Probing the Environment of Signal−Anchor Sequences during Topogenesis in the Endoplasmic Reticulum†

Marie Higy, Stefan Gander, and Martin Spiess*

Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Received September 20, 2004; Revised Manuscript Received November 19, 2004

ABSTRACT: Signal sequences for insertion of protein into the mammalian endoplasmic reticulum orient themselves in the translocon on the basis of their flanking charges. It has recently been shown that hydrophobic N-terminal signals initially insert head-on before they invert their orientation to translocate the C-terminus. The rate of inversion is reduced with the increasing hydrophobicity of the signal due to an increased affinity for the initial bound state at the translocon. To probe the environment of the signal while its orientation is determined, different hydrophobic residues were inserted at various positions throughout a uniform oligoleucine signal sequence and the constructs were expressed in transfected COS-7 cells. The resulting topologies revealed a strikingly symmetric position dependence specifically for bulky aromatic amino acids, reflecting the structure of a lipid bilayer. Maximal N-translocation was observed when the guest residues were placed at the N- or C-terminus of the hydrophobic sequence or in the very center, corresponding to the positions of highest expected affinity of the signal sequence as a membrane-spanning helix for the bilayer. The results support the model that during topogenesis in vivo the signal sequence is exposed to the lipid membrane.

Hydrophobic signal sequences target secretory and membrane proteins to the endoplasmic reticulum (ER)[1] for translocation into the lumen or for integration into the lipid bilayer (*1*−*3*). The signal is first recognized by the signal recognition particle (SRP), which directs the nascent chain−ribosome complex to the ER membrane by interaction with the SRP receptor (*4*). Both SRP and SRP receptor are GTPases which control specific docking to the translocation machinery. The actual translocation pore or translocon is created by the Sec61 complex composed of an α-subunit that spans the membrane 10 times, and a single-spanning β- and γ-subunit. This pore complex allows hydrophilic sequences to pass through the membrane and permits signal and transmembrane sequences to laterally exit into the lipid bilayer.

A signal sequence can insert into the translocon and subsequently into the membrane in two orientations (*5*). Cleavable signals and signal−anchor sequences of type II membrane proteins translocate their C-terminal end, acquire an N_{cyt}/C_{exo} orientation (cytoplasmic N-terminus and exoplasmic C-terminus), and they initiate cotranslational transfer of the growing polypeptide across the membrane. In contrast, reverse signal−anchor sequences of type III proteins (also classified as type Ia) insert with an N_{exo}/C_{cyt} orientation and induce translocation of the N-terminus. Examples for type III proteins are synaptotagmin, neuregulin, and the family of cytochromes P450.

Several factors have been shown to determine the orientation of the signal in the membrane. Most prominently, charged residues flanking the hydrophobic core of the signal influence orientation: the more positive end is generally cytosolic, a phenomenon known as the "positive-inside rule" (*6*−*9*). Charge interactions at the translocon, including residues of the Sec61 complex itself (*10*), are responsible for orienting the signal sequence. In addition, the folding state of hydrophilic sequences N-terminal to a signal may sterically hinder N-terminal translocation irrespective of the flanking charges (*11*). A third determinant is the hydrophobicity of the core of the signal sequence (the h-domain) itself. Strongly hydrophobic signals were observed to insert with an N_{exo}/C_{cyt} orientation even when the charge distribution was more positive at the N-terminus (*12*−*14*).

How hydrophobicity influences signal orientation was not obvious. A recent in vivo study indicated that hydrophobicity slows the kinetics of signal orientation in the translocation machinery (*15*). An N-terminal signal−anchor sequence with an h-domain of 22 consecutive leucine residues inserted with mixed topologies despite a positive N-terminus. Surprisingly, the topology depended on the total length of the protein: the fraction of polypeptides with an N_{cyt}/C_{exo} orientation was lowest for the shortest constructs and increased up to a length of ∼300 residues following the signal sequence. This result indicated that N-terminal signals initially insert "head-on" into the translocon to yield an N_{exo}/C_{cyt} orientation. Then they may invert over time until protein synthesis is completed or until further reorientation is stopped by an as yet unknown mechanism after approximately 40−50 s. Inversion appears

† This work was supported by Grant 31-061579.00 from the Swiss National Science Foundation.

* To whom correspondence should be addressed: Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland. Phone: +41-61-2672164. Fax: +41-61-2672148. E-mail: Martin.Spiess@unibas.ch.

[1] Abbreviations: ER, endoplasmic reticulum; N_{cyt}/C_{exo} and N_{exo}/C_{cyt}, cytoplasmic N-terminus and exoplasmic C-terminus and vice versa, respectively; SRP, signal recognition particle.

10.1021/bi047976z CCC: $30.25
Published on Web 01/20/2005

to be driven by electrostatic interactions, since the topology change was accelerated by increasing the N-terminal positive charge and slowed by reducing it. The hydrophobicity of the h-domain also affected reorientation: increased hydrophobicity diminished the rate of inversion, whereas reduced hydrophobicity allowed more rapid inversion. This effect could be explained by hydrophobicity stabilizing the binding of the signal to an interaction site at the translocation apparatus.

The hydrophobic core of a signal sequence thus affects topology by influencing the kinetics of signal reorientation, rather than by changing its preference for a final topology. This opens the possibility of exploring the environment of the signal during the orientation process using a series of constructs with identical charge distribution and overall hydrophobicity, but containing guest residues in the h-domain at various positions through the sequence. Here we tested the effect of the large hydrophobic amino acids tryptophan, phenylalanine, and tyrosine, as well as the small hydrophobic residues valine and alanine inserted at various positions in an oligoleucine sequence. Bulky hydrophobic amino acids showed a striking position dependence of topogenesis in a conspicuously symmetric pattern not observed with the other similarly hydrophobic amino acid. The results suggest the exposure of the signal to the lipid bilayer during topogenesis in vivo.

MATERIALS AND METHODS

DNA Constructs. The starting construct encoding H1ΔQ-Leu16 has been previously described (*13*). Pairs of leucine residues in the hydrophobic oligoleucine core of the signal sequence were replaced with other amino acids by polymerase chain reaction using Vent polymerase (New England Biolabs) and appropriate mutagenic oligonucleotide primers similar to the procedure used in ref *14*. For example, to construct H1ΔQL16WW1/2, the sense oligonucleotide CGGGGTACCATGGGACCGCAGTGGTGGCTTTTGC-TGCTGCTC was used (the KpnI cloning site and the mutated codons for tryptophan are underlined) in combination with a reverse primer corresponding to a sequence in the plasmid vector and with the cDNA of H1ΔQL16 as the template. Mutations in the C-terminal half of the signal sequence were generated with an antisense primer such as CCGGGATCCCAAGAGCAACAGCAGGAGCCTCCTG-AGGAGCAGC for H1ΔQL16WW9/10 (the BamHI cloning site and the mutated codons for tryptophan are underlined) in combination with an upstream primer complementary to the vector sequence. With the BamHI site, the polymerase chain reaction products were ligated to the downstream cDNA sequence of H1. The final constructs were subcloned into the expression vector pECE (*16*). To test the dependence of protein topology on the length of the protein, the sequence C-terminal to the signal sequence of normally 230 amino acids in H1ΔQLeuWW1/2, H1ΔQLeuWW7/8, and H1ΔQ-LeuWW9/10 was replaced with truncated or extended versions of 110, 170, 290, or 460 residues as described by Goder and Spiess (*15*). All constructs were verified by sequencing.

Cell Culture, Transfection, and Immunoprecipitation. Cell culture reagents were from Life Technologies, Inc. COS-7 cells were grown in modified Eagle's minimal essential medium supplemented with 10% fetal calf serum, 2 mM L-glutamine, 100 units/mL penicillin, and 100 μg/mL streptomycin at 37 °C with 7.5% CO_2. Transient transfection was performed in six-well clusters with lipofectin (Life Technologies, Inc.) according to the manufacturer's instructions. 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 40 min at 37 °C with 100 μCi/mL [^{35}S]methionine, transferred to 4 °C, washed twice with phosphate-buffered saline, and finally lysed and immunoprecipitated using a rabbit antiserum directed against a synthetic peptide corresponding to residues 277−287 near the C-terminus of ASGP receptor H1 (anti-H1C). The immune complexes were isolated with protein A−Sepharose (Amersham Pharmacia Biotech) and analyzed by SDS−polyacrylamide gel electrophoresis and autoradiography. Quantitation was performed using a phosphorimager (Molecular Dynamics Inc.). To determine the fraction of products with the N_{cyt}/C_{exo} topology, the intensity of the glycosylated forms with one and two glycans in the percentage of the total of all glycosylated and unglycosylated forms was calculated. This value proved to be independent of transfection efficiency, which may vary somewhat between experiments.

Alkaline Extraction and Protease Protection Assays. Alkaline extraction was performed as previously described (*17*). To reduce the viscosity of the sample, the cells suspended in alkaline solution were pipetted up and down through a 25 gauge needle to shear the DNA before it was loaded onto the sucrose cushion. For the protease protection assay, labeled cells were incubated at 4 °C with hypotonic swelling buffer [15 mM Hepes/KOH (pH 7.2) and 15 mM KCl] and scraped with a rubber policeman. Aliquots were incubated without protease or with 100 μg/mL trypsin in the presence or absence of 0.5% Triton X-100 for 30 min at 4 °C. Trypsin was then inhibited by addition of 500 μg/mL soybean trypsin inhibitor before immunoprecipitation and analysis by SDS gel electrophoresis and autoradiography. As control constructs, HC, encoding the cleavable signal of influenza hemagglutinin fused to the C-terminal portion of H1 (*18*), and wild-type H1 were used.

RESULTS

Protein Topology Is Strongly Dependent on the Position of Double Tryptophans in an H1ΔQLeu16 Host Sequence. To explore the properties of the signal's environment while its orientation is determined, we tested the behavior of model proteins based on H1ΔQLeu16. This protein is derived from the H1 subunit of the asialoglycoprotein receptor, a typical type II membrane protein. It has been modified at the N-terminus to start with an artificial signal sequence of 16 consecutive leucine residues with a single positive charge at the N-terminus (the α-amino group) and a net negative C-terminal flanking sequence (Figure 1B). Because of its considerably hydrophobic core, only ∼35% of the molecules are able to translocate the C-terminus, before the polypeptide of 230 amino acids is completed and further reorientation of the protein is terminated (ref *15*, and autoradiograph in Figure 1A). To test the position effect of guest residues throughout the uniform h-domain of this construct, we replaced pairs of leucine residues with tryptophans, the most voluminous amino acid. Two residues were simultaneously

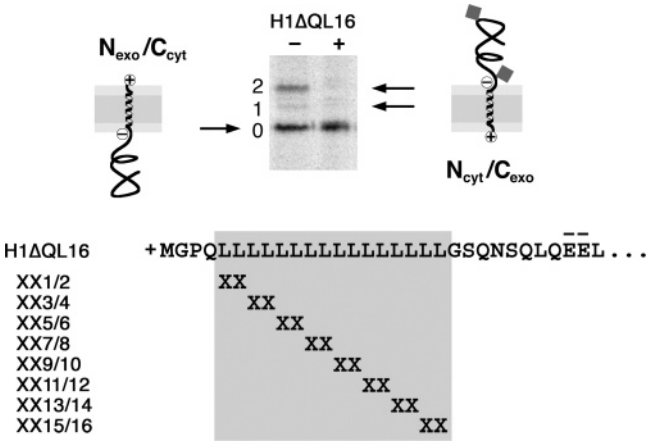

FIGURE 1: Model proteins for studying the position effect of guest residues in an oligoleucine host signal−anchor sequence. (Top) The signal−anchor protein H1ΔQL16 was expressed in COS-7 cells, labeled with [^{35}S]methionine, immunoprecipitated, and analyzed by SDS gel electrophoresis and autoradiography. Three forms were produced corresponding to the unglycosylated protein (0) with the N_{exo}/C_{cyt} orientation, or with one (1) or two (2) glycans of the N_{cyt}/C_{exo} orientation, as schematically illustrated. The latter two forms are sensitive to deglycosylation by endoglycosidase H (+). One of the glycosylation sites is close to the transmembrane sequence and is not completely modified, resulting in some products that are glycosylated only once. (Bottom) To probe the signal's environment during topogenesis, two identical guest residues (XX) were systematically introduced at the indicated positions into the hydrophobic domain (gray background) of the host sequence H1ΔQL16. The N-terminal sequence of H1ΔQL16 is shown in single-letter code (charged residues are denoted with + or −).

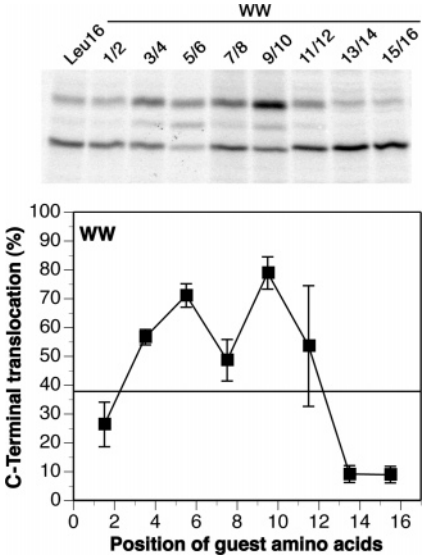

FIGURE 2: Topogenic effect of tryptophan guest residues in an oligoleucine signal−anchor sequence. A series of constructs with two tryptophans (WW) replacing two leucines throughout the oligoleucine sequence in H1ΔQL16 (as shown in Figure 1) was expressed in COS-7 cells labeled with [^{35}S]methionine, immunoprecipitated, and analyzed by SDS gel electrophoresis and autoradiography. The glycosylated and unglycosylated forms representing N_{cyt}/C_{exo} and N_{exo}/C_{cyt} orientations, respectively, were quantified and expressed as the fraction of polypeptides with a translocated C-terminus (N_{cyt}/C_{exo} orientation) vs the position of the guest residues in the oligoleucine sequence. The average and standard deviation of four independent experiments are shown. The horizontal line represents the topology distribution of the original host sequence (Leu16).

altered to potentially generate more significant effects. When two tryptophans had previously been placed at positions 8 and 13 of H1ΔQLeu16, the distribution of topologies had not significantly changed in comparison to the parental construct (*14*).

The constructs containing double tryptophans placed throughout the signal sequence were expressed in transfected COS-7 cells and labeled with [^{35}S]methionine for 40 min. The products were immunoprecipitated and analyzed by SDS gel electrophoresis and autoradiography (Figure 2). Topology was derived from the glycosylation pattern, since modification of the N-glycosylation sites at positions 40 and 108 indicates C-terminal translocation, whereas the unglycosylated form indicates an N_{exo}/C_{cyt} orientation (Figure 1A and refs *13*−*15*). The results showed a dramatic position dependence for the tryptophans in a surprisingly symmetric pattern. When positioned at either end of the h-domain, the tryptophan-containing sequences inserted predominantly with an N_{exo}/C_{cyt} orientation. In contrast, tryptophans further inside the h-domain favored C-terminal translocation, except when placed in the center at positions 7 and 8 where the fraction of N_{exo}/C_{cyt} molecules with ∼50% was again significantly increased. Constructs with single tryptophans inserted into H1ΔQL16 produced a similar pattern with smaller deviations (data not shown).

Interpretation of the glycosylation pattern in terms of protein topology depends on the assumption that all products are integrated into the membrane and that glycosylation efficiency is not affected by insertion of the guest residues. Glycosylation efficiency at one of the two glycosylation sites (most likely the one closest to the membrane) is indeed somewhat dependent on the sequence, probably reflecting the protein's position in the membrane when the glycans are transferred (*19*). To test for complete membrane integration, WW1/2, WW7/8, and WW9/10, i.e., constructs covering the entire spectrum of different topology distributions, were expressed in COS-7 cells, labeled, and subjected to alkaline extraction and centrifugation (Figure 3A). Glycosylated and unglycosylated forms of all constructs were equally and almost completely recovered in the membrane pellet, whereas a control protein consisting of the cleavable signal of influenza hemagglutinin fused to the C-terminal portion of H1 was extracted into the supernatant. All signal−anchor constructs were thus efficiently targeted to the ER and integrated into the membrane irrespective of the position of the tryptophan residues.

To test whether all unglycosylated products have an N_{exo}/C_{cyt} orientation, cells were labeled, ruptured by swelling and scraping, and incubated at 4 °C with or without trypsin (Figure 3B). Wild-type H1, which was analyzed as a control protein, shifted its position in SDS gel electrophoresis upon trypsin treatment because of digestion of its 40-amino acid cytoplasmic domain. The unglycosylated forms of all three WW constructs were efficiently digested by protease, whereas the glycosylated ones were resistant. Upon permeabilization of membranes by detergent, all products were sensitive. These control experiments confirm that the glycosylation state of the proteins as analyzed in Figure 2 directly represents their final orientation in the membrane.

Topologies at a Fixed Protein Length Reflect the Kinetics of Signal Reorientation. To analyze whether the observed topologies of different constructs are a measure of the inversion kinetics of the signal−anchor sequence in the

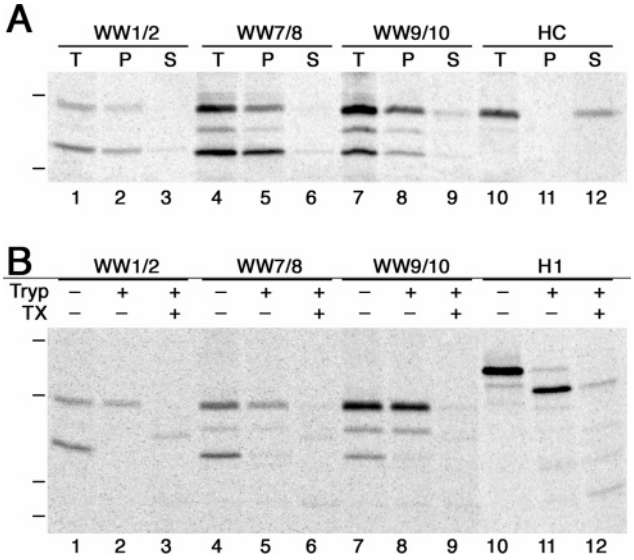

FIGURE 3: Glycosylation patterns reflect protein topology. (A) COS-7 cells expressing selected constructs were labeled with [^{35}S]-methionine and subjected to alkaline extraction. After centrifugation, the membrane pellet (P) and the supernatant (S), as well as an equal aliquot of the total starting material (T), were analyzed by immunoprecipitation, SDS gel electrophoresis, and autoradiography. The control construct HC consists of the cleavable signal sequence of influenza hemagglutinin fused to the C-terminal domain of H1. The positions of molecular mass markers of 26 and 37 kDa are indicated. (B) COS-7 cells expressing selected constructs or wild-type H1 were labeled with [^{35}S]methionine, permeabilized by swelling and scraping, and incubated with or without trypsin (Tryp) in the presence or absence of Triton X-100 (TX). The products were then analyzed by immunoprecipitation, SDS gel electrophoresis, and autoradiography. The positions of molecular mass markers of 20, 26, 37, and 50 kDa are indicated.

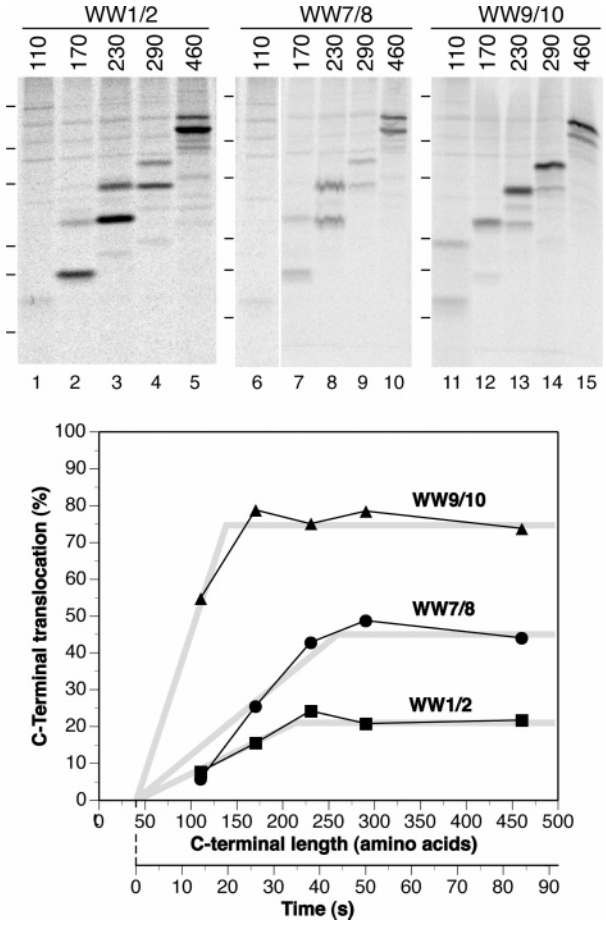

FIGURE 4: Topologies reflect the inversion kinetics of the signal−anchor sequences. For each of the constructs WW1/2, WW7/8, and WW9/10, a series was constructed with polypeptides of 110, 170, 230, 290, and 460 residues following the signal−anchor sequence. These constructs were expressed in COS-7 cells, labeled with [^{35}S]-methionine, immunoprecipitated, and analyzed by SDS gel electrophoresis. The positions of molecular mass markers of 15, 20, 26, 37, 50, and 64 kDa are indicated. The glycosylated and unglycosylated forms representing the N_{cyt}/C_{exo} and N_{exo}/C_{cyt} orientations, respectively, were quantified and plotted as the fraction of polypeptides with a translocated C-terminus (N_{cyt}/C_{exo} orientation) vs the length of the C-terminal domain and the time of translation from when the signal completely emerged from the ribosome to termination. The time scale is based on a translation rate of 5 amino acids/s as determined for cultured cell lines (*20*) and starts at 40 amino acids, since 30−40 residues are hidden within the ribosome. The average of three independent experiments (including the one presented at the top) is shown.

translocon as shown before for the parental construct H1ΔQL16 (*15*), we generated a series of constructs with 110, 170, 230, 290, and 460 residues following the signal−anchor sequences of WW1/2, WW7/8, and WW9/10. Upon expression in COS-7 cells, labeling, and immunoprecipitation, the fraction of glycosylated products corresponding to N_{cyt}/C_{exo} polypeptides was determined. With the increasing length of the proteins, the C-terminally translocated fraction increased (Figure 4). On the basis of a translation rate of ∼5 amino acids/s as determined for cultured mammalian cells (*20*), the length of the protein can be converted into the time of translation from the moment the signal has fully emerged from the ribosome (with ∼40 residues still hidden within the ribosome) to termination (bottom scale in Figure 4). The rate of increase of N_{cyt}/C_{exo} products was lowest for the WW1/2 series of constructs, highest for WW9/10, and intermediate for WW7/8. The topologies observed in Figure 2 for a fixed length of the protein of 230 amino acids downstream of the signal−anchor sequence therefore reflect the signals' inversion kinetics. Thus, tryptophans at the end of the signal core inhibit inversion compared to the oligoleucine host sequence, whereas tryptophans between enhance inversion, except in the very center.

Other Bulky Hydrophobic Guest Residues Generate a Similar Position Dependence of Protein Topology. The signal's inversion kinetics could be affected either by changing the flexibility of the polypeptide in its free state in the translocon in comparison to the host sequence or by altering the affinity to the initial bound state. Tryptophan has a bulky aromatic side chain, and it appears to be unlikely that the exchange of leucines with tryptophans at any position would make the polypeptide more flexible. To explain the increased inversion kinetics for guest tryptophans at positions 3−12, it is more plausible that tryptophans at these positions reduce the affinity for the signal binding site in comparison to the parental oligoleucine sequence.

Because of the nitrogen in its side chain, tryptophan also has polar properties. To address whether the observed position dependence of tryptophans in the host signal is caused by the size of the side chain or its polar contribution, two additional series of constructs with two phenylalanines or two tyrosines as guest residues in the oligoleucine sequence were prepared and analyzed in transfected COS-7 cells as described above. Both phenylalanine and tyrosine are hydrophobic, aromatic, and rigid. They are large in

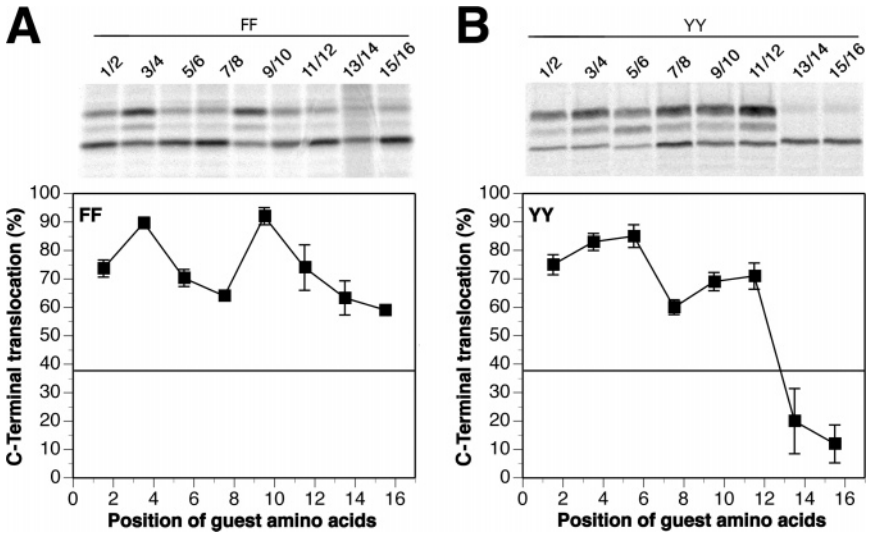

FIGURE 5: Position dependence of other bulky aromatic guest residues: phenylalanines and tyrosines. Construct series with two phenylalanines [FF (A)] or tyrosines [YY (B)] replacing leucines throughout the oligoleucine sequence in H1ΔQL16 were expressed in COS-7 cells, labeled with [^{35}S]methionine, immunoprecipitated, and analyzed by SDS gel electrophoresis and autoradiography. The glycosylated and unglycosylated forms were quantified and plotted as the fraction of polypeptides with a translocated C-terminus (N_{cyt}/C_{exo} orientation) vs the position of the guest residues in the oligoleucine sequence. The average and standard deviation of three independent experiments are shown. The horizontal line represents the topology distribution of the original host sequence.

comparison to most other amino acids, although somewhat smaller than tryptophan. The side chain of phenylalanine lacks a polar group entirely, whereas that of tyrosine contains a polar hydroxyl group.

The resulting topologies for the constructs with double phenylalanines (Figure 5A) presented almost the same symmetric position dependence with minimal N_{cyt}/C_{exo} orientation for phenylalanines placed at the beginning, at the end, or in the center of the signal−anchor sequence. The most notable difference is that phenylalanines at the ends of the signal yielded more N_{cyt}/C_{exo} orientation than the tryptophans. The polar characteristics of tryptophan might be responsible for the strongly reduced inversion kinetics when positioned at the ends of the signal.

The corresponding series of tyrosine-containing constructs again showed the same basic pattern of topologies with relative minima of N_{cyt}/C_{exo} topology with tyrosines on either end or in the center of the h-domain (Figure 5B). Tyrosines at the C-terminal end inserted almost completely with an N_{exo}/C_{cyt} orientation like the corresponding tryptophan constructs, consistent with tyrosine having polar characteristics as well. However, N-terminal tyrosines (YY1/2) produced relatively high levels of the N_{cyt}/C_{exo} orientation (see Discussion).

Symmetric Position Dependence Correlates with Side Chain Bulkiness. As a control, the effect of valine guest residues was tested with the constructs VV1/2 to VV15/16 (Figure 6A). Valine is smaller than leucine, but similar in hydrophobicity. The position effects were less distinctive and entirely different from those of the bulky aromatic residues. The fraction of products with the N_{cyt}/C_{exo} orientation was generally somewhat lower than that of the parental sequence H1ΔQLeu16, ∼30%, but was slightly increased to 40% when the valines were positioned at the N-terminus and decreased to ∼15% at positions 5 and 6 and positions 7 and 8. The pattern was again different when alanines were used as guest residues (Figure 6B). Alanine is small and considerably less hydrophobic than the other residues that were tested. The major effect on topology of the host sequence was a significant increase in the amount of N_{cyt}/C_{exo} orientation when the alanines were placed at positions 9 and 10 and positions 11 and 12. The results with alanine and valine guest residues show that the symmetric M-shaped position dependence of topology is specific for amino acids with bulky side chains.

DISCUSSION

Signal sequences are first recognized by SRP on a hydrophobic surface created by a cluster of methionines on the 54 kDa subunit (SRP54) (*21*−*23*). Once the ribosome−nascent chain−SRP complex has docked, the signal is transferred to a second recognition site in the translocon (*24*−*26*). Topogenic determinants control how the signal is positioned in the translocon: electrostatic forces act on the flanking charges (*6*, *7*, *10*), and an N-terminal hydrophilic extension sterically hinders insertion of the N-terminus (*11*). Our previous in vivo studies for N-terminal, very hydrophobic signals showed that their orientation changes with the time of translation (*15*). Initial insertion leads to an N_{exo}/C_{cyt} orientation, but C-terminal translocation is acquired gradually in a process that is accelerated by an increasing charge difference Δ(N−C) and slowed by the increasing hydrophobicity of the signal. With constant charge and overall hydrophobicity, the effect on topology of guest residues inserted at various positions in a generic oligoleucine signal is expected to reflect changes in the affinity of the signal for the initial bound state and in the flexibility in the free state. The position effect of guest residues on topology may thus provide information about the situation of the signal sequence during topogenesis in vivo.

The position effects of valines and alanines may be best explained by their effects on the flexibility of the polypeptide chain in comparison to the parental oligoleucine sequence. Valine is a β-branched amino acid that reduces flexibility, whereas the small side chain of alanine allows more conformational freedom. Reduced or increased flexibility

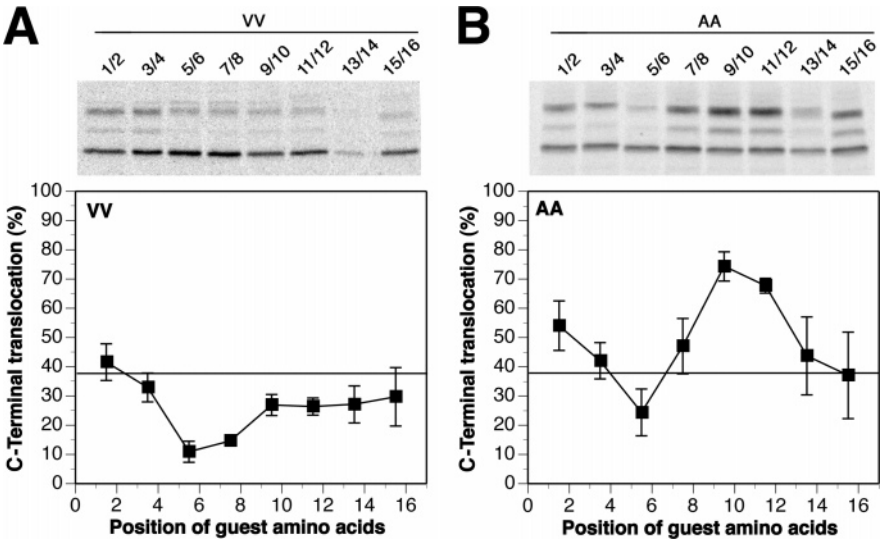

FIGURE 6: Topogenic effect of nonbulky valine and alanine guest residues. Construct series with two valines [VV (A)] or alanines [AA (B)] replacing leucines throughout the oligoleucine sequence in H1ΔQL16 were expressed in COS-7 cells, labeled with [^{35}S]methionine, immunoprecipitated, and analyzed by SDS gel electrophoresis and autoradiography. The glycosylated and unglycosylated forms were quantified and plotted as the fraction of polypeptides with a translocated C-terminus (N_{cyt}/C_{exo} orientation) vs the position of the guest residues in the oligoleucine sequence. The average and standard deviation of three independent experiments are shown. The horizontal line represents the topology distribution of the original host sequence.

particularly at central positions in the apolar sequence will hinder or facilitate, respectively, the inversion of the free polypeptide. Bulky hydrophobic residues generated a strikingly symmetric pattern that cannot be explained by variations in polypeptide flexibility. It is more likely that the properties of the signal binding site dominate the changes in inversion kinetics for different positions of the guest residues. In a proteinaceous signal binding site, bulky side chains might sterically hinder binding or contribute additional favorable contacts, therefore reducing or increasing the binding affinity, resulting in higher or lower inversion rates, respectively. Without a detailed structure of such a binding site, the position dependence of guest residues is not predictable. Symmetry would be accidental. If the bound signal had a helical conformation and were bound on one side to a protein surface in a fixed position, one would expect to find a periodicity of three to four residues. This is clearly not the case. In addition, the oligoleucine sequence is uniform and therefore might always position itself in such a way that the bulky residues are facing away from the protein surface.

However, the symmetry of the position dependence for bulky aromatic guest residues parallels the symmetry of the lipid bilayer (illustrated in Figure 7A) and thus suggests contact of the signal with the lipid membrane during topogenesis. The low fraction of the N_{cyt}/C_{exo} orientation obtained for tryptophans positioned at the ends of the h-domain could be explained by their favorable interaction with the interphase between the apolar core and the headgroup regions of the lipid bilayer (*27*). Statistically, tryptophan is enriched in the interphase regions of transmembrane helices (*28*, *29*) and of β-barrel proteins (*30*, *31*), probably stabilizing the position of the proteins in the membrane. This is also the case for tyrosine, whereas there is no preferred position for phenylalanine in transmembrane sequences. The highest fractions of N_{cyt}/C_{exo} orientation and thus the lowest apparent affinity for the bound state in the translocon are observed for WW5/6 and WW9/10, in which the tryptophans in a transmembrane helix would be positioned in the center of the acyl chain regions of the two lipid layers, where the membrane is most tightly packed and accommodation of a large and stiff side chain is least favorable. In contrast, at the center of the bilayer, order and density are lowest and bulky side chains are more easily accommodated. Correspondingly, tryptophans at the center of the signal in WW7/8 showed a reduced amount of the N_{cyt}/C_{exo} orientation, i.e., increased affinity in comparison to WW5/6 and WW9/10.

The behavior of phenylalanine guest residues supports this interpretation. Because the phenylalanine side chain is similarly bulky, they produce the same basic pattern as tryptophans. Since the side chain is nonpolar and cannot take advantage of the polarity change at the interphase regions of the bilayer, phenylalanines at the ends of the oligoleucine sequence showed more N_{cyt}/C_{exo} orientation. In contrast, tyrosines should behave more like tryptophans. They do, except at the very N-terminus where tyrosines yielded a higher fraction of N_{cyt}/C_{exo} topology and the entire topology pattern for tyrosine guest residues appears to be tilted. A potential explanation might be that the polarity axis of the tyrosine side chain, unlike that of the tryptophan side chain, cannot reach into the interphase as steeply at the N-terminus of a helix as at the C-terminus.

Sixteen apolar residues are not quite long enough to completely span the hydrophobic core region of the membrane. They are, however, flanked on both sides by several polar but uncharged residues, which can also be found within natural signal and transmembrane sequences. Part of these flanking sequences thus must be pulled into the apolar phase (*32*). The symmetry of topology patterns for all three aromatic residues seems to be slightly offset toward the N-terminus, suggesting that the N-terminus of the signal is pulled into the hydrophobic phase more easily. It has also been shown by glycosylation mapping that tryptophans can influence the positioning of a transmembrane helix according to its interface preference (*33*). However, the presence of aromatic residues is not likely to affect the topogenesis of

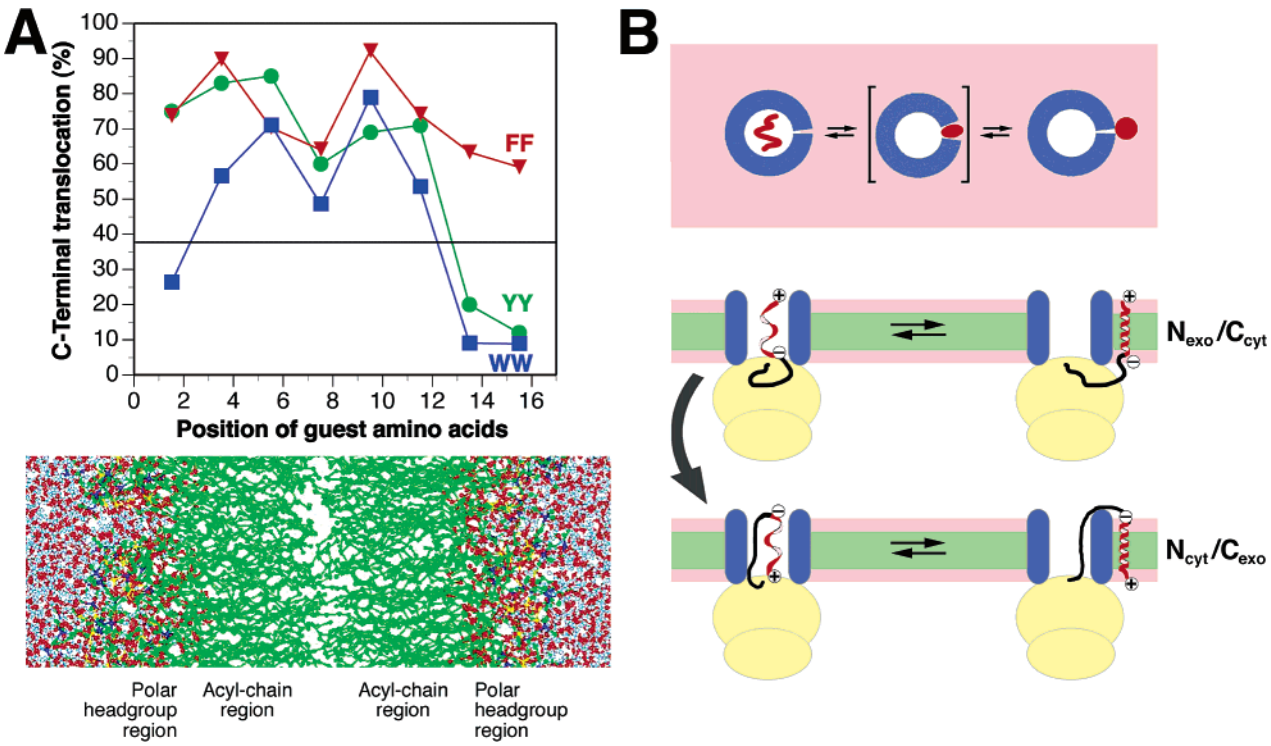

FIGURE 7: Position dependence of bulky guest residues in the signal−anchor sequence of H1ΔLeu16 that reflects the symmetry of the lipid bilayer. (A) The dependence of topology on the position of double tryptophans [WW (blue squares), from Figure 2], phenylalanines [FF (red triangles), from Figure 5A], and tyrosines [YY (green circles), from Figure 5B] are shown with a section of a dipalmitoylphosphatidylcholine bilayer upon molecular dynamics simulation (*39*), clearly visualizing the low density in the center of the membrane and the polar−nonpolar transition at the headgroup interphases (courtesy of A. Lyubartsev, Stockholm University, Stockholm, Sweden). (B) The signal during topogenesis at the translocon is schematically shown from the top (top panel) or from the side (bottom panel). The signal is proposed to be in an equilibrium between a free state in the pore and a bound state in the lipid membrane outside the translocon. The situation of the signal intercalating between the surfaces of the exit site is likely to be an unstable transition state (indicated by brackets). Only the free state is able to invert in a manner driven by charge interactions. See the text also.

natural signal sequences, because they orient themselves much more rapidly than the highly hydrophobic model signals used here. They have completely inverted their orientation before topogenesis is terminated.

In summary, the easiest explanation of our results is that the signal contacts the lipid bilayer during topogenesis in vivo. The concept that the signal upon insertion into the translocon is in contact with lipids is of course not new. In vitro, arrested nascent chains that are just long enough for the signal to enter the translocon could be photo-cross-linked not only to Sec61α but also to lipids (*26*, *34*). Since the extent of cross-linking to the lipid was higher for a more hydrophobic signal−anchor sequence than for a short cleavable signal, it was further proposed that the translocon might open more or less toward the lipid membrane depending on the hydrophobicity of the signal (*34*). However, cross-linking experiments provide a snapshot of a dynamic situation, and the possibility could not be ruled out that some signals were cross-linked to the protein while still within the translocon and others had already been integrated into the bilayer reacting with lipids. Cross-linking patterns obtained with a reactive side chain in different positions in the h-domain suggested that the signal was in a helical conformation in stable contact with transmembrane helices 2 and 7 of Sec61α on one side and with the lipid on the other (*26*). This might reflect the state after the signal had left the translocon upon completion of topogenesis. In vivo, signal reorientation was found to be terminated at the latest approximately 50 s after the signal emerged from the ribosome, even if translation was not yet completed (ref *15* and Figure 4). This period of signal orientation has certainly passed by the time of in vitro cross-linking. Stop-transfer sequences were similarly found to be cross-linkable in defined positions at the interface of the translocon and lipid membrane as long as the nascent chain remained attached to the ribosome (*25*). Different sequences were detected in different positions (in some cases adjacent also to TRAM), suggesting that transmembrane segments tethered to the translocation complex associate at various places with the outside of the pore complex (*25*). In contrast, our current results support the notion that the signal is in contact with the lipid bilayer during topogenesis, i.e., while signal reorientation takes place.

Recently, the crystal structure of the SecYEβ translocation complex of *Methanococcus jannaschii* has been determined (*35*), which is homologous to the mammalian Sec61α$\beta\gamma$ complex. It suggests that the translocation pore is formed by a single Sec61 complex, rather than by three or four complexes as previously proposed on the basis of electron microscopy of the yeast and mammalian translocons (*36*, *37*). As a consequence, the hydrophilic pore generated upon opening of the channel is likely to be less spacious than previously expected (*38*). In any case, however, there is no obvious hydrophobic surface lining the inside of the pore that could serve as a static binding site for apolar signal sequences.

The translocon is organized in two halves (transmembrane helices 1−5 and 6−10) with a lateral exit site toward the lipid membrane between helices 2 and 3 and helices 7 and

8 (*35*). It seems that a hydrophobic environment becomes accessible to an entering signal only when the exit site opens, for example, due to thermal motion within the structure. At such a moment, the h-domain of the signal might exit into the surrounding lipid where it will form a helix, optimizing intramolecular hydrogen bonds. Similarly, the signal might return into the hydrophilic channel pore where the peptide is unlikely to remain helical. As a flexible chain, it may invert its orientation due to charge interactions. We propose that the signal is in an equilibrium between a "bound state" as a transmembrane helix outside the translocon and a flexible, free state within the translocon (Figure 7). An intermediate state of a signal helix intercalating between the transmembrane domains of the translocon's exit site is likely to be unstable, particularly also considering the multitude of possible signal sequences. The translocon therefore allows lateral equilibration of the signal between an aqueous and a transmembrane environment. The translating ribosome may facilitate the transient lateral opening of the pore. Upon termination of translation, re-entry of the signal may be inhibited, resulting in the observed block of further topology changes.

ACKNOWLEDGMENT

We thank Drs. Adriana Pagano and Pascal Crottet for critically reading the manuscript and Dr. Torsten Schwede for helpful discussions.

REFERENCES

1. Walter, P., and Johnson, A. E. (1994) Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane, *Annu. Rev. Cell Biol. 10*, 87−119.
2. Rapoport, T. A., Jungnickel, B., and Kutay, U. (1996) Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes, *Annu. Rev. Biochem. 65*, 271−303.
3. Johnson, A. E., and van Waes, M. A. (1999) The translocon: a dynamic gateway at the ER membrane, *Annu. Rev. Cell Dev. Biol. 15*, 799−842.
4. Keenan, R. J., Freymann, D. M., Stroud, R. M., and Walter, P. (2001) The signal recognition particle, *Annu. Rev. Biochem. 70*, 755−775.
5. Goder, V., and Spiess, M. (2001) Topogenesis of membrane proteins: determinants and dynamics, *FEBS Lett. 504*, 87−93.
6. von Heijne, G. (1986) The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology, *EMBO J. 5*, 3021−3027.
7. Hartmann, E., Rapoport, T. A., and Lodish, H. F. (1989) Predicting the orientation of eukaryotic membrane-spanning proteins, *Proc. Natl. Acad. Sci. U.S.A. 86*, 5786−5790.
8. Beltzer, J. P., Fiedler, K., Fuhrer, C., Geffen, I., Handschin, C., Wessels, H. P., and Spiess, M. (1991) Charged residues are major determinants of the transmembrane orientation of a signal-anchor sequence, *J. Biol. Chem. 266*, 973−978.
9. Parks, G. D., and Lamb, R. A. (1991) Topology of eukaryotic type-II membrane proteins: Importance of N-terminal positively charged residues flanking the hydrophobic domain, *Cell 64*, 777−787.
10. Goder, V., Junne, T., and Spiess, M. (2004) Sec61p contributes to signal sequence orientation according to the positive-inside rule, *Mol. Biol. Cell 15*, 1470−1478.
11. Denzer, A. J., Nabholz, C. E., and Spiess, M. (1995) Transmembrane orientation of signal-anchor proteins is affected by the folding state but not the size of the aminoterminal domain, *EMBO J. 14*, 6311−6317.
12. Sakaguchi, M., Tomiyoshi, R., Kuroiwa, T., Mihara, K., and Omura, T. (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. U.S.A. 89*, 16−19.
13. Wahlberg, J. M., and Spiess, M. (1997) Multiple determinants direct the orientation of signal-anchor proteins: The topogenic role of the hydrophobic signal domain., *J. Cell Biol. 137*, 555−562.
14. Rösch, K., Naeher, D., Laird, V., Goder, V., and Spiess, M. (2000) The topogenic contribution of uncharged amino acids on signal sequence orientation in the endoplasmic reticulum, *J. Biol. Chem. 275*, 14916−14922.
15. Goder, V., and Spiess, M. (2003) Molecular mechanism of signal sequence orientation in the endoplasmic reticulum, *EMBO J. 22*, 3645−3653.
16. Ellis, L., Clauser, E., Morgan, D. O., Edery, M., Roth, R. A., and Rutter, W. J. (1986) Replacement of insulin receptor tyrosine residues 1162 and 1163 compromises insulin-stimulated kinase activity and uptake of 2- deoxyglucose, *Cell 45*, 721−732.
17. Wessels, H. P., Beltzer, J. P., and Spiess, M. (1991) Analysis of protein topology in the endoplasmic reticulum, *Methods Cell Biol. 34*, 287−302.
18. Schmid, S. R., and Spiess, M. (1988) Deletion of the amino-terminal domain of asialoglycoprotein receptor H1 allows cleavage of the internal signal sequence, *J. Biol. Chem. 263*, 16886−16891.
19. Nilsson, I. M., and von Heijne, G. (1993) Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane, *J. Biol. Chem. 268*, 5798−5801.
20. Hershey, J. W. B. (1991) Translational control in mammalian cells, *Annu. Rev. Biochem. 60*, 717−755.
21. Zopf, D., Bernstein, H. D., Johnson, A. E., and Walter, P. (1990) The methionine-rich domain of the 54 kDa protein subunit of the signal recognition particle contains an RNA binding site and can be cross-linked to a signal sequence, *EMBO J. 9*, 4511−4517.
22. Lutcke, H., High, S., Romisch, K., Ashford, A. J., and Dobberstein, B. (1992) The methionine-rich domain of the 54 kDa subunit of signal recognition particle is sufficient for the interaction with signal sequences, *EMBO J. 11*, 1543−1551.
23. Keenan, R. J., Freymann, D. M., Walter, P., and Stroud, R. M. (1998) Crystal structure of the signal sequence binding subunit of the signal recognition particle, *Cell 94*, 181−191.
24. Do, H., Falcone, D., Lin, J., Andrews, D. W., and Johnson, A. E. (1996) The cotranslational integration of membrane proteins into the phospholipid bilayer is a multistep process, *Cell 85*, 369−378.
25. McCormick, P. J., Miao, Y., Shao, Y., Lin, J., and Johnson, A. E. (2003) Cotranslational protein integration into the ER membrane is mediated by the binding of nascent chains to translocon proteins, *Mol. Cell 12*, 329−341.
26. Mothes, W., Jungnickel, B., Brunner, J., and Rapoport, T. A. (1998) Signal sequence recognition in cotranslational translocation by protein components of the endoplasmic reticulum membrane, *J. Cell Biol. 142*, 355−364.
27. Wimley, W. C., and White, S. H. (1996) Experimentally determined hydrophobicity scale for proteins at membrane interfaces, *Nat. Struct. Biol. 3*, 842−848.
28. Landolt-Marticorena, C., Williams, K. A., Deber, C. M., and Reithmeier, R. A. (1993) Non-random distribution of amino acids in the transmembrane segments of human type I single span membrane proteins, *J. Mol. Biol. 229*, 602−608.
29. Arkin, I. T., and Brunger, A. T. (1998) Statistical analysis of predicted transmembrane α-helices, *Biochim. Biophys. Acta 1429*, 113−128.
30. Weiss, M. S., Abele, U., Weckesser, J., Welte, W., Schiltz, E., and Schulz, G. E. (1991) Molecular architecture and electrostatic properties of a bacterial porin, *Science 254*, 1627−1630.
31. Schirmer, T., Keller, T. A., Wang, Y. F., and Rosenbusch, J. P. (1995) Structural basis for sugar translocation through maltoporin channels at 3.1 Å resolution, *Science 267*, 512−514.
32. Killian, J. A., and von Heijne, G. (2000) How proteins adapt to a membrane-water interface, *Trends Biochem. Sci. 25*, 429−434.
33. Braun, P., and von Heijne, G. (1999) The aromatic residues Trp and Phe have different effects on the positioning of a transmembrane helix in the microsomal membrane, *Biochemistry 38*, 9778−9782.
34. Martoglio, B., Hofmann, M. W., Brunner, J., and Dobberstein, B. (1995) The protein-conducting channel in the membrane of the endoplasmic reticulum is open laterally toward the lipid bilayer, *Cell 81*, 207−214.
35. van den Berg, B., Clemons, W. M., Jr., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C., and Rapoport, T. A. (2004) X-ray structure of a protein-conducting channel, *Nature 427*, 36−44.

36. Beckmann, R., Spahn, C. M., Eswar, N., Helmers, J., Penczek, P. A., Sali, A., Frank, J., and Blobel, G. (2001) Architecture of the protein-conducting channel associated with the translating 80S ribosome, *Cell 107*, 361−372.
37. Menetret, J. F., Neuhof, A., Morgan, D. G., Plath, K., Radermacher, M., Rapoport, T. A., and Akey, C. W. (2000) The structure of ribosome-channel complexes engaged in protein translocation, *Mol. Cell 6*, 1219−1232.
38. Hamman, B. D., Chen, J. C., Johnson, E. E., and Johnson, A. E. (1997) The aqueous pore through the translocon has a diameter of 40−60 Å during cotranslational protein translocation at the ER membrane, *Cell 89*, 535−544.
39. Lyubartsev, A. P., and Laaksonen, A. (2000) M.DynaMix: A scalable portable parallel MD simulation package for arbitrary molecular mixtures, *Comput. Phys. Commun. 128*, 565−589.

BI047976Z