Topology Scanning and Putative Three-Dimensional Structure of the Extracellular Binding Domains of the Apical Sodium-Dependent Bile Acid Transporter (SLC10A2)†

Eric Y. Zhang,[||,⊥] Mitch A. Phelps,[‡] Antara Banerjee,[||] Chandra M. Khantwal,[||] Cheng Chang,[‡] Freek Helsper,[||] and Peter W. Swaan*[,‡,§,||]

Ohio State Biophysics Program, 484 West 12th Avenue, Columbus, Ohio 43210, Core Laboratory for Bioinformatics and Computational Biology, Dorothy M. Davis Heart and Lung Research Institute, 473 West 12th Avenue, Columbus, Ohio 43210, and Department of Pharmaceutical Sciences, University of Maryland, 20 Penn Street, Baltimore, Maryland 21201

Received April 13, 2004; Revised Manuscript Received June 9, 2004

ABSTRACT: The apical sodium-dependent bile acid transporter (ASBT, SLC10A2) facilitates the enterohepatic circulation of bile salts and plays a key role in cholesterol metabolism. The membrane topology of ASBT was initially scanned using a consensus topography analysis that predominantly predicts a seven transmembrane (TM) domain configuration adhering to the "positive inside" rule. Membrane topology was further evaluated and confirmed by N-glycosylation-scanning mutagenesis, as reporter sites inserted in the putative extracellular loops 1 and 3 were glycosylated. On the basis of a 7TM topology, we built a three-dimensional model of ASBT using an approach of homology-modeling and remote-threading techniques for the extramembranous domains using bacteriorhodopsin as a scaffold for membrane attachment points; the model was refined using energy minimizations and molecular dynamics simulations. Ramachandran scores and other geometric indicators show that the model is comparable in quality to the crystal structures of similar proteins. Simulated annealing and docking of cholic acid, a natural substrate, onto the protein surface revealed four distinct binding sites. Subsequent site-directed mutagenesis of the predicted binding domain further validated the model. This model agrees further with available data for a pathological mutation (P290S) because the mutant model after in silico mutagenesis loses the ability to bind bile acids.

The enterohepatic circulation effectively conserves the human bile acid pool. Vectorial transport of conjugated bile acids in hepatic and intestinal membranes is achieved efficiently by two related sodium-dependent bile acid transporters (SBATs).[1] The intestinal transporter, apical sodium-dependent bile acid transporter (ASBT, SLC10A2) (*1*−*3*), is expressed on the apical membrane of ileocytes and cholangiocytes (*4*). ASBT is a 41-kDa glycoprotein, consisting of 348 amino acids, functional as a monomer, although there is evidence that dimers of approximately 90 kDa also occur (*5*). The liver Na^+-dependent taurocholate-cotransporting polypeptide, Ntcp (SLC10A1) (*6*, *7*), is localized selectively on the basolateral (sinusoidal) domain of hepatocytes and is the protein principally responsible for bile acid transport from portal blood into the hepatocyte (*8*). ASBT and Ntcp share 35% identity and 63% amino acid sequence similarity and an almost overlapping hydropathy profile, and similar transmembrane (TM) topologies have been suggested (*5*). At the present time, however, a definitive ASBT membrane topology remains controversial and needs further clarification. Topography of polytopic integral-membrane proteins can be studied by various techniques such as in vitro and in vivo translation of constructs containing transmembrane segments (TMSs) (*9*), N-glycosylation-scanning mutagenesis (*10*), and epitope localization (*11*).

Hydropathy analysis of SBATs suggests seven to nine putative TM domains (Figure 1), depending on the prediction algorithm, hydrophobicity scale, and calculation stringency, but these estimates require experimental validation. Previous studies have confirmed an exoplasmic glycosylated N terminus and a cytoplasmic C terminus (N_{exo}/C_{cyt}) (*6*), suggesting an odd number of TM segments. In fact, most early studies and a recent report by Sun and colleagues (*12*) adopted a 7TM model. However, Hallén and colleagues (*9*, *13*) suggested the possibility of a 9TM model based on membrane-insertion-scanning analysis. This topology predicts two membrane-spanning regions of only 9−10 amino acids each (TM3 and TM4) and two TM regions (TM2 and TM3)

† This research was supported by a new investigator award from the PhRMA Foundation (to P.W.S.) and a grant from the National Institutes of Health, National Institute for Digestive Diseases and Kidney DK061425 (to P.W.S).

* To whom correspondence should be addressed: University of Maryland, 20 Penn St., HSF2-621, Baltimore, MD 21201. Telephone: 410-706-0103. Fax: 410-706-1590. E-mail: pswaan@rx.umaryland.edu.

‡ Ohio State Biophysics Program.

§ Dorothy M. Davis Heart and Lung Research Institute.

|| University of Maryland.

⊥ Present address: Pharmacokinetics, Dynamics, and Metabolism, St. Louis Laboratories, Pfizer Inc., Chesterfield, MO 63017.

[1] Abbreviations: ASBT, apical sodium-dependent bile acid transporter; EL, extracellular loop; IL, intracellular loop; ILBP, ileal lipid-binding protein; Ntcp, Na^+-taurocholate cotransporting peptide; SBAT, sodium-dependent bile acid transporter; TM, transmembrane; TMS, transmembrane segment; TCA, taurocholate.

10.1021/bi049270a CCC: $27.50
Published on Web 08/20/2004

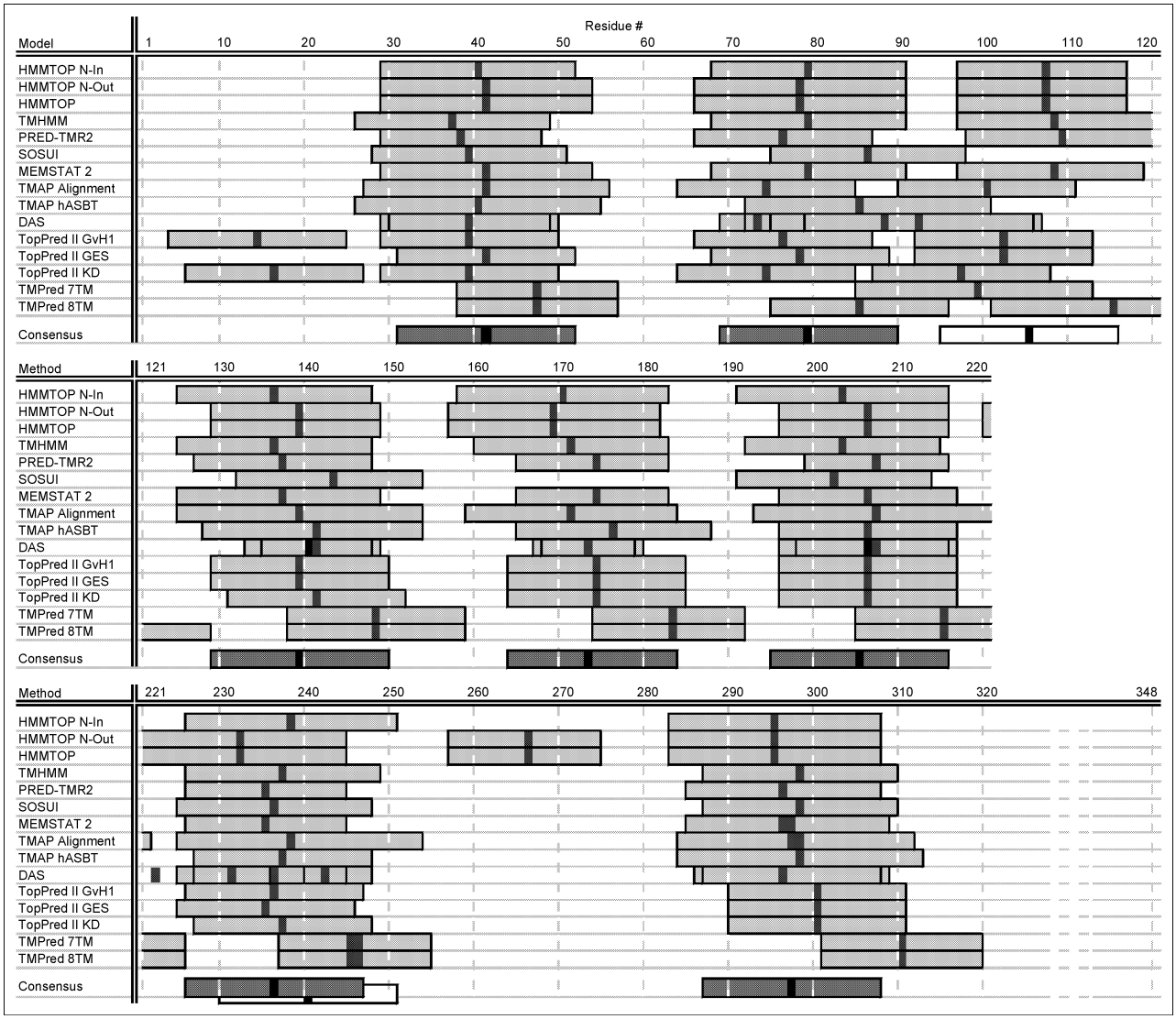


FIGURE 1: Schematic representation of TMS locations predicted by each of the 15 models and the consensus-scanning method. Light gray bars indicate predicted TMS regions for the 15 models and for the 7 TMS regions predicted by 100% (28/28) of the consensus-scanning scenarios. Midpoints of these 7 TMSs averaged over all 28 scenarios are residues 41, 79, 139, 173, 205, 236, and 297. Two additional white bars with midpoints at residues 105 and 240 indicate TMSs predicted by only 60.7% (16/28) and 10.7% (3/28) of the consensus-scanning scenarios. Black bands represent TMS approximate midpoints.

that are linked by merely two extramembranous amino acids. The amino acid length of TM3 and TM4 would indicate a β-sheet conformation for these segments to span the lipid bilayer membrane; a combination of α-helices and β-sheet TM segments is uncommon for polytopic membrane proteins, prompting further investigation.

To differentiate effectively between a seven and a nine TM model for ASBT, we report an N-glycosylation-scanning mutagenesis approach based on the divergent localization of specific protein segments within the two hypothetical models (Figure 2). Because N-glycosylation occurs only on the luminal side of the endoplasmic reticulum, this method has been used successfully to localize extracellular domains of receptors, channels, and transporters (*10*). After confirming that wild-type human ASBT contains one N-linked glycosylation site (Gln^{10}), we engineered the N-glycosylation consensus sequence (*NXS/T*) into a biologically functional aglyco-ASBT (N10D mutant). Our data show that the newly introduced consensus sequences at residues 113−118 and 266−272 can be glycosylated, strongly supporting a seven TM helical model of ASBT.

On the basis of a 7TM topology, we built a homology model utilizing the membrane attachment points of bacteriorhodopsin (BR) as a scaffold. The experimental data presented here further confirm the existence of a 7TM topology for ASBT and suggest possible ligand-binding domains in a three-dimensional (3D) homology model that may be used for further mutagenesis studies and rational inhibitor design.

EXPERIMENTAL PROCEDURES

Materials. ^{3}H-Taurocholate (50 Ci/mmol, TCA) was from American Radiolabeled Chemicals, Inc. (St. Louis, MO). Sulfo-NHS-LC-Biotin was from Pierce (Rockford, IL). Unlabeled TCA and other reagents were from Sigma (St. Louis, MO). Restriction enzymes and cell-culture supplies were obtained from Life Technologies, Inc. (Rockville, MD). COS-1 cells were obtained from the American Type Culture Collection (CRL-1650) and maintained and propagated in DMEM containing 10% fetal calf serum, 4.5 g/L glucose, 100 units/mL penicillin, and 100 μg/mL streptomycin. Calnexin antibodies were obtained from StressGen (Victoria, BC, Canada).

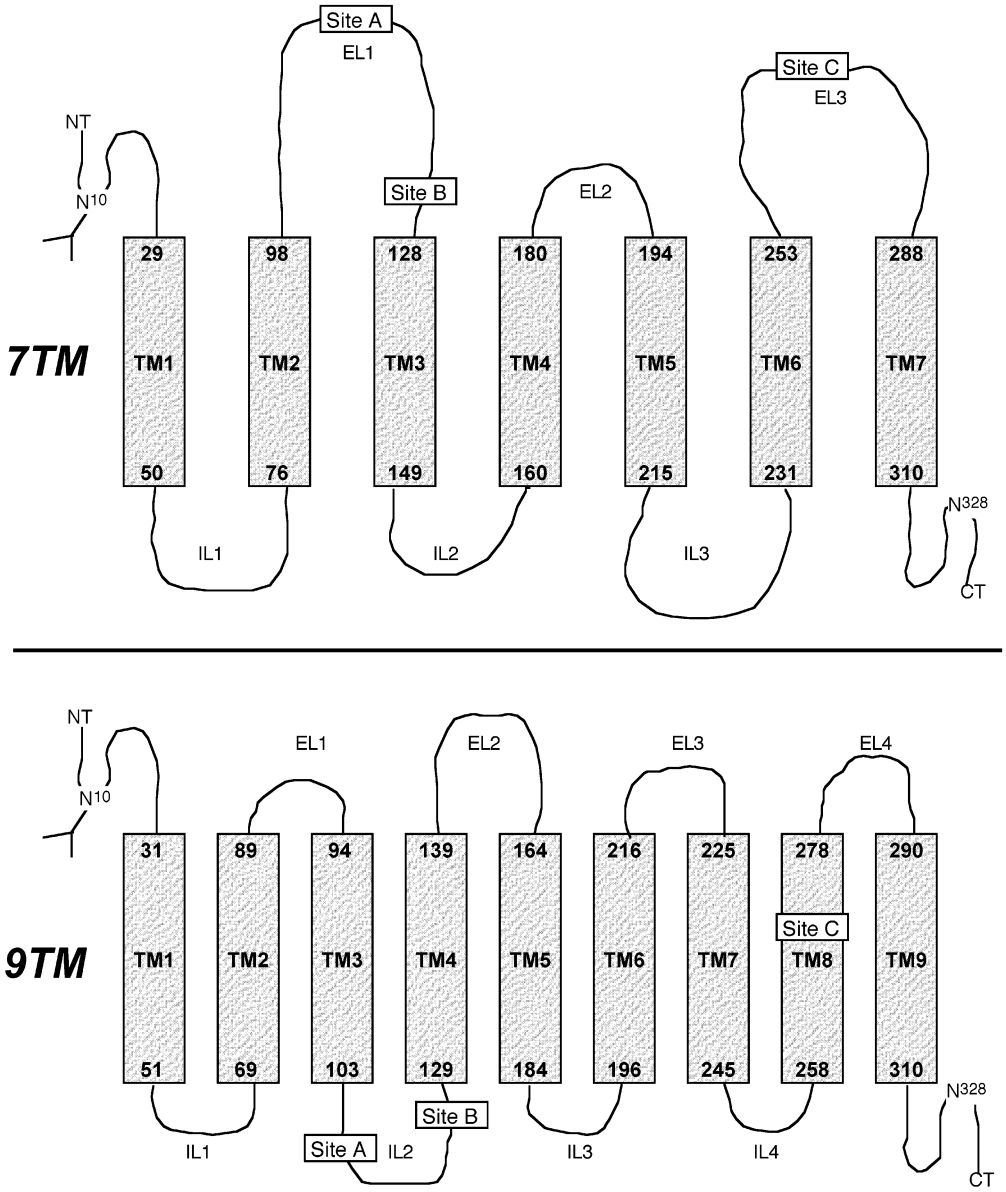


FIGURE 2: 7TM versus 9TM membrane topology model for ASBT and the strategic locations of the inserted new N-glycosylation sites. Site A, B, and C indicate the approximate locations of the engineered N-glycosylation sites. Their exact positions are delineated in Table 1. EL, extracellular loop; IL, intracellular loop; NT, N terminal; CT, C terminal; and TM, transmembrane domain. Numbers indicate the putative residues on the membrane border.

Antibody Preparation. Rabbit polyclonal antiserum was raised against the C terminus of human ASBT (S^{335}FQETNKGFQPDEK^{348}). Cys was added to the N terminus to allow conjugation to the carrier protein keyhole limpet hemocyanin (KLH). IgG was purified from rabbit serum by Protein A-agarose affinity chromatography.

Site-Directed Mutagenesis. Mutations were introduced using a QuickChange kit from Stratagene (La Jolla, CA). The pCMV5 vector containing hASBT cDNA was a kind gift of Dr. Paul Dawson, Wake Forest University (Winston-Salem, NC). Mutagenesis primers were synthesized by Operon (Alameda, CA). Plasmids were purified using a purification kit from Roche (Indianapolis, IN). All mutations were verified by DNA sequencing.

The endogenous N-glycosylation site (Gln^{10}) of human wild-type ASBT was eliminated by the mutation N10D to simplify interpretation. The resultant mutant plasmid was used as a template for the introduction of novel N-glycosylation sites. Additional Ala-substituted mutations were introduced at residues identified by both docking studies and multiple sequence alignment for their potential involvement in substrate binding (Figure 9). Transient transfections in COS-1 cells were conducted using the LipofectAMINE PLUS reagent (Invitrogen, Carlsbad, CA) and after 48 h processed for TCA uptake assays, Western Blot analysis, and cell-surface biotinylation (*3*). The specificity of the biotinylation reaction was verified by the fact that no band could be detected when COS-1 cells transfected with wild-type (wt)ASBT were subjected to all labeling steps in the absence of Sulfo-NHS-LC-Biotin. To ensure that Sulfo-NHS-LC-Biotin is indeed a membrane-impermeable reagent in our experimental protocol and specifically labels proteins on the cell surface, parallel studies were carried out followed by immunoprecipitation using anticalnexin. No biotinylated

calnexin could be detected (data not shown), whereas this abundant protein was easily detectable in control experiments with whole-cell lysate.

Consensus Topography Analysis of Primary Sequence. Putative membrane-spanning regions for the human ASBT (hASBT) were identified by a consensus prediction algorithm using simple majority voting and statistical analysis of 9 individual TM topology prediction methods for eukaryotic proteins: TMPred, TopPred2, DAS, TMAP, MEMSAT 2, SOSUI, PRED-TMR2, and the hidden Markov model-based programs TMHMM 2.0 and HMMTOP 2.0. Where available, input parameters were varied to generate additional prediction scenarios. TopPred2 was implemented with the "positive-inside rule" (*14*), the Kyte−Doolittle, and GES scales, to evaluate the validity of topology models derived from the hydropathy analysis. TMAP predictions were generated both with the single hASBT sequence and with a multiple sequence alignment of human, hamster, rat, and rabbit ASBT sequences. Even though it has been well-established that the N terminus of ASBT is extracellular, HMMTOP predictions were generated with default parameters and by specifying that the N terminus is positioned as both cytosolic and extracellular. Additionally, TMPRED generated two models for ASBT topology, one with eight TMSs and one with seven TMSs. In sum, a total of 15 models were generated. Then, a consensus-scanning approach initially investigated by Ikeda and colleagues (*15*) was employed to determine the most probable TMS for ASBT. Because each individual method has it strengths and weaknesses, the consensus-scanning approach was shown consistently to have a 10−20% higher accuracy of prediction for eukaryotic proteins compared to the single best prediction method (*15*).

Homology Modeling of Membrane Attachment Points and Extracellular Loops (ELs). The crystal structure of BR at 3.0 Å (*16*) was used as a scaffold for modeling the attachment points for extracellular regions in ASBT. For consistency and visualization purposes, only the TM domains of ASBT were modeled using BR as a template, substituting all ASBT amino acids for BR residues. Amino acid residues of Protein Data Bank (PDB) entry 1AT9 (*16*) were replaced by their predicted TM-domains corresponding residues within ASBT. All other residues were removed from the file. The obtained TM model was minimized until the root-mean-square (rms) value of the conjugate gradient was <0.1 kcal/(mol Å). To improve side-chain conformations, a molecular dynamics simulation was carried out (for 50 ps at 300 K), with fixed TM backbone atoms. ELs were modeled using the Sybyl LoopSearch routine.

Fold Prediction for N- and C-Terminal Domains. Protein domains that lacked significant homology with PDB entries, namely, the N- and C-terminal domains of ASBT, were predicted via remote homology threading using THREADER2 (*17*). All command line options were kept at default, except the number of possible paths, which was set to 120 instead of the default 60. The results were sorted by *Z* scores of pairwise and solvation energies, allowing for the percentage of sequence and structure matched (column 13 in the standard THREADER2 output file). All folds with a *Z* score > 2.7 (significant) and *Z* score > 2.5 (borderline significant) were investigated for biological significance (*18*). The three structures with the highest *Z* score were selected for incorporation in the final model and subjected to energy optimization. Amino acids 29−32 of the N-terminal were superimposed on amino acids 1−4 of the TM model, which were then removed, and a connection was made between Asn^{28} and Ile^{29}. Hydrogen atoms and Kollmann charges were added to all atoms before initiation of refinement procedures. A similar procedure was carried out for connecting the C-terminal protein segment.

Structure-Refinement Protocol. The model with frozen TM domains was energy-minimized until the conjugate gradient was <0.1 kcal mol^{-1} $Å^{-1}$. The structure was further refined using a simulated annealing protocol (*19*) with minor modifications. The structure was heated to 1500 K followed by 10 ps of constant temperature molecular dynamics simulation at 1500 K. A total of 10 structures was extracted from the trajectory by sampling every 1 ps. Each structure was cooled to 300 K in 50 ps, followed by 50 ps of constant temperature molecular dynamics simulation at 300 K. Hydrogen bond lengths were constrained using the SHAKE algorithm. The resulting structures were fully energy-minimized. Of the 10 structures obtained, 4 were considered biologically relevant. The other 6 structures were rejected based on the movement of ELs into the TM region. The average structure of the 4 final results was taken and the obtained structure was submitted to 50 ps of a constant temperature molecular dynamics simulation at 300 K.

In Silico Mutagenesis. Wong and colleagues (*3*) previously identified a dysfunctional ASBT isoform in a patient diagnosed with Crohn's disease. DNA sequencing and comparison to wtASBT revealed a single C−T transition resulting in a proline to serine substitution at amino acid position 290. In transfected COS-1 cells, the single amino acid change abolished TCA transport activity but did not alter the synthesis or subcellular distribution of the transporter, indicating that this polymorphism directly affects the transporter function. To further confirm the physiological validity of our model, we simulated the P290S point mutation in our ASBT 3D model by in silico mutagenesis, substituting S for P^{290} using standard bond distances and angles. Subsequently, the mutant protein was subjected to a molecular dynamics protocol as described above in the refinement procedures.

Ligand Docking. To identify potential protein motifs involved in the affinity and transport of natural ligands by ASBT, we employed an automated docking routine (*17*). AutoDock version 2.4 uses a Monte Carlo simulated annealing approach to probe the surface of a macromolecule with a small ligand, while an energy scoring function reports the interaction (docking) energy. Cholic acid, a natural bile acid substrate of ASBT was used as the probe ligand in the docking protocol. A positive and negative control were used in these studies; the NMR structure of ileal lipid binding protein (ILBP, PDB 1EAL) (*20*), a well-characterized cytosolic bile acid binding protein, was used as a positive control template, whereas the dysfunctional mutant P290S served as a negative control. Ranked cluster analysis was performed to identify positive interactions between the ligand and macromolecule.

RESULTS

Consensus Scanning Favors a 7TM hASBT Topology Adhering to the "Positive Inside" Rule. A total of 9 topology prediction programs generated a total of 15 models (Figure

1). 13 models predicted between 7 and 9 TMS, whereas 1 model (SOSUI) predicted 6 TMS and another (DAS) 17, although TMS overlap effectively indicated a 7 or 8 TMS topology (Figure 1). The first intermodel variation appears at the N terminus (residues 1−30), where only TopPred II predicts a TMS. A second variation is a TMS (residues 255−280) predicted only by HMMTOP. The third and most interesting incongruous region (residues 65−120) is a highly hydrophobic region that is sufficiently long to encompass 2 TMS. This region contains the widest spread in predicted TMS midpoints among the 15 models.

Analogous to the 7TM topology initially reported by Wong and colleagues (*1*), the consensus topology presented here adheres to the "positive inside" rule (*14*). The total charge for the extracellular portion of the protein is −3: N terminus (−3), EL1 (−3), EL2 (+3), and EL3 (0). Whereas, the cytosolic domains contribute to a net positive charge of +4: IL1(+3), IL2 (0), IL3(+1), and C terminus (0). Interestingly, a previously reported 7TM model by Oelkers and colleagues (*21*) [subsequently refuted by Hallén et al. (*9*)] violated this rule, illustrating the need for multiple topography analysis approaches.

N-Glycosylation Mutagenesis Supports a 7TM Topography for ASBT. hASBT contains two N-linked glycosylation epitopes at Asn^{10} and Asn^{328} (Figure 2). Two single mutants (N10D and N328D) and one double mutant (N10/328D) were prepared (Table 1A) and transfected in COS-1 cells, along with wtASBT. wtASBT (lane 2) is detected mainly in two equally stained bands (Figure 3A), both the unglycosylated (38 kDa) and glycosylated (41 kDa) forms, which can be removed completely after treatment with tunicamycin (TNM, lane 3). The mobility of N10D (lanes 4 and 5) is identical to that of the wt treated with TNM, whereas N328D (lanes 6 and 7) migrates identical to wt with or without TNM treatment. In addition, the mobility of the double mutant N10/328D (lane 8) is similar to that of N10D. Cell-surface expression was ascertained by biotinylation (Figure 3B). Densitometric analysis suggests that the band intensity of N10D was roughly equal to the combined intensity of the two bands of wtASBT (parts A and B of Figure 3), suggesting that deglycosylation may not significantly alter the total or membrane-surface expression. Cell-surface expression of N328D is comparable to that of the wt and N10D, because the densitometric ratio between the ASBT and calnexin bands (Figure 3C) is equal to that of wtASBT.

N10D retained 48% of the wt TCA uptake activity (Figure 3C). N10/328D was comparable to N10D, whereas the single mutant N328D was comparable to the wt. Kinetic analysis revealed that K_T was identical for wt and N10D (11.3 ± 1.9 and 12.8 ± 2.1 μM, respectively), whereas J_{max} was significantly reduced by 50% (312.5 ± 12.8 and 156.5 ± 6.7 pmol min^{-1} mg $protein^{-1}$, respectively).

When taken together, our data suggest that (1) Asn^{10} is the only N-glycosylation site in ASBT and the removal of the sugar moiety may not influence transport activity and (2) Asn^{328} is not an N-glycosylation site nor is it likely located in a functionally important region. This is in good agreement with the cytoplasmic location of the C terminus.

N-Glycosylation-Scanning Mutagenesis Suggests a 7TM Topology. To distinguish between a 7- and 9TM topology, we engineered N-glycoslyation sites at strategic locations in hASBT N10D mutant (Figure 2). In the 7TM model, the inserted potential N-glycosylation sites are in EL1 (site A = N^{113}ISNYS^{118} and site B = N^{124}LS) and EL3 (site C = N^{266}TSLNST^{272}); in the 9TM model, however, these sites would be present in IL1 and TM8 and not accessible to N-glycosyl transferase.

Table 1: Sequences and Design of Mutagenesis Primers[a]

A	wt	P N S C V D N^{10} A T V
		ccg aac agc tgt gtg gac aat gca aca gtt tg
	N10D	P N S C V D ***D^{10}*** A T V
		ccg aac agc tgt gtg gac ***g***at gca aca gtt tg
	wt	P E S K E N^{328} G T E P E
		cca gag agc aaa gaa aat gga acg gag cca gag
	N328D	P E S K E ***D^{328}*** G T E P E
		cca gag agc aaa gaa ***gac*** gga acg gag cca gag
B	N10D	G T A S N^{113} I L A Y W V D G
		gga act gcc tcc aat atc ttg gcc tat tgg gtc gat ggc
	Site A	G T A S N^{113} I ***S N^{116}*** Y ***S*** V D G
		gga act gcc tcc aat atc t***cc aa***c tat t***c***g gtc gat ggc
	N10D	D G D^{124} M D L S V S
		gat ggc gac atg gac ctg agc gtc agc
	Site B	D G D^{124} M ***N^{124}*** L S V S
		gat ggc gac atg ***a***ac ctg agc gtc agc
	N10D	G M Q N^{266} T Q L C S T I V Q
		ggg atg cag aac acg cag cta tgt tcc acc atc gtt cag
	Site C	G M Q N^{266} T ***S*** L ***N^{270}*** S T I V Q
		ggg atg cag aac acg ***tc***g cta ***aa***t tcc acc atc gtt cag
C	E281A	ctc tcc ttc act cct g***ca*** ga***a*** ctc aat gtc gta ttc ac
	E282A	c tcc ttc act cct gag g***cc*** ctc aat gtc gta ttc ac
	E282D	ctc tcc ttc act cct gag **gac** ctc aat gtc gta ttc acc
	E282K	ctc tcc ttc act cct gag **aag** ctc aat gtc gta ttc ac

[a] (A) Primers for the removal of the endogenous N-glycosylation sites. The numbers indicate the endogenous N-glycosylation sites. (B) Primers for the introduction of the strategic N-glycosylation consensus sites. The numbers indicate the inserted potential N-glycosylation sites. Italics indicate the mutations. Site A is located between residue 113 and 116 in EL1; site B is located on residue 124 in EL1; and site C is located between residue 266 and 270 in EL3. (C) Primers for amino acid substitutions in putative binding domain 1.

Site A mutants have a high molecular-weight band (Figure 4) corresponding to glycosylated hASBT. This band is absent after preincubation with TNM, suggesting successful glycosylation of Asn^{113} or Asn^{116} by insertion of a consensus sequence at site A. Site B mutants are unglycosylated in the presence and absence of TNM. According to a 7TM model, Asn^{124} localizes extracellularly on EL1 and could be glycosylated. However, the active site of oligosaccharyl transferase has a sharp distance threshold above the membrane, and N-glycosylation rarely occurs if the EL is shorter than 30 residues or if the acceptor site is within 10 residues of the membrane (*22*). Although EL1 and EL3 contain 31 and 33 amino acid residues, respectively, Asn^{124} is just 6 amino acids away from the putative

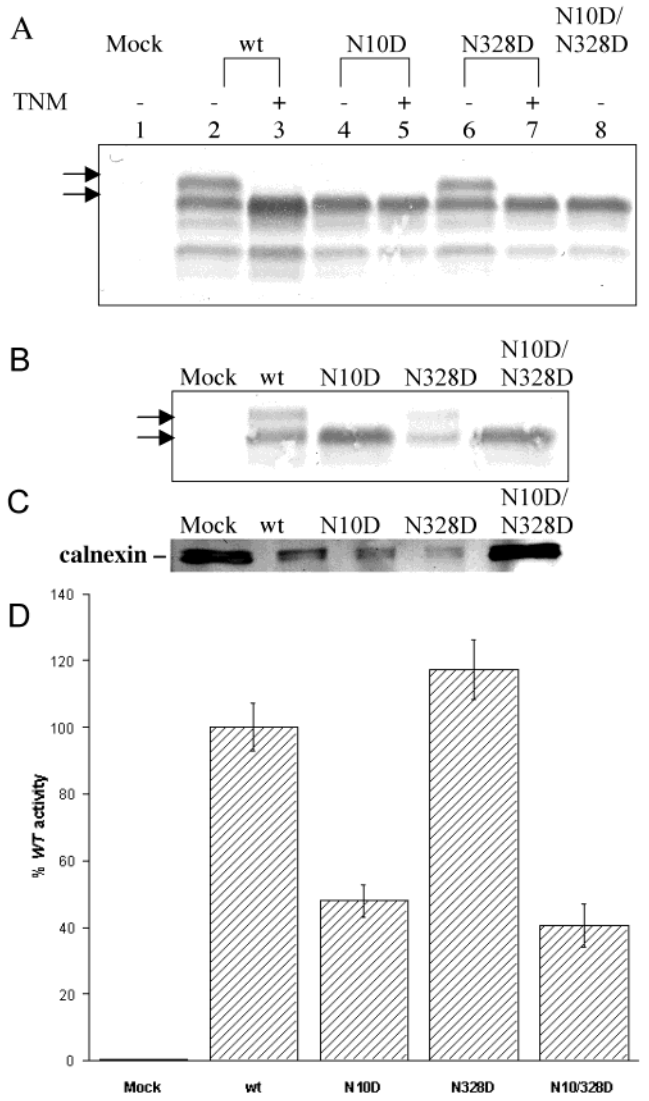


FIGURE 3: ASBT has only one N-glycosylation site occurring at Asn10. (A) N-glycosylation of ASBT. Transient transfection was performed as described in the Experimental Procedures. The transfected COS-1 cells were incubated at 37 °C for 48 h in the presence (+) or absence (−) of N-glycosylation inhibitor TNM. Lysates were subjected to SDS−PAGE, electroblotted onto PVDF membrane, probed with rabbit anti-hASBT antibody, and visualized by a chromogenic Western Blotting kit (Vector Lab). (B) Cell-surface expression of wtASBT and mutants detected by streptavidin-conjugated alkaline phosphatase after cell-surface biotinylation, immunoprecipitation, electrophoresis, and blotting. Arrows indicate the expected molecular weight of glycosylated (41 kDa) and unglycosylated (38 kDa) ASBT. (C) Blots from B were stripped and probed with anti-calnexin. (D) Uptake activity of ASBTs from which the endogenous N-glycosylation consensus sites have been removed. ^{3}H-Taurocholate uptake in COS-1 cells transfected with pCMV5 vector containing the indicated cDNA constructs. Uptake medium was spiked with 5 μM ^{3}H-TCA (0.2Ci/mmol). Results were given as a percentage of the mutant uptake relative to the wt (81.3 $\pm$ 5.7 pmol min^{-1} mg protein^{-1}). Each bar is the mean $\pm$ standard error of three to four different measurements. Mock = pCMVβ.

membrane border, explaining its failed glycosylation. Site C falls within EL3 according to a 7TM model and within TM8 according to the 9TM model (Figure 2). Glycosylation is more efficient at this position relative to site A, confirming extracellular localization. Cell-surface biotinylation (Figure 4B) confirms that, although none of the mutants were functional (data not shown), they were all expressed at the cell surface at densities comparable to that of the wt,

Table 2: Comparative Modeling of the EL and IL ASBT Domains[a]

domain	residues	best fit[b]	percent homology	rms value	net charge
EL1	99−129	1CPC:K/Ser127	41.6	0.42	−3
EL2	181−190	1ZYM:B/Ala112	34.2	0.23	+3
EL3	254−286	2MMZ:D/Ala71	50.0	0.31	−1
IL1	51−72	1GSE:B1/Asp101	41.6	0.25	+3
IL2	150−159	6TAA:Ala302	39.4	0.25	0
IL3	216−226	1BHF:A/Thr425	41.9	0.24	+1

[a] Loopsearch Subroutine in Sybyl 6.3 was used to identify sequence homology against structures in the Protein Data Bank. Results were scored by their rms values and percent homology. [b] Protein with the best possible fit is indicated by its PDB entry code, followed by the chain identification code and the first amino acid residue of the sequence used to replace the target sequence in ASBT.

suggesting that protein folding and sorting was not affected. When taken together, the fact that both site A and C are capable of incorporating a sugar moiety confirms their extracellular localization as predicted by the 7TM model.

Homology Modeling. An initial BLAST search of ASBT segments against the PDB, using the WU-BLAST routine, retrieved, as expected, no significant hits. Experimental data suggested that ASBT contains 7 TM domains, thereby sharing topological (but not mechanistic) similarity with the 7TM-domain photoreceptor BR. Sequence alignment of the 170 putative ASBT TM residues against BR TM regions using the Multalin routine revealed that 15.3% (26) of all residues were highly conserved ($N > 90\%$) and 75.3% (128) of residues were significantly conserved ($50\% < N < 90\%$), indicating that BR is not an ideal scaffold for modeling ASBT TM regions but can provide suitable attachment points for modeling the extracellular protein domains (Table 2, Figure 5).

PROCHECK (*23*) was used to calculate quality assessment data and Ramachandran scores for the model (Table 4). This was compared to structural data of other key membrane proteins determined by crystallography (e.g., aquaporin, KcsA), NMR (sodium channel), or homology modeling (Glut1). Tables 4 and 5 show that the ASBT structure is at least as good as or, in some cases, better than experimentally determined structures.

Modeling of the N- and C-Terminal Folds Reveals High Structural Similarities. Mutational studies on the glucose transporter, SGLT1, demonstrate that extracellular terminal domains play an essential role in initial substrate binding (*24*), and a similar importance has been ascribed to the terminal protein domains of other transporters. Obtaining a satisfactory model for this pivotal protein domain is imperative. No sequence homology of the N- or C-terminal domains with any structure in the PDB was retrieved using WU-BLAST (data not shown), indicating that homology models were unfeasible. We used the THREADER2 algorithm (*17*) to predict protein folding of terminal domains (Table 3). Both 1ERD,0,0 (N terminal) and 1DTR,0,2 (C terminal) have a mainly α, nonbundle structure (CATH classification) and were used as templates to model the terminal domains. The resulting structures for the extracellular N-terminal domain (Met^1-Val^{32}) and the cytosolic C-terminal domain (Phe^{304}-Lys^{348}) were modeled using MODELLER4 (*25*) (Figure 6). Recently, Sun and colleagues (*12*) reported a β-turn structure in the C-terminal domain to be required for apical membrane sorting of ASBT. The location of this β

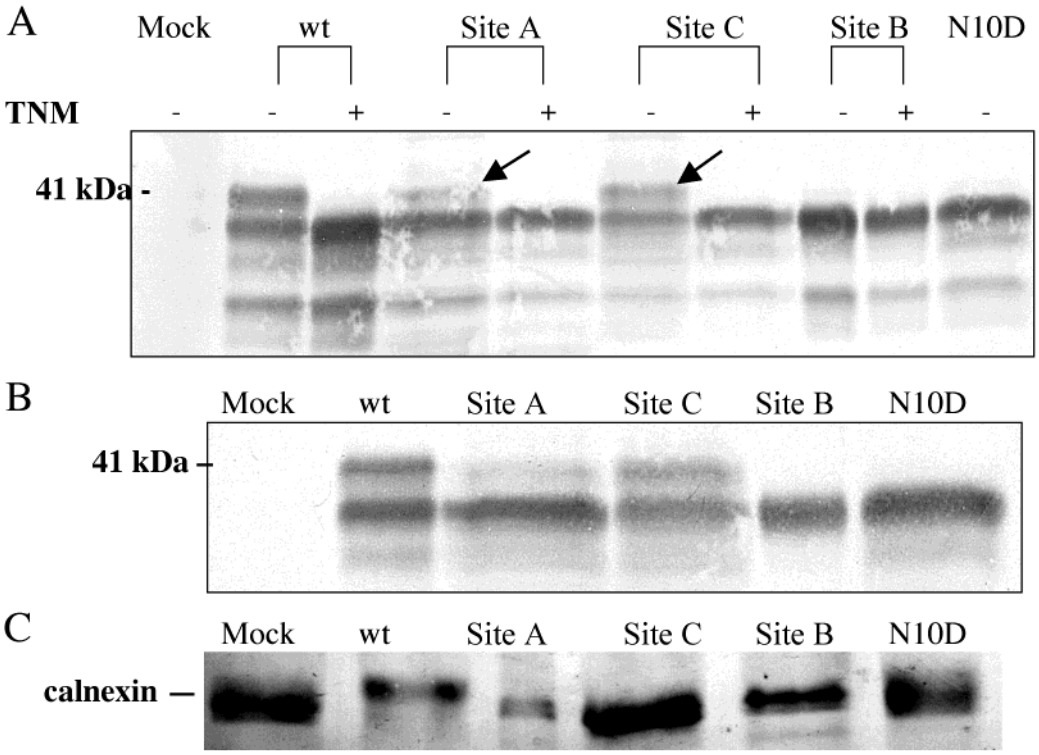


FIGURE 4: Analysis of aglyco-ASBT with the engineered N-glycosylation sites. (A) Representative immunoblotting results on whole-cell lysates. (B) Cell-surface expression of wt hASBT and mutants. Arrows indicate that N-glycosylation occurs on the inserted site. (C) Anticalnexin control blots.

Table 3: THREADER Fold Prediction Results for the N- and C-Terminal Domains[a]

Z score	PDB code, chain, domain	fold target name	CATH classification
		N Terminal	
2.75[b]	1ERD, 0, 0	pheromone er-2	mainly α, nonbundle
0.89	3MON, A, 0	monellin	not classified yet
0.69	1TME, 4, 0	encephalomyelitis virus (Da strain)	few structures, irregular
		C Terminal	
3.26[b]	1DRT, 0, 2	diphtheria toxin repressor	mainly α, nonbundle
2.81[b]	1FXD, 0, 0	ferredoxin II	α/β, 2 layer sandwich
2.60	1PNR, A, 1	repressor hypoxanthine-purf-operator	mainly α, nonbundle

[a] Top three *Z* scores (pairwise solvation/energy) for both terminals are shown. The N-terminal shows a borderline significant hit, and the C-terminal shows a significant hit. [b] Significant hits.

Table 4: PROCHECK Ramachandran Scores for ASBT and Various Membrane-Spanning Proteins (Channels, ABC, and Solute Transport Proteins)

name[a]	PDB access code	ref	Ramachandran plot (%) core	allowed	generous	disallowed
ASBT			83.3	13.4	3.3	0.0
BACR	1AT9	*16*	88.3	9.6	2.0	0.0
OPSD	1f88	*36*	80.9	17.2	1.6	0.2
sodium channel[b]	1QG9	*37*	72.2	27.8	0.0	0.0
AQP1	1IH5	*38*	70.7	20.9	8.4	0.0
KcsA	1BL8	*39*	74.7	24.1	1.2	0.0
Glut1[c]	NA	*40*	81.5	16.9	1.7	0.0
lac permease[d]	1PV6	*42*	79.0	19.1	1.4	0.5
G3PT[e]	1PW4	*43*	85.4	13.5	1.1	0.0

[a] ASBT, apical sodium-dependent bile acid transporter; BACR, bacteriorhodopsin (3.0 Å); OPSD, opsin photoreceptor (bovine rhodopsin) (2.8 Å); AQP, aquaporin (3.70 Å); KcsA, bacterial potassium channel (3.20 Å); Glut, facilitative glucose transporter. [b] NMR solution structure. [c] Homology model based on bacterial lac-permease structure−function studies. [d] *E. coli* lactose permease (3.50 Å). [e] Glycerol-3-phosphate transporter (3.30 Å).

turn (Asn^{340}-Phe^{343}) coincides with a predicted turn in the α-helical bundling of our C-terminal domain (Figure 6C).

Two Putative Ligand-Binding Sites Are Identified Using Simulated Annealing Docking. The protein surface of ASBT was probed with its natural substrate, cholic acid, for energetically favorable docking regions utilizing a simulated annealing and docking algorithm within AutoDock 2.4 (*17*). Two specific binding sites were detected (parts A and B of Figure 7). The first binding site (Figure 7A) is considered highly significant, because a cluster of three ligands was reported to form an energetically favorable interaction with ASBT outside EL3. Glu^{282} and possibly Leu^{283} are involved in the formation of dynamic hydrogen bonds with the 12α-hydroxyl group on the ligand, whereas residues in the vicinity of Asn^2 on the N-terminal domain are capable of forming dynamic hydrogen bonds with either the 3- or 7α-hydroxyl groups (Figure 8). However, previous 3D-QSAR studies indicate that the 3-hydroxyl group is not essential for transport by hASBT (*26*). Therefore, we speculate Asn^2 may provide increased affinity for hASBT to substrates containing a 3- or 7α-OH moiety but may not play an essential role in the overall substrate translocation process. The exact role of this residue remains to be further defined.

The backbone of Leu^{283} can form an additional bond with the ligand 7α-OH group. Interestingly, the 3α- and 12α-hydroxyl groups on cholic acid were primarily involved in hydrogen-bond formation with the protein, which corroborates the established structure−activity relationships for ASBT substrates. It is well-established that at least one

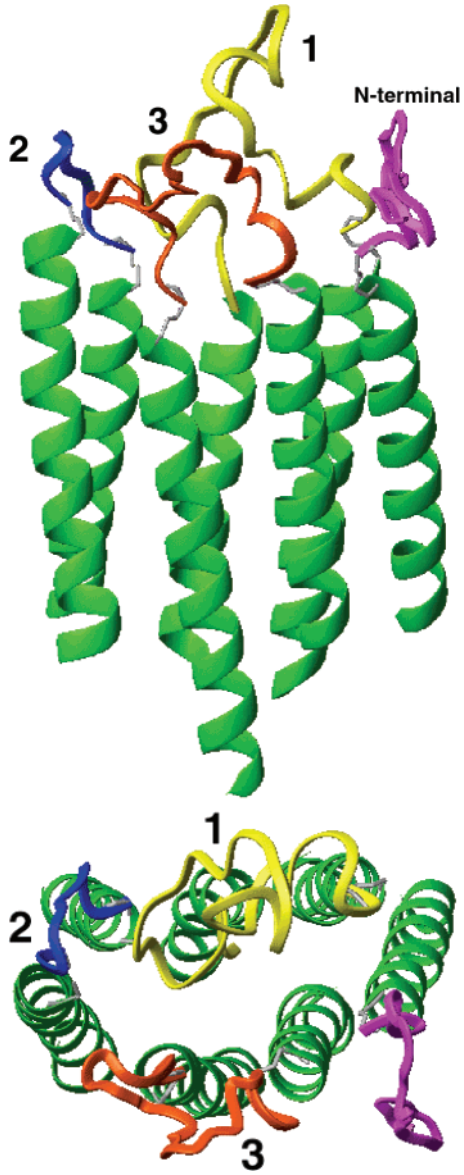


FIGURE 5: Side- (A) and top- (B) view models for the 3D structure of ASBT consisting of 7TM domains. Numbering of these domains is counterclockwise, viewed from the extracellular side, analogous to the numbering of the TM regions of BR.

hydroxyl group on the sterol nucleus of bile acids is essential, regardless of its position (3, 7, or 12). Affinity increases by the presence of two hydroxyl groups but does not significantly increase by adding a third moiety (*27*). The current binding model confirms this hydroxyl group topology.

The molecular and electrostatic surface potentials (parts A and B of Figure 7) suggest the formation of an opening between the N-terminal domain and EL3, referred to as the "cleft area". The energetically most favorable binding pocket for cholic acid (site B) is located within the cleft area between EL3 and the N-terminal domain (Figure 7B). The carboxylic acid moiety of cholic acid is capable of forming a stable interaction with Val^{21} and Asn^{27}, and the 3- or 7α-hydroxyl groups can form hydrogen bonds with Asn^{10} and Ala^{11}. The N-glycosylation site for ASBT (Asn^{10}) is located on the second helix of the N-terminal domain directly opposite EL3 (Figure 5) and facing away from the cleft region. The orientation of this moiety would suggest that the glycosylation entity does not play an active role in ligand recognition.

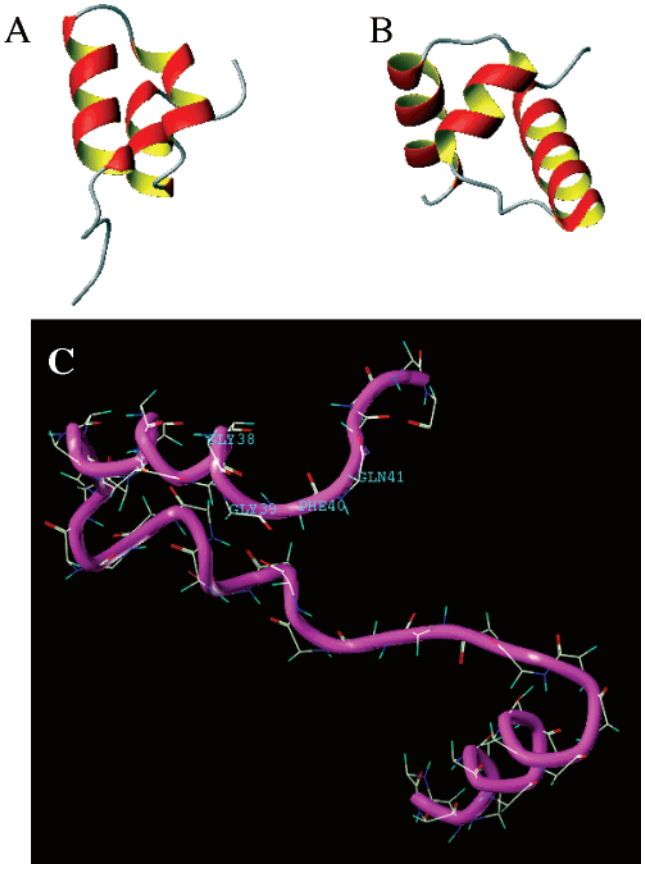


FIGURE 6: Remote homology threading predictions for the N- (A) and C- (B) terminal domains of ASBT, illustrated with MOLMOL. (C) Visualization of the β-turn structure in the C terminus of hASBT as observed by Sun and co-workers in rat ASBT (*12*).

It should be noted that this conformation is further stabilized by hydrophobic interactions between amino acids on EL3 and the two methyl groups on the sterol backbone.

Dimeric bile acid analogues, where a bile acid carboxylic acid moiety is coupled via a spacer to the 3α-hydroxyl group of a second bile acid molecule, are high-affinity inhibitors of ASBT (*28*). The rationale for using dimeric bile acids was to occupy "more than one transporter site on the ileal-bile-salt-transporter complex" (*28*). Interestingly, the most active compound has a 13-atom spacer, which has a calculated interatomic distance between the 3α-hydroxyl oxygen atom and the carboxylic carbon atom of 14.2 Å (data not shown). The respective interatomic distance between cholic acid in site A (Figure 7A) and site B (Figure 7B) is 13.4 Å, suggesting that dimeric bile acids could bind simultaneously to the two sites proposed in our model. These results can be taken as additional biological validity and relevance of this model. It should be noted that the cleft area is not longitudinally aligned with the TM core, suggesting that the binding site does not function as a "channel-like" selectivity filter and could favor a translational or vectorial movement of the substrate involving multiple consecutive binding sites. We speculate that subtle conformational changes in the protein elicited by either sodium or ligand interactions would expose consecutive binding sites to the ligand as it traverses the membrane.

Model Validation: In Silico Mutagenesis of a P290S Polymorphism Confirms the Loss of Activity of the Mutant. We next sought to determine whether our model agrees with

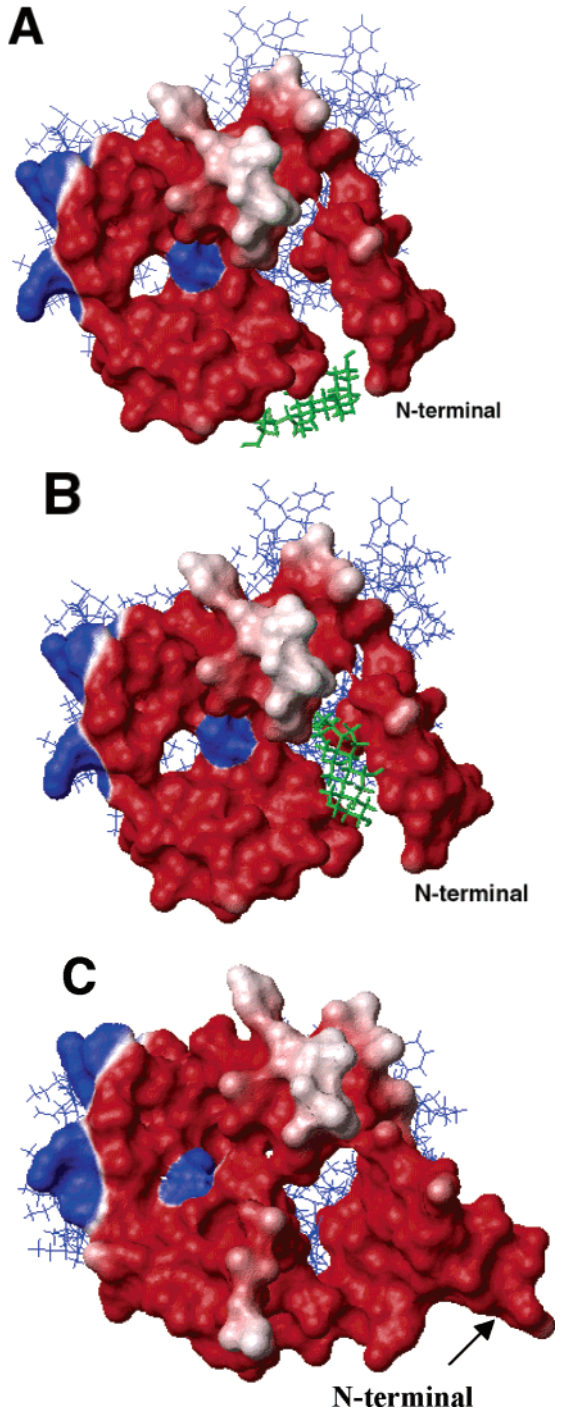


FIGURE 7: AutoDock results superimposed onto electrostatic surface-potential plots of hASBT. Ligand-binding domains 1 (A, predominantly bound to EL3) and 2 (B, substrate-localized within the cleft area) are visualized using an isopotential surface plot of the extracellular ASBT regions. (C) Simulation of a polymorphism, P290S, associated with bile acid malabsorption syndrome. The high-energy binding site between EL3 and the N-terminal domain undergoes a conformational change preventing the substrates from entering the cleft region. Electrostatic potential fields were calculated using MolMol, which solves the Poisson−Boltzmann equation and visualizes the electrostatic potential mapped on rendered molecular surfaces (*41*). The proteins were analyzed at pH 7.0, considering all Lys and Arg residues protonated and all carboxyl groups dissociated. A charge of +1e (e = proton charge) was assigned to the ϵ nitrogen of each lysine, and +0.5e was assigned to each guanidinium nitrogen. The endstanding α-amino group was given a charge of +1.0e, and the oxygen atoms of all aspartic and glutamic acid residues as well as the endstanding carboxylate group have a charge of −0.5e each.

experimentally determined functional and biochemical characteristics. The model was further validated by in silico mutagenesis by simulating a pathogenic polymorphism that leads to bile acid malabsorption syndrome (*21*). To study the conformational consequences of this polymorphism, we replaced amino acid residue Pro^{290} with Ser and re-equilibrated the mutant by subjecting it to constant temperature molecular dynamics simulations. Proline residues are known to perturb the structure of helices by introducing a kink between its preceding and trailing segments. It is well-established that local flexibility of a proline kink can confer an important role on the proline-containing helix in the conformational changes related to the function of the protein. A surface plot of the resulting structure (Figure 7C) indicates that the conformation of the cleft area between the N-terminal domain and EL3 has significantly changed, effectively closing the cleft area and highly modifying binding domain A for cholic acid. Simulated annealing of cholic acid on the surface of the P290S mutant resulted in a significantly decreased binding energy at site A and loss of binding at site B. These results suggest that binding site B is essential in the molecular-transport mechanism of ASBT. When these data are combined, they show that our model can explain the conformational and functional consequences of a hereditary defect leading to bile acid malabsorption.

Model Validation: Site-Directed Mutagenesis of Putative Binding Domain A Confirms the Importance of Glu282 in Ligand−Protein Interactions. To further characterize the role of putative binding domains in the ASBT−substrate interaction, we performed mutagenesis on relevant amino acid residues in binding domain A, specifically Glu^{282} (Figure 8). Sequence comparison of rodent and hASBT (Figure 9A) indicates that hASBT uniquely features a Glu instead of Asp at position 282. E/D282 is flanked by Glu281, creating a strong negatively charged pocket. Our docking experiments strongly suggested a critical role for E282 in ligand−protein interactions. Uptake studies of the mutants in the presence or absence of sodium (Figure 9B) reveal that E282A or E282K leads to an inactive protein that is expressed at the cell surface (data not shown). Substitution of an Asp for Glu at position 282 reduced activity by 52%. In a mutagenesis study by Hallén and colleagues (*13*), a D282E substitution in the mouse ASBT did not affect TCA uptake activity compared to that of the wt; however, wt mouse ASBT reportedly had a 2-fold higher uptake activity compared to hASBT. Interestingly, E281A is not a loss-of-function mutant, as observed with E282A, but appears to have an effect on TCA uptake. Although E281 was not predicted to play a role in ligand interactions, its charge and close proximity to a proposed critical residue may have an electrostatic resonance effect that would influence the overall binding efficiency to a ligand.

Kinetic values for wt hASBT (Figure 9C) are in good agreement with those (13−17 μM) reported previously (*3*, *29*). The E281A mutant has slightly increased K_t values with 42% of wt J_{max} values, respectively. The E282D K_t is comparable to that of the wt but at approximately half of its J_{max}. Thus, we could speculate that the reduction in TCA uptake (Figure 9B) could be attributed to a reduction in the maximal transport velocity. Immunoblotting confirmed cell-surface expression (data not shown). Therefore, we assumed that the inactivity of these mutants is not due to synthesis

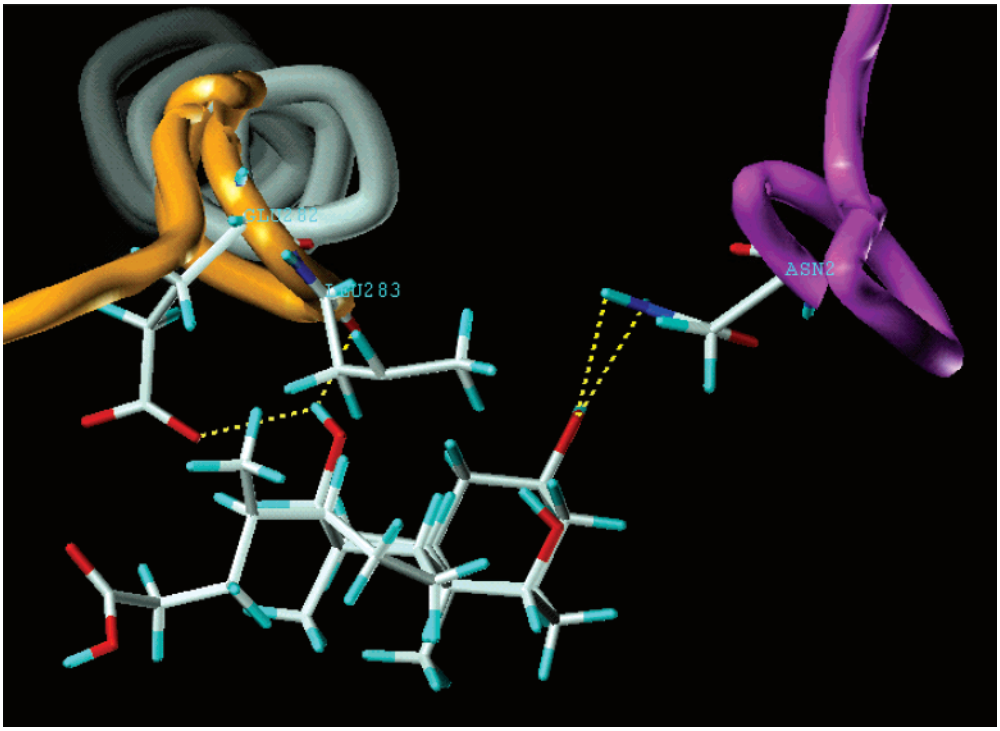


FIGURE 8: Representative visualization of putative binding domain 1. ASBT amino acid residues are shown as a tube ribbon plot of the main-chain carbon atoms; side groups are only shown for those residues that display dynamic intramolecular hydrogen bonds with the natural ligand cholic acid, namely, Glu282, Leu283, and Asn2. In this representation, Asn2 forms hydrogen bonds with the 3α-OH, although alternate conformations involving 7α-OH were observed.

and trafficking defects but to mutation-induced changes on functionally relevant regions of ASBT.

DISCUSSION

ASBT plays an integral role in the enterohepatic circulation of bile acids in the gastrointestinal tract. Its inhibition elicits compensatory hepatic bile acid synthesis from cholesterol, thereby lowering plasma cholesterol levels (*30*). Drug intervention aimed at blocking ASBT offers a specific therapeutic strategy against hypercholesterolemia and related cardiovascular diseases, and at present, several ASBT inhibitors are in clinical trials (*31*). Before rational drug design of novel inhibitors can be consistently applied to ASBT, molecular insights into its structure must be obtained. The aim of the current study was to develop a comprehensive 3D model of ASBT using a homology-modeling approach.

To determine the most suitable template and TM orientation for ASBT homology modeling, it was essential to confirm its membrane topology. We employed a two-tiered bioinformatics and glycosylation-scanning approach to effectively distinguish between a 7TM and 9TM topology. A recently developed consensus scanning technique was applied to predict the membrane topology of ASBT (*15*) using 15 different computational scenarios (Figure 1). This analysis effectively highlights the foundation for the 7TM versus 9TM controversy; namely, the regions between ASBT residues 65 and 120 are sufficiently hydrophobic for most algorithms to predict two individual TM segments in this region resulting in an 8TM configuration (Figure 1). Given that prior experimental evidence has confirmed an odd number of TM segments, the probable topologies for ASBT are either 7TM or 9TM. A forward scanning approach, which combines all methods and loosely simulates the cotranslational insertion of proteins into membranes, favored a 7TM configuration.

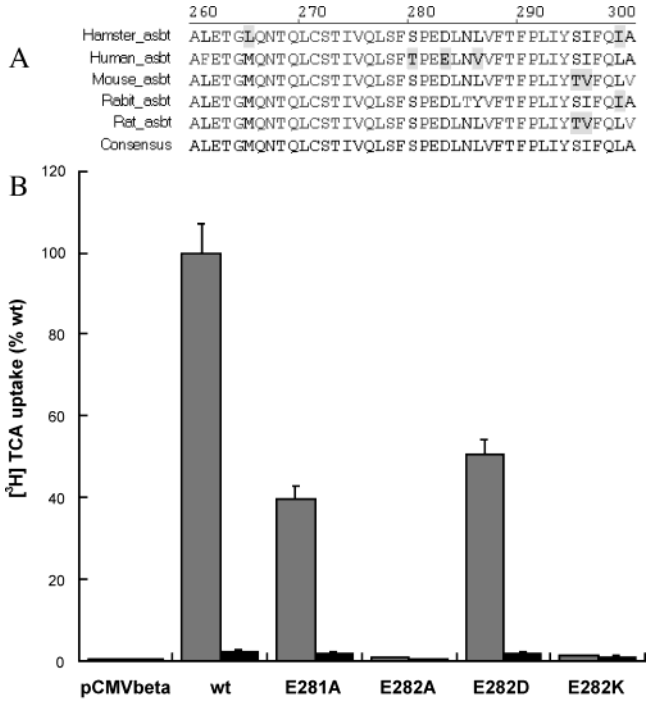
Hallén and co-workers have presented biochemical data that support a postulated 9TM model over a 7TM model (*9*, *13*, *32*). We applied an N-glycosylation-scanning approach to effectively differentiate between the putative 7TM and 9TM models for ASBT (Table 1 and Figure 2). Our results clearly indicate that N-glycosylation occurs at two of the three engineered sites (Figure 3), providing support for a classic 7TM model and challenging a 9TM model. However, we cannot rule out the possibility that a substrate-sensitive reentrant loop domain, as present in the glutamate transporter (*33*), may be responsible for the dynamic coexistence of both 7TM and 9TM models.

N-glycosylation-scanning analysis on NTCP by Hallén and colleagues (*34*) revealed that L265N has a small but significant degree of glycosylation. According to the 9TM model, this site would be buried within the membrane in TM8; contrarily, this site is exposed on EL3 in the present 7TM model, adequately far away from the membrane to be accessible by the glycosylation machinery. Other mutants in this study, e.g., E89S and G97N, were not glycosylated, and these results are consistent with a classical 7TM topology but not with the 7TM model of Oelkers et al. (*21*). However, negative glycosylation data may be inconclusive at best, and only positive results can confirm the topological location of a protein segment.

There are subtle but important differences between the 7TM model contested by Hallén and colleagues (*9*, *21*) and the model proposed in this study that closely follows the model by Wong and co-workers (*1*), primarily at the (putative) membrane entry and exit signal sequences. The data generated by the Hallén group were evaluated and interpreted with respect to the 7TM model of Oelkers et al. (*21*) and their own HMMTOP-derived 9TM model of ASBT/Ntcp. Although we agree that their arguments and results

Table 5: Summary of PROCHECK Quality Assessment Data

structure	hydrogen-bond energy standard deviation	bad contacts per 100 residues	χ^{-1} pooled standard deviation	stereochemical quality indicators: ϕ, ψ distribution	stereochemical quality indicators: χ_1 standard deviation	stereochemical quality indicators: HB energy
ASBT	0.8	0.6	13.0	83.3 (1)	15.8 (2)	2
BACR	0.55	10	28.1	88.3 (1)	26.3 (4)	1
OPSD	0.83	3.3	16.3	80.9 (1)	16.6 (2)	2
sodium channel	0.84	0	23.3	72.2 (2)	25.2 (4)	2



C

	K_t (μM)	J_{max} (pmol·min^{-1}·mg protein^{-1}
wt	11.3 ± 1.9	312.5 ± 12.8
E281A	19.6 ± 4.5	131.6 ± 9.0
E282D	12.3 ± 3.4	164.4 ± 7.5

FIGURE 9: Uptake kinetics of amino acid mutations around putative binding domain 1. (A) Interspecies multiple protein alignment of ASBT, highlighting the region between amino acid residues 260 and 300. Note that hASBT differs from all other species at position 282 (E instead of D). (B) TCA uptake in the presence (open bars) or absence (solid bars) of sodium in COS-1 cells transfected with pCMV5 vector containing the indicated cDNA constructs. Uptake medium was spiked with 5 μM ^{3}H-TCA (0.2Ci/mmol). Results are presented as a percent of transport of the mutants relative to that of the wt (81.3 ± 5.7 pmol min^{-1} mg protein^{-1}). Each bar is the mean ± standard error of three to four different measurements. (C) Kinetic transport parameters.

contest the validity of the 7TM model proposed by Oelkers and colleagues (*21*), these same data provide strong support for the classic 7TM model (*1*) as well as our proposed 7TM model, in many instances more so than a putative 9TM model. Contrary to the claim by Hallén and co-workers (*9*), it is of particular importance to note that the classic 7TM model (Figure 2) strictly adheres to the "positive inside" rule (−4/+4 = out/in charge ratio), which can be taken as additional confirmation of its biological legitimacy.

On the basis of a 7TM topology, we built a homology model for ASBT using BR as a scaffold for positioning the extramembranous domains. The resulting structure was carefully validated by geometrical comparison to comparable proteins (Tables 4 and 5). Of particular note is the strong structural similarity between the N- and C-terminal domains in ASBT as predicted by a remote-threading algorithm. The nonbundled α-helical conformation for these two domains comprises three α helices connected through two distinct β turns (parts A and B of Figure 6). Sun and colleagues recently investigated the structure of the C-terminal domain of rat ASBT. NMR studies on a 20-mer peptide corresponding to cytoplasmic rASBT tail revealed a distinct β-turn conformation spanning residues N340, K341, G342, and F343. As shown in Figure 6C, this β-turn conformation coincides with the predicted conformational turn involving hASBT residues N340, G341, G342, and F343. These studies provide an experimental basis for the existence of the tertiary structure in the C-terminal domain of ASBT. On the basis of threading analysis, we expect that these structural features would translate to the N-terminal domain as well.

Next, we determined potential ligand−protein interaction sites. Two putative binding sites were discovered (parts A and B of Figure 7). The significance of Glu282 in binding site A was further investigated using Ala-scanning mutagenesis (Figures 8 and 9). Our model confirms the following experimental observations: (1) E282A constitutes a loss-of-function mutant, whereas E281A diminishes transporter affinity, likely because of electrostatic interaction with E282; (2) A282D rescues ASBT affinity, yet transporter kinetics are now comparable to hamster ASBT, where an Asp residue naturally occurs in this position (Figure 9).

The homology model and its putative substrate-interaction sites are in good agreement with 3D-QSAR studies from our laboratory (*35*) and a pharmacophore model by Baringhaus and colleagues (*26*). For example, the pharmacophore model indicates that the 3α-hydroxy group in the cis-oriented ring A of the steroid nucleus is not specifically required for transport (*26*). Likewise, our model suggests that either the 3- or 7α-OH moieties of the docked cholic acid ligand may form dynamic hydrogen bonds with hASBT Asn[2] or alternative residues in its vicinity. Furthermore, in accordance with previous models, the 12α-hydroxyl group can form multiple hydrogen bonds with hASBT. Another important interaction point with hASBT of bile acids is a negatively charged side chain, which has been suggested to fulfill a hydrogen-bond acceptor function (*26*). This moiety may interact with positively charged side chains or strong hydrogen-bond donating residues. In our model, residues R254 and R256 in binding domain 2 (EL3) are positioned to potentially fulfill this role. Thus, the advantage of our model over previous pharmacophore and 3D-QSAR models is the possibility of proposing specific interactions of the substrate with amino acid residues in the transport protein. In turn, the model can be utilized to guide the rational design of site-directed mutagenesis experiments.

The validity of our topology vis-à-vis previously suggested 7−9TM configurations is further illustrated by the accessibility of membrane-impermeant methanethiosulfonate (MTS) reagents to Cys residues in ASBT. A fully conserved Cys residue at position 270 (ASBT) and 266 (Ntcp) was the major site of inactivation (although this residue is not essential for bile acid transport function), suggesting that this residue must fall within the extracellular milieu (*32*). As the 7TM model indicates, this residue is highly exposed in EL3 (Figure 2). Mutagenesis of C245, C250, and C260 in Ntcp has minimal effects on thiol-reactive inhibition (*32*). Although ASBT has only a single Cys in this region at position 255, our model indicates that it is near the membrane boundary and would likely have only limited accessibility toward Cys-modifying reagents. Assuming that both Ntcp and ASBT have a mostly similar membrane topology, residues 245 and 250 fall within TM6, whereas the 7TM model of Oelkers et al. would position these residues within the extracellular environment (*32*). Thus, the accessibility of Cys266 and the inaccessibility of cysteine residues 245 and 250 to MTS reagents further illustrate that the 7TM topology model proposed in the present study is corroborated by biochemical data.

In conclusion, we present a model for the three-dimensional structure of ASBT that is consistent with empirical and disease-related data. The localization of potential substrate-binding sites may aid future development of more specific therapeutic strategies against hypercholesterolemia and related cardiovascular diseases.

REFERENCES

1. Wong, M. H., Oelkers, P., Craddock, A. L., and Dawson, P. A. (1994) Expression cloning and characterization of the hamster ileal sodium-dependent bile acid transporter, *J. Biol. Chem. 269*, 1340−1347.
2. Shneider, B. L., Dawson, P. A., Christie, D. M., Hardikar, W., Wong, M. H., and Suchy, F. J. (1995) Cloning and molecular characterization of the ontogeny of a rat ileal sodium-dependent bile acid transporter, *J. Clin. Invest. 95*, 745−754.
3. Wong, M. H., Oelkers, P., and Dawson, P. A. (1995) Identification of a mutation in the ileal sodium-dependent bile acid transporter gene that abolishes transport activity, *J. Biol. Chem. 270*, 27228−27234.
4. Lazaridis, K. N., Pham, L., Tietz, P., Marinelli, R. A., deGroen, P. C., Levine, S., Dawson, P. A., and LaRusso, N. F. (1997) Rat cholangiocytes absorb bile acids at their apical domain via the ileal sodium-dependent bile acid transporter, *J. Clin. Invest. 100*, 2714−2721.
5. Dawson, P. A., and Oelkers, P. (1995) Bile acid transporters, *Curr. Opin. Lipidol. 6*, 109−114.
6. Hagenbuch, B., and Meier, P. J. (1994) Molecular cloning, chromosomal localization, and functional characterization of a human Na^+ bile acid cotransporter, *J. Clin. Invest. 93*, 1326−1331.
7. Hagenbuch, B., Stieger, B., Montserrat, F., Lübbert, H., and Meier, P. J. (1991) Functional expression cloning and characterization of the hepatocyte Na^+/bile acid cotransport system, *Proc. Natl. Acad. Sci. U.S.A. 88*, 10629−10633.
8. Meier, P. J., Jacquemin, E., and Hagenbuch, B. (1992) in *Bile Acids and the Hepatobiliary System. Falk Symposium 68* (Paumgartner, G., Stiehl, A., and Gerok, W., Eds.) pp 95−102, Kluwer Academic Publishers, Dordrecht, The Netherlands.
9. Hallén, S., Branden, M., Dawson, P. A., and Sachs, G. (1999) Membrane insertion scanning of the human ileal sodium/bile acid co-transporter, *Biochemistry 38*, 11379−11388.
10. Turk, E., Kerner, C. J., Lostao, M. P., and Wright, E. M. (1996) Membrane topology of the human Na^+/glucose cotransporter SGLT1, *J. Biol. Chem. 271*, 1925−1934.
11. Covitz, K. M., Amidon, G. L., and Sadee, W. (1998) Membrane topology of the human dipeptide transporter, hPEPT1, determined by epitope insertions, *Biochemistry 37*, 15214−15221.
12. Sun, A. Q., Salkar, R., Sachchidanand, Xu, S., Zheng, L., Zhou, M. M., and Suchy, F. J. (2003) A 14 amino acid sequence with a β-turn structure is required for apical membrane sorting of the rat ileal bile acid transporter, *J. Biol. Chem. 278*, 4000−4009.
13. Hallen, S., Bjorquist, A., Ostlund-Lindqvist, A. M., and Sachs, G. (2002) Identification of a region of the ileal-type sodium/bile acid cotransporter interacting with a competitive bile acid transport inhibitor, *Biochemistry 41*, 14916−14924.
14. von Heijne, G. (1992) Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule, *J. Mol. Biol. 225*, 487−494.
15. Ikeda, M., Arai, M., Lao, D. M., and Shimizu, T. (2002) Transmembrane topology prediction methods: A re-assessment and improvement by a consensus method using a dataset of experimentally characterized transmembrane topologies, *In Silico Biol. 2*, 19−33.
16. Kimura, Y., Vassylyev, D. G., Miyazawa, A., Kidera, A., Matsushima, M., Mitsuoka, K., Murata, K., Hirai, T., and Fujiyoshi, Y. (1997) Surface of bacteriorhodopsin revealed by high-resolution electron crystallography, *Nature 389*, 206−211.
17. Morris, G. M., Goodsell, D. S., Huey, R., and Olson, A. J. (1996) Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4, *J. Comput.-Aided Mol. Des. 10*, 293−304.
18. Jones, D. T., Tress, M., Bryson, K., and Hadley, C. (1999) Successful recognition of protein folds using threading methods biased by sequence similarity and predicted secondary structure, *Proteins 37*, 104−111.
19. Moro, S., Hoffmann, C., and Jacobson, K. A. (1999) Role of the extracellular loops of G protein-coupled receptors in ligand recognition: A molecular modeling study of the human P2Y1 receptor, *Biochemistry 38*, 3498−3507.
20. Lucke, C., Zhang, F., Ruterjans, H., Hamilton, J. A., and Sacchettini, J. C. (1996) Flexibility is a likely determinant of binding specificity in the case of ileal lipid binding protein, *Structure 4*, 785−800.
21. Oelkers, P., Kirby, L. C., Heubi, J. E., and Dawson, P. A. (1997) Primary bile acid malabsorption caused by mutations in the ileal sodium-dependent bile acid transporter gene (SLC10A2), *J. Clin. Invest. 99*, 1880−1887.
22. Nilsson, I. M., and von Heijne, G. (1993) Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane, *J. Biol. Chem. 268*, 5798−5801.
23. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR 8*, 477−486.
24. Panayotova-Heiermann, M., Loo, D. D., Kong, C. T., Lever, J. E., and Wright, E. M. (1996) Sugar binding to Na^+/glucose cotransporters is determined by the carboxyl-terminal half of the protein, *J. Biol. Chem. 271*, 10029−10034.
25. Sanchez, R., and Sali, A. (1997) Evaluation of comparative protein structure modeling by MODELLER-3, *Proteins*, Suppl 1, 50−58.
26. Baringhaus, K. H., Matter, H., Stengelin, S., and Kramer, W. (1999) Substrate specificity of the ileal and the hepatic Na^+/bile acid cotransporters of the rabbit. II. A reliable 3D QSAR pharmacophore model for the ileal Na^+/bile acid cotransporter, *J. Lipid Res. 40*, 2158−2168.
27. Zhang, E. Y., Phelps, M. A., Cheng, C., Ekins, S., and Swaan, P. W. (2002) Modeling of active transport systems, *Adv. Drug Delivery Rev. 54*, 329−354.
28. Kramer, W., Wess, G., Bewersdorf, U., Corsiero, D., Girbig, F., Weyland, C., Stengelin, S., Enhsen, A., Bock, K., Kleine, H., Le Dreau, M. A., and Schafer, H. L. (1997) Topological photoaffinity labeling of the rabbit ileal Na^+/bile-salt-cotransport system, *Eur. J. Biochem. 249*, 456−464.
29. Craddock, A. L., Love, M. W., Daniel, R. W., Kirby, L. C., Walters, H. C., Wong, M. H., and Dawson, P. A. (1998) Expression and transport properties of the human ileal and renal sodium-dependent bile acid transporter, *Am. J. Physiol. 274*, G157−G169.
30. Buchwald, H., Varco, R. L., Matts, J. P., Long, J. M., Fitch, L. L., Campbell, G. S., Pearce, M. B., Yellin, A. E., Edmiston, W. A., Smink, R. D., Jr., et al. (1990) Effect of partial ileal bypass surgery on mortality and morbidity from coronary heart disease in patients with hypercholesterolemia. Report of the Program on the Surgical Control of the Hyperlipidemias (POSCH), *N. Engl. J. Med. 323*, 946−955.

31. Takashima, K., Kohno, T., Mori, T., Ohtani, A., Hirakoso, K., and Takeyama, S. (1994) The hypocholesterolemic action of TA-7552 and its effects on cholesterol metabolism in the rat, *Atherosclerosis 107*, 247−257.
32. Hallén, S., Fryklund, J., and Sachs, G. (2000) Inhibition of the human sodium/bile acid cotransporters by side-specific methanethiosulfonate sulfhydryl reagents: Substrate-controlled accessibility of site of inactivation, *Biochemistry 39*, 6743−6750.
33. Slotboom, D. J., Sobczak, I., Konings, W. N., and Lolkema, J. S. (1999) A conserved serine-rich stretch in the glutamate transporter family forms a substrate-sensitive reentrant loop, *Proc. Natl. Acad. Sci. U.S.A. 96*, 14282−14287.
34. Hallen, S., Mareninova, O., Branden, M., and Sachs, G. (2002) Organization of the membrane domain of the human liver sodium/bile acid cotransporter, *Biochemistry 41*, 7253−7266.
35. Swaan, P. W., Szoka, F. C., Jr., and Oie, S. (1997) Molecular modeling of the intestinal bile acid carrier: A comparative molecular field analysis study, *J. Comput.-Aided Mol. Des. 11*, 581−588.
36. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., and Miyano, M. (2000) Crystal structure of rhodopsin: A G protein-coupled receptor, *Science 289*, 739−745.
37. Doak, D. G., Mulvey, D., Kawaguchi, K., Villalain, J., and Campbell, I. D. (1996) Structural studies of synthetic peptides dissected from the voltage-gated sodium channel, *J. Mol. Biol. 258*, 672−687.
38. Ren, G., Reddy, V. S., Cheng, A., Melnyk, P., and Mitra, A. K. (2001) Visualization of a water-selective pore by electron crystallography in vitreous ice, *Proc. Natl. Acad. Sci. U.S.A. 98*, 1398−1403.
39. Doyle, D. A., Morais Cabral, J., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T., and MacKinnon, R. (1998) The structure of the potassium channel: Molecular basis of K^+ conduction and selectivity, *Science 280*, 69−77.
40. Zuniga, F. A., Shi, G., Haller, J. F., Rubashkin, A., Flynn, D. R., Iserovich, P., and Fischbarg, J. (2001) A three-dimensional model of the human facilitative glucose transporter Glut1, *J. Biol. Chem. 276*, 44970−44975.
41. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures, *J. Mol. Graphics 14*, 51−55, 29−32.
42. Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H. R., and Iwata, S. (2003) Structure and mechanism of the lactose permease of *Escherichia coli*, *Science 301*, 610−615.
43. Huang, Y., Lemieux, M. J., Song, J., Auer, M., and Wang, D. N. (2003) Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*, *Science 301*, 616−620.

BI049270A