Glycoconj J (2009) 26:1029–1045
DOI 10.1007/s10719-008-9222-9

Interaction of arylsulfatase-A (ASA) with its natural sulfoglycolipid substrates: a computational and site-directed mutagenesis study

Matthias Schenk · Chaitanya A. K. Koppisetty · Daniela Costa Santos · Euridice Carmona · Smita Bhatia · Per-Georg Nyholm · Nongnuj Tanphaichitr

Received: 24 June 2008 / Revised: 6 November 2008 / Accepted: 9 December 2008 / Published online: 18 April 2009

Abstract Arylsulfatase A (ASA) hydrolyzes sulfate esters with a pH optimum of 5. Interactions between *p*-nitrocatechol sulfate (NCS, artificial substrate) and active site residues of ASA are revealed from their co-crystal structure. Since equivalent ASA interactions with its natural substrates, sulfogalactosylceramide (SGC) and sulfogalactosylglycerolipid (SGG), are not known, we computationally docked SGC/SGG to the ASA crystal structure. Our dockings suggested that Cys69 was the active site residue, and Lys302 & Lys123 as residues anchoring the sulfate group of SGC/SGG to the active site, as observed for NCS. We further confirmed these results using 2 recombinant ASA mutants: C69A and CKK (Cys69, Lys302 and Lys123-all mutated to Ala). Both ASA mutants failed to desulfate SGC/SGG, and CKK showed minimal binding to [^{14}C]SGC, although C69A still had affinity for this sulfoglycolipid. However, our dockings suggested additional intermolecular hydrogen bonding and hydrophobic interactions between ASA and SGC/SGG, thus contributing to the specificity of SGC/SGG as natural substrates.

Keywords Sulfatase · Arylsulfatase-A · Sulfogalactosylglycerolipid · Sulfogalactosylceramide · Galactosyl sulfate · Structural docking · pK_a calculation · Enzyme–substrate interaction

M. Schenk · D. C. Santos · N. Tanphaichitr (✉)
Chronic Disease,
Ottawa Hospital Research Institute,
Ottawa, ON, Canada
e-mail: ntanphaichitr@ohri.ca

M. Schenk · C. A. K. Koppisetty · P.-G. Nyholm
Biognos AB,
Göteborg, Sweden

E. Carmona
Maisonneuve-Rosemont Hospital Research Centre
and Department of Medicine, University of Montreal,
Montreal, Canada

S. Bhatia
Institute for Biological Sciences,
National Research Council of Canada,
Ottawa, ON, Canada

P.-G. Nyholm (✉)
Department of Medical Biochemistry, Göteborg University,
Göteborg, Sweden
e-mail: nyholm@medkem.gu.se

N. Tanphaichitr
Departments of Obstetrics/Gynecology and Biochemistry/
Microbiology/Immunology, University of Ottawa,
Ottawa, ON, Canada

Arylsulfatase-A (ASA; *EC 3.1.6.8*) is known as a lysosomal/acrosomal enzyme and as belonging to the family of sulfatases. Arylsulfatase-A has the specific ability to desulfate small arylsulfates, such as *p*-nitrocatechol sulfate (NCS), as well as sulfoglycolipids. *In vivo*, sulfogalactosylceramide (SGC; also known as cerebroside sulfate) has been revealed as a natural substrate of ASA by its accumulation in the brain of individuals suffering from the neurological disorder metachromatic leukodystrophy (MLD), due to genetic deficiency of ASA [1, 2]. The deficiency of ASA's activator protein, saposin-B, also leads to MLD [3]. An SGC-analog, sulfogalactosylglycerolipid (SGG, also termed seminolipid), is found on the sperm head anterior surface overlying the acrosome and its involvement in sperm-egg interaction

has been demonstrated [4, 5]. In acrosome-intact sperm, desulfation of SGG has not yet been reported, but it is likely to occur at some stage during acrosomal exocytosis [6]. *In vitro*, ASA can desulfate detergent-presolubilized SGC and SGG, with the pH optimum of 4.5–5.5. At this pH, ASA exists as a highly stable octamer, whereas it is mainly present as a homodimer at neutral pH [7]. X-ray crystallography has been performed on human recombinant (rec) wild-type and mutant ASA, as well as on the native form [8–10]. The crystal structure of recASA complexed with the artificial substrate NCS was achieved by using an inactive mutant, with the active site residue Cys69 replaced with alanine (termed C69A). This mutation still allows NCS binding, but fully prevents its desulfation, thus retaining the intact substrate in the active site pocket. In wild-type ASA, this cysteine is post-translationally transformed into formylglycine [11], which—in its oxidized form—is crucial for the hydrolysis of the sulfate group of the substrate and its subsequent release from the active site pocket. All sulfatases studied have an equivalent cysteine residue (or serine in some prokaryotes), and all are converted into formylglycine by the formylglycine-generating enzyme (FGE) [12, 13]. Apart from Cys69, the following active site amino acids have been suggested to interact with the sulfate group of NCS: Lys123, His229, Lys302, and Ser150. The individual alanine-mutants of the two lysines, Lys123 and Lys302, have been shown to have a much higher K_m than wild-type ASA, suggesting that these two residues are important for NCS binding [14]. The coordination of a magnesium ion in the center of the active site is also of great importance. Interestingly, calcium rather than magnesium is found in the crystal structure of native human placental ASA [10], which is in agreement with other crystallized sulfatases. To date, it is not known whether the same interactions are involved in binding the sulfate groups of SGC and SGG. These interactions need to be discerned, given the fact that the sulfoglycolipids possess hydrocarbon chains, which are absent in NCS, and their galactosyl sulfate head group is significantly different to the sulfated aryl ring of NCS. The co-crystallization of ASA and SGC/SGG has not yet been reported and might be difficult to achieve. Therefore, computational docking studies supported by site-directed mutagenesis of ASA are important approaches for understanding molecular interactions between ASA and SGG/SGC. Only recently, the interaction of the sulfated galactose head group with the active site pocket of ASA has been addressed through a computational approach by Lingwood's group [15]. However, the docking of galactosyl sulfate to ASA was only achieved by forcing the sulfate group into the same position as that of the co-crystallized NCS. Proper docking would require the treatment of SGC/SGG with a higher degree of freedom and be performed at the acidic pH where ASA is enzymatically active. Therefore, any pH-dependent dockings have to be preceded by detailed pK_a calculations for the titratable amino acid residues around the active site in order to obtain accurate charges and protonation states. We took these points into account in our computational study described herein, and our results were further supported by site-directed mutagenesis of ASA. We anticipate that our study will lead to a discovery of further natural substrates as well as enzyme inhibitors, which could be used to answer the question of whether ASA activity is essential for fertilization.

Experimental procedures

Molecular structures

The ASA crystal structures at pH 5 were obtained from the RCSB Protein Data Bank (PDB). The structure of the inactive Cys69Ala mutant (C69A) (*1E2S*), co-crystallized with the artificial substrate *p*-nitrocatechol sulfate (NCS) [9] was used for all dockings. This mutant structure contains alanine in place of formylglycine at position 69. Formylglycine (FGly) is a non-standard residue and could not be treated appropriately in all the computational protocols used for this study. Also, the position of NCS in the C69A mutant structure could be used as a reference for ligand docking. As an initial preparation step, all ligand and water molecules were removed from the PDB file. Polar hydrogen atoms and charges were obtained using the Sybyl program (Tripos Inc., St. Louis, MO, USA). However, the crystallized human placenta ASA (*1N2K*), which is the only ASA structure possessing the complete FGly residue [10], was used for evaluation of the dockings with *1E2S*. In this structure, due to the pre-soaking of the ASA crystal with 4-methylumbelliferyl phosphate (a known inhibitor of ASA), a phosphate group is covalently bound to FGly69 [10]; a sulfate group from ASA's substrate would be expected to be bound and oriented in a very similar fashion during the hydrolysis.

Various ligands were prepared for docking into the ASA active site as follows. The NCS molecule was obtained from the mutant ASA structure (C69A [9] and set up for docking using AutoDockTools (ADT). A sulfated galactosyl ring, the head group of SGC and SGG, was built starting from a galactose molecule, obtained from the CERMAV monosaccharide database (www.cermav.cnrs.fr/glyco3d/index.php). The modeling of the sulfate group and basic energy minimization was performed in Sybyl. To obtain the structures of SGC or SGG, the sulfated galactose was extended with a backbone and short lipid chains (see Figs. 6 and 7) using Sybyl. Further ligand preparation, including the distribution of atom charges, was performed in ADT.

pK_a calculations

pK_a values (defined as the pH values at which a titrating group is 50% protonated) of titratable amino acids in the active site of ASA were calculated taking into account the influence of the surrounding amino acids. Hence, some titratable residues might have pK_a values that deviate strongly from their amino acid standard pK_a, due to charge interaction with surrounding residues. The program WhatIf [16] was used in conjunction with DelPhi [17] to obtain the pK_a values for the titratable residues of ASA. Resulting pK_a values were expected to have an accuracy of ±1.5 pH unit (according to the introduction to pK_a calculations on http://enzyme.ucd.ie/Science/pK_a/). The calculations were done with the default solvent dielectric constant 80 and internal dielectric constant 8, an ion exclusion layer of 2.0 Å, an ionic strength of 0.144 M, and a surface probe with a 1.4 Å radius. The ASA structure used for the calculations was the mutant form C69A, as WhatIf does not allow the inclusion of non-standard residues such as formylglycine. In the initial run (defined as "default run" in Table 1), the magnesium ion was not included. Calculations were later performed with the bound active site Mg^{2+} ion treated as a ligand (defined as "run with Mg^{2+}" in Table 1).

Charge sets

Based on the obtained pK_a results, the protonation states of all active site residues were defined before applying atom charges and protonating hydrogens obtained using PDB2PQR [18]. To obtain an accurate charge distribution for pH 5, only those titrating groups that had a calculated pK_a value above 5 were assigned to be protonated. Similarly, a pK_a value above 7 was required for protonation at neutral pH. The Poisson–Boltzmann solving program APBS [19–21] was used to calculate the electrostatic potentials of ASA at both pH 5 and pH 7. The resulting potentials were displayed on the ASA surface within Pymol [22] to compare the surface charges.

When preparing the macromolecule file for the dockings, Sybyl was used to apply Kollman charges. Solvent parameters were added using the Addsol function of the AutoDock suite.

Table 1 Calculated pK_a values for residues around the active site pocket

Group	Default run	Run with Mg^{2+}[a]
Asp29	−10.36	**−15.83**
Asp30	−12.99	**−24.29**
Arg73	27.74	**23.05**
Lys123	22.19	**18.35**
His125	−0.38	**−2.71**
Asp152	−0.60	**−0.72**
His229	**7.34**	3.95
Asp281	14.04	**−30.13**
Glu285	−3.84	**−3.95**
Arg288	18.54	**18.75**
Lys302	22.50	**23.07**
His405	5.03	**5.16**

pK_a values used for the dockings are highlighted in bold
[a] Magnesium was included as a ligand in the WhatIf/DelPhi run

Dockings

Initially, AutoDock (AD) version 3.05 [23] was used for all dockings. By default, AutoDock uses atom charges at neutral pH, whereas desulfation of substrates by ASA occurs preferentially at pH ~5. Therefore, electrical charges of all residues were first assigned according to their calculated pK_a values as described above.

The Lamarckian genetic algorithm (LGA) simulation was used in all of the AD calculations. After having scanned the whole protein surface (as a monomer and a dimer) using AD, the major predicted binding site for each ligand was identical with the known active site. The site-specific docking runs were then performed in two cycles. For the first cycle, the grid was assigned to have the dimensions 45/45/45 (or 55/55/55 for the larger ligands SGC and SGG) at 0.375 Å per unit, and it was centered on the co-crystallized NCS molecule. The maximal translational step was kept at 2.0 Å, and the quaternion angle step and the dihedral angle were both changed to 25°. The population size was set to 100 individuals and the number of evaluations was to five million (or 7.5 million for SGC and SGG) for 200 GA (genetic algorithm) runs. The best resulting conformation (based on lowest docking energy and cluster size) was picked as the ligand starting conformation for the second cycle.

Cycle 2 was run with a grid size of 65/65/65 (or 81/81/81 for SGC and SGG) at 0.2 Å per unit, and it was centered on the top docked pose from the first run. The translational step was reduced to 1.5 Å. The quaternion angle step as well as the dihedral angle step was set to 10° with a reduction factor of 0.99. The population size was set to 100 individuals. Two hundred GA runs were performed with 7.5 million (or 10 million for SGC and SGG) energy evaluations per run.

Docking results were analyzed based on the calculated binding energies, using automatic clustering with ADT. Molecular graphics images were produced using either Pymol or the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081) [24]. Distances between atoms were

measured using Chimera and Sybyl (Tripos Inc., St Louis USA). The Glide quantum mechanics (QM) polarized ligand docking (QPLD) [25] method was then used to complement the AD results. QPLD involves a dynamic estimation of charges on the ligand atoms from QM/Molecular mechanics (MM) calculations in the presence of protein instead of using fixed charges from a force field. With QPLD, the charges on the ligands were calculated using density functional theory (DFT), a robust *ab initio* quantum chemical approach. In this approach, the polarization effects on the ligand, due to the protein environment, were considered and the charges on the ligand were recalculated at intervals during the docking. ASA has a magnesium ion with a positive +2 charge in its active site and it could have significant polarization effects on the substrates. Only the QPLD calculation, and not the AD methodology, could account for such a dynamic variation.

Hydrogens were added to the ASA substrates, NCS, galactose sulfate, SGC and SGG and the epik [26] program was used to predict the protonation states at pH 5 and the geometry was energy minimized. The above steps for substrate structure preparation were done using the ligprep program in the interface Maestro software (Schrödinger LLC, New York). The macromolecule, ASA was protonated in Maestro for pH 5, considering the pK_a values of its amino acids. The charges from the optimized potential for liquid simulations 2001 (OPLS-2001) [27] force field were assigned and the hydrogens and the protein side chain atoms were optimized using restrained prime-impref minimization methodology (Prime 2.0, Schrödinger LLC) with maximum permissible RMSD of 0.32 Å from the crystal coordinates. Initial Glide dockings were done with a grid size of 12 Å using the Glide–SP docking methodology [28, 29]. The top two poses from the initial dockings were considered for recalculation of charges through QPLD using the Qsite 5.0 quantum mechanics package from Schrödinger LLC. In the subsequent docking steps also with Glide-SP/QPLD, the dynamically recalculated charges together with parameters from the OPLS2001 force field were used.

Site-directed mutagenesis

Our cloned pig testis ASA cDNA (Gene Bank AF316108.2; GI: 45686370) [30] was used in site directed mutagenesis and subsequent mammalian cell transfection for rec ASA production. Its overall identity to human ASA (Gene Bank NM_000487; GI: 7262293) [31] is 91%, whereas the identity of the amino acids forming the active site pocket of pig testis and human ASA is 100%. Pig testis ASA cDNA in a pBluescript II SK(−) plasmid (UniZAP system from Stratagene, La Jolla, CA, USA) was sub-cloned into pcDNA3.1(+) (Invitrogen, Carlsbad, CA, USA) for mammalian cell expression, using appropriate restriction endonucleases. The inserted ASA cDNA was then entirely sequenced using an ABI DNA analyzer (Applied Biosystems, Foster City, CA, USA) at the Ontario Genomics Innovation Centre, Ottawa, Ontario, Canada. Site-directed mutagenesis was performed, according to the manufacturer's instructions, using the QuickChange XL kit from Stratagene, which allows site-specific mutation in a double stranded plasmid. For the single mutant C69A, sense and antisense primers of the following sequence were used to replace the cysteine codon, TGC, by the alanine codon, GCC: 5'CGGTGTCTCTG*GCC* ACACCCTCCCG3'. The obtained C69A plasmid was then used as the template for the double mutation C69A/K123A. Primers for the sequence 5'GGGATGGCTGGC*GCC*TGG CACCTTGGG3' were used to replace the lysine codon, AAG, by the alanine codon. The triple mutant C69A/K123A/K302A (CKK) was finally obtained using this C69A/K123A plasmid as the template with sense and antisense primers for the sequence 5'CCTACGATGCG GA*GCC*GGAACCACTTTTG3', also replacing the AAG lysine codon by the alanine codon. For all mutations, positive constructs, screened by differential restriction patterns, were re-sequenced to confirm nucleotide replacement.

Mammalian cell expression and purification of recombinant pig testis arylsulfatase-A and mutants

Chinese hamster ovary-K1 (CHO-K1) cells (American Type Culture Collection, Manassas, VA) were grown and maintained in "complete medium", containing Gibco OptiMEM-I medium (Invitrogen, Burlington, ON, Canada), supplemented with 2% fetal bovine serum (FBS), 1% L-glutamine, 0.1% penicillin/streptomycin and 0.1% gentamycin. Wild type ASA cDNA, and mutants C69A and CKK, inserted into the pcDNA3.1 expression vector, as well as the control empty pcDNA3.1, were used to transfect CHO-K1 cells. Transient transfection was performed in 100-mm culture dishes, having 70% confluent cells, with 3.5 μg of plasmid DNA and 20 μl lipofectamine reagent (Invitrogen) in OptiMEM-I medium for 5 h, at 37°C, under 5% CO_2. After this period, the incubation medium was replaced by the complete one and the cells were allowed to grow for 48 h, at 37°C, under 5% CO_2. In order to induce secretion of ASA (a lysosomal protein) into the culture medium, cells were washed with phosphate buffered saline (PBS), cultured in serum-free medium for 6 h and then incubated with new serum-free medium containing 10 mM ammonium chloride for 18 h [32]. The resulting culture medium, containing mainly over-expressed rec ASA (called wild type, C69A, CKK and pcDNA from cells transfected with pcDNA3.1 containing unaltered ASA cDNA, C69A construct, CKK construct and blank pcDNA3.1, respectively), was collected in tubes containing protease inhibitor cocktail (Complete Mini; Roche,

Laval, QC, Canada), centrifuged at low speed to remove contaminating cells; an aliquot was then assayed for NCS desulfation activity. Notably, the amount of total proteins in the medium of transfected CHO-K1 cells was the same for all plasmids used for transfection. In the initial set of experiments, transfected cells were scraped from the plate in PBS and homogenized in SDS-PAGE sample buffer [33]. Both the collected medium and the SDS-solubilized cells were subjected to immunoblotting for ASA detection (see below).

In the subsequent experiments, ASA was further purified from the collected medium. The medium was dialyzed against 10 mM ammonium bicarbonate using Spectra/Por tubing (molecular weight cutoff: 50 kDa; VWR International, Mississauga, ON). The dialyzed sample was then lyophilized and stored at −20°C. For further purification, the sample was first resuspended in 20 mM Tris-maleate, pH 6.1, and then incubated with Mono-Q-Sepharose Fast Flow resin (Amersham Bioscience, Uppsala, Sweden) at 4°C for 1 h with gentle shaking. Unbound proteins were removed in four washes of 20 mM Tris–maleate, pH 6.1, and bound proteins were eluted from the resin using 2 M sodium chloride in 20 mM Tris–maleate, pH 6.1. Hereafter, the sample was concentrated using Microcon YM-100 centrifugal filter device (molecular weight cutoff: 100 kDa; Amicon Bioseparations, Millipore Corporation, Bedford, MA, USA). The retentate containing rec ASA was collected from the membrane in 30–50 μl of 0.25 M sodium acetate, pH 5, and an aliquot was quantified for proteins using Bio-Rad Protein Assay Solution (Bio-Rad Laboratories, Hercules, CA, USA). Another aliquot of the retentate was assessed for the SGG/SGC desulfation activity (see below). The remainder of the retentate was stored at −20°C until further use.

The retentate of C69A and CKK was further purified for the experiments on SGC binding (see below). It was subjected to Sephacryl S-300 chromatography (Amersham Bioscience) using a mini-column (size, inner diameter, 6 mm; length, 5 cm). After equilibrating the column with 150 mM NaCl, 2 mM EDTA in 50 mM sodium acetate, pH 5, the ASA sample in 50 μl was loaded and the column was eluted with 1 ml of the same buffer at the rate of 25 μl/min. Fractions (45 μl each) were collected and probed for the presence of ASA by immunoblotting. The fractions containing the mutant ASA were pooled and concentrated with centrifugation through a Microcon YM-30, and an aliquot of the retentate was subjected to SDS-PAGE (10% acrylamide) [33], followed by silver staining using the Bio-Rad Silver Stain kit. The amount of the purified mutant ASA was also assessed by staining the gel with RAPIDstain (G-Biosciences, St. Louis, MO, USA) and comparing its staining intensity with that of a bovine serum albumin (BSA) standard of various amounts.

NCS desulfation activity and immunoblotting of ASA

Both of these procedures were performed as described previously [30]. One unit of ASA activity was defined as 1 μmol of NCS hydrolyzed per hour. ASA activity was measured for the culture medium collected from the transfected cells as well as for each step of the purification. Rabbit polyclonal antiserum monospecific to ASA, produced as previously described [30], was used for immunoblotting.

SGC/SGG desulfation

SGC/SGG desulfation by native sperm surface ASA (positive control, purified as previously described [30]) and partially purified rec wild-type and mutant ASA (C69A and CKK) was carried out *in vitro* in a 50 μl reaction containing 8 μg of SGC (a mixture of hydroxylated and non-hydroxylated forms purified from bovine brain; Sigma, St. Louis, MO) or SGG (purified in our lab as previously described [34], 50 μg of taurodeoxycholate and 2 μg of fatty-acid free BSA in 250 mM sodium acetate, pH 5. For native ASA and wild type rec ASA, 0.5 U (as determined by NCS desulfation) of the enzyme was added into the reaction mixture. An equivalent amount of the ASA mutant (C69A or CKK) to that of wild type ASA, as assessed by immunoblotting (~14 μg of total proteins of the partially purified ASA samples), was used in the reaction mixture. Due to the presence of a minimum amount of endogenous ASA in CHO-K1 cells, a reaction mixture was also set up for pcDNA, using the medium collected from the same number of cells as those transfected with wild type-, C69A- and CKK-containing pcDNA3.1, and processed through the same purification methods. The reaction mixture was incubated at 37°C for 16 h, and then subjected to lipid extraction according to the modified [35] Bligh and Dyer's [36] method. The extracted lipids were concentrated under a stream of N_2 and separated by high performance thin layer chromatography (HPTLC; using a silica plate with 200 μm mesh size, with a dimension of 10×10 cm; EMD Chemicals Inc., Gibbstown, Germany); the solvent system was chloroform/methanol/water (65:25:4, *v*/*v*/*v*) [37]. Purified SGG and SGC as well as their desulfated products, galactosylglycerolipid (GG, prepared in our lab as previously described [34]) and galactosylceramide (GC, non-hydroxylated form; Sigma) were co-chromatographed as standards. The HPTLC plate was then stained with 0.2% orcinol in 50% H_2SO_4 at 100°C [35].

SGC-ASA binding

One microgram of rec ASA (C69A or CKK) or transferrin (purified from OptiMEM-I medium following the same protocol as used for rec ASA purification) was incubated

with 818 pmol (~45 nCi) of radiolabeled [^{14}C]SGC (*N*-[1′-^{14}C]stearoyl-(3″-SGC)) (American Radiolabeled Chemicals, St. Louis, MO, USA, a generous gift from P. Manowitz) in 50 µl of 0.1% taurodeoxycholate in 0.25 M sodium acetate (NaAc-T), pH 5 for 2 h at 37°C. A blank was also set up by omitting protein in the incubation mixture with [^{14}C]SGC. The reaction mixture was then transferred to a Microcon YM-100 centrifugal filter device and the volume was completed to 400 µl using NaAc-T. The Microcon YM-100 was subjected to centrifugation (14,000×*g*, 10 min, 4°C), followed by five further such washes, each with 400 µl NaAc-T. The final retentate was collected in 150 µl NaAc-T, and subjected to scintillation counting for the amount of [^{14}C]SGC, using a Scintiverse BD Cocktail (Fisher Scientific, Ottawa, ON, Canada) in a Beckman LS1700 scintillation counter (Beckman Coulter Canada, Inc., Mississauga, ON, Canada).

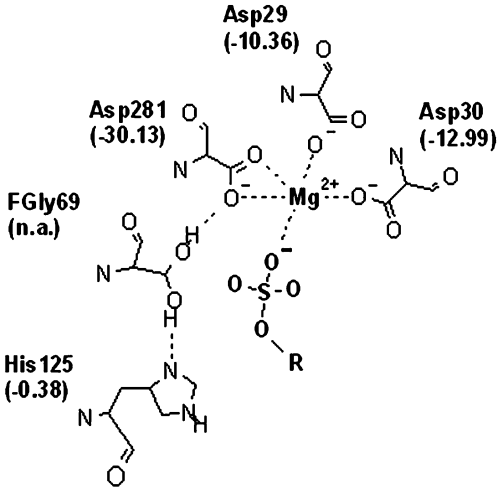

Fig. 1 Scheme displaying the residues around the metal ion in the ASA active site with their calculated pK_a values. The magnesium ion is coordinated by the three active site aspartates, which all have a highly negative pK_a and are thus negatively charged at pH 5. His125, which forms a hydrogen bond with a hydroxyl group of FGly69, has a pK_a just under 0 and would thus be found mainly in its neutral form

Results

Calculated pK_a values of ASA amino acid residues

The amino acid pK_a values calculated by WhatIf/DelPhi ranged from −13 to +27. Negative values represent titrating groups that would normally not get protonated, whereas groups that are practically always protonated have extremely high numbers. Apart from aspartates, glutamates and histidines, all residues had a pK_a value well above 7, as expected, and would not change their protonation state around the optimum pH of ASA.

Table 1 lists the calculated pK_a values for titratable residues around the active site. In the standard calculations ("default run" column), the lowest values were obtained for Asp29 and Asp30 (−10 and −13, respectively). This was expected based on the ASA crystal structure [8], as these two residues require a negative charge in order to assist in the coordination of the magnesium ion (Fig. 1). In contrast, a third aspartate, Asp281, had a calculated pK_a value of +14. This value was conflicting with results from the ASA crystal structures, which strongly suggest that this aspartate coordinates the metal ion with its full valence using both its oxygens [8], as shown in Fig. 1. Magnesium had not been included in the standard WhatIf/DelPhi run, and its absence, in conjunction with the negative electrostatic potential of the other two aspartates, would have severely increased the pK_a of Asp281 and could thus explain this unusually high value. After repeating the calculations with Mg^{2+} treated as a ligand, a major pK_a shift affected several residues, most importantly Asp281, which now had an extremely low value of −30. This confirmed its major role in binding the magnesium ion. Significant shifts were also observed for other amino acids around the active site (Asp29, Asp30, Arg73, Lys123, His125, and His229; Table 1, right column). However, except for His229, these shifts would not affect the electrical charges of these amino acids at pH 5, since the pK_a values of both aspartates and His125 remained less than 5, and those of Arg73 and Lys123 were significantly higher than 5. By including Mg^{2+} in the calculation, the pK_a value for His229 dropped from 7.34 to 3.95. The new value would suggest a neutral charge at pH 5, which would be in conflict with results previously obtained from the crystallography studies. The crystal structure of C69A-ASA complexed with NCS [9] revealed an inter-atomic distance of 2.4 Å between His229-Nε and the ester oxygen OS4, indicating strong interaction between His229 and the sulfate group; this could occur only with His229 fully protonated [9]. In addition, the generation of the alcoholate product from the enzyme-substrate intermediate requires proton donation from His229 (see Fig. 2). The distance between the Mg^{2+} ion and His229 is 4.3 Å, a value that argues against strong influence of Mg^{2+} on the electrical charge status of the amino acid. Based on these considerations, we used the fully protonated form of His229 (pK_a value of 7.34) from the standard WhatIf/DelPhi run for the ensuing docking of SGG/SGC.

While most histidines, including His229, would be positively charged at pH 5, the calculated pK_a values for His125, His151, His226, and His397 were below 5 for both calculations, with and without Mg^{2+}. Thus, these residues were treated as neutral in the subsequent dockings. A further histidine, His405, had a pK_a value of 5.16 and could be treated either way at pH 5. All aspartates and most

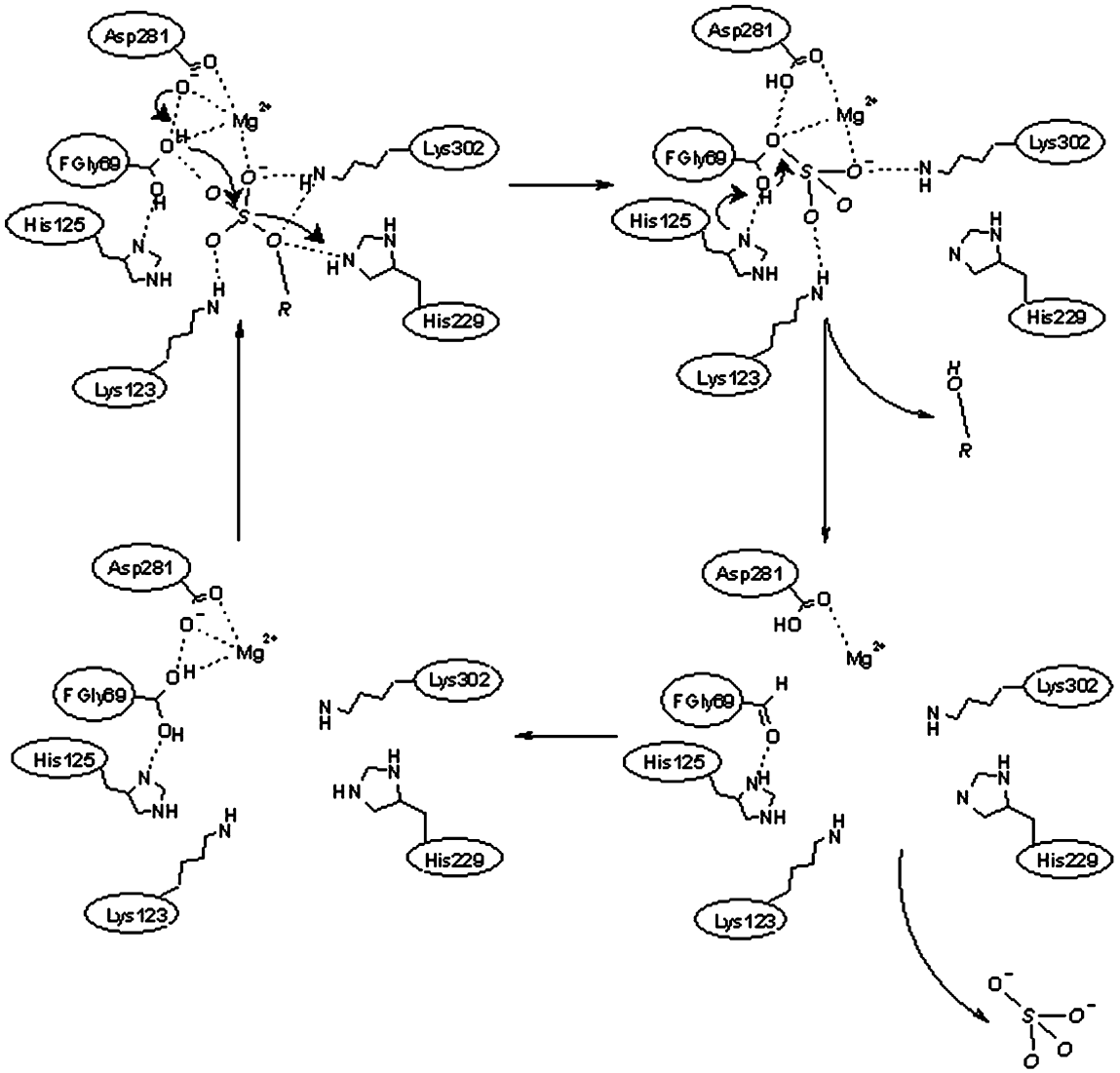

Fig. 2 Desulfation mechanism of ASA. The illustration, graphically adapted from Von Bulow *et al*. [8], shows relevant amino acids in the active site, including the two lysines (Lys302 and Lys123) that are involved in this mechanism. Once the sulfate group of the substrate is positioned close enough to one of the hydroxyl groups of FGly69 through interactions with the binding residues, it sets off an S_N2 substitution reaction that covalently binds the sulfate group to FGly69

glutamates were left deprotonated, apart from Glu109 and Glu312.

Using the program APBS, the electrostatic potentials for ASA were calculated at pH 5 and pH 7, based on the calculated pK_a values. The resulting potentials were projected onto the surface of ASA using the APBS plug-in within Pymol, as shown in Fig. 3. The colors blue and red represent positive and negative potentials, respectively. From pH 7 to pH 5, an increase of positive charges was observed all over the surface, but the strongest change was around the active site. At pH 7, there was only a slight positive charge at the active site, which was mainly attributed to the magnesium ion. This was confirmed by calculating the electrostatic potential without a metal cation; the calculation yielded a negative charge potential at the active site (data not shown).

Dockings

Dockings at pH 5 were performed with the protonation states of all titratable residues set according to their calculated pK_a values, as described above and shown in Table 1 (bold font). Since the calculated pK_a of His405 was close to 5, it could be treated as either neutral or positive. Initial dockings of NCS with AD were performed with either the neutral or positively charged form of His405; only the neutral form resulted in top docked conformations with the sulfate group oriented towards Mg^{2+}, as expected

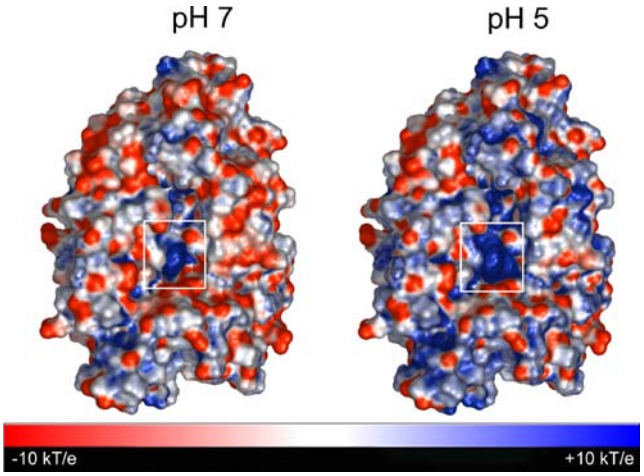

Fig. 3 Electrostatic surface of ASA at pH 7 (*left*) and pH 5 (*right*). Based on the pKa values calculated by WhatIf and Delphi, the electrostatic potential of ASA was calculated at pH 7 and 5 using APBS. The surfaces were displayed using Pymol, with negative and positive potential values in *red* and *blue*, respectively (in energy units of kT/e). ASA is more positively charged at pH 5, as expected. Especially the active site (surrounded by a *white box*) is significantly more positive at pH 5. At pH 7, the positive charge of the active site is entirely due to magnesium; calculating the surface at pH 7 excluding the metal resulted in a negative charge (data not shown). This suggests that electrostatic interactions play a crucial role in substrate binding

based on the crystal structure [9]. Therefore, His405 was considered as neutral throughout subsequent dockings.

For the NCS dockings with AD, the best three clusters of conformations displayed similar docking energies but showed different orientations of NCS. While the third cluster highly resembled the co-crystallized NCS with a root mean square deviation (rmsd) of around 1 Å, the first and second clusters had poses that would not allow desulfation to occur and would interfere with FGly69. On the other hand, the AD dockings of the other three ligands (galactosyl sulfate, SGG and SGC) revealed that their sulfate group bound to the active site in the same manner as the co-crystallized NCS. Out of the two crystallized NCS conformations (A and B, as described in the PDB structure of ASA), all dockings converged towards conformation B. The QPLD dockings showed the sulfate group of the docked NCS in the same position as observed in the B conformation of crystal structure (Fig. 4). Except for the precise placement of sulfate group, the contacts made by AD poses were conserved in the poses suggested from the QPLD methodology. However, with the QPLD calculations, the charges on the sulfate substrate atoms were less negative, as compared with the AD calculations (Table 2). Therefore, it is expected that the substrate would be farther from Ala69, when the QPLD docking was done, a situation more subtly resembling the crystal structure. In fact, comparison of atomic distances obtained by QPLD and AD dockings showed that OS2 and OS3 of NCS were slightly

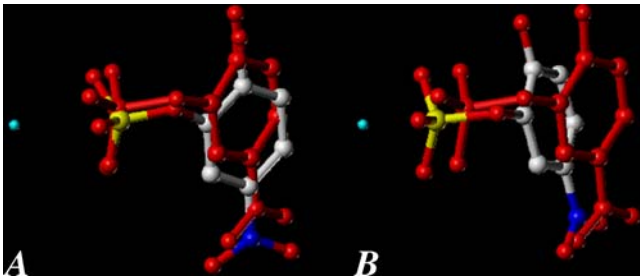

Fig. 4 Orientation of the sulfate group of NCS from AutoDock and QPLD docking, compared with the orientation seen in the crystal structure. Mg^{2+} ion is shown as a blue sphere. **a** QPLD top pose superimposed over the crystal structure pose (all in *red*). An RMSD of 0.2 Å was observed. **b** AutoDock pose superimposed over the crystal structure pose (all in *red*). The sulfate group was seen closer to the Mg^{+2} ion by AD docking, as compared with its position in the crystal structure

farther from C-β of Ala69 in *1E2S* in QPLD docking, a result more similar to the situation in the crystal structure (Table 3). The same trend was observed for the atomic distance between *OS1* and Mg^{2+}. However, the same atomic distance values of the interaction between OS1 and C-β of Ala69 were obtained from QPLD and AD dockings. Therefore, we believe that the two dockings, by and large, gave the same trend of results.

Table 3 lists all other atomic interactions between the active site residues and the ligands, as revealed by both AD and QPLD dockings. For the NCS dockings with AD and QPLD, hydrogen bonds were seen between the sulfate oxygens and atoms in the side chains of residues Lys123, Lys302, and His229, as well as the backbone nitrogen of Ala69. Thus, the dockings represented most of the hydrogen bonds to the sulfate group that had been observed in the crystallized complex [9] (Table 3, and Fig. 5), even though some atomic distances were slightly different and some were slightly higher than the cutoff 3.5 Å value. An exception was the hydrogen bond between Ser150 and the sulfate oxygen OS2, which was found in the crystal structure [38] but not in the docked complex from either of the docking methods. Even though Ser150 has been

Table 2 Charges of atoms in the NCS sulfate group used in AD and QPLD docking

Atom	QPLD charges	AD charges
S	1.71	0.22
OS1	−0.89	−0.60
OS2	−0.79	−0.60
OS3	−0.80	−0.60
OS4	−0.48	−0.24
Total	**−1.25**	**−1.82**

Table 3 Atomic distances between ASA residues and substrate atoms according to AD and QPLD dockings

ASA residues		Ligands								
		NCS[a]	NCS		Gal3S		SGG		SGC	
Amino acid	Atom	Atom	Distance (in Å)							
			AD	QPLD	AD	QPLD	AD	QPLD	AD	QPLD
	Mg^{2+}	*OS1*	*2.7*	1.7	3.1	1.7	3.2	1.7	3.0	1.7	3.2
A69	C-β	*OS1*	*3.3*	3.5	3.5	3.7	3.4	3.6	3.4	3.7	3.4
A69	C-β	*OS2*	*3.5*	2.6	3.5	2.6	3.1	2.6	3.0	2.5	3.1
A69	C-β	*OS3*	*3.4*	2.7	3.2	2.8	3.2	2.7	3.1	2.7	3.3
K123	N-ζ	*OS1*	***3.2***[b]	3.9	3.8	3.6	3.6	3.6	3.6	3.6	3.8
K302	N-ζ	*OS1*	***3.5***	**3.4**	**3.1**	**3.5**	**3.2**	3.6	**3.3**	3.6	**3.2**
H229	N-ε	*OS1*	***2.9***	3.6	**3.1**	**3.3**	**3.0**	**3.3**	**3.0**	**3.3**	**3.1**
K123	N-ζ	*OS2*	***3.3***	**2.6**	**3.1**	**2.6**	**2.7**	**2.5**	**2.7**	**2.6**	**2.7**
S150	O-γ	*OS2*	***3.2***	4.0	4.3	4.0	4.0	3.9	4.1	4.1	4.2
A69	N	*OS3*	***3.3***	**2.5**	**3.4**	**2.6**	**3.4**	**2.6**	**3.3**	**2.5**	**3.4**
K302	N-ζ	*OS4*	***2.8***	**2.5**	**2.6**	**2.6**	**2.9**	**2.5**	**2.8**	**2.6**	**2.9**
H229	N-ε	*OS4*	***2.4***	**2.6**	**2.7**	**2.7**	**2.8**	**2.7**	**2.8**	**2.7**	**2.7**
S150	O-γ	*O7*[c]	***3.2***	**2.3**	**3.4**	–		–		–	
H229	N-ε	*O7*[c]	***3.2***	**2.8**	**3.3**	–		–		–	
K302	N-ζ	*O4′*	–	–		**3.1**	3.9	**3.1**	**3.5**	**2.9**	4.5
H229	N-ε	*O4′*	–	–		**2.9**	**3.5**	**2.8**	**2.8**	**2.9**	**3.4**
D152	O-δ1	*O6′*	–	–		**3.0**	**3.4**	**3.3**	4.7	**3.2**	4.8
D152	O-δ2	*O6′*	–	–		**2.7**	**3.4**	**2.9**	4.4	**2.6**	5.0
V91	N	*O1*[d]	–	–		–		**3**	**3.3**	–	
H405	N-ε	*N2*[d]	–	–		–		–		**2.9**	**3.4**
Y379	O-ω	*O1*[e]	–	–		–		**3.5**	**3.1**	3.7	**3.1**

[a] Data in this column were from the co-crystal structures of ASA and NCS (1E2S)
[b] Data in bold font represent hydrogen bonds with an applied cutoff of 3.5 Å. Primed atom numbers denote atoms that are part of the galactose
[c] Denotes the oxygen at the ring position 2 in NCS
[d] denotes atoms in the lipid backbone (glycerol for SGG and sphingosine for SGC)
[e] denotes the carbonyl oxygen of the acyl group

suggested to play a role in NCS binding [9], a point-mutation of this serine to alanine had no significant effect on the K_m [9]. Hence, Ser150 cannot be a major residue for substrate binding, and it might not provide a significant hydrogen bond to the ligand. Interestingly, when comparing the ASA crystal structures *1E2S* and *1N2K*, Ser150 is the only active site residue with different orientations in the two crystal structures, further implying that its positioning might not be crucial for substrate binding.

Using the same setup and parameters as for the NCS dockings, galactosyl sulfate, SGG, and SGC were also docked into ASA through both AD and QPLD. When comparing the four docked ligands with each other, almost all atom distances for polar interactions remained the same (Table 3), and six hydrogen bonds to the sulfate groups (as mentioned above) were found throughout. Lys123 and Lys302 bound the sulfate group with a total of three hydrogen bonds of which two were strong. This indicated the significant roles of these lysines in the binding of the sulfate group of the substrate to the ASA active site pocket. From the ASA crystal structure (*1E2S*) and from the docking results shown herein, His229 also provided two hydrogen bonds for the capture of the sulfate group. However, His229 may act beyond the sulfate anchoring; it may participate in the desulfation mechanism by acting as a proton donor, thus generating the R-OH product of the reaction [9] (Fig. 2).

Residues Asp152, Asp173, and Arg288 have previously been suggested to be involved in the binding of the galactosyl sulfate head group of the sulfoglycolipids [9]. However, the dockings of galactosyl sulfate indicated that no interaction with Asp173 or Arg288 could occur. On the other hand, Asp152 did indeed form a hydrogen bond to O6′ of the galactosyl ring of galactosyl sulfate in the poses suggested from AD and QPLD dockings. The binding of galactosyl sulfate was stabilized by interactions between O4′ and Lys302/His229. Similar polar interactions were also found for the docked sulfoglycolipids SGG and SGC, although their O6′ seemed to have weaker interaction with Asp152 as revealed by QPLD. However, additional

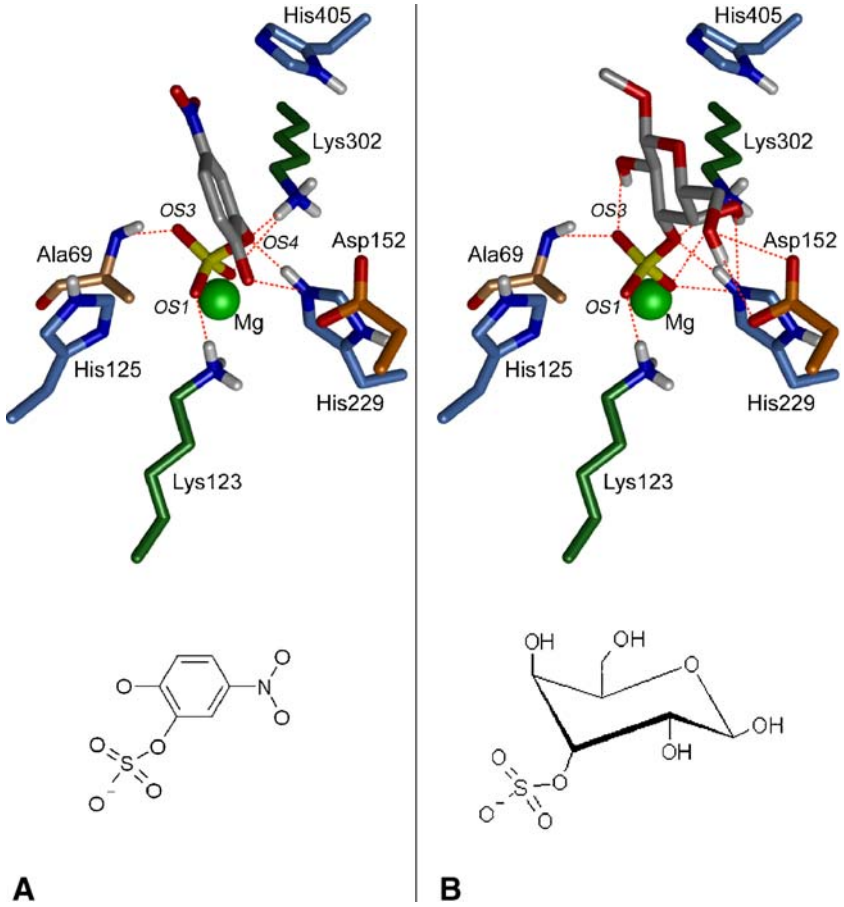

Fig. 5 NCS (**a**) and galactosyl sulfate (**b**) docked with AD into the ASA active site surrounded by active site residues. Mg^{2+} is shown as a *green sphere*, hydrogen bonds as *lines*. The sulfate groups of NCS and galactosyl sulfate were positioned in exactly the same position and orientation. Chimera was used to display the dockings in this figure and also in Figs. 5 and 6

contacts were present between SGG/SGC and active site residues (Table 3 and Fig. 6). There was a strong hydrogen bond seen from both docking methods between O1 on the glycerol backbone of SGG and the backbone nitrogen of Val91. The binding of SGC was strengthened by a hydrogen bond between N2 of the sphingosine backbone and His405. Both bound sulfoglycolipids were also reinforced through interaction between O1 of the glycerol/sphingosine backbone and the hydroxyl group of Tyr379 at a distance of 3.5 and 3.7 Å for SGG and SGC, respectively, from AD docking and 3.1 and 3.1, respectively, from QPLD results. However, when considering all top orientations up to the third cluster in AD, the observed distance was as low as 3.1 and 2.7 Å for SGG and SGC, respectively, indicating a stable hydrogen bond. One intramolecular hydrogen bond between the sulfate oxygen OS3 and the galactosyl oxygen O2′ helped to further stabilize the bound conformation of SGG (Fig. 6).

Both sulfoglycolipids found further binding support through hydrophobic interactions between atoms in their hydrocarbon chains and residues Leu68, Val91, and Val93, as listed in Table 4. Very similar observations were made from both the docking methods. Additional hydrophobic interactions are expected in the natural sulfoglycolipids due to their longer hydrocarbon chains. Figure 7 shows the dockings of SGG and SGC to ASA with the protein surface colored according to lipophilicity. The sulfated galactosyl headgroups were completely accommodated within the active pocket, leaving just enough space for the sulfoglycolipid backbones and the proximal parts of the hydrocarbon chains. This may thus contribute to the substrate specificity of ASA. Tyr379, which was also found to contribute to substrate specificity by providing a hydrogen bond (Fig. 7), protrudes as a bump, forcing the hydrocarbon chains to split in two directions (Fig. 7). A comparable substrate specific residue has been revealed in the active site of arylsulfatase C; its Arg98 provides a hydrogen bond stabilizing the bound substrate, estrone sulfate [39]. The two clefts just outside of the ASA active site suggest the fitting of the proximal parts of the hydrocarbon chains of the sulfoglycolipids. Several hydrophobic residues exist in these clefts might support the binding of extended hydrocarbon chains, but the binding of the complete chains is unlikely.

Site-directed mutagenesis of Cys69, Lys302 and Lys123 to verify their roles in sulfoglycolipid desulfation/binding

Results from computational studies described above suggested that Lys302 and Lys123 are important for binding

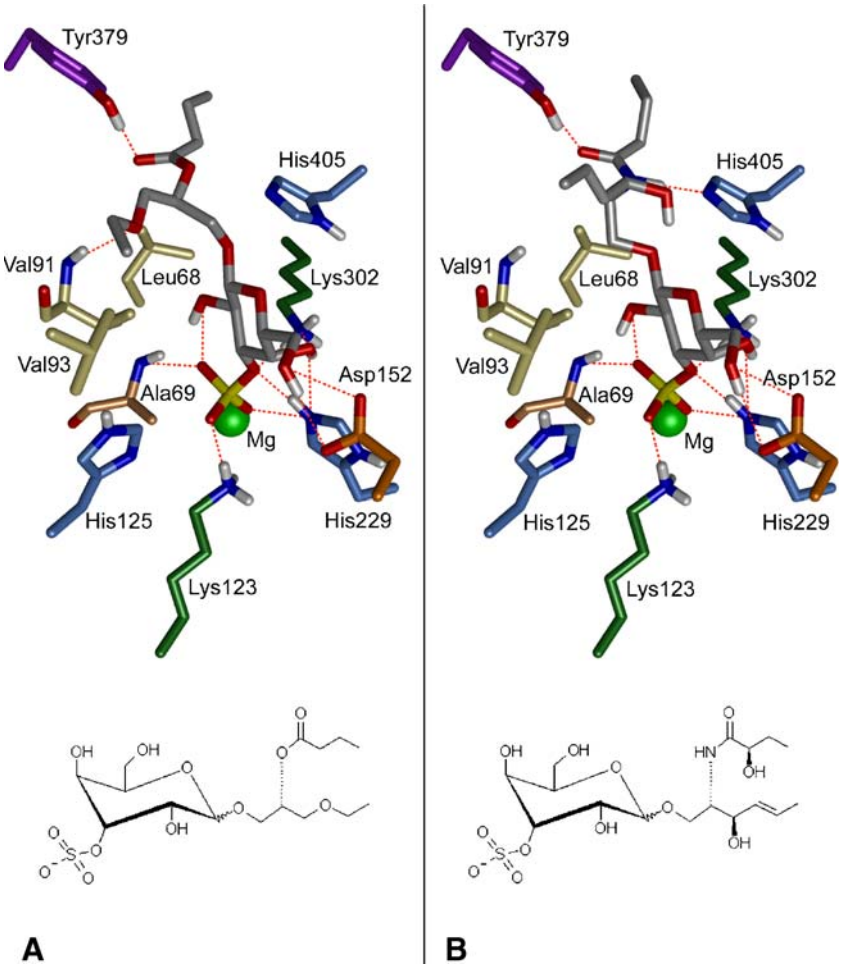

Fig. 6 SGG (**a**) and SGC (**b**) docked with AD into the ASA active site surrounded by active site residues. Mg^{2+} is displayed as a *green sphere*, hydrogen bonds as *lines*. The sulfate groups and galactose rings of SGG and SGC overlapped precisely with the docked galactosyl sulfate

the sulfate group of SGG/SGC. On the other hand, accumulated evidence, as described previously [8, 9] and herein, has pointed to the significance of FGly69 (Cys69) as the active site amino acid in the arylsulfate desulfation mechanism. The involvement of FGly69 in NCS desulfation has been proven by site-directed mutagenesis experiments [40]. To further prove that this was also the case for SGG/SGC, a C69A mutant ASA, along with wild type ASA, was produced as a secreted protein by CHO-K1 cells (see "Experimental procedures"). The significance of the two lysines in terms of the substrate binding was also studied by producing the CKK mutant. Immunoblotting shown in Fig. 8 revealed that all three rec ASAs were equally well expressed by CHO-K1 cells, and both C69A and CKK mutants were secreted into the medium by NH_4Cl-treated cells in the same manner as wild type ASA. The expression levels of wild type ASA, C69A, and CKK were much higher than the endogenous ASA (the pcDNA sample). These rec ASAs had the same apparent molecular mass (~66 kDa) on SDS-PAGE as native pig sperm ASA, as demonstrated by immunoblotting (data not shown). Both native and rec human ASA are in the dimer and octamer form in solution at neutral and acidic pH, respectively [8]. Our rec ASAs expressed here also existed in an oligomer form; they were present only in the retentate after Microcon YM-100 centrifugation at both pH 5 and 7 (data not shown).

As expected, our expressed wild type rec ASA contained NCS desulfation activity. In the medium of NH_4Cl-treated CHO-K1 transfected cells, wild type ASA had a specific activity of ~10 U/mg proteins (Fig. 8). After partial purification (see "Experimental procedures"), its specific activity was increased to ~80 U/mg protein. Partially purified wild type ASA possessed the ability to desulfate the natural substrates, SGC and SGG (Fig. 9). In contrast, C69A and CKK did not have NCS desulfation activity (Fig. 8). Also, SGC/SGG desulfation by C69A (Fig. 9) and CKK (data not shown) could not be achieved. These results demonstrated the necessity of Cys69 (FGly69) in the desulfation process of SGC and SGG, the natural substrates of ASA.

Docking results shown in Fig. 6 suggest that Lys302 and Lys123 are involved in the initial binding of the sulfate group of the substrate to the active site pocket of ASA (see scheme in Fig. 2). This postulate was supported by results

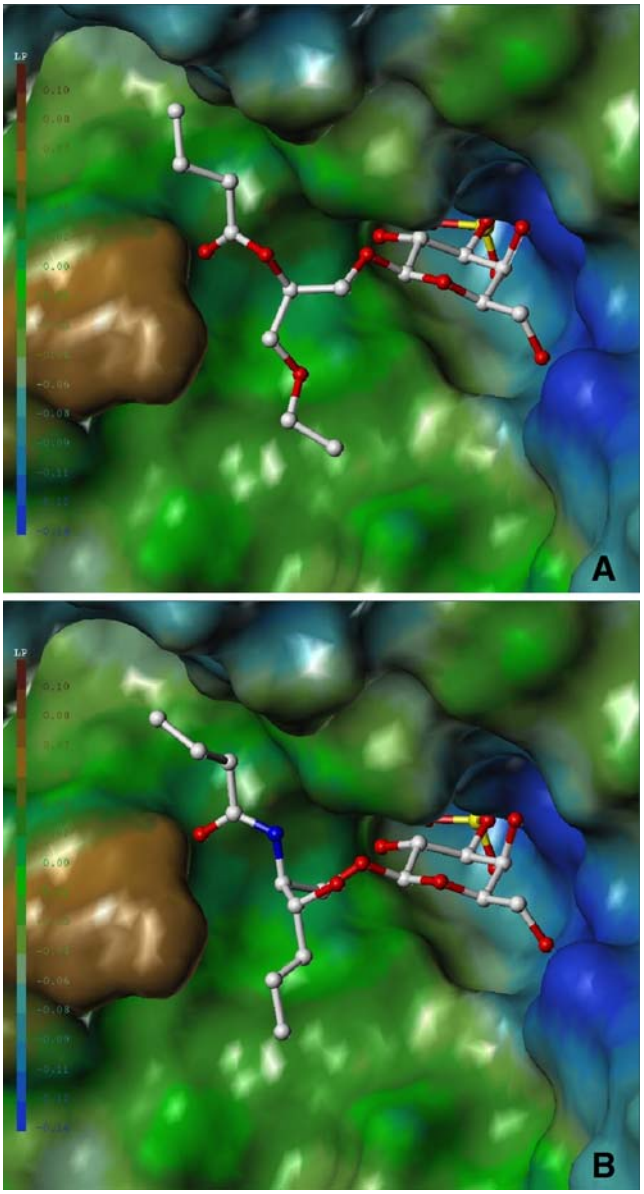

Fig. 7 Ligands, SGG (**a**) and SGC (**b**), docked with AD at the molecular surface of ASA. The surface is colored according to lipophilic potential in Sybyl; *brown* indicates high lipophilicity and low hydrophilicity. The sulfated galactosyl headgroups are positioned inside the active pocket, with the hydrocarbon chains splitting at a bump, which is Tyr379

displayed in Fig. 10. When C69A or CKK was incubated with [^{14}C]SGC at pH 5, C69A showed a significant affinity for the sulfoglycolipid. The complex of C69A and [^{14}C]SGC, remaining in the Microcon YM-100 retentate (see "Experimental procedures"), revealed a 1:4 molar ratio of SGC/C69A. In contrast, the molar ratio of SGC to CKK in the Microcon YM-100 retentate was only 1:25, indicating a low affinity of CKK for the sulfoglycolipid. Transferrin, which was co-purified with rec ASA (see Fig. 10, inset for purity of C69A and CKK used for this binding experiment), had no affinity for SGC (only background radioactive SGC remained with transferring in the Microcon YM-100 retentate). Since Lys302 and Lys123 were present only in C69A and not in CKK, these results indicated that these two lysines were involved in the binding of the sulfoglycolipid substrate.

Discussion

Even though ASA has been crystallized as a complex with NCS, to date a crystal structure containing the natural substrates has not been achieved. Therefore, the details of interactions between the sulfoglycolipids and ASA are not known. Our computational dockings, described herein, gave a detailed prediction of these interactions. The main interactions known for the NCS sulfate group with amino acids and the metal ion in the ASA active site pocket were reproduced by docking methods. Although AD docking yielded the atomic distances, which deviated slightly from those seen in the crystal structure, the atomic distances as revealed by QPLD docking were very similar to those in the ASA crystal. The slight differences in the orientation of the sulfate group suggested that there were polarization effects on the substrate atoms from the protein environment, which were taken into account in the QPLD docking. Nonetheless, the amino acid residues interacting with the substrates were seen in common in both the docking procedures, thus validating their use in defining amino acids of ASA that SGG and SGC interact with. The significance of two basic amino acids, Lys302 and Lys123, in capturing SGG and SGC into the active site pocket, as shown by these dockings, was further confirmed by site directed mutagenesis experiments.

The docking results confirmed Asp281 as the major metal-coordinating residue. Although its pK_a value was positive in the initial WhatIf run, the inclusion of Mg^{2+} resulted in a strong pK_a shift from +14 to −30. Therefore, Asp281 would be negatively charged, as long as the metal ion is present. His125 had a negative pK_a value and would be neutral at pH 5, whereas His229 was considered to be fully protonated (see Results section). In a model derived from the crystal structure of rec human ASA, Lukatela *et al.* [8] suggested that a deprotonated His125 could stabilize the formylglycine residue by hydrogen bonding with one of its hydroxyl groups. The recently resolved crystal structure of native human ASA [10] contains the full formylglycine residue. In this structure, the proposed hydrogen bond exists between His125 and FGly69. A hydrogen bond from the hydrogen on His125N-ε to the backbone oxygen of Leu92 also suggests that the resting state of His125 is the epsilon form. His229, on the other hand, would be fully protonated at pH 5 until the sulfate group of a substrate

Table 4 Interacting atomic distances between hydrophobic residues and substrate atoms according to AD and QPLD dockings

ASA residue		SGG			SGC		
Amino acid	Atom	Atom	Distance (Å)		Atom	Distance (Å)	
			AD	QPLD		AD	QPLD
L68	*C*-δ1	C2	3.7	4.0	C1	3.4	3.8
L68	*C*-δ1	C1″	3.3	3.8	C2	3.6	3.9
L68	*C*-δ1	–	–	–	C1″	3.2	3.4
V91	*C*-γ*1*	–	–	–	C1	3.6	3.9
V91	*C*-γ*2*	C1′ ″	3.6	3.9	C1′	3.6	3.7
V91	*C*-γ*2*	C2′ ″	3.4	3.8	C1	3.4	3.5
V93	*C*-γ*1*	C2′ ″	3.2	3.7	–	–	–

Primed atom numbers denote atoms that are part of the galactose. Double- and triple-primed numbers denote atoms in the acyl and alkyl chains, respectively

enters the site, at which point the simultaneous protonation of Asp281 and deprotonation of His229 would set off the cleavage of the sulfate group (Fig. 2). Our dockings suggested that the initial binding of the sulfate group of the substrate relies mainly on Lys123 and Lys302, whereas His229 mainly acts as a proton donor. Also, the dockings implied that the sulfur atom might be positioned very closely to a hydroxyl group of FGly69. Only on this basis, can a covalent bond be formed between FGly69 and the sulfate group. The subsequent release of the cleaved sulfate would involve the protonation of His125, as presented in Fig. 2, an illustrative adaption of the model by von Bulow *et al*. [9].

Because His405 had a pK_a value close to 5, dockings were tested with His405 fully protonated or neutral. When docked using the fully protonated form, a sulfate oxygen of the ligand was positioned within hydrogen bonding distance to the protonated His405N-ε, but away from FGly69 and Mg^{2+}. This indicated that His405 should be in its neutral form. Our closer assessment of the crystal structure of ASA revealed a hydrogen bond between His405N-δ and the non-protonated Glu285. Using His405 in its neutral delta form, our dockings revealed that the

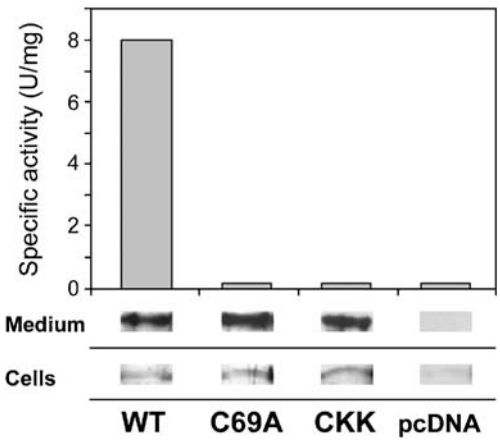

Fig. 8 Expression of rec ASAs and their NCS desulfation activity. RecASA (WT, C69A, CKK) secreted into the medium was confirmed to be ASA by immunoblotting. All samples were from one 100-mm dish of transfected cells (see "Experimental procedures"). The amount of ASA remaining inside the cells was minimal, indicating the prevalent targeting of ASA into the medium. Under the same conditions, ASA was not detectable in the pcDNA sample (from cells transfected by an empty plasmid). While WT had high NCS desulfation activity, the mutants, C69A and CKK, were inactive, showing background activity values as observed for pcDNA. Results shown were representative of triplicate experiments

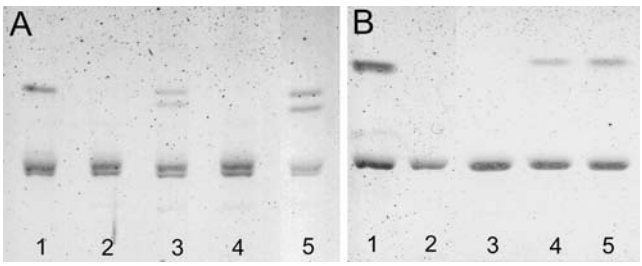

Fig. 9 HPTLC showing SGG/SGC desulfation by native and wild type recASA but not C69A mutant ASA. Reactions for SGC and SGG desulfation by recASAs and native ASA, as well as pcDNA were carried out as described in the Experimental procedures. **a** SGC desulfation. *Lane 1* standard lipids, SGC (mixture of hydroxylated and non-hydroxylated forms) and GC (non-hydroxylated form); *lane 2* pcDNA sample was used in the reaction to reveal the level of endogenous ASA in CHO-K1 cells; *lanes 3* and *4*: wild type recASA and C69A mutant ASA, respectively, were used in the reaction; *lane 5*: native pig sperm ASA was used as a positive control in the reaction [30]. **b** SGG desulfation. *Lane 1* standard lipids, SGG and GG; *lane 2* pcDNA was used in the reaction. *Lanes 3* and *4* C69A mutant and wild type recASA, respectively, were used in the reaction; *lane 5* native sperm ASA was used as a positive control. Results shown were representative of duplicate experiments

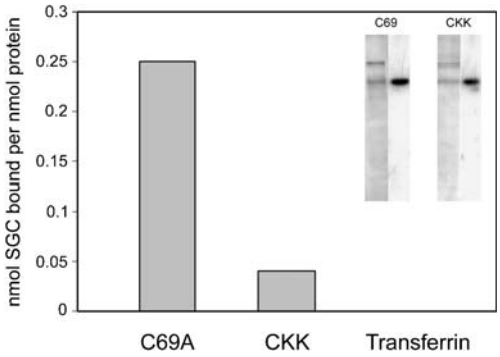

Fig. 10 Binding of SGC to C69A but not to CKK. C69A, CKK or transferrin (1 μg each) was incubated with [^{14}C]SGC in NaAc-T buffer at pH 5 as described in "Experimental procedures". A "blank" was also set up by omitting protein in the radioactive incubation mixture. At the end of the incubation period, the incubation mixture was subjected to Microcon YM-100 centrifugation and washed successively with NaAc-T, before the radioactivity in the retentate was quantified by scintillation counting. Under these conditions, the retentate of the blank sample had less than 0.1% of the [^{14}C]SGC used in the incubation mixture. *Inset* Silver stained SDS-PAGE (*left*) and corresponding immunublotting with anti-ASA (*right*), showing the purity of C69A and CKK used for incubation with [^{14}C]SGC. Note that transferrin was the only protein co-purified with ASA. Results shown are representative from duplicate experiments

sulfate group was positioned close to FGly69 and Mg^{2+}, as expected. The dockings of SGC also showed that His405 in its neutral state can form a hydrogen bond with the NH group of the sphingosine backbone, thus contributing to substrate specificity.

Even though the rmsd between the docked and co-crystallized NCS (in *1E2S*) was around 1 Å in AD docking, some differences in atom-atom distances were seen in these AD results with regard to the sulfate position. This can be explained by the fact that AD does not account for polarization effects. In fact, a more accurate pose was obtained from the QPLD method, where the rmsd of the docked pose from the crystal structure was only 0.2 Å. However, in wild type ASA, the sulfur atom of the sulfate group should be closer to FGly69 in order for the cleavage and subsequent covalent bonding of the sulfate group to FGly69 to occur. Our AD docking suggests that the sulfur atom of the sulfate positions within 1.5 Å from one hydroxyl group of FGly69 (data not shown). Due to this proximity, an S_N2 substitution would be forced onto the sulfate group, thus simultaneously cleaving it from the substrate and covalently binding to FGly69. The sulfate oxygens would flip up, as suggested in Fig. 11. In this figure, the orientations of FGly69 and the bound sulfate group are based on the crystal structure of human placenta ASA, which contains a covalently bound phosphate group at this location [10]. Interactions are shown between the sulfate group and FGly69 and its surrounding residues (for visibility, Lys123 was left out) upon the entrance of NCS into the active site (Fig. 11a). After the substrate is oriented correctly and the sulfur atom positioned within about 1.5 Å to one oxygen of FGly69, a covalent bond is formed to the sulfate, when its bond to nitrocatechol is cleaved (Fig. 11b). By forming a covalent bond to the sulfur, the FGly69 oxygen would lose its hydrogen to Asp281. This aspartate would readily accept the hydrogen, while continually maintaining the coordination of the metal ion with its other oxygen. Once Asp281 is protonated and the sulfate group is covalently bound to FGly69 (Fig. 11c), the second oxygen of FGly69 can give up its proton to His125 in the process of releasing the sulfate group (as described in Fig. 2). Our results thus support the desulfation mechanism presented in the model by von Bulow *et al*. [9]. The involvement of Lys123, Lys302 and His229 in the binding of the sulfate

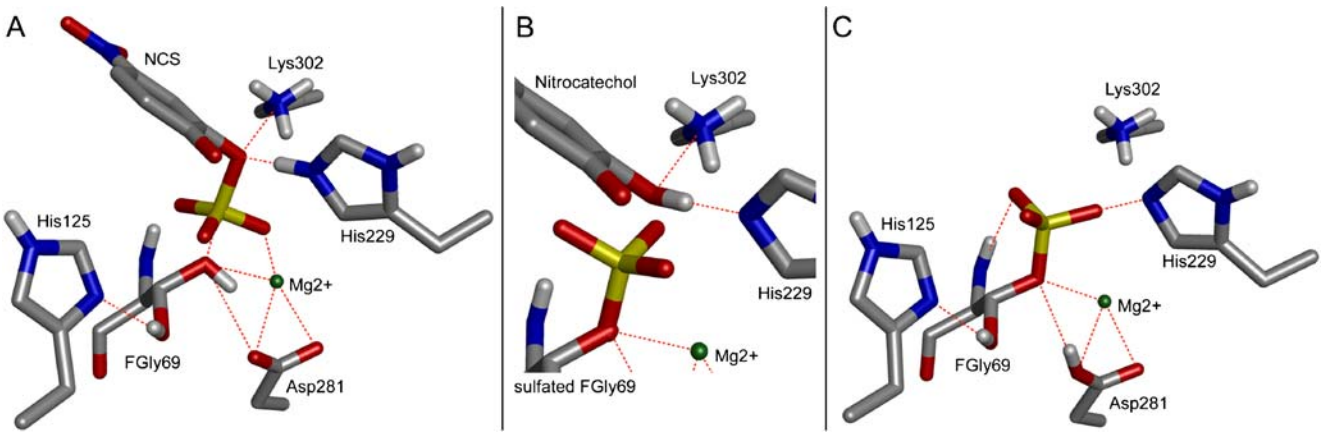

Fig. 11 Desulfation of a sulfated substrate by ASA is an S_N2 substitution reaction. NCS is shown as a representative substrate. **a** Upon entrance into the active site, the sulfate group of the substrate interacts with a distance of 1.5 Å between sulfur and FGly69. **b** A covalent bond is formed between the sulfate group and one of the oxygens of FGly69, simultaneously with the cleavage of the nitrocatechol moiety. This reaction leads to the sulfate oxygens flipping up. **c** After the nitrocatechol moiety gone, the covalently bound sulfate group is in a state that allows it to be released when the second hydroxyl group of FGly69 donates its proton to His125

group was confirmed by our docking studies, while Ser150 seems to be of less importance. Whereas His229 does display several interactions with the sulfate group, its main purpose is most likely to act as a proton donor during hydrolysis.

Lys123 and Lys302, on the other hand, are only responsible for the binding of the sulfate group. The importance of Lys302 and Lys123 in capturing NCS into the active site pocket of ASA has previously been implicated [14]. Site directed mutagenesis of each of these lysines leads to an increase in the K_m value of NCS desulfation. In this report, we also used the site directed mutagenesis approach to validate the combined role of Lys302 and Lys123 in the initial anchoring of SGC to ASA's active site pocket, as suggested by the docking results in Figs. 5 and 6. We also eliminated the subsequent hydrolysis step by substituting Cys69 with an alanine, and as expected this substitution alone led to inactivation of the enzyme activity. C69A could not desulfate NCS or the natural sulfoglycolipid substrates, SGC and SGG, in contrast to WT ASA, which exhibited sulfatase activities (Figs. 8 and 9). The impairment in this sulfate hydrolysis was not due to the inability of this mutant ASA to capture SGC. As shown in Fig. 10, C69A bound to [^{14}C]SGC at a significant level (0.25 nmol per 1 nmol of C69A). In contrast, CKK (Cys69, Lys302 and Lys123 all mutated to Ala) showed minimal binding to [^{14}C]SGC. Our experiments showed that the marked decrease in SGC binding was not due to changes of overall conformation of CKK. It remained in the Microcon YM-100 retentate like C69A and WT ASA, and was also eluted from the Sephacryl S-300 column at the same position as the other two rec ASAs (data not shown). CKK was also targeted extracellularly from cells cultured in the presence of ammonium chloride in the same manner as C69A and WT ASA (Fig. 8). All of these results indicated for the first time the direct involvement of Lys302 and Lys123 in the binding to its natural sulfoglycolipid substrate, SGC.

The final intermolecular energies calculated by AD were around −11 kcal/mol for NCS, −12 kcal/mol for galactosyl sulfate, and between −15 and −16 kcal/mol for each of the two sulfoglycolipids. The higher binding energy for NCS was consistent with its high K_m value. In spite of its high hydrolysis rate by ASA, NCS is an artificial substrate and does not bind very well to the ASA active site—interactions are mainly between active site residues and the sulfate, and there are no hydrophobic interactions to the aryl ring. The presence of the two alternate crystallized conformations (A and B) [9] further imply that—apart from the sulfate group—NCS does not bind well to ASA. This was also reflected by the fact that the dockings resulted in several ligand conformations besides the main one. All these points might explain the anomalous kinetics known for NCS desulfation by ASA [41]. The calculated docking energy was only slightly lower for galactosyl sulfate than for NCS, as the carbohydrate ring provided a few additional binding interactions. However, SGG and SGC bound significantly better to ASA than galactosyl sulfate or NCS. There were more polar interactions for the sulfoglycolipids, and hydrophobic interactions would stabilize the positioning of their backbones. The typical aromatic ring-galactose stacking interactions were not seen in the docking results. The reason appears to be that the aromatic surface of the residues were not sufficiently exposed to the solvent accessible surface (Y230 and Y149). Also, the strong co-ordination of the Mg^{2+} seems to dominate the interactions. The dockings revealed that, additional to any interactions with the sulfate group, hydrogen bonds existed between SGG/SGC and Val91, Tyr379, and His405. Therefore, these three amino acid residues of ASA are likely to contribute to the substrate specificity for sulfoglycolipids. It is not clear which of the two sulfoglycolipids binds more tightly, since the docking energies are similar, the number of hydrogen bonds is comparable, and previous experimental data have shown that their K_m values are in the same range [30].

The lipid chains of SGC and SGG were kept short in the modeling, because a large number of connected carbon atoms would give rise to too many degrees of freedom that are unlikely to be evaluated accurately during the docking runs. However, this should not significantly affect the results, since the main interactions would occur with the galactosyl sulfate head group and the glycerol/sphingosine backbone. Additionally, the lipid chains might only partly bind to the surface of ASA and are otherwise extending outwards. This postulate is based on the fact that *in vivo*, SGC/SGG has to be extracted from membranes by the transporter protein saposin B and then transferred to the active site of ASA [42]. The crystal structure of saposin B [3] indicates that it exists as a dimer with an unusually large hydrophobic cavity (900 Å^3) that would interact with the hydrocarbon chains of SGC/SGG. Thus, the lipid chains might remain within the saposin B cavity while the sulfate group is cleaved off.

The docking results together with an interpretation of the pK_a values enabled us to define more detailed roles of individual residues. Essentially, the residues that had previously been suggested to bind NCS were confirmed by the docking methods, and the same residues were also involved in binding SGC and SGG. Most importantly, the two lysine residues in the active site, Lys123 and Lys302, provided the main interaction for substrate binding, more so for the sulfated sugar head group of SGC and SGG, than for NCS. The significance of these two lysine residues for the overall binding capability of ASA to the sulfoglycolipids was in fact verified by results from our site directed mutagenesis experiments.

Acknowledgment This work was supported by an International Creative Research Initiative Grant from the University of Ottawa and the Canadian Institutes of Health Research (CIHR #62945), both given to NT, as well as a grant from the Swedish Medical Research Council to PGN (grant no. 006).

References

1. Kolodny, E.H., Fluharty, A.L.: In: Scriver, C.R. (ed.) The Metabolic and Molecular Bases of Inherited Disease, pp. 2693–2741. McGraw-Hill, New York (1995)
2. Gieselmann, V.: Metachromatic leukodystrophy: genetics, pathogenesis and therapeutic options. Acta. Paediatr. Suppl. **97**, 15–21 (2008). doi:10.1111/j.1651-2227.2008.00648.x
3. Ahn, V.E., Faull, K.F., Whitelegge, J.P., Fluharty, A.L., Prive, G. G.: Crystal structure of saposin B reveals a dimeric shell for lipid binding. Proc. Natl. Acad. Sci. U. S. A. **100**, 38–43 (2003). doi:10.1073/pnas.0136947100
4. White, D., Weerachatyanukul, W., Gadella, B., Kamolvarin, N., Attar, M., Tanphaichitr, N.: Role of sperm sulfogalactosylglycerolipid in mouse sperm-zona pellucida binding. Biol. Reprod. **63**, 147–155 (2000). doi:10.1095/biolreprod63.1.147
5. Weerachatyanukul, W., Rattanachaiyanont, M., Carmona, E., Furimsky, A., Mai, A., Shoushtarian, A., Sirichotiyakul, S., Ballakier, H., Leader, A., Tanphaichitr, N.: Sulfogalactosylglycerolipid is involved in human gamete interaction. Mol. Reprod. Dev. **60**, 569–578 (2001). doi:10.1002/mrd.1122
6. Tanphaichitr, N., Smith, J., Kates, M.: Levels of sulfogalactosylglycerolipid in capacitated motile and immotile mouse sperm. Biochem. Cell Biol. **68**, 528–535 (1990)
7. Nichol, L.W., Roy, A.B.: The sulfatase of ox liver. IX. The polymerization of sulfatase A. Biochemistry **4**, 386–396 (1965). doi:10.1021/bi00879a002
8. Lukatela, G., Krauss, N., Theis, K., Selmer, T., Gieselmann, V., von Figura, K., Saenger, W.: Crystal structure of human arylsulfatase A: the aldehyde function and the metal ion at the active site suggest a novel mechanism for sulfate ester hydrolysis. Biochemistry **37**, 3654–3664 (1998). doi:10.1021/bi9714924
9. Von Bulow, R., Schmidt, B., Dierks, T., von Figura, K., Uson, I.: Crystal structure of an enzyme–substrate complex provides insight into the interaction between human arylsulfatase A and its substrates during catalysis. J. Mol. Biol. **305**, 269–277 (2001). doi:10.1006/jmbi.2000.4297
10. Chruszcz, M., Laidler, P., Monkiewicz, M., Ortlund, E., Lebioda, L., Lewinski, K.: Crystal structure of a covalent intermediate of endogenous human arylsulfatase A. J. Inorg. Biochem. **96**, 386–392 (2003). doi:10.1016/S0162-0134(03)00176-4
11. Schmidt, B., Selmer, T., Ingendoh, A., von Figura, K.: A novel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency. Cell **82**, 271–278 (1995). doi:10.1016/0092-8674(95)90314-3
12. Dierks, T., Schmidt, B., von Figura, K.: Conversion of cysteine to formylglycine: a protein modification in the endoplasmic reticulum. Proc. Natl. Acad. Sci. USA **94**, 11963–11968 (1997). doi:10.1073/pnas.94.22.11963
13. Dierks, T., Miech, C., Hummerjohann, J., Schmidt, B., Kertesz, M.A., von Figura, K.: Postranslational formation of formylglycine in prokaryotic sulfatases by modification of either cysteine or serine. J. Biol. Chem. **273**, 25560–25564 (1998). doi:10.1074/jbc.273.40.25560
14. Waldow, A., Schmidt, B., Dierks, T., Von Bulow, R., von Figura, K.: Amino acid residues forming the active site of arylsulfatase A. J. Biol. Chem. **274**, 12284–12288 (1999). doi:10.1074/jbc.274.18.12284
15. Lingwood, C., Mylvaganam, M., Minhas, F., Binnington, B., Branch, D.R., Pomes, R.: The sulfogalactose moiety of sulfoglycosphingolipids serves as a mimic of tyrosine phosphate in many recognition processes. Prediction and demonstration of Src homology 2 domain/sulfogalactose binding. J. Biol. Chem. **280**, 12542–12547 (2005). doi:10.1074/jbc.M413724200
16. Vriend, G.: WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. **8**, 52–6, 29 (1990)
17. Honig, B., Nicholls, A.: Classical electrostatics in biology and chemistry. Science **268**, 1144–1149 (1995). doi:10.1126/science.7761829
18. Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., Baker, N.A.: PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. Nucleic Acids Res. **32**, W665–W667 (2004). doi:10.1093/nar/gkh381
19. Baker, N.A., Sept, D., Joseph, S., Holst, M.J., McCammon, J.A.: Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. U. S. A. **98**, 10037–10041 (2001). doi:10.1073/pnas.181342398
20. Holst, M., Saied, F.: Multigrid solution of the Poisson–Boltzmann equation. J. Comput. Chem. **14**, 105–113 (1993). doi:10.1002/jcc.540140114
21. Holst, M., Saied, F.: Numerical solution of the nonlinear Poisson–Boltzmann equation: Developing more robust and efficient methods. J. Comput. Chem **16**, 337–364 (1995). doi:10.1002/jcc.540160308
22. DeLano, W.L.: The PyMOL Molecular Graphics System. [v0.98]. DeLano Scientific, San Carlos, CA, USA (2002)
23. Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W. E., Belew, R.K., Olson, A.J.: Automated docking using a Lamarckian genetic algorithm and empirical binding free energy function. J. Comput. Chem. **19**, 1639–1662 (1998). doi:10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B
24. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Couch, G.S., Meng, E.C., Ferrin, T.E.: UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. **25**, 1605–1612 (2004). doi:10.1002/jcc.20084
25. Cho, A.E., Guallar, V., Berne, B.J., Friesner, R.: Importance of accurate charges in molecular docking: quantum mechanical/molecular mechanical (QM/MM) approach. J. Comput. Chem. **26**, 915–931 (2005). doi:10.1002/jcc.20222
26. Shelley, J.C., Cholleti, A., Frye, L.L., Greenwood, J.R., Timlin, M.R., Uchimaya, M.: Epik: a software program for p*K*(a) prediction and protonation state generation for drug-like molecules. J. Comput. Aided Mol. Des. **21**, 681–691 (2007). doi:10.1007/s10822-007-9133-z
27. Jorgensen, W.L., Maxwell, D.S., Tirado-Rives, J.: Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J. Am. Chem. Soc. **118**, 11225–11236 (1996). doi:10.1021/ja9621760
28. Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, M.P., Repasky, M.P., Knoll, E.H., Shelley, M., Perry, J.K., Shaw, D.E., Francis, P., Shenkin, P.S.: Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J. Med. Chem. **47**, 1739–1749 (2004). doi:10.1021/jm0306430
29. Halgren, T.A., Murphy, R.B., Friesner, R.A., Beard, H.S., Frye, L.L., Pollard, W.T., Banks, J.L.: Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J. Med. Chem. **47**, 1750–1759 (2004). doi:10.1021/jm030644s
30. Carmona, E., Weerachatyanukul, W., Soboloff, T., Fluharty, A.L., White, D., Promdee, L., Ekker, M., Berger, T., Buhr, M., Tanphaichitr, N.: Arylsulfatase A is present on the pig sperm surface and is involved in sperm-zona pellucida binding. Dev. Biol. **247**, 182–196 (2002). doi:10.1006/dbio.2002.0690

31. Stein, C., Gieselmann, V., Kreysing, J., Schmidt, B., Pohlmann, R., Waheed, A., Meyer, H.E., O'Brien, J.S., von Figura, K.: Cloning and expression of human arylsulfatase A. J. Biol. Chem. **264**, 1252–1259 (1989)
32. Chang, P.L., Ameen, M., Yu, C.Z., Kelly, B.M.: Effect of ammonium chloride on subcellular distribution of lysosomal enzymes in human fibroblasts. Exp. Cell Res. **176**, 258–267 (1988). doi:10.1016/0014-4827(88)90329-1
33. Laemmli, U.K.: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**, 680–685 (1970). doi:10.1038/227680a0
34. Tupper, S., Wong, P.T.T., Kates, M., Tanphaichitr, N.: Interaction of divalent cations with germ cell specific sulfogalactosylglycerolipid and the effects on lipid chain dynamics. Biochemistry **33**, 13250–13258 (1994). doi:10.1021/bi00249a011
35. Kates, M.: In: Burdon, R.H. (ed.) Laboratory Techniques in Biochemistry and Molecular Biology, pp. 100–278. Elsevier, New York (1986)
36. Bligh, E.G., Dyer, W.J.: A rapid method of total lipid extraction and purification. Can. J. Biochem. Physiol. **31**, 911–917 (1959)
37. Levine, M., Bain, J., Narashimhan, R., Palmer, B., Yates, A.J., Murray, R.K.: A comparative study of the glycolipids of human, bird and fish testes and of human sperm. Biochim. Biophys. Acta. **441**, 134–145 (1976)
38. Boltes, I., Czapinska, H., Kahnert, A., Von Bulow, R., Dierks, T., Schmidt, B., von Figura, K., Kertesz, M.A., Uson, I.: 1.3 A structure of arylsulfatase from *Pseudomonas aeruginosa* establishes the catalytic mechanism of sulfate ester cleavage in the sulfatase family. Structure **9**, 483–491 (2001). doi:10.1016/S0969-2126(01)00609-8
39. Ghosh, D.: Three-dimensional structures of sulfatases. Methods Enzymol. **400**, 273–293 (2005). doi:10.1016/S0076-6879(05)00016-9
40. Recksiek, M., Selmer, T., Dierks, T., Schmidt, B., von Figura, K.: Sulfatases, trapping of the sulfated enzyme intermediate by substituting the active site formylglycine. J. Biol. Chem. **273**, 6096–6103 (1998). doi:10.1074/jbc.273.11.6096
41. Roy, A.B., Mantle, T.J.: The anomalous kinetics of sulphatase A. Biochem. J. **261**, 689–697 (1989)
42. Louis, A.I., Fluharty, A.L.: Activator-dependant hydrolysis of myelin cerebroside sulfate by arylsulfatase A. Dev. Neurosci. **13**, 41–46 (1991). doi:10.1159/000112139