J. Membrane Biol. 192, 149–157 (2003)
DOI: 10.1007/s00232-002-1071-z

The Journal of
Membrane
Biology

Topical Review

Understanding the Biogenesis of Polytopic Integral Membrane Proteins

R.J. Turner
Membrane Biology Section, Gene Therapy and Therapeutics Branch, National Institute of Dental and Craniofacial Research, National Institutes of Health, DHHS, 10 Center Drive MSC 1190, Bethesda MD 20892, USA

Received: 17 September 2002

Introduction

Experiments carried out over the past decade have dramatically advanced our understanding of the biogenesis of integral membrane proteins. It is now clear that the membrane integration and assembly of most of these proteins occurs with the assistance of a large complex of translocation/insertion machinery. An appreciation of how integral membrane proteins are recognized and processed by this machinery provides considerable insight into their topology, folding and structure. The purpose of this article is to review our current understanding of these events with emphasis on recent studies. For more detailed information on many related topics that could only be mentioned or discussed briefly here, the reader is referred to a number of previous reviews (von Heijne, 1996, 1999; Hegde & Lingappa, 1997, 1999; Matlack, Mothes & Rapoport, 1998; van Geest & Lolkema, 2000; Goder & Spiess, 2001; Chin, von Heijne & de Gier, 2002).

The evidence available to date indicates that the membrane-spanning domains of integral membrane proteins are mainly composed of either transmembrane α helices or amphipathic β strands forming a closed β barrel. Most integral membrane proteins appear to be of the former type, that is, they consist of one or more α-helical membrane spanning segments (MSSs) connected by alternating intracellular and extracellular peptide loops. Proteins that contain a single MSS, and thus cross the membrane only once, are referred to as *bitopic* while those with multiple MSSs are called *polytopic*. The integration and folding of α-helical membrane proteins into the eukaryote endoplasmic reticulum (ER) and bacterial plasma membrane have been studied in considerable detail and it is this process that concerns us here. The folding of β-barrel membrane proteins, which have only been found in mitochondria, chloroplasts and the outer membranes of gram-negative bacteria, is less well characterized (Buchanan, 1999) and will not be discussed.

In what follows I begin with a simplified description of the way in which membrane proteins of the α-helical bundle type are integrated into the lipid bilayer. I then return to the various steps in this process and discuss observations that have led to the current refinements in our understanding.

For simplicity I will also refer mainly to experiments dealing with the integration of proteins into the membrane of the ER; a very similar process, involving closely evolutionarily related integration machinery, operates in the bacterial plasma membrane (von Heijne, 1996, 1999; Matlack et al., 1998; van Geest & Lolkema, 2000).

Cotranslational Integration, a Simplified Model of the Biogenesis of a Polytopic Membrane Protein

Most polytopic membrane proteins are thought to integrate into the membrane of the ER cotranslationally essentially as follows (Hegde & Lingappa, 1997; Matlack et al., 1998). As the first hydrophobic MSS of the protein emerges from the ribosome it is recognized by a cytosolic ribonucleoprotein complex referred to as a *signal recognition particle* (SRP). The SRP binds to the ribosome/nascent chain accomplishing two things: first, the translation of the nascent chain is stopped (or more correctly, paused) preventing the synthesis of the (insoluble) membrane protein in the cytoplasm, and second, the SRP-ribosome–nascent chain complex is targeted to the SRP receptor on the ER. Here the nascent chain and its ribosome are released from the SRP and transferred to a large membrane-embedded protein complex referred to as a *translocon* (Johnson

Correspondence to: R.J. Turner; email: rjturner@nih.gov

Key words: Translocon — Membrane protein folding — Transmembrane topology

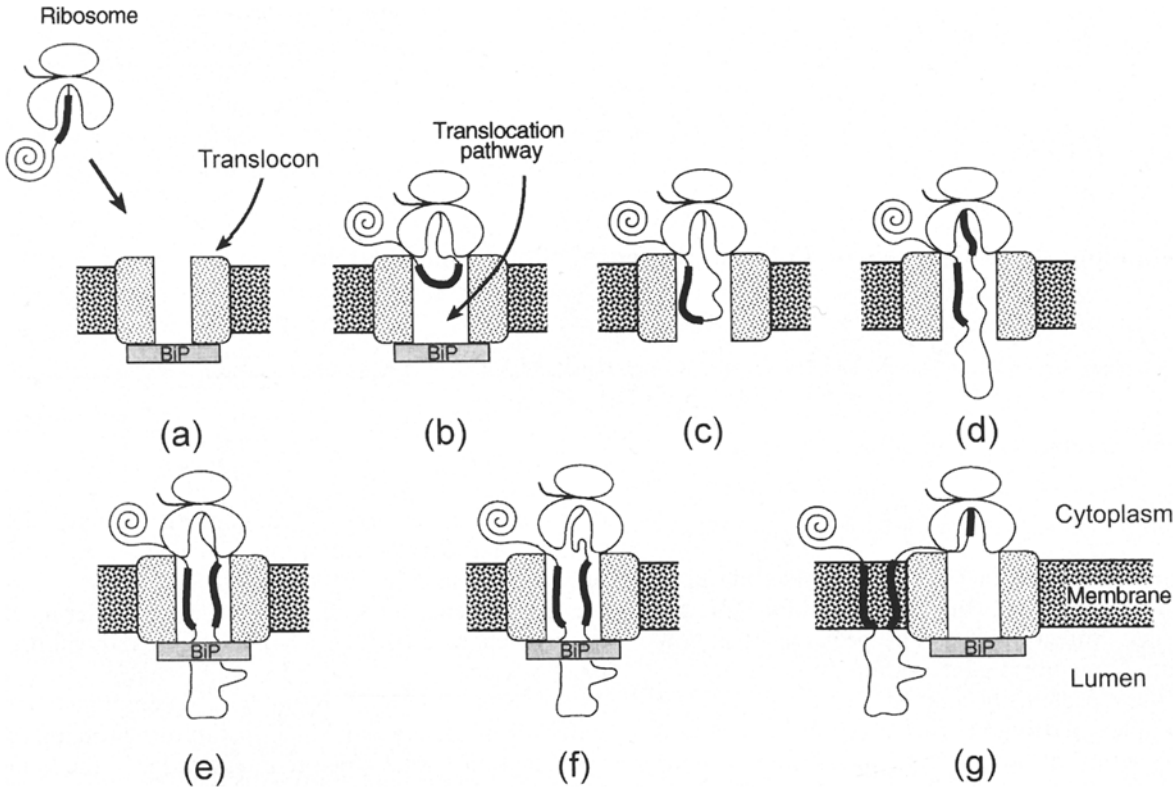

Fig. 1. A schematic representation of the interactions of the ribosome, translocon, and nascent polypeptide chain during the integration of MSSs into the lipid bilayer (*see* text for details), (*a*) The appearance of the first MSS in the nascent chain results in the targeting of the ribosome/nascent chain to the translocon. The empty translocon is sealed at its luminal side by BiP. (*b*) The ribosome forms a tight seal with the translocon and the MSS is fed into the translocon pore. The MSS could in principle adopt an N_{cyt}/C_{lum} or an N_{lum}/C_{cyt} orientation, but here it is assumed to be constrained to the former by the folding of the N terminus of the nascent chain. (*c–d*) The luminal end of the translocon pore opens, allowing the nascent chain to be extruded into the ER lumen while the first MSS is retained by the translocon. (*e*) A second MSS is recognized and retained by the translocon in an N_{lum}/C_{cyt} orientation. The luminal end of the translocon pore is closed by BiP. The amino acids in the nascent chain that follow the second MSS remain on the cytosolic side of the membrane. (*f–g*) The ribosome-translocon seal is broken while the BiP-dependent luminal seal remains intact, allowing the intracellular domains of the protein to enter the cytosol. The MSSs exit the translocon laterally into the lipid bilayer. (*g*) A tight ribosome-translocon seal reforms as a third MSS emerges from the ribosome and the process is repeated from (*b*).

& van Waes, 1999) that subsequently mediates the integration of the protein into the lipid bilayer of the ER (Fig. 1*a*). Translation of the nascent chain resumes after release from the SRP and the hydrophobic MSS comes into contact with and is recognized by the translocon, which opens to form an aqueous channel across the ER membrane (Fig. 1*b–c*; the significance of the protein labeled BiP in Fig. 1 will be explained later).

In Fig. 1 it is assumed that the N-terminus of the nascent chain has folded into a sufficiently large structure that cannot pass through the ribosome–translocon junction and therefore remains in the cytoplasm. As the length of the nascent chain increases, the first MSS inserts into and is retained by the translocon (Fig. 1*c–d*) with its N-terminus facing the cytoplasmic side of the ER and its C terminus toward the ER lumen (N_{cyt}/C_{lum}). An MSS in this configuration is referred to as a *type II signal anchor sequence* (*type II* refers to its N_{cyt}/C_{lum} orientation; if the first MSS instead inserted in a N_{lum}/C_{cyt} orientation, it would be a *type I signal anchor sequence*). The amino acids in the nascent chain that follow this signal-anchor sequence are then extruded through the translocon into the interior of the ER as they are synthesized (Fig. 1*d*). This translocation process is subsequently stopped by the appearance of a second MSS that acts as a *stop transfer sequence* by associating with, and being retained by, the translocon complex (Fig. 1*e–f*) in the opposite orientation to that of the preceding signal anchor.[1] The amino acids in the

[1]The distinction between a 'type I signal anchor sequence' and a 'stop transfer sequence' is that the former can insert into the membrane (with the aid of the translocon) independent of any preceding MSS while the latter acts only to stop the translocation of the nascent chain initiated by a preceding type II signal anchor. Thus a type I signal anchor sequence can act as a stop transfer sequence but a stop transfer sequence may not necessarily be able to function as a type I signal anchor.

nascent chain that follow this stop-transfer sequence will then remain on the cytosolic side of the membrane until the appearance of a new signal anchor sequence. Additionally, at appropriate times one or more of the MSSs residing in the translocon are transferred laterally into the ER membrane (Fig. 1*f–g*).

According to the above model of membrane-protein biogenesis, the final topology of a polytopic membrane protein is determined by the initial transmembrane orientation of the first MSS; once the first MSS is properly inserted into the membrane, the translocon recognizes each successive MSS and inserts it in the orientation opposite to the one before. It is now known that this simple model is a good first approximation to this process for many integral membrane proteins and that hydrophobicity, as one would have expected, is the main determinant for recognition of MSSs by the translocon. However, more recent studies have made it clear that additional refinements in this simple view are necessary. It has been shown, for example, that the insertion or deletion of a single MSS in a polytopic membrane protein does not always change the orientation of downstream MSSs (Bibi et al., 1991; McGovern, Ehrmann & Beckwith, 1991), and that some MSSs have a sufficiently strongly preferred orientation that they can force other hydrophobic segments out of the membrane (McGovern et al., 1991; Gafvelin & von Heijne, 1994). This is because hydrophobicity is not the only *topogenic signal* that can direct membrane-protein integration. These additional signals as well as recent experiments that clarify the operation of the translocon complex are discussed below.

It should also be noted that secreted proteins are transferred into the lumen of the ER via the translocon complex as well (Wilkinson, Regnacq & Stirling, 1997). These proteins are targeted to the ER by an N-terminal hydrophobic *cleavable signal sequence* typically consisting of 7–14 predominantly apolar amino acids (Nilsson, Whitley & von Heijne, 1994). This signal sequence is recognized by the translocon and integrated into the ER membrane in an N_{cyt}/C_{lum} orientation. The remainder of the protein is then threaded through the translocon into the lumen of the ER where signal peptidase cleaves off the signal sequence, releasing the soluble secreted protein.

The Translocon

The core of the eukaryotic translocon is the heterotrimeric Sec61p complex consisting of the integral membrane proteins Sec61α, Sec61β and Sec61γ (Johnson & van Waes, 1999). Structural studies (Hanein et al., 1996; Menetret et al., 2000; Beckmann et al., 2001), including a recent cryoelectron microscopic reconstruction of a ribosome–translocon–nascent chain complex at 15.4 Å resolution (Beckmann et al., 2001), indicate that three to four Sec61p heterotrimers form the translocation channel. These structures also show that the translocon is positioned such that its central pore aligns with the exit site for the nascent chain in the large ribosomal subunit, consistent with the general ideas presented above and in Fig. 1. Reconstitution experiments have shown that the Sec61p complex alone is sufficient for the processing of some integral membrane proteins while the addition of the translocating chain-associated membrane protein (TRAM) facilitates or is essential to the correct membrane integration of many others (Gorlich & Rapoport, 1993; Voigt et al., 1996). Additional, as yet unidentified elements may also be involved in the processing of some membrane proteins (Hegde, Voistz Lingappa, 1998b; Falk & Gilula, 1998; Ukaji et al., 2002; Dohke & Turner, 2002). A number of other proteins are known to interact with nascent translocating polypeptides and thus to be directly or indirectly associated with the translocon complex (Johnson & van Waes, 1999). These include proteins with known enzymatic properties such as oligosaccharyl transferase and signal peptidase as well as other ER luminal proteins that are thought to be molecular chaperones but whose roles are not well understood.

Three general models have been suggested to describe ribosome–translocon–nascent chain behavior during the integration of a polytopic membrane protein into the lipid bilayer. In the first (Blobel, 1980), the ribosome remains membrane-bound only while synthesizing an MSS or luminal domain (Fig. 1*b–e*) and releases completely from the translocon while synthesizing a cytosolic domain (in contrast to the depiction in Fig. 1*f*). In the second model (Borel & Simon, 1996; Do et al., 1996), the ribosome remains membrane-bound throughout protein synthesis and the MSSs accumulate within the translocon only to be released upon termination of translation. Finally, in the third model, the ribosome remains membrane-bound throughout synthesis but MSSs can exit laterally into the lipid before the termination of translation.

As anticipated in Fig. 1, recent results support the third model of translocon function. A number of studies now indicate the ribosome remains membrane-bound throughout protein synthesis (Do et al., 1996; Mothes et al., 1997; Potter, Seiser & Nicchitta, 2001; Potter & Nicchitta, 2002), arguing strongly against the first model. In fact, Potter and Nicchitta (2002) found that ribosomes synthesizing membrane proteins remain stably bound to the translocon even after the termination of translation and the release of the protein into the lipid bilayer. In addition, membrane protein synthesis could be reinitiated from these same membrane-bound ribosomes. Using

chemical crosslinking studies Heinrich et al. (2000) have studied the association of MSSs with the translocon, membrane lipids and TRAM during their integration into the bilayer. In contrast to the predictions of the second model of translocon function, their results indicated that MSSs could enter the lipid phase while still attached to the ribosome. They found that a strongly hydrophobic MSS entered the bilayer as soon as it extended far enough outside the ribosome to span the membrane. The same MSS containing one added positive charge also entered the lipid phase but less readily than the wild type, and an MSS with two added positive charges entered the lipid less readily again. On the basis of these results Heinrich et al. (2000) proposed the following general model of translocon function: The translocon acts to provide an environment in which hydrophobic MSSs can partition laterally into the lipid phase of the membrane, thus avoiding the energetic barrier presented by the polar head groups of the membrane phospholipids (the mechanism of this lateral gating remains unclear). A sufficiently hydrophobic MSS leaves the channel immediately and completely partitions into the lipid environment, while MSSs with more polar or charged residues favor the amphipathic interface between channel and lipid. In a polytopic membrane protein, two or more such MSS could assemble into a structure in which their hydrophilic residues are shielded from their surface before they can easily enter the lipid environment. Since Heinrich et al. also found that only less hydrophobic MSSs were crosslinked to TRAM, they further suggested that TRAM may be involved in the retention of these MSSs within the translocon until they can equilibrate into the lipid phase.

Another important and intriguing question concerning the biogenesis of integral membrane proteins is how the MSSs and their associated luminal and cytoplasmic loops are released from the translocon without compromising the permeability barrier of the ER membrane. Johnson and collaborators (Liao et al., 1997; Hamman, Hendershat & Johnson, 1998; Haigh & Johnson, 2002) have studied the opening and closing of the translocon pore during cotranslational protein integration by incorporating fluorescent probes into nascent transmembrane peptides and examining their accessability to quenching reagents applied from the cytosolic or luminal sides of isolated microsomes. Briefly stated, their results suggest the following scenario. Before ribosome binding, the translocon pore is closed at its luminal end by BiP (Fig. 1*a*), a luminal Hsp70. Upon binding, the ribosome forms a tight seal with the translocon so that the integrating nascent chain is inaccessible from the cytoplasm (Fig. 1*b*). Then the luminal end of the translocon pore opens and the nascent chain is fed through the pore so that the luminal domain of the protein enters the ER lumen (Fig. 1*c–d*). Next, the pore is resealed at the luminal side in a BiP-dependent manner (Fig. 1*e*); this appears to occur while the nascent chain is still inside the translocon pore, so this mechanism of sealing by BiP may differ from the way it closes ribosome-free translocons. The ribosome–translocon seal is then broken while the BiP-dependent luminal seal remains intact (Fig. 1*f*). This release of the ribosome–translocon seal presumably allows the intracellular domains of the protein to enter the cytosol (Fig. 1*f–g*). Surprisingly, the results of these authors also suggest that it is signals from the ribosome rather than from the translocon that are responsible for coordinating the luminal and cytosolic gating of the translocon channel (Liao et al., 1997). Although additional studies are required to clarify these (and many other) aspects of translocon function, it is clear that this mechanism of alternately opening the luminal and cytosolic ends of the translocon pore can account for the maintenance of the permeability barrier of the ER during membrane protein integration.

The above studies must also be reconciled with the recent cryoelectron microscopic reconstructions of the ribosome–translocon complex, which indicate the presence of significant gaps (~15 Å) in the region of contact between the ribosome and translocon (Menetret et al., 2000; Beckmann et al., 2001). It has been suggested (Menetret et al., 2000; Beckmann et al., 2001) that these gaps might allow the exodus of the intracellular domains of the nascent chain into the cytoplasm without requiring the breaking of the ribosome/translocon seal (*cf.*, Fig. 1*f*). However, it is not clear at this point whether these gaps could exist in native membranes without compromising the permeability barrier of the ER, whether they may be artifacts of the preparation procedure for cryoelectron microscopy, or whether they may actually represent the broken ribosome–translocon seal documented by Johnson and collaborators (Liao et al., 1997; Hamman et al., 1998; Haigh and Johnson, 2002).

It is now known that the translocon can also operate in reverse, that is, that it can retro-translocate proteins from the membrane or the ER lumen into the cytoplasm (Tsai, Ye & Rapoport, 2002). This is the major pathway for removal of misfolded proteins from the ER. Many aspects of this process are still uncertain, however, most of these retro-translocated proteins are thought to be polyubiquitinated, extruded into the cytoplasm via the translocon, then destroyed by the proteosome (Tsai et al., 2002). This bidirectional protein movement is consistent with the proposal of Heinrich et al. (2000) presented above, that the translocon provides an environment in which MSSs can partition between the aqueous environment of the translocation pore and the lipid phase. As discussed in more detail below, this dynamic property of the translocon suggests that it may also be in-

volved in the retrieval and refolding of some integral membrane proteins.

Topogenic Signals, Integration *vs.* Frustration, and Topology "Rules"

As indicated earlier, although hydrophobicity is obviously of great importance, it is not the only topogenic signal encoded into the sequence of polytopic membrane proteins. A number of other factors that have been shown to affect the integration and/or orientation of MSSs are discussed below.

Upstream Folding

Upstream folding can be a strong determinant of integral membrane protein topology. Thus, for example, if the N-terminus of a membrane protein folds into a sufficiently large and stable structure so that it cannot be pulled through the ribosome-translocon complex, one would expect the first MSS to adopt an N_{cyt}/C_{lum} orientation (*cf.* the hypothetical protein illustrated in Fig. 1). Experiments have shown, in fact, that it is the structure rather than the length of the N-terminus that favors its retention in the cytoplasm (Denzer et al., 1995). There is also an obvious tendency for each successive MSS to be inserted into the bilayer in the opposite orientation to the one preceding it; but as already alluded to, there are exceptions to this rule, as will be discussed in more detail shortly.

Flanking Charges

The most thoroughly studied determinant of the orientation of a MSS is the distribution of charged amino acids in its flanking extramembrane loops. The 'positive inside rule' *viz.*, the observation that positively charged amino acids tend to be more prevalent in the cytoplasmic than the extra-cytoplasmic segments of integral membrane proteins, was first discovered by von Heijne (1986). In a study of bacterial inner membrane proteins he showed that positive residues were four times more abundant in cytosolic than in periplasmic loops (von Heijne, 1986, 1994). A similar but somewhat weaker correlation is found for eukaryotic membrane proteins (Sipos & von Heijne, 1993; Gafvelin et al., 1997). Subsequent experiments have confirmed that the presence of positive amino acids in the sequence preceding an MSS confers a tendency on it to adopt an N_{cyt}/C_{lum} orientation, while the presence of these charges in the sequence following the MSS tends to result in it adopting the opposite orientation (Gafvelin et al., 1997). It is thought that the topogenic effect of these charged amino acids is mainly due to their interactions with charged residues in the translocon complex and/or with anionic membrane lipids that interfere with their transmembrane translocation. In fact, van Klompenburg et al. (1997) have demonstrated that changing the anionic phospholipid content of *E. coli* can affect the transmembrane orientation of model membrane proteins. More specifically, they found that increasing the content of anionic phospholipid enhanced the efficacy of the positive-inside rule and *vice versa*, consistent with the idea that the effect of positive charge may arise at least in part from interactions with lipid. Related studies on eukaryotic cells have not been carried out. Experiments in *E. coli* have also shown that the tendency of positively charged extramembrane loops to remain cytosolic is enhanced by the membrane potential (Andersson & von Heijne, 1994). This observation may account for the somewhat weaker applicability of the positive-inside rule to eukaryotic membrane proteins since the potential difference across the ER is much smaller than across bacterial plasma membranes. Finally, it has been noted that the positive-inside rule seems to be less important for longer extramembrane loops (von Heijne, 1994), possibly because their secondary structure is a stronger topological determinant.

Several hypothetical examples of membrane proteins folding according to the positive-inside rule are illustrated schematically in Fig. 2. In Fig. 2*a* the presence of a positively charged region in the N-terminal tail of the protein is responsible for its retention in the cytosol, and positive charges in the loops following the second and fourth MSSs stabilize these in an N_{lum}/C_{cyt} orientation. Placing positive charges instead in the loops following MSSs 1 and 3 results in an inverted topology (Fig. 2*b*). The hypothetical protein shown in Fig. 2*c* illustrates an interesting consequence of the positive-inside rule. Here the protein shown in Fig. 2*a* has been modified by moving the positive charges following MSS 4 to the loop following MSS 3. The presence of conflicting topological signals on either side of MSS 3 results in its being excluded from the membrane. Gafvelin and von Heijne (1994) have referred to this phenomenon as "topological frustration". A complementary effect is also possible; downstream positive charges can enhance the stop transfer activity of an otherwise weakly hydrophobic MSS and constrain it to the membrane (Dohke and Turner, 2002).

Hermansson, Monne and von Heijne (2001) have recently demonstrated that downstream charges can also influence the formation of so-called *helical hairpins*, two closely spaced MSSs separated by a short turn. These authors studied the membrane insertion of a 40-residue hydrophobic sequence that normally integrated into isolated microsomes as a single MSS in an N_{cyt}/C_{lum} orientation. Adding clusters of positively charged amino acids downstream of this sequence caused it to adopt an N_{cyt}/C_{cyt} hairpin conformation. Somewhat surprisingly, downstream

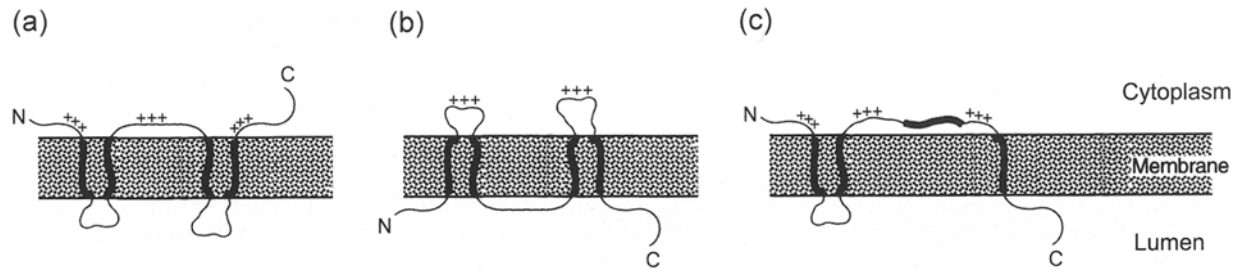

Fig. 2. Schematic representations of the effects of the positive-inside rule. *See* text for details.

clusters of negatively charged amino acids had a similar effect. Downstream charges also enhanced the effects of amino acids that had previously been shown to promote hairpin formation when placed in the middle of the hydrophobic sequence (Monne, Hermansson & von Heijne, 1999a; Monne et al., 1999b).

Intrinsic Orientation, Intramolecular Interactions and Other Topological Determinants

Experiments with model MSSs indicate that long hydrophobic sequences tend to insert in an N_{lum}/C_{cyt} orientation (Wahlberg & Spiess, 1997; Harley et al., 1998; Rosch et al., 2000), although it was found that preceding the MSS with a long hydrophilic sequence could overcome this effect, as could appropriate flanking charged amino acids. Ota et al. (1998b) have studied a particularly dramatic example of preferred MSS orientation. These authors were able to demonstrate that a naturally occurring MSS from the Cl^-/HCO_3^- exchanger, band 3, had such a strong tendency to assume an N_{lum}/C_{cyt} orientation that it could pull a preceding hydrophilic sequence into the membrane to form an N_{cyt}/C_{lum} MSS. There are also a number of other examples in the literature providing evidence that the integration and/or final transmembrane orientation of certain MSSs requires the presence of neighboring or more distantly downstream MSSs (Skach & Lingappa, 1993; Lin & Addison, 1995; Guo et al., 1996; Wilkinson, Critchley & Stirling, 1996; Lu et al., 1998; Ota et al., 1998a; Nilsson et al., 2000; Dohke & Turner, 2002; Sato et al., 2002). Some of these effects appear to be due to the strongly preferred orientation of a downstream MSS while others apparently arise from direct interactions between MSSs. In many cases the effect of a neighboring MSS is to cause the incorporation of a weakly hydrophobic MSS into the membrane.

Goder, Bieri and Spiess (1999) studied the behavior of a model protein containing two MSSs that was synthesized in two topological forms, one with the N- and C-termini in the ER lumen and the other with the termini in the cytosolic compartment. They found that when they engineered a glycosylation site into the extramembrane loop between the two MSSs, the proportion of proteins with cytosolic termini dramatically increased. These experiments suggest that glycosylation of this extramembrane loop in the lumen of the ER results in its being trapped there, apparently because the attached sugar chain is excluded from the translocon. Thus, glycosylation can apparently also act as topological determinant. Even more significantly, however, these experiments are consistent with the idea that membrane proteins can dynamically reorient within the translocon so that they can explore all of their possible topological conformations before a final topology is decided and they are released into the lipid phase.

Bogdanov et al. (Bogdanov, Heacock & Dowhan, 2002) have recently demonstrated a topological role for membrane lipids in addition to their possible involvement in the positive-inside rule (*see* above). In *E. coli* mutants lacking phosphatidylethanolamine these authors found that the N-terminal half of the lactose transporter LacY adopts an inverted topology and is incapable of active transport. In these mutants the first 6 MSSs of LacY are inserted in the orientation opposite to that found in wild-type *E. coli* and the 7th MSS appears to be left out of the membrane ("frustrated"). Why this occurs is still unclear, however, it does not seem to be due to a general effect on the translocon because the cells are viable and other membrane proteins are inserted correctly. The mutants also have a normal membrane potential and anionic lipid content. Interestingly, the induction of phosphatidylethanolamine synthesis in these bacteria resulted in conversion of LacY to its wild-type topology and restoration of active transport. This result indicates once again the apparent dynamic character of membrane protein topology and suggests that the translocon may be involved in post-assembly proofreading and correction of misfolded structures.

The existence of so-called *stop transfer effectors* have also been documented in studies of the prion protein and murine IgM (Yost et al., 1990; Falcone et al., 1999), but their properties are as yet poorly understood. These are short sequences that precede a relatively non-hydrophobic MSS and facilitate its integration into the membrane as a stop transfer sequence, possibly by interacting with receptors on the ER.

Membrane Protein Folding and Misfolding

It is clear from the above discussion that the final topology of a polytopic membrane protein is the net result of the influences of a variety of topogenic signals embedded in the protein sequence and interpreted via the translocon complex. Owing to the nature of this process it also seems clear that there is no a priori guarantee that a membrane protein will fold into its lowest energy state. Rather, it may fold into an energetically and/or conformationally trapped state whose characteristics can only be understood with reference to the way in which membrane proteins interact with and are handled by the translocon complex. These aspects of membrane protein folding could explain why many of these proteins have proven difficult to refold after denaturation.

Perhaps the most surprising result that emerges from our present understanding of this process is that hydrophobicity is not an absolute characteristic of an MSS. As indicated above, under some conditions a hydrophobic region can actually be excluded from the membrane, while under others a relatively hydrophilic region can form an MSS. Such hydrophilic transmembrane segments may have considerable freedom to move relative to the membrane and accordingly may play important roles in protein function.

As also indicated above, a number of recent observations indicate a dynamic relationship between the translocon and the nascent polypeptide chain. Thus, while one or more MSSs are within the translocon, they appear to have considerable freedom to sample their conformational space. In addition, it seems that groups of MSSs or even previously released complete proteins can return to the translocon and possibly be refolded. This latter idea is in keeping with the recent evidence that the translocon is involved in the retro-translocation and removal of membrane proteins from the ER (*see* above). Membrane protein folding is thus a complex process involving synthesis, interactions of MSSs and their flanking sequences with the translocon, packing and possibly repacking of MSSs within the translocon, folding of extramembrane domains, and perhaps also some repacking of MSSs as they move from the aqueous environment of the translocon to the lipid bilayer.

The interplay and competition among topogenic signals can be complex. However, considerable progress has been made in the prediction of membrane protein topology, particularly with the use of machine learning approaches (Moller, Croning & Apweiler, 2001). To further complicate matters, a number of naturally occurring proteins with topological heterogeneity have been observed. One of these is ductin, a protein with 4 MSSs that can act as a component of a connexon channel as well as a subunit of a V-type ATPase. These two functions are carried out by identical molecules with opposite membrane topologies (Dunlop, Jones & Finbow, 1995). In addition to its secreted form found in normal brain, the prion protein PrP can also be synthesized in two oppositely oriented transmembrane forms (Hegde et al., 1998a). One of these triggers spontaneous neurodegeneration when overexpressed (Hegde et al., 1998a) and appears to be induced in infectious prion disease (Hegde et al., 1999). The cystic fibrosis transmembrane regulator (CFTR) and the multidrug-resistance protein (MDR or P-glycoprotein) have also been shown to exist in mixed topologies where particular hydrophobic segments are left out of the membrane in some molecules (Han & Zhang, 1998; Zhang et al., 1998; Chen & Zhang, 1999). But at the present time it is not clear whether these "frustrated" proteins are physiologically relevant or simply misfolded. In this regard it has been suggested that CFTR may be prone to misfolding and that this might account for the disease-associated effects of some apparently conservative mutations (Sanders & Nagy, 2000).

Concluding Remarks

Although many questions and details remain to be resolved, we now appear to have a good basic understanding of the way in which α-helical integral membrane proteins are produced and assembled by the cell. Considerable progress has also been made in the understanding of many other aspects of the biochemistry and biophysics of these proteins: Molecular methods have made it possible to identify and characterize the important functional regions of a number of integral membrane proteins, most of which were resistant to analysis using classical biochemical approaches. In vitro studies of membrane protein folding have led to a better understanding of the forces that drive this process and the roles of lipids and detergents (Booth et al., 2001). Recent studies have also investigated possible amino-acid motifs involved in helix-helix interactions in a lipid environment (Fleming & Engelman, 2001; Gratkowski et al., 2001; Senes, Ubarretxena-Belandia & Engelmen, 2001; Zhou et al., 2001). The number of available crystal structures of membrane proteins, although still small, is gradually increasing and methods are being developed to improve membrane protein purification, yield and stability. Computational approaches to the study of membrane protein structure and function are still in their infancy but considerable progress in this area is also anticipated over the next few years.

I thank Drs. Bruce Baum and Youngsuk Oh for critical readings of the manuscript and the members of the laboratory for numerous discussions and insights.

References

Andersson, H., von Heijne, G. 1994. Membrane protein topology: effects of delta mu H+ on the translocation of charged residues explain the 'positive inside' rule. *EMBO J.* **13**:2267–2272

Beckmann, R., Spahn, C.M., Eswar, N., Helmers, J., Penczek, P.A., Sali, A., Frank, J., Blobel, G. 2001. Architecture of the protein-conducting channel associated with the translating 80S ribosome. *Cell* **107**:361–372

Bibi, E., Verner, G., Chang, C.Y., Kaback, H.R. 1991. Organization and stability of a polytopic membrane protein: deletion analysis of the lactose permease of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **88**:7271–7275

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

Bogdanov, M., Heacock, P.N., Dowhan, W. 2002. A polytopic membrane protein displays a reversible topology dependent on membrane lipid composition. *EMBO J.* **21**:2107–2116

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

Booth, P.J., Templer, R.H., Meijberg, W., Alien, S.J., Curran, A.R., Lorch, M. 2001. In vitro studies of membrane protein folding. *Crit. Rev. Biochem. Mol. Biol.* **36**:501–603

Buchanan, S.K. 1999. Beta-barrel proteins from bacterial outer membranes: structure, function and refolding. *Curr. Opin. Struct. Biol.* **9**:455–461

Chen, M., Zhang, J.T. 1999. Topogenesis of cystic fibrosis transmembrane conductance regulator (CFTR): regulation by the amino terminal transmembrane sequences. *Biochemistry* **38**: 5471–5477

Chin, C.N., von Heijne, G., de Gier, J.W. 2002. Membrane proteins: shaping up. *Trends Biochem. Sci.* **27**:231–234

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

Do, H., Falcone, D., Lin, J., Andrews, D.W., Johnson, A.E. 1996. The cotranslational integration of membrane proteins into the phospholipid bilayer is a multistep process. *Cell* **85**:369–378

Dohke, Y., Turner, R.J. 2002. Evidence that the transmembrane biogenesis of aquaporin 1 is cotranslational in intact mammalian cells. *J. Biol. Chem.* **277**:15215–15219

Dunlop, J., Jones, P.C., Finbow, M.E. 1995. Membrane insertion and assembly of ductin: a polytopic channel with dual orientations. *EMBO J.* **14**:3609–3616

Falcone, D., Do, H., Johnson, A.E., Andrews, D.W. 1999. Negatively charged residues in the IgM stop-transfer effector sequence regulate transmembrane polypeptide integration. *J. Biol. Chem.* **274**:33661–33670

Falk, M.M., Gilula, N.B. 1998. Connexin membrane protein biosynthesis is influenced by polypeptide positioning within the translocon and signal peptidase access. *J. Biol. Chem.* **273**:7856–7864

Fleming, K.G., Engelman, D.M. 2001. Specificity in transmembrane helix-helix interactions can define a hierarchy of stability for sequence variants. *Proc. Natl. Acad. Sci. USA* **98**:14340–14344

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

Gafvelin, G., von Heijne, G. 1994. Topological "frustration" in multispanning *E. coli* inner membrane proteins. *Cell* **77**:401–412

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

Goder, V., Spiess, M. 2001. Topogenesis of membrane proteins: determinants and dynamics. *FEES Lett.* **504**:87–93

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

Gratkowski, H., Lear, J.D., DeGrado, W.F, 2001. Polar side chains drive the association of model transmembrane peptides. *Proc. Natl. Acad. Sci. USA* **98**:880–885

Guo, D., Liu, J., Motlagh, A., Jewell, J., Miller, K.W. 1996. Efficient insertion of odd-numbered transmembrane segments of the tetracycline resistance protein requires even-numbered segments. *J. Biol. Chem.* **271**:30829–30834

Haigh, N.G., Johnson, A.E. 2002. A new role for BiP: closing the aqueous translocon pore during protein integration into the ER membrane. *J. Cell. Biol.* **156**:261–270

Hamman, B.D., Hendershot, L.M., Johnson, A.E. 1998. BiP maintains the permeability barrier of the ER membrane by sealing the lumenal end of the translocon pore before and early in translocation. *Cell* **92**:747–758

Han, E.S., Zhang, J.T. 1998. Mechanism involved in generating the carboxyl-terminal half topology of P-glycoprotein. *Biochemistry* **37**:11996–12004

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

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

Hegde, R.S., Lingappa, V.R. 1997. Membrane protein biogenesis: regulated complexity at the endoplasmic reticulum. *Cell* **91**:575–582

Hegde, R.S, Lingappa, V.R. 1999. Regulation of protein biogenesis at the endoplasmic reticulum membrane. *Trends Cell. Biol.* **9**:132–137

Hegde, R.S., Mastrianni, J.A., Scott, M.R., DeFea, K.A., Tremblay, P., Torchia, M., DeArmond, S.J., Prusiner, S.B., Lingappa, V.R. 1998a. A transmembrane form of the prion protein in neurodegenerative disease. *Science* **279**:827–834

Hegde, R.S., Tremblay, P., Groth, D., DeArmond, S.J., Prusiner, S.B., Lingappa, V.R. 1999. Transmissible and genetic prion diseases share a common pathway of neurodegeneration. *Nature* **402**:822–826

Hegde, R.S., Voigt, S., Lingappa, V.R. 1998b. Regulation of protein topology by trans-acting factors at the endoplasmic reticulum. *Mol. Cell* **2**:85–91

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

Hermansson, M., Monne, M., von Heijne, G. 2001. Formation of helical hairpins during membrane protein integration into endoplasmic reticulum membrane. Role of N and C terminal flanking regions. *J. Mol. Biol.* **313**:1171–1179

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

Liao, S., Lin, J., Do, H., Johnson, A.E. 1997. Both lumenal and cytosolic gating of the aqueous ER translocon pore are regulated from inside the ribosome during membrane protein integration. *Cell* **90**:31–41

Lin, J., Addison, R. 1995. A novel integration signal that is composed of two transmembrane segments is required to integrate

the Neurospora plasma membrane H^+-ATPase into microsomes. *J. Biol. Chem.* **270:**6935–6941

Lu, Y., Xiong, X., Helm, A., Kimani, K., Bragin, A., Skach, W.R. 1998. Co- and posttranslational translocation mechanisms direct cystic fibrosis transmembrane conductance regulator N terminus transmembrane assembly. *J. Biol. Chem.* **273:**568–576

Matlack, K.E.S., Mothes, W., Rapoport, T.A. 1998. Protein translocation: tunnel vision. *Cell* **92:**381–390

McGovern, K., Ehrmann, M., Beckwith, J. 1991. Decoding signals for membrane protein assembly using alkaline phosphatase fusions. *EMBO J.* **10:**2773–2782

Menetret, J.F., Neuhof, A., Morgan, D.G., Plath, K., Radermacher, M., Rapoport, T.A., Akey, C.W. 2000. The structure of ribosome-channel complexes engaged in protein translocation. *Mol. Cell* **6:**1219–1232

Moller, S., Croning, M.D., Apweiler, R. 2001. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* **17:**646–653

Monne, M., Hermansson, M., von Heijne, G. 1999a. A turn propensity scale for transmembrane helices. *J. Mol. Biol.* **288:**141–145

Monne, M., Nilsson, I., Elofsson, A., von Heijne G. 1999b. Turns in transmembrane helices: determination of the minimal length of a "helical hairpin" and derivation of a fine-grained turn propensity scale. *J. Mol. Biol.* **293:**807–814

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

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

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

Ota, K., Sakaguchi, M., Hamasaki, N., Mihara, K. 1998a. Assessment of topogenic functions of anticipated transmembrane segments of human band 3. *J. Biol. Chem.* **273:**28286–28291

Ota, K., Sakaguchi, H., 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

Ota, K., Sakaguchi, M., von Heijne, G., Hamasaki, N., Mihara, K. 1998b. Forced transmembrane orientation of hydrophilic polypeptide segments in multispanning membrane proteins. *Mol. Cell* **2:**495–503

Potter, M.D., Nicchitta, C.V. 2002. Endoplasmic reticulum-bound ribosomes reside in stable association with the translocon following termination of protein synthesis. *J. Biol. Chem.* **277:**23314–23320

Potter, M.D., Seiser, R.M., Nicchitta, C.V. 2001. Ribosome exchange revisited; a mechanism for translation-coupled ribosome detachment from the ER membrane. *Trends Cell Biol.* **11:**112–115

Rosch, K., Naeher, D., Laird, V., Goder, V., Spiess, M. 2000. The topogenic contribution of uncharged ammo acids on signal sequence orientation in the endoplasmic reticulum. *J. Biol. Chem.* **275:**14916–14922

Sanders, C.R., Nagy, J.K. 2000. Misfolding of membrane proteins in health and disease: the lady or the tiger? *Curr. Opin. Struct. Biol.* **10:**438–442

Sato, Y., Sakaguchi, M., Goshima, S., Nakamura, T., Uozumi, N. 2002. Integration of Shaker-type K+ channel, KAT1, into the endoplasmic reticulum membrane: synergistic insertion of voltage-sensing segments, S3-S4, and independent insertion of pore-forming segments, S5-P-S6. *Proc. Natl. Acad. Sci. USA* **99:**60–65

Senes, A., Ubarretxena-Belandia, I., Engelman, D.M. 2001. The Calpha—H...O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc. Natl. Acad. Sci. USA* **98:**9056–9061

Sipos, L., von Heijne, G. 1993. Predicting the topology of eukaryotic membrane proteins. *Eur. J. Biochem.* **213:**1333–1340

Skach, W.R., Lingappa, V.R. 1993. Amino-terminal assembly of human P-glycoprotein at the endoplasmic reticulum is directed by cooperative actions of two internal sequences. *J. Biol. Chem.* **268:**23552–23561

Skach, W.R., Lingappa, V.R. 1994. Transmembrane orientation and topogenesis of the third and fourth membrane-spanning regions of human P-glycoprotein (MDR1). *Cancer Res.* **54:**3202–3209

Tsai, B., Ye, Y., Rapoport, T.A. 2002. Retro-translocation of proteins from the endoplasmic reticulum into the cytosol. *Nat. Rev. Mol. Cell Biol.* **3:**246–255

Ukaji, K., Ariyoshi, N., Sakaguchi, M., Hamasaki, N., Mihara, K. 2002. Membrane topogenesis of the three amino-terminal transmembrane segments of glucose-6-phosphatase on endoplasmic reticulum. *Biochem. Biophys. Res. Commun.* **292:**153–160

van Geest, M., Lolkema, J.S. 2000. Membrane topology and insertion of membrane proteins: search for topogenic signals. *Microbiol. Mol. Biol. Rev.* **64:**13–33

van Klompenburg, W., Nilsson, I., von Heijne, G., de Kruijff, B. 1997. Anionic phospholipids are determinants of membrane protein topology. *EMBO J.* **16:**4261–4266

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

von Heijne, G. 1986. The distribution of positively charged residues in bacterial inner membrane proteins correlates with the transmembrane topology. *EMBO J.* **5:**3021–3027

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

von Heijne, G. 1996. Principles of membrane protein assembly and structure. *Prog. Biophys. Mol. Biol.* **66:**113–139

von Heijne, G. 1999. Recent advances in the understanding of membrane protein assembly and structure. *Q. Rev. Biophys.* **32:**285–307

Wahlbereg, 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

Wilkinson, B.M., Critchley, A.J., Stirling, C.J. 1996. Determination of the transmembrane topology of yeast Sec61p, an essential component of the endoplasmic reticulum translocation complex. *J. Biol. Chem.* **271:**25590–25597

Wilkinson, B.M., Regnacq, M., Stirling, C.J. 1997. Protein translocation across the membrane of the endoplasmic reticulum. *J. Membrane Biol.* **155:**189–197

Yost, C.S., Lopez, C.D., Prusiner, S.B., Myers, R.M., Lingappa, V.R. 1990. Non-hydrophobic extracytoplasmic determinant of stop transfer in the prion protein. *Nature* **343:**669–672

Zhang, J.T., Chen, M., Han, E., Wang, C., 1998, Dissection of de novo membrane insertion activities of internal transmembrane segments of ATP-binding-cassette transporters: toward understanding topological rules for membrane assembly of polytopic membrane proteins. *Mol. Biol. Cell* **9:**853–863

Zhou, F.X., Merianos, H.J., Brunger, A.T., Engelman, D.M. 2001. Polar residues drive association of polyleucine transmembrane helices. *Proc. Natl. Acad. Sci. USA* **98:**2250–2255