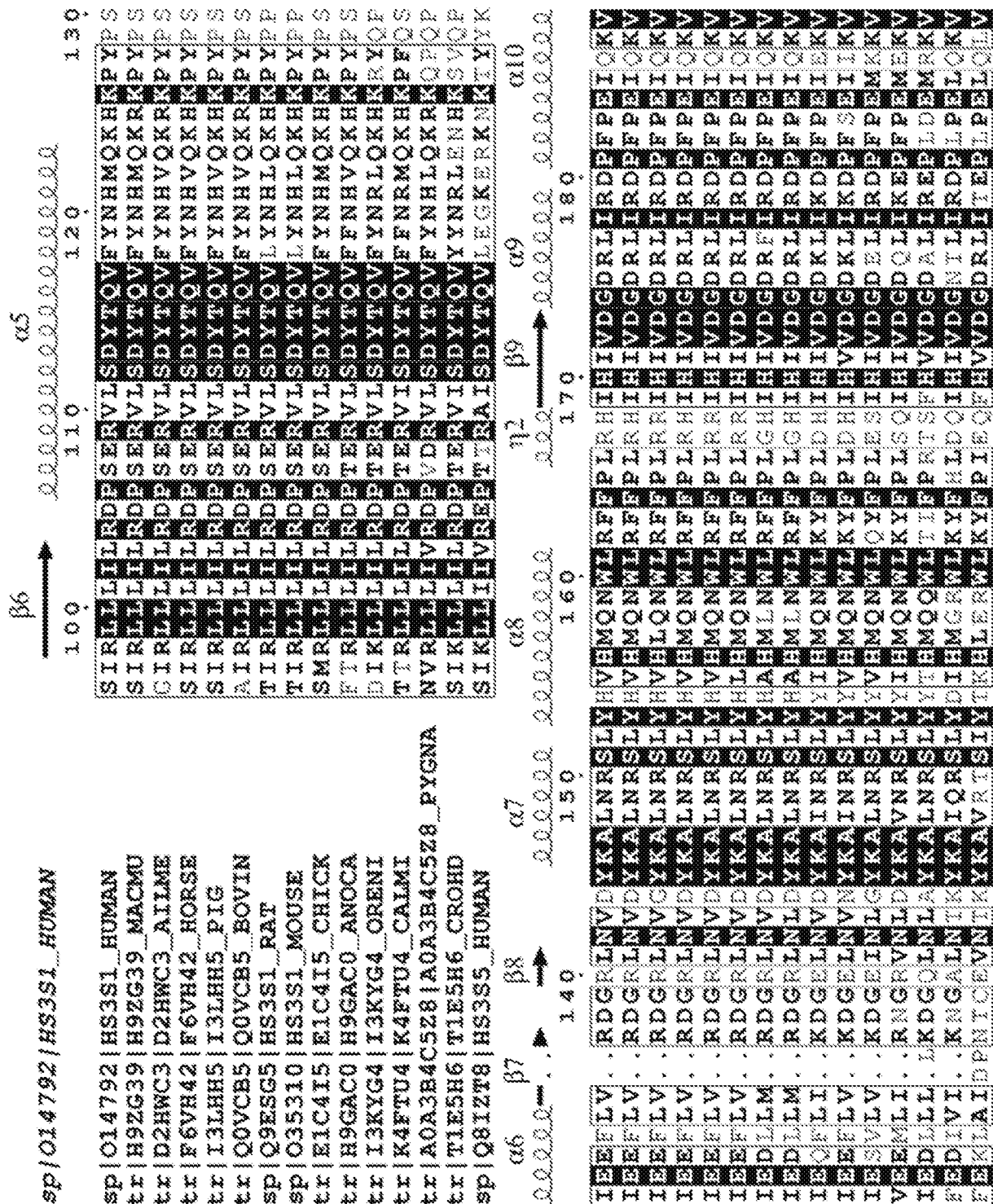
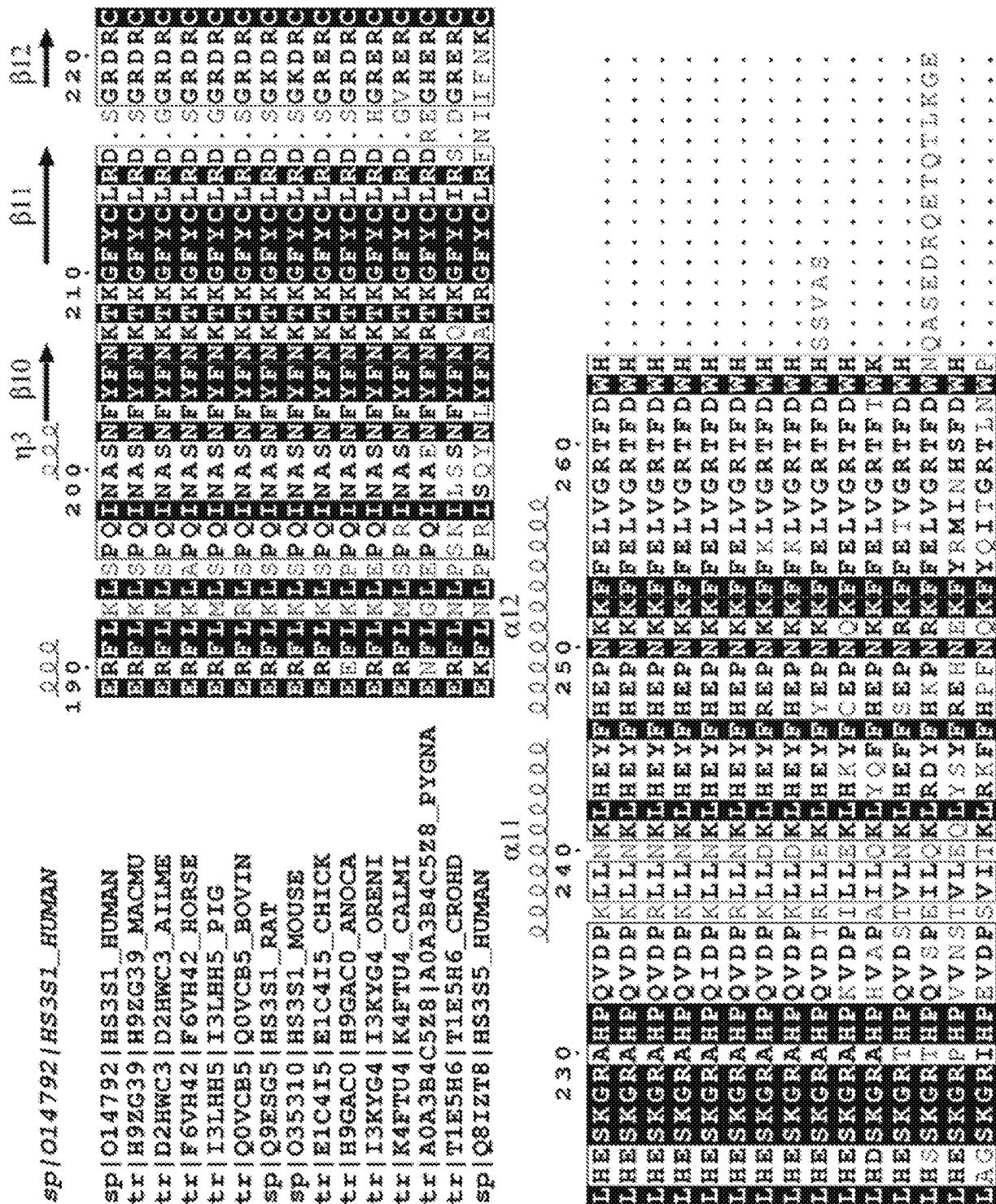


Figure 22







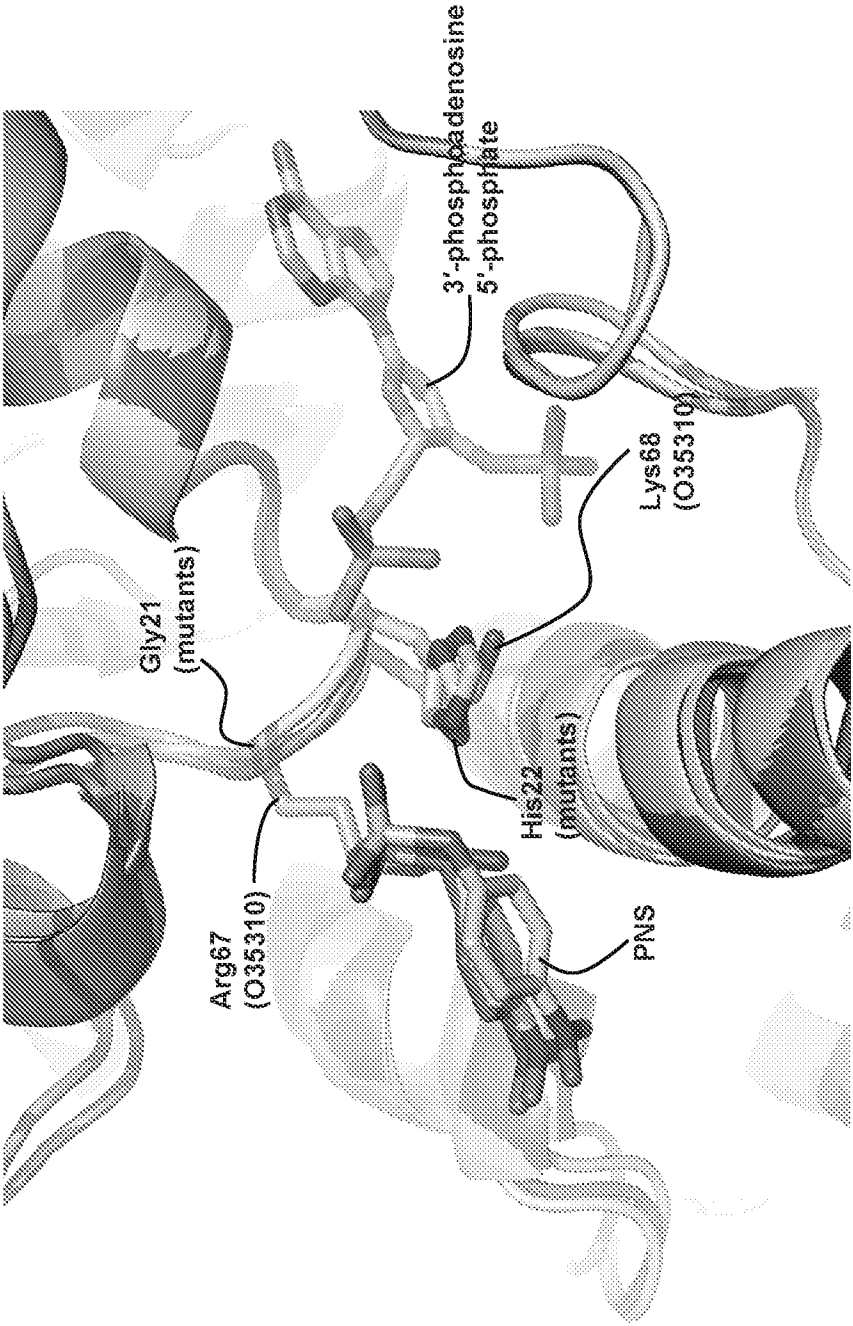


Figure 24

8

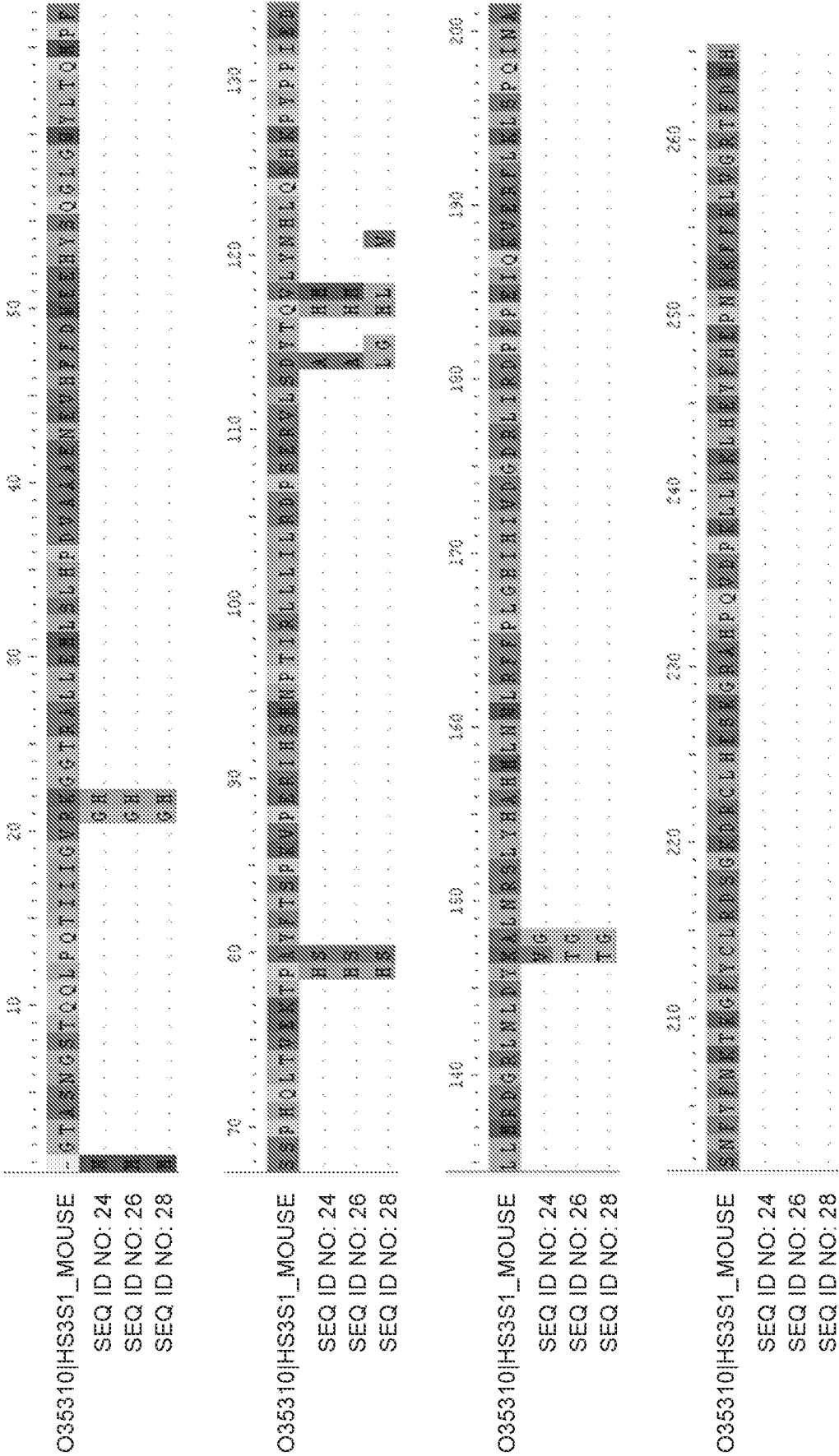


Figure 25

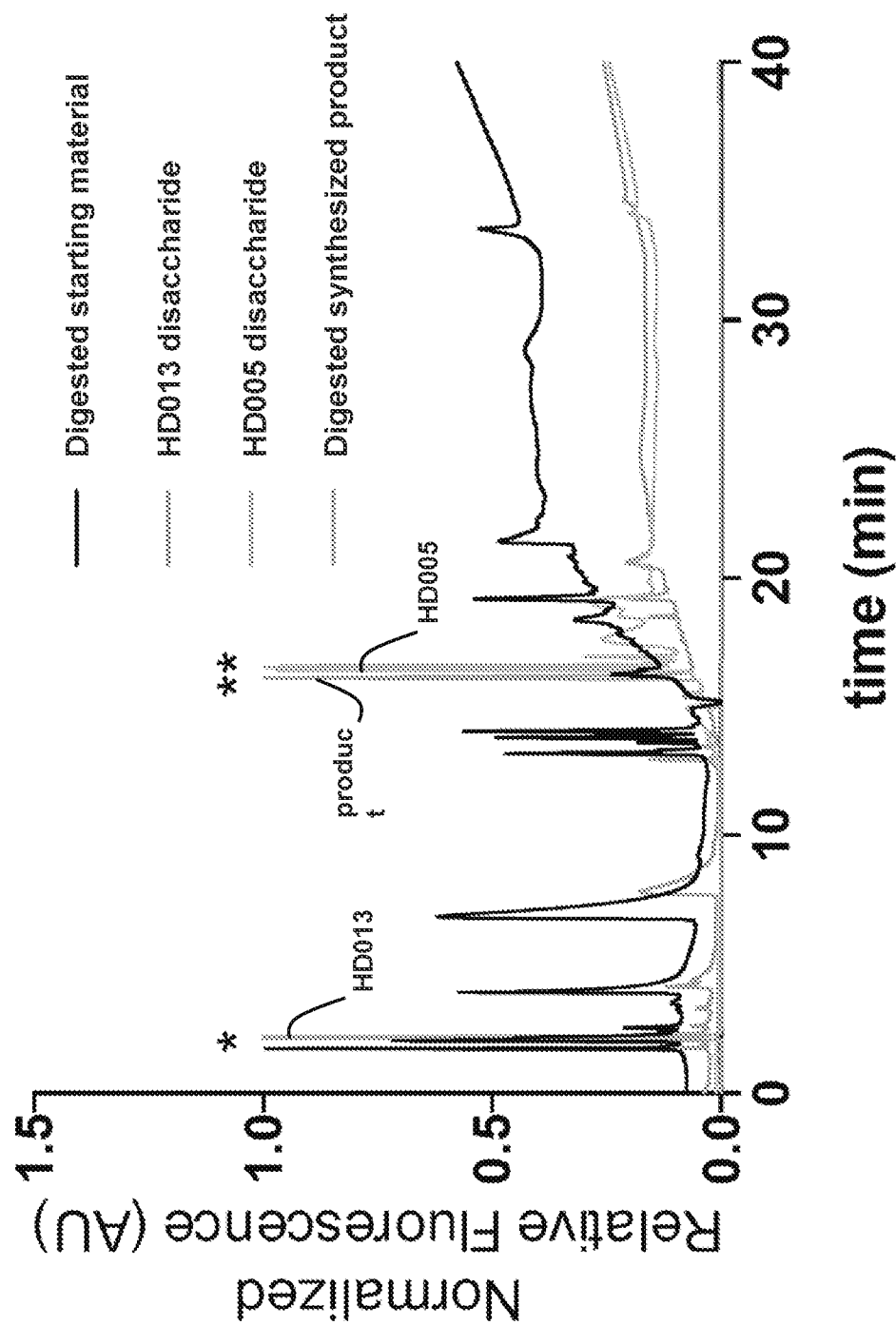


Figure 26

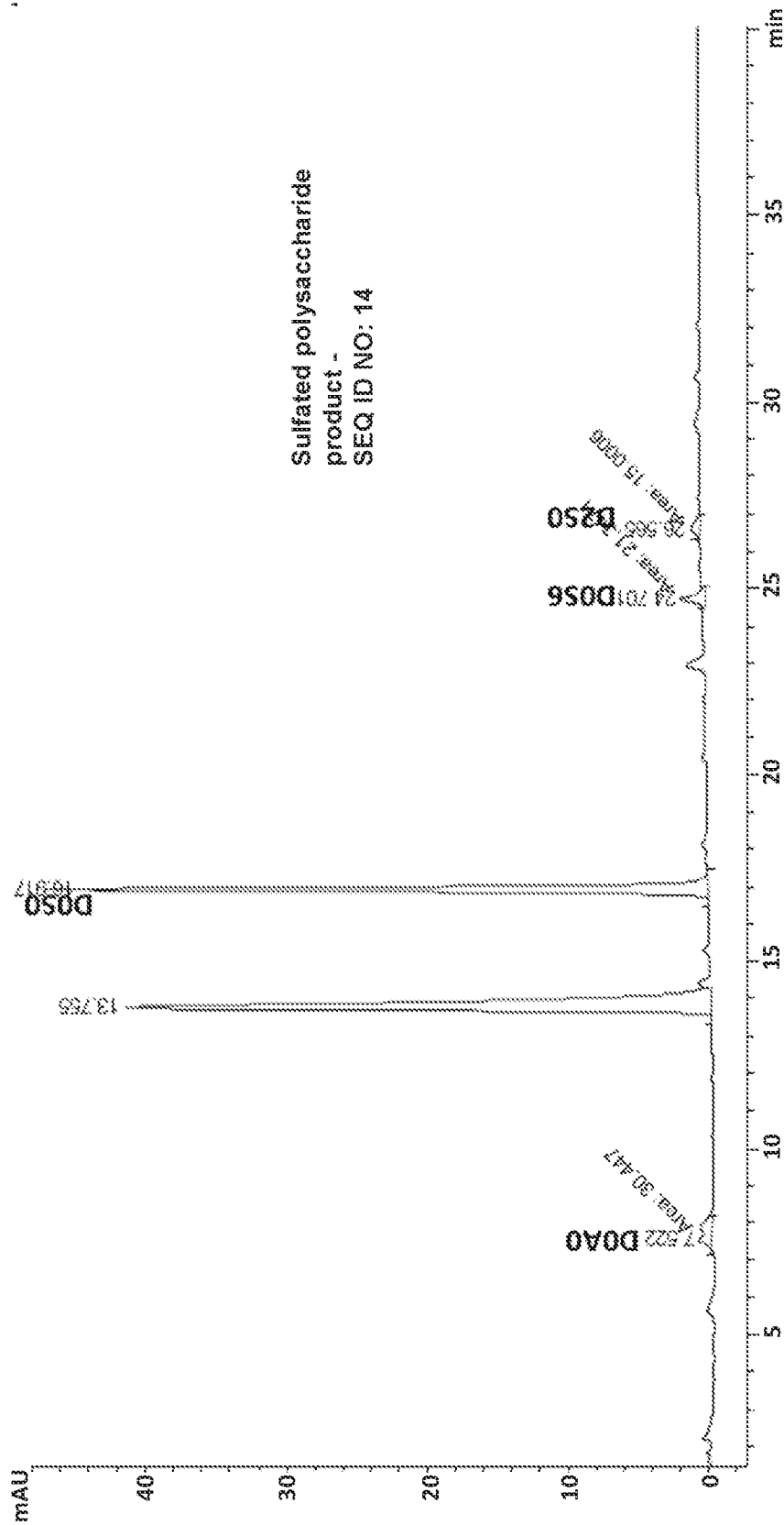


Figure 27A

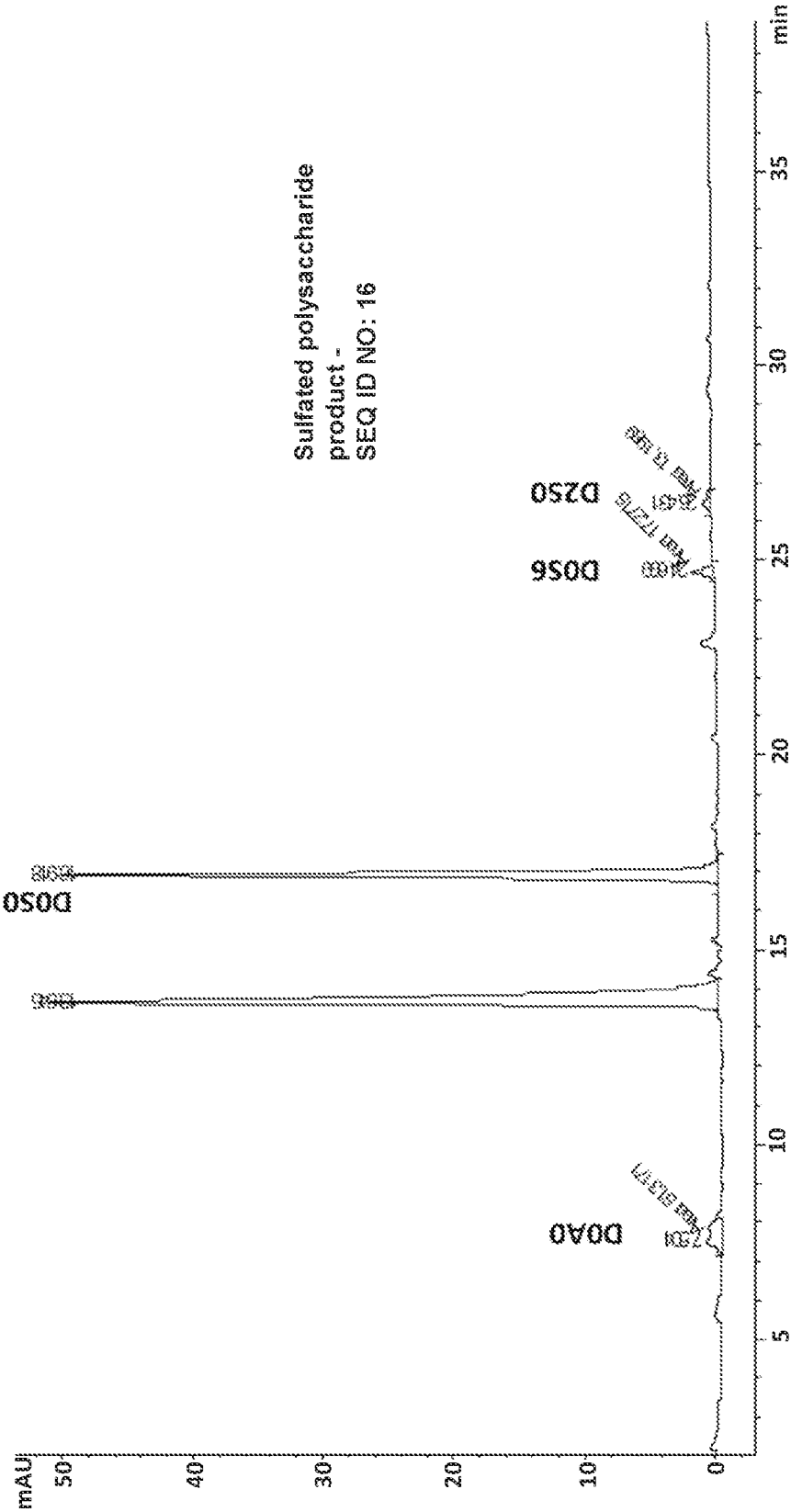


Figure 27B

Sulfated polysaccharide
product -
SEQ ID NO: 16

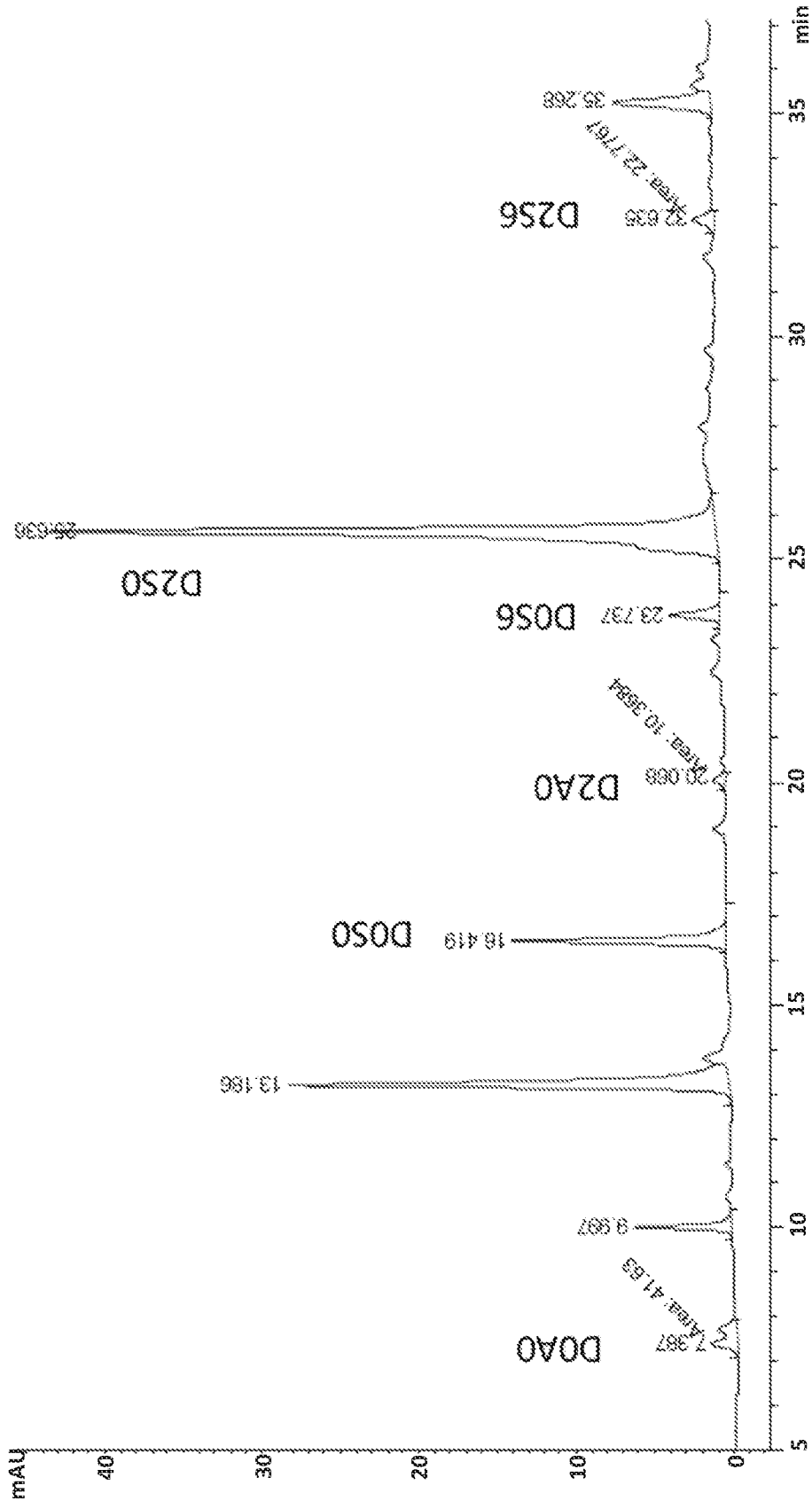


Figure 28A

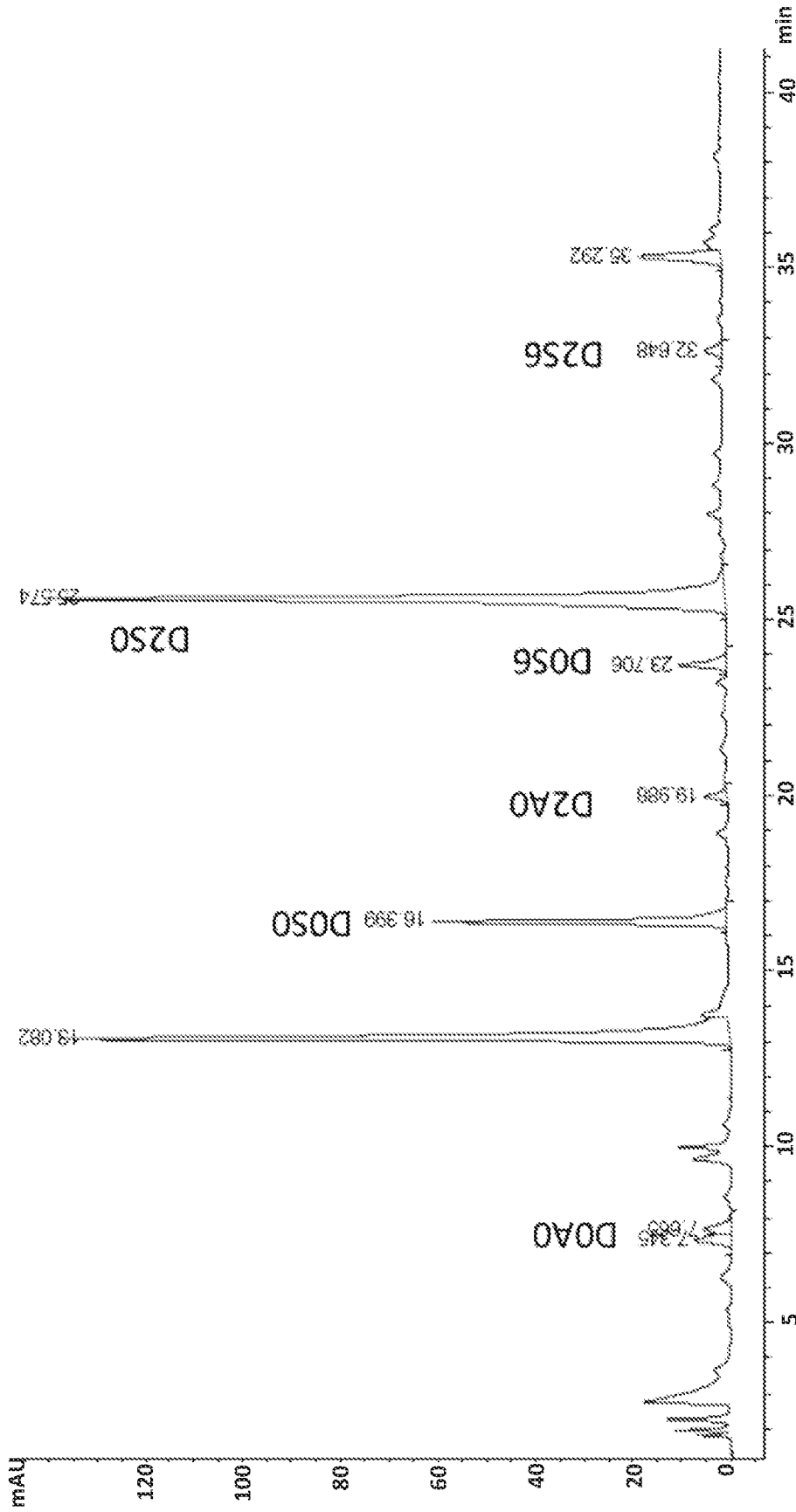


Figure 28B

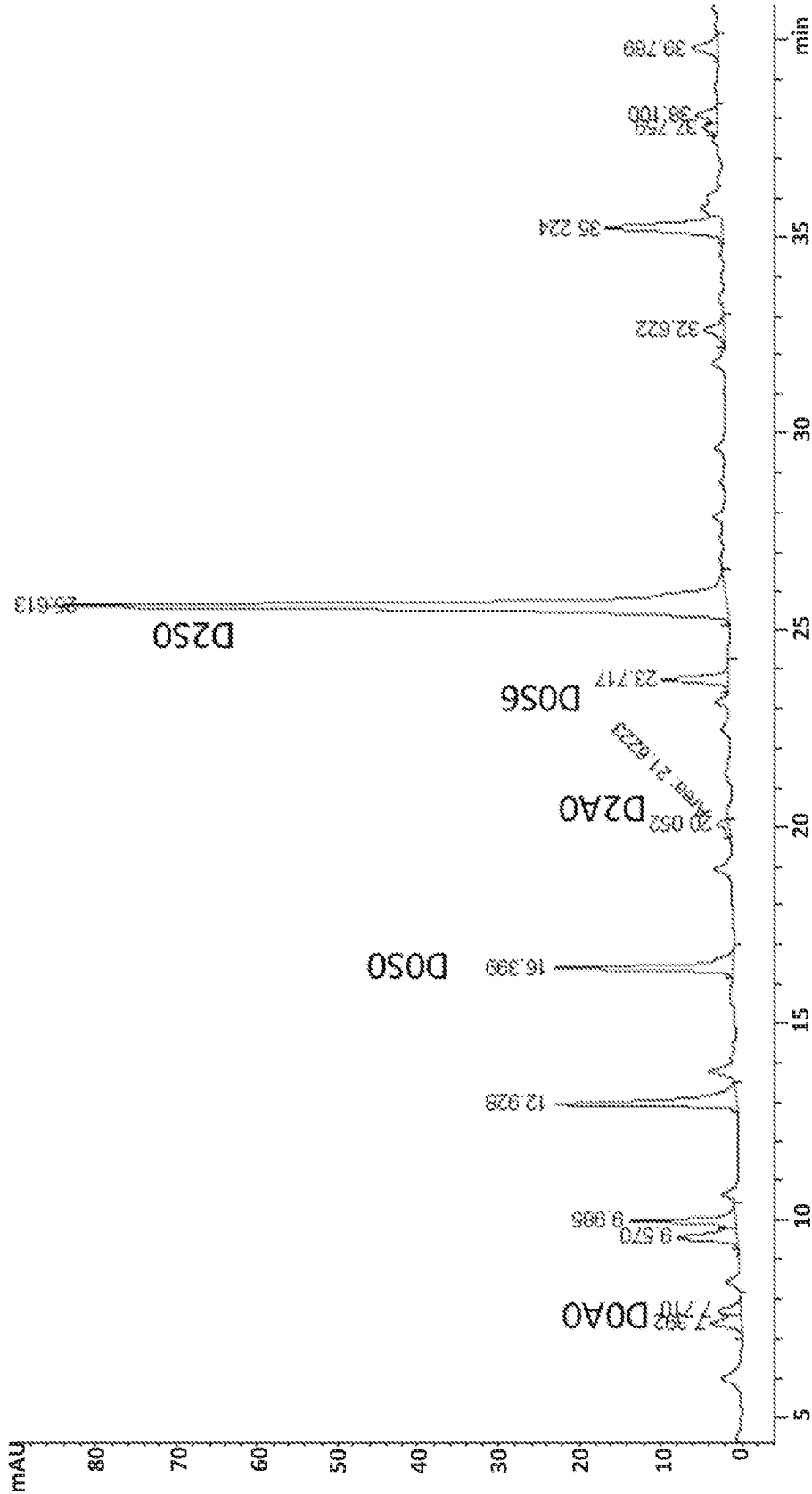


Figure 28C

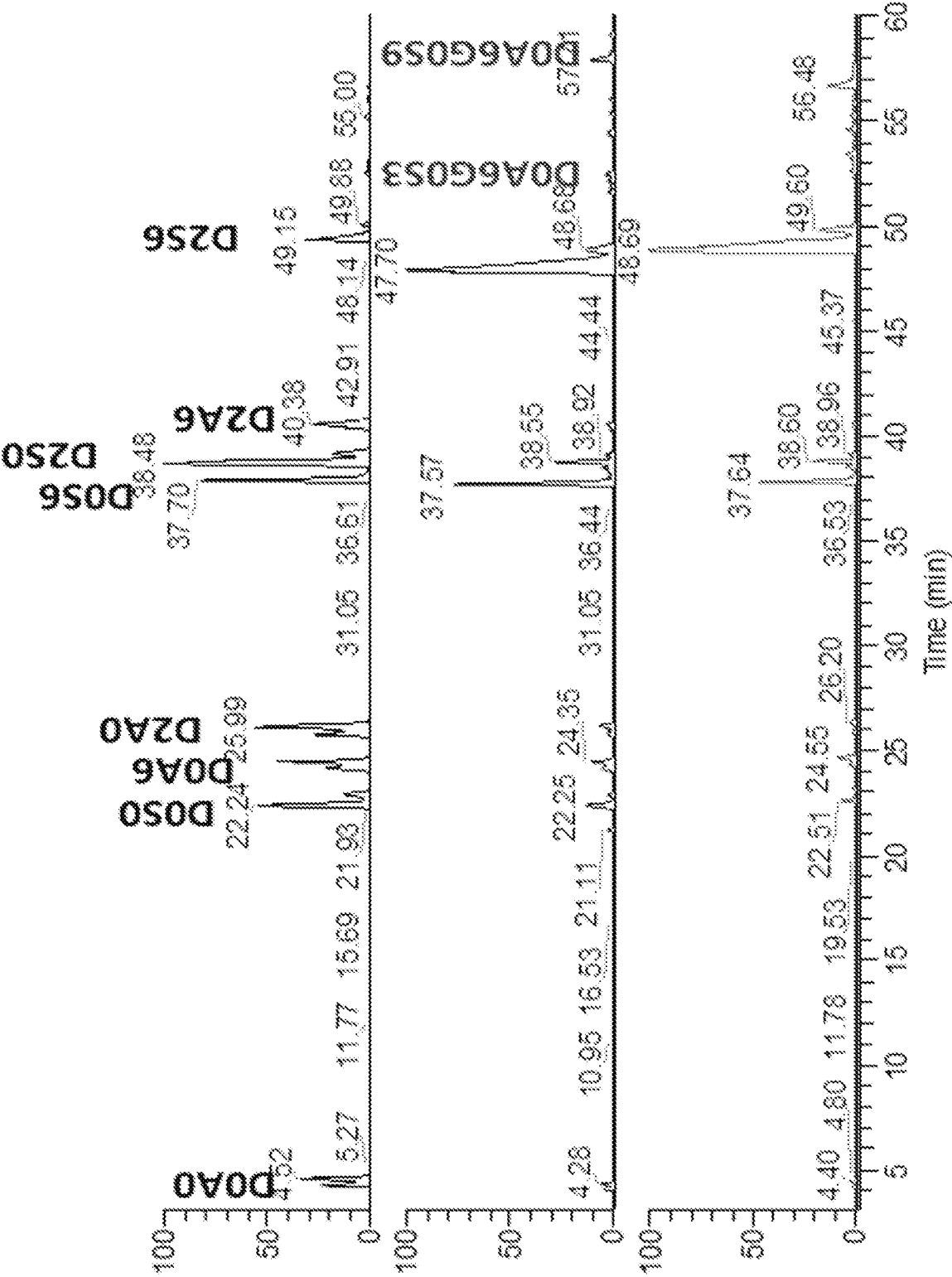


Figure 29A

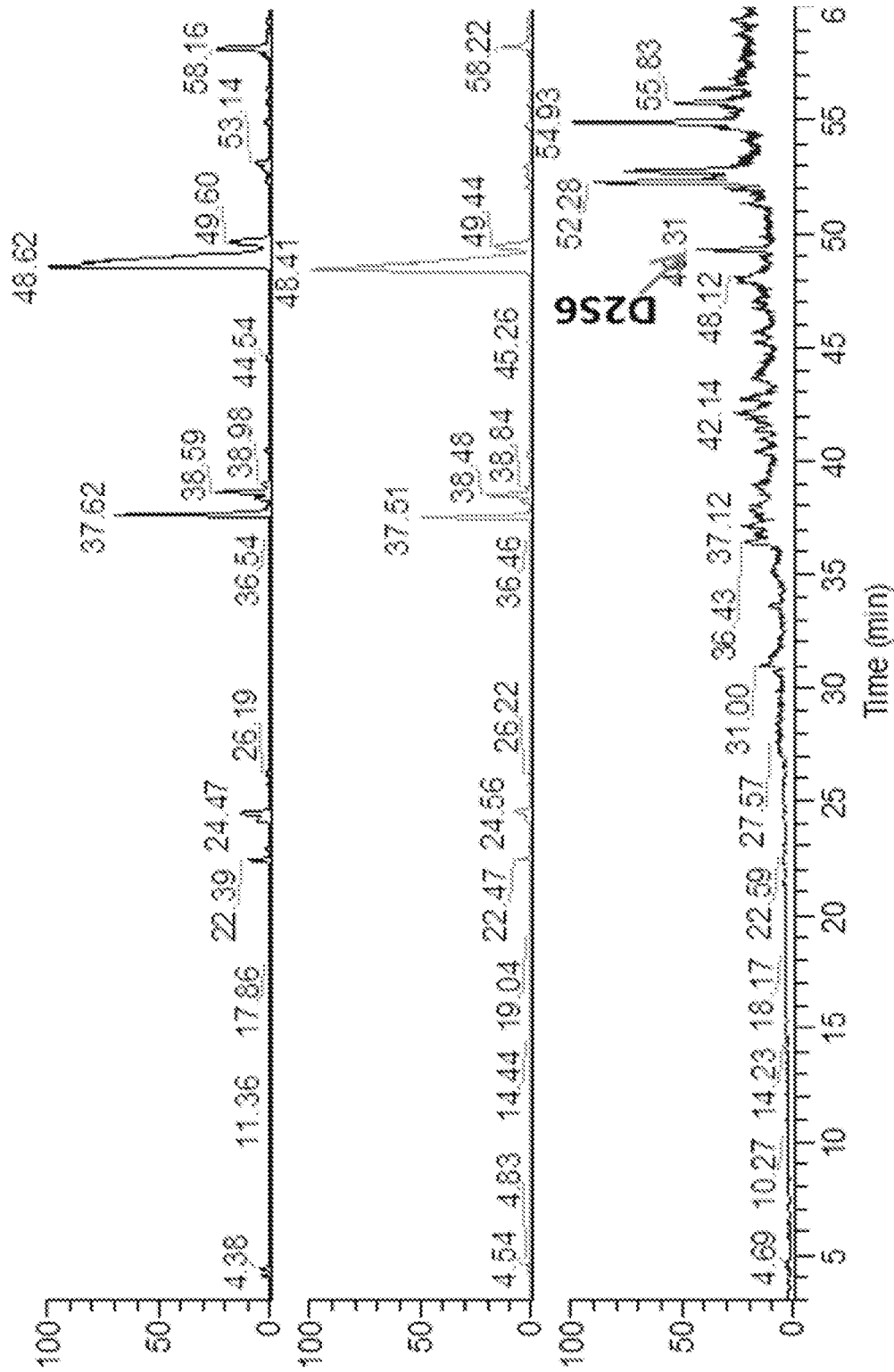


Figure 29B

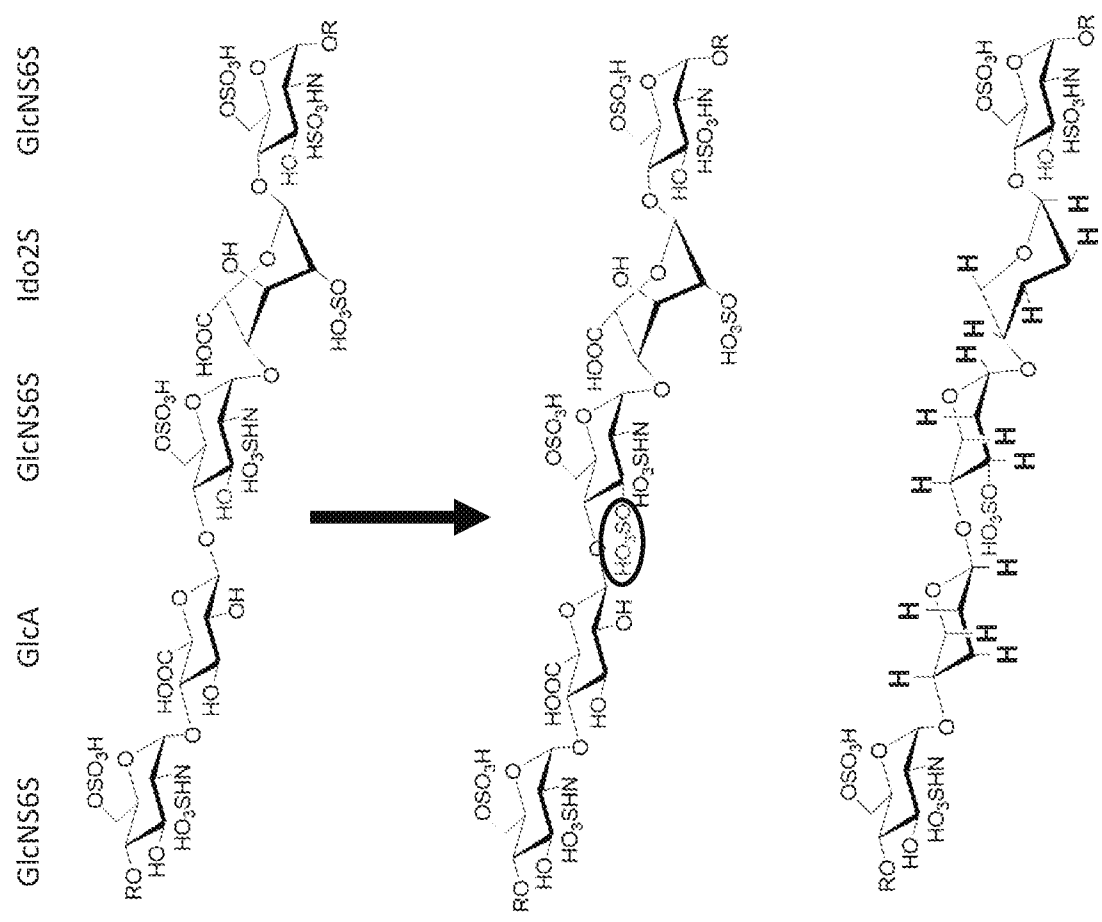


Figure 30

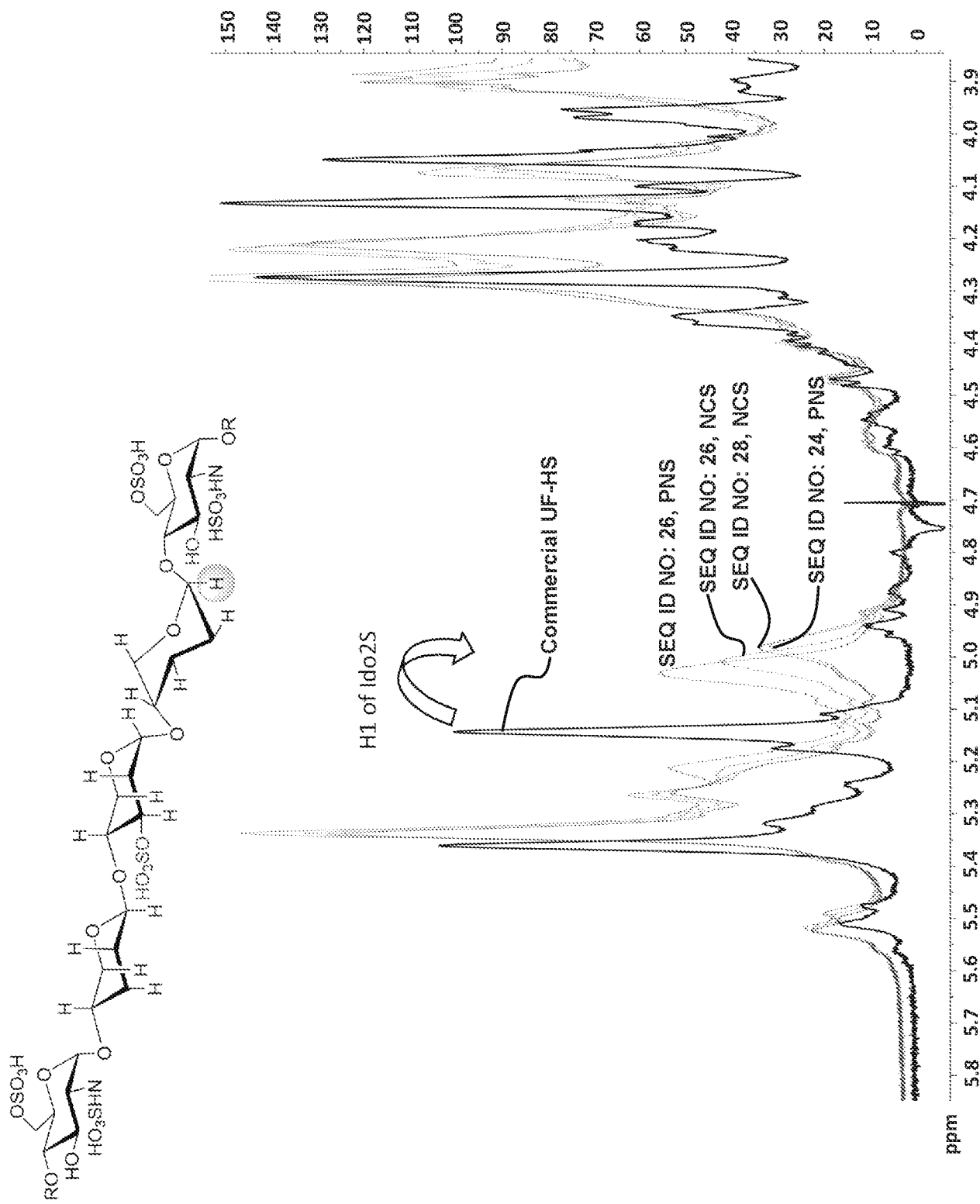


Figure 31

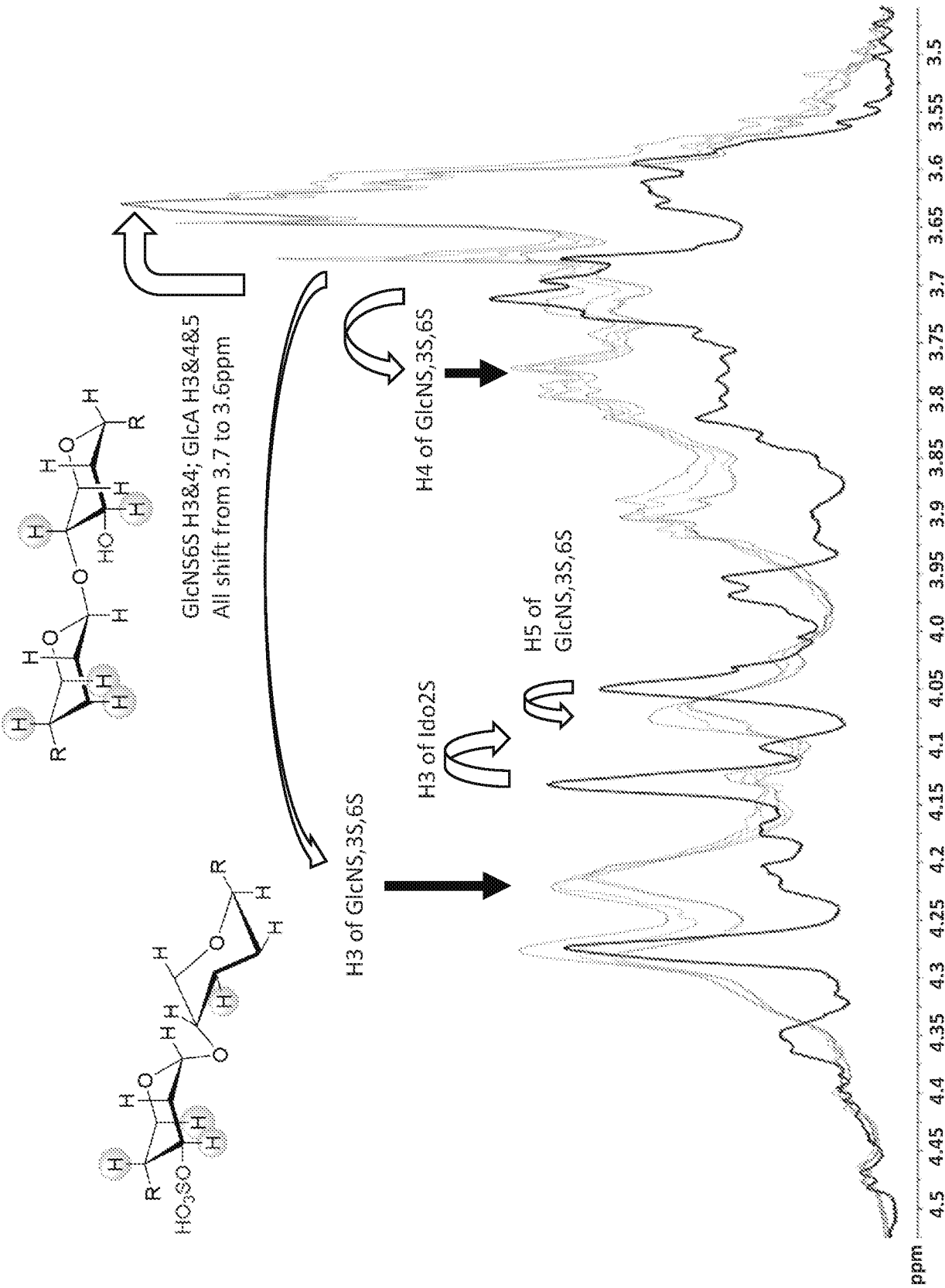


Figure 32

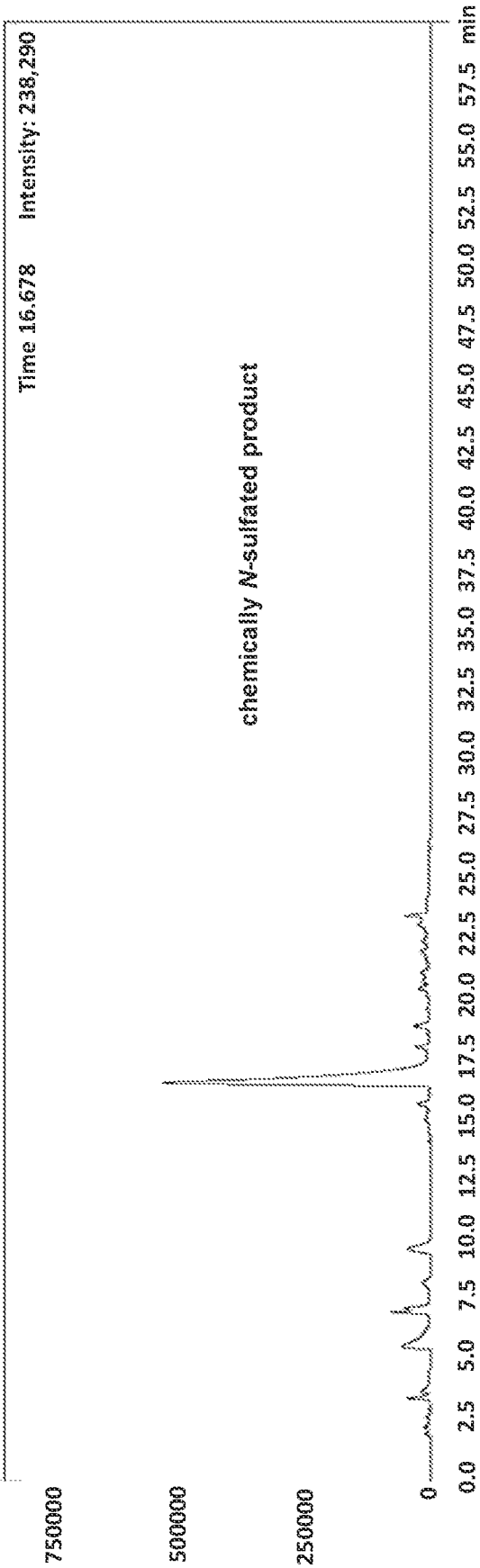
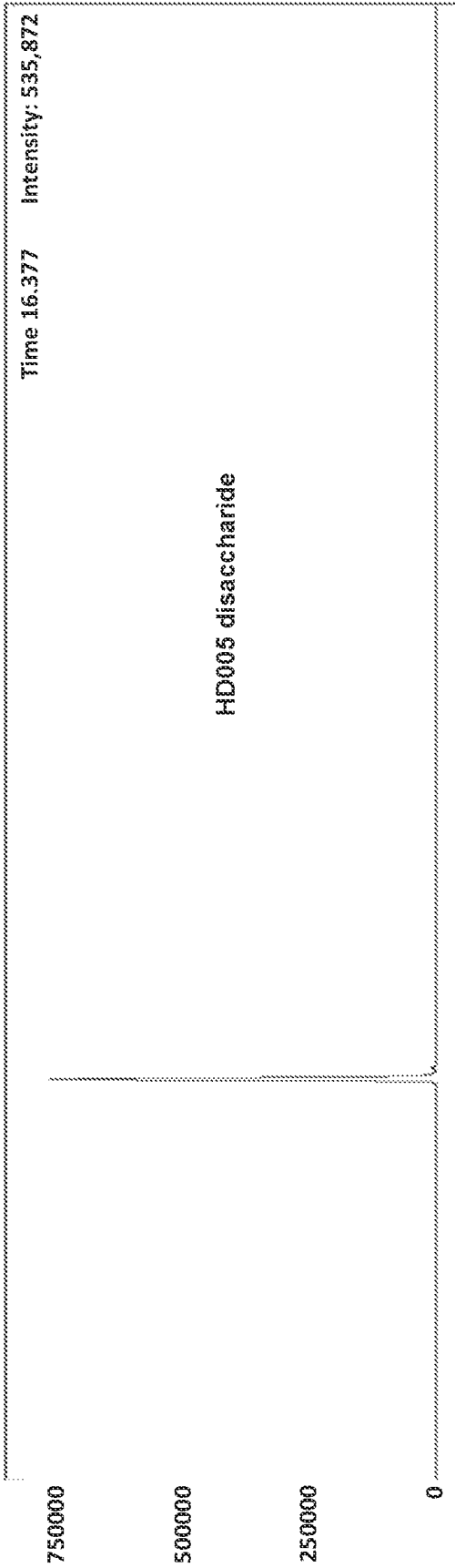


Figure 33

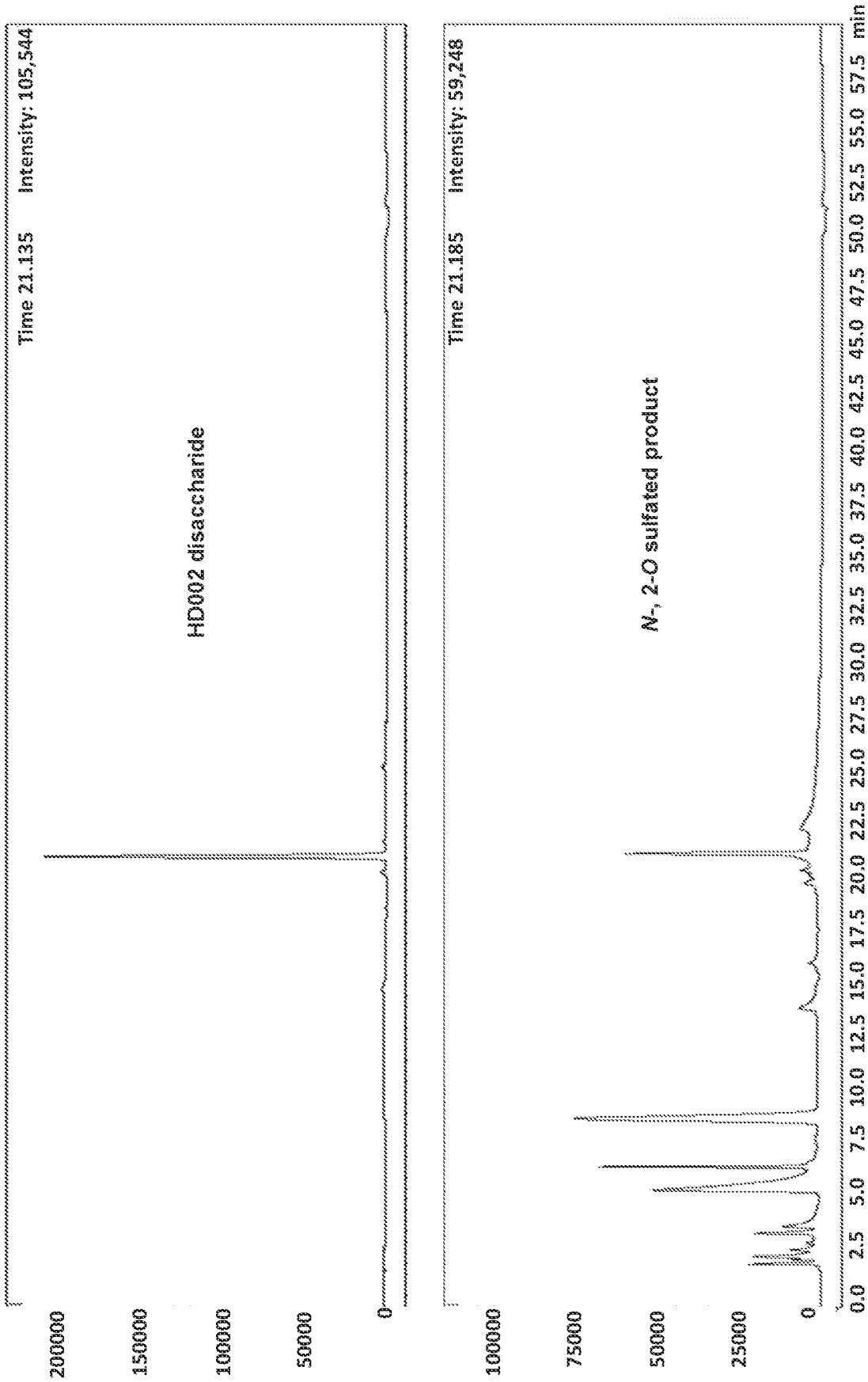


Figure 34

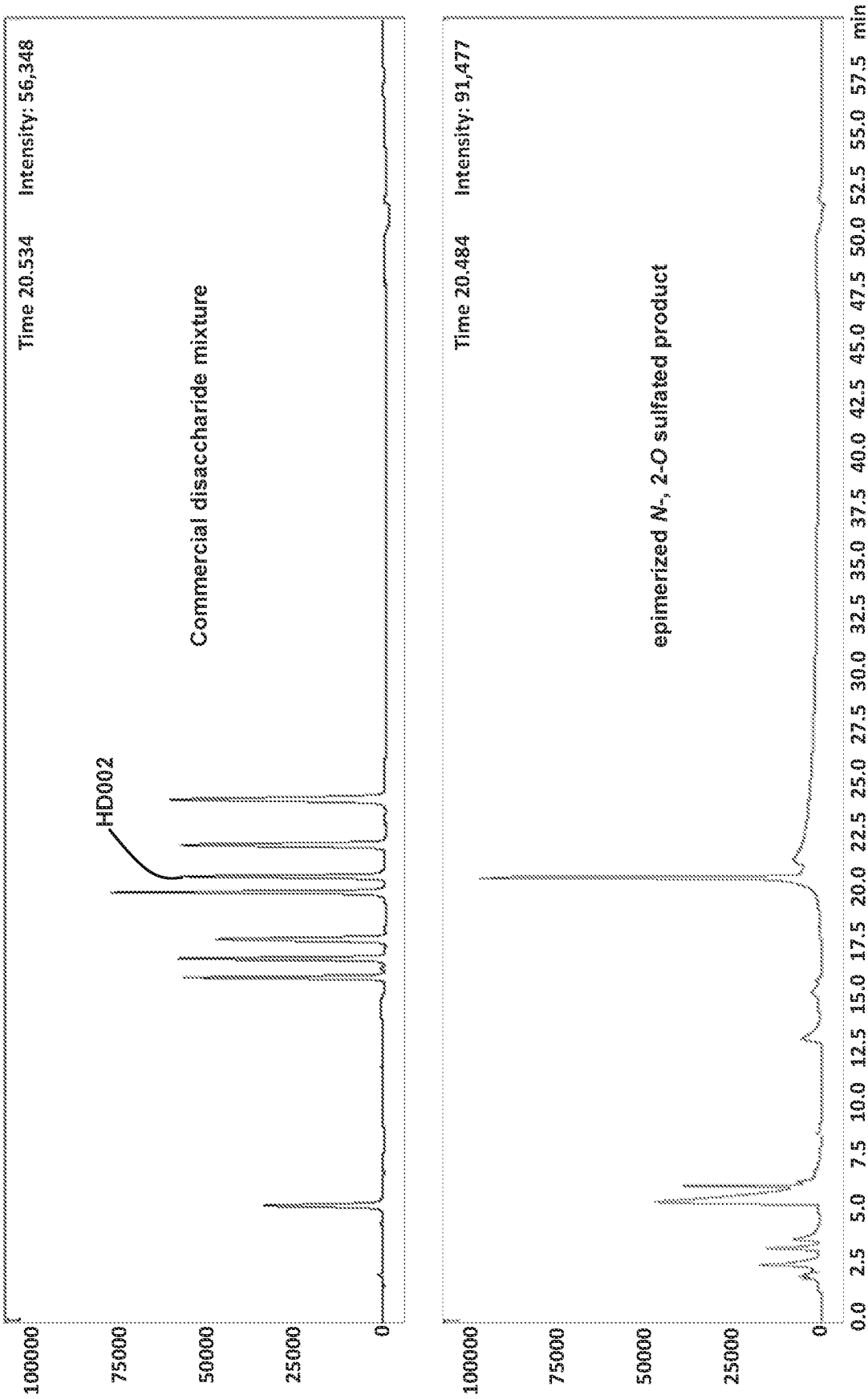


Figure 35

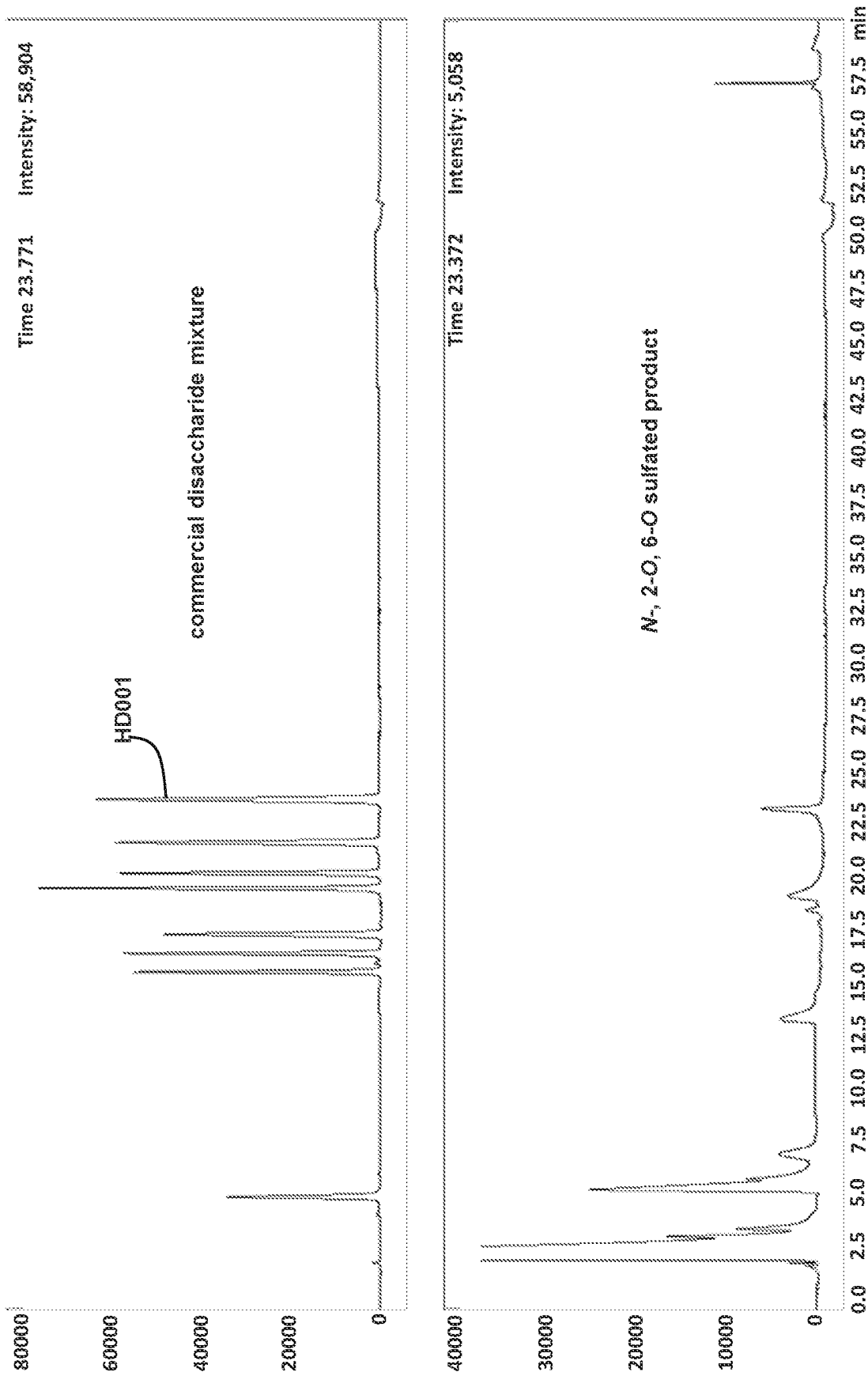


Figure 36

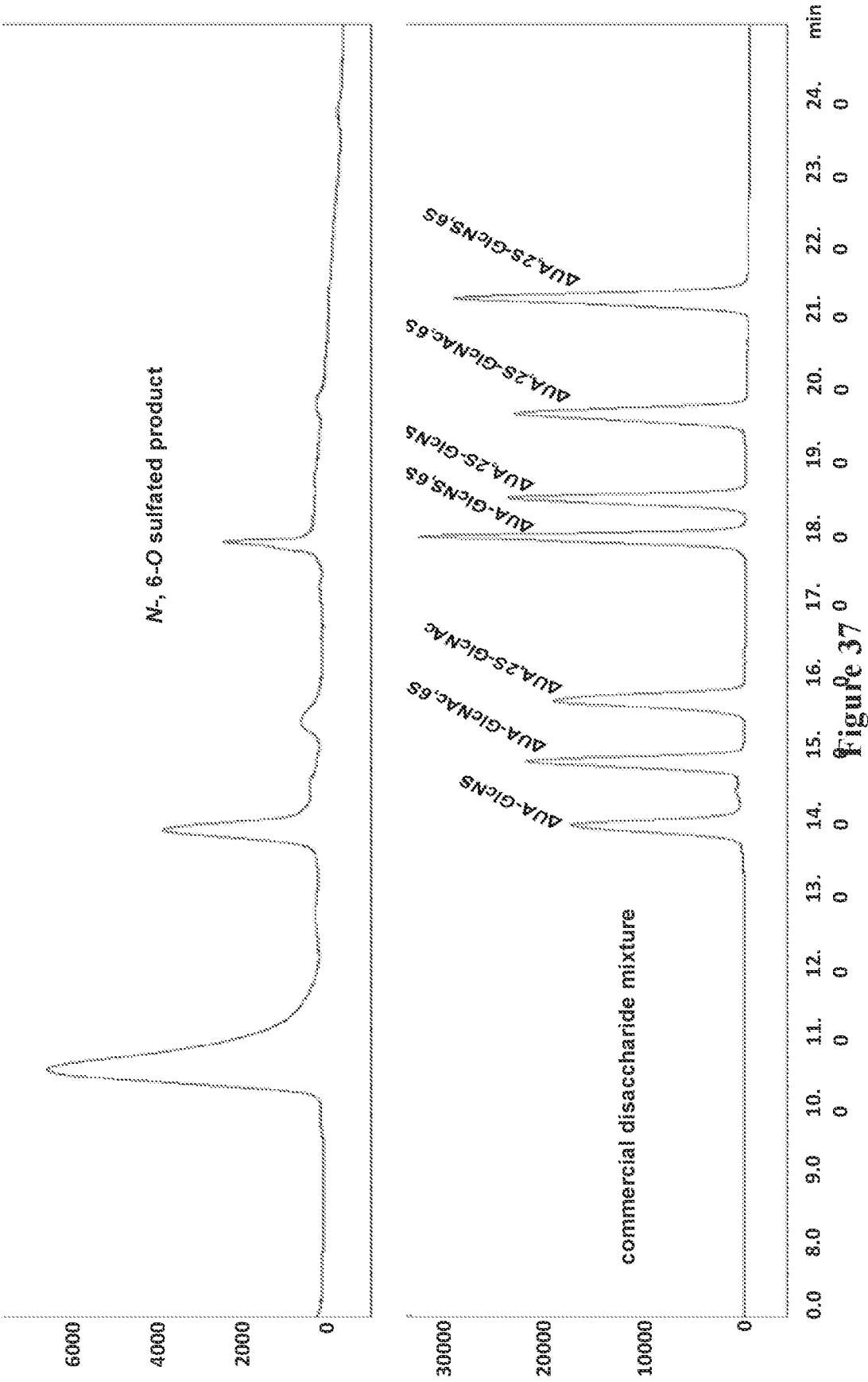


Figure 37

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 21/41537

Box No. 1 Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
- a. ☒ forming part of the international application as filed:
☒ in the form of an Annex C/ST.25 text file.
☐ on paper or in the form of an image file.
- b. ☐ furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
- c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
☐ in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).
☐ on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 21/41537

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 4-16, 21-29
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

---See Supplemental Box ---

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-3

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 21/41537

A. CLASSIFICATION OF SUBJECT MATTER

IPC - C12N 9/10 (2021.01)

CPC - C12N 9/13, C12Y 208/02, C12Y 208/02008, C12Y 208/02023

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

See Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WANG et al., Heparin's anti-inflammatory effects require glucosamine 6-O-sulfation and are mediated by blockade of L- and P-selectins. The Journal of Clinical Investigation. July 2002, Vol. 110, No. 1, pages 127-136; abstract; pg 135, col 1, para 1; pg 132, col 2, para 2; pg 133, col 1, para 1; Figure 1	1-3
A	US 2018/0298411 A1 (AJINOMOTO CO., INC.) 18 October 2018 (18.10.2018) entire document	1-3

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

20 October 2021

Date of mailing of the international search report

DEC 28 2021

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Kari Rodriguez

Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 21/41537

Box No. III Observations where unity of invention is lacking

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-3, directed to a pharmaceutical composition comprising N-sulfated, 6-O-sulfated heparan sulfate (NS6S/HS) polysaccharides that are structural analogs of O-desulfated heparin (ODSH).

Group II, claims 17-20, directed to a method of enzymatically synthesizing an NS6S/HS product comprising NS6S/HS polysaccharides that are structural analogs of O-desulfated heparin (ODSH).

The inventions listed as Groups I-II do not relate to a single special technical feature under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Special technical features:

Group I has the special technical feature of a pharmaceutical composition comprising N-sulfated, 6-O-sulfated heparan sulfate (NS6S/HS) polysaccharides, that is not required by Group II.

Group II has the special technical feature of a method of enzymatically synthesizing an NS6S/HS product comprising NS6S/HS polysaccharides, that is not required by Group I.

Common technical features:

Groups I-II share the common technical feature of:
a pharmaceutical composition comprising N-sulfated, 6-O-sulfated heparan sulfate (NS6S/HS) polysaccharides that are structural analogs of O-desulfated heparin (ODSH), wherein each of the hexuronic acid and glucosamine residues within the NS6S/HS polysaccharides are fully unsulfated at their 2-O and 3-O positions, respectively.

However, this shared technical feature does not represent a contribution over prior art, because this shared technical feature is anticipated by the publication entitled "Heparin's anti-inflammatory effects require glucosamine 6-O-sulfation and are mediated by blockade of L- and P-selectins" by Wang et al. (hereinafter "Wang").

Wang discloses a pharmaceutical composition (abstract - "Such analogs may prove useful as therapeutically effective inhibitors of inflammation"; pg 135, col 1, para 1 - "As a drug, heparin has wide array of potential pharmacological uses") comprising N-sulfated, 6-O-sulfated heparan sulfate (NS6S/HS) polysaccharides (Figure 1 illustrates 6-O-sulfated heparan sulfate, 2/3DS-heparin) that are structural analogs of O-desulfated heparin (ODSH) (Figure 1 illustrates 6-O-sulfated heparan sulfate, 2/3DS-heparin; abstract - "Compared with unfractionated heparin, the modified heparinoids had inhibitory activity in this general order: over-O-sulfated heparin > heparin > 2-O,3-O-desulfated . N-desulfated/N-acetylated heparin carboxyl-reduced heparin . N-,2-O,3-O-desulfated heparin >> 6-O-desulfated heparin...Such analogs may prove useful as therapeutically effective inhibitors of inflammation"), wherein each of the hexuronic acid and glucosamine residues within the NS6S/HS polysaccharides are fully unsulfated at their 2-O and 3-O positions, respectively (abstract - "Compared with unfractionated heparin, the modified heparinoids had inhibitory activity in this general order: over-O-sulfated heparin > heparin > 2-O,3-O-desulfated . N-desulfated/N-acetylated heparin carboxyl-reduced heparin . N-,2-O,3-O-desulfated heparin >> 6-O-desulfated heparin").

As the technical features were known in the art at the time of the invention, they cannot be considered special technical features that would otherwise unify the groups.

Therefore, Group I-II inventions lack unity under PCT Rule 13 because they do not share the same or corresponding special technical feature.

*Continuation of number 4 above: claims 4-16, 21-29 are held unsearchable because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Note, Claims 21-22 which depend from "the method according to any of Claims 1-5", is objected to, because Claims 1-5 do not disclose the antecedent "N-sulfation agent" specified in Claims 21-22. Claim 21-22 are reconstructed as though depending from "the method according to any of Claims 17-20".