Characterization and Mutagenesis of Gal/GlcNAc-6-*O*-sulfotransferases†

Jocelyn R. Grunwell,‡ Virginia L. Rath,§ Jytte Rasmussen,|| Zeljka Cabrilo,|| and Carolyn R. Bertozzi*,‡,||,⊥

Departments of Chemistry and Molecular and Cell Biology and Howard Hughes Medical Institute, University of California, Berkeley, California 94720, and Thios Pharmaceuticals, Inc., 5980 Horton Street, Emeryville, California 94608

Received October 7, 2002

ABSTRACT: The installation of sulfate groups on the carbohydrate residues of glycoproteins, glycolipids, and glycosaminoglycans is a critical posttranslational modification that occurs in all higher eukaryotes. The Gal/GalNAc/GlcNAc-6-*O*-sulfotransferases (GSTs) are a recently discovered family of carbohydrate sulfotransferases that share significant sequence homology at the amino acid level and mediate a number of different biological processes such as leukocyte adhesion at sites of chronic inflammation. Structural and mechanistic studies of this family of sulfotransferases have been hindered by the lack of a productive recombinant protein expression system. We developed a baculovirus expression system for five of the seven cloned GSTs and determined their kinetic parameters using both thin-layer chromatography and a recently developed polymer dot-blot assay. We used these tools to perform the first site-directed mutagenesis study of a member of this sulfotransferase family, GST2. Using sequence alignments with other carbohydrate and cytosolic sulfotransferases, we selected residues within the putative binding regions for 3′-phosphoadenosine 5′-phosphosulfate (PAPS) and the carbohydrate substrate for mutagenesis. Kinetic analysis of the mutants identified residues that are essential for catalytic activity. These results should facilitate mechanistic studies and the development of small molecule inhibitors of this enzyme family to ameliorate chronic inflammatory diseases.

Sulfates on the carbohydrate residues of glycoproteins, glycolipids, and glycosaminoglycans (GAGs)[1] are critical posttranslational modifications that modulate a variety of biological responses (reviewed in refs *1*−*3*) including viral and bacterial invasion of a host cell (*4*−*11*), cell−cell adhesion (*2*, *3*, *12*−*15*), glycoprotein hormone activity (*16*−*18*), blood clotting (*19*−*27*), and modulation of cytokine and growth factor activity (*28*−*32*). For example, glycoproteins displaying the L-selectin binding epitope 6-sulfo sialyl Lewis x (sLe^x) are instrumental in recruiting leukocytes to sites of chronic inflammation and are essential components of the lymphocyte recirculation system (*33*−*38*). The dysregulation of leukocyte homing to nonlymphoid tissue leads to chronic diseases such as type I diabetes, psoriasis, asthma, and rheumatoid arthritis (*39*). The undersulfation of GAG chains can lead to skeletal and connective tissue deficiencies that result in severe structural deformities (*40*−*43*). More specifically, the lack of sulfation on the 6-position of GlcNAc residues within corneal keratan sulfate results in opacity of the cornea and leads to eventual blindness in patients suffering from macular corneal dystrophy (MCD) (*44*−*49*).

There are two major classes of sulfotransferases (STs): the cytosolic and the Golgi-resident STs. The cytosolic STs sulfate small molecules such as the hormones estrogen and dihydroxyepiandosterone (DHEA) and phenol-based xenobiotics found in the environment. The membrane-bound Golgi-resident STs catalyze the transfer of sulfate to the hydroxyl or amino groups of sugars on glycoproteins, glycolipids, and GAGs (*1*) and to the tyrosine residues of membrane proteins such as P-selectin glycoprotein ligand-1

† Financial support for this work was provided by the National Institutes of Health (Grant GM59907).

* To whom correspondence should be addressed at the Department of Chemistry, University of California, Berkeley, Berkeley, CA 94720. Tel: (510) 643-1682. Fax: (510) 643-2628. E-mail: bertozzi@cchem.berkeley.edu.

‡ Department of Chemistry, University of California.

§ Thios Pharmaceuticals, Inc.

|| Howard Hughes Medical Institute, University of California.

⊥ Department of Molecular and Cell Biology, University of California.

[1] Abbreviations: GAG, glycosaminoglycans; sLe^x, sialyl Lewis x; GlcNAc, *N*-acetylglucosamine; MCD, macular corneal dystrophy; ST, sulfotransferase; DHEA, dihydroxyepiandosterone; PSGL-1, P-selectin glycoprotein ligand-1; GST, *N*-acetylgalactosamine/galactose/*N*-acetylglucosamine-6-*O*-sulfotransferase; HEC, high endothelial cell; PAPS, 3′-phosphoadenosine 5′-phosphosulfate; 5′-PSB, 5′-phosphosulfate binding site; SBS, substrate binding site; 3′-PB, 3′-phosphate binding site; NDST-1, the sulfotransferase domain of human *N*-deacetylase/*N*-sulfotransferase-1; HNK-1ST, human natural killer cell-1 sulfotransferase; PCR, polymerase chain reaction; FBS, fetal bovine serum; HRP, horseradish peroxidase; TLC, thin-layer chromatography; pen/strep, penicillin/streptomycin; Kan, kanamycin; Amp, ampicillin; BSA, bovine serum albumin; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; CV, column volumes; *β*BnO-GlcNAc, *β*-*O*-benzyl-*N*-acetylglucosamine; KS, keratan sulfate; Hepes, 2-[4-(2-hydroxyethyl)-1-piperazinyl]ethanesulfonic acid; MOPS, 3-(*N*-morpholino)propanesulfonic acid; PI, postinfection; PVDF, polyvinylidene fluoride; GlcNAc-PAA, *N*-acetylglucosamine polyacrylamide polymer; mEST, mouse estrogen sulfotransferase; hEST, human estrogen sulfotransferase; hSULT1A1, human phenol sulfotransferase; hSULT1A3, human monoamine sulfotransferase; F3ST, flavonol 3-sulfotransferase; hHST, human hydroxysteroid sulfotransferase; rat ST-40, rat sulfotransferase-40; GalCerST, galactosylceramide sulfotransferase; WT, wild type; PSIPRED, protein secondary structure prediction based on position-specific scoring matrices; PAP, 3′-phosphoadenosine 5′-phosphate; APS kinase, adenosine 5′-phosphosulfate kinase; TMD, transmembrane domain.

10.1021/bi0269557 CCC: $22.00
Published on Web 12/07/2002

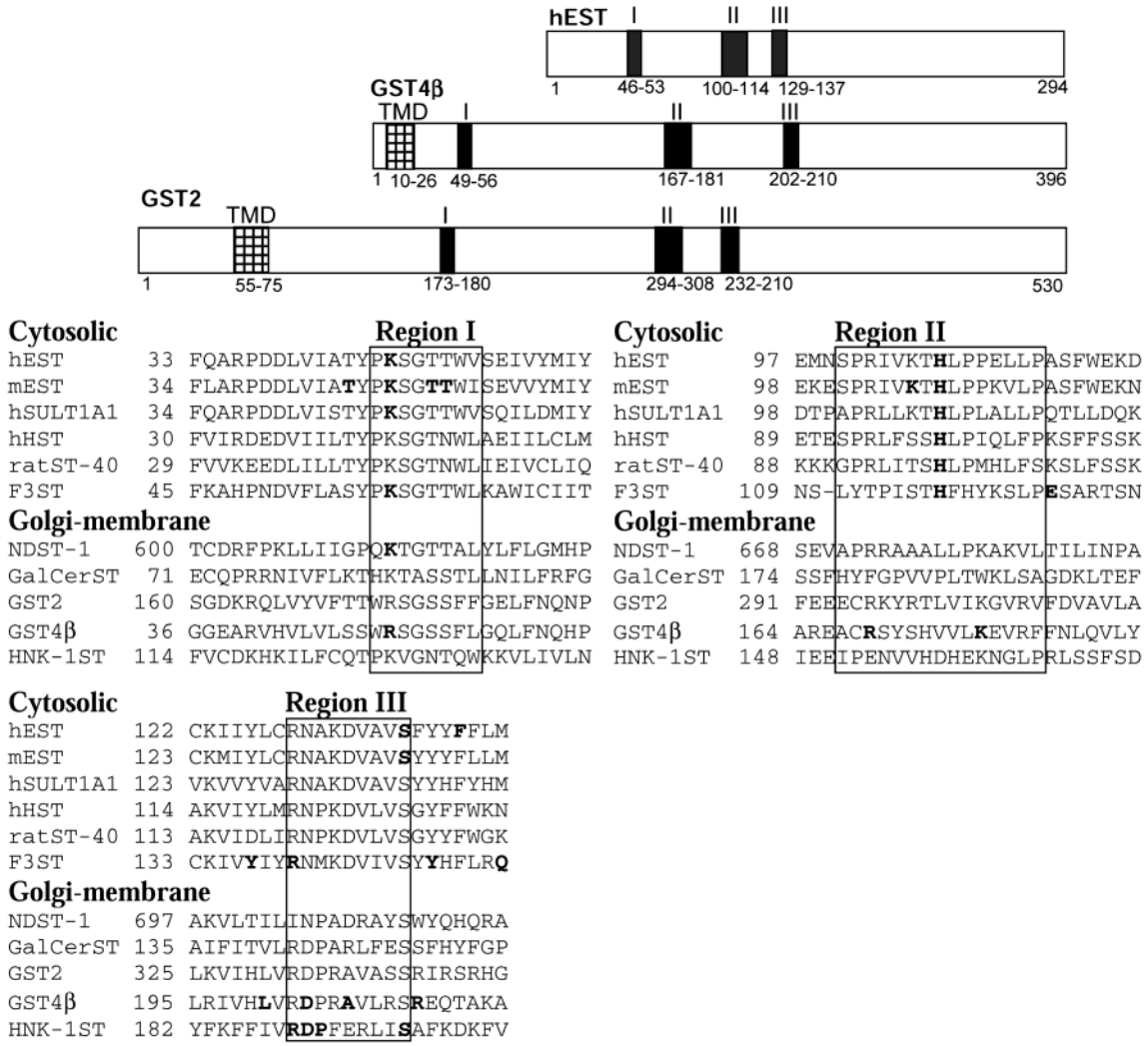

FIGURE 1: Sequence alignments of several cytosolic and carbohydrate STs. High homology exists between the cytosolic and carbohydrate STs within region I, the 5′-phosphosulfate binding site (5′-PSB), and region III, the 3′-phosphate binding site (3′-PB). GST1, GST2, GST3, GST4α, and GST5 were expressed in this study. GST4*β* is shown in this alignment because mutations within this enzyme result in a patient population suffering from MCD as discussed in the text. Residues in bold have been mutated and affect enzymatic activity. Within region I, the cytosolic STs have a conserved Lys residue that is replaced by an Arg residue within the carbohydrate STs. Of note is the RDP motif in region III of the carbohydrate STs indicated in boldface type for HNK-1ST. A Ser residue defines the C-terminus of this region for both the cytosolic and carbohydrate STs. Region II is defined as the substrate binding site (SBS), and as expected due to the differences between the hydrophobic steroid substrates of the cytosolic STs and the hydrophilic sugar substrates of the carbohydrate STs, there is very little homology within this region for the two classes of STs. Above the sequence alignments is a scaled scheme of regions I, II, and III within the context of the length of the ST protein. Abbreviations: TMD, transmembrane domain; hEST, human estrogen ST; mEST, mouse estrogen ST; hSULT1A1, human phenol ST; hHST, human hydroxysteroid ST; rat ST-40, rat aryl ST-40; F3ST, flavonol 3-ST; NDST-1, the sulfotransferase domain of heparan sulfate *N*-deacetylase/*N*-sulfotransferase-1; GST2, *N*-acetylglucosamine-6-*O*-ST-2; GST4*β*, *N*-acetylglucosamine-6-*O*-ST-4*β*; HNK-1ST, human natural killer cell-1ST; GalCerST, galactosylceramide ST.

(PSGL-1) (*50*). The enzymes in the "GalNAc/Gal/GlcNAc ST" (GST) family are responsible for sulfation of the 6-position of *N*-acetylgalactosamine, galactose, or *N*-acetylglucosamine residues. Briefly, GST0 (also termed C6ST or CHST) (*51*, *52*) and GST1 (also termed KSGal6ST, C6ST, or CHST-1) sulfate GalNAc residues in chondroitin sulfate and the Gal residues in keratan sulfate, respectively (*53*−*55*). GST2 (also termed GlcNAc6ST-1 or CHST-2) (*53*, *56*), GST3 (also termed HEC-GlcNAc6ST, LSST, or GlcNAc6ST-2) (*57*), GST4α/*β* (also termed GlcNAc6ST-3, I-GlcNAc6ST/GlcNAc6ST-5, or C-GlcNAc6ST) (*58*, *59*), and GST5 (also termed GlcNAc6ST-4 or C6ST-2) (*60*−*62*) transfer a sulfate group to nonreducing GlcNAc residues to create a variety of cell surface epitopes, including the L-selectin binding epitope 6-sulfo sLe^x. Of these seven carbohydrate 6-*O*-sulfotransferases, three have highly restricted tissue distributions. GST3, GST4α, and GST4*β* are expressed primarily in peripheral lymph node high endothelial cells (HEC) (*57*), within the intestines (*58*), and within the cornea (*59*), respectively. MCD patients have mutations within the gene encoding GST4*β* (*44*, *45*, *48*).

The Golgi-resident STs are all single span type II transmembrane proteins with short N-terminal cytosolic tails and C-terminal catalytic domains that reside in the Golgi lumen (Figure 1). While the cytosolic and Golgi STs differ in sequence characteristics and sulfate a diverse array of substrates, they have in common the requirement of the activated sulfate donor 3′-phosphoadenosine 5′-phosphosulfate (PAPS). Two regions of homology exist within the cytosolic and the Golgi-resident STs: a 5′-phosphosulfate

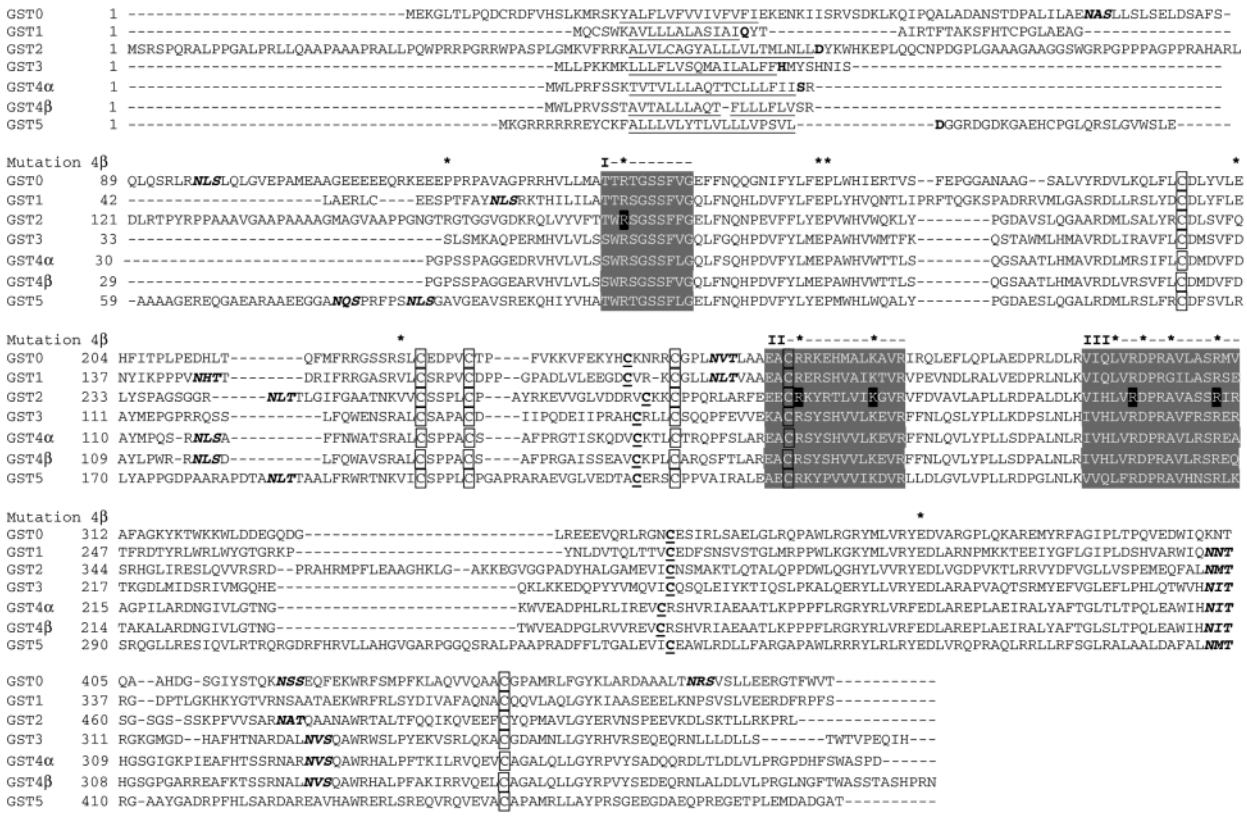

FIGURE 2: Sequence alignment of the seven cloned GSTs generated using ClustalW (*93*) and BoxShade. The transmembrane domains (TMDs) are underlined. Sites of potential N-linked glycosylation are in bold italics. Regions of high homology are indicated by overscoring, where region I refers to the 5′-phosphosulfate binding site, region II is the putative carbohydrate binding site, and region III is the 3′-phosphate binding site. Asterisks above the sequences correspond to mutations found in the GST4*β* gene in patients with MCD. Conserved cysteines are boxed. Additional cysteines, found in similar regions, are underlined and in bold. Mutations made in this study to the GST2 sequence are highlighted by a black box with white text. The residue with which the soluble secreted constructs begin is indicated in bold following the underlined TMD sequence.

binding site (5′-PSB, region I) and a 3′-phosphate binding site (3′-PB, region III) (Figure 1). As expected, a putative substrate binding site (SBS, region II) is very different between these two classes. Residues within this region of cytosolic STs have been shown to interact with their hydrophobic substrates. A high level of conservation in region II of the GSTs suggests those residues also contribute to substrate binding (Figure 2). Mutations found in MCD patients are found in all three regions.

A substantial amount of work has been published characterizing the structure and function of the cytosolic sulfotransferases by crystallography (*63*−*69*), mutagenesis (*70*−*76*), photoaffinity labeling (*73*, *74*), and kinetic analysis (*77*−*82*). In contrast, information on the membrane-bound Golgi-resident carbohydrate sulfotransferases (including the GSTs) is lacking, in part due to the difficulty in overexpression of these proteins. The mammalian expression systems currently employed for GSTs, such as transient expression in COS-7 cells, suffer from poor yields and lack of purity, and as a result, clarified conditioned media (*83*), beads with immobilized protein (*84*), or homogenized cells (*85*) have been used for in vitro assays.

In this study, we report the development of a productive baculovirus expression system for soluble versions of five of the seven cloned enzymes within the GST family. We optimized the in vitro reaction conditions for each enzyme and characterized the first- and second-order rate constants for the substrate PAPS. In the first site-directed mutagenesis studies of GST2, we used this system to define residues required for catalytic activity. Five mutations were made at basic residues (Arg or Lys) predicted to be within the 5′-PSB, the SBS, and the 3′-PB regions on the basis of sequence alignment with the human cytosolic STs, the sulfotransferase domain of human heparan sulfate *N*-deacetylase/*N*-sulfotransferase-1 (NDST-1) (*86*), and human natural killer cell-1 sulfotransferase (HNK-1ST) (*18*). Mutations were made at sites that correspond to known polymorphisms found in GST4*β* of MCD patients. Several mutations decreased or eliminated catalytic activity, suggesting a direct involvement in catalytic function.

EXPERIMENTAL PROCEDURES

Construction of a His$_6$-Tagged Secretion Vector for Use in the Bac-to-Bac Expression System. To generate soluble, secreted versions of the enzymes, we replaced the N-terminal cytosolic tail and transmembrane domain with a hexahistidine (His_6) tag and a honeybee melittin signal sequence (Sigma-Genosys). The residues that initiate the truncated sequences are Gln19 for GST1, Asp75 for GST2, His25 for GST3, Ser28 for GST4α, and Asp33 for GST5 (Figure 2). The sense primer sequence was 5′-GATCCATGAAATTCTTAGT-CAACGTTGCCCTTGTTTTTATGGTCGTCTACATTT-CTTACATCTATGCGGATCCAAGCCCGG-3′, and the antisense primer sequence was 5′-CGCGCCGGGCTTGG-ATCCGCATAGATGTAAGAAATGTAGACGACCATAAAA

ACAAGGGCAACGTTGACTAAGAATTTCATG-3′. The primers were annealed and ligated into the *Bam*HI and *Bss*HII restriction sites of a digested pFASTBac vector (Invitrogen). Ligation of the secretion signal sequence into the vector was confirmed by polymerase chain reaction (PCR) using the pFASTBac forward sequencing primer. The sequence encoding for a His_6-tagged protein with a thrombin protease cleavage site was cloned in-frame with the secretion signal using the *Bss*HII and *Eco*RI sites for GST1, GST2, GST4α, and GST5 and the *Bss*HII and *Sal*I sites for GST3. The sense primer used was 5′-CGCGCCATCATCACCAC-CATCACTCCGCGGGTTTGGTGCCCAGAGGCAGCG-3′, and the antisense primer was either 5′-AATTCGCTGC-CTCTGGGCACCAAAACCGCGGAGTGATGGTGGTG-ATGATGG-3′ for the *Eco*RI site or 5′-TCGACGCTGC-CTCTGGGCACCAAAACCGCGGAGTGATGGTGGTG-ATGATGG-3′ for the *Sal*I cloning site. The final donor plasmid construct was confirmed by sequencing (Davis Sequencing, Davis, CA).

Cloning of the Sulfotransferases. The genes for the GSTs were subcloned into the pFASTBac donor vector using PCR with template DNA originating from the cDNA cloned into mammalian expression vectors (a generous gift from Professor S. D. Rosen, UCSF). The following primers were used to amplify the GST genes and introduce the restriction sites *Eco*RI and *Kpn*I for GST1, GST2, GST4α, and GST5 or *Sal*I and *Kpn*I for GST3. GST1 primers were 5′-CCG-GAATTCATGCAGTACACGGCCATCCGCACC-3′ and 5′-CGGGGTACCTCACGAGAAGGGGCGGAAGTC-3′. GST2 primers were 5′-CCGGAATTCATGGACTACAAGTGGC-ATAAGGAG-3′ and 5′-CGGGGTACCTTAGAGACGGGG-CTTCCGAAG-3′. GST3 primers were 5′-ACGCGTCGA-CATGCACATGTACAGCCACACTC-3′ and 5′-CGGGG-TACCTTAGTGGATTTGCTCAGGGAC-3′. GST4α primers were 5′-CCGGAATTCATGTCCCGGCCAGGGCCC-TCATCC-3′ and 5′-CGGGGTACCTCAGTCAGGCGAT-GCCCAGCT-3′. GST5 primers were 5′-CCGGAATTCATG-GACGGCGGCCGCGACGGGGAC-3′ and 5′-CGGGGTA-CCCTACGTGGCGCCGTCGGCATC-3′. PCR reactions were performed in an MJ Research thermocycler (South San Francisco, CA) by adding the following components to thin-walled PCR tubes: 1 μL of a 10 μM stock solution of forward primer in sterile-filtered H_2O, 1 μL of a 10 μM stock solution of reverse primer in sterile H_2O, 5 μL of 10× *Pfu* buffer (Stratagene), 1 μL of template DNA, 2.5 μL of DMSO (5% final concentration), 1 μL of dNTPs (10 mM each), 1 μL of *Pfu* polymerase (Stratagene), and water to achieve a final reaction volume of 50 μL. The reactions were first heated to 94 °C for 3 min. One cycle consisted of 94 °C for 45 s and 45 °C for 45 s, followed by 72 °C for 3 min, for 30 cycles. Finally, the reactions were incubated at 72 °C for 5 min. The desired PCR products were gel purified (Qiagen).

The GST genes were ligated into the digested pFASTBac donor vector using T4 DNA ligase (NEB). Several colonies were used to inoculate 3 mL of LB media containing 100 μg/mL ampicillin (Amp) and grown overnight. Plasmid was isolated (Qiagen), and positive clones were identified either by PCR as described above or by restriction digest and confirmed by sequencing (Davis Sequencing).

Site-Directed Mutagenesis. Mutations were introduced into the GST2 gene using the QuickChange mutagenesis protocol (Stratagene) and gel-purified primers (Sigma-Genosys). The sequences of the primers for each mutant are listed below, where the forward primer precedes the reverse primer. Oligonucleotides used: for R174A, 5′-GGTGTACGTG-TTCACCACGTGGgcCTCTGGaTCcTCGTTCTTCGGCGA-GCTATTC-3′ and 5′-GAATAGCTCGCCGAAGAACGAg-GAtCCAGAGgcCCACGTGGTGAACACGTACACC-3′; for R296A, 5′-CGCGTTTCGAGGAGGAGTGCgcCAAGTAC-CGCACcCTAGTCATAAAGGGTGTGCG-3′ and 5′-CGCA-CACCCTTTATGACTAGgGTGCGGTACTTGgcGCACTCCT-CCTCGAAACGCG-3′; for K304A, 5′-GGAGTGCCGC-AAGTACCGCACcCTAGTCATAgcGGGTGTGCGCGTC-TTCGACG-3′ and 5′-CGTCGAAGACGCGCACACCCgc-TATGACTAGgGTGCGGTACTTGCGGCACTCC-3′; for R332A, 5′-CAAGGTCATCCACTTGGTGgcTGATCCtC-GaGCGGTGGCGAGTTCACGGATCC-3′ and 5′-GGATC-CGTGAACTCGCCACCGCtCGaGGATCAgcCACCAA-GTGGATGACCTTG-3′; and for R341A, 5′-CCCGCGC-GGTGGCGAGTTCAgcGATCCGCTCGCGCCACGG-CC-3′ and 5′-GGCCGTGGCGCGAGCGGATCgcTGAAC-TCGCCACCGCGCGGG-3′, where the lower case letters indicate the altered nucleotides. PCR was performed using 5 μL of 10× *Pfu* reaction buffer, 2.5 μL of DMSO (5%), 0.3 μL of vector (173 ng/μL), 1.25 μL of the forward primer (100 ng/μL), 1.25 μL of the reverse primer (100 ng/μL), 1 μL of dNTP (10 mM), 1 μL of *Pfu* polymerase (2.5 units/μL) (Stratagene), and sterile water to a total volume of 50 μL. The sample was heated to 95 °C for 30 s, and 18 PCR cycles consisting of heating the sample to 95 °C for 30 s, 45 °C for 1 min, and 68 °C for 15 min were performed. XL-10 Gold Ultracompetent cells (100 μL) were transformed with 4 μL of the *Dpn*I-digested PCR reactions (to remove remaining wild-type vector) according to the manufacturer's protocol (Stratagene). Plasmid isolated from positive clones was confirmed by sequencing (Davis Sequencing).

Generation of Recombinant Bacmid DNA. DH10Bac cells (100 μL) were transformed by heat shock with 5 μL of recombinant pFASTBac donor plasmid containing the GST wild-type and mutant genes, and the bacmid DNA was isolated according to the manufacturer's protocol (Invitrogen). The insertion of the GST gene into the bacmid DNA was confirmed by PCR using the M13/pUC forward and reverse primers flanking the transposition target sequences in the bacmid.

Transfection of Bacmid DNA into Sf9 Cells. Sf9 cells were transfected with bacmid DNA using CellFECTIN reagent (Gibco) according to the manufacturer (Invitrogen). Three days posttransfection, when the cells began to swell and detach from the plate, the virus-containing media were harvested and clarified by centrifuging at 500*g* for 5 min.

Amplification of Recombinant Virus. Virus-containing media (0.5 mL) were added to a 10 cm plate seeded with 2 × 10^6 cells/mL in 11.5 mL of complete media [Excell-400, 2% fetal bovine serum (FBS), 50 units/mL penicillin (base), and 50 μg/mL streptomycin (Na salt) (pen/strep, Invitrogen)]. Cells were incubated at 27 °C under an upside-down tray with damp paper towels for 5 days to generate a P1 stock. P1 virus stock (10 mL) was harvested from the plates, clarified by centrifugation, filtered through a 0.2 μm filter, aliquoted, and stored at −20 °C. P2 stock was generated by infecting 100 mL of Sf9 cells at a cell density of 1 × 10^6 cells/mL in Excell-400 serum-free media containing pen/strep with 1 mL of filtered P1 stock. The cells were incubated

Table 1: Day Postinfection (PI) at Which the Maximal Protein Is Produced[a]

GST	1	2	3	4α	5
day of maximal protein yield	6	2 or 3	6	3	2

[a] High-Five cells were infected with enough P2 viral stock to halt cell growth to <25% within 24 h PI. Time points were taken each day, and equal amounts of conditioned media were probed by Western blot with an anti-pentaHis-HRP conjugated antibody and a chemiluminescent substrate.

at 27 °C for 4−6 days and centrifuged, and the supernatant was filtered through a 0.2 μm unit. The P2 stock was supplemented with 2% FBS, aliquoted, and stored at −70 °C. Viral stock was amplified using 100 mL of Sf9 cultures as described to generate P2 stock until a viral stock of sufficient titer was obtained to arrest cell growth within 24 h postinfection (PI). Supernatant (10 μL) was used to detect protein production by activity assay and Western blot using horseradish peroxidase (HRP) conjugated anti-pentaHis antibody (Qiagen).

Expression of Recombinant Sulfotransferases in Baculovirus-Infected Cells. P2 viral stock solution (10 mL) was added to a 1 L culture of High-Five cells at ~1 × 10^6 cells/mL. Cells were incubated at 27 °C with shaking for the optimal number of days for maximal protein production (Table 1) as determined by performing Western blots of conditioned media at various days PI. Cells were centrifuged at 6000*g* for 15 min at 4 °C and filtered through a 0.2 μm filter unit, and the filtrate was directly loaded onto a 5 mL HiTrap chelating column (AP Biotech) equilibrated with 50 mM sodium phosphate, pH 7.4, containing 500 mM NaCl and 10 mM imidazole. The column was washed with equilibration buffer, and eluted using a 0−60% linear gradient of a solution of 50 mM sodium phosphate, pH 7.4, containing 500 mM NaCl and 500 mM imidazole, followed by a step of 500 mM imidazole in buffer. Fractions (5 mL) containing GST were identified by Western blot using an anti-pentaHis-HRP antibody.

Synthesis of ^{35}S-PAPS. Radiolabeled PAPS was synthesized by incubating 2.5 mCi of $^{35}SO_4$ with 130 mM ATP, 20 units of inorganic pyrophosphatase (60% *Escherichia coli*, Sigma), 2 units of ATP sulfurylase (Sigma), and 7 milliunits of APS kinase in PAPS synthesis buffer (50 mM Tris-HCl, pH 8.0, 30 mM KCl, 40 mM $MgCl_2$, 1 mM EDTA, and 1 mM DTT) in a total volume of 300 μL at room temperature for 3 h. The amount of $^{35}SO_4$ incorporated into ^{35}S-PAPS was determined by PEI−cellulose thin-layer chromatography (TLC) analysis using 0.9 M LiCl as eluent. Yields were typically better than 90%.

Cloning and Expression of Recombinant APS Kinase. The gene encoding APS kinase was obtained from *E. coli* DH5α genomic DNA by PCR amplification (using *Pfu* polymerase). The forward primer had the sequence 5′-CGCGGATCCATGGCGCTGCATGACGAAAACGTCG-3′ with the *Bam*HI restriction site. The reverse primer had the sequence 5′-CCGCCGCTCGAGGGATCTGATAATATCGTTCTGTCTC-3′ with the *Xho*I restriction site. The PCR product and the expression vector, pET24b(+) (Novagen), were purified and digested with the restriction enzymes *Bam*HI and *Xho*I and ligated with T4 DNA. The resulting plasmid was isolated using a mini-prep kit (Qiagen) and digested with *Bam*HI and *Xho*I to confirm insertion of the APS kinase gene. Two positive plasmids were sequenced (Elim Biopharma, Hayward, CA). BL21(DE3) cells were transformed with the recombinant pET24b vector and plated onto LB plates containing 50 μg/mL kanamycin (Kan).

Table 2: Optimization of the Sulfotransferase Reaction Conditions

GST	pH	buffer (50 mM)	[$Mg(OAc)_2$] (mM)	[ATP] (mM)
1	7.5	Bis-Tris propane	5	0.1
2	7.0	Hepes	5	1
3	7.0	Hepes	5	0.5
4α	7.0	Bis-Tris propane	2	0.05
5	7.0	Bis-Tris propane	2	0.1

A single colony was used to inoculate 10 mL of LB/Kan which served as the inoculant for 3 L of LB/Kan media. The culture was grown to an OD_{600} of 1.0, induced with 0.5 mM IPTG, and incubated at 27 °C for 20 h. The cells were collected by centrifugation and lysed (BugBuster, Novagen). The filtered lysate was loaded onto a 1 mL HiTrap chelating column (Amersham), washed with buffer 1 (50 mM NaH_2PO_4, pH 8, 300 mM NaCl, 10% glycerol, 5 mM 2-mercaptoethanol), followed by buffer 2 (50 mM NaH_2PO_4, pH 6, 300 mM NaCl, 10% glycerol, 5 mM 2-mercaptoethanol), and eluted using a linear gradient of 0−100% elution buffer (50 mM NaH_2PO_4, pH 8, 300 mM NaCl, 10% glycerol, 5 mM 2-mercaptoethanol, 0.5 M imidazole). Fractions (1 mL) were collected, and protein was visualized by Coomassie Blue staining and confirmed by Western blot using an HRP-conjugated anti-pentaHis antibody (Qiagen). The peak was pooled (typically fractions 5−13), concentrated in a Centriprep YM-30 filter unit (Millipore), dialyzed against buffer 1, flash frozen, and stored at −80 °C.

APS Kinase Activity. The activity of recombinant APS kinase was tested using the pyruvate kinase−lactate dehydrogenase coupled assay of Burnell and Whatley (*87*). The activity of the recombinant APS kinase was 0.35 unit/mL (1 unit = 1 μmol/min).

Sulfotransferase Assays. GST activity was measured by both the TLC (*88*) and the dot-blot assay (*89*). In the TLC assay, the carbohydrate substrate used was β-*O*-benzyl-*N*-acetylglucosamine (βBnO-GlcNAc), while for the dot-blot assay, the substrate was a GlcNAc-conjugated polyacrylamide polymer (GlcNAc-PAA; Glycotech, Rockville, MD). GST1 activity was measured using this enzyme's native substrate keratan sulfate (KS) (Sigma). Reactions were performed in a total volume of 25 μL using the optimized reaction conditions (Table 2). The final concentrations of the components of the reaction were 50 mM Bis-Tris propane (or Hepes), 10 mM NaF, pH 7.0, 2 or 5 mM $Mg(OAc)_2$, 0.05−1 mM ATP, 0.1% Triton X-100, 10% glycerol, 1−4 μL of enzyme (~5−15 ng/μL), 2 μM PAPS, 0.2 μCi of ^{35}S-PAPS, and either 250 ng of KS, 62.5 ng of GlcNAc-PAA, or a 1 mM solution of βBnO-GlcNAc. Reactions were run at room temperature to less than 15% completion (typically 1−4 h). The following buffers were tested at pH 7.0: 50 mM MOPS, 50 mM Hepes, 25 mM NaH_2PO_3, 25 mM sodium citrate, and 25 mM Bis-Tris propane/25 mM glycine, each with 1 mM ATP and 4 mM $Mg(OAc)_2$. The optimal pH was determined by performing the sulfotransferase reactions in a 25 mM Bis-Tris propane/25 mM glycine buffer

with a pH ranging from 5 to 9 in 0.5 unit increments and 50 mM Hepes/10 mM NaF or a 50 mM Bis-Tris propane/10 mM NaF buffer at pH 6.5, 7.0, and 7.5 in the presence of 1 mM ATP and 4 mM $Mg(OAc)_2$. $Mg(OAc)_2$ levels were optimized by performing the assays in the optimal buffer and pH for that enzyme at 1 mM ATP with $Mg(OAc)_2$ concentrations of 0, 0.01, 0.05, 0.1, 0.5, 1, 2, 5, and 10 mM. ATP levels were optimized by performing the assays in the optimal buffer and pH for that enzyme at 4 mM $Mg(OAc)_2$ and ATP concentrations of 0, 0.01, 0.05, 0.1, 0.5, 1, 5, and 10 mM.

K_M values for PAPS were determined using 250 ng of KS, 62.5 ng of GlcNAc-PAA or a 1 mM solution of βBnO-GlcNAc, and PAPS concentrations of 0.25, 0.5, 0.75, 1, 2, 3, and 5 μM for GST1; 0.5, 1, 2, 3, 5, 7, 10, and 15 μM for GST2; 0.1, 0.2, 0.4, 0.6, 0.8, 1.5, 3, and 5 μM for GST3; 5, 10, 15, 20, 25, 30, and 120 μM for GST4α; and 0.3, 0.5, 1, 2, 3, 5, 7, 12, and 20 μM for GST5. All of the PAPS K_M data were obtained using the dot-blot assay except for GST4α, in which case the TLC assay was used with βBnO-GlcNAc as a substrate. The concentration of PAPS was fixed at 3.15 μM, and the concentrations of βBnO-GlcNAc were 4.52, 2.26, 1.13, 0.56, and 0.28 mM. GST2 mutant R296A also required βBnO-GlcNAc as a substrate, and concentrations of 45.2, 22.6, 11.3, 5.65, 2.83, and 1.41 mM were used in the assay. The molecular masses of the proteins used to calculate k_{cat} values are as follows: GST1, 47480 Da; GST2, 52420 Da; GST3, 44959 Da; GST4α, 43420 Da; and GST5, 53114 Da.

TLC Assay Analysis. Reactions processed by the TLC method were quenched with 25 μL of methanol, and 20 μL of the quenched reaction was spotted onto a silica gel TLC plate. The samples were eluted in a solvent mixture of 8:2:1 1-butanol−ethanol−water at 4 °C in a closed TLC chamber, air-dried, and analyzed on a phosphorimager (Storm, Molecular Dynamics). The amount of product produced was calculated according to the equation:

$$\left(\frac{\text{density of product}}{\text{density of total}}\right) \times \left(\frac{[\text{PAPS}]}{\text{time of reaction}}\right) \qquad (1)$$

Dot-Blot Assay Workup and Analysis. Reactions to be processed by the dot-blot method were quenched with 100 μL of methanol, and 25 μL of the quenched sample was spotted onto a wet membrane (water for the Hybond N and methanol for the PVDF membranes) in a 96-well plate vacuum apparatus (Bio-Rad). Each of the wells was washed twice with 50 μL of 1.25 M $(NH_4)_2SO_4$, pH 3.0. The membrane was removed from the vacuum apparatus, washed twice in 1.25 M $(NH_4)_2SO_4$, pH 3.0, for 30 s, and rinsed with water for 30 s. The membrane was blotted onto a paper towel, air-dried, wrapped in cellophane, and analyzed with a phosphorimager. To relate the measured density to the absolute amount of radioactive product, a standard curve was generated as follows. From a stock solution of PAPS (typically 200 μM PAPS, ~1 μCi/μL ^{35}S-PAPS, specific activity = 4128 Ci/mmol), 1 μL was quantified by scintillation counting and another 1 μL spotted onto PEI−cellulose for analysis on the phosphorimager. The counts per minute were converted into disintegrations per minute (dpm) (estimate 95% counting efficiency) and then into microcuries based on 1×10^6 dpm = 2.22 μCi. A number of 2-fold serial dilutions of the stock were measured and used to create a calibration curve by linear regression through the origin. The phosphorimager density was converted to a microcurie value using the slope of this calibration line. A rate was determined by dividing this value by duration of the enzyme reaction. The V_{max} value (microcuries per minute per nanogram) can be converted into the number of moles per minute per nanogram since the ratio of radiolabeled:unlabeled PAPS is known from scintillation counting of the PAPS stock solution. The K_M and k_{cat} values were determined by nonlinear regression least-squares fit using a Michaelis−Menten kinetic model (GraFit4, Erithacus Software Ltd.). Each data point was acquired in quadruplicate, with the entire experiment repeated at least twice using the wild-type enzymes.

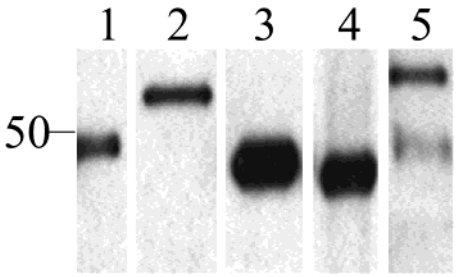

FIGURE 3: Western blot of the five purified GST1, GST2, GST3, GST4α, and GST5. The 50 kDa molecular mass marker is indicated to the left of the blot.

RESULTS

Expression of Recombinant GST Proteins. Genes encoding the Golgi lumenal domains of GST1, GST2, GST3, GST4α, and GST5 were subcloned into the plasmid pFASTBac that was modified to contain a honey bee melittin secretion signal fused to a thrombin cleavable N-terminal His_6 affinity tag (*90*). To produce recombinant GSTs, Sf9 or High-Five cells were infected with amplified baculovirus sufficient to halt cell growth to <25% within the first 24 h postinfection (PI). Cultures were incubated for the time required for optimal protein yield as listed in Table 1. Conditioned media containing secreted protein were obtained and processed as described in Experimental Procedures. Figure 3 shows Western blots for each of the five GSTs. Lower molecular weight bands are visible for GST5 and most likely correspond to C-terminal degradation products since the His_6 tag is fused to the N-terminus of the protein. Protein concentration was determined by Western blot using a known amount of a His_6-tagged protein that was serially diluted to generate a standard curve. Yields were in the hundreds of micrograms per liter of culture for the GSTs. This is the first report in which the expression of five GSTs has been achieved with reliable and quantifiable amounts of protein.

Activity Assays and Optimization of in Vitro Enzyme Reaction Conditions. Traditionally, GST activity has been measured by quantifying the transfer of radiolabeled $^{35}SO_3$ from ^{35}S-PAPS to a GlcNAc-terminated oligosaccharide using a thin-layer chromatography (TLC) assay (*83*−*85*, *88*). We employed a similar assay for analysis of GST2, GST3, GST4α, and GST5 using βBnO-GlcNAc as an artificial substrate. Aliquots of the reactions were adsorbed onto a semipreparative TLC plate, and radioactivity was detected by phosphorimaging. The amount of product was determined

by multiplying the total amount of PAPS used in the reaction by the percentage of radioactive sulfate incorporated into the higher eluting sulfated carbohydrate product. Recently, a more efficient dot-blot method for analyzing sulfotransferase reactions has been developed (*89*). This assay uses a synthetic GlcNAc-polyacrylamide polymer substrate for GST2, GST3, and GST5 or a natural keratan sulfate polymer for GST1. Reactions are performed in 96-well plates with the sulfated polymer captured onto either a PVDF (polyacrylamide polymer) or a positively charged membrane (keratan sulfate polymer). The amount of radioactive sulfate incorporated into the polymer substrate is reported as the phosphorimager signal density converted to microcuries on the basis of a standard curve (*89*). The polymer assay enables the processing of large numbers of samples, as required for the screening of small molecule inhibitor libraries or for the rapid optimization of several enzyme assays simultaneously.

Buffer, pH, and ATP and Mg^{2+} concentrations were optimized for each enzyme (Table 2). GST activity was optimized by testing enzymatic activity as a function of pH using Bis-Tris propane/glycine, Hepes, or Bis-Tris propane buffers ranging from pH 5 to pH 9 in 0.5 unit increments. Enzymes were not active in sodium phosphate−citrate buffer. Magnesium acetate and ATP concentrations were also optimized for each enzyme. Interestingly, GST4α appears to require significantly less ATP (0.05 mM) than the other GSTs, while GST2 requires the most ATP (2 mM). All GSTs required similar Mg^{2+} concentrations for optimal activity (2 or 5 mM; Table 2).

Recombinant Wild-Type GST Activity toward PAPS. Using the optimized conditions, the K_M of PAPS and the first- (k_{cat}) and second-order (k_{cat}/K_M) rate constants were determined for each enzyme (Table 3, section a, and Figure 4). Each data point shown in Figure 4 is the average of at least three data points, with the entire experiment repeated twice. Dot-blot assays were used for all the GSTs, except in the case of GST4α which did not sulfate the GlcNAc-polyacrylamide polymer, in which case the TLC assay using βBnO-GlcNAc as a substrate was used. The K_M for PAPS is in the micromolar range but varies about 10-fold among GST1, GST2, GST3, and GST5; in contrast, GST4α has a K_M for PAPS of 75 μM, a value that is approximately 20-fold higher than for the other GSTs measured. The k_{cat} varies 100-fold from 0.01 to 1.1 min^{-1} among the enzymes.

Mutagenesis of GST2. Five conserved residues within the putative PAPS or carbohydrate binding site were chosen for site-directed mutagenesis of GST2 (Figure 2). As indicated by boldface residues within the sequence alignment of Figure 2, we generated proteins containing the mutation R174A within the 5′-PSB region, the mutations R296A and K304A within the putative carbohydrate binding site, and the mutations R332A and R341A within the 3′-PB region. Four of the five mutations correspond to mutations found within GST4β in patients with MCD. All of the mutant proteins were expressed at a level comparable to that of wild-type GST2 (inset to Figure 5) and purified in the same manner. All mutants appeared to be stably folded and behaved like wild-type enzyme during purification. We found that mutations K304A and R332A abrogated enzyme activity, while mutations R174A, R296A, and R342A showed residual activities of 22%, 34%, and 34%, respectively, relative to wild type (Figure 5). In addition to activity, the K_M and k_{cat} values for PAPS and for the carbohydrate substrate were determined (Tables III, sections b and c). Each mutation is considered separately in the following sections.

Table 3: Kinetic Parameters for the Purified GSTs[a]

Section a

GST	K_M [PAPS] (μM)	k_{cat} (min^{-1})	$k_{cat}/K_M \times 10^4$ (M^{-1} min^{-1})
1[b]	0.4± 0.061	1.1 ± 0.046	280 ± 43
2[c]	3.9 ± 1.61	0.06 ± 0.01	1.5 ± 0.7
3[c]	0.47 ± 0.24	0.29 ± 0.035	61 ± 32
4α[d]	75 ± 18.5	0.31 ± 0.044	0.41 ± 0.12
5[c]	1.4 ± 0.52	0.01 ± 0.0003	0.48 ± 0.18

Section b

GST2	K_M [PAPS] (μM)	rel K_M	k_{cat} (min^{-1})	$k_{cat}/K_M \times 10^3$ (M^{-1} min^{-1})	mutant/WT rel k_{cat}/K_M
WT[c]	3.9 ± 1.6	1.0	0.06 ± 0.001	15 ± 7	1
R174A[d]	21 ± 9.9	5.4	0.12 ± 0.017	5.6 ± 2.8	0.37
R296A[c]	11.9 ± 3.1	3.1	0.017 ± 0.0025	1.4 ± 4.27	0.09
K304A	NA	NA	NA	NA	NA
R332A	NA	NA	NA	NA	NA
R341A[d]	43.2 ± 13.6	11.2	0.22 ± 0.025	5.1 ± 1.7	0.34

Section c

GST2	K_M [βBnO-GlcNAc] (mM)	rel K_M	k_{cat} (min^{-1})	k_{cat}/K_M (M^{-1} min^{-1})	mutant/WT rel k_{cat}/K_M
WT	1.4 ± 0.24	1.0	0.095 ± 0.007	69.9 ± 13.1	1
R174A	2.5 ± 0.78	1.8	0.048 ± 0.009	19.3 ± 6.9	0.28
R296A	1.6 ± 0.83	1.2	0.008 ± 0.0009	4.9 ± 2.6	0.07
K304A	NA	NA	NA	NA	NA
R332A	NA	NA	NA	NA	NA
R341A	2.4 ± 0.44	1.7	0.075 ± 0.0072	31.9 ± 6.7	0.46

[a] (a) Data reported are representative of at least two independent experiments where each concentration of PAPS was tested in either triplicate or quadruplicate. (b) Kinetic parameters for mutants of GST2 with respect to PAPS. The concentration of βBnO-GlcNAc or GlcNAc-PAA was held constant, and the amount of PAPS was varied in the assay. (c) Kinetic parameters for mutants of GST2 with respect to the substrate βBnO-GlcNAc. The concentration of PAPS was held constant, and the amount of βBnO-GlcNAc was varied in the assay. Error was determined as the standard deviation from at least three replicate experiments. Mutants K304A and R332A were expressed and purified but were not active (NA) (b and c). [b] TLC assay using keratan sulfate as a substrate. [c] Dot-blot assay using GlcNAc-PAA as a substrate. [d] TLC assay using βBnO-GlcNAc as a substrate.

Region I: 5′-PSB. The mutation made at R174A results in a k_{cat}/K_M value 28% that of wild-type GST2 for the acceptor substrate βBnO-GlcNAc and 37% that of wild-type GST2 for the donor substrate PAPS. The K_M for PAPS was increased approximately 5.5-fold while the K_M of the carbohydrate substrate was 2-fold higher than the wild-type value, suggesting that this residue is most important for PAPS binding, as predicted on the basis of sequence alignments with the cytosolic STs.

Region II: Carbohydrate Substrate Binding Region. Residues R296A and K304A are predicted to lie within the putative carbohydrate binding site. We chose these two residues on the basis of their conservation within the GST family and because mutations R166P and K174R are found within GST4β of MCD type I patients. As stated above, K304A had no activity under standard ST reaction conditions (4 μM PAPS and 1 mM βBnO-GlcNAc). R296A had k_{cat}/K_M values 9% and 7% that of the wild-type enzyme for PAPS and βBnO-GlcNAc, respectively. The K_M of R296A for PAPS was approximately 3-fold higher than that of the wild-

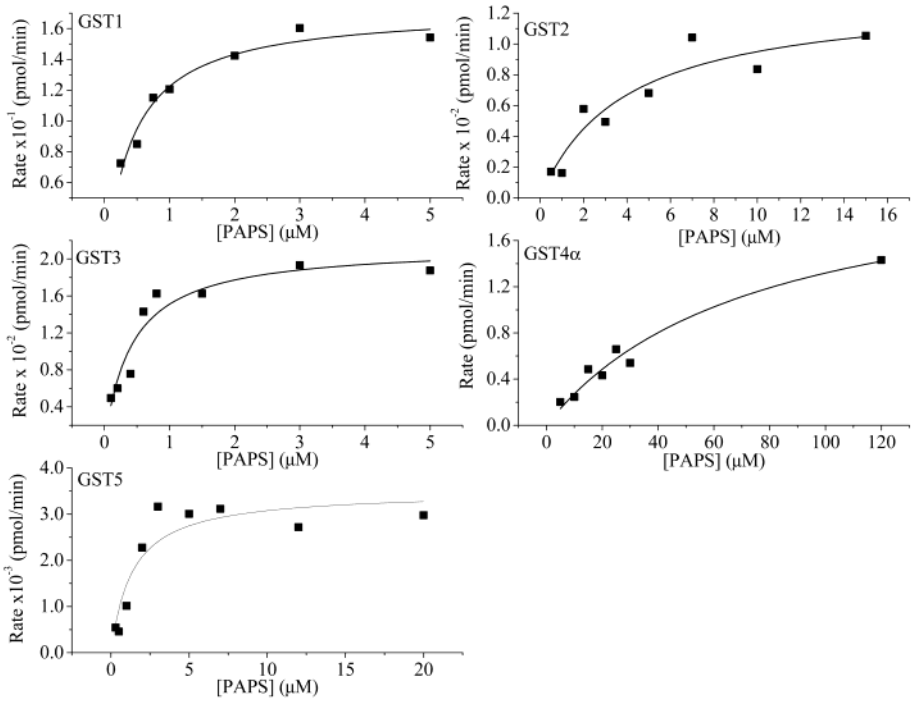

FIGURE 4: Dependence of GST activity on the concentration of PAPS as determined by the dot-blot assay for GST1, GST2, GST3, and GST5. GST4α was assayed using the TLC method.

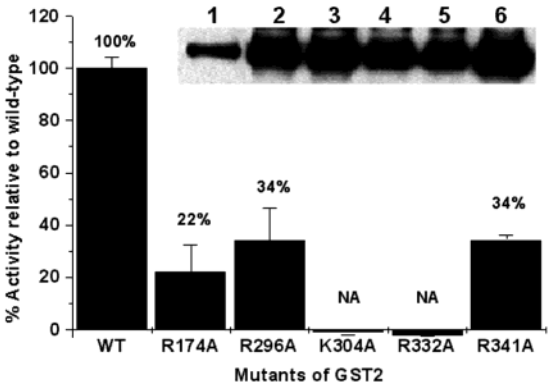

FIGURE 5: Comparison of the enzymatic activities of wild-type GST2 and mutants. Mutation R174A lies within the 5′-PSB region; R296A and K304A are within the putative carbohydrate binding site; R332A and R341A are within the 3′-PB region. Mutants K304A and R332A were not active under these reaction conditions. Inset: Western blot of wild-type GST2 and mutants: lane 1, wild-type GST2 (WT); lane 2, R174A; lane 3, R296A; lane 4, K304A; lane 5, R332A; lane 6, R341A.

type enzyme; however, the K_M for the carbohydrate substrate was close to that of the wild-type enzyme. Thus, this mutation has a greater effect on activity than binding of the substrate used in our assay, suggesting that the carbohydrate binding site identified by sequence analysis also includes residues involved in catalysis.

Region III: 3′-PB. The last two mutations are within the 3′-PB region. R332A (R204 in GST4β) is a conserved residue in the GST family, and the resulting protein was not active. One possibility is that the arginine side chain plays a role in neutralizing the charge on PAPS; however, from our results, we cannot distinguish between an effect on substrate binding and catalysis. The final mutation, R341A, had second-order rate constants (k_{cat}/K_M) that were 46% (PAPS) and 34% (βBnO-GlcNAc) that of the wild-type enzyme. The K_M of PAPS was 11-fold higher, and the K_M for βBnO-GlcNAc was 1.5-fold higher than that of the wild-type enzyme, suggesting that the arginine contributes to PAPS binding but is not essential for it.

DISCUSSION

Protein Expression. The GSTs have several features that make them difficult candidates for heterologous recombinant protein expression. They are membrane-bound Golgi enzymes that require removal of the N-terminal transmembrane domain to produce soluble protein. These enzymes contain at least eight cysteine residues which may form disulfide bonds necessary for proper folding. In addition, the GSTs have four to six potential N-linked glycosylation sites, and it is unknown whether glycosylation of these sites is required for folding or activity. The constructs made in this study result in proteins that start at the first residue after the predicted transmembrane domain. Successful expression of microgram quantities of all the enzymes was achieved using baculovirus-infected insect cells, which are able to glycosylate the GSTs, although we did not confirm the glycosylation state of the resulting protein.

Kinetic Analysis. Several groups have reported the substrate specifities of GST2 (*83*−*85*), GST3 (*83*−*85*), GST4α (*83*, *84*), and GST5 (*83*, *91*). It is difficult to compare our results with the previous studies, in part because we have used different artificial substrates (βBnO-GlcNAc or a polyacrylamide derivative of it). For example, the study of GST2, GST3, and GST4α by Uchimura and colleagues (*84*) used higher order carbohydrate substrates and unquantified amounts of enzyme immobilized on IgG−Sepharose. Despite these differences, the K_M values for PAPS we report for GST2 and GST4α are in good agreement with theirs (*84*), and our value for GST1 is in good agreement with a previously reported value (*54*). In contrast, we report a 10-fold lower K_M value for PAPS for GST3 than Uchimura et al. [0.47 μM compared to 5.9 μM (*84*)].

Mutagenesis. The mutagenesis results described provide the first in vitro data addressing the roles of specific residues in binding substrates and catalysis. The mutations were selected because they belong to three regions of sequence homology between the cytosolic and Golgi STs and because they correspond to mutations in GST4β that result in undersulfation of keratan sulfate in the cornea, a deficiency that leads to eventual blindness. Within the 5′-PSB, we mutated an arginine residue in GST2 originally identified by mutagenesis studies of the cytosolic STs and two Golgi-resident carbohydrate STs [heparan sulfate NDST-1 (*76*) and HNK-1ST (*75*)] in which the corresponding residue is a lysine. The mutation of the lysine in mouse estrogen ST (K48M, mEST) (*69*), human phenol ST (K48M, hSULT1A1), and human monoamine ST (K48M, hSULT1A3) (*72*) or the equivalent lysines in flavonol 3-ST (K59A, F3ST) (*73*, *92*), HNK-1ST (K128A) (*75*), and NDST-1 (K614A) (*76*) result in inactive or nearly inactive enzymes. In our experiments, we found that the corresponding mutation R174A in GST2 resulted in a 78% decrease in enzymatic activity compared to wild-type enzyme. We note that R174 corresponds to the R50C mutation in GST4β in MCD patients. It is not known whether these patients produce inactive protein or none at all.

We attempted to model the 5′-PSB of GST2 on the basis of the only known cyrstal structure of a Golgi ST, the sulfotransferase domain of NDST-1. Although there is no overall homology between these two Golgi STs, the 5′-PSB can be predicted on the basis of a motif search alone owing to its strong internal conservation. The identification of the 5′-PSB in GST2 is further supported by the coincidence of the observed secondary structure in the crystal with the secondary structure predicted (PSIPRED, 99) for GST2 (a sheet−turn−helix motif). On the basis of these observations, we propose a role for R174 in binding the 5′-phosphosulfate group (Figure 6) similar to that seen in the NDST-1 X-ray structure.

The NDST-1 crystal structure shows that the terminal amino group of the equivalent Lys614 side chain forms a hydrogen bond to the oxygen atom of the 5′-phosphate group of PAP (2.9 Å). In the mEST crystal structure, a similar interaction is formed between the terminal amino group of Lys48 and an equivalent oxygen (2.9 Å) in the 5′-phosphate group. Replacement of Lys with an Arg in GST2 should allow for retention of these contacts and also permit an additional hydrogen bond, perhaps to the 3′-phosphate group that is 5 Å away. This additional hydrogen bond could be direct if either the Arg side chain or the phosphate group position was adjusted, or it could be mediated by an intervening water molecule. There is an additional hydrogen bond formed between the backbone amide of K614 in NDST-1 and one of the 5′-phosphate oxygen atoms of PAP, and this backbone contact would not necessarily be altered in the GST2 mutant. The fact that 22% residual activity remains for the R174A mutation made in GST2 suggests that one of these hydrogen bonds remains.

The 3′-PB is less well conserved across the Golgi and cytosolic STs compared to the 5′-PSB and is therefore more difficult to identify with certainty in GST2. However, R332 within the GST2 sequence corresponds to the first residue of the conserved RDP sequence that defines the N-terminal start of the 3′-PB region and corresponds to mutations made

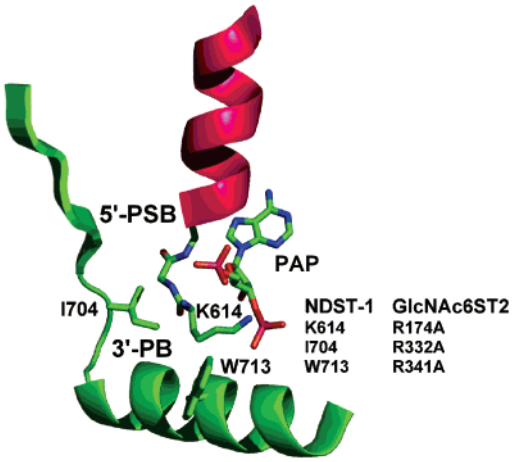

FIGURE 6: Key residues interacting with PAP, from the crystal structure of NDST-1 (*67*). The secondary structural elements that form the 3′-PB are shown in green; those of the 5′-PSB are shown in magenta. Shown are the modeled locations of the mutations made in GST2 based on sequence and secondary structure alignment using the crystal structure of NSDT-1. Side chains of K614 (R174A mutation in GST2), I704 (R332A mutation in GST2), and W713 (R341A mutation in GST2) are shown colored by atom type (blue, nitrogen; red, oxygen; green, carbon). The figure was prepared using PyMOL (*94*).

at R189A within HNK-1ST (*75*) and at R140S within F3ST which eliminate activity (*74*, *75*). Similarly, the R332A mutation in GST2 showed no activity, indicating a significant role in catalysis for an arginine at this position. The second mutation in the 3′-PB, R341A [R211W in GST4β in MCD type I patients (*44*)], resulted in a 66% loss in enzymatic activity, which is in part due to a 10-fold increase in the K_M for PAPS, supporting a direct role in binding. In NDST-1, the 3′-PB consists of β sheet 2 (residues 698−703) and the INP turn, followed by helix α6 (residues 707−719) (*67*). The only contact to PAP in the 3′-PB in the crystal structure is between the highly conserved S712 hydroxyl group and one of the 3′-phosphate oxygen atoms. Within the Golgi ST sequences, the putative 3′-PB regions are nearly identical, as are the predicted secondary structures which consist of a 5-residue sheet (326−330) followed by the analogous RDP turn ending in a 21-residue helix (residues 334−355), which matches the secondary structure of NDST-1 reasonably well. However, the exact locations and structural consequences of the mutations made in the 3′-PB are more difficult to predict from the NDST-1 crystal structure. If our secondary structure overlap is correct, both R332 and R341 should lie within the terminal helix which forms the bottom of the PAPS binding site (Figure 6), and their mutation could affect activity by perturbing the architecture of this part of the catalytic site.

In summary, only one other carbohydrate ST, HNK-1ST, has been studied by site-directed mutagenesis, and we present the first such study on a GST family member. These studies were made possible by the development of a baculovirus system to recombinantly express the enzymes in sufficient quantities for biochemical analysis. Future mutagenesis studies to define the carbohydrate binding pocket and to determine the number of cysteines necessary for the proper folding and activity of the GSTs are of great interest. Improving the yield of expressed protein through scale up or by devising a stable cell line that continuously secretes

the GSTs into the culture media is another immediate goal and may enable structural studies. The discovery of inhibitors of these enzymes, found by screening libraries of small molecules against the recombinant enzyme, may enable the chemical "knock-out" of their activity to study their in vivo functions.

ACKNOWLEDGMENT

We thank Jennifer Kohler for a critical reading of the manuscript.

REFERENCES

1. Armstrong, J. I., and Bertozzi, C. R. (2000) *Curr. Opin. Drug Discuss. Dev. 3*, 502−515.
2. Bowman, K. G., and Bertozzi, C. R. (1999) *Chem. Biol. 6*, R9−R22.
3. Hooper, L. V., Manzella, S. M., and Baenziger, J. U. (1996) *FASEB J. 10*, 1137−1146.
4. Alvarez-Dominguez, C., Vazquez-Boland, J. A., Carrasco-Marin, E., Lopez-Mato, P., and Leyva-Cobian, F. (1997) *Infect. Immun. 65*, 78−88.
5. Chen, Y., Maguire, T., Hileman, R. E., Fromm, J. R., Esko, J. D., Linhardt, R. J., and Marks, R. M. (1997) *Nat. Med. 3*, 866−871.
6. Feyzi, E., Trybala, E., Bergstrom, T., Lindahl, U., and Spillmann, D. (1997) *J. Biol. Chem. 272*, 24850−24857.
7. Hoffman, J. A., Badger, J. L., Zhang, Y., Huang, S. H., and Kim, K. S. (2000) *Infect. Immun. 68*, 5062−5067.
8. Jansen, H. J., Hart, C. A., Rhodes, J. M., Saunders, J. R., and Smalley, J. W. (1999) *J. Med. Microbiol. 48*, 551−557.
9. Patel, M., Yanagishita, M., Roderiquez, G., Bou-Habib, D. C., Oravecz, T., Hascall, V. C., and Norcross, M. A. (1993) *AIDS Res. Hum. Retroviruses 9*, 167−174.
10. Su, H., Raymond, L., Rockey, D. D., Fischer, E., Hackstadt, T., and Caldwell, H. D. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 11143−11148.
11. Summerford, C., and Samulski, R. J. (1998) *J. Virol. 72*, 1438−1445.
12. Fuxe, K., Tinner, B., Staines, W., David, G., and Agnati, L. F. (1997) *Brain Res. 746*, 25−33.
13. Kjellen, L., and Lindahl, U. (1991) *Annu. Rev. Biochem. 60*, 443−475.
14. Sugahara, K., and Kitagawa, H. (2000) *Curr. Opin. Struct. Biol. 10*, 518−527.
15. Tanaka, Y., Fujii, K., Hubscher, S., Aso, M., Takazawa, A., Saito, K., Ota, T., and Eto, S. (1998) *Arthritis Rheum. 41*, 1365−1377.
16. Baenziger, J. U. (1995) *Glycobiology 5*, 459.
17. Skelton, T. P., Hooper, L. V., Srivastava, V., Hindsgaul, O., and Baenziger, J. U. (1991) *J. Biol. Chem. 266*, 17142−17150.
18. Xia, G., Evers, M. R., Kang, H. G., Schachner, M., and Baenziger, J. U. (2000) *J. Biol. Chem. 275*, 38402−38409.
19. Barbucci, R., Lamponi, S., Magnani, A., Poletti, L. F., Rhodes, N. P., Sobel, M., and Williams, D. F. (1998) *J. Thromb. Thrombolysis 6*, 109−115.
20. Akashi, M., Sakamoto, N., Suzuki, K., and Kishida, A. (1996) *Bioconjugate Chem. 7*, 393−395.
21. Anderson, J. A., Fredenburgh, J. C., Stafford, A. R., Guo, Y. S., Hirsh, J., Ghazarossian, V., and Weitz, J. I. (2001) *J. Biol. Chem. 276*, 9755−9761.
22. Dol, F., Caranobe, C., Dupouy, D., Petitou, M., Lormeau, J. C., Choay, J., Sie, P., and Boneu, B. (1988) *Thromb. Res. 52*, 153−164.
23. White, G. C., Jr., Pickens, E. M., Liles, D. K., and Roberts, H. R. (1998) *Transfus. Sci. 19*, 177−189.
24. Uchiyama, H., Metori, A., Ogamo, A., and Nagasawa, K. (1990) *J. Biochem. (Tokyo) 107*, 377−380.
25. Pittman, D. D., Tomkinson, K. N., Michnick, D., Selighsohn, U., and Kaufman, R. J. (1994) *Biochemistry 33*, 6952−6959.
26. Ogamo, A., Metori, A., Uchiyama, H., and Nagasawa, K. (1990) *J. Biochem.* (*Tokyo*) *108*, 588−592.
27. Michnick, D. A., Pittman, D. D., Wise, R. J., and Kaufman, R. J. (1994) *J. Biol. Chem. 269*, 20095−20102.
28. Lipscombe, R. J., Nakhoul, A. M., Sanderson, C. J., and Coombe, D. R. (1998) *J. Leukocyte Biol. 63*, 342−350.
29. Selvan, R. S., Ihrcke, N. S., and Platt, J. L. (1996) *Ann. N.Y. Acad. Sci. 797*, 127−139.
30. Vlodavsky, I., Miao, H. Q., Medalion, B., P., D., and Ron, D. (1996) *Cancer Metastasis Rev. 15*, 177−186.
31. Guimond, S., Maccarana, M., Olwin, B. B., Lindahl, U., and Rapraeger, A. C. (1993) *J. Biol. Chem. 268*, 23906−23914.
32. Rapraeger, A. C. (1995) *Chem. Biol. 2*, 645−649.
33. Imai, Y., and Rosen, S. D. (1993) *Glycoconjugate J. 10*, 34−39.
34. Rosen, S. D. (1993) *Res. Immunol. 144*, 699−703.
35. Rosen, S. D., and Bertozzi, C. R. (1994) *Curr. Opin. Cell Biol. 6*, 663−673.
36. Rosen, S. D., Hwang, S. T., Giblin, P. A., and Singer, M. S. (1997) *Biochem. Soc. Trans. 25*, 428−433.
37. Rosen, S. D. (1999) *Am. J. Pathol. 155*, 1013−1020.
38. Varki, A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 7390−7397.
39. Hemmerich, S. (2001) *Drug Discovery Today 6*, 27−35.
40. Rossi, A., Bonaventure, J., Delezoide, A. L., Cetta, G., and Superti-Furga, A. (1996) *J. Biol. Chem. 271*, 18456−18464.
41. Sugahara, K., and Schwartz, N. B. (1982) *Arch. Biochem. Biophys. 214*, 589−601.
42. Pennypacker, J. P., Kimata, K., and Brown, K. S. (1981) *Dev. Biol. 81*, 280−287.
43. Superti-Furga, A. (1994) *Am. J. Hum. Genet. 55*, 1137−1145.
44. Akama, T. O., Nishida, K., Nakayama, J., Watanabe, H., Ozaki, K., Nakamura, T., Dota, A., Kawasaki, S., Inoue, Y., Maeda, N., Yamamoto, S., Fujiwara, T., Thonar, E. J., Shimomura, Y., Kinoshita, S., Tanigami, A., and Fukuda, M. N. (2000) *Nat. Genet. 26*, 237−241.
45. Liu, N. P., Dew-Knight, S., Rayner, M., Jonasson, F., Akama, T. O., Fukuda, M. N., Bao, W., Gilbert, J. R., Vance, J. M., and Klintworth, G. K. (2000) *Mol. Vision 6*, 261−264.
46. Hasegawa, N., Torii, T., Nagaoka, I., Nakayasu, K., Miyajima, H., and Habuchi, O. (1999) *J. Biochem.* (*Tokyo*) *125*, 245−252.
47. Hasegawa, N., Torii, T., Kato, T., Miyajima, H., Furuhata, A., Nakayasu, K., Kanai, A., and Habuchi, O. (2000) *Invest. Ophthalmol. Vis. Sci. 41*, 3670−3677.
48. El-Ashry, M. F., El-Aziz, M. M., Wilkins, S., Cheetham, M. E., Wilkie, S. E., Hardcastle, A. J., Halford, S., Bayoumi, A. Y., Ficker, L. A., Tuft, S., Bhattacharya, S. S., and Ebenezer, N. D. (2002) *Invest. Ophthalmol. Vis. Sci. 43*, 377−382.
49. Akama, T. O., Nakayama, J., Nishida, K., Hiraoka, N., Suzuki, M., McAuliffe, J., Hindsgaul, O., Fukuda, M., and Fukuda, M. N. (2001) *J. Biol. Chem. 276*, 16271−16278.
50. Kehoe, J. W., and Bertozzi, C. R. (2000) *Chem. Biol. 7*, R57−R61.
51. Tsutsumi, K., Shimakawa, H., Kitagawa, H., and Sugahara, K. (1998) *FEBS Lett. 441*, 235−241.
52. Habuchi, O., Suzuki, Y., and Fukuta, M. (1997) *Glycobiology 7*, 405−412.
53. Li, X., and Tedder, T. F. (1999) *Genomics 55*, 345−347.
54. Fukuta, M., Inazawa, J., Torii, T., Tsuzuki, K., Shimada, E., and Habuchi, O. (1997) *J. Biol. Chem. 272*, 32321−32328.
55. Mazany, K. D., Peng, T., Watson, C. E., Tabas, I., and Williams, K. J. (1998) *Biochim. Biophys. Acta 1407*, 92−97.
56. Uchimura, K., Muramatsu, H., Kaname, T., Ogawa, H., Yamakawa, T., Fan, Q. W., Mitsuoka, C., Kannagi, R., Habuchi, O., Yokoyama, I., Yamamura, K., Ozaki, T., Nakagawara, A., Kadomatsu, K., and Muramatsu, T. (1998) *J. Biochem. (Tokyo) 124*, 670−678.
57. Bistrup, A., Bhakta, S., Lee, J. K., Belov, Y. Y., Gunn, M. D., Zuo, F. R., Huang, C. C., Kannagi, R., Rosen, S. D., and Hemmerich, S. (1999) *J. Cell Biol. 145*, 899−910.
58. Lee, J. K., Bhakta, S., Rosen, S. D., and Hemmerich, S. (1999) *Biochem. Biophys. Res. Commun. 263*, 543−549.
59. Bartes, A., Bhakta, S., and Hemmerich, S. (2001) *Biochem. Biophys. Res. Commun. 282*, 928−933.
60. Kitagawa, H., Fujita, M., Ito, N., and Sugahara, K. (2000) *J. Biol. Chem. 275*, 21075−21080.
61. Uchimura, K., Kadomatsu, K., Nishimura, H., Muramatsu, H., Nakamura, E., Kurosawa, N., Habuchi, O., El-Fasakhany, F. M., Yoshikai, Y., and Muramatsu, T. (2002) *J. Biol. Chem. 277*, 1443−1450.
62. Bhakta, S., Bartes, A., Bowman, K. G., Kao, W. M., Polsky, I., Lee, J. K., Cook, B. N., Bruehl, R. E., Rosen, S. D., Bertozzi, C. R., and Hemmerich, S. (2000) *J. Biol. Chem. 275*, 40226−40234.
63. Bidwell, L. M., McManus, M. E., Gaedigk, A., Kakuta, Y., Negishi, M., Pedersen, L., and Martin, J. L. (1999) *J. Mol. Biol. 293*, 521−530.

64. Dajani, R., Cleasby, A., Neu, M., Wonacott, A. J., Jhoti, H., Hood, A. M., Modi, S., Hersey, A., Taskinen, J., Cooke, R. M., Manchee, G. R., and Coughtrie, M. W. (1999) *J. Biol. Chem. 274*, 37862−37868.
65. Kakuta, Y., Pedersen, L. G., Carter, C. W., Negishi, M., and Pedersen, L. C. (1997) *Nat. Struct. Biol. 4*, 904−908.
66. Yoshinari, K., Petrotchenko, E. V., Pedersen, L. C., and Negishi, M. (2001) *J. Biochem. Mol. Toxicol. 15*, 67−75.
67. Kakuta, Y., Sueyoshi, T., Negishi, M., and Pedersen, L. C. (1999) *J. Biol. Chem. 274*, 10673−10676.
68. Pedersen, L. C., Petrotchenko, E., Shevtsov, S., and Negishi, M. (2002) *J. Biol. Chem. 7*, 17928−17932.
69. Kakuta, Y., Petrotchenko, E. V., Pedersen, L. C., and Negishi, M. (1998) *J. Biol. Chem. 273*, 27325−27330.
70. Hempel, N., Barnett, A. C., Bolton-Grob, R. M., Liyou, N. E., and McManus, M. E. (2000) *Biochem. Biophys. Res. Commun. 276*, 224−230.
71. Homma, H., Ogawa, K., Hirono, K., Morioka, Y., Hirota, M., Tanahashi, I., and Matsui, M. (1996) *Biochim. Biophys. Acta 1296*, 159−166.
72. Liu, M. C., Suiko, M., and Sakakibara, Y. (2000) *J. Biol. Chem. 275*, 13460−13464.
73. Marsolais, F., and Varin, L. (1995) *J. Biol. Chem. 270*, 30458−30463.
74. Marsolais, F., and Varin, L. (1997) *Eur. J. Biochem. 247*, 1056−1062.
75. Ong, E., Yeh, J. C., Ding, Y., Hindsgaul, O., Pedersen, L. C., Negishi, M., and Fukuda, M. (1999) *J. Biol. Chem. 274*, 25608−25612.
76. Sueyoshi, T., Kakuta, Y., Pedersen, L. C., Wall, F. E., Pedersen, L. G., and Negishi, M. (1998) *FEBS Lett. 433*, 211−214.
77. Falany, C. N. (1997) *FASEB J. 11*, 206−216.
78. Varin, L., and Ibrahim, R. K. (1992) *J. Biol. Chem. 267*, 1858−1863.
79. Whittemore, R. M., Pearce, L. B., and Roth, J. A. (1985) *Biochemistry 24*, 2477−2482.
80. Zhang, H., Varlamova, O., Vargas, F. M., Falany, C. N., Leyh, T. S., and Varmalova, O. (1998) *J. Biol. Chem. 273*, 10888−10892.
81. Duffel, M. W., and Jakoby, W. B. (1981) *J. Biol. Chem. 256*, 11123−11127.
82. Barnes, S., Waldrop, R., Crenshaw, J., King, R. J., and Taylor, K. B. (1986) *J. Lipid Res. 27*, 1111−1123.
83. Bowman, K. G., Cook, B. N., de Graffenried, C. L., and Bertozzi, C. R. (2001) *Biochemistry 40*, 5382−5391.
84. Uchimura, K., El-Fasakhany, F. M., Hori, M., Hemmerich, S., Blink, S. E., Kansas, G. S., Kanamori, A., Kumamoto, K., Kannagi, R., and Muramatsu, T. (2002) *J. Biol. Chem. 277*, 3979−3984.
85. Seko, A., Nagata, K., Yonezawa, S., and Yamashita, K. (2002) *Glycobiology 12*, 379−388.
86. Dixon, J., Loftus, S. K., Gladwin, A. J., Scambler, P. J., Wasmuth, J. J., and Dixon, M. J. (1995) *Genomics 26*, 239−244.
87. Burnell, J. N., and Whatley, F. R. (1975) *Anal. Biochem. 68*, 281−288.
88. Cook, B. N., Bhakta, S., Biegel, T., Bowman, K. G., Armstrong, J. I., Hemmerich, S., and Bertozzi, C. R. (2000) *J. Am. Chem. Soc. 122*, 8612−8622.
89. Verdugo, D. A., and Bertozzi, C. R. (2002) *Anal. Biochem. 307*, 330−336.
90. Tessier, D. C., Thomas, D. Y., Khouri, H. E., Laliberte, F., and Vernet, T. (1991) *Gene 98*, 177−183.
91. Uchimura, K., Fasakhany, F., Kadomatsu, K., Matsukawa, T., Yamakawa, T., Kurosawa, N., and Muramatsu, T. (2000) *Biochem. Biophys. Res. Commun. 274*, 291−296.
92. Marsolais, F., Laviolette, M., Kakuta, Y., Negishi, M., Pedersen, L. C., Auger, M., and Varin, L. (1999) *Biochemistry 38*, 4066−4071.
93. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−4680.
94. DeLano, W. L. (2002) DeLano Scientific, San Carlos, CA.

BI0269557