REVIEWS

Life at the border: Adaptation of proteins to anisotropic membrane environment

Irina D. Pogozheva, Henry I. Mosberg, and Andrei L. Lomize*

Department of Medicinal Chemistry, College of Pharmacy, University of Michigan, Ann Arbor, Michigan 48109-1065

Received 2 June 2014; Accepted 18 June 2014
DOI: 10.1002/pro.2508
Published online 20 June 2014 proteinscience.org

Abstract: This review discusses main features of transmembrane (TM) proteins which distinguish them from water-soluble proteins and allow their adaptation to the anisotropic membrane environment. We overview the structural limitations on membrane protein architecture, spatial arrangement of proteins in membranes and their intrinsic hydrophobic thickness, co-translational and post-translational folding and insertion into lipid bilayers, topogenesis, high propensity to form oligomers, and large-scale conformational transitions during membrane insertion and transport function. Special attention is paid to the polarity of TM protein surfaces described by profiles of dipolarity/polarizability and hydrogen-bonding capacity parameters that match polarity of the lipid environment. Analysis of distributions of Trp resides on surfaces of TM proteins from different biological membranes indicates that interfacial membrane regions with preferential accumulation of Trp indole rings correspond to the outer part of the lipid acyl chain region—between double bonds and carbonyl groups of lipids. These "midpolar" regions are not always symmetric in proteins from natural membranes. We also examined the hydrophobic effect that drives insertion of proteins into lipid bilayer and different free energy contributions to TM protein stability, including attractive van der Waals forces and hydrogen bonds, side-chain conformational entropy, the hydrophobic mismatch, membrane deformations, and specific protein–lipid binding.

Keywords: protein–lipid interactions; hydrophobic thickness; polarity; database; protein stability; protein folding; membrane protein

Abbreviations: 3D, three-dimensional; BAM, β-barrel assembly machinery; BR, bacteriorhodopsin; CL, cardiolipin; ECF, energy-coupling factor; ER, endoplasmic reticulum; GPCR, G protein-coupled receptors; IM, inner membrane; LPS, lipopolysaccharide; MD, molecular dynamics; MIM, mitochondrial inner membrane; MOM, mitochondrial outer membrane; OM, outer membrane; OPM, Orientations of Proteins in Membranes database; PDB, Protein Data Bank; PG, phosphatidylglycerol; PM, plasma membrane; PPM, Positioning of Proteins in Membranes method and web-server; rmsd, root mean square deviation of Cα atom coordinates; SRP, signal recognition particle; TA, C-tail anchored proteins; TM, transmembrane; VDAC, voltage-dependent anion channel.

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: National Science Foundation; Grant number: 1145367. Grant sponsor: National Institute of Health; Grant number: 5R01DA003910.

*Correspondence to: Andrei Lomize, 428 Church St., Ann Arbor, MI 48109. E-mail: almz@umich.edu

Introduction

Biological membranes separate cells from the external environment by creating selective permeability barriers. In eukaryotic cells, membranes also divide cells into specialized compartments or organelles, such as nucleus, mitochondria, chloroplasts, endoplasmic reticulum, Golgi apparatus, lysosomes, peroxisomes, vacuoles, and transport vesicles. Different types of biological membranes grow, divide, evolve and function as complex macromolecular assemblies with specific lipid, protein, and carbohydrate compositions and distinct physical properties.

The lipid bilayer of cellular membranes is crowded with integral membrane proteins, with lipid:protein ratios ranging from 4:1 to 1:4 by mass.[1] Membrane proteins may be classified as transmembrane (TM), which span the entire lipid bilayer, monotopic, which are permanently inserted into the membrane from one side, and peripheral proteins that bind to the membrane transiently. Membrane proteins participate in all vital cellular process, including protein and lipid biogenesis, cell shape regulation, transport, cell recognition and adhesion, energy production and homeostasis, signal transduction, and generation and propagation of electric impulse.

Because of the technical challenges, the pace of elucidation of three-dimensional (3D) structures of TM proteins is much slower than that of water-soluble proteins. Nevertheless, recent process in protein expression, crystallization, and the emerging crystallography on submicrometre-sized crystals has significantly accelerated membrane protein structure determination.[2,3] In addition, improved NMR techniques have helped study structure and dynamics of TM β-barrels[4,5] and multi-helical proteins,[6,7] and allowed structural determination of dozens of oligomers formed by single-spanning TM proteins, for which crystallization usually fails to yield diffraction quality crystals.[8–10]

3D structures of TM proteins are provided by the Protein Data Bank (PDB).[11] To date, more than 1000 distinct structures of more than 500 TM proteins and protein complexes have been determined. They are represented by ~2000 PDB entries. Many of these structures are listed in Stephen White's database,[12] the Membrane Proteins Data Bank[13] and the transporter classification database.[14] Three other specialized databases provide derivative PDB files with calculated locations of hydrophobic membrane boundaries: Protein Data Bank of Transmembrane Proteins,[15] the Goarse-grained molecular dynamics simulation database,[16,17] and the orientations of proteins in membranes (OPM) database.[18] Collection of membrane proteins in specialized databases facilitates comparative analysis of membrane proteins. OPM is the most convenient resource for such analysis, because it simultaneously provides spatial positions, classification, biological origin, different conformations and oligomeric states, topology, and intracellular localization of TM, monotopic, and peripheral proteins. Spatial positions of proteins in membranes were determined in OPM by optimizing transfer energy of molecules from water to the anisotropic lipid environment.[19] In addition, the underlying positioning of proteins in membranes (PPM 2.0) method provides methodology for analysis of polarity of protein surfaces.[20]

Adaptation of TM proteins to the lipid bilayer leads to a number of structural features which distinguish them from water-soluble proteins. Major features are briefly reviewed here, including the following: (a) limitations imposed by the lipid bilayer on protein architecture; (b) hydrophobic thickness and spatial arrangement of TM proteins in the lipid bilayer, (c) changes in TM protein surface polarity along the bilayer normal that can be described by polarity profiles; (d) the chaperone-assisted insertion and folding of proteins into membranes; (e) role of charged residues in topogenesis of TM proteins; (f) significant oligomerization propensity of proteins in the two-dimensional lipid matrix; (g) large-scale conformational transitions of proteins during their association with membranes and transport function. We also discuss physical forces that drive membrane-protein integration and specific interactions of TM proteins with each other and membrane lipids. The analysis uses 483 unique structures of TM proteins and other information provided by the OPM database (Supporting Information: Tables S1 and S2).

Structural Architectures of TM Proteins

The structural diversity of integral membrane proteins is limited by the restrictions imposed by their hydrophobic environment. TM α-helices, β-barrels, and β-helices are the only known protein architectures that fulfill the requirement to saturate the hydrogen bonding potential of their polar main-chain groups buried in membrane from water. Because of the limitations on protein structure, the probable number of membrane protein families was estimated as about 10 times fewer than that for soluble proteins.[21] A recent estimate suggests that 80% of the sequence space for all polytopic membrane proteins may be covered by only ~700 structures, while 70–80% of soluble protein domain sequences may be covered by ~25,000 structures.[22]

TM α-helical proteins

The vast majority of TM proteins are α-helical (Fig. 1). These proteins are found in all types of biological membranes and are encoded by ~25–30% genes in sequenced organisms.[23] TM proteins are known as bitopic or polytopic if they cross the membrane only one or multiple times, respectively. Bitopic proteins represent the largest membrane protein class

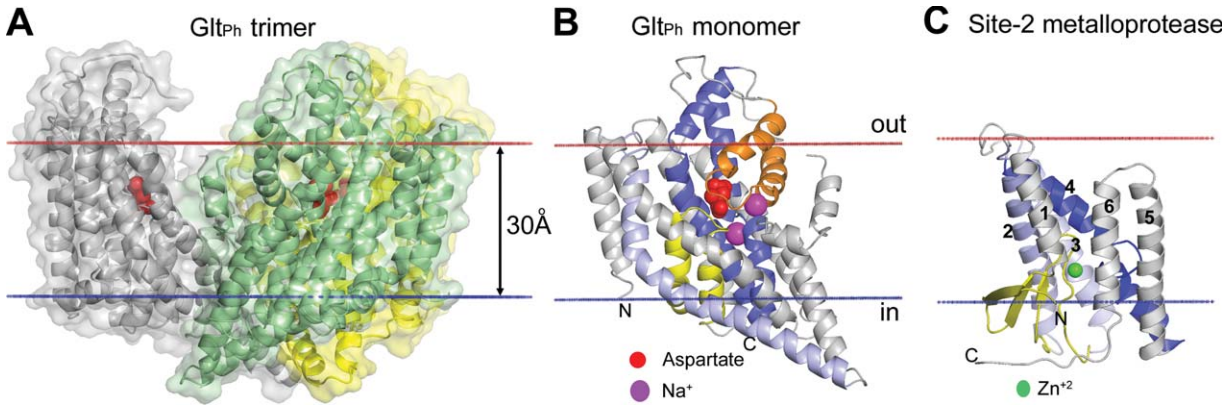

Figure 1. Complex structure of α-helical membrane proteins. (A) A trimer of sodium (Na^+)-coupled aspartate transporter Glt_{Ph} from *Pyrococcus horikoshii* in complex with aspartate (2NWL) has a bowl-shaped architecture with a crevice reaching halfway the lipid bilayer and open to the aqueous environment. Individual subunits in a trimer are colored grey, yellow, and green, aspartate molecules are shown by red spheres. (B) A monomer (2NWX, subunit A) has 8 TM helices and 2 reentrant loops from both membrane sides (yellow and orange). The highly tilted and bended 46 residue-long TM5 extends beyond the lipid bilayer on the cytoplasmic side (light blue). Reentrant loops and two marginally hydrophobic helices (discontinuous TM7 and TM8, blue) form the translocation pathway for substrate-aspartate (red spheres). The unwound region of TM4 defines the binding site for sodium ions (purple balls). (C) Site-2 Zn-metalloprotease from *Methanocaldococcus jannaschii* (3B4R) represents the TM six-helical α-bundle with three short helices (TM1, TM5, TM6) and a discontinuous TM4 helix. TMs 2–3 (light blue) and TM4 (blue) form protein core with a catalytic Zn^{2+}ion (green sphere) in the active site, which is coordinated by two histidines and an aspartate. An unusual structural feature is a membrane-inserted four-stranded β-sheet (yellow) that occludes the active site. Calculated hydrophobic membrane boundaries from the OPM database are shown by lines: blue at the cytoplasmic side and red at the periplasmic/extracellular side.

covering 15–18% and 30–47% of membrane proteomes in prokaryotes and eukaryotes, respectively.[24] Genome-wide analysis of membrane proteins demonstrates that while proteins with small number of TM segments are prevalent in all organisms, prokaryotes also have a higher preference for proteins with 6 and 12 TM α-helices, and eukaryotes for proteins with 7 TM α-helices.[25] It is noteworthy that proteins with even number of helices and with N- and C-termini located at the cytoplasmic membrane side are strongly overrepresented in eukaryotic and bacterial proteomes.[26]

To date more than 400 unique 3D structures of TM α-helical proteins and protein complexes are available for structural analysis (Table S1). However, the actual number of different structures is much greater due to the presence of PDB entries corresponding to different conformational, mutational, and oligomeric states, and complexes with distinct ligands or cofactors. Most structures represent prokaryotic proteins from archaebacteria (32 structures), Gram-negative bacteria (151 structures), and Gram-positive bacteria (33 structures). There are also 165 unique structures of eukaryotic TM α-helical proteins, most of which are from the plasma membrane and the endoplasmic reticulum, ER (103 and 21 structures, respectively), whereas only 7, 17, 14, and 3 structures represent α-helical proteins from the mitochondrial outer membrane (MOM), mitochondrial inner membranes (MIM), thylakoid, and vacuole/vesicule membranes, respectively.

Structural analysis[27,28] of α-helical membrane proteins indicate the complexity of their architecture. TM α-helices are not necessarily straight, hydrophobic, and long enough to traverse the entire bilayer, rather they can be bent, distorted, or shorter than a half of the membrane thickness.[29,30] Structural details of TM α-helical proteins have been extensively reviewed.[27,31–36] Here we only briefly outline them.

TM α-helices of membrane proteins are generally hydrophobic. However, more than 30% of TM α-helices in polytopic proteins have low hydrophobicity.[37–39] Such marginally hydrophobic TM segments are frequently functionally important and present in ion channels and transporters.[40] Computational methods of TM helix prediction are based on search of long hydrophobic stretches and therefore often fail to recognize such helices.[41,42] Though single-spanning helices of bitopic proteins are generally more hydrophobic than TM segments of polytopic proteins,[43,44] around 20% of TM helices in bitopic proteins contain polar residues.[24] The average number of polar residues per helix in bitopic proteins decreases from 4 in prokaryotic to 2 in eukaryotic proteins.[24] The presence of polar and ionizable residues in single-spanning TM proteins was shown to promote helix–helix association.[45–47]

Lengths of TM α-helices range from 15 to 39 residues with an average length of ~26 residues and preference for lengths greater than 20 residues. This differs from helices in water-soluble α-bundles that

have an average length of ~15–18 residues per helix and preference for shorter helices.[32,33,48] Helices usually traverse the entire bilayer with tilt angle ranging from 5° to 35° relative to the membrane normal,[31] and an average tilt 17° ± 11°.[33] However, more tilted helices (tilt >45°) exist, for example, in structures of chloride-protein antiporter ClC (1KPL), sodium-hydantoin transporter Mhp1 (2JLN), sodium-dependent glutamate transporter GltPh (2NWL) [Fig. 1(A,B)], ammonia channel (1U7G), SecYEG translocon (subunit SecE, 3DIN), complex III (2FYU), ECF-transporter complex (subunit EcfS, 4HUQ, 4HZU). The very long and tilted helices usually form extensive contacts with 3–5 helices thus stabilizing protein structure.

TM α-helices may be straight or distorted due to the presence of kinks induced by Pro, tight turns of 3_{10}-helix, or wide turns of π-helix induced by insertions of a single residue into α-helix.[34] Almost 40% of TM helices have irregularities, as compared to 19% in water-soluble proteins.[32] TM α-bundle proteins may include a β-sheet, as it was found in structure of the site-2 meltalloproteinase (3B4R) [Fig. 1(C)]. Some α-helices are discontinuous (e.g., 3RKO), or penetrate only halfway into the membrane, thus forming reentrant loops.[27,35,49] Reentrant loops constitute a signature element in aquaporins (e.g., 3C02), voltage-gated potassium channels (e.g., 2R9R), rhomboid proteases (e.g., 4H1D), ClC H^+/Cl^- exchange transporters (e.g., 1kpl), H^+-Glu^- symporter (e.g., 3KBC), SecY translocases (e.g., 1RHZ). Exposed main-chain in discontinuous helices and non-helical regions of reentrant loops are often associated with binding sites of ions or substrates and regions with increased conformational flexibility.[50]

Many membrane proteins have repeated TM domains with similar tertiary structure.[27] These tandem domains may have either parallel or antiparallel topology in the lipid bilayer. In the latter case, they are named inverted repeat domains. Proteins with duplicated TM domains in parallel orientation include ABC-transporters expressed as a single polypeptide (e.g., P-glycoprotein, 3G5U), and numerous secondary transporters: mitochondrial ATP-ADP carrier formed by three helical hairpins with a pseudo threefold symmetry (1OKC), different transporters from the major facilitator superfamily (1PW4, 2CFQ, 2XUT, 3WDO, 3O7Q, 3H5M, 4GC0, 4J29, 4M64) and from the resistance-nodulation-cell division[51] family (2V50, 3NE5), where two halves of 12 TM α-bundle are related by a pseudo twofold axis. Among proteins with inverted repeat domains are aquaporins (1PW4), glycerol-conducting channel GlpF (1LDF), ammonia channels Amt1 (2B2F), SeqY translocase (1RHZ, 3DIN), transporters BtuCD (1L7V), ClC H^+/Cl^- exchange transporter (1KPL), Leu transporter LeuT (2A65), Na^+/H^+ antiporter NhaA (4BWZ).[27,36] These proteins probably evolutionarily arise from duplication of the corresponding genes, which may lead to formation of covalently linked dual topology dimers.[36,52]

TM β-barrels

The second common type of TM proteins is a β-barrel (Fig. 2). TM β-barrels are abundant in outer membranes (OM) of Gram-negative bacteria, mycobacteria, mitochondria and plant plastids. In addition, various toxins create β-barrel pores in host cell membranes. A TM β-barrel can be formed either by a single polypeptide chain or by multiple chains. Currently there are 102 unique structures of TM β-barrels, among them 90 structures represent single-chain barrels and 12 structures are multi-chain barrels (Table S2). All known TM β-Barrels have the simplest sequential up-and-down topology of antiparallel β-strands. It was suggested that a β-hairpin serves as a principal evolutionary building block of such proteins.[53] All TM β-barrels can be divided into three groups based on their topology and localization.

The first and the most numerous group includes single-chain β-barrels from the bacterial OM [Fig. 2(A,B)]. Around 2–3% proteins encoded by genomes of *Escherichia coli* and *Pseudomonas aeruginosa* are estimated as probable OM β-barrels.[26,54,55] Structural features of OM β-barrels have been extensively reviewed[56–60] and will be only mentioned here.

OM β-barrels are composed of an even number of antiparallel β-strands, and their N- and C-termini face the periplasm. β-Strands are tilted by ~35–50° and connected *via* long loops on the extracellular side and short turns at the periplasmic side of the membrane. The internal pore has an oval shape with hydrophobic and aromatic residues being exposed to the lipid bilayer, while polar and charged residues face the interior of the pore. The most abundant OM proteins are porins that have 16 or 18 β-strands. Porins facilitate passive diffusion of small molecules across the membrane and can discriminate these molecules based on their size, charge, and lipohilicity. Some 8- to 12-stranded β-barrels act as enzymes, receptors, or adhesion molecules. Larger, 12- to 24-stranded β-barrels function as transporters of proteins involved in adhesion or biofilm formation. Because the OM lacks proton motive force, most β-barrels do not use energy. However, some of them can actively import molecules (e.g., iron complexes, nickel chelates, vitamin B12, and carbohydrates) against concentration gradients because they form complexes with active transporters from the bacterial inner membrane (IM). For example, 22-stranded TonB-dependent FhuA (1QFG) and BtuB (2GSK) transporters can use active transport mechanisms through interactions with TonB-ExbB-ExbD complex from the bacterial IM.[61]

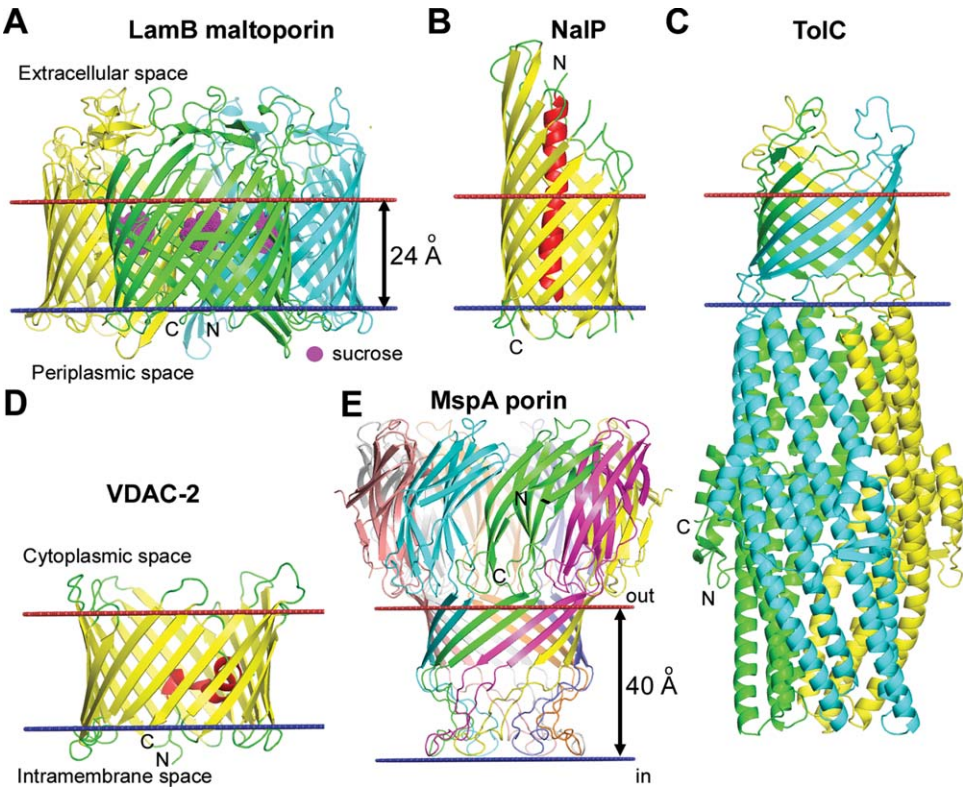

Figure 2. Different types of TM β-barrels. (A) Trimer of maltoporin LamB from the OM of *E. coli* with sucrose in the hydrophilic channel (1AF6). Each monomer of the sugar transporter represents a 18-stranded single-chain TM β-barrel with a channel that selects sugars through interaction with aromatic residues. (B) Monomer of NalP autotransporter from the OM of *Neisseria meningitides* (1UYO). NalP is a 12-stranded single-chain TM β-barrel with the N-terminal passenger domain (red helix) inside the pore that can be released after the self-cleavage. (C) Multichain β-barrel of TolC from the OM of *E. coli* (1EK9). 12-Stranded β-barrel of TolC is formed by homotrimer. (D) Single chain β-barrel of mitochondrial voltage dependent anion channel VDAC-2 from the MOM of *Danio rerio* (4BUM). VDAC-2 represents a 19-stranded TM β-barrel with N-terminal amphiphilic α-helix located inside the circular channel, which is involved in voltage sensing. (E) Multichain β-barrel of MspA porin from the OM of *Mycobacterium smegmatis* (1UUN). The stem domain of MspA represents a 16-stranded β-barrel formed by homo-octamer. Calculated membrane boundaries from the OPM database are shown by red and blue lines.

The second group is represented by single-chain β-barrels formed by isoforms of the mitochondrial voltage-dependent anion channel (VDAC) [Fig. 2(D)], the most abundant proteins in the MOM.[62] VDAC has an unusual topology and is not evolutionarily related to bacterial porins.[53] Unlike bacterial OM β-barrels, VDAC-1 (3EMN) and VDAC-2 (4BUM) have an odd number of antiparallel β-strands that are closed through a pair of parallel β-strands (1st and 19th).[5,63–65] Both isoforms have short loops at both membrane sides and N-terminal amphiphilic α-helix inserted into the round-shape channel.

The third group includes multi-chain β-barrels. Unlike single-chain β-barrels, they are obligatory homo- or hetero oligomers, where each subunit provides several β-strands to a β-barrel that encloses a channel with a relatively large inner diameter. The most extensively studied multi-chain β-barrels are TolC-like proteins (1EK9, 1YC9, 3PIK, 1WP1), the OM components of the bacterial Type I secretion system involved in extrusion of drugs and proteins[66,67] [Fig. 2(C)]. Homotrimeric 12-stranded β-barrels were also found in bacterial autotransporter adhesins Hia (2GR7)[68] and YadA (2LME).[4] Another multi-chain β-barrel is the goblet-like mycobacterial MspA porin (1UUN), the most abundant protein from the OM of *Micobacterium smegmatis* that enable the transport of hydrophilic nutrients into mycobacteria[69] [Fig. 2(E)].

Multi-chain TM β-barrels are also formed by diverse pore-forming proteins, such as mammalian perforins or certain toxins from bacteria, fungi, plants, and eukaryotic parasites.[70,71] The known structures of TM β-barrels of bacterial toxins represent mushroom-shaped homo-heptamers (7AHL, 3O44, 4H56)[72–74] or bicomponent octamers (3B07)[75] [Fig. 3(C)].

TM β-helices

The β-Helix is another type of regular structure which can be stable in the lipid bilayer. It can be formed by peptides with alternating L- and D-amino acids, such as non-ribosomal antibiotics, gramicidin A, B, C.[76,77] The cation-conducting structure of

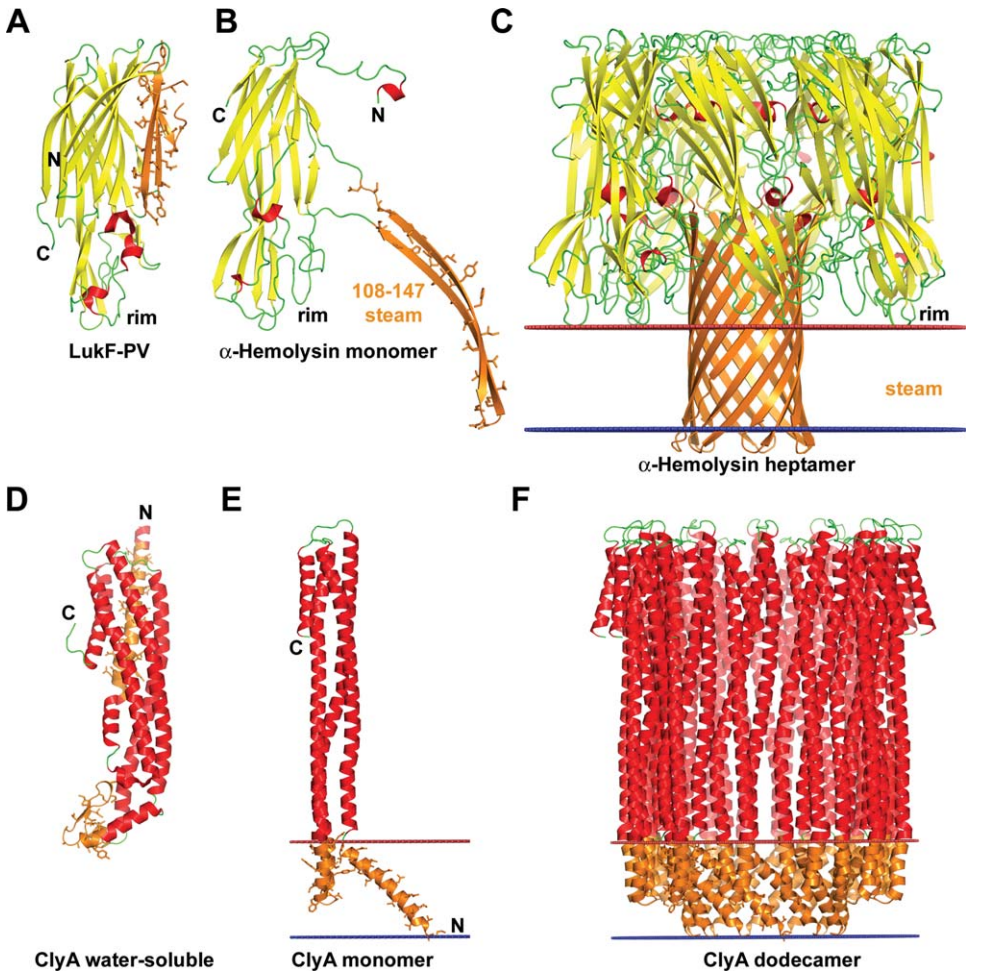

Figure 3. Conformational rearrangements of bacterial pore-forming toxins (PFT) upon binding to membranes. (A–C) Structures of β-PFT from *Staphylococcus aureus*: (A) soluble form of LukF monomer, a homologue of α-hemolysin (1PVL), (B) monomer of α-hemolysin from the oligomer (7AHL), (C) α-hemolysin heptamer (7AHL). Conformational changes in α-hemolysin include release of an amphiphilic β-hairpin (residues 108–147), oligomerization of seven molecules to create a stem domain with the central hydrophilic channel, and insertion of the stem domain into the membrane. (D–F) Structures of α-PFT from *E. coli*: (D) soluble form of ClyA monomer from *E. coli* (1QOY), (E) monomer of ClyA from the oligomer (2WCD), (F) ClyA dodecamer (2WCD). Conformational transitions of the ClyA toxin include detachment of β-tongue domain (residues 176–202) from the α-bundle, rearrangement of N-terminal α-helix, and formation of oligomers (octamers, dodecamers, or tridecamers) with circular pore inside the hydrophobic membrane core. The β-sheet is shown in yellow, the α-helices are shown in red, membrane-bound segments are shown in orange with side chains shown by sticks. Location of hydrophobic membrane boundaries are shown by lines: inner membrane leaflet is colored blue, outer leaflet is colored red.

gramicidin A (1GRM) represents a TM head-to-head dimer of two right-handed single-stranded $\beta^{6.3}$-helices (with 6.3 residues per helical turn)[78–80] [Fig. 4(A)]. The size of the channel pore (diameter ~4 Å)[81] is large enough to enable the passage of monovalent cations.[81,82] The gramicidin A channel is anchored to the membrane-water interface by four Trp near the C-terminus of each protomer. The hydrogen-bonding ability of Trp and dipole moment of the indole ring are essential for stabilization of the structure of the ion-conducting channel.[83]

Changing the gramicidin A sequence (Trp substitutions) or environmental conditions (using organic solvents or lipid bilayers with acyl chains <10 or >20 carbons) promotes formation of very different double-stranded intertwined helices, which can be right- or left-handed with parallel- or anti-parallel chains (2XDC, 2IZQ, 1MIC) [Fig. 4(B,C)]. Double-stranded helices can traverse the membrane but are unable to conduct ions.[76,77,84]

Folding and Targeting of TM Proteins

At appropriate conditions, a number of hydrophobic peptides,[85,86] some α-helical proteins[87–89] and β-barrels[60] can fold and insert into the lipid bilayer independently of proteinaceous chaperones, assembly cofactors, and energy sources. However, such *in vitro* assembly is significantly slower and less efficient than similar processes occurring *in vivo*.

In vivo TM proteins fold and insert into membranes either co-translationally or post-translationally. Membrane protein insertion pathways are especially

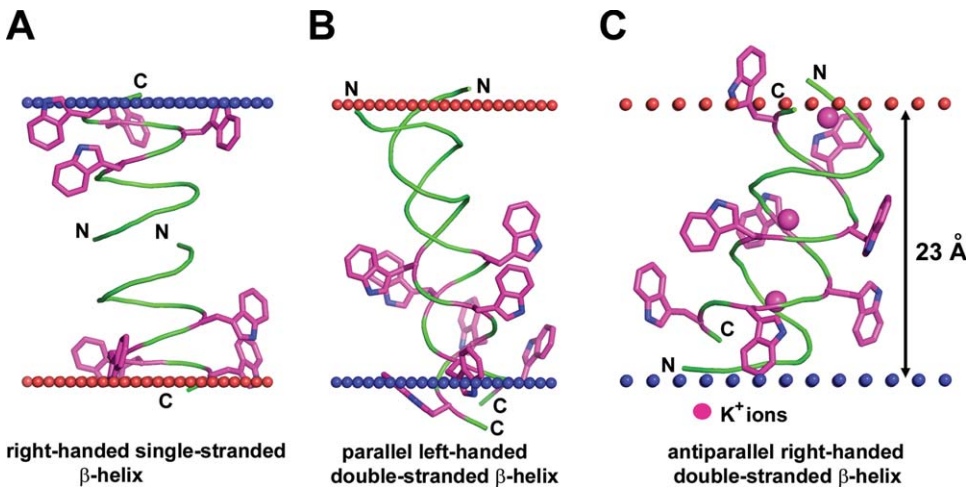

Figure 4. Different structures of antibiotic pentadecapeptide, gramicidin from *Bacillus brevis*. (A) Single-stranded right-handed β-helix of gramicidin A in the lipid bilayer (1MAG) forming a narrow ion channel with circular cross-section (diameter 4 Å). (B) Parallel left-handed double-stranded β-helix of gramicidin A in Ca^{+2}–methanol solution (1MIC). (C) Antiparallel right-handed double-stranded β-helix of gramicidin D crystallized from methanol in complex with K^{+}-ions (2IZQ). Insertion of gramicidin into the lipid bilayer decreases local hydrophobic thickness of the bilayer up to 20 Å and induces leakage of monovalent ions. Four Trp residues in each pentadecapeptide are shown by sticks colored purple (C-atoms) and blue (N-atoms). Calculated membrane boundaries from the OPM database are shown by red and blue spheres.

complex in eukaryotes, because the newly synthesized proteins must be targeted from ER or cytosol to the plasma membrane (PM) or different intracellular organelles. Biogenesis of membrane proteins *in vivo* is assisted by complex protein machineries aimed at recognition, translocation and integration of proteins to specific membranes in correct topology, conformation and oligomeric state. Here we briefly describe key TM chaperones with available 3D structures, which are involved in protein insertion into membranes.

Cotranslational protein folding and insertion into membranes

Most TM α-helical proteins insert into membranes through the co-translational pathway that involves the signal recognition particle (SRP) and Sec-translocon complex. This SRP/Sec-translocon system exists in ER of eukaryotic cells, cell membranes of Archaea and Eubacteria, thylakoid membranes, but is absent in mitochondria.[90,91]

SRP recognizes the cleavable N-terminal hydrophobic signal sequence or the first TM helix while it emerges from the ribosome. By interacting with the SRP receptor and the Sec-translocon, SRP targets the ribosome-nascent chain complex to the eukaryotic ER or the prokaryotic cell membrane. Structures of bacterial SecYEG (3DIN),[92] archaebacterial SecYEβ (1RHZ),[93] and canine Sec61αβγ (4CG5)[94] demonstrate structural conservation of the SRP/Sec-translocon systems (Fig. 5). Ribosome binds directly to the Sec61α/SecY pore at the cytoplasmic site and pushes the nascent polypeptide chain through the protein translocation channel. GTP hydrolysis provides energy for chain elongation by the ribosome. A lateral gate located between TM2b and TM7 of Sec61α/SecY-translocon presumably provides exit of folded helical segments one by one or in pairs into the lipid environment.[93,95–97] Insertion of chloroplast-encoded proteins (e.g., proteins from the reaction center of photosystem II) occurs from stroma into thylakoid membranes through similar cpSRP-pathway.[98]

SecA ATPase assists SecYEG [Fig. 5(C)] in translocation of secretory proteins and membrane proteins with large periplasmic domains across the membrane.[97] SecA-mediated translocation is stimulated by 12-helical SecDF dimer (3AQP), a member of the RND family of transporters that uses the proton-motive force[99] [Fig. 6(B)].

Bacterial SecYEG translocase co-operates with YidC chaperone [Fig. 6(A)] during folding and topogenesis of multi-spanning proteins, proteins complexes[100–102] and proteins with large periplasmic domains.[91,103] YidC can also function as an independent insertase of single and double membrane-spanning proteins (e.g., subunit C of F_1F_0-ATP syntase).

YidC homologues were also found in archaebacteria, mitochondria (Oxa1) and chloroplasts (Alb3). The YidC/Oxa1/Alb3 protein family mediates assembly of major energy-transducing protein complexes.[102] Mitochondrial Oxa1, which functions as a voltage-gated and substrate-dependent translocase,[104] is involved in co-translational insertion of polytopic TM proteins encoded by mitochondrial genome (13 core subunits of respiratory complexes I, III, IV and V) from the matrix to the MIM.[105]

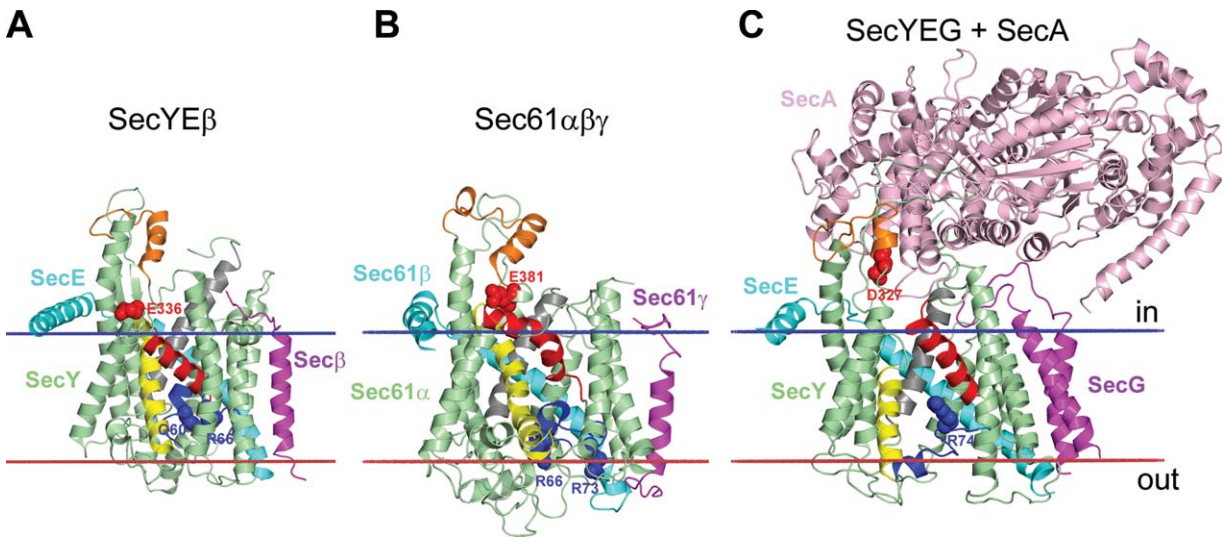

Figure 5. Protein translocation channels: (A) Archaebacterial SecYEβ from *Methanococcus jannaschii* (1RHZ); (B) canine Sec61αβγ (4CG5); (C) complex of SecYEG with SecA from *E. coli* (3DIN). The channel-forming Sec61α/SecY subunit has 10 TM α-helices arranged as two five-helix inverted repeats forming an hourglass-shaped channel between two repeats. The loop5–6 and loop8–9 (colored orange) contact with SecA or ribosome. The channel has a cytoplasmic funnel, the central hydrophobic ring, and external funnel filled by the TM2a helical plug (colored blue) with basic/polar residues (shown by blue spheres) that can electrostatically repel positive charges of translocated segments thus assisting in topology decision.[322] Negatively charged residues from TM8 (shown by red spheres) also play topological role. In resting state (A, B), the lateral gate between TM2b (colored red) and TM7 (colored yellow) is closed, TM10 (colored gray) is moved outward, and the TM2a plug seals the pore ring, thus preventing ion leakage. In the SecYEG-SecA complex (C), insertion of two-helix finger of SecA into cytoplasmic funnel of SecY causes movement of the TM2a plug toward the periplasm and the displacement of the TM2b away from the TM7, thus creating a 5 Å-gap between these helices and providing partial opening of the translocation channel.[92] In the open state, lateral gate between TM2b and TM7 is open and TM10 moves inward to the channel, which helps to release the translocated segment from channel into the lipid environment.[94] Calculated membrane boundaries from the OPM database are shown by red and blue lines.

In addition, the SecYEG-YidC machinery may co-operate with the TatABC-system [Fig. 6(D,E)] during integration of polytopic TM proteins with large globular domains bearing diverse cofactors.[106] It was proposed that some membrane proteins, such as Rieske iron-sulfur protein, are initially inserted into membrane by the Sec-machinery and then use Tat-complex to translocate folded globular domain containing Fe-S cluster.[107]

Post-translational targeting, folding and membrane insertion

In contrast to the majority of TM α-helical proteins, which insert co-translationally, β-barrels fold and insert into membranes *via* post-translational pathways. The post-translational integration with membranes also occurs for C-tail anchored proteins (TA) and nuclear-encoded TM α-helical proteins targeted to mitochondria and chloroplasts.

In Gram-negative bacteria, the biogenesis of OM β-barrel proteins includes several steps[108,109]: (i) SecYEG-SecA-dependent translocation across the IM assisted by SecB chaperone; (ii) transit through the periplasm assisted by chaperones Scp, DEgP and SurA; and (iii) folding and insertion into the OM by the β-barrel assembly machinery (BAM) formed by BamA 16-stranded TM β-barrel and four accessory lipoproteins (BamBCDE). Comparison of available BamA structures (4K3B and 4K3C) suggests that the lateral gate from the protein interior to the lipid bilayer may be open due to transient separation of the 1st from the 16th β-strand [Fig. 6(C)].

C-tail anchored proteins have a large water-soluble N-terminal domain and a single C-terminal TM α-helix. Because the C-terminal membrane anchor emerges from the ribosome after termination of translation, it cannot interact with SRP and Sec-system and uses post-translational insertion pathways that involve cytosolic chaperones and specific membrane receptors.[110,111] It was also suggested that TA proteins may require chaperone assistance only for delivery to the targeted membrane in an insertion-competent form (to prevent aggregation), but not for insertion into membranes.[110]

The nuclear-encoded mitochondrial preproteins are targeted to the organelle by either N-terminal cleavable presequences or internal targeting signals with medium hydrophobicity and positively charged flanking residues.[112–115] Proteins targeted to the MOM usually have internal targeting signals, proteins targeted to the intermembrane space have "classical" amphipathic signals, while proteins directed to the MIM or matrix have either bipartite signals (amphipathic presequence followed by

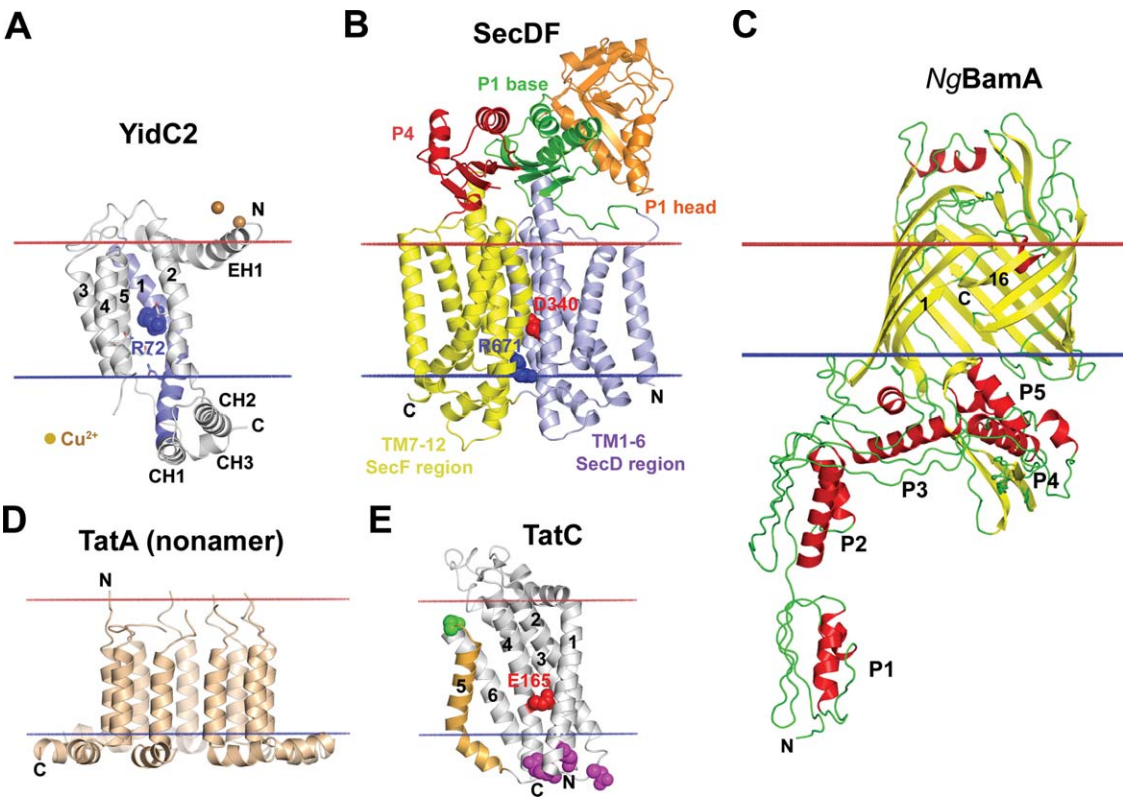

Figure 6. Structures of TM chaperones assisting membrane insertion of TM proteins. (A) YidC2 fragment (27–267) from *Bacillus halodurans*, (3WO7), an independent insertase and a membrane chaperone that cooperates with translocon SecYEG.[196] $YidC_{27–267}$ has five-helical TM core, N-terminal amphiphilic EH1 helix, and CH1, CH2, CH3 hydrophilic cytoplasmic helices. The long TM1 (colored blue) is bent in the middle at P77 and forms a continuous helix with the CH1 helix. TMs 1–5 form a positively charged hydrophilic grove with central R72 (blue sphere) and several solvent-exposed polar residues (shown by sticks). The grove is open to the lipids and cytoplasm and sealed toward the extracellular side. It was suggested that positively charged R72 in the grove (shown by blue spheres) transiently captures hydrophilic regions with acidic residues of translocated helices, thus facilitating their TM insertion.[196] (B) Structure of SecDF (F-form) from *Thermus thermophilus* (3aqp), a chaperone powered by proton motive force that assists SecYEG. D340 (shown by red spheres) and R671 (shown by blue spheres) are residues involved in proton conduction.[99] (C) BamA protein from *Neisseria gonorrhoeae,* the central component of BAM complex (4K3B). BamA, which has TM 16-stranded β-barrel and five periplasmic POTRA domains (P1-P5), coordinates recognition, folding and membrane insertion of TM β-barrels in the bacterial OM. BamA can undergo lateral opening between 1st and 16th β-strands, providing route from the interior channel to the lipid environment.[323] (D) Nonamer of TatA from *E. coli* (2LZS), a protein-translocating element of Tat-system. TatA binds to TatBC/substrate complex and polymerizes to form substrate translocation channel. (E) TatC from *Aquifex aeolicus* (4B4A), a six-helical substrate-binding component of Tat-system. TatC forms complex with bitopic TatB protein *via* TM5 (colored orange with green spheres). Tween-arginine signal sequence of a substrate is recognized by TatC (contact region is shown by purple spheres) within TatBC complex. Bitopic TatA protein binds to negatively charged concave face of TatC with solvent-exposed E165 (shown by red spheres). Calculated membrane boundaries from the OPM database are shown by red ("out" side) and blue ("in" side) lines.

hydrophobic sorting sequence) or internal signals.[114] Proteins in unfolded state are translocated across MOM by the Translocase of Outer Mitochondrial membrane (TOM)-complex composed of several receptors and TOM40 β-barrel channel.[114,116,117] Almost all mitochondria-targeted proteins are recognized and transferred across the MOM by TOM complex, except TA and signal-anchored proteins, which insert into MOM spontaneously. Membrane insertion of these proteins depends on the lipid composition of target membranes, for example, on the presence of ergosterol and cardioplipin.[118–120] TOM complex also assists insertion into MOM of some bitopic and polytopic proteins.[117] β-barrels (VDAC, TOM40, and SAM50) are inserted into MOM from the intermembrane space by the Sorting and Assembly Machinery (SAM)-complex which is similar to bacterial BAM-complex.[114,117]

Protein translocation and insertion into MIM is assisted by translocases of the inner membrane (TIM) complexes. TIM22 complex inserts multispanning proteins containing internal signals, such as ADP/ATP and phosphate carriers.[116] The twin-pore carrier TIM22 is an essential integral protein of the complex. It forms a hydrophilic high conductance channel activated by the inner membrane

potential $\Delta\psi$ and regulated by internal targeting peptides.[121] TIM23 complex interacts with presequence-translocase-associated import motor (PAM) complex. Together, they translocate proteins with cleavable presequences into mitochondrial matrix or insert them into MIM. TIM23 complex uses both the electric membrane potential $\Delta\psi$ and the ATP-dependent activity of PAM as driving forces for translocation of preproteins across MIM.[120] Oxa1 may assist TIM23 complex in import of multi-spanning MIM proteins.[114]

Plant chloroplasts have three distinct membranes: the outer, the inner and the thylakoid ones. Targeting to chloroplasts is guided by cleavable N-terminal transit sequences or by internal sequences included into mature proteins.[115,122] Proteins targeted to thylakoid lumen have bipartite presequences that contain a tandem of stromal and luminal sorting signals. Interior chloroplast proteins are translocated to stroma across both the outer and inner envelope membranes by the Translocons of the Outer and Inner envelopes of Chloroplast (TOC and TIC). They use GTP and ATP as energy sources.[123,124] After translocation to the stroma and cleavage of the transit sequence, TM proteins can insert into thylakoid or inner envelope membranes. Proteins destined to thylakoid membranes can use one of three chaperone-assisted pathways (cpSec-, cpSRP-, or DpH-) or insert spontaneously.[125,126] The first, cpSec-pathway requires SecYE, SecA proteins and ATP for translocation of proteins targeted to the lumen or TM proteins with large globular domains. The second, cpSRP-pathway works post-translationally and independently of cpSecY. It requires two SRPs with their receptors and Alb3.1 protein.[127] This pathway is engaged in transport of the highly abundant nuclear-encoded light harvesting chlorophyll a/b-binding proteins (LHCPs) (1rtw).[128] The third, DpH-dependent pathway uses the proton gradient (ΔpH and $\Delta\psi$) as a driving force. It provides translocation of presequences with twin-arginine motif to the lumenal side of thylakoids (e.g., PsaN and PsbT from Photosystem II).[129,130] DpH-system components are homologous to components of the bacterial Tat-system.[129] Finally, the spontaneous insertion is restricted to a few single-spanning membrane proteins (CF_oII of the ATPase, and PsbX, PsbW of the photosystem II) and some polytopic proteins (Elip2, PsbS).[90]

Spatial Arrangement in Membranes and Hydrophobic Thickness of TM Proteins

Each TM protein has large continuous non-polar surface that penetrates to the lipid acyl chain region. To provide burial of this non-polar surface from water, a TM protein acquires a certain spatial position in the lipid bilayer. This position can be investigated experimentally or calculated by minimizing transfer energy of the protein structure from water to membrane.[131]

Each TM protein has an additional important geometrical parameter: its intrinsic hydrophobic thickness, which can be defined as the width of a slice enclosed by two parallel planes separating the non-polar and polar regions. Localization and concentration of proteins with similar hydrophobic thicknesses in a certain membrane type may regulate the membrane width.[132] Therefore, average hydrophobic thickness of TM proteins is an appropriate characteristic of a particular membrane type: proteins from different membrane types tend to have distinct intrinsic hydrophobic thicknesses (Fig. 7). For example, average hydrophobic thicknesses of α-helical proteins from eubacterial and archaebacterial cell membranes, eukaryotic ER, and thylakoid membranes are close to 31 Å. Average hydrophobic thicknesses are significantly larger for eukaryotic PM proteins ($\sim$33 Å), but smaller for proteins from MIM ($\sim$28 Å). This is in agreement with X-ray scattering data.[132] The gradual decrease of hydrophobic thicknesses of proteins targeted to different membranes (PM>ER>MIM) may facilitate protein sorting during membrane biogenesis. Indeed, the post-translational targeting and unassisted insertion of TA proteins into MOM requires short and moderately hydrophobic helices with positively charged flanking residues, while ER-targeting is driven by longer and more hydrophobic helices.[110,133]

All β-barrel proteins from the OM of Gram-negative bacteria and the MOM have significantly smaller hydrophobic thickness ($\sim$24 Å) (Figs. 2 and 7), in agreement with NMR studies of detergent-embedded residues of OmpX, analysis of the thickness of detergent belt in crystals of OMPLA, and neutron scattering studies of LPS bilayers.[134–138] In contrast, the stem domain of β-barrel porin MspA from the outer mycobacterial membrane has unusually large hydrophobic thickness of $\sim$40 Å [Fig. 2(E)], consistent with the expected thickness of this membrane, as it contains up to $\sim$60% of long-chain (C_{30} to C_{90}) mycolic acids.[139]

Polarity Profiles, Membrane Interface, and Midpolar Regions

The lipid bilayer cannot be viewed merely as a hydrocarbon slab. This is a complex anisotropic environment that can be described by profiles of polarity and lateral pressure along the bilayer normal.[140] While a protein in water is under isotropic pressure, a TM protein experience negative pressure at the level of lipid carbonyls, which is counterbalanced by positive pressure at the lipid head-group and acyl chain regions.[140] The lateral pressure profiles in membranes cannot be directly measured; however, they can be estimated from MD simulation of artificial lipid bilayers.[141–143]

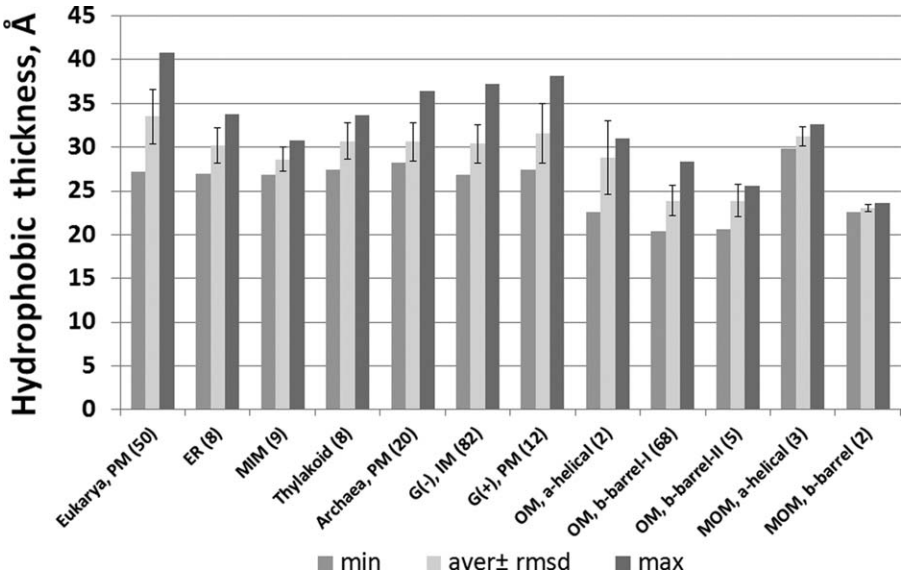

Figure 7. Intrinsic hydrophobic thicknesses of α-helical and β-barrel TM proteins from different membrane types: eukaryotic plasma membrane (PM), endoplasmic reticulum (ER), mitochondrial outer (MOM) and inner (MIM) membranes, PM of Archaebacteria and Gram-positive bacteria, outer (OM) and inner (IM) membranes of Gram-negative bacteria. Numbers of protein structures in each protein set are indicated in parenthesis (adopted from Ref. 20).

Polarity, elastic, and geometrical properties of artificial lipid bilayers can be experimentally obtained using scattering techniques.[144–147] For example, the polarity of a fluid lipid bilayer can be quantified by a few parameters commonly used to describe solubility of molecules in organic solvents: hydrogen bonding donor and acceptor capacities (α and β) and dipolarity/polarizability parameter (π*).[19] The transbilayer profiles of these parameters have been recently calculated for several single- and multi-component lipid bilayers using distribution of lipid fragments obtained from neutron diffraction and X-ray scattering data[20] [Fig. 8(A,B)].

Unfortunately, the polarity of natural biological membranes cannot be determined in the same way, because they have a highly diverse and yet unidentified lipid and protein composition. Instead, one can analyze hydrophobic thickness and polarity of surfaces in 3D structures of TM proteins, which is expected to match polarity of surrounding lipids. Indeed, all TM proteins have a central hydrophobic zone composed of aliphatic residues and two polar zones that contain numerous ionizable residues and crystallized water. Arg and Lys residues are embedded in the lipid head-group area, where they can form H-bonds and ion pairs with lipids. The intermediate region is marked by belts of Tyr and Trp residues that favorably interact with the membrane interface.[148–153]

Hence the polarity of protein surfaces may be described by distributions of different types of amino acid residues, such as polar, non-polar, aromatic, charged, and so forth. However, it can be described more quantitatively by calculating profiles of polarity parameters α, β, and π*. Unlike distributions of residues, these polarity profiles represent integral quantitative characteristics that account for hydrogen-bonding capacities and dipolarity/polarizability of atoms on the protein surface.

The corresponding polarity profiles were recently obtained for TM proteins from different biological membranes, including bacterial, archaebacterial and eukaryotic plasma membranes, inner mitochondrial, thylakoid, and bacterial outer membranes.[20] As expected, the polarity of protein surfaces (described by α, β, and π*) changes gradually [Fig. 8(D,F)] and correlates with properties of the surrounding lipid bilayer. Polarity profiles can be approximated by sigmoidal curves whose midpoints coincide with midpoints in distributions of co-crystallized water molecules, with maxima in distributions of carbonyl groups from co-crystallized lipids, and match positions of calculated hydrophobic membrane boundaries. Interestingly, polar groups of Trp and Tyr residues of TM proteins often point to membrane boundaries where they may form hydrogen bonds with glycerol groups of co-crystallized lipids. However, distributions of aromatic rings of Tyr and Trp residues do not coincide: while maxima of distribution of Tyr benzene rings correspond to the lipid carbonyl region, maxima of distributions of Trp indole rings are shifted by 3–5 Å closer to the membrane center [Fig. 8(C,E)]. This tendency was observed for all proteins from different membrane types, including α-helical and β-barrel TM proteins.

Polarity profiles calculated for proteins from different membrane types appear to be quite similar at the center of membranes (Fig. 9). Thus, central regions of proteins derived from membranes with dissimilar composition demonstrate comparable hydrophobicity. This observation may explain the

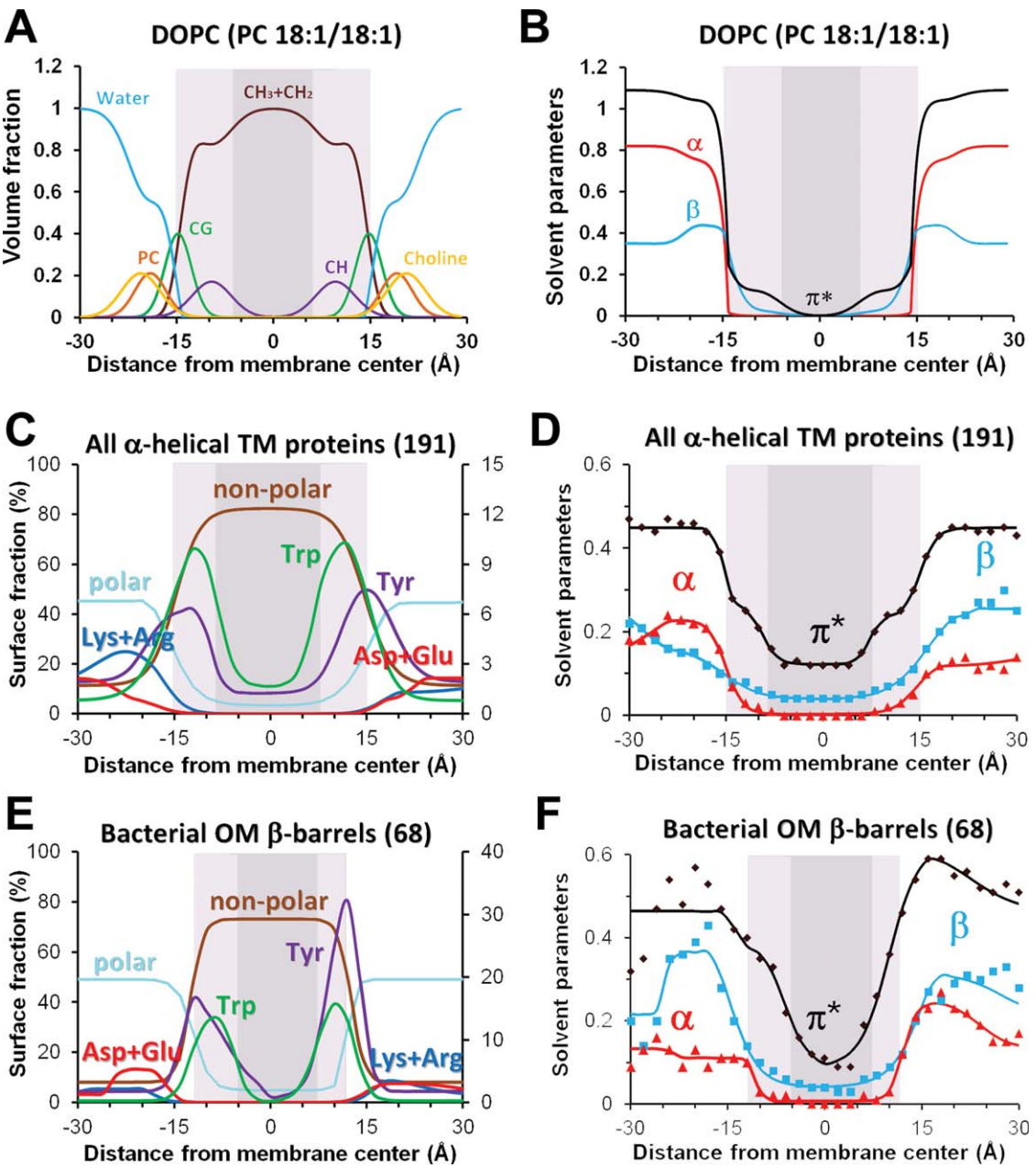

Figure 8. Distributions of chemical groups and polarity profiles in artificial lipid bilayers (A, B) and at the lipid-facing surfaces of TM proteins (C-F). (A) Volume fractions of lipid segments determined by X-ray and neutron scattering for fluid DOPC membrane.[324] (B) Profiles of hydrogen-bonding donor (α) and acceptor (β) capacities and dipolarity/polarizability parameter (π*) calculated for DOPC bilayers.[19] (C, E) Distribution of lipid-facing atoms in structures of 191 TM α-helical proteins from all membranes (C) and 68 OM β-barrels (E): polar atoms (N- and O-atoms) of all residues, non-polar atoms (C- and S- atoms of Val, Leu, Ile, Met, Cys, Phe, Tyr, and Trp), aromatic atoms and C-atoms from benzene ring of Tyr and indole ring of Trp, charged groups of Asp/Glu and Arg/Lys. (D, F). Transbilayer profiles of polarity parameters (α, β, and π*) were calculated for lipid-facing atoms on surface of TM α-helical (D) and OM β-barrel proteins (F) (see Ref. 20 for details). Midpolar regions (colored light gray) were mapped as area of preferential accumulation of Trp indole rings in TM proteins (C–F), and based on locations lipid double bond in artificial bilayers (A,B). Central hydrophobic regions are colored dark gray.

well-known tolerance of TM proteins to alteration of the lipid composition in native cells, in artificial bilayers, and during protein expression in different host organisms.[154] Larger distinctions in polarity profiles are observed near hydrophobic boundaries and in the lipid head group areas. These dissimilarities may correlate with the large diversity of the lipid head-groups in different membranes though may also reflect biases associated with mechanisms of protein insertion into membranes.

A closer examination of polarity profiles in TM proteins (Figs. 8 and 9) shows that the central, 30–31 Å-wide, non-polar region is not a uniform environment, but includes two peripheral, 5–8 Å-wide

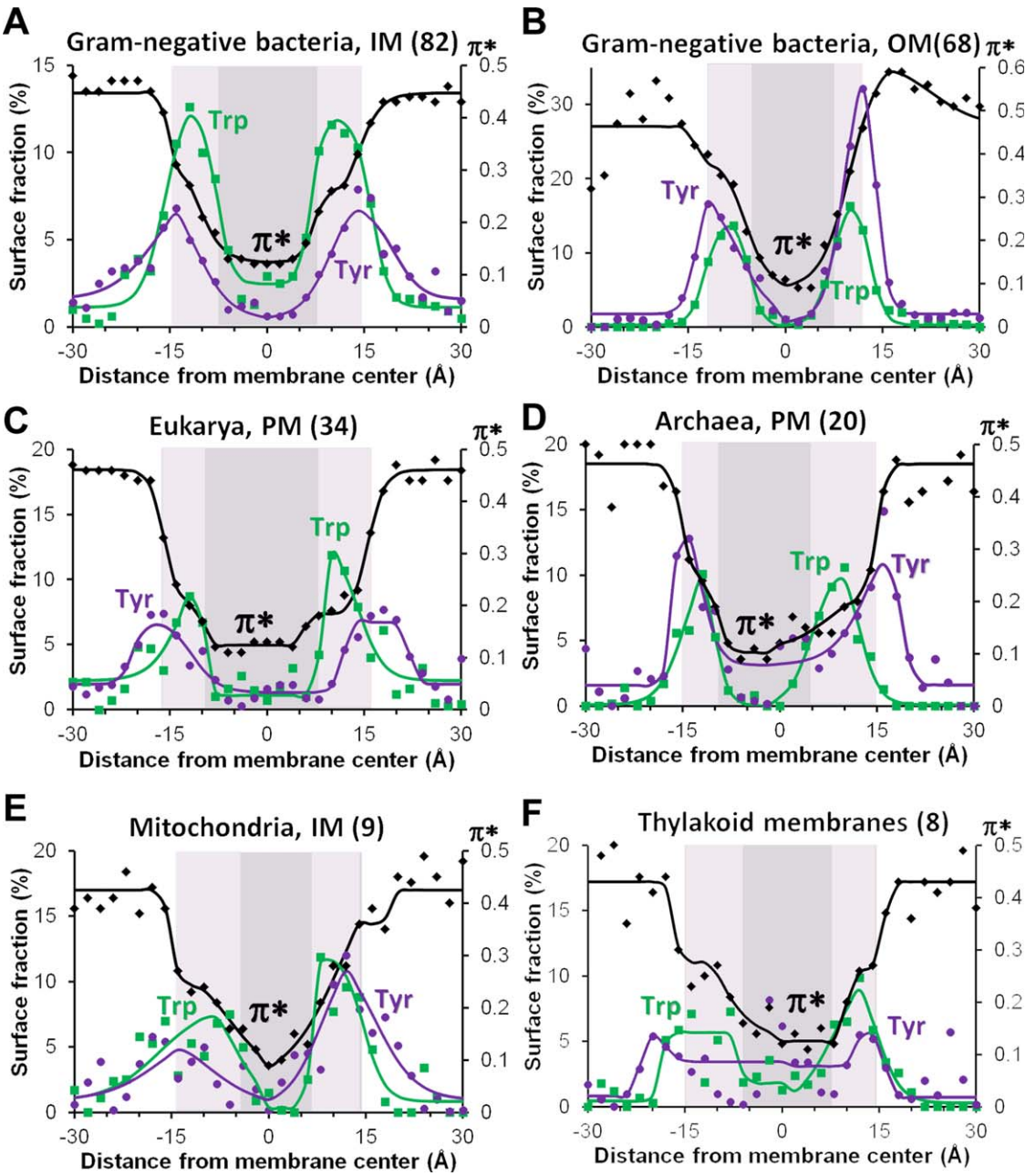

Figure 9. Comparison of surface polarity in TM proteins from different biological membranes: IM (A) and OM (B) of Gram-negative bacteria, PM of eukaryotic (C) and archaebacterial (D) cells, mitochondrial inner membranes (E) and thylakoid membranes (F). Transbilayers distributions of aromatic and C-atoms of benzene ring of Tyr (purple lines and dots) and indole ring of Trp (green lines and squares) are compared with profiles of dipolarity/polarizability parameter π* (black lines and diamonds) calculated from distributions of all solvent-exposed atoms in proteins from different membranes. Numbers of structures in each protein set are indicated in parenthesis. Midpolar regions mapped by the preferential accumulation of Trp indole rings are colored light gray, central hydrophobic regions are colored dark gray (see the set of proteins and details of calculations in Ref. 20).

zones with increased dipolarity/polarizability parameter (π*) and hydrogen-bonding acceptor capacity (β). These peripheral zones with intermediate polarity can be named "midpolar" regions.[155] In native biological membranes, location of "midpolar" regions can be deduced from the changes of polarity parameter π*, or mapped by distribution of Trp indole rings in TM proteins, which can be regarded as a reporter group. Preferential accumulation of the Trp indole rings in the area between lipid carbonyl groups and double bonds indicates a relatively small electrostatic penalty for the Trp dipole and a higher dielectric constant there, which may be partly attributed to the presence of residual water.

Depths of "midpolar" regions, as deduced from locations of Trp residues, depend on the type of membrane and may be different on the inner and outer sides. For example, the polarity profiles of proteins

from the bacterial OM demonstrate significant asymmetry of their "midpolar" regions. This correlates with asymmetric lipid composition in OM, where the inner leaflet is composed of phospholipids, but the outer leaflet is formed by lipopolysaccharides (LPS).[156] In contrast, proteins from the bacterial IM, which are composed entirely of phospholipids, show symmetric polarity profiles. Some degree of asymmetry in the "midpolar" regions was also observed for proteins from the eukaryotic and archaebacterial PM, MIM, and thylakoid membranes (Fig. 9). This observed asymmetry of polarity in eukaryotic PM proteins correlates with non-uniform distribution of different lipid species and cholesterol between leaflets of eukaryotic cell membranes.[157,158]

Topogenesis of TM α-Helical Proteins and Role of Charged Residues

In addition to the spatial arrangement in membranes and hydrophobic thickness, an important structural parameter of TM proteins is topology, or placement of their termini on the certain side of the membrane. This is another feature that distinguishes them from water-soluble proteins. TM proteins usually acquire unique topology, although there are rare examples of proteins with dual topology.[37] For example, the bacterial multidrug transporter EmrE functions as an antiparallel homodimer of monomers with opposite topologies (3B5D).[159,160] Based on their orientation in membranes, bitopic proteins were classified to topological classes I (with N-terminus facing to the outer space and N-terminal cleavable signal peptide), II (with N-terminus facing the inner space), and III (with N-terminus facing to the outer space), while all polytopic proteins are assigned to class IV regardless of their topology.[1,161]

The topology of TM α-helical proteins is usually defined during their co-translational insertion, folding, and assembly in membranes.[37] Topogenesis of multi-spanning proteins may involve significant rearrangement and reorientation of helices during protein folding[162–166] and post-translational insertion of reentrant loops.[163] For example, marginally hydrophobic helices can be released from Sec-translocon peripherally in the lipid head group region and be refolded later in the course of helix assembly.[163] Membrane insertion and spatial orientation of such helices may be promoted by adding charged or hydrophobic residues to the flanking regions, or guided by interactions with adjacent TM helices and their topological preferences.[38,40,44,167,168] For example, insertion of highly charged S4 helix of voltage-sensing domain of the Shaker and KAT1 voltage-gated K-channels is assisted by electrostatic interactions with neighboring helices.[169]

The topology of proteins with co-translational membrane insertion depends on a number of factors. Among them are properties of TM segments (length, hydrophobicity, hydrophobic gradient of helical segment, position of aromatic residues near TM helix ends, charge difference at both ends of helix), folding and glycosylation of extramembrane domains, membrane electrochemical potential, and protein–lipid interactions.[170–172]

However, two major topological determinants are the hydrophobicity of TM segments and the asymmetric distribution of positively charged residues (Arg and Lys) in polar regions located at both membrane sides, the so-called positive-inside rule.[173,174] This rule was suggested based on statistical analyses of amino acid residue distributions in TM α-helical proteins from Eubacteria, Archaea, and Eukarya (including proteins form mitochondria and thylakoids).[161,174–181] Such analysis exposed the significant prevalence of positively charged residues in loops located on the side of the membrane from which proteins are inserted (the *cis* side or "inside"), which is a cytoplasmic side of bacterial IM, eukaryotic PM and ER, the stromal side of thylakoids, and the matrix side of MIM. The "inside" bias of positively charged residues appear to be stronger for prokaryotic proteins.[174] Interestingly, the positive-inside rule is applicable to mitochondria-encoded MIM proteins, but not to nuclear-encoded MIM proteins which show more symmetric distribution of positively charged residues on both membrane sides, but significant "outside" bias for negatively charged Glu residues.[180]

Distributions of positive charges in 3D structures of TM proteins are consistent with the "positive inside" rule for α-helical proteins from all membranes, no matter whether one considers all protein residues[182,183] or only solvent-accessible ones.[20] Moreover, not only the distributions of cationic residues, but also distributions of net charge follow this pattern in protein structures from nearly all membranes: it is positive on the inner membrane side and close to zero or slightly negative on the outer side. Two exceptions are proteins from MIM and thylakoid membranes: MIM proteins have a positive net charge on both membrane sides, though it is larger at the inner leaflet, while thylakoid TM proteins have rather similar numbers of positively and negatively charged residues at each side, and therefore no pronounced maxima for net charge (Fig. 10).

Maxima of distributions of positively charged groups are located at approximately 20–22 Å-distance from the bilayer center in both membrane leaflets. This corresponds to locations of phosphodiester groups of phospholipids that may form ionic pairs with Arg and Lys residues. The unusual lack of maxima of net positive charges in proteins from thylakoid membranes may be explained by the unique membrane composition of these membranes.

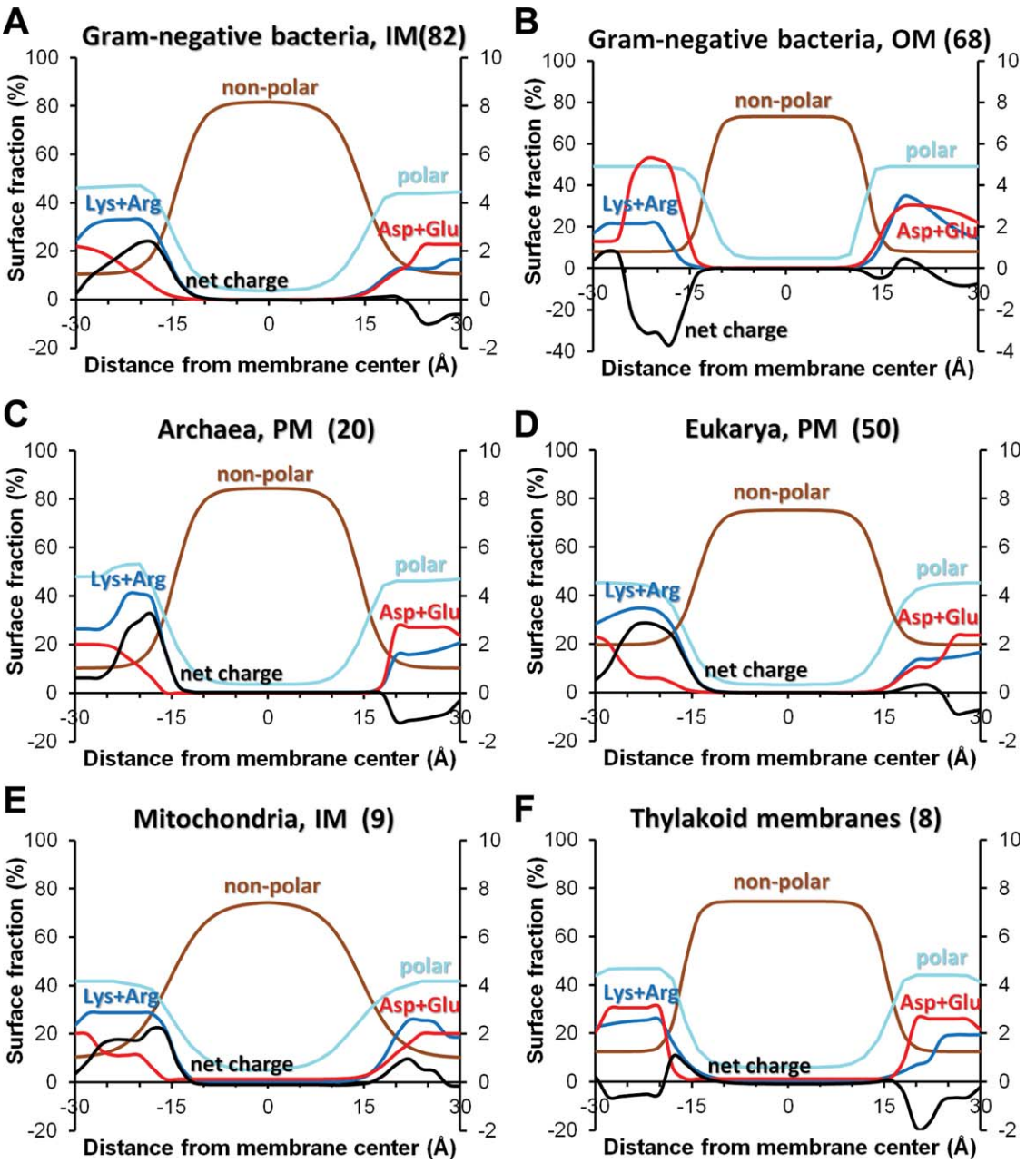

Figure 10. Distribution of lipid-facing atomic groups and charges in TM proteins from different biological membranes: IM (A) and OM (B) of Gram-negative bacteria, PM of archaebacterial (C) and eukaryotic (D) cells, mitochondrial inner membranes (E) and thylakoid membrane (F). Distribution of polar atoms (N- and O-atoms) of all residues are shown by light blue lines, non-polar atoms (C- and S- atoms of Val, Leu, Ile, Met, Cys, Phe, Tyr, and Trp) are shown by brown lines, positively charged groups of Arg and Lys are shown by blue lines, negatively charged groups of Asp and are shown by red lines, net charges are shown by black lines. Numbers of structures in each protein set are indicated in parenthesis (see the set of proteins and details of calculations in Ref. 20).

In thylakoid membranes, phospholipids are mainly substituted by non-phosphorous glycolipids,[184] which significantly decreases the amount of lipid groups involved in ionic interactions with basic protein residues.

The role of positive charges as topological determinants was experimentally supported for both signal peptides and integral membrane proteins. For example, it was found that adding or removing of a single basic residue at critical locations can invert the topology of individual TM segments,[185] or entire proteins.[164] Placing charged residues downstream of a signal sequence arrested protein chain movement through Sec-translocon[186] and blocked protein secretion.[187] This "charge block" effect was more pronounced in prokaryotic than eukaryotic systems, which can be attributed to the presence of an electrical potential (positive outside) across the bacterial IM.[187]

Acidic residues may also affect topology of proteins inserted by Sec-translocon, but to a lesser

extent than basic residues, and only if present in high numbers.[185] It was suggested that negative charges around hydrophobic segments play only a secondary role in fine-tuning the topology in Sec-dependent membrane insertion.[185] In contrast, acidic residues in periplasmic regions of double-helical M13 and single-helical P3f coat proteins appear to be important for Sec-independent translocation/insertion that requires the proton motive force.[188–190] Consistently with smaller topological role of acidic residues, it was found that distributions of these residues in sequences of membrane proteins is generally symmetric[174] and show only moderate preference of Glu on the *trans* ("outside") side of the MIM.[161,180] Likewise, analysis of 3D structures of membrane proteins showed almost symmetric distribution of negatively charged residues at the both membrane sides.[20,182,183]

Physical mechanisms behind the positive-inside topological rule are not completely understood. It may be possible that the retention of positively charged residues on the cytoplasmic side partially results from their interactions with acidic phospholipids at the cytoplasmic leaflet[191] or from the influence of the electrochemical potential on the bacterial IM (inside negative).[190,192] However, the same positive-inside bias was observed for proteins that reside in membranes of acidophilic archaebacteria with reversed $\Delta\psi$ (inside positive).[193] Hence the positive-inside rule is not strictly connected to the sign of the transmembrane potential. Instead, it is likely attributed to structural details of the Sec-translocation channel and YidC chaperone. Indeed, mutagenesis studies of the Sec61p-translocon demonstrated that orientation of signal sequence can be changed by reversing the positively charged residues in the TM2a plug (R67E, R74E) or by elimination of a negatively charged residue at the cytoplasmic end of the TM8 (E382Q)[194,195] (Fig. 5). The insight on the topological role of acidic residues of single-spanning proteins was obtained from the inspection of the crystal structure of the bacterial YidC[196]: the translocation of N-terminal acidic residues flanking the single TM helix may be directed by conserved positive charge located in the hydrophilic grove of the YidC [Fig. 6(A)].

Unlike TM α-helical proteins, single-chain TM β-barrels from the bacterial OM demonstrate opposite "positive-outside"[197] trend in distribution of solvent-exposed Lys and Arg residues [Fig. 10(B)]. Basic residues are highly abundant in outside-facing loops, where they interact with negatively charged LPS from the outer leaflet. In contrast, acidic residues occur much more frequently in periplasmic turns of β-barrels ("negative-inside" rule),[20] where they interact with cationic periplasmic chaperones (e.g., Skp) that assist in folding and insertion of β-barrels into the OM.[198]

Oligomerization of TM Proteins

TM α-helical proteins frequently form dimers or higher order oligomers in membranes.[199] The majority of membrane protein complexes (50–70%) are homo-oligomers,[200,201] where homo-dimers represent the largest fraction (>65%).[202] The prevalence of homomers is also observed in available structures of α-helical TM proteins, where homo-dimers and higher order homo-oligomers account for ~20% and ~30% of unique protein structures, respectively (Table S1), while hetero-oligomers account for ~19% of structures. Large membrane protein complexes may also include peripheral proteins, for example in signaling cascades and electron-transfer chains.

Though oligomerization is a common property of both water-soluble and membrane proteins,[203] the presence of a lipid bilayer reduces the area accessible to diffusion of proteins and therefore promotes non-specific aggregation due to several factors: increased concentration of proteins in two-dimensional space, microphase separation and lateral segregation of lipids, and formation of lipid rafts. The lipid bilayer also restrains positions of TM helix termini and forces helices to adopt nearly parallel or anti-parallel arrangements.

All functional protein–protein association in membrane is driven primarily by specific interactions, rather than these non-specific factors. Specific interactions may involve non-covalent binding of extramembrane and/or TM domains, domain swapping and even covalent linking.[199,203–206]

High affinity lateral association of TM α-helices critically depends on packing of complementary helical interfaces that may contain specific dimerization motifs or polar residues involved in interhelical salt bridges or hydrogen bonds.[45,207–211] The extent and structure of TM protein aggregates may also depend on membrane curvature,[212] be modulated by physico-chemical properties of the lipid bilayer[213] and affected by the environmental conditions (pH, presence of ions, organic or macromolecular ligands, lipids, and detergents).[214–216]

While the physiological impact of experimentally observed transient oligomerization of some proteins, such as GPCRs, is not completely understood,[217] in many cases, the functional significance of stable multi-protein complexes has been well-established. It is known that protein association increases protein stability, generates TM conductive pores (ion channels, bacterial secretion systems, toxins), forms ligand-binding sites at the oligomer interface, allows cooperativity between subunits, and provides an additional level of regulation.[218]

Some structural units are stable and functional in membranes only as oligomers.[199,218] Among them are: α-helical dimers of potassium uptake proteins (3UM7), ABC transporters (2QI9, 3G5U, 3PUW, 4AYT), EmrE multidrug transporter (3B5D), MgtE

transporter (2YVX), trimers of acid-sensing and ATP-gated channels (3H9V, 4NTX), tetramers of voltage-, ligand-, Ca^{2+}-, or H^{+}-gated (2R9R, 3EFF, 3LDC) potassium channels, pentamers of ligand-gated channels (3rhw), hexameric Gap junction-forming connexins (2ZW3) and CRAC-channels (4HKR), 4-, 5-, 7-meric mechanosensitive channels (2OAR, 2VV5, 3HZQ), F-, V-, and A-type proton and sodium-translocating ATPases with 10 to 15 subunits (1YCE, 2BL2, 2X2V, 2XQU, 3V3C, 4B2Q), octameric α-helical barrels of Wza (2J58), 14-meric α-helical core complex of type IV secretion system (3JQO), trimeric β-barrels of TolC-like bacterial proteins (1EK9), 12-meric α-pore-forming ClyA toxin (2WCD), and heptameric haemolysin (7AHL) or octameric (3B07) β-pore-forming toxins [Figs. 2(C,E) and 3(C,F)].

Large-Scale Conformational Transitions

TM proteins seem to undergo significant conformational transitions more frequently than water-soluble proteins due to several reasons. First, co-translational and post-translational integration of proteins in membranes frequently leads to significant conformational changes. These changes have been extensively studied for antimicrobial, cell-penetrating and fusogenic peptides, bitopic proteins, and pore-forming toxins.[8,205,219–223] For example, membrane binding and insertion of many biologically-active peptides is accompanied by folding and stabilization of their regular secondary structure, the so-called partitioning-folding coupling.[224,225] Pore-forming toxins also undergo large conformation rearrangements and oligomerization when converted from the water-soluble form to membrane-integrated pores.[71] After binding to specific receptors (e.g., lipids, proteins, sugars, GPI anchors) in target cell membranes, these water-soluble toxins undergo activation, oligomerization, and folding into α-helical or β-barrel structures enclosing large (up to ~26 Å inner diameter) water-filled channels.[71] Bacterial, viral, and eukaryotic toxins can form different types of pores that either destroy membranes or facilitate toxin delivery into host cells. Though molecular details of toroidal pores that can be formed by actinoporin-like toxins are still poorly understood, structures of pores formed by α-helical dodecamer of *E. coli* ClyA toxin (2WCD) or by multi-chain β-barrels of bacterial α-, and γ-hemolysins (7AHL, 3O44, 4H56)[72–74] and bicomponent leukoltoxin (Luk) (3B07)[75] have been resolved. Comparison of structures of water-soluble and membrane-bound forms of pore-forming toxins uncovers the magnitude and nature of structural rearrangements during protein activation, oligomerization and insertion into the lipid bilayer (Fig. 3).

Furthermore, the large-scale conformational transitions between multiple states ("open," "closed," "inward-facing," "outward-facing") is a signature of proteins involved in transfer of various substances or signals across the membrane: receptors, channels, and transporters. Conformational transitions may include significant movements of "plug" and other structural elements and changes in positions of individual subunits, domains, or entire protein complexes with respect to the lipid bilayer. Structural transitions in secondary and primary transporters induced by binding of ligands, lipids, and nucleotides involve domain reorientation and movement of helical pairs.[27,226–229] Rigid-body helix movements can be depicted as translation, piston, rotations parallel (pivot) and perpendicular to the membrane.[230]

For example, large protomer reorganization and reorientation in membranes occurs in energy-coupling factor (ECF), a member of ABC transporter superfamily. ECF transporters are composed of two TM domains (EcfT and EcfS) and two nucleotide-binding cytoplasmic domains (EcfA and EcfA').[231,232] Substrate-binding domains EcfS are individually stable in membranes and may exist either independently (3RLB, 3P5N, 4DVE) or as part of a tetrameric complex (4HZU). Surprisingly, TM helices of EcfS are oriented along the membrane normal in isolated subunit, but almost perpendicular to membrane normal in tetramer (Fig. 11).

Conformational transitions may involve changes in subunit composition and quaternary structure of protein complexes. One notable example of substrate-dependent protomer rearrangement is the Tat-system found in bacterial, archaebacterial, and thylakoid membranes.[233] The Tat-system is composed of three small proteins, single-spanning TatA (2LZS) and TatB (2MI2) and a hexa-helical TatC (4B4A). These subunits form large multimeric complexes for translocation across membranes of folded proteins with twin-arginine signal sequence. After binding of a translocated protein to TatBC complex, this complex recruits TatA protomers and oligomerize into a TM pore-like channels[228,233,234] of variable diameter (30–70 Å) depending on the number of protomers in the complex (85–130 particles)[235] [Fig. 6(D,E)].

Energetics of Protein–Membrane Interactions

3D structures of both water-soluble and membrane proteins correspond to a minimum of Gibbs free energy in their native environment.[236] Physical forces and factors that contribute to structural stability of membrane proteins were discussed in the literature.[237–243] A predictive and verifiable theory of membrane protein folding must account for all these forces and reproduce the experimentally determined membrane folding pathways and thermodynamic stabilities of TM α-helical and β-barrel proteins. However, development of such theory still remains a challenge.

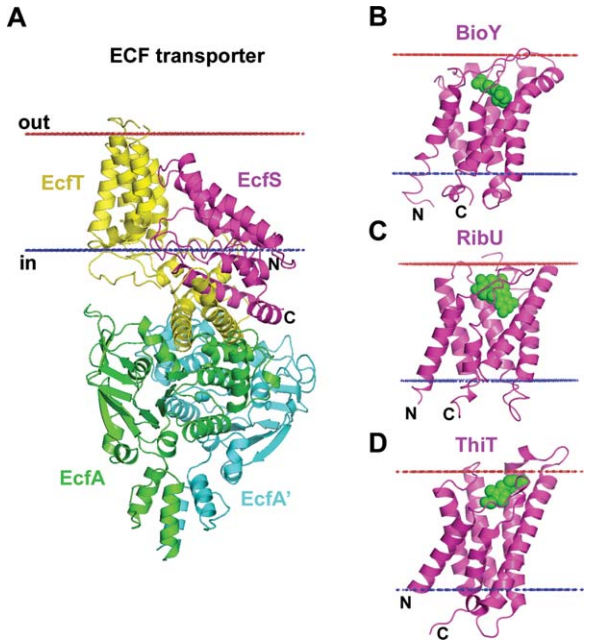

Figure 11. Reorientations of TM domains in energy coupling factor (ECF) transporters. (A) ECF transporter tetramer from *Lactobacillus brevis* (4HZU) is composed of TM energy-coupling component (EcfT) that links cytosolic ATPases, EcfA-EcfA', to the substrate-binding or S-protein (EcfS). In context of the ECF transporter tetramer, helices of the EcfS are highly tilted relative to the membrane plane (>45°) with the opening of the substrate-binding pocket approaching the cytoplasmic membrane interface. (B–D) Individual substrate-binding proteins are stable in membranes and are oriented along the membrane normal with substrate-binding pocket facing the extracellular space: (B) Biotin transporter S-protein, BioY, from *Lactococcus lactis* (4DVE); (C) Riboflavin transporter S-protein, RibU, from *Staphylococcus aureus* (3P5N); and (D) Thiamine transporter S-protein, ThiT, from *Lactococcus lactis* (3RLB). Substrates inside the binding pockets of S-proteins are shown by green spheres. Calculated hydrophobic membrane boundaries from the OPM database are shown by lines: blue at the cytoplasmic side and red at the extracellular side.

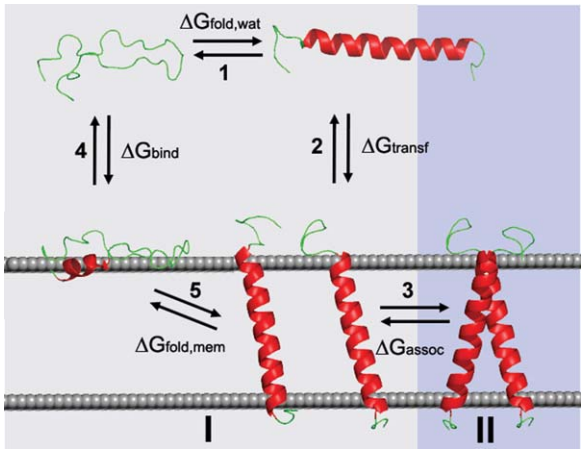

Figure 12. Folding, membrane insertion and oligomerization of an α-helical peptide. The five states of the peptide are: a coil, an α-helix in aqueous solution, a partially folded peptide at the surface of membrane, TM α-helix and TM α-helical dimer. Numbers in the thermodynamic cycle indicate different processes: (1) helix–coil transition in water, (2) transfer of folded α-helices into the lipid environment; (3) helix–helix association in membrane, (4) polypeptide chain binding and partial folding at membrane interface, (5) membrane insertion/folding of interface-bound polypeptide chain. Processes (4+5) and (3) correspond to steps (I) and (II) in two-stage model of membrane protein assembly *in vitro*:[325] (I) α-helix folding and insertion; and (II) lateral association in the lipid bilayer.

Folding of α-helical peptides and proteins in membranes may be described by a thermodynamic cycle that includes several processes: (1) helix-coil transition in water, (2) transfer of folded α-helices into the lipid environment; (3) helix-helix association in membrane, (4) polypeptide chain binding and partial folding at membrane interface, (5) membrane insertion/folding of interface-bound polypeptide chain (Fig. 12). A similar cycle can be drawn for β-barrel proteins. Free energy changes during these processes were determined in experimental studies of protein folding and helix-coil transition in aqueous solution ($\Delta G_{\mathrm{fold,wat}}$),[244–246] peptide binding to membrane surface (ΔG_{bind}),[247–249] transmembrane insertion of peptides ($\Delta G_{\mathrm{fold,wat}} + \Delta G_{\mathrm{transf}}$),[250–252] membrane binding of folded peripheral proteins ($\Delta G_{\mathrm{transf}}$),[253] and energies of α-helix association in membranes and micelles ($\Delta G_{\mathrm{assoc}}$).[254] In addition, equilibrium unfolding energies were measured for many TM α-helical and β-barrel proteins.[255]

These experimental free energy values can be used for development and verification of theoretical models that describe energetics of protein folding and association in membanes. For example, formation of a TM α-helical dimer can be represented as the sequence of steps (1), (2), and (3), that is, formation of an α-helix in water, helix transfer from water into the membrane, and association of helices in the membrane. Though these steps do not necessarily describe real events during protein folding and insertion in membranes, this cycle remains valid for energy calculations based on experimentally-derived parameters. Energy of helix-coil transition (step 1) can be calculated using experimental α-helix propensities, entropic and enthalpic contributions for polypeptide backbone, and energies of interactions between side-chains in water.[244–246] Transfer energy of a protein from water to the fluid lipid bilayer (step 2) can be estimated using an implicit solvent model of the lipid bilayer.[19,256] Energy of specific helix-helix association (step 3) can be evaluated based on empirical parameterization of hydrogen-bonds and van der Waals forces in non-polar media.[257] The corresponding physical processes of transfer and association in membrane are described in more detail below.

Non-specific solvation of membrane proteins by the lipid bilayer

Insertion of membrane protein into the lipid bilayer is driven by the hydrophobic effect, as proteins tend to bury their large hydrophobic surfaces from water to non-polar environment. However, the hydrophobic interactions are unimportant for helix–helix association inside the lipid bilayer. This situation differs from folding of water-soluble proteins, where burial of non-polar residues from water drives the entire folding process.

The energy of insertion of TM proteins into membranes can be evaluated based on various hydrophobicity scales. Whole-residue hydrophobicity scales were derived from studies of partitioning of peptides or amino acid analogues between water and more hydrophobic organic solvents or lipid vesicles,[258] from mutagenesis studies,[259] from statistical analysis of residue distribution along the membrane normal,[182,183,260–263] or MD simulations.[264–267] All these scales correlate with each other, though absolute values vary by several fold.[39,268,269] Both experimental and theoretical studies indicate the favorable insertion into the lipid bilayer of non-polar residues, but the high cost of insertion of polar and charged residues.

Recently, the so-called "biological" hydrophobicity scales have been derived based on of the efficiency of the insertion of hydrophobic segments into the ER membranes (mammalian and yeast), bacterial IM, and mitochondrial inner membranes, which were mediated by Sec61, SecYEG, YidC, and TIM23 translocons, respectively.[43,270–273] It was found that "biological" scales developed for different systems are rather similar, though absolute difference between contribution for the most hydrophobic and most hydrophilic residues is larger in bacterial system and smaller in yeast than for the mammalian ER membranes. "Biological" scales appear to be rather similar to a "knowledge-based" scale derived from relative abundance of residues in membrane proteins[182] and to the White-Wimley octanol scale.[274]

Although whole-residue hydrophobicity scales may be helpful for predicting TM α-helical segments and their topology,[275,276] or positioning of proteins in membranes,[183,262] they cannot properly account for the solvent exposure of protein residues. For example, the solvent-accessible surface area and the corresponding solvation energies are much greater for amino acid analogues than for less exposed residues included in an α-helix, a β-barrel, or the whole protein structure.[19] Therefore, the correct estimation of solvation energy requires the full-atomic description of membrane proteins.

The estimation of transfer energies of proteins from water to membrane requires adequate representation of the lipid bilayer, where energy cost of residue insertion strongly depends on membrane penetration depth.[43] It's noteworthy that even hydrocarbon slab representation of the lipid can produce very reasonable results for positioning of TM and peripheral proteins in membranes.[253] Nevertheless, more complex description of anisotropic membrane properties and more advanced solvation models are desirable to account not only for the depth-dependent hydrophobic interactions, but also for electrostatic, hydrogen bonding, ionic, and other energy contributions that change along the membrane normal. For example, the recently proposed PPM method[19,256] uses a new implicit solvent model and represents the anisotropic lipid bilayer by profiles of polarity parameters (α, β, ϵ, and π*) calculated from X-ray and neutron scattering data on lipid bilayers. This approach also allows characterization of polarity profiles for different biological membranes, including their hydrophobic thicknesses and polarity of interfacial regions.[20]

Folding and insertion of TM α-helical proteins in artificial membranes also depend on mechanical properties of the bilayer. These properties can be regulated by changing lipid composition and vesicle size. In particular, it was shown that curvature stress translated in increased lateral pressure at the center of the lipid bilayer enhanced folding of OmpA and galactose transporter GalP, but decreased folding of bacteriorhdopsin, diacylglycerol kinase, and EmrE transporter.[277] Therefore, calculations of protein transfer energy into membranes, as well as the energy of protein association (see below), should account for some additional contributions from protein–lipid and lipid–lipid interactions, including curvature stress and mechanical deformation of the lipid bilayer.[140] The energetic cost of bilayer deformations depends on material properties of the bilayer, such as thickness, intrinsic lipid curvature, and the elastic compression and bending moduli.[278–280]

Helix–helix association in the lipid bilayer

The association of TM α-helices within the non-polar lipid bilayer cannot be driven by hydrophobic forces, as for water-soluble proteins. Instead, it is primarily driven by van der Waals interactions of tightly packed helices and, to a smaller degree, by interhelical hydrogen bonds and ionic interactions.[204,205] To describe energetics of helix association in membranes, the following interactions should be considered: (1) attractive van der Waals forces and hydrogen bonds in the lipid environment, (2) side-chain conformational entropy changes which oppose helix–helix association; and (3) additional "solvation" energy contributions which come from protein–lipid and lipid–lipid interactions, including those responsible for mechanical deformation of membrane,[281] and (4) electrostatic interactions of helix

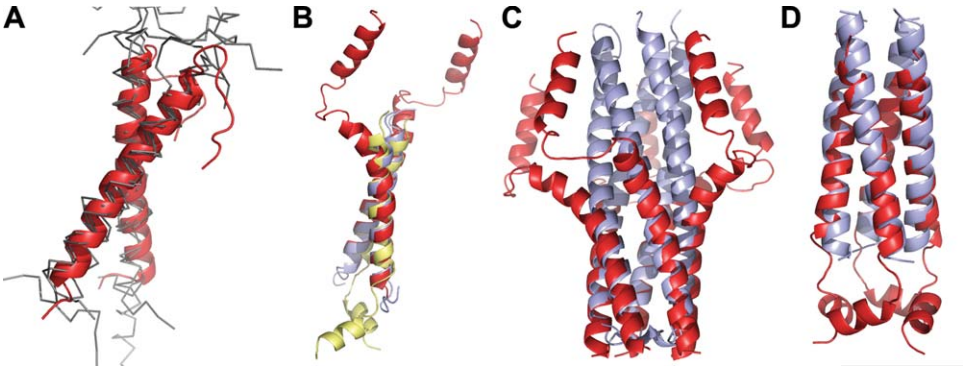

Figure 13. Structural similarity of TM and water-soluble α-helical oligomers. (A) Superposition (rmsd 1.1 to 2.2 Å) of right-handed (crossing angle ∼−40°) TM dimers of bitopic proteins with GxxxG motif: glycophorin A (1AFO, cartoon, colored red) and other unrelated TM homo- and hetero-dimers (2J5D, 2JWA, 2K1K, 2K9J, 2J7A, thin lines, colored grey). (B) Superposition (rmsd 1.1 to 1.2 Å) of left-handed (crossing angle ∼20°) TM helix pairs taken from bitopic protein complexes with heptad packing motif: zeta dimer (2HAC, colored blue), phospholamban pentamer (1ZLL, colored red), and M2 tetramer (2RLF, colored yellow). (C–D) Superposition (rmsd ∼1 Å) of TM oligomers with water-soluble coiled-coil proteins: (C) phospholamban (1ZLL, colored red) and water-soluble pentameric cartilage oligomeric matrix protein (1VDF, colored blue); (D) tetramer M2 channel (2RLF,red) and water-soluble tetrameric CGN4 leucine zipper mutant (1GCL, colored blue).

macrodipoes and rigid-body immobilization entropy of interacting helices.[257,282]

All intermolecular forces, including van der Waals interactions and H-bonds, are of electrostatic nature, and therefore depend on dielectric properties of the environment.[283] The magnitudes of these interactions are similar in the non-polar interiors of water-soluble and TM proteins, ∼ −1 to −1.5 kcal mol^{-1} for hydrogen-bond and ∼0.02 kcal mol^{-1} $Å^2$ for van der Waals interactions.[257,284,285] It is noteworthy that helix-helix association is similar to crystallization or liquid-solid state transition.[285,286] Therefore, these interactions are of the same nature as enthalpy of fusion of molecular crystals; they follow "like dissolves like" rule, and their energy is much weaker than "potential energy" in molecular mechanics and dynamics, which is parameterized using enthalpy of sublimation. This requires novel parameterization of interatomic potentials based on data for condensed media, such as $\Delta\Delta G$ values for protein mutants.[257,285]

The importance of van der Waals interactions for association of TM helices is supported by many observations. The mutual arrangements of TM α-helixes are limited to a few discrete packing modes, which provide knobs-into-holes intercalation of their side chains, as required for close packing of the quasi-parallel helices. These packing modes include parallel and antiparallel helix arrangement, with left-handed and right-handed crossing motifs.[287] In parallel helices, the left-handed packing with positive crossing angles (∼+20°) is related to the presence of a heptad repeat, while the right-handed packing with negative crossing angles (∼−40°) is related to the presence of a tetrad repeat and GxxxG-like sequence motif.[207,208] The structure of parallel TM dimers and symmetric oligomers are well superimposible (rmsd of 1–2 Å) with corresponding templates taken from water-soluble proteins (Fig. 13).

Furthermore, small and polar residues (Gly, Ala, Ser, Thr) occur more frequently at helix packing interface in TM proteins[207] than in water-soluble α-bundle proteins.[32,288] The presence of small residues (Gly>Ala>Va>Ile) is required for efficient helix association in membranes, because it allows close packing and significantly increases van-der Waals interactions of anti-parallel helices with heptad motifs.[289] In contrast, efficient association of similar helices in water, which is driven by hydrophobic forces, required presence of larger side chains on the contact surface and show higher preference for the Ile side chain vs. Gly (Ile>Va>Ala>Gly).[289]

The important role of hydrogen-bonding and ionic interactions was demonstrated for self-assembly of hydrophobic TM α-helices bearing centrally located polar amino-acid residues (Asn, Asp, Gln, Glu, and His).[290–292] It was shown that the affinity of oligo-Leu TM segments and membrane-soluble MS1 peptides in the lipid bilayer dramatically increased upon placement of an Asn residue within the TM segment,[290,292] which was due to formation of intermolecular hydrogen bonds between Asn side chains.[290] The role of ionic interactions between charged residues embedded into the lipid bilayer was demonstrated in a study of the T-cell receptor complex.[293] Formation of α-helical complexes is also affected by electrostatic interactions of α-helix macrodipoles in membrane, which are stabilizing for antiparallel, but destabilizing for parallel α-helices.[257,282]

In addition to enthalpic contributions, the free energy of helix-helix interactions in membrane depends on entropic contributions that oppose the

association. The most important of these represents the conformational entropy changes of flexible side-chains due to their immobilization during association.[257] The loss of side-chain entropy is smaller for residues that are conformationally restricted in isolated α-helices (Gly, Ala, Val, Thr, Ile).

Another important factor for association of proteins in membranes is the hydrophobic mismatch, that is, the difference between hydrophobic thicknesses of TM proteins and surrounding lipid acyl chain region.[294] A hydrophobic mismatch can be alleviated through changes of protein tilt angles, conformational transitions, oligomerization, or by thinning or stretching the lipid bilayer.[295] Bilayer-mediated elastic forces caused by the hydrophobic mismatch can significantly affect specific and non-specific association of TM domains.[209,211,212,296–299] For example, the hydrophobic mismatch between mitochondrial F_1F_0 ATP synthase (hydrophobic thickness 32.8 Å) and inner mithochondrial membrane (average hydrophobic thickness 28.6 ± 1.4 Å)[20] may promote oligomerization of this protein. Indeed, oligomers of mitochondrial F_1F_0-ATP synthase were observed by EM as ribbon-like structures stacked in parallel along the crystae.[300]

Specific lipid–protein interactions

Insertion, positioning and 3D structure of proteins in membranes can be significantly affected by direct protein–lipid interactions.[149,150,301–303] The strength of protein–lipid association can vary from high affinity specific binding to rotational immobilization of first-shell (or annular) lipids that undergo fast exchange with bulk lipids.[279] Many TM proteins were co-crystallized with annular lipids or surrounding detergents, as well as with specifically bound bound lipids that usually originate from native membranes. Here we mention several such examples.

Acyl chains of annular lipids usually fill hydrophobic cavities on protein surfaces and form non-specific van der Waals interactions. Surprisingly, it was observed that hydrocarbon tails of different lipids or detergents in tetramers of mammalian aquaporin AQP0 (2b60, 3m9i, 1ymg) occupy similar cavities on protein surface, while their head-groups interact differently with polar residues of the protein.[304,305]

The locations of strongly bound lipids are also rather similar in structures of homologous proteins, such as mitochondrial (2DYR) or bacterial (1M56, 1QLE) cytochrome c oxidases.[306] The structural conservation of lipid binding sites were reported for cholesterol in β2-adrenergic receptor (2RH1) and Na^+-K^+-ATPase (2ZXE), for cardiolipin (CL) in formate dehydrogenase–N (1KQF), and for phosphatidylglycerol (PG) in KcsA potassium channel (1K4C).[303]

The specific binding of lipids and other hydrophobic molecules, such as carotinoids, chlorophyll and other cofactors, can stabilize loosely packed TM α-bundles and large multimeric protein complexes.[150,303] For example, folding, stability, and function of the plant light harvesting complex II, which is composed of loosely packed marginally hydrophobic TM helices (1RWT, 2BHW), requires presence the of lipids and numerous cofactors (chlorophylls and carotenoids).[307] Lipids co-crystallized with cyanobacteria Photosystem II (3BZ1) include 11 monogalactosyl diacylglycerol, 7 digalactosyl diacylglycerol, 5 sulfoquinovosyl diacylglycerol, and 2 PG. This lipid composition of the preferential location of acidic lipids at the stroma leaflet reflects the relative contents (~45 mol%, ~25 mol%, ~15–25 mol%, 5–15 mol%, respectively) and asymmetric distribution of these mostly non-phosphorus lipids in thylakoid membranes.[308] Coordination of lipid molecules with photosynthetic proteins was shown to play an important structural role: lipids fill internal cavities near cofactors, stabilize oligomers, "lubricate" subunits that need to be structurally flexible, and, possibly, influence the biophysical properties of cofactors that affect the rate of electron transfer.[309]

Natural lipids and cofactors also stabilize trimeric complexes of archaebacterial retinal-containing bacteriorhodopsin (BR)-like ion pumps.[310] Different cryo-electron microscopy and crystal structures of BR show the presence of up to 30 lipid-binding sites per trimer that are occupied by archaeols (*sn*-2,3-diphitanylglycerol) consisting of two C_{20} isoprenoid chains connected *via* an ether linkage to glycerol.[311] Besides, in each trimeric assembly of all BR-like proteins, except BR itself, carotenoid bacterioruberin with C_{50} carbon atoms stabilizes the trimer by binding to crevices between adjacent protomers; while in the deltarhodopsin structure (4FBZ), three extended squalene molecules together with lipid molecules fill the central opening of the trimer [Fig. 14(A–C)].

The crystal structure of dimeric mitochondrial cytochrome c oxidase (2DYR) includes 26 tightly bound phospholipids.[312] Among these lipids, there are four CL molecules: two CLs facing the intermembrane side that supposedly participate in dimer stabilization, and two CLs facing matrix side that may function as proton traps that facilitate the proton translocation along the protein surface[313] [Fig. 14(D)]. CL bearing four acyl chains and two negative charges is the signature phospholipid of mitochondrial membranes: it accounts up to 20 mol% of the inner membrane.[314,315] CL is known to be critically required for biogenesis and stabilization of respiratory chain supercomplexes[316,317] and for regulation of many other mitochondrial functions.[318] Another highly abundant mitochondrial protein, the ADP/ATP carrier, represents a TM six-helical bundle

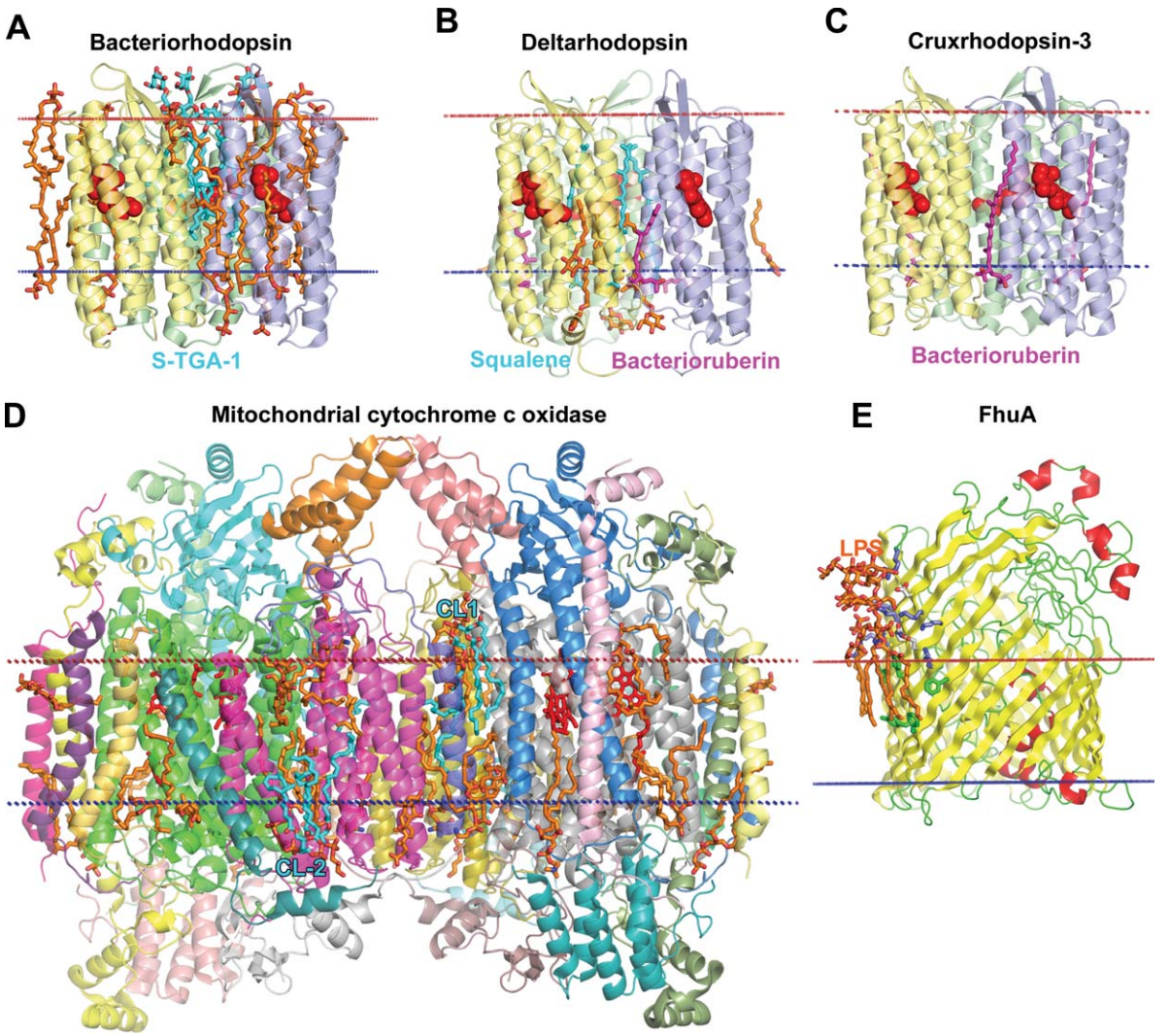

Figure 14. Specific protein–lipid interactions. (A) Bacteriorhodopsin trimer from *Halobacterium salinarium* (1IW6) with co-crystallized archaeol lipids (colored orange), including S-TGA-1 sulfoglycolipid (3-sulfate-Galpβ1–6Manpα1–2Glcpα1-archeol) colored cyan. (B) Deltarhodopsin trimer from *Haloterrigena thermotolerans* (4FBZ) with co-crystallized lipids/detergents (colored orange), 3 squalene molecules (colored cyan) occupying intertrimer space, and 3 bacterioruberin moleculs (colored purple) located at monomer interfaces. (C) Cruxrhodopsin-3 trimer from *Haloarcula vallismortis* (4JR8) with 3 co-crystallized bacterioruberin moleculs (colored purple) located at monomer interfaces. (D) Dimeric mitochondrial cytochrome c oxidase from bovine heart (2DYR) co-crystallysed with 26 lipids (colored orange) including four hemes A (colored red) and four CLs (colored cyan): two CLs (CL-1) at the intermembrane side that likely participate in dimer stabilization, two CLs (CL-2) molecules at the matrix side that may participate in the proton translocation along the protein surface.[313] In the latter case CL in each monomer interacts with basic residues from subunits 3 and 7A1. (E) Ferrix hydroxamate uptake receptor FhuA *from E. coli* (colored by secondary structure: yellow β-strands, red helices, and green loops) in complex with LPS (colored orange) (1QFG). Lipid A of LPS forms specific interactions with a number of basic (shown as blue sticks), polar (white sticks) and aromatic residues (green sticks) of FhuA.

formed by three repeated α-helical hairpins, each bound to a CL molecule.[319] It was shown that CL is critical for carrier folding and translocation.[320]

Specific lipid-protein interactions may involve contacts of lipid hydrocarbon chains with aromatic residues and ionic and/or hydrogen bonding interactions of lipid phosphodiester groups with Tyr, Asn, Lys, His, Arg, Trp, Ser, and Gln located near lipid head groups.[149,150] For example, a conserved structural motif for LPS recognition by the ferric hydroxamate uptake receptor FhuA, an OM β-barrel, [Fig. 14(E)] includes cluster of eight basic residues (K280, K306, K351, R382, R384, K439, K441, R474) and three polar residues (E304, Q353, D386) that form ionic and electrostatic interactions with the phosphate moieties. In addition, six aromatic residues of FhuA (F231, F235, Y284, F302, F355, F380) interact with six acyl chains of the lipid A.[303,321]

These examples demonstrate the essential role of annular and specifically bound lipids for structure, stability, and function of TM proteins and their complexes. Thus, both non-specific and specific protein–

lipid interactions should be taken into account during computational modeling of membrane proteins.

Summary

This review summarizes our current understanding of membrane protein structures, different pathways of protein biogenesis, physical forces that define protein folding and integration into membranes, as well as common and distinctive properties of TM proteins targeted to different biological membranes. The available spatial architectures of TM proteins are limited to α-helical bundles, single-chain and multi-chain β-barrels and β-helices. However, their 3D structures are highly complex and include significant irregularities in secondary structure (3_{10}-helices, kinks, interruptions, β-bulges, turns of π-helix), reentrant loops, and repeated domains with identical or different topologies. TM proteins frequently form large complexes, which are crucial for their biological function. Structural examination demonstrates several helix-helix interaction motifs, suggests the role of ligands and cofactors, and highlights ability of these proteins to switch between different conformational states and oligomerization modes. The most significant conformational transitions are observed in pore-forming toxins that exist in water-soluble and membrane-bound forms and transporters that switch between multiple conformations. Analysis of polarity of TM protein surfaces uncovers the presence of a "midpolar" region inside the hydrocarbon region of the lipid bilayer. This membrane region is characterized by the increased dipolarity/polarizability and the hydrogen-bonding acceptor capacity and can be mapped by the preferential accumulation of indole rings of solvent-exposed Trp residues. The observed asymmetry of "midpolar" regions in proteins from different membranes correlates with asymmetric lipid composition of corresponding membranes. Comparison of TM proteins from different types of natural biological membranes shows protein adaptation to distinct lipid environments and correlation between properties of protein surfaces and thicknesses and polarity asymmetry of target membranes. Distinct hydrophobic thicknesses and polarity profiles obtained from analysis of TM proteins from different membranes can be used to describe anisotropic lipid environment in corresponding natural membranes. These models of biological membranes may be useful for development of advanced computational methods aimed at structural modeling and prediction of spatial localization, topology, and self-translocation ability of peptides, proteins and small drug-like molecules. Analysis of forces and factors that define folding and stability of TM proteins highlights the essential role of the hydrophobic effect that drives protein insertion into membrane and important contributions into protein stability of van der Waals attraction forces, hydrogen bonding and ionic interactions, side-chain conformational entropy, the hydrophobic mismatch, membrane deformations, and specific protein-lipid binding.

References

1. Tan S, Tan HT, Chung MC (2008) Membrane proteins and membrane proteomics. Proteomics 8:3924–3932.
2. Bill RM, Henderson PJF, Iwata S, Kunji ERS, Michel H, Neutze R, Newstead S, Poolman B, Tate CG, Vogel H (2011) Overcoming barriers to membrane protein structure determination. Nat Biotechnol 29:335–340.
3. Johansson LC, Arnlund D, White TA, Katona G, Deponte DP, Weierstall U, Doak RB, Shoeman RL, Lomb L, Malmerberg E, Davidsson J, Nass K, Liang M, Andreasson J, Aquila A, Bajt S, Barthelmess M, Barty A, Bogan MJ, Bostedt C, Bozek JD, Caleman C, Coffee R, Coppola N, Ekeberg T, Epp SW, Erk B, Fleckenstein H, Foucar L, Graafsma H, Gumprecht L, Hajdu J, Hampton CY, Hartmann R, Hartmann A, Hauser G, Hirsemann H, Holl P, Hunter MS, Kassemeyer S, Kimmel N, Kirian RA, Maia FR, Marchesini S, Martin AV, Reich C, Rolles D, Rudek B, Rudenko A, Schlichting I, Schulz J, Seibert MM, Sierra RG, Soltau H, Starodub D, Stellato F, Stern S, Struder L, Timneanu N, Ullrich J, Wahlgren WY, Wang X, Weidenspointner G, Wunderer C, Fromme P, Chapman HN, Spence JC, Neutze R (2012) Lipidic phase membrane protein serial femtosecond crystallography. Nat Methods 9:263–265.
4. Shahid SA, Bardiaux B, Franks WT, Krabben L, Habeck M, van Rossum BJ, Linke D (2012) Membrane-protein structure determination by solid-state NMR spectroscopy of microcrystals. Nat Methods 9: 1212–1217.
5. Hiller S, Garces RG, Malia TJ, Orekhov VY, Colombini M, Wagner G (2008) Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. Science 321:1206–1210.
6. Gautier A (2014) Structure determination of alpha-helical membrane proteins by solution-state NMR: emphasis on retinal proteins. Biochim Biophys Acta 1837:578–588.
7. Van Horn WD, Kim HJ, Ellis CD, Hadziselimovic A, Sulistijo ES, Karra MD, Tian C, Sonnichsen FD, Sanders CR (2009) Solution nuclear magnetic resonance structure of membrane-integral diacylglycerol kinase. Science 324:1726–1729.
8. Call ME, Chou JJ (2010) A view into the blind spot: solution NMR provides new insights into signal transduction across the lipid bilayer. Structure 18:1559–1569.
9. Gayen S, Li Q, Kang C (2012) Solution NMR study of the transmembrane domain of single-span membrane proteins: opportunities and strategies. Curr Protein Pept Sci 13:585–600.
10. Bocharov EV, Volynsky PE, Pavlov KV, Efremov RG, Arseniev AS (2010) Structure elucidation of dimeric transmembrane domains of bitopic proteins. Cell Adh Migr 4:284–298.
11. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Iype L, Jain S, Fagan P, Marvin J, Padolla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C (2002) The Protein Data Bank. Acta Crystallogr D Biol Crystallogr 58:899–907.
12. White SH, Snaider C. Database. Available at: http://blanco.biomol.uci.edu/mpstruc White laboratory at UC Irvine. (2014).

13. Raman P, Cherezov V, Caffrey M (2006) The membrane Protein Data Bank. Cell Mol Life Sci 63:36–51.
14. Saier MH, Yen MR, Noto K, Tamang DG, Elkan C (2009) The transporter classification database: recent advances. Nucleic Acids Res 37:D274–D278.
15. Tusnady GE, Dosztanyi Z, Simon I (2005) PDB_TM: selection and membrane localization of transmembrane proteins in the protein data bank. Nucleic Acids Res 33:D275–D278.
16. Sansom MSP, Scott KA, Bond PJ (2008) Coarse-grained simulation: a high-throughput computational approach to membrane proteins. Biochem Soc Trans 36:27–32.
17. Chetwynd AP, Scott KA, Mokrab Y, Sansom MSP (2008) CGDB: a database of membrane protein/lipid interactions by coarse-grained molecular dynamics simulations. Mol Membr Biol 25:662–669.
18. Lomize MA, Pogozheva ID, Joo H, Mosberg HI, Lomize AL (2012) OPM database and PPM web server: resources for positioning of proteins in membranes. Nucleic Acids Res 40:370–376.
19. Lomize AL, Pogozheva ID, Mosberg HI (2011) Anisotropic solvent model of the lipid bilayer. 2. Energetics of insertion of small molecules, peptides, and proteins in membranes. J Chem Inf Model 51:930–946.
20. Pogozheva ID, Tristram-Nagle S, Mosberg HI, Lomize AL (2013) Structural adaptations of proteins to different biological membranes. Biochim Biophys Acta 1828:2592–2608.
21. Liu W, Eilers M, Patel AB, Smith SO (2004) Helix packing moments reveal diversity and conservation in membrane protein structure. J Mol Biol 337:713–729.
22. Oberai A, Ihm Y, Kim S, Bowie JU (2006) A limited universe of membrane protein families and folds. Protein Sci 15:1723–1734.
23. Almen MS, Nordstrom KJ, Fredriksson R, Schioth HB (2009) Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. BMC Biol 7:50.
24. Worch R, Bokel C, Hofinger S, Schwille P, Weidemann T (2010) Focus on composition and interaction potential of single-pass transmembrane domains. Proteomics 10:4196–4208.
25. Wallin E, von Heijne G (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. Protein Sci 7:1029–1038.
26. Elofsson A, von Heijne G (2007) Membrane protein structure: prediction versus reality. Annu Rev Biochem 76:125–140.
27. Vinothkumar KR, Henderson R (2010) Structures of membrane proteins. Q Rev Biophys 43:65–158.
28. McLuskey K, Roszak AW, Zhu Y, Isaacs NW (2010) Crystal structures of all-alpha type membrane proteins. Eur Biophys J 39:723–755.
29. Arce J, Sturgis JN, Duneau J-P (2009) Dissecting membrane protein architecture: an annotation of structural complexity. Biopolymers 91:815–829.
30. Hedin LE, Illergard K, Elofsson A (2011) An introduction to membrane proteins. J Proteome Res 10:3324–3331.
31. Bowie JU (1997) Helix packing in membrane proteins. J Mol Biol 272:780–789.
32. Gimpelev M, Forrest LR, Murray D, Honig B (2004) Helical packing patterns in membrane and soluble proteins. Biophys J 87:4075–4086.
33. Eyre TA, Partridge L, Thornton JM (2004) Computational analysis of alpha-helical membrane protein structure: implications for the prediction of 3D structural models. Protein Eng Des Sel 17:613–624.
34. Popot JL, Engelman DM (2000) Helical membrane protein folding, stability, and evolution. Annu Rev Biochem 69:881–922.
35. Yan C, Luo J (2010) An analysis of reentrant loops. Protein J 29:350–354.
36. Pornillos O, Chang G (2006) Inverted repeat domains in membrane proteins. FEBS Lett 580:358–362.
37. von Heijne G (2006) Membrane-protein topology. Nat Rev Mol Cell Biol 7:909–918.
38. Ojemalm K, Halling KK, Nilsson I, von Heijne G (2012) Orientational preferences of neighboring helices can drive er insertion of a marginally hydrophobic transmembrane helix. Mol Cell 45:529–540.
39. Zhao G, London E (2006) An amino acid "transmembrane tendency" scale that approaches the theoretical limit to accuracy for prediction of transmembrane helices: relationship to biological hydrophobicity. Protein Sci 15:1987–2001.
40. Hedin LE, Ojemalm K, Bernsel A, Hennerdal A, Illergard K, Enquist K, Kauko A, Cristobal S, von Heijne G, Lerch-Bader M, Nilsson I, Elofsson A (2010) Membrane insertion of marginally hydrophobic transmembrane helices depends on sequence context. J Mol Biol 396:221–229.
41. Rath EM, Tessier D, Campbell AA, Lee HC, Werner T, Salam NK, Lee LK, Church WB (2013) A benchmark server using high resolution protein structure data, and benchmark results for membrane helix predictions. BMC Bioinform 14:111.
42. Rath A, Deber CM (2012) Protein structure in membrane domains. Annu Rev Biophys 41:135–155.
43. Hessa T, Meindl-Beinker NM, Bernsel A, Kim H, Sato Y, Lerch-Bader M, Nilsson I, White SH, von Heijne G (2007) Molecular code for transmembrane-helix recognition by the Sec61 translocon. Nature 450:1026–1030.
44. von Heijne G (2011) Membrane proteins: from bench to bits. Biochem Soc Trans 39:747–750.
45. Bano-Polo M, Baeza-Delgado C, Orzaez M, Marti-Renom MA, Abad C, Mingarro I (2012) Polar/ionizable residues in transmembrane segments: effects on helix–helix packing. PLoS One 7:e44263.
46. Herrmann JR, Fuchs A, Panitz JC, Eckert T, Unterreitmeier S, Frishman D, Langosch D (2010) Ionic interactions promote transmembrane helix-helix association depending on sequence context. J Mol Biol 396:452–461.
47. Meindl-Beinker NM, Lundin C, Nilsson I, White SH, von Heijne G (2006) Asn- and Asp-mediated interactions between transmembrane helices during translocon-mediated membrane protein assembly. EMBO Rep 7:1111–1116.
48. Baeza-Delgado C, Marti-Renom MA, Mingarro I (2013) Structure-based statistical analysis of transmembrane helices. Eur Biophys J 42:199–207.
49. Viklund H, Granseth E, Elofsson A (2006) Structural classification and prediction of reentrant regions in alpha-helical transmembrane proteins: application to complete genomes. J Mol Biol 361:591–603.
50. Screpanti E, Hunte C (2007) Discontinuous membrane helices in transport proteins and their correlation with function. J Struct Biol 159:261–267.
51. Kazi A, Sun JZ, Doi K, Sung SS, Takahashi Y, Yin H, Rodriguez JM, Becerril J, Berndt N, Hamilton AD, Wang HG, Sebti SM (2011) The BH3 alpha-helical mimic BH3-M6 disrupts Bcl-X(L), Bcl-2, and MCL-1 protein-protein interactions with Bax, Bbak, Bad, or

Bim and induces apoptosis in a Bax- and Bim-dependent manner. J Biol Chem 286:9382–9392.
52. Rapp M, Seppala S, Granseth E, von Heijne G (2007) Emulating membrane protein evolution by rational design. Science 315:1282–1284.
53. Zeth K, Thein M (2010) Porins in prokaryotes and eukaryotes: common themes and variations. Biochem J 431:13–22.
54. Wimley WC (2002) Toward genomic identification of beta-barrel membrane proteins: composition and architecture of known structures. Protein Sci 11:301–312.
55. Zhai YF, Saier MH (2002) The beta-barrel finder (BBF) program, allowing identification of outer membrane beta-barrel proteins encoded within prokaryotic genomes. Protein Sci 11:2196–2207.
56. Buchanan SK (1999) Beta-barrel proteins from bacterial outer membranes: structure, function and refolding. Curr Opin Struct Biol 9:455–461.
57. Schulz GE (2000) Beta-barrel membrane proteins. Curr Opin Struct Biol 10:443–447.
58. Schulz GE (2002) The structure of bacterial outer membrane proteins. Biochim Biophys Acta 1565:308–317.
59. Fairman JW, Noinaj N, Buchanan SK (2011) The structural biology of beta-barrel membrane proteins: a summary of recent reports. Curr Opin Struct Biol 21: 523–531.
60. Buchanan SK, Yamashita Y, Fleming KG. Structure and folding of outer membrane proteins. In: Engelman EH, Tamm LK, Eds. (2012) Comprehensive biophysics. Oxford: Academic Press, pp 139–163.
61. Noinaj N, Guillier M, Barnard TJ, Buchanan SK (2010) TonB-dependent transporters: regulation, structure, and function. Annu Rev Microbiol 64:43–60.
62. Bay DC, Hafez M, Young MJ, Court DA (2012) Phylogenetic and coevolutionary analysis of the beta-barrel protein family comprised of mitochondrial porin (VDAC) and Tom40. Biochim Biophys Acta 1818: 1502–1519.
63. Bayrhuber M, Meins T, Habeck M, Becker S, Giller K, Villinger S, Vonrhein C, Griesinger C, Zweckstetter M, Zeth K (2008) Structure of the human voltage-dependent anion channel. Proc Natl Acad Sci USA 105:15370–15375.
64. Ujwal R, Cascio D, Colletier JP, Faham S, Zhang J, Toro L, Ping PP, Abramson J (2008) The crystal structure of mouse VDAC1 at 2.3 angstrom resolution reveals mechanistic insights into metabolite gating. Proc Natl Acad Sci USA 105:17742–17747.
65. Schredelseker J, Paz A, Lopez CJ, Altenbach C, Leung CS, Drexler MK, Chen JN, Hubbell WL, Abramson J (2014) High-resolution structure and double electron-electron resonance of the zebrafish voltage dependent anion channel 2 reveal an oligomeric population. J Biol Chem 289:12566–12577.
66. Andersen C (2003) Channel-tunnels: outer membrane components of type I secretion systems and multidrug efflux pumps of Gram-negative bacteria. Rev Physiol Biochem Pharmacol 147:122–165.
67. Andersen C, Hughes C, Koronakis V (2001) Protein export and drug efflux through bacterial channel-tunnels. Curr Opin Cell Biol 13:412–416.
68. Meng G, Surana NK, St Geme JW, III, Waksman G (2006) Structure of the outer membrane translocator domain of the haemophilus influenzae Hia trimeric autotransporter. EMBO J 25:2297–2304.
69. Faller M, Niederweis M, Schulz GE (2004) The structure of a mycobacterial outer-membrane channel. Science 303:1189–1192.
70. Bischofberger M, Iacovache I, van der Goot FG (2012) Pathogenic pore-forming proteins: function and host response. Cell Host Microbe 12:266–275.
71. Iacovache I, Degiacomi MT, van der Goot FG. Pore-forming toxins. In: Egelman EH, Ed. (2012) Comprehensive biophysics. Elsevier B.V, Academic Press, New York, v.5, pp 164–188.
72. De S, Olson R (2011) Crystal structure of the vibrio cholerae cytolysin heptamer reveals common features among disparate pore-forming toxins. Proc Natl Acad Sci USA 108:7385–7390.
73. Song L, Hobaugh MR, Shustak C, Cheley S, Bayley H, Gouaux JE (1996) Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. Science 274:1859–1866.
74. Savva CG, Fernandes da Costa SP, Bokori-Brown M, Naylor CE, Cole AR, Moss DS, Titball RW, Basak AK (2013) Molecular architecture and functional analysis of NetB, a pore-forming toxin from clostridium perfringens. J Biol Chem 288:3512–3522.
75. Yamashita K, Kawai Y, Tanaka Y, Hirano N, Kaneko J, Tomita N, Ohta M, Kamio Y, Yao M, Tanaka I (2011) Crystal structure of the octameric pore of staphylococcal gamma-hemolysin reveals the beta-barrel pore formation mechanism by two components. Proc Natl Acad Sci USA 108:17314–17319.
76. Kelkar DA, Chattopadhyay A (2007) The gramicidin ion channel: a model membrane protein. Biochim Biophys Acta 1768:2011–2025.
77. Wallace BA (1990) Gramicidin channels and pores. Annu Rev Biophys Biophy Chem 19:127–157.
78. Arseniev AS, Barsukov IL, Bystrov VF, Lomize AL, Ovchinnikov Yu A (1985) ^{1}H NMR study of gramicidin a transmembrane ion channel. Head-to-head right-handed, single-stranded helices. FEBS Lett 186:168–174.
79. Lomize AL, Orekhov V, Arseniev AS (1992) Refinement of the spatial structure of the gramicidin a ion channel. Bioorg Khim 18:182–200.
80. Ketchem RR, Lee KC, Huo S, Cross TA (1996) Macromolecular structural elucidation with solid-state NMR-derived orientational constraints. J Biomol NMR 8:1–14.
81. Smart OS, Goodfellow JM, Wallace BA (1993) The pore dimensions of gramicidin a. Biophys J 65:2455–2460.
82. Urry DW, Goodall MC, Glickson JD, Mayers DF (1971) The gramicidin a transmembrane channel: characteristics of head-to-head dimerized (l,d) helices. Proc Natl Acad Sci USA 68:1907–1911.
83. Sun H, Greathouse DV, Andersen OS, Koeppe RE (2008) The preference of tryptophan for membrane interfaces—insights from *N*-methylation of tryptophans in gramicidin channels. J Biol Chem 283: 22233–22243.
84. Koeppe RE, Schmutzer SE, Andersen OS. Gramicidin channels as cation nanotubes. In: Hayden O, Nielsch K, Eds. (2011) Molecular- and nano-tubes. Berlin: Springer-Verlag Berlin, pp 11–30.
85. Killian JA (2003) Synthetic peptides as models for intrinsic membrane proteins. FEBS Lett 555:134–138.
86. Deber CM, Liu LP, Wang C (1999) Perspective: peptides as mimics of transmembrane segments in proteins. J Pept Res 54:200–205.
87. Booth PJ, Curnow P (2006) Membrane proteins shape up: understanding in vitro folding. Curr Opin Struct Biol 16:480–488.

88. Harris NJ, Booth PJ (2012) Folding and stability of membrane transport proteins in vitro. Biochim Biophys Acta 1818:1055–1066.
89. DiBartolo ND, Booth PJ. The membrane factor: biophysical studies of alpha helical transmembrane protein folding. In: Egelman EH, Ed. (2012) Comprehensive biophysics. Elsevier B.V., Academic Press, New York, v.3, pp 290–316.
90. Dalbey RE, Kuhn A (2000) Evolutionarily related insertion pathways of bacterial, mitochondrial, and thylakoid membrane proteins. Annu Rev Cell Dev Biol 16:51–87.
91. Dalbey RE, Wang P, Kuhn A (2011) Assembly of bacterial inner membrane proteins. Annu Rev Biochem 80:161–187.
92. Zimmer J, Nam Y, Rapoport TA (2008) Structure of a complex of the ATPase SecA and the protein-translocation channel. Nature 455:936–943.
93. Van den Berg B, Clemons WM, Jr, Collinson I, Modis Y, Hartmann E, Harrison SC, Rapoport TA (2004) X-ray structure of a protein-conducting channel. Nature 427:36–44.
94. Gogala M, Becker T, Beatrix B, Armache JP, Barrio-Garcia C, Berninghausen O, Beckmann R (2014) Structures of the Sec61 complex engaged in nascent peptide translocation or membrane insertion. Nature 506:107–110.
95. White SH, von Heijne G (2004) The machinery of membrane protein assembly. Curr Opin Struct Biol 14:397–404.
96. Bowie JU (2005) Solving the membrane protein folding problem. Nature 438:581–589.
97. Park E, Rapoport TA (2012) Mechanisms of Sec61/SecY-mediated protein translocation across membranes. Annu Rev Biophys 41:21–40.
98. Zhang L, Paakkarinen V, Suorsa M, Aro EM (2001) A SecY homologue is involved in chloroplast-encoded D1 protein biogenesis. J Biol Chem 276:37809–37814.
99. Tsukazaki T, Mori H, Echizen Y, Ishitani R, Fukai S, Tanaka T, Perederina A, Vassylyev DG, Kohno T, Maturana AD, Ito K, Nureki O (2011) Structure and function of a membrane component SecDF that enhances protein export. Nature 474:235–238.
100. Kol S, Nouwen N, Driessen AJ (2008) Mechanisms of YidC-mediated insertion and assembly of multimeric membrane protein complexes. J Biol Chem 283: 31269–31273.
101. van der Laan M, Bechtluft P, Kol S, Nouwen N, Driessen AJ (2004) F1F0 ATP synthase subunit c is a substrate of the novel YidC pathway for membrane protein biogenesis. J Cell Biol 165:213–222.
102. van der Laan M, Nouwen NP, Driessen AJ (2005) YidC—an evolutionary conserved device for the assembly of energy-transducing membrane protein complexes. Curr Opin Microbiol 8:182–187.
103. Kiefer D, Kuhn A (2007) YidC as an essential and multifunctional component in membrane protein assembly. Int Rev Cytol 259:113–138.
104. Kruger V, Deckers M, Hildenbeutel M, van der Laan M, Hellmers M, Dreker C, Preuss M, Herrmann JM, Rehling P, Wagner R, Meinecke M (2012) The mitochondrial oxidase assembly protein1 (Oxa1) insertase forms a membrane pore in lipid bilayers. J Biol Chem 287:33314–33326.
105. Herrmann JM, Neupert W, Stuart RA (1997) Insertion into the mitochondrial inner membrane of a polytopic protein, the nuclear-encoded Oxa1p. EMBO J 16:2217–2226.
106. Jermy A (2012) Bacterial secretion: Sec and Tat collaborate in a Rieske business. Nat Rev Microbiol 10:801.
107. Keller R, de Keyzer J, Driessen AJ, Palmer T (2012) Co-operation between different targeting pathways during integration of a membrane protein. J Cell Biol 199:303–315.
108. Hagan CL, Silhavy TJ, Kahne D (2011) Beta-barrel membrane protein assembly by the Bam complex. Annu Rev Biochem 80:189–210.
109. Rigel NW, Silhavy TJ (2012) Making a beta-barrel: assembly of outer membrane proteins in gram-negative bacteria. Curr Opin Microbiol 15:189–193.
110. Borgese N, Fasana E (2011) Targeting pathways of c-tail-anchored proteins. Biochim Biophys Acta 1808: 937–946.
111. Hegde RS, Keenan RJ (2011) Tail-anchored membrane protein insertion into the endoplasmic reticulum. Nat Rev Mol Cell Biol 12:787–798.
112. Waizenegger T, Stan T, Neupert W, Rapaport D (2003) Signal-anchor domains of proteins of the outer membrane of mitochondria: structural and functional characteristics. J Biol Chem 278:42064–42071.
113. Kanaji S, Iwahashi J, Kida Y, Sakaguchi M, Mihara K (2000) Characterization of the signal that directs Tom20 to the mitochondrial outer membrane. J Cell Biol 151:277–288.
114. Fox TD (2012) Mitochondrial protein synthesis, import, and assembly. Genetics 192:1203–1234.
115. von Heijne G, Steppuhn J, Herrmann RG (1989) Domain structure of mitochondrial and chloroplast targeting peptides. Eur J Biochem 180:535–545.
116. Rehling P, Brandner K, Pfanner N (2004) Mitochondrial import and the twin-pore translocase. Nat Rev Mol Cell Biol 5:519–530.
117. Dukanovic J, Rapaport D (2011) Multiple pathways in the integration of proteins into the mitochondrial outer membrane. Biochim Biophys Acta 1808:971–980.
118. Merklinger E, Gofman Y, Kedrov A, Driessen AJ, Ben-Tal N, Shai Y, Rapaport D (2012) Membrane integration of a mitochondrial signal-anchored protein does not require additional proteinaceous factors. Biochem J 442:381–389.
119. Kemper C, Habib SJ, Engl G, Heckmeyer P, Dimmer KS, Rapaport D (2008) Integration of tail-anchored proteins into the mitochondrial outer membrane does not require any known import components. J Cell Sci 121:1990–1998.
120. Dimmer KS, Rapaport D (2012) Unresolved mysteries in the biogenesis of mitochondrial membrane proteins. Biochim Biophys Acta 1818:1085–1090.
121. Kovermann P, Truscott KN, Guiard B, Rehling P, Sepuri NB, Muller H, Jensen RE, Wagner R, Pfanner N (2002) Tim22, the essential core of the mitochondrial protein insertion complex, forms a voltage-activated and signal-gated channel. Mol Cell 9:363–373.
122. Lee DW, Kim JK, Lee S, Choi S, Kim S, Hwang I (2008) Arabidopsis nuclear-encoded plastid transit peptides contain multiple sequence subgroups with distinctive chloroplast-targeting sequence motifs. Plant Cell 20:1603–1622.
123. Wickner W, Schekman R (2005) Protein translocation across biological membranes. Science 310:1452–1456.
124. Soll J, Schleiff E (2004) Protein import into chloroplasts. Nat Rev Mol Cell Biol 5:198–208.
125. Robinson C, Thompson SJ, Woolhead C (2001) Multiple pathways used for the targeting of thylakoid proteins in chloroplasts. Traffic 2:245–251.
126. Skalitzky CA, Martin JR, Harwood JH, Beirne JJ, Adamczyk BJ, Heck GR, Cline K, Fernandez DE

(2011) Plastids contain a second Sec translocase system with essential functions. Plant Physiol 155:354–369.
127. Jaru-Ampornpan P, Shen K, Lam VQ, Ali M, Doniach S, Jia TZ, Shan SO (2010) ATP-independent reversal of a membrane protein aggregate by a chloroplast Srp subunit. Nat Struct Mol Biol 17:696–702.
128. Tu CJ, Schuenemann D, Hoffman NE (1999) Chloroplast ftsy, chloroplast signal recognition particle, and GTP are required to reconstitute the soluble phase of light-harvesting chlorophyll protein transport into thylakoid membranes. J Biol Chem 274:27219–27224.
129. Cline K, Dabney-Smith C (2008) Plastid protein import and sorting: different paths to the same compartments. Curr Opin Plant Biol 11:585–592.
130. Kapazoglou A, Sagliocco F, Dure L, III (1995) PSII-T, a new nuclear encoded lumenal protein from photosystem II. Targeting and processing in isolated chloroplasts. J Biol Chem 270:12197–12202.
131. Lomize AL, Pogozheva ID, Lomize MA, Mosberg HI (2006) Positioning of proteins in membranes: a computational approach. Protein Sci 15:1318–1333.
132. Mitra K, Ubarretxena-Belandia T, Taguchi T, Warren G, Engelman DM (2004) Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. Proc Natl Acad Sci USA 101:4083–4088.
133. Borgese N, Brambillasca S, Soffientini P, Yabal M, Makarow M (2003) Biogenesis of tail-anchored proteins. Biochem Soc Trans 31:1238–1242.
134. Fernandez C, Hilty C, Wider G, Wuthrich K (2002) Lipid-protein interactions in DHPC micelles containing the integral membrane protein OmpX investigated by NMR spectroscopy. Proc Natl Acad Sci USA 99:13533–13537.
135. Snijder HJ, Timmins PA, Kalk KH, Dijkstra BW (2003) Detergent organisation in crystals of monomeric outer membrane phospholipase A. J Struct Biol 141:122–131.
136. Abraham T, Schooling SR, Nieh MP, Kučerka N, Beveridge TJ, Katsaras J (2007) Neutron diffraction study of pseudomonas aeruginosa lipopolysaccharide bilayers. J Phys Chem B 111:2477–2483.
137. Kucerka N, Papp-Szabo E, Nieh MP, Harroun TA, Schooling SR, Pencer J, Nicholson EA, Beveridge TJ, Katsaras J (2008) Effect of cations on the structure of bilayers formed by lipopolysaccharides isolated from pseudomonas aeruginosa PAO1. J Physl Chem B 112: 8057–8062.
138. Tamm LK, Hong H, Liang BY (2004) Folding and assembly of beta-barrel membrane proteins. Biochim Biophys Acta 1666:250–263.
139. Niederweis M, Danilchanka O, Huff J, Hoffmann C, Engelhardt H (2009) Mycobacterial outer membranes: in search of proteins. Trends in Microbiol 18:109–116.
140. Fiedler S, Broecker J, Keller S (2010) Protein folding in membranes. Cell Mol Life Sci 67:1779–1798.
141. Marsh D (2007) Lateral pressure profile, spontaneous curvature frustration, and the incorporation and conformation of proteins in membranes. Biophys J 93: 3884–3899.
142. Cantor RS (1997) Lateral pressures in cell membranes: a mechanism for modulation of protein function. J Phys Chem B 101:1723–1725.
143. Cantor RS (1999) Lipid composition and the lateral pressure profile in bilayers. Biophys J 76:2625–2639.
144. Pabst G, Kucerka N, Nieh MP, Rheinstadter MC, Katsaras J (2010) Applications of neutron and X-ray scattering to the study of biologically relevant model membranes. Chem Phys Lipids 163:460–479.
145. Raghunathan M, Zubovski Y, Venable RM, Pastor RW, Nagle JF, Tristram-Nagle S (2012) Structure and elasticity of lipid membranes with genistein and daidzein bioflavinoids using X-ray scattering and MD simulations. J Phys Chem B 116:3918–3927.
146. Boscia AL, Treece BW, Mohammadyani D, Klein-Seetharaman J, Braun AR, Wassenaar TA, Klosgen B, Tristram-Nagle S (2014) X-ray structure, thermodynamics, elastic properties and MD simulations of cardiolipin/dimyristoylphosphatidylcholine mixed membranes. Chem Phys Lipids 178:1–10.
147. Boscia AL, Akabori K, Benamram Z, Michel JA, Jablin MS, Steckbeck JD, Montelaro RC, Nagle JF, Tristram-Nagle S (2013) Membrane structure correlates to function of LLP2 on the cytoplasmic tail of HIV-1 gp41 protein. Biophys J 105:657–666.
148. Killian JA, von Heijne G (2000) How proteins adapt to a membrane-water interface. Trends Biochem Sci 25:429–434.
149. Lee AG (2003) Lipid–protein interactions in biological membranes: a structural perspective. Biochim Biophys Acta 1612:1–40.
150. Palsdottir H, Hunte C (2004) Lipids in membrane protein structures. Biochim Biophys Acta 1666:2–18.
151. Yau WM, Wimley WC, Gawrisch K, White SH (1998) The preference of tryptophan for membrane interfaces. Biochemistry 37:14713–14718.
152. Hong HD, Park S, Jimenez RHF, Rinehart D, Tamm LK (2007) Role of aromatic side chains in the folding and thermodynamic stability of integral membrane proteins. J Am Chem Soc 129:8320–8327.
153. Liu W, Caffrey M (2006) Interactions of tryptophan, tryptophan peptides, and tryptophan alkyl esters at curved membrane interfaces. Biochemistry 45:11713–11726.
154. Sanders CR, Mittendorf KF (2011) Tolerance to changes in membrane lipid composition as a selected trait of membrane proteins. Biochemistry 50:7858–7867.
155. Hanshaw RG, Stahelin RV, Smith BD (2008) Noncovalent keystone interactions controlling biomembrane structure. Chemistry 14:1690–1697.
156. Nikaido H. Outer membranes, Gram-negative bacteria. In: Schaechter M, Ed. (2009) Encyclopedia of microbiology. Oxford, UK: Academic Press, pp 439–452.
157. van Meer G, Voelker DR, Feigenson GW (2008) Membrane lipids: where they are and how they behave. Nat Rev Mol Cell Biol 9:112–124.
158. Maxfield FR, van Meer G (2010) Cholesterol, the central lipid of mammalian cells. Curr Opin Cell Biol 22: 422–429.
159. Chen YJ, Pornillos O, Lieu S, Ma C, Chen AP, Chang G (2007) X-ray structure of EmrE supports dual topology model. Proc Natl Acad Sci USA 104:18999–19004.
160. Korkhov VM, Tate CG (2009) An emerging consensus for the structure of EmrE. Acta Crystallogr D 65:186–192.
161. von Heijne G, Gavel Y (1988) Topogenic signals in integral membrane proteins. Eur J Biochem 174:671–678.
162. Kauko A, Hedin LE, Thebaud E, Cristobal S, Elofsson A, von Heijne G (2010) Repositioning of transmembrane alpha-helices during membrane protein folding. J Mol Biol 397:190–201.
163. Pitonzo D, Skach WR (2006) Molecular mechanisms of aquaporin biogenesis by the endoplasmic reticulum Sec61 translocon. Biochim Biophys Acta 1758:976–988.
164. Seppala S, Slusky JS, Lloris-Garcera P, Rapp M, von Heijne G (2010) Control of membrane protein topology

by a single C-terminal residue. Science 328:1698–1700.
165. Ojemalm K, Watson HR, Roboti P, Cross BC, Warwicker J, von Heijne G, High S (2013) Positional editing of transmembrane domains during ion channel assembly. J Cell Sci 26:464–472.
166. Lu Y, Turnbull IR, Bragin A, Carveth K, Verkman AS, Skach WR (2000) Reorientation of aquaporin-1 topology during maturation in the endoplasmic reticulum. Mol Biol Cell 11:2973–2985.
167. Norholm MH, Cunningham F, Deber CM, von Heijne G (2011) Converting a marginally hydrophobic soluble protein into a membrane protein. J Mol Biol 407:171–179.
168. Bano-Polo M, Martinez-Gil L, Wallner B, Nieva JL, Elofsson A, Mingarro I (2013) Charge pair interactions in transmembrane helices and turn propensity of the connecting sequence promote helical hairpin insertion. J Mol Biol 425:830–840.
169. Zhang L, Sato Y, Hessa T, von Heijne G, Lee JK, Kodama I, Sakaguchi M, Uozumi N (2007) Contribution of hydrophobic and electrostatic interactions to the membrane integration of the shaker K+ channel voltage sensor domain. Proc Natl Acad Sci USA 104: 8263–8268.
170. Dowhan W, Bogdanov M (2009) Lipid-dependent membrane protein topogenesis. Ann Rev Biochem 78: 515–540.
171. von Heijne G (2007) Formation of transmembrane helices in vivo—is hydrophobicity all that matters? J Gen Physiol 129:353–356.
172. van Dalen A, de Kruijff B (2004) The role of lipids in membrane insertion and translocation of bacterial proteins. Biochim Biophys Acta 1694:97–109.
173. von Heijne G (1992) Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. J Mol Biol 225:487–494.
174. Nilsson J, Persson B, von Heijne G (2005) Comparative analysis of amino acid distributions in integral membrane proteins from 107 genomes. Proteins Struct Funct Bioinform 60:606–616.
175. von Heijne G (1989) Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. Nature 341:456–458.
176. von Heijne G (1986) The distribution of positively charged residues in bacterial inner membrane-proteins correlates with the trans-membrane topology. EMBO J 5:3021–3027.
177. Dalbey RE (1990) Positively charged residues are important determinants of membrane protein topology. Trends Biochem Sci 15:253–257.
178. Sipos L, von Heijne G (1993) Predicting the topology of eukaryotic membrane proteins. Eur J Biochem 213: 1333–1340.
179. Gavel Y, Steppuhn J, Herrmann R, von Heijne G (1991) The "positive-inside rule" applies to thylakoid membrane proteins. FEBS Lett 282:41–46.
180. Gavel Y, von Heijne G (1992) The distribution of charged amino acids in mitochondrial inner-membrane proteins suggests different modes of membrane integration for nuclearly and mitochondrially encoded proteins. Eur J Biochem 205:1207–1215.
181. Nakashima H, Nishikawa K (1992) The amino acid composition is different between the cytoplasmic and extracellular sides in membrane proteins. FEBS Lett 303:141–146.
182. Ulmschneider MB, Sansom MSP, Di Nola A (2005) Properties of integral membrane protein structures: derivation of an implicit membrane potential. Proteins Struct Funct Bioinform 59:252–265.
183. Schramm CA, Hannigan BT, Donald JE, Keasar C, Saven JG, Degrado WF, Samish I (2012) Knowledge-based potential for positioning membrane-associated structures and assessing residue-specific energetic contributions. Structure 20:924–935.
184. Benning C (2009) Mechanisms of lipid transport involved in organelle biogenesis in plant cells. Annu Rev Cell Develop Biol 25:71–91.
185. Nilsson I, von Heijne G (1990) Fine-tuning the topology of a polytopic membrane protein: role of positively and negatively charged amino acids. Cell 62:1135–1141.
186. Fujita H, Yamagishi M, Kida Y, Sakaguchi M (2011) Positive charges on the translocating polypeptide chain arrest movement through the translocon. J Cell Sci 124:4184–4193.
187. Johansson M, Nilsson I, von Heijne G (1993) Positively charged amino acids placed next to a signal sequence block protein translocation more efficiently in *Escherichia coli* than in mammalian microsomes. Mol Gen Genet 239:251–256.
188. Kiefer D, Hu X, Dalbey R, Kuhn A (1997) Negatively charged amino acid residues play an active role in orienting the Sec-independent Pf3 coat protein in the *Escherichia coli* inner membrane. EMBO J 16:2197–2204.
189. Delgado-Partin VM, Dalbey RE (1998) The proton motive force, acting on acidic residues, promotes translocation of amino-terminal domains of membrane proteins when the hydrophobicity of the translocation signal is low. J Biol Chem 273:9927–9934.
190. Schuenemann TA, Delgado-Nixon VM, Dalbey RE (1999) Direct evidence that the proton motive force inhibits membrane translocation of positively charged residues within membrane proteins. J Biol Chem 274: 6855–6864.
191. van Klompenburg W, Nilsson I, von Heijne G, de Kruijff B (1997) Anionic phospholipids are determinants of membrane protein topology. EMBO J 16: 4261–4266.
192. Andersson H, von Heijne G (1994) Membrane-protein topology—effects of delta-mu(h)+ on the translocation of charged residues explain the positive inside rule. EMBO J 13:2267–2272.
193. van de Vossenberg JL, Albers SV, van der Does C, Driessen AJ, van Klompenburg W (1998) The positive inside rule is not determined by the polarity of the delta psi (transmembrane electrical potential). Mol Microbiol 29:1125–1127.
194. 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.
195. Junne T, Schwede T, Goder V, Spiess M (2007) Mutations in the Sec61p channel affecting signal sequence recognition and membrane protein topology. J Biol Chem 282:3201–33209.
196. Kumazaki K, Chiba S, Takemoto M, Furukawa A, Nishiyama K-i, Sugano Y, Mori T, Dohmae N, Hirata K, Nakada-Nakura Y, Maturana AD, Tanaka Y, Mori H, Sugita Y, Arisaka F, Ito K, Ishitani R, Tsukazaki T, Nureki O (2014) Structural basis of Sec-independent membrane protein insertion by YidC. Nature 509:516–520.
197. Jackups R, Liang J (2005) Interstrand pairing patterns in beta-barrel membrane proteins: the positive-outside rule, aromatic rescue, and strand registration prediction. J Mol Biol 354:979–993.
198. Qu J, Behrens-Kneip S, Holst O, Kleinschmidt JH (2009) Binding regions of outer membrane protein a

in complexes with the periplasmic chaperone Skp. A site-directed fluorescence study. Biochemistry 48: 4926–4936.
199. Cymer F, Schneider D (2012) Oligomerization of polytopic alpha-helical membrane proteins: causes and consequences. Biol Chem 393:1215–1230.
200. Pereira-Leal JB, Levy ED, Teichmann SA (2006) The origins and evolution of functional modules: lessons from protein complexes. Philos Trans R Soc Lond B Biol Sci 361:507–517.
201. Venkatakrishnan AJ, Levy ED, Teichmann SA (2010) Homomeric protein complexes: evolution and assembly. Biochem Soc Trans 38:879–882.
202. Nishi H, Hashimoto K, Madej T, Panchenko AR (2013) Evolutionary, physicochemical, and functional mechanisms of protein homooligomerization. Prog Mol Biol Transl Sci 117:3–24.
203. Marianayagam NJ, Sunde M, Matthews JM (2004) The power of two: protein dimerization in biology. Trends Biochem Sci 29:618–625.
204. Ridder A, Langosch D Transmembrane domains in membrane protein folding, oligomerization, and function. In: Buchner J, Kiefhaber T , Eds. (2008) Protein Folding Handbook. Wiley-VCH Verlag GmbH, Weinheim, Germany, pp 876–918.
205. Langosch D, Arkin IT (2009) Interaction and conformational dynamics of membrane-spanning protein helices. Protein Sci 18:1343–1358.
206. Schneider D, Finger C, Prodohl A, Volkmer T (2007) From interactions of single transmembrane helices to folding of alpha-helical membrane proteins: analyzing transmembrane helix–helix interactions in bacteria. Curr Protein Pept Sci 8:45–61.
207. 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.
208. Curran AR, Engelman DM (2003) Sequence motifs, polar interactions and conformational changes in helical membrane proteins. Curr Opin Struct Biol 13:412–417.
209. Cymer F, Veerappan A, Schneider D (2012) Transmembrane helix–helix interactions are modulated by the sequence context and by lipid bilayer properties. Biochim Biophys Acta 1818:963–973.
210. Melnyk RA, Kim S, Curran AR, Engelman DM, Bowie JU, Deber CM (2004) The affinity of GxxxG motifs in transmembrane helix–helix interactions is modulated by long-range communication. J Biol Chem 279:16591–16597.
211. Li E, Wimley WC, Hristova K (2012) Transmembrane helix dimerization: beyond the search for sequence motifs. Biochim Biophys Acta 1818:183–193.
212. Parton DL, Klingelhoefer JW, Sansom MS (2011) Aggregation of model membrane proteins, modulated by hydrophobic mismatch, membrane curvature, and protein class. Biophys J 101:691–699.
213. Hong H, Bowie JU (2011) Dramatic destabilization of transmembrane helix interactions by features of natural membrane environments. J Am Chem Soc 133: 11389–11398.
214. Cross TA, Sharma M, Yi M, Zhou HX (2011) Influence of solubilizing environments on membrane protein structures. Trends Biochem Sci 36:117–125.
215. Grigoryan G, Moore DT, DeGrado WF (2011) Transmembrane communication: general principles and lessons from the structure and function of the M2 proton channel, K(+) channels, and integrin receptors. Annu Rev Biochem 80:211–237.
216. Ubarretxena-Belandia I, Stokes DL (2012) Membrane protein structure determination by electron crystallography. Curr Opin Struct Biol 22:520–528.
217. Gurevich VV, Gurevich EV (2008) How and why do GPCRs dimerize? Trends Pharmacol Sci 29:234–240.
218. Meng GY, Fronzes R, Chandran V, Remaut H, Waksman G (2009) Protein oligomerization in the bacterial outer membrane. Mol Membr Biol 26:136–145.
219. Bechinger B (2008) A dynamic view of peptides and proteins in membranes. Cell Mol Life Sci 65:3028–3039.
220. Song C, Weichbrodt C, Salnikov ES, Dynowski M, Forsberg BO, Bechinger B, Steinem C, de Groot BL, Zachariae U, Zeth K (2013) Crystal structure and functional mechanism of a human antimicrobial membrane channel. Proc Natl Acad Sci USA 110:4586–4591.
221. Bechinger B, Resende JM, Aisenbrey C (2011) The structural and topological analysis of membrane-associated polypeptides by oriented solid-state nmr spectroscopy: established concepts and novel developments. Biophys Chem 153:115–125.
222. Poschner BC, Fischer K, Herrmann JR, Hofmann MW, Langosch D (2011) Structural features of fusogenic model transmembrane domains that differentially regulate inner and outer leaflet mixing in membrane fusion. Mol Membr Biol 27:1–10.
223. Iacovache I, van der Goot FG, Pernot L (2008) Pore formation: an ancient yet complex form of attack. Biochim Biophys Acta 1778:1611–1623.
224. White SH, Wimley WC, Ladokhin AS, Hristova K (1998) Protein folding in membranes: determining energetics of peptide-bilayer interactions. Energ Biol Macromol B 295:62–87.
225. Wimley WC, White SH (1996) Experimentally determined hydrophobicity scale for proteins at membrane interfaces. Nat Struct Biol 3:842–848.
226. Marcoux J, Wang SC, Politis A, Reading E, Ma J, Biggin PC, Zhou M, Tao H, Zhang Q, Chang G, Morgner N, Robinson CV (2013) Mass spectrometry reveals synergistic effects of nucleotides, lipids, and drugs binding to a multidrug resistance efflux pump. Proc Natl Acad Sci USA 110:9704–9709.
227. Thogersen L, Nissen P (2012) Flexible P-type ATPases interacting with the membrane. Curr Opin Struct Biol 22:491–499.
228. Rollauer SE, Tarry MJ, Graham JE, Jaaskelainen M, Jager F, Johnson S, Krehenbrink M, Liu SM, Lukey MJ, Marcoux J, McDowell MA, Rodriguez F, Roversi P, Stansfeld PJ, Robinson CV, Sansom MS, Palmer T, Högbom M, Berks BC, Lea SM (2012) Structure of the TatC core of the twin-arginine protein transport system. Nature 492:210–214.
229. Shi Y (2013) Common folds and transport mechanisms of secondary active transporters. Annu Rev Biophys 42:51–72.
230. Matthews EE, Zoonens M, Engelman DM (2006) Dynamic helix interactions in transmembrane signaling. Cell 127:447–450.
231. Wang T, Fu G, Pan X, Wu J, Gong X, Wang J, Shi Y (2013) Structure of a bacterial energy-coupling factor transporter. Nature 497:272–276.
232. Xu K, Zhang M, Zhao Q, Yu F, Guo H, Wang C, He F, Ding J, Zhang P (2013) Crystal structure of a folate energy-coupling factor transporter from *Lactobacillus brevis*. Nature 497:268–271.
233. Palmer T, Berks BC (2012) The twin-arginine translocation (Tat) protein export pathway. Nat Rev Microbiol 10:483–496.
234. Frobel J, Rose P, Muller M (2012) Twin-arginine-dependent translocation of folded proteins. Philos Trans R Soc Lond B Biol Sci 367:1029–1046.

235. Gohlke U, Pullan L, McDevitt CA, Porcelli I, de Leeuw E, Palmer T, Saibil HR, Berks BC (2005) The tata component of the twin-arginine protein transport system forms channel complexes of variable diameter. Proc Natl Acad Sci USA 102:10482–10486.
236. White SH (2003) Translocons, thermodynamics, and the folding of membrane proteins. FEBS Lett 555: 116–121.
237. Haltia T, Freire E (1995) Forces and factors that contribute to the structural stability of membrane proteins. Biochim Biophys Acta 1241:295–322.
238. White SH, Wimley WC (1999) Membrane protein folding and stability: physical principles. Annu Rev Biophys Biomol Struct 28:319–365.
239. Booth PJ, Templer RH, Meijberg W, Allen SJ, Curran AR, Lorch M (2001) In vitro studies of membrane protein folding. Crit Rev Biochem Mol Biol 36:501–603.
240. Minetti CA, Remeta DP (2006) Energetics of membrane protein folding and stability. Arch Biochem Biophys 453:32–53.
241. MacKenzie KR (2006) Folding and stability of alpha-helical integral membrane proteins. Chem Rev 106: 1931–1977.
242. Stanley AM, Fleming KG (2008) The process of folding proteins into membranes: challenges and progress. Arch Biochem Biophys 469:46–66.
243. Bowie JU (2011) Membrane protein folding: how important are hydrogen bonds? Curr Opin Struct Biol 21:42–49.
244. Chakrabartty A, Baldwin RL (1995) Stability of alpha-helices. Adv Protein Chem 46:141–176.
245. Lomize AL, Mosberg HI (1997) Thermodynamic model of secondary structure for alpha-helical peptides and proteins. Biopolymers 42:239–269.
246. Makhatadze GI (2005) Thermodynamics of alpha-helix formation. Adv Protein Chem 72:199–226.
247. Ladokhin AS, White SH (2001) Protein chemistry at membrane interfaces: non-additivity of electrostatic and hydrophobic interactions. J Mol Biol 309:543–552.
248. Seelig J (2004) Thermodynamics of lipid-peptide interactions. Biochim Biophys Acta 1666:40–50.
249. Ziegler A (2008) Thermodynamic studies and binding mechanisms of cell-penetrating peptides with lipids and glycosaminoglycans. Adv Drug Deliv Rev 60:580–597.
250. Reshetnyak YK, Andreev OA, Segala M, Markin VS, Engelman DM (2008) Energetics of peptide (PHLIP) binding to and folding across a lipid bilayer membrane. Proc Natl Acad Sci USA 105:15340–15345.
251. Wimley WC, White SH (2000) Designing transmembrane alpha-helices that insert spontaneously. Biochemistry 39:4432–4442.
252. Kyrychenko A, Rodnin MV, Posokhov YO, Holt A, Pucci B, Killian JA, Ladokhin AS (2012) Thermodynamic measurements of bilayer insertion of a single transmembrane helix chaperoned by fluorinated surfactants. J Mol Biol 416:328–334.
253. Lomize AL, Pogozheva ID, Lomize MA, Mosberg HI (2007) The role of hydrophobic interactions in positioning of peripheral proteins in membranes. BMC Struct Biol 7:44.
254. MacKenzie KR, Fleming KG (2008) Association energetics of membrane spanning alpha-helices. Curr Opin Struct Biol 18:412–419.
255. Booth PJ, Curnow P (2009) Folding scene investigation: membrane proteins. Curr Opin Struct Biol 19:8–13.
256. Lomize AL, Pogozheva ID, Mosberg HI (2011) Anisotropic solvent model of the lipid bilayer. 1. Parameterization of long-range electrostatics and first solvation shell effects. J Chem Inf Model 51:918–929.
257. Lomize AL, Pogozheva ID, Mosberg HI (2004) Quantification of helix-helix binding affinities in micelles and lipid bilayers. Protein Sci 13:2600–2612.
258. White SH, Wimley WC (1998) Hydrophobic interactions of peptides with membrane interfaces. Biochim Biophys Acta 1376:339–352.
259. Moon CP, Fleming KG (2011) Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. Proc Natl Acad Sci USA 108: 10174–10177.
260. Adamian L, Nanda V, DeGrado WF, Liang J (2005) Empirical lipid propensities of amino acid residues in multispan alpha helical membrane proteins. Proteins Struct Funct Bioinform 59:496–509.
261. Koehler J, Woetzel N, Staritzbichler R, Sanders CR, Meiler J (2009) A unified hydrophobicity scale for multispan membrane proteins. Proteins Struct Funct Bioinform 76:13–29.
262. Senes A, Chadi DC, Law PB, Walters RFS, Nanda V, DeGrado WF (2007) E-z, a depth-dependent potential for assessing the energies of insertion of amino acid side-chains into membranes: derivation and applications to determining the orientation of transmembrane and interfacial helices. J Mol Biol 366:436–448.
263. Hsieh D, Davis A, Nanda V (2012) A knowledge-based potential highlights unique features of membrane alpha-helical and beta-barrel protein insertion and folding. Protein Sci 21:50–62.
264. Schow EV, Freites JA, Cheng P, Bernsel A, von Heijne G, White SH, Tobias DJ (2011) Arginine in membranes: the connection between molecular dynamics simulations and translocon-mediated insertion experiments. J Membr Biol 239:35–48.
265. Johansson ACV, Lindahl E (2008) Position-resolved free energy of solvation for amino acids in lipid membranes from molecular dynamics simulations. Proteins Struct Funct Bioinform 70:1332–1344.
266. MacCallum JL, Bennett WFD, Tieleman DP (2007) Partitioning of amino acid side chains into lipid bilayers: results from computer simulations and comparison to experiment. J Gen Physiol 129:371–377.
267. MacCallum JL, Bennett WF, Tieleman DP (2008) Distribution of amino acids in a lipid bilayer from computer simulations. Biophys J 94:3393–3404.
268. Zhao G, London E (2009) Strong correlation between statistical transmembrane tendency and experimental hydrophobicity scales for identification of transmembrane helices. J Membr Biol 229:165–168.
269. MacCallum JL, Tieleman DP (2011) Hydrophobicity scales: a thermodynamic looking glass into lipid–protein interactions. Trends Biochem Sci 36:653–662.
270. Hessa T, Kim H, Bihlmaier K, Lundin C, Boekel J, Andersson H, Nilsson I, White SH, von Heijne G (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. Nature 433:377–381.
271. Hessa T, Reithinger JH, von Heijne G, Kim H (2009) Analysis of transmembrane helix integration in the endoplasmic reticulum in *S. cerevisiae*. J Mol Biol 386:1222–1228.
272. Xie K, Hessa T, Seppala S, Rapp M, von Heijne G, Dalbey RE (2007) Features of transmembrane segments that promote the lateral release from the translocase into the lipid phase. Biochemistry 46: 15153–15161.
273. Botelho SC, Osterberg M, Reichert AS, Yamano K, Bjorkholm P, Endo T, von Heijne G, Kim H (2011) Tim23-mediated insertion of transmembrane alpha-

helices into the mitochondrial inner membrane. EMBO J 30:1003–1011.
274. White SH, von Heijne G (2008) How translocons select transmembrane helices. Annu Rev Biophys 37: 23–42.
275. Fagerberg L, Jonasson K, von Heijne G, Uhlen M, Berglund L (2010) Prediction of the human membrane proteome. Proteomics 10:1141–1149.
276. Bernsel A, Viklund H, Falk J, Lindahl E, von Heijne G, Elofsson A (2008) Prediction of membrane-protein topology from first principles. Proc Natl Acad Sci USA 105:7177–7181.
277. Mitchell DC (2012) Progress in understanding the role of lipids in membrane protein folding. Biochim Biophys Acta 1818:951–956.
278. Andersen OS, Koeppe RE (2007) Bilayer thickness and membrane protein function: an energetic perspective. Annu Rev Biophys Biomol Struct 36:107–130.
279. Marsh D (2008) Protein modulation of lipids, and vice-versa, in membranes. Biochim Biophys Acta 1778:1545–1575.
280. Phillips R, Ursell T, Wiggins P, Sens P (2009) Emerging roles for lipids in shaping membrane-protein function. Nature 459:379–385.
281. Yano Y, Yamamoto A, Ogura M, Matsuzaki K (2011) Thermodynamics of insertion and self-association of a transmembrane helix: a lipophobic interaction by phosphatidylethanolamine. Biochemistry 50:6806–6814.
282. 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.
283. Israelachvili JN (1992) Intermolecular and surface forces. London: Academic Press.
284. Faham S, Yang D, Bare E, Yohannan S, Whitelegge JP, Bowie JU (2004) Side-chain contributions to membrane protein structure and stability. J Mol Biol 335: 297–305.
285. Lomize AL, Reibarkh MY, Pogozheva ID (2002) Interatomic potentials and solvation parameters from protein engineering data for buried residues. Protein Sci 11:1984–2000.
286. Shakhnovich EI, Finkelstein AV (1989) Theory of cooperative transitions in protein molecules. I. Why denaturation of globular protein is a first-order phase transition. Biopolymers 28:1667–1680.
287. Grigoryan G, DeGrado WF (2011) Probing designability via a generalized model of helical bundle geometry. J Mol Biol 405:1079–1100.
288. Eilers M, Patel AB, Liu W, Smith SO (2002) Comparison of helix interactions in membrane and soluble alpha-bundle proteins. Biophys J 82:2720–2736.
289. Zhang Y, Kulp DW, Lear JD, DeGrado WF (2009) Experimental and computational evaluation of forces directing the association of transmembrane helices. J Am Chem Soc 131:11341–11343.
290. Zhou FX, Cocco MJ, Russ WP, Brunger AT, Engelman DM (2000) Interhelical hydrogen bonding drives strong interactions in membrane proteins. Nat Struct Biol 7:154–160.
291. Zhou FX, Merianos HJ, Brunger AT, Engelman DM (2001) Polar residues drive association of polyleucine transmembrane helices. Proc Natl Acad Sci USA 98: 2250–2255.
292. Choma C, Gratkowski H, Lear JD, DeGrado WF (2000) Asparagine-mediated self-association of a model transmembrane helix. Nat Struct Biol 7:161–166.
293. Call ME, Pyrdol J, Wiedmann M, Wucherpfennig KW (2002) The organizing principle in the formation of the T cell receptor-CD3 complex. Cell 111:967–979.
294. Marsh D (2008) Energetics of hydrophobic matching in lipid–protein interactions. Biophys J 94:3996–4013.
295. Holt A, Killian JA (2010) Orientation and dynamics of transmembrane peptides: the power of simple models. Eur Biophys J 39:609–621.
296. Ursell T, Huang KC, Peterson E, Phillips R (2007) Cooperative gating and spatial organization of membrane proteins through elastic interactions. PLoS Comput Biol 3:e81.
297. de Planque MR, Killian JA (2003) Protein–lipid interactions studied with designed transmembrane peptides: role of hydrophobic matching and interfacial anchoring. Mol Membr Biol 20:271–284.
298. Benjamini A, Smit B (2012) Robust driving forces for transmembrane helix packing. Biophys J 103:1227–1235.
299. Sparr E, Ash WL, Nazarov PV, Rijkers DT, Hemminga MA, Tieleman DP, Killian JA (2005) Self-association of transmembrane alpha-helices in model membranes: importance of helix orientation and role of hydrophobic mismatch. J Biol Chem 280:39324–39331.
300. Thomas D, Bron P, Weimann T, Dautant A, Giraud MF, Paumard P, Salin B, Cavalier A, Velours J, Brethes D (2008) Supramolecular organization of the yeast F1Fo-ATP synthase. Biol Cell 100:591–601.
301. Lee AG (2011) Lipid–protein interactions. Biochem Soc Trans 39:761–766.
302. Findlay HE, Booth PJ (2006) The biological significance of lipid–protein interactions. J Phys Condens Matter 18:S1281–S1291.
303. Adamian L, Naveed H, Liang J (2011) Lipid-binding surfaces of membrane proteins: evidence from evolutionary and structural analysis. Biochim Biophys Acta 1808:1092–1102.
304. Hite RK, Gonen T, Harrison SC, Walz T (2008) Interactions of lipids with aquaporin-0 and other membrane proteins. Pflugers Arch 456:651–661.
305. Hite RK, Li Z, Walz T (2010) Principles of membrane protein interactions with annular lipids deduced from aquaporin-0 2D crystals. EMBO J 29:1652–1658.
306. Qin L, Hiser C, Mulichak A, Garavito RM, Ferguson-Miller S (2006) Identification of conserved lipid/detergent-binding sites in a high-resolution structure of the membrane protein cytochrome c oxidase. Proc Natl Acad Sci USA 103:16117–16122.
307. Horn R, Paulsen H (2002) Folding in vitro of light-harvesting chlorophyll a/b protein is coupled with pigment binding. J Mol Biol 318:547–556.
308. Kern J, Guskov A (2011) Lipids in photosystem II: multifunctional cofactors. J Photochem Photobiol B Biol 104:19–34.
309. Jones MR (2007) Lipids in photosynthetic reaction centres: structural roles and functional holes. Prog Lipid Res 46:56–87.
310. Krebs MP, Isenbarger TA (2000) Structural determinants of purple membrane assembly. Biochim Biophys Acta 1460:15–26.
311. Pebay-Peyroula E, Rosenbusch JP (2001) High-resolution structures and dynamics of membrane protein–lipid complexes: a critique. Curr Opin Struct Biol 11: 427–432.
312. Shinzawa-Itoh K, Aoyama H, Muramoto K, Terada H, Kurauchi T, Tadehara Y, Yamasaki A, Sugimura T, Kurono S, Tsujimoto K, Mizushima T, Yamashita E, Tsukihara T, Yoshikawa S (2007) Structures and physiological roles of 13 integral lipids of bovine heart cytochrome c oxidase. EMBO J 26:1713–1725.
313. Arnarez C, Marrink SJ, Periole X (2013) Identification of cardiolipin binding sites on cytochrome c

oxidase at the entrance of proton channels. Sci Rep 3: 1263.
314. de Kroon AI, Dolis D, Mayer A, Lill R, de Kruijff B (1997) Phospholipid composition of highly purified mitochondrial outer membranes of rat liver and neurospora crassa. Is cardiolipin present in the mitochondrial outer membrane? Biochim Biophys Acta 1325: 108–116.
315. Cosentino K, Garcia-Saez AJ (2014) Mitochondrial alterations in apoptosis. Chem Phys Lipids 181:62–75.
316. Zhang M, Mileykovskaya E, Dowhan W (2002) Gluing the respiratory chain together. Cardiolipin is required for supercomplex formation in the inner mitochondrial membrane. J Biol Chem 277:43553–43556.
317. Schagger H (2002) Respiratory chain supercomplexes of mitochondria and bacteria. Biochim Biophys Acta 1555:154–159.
318. Schlame M, Rua D, Greenberg ML (2000) The biosynthesis and functional role of cardiolipin. Prog Lipid Res 39:257–288.
319. Pebay-Peyroula E, Dahout-Gonzalez C, Kahn R, Trezeguet V, Lauquin GJ, Brandolin G (2003) Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. Nature 426:39–44.
320. Hoffmann B, Stockl A, Schlame M, Beyer K, Klingenberg M (1994) The reconstituted ADP/ATP carrier activity has an absolute requirement for cardiolipin as shown in cysteine mutants. J Biol Chem 269:1940–1944.
321. Ferguson AD, Welte W, Hofmann E, Lindner B, Holst O, Coulton JW, Diederichs K (2000) A conserved structural motif for lipopolysaccharide recognition by procaryotic and eucaryotic proteins. Structure 8:585–592.
322. 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.
323. Noinaj N, Kuszak AJ, Gumbart JC, Lukacik P, Chang H, Easley NC, Lithgow T, Buchanan SK (2013) Structural insight into the biogenesis of beta-barrel membrane proteins. Nature 501:385–390.
324. Kucerka N, Nagle JF, Sachs JN, Feller SE, Pencer J, Jackson A, Katsaras J (2008) Lipid bilayer structure determined by the simultaneous analysis of neutron and X-ray scattering data. Biophys J 95:2356–2367.
325. Popot JL, Engelman DM (1990) Membrane-protein folding and oligomerization—the 2-stage model. Biochemistry 29:4031–4037.