

# The Sulfogalactose Moiety of Sulfoglycosphingolipids Serves as a Mimic of Tyrosine Phosphate in Many Recognition Processes

PREDICTION AND DEMONSTRATION OF Src HOMOMOLOGY 2 DOMAIN/SULFOGALACTOSE BINDING\*

Received for publication, December 6, 2004, and in revised form, January 4, 2005  
Published, JBC Papers in Press, January 5, 2005, DOI 10.1074/jbc.M413724200

Clifford Lingwood<sup>‡§¶</sup>, Murugesapillai Mylvaganam<sup>‡</sup>, Farah Minhas<sup>‡</sup>, Beth Binnington<sup>‡</sup>,  
Donald R. Branch<sup>\*\*‡§§</sup>, and Régis Pomès<sup>‡§</sup>

From the <sup>‡</sup>Research Institute, The Hospital for Sick Children, Toronto, Ontario M4G 1X8, Canada, <sup>§</sup>Departments of Biochemistry, <sup>¶</sup>Laboratory Medicine and Pathobiology, and <sup>\*\*</sup>Medicine, University of Toronto, Toronto, Ontario M5S 1A8, Canada, <sup>‡‡</sup>Division of Cell and Molecular Biology, Toronto General Research Institute of the University Health Network and <sup>§§</sup>Canadian Blood Services, Toronto, Ontario M5G 2M1, Canada

**Multiple ligand co-recognition of 3'-sulfogalactosylceramide (SGC) and sulfotyrosine initiated the comparison of SGC and sulfotyrosine and, subsequently, phosphotyrosine (pY) binding. SGC is a receptor for ligands involved in cell adhesion/microbial pathology. pY forms a Src homology domain 2 recognition motif in intracellular signaling. Using hsp70, anti-SGC, and anti-pY antibodies, ligand binding is retained following phosphate/sulfate and tyrosine/galactose substitution in SGC and sulfate/phosphate exchange in pY. Remarkable lipid-dependent binding to phosphatidylethanolamine-conjugated sulfotyrosine suggests "microenvironmental" modulation of sulfotyrosine-containing receptors, similar to glycosphingolipids. Based on an aryl substrate-bound co-crystal of arylsulfatase A, a sulfogalactose and phosphotyrosine esterase, modeling provides a solvation basis for co-recognition. c-Src/ Src homology domain 2:SGC/phosphogalactosylceramide binding confirms our hypothesis, heralding a carbohydrate-based approach to regulation of phosphotyrosine-mediated recognition.**

Sulfogalactolipids (SGLs)<sup>1</sup> play a central role in spermatogenesis and nerve function (1) and are highly expressed in gastrointestinal and renal tissue (2). Our glycolipid receptor studies showed that 3'-sulfogalactosylceramide (sulfatide, SGC) bound hsp70 family members (3–7). Interaction of the N-terminal domain of hsp70 with SGC or its derivatives results in the inhibition of hsp70 ATPase activity (8), suggesting that hsp70/SGL binding might modulate chaperone function. Our finding that cytosolic hsp70s showed a similar SGL binding specificity (9) yet were unlikely co-localized with membrane SGC in cells, questioned the physiological significance of this mechanism of regulation of ATPase activity of such chaperones.

\* This work was supported by CIHR Grant MT13073. 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. Tel.: 416-813-5998; Fax: 416-813-5993; E-mail: cling@sickkids.ca.

<sup>1</sup> The abbreviations used are: SGL, sulfogalactolipid; SGC, 3'-sulfogalactosylceramide (sulfatide); PGC, phosphogalactosylceramide; SH2, Src homology domain; pY, phosphotyrosine; PE, phosphatidylethanolamine; GSL, glycosphingolipid; nt, nucleotide; TLC, thin layer chromatogram; pNCS, p-nitrophenyl catachol sulfatase; GST, glutathione S-transferase; GM3, N-acetylneuraminylgalactosylceramide; DHD, dihexadecyl; mAb, monoclonal antibody; YopH, *Yersinia* tyrosine phosphatase; GalCer, galactosylceramide.

In addition, the similar SGL binding of hsp70s from prokaryotes, which contain no SGLs, questions whether SGL binding is an evolutionarily conserved or fortuitous function. hsp70 binding to the sulfogalactosyl residue is dependent on (6) and modulated by (9, 10) the lipid moiety. Lipid modulation of glycosphingolipid (GSL) receptor function is observed frequently (11).

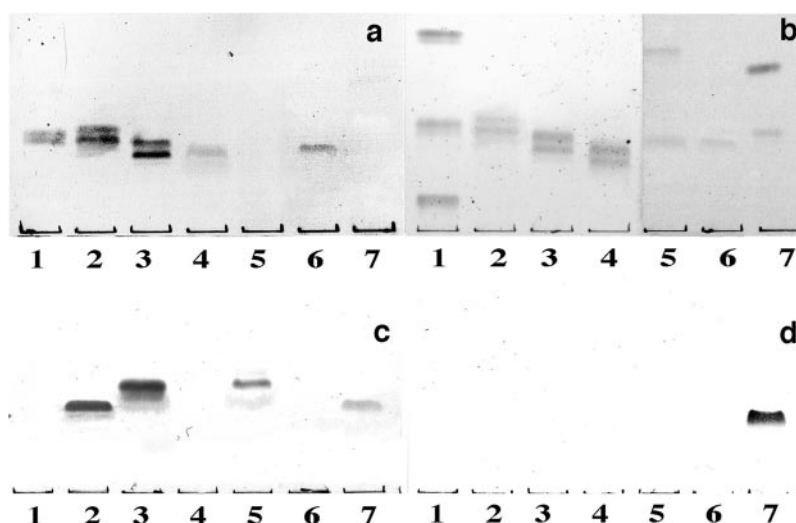
Three other proteins, which specifically bind SGC, *i.e.* the human immunodeficiency virus adhesin (glycoprotein 120) (12), the coagulation protein (von Willebrand factor VIII) (13), and the eukaryotic selectin adhesin family (14, 15), also bind tyrosine sulfate (16–18). Therefore, we considered whether tyrosine sulfate and SGC recognition were related and, subsequently, whether phosphotyrosine and sulfogalactose recognition could be related. Peptide mimics of carbohydrate epitopes are well known (19). Therefore, the reverse was also considered possible. Considerable evidence in the literature is consistent with this concept.

## MATERIALS AND METHODS

Tyrosine sulfate conjugated to dihexadecyl (DHD) phosphatidylethanolamine (PE) prepared as described previously (18) was kindly supplied by Dr. T. Feizi (Imperial College, Harrow, United Kingdom). Tyrosine and tyrosine sulfate coupled to PE were prepared by the same procedure in our laboratory using both DHD and *Escherichia coli* diacyl-PE. 4'-SGC, 6'-SGC, and 6'-phosphogalactosylceramide (6'-PGC) were synthesized from GalCer to retain ceramide heterogeneity. The structure of these compounds was verified by mass spectrometry, and the synthetic procedure will be described elsewhere. The binding of mAb anti-SGC (Sulf1), kindly provided by Dr. P. Fredman (University of Goteborg, Goteborg, Sweden), and DnaK by TLC overlay assay was detected immunologically as described previously (9, 20), and bound rabbit polyclonal anti-pY (Upstate Biotechnology Inc., Lake Placid, NY) were detected according to the supplier's instructions.

Recombinant fusion gene constructs between SH2 domains from human c-Src kinase (residues 440–766) or the tyrosine phosphatase SHP-1 and N-terminal (residues 1–345), C-terminal (residues 328–654), and SHP-1 domains (C+N-terminal SH2 domains) (21) were generated.

Human full-length c-Src cDNA was generated by transposing residues 1–875 of the 5'-cDNA sequence from human c-Src cDNA in the Bluescript plasmid (a kind gift of Dr. D. Fujita, University of Calgary, Calgary, Alberta, Canada) and residues 876–1611 of the 3'-cDNA sequence isolated from human MCF-10A cells using PCR cloning based on the published sequence (22, 23). The two c-Src fragments were recombined at a KpnI site, the construct was cloned using TA-cloning vector, and the entire sequence was verified (ACGT, Toronto, Ontario, Canada). The cDNA then was introduced into pcDNA3.1/His A (Invitrogen) vector and used as a basic plasmid for construction of GST-c-SrcSH2. cDNA for c-SrcSH2 was generated from pcDNA3 using the 5'-primers Src-SH2-F (5'-GAATTCAGGCTGAGGAGTGGTATTTTG-3' corresponding to nt 440–461 of c-Src and the 3'-primer Src-SH2-R (5'-CTCGAGTCTGCGGCTTGACGTTGGGCA-3' complementary to nt 742–766 of c-Src).



**FIG. 1. Effect of phosphate or sulfate substitution in galactolipid and tyrosyl-lipid protein recognition: a comparison of DnaK, mAb anti-SGC, and anti-phosphotyrosine antibody binding by TLC overlay.** Panel a, DnaK binding to 3'-SGC (GalCer, GM3 present, not bound) (lane 1), 6'-SGC (doublet due to hydroxy and non-hydroxy fatty acids) (lane 2), 4'-SGC (lane 3), 6'-PGC (lane 4), tyrSO<sub>4</sub>-*E. coli* PE (lane 5), tyrSO<sub>4</sub>-DHD PE (lane 6) and *E. coli* PE/phosphatidylcholine and cholesterol sulfate (lane 7). Panel b, chemical detection. Lanes 1–4, orcinol spray for carbohydrate. Lanes 5–7, molybdenum blue spray for phosphate. Lane 1 (from top) GalCer, 3'-SGC, and GM3 ganglioside. Lane 2, 6'-SGC. Lane 3, 4'-SGC. Lane 4, 6'-PGC. Lane 5, tyr-*E. coli* PE (upper), tyrSO<sub>4</sub>-*E. coli* PE. Lane 6, tyrSO<sub>4</sub>-DHD PE. Lane 7, *E. coli* PE (upper), phosphatidylcholine (lower), and cholesterol sulfate (not stained). Panel c, mAb anti-SGC binding. Panel d, polyclonal rabbit anti-pY binding. Lane 1, *E. coli* PE. Lane 2, tyrSO<sub>4</sub>-*E. coli* PE. Lane 3, 3'-SGC. Lane 4, 6'-PGC. Lane 5, 6'-SGC. Lane 6, 4'-SGC. Lane 7, tyrSO<sub>4</sub>-DHD PE. Each lane contains 1  $\mu$ g of lipid.

The full-length SHP-1 from peripheral blood mononuclear cells encoding amino acids 1–595 (bb) were amplified with oligonucleotides SHPA (5'-GATCAGGAATCCCATGGTGAGGTGGT-3') and SHPB (5'-TCGCTCGAGCCACCTGAGGACAGCACCGCT-3'). For making SHP-1/SH2N, SHP-1/SH2C, and SHP-1/SH2N-SH2C DNA fragments, a jumping PCR approach was applied from above full-length SHP-1 DNA template using the 5'-primer SHPA, corresponding to nt 1–16 of SHP-1, and SHP-1 SH2C-F (5'-GATCAGGAATCCCATGGTGAGGTGGTACCATTGGCCACATC-3'), corresponding to nt 1–9 and 328–345 of SHP-1, respectively, and the 3'-primer SHP-1 SH2N-R (5'-CTCGAGCCTCTCACTAGTGGGATCGGA-3'), complementary to nt 307–327 of SHP-1, and SHP-1 SH2C-R (5'-CTCGAGCACCC TCGTGGCATAGTACGG-3'), complementary to nt 634–654 of SHP-1).

PCRs were performed in 50  $\mu$ l of reaction mixture using 50 ng of primers and 2.5 units of *Taq* polymerase (Invitrogen) according to the recommendation of manufacturer. The PCR products contained an *Eco*RI as well as *Xho*I restriction sites at the N and C termini. These DNAs were digested with *Eco*RI and *Xho*I restriction enzymes and cloned within the pGEX-4T-2 (Amersham Biosciences) vector at the same sites located downstream of glutathione *S*-transferase (GST) driven by *tac* promoter and in-frame with GST. The region that contained the truncations was completely sequenced to assure the effectiveness of the deletion and that no other mutations were introduced during deletion or cloning.

GST-c-Src/SH2, SHP-1/SH2 C, SHP-1/SH2 N, and SHP-1/SH2 C+N proteins were overexpressed in *E. coli* strain BL21(DE3) (Novagen) and transformed with the pGEX-4T-2 constructs, and the GST fusion proteins were produced as described previously (24). A 500-ml culture of *E. coli* cells expressing each pGEX-4T-2/SH2 construct was grown at 37 °C until an optical density of 0.6 at 600 nm was reached. The culture was then induced with 1 mM isopropyl- $\beta$ -D-thiogalactopyranoside, and the incubation continued for an additional 2 h. The cells were harvested and disrupted by sonication (six cycles of 50 s with a 1-min pause in between on ice). Triton X-100 to final concentration of 1% was added to the disrupted cells, and then cells were centrifuged at 10,000  $\times$  g for 5 min at 4 °C. 1 ml of a 50% slurry of glutathione-Sepharose 4B beads was added to the supernatant and mixed gently for 5 min at room temperature. The beads were washed three times by adding 50 ml of ice-cold phosphate-buffered saline and centrifuging for 10 s at 500  $\times$  g. The fusion protein was eluted by adding 1 ml of 50 mM Tris-HCl, pH 8.0, and 5 mM reduced glutathione, and the purified GST protein fractions were analyzed by SDS-PAGE for appropriate size.

GSL binding of the SH2 domain fusion proteins was determined by TLC overlay, either by dot blot or chromatographically separated species. TLC chips or plates were blocked with bovine serum albumin/phosphate-buffered saline and incubated with fusion proteins (~1  $\mu$ g/

ml) overnight at room temperature. After washing, bound SH2-GST was detected using an anti-GST (9)/peroxidase/chloronaphthol system. GST alone was used as control.

In the molecular modeling studies, the minimum energy conformations of tyrosine phosphate and of 3'-sulfogalactose were obtained from *ab initio* Restricted Hartree-Fock/6-31G\* geometry optimization using the GAUSSIAN98 program (25). The initial structure of 3'-sulfogalactose was derived from Protein Data Bank structure 1ONQ of a CD1-sulfatide complex (26) followed by initial geometry optimization at the Slater Type Orbital (STO-3G) level. Point charges for the O atoms of the sulfate and phosphate groups were obtained from the CHELPG (charges from electrostatic potentials using a grid) method (27).

Galactose 3'-sulfate was docked into the binding site of *p*-nitrophenyl catachol sulfate co-crystallized within the arylsulfatase substrate binding site (Protein Data Bank code 1E2S) (28). The sulfate group of galactose 3'-sulfate was forced to occupy the same position as that of *p*NCS with harmonic restraints, and the potential energy of the system *in vacuo* was minimized with a fully flexible ligand and binding pocket using the program CHARMM (Chemistry at Harvard Molecular Mechanics) (29) together with the CHARMM22 force field (30) and carbohydrate parameters developed by Brady and co-workers (31, 67).

## RESULTS AND DISCUSSION

**Substitution of Phosphate for Sulfate and Tyrosine for Galactose: DnaK and mAb Anti-SGC Sulfogalactolipid Binding—** Positional isomers of SGC were made together with the phosphate analogue of 6'-SGC, 6'-PGC. To address the effect of tyrosine substitution for galactose on ligand binding, tyrosine and tyrosine sulfate derivatives of PE, originally generated to demonstrate P and L-selectin binding (18), were constructed. Fig. 1a shows that, in addition to 3'-SGC, both (non-physiological) 4'- and 6'-SGC are effectively bound by DnaK (*E. coli* hsp70). A comparison of the binding of 6'-SGC with 6'-PGC (Fig. 1a) shows the retention of DnaK binding following phosphate-sulfate substitution at this position. However, the tolerance for phosphate substitution within SGC is ligand-selective. The mAb anti-3'-SGC antibody, Sulf1 (20), while effectively binding 6'-SGC (although less than 3'-SGC), does not recognize 6'-PGC (Fig. 1c). Thus, in this case, phosphate cannot substitute for sulfate within the galactosphingolipid. Unlike DnaK, Sulf1 does not bind 4'-SGC (Fig. 1c). The doublet for 6'-SGC on TLC (hydroxyl and non-hydroxylated fatty acids) is bound by

DnaK, whereas only the upper (non-hydroxylated) species is bound by Sulfl, a clear demonstration of lipid modulation of SGC binding. Both DnaK and Sulfl tolerate galactose substitution by tyrosine.

Commercially available anti-pY antibodies are routinely used to screen for phosphotyrosine proteins (32) following up-regulation of intracellular signal transduction pathways. These antibodies do not bind phosphoserine or phosphothreonine (33) but may (34) or may not (35, 36) bind sulfotyrosine. The target antigens are specific pY residues within proteins in signal transduction cascades generated by a balance between cytosolic tyrosine kinases (37) and an equivalent family of phosphatases (38). DnaK, Sulfl, and anti-pY antibodies show significant binding to (DHD) PE-sulfotyrosine (Fig. 1, a, c, and d). However, Sulfl preferentially bound sulfotyrosine-*E. coli* PE (diacyl, containing a complex mixture of fatty acid (C16:0, C16:1, and C18:1) esters (39)). Remarkably, the binding of both anti-pY and DnaK to sulfotyrosine-*E. coli* PE was not observed, whereas Sulfl showed a significant preference for sulfotyrosine-*E. coli* PE (Fig. 1). Thus, the character of sulfotyrosine-PE binding is typical of GSL receptors in that the lipid moiety can have a major effect on recognition.

Thus, for DnaK, sulfate can be replaced by phosphate and galactose can be substituted by tyrosine. For mAb anti-SGC, galactose can be replaced by tyrosine but sulfate cannot be replaced by phosphate, at least in the 6' position. For anti-pY, phosphate can be replaced sulfate (in a select "chemoenvironment") but galactose cannot replace tyrosine.

Recognition of the phosphotyrosine motifs within signaling proteins by downstream signaling molecules is a highly selective process that involves not only the phosphotyrosine moiety itself but its molecular environment as well (adjacent amino acids) (40). Particularly, tyrosine residues within the cytosolic domains of transmembrane proteins are prevalent adjacent to the membrane domain (41). Thus, the membrane environment may play a role in the receptor function of such residues when phosphorylated, as we have observed for GSLs (42). The recognition of SGLs by their ligands is modulated by their "molecular environment" (the lipid moiety of the SGL) (9, 10). Both sulfotyrosine (Fig. 1) and sulfogalactose (10) ligand binding is similarly modulated by lipid conjugation, supporting a structural relationship between these epitopes.

**Comparison of Galactose 3'-Sulfate and Phosphotyrosine Conformation**—Molecular superimposition of the minimal energy conformers of phosphotyrosine and 3'-sulfogalactose shows an essential overlap of the rings and a significant coincidence of the orientation of the sulfate and phosphate moieties. The energy-minimized conformation and charge distribution for phosphotyrosine were obtained from *ab initio* calculations (Fig. 2A). The conformation of 3'-sulfogalactose from the structure of the SGC/CD1a co-crystal (26) was used as the starting point for geometry optimization (Fig. 2B). The charge distribution within the phosphate oxygen atoms is not greatly dissimilar from that of the sulfate oxygen atoms, despite the phosphate double and the sulfate single charge. This is explained by the partial delocalization of the phosphate charge within the tyrosine ring.

**Arylsulfatase Provides a Model for Phosphotyrosine-Sulfogalactose Equivalence**—Precedent for the structural convergence of 3'-sulfogalactose and phosphotyrosine with regard to ligand binding can be found from a well studied example in the literature. The enzyme responsible for the degradation of SGC *in vivo* is termed arylsulfatase A. Deficiency in this enzyme activity results in the lysosomal storage disease, metachromatic leukodystrophy, due to intracellular SGC accumulation (43). The name "arylsulfatase" is used because the enzyme will also desul-

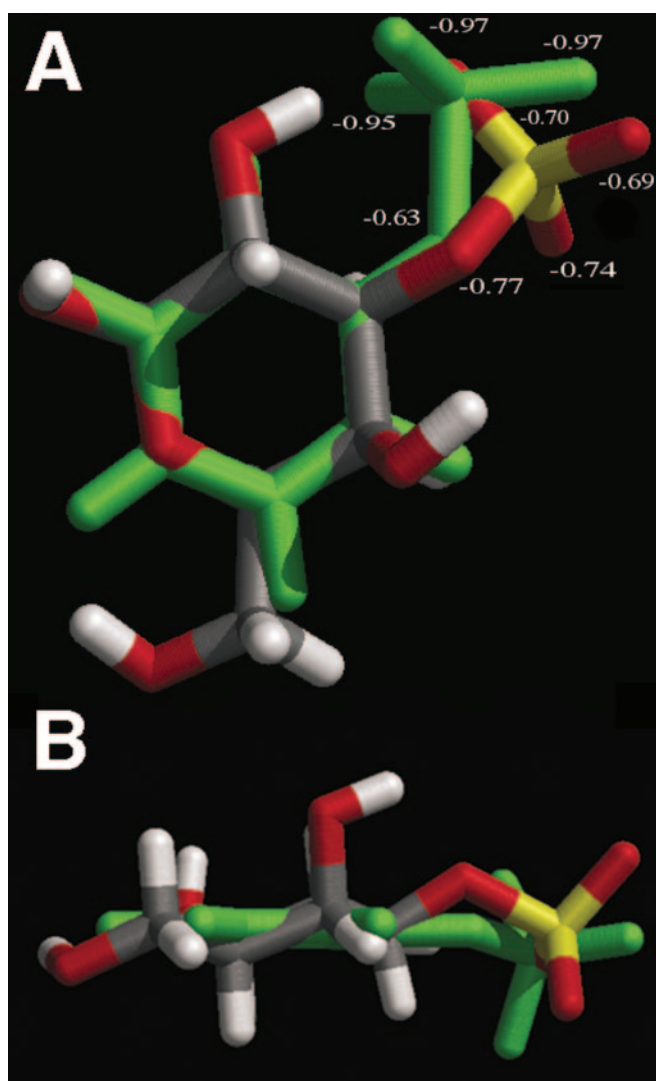


FIG. 2. Comparison of minimum energy conformations of 3'-sulfogalactose and phosphotyrosine. The minimum energy conformations of phosphotyrosine (green) and of 3'-sulfogalactose are superimposed. Point charges are shown for the O atoms of the sulfate and phosphate groups.

fate various aryl substrates. Indeed, this activity was demonstrated before the natural substrate (SGC) was identified (44). The several aryl sulfate substrates have a substituted tyrosine-like ring structure. Arylsulfatase has recently been found to have tyrosine phosphatase activity (45). Phosphotyrosine inhibits arylsulfatase-mediated SGC desulfation (46). Phosphotyrosine and methylumbelliferyl phosphate can bind within the aryl sulfate/SGC binding site to inhibit desulfation. Moreover, the phosphorylated enzyme intermediate in phosphotyrosine cleavage has been crystallized. The covalently bound phosphate was localized in the same position as the sulfate in the enzyme *p*-nitrocatechol sulfate co-crystal (47). The substrate binding site activity (45) is in a cleft lined primarily with polar residues (Fig. 3). Surprisingly, a non-polar domain within the cleft (Leu<sup>68</sup> and Val<sup>91</sup>) is uninvolved in binding the hydrophobic aryl substrates. Since SGC has not been crystallized with the enzyme, we have modeled 3'-sulfogalactose within the substrate binding site as defined by the co-crystal (Fig. 3) (28). The galactose ring occupies a position intermediate between the two conformations of the aryl ring in the co-crystal, and its hydroxyl groups can form several hydrogen bonds with the protein (Fig. 3, Table I). In addition, the 6'-CH<sub>2</sub> makes hydrophobic contact with Val<sup>91</sup> in the hydrophobic domain, providing additional binding energy for

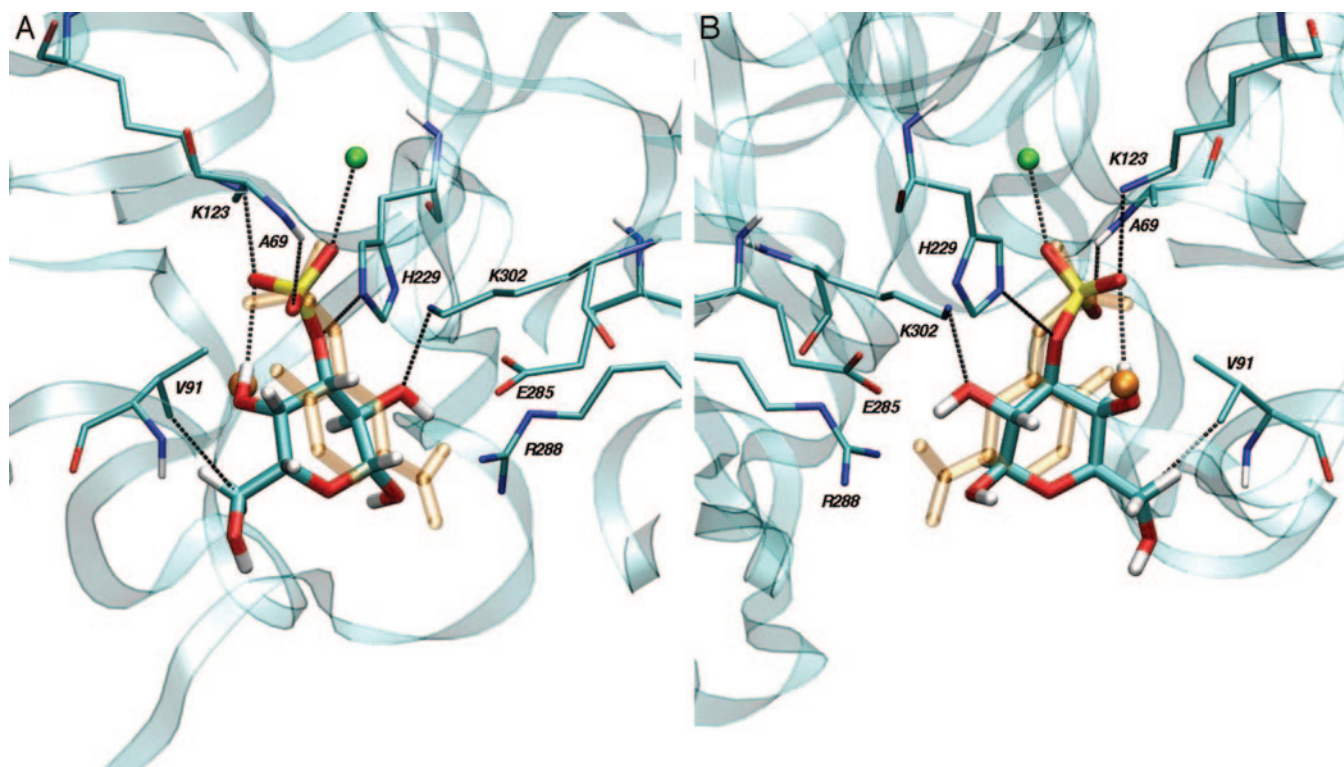


FIG. 3. Superimposition of 3'-sulfogalactose in the substrate binding site of arylsulfatase. Galactose 3'-sulfate was docked into the binding site of pNCS (pale orange) co-crystallized within the arylsulfatase substrate binding site. A, non-polar carbohydrate face. B, polar face view. Hydroxyl groups 2 and 4 can form several hydrogen bonds as highlighted with dashed lines (Table I). In addition, the methylene group at position 6 forms a hydrophobic contact with Val<sup>91</sup>, whereas the hydroxyl groups 2 and 6 are at least partly exposed to solvent. A water molecule immobilized in the pNCS co-crystal is shown in B, which is virtually coincident with the 4'-OH of the docked galactose sulfate. The mean position of the ring is intermediate between the two alternate orientations of the pNCS ring proposed in the crystallographically determined structure (28).

TABLE I  
Hydrogen bonds (HB) and non-polar contacts involving galactose sulfate docked in arylsulfatase A

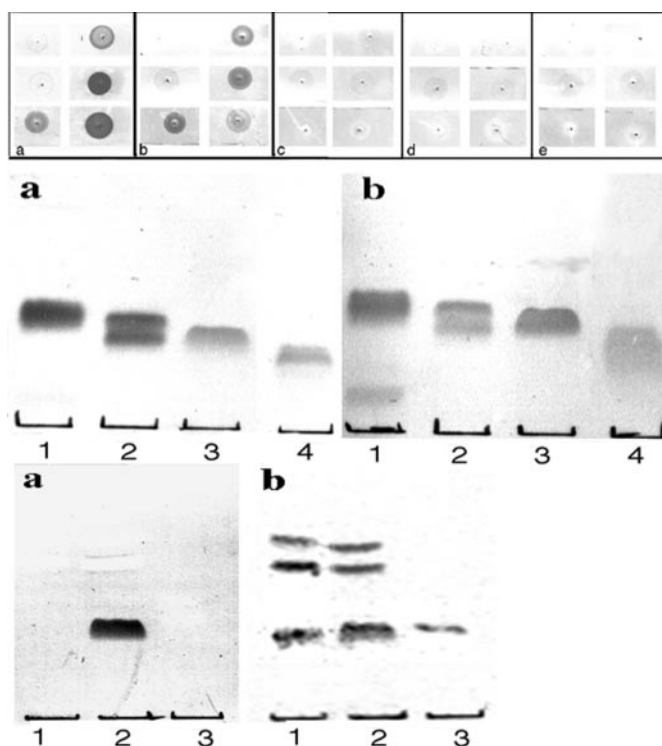
Atom 1	Atom 2	Separation	Comments
		nm	
Galactose atom (CHARMM)			
C-6	Val <sup>91</sup> , C $\gamma$ 2	0.32	Hydrophobic contact
O-4	His <sup>125</sup> , Ne2	0.37	Weak HB
O-4	Ser <sup>150</sup> , O $\gamma$	0.36	Weak HB
O-4	Sulfate OS3	0.31	Intramolecular HB
O-1	Arg <sup>288</sup> , N $\eta$ 2	0.35	HB
O <sub>2</sub>	Glu <sup>285</sup> , O $\epsilon$ 2	0.29	HB
O <sub>2</sub>	Lys <sup>302</sup> , N $\zeta$	0.27	HB
O-3 (sulfate)	Lys <sup>302</sup> , N $\zeta$	0.31	HB
O-3 (sulfate)	His <sup>229</sup> , Ne2	0.27	HB
OS1 (sulfate)	Lys <sup>123</sup> , N $\zeta$	0.27	HB
OS2 (sulfate)	Lys <sup>302</sup> , N $\zeta$	0.30	HB
OS2 (sulfate)	Ala <sup>69</sup> , NH (backbone amide)	0.34	HB

the natural SGL substrate (Fig. 3B). Thus, although the aryl ring provides the more non-polar substrate, it is the galactosyl substrate that utilizes the hydrophobic patch within the arylsulfatase substrate binding site. In the catechol sulfate/arylsulfatase co-crystal, a water molecule is resolved adjacent to the aryl ring, which is H-bonded to the sulfate oxygen to assist in the coordination of the bound sulfate. In our model of the sulfogalactosyl complex, the 4'-O of the polar surface of the galactose ring is superimposed on this water molecule and forms an intramolecular hydrogen bond with a sulfate oxygen (Fig. 3B and Table I). Thus, in the aryl substrate, a water molecule substitutes for the more hydrophilic character of the galactose as opposed to the catechol (or tyrosine?) ring. Thus, the hydrophobic difference between the aryl and the galactopyranose ring may be reduced, surprisingly by hydration of the aryl ring in the ligand-bound form to more closely mimic the asymmetric polarity of the galac-

topyranose. This solvation of the ligand-bound aryl ring is similar to the effect we have ascribed to the lipid moiety on GSL solvation for binding (48) and could similarly provide a basis for the effect of the lipid moiety on sulfotyrosine-PE binding.

Both the 3'- and (non-physiological) 4'-SGC are effectively bound by hsp70 (4' < 3') (10). Similarly, both catechol 3'- and 4'-sulfate are effectively bound by ligand (4' < 3') (49). This finding suggests that vicinal sulfates in either an aryl or sugar ring can present sulfate in an equivalent format for ligand binding.

In terms of hsp70/SGL binding, our model would suggest that the ATPase activity of the N-terminal domain of cytosolic (and other) hsp70s might be modulated by interaction with unknown cytosolic phosphotyrosine (or sulfotyrosine) containing polypeptides. This may provide an as yet undetected means to regulate hsp70 function. Tyrosine phosphorylation is an

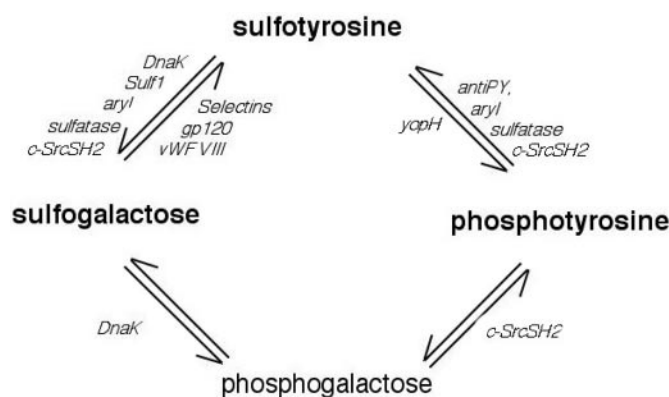


**FIG. 4. Sulfogalactolipid binding by SH2 domains.** Upper panels, dot blot TLC binding comparing Sulfl mAb anti-SGC (panel a) and GST fusion proteins of c-Src (panel b), SHP-1 C-terminal domain (panel c), SHP-1 N-terminal domain (panel d), and SHP-1 C+N-terminal SH2 domain (panel e). Left column from top, lipids are cholesterol sulfate, GM3 ganglioside, tyrSO<sub>4</sub>-DHD PE. Right column, 6'-SGC, 3'-SGC, tyrSO<sub>4</sub>-*E. coli* PE. Middle panels, a comparison of DnaK (panel a) and c-SrcSH2 domain (panel b) TLC overlay binding. Lane 1, GalCer, 3'-SGC and GM3. Lane 2, 4'-SGC. Lane 3, 6'-SGC. Lane 4, 6'-PGC. Each lane contains 1  $\mu$ g of GSL. Lower panels, c-SrcSH2 domain TLC overlay binding to tyrosine-PE derivatives. Panel a, c-SrcSH2 binding. Panel b, molybdenum blue staining for phosphate. Lane 1, from top, *E. coli* PE, tyr-*E. coli* PE, and tyrSO<sub>4</sub>-*E. coli* PE. Lane 2, from top, DHD PE, tyr-DHD PE, and tyrSO<sub>4</sub>-DHD PE. Lane 3, phosphatidylcholine + cholesterol sulfate.

older post-translational modification than sulfogalactosylation. Interaction of prokaryote hsp70s with phosphotyrosine motifs could provide an evolutionary basis of the conserved SGL binding that we have observed (9). Thus, under appropriate circumstances, sulfogalactose-based analogues might modify tyrosine kinase/phosphatase activity and signaling pathways.

**SH2 Domain Sulfogalactolipid Binding**—SH2 domains provide the central recognition mechanism for binding pY motifs mediating protein:phosphoprotein interactions in signal transduction (50). Recombinant fusion proteins of several SH2 domains (c-Src kinase and the tyrosine phosphatase, SHP-1, and N-terminal, C-terminal, and C+N-terminal SH2 domains) and GST were tested for GSL binding by TLC overlay (Fig. 4). Strong binding (equivalent to mAb anti-SGC and DnaK) to SGC (3' > 4'-SGC = 6'-SGC) was seen only for c-SrcSH2. This chimera was also the only SH2 domain to bind the sulfotyrosine-PE. As with the other ligands, c-SrcSH2/sulfotyrosine-PE binding was affected by the lipid. Binding was greater to the ether-PE than the acyl-PE conjugate (similar to DnaK and anti-pY). In the TLC-separated assay (Fig. 4, lower panel a), no binding to acyl PE-sulfotyrosine was seen. Significant binding of c-SrcSH2 to 6'-PGC (Fig. 4, middle panel b, lane 4) showed that, in this case, sulfogalactose can in part mimic pY receptor function.

These results validate the prediction that sulfogalactose can substitute for phosphotyrosine recognition. Demonstration of c-SrcSH2 as a carbohydrate binding motif identifies a potential



**SCHEME 1. Receptor mimicry between sulfogalactose and phosphotyrosine moieties: relationship between “intermediate” structures and binding ligands from the literature and the present work.**

new avenue for the generation of inhibitors of SH2 domain/phosphoprotein binding. In the c-SrcSH2-activated platelet-derived growth factor receptor co-crystal (51), a water molecule is coordinated in the binding of the (tyrosine) phosphate in a manner analogous to the aromatic sulfate that we highlighted in arylsulfatase. This could explain the selective binding of SGC by this SH2 domain. Solvation of the “receptor ring” may provide a link between GSL and pY recognition.

Furthermore, the co-crystal structure of a *Yersinia* tyrosine phosphatase (YopH) complexed with a small molecule inhibitor has been solved (52). This inhibitor is *p*-nitrocatachol sulfate, the same compound that was co-crystallized in the substrate binding site of arylsulfatase (Fig. 3). YopH is a homologue of SHP-1 and pNCS was found to be a competitive inhibitor of many such SHP-1 homologous tyrosine phosphatases (52). Because pNCS is bound by both arylsulfatase and YopH, we predict SGC binding for YopH. pNCS binds in a solvent-exposed positively charged cleft in YopH where solvation is likely to play a significant role. Indeed, Sun *et al.* (52) identify a bound water coordinating the charge in much the same way as is seen for arylsulfatase (and c-Src) and suggest that hydroxyl groups might be used as an advantage in the generation of more effective inhibitors. The hydrophobic interactions of the aryl ring with YopH occur from the more protein-adjacent face, whereas the polar and solvent-mediated interactions (with the exception of the direct coordination of the sulfate) occur from the other more solvent-exposed face, generating a polarity asymmetry such as might be mimicked by a sugar ring.

The overall relationship that we propose is summarized in Scheme 1. Our data establish a continuum of ligand binding from sulfogalactose to phosphotyrosine, either by tyrosine substitution for galactose or by phosphate substitution for sulfate. In terms of the latter, only one of the carbohydrate positional isomers has been studied thus far, the 6'-sulfo- and 6'-phosphogalactosylceramide. Phosphogalactose is an uncommon motif, and therefore limited physiological inferences can be drawn from its ligand binding activity. We propose primarily a structural relationship between sulfogalactose and phosphotyrosine. Although arylsulfatase sets a well studied precedent for binding the extremities of this spectrum, the c-SrcSH2 domain provides the first experimentally defined system that supports this “cross-reactive” relationship. While of central pharmacological interest, since carbohydrate may provide a scaffold for a new family of SH2 domain ligands/inhibitors, the physiological relevance of this molecular mimicry may be limited as GSLs and tyrosine phosphoproteins are generally found in separate cell compartments. However, cytosolic SGC has been reported previously (53). Increased SGC is associated with gastrointes-

tinal cancer (54), correlating with the up-regulation of Src in this disease (55). Because SGC can function as a partial structural mimic of phosphotyrosine, the effects microbial pathogens which bind SGC (4, 56), have on host cell tyrosine phosphorylation (57, 58) should be considered. Such a SrcSH2-mediated effect may have already been reported (59). Similarly, the effect of human immunodeficiency virus glycoprotein 120 on host cell tyrosine phosphorylation (60) could relate to its SGL binding.

Tyrosine sulfate is a more common cell surface post-translational modification. As indicated in Scheme 1, there are more examples, both from the literature and our present work, that indicate this may be the more physiologically relevant mimicry with SGLs. The fact that, like SGL biosynthesis (1, 61), tyrosine sulfation is important in male fertility (62) and that both tyrosine sulfation (17, 63) and SGC (64, 65) play important roles in platelet aggregation suggests arenas in which a search for a functional consequence of this relationship may be fruitful.

## REFERENCES

- Honke, K., Hirahara, Y., Dupree, J., Suzuki, K., Popko, B., Fukushima, K., Fukushima, J., Nagasawa, T., Yoshida, N., Wada, Y., and Taniguchi, N. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99**, 4227–4232
- Lingwood, C. A., Hay, G., and Schachter, H. (1981) *Can. J. Biochem.* **59**, 556–563
- Boulanger, J., Faulds, D., Eddy, E. M., and Lingwood, C. A. (1995) *J. Cell. Physiol.* **165**, 7–17
- Huesca, M., Borgia, S., Hoffman, P., and Lingwood, C. A. (1996) *Infect. Immun.* **64**, 2643–2648
- Hartmann, E., and Lingwood, C. A. (1997) *Infect. Immun.* **65**, 1729–1733
- Mamelak, D., and Lingwood, C. (1997) *Glycoconj. J.* **14**, 715–722
- Mamelak, D., and Lingwood, C. (2001) *J. Biol. Chem.* **276**, 449–456
- Whetstone, D., and Lingwood, C. (2003) *Biochemistry* **42**, 1611–1617
- Mamelak, D., Mylvaganam, M., Whetstone, H., Hartmann, E., Lennarz, W., Wyrick, P., Raulston, J., Han, H., Hoffman, P., and Lingwood, C. (2001) *Biochemistry* **40**, 3572–3582
- Mamelak, D., Mylvaganam, M., Tanahashi, E., Ito, H., Ishida, H., Kiso, M., and Lingwood, C. (2001) *Carbohydr. Res.* **335**, 91–100
- Lingwood, C. A. (1996) *Glycoconj. J.* **13**, 495–503
- Bhat, S., Mettus, R., Reddy, E., Ugen, K., Srikanthan, V., Williams, W., and Weiner, D. (1993) *AIDS Res. Hum. Retroviruses* **9**, 175–181
- Roberts, D. D., Williams, S. B., Gralnick, H. R., and Ginsburg, V. (1986) *J. Biol. Chem.* **261**, 3306–3309
- Bajorath, J., Hollenbaugh, D., King, G., Harte, W., Jr., Eustice, D. C., Darveau, R. P., and Aruffo, A. (1994) *Biochemistry* **33**, 1332–1339
- Konno, A., Nunogami, K., Wada, T., Yachie, A., Suzuki, Y., Takahashi, N., Suzuki, T., Miyamoto, D., Kiso, M., Hasegawa, A., and Miyawaki, T. (1996) *Int. Immunol.* **8**, 1905–1913
- Farzan, M., Vasilieva, N., Schnitzler, C. E., Chung, S., Robinson, J., Gerard, N. P., Gerard, C., Choe, H., and Sodroski, J. (2000) *J. Biol. Chem.* **275**, 33516–33521
- Dong, J. F., Li, C. Q., and Lopez, J. A. (1994) *Biochemistry* **33**, 13946–13953
- Galustian, C., Childs, R. A., Stoll, M., Ishida, H., Kiso, M., and Feizi, T. (2002) *Immunology* **105**, 350–359
- Cunto-Amesty, G., Luo, P., Monzavi-Karbassi, B., and Kieber-Emmons, T. (2001) *Int. Rev. Immunol.* **20**, 157–180
- Fredman, P., Mattsson, L., Andersson, K., Davidsson, P., Ishizuka, I., Jeansson, S., Mansson, J.-E., and Svennerholm, L. (1988) *Biochem. J.* **251**, 17–22
- Tenev, T., Keilhack, H., Tomic, S., Stoyanov, B., Stein-Gerlach, M., Lammers, R., Krivtsov, A. V., Ullrich, A., and Bohmer, F. D. (1997) *J. Biol. Chem.* **272**, 5966–5973
- Anderson, S. K., Gibbs, C. P., Tanaka, A., Kung, H. J., and Fujita, D. J. (1985) *Mol. Cell. Biol.* **5**, 1122–1129
- Tanaka, A., Gibbs, C. P., Arthur, R. R., Anderson, S. K., Kung, H. J., and Fujita, D. J. (1987) *Mol. Cell. Biol.* **7**, 1978–1983
- Ma, X. Z., Jin, T., Sakac, D., Fahim, S., Zhang, X., Katsman, Y., Bali, M., and Branch, D. R. (2003) *Exp. Hematol.* **31**, 131–142
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakrzewski, V. G., Montgomery, J. A., Stratmann, J. R. E., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Farkas, M. C. S. O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Baboul, A. G., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P. M. W., Johnson, B. G., Chen, W., Wong, M. W., Andres, J. L., Head-Gordon, M., Replogle, E. S., and Pople, J. A. (1998) *Gaussian*, Gaussian, Inc., Wallingford, CT
- Zajonc, D. M., Elsliger, M. A., Teyton, L., and Wilson, I. A. (2003) *Nat. Immunol.* **4**, 808–815
- Breneman, C. M., and Wilberg, K. B. (1990) *J. Comput. Chem.* **11**, 361–373
- von Bulow, R., Schmidt, B., Dierks, T., von Figura, K., and Uson, I. (2001) *J. Mol. Biol.* **305**, 269–277
- Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem.* **4**, 187–217
- MacKerell, A. D. J., Bashford, D., Bellot, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reither, W. E., Roux, I. B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., and Karplus, M. (1998) *J. Phys. Chem. B* **102**, 3586–3616
- Ha, S. N., Giammona, A., Field, M., and Brady, J. W. (1988) *Carbohydr. Res.* **180**, 207–221
- Kamps, M. P. (1991) *Methods Enzymol.* **201**, 101–110
- Glennay, J. R., Jr., Zokas, L., and Kamps, M. P. (1988) *J. Immunol. Methods* **109**, 277–285
- Donella-Deana, A., Varro, A., Dockray, G. J., and Pinna, L. A. (1991) *Biochim. Biophys. Acta* **1095**, 75–77
- Jones, J. A., and Cushley, W. (1985) *Biosci. Rep.* **5**, 707–714
- Jones, J. A., Wood, A., and Cushley, W. (1986) *Biosci. Rep.* **6**, 265–273
- Craven, R. J., Lightfoot, H., and Cance, W. G. (2003) *Surg. Oncol.* **12**, 39–49
- Wang, W. Q., Sun, J. P., and Zhang, Z. Y. (2003) *Curr. Top. Med. Chem.* **3**, 739–748
- Urban, J. E., Klopfenstein, W. E., Ahmad, K., and Baines, J. D. (1981) *Can. J. Microbiol.* **27**, 1283–1289
- Zhang, Z. Y., Thieme-Seifler, A. M., Maclean, D., McNamara, D. J., Dobrusin, E. M., Sawyer, T. K., and Dixon, J. E. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 4446–4450
- Sipos, L., and von Heijne, G. (1993) *Eur. J. Biochem.* **213**, 1333–1340
- Lingwood, C. A. (1999) *Trends Glycosci. Glycotech.* **11**, 1–9
- Gieselmann, V., Matzner, U., Hess, B., Lullmann-Rauch, R., Coenen, R., Hartmann, D., D'Hooge, R., DeDeyn, P., and Nagels, G. (1998) *J. Inherit. Metab. Dis.* **21**, 564–574
- Roy, A. B. (1979) *CIBA Found. Symp.* 177–190
- Chruszcz, M., Laidler, P., Monkiewicz, M., Ortlund, E., Lebioda, L., and Lewinski, K. (2003) *J. Inorg. Biochem.* **96**, 386–392
- Stoecker, P. W., Manowitz, P., Harvey, R., and Yacynych, A. M. (1998) *Anal. Biochem.* **258**, 103–108
- Boltes, I., Czapsinska, H., Kahmert, A., von Bulow, R., Dierks, T., Schmidt, B., von Figura, K., Kertesz, M. A., and Uson, I. (2001) *Structure (Camb.)* **9**, 483–491
- Mylvaganam, M., and Lingwood, C. A. (2003) in *Carbohydrate-based Drug Discovery* (Wong, C.-H., ed) Vol. 2, pp. 761–780, Wiley-VCH Press, Weinheim, Germany
- Strobel, G., Werle, E., and Weicker, H. (1990) *Biochem. Int.* **20**, 343–351
- Pawson, T. (2004) *Cell* **116**, 191–203
- Lubman, O. Y., and Waksman, G. (2003) *J. Mol. Biol.* **328**, 655–668
- Sun, J. P., Wu, L., Fedorov, A. A., Almo, S. C., and Zhang, Z. Y. (2003) *J. Biol. Chem.* **278**, 33392–33399
- Berntson, Z., Hansson, E., Rönnbäck, L., and Fredman, P. (1998) *J. Neurosci. Res.* **52**, 559–568
- Morichika, H., Hamanaka, Y., Tai, T., and Ishizuka, I. (1996) *Cancer* **78**, 43–47
- Malek, R. L., Irby, R. B., Guo, Q. M., Lee, K., Wong, S., He, M., Tsai, J., Frank, B., Liu, E. T., Quackenbush, J., Jove, R., Yeatman, T. J., and Lee, N. H. (2002) *Oncogene* **21**, 7256–7265
- Lingwood, C. A., Quinn, P. A., Wilansky, S., Nutikka, A., Ruhnke, H., and Miller, R. B. (1990) *Biol. Reprod.* **43**, 694–697
- Rawadi, G., and Roman-Roman, S. (1996) *Infect. Immun.* **64**, 637–643
- Mitchell, D. J., Huynh, H. Q., Ceponis, P. J., Jones, N. L., and Sherman, P. M. (2004) *Infect. Immun.* **72**, 537–545
- Azuma, T., Yamazaki, S., Yamakawa, A., Ohtani, M., Muramatsu, A., Suto, H., Ito, Y., Dojo, M., Yamazaki, Y., Kuriyama, M., Keida, Y., Higashi, H., and Hatakeyama, M. (2004) *J. Infect. Dis.* **189**, 820–827
- Phipps, D. J., Read, S. E., Piovesan, J. P., Mills, G. B., and Branch, D. R. (1996) *Aids* **10**, 1191–1198
- Weerachatanukul, W., Rattanachaiyanont, M., Carmona, E., Furimsky, A., Mai, A., Shoushtarian, A., Sirichotiyaku, S., Ballakier, H., Leader, A., and Tanphaichitr, N. (2001) *Mol. Reprod. Dev.* **9999**, 1–10
- Moore, K. L. (2003) *J. Biol. Chem.* **278**, 24243–24246
- Andrews, R. K., Lopez, J. A., and Berndt, M. C. (1997) *Int. J. Biochem. Cell Biol.* **29**, 91–105
- Merten, M., and Thiagarajan, P. (2001) *Circulation.* **104**, 2955–2960
- Borthakur, G., Cruz, M. A., Dong, J. F., McIntire, L., Li, F., Lopez, J. A., and Thiagarajan, P. (2003) *J. Thromb. Haemostasis* **1**, 1288–1295
- Banville, D., Stocco, R., and Shen, S. H. (1995) *Genomics*, 165–173
- Palma, R., Zuccato, P., Himmel, M. E., Liang, G., and Brady, J. W. (2001) in *Glycosyl Hydrolases for Biomass Conversion* (Himmel, M. E., Baker, J. O., and Saddler, J. N., eds) ACS Symposium Series 769, pp. 112–130, American Chemical Society, Washington, D. C.