Published October 18, 1999

Glycosylation Can Influence Topogenesis of Membrane Proteins and Reveals Dynamic Reorientation of Nascent Polypeptides within the Translocon

Veit Goder, Christoph Bieri, and Martin Spiess

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

Abstract. The topology of multispanning membrane proteins in the mammalian endoplasmic reticulum is thought to be dictated primarily by the first hydrophobic sequence. We analyzed the in vivo insertion of a series of chimeric model proteins containing two conflicting signal sequences, i.e., an NH_2-terminal and an internal signal, each of which normally directs translocation of its COOH-terminal end. When the signals were separated by more than 60 residues, linear insertion with the second signal acting as a stop-transfer sequence was observed. With shorter spacers, an increasing fraction of proteins inserted with a translocated COOH terminus as dictated by the second signal. Whether this resulted from membrane targeting via the second signal was tested by measuring the targeting efficiency of NH_2-terminal signals followed by polypeptides of different lengths. The results show that targeting is mediated predominantly by the first signal in a protein. Most importantly, we discovered that glycosylation within the spacer sequence affects protein orientation. This indicates that the nascent polypeptide can reorient within the translocation machinery, a process that is blocked by glycosylation. Thus, topogenesis of membrane proteins is a dynamic process in which topogenic information of closely spaced signal and transmembrane sequences is integrated.

Key words: endoplasmic reticulum • glycosylation • integral membrane protein • signal recognition particle • signal sequence

In higher eukaryotic cells, most membrane and secretory proteins are targeted by a hydrophobic signal sequence to the ER membrane in a cotranslational process involving the signal recognition particle (SRP)[1] and the SRP receptor (for reviews see Walter and Johnson, 1994; High and Laird, 1997). Membrane insertion and translocation of polypeptides is mediated by the Sec61 complex, which forms a gated pore (Görlich and Rapoport, 1993; Hanein et al., 1996; Hamman et al., 1997) and specifically recognizes signal sequences (Mothes et al., 1998). Additional components that contribute to the insertion and translocation process are the ribosome, which binds to the translocation pore and largely seals it towards the cytosol, TRAM (translocating chain-associating membrane protein) (Görlich and Rapoport, 1993; Voigt et al., 1996; Hegde et al., 1998), and BiP, a chaperone that binds to the unfolded polypeptide on the lumenal side of the membrane and drives rapid transfer of hydrophilic sequences through the membrane (Brodsky et al., 1995; Matlack et al., 1999). The signal peptidase complex and the oligosaccharyl transferase complex are also closely associated with the translocon, since they are able to cotranslationally modify the translocating polypeptide (Habener et al., 1976; Rothman and Lodish, 1977; Maurer and McKean, 1978). To what extent, and how these components influence the insertion process and topogenesis is still largely unknown.

A signal sequence may be inserted into the translocon and subsequently the membrane, with either its NH_2 or its COOH terminus facing the cytosol (Spiess, 1995). Cleaved signals and signal-anchor sequences of type II membrane proteins assume an N_{cyt}/C_{exo} orientation, whereas the reverse signal-anchors of type III membrane proteins insert with an N_{exo}/C_{cyt} orientation. Several features determine which end of the signal is translocated across the membrane. The most established feature is the distribution of charged residues flanking the hydrophobic core of the signal. In natural proteins, positive charges are statistically enriched on the side remaining cytosolic (the "positive-inside rule" of von Heijne, 1989; and the "charge differ-

Address correspondence to Martin Spiess, Department of Biochemistry, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland. Tel.: 41-61-2672164. Fax: 41-61-2672149. E-mail: spiess@ubaclu.unibas.ch

1. *Abbreviations used in this paper:* ASGP, asialoglycoprotein; endo H, endo-β-D-N-acetyl glucosaminidase H; SRP, signal recognition particle.

Published October 18, 1999

ence rule" of Hartmann et al., 1989). The flanking charges were shown by mutagenesis to be important for orienting the signal sequence; however, the mutant proteins did not strictly follow the charge rules and often inserted with mixed topologies (Beltzer et al., 1991; Parks and Lamb, 1991; Andrews et al., 1992), indicating that additional features codetermine the insertion process. These features include the length and hydrophobicity of the signal sequence. Long, hydrophobic sequences favor translocation of the NH_2 terminus (Sakaguchi et al., 1992; Wahlberg and Spiess, 1997; Eusebio et al., 1998; Harley et al., 1998). Furthermore, rapid folding of the NH_2-terminal hydrophilic segment preceding the apolar signal sequence has been shown to inhibit NH_2-terminal translocation (Denzer et al., 1995).

Multispanning membrane proteins are believed to be targeted to the ER membrane by their first hydrophobic signal, which is either a cleaved signal peptide or the first transmembrane segment of the protein. The subsequent transmembrane domains insert with alternating orientations. According to the simplest model, the initial signal defines its own orientation as well as the orientations of all subsequent transmembrane segments. The latter do not require any additional information, but will simply follow the lead of the first signal. Evidence for this "linear insertion model" (initially proposed by Blobel, 1980) has been provided by in vitro studies using chimeric proteins with two to four transmembrane segments separated by ~50–200 residues from each other (Wessels and Spiess, 1988; Lipp et al., 1989). The results showed that signal-anchors insert as stop-transfer sequences depending only on their position relative to the preceding hydrophobic segments.

However, statistics show that internal transmembrane domains of natural multispanning proteins also follow the charge rules, although less stringently than the most NH_2-terminal signal (von Heijne, 1989). This suggests that the subsequent transmembrane segments also possess topological information. In support of this, insertion of clusters of positive charges into short exoplasmic loops of model proteins resulted in individual hydrophobic domains not spanning the membrane at all ("frustrated" topologies; Gravelin et al., 1997). Similarly, inversion of the charge difference of the first signal of the glucose transporter Glut1 did not affect the topology of the rest of the molecule, but prevented insertion of the first signal (Sato et al., 1998). These studies showed that charged residues could also affect the insertion process at internal positions, and suggested that multispanning proteins contain topogenic information throughout their sequence.

Interestingly, the hydrophilic sequences separating transmembrane segments in natural proteins are frequently much shorter than those used in the studies supporting the linear insertion model. Here, we systematically analyzed the topogenic contribution of an internal hydrophobic sequence in relation to its distance from an NH_2-terminal signal sequence. Depending on the characteristics of the signals used, nonlinear insertion was observed only with spacer sequences of <80 residues. Topogenesis was found to be affected by glycosylation, revealing surprising conformational dynamics of the polypeptide during the orientation process within the translocon.

Materials and Methods

DNA Constructs

All constructs were made by site-directed mutagenesis using PCR with Vent polymerase (New England Biolabs). The final constructs for in vivo expression were subcloned into the vector pECE (Ellis et al., 1986) and verified by sequencing. To construct chimeric proteins with conflicting signals, the coding sequence for the 98 NH_2-terminal residues of influenza hemagglutinin (strain A/Jap/305/57 [H2N2]) (Gething and Sambrook, 1981) was fused via a linker sequence (Gly-Ser) in front of codon 4 of the cDNA of the asialoglycoprotein (ASGP) receptor subunit H1 (Spiess et al., 1985). The codon for methionine in position 95 was mutated to one for isoleucine to eliminate potential internal translation initiation. The resulting sequence encoded a chimeric protein with two ER targeting signals separated by a spacer of 128 residues. By PCR with this sequence as a template, 5′ truncations were prepared consisting of a 5′ Bgl II site, the COOH-terminal 100, 80, 60, 40, or 20 residues of the spacer sequence, and the signal and the COOH-terminal remainder of H1. In front of these sequences, the cDNAs of the signal sequences of hemagglutinin (H), human prepro-vasopressin-neurophysin II (V), bovine preprolactin (P), and the ASGP receptor H1 (with the wild-type hydrophilic sequence, A, or with a 30 residue deletion, ΔA) (Beltzer et al., 1991), as shown in Fig 1 C, were ligated by BamH I/Bgl II (Gly-Ser) to produce five series of plasmids. The constructs (schematically shown in Fig. 1 A) were named to indicate the origin of the first signal, the length of the hydrophilic spacer sequence, and the origin of the second signal (e.g., V40A). The series without the second signal (H20− etc.) was generated by deleting the entire coding region of the H1 signal from the premade constructs harboring two conflicting signals (H20A etc.) by PCR.

To generate potential *N*-glycosylation sites in the spacer sequence, the codons for IleThrLeu (ITL), MTM, or GS were inserted into the spacer sequence after the codon for Asn-13 of the H1 sequence of the spacer sequence between the signals. To destroy the glycosylation site NITL, the codon for Thr was mutated to one for Asn. To test a different spacer sequence with a natural glycosylation site (NFT), the residues 26–59 of V40A were replaced by residues 61–100 of H1, resulting in V45(NTF)A. The glycosylation site was mutated to QTF in V45(QTF)A.

To generate the truncated constructs H55, H75, H95, and H115, the 5′ sequences of constructs H40(NIT)A, H60(NIT)A, H80(NIT)A, and H100(NIT)A encoding the hemagglutinin signal and the spacer sequences were fused via two methionine codons and a Kpn I site encoding Gly-Thr to the 14 COOH-terminal residues of H1, which are recognized by the antibody anti-H1C. Constructs V55, V75, V95, and V115 were constructed accordingly.

In Vivo Expression and Analysis of Receptor Constructs

Cell culture reagents were from Life Technologies, Inc. COS-1 cells were grown in modified MEM supplemented with 10% FCS, 2 mM L-glutamine, 100 U/ml penicillin, and 100 μg/ml streptomycin at 37°C with 7.5% CO_2. Transient transfection was performed with lipofectin (Life Technologies, Inc.) according to the manufacturer's instructions in 6-well clusters, and the cells were processed the second day after transfection. For in vivo labeling, transfected cells were incubated for 40 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, washed twice with PBS, lysed, and 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 (Amersham Pharmacia Biotech) and analyzed by SDS-PAGE and fluorography. For deglycosylation, the immune complexes were released from protein A–Sepharose by boiling in 50 mM Na citrate, pH 6, 1% SDS, and incubated with 1 mU endo-β-D-N-acetyl glucosaminidase H (endo H) for 5 h at 37°C, before gel electrophoresis. Quantitation was performed using a PhosphorImager (Molecular Dynamics, Inc.). For constructs with the vasopressin signal, the results were corrected for the loss of a methionine within the signal peptide upon signal cleavage.

Alkaline extraction was performed as described previously (Gilmore and Blobel, 1985; Wessels et al., 1991). In brief, cells were homogenized and incubated in Hepes buffer, pH 11.5, for 15 min on ice. One half of each sample was centrifuged through a sucrose cushion, and the membrane pellet and the supernatant were immunoprecipitated separately and

Published October 18, 1999

compared with the immunoprecipitate of the untreated second half of the sample (total). In vitro translation was performed with standard procedures using rabbit reticulocyte lysate and dog pancreas microsomes (Wessels et al., 1991). Protease protection assay was performed as described by Wahlberg and Spiess (1997), except that the scraped cells were homogenized in a cell cracker (from EMBL) as described by Leitinger et al. (1995).

Results

Insertion of Polypeptides with Conflicting Signals

To characterize the in vivo topogenesis of membrane proteins and to test the linear insertion model, we constructed a series of chimeric polypeptides containing an NH_2-terminal cleaved signal sequence and an internal type II signal-anchor sequence. An NH_2-terminal portion of influenza virus hemagglutinin with its cleavable signal sequence was fused to the NH_2 terminus of subunit H1 of the ASGP receptor, a type II membrane protein. By successive truncation of the hydrophilic sequence separating the two hydrophobic signals, a series of constructs with spacers of ~20, 40, 60, 80, and 100 residues was created (H20A, H40A, etc.) (Fig. 1 A). In the wild-type context, each of the two signals efficiently directs translocation of its COOH-terminal end across the membrane. According to the linear insertion model, the NH_2-terminal hemagglutinin signal will target the polypeptide to the ER membrane and induce translocation of the spacer sequence. The internal H1 signal will then enter the translocon as a stop-transfer sequence and halt further translocation, leaving the COOH terminus of the polypeptide, including the only two potential sites for *N*-glycosylation, in the cytosol (Fig. 1 B, left). If, in contrast, the topogenic information of the internal signal was dominant, the COOH-terminal portion of H1 would be translocated and glycosylated, whereas the NH_2-terminal signal would be forced to insert with the opposite N_{exo}/C_{cyt} orientation or would fail to insert at all (Fig. 1 B, right).

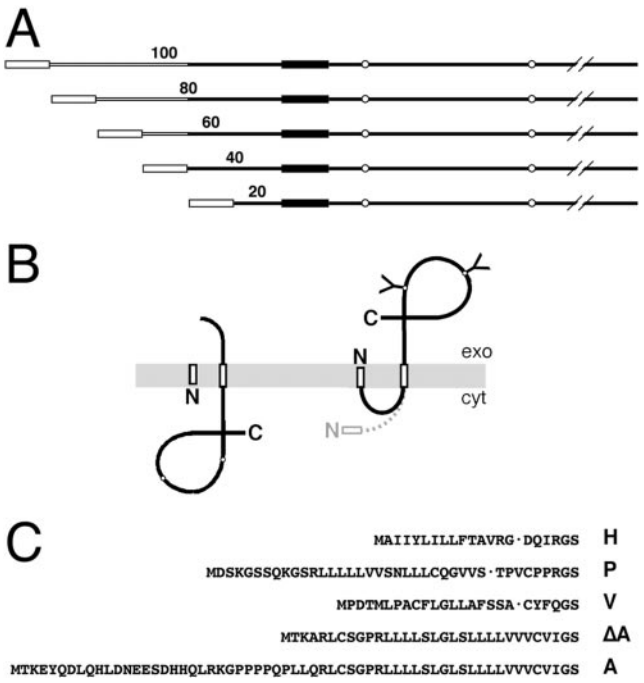

Figure 1. Chimeric proteins and their potential insertion patterns. (A) Schematic representation of the fusion proteins consisting of an NH_2-terminal signal sequence (open boxes), a spacer sequence of 20–100 residues, the internal signal-anchor (filled boxes) and the entire COOH-terminal domain of the ASGP receptor H1 with two potential *N*-glycosylation sites (open circles). Sequences derived from hemagglutinin and the ASGP receptor are shown as open and filled lines, respectively. (B) Possible topologies of the chimeras in the ER membrane. The linear insertion model predicts a cleaved NH_2-terminal signal, a translocated spacer, and a cytosolic, unglycosylated COOH terminus (left). Alternatively, the COOH terminus is translocated and glycosylated, the spacer remains cytosolic, and the NH_2-terminal signal is uncleaved and either integrated or cytosolic (right). cyt, cytoplasmic side; exo, exoplasmic side. (C) The sequences of the different NH_2-terminal signals used. H, hemagglutinin; P, preprolactin; V, prepro-vasopressin-neurophysin II; A and ΔA, ASGP receptor H1 with its complete and a truncated hydrophilic, cytoplasmic domain, respectively. Signal cleavage sites are indicated by a dot.

The constructs were transiently expressed in COS-1 cells, labeled for 30 min with [^{35}S]methionine, immunoprecipitated, and analyzed by SDS-gel electrophoresis and fluorography (Fig. 2 A). Consistent with the linear insertion model, the constructs H100A and H80A yielded single products which were not glycosylated, as revealed by their resistance to deglycosylation by endo H. However, the constructs with shorter spacers between the signals, H60A, H40A, and H20A, produced an increasing fraction of an additional species of lower electrophoretic mobility. The shift in mobility was due to glycosylation, as shown by endo H digestion. Both the glycosylated and the unglycosylated products were stably integrated into the ER membrane, since they were recovered in the membrane pellet after alkaline extraction (Fig. 2 B). Under the same conditions, a control protein lacking a membrane anchor (H20−), was completely extracted into the soluble fraction. The results demonstrate that the internal signal, when separated by <80 residues from the NH_2-terminal signal, did not always function as a stop-transfer sequence, but was able to direct translocation of its COOH terminus across the membrane and to anchor the protein in a N_{cyt}/C_{exo} orientation.

Whereas the sequence context of the internal H1 signal is the same in all constructs, that of the hemagglutinin signal is always different. As a result, the flanking charges, which are known to influence the topological preference of signal sequences, are not constant. The charge difference between the 15 COOH-terminal and the 15 NH_2-terminal flanking residues, Δ(C–N), as defined by Hartmann et al. (1989), is −0.5 for the hemagglutinin signal in its wild-type context. Δ(C–N) is more negative in H40A (−1.5), H60A (−1), and H80A (−1). Thus, the observed insertion pattern cannot be explained simply by an effect of the flanking charges on the orientation of the NH_2-terminal signal. Only in H20A, the construct with the shortest spacer, in which the flanking segments of the two signals largely overlap, is Δ(C–N) positive for the hemagglutinin signal (+2) and might favor translocation of the NH_2 terminus. To test experimentally the insertion behavior of the hemagglutinin signal in the different chimeras without the potential interference by a second signal, a series of

Published October 18, 1999

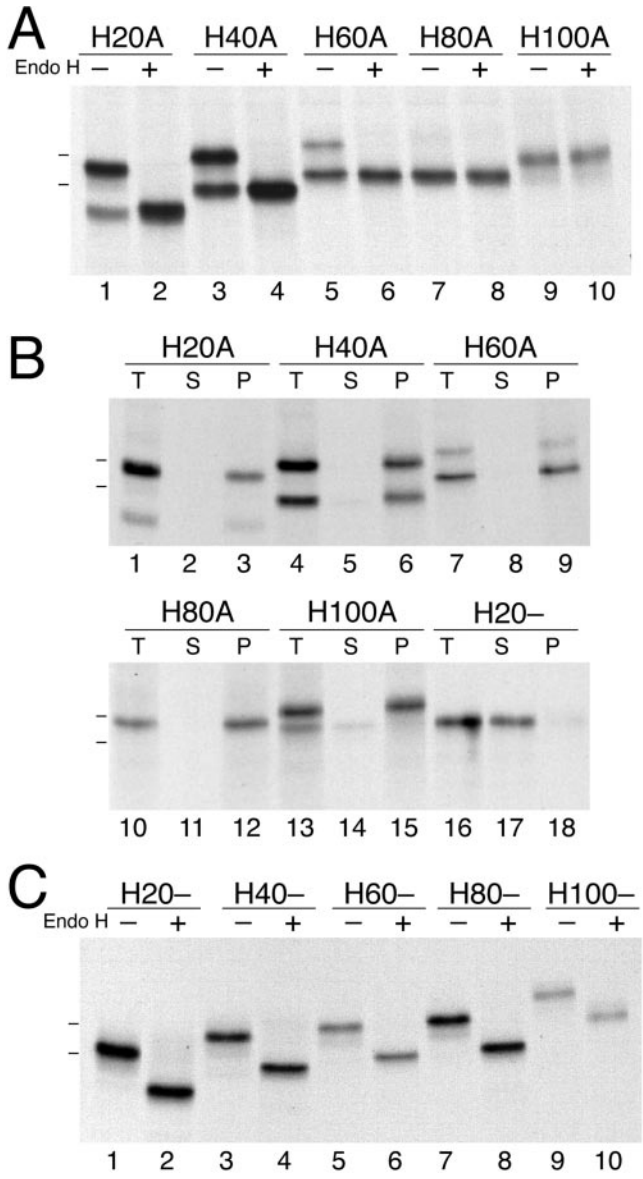

Figure 2. ER insertion of chimeric proteins containing the NH_2-terminal signal of hemagglutinin and the internal signal of the ASGP receptor H1. (A) Transfected COS cells were labeled with [^{35}S]methionine for 30 min. The chimeric proteins were immunoprecipitated and analyzed by SDS-gel electrophoresis and fluorography after deglycosylation with endo H (+) or without treatment (−). (B) Membrane integration tested by alkaline extraction. Transfected COS cells were metabolically labeled and scraped. One half was subjected to alkaline extraction. The resulting supernatant (S) and pellet fractions (P), as well as the other half (the total, T) were then subjected to immunoprecipitation and analyzed by gel electrophoresis and fluorography. As a soluble control, H20−, which lacks a second, internal signal and is secreted into the ER lumen, was analyzed. (C) Translocation efficiency of the hemagglutinin signal in the absence of a second signal sequence. From the constructs H#A, the segment encoding the internal signal of the ASGP receptor H1 was deleted. The resulting H#− constructs were expressed in COS cells, metabolically labeled, and analyzed by immunoprecipitation, gel electrophoresis, and fluorography. The position of molecular weight markers of 35 and 40 kD are indicated.

constructs was made lacking the hydrophobic part of the internal signal (H20−, H40−, etc.). In all cases, the COOH-terminal sequence was completely translocated and glycosylated (Fig. 2 C), demonstrating that different flanking sequences did not affect the functionality of the hemagglutinin signal as a secretory signal. It required the competition by the second signal to cause insertion of H20A, H40A, and H60A with a translocated COOH terminus.

The Ratio of Topologies Depends on the Characteristics of the First Signal

To determine the contribution of the NH_2-terminal signal on topogenesis, the hemagglutinin signal in the H#A constructs was replaced by the cleavable signals from preprolactin (P#A) or the vasopressin precursor protein (V#A), or by the signal-anchor sequence of the ASGP receptor H1, either with (A#A) or without the complete 40-residue NH_2-terminal, hydrophilic domain (ΔA#A) (Fig. 1 C). The constructs were expressed in COS cells and analyzed by gel electrophoresis and fluorography (Fig. 3). The fraction of glycosylated products corresponding to the topology with a translocated COOH terminus was quantified (Fig. 4). The constructs with the preprolactin signal yielded

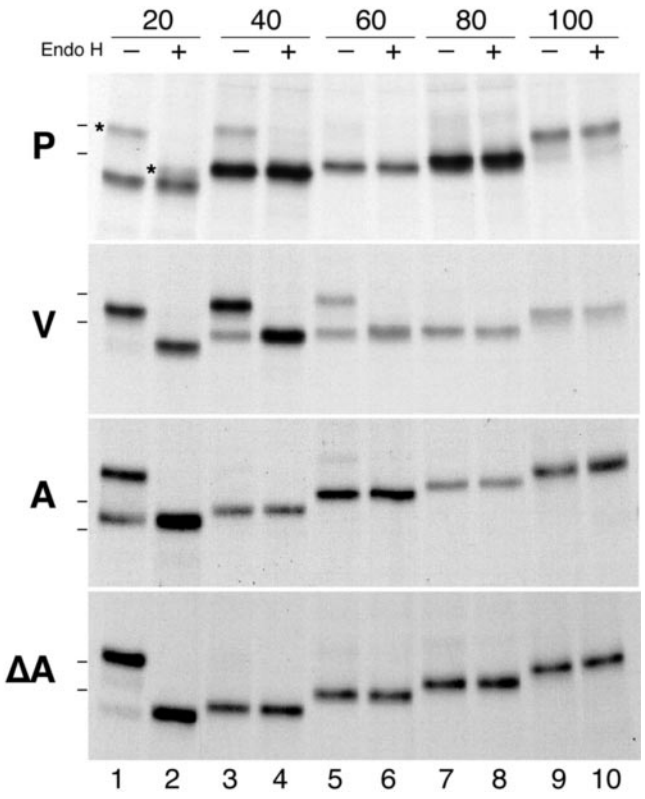

Figure 3. Insertion of chimeric proteins with different NH_2-terminal signals. Chimeric proteins P#A, V#A, A#A, and ΔA#A, which contain the NH_2-terminal signals of preprolactin, preprovasopressin-neurophysin II, and the full-size or truncated ASGP receptor H1, respectively, were expressed in COS cells and analyzed as described in Fig. 2 A. Asterisks indicate the uncleaved population of P20A. The position of molecular weight markers of 35 and 40 kD are indicated.

Published October 18, 1999

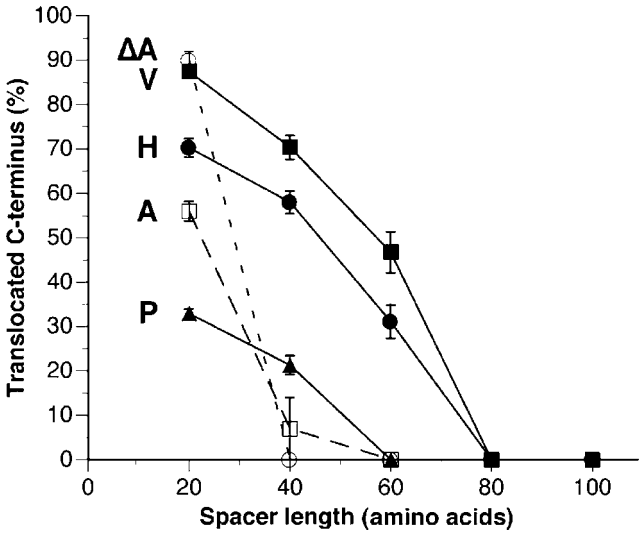

Figure 4. Quantitation of transmembrane topologies. Multiple experiments like those shown in Fig. 2 A and Fig. 3 were quantified using a PhosphorImager. The fraction of polypeptides with a glycosylated and thus translocated COOH terminus are plotted versus the length of the spacer separating the two signals. The curves are labeled with the abbreviation for the first signal sequence: H, hemagglutinin (filled circles); P, preprolactin (filled triangles); V, prepro-vasopressin-neurophysin II (filled squares); A, ASGP receptor H1 including the NH_2-terminal hydrophilic sequence (open squares); ΔA, ASGP receptor H1 with a truncated NH_2-terminal domain (open circles). The means and SDs of three to six determinations are shown.

fewer COOH-terminally translocated proteins than those with the hemagglutinin signal: a maximum of 35% for P20A and none for P60A. Since the presequence of preprolactin is relatively long, signal cleavage could be observed for P20A. The unglycosylated, cleaved products displayed a slightly higher electrophoretic mobility than the glycosylated, uncleaved products (indicated by an asterisk in Fig. 3, panel P) after endo H digestion, consistent with their expected topologies (Fig. 1 B).

The constructs with the signal of vasopressin generated more glycosylated products than the other series: ~90% for V20A and ~50% with a spacer of 60 residues (V60A). However, with spacers of 80 or 100 residues, no glycosylated products were generated for any of the constructs. Based on these results, the strength of the different secretory signals, i.e., the ability to dominate the insertion process, can be ranked P > H > V.

In the construct series A#A and ΔA#A, the first copy of the internal signal-anchor of the ASGP receptor H1 within the polypeptide appeared to dictate the topology, since glycosylated products were only generated with the shortest spacer length. In A20A and ΔA20A, the positive flanking charges at the NH_2 terminus of the second signal are also part of the COOH-terminal flanking region of the first, and are thus likely to weaken the first signal. Deletion of most of the 40-amino acid hydrophilic NH_2 terminus in ΔA20A increased the fraction of glycosylated products from ~55 to ~90%. Thus, the larger NH_2 terminus of A20A inhibits the topology with a translocated COOH terminus.

Signal Competition for SRP?

Two mechanisms could explain the observed behavior of conflicting signals in the insertion process: competition between the signals for the recruitment of SRP in the cytosol (Fig. 5 A), and/or competition for the preferred topology within the translocon (Fig. 5 B). According to the first model, the kinetics of SRP binding to the first signal are slow enough that the second signal may emerge from the ribosome before SRP was recruited to the first. The two signals then compete directly for SRP binding. If the first signal binds SRP, it will initiate translocation of the spacer sequence. If the second signal does, it will induce translocation of the COOH terminus, whereas the first signal will remain in the cytosol or will subsequently insert into the membrane with an N_{exo}/C_{cyt} orientation. This model predicts, in agreement with our results, that the ability of the second signal to compete for SRP decreases with the length of the spacer, and depends on the relative affinities of the two signals for SRP.

This model requires that at the time when the second

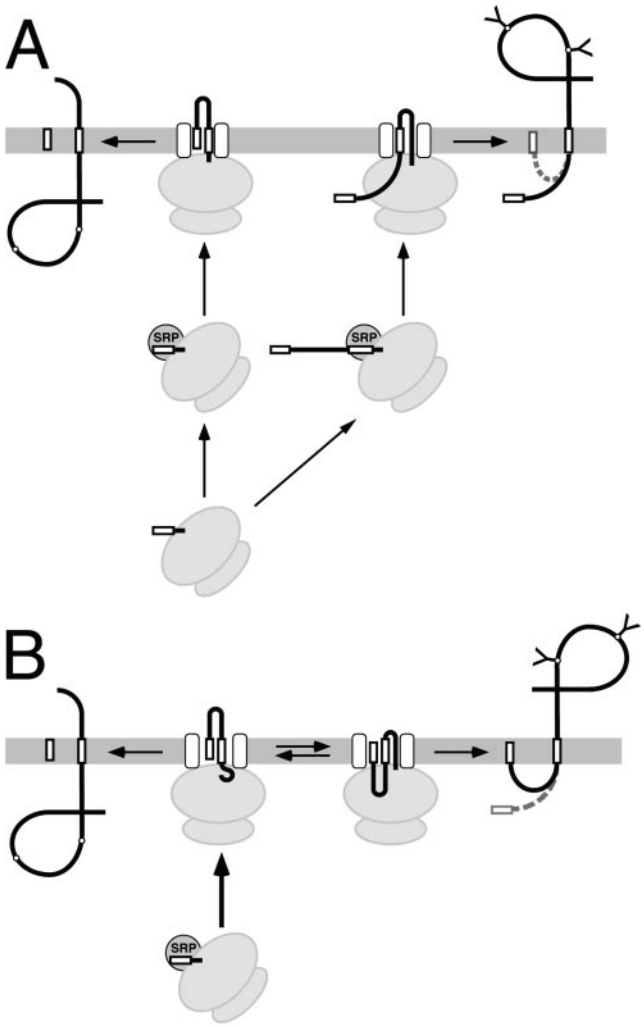

Figure 5. Two mechanisms for topogenic competition. Competition between the signals for SRP in the cytosol (A) or for the preferred orientation in the translocon (B). See text for details.

Published October 18, 1999

signal emerges from the ribosome, there are still polypeptides that have not been targeted to the ER. To test this condition, we generated the constructs consisting of the signal of vasopressin or hemagglutinin, followed by 55, 75, 95, or 115 residues, including a diagnostic glycosylation site (Fig. 6 A) (V55, H55, etc.). Since at least 35 residues of a nascent polypeptide are hidden within the ribosome and inaccessible to SRP, translation will reach the stop codon and trigger disassembly of the ribosome when the first 20, 40, 60, or 80 residues following the signal sequence have emerged from the ribosome. This corresponds to a moment in the translation of the constructs V20A, V40A, V60A, and V80A (or of the corresponding constructs H#A) when the second hydrophobic sequence is just beginning to emerge from the ribosome (Fig. 6 A). When the stop codon is reached, only polypeptides that have already been targeted to the ER membrane will produce glycosylated products; all others will be released into the cytosol and remain unglycosylated. In vitro translation in reticulocyte lysate with dog pancreas microsomes added either during or after translation confirmed that these constructs could not be posttranslationally translocated (data not shown).

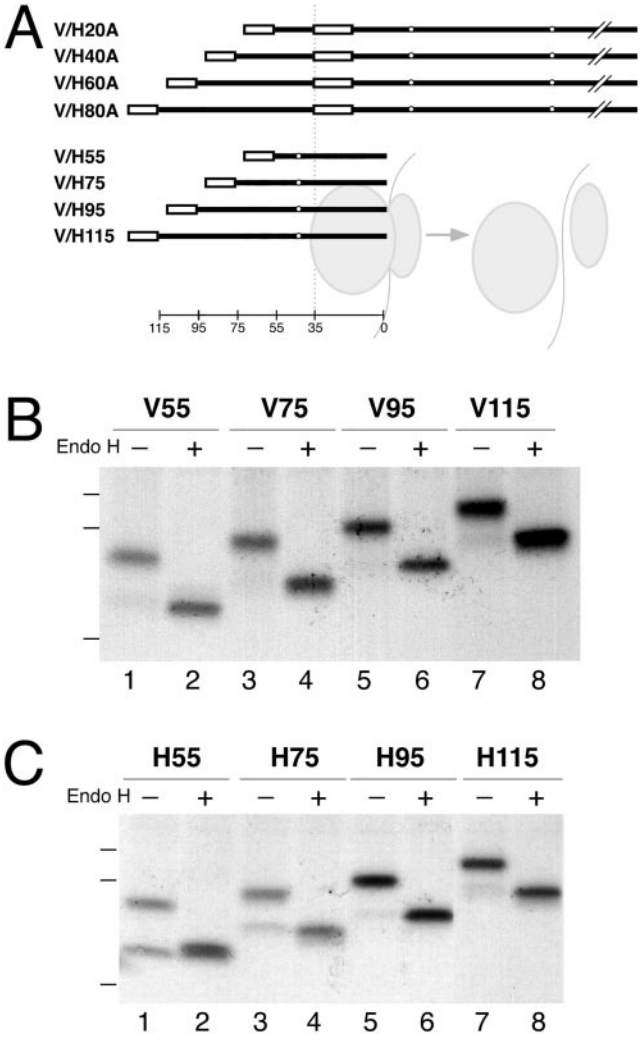

Figure 6. Kinetics of SRP recruitment and targeting by the vasopressin and hemagglutinin signals. Polypeptides consisting of the signal sequence of hemagglutinin or vasopressin followed by 55, 75, 95, or 115 residues, including a glycosylation site (H55, H75, etc. and V55, V75, etc.; schematically shown in A) were expressed in COS cells, metabolically labeled for 40 min, immunoprecipitated, incubated with (+) or without endo H (−), and analyzed by SDS-gel electrophoresis and fluorography (B and C). The unglycosylated fraction corresponds to polypeptides that had not been targeted to the ER, yet when the ribosome reached the stop codon and disassembled. The position of molecular weight markers of 9, 14, and 20 kD are indicated.

As shown in Fig. 6 B, even the shortest construct with a vasopressin signal, V55, was >90% glycosylated (lanes 1 and 2), indicating that the binding of SRP to the vasopressin signal and its targeting to the ER membrane was fast. Only few polypeptides of the constructs V#A thus had the opportunity to be targeted by the second signal. In contrast, the polypeptides with a hemagglutinin signal produced significant fractions of unglycosylated products: ~40% of H55, ~25% of H75, and still ~10% of H95. Therefore, SRP recruitment is considerably slower than with a vasopressin signal. This suggests that the second signal in the H#A constructs might be able to compete for SRP binding. However, the numbers of untargeted polypeptides cannot account for the observed populations of H20A, H40A, and H60A with a translocated COOH terminus (Fig. 4). Most importantly, a larger fraction of the V#A than of the H#A constructs inserted with a translocated COOH terminus, even though targeting by the vasopressin signal is faster. Therefore, the different topologies are more likely the result of competition between the signals within the translocon (Fig. 5 B) than of the kinetics of SRP binding and targeting by the first signal.

Glycosylation in the Spacer Sequence Affects Topogenesis

According to the second model (Fig. 5 B), the spacer sequence of all polypeptides is exposed to the ER lumen, at least transiently. To probe the lumenal exposure of the spacer, a diagnostic *N*-glycosylation site was introduced. In the construct V40(NIT)A, a potential glycosylation site NIT was generated by insertion of three additional codons at asparagine 13 of the spacer segment. As shown in Fig. 7 A, lanes 3 and 4, the site was efficiently modified, since all polypeptides were now glycosylated either once or twice, demonstrating that they had translocated either the spacer sequence (one glycosylation) or the COOH-terminal domain (two glycosylations). However, in comparison to V40A (Fig. 7 A, lanes 1 and 2), the fraction of twice glycosylated proteins dropped significantly. In a protease protection assay, it was confirmed that the efficiency of glycosylation in the translocated COOH terminus was not reduced (data not shown). Therefore, the ratio between the two topologies was significantly altered by addition of the glycosylation site in favor of polypeptides with a translocated spacer sequence. The fraction of products with a translocated COOH terminus decreased from ~75 to ~30% (Fig. 7 B). A similar effect was seen for the longer spacer of 60 residues, where insertion of the glycosylation site reduced this fraction from ~50% in V60A to <20% in V60(NIT)A (Fig. 7 A, lanes 9–12, and Fig. 7 B). This showed that topogenesis is affected by glycosylation, a lu-

Published October 18, 1999

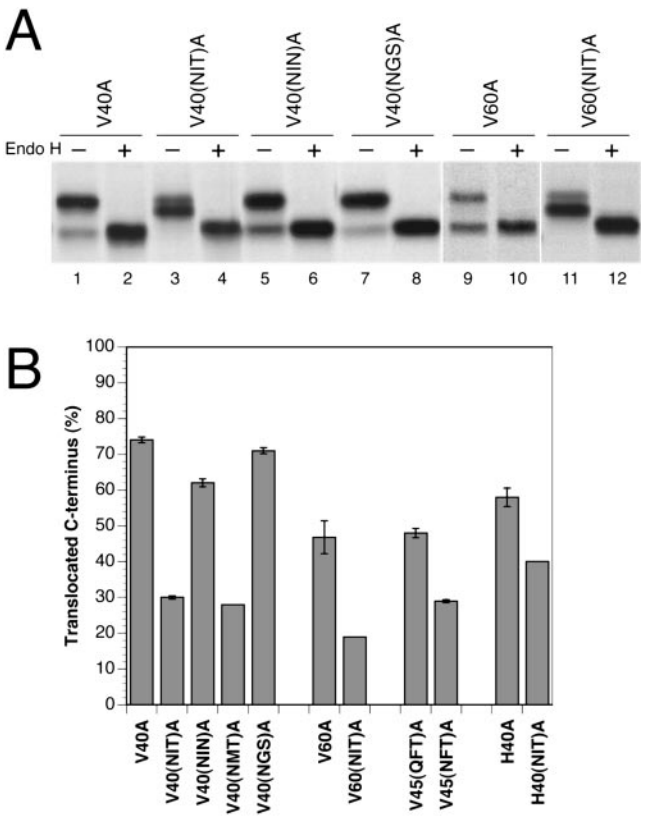

Figure 7. Effect of glycosylation in the spacer sequence on protein topology. (A) The indicated constructs were expressed in COS cells and analyzed with (+) or without (−) endo H digestion. The ratio of proteins with a translocated COOH terminus (twice glycosylated) to proteins with a translocated spacer sequence (unglycosylated or once glycosylated) shifted in favor of the latter upon insertion of functional glycosylation sites into the spacer segment. (B) Quantitation of experiments as shown in A. The means and SDs of three to five determinations are plotted, except for V40(NMT)A, V60(NIT)A, and H40(NIT)A of which single measurements are shown.

menal modification, which clearly takes place after targeting.

To exclude the possibility that the mutation exerted its effect by altering the conformational properties of the spacer sequence rather than by glycosylation itself, the sequence NIT was mutated to NIN, which is not modified by oligosaccharyl transferase (V40[NIN]A; Fig. 7 A, lanes 5 and 6, and Fig. 7 B): the effect on the topology was largely reverted. A different consensus sequence, NMT, which was also efficiently glycosylated, showed the same increase of polypeptides with translocated spacer sequences as NIT (Fig. 7 B). The sequence NGS, in contrast, did not affect topology, but also produced hardly any polypeptides with a single glycan (Fig. 7 A, lanes 7 and 8, and Fig. 7 B). This confirms that a consensus sequence N-X-S/T is not sufficient for the topogenic effect, but that efficient glycosylation is required. (Interestingly, NGS was reported to be the most efficient N-X-S consensus sequence for *N*-glycosylation in an in vitro study [Shakin-Eshleman et al., 1996]. Apparently, the sequence context dramatically affects the relative efficiencies of potential glycosylation sites). The effect of glycosylation on topology was also reproduced with an entirely different spacer sequence of 45 amino acids corresponding to a segment of the exoplasmic portion of H1 with a natural glycosylation site NFT. Mutation to QFT caused the fraction of COOH-terminally translocated polypeptides to increase from ~30% (V45[NFT]A) to ~50% (V45[QFT]A) (Fig. 7 B). As expected, the topogenic influence of glycosylation was also observed with hemagglutinin as the NH_2-terminal signal sequence (H40[NIT]A vs. H40A) (Fig. 7 B). Glycosylation affects topogenesis and thus takes place before the topology of the polypeptides is determined, most likely by trapping the spacer sequence on the lumenal side, preventing its return to the cytosolic side of the membrane (as illustrated in Fig. 8).

Discussion

The insertion of multispanning membrane proteins is a complex process that is still poorly understood. To test how successive topogenic determinants in a polypeptide chain are decoded by the targeting and translocation machinery of the mammalian ER, we expressed a series of chimeric proteins containing two conflicting signal sequences in COS cells. Only when the two signals are sufficiently separated from each other can the results be explained by a simple linear insertion process, in which the most NH_2-terminal signal sequence determines its own orientation as well as that of a subsequent transmembrane domain. The required spacer length depends on the characteristics (the strength) of the two signals. The weakest NH_2-terminal signal tested, that of vasopressin, completely forced the internal signal-anchor of the ASGP receptor H1 into a stop-transfer orientation with a spacer of 80 residues or longer, whereas the H1 signal-anchor itself could do so with a spacer of only 40 amino acids. With shorter spacers, mixed topologies were observed, indicating that topogenic information in the second signal of the polypeptide codetermined the insertion process. Hence, insertion did not occur in a strictly linear manner from the NH_2 to the COOH terminus.

The situation is thus comparable to that found in bacteria by Coleman et al. (1985) using proteins with two copies of the prolipoprotein signal separated by either 13 or 27 residues. The spacer was predominantly translocated with the longer spacer, and the COOH terminus with the shorter one. However, a point mutation that shifted the first signal from a cotranslational to a posttranslational mode of action caused COOH-terminal translocation also with the longer spacer, suggesting that targeting and insertion was now directed by the second signal. Similarly, if in

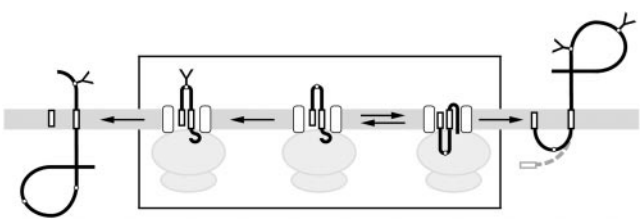

Figure 8. Schematic model of how glycosylation affects topogenesis. See text for details.

Published October 18, 1999

our system the binding kinetics of SRP to the NH_2-terminal signal are sufficiently slow, SRP binding and targeting by the second signal might explain our results. Experiments by Johnsson et al. (1994) provided evidence that in yeast many signal sequences are surprisingly slow in targeting to the ER membrane. The signal sequences tested were of different efficiencies, allowing for between ~100 (invertase) and ~300 residues (carboxypeptidase Y and Ste6) to be exposed to the cytosol for 50% of the polypeptides. However, most of the weak signals, such as that of caboxypeptidase Y, were later shown to function independently of SRP (Ng et al., 1996), and even invertase secretion was only slightly affected in an SRP-defective strain (Rothe and Lehle, 1998). The question of SRP-dependent targeting kinetics in mammalian cells, where only few proteins are known to be targeted independently of SRP (Zimmermann et al., 1990; Kutay et al., 1995), was still open.

In this study, we analyzed the kinetics of SRP-dependent targeting in COS cells using the hemagglutinin and vasopressin signals followed by reporter polypeptides of increasing length. Targeting was clearly faster (relative to translation) than in all the examples tested in yeast. 60% of the chains with a hemagglutinin signal and >90% with a vasopressin signal had been targeted by the time 55 residues past the signal had been translated. The relative rate of targeting might reflect that pro-vasopressin–neurophysin II is a relatively short protein of 145 residues in comparison to hemagglutinin with 547 amino acids, and thus needs to be targeted more rapidly. From the sequence, it is not clear what makes the vasopressin signal more efficient for SRP recruitment. For our constructs with conflicting signal sequences, the targeting rates of the hemagglutinin and vasopressin signals do not provide an explanation for the observed topologies, since there is no correlation with COOH-terminal translocation. Although a fraction of some of the proteins, e.g., with a hemagglutinin signal and with the shortest spacer, may be targeted by the second signal, the topogenic influence of the second signal appears to be largely independent of how the protein was brought to the translocon.

Glycosylation sites introduced into the spacer segment between the signal sequences provided a surprising insight into the dynamics of the orientation process in the translocon. *N*-glycosylation significantly affected the equilibrium of topologies in favor of that with a translocated spacer sequence and a cytosolic COOH terminus. This implies that with or without a glycosylation site, the spacer segment destined for cytosolic localization is transiently exposed to the ER lumen. With the glycosylation site, the spacer is modified by oligosaccharyl transferase when it appears in the lumen and is trapped there (Fig. 8). The results indicate that the polypeptides dynamically reorient within the translocon, exploring the possible topologies favored by either of the two signals. Hydrophilic sequences of up to 60 residues between transmembrane domains move in and out of the ER before the topology is decided. A likely mechanism for this glycosylation effect is steric hindrance of the modified segment to return to the cytosolic side. It has been previously observed in vitro that a glycosylated segment of a model membrane protein was finally exposed to the cytosol (Lipp et al., 1989). We did not recover any threefold glycosylated products indicative of a lumenal COOH terminus and a glycosylated spacer, suggesting either that in our system glycosylated sequences cannot flip back to the cytosolic side or, if they can, that they are rapidly deglycosylated in the cytosol (Suzuki et al., 1998).

For the insertion process of natural multispanning membrane proteins, our results illustrate that successive membrane-spanning sequences insert in a coordinated manner, if sufficiently close within the polypeptide. What determines the topogenic strength of a signal is not obvious from the few examples analyzed. It does not simply correlate with the charge difference, since Δ(C–N) of the constructs with a spacer of 40 residues, for example, are −6.5 (P40A), −1.5 (H40A), and −5 (V40A), listed in the order of decreasing topogenic strength. Other criteria shown to influence the orientation of individual signals, such as hydrophobicity (Wahlberg and Spiess, 1997) and the size of the NH_2-terminal hydrophilic segment (Figs. 3 and 4, A20A vs. ΔA20A; Denzer et al., 1995), are likely to contribute as well. In addition, longer and more hydrophobic signals, like that of H1 and, to a lesser extent, that of preprolactin, may exit the translocation pore more rapidly (Martoglio et al., 1995; Mothes et al., 1997) and might thereby limit the influence of the second signal.

Spacer segments of 80 or more residues may be sterically unable to reorient themselves within the translocon. An additional mechanism that could limit an effect of the second topogenic sequence is the binding of lumenal chaperones, particularly BiP. The yeast homologue of BiP, Kar2p, has been shown to promote posttranslational translocation, but probably acts also in cotranslational transport (Brodsky et al., 1995; Panzner et al., 1995; Pilon et al., 1998; Zimmermann, 1998; Matlack et al., 1999). It is likely that spacer sequences that are long enough to engage with lumenal chaperones will be anchored by this interaction, thus determining the topology.

In summary, topologenesis of membrane proteins appears to be a dynamic process in which topogenic information of closely spaced signal and transmembrane sequences throughout the polypeptide is integrated. Enzymes associated with the translocation machinery and acting on the protein to be inserted, such as oligosaccharyl transferase and potentially signal peptidase and chaperones, can affect the process and thereby contribute to the efficient and uniform insertion of natural membrane proteins.

We thank Nicole Beuret for technical assistance, and Dr. N. Kralli and our colleagues in the laboratory for critical reading of the manuscript.

This work was supported by Grant 31-43483.95 from the Swiss National Science Foundation.

Submitted: 21 June 1999
Revised: 8 September 1999
Accepted: 14 September 1999

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.

Blobel, G. 1980. Intracellular protein topogenesis. *Proc. Natl. Acad. Sci. USA.* 77:1496–1500.

Brodsky, J.L., J. Goeckeler, and R. Schekman. 1995. BiP and Sec63p are required for both co- and posttranslational protein translocation into the yeast endoplasmic reticulum. *Proc. Natl. Acad. Sci. USA*. 92:9643–9646.

Coleman, J., M. Inukai, and M. Inouye. 1985. Dual functions of the signal peptide in protein transfer across the membrane. *Cell.* 43:351–360.

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.

Eusebio, A., T. Friedberg, and M. Spiess. 1998. The role of the hydrophobic domain in orienting natural signal sequences within the ER membrane. *Exp. Cell Res.* 241:181–185.

Gething, M.J., and J. Sambrook. 1981. Cell-surface expression of influenza haemagglutinin from a cloned DNA copy of the RNA gene. *Nature.* 293: 620–625.

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.

Görlich, D., and T.A. Rapoport. 1993. Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell.* 75:615–630.

Gravelin, G., M. Sakaguchi, H. Andersson, and G. von Heijne. 1997. Topological rules for membrane protein assembly in eukaryotic cells. *J. Biol. Chem.* 272:6119–6127.

Habener, J.F., J.T. Potts, Jr., and A. Rich. 1976. Pre-proparathyroid hormone. Evidence for an early biosynthetic precursor of proparathyroid hormone. *J. Biol. Chem.* 251:3893–3899.

Hamman, B.D., J.-C. Chen, E.E. Johnson, and A.E. Johnson. 1997. The aqueous pore through the translocon has a diameter of 40-60 A during cotranslational protein translocation at the ER membrane. *Cell.* 89:535–544.

Hanein, D., K.E.S. Matlack, B. Jungnickel, K. Plath, K.-U. Kalies, K.R. Miller, T.A. Rapoport, and C.W. Akey. 1996. Oligomeric rings of the Sec61p complex induced by ligands required for protein translocation. *Cell.* 87:721–732.

Harley, C.A., J.A. Holt, R. Turner, and D.J. Tipper. 1998. Transmembrane protein insertion orientation in yeast depends on the charge difference across transmembrane segments, their total hydrophobicity, and its distribution. *J. Biol. Chem.* 273:24963–24971.

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.

Hegde, R.S., S. Voigt, T.A. Rapoport, and V.R. Lingappa. 1998. TRAM regulates the exposure of nascent secretory proteins to the cytosol during translocation into the endoplasmic reticulum. *Cell.* 92:621–631.

High, S., and V. Laird. 1997. Membrane protein biosynthesis-all sewn up? *Trends Cell Biol.* 7:206–209.

Johnsson, N., and A. Varshavsky. 1994. Ubiquitin-assisted dissection of protein transport across membranes. *EMBO (Eur. Mol. Biol. Organ.) J.* 13:2686–2698.

Kutay, U., G. Ahnert-Hilger, E. Hartmann, B. Wiedenmann, and T.A. Rapoport. 1995. Transport route for synaptobrevin via a novel pathway of insertion into the endoplasmic reticulum membrane. *EMBO (Eur. Mol. Biol. Organ.) J.* 14:217–223.

Leitinger, B., A. Hille-Rehfeld, and M. Spiess. 1995. Biosynthetic transport of the asialoglycoprotein receptor H1 to the cell surface occurs via endosomes. *Proc. Natl. Acad. Sci. USA.* 92:10109–10113.

Lipp, J., N. Flint, M.T. Haeuptle, and B. Dobberstein. 1989. Structural requirements for membrane assembly of proteins spanning the membrane several times. *J. Cell Biol.* 109:2013–2022.

Martoglio, B., M.W. Hofmann, J. Brunner, and B. Dobberstein. 1995. The protein-conducting channel in the membrane of the endoplasmic reticulum is open laterally toward the lipid bilayer. *Cell.* 81:207–214.

Matlack, K.E.S., B. Misselwitz, K. Plath, and T.A. Rapoport. 1999. BiP acts as a molecular ratchet during posttranslational transport of prepro-α factor across the ER membrane. *Cell.* 97:553–564.

Maurer, R.A., and D.J. McKean. 1978. Synthesis of preprolactin and conversion to prolactin in intact cells and a cell-free system. *J. Biol. Chem.* 253: 6315–6318.

Mothes, W., S.U. Heinrich, R. Graf, I. Nilsson, G. van Heijne, J. Brunner, and T. Rapoport. 1997. Molecular mechanism of membrane protein integration into the endoplasmic reticulum. *Cell.* 89:523–533.

Mothes, W., B. Jungnickel, J. Brunner, and T.A. Rapoport. 1998. Signal sequence recognition in cotranslational translocation by protein components of the endoplasmic reticulum membrane. *J. Cell Biol.* 142:355–364.

Ng, D.T., J.D. Brown, and P. Walter. 1996. Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J. Cell Biol.* 134:269–278.

Panzner, S., L. Dreier, E. Hartmann, S. Kostka, and T.A. Rapoport. 1995. Posttranslational protein transport in yeast reconstituted with a purified complex of Sec-proteins and Kar2p. *Cell.* 81:561–570.

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.

Pilon, M., K. Romisch, D. Quach, and R. Schekman. 1998. Sec61p serves multiple roles in secretory precursor binding and translocation into the endoplasmic reticulum membrane. *Mol. Biol. Cell.* 9:3455–3473.

Rothe, C., and L. Lehle. 1998. Sorting of invertase signal peptide mutants in yeast dependent and independent on the signal-recognition particle. *Eur. J. Biochem.* 252:16–24.

Rothman, J.E., and H.F. Lodish. 1977. Synchronised transmembrane insertion and glycosylation of a nascent membrane protein. *Nature.* 269:775–780.

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.* 89:16–19.

Sato, M., R. Hresko, and M. Mueckler. 1998. Testing the charge difference hypothesis for the assembly of a eucaryotic multispanning membrane protein. *J. Biol. Chem.* 273:25203–25208.

Shakin-Eshleman, S.H., S.L. Spitalnik, and L. Kasturi. 1996. The amino acid at the X position of an Asn-X-Ser sequon is an important determinant of N-linked core-glycosylation efficiency. *J. Biol. Chem.* 271:6363–6366.

Spiess, M. 1995. Head or tails—what determines the orientation of proteins in the membrane. *FEBS Lett.* 369:76–79.

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.

Suzuki, T., H. Park, K. Kitajima, and W.J. Lennarz. 1998. Peptides glycosylated in the endoplasmic reticulum of yeast are subsequently deglycosylated by a soluble peptide: N-glycanase activity. *J. Biol. Chem.* 273:21526–21530.

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. 1989. Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. *Nature.* 341:456–458.

Wahlberg, J.M., and M. Spiess. 1997. Multiple determinants direct the orientation of signal-anchor proteins: the topogenic role of the hydrophobic signal domain. *J. Cell Biol.* 137:555–562.

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.

Wessels, H.P., and M. Spiess. 1988. Insertion of a multispanning membrane protein requires only one signal sequence. *Cell.* 55:61–70.

Wessels, H.P., J.P. Beltzer, and M. Spiess. 1991. Analysis of protein topology in the endoplasmic reticulum. *Methods Cell Biol.* 34:287–302.

Zimmermann, R. 1998. The role of molecular chaperones in protein transport into the mammalian endoplasmic reticulum. *Biol. Chem.* 379:275–282.

Zimmermann, R., M. Zimmermann, H. Wiech, G. Schlenstedt, G. Müller, F. Morel, P. Klappa, C. Jung, and W.W.E. Cobet. 1990. Ribonucleoparticle-independent transport of proteins into mammalian microsomes. *J. Bioenerg. Biomembr.* 22:711–723.