



# Energetics of side-chain partitioning of $\beta$ -signal residues in unassisted folding of a transmembrane $\beta$ -barrel protein

Received for publication, April 3, 2017, and in revised form, June 2, 2017. Published, Papers in Press, June 7, 2017, DOI 10.1074/jbc.M117.789446

Bharat Ramasubramanian Iyer<sup>1,2</sup>, Punit Zadafiya<sup>1</sup>, Pallavi Vijay Vetal, and  Radhakrishnan Mahalakshmi<sup>3</sup>

From the Molecular Biophysics Laboratory, Department of Biological Sciences, Indian Institute of Science Education and Research, Bhubaneswar, Bhopal 462066, India

Edited by Wolfgang Peti

The free energy of water-to-interface amino acid partitioning is a major contributing factor in membrane protein folding and stability. The interface residues at the C terminus of transmembrane  $\beta$ -barrels form the  $\beta$ -signal motif required for assisted  $\beta$ -barrel assembly *in vivo* but are believed to be less important for  $\beta$ -barrel assembly *in vitro*. Here, we experimentally measured the thermodynamic contribution of all 20 amino acids at the  $\beta$ -signal motif to the unassisted folding of the model  $\beta$ -barrel protein PagP. We obtained the partitioning free energy for all 20 amino acids at the lipid-facing interface ( $\Delta\Delta G^0_{w,i(\phi)}$ ) and the protein-facing interface ( $\Delta\Delta G^0_{w,i(\pi)}$ ) residues and found that hydrophobic amino acids are most favorably transferred to the lipid-facing interface, whereas charged and polar groups display the highest partitioning energy. Furthermore, the change in non-polar surface area correlated directly with the partitioning free energy for the lipid-facing residue and inversely with the protein-facing residue. We also demonstrate that the interface residues of the  $\beta$ -signal motif are vital for *in vitro* barrel assembly, because they exhibit a side chain-specific energetic contribution determined by the change in nonpolar accessible surface. We further establish that folding cooperativity and hydrophobic collapse are balanced at the membrane interface for optimal stability of the PagP  $\beta$ -barrel scaffold. We conclude that the PagP C-terminal  $\beta$ -signal motif influences the folding cooperativity and stability of the folded  $\beta$ -barrel and that the thermodynamic contributions of the lipid- and protein-facing residues in the transmembrane protein  $\beta$ -signal motif depend on the nature of the amino acid side chain.

Membrane proteins constitute one-third of the proteome of any living organism. They are vital players in important biological processes, including transport, signaling, and cell-cell communication. There is a recent interest in understanding the

biogenesis, folding, and turnover of membrane proteins, particularly the transmembrane  $\beta$ -barrels found in the outer membrane of mitochondria, chloroplast, and Gram-negative bacteria (1–4). Collectively known as outer membrane proteins (OMPs),<sup>4</sup> transmembrane  $\beta$ -barrels share a very simple structural framework consisting of 8–26  $\beta$ -strands embedded into the membrane, arranged largely in an antiparallel fashion (5). The major contributing factor to the thermodynamic stability of OMPs is the free energy of water-to-bilayer partitioning of the lipid-facing amino acid (6). This transfer free energy from water to the bilayer midplane has been measured experimentally using an OMP for the 20 most abundant amino acids (7). This *whole-protein scale* has also been validated computationally (8). The other experimental partitioning scales derived using protein systems are available from transmembrane helices. These include the translocon scale (also considered the *biological hydrophobicity scale*), which measures the translocon-to-bilayer free energy change (9, 10), and the dsTBL system deriving insertion energetics for a single-pass transmembrane helix (11). Most of these scales report the partitioning free energy for residues to the bilayer midplane.

A reliable measure of the thermodynamics at the membrane interface is available from both experimental (11, 12) and computational methods (13). Most of our current understanding of the amino acid partitioning free energy at the interface is derived from the seminal work of Wimley and White (12). The Wimley–White scale, derived from the water-to-bilayer interface partitioning free energy of model pentapeptides, has been widely employed to describe the bilayer interface energetics. The bilayer interface is a complex environment that bridges the hydrophilic solvated exterior with the hydrophobic lipid core (14, 15). It has been argued that the unique amphiphilic chemical milieu presented at the interface demands equally complex residues in the protein structure to be positioned here (14, 16–18). Indeed, aromatic and positively charged residues are abundant at the bilayer interface, where they define the protein boundary during cotranslational folding (19), serve as post-folding protein anchors (14), and facilitate protein–protein

This work was supported by Wellcome Trust/DBT India Alliance Grant IA/I/14/1/501305 (to R. M.). The authors declare that they have no conflicts of interest with the contents of this article.

✂ Author's Choice—Final version free via Creative Commons CC-BY license.

This article contains supplemental Tables S1 and S2 and Figs. S1–S20.

<sup>1</sup> Both authors contributed equally to this work.

<sup>2</sup> Recipient of a senior research fellowship from the University Grants Commission, Government of India.

<sup>3</sup> A Wellcome Trust/DBT India Alliance Intermediate Fellow. To whom correspondence should be addressed: Molecular Biophysics Laboratory, Dept. of Biological Sciences, Indian Institute of Science Education and Research, Rm. 324, III Fl., Block C, Academic Bldg. 3, Bhubaneswar, Bhopal 462066, India. Tel.: 91-755-6691423; Fax: 91-755-6692392; E-mail: maha@iiserb.ac.in.

<sup>4</sup> The abbreviations used are: OMP, outer membrane protein; ASA, accessible surface area; ASP, atomic solvation parameter; BAM,  $\beta$ -barrel assembly machinery;  $C_m$ , chemical denaturation midpoint; DLPC, 1,2-dilauroyl-*sn*-glycero-3-phosphocholine; DPC, *n*-dodecyl phosphocholine; DPR, detergent/protein ratio;  $f_U$ , unfolded fraction;  $\Delta G^0$ , Gibbs free energy of folding; GdnHCl, guanidine hydrochloride; LUV, large unilamellar vesicle;  $m$ , folding cooperativity; pNPP, *p*-nitrophenyl palmitate; SUV, small unilamellar vesicle; Aro, aromatic residue; PDB, Protein Data Bank.

## Side-chain interface energetics in PagP folding

interactions (20, 21). In Gram-negative bacteria, an additional functional importance of the interface residues at the C terminus is in OMP biogenesis.

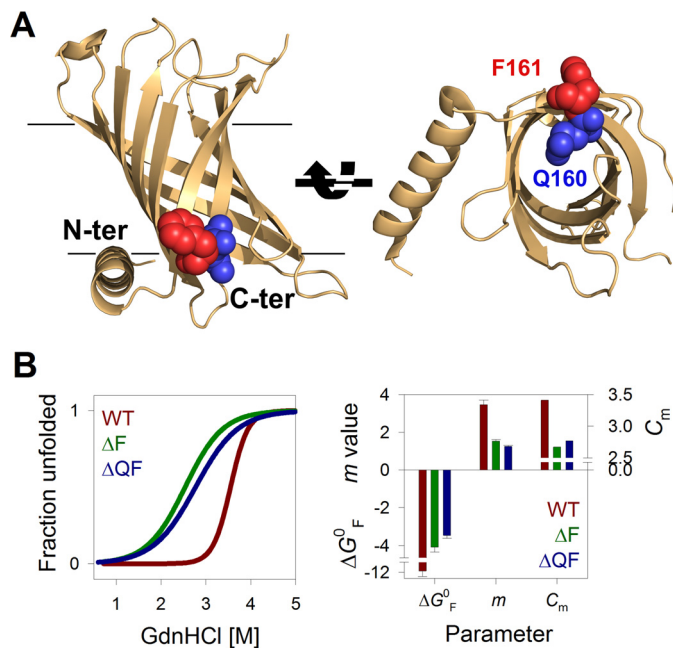
The C-terminal interface residues of transmembrane  $\beta$ -barrels are of particular interest in OMP folding. They bear the  $\beta$ -signal with the consensus Aro-Xaa-Aro motif (where Aro represents an aromatic amino acid and Xaa is any amino acid) (supplemental Fig. S1) and are well conserved across all known bacterial OMPs (22, 23). The  $\beta$ -signal is believed to serve as the recognition sequence for OMP assembly factors such as SurA and Skp in prokaryotes (22) and their eukaryotic analogs (tiny translocases of the inner membrane (24)) during OMP biogenesis. In the model OMP PagP (PhoPQ-activated gene P), the C-terminal residues Gln<sup>160</sup> and Phe<sup>161</sup> form a part of the  $\beta$ -signal. These residues map to the interface in the folded  $\beta$ -barrel topology (25–27). In folded PagP, the side chain of Gln<sup>160</sup> faces the barrel interior, whereas Phe<sup>161</sup> points to the water-lipid interface (Fig. 1A). NMR measurements of PagP in *n*-dodecylphosphocholine (DPC) micelles showed that the terminal Phe<sup>161</sup> represents a position of moderate hydrophobicity (28). The results from this study, which used oxygen and water contacts as probes of hydrophobicity and topology, indicate that Phe<sup>161</sup> exhibits NMR characteristics that suggest an interface location, whereas the inward-facing side chain of Gln<sup>160</sup> displays relaxation rates indicative of a hydrophilic region. Gln<sup>160</sup> and Phe<sup>161</sup> also show higher dynamicity than other strand residues (25). Additionally, Gln<sup>160</sup> exhibits a high  $\phi$ -value and therefore appears to be well structured in the transition state ensemble of PagP (29). Gln<sup>160</sup> is also involved in hydrogen bonding with Trp<sup>60</sup> in the interior of the PagP barrel and is crucial to the formation of the folded structure. Hence, Gln<sup>160</sup> can be considered a part of the PagP folding nucleus (29).

Currently, there is no thermodynamic measurement available for the contribution of these  $\beta$ -signal residues to OMP stability, post-assembly. Here, we measured the thermodynamic contribution of all 20 residues at the  $\beta$ -signal motif, using the C-terminal  $\beta$ -signal residues of PagP. PagP is one of the few thermodynamically stabilized OMPs to exhibit reversible folding (6, 29, 30). We measured the folding free energy for all residues at the protein-facing and lipid-facing membrane interface by systematically substituting the C-terminal interface residues PagP-Q<sup>160</sup> and PagP-F<sup>161</sup>. We found that the residue preference at the amphiphilic interface (position 161) was in favor of side-chain hydrophobicity. Further, we obtained residue dependence for  $\beta$ -barrel assembly from the side-chain preference at the penultimate residue (position 160). Our measured interface energetics for the  $\beta$ -signal residues correlates conditionally with the biological hydrophobicity scale, whereas we obtained strong correlations with the whole-protein and Wimley–White scales, with interesting deviations.

## Results

### Interface residues Gln<sup>160</sup>-Phe<sup>161</sup> stabilize PagP through backbone hydrogen bonds

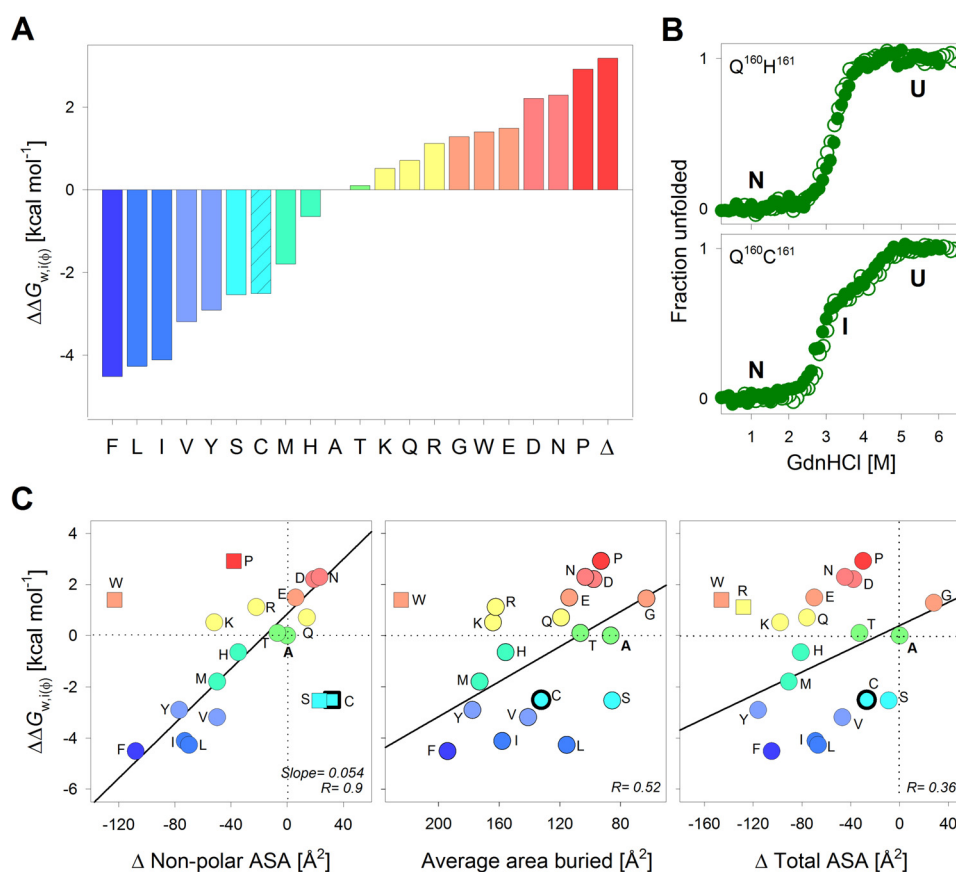
The folding of PagP *in vitro* is driven exclusively by inter- and intramolecular interactions, without assistance from accessory proteins (25, 28, 29), and would therefore not require the  $\beta$ -sig-



**Figure 1. C-terminal residues affect PagP equilibrium folding free energy.** *A*, schematic representations of *E. coli* PagP (PDB entry 3GP6) highlighting the C-terminal (Phe<sup>161</sup>; maroon) and penultimate (Gln<sup>160</sup>; blue) residues belonging to the  $\beta$ -signal of the 8-stranded transmembrane barrel. The structure on the right highlights how the Phe<sup>161</sup> side chain projects toward the lipid environment, whereas that of Gln<sup>160</sup> faces the protein interior. *B*, two-state equilibrium folding data of PagP-WT (WT; brown) and C-terminal deletion mutants PagP- $\Delta$ F<sup>161</sup> ( $\Delta$ F; green) and PagP- $\Delta$ Q<sup>160</sup>F<sup>161</sup> ( $\Delta$ QF; blue) obtained from GdnHCl-mediated chemical denaturation. Fluorescence intensities at the emission maximum were plotted as unfolded fraction from 0 to 1 and fitted to a two-state unfolding equation (fits are shown in the left panel; see supplemental Fig. S3 for the data) to derive the thermodynamic parameters  $\Delta G^0$ ,  $m$  value, and  $C_m$  (right). Error bars, goodness of fit.

nal. To check this, we measured the thermodynamic parameters for wild-type PagP (PagP-WT) and two mutants that lacked the terminal Phe<sup>161</sup> (PagP- $\Delta$ F<sup>161</sup>) and Gln<sup>160</sup>-Phe<sup>161</sup> (PagP- $\Delta$ Q<sup>160</sup>F<sup>161</sup>) using equilibrium chemical denaturation measurements. We used the untagged full-length PagP overproduced as inclusion bodies in *Escherichia coli*. It is widely known that membrane proteins are rarely amenable to reversible equilibrium measurements and often exhibit hysteresis between the folding and unfolding arms (6, 31). In lipidic systems, such as 1,2-dilauroyl-*sn*-glycero-3-phosphocholine (DLPC) vesicles, PagP lacking a His<sub>6</sub> tag (32) also shows hysteresis in various buffer and pH conditions. We could achieve reversible folding of untagged PagP only at pH 9.5 in DPC micelles. Other buffer systems and pH conditions (including acidic pH) showed path dependence and depressed values for folding/unfolding cooperativity. Hence, our experiments are carried out at pH 9.5. Whereas recent studies suggest that properties of membrane proteins measured *in vitro* using micellar systems, such as DPC, can represent their native behavior (33–36), whether our results from such a simplified system at non-native pH can be extrapolated to conditions *in vivo* is unclear.

The deletion of Phe<sup>161</sup> is sufficient to decrease the Gibbs free energy ( $\Delta G_{F,H_2O}^0$ , indicated henceforth as  $\Delta G^0$ ) of PagP-WT by  $\sim 8$  kcal mol<sup>-1</sup> (Fig. 1B). This 3-fold reduction in the stability of PagP upon deletion of residue 161 suggests that phenylalanine plays an anchoring role in PagP at the membrane interface, as seen for interface aromatics in several  $\beta$ -barrels (37–39).



**Figure 2. Interface energetic contributions of the PagP terminal lipid-facing residue depends on non-polar ASA.** *A*, water-to-interface partitioning free energy calculated for all 20 residues at the lipid-facing amphiphilic position 161, normalized with respect to alanine. Amino acids are represented by their single-letter codes;  $\Delta$ , PagP- $\Delta F^{161}$ . Histograms are colored from blue to red in decreasing order of  $\Delta\Delta G_{w,i(\phi)}^0$ . PagP- $C^{161}$  showed a three-state folding profile (shown in *B*), and the free energy change for the more distinct transition ( $\Delta G_{I \rightarrow N}^0$ ; patterned fill in histogram) is used here for comparison. *B*, chemical denaturation profiles exhibit no hysteresis, as seen from the overlay of folding (open circles) and unfolding (filled circles) profiles for two representative mutants. See supplemental Figs. S3 and S4 for the complete data. The  $\Delta G^0$  derived from the folding titrations was used to calculate the partitioning free energies shown in *A*. *C*, correlations between  $\Delta\Delta G_{w,i(\phi)}^0$  and empirical parameters describing the change in ASA for each amino acid are presented as scatter plots. Linear fits to the correlation are shown as solid lines. Values excluded from the fits are shown as square symbols. Variants with three-state profiles are shown as symbols with thicker edges. The best correlation was observed with the change in non-polar ASA (left). Other parameters, such as average area buried upon folding (middle) and change in total ASA (right) do not correlate well with the interface partitioning free energy. The color code for the scatter plots is retained from *A*.

To assess whether PagP requires Phe<sup>161</sup> to fold in micellar systems *in vitro*, we generated a PagP- $X^{161}$  (where  $X$  represents any amino acid) mutant library. We found that all mutants were >85% folded, resisted proteolysis, and were enzymatically active (supplemental Fig. S1). Hence, substitution but not deletion of the C-terminal motif residue preserves the ability of PagP to fold in lipidic environments. Whereas previous studies (22, 23) have associated the Aro-Xaa-Aro motif with its role in interacting with the barrel assembly machinery, we found evidence that the C-terminal residue anchors the OMP to the membrane (Fig. 1) and facilitates barrel assembly through backbone hydrogen bonding (Fig. 1*B*).

#### Thermodynamic contribution of the interface residue correlates with change in non-polar accessible surface area (ASA)

The interface residues of membrane proteins bridge the complex hydrophobic interior of the lipid bilayer with the highly polar aqueous exterior and usually have an amphiphilic nature (40, 41). The PagP- $X^{161}$  library serves as an excellent model to experimentally determine the water-to-interface par-

tioning free energy for all 20 residues at the  $\beta$ -signal motif. To validate whether residue 161 is indeed located at the interface in the PagP-DPC assembly, we carried out an all-atom molecular dynamics simulation and determined the local vicinity for Phe<sup>161</sup>. The 5- $\text{\AA}$  vicinity of Phe<sup>161</sup> shows a moderate occupancy for water, lipid headgroup, and acyl tail, which is characteristic of an interface location for this residue (for details, see supplemental Fig. S2).

We determined the  $\Delta G^0$  for each PagP mutant using equilibrium chemical denaturation measurements (Fig. 2 (*A* and *B*) and supplemental Figs. S3–S5). Except for PagP- $C^{161}$ , all of the 19 PagP- $X^{161}$  constructs showed a two-state transition between the folded (F) and unfolded (U) protein states, which we fitted to a two-state linear extrapolation model (42) to obtain the  $\Delta G^0$  for each mutant. From the  $\Delta G^0$ , we derived the energetic cost of transferring the guest amino acid at Xaa<sup>161</sup> ( $\Delta G_{X}^0$ ) to the interface with reference to alanine ( $\Delta G_A^0$ ), as  $\Delta\Delta G_{w,i(\phi)}^0 = \Delta G_A^0 - \Delta G_X^0$  (where  $\phi$  represents the hydrophobic lipid-facing interface).

We found that the  $\Delta\Delta G_{w,i(\phi)}^0$  follows a hydrophobicity scale with hydrophobic residues at one end and charged residues at the other end, separated by alanine. The two mutants

## Side-chain interface energetics in PagP folding

PagP- $\Delta F^{161}$  and PagP-P<sup>161</sup> and  $\Delta P^{161}$  are highly destabilized because they lack a backbone hydrogen bond(s). We anticipated that residues with amphiphilic or polar nature would complement the amphiphilic nature of the interface and show the highest  $\Delta\Delta G^0_{w,i(\phi)}$ . Surprisingly, we obtained the highest stabilizing contribution at the interface for hydrophobic residues (Fig. 2A). Furthermore, aliphatic residues were marginally more stabilizing than their aromatic counterparts (Leu  $\sim$  Phe; Ile  $>$  Tyr; Fig. 2A). Hence, free energy contributions at the  $\beta$ -signal interface show similar dependences as the hydrophobic effect seen at the bilayer midplane (7).

To identify the factors that contribute to side-chain partitioning in PagP, we correlated the  $\Delta\Delta G^0_{w,i(\phi)}$  with the ASA (43, 44) (Fig. 2C). We found that the  $\Delta\Delta G^0_{w,i(\phi)}$  correlated directly with the difference in non-polar buried ASA for each amino acid ( $r = 0.9$ ; Fig. 2C). The slope of this linear correlation is higher (slope = 0.054) than previous reports (slope = 0.023) (12, 45) and instead resembles non-polar ASA changes for the bilayer midplane (7). This may be because our measured  $\Delta\Delta G^0_{w,i(\phi)}$  at the interface was overestimated. The dynamicity and increased hydration of the DPC micelle surface might promote further occlusion of the non-polar surface of amino acid side chains, giving rise to the overestimated  $\Delta\Delta G^0_{w,i(\phi)}$ . This is a limitation in our  $\Delta G^0$  measurements. Although we found that the micelle structure was largely intact within the denaturant range that we used in our experiment, the amount of DPC bound to PagP may vary. Hence, we limited our comparisons to the thermodynamics across mutants and derived the  $\Delta\Delta G^0_{w,i(\phi)}$  with respect to alanine. In this process, we assumed that the energetic contributions were normalized across the mutants for the minor artifacts arising from the use of micelles as the model system. The correlations to the average per-residue buried area and total ASA are poor (Fig. 2C). Hence, a major stabilizing factor at the interface residue 161 of PagP appears to be the occlusion of the non-polar region of its side chain.

We found no reorganization and restructuring in the DPC micellar system across the denaturant range used in our equilibrium folding measurements, yet we cannot rule out contributions from local structural organization in the protein-micelle system to the energetic measurements. Changes in  $m$  values across mutants suggest structural heterogeneity in the folded state of PagP (discussed below; see Fig. 6). Therefore,  $\Delta\Delta G^0_{w,i(\phi)}$  is a reliable, if not absolute, measure for the partitioning free energy of each residue at the interface.

Although the  $\Delta\Delta G^0_{w,i(\phi)}$  values of the majority of residues correlated well with the non-polar ASA, we obtained interesting outliers, such as Cys, Ser, and Trp, whose  $\Delta G^0$  was heavily skewed. These anomalies might be due to under- and overrepresentation in the non-polar surface area calculations, respectively. The tryptophan side chain appears to be more hydrophilic than predicted, despite its bulky aromatic nature. Indeed, a detailed mutational analysis on the energetic contribution of Trp to the  $\beta$ -barrel OmpA revealed that the indole had a destabilizing effect at the interface (38), which is in line with our findings from PagP. Other studies have also reported a wide range of hydrophobicities for the indole

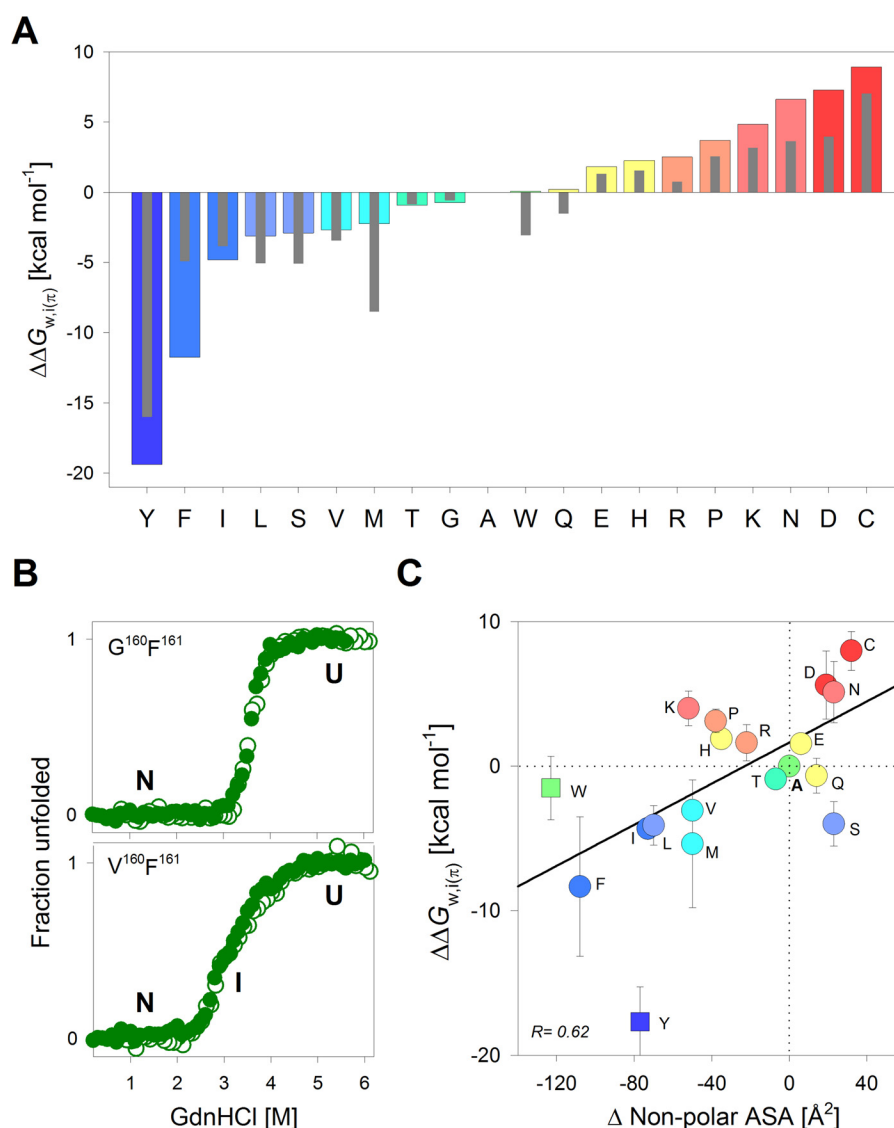
side chain, from most hydrophobic (Nozaki–Tanford (46), Wimley–White octanol (45), and interface (12) scales) to moderately polar (Kyte–Doolittle (47), Eisenberg–Weiss (48), Engelman–Steiz–Goldman (49), Hessa–von Heijne (9), and Moon–Fleming (7)). Considering that the tryptophan side chain exerts an energetic barrier for incorporation at the interface ( $\Delta\Delta G^0_{w,i(\phi)} = 1.4 \text{ kcal mol}^{-1}$ ), we surmise that this energetic cost could be offset by optimal binding to assembly factors and chaperones.

PagP-C<sup>161</sup> is the only mutant that deviates from the two-state folding transitions seen for all other PagP- $X^{161}$  mutants (Fig. 2B). We ruled out the possibility of oxidative modifications of cysteine by carrying out all experiments for these mutants under reducing environments (see gel image in [supplemental Fig. S5](#)). Further, existing hydrophobicity scales support both a hydrophilic (supported by the works of Kyte–Doolittle (47), Wolfenden (50), and Wimley–White (45) due to its polar side chain) and hydrophobic (from the Rose (51) and Janin (52) scales, due to its occurrence in hydrophobic regions in native protein structures and correlation with the non-polar ASA) nature for this residue. Such duality in the nature of cysteine, and the formation of additional interactions of the thiol with the detergent headgroup, could account for its anomalous  $\Delta\Delta G^0_{w,i(\phi)}$  at the membrane. Interestingly, the behavior of serine is analogous to that of cysteine, not threonine. Although we do not have a convincing explanation for this anomaly, we speculate that similar local interactions formed by oxygen (in Ser) and sulfur (in Cys) might give rise to the observed deviations in the  $\Delta G^0$  of PagP-S<sup>161</sup> and PagP-C<sup>161</sup>.

### Thermodynamic contribution of the penultimate interface residue to barrel folding

The residue Gln<sup>160</sup> is located at the interface in PagP, and the side chain points to the barrel interior (Fig. 1A). Gln<sup>160</sup> is also a likely part of the PagP folding nucleus (29). A systematic substitution at residue 160 (PagP- $X^{160}$  library) would allow us to measure side-chain contributions during formation of the PagP folding nucleus. Here, we could explain the equilibrium folding transitions for some PagP- $X^{160}$  mutants using a two-state model (yielding the  $\Delta G^0$ ). However, others required a three-state model to derive the total  $\Delta G^0$  of the folding process (Fig. 3 (A and B) and [supplemental Figs. S3, S6, and S7](#)). In the latter case, the  $\Delta G^0$  is the sum of the change in free energy from the unfolded to the intermediate (I) state ( $\Delta G_{U \rightarrow I}$ ) and the intermediate to the native state ( $\Delta G_{I \rightarrow N}$ ).

The energetic cost of partitioning of each residue at position 160 ( $\Delta\Delta G^0_{w,i(\pi)}$  with respect to alanine ( $\Delta\Delta G^0_{w,i(\pi)} = \Delta G^0_A - \Delta G^0_X$ ) showed a heterogeneous distribution of polar, hydrophobic, and charged residues (Fig. 3A). Unlike the  $\Delta\Delta G^0_{w,i(\phi)}$  at residue 161, the  $\Delta\Delta G^0_{w,i(\pi)}$  at residue 160 represents the energy required to partition an amino acid side chain from the aqueous solvent to the polar protein interior at the interface. Here, we found a skewed polarity scale, with hydrophobic residues (Tyr, Ile, and Phe) favored over hydrophilic residues. In particular, Tyr deviates considerably as it displays the highest  $\Delta\Delta G^0_{w,i(\pi)}$ . We speculate that the -O<sup>δ</sup>H of Tyr, which adds a polar and hydrophilic nature to the phenol ring, may allow Tyr to form both polar and non-polar interactions in PagP, thereby increas-



**Figure 3. Skewed polarity for partitioning of penultimate protein-facing interface residue of PagP.** *A*, water-to-interface partitioning free energy values calculated for all 20 residues at the protein-facing hydrophilic interface position 160, normalized with respect to alanine. *Colored bars and thin gray bars* represent the free energies generated from the PagP- $\chi^{160}\text{F}^{161}$  and PagP- $\chi^{160}\text{L}^{161}$  mutant libraries, respectively. Histograms are *colored from blue to red* in decreasing order of partitioning free energy values for PagP- $\chi^{160}\text{F}^{161}$ . *B*, representative equilibrium chemical denaturation profiles showing overlay of folding (*open circles*) and unfolding (*filled circles*) curves for two mutants (see [supplemental Figs. S3 and S6](#) for the complete data). *C*, correlation between  $\Delta\Delta G_{w,i(\pi)}$  and the change in non-polar ASA is shown as a scatter plot.  $\Delta\Delta G_{w,i(\pi)}$  values are averaged from PagP- $\chi^{160}\text{F}^{161}$  and PagP- $\chi^{160}\text{L}^{161}$  mutant libraries (*error bars* represent deviation from this average). Linear fits to the correlation are represented as *solid black lines*. Points that are excluded from the fits are shown as *square symbols*. The *color code* for the scatter plots is retained from *A*. Also see [supplemental Fig. S9](#).

ing the  $\Delta\Delta G_{w,i(\pi)}^0$ . Further, the  $\Delta\Delta G_{w,i(\pi)}^0$  is not affected by the aromatic nature of residue 161. We infer this by comparing the  $\Delta\Delta G_{w,i(\pi)}^0$  derived for PagP- $\chi^{160}$  mutants bearing an additional Phe<sup>161</sup> → Leu substitution (PagP- $\chi^{160}\text{L}^{161}$  mutants) ([supplemental Fig. S8](#)). We found that the  $\Delta\Delta G_{w,i(\pi)}^0$  was largely similar across all mutants, and it was only corrected for the  $\Delta G^0 \sim 0.3$  kcal mol<sup>-1</sup> for Phe<sup>161</sup> → Leu (Fig. 3A). Further, the  $\Delta\Delta G_{w,i(\pi)}^0$  did not correlate well with the change in amino acid ASA (Fig. 3C and [supplemental Fig. S9](#)).

Although we found that hydrophobic residues are more stabilizing at position 160, polar residues are preferred evolutionarily. To address this, we reexamined the folding profiles. We found that when Xaa<sup>160</sup> = Trp, Tyr, Phe, Leu, Ile, Val, or Met (hydrophobic residues), a stable thermodynamic intermediate was seen in the equilibrium folding profiles. We surmise that

hydrophobic residues at position 160 might promote alternative folding pathways for PagP. Gln<sup>160</sup> establishes multiple polar contacts in the folded PagP barrel, which can be disrupted by the formation of non-native interactions by hydrophobic residues at position 160. Due to the role of Gln<sup>160</sup> in forming the folding nucleus, the PagP- $\chi^{160}$  mutant library is well suited to provide insight into the folding mechanism of PagP. Based on our results, we speculate that the choice of the penultimate residue (Gln<sup>160</sup> in PagP) may be determined by a balance between achieving a smooth folding landscape and a stable  $\beta$ -barrel.

An interesting observation from the  $\Delta\Delta G_{w,i(\pi)}^0$  of the PagP- $\chi^{160}$  mutants is the remarkable difference of  $\sim 5.5$  kcal mol<sup>-1</sup> between Glu and Asp (or Gln and Asn). Both Glu and Asp have similar chemical properties, except for the longer side chain

## Side-chain interface energetics in PagP folding

(additional C<sup>γ</sup>H<sub>2</sub>) in Glu and Gln. To account for this difference in  $\Delta\Delta G^0_{w,i(\pi)}$ , we examined the crystal structure of PagP (PDB entry 1THQ) (26). We found that Gln<sup>160</sup> O<sup>ε1</sup> and N<sup>ε2</sup> establish weak electrostatic interactions with three spatially proximal residues, His<sup>22</sup> (3.5 Å), Asp<sup>24</sup> (4.7 Å), and Trp<sup>60</sup> N<sup>ε1</sup> (4.3 Å) (supplemental Fig. S10). Mutating Gln<sup>160</sup> to Asp<sup>160</sup> *in silico* abolishes these interactions and additionally places the carboxyl side chain of Asp<sup>160</sup> within 1.7 Å of Asp<sup>24</sup> (supplemental Fig. S10), giving rise to possible charge–charge repulsion. Such unfavorable interactions are less pronounced in PagP-E<sup>160</sup> or PagP-Q<sup>160</sup> (PagP-WT) (see supplemental Fig. S10). The destabilizing effect of Asp<sup>160</sup> is also evident from the lack of a prominent gel mobility shift in the folded PagP-D<sup>160</sup> protein, although the protein is resistant to proteolysis (supplemental Fig. S10). Hence, the additional C<sup>γ</sup>H<sub>2</sub> group of Glu and Gln positions the side-chain functional groups at residue 160 away from the carboxyl group of Asp<sup>24</sup>, giving rise to a more favorable  $\Delta G^0$  than its Asp and Asn counterparts, accounting for the measured difference in  $\Delta\Delta G^0_{w,i(\pi)}$  values.

To confirm that the destabilization is indeed due to Asp<sup>160</sup>, we compared the results from PagP-Q<sup>160</sup>F<sup>161</sup> and PagP-Q<sup>160</sup>L<sup>161</sup> with two additional mutants, PagP-D<sup>160</sup>G<sup>161</sup> and PagP-D<sup>160</sup>H<sup>161</sup>, as well as their parent constructs PagP-Q<sup>160</sup>G<sup>161</sup> and PagP-Q<sup>160</sup>H<sup>161</sup> (supplemental Fig. S10). This allowed us to rule out the influence of side-chain and backbone contributions as well as conformational constraints imposed by the lipid-facing residue on position 160. We found that in all cases, when Asp<sup>160</sup> was present, the  $\Delta G^0$  was lowered, independent of the terminal residue. These results support our conclusion that a residue-dependent contribution to the folding free energy of PagP at the penultimate position (residue 160) depends on the chemical nature and the length of the amino acid side chain. A similar destabilizing effect was seen in PagP-C<sup>160</sup> (Fig. 3A), wherein the thiol side chain might acquire a negative charge at our experimental pH of 9.5 (pK<sub>a</sub> of the thiol group of Cys = 8.3).

### Influence of hydrophobicity of interface residue 160 on folding intermediates

Under thermodynamic equilibrium, the energetics we measure for the protein-facing PagP-X<sup>160</sup> mutants differ in the nature of the folding transitions, because the presence of a hydrophobic residue gives rise to a prominent folding intermediate (I). Consequently, the total  $\Delta G^0_{U\rightarrow N}$  is a sum of  $\Delta G^0_{U\rightarrow I}$  (change in  $\Delta G^0$  from U to I) and  $\Delta G^0_{I\rightarrow N}$  (change in  $\Delta G^0$  from I to N). Because the  $\Delta G^0$  or free energy is a state function, the presence of the intermediate should not affect the free energy calculations. However, the nature of Xaa<sup>160</sup> affects the final folded state of PagP (detailed below). Assessment of the residue contribution at the protein-facing interface ( $\Delta\Delta G^0_{w,i(\pi)}$ ) can now vary based on whether  $\Delta G^0_{U\rightarrow I}$  or  $\Delta G^0_{I\rightarrow N}$  represents the true folding transition (supplemental Fig. S11). Hence, we examined the source of the rougher PagP folding landscape by comparing the folding behavior in DLPC small unilamellar vesicles (SUVs) (supplemental Fig. S12).

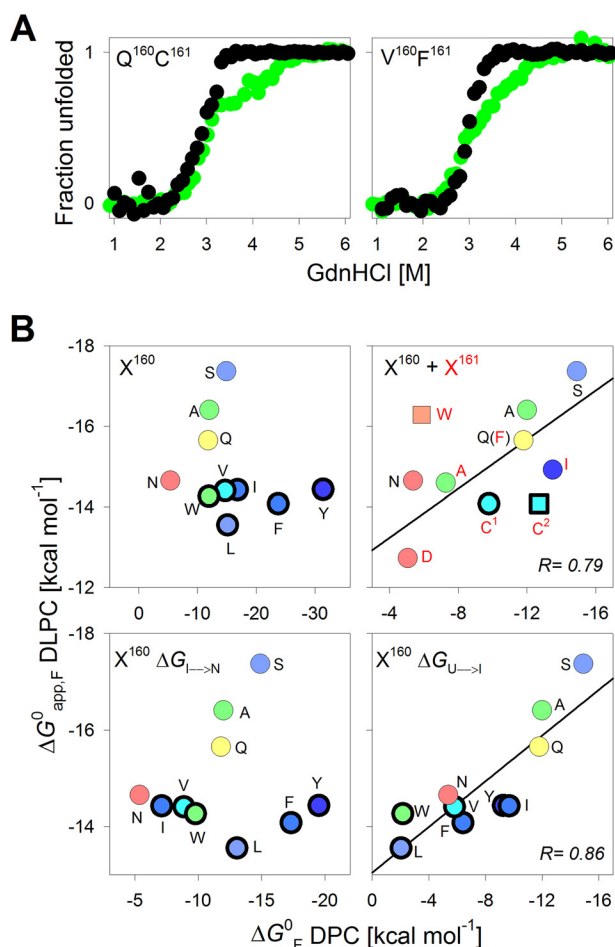
In membrane protein folding studies, large unilamellar vesicles (LUVs) are considered as a more reliable membrane mimetic as compared with SUVs. SUVs are not preferred due to

their heterogeneity and non-equilibrium nature (53). However, we believe that this did not interfere in our measurements, because we used SUVs solely to compare the behavior of PagP across residue types. We found that full-length PagP exhibited hysteresis in both LUVs and SUVs despite prolonged incubation. The system was no longer under equilibrium, and exhibited path dependence. The measured free energy in such systems is considered an apparent value ( $\Delta G^0_{app}$ ) (54). Therefore, we derived the apparent free energy of folding ( $\Delta G^0_{app,F}$ ) for PagP in DLPC vesicles and correlated it with the  $\Delta G^0$  obtained in DPC micelles.

In DLPC, the folding profiles of all of the mutants that we tested showed a two-state transition with a global *m* value of  $\sim 4.80 \text{ kcal mol}^{-1} \text{ M}^{-1}$ , which is similar to the  $\sim 5.4 \text{ kcal mol}^{-1} \text{ M}^{-1}$  reported for PagP-WT in DLPC LUVs (6). We also observed no detectable intermediates (Fig. 4A). A similar result is seen in the high detergent/protein ratio (DPR) ( $\sim 25,000:1$ ) of DPC (supplemental Fig. S13). Hence, DLPC and high DPR of DPC allow for a concerted folding of PagP-X<sup>160</sup> mutants (supplemental Fig. S12), and the intermediate seen when Xaa<sup>160</sup> is a hydrophobic residue is due to the low DPR conditions of DPC used in this study. Because sufficient DPC molecules exist as micelles in our experimental conditions, the roughness of the PagP folding landscape and the likely formation of non-native structures seems to correlate directly with the hydrophobicity of residue 160 (supplemental Fig. S13).

We compared the total  $\Delta G^0$  with the  $\Delta G^0_{app,F}$  obtained from DPC and DLPC, respectively, for the PagP-X<sup>160</sup> and PagP-X<sup>161</sup> mutants. Interestingly, despite differences in the absolute  $\Delta G^0$  values, only the mutants showing a two-state transition showed an excellent linear correlation for the total  $\Delta G^0$  in DPC and DLPC (Fig. 4B). We observed no such correlation for mutants that showed a three-state folding transition. In the latter case, we obtained a strong linear correlation between  $\Delta G^0_{U\rightarrow I}$  and  $\Delta G^0_{app,F}$  ( $r = 0.86$ ), but not between  $\Delta G^0_{I\rightarrow N}$  and  $\Delta G^0_{app,F}$  (Fig. 4B). Similarly, the  $\Delta G^0$  in high DPR of DPC was close to the  $\Delta G^0_{U\rightarrow I}$  observed in low DPR. These results suggest that the U  $\rightarrow$  I transition in DPC ( $\Delta G^0_{U\rightarrow I}$ ) is a better measure of the global folding process, whereas the I  $\rightarrow$  N transition ( $\Delta G^0_{I\rightarrow N}$ ) might represent the formation of non-native interactions or an alternate folding pathway that is populated in mutants where Xaa<sup>160</sup> is a hydrophobic residue (supplemental Fig. S13). We found that folding of PagP was hampered when a hydrophobic residue was introduced at position 160, whereas it was two-state when a polar residue was present.

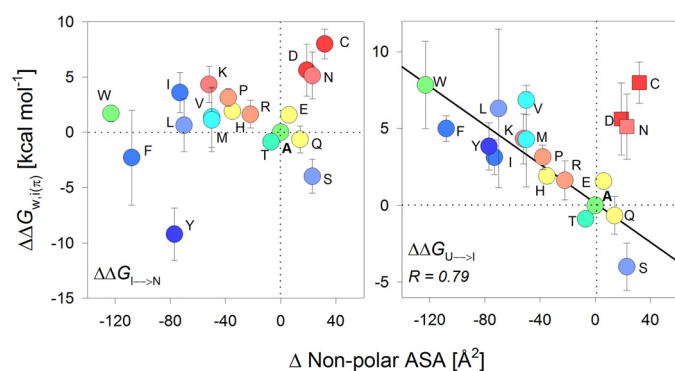
The data for PagP-W<sup>161</sup> and PagP-C<sup>161</sup> in DLPC merit discussion. Upon comparing the DPC data for PagP-W<sup>161</sup> with the  $\Delta G^0_{app,F}$  measured in DLPC (Fig. 4B), we observe that although the Trp side chain is favored at the interface in DLPC vesicles, the magnitude of stabilization does not measure up to the values presented by established interface hydrophobicity scales (10, 12). Hence, the contribution of the indole might be protein-dependent. PagP-C<sup>161</sup> is the only mutant in the PagP-X<sup>161</sup> library that exhibits a three-state folding transition. Hence, we compared the  $\Delta G^0_{U\rightarrow I}$  and  $\Delta G^0_{I\rightarrow N}$  in DPC for PagP-C<sup>161</sup> with the  $\Delta G^0_{app,F}$  measured in DLPC (Fig. 4B). Surprisingly, we found here that the  $\Delta G^0_{I\rightarrow N}$ , but not  $\Delta G^0_{U\rightarrow I}$ , shows a good correlation with  $\Delta G^0_{app,F}$  (DLPC). Although this is in



**Figure 4. Energetics in DLPC demarcate the two transitions observed for PagP- $X^{160}$  mutants with hydrophobic substituents.** A, overlay of representative folding profiles for two PagP variants that exhibit three-state equilibrium folding transitions in 10 mM DPC (green) and the corresponding data in 0.8 mM DLPC (black). B, scatter plots correlating the  $\Delta G^0$  of PagP in DPC and  $\Delta G^0_{app,F}$  in DLPC. The Xaa<sup>161</sup> mutants (red) and the Xaa<sup>160</sup> mutants (black) are represented by their single-letter codes. Variants with three-state profiles in DPC are shown as symbols with thick edges. Along the x axis, the two upper panels depict the total  $\Delta G^0$  ( $\Delta G^0_{U \rightarrow N}$  for the two-state transitions or  $\Delta G^0_{U \rightarrow I} + \Delta G^0_{I \rightarrow N}$  for the three-state transitions) determined for PagP- $X^{160}$  mutants in DPC micelles. Along the x axis, the two bottom panels show  $\Delta G^0_{I \rightarrow N}$  (left) and  $\Delta G^0_{U \rightarrow I}$  (right) for the three-state transitions in DPC micelles and are presented along with the  $\Delta G^0$  (DPC) for the two-state transitions. Linear fits to the correlation are represented as solid black lines. A good correlation ( $r = 0.79$ ) is observed between  $\Delta G^0$  (DPC) and  $\Delta G^0_{app,F}$  (DLPC) only for the PagP- $X^{160}$  and PagP- $X^{161}$  mutants showing a two-state transition (top right). The total  $\Delta G^0$  in DPC of PagP-C<sup>161</sup> is split into  $\Delta G^0_{I \rightarrow N}$  (C<sup>1</sup>) and  $\Delta G^0_{U \rightarrow I}$  (C<sup>2</sup>).

contrast to the correlation seen for the PagP- $X^{160}$ , the reason for such a deviation is currently unclear. A likely possibility is the formation of non-native hydrophobic contacts by the thiol group and lipidic micelles in the unfolded state of PagP. However, further experiments are required to validate this argument.

We also found a strong inverse correlation between the  $\Delta \Delta G^0_{w,i(\pi)}$  and the non-polar ASA for each amino acid side chain when we consider  $\Delta G^0_{U \rightarrow N}$  (two-state) and  $\Delta G^0_{U \rightarrow I}$  (three-state) (Fig. 5). This inverse correlation was the opposite of the direct linear correlation we obtained for the lipid-facing residue (Fig. 2C) and is expected for the polar environment of the residue at position 160.



**Figure 5. Partitioning free energy at the protein-facing interface correlates inversely with the change in non-polar ASA of the guest residue.** Correlation plots were generated by mapping the change in non-polar ASA with the total change in folding free energy  $\Delta \Delta G^0_{U \rightarrow N}$  for the PagP- $X^{160}$  mutants that exhibit two-state profiles. The plot also includes the  $\Delta \Delta G^0_{w,i(\pi)}$  for the PagP- $X^{160}$  mutants exhibiting three-state folding profiles (W, Y, F, L, I, V, and M) divided into  $\Delta \Delta G^0_{I \rightarrow N}$  (left) and  $\Delta \Delta G^0_{U \rightarrow I}$  (right). A linear (inverse) fit with a high regression coefficient (solid line;  $r = 0.79$ ) could be obtained only when the  $\Delta \Delta G^0_{U \rightarrow I}$  was considered (right). Points that are excluded from the fits are shown as square symbols. The color code for the scatter plots is retained from Fig. 3A.

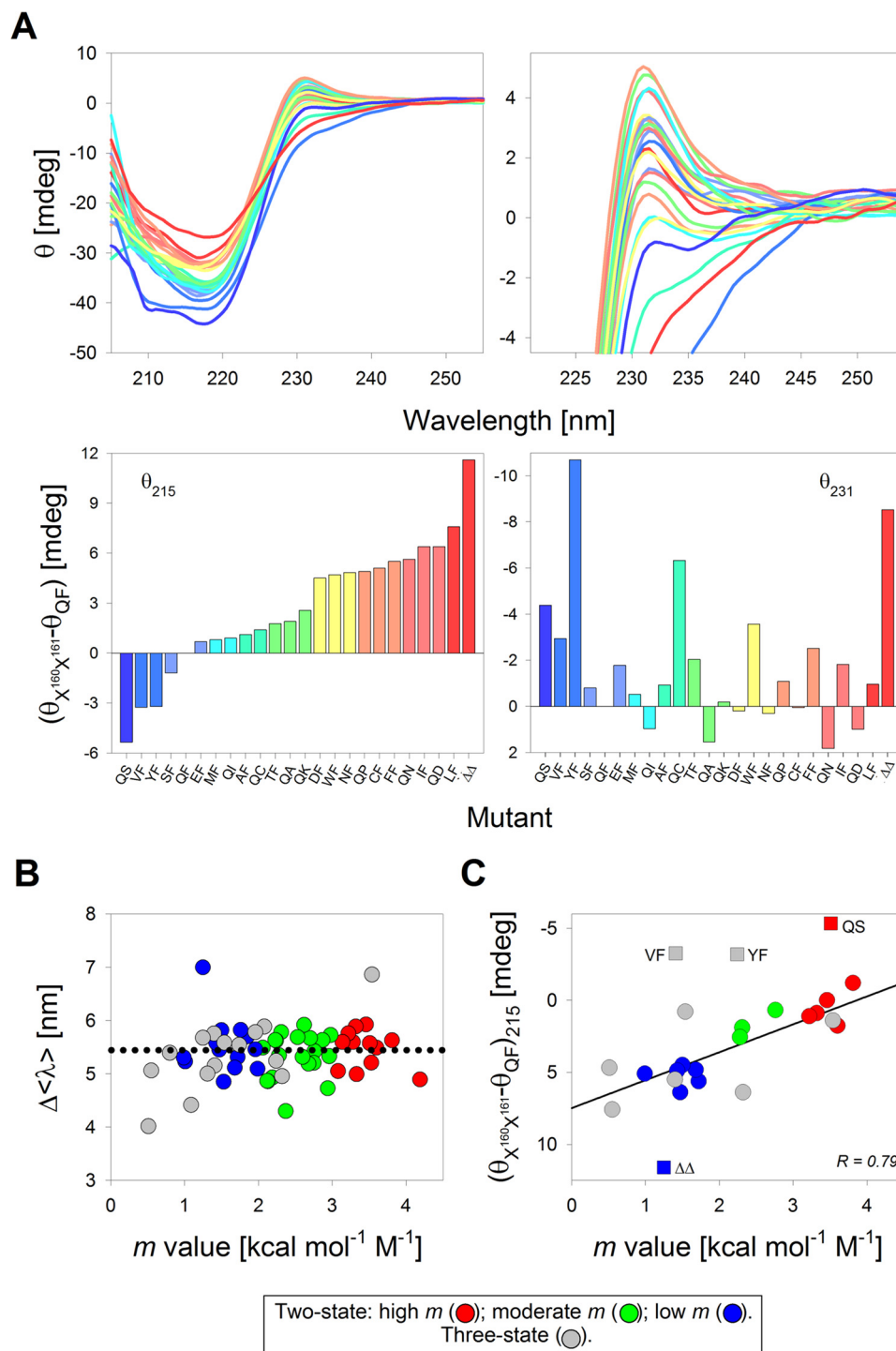
#### Correlating partitioning free energy with atomic solvation, $m$ value, and PagP folding landscape in micelles

The process of protein folding (*i.e.* the change from U to N state) is a highly cooperative process. This is accompanied by the change in the solvent-accessible surface area as the protein folds and is reflected in the important thermodynamic parameter termed the  $m$  value. By definition, the  $m$  value is the difference in the solvent-accessible surface area ( $\Delta ASA$ ) between the U and N states, and it increases proportionately with the size of the protein. However, membrane proteins, including PagP, bury a larger area of their accessible surface upon folding; hence, PagP shows a higher  $m$  value than soluble proteins of a similar size (6, 55).

Our thermodynamic analysis reveals that PagP- $X^{160}$  and PagP- $X^{161}$  mutants exhibit a range of  $m$  values for the two-state ( $1-4 \text{ kcal mol}^{-1} \text{ M}^{-1}$ ) and three-state ( $0.5-8.0 \text{ kcal mol}^{-1} \text{ M}^{-1}$ ) folding transitions (supplemental Fig. S14). This could arise from a change in the structure of PagP upon folding or the presence of residual structure in the unfolded state. Point mutations can alter the residual structure in the unfolded protein state without affecting the folded state and thereby change the  $m$  value (56). In such a case, mutants with residual structure in the unfolded state have a lower  $m$  value, whereas the mutants without any residual structure have a higher  $m$  value (37, 56). A comparison of the emission profiles of the unfolded PagP mutant proteins ( $\langle \lambda_U \rangle$ ) suggests that the unfolded protein state might not possess any significant residual structure (supplemental Fig. S15). We do not rule out the possible existence of structural heterogeneity in the unfolded state that we are unable to measure using Trp fluorescence.

Interestingly, far-UV and near-UV CD structural analysis at 215 nm ( $\theta_{215}$ ) and 231 nm ( $\theta_{231}$ ) (57) show that the increase in the secondary structure content ( $\theta_{215}$ ) of folded PagP can account for the observed increase in  $m$  values (Fig. 6 and supplemental Fig. S16). Comparable values of tryptophan lifetimes (average value of  $\sim 3.39$  ns) and anisotropy ( $\sim 0.098$ ) obtained for  $\sim 20$  representative mutants in their

## Side-chain interface energetics in PagP folding



**Figure 6. Secondary structure content of the PagP- $X^{160}X^{161}$  mutants correlates with the measured  $m$  value.** A, far-UV CD wavelength scans recorded to examine the secondary structure content (top left) and tertiary aromatic interactions attained upon barrel assembly (top right) for a subset of PagP variants (PagP- $X^{160}$  and PagP- $X^{161}$ ) displaying a range of  $m$  values in DPC. The corresponding raw ellipticity values ( $\theta$ , in millidegrees) at 215 and 231 nm normalized with respect to PagP-WT (WT) are presented as histograms in the bottom panels. The mutants are arranged and color-coded in decreasing order of  $\theta_{215}$ , from dark blue (highest  $\theta_{215}$ ) to dark red (lowest  $\theta_{215}$ ) in the two left panels. This color scheme is retained in the right panels. Mutants are named with single-letter codes for residues present at the terminal two positions; PagP- $\Delta Q^{160}F^{161}$  is depicted as  $\Delta\Delta$ . B, correlation plot between the change in  $\theta_{215}$  and the  $m$  value for all of the mutants shown in A (also see supplemental Fig. S16). The scatter plot is colored based on the  $m$  value of the mutants, with the two-state mutants divided into high (red circles), moderate (green circles), and low (blue circles)  $m$  values and the  $m_{U \rightarrow 1}$  for the three-state mutants represented as gray circles. C, correlation plot between the change in  $\theta_{215}$  and the  $m$  value for all of the mutants shown in A (also see supplemental Fig. S16). The color code is retained from B. A linear fit to the correlation is shown as a solid black line ( $r = 0.79$ ). Points excluded from the fit are shown as square symbols.

folded state (supplemental Table S2) support the formation of a well-structured PagP  $\beta$ -barrel scaffold. Hence, the difference in the  $m$  value appears to be a function of the secondary

structure attained in the folded state (Fig. 6C). Additionally, the  $m$  value shows a strong linear correlation with the difference in non-polar ASA of each lipid- or protein-facing resi-



due (supplemental Fig. S17). The  $m$  value is expected to change under conditions of hydrophobic mismatch between the transmembrane region and the lipid bilayer (58). However, DPC micelles impose little or no lateral pressure during PagP folding and are highly dynamic entities when compared with DLPC. Hence, DPC might allow structural heterogeneity in the folded state of mutants, thereby altering the  $m$  value in a residue-dependent manner.

A quantitative indication of the energetics involved in the partitioning process between two phases is provided by the atomic solvation parameter (ASP). At the water-lipid interface, the ASP value that we derive for the insertion of hydrophobic residues is  $0.043 \pm 0.01 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$  ( $r = 0.90$ ) (Fig. 2C), which is  $>3$  times higher than the ASP of  $\sim 0.0131 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$  for the Wimley–White interface scale (12) but closer to the  $\sim 0.023 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$  obtained for folding of the OmpLA  $\beta$ -barrel into phosphocholine vesicles (7). The abnormally large value of ASP may arise from ASA ( $m$  value) changes between the folded and unfolded states of PagP mutants in DPC micelles. However, we obtain a strong linear correlation (Fig. 4B,  $r = 0.79$ , slope = 0.31) for our data from DPC micelles and DLPC vesicles. Moreover, DPC reorganization and local structural heterogeneity do not influence our measurements. Therefore, the structural variability of the folded state, which appears to be the major drawback of the PagP–DPC system, prevents quantitative assessment of side-chain  $\Delta\Delta G^0$  but provides a reliable qualitative readout for interface energetics.

The high value of ASP that we observed might also be explained as follows. When we compared the free energy values for PagP–WT in DPC micelles and DLPC SUVs, we observed a difference of  $\sim 4 \text{ kcal mol}^{-1}$  (see Fig. 4B). The process of folding in vesicles yielded a higher free energy value. This could be due to one of two reasons: (i) folded PagP establishes better protein–lipid contacts in DLPC vesicles than in DPC micelles due to matching hydrophobic thickness, or (ii) the native state is identical in both systems, but the denatured state ensemble is not completely unfolded in both cases. By considering either possibility, we can conclude that in DLPC vesicles, we are able to capture the thermodynamic transition of a higher magnitude. And yet, when we correlate the  $\Delta G^0$  for DPC micelles with that of the DLPC SUVs, it is the micellar reaction that exhibits the higher magnitude (slope = 0.31 for DLPC *versus* DPC) (Fig. 4B). This would imply that the hydrophobic contribution per residue is magnified in the case of DPC micelles. Therefore, what works as an inhibitory factor in the computation of absolute free energy values (*e.g.* dynamicity of DPC micelles and the occurrence of a rougher folding landscape that probably promotes non-native interactions) proves to be an amplifying factor when the side-chain hydrophobicity is measured.

We found that the physico-chemical property of the residue 160 side chain regulates the  $\beta$ -barrel thermodynamic free energy landscape. Gln<sup>160</sup> is a part of the PagP folding nucleus (29); therefore, side-chain partitioning energetics derived from the PagP–X<sup>160</sup> library must be interpreted with caution. Here, we found that polar side chains are preferred over large bulky groups, because the latter influences the PagP folding landscape by stabilizing folding intermediates. On the other hand, the selection of Phe<sup>161</sup> has previously been attributed as important

for both its hydrophobicity and aromaticity (59). However, we found that Tyr and Trp are less stabilizing than Phe, whereas aliphatic hydrophobic residues, such as Ile, Val, and Leu, are preferred. Hence, we surmise that the hydrophobic contribution of Phe<sup>161</sup> is more pronounced than its aromaticity. Furthermore, extrapolating the linear correlation between DPC and DLPC provides us with an energy contribution of  $\sim 2 \text{ kcal/mol}$  for the interface hydrogen bond in DLPC. In comparison, an interface hydrogen bond contributes  $\sim 0.6 \text{ kcal mol}^{-1}$  to the stability of transmembrane helices (60). Hence, the contribution of hydrogen bonding appears to be significant for PagP folding, especially when we consider that strand assembly and barrel closure in PagP is driven by interstrand hydrogen bonding.

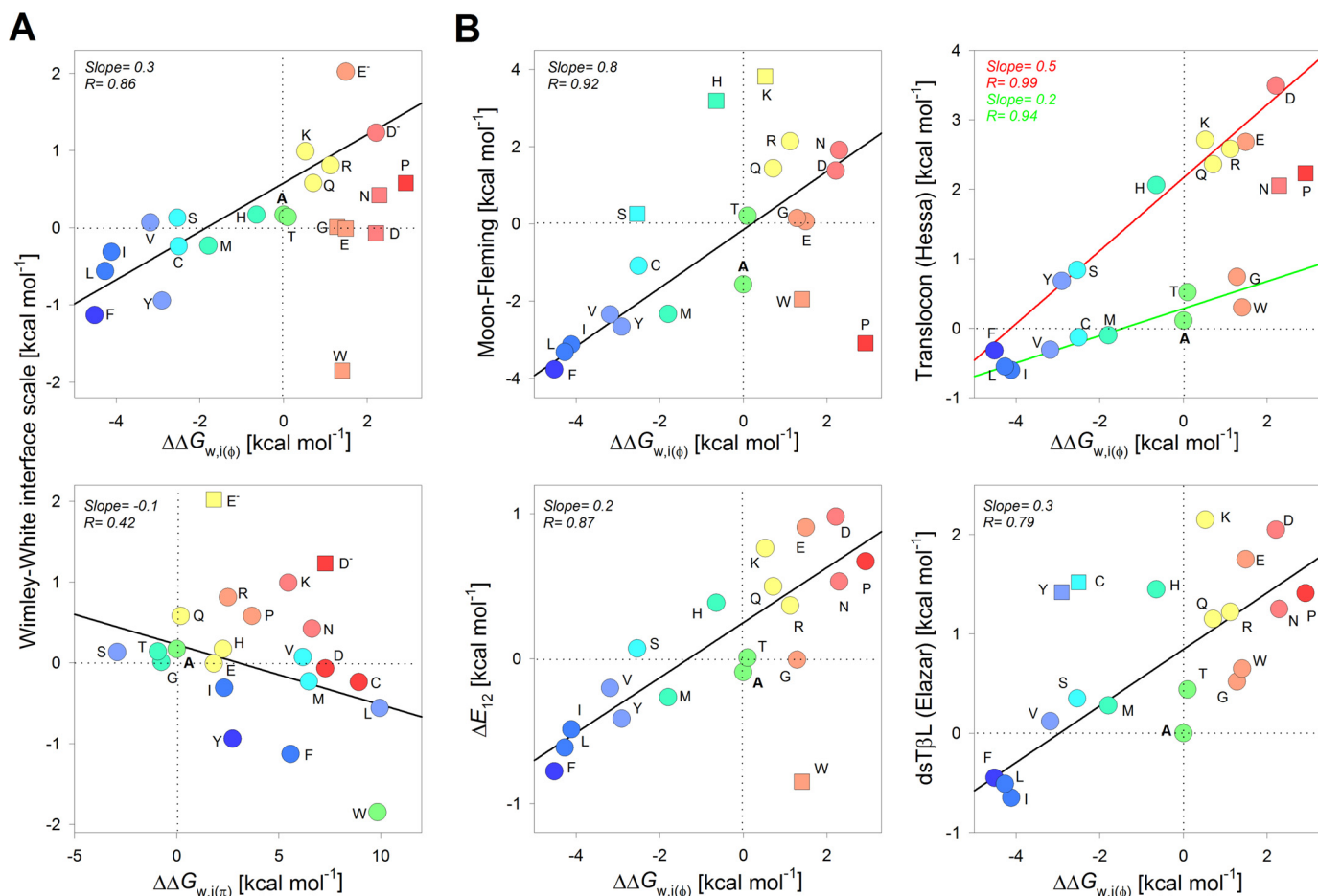
### Correlating side-chain transfer free energy for $\beta$ -signal residues with reported scales

A major driving factor in membrane protein folding is the partitioning of hydrophobic residues from the aqueous milieu into the lipid bilayer. Not surprisingly, a preferential distribution of residues was observed in membrane proteins, as the structure traverses the bilayer (14, 18). Several theoretical and experimental studies have correlated the hydrophobicity of an amino acid side chain with the free energy of its transfer from water to an apolar environment (7, 9, 11, 12, 45, 47). Hence, we asked whether the  $\Delta\Delta G^0_{w,i}$  that we measured for each residue at the  $\beta$ -signal correlates with established hydrophobicity scales (Fig. 7 and supplemental Figs. S18 and S19). Although our thermodynamic measurements with the PagP–X<sup>161</sup> interface mutants were in DPC micelles (Fig. 2A), and results from such a simplified system cannot be directly extrapolated to conditions *in vivo*, considering the limited information available for the interface, a comparison of our  $\Delta\Delta G^0_{w,i}$  values with existing scales might provide interesting insights.

For the lipid-facing interface ( $\Delta\Delta G^0_{w,i(\phi)}$ ; Fig. 7), fits of the correlation plots with the Wimley–White interface scale (12) gave us a slope of 0.3 ( $r = 0.86$ ), the Moon–Fleming scale (7) gave a slope of 0.8 ( $r = 0.92$ ), the  $E_z$  potential (13) at 12 Å from the midplane (as an indicator of interface position) gave a slope of 0.2 ( $r = 0.87$ ), and the dsT $\beta$ L scale (11) gave a slope of 0.3 ( $r = 0.79$ ). Although we observe good correlations for the  $\Delta\Delta G^0_{w,i(\phi)}$  with all of the scales, the low value of slope of  $\leq 0.3$  indicates that the energetic cost of partitioning is overestimated in our measurements. Moreover, the structural heterogeneity of the folded state (Fig. 6C) and associated changes in ASA pose major limitations to our current analysis. The Wimley–White interface scale includes the contribution of the peptide bond in the partitioning free energy calculations and might account for the low value of slope. Nevertheless, one similarity in the correlations is that the positively charged residues are more favorable at the water–bilayer interface compared with negatively charged residues.

The high slope (0.8) and correlation ( $r = 0.92$ ) for the  $\Delta\Delta G^0_{w,i(\phi)}$  with the Moon–Fleming scale (Fig. 7B) might suggest that the lipid-facing  $\beta$ -signal residue of PagP in DPC favors hydrophobicity for partitioning at the interface. Most known scales use transmembrane helices as their model systems for partitioning free energy measurements (note how R is high for

## Side-chain interface energetics in PagP folding



**Figure 7. Comparison of interface energetics for PagP  $\beta$ -signal residues with previously reported scales.** A, correlation plots generated by comparing the  $\Delta\Delta G_{w,i(\phi)}^0$  (derived from the lipid-facing interface position; top) and the  $\Delta\Delta G_{w,i(\pi)}^0$  transfer free energy (derived from the polar protein-facing interface position; bottom) with the Wimley–White interface scale (12). Here, the  $\Delta\Delta G_{w,i(\pi)}^0$  scale is derived from the  $\Delta\Delta G_{U \rightarrow I}^0$  (W, Y, F, L, I, V, and M) or  $\Delta\Delta G_{U \rightarrow N}^0$  observed at the penultimate residue. B, comparison of  $\Delta\Delta G_{w,i(\phi)}^0$  with the whole-protein (7), translocon (9),  $\Delta E_{12}$  (13), and dsTβL (11) scales. Linear fits to the correlation are represented as solid lines. Points excluded from the fits are shown as square symbols. Color codes for residues are retained from Fig. 2A for  $\Delta\Delta G_{w,i(\phi)}^0$  and Fig. 3A for  $\Delta\Delta G_{w,i(\pi)}^0$ . The slope and corresponding regression coefficients are indicated within each panel and are color-coded for the fits to the translocon scale in B.

the correlation with both the  $\Delta E_{12}$  and dsTβL scales in Fig. 7B, whereas the slope is low). Unlike helical proteins,  $\beta$ -sheet structures occlude more non-polar surface area (61). The Moon–Fleming scale uses a  $\beta$ -barrel protein, which could also account for why we obtain the higher slope of 0.8 (Fig. 7B). The noticeable deviation is for partitioning of charged residues; whereas the Moon–Fleming scale assigns a higher energy penalty at bilayer midplane, it appears to be more energetically favorable at the interface (Fig. 7B, top left).

The hydrophobicity of tryptophan exhibits a noticeable difference in most of the correlations. Whereas all scales (except dsTβL) assign an energetic preference for tryptophan at the interface due to its aromaticity and amphiphilic nature, our measurements seem to account for only the hydrophobicity of the indole (Fig. 7). As a result, the partitioning of Phe and Tyr at the interface are more favorable than that of Trp. Upon correlation with the depth-dependent hydrophobicity scale, the  $E_z$  potential (13), we observe that Trp is an outlier for both the midplane and interface positions (Fig. 7B and supplemental Figs. S19 and S20). This deviation is not surprising, because studies using OmpA have shown that Trp exhibits dual behavior (hydrophobicity and amphiphilic nature) in membrane pro-

teins (37). Similarly, PagP-W<sup>161</sup> exhibits a destabilizing effect, because Trp appears to be more hydrophilic than its calculated non-polar ASA at the interface. Electrostatic interactions of the indole with DPC may explain the reduced hydrophobicity of the indole. The contribution of an interface Trp to membrane protein stability may depend on its local environment and, therefore, be protein-specific. Such a protein-specific contribution could explain why the  $\Delta\Delta G_{w,i(\phi)}^0$  that we measure for Trp correlates poorly with established scales (Fig. 7).

The  $\Delta\Delta G_{w,i(\phi)}^0$  also correlates conditionally with the translocon scale (9), wherein hydrophobic residues (Ile, Phe, Leu, Val, Cys, Met, Ala, Trp, and Gly) correlate with a slope of 0.2 ( $r = 0.94$ ), and hydrophilic residues (Ser, Gln, His, Leu, Arg, Glu, Asp, and Tyr) fit to the linear regression with a slope of 0.5 ( $r = 0.99$ ) (Fig. 7B). Notably, Tyr correlates with other hydrophilic residues. Such a conditional correlation brings out an interesting insight into our current understanding of membrane energetics. It is believed that the magnitude of the partitioning free energy in the translocon scale is underestimated (62), because it is believed to represent the side-chain transfer free energy from the interior of the translocon to the interior of the membrane instead of transfer from water to the membrane. As a result,

the translocon scale also shows a depressed value for the ASP (62). This effect is now evident when we compare only the hydrophobic residues (Fig. 7, *top right, green line*), wherein we observe a slope of 0.2, as opposed to the correlation seen with hydrophilic residues (slope = 0.5). Such a conditional correlation highlights the difference in partitioning energy of hydrophobic and hydrophilic residues into the lipid membrane.

For the protein-facing interface ( $\Delta\Delta G^0_{w,i(\pi)}$ ; Fig. 7A, *bottom*), the correlation shows a negative slope and a poor *r* value, affirming the anticipated inverse relationship between the partitioning free energy at the polar interface,  $\Delta\text{ASA}$  (non-polar) (see Fig. 5), and reported hydrophobicity scales. The main reason for this seemingly anomalous observation is that the protein-facing polar interface prefers small and polar side chains over charged amino acid residues due to steric constraints during scaffold packing (see [supplemental Fig. S10](#)). Further, charged residues, such as Arg and Lys, are placed closer to hydrophobic residues (note how  $\Delta\text{ASA}$  for Arg and Lys is higher than Ala in Fig. 5 (*right*)) in the non-polar ASA calculations than in conventional hydrophobicity scales, mainly due to the aliphatic segments in their side chains. Hence, our  $\Delta\text{ASA}$  calculation (Fig. 5) and  $\Delta\Delta G^0_{w,i(\pi)}$  reflect the tendency of the non-polar region of side chain to occlude itself from water.

## Discussion

The folding of the nascent OMP polypeptide is widely accepted to be a multistep process (63–65). Physico-chemical properties of the lipid (66), chaperone-dependent factors (67), and the energetic contribution of the protein primary sequence (6, 68) together play an important role in OMP assembly. The bilayer interface is amphipathic (zwitterionic), whereas the midplane is hydrophobic (apolar), yet we found that the side-chain partition energetics that we measured for the interface residues correlates surprisingly well to free energy measurements at the bilayer midplane (Moon–Fleming scale (7); Fig. 7B) and shows a matching magnitude. Hence, the energetic contribution for each residue at the  $\beta$ -signal might have considerable contributions from hydrophobic forces and change in non-polar ASA. The correlation that we observe with other reported biological hydrophobicity scales (Fig. 7 and [supplemental Figs. S18 and S19](#)) also indicates that physico-chemical factors that determine the partitioning free energy at the bilayer midplane (for both  $\alpha$ -helical and  $\beta$ -barrel structures) and the interface might indeed be similar and differ only in the magnitude of change in the measured free energy values (*e.g.* see [supplemental Fig. S20](#)). At the same time, the interesting deviations that we and others (38, 45, 47, 51) observe for specific residues, such as tryptophan and cysteine, might reflect a more protein-specific contribution in transmembrane  $\beta$ -barrels. For example, the residue preference at the terminal position, Phe over Trp ([supplemental Fig. S1A](#)), may be due to the higher stability provided by the Phe side chain.

We found that the mere hydrophobicity of an amino acid side chain at the  $\beta$ -signal motif is energetically sufficient to stabilize the barrel scaffold. Hence, we surmise that specific *in vivo* stop transfer signals (14) (Trp, Arg, Lys, etc.) should bear greater

functional relevance than serving as passive markers for membrane protein translocation. Furthermore, in accordance with Anfinsen's hypothesis (69), the asymmetric distribution of hydrophobic side chains at lipid-facing sites, polar residues at the extramembranous regions, and intra- and interprotein contact sites provides a sufficient molecular footprint to fold the nascent polypeptide. The interiors of many OMP barrels are enriched with polar residues that pose a high-energy barrier (7) when partitioned in the hydrophobic bilayer interior. Strong hydrophilic interactions established between charged and polar residues (as seen in transmembrane helices in the membrane (7, 70, 71)), along with interstrand backbone hydrogen bonds, might together contribute to the unusually high stability of OMPs (6).

Our study shows that the PagP folding pathway is strongly modulated by the physical properties of the lipidic environment, with higher lateral bilayer pressure promoting a smoother folding landscape. Similar observations in other model OMPs (58) suggest that *in vivo*, the lateral bilayer pressure works in concert with the barrel assembly machinery (BAM) (3, 67) complex to drive OMP folding. The BAM complex, lipid headgroup, and lateral bilayer pressure may be important to suppress off-pathway kinetic intermediates (66, 72, 73) or alternate folded structures (as seen in some of our PagP mutants). This argument also explains why the specific choice of residues at the C-terminal Aro-Xaa-Aro motif is important for unassisted OMP folding *in vitro*.

In conclusion, we find not only that the PagP C-terminal  $\beta$ -signal motif is a recognition signal for the Skp–BAM complex (22) but that it can considerably influence the folding cooperativity and stability of the folded  $\beta$ -barrel. The thermodynamic contribution of the lipid- and protein-facing interface residues that form a part of the crucial  $\beta$ -signal motif of transmembrane OMPs depends considerably on the nature of the amino acid side chain. Overall, the interface partitioning energetics that we have determined for PagP will instigate further measurements of a whole-protein interface scale for transmembrane proteins in near-native environments.

## Experimental procedures

### Molecular dynamics simulations

All molecular dynamics simulations were performed at a constant temperature of 310 K and a constant pressure of 1 bar, using GROMACS version 5.0.4 (74), using previously reported protocols (30). Using the CHARMM-GUI (75–77) input generator, we embedded the protein molecule (PagP crystal structure, PDB code 1THQ (26)) in a micelle system consisting of 80 DPC molecules. We first equilibrated the protein–micelle complex and followed it up with a production molecular dynamics simulation of 10 ns. Vicinity analysis was performed using VMD (Visual Molecular Dynamics) (78).

### Protein preparation

The wild-type *pagP* gene (accession number NC\_000913) was amplified from the *E. coli* K12 MG1655 genome using specific primers ([supplemental Table S1](#)) and cloned without the signal sequence into pET-3a vector between NdeI and BamHI sites (30). This was used as the template to generate a library of 60 mutants at the C-terminal residues 160 and 161, which

## Side-chain interface energetics in PagP folding

included Gln<sup>160</sup>-Xaa<sup>161</sup>, Xaa<sup>160</sup>-Phe<sup>161</sup>, and Xaa<sup>160</sup>-Leu<sup>161</sup> (where Xaa is any one of 20 naturally occurring amino acids). Further, the residue deletion mutants  $\Delta F^{161}$  and  $\Delta Q^{160}F^{161}$  were also generated. PagP-WT and all of the mutants were expressed without any affinity tags. *E. coli* BL21(DE3) cells were transformed with these plasmids and were used for producing the protein as inclusion bodies. The proteins were processed further to ~95% purity using reported methods (79).

### PagP folding in DPC

For the equilibrium folding and unfolding measurements of PagP, stock protein was generated using reported methods (29, 30). Briefly, the unfolded protein was dissolved in 20 mM Tris-HCl, pH 9.5, containing 8 M urea, at a concentration of ~300  $\mu$ M. This was diluted 10-fold into the folding reaction containing 100 mM DPC prepared in 20 mM Tris-HCl, pH 9.5, at 4 °C. The sample was heated at 70 °C for 3 min (80), immediately transferred to 4 °C, and incubated overnight. The next day, samples were centrifuged at  $16,600 \times g$  for 1 h to remove any trace amounts of protein aggregates. Each preparation was checked for soluble aggregates on a UV spectrophotometer by monitoring the scattering between 300 and 340 nm and electrophoretically using cold SDS-PAGE (66, 80, 81). The final stock contained ~30  $\mu$ M PagP in 100 mM DPC and 20 mM Tris-HCl, pH 9.5. This corresponds to a DPR of ~3300:1. The folding efficiency for this stock was quantified using densitometry analysis of the gel mobility shift (66, 80, 81) and proteolysis by proteinase K (30). The unfolded stock was prepared similarly, with the only difference that all the solutions contained 8 M GdnHCl.

### PagP folding in DLPC

PagP exhibits hysteresis in all of the conditions of protein-lipid ratios, buffer, pH, and temperature that we were able to screen. Hence, we studied only the folding equilibrium of PagP in SUVs of DLPC, prepared by sonication. For the folding measurements of PagP, unfolded protein (2.6 mM PagP in 7 M GdnHCl prepared in Tris-HCl, pH 9.5) was added in a 1:9 ratio to 89 mM DLPC SUVs. This mixture was further 10-fold diluted in 7 M GdnHCl to yield the folding stock containing 26  $\mu$ M protein and 8 mM lipid. This stock was 10-fold diluted into varying GdnHCl concentrations ranging from 0.7 to 7 M, and

these final reactions were used for equilibrium fluorescence measurements.

### Equilibrium (un)folding experiments using steady-state fluorescence

The PagP samples for equilibrium folding and unfolding in DPC were prepared by 10-fold dilution of the respective protein stocks in various GdnHCl concentrations ranging from ~0.7 to ~6.6 M, at an increment of 0.1 M. This gave us a final protein concentration of 3  $\mu$ M and DPC of 10 mM in each reaction. Samples were incubated at 25 °C, and fluorescence measurements were acquired on a microplate reader at the same temperature. We monitored the progress of the reactions using the decrease in tryptophan fluorescence emission intensity, with increase in GdnHCl concentration. Spectra were acquired using a  $\lambda_{ex}$  of 295 nm and  $\lambda_{em}$  of 320–400 nm. For PagP and its mutants, equilibrium for the reaction was achieved within 24 h.

From the fluorescence profiles, we calculated the unfolded fraction ( $f_U$ ) for the 48-h data using the following equation.

$$f_U = \frac{y_O - (y_F + m_F[D])}{(y_U + m_U[D]) - (y_F + m_F[D])} \quad (\text{Eq. 1})$$

Here,  $y_O$  is the observed fluorescence at GdnHCl concentration [D], whereas  $y_F$ ,  $m_F$ ,  $y_U$ , and  $m_U$  are intercepts and slopes of the pre- and post-transition baselines, respectively.

We were able to explain the folding transitions for most of the mutants using the two-state equation (42).

$$f_U = \frac{\exp(-(\Delta G + m[D])/RT)}{1 + \exp[-(\Delta G + m[D])/RT]} \quad (\text{Eq. 2})$$

This equation assumes that the protein folds in a cooperative manner from the unfolded (U) to the folded (F) state, without a detectable folding intermediate. We obtained the thermodynamic parameters  $\Delta G^0$  ( $\Delta G_F^{0,H_2O}$ , folding free energy) and  $m$  value (change in ASA between U and F states) of folding from the fits. The midpoint of chemical denaturation ( $C_m$ ) was calculated as  $C_m = \Delta G/m$ .

The folding transition of some mutants could only be explained using a three-state equation (Equation 3), due to the occurrence of an intermediate (I) (7).

$$f_U = \frac{(y_F + m_F[D]) + \left( \exp\left(-\frac{\Delta G_1 + m_1[D]}{RT}\right)(y_I + m_I[D]) \right) + \left( \exp\left(-\frac{\Delta G_1 + m_1[D]}{RT}\right) \exp\left(-\frac{\Delta G_2 + m_2[D]}{RT}\right)(y_U + m_U[D]) \right)}{1 + \exp\left(-\frac{\Delta G_1 + m_1[D]}{RT}\right) + \exp\left(-\frac{\Delta G_1 + m_1[D]}{RT}\right) \exp\left(-\frac{\Delta G_2 + m_2[D]}{RT}\right)}$$

(Eq. 3)

Here, we obtained  $\Delta G_1$  and  $\Delta G_2$  and their corresponding  $m_1$  and  $m_2$  values for the change in free energy from the first (U  $\rightarrow$  I) and second (I  $\rightarrow$  N) transitions, respectively. The PagP samples for folding in DLPC were prepared as described above. We used the same parameters for data acquisition

in DLPC samples as for DPC. From the fluorescence profiles, we calculated the unfolded fraction ( $f_U$ ) for the 48-h data using Equation 1. The data were fitted globally to Equation 2, assuming a common  $m$  value. Fits of the data to Equation 2 yielded the apparent thermodynamic parameters, namely the

apparent change in free energy  $\Delta G_{\text{app}}$ , apparent ASA change  $m_{\text{app}}$ , and  $C_m$  for each mutant.

### Fluorescence anisotropy, lifetime, and average emission wavelength measurements

For PagP and its mutants, equilibrium for the chemical denaturation was achieved within 24 h. Tryptophan fluorescence anisotropy and lifetimes were measured 24–48 h post-equilibrium. We used a  $\lambda_{\text{ex}}$  of 295 nm. Anisotropy was monitored at 344 nm ( $\lambda_{\text{em-max}}$  of folded PagP) and lifetime at 340 nm. The instrument parameters reported previously (30) were followed here. We fitted the decay curves for lifetime measurements to a triple exponential decay function and calculated the average lifetime as  $\langle\tau\rangle = \sum(\tau_i \alpha_i)$ .

Average emission wavelength ( $\langle\lambda\rangle$ ) was calculated from the emission spectra using Equation 4, as follows (82).

$$\lambda = \frac{\sum(I(\lambda_i)\lambda_i)}{\sum(I(\lambda_i))} \quad (\text{Eq. 4})$$

Here,  $\lambda_i$  is the  $i$ th emission wavelength, and  $I(\lambda_i)$  is the emission intensity at that wavelength. The data were further corrected using a normalization factor ( $Q_R$ ) so that the  $\langle\lambda\rangle$  is now linearly proportional to the unfolded protein fraction in solution (82).

### Enzymatic assay

We monitored the activity of PagP refolded in DPC micelles by measuring the release of *p*-nitrophenol through the hydrolysis of *p*NPP (30). Briefly, we added 0.04  $\mu\text{g}/\mu\text{l}$  (2  $\mu\text{M}$ ) refolded protein to 1 mM of the substrate analogue *p*NPP prepared in 2% Triton X-100. We monitored the release of *p*-nitrophenol spectrophotometrically at 405 nm. The rate of the reaction was monitored for 60 min, following which we used the linear region of the kinetics curve for deriving PagP activity (30).

### Circular dichroism measurements

We recorded the far-UV CD wavelength scans of 0.3  $\mu\text{g}/\mu\text{l}$  protein in 50 mM DPC using a 1-mm quartz cuvette, between 205 and 260 nm at a temperature of 5 °C, as described previously (30). Data were averaged over three accumulations, corrected for buffer contribution, smoothed, and represented as molar ellipticity (degrees  $\text{cm}^2 \text{dmol}^{-1}$ ) values calculated using reported methods (80).

**Author contributions**—R. M. conceived, designed, and coordinated the study and wrote the paper. B. R. I. performed the experiments shown in Figs. 1, 2, 3, 5, and 7 and contributed to writing the paper. P. Z. performed the experiments shown in Figs. 3–7 and contributed to writing the paper. P. V. V. performed the experiments shown in Figs. 2 and 4. All authors analyzed the results and approved the final version of the manuscript.

**Acknowledgments**—We thank Prof. Stephen White (University of California, Irvine, CA), Prof. James U. Bowie (University of California, Los Angeles, CA), and Prof. William F. DeGrado (University of California, San Francisco, CA) for invaluable comments and insights, which have helped to considerably improve our work.

### References

- Rassam, P., Copeland, N. A., Birkholz, O., Tóth, C., Chavent, M., Duncan, A. L., Cross, S. J., Housden, N. G., Kaminska, R., Seger, U., Quinn, D. M.,

- Garrod, T. J., Sansom, M. S., Piehler, J., Baumann, C. G., and Kleanthous, C. (2015) Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria. *Nature* **523**, 333–336
- Shiota, T., Imai, K., Qiu, J., Hewitt, V. L., Tan, K., Shen, H. H., Sakiyama, N., Fukasawa, Y., Hayat, S., Kamiya, M., Elofsson, A., Tomii, K., Horton, P., Wiedemann, N., Pfanner, N., *et al.* (2015) Molecular architecture of the active mitochondrial protein gate. *Science* **349**, 1544–1548
- Bakelar, J., Buchanan, S. K., and Noinaj, N. (2016) The structure of the  $\beta$ -barrel assembly machinery complex. *Science* **351**, 180–186
- Han, L., Zheng, J., Wang, Y., Yang, X., Liu, Y., Sun, C., Cao, B., Zhou, H., Ni, D., Lou, J., Zhao, Y., and Huang, Y. (2016) Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. *Nat. Struct. Mol. Biol.* **23**, 192–196
- Fairman, J. W., Noinaj, N., and Buchanan, S. K. (2011) The structural biology of beta-barrel membrane proteins: a summary of recent reports. *Curr. Opin. Struct. Biol.* **21**, 523–531
- Moon, C. P., Zaccai, N. R., Fleming, P. J., Gessmann, D., and Fleming, K. G. (2013) Membrane protein thermodynamic stability may serve as the energy sink for sorting in the periplasm. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 4285–4290
- Moon, C. P., and Fleming, K. G. (2011) Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10174–10177
- Lin, M., Gessmann, D., Naveed, H., and Liang, J. (2016) Outer membrane protein folding and topology from a computational transfer free energy scale. *J. Am. Chem. Soc.* **138**, 2592–2601
- Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S. H., and von Heijne, G. (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* **433**, 377–381
- Hessa, T., Meindl-Beinker, N. M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M., Nilsson, I., White, S. H., and von Heijne, G. (2007) Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature* **450**, 1026–1030
- Elazar, A., Weinstein, J., Biran, I., Fridman, Y., Bibi, E., and Fleishman, S. J. (2016) Mutational scanning reveals the determinants of protein insertion and association energetics in the plasma membrane. *Elife* **5**, e12125
- Wimley, W. C., and White, S. H. (1996) Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Biol.* **3**, 842–848
- Senes, A., Chadi, D. C., Law, P. B., Walters, R. F., Nanda, V., and Degrad, W. F. (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
- Killian, J. A., and von Heijne, G. (2000) How proteins adapt to a membrane-water interface. *Trends Biochem. Sci.* **25**, 429–434
- Chandler, D. (2005) Interfaces and the driving force of hydrophobic assembly. *Nature* **437**, 640–647
- de Planque, M. R., and Killian, J. A. (2003) Protein-lipid interactions studied with designed transmembrane peptides: role of hydrophobic matching and interfacial anchoring. *Mol. Membr. Biol.* **20**, 271–284
- Granseth, E., von Heijne, G., and Elofsson, A. (2005) A study of the membrane-water interface region of membrane proteins. *J. Mol. Biol.* **346**, 377–385
- MacCallum, J. L., and Tieleman, D. P. (2011) Hydrophobicity scales: a thermodynamic looking glass into lipid-protein interactions. *Trends Biochem. Sci.* **36**, 653–662
- Öjemalm, K., Higuchi, T., Lara, P., Lindahl, E., Suga, H., and von Heijne, G. (2016) Energetics of side-chain snorkeling in transmembrane helices probed by nonproteinogenic amino acids. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10559–10564
- Miller, A. S., and Falke, J. J. (2004) Side chains at the membrane-water interface modulate the signaling state of a transmembrane receptor. *Biochemistry* **43**, 1763–1770
- Sun, H., Greathouse, D. V., Andersen, O. S., and Koeppe, R. E., 2nd (2008) The preference of tryptophan for membrane interfaces: insights

## Side-chain interface energetics in PagP folding

- from *N*-methylation of tryptophans in gramicidin channels. *J. Biol. Chem.* **283**, 22233–22243
22. Robert, V., Volokhina, E. B., Senf, F., Bos, M. P., Van Gelder, P., and Tommassen, J. (2006) Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol.* **4**, e377
23. Walther, D. M., Bos, M. P., Rapaport, D., and Tommassen, J. (2010) The mitochondrial porin, VDAC, has retained the ability to be assembled in the bacterial outer membrane. *Mol. Biol. Evol.* **27**, 887–895
24. Schleiff, E., and Becker, T. (2011) Common ground for protein translocation: access control for mitochondria and chloroplasts. *Nat. Rev. Mol. Cell Biol.* **12**, 48–59
25. Hwang, P. M., Choy, W. Y., Lo, E. I., Chen, L., Forman-Kay, J. D., Raetz, C. R., Privé, G. G., Bishop, R. E., and Kay, L. E. (2002) Solution structure and dynamics of the outer membrane enzyme PagP by NMR. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13560–13565
26. Ahn, V. E., Lo, E. I., Engel, C. K., Chen, L., Hwang, P. M., Kay, L. E., Bishop, R. E., and Privé, G. G. (2004) A hydrocarbon ruler measures palmitate in the enzymatic acylation of endotoxin. *EMBO J.* **23**, 2931–2941
27. Cuesta-Seijo, J. A., Neale, C., Khan, M. A., Mokhtar, J., Tran, C. D., Bishop, R. E., Pomès, R., and Privé, G. G. (2010) PagP crystallized from SDS/cosolvent reveals the route for phospholipid access to the hydrocarbon ruler. *Structure* **18**, 1210–1219
28. Evancsics, F., Hwang, P. M., Cheng, Y., Kay, L. E., and Prosser, R. S. (2006) Topology of an outer-membrane enzyme: measuring oxygen and water contacts in solution NMR studies of PagP. *J. Am. Chem. Soc.* **128**, 8256–8264
29. Huysmans, G. H., Baldwin, S. A., Brockwell, D. J., and Radford, S. E. (2010) The transition state for folding of an outer membrane protein. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 4099–4104
30. Iyer, B. R., and Mahalakshmi, R. (2015) Residue-dependent thermodynamic cost and barrel plasticity balances activity in the PhoPQ-activated enzyme PagP of *Salmonella typhimurium*. *Biochemistry* **54**, 5712–5722
31. McDonald, S. K., and Fleming, K. G. (2016) Negative charge neutralization in the loops and turns of outer membrane phospholipase A impacts folding hysteresis at neutral pH. *Biochemistry* **55**, 6133–6137
32. McMorran, L. M., Bartlett, A. I., Huysmans, G. H., Radford, S. E., and Brockwell, D. J. (2013) Dissecting the effects of periplasmic chaperones on the *in vitro* folding of the outer membrane protein PagP. *J. Mol. Biol.* **425**, 3178–3191
33. Hiller, S., Abramson, J., Mannella, C., Wagner, G., and Zeth, K. (2010) The 3D structures of VDAC represent a native conformation. *Trends Biochem. Sci.* **35**, 514–521
34. Nolandt, O. V., Walther, T. H., Grage, S. L., and Ulrich, A. S. (2012) Magnetically oriented dodecylphosphocholine bicelles for solid-state NMR structure analysis. *Biochim. Biophys. Acta* **1818**, 1142–1147
35. Popot, J. L., and Engelman, D. M. (2016) Membranes do not tell proteins how to fold. *Biochemistry* **55**, 5–18
36. Frey, L., Lakomek, N. A., Riek, R., and Bibow, S. (2017) Micelles, bicelles, and nanodiscs: comparing the impact of membrane mimetics on membrane protein backbone dynamics. *Angew. Chem. Int. Ed. Engl.* **56**, 380–383
37. Hong, H., Park, S., Jiménez, R. H., Rinehart, D., and Tamm, L. K. (2007) Role of aromatic side chains in the folding and thermodynamic stability of integral membrane proteins. *J. Am. Chem. Soc.* **129**, 8320–8327
38. Hong, H., Rinehart, D., and Tamm, L. K. (2013) Membrane depth-dependent energetic contribution of the tryptophan side chain to the stability of integral membrane proteins. *Biochemistry* **52**, 4413–4421
39. Domene, C., Vemparala, S., Klein, M. L., Vénien-Bryan, C., and Doyle, D. A. (2006) Role of aromatic localization in the gating process of a potassium channel. *Biophys. J.* **90**, L01–L03
40. Strandberg, E., and Killian, J. A. (2003) Snorkeling of lysine side chains in transmembrane helices: how easy can it get? *FEBS Lett.* **544**, 69–73
41. Kim, C., Schmidt, T., Cho, E. G., Ye, F., Ulmer, T. S., and Ginsberg, M. H. (2011) Basic amino-acid side chains regulate transmembrane integrin signalling. *Nature* **481**, 209–213
42. Santoro, M. M., and Bolen, D. W. (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl  $\alpha$ -chymotrypsin using different denaturants. *Biochemistry* **27**, 8063–8068
43. Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H., and Zehfus, M. H. (1985) Hydrophobicity of amino acid residues in globular proteins. *Science* **229**, 834–838
44. Miller, S., Janin, J., Lesk, A. M., and Chothia, C. (1987) Interior and surface of monomeric proteins. *J. Mol. Biol.* **196**, 641–656
45. Wimley, W. C., Creamer, T. P., and White, S. H. (1996) Solvation energies of amino acid side chains and backbone in a family of host-guest pentapeptides. *Biochemistry* **35**, 5109–5124
46. Nozaki, Y., and Tanford, C. (1971) The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions: establishment of a hydrophobicity scale. *J. Biol. Chem.* **246**, 2211–2217
47. Kyte, J., and Doolittle, R. F. (1982) A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **157**, 105–132
48. Eisenberg, D., Weiss, R. M., and Terwilliger, T. C. (1984) The hydrophobic moment detects periodicity in protein hydrophobicity. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 140–144
49. Engelman, D. M., Steitz, T. A., and Goldman, A. (1986) Identifying non-polar transbilayer helices in amino acid sequences of membrane proteins. *Annu. Rev. Biophys. Biophys. Chem.* **15**, 321–353
50. Wolfenden, R., Andersson, L., Cullis, P. M., and Southgate, C. C. (1981) Affinities of amino acid side chains for solvent water. *Biochemistry* **20**, 849–855
51. Rose, G. D. (1978) Prediction of chain turns in globular proteins on a hydrophobic basis. *Nature* **272**, 586–590
52. Janin, J. (1979) Surface and inside volumes in globular proteins. *Nature* **277**, 491–492
53. Schmidt, C. F., Lichtenberg, D., and Thompson, T. E. (1981) Vesicle-vesicle interactions in sonicated dispersions of dipalmitoylphosphatidylcholine. *Biochemistry* **20**, 4792–4797
54. Pocsan, C. L., Popot, J. L., and Kleinschmidt, J. H. (2013) Folding and stability of outer membrane protein A (OmpA) from *Escherichia coli* in an amphipathic polymer, amphipol A8-35. *Eur. Biophys. J.* **42**, 103–118
55. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–2148
56. Wrabl, J., and Shortle, D. (1999) A model of the changes in denatured state structure underlying *m* value effects in staphylococcal nuclease. *Nat. Struct. Biol.* **6**, 876–883
57. Khan, M. A., Neale, C., Michaux, C., Pomès, R., Privé, G. G., Woody, R. W., and Bishop, R. E. (2007) Gauging a hydrocarbon ruler by an intrinsic exciton probe. *Biochemistry* **46**, 4565–4579
58. Hong, H., and Tamm, L. K. (2004) Elastic coupling of integral membrane protein stability to lipid bilayer forces. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 4065–4070
59. Struyvé, M., Moons, M., and Tommassen, J. (1991) Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. *J. Mol. Biol.* **218**, 141–148
60. Joh, N. H., Min, A., Faham, S., Whitelegge, J. P., Yang, D., Woods, V. L., and Bowie, J. U. (2008) Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. *Nature* **453**, 1266–1270
61. Chothia, C. (1976) The nature of the accessible and buried surfaces in proteins. *J. Mol. Biol.* **105**, 1–12
62. Öjemalm, K., Higuchi, T., Jiang, Y., Langel, Ü., Nilsson, I., White, S. H., Suga, H., and von Heijne, G. (2011) Apolar surface area determines the efficiency of translocon-mediated membrane-protein integration into the endoplasmic reticulum. *Proc. Natl. Acad. Sci. U.S.A.* **108**, E359–E364
63. Hagan, C. L., Silhavy, T. J., and Kahne, D. (2011)  $\beta$ -Barrel membrane protein assembly by the Bam complex. *Annu. Rev. Biochem.* **80**, 189–210
64. Noinaj, N., Kuszak, A. J., Gumbart, J. C., Lukacik, P., Chang, H., Easley, N. C., Lithgow, T., and Buchanan, S. K. (2013) Structural insight into the biogenesis of  $\beta$ -barrel membrane proteins. *Nature* **501**, 385–390
65. Leyton, D. L., Belousoff, M. J., and Lithgow, T. (2015) The  $\beta$ -barrel assembly machinery complex. *Methods Mol. Biol.* **1329**, 1–16
66. Gessmann, D., Chung, Y. H., Danoff, E. J., Plummer, A. M., Sandlin, C. W., Zaccai, N. R., and Fleming, K. G. (2014) Outer membrane  $\beta$ -barrel protein

- folding is physically controlled by periplasmic lipid head groups and BamA. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 5878–5883
67. Plummer, A. M., and Fleming, K. G. (2016) From chaperones to the membrane with a BAM! *Trends Biochem. Sci.* **41**, 872–882
  68. Fleming, K. G. (2014) Energetics of membrane protein folding. *Annu. Rev. Biophys.* **43**, 233–255
  69. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science* **181**, 223–230
  70. Jiang, Y., Lee, A., Chen, J., Ruta, V., Cadene, M., Chait, B. T., and MacKinnon, R. (2003) X-ray structure of a voltage-dependent K<sup>+</sup> channel. *Nature* **423**, 33–41
  71. Ruta, V., Chen, J., and MacKinnon, R. (2005) Calibrated measurement of gating-charge arginine displacement in the KvAP voltage-dependent K<sup>+</sup> channel. *Cell* **123**, 463–475
  72. Noinaj, N., Gumbart, J. C., and Buchanan, S. K. (2017) The  $\beta$ -barrel assembly machinery in motion. *Nat. Rev. Microbiol.* **15**, 197–204
  73. Fleming, K. G., Costello, S. M., Plummer, A. M., and Fleming, P. J. (2017) Periplasmic chaperones play hot potato with unfolded outer membrane proteins. *FASEB J.* **31**, 604.8
  74. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. (2005) GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718
  75. Jo, S., Kim, T., Iyer, V. G., and Im, W. (2008) CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865
  76. Brooks, B. R., Brooks, C. L., 3rd, Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., *et al.* (2009) CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614
  77. Cheng, X., Jo, S., Lee, H. S., Klauda, J. B., and Im, W. (2013) CHARMM-GUI micelle builder for pure/mixed micelle and protein/micelle complex systems. *J. Chem. Inf. Model.* **53**, 2171–2180
  78. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38, 27–28
  79. Gupta, A., Iyer, B. R., Chaturvedi, D., Maurya, S. R., and Mahalakshmi, R. (2015) Thermodynamic, structural and functional properties of membrane protein inclusion bodies are analogous to purified counterparts: case study from bacteria and humans. *RSC Adv.* **5**, 1227–1234
  80. Maurya, S. R., Chaturvedi, D., and Mahalakshmi, R. (2013) Modulating lipid dynamics and membrane fluidity to drive rapid folding of a trans-membrane barrel. *Sci. Rep.* **3**, 1989
  81. Burgess, N. K., Dao, T. P., Stanley, A. M., and Fleming, K. G. (2008)  $\beta$ -Barrel proteins that reside in the *Escherichia coli* outer membrane *in vivo* demonstrate varied folding behavior *in vitro*. *J. Biol. Chem.* **283**, 26748–26758
  82. Moon, C. P., and Fleming, K. G. (2011) Using tryptophan fluorescence to measure the stability of membrane proteins folded in liposomes. *Methods Enzymol.* **492**, 189–211