Title
Membrane Protein Topology Acquisition and Folding Through the Transition State

Permalink
https://escholarship.org/uc/item/6z7644mj

Author
Woodall, Nicholas

Publication Date
2016

Peer reviewed|Thesis/dissertation
Membrane Protein Topology Acquisition and Folding Through the Transition State

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Biochemistry and Molecular Biology

by

Nicholas B. Woodall
ABSTRACT OF DISSERTATION

Membrane Protein Topology Acquisition and the Folding Through the Transition State

by

Nicholas B. Woodall
Doctor of Philosophy in Biochemistry and Molecular Biology
University of California, Los Angeles, 2016
Professor James U. Bowie, Chair

The folding mechanisms of helical membrane proteins remain largely uncharted. Here we examine two of the primary events during membrane protein folding: insertion/topology acquisition into the lipid bilayer and the assembly of alpha helices into their native state. Towards this goal we characterize the kinetics of bacteriorhodopsin folding and employ φ-value analysis to explore the folding transition state. We were able to obtain reliable φ-values for 16 mutants of bacteriorhodopsin well distributed throughout the protein. Every φ-value was less than 0.4, indicating the transition state is not uniquely structured. We suggest that the transition state is a loosely organized ensemble of conformations. Before the final folding of the alpha helices, the transmembrane segments must be properly aligned in the lipid bilayer. The topology of helical membrane proteins is generally defined during insertion of the transmembrane helices. It is now clear, however, that topology can remain malleable in some cases. We show that the entire EmrE protein can indeed invert topology after the full protein is translated and inserted. Wholesale topology flipping is remarkable given the physical constraints of the membrane and exposes new avenues to envision folding pathways, both productive and detrimental.
The dissertation of Nicholas B. Woodall is approved.

Robert T. Clubb

Pascal F. Egea

James U. Bowie, Committee Chair

University of California, Los Angeles

2016
Table of Contents

Abstract of Dissertation  ii
Table of Contents  iv
Acknowledgements  vi
VITA  vii
Publications  viii

Chapter 1: Introduction to Membrane Protein Folding  1

1.1 Introduction to Membrane Environment  2
1.2 α-Helical Membrane Proteins  3
1.3 Membrane Protein Folding: Stage 1  4

1.3.1 Insertion of α-helices  4
1.3.2 Membrane Protein Topology  5

1.4 Membrane Protein Folding: Stage 2  8

1.4.1 The Denatured State Ensemble  8
1.4.2 The Native State  9
1.4.3 The Transition State  10

Chapter 2: Bacteriorhodopsin Folds through a Poorly Organized Transition State  23

2.1 Abstract  24
2.2 Introduction  24
2.3 Results and Discussion  27
2.4 Conclusions  37
2.5 Experimental Section  40
2.6 References  56

Chapter 3: Dual-Topology Insertion of a Dual-Topology Protein  59

3.1 Abstract  60
3.2 Results  62
3.3 Conclusion  68
3.4 Methods  70
3.5 References  89

Chapter 4: Complete Topology Inversion  92

4.1 Abstract  93
4.2 Introduction  93
4.3 Results and Discussion  95
4.4 Conclusion  101
4.5 Methods  102
4.6 References  114
List of Figures and Tables

Chapter 1: Introduction to the Membrane Environment

Figure 1-1 Structure of the Bilayer 11
Figure 1-2 α-helices in the Bilayer 12
Figure 1-3 Spontaneous Insertion 13
Figure 1-4 Ribosomal Insertion 14
Figure 1-5 Positive-Inside Rule 15
Figure 1-6 Folding Funnel 16
Figure 1-7 Multiple States Folding Funnel 17
Figure 1-8 Transition State Energy Landscape 18

Chapter 2: Bacteriorhodopsin Folds through a Poorly Organize Transition State

Figure 2-1 Conformational Relaxation Kinetics of Br 44
Figure 2-2 Curve-fitting Residuals 45
Figure 2-3 Retinal Hydrolysis 46
Figure 2-4 Kinetic Correlations 47
Figure 2-5 M-Value Change in Pathway 48
Figure 2-6 Trace Change in Pathway 49
Figure 2-7 Chevron Plots 50
Figure 2-8 Kinetic and Equilibrium Comparison 53
Table 2-1 Φ-value of bR 54
Figure 2-9 Φ-value map of bR 55

Chapter 3: Dual Topology Insertion of a Dual Topology Protein

Figure 3-1 Dual Topology Insertion Models 76
Figure 3-2 Indirect Topology Assay of EmrE-C+ and HA-EmrE-C+ 77
Figure 3-3 Ethidium Bromide Resistance of HA-EmrE 78
Figure 3-4 FtsH Dependent Proteolysis of EmrE 79
Figure 3-5 Cysteine-Accessibility Method 80
Figure 3-6 Cysteine-Free Proteins are not Reactive 81
Figure 3-7 Protein Amount is Linearly Related to Band Intensity 82
Figure 3-8 Direct Topology Assay of EmrE-C+ and HA-EmrE-C+ 83
Figure 3-9 AMS Reactivity Tests for EmrE 84
Figure 3-10 Disulfide Formation at the N-terminus 85
Figure 3-11 Time Dependence of AMS Reactivity 86
Figure 3-12 DNA Sequences of EmrE Constructs 87
Figure 3-13 Uncropped Blot of HA-EmrE<sub>Class</sub> C+ F27C 88

Chapter 4: Complete Topology Inversion can be Part of Normal Membrane Protein Biogenesis

Figure 4-1 EmrE Constructs and Topology Models 107
Figure 4-2 Topology Changes Assessed by Ethidium Bromide Resistance 108
Figure 4-3 Toxicity of the N<sub>out</sub>-HA-EmrE-C+ Construct 109
Figure 4-4 Substituted Cysteine Accessibility Method (SCAM) 110
Figure 4-5 Background Biotinylation Control 111
Figure 4-6 Analysis of C-terminal Topology Changes 112
Figure 4-7 Full Topology Inversion of EmrE 113
Acknowledgements

Any incremental knowledge obtained from this work is meaningless without the efforts of those who came before and those who will follow. I salute you.

On a personal note, the people in my life make it worth living. Thank you.
VITA

2007-2010 Bachelors of Science
Georgia Institute of Technology

2008-2010 Research Assistant
Georgia Institute of Technology
Advisor: Professor Andreas Bommarius

2011 Research Assistant
University of Tennessee Knoxville
Advisor: Professor Albrecht von Arnim

2011-2016 Graduate Student Researcher
Department of Biochemistry Molecular and Structural Biology
University of California Los Angeles
Advisor: James Bowie

2012-2015 Chemistry Biology Interface Training Grant Awardee
University of California Los Angeles

2014 Anatrace Travel Award
Membrane Protein Structural Dynamics Conference Chicago

2016 Fowler Fellowship Award
University of California Los Angeles

2016 Dissertation Year Fellowship
University of California Los Angeles
Publications

**Woodall N.,** Hadley S., Yin Y., Bowie J., Complete Topology Changes can be a Normal Part of Membrane Protein Biogenesis (2016) Submitted


*Co-first Author

Chapter 1:
Introduction to Membrane Protein Folding
1.1 Introduction to Membrane Environment

The cell membrane separates the inside of the cell from the chaotic outside world. Membrane proteins residing in the membrane connect the cell to the outside world by importing nutrients, exporting wastes, anchoring cellular structures, acting as sensors as well as defending the cell through the active removal of toxins. Unfortunately, studying this important class of proteins has been hampered by the complexity of the membrane environment. Determining the critical mechanical properties of biological membranes either by experiment or simulation is still fraught with uncertainty for even single-component systems\(^1\). Adding a complex polymer such as an \(\alpha\)-helical membrane protein into a multi-component biological membrane only deepens the mystery. Furthermore, using native-like environments to study membrane proteins interferes with many common laboratory procedures. As a result, the majority of membrane protein studies are performed in cell membrane mimetics, detergents, that are only sporadically effective proxies\(^2\). Finally, few membrane proteins express to a high level and are often toxic when over-expressed\(^3\). Understanding processes in membrane protein biogenesis such as topology acquisition and folding may reduce the complications embedded in the study of membrane proteins allowing us to gain another foothold into this fascinating world.

A hydrophobic stretch of alkyl chains composes the kinetic barrier of the cell membrane preventing the free crossing of polar molecules and ions via the dielectric effect. The major component of the barrier are phospholipids which consist of a hydrophilic headgroup and two fatty acid tails (Fig. 1-1 A). When placed in water, phospholipids spontaneously form bilayers with the hydrophilic head groups positioned
towards water and the alkyl groups sandwiched in between (Fig. 1-2A). Despite the energetic cost to entering the hydrophobic barrier region presented by the alkyl chains, the permeation of water across lipid bilayers still occurs to a significant degree⁴. Mechanistically, water molecules need to gain enough energy to desolvate from the ionic lipid headgroups and diffuse into the alkyl chains to cross the lipid bilayer. Some crossing events may also be catalyzed by the formation of transient defects in the membrane. Even with simple processes in the bilayer, the mechanistic details and influence of critical bilayer properties to water permeation are still being debated⁵. A rough calculation based on the water permeability coefficient states that in a standard phospholipid bilayer roughly 4000 water molecules rush pass each individual phospholipid every second. For small ions like sodium, however, each phospholipid lets only one sodium ion through every 70 hours⁴. It is important to note that changes in bilayer composition can change any of the bilayer properties like permeability by orders of magnitude⁵. In general, lipid bilayers are quite impermeant to ions.

1.2 α-Helical Membrane Proteins

The simple hydrophobic barrier of the lipid bilayer gains the true functionality of a biological membrane through the use of proteins. Proteins are sophisticated polymers composed of amino acids that fold into intricate 3-dimensional shapes to perform complex functions such as the selective permeation of a single type of ion or actively pumping out wastes. Proteins that perform functions such as these must interact with the full range of the bilayer and are denoted as integral membrane proteins. Their significant hydrophobicity ensures that these proteins are a permanent fixture in the
bilayer. The vast number of functions these proteins perform causes them to occupy ~25% of the coding genes in all organisms. By far the largest class of integral membrane proteins are α-helical membrane proteins. Composed of hydrophobic helices that span the bilayer, these transmembrane α-helices are the exclusive type of secondary structure present in the plasma membrane of both prokaryotes and eukaryotes (Fig 1-2).

The transmembrane helices are connected together by hydrophilic extra-membrane loops resulting in a polypeptide chain composed of alternating hydrophobic and hydrophilic segments. With hydrophilic segments split across the membrane, some of the segments must traverse the kinetic barrier of the membrane for insertion into the bilayer (Fig. 1-3). Some small proteins can spontaneously insert by coupling the energetic cost of desolvating polar groups with the favorable energy of the hydrophobic segments entering the apolar interior of the bilayer. For most proteins, life must catalyze this process.

1.3 Membrane Protein Folding: Stage 1
   1.3.1 Insertion of α-helices

   The two-stage model as proposed 16 years ago by Popot and Engleman has correctly guided thought on membrane protein folding. In the first stage, hydrophobic helices are inserted into the membrane and exist as separated α-helices. In the second stage, the α-helices coalesce to form their native structure. The first insertion stage is accomplished through the use of the Sec-translocon. As the first transmembrane helix exits the ribosome, a complex system stalls translation and delivers the ribosome nascent-chain complex to the translocon. The translocon consists of a gated
transmembrane-helix size pore through the membrane. Once the translating ribosome binds the translocon, the first transmembrane helix enters the translocon pore. As translation continues, a lateral gate opens in the side of the translocon allowing ~20 residues of the nascent polypeptide to sample the hydrophobic lipid bilayer (Fig. 1-4). The very hydrophobic stretches of each transmembrane helix signal the insertion into the bilayer. Other chaperone proteins surrounding the translocon also assist with the insertion and folding process.

The α-helical secondary structure of each transmembrane helix is stabilized by the hydrophobic environment. While the side chains of each amino acid contribute the large hydrophobicity of each transmembrane segment, the polar amide and carbonyl groups composing the peptide backbone must be satisfied by hydrogen bonding to each other to form the α-helix. The energetic cost of unfolding an inserted α-helix to a worm-like chain is greater than the insertion energy for the helix resulting in very stable secondary structure within the bilayer. The formation of alternative hydrogen-bonding interactions among the backbone allows flexibility in the helices.

### 1.3.2 Membrane Protein Topology

The directionality of α-helix insertion is another key component to the folding process of membrane proteins. Given the asymmetry on either side of the membrane plane, there are essentially two ways to place a transmembrane helix into the bilayer: the N-terminal end facing the cytoplasm or the N-terminal end facing the periplasm. Some proteins like ion channels, symporters, and anti-porters depend only on concentration gradients and function equivalently regardless of their inserted topology. For other proteins the final topology is absolutely essential. Pumps that utilize the
energy of ATP hydrolysis must have their binding domains in the cytoplasm while the translocon would be useless with its ribosome binding site on the outside of the cell.

Membrane proteins have their topology encoded in their amino acid sequence through the placement of positively-charged residues on the extra-membrane loops and termini (Fig. 1-5). Known as the positive-inside rule, this universal correlation states that positive-charges tend to be in the cytoplasm\textsuperscript{14}. The exact mechanism of the positive-inside rule is still unclear but many factors have been shown to contribute to membrane protein topology. In prokaryotes, proteins translation occurs through the translocon which exists across a proton gradient. Translocating positively-charged residues, like lysine and arginine, across this gradient therefore induces an energetic cost preferentially orienting the positive-charges in the membrane\textsuperscript{15}. The translocon is preserved across all domains of life but protein synthesis in eukaryotes occurs across a very different membrane of the endoplasmic reticulum (ER). Notably, the positive-inside rule is weaker for eukaryotes. The ER lacks a proton gradient, therefore the positive-inside rule must exert its force through another mechanism\textsuperscript{16}. Negative charges on the translocon and ribosome have also been shown to affect the topology of membrane proteins which could be essential for topology generation in ER. The negative charges are oriented towards the cytoplasm such that positive-charges on membrane protein loops and termini interacting with these negative charges would orient the entire transmembrane helix in a positive-inside fashion\textsuperscript{17}.}

Interesting to note, the next step in membrane protein biogenesis for eukaryotes, the golgi apparatus, does contain a pH gradient. As such, it is possible that topology acquisition for some proteins like
aquaporin I could continue into this organelle\textsuperscript{18}. Additionally, the charges on lipid headgroups have also been able to produce changes in membrane protein topology\textsuperscript{19}.

Beyond the positive-inside rule, the complicated process of translation from the ribosome into the translocon and insertion into the bilayer contains many small features that can also exert effects on membrane protein topology. Since co-translational insertion is a kinetic process the time for helices to insert and orient are at the whims of many processes including variations translation speed\textsuperscript{20}, the pulling force of insertion\textsuperscript{21}, the initial insertion of the nascent chain into the translocon\textsuperscript{22}, the lack of insertion of a marginally hydrophobic transmembrane helix\textsuperscript{18}, previously inserted helices and chaperones interacting with the current helix, co-translational folding events\textsuperscript{23} as well as any other effect occurring during translation. In general, these effects will dominated by the distribution of positive charges but they can still be significant.

Membrane proteins also have the opportunity to be dual topology. Nearly all membrane proteins have a large positive-charge bias across their termini and loops to ensure insertion into a single topology. Without a bias, proteins like EmrE insert into the membrane in both topologies\textsuperscript{24}. Two opposite topology monomers then come together to form a dual topology dimer in the case of EmrE causing inverted symmetry. Inverted symmetry of transmembrane helices within membrane proteins is a common feature of anti-porters. Some have speculated that this inverted symmetry first evolved through dual topology proteins that underwent a genetic duplication and fusion event\textsuperscript{25}. Before the evolution of the translocon, extremely hydrophobic dual topology proteins would need to insert into the bilayer without assistance. In this case, EmrE could be viewed as an example of an extremely ancient protein.
In the presentation of the two-stage model, Popot and Engeleman offered the assumption of static membrane topology. After insertion the hydrophilic loops and termini would have to cross the kinetic barrier of the membrane in order for topology changes to occur. With increasing numbers of helices, increasing numbers of extra-membrane loops would need traverse the membrane in at least a mildly cooperative fashion due to the covalent connection between transmembrane helices making topology changes appear more infeasible. Calculations have presented a cost of 60 kJ/mol for the positively-charged amino acid arginine to enter the lipid bilayer agreeing with the notion that topology changes would be difficult. An experimental proxy for topology changes would be the process of lipid flip-flop. Lipids can move from one leaflet of the bilayer to the other with a half-life of hours to days in physiological bilayers. Extrapolating to the multiple charges and multiple loops of most membrane proteins, topology changes would seem out of reach. In spite of the evidence against topology changes, data has been mounting for their occurrence. A further background and my own experiments are presented in Chapters 3 and 4.

1.4 Membrane Protein Folding: Stage 2

1.4.1 The Denatured State Ensemble

The denatured state ensemble is the collection of all the non-functional states (mostly random coil) a protein can adopt. The number of states a protein with 300 amino acids can adopt is beyond astronomical at $10^{143}$ confirmations compared to the estimated $10^{82}$ atoms in the observable universe. If a protein were to sample each confirmation for the native state randomly, the search would take far longer than the life
of the universe. This Levinthal paradox resolves with the answer that the search is not random.

1.4.2 The Native State

In the final stage of folding, membrane proteins achieve their native state. While this state is the crowning achievement of protein folding, the definition of the native state can be rather dubious. A general starting point is the 3-dimensional structure that spontaneously forms as the polypeptide explores the energy landscape. The folding funnel in Figure 1-6, a common visualization for the folding process, shows how a protein “falls” down the energy landscape to the lowest energy state, a familiar folded domain. In physiological conditions, the spontaneous structure evolved to be the functional structure. In other conditions like high temperature, the polypeptide will generally form a random coil due to the effect of entropy. In this case, the folding funnel would be reversed from Figure 1-6 with the denoted native state at the highest energy and the denatured state ensemble occupying the lower energy region. Adding to this difficulty, defining the native state is more complicated than simply identifying the lowest energy state in an appropriate condition. Often times the native state is not the lowest energy-state of the protein. This honor usually belongs to amorphous aggregates and amyloid fibers. A more accurate folding landscape is shown in Figure 1-7. The question then remains, if the native state is not the lowest energy state how does the protein find the native state (or physiological functional state)? The search from the extremely large denatured state ensemble tends towards the characteristic and functional native state in physiological conditions, but many times chaperones are required. Some proteins can be refolded in simpler laboratory conditions, but this is
almost always not the case for membrane proteins. A deeper understanding of the folding process is required before we can reliably manipulate membrane proteins in the lab.

1.4.3 The Transition State

The simplest form of folding involves a single kinetic transition from the denatured state to the native state without any long-lived intermediates. Proteins that undergo this type of folding pathway are known as two-state folders. Shown in Figure 1-8, the denatured protein must move over an energy barrier to get to the native state. At the highest energy point on the barrier, the protein adopts its transition state structure. In the simplest case, the transition state exists as an amalgam of disordered regions and contacts that are made in the native protein structure. A good first step towards understanding the process of protein folding is determining the structure of the transition state. Two key problems exist for studying folding in membrane proteins: most membrane proteins do not refold reversibly and using denaturant changes the lipid environment significantly unlike soluble protein folding studies. Despite these challenges, we have been able to map the transition state structure of the infamous membrane protein bacteriorhodopsin using the \( \phi \)-value analysis method described in Chapter 2.
Figure 1-1 (A) Chemical structure of the phosphatidylcholine lipid. (B) Organization of the lipid bilayer.
Figure 1-2 (A) Single transmembrane helix spanning the bilayer (B) α-helices in their final folded form.
Figure 1-3  On the left, a simple membrane protein bound to the lipid bilayer in a high energy configuration. On the right, the membrane protein after insertion. The polar extra-membrane loop must cross the lipid bilayer to achieve this confirmation.
Figure 1-4 The translocon (purple) catalyzes the insertion of a complex membrane protein during translation from the ribosome (brown).
The positive-inside rule states the positive-charges tend to be in the cytoplasm. Given the same sequence the distribution of positive-charges control topology. More positive charges on the extra-membrane loop give the topology on left. More positive charges on the termini result in the topology on the right.
Figure 1-6 Folding funnel of a two-state folding protein.
Figure 1-7 Folding funnel of protein with considering many possible states.
**Figure 1-8** Folding energy diagram of a two-state folding protein. A single kinetic transition can be measured.
References


   Rapid flip-flop in polyunsaturated (docosahexaenoate) phospholipid membranes. 


Chapter 2:

Bacteriorhodopsin Folds through a Poorly Organized Transition State
2.1 Abstract

The folding mechanisms of helical membrane proteins remain largely uncharted. Here we characterize the kinetics of bacteriorhodopsin folding and employ φ-value analysis to explore the folding transition state. First, we developed and confirmed a kinetic model that allowed us to assess the rate of folding from SDS-denatured bacteriorhopsin (bRu) and provides accurate thermodynamic information even under influence of retinal hydrolysis. Next, we obtained reliable φ-values for 16 mutants of bacteriorhodopsin with good coverage across the protein. Every φ-value was less than 0.4, indicating the transition state is not uniquely structured. We suggest that the transition state is a loosely organized ensemble of conformations.

2.2 Introduction

To understand membrane protein folding, it is important to explore the nature of the transition state in the folding process. Φ-value analysis discerns the structure of the transition state by determining whether individual residues have their native contacts or not in the transition state. Destabilizing point mutations are introduced by deleting interactions made in the folded state and the φ-value measures what fraction of the total destabilization also occurs in the transition state. When φ = 1, the full destabilization seen in the folded state is also realized in the transition state, suggesting that the side chain makes the same interactions in the transition state as in the native state. In other words, that part of the protein is structured in the transition state. Alternatively, when φ = 0, it implies that the mutated side chain has no effect on transition-state stability and
therefore does not have native-like contacts in the transition state. By measuring $\phi$-values for many residues, it is possible to map the transition state structure.

Although $\Phi$-value analysis has been used extensively to study the folding of small water-soluble proteins$^{2-8}$, the applications to membrane proteins have been limited. The Radford and Brockwell groups obtained extensive $\phi$-values for the coupled folding/insertion of the $\beta$-barrel outer membrane protein PagP$^9$. Their results suggest that the protein inserts via a partially folded structure that is tilted sideways with respect to the membrane normal. Otzen obtained $\varphi$-values for the helical membrane protein DsbB, suggesting the early formation of structure near the tips of several transmembrane helices may nucleate a wave of structure formation$^{10}$. Previously, the mechanism of bR folding has been studied with unfolded bacterioopsin (bO) as the initial state$^{11-14}$. Folding from unfolded bO involves formation of multiple intermediates and binding of retinal, which makes the kinetics unsuitable for $\varphi$-value analysis of a single transition state. However, refolding from unfolded bR (bR$_U$), which still retains retinal, occurs with a single dominant rate-limiting step and allows investigation of the transition state by $\varphi$-value analysis. Using this approach, the Booth group obtained $\varphi$-values for two helices in bR$^{15,16}$. These results indicated that helix B is structured in the transition state, forming a nucleus for consolidation of the rest of the protein structure.

Subsequent to the initial $\varphi$-value analyses of bR folding, we found that the kinetics of bR folding was significantly altered by bulk solution concentrations of detergent$^7$, which would likely confound $\varphi$-value measurements. For folding studies of bR, the protein is commonly solubilized in mixed micelles composed of a phospholipid (DMPC), a non-denaturing detergent (CHAPSO or CHAPS) and various concentrations
of the denaturing detergent SDS to alter the stability of the protein. The native bR (bR\textsubscript{F}) unfolds to SDS-denatured bR (bR\textsubscript{U}) in a sigmoidal transition as the concentration of SDS is increased. Similarly, bR\textsubscript{U} refolds to bR\textsubscript{F} as the SDS concentration is decreased. Since the protein folds within a micelle environment, it is typical to express SDS concentration as the mole fraction of total detergent (X\textsubscript{SDS}). The X\textsubscript{SDS} can be readily adjusted to unfold the protein by adding more total SDS or to refold the protein by adding more total DMPC/CHAPSO or DMPC/CHAPS. When folding and unfolding kinetic measurements are performed in this way the bulk detergent concentration changes, however. Much to our surprise, we found that a higher DMPC/CHAPSO concentration increases the folding rate of bR by an unknown mechanism\textsuperscript{17}. Thus, if one refolds bR by dilution from high X\textsubscript{SDS} into low X\textsubscript{SDS} by simply adding a high concentration of DMPC/CHAPSO, the folding rate will increase as X\textsubscript{SDS} is lowered by both the increased stability of the folded state and also the higher detergent concentration. As a result, the sensitivity of the rate constants with X\textsubscript{SDS} will be a complex function of both folding and detergent properties.

Given the unexpected discovery of the detergent dependence on folding rates\textsuperscript{17}, we re-measured a few key \(\varphi\)-values holding the total DMPC/CHAPSO concentration constant. Our initial results deviated substantially from the prior measurements. We therefore repeated and further extended the \(\varphi\)-value analysis to obtain a more comprehensive view of the transition state for bR folding.
2.3 Results and Discussion

*Model for bR folding kinetics*

Increasing the mole fraction of SDS results in the conversion of bRF to a bRU state, which we refer to as unfolding. In the bRU state, tertiary structure is lost, but most of the secondary structure remains\(^\text{18}\). The characteristic absorbance of the retinal cofactor in native bR serves as a convenient probe to follow this transition in the protein. Yet the cofactor can also complicate efforts to investigate the folding of this protein. When bRF is unfolded to bRU, the retinal Schiff base is susceptible to hydrolysis, which results in the formation of bacterioopsin (bO). Therefore, the complete reaction of bR unfolding occurs as shown in the following scheme:

\[
\begin{align*}
\text{bRF} & \rightleftharpoons \text{bRU} \rightarrow \text{bO + retinal,} \\
\text{Scheme 1}
\end{align*}
\]

where \(k_f\), \(k_u\), and \(k_h\) are the rate constants for folding, unfolding, and hydrolysis, respectively. Retinal hydrolysis is reversible but can be treated as irreversible for simplicity in the range of \(X_{SDS}\) where the equilibrium between bRF and bRU is investigated \(^\text{17}\). In this kinetic scheme, the equilibrium between bRF and bRU can be achieved experimentally only when the conformational relaxation is much faster than retinal hydrolysis \((k_u + k_f \gg k_h)\). We previously demonstrated that bR folding is significantly faster in 29 mM DMPC and 31 mM CHAPSO than in 15 mM DMPC and 16
mM CHAPSO (or CHAPS),\textsuperscript{17} which is the typical condition previously employed for the investigation of folding energetics of bR.\textsuperscript{14–16,19,20} Faster folding of bR in 29 mM DMPC and 31 mM CHAPSO is beneficial in suppressing the influence of the retinal hydrolysis on the folding kinetics. However, even in 29 mM DMPC and 31 mM CHAPSO, retinal hydrolysis still interferes with the folding and unfolding of bR in the transition zone. Therefore, the hydrolysis reaction must be accounted for as shown in Scheme 1 in order to determine the rate constants for folding and unfolding.

When the folding and unfolding of bR occurs as shown in Scheme 1, the concentration of $\text{bR}_F$ is expressed as function of time in a double-exponential equation\textsuperscript{21}:

$$[\text{bR}_F] = a_1e^{-\lambda_1t} + a_2e^{-\lambda_2t}$$

(1)

where $\lambda_1$ and $\lambda_2$ are the macroscopic rate constants and $a_1$ and $a_2$ are the amplitudes for the fast phase and the slow phase, respectively. According to the kinetic model in Scheme 1, $\lambda_1$ and $\lambda_2$ are the solutions of the following quadratic equation (see Supporting Information)\textsuperscript{21}:

$$\lambda^2 - (k_f + k_u + k_h)\lambda + k_u k_h = 0$$

(2)

Whether folding or unfolding is monitored, the change in $\text{bR}_F$ concentration is described with the same rate constants, $\lambda_1$ and $\lambda_2$. According to the initial conditions, $a_1 + a_2$ is
zero for a folding reaction and the initial total bRF concentration for an unfolding reaction.

Based on the property of a quadratic equation, we can extract the relationship between the macroscopic rate constants, \( \lambda_1 \) and \( \lambda_2 \), and the three elementary rate constants for folding, unfolding, and hydrolysis from Eq. 2 as follows:

\[
\begin{align*}
\lambda_1 \lambda_2 &= k_u k_h \\
\lambda_1 + \lambda_2 &= k_f + k_u + k_h
\end{align*}
\]  

(3)  

(4)

From these identities, the rate constants for folding and unfolding under a given condition can be extracted from the experimentally determined values of \( \lambda_1 \), \( \lambda_2 \), and \( k_h \). The relative amplitudes of the two phases observed in folding and unfolding reactions (\( a_1 \) and \( a_2 \) in Eq. 1) are also dependent on \( k_f \) and \( k_u \), but we did not use the amplitudes in determining \( k_f \) and \( k_u \) due to the limited accuracy with which the amplitudes could be determined.

**Validation of the kinetic model**

To assess whether the observed relaxation kinetics are consistent with the kinetic model (Scheme 1), we first investigated the folding and unfolding of bR in 29 mM DMPC and 31 mM CHAPSO at \( X_{SDS} = 0.73 \). At this \( X_{SDS} \), bRF shifts instantaneously to a previously described 600-nm “blue form” before slowly unfolding to the SDS-denatured state at 440-nm\(^{14} \). The rapid shift does not interfere with our kinetic measurements that are taken over the course of two hours. The 600-nm “blue form” likely results from the
protonation of D85 and the isomerization of the retinal cofactor\(^{14}\) which only negligibly alters both the structure and energetics compared to the unfolding of bacteriorhodopsin. Data collected at both 560-nm and 600-nm gives indistinguishable kinetic rates.

As shown in Fig. 2-1A, refolding features an initial increase in \(A_{600}\) followed by a slow decay driven by retinal hydrolysis as expected from our kinetic model. Unfolding also features two kinetic phases, which is obvious from non-random curve fitting residuals for a single exponential fit (Fig. 2-2). The time scale of the slow phase is similar to the time scale of retinal hydrolysis \((t_{1/2} \sim 10\ \text{min})\),\(^{14,17}\) suggesting that the biphasic kinetics result from retinal hydrolysis, consistent with the model in Scheme 1.

Based on this kinetic model, we determined \(\lambda_1\) and \(\lambda_2\) of wild-type bR at varying \(X_{\text{SDS}}\) using both refolding and unfolding reactions \((X_{\text{SDS}} = 0.67 - 0.76)\) (Fig. 2-1B). When plotted against \(X_{\text{SDS}}\), the natural logarithm of the observed rates constants \((\ln\lambda_1\) and \(\ln\lambda_2)\) make a continuous trend throughout the entire range of \(X_{\text{SDS}}\). As predicted by Eq. 2 (see Supporting Information), \(\lambda_1\) is greater than \(k_h\), and \(\lambda_2\) is less than \(k_h\). The kinetic model predicts that the \(\lambda_1\) and \(\lambda_2\) values observed in the folding reaction are identical to those observed in the unfolding reaction at the same \(X_{\text{SDS}}\). At \(X_{\text{SDS}} = 0.73\), where kinetics in both directions is accessible, we found \(\lambda_1\) and \(\lambda_2\) to be \((4.5 \pm 0.3) \times 10^{-3}\ \text{s}^{-1}\) and \((3.7 \pm 0.8) \times 10^{-4}\ \text{s}^{-1}\), respectively, when measured in the folding direction and \((3.5 \pm 0.3) \times 10^{-3}\ \text{s}^{-1}\) and \((3.7 \pm 0.2) \times 10^{-4}\ \text{s}^{-1}\), respectively, when measured in the unfolding direction (Fig. 2-1B). The similarity in the kinetic constants is consistent with the model of folding under the influence of retinal hydrolysis.

Using Eqs. 3 and 4, we calculated \(k_f\) and \(k_u\) values from the experimentally determined \(\lambda_1\) and \(\lambda_2\) within the transition zone from the chevron plots shown in Fig. 2-
1B. We determined $k_h$ experimentally to be $(8.3 \pm 0.4) \times 10^{-4}$ s$^{-1}$ by monitoring free retinal release from bR$_U$ at $X_{SDS} = 0.83$ (Fig. 2-3). As seen in Fig. 2-1C, the $\ln k_i$ and $\ln k_u$ values are, within experimental error, linearly dependent on $X_{SDS}$ in the transition zone of the equilibrium unfolding curves ($X_{SDS} = 0.67 - 0.76$). The values for $k_i$ obtained from unfolding kinetic traces are particularly unreliable because the error in the $k_i$ measurement originates from both the error in $k_u$ and in $k_h$. Consequently, when $k_i$ is small, the absolute errors in both the larger $k_u$ and $k_h$ values overwhelm the measurement.

The kinetic model predicts that the kinetic $C_m$ value, where the $k_i$ and $k_u$ are equal, should be the same when the values are obtained from the folding reaction or the refolding reaction. Indeed, the kinetic $C_m$ value for wild-type bR (the intersection of the linear fits of $\ln k_i$ and $\ln k_u$) is 0.73 in $X_{SDS}$ when the values are obtained from refolding experiments and 0.73 in $X_{SDS}$ when the parameters are obtained by unfolding experiments. Overall, the kinetic scheme fits the observed behavior well.

*Effect of mutations on kinetics of bR folding*

To survey the effect of mutations on the folding kinetics of bR, we tested 30 mutations of buried residues. For each mutant, we determined kinetic parameters at a range of $X_{SDS}$ around the transition zone. The dependence of $\ln k_u$ and $\ln k_f$ were generally well described by linear fits. The mutations L66A, F71A, D115A, R134A, W138A, S141A, M145A, and L149A did not destabilize bR more than 0.5 kcal/mol according to our kinetic measurements and therefore were not suitable for $\phi$-value analysis. Other
mutations appeared to alter the folding pathway of bR precluding the application of $\varphi$-value analysis. For example, the $m_{\text{t-u}}$ for the mutants I119A and L93A, T47A were near zero (Fig. 2-4). Such a drastic change in the response to denaturant implies an altered folding pathway. Mutants L94A, W86F, I148A, W189F, and F219A displayed altered kinetic traces as compared to the wild-type bR (Fig. 2-5). Specifically, our kinetic model predicts that the irreversible hydrolysis of retinal in the transition zone of $bR_F$ to $bR_U$ will eventually convert all the protein to $bO$, but these mutants had a significant residual absorbance that lasted for many hours, perhaps reflecting residual structure in the unfolded state. In the end, we chose sixteen bR mutants that we considered suitable for $\varphi$-value analysis. An example (L152A) is shown in Fig. 2-1C, and all the chevron plots are present in Fig. 2-6. The results from the kinetic analysis of the selected mutants are summarized in Table 2-1.

**Tanford $\beta$ value**

The slopes of the plots of $\ln k_f (m_{\text{t-u}})$ and $\ln k_u (m_{\text{t-n}})$ versus $X_{\text{SDS}}$ are measures of the protein’s sensitivity to denaturant, and reflect how close in their interaction with SDS the folding transition state is to the folded or unfolded state on the reaction coordinate. For wild-type bR $m_{\text{t-u}} = -12.8 \pm 0.7$ kcal/mol and $m_{\text{t-n}} = 18.9 \pm 0.7$ kcal/mol. The slope in the unfolding direction is larger than the slope in the folding direction, indicating that the transition state is closer to the unfolded state. This is generally quantified by the Tanford $\beta$-value:
\[ \beta = \frac{m_{\text{f-u}}}{(m_{\text{f-u}} - m_{\text{f-n}})} \]

We obtain a \( \beta \)-value of 0.4, suggesting that the transition state for unfolding is a little more than halfway to the fully unfolded protein. This value contrasts with the previously determined \( \beta \)-value of 0.1 \(^{15}\), most likely because the slopes were previously altered by the changing detergent concentrations. The previous study also reported that \( \beta \)-values of bR mutants correlate with the change in the energy of the transition state (\( \Delta \Delta G_{\text{f-u}} \)), indicating the movement of the transition state along the reaction coordinate upon destabilization of the protein \(^{16}\). However, we do not observe any significant correlation either between the \( \beta \)-values and \( \Delta \Delta G_{\text{f-u}} \) or between the kinetic \( m \)-values and \( \Delta \Delta G_{\text{unf}} \) (Fig. 2-7). The lack of the dependence suggests that destabilization of the native structure by mutation does not alter the degree of the interaction between the transition state and SDS relative to bRF and bRU.

**Correspondence of kinetic and equilibrium measurements**

The ratio of the rate constants \( k_u/k_i \) corresponds to the equilibrium constant for the bRF to bRU transition, \( K_{\text{unf}}(\text{bR}) \). To further validate the model, it would be useful to compare the equilibrium constant calculated with the rate constants with the value obtained from equilibrium measurements. Direct comparison of kinetically inferred equilibrium constants to experimentally determined equilibrium constants is not possible, however, because the bRF to bRU transition is not fast enough to eliminate the interference from retinal hydrolysis \(^{22}\). Nevertheless, we can obtain equilibrium constant, \( K_{\text{unf}}(\text{bO}) \), for the
overall reaction$^{22}$: $\text{bR}_f \rightleftharpoons \text{bO} + \text{retinal}$. We therefore decided to compare the mutational effects on $K_{\text{unf}}(\text{bO})$ from equilibrium measurements with their effects on $K_{\text{unf}}(\text{bR})$ obtained from kinetics. Since linear free energy relationships do not appear to hold over a large range of $X_{\text{SDS}}$,$^{23}$ we wanted to minimize extrapolations. We therefore compare values at $X_{\text{SDS}} = 0.67$, near the lower edge of the transition zone for the wild-type protein. Since all the mutants in this work destabilize the protein, this is a point where we can generally either make a direct measurement or where only a short extrapolation is required.

For the mutants used in $\varphi$-value analysis, we compared $\Delta \Delta G_{\text{unf}}^\circ$ obtained by kinetics with those obtained from the equilibrium measurements. As shown in Fig. 2-8, we see a linear correlation with a slope of 0.97, very closed to ideal expected slope of 1.00. The close agreement despite a difference in reference state and a large shift in $C_m$ between the two methods further supports the validity of our kinetic model.

**$\Phi$-value Analysis of Bacteriorhodopsin**

With the kinetic model that can reliably determine $k_l$ and $k_u$ for bR, we set out to map the transition state of folding by $\varphi$-value analysis. $\Phi$-values are the ratio of the change in free energy of the transition state upon mutation, $\Delta \Delta G_{\text{t-U}}$, (derived from $k_l$) to $\Delta \Delta G_{\text{unf}}^\circ$ (here derived from $k_l$ and $k_u$). The ideal mutation for $\varphi$-value analysis destabilizes the protein by at least 1 kcal/mol to provide sufficient signal to noise$^{24}$, without introducing any new interactions or altering the kinetic behavior$^1$. The mutants for which we have determined $\varphi$-values (Table 1) were selected based on these criteria. Also, for $\varphi$-value
determination we used kinetic constants at $X_{\text{SDS}} = 0.67$, which either required no extrapolation or only small extrapolations from the experimentally observable range of $X_{\text{SDS}}$.

The 16 reliable $\phi$-values provide good coverage across the protein (Fig. 2-9). Including the residues that make interhelical contacts with the mutated residues, this analysis covers the structure of bR extensively. The mutations that are excluded from $\phi$-values does not show any apparent pattern in their distribution on the structure (Fig. 2-9B), which indicates that our survey is not biased to a specific region by their exclusion. Every $\phi$-value measured was low. In particular, 14 of the 16 $\phi$-values measured were at or below 0.3 and the other two (F27A and E204A) were $\sim0.4$.

The globally distributed low $\phi$-values indicate that development of a localized folding nucleus in the transition state is not likely. Even the two residues with relatively higher $\phi$-values are distant from each other in the native structure, unlikely to coalesce in a folding nucleus. Considering most of the mutated residues contain extensive interhelical contacts in the native structure, we can picture that interhelical packing is minimal in the transition state. These results suggest that there is not a high degree of unique structure in the transition state.

*Mechanism of folding from SDS-denatured bacteriorhodopsin.*

The kinetics of refolding from the bO state has been investigated in the pioneering spectroscopic studies of the Booth group$^{12,13,25,26}$ and more recently by elegant studies by the Lanyi group$^{27}$ using EPR methods and the Konermann group using pulsed
oxidation or H/D exchange reactions\textsuperscript{28,29}. All see a series of intermediate states prior to the slow incorporation of retinal. In our case, however, we are examining folding from the bR\textsubscript{U} state for which the retinal is already covalently attached. Folding from bR\textsubscript{U} is less complex and is well described by a two-state model\textsuperscript{14,22} as in Scheme 1.

The structure of bR\textsubscript{U} in SDS has been investigated by the Lanyi and Langen labs using DEER spectroscopy\textsuperscript{18}. Their results indicate that the transmembrane helix structure remains, but is heterogeneously frayed, consistent with the \textasciitilde40\% reduction in helical content seen by far-UV CD spectroscopy\textsuperscript{11}. Results from the Konermann group using pulsed H/D exchange are also consistent with fluctuating helical secondary structure\textsuperscript{29}. In contrast, the tertiary structure is largely lost in the SDS unfolded state. Thus, folding of bR from an SDS denatured state (bR\textsubscript{U}) involves the assembly of largely preformed transmembrane helices as envisioned in the two-stage model of helical membrane protein folding\textsuperscript{30}. But how does this assembly occur?

Globally distributed low $\varphi$-values indicate that bR does not have significant interhelical packing in the transition state. Based on our results, we envision an expanded, poorly-ordered transition state in which partially folded, fluctuating transmembrane helices have roughly correct orientation/topology and are ready to coalesce in a cooperative fashion. This model is similar to the topomer search model proposed for soluble proteins\textsuperscript{31}. While achieving an approximately topologically correct structure at random without any energetic guidance from native contacts is somewhat improbable, finding this state would be facilitated by the short loops of bR and the largely preformed helices. If achieving a topologically correct structure is the rate-limiting step, the energetic barrier to reach the transition state would be largely entropic.
Reducing the search space significantly, the short loops and the preformed helices may decrease the entropic cost to achieve the topologically correct transition state.

Another structural change we may envision to occur in the transition state is shedding of SDS molecules. The β-value of 0.4 indicates that interaction with SDS in the transition state is somewhat distinct from that in the unfolded state. In bR\textsubscript{U} state, SDS is believed to interact with individual helices extensively. The lack of native-like contacts in the transition state suggests that some SDS may still remain between helices in the transition state.

The low fractional φ-values may also result from a number of relatively iso-energetic pathways. In this case, the transition state is an ensemble of structures that contain native-like interhelical contacts at different locations. It is easy to envision a folding mechanism involving an initial collection of helices as the transition state followed by the rapid accrual of the additional helices. Such a mechanism could occur by many different pathways. To have similar low φ-values throughout the protein structure, however, each pathway needs to contribute equally to the overall folding of bR, which is somewhat hard to imagine considering cooperative nature of protein folding in general.

2.4 Conclusions
Refolding of bacteriorhodopsin from SDS-denatured bR\textsubscript{U} state offers a valuable opportunity to investigate folding of helical membrane in a great detail. Though retinal
hydrolysis complicates the kinetics, we demonstrate here that, using a proper kinetic model, we can reliably determine folding and unfolding kinetic parameters. Refolding to bRF followed by slow retinal hydrolysis allowed us to determine folding and unfolding kinetic constants simultaneously in an extended region near the transition zone (Fig. 2-1C), which is not possible in most proteins. We confirmed the validity of the kinetic model and the folding and unfolding kinetic constants from the model by comparing of $k_f$ and $k_u$ determined from unfolding reactions and those determined from folding reactions (Fig. 2-1A) and also by comparing $\Delta \Delta G_{\text{unf}}^\circ$ values determined from kinetics with those determined from equilibrium unfolding.

Our φ-value analysis of the bRU-to-bRF transition indicates that the structure of the transition state is poorly organized without any significant interhelical native contacts. This finding is distinct from the previous report of helix B being the folding nucleus of bR folding$^{16}$, demonstrating the importance of eliminating the effect of the change in detergent concentrations on folding rates. With the comprehensive φ-value analysis, bR is a rather unique membrane protein since we now have considerable information about the folded state structure$^{32}$, the unfolded state structure$^{18}$, and the transition state. Based on the known structure of bRU and our φ-value analysis, we propose that the rate-limiting step of bR refolding from SDS-denatured form is finding topologically correct helical arrangement from the ensemble of rapidly interconverting conformations with largely preformed helices. As the orientation of the transmembrane helices is limited in a bilayer, bR may not have this topology search problem under natural conditions. Thus, it is possible that the folding transition state is different under
native conditions, but investigation of that possibility will require developing new technology for studying the folding process in bilayers$^{23,33}$. 
2.5 Experimental Section

Materials

Wild-type and mutant bR’s were expressed and purified as previously described.\textsuperscript{34,35,36} 1,2-Dimyristoyl-sn-glycerol-3-phosphocholine (DMPC) was purchased from Avanti Polar Lipids (Alabaster, AL). 3\{[3-cholamidopropyl]dimethylammonio\}-2-hydroxy-1-propanesulfonate (CHAPSO) was purchased from Affymetrix (Maumee, OH). Bio-Xtra sodium dodecyl sulfate was purchased from Sigma Aldrich.

Refolding kinetics

The bR protein was equilibrated in 10 mM sodium phosphate buffer (pH 6.0) containing 15 mM DMPC and 16 mM CHAPSO at 25°C for at least an hour prior to refolding experiments. The protein was first unfolded by incubating 1.0 mg/ml bR in 10 mM sodium phosphate buffer (pH 6.0) containing 15 mM DMPC/16 mM CHAPSO and SDS at $X_{SDS} = 0.82$ at room temperature for three minutes. Refolding was then initiated by diluting 10-fold into a 10 mM sodium phosphate buffer (pH 6.0) containing varying concentrations of SDS, DMPC, and CHAPSO at 25°C. The final refolding reactions contained 29 mM DMPC, 31 mM CHAPSO and 0.10 mg/mL bR. The absorption at 600 nm was monitored on a Molecular Devices Spectra Max M5 plate reader for two hours. The rate constants for each refolding reaction were determined by fitting a plot of the $A_{600}$ versus time to Eq. 1. Because folding and unfolding rate constants are not sensitive
to small differences in $k_h$ (data not shown), we assumed $k_h$ in the mutants is the same as that of wild-type bR. Mutants Y83A and Y185A had altered $\lambda_{\text{max}}$ values of 520 nm and 530 nm, respectively, when solubilized in DMPC/CHAPSO. The absorption at 560 nm was used to monitor the refolding of these mutants.

*Unfolding kinetics*

bR was equilibrated in 10 mM sodium phosphate buffer (pH 6.0) containing 15 mM DMPC and 16 mM CHAPSO at 25°C for at least an hour prior to unfolding. Unfolding was initiated at 25°C by a 10-fold dilution into a 10 mM sodium phosphate buffer (pH 6.0) containing 30 mM DMPC, 32 mM CHAPSO, and varying concentrations of SDS. The final unfolding reactions contained 29 mM DMPC and 31 mM CHAPSO. The final protein concentration was 0.10 mg/mL. The absorption at 600 nm was monitored on a Molecular Devices Spectra Max M5 plate reader for two hours. Observed rate constants for each unfolding reaction were determined by fitting a plot of the $A_{600}$ versus time according to Eq. 1.

*Determination of consensus values for the observed rate constants in the transition zone*

To ensure accurate determination of $\lambda_1$ and $\lambda_2$ for bR, we measured the kinetic constants for both folding and unfolding reactions in triplicate at multiple $X_{\text{SDS}}$ concentrations near the $C_m$. We calculated $k_f$ and $k_u$ from $\lambda_1$, $\lambda_2$, and $k_h$ using Eqs. 3 and 4. For $k_h$, we use $(8.3 \pm 0.4) \times 10^{-4}$ s$^{-1}$, which we determined by monitoring free retinal
release from wild-type bRu at $X_{\text{SDS}} = 0.83$. We assumed mutations do not affect $k_h$. To validate this assumption, we investigated the kinetics of hydrolysis step of two mutants. The $k_h$ values of L11A bR and E204A bR were $(7.5 \pm 0.2) \times 10^{-4}$ s$^{-1}$ and $(7.7 \pm 0.6) \times 10^{-4}$ s$^{-1}$, respectively, which were not significantly different from that of wild-type bR, $(8.3 \pm 0.4) \times 10^{-4}$ s$^{-1}$. Moreover, the use of the individual $k_h$ values does not affect the $\phi$-values. The dependence of the natural logarithm of $k_i$ on the $X_{\text{SDS}} (m_{i\rightarrow u})$ and the natural logarithm of $k_u$ on the $X_{\text{SDS}} (m_{\text{f} \rightarrow t})$ were determined by fitting the linear portions of the natural logarithms of the consensus rate constants for folding and unfolding against the $X_{\text{SDS}}$ in the transition zone.

Calculation of phi Values

$\Phi$-values are the ratio of the free energy change of the transition state upon mutation ($\Delta \Delta G_{\text{f} \rightarrow \text{U}}$) and to the free energy change of unfolding upon mutation ($\Delta \Delta G_{\text{unf}}^\circ$).

$$\phi_F = \frac{\Delta \Delta G_{\text{f} \rightarrow \text{U}}} {\Delta \Delta G_{\text{unf}}^\circ}$$

$\Delta \Delta G_{\text{f} \rightarrow \text{U}}$ is calculated from the wild type and mutant folding rates as follows:

$$\Delta \Delta G_{\text{f} \rightarrow \text{U}} = RT \ln \frac{k_f \text{ wt}} {k_f \text{ mut}}$$  (5)

$\Delta \Delta G_{\text{unf}}^\circ$ is calculated from the wild type and mutant folding rates and unfolding rates as follows:
$$\Delta \Delta G_{unf} = RT \ln \frac{k_f \text{mut}}{k_u \text{mut}} - RT \ln \frac{k_f \text{wt}}{k_u \text{wt}}$$  \hspace{1cm} (6)$$

The $k_u$ and $k_f$ kinetic rates were calculated from the measured $\lambda_1$ and $\lambda_2$ values as described in the text. The value of $k_u$ and $k_f$ was interpolated or extrapolated to $X_{SDS} = 0.67$ using a weighted linear fit of $\ln(k_u)$ or $\ln(k_f)$ by the predict function in the R statistics program. These values were then used to calculate the $\phi$-value. The error for the $\phi$-value was propagated from the triplicate measurement of $\lambda_1$ and $\lambda_2$. The error for the interpolation or extrapolation to $X_{SDS} = 0.67$ used the 95% confidence interval for the predicted value.

**Unfolding Equilibrium measurements**

The bR$_{F}$-to-bO$_{U}$ unfolding equilibrium was measured and calculated as described previously$^{22}$ except the conditions were altered to match the conditions used for the refolding kinetics of bR. The final conditions were 0.1mg/mL bR, 30 mM DMPC, 32 mM CHAPSO, 10 mM sodium phosphate pH 6.0, 9.1 $\mu$M all-trans RET and varying SDS concentrations. After the samples were equilibrated in dark at room temperature for ~ 4 days in a 96-well UV-star micro-plate (Greiner Bio-One), the absorbance at 553 nm was measured by SpectraMax M5 plate reader (Molecular Devices). All measurements were done in triplicate.
Figure 2-1  Conformational relaxation kinetics of bR in the folding transition zone. (A) The refolding of bRu (black line) and unfolding of bRf (red line) monitored at $X_{\text{SDS}} = 0.73$ by the absorbance of the folded protein at 600 nm ($A_{600}$). (B) The natural logarithms of the observed rate constants of the fast ($\lambda_1$ circles) and slow ($\lambda_2$ squares) phases of conformational relaxation of wild-type against $X_{\text{SDS}}$. Rate constants from refolding are shown as filled symbols and rate constants from unfolding are shown as open symbols. The rate of retinal hydrolysis ($k_h$; black line) is plotted for reference. (C) The natural logarithms of the extracted rate constants for folding and for unfolding of wild-type (black) and L152A (blue) bR's are plotted against the $X_{\text{SDS}}$. Whether the values were obtained from unfolding or refolding is indicated in the figure.
Figure 2-2  Residuals from curve-fitting of bR unfolding (A) The residuals from fitting the unfolding of bRF at X_{SDS} = .73 with a double exponential (B) The residuals from fitting the unfolding of bRF at X_{SDS} = .73 with a single exponential
Figure 2-3  A single exponential fit (black) of the hydrolysis of retinal from bR_U as monitored by the increase at $A_{390}$ (blue squares). The best fit exponential shown below with the rate constant in seconds.

$$A_{390} = 0.11 - 0.045 e^{-0.0052t}$$
Figure 2-4  No correlation between the degree of destabilization and other kinetic parameters.
Figure 2-5  Change in the folding pathway of I119A as determined by a large shift in the m-value. The natural log of $k_f$ (filled blue squares) and $k_u$ (open blue squares) of the bR mutant I119A is plotted with the natural log of $k_f$ (filled black circles) and $k_u$ (open black circles) of WT bR for reference.
Figure 2-6  The change in folding pathway as determined by the change in the kinetic trace. The refolding trace of the bR mutant F219A (red dots) is plotted with WT bR (black dots) for reference.
The extracted rates constants of the bR mutants used for φ-value analysis. The natural log of $k_f$ (filled blue squares) and $k_u$ (open blue squares) of the bR mutants are plotted with the natural log of $k_f$ (filled black circles) and $k_u$ (open black circles) of WT bR for reference. (A) L13A (B) M20A (C) F27A (D) K41A (E) F42A (F) T46A (G) M60A (H) Y83A (I) L97A (J) L100A (K) L111A (L) L152A (M) F171A (N) L174A (O) Y185A (P) E204A
Figure 2-8   **Comparison of $\Delta \Delta G_{\text{unf}}$*(bR)* from kinetics to $\Delta \Delta G_{\text{unf}}$*(bO)* from equilibrium measurements.** The effect of mutation on the free energy of unfolding as measured via the kinetic model to the bR$_U$ reference state or obtained from equilibrium methods to the bO reference state. A weighted linear fit to the data is shown (slope = 0.97, intercept = 0.26 kcal/mol, $R^2 = 0.88$).
<table>
<thead>
<tr>
<th>$\varphi_T$</th>
<th>$\Delta G_{\text{int}}^{{\text{skin}}^1}$ (kcal/mol)</th>
<th>$\Delta G_{\text{int}}^{{\text{pos}}^2}$ (kcal/mol)</th>
<th>$m_{{\text{z-a}}}$ (kcal/mol)</th>
<th>$m_{{\text{z-f}}}$ (kcal/mol)</th>
<th>$m_{{\text{int}}^3}$ (kcal/mol)</th>
<th>$k_t^4$ ($\times 10^{-3}$ s$^{-1}$)</th>
<th>$k_w^4$ ($\times 10^{-3}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>0.23 ± 0.06</td>
<td>-1.8 ± 0.2</td>
<td>-13 ± 0.1</td>
<td>15.0 ± 0.7</td>
<td>18.9 ± 0.7</td>
<td>31.7 ± 0.9</td>
<td>7.8 ± 1.0</td>
</tr>
<tr>
<td>L13A</td>
<td>0.25 ± 0.06</td>
<td>-2.8 ± 0.5</td>
<td>-2.7 ± 0.1</td>
<td>-7.8 ± 0.4</td>
<td>19.2 ± 2.8</td>
<td>27.0 ± 2.8</td>
<td>2.4 ± 0.3</td>
</tr>
<tr>
<td>M20A</td>
<td>0.43 ± 0.07</td>
<td>-2.1 ± 0.2</td>
<td>-2.2 ± 0.2</td>
<td>-14.1 ± 0.7</td>
<td>17.7 ± 1.3</td>
<td>31.8 ± 1.4</td>
<td>1.7 ± 0.2</td>
</tr>
<tr>
<td>K41A</td>
<td>0.13 ± 0.12</td>
<td>-0.9 ± 0.2</td>
<td>-0.7 ± 0.1</td>
<td>-12.6 ± 1.2</td>
<td>25.1 ± 0.7</td>
<td>37.7 ± 1.3</td>
<td>6.4 ± 0.8</td>
</tr>
<tr>
<td>F42A</td>
<td>0.29 ± 0.10</td>
<td>-2.1 ± 0.4</td>
<td>-1.8 ± 0.1</td>
<td>-12.3 ± 1.8</td>
<td>14.4 ± 2.7</td>
<td>26.7 ± 3.3</td>
<td>2.8 ± 0.8</td>
</tr>
<tr>
<td>T46A</td>
<td>0.26 ± 0.09</td>
<td>-2.2 ± 0.5</td>
<td>-2.3 ± 0.1</td>
<td>-11.2 ± 0.7</td>
<td>12.6 ± 2.2</td>
<td>23.8 ± 2.2</td>
<td>3.0 ± 0.7</td>
</tr>
<tr>
<td>M60A</td>
<td>0.10 ± 0.14</td>
<td>-1.1 ± 0.3</td>
<td>-1.1 ± 0.1</td>
<td>-17.4 ± 2.1</td>
<td>15.6 ± 3.6</td>
<td>33.0 ± 4.2</td>
<td>6.6 ± 1.4</td>
</tr>
<tr>
<td>Y83A$^5$</td>
<td>0.25 ± 0.10</td>
<td>-1.7 ± 0.4</td>
<td>-0.8 ± 0.1</td>
<td>-14.3 ± 0.6</td>
<td>15.6 ± 1.9</td>
<td>29.9 ± 2.0</td>
<td>3.9 ± 0.8</td>
</tr>
<tr>
<td>L97A</td>
<td>0.26 ± 0.08</td>
<td>-2.9 ± 0.4</td>
<td>-2.4 ± 0.1</td>
<td>-11.2 ± 1.1</td>
<td>33.4 ± 1.3</td>
<td>44.6 ± 1.7</td>
<td>2.2 ± 0.7</td>
</tr>
<tr>
<td>L100A</td>
<td>0.20 ± 0.07</td>
<td>-3.2 ± 0.3</td>
<td>-2.9 ± 0.2</td>
<td>-9.9 ± 1.2</td>
<td>21.2 ± 1.1</td>
<td>31.1 ± 1.6</td>
<td>2.7 ± 0.9</td>
</tr>
<tr>
<td>L111A</td>
<td>0.25 ± 0.07</td>
<td>-2.0 ± 0.3</td>
<td>-1.8 ± 0.1</td>
<td>-10.8 ± 1.7</td>
<td>27.7 ± 1.9</td>
<td>38.4 ± 2.5</td>
<td>3.3 ± 0.6</td>
</tr>
<tr>
<td>L152A</td>
<td>0.21 ± 0.06</td>
<td>-2.1 ± 0.2</td>
<td>-1.6 ± 0.1</td>
<td>-10.4 ± 1.1</td>
<td>20.1 ± 0.9</td>
<td>31.2 ± 1.4</td>
<td>3.7 ± 0.5</td>
</tr>
<tr>
<td>F171A</td>
<td>0.23 ± 0.09</td>
<td>-1.1 ± 0.2</td>
<td>-0.9 ± 0.1</td>
<td>-8.7 ± 1.2</td>
<td>21.5 ± 1.5</td>
<td>30.2 ± 1.9</td>
<td>5.1 ± 0.4</td>
</tr>
<tr>
<td>L174A</td>
<td>0.28 ± 0.16</td>
<td>-1.8 ± 0.4</td>
<td>-1.8 ± 0.1</td>
<td>-12.8 ± 3.9</td>
<td>9.2 ± 3.2</td>
<td>22.0 ± 5.0</td>
<td>3.4 ± 1.4</td>
</tr>
<tr>
<td>Y185A$^6$</td>
<td>0.29 ± 0.08</td>
<td>-4.2 ± 0.5</td>
<td>-2.9 ± 0.1</td>
<td>-13.6 ± 1.4</td>
<td>23.1 ± 1.8</td>
<td>36.7 ± 2.2</td>
<td>1.0 ± 0.5</td>
</tr>
<tr>
<td>E204A</td>
<td>0.34 ± 0.08</td>
<td>-2.2 ± 0.3</td>
<td>-1.8 ± 0.1</td>
<td>-8.1 ± 0.6</td>
<td>15.8 ± 1.6</td>
<td>24.0 ± 1.7</td>
<td>2.2 ± 0.3</td>
</tr>
</tbody>
</table>

1$\Delta G_{\text{int}}^{{\text{skin}}^1}$ values were calculated from $k_t$ and $k_w$ at $X_{\text{SDS}} = 0.67$ using Eq. 6.
2$\Delta G_{\text{int}}^{{\text{pos}}^2}$ values are the stability of bR determined with bO as the reference state.
3$m_{{\text{z-a}}}$ = $m_{{\text{z-f}}}$ - $m_{{\text{z-a}}}$.
4Kinetic constants at $X_{\text{SDS}} = 0.67$.
5Mutation shifted $\lambda_{\text{max}}$ to 520nm in DMPC/CHAPSO.
6Mutation shifted $\lambda_{\text{max}}$ to 530nm in DMPC/CHAPSO.

Table 2-1  Bacteriorhodopsin $\varphi$-values and associated kinetic parameters
Figure 2-9  Φ-value map of bR.  (A) The side chains of the 16 positions where we obtained reliable φ-values are shown in purple on the structure of bR (PDB 1C3W32). The backbone trace is shown as a green ribbon. (B) The side chains of the positions that did not yield usable φ-values are shown in red, along with the side chains of positions that did yield useable φ-values in purple.  (C) The positions are shown on a secondary structure map below using the same color scheme in panels A and B.
2.6 References


(8) Zarrineafsar, A. Methods 2004, 34, 41.


Chapter 3:
Dual-Topology Insertion of a Dual-Topology Protein
3.1 Abstract

Some membrane transporters are dual topology dimers in which the subunits have inverted transmembrane topology. How a cell manages to generate equal populations of two opposite topologies from the same polypeptide chain remains unclear. For the dual-topology transporter EmrE, the evidence to date remains consistent with two extreme models. A post-translational model posits that topology remains malleable after synthesis and becomes fixed once the dimer forms. A second, co-translational model, posits that the protein inserts in both topologies in equal proportions. Here we show that while there is at least some limited topological malleability, the co-translational model likely dominates under normal circumstances.

3.2 Introduction

A dual topology protein presents a difficult problem for evolution to solve. How can the same polypeptide achieve equal populations of the two topologies? If one orientation is strongly preferred it will lead to many orphan subunits that don’t have an opposite topology partner. A possible solution to this problem would be for topology to remain malleable after synthesis so that topology is set only after an opposite topology partner is found, driven by the stability of the anti-parallel dimer. We will refer to this as post-translational dual topology generation (Fig. 3-1A). A number of studies have shown that membrane protein topology can be altered after synthesis. In particular, the Skach lab showed that aquaporin-1 can insert in a partially incorrect topology and then re-orient after synthesis and work from the Dowhan group revealed that altering the lipid
composition leads to a remarkable topology rearrangement of an entire domain of lactose permease$^{3,4}$.

The most well-studied dual topology transporter is the multidrug resistance protein EmrE$^{1,5,6}$. In one subunit the N- and C-termini are inside the cytoplasm (N$_{in}$/C$_{in}$) and in the other subunit, the N- and C-termini are outside (N$_{out}$/C$_{out}$)$^{7,8}$. The topology of helical membrane proteins can be predicted by the preference for positively-charged amino acids to reside in the cytoplasm, the so-called positive inside rule$^{9,10}$. As a dual topology protein, EmrE does not exhibit a positive-inside rule charge bias. By the strategic introduction of positive charges, however, it is possible to force subunits into either the N$_{in}$/C$_{in}$ topology or the N$_{out}$/C$_{out}$ topology$^{11}$. Subunits forced into the N$_{in}$/C$_{in}$ topology will not form active dimers with themselves, but will form active dimers with subunits forced into the N$_{out}$/C$_{out}$ topology.

Studying the dual topology protein EmrE, the von Heijne group showed that only subunits with N$_{out}$/C$_{out}$ topology would complement an EmrE construct with C-terminal positive charges$^{12}$. These results have been interpreted to imply that topology can remain undefined until after the entire protein is made, consistent with the post-translational model (Fig. 3-1B)$^{12-14}$. Yet these results can also be explained by a co-translational model in which the protein inserts in both topologies, but the C-terminal positive charges simply inactivate subunits initially inserted in the N$_{out}$ topology (Fig. 3-1B). Consequently, without any direct examination of the inserted topologies, the indirect complementation experiments do not demonstrate topological malleability and the mechanism of dual-topology generation remains in question.

The problem with distinguishing between the post-translational and co-
translational topology models is that the final state is the same: dual topology. What is
needed is a way to lock the N-terminus in its initially inserted topology. As described
below, we find that placement of a short hemagglutinin (HA) epitope tag at the N-
terminus of the protein (HA-EmrE, Fig. 3-2A) allows the expression of both topologies of
EmrE, but blocks subsequent topology rearrangement, allowing us to examine the co-
and post-translational topology generation models.

3.3 Results
HA-tagging of EmrE does not impair the ability to generate dual topology: We first
asked whether the HA-tag altered the functional behavior of the protein by attempting to
recapitulate the prior in vivo results from the von Heijne group which showed that C-
terminal positive charges would only generate active N_in/C_in subunits\textsuperscript{12}. As shown in
Figure 3-3A, HA-EmrE imparts strong resistance to ethidium bromide suggesting that
the HA epitope does not greatly impair the final topology generation or function of EmrE.
We then created a construct with a cluster of positive charges (KKKHHHHHHH) at the C-
terminus, HA-EmrE-C+, to mirror the construct used previously\textsuperscript{12} (Fig. 3-2A) and a
control construct, EmrE-C+, without the HA tag, leaving a wild-type N-terminus. We
then tested the ability of these constructs to complement subunits forced into either the
N_in/C_in topology (HA-EmrE\textsubscript{locked_in}) or the N_out/C_out topology (HA-EmrE\textsubscript{locked_out}) by
strategic placement of positive charges as described previously\textsuperscript{11}. EmrE-C+ and HA-
EmrE-C+ effectively complement the N_out/C_out subunit HA-EmrE\textsubscript{locked_out}, but not HA-
EmrE\textsubscript{locked_in} (Fig. 3-2B). Thus, EmrE-C+ and HA-EmrE-C+ produce active subunits, but
only in the N_in/C_in topology, consistent with prior experiments\textsuperscript{12}.
Rapid degradation of orphan EmrE subunits: We next sought to move beyond genetic complementation to directly map the topology of the constructs. We were initially stymied in these efforts, however, because both HA-EmrE-C+ and EmrE-C+ had very low expression levels in the absence of the appropriate N_out/C_out partner subunit. We therefore considered the possibility that orphan subunits might be rapidly proteolyzed. Fig. 3-4A shows the degradation of HA-EmrE-C+ over time in the presence of the incorrectly oriented subunit partner, HA-EmrE_{locked_in}, or in the presence of the correctly oriented subunit partner HA-EmrE_{locked_out}. When dimer formation is precluded, the HA-EmrE-C+ is degraded rapidly (half-life < 30 min), while in the presence of the appropriate partner, there is essential no detectable degradation after 120 min. These results suggest that the C-terminally charged constructs exist in an unstable form that is subject to proteolysis.

Blocking FtsH mediated proteolysis of orphan EmrE subunits: FtsH is an integral membrane protease that is known to preferentially degrade unstable membrane proteins\textsuperscript{15}, so we tested whether FtsH was involved in degrading HA-EmrE-C+. Fig. 3-4B compares the expression levels in a control strain (AR3289) and an FtsH-null strain (AR3291)\textsuperscript{16}. Expression is \~80-fold higher in the FtsH-null strain than in the control strain, indicating that FtsH is at least partially responsible for the degradation of HA-EmrE-C+. The FtsH-null strain became an important tool for mapping the topological variants of EmrE, because it allowed us to map the topology of even unstable forms that might otherwise be lost to degradation.

With the ability to prevent proteolysis, we sought to directly determine the
topology of EmrE-C+ constructs using cysteine-accessibility. Cysteine-accessibility maps the topology of a membrane protein by identifying whether a single introduced cysteine resides in the cytoplasm or the periplasm, by measuring the reactivity of the cysteine to a membrane-impermeable reagent\textsuperscript{17}. The procedure is outlined in Fig. 3-5. We first add the membrane impermeable reagent, 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid (AMS), and later a membrane permeable biotinylation reagent, 3-(N-Maleimidopropionyl)-biocytin (MPB). If AMS has already reacted with the lone cysteine (i.e., cysteine has a periplasmic location) then reaction of MPB will be blocked by AMS. Otherwise MPB will modify the cysteine. The extent of MPB modification can be detected by a gel-shift in the presence of avidin (seen as a loss of the gel band corresponding to free EmrE). To reliably assess the level of biotinylation without AMS, we also perform a reaction with MPB alone. To apply this method, we introduced unique cysteines (3C, F27C, Q81C and T108C) into a cysless background (C39A/C41A/C95A), denoted as EmrE\textsubscript{Cless}. The EmrE\textsubscript{Cless} variants were unreactive with MPB under our conditions and the intensity of the bands on western blots were linearly related to the amount of protein loaded (Fig. 3-6, 3-7). We utilized the FtsH-null strain in all cysteine accessibility topology experiments unless otherwise noted.

**The topology of EmrE-C+ is N\textsubscript{in}/C\textsubscript{in}:** To assess the topology of EmrE\textsubscript{Cless}-C+ with a wild-type N-terminus, we mapped the locations of F27C in the first loop and T108C on the C-terminus. As expected for an N\textsubscript{in}/C\textsubscript{in} topology, F27C is periplasmic as indicated by the large decrease in the avidin dependent gel shift upon prior reaction with AMS (78% ± 13% change compared to no AMS) (Fig. 3-8A). The C-terminally located T108C
showed no change in gel shift after reaction with AMS (10 % ± 14%) indicating a cytoplasmic location, as is expected for an N\textsubscript{in}/C\textsubscript{in} topology (Fig. 3-8A). These results indicate that the C-terminal positive charges do define the topology of the N-terminus as suggested previously based on complementation experiments\textsuperscript{12}.

**Does partial reactivity of F27C in HA-EmrE-C+ indicate a mixed topology?**

Continuing our characterization of the HA-tagged protein, we examined the topology of HA-EmrE\textsubscript{Cless}-C+ using the same F27C and T108C mutations. A cysteine introduced into the first loop (F27C) of HA-EmrE\textsubscript{Cless}-C+ showed only partial blocking (48% ± 13%) by AMS (Fig. 3-8B). In contrast the C-terminal cysteine, T108C, which is adjacent to the C-terminal positive charges showed no response to AMS (1% ± 4%) which indicates that it resides completely in the cytoplasm (Fig. 3-8B).

The partial reactivity of F27C in the HA construct led us to consider the possibility that the N-terminus normally inserts in two different topologies, but that the HA-tag blocks subsequent topology rearrangement, leading to a mixed topology form as illustrated in Figure 3-1B. We first had to evaluate other possible reasons for the partial reactivity, however.

**Partial reactivity is not due to an environmental constraint:** We considered the possibility that F27C is actually located fully in the periplasm in the HA-EmrE-C+ construct, but is only partially reactive to AMS because it is partially buried by protein or membrane (even though this does not occur in the EmrE-C+ construct missing the HA-tag). To test whether F27C in HA-EmrE\textsubscript{Cless}-C+ would be fully blocked by AMS if the
protein were in a pure N\text{in}/C\text{in} topology, we expressed the single cysteine variants of HA-EmrE\text{Cless}.C+ in a wild-type (FtsH+) strain in the presence of HA-EmrE\text{locked}_\text{out}. As illustrated above, when expressed in an FtsH+ strain, the N\text{in}/C\text{in} topology remains resistant to proteolysis when expressed in the presence of HA-EmrE\text{locked}_\text{out} because it can form stable dimers, but orphan, unpaired subunits, are rapidly degraded by FtsH. Thus, only the N\text{in}/C\text{in} topology of HA-EmrE\text{Cless}.C+ will be present, allowing us to test the cysteine’s reactivity in a pure N\text{in}/C\text{in} subunit in the active dimer. As shown in Figure 3-9A, when expressed with EmrE\text{locked}_\text{out}, F27C is fully blocked by AMS (85% ± 20%) whereas T108C remained inaccessible to AMS (5% ±6%). Moreover, Q81C, in the third loop was fully blocked by AMS (94% ± 12%), consistent with the N\text{in}/C\text{in} topology when expressed with HA-EmrE\text{locked}_\text{out}. These results suggest that we could have seen complete blocking of the F27C position if HA-EmrE\text{Cless}.C+ adopted a pure N\text{in}/C\text{in} topology.

It still remained possible that F27C in HA-EmrE\text{Cless}.C+ is more reactive to AMS in an active dimer compared to the orphan form in the absence of an N\text{out}/C\text{out} subunit partner. To test this possibility, we attempted to generate pure normal N\text{in}/C\text{in} and distorted N\text{out}/C\text{in} topology orphan subunits by controlling the topology of the N-terminal helix. To favor an N\text{out} orientation, we introduced a positive charge bias into loop 1 using the mutations G26R and T28R, to produce N\text{out}-HA-EmrE\text{Cless}.C+ (Fig. 3-9B). To favor an N\text{in} orientation, we placed positive charges at the N-terminus by changing the N-terminal sequence from MPNYIY to MRRRYIY, generating N\text{in}-HA-EmrE\text{Cless}.C+ (Fig. 3-9C). Because of the introduced charge biases, the N-terminal helix in these constructs should insert in unique orientations. Indeed for N\text{out}-HA-EmrE\text{Cless}.C+
construct, F27C and T108C no longer show any reaction to the AMS reagent (5% ± 15% and 0 ± 4%, respectively) indicating that F27C and T108C are cytoplasmic (Fig. 3-9B). Conversely, for N\textsubscript{in}-HA-EmrE\textsuperscript{Class}.C+, the F27C position is nearly fully blocked by AMS (78% ± 10%) and 3C is completely unblocked (6 ± 9%), consistent with an N\textsubscript{in} topology (Fig. 3-9C). Q81C shows a less complete AMS response (61 ± 7%), possibly due to an altered conformation in the monomeric subunit. Nevertheless, these results strongly suggest that the partial blockage observed for F27C in HA-EmrE\textsuperscript{Class}.C+ is not due to changes in reactivity of F27C.

We were unable to use cysteine accessibility to map the location of 3C in the HA-EmrE\textsuperscript{Class}.C+ and N\textsubscript{out}-HA-EmrE\textsuperscript{Class}.C+ constructs because of disulfide bond formation of 3C. As shown in Fig. 3-10, a band at a dimer molecular weight is seen in constructs where 3C is expected to reside in the periplasm, but not in constructs where 3C is expected to reside in the cytoplasm. The dimer band is eliminated after reduction by β-mercaptoethanol (β-ME). Although disulfide formation complicates cysteine accessibility analysis, formation of the disulfide bond is indicative of an oxidizing, periplasmic location. In N\textsubscript{in}-HA-EmrE\textsuperscript{Class}.C+ the 3C cysteine has robust reactivity with MPB and does not form disulfide bonds as would be expected for its cytoplasmic location (Fig. 3-10). Thus, the observation of disulfide formation is consistent with the predicted topologies.

**Partial reactivity is not due to topology flipping:** We finally considered the possibility that the HA-tagged N-terminus topology is not fixed in an \( \text{N}_{\text{in}} \) or \( \text{N}_{\text{out}} \) topology, but flips back and forth. If the N-terminal segment (including F27C) of HA-EmrE-C+ can move
from the cytoplasm to the periplasm, then the partial reactivity could be due its presence in the periplasm for only a fraction of the time. If so, the extent of F27C blocking should be time dependent. When we increased the AMS reaction time with HA-EmrE\textsuperscript{Cless.-C+} F27C, however, we saw no increase in AMS blocking. When blocked for 10 minutes or 20 minutes we observed AMS blocking of 48\% ± 13\% and 49\% ± 13\%, respectively (Fig. 3-11). These results are consistent with a static topology of the inserted HA-EmrE-C+, i.e., the protein is either N\textsubscript{in}/C\textsubscript{in} or N\textsubscript{out}/C\textsubscript{in} and does not flip after insertion.

### 3.4 Conclusion

Our results indicate that under normal circumstances EmrE inserts in two opposite topologies at roughly equal frequencies. In particular, with an HA-tag at the N-terminus that blocks subsequent topology flipping, we observe insertion in both an N\textsubscript{in} and an N\textsubscript{out} topology. As the initial insertion occurs in both topologies with roughly equal proportions, it suggests that the primary mechanism for EmrE dual-topology biogenesis is co-translational (Fig. 3-1A). While we cannot rule out global topology flipping in the wild-type protein as predicted by the post-translational mechanism, we only have direct evidence for more limited topology rearrangements when the protein is unnaturally inserted in a distorted topology. Indeed, to our knowledge our work is the first direct demonstration of topological malleability in EmrE since a static topology would produce the same results in the complementation assays used previously\textsuperscript{12} (see Fig. 3-1B). Moreover, we find that incorrectly inserted or unpaired subunits can be cleared rapidly by proteolysis so a mechanism for ensuring equal populations of both topologies by subsequent topology flipping is not necessary. When topological signals are
manipulated by the placement of positive charges at the C-terminus, causing initial 
insertion in a distorted N\textsubscript{out}/C\textsubscript{in} topology, the topology can indeed rearrange to an N\textsubscript{in}/C\textsubscript{in} 
topology as previously suggested by von Heinje and co-workers\textsuperscript{12}. It is fascinating and 
remarkable that even partial topological rearrangements are possible, although they are 
likely driven by distorted topologies that may only rarely occur naturally. Thus, we 
suggest that major topology rearrangements are unlikely to play a major role in dual 
topology generation with the normal EmrE topology signals.
3.5 Methods

Materials

4-acetamido-4’-maleimidylstilbene-2,2’-disulfonic acid (AMS) was purchased from Life Technologies (Grand Island, NY). 3-(N-Maleimidopropionyl)-biocytin (MPB) was purchased from Cayman Chemical (Ann Arbor, MI).

Plasmids and Constructs

The EmrE constructs were cloned into the pBAD/His A plasmid (Invitrogen) using the Ncol and Xhol cut sites, which removes the built in tags. For the co-expression experiments, the EmrE\textsubscript{locked\_out} and EmrE\textsubscript{locked\_in} constructs were cloned into another pBAD vector which contains a chloramphenicol resistance gene and a ClodF13-derived CDF replicon. This vector was created from the pSEL1 vector described previously\textsuperscript{18}. The CDF origin was removed from pCDFDuet-1 vector (Novagen) using the Xbal and Nhel restriction sites and was ligated into the pSEL1 vector replacing the p15A origin using the same restriction sites. EmrE constructs were cloned into the vector using a Ncol and HindIII site. A second Ncol site in the chloramphenicol resistant gene was removed by making a silent mutation of Thr172 from the ACC codon to the ACA codon using PCR quickchange mutagenesis.

HA-EmrE bears an N-terminal HA epitope (YPYDVPDYA) before the WT \textit{E.coli} EmrE sequence with an additional glycine downstream of the initial methionine. The amino acids KKKHHHHHHH were added onto the C-terminus for the constructs HA-EmrE-
C+ after a linker (ENLYFQG). The EmrE-C+ construct has the same protein sequence as HA-EmrE-C+ except with a wild-type N-terminus. A more detailed view of these constructs is given in Figure 3-12.

Surprisingly, HA-EmrE can be induced at full induction in the pBAD system without the cellular toxicity normally associated with the overexpression of EmrE\textsuperscript{11}. Expression of the EmrE-C+ construct lacking the HA-tag, was toxic in our pBAD system. We speculate that the reduced toxicity of HA-EmrE may be due to a predicted RNA hairpin accidentally included in the HA-tag that reduces expression (Fig. 3-12). To use EmrE-C+ in our system, we therefore engineered a predicted RNA hairpin into EmrE-C+ by altering the DNA but not the protein sequence. The DNA mutations A-7T (in the pBad His A plasmid) and C6T prevent the toxic overexpression of EmrE-C+ in our pBAD system (Fig. 3-12C, 3-12D). The N-terminal portion of the WT-EmrE construct contains the same DNA sequence as EmrE-C+.

All of the point mutations were introduced by quickchange PCR mutagenesis. The HA-EmrE\textsubscript{locked\_out} construct based on a previous work\textsuperscript{11} contains the N-terminal HA-tag of HA-EmrE to prevent toxicity and the mutations T28R, L85R and R106A. The HA-EmrE\textsubscript{locked\_in} construct has the same HA-tag to prevent toxicity and was also based on previous work\textsuperscript{11} but contains an additional mutation (M1K) to ensure complete topological locking in our system (M1K, R29G, R82S, S107K)\textsuperscript{11}.

In all constructs used with cysteine accessibility experiment (N\textsubscript{in}-HA-EmrE\textsubscript{Cless, C+}, N\textsubscript{out}-HA-EmrE\textsubscript{Cless, C+}, HA-EmrE\textsubscript{Cless, C+}, EmrE\textsubscript{Cless, C+}) the native cysteine residues were changed to alanine (C39A, C41A, C95A). The added cysteines F27C, Q81C and T108C were then singly introduced into each of the cysteine null constructs. The mutation 3C
adds the amino acids ‘CG’ after the initial two N-terminal residues, MG, which are upstream of the HA epitope. To form the $N_{\text{in}}$-HA-Emr$E^\text{Cless}$-C+ construct, the HA-Emr$E^\text{Cless}$-C+ construct had the mutations N2R, P3R and additional arginine inserted between the two residues. To form the $N_{\text{out}}$-HA-Emr$E^\text{Cless}$-C+ construct, the HA-Emr$E^\text{Cless}$-C+ construct had the mutations G26R and T28R incorporated into sequence.

**Ethidium bromide-resistance Assay**

The ethidium bromide resistance assay was performed in a similar manner as described previously. E. coli BL21Pro (Addgene/ClonTech) cells with the relevant plasmids were grown to saturation (~10 hours) and then serially diluted 10-fold six times. 5µL of each dilution was spotted onto plates with 0.2% arabinose, 34µg/mL chloramphenicol, 100µg/mL ampicillin, and the indicated concentration of EtBr. The plated cells were grown at 37°C for 18 hours.

**Cysteine-Accessibility Topology Assay**

AR3291 (FtsH null) cells containing the desired EmrE construct were grown to ~0.8 OD$_{600}$ and then induced with .2% arabinose at 30°C for 2 hours. Cells from 50 mL of cell culture were collected by centrifugation and suspended in 500µL of 50mM phosphate buffer with 17 mM NaCl at pH 8.0. 200µL of resuspended cells were incubated in a final concentration of 2 mM AMS for 10 minutes, rotating in the dark at room temperature. The cells were then washed twice with 50 mM Tris-Cl at pH 7.5. The final cell pellet was resuspended to a final volume of ~200uL. Samples not reacted with AMS underwent the
same procedure without the AMS reagent. The cells were further incubated with 4mM MPB in 4% DMSO, .5% toluene and 50 mM Tris-Cl pH 7.5 for 1 hour while rotating in the dark at room temperature. The cells were then lysed by sonication. Cell debris was removed by centrifugation at 16,000 g for 10min and then membranes were isolated by ultracentrifugation of the supernatant in a Beckman Coulter Airfuge at 160,000g for 1hr at room temperature. Membranes were resuspended in 20mM Tris-Cl pH 7.5. Total protein was determined by the DC Protein Assay (Bio-rad). For each lane, 24µg of protein was mixed with 4X SDS loading buffer. Either 15 µg of avidin (Sigma BioUltra) in 20mM Tris-Cl pH 7.5 or 20mM Tris-Cl pH 7.5 was added to the samples and the samples were run on a 12% NuPAGE Bis-Tris (Life Technologies) gel using MES running buffer. Electrophoresis was carried out at 40V for 25 minutes and then 100V for 135 minutes. The samples were then transferred to a (0.2 µm) PVDF membrane at 95mA for 1 hour using Towbin transfer buffer with 20% methanol.

For the HA-antibody, the membrane was blocked with 5% nonfat milk in TBS-T (20mM Tris-Cl pH 7.5, 150mM NaCl, .1% Tween 20) followed by an hour incubation in a 1:1,000 dilution of 1mg/mL monoclonal HA antibody (Sigma) in TBS-T with .5% nonfat milk. The blot was then washed four times in 20mM Tris-Cl pH 7.5, 500mM NaCl, 0.2% Triton X-100, and 0.05% Tween 20 buffer (TBS-Tween/Triton). The blot was incubated for one hour in a 1:3,000 dilution of anti-mouse IgG peroxidase conjugate (Sigma) in TBS-T with 0.5% nonfat milk and then washed six times for 10 minutes in TBS-Tween/Triton buffer. Western blots for the histidine tag were performed as described above except that the membrane was blocked and blotted with the supplied blocking buffer and the pentahis
HRP conjugate antibody (Qiagen) as described in the QIAexpress detection and assay handbook.

The blots were visualized on a FluorChem FC2 (Alpha Innotech) CCD imager using the Amersham ECL Prime detection reagent (GE Healthcare) according the recommended protocol. The intensity of the bands was quantified using ImageJ. A general background subtraction obtained from a dark area of the blot was applied to each band. The amount of biotinylation was determined by the decrease in band intensity between the lanes with and without avidin. The percent change in biotinylation between the reaction with MPB only and the reaction with both AMS and MPB (labeled AMS response) was used to determine the location of each cysteine. Each topology assay was performed in at least triplicate.

**CCL EmrE Degradation Assay**

BL21Pro *E. coli* cells co-expressing HA-EmrE<sup>Cless</sup>-C+ with either N<sub>out</sub>-HA-EmrE<sup>Cless</sup>-C+ or N<sub>in</sub>-HA-EmrE<sup>Cless</sup>-C+ were grown to ~0.8 OD<sub>600</sub> and induced with 0.2% arabinose for 1 hour at 30°C. The cells were washed twice with LB media and then switched into LB media containing 350 µg/mL erythromycin and 0.2% glucose to stop further protein synthesis. Samples were taken at time zero (after the media switch) and at the other indicated time points.

**FtsH null Expression Test**
AR3291 cells (FtsH null) or AR3289 (control cells) obtained from the Ogura Lab\textsuperscript{16} were grown to \(\sim 0.8 \text{ OD}_{600}\) and then induced with 0.2% arabinose for 2 hours at 30°C. The cells were collected by centrifugation and lysed by sonication. The membranes were isolated and the samples were visualized by western blot for the HA-epitope as described previously.

\textit{Disulfide Bond Determination}

AR3291 (FtsH null) cells containing the desired EmrE construct were grown to \(\sim 0.8 \text{ OD}_{600}\) and then induced with .2% arabinose at 30°C for 2 hours. Cells from 50 mL of cell culture were collected by centrifugation and suspended in 500µL of 50mM Tris-Cl buffer pH 7.5. 200µL of resuspended cells were incubated in a final concentration of 5 mM iodoacetamide for 15 minutes, rotating in the dark at room temperature. The cells were then washed twice with 50 mM Tris-Cl at pH 7.5 and lysed by sonication. Cell debris was removed by centrifugation at 16,000 g for 10min and then membranes isolated by ultracentrifugation of the supernatant in a Beckman Coulter Airfuge at 160,000g for 1hr at room temperature. Isolated membranes were resuspended in 20mM Tris-Cl pH 7.5. 24µg of the samples were then mixed in SDS loading buffer with 2-mercaptoethanol (\(\beta\text{ME}\)) at 170mM and without. The protein was visualized by western blot using the HA-epitope as described earlier.
Figure 3-1 Dual Topology Insertion Models (A) Post-translational Dual Topology Model: EmrE inserts into the membrane in a preferred topology and then changes topology to form the stable anti-parallel dimer. Co-translational Dual Topology Model: EmrE obtains dual topology by inserting equally into both topologies. Extra subunits are degraded. (B) Topology of EmrE with C-terminal positive-charges if the N-terminal helix has a malleable or static topology. Regardless of the mechanism, no active dimer can be formed with the N\textsubscript{in}/C\textsubscript{in} EmrE. The frustrated N\textsubscript{out}/C\textsubscript{in} topology shown is one of many possible topologies.
Figure 3-2 Indirect Topology Assay of EmrE-C+ and HA-EmrE-C+. (A) Diagram of the EmrE constructs used. (B) The growth of 10-fold dilutions of stationary phase cell culture expressing the HA-EmrE-C+ or EmrE-C+ constructs with a single topology EmrE mutant (HA-EmrE\textsubscript{locked\_in} or HA-EmrE\textsubscript{locked\_out}), spotted on agar plates in the presence of 190µg/mL ethidium bromide.
Figure 3-3 Ethidium Bromide Resistance of HA-EmrE. (A) The growth of 10-fold dilutions of stationary phase BL21Pro cell culture expressing the indicated constructs spotted on agar plates in the presence of 220µg/mL ethidium bromide. (B) The growth of 10-fold dilutions of stationary phase BL21Pro cell culture with the indicated constructs on an arabinose inducible plasmid spotted on agar plates. The induction of the constructs at .2% arabinose is not toxic.
Figure 3-4  FtsH Dependent Proteolysis of EmrE  (A) The loss of HA-EmrE-C+ co-expressed with either HA-EmrE_{locked\_in} or HA-EmrE_{locked\_out} after protein synthesis has been stopped at time zero. HA-EmrE –C+ was visualized by western blotting for the penta-histidine epitope. (B) The expression of HA-EmrE and HA-EmrE-C+ in wild-type and FtsH knockout cells. The protein was visualized by western blotting using the HA-epitope.
Figure 3-5 Cysteine-Accessibility Method. Illustration of our specific application of the cysteine accessibility topology determination method. We determine whether a single cysteine in EmrE resides in either the cytoplasm or periplasm by its reactivity with a membrane-impermeable maleimide, AMS. Since the reaction of AMS is not easily detectable, we assess the amount of AMS that is reacted with EmrE by its ability to block a second reaction with a biotin-maleimide (MPB). The amount of biotinylation is then quantified by the amount of protein that is gel-shifted upon the addition of avidin (detected as a loss of the gel band corresponding to free EmrE). For each trial, the level of biotinylation without the AMS pre-reaction is determined by a reaction with the biotin-maleimide alone. The percent difference in biotinylation between the biotin-maleimide reaction alone and the AMS/biotin maleimide series is termed “AMS response.” (A) In the case of a cytoplasmic cysteine, AMS cannot react with the cytoplasmic cysteine due to its inability to cross the membrane. The subsequently added biotin-maleimide can therefore react freely with the unblocked cysteine. Thus, for a cytoplasmic cysteine, we see near complete biotinylation with the biotin-maleimide alone (~95% of the band lost upon addition of avidin) and again near complete biotinylation when pre-reacted with AMS (~95% of the band lost upon addition of avidin). The low AMS response, 0%, is indicative of a cytoplasmic cysteine. (B) In the case of a periplasmic cysteine, AMS reacts with the lone cysteine. The biotin-maleimide added next cannot react with the protein as the cysteine has already reacted with AMS. The final result is near complete biotinylation with the biotin-maleimide alone (~95% of the band lost upon addition of avidin) and limited biotinylation when pre-reacted with AMS (~20% of the band lost upon addition of avidin). The high AMS response, 79% = 100*((95-20)/95), is indicative of a periplasmic cysteine.
Figure 3-6 Cysteine-Free Proteins are not Reactive. HA-EmrE<sup>Cless</sup> and HA-EmrE<sup>Cless-C+</sup> are not biotinylated by MPB as shown by the absence of a gel shift with avidin. The proteins were visualized by western blot with a HA-antibody.
Figure 3-7 Protein Amount is Linearly Related to Band Intensity. (A) A membrane preparation containing HA-EmrE-C+ was loaded into each well with the specified amount of total protein. The protein was visualized by western blot using the HA-epitope. (B) The band intensity from (A) plotted against the total protein loaded in the lane.
Figure 3-8 Direct Topology Assay of EmrE-C+ and HA-EmrE-C+. Proposed topologies are shown on the left and cysteine reactivity data are shown on the right. Added positively charged residues are shown in red. Invariant positively charged residues present in the wild-type protein are shown in light gray. AMS response is the percent change in biotinylation due to the reaction with AMS. (A) Results for the EmrE\textsubscript{Cless}-C+ constructs. The high AMS response of F27C is indicative of a periplasmic cysteine and the low AMS response of T108C is indicative of a cytoplasmic location. (B) Results for the HA-EmrE-C+ construct. While T108C shows a low AMS response similar to the EmrE\textsubscript{Cless}-C+ construct that does not possess and HA-tag, F27C shows much lower AMS response than seen in the EmrE\textsubscript{Cless}-C+ construct. EmrE\textsubscript{Cless}-C+ was visualized by western blotting for the penta-histidine epitope and HA-EmrE-C+ was visualized by western blotting for the HA epitope. The western blots shown are representative of assays performed in triplicate.
Figure 3-9 AMS Reactivity Tests for EmrE Constructs. Proposed topologies are shown on the left and cysteine reactivity data are shown on the right. Added positively charged residues are shown in red. Invariant positively charged residues present in the wild-type protein are shown in light gray. AMS response is the percent change in biotinylation due to the reaction with AMS. (A) AMS response of different cysteines in the HA-EmrECless-C+ construct, when paired with HA-EmrElocked_out to generate an antiparallel dimer. Unpaired HA-EmrECless-C+ subunits are rapidly degraded by FtsH. (B) Topology mapping for the construct N_{out}-HA-EmrECless-C+. The frustrated N_{out}/C_{in} topology shown is one of many possible topologies. (C) Topology mapping for the construct N_{in}-HA-EmrECless-C+. EmrE is visualized by western blotting for the HA epitope and the results shown are representative of assays performed in triplicate.
Figure 3-10 Disulfide formation at the N-terminus. A 3C mutant in the context of the HA-EmrE<sup>Cless-C+</sup> and N<sub>out</sub>-HA-EmrE<sup>Cless-C+</sup> form dimers that are sensitive to the addition of 2-mercaptoethanol suggest that the 3C can form disulfide-linked dimers. No dimers were seen for the N<sub>in</sub>-HA-EmrE<sup>Cless-C+</sup> construct. These results are consistent with a 3C location in the oxidizing periplasm for EmrE<sup>Cless-C+</sup> and N<sub>out</sub>-HA-EmrE<sup>Cless-C+</sup> constructs and a 3C location in the reducing cytoplasm for the N<sub>in</sub>-HA-EmrE<sup>Cless-C+</sup> construct. Further disulfide formation was blocked by the addition of iodoacetamide prior to cell lysis. EmrE is visualized by western blotting for the HA epitope.
Figure 3-11  Time dependence of AMS reactivity. The extent of AMS blocking of HA-EmrEClass-C+ F27C and T108C for 10 min and 20 min is shown. Essentially no difference is observed, indicating that 10 min is sufficient to block all the available thiols in F27C and that the extent of reaction reflects an end point rather than a slowed rate of modification. EmrE is visualized by western blotting for the HA epitope.
Figure 3-12 DNA Sequences of EmrE Constructs. DNA sequence at the N- or C-termini for the indicated constructs. Regions that have a strong predicted RNA-hairpin are shown in a red font (A) HA-EmrE (B) HA-EmrE-C+ (C) EmrE-C+ (toxic) (D) EmrE-C+ hairpin (non-toxic) Mutations that do not alter the protein sequence to create a predicted RNA-hairpin are shown in a bold typeface.
Figure 3-13 Uncropped Blot of HA-EmrE_{Cless-C+} F27C. Uncropped blot of cysteine accessibility assay for HA-EmrE_{Cless-C+} F27C.
Acknowledgements

This work was supported by a NIH Grant RO1 GM063919 to JUB and an NIH Chemistry/Biology Interface Training Grant to NBW. We thank Teru Ogura for supplying the AR3289 and AR3291 strains.

3.6 References


Chapter 4

Complete Topology Inversion can be Part of Normal Membrane Protein Biogenesis
4.1 Abstract

The topology of helical membrane proteins is generally defined during insertion of the transmembrane helices, yet it is now clear that it is possible for topology to change under some circumstances. It remains unclear, however, if topology reorientation is part of normal biogenesis. For dual topology dimer proteins such as the multidrug transporter EmrE, there may be evolutionary pressure to allow topology flipping so that the populations of both orientations can be equalized. We previously demonstrated that when EmrE is forced to insert in a distorted topology, topology flipping of the first transmembrane helix can occur during translation. Here we show that topological malleability also extends to the C-terminal helix and that even complete topology inversion of the entire EmrE protein can occur after the full protein is translated and inserted. Thus, topology rearrangements are possible during normal biogenesis. Wholesale topology flipping is remarkable given the physical constraints of the membrane and expands the range of possible membrane protein folding pathways, both productive and detrimental.

4.2 Introduction

The lipid bilayer presents an apolar barrier to the transport of polar molecules. Even the simple process of lipid flip-flop, moving the lipid headgroup from one bilayer leaflet to another, is associated with a high energetic barrier. For example, the half-life for flip-flop of phosphatidylcholine lipids can range from hours to days\(^1\). Thus, models of membrane protein biogenesis generally assume that transmembrane helix topology is fixed upon initial insertion by the translocon, largely defined by the positive inside rule\(^2,3\). Nevertheless, it is now clear that topology changes are possible.
To our knowledge the first evidence for post-insertion topology changes came from the Skach group, who employed an *in vitro* transcription/translation system to study the topology of aquaporin-1 during biogenesis\(^4\). When aquaporin-1 was expressed in truncated forms as a proxy for early insertion intermediates, the protein was found inserted in a non-native topology that would need to be subsequently resolved upon insertion of the full protein. In these experiments, however, it remains unclear whether the incorrect topology of these truncated forms reflect true kinetic intermediates along the natural folding pathway or the accumulation of an off-pathway form when translation is halted prematurely.

The Dowhan group has shown that the topology of lactose permease undergoes truly remarkable changes upon variation of phosphatidylethanolamine in the membrane, involving complete topology changes of six transmembrane helices in the N-terminal domain\(^5\)–\(^9\). Moreover, these dramatic topology changes are reversible, depending on the lipid composition.

EmrE is a multi-drug resistance transporter that also exhibits topological malleability. EmrE consists of four transmembrane helices, and the active form of EmrE is a dual topology dimer in which the subunits have opposite topologies (Fig. 4-1A), with one subunit in an N-terminal inside/C-terminal inside (N\(_{\text{in}}/C_{\text{in}}\)) orientation and the other in an N\(_{\text{out}}/C_{\text{out}}\) orientation\(^10,11\). A parallel dimer can also form, however\(^12\)–\(^18\). The wild-type sequence of EmrE exhibits a weak positive-charge bias across its transmembrane helices that allows biogenesis in both topologies in accordance with the positive-inside rule that positive-charges are preferred in the cytoplasm. The von Heijne group showed that when positively-charged residues were placed at the C-terminus of EmrE, only an N\(_{\text{in}}/C_{\text{in}}\) topology remained active, suggesting that distant topology signals at the C-terminus could influence the topology of the N-terminal helices after cotranslational insertion\(^19\). Following on this work, we showed that the first transmembrane helix of EmrE normally inserts in both N\(_{\text{in}}\) and N\(_{\text{out}}\) orientations\(^20\), so that C-terminal positive charges direct both an N\(_{\text{in}}/C_{\text{in}}\) orientation and a distorted N\(_{\text{out}}/C_{\text{in}}\) orientation. The distorted N\(_{\text{out}}/C_{\text{in}}\)
orientation subsequently flips the N-terminus to a normal $N_{in}/C_{in}$ topology\textsuperscript{20}. Thus, there is some topological malleability.

While amazing topology changes are clearly possible, it remains unclear whether they could be part of a normal membrane protein biogenesis process or if topology rearrangements only happen when membrane proteins are forced into unusual membrane environments or distorted topologies. To test whether topology flipping can occur after EmrE synthesis in a natural membrane environment, we added a set of positive charges at the end of EmrE that could be subsequently removed by TEV protease. The C-terminal charges provide a strong topological signal that directs the C-terminus into the cytoplasm ($N_{in}/C_{in}$ or $N_{out}/C_{in}$ topology)\textsuperscript{20}. The ability to remove the positive charges after the protein is made allows us to observe whether the initially set topology can change after cleavage. We find that after subsequent cleavage of the C-terminal positive charges, the entire protein is indeed free to invert in the membrane.

4.3 Results and Discussion

**Protein constructs employed in this work:** To probe the topological malleability of EmrE, we employed a set of protein constructs shown in Fig. 4-1. The central experimental construct, EmrE-C+, adds a TEV cleavage site, followed by a C-terminal positive-charge tag, KKKHHHHHH (Fig. 4-1B). The charges provide a strong topology signal that directs both the N and C-termini into the cytoplasm ($N_{in}/C_{in}$ topology)\textsuperscript{20}. The TEV site provides a method to cut the tag off, thereby removing the topology signal and allowing us to test whether topology flipping can occur after charge removal. HA-EmrE-C+ adds an HA epitope tag at the N-terminus (Fig. 4-1C). As shown previously, this construct inserts into the membrane in both a normal $N_{in}/C_{in}$ topology and a distorted $N_{out}/C_{in}$ topology, and the HA tag blocks subsequent
topology flipping at the N-terminus\textsuperscript{20}. N\textsubscript{in}-HA-EmrE-C+ adds a few N-terminal positive charges to generate a pure N\textsubscript{in}/C\textsubscript{in} topology (Fig. 4-1D) while N\textsubscript{out}-HA-EmrE-C+ adds positive charges to the first extra-membrane loop to generate a pure N\textsubscript{out}/C\textsubscript{in} distorted topology (Fig. 4-1E)\textsuperscript{20}.

**Topological malleability at the C-terminus:** We had previously shown that an N\textsubscript{out}/C\textsubscript{in} topology can be resolved to an N\textsubscript{in}/C\textsubscript{in} topology indicating that the N-terminal helix of EmrE can flip. To further explore the topological malleability of EmrE, we examined whether the C-terminus can flip from a distorted N\textsubscript{out}/C\textsubscript{in} topology to a regular N\textsubscript{out}/C\textsubscript{out} topology. We employed the HA-EmrE-C+ construct that contains an HA epitope which blocks N-terminal topology rearrangements so that it inserts in two locked topologies: N\textsubscript{in}/C\textsubscript{in} and distorted N\textsubscript{out}/C\textsubscript{in} (Fig. 4-1C).

The presence of functional EmrE can easily be assessed *in vivo* by resistance of *E. coli* cells to ethidium bromide (EtBr), which is pumped out of the cell by an active EmrE. Since EmrE requires both topological forms to be functional, the topology of a particular construct can be assessed by co-expression with variants that are locked in a single topology by the strategic placement of positively charged residues (N\textsubscript{out}/C\textsubscript{out}: HA-EmrE\textsuperscript{OUT} and N\textsubscript{in}/C\textsubscript{in}: HA-EmrE\textsuperscript{IN})\textsuperscript{19,20}. EtBr resistance occurs when both topologies are present, creating an active dimer.

As shown in Fig. 4-2A, HA-EmrE-C+ by itself shows low EtBr resistance which is not complemented by the HA-EmrE\textsuperscript{IN} construct, but is complemented by the HA-EmrE\textsuperscript{OUT} construct, indicating that HA-EmrE-C+ cannot generate a proper N\textsubscript{out}/C\textsubscript{out} topology. When HA-EmrE-C+ is co-expressed with TEV protease, however, EtBr resistance is dramatically enhanced as would be expected if the C-terminus of the N\textsubscript{out}/C\textsubscript{in} topological form flipped to an N\textsubscript{out}/C\textsubscript{out} topology and dimerizes with the N\textsubscript{in}/C\textsubscript{in} topology of HA-EmrE-C+ (Fig. 4-2A).
To assess which topological form of HA-EmrE-C+ was undergoing a topology change, we examined the ability of TEV protease cleavage to activate the topology locked variants $N_{in}$-HA-EmrE-C+ and $N_{out}$-HA-EmrE$^{\text{Class-}}$-C+. In $N_{out}$-HA-EmrE$^{\text{Class-}}$-C+ we changed the native cysteine residues to alanine because, for unknown reasons, the native cysteine residues were toxic in this construct alone (Fig. 4-3). As shown in Fig. 4-2, TEV protease co-expression failed to activate these constructs when expressed alone as the mutated positive charges direct each protein to a single topological form$^{20}$. The $N_{in}$/C$_{in}$ topology of $N_{in}$-HA-EmrE-C+ only restores EtBr resistance when co-expressed with HA-EmrE$^{\text{OUT}}$ regardless of TEV protease expression suggesting that the $N_{in}$/C$_{in}$ of HA-EmrE-C+ does not change topology (Fig. 4-2B). However, after TEV protease co-expression and removal of the C-terminal positive charges, the $N_{out}$/C$_{in}$ distorted topology of $N_{out}$-HA-EmrE$^{\text{Class-}}$-C+ appears to regain a viable $N_{out}$/C$_{out}$ topology since it can be complemented with HA-EmrE$^{\text{IN}}$ (Fig. 4-2C). These results suggest that the C-terminus can be driven to move from the cytoplasm to the periplasm.

**Direct Observation of Topology Changes after Cleavage** – To validate the indirect topology measurements from the ethidium bromide resistance phenotypic assay, we used the substituted cysteine accessibility method (SCAM) to directly measure topology$^{21}$. SCAM ascertains whether a single introduced cysteine resides in either the cytoplasm or the periplasm by assessing its reaction with a membrane-impermeable reagent, in this case, 4-acetamido-4′-maleimidylstilbene-2,2′-disulfonic acid (AMS). AMS is first incubated with whole cells, allowing it to react with periplasmic cysteine residues. After AMS is washed away, membranes are solubilized in detergent and a biotinylation reagent, 3-($N$-maleimido-propionyl)-biocytin (MPB) is added to react with any free cysteine residues that have not been blocked by AMS. In this way, cysteine residues in the periplasm will be modified by AMS and cysteine residues in the cytoplasm will be biotinylated with MPB. We can then differentiate the two cases, by observing
whether the EmrE construct gel-shifts with the addition of avidin. Since the biotinylation reaction is somewhat variable, we run a biotinylation only control with every sample to assess the percent change in biotinylation due to the AMS pre-reaction, which we term AMS response (see Methods). Control experiments indicate that a cytoplasmic cysteine can be fully protected from AMS (0% AMS response), fully periplasmic cysteines are highly, but not completely reactive with AMS (~75% response), and mixed topologies generate an intermediate response20 (Fig. 4-4).

The HA epitope cannot cross the membrane, trapping the N-terminus of EmrE in these experiments in its initially inserted topology20. As such, we can monitor the movement of the C-terminal half of EmrE after TEV protease cleavage. As described previously, we changed all the native cysteine residues in EmrE to alanine (C39A, C41A, C95A) creating EmrE\textsubscript{Cless}, which eliminates background reactions with the maleimide reagents (Fig. 4-5). The T108C mutation allows us to monitor the movement of the C-terminus of EmrE in the constructs HA-EmrE\textsubscript{Cless}-C+, N\textsubscript{in}-HA-EmrE\textsubscript{Cless}-C+ and N\textsubscript{out}-HA-EmrE\textsubscript{Cless}-C+. In our previous work, we established that when T108C is adjacent to the C-terminal positive charges, it is always cytoplasmic20.

When the EmrE constructs are co-expressed with TEV protease, the cleaved form of EmrE is resolved on the gel from the full length form of EmrE allowing both the cut and full length topologies to be assessed at the same time. In line with the previous results T108C is cytoplasmic in the full length constructs as indicated by the low AMS response in the constructs HA-EmrE\textsubscript{Cless}-C+ T108C (-3% ± 8%), N\textsubscript{in}-HA-EmrE\textsubscript{Cless}-C+ T108C (3% ± 2%) and N\textsubscript{out}-HA-EmrE\textsubscript{Cless}-C+ T108C (10 ± 8%) (Fig. 4-5). Once cleaved the distorted topology of N\textsubscript{out}-HA-EmrE\textsubscript{Cless}-C+ T108C re-orient to move T108C into the periplasm as shown by the increase in the AMS response for N\textsubscript{out}-HA-EmrE\textsubscript{Cless}-C+ T108C after cleavage (47% ± 17%) (Fig. 4-5A). With the N-terminus cytoplasmic in the construct N\textsubscript{in}-HA-EmrE\textsubscript{Cless}-C+ T108C, T108C is already in line with a proper topology for EmrE and we observe no movement across the membrane.
after cleavage, indicated by an indistinguishable AMS response (1% ± 4%) compared to the full length protein (3% ± 2%) (Fig. 4-5B). HA-EmrE\textsuperscript{Cless} C+ T108C which is composed of equal portions N\textsubscript{in} and N\textsubscript{out} shows movement into periplasm as well with, an increased AMS response upon cleavage (increase from -3% ± 8% to 38% ± 12%), most likely originating from the N\textsubscript{out}/C\textsubscript{in} topology monomers (Fig. 4-5C).

**Full topology inversion:** The results so far and our prior work\textsuperscript{20} indicate that the N- and C-terminal ends of EmrE can change topologies in response to a distorted topology at the other end of the protein. But if the protein is inserted in a normal topology with no topology distortions to drive changes, would it still flip? To test this possibility we removed the HA epitope which blocks topology flipping at the N-terminus\textsuperscript{20}, creating the construct EmrE-C+ (Fig. 4-1). EmrE-C+ adopts a uniform N\textsubscript{in}/C\textsubscript{in} topology\textsuperscript{20}. Removing the C-terminal positive charges by TEV protease cleavage would thereby eliminate any topological determinants and free the protein to flip in the membrane, if possible.

As seen in Fig. 4-6A, cells expressing EmrE-C+ grow poorly on EtBr, consistent with insertion in a single N\textsubscript{in}/C\textsubscript{in} topology without an N\textsubscript{out}/C\textsubscript{out} partner to form an active dimer\textsuperscript{20}. Co-expressing EmrE-C+ with an N\textsubscript{in}/C\textsubscript{in} single topology EmrE mutant, HA-EmrE\textsuperscript{IN}, does not restore growth on EtBr, further indicating that EmrE-C+ and HA-EmrE\textsuperscript{IN} are in the same topology (Fig. 4-6A). However, when EmrE-C+ was co-expressed with the single topology HA-EmrE\textsuperscript{OUT} variant, EtBr resistance is restored as observed previously\textsuperscript{19,20}. Taken together, these results indicate that without removal of the C-terminal positive charges, EmrE-C+ adopts an N\textsubscript{in}/C\textsubscript{in} topology only.

As shown in Fig. 4-6A, when the TEV protease is co-expressed with EmrE-C+ alone, the EtBr resistance increases dramatically, suggesting that both topologies are present. If so, it
would require that cut EmrE-C+ can flip in the membrane after the protein is made in the N_{in}/C_{in} topology. Co-expression with HA-EmrE^{IN} or HA-EmrE^{OUT}, does not change the EtBr resistance as EmrE-cut is active alone (Fig. 4-6A).

To further evaluate whether the gain of function occurs simply by removal of the C-terminal positive charges, we employed the active site mutation E14D. The E14D mutation is inactive by itself, but can form an active dimer when paired with a wild-type subunit^{19,22}. Thus, we reasoned that if activation occurs upon TEV protease cleavage of a subunit containing an E14D mutation, it must be due to proper pairing with another subunit and not activation of the cleaved subunit itself. As shown in Fig. 4-6B, both EmrE-C+ E14D and EmrE-cut E14D are inactive since only E14D monomers are present. When EmrE-C+ E14D is co-expressed with HA-EmrE^{OUT}, however, a functional dimer of N_{in}/C_{in} EmrE-C+ E14D and HA-EmrE^{OUT} is formed (Fig. 4-6B). The combination of EmrE-C+ E14D and HA-EmrE^{IN}, fails to yield a productive interaction, as expected, since both subunits are in an N_{in}/C_{in} orientation. When EmrE-C+ E14D is cleaved by the TEV protease to EmrE-cut E14D, however, activity is restored in the presence of HA-EmrE^{IN}, consistent with the ability of EmrE-cut E14D to flip topologies and form an active dimer with HA-EmrE^{IN} (Fig. 4-6B).

While the biological phenotypes strongly suggest the EmrE possesses the ability to completely flip topology, we wanted to test the ability at the protein level. The absence of an antibody tag in the EmrE-C+ construct after TEV cleavage, however, precluded a SCAM analysis employed above. We therefore employed an indirect assay that detects anti-parallel dimer formation by protection from intracellular proteolysis. As shown in Fig. 4-6C, HA-EmrE^{IN} construct is degraded at a much higher rate than our HA-EmrE^{OUT} mutant despite a difference of only a few topology defining mutations, because the cytoplasmically located FtsH protease can grasp the termini of HA-EmrE^{IN}, but not HA-EmrE^{OUT}. The half-life of the HA-EmrE^{IN} construct is less than 30 min while the EmrE^{OUT} construct shows no obvious degradation after 120 min.
When HA-EmrE\textsuperscript{IN} is expressed in FtsH null cells, its degradation is essentially eliminated, indicating that FtsH is the primary protease responsible for degrading HA-EmrE\textsuperscript{IN} (Fig. 4-6C).

We previously showed that HA-EmrE-C+ was protected from FtsH proteolysis if it could form an active anti-parallel dimer, presumably because the N\textsubscript{in}/C\textsubscript{in} topological form is stabilized in the dimer, making it immune to FtsH proteolysis\textsuperscript{23}. Thus, if EmrE-C+ can change topologies to an N\textsubscript{out}/C\textsubscript{out} state after TEV cleavage, then it should stabilize the N\textsubscript{in}/C\textsubscript{in} construct, HA-EmrE\textsuperscript{IN}, which we can monitor using the HA epitope. Indeed, as seen in Fig. 4-6D, HA-EmrE\textsuperscript{IN} has increased stability when co-expressed with EmrE-C+ and the TEV protease (17% ± 6% remaining after 2 hours of degradation compared to 1% ± 1% without the TEV protease).

4.4 Conclusion

Our results indicate that EmrE exhibits a high degree of topological malleability. Both the N- and C-termini can change topology independently and the entire protein can flip completely in the membrane. While it is possible that the entire protein flips in a concerted manner, we suggest that the flipping may occur in a more piecemeal process. Our results are largely consistent with the kinetic annealing model of Van Lehn et al., in which EmrE can insert initially in a variety of topologies which can be subsequently resolved into either an N\textsubscript{in}/C\textsubscript{in} or C\textsubscript{out}/N\textsubscript{out} topology\textsuperscript{24}. What was not anticipated in the model, however, was that the energy barrier for the kinetic annealing process may be low enough for EmrE to continue in reverse. In particular, if the distorted topologies such as N\textsubscript{out}/C\textsubscript{in} are not too unfavorable energetically, it is possible that they could be explored with reasonable probability even after correct topologies are achieved, ultimately resolving in an inverted orientation. Our results suggest that the final topology of EmrE only becomes fixed upon stabilization in the anti-parallel dimer.
Evolutionary pressure for topological malleability makes sense for dual topology proteins like EmrE that need both topologies in equal amounts. Thus, EmrE maintains relatively short loops and is relatively hydrophobic overall. Whether such folding flexibility occurs in other proteins as part of the natural folding process is an open question. Nevertheless, our results indicate that topology flipping not only occurs under aberrant conditions, but also as part of a normal biogenesis of EmrE, so it is reasonable to suppose that topological changes are part of the natural folding process of other membrane proteins as well. Moreover, we must consider the possibility that disease-causing mutations could act by blocking topological malleability required during the folding process. Topological changes after insertion seem particularly likely to occur in proteins with re-entrant loops like ClC channels\textsuperscript{25} and Aquaporins\textsuperscript{4} as found originally by the Skach group.

### 4.5 Methods

**Strains and Plasmids:** The primary EmrE variant constructs were prepared and expressed in pBAD/His A plasmids (Invitrogen) using NcoI/ XhoI cut sites as described previously\textsuperscript{20}. HA-EmrE\textsuperscript{IN} and HA-EmrE\textsuperscript{OUT} were co-expressed in a separate pBAD based vector that bears chloramphenicol resistance and a ClodF13-derived CDF replicon (pBAD\textsubscript{CDF}) described previously\textsuperscript{20}. The pRK603 plasmid containing the TEV protease for \textit{in vivo} expression and cleavage of substrates was a gift from David Waugh (Addgene plasmid # 8831) as well as the BL21Pro cells that constitutively express the Tet repressor to control TEV cleavage\textsuperscript{26}. AR3291 (FtsH null) cells were a gift from the Ogura lab\textsuperscript{27}. 
**Ethidium Bromide Resistance Assay**  
*E. coli* BL21Pro cells containing the desired EmrE construct (pBad HisA plasmid), the complementing EmrE construct (Empty vector, HA-EmrE\textsuperscript{IN}, HA-EmrE\textsuperscript{OUT}) in pBad\textsubscript{CDF}, and TEV protease on the pRK603 plasmid were grown to saturation over \(\sim 8\) hours at 37°C in LB media containing 100 \(\mu\text{g-mL}^{-1}\) ampicillin, 50 \(\mu\text{g-mL}^{-1}\) kanamycin, and 34 \(\mu\text{g-mL}^{-1}\) chloramphenicol. The saturated cultures were diluted as indicated into LB broth and 5\(\mu\text{L}\) of each dilution was spotted onto an LB agar plate containing .2% (w/v) arabinose, 100 \(\mu\text{g-mL}^{-1}\) ampicillin, 50 \(\mu\text{g-mL}^{-1}\) kanamycin, 34 \(\mu\text{g-mL}^{-1}\) chloramphenicol, 100 ng-\(\text{mL}^{-1}\) anhydrotetracycline (if the TEV protease was induced) and 225 \(\mu\text{g-mL}^{-1}\) of ethidium bromide. For E14D mutant complementation assays, the ethidium bromide concentration was reduced to 150 \(\mu\text{g-mL}^{-1}\). The plates were grown at 37°C for 18 hours before they were imaged on a Gel Doc XR+ (Bio-Rad) using the UV light illumination.

**Cysteine-accessibility topology assay**  
AR3291 (FtsH null) cells bearing the desired plasmids were grown at 25°C in LB media containing 100 \(\mu\text{g-mL}^{-1}\) ampicillin, 50 \(\mu\text{g-mL}^{-1}\) kanamycin, and 34 \(\mu\text{g-mL}^{-1}\) chloramphenicol to an \(\text{OD}_{600}\) of \(\sim 0.8\), then induced by the addition of arabinose to 0.2% and incubated at 30°C to express the desired EmrE construct and/or TEV protease. EmrE constructs alone were induced for 2 hours and when both the EmrE construct and the TEV protease were present, induction proceed for 3 hours. 50\(\text{mL}\) of the cell culture was collected by centrifugation and the resuspended in 500\(\mu\text{L}\) of 50\(\text{mM}\) phosphate buffer [pH 8.0] with 17\(\text{mM}\) NaCl. 200 \(\mu\text{L}\) of the resuspension was incubated for 10 minutes with or without 2\(\text{mM}\) 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid (AMS) (Thermo Fisher Scientific) with gentle mixing the dark. The cells were pelleted by centrifugation and washed twice with \(\sim 1\text{mL}\) of 50\(\text{mM}\) phosphate [pH 7.5]. The final pellet was resuspended in \(\sim 475\mu\text{L}\) in 50\(\text{mM}\) phosphate [pH 7.5]. The cells were then lysed by sonication and then centrifuged at 16,000 \(g\) for 10 minutes. Membranes were isolated from the supernatant by centrifugation at 160,000 \(g\) in a
Beckman Coulter Airfuge for 30 minutes. The membranes were solubilized in 1% SDS, 50mM phosphate [pH 7.0]. The total protein concentration for each sample was determined by the DC protein assay (Bio-Rad) using bovine serum albumin as a standard. An aliquot of the membranes was biotinylated in 200 µl of 150µM 3-(N-maleimido-propionyl)-biocytin (MPB) (Thermo Fisher Scientific) in 50mM phosphate [pH 7.0], 1% SDS and 0.5% dimethyl sulfoxide and a final protein concentration of 0.5 mg·mL⁻¹, with gentle mixing for 1 hour at 25°C. The protein was then precipitated with ~1.2mL of acetone and then centrifuged at 16,000g to remove unreacted MPB. The pelleted was then air-dried and resuspended in 200µL of 1% SDS in 50mM phosphate [pH 7.0].

To visualize the AMS and MPB labeling by an avidin gel shift, 30µL of the labeled sample was mixed with 10µL of 4X loading dye (250mM tris [pH 6.75], 40% glycerol, 170mM β-mercaptoethanol). The sample was then split into two 20µL aliquots and 2µL of either 20mM tris [pH 7.5] or 10mg/mL avidin (Sigma BioUltra) in 20mM tris [pH 7.5] was added. 12µL of each sample was then loaded onto a NuPAGE™ 12% Bis-Tris gel (Thermo Fisher Scientific) using Accuruler prestained protein ladder (Lambda Biotech). The gel was resolved for 25 minutes at 40 volts followed by 145 minutes at 100 volts. The gel was then washed twice for 15 minutes in distilled water and then transferred to a PVDF membrane using a Pierce Power Blot Cassette (Thermo Scientific) with Pierce 1-Step Transfer Buffer (Thermo Scientific) for 5 minutes at 1.3 Amps. The blot was rinsed in water and placed into the iBind (Life Technologies) using the associated iBind Solution Kit (Life Technologies) according the iBind instructions. A 1 mg·mL⁻¹ stock of monoclonal HA antibody (Sigma #H3663) and a 1 mg·mL⁻¹ stock of anti-mouse IgG peroxidase conjugate (Sigma #4416) was diluted 1:1250 in iBind solution. After the iBind step (~2.5 hours), the blot was rinsed with water and was visualized on a FluorChem FC2 (Alpha Innotech) CCD imager using the Amersham ECL Prime detection reagent (GE Healthcare) according to the recommended protocol. The intensity of each band was integrated using
ImageJ software with general background subtraction taken from a blank area of the blot. The amount of biotinylation for a given sample was determined from the percent difference in integrated intensity from the lane with avidin compared to the lane without avidin. The percent change in biotinylation of the sample with AMS as compared to the sample without is quantified as AMS Response:

$$\text{AMS Response} = \frac{[\%\text{biotinylated no AMS}] - [\%\text{biotinylated AMS}]}{[\%\text{biotinylated no AMS}]} \times 100$$

Each topology assay was run in triplicate and the error quantified as the standard deviation between the three measured AMS responses.

**EmrE Stability Assay** - *E. coli* BL21Pro cells containing the HA-EmrE\textsuperscript{IN} in the pBAD\textsuperscript{CDF} plasmid, EmrE-C+ in the pBAD His A plasmid and the TEV plasmid on the pRK603 vector were grown in LB media to ~0.6 OD\textsuperscript{600} and induced at 30°C for 2 hours. To stop protein synthesis, 350 µg·mL\textsuperscript{-1} of erythromycin was then added to stop protein synthesis, denoting time zero in the degradation assay. Aliquots of each culture were taken at the respective time points. The cells were then lysed by sonication in 50mM phosphate buffer [pH 7.5] and then centrifuged at 16,000g for 10 minutes. Membranes were isolated by the centrifuging the supernatant at 160,000g in the Beckman Coulter Airfuge for 30 minutes. The membranes were solubilized in 1% SDS, 50mM phosphate [pH 7.0]. The total protein concentration for each sample was determined by the DC protein assay (Bio-Rad) using bovine serum albumin as the standard. The total protein concentration was adjusted to 1 mg·mL\textsuperscript{-1} and a 15µL aliquot was mixed with 5µL 4X SDS loading dye (250mM tris [pH 6.75], 40% glycerol, 170mM β-mercaptoethanol) and then 15µL of the mixture was loaded into each well. The gel separation and western blot were performed as described above. The HA-EmrE\textsuperscript{IN}/EmrE-C+ combinations with or without the TEV protease were
performed in triplicate. The error for the percent of HA-EmrEmr remaining after 120 minutes is expressed as the standard deviation.
Figure 4-1. EmrE Constructs and Topology Models. (A) Wild-Type EmrE is composed of opposite topology monomers with four transmembrane helices that form a dual-topology dimer. (B) The construct EmrE-C+ consists of a wild-type N-terminus and a C-terminal positive-charge tag behind a TEV protease cleavage site. The topology of this construct has been previously shown to be N_{in}/C_{in}.20 (C) HA-EmrE-C+ adds an HA-epitope tag to the N-terminus of EmrE-C+. The addition of the N-terminal HA epitope traps the N-terminus of EmrE in its initially inserted topology20. The C-terminal positive charges enforce a cytoplasmic C-terminus. The distorted N_{out}/C_{in} topology model shown is one of many possible ways to generate an N_{out}/C_{in} topology. (D) N_{in}-HA-EmrE-C+ adds three arginine residues to the N-terminus of HA-EmrE-C+ driving the formation of a homogenous N_{in}/C_{in} topology20. (E) N_{out}-HA-EmrE-C+ adds two arginine mutations in the first loop of HA-EmrE-C+ driving the formation of a homogenous distorted N_{out}/C_{in} topology of N_{out}-HA-EmrE-C+.20
Figure 4-2. Topology Changes Assessed by Ethidium Bromide Resistance. The ability of various combinations of EmrE constructs to grow on plates containing ethidium bromide, in the presence or absence of TEV protease co-expression. Resistance to ethidium bromide is assessed by growth of the indicated dilution of a stationary phase culture. Cartoons at the right present topological models consistent with the results. (A) Analysis of HA-EmrE-C+ (B) Analysis of N_{in}-HA-EmrE-C+ (C) Analysis of N_{out}-HA-EmrE_{Cless}-C+.
Figure 4-3 Toxicity of the \( N_{\text{out}}-\text{HA-EmrE}^{\text{Class}}-C^{+} \) Construct. Agar plates were spotted with dilutions of saturated culture in the same manner as the ethidium bromide resistance assay. On the left, expression of the EmrE protein constructs was not induced. On the right marked "+inducer", .2% arabinose was added to induce expression of the indicated EmrE construct. Expression of the construct \( N_{\text{out}}-\text{HA-EmrE}-C^{+} \) prevents cell growth but not \( N_{\text{out}}-\text{HA-EmrE}^{\text{Class}}-C^{+} \).
Figure 4-4 Substituted Cysteine Accessibility Method (SCAM). An illustration of the SCAM method. (A) Cytoplasmic cysteine case. AMS is added to whole cells. AMS cannot react with the cysteine because AMS cannot cross the membrane. The AMS is washed away and isolated membranes are solubilized in detergent and a biotin maleimide reagent added, which reacts with any free cysteine. A no-AMS control is also performed to assess the degree biotinylation without AMS pre-reaction. To determine the amount of biotinylation, we add avidin which gel shifts biotinylated EmrE. The loss of protein at EmrE’s normal ~12 kDa band reflects the amount of biotinylated protein. (B) Periplasmic cysteine case. AMS reacts with the accessible periplasmic cysteine upon its addition to whole cells. The AMS then blocks the subsequent reaction with the biotin maleimide. As such the protein will no longer gel shift with avidin in the AMS pre-reaction lanes. An example calculation for AMS response is shown. Based on control experiments, ~75% AMS response represents a fully periplasmic cysteine.
**Figure 4-5 Background Biotinylation Control.** The construct $N_{in}$-HA-EmrE$^{Cless}$-C+ does not gel shift with avidin after reaction with MPB because it contains no cysteine residues. The protein is visualized by a western blot for the HA epitope.
Figure 4-6 Analysis of C-terminal Topology Changes. SCAM analysis of the C-terminal topology of EmrE constructs before and after cleavage by TEV protease. The representative western blots show the ability of avidin to gel-shift biotinylated EmrE with or without prior reaction with AMS as indicated. Protection from biotinylation is quantified as the AMS response (see Methods). The mean AMS response and standard deviation of three separate experiments are shown below each representative blot. Proposed topology models consistent with the cysteine SCAM data are shown on the right. (A) Analysis of the N<sub>out</sub>-HA-EmrE<sup>C≤ss</sup>-C+ construct. In the full length distorted topology construct N<sub>out</sub>-HA-EmrE<sup>C≤ss</sup>-C+, T108C is completely cytoplasmic as indicated by the negligible AMS response. After TEV protease cleavage, N<sub>out</sub>-HA-EmrE<sup>C≤ss</sup>-C+ undergoes a topology change with some of the T108C residues moving into periplasm as indicated by the increased AMS response of the cut form. (B) Analysis of the N<sub>n</sub>-HA-EmrE<sup>C≤ss</sup>-C+ construct. In the full length and cut versions of N<sub>n</sub>-HA-EmrE<sup>C≤ss</sup>-C+, T108C is completely cytoplasmic as indicated by the negligible AMS response. Topology changes upon cleavage are not observed. (C) Analysis of the HA-EmrE<sup>C≤ss</sup>-C+, T108C construct. HA-EmrE<sup>C≤ss</sup>-C+ adopts a mixed topology, composed equally of N<sub>n</sub> and N<sub>out</sub> forms<sup>20</sup>. In the full length construct, T108C is cytoplasmic as measured by the low AMS response for both topological forms. After TEV protease cleavage, however, a fraction of the T108C moves into the periplasm as indicated by the increase in AMS response.
Figure 4-7 Full Topology Inversion of EmrE-C+. (A) EmrE-C+ topology before and after cleavage by TEV protease assessed by ethidium bromide resistance, as described in Fig. 2. A topological interpretation of the results is shown below the growth results. (B) EmrE-C+ E14D topology before and after cleavage by TEV protease assessed by ethidium bromide resistance, as described in Fig. 2. (C) Degradation of EmrE constructs by FtsH protease. Western blots show EmrE levels after protein synthesis is halted by the addition of erythromycin (degradation time zero). HA-EmrE\textsuperscript{IN} and HA-EmrE\textsuperscript{OUT} were expressed in FtsH\textsuperscript{+} cells for the left two panels. The third panel shows that HA-EmrE\textsuperscript{IN} degradation is largely eliminated in ΔFtsH cells. (D) Degradation of HA-EmrE\textsuperscript{IN} in the presence of uncut EmrE-C+ or cut EmrE-C+. HA-EmrE\textsuperscript{IN} is stabilized when EmrE-C+ is expressed with TEV protease, consistent with its ability to change topology after cleavage and bind to HA-EmrE\textsuperscript{IN}. The blot shown is representative. The percent of HA-EmrE\textsuperscript{IN} remaining after 120 minutes is indicated at the bottom (mean and standard deviation of triplicates).
References


