Membrane Insertion Scanning of the Human Ileal Sodium/Bile Acid Co-transporter†

Stefan Hallén,‡ Magnus Brändén,‡ Paul A. Dawson,§ and George Sachs*,‡

UCLA and Wadsworth Veterans Administration Hospital, Los Angeles, California 90073, and Department of Internal Medicine, Division of Gastroenterology, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157

Received March 9, 1999; Revised Manuscript Received May 10, 1999

ABSTRACT: Mammalian sodium-dependent bile acid transporters (SBATs) responsible for bile salt uptake across the liver sinusoidal or ileal/renal brush border membrane have been identified and share approximately 35% amino acid sequence identity. Programs for prediction of topology and localization of transmembrane helices identify eight or nine hydrophobic regions for the SBAT sequences as membrane spanning. Analysis of *N*-linked glycosylation has provided evidence for an exoplasmic N-terminus and a cytoplasmic C-terminus, indicative of an odd number of transmembrane segments. To determine the membrane topography of the human ileal SBAT (HISBAT), an *in vitro* translation/translocation protocol was employed using three different fusion protein constructs. Individual HISBAT segments were analyzed for signal anchor or stop translocation (stop transfer) activity by insertion between a cytoplasmic anchor (HK M0) or a signal anchor segment (HK M1) and a glycosylation flag (HK β). To examine consecutive HISBAT sequences, sequential hydrophobic sequences were inserted into the HK M0 vector or fusion vectors were made that included the glycosylated N-terminus of HISBAT, sequential hydrophobic sequences, and the glycosylation flag. Individual signal anchor (SA) and stop transfer (ST) properties were found for seven out of the nine predicted hydrophobic segments (H1, H2, H4, H5, H6, H7, and H9), supporting a seven transmembrane segment model. However, the H3 region was membrane inserted when translated in the context of the native HISBAT flanking sequences. Furthermore, results from translations of sequential constructs ending after H7 provided support for integration of H8. These data provide support for a SBAT transmembrane domain model with nine integrated segments with an exoplasmic N-terminus and a cytoplasmic C-terminus consistent with a recent predictive analysis of this transporter topology.

The enterohepatic circulation of bile acids in mammals involves a number of discrete steps that include transport across the ileal enterocyte and hepatocyte plasma membranes (*1*). Bile acid uptake from the intestinal lumen or hepatic sinusoidal space is mediated by sodium-dependent transporters (SBATs)[1] that have been cloned and sequenced from a number of species, including hamster (ileal), rat (ileal/liver), rabbit (ileal/liver), mouse (ileal/liver), and human (ileal/liver) (*2*−*10*).

The SBATs are polytopic integral membrane glycoproteins containing 347−362 amino acids and have an overall sequence identity of 35% as well as a limited sequence identity with the P3 housekeeping protein whose function is still unknown (*11*). However, the SBATs do not exhibit any significant similarity to the other sodium-dependent co-transporters such as the sodium/glucose co-transporter family or sodium/neurotransmitter co-transporter family (*12*, *13*).

A topography model with seven transmembrane (7 TM) regions has been suggested for this new family of sodium/substrate co-transporters (*4*, *10*, *14*). This model is based on experimental evidence for a *trans* position of the N- and C-termini (N_{exo}/C_{cyt}). However, it does not accommodate the eight or nine membrane spanning segments predicted from several topology/hydropathy analyses and violates the "positive inside" rule (*15*), originally developed for bacterial proteins but now also shown to apply to eukaryotic membrane proteins (*16*, *17*). In the 7 TM model, the second and third hydrophobic sequences are predicted to form 1 long intra-membranal segment (*14*) containing as many as 35 amino acids.

There are various approaches to study the topography of integral membrane proteins in the absence of a resolved crystal structure. *In vitro* and *in vivo* translation of constructs containing one or more transmembrane sequences (*17*−*22*), epitope localization, reaction with sided reagents, and proteolysis of the purified membrane-inserted protein (*23*) have all been used individually or in combination. Each has advantages and disadvantages, and only rarely does one method enable conclusive topographic analysis of the membrane-embedded segments of a polytopic integral membrane protein.

† Supported by USVA SMI and NIH Grants DK46917, 53462, 41301, and 17294.

* To whom correspondence should be addressed at the 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.

‡ UCLA and Wadsworth Veterans Administration Hospital.

§ Wake Forest University School of Medicine.

[1] Abbreviations: ER, endoplasmic reticulum; SA, signal anchor; ST, stop transfer; SBAT, sodium bile/acid co-transporter; HISBAT, human ileal sodium/bile acid co-transporter; H, hydrophobic segments; K−D, Kyte−Doolittle; GES, Goldman, Engelman, and Steitz; HMM, hidden Markov model; TM, transmembrane; SRP, signal recognition particle; CMM, canine microsomal membrane; PAGE, polyacrylamide gel electrophoresis; kDa, kilodalton(s); PCR, polymerase chain reaction; OST, oligosaccharide transferase.

10.1021/bi990554i CCC: $18.00 © 1999 American Chemical Society
Published on Web 08/10/1999

Based on hydropathy plots and topography predictions, we have explored the use of a cell-free system to analyze the topography of the human ileal sodium/bile acid co-transporter. Several *in vitro* translation and translocation methods were designed to investigate the membrane insertion properties of individual as well as sequential sequences of the human ileal SBAT. Insertion and orientation was probed as a molecular weight shift caused by microsomal *co-* or *post*-translational *N*-linked glycosylation in combination with a membrane extraction assay.

The ability of individual HISBAT segments to function as signal anchors was tested by insertion between a cytoplasmic anchor encompassing the first 101 amino acids of the rabbit H^+,K^+-ATPase (HK M0) α-subunit and a glycosylation flag sequence consisting of 5 *N*-linked glycosylation sites located in the C-terminal 177 amino acids of the rabbit H^+,K^+-ATPase (HK M0) *β*-subunit (*18*). Stop transfer properties of hydrophobic sequences were examined in a similar manner using the first 139 N-terminal amino acids of the rabbit H^+,K^+-ATPase α-subunit (HK M1) containing the first membrane sequence of the H^+,K^+-ATPase as a signal anchor upstream of individual predicted HISBAT TM sequences linked to the *β*-glycosylation flag. The HK M0/M1 vectors are illustrated in the upper panel of Figure 1.

A series of HISBAT truncations were also constructed that initiated at the exoplasmic and glycosylated HISBAT N-terminus and terminated with the *β*-glycosylation flag in order to examine the possible topogenic effects due to interaction between putative membrane segments. These constructs were used to supplement data from sequential HISBAT inserts in the HK M0 vector, and this vector is displayed in the lower panel of Figure 1. The results from these *in vitro* translation/translocation studies support a SBAT transmembrane domain model holding an exoplasmic N-terminus, a cytoplasmic C-terminus, and nine membrane-integrated segments.

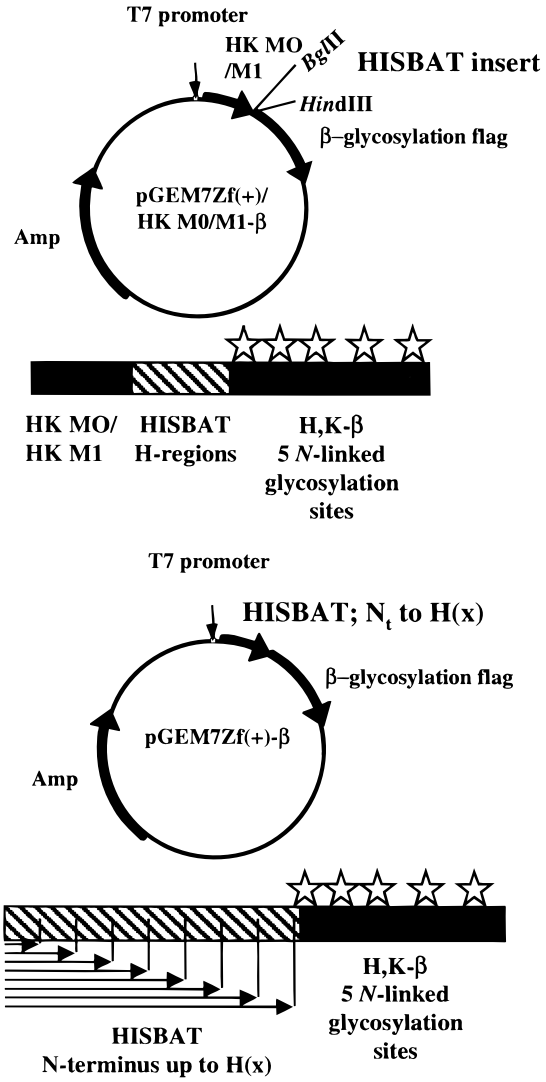

FIGURE 1: Plasmid constructs for topography investigation of the HISBAT sequences. All the plasmid constructs used are based on a modified pGEM7zf(+) vector as described (*18*). Signal anchor activity was investigated by inserting the predicted HISBAT segment between a cytoplasmic anchor (HK M0, the 101 N-terminal amino acids from the H^+,K^+-ATPase α-subunit) and a C-terminal glycosylation flag (the 177 C-terminal amino acids from the H^+,K^+-ATPase *β*-subunit) containing 5 consensus sites for *N*-linked glycosylation. In a second vector, stop transfer activity was identified by the ability of the inserted segment to block translocation/glycosylation of the *β* flag initiated by the first signal anchor sequence (HK M1) of the H^+,K^+-ATPase α-subunit (131 N-terminal amino acids). Consecutive HISBAT sequences were investigated both in the HK M0 vector and in a third type of construct starting at the natural N-terminus with truncations at desired C-terminal positions upstream of the *β*-flag (lower).

EXPERIMENTAL PROCEDURES

Analysis of Primary Sequences. The putative membrane spanning regions and sequence topography for the human ileal sodium/bile acid co-transporter were analyzed using the TopPred II program (*24*) (http://www.biokemi.su.se), the PhdTopography neural network system developed by Rost et al. (*25*) (http://www.embl-heidelberg.de/predictprotein), and the hidden Markov model recently described by Tusnády and Simon (*26*) (http://www.enzim.hu/hmmtop). The TopPred II program was used with the default settings for eukaryotic proteins, and the GES (*27*) and Kyte−Doolittle (*28*) hydrophobicity scales were applied.

Vector Construction. The vectors used for the topography investigation are summarized in Figure 1. The constructs used for signal anchor (SA) (HK M0)/stop transfer (ST) (HK M1) screening and analysis of the C-terminal truncated HISBAT sequences were all based on the pGEM7zf+(Δ*Hin*dIII) plasmid (Promega, Madison, WI) (*18*). Sequences to be investigated for SA/ST activity were cloned into a *Bgl*II/*Hin*dIII linker region between the first 101/139 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 (*18*). The plasmids encoding the C-terminal truncated HISBAT−*β* fusion sequences were constructed by excising the M0 sequence with the restriction enzymes *Xho*I/*Hin*dIII and replacing it with the indicated HISBAT sequences. The full-length natural HISBAT sequence was inserted behind the T7 promotor in the pcDNA3 expression vector from 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 included a *Bgl*II site, and the primers

for analysis of the native sequences contained a *Xho*I site as well as the Kozak -ACC- consensus (*29*) in the corresponding position. All antisense primers have a *Hin*dIII site on the 3′ end. The amplification reactions (100 μL) reactions contained 1 ng of template, pCMV-HISBAT, 1 μM of each sense and antisense primer, 200 μM of each deoxynucleotide (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/*Hin*dIII (HK M0/M1 constructs) or *Xho*I/*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 the gel purification (calf intestine phosphatase, Pharmacia). The inserts were ligated into the appropriate vectors and subsequently transformed into *E. coli* JM109 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 analyzed 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 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, plus 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 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 an 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 *N*-glycosidase F (PNGase F; Boehringer Mannheim). For deglycosylation, 6 μL of the translation reaction was dissolved in buffer, yielding a final concentration of 50 mM Tris-HCl, pH 7.8, 2 mM EDTA, 0.1% SDS, and incubated with 0.2 unit of PNGase F for 2 h at room temperature.

SDS−PAGE and Phosphor-imaging. Samples were mixed with an equal volume of 2× sample buffer containing 0.125 M Tris-HCl, pH 6.8, 4% SDS, 20% glycerol, 5% β-mercaptoethanol with bromophenol blue as marker dye and then separated on either 4−12% or 10% Tris/glycine gels from Novex (San Diego, CA). After electrophoresis, the gels were fixed/stained in 45% methanol, 10% glacial acetic acid ± 0.1% Coomassie Brilliant Blue for 20 min. The Coomassie-stained gels were destained in two steps: 30 min with 45% methanol, 10% acetic acid followed by overnight incubation in 5% methanol, 5% acetic acid. Other gels were fixed and washed in 45% methanol, 10% acetic acid overnight. The Tris/glycine gels were allowed to swell for 45 min in deionized water prior to drying. Relative molecular weights were determined using the AMBIS/RFLP scan software package calibrated with the prestained low-range molecular weight markers from Bio-Rad. Gels were exposed to phosphor-imaging screens, and radio-imaging was carried out on a Molecular Dynamics PhosphorImager.

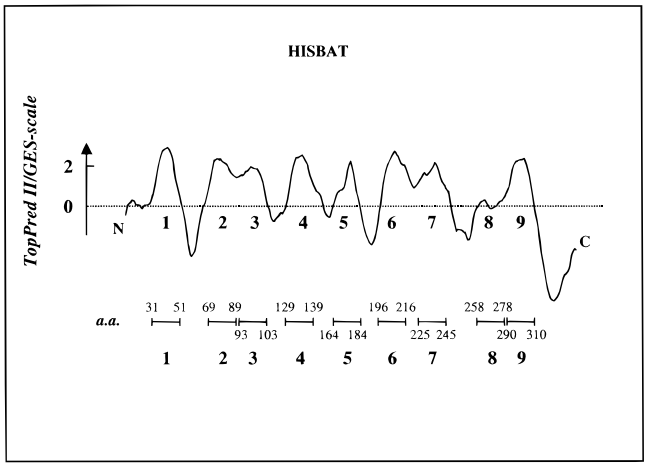

FIGURE 2: Hydropathy plot of the HISBAT sequence. The hydrophobicity profile of the HISBAT sequence calculated using a 21 amino acid window and the GES-scale (TopPred II program) with putative transmembrane regions indicated.

Buffers and Reagents. All chemicals used were of analytical or higher grade.

RESULTS

Prediction of Topography and Positions of Transmembrane Helices. Hydropathy and topography analyses of the HISBAT sequence were performed using a variety of predictive algorithms. Both the TopPred II program as well as the neural network analysis predict structures with eight transmembrane segments and exoplasmic positions for both the N- and C-termini. The *exo*-position for the N-terminus is anticipated by these algorithms due to the positive charges located just downstream of H1, while the C-terminus was assigned its orientation based on an even number of predicted TM segments. The amino acid distribution analysis performed by the hidden Markov model (HMMTOP) resulted in a predicted topography consisting of an exoplasmic N-terminus, nine transmembrane helices, and a cytoplasmic C-terminus. In this algorithm, an additional region, H8, is predicted to be membrane-integrated on the basis of amino acid composition rather than the overall hydrophobicity. Figure 2 shows the HISBAT hydropathy profile with the nine predicted segments indicated, along with the amino acids in these segments. Figure 4 (upper panel) displays the amino acid sequence with the consensus predictions of the applied analysis programs boxed in dark gray and amino acids predicted to be within the lipid bilayer by any of the algorithms boxed in light gray.

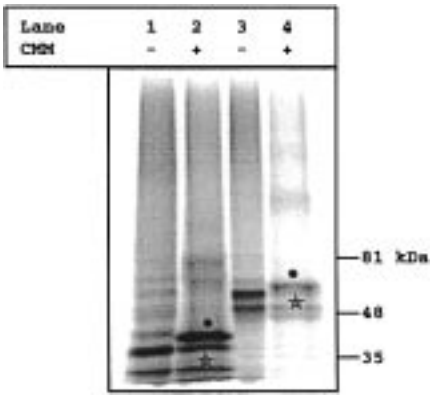

FIGURE 3: Orientation of the N- and C-termini. The native HISBAT sequence was translated with or without the C-terminal β-glycosylation flag. The translation mixes were separated on a 4−12% Tris/glycine gel and exposed to a phosphor-image screen. Products with N-terminal glycosylation in the presence of canine microsomal membranes (CMM) are indicated with (●) (N_{exo}/C_{cyt}). Molecular mass shifts due to the glycosylation are ~3 kDa for both HISBAT translations. Bands with lower molecular mass that originate from "false" downstream translation initiations are marked (☆).

Orientation of the N- and the C-Terminus. To determine the location of the N- and the C-terminus, the β-glycosylation flag was appended to the C-terminus of the HISBAT natural sequence. The HISBAT−β fusion protein produced has one endogenous glycosylation site at the N-terminus and the five consensus sites on the β-flag at the C-terminus. For comparison, the native sequence was expressed without the C-terminal glycosylation flag in a second vector. As shown in Figure 3 (lanes 2, 4), the same glycosylation pattern is seen with both types of constructs. The molecular mass shift of 3 kDa seen in the presence of canine microsomes corresponds well to core glycosylation of the single NAT site on the HISBAT N-terminus. This result also confirms earlier observations from immunoprecipitation and immunoblotting analyses of various SBATs (*30*) and SBAT chimeras (*9*). The data provide evidence for an odd number of transmembrane regions between an exoplasmic N-terminus and a cytoplasmic C-terminus. As seen in Figures 3 and 6, all constructs starting at the natural N-terminus exhibit multiple translation initiation sites (lower molecular weight products marked with unfilled stars). The apparent molecular weights of these smaller products are similar to those predicted for utilization of downstream ATG sequences in the HISBAT sequences. This apparent downstream initiation was not hindered by introduction of part of the Kozak consensus sequence (ACCATG) (*29*) at the first start codon or by capping the transcribed mRNA (data not shown). Systematic mutations of candidate start codons revealed that deleted initiation sites were readily replaced with neighboring ones. This indicates that the translation start is most likely directed by structural effects of the mRNA (*29*).

Signal Anchor Activity of HISBAT Hydrophobic Sequences. Based on hydropathy predictions, membrane insertion of individual hydrophobic sequences of the protein was tested in the HK M0 vector. The predicted nine transmembrane segments are shown in the upper panel of Figure 4. Sequences tested in the HK M0 construct are indicated with arrows. Figure 4, lower panel, displays the phosphor-images of the *in vitro* translation gels. HISBAT H1, H2, H4, H5, H6, H7, and H9 all behave as signal anchor sequences. In fact, these segments behaved as signal anchors even though several of these constructs require the region to insert in an orientation opposite to that in the mature protein. The translocation yield for the β-flag varies between the constructs, possibly reflecting differences in SRP binding and/or the intrinsic membrane insertion properties. All the observed molecular mass shifts, postulated to be due to core glycosylation of the β-flag (~10 kDa), were reversed by *post*-translational deglycosylation. The HISBAT H3 and H8 segments were not able to insert as independent signal anchor sequences in this M0 vector. The background bands seen in Figure 4, lanes 4, 10, and 13, were all due to high translation yields in the absence of microsomes causing overloading of the gels. These translation products are not associated with the microsomal membranes as demonstrated by alkaline extraction (data not shown) and should not be confused with the β-glycosylation seen in lanes 5, 11, and 14.

Stop Transfer Properties of HISBAT Hydrophobic Sequences. Figure 5 summarizes the stop transfer properties of HISBAT hydrophobic sequences as reflected by their ability to prevent translocation of the β-flag initiated by the M1 signal anchor sequence from the H^+,K^+-ATPase α-subunit. The HISBAT sequences were inserted between the HK M1 and β-flag and translated as described above. HISBAT regions competent to block translocation resulted in nonglycosylated products whose membrane insertion was confirmed by alkaline extraction of the microsomal membrane. Recovery of membrane-inserted segments after precipitation was verified by the translation and alkali extraction of the HK M1 vector as an internal standard and was normally above 50% (Figure 5A). As a control for possible adsorption to the lipid bilayer, microsomal membranes were added after translation and subsequently extracted as described above. Stop transfer activity of a specific sequence was then defined as the percent recovery of product minus percent background adsorption to microsomes. Inability to block translocation resulted in glycosylation of the β-flag and was consequently detected as molecular weight shift after SDS−PAGE. This was seen with the HISBAT H3 and H8 sequences (Figure 5B, lanes 11, 12, 31, and 32) that did not behave as a stop transfer segments when inserted downstream of the H^+,K^+-ATPase M1 sequence.

Sequences with incomplete, but significant, stop transfer activities, HISBAT H2 and H5 are seen in lanes 8 and 20 (Figure 5B). These sequences prevented glycosylation of the β-flag in 50 and 74% of the membrane-inserted translation products, respectively. Only H1, H7, and H9 exhibited highly efficient stop transfer activity and high recovery after alkaline extraction. HISBAT H4 and H6 both act as stop transfer signals, but the translation products also show a lower efficiency for membrane insertion. The reason could be a diminished SRP recognition of HK M1 due to negative interactions with the downstream HISBAT region. This would result in a lower yield of membrane-inserted protein versus product translated in the cytoplasm.

Consecutive HISBAT Constructs: C-Terminal Truncations. To examine the insertion properties of the predicted transmembrane segments in context of their native sequences, consecutive HISBAT segments were ligated into the pGEM vector upstream of the H^+,K^+-ATPase β-flag and translated as described above. C-terminal truncations were made at various positions in the native sequence. HISBAT has an

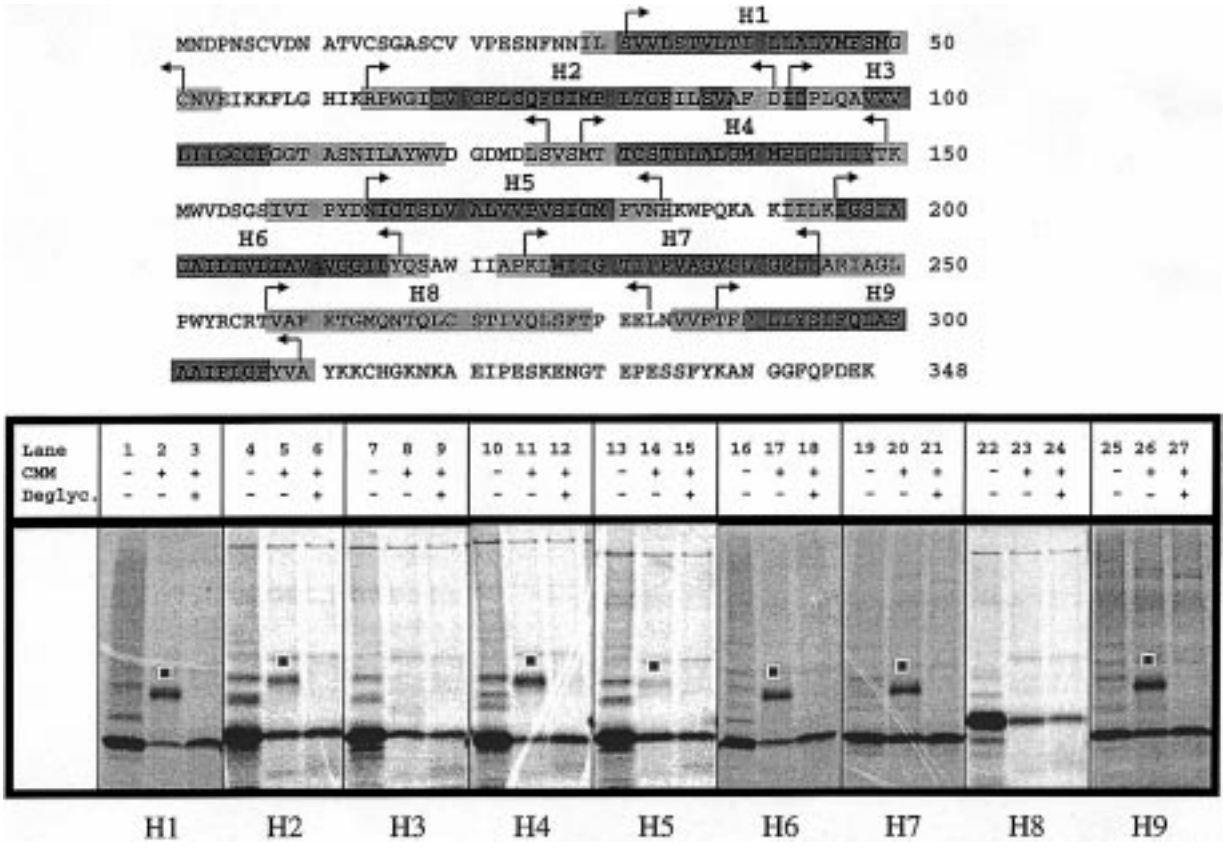

FIGURE 4: Signal anchor activity of predicted HISBAT TM regions. Hydrophobic HISBAT segments were tested in the HK M0 vector for independent insertion as indicated by translocation of the C-terminal β-flag in the presence of canine microsomal membranes (CMM). In the upper panel, segments shaded in dark gray constitute the consensus output from the computer analyses while amino acids shaded in light gray are predicted by any of the algorithms to reside within the lipid bilayer. The H8 sequence above is only predicted by the HMMTOP algorithm (*26*). Amino acids included in the constructs are marked with arrows. The molecular mass shift on the 10% Tris/glycine gels (phosphor-image), due to glycosylation of the five consensus sites on the β-flag, is ~10 kDa and was reversed by endoglycosidase treatment. Glycosylated products are marked (■) in the phosphor-images.

endogenous consensus site for *N*-linked glycosylation preceding its first transmembrane region that serves as an intrinsic topogenic marker for the N-terminal domain. Glycosylation of this *N*-linked site results in a 3 kDa molecular mass shift as shown in Figure 3, while glycosylation of the C-terminal β-flag alone produces a 10 kDa migration shift when analyzed by SDS−PAGE (Figure 5A). Consequently, insertion with N_{exo}/C_{cyt}, N_{cyt}/C_{exo}, and N_{exo}/C_{exo} orientations can be probed as molecular mass shifts of 3, 10, and 13 kDa, respectively.

Translations of native sequences with C-terminal truncations are shown in Figure 6. The upper panel displays the predicted transmembrane sequences along with arrows indicating the position of the β-flag. In the lower panel, the phosphor-images of the translation products separated by SDS−PAGE are shown with symbols indicating the orientation of the N- and C-termini.

Translation/translocation of the native sequence truncated after the positive charges in the H1−H2 loop results in a major product with N-terminal glycosylation (Figure 6, Nt-A, lane 2, band 3). Bands at higher molecular weight (numbers in the figure) represent background (lane 2, band 1) and N_{cyt}/C_{exo} orientated product (lane 2, band 2), while bands 4 and 5 correspond to nonglycosylated Nt-A product and false downstream initiation, respectively. When H2 is included downstream of the H1−H2 loop, a new band corresponding to glycosylation of both the N-terminus and the β-flag appears on the gel (Figure 6, Nt-B, lane 4, band 1). The observation that H1 and H2 can form a TM pair with N_{exo}/C_{exo} orientation is in agreement both with the "positive inside" rule and with the signal anchor/stop transfer data obtained for these segments (Figures 4 and 5). Furthermore, it demonstrates that H2 can function as a signal anchor sequence in its native context, independent of the downstream H3 segment. The presence of a product glycosylated only at the N-terminus may reflect a low efficiency of H2 insertion in the *in vitro* system (Figure 6, Nt-B, lane 4, band 2). In the same lane, band 3 represents nonglycosylated Nt-B product, and bands 4 and 5 are both due to false downstream initiation. When the truncation was made at the C-terminal side of H3 (Figure 6, N_t-C), glycosylation of the β-flag was fully prevented, yielding a product with N_{exo}/C_{cyt} orientation (Figure 6, lane 6, band 2). Whereas the H3 segment does not function as a stop transfer sequence in the HKM1 vector, the H3 segment in its native context is capable of efficiently preventing the translocation initiated by H2 (see Figure 5B, lanes 11 and 12). The molecular weight of band 1 in lane 6, Figure 6, fits well with β-flag glycosylation of the major downstream-initiated Nt-C product (band 4).

The HISBAT H4 and H5 regions act as both signal anchor and stop transfer sequences when tested individually in the HK M0/M1 vectors (Figures 4 and 5). Translation of constructs with truncations made in the native sequence

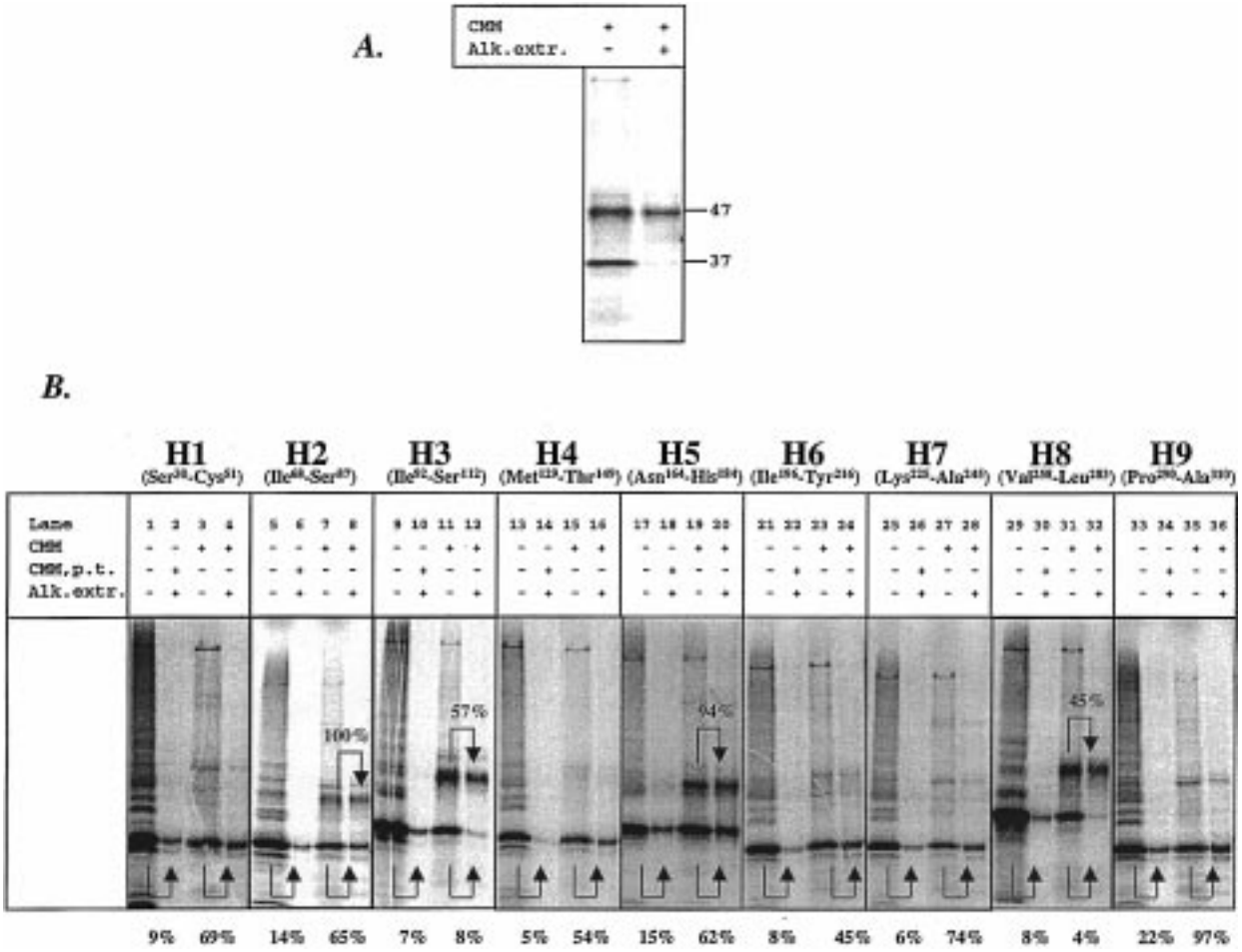

FIGURE 5: Stop transfer activity of HISBAT segments in the HK M1 vector. Panel A shows the molecular weight shift caused by translocation and glycosylation of the β-flag induced by the HK M1 sequence in the presence of canine microsomal membranes (CMM). The recovery of membrane-anchored versus cytoplasmic-translated product (HK M1-β) after extraction with 0.1 M Na_2CO_3 (Alk.extr.) is seen in the right lane. Panel B displays the behavior of nine hydrophobic HISBAT segments as stop transfer sequences. Membrane insertion was verified by alkaline extraction, and adhesion of products to the microsome surface was determined by *post*-translational addition of the microsomes (CMM,p.t.) prior to extraction. Recovery of product (indicated by arrows) after alkaline extraction was calculated from the integrated optical density of the bands in the phosphor-image and is shown as percent of control (the nonextracted product) underneath the lanes. Segments with partial stop transfer activity are seen in lanes 8 and 20 (Figure 5B).

downstream of both H4 and H5 (Figure 6, Nt-D and Nt-E) generates major products with N_{exo}/C_{cyt} orientation (Figure 6, lanes 8 and 10, bands 2 and 1, respectively), although there is a weak glycosylation signal from the β-flag in the construct terminating with H4 (lane 8, band 1). This indicates that the H4−H5 region either usually inserts as a pair or is not integrated into the membrane when tested in context with preceding HISBAT sequence. Nonglycosylated products from translation of the Nt-D (lane 8) and Nt-E (lane 10) constructs correspond to bands 3 and 2 in respective lanes, while band 4 in lane 8 and bands 3 and 4 in lane 10 are all due to false downstream initiation.

The N_{exo}/C_{cyt} orientation obtained with the N_t-H5 construct is shifted to a mixed N_{exo}/C_{cyt} and N_{exo}/C_{exo} topography when the native sequence is truncated after H6 (Figure 6, N_t-F, lane 12, bands 1 and 3). Accordingly, H6 acts as a signal anchor in the HK M0 and partially as such in the C-terminal truncated constructs. A weak band corresponding to β-glycosylation of downstream-initiated Nt-F product (band 2) can also be seen in lane 12 along with nonglycosylated product (band 4) and downstream initiation products (bands 5 and 6). When the native sequence ends on the C-terminal side of H7 (Figure 6, Nt-G), the major translation product once again has predominantly N-terminal glycosylation (Figure 6, lane 14, band 2), verifying the stop transfer properties seen for H7 in the HK M1 vector (Figure 5B, lanes 27 and 28). Bands 1 and 3 in lane 12 represent N_{exo}/C_{exo} oriented and nonglycosylated products, respectively, as downstream initiated Nt-G products are seen as bands 4 and 5. The N_{exo}/C_{cyt} topography initiated by the H7 sequence remains unchanged when the truncation is made after the next predicted transmembrane segments, H8 (Figure 6, Nt-H, lane 16, band 2) and H9 (Figure 6, Nt-I, lane 18, band 1). A small fraction of N_{exo}/C_{exo}-inserted Nt-H product is represented by band 1 in lane 16, and in the same lane, with lower molecular weight, band 3 and bands 4 and 5 correspond to nonglycosylated and downstream initiated products, respectively. In lane 18, nonglycosylated Nt-I product corresponds to band 2, and the downstream initiated products are seen as bands 3 and 4. The exclusive N_{exo}/C_{cyt} orientation seen for the Nt-I (H9) construct is in agreement with the cytoplasmic position found for the native HISBAT C-terminus (Figure 3, lane 4). The data also suggest that H8 and H9 insert as a transmembrane pair or do not insert at all. The strong signal anchor/stop transfer properties of H9 in the HK M0/M1 vectors support the former possibility.

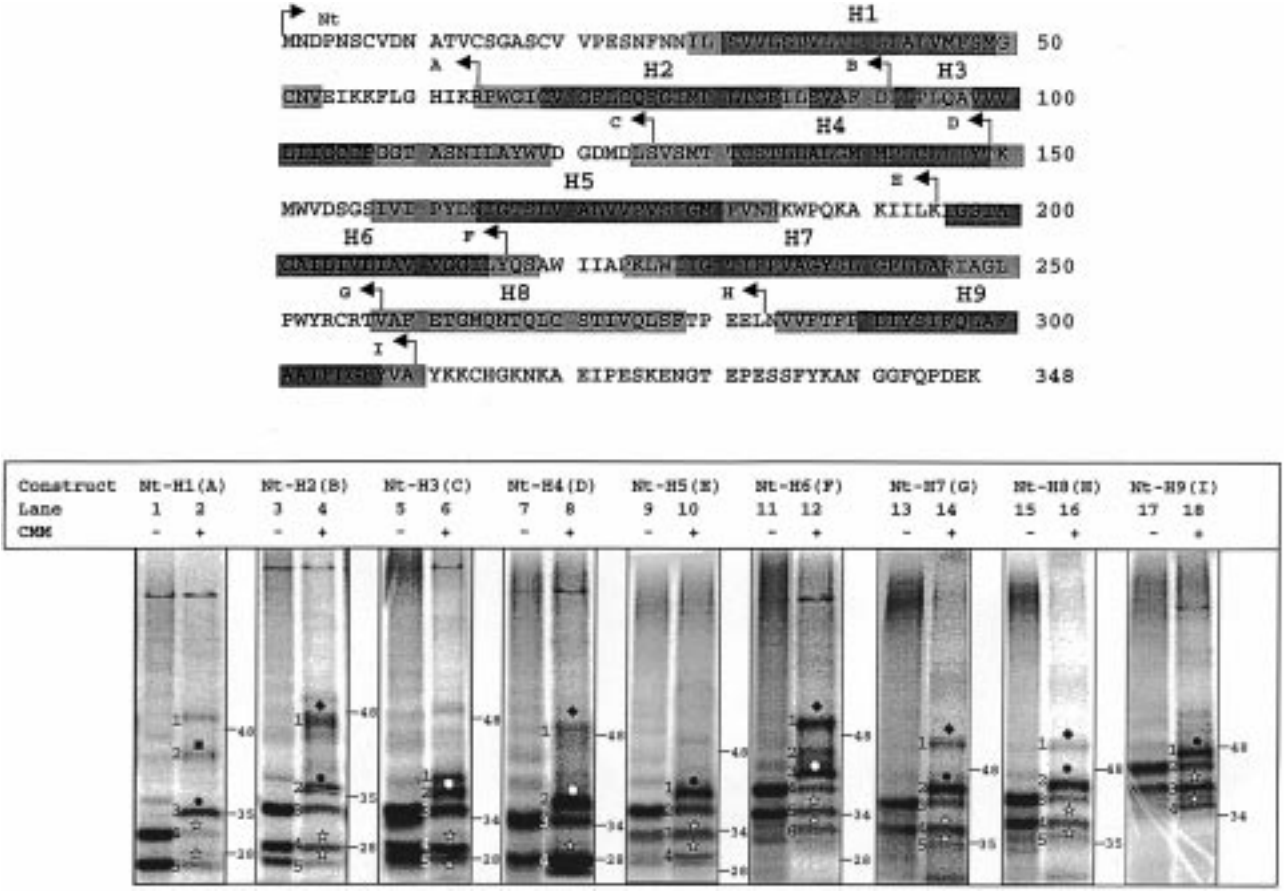

FIGURE 6: Translation of C-terminal truncated native HISBAT sequences. The β-flag was positioned downstream of the native N-terminus as indicated in the upper panel. Segments shaded in dark gray constitute the consensus output from the topology/hydropathy analyses while amino acids shaded in light gray are predicted by any of the algorithms to reside within the lipid bilayer. Products from the translations were separated on 10% Tris/glycine gels, and the phosphor-images are presented in the lower panel along with markers for molecular mass in kDa. Phosphor-imager bands corresponding to glycosylation of the N- and/or the C-terminus after insertion of translation products into the microsome membranes (CMM) are marked for the following orientations: (◆) for N_{exo}/C_{exo} (~13 kDa shift); (●) for N_{exo}/C_{cyt} (~3 kDa shift); and (■) for N_{cyt}/C_{exo} (~10 kDa shift). Bands with lower molecular mass that originate from "false" downstream translation initiations are marked (☆). Numbers to the right indicate molecular mass, and numbers to the left, the different translation products discussed in text.

Consecutive HISBAT Constructs Tested in the HK M0 Vector. As shown above, H3 acts as a stop transfer sequence when translated downstream of the HISBAT sequence starting at the native N-terminus. The same trend is seen for the HK M0-H23 constructs. Glycosylation of the β-flag initiated by HISBAT H2 [Figure 7, H2 (Arg64−Asp91), lane 2 (■)] was partially inhibited when the sequence was truncated at the end of the downstream H3 segment [Figure 7, H23S (Arg64−Pro107), lane 4] and fully blocked when the loop between H3 and H4 was included in the construct [Figure 7, H23L (Arg64−Ser126), lane 6]. This result demonstrates again that the third hydrophobic HISBAT sequence is able to act as a stop transfer segment when translated in context with the H2 sequence.

In conflict with the signal anchor activity seen for H4 in the HK M0 vector, this region was able to insert only very weakly when translated downstream of native sequence starting at the N-terminus. The N_{exo}/C_{cyt} orientation seen with translation of the Nt-H5 construct indicated that the H4−H5 sequence either inserts as a pair or is not part of the transmembrane domain. To test this hypothesis, a construct was made with the H4−H5 sequence (Met129−His184) inserted into the HK M0 vector. Translation of this construct is shown in Figure 7, lanes 7−10. Only very weak β-flag glycosylation can be seen in the presence of microsomes [lanes 9 and 10, (■)]; the major translation product is nonglycosylated product that is membrane-embedded as confirmed by alkaline extraction (lane 10). Consequently, the H4−H5 region behaves as a transmembrane pair with N_{cyt}/C_{cyt} orientation when tested in the HK M0 vector.

In the C-terminal truncated constructs, H6 initiates translocation of the β-flag, and H7 acts as a stop transfer sequence (Figure 6). This suggests that the H6−H7 region inserts with a N_{cyt}/C_{cyt} orientation into the microsomal membrane. To further investigate the integration of H6−H7 as a transmembrane pair, the segment (Ile196−Ala245) was tested in the HK M0 vector and the translation product subject to alkaline extraction. Figure 7, lanes 13−14, displays the fraction of HK M0-H67 translated in the presence of ER microsomes that is recovered after alkaline extraction compared to the control seen in lanes 11 and 12. Glycosylation of the β-flag initiated by H6 (see Figure 4, lane 17) is clearly prevented by the downstream addition of H7 (Figure 7, lane 13), and the nonglycosylated HK M0-H6−H7 product is membrane-integrated as confirmed by alkaline extraction (Figure 7, lane 14). Hence, the signal anchor and stop transfer properties of H6 and H7, respectively, can be confirmed both in the truncated native sequences and in the consecutive HK M0-H67 construct.

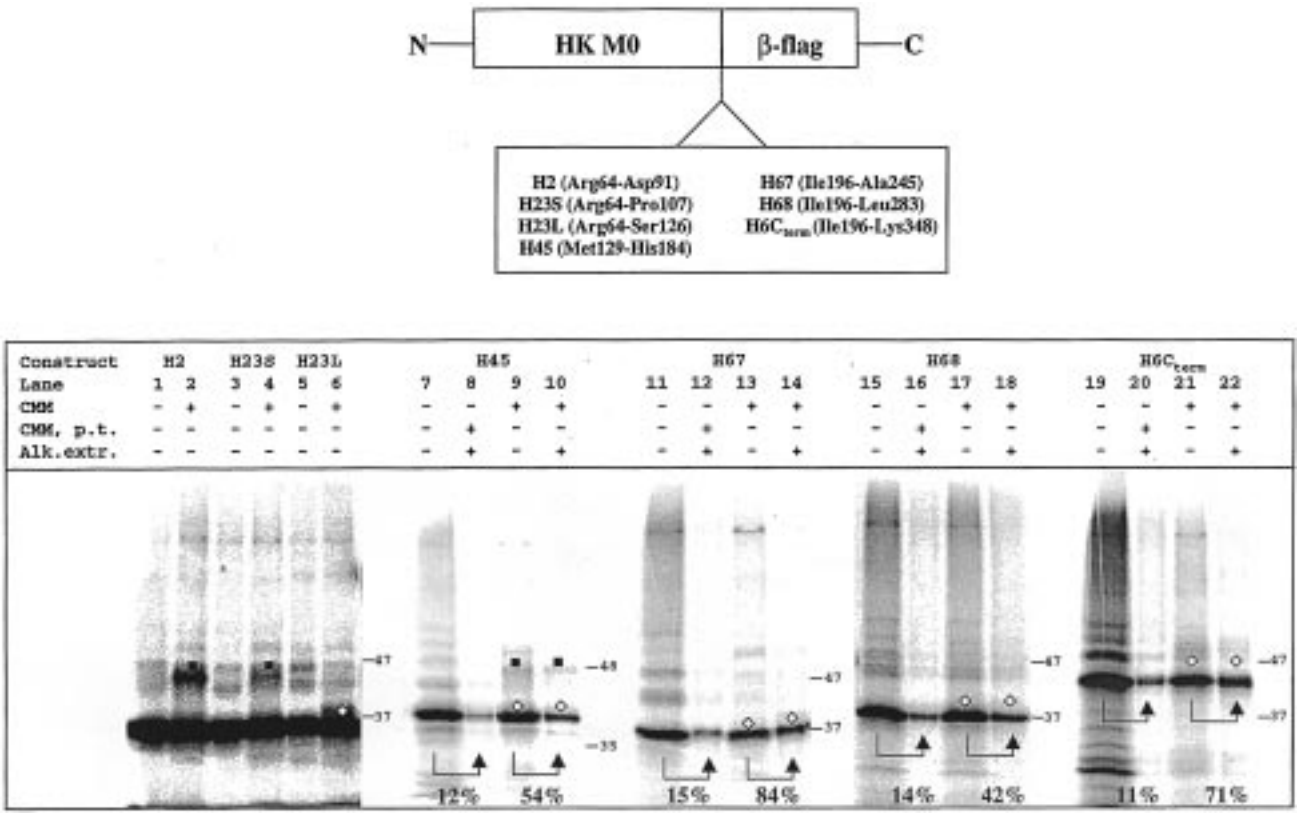

FIGURE 7: Consecutive HISBAT sequences translated in the HK M0 vector. Sequential HISBAT segments tested in the HK M0 vector are outlined in the upper panel. The translation products were separated on 10% Tris/glycine SDS gels, and the corresponding phosphorimages are shown in the lower panel along with markers for molecular mass. Bands corresponding to membrane insertion with N_{cyt}/C_{cyt} and N_{cyt}/C_{exo} orientations are marked (◇) and (■), respectively. Membrane insertion of the H4−5, H6−7, H6−8, and H6−C_{term} products was verified by alkaline extraction as for the translations in the HK M1 vector (Figure 5).

In addition, constructs starting at H6 (Ile196) and ending either after H8 (Leu283) or at the native C-terminus (Lys348) were inserted and translated in the HK M0 vector (Figure 7, H68 and H6C_{term}). A nonglycosylated membrane-integrated product was formed after translation of both vectors as can be seen in Figure 7, lanes 15−18 (H68) and lanes 19−22 (H6C_{term}). The sequence starting downstream of H7 and ending at the C-terminus is therefore acting either as a transmembrane pair or as an noninserted segment when translated in HK M0 (compare with Nt-H7, Nt-H8, and Nt-H9 in Figure 6). The fact that H9 is clearly inserted as a transmembrane segment in both the HK M0 and HK M1 vectors (Figure 4, lane 23, and Figure 5, lanes 31−32, respectively) indicates integration of the H8/H9 pair. Although H8 was neither a signal anchor nor a stop transfer sequence in the HK M0/M1 vectors (Figures 4 and 5B), it behaves as an integrated segment when translated in its native context with the neighboring hydrophobic sequences, H7 and H9.

DISCUSSION

The mammalian SBATs exhibit no significant homology with the other families of sodium co-transporters. Therefore, little structural information can be obtained by direct comparison with well-characterized carriers such as the sodium/glucose co-transporter or sodium/neurotransmitter co-transporters (*12*, *13*). Earlier work using glycosylation or antibody epitope localization suggested the SBATs possess an odd number of transmembrane domains giving a N_{exo}/C_{cyt} topography. Therefore, the eight transmembrane topography model, anticipated by three out of the four predictive algorithms used here, must be modified by addition or deletion of an odd number of membrane segments. In earlier 2D models for the SBATs, the latter was accomplished either by joining the second and third hydrophobic sequences into one long segment (*14*) or by excluding the fourth predicted sequence from the transmembrane domain (*4*, *10*). These models were, however, not founded on any detailed experimental topography investigation.

A new algorithm (HMMTOP) mapping topography and transmembrane helices based on the difference in the distribution of amino acids in structural domains predicts an additional integrated segment starting at Val258 and ending at Thr279. This region is highly conserved among the SBATs and contains a number of polar residues that could serve in solute binding and translocation.

Here, we have used three different types of constructs with an *in vitro* translation system to identify integrated membrane segments in the human ileal SBAT. In these vectors, signal anchor and stop transfer properties were tested for individual and consecutive HISBAT segments. The orientation of these hydrophobic sequences when inserted in the HK M0 and HK M1 constructs is fully determined by the topogenic information given by the M0 cytoplasmic anchor and M1 signal anchor sequences. As a consequence, no knowledge about the native orientation of the tested segments could be obtained, just their ability to be targeted to the translocon and to insert as type II signal anchors (in the HK M0 vector) or to act as stop transfer signals (in the HK M1 vector). Independent of their possible orientation in the mature

protein, seven of the nine putative segments insert into the membrane both as independent signal anchor and as stop transfer sequences (see Figures 4 and 5). Only the H3 and H8 sequences show no integration properties as individual regions when translated in the HK M0/M1 vectors. Thus, the data obtained from individual segment insertion scanning by *in vitro* translation agree with a 7 TM model.

However, transmembrane regions in polytopic membrane proteins may require flanking topogenic information to become integrated into the lipid bilayer (*31*). It has been demonstrated here that HISBAT H5 is a more efficient stop transfer signal when placed downstream of the native H4 region (Figure 7) compared to the test signal anchor sequence in the HK M1 vector (Figure 5). Furthermore, the M5−M6 region in the H^+,K^+- and SR Ca^{2+}-ATPases does not insert when translated in the HK M0/M1 constructs, although data from various other techniques show this pair to be membrane-embedded. Therefore, we analyzed constructs encompassing the HISBAT H3 and H8 segments in the context of the native flanking sequence.

The common algorithms predict a transmembrane pair arrangement of H2 and H3. In Figures 6 and 7, we show that the H3 segment acts as a stop transfer signal when preceded by the native H2 sequence. The HISBAT H3 completely prevents glycosylation of the β-flag in both constructs used. The finding that the H2−H3 region acts as a membrane pair inverts the orientation of all downstream segments compared with the 7 TM model. As a result, the positively charged clusters between H5−H6 and H7−H8 are now cytoplasmic rather than extra-cytoplasmic, consistent with the "positive inside" rule. This orientation is further supported by the data obtained from translations of the C-terminal truncated Nt-H5, Nt-H6, and Nt-H7 constructs (Figure 6). In addition, the observation that H5 functions as a stop transfer in the native sequence provides an explanation for the weak signal anchor activity seen with this segment in the HK M0 vector (Figure 4).

From the data presented in Figures 6 and 7, it can be concluded either that H8 and H9 are excluded from the transmembrane domain or that they form a transmembrane pair. The former is less likely as HISBAT H9 is acting as a strong transmembrane region in both the HK M0 and HK M1 vectors. The HISBAT H8 segment is only moderately hydrophobic, and it is unable to insert as an individual region in the HK M0 and HK M1 vectors. Nevertheless, its presence in the HK M0 H6−C_{term} and N_t-H9 vectors resulted in a product consistent with an additional membrane-inserted sequence. Membrane insertion of H8 allows H9 to have a N_{exo}/C_{cyt} orientation as required from translation of the native sequence (Figure 3). Integration of hydrophilic segments preceding strong type I signal anchor sequences (like HISBAT H9) has recently been demonstrated for the human band 3 protein along with a set of test constructs (*32*). A general mechanism where the N_{exo}/C_{cyt} orientation of the signal anchor forces the upstream hydrophilic segment into a transmembrane position was proposed and could also provide an explanation for the insertion behavior of the HISBAT H8 segment.

Based on computer predictions and the *in vitro* translation/translocation data, we have found evidence for an odd number of transmembrane regions and therefore seven or nine membrane spanning regions in the core structure of the human ileal transporter. With the finding that H3 is a membrane-inserted segment, the model now enables the downstream sequence to conform to the "positive inside" rule and also follows hydropathy predictions for HISBAT. Exclusion of H9 as a transmembrane segment seems unlikely, and the data therefore suggest a nine transmembrane segment arrangement for the intestinal sodium-dependent bile acid transporter.

ACKNOWLEDGMENT

We thank Drs. Krister Bamberg, Denis Bayle, Jai Moo Shin, and Greg Shelness for helpful discussions.

REFERENCES

1. Hofmann, A. F. (1993) *Gastrointestinal disease. Pathophysiology, diagnosis, management* (Sleisenger, M. H., and Fordtran, J. S., Eds.) 5th ed., pp 127−150, W. B. Saunders Co., Philadelphia.
2. Wong, M. H., Oelkers, P., Craddock, A. L., and Dawson, P. A. (1994) *J. Biol. Chem. 269*, 1340−1347.
3. Christie, D. M., Dawson, P. A., Thevananther, S., and Shneider, B. L. (1996) *Am. J. Physiol. 271*, G377−G385.
4. Hagenbuch, B., Stieger, B., Foguet, M., Lubbert, H., and Meier, P. J. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 10629−10633.
5. Stengelin, S., Apel, S., Becker, W., Maier, M., Rosenberger, J., Wess, G., and Kramer, W. (1995) GeneBank accession no. Z54357.
6. Stengelin, S., Becker, W., Maier, M., Noll, R., and Kramer, W. (1998) GeneBank accession no. AJ131361.
7. Saeki, T., and Matoba, K. (1997) Genebank accession no. AB002693.
8. Saeki, T. (1997) GeneBank accession no. AB003303.
9. Wong, M. H., Oelkers, P., and Dawson, P. A. (1995) *J. Biol. Chem. 270*, 27228−27234.
10. Hagenbuch, B., and Meier, P. J. (1994) *J. Clin. Invest. 93*, 1326−1331.
11. Reizer, J., Reizer, A., and Saier, M. H., Jr. (1994) *Biochim. Biophys. Acta 1197*, 133−166.
12. Wright, E. M., Loo, D. D., Turk, E., and Hirayama, B. A. (1996) *Curr. Opin. Cell Biol. 8*, 468−473.
13. Nelson, N. (1998) *J. Neurochem. 71*, 1785−1803.
14. Oelkers, P., Kirby, L. C., Heubi, J. E., and Dawson, P. A. (1997) *J. Clin. Invest. 99*, 1880−1887.
15. von Heijne, G., and Gavel, Y. (1988) *Eur. J. Biochem. 174*, 671−678.
16. Wallin, E., and von Heijne, G. (1998) *Protein Sci. 7*, 1029−1038.
17. Bayle, D., Weeks, D., and Sachs, G. (1997) *J. Biol. Chem. 272*, 19697−19707.
18. Bamberg, K., and Sachs, G. (1994) *J. Biol. Chem. 269*, 16909−16919.
19. Bayle, D., Weeks, D., and Sachs, G. (1995) *J. Biol. Chem. 270*, 25678−25684.
20. Chen, M., and Zhang, J. T. (1996) *Mol. Membr. Biol. 13*, 33−40.
21. Meera, P., Wallner, M., Song, M., and Toro, L. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 14066−14071.
22. Melchers, K., Weitzenegger, T., Buhmann, A., Steinhilber, W., Sachs, G., and Schafer, K. P. (1996) *J. Biol. Chem. 271*, 446−457.
23. Shin, J. M., Besancon, M., Bamberg, K., and Sachs, G. (1997) *Ann. N.Y. Acad. Sci. 834*, 65−76.
24. Claros, M. G., and von Heijne, G. (1994) *Comput. Appl. Biosci. 10*, 685−686.
25. Rost, B., Fariselli, P., and Casadio, R. (1996) *Protein Sci. 5*, 1704−1718.
26. Tusnady, G. E., and Simon, I. (1998) *J. Mol. Biol. 283*, 489−506.
27. Engelman, D. M., Steitz, T. A., and Goldman, A. (1986) *Annu. Rev. Biophys. Biophys. Chem. 15*, 321−353.

28. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105−132.
29. Kozak, M. (1991) *J. Biol. Chem. 266*, 19867−19870.
30. Dawson, P. A., and Oelkers, P. (1995) *Curr. Opin. Lipidol. 6*, 109−114.
31. Hegde, R. S., and Lingappa, V. R. (1997) *Cell 91*, 575−582.
32. Ota, K., Sakaguchi, M., von Heijne, G., Hamasaki, N., and Mihara, K. (1998) *Mol. Cell 2*, 495−503.

BI990554I