

Viral Membrane Protein Topology Is Dictated by Multiple Determinants in Its Sequence

Ana Saurí¹, Silvia Tamborero¹, Luis Martínez-Gil¹,
Arthur E. Johnson^{2,3,4} and Ismael Mingarro^{1*}

¹*Departament de Bioquímica i Biologia Molecular, Universitat de València, E-46 100 Burjassot, Spain*

²*Department of Molecular and Cellular Medicine, Texas A&M Health Science Center, College Station, TX 77843-1114, USA*

³*Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, USA*

⁴*Department of Chemistry, Texas A&M University, College Station, TX 77843, USA*

Received 2 January 2009;
received in revised form
30 January 2009;
accepted 30 January 2009
Available online
4 February 2009

The targeting, insertion, and topology of membrane proteins have been extensively studied in both prokaryotes and eukaryotes. However, the mechanisms used by viral membrane proteins to generate the correct topology within cellular membranes are less well understood. Here, the effect of flanking charges and the hydrophobicity of the N-terminal hydrophobic segment on viral membrane protein topogenesis are examined systematically. Experimental data reveal that the classical topological determinants have only a minor effect on the overall topology of p9, a plant viral movement protein. Since only a few individual sequence alterations cause an inversion of p9 topology, its topological stability is robust. This result further indicates that the protein has multiple, and perhaps redundant, structural features that ensure that it always adopts the same topology. These critical topogenic sequences appear to be recognized and acted upon from the initial stages of protein biosynthesis, even before the ribosome ends protein translation.

© 2009 Elsevier Ltd. All rights reserved.

Edited by J. Bowie

Keywords: endoplasmic reticulum; membrane topology; translocon; trans-membrane segment; viral protein

Introduction

The cotranslational integration of membrane proteins requires the concerted actions of translating ribosomes that are complexed with membrane-embedded proteins that form a protein-conducting channel, the translocon.^{1,2} During integration, nascent membrane

proteins have to adopt the correct orientation in the lipid bilayer, fold properly, and assemble into their native state. For some time, researchers have endeavored to understand what structural features of a membrane protein dictate its topology,³ and several topological parameters have been proposed to govern the orientation of transmembrane (TM) segments in the membrane.

The primary determinant appears to be the distribution of positively charged amino acid side chains, with positive charges more frequently found in the protein segments located in the cytosol. This 'positive-inside rule' was first identified for bacterial proteins,⁴ where bacteria maintain a net negative-inside electrical potential across the cytoplasmic membrane. A similarly skewed distribution was also observed for membrane proteins in the endoplasmic reticulum (ER),⁵ in the mitochondrial inner membrane,⁶ and in the thylakoid membrane.⁷

*Corresponding author. E-mail address:

Ismael.Mingarro@uv.es.

Present address: A. Sauri, Molecular Microbiology Department, Faculty of Earth and Life Sciences, Vrije Universiteit, 1081HV Amsterdam, The Netherlands.

Abbreviations used: TM, transmembrane; ER, endoplasmic reticulum; N-t, N-terminus; OST, oligosaccharyl transferase; C-t, C-terminus; Lep, leader peptidase; SRP, signal recognition particle; IP, immunoprecipitation; GFP, green fluorescent protein; bbe, berberine bridge enzyme; RNC, ribosome nascent chain.

Another structural feature that influences the orientation of individual TM sequences in the membrane bilayer is their hydrophobicity, a topogenic determinant that has been demonstrated both *in vitro*⁸ and *in vivo*.⁹ TM helices with higher hydrophobicity are prone to insertion, with the N-terminal extramembranous region facing the noncytosolic exterior, while less hydrophobic TM segments are typically oriented with their N-terminus (N-t) in the cytosol.

Most membrane proteins are cotranslationally inserted into the lipid bilayer through the translocon, an evolutionarily conserved channel responsible for both the translocation of secretory proteins and the integration of membrane proteins.^{2,10,11} Photocrosslinking studies showed that, upon entering the translocon, the TM sequence was immediately adjacent to phospholipids,^{12–14} and this led Martoglio *et al.* to postulate that the TM sequence in the aqueous translocon pore was exposed to and partitioned directly into the lipid core of the bilayer. Consistent with this model, subsequent studies indicated that TM sequence insertion was dependent primarily on its hydrophobicity, but also noted that amino acid distribution within the TM sequence also influenced insertion.^{15,16} The crystal structure of an archaeal Sec YEF monomer has been reported,¹⁷ and the authors proposed that TM sequences would move laterally into the membrane after the closed monomer had flexed and opened to expose the lipid core. Moreover, some charged residues in the translocon itself appear to be involved in establishing the topology of membrane proteins.^{18,19}

Two models for how the N-t is positioned in the cytosol also reflect the above uncertainty in the extent to which translocon proteins are directly involved in membrane protein integration. In one model, the nascent chain is proposed to arrive at the translocon folded in a hairpin structure that locates the N-terminal end of the nascent membrane protein in the cytosol without further manipulation.²⁰ A second model suggests that the N-terminal hydrophobic segment initially inserts directly into the translocon and then dynamically inverts within the translocon to position the N-t in the cytosol.²¹ Subsequent hydrophobic TM domains may stop or reinitiate protein translocation and integration into the membrane to generate multispanning membrane proteins by alternately directing hydrophilic segments to the cytosol or lumen of the ER in response to the sequential presentation of topogenic determinants.²² Yet this model also raises a number of unanswered questions since some multispanning membrane protein integration is complex, involving, among other things, interactions between TM sequences and reorientation of one or more TM segments after reaching an initial protein topology.^{21,23–27}

Viral membrane proteins provide further examples of unusual topological behavior. For instance, membrane proteins that have dual topologies and can be found in two different orientations in the membrane have been reported for an increasing number of viral membrane proteins such as the hepatitis B large

envelop protein,²⁸ hepatitis C virus glycoproteins E1 and E2,²⁹ the Newcastle disease virus fusion protein,³⁰ the nonstructural protein 4B from hepatitis C virus,³¹ and the adenovirus E3-6K protein, where some topologies even constitute exceptions to the positive-inside rule.³² These unusual behaviors are not restricted to animal viruses, since integration of an amphipathic helix of the Tomato ringspot nepovirus NTB-VPg protein has been suggested to occur through protein oligomerization.³³

Plant viruses spread from infected cells to healthy cells using intracellular connections—plasmodesmata—that have evolved for cell-to-cell traffic in-between plant cells with rigid cell walls. Plasmodesmata are structural channels that provide cytoplasmic and ER membrane continuity between adjacent cells. Many viruses have evolved specific viral proteins—the so-called movement proteins—to transport uncoated viral genomes through plasmodesmata into adjacent cells. To accomplish this function, the carnation mottle virus, a positive-sense single-stranded RNA virus of the *Carmovirus* genus, encodes two movement proteins, p7 and p9. p7 was characterized as a soluble protein with RNA-binding capacity,^{34,35} and p9 was characterized as a double-spanning membrane protein with a defined topology.³⁶ Since both proteins are essential for the local transport of the viral genome, proper membrane integration and orientation of p9 are critical for the local transport of the viral RNA and the spread of viral infection.

Since p9 is essential and must have the correct topology to function, it follows that this viral protein must have evolved to maximally ensure its proper orientation during integration. We have therefore examined systematically the effect of modifying either the flanking charges or the hydrophobicity of the N-terminal hydrophobic segment on its topogenesis. Surprisingly, the data reveal that altering the classical topological determinants has only a minor effect on protein topology. Thus, it appears that multiple structural features are recognized to ensure that p9 topology is maintained and not subject to singular changes in the p9 sequence. Furthermore, the topogenic signals are detected early in protein biosynthesis, even before translation of this short protein is terminated.

Results

Distribution of extramembranous charged residues

The distribution of charged residues near the hydrophobic core of TM segments is the best-established topological determinant.³⁷ p9 is a double-spanning membrane protein with an even net charge distribution on both sides of the membrane, consisting in a small N-terminal tale (1–6), the first TM segment (TM1; 7–24), a translocated hydrophilic loop interconnecting both TM segments (25–38), the second TM segment (TM2; 39–59), and a

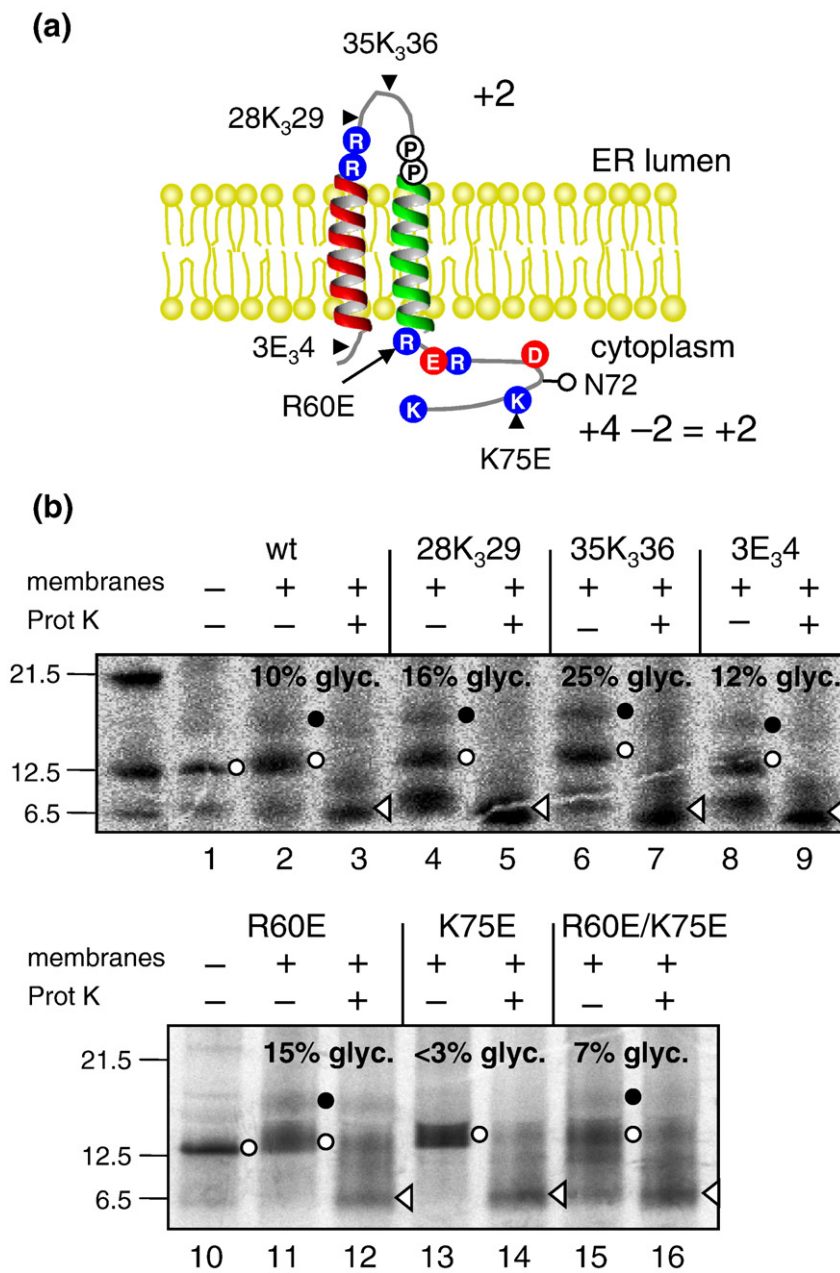


Fig. 1. Influence of charged residues on p9 topology. (a) Topology of plant viral p9 protein. The positively charged residues (arginine and lysine) and the negatively charged residues (aspartate and glutamate) are circled in blue and red, respectively. Lysine and glutamate insertions, as well as single-point mutations, are depicted. The charge differences on both sides of the ER membrane are also shown. The glycosylation site at Asn72 (N72) is indicated. (b) Radiolabeled p9 protein translations were performed in the presence (+) or in the absence (-) of microsomal membranes and proteinase K, as indicated. The filled circle identifies glycosylated p9, while the open circle indicates nonglycosylated p9 polypeptides. An open triangle indicates the protected p9 fragment.

C-terminal cytoplasmic domain (60–87) (Fig. 1a). There are only eight charged residues along the protein sequence with an even distribution on both sides of the membrane. Hence, there are two positively charged residues in the translocated loop and four positively charged residues plus two negatively charged residues in the cytosolic C-terminal domain (Fig. 1a). To monitor topology, we used the glycosylation activity of oligosaccharyl transferase (OST). Glycosylation of a protein translated *in vitro* in the presence of microsomal membranes indicates the exposure of the polypeptide nascent chain to the OST active site on the luminal side of the ER membrane.³⁸ When translated in a cell-free system in the presence of [³⁵S]methionine and canine pancreatic microsomes, the p9 C-terminal domain was

barely glycosylated (~10%; Fig. 1b, lane 2), as previously reported.³⁶ To examine in detail the influence of charge residues distribution on both sides of the membrane, a series of constructs was designed to skew charge balance. Hence, three lysine residues were inserted into two different positions in the translocated loop, just after TM1 (mutant 28K₃₂₉) and downstream of the loop sequence (mutant 35K₃₃₆). Upon synthesis in the presence of microsomal membranes, a modest increase in glycosylated band was observed in both mutants (Fig. 1b, lanes 4 and 6). Thus, even though the addition of three lysine residues in the loop region provides a net balance of five positive charges at the loop and only a net balance of two positive charges at the C-terminus (C-t), no inversion of the topology was

observed for more than 75% of the molecules. This topology was probed by proteinase K treatment of the constructs, where the protease-protected fragment (~6.5 kDa, corresponding to the TM1-loop-TM2 hairpin) was indicative of wild-type $N_{\text{cyt}}\text{-}C_{\text{cyt}}$ orientation (Fig. 1b, lanes 3, 5, and 7). These findings suggest that increasing the number of positively charged residues at the protein loop is not a sufficiently strong feature to invert protein topology.

As regards eukaryotic cells, it has been proposed that rather than the presence of positively charged residues, it is the charge difference that governs overall protein topology. For instance, a net positively charged region dictates a cytoplasmic disposition, while a net negatively charged region dictates the opposite (extracytoplasmic) disposition.³⁹ To test this hypothesis, we commenced by designing a p9 mutant in which we inserted three glutamic acids between Tyr3 and Gly4 at the N-t of the protein preceding TM1, which led to a negative balance that should favor N-t translocation. However, the *in vitro* translation of this mutant did not increase protein glycosylation, and proteinase K digestion resulted in the 6.5-kDa hairpin indicative of $N_{\text{cyt}}\text{-}C_{\text{cyt}}$ topology (Fig. 1b, lanes 8 and 9).

To analyze the effect of charge difference in the C-terminal domain of the protein, we designed a series of replacement mutants in which one or two positively charged residues were mutated to glutamic acid, thus converting the net charge balance of +2 of the wild-type sequence into 0 or -2, respectively. As observed, replacement of Arg60 and/or Lys75 with glutamic acid resulted in a glycosylation and protease K digestion pattern similar to that of the wild-type sequence (Fig. 1b, lanes 11–16), consistent with a predominant $N_{\text{cyt}}\text{-}C_{\text{cyt}}$ topology.

In order to investigate the influence of charged residues on TM1 orientation, we used a previously constructed fusion (TM1/P2) where the leader peptidase (Lep) P2 domain (with a glycosylation site) was fused after Leu36.³⁶ Lep is an inner membrane protein with two TM helices, H1 and H2, connected by a highly charged cytoplasmic loop (P1), and with both N-t and C-t facing the bacteria periplasm. According to the reported data (compare Fig. 2, lane 2, with Fig. 4 in Vilar *et al.*³⁶), this construct is strongly glycosylated (~65%), and this result is consistent with an $N_{\text{cyt}}\text{-}C_{\text{out}}$ orientation for TM1. The insertion of three lysine residues just after

TM1 significantly reduced protein glycosylation (~20%; Fig. 2, lane 5), indicative of protein inversion as proved by proteinase K treatment (Fig. 2, lane 6). When the three lysines were inserted downstream between residues 35 and 36, the charge effect in protein inversion was diminished, as suggested by an increase in glycosylated molecules (~55%; Fig. 2, lanes 7 and 8 *versus* lanes 5 and 6). Interestingly, the position effect of lysine clusters on membrane orientation has been systematically described,^{40,41} and this effect gradually decreased as the distance from the TM segment increased, in good agreement with our results.

Collectively, these results indicate that the influence of positively charged residues on viral membrane protein topology dramatically decrease when full-length proteins are translated (compare Fig. 1b, lanes 1–7, with Fig. 2), thus suggesting the idea that topological determinants could be distributed through the overall protein sequence.

Hydrophobicity of the first TM segment

The topological studies related to signal sequence hydrophobicity showed a preferential orientation depending on their total hydrophobicity. According to these studies, strongly hydrophobic signals correlate to an N-t extracytoplasmic orientation (N_{out}), while less hydrophobic signals tend to adopt an N-t cytosolic orientation (N_{cyt}).^{9,21,42} Although p9 is a viral membrane protein lacking a cleavable signal sequence, TM1 accomplishes the signal sequence function in it.⁴³ The overall hydrophobicity of TM1 was found to be -4.41 kcal/mol (calculated as the free energy required to adopt a transbilayer helical conformation: $\Delta G_{\text{woc}} - \Delta G_{\text{wif}}$ using the MPEx algorithm⁴⁴). This moderate hydrophobicity could explain the N_{cyt} orientation of the protein according to the recently proposed mechanism.²¹ To investigate whether increasing the hydrophobicity of the first TM segment changes the topology of p9, an additional Leu-Ile-Leu sequence was introduced roughly into the middle of TM1 sequence (mutant p9-LIL), between Thr15 and Gly16. The insertion of these hydrophobic residues increased the overall hydrophobicity of TM1 to -6.55 kcal/mol (Fig. 3a). The *in vitro* translation of p9-LIL in the presence of microsomal membranes did not increase the glycosylation levels of the wild-type sequence (Fig. 3b, compare lanes 2 and 4). The

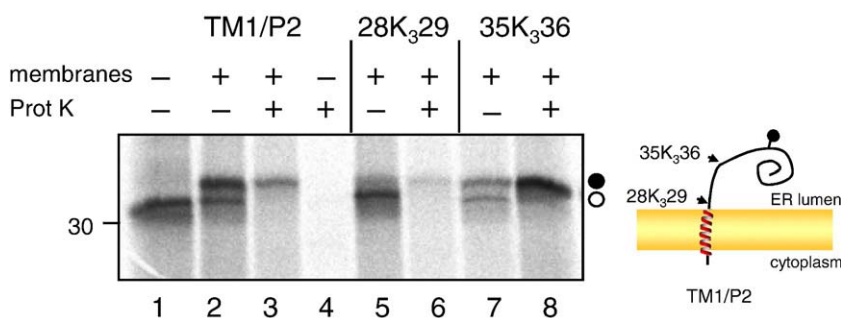


Fig. 2. Influence of positively charged residues on TM1 orientation. Residues 1–81 of Lep were replaced with residues 1–36 from p9. As in Fig. 1, *in vitro* translations were performed in the presence of microsomal membranes and proteinase K, as indicated. Lysine insertions are depicted. The filled circle represents the glycosylated proteins, while the open circle indicates nonglycosylated polypeptides.

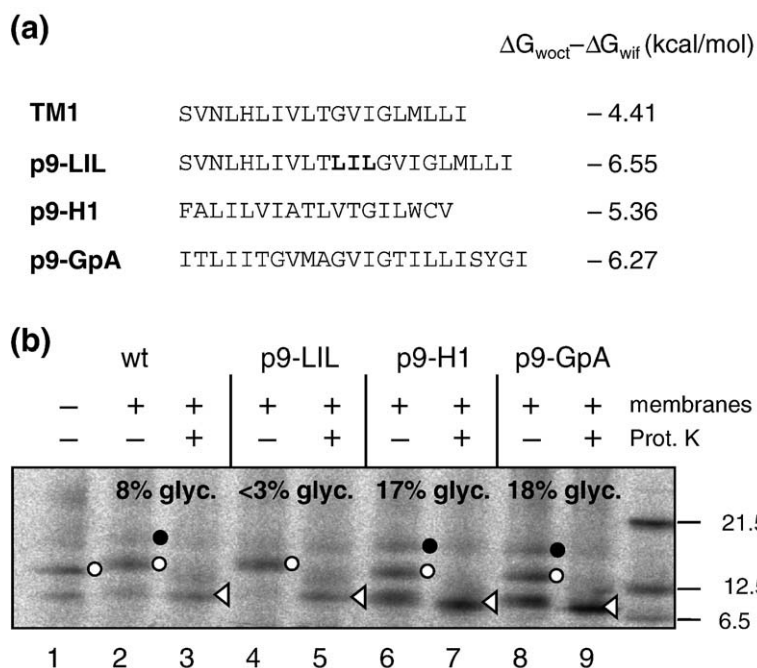


Fig. 3. Glycosylation of p9 mutants with increased hydrophobicity of the first TM segment. (a) Sequences studied as first TM segment and their calculated free energy required to adopt a transbilayer helical conformation using the MPEx algorithm.⁴⁴ (b) As before, the indicated constructs were translated *in vitro* in the presence of microsomal membranes and proteinase K, as indicated. The filled circle identifies glycosylated proteins, the open circle represents nonglycosylated polypeptides, and an open triangle depicts the protected p9 fragment.

proteinase K digestion of the protein resulted in the 6.5-kDa protected form that corresponds to the TM1-loop-TM2 hairpin and correlates with an $N_{\text{cyt}}-C_{\text{cyt}}$ orientation of p9-LIL.

Additionally, an alternative approach was designed to study the role of TM domain hydrophobicity in membrane topology. This approach consists in swapping the p9 TM1 domain with other TM sequences with a higher overall hydrophobicity. We selected the first TM domain of Lep from *Escherichia coli*, the H1 domain of Lep (p9-H1), and the TM segment of the single-spanning membrane protein glycophorin A from human erythrocytes (p9-GpA). Both TM segments are more hydrophobic than TM1: -5.36 kcal/mol for H1 and -6.27 kcal/mol for GpA-TM (Fig. 3a). Moreover, both TM segments adopted an N-t extracytoplasmic orientation in their respective native proteins (i.e., opposite orientation to p9 TM1). The *in vitro* translation of both p9 derivatives resulted in only a moderate increment in the glycosylation levels of the proteins (Fig. 3b, lanes 6 and 8), and the proteinase K digestion mainly generated the 6.5-kDa hairpin form (Fig. 3b, lanes 7 and 9). These results indicate that despite the higher hydrophobicity and their reversed orientation in their native context, the topology of these p9 derivatives (p9-H1 and p9-GpA) was not substantially modified. Therefore, further factors and/or regions in the protein could contribute to the acquisition of the defined topology.

Influence of the loop region

Next we examined the contribution of the loop connecting both TM segments in protein orientation. We previously observed that this short loop influenced the integration mechanism of the protein.^{43,45} In order to test the influence of this domain in the

topology of the protein, we replaced the 14 residues of this translocated region with the 42 residues from the cytoplasmic P1 domain of Lep (p9-P1). The P1 domain cannot be translocated across the inner *E. coli* membrane, presumably because the positively charged amino acids immediately downstream of the first hydrophobic domain (H1) prevent translocation.⁴⁶ In contrast, the expression of certain Lep-derived constructs in the presence of microsomal membranes allowed P1 translocation.⁵ The *in vitro* translation of p9-P1 chimera did not increase the glycosylation of the protein (Fig. 4a, compare lanes 2 and 5). Proteinase K analysis of this construct rendered a band with a molecular weight corresponding to the TM1P1TM2 hairpin (Fig. 4a, lane 6). The higher intensity of the proteinase K digestion product in the p9-P1 construct is due to the presence of two additional methionines at the inserted P1 sequence. These results are consistent with a predominantly $N_{\text{cyt}}-C_{\text{cyt}}$ topology even when its native loop is replaced by a conversely oriented loop. A parallel analysis of a Lep chimera, where the P1 domain was replaced with the p9 loop stretch (Lep-p9 loop), was performed to test the topological contribution of this region out of its native context. As Fig. 4b shows, replacing P1 with the p9 loop did not affect the overall topology of Lep. It should be noted that the proteinase K digestion of Lep-p9 loop was not complete probably because the short length of this loop makes it a poor substrate for protease activity. Given these observations, we realized that the p9 loop is not sufficient for its translocation in the Lep context.

Interestingly, tandem proline residues, bridging the hydrophilic loop and the downstream TM segment (Fig. 1a), are highly conserved in many *Carmovirus* homologues,^{36,47} and the presence of proline residues preceding hydrophobic regions has

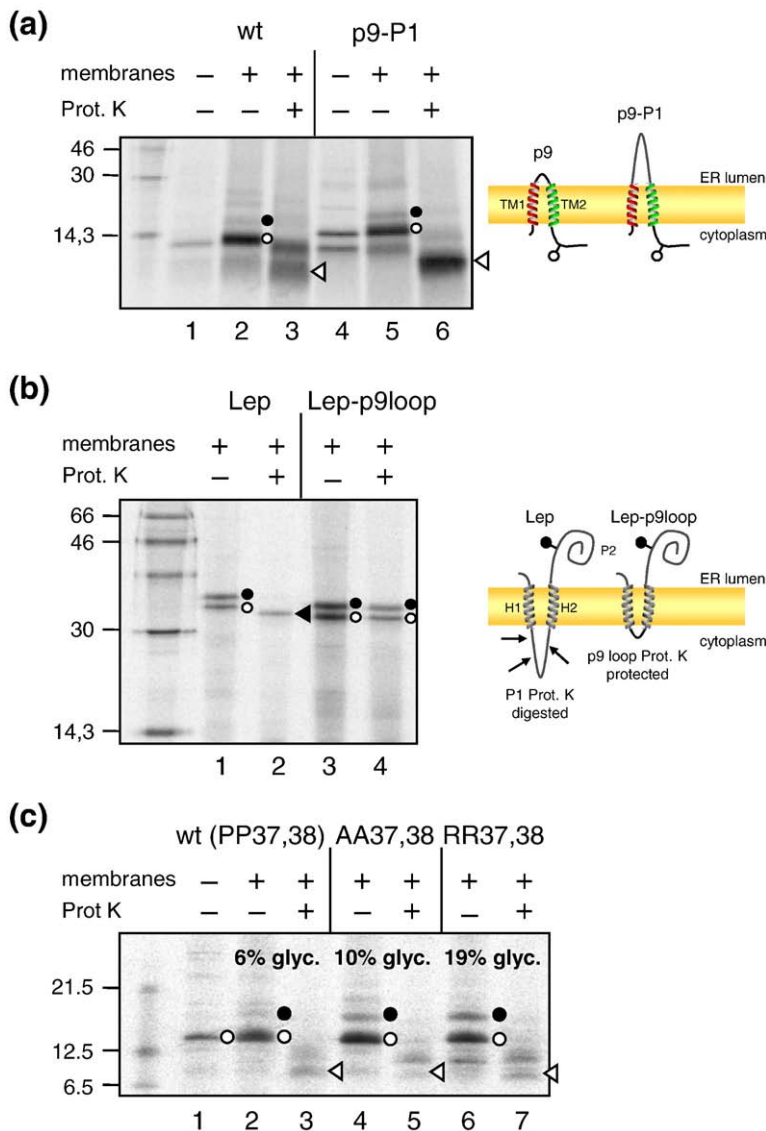


Fig. 4. Topogenic effect of the loop between the two TM segments. (a) Glycosylation of a p9 chimera containing the loop (P1 region) of Lep. (b) Glycosylation of a Lep chimera containing the loop of p9. (c) Topogenic effect of the tandem proline residues preceding the second TM segment. Open and filled circles indicate nonglycosylated and glycosylated polypeptides, respectively. Open and filled triangles indicated nonglycosylated and glycosylated protected fragments, respectively.

been found to be crucial to the translocation of these regions.⁴⁸

To elucidate the importance of the tandem proline residues in the translocation of this domain, these prolines were separately replaced with alanine and arginine residues (Fig. 4c). The alanine mutant, in which two proline residues were replaced with alanines, was glycosylated by about twice the wild-type (proline containing) sequence (Fig. 4c, compare lanes 2 and 4), whereas the arginine mutant (with the two prolines replaced with arginines) was glycosylated by more than three times the wild-type protein (Fig. 4c, lanes 2 and 6). The additive effect of the positive charges in this latter construct confirmed the mild topogenic function of the tandem proline residues.

TM2 insertion and topology

If p9 topology is not directed by the loop domain, it may be governed by a sequential alternation of hydrophobic segments (i.e., TM1 inserts in an N_{cyt} orientation, which in turn forces TM2 to adopt an

N_{out} orientation). Alternatively, the orientation of individual TM segments may be strictly independent of each other. To analyze the contribution of TM2 in the absence of the upstream domains of p9, a construct was designed in which both TM1 and the loop region were deleted (TM2/P2). In this construct, the P2 domain from Lep was fused downstream of the p9 sequence as a reporter gene to permit SDS-PAGE analysis. The *in vitro* translation of TM2/P2 reveals that the protein is not glycosylated and is sensitive to proteinase K digestion (Fig. 5, lanes 2 and 3), therefore indicating that TM2 precludes the translocation of the C-terminal region, as with the native p9 protein. This result demonstrates that TM2 contains topological information to adopt its native orientation, even in the absence of TM1.

Next we studied whether TM2 could govern protein topology. We swapped TM1 with TM2, and the P2 domain was fused at the C-t (TM2TM2/P2 mutant) to compare it with the previous p9 derivative. The *in vitro* translation of this chimera showed that the protein is not glycosylated and is sensitive to

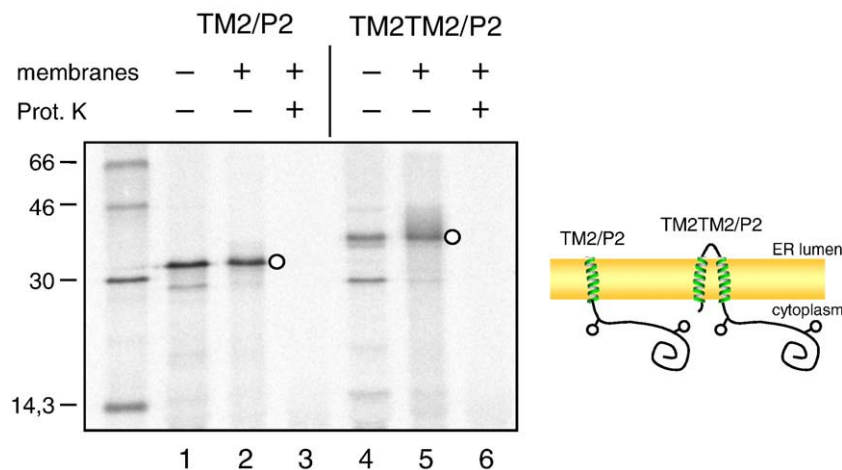


Fig. 5. Glycosylation of TM2-derived constructs. Residues 1–81 of Lep were replaced with residues 39–87 from p9, which includes TM2 plus the C-t of p9 to generate the TM2/P2 construct, and residues 1–29 of p9 (TM1) were replaced with residues 39–59 of p9 to generate the TM2TM2/P2 construct. The open circle denotes non-glycosylated proteins.

proteinase K treatment (Fig. 5, lanes 5 and 6), indicative of an $N_{\text{cyt}}\text{--}C_{\text{cyt}}$ topology, as with the native p9 protein, suggesting that TM2 behaves as TM1 when placed as a first TM segment.

We have previously shown that p9 is a double-spanning viral membrane protein that lacks a cleavable signal sequence, and that TM1 is recognized as a functional signal by cellular components.⁴³ We performed photocrosslinking experiments to study whether these chimeric proteins, where TM1 is absent, are properly targeted and inserted into the ER membrane. In eukaryotic cells, signal recognition particle (SRP) is responsible for the cotranslational targeting of membrane proteins to the ER membrane. In order to analyze the SRP role in TM2TM2 targeting, 70-residue nascent polypeptides were synthesized in a cell-free system by translating truncated mRNAs (Fig. 6a) in the presence of [³⁵S]methionine. In this system, the mRNAs transcribed from truncated cDNA possess no in-frame termination codon; therefore, nascent polypeptides are retained on ribosomes as peptidyl-tRNA, and the first hydrophobic segment remains outside the ribosome tunnel and is accessible for putative recognition by the SRP54, the subunit responsible for signal recognition. We introduced a photoreactive amino acid (eANB-Lys-tRNA^{amb}) roughly into the middle of the first TM2 sequence at position 13 (TAG13) (Fig. 6a). After UV radiation of the product translated *in vitro*, we detected a photoadduct band with a molecular weight equivalent to that of a complex formed by the SRP54 and the 70-residue nascent chains (Fig. 6b, lane 4). Hence, the SRP complex is responsible for the targeting of this p9 derivative, TM2TM2, as with the native protein⁴³ (Fig. 6b, compare lanes 2 and 4). To dissect the process of membrane integration, nascent chains of various lengths were synthesized in the presence of microsomal membranes and irradiated with UV light. Irradiated samples were split, and half of each sample was subjected to immunoprecipitation (IP) with antibodies against Sec61 α (Fig. 6c), the central component of the eukaryotic translocon through which TM segments are supposed to integrate into the membrane. The other half was immu-

noprecipitated with antibodies against the TRAM (translocating chain-associating membrane) protein (Fig. 6d), a translocon-associated protein that may be required as a chaperone to increase the efficiency of the insertion of certain TM segments. Strong cross-links to Sec61 α were detected for those nascent chains longer than 70 residues (Fig. 6c, lanes 6–13), demonstrating that when TM2 is used as a signal sequence, the protein integrates into the ER membrane through the protein-conducting channel. Moreover, crosslinks to TRAM proteins (Fig. 6d, lanes 15–22) indicate that the integration of the protein occurs with the assistance of this proposed membrane protein chaperone, as shown with the native p9 protein.⁴⁵

In order to study the effect of charge distribution on these TM2-derived constructs, we performed Arg60Glu and Lys75Glu replacements in the TM2/P2 construct (numbering according to the full-length p9 protein). As shown in Fig. 7a, the P2 domain was only barely glycosylated in the Arg60Glu construct (see lanes 2, 4, 6, and 8), even though this domain was translocated in its native Lep context (see scheme in Fig. 4b). Since the proteinase K treatment of these constructs renders a very small fragment (only TM2) that could not be detected in the SDS-PAGE analyses, we sought to report the membrane topology in *E. coli* cells by green fluorescent protein (GFP) C-terminal tagging. GFP is efficiently folded in the reducing environment of the cytoplasm,⁵⁰ whereas its translocation to the periplasm renders no fluorescent chimeras,⁵¹ allowing membrane protein topology analysis. When tested under a fluorescence microscope, the wild-type TM2-derived construct (TM2/GFP) was found to be highly fluorescent (Fig. 7b), as expected. Replacement of the positively charged residue, just flanking the TM2 hydrophobic region by a negatively charged residue (construct Arg60Glu), significantly reduced the fluorescence signal. This result is indicative of C-t translocation and is in good agreement with the stronger effect of the positive-inside rule in bacteria.³⁷ When the positively charged residue at position 75 was replaced with a negatively charged residue (Lys75Glu), the

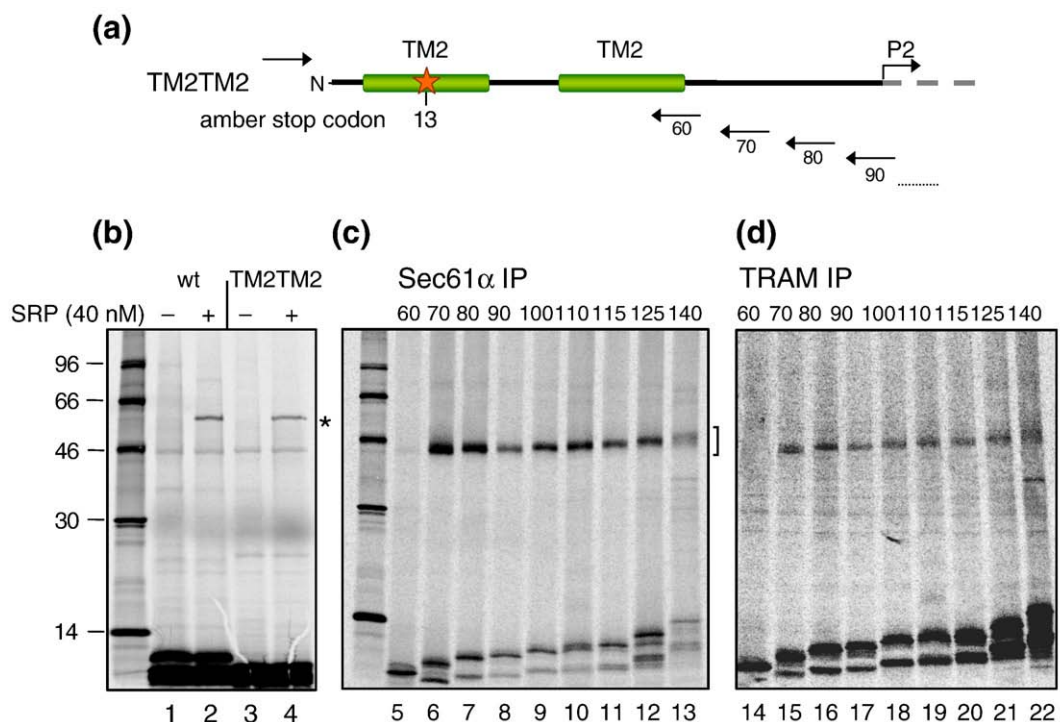


Fig. 6. Targeting and insertion of the TM2 segment. (a) Structural organization of the TM2TM2 construct. A single photoreactive probe was incorporated by positioning an amber stop codon at position 13. (b) Photocrosslinking of p9 (wt) and TM2TM2 nascent chains to SRP. Ribosome nascent chains (RNCs) containing 70-residue radioactive nascent chains carrying a photoreactive probe were photolysed and analyzed by SDS-PAGE. Where indicated, the translation and subsequent photocrosslinking were performed in the presence of 40 nM purified SRP (+ lanes). An asterisk (*) indicates RNC-SRP photoadducts. (c) Photocrosslinking of TM2TM2 nascent chains to Sec61 α . Integration intermediates containing radioactive nascent chains of 60 or 140 residues were photolysed and examined by IP using Sec61 α antibody. (d) Photocrosslinking of TM2TM2 nascent chains to TRAM. Integration intermediates containing radioactive nascent chains of 60 or 140 residues were photolysed and examined by IP as in (c), but using antibodies specific for TRAM. Brackets signal the photoadducts.

TM2/GFP protein remained, as in the wild-type protein, with a C_{cyt} orientation (Fig. 7b, third panel). In good agreement with the positive-inside rule, the double mutant Arg60Glu/Lys75Glu mainly translocated the GFP moiety, consistent with the importance of arginine replacement on protein orientation. Fusion proteins were over-expressed, purified using Ni²⁺-NTA chromatography, and analyzed by SDS-PAGE to verify the expression of all the constructs (Fig. 7c). The presence of the GFP moiety in all fusions was confirmed by immunoblotting with a GFP antibody (not shown). The amount of protein synthesis was used to normalize the activity of the GFP fusion proteins in a fluorometer. The typical GFP fluorescence emission spectra from liquid cultures (whole cells) expressing TM2/GFP derivatives are shown in Fig. 7d, and a control nonfluorescent chimera previously shown to fully translocate the GFP moiety was included.⁴⁹ When cultures were grown at 37 °C, only cells expressing the wild-type TM2/GFP and K75E mutation fusions displayed typical emission at 508 nm (Fig. 7d), thus confirming the results observed under microscopy and suggesting a greater effect of the charge balance on TM2 orientation in the bacterial system than in ER-derived membranes.

Topology is acquired at the initial stages of protein biogenesis

In order to monitor the topology at the initial stages of protein biosynthesis, we enlarged the loop between the two TM segments in our viral membrane protein by inserting the N-terminal domain (without the signal sequence) from the berberine bridge enzyme (bbe) from opium poppy.⁵² This large fragment facilitated topology analysis by virtue of both its length and its N-linked glycosylation site (Fig. 8, top). This plant protein domain has been demonstrated to be capable of canine microsomal translocation with concomitant glycosylation.⁵³ First, we demonstrated that the insertion of this domain did not alter the overall protein topology. The *in vitro* translation of a construct carrying the glycosylation site in the loop region was efficiently glycosylated when compared with a control construct where the glycosylation site was erased by an asparagine-to-glutamine replacement (Fig. 8, compare lanes 2 and 5). These results demonstrated that the insertion of this hydrophilic domain did not modify the native p9 topology, and that its p9 derivative forms maintained an N_{cyt}-C_{cyt} topology.

To study the topogenesis of these p9 variants in greater detail, we generated a series of integration

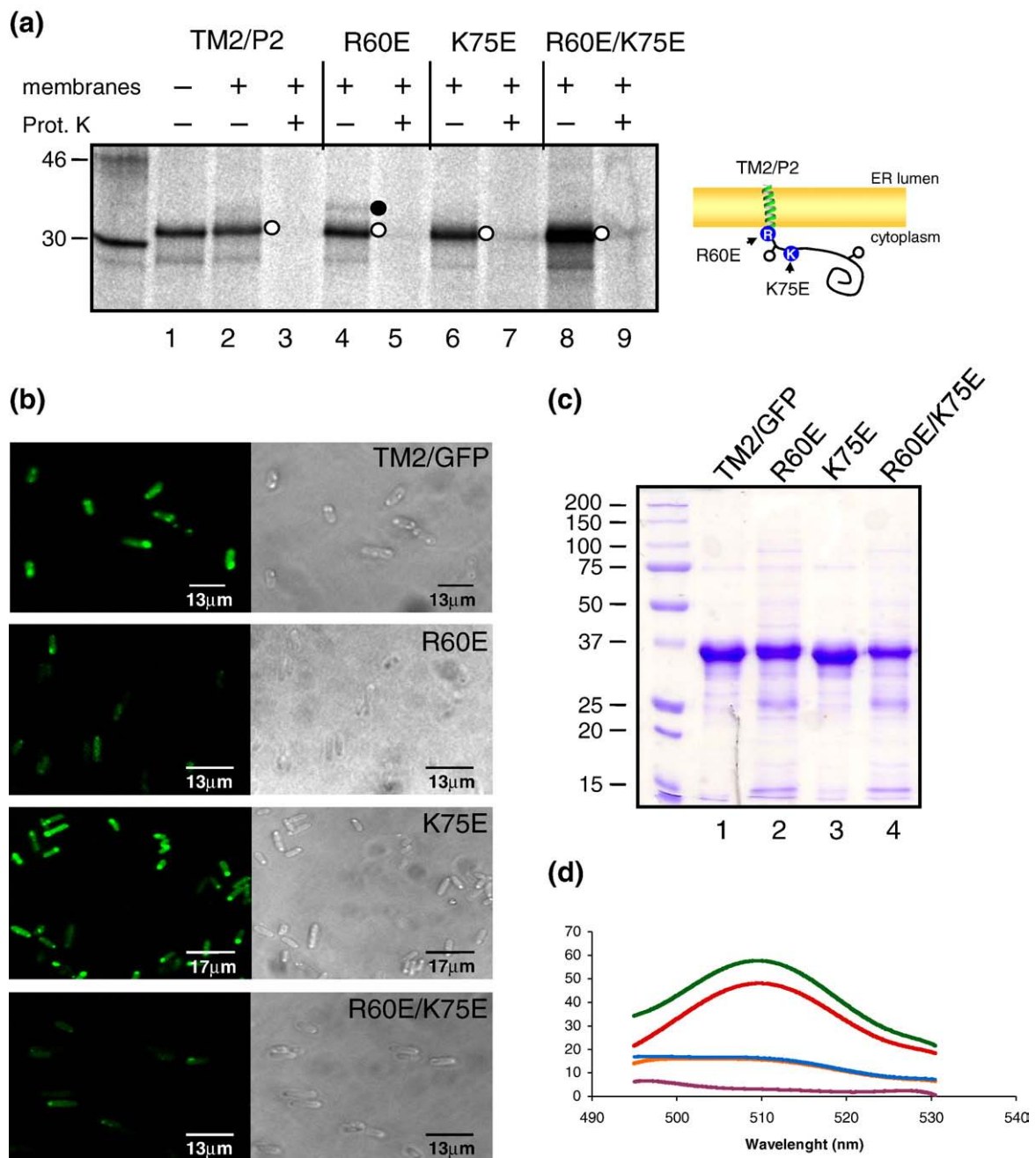


Fig. 7. Influence of charged residues in TM2 orientation *in vitro* and in *E. coli* cell membranes. (a) *In vitro* translation of the indicated constructs. Open and filled circles indicate nonglycosylated and glycosylated proteins, respectively. (b) Fluorescence microscope analysis of *E. coli* cells expressing the GFP fused to the C-t of the analyzed proteins. Cells were prepared for analysis as described in [Materials and Methods](#). White light exposures of the cells and UV-excited images are shown to the right and left, respectively. Cells containing TM2/GFP fusions with the indicated mutations are shown. (c) Overexpression and purification of GFP fusion proteins. Coomassie staining of SDS-PAGE gels representing fusion proteins purification by Ni^{2+} -NTA chromatography. Molecular weight markers are shown (left). (d) GFP fluorescence emission spectra of cells expressing the GFP fusion proteins grown at 37 °C. TM2/GFP spectrum is shown in green, R60E is shown in orange, K75E is shown in red, and the double mutant R60E/K75E is shown in blue. A control construct with the fully translocated GFP moiety (p7B/GFP in Martinez-Gil *et al.*⁴⁹) is also included, and the spectrum is shown in purple.

intermediates and used glycosylation as a reporter for ER lumen translocation since this modification takes place during the insertion and translocation of nascent chains. Parallel analyses of the construct in which the glycosylation site was erased (p9-bbe-QFT) were used as control. Integration intermediates were generated by the translation of truncated

mRNAs as described above. Translation in the presence of microsomal membranes generates integration intermediates of increasing length, each with its C-t still associated with the tRNA in the ribosome. To increase the incorporation of radioactivity, the last two amino acids of each integration intermediate were changed to methionines. Samples were

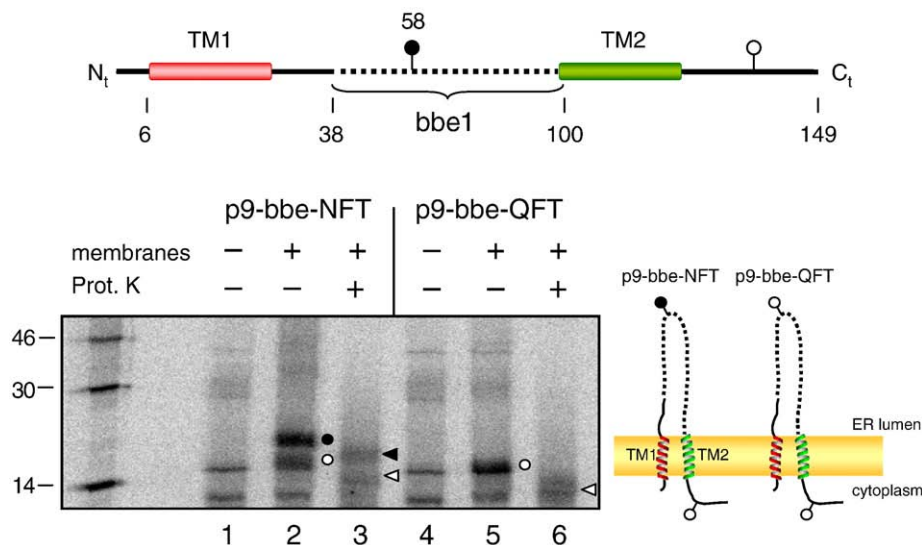


Fig. 8. Glycosylation of p9-bbe constructs. Residues 21–83 of the bbe1 template were inserted between the p9 loop region and the TM2 segment. The structural organization of these constructs is shown at the top. A glycosylation site is included in the bbe1 sequence (position 58), and all experiments were performed in parallel with a mutant in which the target asparagine was replaced with a nonglycosylable glutamine. Open and filled circles denote nonglycosylated and glycosylated chimeras, respectively. Open and filled triangles represent protected nonglycosylated and glycosylated fragments, respectively.

split into thirds, and one-third was analyzed directly by SDS-PAGE, while the rest was treated with puromycin, a reagent that releases nascent chains from ribosomes. Then, half of the puromycin-released sample was subjected to proteinase K treatment prior to SDS-PAGE analysis. To ensure proper protein targeting and integration, we used nascent chains that were long enough to allow complete membrane targeting and insertion of TM1.⁴³

The nascent chains of 99 residues (mRNA was truncated so that protein synthesis halts 41 residues after the asparagine residue at the glycosylation site; Fig. 9a) in Fig. 9b were truncated just before TM2. As expected, the main product in lane 1 corresponded to the unglycosylated nascent chain (empty circle), since 41 residues of a translocating peptide are not enough to bridge the distance between the P-site in ER-bound ribosomes and the lumenally disposed active site of the OST.⁵⁴ In addition, this polypeptide form was fully protected from proteinase K treatment (Fig. 9b, lane 3), indicating that this translation intermediate had been targeted to the membrane, suggesting an N_{cyt} orientation of the TM1 with a translocated loop region. The higher-molecular-weight product (indicated by an asterisk in Fig. 9) was identified by treating the translation reaction with puromycin, an amino-acyl-tRNA analogue that was incorporated into nascent peptide chains by the ribosomal peptidyl transferase, thereby dissociating the nascent chain-tRNA bond (Fig. 9b, lanes 2, 3, 5, and 6).⁵⁵ Thus, nascent peptide attached to a tRNA molecule yielded the observed larger molecular mass band, consistent with measurements of molecular mass for tRNAs based on its mobility in SDS-PAGE.⁵⁶ Therefore, the low level of glycosylation observed in the puromycin-released sample (lane 2)

must be due to the expected high level of flexibility of this C-terminal hydrophilic nascent chain region (approximately 75 residues from the end of the TM1 segment to the very end of the nascent chain), which would affect OST efficiency.

Nascent chains of 112 amino acids, in which only the first half of TM2 was synthesized, were essentially not glycosylated either (Fig. 9c, lane 1). At this length, the glycosylation site had emerged from the ribosome (there are 54 residues from the P-site to the glycosylation site; see scheme in Fig. 9c), but it had still not reached the active site of the OST. Interestingly, puromycin-released nascent chains showed an increased level of glycosylation, which is very likely due to the membrane-anchoring capacity of the TM2-synthesized portion that increases OST efficiency. These results correlate with an N_{cyt} orientation of the nascent chain.

Similar experiments were performed with nascent chains of 132 residues (Fig. 9d). At this length, TM2 is fully synthesized but still remains inside the ribosomal exit tunnel, while the distance between the P-site and the glycosylation site (74 residues) is sufficient to observe efficient glycosylation, even for sequences with strong helix-forming propensity located inside the ribosomal exit tunnel.⁵⁷ As shown in lane 1, nascent chains are nicely glycosylated, indicating that the loop region is exposed to the lumen of the ER membrane before protein translation termination. The glycosylation level achieved did not increase with puromycin treatment (lane 2), corresponding to the utmost achievable glycosylation level. Furthermore, the glycosylated form of the nascent chain (filled circle, lane 3) was resistant to proteinase K treatment, indicative of proper integration into the ER mem-

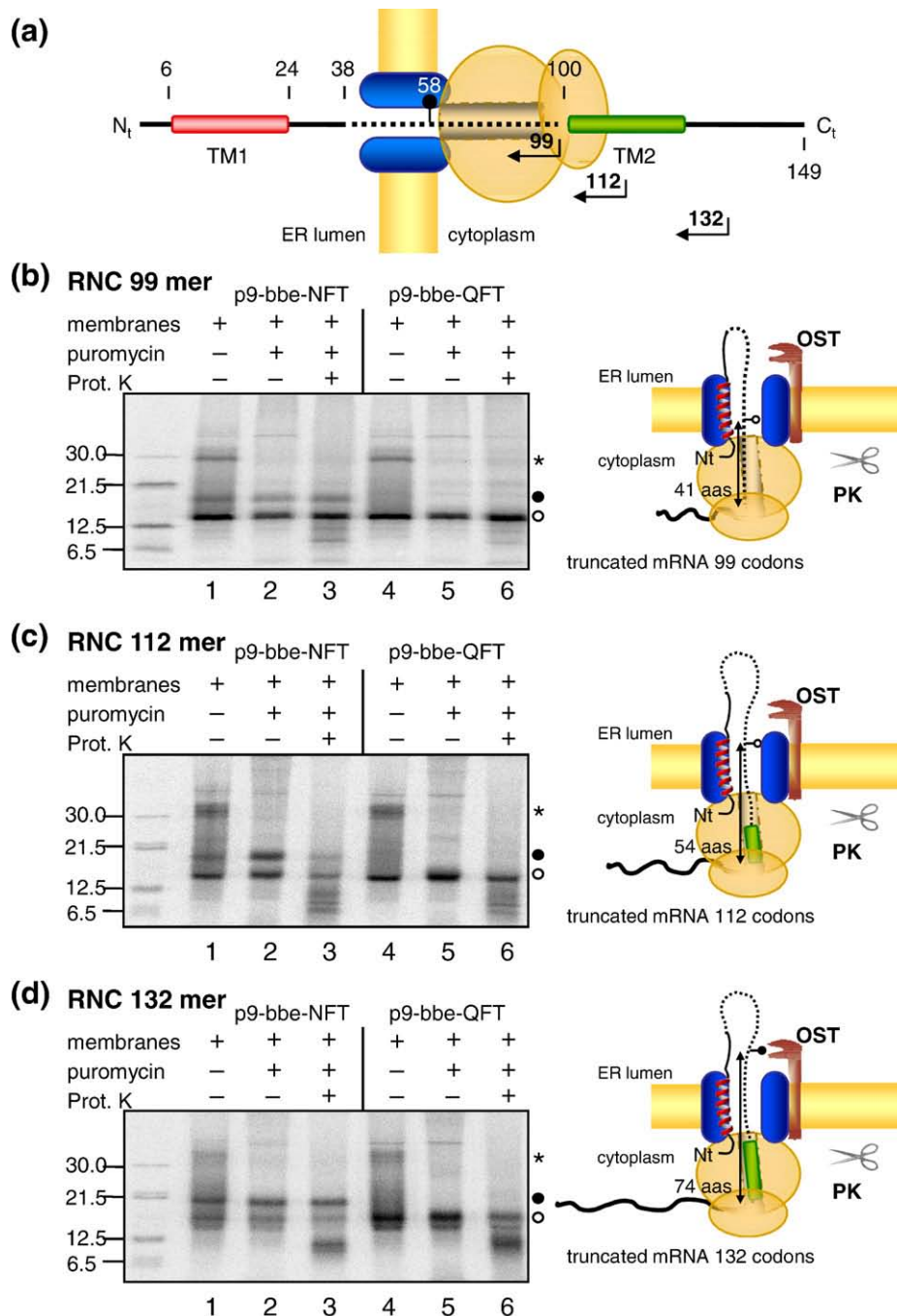


Fig. 9. Topology is acquired during protein biogenesis. (a) Schematic representation of the RNCs used in these experiments. Region 21–83 of the secreted bbe1 protein (dotted line) was inserted preceding TM2 (striped box). A filled circle indicates the glycosylation site. Numbers indicate the amino acid residues at each position. Open and filled circles denote nonglycosylated and glycosylated polypeptides, respectively. Translation *in vitro* in the presence of microsomal membranes of truncated mRNAs of different lengths: 99 residues (b), 112 residues (c), and 132 residues (d). One aliquot was directly analyzed (lanes 1 and 4) to monitor protein synthesis and glycosylation. After incubation with puromycin (lanes 2 and 5), proteinase K protection assays were performed (lanes 3 and 6). RNCs with a glycosylation site (p9-bbe-NFT) and a nonglycosylable construct (p9-bbe-QFT) were processed in parallel. An asterisk (*) indicates nascent chains bound to tRNA. Open and filled circles denote nonglycosylated and glycosylated nascent chains, respectively.

brane after release of the nascent chain with puromycin treatment. These results demonstrate that TM1 adopts an $N_{\text{cyt}}-C_{\text{out}}$ topology before other downstream topological parameters reach the translocon channel.

Discussion

Membrane protein topology is assumed to be established while the nascent chains pass through the

translocon.³⁷ Among the topological determinants that govern membrane protein topology, the prominence of positively charged residues flanking the hydrophobic domains on the cytoplasmic face has been, by far, the more extensively demonstrated (positive-inside rule). For bacterial membrane proteins, this has been generally attributed to the TM potential and also to the fact that negatively charged lipids are more abundant in the cytoplasmic bilayer leaflet. As regards eukaryotic membrane proteins, it has been assumed that rather than the positive-inside rule, it is the net charge difference between the extramembranous regions flanking the hydrophobic TM domains that drives membrane protein topology. Additional determinants are the length and hydrophobicity of the apolar region. The analyses of model proteins showed that very hydrophobic TM segments integrate into the membrane with an N_{out} topology.^{42,58} Recently, the translocon has been implicated in the recognition of both topological determinants (the presence of charges in TM flanking regions and the hydrophobicity of the apolar domain).^{18,19,21}

In this study, we carried out a systematic analysis of the topologies adopted by model viral membrane proteins with two TM segments and different distributions of positively charged residues along the extramembranous regions. We also analyzed constructs in which the hydrophobicity and identity of the first TM segment varied, as had the length of the loop connecting both TM segments. The results are summarized as follows.

Firstly, the wild-type protein displayed a defined topology despite the absence of a charge difference bias between both sides of the membrane. Thus, the positive-inside rule could not be the only topological determinant present in this viral membrane protein. Furthermore, mutants 28K₃29, 35K₃36, 3E₃4, R60E, K75E, and R60E/K75E maintained the wild-type N_{cyt} - C_{cyt} orientation despite a clearly skewed charge distribution (Fig. 1). Notably, when testing the charge mutations on isolated parts of the protein, mutant 28K₃29 showed the opposite N_{out} orientation in the TM1/P2 construct (Fig. 2), as would be predicted by the positive-inside rule. Consistent with this prediction, our data demonstrate that the presence of positively charged residues flanking TM1 is important in determining the topology of this segment in the absence of TM2, although positive charges were somewhat less influential on full-length protein topology. This effect became even more evident when topological studies were performed in *E. coli* membranes, where reversion of C-t-positive charges (mutants R60E and R60E/K75E) dramatically reduced fluorescence emission (Fig. 7b and d), probably due to the existence of a TM potential across the inner membrane and the biased anionic phospholipids distribution in prokaryotic cells.

Secondly, the proteins in constructs with an increased hydrophobicity of TM1 retained the original N_{cyt} orientation, and only the replacement of this first hydrophobic segment with TM segments

with a reversed N_{out} orientation had a slightly weak effect (Fig. 3), with the majority of the polypeptides remaining in a N_{cyt} - C_{cyt} orientation. These results clearly contrasted with that found for nonviral proteins where charge distribution exerted a major influence on the orientation of sequences with different hydrophobicities.⁹

Thirdly, in those constructs where the p9 luminal loop was swapped with the cytosolic loop of Lep, both proteins maintained their original topology, suggesting that the p9 loop was not a prominent determinant, although some effect on the reversion of topology was observed by mutating the tandem proline residues present in this loop region to arginines (Fig. 4), probably due to the additive contribution of the positively charged residues.

Fourthly, TM2 in the absence of TM1 adopts the same topology as in full-length constructs (Fig. 5). In addition, crosslinking experiments (Fig. 6) suggest that TM2 is targeted to the ER-derived membranes by the SRP and is integrated into the membrane through the translocon (Sec61 α) with the assistance of TRAM in the absence of any signal sequence and TM1. As previously observed,^{43,45} a striking feature of these viral TM segments is its constant association with both the Sec61 complex and TRAM throughout nascent chain biosynthesis. Even when integration intermediates were artificially extended by the addition of the P2 domain of Lep at the C-t, TM2 remained close to the translocon components. A similar behavior was recently described for opsin TM7,⁵⁹ and it was postulated that specific properties and assembly requirements, mainly in hydrophobicity terms, could account for either the rapid release into the lipid bilayer or the selective retention of TM segments at the ER translocon. Whether the viral nature of the sequences assayed in the present work is responsible for translocon retention remains to be established.

Finally, topogenesis experiments using truly integration intermediates demonstrate that the protein adopts its native N_{cyt} orientation with the translocated loop even before TM2 reaches the translocon. This means that a crosstalk is dispensable between the two TM segments in order to acquire the final protein topology.

Therefore, the collective results indicate that no single domain in p9 is solely responsible for protein topology, but that multiple contributions along the entire p9 molecule collaborate in acquiring a defined membrane protein topology from the initial stages of protein biogenesis, thus indicating that the determinants of viral membrane protein topology are more complex than the positive-inside rule and the hydrophobicity of the first TM segment initially suggested. In addition, it cannot be ruled out that other membrane components play a relevant role in viral membrane protein topogenesis. For instance, it has been established that lipid composition plays an active role in determining the topological organization of Lep-derived constructs⁶⁰ and in multi-spanning membrane proteins.⁶¹ Furthermore, the presence of membrane protein chaperones such as

TRAM, which has been found adjacent to several integration intermediates of a viral origin,^{45,62} could perform relevant functions in properly orienting these viral components. In that sense, it is tempting to think that viral membrane protein topology could have coevolved with the lipid and protein environment of its natural hosts, thus emphasizing the importance of studying not only membrane protein topology but also its assembly and structure in a native context.

Materials and Methods

Enzymes and chemicals

Unless otherwise stated, all enzymes, as well as plasmid pGEM1, RiboMAX SP6 RNA polymerase system, and rabbit reticulocyte lysate, were purchased either from Promega (Madison, WI) or Roche (Basel, Switzerland). [³⁵S]Met and ¹⁴C-methylated marker proteins were obtained from GE Healthcare (Uppsala, Sweden). The PCR purification and RNeasy RNA clean up kits were obtained from Qiagen (Hilden, Germany). The PCR mutagenesis kit QuikChange was obtained from Stratagene (La Jolla, CA), while the oligonucleotides were obtained from Isogen (Maarsse, The Netherlands).

Constructs

All the constructs used were obtained from the plasmid pGEM1-p9, as previously described.³⁶ For the p9-H1, p9-GpA, and TM2TM2 constructions, a BglII restriction site (AGATCT) was introduced immediately after the TM1 of p9 at nucleotides 63–68 (p9-BglII). With an NcoI/BglII digestion, residues 1–29 of p9 (TM1) were replaced with residues 1–22 of Lep (p9-H1), residues 73–95 of glycophorin A (p9-GpA),⁶³ and residues 39–59 of p9 (TM2TM2). For the p9-P1 construct, a SpeI restriction site (ACTAGT) before TM2 was used, as previously described.⁴³ For the p9-bbe construct, the SpeI restriction site was used. After SpeI digestion, 62 residues from a plant secretion protein (residues 21–83), bbe,^{52,53} were inserted downstream of the natural p9 loop. Site-directed mutagenesis was performed using the QuikChange mutagenesis kit from Stratagene in accordance with the manufacturer's instructions. Overall, we followed the same protocol for all the constructs. Briefly after PCR amplification, the PCR products were purified, digested, and ligated to the corresponding p9 vectors, digested with the same enzymes. Constructs were confirmed by DNA sequencing.

In vitro transcription and translation

In vitro transcription of p9 derivatives was performed as previously described.³⁶ *In vitro* translation of purified mRNA was performed in the presence of reticulocyte lysate, [³⁵S]Met, and dog pancreas microsomes, when indicated. After completing the translation, the sample was split into two: half of the sample was digested with 1 mg/ml proteinase K for 30 min on ice, and the remaining sample was kept as untreated control. Afterwards, samples were alkaline-extracted and ultracentrifuged (100,000g for 20 min) onto a sucrose cushion. Pellets were redissolved in 50 µl of sample buffer, and after they had been boiled at 95 °C for 5 min, they were electro-

phoresed on 16% SDS-PAGE and analyzed by Phosphor-Imaging. The glycosylation levels of the p9 derivative proteins were quantified using the QuantityOne software (BioRad) by comparing the glycosylated and nonglycosylated forms for each analyzed lane.

Photocrosslinking experiments

Truncated mRNAs were generated by PCR using different reverse primers that lacked a stop codon to obtain nascent chains of a specific length. PCR products were transcribed *in vitro* using purified SP6 RNA polymerase, as described above. For the SRP photocrosslinking experiments, an *in vitro* translation of a 70-residue-long truncated mRNA was performed as before^{12,43} in a wheat germ cell-free extract containing 40 nM SRP, 10 mCi of [³⁵S]Met, and 32 pmol of εANB-Lys-tRNA^{amb}. After translation, samples were photolysed, and SDS-PAGE was analyzed as before.^{43,45}

To assess Sec61α and TRAM photocrosslinking, truncated mRNAs were translated as described above, with the exception that samples contained 8 Eq of column washed rough ER microsomes. Samples were photolysed and sedimented as described above prior to analysis by IP.

Immunoprecipitation

Pelleted membranes were resuspended in 50 µl of 1% (wt/vol) SDS and 100 mM Tris-HCl (pH 7.5) and incubated at 55 °C for 30 min. Samples were washed three times with 150 µl of buffer A [140 mM NaCl, 10 mM Tris-HCl (pH 7.5), and 2% (vol/vol) Triton X-100] for Sec61α IPI, and with buffer B [150 mM NaCl, 50 mM Tris-HCl (pH 7.5), 2% (vol/vol) Triton X-100, and 0.2% (wt/vol) SDS] for TRAM IP. Samples were precleared by rocking with 40 µl of buffer-A/buffer-B-washed protein A-Sepharose (Sigma) at room temperature for 1 h. After removal of the beads by centrifugation at room temperature, the supernatants were incubated with affinity-purified rabbit antisera specific for Sec61α or TRAM (Research Genetics, Huntsville, AL) overnight at 4 °C. For IP, 40 µl of protein A-Sepharose, which was previously equilibrated with buffer A or buffer B, was then added and incubated for 4 h at 4 °C. IPs were harvested by sedimentation and washed twice with 750 µl of buffer A or buffer B, followed by a final washing in the same buffer without detergent. Samples were prepared for SDS-PAGE analysis by adding 25 µl of sample buffer and incubating at 37 °C for 10 min. The results were visualized and processed by PhosphorImaging.

Expression of GFP fusion proteins in *E. coli* membranes

Fluorescence microscope analysis of *E. coli* cells expressing the GFP fused to the C-t of the analyzed proteins was performed as described previously.⁴⁹ Briefly, a single colony was grown overnight at 37 °C in the presence of antibiotics, and cultures were diluted and grown at 37 °C. When optical density reached 0.4–0.5, IPTG was added at a final concentration of 0.4 mM. Cells were grown for another 3 h and then harvested. After resuspension, cells were stored at room temperature for 2 h to enhance GFP folding and used directly to collect emission spectra between 495 and 530 nm, with excitation at 485 nm determined in a Perkin-Elmer LS 50 fluorometer. For a more detailed protocol, see Drew *et al.*⁶⁴ In order to visualize

whole-cell cultures, 10 µl of cell suspension was fixed. Samples were initially observed under white light, and the same field was then analyzed by using UV laser excitation at 488 nm and by recording emission wavelengths between 500 and 600 nm under a Leica TCS-SP confocal microscope (SCSIE; Universitat de València).

Generation and translation of protein truncates

For the topological studies performed during p9 biogenesis, the protein truncates of p9-bbe-NFT and p9-bbe-QFT of 99, 112, and 132 residues were generated by PCR using the corresponding backward oligonucleotides lacking stop codons. Purified PCR products were transcribed *in vitro* using the SP6 RiboMAX system. The generated mRNAs were translated *in vitro* for 30 min at 30 °C, as described above. The translation reaction was inhibited by the addition of 1 mg/ml cycloheximide for 10 min at 30 °C. Samples were split into three equivalent fractions. One aliquot was kept as untreated control, while the other two fractions were treated with 2 mM puromycin for 30 min at 30 °C. After puromycin treatment, only one of the latter two fractions was digested with 1 mg/ml proteinase K for 30 min on ice. Samples were ultracentrifuged on sucrose cushion as described above and analyzed by 16% SDS-PAGE.

Acknowledgements

We wish to thank David Bird for donating the bbe1 DNA template. This work was supported by grant BFU2006-01532 from the Spanish MEC and grant ACOMP07-119 from the Generalitat Valenciana (to I.M.), and by National Institutes of Health grant R01 GM26494 and by the Robert A. Welch Foundation (to A.E.J.). A.S. was a recipient of a predoctoral FPU fellowship from the Spanish MEC. S.T. was a recipient of a predoctoral fellowship from the University of Valencia (V Segles).

References

1. Alder, N. N. & Johnson, A. E. (2004). Cotranslational membrane protein biogenesis at the endoplasmic reticulum. *J. Biol. Chem.* **279**, 22787–22790.
2. Johnson, A. E. & van Waes, M. A. (1999). The translocon: a dynamic gateway at the ER membrane. *Annu. Rev. Cell Dev. Biol.* **15**, 799–842.
3. von Heijne, G. (2006). Membrane-protein topology. *Nat. Rev. Mol. Cell. Biol.* **7**, 909–918.
4. 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.
5. Gafvelin, G., Sakaguchi, M., Andersson, H. & von Heijne, G. (1997). Topological rules for membrane protein assembly in eukaryotic cells. *J. Biol. Chem.* **272**, 6119–6127.
6. Gavel, Y. & von Heijne, G. (1992). The distribution of charged amino acids in mitochondrial inner membrane proteins suggests different modes of membrane integration for nuclearly and mitochondrially encoded proteins. *Eur. J. Biochem.* **205**, 1207–1215.
7. Gavel, Y., Steppuhn, J., Herrmann, R. & von Heijne, G. (1991). The positive-inside rule applies to thylakoid membrane proteins. *FEBS Lett.* **282**, 41–46.
8. Sato, T., Sakaguchi, M., Mihara, K. & Omura, T. (1990). The amino-terminal structures that determine topological orientation of cytochrome-P-450 in microsomal membrane. *EMBO J.* **9**, 2391–2397.
9. Wahlberg, J. M. & 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.
10. Rapoport, T. A. (2008). Protein transport across the endoplasmic reticulum membrane. *FEBS J.* **275**, 4471–4478.
11. Higgy, M., Junne, T. & Spiess, M. (2004). Topogenesis of membrane proteins at the endoplasmic reticulum. *Biochemistry*, **43**, 12716–12722.
12. McCormick, P. J., Miao, Y., Shao, Y., Lin, J. & 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.
13. Heinrich, S. U., Mothes, W., Brunner, J. & Rapoport, T. A. (2000). The Sec61p complex mediates the integration of a membrane protein by allowing lipid partitioning of the transmembrane domain. *Cell*, **102**, 233–244.
14. Martoglio, B., Hofmann, M. W., Brunner, J. & 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.
15. Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H. *et al.* (2005). Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature*, **433**, 377–381.
16. Hessa, T., Meindl-Beinker, N. M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M. *et al.* (2007). Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature*, **450**, 1026–1030.
17. Van den Berg, B., Clemons, W. M., Jr, Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C. & Rapoport, T. A. (2004). X-ray structure of a protein-conducting channel. *Nature*, **427**, 36–44.
18. Goder, V., Junne, T. & Spiess, M. (2004). Sec61p contributes to signal sequence orientation according to the positive-inside rule. *Mol. Biol. Cell*, **15**, 1470–1478.
19. Junne, T., Schwede, T., Goder, V. & Spiess, M. (2007). Mutations in the Sec61p channel affecting signal sequence recognition and membrane protein topology. *J. Biol. Chem.* **282**, 33201–33209.
20. Shaw, A. S., Rottier, P. J. & Rose, J. K. (1988). Evidence for the loop model of signal-sequence insertion into the endoplasmic reticulum. *Proc. Natl Acad. Sci. USA*, **85**, 7592–7596.
21. Goder, V. & Spiess, M. (2003). Molecular mechanism of signal sequence orientation in the endoplasmic reticulum. *EMBO J.* **22**, 3645–3653.
22. Blobel, G. (1980). Intracellular protein topogenesis. *Proc. Natl Acad. Sci. USA*, **77**, 1496–1500.
23. Goder, V., Bieri, C. & Spiess, M. (1999). Glycosylation can influence topogenesis of membrane proteins and reveals dynamic reorientation of nascent polypeptides within the translocon. *J. Cell Biol.* **147**, 257–266.
24. Sato, Y., Sakaguchi, M., Goshima, S., Nakamura, T. & Uozumi, N. (2003). Molecular dissection of the contribution of negatively and positively charged residues in S2, S3, and S4 to the final membrane

- topology of the voltage sensor in the K⁺ channel, KAT1. *J. Biol. Chem.* **278**, 13227–13234.
25. Ota, K., Sakaguchi, M., Hamasaki, N. & Mihara, K. (2000). Membrane integration of the second transmembrane segment of band 3 requires a closely apposed preceding signal-anchor sequence. *J. Biol. Chem.* **275**, 29743–29748.
 26. Lu, Y., Turnbull, I. R., Bragin, A., Carveth, K., Verkman, A. S. & Skach, W. R. (2000). Reorientation of aquaporin-1 topology during maturation in the endoplasmic reticulum. *Mol. Biol. Cell*, **11**, 2973–2985.
 27. Nilsson, I., Witt, S., Kiefer, H., Mingarro, I. & von Heijne, G. (2000). Distant downstream sequence determinants can control N-tail translocation during protein insertion into the endoplasmic reticulum membrane. *J. Biol. Chem.* **275**, 6207–6213.
 28. Lambert, C. & Prange, R. (2001). Dual topology of the hepatitis B virus large envelope protein: determinants influencing post-translational pre-S translocation. *J. Biol. Chem.* **276**, 22265–22272.
 29. Cocquerel, L., Op de Beeck, A., Lambot, M., Roussel, J., Delgrange, D., Pillez, A. *et al.* (2002). Topological changes in the transmembrane domains of hepatitis C virus envelope glycoproteins. *EMBO J.* **21**, 2893–2902.
 30. McGinnes, L. W., Reitter, J. N., Gravel, K. & Morrison, T. G. (2003). Evidence for mixed membrane topology of the Newcastle disease virus fusion protein. *J. Virol.* **77**, 1951–1963.
 31. Lundin, M., Lindstrom, H., Gronwall, C. & Persson, M. A. (2006). Dual topology of the processed hepatitis C virus protein NS4B is influenced by the NS5A protein. *J. Gen. Virol.* **87**, 3263–3272.
 32. Moise, A. R., Grant, J. R., Lippe, R., Gabathuler, R. & Jefferies, W. A. (2004). The adenovirus E3-6.7K protein adopts diverse membrane topologies following post-translational translocation. *J. Virol.* **78**, 454–463.
 33. Zhang, S. C., Zhang, G., Yang, L., Chisholm, J. & Sanfacon, H. (2005). Evidence that insertion of Tomato ringspot nepovirus NTB-VPg protein in endoplasmic reticulum membranes is directed by two domains: a C-terminal transmembrane helix and an N-terminal amphipathic helix. *J. Virol.* **79**, 11752–11765.
 34. Vilar, M., Sauri, A., Marcos, J. F., Mingarro, I. & Perez-Paya, E. (2005). Transient structural ordering of the RNA-binding domain of carnation mottle virus p7 movement protein modulates nucleic acid binding. *ChemBioChem*, **6**, 1391–1396.
 35. Vilar, M., Esteve, V., Pallas, V., Marcos, J. F. & Perez-Paya, E. (2001). Structural properties of carnation mottle virus p7 movement protein and its RNA-binding domain. *J. Biol. Chem.* **276**, 18122–18129.
 36. Vilar, M., Sauri, A., Monne, M., Marcos, J. F., von Heijne, G., Perez-Paya, E. & Mingarro, I. (2002). Insertion and topology of a plant viral movement protein in the endoplasmic reticulum membrane. *J. Biol. Chem.* **277**, 23447–23452.
 37. Goder, V. & Spiess, M. (2001). Topogenesis of membrane proteins: determinants and dynamics. *FEBS Lett.* **504**, 87–93.
 38. Nilsson, I., Kelleher, D. J., Miao, Y., Shao, Y., Kreibich, G., Gilmore, R. *et al.* (2003). Photocross-linking of nascent chains to the STT3 subunit of the oligosaccharyltransferase complex. *J. Cell Biol.* **161**, 715–725.
 39. Hartmann, E., Rapoport, T. A. & Lodish, H. F. (1989). Predicting the orientation of eukaryotic membrane proteins. *Proc. Natl Acad. Sci. USA*, **86**, 5786–5790.
 40. Kida, Y., Morimoto, F., Mihara, K. & Sakaguchi, M. (2006). Function of positive charges following signal-anchor sequences during translocation of the N-terminal domain. *J. Biol. Chem.* **281**, 1152–1158.
 41. Lerch-Bader, M., Lundin, C., Kim, H., Nilsson, I. & von Heijne, G. (2008). Contribution of positively charged flanking residues to the insertion of transmembrane helices into the endoplasmic reticulum. *Proc. Natl Acad. Sci. USA*, **105**, 4127–4132.
 42. Sakaguchi, M., Tomiyoshi, R., Kuroiwa, T., Mihara, K. & 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. USA*, **89**, 16–19.
 43. Sauri, A., Saksena, S., Salgado, J., Johnson, A. E. & Mingarro, I. (2005). Double-spanning plant viral movement protein integration into the endoplasmic reticulum membrane is signal recognition particle-dependent, translocon-mediated, and concerted. *J. Biol. Chem.* **280**, 25907–25912.
 44. Jaysinghe, S., Hristova, K., Wimley, W. C., Snider, C. & White, S. H. (2006). <http://blanco.biomol.uci.edu/mpex>.
 45. Sauri, A., McCormick, P. J., Johnson, A. E. & Mingarro, I. (2007). Sec61alpha and TRAM are sequentially adjacent to a nascent viral membrane protein during its ER integration. *J. Mol. Biol.* **366**, 366–374.
 46. von Heijne, G., Wickner, W. & Dalbey, R. E. (1988). The cytoplasmic domain of *Escherichia coli* leader peptidase is a translocation poison: sequence. *Proc. Natl Acad. Sci. USA*, **85**, 3363–3366.
 47. Rico, P. & Hernandez, C. (2004). Complete nucleotide sequence and genome organization of *Pelargonium* flower break virus. *Arch. Virol.* **149**, 641–651.
 48. Kida, Y., Sakaguchi, M., Fukuda, M., Mikoshiba, K. & Mihara, K. (2000). Membrane topogenesis of a type I signal-anchor protein, mouse synaptotagmin II, on the endoplasmic reticulum. *J. Cell Biol.* **150**, 719–730.
 49. Martinez-Gil, L., Sauri, A., Vilar, M., Pallas, V. & Mingarro, I. (2007). Membrane insertion and topology of the p7B movement protein of Melon Necrotic Spot Virus (MNSV). *Virology*, **367**, 348–357.
 50. Feilmeier, B. J., Iseminger, G., Schroeder, D., Webber, H. & Phillips, G. J. (2000). Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli*. *J. Bacteriol.* **182**, 4068–4076.
 51. Drew, D., Sjostrand, D., Nilsson, J., Urbig, T., Chin, C. N., de Gier, J. W. & von Heijne, G. (2002). Rapid topology mapping of *Escherichia coli* inner-membrane proteins by prediction and PhoA/GFP fusion analysis. *Proc. Natl Acad. Sci. USA*, **99**, 2690–2695.
 52. Facchini, P. J., Penzes, C., Johnson, A. G. & Bull, D. (1996). Molecular characterization of berberine bridge enzyme genes from opium poppy. *Plant Physiol.* **112**, 1669–1677.
 53. Abell, B. M., High, S. & Moloney, M. M. (2002). Membrane protein topology of oleosin is constrained by its long hydrophobic domain. *J. Biol. Chem.* **277**, 8602–8610.
 54. Whitley, P., Nilsson, I. M. & von Heijne, G. (1996). A nascent secretory protein may traverse the ribosome/ER translocase complex as an extended chain. *J. Biol. Chem.* **271**, 6241–6244.
 55. Borel, A. C. & Simon, S. M. (1996). Biogenesis of polytopic membrane proteins: membrane segments assemble within translocation channels prior to membrane integration. *Cell*, **85**, 379–389.
 56. Borel, A. C. & Simon, S. M. (1996). Biogenesis of polytopic membrane proteins: membrane segments of

- P-glycoprotein sequentially translocate to span the ER membrane. *Biochemistry*, **35**, 10587–10594.
57. Mingarro, I., Nilsson, I., Whitley, P. & von Heijne, G. (2000). Different conformations of nascent polypeptides during translocation across the ER membrane. *BMC Cell Biol.* **1**, 3.
 58. Rosch, K., Naeher, D., Laird, V., Goder, V. & Spiess, M. (2000). The topogenic contribution of uncharged amino acids on signal sequence orientation in the endoplasmic reticulum. *J. Biol. Chem.* **275**, 14916–14922.
 59. Ismail, N., Crawshaw, S. G., Cross, B. C., Haagsma, A. C. & High, S. (2008). Specific transmembrane segments are selectively delayed at the ER translocon during opsin biogenesis. *Biochem. J.* **411**, 495–506.
 60. van Klompenburg, W., Nilsson, I. M., von Heijne, G. & de Kruijff, B. (1997). Anionic phospholipids are determinants of membrane protein topology. *EMBO J.* **16**, 4261–4266.
 61. Bogdanov, M., Xie, J., Heacock, P. & Dowhan, W. (2008). To flip or not to flip: lipid–protein charge interactions are a determinant of final membrane protein topology. *J. Cell Biol.* **182**, 925–935.
 62. Saksena, S., Shao, Y., Braunagel, S. C., Summers, M. D. & Johnson, A. E. (2004). Cotranslational integration and initial sorting at the endoplasmic reticulum translocon of proteins destined for the inner nuclear membrane. *Proc. Natl Acad. Sci. USA*, **101**, 12537–12542.
 63. Orzaez, M., Perez-Paya, E. & Mingarro, I. (2000). Influence of the C-terminus of the glycophorin A transmembrane fragment on the dimerization process. *Protein Sci.* **9**, 1246–1253.
 64. Drew, D., Lerch, M., Kunji, E., Slotboom, D. J. & de Gier, J. W. (2006). Optimization of membrane protein overexpression and purification using GFP fusions. *Nat. Methods*, **3**, 303–313.