Redirecting the substrate specificity of heparan sulfate 2-*O*-sulfotransferase by structurally guided mutagenesis

Heather N. Bethea[a], Ding Xu[a], Jian Liu[a,1], and Lars C. Pedersen[b]

[a]Division of Medicinal Chemistry and Natural Products, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599; and [b]Laboratory of Structural Biology, National Institute on Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709

Edited by Chi-Huey Wong, Academia Sinica, Taipei, Taiwan, and approved October 17, 2008 (received for review July 19, 2008)

Heparan sulfate (HS) is a polysaccharide involved in essential physiological functions from regulating cell growth to blood coagulation. HS biosynthesis involves multiple specialized sulfotransferases such as 2-*O*-sulfotransferase (2OST) that transfers the sulfo group to the 2-OH position of iduronic acid (IdoA) or glucuronic acid (GlcA) within HS. Here, we report the homotrimeric crystal structure of 2OST from chicken, in complex with 3′-phosphoadenosine 5′-phosphate. Structural based mutational analysis has identified amino acid residues that are responsible for substrate specificity. The mutant R189A only transferred sulfates to GlcA moieties within the polysaccharide whereas mutants Y94A and H106A preferentially transferred sulfates to IdoA units. Our results demonstrate the feasibility for manipulating the substrate specificity of 2OST to synthesize HS with unique sulfation patterns. This work will aid the development of an enzymatic approach to synthesize heparin-based therapeutics.

crystal structure | heparin | protein engineering | maltose-binding protein chimera | trimeric proteins

Heparan sulfate (HS) is a highly sulfated linear polysaccharide ubiquitously present on the cell surface and in the extracellular matrix. HS participates in a wide range of physiological and pathophysiological functions, including embryonic development, inflammatory response, blood coagulation, and assisting viral/bacterial infections (1). Heparin, a special form of HS, is a commonly used anticoagulant drug. The sulfation pattern of the HS polysaccharide governs its functional selectivity (2). Understanding the structure–function relationship of HS may allow us to manipulate HS biosynthesis to design HS/heparin with improved anticoagulant efficacy and exploit heparin or heparin-like molecules for the development of anticancer and antiviral drugs (3). However, obtaining HS oligosaccharides or polysaccharides with defined structures remains a challenge. Despite many examples of success with chemical synthesis of short-HS fragments, the synthesis of molecules larger than hexasaccharides is extremely difficult. Using HS biosynthetic enzymes to prepare biologically active polysaccharides and oligosaccharides has offered a promising alternative approach (4–7).

The biosynthesis of HS involves a series of specialized enzymes, including glycosyl transferases, an epimerase, and various sulfotransferases [supporting information (SI) Fig. S1]. 2-*O*-sulfotransferase (2OST) transfers a sulfo group to the 2-OH position of iduronic acid (IdoA) or GlcA within the HS chain to form IdoA2S and GlcA2S respectively (Fig. 1) (8). The functions of 2-*O*-sulfated iduronic acid (IdoA2S) in HS have been demonstrated by in vitro and in vivo studies. Unlike IdoA2S, GlcA2S is much less abundant in the HS isolated from natural sources, and thus the function of GlcA2S is unknown. The IdoA2S unit is necessary for binding to numerous fibroblast growth factors and is crucial for triggering fibroblast growth factor-mediated signal transduction pathways (9). Only one isoform of 2OST is present in most vertebrate genomes, and the sequence is highly conserved across species. Human 2OST shares 97% sequence identity with the mouse ortholog and 92% identity with the chicken. The high degree of sequence identity suggests a central physiological role in these organisms. The 2OST-knockout mice demonstrate renal agenesis and die in the neonatal period, and they show retardation of eye development and skeletal overmineralization (10). In *Caenorhabditis elegans*, 2OST has been determined to be essential for axon migration/guidance and nervous system development (11, 12). Dependence on 2-*O*-sulfation is notably less critical in *Drosophila* fibroblast growth factor signaling. It was suggested that the *2OST*$^{-/-}$ mutant synthesizes HS with a higher level of 6-*O*-sulfation, which may compensate for the reduced 2-*O*-sulfation in *Drosophila* (13). Here, we report the crystal structure of chicken 2OST as a maltose-binding protein (MBP) fusion protein in complex with 3′-phosphoadenosine 5′-phosphate (PAP). Our results provide structural and biochemical information for delineating the mechanism of action of 2OST and evidence for altering the substrate specificity of 2OST for the synthesis of HS with unique biological functions.

Results and Discussion

Comparison with Other Sulfotransferases. A crystal structure of MBP and the catalytic domain of chicken 2OST (D69–N356) was solved at the resolution of 2.65 Å (Fig. S2 and Table S1). The 2OST catalytic domain comprises a conical sulfotransferase α/β motif that contains the PAPS-binding loop (Fig. 2*A* and Fig. S2 *b* and *c*). Despite their global similarities, the crystal structure of 2OST has distinct differences from other HS sulfotransferases, including *N*-sulfotransferase domain of *N*-deacetylase/*N*-sulfotransferase-1 (NST) and the 3OST isoforms 1, 3, and 5 for which crystal structures exist (14–17). Sequence alignments of 2OST to NST and 3OST isoforms only show ≈12–13% sequence identity compared with 40–50% among the 3OST isoforms and ≈28% between NST and the 3-*O*-sulfotransferase (3OST) isoforms. The degree of sequence similarity is consistent with structural comparisons because NST and 3OST-3 share 234 structurally equivalent Cαs with a rmsd of 1.3 Å, whereas structural comparisons of 2OST with 3OST-3 reveal only 103 structurally equivalent residues with a rmsd of 2.1 Å. Noticeably absent in 2OST vs. NST and the 3OSTs is an exterior 3-stranded antiparallel β-sheet containing a disulfide bond and a small loop off the last strand of the central parallel β-sheet where the proposed catalytic glutamic acid base (Glu-184 in

Author contributions: J.L. and L.C.P. designed research; H.N.B., D.X., J.L., and L.C.P. performed research; H.N.B., D.X., J.L., and L.C.P. analyzed data; and H.N.B., J.L., and L.C.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 3F5F).

[1]To whom correspondence should be addressed at: Room 309, Beard Hall, University of North Carolina, Chapel Hill, NC 27599. E-mail: jian_liu@unc.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0806975105/DCSupplemental.

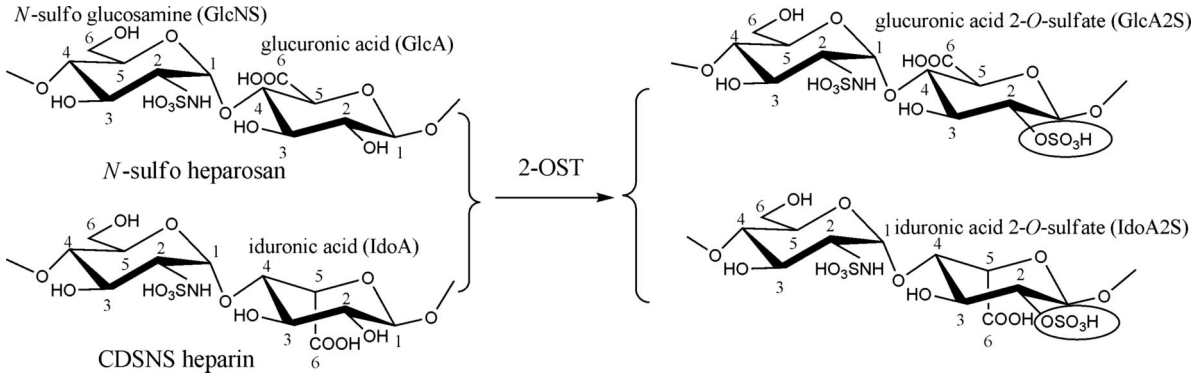

Fig. 1. Reactions catalyzed by 2OST. 2OST sulfates either the GlcA unit present in *N*-sulfo heparosan or the IdoA unit present in CDSNS heparin. If the polysaccharide contains both GlcA and IdoA units, 2OST prefers IdoA.

3OST-3) resides. However, the crystal structure of 2OST exhibits a unique elongated C-terminal helix followed by a 20-aa tail extending away from the catalytic domain (Fig. S2 *b* and *c*).

HS sulfotransferases are type II transmembrane Golgi-associated enzymes. Cytosolic sulfotransferases differ from their Golgi-associated counterparts in that they have a buried hydrophobic binding pocket for the acceptor substrate, such as a hydrophobic steroid (18). A structural comparison of 2OST with the cytosolic enzyme, estrogen sulfotransferase (EST), shows a rmsd of 2.0 Å for 107 structurally equivalent cαs but with a slightly higher sequence identity (≈16.5%) (19). Interestingly, 7 of the 45 sequence conserved residues come from the long α-helix (V182–L188) equivalent to helix 6 in EST, that runs across the top of both the donor and acceptor substrate-binding pockets (Fig. 2*A*). In addition, like 2OST, cytosolic sulfotransferases lack a catalytic glutamate. Instead, the cytosolic enzymes use a histidine at the end of the 4th strand of their central β-sheet (19). A structurally equivalent histidine (His-142) is present in 2OST (Fig. 2*A* and *B*). Xu *et al.* (20) had identified this histidine as a potential catalytic residue based on comparisons with the cytosolic enzymes by using sophisticated sequence alignments and mutational analysis. In the 2OST structure, His-142 has 2 apparent conformations. One conformation positions the side chain away from the substrate binding cleft (conformation A). The other conformation (conformation B) superimposes very well with His-108 of EST, positioning the side chain in the active-site cleft for a possible role in catalysis (Fig. 2*A*).

Trimeric Complex. The most distinct structural difference between 2OST and the other HS sulfotransferases is that 2OST appears to function as a trimer. In the crystal structure of MBP-2OST, a 3-fold crystallographic axis runs through the middle of the trimer (Fig. 3 *A* and *B*). This arrangement buries 24% of the surface area of any 1 molecule. Supporting this quaternary structure, the fusion protein appears to behave as a trimer in solution as determined by gel filtration chromatography (Fig. S3). The molecular mass of monomeric MBP-2OST is ≈70 kDa. Using a standard curve generated

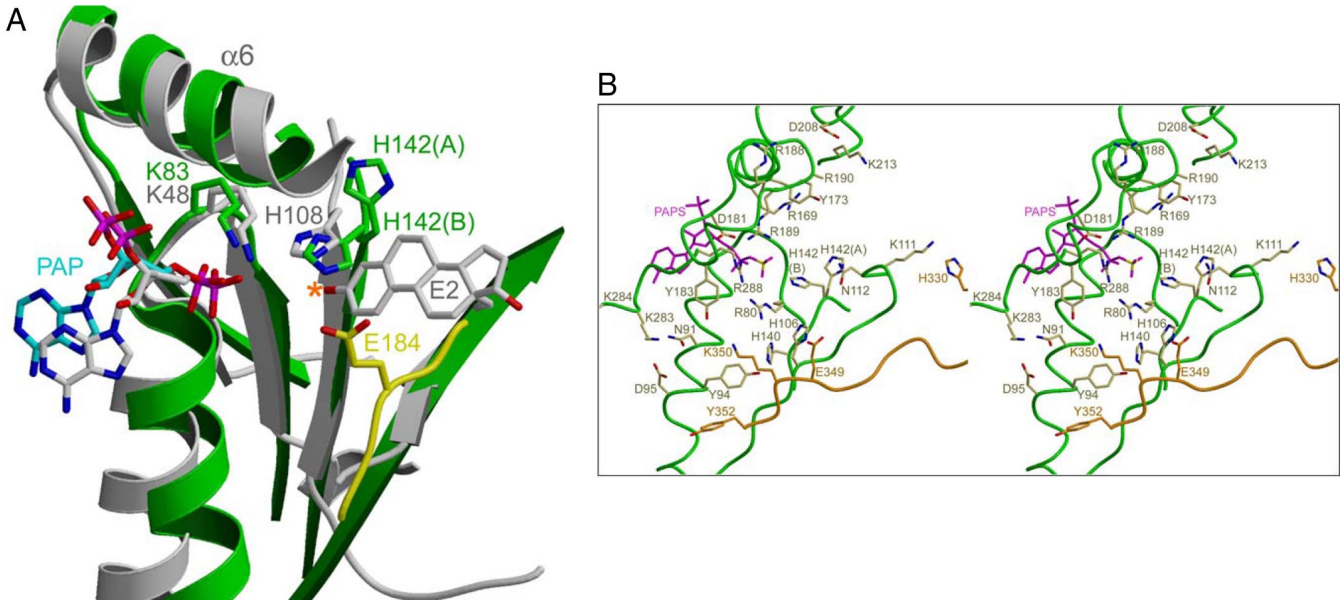

Fig. 2. Active site residues of 2OST. (*A*) Superposition of active sites of 2OST (green) with PAP (cyan) and the cytosolic EST enzyme (gray) with PAP and acceptor substrate 17β-estradiol (E2). The figure displays the position of the acceptor OH of E2 with respect to the proposed catalytic base His-108 in EST. The structurally equivalent His-142 of 2OST is shown in the 2 conformations displayed in the crystal structure. In contrast, the position of the proposed catalytic base in 3OST-3 (yellow), Glu-184, is shown originating from a different location in the structure, yet the acceptor (data not shown) location is similar to that of the OH of E2 in EST (marked by *). (*B*) Stereo diagram of the proposed acceptor-binding cleft of a 2OST molecule (green) and the adjacent C-terminal tail of another 2OST molecule in the trimer (orange). PAPS (purple) is based on a superposition of PAPS from an EST structure with bound PAPS (Protein Data Base ID code 1HY3) onto the PAP molecule in 2OST. Residues shown lining the cleft were mutated in this work (Table 1).

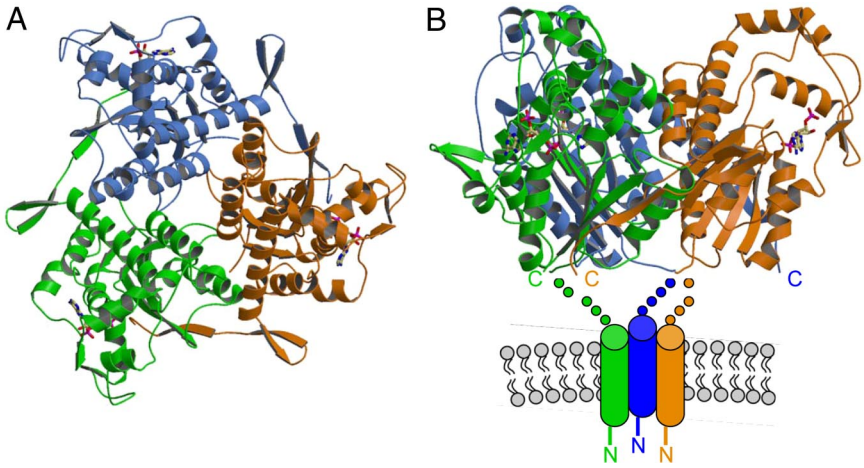

Fig. 3. Trimeric complex of the catalytic domain of 2OST. (*A*) View of the 2OST trimer looking down the 3-fold crystallographic axis (monomers are colored green, blue, and orange). The active sites containing PAP (khaki) are located on the outer surface of the trimer. (*B*) Side view of the trimer based on a ≈90° rotation of *A*, along an axis in the plane of the paper. The hypothetical positions of the transmembrane N-terminal helices and the Golgi membrane with respect to the trimer have been sketched into the figure.

from elution volumes and molecular mass of protein standards, the MBP-2OST fusion protein used for crystallization trials eluted with a calculated molecular mass of 210 kDa, consistent with that of a trimer.

A low-resolution crystallographic dataset collected from crystals of the catalytic domain of hamster 2OST, lacking a fused MBP molecule, revealed similar trimer formation. Using the structure of the chicken 2OST catalytic domain trimer as a search model, we were able to obtain phases by molecular replacement for data collected to 3.3 Å on hamster 2OST. The solution revealed 2 trimers in the asymmetric unit that used protein–protein interactions very similar to those observed in the chicken 2OST structure. These results are consistent with a study by Kobayashi *et al.* (21) that demonstrated that 2OST isolated from Chinese hamster ovary cells migrated as an oligomer by gel filtration chromatography (21).

The orientation of the monomers in the trimeric complex positions the N termini of all 3 molecules on one side of the trimer in close proximity to one another (Fig. 3*B*). This orientation is consistent with a type 2 membrane-bound protein where all 3 molecules would be anchored to the membrane via their N-terminal transmembrane domains (L12–L27) (22), although this region is not present in the crystal structure. Positioning of the trimers separates the active sites from one another, suggesting that the individual active sites may function independently (Fig. 3 *A* and *B*).

One interesting feature of this trimer is the extended C-terminal tail of 2OST, which forms an antiparallel strand (N345–Y352) arrangement with the 5th strand (V104–T110) of the central β-sheet of another molecule (Fig. 3 *A* and *B* and Fig S2*b*). This arrangement situates C-terminal residues Lys-350 and Glu-349 of 1 molecule near the substrate binding site of a second molecule, in a position to potentially stabilize the active site and participate in substrate binding (Fig. 3*B*). It has been reported that C5-epimerase forms a complex with 2OST (23). However, it is unlikely that C5-epimerase could occupy one of the trimeric sites of 2OST. In fact, the purified 2OST from Chinese hamster ovary cells, although in a trimeric form, did not contain C5-epimerase (21). The proposed in vivo complex perhaps involves a supramolecular structure between trimeric 2OST and C5-epimerase.

Mutational Analysis of 2OST. Site-directed mutagenesis studies were performed to identify the roles of specific amino acid residues in the

Table 1. Activity of 2OST WT and mutants toward polysaccharide substrates

	Sulfotransferase activity to polysaccharide substrates			Sulfotransferase activity to polysaccharide substrates	
2OST mutants*	CDSNS-heparin (-IdoA-GlcNS-)$_n$† , %	*N*-Sulfo heparosan (-GlcA-GlcNS-)$_n$†, %	2OST mutants*	CDSNS-heparin (-IdoA-GlcNS-)$_n$†, %	*N*-Sulfo heparosan (-GlcA-GlcNS-)$_n$†, %
WT	100	100	WT	100	100
R80A	1	2	R188A	100	100
N91A	51	42	R189A	0.3	100
Y94A	90	7	R190A	100	74
D95A	80	32	D208A	76	42
H106A	51	0.4	K213A	100	56
K111A	100	62	K283A	100	100
N112A	97	54	K284A	68	42
H140A	7	<0.1	R288A	5	10
H140N	52	14	H330A	100	100
H142A	<0.1	<0.1	E349A	31	77
H142N	20	3	K350A	22	7
R169A	1	0.3	Y352A	69	43
Y173A	80	50	V332STP	2	0.5
D181A	20	11	L342STP	11	3
Y183A	8	2			

*The 2OST mutants were prepared using a modified site-directed mutagenesis protocol from Stratagene. The protein expression level of the mutant proteins is comparable with the wild-type (WT) protein as determined by the intensity of the Coomassie blue-stained protein band migrated at 70 kDa on SDS/PAGE.
†The activity of 2OST was assayed by incubating the purified mutant proteins with either completely de-*O*-sulfated *N*-sulfated heparin (CDSNS-heparin) [(-IdoA-GlcNS-)$_n$] or *N*-sulfo heparosan [(-GlcA-GlcNS-)$_n$] and [^{35}S]PAPS, and the resultant ^{35}S-substrate was quantified by DEAE chromatography, where 100% activity represents the transfer of 155 pmol of sulfate per μg of protein for CDNS-heparin and 65 pmol of sulfate per μg of protein for NS-heparosan under the standard assay conditions.

catalytic function and substrate specificity of 2OST (Table 1). Amino acid residues that are involved in binding to PAP/PAPS are not the focus of this work because they have been reported in a previous publication (20). Instead, amino acid residues lining the proposed substrate-binding cleft were chosen for mutational analysis. Mutant enzymes were exposed to 2 polysaccharide substrates differing in the structure of the disaccharide repeating unit. Completely desulfated *N*-sulfated (CDSNS) heparin predominantly has the repeating disaccharide of -IdoA-GlcNS-, whereas *N*-sulfo heparosan consists of a repeating structure of -GlcA-GlcNS- (Fig. 1). Testing the activity of 2OST mutants with these 2 different substrates has permitted us to identify residues that differentiate between IdoA and GlcA acceptor substrates in the polysaccharide chain.

Residues involved in catalysis. Based on the crystal structure, 4 residues (Arg-80, His-140, His-142, and Arg-288) are located near the point of sulfo transfer and thus may be involved in catalysis and/or acceptor substrate binding. Consistent with previous alanine scanning mutagenesis experiments, our results demonstrate that these residues are important for activity (20). In the crystal structure, Arg-80 and Arg-288 are located near the proposed position of the sulfo moiety on the PAPS molecule and therefore may be involved in stabilization of the transition state and/or substrate binding (Fig. 2*B*). The R80A and R288A mutants display 10% or less activity toward polysaccharide substrates (Table 1).

Previously, it was demonstrated that the mutations H140A and H142A greatly reduced the activity of 2OST and that the double mutant H140A/H142A resulted in complete loss of activity (20). The crystal structure reveals that His-142 is in position potentially to serve as a catalytic base facilitating the deprotonation of the acceptor substrate, whereas His-140 is located more distal to the point of sulfo transfer and is in position to form a hydrogen bond with Arg-80. The H142A mutant of chicken 2OST had no detectable activity with the 2 substrates, supporting its essential role as a catalytic base, whereas 2OST H140A retains ≈7% of wild-type activity against IdoA containing substrate (CDSNS heparin). To probe the functions of His-140 and His-142 further, mutants H140N and H142N were prepared. The H140N mutant maintained 52% activity against IdoA-containing substrate and 14% against GlcA-containing substrate (*N*-sulfo heparosan). Based on the crystal structure, it appears that the asparagine substitution could still form a hydrogen bond with Arg-80. Thus, the major role of His-140 may be to position Arg-80 properly. Surprisingly, the H142N mutant maintains 20% and 3% of wild-type activity against IdoA- and GlcA-containing substrates, respectively. In contrast, the equivalent mutation of His-107 of human EST to asparagine abolished the sulfotransferase activity (19). One possible explanation for the residual activity of the H142N mutant vs. H142A is that a water molecule could occupy the location of the $N^{\varepsilon}2$ atom of the histidine side chain, stabilized by interactions with Asn-142, and substitute as a weak proton acceptor in the absence of the histidine side chain. An additional possibility is that the reaction for 2OST is more dissociative in nature, compared with EST, with greater emphasis on bond breaking and stabilization of the leaving group rather than deprotonation of the nucleophile, decreasing the need for a strong base.

Residues involved in trimer formation. Several residues along the C-terminal tail were subjected to mutational analysis, including His-330, Glu-349, Lys-350, and Tyr-352 (Fig. 2*B*). The only residue showing a loss of activity >70% was the Lys-350 mutant (Table 1). The side chain of this residue is extended into the active site of the adjacent molecule, suggesting that it could play a minor role in substrate binding. Although no single point mutation from this region caused a severe decrease in activity, introduction of a premature stop codon after either Val-332 or Leu-342 resulted in a drastic loss in activity toward both substrates. These data suggest that deletion of residues along the C terminus eliminates the intermolecular interactions between the protein backbone atoms, disrupting the adjacent active site and potentially destabilization of

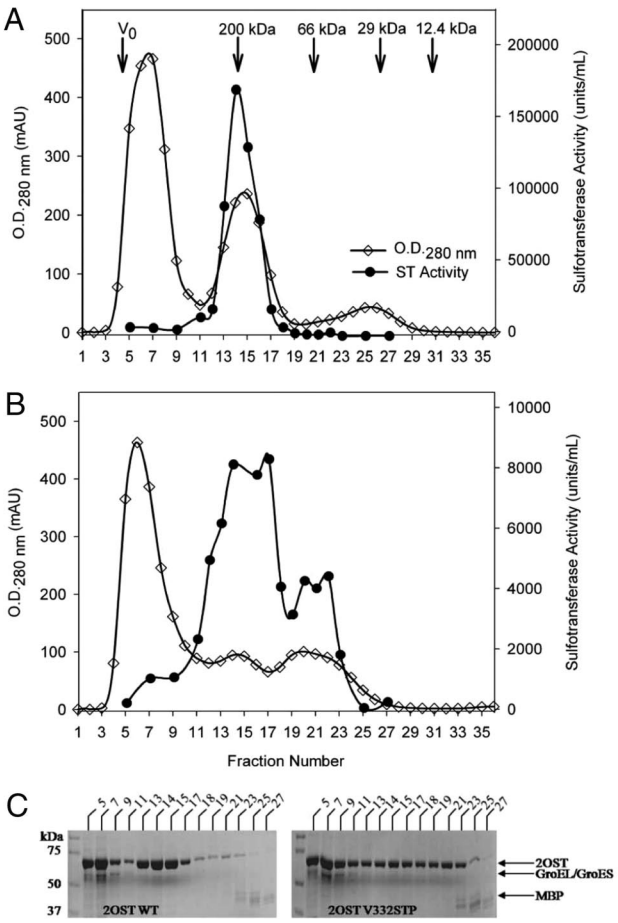

Fig. 4. Comparison of the oligomeric state of 2OST wild-type with 2OST V332STP. Approximately 7 mg of amylose–agarose column-purified wild-type or premature stop codon mutant (V332STP) 2OST was analyzed by gel permeation chromatography. The sulfotransferase activity was also determined by using CDSNS heparin as a substrate. The UV absorbance and sulfotransferase activity are plotted against fraction number for wild-type (*A*) and mutant (*B*) 2OST. The molecular mass standards resolved under the same experimental conditions are indicated by arrows. One unit is equivalent to 1 pmol of sulfate transferred to the substrate under standard reaction conditions. (*C*) Fractions that were analyzed by SDS/PAGE for both the 2OST wild-type and 2OST V332STP. The fractions analyzed are labeled with a number above each lane of the SDS gel. SDS/PAGE analysis of these fractions from gel permeation chromatography confirmed the presence of 2OST within all fractions from 5 to 23. Both 2OST WT and 2OSTV332STP are MBP fusion proteins.

the protein and/or the protein trimeric complex. To determine whether the truncation mutant disrupted the trimeric complex, the mutant protein was analyzed by gel filtration chromatography. The chromatogram for the wild-type MBP-2OST revealed 3 distinct peaks (Fig. 4*A*) representing aggregated MBP-2OST (peak 1, fractions 3–9) in the void volume, trimeric MBP-2OST (peak 2, fractions 12–16) (sample used for crystallization trials), and MBP (peak 3, fractions 23–28), respectively. Of these peaks, the only one containing sulfotransferase activity was that consistent with the trimer.

The mutant (2OST V332STP) revealed a starkly different elution profile from the wild-type MBP-2OST construct. The aggregated protein elutes with intensity and position similar to the normal wild-type MBP-2OST; however, there is a substantial decrease in the intensity of the trimeric 2OST peak and the appearance of an additional peak before MBP between fractions 19 and 23 (Fig. 4*B*). This peak corresponds to a molecular mass of ≈70 kDa and represents monomeric MBP-2OST (Fig. 4*C*). Sulfotrans-

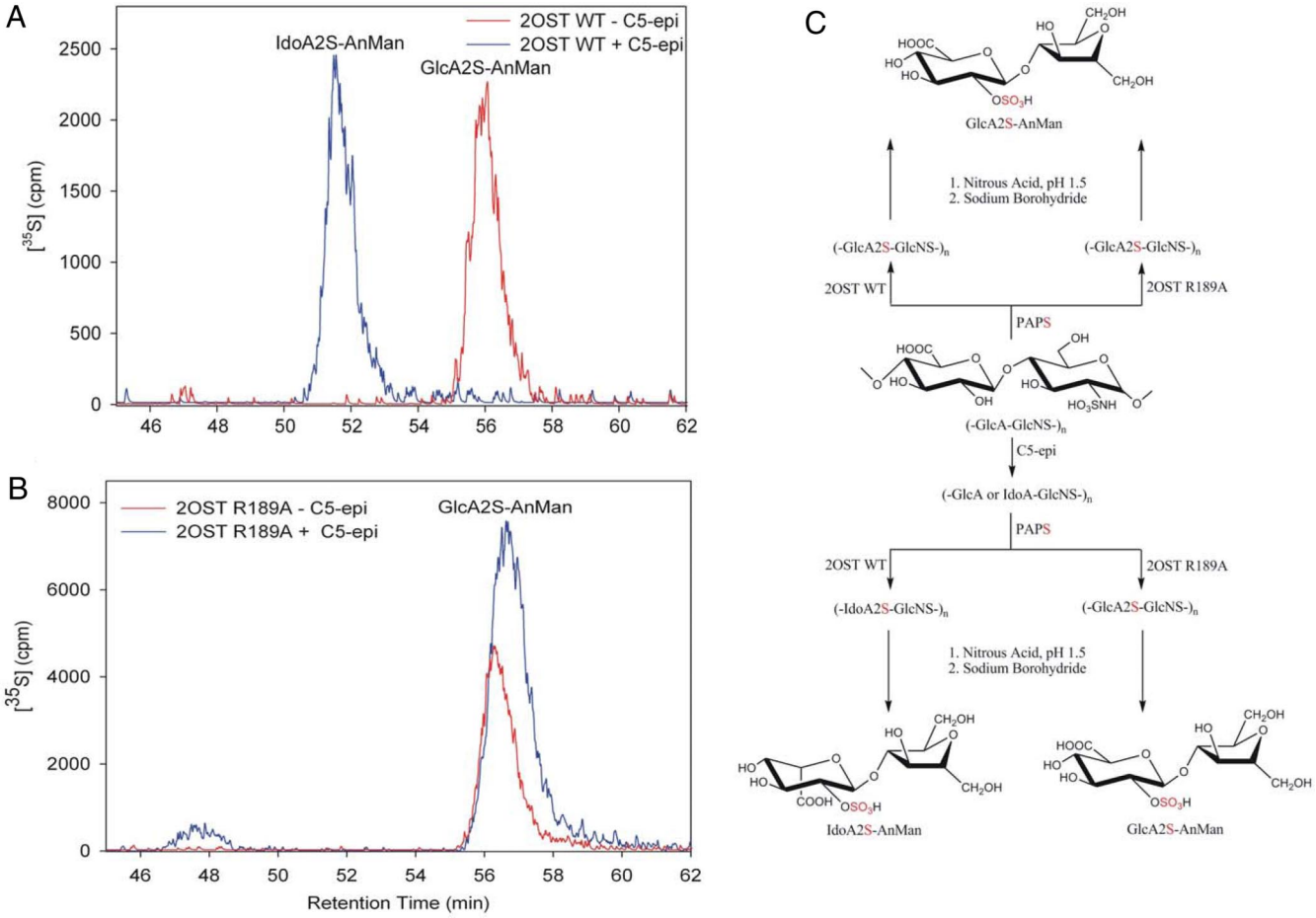

Fig. 5. Reverse-phase ion pairing-HPLC chromatograms of the disaccharide analysis of 2OST wild-type and 2OST R189A-modified polysaccharides. (*A*) Chromatogram of nitrous acid degraded 2-*O*-[^{35}S]sulfated *N*-sulfo heparosan using 2OST wild-type enzyme in the presence (blue line) or absence (red line) of C5-epimerase modification. (*B*) Chromatogram of nitrous acid degraded 2-*O*-[^{35}S]sulfated *N*-sulfo heparosan using 2OST R189A enzyme with (blue line) or without (red line) C5-epimerase modification. (*C*) Reactions involved in preparing 2OST-modified polysaccharide with or without C5-epimerase and degradation by nitrous acid. AnMan represents 2,5-anhydromannitol.

ferase activity was detected predominantly within the trimeric peak and the monomeric 2OST peak, but at a lower level. The activity for the mutant trimer is substantially lower than for the wild-type protein, suggesting that the extended C-terminal tail influences trimer formation and is crucial for the activity.

Residues involved in substrate binding and specificity. Residues Tyr-94, His-106, Arg-169, Asp-181, Tyr-183, Arg-189, and Lys-350 are located in the substrate-binding cleft but distal to the catalytic core, implicating them in acceptor substrate binding (Fig. 2*B*). Mutations of these residues showed a noticeable effect on enzymatic activity (Table 1). Of these mutations, Y183A and R169A resulted in a substantial loss of activity toward both CDSNS heparin and *N*-sulfo heparosan. Tyr-183 is located toward the predicted nonreducing end of the binding pocket based on comparisons with the crystal structure of 3OST-3 (16). The side chain of Tyr-183 extends into the solvent, suggesting that it does not play a structural role, but rather may play a significant role in substrate binding. The side chain of Arg-169 is located on the reducing end of the binding pocket. Most of this side chain is buried with the guanidinium moiety on the surface of the binding pocket. Arg-169 is conserved in NST and the 3OST isoforms as well. In the 3OST-3 structure with a tetrasaccharide bound, this residue (Arg-248) does not directly interact with the substrate. The main role of Arg-169 may be to stabilize the PSB loop because the guanidinium moiety is within hydrogen-bonding distance to the backbone carbonyl oxygen of Pro-82 in this loop.

Also lining the pocket, but largely buried, is the side chain of Asp-181. This residue is in position to form hydrogen bonds with Arg-189 (2.8 Å) and/or Arg-288 (2.9 Å). Such interactions suggest that Asp-181 may help position Arg-288 and Arg-189 for substrate binding and/or catalysis. As with Arg-288, replacing Asp-181 with alanine resulted in substantial loss of the sulfotransferase activity toward both CDSNS heparin and *N*-sulfo heparosan (Table 1).

Although wild-type 2OST sulfates both GlcA and IdoA in HS substrates, the mutants H106A, Y94A, and R189A showed preference for one substrate over the other. Both Y94A and H106A mutants sulfated CDSNS heparin, the IdoA-containing substrate, but lost activity with *N*-sulfo heparosan, the GlcA-containing substrate, displaying 10-fold and 100-fold differences in percentage activity, respectively. These results suggest that Tyr-94 and His-106 allow recognition of the GlcA-containing substrate. Thus, replacing either residue with alanine resulted in a substantial preference for the polysaccharide substrate containing IdoA.

Most impressive was the selective preference of R189A for the GlcA-containing substrate over IdoA. The mutant R189A retained full activity with the GlcA-containing substrate but showed virtually no activity with the IdoA-containing substrate. This observation suggested that residue Arg-189 is likely important for recognition of IdoA-containing substrate, and replacing the arginine residue with an alanine does not allow the enzyme to exhibit substrate flexibility.

To examine the role of Arg-189 in substrate recognition further, we reconstituted the biosynthesis of IdoA2S in vitro. The *N*-sulfo heparosan, similar to the unepimerized polysaccharide substrate

(precursor of HS) (24), was incubated with wild-type 2OST and 2OST R189A in the presence or absence of C5-epimerase, where C5-epimerase converts part of the GlcA to IdoA (as illustrated in Fig. 5*C*). We characterized the resultant polysaccharides by disaccharide analysis of the products (Fig. 5 *A* and *B*). As expected, the product modified by wild-type 2OST predominantly contains GlcA2S in the absence of C5-epimerase (Fig. 5*A*, red line) or IdoA2S in the presence of C5-epimerase (Fig. 5*A*, blue line) (8). In contrast, the products modified by 2OST R189A contain only GlcA2S regardless of the absence or presence of C5-epimerase (Fig. 5*B*), suggesting that the mutant protein does not recognize IdoA as an acceptor unit in the polysaccharide substrate.

The roles of Tyr-94, His-106, and Arg-189 in directing the substrate specificity are currently unknown because the structure of a ternary complex containing an appropriate oligosaccharide substrate is not available. However, based on gel filtration chromatography analysis, the Y94A, H106A, and R189A mutants migrate as a trimer similar to wild-type enzyme (Fig. S4). Thus, these mutations do not lead to large-scale structural changes that disrupt the active trimer, supporting a local effect on substrate recognition as the cause for differences in substrate specificity.

Conclusions

Unlike previously characterized HS sulfotransferases, 2OST is present in a trimeric form that appears to be essential for enzymatic activity. Based on the structure, we have identified several amino acid residues that are involved in polysaccharide substrate binding. Mutation of these amino acid residues alters the spectrum of substrate specificity of 2OST. A mutant protein with altered substrate specificity offers the possibility to synthesize selected HS products that cannot be achieved by wild-type 2OST. For example, 2OST R189A selectively sulfates the GlcA unit in the presence of IdoA, producing GlcA2S- but not IdoA2S-containing HS polysaccharides. One potential application of this mutant or perhaps another similarly engineered sulfotransferase would be to synthesize therapeutic heparin with reduced side effects. Heparin-induced thrombocytopenia is a major side effect of therapeutic heparin, caused in part by heparin-binding platelet factor 4. It is known that the IdoA2S unit, the product of wild-type 2OST modification, is involved in heparin binding to platelet factor 4 (25). Using the 2OST R189A mutant to synthesize heparin, we will be able to replace the IdoA2S unit by GlcA2S, hopefully reducing the side effect. Another potential application of engineered 2OST will be to allow us to produce HS and heparin with highly enriched GlcA2S units for probing the biological function in glycomics studies (26).

Methods

Cloning, Expression, and Purification of Chicken 2OST. The MBP-2OST fusion protein was created by using a modified pMAL-c2x vector (New England Biolabs). The amino acid sequence of MBP was truncated at Asn-367 and contained the mutation E359A (27). The linker region encodes 3 alanine residues (A368–A370) and contained an NotI site for cloning. The catalytic domain of chicken 2OST (D69–N356) was cloned into the vector by using the NotI and BamHI sites. The MBP-2OST was expressed in Origami B(DE3) cells (Novagen). Cells were grown on a shaker at 37 °C in LB medium and induced with isopropyl-β-D-thiogalactopyranoside. Cells were allowed to shake overnight at 18 °C. Cells were pelleted, resuspened in 25 mM Tris (pH 7.5), 500 mM NaCl, and 1 mM DTT, and then lysed by sonication. MBP-2OST was bound to amylose resin (New England Biolabs), eluted with maltose, then loaded onto a HR16/60 Superdex 200 (Amersham) column preequilibrated in the sonication buffer containing 40 mM maltose. The purified protein was then dialyzed overnight against 25 mM Tris (pH 7.5), 75 mM NaCl, 5 mM maltose, and 1 mM DTT. PAP was added to 1 mM. The sample was concentrated to 19 mg/mL, followed by the addition of more PAP for a final concentration of 4 mM.

In addition, a C-terminal $(His)_6$-tagged 2OST expression construct was prepared by cloning the catalytic domain of hamster 2OST (R63–N356) into the pET-21b vector (Novagen) by using NdeI and HindIII sites. The procedures for the expression and purification of hamster 2OST are described in ref. 20.

Determination of the Substrate Specificity of 2OST R189A. Because *N*-sulfo heparosan contains no IdoA unit, it was incubated with recombinant C5-epimerase to synthesize the IdoA unit followed by 2OST modification. To this end, *N*-sulfo heparosan (9 μg) was incubated with C5-epimerase (7 μg) in a buffer containing 50 mM Mes (pH 7.0), 1% Triton X-100, and 2 mM $CaCl_2$ for 1 h at 37 °C. To this reaction, 10 μg of 2OST R189A or 2OST WT and 1.5×10^6 cpm of [^{35}S]PAPS were added. The 2-*O*-[^{35}S]sulfated polysaccharide was purified by a DEAE column. The position of the 2-*O*-[^{35}S]sulfo group in the polysaccharide was determined by disaccharide analysis. Briefly, the resultant ^{35}S-labeled polysaccharide was degraded to disaccharides with nitrous acid at pH 1.5 followed by reduction with sodium borohydride. The disaccharide products were then analyzed by reverse-phase ion-pairing HPLC. The identities of the resultant ^{35}S-labeled disaccharides were determined by coelution with appropriate disaccharide standards (28).

Additional experimental procedures are presented under *SI Text*.

ACKNOWLEDGMENTS. We thank H. Kohn, T. Hall, and A. Moon for reviewing the manuscript. This work was supported in part by National Institutes of Health Grant AI50050 (to J.L.) and by the Intramural Research Program of the National Institute of Environmental Health Sciences, National Institutes of Health (L.C.P.). D.X. is a recipient of a predoctoral fellowship from the American Heart Association, Mid-Atlantic Affiliate.

1. Bishop J, Schuksz M, Esko JD (2007) Heparan sulphate proteoglycans fine-tune mammalian physiology. *Nature* 446:1030–1037.
2. Gama C, *et al*. (2006) Sulfation patterns of glycosaminoglycans encode molecular recognition and activity. *Nat Chem Biol* 2:467–473.
3. Fuster MM, Esko JD (2005) The sweet and sour of cancer: Glycans as novel therapeutic targets. *Nat Rev Cancer* 5:526–542.
4. Chen J, Jones CL, Liu J (2007) Using an enzymatic combinatorial approach to identify anticoagulant heparan sulfate structures. *Chem Biol* 14:986–993.
5. Muñoz E, *et al*. (2006) Enzymatic synthesis of heparin-related polysaccharides on sensor chips: Rapid screening of heparin–protein interactions. *Biochem Biophys Res Commun* 339:597–602.
6. Chen J, *et al*. (2005) Enzymatically redesigning of biologically active heparan sulfate. *J Biol Chem* 280:42817–42825.
7. Copeland RJ, *et al*. (2008) Using a 3-*O*-sulfated heparin octasaccharide to inhibit the entry of herpes simplex virus 1. *Biochemistry* 47:5774–5783.
8. Rong J, Habuchi H, Kimata K, Lindahl U, Kusche-Gullberg M (2001) Substrate specificity of the heparan sulfate hexuronic acid 2-*O*-sulfotransferase. *Biochemistry* 40:5548–5555.
9. Kreuger J, Salmivirta M, Sturiale L, Gimenez-Gallego G, Lindahl U (2001) Sequence analysis of heparan sulfate epitopes with graded affinities for fibroblast growth factors 1 and 2. *J Biol Chem* 276:30744–30752.
10. Bullock SL, Fletcher JM, Beddington RS, Wilson VA (1998) Renal agenesis in mice homozygous for a gene trap mutation in the gene encoding heparan sulfate 2-sulfotransferase. *Genes Dev* 12:1894–1906.
11. Bulow HE, Hobert O (2004) Differential sulfations and epimerization define heparan sulfate specificity in nervous system development. *Neuron* 41:723–736.
12. Kinnunen T, *et al*. (2005) Heparan 2-*O*-sulfotransferase, hst-2, is essential for normal cell migration in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* 102:1507–1512.
13. Kamimura K, *et al*. (2006) Specific and flexible roles of heparan sulfate modifications in *Drosophila* FGF signaling. *J Cell Biol* 174:773–778.
14. Kakuta Y, Sueyoshi T, Negishi M, Pedersen LC (1999) Crystal structure of the sulfotransferase domain of human heparan sulfate *N*-deacetylase/*N*-sulfotransferase 1. *J Biol Chem* 274:10673–10676.
15. Edavettal SC, *et al*. (2004) Crystal structure and mutational analysis of heparan sulfate 3-*O*-sulfotransferase isoform 1. *J Biol Chem* 279:25789–25797.
16. Moon A, *et al*. (2004) Structural analysis of the sulfotransferase (3OST-3) involved in the biosynthesis of an entry receptor of herpes simplex virus 1. *J Biol Chem* 279:45185–45193.
17. Xu D, Moon A, Song D, Pedersen LC, Liu J (2008) Engineering sulfotransferases to modify heparan sulfate. *Nat Chem Biol* 4:200–202.
18. Negishi M, *et al*. (2001) Structures and functions of sulfotransferases. *Arch Biochem Biophys* 390:149–157.
19. Pedersen LC, Petrotchenko E, Shevtsov S, Negishi M (2002) Crystal structure of the human estrogen sulfotransferase–PAPS complex: Evidence for catalytic role of Ser^{137} in the sulfuryl transfer reaction. *J Biol Chem* 277:17928–17932.
20. Xu D, Song D, Pedersen L, Liu J (2007) Mutational study of heparan sulfate and chondroitin sulfate 2-*O*-sulfotransferases. *J Biol Chem* 282:8356–8367.
21. Kobayashi M, Habuchi H, Habuchi O, Saito M, Kimata K (1996) Purification and characterization of heparan sulfate 2-sulfotransferase from cultured Chinese hamster ovary cells. *J Biol Chem* 271:7645–7653.
22. Kobayashi M, Habuchi H, Yoneda M, Habuchi O, Kimata K (1997) Molecular cloning and expression of Chinese hamster ovary cell heparan sulfate 2-sulfotransferase. *J Biol Chem* 272:13980–13985.
23. Pinhal MA, *et al*. (2001) Enzyme interactions in heparan sulfate biosynthesis: Uronosyl 5-epimerase and 2-*O*-sulfotransferase interact in vivo. *Proc Natl Acad Sci USA* 98:12984–12989.
24. Esko JD, Selleck SB (2002) Order out of chaos: Assembly of ligand-binding sites in heparan sulfate. *Annu Rev Biochem* 71:435–471.
25. Mikhailov D, Young HC, Linhardt RJ, Mayo KH (1999) Heparin dodecasaccharide binding to platelet factor-4 and growth-related protein-α. *J Biol Chem* 274:25317–25329.
26. Shipp E, Hsieh-Wilson LC (2007) Profiling the sulfation specificities of glycosaminoglycan interactions with growth factors and chemotactic proteins using microarrays. *Chem Biol* 14:195–208.
27. Center RJ, *et al*. (1998) Crystallization of a trimeric human T cell leukemia virus type 1 gp21 ectodomain fragment as a chimera with maltose-binding protein. *Protein Sci* 7:1612–1619.
28. Xia G, *et al*. (2002) Heparan sulfate 3-*O*-sulfotransferase isoform 5 generates both an antithrombin-binding site and an entry receptor for herpes simplex virus, type 1. *J Biol Chem* 277:37912–37919.