Organization of the Membrane Domain of the Human Liver Sodium/Bile Acid Cotransporter†

S. Hallén, O. Mareninova, M. Brändén, and G. Sachs*

University of California at Los Angeles and Wadsworth Veterans Administration Hospital, Los Angeles, California 90073

Received December 12, 2001; Revised Manuscript Received March 8, 2002

ABSTRACT: Mammalian sodium/bile acid cotransporters (SBATs) are glycoproteins with an exoplasmic N-terminus, an odd number of transmembrane regions, and a cytoplasmic C-terminus. Various algorithms predict eight or nine membrane-embedded regions derived from nine hydrophobic stretches of the protein (H1−H9). Three methods were used to define which of these were transmembrane or membrane-associated segments in the liver bile acid transporter. The first was in vitro translation/insertion scanning using either single hydrophobic sequences between the N-terminal domain of the α-subunit of the gastric H,K-ATPase and the C-terminal domain of the β-subunit that contains five N-linked glycosylation exoplasmic flags or using constructs beginning with the N-terminus of the transporter of various lengths and again ending in the C-terminus of the H,K-ATPase β-subunit. Seven of the predicted segments, but not the amphipathic H3 and H8 sequences, insert as both individual signal anchor and stop transfer sequences in the reporter constructs. These sequences, H3 and H8, are contained within two postulated long exoplasmic loops in the classical seven-transmembrane segment model. The H3 segment acts as a partial stop transfer signal when expressed downstream of the endogenous H2. In a similar manner, the other amphipathic segment, H8, inserts as a signal anchor sequence when translated in the context with the upstream transporter sequence in two different glycosylation constructs. Alanine insertion scanning identified regions of the transporter requiring precise alignment of sequence to form competent secondary structures. The transport activity of these mutants was evaluated either in native protein or in a yellow fluorescent protein (YFP) fusion protein construct. All alanine insertions in H3 and H8 abolished taurocholate uptake, suggesting that both these regions have structures with critical intramolecular interactions. Moreover, these insertions also prevented trafficking to the plasma membrane as assessed by confocal microscopy with a polyclonal antibody against either the C-terminus of the transporter or the YFP signal of the YFP−transporter fusion protein. Two glycosylation signals inserted in the first postulated loop region and four of five such signals in the second postulated loop region were not recognized by the oligosaccharide transferase, and the L256N mutation exhibited 10% glycosylation and was inactive. These findings support a topography with nine membrane-spanning or membrane-associated segments.

Sodium/bile acid cotransporters (SBATs)[1] are polytopic integral membrane proteins that participate in the enterohepatic circulation of bile acids in mammals. Two homologous transporters are involved in the reabsorption of bile acids, one absorbing from the intestinal lumen, the bile duct, and the kidney with an apical localization and the other being found in the basolateral membranes of hepatocytes. These transporters have been cloned from a number of species, including hamsters, rats, rabbits, mice, and humans (*1*−*9*). The apical and basolateral transporters share Na^+-dependent bile acid uptake and ∼35% identity as well as analogous hydrophobicity profiles. Several independent observations have established that they have an exoplasmic, glycosylated N-terminus and a cytoplasmic C-terminus, resulting in an odd number of transmembrane sequences (*9*−*11*). Existing two-dimensional (2D) models of the transporter postulate seven membrane-spanning regions with two long exoplasmic loops (*9*, *10*). Both these long segments hold motifs that are fully conserved among all identified SBATs, and their amphipathic character makes them suitable for binding and translocation of sodium and bile acid. However, algorithms for membrane protein topography generally predict eight to nine membrane-spanning segments; therefore, either one or both of the long exoplasmic loops are predicted to contain transmembrane domains (*12*−*14*). The basis for the different models investigated here is shown in Figure 1 which shows the Kyte−Doolittle 11-amino acid moving average hydropathy profile of the protein. The long exoplasmic loops of the seven-membrane segment model are between amino acids 77 and 119 and between amino acids 245 and 286 containing H3 and H8 of Figure 1. These regions contain either two

† Supported by the U.S. Veterans Administration and NIH Grants DK46917, -53462, -41301, and -17294.

* To whom correspondence should be addressed: Membrane Biology Laboratory, West Los Angeles VA Medical Center, 11301 Wilshire Blvd., Bldg. 113, Rm 324, Los Angeles, CA 90073. Telephone: (310) 268-4672. Fax: (310) 312-9478. E-mail: gsachs@ucla.edu.

[1] Abbreviations: ER, endoplasmic reticulum; SA, signal anchor; ST, stop transfer; SBAT, sodium bile/acid cotransporter; Hum NTCP, human liver (basolateral) sodium taurocholate cotransporting protein; H, hydrophobic segments; HMMTOP, hidden Markov model topology prediction; TM, transmembrane; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; OST, oligosaccharide transferase; PNGase F, endoglycosidase F.

10.1021/bi012152s CCC: $22.00
Published on Web 05/15/2002

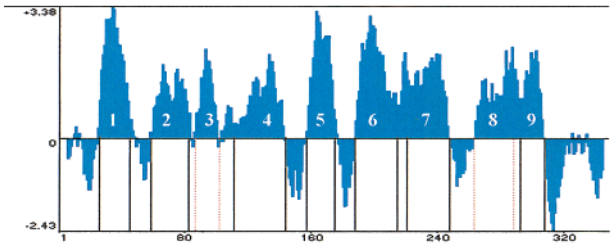

FIGURE 1: 2D maps of the human liver bile acid transporter. The seven- and nine-transmembrane segment models are illustrated in relation to the Kyte−Doolittle hydropathy analysis. It can be seen that more than one type of transmembrane segment arrangement can be derived from this analysis, and experimental verification is required.

membrane-associated segments or two membrane-spanning segments in the models based on the data given below.

The uncertain prediction of membrane insertion of polytopic proteins, especially transporters, often requires multiple experimental methods for defining the 2D arrangement of the membrane domain. There are numerous protocols for topology scanning of membrane proteins utilizing, for example, side- and site-specific reagents, glycosylation flag insertion, cysteine scanning, and enzyme fusions (for a review, see ref *15*). One method used here was examination of glycosylation flag constructs using in vitro transcription/translation/membrane insertion scanning (*16*) to find membrane-inserted segments. Another method that was employed was alanine insertion scanning with transient expression of the normal or YFP fusion liver transporter to search for domains forming well-defined tertiary structures. This has been used to detect secondary structures with highly ordered packing such as transmembrane α-helices or β-sheets (*17*). Insertion of a single residue in any such highly ordered domain is expected to disturb structure and function, while an addition within a random coil or loop region is predicted to be more permissive. Hence, alanine insertion into the long exoplasmic loops would not be expected to disrupt plasma membrane trafficking since trafficking signals are cytoplasmic and not exoplasmic. Then, natural or mutated N-linked glycosylation sites in the regions of the postulated exoplasmic loops were investigated for their competence in translation or expression of the protein.

Translation/insertion scanning, alanine insertion, and glycosylation site mutagenesis support membrane integration or association of all nine sequences predicted by the HMMTOP algorithm. The data do not support a model with seven transmembrane segments and two large exoplasmic loops. On the basis of the known positions for the N- and C-termini, a 2D structure encompassing nine membrane-associated segments is supported by the data presented here.

EXPERIMENTAL PROCEDURES

Topographic Analysis of Primary Sequence. Putative membrane-spanning regions for the human basolateral sodium/bile acid cotransporter were identified using two programs based on a hidden Markov model, TMHMM (*13*), HMMTOP (*14*), and the neural network system, PhdTopology, developed by Rost et al. (*12*).

Vector Construction. The constructs used for signal anchor (SA) (HK M0)/stop transfer (ST) (HK M1) screening and analysis of the C-terminal truncated SBAT sequences were all based on the pGEM7zf+ (Δ*Hin*dIII) plasmid (Promega, Madison, WI) (*16*). Sequences to be investigated for SA/ST activity were cloned into a *Bgl*II/*Hin*dIII linker region between the first 101 and 139 amino acids from the N-terminus of the rabbit H,K-ATPase α-subunit and the 177 C-terminal amino acids of the rabbit H,K-ATPase β-subunit, containing five N-linked glycosylation consensus sites (*16*). The plasmids encoding the C-terminal truncated SBAT sequences were constructed by excising the M0 sequence with restriction enzymes *Xho*I and *Hin*dIII and replacing it with the indicated transporter sequences downstream of the T7 promoter in the pcDNA3 expression vector (Invitrogen, Carlsbad, CA).

PCR. PCR primers were selected using the OLIGO 4.0 primer analysis program (National Biosciences, Inc., Plymouth, MN) or the Primer Select software included in the Laser Gene package. Sense primers for insertion into the HK M0/M1 vectors were 5′ flanked with a *Bgl*II site, and the primers for analysis of the native sequences contained an *Xho*I site as well as the Kozak ACC consensus in the corresponding position. All antisense primers have a *Hin*dIII site on the 3′ end. The amplification reaction mixtures (100 μL) contained 1 ng of template, pCMV5-humNTCP, each sense and antisense primer at 1 μM, each deoxynucleotide at 200 μM (Promega), 1 unit of VentR DNA polymerase, and 1× ThermoPol reaction buffer (New England Biolabs).

Ligation, Cloning, and Plasmid Purification. The PCR products were purified using the PCR product Resin/Wizard Minicolumns (Promega) and then digested with the restriction enzymes (Promega) *Bgl*II and *Hin*dIII (HK M0/M1 constructs) or *Xho*I and *Hin*dIII (C-terminal truncated native sequences). Finally, the inserts were gel purified (0.8% agarose gel) and recovered by centrifugation in a spin column (Spin-X, Costar). The HK M0/M1 vectors were digested and purified using the same protocol, but were dephosphorylated prior to gel purification (calf intestine phosphatase, Pharmacia). The inserts were ligated into the appropriate vectors and subsequently transformed into *Escherichia coli* JM109 cells by a standard electroporation procedure (Bio-Rad). Plasmids were purified using anion exchange columns (Qiagen, Chatsworth, CA), and all inserts were verified by dideoxy sequencing.

In Vitro Transcription/Translation Reactions. Membrane insertion properties of all the investigated sequences were probed using the TNT T7 coupled rabbit reticulocyte lysate system in the presence of canine pancreatic microsomes (Promega). All reaction components were mixed on ice; microsomal membranes were added, and the transcription/translation reaction was initiated (on ice) by addition of the vector. Transcription/translation reactions were carried out at 30 °C for 2 h in the presence of [^{35}S]methionine (ICN). Reaction mixtures lacking the microsomal membranes constituted the negative controls.

Alkaline Extraction of Microsomal Membranes. Insertion into the microsomal membranes was confirmed by extracting the membranes with 0.1 M Na_2CO_3 on ice for 30 min. The microsomes were recovered by centrifugation in a Beckman Airfuge at 25 psi for 30 min and resuspended in 1 volume of 100 mM Tris-HCl (pH 7.4) with 1 volume of 2× sample buffer. As a control for adhesion rather than insertion, the microsomal membranes were added after translation and extracted as described above. The degree of SRP-dependent

insertion was calculated using the Ambis/RFLP software by comparing the amount of membrane-bound product after a translation reaction in the presence of microsomes with that of the control described above. Total microsomal protein was used as internal standard for membrane recovery and estimated from the Coomassie-stained gels.

Deglycosylation. Core glycosylation of translocated consensus sites was validated by post-translational deglycosylation using endoglycosidase F (Boehringer Mannheim). For deglycosylation, 6 μL of the translation reaction mixture was dissolved in buffer, yielding a final concentration of 50 mM Tris-HCl (pH 7.8), 2 mM EDTA, and 0.1% SDS, and incubated with 0.2 unit of *N*-glycosidase F for 2 h at room temperature. A control reaction was always run in parallel without the enzyme.

Site-Directed Mutagenesis, Cloning Procedures, and DNA Sequencing. Site-directed mutations in the basolateral human sodium/bile acid cotransporter were introduced by using the QuikChange kit from Stratagene. Mutant oligonucleotides for alanine insertions were constructed using the Primer Select software and purchased from Gibco Life Technologies. cDNA for the human basolateral SBAT was a kind gift of P. Dawson (Wake Forest University, Winston-Salem, NC), and the transporter cDNA was cloned into the pcDNA3.1 (+)-vector from Invitrogen for use as a template in the mutation PCR. The DnaI-digested PCR products (1−2 μL) following the QuikChange protocol were transformed into JM109 cells by electroporation using Bio-Rad's electropulser at 1.5 kV, 800 Ω, and 25 μF. Plasmids were purified using anion exchange columns (Qiagen, Chatsworth, CA), and all mutations were verified by dideoxy sequencing.

Construction of the YFP Fusion Protein. The sequence of human basolateral SBAT was fused with an *Xho*I site at the 5′′ end and a *Sac*II site at the 3′′ end by PCR using full-length transporter cDNA as a template. These restriction sites were compatible with the pEYFP-N1 vector (Clontech, Palo Alto, CA). The PCR products were gel-purified, digested at the sites incorporated at the end of PCR products, and inserted into pEYFP-N1 so that the YFP would be located at the C-terminus of the human basolateral SBAT. All plasmid constructs were identified by restriction digestion analysis. The positive clones containing the transporter cDNA insert were verified by sequencing.

Expression of Alanine Insertion Mutants into HEK 293 Cells and Measurement of Transport Activity. HEK 293 cells grown to 70−80% confluence in Dulbecco's modified Eagle's 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) were trypsinized by washing the monolayer with PBS before adding a 0.05% trypsin/53 mM EDTA mixture (Gibco Life Technologies), and single cells were recovered by centrifugation. Approximately 3×10^4 cells in 100 μL of DMEM were then added to each well in a Cytostar SPA 96-well plate. After 1 day, cells were transfected with 0.075 μg of DNA/well using the FuGENE6 transfection reagent from Roche Molecular Biochemicals according to the manufacturer's recommendations. On day 2−3 (24−48 h post-transfection), cells were 90−100% confluent in the Cytostar-T 96-well plate, and uptake of [^{14}C]-taurocholic acid (NEN Life Science Products, Inc.) was assessed using a Micro-β scintillation plate reader (EG&G/Wallac). After the transport assay, HEK 293 cells were lysed in a buffer containing 15% SDS, 8 M urea, 10% sucrose, 62.5 mM Tris-HCl (pH 6.8), 10 mM EDTA, and 5 mM DTT. The approximate expression levels of the human basolateral SBATs were typically detected by Western blotting of 10 μg of total cell lysate separated on a precast 1.5 mm 10% Tris/glycine gel (Novex) and then electrotransfer to PVDF membrane (Millipore Immobilon-P, 0.22 μM). The efficiency of the electrotransfer was verified by Coomassie staining of the membrane prior to immunoblotting. The primary antibody for detection of the human basolateral SBAT was an affinity-purified polyclonal antibody from rabbit raised against the C-terminal sequence, GNGTYKGEDCSPCTA (Genemed Inc., San Francisco, CA). The secondary antibody was a goat anti-rabbit HRP-conjugated polyclonal antibody from American Qualex. Detection was with the ECL-Plus chemiluminescence reaction kit from Amersham/Pharmacia.

In some instances, when an YFP fusion protein was used, anti-GFP antibody (Roche Molecular Biochemicals) was used in place of the transporter antibody. This anti-GFP antibody was able to recognize unglycosylated protein, protein that was core glycosylated, and fully glycosylated protein on a Western blot.

Construction of N-Linked Glycosylation Sites in the H3 and H8 Regions of the Transporter. The mutant lacking the natural N-terminal glycosylation sites was created by site-directed substitutions of Asn6 and Asn11 with glutamine using the YFP fusion wild-type Ntcp construct as a template. This mutant then was used as a template for generating artificial glycosylation sites by site-directed substitutions of E89 with S and G97 with N in the H3 region and R249 with T, M256 with N, Q264 with T, L265 with N, and A273 with S in the H8 region of the transporter. These artificial glycosylation sites are at positions likely to be substrates for the oligosaccharide transferase if they extend into a long exoplasmic loop. Glycosylation of these mutants was detected on Western blots of transiently transfected HEK 293 cells using the anti-GFP antibody as described above and membrane trafficking as described in detail below.

Confocal Localization of the Human Basolateral Wild-Type and Mutant Transporters Expressed in HEK 293 Cells. (a) Immunostaining. HEK 293 cells were grown as described above to 70−80% confluence on Fisher Cover Glasses precoated with polylysine and placed in a six-well tissue culture plate containing 2 mL of DMEM per well. Cells were cotransfected on day 1 with 0.9 μg of human basolateral SBAT wild-type or mutant cDNA and 0.1 μg of CFP-Mem cDNA (Clonetech) using the FuGENE6 reagent. On day 2, the cells were washed twice with 2 mL of T-TBS before fixation in 100% MeOH (−20 °C) for 7 min. The cells were then rehydrated in 2 mL of T-TBS prior to preblocking in a 5% nonfat dry milk/T-TBS mixture for 30 min at room temperature using an orbital shaker for gentle mixing. The primary antibody (the same as described above) was used at a 1:500 dilution (5% nonfat dry milk/T-TBS), and the fixed cells were incubated for 2 h at room temperature. The samples were subsequently washed four times with 2 mL of T-TBS (quick rinse and three 10 min steps) prior to adding a goat anti-rabbit IgG (H+L) F(ab′)2 fragment TRITC antibody from American Qualex Inc. diluted 1:500 in a 5% nonfat dry milk/T-TBS mixture. After being washed as described above, the samples were mounted on cover slides

using Dako fluorescence mounting medium and images recorded using a Zeiss LSM510 confocal microscope with filter settings for rhodamine and cyan fluorescent protein. Image resolution using a 100× Neofluor objective and LSM510 software was 1024 × 1024 pixels.

(b) YFP Fluorescence. Confocal microscopy of YFP fusion protein was performed as described above, but the YFP fusion proteins were used to transfect the cells. Fixed cells were observed directly, and YFP fluorescence was detected using excitation and emission wavelengths of 514 and 530 nm, respectively.

RESULTS

Predictions of Transmembrane Regions and Overall Topography. Various algorithms (HMMTOP, TMHMM, and PhDTopology), which were used to define membrane-inserted segments, predict either eight or nine membrane-inserted sequences. The discrepancy is due to the prediction of the location of the region defined at H8, the last long hydrophobic sequence in the protein. HMMTOP predicts that a sequence in this region will be membrane-inserted and, therefore, that the protein will have nine transmembrane segments, with the N-terminus exoplasmic and the C-terminus cytoplasmic. The others, although predicting similar locations for the N- and C-termini, do not define this region as intramembranal, resulting in eight membrane segments. This is not compatible with an uneven number of membrane-spanning segments so, in addition, H3 is left out of the membrane in the seven-membrane segment model. Figure 1 shows a Kyte−Doolittle hydropathy profile (11-amino acid moving average) and defines the nine hydrophobic segments of the transporter.

Positions of the N- and C-Termini. The human basolateral SBAT is known to utilize only the two N-terminal consensus sites for N-linked glycosylation, although there is an N-linked site within the transporter at Asn103 close to the outer edge of the predicted exoplasmic loop between TM2 and TM3 in the seven-transmembrane topography but at the cytoplasmic edge of TM3 in the nine-segment model. Two additional N-linked sites are present, one at the exoplasmic boundary of TM3 at Asn117 in the seven-TM model and one in the cytoplasmic tail at Asn336. The N-linked site at Asn103 in the seven-membrane segment model is far enough away from the hydrophobic sequences to be recognized by the oligosaccharide transferase (*24*) if this region were exoplasmic. Immunostaining has shown that the C-terminus has a cytoplasmic localization (*9*).

The wild-type transporter was expressed in vitro with or without a glycosylation reporter sequence fused to the C-terminus. The fusion protein construct is outlined in Figure 2A, and the phosphorimages from in vitro expression of the wild-type transporter along with the glycosylation fusion construct are shown in Figure 2B. Both products exhibit a molecular mass increase of ∼6 kDa in the presence of ER microsomes (Figure 2B, lanes 2 and 4, as indicated by the diamond on the left). This shift corresponds in both cases only to glycosylation of the endogenous N-terminal N-linked glycosylation sites, N5AS and N11FT. The C-terminal flag is not glycosylated, and this result confirms that the SBAT C-terminus is cytoplasmic. These experiments confirm the presence of an odd number of transmembrane segments in the SBATs. Moreover, the 6 kDa shift in relative molecular mass (MWt) shows that only the two N-terminal sites of the five potential N-linked glycosylation sites in the native transporter are utilized. In vitro translation of the SBAT-β fusion protein inserted into the pGEM vector results also in a lower-MWt product marked by an arrow in lanes 3 and 4 Figure 2B. The apparent molecular mass of this smaller product may be due to utilization of the possible downstream ATG initiation sites.

Insertion Properties of the Predicted Sequences Expressed in the HK M0 Vector. The segments predicted to be transmembranous by any of the algorithms were tested for signal anchor activity in a construct starting with a cytoplasmic anchor consisting of the first 101 amino acids of the H,K-ATPase α-subunit and the SBAT membrane segment sequence and ending with a C-terminal 177-amino acid glycosylation reporter sequence from the H,K-ATPase β-subunit. The HK M0 construct along with the tested SBAT sequences is illustrated in Figure 3A. Sequences that act as signal anchors will initiate translocation of the glycosylation sequences into the ER microsome and expose them to the oligosaccharide transferase. The resulting molecular mass shift (∼10 kDa) is detected after SDS−PAGE and phosphorimaging. The HK M1 construct, identical to HK M0 but also containing the first transmembrane region of the H,K-ATPase α-subunit upstream of the glycosylation flag, was used as a control for translocation and glycosylation efficiency. Translation of HK M1 in the presence of ER microsomes results in a product with a molecular mass shift corresponding to the core glycosylation of the five NXS/T sites present in the reporter sequence (Figure 3B, lane 1) as confirmed by treating half the sample with PNGase F (Figure 3B, lane 2).

The results of the experiments using the HK M0 constructs encompassing the nine putative transmembrane segments of the human basolateral SBAT, H1−H9, are shown in lanes 3−11 of Figure 3B. All translation products except for the HK M0-H3 and H8 constructs (lanes 5 and 10) initiate translocation and glycosylation of the reporter sequence. Sequences H1, H2, H4−H7, and H9 are all capable of acting as signal anchor sequences in this construct. H3 and H8 lack one or both of these properties when expressed in the HK M0 vector (lane 5 and 10). Molecular mass shifts due to glycosylation of translation products were all confirmed by PNGase F deglycosylation as shown, for example, for the HK M1 translation in lane 2 of Figure 3B. The deglycosylated products of all tested constructs exhibited the same molecular masses as the proteins translated in the absence of ER microsomes (data not shown).

Stop Transfer Activity of Putative Transmembrane Segments Tested in the HK M1 Vector. Figure 4A displays the construct used for examination of stop transfer properties of the SBAT sequences analyzed above. The ability to block the translocation of the reporter sequence initiated by M1 was monitored by the disappearance of glycosylated product and confirmed by alkaline extraction of the nonglycosylated product. In vitro translation of the HK M1 construct in the presence of ER microsomes results in a membrane-inserted and glycosylated product that withstands subsequent alkaline extraction (Figure 4B, lanes 1 and 2).

Of the tested SBAT hydrophobic sequences, H1, H2, H4−H7, and H9 prevented glycosylation of the reporter sequence.

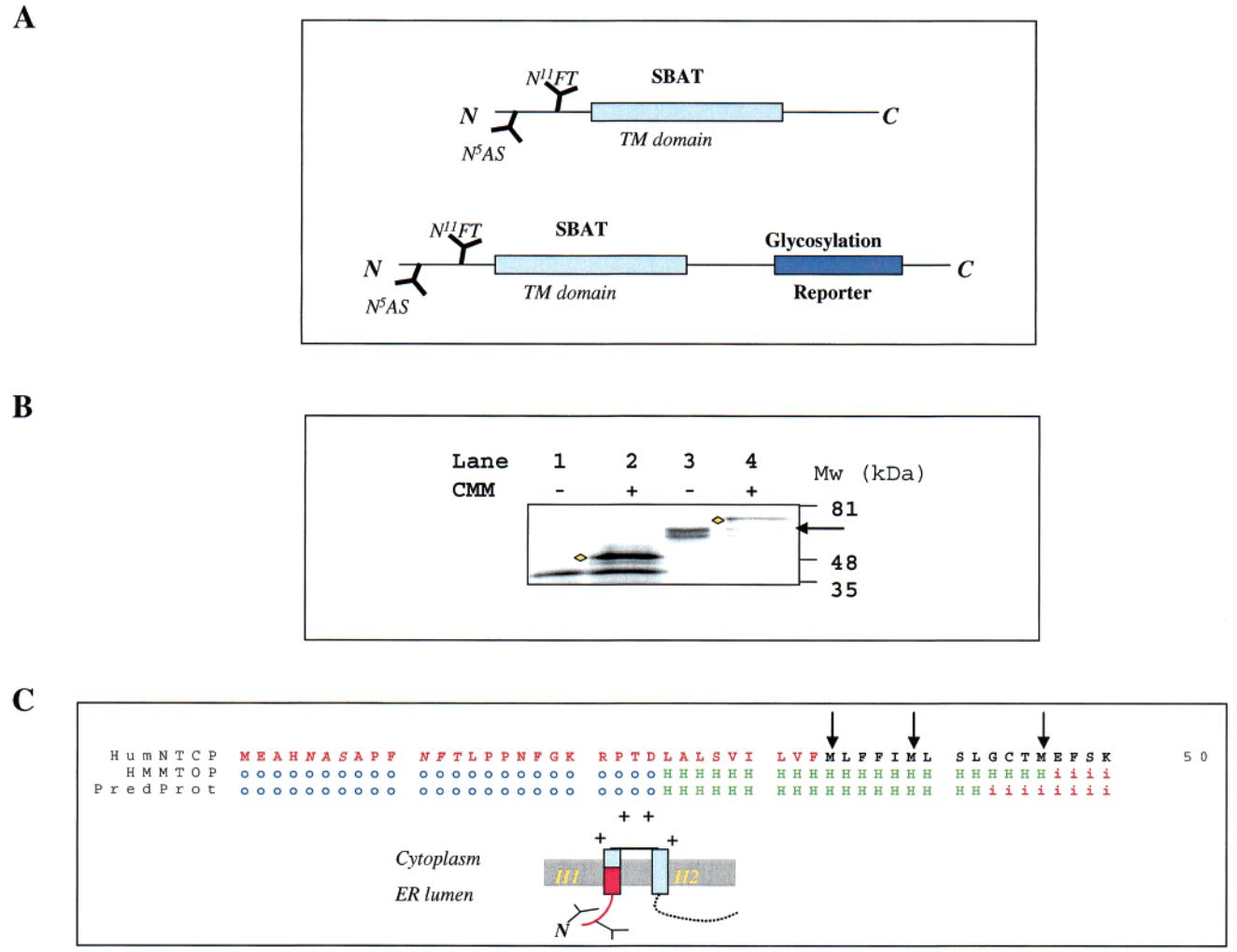

FIGURE 2: Positions for the N- and C-termini. In panel A, the constructs for investigating the cis or trans arrangement of the amino and carboxy termini is outlined. The wild-type transporter is expressed with or without a glycosylation reporter fused to the C-terminus. Phosphorimages of SDS gels separating the products from in vitro translation of these constructs in the presence or absence of canine ER microsomes (CMM) are shown in panel B. The yellow symbol designates glycosylated protein. Lanes 1 and 2 represent translation of the transporter in the absence and presence of microsomes, respectively, and lanes 3 and 4 represent translation of the transporter inserted between the H,K-ATPase sequences in the absence and presence of microsomes, respectively. A downstream initiation product is indicated with an arrow along the phosphorimage, and the putative initiation sites are assigned in panel C along with a schematic drawing of the effect on the topology of the N-terminus.

Subsequent alkaline extraction of the microsomes confirmed the presence of membrane insertion of the nonglycosylated products (Figure 4B, lanes 4, 6, 10, 12, 14, 16, and 20). Neither H3 nor H8 demonstrated stop transfer activity in the HK M1 construct as also seen with the negative control sequence (Ct) consisting of part of the hydrophilic C-terminus (Figure 4B, lanes 8, 18, and 22). Thus, the same set of hydrophobic segments that acted as signal anchor sequences are also shown to be efficient stop transfer signals. The analysis of individual segments is therefore compatible with a membrane domain containing seven transmembrane segments. This topography does not account for an N-linked glycosylation site at Asn103 that is not utilized, although in principle it should be accessible to the OST (*24*). The relatively short loop between M4 and M5 amino acids 173−192 contains five uncompensated positive charges and yet is placed in an exoplasmic position. Hence, a more detailed examination of the membrane domain of this protein was undertaken.

Stop Transfer Properties of H3 Expressed in the Context of Upstream Native Sequence. It has been demonstrated that some polytopic membrane proteins require folding within the translocon prior to membrane insertion and, therefore, do not follow a simple sequential mechanism for the lateral transfer of hydrophobic segments into the lipid bilayer (*19*). To test for the importance of the neighboring native sequence for the membrane integration of H3, a series of HK M0 constructs starting with the wild-type upstream signal anchor segment, H2, was made (Figure 5A). In vitro translation in the presence of ER microsomes with subsequent alkaline extraction demonstrated that translocation of the reporter sequence initiated by SBAT H2 (glycosylated product marked with an arrow) was not blocked by adding only part of the downstream H3 sequence (Figure 5B, lanes 3 and 4). In contrast, expression of both the longer HK M0−H2−H3 and H2−H3L constructs in the presence of microsomes showed a large reduction in the amount of glycosylated product, paralleling a marked increase in the nonglycosylated fraction recovered after alkaline extraction (Figure 5B, lanes 5−8). The fraction of nonglycosylated product was significantly higher than in the negative control where microsomes were added post-translationally and extracted along with the

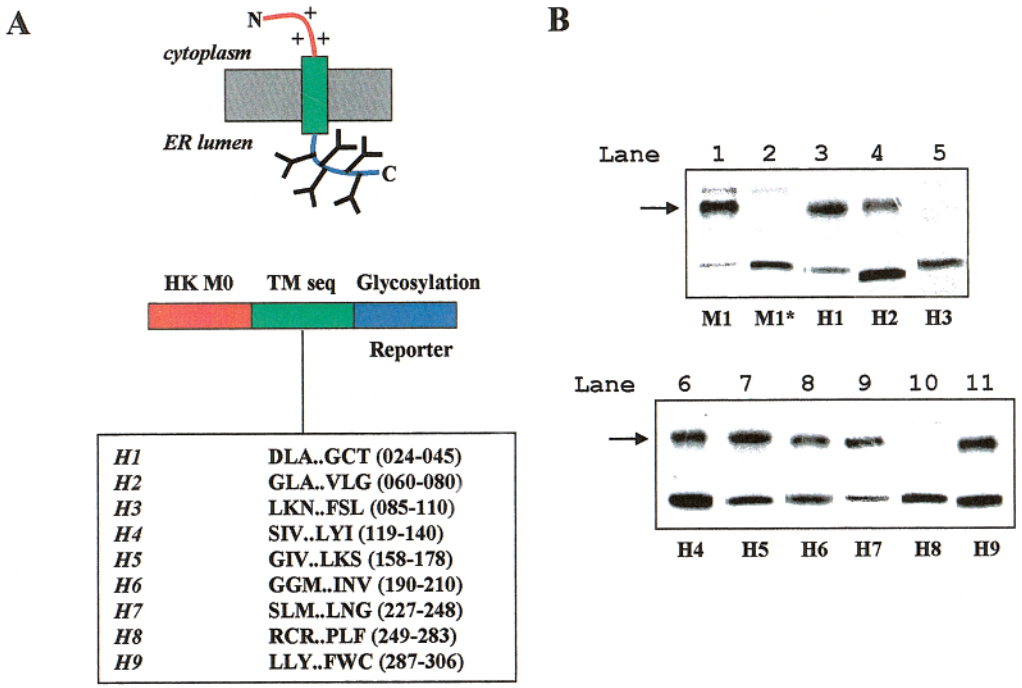

FIGURE 3: Signal anchor properties of putative transmembrane regions in the human basolateral SBAT. The tested sequences are summarized in panel A in context with a representation of the HK M0 construct. Panel B shows the SDS gel phosphorimages from translation of the HK M0 fusion proteins in the presence of canine ER microsomal membranes. The arrows denote the glycosylated product, and the asterisk in lane 2 marks the HK M1 product that has been deglycosylated post-translation using PNGase F.

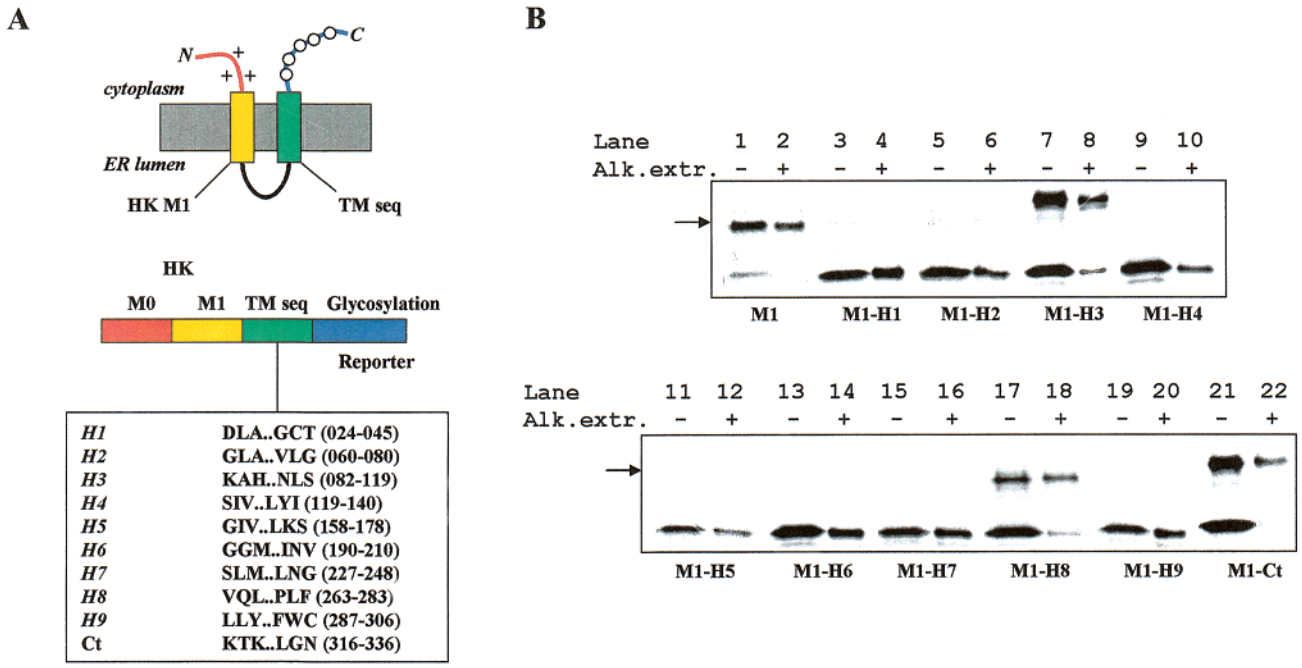

FIGURE 4: Stop transfer (ST) activities for the predicted hydrophobic regions. The top part of panel A displays the nonglycosylated product expected in the case of ST activity in the transporter sequence. All the tested sequences, including a negative control, Ct, are shown in the context with the HK M1 construct. Panel B shows the phosphorimages of SDS gels separating products from in vitro translations of the HK M1 constructs in the presence of canine ER microsomal membranes (CMM). After the translation/translocation reaction, all the products were subject to alkaline extraction to probe for membrane insertion of the nonglycosylated product.

translocation products (data not shown). Thus, while H3 exhibits no detectable stop transfer activity in the HK M1 vector as previously shown in lanes 7 and 8 of Figure 4B, it is able to block translocation of the reporter sequence when expressed downstream of the native H2 signal anchor segment as shown in Figure 5. Hence, in the context of the natural sequence, H3 is able to act as a stop transfer sequence.

Topography of the Natural Glycosylation Sites in the H3 Region Using C-Terminal Truncated Constructs. The human basolateral SBAT has five consensus sites for N-linked glycosylation, two of which are utilized in vivo, N5AS and N11FT (confirmed by site-directed deletions as shown later in the YFP fusion protein illustrated in Figure 9B). Two sites, Asn103 and Asn117, are located in the first long exoplasmic

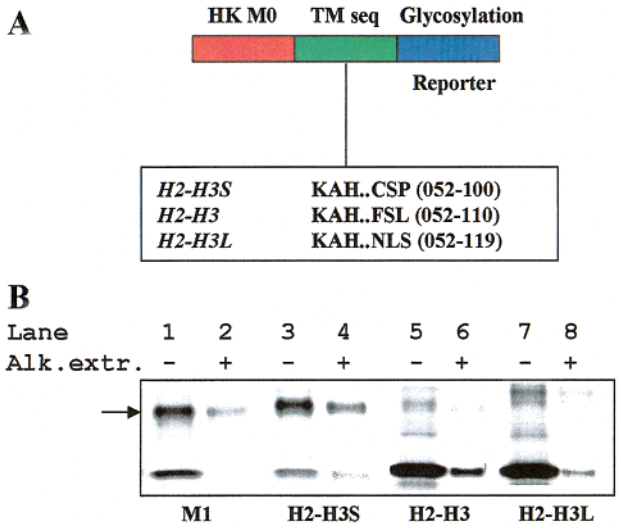

FIGURE 5: Stop transfer properties of H3 in the context of natural sequence. The ability to block translocation initiated by the human basolateral SBAT H2 was tested for H3 in the HK M0 vector as displayed in panel A. Three sequences with different C-terminal truncations in H3 were translated in the presence of canine ER microsomal membranes and subject to alkaline extraction, SDS−PAGE, and phosphorimaging. The resulting images are shown in panel B with the arrow denoting the product with glycosylation of the reporter sequence.

loop of the seven-membrane segment model or at the inner boundary of TM3 and in the TM3−TM4 cytoplasmic loop in the nine-segment model. There is also one site in the C-terminus. To test whether the consensus sequences in the H3 segment are substrates for the OST and accessed during partial biogenesis of the SBAT, C-terminal truncated constructs starting at the native N-terminus and ending at different positions in the H2−H4 region were made. The constructs were expressed either with or without the glycosylation reporter sequence fused to the C-terminal position (top panels of Figure 6A,B).

The first set of native truncated constructs lacking the β-flag was translated in the presence of ER microsomes and subjected to alkaline extraction and half the sample enzymatically deglycosylated. The samples were subsequently separated on a 10−20% Tris/Tricine gradient gel; the gel was dried and exposed in the phosphorimager. The bottom panel of Figure 6A shows the molecular mass shifts due to glycosylation of the C_t2, C_t3, and C_t4 constructs. The C_t2 translocation products have molecular masses corresponding to core glycosylation only of the two N-terminal sites (lanes 2 and 6, as marked on the figure). Expression of C_t3 in the presence of ER microsomes gives rise to additional higher-molecular mass products (lane 4). The migration shifts seen for these upper bands are compatible with utilization of all four consensus sites present in the construct and therefore demonstrate that the N103LS and N117LS sites can be substrates for the oligosaccharide transferase and are on the inside of the ER after translation of this truncated construct. However, when H4 is added as in the C_t4 construct, only the two N-terminal sites are glycosylated. Since the N-terminal boundary of H4 is more than 12−14 amino acids away from the N103LS site, and hence accessible to the oligosaccharide transferase (*24*), this observation suggests that H3 is membrane-inserted or -associated when in context with H4. This would account for the lack of glycosylation of the sites at Asn103 and Asn117.

The same C-terminal truncated transporter sequences were also tested upstream of the H,K-ATPase glycosylation flag region to monitor the orientation of the inserted products. Figure 6B shows the construction of these fusion proteins along with the data from the in vitro translation experiments. The molecular mass shifts of the products from the in vitro translations performed in the presence of ER microsomes due to glycosylation are displayed in the bottom panel of Figure 6B. The lower bands in lanes 1−3 (empty diamonds) are similar to the nonglycosylated products obtained in the absence of ER microsomes (data not shown). Translation of the C_t2-β construct yields a glycosylated product with a molecular mass corresponding to utilization of both the endogenous N-terminal sites and the five consensus sites in the reporter flag (blue square). Thus, the two first predicted segments, H1 and H2, insert as a pair with the exoplasmic location of both termini, and the flag also crosses the membrane. When the downstream H3 sequence is included, the glycosylation pattern for the translocated products indicates a split between the N_{exo}/C_{cyt} (yellow diamond) and N_{exo}/C_{exo} orientations (red square). The molecular mass of the upper band is that expected with core glycosylation of all the NXS/T sites in the C_t3-β construct, including the two sites in the H3 region; hence, the H3 region remains in the lumen of the microsome. The size of the lower band indicates N-terminal glycosylation alone. Hence, H3 is able to cross the membrane to produce this peptide and therefore is able to integrate into the membrane and act as a stop transfer sequence in contrast to when it is alone in HK translation modules. The intermediate band (filled star) has a molecular weight that fits with a product that originates from any of the downstream ATG sites described in Figure 1C, thus lacking the N-terminal sequence and consequently the corresponding glycosylation sites. In contrast to data obtained with the C_t2-β and C_t3-β constructs, the in vitro translation/translocation reaction of C_t4-β results in a product with a dominating N_{exo}/C_{cyto} orientation (Figure 6B, bottom panel, lane 3). Hence, H3 and H4 do not form an efficient membrane pair in this construct and only span the membrane once when they are translated in tandem. Since the C-terminus is cytoplasmic and, in this analysis, the end of H4 is also cytoplasmic preceded by three membrane-inserted segments, there should be an even number of membrane segments between the end of H4 and the C-terminus, as in the seven-membrane segment model.

H8 Acts as a Transmembrane Segment When Expressed with the Neighboring Native Sequence. To further investigate the insertion properties of the eighth putative membrane-spanning region, it was tested in context with the natural sequence. Two different types of constructs were made. One had native sequence starting at H6 with truncations in the C-terminal part of the transporter inserted between the HK M0 sequence and the glycosylation reporter. A second set of fusion proteins began at the native N-terminus with C-terminal truncations upstream and downstream of the H8 region and ended with the β-glycosylation flag sequence (top panels of Figure 7A,B).

As seen in Figure 7A (bottom panel, lanes 1 and 2), the in vitro translation/translocation reaction of H6 and H7 in the HK M0 vector results in a nonglycosylated product that resists alkaline extraction, compatible with insertion as a

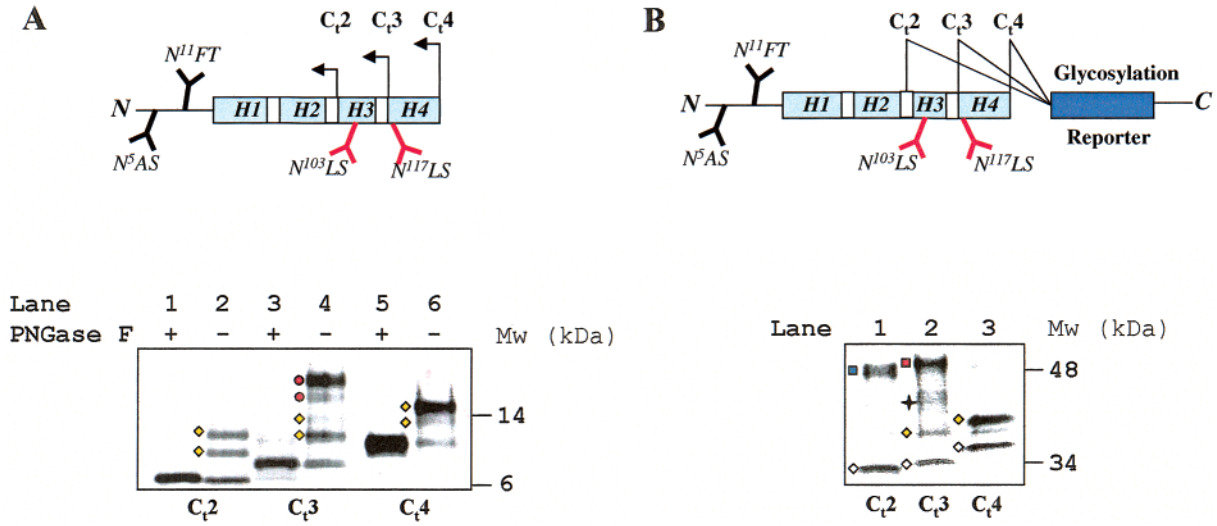

FIGURE 6: Translation of C-terminal truncations in the H2−H4 region. The positions of the endogenous N-linked glycosylation sequences, N103LS and N117LS, and the C-terminus were probed in two different constructs. The natural sequence with truncations downstream of H2, H3, and H4 was translated in the presence of canine ER microsomes and alkaline extracted, and half the product was deglycosylated using PNGase F before separation on a 10 to 20% SDS gel (Tris/Tricine). The phosphorimages from these experiments are shown in the lower half of panel A. Filled diamonds indicate glycosylation of only the N-terminal sites, N5AS and N11FT, whereas filled circles mark products with a combination of N-terminal and H3 (N103LS and N117LS) glycosylation. Panel B outlines the corresponding experiments where the glycosylation reporter has been fused onto the C-terminus of the respective constructs. The bottom half displays the phosphorimages of SDS gels separating the products from translation of these constructs in the presence of microsomes. The empty diamond denotes the nonglycosylated product. Red filled squares mark glycosylation of both the N-terminal sites and the reporter sequence. Blue filled squares indicate glycosylation of N-terminal, H3, and reporter glycosylation sites. The filled star marks the product from downstream initiation.

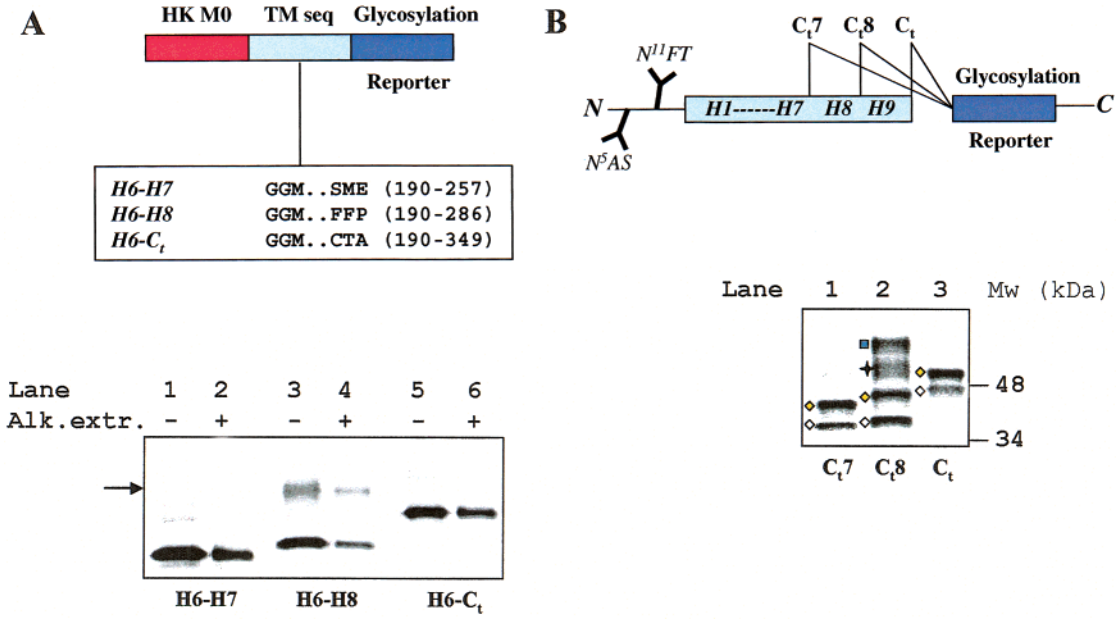

FIGURE 7: Membrane insertion properties of H8 along with the neighboring natural sequence. H8 was tested in two different constructs, either with the flanking transporter sequence in the HK M0 vector (A) or in native transporter constructs with C-terminal truncation downstream of H7 or H8 ending with the glycosylation reporter (B). In panel A, the sequences expressed in the HK M0 construct are displayed along with the phosphorimage of the products from the translation/translocation reactions separated by SDS−PAGE. We tested for membrane insertion of nonglycosylated products by alkaline extraction. The C-terminal truncated fusion proteins translated in the presence of canine ER microsomes were subjected to SDS−PAGE, and the resulting phosphorimages are outlined in the lower half of panel B. Empty and filled diamonds indicate the nonglycosylated product and glycosylation of only the N-terminal sites, N5AS and N11FT, respectively, whereas filled squares mark glycosylation of both the N-terminal sites and the reporter sequence. The product from downstream initiation is indicated with a filled star.

transmembrane pair as expected from their individual properties as signal anchor stop transfer sequences. Expression in the absence of ER microsomes yields a product with the same molecular mass, thus confirming the cytoplasmic location of the β-flag (data not shown). This cytoplasmic location of the H7 C-terminus is also found with translation of the construct starting at the native N-terminus with the glycosylation reporter fused to a position downstream of H7 in the presence of ER microsomes (Figure 7B, bottom panel, lane 1). The molecular mass shift observed is that expected for a glycosylated product with the N_{exo}/C_{cyt} orientation (yellow diamond).

In contrast, the in vitro translation/translocation reaction of HK M0−H6−H8 and C_t8 constructs both produce proteins

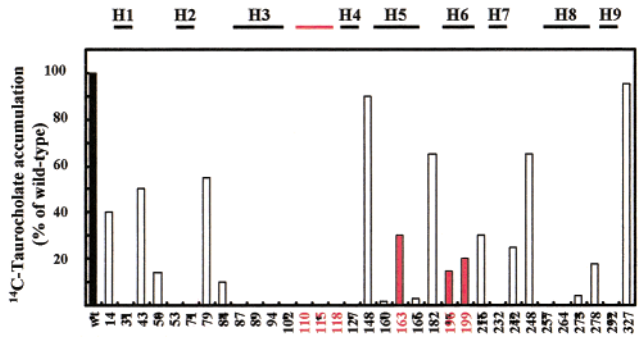

FIGURE 8: Alanine insertion scanning of the human basolateral SBAT. Site-directed alanine insertions were made at the positions placed in a seven- or nine-transmembrane segment model as illustrated at the top. The insertion mutants were transiently expressed in HEK 293 cells along with the wild-type transporter, and uptake was assessed as the level of accumulation of 50 μM [^{14}C]taurocholate. In the bottom panel, transport activities are displayed as the percentage of wild-type taurocholate uptake. Insertions resulting in activity not in agreement with the calculated predictions are indicated in red.

with partial glycosylation of the C-terminal reporter sequence (Figure 7A, bottom panel, lanes 3 and 4, and Figure 7B, bottom panel, lane 2). Although the data in both experiments reflect a mixed glycosylation pattern, H8 clearly acts as a signal anchor in part of the expressed product. Therefore, even though the eighth predicted sequence does not behave as an independent membrane-inserted segment, it is capable of integrating into the lipid bilayer when expressed downstream of the native sequence.

When the constructs are extended to have the reporter sequence fused onto the native C-terminus, no molecular mass shifts corresponding to glycosylation of the β-flag are seen (Figure 7A, bottom panel, lanes 5 and 6, and Figure 7B, bottom panel, lane 3). This is consistent with the known cytoplasmic location of the transporter C-terminus and the strong insertion properties of H9 seen in experiments using the HK M0 and HK M1 vectors. Moreover, the exoplasmic location of the native amino terminus and the cytoplasmic position of the C-terminus downstream of both the seventh putative transmembrane and the full-length transporter confirm an odd number of transmembrane segments between the N-terminus and the end of H7 as well as an even number of transmembrane segments between H7 and the carboxy terminus. When the insertion properties observed for H8 in this set of experiments are considered, it is more likely that the H8−H9 region folds as a transmembrane pair or that H8 is membrane-associated rather than part of a large cytoplasmic loop.

Alanine Insertion Scanning. The addition of a single alanine into membrane α-helical segments will result in a 120° rotation of the helix after the insertion and loss of critical contact points. Similarly, insertion of a single alanine into a β-sheet structure will distort lateral interactions in the β-sheet after the insertion. An insertion mutagenesis protocol was used to map regions of the transporter containing ordered regions. Systematic site-directed alanine insertions were made throughout the primary sequence and the mutant transporters expressed in HEK 293 cells for functional assay and plasma membrane targeting analysis using antibody staining of the mutant protein or visualization of YFP fusion proteins with confocal microscopy.

When the various alanine insertion mutants were analyzed for transport activity and glycosylation, the results depicted in Figure 8 were obtained. Displayed is the taurocholate transport as a percentage of wild-type activity with the different insertion mutants. The sites of alanine insertion are indicated on the seven (left)- and nine (right)-segment models at the top of the figure.

Addition of alanines in the N-terminal or C-terminal region results in the retention of transport activity, whereas insertion of alanine into the first hydrophobic sequence results in a loss of activity. Insertion of alanine at the cytoplasmic end of the first hydrophobic segment results in the retention of activity. Alanine insertion into any of the hydrophobic regions resulted in a loss of activity except for alanine insertion at position 163 in the middle of putative membrane-spanning segment H5 and positions 196 and 199 in putative H6. As found with proteins such as the SERCA Ca ATPase, there can be local unwinding of a membrane helix, and hence, this could explain a lesser effect of Ala insertion into such a location (*21*).

In general, alanine insertion into extramembranal domains allowed retention of transport activity. An exception to this was seen with alanine insertions into the region between positions 108 and 119, predicted to be exoplasmic in the seven-segment model and to be cytoplasmic between H3 and H4 in the nine-segment model. As discussed below, these alanine additions result in a loss of trafficking to the plasma membrane.

In contrast to the alanine insertions, mutation of Ser109 to Ala or Lys113 or Ser119 to Ala results in a functional transporter. However, mutation of Asp115 to Ala resulted in loss of transporter activity. Addition of alanine in this region inactivates transport, and substitution with alanine in general results in a functional transporter. This loop appears to be very close to the plasma membrane surface. Thus, despite the hydrophilicity and the high probability that this sequence is an extramembranal loop, it has strict structural properties since addition of a single amino acid results in a loss of trafficking.

Confocal Analysis of Plasma Membrane Trafficking. HEK 293 cells coexpressing the alanine insertion mutants along with a plasma membrane marker, CFP-Mem (palmitoylated CFP fusion protein), were immunostained using an antibody raised against the human basolateral SBAT C-terminus. Figure 9a shows the images recorded after immunostaining of the wild-type transporter (A−C) along with two alanine insertion mutants, Ala94 (D−F) and Ala278 (G−I). Panels A, D, and G display the rhodamine channel (antibody); panels B, E, and H show the channel for the cyan fluorescent protein, and panels C, F, and I are the merged images. Both the wild-type and Ala278 transporters colocalize in the plasma membrane as seen with wild-type protein, while the H3 mutant, Ala94, exhibits ER retention and little or no overlap with the CFP-Mem plasma membrane probe. These results are representative of the data obtained using the antibody generated against the liver bile acid transporter.

Further observations were made using the YFP fusion protein by determining the presence of core or full glycosylation of the expressed protein. The YFP fusion protein was fully transport competent, whereas the mutant lacking N-terminal glycosylation had 60% of wild-type activity (data not shown). Western analysis showed that transient expres-

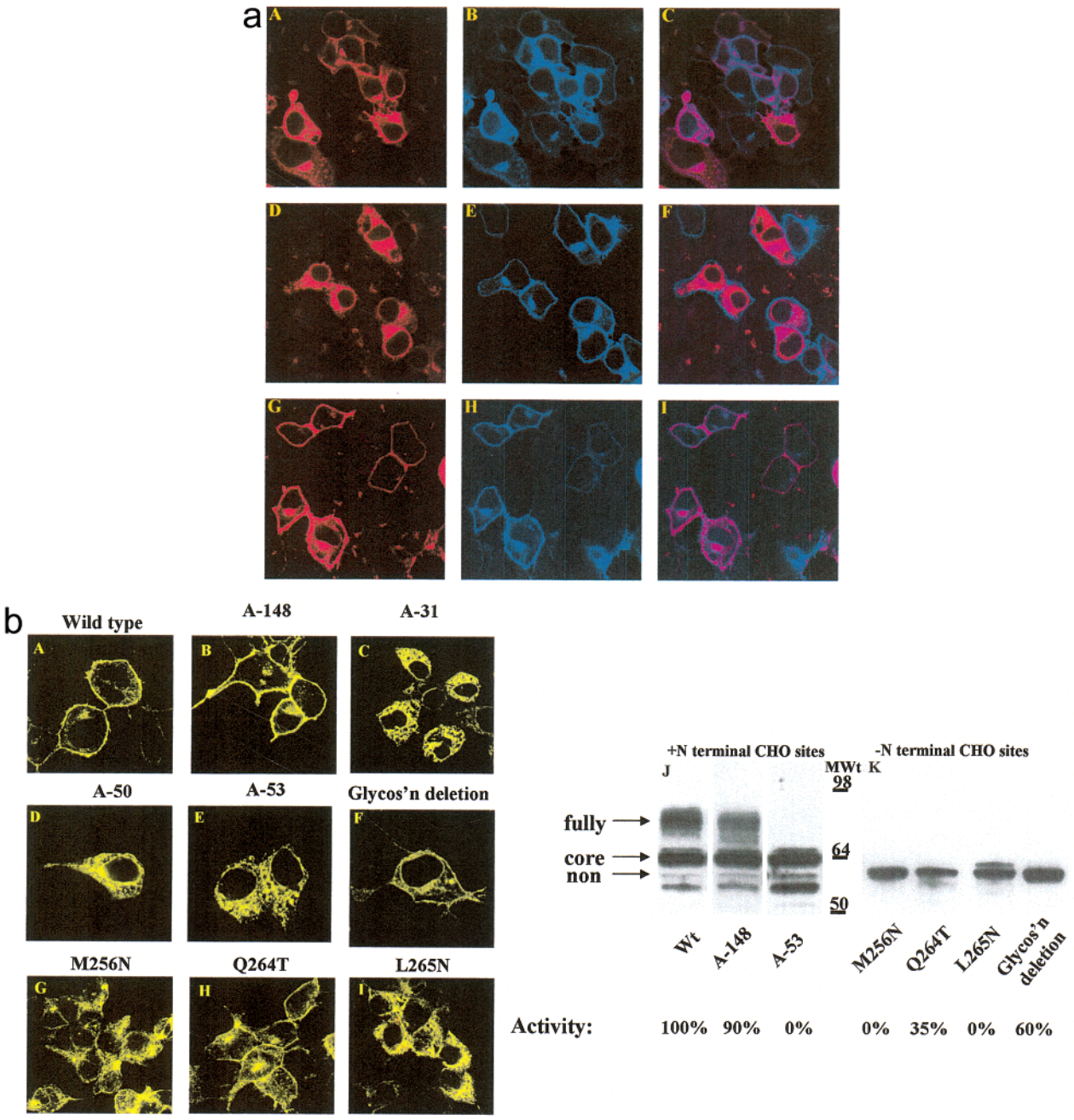

FIGURE 9: (a) Immunostaining of the human basolateral SBAT and alanine insertion mutants expressed in HEK 293 cells. HEK 293 cells were cotransfected with the wild-type or alanine constructs and the plasma membrane probe, CFP-Mem. The cells were fixed and immunostained 24 h post-translation using an anti-Hum NTCP antibody as a primary antibody and a goat anti-rabbit IgG (H+L) F(ab′)2 fragment TRITC as a secondary antibody. Confocal images were recorded with channel settings for rhodamine and CFP. Panels A, D, and G display the rhodamine channel for the wild type, Ala94, and Ala278, respectively, while panels B, E, and H show the corresponding CFP images. Merged channels for the same set of constructs are displayed in panels C, F, and I. (b) Expression of wild-type YFP fusion and glycosylation negative protein with alanine and glycosylation mutants in HEK 293 cells. The YFP fusion proteins (wild type or N-terminal N-targeted mutations, and Ala53 and -148 insertion mutants) were expressed in HEK 293 cells, and the crude membrane fraction was separated by SDS−PAGE as described in detail in Experimental Procedures for Western analysis using an anti-GFP antibody. Cells were transfected with the wild-type YFP fusion and mutant YFP fusion protein vectors and viewed using confocal microscopy as described in detail in Experimental Procedures. The plasma membrane expression of the N-terminal glycosylated YFP fusion protein is shown in panels A−E with plasma membrane expression of the wild type and Ala148 insertion mutant (A and B, respectively) and no plasma membrane expression of the Ala31, Ala50, and Ala53 insertion mutants (C−E, respectively). Western blotting with an anti-GFP antibody shows that the wild type and Ala148 are expressed as both fully glycosylated and core glycosylated protein with a minor component that is nonglycosylated. Both of these fusion proteins catalyze bile salt uptake The other Ala insertion mutants show only core glycosylation (panel J) and no transport activity. The mutant where the two N-linked sites were removed is expressed as a nonglycosylated band (panel K) but is localized to the plasma membrane (panel F) as is the glycosylation insertion mutant Q264T (panel G), and both of these exhibit transport activity. The M256N and L265N mutants are not membrane-expressed and are inactive in transport (panels H and I, respectively).

sion of the wild-type YFP fusion protein resulted in a mixture of core [approximately 65 kDa, as expected for the fusion of the human core glycosylated transporter (41 kDa) and the YFP protein (24 kDa)] and fully glycosylated protein [approximately 75 kDa, as expected for human fully glycosylated transporter (51 kDa) and the YFP protein (24 kDa)].

Table 1: Summary of Plasma Membrane Trafficking of the Alanine Insertions in the Human Basolateral Bile Acid Transporter[a]

Ala insertion position	TCA uptake (% of control)	cell surface expression	Ala insertion position	TCA uptake (% of control)	cell surface expression
14	50	yes	160	2	yes
31	0	no	163	30	yes
43	50	yes	166	3	yes
50	14	no	182	65	yes
53	0	no	196	15	yes
71	0	no	199	20	yes
79	55	yes	216	30	yes
84	10	no	232	0	no
87	17	no	242	16	no
89	0	no	248	65	yes
94	0	no	257	0	no
102	0	yes	264	0	no
110	0	no	273	4	yes
115	0	no	278	18	yes
118	0	no	292	0	no
127	0	no	327	95	yes
148	90	yes			

[a] Uptake activities are shown along with the cell surface expression as analyzed by immunostaining. Cell surface expression is defined as colocalization with the plasma membrane probe, CFP-Mem, or YFP fusion protein localization.

The mutant lacking the two N-terminal glycosylation sites when expressed as a nonglycosylated YFP fusion protein had the expected MWt of 57−58 kDa. Digestion of the wild-type protein with endoglycosidase F produced a pattern identical to that of the glycosylation-less mutant. Expression of the alanine insertion mutant at position 53 resulted in a protein pattern showing only core glycosylation, whereas the Ala insertion mutant at position 148 having 90% of wild-type activity showed both core and full glycosylation. This mutant was expressed on the plasma membrane. The Ala31, -50, and -53 inserts in the YFP construct did not access the plasma membrane, as shown in Figure 9b, whereas the glycosylation-less mutant protein did move to the plasma membrane, as expected from its transport competence, despite the absence of core and full glycosylation. A possible explanation for the cell surface expression of the latter mutant may be its lack of recognition by glycosylation-dependent chaperones involved in quality control during Golgi transit to the plasma membrane.

Although the H3 and H8 mutants, with Ala insertions in positions 102 and 273, as well as Ala160 and Ala166 mutants of the H5 were shuttled almost entirely to the plasma membrane, the transport activity of these mutants was completely lost. These alanine insertions therefore disturb the structure−function properties of the bile acid transporter, indicating that the H3 and H8 regions may play an important role in the transport activity of the protein.

A number of the inactive mutants exhibit a marked retention in the ER of the HEK 293 cells. Among those are the H3−H4 loop mutants, Ala110, -115, and -118, as well as the H3 and H8 mutants, Ala94 and Ala257 and -264. The inability to be shuttled to the plasma membrane seen with some of the insertion mutants is likely to be caused either by a loss of trafficking signal or by misfolding followed by abnormal glycosylation resulting in a subsequent failure to transit the quality control machinery in the Golgi (*20*).

Table 1 summarizes the data from immunostaining of all the alanine insertion mutants along with their bile acid transport activities. As expected, all transporters showing taurocholate uptake activity are also shuttled to the plasma membrane. No activity was observed unless plasma membrane trafficking was present.

Glycosylation Insertion Mutants. All of the glycosylation insertion mutants discussed here were expressed as nonglycosylated YFP fusion proteins (MWt = 57−58 kDa). Substitutions in the H3 region (E89S and G97N) resulted in a loss of transport activity. Both of these mutants exhibited intracellular accumulation of the protein and no plasma membrane expression when expressed in HEK 293 cells (confocal data). None of the mutants showed any band due to additional glycosylation when analyzed by Western blotting.

The H8 glycosylation insertion mutants, R249T, Q264T, and A273S, were functional with 70, 60, and 95% of control uptake in comparison with the glycosylation-less mutant as background, and 50, 35, and 60% of wild-type activity, respectively. Confocal microscopy analysis showed that R249T, Q264T (Figure 9b, panel G), and A273S mutants were localized almost entirely on the plasma membrane, consistent with their functional activity.

The mutations M256N (Figure 9b, panel G) and L265N (Figure 9b, panel I) resulted in complete loss of transport activity of the protein, and these mutants were not located in the membrane but retained in the endoplasmic reticulum. Hence, these mutations caused functional and structural disruption. Of these two mutants, only the L265N mutation showed a higher-MWt band accounting for ∼10% of the expressed protein (MWt = 64−65 kDa) as shown in panel K of Figure 9b. Treatment with PNGase F removed the extra band, confirming that this band represents glycosylated protein. This partial glycosylation reflects a small fraction of the misfolded protein.

Figure 10 illustrates the data obtained schematically showing the seven-transmembrane segment model, the same model with two membrane-associated helices, and the nine-transmembrane segment model. The data presented are not compatible with the seven-transmembrane segment model with long exoplasmic loops.

DISCUSSION

The sodium/bile acid cotransporters are part of the transport components of the enterohepatic circulation in mammals. Their most important physiological function is in the reabsorption of bile acids from the small intestine and the portal blood. Interruption of the enterohepatic circulation of bile acids in the small intestine causes compensatory conversion of cholesterol into bile acid in the liver (*22*). Drug intervention that prevents reabsorption of bile acids by the ileal SBAT may provide a new promising therapy for lowering of cholesterol (*23*).

A recent investigation of the 2D structure of the human apical SBAT (ileal) by membrane insertion scanning generated data perhaps suggestive of nine rather than seven membrane-integrated regions (*11*). To complement the findings for the human apical transporter, we have now investigated the membrane insertion properties of the homologous human basolateral (liver) transporter in considerably more detail. The two transporters are ∼35% identical but have similar hydrophobicity profiles and are expected to share the same transmembrane core structure.

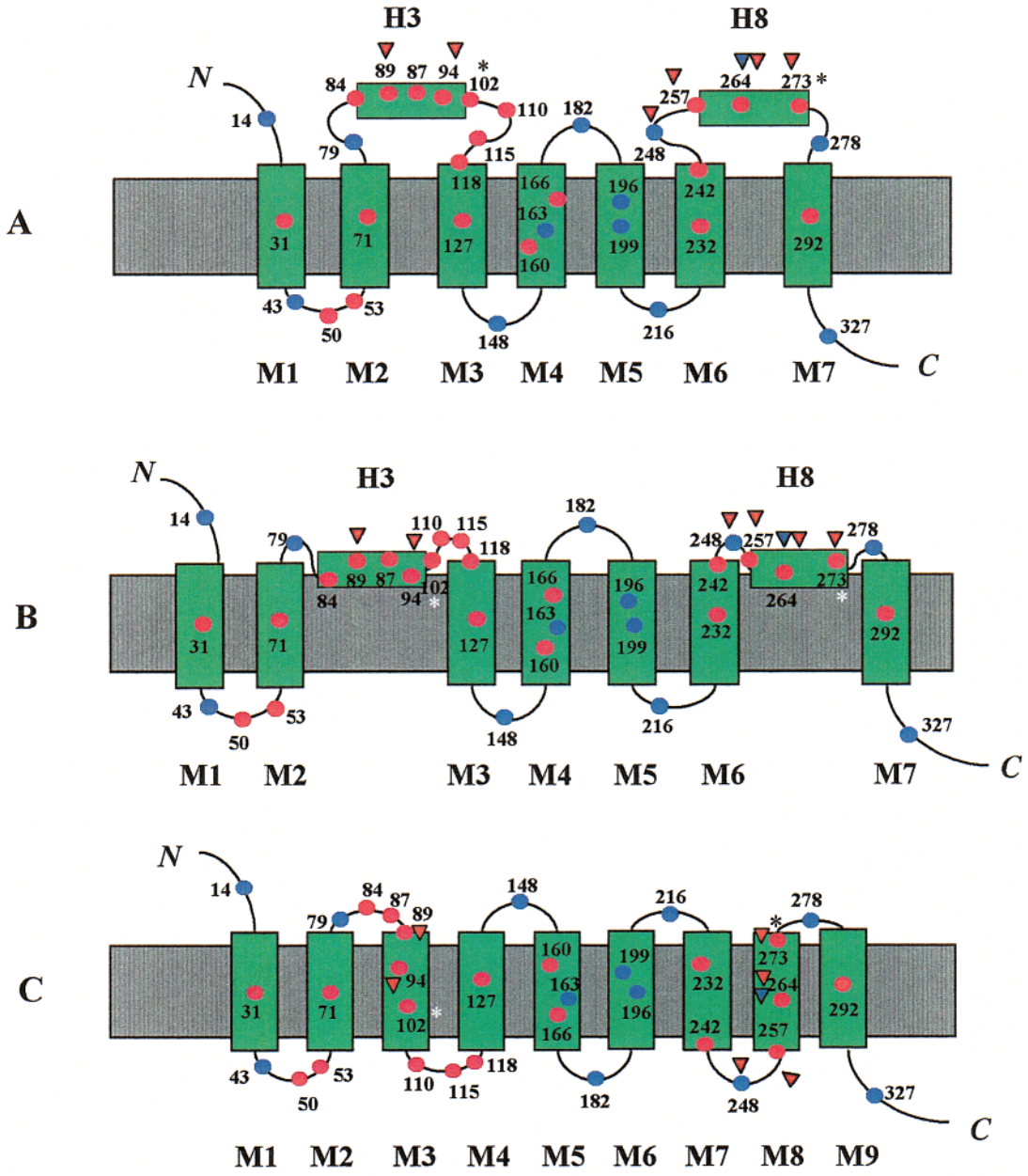

FIGURE 10: Topography models of the liver bile acid transporter incorporating results of alanine insertion and glycosylation insertion mutations. The results of alanine insertion scanning are displayed on a model of the transporter with either seven transmembrane segments (A), seven transmembrane segments with two membrane-associated regions (B), or nine transmembrane segments (C) showing alanine insertions that retained activity (blue circles) and that lost activity (red circles) and also trafficking. The asterisks denote two alanine insertions where trafficking is retained but function is lost. Arrowheads indicate mutations that introduce glycosylation sites; red are inactive, and the blue ones show that the mutation at position 264 retains activity. The data exclude the model illustrated in panel A, but are compatible with the predictions of the other two models, particularly model C, a nine-membrane segment model.

When the hydrophobic stretches of the protein were tested for signal anchor activity in the HK M0 vector, seven regions all acted as independent insertion sequences while the more amphipathic H3 and H8 segments failed to initiate translocation of the glycosylation reporter. An identical result is seen when the same set of sequences is expressed in the HK M1 vector for analysis of stop transfer properties. All regions, except the third and the eighth, are capable of blocking the import of the β-flag into the ER microsomes. Thus, despite the lack of topogenic information for prefolding and orientation, seven putative transmembrane segments insert as both signal anchor or stop transfer sequences.

Moreover, the addition of an alanine into H1, H2, H4, H5, H7, and H9 resulted in partial or total loss of taurocholate uptake in the HEK 293 cells. Thus, signal anchor/stop transfer scanning and alanine insertion mutagenesis identify H1, H2, H4−H7, and H9 as being membrane-integrated with secondary structures sensitive to positional displacements. These results reflect the consensus for at least seven membrane-spanning segments.

To further investigate the H3 and H8 regions, various constructs were made encoding these segments in the context of the neighboring natural sequence. The H3 sequence was shown to act as a partial stop transfer sequence when placed downstream of the native H2 in two independent constructs. In addition, two consensus sites for N-linked glycosylation, N103LS and N117LS, were demonstrated to be utilized when expressed in a C-terminal truncated construct ending after

H3, but not when the protein was extended to include H4. Thus, these two sites are competent endogenous glycosylation reporter sites as substrates for the oligosaccharide transferase, but are not exposed to the enzyme in either the native protein or the C-terminal truncated protein ending downstream of H4. The minimal distance for the OST to access a NXS/T site is approximately 12 amino acids away from a preceding TM segment and 14 amino acids distant from a downstream membrane-embedded sequence (*25*). Therefore, the disappearance of the higher level of glycosylation in the C_t4 construct suggests insertion of H3 and/or H4 rather than a long exoplasmic loop structure.

Alanine insertion into the rather short connecting loop between H3 and H4 results in an inactive protein due to loss of trafficking to the plasma membrane measured by confocal microscopy (Ala110, -115, and -118). Since this part of the transporter is likely (probability of 0.8) to be an inside loop region, if the nine-membrane segment model is correct, a possible explanation for the lack of cell surface expression is the direct involvement of this sequence as a trafficking signal. Alanine mutations rather than insertions did not in general disturb function. Hence, distortion of the sequence by alanine addition was fatal for trafficking of the protein, whereas alanine mutations were not, with the exception of substitution of Asp114.

Alanine insertions, Ala94 and Ala102, in H3 abolished taurocholate uptake in transfected HEK 293 cells. The Ala102 mutant was present on the plasma membrane, whereas Ala94 failed to be trafficked to the membrane. The loss of functional activity with normal trafficking due to Ala102 insertion is probably due to a structural effect in an α-helix involved in transport or in critical interactions between neighboring α-helices. The Ala94 insertion presumably perturbs the helix of H3 of the nine-membrane-associated segment model, preventing plasma membrane localization. This is consistent with membrane localization and not with exoplasmic localization of H3 since this would not be expected to interfere with trafficking. This is additional evidence that H3 is a part of the membrane-embedded domain rather than part of a large, flexible exoplasmic loop.

The other amphipathic sequence, H8, was also demonstrated to be able to be membrane-integrated when expressed in the context with the neighboring native sequence. It acts as a signal anchor sequence in both the C-terminal truncated fusion protein and the H,K vector. Furthermore, alanine insertion mutagenesis in this region results in a complete loss of taurocholate uptake in HEK 293 cells. The loss of transport activity correlates with the loss of expression on the plasma membrane for Ala257 and Ala264 mutants. Disruption of the helix predicted to be H8 prevents normal trafficking. This region is predicted to be exoplasmic in the seven-segment model and therefore should not be involved in plasma membrane trafficking since this depends on cytoplasmic sequence interactions as also pointed out above in terms of the effect of insertions in the first long exoplasmic loop postulated in the seven-membrane segment model. Alanine insertion at position 273 affects only activity, but not delivery of the protein to the plasma membrane, indicating the possibility of disruption of interactions between α-helices that are important for the functional activity or a direct effect on the transport mechanism. These results suggest a role for H8 as a membrane-associated sequence in the native transporter.

Additional support for the membrane association of H3 comes from generating sequences for N-linked glycosylation at two different positions, E89S and G97N, within this domain. None of these sites were glycosylated when the mutants were expressed in HEK 293 cells, and these mutations resulted in a loss of function. Three glycosylation sequences engineered into the second putative exoplasmic loop, R249T, Q264T, and A273S, showed function and plasma membrane trafficking. Two others showed complete loss of function and plasma membrane expression, M256N and L265N. Only the L265N mutant exhibited any glycosylation, and this did not exceed 10% of total protein expressed. The lack of glycosylation at several N-linked sites, whether natural or engineered, suggests that the two long exoplasmic loops in the seven-transmembrane segment model are not exposed to the oligosaccharide transferase. This is indicative of retention of these structures in the translocon during translation of downstream sequences.

Figure 10 summarizes the alanine and glycosylation site insertion data placed on a seven-membrane segment model without or with two membrane re-entrant loops or a nine-transmembrane segment model of the human liver bile acid transporter. The mutations where activity and trafficking are retained are shown in blue; inactive mutations in red. Asterisks indicate insertions with disrupted function but retained appearance at the plasma membrane. Membrane segments in the seven- and nine-segment model are indicated as M1, M2, etc. H3 and H8 are placed in association with the membrane in the seven-transmembrane segment model containing two associated membrane segments. Inspection of the figure shows that all changes in H3 and H8 result in inactivation of transport and of plasma membrane expression. These data are contrary to what is expected of the seven-transmembrane segment model shown in Figure 10A. These are also somewhat unexpected of the membrane-associated sequences of Figure 10B and are most compatible with membrane insertion of H3 and H8 as illustrated in the nine-segment model shown in Figure 10C.

Polytopic membrane proteins that function in ion movements are a priori expected to have polar ligands for the binding, occlusion, and translocation of charged molecules. A common property of ion transporters and channels is the presence of polar or charged residues within the transmembrane domain of the protein that participate in transport. Candidate regions for this function in the sodium-dependent bile acid transporters are the highly conserved and amphipathic H3 and H8 sequences.

A number of independent observations have suggested that the H8−H9 domain interacts with bile acids in a sodium-dependent manner. Labeling of the rabbit apical SBAT using radioactive photoaffinity derivatives of bile acids identified a 56−67-amino acid C-terminal fragment as one site of interaction between the substrate and transporter (*26*). Furthermore, inactivation of the SBATs by side-specific methanethiosulfonates is due to alkylation of the fully conserved cysteine in the CSTI motif in H8. The accessibility of this site is regulated by sodium and bile acid, the former increasing the sensitivity and the latter preventing inactivation (*27*).

The SBATs have a highly conserved cluster of positive charges residing in the short loop between H5 and H6.

Positive charges surrounding the N-terminus and those residing in short loops have been shown to be particularly important in determining the orientation of membrane proteins, anchoring the polypeptide chain on the inside of the plasma membrane (*24*, *28*). A cytoplasmic location for the H5−H6 loop requires an odd number of transmembrane segments upstream and an even number of transmembrane segments downstream of this region. This is compatible with the nine-, not the seven-, transmembrane segment model, if all the seven independent signal anchor/stop transfer sequences do span the membrane since re-entrant loops do not change the position of this region.

The weight of evidence excludes the 2D structure encompassing seven membrane-spanning segments with long loops between positions 77 and 119 and positions 244 and 286. Thus, the data from insertion scanning of the human SBATs support membrane integration of all nine hydrophobic sequences. It is not clear whether they all insert as transmembrane segments or if two of them fold as lipid-buried re-entrant loops. Such re-entrant motifs have recently been identified using cysteine scanning protocols to investigate the topology of members of the glutamate transporter family (*31*−*33*). Systematic substitution of cysteines in the human basolateral SBAT, however, results in a stepwise loss of activity down to less than 10% of wild-type activity with several cysteines still to be replaced (S. Hallén and G. Sachs, unpublished data), so this approach to the bile acid transporter is not feasible.

In summary, translation scanning and alanine insertion mutagenesis both indicate an involvement of the regions containing H3 and H8 (between amino acids 77 and 119 and between amino acids 245 and 286, respectively) in membrane and structural intramolecular interactions of the transporter. The bile acid transporters therefore have a 2D structure consisting of either nine transmembrane segments or seven transmembrane segments with two re-entrant exoplasmic membrane-associated sequences.

ACKNOWLEDGMENT

We thank Dr. Paul Dawson for the kind gift of the SBAT cDNA.

REFERENCES

1. Wong, M. H., Oelkers, P., Craddock, A. L., and Dawson, P. A. (1994) *J. Biol. Chem. 269* (2), 1340−1347.
2. Christie, D. M., Dawson, P. A., Thevananther, S., and Shneider, B. L. (1996) *Am. J. Physiol. 271* (2, Part 1), G377−G385.
3. Hagenbuch, B., Stieger, B., Foguet, M., Lubbert, H., and Meier, P. J. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88* (23), 10629−10633.
4. Stengelin, S., Apel, S., Becker, W., Maier, M., Rosenberger, J., Wess, G., and Kramer, W. (1995) GenBank accession no. Z54357.
5. Stengelin, S., Becker, W., Maier, M., Noll, R., and Kramer, W. (1998) GenBank accession no. AJ131361.
6. Saeki, T., and Matoba, K. (1997) Genbank accession no. AB002693.
7. Saeki, T. (1997) GenBank accession no. AB003303.
8. Wong, M. H., Oelkers, P., and Dawson, P. A. (1995) *J. Biol. Chem. 270* (45), 27228−27234.
9. Hagenbuch, B., and Meier, P. J. (1994) *J. Clin. Invest. 93* (3), 1326−1331.
10. Dawson, P. A., and Oelkers, P. (1995) *Curr. Opin. Lipidol. 6* (2), 109−114.
11. Hallén, S., Branden, M., Dawson, P. A., and Sachs, G. (1999) *Biochemistry 38* (35), 11379−11388.
12. Rost, B., Fariselli, P., and Casadio, R. (1996) *Protein Sci. 5* (8), 1704−1718.
13. Sonnhammer, E. L., von Heijne, G., and Krogh, A. (1998) *Proc. Int. Conf. Intell. Syst. Mol. Biol. 6*, 175−182.
14. Tusnady, G. E., and Simon, I. (1998) *J. Mol. Biol. 283* (2), 489−506.
15. van Geest, M., and Lolkema, J. S. (2000) *Microbiol. Mol. Biol. Rev. 64* (1), 13−33.
16. Bamberg, K., and Sachs, G. (1994) *J. Biol. Chem. 269* (24), 16909−16919.
17. Braun, P., Persson, B., Kaback, H. R., and von Heijne, G. (1997) *J. Biol. Chem. 272* (47), 29566−29571.
18. Kozak, M. (1991) *J. Biol. Chem. 266* (30), 19867−19870.
19. Borel, A. C., and Simon, S. M. (1996) *Cell 85* (3), 379−389.
20. Ellgaard, L., and Helenius, A. (2001) *Curr. Opin. Cell Biol. 13* (4), 431−437.
21. Toyoshima, C., Nakasako, M., Nomura, H., and Ogawa, H. (2000) *Nature 405* (6787), 647−655.
22. Hofmann, A. F. (1993) in *Gastrointestinal disease. Pathophysiology, diagnosis, management* (Sleisenger, M. H., and Fordtran, J. S., Eds.) 5th ed., pp 127−150, W. B. Saunders Co., Philadelphia.
23. Kramer, W., and Wess, G. (1996) *Eur. J. Clin. Invest. 26* (9), 715−732.
24. Wallin, E., and von Heijne, G. (1998) *Protein Sci. 7* (4), 1029−1038.
25. Nilsson, I. M., and von Heijne, G. (1993) *J. Biol. Chem. 268* (8), 5798−5801.
26. Kramer, W., Girbig, F., Glombik, H., Corsiero, D., Stengelin, S., and Weyland, C. (2001) *J. Biol. Chem. 276* (38), 36020−36027.
27. Hallén, S., Fryklund, J., and Sachs, G. (2000) *Biochemistry 39* (22), 6743−6750.
28. von Heijne, G., and Gavel, Y. (1988) *Eur. J. Biochem. 174* (4), 671−678.
29. Nilsson, I., and von Heijne, G. (1998) *J. Mol. Biol. 284* (4), 1185−1189.
30. Tieleman, D. P., Shrivastava, I. H., Ulmschneider, M. R., and Sansom, M. S. (2001) *Proteins 44* (2), 63−72.
31. Grunewald, M., Bendahan, A., and Kanner, B. I. (1998) *Neuron 21* (3), 623−632.
32. Slotboom, D. J., Sobczak, I., Konings, W. N., and Lolkema, J. S. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96* (25), 14282−14287.
33. Grunewald, M., and Kanner, B. I. (2000) *J. Biol. Chem. 275* (13), 9684−9689.

BI012152S