Untangling the complexity of membrane protein folding

Heedeok Hong, *Michigan State University*
Hyun-Kyu Choi, *Emory University*
Tae-Young Yoon, *Seoul National University*

**Journal Title:** Current Opinion in Structural Biology

**Volume:** Volume 72

**Publisher:** Elsevier | 2022-02-01, Pages 237-247

**Type of Work:** Article | Post-print: After Peer Review

**Publisher DOI:** 10.1016/j.sbi.2021.11.013

**Permanent URL:** [https://pid.emory.edu/ark:/25593/w3c4s](https://pid.emory.edu/ark:/25593/w3c4s)


**Copyright information:**

This is an Open Access work distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License ([https://creativecommons.org/licenses/by-nc-nd/4.0/](https://creativecommons.org/licenses/by-nc-nd/4.0/)).

Accessed June 8, 2024 4:13 PM EDT
Untangling the complexity of membrane protein folding

Heedeok Hong1,*, Hyun-Kyu Choi2,*, Tae-Young Yoon3

1Department of Chemistry, Michigan State University, East Lansing, MI 48824, USA
2Wallace H. Coulter Department of Biomedical Engineering and Parker H. Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology and Emory University, Atlanta, GA 30332, USA
3School of Biological Sciences and Institute for Molecular Biology and Genetics, Seoul National University, Seoul 08826, South Korea

Abstract

Delineating the folding steps of helical-bundle membrane proteins has been a challenging task. Many questions remain unanswered, including the conformation and stability of the states populated during folding, the shape of the energy barriers between the states, and the role of lipids as a solvent in mediating the folding. Recently, theoretical frames have matured to a point that permits detailed dissection of the folding steps, and advances in experimental techniques at both single-molecule and ensemble levels enable selective modulation of specific steps for quantitative determination of the folding energy landscapes. We also discuss how lipid molecules would play an active role in shaping the folding energy landscape of membrane proteins, and how folding of multi-domain membrane proteins can be understood based on our current knowledge. We conclude this review by offering an outlook for emerging questions in the study of membrane protein folding.

Keywords

Membrane protein folding; Folding energy landscape; Single-molecule force spectroscopy; Magnetic tweezers; Steric trapping; Cooperativity; Lipophobic effect

The energetics of membrane protein folding

Membrane proteins carry out essential cellular functions including the synthesis of ATP, material exchange, and eliciting cell signaling, to just name a few. To perform a wide variety of functions, membrane proteins should correctly fold into native three-dimensional structures in the lipid bilayers. Cells invest significant resources for homeostasis of membrane proteins, and failures in their correct folding are implicated in various diseases [1–6]. The folding energy landscape of proteins is generally determined by the amino acid sequence of polypeptide chains and their interaction with the solvent [7]. In the case of globular proteins, the hydrophobic effect provides a critical driving force for folding by
inducing the collapse of nonpolar residues and exclusion of water from the protein interior [8]. Due to the distinct environmental constraints provided by the lipid bilayers, the physical principles of folding are expected to be substantially different for the membrane proteins.

A “two-stage” model has been a prominent model for describing the folding of helical-bundle membrane proteins, in which each folding stage is driven by a distinct set of molecular forces [9,10]. In the first stage, polypeptide segments insert into the lipid bilayer as stable transmembrane helices (TMHs), mainly driven by the hydrophobic nature of the amino acids belonging to TMHs and the high desolvation cost of unpaired backbone hydrogen bonding (H-bonding) in the nonpolar bilayer (Figure 1a) [11,12]. In the second stage, the TMHs associate into a compact native structure through side-to-side interactions. The van der Waals (vdW) packing and polar interactions drive this stage largely replacing the hydrophobic effect due to a paucity of water molecules in the bilayer core [13–19]. Interestingly, a recent NMR relaxation study indicates that membrane proteins display more pronounced fast (sub-ns) side-chain motions than globular proteins in the interior. Although the contribution of slower motions to an overall side-chain entropy remains unexplored, this result raises a possibility that the side-chain entropic cost can be relatively small upon folding of membrane proteins in the bilayer [20].

Although the two-stage model still serves as a useful, conceptual framework, there is an accumulating body of evidence that the folding of helical membrane proteins is more complex than expected from the canonical two-stage model [21–26]. Wimley and White further dissected the helical membrane protein folding into four steps (Figure 1b), each consisting of an experimentally accessible sequential step with a corresponding free energy contribution [27]:

\[
\Delta G_0 = \Delta G^c_{\text{ic}} + \Delta G^h_{\text{ih}} + \Delta G^m_{\text{im}} + \Delta G^{mb}_{\text{hm}}
\]

Here, \(\Delta G_0\) denotes the total free energy difference between an unfolded coil state in water and the final folded state in the membrane. \(\Delta G^c_{\text{ic}}\) and \(\Delta G^h_{\text{ih}}\) reflect the free energy differences caused upon transfer of an unfolded coil from the water phase to a water–membrane interface and the coil–helix transition at the interface, respectively. \(\Delta G^m_{\text{im}}\) and \(\Delta G^{mb}_{\text{hm}}\) denote the respective changes accompanying insertion of the interfacial helix to the membrane, and the helix–helix association to make helical bundles in the membrane. These free-energy changes have been extensively characterized using model peptides [27–29].

A multitude of studies have determined the range of free energy changes in each step. By tracking partitioning of unstructured peptides and formation of the backbone hydrogen bonds, \(\Delta G^c_{\text{ic}}\) and \(\Delta G^h_{\text{ih}}\) have been estimated as approximately −0.4 \(k_B T\)/residue [30,31] and −0.7 to −0.2 \(k_B T\)/residue [32], respectively. The late-stage values of \(\Delta G^m_{\text{im}}\) and \(\Delta G^{mb}_{\text{hm}}\) are alleged to span −0.8 to −0.6 \(k_B T\)/residue [33–35] and −0.4 to −0.1 \(k_B T\)/residue from TM helix dimerization of glycorphorin A [13,36,37], respectively. Notably, in calculating \(\Delta G^h_{\text{im}}\), the membrane insertion of a polar H-bonded peptide backbone...
is unfavorable with a penalty of $-2.0 \, k_B T/residue$, causing a struggle that has to be compensated by the favorable membrane partitioning of hydrophobic side chains [38]. The membrane insertion of helices also depends on the bilayer properties including thickness and area-compressibility modulus [35]. Collectively, a back-of-envelope calculation yields $-2.3$ to $-1.3 \, k_B T/residue$ for $\Delta G_0$, indicating that membrane protein folding is a spontaneous, thermodynamically down-hill process.

**Emerging tools for dissecting membrane protein folding**

Delineating the folding energy landscape of helical membrane proteins is a daunting task because a reversible control of their unfolding and refolding is notoriously difficult while maintaining the intact bilayer environments. For example, sodium dodecylsulfate (SDS) – a strong anionic detergent used as a denaturant of membrane proteins – inevitably disrupts the bilayer structure at high concentrations, yielding denatured states with non-native topology and interactions [39–43]. Recently, novel methods have been devised for selective control of membrane protein unfolding/refolding, enabling dissection of the folding steps and underlying energetics in unprecedented details.

For water-soluble proteins, methods such as hydrogen/deuterium (H/D) exchange and NMR relaxation dispersion have revealed the dynamic and multistate nature of the native state, which is essential for folding, function and quality control [44–48]. Steric trapping delves into membrane protein folding by coupling spontaneous denaturation of a double biotin-tagged protein to competitive binding of bulky monovalent streptavidin (mSA, $\sim 52$ kDa [49]), retaining native lipid–protein and water–protein interactions (Figure 2a) [36, 37, 50–52]. When biotin tags are attached to two specific residues that are spatially close in the folded state but distant in the amino acid sequence, a first mSA molecule binds either of the biotin tags with an intrinsic binding affinity of $\Delta G_{\text{Bind}}$. Due to steric hindrance, a second mSA binds only when the membrane protein is transiently denatured. Steric trapping thus selectively captures disruption of the tertiary interactions in the region between two biotin tags minimally affecting their secondary structures. By adjusting the biotin affinity of mSA by mutation, the denaturation and binding can be reversibly controlled analogous to the EX2 condition in H/D exchange (i.e., the exchange is limited by the equilibrium constant between the native and denatured states). Protein stability ($\Delta G_{\text{N-D}}$) is obtained from the attenuated second binding of mSA, which is coupled to denaturation ($\Delta G_{\text{N-D}} + \Delta G_{\text{Bind}}$). Analogous to the EX1 condition in H/D exchange (i.e., the exchange is limited by the denaturation rate of a protein), wild-type mSA with a high biotin binding affinity and slow dissociation ($K_d = 10^{-14} \, \text{M}$ and $\tau_{\text{off}} = \text{days}$) can irreversibly trap the denatured state, allowing the determination of a spontaneous denaturation rate [53–55].

To render this method broadly applicable, novel thiol-reactive biotin probes have been developed combining all essential features for steric trapping [52]. These probes contain a fluorophore to detect binding of quencher-labeled mSA or a spin label to measure interspin distances in the native and denatured states using double electron-electron resonance spectroscopy [52]. Steric trapping can be transformed into various formats for investigating several key elements in membrane protein folding [52]: 1) The method captures transient disruption of the tertiary contacts at a specific biotin pair. Thus, by moving the location...
of the biotin pair, local thermodynamic and kinetic stability can be measured. 2) By introducing a point mutation that perturbs local residue interaction and measuring the mutation-induced stability changes at different biotin pairs, it can be evaluated whether the perturbed residue interactions are localized or cooperatively propagated through the surrounding (called “cooperativity profiling”), and 3) Detailed conformational features (e.g., the flexibility and compactness) of the trapped denatured state can be studied under native conditions.

Single-molecule force spectroscopy is a unique tool for studying membrane protein folding because its signals can be directly translated into structural information. This salient feature was well demonstrated in the pioneering works by Müller, Gaub and co-workers, in which atomic force microscopy (AFM) was used to pull single membrane proteins to induce their unfolding [56–58]. Tracking trajectories of the AFM tip as a function of applied force directly reveals the numbers of amino acid residues in unfolded parts of single membrane proteins, allowing to trace unfolding and refolding events in a time-resolved manner. Recently, Perkins and co-workers realized 1 μs resolution in their AFM measurement, which led to unveiling of unfolding intermediates of bacteriorhodopsin at a resolution of 3 amino acid residues [59,60]. However, the loading rates – by which applied force increases and decreases with time – typically reach a few nN/s in AFM studies, driving unfolding and refolding events to occur in a deep non-equilibrium regime. In addition, the unfolded parts are inevitably pulled out of lipid bilayers while losing their secondary structures, making it difficult to resolve respective contributions to the total free energy differences.

Single-molecule magnetic tweezer technique has recently been used in an effort to observe the late-stage folding events (Figure 2b) [61]. Because magnetic tweezing and biochemical processes in question can be completely decoupled, unfolding and refolding events of single membrane proteins can be observed under essentially constant force in a magnetic tweezer setup. Helical structures were observed to be restored in a force range from 20 to 10 pN for all membrane proteins examined, confirming the validity of the two- and four-stage models, in which the coil-to-helix transition (e.g., $\Delta G_{\text{w}}^c \rightarrow \text{i}$ and $\Delta G_{\text{c}}^h \rightarrow \text{h}$) precedes membrane insertion of TMHs and tertiary structure formation. When the magnetic force was further relaxed down to ~5 pN, the single membrane proteins showed discrete transitions between intermediates before reaching the native folded states, revealing crucial on-pathway intermediates of membrane protein folding (Figure 2b). Of note, these transitions are found to involve both membrane insertion and tertiary structure formation that are intricately mixed with one another (e.g., $\Delta G_{\text{i}}^h \rightarrow \text{m}$ and $\Delta G_{\text{m}}^h \rightarrow \text{hb}$).

One emerging question is which energetics out of membrane insertion or tertiary contact formation drives these late-stage transitions. The magnetic tweezers data yielded mixed implications to this question. On one hand, single-point mutations – out of more than two hundreds of amino acid residues and presumably affecting membrane insertion to a marginal extent – still markedly changed the kinetics of state transitions, an indicative that tertiary contact formation is a determining factor in membrane protein folding. On the other hand, adopting vesicle membranes, instead of more permeable bicelles, seriously impedes the observed structural transitions to native folded states. Also, for a transporter protein with a
weaker folding propensity, dedicated insertase complexes need to be added to facilitate their folding (H.-K.C. and T.-Y.Y., unpublished data). These latter observations strongly point to a hypothesis that membrane insertion also presents serious energy barriers, at least, for a certain class of membrane proteins exhibiting higher insertion energy costs for some of their TMHs.

Native nano-electrospray ionization ion-mobility/mass spectrometry is also addressing several key problems in membrane protein folding, including identification of the lipids that tightly bind to proteins [62], the affinity and cooperativity of lipid-protein interactions [63,64], and the effect of bound lipids on protein stability. Single-molecule photobleaching is especially powerful for measuring strengths of protein-protein interactions in a lipid bilayer [65]. This method has been applied for determining the dimer affinity of the H⁺/Cl⁻ antiporter of *E. coli* (CLC-ec1) in the bilayer [65], revealing one of the strongest membrane protein interactions ever measured and the lipid effect on the dimerization equilibrium [66].

**Dissecting the folding pathways of helical membrane proteins: Lessons from GlpG**

One of the major questions regarding the folding of helical-bundle membrane proteins is whether these proteins fold cooperatively or via partially folded intermediates. Cooperativity links the behavior of different residues within a protein [67], manifested as all-or-none (i.e., two-state) folding and a single predominant free energy barrier separating the native and the denatured states [68]. While theoretical tests with membrane proteins have reached diverse conclusions, with some supporting cooperative folding to attain the native state as the global free energy minima, and others predicting multiple or more rugged folding pathways with local energy minima [69,70]. Recently, the TM domain of GlpG of *E. coli*, a six-transmembrane helical bundle membrane protein that belongs to the rhomboid protease family, has emerged as a tractable model due to its folding reversibility [71], monomeric character [72], and biological importance of the protein family [73]. Multiple experimental [43,52,53,61,71,74,75] and computational [69,70,76,77] studies with GlpG have provided a wholistic picture, pointing to a complex folding energy landscape even for the single-domain membrane protein.

With the ability to control folding and denaturation in a native environment, steric trapping has resolved two subdomains with distinct folding properties in micelles (Figure 2c) [52,53]. A N-terminal subdomain consisting of TMHs 1–3 has the higher thermodynamic (ΔG_N,D = 9.8 k_BT) and kinetic stability (E_a,D = 62.0 k_BT; E_a,D denotes the activation energy of denaturation), whose disruption leads to global denaturation. The remaining subdomain (TMHs 4–6), harboring the catalytic dyad, is less stable (ΔG_N,D = 7.9 k_BT and E_a,D = 48.5 k_BT) and undergoes subglobal denaturation. Cooperativity profiling based on steric trapping reveals multiple clusters of cooperative interactions in GlpG (Figure 2d) [52]. Notably, the cooperative clusters overlap with the stability hotspots determined by thermal denaturation [71], as well as the folding nucleus (a membrane-buried packing core consisting of N-terminal regions) suggested from the ϕ-value analysis [43]. This distributive nature of cooperative interactions in GlpG is markedly distinguished from folding of soluble, globular proteins.
proteins that often involves cooperative interactions in the hydrophobic core working as a single folding nucleus [44,67,78,79].

Single-molecule force spectroscopy using magnetic tweezers have also resolved detailed conformational and energetic features of the intermediates populated during GlpG folding in the lipid bilayer environment [61,74]. By relaxing force to low-pN levels, Choi, Yoon and co-workers were able to observe two intermediates repeatedly populated during folding of single GlpG proteins [61]. While the first intermediate \( I_1 \) consists of TMHs 1 and 2 connected by a long linker region, the second intermediate \( I_2 \) involves TMHs 3 and 4 in the growing structure, indicating helical hairpins as a basic folding unit. Because the magnetic tweezer experiments were conducted under essentially constant force levels, the observed energy landscape of state transition was directly converted to parameters describing the energy landscape governing the late-stage folding steps [61] (Figure 2e). Different models yielded consistent results, indicating the robustness of the energy landscape reconstruction.

Remarkably, despite subtle differences in intermediate conformations, the steric trapping and the single-molecule force spectroscopy methods agree on many observations: 1) GlpG has moderate thermodynamic stability with a large unfolding barrier. 2) GlpG has a modular architecture consistent with predictions from the AWSEM (Associative Memory, Water Mediated, Structure and Energy Model)-membrane molecular dynamics simulation [70], in which the more stable N-terminal region serves as a keystone for the global conformational integrity whereas the folding of the C-terminal region is rather dispensable [52,61,70,74,76]. Taken together, these findings suggest that the single-domain helical-bundle membrane protein GlpG has a rugged folding energy landscape during folding within the bilayer. The high kinetic barrier between the native and unfolded state and the strong foldability of the N-terminal region may be beneficial for avoiding misfolding or aggregation in the crowded cellular environments [61,80].

**Does a “lipophobic effect” promote protein folding in the membrane?**

In the second stage of the two-step model, individual TMHs associate into a compact native structure while overcoming their favorable interactions with lipids (Figure 1a). An emerging question is if there is a lipophobic effect that induces coalescence of unassembled TMHs in the lipid bilayers [81,82], analogous to the way the hydrophobic effect drives collapse of hydrophobic residues repelled by water molecules. On the bright side, such lipophobic effect could serve as a folding driving force that can compensate the lack of the hydrophobic effect within the membrane. On the dark side, the same force would facilitate nonspecific collapse of TMHs in the membrane, increasing a chance for misfolding. Interestingly, a number of evidence imply that membrane proteins are vulnerable to misfolding and nonspecific protein-protein interactions in the cell membranes: 1) The cellular folding efficiency of membrane proteins is typically low, ranging from only 20 to 50 % [83,84]; 2) Proper membrane protein biogenesis requires an assistance of quality control machineries (e.g., chaperones and proteases [2,83,85–90]; 3) In the liposomes, the affinity of native TM helix-helix interactions dramatically decreases when the total membrane protein extracts are co-reconstituted [37].
At the molecular level, the lipophobic effect may be manifested as a way to relieve the energetic penalty caused by unfavorable protein-lipid interactions in the denatured state, such as a hydrophobic mismatch between a protein and the bilayer leading to lipid deformation at the protein surface \cite{66,91} and an unfavorable exposure of bulky polar/charged residues to the bilayer core that results in ineffective lipid solvation and burial of the exposed residues into the protein interior \cite{92,93} (Figure 3a). Another possible scenario for the lipophobic effect is that inserted TMHs perturb favorable lipid-lipid interactions \cite{94}, increasing the overall chemical potential of the denatured state relative to the native state (Figure 3b). If so, the tendency to maximal restoration of the perturbed lipid-lipid interactions may provide a driving force for protein compaction within the bilayer regardless of the hydrophobic mismatch and amino acid sequence of the TMHs. Notably, an all-atom MD simulation study predicts that glycorphorin A monomers collapse with non-native contacts in the bilayer as an obligatory step prior to the formation of the native dimer \cite{95}.

**Folding of multi-domain membrane proteins**

Many membrane proteins are characterized by higher-order symmetry properties in their structures, with the pseudo-C2 and -C4 symmetry of transporter and channel proteins being outstanding examples \cite{96}. The symmetry planes divide the tertiary and quaternary structures of membrane proteins (or complexes) into subdomains and subunits, each defining basic units in function and folding. For example, secondary transporters belonging to the major facilitator superfamily (MFS) have two highly symmetric domains with conserved structures (referred to as the MFS fold) \cite{70,97} (Figure 4a). These two domains serve as principal units in the transporters’ rocker switch motions, creating conformational states with alternating access to extra- and intra-cellular spaces \cite{97}.

One interesting observation is many membrane proteins, including the MFS transporters, have evolved to form a connected structure via duplication or fusion of the genes of primordial transporters \cite{98,99}. Such connected structures, however, seemingly pose a higher risk in terms of foldability because a failure in folding of either one domain would jeopardize folding of the entire protein, raising a question about the advantages of the connected structure despite the increased risk for misfolding. Indeed, it is recently suggested that gene duplication likely increases the chance for misfolding by forming a swapped structure between duplicated domains \cite{100}. One possible hypothesis is specialization of each domain in its role, with one domain evolving to contend with functional requirements (but becoming less foldable) and the other domains providing the structural stability of the complex. It is observed that the C-domain of *E. coli* ClC-ec1 is unstable by itself and requires its N-terminal counterpart for maintaining its tertiary structure and likely, proper folding in the endoplasmic reticulum (ER) membrane \cite{101}. The differentiated folding propensity of the N- and C-terminal domains has been observed for GlpG using the steric trapping method (Figure 2). This strict hierarchy in the folding order of respective domains may account for how membrane proteins structure its folding and prevent a domain-swapped structure. It remains to be examined whether this specialization of different domains in foldability and functionality is observed for a broader class of multi-domain membrane proteins.
Most multi-domain membrane protein families are epitomized by the structural folds that have been conserved over a long evolutionary distance from *E. coli* to human. This would suggest that the folding pathway of these structural folds are essentially shared across the family members. One interesting finding is that many of these conserved folds have entangled topologies in their TMH alignment, defying a folding pathway in which TMHs are added onto one another in a simple, interchangeable manner. In the case of MSF folds, TM helix 1 is flanked by TMHs 5 and 6 (Figure 4b). This entanglement suggests that TMHs 5 and 6 should follow TM helix 1 in the folding order, likely but not the other way around. Likewise, for the partial structure composed of TMHs 1 to 4, TMHs 1 and 2 in turn wrap around TMH 4, which would force folding of TMHs 1 and 2 only after that of TMH 4. Thus, by only considering the topology of the structure, one would be able to roughly delineate the folding order of the MFS fold. Presumably, TMHs 4 (and 3) form the inner core of the MFS fold, onto which TMHs 1 and 2 are added. TMHs 5 and 6 follow to wrap around the growing structure to complete folding of the MFS fold. This folding order remains to be confirmed by experimental observations.

**Conclusion**

Despite the daunting complexity of membrane protein biogenesis, theoretical frames have evolved to permit detailed dissection of membrane protein folding into elementary steps. Advances in experimental platforms now allow us to observe and modulate different steps of membrane protein folding in a highly selective manner. These collective efforts suggest what appears to be overarching principles for biogenesis of diverse classes of membrane proteins. One prominent example is an overall hierarchical N- to C-terminal folding pathway. During the translation of membrane proteins, the more stable N-terminal segments are favorably inserted into the membrane and partially folded, protecting the nascent chain from misfolding and aggregation. Upon emergence of more C-terminal TMHs from the ribosome vestibule, the partially folded N-terminal region works as a structural template for subsequent folding guided by the funneled energy landscape. Once folded, the protein is kinetically stabilized to carry out function. This scenario may be adopted by other types of membrane proteins since a similar N-to-C directional folding has been observed for GPCRs such as β₂-adrenergic receptor [61] and rhodopsin [80], which are evolutionary distant from GlpG. Finally, we note that cryo-electron microscopy techniques are unveiling the tertiary and quaternary structures of membrane proteins as well as interactions with lipids with unprecedented speed [102], which will serve as bases for elucidation of their folding pathways and underlying physical principles. The new computational platforms that permit direct structural predictions from sequence information [103,104] will also expedite the discovery of new membrane protein folds and the design and analysis of membrane protein folding experiments. Thus, we expect the coming decades will be an exciting era for the field of membrane protein folding to make previously inconceivable discoveries and chart new territories.

**Acknowledgement**

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (2021R1A3B1071354 to T.-Y.Y.), Postdoctoral Fellowship Program granted by National
References


66. Chadda R, Bernhardt N, Kelley EG, Teixeira SCM, Griffith K, Gil-Ley A, Ozturk TN, Hughes LE, Forsythe A, Krishnamani V, et al.: Membrane transporter dimerization driven by differential lipid solvation energetics of dissociated and associated states. Elife 2021, 10. •• This work demonstrates that membrane-protein-protein interaction can be strongly modulated by the lipid deformation energy caused by the hydrophobic mismatch in both the dissociated and assembled states.


This work suggests that membrane protein folding can be described using the minimally-frustrated, funneled folding energy landscape.


Figure 1. Models for the folding of helical-bundle membrane proteins.

(a) Two-stage model. In Stage I, a nascent unstructured polypeptide chain is inserted into a lipid bilayer forming transmembrane helical secondary structures. In the inset, co-translational insertion is described. In Stage II, inserted helices laterally associate to form a tertiary structure. Recently, it has been suggested that helices can transiently partition to the membrane surface, or their membrane insertion is coupled to the association (denoted as “entangled pathway”). Adapted and modified from Ref [9].

(b) Four-step model. The overall folding is dissected into (1) adsorption of an unstructured coil onto the membrane surface, (2) coil-to-helix transition, (3) insertion of helices into membrane, and (4) helix-helix association towards its native state. Adapted and modified from Ref [27]. GlpG (2XOV) as the model protein was used to visualize.
Figure 2 | Emerging techniques for resolving membrane protein folding.

(a) Illustration of steric trapping method for studying membrane protein folding under native conditions. The principle of the method is described in the main text. \( \Delta G^0 \) Binding: the free energy change upon binding of mSA to biotin; \( \Delta G_{N-D}^0 \): the free energy change upon denaturation (i.e., thermodynamic stability); \( k_{on} \): on-rate constant of mSA binding to biotin; \( k_{off} \): off-rate constant of mSA-biotin complex; \( k_f \): spontaneous denaturation rate; \( k_f^* \): spontaneous folding rate. Adapted and modified from Refs [52,53].

(b) Schematic diagram of single-molecule magnetic tweezer experiments for examining folding of single GlpG proteins (2XOV). Adapted and modified from Ref [61].

(c) Folding energy landscape of GlpG obtained from steric trapping in dodecylmaltoside micelles. \( \Delta G_{N-D}^0 \) and activation energy of denaturation (\( E_{a,D} \)) of each subdomain were determined using the biotin pair conjugated to each subdomain. \( \ddagger \): transition state. Adapted and modified from Ref [53].

(d) The cooperative interaction map of GlpG obtained from steric trapping-based cooperativity profiling method. The degree of cooperativity or localization of each residue interaction is color-coded. Notably, the residues at the TMH4 TMH6 interface are located at C-subdomain but the disruption of these residue interactions more substantially destabilize N-subdomain (i.e., highly overpropagated interactions). Absolutely conserved residues among the rhomboid protease family are marked with asterisks. \( \Delta \Delta \Delta G = \Delta \Delta G_{WT-Mut}^0 \) (N-subdomain) – \( \Delta G_{N-D}^0 \) (C-subdomain). Here, \( \Delta \Delta \Delta G \), the cooperativity parameter, is
defined as the difference between the stability changes induced by a specific point mutation ($\Delta \Delta G_{WT-Mut}$) measured at the biotin pairs in N- and C-subdomains. Adapted and modified from Ref [52]. (e) Reconstructed folding energy landscapes (black for Bell-Zhrukov model, red for Dudko model) for *E. coli* GlpG at zero force. Errors in free energy are from the errors in prefactors (for Bell-Zhrukov) and rates (for Dudko model), and errors in reaction coordinate are from the errors in distances between equilibrium states and transition states. Adapted and modified from Ref [61].
Figure 3 | Proposed modes of the lipophobic effect.
(a) Illustration of the monomer-dimer equilibrium of *E. coli* H^+*/Cl^−* antiporter (ClC-ec1; 1OTS). In the monomeric state of ClC-ec1, the hydrophobic mismatch near the protein-lipid interfacial regions (green and blue) can occur, causing the unfavorable lipid deformation at the dimer interfaces and the penetration of water molecules (inset). These unfavorable lipid-protein interactions can drive dimerization. Adapted and modified from Ref [66]. (b) Illustration of the general lipophobic effect. Relative to the native state, the lipid-lipid interactions in the lipid bilayer are perturbed by the lipid solvation of individual TM segments (*i.e.*, unfavorable lipid-lipid interactions: the chemical potential contribution from this interaction is positive, $\Delta \mu_{\text{Lipid-Lipid}} > 0$, red upward arrow). Here, the favorable intramolecular protein interactions are not yet fully realized (*i.e.*, unfavorable protein-protein interactions; $\Delta \mu_{\text{Protein-Protein}} > 0$, red upward arrow), and the TM segments are highly solvated by lipids (*i.e.*, favorable lipid-protein interaction: $\Delta \mu_{\text{Lipid-Protein}} < 0$, blue downward arrow). Summed up, the chemical potential of the system in the denatured state increases ($\Delta \mu > 0$) relative to the native state. These energetic factors in the denatured state can serve as driving forces for folding into the compact native state.
Figure 4. Folding of multi-domain membrane proteins.

(a) Structure of human glucose transporter 3 with pseudo-C2 symmetry (GLUT3; 4ZW9). Left and right insets show the projected structures from the extracellular view and intracellular view, respectively. (b) A representative schematic for a possible folding pathway of the MFS fold. The N-terminal domain of human GLUT3 is used as a model. MFS; major facilitator superfamily.