nature structural & molecular biology

Cellular mechanisms of membrane protein folding

William R Skach

The membrane protein–folding problem can be articulated by two central questions. How is protein topology established by selective peptide transport to opposite sides of the cellular membrane? And how are transmembrane segments inserted, integrated and folded within the lipid bilayer? In eukaryotes, this process usually takes place in the endoplasmic reticulum, coincident with protein synthesis, and is facilitated by the translating ribosome and the Sec61 translocon complex (RTC). At its core, the RTC forms a dynamic pathway through which the elongating nascent polypeptide moves as it is delivered into the cytosolic, lumenal and lipid compartments. This Perspective will focus on emerging evidence that the RTC functions as a protein-folding machine that restricts conformational space by establishing transmembrane topology and yet provides a permissive environment that enables nascent transmembrane domains to efficiently progress down their folding energy landscape.

The process of polytopic (multispanning) membrane protein folding can be viewed as a series of sequential but potentially overlapping steps: (i) formation, orientation and integration of transmembrane helices in the lipid bilayer; (ii) helical packing within the membrane; (iii) localization and folding of cytosolic and extracytoplasmic domains; and, for many proteins, (iv) quaternary organization into functional oligomers. Structural intermediates that populate the folding pathway are therefore conceptually (although not physically) analogous to 'molten globules', where collapse of transmembrane domains during helical packing would presumably replace lipid contacts with more energetically favorable helix-helix interactions. However, the folding environment, physical forces, and energetics that give rise to membrane protein structure differ strikingly from the aqueous environment of globular protein folding[1,2]. In addition, membrane protein folding in eukaryotes is primarily compartmentalized in the endoplasmic reticulum (ER) and facilitated by a complex set of folding machinery[3,4]. Therefore, understanding how membrane proteins fold in cells requires detailed knowledge of how cellular factors affect the folding landscape within the hydrophobic environment of the lipid bilayer.

Cellular basis of membrane protein folding

Nearly three decades ago, Günter Blobel proposed that polytopic proteins might be generated by internally encoded topogenic determinants that interact with cytosolic and membrane-bound receptors and a multisubunit protein 'translocator' to co-translationally insert the nascent polypeptide to the ER membrane[5]. Several aspects of this prescient prediction are particularly noteworthy. Experimental evidence for the underlying signal sequence hypothesis was just emerging, and few receptors had yet been identified. The existence of the putative ER protein translocator was entirely speculative and stimulated a heated debate that lasted for nearly two decades. And skeptics argued that the biological complexity of polytopic proteins was too great to be accommodated by a singular protein machine.

Over the ensuing years, biochemical, genetic and structural studies confirmed both the presence and identities of proteins involved in virtually all aspects of targeting to, translocation across and integration into the ER and other organellar membranes[3,4,6,7]. Applying these principles to polytopic proteins, however, has proven challenging, in part owing to technical constraints, but perhaps more importantly because of the paucity of solved membrane protein structures with which to test specific hypotheses. Improved methodologies for interrogating nascent polypeptides and expansion of the membrane protein structure database have provided new insights into thermodynamics of protein folding and the molecular structure of cellular machineries that guide the folding process. These topics have been the subject of several recent excellent reviews[7–10] and will therefore not be covered in detail here. Rather, this Perspective will focus primarily on our emerging understanding of how the ER translocation machinery interacts with, and is in turn controlled by, the nascent polypeptide as it facilitates specific topogenesis and folding events.

Membrane protein folding is an extension of translocation

From a mechanistic standpoint it is instructive to consider early steps of polytopic protein folding as an extension of protein secretion, many principles of which are relatively well established[3,6,10–12]. Protein secretion is initiated when an ER signal sequence emerges from the ribosome, binds the cytosolic signal recognition particle (SRP), and targets the ribosome–nascent chain complex (RNC) to the ER membrane[13]. GTP hydrolysis by SRP and its receptor releases the signal sequence and transfers the RNC to a large protein-conducting channel formed by the Sec61αβγ heterotrimer[14] and numerous associated proteins including translocation-associated membrane protein (TRAM)[15], the translocation-associated protein complex (TRAP)[16], oligosaccharyl transferase[17], signal peptidase complex and others[18]. Although the precise stoichiometry and structure of actively engaged Sec61 complexes remain unknown, for the purposes of this Perspective I will refer to the fully assembled and functional protein complex as the ER translocon[13,19] to distinguish it from the actual channel core, which is formed by one or more copies of Sec61αβγ[20–22].

William R Skach is in the Department of Biochemistry & Molecular Biology, Oregon Health & Sciences University, Portland, Oregon, USA. Correspondence should be addressed to W.R.S. (skachw@ohsu.edu).

Published online 3 June 2009; doi:10.1038/nsmb.1600

As the RNC docks onto the ER membrane, the exit tunnel of the large ribosomal subunit is aligned with the axial translocon pore[21,23], and the signal sequence engages a binding site within Sec61α. Biochemical and fluorescence-quenching studies have shown that this establishes a tight association between the ribosome and translocon[24,25] that shields the nascent chain from the cytosol[26]. The signal sequence also opens the lumenal gate of the translocon to create a continuous aqueous translocation pathway from the ribosome exit tunnel through the translocon pore and into the ER lumen[26,27]. Most secretory and transmembrane proteins move through this pathway coincident with peptide elongation, although the efficiency of translocation is dependent upon the nature of the signal sequence, the passenger domain and, potentially, the presence of regulatory translocon factors[28,29].

Bitopic membrane proteins require at least two additional steps that are mediated by so-called stop-transfer sequences: peptide movement into the ER lumen must be terminated; and the hydrophobic transmembrane segment must be transferred laterally from the proteinaceous environment of the translocon into the surrounding lipid bilayer, a step commonly referred to as membrane integration[30]. Again, analysis of fluorescent probes incorporated into nascent integration intermediates indicate that these events are orchestrated by dynamic regulation of the RTC. Synthesis and compaction of the transmembrane segment (TMS) within the ribosome exit tunnel[31] alter the translocation pathway by closing the lumenal gate of the translocon[32,33], and subsequently relaxing the ribosome-translocon junction. Thus, instead of continuing to move into the ER lumen, the elongating nascent polypeptide is redirected beneath the base of the ribosome and into the cytosol[33,34].

An important and currently unresolved issue in membrane protein folding is how and when the TMS is released from Sec61. One view supported by lipid cross-linking studies is that the TMS continuously samples the hydrophobic membrane environment and moves into the bilayer by passive thermodynamic partitioning through a lateral cleft in the Sec61α subunit[4,20,35,36]. Consistent with this, the ability of a TMS in bitopic proteins (and, to a lesser extent, polytopic proteins) to terminate translocation and adopt a membrane-spanning topology seems to be primarily determined by its overall hydrophobicity[37–39]. Alternatively, release of the TMS from the translocon can be delayed and mechanistically triggered at specific stages of synthesis during or at the end of translation. This is presumably accomplished by a conformational change in the RTC that provides access to bulk lipid[40], although the precise mechanism remains unknown. In support of this latter model, chemical and photo–cross-linking studies have shown that many and possibly most native TMSs seem to reside within the translocon for prolonged periods of time and progress through different proteinaceous environments before integration[30,40–44]. A remarkable finding is that some TMSs can be actively retained within the translocon via specific polar interactions even after peptidyl tRNA bond cleavage and released into the membrane in an ATP-dependent manner[30]. This last finding is surprising as there are no known eukaryotic translocon components that contain ATPase activity.

It should be also noted that, although the basic principles of translocation are well established, many details are not universally accepted. For example, the cytosolic inaccessibility of nascent chains in functionally engaged translocons as determined by fluorescence collisional quenching has been challenged by cryo-EM studies of detergent-solubilized RTC complexes lacking substrate. Results of the latter show a constitutive 12–17-Å gap between the ribosome base and the translocon, which would be sufficient to expose the translocating nascent chain to the cytosolic environment[4,21,22]. Studies of functional translocons have also implicated a large translocon pore that is gated at its lumenal end by the action of BiP[45], whereas the crystal structure of an archaebacterial Sec61 homolog suggests that gating of a small pore is accomplished by displacement of a short helical plug[20]. Although further studies are clearly needed to resolve these and other issues, this Perspective will focus primarily on data obtained using functionally intact translocons that contain a translocating substrate, and I will then discuss potential ways in which these results might affect consideration of alternative structural models.

The co-translational topogenesis model

How then does the RTC direct polytopic protein biogenesis when multiple TMSs are presented in rapid succession? The simplest biogenesis mechanism would involve a series of iterative co-translational translation initiation and termination events similar to those used by secretory and bitopic proteins (**Fig. 1a**). Such a model predicts that the first TMS of a polytopic protein would function as a signal (or signal anchor) sequence to target the RNC to the membrane, gate open the translocon and initiate translocation into the ER lumen. The second TMS would function as a stop-transfer sequence to terminate translocation, close the lumenal gate, relax the ribosome junction and direct the downstream peptide loop into the cytosol. The third TMS would then re-establish the ribosome translocon junction, reopen the translocon pore, reinitiate translocation and so forth. Alternate gating of the translocon channel and ribosome junction by sequential topogenic determinants would therefore direct peptide loops into their proper cellular compartment and establish topology of each TMS as the nascent chain emerged from the ribosome[11,12].

The co-translational model imposes three major requirements. First, the direction of nascent chain movement, and hence transmembrane topology, is determined by the functional state of the translocation pathway. For example, when the ribosome is tightly bound and the translocon gate is open, the nascent peptide loop can translocate only into the ER lumen. If, however, the translocon is closed and the ribosome junction is relaxed, then the growing polypeptide must move into the cytosol. Second, the nature of the translocation pathway is strictly controlled by topogenic information encoded within the nascent polypeptide. To achieve the correct topology, each TMS must be recognized by the RTC as either a signal (anchor) sequence or a stop-transfer sequence and reset the translocation pathway for proper delivery of the next hydrophilic peptide loop. Third, the translocation pathway must be highly dynamic and precisely coordinated to change the direction of peptide movement every few seconds as TMSs are synthesized in rapid succession.

Although substantial effort has been directed at testing whether the co-translational model accurately describes native polytopic protein biogenesis, few studies have directly measured RTC gating. Instead, functional properties of topogenic determinants have been inferred by their ability to direct translocation and integration of heterologous reporter domains. A common method involves analysis of truncated fusion proteins in which a passive translocation reporter is ligated C-terminally to each TMS to provide a simple readout for topology of sequential loops[46]. This approach identified signal anchor and stop-transfer activities in some native polytopic proteins, but many topogenic determinants were either arranged in unexpected patterns or showed unusual topogenic properties[47–52]. The C-terminal reporter approach therefore failed to confirm a unified biogenesis mechanism and often resulted in ambiguous and/or conflicting topologies. Although initially confusing from a topological standpoint, these

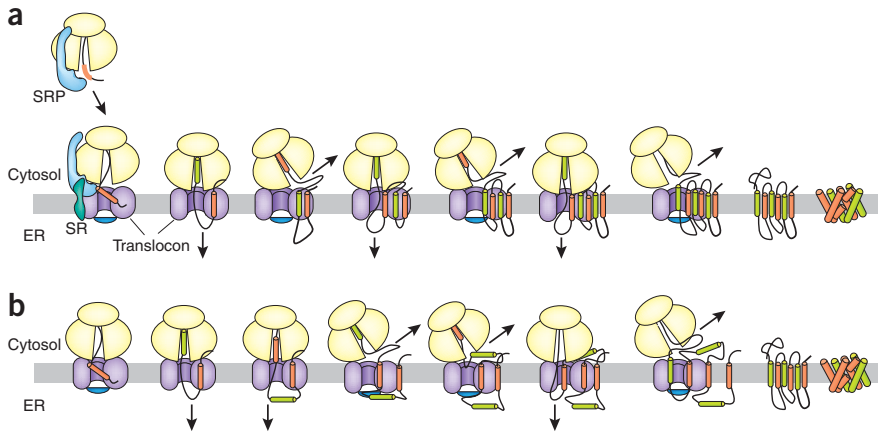

Figure 1 Models of polytopic protein biogenesis. (**a**) Co-translational biogenesis is initiated as the signal recognition particle (SRP) interacts with a signal sequence, binds its receptor (SR) at the ER membrane and transfers the ribosome nascent chain complex to the Sec61 the translocon. Signal sequences (orange cylinders) stimulate ribosome binding and open the gate of the translocon pore (blue disc) to initiate peptide movement into the ER lumen. Stop-transfer sequences (green cylinders) terminate translocation and redirect the elongating nascent chain beneath the ribosome and into the cytosol. Each sequential TMS therefore alters the direction of nascent chain movement through the RTC (arrow) to establish transmembrane topology from the N to the C terminus, one helix at a time. (**b**) During AQP1 biogenesis, TM2 fails to terminate translocation, TM3 is initially inserted into the translocon in a type I topology and TM4 transiently resides on the cytosolic face of the membrane. This sequence of events generates a four-spanning intermediate that is converted to a six-spanning topology during or after synthesis of TM5 and TM6.

results provided important insight into folding mechanisms that often went unappreciated. A well-studied example involved two homologous aquaporin water channels, AQP1 and AQP4, where minor differences in primary sequence dramatically change the folding pathway but have no effect on the final folded structure[11].

AQP4 as a paradigm for co-translational biogenesis

Aquaporins constitute a large family of proteins that facilitate passive water and glycerol transport across biological membranes[53–55]. High-resolution structures have demonstrated a conserved topological fold in which six membrane-spanning helices and two half helices are arranged in the membrane in an inverted two-fold pseudo symmetry[56,57]. Although each monomer contains an intact pore, mammalian aquaporins are believed to function exclusively as homotetramers that fold and assemble in the ER before transport to the cell surface[11,58–60]. Their small size, known structure and high degree of homology make aquaporins ideal candidates to investigate membrane protein–folding mechanisms.

Topological analysis of truncated AQP4 fusion proteins confirmed the presence of alternating signal (TM1, TM3 and TM5) and stop-transfer (TM2, TM4 and TM6) sequences and provided some of the first evidence that an entire native polytopic protein could be generated via a co-translational topogenesis mechanism. Specifically, each AQP4 TMS was found to act independently to establish the expected six-spanning topology from the N to the C terminus, one transmembrane helix at a time[61] (**Fig. 1a**). Photo–cross-linking studies using a series of truncated integration intermediates, which provide static snapshots of the nascent chain environment, also demonstrated that each TMS enters a binding site within Sec61 as it exits the ribosome, remains in this location until entry of the next TMS and then progresses through the translocon in a unique and highly ordered manner[41]. Some TMSs (that is, TM2 and TM4) show a brief, well-defined interval of cross-linking, whereas others (TM1 and TM3) show several phases of cross-linking at different stages of synthesis. As a result of this behavior, multiple TMSs were found to accumulate within or in close proximity to Sec61α before their release into the lipid bilayer.

AQP1, an exception to the co-translational rule

Surprisingly, AQP1-TMSs control the translocation pathway in a very different manner from their AQP4 counterparts[50] (**Fig. 1b**). For example, AQP1-TM2 does not efficiently terminate translocation and therefore transiently passes through the translocon into the ER lumen as it exits the ribosome. Because TM2 does not close the translocation pathway, TM3 encounters an open translocon, where it terminates translocation and directs its C-terminal flanking residues into the cytosol. This behavior results in a mixture of topologies in which most nascent AQP1 polypeptides co-translationally span the membrane only four times[11,50]. However, the initial four-spanning structure actually represents a folding intermediate that is converted into a mature six-spanning topology during and/or after the completion of synthesis[62]. AQP1 maturation therefore involves a 180° rotation of TM3 from a type I to a type II topology, which transfers the TM2-3 peptide loop from the ER lumen to the cytosol, the TM3-4 loop from the cytosol to the ER lumen, and inserts TM2 and TM4 into the plane of the membrane. These events represent a striking exception to the co-translational model and raise two obvious questions. Why would two highly homologous proteins with similar structure and function use such different folding pathways? And how is AQP1 converted from a four-spanning to a six-spanning topology?

AQP1 folding illustrates the potential conflict between structural determinants that direct early topogenesis events within the RTC, and the role of the same residues during subsequent stages of protein maturation. A principal difference between AQP1 and AQP4 topogenesis is that AQP1-TM2 fails to terminate translocation as it exits the ribosome. This divergent behavior is caused by two residues in TM2, Asn49 and Lys51 in AQP1 versus Met48 and Leu50 in AQP4 (ref. 63). Exchanging these residues converts AQP1-TM2 to an efficient stop-transfer sequence but completely disrupts water channel function. Hydrophilic residues are therefore needed to generate a functional channel, but their presence forces AQP1 to deviate from a co-translational biogenesis mechanism. Such behavior seems to be a relatively common feature of eukaryotic polytopic proteins to which rules of co-translational topogenesis do not strictly apply[64,65].

What then is the advantage for polytopic proteins to utilize a non–co-translational folding pathway? For AQP1, Asn49 and Lys51 both interact with a partially buried aspartic acid residue, Asp185, near the C terminus of TM5. Asn49 forms an intramolecular hydrogen bond that is needed for proper helical packing of the monomer, whereas Lys51 forms a nonessential ionic bond that stabilizes the AQP1

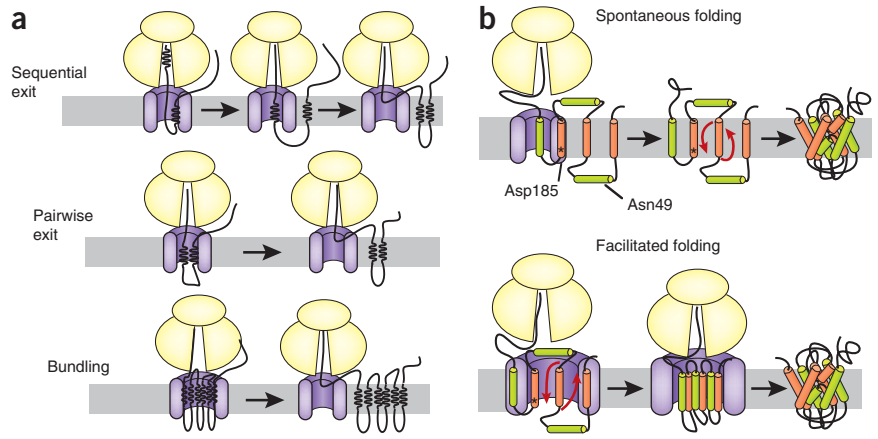

Figure 2 Mechanism of TM integration. (**a**) During membrane protein biogenesis, the TMS is transferred laterally from the proteinaceous environment of the translocon into the lipid bilayer. This may occur in a sequential fashion with each TMS integrating independently, in a pairwise fashion, or in groups. (**b**) The timing of TMS integration will determine, in part, whether helical packing takes place primarily in a lipid or proteinaceous environment. Two potential folding models are shown for AQP1. If TMSs are sequentially released from the translocon into bulk lipid (above), then topological maturation of TMs 2–4 will involve spontaneous TMS insertion, rotation and movement of two hydrophilic loops across the lipid bilayer. An intramolecular hydrogen bond between Asn49 in TM2 and Asp185 on TM5 is required for proper helical packing of the AQP1 monomer[60]. Alternatively, retention of the TMS within or in close proximity to translocon proteins (below) could potentially reduce energy barriers for peptide transfer and topological maturation.

tetramer[60]. Because proteins emerge from the ribosome vectorially from the N to the C terminus, TM2 must enter the translocon before TM5 is synthesized. At early stages of synthesis, the unpaired hydrophilic residues therefore prevent TM2 from terminating translocation. This could be caused by a reduction in hydrophobicity that interferes with the translocon interactions needed to terminate translocation[38], or, alternatively, a delay in helix formation within the ribosome exit tunnel, which would potentially be needed to relax the ribosome junction and allow TM2 to properly span the membrane[31,33]. Current studies are underway to distinguish these possibilities. Importantly, the same polar residues that dictate these early events of topogenesis are also needed to stabilize a polar residue (Asp185) in a distal C-terminal TMS. Failure of Asp185 to interact productively with Asn49 in TM2 leads to sequestration of AQP1 in a large complex or aggregate, a finding consistent with the strong energetic potential of aspartate residues to stimulate helix-helix interactions in hydrophobic environments[66,67]. AQP1 biogenesis therefore illustrates the complex ways in which subtle variations in primary sequence can alter multiple steps along the folding pathway: (i) through co-translational translocon interactions that direct early topogenesis; (ii) through tertiary folding and helical packing within the lipid bilayer; and (iii) by quaternary stabilization of AQP1 tetramers.

Functional implications of the AQP1 folding pathway

In contrast to initial predictions[5], delivery of the nascent chain into its proper cellular compartment is not necessarily a constitutive process. TMSs that lack strong topogenic properties may allow the nascent chain to enter both the cytosol and ER lumen as they exit the ribosome, an observation that has also been reported for other native and engineered substrates[51,68,69]. It is currently unknown whether the RTC provides access to both compartments simultaneously, or whether access is achieved stochastically by alternative translocon conformations. Such findings contrast with the view of a rigid, cytosolically inaccessible translocation pathway and suggest that protein movement through the RTC can show substantial variation depending on topogenic information present in the nascent polypeptide. These aspects of polytopic protein topogenesis are consistent with recent observations that the translocation efficiency of secretory proteins can also be regulated by specific properties of the signal sequence, the availability of translocon-associated proteins[29] and/or presence of small-molecule inhibitors[70]. Protein secretion and membrane protein folding may therefore share similar molecular mechanisms of translocon control. Further work is needed to understand the molecular basis for these observations, particularly the manner in which the ribosome junction and translocon channel gating are controlled during different states of translocation. An obvious advantage of this nonconstitutive translocation behavior is that it enables polytopic proteins to acquire variations in primary sequence (that is, polar residues) that would not be permitted by a strict co-translational topogenesis mechanism.

General implications for membrane protein folding

Another remarkable feature of AQP1 biogenesis is that there must be a mechanism in the ER for reorienting TMSs and peptide loops that are initially (co-translationally) directed into the wrong cellular compartment. The precise mechanism that drives AQP1 reorientation remains unknown, but the folded monomer is probably stabilized in part by alignment of TM2 and TM5 in the membrane and formation of the partially buried hydrogen bond between Asn49 and Asp185 in the relatively apolar core of the protein[60]. The nascent chain must therefore retain sufficient conformational flexibility during early stages of folding to allow for 'topological editing' while synthesis of downstream peptide regions provides additional folding information. This behavior raises a fundamental question as to the extent to which conformational flexibility is allowed, and where and how it is achieved in the context of the RTC and/or the ER membrane.

Movement of charged residues and/or hydrophilic peptide loops directly across the core of the lipid bilayer imposes a substantial energy barrier that would probably limit the transfer rate and hence slow the kinetics of AQP1 folding from a four- to a six-spanning topology[2]. The magnitude of such a barrier is difficult to estimate for polytopic proteins because transfer of a given peptide region would take place in an environment formed by both lipids and other TMSs within the protein. Studies of prokaryotic transporters indicate that peptide loops and even helical bundles may be spontaneously transferred across the membrane simply by changes in phospholipid composition[71]. Relatively large peptide domains in tail-anchored proteins have also been shown to translocate unassisted across the ER membrane[72]. Thus, it is possible that AQP1 topological maturation might take place after the nascent chain has been released into the bulk lipid of the ER (**Fig. 2**).

An alternative scenario is that adjacent translocon proteins might contribute to the folding environment and thereby facilitate peptide reorientation. This intriguing possibility is supported by numerous cross-linking and membrane-extraction studies demonstrating that

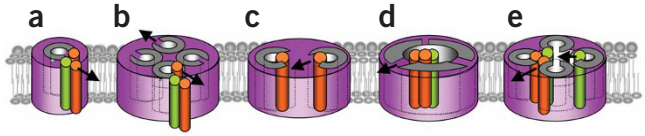

Figure 3 Potential arrangement of Sec61αβγ heterotrimers (gray cylinders) in the assembled translocon (purple disc) and implications for co-translational folding. Cryo-EM analyses of empty, solubilized mammalian ER RTCs suggest that Sec61 may be present beneath the ribosome in a single copy[22] (**a**) or in a back-to-back tetramer configuration[21,82] (**b**). Both models propose that only one Sec61 protein is used for translocation, providing transmembrane helices (orange and green cylinders) with a single lateral exit site to the translocon periphery. Alternative arrangements include a front-to-front Sec61 dimer configuration observed in cryo-EM structures of the *E. coli* SecYEG complex[83] (**c**), a large central pore derived from fluorescence-quenching experiments of functionally intact ER translocons[84] and supported by early low-resolution EM studies[75–77] (**d**), and a related but hypothetical oligomeric front-to-front configuration[11] in which the TMS could initially exit Sec61 into the translocon interior (**e**). These alternative models provide a potential means to accommodate multiple helices during translocation[85] and before nascent chain movement between subunits into the bilayer.

multiple TMSs can accumulate in close proximity to translocon proteins and be released into the bilayer in pairs or groups[42–44,49,73,74] (**Fig. 2a**). In the case of AQP1, the proximity of adjacent translocon proteins could potentially reduce the free-energy barrier imposed by TM3 reorientation and thereby provide the nascent polypeptide access to an increased conformational space needed to establish proper helical packing. Once formed, the mature six-spanning structure would presumably be stabilized by proper helix-helix contacts and by formation of the hydrogen bond between TM2 and TM5. The net outcome would be to improve the efficiency with which the nascent chain could progress down its folding energy landscape to a lower free-energy state than was achieved in the original, co-translational four-spanning topology. Currently, however, the precise mechanism of TMS reorientation, the temporal sequence of helical packing and the role of lipids and/or translocon or other ER proteins in carrying out these processes remain unknown for AQP1 or any other eukaryotic polytopic protein of similar complexity.

Cross-linking results suggest two possible mechanisms by which the translocon might facilitate helical packing. TMSs could potentially be released from the translocation channel but remain associated with accessory proteins at the translocon periphery, as has been postulated for TRAM in eukaryotes and the chaperone-like protein YidC in prokaryotes[4,36,40,43]. Indeed, TRAM association with a nascent TMS has been correlated with the presence of charged residues[36]. Depending on the architecture of individual subunits, it is also possible that a TMS might initially exit from the Sec61 channel to another location within the interior of the translocon that may contain either a large hole[23,75–77] or central depression that has been proposed to be filled with intercalated lipid[21,78]. Such an arrangement is appealing, because the translocon interior could shield the nascent chain from bulk lipids and yet provide a relatively hydrophobic environment that would permit transient hydrophobic as well as polar interactions. The nature of such interactions would depend on the primary sequence of the substrate and the available translocon proteins and lipids in the immediate vicinity, each of which could be potentially tailored for substrate folding.

Is the ER translocon a membrane protein chaperone?

Cellular chaperones are typically defined by their ability to assist the folding and assembly of proteins in a catalytic and nonconsumptive manner. The ER translocon was originally viewed as a channel with the capacity to facilitate translocation of secretory proteins from the cytosol to the ER lumen. It is now accepted that the translocon also establishes topology of bitopic and polytopic membrane proteins by recognizing specific sequence determinants and delivering peptide regions into their proper folding compartment. Because localization is crucial for folding, this criterion alone would fulfill the requirements for a membrane protein chaperone. Recent studies now raise the more profound possibility that, in addition to substrate localization, the translocon may provide a specialized environment that increases conformational flexibility needed for tertiary folding of helical transmembrane domains. If this is the case, then the chaperone functions of the translocon might be conceptually analogous to the folding chamber formed by the Hsp60 protein family, in which GroEL and the TRIC chaperonin sequester small globular proteins from the bulk cytosol as they acquire their proper tertiary structure[79]. Recent observations that the translocon can actively retain a TMS and release it in an ATP-dependent manner further resemble these chaperone functions and suggests that substrate release into the bilayer might also be an active and regulated process[30,80].

The extent to which eukaryotic membrane protein folding takes place in the bilayer, the translocon or both will obviously depend on the composition, stoichiometry and precise architecture of ribosome-bound, functional ER translocons (**Fig. 3**). Unfortunately, none of these parameters is currently known with precision, and hence any model of membrane protein folding is inherently speculative. However, the crystal structure of a Sec61 homolog from *Methanococcus janaschii* (SecYEβ) together with cryo-EM studies of solubilized ribosome-bound translocons has provided some important clues[20,21]. SecYEβ is a cuboidal structure roughly 40 Å in size that contains a constricted ~8-Å diameter central pore and a lateral opening through which protein translocation and integration, respectively, are predicted to occur. On the basis of this structure, several models have been proposed to explain how the eukaryotic homolog Sec61αβγ might facilitate co-translational membrane protein topogenesis[4,7,81]. As this topic has been covered extensively elsewhere[10], the goal here has been to provide a different viewpoint of the membrane protein–folding problem based on novel folding behaviors and the requirements of native substrates. Such an approach raises several questions that have been less well appreciated in structurally based models. For example, the predicted pore formed by Sec61αβγ is clearly too small to accommodate multiple TMSs, rotation of a single TMS, or topological maturation and folding of large protein domains. Thus, these events must either occur outside of the Sec61 channel in the bilayer or within the confines of a larger, fully assembled translocon. In this regard, Sec61 has been shown to form ring-like oligomers with a large central pore-like region[23,75,77]. However, higher-resolution studies have suggested alternative oligomeric arrangements[21,78] and, most recently, the possibility that a single Sec61αβγ heterotrimer might reside beneath the ribosome[22] (**Fig. 3**). Unfortunately, these structures lack translocon-associated subunits such as TRAM, oligosaccharyltransferase, signal peptidase complex and others. The specific arrangement of Sec61αβγ heterotrimers and/or additional associated proteins may therefore have crucial roles in TMS retention, rotation and topological maturation observed for AQP1 and potentially other native substrates. Solving this difficult problem will undoubtedly require complementary techniques that consolidate structural and functional information from fully assembled RTCs that are actively

engaged with substrate. A general model derived from such studies that adequately explains membrane protein folding at a molecular level will solve one of the long-standing and persistent questions in modern biology.

ACKNOWLEDGMENTS

I would like to thank B. Conti and other members of the Skach laboratory for helpful comments on this manuscript and the US National Institutes of Health (R01GM53457 and R01 DK51818) for supporting our research in this area.

Published online at http://www.nature.com/nsmb/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Popot, J.L. & Engelman, D. Helical membrane protein folding, stability and evolution. *Annu. Rev. Biochem.* **69**, 881–922 (2000).
2. Bowie, J.U. Solving the membrane protein folding problem. *Nature* **438**, 581–589 (2005).
3. Johnson, A.E. & van Waes, M. The translocon: a dynamic gateway at the ER membrane. *Annu. Rev. Cell Dev. Biol.* **15**, 799–842 (1999).
4. Rapoport, T.A., Goder, V., Heinrich, S. & Matlack, K. Membrane-protein integration and the role of the translocation channel. *Trends Cell Biol.* **14**, 568–575 (2004).
5. Blobel, G. Intracellular protein topogenesis. *Proc. Natl. Acad. Sci. USA* **77**, 1496–1500 (1980).
6. Schnell, D.J. & Hebert, D. Protein translocons: multifunctional mediators of protein translocation across membranes. *Cell* **112**, 491–505 (2003).
7. White, S.H. & von Heijne, G. The machinery of membrane protein assembly. *Curr. Opin. Struct. Biol.* **14**, 397–404 (2004).
8. Johnson, A. Fluorescence approaches for determining protein conformations, interactions and mechanisms at membranes. *Traffic* **6**, 1078–1092 (2005).
9. von Heijne, G. Membrane-protein topology. *Nat. Rev. Mol. Cell Biol.* **7**, 909–918 (2006).
10. Osborne, A.R., Rapoport, T.A. & van den Berg, B. Protein translocation by the Sec61/SecY channel. *Annu. Rev. Cell Dev. Biol.* **21**, 529–550 (2005).
11. Pitonzo, D. & Skach, W. Molecular mechanisms of aquaporin biogenesis by the endoplasmic reticulum Sec61 translocon. *Biochim. Biophys. Acta* **1758**, 976–988 (2006).
12. Sadlish, H. & Skach, W. Biogenesis of CFTR and other polytopic membrane proteins; new roles for the ribosome-translocon complex. *J. Membr. Biol.* **202**, 115–126 (2004).
13. Walter, P. & Lingappa, V.R. Mechanisms of protein translocation across the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.* **2**, 499–516 (1986).
14. Song, W., Raden, D., Mandon, E. & Gilmore, R. Role of Sec61α in the regulated transfer of the ribosome-nascent chain complex from the signal recognition particle to the translocation channel. *Cell* **100**, 333–343 (2000).
15. Görlich, D., Hartmann, E., Prehn, S. & Rapoport, T. A protein of the endoplasmic reticulum involved in early polypeptide translocation. *Nature* **357**, 47–52 (1992).
16. Wiedmann, M., Kurzchalia, T., Hartmann, E. & Rapoport, T. A signal sequence receptor in the endoplasmic reticulum membrane. *Nature* **328**, 830–833 (1987).
17. Kelleher, D.J. & Gilmore, R. An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* **16**, 47R–62R (2006).
18. Shibatani, T., David, L., McCormack, A., Frueh, K. & Skach, W. Proteomic analysis of mammalian oligosaccharyltransferase reveals multiple subcomplexes that contain Sec61, TRAP and two potential new subunits. *Biochemistry* **44**, 5982–5992 (2005).
19. Snapp, E.L., Reinhart, G., Bogert, B., Lippincott-Schwartz, J. & Hegde, R. The organization of engaged and quiescent translocons in the endoplasmic reticulum of mammalian cells. *J. Cell Biol.* **164**, 997–1007 (2004).
20. van den Berg, B. *et al.* X-ray structure of a protein-conducting channel. *Nature* **427**, 36–44 (2004).
21. Ménétret, J.-F. *et al.* Architecture of the ribosome-channel complex derived from native membranes. *J. Mol. Biol.* **348**, 445–457 (2005).
22. Ménétret, J.F. *et al.* Single copies of Sec61 and TRAP associate with a nontranslating mammalian ribosome. *Structure* **16**, 1126–1137 (2008).
23. Beckmann, R. *et al.* Architecture of the protein-conducting channel associated with the translating 80S ribosome. *Cell* **107**, 361–372 (2001).
24. Belin, D., Bost, S., Vassalli, J.-D. & Strub, K. A two step recognition of signal sequences determines the translocation efficiency of proteins. *EMBO J.* **15**, 468–478 (1996).
25. Jungnickel, B. & Rapoport, T. A posttargeting signal sequence recognition event in the endoplasmic reticulum membrane. *Cell* **82**, 261–270 (1995).
26. Crowley, K.S., Liao, S., Worrell, V., Reinhart, G. & Johnson, A. Secretory proteins move through the endoplasmic reticulum membrane via an aqueous, gated pore. *Cell* **78**, 461–471 (1994).
27. Crowley, K.S., Reinhart, G. & Johnson, A. The signal sequence moves through a ribosomal tunnel into a noncytoplasmic aqueous environment at the ER membrane early in translocation. *Cell* **73**, 1101–1115 (1993).
28. Kim, S.J., Mitra, D., Salerno, J. & Hegde, R. Signal sequences control gating of the protein translocation channel in a substrate-specific manner. *Dev. Cell* **2**, 207–217 (2002).
29. Hegde, R.S. & Kang, S.-W. The concept of translocational regulation. *J. Cell Biol.* **182**, 225–232 (2008).
30. Pitonzo, D., Yang, Z., Matsumura, Y., Johnson, A. & Skach, W. Sequence-specific retention and regulated integration of a nascent membrane protein by the ER Sec61 translocon. *Mol. Biol. Cell* **20**, 685–698 (2009).
31. Woolhead, C.A., McCormick, P. & Johnson, A. Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. *Cell* **116**, 725–736 (2004).
32. Alder, N.N., Shen, Y., Brodsky, J., Hendershot, L. & Johnson, A. The molecular mechanisms underlying BiP-mediated gating of the Sec61 translocon of the endoplasmic reticulum. *J. Cell Biol.* **168**, 389–399 (2005).
33. Haigh, N.G. & Johnson, A. A new role for BiP: closing the aqueous translocon pore during protein integration into the ER membrane. *J. Cell Biol.* **156**, 261–270 (2002).
34. Liao, S., Lin, J., Do, H. & Johnson, A. Both lumenal and cytosolic gating of the aqueous translocon pore are regulated from inside the ribosome during membrane protein integration. *Cell* **90**, 31–41 (1997).
35. Martoglio, B., Hofmann, M., Brunner, J. & Dobberstein, B. The protein-conducting channel in the membrane of the endoplasmic reticulum is open laterally toward the lipid bilayer. *Cell* **81**, 207–214 (1995).
36. Heinrich, S.U., Mothes, W., Brunner, J. & Rapoport, T. The Sec61p complex mediates the integration of a membrane protein by allowing lipid partitioning of the transmembrane domain. *Cell* **102**, 233–244 (2000).
37. Hessa, T. *et al.* Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* **433**, 377–381 (2005).
38. Hessa, T. *et al.* Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature* **450**, 1026–1030 (2007).
39. Enquist, K. *et al.* Membrane-integration characteristics of two ABC transporters, CFTR and P-glycoprotein. *J. Mol. Biol.* **387**, 1153–1164 (2009).
40. Do, H., Falcone, D., Lin, J., Andrews, D. & Johnson, A. The cotranslational integration of membrane proteins into the phospholipid bilayer is a multistep process. *Cell* **85**, 369–378 (1996).
41. Sadlish, H., Pitonzo, D., Johnson, A. & Skach, W. Sequential triage of transmembrane segments by Sec61α during biogenesis of a native multispanning membrane protein. *Nat. Struct. Mol. Biol.* **12**, 870–878 (2005).
42. McCormick, P.J., Miao, Y., Shao, Y., Lin, J. & Johnson, A. Cotranslational protein integration into the ER membrane is mediated by the binding of nascent chains to translocon proteins. *Mol. Cell* **12**, 329–341 (2003).
43. Meacock, S.L., Lecomte, F., Crawshaw, S. & High, S. Different transmembrane domains associate with distinct endoplasmic reticulum components during membrane integration of a polytopic protein. *Mol. Biol. Cell* **13**, 4114–4129 (2002).
44. Ismail, N., Crawshaw, S., Cross, B., Haagsma, A. & High, S. Specific transmembrane segments are selectively delayed at the ER translocon during opsin biogenesis. *Biochem. J.* **411**, 495–506 (2008).
45. Hamman, B.D., Hendershot, L. & Johnson, A. 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 (1998).
46. Skach, W. Topology of P-glycoproteins. *Methods Enzymol.* **292**, 265–278 (1998).
47. Xie, Y., Langhans-Rajasekaran, S., Bellovino, D. & Morimoto, T. Only the first and the last hydrophobic segments in the COOH-terminal third of the Na,K-ATPase α subunit initiate and halt, respectively, membrane translocation of the newly synthesized polypeptide. *J. Biol. Chem.* **271**, 2563–2573 (1996).
48. Skach, W.R., Calayag, M.C. & Lingappa, V. Evidence for an alternate model of human P-glycoprotein structure and biogenesis. *J. Biol. Chem.* **268**, 6903–6908 (1993).
49. Skach, W.R. & Lingappa, V. Amino terminus assembly of human P-glycoprotein at the endoplasmic reticulum is directed by cooperative actions of two internal sequences. *J. Biol. Chem.* **268**, 23552–23561 (1993).
50. Skach, W.R. *et al.* Biogenesis and transmembrane topology of the CHIP28 water channel in the endoplasmic reticulum. *J. Cell Biol.* **125**, 803–815 (1994).
51. Moss, K., Helm, A., Lu, Y., Bragin, A. & Skach, W. Coupled translocation events generate topologic heterogeneity at the endoplasmic reticulum membrane. *Mol. Biol. Cell* **9**, 2681–2697 (1998).
52. Audigier, Y., Friedlander, M. & Blobel, G. Multiple topogenic sequences in bovine opsin. *Proc. Natl. Acad. Sci. USA* **84**, 5783–5787 (1987).
53. Agre, P. *et al.* Aquaporin water channels—from atomic structure to clinical medicine. *J. Physiol. (Lond.)* **542**, 3–16 (2002).
54. King, L.S., Kozono, D. & Agre, P. From structure to disease, the evolving tale of aquaporin biology. *Nat. Rev. Mol. Cell Biol.* **5**, 687–698 (2004).
55. Fujiyoshi, Y. *et al.* Structure and function of water channels. *Curr. Opin. Struct. Biol.* **12**, 509–515 (2002).
56. Sui, H., Han, B.-G., Lee, J., Walian, P. & Jap, B. Structural basis of water specific transport through the AQP1 water channel. *Nature* **414**, 872–878 (2001).
57. Fu, D. *et al.* Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* **290**, 481–486 (2000).
58. Verbavatz, J.M. *et al.* Tetrameric assembly of CHIP28 water channels in liposomes and cell membranes. A freeze-fracture study. *J. Cell Biol.* **123**, 605–618 (1993).
59. Smith, B.L. & Agre, P. Erythrocyte Mr 28,000 transmembrane protein exists as a multisubunit oligomer similar to channel proteins. *J. Biol. Chem.* **266**, 6407–6415 (1991).
60. Buck, T.M., Wagner, J., Grund, S. & Skach, W. A novel tripartite structural motif involved in aquaporin topogenesis, monomer folding and tetramerization. *Nat. Struct. Mol. Biol.* **14**, 762–769 (2007).

61. Shi, L.-B., Skach, W., Ma, T. & Verkman, A. Distinct biogenesis mechanisms for water channels MIWC and CHIP28 at the endoplasmic reticulum. *Biochemistry* **34**, 8250–8256 (1995).
62. Lu, Y. *et al.* Reorientation of Aquaporin-1 topology during maturation in the endoplasmic reticulum. *Mol. Biol. Cell* **11**, 2973–2985 (2000).
63. Foster, W. *et al.* Identification of sequence determinants that direct different intracellular folding pathways for AQP1 and AQP4. *J. Biol. Chem.* **275**, 34157–34165 (2000).
64. Tu, L.W., Wang, J., Helm, A., Skach, W. & Deutsch, C. Transmembrane biogenesis of Kv1.3. *Biochemistry* **39**, 824–836 (2000).
65. Lu, Y. *et al.* Co- and posttranslational mechanisms direct CFTR N-terminus transmembrane assembly. *J. Biol. Chem.* **273**, 568–576 (1998).
66. Gratkowski, H., Lear, J. & DeGrado, W. Polar side chains drive the association of model transmembrane peptides. *Proc. Natl. Acad. Sci. USA* **98**, 880–885 (2001).
67. Zhou, F.X., Cocco, M., Russ, W., Brunger, A. & Engelman, D. Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nat. Struct. Biol.* **7**, 154–160 (2000).
68. Gafvelin, G. & von Heijne, G. Topological "frustration" in multispanning *E. coli* inner membrane. *Cell* **77**, 401–412 (1994).
69. Goder, V., Bieri, C. & Spiess, M. Glycosylation can influence topogenesis of membrane proteins and reveals dynamic reorientation of nascent polypeptides within the translocon. *J. Cell Biol.* **147**, 257–266 (1999).
70. Garrison, J., Kunkel, E., Hegde, R. & Taunton, J. A substrate-specific inhibitor of protein translocation into the endoplasmic reticulum. *Nature* **436**, 285–289 (2005).
71. Bogdanov, M., Xie, J., Heacock, P. & Dowhan, W. To flip or not to flip: lipid-protein charge interactions are a determinant of final membrane protein topology. *J. Cell Biol.* **182**, 925–935 (2008).
72. Brambillasca, S., Yabal, M., Makarow, M. & Borgese, N. Unassisted translocation of large polypeptide domains across phospholipd bilayers. *J. Cell Biol.* **175**, 767–777 (2006).
73. Lin, J. & Addison, R. 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 (1995).
74. Ismail, N., Crawshaw, S. & High, S. Active and passive displacement of transmembrane domains during opsin biogenesis at the Sec61 translocon. *J. Cell Sci.* **119**, 2826–2836 (2006).
75. Beckmann, R. *et al.* Alignment of conduits for the nascent polypeptide chain in the ribosome-Sec61 complex. *Science* **278**, 2123–2126 (1997).
76. Ménétret, J.F. *et al.* The structure of ribosome-channel complexes engaged in protein translocation. *Mol. Cell* **6**, 1219–1232 (2000).
77. Hanein, D. *et al.* Oligomeric rings of the Sec61p complex induced by ligands required for protein translocation. *Cell* **87**, 721–732 (1996).
78. Morgan, D.G., Menetret, J., Neuhof, A., Rapoport, T. & Akey, C. Structure of the mammalian ribosome-channel complex at 17Å resolution. *J. Mol. Biol.* **324**, 871–886 (2002).
79. Young, J.C., Agashe, V., Siegers, K. & Hartl, F. Pathways of chaperone-mediated protein folding in the cytosol. *Nat. Rev. Mol. Cell Biol.* **5**, 781–791 (2004).
80. Oberdorf, J., Pitonzo, D. & Skach, W. An energy-dependent maturation step is required for release of the cystic fibrosis transmembrane conductance regulator from early endoplasmic reticulum biosynthetic machinery. *J. Biol. Chem.* **280**, 38193–38202 (2005).
81. Clemons, W.M., Menetret, J.-F., Akey, C. & Rapoport, T. Structural insight into the protein translocation channel. *Curr. Opin. Struct. Biol.* **14**, 390–396 (2004).
82. Breyton, C., Haase, W., Rapoport, T., Kuehlbrandt, W. & Collinson, I. Three-dimensional structure of the bacterial protein-translocation complex SecYEG. *Nature* **418**, 662–665 (2002).
83. Mitra, K. *et al.* Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. *Nature* **438**, 318–324 (2005).
84. Hamman, B.D., Chen, J.-C., Johnson, E. & Johnson, A. The aqueous pore through the translocon has a diameter of 40–60Å during cotranslational protein translocation at the ER membrane. *Cell* **89**, 535–544 (1997).
85. Kida, Y., Morimoto, F. & Sakaguchi, M. Two translocating hydrophilic segments of a nascent chain span the ER membrane during multispanning protein topogenesis. *J. Cell Biol.* **179**, 1441–1452 (2007).