Multiple Determinants Direct the Orientation of Signal–Anchor Proteins: The Topogenic Role of the Hydrophobic Signal Domain

Johanna M. Wahlberg and Martin Spiess

Biozentrum, University of Basel, CH-4056 Basel, Switzerland

Abstract. The orientation of signal–anchor proteins in the endoplasmic reticulum membrane is largely determined by the charged residues flanking the apolar, membrane-spanning domain and is influenced by the folding properties of the NH_2-terminal sequence. However, these features are not generally sufficient to ensure a unique topology. The topogenic role of the hydrophobic signal domain was studied in vivo by expressing mutants of the asialoglycoprotein receptor subunit H1 in COS-7 cells. By replacing the 19-residue transmembrane segment of wild-type and mutant H1 by stretches of 7–25 leucine residues, we found that the length and hydrophobicity of the apolar sequence significantly affected protein orientation. Translocation of the NH_2 terminus was favored by long, hydrophobic sequences and translocation of the COOH terminus by short ones. The topogenic contributions of the transmembrane domain, the flanking charges, and a hydrophilic NH_2-terminal portion were additive. In combination these determinants were sufficient to achieve unique membrane insertion in either orientation.

MOST proteins are targeted and inserted into the endoplasmic reticulum membrane by a mechanism involving signal recognition particle, single recognition particle receptor, the Sec61 translocation complex, and translocating chain-associating membrane protein (TRAM;[1] for review see Walter and Johnson, 1994). Three types of single-spanning membrane proteins are generated by this machinery (von Heijne and Manoil, 1990; Spiess, 1995). Insertion of type I membrane proteins is initiated by an NH_2-terminal, cleavable signal sequence that directs the transfer of the COOH-terminal sequence across the membrane. Translocation is terminated by a hydrophobic stop–transfer sequence that anchors the protein in the membrane with an exoplasmic NH_2 terminus and a cytoplasmic COOH terminus (N_{exo}/C_{cyt} orientation). Typically, cleaved signals are composed of a short, positively charged hydrophilic segment followed by a hydrophobic domain of 7–15 uncharged residues. In proteins of types II and III, the signal is not necessarily located at the very NH_2 terminus, it is not cleaved, and it anchors the protein in the membrane. These hydrophobic domains are longer, usually 19–27 residues (Nilsson et al., 1994). Type II signal anchors, like cleavable signals, initiate translocation of their COOH-terminal sequence generating an N_{cyt}/C_{exo} orientation of the protein. In contrast, type III signal–anchor sequences (or "reverse signal anchors") promote translocation of their NH_2-terminal sequence and produce an N_{exo}/C_{cyt} topology.

The most prominent feature that determines which end of the signal is translocated is the distribution of charged residues flanking the signal anchor sequence. Statistically, positive charges are enriched on the cytosolic side and depleted from the exoplasmic side of signal–anchor sequences (the "positive inside" rule; von Heijne, 1986; von Heijne and Gavel, 1988). For eukaryotic proteins, the charge difference between the two flanking segments (more positive on the cytosolic side) rather than the positive residues per se has been shown to correlate with protein orientation (Hartmann et al., 1989). The topogenic role of the flanking charges has been experimentally confirmed by site-directed mutagenesis. The topology of bacterial membrane proteins (e.g., von Heijne, 1989) as well as of eukaryotic ones was shown to be affected by charge mutations. The type III protein cytochrome P-450 was converted to a type II protein by insertion of positively charged residues into its short NH_2-terminal domain (Monier et al., 1988; Szczesna-Skorupa et al., 1988; Szczesna-Skorupa and Kemper, 1989; Sato et al., 1990). The asialoglycoprotein (ASGP) receptor subunit H1 and the paramyxovirus hemagglutinin–neuraminidase, two type II proteins, were induced to insert partially in the opposite type III orientation by mutation of flanking charges (Beltzer et al., 1991; Parks and

Lamb, 1991, 1993). Positive charges had a stronger effect on topogenesis than negative ones and were more effective the closer they were to the hydrophobic segment. However, in these and other studies (e.g., Andrews et al., 1992) the asymmetric distribution of flanking charges in recombinant proteins was not sufficient to generate a unique topology. Additional requirements must be met to efficiently direct the polypeptide to insert with a single topology as generally observed for natural membrane proteins.

Another topologically important feature is the folding state of the NH_2-terminal hydrophilic domain. This segment is already synthesized when the signal–anchor sequence emerges from the ribosome. For the ASGP receptor H1 it was shown that this domain needs to be unfolded for translocation and that translocation of the NH_2 terminus is hindered or even prevented by a rapidly and stably folding domain but is facilitated by destabilizing mutations (Denzer et al., 1995).

An influence of the hydrophobic segment on the function of signal and signal–anchor sequences was suggested by in vitro experiments by Sato et al. and Sakaguchi et al. Short deletions within the hydrophobic segment of the N_{exo}/C_{cyt} signal–anchor sequence of cytochrome P-450 resulted in translocation of the COOH-terminal reporter sequence for a fraction of the product proteins (Sato et al., 1990). Similarly, an artificial signal sequence with a hydrophobic segment of fewer than 12 leucines and a negative NH_2-terminal net charge was found to translocate the COOH-terminal reporter sequence, whereas with longer hydrophobic segments of 13 or 15 leucines, a fraction of the polypeptides was anchored in the microsomal membrane from the cytosolic side as type III proteins (Sakaguchi et al., 1992).

In the present study we tested in vivo the role of the hydrophobicity and the length of the apolar domain of a type II signal–anchor sequence, either close to the NH_2 terminus or in an internal position and in the context of altered flanking charges. We found that long, hydrophobic signal domains promote NH_2-terminal translocation, whereas short ones promote translocation of the COOH-terminal sequence. This effect adds to the topogenic influence of mutations in the charged flanking residues. By combining appropriate flanking charges and apolar domains it is possible to design polypeptides with essentially unique N_{cyt}/C_{exo} or N_{exo}/C_{cyt} topology.

Materials and Methods

DNA Constructs

The plasmids encoding the wild-type ASGP receptor subunit H1 (Spiess et al., 1985; Spiess and Lodish, 1986) and the mutant constructs H1-4 (Beltzer et al., 1991), HC (Schmid and Spiess, 1988), H1Δ (= H1[Δ2-37]), and H1ΔLeu19 (Wahlberg et al., 1995) have been described previously.

H1ΔLeu#. The sequence encoding Leu_{19} in H1ΔLeu19 ends in a BamHI site. The 5′ sequence was sequentially shortened or extended by three codons at a time by PCR using antisense oligonucleotide primers that corresponded to part of the Leu_{19} sequence and the BamHI site either lacking several leucine codons or containing additional ones. A second primer corresponding to a sequence in the plasmid vector was used. PCR products were digested with HindIII and BamHI and ligated to the 3′ BamHI–EcoRI portion of the H1 cDNA in the expression vector pECE (Ellis et al., 1986).

H1ΔQLeu#. The arginine codon CGC preceding the transmembrane portion of the H1ΔLeu# constructs was mutated to CAG for glutamine by PCR using the mutagenic primer CGGGTACCATGGGACCGCAGCT-GTTGC, which also encodes a 5′ KpnI site for subcloning. H1ΔQ was similarly generated using the primer CGGGTACCATGGGACCTCAG-CTCCTCC and H1Δ cDNA as the template.

H1Leu#. The KpnI site immediately preceding the initiation codon in the cDNAs of H1ΔLeu# was blunted and ligated to the 5′ HindIII–NruI fragment of pSA11/5 (Spiess and Handschin, 1987) to extend the 5′ end by that of wild-type H1.

H1-4Leu#. The plasmids encoding H1ΔLeu# were used as templates to amplify the DNAs encoding the transmembrane and COOH-terminal segments with a 5′ GAT codon for aspartic acid and an overlapping BglII site, using the mutagenic primer GACCAGATCTGTTGCTTTTGCT-GCTG. The DNA encoding the cytosolic portion of H1-4 with a matching BglII site was amplified using the primer CAACAGATCTGGTCCG-GAGCAGAGAT. The fragments were combined at the BglII site and subcloned into pECE. Finally, the 3′ BamHI–EcoRI fragments were replaced by the corresponding fragment of pSA1-4 (Beltzer et al., 1991).

H1-4gLeu#. The plasmid pSA1-4g (Beltzer et al., 1991) encoding H1-4 having a potential glycosylation site in the cytoplasmic domain was used as a template to amplify the DNAs encoding the NH_2-terminal hydrophilic segment with a 3′ BglII site, using the mutagenic primer CAACA-GATCTGGTCCGGAGCAGAGAT. The resulting HindIII–BglII fragment was ligated to the BglII–EcoRI 3′ portion of pSA1-4Leu# followed by subcloning of the combined cDNA (HindIII–EcoRI) into pECE. The sequence of the initial H1ΔLeu# constructs was confirmed by sequencing and of the derivatives by diagnostic restriction analysis.

In Vivo Expression and Analysis of Receptor Constructs

Cell culture reagents were purchased from GIBCO BRL (Gaithersburg, MD). COS-7 cells were grown in MEM with 10% fetal calf serum at 37°C with 7.5% CO_2. The media were supplemented with 2 mM L-glutamine, 100 U/ml penicillin, and 100 μg/ml streptomycin. Transfection of COS-7 cells was performed according to Cullen (1987) in 6-well clusters, and the cells were processed the second day after transfection. For in vivo labeling, transfected cells were incubated for 30 min in methionine-free medium, labeled for 30 min at 37°C with 100 μCi/ml [^{35}S]methionine in starvation medium, transferred to 4°C, and washed twice with PBS. To extract cytoplasmic proteins, the cells were incubated with 500 μl of 0.1% saponin in PBS for 30 min at 4°C. The saponin extract was removed, and the cells were lysed. Both fractions were immunoprecipitated using a rabbit antiserum directed against a synthetic peptide corresponding to residues 277–287 near the COOH terminus of the ASGP receptor H1 (anti-H1C). The immune complexes were isolated with protein A–Sepharose (Pharmacia, Upsala, Sweden) and analyzed by SDS–polyacrylamide gel electrophoresis. Alkaline extraction was essentially performed as described by Gilmore and Blobel (1985).

For analysis with endo-β-N-acetylglucosaminidase H (endo H; Boehringer Mannheim Corp., Mannheim, Germany), the immune complexes were isolated with protein A–Sepharose and boiled for 5 min in 50 μl, 50 mM Na-citrate, pH 6, 1% SDS. Aliquots were incubated with 1 mU endo H for 2 h at 37°C. Finally, samples were boiled in SDS–sample buffer and analyzed by SDS–polyacrylamide gel electrophoresis. The gels were fixed, soaked in 1 M Na-salicylate containing 1% glycerol, dried, and fluorographed on Kodak BioMax MR films. To quantify the relative abundance of polypeptides with different topologies, the fluorographs were densitometrically scanned, and the intensities of differently glycosylated species were expressed as percentages of the total of all forms.

For protease protection assays, the transfected cells were scraped into 200 μl PBS per well and pipetted up and down 10 times. Equal aliquots were incubated with or without 0.1 vol of 0.5 mg/ml TPCK-treated trypsin for 30 min on ice in the presence or absence of 1% Triton X-100. The reaction was terminated by addition of 0.1 vol of 1 mg/ml chicken egg white trypsin inhibitor and 10 mM PMSF. The samples were then subjected to immunoprecipitation using anti-H1C antiserum in the presence of 0.1 mg/ml trypsin inhibitor and 1 mM PMSF. The half lives of the glycosylated and unglycosylated forms of H1ΔLeu19, H1ΔLeu22, and H1ΔLeu25 were determined by pulse labeling for 30 min followed by a chase of up to 5 h. Products were immunoprecipitated and analyzed by gel electrophoresis and fluorography.

Results

H1Δ with a Leu₁₉ Sequence Replacing Its Hydrophobic Domain Is Partially Inserted in an Inverted N_{exo}/C_{cyt} Orientation

The ASGP receptor subunit H1 is a type II membrane protein with an internal signal–anchor sequence. The hydrophobic core of the signal consists of 19 mostly apolar residues starting after arginine-40 and followed by 8 mainly polar but uncharged amino acids and two glutamic acid residues (positions 68 and 69). Upon transfection of H1 cDNA into COS-7 cells, 30 min labeling with [^{35}S]methionine, and immunoprecipitation, a single polypeptide of 40 kD was recovered (Fig. 1 A, lane 1). Truncation of the NH_2-terminal domain to only 4 residues in H1Δ resulted in a product of ~36 kD (Fig. 1 A, lane 2) consistent with its reduced size. Digestion of an aliquot of the immunoprecipitates with endo H showed that H1Δ, like wild-type H1, was glycosylated at two sites in the COOH-terminal portion of the protein (Fig. 1 B, lanes 1–4; Beltzer et al., 1991). Deletion of the NH_2-terminal portion thus had no effect on the topology of the protein in the membrane. This was to be expected according to the charge difference criterion (Hartmann et al., 1989): the charge difference between the 15 COOH-terminal and the 15 NH_2-terminal flanking residues, Δ(C–N), was −3 for both H1 and H1Δ, statistically correlating with an N_{cyt}/C_{exo} orientation.

In a first construct to analyze the topogenic importance of the hydrophobic domain, the 19–amino acid apolar domain of H1Δ was replaced by a generic sequence of 19 leucine residues. Upon expression of this mutant H1ΔLeu19, two major forms were generated: one of 36 kD, with the same electrophoretic mobility as H1Δ, and a smaller one of ~30 kD (Fig. 1 A, lane 4). This latter form was resistant to endo H treatment and indistinguishable from the endo H digestion product of the higher molecular weight form (Fig. 1 B, lanes 5 and 6). This indicated that the 30-kD expression product was an unglycosylated form of H1ΔLeu19.
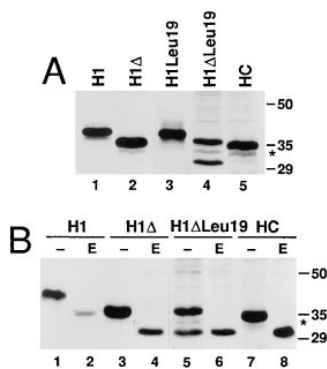

Figure 1. Mutant H1ΔLeu19 is expressed as two differently glycosylated forms. (A) COS-7 cells were transfected with cDNAs of H1, H1Δ, H1Leu19, H1ΔLeu19, and HC (the exoplasmic portion of H1 equipped with the cleavable signal sequence of influenza virus hemagglutinin) as indicated. The transfected cells were labeled for 30 min with [^{35}S]methionine, solubilized, and subjected to immunoprecipitation using an antiserum directed against the COOH-terminal sequence of H1. The immunoprecipitates were analyzed by gel electrophoresis and fluorography. (B) Immunoprecipitates were treated without (−) or with endoglycosidase H (E) before analysis by gel electrophoresis. The positions of marker proteins are shown with their molecular weights indicated in kD. The band indicated by an asterisk represents a partially glycosylated species (see text).

Lack of glycosylation could be due to a reduced efficiency of the mutated signal sequence in targeting the protein to the ER membrane, resulting in a soluble, cytosolic polypeptide. Alternatively, all products may still be integrated in the membrane, but some with the opposite N_{exo}/C_{cyt} orientation leaving the COOH-terminal portion with the glycosylation sites in the cytosol. Membrane association of the unglycosylated form was analyzed by saponin and alkaline extraction of expressing COS-7 cells. After labeling with [^{35}S]methionine, the cells were incubated with 0.1% saponin for 30 min at 4°C to release soluble proteins into the supernatant. Extracted and membrane-associated proteins were then immunoprecipitated from the saponin extract (S) and from the residual cells (C), respectively (Fig. 2 A). A secretory form of the exoplasmic portion of H1 with a cleavable signal sequence (HC) was efficiently extracted (Fig. 2 A, lanes 10–12), whereas wild-type H1 and H1Δ were completely retained with the cellular membranes (Fig. 1 A, lanes 1–6). Both the glycosylated and the unglycosylated forms of H1ΔLeu were clearly membrane associated in this assay (Fig. 1 A, lanes 7–9). For alkaline extraction, the transfected and labeled cells were homogenized, exposed to pH 11.5, and then separated into soluble fraction (S) and membrane pellet (P) by centrifugation through a sucrose cushion (Fig. 2 B). Like H1Δ (Fig. 2 B, lanes 4–6), both forms of H1ΔLeu19 were recovered in the pellet fraction (Fig. 2 B, lanes 7–9). These results indicate that the unglycosylated form of H1ΔLeu19 is integrated in the membrane and argues against a reduced ability of the Leu19 domain to function as a signal sequence.

The disposition of the H1ΔLeu19 products with respect to the membrane was analyzed by protease digestion of homogenized labeled cells. The glycosylated 36-kD form of H1ΔLeu19 was resistant to trypsin, unless the membranes were disrupted with detergent (Fig. 2 C, lanes 4–6). This confirms its type II topology with only four NH_2-terminal residues exposed to the cytoplasm. In contrast, the unglycosylated 30-kD protein was sensitive to trypsin also in the absence of detergent, indicating that the COOH-terminal portion is cytoplasmically exposed and that the protein was inserted into the membrane in an N_{exo}/C_{cyt} orientation. Thus, exchanging the apolar signal domain of H1Δ by an artificial, even more hydrophobic sequence of identical length resulted in a significant fraction of the polypeptides inserting as type III proteins (Fig. 2 D).

If H1ΔLeu19 was NH_2-terminally extended by the NH_2-terminal hydrophilic portion of H1, the resulting construct H1Leu19 was expressed uniquely as a glycosylated protein of 40 kD like wild-type H1 (Figs. 1 A, lane 3, and 2 B, lanes 1–3), indicative of type II topology. This was confirmed by trypsin digestion, which generated a resistant COOH-terminal fragment of 36 kD (Fig. 2 C, lanes 1–3). In an internal position, the Leu₁₉ sequence thus did not affect the orientation of the protein.

Expression of all the constructs with a truncated NH_2-terminal domain produced a small amount of protein of ~33 kD, with an electrophoretic mobility intermediate between that of the twice glycosylated and the unglycosylated forms (indicated by asterisks in Figs. 1, 2, and 4). Upon endo H digestion, this material shifted to the position of the 30-kD unglycosylated form (Fig. 1 B); and in a protease protection assay, it was equally resistant as the
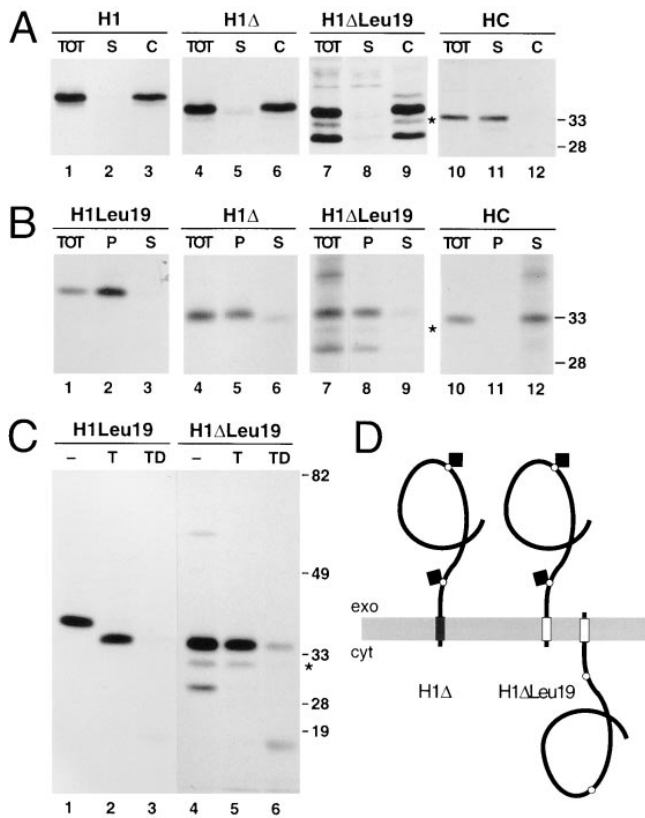
Figure 2. The unglycosylated form of mutant H1ΔLeu19 is inserted in the membrane in an inverted orientation. (*A*) Saponin extraction: cells were transfected with the indicated constructs, labeled, and extracted with 0.1% saponin. The saponin extract (*S*) and the remaining cells (*C*) were separately immunoprecipitated and analyzed by gel electrophoresis and fluorography. Untreated cells were solubilized and immunoprecipitated as a measure of the total material (*TOT*). (*B*) Alkaline extraction: transfected and labeled cells were homogenized and incubated at pH 11.5. The samples were either immunoprecipitated directly (*TOT*) or after separation into pellet (*P*) and supernatant fractions (*S*). (*C*) Protease protection: transfected cells were labeled, homogenized, and incubated without (−) or with trypsin (*T*) or with trypsin in the presence of detergent (*TD*). Immunoprecipitates were analyzed by SDS–gel electrophoresis and fluorography. The band indicated by an asterisk represents the partially glycosylated species. (*D*) Schematic representation of the membrane orientation of constructs H1Δ and H1ΔLeu19. The H1 sequence is shown in black and the Leu$_{19}$ domain as an empty rectangle. The glycosylation sites as indicated by open circles and N-linked glycans by closed squares.

twice glycosylated polypeptides (Fig. 2 *B*). These results indicate that the 33-kD species corresponds to type II polypeptides that were glycosylated only once. Incomplete glycosylation was generally not observed for constructs with the complete NH$_2$-terminal domain of H1. Most likely, glycosylation at the site near the membrane (position 79 of the wild-type sequence) is slightly influenced by the presence or absence of the NH$_2$-terminal domain.

Long Hydrophobic Signal Domains Near the NH$_2$ Terminus Promote N$_{exo}$/C$_{cyt}$ Orientation

Replacing the transmembrane domain of H1 by a se-

quence of 19 leucine residues increased its hydrophobicity without altering the physical length of the domain. To further explore the topogenic role of the hydrophobic domain, we prepared mutant constructs of H1 and H1Δ with artificial transmembrane domains consisting of 7–25 leucine residues. Thus the overall hydrophobicity and the length of the apolar domain were varied in parallel (Fig. 3).

For the constructs with a short NH$_2$-terminal domain of only four residues (H1ΔLeu#), the amount of the unglycosylated form increased with increasing number of leucines in the signal anchor and reached ∼80% of the total translation products for H1ΔLeu25 (Figs. 4 *B*, lanes *2–8*, and quantified in 5, *closed squares*). This form was essentially absent in the constructs with transmembrane domains of 16 leucines or fewer. Increasing length and/or hydrophobicity of the signal anchor sequence thus promotes translocation of the NH$_2$-terminal sequence across the ER membrane despite a negative charge difference Δ(C–N), when the signal is positioned near the NH$_2$ terminus of the protein. In contrast, the same constructs extended by the NH$_2$-terminal domain of H1 (H1Leu#) were all expressed exclusively as glycosylated type II membrane proteins of 40 kD (Figs. 4 *A*, lanes *2–8*, and 5, *open triangles*). Even a hydrophobic core of 25 consecutive leucine residues in an internal position did not affect the protein's topology.

Combination of a Long Hydrophobic Domain with a Reduced NH$_2$-Terminal Charge Produces Unique N$_{exo}$/C$_{cyt}$ Insertion

The best documented determinant for the topology of membrane spanning proteins are the charged residues flanking the transmembrane domain. To test whether the length/hydrophobicity and the charge difference of the signal anchor affect the insertion process in an additive manner, the NH$_2$-terminal arginine directly preceding the hydrophobic core of the signal anchor was mutated to glutamine. In the resulting constructs H1ΔQ and H1ΔQLeu#, the NH$_2$-terminal charge was reduced from +2 to +1, and the charge

Figure 3. Amino acid sequence of the signal–anchor domain of H1 mutant constructs. The hydrophobic transmembrane segments and their flanking sequences are listed. H1-4g and H1-4gLeu# are identical to H1-4 and H1-4Leu# except for the insertion of the tripeptide sequence MTM following asparagine-13 in the NH$_2$-terminal portion, which creates a potential glycosylation site.
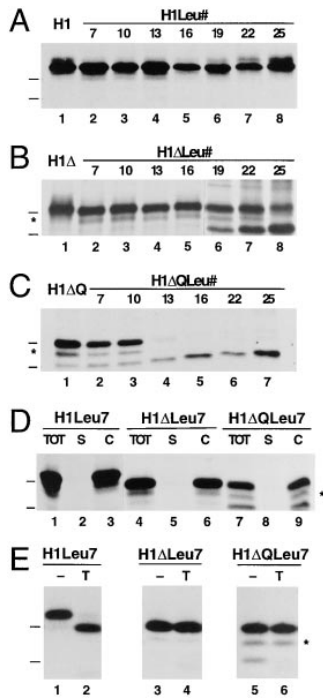
Figure 4. Effect of different hydrophobic domains on membrane insertion of H1, H1Δ, and H1ΔQ. The constructs H1 (*A*), H1Δ (*B*), and H1ΔQ (*C*) with the wild-type transmembrane domain of H1 (lane *1*) or with hydrophobic segments consisting of 7–25 leucine residues (lanes *2–8*) were expressed in COS-7 cells, labeled, immunoprecipitated, and analyzed by gel electrophoresis and fluorography. Membrane integration assessed by saponin extraction (*D*) and protease sensitivity (*E*) is shown for the constructs with the shortest hydrophobic segments of 7 leucines (see legend to Fig. 2). The position of the marker proteins of 29 and 35 kD are indicated.

difference changed from -3 to -2; this is still characteristic for many natural type II membrane proteins.

For H1ΔQ, which contains the wild-type transmembrane domain, this charge mutation resulted in ~25% of type III insertion (Fig. 4 *C*, lane *1*). For the H1ΔQLeu# constructs (Figs. 4 *C*, lanes *2–7*, and 5, *closed circles*), the mutation resulted in an increased fraction of unglycosylated N_{exo}/C_{cyt} form in comparison to the H1ΔLeu# series. Whereas the construct with the shortest apolar domain of only 7 leucines still mostly translocated the COOH-terminal sequence across the ER membrane (N_{cyt}/C_{exo}), those with the longest of 22 and 25 leucines almost exclusively exposed their COOH-terminal domain to the cytoplasm (N_{exo}/C_{cyt}).

These results show that both the flanking charges and the length/hydrophobicity of the signal anchor contribute to the topology decision. Reducing the NH₂-terminal positive charge and increasing the length/hydrophobicity of the signal anchor were sufficient to achieve unique type III (N_{exo}/C_{cyt}) insertion of an NH₂-terminally truncated form of H1.

Membrane association of the products of the H1Leu#, H1ΔLeu#, and H1ΔQLeu# constructs was confirmed by saponin extraction, as is shown for those with the shortest apolar segment of only 7 leucines in Fig. 4 *D*. The disposition of the COOH terminus was tested by protease sensitivity in Fig. 4 *E* for these constructs. The partial sensitivity of H1Leu7 to exogenous trypsin indicates a transmembrane topology. Since seven residues are too short to span a membrane as a helix, it is likely that the following eight residues in the sequence, which are largely polar but uncharged (Fig. 3), are also part of the membrane-embedded segment.

The observed ratio of type II and type III products after 30-min labeling with [³⁵S]methionine also depends on the degradation rate of the two species. As determined by

pulse–chase experiments for selected constructs (not shown), the type II species displayed half-times of degradation of up to 5 h. In contrast, type III products, which cannot form disulfide bonds in their cytosolic COOH-terminal portion, were more rapidly degraded with half-times of only ~0.5 h. For the extreme example of H1ΔLeu19 for which the degradation rates of the two forms differ by a factor of 10, the true insertion ratio of type II to type III products, corrected on the basis of measured degradation rates, were 59 to 41% in comparison to the apparent ratio of 66 to 34% after 30 min pulse labeling. The different stabilities of the two forms thus cause a bias in favor of the more stable type II forms. However, the observed ratios in our experimental setup are close estimates of the initial insertion ratios of the tested constructs.

A Long Internal Signal and Inverted Flanking Charges Are Sufficient to Completely Convert H1 from a Type II to a Type III Protein

In the constructs H1Leu#, described above, the leucine sequences had no effect on the topology of the proteins in the ER membrane. This could be due to the charge difference of -4.5 (as calculated according to Hartmann et al., 1989), which is more negative than in the H1ΔLeu# series. In addition, the presence of a sizable NH₂-terminal hydrophilic domain is likely to disfavor N_{exo}/C_{cyt} insertion. To test whether a long, hydrophobic signal domain in an internal position can influence insertion at all, we used H1-4, a previously described mutant of H1 in which the two NH₂-terminal and the two COOH-terminal charges flanking the transmembrane domain had been mutated to residues of opposite charge (Beltzer et al., 1991; Fig. 3). Even though the charge difference of this mutant was strongly positive ($+5$), only half of the polypeptides were inserted as type III proteins (Fig. 6 *A*, lane *1*). In this mutant con-
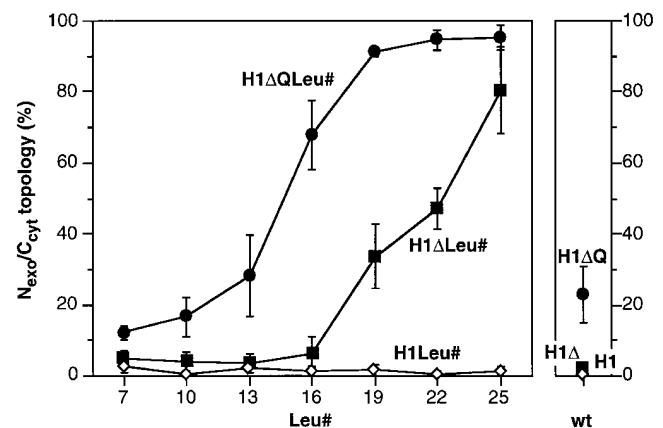

Figure 5. Topology of signal–anchor mutants. Insertion experiments, including those shown in Fig. 4, were quantified by densitometric scanning of fluorographs. The fraction of unglycosylated protein, i.e., with N_{exo}/C_{cyt} orientation, is presented as percent of the total of all forms. The values for constructs with polyleucine domains are plotted as a function of the number of leucines in this segment (*Leu*#). Corresponding constructs with the transmembrane segment of wild-type H1 are shown to the right (*wt*). The values represent the mean of three or more experiments with standard deviations.
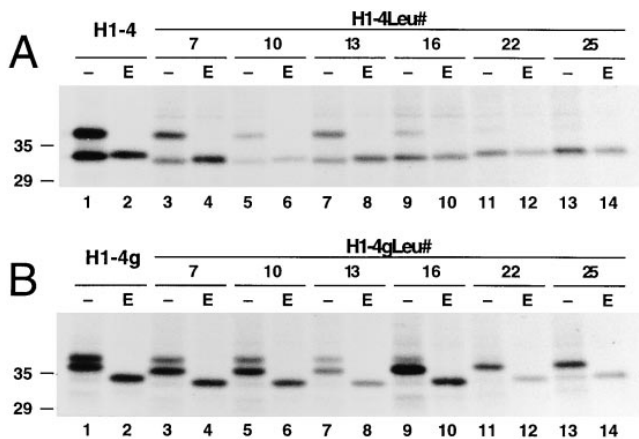
Figure 6. Effect of different hydrophobic domains on membrane insertion of H1-4 and H1-4g. The constructs H1-4 (*A*) and H1-4g (*B*) with the wild-type transmembrane domain of H1 (lanes *1* and *2*) or with hydrophobic segments consisting of 7–25 leucine residues (lanes *3–14*) were expressed in COS-7 cells, labeled, and immunoprecipitated. Samples were treated without ($-$) or with endoglycosidase H (*E*) before analysis by gel electrophoresis and fluorography.
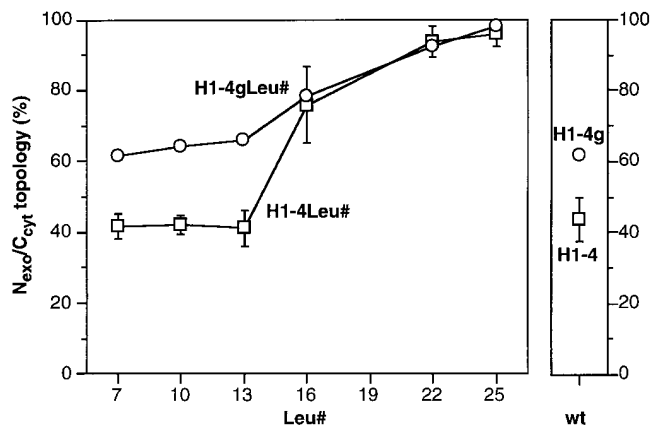

Figure 7. Quantitation of the topology of the constructs H1-4Leu# and H1-4gLeu#. The insertion experiments including those shown in Fig. 6 were quantified by densitometric scanning of the fluorographs. The fraction of once glycosylated protein, i.e., with N_{exo}/C_{cyt} orientation, is presented as percent of the total of all forms as described in the legend to Fig. 5. The values for H1-4Leu# represent the mean of three or more experiments with standard deviations; those for H1-4gLeu# represent single determinations performed simultaneously.

struct, the 19 hydrophobic amino acids of the signal anchor were replaced by 7–25 consecutive leucine residues, thus generating the mutant series H1-4Leu#.

As is shown in Fig. 6 *A* and quantified in Fig. 7 (*open squares*), transmembrane segments >13 leucine residues showed increased translocation of the NH_2 terminus and were thus topogenically active even in an internal position. Virtually unique type III insertion was observed for H1-4Leu22 and H1-4Leu25. Whereas the longer leucine sequences favored type III insertion, shorter ones did not shift the balance towards type II insertion. H1-4Leu13, H1-4Leu10, and H1-4Leu7 did not insert as type II proteins to any larger extent than H1-4. Shortening the oligo–leucine domain more effectively increased the fraction of N_{cyt}/C_{exo} polypeptides in constructs with the signal domain located close to the NH_2 terminus than further inside the protein.

To obtain direct evidence for the translocation of the NH_2 terminus in those forms that were not glycosylated in their COOH-terminal domain, we constructed the series H1-4gLeu# which contains an additional glycosylation site in the NH_2-terminal portion. Like for H1-4g, the corresponding construct with the original apolar sequence of H1 (Beltzer et al., 1991), two types of products were expressed in transfected COS cells (Fig. 6 *B*): polypeptides twice glycosylated in the COOH-terminal portion and polypeptides with electrophoretic mobility intermediate between the twice glycosylated and the deglycosylated proteins, indicative of a single glycosylation in the NH_2-terminal domain. For the shorter polyleucine transmembrane domains, the NH_2-terminal glycosylation site resulted in an increased type III insertion, as has previously been observed for H1-4g versus H1-4. A likely explanation is that the tripeptide insertion to generate the glycosylation site affected the folding of the NH_2-terminal domain, thus facilitating its translocation through the membrane.

Discussion

The best established determinant of the orientation of signal–anchor sequences is the distribution of charged residues flanking the hydrophobic transmembrane domain (Dalbey, 1990; von Heijne, 1994; Spiess, 1995). In addition, the NH_2-terminal hydrophilic portion and its folding properties were shown to play a topogenic role, because the polypeptide needs to be unfolded for translocation through the membrane (Denzer et al., 1995). Deletion constructs by Sakaguchi et al. (1992) and Sato et al. (1990) also suggested an influence of the hydrophobic segment on the topology. Whereas these criteria were shown by site-directed mutagenesis to direct the insertion process, they were not generally sufficient to generate a unique topology in recombinant proteins. In the present study we have systematically analyzed the contribution of the hydrophobic segment of the signal on topogenesis. By combining these different determinants it was possible to generate uniquely one or the other orientation in the membrane.

In the context of the wild-type sequence of H1, where the charge distribution and the sizable NH_2-terminal domain favor translocation of the COOH-terminal sequence, replacing the transmembrane segment by sequences of 7–25 consecutive leucines (H1Leu#) did not affect the insertion behavior. Truncation of the NH_2-terminal domain (H1Δ) alone also had no effect. However, when in addition the transmembrane segment was exchanged by a stretch of 19 or more leucines, significant fractions of the polypeptides inserted in the opposite N_{exo}/C_{cyt} orientation (H1ΔLeu19/22/25). This fraction increased with the length of the polyleucine segment. The effect was further enhanced by reducing the positive charge at the NH_2 terminus. Even though the NH_2 terminus still carried a net positive charge, significant N_{exo}/C_{cyt} products were generated by all constructs, including that containing the wild-type transmem-

brane segment of H1 (H1ΔQ) as well as those containing stretches of 7–25 leucines (H1ΔQLeu#). Constructs with polyleucine sequences of 19 or more residues inserted essentially with unique type III topology.

The topogenic effect of long polyleucine transmembrane domains is not limited to proteins with very short NH_2-terminal hydrophilic portions. An inverted charge distribution flanking the apolar segment of H1 resulted only in a partial change of orientation in the membrane (H1-4). In combination with a segment of ≥16 leucines, type III insertion was significantly enhanced, and with 22 or 25 leucines, unique type III insertion was achieved.

The topology decision of signal–anchor sequences is thus a result of the combined effects of at least three different features: the charge distribution in the vicinity of the signal sequence, the presence or absence of folded NH_2-terminal segments, and the length and hydrophobicity of the apolar segment of the signals. Each of these features may be involved in additional functions besides membrane insertion. For example, charged residues close to the membrane surface might be involved in functional interactions of the protein with particular lipids, and the length of the membrane-spanning domain has been proposed to affect intracellular protein sorting (Munro, 1995; Parks, 1996). Multiple determinants for protein orientation allow variability in each of them without compromising insertion in a unique topology. This may explain how the exceptional proteins that do not conform to the charge distribution rule are correctly inserted. Unlike the charge distribution, the length and the total hydrophobicity of the transmembrane segment do not clearly correlate with the orientation of natural signal anchor proteins. This suggests that the flanking charges are the most general criterion in determining signal anchor orientation.

Changing the number of leucine residues in the signal anchor sequence of a protein varies both the length of this domain and its total hydrophobicity. It is therefore not possible to distinguish between these two factors with respect to their influence on topogenesis. The wild-type transmembrane segment of H1 has the same length as the Leu_{19} sequence and the same hydrophobicity as the Leu_{16} sequence (as estimated by summing up the hydropathy indices of each residue; Kyte and Doolittle, 1982). According to our results, NH_2-terminal translocation is promoted by these transmembrane segments in the order $Leu_{19} > Leu_{16} >$ wild type. It is thus neither the length of the apolar segment of the signal nor its hydrophobicity alone that is responsible for the observed effects on insertion behavior. It is obvious that two different sequences of the same length also differ in other properties, e.g., the shape of the molecule and the propensity to assume α-helical conformation.

Cleaved signals of secretory and type I membrane proteins resemble type II signal anchors, since they promote translocation of the COOH-terminal sequence through the membrane and typically have a more positive NH_2 terminus. Cleaved signals differ from signal anchors by consistently shorter apolar segments (7–17 vs. 17–28 residues, respectively; Nilsson et al., 1994). Since they generally lack sizable hydrophilic NH_2-terminal domains, translocation of the NH_2 terminus is not sterically hindered. Our results are consistent with a role of short apolar signal domains in keeping the NH_2 terminus cytosolic and can provide an explanation for those eukaryotic signal sequences that do not conform to the "positive inside" and the charge difference rules. Short polyleucine stretches in an NH_2-terminal position (constructs H1ΔLeu# and H1ΔQLeu#) promoted translocation of the COOH-terminal sequence. In contrast, short polyleucine sequences in an internal position (constructs H1Leu# and H1-4Leu#) had only a limited effect on the topology of the proteins. H1-4Leu7, H1-4Leu10, and H1-4Leu13 inserted to the same extent of ~50% with N_{exo}/C_{cyt} orientation. The translocation machinery thus seems to recognize the hallmarks of a cleavable signal: a short hydrophobic segment close to the NH_2 terminus. If these characteristics are present, there is a strong preference for COOH-terminal translocation. Studies by Nilsson et al. (1994) have provided evidence that N_{cyt}/C_{exo} signals with short apolar segments are positioned differently in the translocation machinery. The minimal distance between the apolar segment of a signal and a potential glycosylation site required to reach the active site of oligosaccharyl transferase was found to be different for short apolar segments of up to 14 leucine residues than for longer apolar segments. TRAM may be the component of the translocation machinery involved in distinguishing these signals, since in reconstitution experiments TRAM independence appeared to correlate both with the size of the NH_2-terminal hydrophilic portion of a signal and with the length of the hydrophobic segment (Voigt et al., 1996).

References

Andrews, D.W., J.C. Young, L.F. Mirels, and G.J. Czarnota. 1992. The role of the N-region in signal sequence and signal-anchor function. *J. Biol. Chem.* 267:7761–7769.

Beltzer, J.P., K. Fiedler, C. Fuhrer, I. Geffen, C. Handschin, H.P. Wessels, and M. Spiess. 1991. Charged residues are major determinants of the transmembrane orientation of a signal-anchor sequence. *J. Biol. Chem.* 266:973–978.

Cullen, B.R. 1987. Use of eukaryotic expression technology in the functional analysis of cloned genes. *Methods Enzymol.* 152:684–704.

Dalbey, R.E. 1990. Positively charged residues are important determinants of membrane protein topology. *Trends Biochem. Sci.* 15:253–257.

Denzer, A.J., C.E. Nabholz, and M. Spiess. 1995. Transmembrane orientation of signal-anchor proteins is affected by the folding state but not the size of the N-terminal domain. *EMBO (Eur. Mol. Biol. Organ.) J.* 14:6311–6317.

Ellis, L., E. Clauser, D.O. Morgan, M. Edery, R.A. Roth, and W.J. Rutter. 1986. Replacement of insulin receptor tyrosine residues 1162 and 1163 compromises insulin-stimulated kinase activity and uptake of 2-deoxyglucose. *Cell.* 45:721–732.

Gilmore, R., and G. Blobel. 1985. Translocation of secretory proteins across the microsomal membrane occurs through an environment accessible to aqueous perturbants. *Cell.* 42:497–505.

Hartmann, E., T.A. Rapoport, and H.F. Lodish. 1989. Predicting the orientation of eukaryotic membrane-spanning proteins. *Proc. Natl. Acad. Sci. USA.* 86:5786–5790.

Kyte, J., and R.F. Doolittle. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157:105–132.

Monier, S., P. Van Luc, G. Kreibich, D.D. Sabatini, and M. Adesnik. 1988. Signals for the incorporation and orientation of cytochrome P450 in the endoplasmic reticulum membrane. *J. Cell Biol.* 107:457–470.

Munro, S. 1995. An investigation of the role of transmembrane domains in Golgi protein retention. *EMBO (Eur. Mol. Biol. Organ.) J.* 14:4695–4704.

Nilsson, I., P. Whitley, and G. von Heijne. 1994. The COOH-terminal ends of internal signal and signal-anchor sequences are positioned differently in the ER translocase. *J. Cell Biol.* 126:1127–1132.

Parks, G.D. 1996. Differential effects of changes in the length of a signal/anchor

domain on membrane insertion, subunit assembly, and intracellular transport of a type II integral membrane protein. *J. Biol. Chem.* 271:7187–7195.

Parks, G.D., and R.A. Lamb. 1991. Topology of eukaryotic type-II membrane proteins—importance of N-terminal positively charged residues flanking the hydrophobic domain. *Cell.* 64:777–787.

Parks, G.D., and R.A. Lamb. 1993. Role of NH$_2$-terminal positively charged residues in establishing membrane protein topology. *J. Biol. Chem.* 268: 19101–19109.

Sakaguchi, M., R. Tomiyoshi, T. Kuroiwa, K. Mihara, and T. Omura. 1992. Functions of signal and signal-anchor sequences are determined by the balance between the hydrophobic segment and the N-terminal charge. *Proc. Natl. Acad. Sci. USA.* 88:16–19.

Sato, T., M. Sakaguchi, K. Mihara, and T. Omura. 1990. The amino-terminal structures that determine topological orientation of cytochrome-P-450 in microsomal membrane. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:2391–2397.

Schmid, S.R., and M. Spiess. 1988. Deletion of the amino-terminal domain of asialoglycoprotein receptor H1 allows cleavage of the internal signal sequence. *J. Biol. Chem.* 263:16886–16891.

Spiess, M. 1995. Heads or tails—what determines the orientation of proteins in the membrane. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 369:76–79.

Spiess, M., and H.F. Lodish. 1986. An internal signal sequence: the asialoglycoprotein receptor membrane anchor. *Cell.* 44:177–185.

Spiess, M., and C. Handschin. 1987. Deletion analysis of the internal signal-anchor domain of the human asialoglycoprotein receptor H1. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:2683–2691.

Spiess, M., A.L. Schwartz, and H.F. Lodish. 1985. Sequence of human asialoglycoprotein receptor cDNA. An internal signal sequence for membrane insertion. *J. Biol. Chem.* 260:1979–1982.

Szczesna-Skorupa, E., and B. Kemper. 1989. NH$_2$-terminal substitutions of basic amino acids induce translocation across the microsomal membrane and glycosylation of rabbit cytochrome P450IIC2. *J. Cell Biol.* 108:1237–1243.

Szczesna-Skorupa, E., N. Browne, D.A. Mead, and B. Kemper. 1988. Positive charges at the NH$_2$ terminus convert the membrane-anchor signal peptide of cytochrome P-450 to a secretory peptide. *Proc. Natl. Acad. Sci. USA.* 85: 738–742.

Voigt, S., B. Jungnickel, E. Hartmann, and T.A. Rapoport. 1996. Signal sequence-dependent function of the TRAM protein during early phases of protein transport across the endoplasmic reticulum membrane. *J. Cell Biol.* 134:25–35.

von Heijne, G. 1986. The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:3021–3027.

von Heijne, G. 1989. Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. *Nature (Lond.).* 341:456–458.

von Heijne, G. 1994. Membrane proteins: from sequence to structure. *Annu. Rev. Biophys. Biomol. Struc.* 23:167–192.

von Heijne, G., and Y. Gavel. 1988. Topogenic signals in integral membrane proteins. *Eur. J. Biochem.* 174:671–678.

von Heijne, G., and C. Manoil. 1990. Membrane proteins: from sequence to structure. *Protein Eng.* 4:109–112.

Wahlberg, J.M., I. Geffen, F. Reymond, T. Simmen, and M. Spiess. 1995. *trans*-Golgi retention of a plasma membrane protein: mutations in the cytoplasmic domain of the asialoglycoprotein receptor subunit H1 result in *trans*-Golgi retention. *J. Cell Biol.* 130:285–297.

Walter, P., and A.E. Johnson. 1994. Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.* 10: 87–119.