Published in final edited form as:

ACS Chem Biol. 2017 January 20; 12(1): 73–82. doi:10.1021/acscchembio.6b00841.

Structure Based Substrate Specificity Analysis of Heparan Sulfate 6-O-Sulfotransferases

Yongmei Xu¹, Andrea F Moon², Shuqin Xu^{1,3}, Juno M. Krahn², Jian Liu^{1,*}, and Lars C. Pedersen²

¹Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599, USA

²Genome Integrity and Structural Biology Laboratory, NIEHS, National Institutes of Health, Research Triangle Park, NC 27709, USA

³School of Pharmaceutical Science, Jiangnan University, Wuxi 214122, China

Abstract

Heparan sulfate (HS) is a sulfated polysaccharide exhibiting essential physiological functions. HS 6-O-sulfotransferase (6-OST) transfers a sulfo group to the 6-OH position of glucosamine units to confer a variety of HS biological activities. There are three different isoforms of 6-OST in the human genome. Here, we report crystal structures of the ternary complex of 6-OST with the sulfo donor analog 3'-phosphoadenosine 5'-phosphate and three different oligosaccharide substrates at 1.95 to 2.1 Å resolutions. Structural and mutational analyses reveal amino acid residues that contribute to catalysis and substrate recognition of 6-OST. Unexpectedly, the structures reveal 6-OST engages HS in a completely different orientation than other HS sulfotransferases, and sheds light on the basic HS requirements for specificity. These findings also contribute structural information to understand mutations in human 6-OST isoform 1 associated with the human genetic disease idiopathic hypogonadotrophic hypogonadism characterized by delayed or lack of puberty.

Keywords

oligosaccharides; heparin; sulfated carbohydrates; crystal structure

Heparan sulfates (HS) are sulfated polysaccharides that are found on the surface of mammalian cells attached to core syndecan and glypican proteins and are secreted into the extracellular matrix, associated with perlecan, agrin and collagen XVIII^{1, 2}. HS contains a disaccharide repeating unit of glucuronic acid (GlcA) or iduronic acid (IdoA) and glucosamine (GlcN). Both IdoA and GlcN are capable of carrying sulfo groups, while GlcA can be sulfated to a lesser extent. HS play important roles in a wide range of physiological

*To whom correspondence should be addressed: Rm 1044, Genetic Medicine Building, 120 Mason Farm Road, University of North Carolina, NC 27599. Tel: 919-843-6511; jian_liu@unc.edu.

Accession Codes. Coordinates have been deposited in the Protein Data Bank and assigned PDB ID codes of 5T03, 5T05, and 5T0A for substrates **I**, **II**, and **III** respectively.

Supporting Information Available: This material is available free of charge via the Internet.

and pathophysiological functions including assisting in viral/bacterial infection, inflammatory responses, blood coagulation, angiogenesis, and embryonic development³. Notably, HS is implicated in the formation of amyloid plaques contributing to Alzheimer's disease⁴. HS carry out their wide range of functions through interactions with proteins such as growth factors, protease inhibitors, proteases, cytokines, chemokines and morphogens^{1, 2, 5}. Many of these interactions rely on distinctive sulfated saccharide sequences in HS for high specificity.

The HS 6-*O*-sulfotransferase isoforms (6-OSTs) are members the HS biosynthetic enzyme family that transfer a sulfo group from 3'-phosphoadenosine 5'-phosphosulfate (PAPS) to the 6-OH of GlcN. HS biosynthesis occurs in the Golgi and endoplasm reticulum and involves glycosyltransferases, an epimerase and several sulfotransferases. A backbone polysaccharide with a disaccharide repeating unit of GlcA and *N*-acetyl glucosamine (-GlcA-GlcNAc-) is initially synthesized by HS polymerase. The backbone polysaccharide then undergoes a series of modifications (Supplementary Figure 1)⁶. In general, modification begins with *N*-deacetylation/*N*-sulfation by the *N*-deacetylase/*N*-sulfotransferases (NDSTs) to convert GlcNAc residues to *N*-sulfo GlcN (GlcNS) residues. This modification allows for C₅-epimerization of neighboring GlcA residues to form an IdoA residue, followed by sulfation at the 2-OH by the 2-*O*-sulfotransferase (2-OST). HS is further modified by 6-*O*-sulfation on GlcNS/GlcNAc residues by 6-OSTs. The final modification step involves the introduction of 3-*O*-sulfo groups on GlcN residues by 3-*O*-sulfotransferases (3-OSTs). The crystal structures of several members of the HS biosynthetic enzymes have been solved, including the *N*-sulfotransferase domain of NDST-1, 2-OST, 3-OST (isoforms 1,3 and 5) and C₅-epimerase⁷⁻¹³. The structural information, especially those complexes with bound oligosaccharide substrates, has revealed the mechanisms used by different enzymes to recognize saccharide residues in HS. Until now, the crystal structure of 6-OST has remained elusive.

6-*O*-sulfation participates in a wide range of biological functions and has been associated with a variety of human diseases. In particular, 6-*O*-sulfation is required for the formation of a ternary complex between fibroblast growth factor (FGF)/HS and the fibroblast growth factor receptor (FGFR1), which initiates signaling processes. The role of 6-*O*-sulfation is to interact with the FGFR, but not with FGF¹⁴. The dysfunction of 6-*O*-sulfation on HS has been linked with rheumatoid arthritis¹⁵, idiopathic hypogonadotropic hypogonadism¹⁶ and various cancers, as demonstrated in cell-based assays¹⁷⁻¹⁹.

Three isoforms of 6-OST are present in the human and mouse genomes. 6-OST isoform 1 knockout is lethal to mice at the late embryonic stage, displaying systematic skeletal growth retardation, aberrant lung morphology, and reduced placental angiogenesis, whereas 6-OST isoform 2 knockout mice appear to be normal and fertile²⁰. The cellular concentration of 6-*O*-sulfation is influenced by expression levels of 6-OST and endoglucosamine-6-sulfatases which selectively hydrolyze the 6-*O*-sulfo group present in the highly sulfated domains of HS containing the disaccharide repeating unit -IdoA2S-GlcNS6S-²¹.

The biosynthesis of HS is a non-template driven biosynthetic process. The control of the structures of HS largely depends on the substrate specificity of each enzyme in the

biosynthetic pathway to ensure the structural consistency between iterations. Understanding how HS biosynthesis enzymes function and interact with their substrates plays an essential role in uncovering the cellular control of HS biosynthesis. Despite extensive efforts, crystallization of human and mouse isoforms of 6-OST was unsuccessful. Here, we report ternary complex crystal structures of zebrafish 6-OST isoform 3 as a maltose-binding protein (MBP) fusion protein (zf6-OST) at 1.95 – 2.1 Å resolutions in the presence of three different homogeneous HS oligosaccharide substrates and the cofactor product, 3'-phosphoadenosine 5'-phosphate (PAP) (Figure.1, Figure 2 and Supplementary Table 1). The zf6-OST shows >70% sequence identity to the three human isoforms, offering a good model for understanding the functionality of the human orthologs. Based on these crystal structures, mutants were created and assayed against different oligosaccharide substrates to probe the roles of individual amino acids in catalysis and saccharide substrate recognition. This study reveals how 6-OST engages its substrate in an unpredicted orientation, defines key residues involved in substrate binding and catalysis, addresses substrate specificity between the three mammalian isoforms, and provides structural insight for mutants associated with human disease. As a result, several mutants were generated which may represent useful tools for chemoenzymatic methods of producing HS analogs for drug development.

Results and Discussion

Overall structure

The catalytic domain (Lys75-Ser395) of zf6-OST was crystallized as an N-terminal MBP fusion protein in ternary complexes with the cofactor product PAP and three oligosaccharides (substrates **I**, **II** and **III**) (Figure 1, Figure 2 and Supplementary Figure 4). All three structures were solved in the same crystal form (Supplementary Table 1), containing two MBP-fusion molecules in the asymmetric unit (referred to as molecule A and molecule B), with each zf6-OST domain binding simultaneously to PAP and an oligosaccharide substrate (Supplementary Figure 2). The two zf6-OST domains in the asymmetric unit superimpose well with an RMSD of 0.46 Å over 287 Ca's (Supplementary Figure 2b). The largest difference is in the position of the extended-flexible C-terminal helix (helix 11). In molecule B, a few more residues can be modeled to weak electron density with the α -helix slightly unwound at residue Gln374 and its position stabilized by lattice contacts.

Despite limited sequence identity to catalytic domains of the other HS sulfotransferases (17% to the *N*-sulfotransferase domain of NDST-1, 16% to 3-OST-1, and 23% to 2-OST), the catalytic domain of zf6-OST retains elements of the canonical α/β sulfotransferase domain fold, consisting of a central four-stranded parallel β -sheet flanked on either side by α -helices (Figure 2a). At the heart of this domain is a 5'-phosphosulfate binding-loop consisting of a strand-loop-helix motif (β 1– α 2) involved in binding and proper positioning of the cofactor PAPS for catalysis. In addition, a conserved α -helix (α 4) runs across the top of the substrate binding site interacting with both PAP and the saccharide substrate. Unique to zf6-OST is a 20 residue N-terminal coil that cradles the C-terminal helix and runs along

the N-terminal edge of the central β -sheet (Figure 2a, orange). The zf6-OST also contains five conserved disulfide bonds that are not found in the other HS sulfotransferases.

Protein-Substrate Interactions

Despite the substrates having three potential glucosamines that could bind in the active site, all three substrates bind in a similar orientation, with the nonreducing end GlcNS residue (residue 1) bound in the active site near the PAP (Figure 2, Supplementary Figure 3). Substrate **II** forms the same interactions with zf6-OST as substrate **III** however additional binding information can be obtained from the ternary structure of zf6-OST/PAP/substrate **III** due to the presence of an extra GlcA (0) on the nonreducing end of substrate **III**. Therefore, the subsequent analysis focuses on the binding of zf6-OST to substrate **III** (Figure 2a, b, and c). In this structure, the majority of interactions between zf6-OST and the substrates are with the acceptor GlcNS (1) and the adjacent disaccharide GlcA-GlcNS (residues 2 and 3) on the reducing side. The acceptor 6-hydroxyl group of GlcNS (1) lies 4.7 Å from the leaving group oxygen of the PAP molecule, displaying the anticipated orientations and distance between acceptor and donor groups for the in-line transfer of the sulfo group from PAPS (Figure 2b). Residue His158 forms a hydrogen bond with this 6-hydroxyl group, likely serving as the catalytic base for zf6-OST. In addition, the sidechain of Trp210 sits perpendicular to the face of GlcNS (1) and is also in position to form a hydrogen bond with the 6-hydroxyl group. The *N*-sulfo moiety of GlcNS (1) is within hydrogen bonding distance to backbone amides from both residues Thr209 and Trp210 and forms water mediated hydrogen bonds with Thr209 and Thr212. Residues Arg148 and Trp153 are both within hydrogen bonding distance of the carboxylate group from the adjacent GlcA (2). In addition, Lys131 may form a hydrogen bond with the carboxylate group. After GlcA (2), there is variability between the positions of the substrates with respect to the two molecules in the asymmetric unit (A and B) for each structure, and between the different substrates (Figures 2c, and 2d, Supplementary Figure 2b). However, the key interactions between the protein and substrate **II** and **III** are similar in all cases: Lys131 is within hydrogen bonding distance to the 3-hydroxyl of the next GlcNS (3) sugar; Lys 132 is within hydrogen bonding distance to both sulfo moieties from GlcNS (3) and IdoA2S (4) residues. The protein displays no interactions with the saccharides at the 5th and 6th positions.

The nonreducing end GlcA residue of substrate **III** (0) also displays interactions with the protein, revealing how the enzyme interacts with substrates beyond the nonreducing end of the acceptor GlcN, as would be the case for a longer HS substrate. In molecule A of the asymmetric unit, there is no electron density for this saccharide, suggesting it is disordered. However, in molecule B, this saccharide is refined with an occupancy of ~ 75%, consistent with the partial difference density observed in the electron density omit map (Supplementary Figure 3c). In this position, the 2-hydroxyl is in position to form a hydrogen bond with Arg148, while the carboxylate can form a hydrogen bond with Arg206 (Figure 2b). Manually modeling an additional GlcNS on the nonreducing end of substrate **III** suggests the binding path for larger HS chains lies across the face of the enzyme, crossing over the PAP and effectively burying it within the active site (Supplementary Figure 4a). The *N*-sulfo moiety of the modeled GlcNS is in position to bind to Arg112, while the 6-hydroxyl group lies near residues Arg206 and Arg329. In all three crystal structures, a tartrate molecule is

observed in molecule B of zf6-OST at this position, forming salt bridges with Arg112 that potentially mimic the proposed interaction with the *N*-sulfo group on the modeled GlcNS (Supplementary Figure 4a).

Substrate **I** differs from **II** and **III** as it contains only GlcA. The binding position of substrate **I** is similar to that of **II** and **III** for sugar residues 1 thru 3 (Figure 2d). However, substrate **I** has a GlcA residue in position 4, causing changes in the substrate orientation since the GlcA is observed in the 4C_1 conformation rather than the 2S_0 skew boat conformation of the IdoA2S in substrates **II** and **III** (Figure 2d). In molecule A of zf6-OST/PAP/substrate **I** structure, Lys131 forms similar interactions with the GlcA (2) as with substrates **II** and **III**, while Lys132 is within hydrogen bonding distance to only the sulfo group of GlcNS (3) due to the substitution of GlcA for IdoA2S at residue 4. In molecule B of the substrate **I** ternary complex, Lys131 is mostly disordered and Lys132 points away from the oligosaccharide substrate (Figure 2d, Supplementary Figure 4b). In molecule B of all three structures, there is a small amount of density consistent with a low occupancy orientation of the saccharides from residue 3 to 6 that may represent 10–20% of the bound substrate. However, this orientation cannot be modeled accurately due to its low occupancy.

Comparison of zf6-OST to other HS Sulfotransferases

The binding orientation of these substrates to zf6-OST differs tremendously from structures of the other HS sulfotransferase/substrate complexes; 3-OST-1 and -3, and 2-OST. In these sulfotransferases, the saccharide substrates are bound in an open cleft that extends across the face of the enzyme perpendicular to the PAP molecule (Figure 3a, orange). However, this open cleft is occluded in zf6-OST by a coil containing residues Thr209 and Trp210 (Figure 3a, cyan). Instead, the substrates bind to zf6-OST in an orientation perpendicular to that of the other HS sulfotransferases, likely passing over the co-factor, potentially burying PAPS within the active site (Figure 3a, green).

In spite of the differences in the global substrate binding orientation, the position of the key catalytic amino acid residues in the active sites and the orientation of the PAP and acceptor saccharide hydroxyl group from 2-OST and zf6-OST are remarkably similar (Figure 3b). Both 2-OST and zf6-OST appear to utilize a histidine residue as a general base with a Lys residue from the 5'-phosphosulfate binding loop assisting in catalysis. This differs from the NDST-1 and 3-OST isoforms that utilize a glutamate as their catalytic bases^{10, 13, 22}. The position of the acceptor 6-hydroxyl from the GlcNS (1) (acceptor residue for zf6-OST) and the 2-hydroxyl of IdoA (acceptor residue for 2-OST) superimpose perfectly in the active site and are found within hydrogen bonding distance to their proposed catalytic bases (Figure 3b). This conservation in active site geometry and key catalytic residues suggests that these enzymes utilize a similar mechanism.

The different binding orientation of substrate may allow 6-OST to exhibit a more relaxed substrate specificity than either the 2-OST or 3-OSTs, with regards to length, sequence and sulfation. 6-OST can utilize substrates as short as trisaccharides and sulfates both GlcNS and GlcNAc residues surrounded by GlcA, IdoA as well as IdoA2S residues, although the susceptibility of GlcNAc to 6-OST modification is somewhat lower²³. 2-OST demonstrates a higher level of specificity, requiring at least a pentasaccharide with *N*-sulfation, but

excluding substrates with 6-*O*-sulfation^{13, 24}. 3-OST demonstrates the greatest substrate specificity among HS sulfotransferases, requiring at least a tetrasaccharide with specific saccharide sequences and sulfation patterns within adjacent residues^{23, 25}. Furthermore, 6-OST is the only HS sulfotransferase that transfers the sulfo group to a primary alcohol; whereas 2-OST and 3-OST sulfate secondary alcohols. The fact that 6-OST sulfates a primary hydroxyl allows the bulk of the substrate to sit an additional bond length away from the active site (Figure 3b), potentially reducing the number of sugar residues that come in contact with the protein. Thus, the distinct substrate binding orientation in 6-OST may allow the enzyme to be more promiscuous, consequently generating a larger pool of diverse HS products for interaction with a wider variety of target proteins.

Mutagenesis Analysis

Site-directed mutagenesis experiments were conducted to determine the contribution of individual residues to catalysis and substrate recognition for zf6-OST using five specially designed oligosaccharide substrates (substrates **IV**, **V**, **VI**, **VII** and **VIII** (Figure 1b)). To match the mutagenesis work with the crystallography, oligosaccharides were synthesized that only allowed for sulfation at the nonreducing end GlcN site by placement of sulfo groups on the 6-hydroxyls of the remaining GlcN.

Wild type (WT) zf6-OST exhibits different levels of activity towards substrate **IV** to **VIII** (Table 1). Among them, the least reactive substrate is substrate **VI** (contains a GlcNAc residue at the acceptor), and the most reactive substrate is **VIII** (GlcNS residue as the acceptor), consistent with the fact that 6-OSTs are more reactive towards GlcNS than GlcNAc. This observation correlates with the structures, since zf6-OST forms extensive interactions with the *N*-sulfo group via residues Thr209, Trp210 and Thr212 (Figure 2b).

The relative activities of the mutants towards these substrates were also measured and normalized to levels exhibited by the wildtype enzyme (Table 1). Mutations of amino acid residues within the active site significantly reduced the activities to all tested substrates (Table 1). Mutant zf6-OSTs H101A and K104A displayed 20- to 100-fold reduction in the sulfotransferase activity towards substrates **IV** to **VII**. In the crystal structures, these residues interact with the 5'-phosphate of the PAP molecule, suggesting involvement in PAP binding and/or stabilization of the transition state (Figure 2b). Based on the crystal structure, His158 is the proposed catalytic base and directly interacts with the 6-OH acceptor of GlcNS (1). As expected, the H158A mutant displays no enzymatic activity. Mutants W210A and H203A also displayed significant reduction in the activity, as these two residues interact with GlcNS (1) and are likely involved in properly orienting the acceptor for catalysis (Figure 2b).

Mutational analysis revealed intriguing molecular interactions pertaining to substrate recognition. Based on the structures with substrates **II** and **III**, Lys132 has a unique role in the substrate recognition by interacting with an IdoA2S (4) on the reducing side of the acceptor GlcNS (1). Mutants K132A and K131A/K132A reduced the activities to substrates **IV** to **VII** (containing IdoA2S) to 4 to 8% of that for wt zf6-OST. Furthermore, substitution of Lys132 with a negatively charged amino acid residue (K132E) had an even more profound impact on substrate reactivity than that of the substitution with a neutral amino

acid residue (K132A). However, using a substrate (substrate **VIII**) with no IdoA2S residues, we found that the mutants maintained 91% and 28% reactivity for K132A and K132E, respectively. These observations support the role of K132 in enhancing binding of substrates containing IdoA2S, mediated by interaction with the 2-*O*-sulfo group (Figure 2b and 2c). The lower activity for the K132E vs K132A mutant on substrate **VIII** is likely due to repulsion with the neighboring NS group on GlcN (3).

Several mutants, including K202E, R112E, R116E and R329E, displayed different degrees of increased activity compared with wt zf6-OST. The increase is particularly profound towards Substrate **VI**. For example, the activity of R112E was 7-fold higher than that for the wild type protein. The side chain of Arg112, as well as other amino acid residues in this group, possibly interacts with saccharide substrates beyond the nonreducing end of the acceptor GlcN (1) present in the crystal structures (Figure 2b and Supplementary Figure 4a). Interactions with wild type residues may exhibit a negative impact on the activity when the acceptor is a GlcNAc residue, a known poor substrate for 6-OST. For substrate **VI**, GlcNS6S residues may compete for binding to the acceptor site placing other GlcN residues near Arg112 (modeled GlcNS, Supplementary Figure 4a) thus inhibiting productive binding. The R112A mutation could reduce this interaction while R112E would repel binding in this orientation allowing for more substrate to bind with the acceptor GlcNAc group in the active site. This effect would be reduced for substrates with GlcNS acceptor residues (substrates **IV** and **V**) as the –NS group would have enhanced binding to the active site via the hydrogen bonding network to residues T209, W210 and T212 as well as substrate **VII** due to the additional interactions with the GlcA on the nonreducing end.

Substrate Specificity of 6-OST Isoforms

6-OST is present in three different isoforms in the human genome. The zf6-OST shares 71%, 72% and 81% sequence identities with the human 6-OST isoforms 1, 2 and 3, respectively (Figure 4). Whether different 6-OST isoforms display distinct substrate specificities has been controversial. Previous reports using mouse 6-OSTs and heterogeneous polysaccharide substrates suggested these enzymes have different but somewhat overlapping substrate preferences²⁶. In contrast, later work observed that these enzymes have “qualitatively similar substrate specificities, with minor differences in target preferences²⁷.” To confirm the substrate specificity between isoforms, we measured the reactivity of m6-OST-1, h6-OST-2 and h6-OST-3 towards five oligosaccharide substrates **IV** to **VIII**. Our results show very similar trends of selectivity towards substrates **IV** to **VIII** (Supplementary Table 2). For example, Substrate **VIII** is the most reactive substrate and Substrate **IV** is the least reactive substrate for all isoforms. The data suggest that with the substrates tested, there are no major difference in substrate specificities among isoforms. This conclusion is also supported by the crystal structure. Mapping of conserved residues of the human 6-OST isoforms onto the zf6-OST structure reveals that all residues that interact with HS in these structures and surround the HS binding pocket are totally conserved in all three human isoforms (Figure 3c). The major sequence differences between the enzymes reside in the partially disordered C-terminal tail extending away from the catalytic domain, distal from the substrate binding site in the crystal structures.

IHH Associated Mutations

6-OST is a critical enzyme in the HS biosynthetic pathway producing 6-*O*-sulfated GlcN residues in HS to display various biological functions. Mutations in h6-OST-1 have been associated with idiopathic hypogonadotrophic hypogonadism (IHH), a genetic disease manifested by delayed or lack of puberty¹⁶. Mutations R306W, R306Q, R323Q, R382W and M404V in 6OST-1 were reportedly found in some IHH patients¹⁶. In a study by Tornberg *et al*, all four mutants exhibited modest reduction in sulfotransferase activity in an *in vitro* assay¹⁶. Three of the four disease-related residues (except M404) are conserved in zf6-OST and map to the structure, providing some insight for the reduction in their activity (Figure 3d and Fig 4.). Most notable is the R323Q mutation, which is equivalent to Arg329 in zf6-OST. Tornberg and colleagues suggested that Arg323 may interact with an *O*-sulfo moiety from HS. The zebrafish equivalent (Arg329) is located in the substrate binding pocket near the nonreducing end of substrate **III**. Modeling an additional GlcNS on to the end of substrate **III** positions the 6-hydroxyl of the modeled GlcNS in close proximity to Arg329 (Supplementary Figure 4a). Based on the modeling, it is plausible that this Arg residue could interact with a 6-*O*-sulfo group on HS, possibly contributing to creation of contiguously 6-*O*-sulfated NS domains. This is supported by Tornberg *et al*'s *in vitro* work which displayed a reduction in activity for R323Q against HS but maintained similar activity to wild type enzyme for substrate lacking 2-*O*- and 6-*O*- sulfo groups. In our mutational analysis, R329A of zf6OST-3 has very similar sulfotransferase activity to that of wild type protein (Table 1). This result is somewhat unsurprising as this residue does not interact with our substrates (based on the crystal structures) due to the lack of an extended nonreducing end in the synthesized substrates. To prove this hypothesis, a series of specially designed oligosaccharide substrates would be required. The synthesis of the oligosaccharide substrates with the GlcNS6S residue at the nonreducing end but GlcNS residue at the reducing end is currently beyond our synthetic capability. Zebrafish residue Arg312 (human Arg306) is found in the catalytic core away from the substrate binding site and forms a hydrogen bond with Glu372 (human Glu366) of the C-terminal α -helix 11 suggesting it is not involved in binding substrate, but potentially affects enzyme stability. Arg388 (human Arg382) is found at the very C-terminus of our structure far from the active site, and forms no interaction with the rest of the protein (Figure 3d). This makes it difficult to interpret possible effects of the R382W mutation on activity and disease.

Improving Chemoenzymatic Synthesis of HS

A cost-effective method to prepare diversified HS oligosaccharides with different sulfation patterns is critical for developing HS-based therapeutics with high specificity for different targets. One particular example is to prepare synthetic heparins for anticoagulant therapy²⁸. The chemoenzymatic method which utilizes 6-OST is becoming a promising approach to synthesize structurally defined HS oligosaccharides²⁹. The structure of zf6-OST revealed potential locations to engineer amino acid changes for synthesis of specific HS oligosaccharides. For example, mutant K132E effectively sulfates the substrates containing GlcA, but has very low reactivity towards IdoA2S-containing oligosaccharides. The K132 mutants offer the possibility to engineer the substrate specificity of zf6-OST for synthesis of oligosaccharides containing GlcA residues, but no IdoA2S residues, around the modification

site. In addition, several mutants, K202E, R112E, R116E and R329E, were identified to display significantly increased activity towards GlcNAc residues. These serendipitous mutations provide new tools to generate greater yields of GlcNAc6S-containing saccharides. Thus, in addition to expanding our general knowledge of the HS biosynthesis pathway, the results from this study clearly provide the structural information to engineer 6-OST with expanded capability to synthesize HS based therapeutics.

Methods

Crystallization/Structure Determination

Zf6-OST was expressed and purified as a fusion with the maltose binding protein as described in the Supporting Information. Crystals of zf6-OST in the presence of 10 mM substrate **II** were grown by hanging drop vapor diffusion with 2 μ l of protein solution at 13 mg ml⁻¹ and 2 μ l of reservoir solution containing 100 mM Tris pH 7.5, 200 mM disodium tartrate and 18–19% PEG3350 (w/v) at 20° C. Crystals containing substrates **I** and **III** were grown by streak-seeding hanging drops of zf6-OST with 10mM of substrate **I** or **III** with seeds from crystals of protein with substrate **II** at 20° C. For data collection, 1–2 μ l of cryoprotectant solution (100 mM Tris pH 7.5, 125 mM NaCl, 200mM disodium tartrate, 21% PEG3350 (w/v), 1 mM PAP, 4 mM maltose, 15% ethylene glycol (v/v) and 1 mM of the appropriate substrate (**I**, **II**, or **III**), was added to the crystallization hanging drop. The crystals were then transferred directly to cryoprotectant solution and flash frozen in liquid nitrogen. Data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory at a wavelength of 1.0 Å. All data were processed with HKL2000³⁰. To solve the structure of zf6-OST with substrate **II** present, molecular replacement was carried out in Phenix³¹ with Phaser³² using a model of MBP (PDB code 3DM0) along with a trimmed polyalanine model of 2-OST (PDB code 3F5F) containing only well conserved secondary structure. The structure was refined in Phenix using Cartesian simulated annealing, positional, ADP refinement, and manual model building in Coot^{33, 34}. The structures of zf6-OST containing substrates **II** and **III** were subsequently solved using the structure of zf6-OST with substrate **II**. The same R_{free} flagged reflections were maintained for all three data sets. The structures possessed good geometry, with all three structures having greater than 97.8% of the residues in the Ramachandran favored region, with 0.1% outliers and overall refinement scores of 1.1, as assessed by MolProbity³⁵ (Supplementary Table 1).

Preparation of 6-OSTs Mutants and Isoforms

Mutants were generated on the pMALX MBP-fusion zf6-OST construct used for crystallization, using QuikChange (Agilent) mutagenesis. The mutant plasmids were confirmed by DNA sequencing analysis. The resultant mutants were expressed and purified in 1L scale following the same protocol as used for the wild type protein. Mammalian isoforms of mouse 6-OST-1 and human 6-OST isoforms 2 and 3 were expressed and purified from SF9 insect cells and is described in detail in the Supporting Information.

Determination of the Enzymatic Activity of zf6-OST Mutants and Mammalian Isoforms

The synthesis of substrates were completed according to the chemoenzymatic method published previously^{36, 37} and are described in more detail in Supporting Information. Activity measurement for wildtype and mutant proteins were carried out by incubating 1 µg of wildtype or mutant zf6-OST with 0.1 nmol of substrate (**IV** to **VIII**) and 1×10^5 cpm of [³⁵S]PAPS (0.1nmol) in 100 µL of buffer containing 50 mM MES, pH 7.2 at 37 °C for 2.5 hours. Reactions were quenched by the addition of 6 M urea and 100 mM EDTA then subjected to a 200-µL DEAE-Sephareose chromatography to purify the ³⁵S-labeled oligosaccharide. The quantity of [³⁵S] sulfated oligosaccharide was determined by liquid scintillation counting. The negative control contained all of the components without enzyme.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This research was supported in part by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences; Grant 1ZIA ES102645 (to L.C.P.), and the US Department of Energy, Office of Science, Office of Basic Energy Sciences Contract W-31-109-Eng-38; and by grants from National Institutes of Health (GM102137 and HL094463, to J.L.) and by a grant from Eshelman Institute for Innovation (to J.L.). The authors would like to thank M. Sobhany and P. Tumbale for critical reading of the manuscript.

References

- (1). Lindahl U, Li JP. Interactions between heparan sulfate and proteins-design and functional implications. *Int Rev Cell Mol Biol.* 2009; 276:105–159. [PubMed: 19584012]
- (2). Sarrazin S, Lamanna WC, Esko J. Heparan sulfate proteoglycans. *Cold Spring Harb. Perspect. Biol.* 2011; 3:a004952. [PubMed: 21690215]
- (3). Bishop J, Schuksz M, Esko JD. Heparan sulphate proteoglycans fine-tune mammalian physiology. *Nature.* 2007; 446:1030–1037. [PubMed: 17460664]
- (4). Liu CC, Zhao N, Yamaguchi Y, Cirrito JR, Kanekiyo T, Holtzman DM, Bu G. Neuronal heparan sulfates promotes amyloid pathology by modulating brain amyloid- β clearance and aggregation in Alzheimer's disease. *Sci. Transl. Med.* 2016; 8:1–11.
- (5). Xu D, Esko J. Demystifying heparan sulfate-protein interactions. *Annu Rev Biochem.* 2014; 83:129–157. [PubMed: 24606135]
- (6). Esko JD, Selleck SB. Order out of chaos: Assembly of ligand binding sites in heparan sulfate. *Ann. Rev. Biochem.* 2002; 71:435–471. [PubMed: 12045103]
- (7). Kakuta Y, Sueyoshi T, Negishi M, Pedersen LC. Crystal structure of the sulfotransferase domain of human heparan sulfate N-deacetylase/N-sulfotransferase 1. *J. Biol. Chem.* 1999; 274:10673–10676. [PubMed: 10196134]
- (8). Bethea HN, Xu D, Liu J, Pedersen LC. Redirecting the substrate specificity of heparan sulfate 2-O-sulfotransferase by structurally guided mutagenesis. *Proc. Natl. Acad. Sci. USA.* 2008; 105:18724–18729. [PubMed: 19022906]
- (9). Edavettal SC, Lee KA, Negishi M, Linhardt RJ, Liu J, Pedersen LC. Crystal structure and mutational analysis of heparan sulfate 3-O-sulfotransferase isoform 1. *J. Biol. Chem.* 2004; 279:25789–25797. [PubMed: 15060080]
- (10). Moon A, Edavettal SC, Krahn JX, Munoz EM, Negishi M, Linhardt RJ, Liu J, Pedersen LC. Structural analysis of the sulfotransferase (3-OST-3) involved in the biosynthesis of an entry receptor of herpes simplex virus 1. *J. Biol. Chem.* 2004; 279:45185–45193. [PubMed: 15304505]
- (11). Xu D, Moon A, Song D, Pedersen LC, Liu J. Engineering sulfotransferases to modify heparan sulfate. *Nat. Chem. Biol.* 2008; 4:200–202. [PubMed: 18223645]

- (12). Qin Y, Ke J, Gu X, Fang J, Wang W, Cong Q, Li J, Tan J, Brunzelle JS, Zhang C, Jiang Y, Melcher K, Li JP, Xu HE, Ding K. Structural and functional study of D-glucuronyl C5-epimerase. *J. Biol. Chem.* 2015; 290:4620–4630. [PubMed: 25568314]
- (13). Moon AF, Xu Y, Woody S, Krahn JM, Linhardt RJ, Liu J, Pedersen LC. Dissecting substrate recognition mode of heparan sulfate 3-O-sulfotransferases for the biosynthesis of anticoagulant heparin. *Proc. Natl. Acad. Sci. USA.* 2012; 109:5256–5270.
- (14). Schlessinger J, Plotnikov AN, Ibrahimi OA, Eliseenkova AV, Yeh BK, Yayon A, Linhardt RJ, Mohammadi M. Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol. Cell.* 2000; 6:743–750. [PubMed: 11030354]
- (15). Sabol JK, Wei W, López-Hoyos M, Seo Y, Andaya A, Leary JA. Heparan sulfate differences in rheumatoid arthritis versus healthy sera. *Matrix Biol.* 2014; 40:54–61. [PubMed: 25217862]
- (16). Tornberg J, Sykiotis GP, keefe K, Plummer L, Hoang X, Hall JE, Quinton R, Seminara SB, Hughes V, AVan Vliet G, Van Uum S, Crowley WF, Habuchi H, Kimata K, Pitteloud N, Bulow HE. Heparan sulfate 6-O-sulfotransferase 1, a gene involved in extracellular sugar modifications, is mutated in patients with idiopathic hypogonadotropic hypogonadism. *Proc. Natl. Acad. Sci. USA.* 2011; 108:11524–11529. [PubMed: 21700882]
- (17). Jayson GC, Lyon M, Paraskeva C, Turnbull JE, Deakin JA, Gallagher JT. Heparan sulfate undergoes specific structural changes during the progression from human colon adenoma to carcinoma in vitro. *J. Biol. Chem.* 1998; 273:51–57. [PubMed: 9417046]
- (18). Delcommenne M, Klingemann HG. Detection and characterization of syndecan-1-associated heparan sulfate 6-O-sulfated motifs overexpressed in multiple myeloma cells using single chain antibody variable fragments. *Hum. Antibodies.* 2012; 21:29–40. [PubMed: 22885958]
- (19). Ferreras C, Rushton G, Cole CL, Babur M, Telfer BA, van Kuppevelt TH, Gardiner JM, Williams KJ, Jayson GC, Avizienyte E. Endothelial heparan sulfate 6-O-sulfation levels regulate angiogenic responses of endothelial cells to fibroblast growth factor 2 and vascular endothelial growth factor. *J. Biol. Chem.* 2012; 287:36132–36146. [PubMed: 22927437]
- (20). Habuchi H, Kimata K. Mice deficient in heparan sulfate 6-O-sulfotransferase-1. *Prog. Mol. Biol. Trans. Sci.* 2010; 93:79–111.
- (21). Pompe E, Burch TC, Law CJ, Liu J. Substrate specificity of 6-O-endosulfatase 2 and its implications in synthesizing anticoagulant heparan sulfate. *Glycobiology.* 2012; 22:1353–1362. [PubMed: 22692045]
- (22). Kakuta Y, Li L, Pedersen LC, Pedersen LG, Negishi M. Heparan sulphate N-sulphotransferase activity: reaction mechanism and substrate recognition. *Biochem. Soc. Trans.* 2003; 31(pt2):331–334. [PubMed: 12653630]
- (23). Liu R, Liu J. Enzymatical placement of 6-O-sulfo groups in heparan sulfate. *Biochemistry.* 2011; 50:4382–4391. [PubMed: 21506605]
- (24). Liu C, Sheng J, Krahn JM, Perera L, Xu Y, Hsieh P, Duo W, Liu J, Pedersen LC. Molecular mechanism of substrate specificity for heparan sulfate 2-O-sulfotransferase. *J. Biol. Chem.* 2014; 289:13407–13418. [PubMed: 24652287]
- (25). Xu D, Song D, Pedersen L, Liu J. Mutational study of heparan sulfate and chondroitin sulfate 2-O-sulfotransferases. *J. Biol. Chem.* 2007; 282:8356–8367. [PubMed: 17227754]
- (26). Habuchi H, Tanaka M, Habuchi O, Yoshida K, Suzuki H, Ban K, Kimata K. The occurrence of three isoforms of heparan sulfate 6-O-sulfotransferase having different specificities for hexuronic acid adjacent to the targeted N-sulfoglucosamine. *J. Biol. Chem.* 2000; 275:2859–2868. [PubMed: 10644753]
- (27). Smeds E, Habuchi H, Do A-T, Hjertson E, Grundberg H, Kimata K, Lindahl U, Kusche-Gullberg M. Substrate specificities of mouse heparan sulphate glucosaminyl 6-O-sulfotransferases. *Biochem. J.* 2003; 372:371–380. [PubMed: 12611590]
- (28). Linhardt RJ, Liu J. Synthetic heparin. *Curr Opin Pharmacol.* 2012; 12:217–219. [PubMed: 22325855]
- (29). Liu J, Linhardt RJ. Chemoenzymatic synthesis of heparan sulfate and heparin. *Nat. Prod. Rep.* 2014; 31:1676–1685. [PubMed: 25197032]
- (30). Otwinowski Z, Minor V. Processing of x-ray diffraction data collected in oscillation mode. *Meth. Enzymol.* 1997; 276:307–326.

- (31). Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 2010; 66:213–221. [PubMed: 20124702]
- (32). McCoy AJ. Solving structures of protein complexes by molecular replacement with Phaser. *J. Appl. Crystallogr.* 2007; 40:658–674. [PubMed: 19461840]
- (33). Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D. Biol Crystallogr.* 2004; 60(pt 12 Pt 1):2126–2132. [PubMed: 15572765]
- (34). Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 2010; 66:486–501.
- (35). Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D. Biol Crystallogr.* 2010; 66:12–21. [PubMed: 20057044]
- (36). Xu Y, Masuko S, Takieddin M, Xu H, Liu R, Jing J, Mousa SA, Linhardt RJ, Liu J. Chemoenzymatic synthesis of homogeneous ultra-low molecular weight heparin. *Science.* 2011; 334:498–501. [PubMed: 22034431]
- (37). Xu Y, Cai C, Chandarajoti K, Hsieh P, Lin Y, Pham TQ, Sparkenbaugh EM, Sheng J, Key NS, Pawlinski RL, Harris EN, Linhardt RJ, Liu J. Homogeneous and reversible low-molecular weight heparins with reversible anticoagulant activity. *Nat. Chem. Biol.* 2014; 10:248–250. [PubMed: 24561662]

Figure 1.

Structures of oligosaccharides used in this study. a) Chemical structures of oligosaccharides used in the crystallization experiments. The acceptor GlcNS residue found in the active site of *zf6*-OST is depicted in a blue box. The IdoA2S residue (4) in substrate **II** and **III** is presented in the 2S_0 skew boat conformation as found in the co-crystal structures. b) Chemical structures of oligosaccharides used to probe substrate specificities of mutants. The 6-*O*-sulfation site in each substrate is circled. The IdoA2S residue in substrate **IV** to **VII** is presented in 1C_4 conformation. Representative electrospray ionization mass spectrometry spectra are shown in Supplementary Figure 5. The abbreviated sequences for each substrate is: Substrate **I** (6-mer), GlcNS-GlcA-GlcNS-GlcA-GlcNS-GlcA-pNP; Substrate **II** (6-mer), GlcNS-GlcA-GlcNS-IdoA2S-GlcNS-GlcA-pNP; Substrate **III** (7-mer), GlcA-GlcNS-GlcA-GlcNS-IdoA2S-GlcNS-GlcA-pNP; Substrate **IV** (8-mer), GlcNS-GlcA-GlcNS6S-IdoA2S-GlcNS6S-IdoA2S-GlcNS6S-GlcA-pNP; Substrate **V** (9-mer), GlcA-GlcNS-GlcA-GlcNS6S-IdoA2S-GlcNS6S-IdoA2S-GlcNS6S-GlcA-pNP; Substrate **VI** (8-mer), GlcNAc-GlcA-GlcNS6S-IdoA2S-GlcNS6S-IdoA2S-GlcNS6S-GlcA-pNP; Substrate **VII** (9-mer), GlcA-GlcNAc-GlcA-GlcNS6S-IdoA2S-GlcNS6S-IdoA2S-GlcNS6S-GlcA-pNP; Substrate **VIII** (9-mer), GlcA-GlcNS-GlcA-GlcNS6S-GlcA-GlcNS6S-GlcA-GlcNS6S-GlcA-pNP.

Figure 2.

Crystal structure of zf6-OST. a) Catalytic domain of zf6OST (blue) with substrate **III** (green) and the cofactor product PAP (magenta) bound to the active site of molecule B. Disulfide bonds are displayed in yellow and the N-terminal tail unique to the zf6-OST enzymes is shown in orange. The C-terminal α -helix extending away from the subdomain structure has been cropped in this image. b) Interactions between substrate **III** and zf6-OST. Hydrogen bonds between substrate **III** or PAP and zf6-OST are shown in dashed black lines. Waters are shown as red spheres. A red asterisk marks the acceptor 6-hydroxyl. c) Superposition of molecules A (gray) and B (blue) of zf6-OST in the asymmetric unit with substrate **III** bound. Hydrogen bonds between molecule A and B and their respective substrates (white and green respectively) are displayed as pink and black dashed lines respectively. d) Superposition of zf6-OST binding substrate **I** (pink) with substrate **III** (protein in blue, substrate green) in molecule B for each structure. The SDS-PAGE gel picture of zf6-OST is shown in Supplementary Figure 6.

Figure 3.

Structural analysis of the zf6-OST crystal structure. a) Superposition of the crystal structure of 2-OST (gray, PDB code:4NDZ) with bound substrate (orange) with zf6-OST (blue) with substrate **III** (green). The loop in zf6-OST that blocks the substrate channel in other HS sulfotransferases is colored cyan. b) Superposition of catalytic active site residues and acceptor saccharide residues of 2-OST and zf6-OST with hydrogen bonds shown (pink and black respectively). c) Conserved residues between human 6-OST-1, 2, and 3 isoforms mapped onto the structure of the catalytic domain of zf6-OST with bound substrate **III** (green). Residues that are totally conserved are colored gray, residues conserved between isoforms 1 and 2 are shown colored cyan, 1 and 3 in pink and 2 and 3 in orange. Residues that differ in all isoforms are colored purple. d) Positions of mutations (cyan) found in IHH patients mapped onto the zf6-OST structure with bound substrate **III**. Numbering is for the human 6-OST-1 enzyme.

Figure 4.

Alignment of catalytic domain sequences from zf6-OST with that of the three human isoforms h6-OST-1, -2, and -3. Secondary structural elements from the crystal structure are displayed above the sequence with β-strands depicted as blue arrows and α-helices depicted as teal rectangles. Conserved catalytic residues are colored red, His101 which may stabilize the transition state is colored blue, residues that interact with the oligosaccharides in the crystal structure are green, and residues that could possibly form additional interactions with longer substrates are colored magenta. Missense mutations suggested to contribute to idiopathic hypogonadotropic hypogonadism (IHH) are boxed in cyan. N-terminal region is boxed in orange.

Table 1

Determination of zf6-OST mutants to different oligosaccharide substrates

	Mutants	Substrate IV ^a	Substrate V ^b	Substrate VI ^c	Substrate VII ^d	Substrate VIII ^e
	WT	100 ± 3.0% ^f	100 ± 6.5%	100 ± 31%	100 ± 8.6%	100 ± 3.7%
	Relative activity to substrate IV	-	85%	17%	41%	149%
Catalytic Active Site Residues	H101A	1.3 ± 0.2%	1.0 ± 0.1%	2.5 ± 0.2%	1.2 ± 0.5%	ND ^g
	K104A	2.8 ± 0.7%	3.8 ± 0.4%	5.1 ± 1%	3.0 ± 0.2%	ND ^g
	H158A	0.3 ± 0.2%	0.3 ± 0.0%	1.2 ± 0.5%	0.4 ± 0.0%	ND ^g
	H203A	10 ± 0.1%	9.8 ± 0.5%	13 ± 2.1%	7.9 ± 0.5%	ND ^g
	W210A	0.6 ± 0.1%	0.7 ± 0.2%	2.7 ± 2.0%	0.4 ± 0.1%	ND ^g
Extended Saccharide Binding Residues	K131A	62 ± 9.0%	37 ± 2.2%	35 ± 4.6%	35 ± 0.7%	ND ^g
	K132A	8.0 ± 0.2%	4.4 ± 0.1%	6.4 ± 0.7%	4.3 ± 0.3%	91 ± 15%
	K132E	0.9 ± 0.0%	0.9 ± 0.0%	2.3 ± 0.5%	1.0 ± 0.2%	28 ± 6.4%
	K131A/K132A	2.4 ± 0.2%	1.6 ± 0.4%	2.9 ± 0.2%	1.4 ± 0.1%	80 ± 4.5%
	R148A	17 ± 0.2%	46 ± 1.7%	9.5 ± 0.4%	12 ± 0.3%	ND ^g
	W153A	34 ± 1.1%	41 ± 1.7%	40 ± 2.4%	86 ± 10.4%	ND ^g
	R206A	163 ± 16%	70 ± 3.3%	126 ± 21%	33 ± 8.6%	100 ± 3.2%
	R206E	110 ± 25%	155 ± 4.0%	497 ± 13%	87 ± 1.6%	ND ^g
Potential Nonreducing End HS Binding Residues	K202A	132 ± 1.5%	101 ± 3.9%	85 ± 3.4%	114 ± 0.7%	100 ± 1.4%
	K202E	110 ± 25%	155 ± 4%	497 ± 13%	87 ± 1.6%	ND ^g
	R112A	127 ± 2.6%	206 ± 3.9%	310 ± 27%	276 ± 17%	ND ^g
	R112E	109 ± 11%	194 ± 12%	700 ± 17%	349 ± 27%	ND ^g
	R116A	130 ± 6.8%	191 ± 3.9%	305 ± 39%	231 ± 24%	ND ^g
	R116Q	136 ± 1.0%	173 ± 4.4 %	231 ± 15%	228 ± 45%	ND ^g
	R116E	125 ± 5.4%	204 ± 0.4%	512 ± 109%	337 ± 6.3%	ND ^g
	R329A	159 ± 12%	141 ± 2.3%	144 ± 9.2%	165 ± 0.3%	103 ± 0.0%
	R329E	144 ± 4.3%	150 ± 26%	450 ± 75%	263 ± 15%	ND ^g

^aThe reactivity of wild type enzyme to substrate **IV** is 8,368 ± 918 pmol mg protein⁻¹ h⁻¹ (n = 6).^bThe reactivity of wild type enzyme to substrate **V** is 7,088 ± 1,403 pmol mg protein⁻¹ h⁻¹ (n = 6).^cThe reactivity of wild type enzyme to substrate **VI** is 1,415 ± 223 pmol mg protein⁻¹ h⁻¹ (n = 6).^dThe reactivity of wild type enzyme to substrate **VII** is 3,460 ± 282 pmol mg protein⁻¹ h⁻¹ (n = 6).^eThe reactivity of wild type enzyme to substrate **VIII** is 12,514 ± 3,949 pmol mg protein⁻¹ h⁻¹ (n = 4).^fThe values for activity for each mutant are normalized to levels exhibited by the wildtype enzyme, and represent the mean ± standard deviation.

^gNot determined

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript