



Review

# Folding and Insertion of Transmembrane Helices at the ER

Paul Whitley <sup>1</sup>, Brayán Grau <sup>2</sup>, James C. Gumbart <sup>3</sup>, Luis Martínez-Gil <sup>2</sup> and Ismael Mingarro <sup>2,\*</sup>

<sup>1</sup> Department of Biology and Biochemistry, Centre for Regenerative Medicine, University of Bath, Bath BA2 7AY, UK; bssprw@bath.ac.uk

<sup>2</sup> Department of Biochemistry and Molecular Biology, Institute of Biotechnology and Biomedicine (BIOTECMED), Universitat de València, E-46100 Burjassot, Spain; Brayán.Grau@uv.es (B.G.); Luis.Martinez-gil@uv.es (L.M.-G.)

<sup>3</sup> School of Physics, School of Chemistry and Biochemistry, Parker H. Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, GA 30332, USA; gumbart@physics.gatech.edu

\* Correspondence: mingarro@uv.es; Tel.: +34-963543796

**Abstract:** In eukaryotic cells, the endoplasmic reticulum (ER) is the entry point for newly synthesized proteins that are subsequently distributed to organelles of the endomembrane system. Some of these proteins are completely translocated into the lumen of the ER while others integrate stretches of amino acids into the greasy 30 Å wide interior of the ER membrane bilayer. It is generally accepted that to exist in this non-aqueous environment the majority of membrane integrated amino acids are primarily non-polar/hydrophobic and adopt an  $\alpha$ -helical conformation. These stretches are typically around 20 amino acids long and are known as transmembrane (TM) helices. In this review, we will consider how transmembrane helices achieve membrane integration. We will address questions such as: Where do the stretches of amino acids fold into a helical conformation? What is/are the route/routes that these stretches take from synthesis at the ribosome to integration through the ER translocon? How do these stretches ‘know’ to integrate and in which orientation? How do marginally hydrophobic stretches of amino acids integrate and survive as transmembrane helices?

**Keywords:** folding; insertion; membrane protein; translocon; ribosome; transmembrane segment



**Citation:** Whitley, P.; Grau, B.; Gumbart, J.C.; Martínez-Gil, L.; Mingarro, I. Folding and Insertion of Transmembrane Helices at the ER. *Int. J. Mol. Sci.* **2021**, *22*, 12778. <https://doi.org/10.3390/ijms222312778>

Academic Editor:  
Richard Zimmermann

Received: 14 October 2021  
Accepted: 23 November 2021  
Published: 26 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The majority of, if not all, integral membrane proteins distributed throughout the endomembrane network in eukaryotic cells first assemble into the endoplasmic reticulum (ER) membrane. Following this assembly, these proteins are distributed to their intended destinations via specific trafficking pathways. The signals and machineries that direct complex endomembrane trafficking pathways are beyond the scope of this article. Thus, we will concentrate on the initial folding and assembly of proteins into the ER membrane and not their subsequent cellular distribution.

Most integral membrane proteins require proteinaceous machinery known as an integrase for their insertion into the ER [1]. The Sec61 translocon/integrase is the primary integration machinery of the ER although there are others, notably the Get1/Get2 integrase that is responsible for tail-anchored (TA) protein integration [2,3]. A further integrase called the ER membrane protein complex (EMC) is also implicated in post-translational membrane integration of a subset of TA proteins but also seems to play additional roles in co-translational membrane protein assembly [4]. These integrases have been evolutionarily conserved in eukaryotes [5] with some of them, such as Sec61 and Get1/2, having homologs in prokaryotes. In this article, we will exclusively consider membrane protein assembly mediated by the Sec61 translocon.

## 2. Structure-Function of the Translocon

The mammalian Sec61 translocon consists of a core heterotrimeric Sec61 $\alpha$ ,  $\beta$ , and  $\gamma$  complex which have 10, 1, and 1 transmembrane (TM) domains respectively [6]. In ad-

dition, numerous accessory proteins are associated with this core complex which may modulate translocon activity or provide functionality that complements its translocation and integrase function. Accessory protein complexes such as TRAP seems to be involved in membrane protein topogenesis [7–9] and oligosaccharyltransferase, which adds oligosaccharides to asparagine residues in the lumen of the ER (at NXS/T sequons), are commonly found associated with the translocon core and/or are present in approximately stoichiometric amounts [10,11]. Others such as signal peptidase complex, TRAM, or Sec62/63 are present in sub-stoichiometric amounts and may only associate transiently or under certain circumstances when particular substrates are present in the translocon [11]. Recently, it was discovered that the EMC is also present as a cooperative partner of the Sec61 translocon machinery during co-translational membrane-protein insertion [12]. Thus, the translocon should be considered as a dynamic rather than a well-defined complex.

The structure of the core Sec61 translocon and homologous complexes such as SecYEG of *Escherichia coli* and SecYE $\beta$  of the archaeal *Methanococcus janaschii* have been extensively studied [6]. The structures are all fundamentally similar, suggesting a common mode of action in facilitating the translocation of polypeptide chains across membranes and the integration of appropriate stretches of amino acids into the lipid bilayer. The structure/function relationship of all translocons that is most widely accepted is as follows: Looking from the top, onto the membrane the 10 TM subunit (Sec61 $\alpha$ /SecY) is pseudosymmetrical with TM domains 1–5 and 6–10 forming the two halves in a clamshell arrangement (Figure 1A–C); A hinge region at one interface between the two halves consisting of TM5 and TM6 allows for the opening and closing (breathing) of a lateral gate between TM2b and TM7 at the opposite side of the clamshell; this breathing is proposed to allow lateral movement of polypeptides out of the translocon and into the membrane providing the integrase function. From a different view, looking into the membrane from the side (Figure 1D) there appears to be a continuous channel through the membrane that is constricted at the center of the membrane and wider at both extremities, resembling an hourglass. The constriction is lined with six hydrophobic amino acid residues [13]. Furthermore, in structures without a translocating polypeptide, there is a small ‘plug’ helix (on the luminal/non-cytoplasmic side) blocking the continuous channel through the membrane. It is proposed that this small helix prevents the leakage of ions when the translocon is not occupied by a translocating polypeptide chain [14]. When there is a translocating or inserting polypeptide chain it is proposed that the plug helix is displaced so that it no longer blocks the continuous channel [15] (Figure 1C). Ion leakage is still prevented by the presence of the translocating polypeptide chain and the hydrophobic amino acids in the central constriction, which has been proposed to act as a greasy gasket forming a tight seal around the translocating chain [16,17]. However, the gasket role of the central constriction is seemingly dispensable as yeast expressing mutant Sec61 with the hydrophobic residues replaced with even charged residues can grow and perform translocation efficiently [18]. Regarding the  $\beta$ - and  $\gamma$ -subunits of the Sec61 translocon, the former is not essential for TM insertion, and the latter acts as a clamp that brings both halves of Sec61 $\alpha$  together [19].

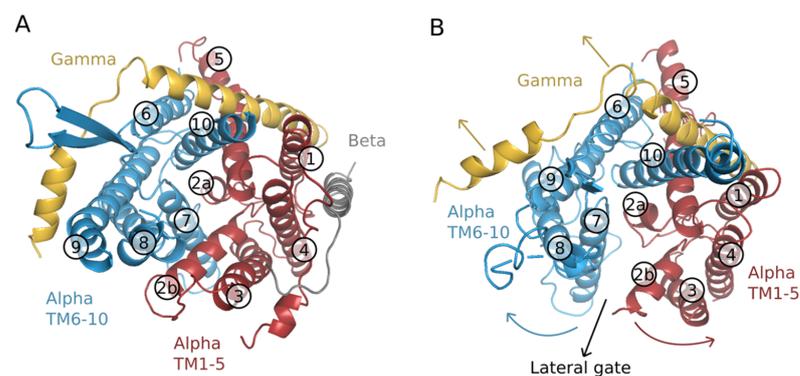
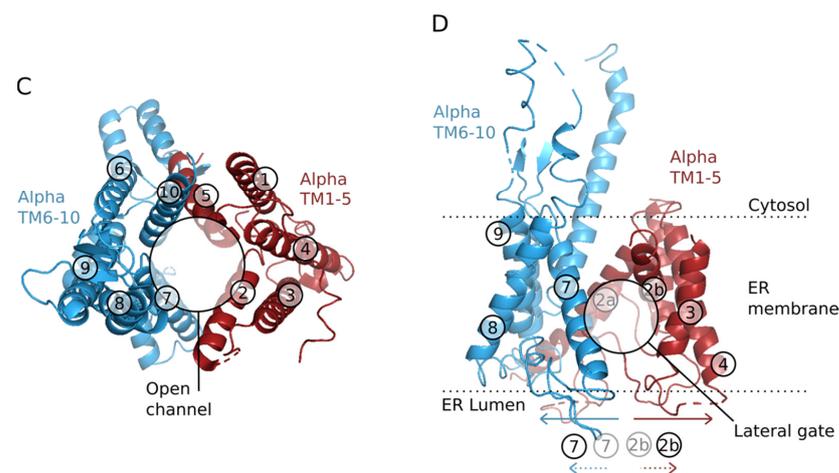


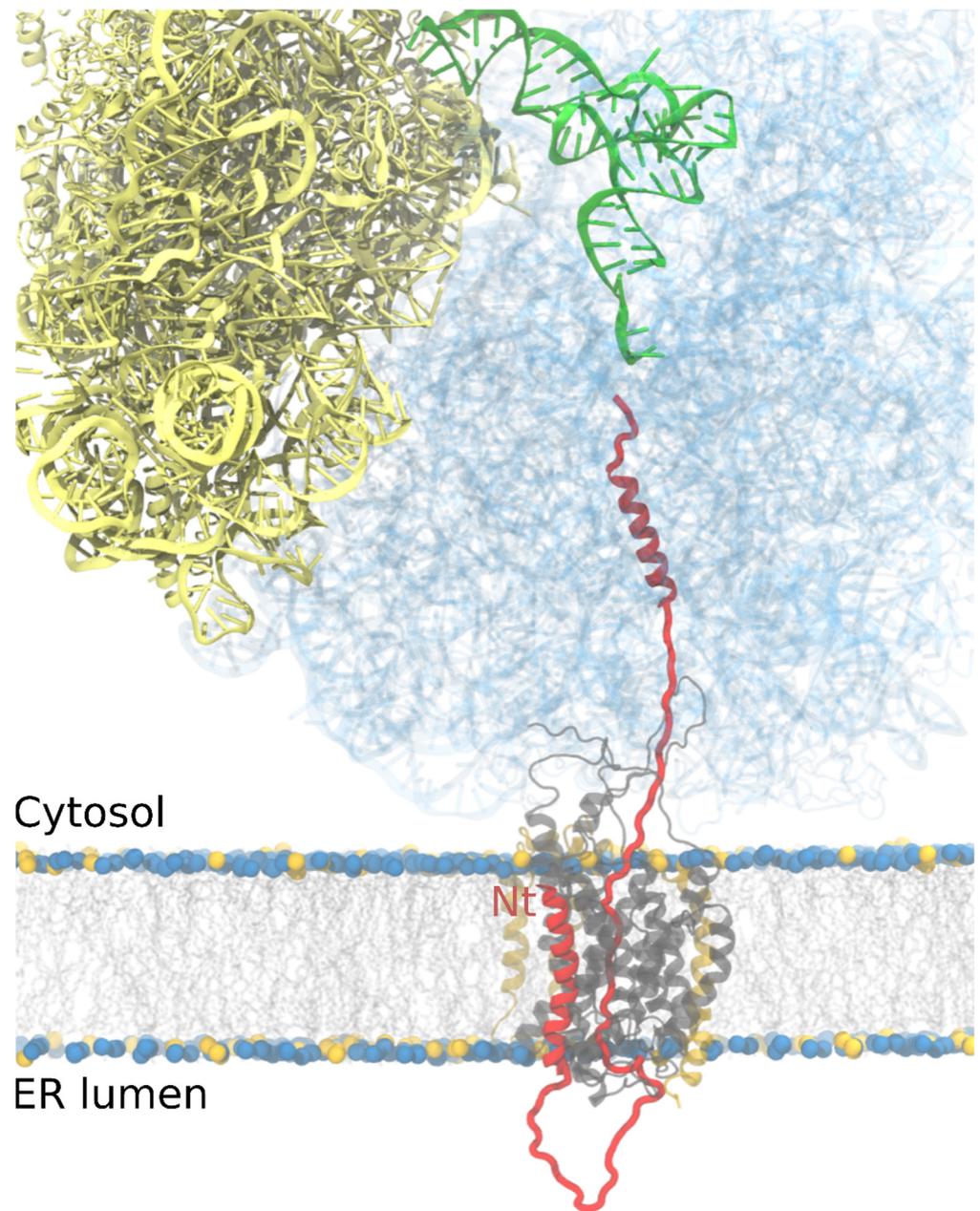
Figure 1. Cont.



**Figure 1.** Structure of the translocon. Structure-based cartoon representations of the translocon. All TM segments of Sec61 $\alpha$  are colored (red and blue for each half) except for the  $\beta$ -subunits and  $\gamma$ -subunits, which are shown in gray and yellow respectively. All TM segments are numbered for easy comparison between the open and closed structures. **(A)** Top view of the translocon at a resting/close position from *M. jannaschii* (PDB: 1RHZ). **(B)** Top view of the partially open structure of the translocon from *P. furiosus* (PDB: 3MP7). The colored arrows in red and blue indicate the helix displacements required for the widening of the channel and opening of the lateral gate. The black arrow shows the lateral gate exit pathway of a TM segment from the interior of the channel into the membrane. **(C)** Top view of the translocon with the plug (TM 2a) open from *Saccharomyces cerevisiae* (PDB: 7KAL). The open channel facilitates translocation into the ER lumen of secreted and TM proteins segments. **(D)** Lateral view of the partially open structure of the translocon from *P. furiosus* (PDB: 3MP7). The colored arrows (red and blue) indicate the helix displacements required for the widening of the channel and opening of the lateral gate. The dotted colored arrows at the base of the panel indicate the movement of helices 2b and 7 (compared to a close state, panel A). The limits of the membrane are shown with a dotted line, data from the OPM database (<https://opm.phar.umich.edu/> (accessed on 18 November 2021)).

### 3. Targeting to the Translocon

In higher eukaryotes such as mammals, the majority of proteins utilizing the Sec61 translocon for entry into the ER do so in a co-translational mode [20–22]. It should be noted, however, that there is also a post-translational mode of targeting to the ER that is more prevalent in lower eukaryotes such as fungi that will not be considered here [22]. In co-translational mode, proteins that are to be targeted to the ER are synthesized with a hydrophobic N-terminal sequence. This sequence can either be a signal peptide (signal sequence) that is eventually cleaved from the mature protein once it reaches the ER or an uncleaved TM segment (signal anchor) which comes in two different flavors, conventional (N-terminal cytoplasmic) (Figure 2) or reverse (N-terminal ER lumen). Upon emerging from the ribosome exit tunnel, the N-terminal hydrophobic sequence on the newly synthesizing (nascent) protein is recognized by a ribonucleoprotein chaperone complex called signal recognition particle (SRP). The binding of SRP arrests translation by the ribosome and a ribosome/nascent chain/SRP complex is formed. Targeting to the ER membrane is facilitated by a specific interaction between SRP and an SRP receptor, which is integral in the ER membrane. SRP releases from the complex upon hydrolysis of GTP, reversing the arrest in translation with the ribosome/nascent chain situated on top of a translocon (Figure 2).



**Figure 2.** Pathway of a nascent protein. The ribosome is shown in yellow (small subunit) and light blue (large subunit, transparent). A P-site tRNA is shown in green. The nascent chain (red) traverses the ribosome's exit tunnel before encountering the membrane-bound translocon complex (transparent orange and grey). The nascent chain has been shown to form a secondary structure in the translocon as well as at the ribosome's exit tunnel. Most of the structure shown is taken from PDB 4V6M [23], although the nascent-chain helix inside the exit tunnel is modeled.

#### 4. Getting across the Membrane

When translation resumes in the vicinity of the translocon, signal sequences/anchors have to orient themselves appropriately within the translocon. Positively charged residues positioned at one end of the hydrophobic stretch of amino acids play an important role in establishing the orientation of this initial interaction with the translocon together with other ER components, i.e., lipids [24] or potentially accessory proteins. Positive charges at the N-terminus of signal sequences and signal anchors position these types of topogenic sequences with their N-termini in the cytoplasm and direct translocation of C-terminal sequences into the lumen of the ER ( $N_{in}$ - $C_{out}$ ) (Figure 2). When situated at the C-terminus

of a hydrophobic sequence, a C-terminus cytoplasmic orientation is favored and the N-terminus is translocated ( $N_{out}-C_{in}$ ) [25]. It should be noted that other factors such as the length of hydrophobic stretch of amino acids or folding of previous sequence domains can affect this initial topogenic decision [26,27].

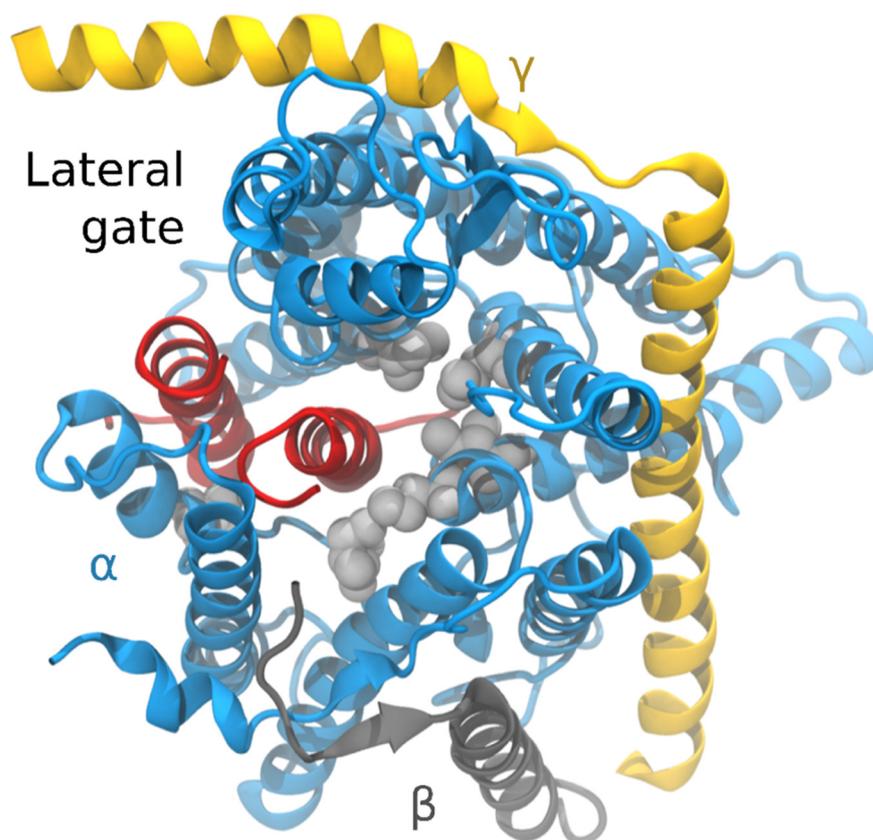
A cryo-EM structure of a ribosome/Sec61 complex in the process of translating/translocating pre-prolactin stalled at 86 residues shows the signal sequence positioned next to TM2 and TM7 in the open lateral gate of Sec61 $\alpha$  and density in the translocon channel and ribosome tunnel are consistent with a looped configuration of the nascent chain [28]. If translation were to proceed, the loop would get longer on the luminal side of the membrane as the polypeptide chain extends (Figure 2), at least until the signal sequence is cleaved and the translocating sequence is no longer a loop.

It has been shown using glycosylation mapping that translocating polypeptides can span the Sec61 translocon in an extended conformation but that  $\alpha$ -helices can also be accommodated in the translocon [29,30]. It is broadly accepted that more hydrophobic polypeptides in an  $\alpha$ -helical conformation move from the translocon channel into the lipid bilayer more easily than less hydrophobic segments in an extended conformation that will be translocated to the ER lumen. The potential routes that topogenic signals in nascent chains take into the translocon are the topic of later sections in this article.

## 5. Integrating into the Membrane

Membrane-spanning domains are typically made up of around 20 hydrophobic amino acids in an  $\alpha$ -helical conformation. This is sufficient to span the 30 Å hydrophobic core of a model lipid bilayer as each amino acid contributes 1.5 Å to the length of a helix. Not only does this feature of TM domains make them stable in a hydrophobic rather than aqueous environment, but it is also apparent that high hydrophobicity is a major driving force for integration [31]. There seems to be a hydrophobicity threshold above which stretches of amino acids partition into the membrane from the translocon, and below which they continue their translocation through the translocon and into the lumen of the ER. The majority of single-span membrane proteins have TM domains that have a hydrophobicity above this threshold [32] and therefore are theoretically capable of integration in an autonomous (unassisted) way.

This all seems very straightforward. A sufficiently hydrophobic sequence will integrate while others will not. However, things are not that simple. Depending on the topology of the nascent polypeptide, inversions can occur via two sequential energetic transitions of the TM segment: first, the insertion, driven by the hydrophobic effect, and second, the inversion that has been proposed to be driven by electrostatic interactions between the nascent chain and the translocon (or associated proteins) and/or membrane lipids [33]. Furthermore, in multi-spanning proteins, at least 25% of TM helices from proteins of known structure do not reach the threshold that would allow them to integrate autonomously by a simple hydrophobic partitioning [32,34]. Low-hydrophobic segments could be helped by interacting with upstream or downstream hydrophobic segments in the nascent chain to adopt a TM disposition [35] since photocrosslinking experiments have shown that the translocon can simultaneously accommodate more than one TM helix [36–38] (Figure 3).



**Figure 3.** The translocon is shown from the cytoplasmic side in light blue (Sec61 $\alpha$ ), orange (Sec61 $\gamma$ ), and grey (Sec61 $\beta$ ). Two helices of a nascent chain are shown in red, one residing in the translocon channel and one at the open lateral gate. Pore ring residues are shown in a white space-filling representation. The structure was modeled based on MD simulations of PDB 1RHZ [13].

## 6. Where Does Folding of TM Helices Occur?

First, we want the reader to reflect on why TM domains of the type of integral membrane protein we are considering in this review are exclusively  $\alpha$ -helical. The answer is that a regular pattern of hydrogen bonding between carbonyl oxygens and amide hydrogens in peptide bonds in an  $\alpha$ -helix secondary structure ensures maximal hydrophobicity by effectively neutralizing polar groups in the peptide bonds. It has been calculated that a stretch of hydrophobic amino acids in an extended conformation could not reach the hydrophobic threshold to partition through the lateral gate into the lipid environment [39–42]. Thus, we assume that hydrophobic stretches of amino acids must be in an  $\alpha$ -helical conformation to partition into the lipid environment in a thermodynamically favorable way. Hydrophobic stretches of amino acids may fold into  $\alpha$ -helices in the translocon tunnel as they enter. However, it is also feasible that the secondary structure is formed before entering the translocon (Figure 2).

During co-translational membrane insertion secondary structures can be acquired in the ribosome specifically in the ribosome exit tunnel. This tunnel accommodates nascent peptides from the ribosomal peptidyl transferase site (P-site) to the ribosomal exit site, providing a protective shield. The first evidence for folding in the ribosome before insertion came more than two decades ago, indicating that the constrained environment of the ribosome-translocon complex has an active role in the propensity of some nascent chains to acquire an extended or a more compact conformation [29]. Furthermore, the authors determined that this compaction was acquired co-translationally in an amino acid-dependent manner. A few years later, it was discovered that the ribosome exit tunnel itself provides enough space to allow the folding of secondary structures as  $\alpha$ -helices as Johnson and colleagues demonstrated attaching fluorescent FRET partners at different positions in nascent

peptides. The results indicated that folding of TM sequences in an  $\alpha$ -helix-like structure was induced and stabilized far inside the ribosome exit tunnel and close to the ribosomal P-site, whereas a nascent secretory protein remained in an extended conformation within the ribosome exit tunnel [43]. These results were further supported by biochemical assays based on pegylation of cysteine reporters [44]. These studies also established that there are seemingly different folding zones within the ribosome exit tunnel where the secondary structure formation can occur, as alanine-replaced peptide segments used for the pegylation experiments folded more compactly when located near the P-site than when located more distally in the tunnel [45].

Following these biochemical studies where compaction could be inferred, atomic structures of  $\alpha$ -helical nascent polypeptide chains visualized within the ribosome exit tunnel were quick to follow. Utilizing single-particle cryo-electron microscopy reconstitutions of eukaryotic 80S ribosomes containing nascent chains with an  $\alpha$ -helical propensity, Beckmann and collaborators visualized helix density inside the tunnel as well as interaction sites with the tunnel wall components [46].

It is thought that  $\alpha$ -helix formation occurs first wherever it is possible and that the ribosome may be selective in the types of sequences that it allows to form secondary structures. Using glycosylation mapping and molecular dynamics (MD) simulations [47] it was shown that in combination, hydrophobicity, helical propensity, and length of hydrophobic stretches of amino acids are major determinants for  $\alpha$ -helix adoption within the ribosome exit tunnel (Figure 2). In these experiments, nascent chains harboring  $\alpha$ -helix sequences from TM segments were able to fold inside the ribosome exit tunnel whereas those with  $\alpha$ -helical sequences of soluble proteins were not.

It is, however, not only secondary structures that can be adopted inside the ribosome tunnel. Small domains with a tertiary structure such as zinc-finger domains have been observed to fold close to the ribosome exit site, also called the vestibule, where the ribosome exit tunnel widens substantially [48], with this folding accelerated and stabilized by the tunnel [49]. Furthermore, helical TM hairpin formation is also possible in the vestibule of the ribosome exit tunnel [50]. Currently, we think of the ribosome/translocon complex as specialized chaperone-like machinery implicated in the formation of secondary or tertiary structure as a platform to overcome the huge energy barrier required to insert unfolded hydrophobic sequences and to prevent their exposure on the cytosolic side of the ER membrane [36,47].

The implication of the ribosome in promoting the secondary structure of TM domains raises a question, 'could this be an evolutionarily conserved function of ribosomes? Presumably, the very first TM proteins had to insert into membranes without the assistance of translocon machinery. There are numerous examples of  $\alpha$ -helical membrane proteins that can insert into simplified biomimetic systems without translocon assistance [51–53]. Recently proteorhodopsin has been shown to spontaneously integrate into a simple lipid membrane in the absence of chaperones or a translocon [54]. Other evidence of the pre-translocon era is that yeast mitochondrial ribosomes permanently attach to the mitochondrial inner membrane which lacks the translocon [55–57]. Nevertheless, other components tethering translating ribosomes to the mitochondrial inner membrane have been identified [55]. The permanent docking of these ribosomes facilitates the insertion of membrane proteins encoded by mitochondrial DNA efficiently. Therefore, not only the ribosome but also the recruitment of translating ribosomes to membranes seem to be crucial to a successful insertion and folding process. All these examples highlight the importance of the ribosome in the process of membrane insertion, however, we should be mindful that it is not only TM domains and ribosomes that have co-evolved. There is evidence that extant biological membranes have arisen as a consequence of a lipid bilayer and membrane protein co-evolution process [58]. Altogether, it is highly probable that the ribosome's ability to assist in  $\alpha$ -helix folding is not a coincidence exploited by membrane proteins for rapid structure acquisition before membrane insertion; rather it seems likely it is the result of a fine-tuned co-evolution process that continues to this day.

## 7. Route Into, Through, and Out of the Translocon

In addition to its insertase activity, the translocon is responsible for the secretion of proteins to the ER lumen. The insertion or secretion of proteins can occur both co-translationally or post-translationally depending on the presence of an N-terminal signal sequence (SS). The SS is characterized by a hydrophobic core, often the first TM segment for membrane proteins, flanked by positively charged amino acid residues and polar but uncharged residues at the N- and C-termini, respectively [59]. If the SS is present, once the ribosome is brought to the membrane, it engages with the Sec61 translocon while the nascent chain is still being translated (Figure 2). As mentioned above, the presence of positively charged residues at the N-terminus of the SS causes inversion of its topology, aided through its interactions with lipids and possibly nearby ribosomal RNA [23,33,60].

Within the ribosome/translocon complex the cytoplasmic entrance of the Sec61 $\alpha$  channel has a diameter of 20–25 Å [13] (Figure 1D), which nicely matches with the diameter of the ribosome exit tunnel vestibule [61], allowing ribosome-acquired helices to enter into the aqueous translocon channel [62]. When the nascent chain arrives at the laterally closed translocon, it has been demonstrated that the presence of the SS triggers channel opening of the translocon at the lateral gate, which is also linked to the partitioning of TM  $\alpha$ -helical structures of membrane proteins into the lipid phase [13,63]. Subsequent hydrophobic regions of the nascent chains will be exposed to the lipid phase and can partition into it as experimental data, structures, and simulations have revealed [64–68] (Figure 3). Recent single-molecule FRET experiments indicate that the lateral gate is highly dynamic even in the absence of a membrane-inserting SS or TM segment, although binding of the ribosome and insertion of TM segments increases the probability of the open state [69]. Indeed, earlier cryo-tomography and fluorescence studies also revealed that binding of the ribosome induces channel opening [70,71].

As noted earlier, for multi-pass membrane proteins, the topology is typically established by the first TM, which adopts either an N<sub>in</sub>-C<sub>out</sub> (N-terminus in the cytoplasm and C-terminus in the ER lumen) or an N<sub>out</sub>-C<sub>in</sub> topology in the translocon. Each subsequent TM segment adopts the opposite topology of the previous one as it enters the membrane. Recent evidence has implicated the EMC in helping to establish the topology of the first TM of many GPCRs and likely other proteins [12,72]. The translocon's lateral gate allows direct sampling of the membrane environment by the TM [73], possibly from the moment of its encounter with the translocon [39,74]. The code for insertion, i.e., what sequences partition to the membrane vs. remain in the channel, has been determined in exquisite detail with extensive, elegant experiments from von Heijne, White, and colleagues in the mid-2000s [31,32]. Nonetheless, in addition to thermodynamics, a role for kinetics has also been found from both simulations [65,75,76] and experiments [18,77,78]. In particular, slowing down the rate of protein synthesis can increase the probability of membrane insertion.

While most membrane proteins have a fixed topology that alternates from one TM to the next, some defy this straightforward expectation and are known as “dual-topology proteins”. These proteins can be found in roughly equal numbers in an N<sub>in</sub>-C<sub>out</sub> or an N<sub>out</sub>-C<sub>in</sub> topology [79]. In particular, the topology of EmrE, a canonical example of a dual-topology protein, has been found to be very sensitive to the presence of charges, not just at its N-terminus but in any of its loops and even at its C-terminus [80]. This surprisingly suggests that even after synthesis is completed, dual-topology proteins can exist in a dynamic equilibrium where TM helices flip in and out of the membrane, probably in the vicinity of the translocon [81]. Coarse-grained simulations indicated this unusual ability arises from a lack of full integration of some TMs until well after the completion of synthesis [82].

## 8. Exploring the Limits of TM Domain Insertion

We have previously described how TM helices are proposed to partition from the translocon into the non-polar core of the lipid bilayer driven by hydrophobicity with, the limits for the insertion of TM segments being explored using computationally designed

segments with naturally occurring amino acid distributions [83]. However, features found in naturally occurring TM domains such as limited length, charged residues, or a low hydrophobicity profile challenge this rather simplistic view of what a TM segment is.

Given that exposure of hydrophobic groups in proteins and lipids to water is highly unfavorable, membrane proteins tend to minimize their free energy by maximizing the match between the hydrophobic width of the bilayer and the length of a TM segment, a phenomenon called hydrophobic matching [84,85]. Indeed, the average length of a TM segment is  $24 \pm 5.6$  residues long (36 Å in a 3,6  $\alpha$ -helix) [86], while for instance, the width of the hydrophobic core of an ER hepatocyte membrane is close to 38 Å [87–89]. A mismatch in length between the hydrophobic section of a membrane-spanning protein and the bilayer in which it is located results in lipid and peptide rearrangements to compensate [90]. Ultimately a major hydrophobic mismatch might prevent insertion into the membrane, but how much is too much?

The minimum hydrophobic length necessary to form a TM segment in lipid bilayers has been investigated using short hydrophobic peptides in dioleoylphosphatidylcholine (DOPC) and 1,2-dierucoyl-*sn*-glycero-3-phosphocholine (DeuPC, a shorter lipid) vesicles [91]. Peptides composed of Leu residues were compared to sequences of the same length containing alternating Leu and Ala residues (which have a hydrophobicity typical of natural TM helices) [91]. The authors observed that peptides composed exclusively of Leu residues were able to adopt a TM disposition with just 11–12 residues. In this case, the bilayer width exceeded the hydrophobic length of the peptide by ~11–12 Å. For the alternating Leu/Ala sequence 13 residues representing a negative mismatch of ~9 Å were required. The minor differences indicate that the minimum length necessary to form a TM segment is only modestly hydrophobicity-dependent, at least for the sequences tested in this study.

In vitro expression and MD simulations have also been used to systematically examine the insertion efficiency of TM segments consisting primarily of leucine residues [92,93]. Depending on the flanking residues the minimum length to achieve a ~100% insertion efficiency varied from 12, with Gly, Asn, or Asp rich flanking sequences, to 10 residues when Lys was used as a flanking residue. The MD simulations suggest that the insertion efficiency of these sequences is determined primarily by the energetic cost of distorting the bilayer in the vicinity of a short TM segment. The presence of Lys residues flanking the hydrophobic core can reduce the energetic cost by extensive hydrogen bonding with water and lipid phosphate groups (snorkeling) and by partial backbone unfolding. The unfolding is stabilized in the simulation by water molecules entering the bilayer along the peptide backbone.

The studies cited above utilized model hydrophobic sequences composed of only a few different kinds of amino acids. However, TM segments in native membrane proteins vary significantly not only in length but also in composition, revealing more complex scenarios [86]. A computational analysis of the composition and location of amino acids in TM helices found, in membrane proteins of known structure, a strong reverse correlation between the composition/overall hydrophobicity and the required length for their insertion in the lipid bilayer [83]. These results were in accordance with in vitro studies, in which the length dependence of a TM segment (varying from 10 to 25) strongly depends on its amino acid composition [32]. Furthermore, the analysis of naturally occurring residues in TM segments put the focus on the importance of residue positioning, particularly Pro and charged residues.

At the other extreme, the longest TM segment found in naturally occurring membrane proteins is close to 40 residues long [86]. Long TM segments usually adapt their hydrophobic length to the lipid membrane by tilting [86,94]. Accordingly, tilting of long hydrophobic sequences should occur before their insertion in the lipid bilayer, that is, before or during its partition from the translocon into the membrane. An extensively long TM segment might not be easily accommodated within the translocon in its tilted disposition which could prevent its insertion as a single TM domain. Interestingly, a 40-residue long hydropho-

bic sequence could potentially span a membrane twice in a helical hairpin conformation. The transformation from a long TM domain into a helical hairpin depends primarily on the presence of turn propensity residue(s) in the middle of the sequence both in natural and polyLeu sequences [95,96]. The probability of possessing one of these residues increases tremendously as sequence length extends beyond the 40-residue mark. Increasing the number of residues beyond 40 will facilitate the introduction of a turn propensity residue while maintaining the minimum distance to cross the membrane twice.

Another challenge to the idealized view of a TM segment is the presence of charged residues. The prevalence of these residues within the buried sections of membrane proteins is very low, as expected based on their polarity [86]. However, sequence analysis of membrane protein databases shows that ionizable amino acid residues are indeed present in TM domains, often with functional and/or structural roles [86,97]. How then, are TM domains containing charged residues inserted into the hydrophobic core of the lipid bilayer? The strong hydrophobic contribution of the abundant non-polar residues in a TM segment might be greater than the energy penalty of introducing a charge into the membrane. This negative  $\Delta\Delta G$  should be enough to promote the insertion. Interestingly, computer simulation studies suggest that the transfer of four leucines from water to the bilayer interior could be sufficient to compensate for the transfer of a cationic residue [98]. Experiments have also corroborated these predictions [99]. The position within the membrane and the polarity of the residue is also a factor that needs to be taken into consideration when determining the energy necessary for their insertion. Residues close to the water lipid interface can be hydrated due to the presence of “water defects” that reduce the insertion penalty. The closer the polar residue is to the center of the bilayer the larger (and increasingly unfavorable) the water defects become.

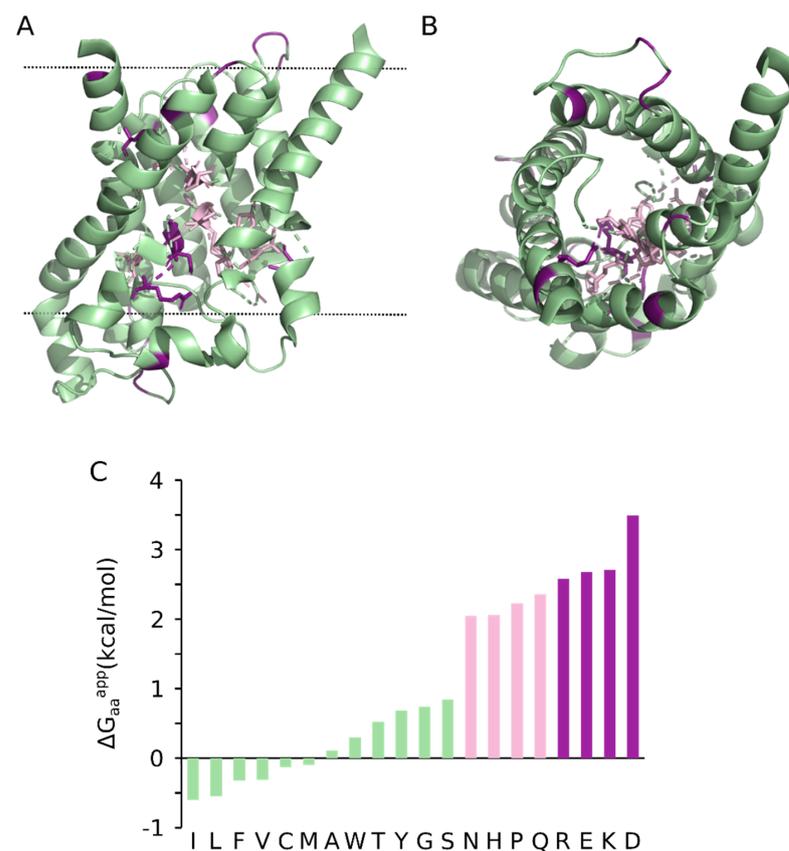
The presence of water molecules within the lipid bilayer negatively affects membrane integrity. Interestingly, destabilization by the presence of water molecules has been exploited by some proteins to permeabilize the membrane. Water defect-inducing residues, that is, charged and polar residues surrounded by hydrophobic amino acids, are indeed found in pore-forming proteins [100–102]. Importantly, once in the membrane, polar residues strongly influence the folding or association of integral membrane proteins [97,103–105] and their activity [106,107].

The presence of polar residues within TM segments cannot always be explained by the large hydrophobicity of the surrounding amino acids compensating for the energy penalty of their membrane integration. As mentioned previously, around 25% of TM segments do not reach the hydrophobicity threshold required for autonomous partitioning into the membrane. In these cases, other forces should facilitate the localization of polar amino acids in the hydrophobic environment of the lipid bilayer. Interaction between TM segments within the translocon (Figure 3) provides an opportunity for polarity masking (see Section 5). Although rare, it has been demonstrated that polar interactions between neighboring helices can facilitate insertion [95,108]. Since these interactions are necessary for the partitioning into the lipid bilayer, a cooperative insertion of the TM segments is predicted which requires the presence at the same time of, at least, two helices in the translocon [109] (Figure 3). Similarly, it has been suggested that salt-bridge formation between residues located on the same face of a single TM domain may reduce the free energy of membrane partitioning.

Of note, in a low dielectric constant environment such as the membrane core, the force of an electrostatic interaction increases tremendously, thus creating very stable associations [103,104]. It has been shown that hydrogen bonding between TM segments gives stronger associations than the packing of surfaces in glycoporphin A helices driven by the GxxxG interaction domain [104,110].

In channel-forming proteins, such as aquaporins (Figure 4), the presence of polar/charged residues close to the membrane hydrophobic core is not explained either by the strong hydrophobicity of the accompanying lipid-facing residues or by electrostatic interactions with other residues. Membrane channels create an amphipathic environment.

Some of the residues in them (the apolar ones), regardless of the depth they are found within the membrane, will be facing the hydrophobic core of the membrane while others (the polar/charged residues) will be exposed to the water-filled tunnel of the channel (Figure 4). The nature of the residues in those two different situations varies accordingly, so the overall amino acid composition of the helices that constitute these channels resemble that of interfacial (amphipathic) helices where one side of the helix is polar while the other is filled with hydrophobic residues. In this case, there is no energy penalty for the inclusion of a charged residue in a TM domain as long as it is lining the aqueous pore of the channel. However, since the amphipathic environment is created by the channel itself, the assembly (or partial formation) of its tertiary structure must presumably involve some co-operativity between TM domains or require chaperones to avoid exposure of the hydrophilic residues to the lipids' hydrocarbon chains. Insertion of multiple helices at once represents a challenge for integrases such as the translocon and associated components that will probably need to expand capacity to accommodate helical bundles. In the case of aquaporin 1 (AQP1) the second TM domain of six is fully translocated into the ER lumen and only adopts a transmembrane orientation after TM4 has been synthesized [111]. Membrane insertion of TM2 also requires a 180 degree flip of TM3 in the membrane presumably, but not necessarily, facilitated by the Sec61 translocon [112]. Such TM domain gymnastics during membrane protein assembly further highlight the potential flexibility of the Sec61 translocon in facilitating different insertion modalities.



**Figure 4.** Crystal structure of human aquaporin 10. Cartoon representation of the human aquaporin 10 (PDB:6F7H), (A) lateral view, (B) top view. Residues are colored based on their membrane insertion propensity: Residues with a low insertion propensity are shown in deep purple and pink. Residues with a mild and high insertion propensity are depicted in green. The side chain of polar and charged residues within the protein's pore are shown in a stick representation. In panel A the limits of the membrane are shown with a dotted line, data from the OPM database. (C) Apparent free energy for the insertion in biological membranes of the 20 natural amino acids. Data from [31].

## 9. Concluding Remarks

Tremendous progress has been made over the last two decades in the field of membrane protein insertion and folding, ranging from biochemical and structural data on the acquisition of secondary structure in nascent chains to the quantitative understanding of the energetic forces for sequence-dependent membrane insertion through the ER translocon and the structures of translocons engaged in nascent chain integration. This does not mean that we have a complete understanding of all aspects of membrane protein assembly and there are still challenges for the field. For example, it will be challenging to explore the dynamics and mechanisms of recruitment of translocon-associated proteins. Thus, it is difficult to envisage how proteins such as TRAM or other accessory components in the ER, which are present in sub-stoichiometric amounts can be present only at those translocons where its suggested chaperone function is needed [113]. Deducing detailed pathways of insertion of marginally hydrophobic TM domains such as TM2 of AQP1 remains another challenge. We hope that the development of new and existing technologies that can answer these remaining questions regarding membrane protein insertion and assembly is not too distant on the horizon. In recent years, new developments in cryo-EM sample preparation and data acquisition have been very fruitful in determining structures of integrase protein complexes [114] and, hopefully, they will be able to shed light on complete dynamics of the insertion process by determining intermediate states of membrane protein assembly through the aforementioned complexes. The increase in experimental data and computational power, in addition to the design of user-friendly interfaces, make MD simulations well positioned to play a very important role in the coming years in the understanding of membrane protein insertion. Simulations of ribosome–translocon complexes on the ns– $\mu$ s time scale have been possible for a decade [65,115], yet full insertion and maturation take place over seconds. Coarse-grained simulations make physiologically realistic time scales achievable, albeit at the cost of detail [67]. Balancing the need for atomic resolution with that for long time scales will be an ongoing effort, made more challenging by the increasing number of players characterized structurally in integrase complexes.

**Author Contributions:** P.W., B.G., J.C.G., L.M.-G. and I.M. jointly wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** Work in the authors' laboratories was supported by grants PID2020-119111GB-I00 from the Spanish Ministry of Science and Innovation (to I.M.), PROMETEU/2019/065 from Generalitat Valenciana (to I.M.) and R01-GM123169 from the US National Institutes of Health (to J.C.G.).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

EMC: ER membrane–protein complex; ER: Endoplasmic reticulum; MD: molecular dynamic; P-site: Peptidyl transferase site; SRP: Signal recognition particle; SS: Signal sequence; TA: Tail anchored; TM: Transmembrane.

## References

1. Ito, K.; Shimokawa-Chiba, N.; Chiba, S. Sec translocon has an insertase-like function in addition to polypeptide conduction through the channel. *F1000Research* **2019**, *8*, 2126. [[CrossRef](#)]
2. Farkas, A.; Bohnsack, K.E. Capture and delivery of tail-anchored proteins to the endoplasmic reticulum. *J. Cell Biol.* **2021**, *220*, e202105004. [[CrossRef](#)] [[PubMed](#)]
3. Mateja, A.; Keenan, R.J. A structural perspective on tail-anchored protein biogenesis by the GET pathway. *Curr. Opin. Struct. Biol.* **2018**, *51*, 195–202. [[CrossRef](#)] [[PubMed](#)]
4. Chitwood, P.J.; Hegde, R.S. The role of EMC during membrane protein biogenesis. *Trends Cell Biol.* **2019**, *29*, 371–384. [[CrossRef](#)] [[PubMed](#)]

5. Wideman, J.G. The ubiquitous and ancient ER membrane protein complex (EMC): Tether or not? *F1000Research* **2015**, *4*, 624. [[CrossRef](#)]
6. Rapoport, T.A.; Li, L.; Park, E. Structural and mechanistic insights into protein translocation. *Annu. Rev. Cell Dev. Biol.* **2017**, *33*, 369–390. [[CrossRef](#)]
7. Bano-Polo, M.; Martinez-Garay, C.A.; Grau, B.; Martinez-Gil, L.; Mingarro, I. Membrane insertion and topology of the translocon-associated protein (TRAP) gamma subunit. *Biochim. Biophys. Acta Biomembr.* **2017**, *1859*, 903–909. [[CrossRef](#)]
8. Pfeffer, S.; Dudek, J.; Schaffer, M.; Ng, B.G.; Albert, S.; Plitzko, J.M.; Baumeister, W.; Zimmermann, R.; Freeze, H.H.; Engel, B.D.; et al. Dissecting the molecular organization of the translocon-associated protein complex. *Nat. Commun.* **2017**, *8*, 14516. [[CrossRef](#)]
9. Sommer, N.; Junne, T.; Kalies, K.U.; Spiess, M.; Hartmann, E. TRAP assists membrane protein topogenesis at the mammalian ER membrane. *Biochim. Biophys. Acta* **2013**, *1833*, 3104–3111. [[CrossRef](#)]
10. Braunger, K.; Pfeffer, S.; Shrimal, S.; Gilmore, R.; Berninghausen, O.; Mandon, E.C.; Becker, T.; Forster, F.; Beckmann, R. Structural basis for coupling protein transport and N-glycosylation at the mammalian endoplasmic reticulum. *Science* **2018**, *360*, 215–219. [[CrossRef](#)]
11. Gemmer, M.; Forster, F. A clearer picture of the ER translocon complex. *J. Cell Sci.* **2020**, *133*, jcs231340. [[CrossRef](#)] [[PubMed](#)]
12. Chitwood, P.J.; Juszkievicz, S.; Guna, A.; Shao, S.; Hegde, R.S. EMC is required to initiate accurate membrane protein topogenesis. *Cell* **2018**, *175*, 1507–1519.e16. [[CrossRef](#)] [[PubMed](#)]
13. Van den Berg, B.; Clemons, W.M., Jr.; Collinson, I.; Modis, Y.; Hartmann, E.; Harrison, S.C.; Rapoport, T.A. X-ray structure of a protein-conducting channel. *Nature* **2004**, *427*, 36–44. [[CrossRef](#)] [[PubMed](#)]
14. Osborne, A.R.; Rapoport, T.A.; van den Berg, B. Protein translocation by the Sec61/SecY channel. *Annu. Rev. Cell Dev. Biol.* **2005**, *21*, 529–550. [[CrossRef](#)]
15. Voorhees, R.M.; Hegde, R.S. Toward a structural understanding of co-translational protein translocation. *Curr. Opin. Cell Biol.* **2016**, *41*, 91–99. [[CrossRef](#)]
16. Gumbart, J.; Schulten, K. The roles of pore ring and plug in the SecY protein-conducting channel. *J. Gen. Physiol.* **2008**, *132*, 709–719. [[CrossRef](#)]
17. Park, E.; Rapoport, T.A. Preserving the membrane barrier for small molecules during bacterial protein translocation. *Nature* **2011**, *473*, 239–242. [[CrossRef](#)] [[PubMed](#)]
18. Junne, T.; Kocik, L.; Spiess, M. The hydrophobic core of the Sec61 translocon defines the hydrophobicity threshold for membrane integration. *Mol. Biol. Cell* **2010**, *21*, 1662–1670. [[CrossRef](#)]
19. Martinez-Gil, L.; Sauri, A.; Marti-Renom, M.A.; Mingarro, I. Membrane protein integration into the endoplasmic reticulum. *FEBS J.* **2011**, *278*, 3846–3858. [[CrossRef](#)]
20. Johnson, A.E.; van Waes, M.A. The translocon: A dynamic gateway at the ER membrane. *Annu. Rev. Cell Dev. Biol.* **1999**, *15*, 799–842. [[CrossRef](#)]
21. Keenan, R.J.; Freymann, D.M.; Stroud, R.M.; Walter, P. The signal recognition particle. *Annu. Rev. Biochem.* **2001**, *70*, 755–775. [[CrossRef](#)] [[PubMed](#)]
22. Shao, S.; Hegde, R.S. Membrane protein insertion at the endoplasmic reticulum. *Annu. Rev. Cell Dev. Biol.* **2011**, *27*, 25–56. [[CrossRef](#)] [[PubMed](#)]
23. Frauenfeld, J.; Gumbart, J.; Sluis, E.O.; Funes, S.; Gartmann, M.; Beatrix, B.; Mielke, T.; Berninghausen, O.; Becker, T.; Schulten, K.; et al. Cryo-EM structure of the ribosome-SecYE complex in the membrane environment. *Nat. Struct. Mol. Biol.* **2011**, *18*, 614–621. [[CrossRef](#)] [[PubMed](#)]
24. Dowhan, W.; Vitrac, H.; Bogdanov, M. Lipid-assisted membrane protein folding and topogenesis. *Protein J.* **2019**, *38*, 274–288. [[CrossRef](#)]
25. Whitley, P.; Mingarro, I. Stitching proteins into membranes, not sew simple. *Biol. Chem.* **2014**, *395*, 1417–1424. [[CrossRef](#)] [[PubMed](#)]
26. Higy, M.; Junne, T.; Spiess, M. Topogenesis of membrane proteins at the endoplasmic reticulum. *Biochemistry* **2004**, *43*, 12716–12722. [[CrossRef](#)]
27. Sauri, A.; Tamborero, S.; Martinez-Gil, L.; Johnson, A.E.; Mingarro, I. Viral membrane protein topology is dictated by multiple determinants in its sequence. *J. Mol. Biol.* **2009**, *387*, 113–128. [[CrossRef](#)] [[PubMed](#)]
28. Voorhees, R.M.; Hegde, R.S. Structure of the Sec61 channel opened by a signal sequence. *Science* **2016**, *351*, 88–91. [[CrossRef](#)]
29. Mingarro, I.; Nilsson, I.; Whitley, P.; von Heijne, G. Different conformations of nascent polypeptides during translocation across the ER membrane. *BMC Cell Biol.* **2000**, *1*, 3. [[CrossRef](#)]
30. Whitley, P.; Nilsson, I.M.; von Heijne, G. A nascent secretory protein may traverse the ribosome/endoplasmic reticulum translocase complex as an extended chain. *J. Biol. Chem.* **1996**, *271*, 6241–6244. [[CrossRef](#)]
31. Hessa, T.; Kim, H.; Bihlmaier, K.; Lundin, C.; Boekel, J.; Andersson, H.; Nilsson, I.; White, S.H.; von Heijne, G. Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* **2005**, *433*, 377–381. [[CrossRef](#)] [[PubMed](#)]
32. Hessa, T.; Meindl-Beinker, N.M.; Bernsel, A.; Kim, H.; Sato, Y.; Lerch-Bader, M.; Nilsson, I.; White, S.H.; von Heijne, G. Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature* **2007**, *450*, 1026–1030. [[CrossRef](#)] [[PubMed](#)]
33. Devaraneni, P.K.; Conti, B.; Matsumura, Y.; Yang, Z.; Johnson, A.E.; Skach, W.R. Stepwise insertion and inversion of a type II signal anchor sequence in the ribosome-Sec61 translocon complex. *Cell* **2011**, *146*, 134–147. [[CrossRef](#)] [[PubMed](#)]

34. Ojemalm, K.; Watson, H.R.; Roboti, P.; Cross, B.C.; Warwicker, J.; von Heijne, G.; High, S. Positional editing of transmembrane domains during ion channel assembly. *J. Cell Sci.* **2013**, *126*, 464–472. [[CrossRef](#)]
35. Ota, K.; Sakaguchi, M.; von Heijne, G.; Hamasaki, N.; Mihara, K. Forced transmembrane orientation of hydrophilic polypeptide segments in multispanning membrane proteins. *Mol. Cell* **1998**, *2*, 495–503. [[CrossRef](#)]
36. Sadlish, H.; Pitonzo, D.; Johnson, A.E.; Skach, W.R. Sequential triage of transmembrane segments by Sec61alpha during biogenesis of a native multispanning membrane protein. *Nat. Struct. Mol. Biol.* **2005**, *12*, 870–878. [[CrossRef](#)]
37. Sauri, A.; McCormick, P.J.; Johnson, A.E.; Mingarro, I. Sec61alpha and TRAM are sequentially adjacent to a nascent viral membrane protein during its ER integration. *J. Mol. Biol.* **2007**, *366*, 366–374. [[CrossRef](#)]
38. Sauri, A.; Saksena, S.; Salgado, J.; Johnson, A.E.; Mingarro, I. Double-spanning plant viral movement protein integration into the endoplasmic reticulum membrane is signal recognition particle-dependent, translocon-mediated, and concerted. *J. Biol. Chem.* **2005**, *280*, 25907–25912. [[CrossRef](#)]
39. Cymer, F.; von Heijne, G.; White, S.H. Mechanisms of integral membrane protein insertion and folding. *J. Mol. Biol.* **2015**, *427*, 999–1022. [[CrossRef](#)]
40. White, S.H.; Ladokhin, A.S.; Jayasinghe, S.; Hristova, K. How membranes shape protein structure. *J. Biol. Chem.* **2001**, *276*, 32395–32398. [[CrossRef](#)]
41. Wimley, W.C.; Creamer, T.P.; White, S.H. Solvation energies of amino acid side chains and backbone in a family of host-guest pentapeptides. *Biochemistry* **1996**, *35*, 5109–5124. [[CrossRef](#)] [[PubMed](#)]
42. Wimley, W.C.; White, S.H. Designing transmembrane alpha-helices that insert spontaneously. *Biochemistry* **2000**, *39*, 4432–4442. [[CrossRef](#)] [[PubMed](#)]
43. Woolhead, C.A.; McCormick, P.J.; Johnson, A.E. Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. *Cell* **2004**, *116*, 725–736. [[CrossRef](#)]
44. Lu, J.; Deutsch, C. Secondary structure formation of a transmembrane segment in Kv channels. *Biochemistry* **2005**, *44*, 8230–8243. [[CrossRef](#)] [[PubMed](#)]
45. Lu, J.; Deutsch, C. Folding zones inside the ribosomal exit tunnel. *Nat. Struct. Mol. Biol.* **2005**, *12*, 1123–1129. [[CrossRef](#)]
46. Bhushan, S.; Gartmann, M.; Halic, M.; Armache, J.P.; Jarasch, A.; Mielke, T.; Berninghausen, O.; Wilson, D.N.; Beckmann, R. alpha-Helical nascent polypeptide chains visualized within distinct regions of the ribosomal exit tunnel. *Nat. Struct. Mol. Biol.* **2010**, *17*, 313–317. [[CrossRef](#)]
47. Bano-Polo, M.; Baeza-Delgado, C.; Tamborero, S.; Hazel, A.; Grau, B.; Nilsson, I.; Whitley, P.; Gumbart, J.C.; von Heijne, G.; Mingarro, I. Transmembrane but not soluble helices fold inside the ribosome tunnel. *Nat. Commun.* **2018**, *9*, 5246. [[CrossRef](#)]
48. Nilsson, O.B.; Hedman, R.; Marino, J.; Wickles, S.; Bischoff, L.; Johansson, M.; Muller-Lucks, A.; Trovato, F.; Puglisi, J.D.; O'Brien, E.P.; et al. Cotranslational protein folding inside the ribosome exit tunnel. *Cell Rep.* **2015**, *12*, 1533–1540. [[CrossRef](#)]
49. Wruck, F.; Tian, P.; Kudva, R.; Best, R.B.; von Heijne, G.; Tans, S.J.; Katranidis, A. The ribosome modulates folding inside the ribosomal exit tunnel. *Commun. Biol.* **2021**, *4*, 523. [[CrossRef](#)]
50. Tu, L.; Khanna, P.; Deutsch, C. Transmembrane segments form tertiary hairpins in the folding vestibule of the ribosome. *J. Mol. Biol.* **2014**, *426*, 185–198. [[CrossRef](#)]
51. Baars, L.; Wagner, S.; Wickstrom, D.; Klepsch, M.; Ytterberg, A.J.; van Wijk, K.J.; de Gier, J.W. Effects of SecE depletion on the inner and outer membrane proteomes of *Escherichia coli*. *J. Bacteriol.* **2008**, *190*, 3505–3525. [[CrossRef](#)]
52. Berhanu, S.; Ueda, T.; Kuruma, Y. Artificial photosynthetic cell producing energy for protein synthesis. *Nat. Commun.* **2019**, *10*, 1325. [[CrossRef](#)] [[PubMed](#)]
53. Matsubayashi, H.; Kuruma, Y.; Ueda, T. In vitro synthesis of the *E. coli* Sec translocon from DNA. *Angew Chem.* **2014**, *53*, 7535–7538. [[CrossRef](#)]
54. Eaglesfield, R.; Madsen, M.A.; Sanyal, S.; Reboud, J.; Amtmann, A. Cotranslational recruitment of ribosomes in protocells recreates a translocon-independent mechanism of proteorhodopsin biogenesis. *iScience* **2021**, *24*, 102429. [[CrossRef](#)] [[PubMed](#)]
55. Doan, K.N.; Grevel, A.; Martensson, C.U.; Ellenrieder, L.; Thornton, N.; Wenz, L.S.; Opalinski, L.; Guiard, B.; Pfanner, N.; Becker, T. The mitochondrial import complex MIM functions as main translocase for alpha-helical outer membrane proteins. *Cell Rep.* **2020**, *31*, 107567. [[CrossRef](#)] [[PubMed](#)]
56. Glick, B.S.; Von Heijne, G. *Saccharomyces cerevisiae* mitochondria lack a bacterial-type sec machinery. *Protein Sci.* **1996**, *5*, 2651–2652. [[CrossRef](#)] [[PubMed](#)]
57. Jia, L.; Dienhart, M.; Schramp, M.; McCauley, M.; Hell, K.; Stuart, R.A. Yeast Oxa1 interacts with mitochondrial ribosomes: The importance of the C-terminal region of Oxa1. *EMBO J.* **2003**, *22*, 6438–6447. [[CrossRef](#)] [[PubMed](#)]
58. Mulikjanian, A.Y.; Galperin, M.Y.; Koonin, E.V. Co-evolution of primordial membranes and membrane proteins. *Trends Biochem. Sci.* **2009**, *34*, 206–215. [[CrossRef](#)]
59. Cross, B.C.; High, S. Dissecting the physiological role of selective transmembrane-segment retention at the ER translocon. *J. Cell Sci.* **2009**, *122*, 1768–1777. [[CrossRef](#)]
60. Park, E.; Rapoport, T.A. Mechanisms of Sec61/SecY-mediated protein translocation across membranes. *Annu. Rev. Biophys.* **2012**, *41*, 21–40. [[CrossRef](#)]
61. Voss, N.R.; Gerstein, M.; Steitz, T.A.; Moore, P.B. The geometry of the ribosomal polypeptide exit tunnel. *J. Mol. Biol.* **2006**, *360*, 893–906. [[CrossRef](#)] [[PubMed](#)]

62. Gumbart, J.; Chipot, C.; Schulten, K. Free energy of nascent-chain folding in the translocon. *J. Am. Chem. Soc.* **2011**, *133*, 7602–7607. [[CrossRef](#)] [[PubMed](#)]
63. Egea, P.F.; Stroud, R.M. Lateral opening of a translocon upon entry of protein suggests the mechanism of insertion into membranes. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 17182–17187. [[CrossRef](#)] [[PubMed](#)]
64. Gogala, M.; Becker, T.; Beatrix, B.; Armache, J.P.; Barrio-Garcia, C.; Berninghausen, O.; Beckmann, R. Structures of the Sec61 complex engaged in nascent peptide translocation or membrane insertion. *Nature* **2014**, *506*, 107–110. [[CrossRef](#)] [[PubMed](#)]
65. Gumbart, J.C.; Teo, I.; Roux, B.; Schulten, K. Reconciling the roles of kinetic and thermodynamic factors in membrane-protein insertion. *J. Am. Chem. Soc.* **2013**, *135*, 2291–2297. [[CrossRef](#)] [[PubMed](#)]
66. Mackinnon, A.L.; Paavilainen, V.O.; Sharma, A.; Hegde, R.S.; Taunton, J. An allosteric Sec61 inhibitor traps nascent transmembrane helices at the lateral gate. *eLife* **2014**, *3*, e01483. [[CrossRef](#)]
67. Niesen, M.J.M.; Zimmer, M.H.; Miller, T.F., 3rd. Dynamics of co-translational membrane protein integration and translocation via the sec translocon. *J. Am. Chem. Soc.* **2020**, *142*, 5449–5460. [[CrossRef](#)]
68. Park, E.; Menetret, J.F.; Gumbart, J.C.; Ludtke, S.J.; Li, W.; Whynot, A.; Rapoport, T.A.; Akey, C.W. Structure of the SecY channel during initiation of protein translocation. *Nature* **2014**, *506*, 102–106. [[CrossRef](#)]
69. Mercier, E.; Wang, X.; Maiti, M.; Wintermeyer, W.; Rodnina, M.V. Lateral gate dynamics of the bacterial translocon during cotranslational membrane protein insertion. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2100474118. [[CrossRef](#)]
70. Ge, Y.; Draycheva, A.; Bornemann, T.; Rodnina, M.V.; Wintermeyer, W. Lateral opening of the bacterial translocon on ribosome binding and signal peptide insertion. *Nat. Commun.* **2014**, *5*, 5263. [[CrossRef](#)]
71. Pfeffer, S.; Burbaum, L.; Unverdorben, P.; Pech, M.; Chen, Y.; Zimmermann, R.; Beckmann, R.; Forster, F. Structure of the native Sec61 protein-conducting channel. *Nat. Commun.* **2015**, *6*, 8403. [[CrossRef](#)] [[PubMed](#)]
72. Shurtleff, M.J.; Itzhak, D.N.; Hussmann, J.A.; Schirle Oakdale, N.T.; Costa, E.A.; Jonikas, M.; Weibezahn, J.; Popova, K.D.; Jan, C.H.; Sinitcyn, P.; et al. The ER membrane protein complex interacts cotranslationally to enable biogenesis of multipass membrane proteins. *eLife* **2018**, *7*, e37018. [[CrossRef](#)] [[PubMed](#)]
73. Bowie, J.U. Solving the membrane protein folding problem. *Nature* **2005**, *438*, 581–589. [[CrossRef](#)] [[PubMed](#)]
74. Nicolaus, F.; Metola, A.; Mermans, D.; Liljenstrom, A.; Krc, A.; Abdullahi, S.M.; Zimmer, M.; Miller Iii, T.F.; von Heijne, G. Residue-by-residue analysis of cotranslational membrane protein integration in vivo. *eLife* **2021**, *10*, e64302. [[CrossRef](#)] [[PubMed](#)]
75. Gumbart, J.C.; Chipot, C. Decrypting protein insertion through the translocon with free-energy calculations. *Biochim. Biophys. Acta* **2016**, *1858*, 1663–1671. [[CrossRef](#)]
76. Zhang, B.; Miller, T.F., 3rd. Hydrophobically stabilized open state for the lateral gate of the Sec translocon. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 5399–5404. [[CrossRef](#)]
77. Duong, F.; Wickner, W. Sec-dependent membrane protein biogenesis: SecYEG, preprotein hydrophobicity and translocation kinetics control the stop-transfer function. *EMBO J.* **1998**, *17*, 696–705. [[CrossRef](#)]
78. Trueman, S.F.; Mandon, E.C.; Gilmore, R. Translocation channel gating kinetics balances protein translocation efficiency with signal sequence recognition fidelity. *Mol. Biol. Cell* **2011**, *22*, 2983–2993. [[CrossRef](#)]
79. Rapp, M.; Granseth, E.; Seppala, S.; von Heijne, G. Identification and evolution of dual-topology membrane proteins. *Nat. Struct. Mol. Biol.* **2006**, *13*, 112–116. [[CrossRef](#)]
80. Seppala, S.; Slusky, J.S.; Lloris-Garcera, P.; Rapp, M.; von Heijne, G. Control of membrane protein topology by a single C-terminal residue. *Science* **2010**, *328*, 1698–1700. [[CrossRef](#)]
81. Seurig, M.; Ek, M.; von Heijne, G.; Fluman, N. Dynamic membrane topology in an unassembled membrane protein. *Nat. Chem. Biol.* **2019**, *15*, 945–948. [[CrossRef](#)] [[PubMed](#)]
82. Van Lehn, R.C.; Zhang, B.; Miller, T.F., 3rd. Regulation of multispinning membrane protein topology via post-translational annealing. *eLife* **2015**, *4*, e08697. [[CrossRef](#)]
83. Baeza-Delgado, C.; von Heijne, G.; Marti-Renom, M.A.; Mingarro, I. Biological insertion of computationally designed short transmembrane segments. *Sci. Rep.* **2016**, *6*, 23397. [[CrossRef](#)] [[PubMed](#)]
84. Andersen, O.S.; Koeppe, R.E., 2nd. Bilayer thickness and membrane protein function: An energetic perspective. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 107–130. [[CrossRef](#)] [[PubMed](#)]
85. Mouritsen, O.G.; Bloom, M. Mattress model of lipid-protein interactions in membranes. *Biophys. J.* **1984**, *46*, 141–153. [[CrossRef](#)]
86. Baeza-Delgado, C.; Marti-Renom, M.A.; Mingarro, I. Structure-based statistical analysis of transmembrane helices. *Eur. Biophys. J.* **2013**, *42*, 199–207. [[CrossRef](#)] [[PubMed](#)]
87. Chen, A.; Moy, V.T. Cross-linking of cell surface receptors enhances cooperativity of molecular adhesion. *Biophys. J.* **2000**, *78*, 2814–2820. [[CrossRef](#)]
88. Mitra, K.; Ubarretxena-Belandia, I.; Taguchi, T.; Warren, G.; Engelman, D.M. Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4083–4088. [[CrossRef](#)]
89. Rawicz, W.; Olbrich, K.C.; McIntosh, T.; Needham, D.; Evans, E. Effect of chain length and unsaturation on elasticity of lipid bilayers. *Biophys. J.* **2000**, *79*, 328–339. [[CrossRef](#)]
90. Killian, J.A. Hydrophobic mismatch between proteins and lipids in membranes. *Biochim. Biophys. Acta* **1998**, *1376*, 401–415. [[CrossRef](#)]
91. Krishnakumar, S.S.; London, E. Effect of sequence hydrophobicity and bilayer width upon the minimum length required for the formation of transmembrane helices in membranes. *J. Mol. Biol.* **2007**, *374*, 671–687. [[CrossRef](#)]

92. Jaud, S.; Fernandez-Vidal, M.; Nilsson, I.; Meindl-Beinker, N.M.; Hubner, N.C.; Tobias, D.J.; von Heijne, G.; White, S.H. Insertion of short transmembrane helices by the Sec61 translocon. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11588–11593. [[CrossRef](#)]
93. Ulmschneider, M.B.; Ulmschneider, J.P.; Schiller, N.; Wallace, B.A.; von Heijne, G.; White, S.H. Spontaneous transmembrane helix insertion thermodynamically mimics translocon-guided insertion. *Nat. Commun.* **2014**, *5*, 4863. [[CrossRef](#)] [[PubMed](#)]
94. Grau, B.; Javanainen, M.; Garcia-Murria, M.J.; Kulig, W.; Vattulainen, I.; Mingarro, I.; Martinez-Gil, L. The role of hydrophobic matching on transmembrane helix packing in cells. *Cell Stress* **2017**, *1*, 90–106. [[CrossRef](#)] [[PubMed](#)]
95. Bano-Polo, M.; Martinez-Gil, L.; Wallner, B.; Nieva, J.L.; Elofsson, A.; Mingarro, I. Charge pair interactions in transmembrane helices and turn propensity of the connecting sequence promote helical hairpin insertion. *J. Mol. Biol.* **2013**, *425*, 830–840. [[CrossRef](#)] [[PubMed](#)]
96. Monne, M.; Hermansson, M.; von Heijne, G. A turn propensity scale for transmembrane helices. *J. Mol. Biol.* **1999**, *288*, 141–145. [[CrossRef](#)] [[PubMed](#)]
97. Bano-Polo, M.; Baeza-Delgado, C.; Orzaez, M.; Marti-Renom, M.A.; Abad, C.; Mingarro, I. Polar/Ionizable residues in transmembrane segments: Effects on helix-helix packing. *PLoS ONE* **2012**, *7*, e44263. [[CrossRef](#)]
98. MacCallum, J.L.; Bennett, W.F.; Tieleman, D.P. Distribution of amino acids in a lipid bilayer from computer simulations. *Biophys. J.* **2008**, *94*, 3393–3404. [[CrossRef](#)] [[PubMed](#)]
99. Martinez-Gil, L.; Perez-Gil, J.; Mingarro, I. The surfactant peptide KL4 sequence is inserted with a transmembrane orientation into the endoplasmic reticulum membrane. *Biophys. J.* **2008**, *95*, L36–L38. [[CrossRef](#)]
100. Garcia-Saez, A.J.; Coraiola, M.; Dalla Serra, M.; Mingarro, I.; Menestrina, G.; Salgado, J. Peptides derived from apoptotic Bax and Bid reproduce the poration activity of the parent full-length proteins. *Biophys. J.* **2005**, *88*, 3976–3990. [[CrossRef](#)]
101. Garcia-Saez, A.J.; Coraiola, M.; Serra, M.D.; Mingarro, I.; Muller, P.; Salgado, J. Peptides corresponding to helices 5 and 6 of Bax can independently form large lipid pores. *FEBS J.* **2006**, *273*, 971–981. [[CrossRef](#)]
102. Koneru, J.K.; Prakashchand, D.D.; Dube, N.; Ghosh, P.; Mondal, J. Spontaneous transmembrane pore formation by short-chain synthetic peptide. *Biophys. J.* **2021**, *120*, 4557–4574. [[CrossRef](#)]
103. Choma, C.; Gratkowski, H.; Lear, J.D.; DeGrado, W.F. Asparagine-mediated self-association of a model transmembrane helix. *Nat. Struct. Biol.* **2000**, *7*, 161–166. [[CrossRef](#)]
104. Zhou, F.X.; Cocco, M.J.; Russ, W.P.; Brunger, A.T.; Engelman, D.M. Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nat. Struct. Biol.* **2000**, *7*, 154–160. [[CrossRef](#)] [[PubMed](#)]
105. Zhou, F.X.; Merianos, H.J.; Brunger, A.T.; Engelman, D.M. Polar residues drive association of poly-leucine transmembrane helices. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 2250–2255. [[CrossRef](#)]
106. Almada, J.C.; Bortolotti, A.; Ruyschaert, J.M.; de Mendoza, D.; Inda, M.E.; Cybulski, L.E. Interhelical H-bonds modulate the activity of a polytopic transmembrane kinase. *Biomolecules* **2021**, *11*, 938. [[CrossRef](#)] [[PubMed](#)]
107. Li, R.; Mitra, N.; Gratkowski, H.; Vilaire, G.; Litvinov, R.; Nagasami, C.; Weisel, J.W.; Lear, J.D.; DeGrado, W.F.; Bennett, J.S. Activation of integrin  $\alpha$ IIb $\beta$ 3 by modulation of transmembrane helix associations. *Science* **2003**, *300*, 795–798. [[CrossRef](#)] [[PubMed](#)]
108. Zhang, L.; Sato, Y.; Hessa, T.; von Heijne, G.; Lee, J.K.; Kodama, I.; Sakaguchi, M.; Uozumi, N. Contribution of hydrophobic and electrostatic interactions to the membrane integration of the Shaker K<sup>+</sup> channel voltage sensor domain. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8263–8268. [[CrossRef](#)]
109. Hou, B.; Lin, P.J.; Johnson, A.E. Membrane protein TM segments are retained at the translocon during integration until the nascent chain cues FRET-detected release into bulk lipid. *Mol. Cell* **2012**, *48*, 398–408. [[CrossRef](#)]
110. Russ, W.P.; Engelman, D.M. The GxxxG motif: A framework for transmembrane helix-helix association. *J. Mol. Biol.* **2000**, *296*, 911–919. [[CrossRef](#)]
111. Lu, Y.; Turnbull, I.R.; Bragin, A.; Carveth, K.; Verkman, A.S.; Skach, W.R. Reorientation of aquaporin-1 topology during maturation in the endoplasmic reticulum. *Mol. Biol. Cell* **2000**, *11*, 2973–2985. [[CrossRef](#)] [[PubMed](#)]
112. Skach, W.R. Cellular mechanisms of membrane protein folding. *Nat. Struct. Mol. Biol.* **2009**, *16*, 606–612. [[CrossRef](#)] [[PubMed](#)]
113. Tamborero, S.; Vilar, M.; Martinez-Gil, L.; Johnson, A.E.; Mingarro, I. Membrane insertion and topology of the translocating chain-associating membrane protein (TRAM). *J. Mol. Biol.* **2011**, *406*, 571–582. [[CrossRef](#)] [[PubMed](#)]
114. Pleiner, T.; Tomaleri, G.P.; Januszzyk, K.; Inglis, A.J.; Hazu, M.; Voorhees, R.M. Structural basis for membrane insertion by the human ER membrane protein complex. *Science* **2020**, *369*, 433–436. [[CrossRef](#)] [[PubMed](#)]
115. Gumbart, J.; Trabuco, L.G.; Schreiner, E.; Villa, E.; Schulten, K. Regulation of the protein-conducting channel by a bound ribosome. *Structure* **2009**, *17*, 1453–1464. [[CrossRef](#)] [[PubMed](#)]