The Heparin/Heparan Sulfate 2-O-Sulfatase from Flavobacterium heparinum

A STRUCTURAL AND BIOCHEMICAL STUDY OF THE ENZYME ACTIVE SITE AND SACCHARIDE SUBSTRATE SPECIFICITY*

Received for publication, November 8, 2002, and in revised form, December 16, 2002
Published, JBC Papers in Press, January 7, 2003, DOI 10.1074/jbc.M211425200

Rahul Raman‡‡, James R. Myette‡‡, Zachary Shriver‡, Kevin Pojasek‡‡, Ganesh Venkataraman**, and Ram Sasisekharan‡ ‡‡ From the §Division of Biological Engineering and **Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

In the previous paper (Myette, J. R., Shriver, Z., Claycamp, C., McLean, M. W., Venkataraman, G., and Sasisekharan, R. (2003) J. Biol. Chem. 278, 12157–12166), we described the molecular cloning, recombinant expression, and preliminary biochemical characterization of the heparin/heparan sulfate 2-O-sulfatase from Flavobacterium heparinum. In this paper, we extend our structure-function investigation of the 2-O-sulfatase. First, we have constructed a homology-based structural model of the enzyme active site, using as a framework the available crystallographic data for three highly related arylsulfatases. In this model, we have identified important structural parameters within the enzyme active site relevant to enzyme function, especially as they relate to its substrate specificity. By docking various disaccharide substrates, we identified potential structural determinants present within these substrates that would complement this unique active site architecture. These determinants included the position and number of sulfates present on the glucosamine, oligosaccharide chain length, the presence of a 4,5-un saturated double bond, and the exolytic versus endolytic potential of the enzyme. The predictions made from our model provided a structural basis of substrate specificity originally interpreted from the biochemical and kinetic data. Our modeling approach was further complemented experimentally using peptide mapping in tandem with mass spectrometry and site-directed mutagenesis to physically demonstrate the presence of a covalently modified cysteine (formylglycine) within the active site. This combinatorial approach of structure modeling and biochemical studies provides insight into the molecular basis of enzyme function.

Heparin and heparin sulfate glycosaminoglycans (HSGAGs) are structurally complex linear polysaccharides (1, 2) composed

* This work was supported by National Institutes of Health (NIH) Grants GM 57073 and CA90940. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
‡‡ To whom correspondence should be addressed: Division of Biological Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139. Tel.: 617-258-9494; Fax: 617-258-9409.
§ Recipient of a Merck/MIT Fellowship.
¶ recipient of NIH/MIT Toxicology Training Grant ST32GM08334.
¶¶ Recipient of NIH Biotechnology Training Grant ST32GM08334 and a Whitaker Foundation predoctoral fellowship.

The abbreviations used are: HSGAG, heparin/heparan sulfate glycosaminoglycan; GAG, glycosaminoglycan; 2-O ΔN1–24, recombinant of repeating disaccharides of uronic acid (α-L-iduronic or β-D-glucuronic) linked 1→4 to α-D-glucosamine. This structural complexity derives principally from the variable chemical modifications made to the polysaccharide chain. Such modifications include acetylation or sulfation at the N-position of the glucosamine, epimerization of glucuronic acid to iduronic acid, and additional O-sulfation at the 2-O-position of the uronic acid in addition to the 3-O, 6-O-position of the adjoining glucosamine. It is a highly variable sulfation pattern, in particular, which ascribes to each GAG chain a unique structural signature. In turn, this signature dictates specific GAG-protein interactions underlying critical biological processes related to cell and tissue function. Given this critical structure-function relationship of GAG sulfation, enzymes that can hydrolyze these sulfates in a structurally specific manner are important tools for the determination of GAG fine structure to better ascertain these structure-function relationships. In the previous paper (21), we described the cloning, recombinant expression, and biochemical characterization of one such sulfatase, the 2-O-sulfatase from Flavobacterium heparinum.

As members of a large enzyme family, the sulfatases hydrolyze a wide array of sulfate esters (for a review, see Refs. 3 and 4). Their respective substrates include sulfated complex carbohydrates such as the glycosaminoglycans (GAGs), steroids, sphingolipids, xenobiogenic compounds, and amino acids such as tyrosine. Additionally, many of these enzymes are able to hydrolyze in vitro smaller synthetic substrates (e.g. 4-nitrophenyl sulfate and catechol sulfate). It is for this reason that these enzymes are often generically described as “arylsulfatases” (even when their preferred in vivo substrate is ill defined). Despite their disparate substrate specificities, the members of this enzyme family share both considerable structural homology and a common catalytic mechanism with one another (5). Each sulfatase possesses a signature catalytic domain toward its amino terminus, which is readily identified by the consensus sequence (C/S)xPXRXXXX(S/T)G. The conserved cysteine (or less commonly serine) within this sulfatase motif is of particular functional importance, since it is covalently modified to an L-Ca-formylglycine (L-2-amino-3-oxopropionic acid) (4, 6, 7). The ubiquitous importance of this chemical modification was first functionally identified by its relationship to the etiology of multiple sulfatase deficiency, a genetically recessive disorder in which there is a complete lack of sulfatase activity

2-O-sulfatase lacking NH$_2$-terminal signal sequence composed of first 24 amino acids; MALDI-MS, matrix-assisted laser desorption ionization spectrometry; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); AT-10, antithrombin-binding oligosaccharide; ΔU, uronic acid with a Δ4,5 double bond.
due to a lack of this critical aldehyde (FGly) within the active site of all expressed sulfatases (8).

The crystal structures of two human lysosomal sulfatases, arylsulfatase A (ceresoidase-3-sulfate 3-sulfohydrolyase) (9, 10) and arylsulfatase B (N-acetylglactosamine-4-sulfate 4-sulfohydrolyase) (11), and a bacterial arylsulfatase from Pseudomonas aeruginosa (12) have been solved. These three sulfatases share an identical alkaline phosphatase-like structural fold (according to the Structural Classification of Proteins Database2 composed of a series of mixed parallel and antiparallel β-strands flanked by long and short α-helices on either side (9–12). In addition to their common structural fold, these sulfatase structures also possess a high degree of homology within their respective active sites, especially in the region localized around the modified cysteine (FGly). Taken together, these crystal structures present a clear and consistent description of conserved active site residues and their potential catalytic role in sulfate ester hydrolysis. At the same time, this strong structural homology is somewhat surprising, considering that at least two of these sulfatases act on notably different substrates (e.g. sulfated sphingolipid versus sulfated GAG). The structural basis for substrate specificity, therefore, remains to be determined. The question of this relationship of enzyme active site structure to substrate specificity would appear to be especially relevant to GAGs where multiple sulfate esters are present.

Using the strong structural and functional similarities that have been observed among the members of the sulfatase family of enzymes, we extend our characterization of the recombinant 2-O-sulfatase in this paper to include a structure-function description of the enzyme’s substrate specificity. To do so, we first constructed a homology-based structural model for the wild type 2-O-sulfatase using the three crystal structures (cited above) as a molecular scaffold. From this model, we present a picture of the enzyme active site from which we have identified specific residues likely to be involved in substrate binding and catalysis. Predictions concerning substrate discrimination made from this model are supported by the biochemical and kinetic data presented in the previous paper. In addition, a specific prediction regarding the sulfatase’s exolytic mode of action is made; this prediction is subsequently validated by the biochemical data. Finally, we use a combination of aldehyde-specific chemical labeling and peptide mapping methods in parallel with site-directed mutagenesis to identify the predicted active site cysteine, which is covalently modified to a formylglycine.

**EXPERIMENTAL PROCEDURES**

**Reagents**—The unsaturated heparin decasaccharide ΔU2SHN3 glucose3, 6-methoxyl, galactosamine (H) and glucosamine. Subscripts 2S, 3S, and 6S indicate 2-O, 3-O, and 6-O-sulfations, respectively. NAc and NS, N-acetylation and N-sulfation of glucosamine.

Protein Data Bank accession numbers are as follows: cerebroside-3-sulfate (1E2S); N-acetylgalactosamine-4-sulfate (1PSU); P. aeruginosa arylsulfatase (1HDH).

with different bacterial and lysosomal sulfatases was performed using ClustalW (14) (Fig. 1). Based on this multiple sequence alignment, three model structures of 2-O-sulfatase were obtained corresponding to its alignment with the other three sulfatases. The models were constructed using the Homology module of the Insight II molecular simulations package (Accelrys, San Diego, CA). The side chain of the critical Cys-82, which is known to undergo posttranslational modification in the active enzyme, was replaced by the geminal diol (C6(OH)2). The potentials for the model structures were assigned using the AMBER force field (15). The deletions in the modeled structures were closed using 200 steps of steepest descent minimization without including charges by keeping most of the structure rigid and allowing the regions close to the deletion move freely. The final refined structure was subjected to 400 steps of steepest descent minimization without charges and 400 steps of conjugate gradient minimization including charges.

**Molecular Docking of Disaccharide Substrates into the Active Site of the Modeled 2-O-Sulfatase**—Heparin-derivated sulfatase with a ΔU at the nonreducing end were modeled as follows. The coordinates of the trisulfated ΔU containing disaccharide (ΔU2S[HNS]3) were obtained from the cis-crystal structure of a heparinase-derived hexasaccharide with fibroblast growth factor 2 (Protein Data Bank code 1BPC). This trisulfated disaccharide structure was used as a reference to generate the structural models for other disaccharides including ΔU2S[HNS]2, ΔU2S[HNS]1, and ΔU2S[HNS]0. The coordinates of trisulfated disaccharides (ΔU2S[HNS]3) containing iduronic acids in the C1 and S6 conformations were also obtained from 1BPC. Similarly, chondroitin sulfate-derived disaccharides ΔU2S[GalNAc,4S] and ΔU2S[GalNAc,6S] were modeled using a reference structure of a chondroitin-4-sulfate disaccharide ΔUGalNAc,4S whose coordinates were obtained from its co-crystal structure with the chondroitinase B enzyme (Protein Data Bank code 1BPC). The potentials for these disaccharides were assigned using the AMBER force field modified to include carbohydrates (15) with sulfate and sulfamate groups (16).

The orientation of the cleavable sulfate group relative to Oy of the geminal diol in the active site of human arylsulfatase A and the bacterial arylsulfatase was identical as observed in their respective crystal structures. This orientation was such that one of the faces of the tetrahedral formed by the three oxygen atoms of SO3 was directed toward Oy, facilitating the nucleophilic attack of the sulfur atom and the transfer of the SO3 group to Oy. This highly specific orientation of the sulfate group helped in positioning the disaccharide substrate relative to the active site of the 2-O-sulfatase. After fixing the orientation of the 2-O-sulfate group, the glycosidic torsion angles and exocyclic torsion angles were adjusted manually to remove unfavorable steric contacts with the amino acids in the active site. The enzyme substrate complexes were minimized using 200 steps of steepest descent followed by 400 steps of Newton-Raphson minimization including charges. Most of the enzyme was kept rigid, and only the loop regions constituting the active site were allowed to move freely. On the other hand, a force field energy of 7000 kcal/mol was applied to constrain angles to fix the ring conformation of the disaccharides during the energy minimization. The manual positioning of the substrates was done using the Viewer module, building of the disaccharide structures from the reference structures was done using the Builder module, and the energy minimization was done using the Discover module of Insight II.

**Recombinant Expression and Protein Purification of the Flavobacterial 2-O-Sulfatase**—The molecular cloning, recombinant expression in Escherichia coli, and subsequent purification of the 6× histidine-tagged 2-O-sulfatase by nickel affinity chromatography is as described in the preceding paper. The sulfatase was expressed as an amino-terminal truncation lacking its putative signal sequence (commonly referred to in papers as 2-O Δ1-24).

**AT-10 Compositional Analyses by Capillary Electrophoresis and MALDI-MS**—Approximately 10 μg of the AT-10 oligosaccharide was incubated with 100 pmol of 2-O Δ1-24 in a 40-μl reaction volume at 30 °C. 15-μl aliquots were removed at 4 and 17 h and heat-inactivated at 95 °C. The oligosaccharide reaction products (along with 15 μl of a mixture of the control) were subjected to an extraction I and III digest and subsequent capillary electrophoresis-based compositional analysis. Desulfation of the decasaccharide was assayed in parallel by MALDI-MS using established methods (17).

**2-O-Sulfatase Active Site Labeling and Peptide Mapping**—Approximately 500 μl of 6× histidine-tagged 2-O Δ1-24 (wild type enzyme and Cys-82 to serine) were directed mutant were lyophilized by freeze-drying and vigorously resuspended in 90 μl of denaturation buffer containing 6 μl guanidinium hydrochloride, 0.1 M Tris-HCl, pH 7.5. Active site aldehydes were fluorescently labeled by adding 25 μl of Texas Red hydrazine made up as a 10 mst stock in dimethyl formamide. Labeling
was carried out for 3 h at room temperature with gentle mixing on a rotating platform. The hydrazono linkage was stabilized by the addition of 10 µl of a fresh 5 mM sodium cyanoborohydride stock made up in 1 N NaOH. Reduction was carried out for 1 h at room temperature. Unreacted fluorohyde was removed by repeated aceton precipitation (added 5:1 v/v). Acetone was prechilled at -20 °C. Samples were chilled at -85 °C for 20 min prior to spinning in a microcentrifuge for 10 min, maximum speed, at 4 °C. Pellets were dried by Speed-Vac centrifugation.

The labeled sulfatase (and unlabeled control) were proteolyzed with sequence grade-modified trypsin for 20 h at 37 °C in digestion buffer that contained 0.1 M Tris-HCl, pH 8.5, 1 mM EDTA, 1 mM dithiothreitol, and 10% acetonitrile (v/v) in a 30-µl reaction volume. Trypsin was first reconstituted as a 2.5% stock in 1% acetic acid and added in a 1:5 ratio (w/w) relative to the target protein. Following trypsin digestion, peptide cysteines were reduced by the addition of 50 mM dithiothreitol (in the dark) by the addition of 150 mM iodoacetamide, added from a 2 M stock made up in 0.1 M Tris-HCl, pH 8.5. This reduction-alkylation cycle was repeated one more time. Molecular masses of select peptides were determined by MALDI-MS as described (18) using 1 µl of a-cyano-4-hydroxycinnamic acid in 50% acetonitrile, 0.3% trifluoroacetic acid as a matrix.

**Site-directed Mutagenesis of the C82A Active Site Mutant—**The site-directed mutant C82A was cloned by recombinant PCR using outside primers 5'-GCT ATG CAT ATG CAA ACC TCA AAA GTA GCA GCT-3′ (forward) and 5'-GT CTC GAG GAT CCT TAT TTT TTT AAT GCA TAA AAC GAC TCC-3′ (reverse) in addition to the following mutagenic primer pair: 5′-C CAG CCG CTC CTT ACA ACC TCA GCA CCT-3′ (forward) and 5′-CG TAA GGA TCT AGA CAT ATG CAA ACC TCA AAA GTA GCA GCT-3′ (reverse). The engineered codon change for each DNA strand is underlined. Subcloning into pET28a, recombinant expression in the E. coli strain BL21 (DE3), and subsequent purification by nickel chelation chromatography using the N-terminal 6× histidine purification tag are as described for 2-O-3′,3′-di-O-sulfate (previous paper (21)).

2-O-Sulfatase Cysteine Modification with 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB) and Dithiothreitol—Sulfatase cysteines were titrated based on molar absorptivity using Ellman’s reagent (DTNB) generally as described (19). 50 µM 2-O-sulfatase (2-O-3′,3′-di-O-sulfate) was reacted with 500 µM DTNB in a 1:4 ml reaction volume containing 0.1 mM sodium phosphate, pH 8.0, and 5 mM EDTA. Absorbance was measured at 412 nm following a 15-min incubation at room temperature with a minus enzyme control used as a blank. For measurements made under denaturing conditions, SDS was added to a final concentration of 2%. Molar extinction coefficients used were as follows: non-denaturating (minus SDS), 14,150 M⁻¹ cm⁻¹; denaturing (2% SDS), 12,500 M⁻¹ cm⁻¹.

**RESULTS AND DISCUSSION**

Structure-based Homology Modeling of the 2-O-Sulfatase Active Site—In comparing the structures of human arylsulfatase A (9, 10), arylsulfatase B (N-acetylgalactosamine-4-sulfatase) (11), and a bacterial arylsulfatase from Pseudomonas aeruginosa (12), we observed an obvious structural homology between each of them, especially as it pertained to a conservation of critical active site residues and their spatial arrangement. By extension, most of these amino acids were likewise conserved in the flavobacterial 2-O-sulfatase as evident by a direct alignment of their primary sequences (Fig. 1). We used this close structural relationship to construct three homology-based models for the flavobacterial 2-O-sulfatase, each one based on one of the three crystal structures examined. We ultimately chose as our representative 2-O-sulfatase structure the homology model constructed using the N-acetylgalactosamine-4-sulfatase (aryl sulfatase B) (Fig. 2). This decision was largely based on arylsulfatase B also being a GAG desulfating enzyme. In this model, we replaced cysteine 82 with a formylglycine (FGly-82). We chose to represent FGly-82 in the hydrated state as a geminal diol (Cβ(OH)₂), consistent with the proposed resting state (before catalysis) of the enzyme (5, 9).

In the structural model, the surface of the active site pocket is composed of many amino acids that can potentially interact with a disaccharide substrate (Fig. 3). These include Lys-107, Lys-175, Lys-238, Gln-237 and -309, Thr-104, Glu-106, and Asp-159. Lysines and glutamines are commonly occurring amino acids in heparin binding sites that interact with the sulfate and carboxylate groups of the GAG chain. Unlike the amino acids proximal to the FGly-82, these residues are not conserved in the other sulfatases that we examined (Table I, denoted in normal type), suggesting a potentially unique role of these amino acids in dictating oligosaccharide substrate specificity. This disparity is particularly true when directly comparing the 2-O-sulfatase and arylsulfatase A; many of the nonconserved amino acids in the 2-O-sulfatase are charged, whereas those in arylsulfatase A are predominantly hydrophobic. This observation is consistent with the obvious structural distinction of their respective substrates (i.e. the highly sulfated HS/GAG substrates of the 2-O-sulfatase versus the long hydrophobic alky chains of cerebroside 3-sulfate substrate of arylsulfatase A).

**Active Site Topology and the Exolytic Action of the 2-O-Sulfatase—** Upon inspection of the 2-O-sulfatase structure, several amino acids that potentially constitute the active site were identified (Table I). There are several structurally conserved basic amino acids in the proximity of FGly-82 including Arg-56, Lys-134, His-136, and Lys-308. The topology of the active site as observed in our structural model indicated that the critical FGly-82 and the basic amino acid cluster are located at the bottom of a deep pocket (Fig. 2B). Such a restrictive access to the active site would appear to impose a clear structural constraint on the substrate as it relates to the position of the 2-O-sulfate group within the oligosaccharide chain (i.e. externally versus internally positioned) upon which the enzyme acts. We would predict from this topology that only a sulfate group present at the nonreducing end of the oligosaccharide will be favorably positioned for catalysis; the juxtaposition of an internal sulfate into the active site would require a substantial bending of the oligosaccharide chain. Such chain distortion would be sterically unfavorable. Based on these constraints, therefore, we predict the sulfatase to hydrolyze 2-O-sulfates in an exclusively exolytic fashion. This exclusivity for the nonreducing end does not necessarily preclude, however, the enzyme acting on longer chain oligosaccharides (i.e. those exceeding a disaccharide in length), provided that they in fact possess sulfates at the terminal 2-OH-position. The model does suggest a likely kinetic preference for disaccharide substrates, since they would most readily diffuse into and out of this narrow active site (see enzyme-substrate structural modeling below).

We addressed this important question using as a substrate the purified heparin-derived AT-10 decasaccharide \( \Delta_{255}HNS,65 \) as a substrate for the 2-O-sulfatase. The 2-O-desulfated decasaccharide was then subjected to heparinase treatment. Capillary electrophoresis-based compositional analyses indicated the disappearance of the disaccharide \( \Delta_{255}HNS,65 \), suggesting that this disaccharide remained after sequential treatment with the 2-O-sulfatase and heparin lyases (Fig. 4). Loss of a single sulfate was independently determined by mass spectrometry. Sequencing analysis of the singularly desulfated product allowed us to assign the original position of the 2-O-sulfate exclusively to the uronic acid present at the nonreducing end. This result was confirmed using other oligosaccharide substrates, including tetra- and hexasaccharides containing both nonreducing end 2-O-sulfates and internal 2-O-sulfates (data not shown). Taken together, these results strongly suggest a sulfatase whose hydrolytic activity proceeds in an exolytic fashion. This conclusion is supported by our model, which,
given the narrow topology of the enzyme active site, predicts a strong preference for the cleavage of sulfates positioned at the nonreducing end.

Enzyme-Substrate Structural Complex: Interactions between 2-O-Sulfatase and Disaccharides—Since the active site can readily accommodate disaccharide substrates, we modeled several unsaturated glycosaminoglycan disaccharides as described under "Experimental Procedures." Our choice of 2-O-sulfatase substrates was logical for two reasons: 1) eliminative cleavage of a HS polysaccharide by the flavobacterial lyases that naturally occurs in vivo results in the formation of disaccharides and other small oligosaccharides, all possessing a 4,5-unsaturated bond at the nonreducing end uronic acid and 2) the obligatory substrate-product relationship between the 2-O-sulfatase and the 4,5-glycuronidase that exists both in vitro (as described in the previous paper (21)) and in vivo. In addition, we directly tested the requirement of a 4,5 uronic acid directly in experiments described below. A representative structural complex involving the trisulfated disaccharide H2S, U2SHNS,6S (Fig. 3) was used to describe the molecular interactions between the enzyme and the substrate. This choice was made based on the biochemical data (shown in the accompanying manuscript (21)) that indicated the trisulfated disaccharide as the kinetically preferred substrate. A description of these interactions and their proposed functional roles is shown in Table II. The functional roles of the conserved active site amino acids (listed in boldface type in Table I) were proposed based on their interactions with the 2-O-sulfate group and/or the geminal.

Flavobacterial 2-O-Sulfatase Structure-Function Studies

![Multiple sequence alignment of the sulfatases using ClustalW.](image)

The flavobacterial enzyme is a member of a large sulfatase family. The putative active site is boxed, with critically modified cysteine noted by an asterisk. Invariant residues are shaded in dark gray, those with partial identity in light gray, and conservative substitutions in charcoal. Multiple sequence alignment was generated by ClustalW using only select sequences identified from a BLASTP search of the protein data base. Enzymes are abbreviated as follows. FH2S, F. heparinum 2-O-sulfatase; PARS, P. aeruginosa arylsulfatase; MDSA, Prevotella sp. MdSA gene; HGaLS, human N-acetylgalactosamine-6-sulfate sulfatase (chondroitin 6-sulfatase); HARSA, human cerebroside-3-sulfate sulfatase (aryl- sulfatase A); HARSB, human N-acetylgalactosamine-4-sulfate sulfatase (aryl- sulfatase B); H2S, human idurionate-2-sulfate sulfatase; cons, consensus sequence. The GenBank protein accession numbers for sulfatases listed in Fig. 1 are as follows: CAA88421, P. aeruginosa arylsulfatase; AAF72520, Prevotella sp. MdSA mucin desulfating gene; AAC51350, Homo sapiens N-acetylgalactosamine-6-sulfate sulfatase; AAB03341, H. sapiens cerebroside-3-sulfate sulfatase (aryl- sulfatase A); AAA5784, H. sapiens N-acetylgalactosamine-4-sulfate sulfatase (aryl- sulfatase B); AAA63197, H. sapiens idurionate-2-sulfate sulfatase.

Flavobacterial 2-O-Sulfatase Structure-Function Studies

![Multiple sequence alignment of the sulfatases using ClustalW.](image)
nal diol of the formylglycine at position 82. Identical roles have been proposed for the corresponding amino acids in the three known sulfatase crystal structures (Table I).

The Requirement for an Unsaturated Δ4,5 Nonreducing Terminus—A closer inspection of the modeled enzyme-substrate complex revealed some interesting possibilities pertaining to the role of the nonconserved amino acids in substrate recognition and binding. The planar carboxylate group attached to the C-5 atom of the Δ4,5 uronic acid is oriented in such a manner as to potentially interact with Lys-175 and Lys-238. These amino acids could play an important role, therefore, in favorably orienting the 2-O-sulfate within the active site. We were further interested in this arrangement given the additional constraint imposed upon the planar carboxyl group of the uronic acid by the presence of the C-4–C-5 double bond. This constraint may further influence substrate orientation within the active site. As such, we would predict a substrate discrimination exhibited by the 2-O-sulfatase that is based on the presence of the Δ4,5 double bond at the oligosaccharide nonreducing terminus. In the absence of this double bond, the favorable orientation of the C-5 carboxylate afforded by charge interactions with Lys-175 and Lys-238, respectively, would not

Fig. 2. Structural model of the 2-O-sulfatase and topology of active site. A, ribbon diagram of the proposed 2-O-sulfatase structure constructed using homology modeling of the crystal structure of human arylsulfatase B. The β-strands are colored red, and the α-helices are colored blue. The geminal diol form of the modified cysteine is rendered as CPK (carbon colored green and oxygen colored red). The direction of substrate diffusing into the active site is indicated by an arrow. B, CPK rendering of the top view of the structure shown in A. The modified cysteine is colored purple, and the surrounding basic amino acids (Arg, His, and Lys) are colored dark purple, acidic amino acids (Asp and Glu) are colored red, and Gln and Asn are colored light blue. Note that the active site geminal diol is located in the bottom of a deep cleft.

Fig. 3. Active site amino acids and their interaction with AU_2H_{NSGSH}. A, stereo view of the 2-O-sulfatase active site highlighting important amino acids (shown here by a stick representation). Basic residues are colored blue, acidic amino acids are colored red, Gln is colored light blue, Thr is colored orange, Leu is colored brown, and FGly-82 is colored purple. The docked disaccharide (colored in green) is also shown using a stick representation. The sulfur atom of 2-O-sulfate group is colored yellow, and oxygen atoms of the 2-O-sulfate group and the planar carboxyl group are colored red. B, schematic representation of the amino acids shown in A using the same color scheme. Potential metal ion coordination is also shown, with the divalent cation (M^{2+}) depicted as a gray circle.
Flavobacterial 2-O-Sulfatase Structure-Function Studies

Highly conserved amino acids are listed in boldface type. Nonconserved amino acids are listed in normal type. Amino acids in the 2-O-sulfatase that could be potentially involved in substrate binding are noted by an asterisk. Structural alignment of the modeled 2-O-sulfatase structure with the other sulfatases was obtained based on superposition of their Cα traces using the combinatorial extension algorithm (20). Regions of deletion in the structural alignment are noted with a minus sign.

| 2-O-Sulfatase (P. heparinum) | Arylsulfatase A (human) | Arylsulfatase B (human) | Arylsulfatase (P. aeruginosa) |
|-------------------------------|------------------------|------------------------|-------------------------------|
| Cys-82                        | Cys-69                 | Cys-91                 | Cys-51                        |
| Arg-86                        | Arg-73                 | Arg-95                 | Arg-55                        |
| Lys-134                       | Lys-123                | Lys-145                | Lys-113                       |
| His-136                       | His-125                | His-147                | His-115                       |
| Lys-308                       | Lys-302                | Lys-318                | Lys-375                       |
| Gln-237                       | His-229                | His-242                | His-211                       |
| Asp-42                        | Asp-29                 | Asp-53                 | Asp-13                        |
| Gln-43                        | Asp-30                 | Asp-54                 | Asp-14                        |
| Asp-295                       | Asp-281                | Asp-300                | Asp-317                       |
| His-296                       | Asn-282                | Asn-301                | Asn-318                       |
| Lys-238*                      | Tyr-230                | Glu-243                | Trp-212                       |
| Lys-175*                      | Gln-153                | Arg-180                | Pro-161                       |
| Asp-159*                      | His-151                | Ser-172                | Ala-139                       |
| Thr-104*                      | Val-91                 | Ile-113                | —                             |
| Glu-106                       | Val-93                 | Trp-115                | —                             |
| Lys-107*                      | Pro-94                 | Pro-116                | —                             |
| Glu-309*                      | Gly-303                | Trp-319                | Ala-376                       |

We also modeled enzyme-substrate complexes containing two unsaturated chondroitin sulfate disaccharides (ΔU_{2S} GalNAc,4S and ΔU_{2S} GalNAc,6S). In comparison with our original model using the heparin disaccharide substrate, we found interactions with the 2-O-sulfate and carboxyl group of the ΔU monosaccharide that were identical to that of ΔU_{2S}H_{NS,6S}. There were fewer interactions involving the 4-sulfate and 6-sulfate groups, however. This particular model, therefore, is consistent with our kinetic data describing the sulfates present on the adjoining glucosamine.

The ability of the sulfatase to hydrolyze internally positioned 2-O-sulfates within the AT-10 decasaccharide and subsequent compositional analyses of the heparinase-treated product. A, AT-10 decasaccharide sequence with property-encoded nomenclature (22) and outline of experimental design. B, capillary electrophoretagram for both the control and sulfatase-pretreated samples along with their saccharide compositional assignments. Heparinase cleavage products following sulfatase pretreatment are shown as a dashed line (with gray fill). Minus sulfatase control is shown as a white line (as fill). The tetrasulfated tetrasaccharide (+4–7) is also noted. Disappearance of the triSulfated disaccharide (D) by one-third and the corresponding appearance of the 2-O-desulfated product (ΔU_{NS,6S}) are depicted by vertical arrows. The minor tetrasaccharide contaminant is noted by an asterisk.

![Fig. 4. The exolytic activity of the 2-O-sulfatase.](image)

**TABLE I**
Structure-based comparison of sulfatase active site residues

| 2-O-Sulfatase (P. heparinum) | Arylsulfatase A (human) | Arylsulfatase B (human) | Arylsulfatase (P. aeruginosa) |
|-------------------------------|------------------------|------------------------|-------------------------------|
| Cys-82                        | Cys-69                 | Cys-91                 | Cys-51                        |
| Arg-86                        | Arg-73                 | Arg-95                 | Arg-55                        |
| Lys-134                       | Lys-123                | Lys-145                | Lys-113                       |
| His-136                       | His-125                | His-147                | His-115                       |
| Lys-308                       | Lys-302                | Lys-318                | Lys-375                       |
| Gln-237                       | His-229                | His-242                | His-211                       |
| Asp-42                        | Asp-29                 | Asp-53                 | Asp-13                        |
| Gln-43                        | Asp-30                 | Asp-54                 | Asp-14                        |
| Asp-295                       | Asp-281                | Asp-300                | Asp-317                       |
| His-296                       | Asn-282                | Asn-301                | Asn-318                       |
| Lys-238*                      | Tyr-230                | Glu-243                | Trp-212                       |
| Lys-175*                      | Gln-153                | Arg-180                | Pro-161                       |
| Asp-159*                      | His-151                | Ser-172                | Ala-139                       |
| Thr-104*                      | Val-91                 | Ile-113                | —                             |
| Glu-106                       | Val-93                 | Trp-115                | —                             |
| Lys-107*                      | Pro-94                 | Pro-116                | —                             |
| Glu-309*                      | Gly-303                | Trp-319                | Ala-376                       |

To better understand this likely structural constraint, we superimposed substrate disaccharides containing a nonreducing end iduronic acid in either the 1C or 2S conformation (data not shown). The superimposition was such that the S–O–C–C-1 atoms of all of the uronic acids coincided, thereby fixing the orientation of the 2-O-sulfate group. In this model, the carboxylate groups of the iduronic acid containing disaccharide substrates were, in fact, pointing away from the active site pocket and were not positioned to interact as favorably with Lys-175 or Lys-238 as compared with the original disaccharide substrate possessing a planar C-5 carboxylate.

We assessed this prediction experimentally by testing the ability of the 2-O-sulfatase to hydrolyze size-fractionated hexasaccharides derived from the nitrous acid treatment of heparin. Unlike enzymatic cleavage, these chemically derived heparin saccharides do not possess a Δ4,5-unsaturated bond at their respective nonreducing ends. A majority of the resultant tetrasaccharides, however, do contain an I_{3S} at this end. Using MALDI-MS, we were unable to detect any enzyme-dependent desulfation of treated hexasaccharides under conditions that resulted in quantitative elimination of 2-O-sulfate from Δ4,5-containing oligosaccharides. This result strongly suggests a structural requirement for the Δ4,5 bond in addition to the exolytic preference of the enzyme. As we have already pointed out, our model provides a rationale for this requirement (i.e. the obligatory physical connection between this bond and the planar carboxylate at the uronic acid C-5 position). The constraint imposed by a planar C-4–C-5 double bond on the orientation of this carboxylate in turn confers upon the same substrate an orientation of the 2-O-sulfate within the enzyme active site that is favorable for catalysis.

**Enzyme-Sulfate Interactions in the 2-O-Sulfatase Active Site**—In the previous paper (21), we present a biochemical data demonstrating a clear kinetic preference of the enzyme for highly sulfated disaccharide substrates, namely those possessing a glucosamine sulfated at the 6-OH- and 2-N-positions. Our structural model of the sulfatase-trisulfated disaccharide complex also points out key interactions involving these additional
ability of the so-called “heparin/heparan sulfate” 2-O-sulfatase to hydrolyze 2-O-sulfated chondroitin disaccharides. Given a lack of additional favorable contacts between the enzyme and substrate (e.g. with either the 4-O- or 6-O-sulfates), our model also predicts a lower catalytic efficiency for the chondroitin disaccharides relative to the structurally corresponding heparin disaccharides.

**Metal Ion Coordination** — In discussing this model, we must briefly consider the potential role of divalent metal ions. We decided not to include any such metal ions in our theoretical model of the 2-O-sulfatase, since we could find no divalent metal requirement for enzymatic activity (see preceding paper (21)). A divalent metal ion is present, however, in all three sulfatase crystal structures that we examined. In each case, the metal ion coordinates with the oxygen atoms of the sulfate group of the respective substrate. Additionally, a cluster of four highly conserved acidic amino acids has been observed to coordinate this divalent metal ion.

In the case of human arylsulfatase B, for example, the oxygen atoms of Asp-53, Asp-54, Asp-300, and Asn-301 are coordinated with a Ca2⁺ ion. Three of the four corresponding amino acids in the flavobacterial sulfatase model that we have identified as potentially coordinating with a metal ion are Asp-42, Gln-43, and Asp-295 (Table I). The fourth amino acid in the 2-O-sulfatase corresponds spatially to Asn-301 of arylsulfatase B and His-296. The positive charge of this position, however, does not favor the proximal location of a divalent metal cation. It is perhaps this unfavorable charge interaction that interferes with proper metal ion coordination.

**Enzyme-Substrate Model: Mechanism for Catalysis** — Nearly identical mechanisms for the hydrolysis of the sulfate ester bond involving the conserved active site amino acids have been proposed for human arylsulfatases A and B and the bacterial sulfatase from *P. aeruginosa*. Thus, the mechanism proposed for these other sulfatases should logically apply to the 2-O-sulfatase. The resting state of the active sulfatase in each of the crystal structures is proposed to contain the geminal diol, which is stabilized by interactions with basic residues. His-136 and Arg-86 of the flavobacterial 2-O-sulfatase enzyme are positioned appropriately in the active site to do so (Fig. 2B). A critical step in catalysis involves the correct positioning of the 2-O-sulfate group such that the sulfur atom is accessible to the O-γ1 of the geminal diol. We have already described how interactions of specific active site amino acids with the planar carboxyl group of the uronic acid (Lys-175 and -238), with the 6-O-sulfate of the glucosamine (e.g. Lys-107 and possibly Thr-104) and with the 2-O-sulfate itself (Lys-134 and -308) are likely to serve in this capacity (Table II). At the same time, interaction of the 2-O-sulfate group with charged amino acids would also enhance any electron density withdrawal from the oxygen atoms, thereby increasing the electrophilicity of the sulfur center. It has also been suggested that the nucleophilicity of the O-γ1 atom is enhanced by a possible proton donation to a neighboring aspartic acid residue. In our structural model of the 2-O-sulfatase, this residue would correspond to Asp-295.

An S₂O₅ mechanism has been proposed to follow the above steps and eventually lead to the cleavage of the sulfate ester bond (10, 12). In this mechanism, the exocyclic oxygen atom on the leaving substrate may be protonated by water or potentially by neighboring amino acids. In the 2-O-sulfatase active site model, Lys-308 is juxtaposed to protonate the leaving group (Fig. 2). The resulting sulfate group on the geminal diol is subsequently eliminated by abstraction of a proton from O-γ2, regenerating the formylglycine. His-136 is positioned to abstract this proton. Our functional assignment of these lysine
and histidine residues is based entirely on homologous positions found in the other sulfatases examined and their proposal to catalytically function in the fashion described (5).  

2-O-Sulfatase Peptide Mapping and Chemical Modification of Active Site Formylglycine—Finally, in describing the structure-function relationship of the 2-O-sulfatase active site, we come to the central catalytic player itself, the formylglycine at position 82. The recombinant expression of catalytically active 2-O-sulfatase in E. coli functionally argues for this covalent modification of the active site in vitro. We further established the catalytic function of Cys-82 by site-directed mutagenesis. The mutant (C82A) was recombinantly expressed and purified as a histidine-tagged protein in the same manner employed for the wild type enzyme. Comparable expression levels of soluble protein were achieved. The C82A mutant, however, was completely inactive (data not shown). Both the wild type and mutant possessed the same secondary structure as exhibited by their virtually superimposable CD spectra (data not shown), arguing against any adverse global conformational changes induced by the molecular replacement of the cysteine by alanine. 

We also set out to demonstrate the physical presence of the FGY at position 82 by the tandem use of protein chemistry and mass spectrometry. 10 nmol of wild-type sulfatase (2-O JN1–24) and the C82A mutant were reacted with Texas Red hydrazide (620.74 Da) as described under “Experimental Procedures.” The two sulfatase fractions were subsequently trypsinized under mildly denaturing conditions followed by reductive methylation of the unmodified cysteines. The molecular masses of the resultant peptides were determined by MALDI-MS (Fig. 5). In this experiment, we identified a single ionized species uniquely present in the labeled sulfatase experiment (Fig. 5B) but absent in the active site mutant (Fig. 5C) or in the unlabeled control (Fig. 5A). The empirical mass of this species corresponded most closely to the peptide sequence FTRYACAPFLCTPSR, resulting from a partial trypsin cleavage. This peptide contains the sulfatase consensus sequence CXPR, which includes the critical active site cysteine (underlined) at position 82. The mass of this peptide is consistent with first the conversion of this cysteine to a formylglycine (FGY-82) followed by the covalent hydrazine linkage of the aldehyde-reactive fluorophore at this position. It also takes into account the carbamidomethylation of the second (unmodified) cysteine present in this peptide. These data, taken together with the loss of function observed for the C82A mutant, establish the important structure-function relationship for this active site modification. 

To confirm and extend these findings, we also completed chemical modification experiments on the wild type enzyme using the cysteine-specific reagent DTNB (Ellman’s reagent). Apart from the catalytic Cys-82 that is converted to a formylglycine, none of the remaining seven cysteines appear to be highly conserved among other members of the sulfatase family (Fig. 1). We were unable to effectively inhibit enzyme activity with the addition of DTNB or dithiothreitol (data not shown). This general lack of inhibition by these two cysteine-reactive agents suggests at least two probabilities. First, the 2-O-sulfatase does not require intramolecular disulfide linkages to critically stabilize a catalytically active conformation. This presumption is not surprising, given the fact the native 2-O-sulfatase is a bacterially derived protein. Second, free sulphydryls are not directly participating in catalysis. It is possible, however, that a few of these cysteines are buried and are therefore not accessible to sulphydryl exchange. At least five of the eight cysteines, however, do react with DTNB under nondenaturating conditions (data not shown). This latter fact may suggest an alternate role for these solvent-accessible cysteines (along with specific histidines), perhaps as metal-coordinating thiolates. This possibility is a reasonable one when one makes the comparison between the 2-O-sulfatase and alkaline phosphatase. Both of these enzymes are esterases with similar catalytic mechanisms, including the presumptive formation of a covalent intermediate. The two hydrolytic enzymes also possess structurally related domains, in particular a highly superimposable active site that includes a divalent metal binding pocket. In the case of alkaline phosphatase, it is zinc rather than calcium (or Mg2+) that is coordinated within this pocket. 

In the previous paper, we described the molecular cloning of a HSGAG 2-O-sulfatase from F. heparinum and its recombinant expression in E. coli as a soluble, highly active enzyme in milligram quantities. We also presented a characterization of the enzyme’s biochemical properties as it pertains to optimal in vitro reaction conditions and substrate kinetics. In this paper, we build upon this biochemical characterization of the enzyme by providing an invaluable structural framework to address the molecular basis of enzyme function. To the best of our knowledge, this paper represents one of the first structural descriptions of an HSGAG desulfatizing enzyme. As stated before, our homology model of the 2-O-sulfatase has already provided valuable insight into likely structural determinants for substrate specificity and catalysis. Many of these insights are validated by our biochemical and kinetic evaluation of this specificity. Importantly, we now have in place a meaningful structural framework in which to address additional questions pertaining to oligosaccharide-sulfatase interactions.  

REFERENCES
1. Lindahl, U., Kusche-Gullberg, M., and Kjellen, L. (1998) J. Biol. Chem. 273, 24797–24802
2. Eko, J. D., and Lindahl, U. (2001) J. Clin. Invest. 108, 169–173
3. Parenti, G., Meroni, G., and Gallabio, A. (1997) Curr. Opin. Genet. Dev. 7, 386–391
4. von Figura, K., Schmidt, B., Selmer, T., and Dierks, T. (1998) Bioessays 20, 505–510
5. Waldow, A., Schmidt, B., Dierks, T., von Bulow, R., and von Figura, K. (1999) J. Biol. Chem. 274, 12284–12288
6. Schmidt, B., Selmer, T., Ingendoh, A., and von Figura, K. (1995) Cell 82, 271–278
7. Dierks, T., Miech, C., Hummerjohann, J., Schmidt, B., Kertesz, M. A., and von Figura, K. (1998) J. Biol. Chem. 273, 25560–25564
8. Kolodny, E. H. and Piskitny, A. I. (1995) in The Metabolic and Molecular Bases of Inherited Disease (Scrivener, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds) pp. 2693–2741, McGraw-Hill Inc., New York
9. Lukatela, G., Krauss, N., Theis, K., Selmer, T., Gieselman, V., von Figura, K., and Senger, W. (1998) Biochemistry 37, 3654–3664
10. von Bulow, R., Schmidt, B., Dierks, T., von Figura, K., and Uson, I. (2001) J. Mol. Biol. 305, 269–277
11. Bond, C. S., Clements, P. R., Ashby, S. J., Collyer, C. A., Harrop, S. J., Hopwood, J. J., and Guss, J. M. (1997) Structure 5, 277–289
12. Boltes, I., Czapska, H., Krahnt, A., von Bulow, R., Dierks, T., Schmidt, B., von Figura, K., Kertesz, M. A., and Uson, I. (2001) Structure (Camb.) 9, 483–491
13. Tsuda, T., Hileman, R. E., Smith, A. E., Vlahova, P. I., and Linhardt, R. J. (1996) J. Biol. Chem. 271, 20240–20247
14. Higgins, D. G., Thompson, J. D., and Gibson, T. J. (1996) Methods Enzymol. 266, 383–402
15. Homann, S. W. (1999) Biochemistry 29, 9110–9118
16. Huiue, C. J. M. and Altona, C. (1995) J. Comput. Chem. 16, 56–79
17. Rhomberg, A. J., Ernst, S., Sasisekharan, R., and Biemann, K. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 4716–4718
18. Myette, J. R., Shriver, Z., Liu, J., Venkataraman, G., Rosenberg, B., and Sasisekharan, R. (2002) Biochem. Biophys. Res. Commun. 290, 1206–1213
19. Riddles, P. W., Blakeley, R. L., and Zerner, B. (1983) Methods Enzymol. 91, 118–130
20. Shindyalov, I. N., and Bourne, P. E. (1998) Protein Eng. 11, 739–747
21. Myette, J. R., Shriver, Z., Claycamp, C., McLean, M. W., Venkataraman, G., and Sasisekharan, R. (2003) J. Biol. Chem. 278, 12157–12166
22. Venkataraman, G., Shriver, Z., Raman, R., and Sasisekharan, R. (1999) Science 286, 537–542