Published in final edited form as:

Biochemistry. 2017 January 10; 56(1): 47–60. doi:10.1021/acs.biochem.6b00809.

Novel kinetic intermediates populated along the folding pathway of the transmembrane β -barrel OmpA

Emily J. Danoff and Karen G. Fleming*

T. C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, Maryland 21218, United States

Abstract

We examined the folding of the β -barrel membrane protein OmpA from *E. coli*. Although previous studies identified several intermediate states followed by a concerted translocation mechanism across the bilayer, some aspects of the pathway were still unclear including the extent of secondary structure formation in the intermediate states and how the mechanism gave rise to multiple exponential phases in the folding kinetics. We addressed these questions by investigating the folding kinetics of the OmpA transmembrane β -barrel domain into a range of bilayer thicknesses, allowing us to observe different regions of the folding pathway. The fastest folding into the thinnest bilayers provided information on the later stages of the process, and the slowest folding into thicker bilayers revealed early kinetic steps. Folding was monitored using SDS-PAGE and CD spectroscopy, which provide complementary information about tertiary and secondary structure formation. We globally fit the folding data to kinetic schemes and found that the same core pathway was followed in all lipid conditions. We propose a multi-step folding mechanism for OmpA that includes unstructured surface-adsorbed states converting through a partially inserted state with substantial β -sheet structure to the final natively inserted barrel. Kinetic models show that all steps of the main folding pathway are accelerated by membrane defects that occur as a result of thinning the bilayer or incubation of lipids at the phase transition temperature. In addition to suppressing off-pathway states, BAM-catalyzed folding *in vivo* could accelerate any or all of these main folding steps to ensure efficient OMP biogenesis *in vivo*.

Graphical abstract

karen.fleming@jhu.edu. Phone: (410) 516-7256.

Supporting Information Available

Additional methods, discussion of results, development of kinetic models, tables, figures, and fitting equations. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Membrane proteins serve a variety of essential functions in biological systems, such as transport of metabolites into and out of cells, transduction of signals, and enzymatic reactions. Their importance is demonstrated by the fact that 20–30% of all genes in many organisms' genomes encode membrane proteins,¹ and more than 50% of medical drugs currently on the market target membrane proteins.² In addition, many diseases have been linked to mutations that result in their misfolding or incorrect targeting.³ Therefore, understanding the process by which membrane proteins fold to the correct tertiary structure and insert into their native bilayer environment is of utmost importance.

The need to satisfy all backbone hydrogen bonds within the water-excluding bilayer environment has restricted transmembrane proteins to two types of structures, α -helical and β -barrel. Whereas α -helical membrane proteins are ubiquitous in most biological membranes, β -barrel membrane proteins are found only in the outer membranes of Gram-negative bacteria,⁴ mitochondria,⁵ and chloroplasts.⁶ These outer membrane proteins (OMPs) often serve as porins for small molecules but can have many other functions such as membrane biogenesis, bacterial virulence, and protein import into mitochondria.^{7–11}

With their many functions and connections to disease, it is vital that we understand the folding pathway of β -barrel membrane proteins. In bacteria, OMPs are known to interact with a number of periplasmic chaperones en route to the outer membrane,¹² and assembly into the membrane is catalyzed by the β -barrel assembly machinery, or BAM complex,¹³ but the details of how these proteins facilitate OMP folding are unclear. OMPs are also capable of spontaneously folding to the native conformation in lipid bilayers in the absence of any folding chaperones,¹⁴ but on timescales too slow to be compatible with bacterial growth rates,^{15,16} especially in the presence of lipids native to the bacterial outer membrane.¹⁷ Therefore it appears that the BAM complex does not provide conformational instructions to substrate OMPs, but rather lowers the kinetic barrier to intrinsic OMP folding, possibly by inducing defects in the bilayer that OMPs utilize to insert and fold.^{17,18}

Progress has been made in elucidating the intrinsic mechanism of OMP folding based on studies of the structural protein OmpA from *E. coli*. Monitoring OmpA folding into unilamellar vesicles using a variety of techniques has revealed that the protein folds via a multi-step mechanism with several membrane-associated intermediate states.^{15,16} In addition, it has been shown that β -barrel formation occurs in concert with membrane insertion; i.e., the β -hairpins translocate the bilayer at the same time.¹⁹ However, there are several major aspects of the OMP folding mechanism that had not been determined. Most notably, the conformations of the membrane-bound folding intermediates are unclear, particularly the extent of β -sheet formation. It is also uncertain whether parallel pathways occur in the OMP folding mechanism, as has been proposed for the OMPs FomA and PagP based on observations of multi-exponential kinetics.^{20,21} This can be indicative of independent unrelated reaction pathways. Alternatively, the presence of off-pathway intermediate states (i.e., optional misfolding) can also account for such behavior.²²

To address these questions we investigated the folding kinetics of OmpA into bilayers of different thicknesses and monitored the reaction using SDS-PAGE (to detect tertiary structure) or CD spectroscopy (to detect secondary structure). We conducted our

measurements on OmpA₁₇₁, which comprises the N-terminal β -barrel domain of OmpA.²³ This strategy eliminates contributions to the CD signal from the globular periplasmic domain. Additionally, through our previous studies of the oligomerization of the aqueous unfolded state,²³ we determined conditions under which the unfolded barrel domain does not self-associate. This was a key technical advantage that enabled us to examine the folding reaction without the influence of the competing off-pathway oligomerization reaction of the aqueous state — a condition that more closely resembles the *in vivo* situation where periplasmic chaperones presumably suppress self-association prior to membrane insertion.²⁴

Our SDS-PAGE measurements led us to observe a novel conformation for OmpA that migrates anomalously by gel, which has not been previously discussed in the literature. We include this state in comprehensive kinetic modeling of the folding pathway based on global fitting of time-dependent SDS-PAGE and CD data. Through the modeling process we identified three membrane-associated intermediate states through which the protein sequentially progresses on the way to the native state, and we found that the third intermediate state possesses substantial β -sheet structure. We also determined that the temporal dependence of structure formation is consistent with the population of several off-pathway, presumably misfolded, intermediate states. The presence of parallel pathways is not required. This result is in agreement with the protein folding theory of predetermined pathways with optional errors (PPOE) proposed previously.²² We conclude by interpreting our kinetic model from a structural perspective. Building upon the previous model for OmpA folding,⁸ our data and kinetic modeling enable a more comprehensive description of the multi-step β -barrel formation process and the driving forces for membrane insertion.

Materials and Methods

Preparation of OmpA₁₇₁

The cloning and expression of the transmembrane barrel domain of OmpA (OmpA₁₇₁) has been described previously.²³ The protein was purified in Urea Buffer (8 M urea, 20 mM Tris, pH 8) as described.^{18,23} The protein stock concentration was determined by absorbance at 280 nm; the extinction coefficient was calculated using the values for model amino acid chromophores in 8 M urea measured by Pace and co-workers,²⁵ and was determined to be 45,075 M⁻¹ cm⁻¹. The final protein concentration was typically ~100 μ M. Purified protein was divided into 100 μ l aliquots and stored at -80 °C until use.

Vesicle preparation

Phosphatidylcholine lipids with a range of chain lengths (diC₉PC – diC₁₄PC) were purchased from Avanti Polar Lipids dissolved in chloroform. Lipids were dried to a thin film in glass vials (5 mg per vial) under a gentle stream of nitrogen gas. The lipid films were evacuated overnight to remove residual solvent and stored at -20 °C until use. For vesicle preparation, lipid films were reconstituted in 1 M urea, 20 mM Tris, pH 8 at a concentration of 10 mg ml⁻¹ and incubated at room temperature for at least 30 min, with occasional vortexing. Large unilamellar vesicles (LUVs) were prepared by extruding reconstituted lipids 35 times through a 0.1 μ m filter using a mini-extruder (Avanti). For diC₁₄PC, lipids

were incubated and extruded at 35 °C in order to be well above the phase transition temperature at 24 °C.^{18,26}

Folding kinetics measured by SDS-PAGE

In all experiments, OmpA₁₇₁ was folded into pre-formed LUVs in the standard buffer condition of 1 M urea, 20 mM Tris, pH 8 (Folding Buffer). The final protein concentration was 1 μM and the final lipid concentration was 800 μM (800:1 lipid to protein ratio), in a final volume of 2.2 ml. To measure the time-dependence of folding using SDS-PAGE, OmpA₁₇₁ was rapidly diluted from Urea Buffer into Folding Buffer (8 M urea to 1 M urea) with constant stirring, and then LUVs were added to initiate folding. Folding samples were incubated in a custom-built stirring incubator (Aviv Biomedical) at 25 °C with continual stirring for the duration of the experiment. For folding into diC₁₄PC, measurements were also conducted at 24 °C and 26 °C. Aliquots were removed at specific time points after initiation (typically 5, 12, 30, 90, 180, 360, 720, 1800, 3600, 7200, 10800, 14400, and 18000 s) and folding quenched by mixing with 4X SDS gel-loading buffer to a final concentration of 1X. A duplicate sample was taken at the last time point and boiled at 100 °C in SDS loading buffer for 5 min. All other time points remained at room temperature after mixing with SDS loading buffer. Folding samples were maintained in the stirring incubator for 24 h and then transferred to microcentrifuge tubes at room temperature and long-term time points collected over the course of 7 days.

The extent of folding at each time point was determined by subjecting the samples to SDS-PAGE at 4 °C using pre-cast 12% acrylamide gels (Mini-PROTEAN TGX, BioRad) and staining with Coomassie Blue R-350 (GE Healthcare). All samples were analyzed by SDS-PAGE the same day they were collected. Gels were scanned at 1200 dpi using an Epson 4490 scanner in positive film mode. Densitometry was performed using ImageJ software.

Calculation of fraction folded, fraction unfolded, and fraction elusive

Densitometry was used to determine the intensities of the “folded” and “unfolded” bands at each time point. Fraction folded (f_F) and fraction unfolded (f_U) were calculated as the intensity of the folded (F) or unfolded (U) band divided by the intensity of the unfolded band for the boiled sample (B). We also observed that the protein populates an additional state that migrates with neither the folded nor the unfolded bands, as evidenced by an observed decrease in the sum of the folded and unfolded band intensities over time. We termed this conformation the “elusive” state and quantified the fraction elusive (f_E) as the difference between the boiled band and the sum of the folded and unfolded bands, divided by the boiled band intensity:

$$f_E = \frac{B - (F + U)}{B} \quad (1)$$

For all folding conditions, reported fractional data are the averaged values from at least three independent folding reactions. Reported errors are the standard deviation of the averaged values.

Circular dichroism

Circular dichroism (CD) measurements were conducted using an Aviv Circular Dichroism Spectrometer, Model 410 (Aviv Biomedical), with a custom inset detector to reduce the effects of light scattering. Hellma cuvettes with a path length of 1 cm were used. Samples were equilibrated at 25 °C, unless noted otherwise, with gentle stirring for 10–15 min before each set of measurements.

CD wavelength spectra

Wavelength spectra were recorded between 205 and 280 nm in 1 nm increments using an averaging time of 5 s. For each sample, 3–5 scans were recorded and averaged. When the wavelength spectrum needed to be measured as quickly as possible to capture the initial conformation, the data were collected in 10 nm sections. Fresh folding mixtures were prepared for each wavelength section and the signal recorded in 1 nm increments with a 2 s averaging time. Data were averaged for 4 or 5 separate folding mixtures for each wavelength section, and the final data combined to produce the full spectrum.

Spectra of cuvettes containing only urea/buffer or LUVs/urea/buffer were subtracted from sample spectra to correct for background signal. The presence of 1 M urea (which absorbs light) and LUVs (which scatter light) combined with the low protein concentration (necessary to avoid self-association) resulted in a low signal-to-noise ratio at shorter wavelengths. For this reason, reliable wavelength data were collected only above 212 nm. Even so, this wavelength range was sufficient for assessing the secondary structure and monitoring time-dependent CD signals at the two wavelengths of interest, 216 and 230 nm. LUVs contribute to the measured CD signal, most likely due to differential scattering of left-handed and right-handed circularly polarized light.²⁷ Figure S6 of the Supporting Information shows wavelength spectra for LUVs of diC₉PC – diC₁₄PC and it is evident that the lipids exhibit a peak around 215 nm. It has been demonstrated that this apparent ellipticity arising from light scattering of lipid vesicles can be subtracted from the sample signal as background.^{28,29} We verified this by conducting measurements with unfolded protein and LUVs in separate, side-by-side 1 mm cuvettes, and determined that the signals from the two species are additive (data not shown). Therefore we corrected all of the experimental data for the lipid contribution by subtracting the corresponding LUV spectrum before conversion to mean residue ellipticity.

Determination of mean residue ellipticity for unfolded OmpA₁₇₁

The CD spectrum for unfolded OmpA₁₇₁ was measured in 1 M urea, 20 mM Tris, pH 8, and converted to mean residue ellipticity using the following equation:

$$[\Theta] = \frac{\theta}{10c\ln} \quad (2)$$

where θ is the measured ellipticity in mdeg, c is the concentration in M, l is the path length in cm, and n is the number of residues (172, including the N-terminal methionine²³). The concentration was measured directly by absorbance at 280 nm, using an extinction

coefficient for OmpA₁₇₁ in 1 M urea determined by the Edelhoch method (see Supporting Information).^{30,31}

Time Evolution of CD Spectral Features

OmpA₁₇₁ folding kinetics were measured by CD under the same conditions as for SDS-PAGE: 1 μ M protein and 800 μ M lipid, in Folding Buffer (1 M urea, 20 mM Tris, pH 8). Samples were stirred continuously during measurements. The CD intensity was monitored at either 216 nm or 230 nm using an interval between data points of 10 s and a time constant of 1 s. For each folding reaction, a baseline signal was measured for 3 min of Folding Buffer alone before the addition of protein or LUVs. The protein was then added (and rapidly diluted from 8 M urea to 1 M urea) and this intensity was measured for 2 min to establish a baseline value. LUVs were subsequently added to initiate folding. The LUVs were prepared in 1 M urea ensuring there was no further urea dilution concurrent with the initiation of folding. The folding kinetics were monitored until completion. The reported time-dependent CD data at 216 nm are the average of measurements on 2 or 3 separate samples.

The average buffer-alone signal was used for background correction of the subsequent folding data and the LUV contribution to the folding signal was subtracted using a value measured for a separate LUV-only sample in Folding Buffer. The average protein-alone signal was used to calculate the exact protein concentration of the sample using the previously determined mean residue ellipticity for unfolded OmpA₁₇₁ at that wavelength and the following equation:

$$c = \frac{\theta}{10[\Theta]_{\text{ln}}} \quad (3)$$

The concentration was corrected for the subsequent dilution caused by the addition of LUVs and used to convert the folding data to mean residue ellipticity using Equation 2. The time-values for the kinetics data were also corrected by subtracting the time of LUV addition to obtain a curve with folding beginning at time zero. After the completion of each kinetics experiment, a wavelength spectrum was collected. The spectrum was LUV-corrected and converted to mean residue ellipticity using Equation 2.

Kinetic modeling

By jointly considering the gel and CD data, a kinetic mechanism was developed for each set of folding data corresponding to a particular lipid. This process is extensively described in the Supporting Information and will only be summarized here. Briefly, a kinetic model was developed through simulation of various kinetic schemes and manual adjustment of microscopic rate constants. Complex kinetic schemes were simulated using Euler's method of numerical integration.³² The overall strategy was to find the simplest mathematical model that globally described the three orthogonal sets of gel and CD data as evidenced by randomness of the residuals of the fit and minimization of the chi-squared. Error analysis was performed to obtain estimates for the robustness of each parameter.

Results

OmpA₁₇₁ populates a lipid-mediated “elusive” state that migrates anomalously by SDS-PAGE

OMPs have long been known to exhibit a “heat-modifiability” when undergoing SDS-PAGE.³³ In the absence of heating, native β -barrel membrane proteins remain folded in the presence of SDS and migrate at an apparent molecular weight different than the actual molecular weight. This shift in position is thought to be due to differences in SDS binding and compactness between the two conformations.^{34,35} The identity of the shifted band as the folded form of the protein has been corroborated by spectroscopic features, protease protection, and phage receptor activity.^{33,36} Upon boiling in SDS, the protein becomes unfolded and migrates according to its molecular weight. Under conditions where only part of the protein population is folded, SDS serves to quench the folding reaction and maintain the conformations of folded and unfolded species while undergoing electrophoresis. This results in the appearance of both bands on the gel. Densitometry of the folded and unfolded bands can be used to quantify the extent of folding and this technique has been utilized extensively to measure OMP folding kinetics and stability.^{15,16,20,37–40}

Upon careful quantitation of the gel bands in our folding experiments, it became apparent that the protein was populating an additional state that migrates with neither the canonical folded nor unfolded bands. Figure 1A shows the densitometry results for OmpA₁₇₁ folding into large unilamellar vesicles (LUVs) composed of the lipid diC₁₂PC. It is evident that the unfolded band disappears over time while the folded band appears at a higher apparent molecular weight on the gel. The densitometry plot also shows the sum of the folded and unfolded band intensities (F+U), which noticeably decreases over time. This result indicates that some fraction of the protein is missing from the folded and unfolded bands. To enable references to this population, we designated it as the “elusive” state. Boiling the sample after mixing with SDS causes the reappearance of all of the protein in the unfolded band, as can be observed in the final lane of the gel and the corresponding densitometry value in Figure 1A. Figure S1A of the Supporting Information demonstrates that the elusive state is not formed in the absence of LUVs.

Examination of the entire gel reveals that the elusive protein migrates anomalously by SDS-PAGE. Concomitant with the loss of intensity from the folded and unfolded bands, several distinct bands at higher molecular weights and a continuous smear of protein appear in the upper regions of the latter gel lanes. The higher molecular weight bands and smears can often be faint and difficult to reliably quantify, but densitometry values were obtained for OmpA₁₇₁ folding into LUVs of diC₁₃PC, a condition that exhibits substantial loss of intensity from the folded and unfolded bands (Figures S1B and S1C). The data demonstrate that the calculated amount of intensity lost from the folded and unfolded bands (relative to the boiled band intensity) agrees well with the summed intensities of the high molecular weight bands and smear at each time point. As expected, the smear and bands disappear from the boiled sample lane in conjunction with the full intensity reappearing in the unfolded band. Similarly, no high molecular weight species or smear are observed in the

absence of LUVs (Figure S1A), which is consistent with no loss of intensity from the unfolded band under this condition

Evidently, the elusive state comprises a range of species, including a few distinct molecular weights. Wang and colleagues previously reported that OmpA forms partially folded and unfolded dimers and trimers when folding in limiting amounts of detergent micelles or small lipid vesicles.⁴¹ Our results are consistent with these observations, but further experiments would be required to determine the oligomeric nature of our elusive species. Furthermore, other lipid conditions produce less of this state, which makes it challenging to directly quantify these high molecular weight species on gels. To overcome this challenge, we calculate the elusive population as the amount lost from the monomeric folded and unfolded bands, and we treat it as one observable in the gel analysis.

The discovery of the elusive state significantly impacts the fractions folded and unfolded derived from SDS-PAGE data. Typically, the fractions folded and unfolded have been calculated by dividing by “F+U” because this sum was expected to represent the total protein population in each lane. The existence of the elusive state demonstrates that “F+U” is not a reliable measure of total protein. In contrast, our data show that the intensity of the boiled band is a more accurate measure of the total protein and should be used in calculating fractional populations. This improved protocol has a major impact on the apparent time-dependencies for the formation of the folded and unfolded populations. Dividing the intensity values for folding into diC₁₂PC LUVs by either “F+U” or “B” results in substantially different curves, as shown in Figure 1B. Because the quantity “F+U” decreases over time, the fractions folded and unfolded calculated with this quantity are artificially inflated above the true value, with the fraction folded being more inflated because its raw values are higher. To further demonstrate this point, the data in Figure 1B were fitted to exponential expressions for comparison. The apparent rate constants were found to be moderately different while the fitted amplitudes (which define the final folding efficiency) were substantially different. For an actual kinetic mechanism, both the observed rate constants and the amplitudes of the exponential terms are functions of the microscopic rate constants, and these differences demonstrate that calculation of fractional species using “F+U” leads to incorrect kinetic parameters. For this reason we normalized all of our folding data by the boiled band intensity measured with each dataset.

To our knowledge, the elusive state has not been previously considered in the context of OMP folding in the literature, with the exception of Wang et. al's report of OmpA oligomerization during folding in micelles and vesicles.⁴¹ However, a decrease in “F+U” over time can be observed even “by eye” in the gels of many previous OMP folding studies, ^{20,37,42,43} indicating the presence of elusive state. Therefore any kinetic parameters obtained from analyzing data in these studies are likely incorrect, and conclusions regarding the OMP folding pathway and kinetic behavior in those studies should be revisited.

Folding kinetics and elusive state formation are strongly dependent on bilayer thickness

Bilayer thickness is known to have profound effects on apparent folding kinetics.^{37,40} We utilized this variable to probe different regions of the folding pathway. Because the elusive state has never previously been addressed in the literature, we also examined the effects of

bilayer thickness on elusive state population and kinetics. Figures 2A and S2 show the time-dependence of fraction folded for OmpA₁₇₁ folding into LUVs composed of diC₉PC – diC₁₃PC at 25 °C, which all form vesicles under these conditions (see Figure S3). Consistent with previous observations, folding is generally faster into thinner bilayers.^{37,40} The only exception is diC₉PC, which induces slightly slower kinetics than diC₁₀PC. We have no explanation for this observation at this time. Figures 2B and 2C show the unfolded and elusive species' populations over time, which also exhibit faster time-dependencies in thinner bilayers. It is thought that the acceleration of OMP folding in thinner bilayers is due to a greater incidence of defects in thinner membranes, and we recently corroborated this argument by demonstrating that folding is also accelerated near the defect-prone lipid phase transition region of diC₁₄PC bilayers.¹⁸ Kinetics data measured in diC₁₄PC are shown in Figure S4 (unfolded and elusive species kinetics were not previously reported); it can be observed that the folded, unfolded, and elusive species exhibit the fastest kinetics at the transition temperature of 24 °C, slightly slower kinetics at 25 °C, and the slowest kinetics farther away from the transition, at 26 °C.

For all lipid conditions, the time-dependent evolutions of the three gel species are highly complex and require multiple exponential phases to fit. In addition, a lag phase appears for folding into the longer chain lipids (diC₁₂PC, diC₁₃PC, and diC₁₄PC at 25 and 26 °C). Paradoxically, the time-dependence of elusive state formation is slower in longer chain lipids, yet the population of elusive state formed at long times is higher in longer chain lipids. The increased final fraction elusive observed with increasing lipid chain length is coupled to a decrease in final fraction folded, because, for all lipids, the fraction unfolded approaches zero at long times (long-term time points are shown in Figures S10, S13, and S15-S21). The lipid-dependence of the elusive state kinetics provides further evidence that it is a non-native lipid-induced conformation of the protein. The negative correlation with fraction folded implies that it is formed by an off-pathway reaction that competes with folding. Altogether, this kinetic complexity eliminates two-state folding for OmpA and implies a multi-step mechanism for the folding process.

CD spectra of folded OmpA₁₇₁ exhibit β -sheet structure and an aromatic exciton signal

We first examined the overall CD spectrum of the folded conformation for OmpA₁₇₁. Figure 3A demonstrates the differences between the spectra of the unfolded and folded proteins in diC₁₀PC LUVs. Two distinct changes are evident upon folding: a trough appears at 216 nm indicative of β -sheet structure, and a peak appears at 230 nm. The observed β -sheet intensity of $-5000 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$ is highly consistent with all previous CD data on OMPs.^{14,15,33,37,39,44–48} Although not as common, CD bands in the 230 nm-region of the spectrum have previously been observed in a number of soluble proteins and have been attributed to aromatic residues that interact to give rise to exciton couplets in the spectra.^{49–51} Similarly, the OMP PagP exhibits a peak at 232 nm, which has been attributed to the interaction of a specific pair of aromatic residues brought into close proximity across the interior of the folded β -barrel.⁴⁸ OmpA₁₇₁ contains a large number of aromatic residues, and we propose that a similar interaction occurs, producing a peak of $+700 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$ at 230 nm.

Figure 3B shows the wavelength spectra for OmpA₁₇₁ after folding into lipids with increasing acyl chain lengths. Under all conditions the trough at 216 nm overlays, indicating the same amount of β -sheet structure in all lipids. In contrast, the peak at 230 nm is reduced in intensity with increasing acyl chain length. As shown in Figures 2 and S2, we observed different final fractions folded by gel in different lipids, and we propose that these differences in folding efficiencies could account for the change in CD intensities at 230 nm if the exciton interaction is only formed in the native state. The reduced final amounts of fraction folded correspond to higher final amounts of elusive state as measured by gel, which suggests that the elusive state possesses the same amount of β -sheet structure as the native state in a way that is distinct from the native structure giving rise to the intensity at 230 nm.

Time-evolution of CD signals depends on bilayer thickness and reveals an intermediate state with high β -sheet content

The wavelength spectra for folded OmpA₁₇₁ indicate that the largest changes in ellipticity from the unfolded spectrum occur at 216 and 230 nm. We therefore monitored the folding kinetics at these two wavelengths for all lipids (Figures 4 and S7). Overall the time-evolutions of the CD intensities at both wavelengths follow the same trends observed by SDS-PAGE. In addition, there is more complexity in the kinetics observed by CD because the spectroscopic signal is a combination of the signals from all species present rather than a direct measure of the population of each species (see Supporting Information). Of particular note are the kinetics behavior for folding into the shortest chain lipids, diC₉PC and diC₁₀PC. Under these conditions, the data measured at both 216 and 230 nm begin at more negative ellipticities than the expected unfolded protein signal, shown as a dashed red line in Figure 4 (also see Figures S17 and S18). This decrease within the mixing time to a more negative ellipticity before a subsequent increase to the final value (i.e. loss in negative intensity) suggests the population of an intermediate species with a higher content of β -sheet structure than the final native state. In diC₁₁PC, the intensity decrease is slow enough to be observed at 216 nm, and the signal only decreases slightly below the final intensity value before increasing back to it, indicating that the intermediate is less populated under this condition. Similarly, at 230 nm the intensity decreases slightly below the unfolded signal before increasing to the final value (see Figure S16). In the longer chain lipids diC₁₂PC and diC₁₃PC, the measured data at 216 and 230 nm begin at the baseline unfolded value and decrease (at 216 nm) or increase (at 230 nm) to the final intensity with only a slight indication of the more negative signal in diC₁₂PC at 216 nm (see Figures S13 and S15). Although this could indicate that the β -structured intermediate is not an obligatory intermediate for folding in diC₁₂PC and diC₁₃PC, it is also possible that the protein passes through the intermediate but it is not populated enough to make a strong impact on the overall CD signal. In support of the latter, the β -structured intermediate was observed in diC₁₃PC with the lipids below the phase transition temperature (see below).

The data for diC₁₄PC provide further indication that the β -structured intermediate is populated in the longer chain lipids. The kinetic traces at 24 °C decrease to a more negative value before increasing to the final ellipticities at 216 and 230 nm (Figures S7 and S21).¹⁸ In contrast, the kinetic traces at 25 and 26 °C do not decrease below the native value at 216 nm and only reach a slightly more negative signal at 230 nm before increasing to the final signal

(Figures S7, S19, and S20). Such behavior is consistent with faster folding at 24 °C leading to a higher population of the β -structured intermediate and thus a more negative signal in the observed kinetics at early times.

Wavelength spectra verify population of a β -structured intermediate in diC₉PC bilayers at early times and in diC₁₃PC at 4 °C

To verify the secondary structure content of the hypothesized early intermediate, we measured its full wavelength spectrum. Because the signal reaches the most negative ellipticity in diC₉PC, we captured the spectrum in this lipid condition. Figure 5A shows the wavelength spectrum exhibited by OmpA₁₇₁ immediately after mixing with diC₉PC LUVs. The ellipticity at 216 nm is consistent with the initial value in Figure 4A ($-6700 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$), and the shape of the spectrum is indicative of a β -trough, verifying that the intermediate contains a high content of β -sheet structure. Additionally, the peak at 230 nm associated with the exciton signal is completely absent from the spectrum, demonstrating that the aromatic interaction responsible for this signal has not yet formed in this intermediate state.

In contrast to the spectrum measured in diC₉PC, the spectrum measured immediately after mixing with diC₁₃PC LUVs (Figure 5B) overlays well with the unfolded spectrum. Folding is slow in diC₁₃PC at 25 °C, and the protein does not immediately adopt the β -sheet conformation of the early intermediate under this condition. Rather, the observed spectrum is consistent with the signal in diC₁₃PC beginning at the unfolded values in the kinetics measurements. To determine whether the β -structured early intermediate is in fact populated in diC₁₃PC (but too transiently to be observed at 25 °C), we conducted measurements with more ordered bilayers to trap the intermediate. This strategy was based on previous experiments in which a partially-inserted intermediate form of OmpA was observed in gel phase diC₁₄PC bilayers, presumably because the protein could not fully insert into the membrane when it was in the more ordered phase.^{14,52} The intermediate appeared to have similar β -sheet content to the native state, and we hypothesized that this partially-inserted form could correspond to the observed early intermediate β -structured conformation. The main phase transition for diC₁₃PC occurs at 14 °C, with a pre-transition to a ripple phase occurring at 0 °C.⁵³ We conducted measurements in diC₁₃PC at 4 °C, and therefore utilized bilayers in the partially-gel/partially-fluid ripple phase. In the course of these measurements, we determined that the CD signal of OmpA₁₇₁ exhibits a temperature dependence, which can be normalized using a scalar. This correction factor is discussed further in the Supporting Information and is illustrated in Figure S8.

Figure 5B shows the corrected spectrum that was obtained upon incubating OmpA₁₇₁ with diC₁₃PC LUVs at 4 °C. Strikingly, the spectrum matches very closely to the spectrum of the β -structured intermediate obtained immediately upon mixing with diC₉PC (panel A). Slight differences are observed in the region of 230 nm, which could indicate a small degree of aromatic exciton formation in diC₉PC. Still, the agreement of the signal at 216 nm is a strong indication that the same β -structured intermediate is formed in diC₁₃PC as in diC₉PC, and it can be trapped when the lipid bilayer is in a more ordered phase. We further determined that the β -structured intermediate trapped in diC₁₃PC at 4 °C is on-pathway by

allowing this conformation to form and then raising the temperature to 25 °C.¹⁶ Shown in Figure S9B, the ellipticity at 230 nm measured upon raising the temperature increases to the same final value obtained for folding into diC₁₃PC directly at 25 °C. The final wavelength spectrum also overlays well with the spectrum for OmpA₁₇₁ folded in diC₁₃PC at 25 °C and the same fractions folded, unfolded, and elusive were obtained by SDS-PAGE (data not shown), indicating that the same equilibrium of conformations is reached upon raising the temperature to 25 °C. Moreover, the conversion of the β -structured intermediate to the native state is much faster than that observed for a reaction conducted entirely at 25 °C.

A complex multi-step model is required to describe the formation of OmpA₁₇₁ structure

The kinetic traces measured by SDS-PAGE and CD as a function of bilayer thickness are highly complex and provide a wealth of information. Because the three observables available from these techniques report on different aspects of the folding process, we globally analyzed the data using a series of kinetic models to derive the simplest mathematical model that would describe the time-evolution of OmpA₁₇₁ structure under these folding conditions. The process of simply fitting these data to sums of exponentials was not pursued because the observed rate constants and amplitudes of such exponential terms do not correspond to any kinetic species but are rather functions of the microscopic rate constants for all forward and reverse reactions. Our strategy for reconciling all data into one overarching mechanism assumed that the same basic core pathway was followed by OmpA₁₇₁ under all lipid conditions. With this in mind, distinctions in apparent folding kinetics in different lipids were attributed to species that were kinetically silent under “fast folding” conditions (e.g. thin bilayers) or conversely kinetically pronounced under “slow folding” conditions (thicker bilayers). The analytical strategy we followed was to develop a kinetic model for one lipid condition and then extend or contract this model for the other lipid conditions as necessary to arrive at the simplest model that described the data.

Extensive discussion on kinetic model development is included in the Supporting Information. Briefly, we started with the SDS-PAGE data for each lipid condition and developed the mechanism to adequately represent the time-dependence of the species observed by gel (folded, unfolded, and elusive). The CD kinetics data were then used to assign CD signal values to the species in the mechanism and extend the model as necessary. We will first describe the mechanism obtained for folding into diC₁₂PC. The model is subsequently extended to the other lipids, and finally a general mechanism for OmpA β -barrel folding is presented.

We obtained Scheme 1 (shown in Figure 6) as the simplest model to describe OmpA₁₇₁ folding into diC₁₂PC bilayers. Although this scheme looks complex, simpler kinetic schemes containing fewer species were not sufficient to describe the data, and their shortcomings are detailed in the Supporting Information and Figures S10-S12. In Scheme 1, states **A**, **B**, **C**, and **D** represent species that migrate as unfolded by SDS-PAGE, **F** represents the folded species, and **E**₁, **E**₂, **E**₃, and **E**₄ correspond to species that migrate as elusive state. The fitted curves are included in Figures 2, 4 and S13; the corresponding rate constants are listed in Table S2 along with the CD signals for all species in the mechanism; and a sensitivity analysis of the rate constant values is shown in Figure S14.

Intermediates that migrate as unfolded by gel have a mixture of secondary structure profiles

According to this model, the intermediates **B** and **C** migrate with **A** as unfolded by SDS-PAGE. These species are responsible for the lag phase observed in diC₁₂PC in the formation of the folded species by gel. **A** and **B** both exhibit the unfolded CD signal (−3860 at 216 nm and −1216 at 230 nm), causing a lag phase in the CD kinetics at both wavelengths as well. In contrast, **C** possesses an intensely negative signal at 216 nm (−6700), and corresponds to the β-structured early intermediate. Species **D** stems from an off-pathway reaction with slow “error-correction” that results in a slow phase in the disappearance of the unfolded species by gel. This probabilistic misfolding has been proposed previously as part of a “predetermined pathway with optional errors” (PPOE) view of protein folding, and was shown to be consistent with a variety of kinetic folding data.^{22,54} Interestingly, no slow phase is observed in the CD kinetics at 216 nm, so species **D** must possess the same CD signal at 216 nm as the native state (−5000), and thus the native amount of β-sheet structure. However, **D** does not migrate as folded by gel, implying that it lacks the compact barrel structure.

The exciton peak at 230 nm arises from the formation of native structure

Fitting the CD data at 230 nm verified that the aromatic exciton signal arises from the native state. Species **A**, **B**, **C**, and **D** exhibit negative ellipticities at 230 nm corresponding to the shoulder of the unfolded (**A** and **B**) or β-trough (**C** and **D**) spectra. The folded species, **F**, was determined to possess a signal of +700 at 230 nm, due to the aromatic exciton interaction.

Populations that migrate as elusive have varied amounts of secondary structure

The explicit inclusion of multiple elusive states in the mechanism was required to fit the CD kinetics (see Figure S11). Each elusive state has a structure that resembles a state in the “main” pathway, as we found that **E**₁ and **E**₂ exhibit the unfolded CD signal, **E**₃ has the same signal at 216 nm as the β-structured intermediate, and **E**₄ has the same signal at 216 nm as the native state (and thus the native amount of β-sheet structure). This indicates that the elusive population undergoes conformational changes akin to the main folding pathway in terms of the secondary structure composition. It was also determined that **E**₃ and **E**₄ exhibit slightly elevated signals at 230 nm, suggesting some degree of exciton signal. Evidently these species possess a small positive CD band that increases the ellipticity at 230 nm above the baseline value of the β-trough shoulder, but not as large of a positive band as in the native state. The exact conformations of the elusive states are unknown, but it is possible that they contain aromatic residues brought into proximities that give rise to weak exciton interactions. The interacting residues could be the same as those that cause the signal in the native state, but in a slightly different orientation so as to make the signal weaker, or they could be residues from separate protein molecules that are brought together by the potentially multimeric nature of the elusive state. Verification and identification of the interacting residues would require mutagenesis studies and analysis of the protein variants’ spectra.

Similar models to Scheme 1 describe OmpA₁₇₁ folding into diC₁₃PC, diC₁₁PC, diC₉PC, and diC₁₀PC

As shown in Figures 2 and 4, folding of OmpA₁₇₁ into the thicker diC₁₃PC bilayers is much slower than into diC₁₂PC, and there is a more pronounced lag phase. This manifestation in the data required adaptation of Scheme 1 by introducing another state, **A'**, kinetically before the initial state **A** in the mechanism. The resulting Scheme 2 was required to obtain good fits to the observed data, as demonstrated in Figures 2 and 4 and Figure S15. Conversely, folding into the thinner bilayers of diC₁₁PC was observed to be substantially faster than into diC₁₂PC. We determined that Scheme 3, a truncated form of Scheme 1, adequately describes the data. In this scheme, all of the protein begins in state **B** instead of state **A** because the reaction from **A** to **B** is too fast to be observed. There is a reaction step from **B** directly to **E**₂, due to the need for a path to form the elusive state. Because **A** converts rapidly to **B**, the data suggest that no **E**₁ is formed. In addition, a good fit of the data required that we incorporate another optional misfolding step into the pathway, this time stemming from state **B** to form **B'**. The slow correction of this misfolding to convert back to **B** and continue on the productive folding pathway gives rise to the intermediate phase in the observed kinetics. The fitted curves in Figures 2 and 4 and Figure S16 show the excellent fit that was obtained to the diC₁₁PC data using this model.

Folding was observed to be fastest in the thinnest bilayers of diC₉PC and diC₁₀PC. In both of these lipids, the gel data exhibits substantial burst phases in the folded and unfolded species' kinetics, and the CD data begins at a more negative ellipticity, which led to our discovery of the β -structured intermediate state. Because the observed CD signal for both lipids has already decreased to the more negative values at 216 and 230 nm within the mixing time of the experiment, we concluded that all of the protein begins in the intermediate **C** state. Similarly, **E**₁ and **E**₂ are not populated, and **C** must be able to convert directly to **E**₃. To account for the multiple phases in the gel data (including the burst phase), it was necessary to include an optional misfolding step from species **C** to **C'** as well as the off-pathway intermediate **D** included in the other mechanisms. Thus we constructed Scheme 4 (shown in Figure 6), and found that this model fits both the diC₉PC and diC₁₀PC data well (see Figures 2 and 4, and Figures S17 and S18).

OmpA₁₇₁ folding into diC₁₄PC at and above the phase transition temperature follows the same kinetic mechanism determined for the shorter chain lipids

We showed previously that OmpA₁₇₁ folding into diC₁₄PC is accelerated around the phase transition temperature of 24 °C due to a higher incidence of membrane defects,¹⁸ but it was unclear how exactly defects enhance the β -barrel folding pathway. To address this question, we developed models to describe the diC₁₄PC kinetics data at 24, 25, and 26 °C. A comparison of the diC₁₄PC data to that of the shorter chain lipids reveals that folding in diC₁₄PC at 25 °C and 26 °C (just above the phase transition temperature) displays similar kinetic behavior to folding in diC₁₂PC at 25 °C, and we were able to obtain good fits to Scheme 1 (see Figures S4, S7, S19, and S20). In diC₁₄PC at 24 °C (at the phase transition temperature) the formation of native structure was faster, and these data are well described by a simplified version of Scheme 3 that omits the optional error reaction from **B** to **B'** (shown in Figures S4, S7, and S21). The observed faster folding predominantly results from

the protein population beginning in state **B** instead of **A**. This suggests that the transformation from **A** to **B** involves bilayer insertion because it is accelerated by the higher prevalence of defects at the membrane phase transition temperature. Together, these results suggest that OmpA₁₇₁ folding into diC₁₄PC follows the same general pathway as in shorter chain lipids, but with faster rate constants near the lipid phase transition temperature, resulting in faster apparent folding kinetics. Notably, the protein still transforms through the same intermediate, partially-structured states, based on the secondary structure formation kinetics.

A general mechanism for folding of the OmpA β -barrel

Because the central steps in all the kinetic pathways we developed are shared, we argue that the protein follows the same general mechanism under all these lipid conditions. Differences in apparent kinetics are a result of differing rate constants, parts of the mechanism being kinetically invisible under certain conditions, and distinctions in populations of off-pathway misfolded states. A combination of all possible mechanisms leads to Scheme 5. Although this scheme looks incredibly complex, it is important to keep in mind that not all states are populated in all conditions. As such it can be thought of as a “partition function” for the possibilities available in the folding pathway where different lipid conditions place different weights on the various states. Even so, one predominant barrier to folding is state **D**, which is populated in all lipid conditions and represents a prevalent misfolded species *in vitro* that gives rise to the slow phase of folding. A second common observation is that protein is diverted into the “elusive state” pathway, whose rate of formation and final population both depend strongly on the lipid physical state.

Discussion

Structural mechanism proposed for OmpA β -barrel folding

By globally fitting orthogonal data, we propose a comprehensive kinetic pathway for OmpA β -barrel folding (Scheme 5) that reflects the sum of all possible species. The comprehensive scheme contains numerous off-pathway species that have the ability to strongly impact the observed folding. As discussed above, not all species are significantly populated under all lipid conditions, highlighting the active role that membranes play in the folding of transmembrane proteins. Although seemingly complex, the main productive folding pathway in our scheme is consistent with the previously proposed model of Kleinschmidt and colleagues.^{19,55} Detailed knowledge of the CD and SDS-PAGE kinetic behavior allows refinement of this pathway. Figure 7 shows a cartoon that brings together our new data with previous models for OmpA₁₇₁ folding; the attributes of each state are summarized in Table S10. We previously determined that the unfolded aqueous form of OmpA (**U_{AQ}**) has no regular secondary structure and is expanded.²³ In folding, **U_{AQ}** adsorbs to the membrane surface and forms species **A** (**I_{M1}** in Kleinschmidt’s model) based on tryptophan fluorescence,^{15,16} but with no regular secondary structure, as shown by our CD measurements. This state converts to species **B** (**I_{M2}** in Kleinschmidt’s model) which has slightly penetrated into the membrane, based on the Trps being located in the interfacial region of the bilayer.¹⁹ Our CD data indicate that this species also has no regular secondary structure. State **B** does appear to contain some approximation of a circular arrangement of β -

strands based on site-directed Trp quenching experiments.⁵⁵ This work showed that the *trans* ends of adjacent strands (the ends that translocate the bilayer and are located on the extracellular side in the native state) are in closer proximity than the *cis* (periplasmic) ends, and that the first and last strands that close the barrel are in some proximity to each other in the membrane-adsorbed state. We envision this as a clover-like conformation parallel to the membrane surface, with the *trans* ends of the strands located at the center of the clover.

This loosely organized but unstructured intermediate converts to the partially inserted, β -structured intermediate **C** (I_{M3} in Kleinschmidt's model). Previous work has shown that the tryptophans in this state are deeply buried in the membrane, but that the protein is still accessible to proteolysis and does not migrate at the shifted position by SDS-PAGE, indicating it lacks full membrane protection and the compact barrel structure.^{14,19,52} Previous CD and FTIR spectroscopy showed that this state contains substantial β -sheet structure,^{14,52} and it was proposed that a portion of the protein partially inserts into the membrane and adopts a β -sheet configuration in a closed barrel to satisfy hydrogen bonds within the hydrophobic bilayer.⁵⁵ We have unambiguously determined that state **C** contains a higher content of β -sheet structure than previously indicated. We hypothesize that this could be due to the uninserted sections of the protein engaging in β -structure on the membrane surface in addition to the membrane-inserted portion. The splayed geometry of the uninserted β -strands prevents full H-bonding between β -hairpins, which could explain why this state is sensitive to SDS denaturation and migrates as unfolded by SDS-PAGE. The extracellular loops, which are presumably buried in the membrane in this state, could also be forming β -sheet structure, explaining why the β -signal is more intense than in the native state. Subsequent emergence of the loops into the extracellular environment where there is no longer a driving force to satisfy H-bonds would allow those regions to unfold, resulting in the native state having a slightly lower β -content than the partially inserted intermediate **C**.

Previous work has demonstrated that the OmpA₁₇₁ β -barrel inserts unidirectionally into the membrane;¹⁴ and the conformation of intermediate **B** must be arranged so that only the *trans* ends of the β -strands are at the center of the clover shape and thus insert across the bilayer. In addition, the correct polypeptide segments need to be positioned next to each other so that the correct H-bonds can form and the β -strands will be in register in the final barrel structure. We propose that the formation of the periplasmic turns may be what initiates proper β -hairpin arrangement and subsequent strand formation. Folding of the turn regions would force the polypeptide into the clover conformation (with the turns at the tips of the clover "leaves") and the *trans* ends toward the center where they could penetrate into the bilayer in a concerted fashion. In conjunction with insertion, the β -strands would "zip up" along the inserted and surface-located regions of the chain to form the high β -sheet content observed experimentally. In support of this model, it has been shown that mutations in turn residues result in impaired OmpA assembly *in vivo*.⁵⁶ With this model in mind, it would be of great interest to examine more closely the role of turn formation in OmpA folding.

An additional spectroscopic feature we monitored in our studies was an aromatic exciton interaction that induces a positive peak in the CD signal at 230 nm. In fitting the kinetics data, we determined that only the native barrel possesses this spectroscopic feature. Because it is unknown exactly which residues of the 5 tryptophans and 13 tyrosines in OmpA₁₇₁ are

responsible for the signal, we cannot specify the aromatic interaction that forms in the native state and gives rise to the interaction. In PagP, the aromatic pair identified as responsible for an exciton signal are positioned on either side of the interior of the β -barrel, and thus only interact in the native state.⁴⁸ Similarly, there are two tyrosine residues on the interior of the OmpA β -barrel in the crystal structure,⁵⁷ and these residues are intriguing candidates for the formation an exciton interaction. Complicating a straightforward interpretation is the possibility that the interacting aromatic residues in OmpA₁₇₁ could be externally located, because there are also several pairs of aromatic residues brought into close proximity on adjacent strands on the exterior of the β -barrel. Evidence for the interacting residues being externally located comes from the observation that the native state exhibits a weaker CD contribution at 230 nm in the thicker bilayers of diC₁₄PC than in thinner bilayers, which may indicate that the lipid bilayer environment influences the exciton coupling. It is difficult to predict significant aromatic exciton formation based on side chain geometry and distance,⁵⁰ and experimental mutagenesis studies to investigate this are beyond the scope of the current work.

OmpA is subject to folding errors

The kinetic modeling led us to identify several novel off-pathway intermediate states in the OmpA folding mechanism that give rise to slower phases of folding. Species **D** was necessary for fitting the apparent kinetics in all lipid conditions, indicating that it is a ubiquitous intermediate state. This conformation is formed by state **C** in competition with **F**, and must convert back to **C** to reach the native state. The slower rate of “error correction” is what causes the slow continued appearance of **F** at long times. Under some lipid conditions we observed that species **B** and **C** can also populate the off-pathway states **B'** and **C'**, respectively, which contribute intermediate kinetic phases to the data.

These reaction steps are representative of the “predetermined pathway with optional errors” (PPOE) model for protein folding proposed by the Englander group as an alternative to unrelated parallel pathways.²² Rather than folding via multiple separate pathways within a funnel-shaped free energy landscape, this theory suggests that proteins follow predetermined paths of folding based on the sequential incorporation of structural elements, or “foldons.” Kinetic heterogeneity can be caused by the population of off-pathway intermediate states and does not require the presence of parallel folding routes. It is difficult to envision a β -barrel folding by a substantially different mechanism than the one we have described, there is no direct experimental evidence for the population of significantly different intermediate states, and the presence of off-pathway states is supported by the fitting of our data, so we propose that OMPs also fold by a dominant pathway with the correction of optional errors causing apparent multi-exponential kinetics.

Another major off-pathway species identified in these studies was the elusive state. More likely an ensemble of multimeric conformations, the elusive state denotes the population of OmpA₁₇₁ that migrates anomalously by SDS-PAGE. CD measurements revealed that the elusive state gains secondary structure similar to the species on the main folding pathway and that species **E₃** and **E₄** exhibit a weak aromatic exciton signal. Because elusive states are non-native states that compete with productive folding, they too constitute off-pathway states

in a PPOE model. Although its existence has not been previously discussed, examination of the literature shows that many other OMPs exhibit the formation of elusive state when folding into synthetic bilayers. The elusive state is therefore of broad importance for OMP folding studies and suggests that other OMPs are likely to display the types of complex folding pathways observed here for OmpA. This awaits further investigation, but the common feature shared by many different OMPs – the elusive state – presents as a likely species to be suppressed during folding *in vivo* by the BAM complex.

Bilayer defects drive β -barrel folding and insertion

Membrane properties are well known to strongly influence OMP folding kinetics: thinner bilayers, more curved bilayers, and bilayers undergoing a phase transition all promote faster, more efficient folding, presumably due to a higher prevalence of membrane defects.^{18,37,40} A comparison of the fitted rate constants in different lipid conditions for the main pathway steps in our model shows that the steps are all accelerated by thinning of the lipid bilayer, with the earliest steps affected the most. Furthermore, folding in diC₁₄PC at the phase transition temperature – a condition that produces bilayer in-homogeneities – follows the same basic kinetic model. As shown in Figure 7, progressive bilayer insertion is an important aspect of the transformation of each conformation to the next one. Therefore, it makes sense that folding would occur faster in bilayers that have more defects because these would lower the activation energy barriers.

The fact that all steps in the folding pathway are strongly modulated by membrane structure is of particular relevance to an understanding of OMP biogenesis *in vivo*. While chaperones ensure that nascent OMP chains transit across the aqueous periplasm, insertion into the outer membranes is the final, rate-limiting step in OMP maturation.⁵⁸ This reaction is catalyzed by the multi-protein BAM complex in which the central essential subunit, BamA, is itself a transmembrane β -barrel. By its very nature, BamA must have extensive interactions with membrane lipids. It has been proposed that BamA possesses a lateral gate between β -strands one and sixteen, which could expose its aqueous interior to hydrophobic regions of the membrane.⁵⁹ The membrane must adapt to this situation if it occurs. Indeed, microsecond-long molecular dynamics simulations of BamA showed that it induced bilayer thinning by up to 16 Å near this same β -strand seam in the protein.⁵⁹ A subsequent simulation revealed that the soluble POTRA motifs of BamA diffuse laterally along the membrane surface, which could also alter lipid packing.⁶⁰ If BamA possesses the high thermodynamic stability similar to that measured for other transmembrane β -barrels, part of its folding energy could be used to stabilize sparsely populated, high energy structures of the membrane. With the current kinetic model in mind, BAM-induced bilayer defects – however they occur – represent a mechanism to accelerate all steps of OMP folding.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Source Statement: This work was supported by grants from the National Science Foundation (NSF) (MCB 0919868 and MCB 1412108) and the National Institutes of Health (NIH) (R01 GM079440 and T32 GM008403). E.J.D. is a recipient of an NSF graduate research fellowship (DGE-0707427).

We gratefully acknowledge Michael McCaffery and Erin Pryce of the Johns Hopkins Integrated Imaging Center for assistance with electron microscopy. We also thank members of the Fleming lab and Professor Doug Barrick for many helpful discussions.

Abbreviations

CD	circular dichroism
LUV	large unilamellar vesicle
OMP	outer membrane protein
OmpA	Outer Membrane Protein A
OmpA₁₇₁	the β -barrel domain (residues 1–171) of OmpA
PPOE	predetermined pathway with optional errors
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis

References

1. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 2001; 305:567–580. [PubMed: 11152613]
2. Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? *Nat. Rev. Drug Discov.* 2006; 5:993–996. [PubMed: 17139284]
3. Sanders CR, Myers JK. Disease-related misassembly of membrane proteins. *Annu. Rev. Biophys. Biomol. Struct.* 2004; 33:25–51. [PubMed: 15139803]
4. Koebnik R, Locher KP, Van Gelder P. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol. Microbiol.* 2000; 37:239–253. [PubMed: 10931321]
5. Paschen SA, Neupert W, Rapaport D. Biogenesis of beta-barrel membrane proteins of mitochondria. *Trends Biochem. Sci.* 2005; 30:575–582. [PubMed: 16126389]
6. Hsu SC, Inoue K. Two evolutionarily conserved essential beta-barrel proteins in the chloroplast outer envelope membrane. *Biosci. Trends.* 2009; 3:168–178. [PubMed: 20103843]
7. Wimley WC. The versatile beta-barrel membrane protein. *Curr. Opin. Struct. Biol.* 2003; 13:404–411. [PubMed: 12948769]
8. Tamm LK, Hong H, Liang B. Folding and assembly of beta-barrel membrane proteins. *Biochim. Biophys. Acta.* 2004; 1666:250–263. [PubMed: 15519319]
9. Ruiz N, Kahne D, Silhavy TJ. Advances in understanding bacterial outer-membrane biogenesis. *Nat. Rev. Microbiol.* 2006; 4:57–66. [PubMed: 16357861]
10. Kolodziejek AM, Schnider DR, Rohde HN, Wojtowicz AJ, Bohach GA, Minnich SA, Hovde CJ. Outer membrane protein X (Ail) contributes to *Yersinia pestis* virulence in pneumonic plague and its activity is dependent on the lipopolysaccharide core length. *Infect. Immun.* 2010; 78:5233–5243. [PubMed: 20837715]
11. Hill K, Model K, Ryan MT, Dietmeier K, Martin F, Wagner R, Pfanner N. Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. *Nature.* 1998; 395:516–521. [PubMed: 9774109]

12. Denoncin K, Schwalm J, Vertommen D, Silhavy TJ, Collet JF. Dissecting the Escherichia coli periplasmic chaperone network using differential proteomics. *Proteomics*. 2012; 12:1391–1401. [PubMed: 22589188]
13. Wu T, Malinverni J, Ruiz N, Kim S, Silhavy TJ, Kahne D. Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. *Cell*. 2005; 121:235–245. [PubMed: 15851030]
14. Surrey T, Jähnig F. Refolding and oriented insertion of a membrane protein into a lipid bilayer. *Proc. Natl. Acad. Sci. U.S.A.* 1992; 89:7457–7461. [PubMed: 1502158]
15. Surrey T, Jähnig F. Kinetics of folding and membrane insertion of a beta-barrel membrane protein. *J. Biol. Chem.* 1995; 270:28199–28203. [PubMed: 7499313]
16. Kleinschmidt JH, Tamm LK. Folding intermediates of a beta-barrel membrane protein. Kinetic evidence for a multi-step membrane insertion mechanism. *Biochemistry*. 1996; 35:12993–13000. [PubMed: 8855933]
17. Gessmann D, Chung YH, Danoff EJ, Plummer AM, Sandlin CW, Zaccari NR, Fleming KG. Outer membrane beta-barrel protein folding is physically controlled by periplasmic lipid head groups and BamA. *Proc. Natl. Acad. Sci. U.S.A.* 2014; 111:5878–5883. [PubMed: 24715731]
18. Danoff EJ, Fleming KG. Membrane defects accelerate outer membrane beta-barrel protein folding. *Biochemistry*. 2015; 54:97–99. [PubMed: 25513891]
19. Kleinschmidt JH, den Blaauwen T, Driessen AJ, Tamm LK. Outer membrane protein A of Escherichia coli inserts and folds into lipid bilayers by a concerted mechanism. *Biochemistry*. 1999; 38:5006–5016. [PubMed: 10213603]
20. Pocanschi CL, Apell HJ, Puntervoll P, Høgh B, Jensen HB, Welte W, Kleinschmidt JH. The major outer membrane protein of Fusobacterium nucleatum (FomA) folds and inserts into lipid bilayers via parallel folding pathways. *J. Mol. Biol.* 2006; 355:548–561. [PubMed: 16310217]
21. Huysmans GH, Radford SE, Baldwin SA, Brockwell DJ. Malleability of the folding mechanism of the outer membrane protein PagP: parallel pathways and the effect of membrane elasticity. *J. Mol. Biol.* 2012; 416:453–464. [PubMed: 22245579]
22. Krishna MM, Englander SW. A unified mechanism for protein folding: predetermined pathways with optional errors. *Protein Sci.* 2007; 16:449–464. [PubMed: 17322530]
23. Danoff EJ, Fleming KG. The soluble, periplasmic domain of OmpA folds as an independent unit and displays chaperone activity by reducing the self-association propensity of the unfolded OmpA transmembrane beta-barrel. *Biophys. Chem.* 2011; 159:194–204. [PubMed: 21782315]
24. Schäfer U, Beck K, Müller M. Skp, a molecular chaperone of gram-negative bacteria, is required for the formation of soluble periplasmic intermediates of outer membrane proteins. *J. Biol. Chem.* 1999; 274:24567–24574. [PubMed: 10455120]
25. Pace CN, Vajdos F, Fee L, Grimsley G, Gray T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* 1995; 4:2411–2423. [PubMed: 8563639]
26. Caffrey M, Hogan J. LIPIDAT: a database of lipid phase transition temperatures and enthalpy changes. DMPC data subset analysis. *Chem. Phys. Lipids*. 1992; 61:1–109. [PubMed: 1315624]
27. Bustamante C, Tinoco I Jr, Maestre MF. Circular differential scattering can be an important part of the circular dichroism of macromolecules. *Proc. Natl. Acad. Sci. U.S.A.* 1983; 80:3568–3572. [PubMed: 6574499]
28. Chakraborty H, Lentz BR. A simple method for correction of circular dichroism spectra obtained from membrane-containing samples. *Biochemistry*. 2012; 51:1005–1008. [PubMed: 22264128]
29. Ladokhin AS, Fernandez-Vidal M, White SH. CD spectroscopy of peptides and proteins bound to large unilamellar vesicles. *J. Membr. Biol.* 2010; 236:247–253. [PubMed: 20706833]
30. Gill SC, von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 1989; 182:319–326. [PubMed: 2610349]
31. Edelhoch H. Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry*. 1967; 6:1948–1954. [PubMed: 6049437]
32. Berberan-Santos MN, Martinho JMG. The integration of kinetic rate equations by matrix methods. *J. Chem. Educ.* 1990; 67:375–379.

33. Nakamura K, Mizushima S. Effects of heating in dodecyl sulfate solution on the conformation and electrophoretic mobility of isolated major outer membrane proteins from *Escherichia coli* K-12. *J. Biochem.* 1976; 80:1411–1422. [PubMed: 828162]
34. Reithmeier RAF, Bragg PD. Molecular characterization of a heat-modifiable protein from the outer membrane of *Escherichia coli*. *Arch. Biochem. Biophys.* 1977; 178:527–534. [PubMed: 319761]
35. Ohnishi S, Kameyama K, Takagi T. Characterization of a heat modifiable protein, *Escherichia coli* outer membrane protein OmpA in binary surfactant system of sodium dodecyl sulfate and octylglucoside. *Biochim. Biophys. Acta.* 1998; 1375:101–109. [PubMed: 9767139]
36. Schweizer M, Hindennach I, Garten W, Henning U. Major proteins of the *Escherichia coli* outer cell envelope membrane. Interaction of protein II* with lipopolysaccharide. *Eur. J. Biochem.* 1978; 82:211–217. [PubMed: 340230]
37. Kleinschmidt JH, Tamm LK. Secondary and tertiary structure formation of the beta-barrel membrane protein OmpA is synchronized and depends on membrane thickness. *J. Mol. Biol.* 2002; 324:319–330. [PubMed: 12441110]
38. Hong H, Tamm LK. Elastic coupling of integral membrane protein stability to lipid bilayer forces. *Proc. Natl. Acad. Sci. U.S.A.* 2004; 101:4065–4070. [PubMed: 14990786]
39. Huysmans GH, Radford SE, Brockwell DJ, Baldwin SA. The N-terminal helix is a post-assembly clamp in the bacterial outer membrane protein PagP. *J. Mol. Biol.* 2007; 373:529–540. [PubMed: 17868697]
40. Burgess NK, Dao TP, Stanley AM, Fleming KG. Beta-barrel proteins that reside in the *Escherichia coli* outer membrane in vivo demonstrate varied folding behavior in vitro. *J. Biol. Chem.* 2008; 283:26748–26758. [PubMed: 18641391]
41. Wang H, Andersen KK, Vad BS, Otzen DE. OmpA can form folded and unfolded oligomers. *Biochim. Biophys. Acta.* 2013; 1834:127–136. [PubMed: 22982243]
42. Pocanschi CL, Patel GJ, Marsh D, Kleinschmidt JH. Curvature elasticity and refolding of OmpA in large unilamellar vesicles. *Biophys. J.* 2006; 91:L75–77. [PubMed: 16891370]
43. Ye C, Chai Q, Zhong M, Wei YN. Effect of crowding by Ficolls on OmpA and OmpT refolding and membrane insertion. *Protein Sci.* 2013; 22:239–245. [PubMed: 23225740]
44. Sugawara E, Steiert M, Rouhani S, Nikaido H. Secondary structure of the outer membrane proteins OmpA of *Escherichia coli* and OprF of *Pseudomonas aeruginosa*. *J. Bacteriol.* 1996; 178:6067–6069. [PubMed: 8830709]
45. Eisele JL, Rosenbusch JP. In vitro folding and oligomerization of a membrane protein. Transition of bacterial porin from random coil to native conformation. *J. Biol. Chem.* 1990; 265:10217–10220. [PubMed: 1693917]
46. Dekker N, Merck K, Tommassen J, Verheij HM. In vitro folding of *Escherichia coli* outer-membrane phospholipase A. *Eur. J. Biochem.* 1995; 232:214–219. [PubMed: 7556153]
47. Kramer RA, Zandwijken D, Egmond MR, Dekker N. In vitro folding, purification and characterization of *Escherichia coli* outer membrane protease OmpT. *Eur. J. Biochem.* 2000; 267:885–893. [PubMed: 10651827]
48. Khan MA, Neale C, Michaux C, Pomes R, Prive GG, Woody RW, Bishop RE. Gauging a hydrocarbon ruler by an intrinsic exciton probe. *Biochemistry.* 2007; 46:4565–4579. [PubMed: 17375935]
49. Manning MC, Woody RW. Theoretical study of the contribution of aromatic side chains to the circular dichroism of basic bovine pancreatic trypsin inhibitor. *Biochemistry.* 1989; 28:8609–8613. [PubMed: 2481497]
50. Grishina IB, Woody RW. Contributions of tryptophan side chains to the circular dichroism of globular proteins: exciton couplets and coupled oscillators. *Faraday Discuss.* 1994; 99:245–262.
51. Kuwajima K, Garvey EP, Finn BE, Matthews CR, Sugai S. Transient intermediates in the folding of dihydrofolate reductase as detected by far-ultraviolet circular dichroism spectroscopy. *Biochemistry.* 1991; 30:7693–7703. [PubMed: 1868049]
52. Rodionova NA, Tatulian SA, Surrey T, Jähnig F, Tamm LK. Characterization of two membrane-bound forms of OmpA. *Biochemistry.* 1995; 34:1921–1929. [PubMed: 7849052]

53. Lewis RN, Mak N, McElhaney RN. A differential scanning calorimetric study of the thermotropic phase behavior of model membranes composed of phosphatidylcholines containing linear saturated fatty acyl chains. *Biochemistry*. 1987; 26:6118–6126. [PubMed: 3689765]
54. Bédard S, Krishna MM, Mayne L, Englander SW. Protein folding: independent unrelated pathways or predetermined pathway with optional errors. *Proc. Natl. Acad. Sci. U.S.A.* 2008; 105:7182–7187. [PubMed: 18480257]
55. Kleinschmidt JH, Bulieris PV, Qu JA, Dogterom M, den Blaauwen T. Association of neighboring beta-strands of outer membrane protein A in lipid bilayers revealed by site-directed fluorescence quenching. *J. Mol. Biol.* 2011; 407:316–332. [PubMed: 21256134]
56. Koebnik R, Krämer L. Membrane assembly of circularly permuted variants of the *E. coli* outer membrane protein OmpA. *J. Mol. Biol.* 1995; 250:617–626. [PubMed: 7623380]
57. Pautsch A, Schulz GE. High-resolution structure of the OmpA membrane domain. *J. Mol. Biol.* 2000; 298:273–282. [PubMed: 10764596]
58. Costello SM, Plummer AM, Fleming PJ, Fleming KG. Dynamic periplasmic chaperone reservoir facilitates biogenesis of outer membrane proteins. *Proc. Natl. Acad. Sci. U.S.A.* 2016; 113:E4794–4800. [PubMed: 27482090]
59. Noinaj N, Kuszak AJ, Gumbart JC, Lukacik P, Chang H, Easley NC, Lithgow T, Buchanan SK. Structural insight into the biogenesis of beta-barrel membrane proteins. *Nature*. 2013; 501:385–390. [PubMed: 23995689]
60. Fleming PJ, Patel DS, Wu EL, Qi Y, Yeom MS, Sousa MC, Fleming KG, Im W. BamA POTRA domain interacts with a native lipid membrane surface. *Biophys. J.* 2016; 110:2698–2709. [PubMed: 27332128]
61. Pautsch A, Schulz GE. Structure of the outer membrane protein A transmembrane domain. *Nat. Struct. Biol.* 1998; 5:1013–1017. [PubMed: 9808047]

Figure 1.

OmpA₁₇₁ populates an additional conformation to the canonical folded and unfolded states during SDS-PAGE. (A) Representative gel and corresponding densitometry values for OmpA₁₇₁ folding into diC₁₂PC LUVs. The bands corresponding to the folded and unfolded conformations on the gel are indicated by “F” and “U”, respectively. The gel has been darkened for better clarity but densitometry analysis was performed on the unaltered scan. (B) Fractions folded and unfolded for OmpA₁₇₁ folding into diC₁₂PC. Fractional quantities were calculated by dividing by the sum of the folded and unfolded band intensities (“F+U”)

or by the boiled band intensity (“B”). Solid lines are fits to a quadruple exponential expression.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2.

Time-dependence of species measured by SDS-PAGE for OmpA₁₇₁ folding into LUVs composed of diC₉PC – diC₁₃PC at 25 °C: (A) fraction folded, (B) fraction unfolded, and (C) fraction elusive. Densitometry values were converted to fractional quantities by dividing by the boiled band intensity. The legend is shown in panel A. Solid lines are fits to kinetic models discussed in the text and shown in Figure 6. Data are displayed over two time ranges: 1800 s (30 min) and 18000 s (5 h). Long-term time points can be seen in plots for individual lipid conditions, in Figures S13 and S15-S18.

Figure 3.

CD wavelength spectra of OmpA₁₇₁ in various lipids collected after kinetics measurements. Curves were corrected with the corresponding LUV spectrum (Figure S6), and converted to mean residue ellipticity using the concentration measured before kinetics. Each plot also contains the spectrum for unfolded OmpA₁₇₁ in 1 M urea as a dashed red line. (A) Spectrum after folding into LUVs of diC₁₀PC for 5 h (blue). Arrows indicate the change in signal at 216 nm and 230 nm from the unfolded state. (B) Spectra after folding into LUVs of diC₉PC – diC₁₄PC for 5 h at 25 °C. The trough at 216 nm is indicative of β -sheet structure and the

peak at 230 nm is thought to be due to an exciton interaction between aromatic residues in the native state.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 4.

CD signal kinetics measured at (A) 216 nm, and (B) 230 nm, for OmpA₁₇₁ folding into LUVs composed of diC₉PC – diC₁₃PC at 25 °C, under the same conditions as SDS-PAGE kinetics (Figure 2). Data were converted to mean residue ellipticity using the concentration measured before initiation of folding. The legend is shown in panel A. The signal for unfolded OmpA₁₇₁ at the appropriate wavelength is plotted in each panel as a dashed red line. Black lines are fits to kinetic models discussed in the text and shown in Figure 6. Data are displayed over two time ranges: 1800 s (30 min) and 14400 s (4 h).

Figure 5.

CD wavelength spectra of OmpA₁₇₁ immediately after initiating folding into LUVs of (A) diC₉PC, or (B) diC₁₃PC (dotted lines) at 25 °C, and after incubation with diC₁₃PC at 4 °C (B; orange line). “Initial” wavelength data were collected in 10 nm sections to capture the signal at the earliest times possible, and then combined into a full spectrum. The spectrum at 4 °C (B) was collected at the conclusion of kinetics measurements at 4 °C, and scaled by a factor of 1.25 to correct for the temperature dependence of the signal (see Supporting Information). Also included in each plot for comparison are the spectrum for unfolded

OmpA₁₇₁ (dashed red line), and the spectrum measured after folding was complete in the corresponding lipid at 25 °C (solid light blue and purple lines), from Figure 3B.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Scheme 1

Scheme 3

Scheme 2

Scheme 4

Scheme 5

Figure 6.

Kinetic schemes developed to describe OmpA₁₇₁ folding behavior under different lipid conditions. Species that migrate with the unfolded band by gel are colored red, the natively folded state is colored blue, and the off-pathway “elusive” species that migrate anomalously by gel are colored green.

Figure 7.

Illustration of the proposed on-pathway intermediate states in the folding mechanism for the OmpA β -barrel. The attributes of each species are summarized in Table S10 of the Supporting Information. U_{AQ} is a folding competent, non-membrane-associated unfolded state. Upon membrane association, U_{AQ} converts to species **A**, which contains no regular structure. **A** then converts to species **B**, which has no regular secondary structure but has a loose arrangement of β -hairpins in a “clover-like” shape, oriented parallel to the membrane surface. Site-directed quenching studies have revealed that residues on the translocating (*trans*) portions of adjacent strands (green circles) are in closer proximity in this state than residues on the periplasmic (*cis*) portions of strands (blue circles).⁵⁵ **B** converts to species **C**, which is partially inserted and contains a higher β -sheet content than the native state. For this reason, the portions of the strands remaining on the surface as well as the membrane-embedded extracellular loops are proposed to be engaged in H-bonded β -strand conformations in addition to the membrane-inserted portion of the β -barrel. The closer-proximity *trans* residues are again indicated by green circles and the *cis* residues as blue circles. Full insertion of intermediate **C** leads to formation of the native β -barrel structure, **F** (OmpA₁₇₁ PDB id 1BXW⁶¹). This state exhibits a shifted SDS-PAGE migration and an aromatic exciton signal in the CD spectrum at 230 nm. The interacting residues responsible for this signal are unknown, but the interaction only occurs in the native state and is depicted by a yellow star. The *trans* and *cis* residues used in site-directed quenching experiments are again indicated by green and blue circles, respectively.