Supporting Information

Moon et al. 10.1073/pnas.1117923109

SI Materials and Methods

Expression and Purification of 3-OST-1. The sequence encoding residues Gly49–His311 of HS 3-OST-1 from *Mus musculus* was ligated into the pGEX4T3 expression vector (GE Healthcare) using the EcoRI and NotI restriction sites. The resulting vector was transformed into the BL21 (DE3)-CodonPlusRIL cell line (Stratagene) for expression. The cells were grown in LB medium supplemented with 100 μg/mL ampicillin and 25 μg/mL chloramphenicol at 37 °C, to a cell density where A_{600} = 0.8. The temperature was decreased to 18 °C for 30 min, and protein expression was induced by addition of 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Expression continued overnight at 18 °C. Cells were harvested by centrifugation and lysed by sonication in 1× PBS (pH 7.2), 500 mM NaCl. The lysate was clarified by high-speed centrifugation, and the soluble fraction was bound in-batch to glutathione Sepharose 4B resin (GE Healthcare) for 1 h at 4 °C. Unbound protein was removed by repetitive washing with sonication buffer. 3-OST-1 protein was cleaved from the resin by addition of 500 units of thrombin (Sigma). Cleavage took place overnight at 4 °C with gentle rocking. Cleaved, soluble protein was pooled, concentrated, and purified by size-exclusion chromatography on a Superdex 200 26/60 column (GE Healthcare) equilibrated in 25 mM Tris, pH 7.5, 650 mM NaCl. Fractions containing purified 3-OST-1 were pooled and dialyzed overnight at 4 °C against 25 mM Tris (pH 7.5), 150 mM NaCl. PAP (Sigma) was added to a final concentration of 1 mM, and the protein was concentrated to ∼18–19 mg/mL.

Structure Solution and Refinement. A low-resolution data set for 3-OST-1 in complex with the heptasaccharide substrate was collected on an in-house rotating anode source. Coordinates from the crystal structure of 3-OST-1 (PDB ID code 1VKJ) (1) were used as a starting model for molecular replacement, using MOLREP in the CCP4 suite (2, 3). A preliminary model for the heptasaccharide substrate was built into electron density near the active site of 3-OST-1 using the tetrasaccharide bound to 3-OST-3A (PDB ID code 1T8U) (4) and antithrombin (PDB ID code 3KCG) (5) as structural templates. This model was refined further against the 1.85-Å data reported here while maintaining the same R_{free} set of reflections, using the same techniques. All data were processed using HKL2000 (6), and refined by iterative cycles of model building in COOT (7) and refinement in PHENIX (8). Data collection and refinement statistics are listed in Table S1. Ramachandran statistics were obtaining using MolProbity (9). All structural figures were generated using molecule A from the Protein Data Bank file and MolScript (10), Raster3D (11), and PyMOL (12).

Preparation of 3-OST-1 and 3-OST-3 Mutants. Mutants of murine 3-OST-1 (Gly48–His311) and human 3-OST-3A (Gly139–Gly406) were generated by site-directed mutagenesis in the pGEX4T3 expression vector. The resulting vectors were transformed in the BL21 (DE3)-CodonPlusRIL cell line, and all mutant proteins were expressed and purified as for the wild-type 3-OST-1.

Chemoenzymatic Synthesis of Heptasaccharide Substrate. Synthesis of heptasaccharide substrate included both the backbone synthesis and installation of sulfo groups and IdoA unit. The preparation of backbone heptasaccharide involves two reactions of elongation as described previously (13), including elongation from disaccharide glucuronic acid (GlcA)-anhydromannose (AnMan) to tetrasaccharide GlcA-trifluoroacetylglucosamine (GlcNTFA)-GlcA-AnMan, and elongation from tetrasaccharide to heptasaccharide with a structure of GlcNAc-GlcA-GclcNTFA-GlcA-GclcNTFA-GlcA-AnMan. For elongation reaction from the disaccharide to tetrasaccharide, 6 mg GlcA-AnMan was incubated with 18 μmol UDP-GlcNTFA and 2 mg of KfiA (*N*-acetyl glucosaminyl transferase of *Escherichia coli* K5 strain) in 40 mL buffer containing 25 mM Tris·HCl (pH 7.2) and 10 mM $MnCl_2$ at room temperature overnight. Upon the complete consumption of UDP-GlcNTFA, 2 mg of heparosan synthase 2 from *Pasteurella multocida*, and 27 μmol UDP-GlcUA were added into the reaction mixture and allowed to incubate overnight at room temperature. The product was purified by a Bio-Gel P-2 column (0.75 × 200 cm) that was equilibrated with 0.1 M ammonium bicarbonate at a flow rate of 6 mL/h. The product fraction was located by electrospray ionization MS (ESI-MS) analysis. For the addition of monosaccharide to the tetrasaccharide, the condition was identical to the above. After three additional cycles, the tetrasaccharide was converted to a heptasaccharide.

The preparation of UDP-GlcNTFA was started from glucosamine (Sigma), which was first converted to GlcNTFA by reacting with *S*-ethyl trifluorothioacetate (Sigma-Aldrich) following the protocol described previously (13). The resultant GlcNTFA was converted to GlcNTFA-1-phosphate using *N*-acetylhexosamine 1-kinase (14). The plasmid expressing *N*-acetylhexosamine 1-kinase was a generous gift from Peng Wang (Ohio State University, Columbus, OH), and the expression of the enzyme was carried out in *E. coli* as reported (14). The UDP-GlcNTFA synthesis was completed by transforming GlcNTFA-1-phosphate using GlmU as described (13).

The conversion of heptasaccharide backbone to heptasaccharide substrate for 3-OST involved four steps, including detrifluoroacetylation/*N*-sulfation, C_5-epimerization/2-*O*-sulfation, 6-*O*-sulfation as described previously (15). The enzymes used for the synthesis included *N*-sulfotransferase, C_5-epimerase, 2-*O*-sulfotransferase, 6-*O*-sulfotransferase 1, and 6-*O*-sulfotransferase 3, which were all expressed in *E. coli*. For each sulfation reaction, the heptasaccharide was incubated with reaction mixture containing 500 μM PAPS, 50 mM Mes, pH 7.0, and 0.03 mg/mL enzyme at 37 °C overnight. For C_5-epimerization/2-*O*-sulfation, additional 2 mM $CaCl_2$ was added in the reaction mixture. The product was purified by a Bio-Gel P-2 column (0.75 × 200 cm) that was equilibrated with 0.1 M ammonium bicarbonate at a flow rate of 6 mL/h. The fraction of product was located by ESI-MS analysis. After 6-*O*-sulfation, the product was purified by a DEAE column. The product showed molecular mass of 1,698 Da, which is identical to the calculated value (Fig. S7). A total of 20 mg of the heptasaccharide was prepared for this study.

Enzymatic Activity Assays of 3-OST-1 and 3-OST-3 Mutants. Activity measurement for the mutants using the octasaccharide substrate was determined by incubating 0.1 μg of mutant or wild type 3-OST with 0.5 μg of octasaccharide and 2×10^6 cpm of [^{35}S]PAPS (∼5 μM) in 60 μL of buffer containing 50 mM Mes, pH 7.0, 10 mM $MnCl_2$, and 5 mM $MgCl_2$ at 37 °C overnight. Reactions were quenched by the addition of 6 M urea and 100 mM EDTA, then subjected to a 200-μL DEAE-Sepharose chromatography to purify the ^{35}S-labeled oligosaccharide. The quantity of [^{35}S]sulfated oligosaccharide was determined by liquid scintillation counting. The negative control contained all of the components with the exception of enzyme. For the activity determination of 3-OST mutants using the heptasaccharide substrate, the conditions were essentially the same, the amount of enzyme was 0.05 μg,

the amount of heptasaccharide substrate was 0.3 μg, and reaction time was 30 min.

Kinetic Characterization of 3-OST-1 and 3-OST-3 Mutants. Reactions were carried out at 37 °C and included 50 mM Mes, pH 7.0, 10 mM $MnCl_2$, 5 mM $MgCl_2$, 3-OST wild type or mutants, oligosaccharide substrate, and PAPS (3 nmol unlabeled PAPS mixed with [^{35}S]PAPS to achieve a final specific activity of [^{35}S]sulfate at ~200 cpm/pmol) in total 50-μL reaction volume. Typically for 3-OST-1 wild type and mutants, the amount of protein varied from 0.2 to 3 μg, depending on the activity, the heptasaccharide substrate varied from 0.1 to 3.2 μg, and the reaction time was 5 min. For 3-OST-3 wild type and mutants, the amount of protein varied from 1 to 6 μg, the octasaccharide substrate varied from 0.2 to 6.4 μg, and the reaction time was 15 min. Reactions were quenched by the addition of 6 M urea and 100 mM EDTA, and then subjected to a 200-μL DEAE-Sepharose chromatography to purify the [^{35}S]oligosaccharide. The quantity of [^{35}S]oligosaccharide was determined by liquid scintillation counting. Michaelis–Menten kinetic parameters were determined from the corresponding initial velocity vs. the concentration of substrate curves fitted by a nonlinear least-squares analyses to the equation $V = k_{cat}[E][S]/(K_M + [S])$, where [S] equals the oligosaccharide concentration and [E] equals the enzyme concentration.

DEAE-HPLC Analysis of 3-*O*-Sulfated Oligosaccharides. DEAE-HPLC (TosoHaas) was used to analyze 3-*O*-sulfated oligosaccharides. The DEAE-HPLC column was eluted with a linear gradient of NaCl in 20 mM Tris·HCl buffer (pH 7.5) from 0.3 to 1 M for 40 min.

1. Edavettal SC, et al. (2004) Crystal structure and mutational analysis of heparan sulfate 3-*O*-sulfotransferase isoform 1. *J Biol Chem* 279:25789–25797.
2. Collaborative Computation Project, Number 4 (1994) The CCP4 Suite: Program for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
3. Vagin A, Teplyakov A (1997) MOLREP: An automated program for molecular replacement. *J Appl Cryst* 30:1022–1025.
4. Moon A, et al. (2004) Structural analysis of the sulfotransferase (3-OST-3) involved in the biosynthesis of an entry receptor of herpes simplex virus 1. *J Biol Chem* 279: 45185–45193.
5. Johnson DJ, Langdown J, Huntington JA (2010) Molecular basis of factor IXa recognition by heparin-activated antithrombin revealed by a 1.7-A structure of the ternary complex. *Proc Natl Acad Sci USA* 107:645–650.
6. Otwinowski Z, Minor V (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
7. Emsley P, Cowtan K (2004) COOT: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
8. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221.
9. Lovell SC, et al. (2003) Structure validation by Calpha geometry: Phi, psi and Cbeta deviation. *Proteins* 50:437–450.
10. Kraulis PJ (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Cryst* 24:946–950.
11. Merritt EA, Bacon DJ (1997) Raster3D: Photorealistic molecular graphics. *Methods Enzymol* 277:505–524.
12. DeLano WL (2002) The PyMOL Molecular Graphics System (DeLano Scientific, San Carlos, CA).
13. Liu R, et al. (2010) Chemoenzymatic design of heparan sulfate oligosaccharides. *J Biol Chem* 285:34240–34249.
14. Zhao G, Guan W, Cai L, Wang PG (2010) Enzymatic route to preparative-scale synthesis of UDP-GlcNAc/GalNAc, their analogues and GDP-fucose. *Nat Protoc* 5:636–646.
15. Xu Y, et al. (2011) Chemoenzymatic synthesis of homogeneous ultralow molecular weight heparins. *Science* 334:498–501.

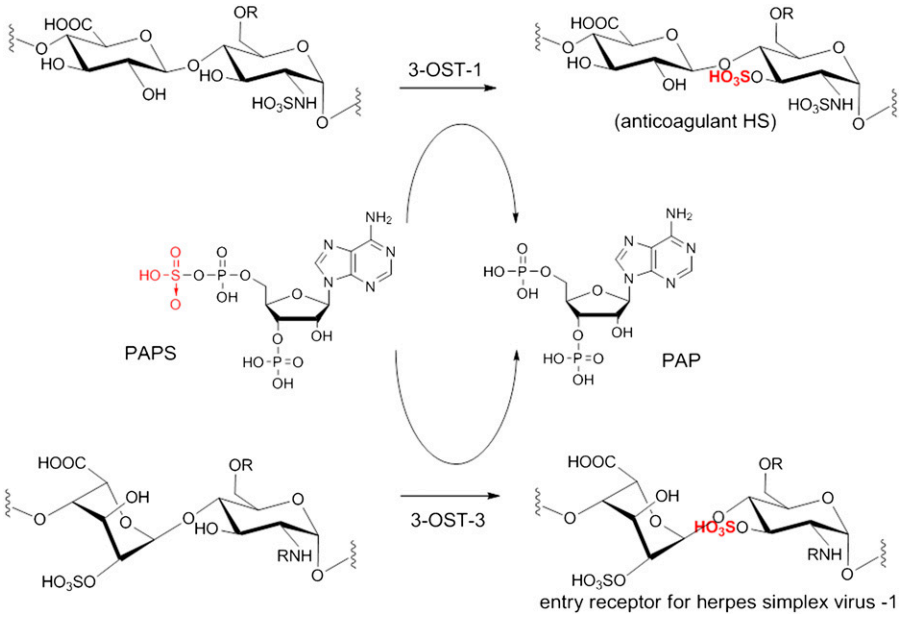

Fig. S1. Reactions catalyzed by 3-OST-1 and 3-OST-3. R = –H or –SO_3H.

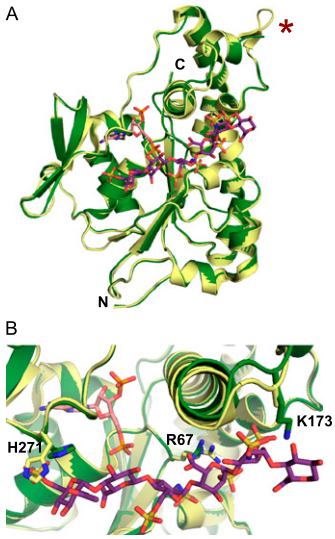

Fig. S2. Structural effects of heptasaccharide substrate binding to 3-OST-1. (*A*) Superposition of the 3-OST-1/PAP binary structure (PDB ID code 1VKJ, wheat) with the 3-OST-1/PAP/heptasaccharide ternary complex (dark green). (*B*) Superposition of 3-OST-1 side chains before (wheat) and after (dark green) heptasaccharide substrate binding. The PAP cofactor is shown in pink, and the heptasaccharide substrate is shown in purple.

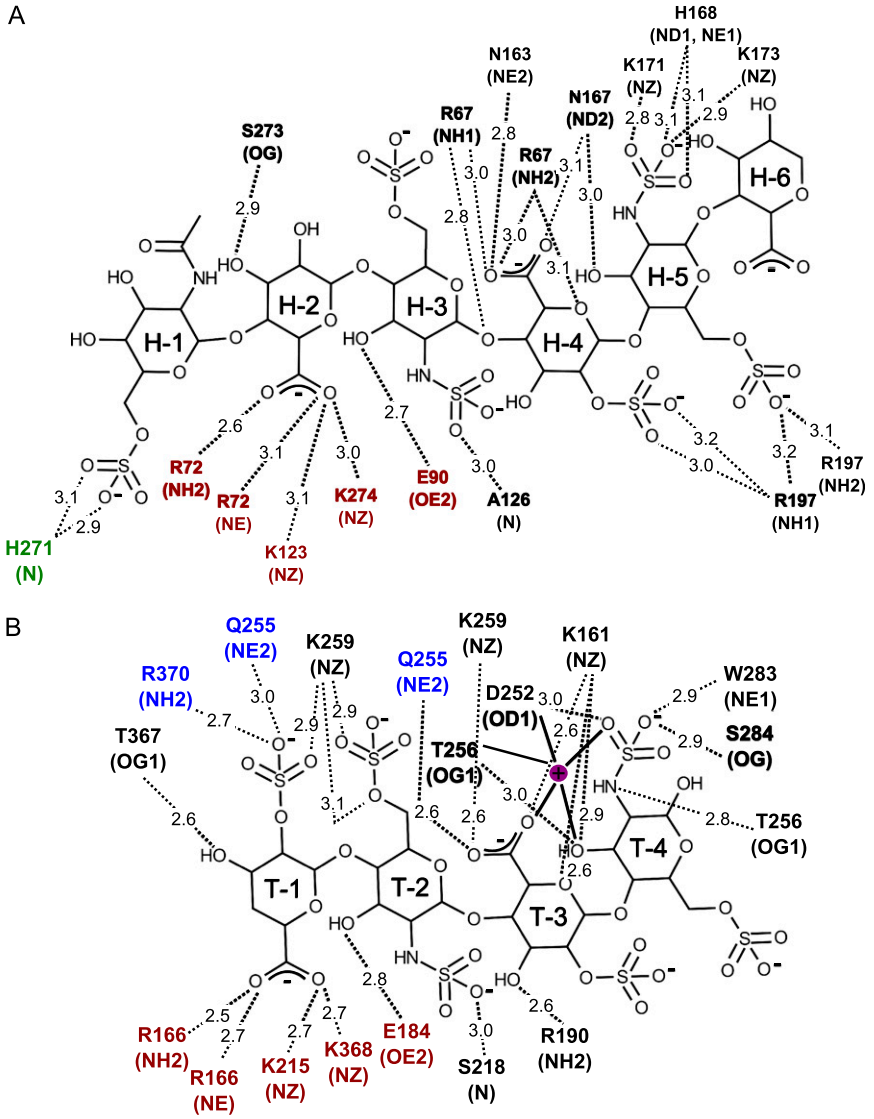

Fig. S3. Comparative analysis of substrate interactions involving the 3-*O*-sulfotransferases. (*A*) Putative hydrogen bonding interactions between 3-OST-1 and its heptasaccharide substrate. (*B*) Putative hydrogen bonding interactions between 3-OST-3 and its tetrasaccharide substrate. Conserved residues and their corresponding interactions are highlighted in red. Residues that are structurally, but not functionally, conserved between the two isoforms are highlighted in blue. Gate residue His271 in 3-OST-1 is highlighted in green. Distances are shown for each potential interaction and are measured in angstroms. The information in this figure was obtained using LIGPLOT (1).

1. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* 8:127–134.

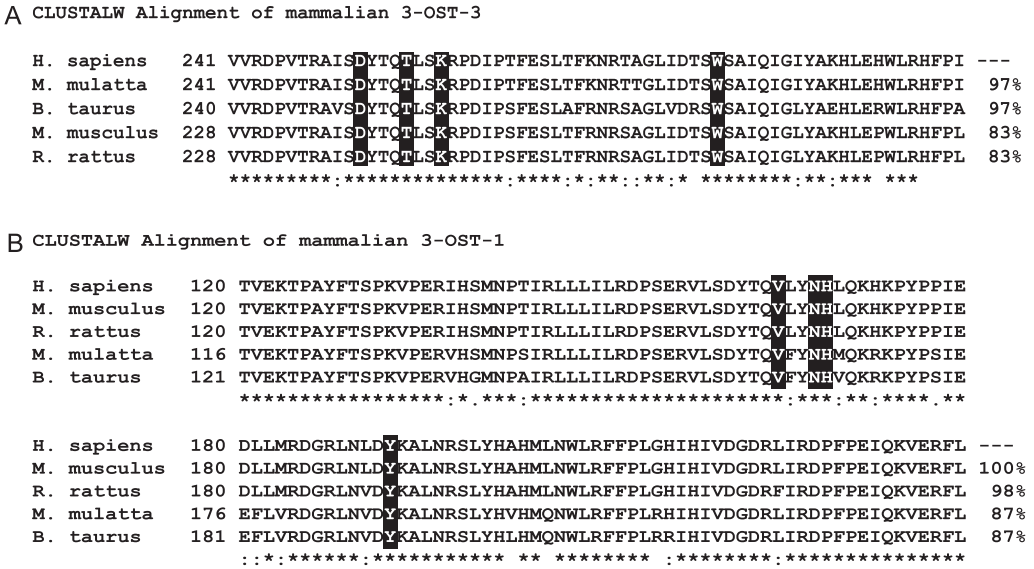

Fig. S4. (*A*) ClustalW sequence alignment (1) of mammalian 3-OST-3 orthologs. Key residues are boxed in black. (*B*) ClustalW sequence alignment (1) of mammalian 3-OST-1 orthologs. Key residues are boxed in black.

1. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680.

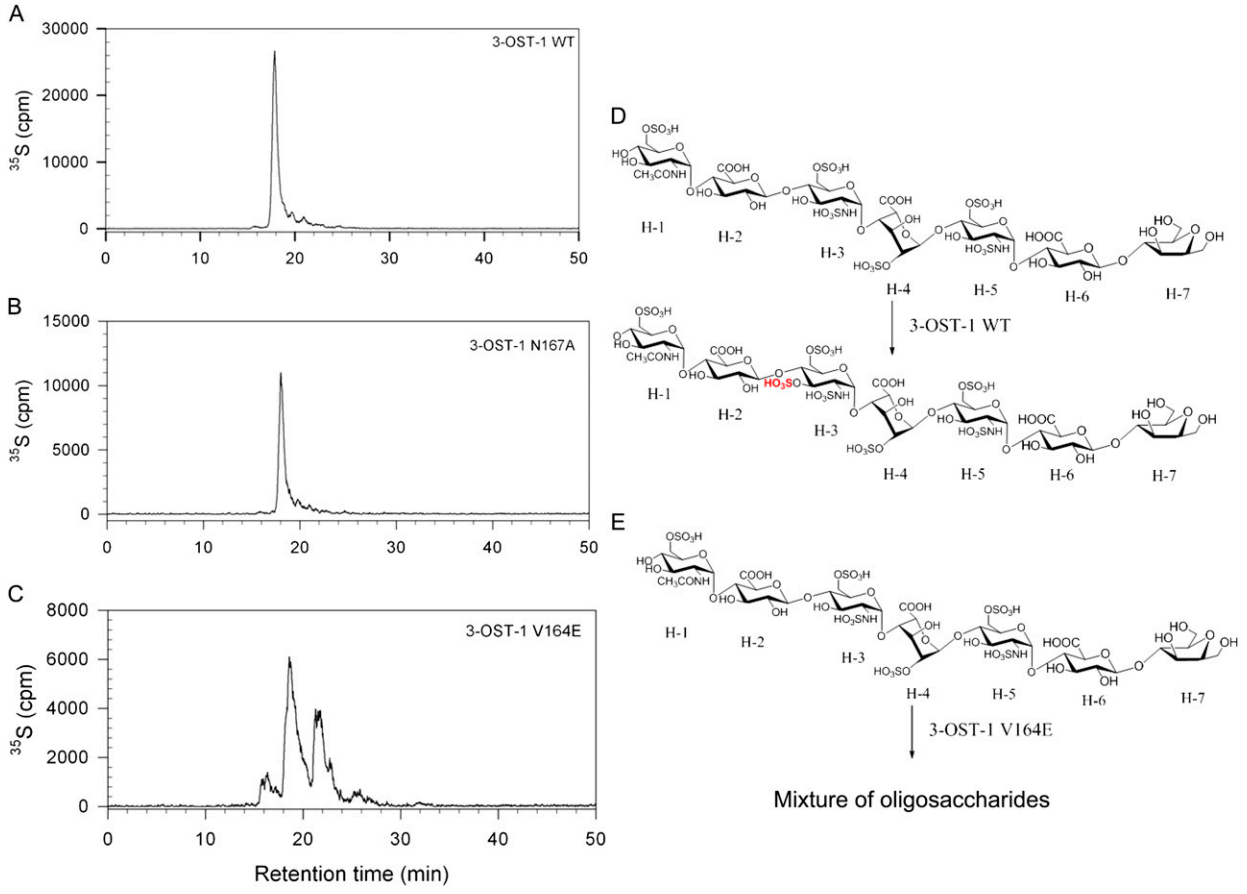

Fig. S5. Analysis of 3-*O*-sulfation products by wild-type and mutant (m)3-OST-1. ^{35}S-sulfated products from the 3-OST-1 sulfation reaction were isolated using a polyamine-based anion exchange column. (*A*) Reaction products from wild-type 3-OST-1. (*B*) Reaction products from the m3-OST-1 N167A mutant. (*C*) Reaction products from the 3-OST-1 V164E mutant. (*D*) Schematic representation of 3-*O*-sulfation by wild-type m3-OST-1. The site of the expected 3-*O*-sulfo group is shown in red. (*E*) Schematic representation of 3-*O*-sulfation by m3-OST-1 V164E mutant.

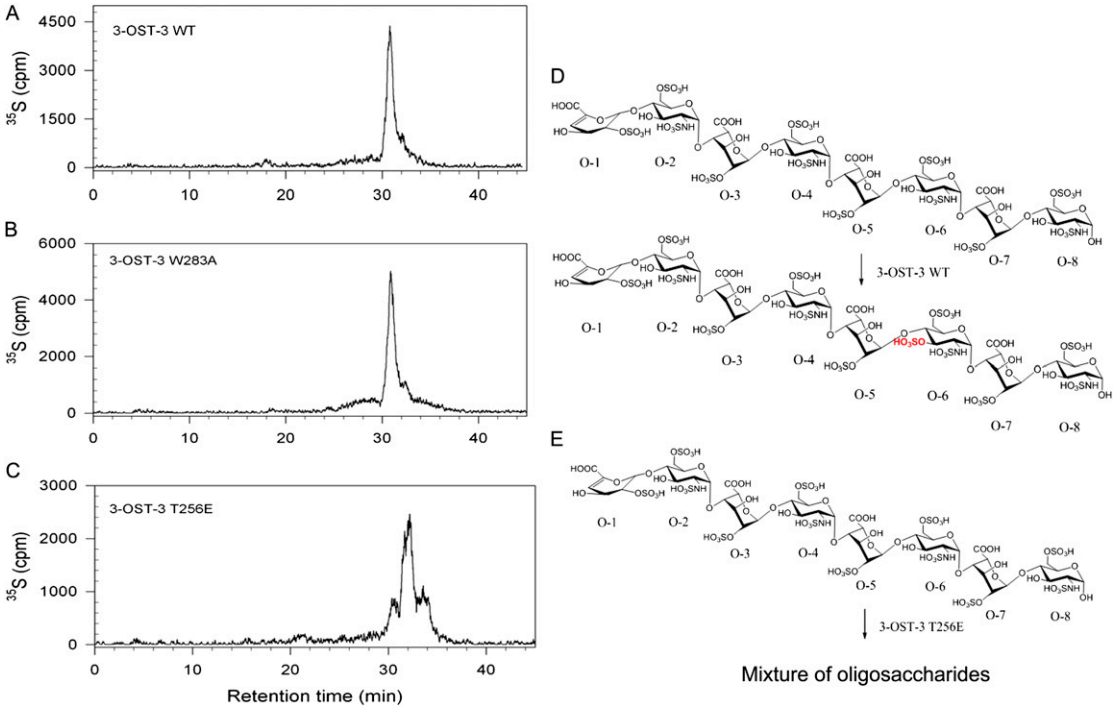

Fig. S6. Analysis of 3-*O*-sulfation products by wild-type and mutant 3-OST-3. ^{35}S-sulfated products from the 3-OST-3 sulfation reaction were isolated using a polyamine-based anion exchange column. (*A*) Reaction products from wild-type 3-OST-3. (*B*) Reaction products from the 3-OST-3 W283A mutant. (*C*) Reaction products from the 3-OST-3 T256E mutant. (*D*) Schematic representation of 3-*O*-sulfation by wild-type 3-OST-3. The site of the expected 3-*O*-sulfo group is shown in red. (*E*) Schematic representation of 3-*O*-sulfation by 3-OST-3 T256E mutant.

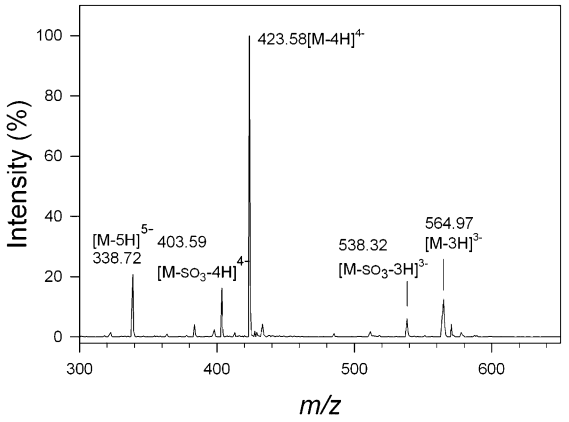

Fig. S7. ESI-MS spectrum of heptasaccharide substrate. The major signal at the *m/z* value of 423.58 represents a quadruply charged ion. The signals at the *m/z* value of 338.72 and 564.97 represents quintuply and triply charged ion, respectively. Signals at 403.59 and 538.32 represent the desulfated quadruply and triply charged ion, respectively. The measured molecular mass from this analysis was 1,698.3 ± 0.4 Da, which is very close to the calculated value of 1,698.1 Da.

Table S1. Data collection and refinement statistics

	3-OST-1 + heptasaccharide (PDB ID code 3UAN)
Data collection	
Space group	$P2_1$
Cell dimensions	
a, b, c, Å	52.88, 62.17, 86.37
α, β, γ, °	90.00, 95.16, 90.00
Resolution, Å	50.00–1.85
No. of molecules/AU	2
No. of unique reflections	46,408
Completeness (%; last shell)	96.9 (84.6)
Redundancy	3.7 (2.9)
R_{sym}*	5.00 (31.4)
Mean I/σI	22.0 (3.1)
Refinement	
R_{work},%†	17.4
R_{free},%‡	20.2
Rmsd§	
Bond length, Å	0.004
Bond angle, °	0.838
No. of atoms	
Protein (MolA/MolB)	2,088/2,086
Ligand (heptasaccharide, MolA/MolB)¶	96/86
Ligand (PAP, MolA/MolB)	26/26
Water	267
Average B-factor, Å^2	
Protein (MolA/MolB)	33.2/37.0
Ligand (heptasaccharide, MolA/MolB)¶	38.2/46.0
Ligand (PAP, MolA/MolB)	23.7/31.1
Water	42.3

*$R_{sym} = \Sigma\,(|I_i - \langle I\rangle|)/\Sigma(I_i)$, where I_i is the intensity of the ith observation and <I> is the mean intensity of the reflection.
†$R_{cryst} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, calculated from working data set.
‡R_{free} is calculated from 5% of data randomly chosen not to be included in refinement.
§Ramachandran results were determined by MolProbity (9).
¶Heptasaccharide substrate bound in molecule B refined with an occupancy of 0.7.

Table S2. Kinetic analysis of 3-OST-1 and 3-OST-3 toward oligosaccharide substrates

Mutant proteins	K_M, μM*	k_{cat}, min^{-1}	$k_{cat}/K_M \times 10^6$, M × min^{-1}	Curve-fitting coefficient, R^{2}†
3-OST-1 WT‡	5.5 ± 1.5	42.5 ± 3.7	7.7	0.97
3-OST-1 N167A	9.7 ± 2.3	26.6 ± 2.4	2.7	0.98
3-OST-1 H168A	4.4 ± 1.4	11.6 ± 1.1	2.6	0.95
3-OST-1 H168E	2.9 ± 0.5	7.3 ± 0.3	2.5	0.98
3-OST-1 K171A	6.3 ± 1.5	35.2 ± 2.8	5.6	0.97
3-OST-1 R268A	18.6 ± 6.3	6.0 ± 0.9	0.3	0.97
3-OST-3 WT§	3.7 ± 1.1	0.77 ± 0.05	0.208	0.98
3-OST-3 T256A	13.3 ± 1.1	0.68 ± 0.04	0.051	0.99
3-OST-3 T256V	12.3 ± 1.2	0.52 ± 0.1	0.042	0.99
3-OST-3 W283A	21.8 ± 7.3	0.10 ± 0.02	0.005	0.97

Although all the mutant proteins can transfer the sulfo group to the substrate, it appears that a mixture of products were obtained for some mutants based on the analysis of DEAE-HPLC. This suggests that some of the mutations altered the substrate specificity of the resulting enzyme (Figs. S5 and S6). Here, the kinetic analysis was only limited to those mutants that generate same products as the wild type proteins. At the present time, we could not determine the structures of the additional peaks resolved by DEAE-HPLC because the amount of the products was generally small and difficult to purify to homogeneity.
*K_M values were measured toward different concentrations of oligosaccharide substrates.
†Curve fitting was conducted by Sigma plot using Michaelis–Menten equation [$V = k_{cat}[E][S]/(K_M + [S])$].
‡Heptasaccharide (Fig. 3*A*) was used for the kinetic analysis of 3-OST-1 mutants.
§Octasaccharide (Fig. 3*C*) was used for the kinetic analysis of 3-OST-3 mutants.