

## A Current Perspective on Membrane Protein Folding

Sankaranarayanan Srinivasan\*

Verna and Marrs McLean, Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030-3498, USA

Membrane protein folding to a fully functional molecule involves two stages, and the very few experimental systems where these proteins fold in a thermodynamically reversible manner have restricted our understanding of this process, in comparison to soluble proteins. Here, we briefly look at our current understanding of membrane protein folding.

Protein folding is an extraordinary process, in which a polypeptide chain folds into a well defined three-dimensional structure having specific functional properties. Kinetic and thermodynamic studies on the folding of water-soluble proteins have significantly increased our understanding in this regard. Advances in theoretical and computational methods have also helped in predicting the native structure of small proteins, accurately. On the other hand, our knowledge regarding folding of membrane proteins is still in its early stage. Membrane protein [MP] is a molecule attached to or embedded in the membrane of a cell, or its organelles. Based on this, they are classified as peripheral or integral membrane proteins [1]. To satisfy the hydrogen bonding criterion of a polypeptide chain, transmembrane domains of MPs either form  $\alpha$ -helices or  $\beta$ -strands. In  $\alpha$ -helices, hydrogen bonds are formed within the same segment.  $\beta$ -strand form hydrogen bonds between adjacent segments, and hence form cylindrical  $\beta$ -barrels to fulfill all their hydrogen bonding requirements. As a result, they are required to traverse the membrane once or multiple times with a string of ~20, predominantly hydrophobic amino acids comprising the transmembrane [TM] portion of the protein.  $\alpha$ -helical MPs that span the membrane with a single TM segment are classified as bitopic proteins [1].

Membrane proteins pose unique experimental challenges. First, their over expression which is a pre-requisite for any meaningful structural characterization, could be toxic to the host. Second is the proper folding in the host membrane, which could be rate limiting. Third, functionally important post-translational modifications might not be reproduced in a heterologous expression system. Fourth, choosing the optimal detergent to extract the protein from the membrane and maintaining it in its native conformation is not straightforward [2]. If all the above challenges are overcome, subsequent crystallization trials or applying nuclear magnetic resonance [NMR] methods require further optimization in choosing the right detergent for structural studies. Until recently, eukaryotic membrane proteins whose three-dimensional structure had been determined were purified from uniquely rich sources. For example, rhodopsin was from eyes [3], or aquaporin from fiber cells of the bovine eye lens [4]. Thus, there is an immediate need for developing novel expression systems for eukaryotic membrane proteins.

At present, structures of 359 unique membrane proteins are available [5]. They have provided a much clearer view of the structural diversity exhibited by membrane proteins. Some of these structures have refuted the long standing belief of TM helices being packed parallel to the membrane normal and spanning the complete length of the membrane. More than 60% of all TM helix packing angles are in the range of  $0^\circ$  to  $40^\circ$  [6]. It has also been observed that 1 in 20 TM helices are half TM helices [7], with about 60% of TM helices accommodating significant bends or other distortions [8]. Loops connecting TM helices

are also not greatly restricted by the membrane. Thus, it can be said that like soluble proteins, membrane proteins have also adapted to fold in such a way that their chains satisfy the functional requirements, albeit within strict confines of the lipid bilayer. Solving the protein-folding problem has been a long standing challenge and the special case of membrane proteins has garnered increasing interest from researchers.

A two-stage model of membrane protein folding was first proposed by Popot and Engelman [9,10], for  $\alpha$ -helical membrane proteins. The first stage involves formation and insertion of helices into the membrane. This step is directed and catalyzed by the translocon complex. The way a protein sequence interacts with the translocon to direct insertion and topology would be the key in understanding this stage. In the second crucial stage, the tertiary and quaternary structures are formed. Here, the helices formed in the first stage assemble and reorient with additional insertion of re-entrant portions of the chain. Subunit oligomerization also happens in this stage [9,10]. Detailed studies to determine specific amino acid preferences for different portions of the bilayer, energetics of interactions within the protein and with the bilayer would provide us a much clearer picture of this stage.

Helical membrane protein sequences are characterized by hydrophobic stretches of approximately 20 residues [9]. They are long enough to span the hydrocarbon core of a typical bilayer in helical conformation. This feature indicates that thermodynamic partitioning between water and the bilayer is important in membrane insertion and maintenance in the bilayer. The high cost of interfacial partitioning of the peptide bond, 1.2 kcal/mol, drives the formation of secondary structures [11].  $\alpha$ -helical formation reduces the cost of interfacial partitioning by 0.4 kcal/mol per residue [12], and 0.5 kcal/mol for a  $\beta$ -sheet [13]. The free energy cost of disrupting hydrogen bonds in the membrane is 4 kcal/mol, on a per residue basis [14]. Hence, it is thermodynamically favorable for polypeptides to satisfy their hydrogen bonding requirements at the interface, before inserting into the membrane. Small such changes can aggregate to be large, thus favoring the folded state more than the unfolded state. In spite of this, it is the side-chain hydrophobicity that drives the equilibrium in favor of insertion into the bilayer. Thus, there exists a threshold of side chain hydrophobicity for a segment to insert in the bilayer [15].

The biological process of nascent chain insertion into the bilayer is catalyzed by the SecY complex in bacteria and the Sec61 complex in eukaryotes [16]. These complexes perform a variety of functions.

**\*Corresponding author:** Sankaranarayanan Srinivasan, Verna and Marrs McLean, Department of Biochemistry and Molecular Biology, Mail Stop BCM125, One Baylor Plaza, Baylor College of Medicine, Houston, TX 77030-3498, USA, Tel: 713-798-6564; Fax: 713-798-8516; E-mail: [ssriniva@bcm.edu](mailto:ssriniva@bcm.edu)

Received October 20, 2012; Accepted October 20, 2012; Published October 24, 2012

**Citation:** Srinivasan S (2012) A Current Perspective on Membrane Protein Folding. Biochem Anal Biochem 1:e123. doi:10.4172/2161-1009.1000e123

**Copyright:** © 2012 Srinivasan S, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

First, they ensure no leakage of ions or small molecules during the passage of the nascent polypeptide chain. Second, they decide if they need to do reverse orientation of the emerging segment of the nascent polypeptide. Third, they also decide if the emerging segment needs to be passed to the aqueous or the membrane phase. Fourth, they also need to ensure that no leakage of ions or small molecules occur, when it opens laterally to release a segment into the membrane. Biochemical studies in conjunction with the structure of the SecY/SecE complex have enabled to propose a general model for protein translocation [16, 17]. The translocon complex consists of three subunits  $\alpha$ ,  $\beta$  and  $\gamma$ . The largest subunit  $\alpha$  is composed of 10 TM helices surrounding the likely translocation pore, an additional TM helix is contributed by the  $\beta$ - subunit and the  $\gamma$  subunit closes the  $\alpha$ -subunit on one end. The subunit contributes the helical plug, which slides out to allow the nascent polypeptide to pass. Hydrophobic residues lining the pore sample the nascent polypeptide and ascribe volition to a protein to let it pass through or let it into the membrane. TM helices enter the membrane by opening the channel laterally to the side opposite to the  $\gamma$ -subunit [16,17].

Topology of membrane inserted segments is determined by interactions between the translocon and primary sequence of the nascent polypeptide chain. It has been now well established that membrane proteins follow the positive-inside rule [18]. This means that the cytoplasmic side of membrane proteins tends to be positively charged. Some reasons for this preference might be the electrostatic potential differences between the two compartments and differences in lipid composition between the two membrane leaflets. Presence of positively charged residues in the plug region of the translocon also might cause a strong repulsion, flipping the orientation of the chain [19]. Other factors such as synthesis rate and the length of the hydrophobic segment also contribute in making the topology decision. Like the topology decision, the decision to insert a certain segment of the nascent polypeptide into the membrane is made through interaction with the translocon. Studies have suggested that frequent opening of the lateral gate in the translocon allows the nascent polypeptide to sample both the aqueous and membrane phases [20]. If the rate of synthesis is slow relative to the sampling process, equilibrium is established and a bilayer favoring configuration enables insertion of the segment in the membrane [20]. The insertion probability is also not only determined by the composition of the residues, but also by positioning of certain polar residues. Studies have shown that, it is more favorable to place an arginine residue at the edge of the hydrocarbon core than at the centre [21].

In soluble proteins, the hydrophobic effect is a primary driving force of folding, and for the association of hydrophobic patches on protein surfaces. Membrane proteins are required thermodynamically to interact with each other to form the final folded structure. This increases the entropy of the total system by reducing lipid-exposed area of the protein. In other words, release of lipid molecules into the lipid pool increases their entropy [22]. The interactions involved in the association step are mostly van der Waals forces, hydrogen bonding and electrostatic interactions [6]. The helices make closely packed left-handed and right-handed contacts with the left handed configuration, at approximately 20°C being the most favored one [6]. Similar to structural studies of membrane proteins, the folding studies have been limited largely due to the difficulty in devising experimental systems, where they can fold in a thermodynamically reversible manner.

A major difference to be considered while studying the folding of membrane proteins is that the unfolded state retains considerable

secondary structure [23]. Hence, most methods used have focused on the second stage of the folding process. Another important difference is the presence of detergents which are required to keep the proteins soluble and sometimes, they have influenced the folding mechanism. Conventionally, techniques that monitor spectroscopic or physical properties that differ between the folded and unfolded state, follow folding of a protein. Since the magnitude of these differences are much smaller in membrane proteins, other methods are used. Specifically, analytical ultracentrifugation, fluorescence resonance energy transfer [FRET] and gel electrophoresis have been widely used. Some of the less frequently used methods are electron paramagnetic resonance, Fourier transform infrared spectroscopy and small angle X-ray scattering [24]. Very recently, pulse-proteolysis has been employed to study protein-unfolding of bacterio-rhodopsin [25].

Formation of tertiary and quaternary structures in membrane proteins occurs in the second stage of the two-stage membrane protein-folding model. This stage largely involves TM helices interacting with each other [9]. Interactions in the extracellular, cytoplasmic loop regions or with other molecules also contribute to formation of the final folded structure. Association of bitopic membrane proteins is significant in the formation of a number of signaling complexes. Similarly, association of TM helices in polytopic membrane proteins led to the formation of functional oligomeric complexes. Thus, association in membrane protein assembly can transform a simple membrane anchor into a biologically active complex. Driving forces behind formation of quaternary structures are primarily of two types [26]. The first force is the covalent bonding between monomers, and the second is non-covalent interactions enabling specific association. Covalent bonds are mostly disulfide bonds between monomers, which are largely formed in the extracellular oxidizing environment. The low dielectric constant of the bilayer prevents formation of a disulfide bond, as it is highly unfavorable for the deprotonation of the thiol group. Hence, membrane protein association mediated by TM regions is largely of non-covalent nature. Van der Waals interactions, salt bridges and interhelical hydrogen bonding are some of the non-covalent forces involved in membrane protein association [26].

A majority of the understanding of TM helical interaction has come from studies on glycoporphin A (GpA). GpA is a single TM helix dimer expressed on the surface of erythrocytes. Mutagenesis and NMR studies have revealed the dimeric interface of the TM domain to be a characteristic right handed motif GxxxG [27]. This motif has also been found to mediate helix association for several other TM domain sequences like that of the erbB receptors, integrins, BH3 only apoptotic proteins, F0F1-ATP synthase and GPCRs. It was not surprising to find this motif to be statistically over-represented in single pass TM helices with pairwise interactions [27]. This observation can largely be attributed to the reason that glycine promotes proximity between helices by minimizing the steric hindrances of the helical backbones and promoting van der Waals interactions. There is also a potential for hydrogen bonding to enhance the stability of interactions between the TM helices [27]. Proline, a conventional helix breaker in soluble proteins is typically accommodated at the N and C-terminal ends of the TM helix. The increased backbone flexibility imparted by the proline residue is important for hinge-bending motions. They are believed to stabilize helix-helix packing. Positioning of a serine or threonine at position  $i$  and proline at position  $i+4$  enables to free the backbone carbonyl oxygen atom for interhelical hydrogen bonding. This hydrogen bonding is believed to provide stabilization energy for association [28]. The above mentioned examples and association motifs clearly indicate the role of many transmembrane domains to

membrane protein association.

Recent reviews by Bowie [29], Booth and Curnow [30] have outlined the points discussed in this article, and would serve as excellent articles for a more detailed outlook. Studies described in here have enabled us to create our current view on membrane protein folding, though it is in its primitive phase. However, the latest rise in the number of high-resolution structures and significant technical advances in membrane protein purification would herald an exciting period in membrane protein folding investigations.

## References

- Chou KC, Elrod DW (1999) Prediction of membrane protein types and subcellular locations. *PROTEINS: Struct Funct Genet* 34:137-153.
- Dobrovetsky E, Lu ML, Andorn-Broza R, Khutoreskaya G, Bray JE, et al. (2005) High-throughput production of prokaryotic membrane proteins. *J Struct Funct Genomics* 6: 33-50.
- Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, et al. (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* 289: 739-745.
- Walz T, Hirai T, Murata K, Heymann JB, Mitsuoka K, et al. (1997) The three-dimensional structure of aquaporin-1. *Nature* 387: 624-627.
- [http://blanco.biomol.uci.edu/Membrane\\_Proteins\\_xtal.html](http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html).
- Bowie JU (1997) Helix packing in membrane proteins. *J Mol Biol* 272: 780-789.
- Cuthbertson JM, Doyle DA, Sansom MS (2005) Transmembrane helix prediction: a comparative evaluation and analysis. *Protein Eng Des Sel* 18: 295-308.
- Yohannan S, Faham S, Yang D, Whitelegge JP, Bowie JU (2004) The evolution of transmembrane helix kinks and the structural diversity of G protein-coupled receptors. *Proc Natl Acad Sci U S A* 101: 959-963.
- Popot JL, Engelman DM (1990) Membrane protein folding and oligomerization: the two-stage model. *Biochemistry* 29: 4031-4037.
- Popot JL, Engelman DM (2000) Helical membrane protein folding, stability, and evolution. *Annu Rev Biochem* 69: 881-922.
- Wimley WC, White SH (1996) Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat Struct Biol* 3: 842-848.
- Ladokhin AS, White SH (1999) Folding of amphipathic alpha-helices on membranes: energetics of helix formation by melittin. *J Mol Biol* 285: 1363-1369.
- Wimley WC, Hristova K, Ladokhin AS, Silvestro L, Axelsen PH, et al. (1998) Folding of beta-sheet membrane proteins: a hydrophobic hexapeptide model. *J Mol Biol* 277:1091-1110.
- Ben-Tal N, Ben-Shaul A, Nicholls A, Honig B (1996) Free-energy determinants of alpha-helix insertion into lipid bilayers. *Biophys J* 70: 1803-1812.
- Liu LP, Li SC, Goto NK, Deber CM (1996) Threshold hydrophobicity dictates helical conformations of peptides in membrane environments. *Biopolymers* 39: 465-470.
- Van den Berg B, Clemons WM Jr, Collinson I, Modis Y, Hartmann E, et al. (2004) X-ray structure of a protein-conducting channel. *Nature* 427: 36-44.
- Rapoport TA, Goder V, Heinrich SU, Matlack KE (2004) Membrane-protein integration and the role of the translocation channel. *Trends Cell Biol* 14: 568-575.
- Heijne G (1986) The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *EMBO J* 5: 3021-3027.
- Goder V, Junne T, Spiess M (2004) Sec61p contributes to signal sequence orientation according to the positive-inside rule. *Mol Biol Cell* 15: 1470-1478.
- Heinrich SU, Mothes W, Brunner J, Rapoport TA (2000) The Sec61p complex mediates the integration of a membrane protein by allowing lipid partitioning of the transmembrane domain. *Cell* 102: 233-244.
- Chamberlain AK, Lee Y, Kim S, Bowie JU (2004) Snorkeling preferences foster an amino acid composition bias in transmembrane helices. *J Mol Biol* 339: 471-479.
- Helms V (2002) Attraction within the membrane. *EMBO Reports* 3:1133-1138.
- White SH, Ladokhin AS, Jayasinghe S, Hristova K (2001) How membranes shape protein structure. *J Biol Chem* 276: 32395-32398.
- DeGrado WF, Gratkowski H, Lear JD (2003) How do helix-helix interactions help determine the folds of membrane proteins? Perspectives from the study of homo-oligomeric helical bundles. *Protein Sci* 12: 647-665.
- Schlebach JP, Kim MS, Joh NH, Bowie JU, Park C (2011) Probing membrane protein unfolding with pulse proteolysis. *J Mol Biol* 406: 545-551.
- Arkin IT (2002) Structural aspects of oligomerization taking place between the transmembrane alpha-helices of bitopic membrane proteins. *Biochim Biophys Acta* 1565: 347-363.
- Senes A, Engel DE, DeGrado WF (2004) Folding of helical membrane proteins: the role of polar, GXXXG-like and proline motifs. *Curr Opin Struct Biol* 14: 465-479.
- Orzáez M, Salgado J, Giménez-Giner A, Pérez-Payá E, Mingarro I (2004) Influence of proline residues in transmembrane helix packing. *J Mol Biol* 335: 631-640.
- Bowie JU (2005) Solving the membrane protein folding problem. *Nature* 438: 581-589.
- Booth PJ, Curnow P (2009) Folding scene investigation: membrane proteins. *Curr Opin Struct Biol* 19: 8-13.