# Topography of the Membrane Domain of the Liver $Na^+$-Dependent Bile Acid Transporter†

Olga Mareninova, Jai Moo Shin, Olga Vagin, Shahlo Turdikulova, Stefan Hallen,‡ and George Sachs*

*Departments of Physiology and Medicine, David Geffen School of Medicine, University of California at Los Angeles, and VA Greater Los Angeles Healthcare System, Los Angeles, California 90073*

*Received July 5, 2005; Revised Manuscript Received August 23, 2005*

ABSTRACT: The ileal apical and liver basolateral bile acid transporters catalyze the $Na^+$-dependent uptake of these amphipathic molecules in the intestine and liver. They contain nine predicted helical hydrophobic sequences (H1−H9) between the exoplasmic N-glycosylated N terminus and the cytoplasmic C terminus. Previous *in vitro* translation and *in vivo* alanine insertion scanning studies gave evidence for either nine or seven transmembrane segments, with H3 and H8 noninserted in the latter model. N-terminal GFP constructs containing either successive predicted segments or only the last two domains of the liver transporter following a membrane anchor signal were expressed in HEK-293 cells, and a C-terminal glycosylation flag allowed detection of membrane insertion. Western blot analysis with anti-GFP antibody after alkali and PNGase treatment showed that H1, H2, H3 behaved as competent transmembrane (TM) sequences. Results from longer constructs were difficult to interpret. H9, however, but not H8 was membrane-inserted. To analyze the intact transporter, a C-terminal YFP fusion protein was expressed as a functionally active protein in the plasma membrane of HEK-293 cells as seen by confocal microscopy. After limited tryptic digestion to ensure the accessibility of only exoplasmic lysine or arginine residues, molecular weight (MW) analysis of the five cleavage products on SDS−PAGE predicted the presence of seven transmembrane segments, H1, H2, H3, H4, H5, H6, and H9, with H7 and H8 exoplasmic. This new method provided evidence for seven membrane segments giving a new model of the membrane domain of this protein and probably the homologous ileal transporter, with H7/H8 as the transport region.

During the enterohepatic circulation of the large, amphipathic bile acids, uptake occurs from the intestinal lumen via the apical ileal sodium-dependent bile acid cotransporter and is delivered via the portal vein to the liver. The bile salts then enter hepatocytes via the basolateral sodium-dependent cotransporter (SBAT)[1] and then are secreted into the bile. The ileal bile acid transporter (IBAT) is a target for drugs designed to inhibit bile acid uptake and thence lowering of cholesterol by stimulation of 7-α hydroxylase activity (*1*, *2*). The ileal and hepatic $Na^+$/bile acid transporters are polytopic integral membrane proteins, which share 35% identity and very similar hydrophobicity profiles. Their secondary structure is therefore likely to be very similar.

A definition of the two-dimensional structure of these transporters is important for understanding of the mechanism of transport of these bile salts. The membrane topography of this family of transporters that contains nine hydrophobic segments (Figure 1) has been difficult to establish because different algorithms and different experimental approaches have provided conflicting data. The N-terminal glycosylation and the antibody-determined cytoplasmic location of the C terminus exclude a model with an even number of membrane segments (*3*, *4*). Membrane insertion algorithms and experimental data have predicted either a seven transmembrane sequence with two long extra-cytoplasmic loops containing the 3rd and 8th hydrophobic sequences in the seven segment model (Figure 1A) or a nine transmembrane segment model (Figure 1B). There is experimental support for each of these, depending upon the techniques used (*4*−*6*).

The topogenic properties of individual hydrophobic segments of the bile acid transporters were analyzed in vectors containing an N-terminal cytoplasmic anchor and a C-terminal glycosylation flag using an *in vitro* transcription/translation system with dog pancreatic microsomes (*5*, *7*). These vectors have been used to determine the topography of several proteins, such as $P_1$- and $P_2$-type ATPases, G7 receptors, and Na/H and $NaHCO_3$ transporters (*8*−*12*). Depending upon the protein, this approach often provided an accurate secondary structure such as for a $P_1$ ATPase of *Helicobacter pylori*, the CCK-A receptor, and the $Na/HCO_3$ cotransporter (*9*, *10*, *12*) or gave results contrary to the known arrangement of transmembrane segments such as for the $P_2$-type ATPases, the gastric H,K-ATPase, and the ER Ca-ATPase (*8*, *13*). The method can also give misleading data indicating membrane insertion of a hydrophobic sequence such as an α helix in cyclooxygenase that is known from the crystal structure not to be membrane-inserted (*8*).

In this *in vitro* translation system, seven of the nine hydrophobic sequences (H1, H2, H4, H5, H6, H7, and H9

† Supported in part by U.S. Veterans Administration and NIH Grants DK46917, 53462, and 58333.

* To whom correspondence should be addressed. E-mail: gsachs@ucla.edu. Telephone: 310-268-3923. Fax: 310-312-9478.

‡ Present address: AstraZeneca R&D Mölndal, S-431 83 Mölndal, Sweden.

[1] Abbreviations: GFP, green fluorescent protein; YFP, yellow fluorescent protein; IBAT, ileal bile acid transporter; SBAT, sodium/bile acid transporter; SBAT−YFP, a fusion protein between SBAT and YFP linked to the C terminus of SBAT.

10.1021/bi051291x CCC: $30.25 
Published on Web 09/30/2005

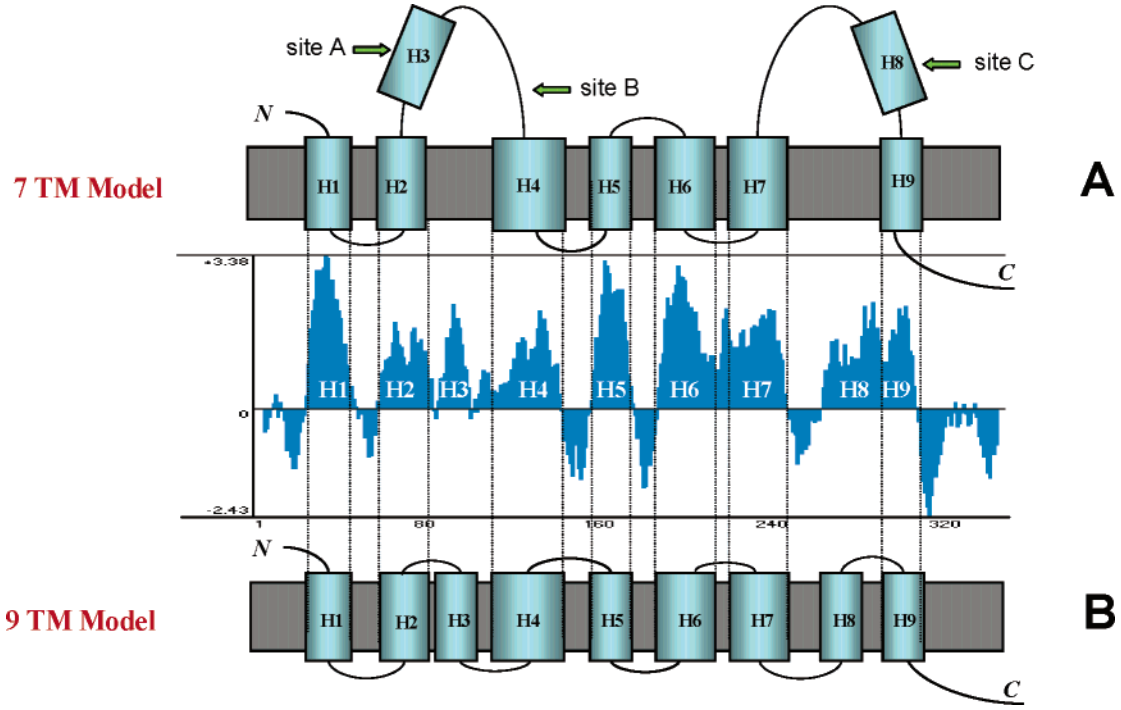


FIGURE 1: Seven (A) and nine (B) transmembrane segment models of the liver basolateral bile salt transporter. The hydrophobicity profile of the transporter was calculated by the Kyte−Doolittle algorithm using an 11 amino acid moving average and shown between models A and B. In the top model, arrows point to regions of engineered glycosylation sites that were introduced by mutagenesis into the IBAT and were apparently glycosylated (*6*), which was interpreted as substantiating the classical seven transmembrane segment arrangement even though these mutations inactivated the protein (A). The lower model (B) illustrates the nine transmembrane segment model for which some evidence had been obtained (*5*) but is now refuted by the data presented in this paper.

but not H3 and H8) of the bile acid transporter were able to act individually as either signal anchor or stop transfer sequences (*4*, *5*), suggesting seven transmembrane-inserted segments. Recently, results of insertion of N-linked glycosylation sites into H3, H3−4, and H8 segments of the cell expressed, N-terminal deglycosylated intestinal bile salt transporter have also been interpreted as evidence for the seven transmembrane segment model shown in Figure 1A similar to the *in vitro* translation data (*6*). One drawback to this model is the exoplasmic orientation of the loop between H5 and H6 that contains several uncompensated positive charges that would be predicted to be cytoplasmic according to the positive inside rule (*14*, *15*). Further, these important glycosylation site insertion mutants were not transport-competent (*6*).

However, when H3 and H8 sequences were expressed in context with the neighboring native sequences in this *in vitro* translation system, they were able to insert into microsomal membranes and the H3 and H8 acted as a stop transfer and signal anchor signals, respectively. These data led to a postulated nine transmembrane segment model as shown in Figure 1B (*4*, *5*).

Additional data, interpreted as supporting a nine transmembrane segment model, were obtained by using alanine insertion scanning (*5*). This method is based on the hypothesis that the addition of an alanine residue in the middle of a transmembrane α helix will disrupt important structural associations with neighboring membrane segments by changing the position of the adjacent α helices (*16*). Thereby, the function of a membrane transporter is expected to be inhibited. Transport activity and plasma-membrane trafficking of the transporter expressed in HEK-293 cells was determined in the mutants with alanine insertions into the different hydrophobic regions of the protein (*5*). Alanine insertion into H1, H2, H4, H5, H7, and H9 resulted in a loss of transport activity as predicted from the translation of these individual sequences. All alanine insertions in H3 and H8 also abolished taurocholate uptake, suggesting that both of these regions have structures with critical intramolecular interactions. Moreover, these insertions also prevented trafficking to the plasma membrane. These data were interpreted as compatible with the putative nine transmembrane model as shown in Figure 1B but now with the positively charged loop between H5 and H6 oriented on the cytoplasmic surface as predicted from the positive inside rule (*14*, *15*).

To try to resolve the contradictory data obtained with different techniques used thus far, we first adapted an *in vivo* translation method previously applied to aquaporin (*17*). GFP-linked constructs were expressed in HEK-293 cells and analyzed by Western blotting for the presence or absence of glycosylation of the C-terminal flag. GFP preceded the first membrane inserted sequence of the H,K-ATPase α sequence followed by successive hydrophobic segments of the liver bile acid transporter and then the glycosylation reporter sequence of the H,K-ATPase $\beta$ subunit. The results for the constructs truncated after H1, H2, and H3 respectively indicated that all three are membrane-integrated segments. A fusion protein containing either only H8 and or H8/9H9 showed that H9 but not H8 is membrane-inserted. Data for longer truncated constructs beginning with H1−H4 were difficult to interpret.

To avoid the problems encountered with the expression of truncated protein or inactive mutants, we expressed a full-length C-terminal YFP fusion protein in HEK-293 cells. The YFP-linked bile acid transporter protein was located on the plasma membrane, as easily seen by fluorescent confocal

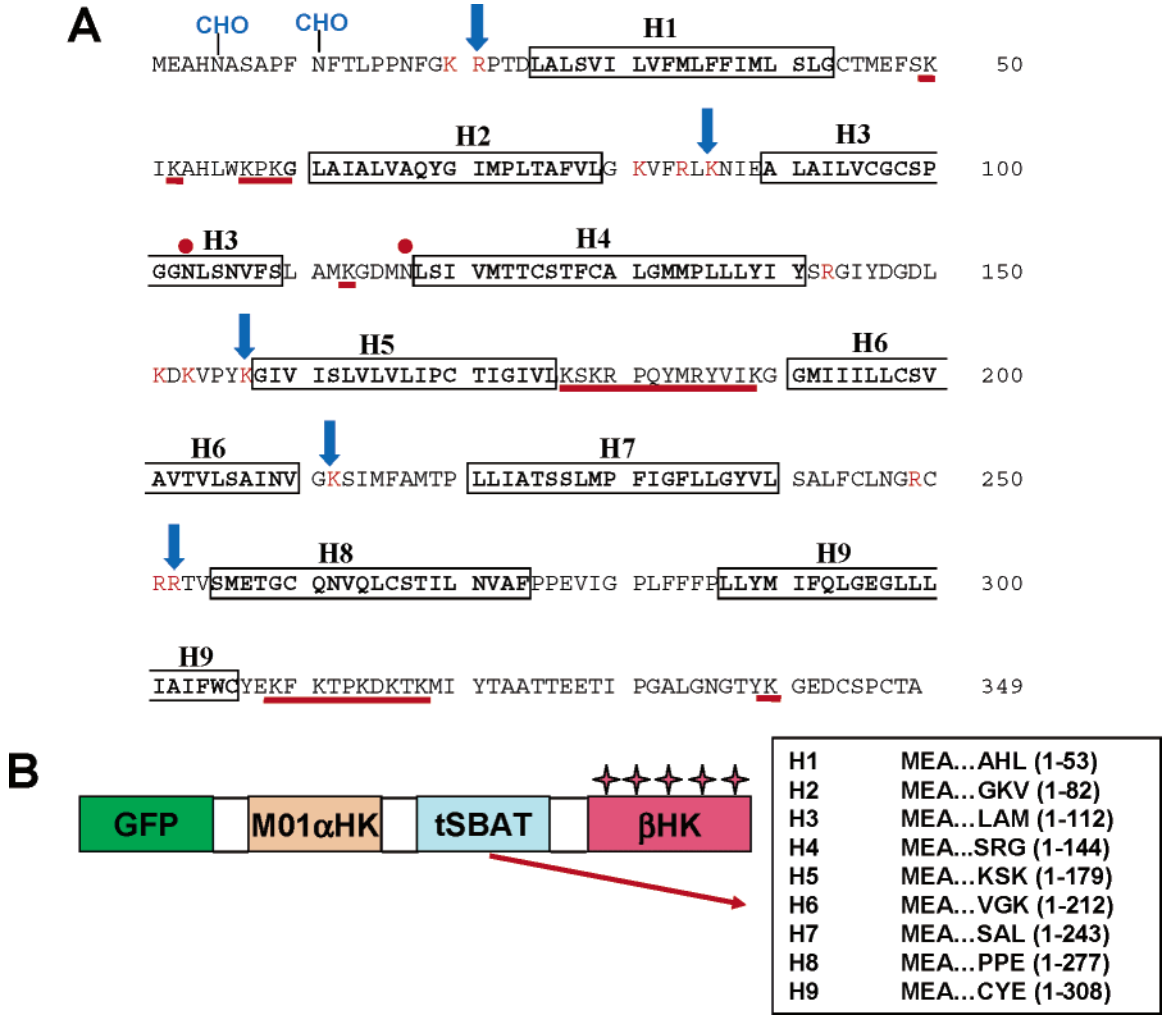


FIGURE 2: (A) Amino acid sequence of the human liver bile acid transporter showing the regions of hydrophobic helical predictions in bold in the rectangles. The observed extracellular sites of tryptic cleavage of the intact protein are indicated by blue arrows, and the corresponding lysine and arginine residues are shown in red. The nonaccessed likely internal possible tryptic cleavage sites are underlined red. The natural glycosylation sites that are not glycosylated in the liver transporter are shown by red circles. CHO, N-linked carbohydrates. (B) Schematic representation of the construction of various fusion vectors. The putative hydrophobic segments (from H1 to H9) were predicted by the HMMTOP algorithm. The sequences used for topology scanning of the N-terminal GFP constructs are shown at the bottom right of the figure in the table. tSBAT, truncated SBAT sequence.

microscopy, and was able to transport bile salts at the same level as wild-type protein (*5*, *18*). These intact cells were then subjected to limited tryptic digestion to define exoplasmic-accessible lysines or arginines. After separation of the fragments on SDS−PAGE and accurate analysis of the relative MW of these fragments on Western blots by using the anti-GFP antibody that also recognized YFP, five luminally exposed tryptic cleavage sites were identified. This analysis yielded a specific model of the plasma-membrane-inserted basolateral liver bile salt transporter, where the first six hydrophobic segments are membrane-inserted, H7 and H8 are exoplasmic, and H9 is membrane-inserted. Hence, according to this method of analysis of the intact, functional transporter, the liver bile acid transporter contains seven transmembrane segments. However, contrary to previous predictions, it is the 7th and 8th hydrophobic sequences that are not membrane-integrated. The sequence and hydropathy similarity of the apical ileal bile salt transporter suggests that it would have similar membrane topography. The 7th and 8th hydrophobic sequences are also in the region of binding of the bile salts in the ileal variant and also the region of binding of ileal transport inhibitors (*5*, *18*). This large region may therefore provide the flexibility necessary for transport of these large amphipathic molecules.

## MATERIALS AND METHODS

*Vector and Plasmid Construction.* The mammalian expression vector, pEGFP-HK M0/M1-$\beta$, was constructed by ligating the HK M0/M1-$\beta$ construct (*13*) into pEGFP-C3 (Clontech) between the *Xho*I and *Eco*RI sites. Sequences coding for human SBAT successive truncation mutants were cloned into the *Bgl*II/*Hin*dIII linker region of pEGFP-HK M0/M1-$\beta$ between the first 139 amino acids from the N terminus of the rabbit H,K-ATPase α-subunit (M0/M1) and 177 C-terminal amino acids of the rabbit H,K-ATPase $\beta$-subunit containing five N-linked glycosylation sites ($\beta$ flag). Each SBAT truncated sequence (H1, H1−H2, and H1−H3) began at the SBAT start codon and ended at the end of each of the putative membrane-spanning regions of SBAT, which were identified using programs based on a hidden Markov model, TMHMM (*19*), HMMTOP (*20*), as shown at the bottom right panel in Figure 2. A second set of GFP-linked sequences was constructed containing only H8 and H8 and H9.

The SBAT-truncated sequences were constructed by PCR using the pcDNA3.1(+) vector containing cDNA for the human liver basolateral bile acid transporter as a template (*5*). PCR primers were designed using the Primer Select software. Sense primers for all PCR reactions were 5′-flanked with a *Bgl*II site, and all antisense primers contained the *Hin*dIII site at the 3′ end. The PCR products were purified using Resin/Wizard minicolumns digested with the restriction enzymes *Bgl*II and *Hin*dIII and then gel-purified (1.2% agarose gel) using the Min Elute Gel Extraction Kit. The inserts were ligated into the pEGFP-HK M0/M1-$\beta$ vector. All plasmid constructs were identified by restriction digestion analysis and verified by sequencing.

*Expression of the Constructs in HEK-293 Cells.* The cell line was purchased from ATCC. The HEK-293 cells were grown to 70−80% confluence in 6-well plates in Dulbecco's modified essential medium (high glucose, supplemented with 10% fetal bovine serum, 100 units/L penicillin, 0.1 mg of streptomycin/L, 2 mM L-glutamine, 0.1 mM nonessential amino acids, and 1.0 mM sodium pyruvate). They were transiently transfected with 1 $\mu$g pf DNA/well using the FuGENE6 transfection reagent (Roche) according to the instructions of the manufacturer.

*Membrane Alkali Extraction.* This was performed to ensure that the segments analyzed were truly membrane-integrated. The membrane fraction was isolated according to previously described methods (*17*), with minor modifications, using an alkaline flotation step to identify truly membrane-integrated fragments and to remove spurious membrane-associated but not integrated peptides. Briefly, transiently transfected HEK-293 cells were washed in phosphate-buffered saline and then homogenized in TEEA buffer (pH 8.0) containing 20 mM Tris-HCl, 1 mM EDTA, 1 mM EGTA, 1 mM phenylmethylsulfonyl fluoride (PMSF), and protease inhibitor cocktail, followed by centrifugation (10 min at 1000*g*). The supernatants were centrifuged at 100000*g* for 30 min, and the resulting particulate fraction was extracted with 100 mM $Na_2CO_3$ (pH 12.0) on ice for 30 min. A total of 1.25 and 0.25 M sucrose were overlaid on the alkaline mixture followed by centrifugation at 100000*g* for 60 min in a Beckman ultracentrifuge. The fraction collected at the 0.25 and 1.25 M sucrose layers and the interface between the 1.25 M sucrose layer and the alkaline mixture was recovered as the membrane fraction and stored at −70 °C prior to SDS−PAGE.

*SDS−PAGE and Western Blot Analysis.* Samples were mixed with 2× sample buffer containing 0.1 M Tris-HCl at pH 6.8, 4% SDS, 20% glycerol, and 5% $\beta$-mercaptoethanol with bromophenol blue as a marker. The samples (5−10 $\mu$g of protein/lane) were separated on 10% Tris/glycine SDS gels and then transferred to PVDF membranes. The recombinant proteins were detected by incubation with a mouse anti-GFP antibody. Alkaline phosphatase (AP)-conjugated goat anti-mouse was used as a secondary antibody. The immunoblots were developed with AP buffer (100 mM Tris, 100 mM NaCl, and 5 mM $MgCl_2$ at pH 9.5) containing 0.3% nitro-blue-tetrazolium solution (w/v) and 0.15% 5-bromo-3-indolyl-1-phosphate solution (w/v) according to the instructions of the manufacturer. The anti-GFP antibody allows clear identification of glycosylated subunits in contrast to the anti-transporter antibodies available, which are much less specific, enabling clear conclusions as to the presence or absence of glycosylation of the $\beta$ subunit N-glycosylation sites.

*PNGase F Treatment To Determine Glycosylation Status.* A 10 $\mu$L aliquot of the membrane fraction was diluted to 20 $\mu$L in 50 mM Tris (pH 7.8), 1% $\beta$-mercaptoethanol, and 1% Nonidet P-40 (final concentrations) and incubated after adding 1 $\mu$L (1000 units) of PNGase F for 1 h at 37 °C. A control sample was run without the enzyme. This proved the presence of glycosylation in the constructs.

*Construction of HEK-293 Stable Cell Lines Expressing SBAT−YFP Fusion Protein.* cDNA encoding the fusion protein between the basolateral SBAT and YFP was constructed by inserting a cDNA for SBAT into pEYFP-N1 vector (Clontech, Palo Alto, CA) so that the YFP would be located at the C terminus of the human basolateral SBAT as described before (*5*). HEK-293 cells were grown on 10 cm plates until 20% confluent and transfected with the SBAT−YFP vector using FuGENE 6 Transfection Reagent (Roche). Stable cell lines expressing SBAT−YFP fusion proteins, were selected by adding the eukaryotic selection marker G-418 at a final concentration of 1.0 mg/mL at 24 h after transfection. This concentration of G-418 was maintained until single colonies appeared. A total of 15−20 colonies were isolated, expanded, and grown in the presence of 0.25 mg of G-418/mL of media in 24-well plate. Two clones, with the highest expression of SBAT−YFP, were selected and expanded for further studies. The cells were shown to be functional using a proximity assay for $Na^+$-dependent $^{14}C$ taurocholate uptake in confluent monolayers (*5*, *21*).

*Confocal Microscopy Studies.* Cells stably expressing SBAT−YFP were grown until becoming confluent on glass-bottom microwell dishes (MatTek Corporation). Confocal microscopic images were acquired using the Zeiss LSM 510 laser-scanning confocal microscope using LSM 510 software, version 3.2.

*Tryptic Digestion of SBAT−YFP Protein and Western Blot Analysis of the Digestion Fragments.* HEK-293 cells expressing SBAT−YFP were grown in T-25 flasks until becoming confluent, rinsed twice with PBS at pH 7.2 (GIBCO, Invitrogen), and digested by incubating cells with 0.5 mL of HBSS buffer containing 0.05% trypsin and 5 mM EDTA−4Na solution in (GIBCO, Invitrogen) at 37 °C for 10 or 60 min in a $CO_2$ incubator. This low concentration of trypsin was used to avoid overdigestion and penetration of the protease to the inside surface of the plasma-membrane-expressed transporter. Control cells were incubated for 60 min in the same buffer but without trypsin. After incubation, 14.5 mL of DMEM media (Mediatech, Cellgro) containing soybean trypsin inhibitor (1 mg/mL) was added to stop the trypsin digestion. We used low concentrations of trypsin (0.05%) and relatively short incubation times and subsequently high concentrations of trypsin inhibitor to ensure that only surface cleavage occurred and that trypsin did not penetrate the membrane. To stop digestion and prevent cleavage during and following extraction and the analytical procedures, a 30× volume of the cell culture media containing fetal bovine serum and soybean trypsin inhibitor were added at the end of incubation. After inhibition of tryptic activity, cells were separated from the supernatant by a short, low-speed centrifugation to remove cells possibly damaged during incubation with trypsin. The cells were collected and spun down at 1000*g* for 1 min. The cell pellet was washed

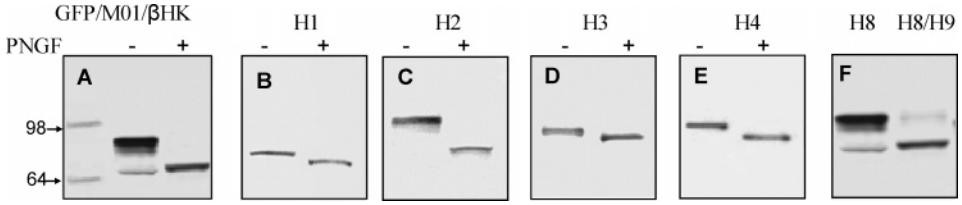


FIGURE 3: Analysis of the membrane topology of SBAT by *in vivo* translation of various truncated constructs of a N-terminal GFP fusion protein followed by the first 139 amino acids of the H,K-ATPase that contain, in addition to its cytoplasmic anchor sequences, its first transmembrane segment (M0M1) then sequential hydrophobic segments of the basolateral liver bile salt transporter terminating in a partial 177 amino acid sequence of the H,K-ATPase $\beta$ subunit containing 5 N-linked glycosylation flag sites. The orientation of the membrane segments of the transporter are in their natural direction. The different constructs were expressed in HEK-293 cells, and the crude membrane fraction, after alkaline extraction, was separated by SDS−PAGE. All recombinant proteins were treated without (−) or with (+) PNGase F and detected on Western blots using an anti-GFP antibody. A is the control without the addition of segments of the transporter, and then, B contains the addition of H1. C shows the effect of the presence of H1 and H2 in the GFP construct, and D shows the effect of the addition of H3 to the construct. In E, the effect of the addition of H4 is shown, and F shows expression of a construct containing only H8 or H8 and H9. The data are representative of three experiments. The similarity of the left band of F and the first lane of A shows that H8 is not inserted but that H9 is inserted into the membrane in this translation system.

twice with PBS buffer. Cells were lysed by incubation with 500 $\mu$L of 0.15 M NaCl in 15 mM Tris at pH 8.0 with 1% Triton X-100 and 4 mM EGTA. Cell lysates were clarified by centrifugation (15000*g* for 5 min). Samples containing 20 $\mu$L of supernatant mixed with 15 $\mu$L of SDS-containing sample buffer were loaded onto SDS−PAGE 10−20% gradient gels and analyzed by Western blotting similar to the procedure described above, with SDS−PAGE using monoclonal anti-GFP antibody. The monoclonal anti-GFP antibody was obtained from Boehringer Mannheim and reacts equally well with YFP. The secondary antibody was a goat anti-mouse HRP-conjugated antibody. Detection was with the ECL chemiluminescence reaction kit. The calibration of the relative MW was performed in each experiment using Sigma MW standards, and the image reflects at least four separate experiments.

*Uptake of Na-taurocholate.* This was done as previously described in detail by measuring the uptake of $^{14}$C-taurocholate in the absence and presence of $Na^+$ using a proximity assay for determination of the radioactivity transiting a confluent monolayer of HEK-293 cells (*5*).

*Materials.* All materials were analytical-grade or higher. All enzymes were from New England Biolabs. pEGFP-C3 plasmid was from Clontech, Palo Alto, CA. The PCR products were purified using Resin/Wizard minicolumns from Promega, Madison, WI, and Mini Elute Gel Extraction Kit from QIAGEN, Chatsworth, CA. HEK-293 cells were grown in medium from Life Technologies. The FuGENE6 transfection reagent, phenylmethylsulfonyl fluoride (PMSF), and protease inhibitor cocktail were obtained from Roche Molecular Biochemicals. Proteins were separated on tricine gradient gels and transferred to PVDF membranes from Millipore (Immobilon-P, 0.22 $\mu$M). The ECL chemiluminescence reaction kit was obtained from Pierce.

## RESULTS

Two suggested 2D models of the bile acid transporters are shown in parts A and B of Figure 1 and were derived from a Kyte−Doolitle hydrophobicity plot (*22*), using an 11 amino acid moving average. The putative transmembrane segments of the nine transmembrane segment model predicted by a new HMMTOP algorithm are labeled as H1, H2, etc. The bile acid transporters belong to the membrane proteins whose glycosylated N terminus is targeted to the ER lumen, and therefore, the first transmembrane sequence acts as a "type III" reverse signal anchor sequence and is inserted into the ER membrane in an $N_{lum}/C_{cyt}$ orientation. To keep the native orientation of the SBAT segments in the constructs, we inserted SBAT sequences, truncated after H1, H2, etc. after the cytoplasmic N-terminal anchor region (M0) and the first transmembrane segment (M1) of the rabbit H,K-ATPase $\alpha$ subunit. The reporter tag, the C-terminal region of the $\beta$-subunit H,K-ATPase containing five N-glycosylation sites, was at the end of the expression vector (Figure 2, bottom). To provide an epitope for detection with Western blot analysis by using the antibody against GFP, we began each construct with the EGFP sequence. The topological data were inferred from the location of the reporter tag relative to the luminal side of the ER. Glycosylation of $\beta$ flag will result in an increase in the apparent molecular mass of approximately 14 kDa on SDS−PAGE electrophoresis. The use of a monoclonal anti-GFP antibody provided very clear Western analysis patterns as to the presence or absence of glycosylation of the C-terminal glycosylation flag in contrast to the less specific polyclonal antibody generated against the transporter (*5*).

The C-terminal truncated sequences were inserted into the GFP vector. The exact sequence of each construct is shown in the table on the lower right of Figure 2 along with the vector sequence for this set of experiments in the intact cells.

The results obtained with various constructs show that all recombinant proteins were expressed as being glycosylated because of the two N-terminal sites of SBAT, which have an extra-cytoplasmic location (Figure 2). The third available site on SBAT was not glycosylated in any construct tested. When the $\beta$ flag was translocated into the ER lumen, these recombinant proteins had additional bands on Western blot because of the additional glycosylation of the reporter sequence. This is illustrated by expression of the M0M1 vector shown in Figure 3A, where the glycosylation of the $\beta$ flag resulted in a large upward shift of MW reversed by PNGase treatment. As seen below, PNGase treatment removed all glycosylation from any of these translated products. Alkali extraction ensured that the expressed constructs were truly membrane-integrated.

*H1.* Expression of the protein truncated after the H1 sequence produced a protein glycosylated only on the two N-terminal sites of SBAT present in this construct. Thus,

glycosylation of the flag reporter sequence was prevented by the membrane insertion of H1 in $N_{exo}/C_{cyt}$ orientation, indicating that the first fragment of SBAT is a strong reverse signal anchor signal. PNGase F treatment resulted in a MW shift of the product because of deglycosylation of two N-terminal sites (Figure 3B).

*H1−H2.* Western analysis of the H1−H2 recombinant protein detected a band with a relative MW corresponding to a protein glycosylated at the two N-terminal sites of SBAT plus glycosylation of the five N-linked sites on the $\beta$-H,K-ATPase flag (~95 kDa) (Figure 3C). Treatment with PNGase F shifted the band down to the size of the nonglycosylated protein, confirming the flag glycosylation of this construct. Therefore, in this GFP N-terminus fragment, the first two hydrophobic sequences are membrane-integrated in HEK-293 cells.

*H1−H3.* When the downstream H3 sequence was included in the expression vector, the C-terminal reporter was now no longer glycosylated, indicating that the H3 segment is able to integrate into the membrane and act as stop transfer signal again shown in Figure 3D. The two N-terminal sites were still glycosylated as shown by the PNGase F treatment. Although there is a natural N-glycosylation site at Asn103, there was no evidence that this site was utilized and the increase in MW was accounted for only by the 30 amino acid increase in length of the protein. Thus, the three first predicted segments, H1, H2, and H3, act as transmembrane segments in this *in vivo* expression system, and the natural N-glycosylation sites in the liver transporter are not utilized. Previous experiments translating individual segments using canine ER microsomes had shown that the H3 sequence was membrane-inserted only when expressed in context with the upstream native sequence (*5*).

*H1−H4.* With the addition of the H4 region after H3, no change was observed in the glycosylation pattern obtained with the H1−H3 GFP construct (Figure 3E). Elongation of the H4 inserted until Pro155 (not shown) did not change the glycosylation pattern. Hence, H4 does not appear to act as a transmembrane sequence in this particular *in vivo* translation system because it was retained on the cytoplasmic side of the translated product without $\beta$-flag glycosylation. H4 had behaved as a weak transmembrane segment, when previously tested for signal anchor/stop transfer properties in ER microsomes (*4*, *5*).

Subsequent hydrophobic sequence insertions showed that either H5 or H6 but not both were able to membrane-insert. The three longer truncated constructs, H1−H7, H1−H8, and H1−H9, did not show C-terminal glycosylation of the $\beta$ flag. Because the location of H5 and H6 was uncertain, the data on these longer constructs could not be interpreted; therefore, these data are not shown in Figure 3.

*H8 and H8/H9.* When a construct was made containing just H8 alone, the major fraction showed glycosylation of the $\beta$ flag (Figure 3A), suggesting that the H8 segment does not insert into the membrane (left lane in Figure 3F). The construct containing the H8/H9 region following M0/M1 of the H,K ATPase now did not show glycosylation of the $\beta$ flag (right lane in Figure 3F). These data show that H8 does not have membrane-insertion properties but that H9 acts as a competent stop transfer sequence as shown before in the *in vitro* translation with dog pancreatic microsomes (*4*, *5*).

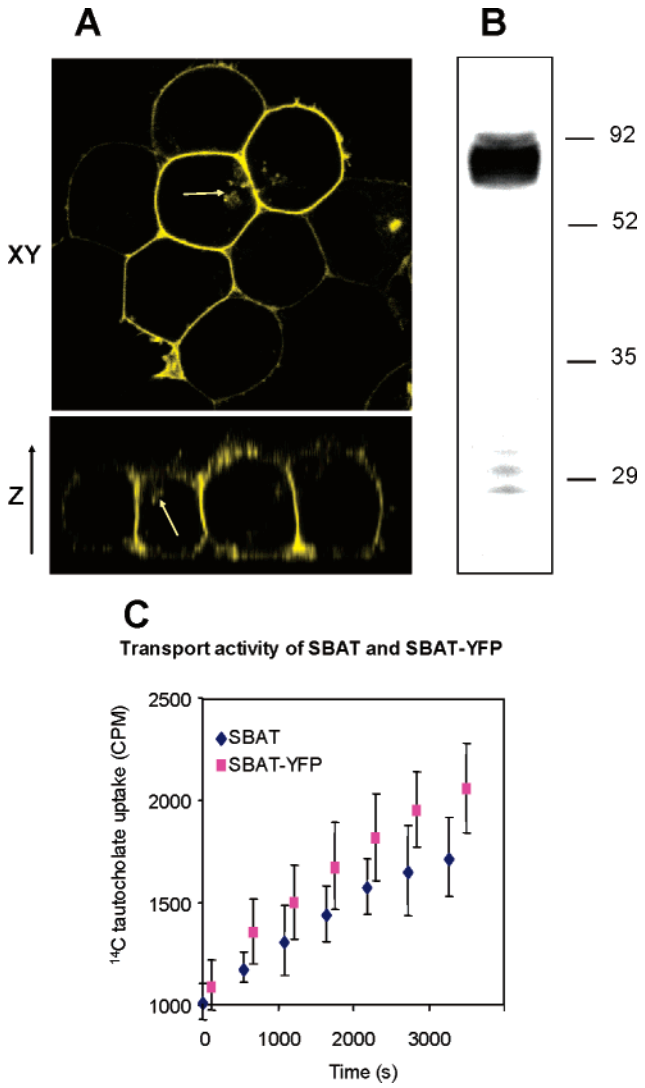


FIGURE 4: (A) Confocal image of HEK-293 cells expressing the C-terminal YFP fusion protein of the liver bile acid transporter in the *x*−*y* and *z* sections showing almost exclusive expression in the plasma membrane but with a minor cytoplasmic component in the two upper heavily stained cells as shown by arrows. A similar cytoplasmic component is shown in the *z* axis particularly in the cell on the left. Most of the protein appears to be expressed in the contact region between the cells, but this is probably due to the high degree of infolding of this region of the membrane. (B) Anti-YFP Western blot using anti-GFP antibody and alkaline phosphatase staining of a solubilized extract of the cells expressing the YFP fusion protein with a large intense band corresponding to the expected MW of the fully glycosylated transporter plus YFP. Some minor fragments are seen in the region of 29 kDa, presumably degradation products of the full-length protein. (C) $^{14}$C-Taurocholate uptake into HEK-293 cells expressing native transporter and the C-terminal YFP fusion protein showing identical activity.

*Stable Expression of the SBAT−YFP Fusion Protein.* Given the results of *in vivo* translation of the N-terminal GFP-truncated constructs, we decided to express the intact C-terminal YFP fusion protein in HEK-293 cells and to digest this surface-expressed protein under mild conditions. The high specificity of the anti-GFP antibody was used to define the exact positions and hence the accurate relative molecular weights of the exoplasmic tryptic cleavage fragments and therefore cleavage sites.

Confocal microscopy (Figure 4A) of the HEK-392 cells expressing this fusion protein, both in the *x*−*y* and *z* sections,

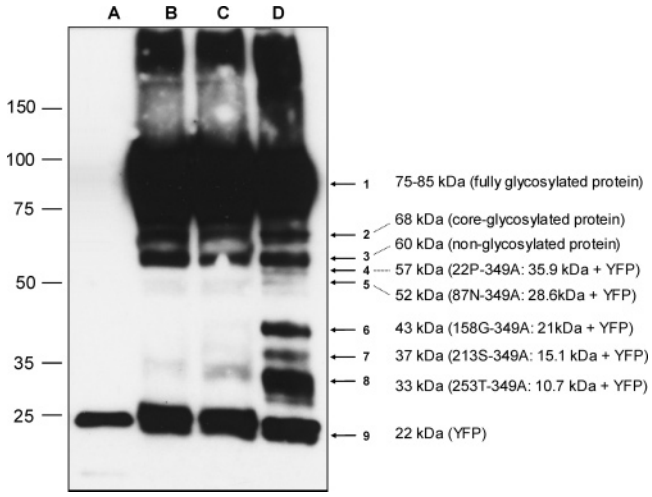


FIGURE 5: Typical ($n = 3$) 10−20% SDS gel separation of the intact and 0.05% tryptic-digested membrane fraction of the HEK-293 cells expressing the YFP fusion protein using Western blot analysis to detect YFP. Lane A shows expression of the YFP protein alone; Lane B shows the blot of the undigested cells; and lanes C and D show the effects of tryptic digestion for 10 and 60 min. The images are typical of at least four separate digestions. Arrow 1, the intact fully glycosylated; 2, core-glycosylated; 3, nonglycosylated fusion protein; 4−8, progressively shorter fragments of the digested protein that still contain the C-terminal YFP tag; and 9, YFP-containing degradation product present in both undigested and digested samples. MW are shown, and the MW of the sequences of the bile salt transporter are calculated after subtraction of the MW of YFP. The fragments increase in intensity with longer trypsin digestion and are absent in the untreated cells. The 9th arrow shows YFP protein expressed at the same level as in untreated cells.

showed that the majority of the protein is expressed in the plasma membrane particularly in the region of contact between the cells in a confluent monolayer. The highly folded nature of the two lateral membranes seen in the image probably determines the observed lateral intensity of high expression of the YFP fusion protein. The predominant surface localization of the product makes this a suitable model for analysis of tryptic cleavage sites with minimal interference from intracellular processing events of an artificial protein. Some YFP fusion protein was still intracellular as shown by the arrows in Figure 4.

Western blot analysis (Figure 4B) of the cell-solubilized extract detected mainly a single band corresponding in MW to the N-terminal glycosylated YFP fusion protein of the basolateral bile acid transporter. Some small fragments were detected below 29 kDa that could represent cellular degradation products of the intact transporter.

As shown in Figure 4C, transport activity by the fusion protein was identical to that of the native transporter as detected by $Na^+$-dependent transport of $^{14}C$-taurocholate.

*Tryptic Digestion.* Figure 5 presents a typical Western blot analysis of the undigested SBAT, and SBAT digested for 10 min and 1 h and is typical of three experiments. From the known sequence and MW analysis, it is possible to define the site of digestion by trypsin of the surface-expressed protein and to show that under the conditions of these experiments, trypsin does not access the cytoplasmic surface of the protein. It is also important that the C terminus is retained after digestion to allow identification based on staining of the YFP fused to the C terminus of the bile acid transporter. To detect all of the bands, ECL detection was used and the final image was overexposed.

Lane A (Figure 5) shows that YFP expressed alone in HEK-293 cells provided a band corresponding to a MW of 22 kDa, exactly the predicted MW of this protein.

Western blot analysis of the extract from cells expressing SBAT−YFP not treated with trypsin shows a major band corresponding to the full-length complex glycosylated YFP fusion protein at 75−85 kDa (lane B of Figure 5). In contrast with the Western blot analysis of SBAT−YFP shown in Figure 4B, this overexposed image reveals other bands besides the full-length complex glycosylated protein. The two bands at 68 and 60 kDa probably reflect the minor intracellular retention of the protein in the ER (core- and nonglycosylated protein), as seen highlighted with arrows in the confocal image of Figure 4A. There are also faint low MW fragments and one broad intense band corresponding to free YFP in this undigested sample, maybe because of degradation of the protein as it is recycled through the endosomal pathway from the plasma membrane. The bands observed in addition to the fully glycosylated SBAT−YFP in this lane are independent of exoplasmic digestion and reflect the fate of the fusion protein when expressed in HEK-293 cells.

Lanes C and D (Figure 5) represent the typical results of 10 and 60 min tryptic digestion of the intact cells. The full-length SBAT−YFP remains largely unchanged, showing that only a minor fraction of the fusion protein was digested in the experimental conditions used that were necessary to avoid penetration of trypsin. The heavily stained bottom band corresponds to the same level of processed free YFP that is seen in the intact cells not exposed to trypsin. The bands at 68 and 60 kDa are unchanged also, and their molecular weights correspond to core- and nonglycosylated fusion protein, respectively. The unchanged quantity of these bands indicates that trypsin did not penetrate the cells at this concentration and time of incubation.

Five additional bands were detected following tryptic digestion with molecular weights of 57, 52, 43, 37, and 33 kDa that increased in intensity with increasing time of exposure to trypsin and reflect tryptic cleavage of the surface protein. These were the only bands observed. After subtraction of the contribution of the C-terminal YFP portion of the protein, these have relative molecular weights of 35.9, 28.6, 21, 15.1, and 10.7 kDa and represent fragments of the bile acid transporter.

As shown in the sequences presented in Figure 6, the 10.7-kDa MW predicts a fragment containing the amino acids from Thr253 to Ala349 and thus the hydrophobic sequences H8 and H9 to the C terminus. The tryptic cleavage site is therefore between H7 and H8 (Arg252), placing it on the outside surface of the plasma membrane. The 15.1-kDa MW band would correspond to a sequence containing the amino acids from Ser213 to the C terminus, with H7, H8, and H9 to the C terminus. Hence, the site between H6 and H7 (Lys212) is also exposed to the outside. There is no band that can represent cleavage between H5 and H6; hence, this positively charged sequence is on the inside surface of the membrane. The 21-kDa fragment would begin at Gly158 and end at the C terminus and thus will contain H5, H6, H7, H6, H8, H9, and the C terminus of the transporter. Hence, this fragment results from cleavage between H4 and H5

- **H8-H9-COOH: 253T-349A (10.68 KDa)**
  TVSMETGCQNVQLCSTILNVAFPPEVIGPLFFFPLLYMIFQLGEGLLLIAIFWCYEKFKTPKDKTKMI
  YTAATTEETI PGALGNGTYK GEDCSPCTA
- **H7-H8-H9-COOH: 213S-349A (15.09 KDa)**
  SIMFAMTPLLIATSSLMPFIGFLLGYVLSALFCLNGRCRRTVSMETGCQNVQLCSTILNVAFPPEVI
  GPLFFFPLLYMIFQLGEGLLLIAIFWCYEKFKTPKDKTKMIYTAATTEETIPGALGNGTYKGEDCSP
  CTA
- **H5-H6-H7-H8-H9-COOH 158G-349A: (21 kDa)**
  GIVISLVLVLIPCTIGIVLKSKRPQYMRYVIKGGMIIILLCSVAVTVLSAINVGKSIMFAMTPLLIATSSL
  MPFIGFLLGYVLSALFCLNGRCRRTVSMETGCQNVQLCSTILNVAFPPEVIGPLFFFPLLYMIFQLG
  EGLLLIAIFWCYEKFKTPKDKTKMIYTAATTEETIPGALGNGTYKGEDCSPCTA
- **H3-H4-H5-H6-H7-H8-H9-COOH 87N-349A: (28.6 kDa)**
  NIEALAILVCGCSPGGNLSNVFSLAMKGDMNLSIVMTTCSTFCALGMMPLLLYIYSRGIYDGDLKD
  KVPYKGIVISLVLVLIPCTIGIVLKSKRPQYMRYVIKGGMIIILLCSVAVTVLSAINVGKSIMFAMTPLLI
  ATSSLMPFIGFLLGYVLSALFCLNGRCRRTVSMETGCQNVQLCSTILNVAFPPEVIGPLFFFPLLY
  MIFQLGEGLLLIAIFWCYEKFKTPKDKTKMIYTAATTEETIPGALGNGTYKGEDCSPCTA
- **H1-H2-H3-H4-H5-H6-H7-H8-H9-COOH 22P-349A: (35.9 kDa)**
  PTDLALSVILVFMLFFIMLSLGCTMEFSKIKAHLWKPKGLAIALVAQYGIMPLTAFVLGKVFRLKNIE
  ANIEALAILVCGCSPGGNLSNVFSLAMKGDMNLSIVMTTCSTFCALGMMPLLLYIYSRGIYDGDLK
  DKVPYKGIVISLVLVLIPCTIGIVLKSKRPQYMRYVIKGGMIIILLCSVAVTVLSAINVGKSIMFAMTPL
  LIATSSLMPFIGFLLGYVLSALFCLNGRCRRTVSMETGCQNVQLCSTILNVAFPPEVIGPLFFFPLL
  YMIFQLGEGLLLIAIFWCYEKFKTPKDKTKMIYTAATTEETIPGALGNGTYKGEDCSPCTA

FIGURE 6: Predicted five major sequences corresponding to the fragments seen in Figure 5 and their amino acid sequence that gives the exact MW seen on the gel.

(Lys157), placing this site on the outside surface. The 28.6-kDa band corresponds to the cleavage site (Lys86) between H2 and H3 beginning at Asn87 and therefore includes H3 to the C terminus. The band with a MW of 35.9 would correspond to a cleavage site at the N terminus (Arg21) and therefore includes all of the hydrophobic regions of the protein. The predicted amino acid sequences of the cleaved products are shown in Figure 6.

The blue arrows of Figure 2 show the five tryptic cleavage sites found in these experiments. The highly positively charged region between H5 and H6 is not cleaved, suggesting a cytoplasmic location for this region of the transporter, contrary to the seven transmembrane segment models proposed previously but consistent with the positive inside rule (*23*, *24*). Also, none of the arginines or lysines between H1 and H2 or between H3 and H4 were cleaved. Similarly, no cleavage at the two lysine residues just after H9 was observed (Figure 2). These are all located therefore on the cytoplasmic surface of the expressed transporter.

## DISCUSSION

The bile acid transporters have generally been predicted to contain seven transmembrane segments with extracytoplasmic location of the N terminus and cytoplasmic location of the C terminus with two hydrophobic sequences located on the outside of the membrane (*3*, *25*, *26*). However, our previous work had suggested that the bile acid transporters may have a 2D structure consisting of either nine or seven transmembrane segments with two re-entrant membrane-associated sequences containing H3 and H8 (*4*, *5*, *18*) but no long extra-cytoplasmic loops. Here, to identify integrated membrane segments in an *in vivo* system, we first tried a cell system expressing GFP N-terminal fusion-truncated protein vectors as described in the Materials and Methods and in Figure 2, using a H,K-ATPase $\beta$-glycosylation flag. The transporter sequences inserted into the expression vector contained successive hydrophobic segments in their natural orientation. The experimental data presented in these studies demonstrated that expression of the native sequence with truncations after the hydrophobic sequences H1, H2, and H3 generated products that suggest that each of these segments is membrane-spanning. Thus, H3 behaved as a membrane-inserted sequence in this *in vivo* system, in contrast to the seven transmembrane segment model (Figure 1A) and in contrast to its translation as an individual segment (*4*, *5*). H4 did not initiate translocation when tested in the GFP-HKMO1-H1-H4-$\beta$-flag setting. However, alanine insertion mutagenesis in the H4 region had resulted in a total loss of taurocholate uptake in HEK-293 cells, and it had both signal anchor and stop transfer properties when assayed as an individual fragment in the M0/M1 *in vitro* translation system (*4*, *5*). When the construct containing H8 and H8−H9 was expressed in the HEK-293 cells (Figure 3F), no membrane integration was found for H8, but H9 did membrane-insert.

There are a number of reports where a transmembrane segment does not integrate properly because there appears to be a need for a downstream hydrophobic sequence (*27*−*29*). Various known transmembrane segments such as TM5 and TM6 of the H,K and ER Ca ATPases (*8*, *13*) also did not show any membrane-insertion properties when assayed as individual segments in the microsomal translation system using the H,K-ATPase vector. A requirement for downstream sequences was also observed for the amino-terminal segment TM1 of the CFTR that also did not membrane-insert when assayed as an individual segment (*30*). Because these moderately hydrophobic sequences were found to be transmembrane segments in the native proteins, their insertion in the membrane must be dependent upon the surrounding hydrophobic segments and perhaps also on *in vivo* chaperones (*31*).

Thus, this newer N-terminal GFP fusion construct approach shows the limitations even in an *in vivo* situation. The membrane or flag sequences or the N-terminal GFP protein used in these vectors may interact with these hydrophobic segments during translocon transit of the liver transporter hydrophobic segments, providing misleading results.

When the C-terminus YFP fusion full-length bile acid transporter is expressed in the plasma membrane as shown by confocal microscopy, it is functional because it is bile salt transport-competent (Figure 4C) and inhibited by inhibitors of this transporter (data not shown). Further, the C-terminal-soluble YFP portion of the fusion protein cannot affect prior membrane integration of H9, and there are no other possibly confounding sequences such as the amino acids of the H,K ATPase.

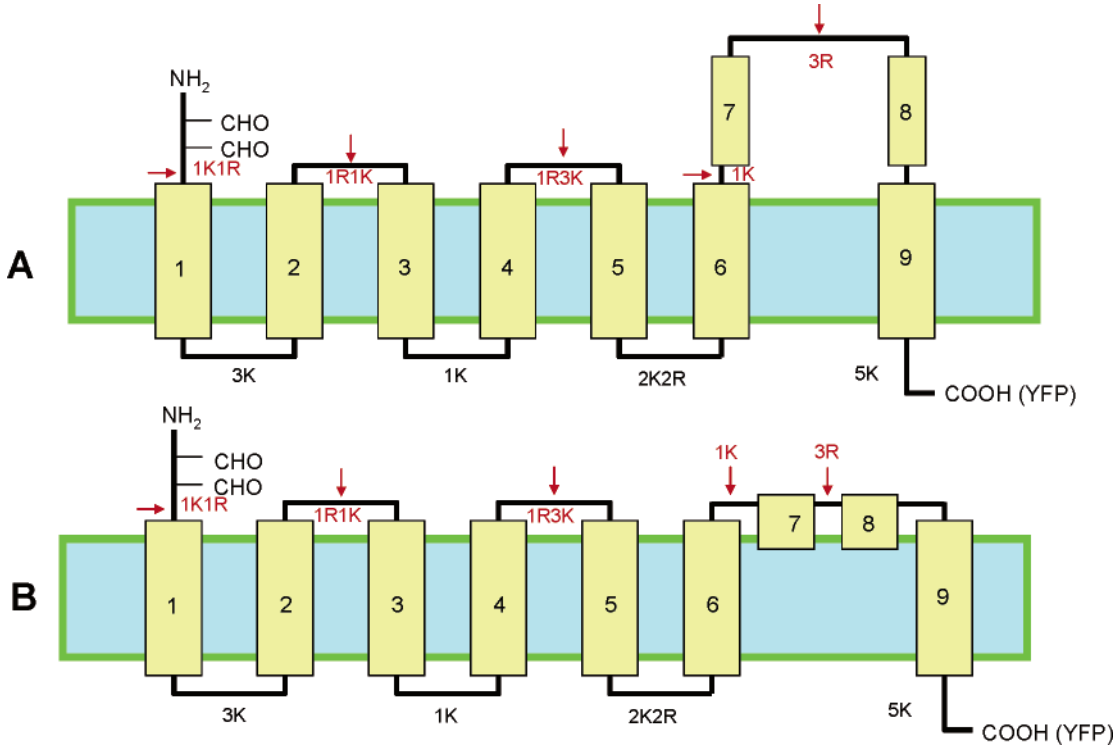


FIGURE 7: Representation of the topographic models resulting from the data presented in this paper. The top model A shows the seven membrane inserted sequences, H1, H2, H3, H4, H5, H6, and H9, with H7 and H8 represented as extending away from the membrane. The bottom model B shows these latter segments as membrane-associated or as re-entrant loops because alanine-scanning insertion into H7 and H8 inactivated the protein, suggesting ordered association between these two segments. Observed extracellular trypsin cleavage sites are shown by red arrows, and the corresponding lysine and arginine residues are shown in red font. Inaccessible intracellular lysine and arginine residues are shown in black font. CHO is an N-linked carbohydrate.

In both untreated and treated cells expressing the SBAT−YFP transporter, there is unchanged expression of core- and unglycosylated protein, because of retention in the endoplasmic reticulum. The finding that these are not cleaved after trypsin exposure shows the absence of trypsin penetration into the cells. The presence of free YFP and perhaps the fragment containing YFP and C-terminal portion of SBAT in untreated cells suggests internal fragmentation of the expressed protein perhaps because of the ER-associated degradation during translation or degradation of the protein as it is recycled through the endosomal pathway from the plasma membrane. Importantly, the intensity of this band stays unchanged after trypsin treatment. This is a clear indication that trypsin does not have intracellular access in the cleavage condition used. Otherwise, the intensity of this band would increase dramatically because of the presence of trypsin cleavage sites in the intracellular C terminus of SBAT (Figure 2A).

There are no trypsin cleavable residues between H8 and H9 (after Arg252) on the outside surface of the protein. With tryptic digestion, the production of two major fragments at 10.7 and 15.1 kDa provides clear evidence for the access of trypsin to Lys212 between H6 and H7 and to Arg252 between H7 and H8. Hence, both H7 and H8 must be exposed on the exoplasmic surface of the protein. The fragment at 21 kDa shows that there is access of trypsin to Lys157 just before H5. The 28.6- and 35.9-kDa fragments suggest tryptic access to Lys86 just before H3 and Arg21 before H1, respectively. The band intensities of tryptic fragments at 28.6 and 35.9 kDa were weak (Figure 5). To intensify these bands, we tried different conditions such as higher trypsin concentrations or a longer time of digestion. However, smaller fragments corresponding to the trypsin digestion at Arg252, Lys212, and Lys157 were always present at high levels relative to the larger 28.6- and 35.9-kDa fragments. These data indicate the greater exposure and higher sensitivity to trypsin of the cleavage sites at Arg252, Lys212, and Lys157. The conditions chosen were optimal to reveal all of the cleavage sites indicated in the model and, at the same time, to avoid penetration of trypsin into cytoplasm. The cleavage of the intact transporter at the N terminus and between H2 and H3, detected as minor bands by cleavage of the intact protein, is also consistent with the results of expression of the fusion proteins containing these segments where H1, H2, and H3 performed as membrane-inserted segments with the N terminus exposed to the outside surface.

Notably, there is no evidence of cleavage of any of the large number of positively charged amino acids in the region between amino acids 177 and 189, strongly suggesting that this region is cytoplasmic as predicted by the positive inside rule (*23*, *24*, *31*). The fragment appearing at 57 kDa is likely due to N-terminal cleavage at the Arg21 on the N-terminal side of H1. The tryptic digest pattern therefore predicts the presence of seven transmembrane segments where the hydrophobic segments H7 and H8 are not membrane-integrated as shown in the models of Figure 7.

Whereas the cellular translation of the GFP-truncated constructs clearly shows the transmembrane orientation of H1, H2, H3, and H9 similar to other truncation methods (*4*, *5*) but then fails to determine the topography of H4−H8, the results of tryptic digestion show membrane integration

of not only H1, H2, and H3 but also importantly a similar topography for H4, H5, H6, and H9 and an exoplasmic location of H7 and H8. Hence, the combination of the truncated construct translation *in vivo* for the more N-proximal membrane segments and tryptic digestion of the C-terminal-flagged intact transporter for the more C-proximal segments provide compelling evidence for the new seven transmembrane segment model proposed.

The novel, seven transmembrane segment models for the liver $Na^+$-dependent bile salt transporter are shown in parts A and B of Figure 7, where H7 and H8 are not transmembrane-segments. This model is likely also to represent the topography of the ileal $Na^+$-dependent bile salt transporter given its homology and hydropathy similarity.

There are two models shown, where either H7 and H8 are either present as long exoplasmic loops (model A) or may be membrane-associated as relatively mobile re-entrant membrane loops (model B), which would be consistent with their interaction with the substrate and the results of alanine-scanning mutagenesis, where insertion into either of these segments inactivated the transporter (*5*, *18*, *21*, *26*).

Recently, data have been presented and interpreted as showing a seven transmembrane segment model using glycosylation insertion mutants into the ileal bile acid transporter, where the natural N-terminal glycosylation sites were deleted by mutagenesis (*6*). These insertions were in the H3 segment, the subsequent sequence connecting H3 to H4, and in the sequence corresponding to H8. These data were thought to confirm the first model (A) of Figure 1, where both H3 and H8 are exoplasmic (*6*). However, these ileal transporter glycosylation site mutated proteins were functionally inactive. Further, the liver bile salt transporter contains two natural N-glycosylation sites, one in the H3 segment and one in the sequence between H3 and H4 (red dots in Figure 2), and these are clearly not glycosylated in the native SBAT (*5*). However, the results of insertion of a glycosylation site in H8 of the ileal transporter (*6*) agree with the model generated from tryptic digestion of the intact, plasma-membrane-expressed functional C-terminal YFP fusion protein.

The work presented in this paper points out the need for multiple approaches to a proper definition of the topography of a polytopic membrane protein and the limitations of either translation *in vitro* or *in vivo* of truncated constructs and the difficulty of interpretation of alanine-insertion scanning. These are avoided by analysis of the intact C-terminal YFP fusion protein expressed largely in the plasma membrane. The resulting seven transmembrane segment model is novel in that the bile salt and inhibitor-binding region are contained within the predicted exoplasmic loop containing both H7 and H8. Their structure seems to be restricted because alanine insertion into H7 and H8 regions resulted in a loss of transport activity of the transporter. This finding leads to the postulate that H7 and H8 are closely associated with the membrane, perhaps as re-entrant loops (model B in Figure 7), and interact specifically with each other, perhaps to facilitate bile salt transport.

This tryptic digest approach with a C-terminal tag that is readily visualized by confocal microscopy and Western analysis may provide the most reliable means of defining the topography of plasma-membrane proteins with a cytoplasmic C terminus. It is important to emphasize the need for limited digestion by using low concentrations of trypsin and rapid inhibition of any residual tryptic cleavage during lysis and accurate analysis of digested fragments.

## ACKNOWLEDGMENT

Thanks are due to Dr. Jeff Kraut and Dr. David Scott for a careful reading of the manuscript.

## REFERENCES

1. Chiang, J. Y., Kimmel, R., and Stroup, D. (2001) Regulation of cholesterol 7α-hydroxylase gene (CYP7A1) transcription by the liver orphan receptor (LXRα), *Gene 262*, 257−265.
2. Spady, D. K., Cuthbert, J. A., Willard, M. N., and Meidell, R. S. (1995) Adenovirus-mediated transfer of a gene encoding cholesterol 7α-hydroxylase into hamsters increases hepatic enzyme activity and reduces plasma total and low-density lipoprotein cholesterol, *J. Clin. Invest. 96*, 700−709.
3. Hagenbuch, B., and Meier, P. J. (1994) Molecular cloning, chromosomal localization, and functional characterization of a human liver $Na^+$/bile acid cotransporter, *J. Clin. Invest. 93*, 1326−1331.
4. Hallen, 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.
5. 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.
6. Zhang, E. Y., Phelps, M. A., Banerjee, A., Khantwal, C. M., Chang, C., Helsper, F., and Swaan, P. W. (2004) Topology scanning and putative three-dimensional structure of the extracellular binding domains of the apical sodium-dependent bile acid transporter (SLC10A2), *Biochemistry 43*, 11380−11392.
7. Audigier, Y., Friedlander, M., and Blobel, G. (1987) Multiple topogenic sequences in bovine opsin, *Proc. Natl. Acad. Sci. U.S.A. 84*, 5783−5787.
8. Bayle, D., Weeks, D., and Sachs, G. (1995) The membrane topology of the rat sarcoplasmic and endoplasmic reticulum calcium ATPases by *in vitro* translation scanning, *J. Biol. Chem. 270*, 25678−25684.
9. Bayle, D., Weeks, D., and Sachs, G. (1997) Identification of membrane insertion sequences of the rabbit gastric cholecystokinin-A receptor by *in vitro* translation, *J. Biol. Chem. 272*, 19697−19707.
10. Melchers, K., Weitzenegger, T., Buhmann, A., Steinhilber, W., Sachs, G., and Schafer, K. P. (1996) Cloning and membrane topology of a P type ATPase from *Helicobacter pylori*, *J. Biol. Chem. 271*, 446−457.
11. Zizak, M., Cavet, M. E., Bayle, D., Tse, C. M., Hallen, S., Sachs, G., and Donowitz, M. (2000) $Na^+/H^+$ exchanger NHE3 has 11 membrane spanning domains and a cleaved signal peptide: Topology analysis using *in vitro* transcription/translation, *Biochemistry 39*, 8102−8112.
12. Tatishchev, S., Abuladze, N., Pushkin, A., Newman, D., Liu, W., Weeks, D., Sachs, G., and Kurtz, I. (2003) Identification of membrane topography of the electrogenic sodium bicarbonate cotransporter pNBC1 by *in vitro* transcription/translation, *Biochemistry 42*, 755−765.
13. Bamberg, K., and Sachs, G. (1994) Topological analysis of $H^+,K^+$-ATPase using *in vitro* translation, *J. Biol. Chem. 269*, 16909−16919.
14. von Heijne, G. (1992) Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule, *J. Mol. Biol. 225*, 487−494.
15. Sipos, L., and von Heijne, G. (1993) Predicting the topology of eukaryotic membrane proteins, *Eur. J. Biochem. 213*, 1333−1340.
16. Braun, P., Persson, B., Kaback, H. R., and von Heijne, G. (1997) Alanine insertion scanning mutagenesis of lactose permease transmembrane helices, *J. Biol. Chem. 272*, 29566−29571.
17. Dohke, Y., and Turner, R. J. (2002) Evidence that the transmembrane biogenesis of aquaporin 1 is cotranslational in intact mammalian cells, *J. Biol. Chem. 277*, 15215−15219.
18. 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.
19. Sonnhammer, E. L., von Heijne, G., and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences, *Proc. Int. Conf. Intell. Syst. Mol. Biol. 6*, 175−182.
20. Tusnady, G. E., and Simon, I. (2001) The HMMTOP transmembrane topology prediction server, *Bioinformatics 17*, 849−850.
21. Hallen, 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.
22. Kyte, J., and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol. 157*, 105−132.
23. Wallin, E., and von Heijne, G. (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms, *Protein Sci. 7*, 1029−1038.
24. von Heijne, G., and Gavel, Y. (1988) Topogenic signals in integral membrane proteins, *Eur. J. Biochem. 174*, 671−678.
25. Hagenbuch, B., Stieger, B., Foguet, M., Lubbert, 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.
26. 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.
27. Skach, W. R., and Lingappa, V. R. (1994) Transmembrane orientation and topogenesis of the third and fourth membrane-spanning regions of human P-glycoprotein (MDR1), *Cancer Res. 54*, 3202−3209.
28. Zhang, J. T. (1996) Sequence requirements for membrane assembly of polytopic membrane proteins: Molecular dissection of the membrane insertion process and topogenesis of the human MDR3 P-glycoprotein, *Mol. Biol. Cell. 7*, 1709−1721.
29. Ota, K., Sakaguchi, M., von Heijne, G., Hamasaki, N., and Mihara, K. (1998) Forced transmembrane orientation of hydrophilic polypeptide segments in multispanning membrane proteins, *Mol. Cell. 2*, 495−503.
30. Lu, Y., Xiong, X., Helm, A., Kimani, K., Bragin, A., and Skach, W. R. (1998) Co- and posttranslational translocation mechanisms direct cystic fibrosis transmembrane conductance regulator N terminus transmembrane assembly, *J. Biol. Chem. 273*, 568−576.
31. van Geest, M., and Lolkema, J. S. (2000) Membrane topology and insertion of membrane proteins: Search for topogenic signals, *Microbiol. Mol. Biol. Rev. 64*, 13−33.

BI051291X