UNIVERSITY OF CALIFORNIA

Los Angeles

Uncovering Membrane Protein Stability Under Native Conditions

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Biochemistry and Molecular Biology

by

Robert Everett Jefferson

2016

ABSTRACT OF THE DISSERTATION

Uncovering Membrane Protein Stability Under Native Conditions

by

Robert Everett Jefferson

Doctor of Philosophy in Biochemistry and Molecular Biology

University of California, Los Angeles, 2016

Professor James U. Bowie, Chair

Membrane proteins are a neglected, but important class of proteins throughout the biological world. They carry out critical roles in the cell due to their unique location, such as transport across a membrane, transduction of exterior signals, and interaction between discrete aqueous regions. Despite the importance of these proteins, understanding of how they fold has lagged far behind that of soluble proteins. One of the primary challenges to studying membrane protein folding is developing methods that interrogate folding in the native environment of the lipid bilayer. Our lab has developed a method for measuring membrane protein stability under native conditions using a secondary protein that preferentially binds the unfolded state, obviating the need for harsh denaturants. Employing this method with a multimeric polytopic membrane protein, we measured an extremely slow unfolding rate, demonstrating that α-helical

membrane proteins can have high kinetic stability under non-denaturing conditions. Efforts were made to expand the steric trap method for single-molecule fluorescence measurements in lipid vesicles, but were ultimately stymied by the inability to preserve the trapped complexes for measurement. Our lab has also applied single-molecule techniques to membrane protein folding. We were able to map the energy landscape of a membrane protein in a lipid bilayer using forced unfolding driven by magnetic tweezers. Further advancements to this technique simplified the attachment chemistry to ready the protein for tweezing. These techniques can be applied to a wide array of membrane proteins in a broad spectrum of membrane environments.

The dissertation of Robert Everett Jefferson is approved.

Pascal Francois Egea

Todd O. Yeates

James U. Bowie, Committee Chair

University of California, Los Angeles

2016

DEDICATION

Dedicated to my parents,

who are an endless font of humor and support.

TABLE OF CONTENTS

Abstract of the Dissertation		ii
Acknowledgements		viii
Vita		ix
Chapter 1:	Introduction	1
	References	14
Chapter 2:	Membrane proteins can have high kinetic stability	19
Chapter 3:	Steric trapping in lipid bilayers for single-molecule measurement	31
	References	53
Chapter 4:	Mapping the energy landscape for second-stage folding of a single membrane protein	55
Chapter 5:	A simple DNA handle attachment method for single molecule mechanical manipulation experiments	83
Chapter 6:	Methods for understanding membrane protein folding	94
	References	121

LIST OF FIGURES

Chapter 1

Figure 1-1 Two-stage model of α-helical membrane protein folding 7

Chapter 3

Figure 3-1 Schemes for preparing steric trapped membrane protein in vesicles for single-molecule measurements 35

Figure 3-2 Comparison of vesicle bursts with and without siliconization 39

Figure 3-3 Dissociation kinetics of mSA_{A35C} variants and biotin-4-fluorescein 40

Figure 3-4 Fluorescent bursts after removing free streptavidin with DMSO-washed biotin-agarose beads 41

Figure 3-5 Removal of free streptavidin by desalting spin columns 43

Figure 3-6 Negative controls for stoichiometry of coincident bursts 44

Figure 3-7 Comparison of burst stoichiometry for titrations of fluorophore-labeled WT mSA_{A35C} 46

Figure 3-8 Fraction of coincident bursts from unfolded bR_{M20A}-BE_2 47

Figure 3-9 Measuring unfolding bR_{M20A}-BF_2 with WT mSA by gel-shift 50

Figure 3-10 Dissociation of mutant mSA from steric trapped complexes 52

Chapter 6

Figure 6-1 Equilibrium Unfolding with SDS 100

Figure 6-2 Steric Trap Unfolding 107

Figure 6-3 Single-molecule Force Spectroscopy for Membrane Protein Unfolding 114

ACKNOWLEDGEMENTS

Chapter 2 is reprinted with permission from the Journal of the American Chemical Society. I would like to acknowledge Tracy M. Blois and Heedeok Hong for pioneering development of the steric trap method and Tracy M. Blois again for laying the groundwork for steric trapping of diacylglycerol kinase. I would also like to thank James U. Bowie for his guidance and conception of the research. The research was supported by National Institutes of Health grant 5R01GM063919 and by the Ruth L. Kirschstein National Research Service Award GM007185.

Chapter 3 is a summary of research efforts to extend the steric trap method into the single-molecule realm. I would like to acknowledge Eitan Lerner, Xavier Michalet, SangYoon Chung, and Yazan Alhadid for training and help with the instrumentation and analysis software for single-molecule measurements. I would also like to thank James U. Bowie and Shimon Weiss for their guidance and conception of the research. I would like to give a special thanks to Scott C. Blanchard for suggesting the use of ultra-stable fluorophores originally developed in his lab and to Heedeok Hong for the gift of specialized fluorescein-conjugated biotin label created in his lab by Ruiqiong Guo.

Chapter 4 is reprinted with permission from Nature Chemical Biology. I would like to acknowledge my co-author Duyoung Min for his expertise and for taking membrane proteins into a new frontier. I would also like to thank James U. Bowie and Tae-Young Yoon for their guidance and conception of the research. The work was supported by the National Creative Research Initiative Program funded by the National Research Foundation of Korea, and Marine Biotechnology Program funded by the Ministry of Oceans and Fisheries of Korea, and supported by US National Institutes of Health grant 2R01GM063919.

Chapter 5 is reprinted with permission from Protein Science. I would like to acknowledge my co-authors Duyoung Min and Mark E. Arbing for further improving upon an already robust technique. I would also like to thank James U. Bowie for his guidance and conception of the research. The work was supported by National Institutes of Health grant R01GM063919.

VITA

2004 - 2008	Bachelors of Science, Biochemistry, Biophysics, and Molecular Biology Whitman College
2006	Undergraduate Research Internship Agilent Technologies
2007 - 2008	Undergraduate Student Researcher Department of Biochemistry, Biophysics, and Molecular Biology Whitman College Advisor: Professor Douglas Juers
2008 - 2016	Graduate Student Researcher Department of Chemistry and Biochemistry University of California, Los Angeles Advisor: Professor James U. Bowie
2009	Teaching Assistant General Chemistry Laboratory; Biochemistry: DNA, RNA, and Protein Synthesis; Physical Biochemistry Department of Chemistry and Biochemistry University of California, Los Angeles
2010 – 2013	Cellular and Molecular Biology Training Grant Fellowship Department of Chemistry and Biochemistry University of California, Los Angeles

2013 – 2014 Teaching Assistant
Biochemistry: Introduction to Structure, Enzymes, and Metabolism
Department of Chemistry and Biochemistry
University of California, Los Angeles

PUBLICATIONS AND PRESENTATIONS

Fan, C., Cheng, S., Liu, Y., Escobar, C. M., Crowley, C. S., Jefferson, R. E., ... & Bobik, T. A. (2010). Short N-terminal sequences package proteins into bacterial microcompartments. *Proceedings of the National Academy of Sciences*, 107(16), 7509-7514.

Jefferson, R. E., Blois, T. M., & Bowie, J. U. (2013). Membrane proteins can have high kinetic stability. *Journal of the American Chemical Society*, 135(40), 15183-15190.

Min, D., Jefferson, R. E., Bowie, J. U., & Yoon, T. Y. (2015). Mapping the energy landscape for second-stage folding of a single membrane protein. *Nature chemical biology*, 11, 981-987.

Min, D., Arbing, M. A., Jefferson, R. E., & Bowie, J. U. (2016). A simple DNA handle attachment method for single molecule mechanical manipulation experiments. *Protein Science*, 25, 1535-1544.

CHAPTER 1

Dissecting the Determinants of Membrane Protein Folding

At the border of every biological compartment lies a membrane, a complex environment at which certain molecules cross, signals are transduced, and interactions are made between otherwise isolated regions. A dismissive observer may see membranes as simply edges that serve to define and differentiate subcellular volumes, but of course these biological borders are composed of a diverse array of lipids in which a functionally broad group of proteins associate with and span the membrane, all to carry out critical functions within, at, and across the lipid bilayer. Those proteins that span the membrane experience unique constraints at the center of this complex where they must interact with the hydrophobic tails of lipids, their charged headgroups, two different aqueous regions, and other proteins both in and out of the lipid bilayer.

How these proteins form and maintain a specific fold in the membrane is a major question of biology whose importance cannot be overstated. Our lab has sought to study membrane protein folding under native conditions by reconstituting purified components in lipid bilayer mimics that closely approximate the natural cellular membrane. Our methods aim to circumvent the use of harsh denaturants to drive unfolding and instead use strategies to unfold membrane proteins under conditions as close as possible to the physiological environment. We seek to quantify the thermodynamics and kinetics of membrane protein folding with these *in vitro* studies. The research discussed herein is concerned with the unfolding kinetics of a trimeric membrane protein and mapping the folding energy landscape of single membrane proteins.

Membrane Protein Folding: In the shadow of soluble protein folding

Protein folding is a fundamental process in biology. Our understanding of this complex structural rearrangement of polypeptide chains opens up avenues to curing misfolding diseases, predicting structure, and engineering proteins for new functions. Many disease-causing mutations are known to affect protein assembly and trafficking rather than function (Sanders and Myers, 2004). Small changes in primary structure, such as single amino acid substitutions, can alter the folding landscape and destabilize the native fold or create kinetically trapped intermediates. Mutations can also affect the folding of other wild-type proteins during oligomerization or aggregation, further illustrating how a small alteration can have drastic effects. Studies of soluble protein folding have developed to where misfolding diseases can be treated with small molecules that can prevent aggregation of specific proteins (Cohen and Kelly, 2003; Hammarström et al., 2003) and alter the overall proteostasis network of the cell (Balch et al., 2008).

In addition to directly addressing misfolding diseases, our ability to predict what secondary, tertiary, and quaternary structures will form from a given primary sequence has advanced substantially for soluble proteins. De novo computational folding has been achieved for small soluble proteins (Bradley et al., 2005) with RMSDs from crystal structures as low as 1 angstrom. Powerful modern computers are now able to simulate the atomistic folding of small proteins (Piana et al., 2012). Only recently have computational algorithms for membrane protein folding developed to accurately predict structure (Kim et al., 2014), but there is still plenty of room for improvement as accurate prediction is limited to small alpha-helical monomers. These models also require a set of experimentally derived known structures and are not directly simulated from quantum mechanics.

A detailed understanding of how proteins fold also opens up the possibility of engineering proteins of new folds. David Baker and colleagues have demonstrated the design of a novel protein fold (Kuhlman et al., 2003) and developed a set of design principles for sculpting soluble proteins (Koga et al., 2012). Protein design has been applied to making monodisperse self-assembling protein cages of various sizes and oligomeric states (King et al., 2012; Lai et al., 2012). These protein scaffolds are open to the wide variety of amino acid chemistries to create custom molecular machines with novel functions. Membrane protein design is still in early stages (Perez-Aguilar and Saven, 2012), but notable examples include functional helical bundles that transport electrons across the membrane (Korendovych et al., 2010) and a Zn^{2+}/H^{+} antiporter (Joh et al., 2014).

Increasing understanding of the principles behind membrane protein folding is the key to solving biological misfolding problems and engineering useful structures. The Bowie lab has sought to create new methods for studying membrane protein folding under native conditions and developing tools for alleviating the challenges of studying membrane proteins *in vitro*. While there are two distinct classes of membrane proteins, those with α-helical transmembrane segments (TMs) and those with β-barrel structures, our research is primarily concerned with α-helical membrane proteins, which make up a much larger fraction of membrane protein genes.

α-helical membrane protein folding can be divided into a two-step process (Popot and Engelman, 1990). In the first step, the membrane protein is cotranslationally inserted via a ribosome-translocon complex, and once the topology is established, the transmembrane helices can fold into a final structure in a second step (Figure 1-1). Once inserted and folded in the membrane, these proteins exist in equilibrium between their unfolded and folded state. The second-stage folding is our primary focus for *in vitro* folding studies (Figure 1-1). While these two steps are not completely separate during insertion, it is reasonable that second-stage folding can inform us about the initial folding of nascent membrane proteins.

One of the major limitations for *in vitro* folding studies is the need for large amounts of purified protein. Attempting to over-express a spectrum of integral membrane proteins has had limited success, and those that do over-express in *E. coli* tend to be smaller proteins with fewer transmembrane segments (Korepanova et al., 2005). The nature of membrane protein expression has an additional set of complexities to those facing soluble proteins stemming from the mechanism of cotranslational insertion. A proteomics study showed that overexpression of several membrane proteins in *E. coli* affected several chaperone and protease systems that overexpression of a soluble protein did not (Wagner et al., 2007). The fact that overexpression of the selected membrane proteins limited expression of other secretory proteins, may indicate that particular secretory chaperone systems and translocation machinery are easily saturated with the overexpression target. In addition to adjusting expression conditions, such as temperature, media, time, and induction levels, progress has been made by engineering bacterial strains specifically tailored to overexpress membrane proteins. Miroux and Walker selected for genomic mutations that improved expression of membrane proteins by limiting their toxic effects (Miroux and Walker, 1996). Later investigations of the genome of those "Walker strains" found that improved expression stemmed from a mutation in the promoter for T7 RNA polymerase (Wagner et al., 2008), leading to a modified strain of BL21(DE3) that can tune T7 RNA polymerase for overexpression of membrane proteins under control of the T7 promoter (Schlegel et al., 2012). The Miroux and Walker selection strategy was further improved upon by Massey-Gendel and colleagues by specifically selecting for mutations that improved the expression of properly trafficked membrane protein up to 75-fold (Massey-Gendel et al., 2009).

Despite advances in overexpression, the purification process can still be disruptive due to the need to extract membrane proteins from cell membranes using solubilizing detergents before incorporating purified protein into a more native environment for *in vitro* studies.

Detergents permit easy manipulation of membrane proteins in solution by masking hydrophobic transmembrane regions with a micellar belt, but without recreating all the structural features of the membrane, such as lateral pressure or topology. Transient dissociation of detergents can expose aggregation-prone hydrophobic regions. Detergent-solubilized membrane proteins are also not constrained to a two-dimensional bilayer, and thus have an entropically favorable unfolded state. Development of amphipathic polymers (amphipols) (Popot, 2010) and tandem facial amphiphiles (Chae et al., 2010) have attempted to alleviate some of these caveats by creating a more stable micelle, but are not typically efficient at solubilizing proteins from the membrane, and thus still must be extracted with solubilizing detergents that are less suitable for long-term stability in solution. Another approach is to make stabilizing mutations in the protein itself (Serrano-Vega et al., 2008), but this of course alters the energy landscape from the wild-type protein.

Cell-free synthesis is a promising alternative that circumvents some of the challenges with cellular overexpression and purification. Expressing membrane proteins in a reconstituted translation system permits incorporation of membrane proteins directly into lipid bilayers (Roos et al., 2013). This strategy eliminates any need to protect against toxicity effects or to purify away from the complex milieu of the cellular membrane. On the other hand, these systems lack the specialized translocation machinery of the cell for membrane protein insertion. Thus cell-free systems rely on finding conditions for which membrane proteins will spontaneously insert properly into a detergent or lipid environment. Cell-free systems have employed exogenous liposomes (Niwa et al., 2015), and the GroEL-ES chaperonin (Chi et al., 2015) to stem misfolding and aggregation. Generating purified membrane proteins in a native bilayer environment is a challenge in itself, and find methods to study how they fold in that environment requires methods tailored to this unique class of proteins.

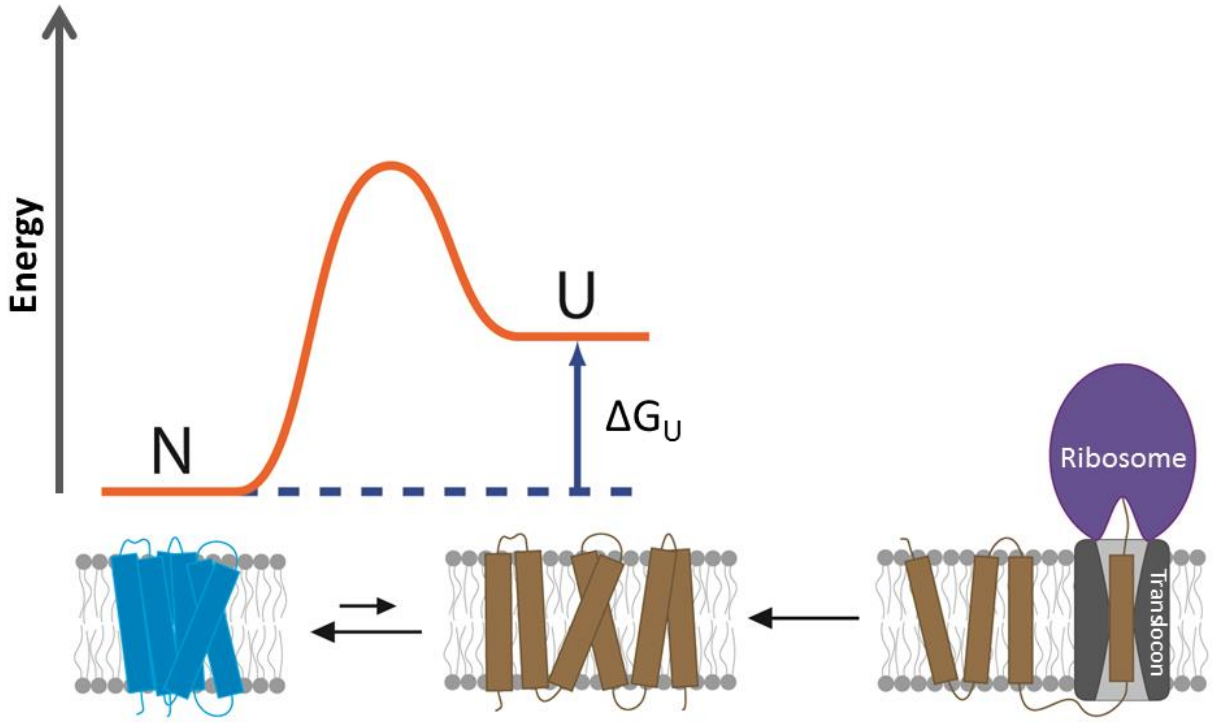


Figure 1-1 Two-stage model of α-helical membrane protein folding. In the first stage, nascent membrane proteins (brown) are cotranslationally inserted into the membrane (gray) via the translocon. In the second-stage, the transmembrane segments associate and fold into the native conformation (blue), existing in an equilibrium between the native state (N) and the inserted unfolded state (U) once dissociated from the translocation machinery. Studies of second-stage membrane protein folding aim to define the kinetics and free energy change of this process.

Methods for Membrane Protein Folding Studies

Chaotropic agents such as urea and guanidine have been commonly used to reversibly shift the folding equilibrium of soluble proteins to the unfolded state. These chemicals have found success with β-barrel membrane proteins that can refold from a soluble denatured state. Urea and guanidine are membrane-compatible and can effectively solubilize β-stranded proteins. The energetics of insertion and folding are coupled in these assays as they have taken advantage of spontaneously refolding outer membrane proteins (Fleming, 2014).

α-helical membrane proteins have not been studied from the completely unfolded state in solution to the folded transmembrane state. They have been recalcitrant to solubilization by chaotropes, and folding studies have focused on the second-stage unfolding to folding transition within the membrane. The denaturing detergent sodium dodecyl sulfate (SDS) has been a useful tool for unfolded helical membrane proteins, while maintaining a hydrophobic environment for transmembrane segments. SDS can be titrated into a micellar phase of non-denaturing detergent to yield mixed micelles with an increasing fraction of SDS (χ_{SDS}). α-helical membrane proteins display a cooperative unfolding curve dependent on χ_{SDS} (Curnow and Booth, 2007; Guo et al., 2016; Lau and Bowie, 1997). Calculating free energies of unfolding in SDS relies on long extrapolation from the transition region where there is a significant portion of unfolded and folded species. Not all helical membrane proteins can reversibly refold from an SDS-denatured state and the need for robust functional assays limits the set of membrane proteins available to study .

To measure second-stage folding of α-helical membrane proteins under native conditions the Bowie lab has developed a method of driving unfolding in a lipid bilayer using a steric trap. This approach utilizes a large protein that binds preferentially to the unfolded state with high affinity. The membrane protein of interest is labeled with biotins in specific sites that are in close proximity in the folded structure, but far apart in the primary sequence, such that a monovalent streptavidin can bind to a single biotin and physically occlude binding of the second biotin unless the protein unfolds. Refolding can be prompted by competing off the streptavidin with free biotin, and the affinity can be modulated through mutations to tune the method for the stability of the membrane protein. The method has been demonstrated with a soluble protein (Blois et al., 2009), a transmembrane helix dimer (Hong and Bowie, 2011; Hong et al., 2010), a trimeric membrane enzyme (Jefferson et al., 2013) (see Chapter 2), a light-driven proton pump (Chang and Bowie, 2014), and an intramembrane protease (Guo et al., 2016). The steric trap

has been used to measure the strength of protein-protein interactions in bilayers, free energies of unfolding, and unfolding kinetics, all without the need for harsh denaturants. Recent advancements in the technique have introduced a generally applicable fluorescence quenching assay to monitor unfolding by steric trapping without exploiting an intrinsic characteristic of the protein for a functional assay.

We have sought to generalize the steric trap method for membrane proteins in lipid vesicles by moving to a single-molecule system. The method measures double-binding of the streptavidin by two-color fluorescence coincidence to detect unfolding and does not rely on the need for a functional assay of the membrane protein of interest. By measuring at the single-molecule level, we also remove any possibility of aggregation and are able to make measurements of membrane protein stability with much less material. These efforts are detailed in Chapter 3.

Another method of studying membrane protein folding is to use mechanical force to drive unfolding. Single-molecule force spectroscopy has come to prominence in recent years. Atomic force microscopy has also been used to study membrane proteins, but these experiments pull the protein orthogonal to the plane of the membrane. We sought to employ force spectroscopy using magnetic tweezers to study the unfolding of a single molecule of the intramembrane protease GlpG. These efforts are detailed in Chapter 4, and efforts to streamline the technique are presented in Chapter 5. Currently available methods for studying membrane protein folding are reviewed in more detail in Chapter 6.

Determinants of Protein Folding and Stability

Though the number of membrane proteins whose folding has been investigated is still relatively low and even fewer measurements have been made under native conditions, the findings from these studies have highlighted some key differences from those for soluble proteins. First of all, membrane proteins are inserted into their particular lipid environment such that the starting point for their second-stage folding is inherently different from that of soluble proteins. That insertion is determined by the favorability of insertion into the membrane from the translocon. Early efforts to predict the thermodynamics of insertion lead a scale calculated by the transfer of amino acids from water to octanol, which serves as a hydrophobic solvent to mimic the hydrophobic core of the membrane (Wimley et al., 1996). This scale can help predict what types of sequences will be committed to insertion and form the TMs of membrane proteins. Unsurprisingly, larger hydrophobic amino acids have more favorable transfer energies and charged amino acids are the least favored for insertion. To check how well this scale predicts insertion of TM sequences in the biological context of the translocon, insertion experiments with full-length TMs were performed in ER microsomes (Hessa et al., 2005). Inserted TMs in a multi-span protein could be detected by glycosylation of a site that was lumenal if translocated or inaccessible in the cytoplasm if not translocated. Placing a single amino acid at the center of a marginally inserting TM offers a method to calculate the transfer free energy of amino acids in a biological context. Comparison with the water-octanol transfer scale show a correlation, with the notable exceptions of tryptophan and proline, which are less likely to insert in the biological context. Additionally, the system allows the study of how positioning of various residues within the TM segment affect insertion. While the position of leucine and phenylalanine does not affect insertion much, the placement of tyrosine and tryptophan towards the interfacial region of the membrane enhances TM insertion. These

results highlight the importance of protein-lipid interactions and the need to study membrane protein folding in the context of the membrane.

Once TMs are inserted, they can fold into their final native conformation, but that process is heavily influenced by the unique context of the membrane. One of the major driving forces of soluble protein folding is collapse of the hydrophobic interior due to the hydrophobic effect, however, in the membrane, while hydrophobicity is a factor for insertion, there is no hydrophobic effect for tertiary folding. This leaves hydrogen bonding and van der Waals packing as the determinants of tertiary membrane protein folding. Ionic bonds are rare due to the dearth of charged residues that end up inserting into the membrane, though hydrogen bonding is prevalent among the backbone and side-chain residues of membrane proteins. It might be expected that hydrogen bonding would play a large role in the low dielectric of the membrane, and the introduction of polar residues are commonly associated with membrane protein misfolding diseases (Partridge et al., 2002, 2004). The actual thermodynamic contribution of side-chain hydrogen bonds was unknown until Joh and colleagues used a double-mutant cycle stability analysis to measure the strength of hydrogen bonds in bacteriorhodopsin while correcting for the background destabilization of individual mutations (Joh et al., 2008). Further studies of membrane protein hydrogen bond strength have shown that straightening of kinked transmembrane helices have a surprisingly modest energetic cost to this conformational shift that breaks traditional i → i+4 backbone hydrogen bonds (Cao and Bowie, 2012). The small energetic cost appeared to be a result of the hydrogen bond network forming compensatory non-canonical i → i+3 hydrogen bonds, illustrating that membrane proteins are held together by a large network of relatively weak interactions that can shift to accommodate disruptive mutations. A point that is highlighted by the shifting hydrogen bond network of the Ca+ ATPase through its functional cycle (Cao and Bowie, 2012).

Without significant contributions from the hydrophobic effect or hydrogen bonds, van der Waals interactions are a major determinant of second-stage membrane protein folding. Tertiary folding of membrane proteins is built upon the association of transmembrane helices. The dimerization of the glycophorin A transmembrane domain has served as a useful model for this fundamental interaction. The glycophorin A dimer features a GxxxG motif (MacKenzie et al., 1997), also known as a glycine zipper, a common structural element of transmembrane helix-helix interactions (Kim et al., 2005) that permits close van der Waals packing between helices. Using an equilibrium sedimentation assay in detergent on a wide array of mutants has teased apart the packing contributions of individual amino acids, and double mutant cycles have revealed a complex network of energetically coupled residues, the strongest of which lie at the distal ends of the helix (Doura and Fleming, 2004; Doura et al., 2004; Fleming, 2002; Fleming and Engelman, 2001; Fleming et al., 1997). While glycophorin A is not a complex multipass membrane protein, its simplicity as a model system has helped break down the fundamental requirements for association of transmembrane segments.

Membrane protein folding is not only governed by the primary sequence of amino acids, but is also influenced by the lateral packing pressure of the bilayer itself. Having developed a reversible folding assay for OmpA, Hong and coworkers measured the effect of membrane thickness and chain saturation on protein stability (Hong and Tamm, 2004). For saturated and mono-unsaturated acyl chains of phosphocholine lipids, there exists a strong correlation between increasing hydrophobic thickness and greater free energy of unfolding. However, for cis-double-unsaturated chains, stability is greatest for short chain lengths. Shorter chain unsaturated lipids have more negative curvature stress, which can be relieved by the insertion and folding of the hourglass-shaped OmpA into the bilayer. In addition to folding, certain proteins respond to changes in the bilayer for their function. Specifically, the mechanosensitive channels sense changes in membrane tension, opening a pore to release water under high

intracellular pressure, a common mechanism among prokaryotic and eukaryotic cells (Perozo et al., 2002; Vásquez et al., 2008). Again, the effects of lateral packing pressure underline the importance of the membrane environment and how the acyl chain length and shape of lipids go hand in hand with protein folding and function.

Beyond the bulk features of the lipid bilayer, membrane proteins also have specific associations with their immediate surrounding ring of lipid molecules, referred to as the annular ring. These lipids are analogous to the solvent shell for soluble proteins. Ordered lipids packing next to membrane proteins were observed in some of the earliest crystal structures (Lee, 2003). The presence of associated native lipids in these structures that are often crystallized in the presence of detergent with few lipid molecules likely indicates that the crystallized lipids are tightly bound in native conditions. A review of membrane protein crystal structure summarized the ordered lipids found within 5 Å of the protein surface (Wiener, 2005). Unsurprisingly, lipids tend to occur near hydrophobic residues. However, these lipids also exhibit a range of unusual acyl chain conformations not seen in pure lipid crystal packing, such as trans-gauche isomerization. In the context of the membrane protein surface, these alternate conformations are energetically favorable and can mediate protein-protein interfaces between subunits. Beyond crystal structures, recent molecular dynamics simulations can aid in elucidating brief, but highly specific, lipid interactions, such as the association of cardiolipin with the c-ring of ATP synthase (Duncan et al., 2016). Many membrane proteins are known to have their function altered by specific lipids, but the nature of where and how those lipids bind is often unclear. To further expand our understanding of membrane protein folding in biologically relevant conditions, we need methods for studying this fundamental process under conditions in which we can manipulate the specific lipid composition and bilayer features.

Further understanding of membrane protein folding relies on developing methods for investigating folding in the native lipid bilayer and interrogating a broad array of targets for a

comprehensive view of how this important class of proteins forms within the complex membrane environment. The steric trap method and single-molecule pulling experiments discussed herein aim to address these challenges and open up new avenues of membrane protein folding studies.

References

Balch, W.E., Morimoto, R.I., Dillin, A., and Kelly, J.W. (2008). Adapting Proteostasis for Disease Intervention. Science *319*, 916–919.

Blois, T.M., Hong, H., Kim, T.H., and Bowie, J.U. (2009). Protein Unfolding with a Steric Trap. J Am Chem Soc *131*, 13914–13915.

Bradley, P., Misura, K.M.S., and Baker, D. (2005). Toward High-Resolution de Novo Structure Prediction for Small Proteins. Science *309*, 1868–1871.

Cao, Z., and Bowie, J.U. (2012). Shifting hydrogen bonds may produce flexible transmembrane helices. PNAS *109*, 8121–8126.

Chae, P.S., Gotfryd, K., Pacyna, J., Miercke, L.J.W., Rasmussen, S.G.F., Robbins, R.A., Rana, R.R., Loland, C.J., Kobilka, B., Stroud, R., et al. (2010). Tandem Facial Amphiphiles for Membrane Protein Stabilization. J. Am. Chem. Soc. *132*, 16750–16752.

Chang, Y.-C., and Bowie, J.U. (2014). Measuring membrane protein stability under native conditions. Proc. Natl. Acad. Sci. U.S.A. *111*, 219–224.

Chi, H., Wang, X., Li, J., Ren, H., and Huang, F. (2015). Folding of newly translated membrane protein CCR5 is assisted by the chaperonin GroEL-GroES. Sci Rep *5*.

Cohen, F.E., and Kelly, J.W. (2003). Therapeutic approaches to protein-misfolding diseases. Nature *426*, 905–909.

Curnow, P., and Booth, P.J. (2007). Combined kinetic and thermodynamic analysis of α-helical membrane protein unfolding. Proceedings of the National Academy of Sciences *104*, 18970–18975.

Doura, A.K., and Fleming, K.G. (2004). Complex Interactions at the Helix–Helix Interface Stabilize the Glycophorin A Transmembrane Dimer. Journal of Molecular Biology *343*, 1487–1497.

Doura, A.K., Kobus, F.J., Dubrovsky, L., Hibbard, E., and Fleming, K.G. (2004). Sequence Context Modulates the Stability of a GxxxG-mediated Transmembrane Helix–Helix Dimer. Journal of Molecular Biology *341*, 991–998.

Duncan, A.L., Robinson, A.J., and Walker, J.E. (2016). Cardiolipin binds selectively but transiently to conserved lysine residues in the rotor of metazoan ATP synthases. PNAS *113*, 8687–8692.

Fleming, K.G. (2002). Standardizing the Free Energy Change of Transmembrane Helix–Helix Interactions. Journal of Molecular Biology *323*, 563–571.

Fleming, K.G. (2014). Energetics of Membrane Protein Folding. Annual Review of Biophysics *43*, 233–255.

Fleming, K.G., and Engelman, D.M. (2001). Specificity in transmembrane helix–helix interactions can define a hierarchy of stability for sequence variants. PNAS *98*, 14340–14344.

Fleming, K.G., Ackerman, A.L., and Engelman, D.M. (1997). The effect of point mutations on the free energy of transmembrane α-helix dimerization1. Journal of Molecular Biology *272*, 266–275.

Guo, R., Gaffney, K., Yang, Z., Kim, M., Sungsuwan, S., Huang, X., Hubbell, W.L., and Hong, H. (2016). Steric trapping reveals a cooperativity network in the intramembrane protease GlpG. Nat Chem Biol *12*, 353–360.

Hammarström, P., Wiseman, R.L., Powers, E.T., and Kelly, J.W. (2003). Prevention of Transthyretin Amyloid Disease by Changing Protein Misfolding Energetics. Science *299*, 713–716.

Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S.H., and von Heijne, G. (2005). Recognition of transmembrane helices by the endoplasmic reticulum translocon. Nature *433*, 377–381.

Hong, H., and Bowie, J.U. (2011). Dramatic Destabilization of Transmembrane Helix Interactions by Features of Natural Membrane Environments. J. Am. Chem. Soc. *133*, 11389–11398.

Hong, H., and Tamm, L.K. (2004). Elastic coupling of integral membrane protein stability to lipid bilayer forces. PNAS *101*, 4065–4070.

Hong, H., Blois, T.M., Cao, Z., and Bowie, J.U. (2010). Method to measure strong protein–protein interactions in lipid bilayers using a steric trap. Proceedings of the National Academy of Sciences *107*, 19802–19807.

Jefferson, R.E., Blois, T.M., and Bowie, J.U. (2013). Membrane Proteins Can Have High Kinetic Stability. J. Am. Chem. Soc. *135*, 15183–15190.

Joh, N.H., Min, A., Faham, S., Whitelegge, J.P., Yang, D., Woods, V.L., and Bowie, J.U. (2008). Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. Nature *453*, 1266–1270.

Joh, N.H., Wang, T., Bhate, M.P., Acharya, R., Wu, Y., Grabe, M., Hong, M., Grigoryan, G., and DeGrado, W.F. (2014). De novo design of a transmembrane Zn2+-transporting four-helix bundle. Science *346*, 1520–1524.

Kim, B.L., Schafer, N.P., and Wolynes, P.G. (2014). Predictive energy landscapes for folding α-helical transmembrane proteins. PNAS *111*, 11031–11036.

Kim, S., Jeon, T.-J., Oberai, A., Yang, D., Schmidt, J.J., and Bowie, J.U. (2005). Transmembrane glycine zippers: Physiological and pathological roles in membrane proteins. PNAS *102*, 14278–14283.

King, N.P., Sheffler, W., Sawaya, M.R., Vollmar, B.S., Sumida, J.P., André, I., Gonen, T., Yeates, T.O., and Baker, D. (2012). Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy. Science *336*, 1171–1174.

Koga, N., Tatsumi-Koga, R., Liu, G., Xiao, R., Acton, T.B., Montelione, G.T., and Baker, D. (2012). Principles for designing ideal protein structures. Nature *491*, 222–227.

Korendovych, I.V., Senes, A., Kim, Y.H., Lear, J.D., Fry, H.C., Therien, M.J., Blasie, J.K., Walker, F.A., and DeGrado, W.F. (2010). De Novo Design and Molecular Assembly of a Transmembrane Diporphyrin-Binding Protein Complex. J. Am. Chem. Soc. *132*, 15516–15518.

Korepanova, A., Gao, F.P., Hua, Y., Qin, H., Nakamoto, R.K., and Cross, T.A. (2005). Cloning and expression of multiple integral membrane proteins from Mycobacterium tuberculosis in Escherichia coli. Protein Science *14*, 148–158.

Kuhlman, B., Dantas, G., Ireton, G.C., Varani, G., Stoddard, B.L., and Baker, D. (2003). Design of a Novel Globular Protein Fold with Atomic-Level Accuracy. Science *302*, 1364–1368.

Lai, Y.-T., Cascio, D., and Yeates, T.O. (2012). Structure of a 16-nm Cage Designed by Using Protein Oligomers. Science *336*, 1129–1129.

Lau, F.W., and Bowie, J.U. (1997). A Method for Assessing the Stability of a Membrane Protein†. Biochemistry *36*, 5884–5892.

Lee, A.G. (2003). Lipid–protein interactions in biological membranes: a structural perspective. Biochimica et Biophysica Acta (BBA) - Biomembranes *1612*, 1–40.

MacKenzie, K.R., Prestegard, J.H., and Engelman, D.M. (1997). A Transmembrane Helix Dimer: Structure and Implications. Science *276*, 131–133.

Massey-Gendel, E., Zhao, A., Boulting, G., Kim, H.-Y., Balamotis, M.A., Seligman, L.M., Nakamoto, R.K., and Bowie, J.U. (2009). Genetic selection system for improving recombinant membrane protein expression in E. coli. Protein Science *18*, 372–383.

Miroux, B., and Walker, J.E. (1996). Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. Journal of Molecular Biology *260*, 289–298.

Niwa, T., Sasaki, Y., Uemura, E., Nakamura, S., Akiyama, M., Ando, M., Sawada, S., Mukai, S., Ueda, T., Taguchi, H., et al. (2015). Comprehensive study of liposome-assisted synthesis of membrane proteins using a reconstituted cell-free translation system. Sci Rep *5*.

Partridge, A.W., Therien, A.G., and Deber, C.M. (2002). Polar mutations in membrane proteins as a biophysical basis for disease. Biopolymers *66*, 350–358.

Partridge, A.W., Therien, A.G., and Deber, C.M. (2004). Missense mutations in transmembrane domains of proteins: Phenotypic propensity of polar residues for human disease. Proteins *54*, 648–656.

Perez-Aguilar, J.M., and Saven, J.G. (2012). Computational Design of Membrane Proteins. Structure *20*, 5–14.

Perozo, E., Cortes, D.M., Sompornpisut, P., Kloda, A., and Martinac, B. (2002). Open channel structure of MscL and the gating mechanism of mechanosensitive channels. Nature *418*, 942–948.

Piana, S., Lindorff-Larsen, K., and Shaw, D.E. (2012). Protein folding kinetics and thermodynamics from atomistic simulation. PNAS *109*, 17845–17850.

Popot, J.-L. (2010). Amphipols, Nanodiscs, and Fluorinated Surfactants: Three Nonconventional Approaches to Studying Membrane Proteins in Aqueous Solutions. Annual Review of Biochemistry *79*, 737–775.

Popot, J.L., and Engelman, D.M. (1990). Membrane protein folding and oligomerization: the two-stage model. Biochemistry *29*, 4031–4037.

Roos, C., Kai, L., Proverbio, D., Ghoshdastider, U., Filipek, S., Dötsch, V., and Bernhard, F. (2013). Co-translational association of cell-free expressed membrane proteins with supplied lipid bilayers. Molecular Membrane Biology *30*, 75–89.

Sanders, C.R., and Myers, J.K. (2004). DISEASE-RELATED MISASSEMBLY OF MEMBRANE PROTEINS. Annual Review of Biophysics and Biomolecular Structure *33*, 25–51.

Schlegel, S., Löfblom, J., Lee, C., Hjelm, A., Klepsch, M., Strous, M., Drew, D., Slotboom, D.J., and de Gier, J.-W. (2012). Optimizing Membrane Protein Overexpression in the Escherichia coli strain Lemo21(DE3). Journal of Molecular Biology *423*, 648–659.

Serrano-Vega, M.J., Magnani, F., Shibata, Y., and Tate, C.G. (2008). Conformational thermostabilization of the β1-adrenergic receptor in a detergent-resistant form. Proceedings of the National Academy of Sciences *105*, 877–882.

Vásquez, V., Sotomayor, M., Cordero-Morales, J., Schulten, K., and Perozo, E. (2008). A Structural Mechanism for MscS Gating in Lipid Bilayers. Science *321*, 1210–1214.

Wagner, S., Baars, L., Ytterberg, A.J., Klussmeier, A., Wagner, C.S., Nord, O., Nygren, P.-A., van Wijk, K.J., and de Gier, J.-W. (2007). Consequences of membrane protein overexpression in Escherichia coli. Mol. Cell Proteomics *6*, 1527–1550.

Wagner, S., Klepsch, M.M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Högbom, M., Van Wijk, K.J., Slotboom, D.J., and Persson, J.O. (2008). Tuning Escherichia coli for membrane protein overexpression. Proceedings of the National Academy of Sciences *105*, 14371–14376.

Wiener, M.C. (2005). A Census of Ordered Lipids and Detergents in X-ray Crystal Structures. In Protein-Lipid Interactions: From Membrane Domains to Cellular Networks, (Wiley-VCH Verlag GmbH & Co. KGaA), pp. 97–117.

Wimley, W.C., Creamer, T.P., and White, S.H. (1996). Solvation Energies of Amino Acid Side Chains and Backbone in a Family of Host–Guest Pentapeptides. Biochemistry *35*, 5109–5124.

Chapter 2

Membrane Proteins Can Have High Kinetic Stability

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

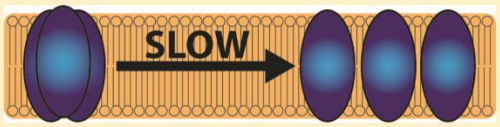
Membrane Proteins Can Have High Kinetic Stability

Robert E. Jefferson, Tracy M. Blois,[‡] and James U. Bowie*

[†]Department of Chemistry and Biochemistry, University of California, Los Angeles-Department of Energy Institute for Genomics and Proteomics, Molecular Biology Institute, University of California, Los Angeles, Los Angeles, California 90095, United States

Supporting Information

ABSTRACT: Approximately 10% of water-soluble proteins are considered kinetically stable with unfolding half-lives in the range of weeks to millenia. These proteins only rarely sample the unfolded state and may never unfold on their respective biological time scales. It is still not known whether membrane proteins can be kinetically stable, however. Here we examine the subunit dissociation rate of the trimeric membrane enzyme, diacylglycerol kinase, from *Escherichia coli* as a proxy for complete unfolding. We find that dissociation occurs with a half-life of at least several weeks, demonstrating that membrane proteins can remain locked in a folded state for long periods of time. These results reveal that evolution can use kinetic stability to regulate the biological function of membrane proteins, as it can for soluble proteins. Moreover, it appears that the generation of kinetic stability could be a viable target for membrane protein engineering efforts.



INTRODUCTION

Some proteins unfold extremely slowly. Wilfredo Colon's lab has identified a large collection of kinetically stable proteins by looking for resistance to sodium dodecyl sulfate (SDS) denaturation. These proteins unfold with apparent half-lives ranging from 79 days to 346 years.[1,2] Pilus protein complexes may have the record for slow unfolding with an estimated half-life of 3 billion years.[3] Most kinetically stable proteins are also thermodynamically stable, but not always. For example, folding of α-lytic protease is catalyzed and driven by a pro-region. Once the pro-region is cleaved off, the enzyme is thermodynamically unstable but remains locked in the folded state because the unfolding half-life is about 1.2 years.[4] Characteristics of kinetically stable water-soluble proteins include a high degree of rigidity, considerable β-sheet structure, and a dearth of monomers.[1,2]

It remains unknown whether the different folding energetics or topology restrictions in the membrane could allow for high kinetic stability. It is particularly questionable for α-helical membrane proteins since almost all the known kinetically stable proteins contain β-sheets, perhaps because of their high contact order.[1] Indeed, the best indication of kinetic stability in membrane proteins comes from unfolding rate studies of the β-barrel porin PagP.[5] For this protein, unfolding rates could be measured at urea concentrations above 8.5 M. Extrapolation back to zero denaturant predicts an unfolding half-life for PagP of more than half a year. Whether the long extrapolation is valid is unclear, however.

There are some indications that helical membrane proteins can be kinetically stable. Yinan Wei's lab found that upon mixing or coexpression of distinguishable subunits of the trimeric membrane protein AcrB, a nonequilibrium distribution is found.[6] This suggests that the oligomers must not exchange completely over the hours needed to express and analyze them. Subunit exchange of dimeric EmrE was also found to take many hours under native conditions.[7] Even the simple glycophorin A transmembrane helix dimer is known to require hours to exchange in certain detergents.[8] A number of membrane proteins have been found to be resistant to SDS denaturation[9−15] and by analogy with Wilfredo Colon's experiments on soluble proteins, this might reflect high kinetic stability. On the other hand, it might also simply indicate high thermodynamic stability as the denaturing power of SDS is likely to be much greater for soluble proteins than membrane proteins, which are already coated by a band of detergent. Extrapolation of SDS-driven unfolding of bacteriorhodopsin to zero SDS suggest a remarkable unfolding half-life of ~20 million years,[16] but extrapolations for SDS unfolding rates are particularly uncertain.

Given the doubts inherent in extrapolating from high denaturant concentrations, it would be ideal to examine unfolding rates under native conditions. To this end, we examined the subunit dissociation kinetics of diacylglycerol kinase (DGK) from *Escherichia coli* as a proxy for unfolding rate. DGK is an obligate trimer with three transmembrane and one amphipathic helix per subunit. A recent crystal structure of the enzyme reveals a structure in which the nine transmembrane helices of the trimer are closely packed around a central axis (Figure 1A).[17] An earlier NMR structure showed a domain-swapped architecture in which the C-terminal transmembrane helix shifts over to an adjacent subunit.[18] It seems clear that the crystal structure is a fully active enzyme, but it is not known if the domain-swapped form can also be active. Subunit mixing experiments[19] and both structures show that the three active sites are shared between subunits. Thus, a monomer is necessarily inactive.

Received: July 20, 2013
Published: September 13, 2013

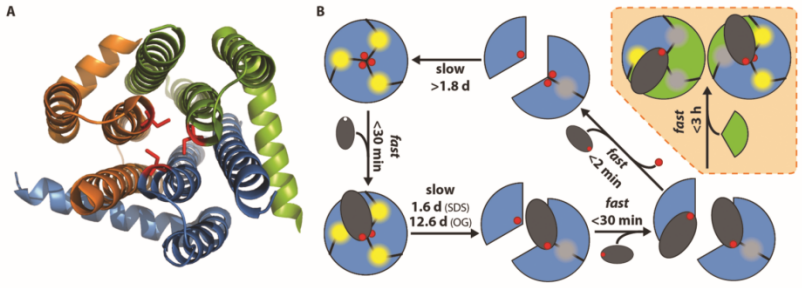


Figure 1. Steric trapping of DGK. (A) Crystal structure of DGK,[17] highlighting the single cysteine introduced at position 53 for biotin labeling. The three distinct subunits are shown in orange, green, and blue. (B) Simple schematic for unfolding and refolding of DGK by the steric trap. The evidence for the reaction scheme investigated in 0.2 X_{SDS} is presented in the text, and the results will only be summarized here. The upper left depicts the DGK trimer. The active sites, depicted in yellow, are shared between subunits. The biotin labels are depicted by the red dots. Initial binding of mSA, depicted in dark gray, is unimpeded and can occur rapidly. The binding of a second mSA cannot occur unless the subunits dissociate due to steric overlap with the initially bound mSA. In 0.2 X_{SDS}, the half-life is 1.6 d, while in pure OG, the half-life of dissociation is at least 12.6 d. Once dissociated, a second mSA can bind rapidly, effectively trapping the protein in a dissociated state, hence the term "steric trapping". If the trapped state is generated with a rapidly dissociating variant of mSA, mSA_{S45A}, it can be rapidly removed by the addition of excess free biotin. The dissociated subunits are then free to reassociate. Reassociation is slow, with a half-life of 1.8 d and is not concentration dependent, suggesting a unimolecular conformational change is rate limiting. The conformational change is depicted by a change in shape of the subunits and inactivation of the active sites by the change from yellow to light gray. Evidence for subunit dissociation during steric trapping is provided by the experiment depicted in the shaded box. In this experiment, an unfolded, inactive mutant shown in green was added to a sterically trapped sample. Rapid recovery of activity was observed due to reconstruction of the folded trimers as shown.

We posit that subunit dissociation must either precede or be concomitant with complete unfolding. In particular, if the subunits can remain folded as monomers, then dissociation precedes unfolding and the dissociation rate reports an upper limit of the unfolding rate. If local unfolding can occur within the intact oligomer, then the rate of subunit dissociation reports the rate of unfolding of a folded oligomeric core so that subunit dissociation is concomitant with complete unfolding. Thus, the dissociation rate provides an upper bound on the overall unfolding rate.

To examine dissociation/unfolding rates under native conditions we took two approaches: steric trapping and subunit exchange measurements. We find that subunit dissociation can be extremely slow, on the order of weeks for DGK, consistent with high kinetic stability of the trimer.

■ MATERIALS AND METHODS

Materials. The pQE80 vector and Ni-NTA resin were obtained from QIAGEN (Valencia, CA). Anagrade *n*-octyl-*β*-D-glucopyranoside (OG) and isopropyl-*β*-D-thiogalactopyranoside (IPTG) were purchased from Affymetrix (Santa Clara, CA). *E. coli* polar lipid extract was purchased from Avanti Polar Lipids, Inc. (Alabaster, AL). N^{α}-(3-propionylmaleimidyl)biocytin (BM) was purchased from Life Technologies (Grand Island, NY).

Cloning, Expression, and Purification of Proteins. The natural Cys residues in the DGK expression plasmid pSD005[20] were changed to Ala using site-directed mutagenesis (QuikChange, Agilent) to create a cysless His_6-DGK construct ($DGK_{Cysless}$). Ile 53 of $DGK_{Cysless}$ was mutated to Cys to create the single-cysteine mutant $DGK_{Cysless/I53C}$. Five glutamate residues were added to $DGK_{Cysless}$ by inserting the sequence, 5′-CGAAGAGAAGAAGAAGAGCT-3′, at the SacI site (ELEEEEE) following the N-terminal His_6-tag to create E_5-$DGK_{Cysless}$. $DGK_{Cysless}$, $DGK_{Cysless/I53C}$, and E_5- $DGK_{Cysless}$ were expressed in WH1061 cells at 37 °C for 3 h after induction with 1 mM IPTG.

For the coexpression of cysless and charged DGK, the genes were subcloned into pQE80 and pBAD/HisA/p15A,[21] respectively. Each vector was modified by insertion of a *Spe*I site (ACTAGT) after the start codon by site-directed mutagenesis. The His_6-tag coding region of pQE80 was replaced with ACTAGT to give pQE80/SpeI. Insertion of $DGK_{Cysless}$ using *Spe*I and *Hin*dIII restriction sites added residues MRGSTS to the N-terminus of the pSD005 construct to give pQ-DGK_{Cyless}. ACTAGT was inserted after the fourth encoded amino acid of pBAD/HisA/p15A just before the His_6 tag to give pBAD/HisA/p15A/SpeI. Insertion of E_5-$DGK_{Cysless}$ using *Spe*I and *Hin*dIII restriction sites added residues MGGSTS to the N-terminus of the pSD005 construct to give pB-E_5-$DGK_{Cysless}$. pQ-$DGK_{Cysless}$ and pB-E_5-$DGK_{Cysless}$ were coexpressed in TOP10 cells at 18 °C overnight after induction with 0.1 mM IPTG and 0.2% arabinose. All DGK proteins were purified in 1.5% OG as previously described.[22]

DGK_{E76L} was created by site-directed mutagenesis of the wild-type DGK in pSD005. To prepare for refolding experiments, DGK_{E76L} purified in 1.5% OG was passed through an Econo-Pac 10DG desalting column (Bio-Rad, Hercules, CA) that had been previously equilibrated with SDS unfolding buffer (50 mM $NaPO_4$ (pH 7.5), 50 mM NaCl, 1% SDS). The peak fraction was diluted to 40 μM with SDS unfolding buffer, snap frozen in liquid nitrogen, and stored at −80 °C.

Monovalent streptavidin variants containing one active subunit per streptavidin tetramer were prepared as described previously.[23]

Chemical Biotinylation of DGK. Biotin labeling was performed by incubating 162 μM $DGK_{Cysless/I53C}$ in 50 mM $NaPO_4$ (pH 7.5), 50 mM NaCl, 1.5% OG, 1 mM TCEP with the addition of solid BM to a final concentration of 9 mM. Samples were incubated for 1 h at room temperature followed by 2 h at 4 °C. Labeling reactions were centrifuged at 14000*g* for 10 min to remove any undissolved label, and the supernatant was applied to 400 μL of Ni-NTA resin. Biotinylated DGK (BM-$DGK_{Cysless/I53C}$) samples were incubated with the resin for 1.5 h at 4 °C and then washed with 30 mM imidazole in 50 mM $NaPO_4$ (pH 7.5), 300 mM NaCl, 1.5% OG, 5 mM β-ME and eluted with 300 mM imidazole in the same buffer. All elution fractions were pooled and then dialyzed against 250 mL of 50 mM $NaPO_4$ (pH 7.5), 50 mM NaCl, 1.5% OG, 1 mM TCEP. Labeled samples were flash frozen in small aliquots using liquid nitrogen and stored at −80 °C. A fresh aliquot was used for all subsequent experiments. Labeling was confirmed by mass spectrometry analysis. Samples were coprecipitated with matrix on a stainless steel MALDI plate (Life Technologies) in a 1:1 (v:v) ratio of protein sample to sinapinic acid matrix solution (saturated solution in 50% acetonitrile, 0.1% trifluoroacetic acid). The dried spot was desalted by washing with 1 μL of water three times. One microliter of sinapinic acid matrix solution was added to the spot

and allowed to dry before acquiring MALDI-MS data on a Voyager DE-STR time-of-flight (TOF) mass spectrometer (Applied Biosystems, Foster City, CA). Mass spectra were collected in linear mode with 1000 shots per spectrum. Spectra were processed using Voyager Data Explorer (Applied Biosystems).

DGK Steric Trap Assays. BM-DGK$_{Cysless/I53C}$ was incubated in 10-fold molar excess mSA or 20-fold molar excess mSA$_{S45A}$ at room temperature in DGK incubation buffer (50 mM $NaPO_4$ (pH 7.5), 50 mM NaCl, 1.5% OG) with 0.2 X_{SDS} (calculated from total bulk concentrations of OG and SDS). To equilibrate DGK and mSA in 0.2 X_{SDS} before steric trapping, the proteins were equilibrated in separate tubes in DGK incubation buffer with 0.2 X_{SDS} for up to 2 h at room temperature. The mSA stock solution in 20 mM $NaPO_4$ (pH 7.0) was diluted with 5x DGK incubation buffer with 1.85% SDS, and then added to DGK after equilibration. At increasing time points, 5 μL aliquots were removed and assayed for activity by the addition of 45 μL of a standard DGK assay mix with either cardiolipin or DMPC as the lipid cofactor as described previously.[22,24] In cases where the DGK concentration was too high to be measured before all the NADH was consumed in the coupled assay, the samples were diluted to 0.2−0.4 μM just prior to the assay. Activity at each time point was normalized to activity of a freshly thawed wild-type DGK aliquot in DGK incubation buffer assayed with the same DGK assay mix. Initial activities of BM-DGK$_{Cysless/I53C}$ after equilibration in 0.2 X_{SDS} were 89−97% as active as the freshly thawed wild-type DGK. Fractional activity was calculated as the fraction of initial normalized activity. Activity plots were fit to either a single or double exponential decay.

For DGK activity recovery experiments after steric trapping, free biotin dissolved in DMSO was diluted 50-fold into aliquots of BM-DGK$_{Cysless/I53C}$ samples trapped with a 20-fold molar excess of mSA$_{S45A}$ to give a 1000-fold molar excess of free biotin over BM-DGK$_{Cysless/I53C}$ subunits. Activity plots for recovered samples were fit to an exponential decay subtracted from the exponential decay of the control sample without mSA.

For reactivation experiments of sterically trapped BM-DGK$_{Cysless/I53C}$ in the presence of an excess of inactive mutant subunits, 0.2 μM BM-DGK$_{Cysless/I53C}$ was first sterically trapped in the presence of 4 μM mSA$_{S45A}$ for 95 h. Two μL of 40 μM SDS-unfolded DGK$_{E76L}$ or SDS unfolding buffer was then added to 80 μL aliquots of sterically trapped DGK$_{Cysless/I53C}$ prior to the addition of 2 μL of 8 mM biotin. Five μL of sample was assayed after an additional 3 h incubation at room temperature. For samples with added biotin, mutant, or buffer, activities were multiplied by a scale factor to account for the slight dilution.

Size-Exclusion Chromatography Binding Experiments. Gel filtration of DGK:mSA complexes was performed on a Superdex 200−10/300 column (GE Healthcare Life Sciences, Pittsburgh, PA). The column was equilibrated with DGK incubation buffer for all runs. BM-DGK$_{Cysless/I53C}$ trimer was first purified on the same column prior to performing the binding experiment to eliminate protein aggregates. 0.32 μM purified BM-DGK$_{Cysless/I53C}$ was incubated with 1 μM mSA for 30 min at room temperature, and free biotin was added to a final concentration of 100 μM to prevent any further binding by mSA. Five hundred microliter of 1 μM mSA alone, 99 μM BM-DGK$_{Cysless/I53C}$ alone or the incubated complex were loaded onto the column separately. 500 μL fractions were collected and assayed for DGK activity.

Preparation of Proteoliposomes. *E. coli* polar lipid extract dissolved in chloroform was dried in glass tubes under a stream of argon gas and residual solvent was removed by vacuum desiccation for 2 h. In 8 mL of 6% OG in DGK vesicle buffer (20 mM Tris-HCl (pH 8.0), 50 mM NaCl) was dissolved 70−144 mg of dried lipids by vortexing and bath sonication. After solubilization, glycerol was added to a final concentration of 20% (v/v). 270 nmol of purified DGK$_{Cysless}$, E_5-DGK$_{Cysless}$, or both were added with 1.5% OG in DGK vesicle buffer to bring the total volume to 16 mL. The mixture was incubated on ice for 30 min before dialysis against DGK vesicle buffer with 20% glycerol over a 24 h period at 4 °C (Spectra/Por 25 kDa cutoff dialysis membrane). After formation of vesicles, the turbid mixture was dialyzed against DGK vesicle buffer to remove glycerol.

The resulting proteoliposomes were purified from aggregated protein by centrifugation over a sucrose cushion. To the proteoliposome samples was added 75% sucrose to a final concentration of 20% and the solution layered over 50% sucrose. The samples were centrifuged at 72000*g* for 16 h at 4 °C. The turbid top layer of proteoliposomes was removed and diluted 2-fold in DGK vesicle buffer before being pelleted twice at 100000*g* for 1.5 h at 4 °C. The proteoliposomes were resuspended in 2 mL of DGK vesicle buffer and passed through Nucleopore track-etched membranes (Whatman) with 200 nm pore sizes using a mini-extruder (Avanti Polar Lipids) 3−5 times. Extruded proteoliposomes were incubated at room temperature overnight to allow recovery of DGK activity.

Anion-Exchange Chromatography of DGK Trimers. For subunit exchange assays in detergent, 500 μL of 70 μM DGK samples, were loaded onto a 1-mL HiTrap Q column (GE Healthcare Life Sciences) pre-equilibrated with 20 mM Tris-HCl (pH 8.0), 1.5% OG. Protein was eluted with an isocratic gradient from 0 to 0.5 M NaCl over 20 mL.

For subunit exchange assays in vesicles, aliquots of proteoliposomes containing DGK$_{Cysless}$, E_5-DGK$_{Cysless}$, or both were incubated at 37 °C. At each time point, aliquots of proteoliposomes containing the same level of DGK activity (activity assayed at t_0) were solubilized with 4−6% OG in DGK vesicle buffer to a final concentration of 3% OG. Anion-exchange chromatography was carried out as in detergent, except the isocratic gradient was extended to 30 mL. Three hundred microliter fractions were collected and assayed for activity.

RESULTS

Monitoring DGK Dissociation Kinetics by Steric Trapping. To monitor dissociation kinetics, we employed a method called steric trapping.[25−27] The steric trapping concept is outlined in Figure 1. In this approach, we label the protein subunits with biotin at sites that are very close in the folded oligomer. When monovalent streptavidin (mSA) is added to the labeled protein, it can bind to one biotinylated subunit with high affinity, but if the biotin labels are sufficiently close, a second mSA cannot bind due to steric overlap, unless the protein dissociates. When the protein dissociates, an additional mSA can bind, effectively trapping the protein in the dissociated state (Figure 1B).

To apply steric trapping to DGK, we introduced a unique Cys residue, I53C, close to the trimeric axis of symmetry (Figure 1A) into a variant of DGK in which the wild-type Cys residues were removed (DGK$_{Cysless}$). Position 53 was then biotinylated using thiol-reactive N^{α}-(3-propionylmaleimidyl) biocytin (BM) to generate BM-DGK$_{Cysless/I53C}$. A MALDI-TOF mass spectrum indicated that ∼58% of the subunits contained a single biotin label and ∼13% contained two labels due to labeling at a secondary site. Thus, ∼71% of the subunits bore a label at position 53. Assuming a perfect binomial distribution of labels, ∼80% of the trimers would have 2 or 3 biotin labels at C53 and could be subject to steric trapping. As the active site of DGK is shared between subunits, dissociation upon the addition of mSA can be conveniently monitored by loss of activity.

To investigate the unfolding rate of DGK, mSA was added to BM-DGK$_{Cysless/I53C}$ in OG micelles at room temperature. As shown in Figure 2A, BM-DGK$_{Cysless/I53C}$ inactivates very slowly (half-life ≈ 12.6 d) in the absence of mSA and the addition of mSA has essentially no effect on the rate of inactivation. When we added 0.2 mol fraction SDS (X_{SDS}), however, the addition of mSA had a dramatic effect on the rate of inactivation (Figure 2B). In the presence of mSA, DGK loses ∼60% activity in the first week and then undergoes a slower inactivation phase. The inactivation curve at 0.2 X_{SDS} in the absence of mSA could be fit

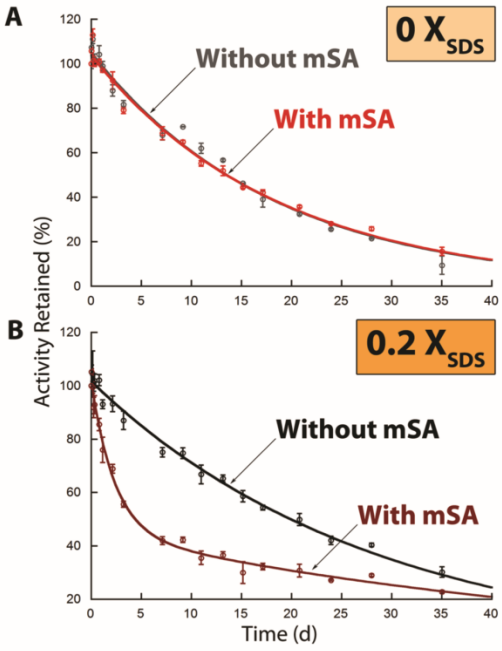


Figure 2. DGK inactivation by steric trapping. (A) The results in pure OG (zero SDS). BM-DGK$_{Cysless/I53C}$ was incubated with (red) or without (gray) mSA. The curves are fits to a single exponential decay. (B) The results in 0.2 X_{SDS}. BM-DGK$_{Cysless/I53C}$ was incubated with (black) or without (dark red) mSA. The curve without mSA is a fit to a single exponential decay. The curve with mSA is a fit to a double exponential decay. Error bars indicate standard deviation of activities measured in triplicate. The specific activity of the enzyme, corresponding to 100% activity, was 71 ± 3 μmol·min^{-1}·mg^{-1}, which is within the reported ranges for wild-type DGK.[20,24]

to a single exponential with a rate constant of k_I = 0.036 ± 0.001 d^{-1}. The inactivation curve in the presence of mSA required a double exponential fit with rate constants of 0.43 ± 0.05 d^{-1} (fast phase) and 0.019 ± 0.004 (slow phase). The similar rate constants for the slow phase and k_I suggest that the slow phase reflects the same inactivation process that occurs in the absence of mSA and that the fast phase reflects inactivation due to steric trapping (k_{ST}).

Investigating the Mechanism of Inactivation by Steric Trapping. The slow rate of inactivation, even in 0.2 X_{SDS}, made an extensive characterization of the mechanism impractical as every experiment requires multiple weeks to perform. Nevertheless, we were able to learn some basic characteristics of the inactivation mechanism. We believe the experiments described below are consistent with the model shown in Figure 1B. In this model, steric trapping causes dissociation primarily to dimers and monomers. The dimers and monomers can reassociate to restore activity, but the rate-limiting reassociation/refolding event is a slow, unimolecular, conformational change. We now present the evidence for this model.

Prior work on the inactivation mechanism at elevated temperatures suggests that the slow, irreversible inactivation observed in the absence of mSA reflects a stable, essentially irreversible conformational change that occurs after subunit dissociation.[28] This is consistent with a slowing of the intrinsic inactivation rate as the DGK concentration increases: in the absence of mSA, the intrinsic inactivation rate in 0.2 X_{SDS} slows from k_I = 0.023 ± 0.004 d^{-1} at 0.2 μM BM-DGK$_{Cysless/I53C}$ to k_I = 0.014 ± 0.003 d^{-1} at 1.8 μM BM-DGK$_{Cysless/I53C}$ (Figure 3, black curves). In this model, the subunits dissociate and then can either reassociate or undergo an irreversible inactivation. In

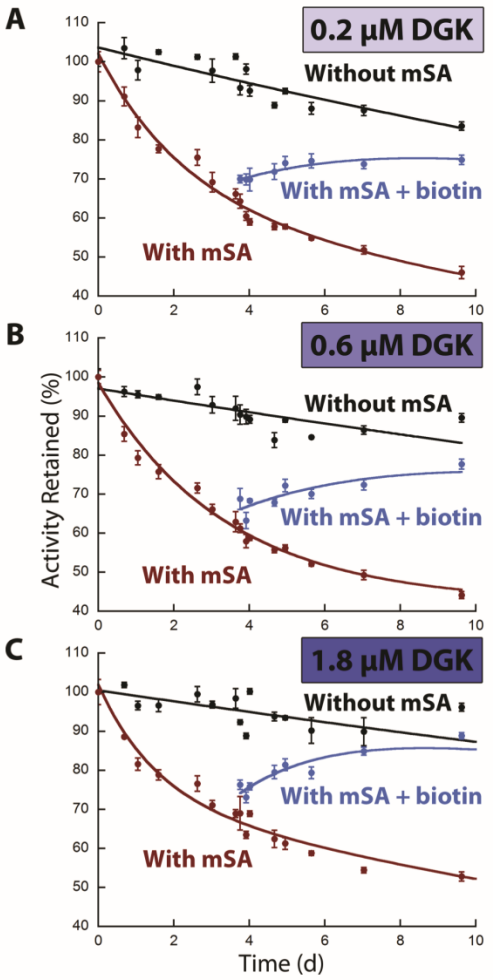


Figure 3. Concentration dependence of steric trap dissociation and reassociation. BM-DGK$_{Cysless/I53C}$ was incubated in 0.2 X_{SDS} without mSA$_{S45A}$ (black) and with mSA$_{S45A}$ (red). Free biotin was added to an aliquot of sterically trapped BM-DGK$_{Cysless/I53C}$ at 89 h (blue). Data are shown for three different concentrations of BM-DGK$_{Cysless/I53C}$: (A) 0.2 μM (B) 0.6 μM, and (C) 1.8 μM. Error bars indicate standard deviation of activities measured in triplicate. The specific activity of the enzyme, corresponding to 100% activity, was 63 ± 4 μmol·min^{-1}·mg^{-1}, which is within the reported ranges for wild-type DGK.[20,24]

the presence of mSA, however, the subunit dissociation becomes irreversible by steric trapping.

As expected for a unimolecular irreversible dissociation in the presence of mSA, the rate of inactivation by steric trapping is not concentration dependent. Figure 3 shows steric trap inactivation at 0.2 μM, 0.6 μM, and 1.8 μM BM-DGK$_{Cysless/I53C}$. Double exponential fits yield steric trap inactivation rates (k_{ST}) of 0.4 ± 0.2 d^{-1}, 0.3 ± 0.1 d^{-1}, and 0.7 ± 0.2 d^{-1} for 0.2, 0.6, and 1.8 μM BM-DGK$_{Cysless/I53C}$, respectively. Thus, in spite of a nearly 10-fold change in concentration, there is very little change in steric trap inactivation rates. To test whether mSA binding might be rate limiting, we mixed mSA and BM-DGK$_{Cysless/I53C}$, incubated for 30 min, and then prevented further binding by the addition of excess biotin. Analysis by size-exclusion chromatography indicates that complete binding to the intact trimer occurs within the 30-min incubation (Figure S1 in Supporting Information) and is therefore too fast to contribute significantly to k_{ST}. Overall, these results are consistent with a rate-limiting dissociation of a subunit from the trimer or a slow conformational change within the trimer that occurs prior to mSA trapping.

If the sterically trapped state is dissociated, it is possible that the reverse reaction (reassociation) would be concentration dependent, unless a conformational change is rate limiting. We therefore examined the concentration dependence of reactivation after the addition of a biotin competitor. To ensure that the rate of mSA release is fast, we employed a mutant of mSA, mSA$_{S45A}$, with a reduced binding affinity. For this mutant, biotin dissociation occurs with a half-life of less than 2 min at room temperature.[29] As shown in Figure S2 in Supporting Information, mSA$_{S45A}$ has sufficient binding affinity at the concentrations used to inactivate BM-DGK$_{Cysless/I53C}$ to the same extent as mSA. Figure 3 shows the progress of activity recovery after biotin addition (blue traces). Activity recovers slowly over the course of many days, asymptotically approaching the activity decay curve for the enzyme that has not been subject to steric trapping (black traces). We fit these curves to a double exponential reflecting both the recovery of activity k_R and the slow inactivation of the untrapped enzyme (k_I, see above). In these fits, k_I was fixed at the observed inactivation rate in the absence of mSA. At 0.2, 0.6, and 1.8 μM BM-DGK$_{Cysless/I53C}$, we found that k_R was 0.20 ± 0.02, 0.21 ± 0.04, and 0.38 ± 0.07 d^{-1}, respectively. Thus, activity recovery from the sterically trapped inactive state is only moderately concentration dependent at best. Thus, from the recovery rates alone it is still possible that sterically trapped DGK is not dissociated or there is a rate-limiting conformational change at some point during either refolding of the monomers or the trimer that is much slower than the reassociation rate.

To test further whether the subunits are dissociated in the sterically trapped state, we asked whether subunit mixing could occur after steric trap inactivation. The experiment is outlined in the box in Figure 1B. We reasoned that if we added dissociated/unfolded subunits bearing a mutation in one-half of the active site, they would be able to freely associate with the sterically trapped monomers and dimers. Reassociation and refolding could then restore complete activity to the sterically trapped subunits, but the mutant subunits would be unable to form active trimers on their own, as illustrated in Figure 1B. For these experiments we first generated a steric trap inactivated BM-DGK$_{Cysless/I53C}$ by incubating with mSA$_{S45A}$ for 4 days until the activity dropped to about 60% of its starting value (Figure 3A). We then mixed with an inactive mutant of DGK, E76L,[19] which had been previously unfolded at high SDS concentration. If the steric trap inactivated BM-DGK$_{Cysless/I53C}$ was dissociated, we expected that the mutant subunits would refold, mix with, and reactivate the BM-DGK$_{Cysless/I53C}$ subunits. As shown in Figure 4, we saw essentially complete recovery of activity within

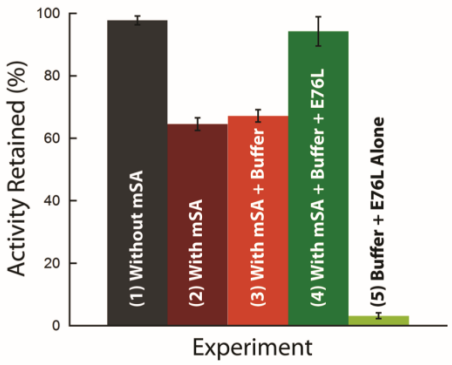


Figure 4. Reactivation of sterically trapped DGK after mixing with the inactive mutant, DGK$_{E76L}$. The experiment is depicted in the shaded box of Figure 1. The relative DGK activities after four different experiments are shown: (1) Incubation of BM-DGK$_{Cysless/I53C}$ for 98 h by itself. The activity of this sample is treated as the maximum possible activity that could be recovered. (2) Incubation of BM-DGK$_{Cysless/I53C}$ for 98 h in the presence of mSA$_{S45A}$ to generate a sterically trapped state. The reduction in activity to ~60% of the value in the absence of mSA indicates the formation of steric trap inactivated enzyme. (3) Addition of the SDS unfolding buffer used to unfold DGK$_{E76L}$ by itself to the sterically trapped sample. (4) Addition of a 5-fold molar excess of a previously SDS-unfolded, inactive mutant DGK$_{E76L}$ to the sterically trapped sample, followed by further incubation for 3 h. (5) Addition of previously SDS-unfolded DGK$_{E76L}$ to the buffer used in steric trapping. Steric trapping was performed in 0.2 X$_{SDS}$. Error bars indicate the standard deviation of activities measured in triplicate.

3 h of mixing with DGK$_{E76L}$. The recovery of activity is considerably faster than reactivation of trapped BM-DGK$_{Cysless/I53C}$ without the SDS-unfolded mutant subunits. Fast refolding from an SDS-unfolded state is consistent with earlier work.[30,31] It therefore seems likely that the SDS and steric trap unfolded states are different. Nevertheless, the SDS unfolded subunits are capable of refolding with the sterically trapped form and restore activity, indicating that the steric trap-inactivated protein is dissociated.

While the results are consistent with the basic model outlined in Figure 1B, in a system this complex, we cannot rule out all possible mechanisms or minor parallel pathways for inactivation and reactivation. Nevertheless, we believe that we can make the following key observations: (1) Steric trapping leads to a dissociated or partially dissociated inactive state as expected. (2) The rate of appearance of the dissociated/unfolded state is extremely slow in OG, with a half-life of at least 12.6 d (the rate of inactivation with or without mSA).

Measuring DGK Unfolding by Subunit Exchange. As an independent measure of subunit dissociation, we examined the kinetics of subunit exchange. We created a subunit that was distinguishable from the wild-type by inserting a pentaglutamate tag at the N-terminus of DGK to create a more negatively charged protein, E$_5$-DGK. To prevent the possibility of disulfide bond formation between subunits that could prevent

exchange, the two native cysteines (C46 and C113) were mutated to alanine to give the constructs used in these experiments: $DGK_{Cysless}$ and E_5-$DGK_{Cysless}$.

As shown in Figure 5A, $DGK_{Cysless}$ and E_5-$DGK_{Cysless}$ can be clearly separated on an ion exchange column. When the two distinct subunits are expressed together in the same cell, we observe mixed subunit compositions, although the distribution is biased toward $DGK_{Cysless}$, which expresses at higher levels than E_5-$DGK_{Cysless}$. To monitor subunit exchange rates, we incubated an equimolar mixture of $DGK_{Cysless}$ and E_5-$DGK_{Cysless}$ at room temperature in OG and monitored subunit exchange by separation on the anion-exchange column. Very little exchange occurred even after 7 days of incubation (Figure 5B). Moreover, as shown in Figure 5C, ~86% of the activity is maintained after 2 days and ~72% after 7 days. Thus, the lack of subunit exchange cannot be explained by rapid irreversible dissociation of subunits. The final ratio of the mixed peaks to the pure $DGK_{Cysless}$ or E_5-$DGK_{Cysless}$ peaks is expected to be 3:1 on the basis of binomial sorting, yet the ratio is only ~0.3 after 7 days, indicating that subunit exchange has only progressed to about 10% of its expected final value. These results imply that the half-life of subunit exchange in OG is at least ~5 weeks.

To compare our experiments in detergent micelles to more physiologically relevant conditions, we measured DGK subunit exchange in vesicles made from *E. coli* lipids. Proteoliposomes containing $DGK_{Cysless}$, E_5-$DGK_{Cysless}$, or a mixture of both in the same vesicles were prepared and incubated at 37 °C (Figure 6). After incubation, the vesicles were solubilized in OG (where only slow exchange occurs) and the subunit distributions analyzed by anion-exchange chromatography as before. It was necessary to use smaller quantities of enzyme for the vesicle experiments, so we detected DGK by activity measurements rather than absorbance at 280 nm. The results of these incubations are shown in Figure 6. The activities and subunit compositions remain largely unchanged over the 2-d incubation period at 37 °C. Longer incubations were not attempted because loss of liposome integrity can be significant after many days.[32] Nevertheless, it is clear that subunit exchange is very slow under physiological conditions, as we found for the detergent-solubilized enzyme.

■ DISCUSSION

Our results indicate that subunit dissociation of DGK is extremely slow. Under the destabilizing conditions of 0.2 X_{SDS}, subunit dissociation occurs with a half-life of about 1.6 d as measured by the steric trap inactivation rate (k_{ST}). In the absence of steric trapping, the protein inactivates much more slowly, and the rate of inactivation is slowed by increased concentration. Previous work on the inactivation mechanism at elevated temperatures also found a protection effect at higher concentration.[28] As suggested previously, we believe this must reflect a partially reversible dissociation step that occurs prior to inactivation of the dissociated subunit. Steric trapping renders this partially reversible dissociation completely irreversible. In pure OG without SDS, steric trapping does not enhance the rate of inactivation. This suggests that dissociation is largely irreversible under these conditions, rendering the inactivation rate with and without mSA the same. It should be noted that OG solubilization can also be considered partially destabilizing conditions because the enzyme is more stable in other, longer-chain detergents.[20] Nevertheless, the inactivation/dissociation rate in OG is very slow, on the order of 2 weeks (12.6 d).

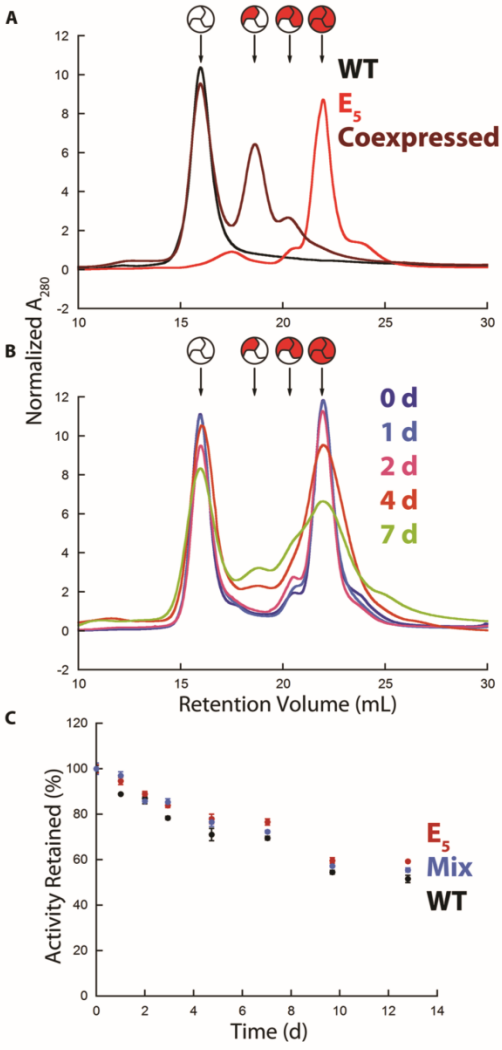


Figure 5. Subunit exchange of DGK in 1.5% OG. (A) Anion-exchange chromatograms of $DGK_{Cysless}$ (black), E_5-$DGK_{Cysless}$ (red), and E_5-$DGK_{Cysless}$ with $DGK_{Cysless}$ coexpressed in the same cells (maroon). All samples were maintained at 4 °C during and after purification. (B) Anion-exchange chromatograms of $DGK_{Cysless}$ and E_5-$DGK_{Cysless}$ that were mixed after purification. The mixed samples were incubated for 0 d (blue), 1 d (light blue), 2 d (pink), 4 d (orange), and 7 d (green) at room temperature. Absorbance at 280 nm was normalized to the average value between elution volumes of 10 and 30 mL. (C) Activity (normalized to the initial value) over the time period of the incubation at room temperature. Error bars indicate standard deviation of activities measured in triplicate.

As an independent measure of subunit dissociation rates, we also examined the rate of subunit exchange. Only ~10% subunit exchange was observed over the course of seven days in OG, and no subunit exchange was observed after 2 days under

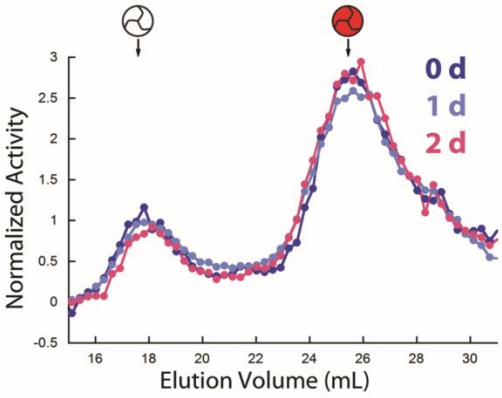


Figure 6. Subunit exchange of DGK trimers in *E. coli* polar lipid vesicles. Anion-exchange chromatograms are shown of a mix of $DGK_{Cysless}$ and E_5-$DGK_{Cysless}$ incubated at 37 °C for 0 d (blue), 1 d (light blue), and 2 d (pink). Prior to chromatography, the vesicles were solubilized in OG, and the chromatography was performed in OG buffer. Fractions were assayed for activity and normalized to the average value.

physiological conditions of *E. coli* lipid vesicles at 37 °C. Thus, on the time scale of an *E. coli* cell with a doubling time of ~30 min, DGK essentially never dissociates.

Although the proximity of residue 53 in the trimer can permit cross-linking of subunits in the I53C mutant in the presence of oxidizing agents,[34] the slow dissociation observed here cannot be due to disulfide bonding. Steric trap inactivation requires multiple labels, precluding disulfide formation in proteins that are subject to the effects of steric trapping. Moreover, we saw no evidence for dimer formation in mass spectra of either the unlabeled or BM-labeled proteins, consistent with the high percentage of labeling observed (~71%). Disulfide formation was not possible in the subunit mixing experiments since they were performed on a cysless variant of DGK.

How might this high kinetic stability be achieved? A likely mechanism is high thermodynamic stability. As yet, there are no measurements for the thermodynamic stability of helical membrane proteins under native conditions that do not involve an uncertain extrapolation from high denaturant concentrations. Nevertheless, theoretical arguments from Deeds et al. suggest that cyclic oligomers should be particularly stable.[35] In the simple case of a cyclic trimer, formation of a dimer must be sufficiently stable to overcome the entropy cost of reducing two freely diffusing protein subunits to a single complex. The completion of the trimer involves the same entropic cost term, but the stabilizing contribution from the interface is doubled because the final subunit forms two interfaces instead of just one, providing a high level of stability. Cyclic oligomers are a natural symmetry for membrane proteins that are constrained to a planar bilayer. Indeed, AcrB is a cyclic trimer and appears to also have a slow subunit dissociation rate.[6] As kinetic stability is likely to be a favorable property for biochemical and structural studies of membrane proteins, it might therefore make sense to look for membrane protein homologues that are higher order oligomers.

In the cell, proteins are often degraded by proteases that recognize unfolded segments. Proteins that do not often sample the unfolded state are resistant to proteolysis,[36] and thus would require some active mechanism of degradation. In *E. coli*, the membrane-bound AAA protease FtsH is known to degrade an unstable mutant of DGK, but not wild-type DGK.[37] It may also be possible for membrane proteins to be cleaved at sites of local unfolding, which may occur on a faster time scale than the unfolding observed in our experiments. Whether active mechanisms exist or are even needed in *E. coli* to degrade kinetically stable proteins such as DGK is unknown.

Evolutionary pressure for high kinetic stability makes sense for an enzyme like α-lytic protease that must operate outside the cell under unforgiving conditions,[4] but the evolutionary advantages of membrane protein kinetic stability is less clear. Diacylglycerol kinase plays an important role in generating membrane-derived oligosaccharides that protect cells against osmotic stress conditions.[38] Kinetic stability may therefore be important for a protein that must remain active under harsh conditions. In the milieu of the membrane where there is a high concentration of protein, it may also be advantageous to avoid unfolding that would provide opportunities for inappropriate docking with other transmembrane helices. Helical bundles can have slow rates of folding and unfolding due to frustrated energy landscapes,[39,40] which may also be a feature of transmembrane helices, given the importance of van der Waals packing[41] that could lead to locally stable incorrect conformations. For example Doura and Fleming showed that the transmembrane helix of glycophorin A can dimerize in multiple ways.[42] Differing structures of DGK solved by solution NMR[18] and X-ray crystallography[17] also suggest that it is possible for the helices to pack in stable alternate conformations. Kinetic stability may be a way to protect against the formation of stable, but inactive structures.

How many kinetically stable membrane proteins exist is not clear at this time. It is entirely possible that most of the membrane proteins of known structure are kinetically stable since proteins need to remain folded for the time required to grow crystals or collect NMR spectra. Now that we know that high kinetic stability can be obtained for membrane proteins, a better understanding of the mechanisms of kinetic stability could be of practical importance. It may allow for the engineering of kinetic stability and expand the number of proteins accessible to detailed scrutiny.

■ ASSOCIATED CONTENT

Supporting Information

Supporting Figures 1 and 2. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

bowie@mbi.ucla.edu.

Present Address

‡Amgen, One Amgen Center Drive, 28-5-B, Thousand Oaks, CA 91320-1799.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by NIH Grant 5R01GM063919 to JUB and by the Ruth L. Kirschstein National Research Service

Award GM007185. We thank Thomas Magliery (OSU) for insightful discussion, and Bowie lab members for critically reading the manuscript.

REFERENCES

(1) Manning, M.; Colón, W. *Biochemistry* **2004**, *43*, 11248−11254.
(2) Xia, K.; Manning, M.; Hesham, H.; Lin, Q.; Bystroff, C.; Colón, W. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 17329−17334.
(3) Puorger, C.; Eidam, O.; Capitani, G.; Erilov, D.; Grütter, M. G.; Glockshuber, R. *Structure* **2008**, *16*, 631−642.
(4) Jaswal, S. S.; Sohl, J. L.; Davis, J. H.; Agard, D. A. *Nature* **2002**, *415*, 343−346.
(5) Huysmans, G. H. M.; Baldwin, S. A.; Brockwell, D. J.; Radford, S. E. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 4099−4104.
(6) Lu, W.; Chai, Q.; Zhong, M.; Yu, L.; Fang, J.; Wang, T.; Li, H.; Zhu, H.; Wei, Y. *J. Mol. Biol.* **2012**, *423*, 123−134.
(7) Rotem, D.; Sal-man, N.; Schuldiner, S. *J. Biol. Chem.* **2001**, *276*, 48243−48249.
(8) Fisher, L. E.; Engelman, D. M.; Sturgis, J. N. *J. Mol. Biol.* **1999**, *293*, 639−651.
(9) Fujii, J.; Maruyama, K.; Tada, M.; MacLennan, D. H. *J. Biol. Chem.* **1989**, *264*, 12950−12955.
(10) Yang, B.; Gonzalez, L.; Prekeris, R.; Steegmaier, M.; Advani, R. J.; Scheller, R. H. *J. Biol. Chem.* **1999**, *274*, 5649−5653.
(11) Hardie, K. R.; Lory, S.; Pugsley, A. P. *EMBO J.* **1996**, *15*, 978−988.
(12) Heginbotham, L.; Odessey, E.; Miller, C. *Biochemistry* **1997**, *36*, 10335−10342.
(13) Sargiacomo, M.; Scherer, P. E.; Tang, Z.; Kübler, E.; Song, K. S.; Sanders, M. C.; Lisanti, M. P. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9407−9411.
(14) Borgnia, M. J.; Kozono, D.; Calamita, G.; Maloney, P. C.; Agre, P. *J. Mol. Biol.* **1999**, *291*, 1169−1179.
(15) Spelbrink, R. E. J.; Kolkman, A.; Slijper, M.; Killian, J. A.; de Kruijff, B. *J. Biol. Chem.* **2005**, *280*, 28742−28748.
(16) Curnow, P.; Booth, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18970−18975.
(17) Li, D.; Lyons, J. A.; Pye, V. E.; Vogeley, L.; Aragão, D.; Kenyon, C. P.; Shah, S. T. A.; Doherty, C.; Aherne, M.; Caffrey, M. *Nature* **2013**, *497*, 521−524.
(18) Van Horn, W. D.; Kim, H.-J.; Ellis, C. D.; Hadziselimovic, A.; Sulistijo, E. S.; Karra, M. D.; Tian, C.; Sonnichsen, F. D.; Sanders, C. R. *Science* **2009**, *324*, 1726−1729.
(19) Lau, F. W.; Chen, X.; Bowie, J. U. *Biochemistry* **1999**, *38*, 5521−5527.
(20) Zhou, Y.; Bowie, J. U. *J. Biol. Chem.* **2000**, *275*, 6975−6979.
(21) Massey-Gendel, E.; Zhao, A.; Boulting, G.; Kim, H.-Y.; Balamotis, M. A.; Seligman, L. M.; Nakamoto, R. K.; Bowie, J. U. *Protein Sci.* **2009**, *18*, 372−383.
(22) Lau, F. W.; Bowie, J. U. *Biochemistry* **1997**, *36*, 5884−5892.
(23) Howarth, M.; Chinnapen, D. J.-F.; Gerrow, K.; Dorrestein, P. C.; Grandy, M. R.; Kelleher, N. L.; El-Husseini, A.; Ting, A. Y. *Nat. Methods* **2006**, *3*, 267−273.
(24) Badola, P.; Sanders, C. R. *J. Biol. Chem.* **1997**, *272*, 24176−24182.
(25) Blois, T. M.; Hong, H.; Kim, T. H.; Bowie, J. U. *J. Am. Chem. Soc.* **2009**, *131*, 13914−13915.
(26) Hong, H.; Blois, T. M.; Cao, Z.; Bowie, J. U. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 19802−19807.
(27) Hong, H.; Bowie, J. U. *J. Am. Chem. Soc.* **2011**, *133*, 11389−11398.
(28) Zhou, Y.; Lau, F. W.; Nauli, S.; Yang, D.; Bowie, J. U. *Protein Sci.* **2001**, *10*, 378−383.
(29) Hyre, D. E.; Stayton, P. S.; Trong, I. L.; Freitag, S.; Stenkamp, R. E. *Protein Sci.* **2000**, *9*, 878−885.
(30) Nagy, J. K.; Lonzer, W. L.; Sanders, C. R. *Biochemistry* **2001**, *40*, 8971−8980.
(31) Lorch, M.; Booth, P. J. *J. Mol. Biol.* **2004**, *344*, 1109−1121.
(32) Elferink, M. G. L.; de Wit, J. G.; Driessen, A. J. M.; Konings, W. N. *Biochim. Biophys. Acta* **1994**, *1193*, 247−254.
(33) Lau, F. W.; Nauli, S.; Zhou, Y.; Bowie, J. U. *J. Mol. Biol.* **1999**, *290*, 559−564.
(34) Nagy, J. K.; Lau, F. W.; Bowie, J. U.; Sanders, C. R. *Biochemistry* **2000**, *39*, 4154−4164.
(35) Deeds, E. J.; Bachman, J. A.; Fontana, W. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 2348−2353.
(36) Park, C.; Zhou, S.; Gilmore, J.; Marqusee, S. *J. Mol. Biol.* **2007**, *368*, 1426−1437.
(37) Herman, C.; Prakash, S.; Lu, C. Z.; Matouschek, A.; Gross, C. A. *Mol. Cells* **2003**, *11*, 659−669.
(38) Van Horn, W. D.; Sanders, C. R. *Annu. Rev. Biophys.* **2012**, *41*, 81−101.
(39) Wensley, B. G.; Batey, S.; Bone, F. A. C.; Chan, Z. M.; Tumelty, N. R.; Steward, A.; Kwa, L. G.; Borgia, A.; Clarke, J. *Nature* **2010**, *463*, 685−688.
(40) Wensley, B. G.; Kwa, L. G.; Shammas, S. L.; Rogers, J. M.; Browning, S.; Yang, Z.; Clarke, J. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 17795−17799.
(41) Oberai, A.; Joh, N. H.; Pettit, F. K.; Bowie, J. U. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 17747−17750.
(42) Doura, A. K.; Fleming, K. G. *J. Mol. Biol.* **2004**, *343*, 1487−1497.

Supporting Information for

Membrane Proteins Can Have High Kinetic Stability

Robert E. Jefferson[†], *Tracy M. Blois*[†‡], *James U. Bowie**[†]

[†]Department of Chemistry and Biochemistry, University of California, Los Angeles-Department of Energy Institute for Genomics and Proteomics, Molecular Biology Institute, University of California, Los Angeles, CA 90095, United States

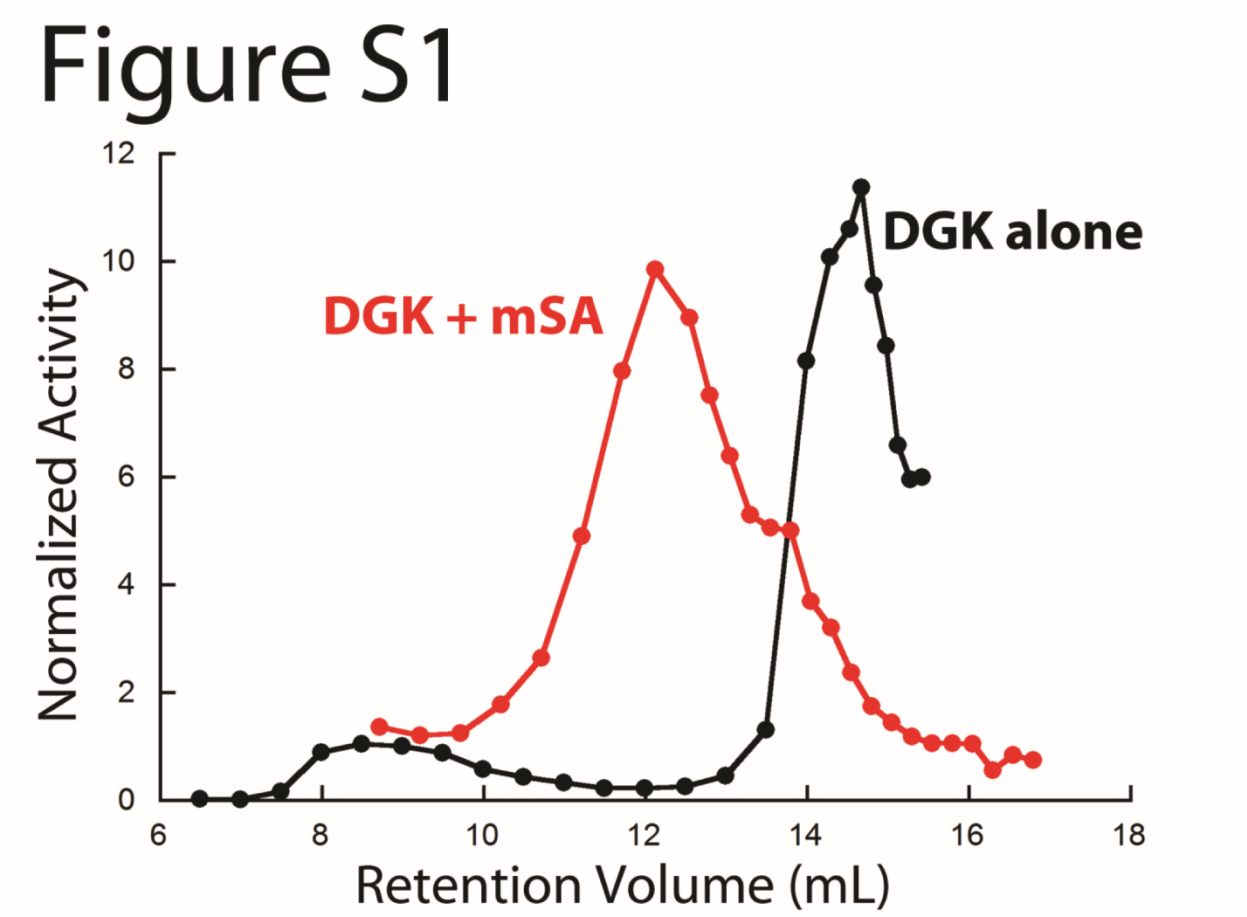


Figure S1. Rapid initial binding of mSA to BM-DGK$_{Cysless/I53C}$. Size exclusion chromatograms are shown of BM-DGK$_{Cysless/I53C}$ before (red) and after (black) incubation with mSA. Prior to chromatography, BM-DGK$_{Cysless/I53C}$ was incubated for 30 min with mSA and then further binding to BM-DGK$_{Cysless/I53C}$ was blocked by the addition of excess biotin.

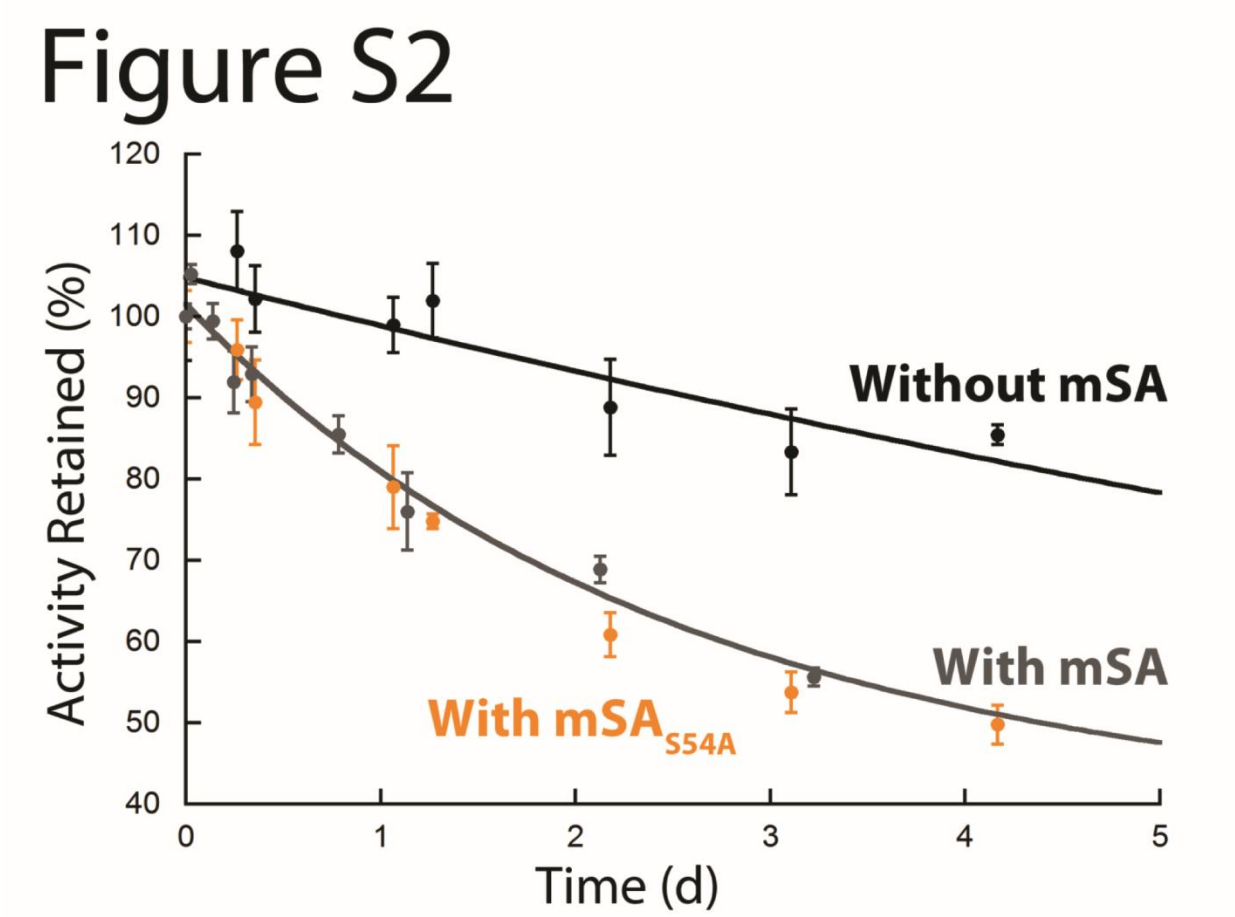


Figure S2. The weaker biotin binding mutant mSA_{S45A} is capable of sterically trapping DGK. 0.2 µM BM-$DGK_{Cysless/I53C}$ in 0.2 X_{SDS} without mSA (black) and with 1.8 µM mSA_{S45A} (orange) or wild-type mSA (gray).

CHAPTER 3

Single-Molecule Measurement of Steric Trapping in Lipid Bilayers

Moving assays into the single-molecule regime can open up new realms of study as well as overcome some of the technical challenges of bulk experiments. Early single-molecule fluorescence experiments measured the unfolding of chymotrypsin inhibitor 2 in denaturing solvent (Deniz et al., 2000). The approach employs bright fluorophores attached to the protein at residues close in the folded structure, but far apart in the primary sequence. The dyes exhibit high Förster resonance energy transfer (FRET) from the donor dye to the acceptor dye under folded conditions, but upon the titration of guanidine the dyes undergo a drastic loss of FRET. Fluorescence from single-molecules is observed from freely diffusing protein at such a low concentration that only a single protein is present in the confocal excitation volume at a time, producing "bursts" of photons detected separately for donor and acceptor emission. While this method can interrogate the unfolding and folding of a protein at a new level of detail, it is also does not require a functional assay, is free of aggregation, and uses only a small amount of protein. These advantages are especially enticing for membrane proteins that often do not have functional assays, are prone to aggregate, and often suffer from low expression levels.

Instead of applying single-molecule FRET-based denaturant titration studies to membrane proteins, we sought to merge a method for studying membrane protein folding under native conditions with single-molecule measurements. The steric trap method utilizes the high affinity of bulky monovalent streptavidin for a protein dual labeled with biotins that are only accessible in the unfolded state (Blois et al., 2009) to drive unfolding within the context of the lipid bilayer. While the steric trap method can measure membrane protein folding thermodynamics and kinetics under native conditions, it has only been demonstrated on membrane proteins with functional assays performed in bulk conditions (Chang and Bowie, 2014; Jefferson et al., 2013). Recent steric trapping studies of the *E. coli* rhomboid protease GlpG have presented a pyrene-based fluorescence quenching assay that is generally applicable to membrane proteins in the bulk condition (Guo et al., 2016).

To translate this method into the single-molecule regime, we proposed monitoring protein unfolding by measuring the fraction of doubly bound species. This new method necessitates that the binding of monovalent streptavidin (mSA) be measured instead of assaying for unfolding through a characteristic change in the membrane protein itself. Our scheme directly labels mSA with two colors of bright single molecule dyes. By using a 1:1 mix of donor and acceptor labeled mSA, half of the doubly bound population will display coincident bursts under saturating conditions. Because the confocal volume is alternately excited by each laser on a μs timescale, photons detected in the acceptor emission channel during donor excitation are attributed to the donor dye. At single-molecule concentrations, concerns regarding scattering from lipid vesicles are irrelevant. The main challenge arises from the need to use a high concentration of streptavidin to drive membrane protein unfolding, which creates a heterogeneous mixture of molecules (Figure 3-1). Provided a sufficient amount of unbound streptavidin can be removed before measurement, then the proteoliposomes containing bound species can be separated from free fluorescent streptavidin by burst duration because the system is diffusion based. The average time a large vesicle resides in the excitation volume will be considerably larger than for monovalent streptavidin. Similar to previous steric trap experiments, the titration of streptavidin will produce an unfolding curve that will be proportional to the increasing fraction of coincident vesicle bursts.

To demonstrate the steric trap method under single-molecule conditions, we used bacteriorhodopsin, a well-studied membrane protein that has been steric trapped in bicelles (Chang and Bowie, 2014). Efforts to steric trap bacteriorhodopsin in lipid vesicles have been challenged by the need to measure absorbance of the retinal chromophore, which requires a high amount of protein and lipid. While bR is not hard to purify in large quantities, the high concentration of vesicles presents a scattering problem for absorbance measurements. To remedy this, the vesicles are solubilized in detergent just before reading. Thus, while the

membrane protein is incubated in a native bilayer, the measurement necessitates the destruction of that environment. By moving this method to a single-molecule measurement we hoped to avoid the challenges of bulk vesicle measurements, as well as generalize the technique for a broader array of targets.

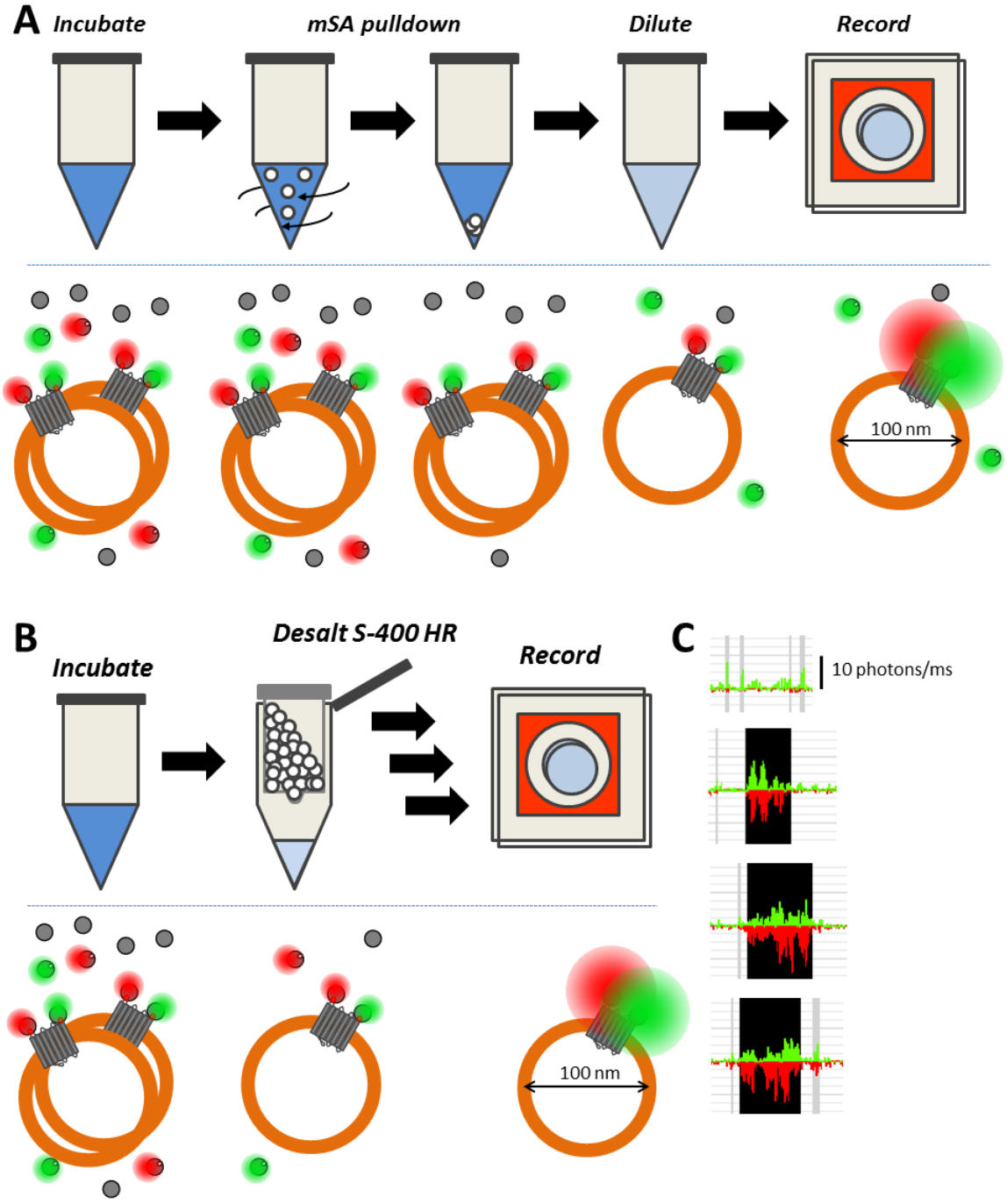


Figure 3-1 Schemes for preparing steric trapped membrane protein in vesicles for single-molecule measurements. **(A)** Scheme for preparing steric trapped vesicles for single-molecule measurement using biotin-agarose beads. Representation of the reaction mixture is shown below. Inactive streptavidin is depicted by gray circles. Fluorophore-labeled monovalent streptavidin is differentiated by acceptor (red) and donor (green) dyes. **(B)** Scheme for preparing steric trapped vesicles for single-molecule measurement using desalting spin columns of Sephacryl S-400 HR resin. Representation of the reaction mixture is shown below. Streptavidin representation is same as in panel A. **(C)** Example single-molecule bursts. Scale bar represents a burst intensity of 10 photons/ms. Photons are binned in 1 ms intervals. Short bursts that do not pass the duration constraint for bound proteoliposomes are marked with a gray background. Long bursts passing the duration constraint are marked with a black background.

Preparation of Single-Molecule Vesicles and Dye-Conjugated Streptavidin for Steric Trapping

Bacteriorhodopsin (bR) was purified as described (Chang and Bowie, 2014). The same double-cysteine background (D36C/F230C) was used for biotin attachment. Wild-type bR is too stable in 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) vesicles, so the protein was destabilized with a point mutation (M20A) to enable steric trapping on a reasonable timescale. 2 µl of purified $bR_{D36C/F230C/M20A}$ in purple membrane was solubilized in 34 µl of 20 mM sodium phosphate (pH 6.0), 1% 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC): 3-[(3-Cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate (CHAPSO) 1:1 (w/w) and treated with 2 µl of 40 mM tris(2-carboxyethyl)phosphine (TCEP) (ThermoFisher Scientific) for final concentrations of ~10-20 µM bR_{M20A} and 2 mM TCEP. After reducing the cysteines for 2 h at room temperature, 2 µl of 20 mM *N*-(biotinoyl)-*N'*-(iodoacetyl)ethylenediamine (Biotium) to a final concentration of 1 mM. The labeling reaction was incubated overnight with shaking at room temperature. Unreacted label was removed by 0.5 ml 40K MWCO Zeba Spin Desalting Columns (ThermoFisher Scientific). Desalting columns were washed three times with 300 µl of 20 mM sodium phosphate (pH 6.0), 1% DMPC:CHAPSO 1:1 (w/w) before desalting the labeling reaction. The eluted biotinylated bR_{M20A} (bR_{M20A}-BE_2) was used promptly for incorporation into vesicles.

To yield vesicles with only a single bR_{M20A}-BE_2, the incorporation of protein was statistically skewed by mixing 1 bR_{M20A}-BE_2 per ten 0.1 µm POPC (Avanti Polar Lipids) vesicles. Most of the resulting vesicles will be empty of protein, but 97% of the non-empty vesicles will only contain a single bacteriorhodopsin protein, according to a Poisson distribution where the probability of incorporation into a vesicle is the number of bR_{M20A}-BE_2 molecules divided by the

volume of the membrane phase, analogous to the probability of capturing a single molecule within a vesicle (Okumus et al., 2004).

In a typical incorporation, 34 nM bR_{M20A}-BE_2 was added to 500 µl of 33 mM POPC dissolved in 20 mM sodium phosphate (pH 6.0), 4% n-octyl-β-D-glucopyranoside (OG). The protein and lipid solution was dialyzed against 3 exchanges of 500 ml of 20 mM sodium phosphate (pH 6.0) in a 50 kD MWCO dialysis bag at 4 °C. The formed multilamellar vesicles were extruded through a 0.1 µm Nucleopore track-etched polycarbonate membrane (Whatman) 15 times in a mini-extruder (Avanti Polar Lipids) to create a monodisperse distribution of unilamellar vesicles.

A cysteine variant of streptavidin (A35C) was created by site-directed mutagenesis for the attachment of fluorescent dyes. Active mSA_{A35C} variants were purified as previously described (Hong et al., 2009), except that 1 mM TCEP was added to all buffers after refolding the streptavidin with inactive subunits to prevent cysteine oxidation. Purified aliquots were snap frozen in liquid nitrogen and stored at -80 °C. Inactive streptavidin (dSA) was made by concentrating a refolded solution of inactive subunits and passing it over a HiTrap Q anion exchange column (GE Healthcare Life Sciences). The flow-through was collected, concentrated again, and stored at 4 °C.

Prior to labeling, ~20 µM streptavidin was treated with 1 mM fresh TCEP for 2 h at room temperature. Iodoacetamide-derivatized dyes (donor: ATTO 532 (ATTO-TEC), acceptor: LD650 (Lumidyne Technologies)) were added to a final concentration of 1 mM in a 40 µl reaction and incubated with shaking overnight at room temperature. Unreacted dye was removed by separation on a Superdex 200 10/300 GL size-exclusion chromatography column (GE Healthcare Life Sciences). Streptavidin labeled with donor and acceptor dyes were mixed to give

an equal concentration of dyes, and thus a maximum probability of doubly bound unfolded bR producing coincident bursts.

To create steric trapped bR_{M20A}-BE_2, single-molecule proteoliposomes were diluted to 1-30 nM and incubated with varying concentrations of dye-labeled streptavidin in 20 mM sodium phosphate (pH 6.0) for 1-2 days to equilibrate. 0.1 mg/ml dSA was added to block nonspecific binding of dye-conjugated streptavidin to membranes and reaction vessels. In addition to dSA, we attempted to prevent loss of bound proteoliposomes to adsorption, either to plastic and glass surfaces or to the membranes themselves. The bright fluorophores used for single-molecule experiments are particularly hydrophobic (Hughes et al., 2014) and at the measured concentrations (low pM) the fraction lost to adsorption is significant. To protect against adsorption, pipette tips, plastic reaction tubes, and the glass chamber used for single-molecule measurements were coated with siliconizing fluid (AquaSil, ThermoFisher Scientific) to create a non-interactive surface. Using siliconized equipment helped preserve fluorescent material and yielded more observed bursts at identical dilutions compared to non-siliconized materials (Figure 3-2).

Single-molecule recordings were made on a previously described setup (Kapanidis et al., 2004). Samples diluted to measureable concentrations were sealed in a chamber made by two glass coverslips around an adhesive silicone gasket. The sample did not contact the gasket material. Excitation intensities were 100 μW at 532 nm for the donor fluorophore and 150 μW at 638 nm for the acceptor fluorophore. The alternation period was 50 μs. Donor emission was detected through a 580 df 60 nm filter. Acceptor emission was detected through a 697 df 45 nm filter. Background rates were determined by fitting the histogram of interval times between subsequent photons to a negative exponential curve. Bursts were defined by a sliding window of 10 photons that achieved a threshold rate of 6 times above the background rate. Data analysis was performed using homebuilt LabVIEW software (National Instruments).

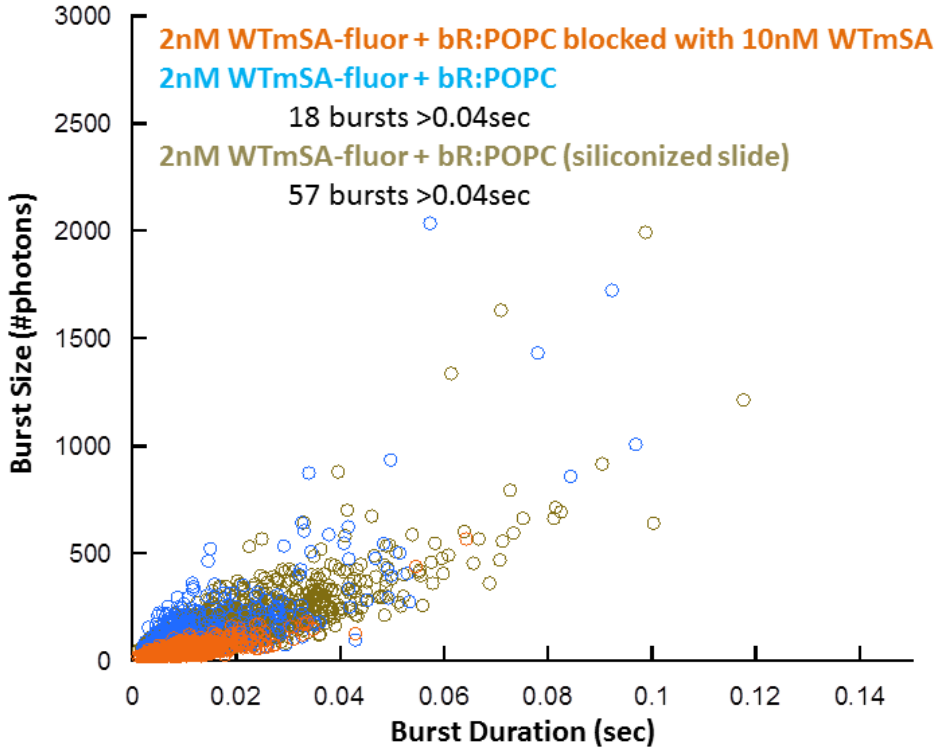


Figure 3-2 Comparison of vesicle bursts with and without siliconization. Orange circles are bursts from fluorophore-labeled mSA_{A35C} in the presence of bR_{M20A}-BE_2 POPC vesicles previously saturated with unlabeled WT mSA. Blue circles are bursts from fluorescent WT mSA_{A35C} bound to bR_{M20A}-BE_2 POPC vesicles. Brown circles are bursts from bR_{M20A}-BE_2 POPC vesicles incubated with the same concentration of fluorescent WT mSA_{A35C}, but incubated, diluted, and recorded with siliconized materials.

Efficient Removal of Free Streptavidin

While single-molecule diffusion experiments can inherently separate vesicle bursts from free protein, it is still crucial to remove most of the free protein so the population of bound species is not drowned out by unbound bursts. The excess mSA must be removed quickly so that doubly bound complexes do not dissociate into singly bound or unbound species due to the off-rate of streptavidin variants. Off-rates were measured using a fluorescein biotin conjugate, biotin-4-fluorescein (ThermoFisher Scientific) that is quenched upon binding. Recovery of the quenched biotin-4-fluorescein upon dissociation was monitored after competing with excess

free biotin. While the wild-type monovalent streptavidin has an extremely slow off-rate with a half-life of ~38 d, mutants with modulated affinity that can be used for binding curves have half-lives from 4.5 h to 11.6 h (Figure 3-3). Though some dissociation will occur over the course of a 20-30 min measurement, the off-rate remains linear within this time frame, which will not alter the final shape of the binding curves.

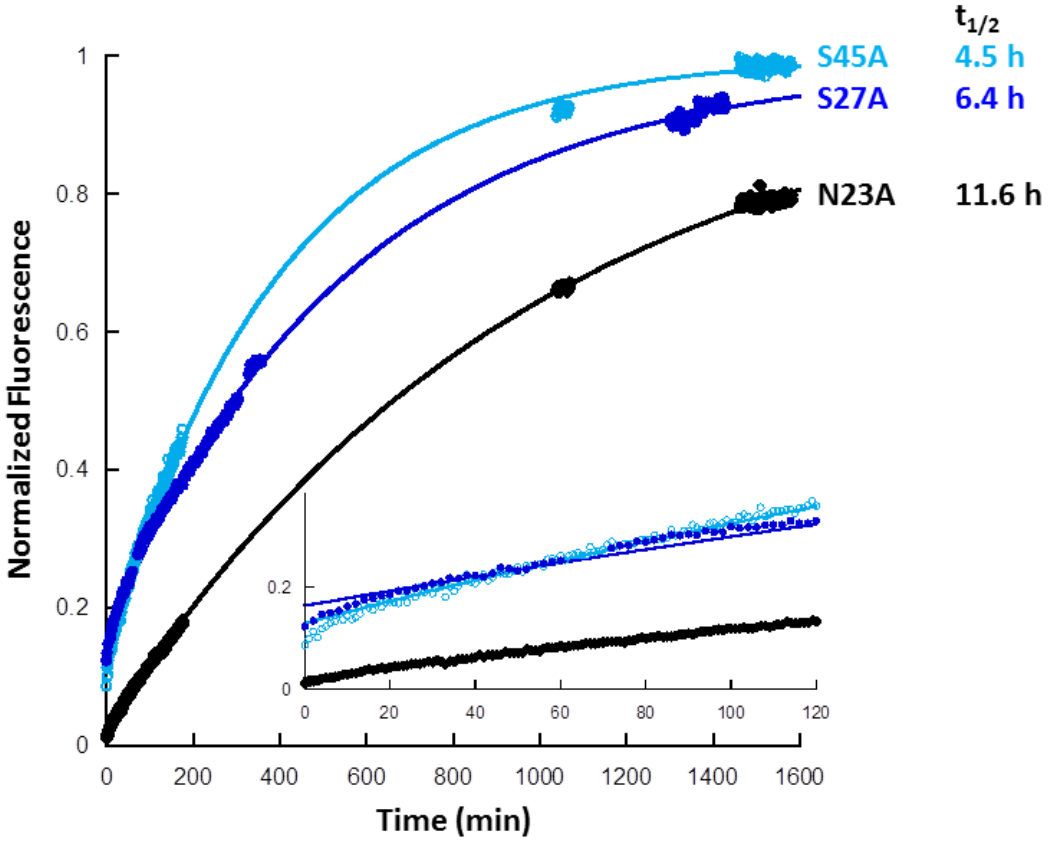


Figure 3-3 Dissociation kinetics of mSA_{A35C} variants and biotin-4-fluorescein. mSA variants were bound to an equimolar amount of biotin-4-fluorescein for 10 min. Fluorescence recovery was monitored after addition of 100-fold excess free biotin and normalized to free biotin-4-fluorescein. **Inset:** Expanded zoom of first two hours of fluorescence recovery after addition of excess free biotin.

Initially, attempts to remove free streptavidin used biotin-agarose beads (Sigma Aldrich) to pull down unbound streptavidin immediately prior to single-molecule measurements. The

beads have a high binding capacity of 30 mg/ml avidin in a 1:1 slurry, so only a small amount of beads needs to be added to the vesicle incubations. 1 part bead slurry was added to 4 parts of the steric trapped vesicles. The beads were suspended during the pull down with a vortex on medium-low speed for 15 min. The reaction volume was 40 uL and the tubes were fastened upright to the vortex and protected from ambient light to prevent photobleaching.

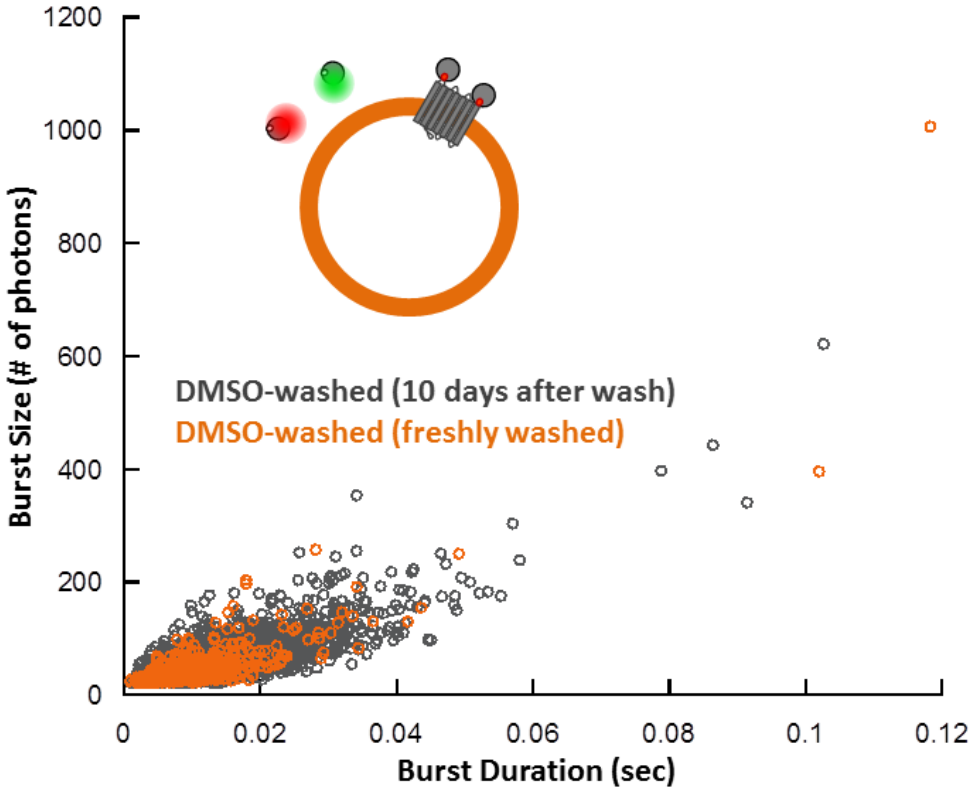


Figure 3-4 Fluorescent bursts after removing free streptavidin with DMSO-washed biotin-agarose beads. Fluorophore-labeled WT mSA_{A35C} were added to bR_{M20A}-BE_2 POPC vesicles saturated with unlabeled WT mSA. Gray circles represent bursts after removing mSA with biotin-agarose washed with DMSO 10 days prior. Orange circles represent bursts after removing mSA with biotin-agarose freshly washed with DMSO.

After optimizing the pulldown conditions, bursts from free dye-labeled streptavidin still remained. Hypothesizing that a small amount of biotin could be dissociating from the agarose and prevents complete removal of free streptavidin, the beads were washed with 20 CVs of

dimethyl sulfoxide (DMSO) using a gravity column. The beads were left in DMSO overnight and then washed again with 20 CVs of DMSO. The DMSO helped solubilize free biotin and wash it away while only biotin attached to the agarose beads remained. The streptavidin removal was much improved and left only a small number of short bursts that could be removed by a duration constraint of ~30 ms (Figure 3-4).

While improved, mSA removal via biotin agarose bead pulldowns was still imperfect. Using a duration constraint to select for vesicle bursts only eliminates many bursts because, although vesicles will diffuse slower due to their relative size, most molecules will skirt the edges of the confocal volume and thus have typically short burst duration times. With better removal protocols, we can include more bursts without separating by duration. Instead of using the affinity of free streptavidin for biotin, we employed desalting spin columns that would remove free streptavidin based on size in a timely manner. Custom spin columns were made with Pierce spin columns and 600 µl of 1:1 Sephacryl S-400 HR resin slurry (GE Healthcare Life Sciences). Spin columns were washed with 300 µl of 20 mM sodium phosphate (pH 6.0) three times before use. To isolate bound single-molecule vesicles, 25 µl of steric trapped vesicles were added to a washed spin column and eluted by centrifugation at 3000 rpm for 2 min. The eluent was discarded and another 25 µl of 20 mM sodium phosphate (pH 6.0) was added to the same spin column to elute the peak fraction of vesicles. To clean up the vesicles for single-molecule measurement, this second elution fraction was passed through two more spin columns of washed Sephacryl S-400 HR resin. Removing free streptavidin by desalting spin columns provided a superior protocol for isolating bound vesicles, leaving only small bursts in the unbound control, which could be removed by a small size constraint of 20 photons without a duration constraint (Figure 3-5).

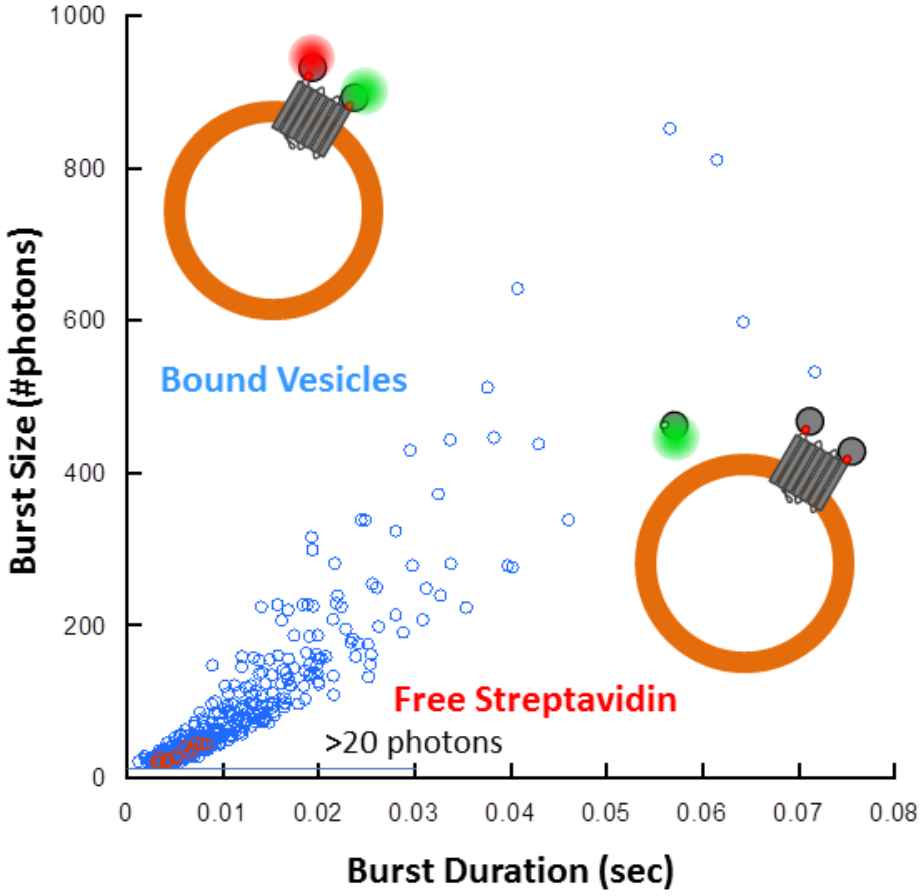


Figure 3-5 Removal of free streptavidin by desalting spin columns. Blue circles represent bursts from bR POPC vesicles bound with fluorophore-labeled WT mSA_{A35C}. Red circles represent bursts recorded after mSA removal with 3 spin columns of Sephacryl S-400 HR resin from the same concentration of fluorescent WT mSA in the presence of bR_{M20A}-BE_2 POPC vesicles previously saturated with unlabeled WT mSA. A size constraint of 20 photons was used to eliminate most of the bursts left after removal.

Defining Coincident Bursts

Calculating the fraction bound from single-molecule measurements of desalted steric trapped $bR_{M20A/D36C/F230C}$ in POPC vesicles was based on simple stoichiometry of the burst. The number of photons from the donor fluorophore is divided by the total number of photons from both fluorophores gives a stoichiometry term for the burst, S. If S=1, the burst is from a donor-only molecule. If S=0, the burst is from an acceptor-only molecule. A stoichiometry of 0.5

indicates a burst from a molecule with both a donor and acceptor fluorophore. Burst stoichiometries are rarely so clean-cut due to photophysical properties of the fluorophores, their attachment to the protein, their trajectory through the excitation volume, and the detectors. To account for the variability of burst stoichiometry we created single-bR vesicles bound with only donor-only or acceptor-only streptavidin. The stoichiometries of those control bursts served as negative controls for coincident bursts, while coincident bursts were defined as those with stoichiometries between the donor-only and acceptor-only extremes (Figure 3-6).

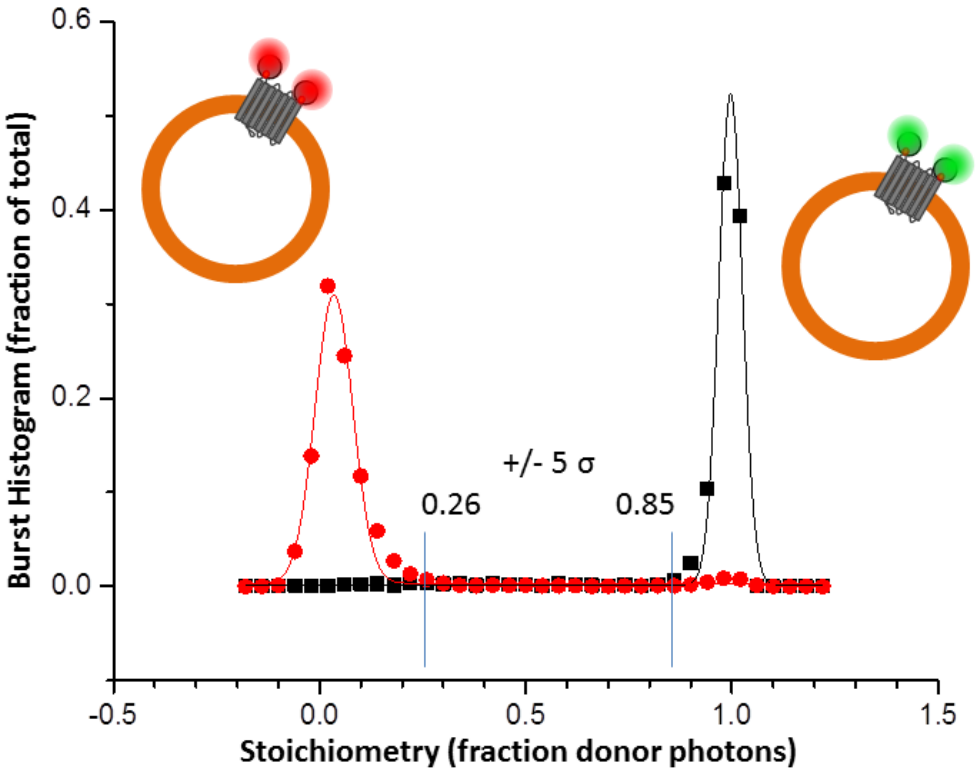


Figure 3-6 Negative controls for stoichiometry of coincident bursts. Histograms of burst stoichiometry from bR_{M20A}-BE_2 POPC vesicles saturated with acceptor fluorophore-labeled mSA_{A35C} (red) or donor fluorophore-labeled mSA_{A35C} (black). Coincident stoichiometries were bounded by the Gaussian fits to acceptor burst stoichiometry +5 σ and donor burst stoichiometry -5 σ.

Employing a stoichiometry-only constraint to define coincident bursts, a binding curve was constructed with wild-type streptavidin. The high affinity of WT mSA (4.8×10^{-14}) (Howarth

et al., 2006) is expected to create a stoichiometric binding curve. The fraction of coincident bursts increases with fluorescent mSA concentration until about 15 nM where the fraction of coincident bursts plateaus. However, due the imprecise measurement of the actual doubly-biotinylated incorporated bR after vesicle formation and extrusion, the protein concentration was overestimated and the resultant curve lacks data points below the maximum observed coincidence. The fraction coincident was averaged from three separate steric trap incubations and showed good agreement, but the maximum fraction of coincident bursts is much lower than expected at 5-7% of vesicle bursts being coincident. While coincidence is not expected to be the ideal 50% due to incomplete fluorophore labeling of mSA_{A35C} and biotinylated bR, the observed plateau was significantly lower than the expected 20-25%.

A second set of steric trap incubations was measured at lower streptavidin concentrations to fill in the unsaturated slope of the binding curve. No stoichiometric binding curve appeared and instead of a plateau of maximum coincidence, the observed coincidence resembled an intrinsic baseline of coincident bursts independent of streptavidin concentration. The burst stoichiometry histograms show a non-zero but essentially constant level of bursts in the range of coincident stoichiometry (Figure 3-7). We hypothesized that the coincident bursts were not actual doubly bound complexes, but rather singly bound folded bR that happened to enter the excitation volume at the same time as another singly bound bR of the opposite color. If this was true, then the bR incorporated into POPC vesicles was not unfolding, perhaps due to the lack of aggregation at single-molecule concentrations that could help destabilize the membrane protein at bulk concentrations. We then sought to create coincident bursts by other means that would help define truly coincident complexes from random coincidence of two different fluorescent vesicles.

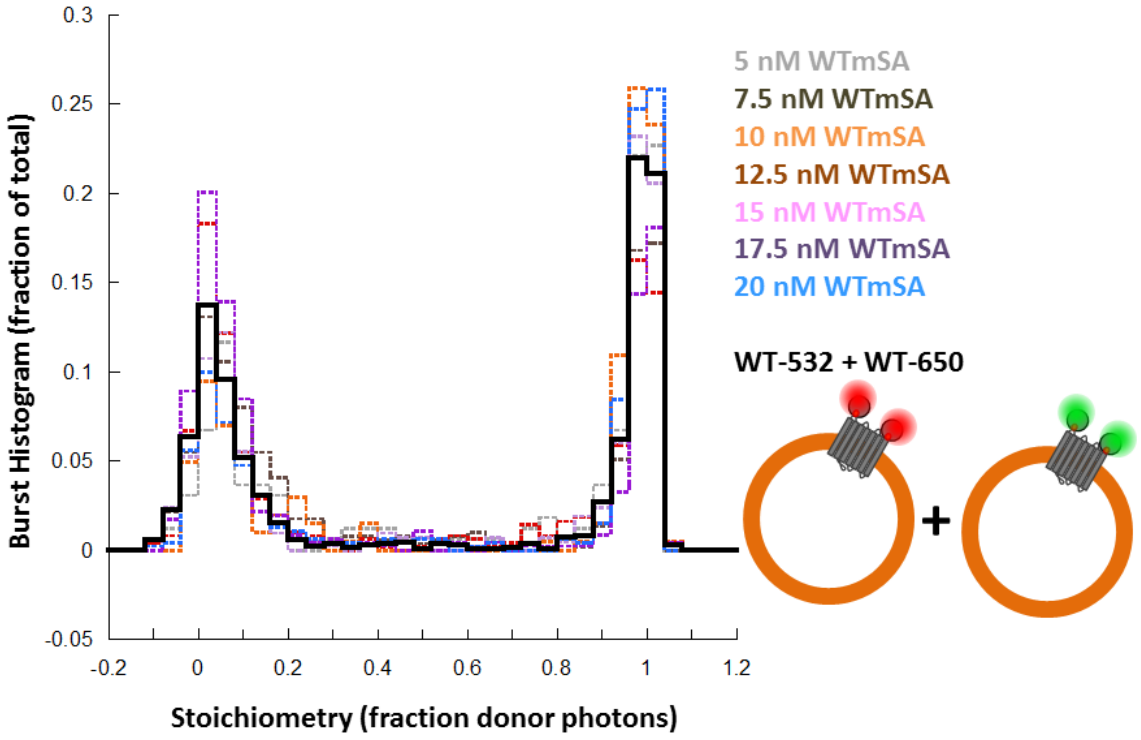


Figure 3-7 Comparison of burst stoichiometry for titrations of fluorophore-labeled WT mSA_{A35C}. bR_{M20A}-BE_2 POPC vesicles were incubated with increasing concentrations of fluorescent WT mSA_{A35C}. As a control for random coincidence of acceptor vesicles with donor vesicles, a 1:1 mixture of bR_{M20A}-BE_2 POPC vesicles saturated with acceptor-labeled WT mSA_{A35C} and donor-labeled WT mSA_{A35C}.

We attempted to create positive controls from unfolded bR, dimerized streptavidin, double-stranded DNA, and direct labeling of the bR cysteine mutant. To ensure an unfolded bR, we titrated in a molar ratio of fluorescent streptavidin to SDS-unfolded bR. Without removal of free streptavidin, the coincidence reaches a maximum when there is one mSA per biotin binding site, and higher ratios of streptavidin drown out the coincidence to give a lower fraction of coincident bursts. (Figure 3-8) While the fractional coincidence increases around the ratio of mSA:bR for binding all biotin sites, the coincidence does not display a very prominent peak beyond the baseline level.

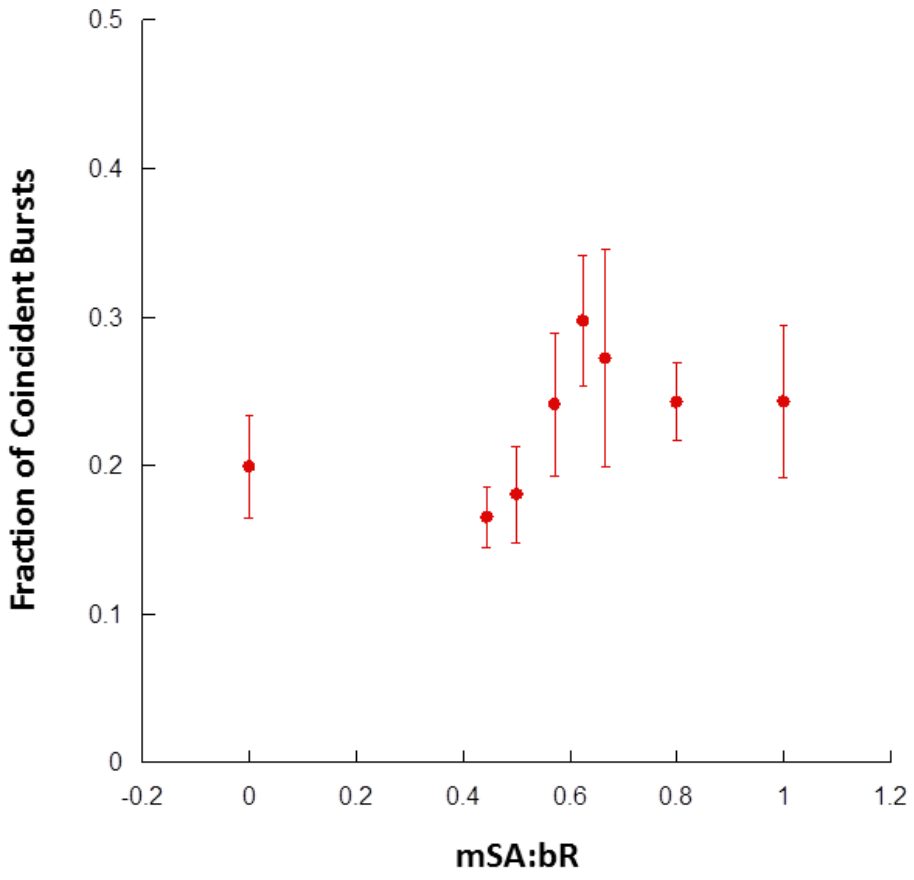


Figure 3-8 Fraction of coincident bursts from unfolded bR_{M20A}-BE_2. bR_{M20A}-BE_2 solubilized in DMPC:CHAPSO micelles was titrated into a constant concentration of fluorophore-labeled WT mSA_{A35C} in 0.63 X_{SDS}. mSA was allowed to bind for 1 h at RT before recording bursts. Error bars are standard error from triplicate measurements. Due to incomplete biotinylation, all biotin sites are expected to be bound at a mSA:bR ratio of 0.59.

To avoid using a mix of streptavidin and biotinylated membrane protein that can yield unsaturated bR_{M20A}-BE_2 or excess free streptavidin, we attempted to create self-dimers of mSA using a bifunctional linked biotin. The formation of dimers was imperfect, even at a 1:2 molar ratio of linker to monovalent streptavidin, but the dimeric species was purified by size exclusion chromatography on a Superdex 200 10/300 GL column. However, even with purified dimers of dye-conjugated monovalent streptavidin, coincident bursts were still only observed at background levels.

Failing coincident bursts from dye-conjugated monovalent streptavidin dimers, we conjugated dyes to complimentary DNA oligos. Separate labeling of the complimentary strands creates double-stranded DNA with one of each donor and acceptor dye. To attach our custom single-molecule fluorophores, we used thiol modified oligos, which limited our labeling efficiency. Even with modest labeling, there was still no clear coincident population of bursts.

Finally, we tried directly attaching both dyes directly to the double cysteine mutant of bR. Labeling with an equimolar ratio of both dyes should yield a mixture of labeled bR similar to bR steric trapped with fluorescent streptavidin. This method suffers from incomplete labeling and differential labeling efficiencies of each cysteine residue, so, perhaps unsurprisingly, did not produce measureable coincident bursts.

Alternate Detection of Steric Trapped Protein

Due to the lack of coincident bursts in steric trapped proteoliposomes and a robust positive control, we sought other means of detecting the unfolded fraction. It is possible to sterically trap under single-molecule conditions and then pool the protein and detect unfolding in a bulk assay, albeit still with high sensitivity. Recently, the Hong lab has advanced the steric trap method with reporter biotin labels conjugated to the nitroxide EPR probe, fluorescein, or pyrene. Though fluorescein is not bright or photostable enough for single-molecule measurements, in aggregate fluorescein is a useful probe. We took advantage of the large size and detergent resistance of streptavidin to separate doubly bound unfolded bR from singly bound folded bR by SDS-PAGE. The gels can then be fluorescence scanned for only bands containing biotin-fluorescein labeled bR, ignoring excess streptavidin and unlabeled bR bands. While significantly less technically complex than observing single molecules in solution, this gel-shift method provides the advantage of measuring multiple incubations in parallel on the same

gel. This method would also offer a quick way to measure membrane protein stability without the need for expensive specialized equipment.

Labeling of $bR_{M20A/D36C/F230C}$ with biotin-fluorescein-iodoacetamide (Gift from H. Hong Lab, Michigan State University) and steric trap incubations were similar to those for single-molecule measurements, with minor changes. The concentration of fluorescein-labeled bR_{M20A} (bR_{M20A}-BF_2) was ultimately too low to be observed with a fluorescence scanner when incorporated at a ratio of 1 bR_{M20A}-BF_2 molecule per 10 vesicles, so ratio of 1:5 was used instead. This ratio still ensured that 94% of the proteoliposomes contained only a single bR_{M20A}-BF_2 molecule. This method requires destruction of the vesicles in SDS for gel electrophoresis, which denatures bR and allows all biotin sites to become accessible. To prevent complete binding of every biotin site during SDS-PAGE, an excess of biotin was added in the gel loading buffer to block all free streptavidin. Gels were run at 140 V on ice to keep the steric trapped complexes intact.

Initial tests with WT mSA contained contaminating fluorescent bands from free streptavidin. The streptavidin used was not fluorescent, so the bands appeared from a small amount of free biotin-fluorescein label that remained after desalting and dialysis during vesicle formation. A size-exclusion chromatography step was added after the desalting column to remove any remaining unreacted label. The label was tightly associated with bR_{M20A}-BF_2 and eluted only marginally later than the protein in 0.2% n-dodecyl-β-D-maltoside on a Superdex 200 10/300 GL column. Lower concentrations of detergent were ineffective at separating the free label. Unreacted dye mixed with bR_{M20A} in detergent separates cleanly, so the limited separation after the labeling reaction is likely due to the label embedding in the protein micelle after the long reaction times. Once purified, the labeled bR_{M20A}-BF_2 was then incorporated into liposomes by dialysis as before. The resulting incubations were free of any fluorescent streptavidin bands.

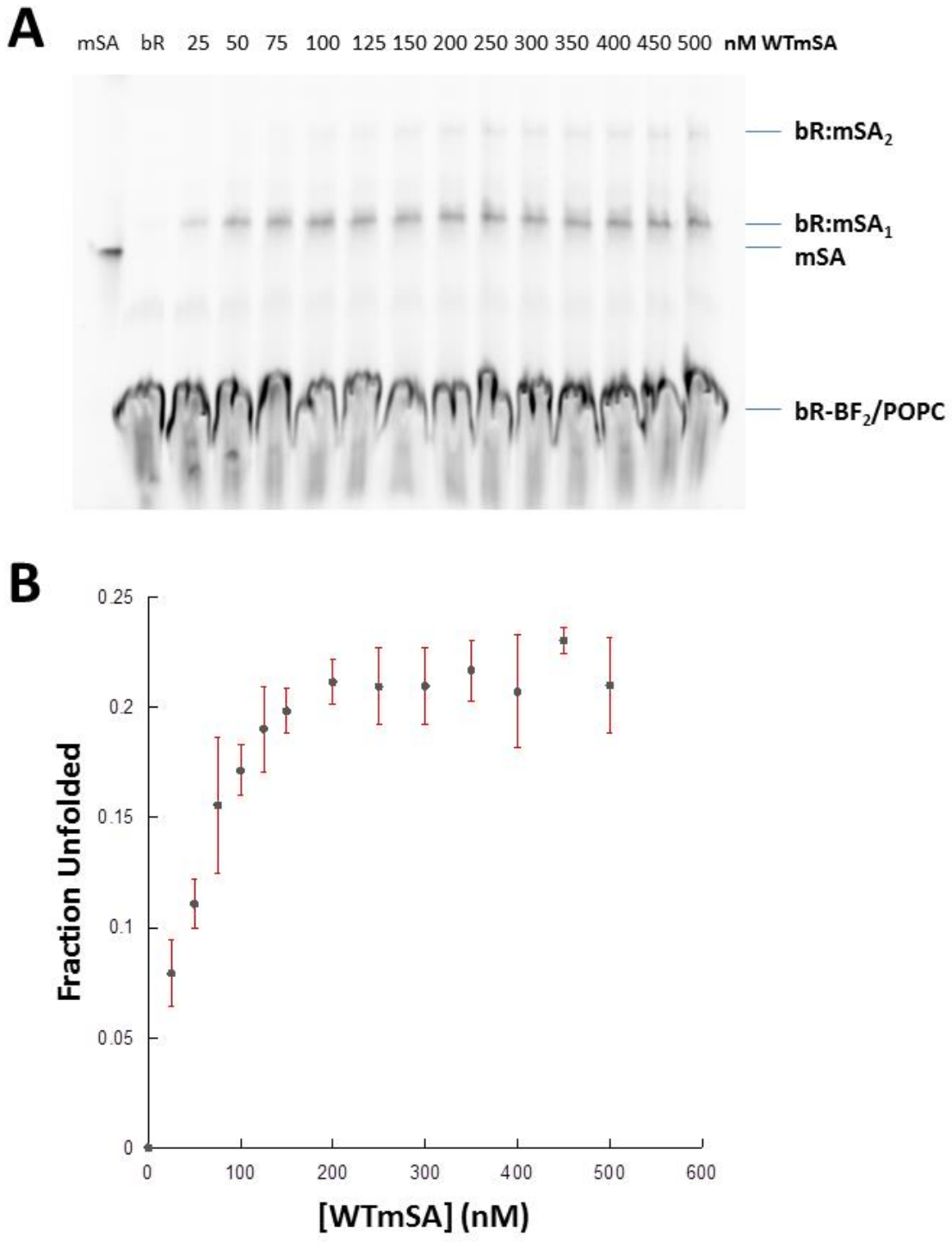


Figure 3-9 Measuring unfolding bR_{M20A}-BF_2 with WT mSA by gel-shift. (A) Fluorescence scan of SDS-PAGE after incubation of bR_{M20A}-BF_2 POPC vesicles with WT mSA. mSA lane is WT mSA bound to biotin-4-fluorescein. bR lane is bR_{M20A}-BF_2 alone. (B) Quantification of the fraction of unfolded bR_{M20A}-BF_2 calculated from band intensities in Image Lab software (Bio-Rad Laboratories). Error bars are standard error from three gels.

WT mSA titrations produced a clear stoichiometric binding curve. (Figure 3-9) There were bands from singly bound and unbound bR_{M20A}-BF_2 even at saturating concentrations of streptavidin. At first, these caveats were attributed to inconsistent labeling of the cysteines that yields bR with a single label that can only be singly bound, and inward facing bR whose biotin labels are inaccessible to streptavidin and ultimately unbound.

However, attempts to create a binding curve with weaker affinity streptavidin mutants were unable to capture the same levels of unfolded bR_{M20A}-BF_2. The S27A mutant of mSA has a half-life of dissociation from free biotin-4-fluorescein of 6.44 h, so most steric trapped complexes should stay intact during the gel electrophoresis. Even with high concentrations of mutant streptavidin as used in bulk experiments, doubly bound unfolded bR_{M20A}-BF_2 was undetectable and singly bound bR_{M20A}-BF_2 bands were significantly reduced in intensity from WT mSA. (Figure 3-10 A) To confirm the streptavidin off-rate in a more relevant context, the dissociation rate from bound bR_{M20A}-BE_2 to free biotin-4-fluorescein was measured in SDS and non-denaturing detergent. Surprisingly, the off-rate of S27A monovalent streptavidin from biotin conjugated to bR was found to be 30.9 times faster in 0.2% n-dodecyl-β-D-maltopyranoside and 74.3 times faster in 2% SDS (Figure 3-10 B). To prevent dissociation, we tried to observe gel shifting with native PAGE in non-denaturing detergent without excess biotin. Though the streptavidin dissociation may still be fast from the membrane protein micelle complex, the complexes could reform without the need for excess free biotin, provided the gel electrophoresis separation did not physically prevent re-association. The slow migration rate of the membrane protein complexes by native PAGE was mitigated with blue-native PAGE, but neither clear-native nor blue-native PAGE displayed doubly bound unfolded bR bands.

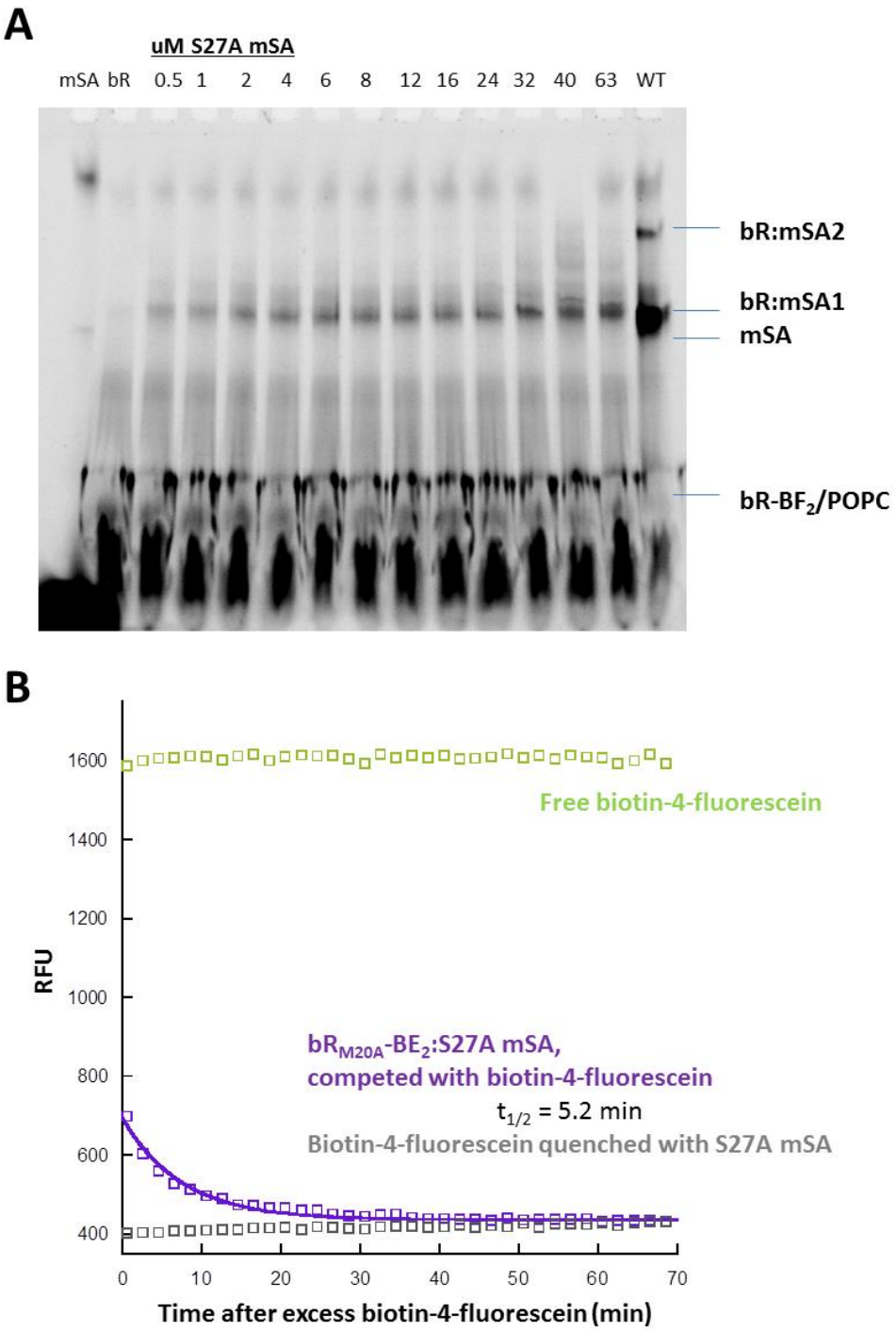


Figure 3-10 Dissociation of mutant mSA from steric trapped complexes. **(A)** Fluorescence scan of SDS-PAGE after incubation of bR_{M20A}-BF_2 POPC vesicles with S27A mSA. mSA lane is WT mSA bound to biotin-4-fluorescein. bR lane is bR_{M20A}-BF_2 alone. WT lane is bR_{M20A}-BF_2 POPC vesicles incubated with 500 nM WT mSA. **(B)** Dissociation of S27A mSA from bR_{M20A}-BE_2. bR_{M20A}-BE_2 was bound with a 2-fold excess of S27A mSA in 2% SDS to ensure all biotin sites were bound. mSA was competed with an excess of biotin-4-fluorescein and the fluorescence quenching was monitored.

The accelerated off-rate of monovalent streptavidin from biotin-conjugated membrane proteins is a significant technical challenge for single-molecule steric trapping in which the steric trapped protein must be isolated from the bulk of excess monovalent streptavidin, whether for single-molecule diffusion measurements or gel electrophoresis. Bulk experiments are measured at equilibrium, and as such the off-rate problem is not a concern. The pyrene quenching steric trap assay developed by the Hong lab provides a sensitive and general method for interrogating membrane protein folding (Guo et al., 2016). In the single-molecule realm, force spectroscopy offers a robust alternative for reversibly unfolding single membrane proteins (Min et al., 2015, 2016; Serdiuk et al., 2016; Thoma et al., 2015).

References

Blois, T.M., Hong, H., Kim, T.H., and Bowie, J.U. (2009). Protein Unfolding with a Steric Trap. J. Am. Chem. Soc. *131*, 13914–13915.

Chang, Y.-C., and Bowie, J.U. (2014). Measuring membrane protein stability under native conditions. Proc. Natl. Acad. Sci. U. S. A. *111*, 219–224.

Deniz, A.A., Laurence, T.A., Beligere, G.S., Dahan, M., Martin, A.B., Chemla, D.S., Dawson, P.E., Schultz, P.G., and Weiss, S. (2000). Single-molecule protein folding: Diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. Proc. Natl. Acad. Sci. *97*, 5179–5184.

Guo, R., Gaffney, K., Yang, Z., Kim, M., Sungsuwan, S., Huang, X., Hubbell, W.L., and Hong, H. (2016). Steric trapping reveals a cooperativity network in the intramembrane protease GlpG. Nat. Chem. Biol. *12*, 353–360.

Hong, H., Joh, N.H., Bowie, J.U., and Tamm, L.K. (2009). Chapter 8 Methods for Measuring the Thermodynamic Stability of Membrane Proteins. In Methods in Enzymology, J.M.H. and G.K.A. Michael L. Johnson, ed. (Academic Press), pp. 213–236.

Howarth, M., Chinnapen, D.J.-F., Gerrow, K., Dorrestein, P.C., Grandy, M.R., Kelleher, N.L., El-Husseini, A., and Ting, A.Y. (2006). A monovalent streptavidin with a single femtomolar biotin binding site. Nat. Methods *3*, 267–273.

Hughes, L.D., Rawle, R.J., and Boxer, S.G. (2014). Choose Your Label Wisely: Water-Soluble Fluorophores Often Interact with Lipid Bilayers. PLoS ONE *9*, e87649.

Jefferson, R.E., Blois, T.M., and Bowie, J.U. (2013). Membrane Proteins Can Have High Kinetic Stability. J. Am. Chem. Soc. *135*, 15183–15190.

Kapanidis, A.N., Lee, N.K., Laurence, T.A., Doose, S., Margeat, E., and Weiss, S. (2004). Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. Proc. Natl. Acad. Sci. U. S. A. *101*, 8936–8941.

Min, D., Jefferson, R.E., Bowie, J.U., and Yoon, T.-Y. (2015). Mapping the energy landscape for second-stage folding of a single membrane protein. Nat. Chem. Biol. *11*, 981–987.

Min, D., Arbing, M.A., Jefferson, R.E., and Bowie, J.U. (2016). A simple DNA handle attachment method for single molecule mechanical manipulation experiments. Protein Sci. *25*, 1535–1544.

Okumus, B., Wilson, T.J., Lilley, D.M.J., and Ha, T. (2004). Vesicle Encapsulation Studies Reveal that Single Molecule Ribozyme Heterogeneities Are Intrinsic. Biophys. J. *87*, 2798–2806.

Serdiuk, T., Balasubramaniam, D., Sugihara, J., Mari, S.A., Kaback, H.R., and Müller, D.J. (2016). YidC assists the stepwise and stochastic folding of membrane proteins. Nat. Chem. Biol. *advance online publication*.

Thoma, J., Burmann, B.M., Hiller, S., and Müller, D.J. (2015). Impact of holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins. Nat. Struct. Mol. Biol. *22*, 795–802.

Chapter 4

Mapping the Energy Landscape for Second-Stage Folding of a Single Membrane Protein

nature chemical biology

ARTICLE

PUBLISHED ONLINE: 19 OCTOBER 2015 | DOI: 10.1038/NCHEMBIO.1939

Mapping the energy landscape for second-stage folding of a single membrane protein

Duyoung Min[1,2,4,5], Robert E Jefferson[3,5], James U Bowie[3]* & Tae-Young Yoon[1,2]*

Membrane proteins are designed to fold and function in a lipid membrane, yet folding experiments within a native membrane environment are challenging to design. Here we show that single-molecule forced unfolding experiments can be adapted to study helical membrane protein folding under native-like bicelle conditions. Applying force using magnetic tweezers, we find that a transmembrane helix protein, *Escherichia coli* rhomboid protease GlpG, unfolds in a highly cooperative manner, largely unraveling as one physical unit in response to mechanical tension above 25 pN. Considerable hysteresis is observed, with refolding occurring only at forces below 5 pN. Characterizing the energy landscape reveals only modest thermodynamic stability (ΔG = 6.5 k_BT) but a large unfolding barrier (21.3 k_BT) that can maintain the protein in a folded state for long periods of time ($t_{1/2}$ ~3.5 h). The observed energy landscape may have evolved to limit the existence of troublesome partially unfolded states and impart rigidity to the structure.

Helical membrane protein folding can be broken down into two major stages[1,2]. The first stage is initial insertion of transmembrane helices, which appears to be largely governed by the water-membrane partitioning of free energy[3]. In the second stage, the protein completes folding to its final native structure. Thus, once insertion occurs, membrane protein folding and unfolding occurs within the membrane. Ideally, studies of the second stage of folding would be performed in a membrane environment, yet folding studies require a means for altering the energy landscape to favor the unfolded state, which is hard to achieve in a membrane. One method, called steric trapping, drives unfolding by using a protein that binds preferentially to the unfolded state[4,5]. Atomic force microscopy (AFM) has been extensively used to study forced unfolding of membrane proteins from bilayers[6–9]. The AFM studies, however, apply force parallel to the membrane normal so that the proteins are physically pulled out of the membrane. To study the more physiological process of folding within a membrane, it is necessary to apply force along the membrane plane.

Here we developed a new method to observe the forced unfolding and refolding behavior of a single membrane protein within a lipid bilayer environment. By adapting techniques pioneered for soluble protein folding[10,11], we hold a single membrane protein in a magnetic trap and provide a bilayer environment for the protein using bicelles, self-assembled bilayer discs wrapped by detergent molecules[12–14]. We use this magnetic trapping strategy to study folding and unfolding of a helical membrane protein, GlpG. GlpG is an *E. coli* rhomboid intramembrane protease that has six transmembrane α-helices[15–18] and cleaves other transmembrane substrates in a lipid bilayer[19–23]. Previously reported extensive bulk equilibrium and kinetic folding studies on GlpG mutants in detergent provide a useful comparison[24,25]. Because GlpG has an even number of helices, the pulling direction is exactly defined along the membrane plane when the N and C termini of GlpG are pulled.

We found a remarkably high degree of cooperativity and a high barrier to unfolding, so large forces were required to unfold the protein at an appreciable rate. To see refolding at a measurable rate, we must return to much lower forces. Thus, we were unable to observe reversible folding directly. Nevertheless, we could construct a putative energy landscape by extrapolating the observed folding and unfolding rates to zero force. We found that GlpG is held close to its native state by a deep energy well near the folded conformation. The energy landscape is ideal for preventing the formation of misfolded states both during insertion and after the protein is synthesized.

RESULTS

Cooperative unfolding and refolding of GlpG in bicelles

Single GlpG proteins were covalently linked to two DNA handles (512 base pairs each) at the N- and C-terminal ends[10,26,27] (**Fig. 1a** and **Supplementary Results**, **Supplementary Fig. 1**). One DNA handle was anchored to a PEG-coated surface via biotin-avidin binding, and the other handle was attached to a magnetic bead. As a pair of magnets approaches, the magnetic bead experiences increasing force of up to tens of pN, which is then delivered to the tweezed GlpG protein[27–33]. The change in the bead height (i.e., the extension value) as a result of the force application can be measured (**Fig. 1a**). With this experimental scheme, we are able to apply tension in a direction vertical to the membrane normal vector, allowing the GlpG protein to unfold and refold within the lipid membrane (**Supplementary Fig. 2**). The experiment is free from nonspecific interactions with the surface because the DNA handles completely separate the GlpG protein from the surface.

Gradual pulling experiments with GlpG, in which the force was slowly increased as the magnets approached the sample at a constant speed (0.1 mm s^{-1}, corresponding to an average force-loading rate of ~0.5 pN s^{-1}), revealed a high degree of unfolding cooperativity (**Fig. 1b**). The GlpG protein remained intact until the magnetic force was increased to ~25 pN and then showed one abrupt unfolding event with a step size of 40 nm (**Fig. 1b**). The observed 40-nm increase was very close to the expected value when a fully folded GlpG was unfolded to a completely unstructured polypeptide at 25 pN (**Supplementary Fig. 3**). Thus, although unfolding was initiated in a bicelle environment, driving unfolding with a reasonable probability required such high force that the entire protein ultimately unraveled. Our observation suggests highly cooperative unfolding of the entire GlpG protein.

[1]National Creative Research Initiative Center for Single-Molecule Systems Biology, KAIST, Daejeon, South Korea. [2]Department of Physics, KAIST, Daejeon, South Korea. [3]Department of Chemistry and Biochemistry, University of California–Los Angeles, Los Angeles, California, USA. [4]Present address: Department of Chemistry and Biochemistry, University of California–Los Angeles, Los Angeles, California, USA. [5]These authors contributed equally to this work. *e-mail: bowie@mbi.ucla.edu or tyyoon@kaist.ac.kr

We observed a large refolding hysteresis. The unfolded GlpG protein only refolded when we decreased the force to a few pN so the unfolding and refolding cycle had a force gap of more than 20 pN (**Fig. 1b**). At these low forces, the transmembrane helical structure could be restored before refolding, allowing refolding within a protein–bicelle complex (**Supplementary Fig. 3**). The unfolding and refolding cycle could be repeated up to tens of times in a very reproducible manner, indicating that, in spite of the hysteresis, the protein completely refolded using our experimental setup.

The experimental system was remarkably robust. We removed the bicelles by buffer exchange, leaving the hydrophobic polypeptide in an aqueous environment (**Fig. 1b** and **Supplementary Fig. 4**). Under these conditions, we saw a single unfolding event at low (5 pN) forces, and the large hysteresis completely vanished (**Fig. 1b** and **Supplementary Fig. 5d**). When the bicelle condition was restored by another round of buffer exchange, however, the unfolding and refolding behavior of the GlpG protein was fully restored (**Fig. 1b**). Moreover, addition of detergent molecules alone, instead of bicelles, substantially decreased the unfolding force and made its distribution much more heterogeneous (**Supplementary Fig. 5**). These observations indicate that the bicelle condition has a crucial role in the cooperative unfolding and refolding of GlpG and also point to the advantage of the single-molecule tethering approach for studying the folding of membrane proteins that are so prone to irreversible aggregation.

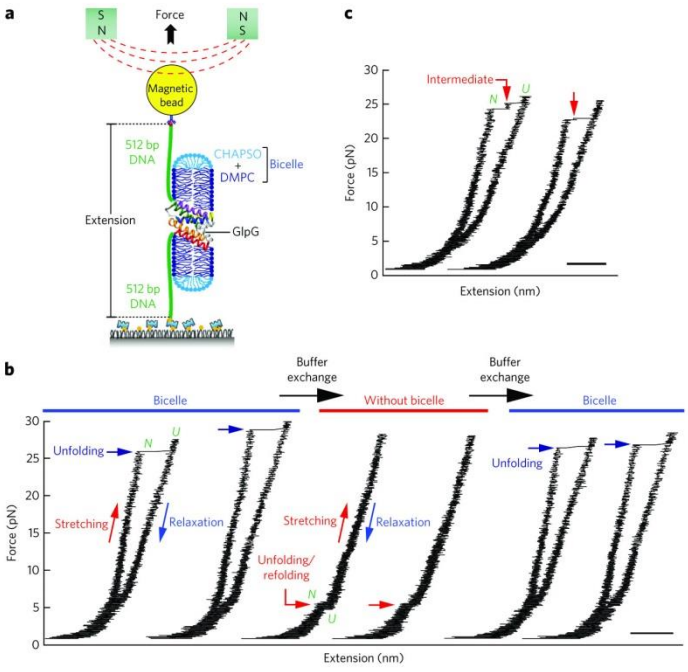


Figure 1 | Cooperative unfolding and refolding of GlpG in bicelles. (**a**) Schematic of the single-molecule magnetic tweezers experiment for studying unfolding and refolding of a single GlpG protein. (**b**) Representative force-extension curves in each buffer condition. After several cycles of unfolding and refolding in bicelles (left), the bicelles were removed and the unfolding and refolding cycles were repeated (middle). In the buffer condition without bicelles, a very small amount of CHAPSO (0.0038%) was added to prevent nonspecific binding. After up to tens of pulling cycles, the bicelle condition could be restored by another round of microfluidic buffer exchange (right), and the unfolding behavior seen previously in bicelles was fully restored. (**c**) Representative force-extension curves showing multiple-step unfolding of single GlpG proteins. In **b** and **c**, scale bar represents 50 nm.

Intermediates in C- to N-terminal unfolding of GlpG

Although unfolding of GlpG was essentially a cooperative process, we noticed transient intermediates, or pauses, during unfolding (**Fig. 1c**). We sought to identify where these unfolding pauses occurred. In the pulling experiment shown in **Figure 1**, however, where mechanical tension was gradually increased, the unfolding events stochastically occurred at different force levels, which precluded direct comparison of observed step sizes. We therefore designed a 'force-jump' experiment where the magnetic force was rapidly increased and maintained at a predetermined value[27] (**Fig. 2**). In such force-jump experiments, unfolding of GlpG was induced at a constant force level, and the observed extension increases could be pooled together to elucidate the structure of unfolding intermediates[34–36].

When we employed force jumps to 21 pN, we were able to observe four different patterns in the unfolding of single GlpG proteins (**Fig. 2a** and **Supplementary Fig. 6**). In about 60% of the total unfolding trajectories (n = 295), no intermediates were resolved with our time resolution. In one-third of the trajectories, one unfolding intermediate was detected. The extension distribution of these intermediates showed two Gaussian peaks with one peak at ~10 nm (I_1) and the other at ~20 nm (I_2). In 7.8% of the trajectories, we observed two intermediates. These two intermediates almost exactly overlap with the I_1 and I_2 intermediates observed for the one-intermediate cases, suggesting that I_1 and I_2 do not represent two independent pathways but two intermediates along one unfolding pathway (**Supplementary Fig. 6**). Finally, we measured the dwell times in I_1 and I_2 (τ_1 and τ_2) and compared these dwell times with the total unfolding time, τ_U, which was the time elapsed between the force jump to 21 pN and the moment of complete unfolding (**Fig. 2b**). Both τ_1 and τ_2 were <2% of τ_U, quantitatively showing that the unfolding process of GlpG had essentially one rate-limiting step and paused only briefly in the I_1 and I_2 intermediate states after the rate-limiting step (**Fig. 2c**). In fact, no intermediates were observed in 60% of the unfolding trajectories with our current time resolution. This dwell-time analysis quantitatively illustrates again that the intermediates are very transient compared to the total unfolding; thus, the unfolding of the entire GlpG protein is highly cooperative.

We reasoned that the observed unfolding is a unidirectional process beginning at either the N or C terminus (**Supplementary Fig. 7**). To determine the directionality of the unfolding process, we examined the unfolding of GlpG^{L155A} and GlpG^{A206G} (refs. 24,25), whose N- and C-terminal parts are respectively destabilized by their mutation (**Fig. 2d**,**e** and **Supplementary Fig. 8**). For the L155A mutant, the observation probability of the I_2 intermediate was selectively reduced compared to the wild type (WT) (**Fig. 2d** and **Supplementary Fig. 8**). Thus, this N-terminal mutation lowered the stability of the region that was unfolded in the I_2 to U step, and the

corresponding unfolding step (from I_2 to U) became accelerated. In contrast, for the A206G mutant, the observation probability of the I_1 intermediate was selectively increased (**Fig. 2d** and **Supplementary Fig. 8**). Thus, the C-terminal mutation of A206G accelerated the N to I_1 step by lowering the stability of the corresponding region. These observations collectively suggest that mechanical unfolding of GlpG starts at the C terminus and propagates toward the N-terminal end.

We next pinpointed the residues comprising the unfolding intermediates. Using the Marko-Siggia formulation of the worm-like chain model (**Supplementary Fig. 9**), the I_1 intermediate was found to extend to approximately residue 221, which corresponds to the unfolding of helices 5 and 6 (**Fig. 2e**,**f**). The I_2 intermediate was found to extend until approximately position 177, corresponding to the unfolding of helices 3 and 4. These results suggest that mechanical unfolding of GlpG at high forces takes place in units of helical hairpins, two helices at a time.

Characterization of folding and unfolding kinetics

Our observations of the unfolding and refolding of GlpG collectively point to the existence of one main energy barrier that separates the folded and unfolded states (**Fig. 2g**), with minor energy barriers separating the I_1 and I_2 intermediates located between the primary energy barrier and the unfolded U state. Crossing of the main energy barrier becomes the rate-limiting step for unfolding, and, once crossed, the unfolding process only briefly pauses in the I_1 and I_2 states.

To characterize the main unfolding energy barrier in a quantitative way, we studied the unfolding and refolding kinetics. We first revisited the gradual pulling experiments of **Figure 1** (**Fig. 3a**). As noted above, each unfolding event stochastically occurred at a different force level, meaning that we could study the unfolded fraction as a function of force (**Fig. 3b**). Fitting this unfolding probability (Online Methods) yielded a kinetic rate for GlpG unfolding at zero tension (k_{u0}) of 5.64×10^{-5} s^{-1} and a distance from the folded state to the transition state ($\Delta x_f^{\dagger}$) of 1.48 nm. To characterize the opposite side of the energy barrier, we repeated the refolding experiments but varied the force levels during refolding (**Fig. 3c**). After waiting 3 min, we checked the folding status by pulling the GlpG protein at 21 pN to determine whether the extension reflected the U or N state. We studied the folded fraction (within the given 3 min) as a function of mechanical tension. By extrapolation, we estimated the kinetic rate for folding at zero tension (k_{f0}) to be 3.91×10^{-2} s^{-1} and the distance from the unfolded state to the transition state ($\Delta x_u^{\dagger}$) to be 3.56 nm (**Fig. 3d** and Online Methods). These data reporting unfolding and refolding kinetics as a function of force are analogous to chevron plots in bulk membrane protein folding studies in detergent that report kinetic parameters as a function of denaturant concentration. Our reaction coordinate (x) is conceptually a thermally averaged end-to-end distance of GlpG measured at zero force along the pulling direction, and the distance (Δx) indicates how x changes[37,38]. Thus, the mechanical tension and distance to the transition state are analogous to the denaturant concentration and its denaturant power (reflected in the m values) used in the bulk folding studies, but the mechanical parameters have direct physical implications.

To test whether the kinetic rates we determined were affected by the specific bicelle conditions, we repeated the unfolding measurements at different lipid/detergent ratios and temperature conditions (**Supplementary Fig. 10** and **Supplementary Tables 1** and **2**).

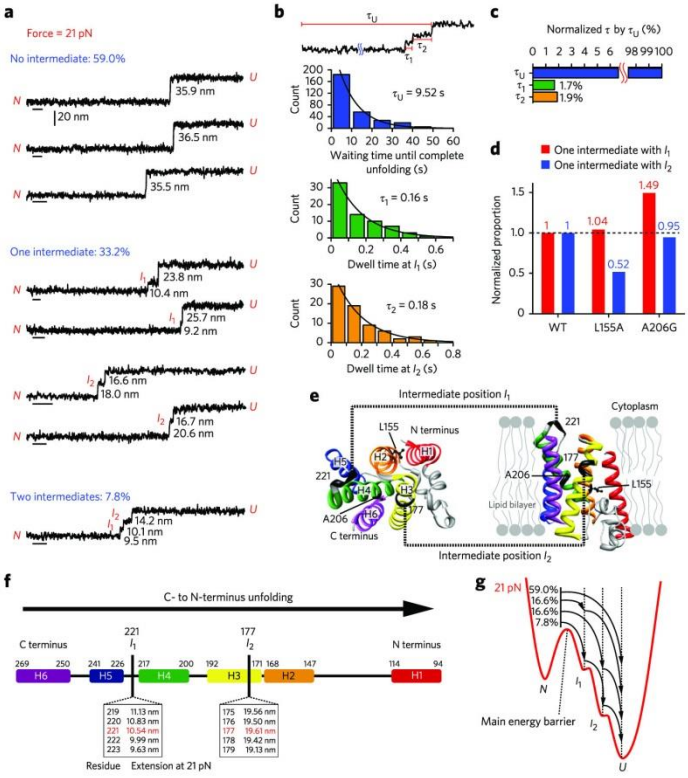


Figure 2 | C- to N-terminus unfolding of single GlpG with two intermediates. (**a**) Representative extension traces at 21 pN for unfolding events (n = 295) with no intermediates (59.0%), one intermediate (33.2%) and two intermediates (7.8%). Statistics of unfolding step sizes are in **Supplementary Figure 6**. Scale bars, 1 s. (**b**) Dwell time analysis (n = 295). τ_U is the waiting time until complete unfolding (blue), and τ_1 and τ_2 are the dwell times in the intermediate states I_1 (green) and I_2 (yellow). (**c**) Dwell times in the intermediates normalized by τ_U. (**d**) Comparison of the normalized proportion of unfolding patterns with one intermediate between the WT and the L155A and A206G mutants. The normalized proportion is defined by $P(X)/P(\mathrm{WT})$, where $P(X)$ means the proportion of each unfolding pattern in the total number of traces for X = WT (n = 295), L155A (n = 81) or A206G (n = 97). The histograms for one intermediate with I_1 or I_2 are shown in red and blue, respectively. (**e**) GlpG structure showing the intermediate positions I_1 and I_2 (black). The mutation sites Leu155 and Ala206 are shown in ball-and-stick representation. Left, cytoplasmic view; right, side view showing a lipid bilayer. (**f**) Schematic diagram showing the mapping of Gaussian peak values to the intermediate residue positions. Arrow indicates unfolding direction. (**g**) Conceptual folding energy landscape at 21 pN. The arrows denote the structural transitions among the native state (N), the intermediate states (I_1 and I_2) and the unfolded state (U).

When the lipid/detergent ratio was increased from 2.2:1 to 2.8:1, the kinetic rates and the distance to the transition state were largely unaffected (with only a 0.6 k_BT difference; **Supplementary Table 1**), indicating that the edge effects of the detergent belt surrounding the bicelle structure were negligible. The gel phase melting temperature of 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC)/1,2-diheptanoyl-*sn*-glycero-3-phosphocholine (DHPC) bicelles (analogous to our DMPC/3-((3-cholamidopropyl)dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO) bicelles) is 21 °C (ref. 39), which is close to the temperature (22 °C) used in our experiments, so we tested whether increasing the temperature would have any effect. When we increased the measurement temperature up to 25 °C, however, we did not see any obvious change in the kinetic rates or the transition state distance (**Supplementary Table 2**). Thus, our results do not seem highly sensitive to small changes in the bicelle conditions.

Folding energy landscape of GlpG

In characterizing the unfolding and refolding kinetics, we needed to use different force ranges (**Fig. 3b**,**d**) because of the large hysteresis observed in the unfolding and refolding cycle of GlpG (**Fig. 1**). Nevertheless, we believe that we can reconstruct an energy landscape at zero force within bicelles by extrapolation if we assume that the transition state for the unfolding induced by high force levels is the same as that of the refolding pathway observed at low force levels. We believe this is a reasonable assumption because: (i) We do not see discontinuities in the unfolding and refolding rates as a

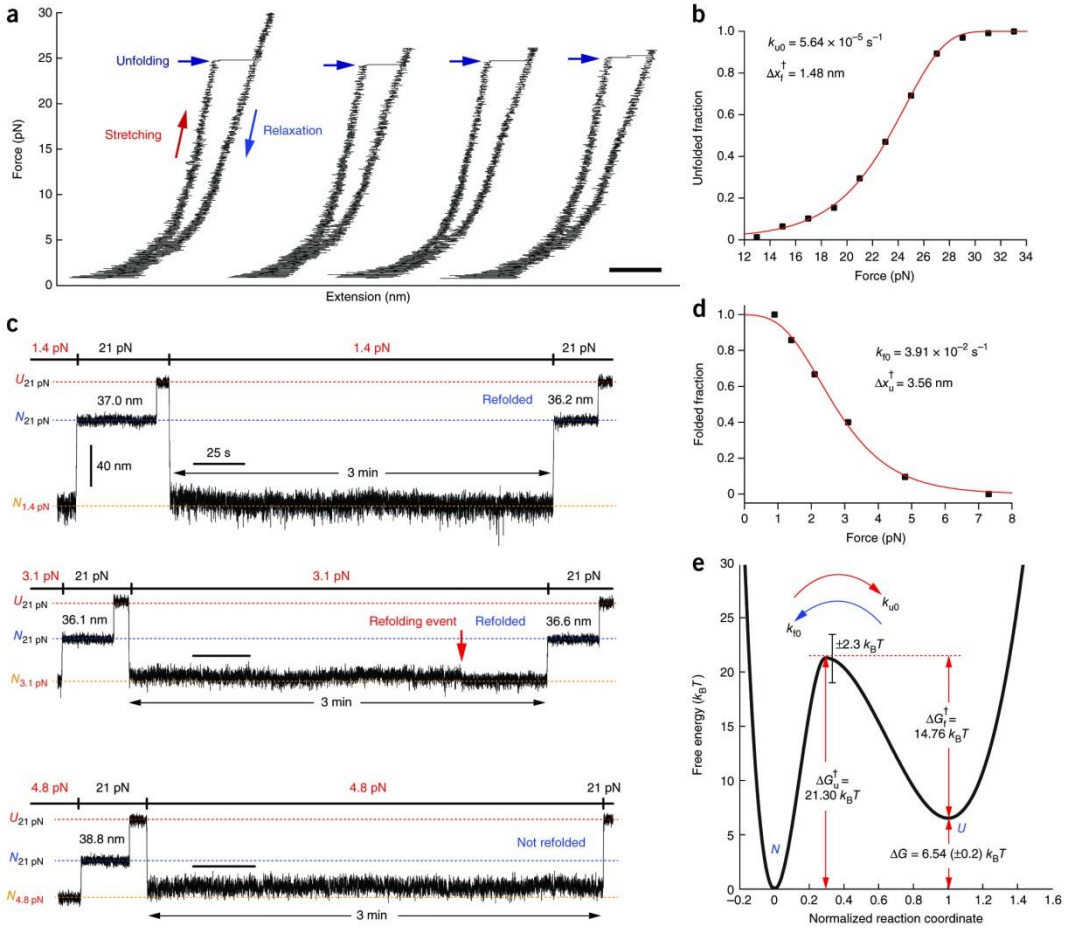


Figure 3 | Folding energy landscape of GlpG. (**a**) Representative gradual pulling experiments measuring the unfolding force of single GlpG proteins. Scale bar, 50 nm. (**b**) Unfolded fraction versus force (n = 233) from which the zero-force unfolding rate (k_{u0}) and the distance from the native state to the transition state ($\Delta x_f^\dagger$) were obtained. (**c**) Representative extension traces in force-cycle experiment for obtaining the folding kinetics. After unfolding at 21 pN, the force was lowered back to 0.9–7.3 pN and maintained for 3 min. The extent of refolding was then determined by restoring the 21-pN force and comparing the observed extension with the extensions observed for the native and unfolded states ($N_{21\,pN}$ and $U_{21\,pN}$ are shown as blue and red dashed lines, respectively). (**d**) Folded fraction versus force (n = 125), which was used to obtain the folding kinetic rate at zero force (k_{f0}) and the distance from the unfolded state to the transition state ($\Delta x_u^\dagger$). (**e**) Putative folding energy landscape of GlpG. The energy difference between the native state and the unfolded state (ΔG) and the energy barriers ($\Delta G_u^\dagger$, $\Delta G_f^\dagger$) are denoted with red arrows. The error of the ΔG represents s.e.m., and the error of the energy barriers represent the error of the frequency factor k_w (Online Methods).

Table 1 | Summary of kinetic and thermodynamic properties of WT and mutant GlpG

	$\Delta x_f^\dagger$ (nm)	$\Delta x_u^\dagger$ (nm)	β_f	k_{u0} (s^{-1})	k_{f0} (s^{-1})	$\Delta G_u^\dagger$ (k_BT)	$\Delta G_f^\dagger$ (k_BT)	ΔG (k_BT)
WT	1.48 (0.03)	3.56 (0.24)	0.29 (0.02)	$5.64\ (0.91) \times 10^{-5}$	$3.91\ (0.54) \times 10^{-2}$	21.30 (2.30)	14.75 (2.30)	6.54 (0.21)
L155A	1.26 (0.09)	2.99 (0.34)	0.29 (0.03)	$1.73\ (0.81) \times 10^{-4}$	$1.37\ (0.18) \times 10^{-2}$	20.18 (2.30)	15.80 (2.30)	4.37 (0.48)
A206G	1.68 (0.10)	3.57 (0.17)	0.32 (0.02)	$1.17\ (0.44) \times 10^{-4}$	$2.42\ (0.19) \times 10^{-2}$	20.57 (2.30)	15.23 (2.30)	5.33 (0.38)

Numbers in parentheses indicate error. For $\Delta G_u^\dagger$ and $\Delta G_f^\dagger$, error represents the error of the frequency factor k_w (Online Methods). All other error values represent s.e.m.

function of force that would imply a change in pathway (**Fig. 3b,d**). The two force regions used for unfolding (13–33 pN) and refolding (1–7 pN) are separated only by 6 pN. (ii) Even though the GlpG–bicelle complex must ultimately become highly distorted as GlpG is unraveled to an unstructured polypeptide at high forces, unfolding is initiated within the bicelle structure, and the distance to the transition state is only 1.5 nm (**Figs. 1b** and **3b**). Thus, unfolding rates reflect unfolding within a bicelle. (iii) Helical structure is restored at low forces, indicating that refolding occurs again within the protein–bicelle complex (**Supplementary Fig. 3**). (iv) Finally, we find that our measured thermodynamic values (ΔG and $\Delta\Delta G$) for the WT and mutants are largely consistent with the values from the bulk equilibrium unfolding experiments described below (**Supplementary Table 3**).

With the assumption we made above, we constructed a putative folding energy landscape of GlpG (**Fig. 3e** and **Table 1**). The ratio of the unfolding and folding rates at zero force led to an unfolding free energy of $\Delta G = -k_BT \times \ln(k_{u0}/k_{f0}) = 6.54\ k_BT$. Bulk SDS unfolding experiments report ΔG values from $7.08\ k_BT$ to $13.88\ k_BT$ (refs. 24,25), which is in reasonable agreement considering the completely different methods for driving unfolding, the different environments and the uncertain extrapolations in SDS unfolding studies[5] (**Supplementary Table 3**). Our measured refolding rate of $3.91 \times 10^{-2}\ s^{-1}$ is very similar to the rate measured in the detergent refolding experiments ($2.7 \times 10^{-2}\ s^{-1}$), a parameter that does not require much extrapolation. The main discrepancy occurs in the unfolding rates ($1.0 \times 10^{-7}\ s^{-1}$ versus $5.64 \times 10^{-5}\ s^{-1}$), but this involves a large extrapolation. Using the Kramer equation (Online Methods), the height of the energy barrier encountered during GlpG folding ($\Delta G_f^\dagger$) was estimated to be $14.76\ k_BT$, rendering the folding process slow ($t_{1/2}$ ~18s). We also mapped the transition state onto the normalized reaction coordinate $x/(\Delta x_f^\dagger+\Delta x_u^\dagger)$, where the $\Delta x_f^\dagger$ and $\Delta x_u^\dagger$ are respectively from the folded and unfolded state to the transition state. Notably, the transition state turned out to be much closer to the native state than to the unfolded state (i.e., $\beta_f \equiv \Delta x_f^\dagger/(\Delta x_f^\dagger+\Delta x_u^\dagger) = 0.29$), consistent with our observations that the six transmembrane helices were tightly coupled and essentially worked as one unit when GlpG was folded and unfolded.

We also studied how the L155A and A206G mutations affected the energy landscape (**Fig. 4** and **Table 1**). As in our previous study of unfolding patterns (**Fig. 2d**), kinetic measurements of the two mutants revealed detailed changes in the unfolding and refolding probabilities as a function of force (**Fig. 4a,b**). These data reconfirm that our measurements do not simply measure disruption and

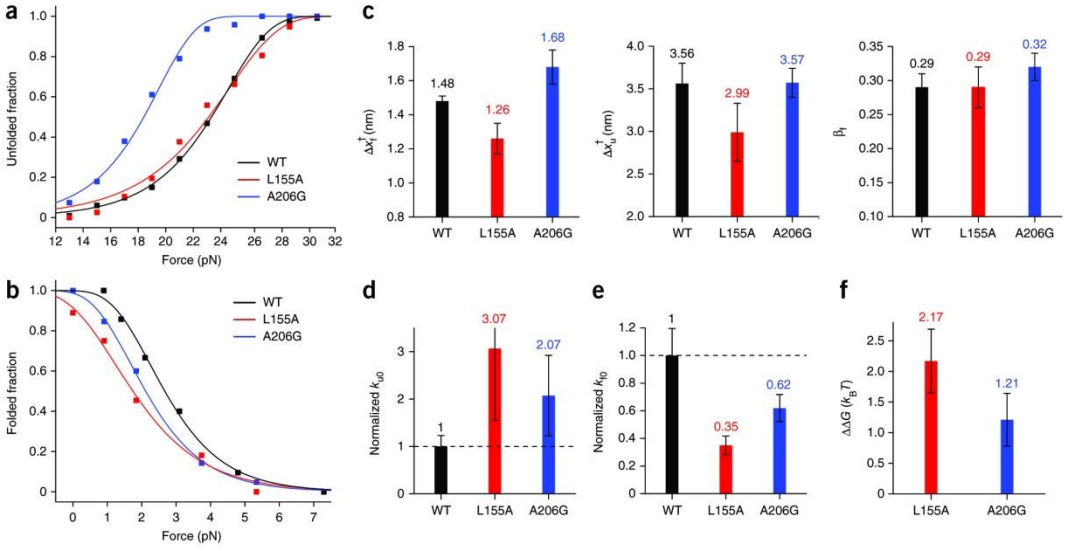


Figure 4 | Comparison of kinetic and thermodynamic properties between WT and mutant GlpG. (**a**,**b**) Unfolded fraction (**a**) and folded fraction (**b**) as a function of force for the WT and mutant GlpG proteins. The total number of unfolding and refolding events are $n = 233$ and $n = 125$ for WT; $n = 77$ and $n = 58$ for the L155A mutant; and $n = 95$ and $n = 87$ for the A206G mutant. Fitting the data (Online Methods) yields kinetic rates for unfolding and folding at zero tension (k_{u0} and k_{f0}) and distances from the folded (and unfolded) state to the transition state ($\Delta x_f^\dagger$ and $\Delta x_u^\dagger$). (**c**) Comparison of the distance values (left, $\Delta x_f^\dagger$; middle, $\Delta x_u^\dagger$) and the transition state positions (right, β_f) of the WT and the mutants. (**d**,**e**) Comparison of the unfolding rate (**d**) and refolding rate (**e**) for the WT and mutant proteins, normalized to the WT rate. (**f**) The change in unfolding free energy of the mutants relative to the WT observed in forced unfolding experiments ($\Delta\Delta G = \Delta G_{WT} - \Delta G_{mutant}$). All error bars represent s.e.m.

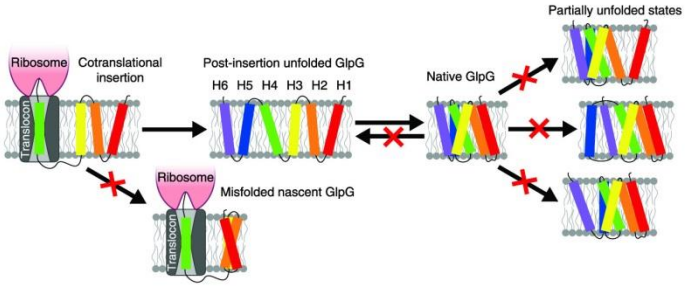


Figure 5 | How the folding energy landscape of GlpG may prevent dangerous misfolded states. Cooperativity can limit the formation of stable off-pathway structures before completion of translation. The high kinetic barrier near the folded state prevents folded GlpG from returning to the unfolded state on a biologically relevant time scale, imparts rigidity and limits the existence of partially unfolded states that might be prone to inappropriate interactions.

association of bicelle complexes but rather reflect subtle differences in the energy landscapes of the mutants. At the same time, however, the general shapes of the energy landscape were essentially preserved for the two mutants. The position of the transition state (β_t) remained at the normalized distance of 0.3, close to the native folded state (**Fig. 4c**). The unfolding and the refolding rates were modestly changed for the two mutants. The unfolding rates at zero force (k_{u0}) were increased by a factor of two or three, corresponding to lowering of the unfolding energy barrier ($\Delta G_u^{\dagger}$) by ~1 k_BT (**Fig. 4d** and **Table 1**). Although the force values reaching 50% unfolding were similar to those of the L155A mutant and the WT (**Fig. 4a**), the difference in the unfolding curve slopes gave a smaller $\Delta x_t^{\dagger}$ for the L155A mutant (**Fig. 4c**), which in turn led to a higher k_{u0} (**Fig. 4d**). The refolding rates were decreased by almost the same factors, indicating that the refolding energy barrier was increased as much as the unfolding energy barrier was decreased (**Fig. 4e** and **Table 1**). The calculated destabilizing extents, $\Delta\Delta G$ (calculated as $\Delta G_{WT} - \Delta G_{mutant}$), were thus in the range of 1–2 k_BT (**Fig. 4f**), consistent with the values obtained from the bulk SDS unfolding experiments[24,25] (**Supplementary Table 3**). Thus, our method of reconstructing energy landscapes is sensitive to modest changes in the intrinsic stability of GlpG.

DISCUSSION

Overall, the primary features of the folding energy landscape for GlpG we observe are (i) high cooperativity, (ii) low thermodynamic stability, (iii) a high kinetic barrier and (iv) a transition state that is structurally closer to the folded state than the unfolded state. As there is still limited information on the folding of large helical membrane proteins, particularly under native conditions, it is unclear how common these characteristics of GlpG folding will be for membrane proteins in general.

In contrast to what we see for GlpG, the transition states found with SDS-driven unfolding of bacteriorhodopsin, DsbB and even GlpG are all placed closer to the unfolded state than the folded state[25,40,41]. It is possible that the difference simply reflects folding and unfolding in the more native-like bicelle. In contrast, it could reflect different requirements for structural flexibility. The close proximity of the energy barrier to the folded state would imply high local curvature of the energy landscape around the folded state, which could impart structural rigidity to GlpG.

The high degree of cooperativity in mechanical unfolding for a helical membrane protein was surprising to us. Individual transmembrane helices are stable within a bilayer[1], so we expected that helices could be pulled off one at a time. Instead, the six transmembrane helices largely behave as one unfolding unit. The folding of bacteriorhodopsin and GlpG also seems to be highly cooperative when studied by SDS unfolding[25,40], so it is possible that this is a common property of membrane proteins. Although cooperativity is not theoretically required, there may also be evolutionary pressure favoring cooperativity in membrane proteins as in soluble proteins[42]. Cooperativity would prevent the formation of structure before complete insertion, thereby limiting the development of stable, albeit incorrect partial structures before the entire protein is available for structure formation (**Fig. 5**). Once formed, these misfolded structures might be difficult to unravel. Thus, it makes evolutionary sense to select an energy function that requires the protein to wait until complete insertion before adopting a stable structure. There is evidence for some structure formation during biological insertion[43,44] and for preferred folding from the N terminus[25,45], but it is unclear whether partially inserted states can generate stable enough structures to direct folding.

A high kinetic barrier for unfolding, as signified by our observation of cooperative unfolding, represents another mechanism for limiting the existence of aggregation-prone unfolded states (**Fig. 5**). Although GlpG is not very thermodynamically stable, once folded, the 3.5-h unfolding half-life (because of the high $\Delta G_u^{\dagger}$ of 21.30 k_BT) implies that an *E. coli* cell will rarely see a GlpG unfold on the time scale of a cell division. Very slow unfolding has also been observed for both diacylglycerol kinase and bacteriorhodopsin[5,40,46]. It is not known how slowly DsbB unfolds under native conditions, but it refolds from an SDS-denatured state on a similar time scale to bacteriorhodopsin and GlpG[47]. Slow folding may reflect a rugged energy landscape for membrane proteins[48].

Like soluble proteins, it is likely that there will be wide variation in the folding behavior of membrane proteins. We need to see more examples of well-characterized folding landscapes under native conditions to learn about structural correlations with folding properties. The approach described here may now allow us to expand our analysis of membrane protein folding to more proteins.

Received 13 July 2015; accepted 14 September 2015;
published online 19 October 2015

METHODS

Methods and any associated references are available in the online version of the paper.

References

1. Engelman, D.M. *et al.* Membrane protein folding: beyond the two stage model. *FEBS Lett.* **555**, 122–125 (2003).
2. Bowie, J.U. Solving the membrane protein folding problem. *Nature* **438**, 581–589 (2005).
3. White, S.H. & von Heijne, G. How translocons select transmembrane helices. *Annu. Rev. Biophys.* **37**, 23–42 (2008).
4. Hong, H., Blois, T.M., Cao, Z. & Bowie, J.U. Method to measure strong protein-protein interactions in lipid bilayers using a steric trap. *Proc. Natl. Acad. Sci. USA* **107**, 19802–19807 (2010).
5. Chang, Y.C. & Bowie, J.U. Measuring membrane protein stability under native conditions. *Proc. Natl. Acad. Sci. USA* **111**, 219–224 (2014).
6. Oesterhelt, F. *et al.* Unfolding pathways of individual bacteriorhodopsins. *Science* **288**, 143–146 (2000).
7. Kedrov, A., Janovjak, H., Sapra, K.T. & Muller, D.J. Deciphering molecular interactions of native membrane proteins by single-molecule force spectroscopy. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 233–260 (2007).
8. Engel, A. & Gaub, H.E. Structure and mechanics of membrane proteins. *Annu. Rev. Biochem.* **77**, 127–148 (2008).

9. Zocher, M. *et al.* Single-molecule force spectroscopy from nanodiscs: an assay to quantify folding, stability, and interactions of native membrane proteins. *ACS Nano* **6**, 961–971 (2012).
10. Cecconi, C., Shank, E.A., Bustamante, C. & Marqusee, S. Direct observation of the three-state folding of a single protein molecule. *Science* **309**, 2057–2060 (2005).
11. Shank, E.A., Cecconi, C., Dill, J.W., Marqusee, S. & Bustamante, C. The folding cooperativity of a protein is controlled by its chain topology. *Nature* **465**, 637–640 (2010).
12. Faham, S. & Bowie, J.U. Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *J. Mol. Biol.* **316**, 1–6 (2002).
13. Joh, N.H. *et al.* Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. *Nature* **453**, 1266–1270 (2008).
14. Dürr, U.H., Gildenberg, M. & Ramamoorthy, A. The magic of bicelles lights up membrane protein structure. *Chem. Rev.* **112**, 6054–6074 (2012).
15. Wang, Y.C., Zhang, Y.J. & Ha, Y. Crystal structure of a rhomboid family intramembrane protease. *Nature* **444**, 179–183 (2006).
16. Wu, Z.R. *et al.* Structural analysis of a rhomboid family intramembrane protease reveals a gating mechanism for substrate entry. *Nat. Struct. Mol. Biol.* **13**, 1084–1091 (2006).
17. Ben-Shem, A., Fass, D. & Bibi, E. Structural basis for intramembrane proteolysis by rhomboid serine proteases. *Proc. Natl. Acad. Sci. USA* **104**, 462–466 (2007).
18. Vinothkumar, K.R. Structure of rhomboid protease in a lipid environment. *J. Mol. Biol.* **407**, 232–247 (2011).
19. Lemberg, M.K. & Freeman, M. Cutting proteins within lipid bilayers: rhomboid structure and mechanism. *Mol. Cell* **28**, 930–940 (2007).
20. Freeman, M. Rhomboid proteases and their biological functions. *Annu. Rev. Genet.* **42**, 191–210 (2008).
21. Ha, Y., Akiyama, Y. & Xue, Y. Structure and mechanism of rhomboid protease. *J. Biol. Chem.* **288**, 15430–15436 (2013).
22. Vinothkumar, K.R. & Freeman, M. Intramembrane proteolysis by rhomboids: catalytic mechanisms and regulatory principles. *Curr. Opin. Struct. Biol.* **23**, 851–858 (2013).
23. Lemberg, M.K. Sampling the membrane: function of rhomboid-family proteins. *Trends Cell Biol.* **23**, 210–217 (2013).
24. Baker, R.P. & Urban, S. Architectural and thermodynamic principles underlying intramembrane protease function. *Nat. Chem. Biol.* **8**, 759–768 (2012).
25. Paslawski, W. *et al.* Cooperative folding of a polytopic α-helical membrane protein involves a compact N-terminal nucleus and nonnative loops. *Proc. Natl. Acad. Sci. USA* **112**, 7978–7983 (2015).
26. Cecconi, C., Shank, E.A., Marqusee, S. & Bustamante, C. in *DNA Nanotechnology: Methods and Protocols* 1st edn. Vol. 749 (eds. Zuccheri, G. & Samorì, B.) 255–271 (Humana Press, 2011).
27. Min, D. *et al.* Mechanical unzipping and rezipping of a single SNARE complex reveals hysteresis as a force-generating mechanism. *Nat. Commun.* **4**, 1705–1714 (2013).
28. Bae, W. *et al.* Programmed folding of DNA origami structures through single-molecule force control. *Nat. Commun.* **5**, 5654–5661 (2014).
29. Gosse, C. & Croquette, V. Magnetic tweezers: micromanipulation and force measurement at the molecular level. *Biophys. J.* **82**, 3314–3329 (2002).
30. Saleh, O.A., Allemand, J.F., Croquette, V. & Bensimon, D. Single-molecule manipulation measurements of DNA transport proteins. *ChemPhysChem* **6**, 813–818 (2005).
31. Kim, K. & Saleh, O.A. A high-resolution magnetic tweezer for single-molecule measurements. *Nucleic Acids Res.* **37**, 136–142 (2009).
32. Lipfert, J., Hao, X.M. & Dekker, N.H. Quantitative modeling and optimization of magnetic tweezers. *Biophys. J.* **96**, 5040–5049 (2009).
33. De Vlaminck, I. & Dekker, C. Recent advances in magnetic tweezers. *Annu. Rev. Biophys.* **41**, 453–472 (2012).
34. Greenleaf, W.J., Frieda, K.L., Foster, D.A.N., Woodside, M.T. & Block, S.M. Direct observation of hierarchical folding in single riboswitch aptamers. *Science* **319**, 630–633 (2008).
35. Alegre-Cebollada, J., Kosuri, P., Rivas-Pardo, J.A. & Fernandez, J.M. Direct observation of disulfide isomerization in a single protein. *Nat. Chem.* **3**, 882–887 (2011).
36. Stigler, J., Ziegler, F., Gieseke, A., Gebhardt, J.C.M. & Rief, M. The complex folding network of single calmodulin molecules. *Science* **334**, 512–516 (2011).
37. Carrion-Vazquez, M. *et al.* Mechanical and chemical unfolding of a single protein: A comparison. *Proc. Natl. Acad. Sci. USA* **96**, 3694–3699 (1999).
38. Liphardt, J., Onoa, B., Smith, S.B., Tinoco, I. Jr. & Bustamante, C. Reversible unfolding of single RNA molecules by mechanical force. *Science* **292**, 733–737 (2001).
39. Beaugrand, M. *et al.* Lipid concentration and molar ratio boundaries for the use of isotropic bicelles. *Langmuir* **30**, 6162–6170 (2014).
40. Curnow, P. & Booth, P.J. Combined kinetic and thermodynamic analysis of α-helical membrane protein unfolding. *Proc. Natl. Acad. Sci. USA* **104**, 18970–18975 (2007).
41. Otzen, D.E. Mapping the folding pathway of the transmembrane protein DsbB by protein engineering. *Protein Eng. Des. Sel.* **24**, 139–149 (2011).
42. Watters, A.L. *et al.* The highly cooperative folding of small naturally occurring proteins is likely the result of natural selection. *Cell* **128**, 613–624 (2007).
43. Cymer, F., von Heijne, G. & White, S.H. Mechanisms of integral membrane protein insertion and folding. *J. Mol. Biol.* **427**, 999–1022 (2015).
44. Kim, S.J. & Skach, W.R. Mechanisms of CFTR folding at the endoplasmic reticulum. *Front. Pharmacol.* **3**, 201 (2012).
45. Curnow, P. *et al.* Stable folding core in the folding transition state of an α-helical integral membrane protein. *Proc. Natl. Acad. Sci. USA* **108**, 14133–14138 (2011).
46. Jefferson, R.E., Blois, T.M. & Bowie, J.U. Membrane proteins can have high kinetic stability. *J. Am. Chem. Soc.* **135**, 15183–15190 (2013).
47. Otzen, D.E. Folding of DsbB in mixed micelles: a kinetic analysis of the stability of a bacterial membrane protein. *J. Mol. Biol.* **330**, 641–649 (2003).
48. Kim, B.L., Schafer, N.P. & Wolynes, P.G. Predictive energy landscapes for folding α-helical transmembrane proteins. *Proc. Natl. Acad. Sci. USA* **111**, 11031–11036 (2014).

Acknowledgments

This work was supported by the National Creative Research Initiative Program (Center for Single-Molecule Systems Biology to T.-Y.Y.) funded by the National Research Foundation of Korea and Marine Biotechnology Program (20150220 to T.-Y.Y.) funded by the Ministry of Oceans and Fisheries of Korea, and supported by US National Institutes of Health grant 2R01GM063919 to J.U.B.

Author contributions

D.M., R.E.J., J.U.B. and T.-Y.Y. designed the experiments. R.E.J. expressed and purified proteins. D.M. prepared the DNA-protein hybrid sample and performed the magnetic tweezers experiments. All of the authors analyzed the data and contributed to writing of the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Correspondence and requests for materials should be addressed to J.U.B. or T.-Y.Y.

ONLINE METHODS

Protein expression and purification. The membrane domain of the *E. coli* GlpG gene (residues 87–276) was amplified from the genome of *E. coli* XL1-Blue strain by PCR. The PCR primers included codons to add cysteine residues at both the N and C termini. The amplification primers were:

FWD: 5′-GGAAAGAGCTCTGTGCCGCCTTGCGTGAACGCG-3′
REV: 5′-CCCTTAAGCTTTTAACATTTTCGTTTTCGCGCATTGAGCG-3′.

The amplified gene was cloned into the pTrcHisB vector at the SacI/HindIII restriction sites, thereby adding an N-terminal His_6 tag. The natural cysteine at position 104 was changed to alanine, and the N-terminal cysteine was shifted two residues away from the N terminus using site-directed mutagenesis with PfuUltra II Fusion HS DNA polymerase (Agilent Technologies) to give the following protein construct, $GlpG_{Cys\text{-}TM\text{-}Cys}$, which includes residues 89–276 of *E. coli* GlpG between the two cysteines.

MGHHHHHHELAACLRERAGPVTWVMMIAAVVVFIAMQILGDQE
VMLWLAWPFDPTLKFEFWRYFTHALMHFSLMHILFNLLWWWYLGG
AVEKRLGSGKLIVITLISALLSGYVQQKFSGPWFGGLSGVVYALMGY
VWLRGERDPQSGIYLQRGLIIFALIWIVAGWFDLFGMSMANGAH
IAGLAVGLAMAFVD SLNARKRKC.

$GlpG_{Cys\text{-}TM\text{-}Cys/L155A}$ and $GlpG_{Cys\text{-}TM\text{-}Cys/A206G}$ were created by site-directed mutagenesis of Leu155 to alanine (residue position 80 in our construct) and Ala206 to glycine (residue position 131 in our construct).

The GlpG protein constructs in BL21-Gold (DE3) were grown in LB medium at 37 °C and induced with 0.5 mM IPTG at 0.7 OD_{600}, and cells were harvested after an additional 3 h incubation. Cells were resuspended in 50 mM Tris-HCl, 1 mM EDTA, 1 mM PMSF, 1 mM DTT and 2 μg/ml DNAse I (pH 8.0) and lysed by two passes through an Avestin EmulsiFlex-C3 at 15,000 psi. Cell debris was removed by centrifugation at 25,000*g* for 15 min. The membrane fraction was collected by centrifugation of the supernatant at 100,000*g* for 90 min at 4 °C. The pellet was resuspended in 25 mM Tris-HCl, 1.25% *n*-decyl-β-D-maltopyranoside (DM, Affymetrix) and 1 mM TCEP (pH 8.0) (5 ml per liter of culture) with the aid of a dounce homogenizer. Membranes were further solubilized with gentle rotation for 45 min at room temperature. The soluble fraction was collected after centrifugation at 100,000*g* for 30 min at 4 °C. 4 M NaCl and 5 M imidazole were added to the supernatant to a final concentration of 300 mM and 10 mM, respectively. The supernatant was incubated for 1 h at 4 °C with Ni-NTA (0.5 ml resin per liter of culture) that had been preequilibrated with 20 mM Tris-HCl, 300 mM NaCl, 10 mM imidazole, 0.2% DM and 1 mM TCEP (pH 8.0). The resin was packed into a column by gravity and, after collecting the flow-through, washed with 5 column volumes of 10 mM and 30 mM imidazole before eluting with 300 mM imidazole in 20 mM Tris-HCl, 300 mM NaCl, 0.2% DM and 1 mM TCEP (pH 8.0). Elution fractions containing protein (detected by absorbance at 280 nm) were pooled, concentrated to 3 ml using a 10,000 MWCO Amicon Ultra centrifugal filter (Millipore) and buffer exchanged into 25 mM Tris-HCl, 0.2% DM, 1 mM TCEP (pH 8.0) using an Econo-Pac 10DG column (BioRad). The protein was then passed over a 1-ml HiTrap Q HP ion exchange column (GE Healthcare Life Sciences) equilibrated with 25 mM Tris-HCl, 0.2% DM, 1 mM TCEP (pH 8.0). The flow-through was collected and bound to Ni-NTA (0.5 ml resin per liter of culture) equilibrated with 20 mM Tris-HCl, 150 mM NaCl, 0.2% DM and 1 mM TCEP (pH 7.5). The resin was packed into a column by gravity and washed with 0.2% DM before washing with 10 column volumes of 0.5% *n*-dodecyl-β-D-maltopyranoside (DDM) to exchange GlpG into DDM micelles. The resin was washed with 0.1% DDM to return to a low concentration of detergent and eluted with 300 mM imidazole. Protein-containing fractions of 1 ml volume were pooled and exchanged into 50 mM Tris-HCl, 150 mM NaCl and 0.1% DDM (pH 7.5) with an Econo-Pac 10DG desalting column to remove TCEP and imidazole. Aliquots of purified $GlpG_{Cys\text{-}TM\text{-}Cys}$ and $GlpG_{Cys\text{-}TM\text{-}Cys/L155A}$ were flash frozen in liquid nitrogen and stored at −80 °C. Fresh aliquots were used for activity assays and 2,2′-dithiodipyridine (DTDP, Sigma-Aldrich) derivatization.

The GlpG substrate, SN-Spitz, was a modified version of SNGpATM, which contains staphylococcal nuclease fused to the transmembrane segment of glycophorin A transmembrane domain and a C-terminal His tag[4]. To convert SNGpATM into a GlpG substrate, the transmembrane segment was modified by Quickchange mutagenesis to include the sequence of Spitz (ASIASGA), which is a known cleavage site for *E. coli* GlpG[49].

MATSTKKLHKEPATLIKAIDGDTVKLMYKGQPMTFRLLLVDT
PETKHPKKGVEKYGPEASAFTKKMVENAKKIEVEFDKGQRTDKYG
RGLAYIYADGKMVNEALVRQGLAKVAYVYKPNNTHEQHLRKSE AQAK
KEKLNIWSEDNADSGPERVQLAHHFSEPGASIASGAVMAGVIGTI
LLISYGIRRLIKKLEHHHHHH.

SN-Spitz was expressed and purified in DDM as previously described for SN-GpA[4].

GlpG activity in detergent and bicelles. Bicelles for activity assays were composed of 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) and 3-((3-cholamidopropyl)dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO, Avanti Polar Lipids) and prepared as described[12,50,51]. To prepare a stock solution of $GlpG_{Cys\text{-}TM\text{-}Cys}$ in bicelles, 53 μM $GlpG_{Cys\text{-}TM\text{-}Cys}$ in 50 mM Tris-HCl, 150 mM NaCl, 0.1% DDM (pH 7.5) was mixed 16.5:1 with 35% DMPC/CHAPSO (2.8:1, w/w) to give a final concentration of 50 μM $GlpG_{Cys\text{-}TM\text{-}Cys}$ in 2% DMPC/CHAPSO. The mixture was then incubated on ice for 30 min followed by 2 h at room temperature before use. We also prepared a stock solution of $GlpG_{Cys\text{-}TM\text{-}Cys}$ in detergent which contained 50 μM $GlpG_{Cys\text{-}TM\text{-}Cys}$ in 50 mM Tris-HCl, 150 mM NaCl, 0.1% DDM (pH 7.5). A stock solution of SN-Spitz was prepared containing 200 μM SN-Spitz in 50 mM Tris-HCl, 150 mM NaCl and 0.1% DDM (pH 7.5). Reactions were initiated by adding 1.6 μl of the $GlpG_{Cys\text{-}TM\text{-}Cys}$ stock, 1 μl of the SN-Spitz stock,17.4 μl of 50 mM Tris-HCl and150 mM NaCl and then were incubated at 37 °C for 18 h. Reactions were stopped by adding 10 μl of 4× SDS sample buffer and heating for 10 min at 65 °C. The cleaved product was visualized by SDS-PAGE using a 4–12% NuPAGE BisTris gradient gel (Life Technologies) run in MES SDS running buffer (**Supplementary Fig. 1a**).

DNA handle attachment to bicelle-incorporated protein. The bicelle stock was made with DMPC lipid (Avanti Polar Lipids) and CHAPSO detergent (Sigma-Aldrich or Affymetrix) in the deionized water[12,51]. The molar ratio of lipid to detergent ranged from 2.2:1 to 2.8:1 with a final bicelle concentration of 8.8%. To dissolve the lipid and detergent, cycles of cooling on ice, brief vortexing, freezing at −80 °C, brief heating to 33 °C and vortexing were performed. Finally a quick spin at 4 °C with a tabletop centrifuge helped remove any remaining powders. The bicelle stock solution was stored at −80 °C.

Purified $GlpG_{Cys\text{-}TM\text{-}Cys}$ was derivatized with DTDP by mixing 250 μl of 27 μM $GlpG_{Cys\text{-}TM\text{-}Cys}$ in 50 mM Tris-HCl, 150 mM NaCl, 0.1% DDM (pH 7.5) with 20 μl of 67 mM DTDP dissolved in acetonitrile for a final concentration of 25 μM GlpG and 5 mM DTDP. The reaction was incubated on a rotator for 1 h at room temperature. Unreacted DTDP was removed using a BioRad Econo-Pac 10DG desalting column equilibrated with 50 mM Tris-HCl, 150 mM NaCl, 0.1% DDM (pH 7.5). Protein was collected in 250-μl fractions. The peak fractions were pooled and concentrated to 200 μl of 26 μM DTDP-derivatized GlpG. Complete labeling of GlpG was verified by LC/MS as described using an orifice potential of 90 V (ref. 52). We added the bicelle stock solution to the DTDP-derivatized GlpG in a 1:4 (v/v) ratio while keeping it on ice and gently pipetted the contents up and down until the solution became clear and homogenous[12,51]. We incubated the mixture on ice for 1 h to allow for complete reconstitution of GlpG into bicelles and kept the bicelle-GlpG mixture on ice until it was used in the next step.

Two types of 512-bp DNA (biotin- and digoxigenin-modified handles) were PCR-amplified using a λDNA template, the forward primer CATGTGGGTGACG CGAAA with a 5′ thiol–modified C6 S-S and the reverse primer TCGCCACCATCATTTCCA with either 5′ biotin or 5′ digoxigenin modification (each 0.4 ml and 2 ml). The thiol modifications of the PCR products were activated by adding 100 mM DTT final concentration and incubating for 1 h at 37 °C. The products were purified using a PCR purification column, eluted into 50 mM Tris-HCl (pH 7.5) and concentrated to 3–10 μM in a final volume of ~30 μl using a 10K Amicon centrifugal filter (Millipore). The DNA handles were stored at −80 °C.

To maximize the likelihood of two different handles attaching to a GlpG protein, the handles were attached sequentially, maintaining the bicelle

concentration as 0.5–2% during the whole attachment reaction. First, about 20-fold excess of the protein (14 μM) was reacted with the biotin-modified DNA handle (0.8 μM) in 40 mM Tris-HCl, 80 mM NaCl, 1.3% bicelle (pH 7.5) for >12 h at room temperature. Repetitive buffer exchange into 50 mM Tris-HCl, 100 mM NaCl, 2% bicelle (pH 7.5) were then performed using a 100K Amicon centrifugal filter. Then the GlpG attached to the biotin-modified handle (0.2 μM) was reacted with about 40-fold excess of digoxigenin-modified DNA handle (7.5 μM) in 50 mM Tris-HCl, 150 mM NaCl and 0.5–1% bicelle (pH 7.5) for >20 h at room temperature. The bicelle-incorporated GlpG covalently linked to the respective DNA handles (bicelle-GlpG-DNA) was diluted tenfold with 50 mM Tris-HCl, 150 mM NaCl and 1.3% bicelle (pH 7.5) and stored at −80 °C. The bicelle-GlpG-DNA sample was analyzed by 6% SDS-PAGE stained with SYBR Safe DNA Gel Stain (Invitrogen) (**Supplementary Fig. 1b**).

Single-molecule magnetic tweezers experiment. A single-molecule magnetic tweezers apparatus was built on an inverted microscope (Olympus, IX73) as previously described[27–33], in which force can be easily controlled by changing the vertical distance of a pair of magnets from the sample. The imaging room of the magnetic tweezers was maintained at constant humidity and constant temperature (22 °C) to prevent an undesirable bicelle phase transition at higher temperature. The sample chamber was a ~20 μl volume channel, constructed by putting together a 24 × 40 mm cleaned coverslip and 24 × 50 mm polyethylene glycol-coated coverslip with double-sided tape. The bicelle-GlpG-DNA sample was injected into the sample channel and then attached to the bottom coverslip by biotin-neutravidin binding and to 2.8-μm magnetic beads by dig-antidig binding. The buffer condition in the sample channel was 50 mM Tris-HCl, 150 mM NaCl, 1.3% bicelle (pH 7.5). We can exchange various buffer solutions by capillary force into the channel. By approaching the pair of magnets to the experiment sample, we can apply a few to tens of pN force to the single GlpG protein and then measure the change of extension, i.e., the end-to-end distance in the bicelle-GlpG-DNA molecule (**Fig. 1a**). The extension change is obtained from the change of diffraction patterns of attached magnetic bead captured in 60 Hz CCD camera (JAI). We corrected vertical drifts of microscope stage by maintaining the vertical position of a nonmagnetic reference bead immobilized directly on the bottom surface every 500 ms.

In the gradual pulling experiments (**Fig. 1** and **Fig. 3a**), the force-loading rate at every moment is far below 1 pN s^{-1}, which is near equilibrium condition during protein unfolding and folding. In the force-jump and force-cycle experiments (**Fig. 2** and **Fig. 3c**), the unfolding step sizes were measured as the difference between arithmetic mean values of extensions over the appropriate intervals before and after the unfolding event and then statistically analyzed as Gaussian distributions by collecting them (further analysis is described below). We can assess the relevant error in the step-size measurement (σ_{step}) with the equation $\sigma_{step} = \sqrt{\sigma_i^2 / N_i + \sigma_f^2 / N_f}$, which illustrates that σ_{step} is a s.e.m. because the s.d. of extension trace (σ_i, σ_f) is divided by the number of data points (N_i, N_f). Because the fluctuation of extension traces is typically less than 5 nm and we include more than 300 data points for each measurement, σ_{step} is less than ~0.4 nm, indicating that we can estimate the step size with an accuracy down to the level of a few Å[27].

Extension analysis for finding intermediate positions. To map the extension values measured in the force-jump experiment to the corresponding residue positions (**Fig. 2**), we analyzed the expected extension when a GlpG protein unfolds from the native state to specific residues[27] (**Supplementary Fig. 9**). The total extension is described as the sum of three terms:

$$x = \sum_{n=1}^{N} x_{p,n} + \sum_{n=1}^{M} x_{h,n} + \Delta d \qquad (1)$$

where $x_{p,n}$ is the extension expected for the n^{th} helix or linker that has lost its secondary structure (i.e., an unstructured polypeptide) as calculated by the Marko-Siggia formula; $x_{h,n}$ is the extension for the n^{th} helix as calculated by Kessler-Rabin formula[53]; and Δd is the axial length change of the tertiary structure between DNA handles calculated from the GlpG structure information[15–18].

The extension for unfolded polypeptide (x_p) is obtained using the Marko-Siggia formula of the worm-like chain (WLC) model:

$$F = \frac{k_B T}{P_p}\left[\frac{1}{4(1 - x_p / L_p)^2} + \frac{x_p}{L_p} - \frac{1}{4}\right] \qquad (2)$$

where F is the applied tension; $k_B T$ is the thermal energy; L_p is the contour length of polypeptide, which is the number of unfolded residues (N_p) times the average residue step size (l_p) of 0.36 nm (ref. 6); and P_p is the persistence length of polypeptide (measured as 0.39 nm; described in the next section). Equation (2) can be applied when the contour length of a polymer is much greater (by at least a factor of five) than its persistence length. Therefore, α-helices are not well described by equation (2) because the persistence length of the helices (tens of nm) is greater than the contour lengths of each helix (a few nm). Thus the extension for the helical part (x_h) is estimated by the Kessler-Rabin formula (KR model), which is applicable for any arbitrary ratio between persistence length and contour length:

$$x_h = -\frac{1}{2f} - \frac{\chi}{f \tanh 2\chi} + \frac{L_h}{\tanh fL_h} - \frac{2\chi^2}{3f}\left(\frac{1}{\tanh fL_h} - \frac{fL_h}{\sinh fL_h} - 1\right) \qquad (3)$$

where $f = F/k_B T$, $\chi = \sqrt{fl_h^2 / 4P_h}$, L_h is the contour length of helix that is the number of residues in the helix (N_h) times the average helical rise per amino acid (l_h) of 0.16 nm, and P_h is the persistence length of helix (measured as 9.17 nm; described in next section). Finally, from the GlpG structure information, the axial length change of Δd can be obtained as $\Delta d = d - d_0$, where d is the axial width of the partially folded structure up to specific residue and d_0 is the axial width of the fully folded structure (**Supplementary Fig. 9a**).

In analyzing the Gaussian peaks of the extension distribution measured in the force-jump experiment (**Fig. 2**), equation (1) can be reduced to

$$x = x_p + \Delta d \qquad (4)$$

because we observe that a helix-coil transition occurs at about 18 pN; thus, all of the helices can be assumed to be unraveled upon unfolding at the 21-pN force used. Thus, with equations (2) and (4), we calculated the expected extension value for a GlpG protein unfolding up to a specific residue position (**Supplementary Fig. 9b**). In this calculation, we compared two versions of extension estimation from two different GlpG structures: one in detergent condition[15] and the other in lipid bilayer condition[18]. The difference in protein structures is reflected in the structural factor Δd in equations (1) and (4), but we did not see any obvious difference between the two estimations (**Supplementary Fig. 9b**).

Helix-coil transition. By pooling the unfolding data (unfolding force F_u and step size ΔL_u) from all traces in the gradual pulling experiments (**Supplementary Fig. 3a**), we produced a scatter plot showing the unfolding force against the step size (**Supplementary Fig. 3b**). When the six transmembrane helices are completely unraveled to polypeptide coils upon unfolding, the data points are expected to be distributed along the line of equation (2) (WLC model; **Supplementary Fig. 3b**). We observed a definite deviation from the WLC model below 20 pN, which indicates that the helix-coil transition in the corresponding force range (**Supplementary Fig. 3b**,**c**). This is further supported by the observation that in the force-jump experiment (**Fig. 2a** and **Supplementary Fig. 6**), the observed step size of ~35 nm corresponds to unfolding of GlpG to the completely unstructured state with no α-helical content (**Supplementary Fig. 9b**).

To obtain the persistence length of the polypeptide (P_p), we fitted the data only for the region over 20 pN with the WLC model,

$$F = \frac{k_B T}{P_p}\left[\frac{1}{4(1 - (x + d_0) / L_p)^2} + \frac{(x + d_0)}{L_p} - \frac{1}{4}\right] \qquad (5)$$

which is derived from equations (2) and (4) with $d = 0$ because there is no tertiary structure (**Supplementary Fig. 3b**). The average residue step size (l_p) of 0.36 nm was used[6]. The persistence length determined for the GlpG polypeptide (P_p = 0.39 nm) is consistent with what was reported for a similar helical membrane protein (P_p = 0.4 nm) (ref. 6).

For the persistence length of helix (P_h), we fitted the data for the region below 17 pN with the WLC-KR model

$$x = x_p - d_0 + 6\times\left[-\frac{1}{2f} - \frac{\chi}{f\tanh 2\chi} + \frac{L_h}{\tanh fL_h} - \frac{2\chi^2}{3f}\left(\frac{1}{\tanh fL_h} - \frac{fL_h}{\sinh^2 fL_h} - 1\right)\right] \quad (6)$$

(**Supplementary Fig. 3b**). Equation (6) is derived from the simplified equation (1), $x \approx x_p + 6{\cdot}x_h - d_0$, and equation (3). The extension for loop regions between helices (x_p) is calculated by equation (2). The estimated persistence length for helices of GlpG (P_h = 9.17 nm) is broadly consistent with the known value for a helix in a coiled coil (P_h = 25 nm) (ref. 54). The fact that the persistence length for the coiled coil is somewhat larger seems reasonable because of the tight association of two helices in the coiled coil.

Quantitative analysis of the folding energy landscape. To obtain a quantitative picture for the folding energy landscape, we measured unfolding and folding kinetics. We used the gradual pulling experiment to obtain the unfolded fraction (U) as a function of force (F) (**Figs. 3a**,**b** and **4a**), from which we determined the kinetic rate constant for unfolding at zero force (k_{u0}) and the distance from the native state to the transition state ($\Delta x_f^\dagger$) (**Table 1**). To this end, we used the following equation,

$$U = 1 - \exp\left(-\frac{k_{u0}k_BT}{A\Delta x_f^\dagger}F^{-1}e^{F\Delta x_f^\dagger/k_BT}\right) \quad (7)$$

where k_BT is the thermal energy and A is the proportional constant of $dF/dt = AF$. The constant A is determined from the data of force calibration with magnet heights, which is approximated as a single exponential function in the analyzed force range. Equation (7) is derived from the first-order rate equation, $dN/dt = -k_uN$, and the Bell equation, $k_u = k_{u0}\exp(-F\Delta x_f^\dagger/k_BT)$, where N represents the folded fraction and k_u represents the unfolding kinetic rate at a given force. Equation (7) can be derived from

$$1 - U = N = \exp\left(-\int dtk_{u0}e^{F\Delta x_f^\dagger/k_BT}\right) = \exp\left(-\int dF\frac{k_{u0}}{\dot{F}}e^{F\Delta x_f^\dagger/k_BT}\right)$$

$$= \exp\left(-\int dF\frac{k_{u0}}{AF}e^{F\Delta x_f^\dagger/k_BT}\right) = \exp\left(-\frac{k_{u0}k_BT}{A\Delta x_f^\dagger}F^{-1}e^{F\Delta x_f^\dagger/k_BT}\right)$$

For folding kinetics, we performed refolding experiments at lower forces ranging from 0 pN to 8 pN (**Figs. 3c**,**d** and **4b**). At these forces, the thermal noise is too high to detect the individual refolding events. Hence after unfolding at 21 pN, we lowered the force to specified levels, waited for 3 min and increased the force to 21 pN to see whether GlpG was refolded during the 3-min waiting time (**Fig. 3c**). From the folding probability (N) as a function of force (**Figs. 3d** and **4b**), we measured the kinetic rate for folding at zero force (k_{f0}) and the distance from the unfolded state to the transition state ($\Delta x_u^\dagger$; **Table 1**). The fitting equation

$$N = 1 - \exp\left(-\Delta tk_{u0}e^{-F\Delta x_u^\dagger/k_BT}\right) \quad (8)$$

in which Δt is the waiting time at specific force for refolding, is likewise derived from the first-order rate equation $dU/dt = -k_fU$, and the Bell equation, $k_f = k_{f0}\exp(-F\Delta x_u^\dagger/k_BT)$, where k_f is the folding kinetic rate at arbitrary force. The formula derivation is developed as

$$1 - N = U = \exp\left(-\int dtk_{f0}e^{-F\Delta x_u^\dagger/k_BT}\right) = \exp\left(-\Delta tk_{f0}e^{-F\Delta x_u^\dagger/k_BT}\right)$$

at constant force.

From the kinetic rate constants (k_{u0}, k_{f0}), we obtained the unfolding free energy (ΔG) and the kinetic energy barriers ($\Delta G_u^\dagger$, $\Delta G_f^\dagger$) (**Fig. 3e** and **Table 1**) by the Kramer equation

$$\Delta G_u^\dagger = -k_BT\ln(k_{u0}/k_w) \quad (9)$$

$$\Delta G_f^\dagger = -k_BT\ln(k_{f0}/k_w) \quad (10)$$

$$\Delta G = \Delta G_u^\dagger - \Delta G_f^\dagger = -k_BT\ln(k_{u0}/k_{f0}) \quad (11)$$

where k_w is the frequency factor in the range of 10^4–10^6 s^{-1} (refs. 55–60), which is why the energy barriers are measured with an error of 2.3 k_BT. The unfolding free energy (ΔG) indicating the thermodynamic stability of protein is more reliably measured with an error of 0.2 k_BT because it is obtained only from the ratio of k_{u0} to k_{f0}, regardless of the frequency factor.

49. Urban, S. & Wolfe, M.S. Reconstitution of intramembrane proteolysis in vitro reveals that pure rhomboid is sufficient for catalysis and specificity. *Proc. Natl. Acad. Sci. USA* **102**, 1883–1888 (2005).
50. Faham, S., Ujwal, R., Abramson, J. & Bowie, J.U. in *Membrane Protein Crystallization* 1st edn. Vol. 63 (eds. DeLucas, L.J.) 109–125 (Academic Press, 2009).
51. Ujwal, R. & Bowie, J.U. Crystallizing membrane proteins using lipidic bicelles. *Methods* **55**, 337–341 (2011).
52. Whitelegge, J.P. *et al.* Toward the bilayer proteome, electrospray ionization-mass spectrometry of large, intact transmembrane proteins. *Proc. Natl. Acad. Sci. USA* **96**, 10695–10698 (1999).
53. Kessler, D.A. & Rabin, Y. Distribution functions for filaments under tension. *J. Chem. Phys.* **121**, 1155–1164 (2004).
54. Schwaiger, I., Sattler, C., Hostetter, D.R. & Rief, M. The myosin coiled-coil is a truly elastic protein structure. *Nat. Mater.* **1**, 232–235 (2002).
55. Schuler, B., Lipman, E.A. & Eaton, W.A. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature* **419**, 743–747 (2002).
56. Yang, W.Y. & Gruebele, M. Folding at the speed limit. *Nature* **423**, 193–197 (2003).
57. Rhoades, E., Cohen, M., Schuler, B. & Haran, G. Two-state folding observed in individual protein molecules. *J. Am. Chem. Soc.* **126**, 14686–14687 (2004).
58. Kubelka, J., Hofrichter, J. & Eaton, W.A. The protein folding 'speed limit'. *Curr. Opin. Struct. Biol.* **14**, 76–88 (2004).
59. Chung, H.S., Louis, J.M. & Eaton, W.A. Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories. *Proc. Natl. Acad. Sci. USA* **106**, 11837–11844 (2009).
60. Gebhardt, J.C.M., Bornschlogla, T. & Rief, M. Full distance-resolved folding energy landscape of one single protein molecule. *Proc. Natl. Acad. Sci. USA* **107**, 2013–2018 (2010).

Supplementary Information

Mapping the energy landscape for second-stage folding of a single membrane protein

Duyoung Min[1,2,4,5], Robert E. Jefferson[3,5], James U. Bowie[3*] and Tae-Young Yoon[1,2*]

[1] National Creative Research Initiative Center for Single-Molecule Systems Biology, KAIST, Daejeon, South Korea

[2] Department of Physics, KAIST, Daejeon, South Korea

[3] Department of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, California, USA

[4] Present address: Department of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, California, USA

[5] These authors contributed equally to this work.

* e-mail: bowie@mbi.ucla.edu or tyyoon@kaist.ac.kr

Supplementary Results

Supplementary Figure 1 | Characterization of GlpG activity and stability.

Supplementary Figure 2 | Comparison of forced unfolding of a single membrane protein in AFM and magnetic tweezers.

Supplementary Figure 3 | Helix-coil transition of GlpG.

Supplementary Figure 4 | Schematic diagram of the single-molecule magnetic tweezers experiment with microfluidic buffer exchanges.

Supplementary Figure 5 | Unfolding forces of GlpG in various reconstitution conditions.

Supplementary Figure 6 | Unfolding step sizes at 21 pN

Supplementary Figure 7 | Argument for unidirectional unfolding of GlpG.

Supplementary Figure 8 | Comparison of the proportion of four different unfolding patterns between wild-type and mutant GlpG.

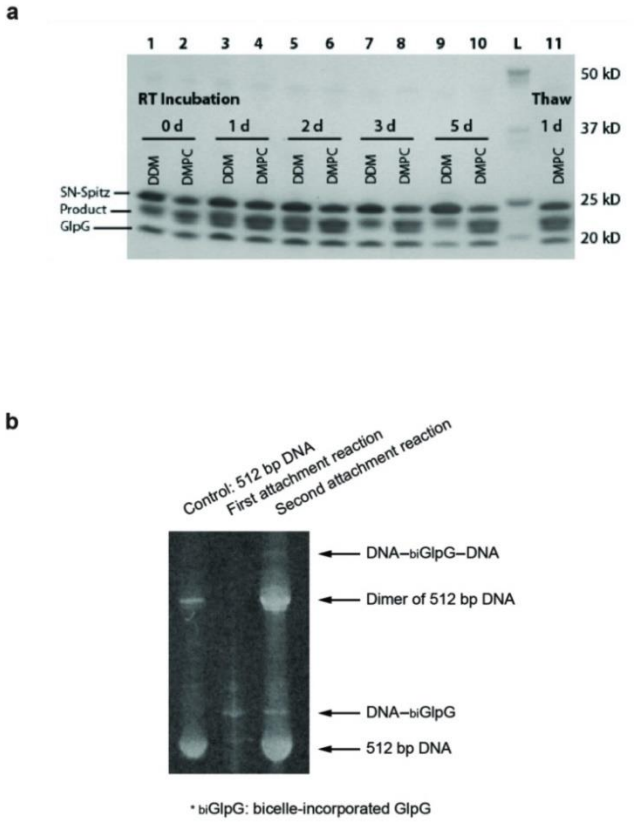
Supplementary Figure 9 | Extension analysis.

Supplementary Figure 10 | Unfolding kinetics of GlpG with different bicelle and temperature conditions.

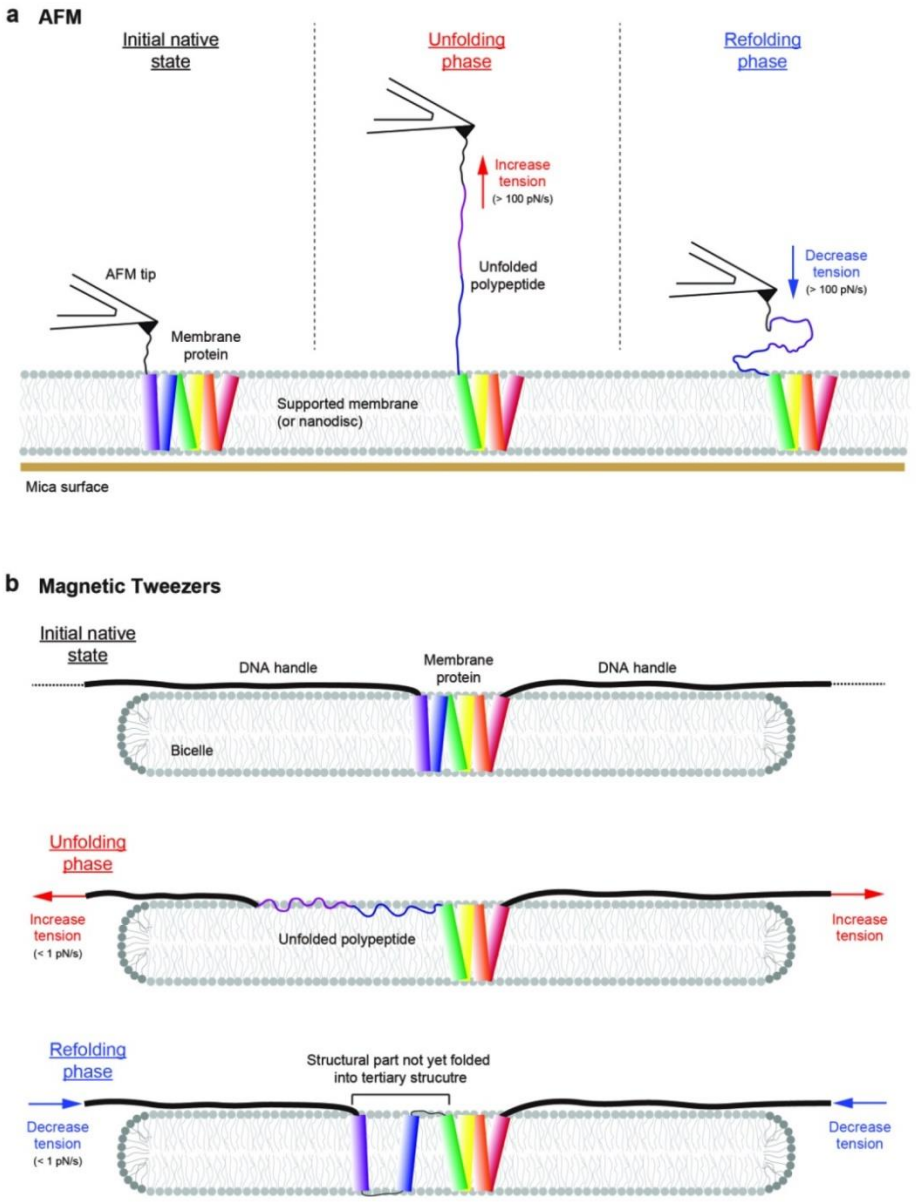
Supplementary Table 1 | Unfolding kinetics of GlpG with different lipid to detergent ratios.

Supplementary Table 2 | Unfolding kinetics of GlpG at different temperatures.

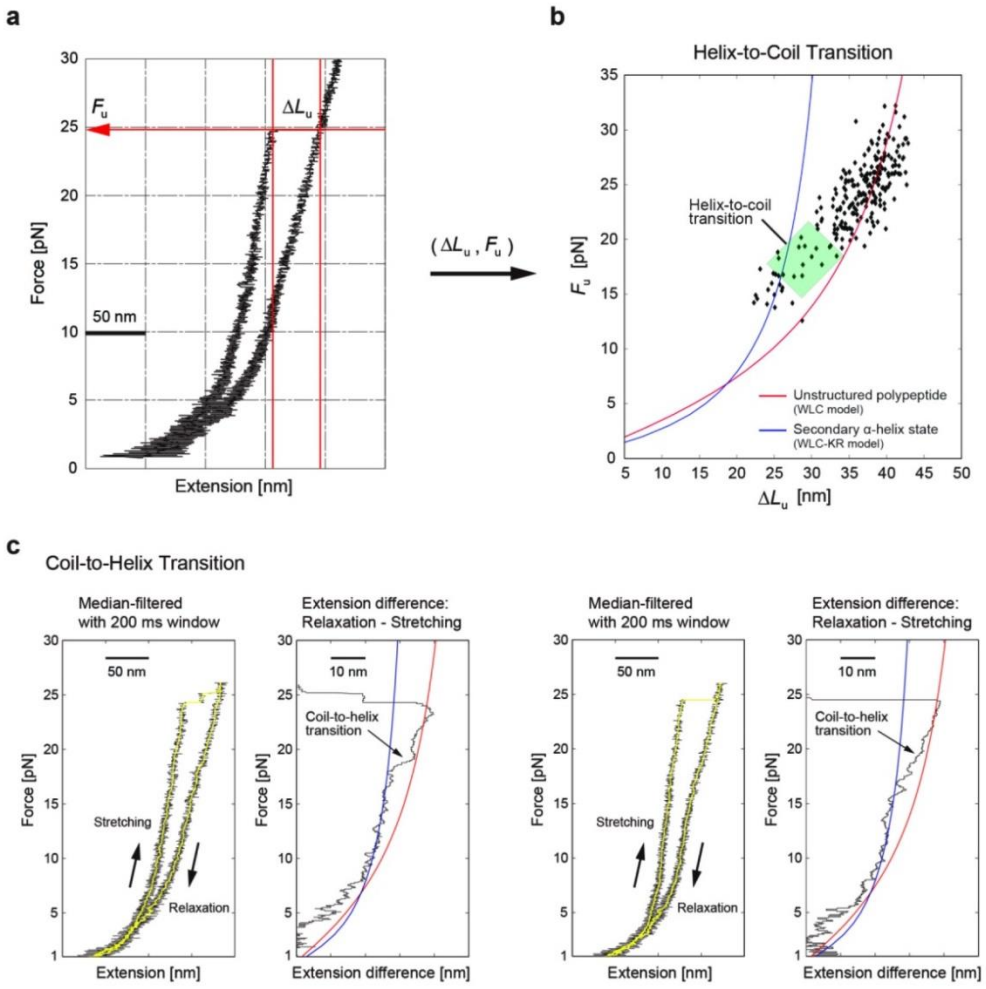
Supplementary Table 3 | Comparison of thermodynamic values of GlpG between the prior bulk unfolding measurements and our single-molecule forced unfolding measurement.



Supplementary Figure 1 | Characterization of GlpG activity and stability. (a) Activity assays of wild-type GlpG incubated in either 0.1% DDM or 2% DMPC:CHAPSO bicelles. Activity was measured after 0, 1, 2, 3, and 5 days of incubation at room temperature (lanes 1-10). The activity is lost after several days of incubation in DDM, but remains fully active after 5 days in bicelles. Thus, the protein is quite stable to inactivation in bicelle environment. Activity of bicelle-incorporated wild-type GlpG was also measured 1 day after flash-freezing in liquid nitrogen and thawing to room temperature (lane 11). The protein remains fully active after freezing. The molecular weight ladder (BioRad Precision Plus Protein Standard) shows bands at 20, 25, 37, and 50 kD (lane L). **(b)** 6% SDS-PAGE for analyzing the crosslinking chemistry of DNA handle attachment to bicelle-reconstituted GlpG protein. To visualize the DNA, this gel was stained with DNA-staining dye (see Online Methods for more details).

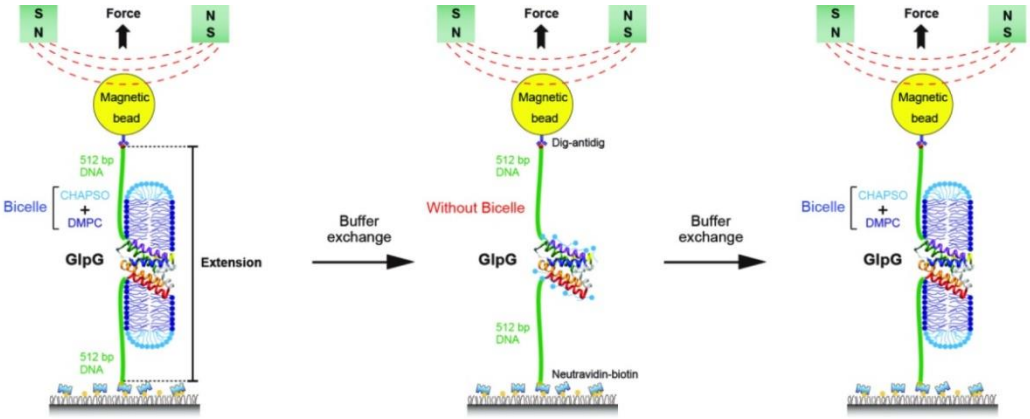


Supplementary Figure 2 | Comparison of forced unfolding of a single membrane protein in AFM and magnetic tweezers. (a,b) Schematic diagrams showing single-molecule unfolding and refolding processes in previous AFM studies (a) and our magnetic tweezers experiment (b). The main differences are the direction of mechanical tension relative to the lipid membrane and its application rate (i.e., force-loading rate).

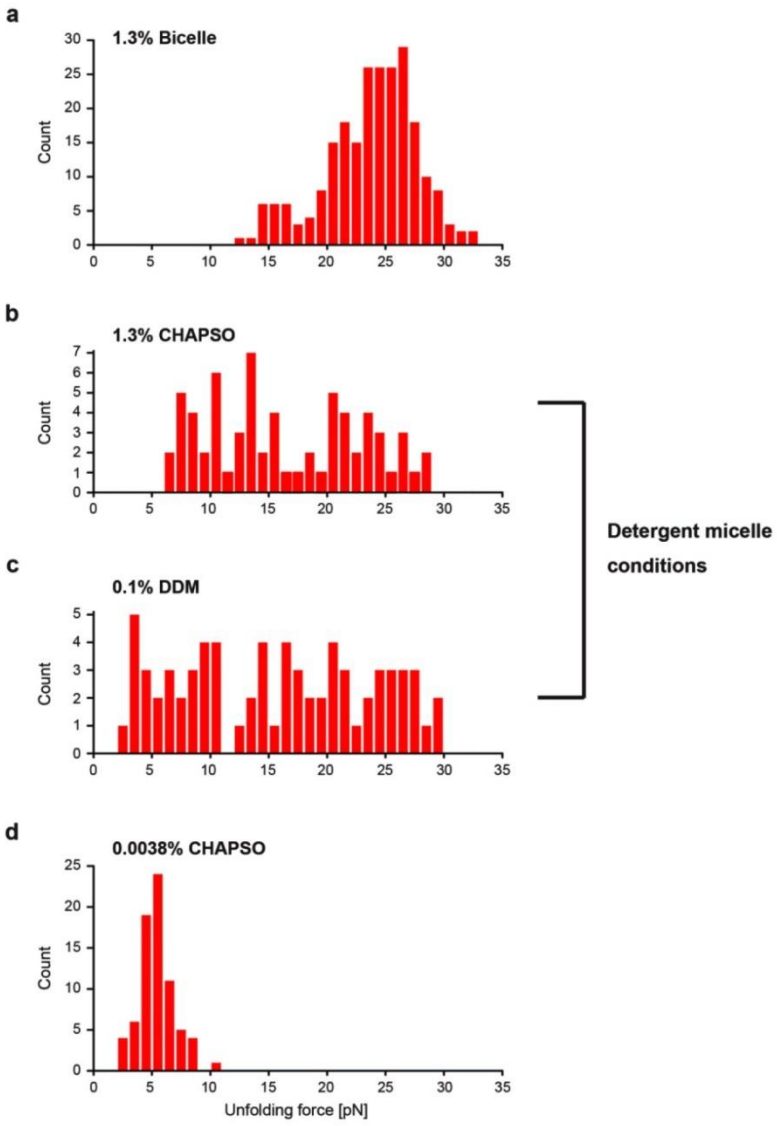


Supplementary Figure 3 | Helix-coil transition of GlpG. (a) Representative gradual pulling experiment showing how the increased extension (ΔL_u) and the unfolding force (F_u) are measured from force-extension traces. **(b)** Scatter plot of the unfolding data in the plane of ΔL_u and F_u (n=233). The fitting lines show the step size expected when GlpG is unfolded to unstructured polypeptide (red line, based on the WLC model) or when GlpG is unfolded with alpha-helical structures retained (blue line, based on the WLC-KR model) (see Online Methods). The distribution of the individual unfolding events between the two fitting lines

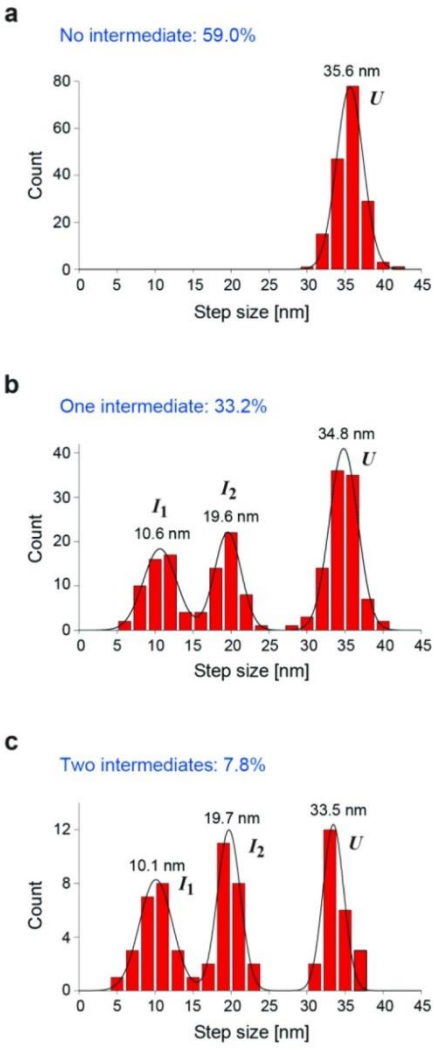
show the helix-to-coil transition at around 18 pN. **(c)** Representative single-molecule traces showing the coil-to-helix transition during the relaxation phase. To visualize the transitions more clearly, the traces were median-filtered with a 200 ms window (left panels) and the extension difference between the relaxation and stretching phases were plotted (right panels). The relatively steep decreases in the extension difference at ~18 pN indicate the coil-to-helix transitions.



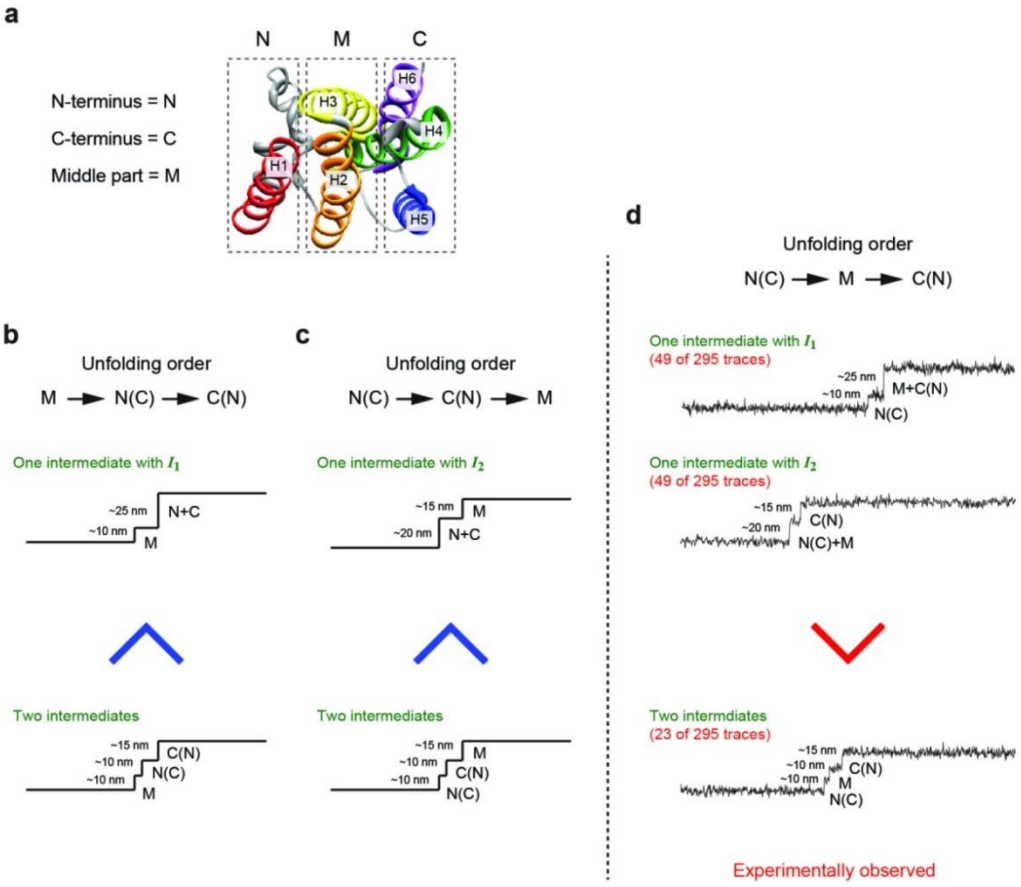
Supplementary Figure 4 | Schematic diagram of the single-molecule magnetic tweezers experiment with microfluidic buffer exchanges. After several cycles of GlpG unfolding and refolding in bicelles (left), the bicelles were removed and the unfolding and refolding cycles were repeated (middle). After up to tens of pulling cycles, the bicelle condition was restored by another round of microfluidic buffer exchange (right). Representative force-extension curves in each buffer condition are shown in Fig. 1b.



Supplementary Figure 5 | Unfolding forces of GlpG in various reconstitution conditions. **(a-d)** Distributions of the unfolding forces when GlpG proteins are reconstituted in 1.3% bicelle (*n*=233) (a), 1.3% CHAPSO micelle (*n*=66) (b), 0.1% DDM micelle (*n*=71) (c) or a minimal amount of detergents (0.0038% CHAPSO) (*n*=74) (d).

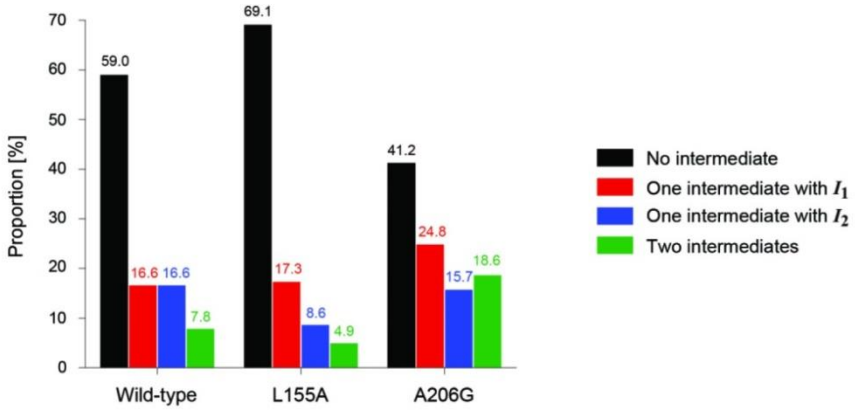


Supplementary Figure 6 | Unfolding step sizes at 21 pN. (a-c) Step size distributions of unfolding events with no intermediates (59.0%) (a), one intermediate (33.2%) (b), and two intermediates (7.8%) (c). Corresponding extension traces are shown in Fig. 2a. The total number of observations is n=295.

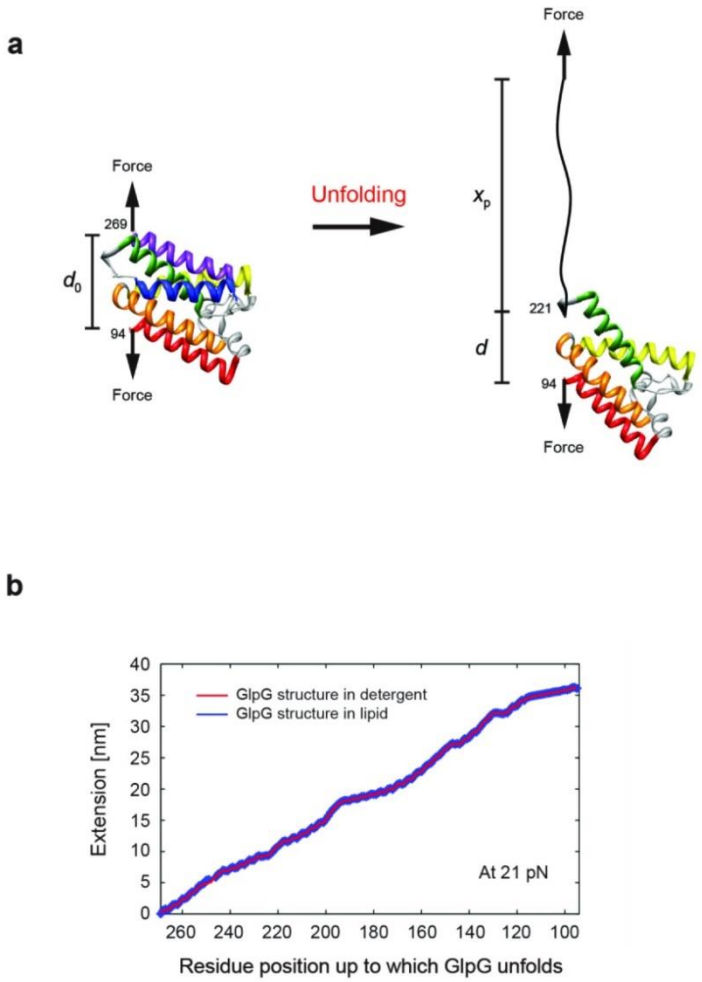


Supplementary Figure 7 | Argument for unidirectional unfolding of GlpG. (a) GlpG structure from a cytoplasmic view showing three distinct domains. Each structural domain is presumed to correspond to one of the three steps observed in the unfolding events. The N-terminal, middle and C-terminal parts of GlpG are denoted by N, M and C, respectively. **(b-d)** Possible sequences of unfolding of the three domains and the consequential observation probability of specific unfolding patterns. For example, in the case where N-domain unfolding is directly followed by that of the C domain (or vice versa), one-intermediate unfolding events require the simultaneous unfolding of the N and C domains within our time resolution of 16 ms, which are structurally distal to each other and stochastic (b,c). Thus, one-intermediate unfolding should be less probable than two-intermediate unfolding (b,c). However, the observation probabilities of one-intermediate unfolding with I_1 (or I_2), turned out to be higher

than that of two-intermediate unfolding (d), thereby defying the scenarios described in (b) and (c). We are only left with the possibility that M-domain unfolding intercalates between unfolding of the N and C domains, which points to unidirectional unfolding of GlpG.

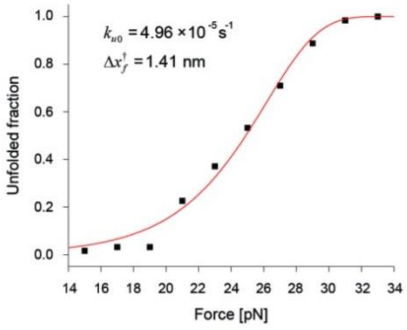


Supplementary Figure 8 | Comparison of the proportion of four different unfolding patterns between wild-type and mutant GlpG. The total number of observations is n=295 for wild-type, n=81 for L155A and n=97 for A206G.

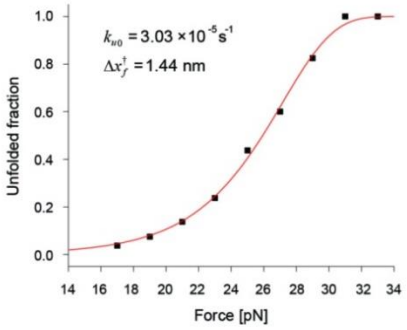


Supplementary Figure 9 | Extension analysis. (a) Cartoon describing an increased extension due to the unfolding of GlpG up to a specific residue. The expected extension increases were estimated as $x = x_p + \Delta d$, where x_p is the extension of unraveled polypeptide calculated with the worm-like chain model and Δd is the axial length change (d-d_0) due to the rotational motion of remaining tertiary part calculated from the known GlpG structure (see Online Methods). **(b)** Extension increase at 21 pN when GlpG is unfolded to the given residue position. The estimations are shown for detergent (red) and lipid (blue) conditions.

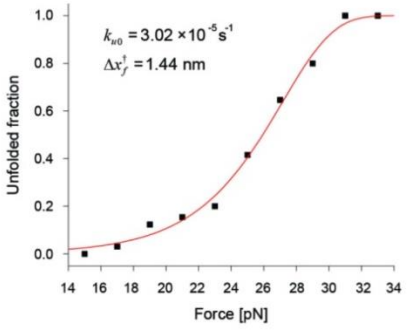
a Lipid : detergent = 2.5 : 1



b Lipid : detergent = 2.8 : 1



c Temperature = 25°C



Supplementary Figure 10 | Unfolding kinetics of GlpG with different bicelle and temperature conditions. (a,b) Unfolded fraction as a function of force when the lipid to detergent ratio is 2.5:1 (n=63) (a) and 2.8:1 (n=80) (b). **(c)** Unfolded fraction as a function of force when the temperature is 25 °C (n=65). Fitting the data (see Online Methods) yields a kinetic rate for unfolding at zero tension (k_{u0}) and a distance from the folded state to the transition state ($\Delta x_f^{\dagger}$). The measured kinetic data are summarized in Supplementary Tables 1 and 2.

Lipid : detergent ratio	$\Delta x_f^\dagger$ [nm]	k_{u0} [s^{-1}]	$\Delta G_u^\dagger$ [k_BT]
2.2 : 1	1.48 (0.03)	5.64 (0.91) × 10^{-5}	21.30 (2.30)
2.5 : 1	1.41 (0.11)	4.96 (2.88) × 10^{-5}	21.42 (2.30)
2.8 : 1	1.44 (0.06)	3.03 (0.99) × 10^{-5}	21.92 (2.30)

Supplementary Table 1 | Unfolding kinetics of GlpG with different lipid to detergent ratios. Numbers in parentheses indicate error. The errors of $\Delta x_f^\dagger$ and k_{u0} represent s.e.m. and the error of $\Delta G_u^\dagger$ represent the error of the frequency factor k_w (Online Methods).

Temperature [°C]	$\Delta x_f^\dagger$ [nm]	k_{u0} [s^{-1}]	$\Delta G_u^\dagger$ [k_BT]
22	1.48 (0.03)	5.64 (0.91) × 10^{-5}	21.30 (2.30)
25	1.44 (0.09)	3.02 (1.53) × 10^{-5}	21.92 (2.30)

Supplementary Table 2 | Unfolding kinetics of GlpG at different temperatures. Numbers in parentheses indicate error. The errors of $\Delta x_f^\dagger$ and k_{u0} represent s.e.m. and the error of $\Delta G_u^\dagger$ represent the error of the frequency factor k_w (Online Methods).

	Methods	ΔG [k_BT]	$\Delta\Delta G$ [k_BT]
Wild-type	Bulk SDS unfolding in detergent micelle[a]	7.08 (1.35)	-
	Bulk SDS unfolding in detergent micelle[b]	13.88 (2.41)	-
	Bulk SDS unfolding in detergent micelle[c]	12.46 (0.34)	-
	Single-molecule forced unfolding in bicelle	6.54 (0.21)	-
L155A	Bulk thermal unfolding in detergent micelle[d]	6.28 (1.60)	0.80 (2.09)
	Bulk SDS unfolding in detergent micelle[c]	10.72 (0.43)	1.74 (0.27)
	Single-molecule forced unfolding in bicelle	4.37 (0.48)	2.17 (0.52)
A206G	Bulk thermal unfolding in detergent micelle[d]	6.99 (1.64)	0.09 (2.12)
	Bulk SDS unfolding in detergent micelle[c]	10.69 (0.39)	1.77 (0.19)
	Single-molecule forced unfolding in bicelle	5.33 (0.38)	1.21 (0.43)

[a] Equilibrium data from ref. 1
[b] Equilibrium data from ref. 2
[c] Kinetic data from ref. 2
[d] Obtained from the fitting curve of T_m versus ΔG in Supplementary Dataset of ref. 1

Supplementary Table 3 | Comparison of thermodynamic values of GlpG between the prior bulk unfolding measurements and our single-molecule forced unfolding measurement. Numbers in parentheses indicate error as s.e.m.

References

1. Baker, R.P. & Urban, S. Architectural and thermodynamic principles underlying intramembrane protease function. *Nat. Chem. Biol.* **8**, 759-68 (2012).
2. Paslawski, W. et al. Cooperative folding of a polytopic alpha-helical membrane protein involves a compact N-terminal nucleus and nonnative loops. *Proc. Natl. Acad. Sci. U S A* **112**, 7978-7983 (2015).

Chapter 5

A simple DNA handle attachment method for single molecule mechanical manipulation experiments

METHODS AND APPLICATIONS

A simple DNA handle attachment method for single molecule mechanical manipulation experiments

Duyoung Min, Mark A. Arbing, Robert E. Jefferson, and James U. Bowie*

Department of Chemistry and Biochemistry, UCLA-DOE Institute, Molecular Biology Institute, Los Angeles, California

Received 20 April 2016; Accepted 23 May 2016
DOI: 10.1002/pro.2952
Published online 24 May 2016 proteinscience.org

Abstract: Manipulating single molecules and systems of molecules with mechanical force is a powerful technique to examine their physical properties. Applying force requires attachment of the target molecule to larger objects using some sort of molecular tether, such as a strand of DNA. DNA handle attachment often requires difficult manipulations of the target molecule, which can preclude attachment to unstable, hard to obtain, and/or large, complex targets. Here we describe a method for covalent DNA handle attachment to proteins that simply requires the addition of a preprepared reagent to the protein and a short incubation. The handle attachment method developed here provides a facile approach for studying the biomechanics of biological systems.

Keywords: forced unfolding; magnetic tweezers; optical tweezers; SpyTag; SpyCatcher; protein folding; membrane protein

Introduction

Single-molecule mechanical manipulation techniques have been widely utilized for studying mechanical properties of proteins in diverse areas such as protein folding, binding, translocation, degradation, and enzymatic catalysis.[1–12] Polyprotein or DNA handles are typically affixed to the target of interest so that the molecule can be readily attached to chamber surface or beads/tips so that force can be applied. The molecular handles prevent undesirable nonspecific adhesion of target proteins to the surfaces, and their well-characterized force responses validate the single molecule conditions. Although polyprotein handles can be conveniently encoded, they are often not suitable for more complex targets like membrane proteins or ribosomes where they may prevent protein expression or assembly.[13–16] Oftentimes, DNA handles are more convenient because they can be attached after expression and assembly. Moreover, DNA handles can provide longer spacers between surfaces or beads and are therefore more appropriate for larger complexes.

A simple approach for coupling DNA handles to proteins is to introduce cysteine residues at the point of attachment, allowing covalent linkage using thiol chemistry.[17–25] Thiol coupling imposes limitations, however, because all the other reactive cysteines must be removed, which is not always possible, and it becomes increasingly difficult with larger proteins that have more cysteines. HaloTag has been used, but requires the addition of a large protein domain to the target, which can limit its utility.[16,26–29] A clever ybbR tag method was devised

Grant sponsor: NIH; Grant number: R01GM063919 (to J.U.B.).

*Correspondence to: James Ulrich Bowie, UCLA, 611 Charles E. Young Dr. E, Los Angeles, CA 90095-1570. E-mail: bowie@mbi.ucla.edu

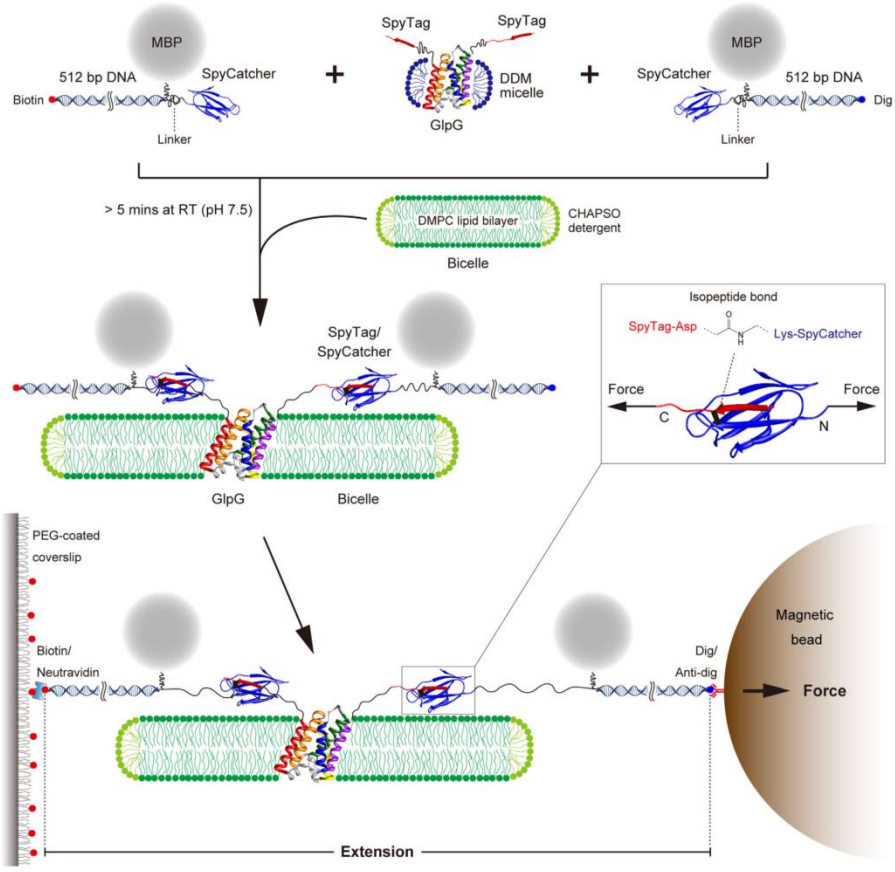


Figure 1. Handle attachment procedure for single molecule manipulations. The DNA handles are first linked to the MBP-SpyCatcher fusion using the traditional thiol chemistry approach. Separately the target protein (in this case SpyTag-GlpG) is prepared with added SpyTag peptides (genetically encoded). Upon simple mixing, SpyTag-GlpG becomes covalently linked to the MBP-SpyCatcher-DNA via an isopeptide bond (see box inset). For unfolding experiments with GlpG, we employ bicelles to provide bilayer conditions, which are fully compatible with the attachment method. The DNA-conjugated GlpG can then be tethered between the glass support and magnetic bead via the biotin/neutravidin and dig/antidig linkages, and subject to mechanical force.

to attach DNA handles to green fluorescent protein or the enormous ribosome.[30–34] In this method, a short ybbR peptide tag was inserted into the target proteins, and Sfp phosphopantetheinyl transferase was then used to connect an ~20 to 30 base oligonucleotide linked to the thiol on the pantothenate moiety of Coenzyme A. Once the oligonucleotide was attached, longer handles could then be appended by ligation. While this method can clearly be used for handle attachment to complex molecules, all the manipulations and incubations required for attachment must be performed on the target itself, which is fine for abundant and highly stable targets like ribosomes, but can be difficult if the target is hard to obtain in large quantities or is unstable as is the case for many membrane proteins.

Here we describe an approach using a SpyTag/SpyCatcher system in which all the difficult DNA attachment chemistry is performed on the easy to obtain SpyCatcher protein. Once the SpyCatcher-DNA conjugate is made, attachment to the target only requires a short incubation at room temperature. Thus, the method should be suitable for DNA handle attachment to a wide variety of complex and sensitive target molecules. We tested the approach on a previously characterized membrane protein GlpG[22] and

show that the attachment method does not significantly alter the mechanical properties of the system.

Results

The handle attachment method is outlined in Figure 1. First SpyCatcher-DNA conjugate reagents are prepared, bearing either biotin or digoxigenin (dig) affinity tags that can be attached to the chamber surface or beads. The target protein encoding the 13 amino acid SpyTag peptide sequence is then mixed with the SpyCatcher-DNA conjugate allowing covalent attachment to the SpyTag via an isopeptide bond between a Lys side chain in SpyCatcher and an Asp side chain in the SpyTag.[35,36]

For convenient purification of the protein and the DNA conjugate, we expressed SpyCatcher as a fusion to maltose binding protein (MBP-SpyCatcher). We introduced a unique Cys into the MBP-SpyCatcher protein to which 512 bp DNA handles could be attached using maleimide chemistry. The MBP-SpyCatcher-DNA conjugate could be conveniently purified away from unconjugated DNA or protein by employing anion exchange chromatography to select for molecules that possess DNA moiety, and then by amylose affinity chromatography to select for molecules that possess MBP moiety. Although MBP is a relatively large affinity tag, we found that DNA conjugation prevented affinity purification using a simple 6xHis tag. Moreover, because both the target protein and DNA attachment points are within the SpyTag/SpyCatcher domain (i.e., MBP is not on the mechanical pulling axis), the MBP purification tag will not experience a force during the pulling experiments, so it is not necessary to remove it. Note that the complex manipulations needed to prepare and purify the MBP-SpyCatcher-DNA conjugate are all performed on the simple and easy to prepare MBP-SpyCatcher protein. Once the conjugate is prepared, it can simply be used as an added reagent to attach to any target possessing a SpyTag sequence.

To test the method, we employed the rhomboid protease GlpG as a target protein. GlpG is a membrane protein with six transmembrane helices. Previously, we had extensively characterized the mechanical unfolding of GlpG using DNA handles coupled using traditional disulfide attachment to introduced Cys residues at the N- and C-termini,[22] allowing us to test whether the new method affected the observed mechanical properties.

SpyTag peptides (AHIVMVDAYKPTK)[35,36] were appended to the N- and C- termini of GlpG (SpyTag-GlpG) to allow for linkage to the MBP-SpyCatcher-DNA conjugate. Figure 2(a) shows the progress of the linkage reaction when we added a *nominal* 13:1 molar ratio of SpyTag-GlpG to MBP-SpyCatcher-DNA (1.2 μM SpyTag-GlpG, 0.09 μM MBP-SpyCatcher-DNA). We obtained ~80% linkage in only 5 min and nearly

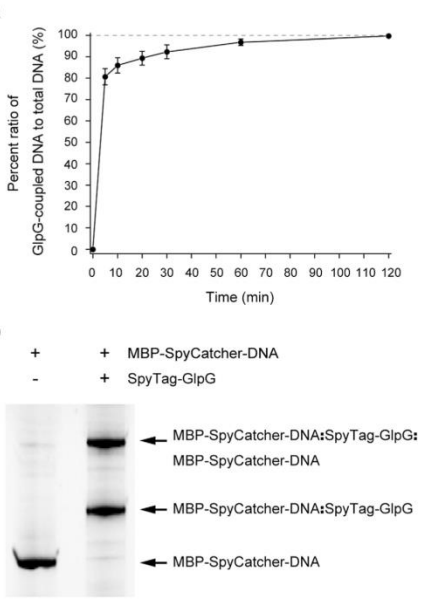


Figure 2. Kinetics and gel analysis for DNA handle attachment to GlpG. (a) Percent ratio of GlpG-coupled DNA to total DNA as a function of time. The percent ratios are estimated from triplicate gels. The reaction was performed at 22°C and pH 7.5. (b) SDS-PAGE gel showing the attachment of DNA handles to GlpG after 1 h incubation. The upward shifts from the MBP-SpyCatcher-DNA band indicate the coupling of GlpG to one or two DNA handles. The gel is stained with nucleic acid gel stain.

quantitative linkage after 1 h at 22°C. Both singly and doubly conjugated SpyTag-GlpG were observed in an approximately 1:1 ratio [Fig. 2(b)]. We were surprised that we needed to add such a large excess of SpyTag-GlpG relative to MBP-SpyCatcher-DNA, but it is possible that at these low concentrations SpyTag-GlpG could stick to microcentrifuge tube surfaces so that the effective concentration is much lower than the nominal concentration. Regardless of the cause, it is a simple matter to perform a titration to determine the appropriate concentration ratios.

There are many possible products from the conjugation reaction. In addition to singly conjugated forms, there are three possible doubly conjugated forms (two biotin tags, two dig tags, or a mixture), yet all are invisible to the magnetic tweezer experiment except the form with both a biotin tag and a dig tag, which can bind to both a neutravidin treated glass surface and an antidig treated magnetic bead (Fig. 1, bottom).

The tethered GlpG proteins were subjected to a slow force ramp (~0.3 pN/s) by approaching a pair of

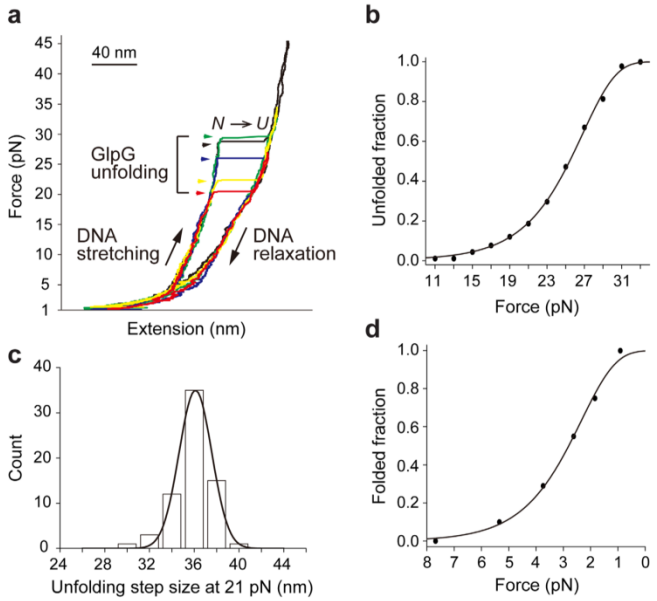


Figure 3. Single-molecule forced unfolding experiments for tethered GlpG. (a) Representative force-extension curves showing the repetitive unfolding events of a single GlpG linked using the SpyTag/SpyCatcher system described here. The symbols *N* and *U* denote the folded and unfolded states respectively. (b) Unfolded fraction as a function of force ($n = 91$). Fitting yields unfolding rate at zero force. (c) Unfolding step size distribution at a constant 21 pN ($n = 67$). (d) Folded fraction as a function of force ($n = 88$). Fitting yields folding rate at zero force.

magnets to magnetic beads. As a response to the force, the extension, i.e., the end-to-end distance of the GlpG-DNA conjugate increased in a gradual manner as the DNA stretches and around 25 pN abrupt jumps in extension occurred, indicating the highly cooperative unfolding of the GlpG protein (Fig. 3).

To test whether the new tethering method altered the observed mechanical unfolding/folding properties of GlpG, we measured the folding and unfolding rates and step size for the unfolding events as described previously[22] (Fig. 3 and Table I). As shown in Table I, the results were similar to those found previously with a direct disulfide tethering method ($P > 0.05$). These results indicate the MBP-SpyCatcher-DNA handles do not significantly alter the observed mechanical unfolding events. Moreover, the SpyTag/SpyCatcher complex has high mechanical stability[35] and did not provide any additional rupture events at forces up to ~45 pN, a typical force regime in protein mechanical responses (Fig. 3).

Discussion

Our approach for protein/DNA conjugation has several advantages for the mechanical manipulation of proteins. (1) The hard protein chemistry of DNA handle coupling and purification steps can be done on the well-behaved and less valuable SpyCatcher protein rather than a precious target protein. (2) The rapid, spontaneous attachment reaction is a major advantage for proteins sensitive to thermal stress and vulnerable to aggregation. (3) The rapid, spontaneous reaction also allows us to use low concentration (tens of nM) of DNA conjugate, which is hard to obtain in large amounts. (4) The binding of SpyTag and SpyCatcher is not a thiol chemistry, and thus we can study native proteins without engineering out the native cysteines of the target proteins. (5) The high mechanical rigidity of the SpyTag/SpyCatcher complex allows for observations of unfolding transitions without complications from additional rupture events from the complex. A limitation of the method is that it would be difficult to attach handles in the middle of a protein for pulling with different geometries. It is not impossible, however, as the SpyTag peptide can also be internally encoded.[35] The new handle attachment method established here should be generally applicable for the mechanical manipulation of highly complex and sensitive systems.

Table I. *Step Size and Rates for GlpG Unfolding/Folding from Two Linkage Methods*

	Unfolding size (nm)	Unfolding rate ($\times 10^{-5}$ s^{-1})	Folding rate ($\times 10^{-2}$ s^{-1})
SpyTag	36.10 (0.05)	8.53 (1.43)	3.56 (0.53)
Disulfide	35.63 (0.09)	5.64 (0.91)	3.91 (0.54)
P	0.1385	0.0900	0.6442

Numbers in parentheses indicate standard deviation for the step size and standard error for the rate constants. The *P* values were assessed by Welch's *t*-test.

Materials and Methods

Cloning, expression, and purification of SpyTag-GlpG

The GlpG protein employed here corresponds to residues 87 to 276 of *E. coli* GlpG.[22] The SpyTag peptide (AHIVMVDAYKPTK)[35,36] and a linker (GSGESG) were added to both N- and C-termini by two sequential PCR amplifications. The prior GlpG construct in a pTrcHisB vector[22] was used as a template for the first PCR reaction. We employed the following primers that include the SpyTag and a linker (annealing regions underlined): FWD: 5'-AATGGTCGATGCGTATAAACCGACGAAAGGTTCAGGAGAGTCAGGCGCCGCCTGTTTGCG−3', REV: 5'-GGCGTCCACCATCACGATGTGGGCACCACTTTCACCACTACCACATTTTCGTTTTCGCGC−3'. The gel-purified product was then used as the template for the second PCR reaction. The second primers completed the SpyTag sequence and include a 24 to 29 bp overlap with SacI/HindIII digested pTrcHisB. Second PC R primers were (annealing regions underlined): FWD: 5'-ATGGGGCATCATCATCATCATCATGAGCTCGCTCATATTGTAATGGTCGATGCGTATAAACC−3', REV: 5'-CTCATCCGCCAAAACAGCCAAGCTTACTCCTTCGTCGGCTTGTAGGCGTCCACCATCACG−3'. The gel-purified PCR product was cloned into pTrcHisB vector at the SacI and HindIII sites, preserving the N-terminal 6xHis-tag to generate the SpyTag-GlpG construct shown in Figure 4(a). The SpyTag-GlpG protein was expressed in *E. coli* BL21-Gold (DE3) and purified as previously described.[22] Aliquots of the purified GlpG (~15 μ*M*) in 50 m*M* Tris (pH 7.5), 150 m*M* NaCl, 0.1% DDM were flash frozen in liquid nitrogen, and stored at −80°C.

Cloning, expression, and purification of MBP-SpyCatcher

A DNA segment encoding the maltose binding protein (MBP) tag was PCR-amplified from the pMAL-c2X vector (New England Biolabs) using the primers, FWD: 5'-TTTAACTTTAAGAAGGAGATATACATATGAAAATCGAAGAAGGTAAACTGGTAATC and REV: 5'-ATGGTGATGGTAGTACGACATGCCATTAGTCTGCGCGTC. The gel-extracted DNA fragment was inserted into NdeI-linearized pDEST14-SpyCatcher vector[35] using the Gibson ISO assembly procedure.[37] The resulting plasmid was then modified using a PCR-based mutagenesis strategy, which replaced the three amino acids (Ala-Met-Val) following the TEV protease site with the sequence Gly-Cys-Gly to allow for DNA conjugation. Briefly, the MBP-containing SpyCatcher plasmid was used as the template for PCR-amplification with the primers, FWD: 5'-GGTGGTTGTGGTGATACCTTATCAGGTTTATCAAGTGAGCAAG and REV: 5'-CTGAAAATACAGGTTTTCGGTCGTTG, the reaction was then treated with DpnI restriction enzyme and subsequently with polynucleotide kinase and T4 DNA ligase, before transformation into competent *E. coli* DH5α. Putative positive transformants were mini-prepped and their sequence confirmed by DNA sequencing (Genewiz) to ensure that the final expression plasmid, pMBP-SpyCatcher, encodes a fusion protein with the following organization of MBP, 6xHis, TEV protease site, Cys-containing linker, and SpyCatcher shown as in Figure 4(b).

The expression plasmid was transformed into *E. coli* BL21(DE3) and an overnight culture was used to inoculate four Ultra Yield flasks (Thomson Instrument Company) each containing 1 L of Terrific broth supplemented with ampicillin (100 μg/mL). The cells were grown with shaking at 37°C until they reached an OD600 of ~1.0 at which point the temperature was lowered to 30°C and protein expression induced with 0.75 m*M* IPTG. Growth was continued for an additional 4 to 5 h and the cells harvested by centrifugation. The cell pellet was frozen pending cell lysis. After thawing, the cell pellet was resuspended in lysis buffer (50 m*M* Tris (pH 8.0), 200 m*M* NaCl, 1 m*M* EDTA, 1 m*M* DTT, 1 m*M* PMSF) supplemented with a cocktail of protease inhibitors (Sigma-Aldrich). The cells were lysed using an EmulsiFlex C3 cell homogenizer (Avestin) and the lysate was centrifuged at 39,000*g* for 40 min at 4°C. The supernatant was decanted, filtered using a 1.0 μm polyethersulfone membrane, and loaded on a 5 mL MBPTrap column (GE Healthcare Bio-Sciences) equilibrated in 50 m*M* Tris (pH 8.0), 200 m*M* NaCl, 1 m*M* TCEP (Buffer A) at 2 mL/min. The column was washed extensively with Buffer A and then bound proteins were eluted with a step gradient to 10 m*M* maltose in Buffer A. The fractions containing the fusion protein were pooled and concentrated, and the fusion protein was further purified by size exclusion chromatography using a Sephacryl S-100 column (GE Healthcare Bio-Sciences) equilibrated in 50 m*M* Tris (pH 8.0), 150 m*M* NaCl, 1 m*M* TCEP at a flow

Figure 4. Amino acid sequences for (a) SpyTag-GlpG and (b) MBP-SpyCatcher. The SpyTag and SpyCatcher are shown in red, GlpG and MBP in blue, 6xHis-tag in green, and TEV protease site in yellow. The unique cysteine residue in MBP-SpyCatcher for DNA handle conjugation is underlined.

rate of 1 mL/min. Fractions containing pure MBP-SpyCatcher proteins were pooled, concentrated to ~150 μ*M*, aliquots flash-frozen with liquid nitrogen, and stored at −80°C pending use.

Conjugation of DNA to MBP-SpyCatcher

A 512 bp DNA possessing a distinct binding tag (biotin or digoxigenin), was PCR-amplified from a λ DNA template (final 2.5 μg/mL; New England Biolabs, N3011S) using forward primer CATGTGGGTGACGCGAAA modified with 5′ amine group, and reverse primer TCGCCACCATCATTTCCA modified with either 5′ biotin or digoxigenin (final 1 μ*M* each; Integrated DNA Technologies). 4 mL of each PCR product was purified using a HiSpeed Plasmid Maxi Kit (Qiagen, 12662), eluted into ~1 mL of 0.1*M* sodium bicarbonate (pH 8.4), and then mixed to total ~2 mL (each ~1.5 μ*M* final concentration).

To attach a maleimide functional group to the amine-modified end, we used the SM(PEG)2 reagent (Thermo Fisher Scientific, 22102), which is a heterobifunctional crosslinker with *N*-hydroxysuccinimide (NHS) ester and the desired maleimide group [Fig. 5(a)]. 8 μL of 250 m*M* SM(PEG)2 crosslinker (dissolved in dimethyl sulfoxide) was added to the ~2 mL DNA mixture and incubated for 20 min at room temperature (RT). The sample was then divided into two ~1 mL samples for purification with the HiSpeed Plasmid Maxi Kit, each eluted into ~1 mL of 0.1*M* sodium phosphate (pH 7.3), 150 m*M* NaCl, and then combined to yield ~2 mL of ~1.1 μ*M* DNA conjugate mixture. To attach the DNA handles to the MBP-SpyCatcher protein, we mixed 200 μL of the purified MBP-SpyCatcher protein to the ~2 mL DNA handle with final concentration of ~14 μ*M* and ~1 μ*M* each, and incubated the mixture for 2 h at RT [Fig. 5(a)].

To separate unconjugated SpyCatcher proteins and DNA in the SpyCatcher/DNA mixture, we first employed an anion exchange column. The mixture (DNA and MBP-SpyCatcher-DNA) was bound to a 1 mL HiTrap Q HP column (GE Healthcare Bio-Sciences, 29-0513-25) equilibrated with 20 m*M* Tris (pH 7.5), and then eluted using a gradient to 1*M* NaCl (gradient volume = 25 mL). The eluted sample was collected in 1 mL fractions, and peak fractions of the DNA constructs were pooled. Unreacted DNA was further removed by employing an amylose affinity resin that only captures the construct with MBP tag, i.e., MBP-SpyCatcher-DNA. We loaded the pooled sample onto a column charged with ~1 mL amylose resin (New England Biolabs, E8021S) and incubated for 2 h at RT with slow tilt rotation. We washed the column with ~60 mL of 50 m*M* Tris (pH 7.4), 200 m*M* NaCl, 20 m*M* imidazole (Buffer B), and then eluted the MBP-SpyCatcher-DNA from the amylose resin by applying 10 m*M* maltose in Buffer B. Using 30K Amicon Centrifugal Filter Unit (EMD Millipore, UFC803024), the MBP-SpyCatcher-DNA construct was concentrated to ~100 n*M* of ~300 μL and stored at −80°C in 10 μL aliquots. The covalent conjugation of MBP-SpyCatcher and DNA was confirmed by separation by 4 to 20% SDS-PAGE (GenScript, M42015) with staining for DNA molecules with GelRed nucleic acid gel stain (Biotium, 41003) [Fig. 5(b)].

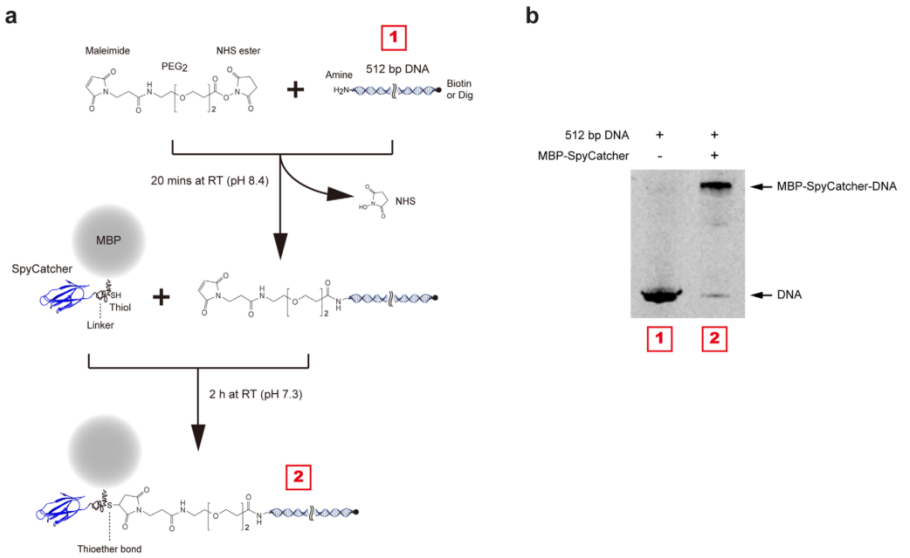


Figure 5. Linking DNA to MBP-SpyCatcher. (a) Schematic diagram of the conjugation process. The amine-modified 512 bp DNA were activated with a maleimide group and then conjugated to MBP-SpyCatcher protein by maleimide-cysteine crosslinking chemistry. The purification of MBP-SpyCatcher-DNA construct was performed sequentially by ion exchange and MBP-tag affinity chromatography. (b) SDS-PAGE gel showing successful conjugation. The upward shift from the DNA band indicates the MBP-SpyCatcher coupling to DNA. The gel was stained with a nucleic acid gel stain.

DNA-SpyTag-GlpG sample preparation

1 μL of ~15 μ*M* SpyTag-GlpG and 11 μL of ~100 n*M* MBP-SpyCatcher-DNA were mixed so that the final solution contained 45 m*M* Tris (pH 7.4), 150 m*M* NaCl, 2 m*M* TCEP, 15 m*M* imidazole, 7 m*M* maltose, and 0.1% DDM. After 1 h incubation at RT, 9 μL of the mixture was diluted to ~800 μL with 50 m*M* Tris (pH 7.5), 150 m*M* NaCl, 2 m*M* TCEP, 1.3% bicelle to make the final DNA concentration ~1.0 n*M*, and stored at −80°C in 10 μL aliquots. The bicelle mixture consisted of DMPC lipid (Avanti Polar Lipids, 850345P) and CHAPSO detergent (Affymetrix, C317) at a 2.5:1 molar ratio as previously described.[22]

Coating of magnetic bead with antidigoxigenin

We used N-hydroxysuccinimide (NHS) ester crosslinking chemistry to coat magnetic beads with antidigoxigenin (antidig). With a magnetic concentrator (Thermo Fisher Scientific, 12321D), 34 μL of 2.8 μm carboxylated magnetic beads (Thermo Fisher Scientific, 14305D) were equilibrated in 1 mL of 0.1*M* MES (pH 6.0), 0.5*M* NaCl, 0.1% Tween 20. To activate the carboxylic acids with NHS esters, we added 100 μL of a solution of 50 m*M* 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide and 100 m*M* sulfo-NHS and incubated for 15 min at RT with gentle mixing on a rotator. The activated magnetic beads were equilibrated in 1 mL of 0.1*M* sodium phosphate (pH 7.3), 150 m*M* NaCl, 0.1% Tween 20 using the magnetic concentrator. We then added 60 μL of 1.1 μ*M* antidig (Sigma-Aldrich, 11333089001) and incubated for 3 h at RT with gentle mixing on a rotator. To quench unreacted NHS esters, we equilibrated the beads in 1 mL of 0.1*M* Tris (pH 8.5), 0.1% Tween 20 and incubated for 10 min at RT with gentle mixing on a rotator. The antidig coated magnetic beads were washed with 0.1*M* sodium phosphate (pH 7.3), 150 m*M* NaCl, 0.01% Tween 20, resuspended in 34 μL of the same buffer, and then stored at 4°C.

PEG/biotin coating of glass coverslip

We coated the surface of glass coverslips (25 × 50 mm, No 1.5; VWR) with a combination of two different NHS ester-functionalized polyethylene glycols (PEG-NHS),[38,39] i.e., 125:1 molar ratio of methylated PEG-NHS (mPEG; Laysan Bio, MPEG-SVA-5000) and biotin-conjugated PEG-NHS (biotin-PEG; Laysan Bio, Biotin-PEG-SVA-5000) using amine-NHS ester crosslinking chemistry. The mPEG molecules prevent nonspecific binding events of beads and proteins whereas the biotin-PEGs specifically bind DNA-conjugated GlpGs via biotin-neutravidin linkage. The coverslips were first cleaned with 1*M* KOH in a sonication bath for 20 min, washed with distilled water, and then functionalized with amine

groups (silanization reaction) using a solution of *N*-(2-aminoethyl)−3-aminopropyltrimethoxysilane (UCT, A0700), acetic acid and methanol in a 1:5:100 volume ratio. The coverslips were incubated with the silanization solution for 12 min at RT, sonicated for 1 min in the sonication bath, and incubated for an additional 20 min at RT. The coverslips were washed with methanol, then distilled water, and then dried with nitrogen gas. PEGs were coupled by incubating 60 μL of a PEG-NHS solution (40 m*M* PEG mixture in 0.1*M* sodium bicarbonate, pH 8.4) between two coverslips for 4 to 5 h at RT in a humidity chamber. The coverslips were washed with distilled water, dried with nitrogen gas, and stored at −20°C.

Single-molecule microscope sample chamber

A sample chamber of ∼15 μL channel volume (≡1 CV) was constructed by putting together a PEG/biotin-coated coverslip (see above) and a KOH-cleaned coverslip (24 × 40 mm, No 1.5; VWR) with double-sided tape. 1 μL of 1.0 μm sized polystyrene beads coated with streptavidin (Polysciences, 24162) were washed and equilibrated in 300 μL of 0.1*M* sodium phosphate (pH 7.4), 150 m*M* NaCl, 0.1% Tween 20 by repeated centrifugation and resuspension. We injected 1 CV of the bead solution into the ∼90 μm thick channel by capillary action and incubated for 90 s at RT. The nonmagnetic polystyrene beads bind to PEG/biotin-coated surface via biotin-streptavidin linkage and are used to correct vertical and lateral drift of sample stage. The chamber surfaces were further passivated by flowing through three CVs of 100 mg/mL bovine serum albumin, and then washed with 50 m*M* Tris (pH 7.5), 150 m*M* NaCl. The sample chamber was then equilibrated in 50 m*M* Tris (pH 7.5), 150 m*M* NaCl, 2 m*M* TCEP, 1.3% bicelle (Buffer C). The bicelle mixture consisted of DMPC lipid and CHAPSO detergent at a 2.5:1 molar ratio as described above.

To bind the DNA-linked GlpG constructs to the PEG/biotin-coated surface, we first attached neutravidin (NTV) molecules to the biotin-modified DNA handles. We added 1 μL of 167 n*M* NTV (Thermo Fisher Scientific, A-2666) to 10 μL of the DNA-linked GlpG stock described above and incubated for 20 min at RT (NTV: biotin-DNA = 30: 1). The NTV-bound GlpG sample was then diluted to 22 μL total volume with Buffer C (i.e., ∼100 pM of NTV-biotin-DNA-GlpG-DNA-digoxigenin), and 1 CV of the diluted sample was injected into the chamber and incubated for 10 min at RT to allow binding to the PEG/biotin-coated glass surface. The vacant biotin binding sites of the neutravidin/streptavidin molecules were blocked by washing with three CVs of short 30 nt DNA oligonucleotides modified by biotin (10 μ*M* in Buffer C) and incubating for 5 min at RT. The chamber was then equilibrated in 50 m*M* Tris-HCl (pH 7.5), 150 m*M* NaCl, 1.3% bicelle (Buffer D). With the magnetic concentrator (the same one used above), 1 μL of antidig-coated magnetic beads were washed with 300 μL of 50 m*M* Tris (pH 7.5), 150 m*M* NaCl once and with 50 μL of Buffer D twice. One CV of the magnetic beads, resuspended in 50 μL Buffer D, were introduced into the chamber, and incubated for 30 min at RT to bind surface-tethered GlpGs having digoxigenin-DNA handles (Fig. 1, bottom).

Magnetic tweezer instrumentation

The magnetic tweezer apparatus was custom-built as previously described.[21,22,40,41] The tweezer setup was constructed on an inverted microscope (Olympus, IX73) with a motorized XY stage (ASI, MS-2000 XY Automated Stage). Magnetic beads in a glass sample chamber were illuminated by 455 nm light-emitting diode (Thorlabs, M455L3) and diffraction pattern images from the beads were captured at 60 Hz frame rate by a charge-coupled device camera (JAI, CM-040GE). We track the lateral (x,y) and vertical (z) motions of the beads by analyzing the diffraction images using customized software programs written in LabView (National Instruments).[21,22,40,41] The lateral movement was tracked by calculating the maximum self-convolution for diffraction pattern intensity profiles ($I(x)$, $I(y)$). The vertical movement (extension change) was tracked by calculating the minimum χ^2 estimate of a radial intensity profile ($I(r)$) with precalibration data, i.e., a stack of the radial intensity profiles as a function of bead height. The calibration data was measured by moving the focal plane in known increments with a nanopositioning piezoelectric stage (Mad City Labs, Nano-F100S). Thermal drift of the microscope stage was corrected by simultaneously tracking the nonmagnetic polystyrene beads immobilized on the chamber surface.

To generate a magnetic field gradient, we employed a pair of two permanent neodymium magnets (10 × 10 × 12 mm) separated by 1 mm with antiparallel alignment of magnetic moments. Vertical and rotational motions of the magnet pair was manipulated with a translation stage (Physik Instrumente, M-126.PD1) and a rotation stage (Physik Instrumente, DT-50). We calibrated the mechanical tension applied to tethered molecules, as a function of the magnet height. At each magnet height, we measured the end-to-end distance (extension) of a tethered molecule (L) and lateral fluctuation of a magnetic bead (δx^2), and then calculated the applied tension from an equation $F_{\mathrm{mag}} = k_{\mathrm{B}}TL/\delta x^2$ where $k_{\mathrm{B}}T$ is the thermal energy. The equation was derived by assuming an inverted pendulum for the trapped molecule in the magnetic potential.[40,42] The force range of our magnetic tweezers is approximately 0.01 pN to 70 pN. We controlled the pair of

magnets with translational speeds of 0.1 mm/s during the gradual force ramp [average ~0.3–0.4 pN/s, Fig. 3(a)]. The representative traces shown in Figure 3(a) were median-filtered with a 50-point window size.

Acknowledgments

Authors thank members of the lab for comments on the manuscript.

References

1. Bustamante C, Chemla YR, Forde NR, Izhaky D (2004) Mechanical processes in biochemistry. Annu Rev Biochem 73:705–748.
2. Kedrov A, Janovjak H, Sapra KT, Muller DJ (2007) Deciphering molecular interactions of native membrane proteins by single-molecule force spectroscopy. Annu Rev Biophys Biomol Struct 36:233–260.
3. Borgia A, Williams PM, Clarke J (2008) Single-molecule studies of protein folding. Annu Rev Biochem 77: 101–125.
4. Alegre-Cebollada J, Perez-Jimenez R, Kosuri P, Fernandez JM (2010) Single-molecule force spectroscopy approach to enzyme catalysis. J Biol Chem 285: 18961–18966.
5. Puchner EM, Gaub HE (2012) Single-molecule mechanoenzymatics. Annu Rev Biophys 41:497–518.
6. Javadi Y, Fernandez JM, Perez-Jimenez R (2013) Protein folding under mechanical forces: a physiological view. Physiology 28:9–17.
7. Zoldak G, Rief M (2013) Force as a single molecule probe of multidimensional protein energy landscapes. Curr Opin Struct Biol 23:48–57.
8. Woodside MT, Block SM (2014) Reconstructing folding energy landscapes by single-molecule force spectroscopy. Annu Rev Biophys 43:19–39.
9. Cordova JC, Das DK, Manning HW, Lang MJ (2014) Combining single-molecule manipulation and single-molecule detection. Curr Opin Struct Biol 28:142–148.
10. Chen Y, Radford SE, Brockwell DJ (2015) Force-induced remodelling of proteins and their complexes. Curr Opin Struct Biol 30:89–99.
11. Olivares AO, Baker TA, Sauer RT (2016) Mechanistic insights into bacterial AAA plus proteases and protein-remodelling machines. Nat Rev Microbiol 14:33–44.
12. Ryu JK, Jahn R, Yoon TY (2016) Progresses in understanding NSF-mediated disassembly of SNARE complexes. Biopolymers 105:518–531.
13. del Rio A, Perez-Jimenez R, Liu R, Roca-Cusachs P, Fernandez JM, Sheetz MP (2009) Stretching single talin rod molecules activates vinculin binding. Science 323:638–641.
14. Junker JP, Ziegler F, Rief M (2009) Ligand-dependent equilibrium fluctuations of single calmodulin molecules. Science 323:633–637.
15. He C, Genchev GZ, Lu H, Li H (2012) Mechanically untying a protein slipknot: multiple pathways revealed by force spectroscopy and steered molecular dynamics simulations. J Am Chem Soc 134:10428–10435.
16. Popa I, Berkovich R, Alegre-Cebollada J, Badilla CL, Rivas-Pardo JA, Taniguchi Y, Kawakami M, Fernandez JM (2013) Nanomechanics of HaloTag tethers. J Am Chem Soc 135:12762–12771.
17. Cecconi C, Shank EA, Bustamante C, Marqusee S (2005) Direct observation of the three-state folding of a single protein molecule. Science 309:2057–2060.
18. Shank EA, Cecconi C, Dill JW, Marqusee S, Bustamante C (2010) The folding cooperativity of a protein is controlled by its chain topology. Nature 465:637–640.
19. Kim J, Zhang CZ, Zhang XH, Springer TA (2010) A mechanically stabilized receptor-ligand flex-bond important in the vasculature. Nature 466:992–995.
20. Gao Y, Sirinakis G, Zhang YL (2011) Highly anisotropic stability and folding kinetics of a single coiled coil protein under mechanical tension. J Am Chem Soc 133: 12749–12757.
21. Min D, Kim K, Hyeon C, Cho YH, Shin YK, Yoon TY (2013) Mechanical unzipping and rezipping of a single SNARE complex reveals hysteresis as a force-generating mechanism. Nat Commun 4:1705–1714.
22. Min D, Jefferson RE, Bowie JU, Yoon TY (2015) Mapping the energy landscape for second-stage folding of a single membrane protein. Nat Chem Biol 11:981–987.
23. Ryu JK, Min D, Rah SH, Kim SJ, Park Y, Kim H, Hyeon C, Kim HM, Jahn R, Yoon TY (2015) Spring-loaded unraveling of a single SNARE complex by NSF in one round of ATP turnover. Science 347:1485–1489.
24. Brenner MD, Zhou RB, Conway DE, Lanzano L, Gratton E, Schwartz MA, Ha T (2016) Spider silk peptide is a compact, linear nanospring ideal for intracellular tension sensing. Nano Lett 16:2096–2102.
25. Jahn M, Buchner J, Hugel T, Rief M (2016) Folding and assembly of the large molecular machine Hsp90 studied in single-molecule experiments. Proc Natl Acad Sci USA 113:1232–1237.
26. Los GV, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, Wood MG, Learish R, Ohana RF, Urh M, Simpson D, Mendez J, Zimmerman K, Otto P, Vidugiris G, Zhu J, Darzins A, Klaubert DH, Bulleit RF, Wood KV (2008) HaloTag: a novel protein labeling technology for cell imaging and protein analysis. ACS Chem Biol 3:373–382.
27. Taniguchi Y, Kawakami M (2010) Application of HaloTag protein to covalent immobilization of recombinant proteins for single molecule force spectroscopy. Langmuir 26:10433–10436.
28. Aubin-Tam ME, Olivares AO, Sauer RT, Baker TA, Lang MJ (2011) Single-molecule protein unfolding and translocation by an ATP-fueled proteolytic machine. Cell 145:257–267.
29. Olivares AO, Nager AR, Iosefson O, Sauer RT, Baker TA (2014) Mechanochemical basis of protein degradation by a double-ring AAA+ machine. Nat Struct Mol Biol 21:871–875.
30. Yin J, Straight PD, McLoughlin SM, Zhou Z, Lin AJ, Golan DE, Kelleher NL, Kolter R, Walsh CT (2005) Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. Proc Natl Acad Sci USA 102:15815–15820.
31. Maillard RA, Chistol G, Sen M, Righini M, Tan J, Kaiser CM, Hodges C, Martin A, Bustamante C (2011) ClpX(P) generates mechanical force to unfold and translocate its protein substrates. Cell 145:459–469.
32. Kaiser CM, Goldman DH, Chodera JD, Tinoco I, Jr, Bustamante C (2011) The ribosome modulates nascent protein folding. Science 334:1723–1727.
33. Pippig DA, Baumann F, Strackharn M, Aschenbrenner D, Gaub HE (2014) Protein-DNA chimeras for nano assembly. ACS Nano 8:6551–6555.
34. Goldman DH, Kaiser CM, Milin A, Righini M, Tinoco I, Jr, Bustamante C (2015) Ribosome. Mechanical force releases nascent chain-mediated ribosome arrest in vitro and in vivo. Science 348:457–460.
35. Zakeri B, Fierer JO, Celik E, Chittock EC, Schwarz-Linek U, Moy VT, Howarth M (2012) Peptide tag

forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. Proc Natl Acad Sci USA 109:E690–E697.
36. Li L, Fierer JO, Rapoport TA, Howarth M (2014) Structural analysis and optimization of the covalent association between SpyCatcher and a peptide Tag. J Mol Biol 426:309–317.
37. Gibson DG (2011) Enzymatic assembly of overlapping DNA fragments. Methods Enzymol 498:349–361.
38. Ha T, Rasnik I, Cheng W, Babcock HP, Gauss GH, Lohman TM, Chu S (2002) Initiation and re-initiation of DNA unwinding by the *Escherichia coli* Rep helicase. Nature 419:638–641.
39. Joo C, Ha T, Single-molecule FRET with total internal reflection microscopy. In: Ha T, Selvin P, Ed. (2012) Single-molecule techniques: A laboratory manual. New York: Cold Spring Harb Lab, pp 3–36.
40. Gosse C, Croquette V (2002) Magnetic tweezers: micromanipulation and force measurement at the molecular level. Biophys J 82:3314–3329.
41. Ribeck N, Saleh OA (2008) Multiplexed single-molecule measurements with magnetic tweezers. Rev Sci Instrum 79:094301.
42. Strick TR, Allemand JF, Bensimon D, Bensimon A, Croquette V (1996) The elasticity of a single supercoiled DNA molecule. Science 271:1835–1837.

Chapter 6

Methods for Understanding How Proteins Fold Within a Membrane

This chapter will be submitted as a review article in a special issue of the Journal of Molecular Biology.

Abstract

Membrane proteins are uniquely constrained by their environment. They must interact with the hydrophobic tails of lipids, charged headgroups, two discrete aqueous regions, and other proteins, while carrying out critical functions. How these proteins maintain a specific fold in the membrane is a major question of biology. Studying how membrane proteins fold in the particular context of the lipid bilayer comes with its own technical challenges. The traditional chaotropic denaturants for unfolding soluble proteins are not generally applicable to membrane proteins and often do not recreate a biologically relevant unfolded state. To confront these challenges, specialized methods have been developed to measure the thermodynamics and kinetics of membrane protein folding under more relevant and native conditions. Chemical denaturation, steric trapping, and force spectroscopy continue to be useful tools for understanding the underlying principles of how full-length membrane proteins fold.

Despite the importance of understanding protein folding, membrane protein folding has been neglected and research lags far behind that of soluble proteins. Membrane proteins are not an unimportant or particularly niche topic, but have been plagued by technical challenges that have impeded experiments. A simple publication search for "protein folding" turns up more than 44,000 results, versus "membrane protein folding", which finds just over 200 (http://www.ncbi.nlm.nih.gov/pubmed). While the search is blunt, the stark contrast in the

number of publications underscores that membrane protein folding studies have a long way to go to catch up with soluble proteins.

While both α-helical and β-barrel membrane proteins have challenges for folding studies, the approaches to examine their relevant folding and unfolding pathways must address each separately. α-helical membrane protein folding can be divided into a two-step process (Popot and Engelman, 1990). In the first step, the membrane protein is cotranslationally inserted via a ribosome-SecYEG translocon complex driven by sequence features of the transmembrane segment (Hessa et al., 2005). Once the topology is established, the transmembrane helices can fold into a final structure in a second step. After insertion and final folding in the membrane, membrane proteins dissociate from the insertion machinery and exist in an equilibrium between the second-stage unfolded and folded state. This second-stage is our primary focus for *in vitro* studies that measure the energetics of folding within the membrane environment. While these two steps are not completely separate during insertion, it is reasonable that second-stage folding can inform us about the tertiary interactions that drive folding of nascent membrane proteins.

In contrast, β-barrel outer membrane proteins in gram-negative bacteria are cotranslationally passed through the SecYEG translocon into the periplasm and kept in an unfolded conformation by chaperones for insertion by specialized outer membrane machinery, referred to as the β-barrel assembly machinery (BAM) (Knowles et al., 2009). Related apparatuses exist in mitochondria and chloroplasts (Paschen et al., 2003). Due to the nature of the final barrel fold, the possibilities for non-sequential interactions are limited, and the two-stage model is a poor descriptor of this folding process. *In vitro* folding studies of β-barrel proteins have commonly combined insertion and folding into a single step.

Membrane protein folding experiments are perhaps most hindered by the difficulty involved in recreating conditions in which we can test membrane protein folding in a membrane. Recapitulating a membrane environment that is compatible with *in vitro* folding studies has been a long-standing challenge. While soluble proteins can be examined in aqueous solutions, membrane proteins require a membrane mimetic. Membrane mimetics range from relatively convenient detergents, to lipid bilayer vesicles, and compromises in between. Detergents permit solubilization of membrane proteins in solution by forming a micellar belt around the hydrophobic surface. While micelles are a fairly easy to manipulate and can readily mix with denaturants, there are caveats to detergent studies that prevent accurate measurements of native folding conditions. Transient dissociation of detergents can expose aggregation prone hydrophobic regions. Unfolded states of membrane proteins in detergent are also not constrained to a 2-dimensional bilayer, but rather the individual helices can occupy a larger space in solution. The increased entropy of the unfolded state can destabilize detergent-solubilized membrane proteins. Furthermore, detergents do not accurately replicate the specific lipid interactions with protein as the polar headgroups are dissimilar and they only have a single hydrophobic acyl chain that is often shorter than a typical membrane lipid.

Bicelles recreate a bilayer environment in a disc bordered by detergent, offering a compromise between a complete membrane and the technical ease of detergents (Sanders and Prosser, 1998). The ratio of lipid to detergent, referred to as the q ratio, determines the radius of the bicelle discs, permitting comparisons of bicelle size on folding and stability. They offer a convenient bilayer alternative to the lipidic cubic phase for crystallization of membrane proteins (Faham et al., 2005), and at low temperature, bicelle mixtures have low viscosity and are easy to manipulate for homogenous incorporation of membrane proteins. However, bicelles transition to a viscous gel at higher temperatures, prohibiting experiments at physiological temperatures.

Though they do form a lipid bilayer, bicelles are still limited to certain lipid/detergent mixtures that use saturated shorter acyl chains.

Ideally, all membrane protein folding studies could be performed in reconstituted membranes. Lipid vesicles can be formed from a wide range of purified lipids to study effects of lipid shape, headgroup charge, and chain length. Purified membrane proteins can be incorporated into preformed vesicles spontaneously, or by detergent removal from a solubilized mixture of protein and lipids (Rigaud et al., 1995). Membrane proteins can also be studied in supported membranes (Sackmann, 1996), but this introduces interactions with the non-aqueous support surface. One of the major roadblocks is establishing an unfolded state without destroying the proteoliposome complex. Additionally, the large membrane structures formed by lipids introduce light scattering that can interfere with spectroscopic assays of protein function.

Though no membrane mimetic is perfect, the methods covered in this review utilize detergent micelles, bicelles, and lipid bilayers. Newer developments are helping to address the challenges of studying membrane protein folding in native environments.

Equilibrium Unfolding with Denaturants

Once a native membrane mimetic has been established, it still requires a mechanism to reversibly drive unfolding. To measure the unfolding energetics of proteins we need a method to shift the equilibrium towards the unfolded state where there is a measurable fraction of unfolded and folded protein, the "transition region." (Figure 6-1) Free energies of unfolding calculated from the transition region display a linear relationship with denaturant concentration, which can be extrapolated back to zero denaturant. Many soluble protein studies have employed chemical denaturation with the chaotropes, urea and guanidium chloride

(Jackson, 1998; Johnson and Fersht, 1995; Pace, 1986). Extrapolating experimentally determined unfolding free energies in the transition zone back to zero denaturant have been validated with theory and other experimental methods, such as hydrogen-deuterium exchange (Huyghues-Despointes et al., 1999). Though urea has been successfully used for β-barrel folding from a soluble unfolded state, these denaturants are not generally applicable for reversible unfolding of membrane proteins. For second-state folding of α-helical membrane proteins sodium dodecyl sulfate (SDS) has become the eminent denaturant for reversible unfolding within a membrane mimetic environment.

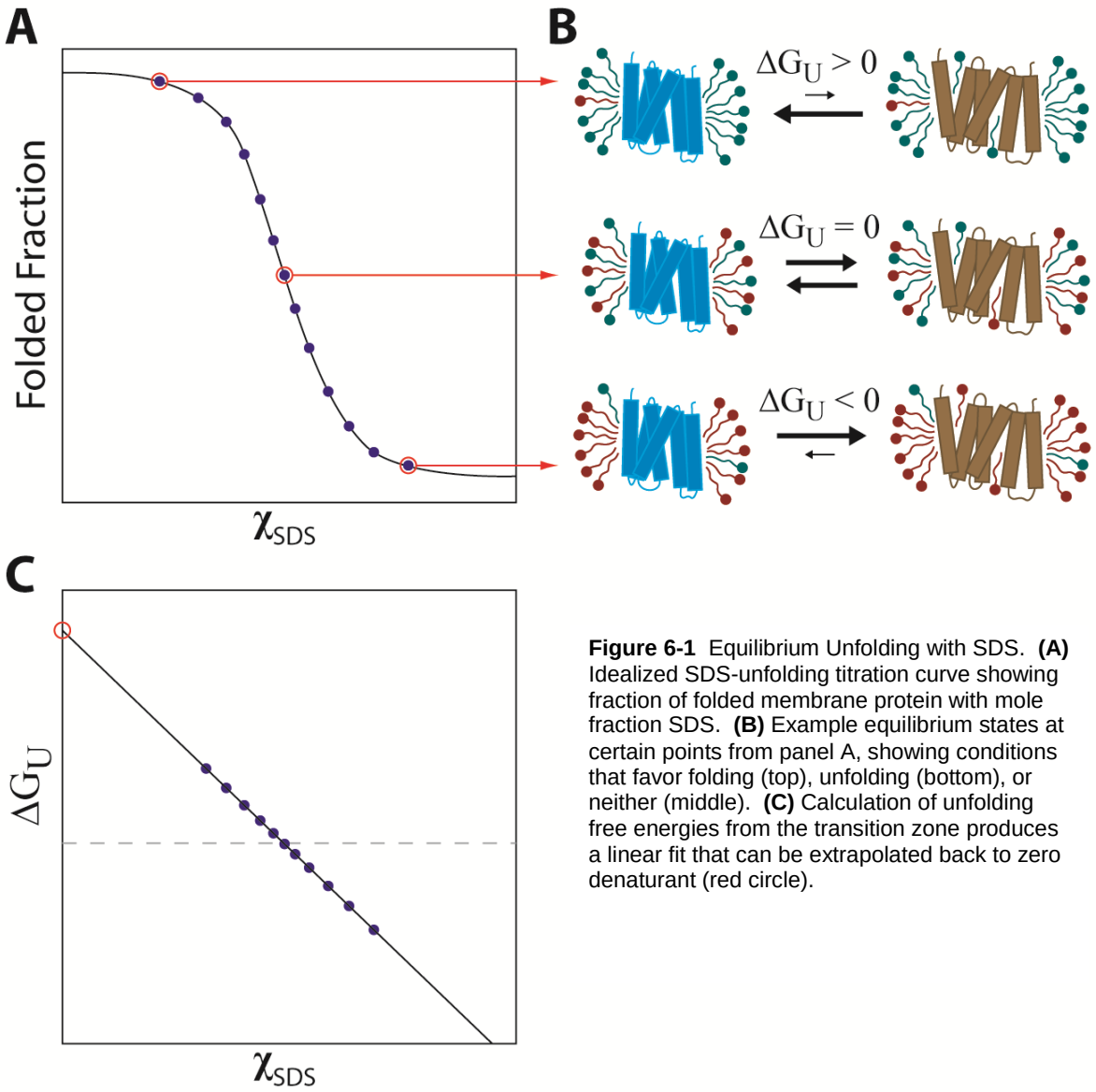


Figure 6-1 Equilibrium Unfolding with SDS. **(A)** Idealized SDS-unfolding titration curve showing fraction of folded membrane protein with mole fraction SDS. **(B)** Example equilibrium states at certain points from panel A, showing conditions that favor folding (top), unfolding (bottom), or neither (middle). **(C)** Calculation of unfolding free energies from the transition zone produces a linear fit that can be extrapolated back to zero denaturant (red circle).

Urea unfolding

Like soluble protein folding, studying membrane protein folding requires a method of shifting the equilibrium to the unfolded state and back to the folded state. While urea is an effective chaotrope to drive unfolding of soluble proteins, its success for reversibly unfolding membrane proteins has been mostly limited to the β-barrel class. Early experiments

demonstrated that β-barrel membrane proteins can insert and adopt a native fold from a urea-denatured state (Surrey and Jähnig, 1992). Further experiments investigated the kinetics of and requirements for spontaneous β-barrel incorporation (Kleinschmidt et al., 1999; Surrey et al., 1996), but it was not until Hong and colleagues examined the insertion of OmpA into lipid vesicles that a fully reversible folding assay was demonstrated for β-barrel proteins (Hong and Tamm, 2004). OmpA completely unfolds in a high concentration of urea and refolds into unilamellar vesicles upon denaturant dilution. The unfolded state is dissociated from the membrane ionic repulsion of negatively charged phosphatidylglycerol (PG) lipids and basic pH to negatively charge OmpA. Though this assay measures refolding from a soluble denatured state, the native state is folded in a membrane, allowing stability comparisons between different lipid compositions. OmpA exhibits a significant dependence on bilayer thickness and chain saturation. Stability increases with lateral pressure from longer acyl chains and unsaturated lipids. Folding of OmpA into its hourglass shape can alleviate the intrinsic lateral pressure in these membranes, demonstrating the importance of the lipid bilayer upon membrane protein folding.

The ability to solubilize unfolded β-barrel membrane proteins with chaotropic agents and spontaneously refold them into bilayers lead to further folding studies of other β-barrel membrane proteins. Later determinations of the free energy of unfolding for OmpW, OmpLA, and PagP employed guanidine as a denaturant (Moon and Fleming, 2011; Moon et al., 2013). Interestingly, stability of PagP was also measured in urea (Huysmans et al., 2010), which created an unfolded state ensemble that was associated with the bilayer, instead of the fully solubilized guanidine-denatured PagP, highlighting the importance of how denaturant choice affects which folding pathway is being measured. The Fleming lab has proposed that the high free energies of unfolding for OmpW, OmpLA, and PagP may help direct targeting of these outer membrane proteins across the periplasm via chaperones, preventing off-pathway misfolding

such as aggregation (Moon et al., 2013). Additionally, high kinetic stability in the outer membrane may be selected for due to the harsh environmental conditions

However, while urea and guanidine is a convenient denaturant for spontaneously incorporating β-barrel membrane proteins, it is problematic for refolding α-helical membrane proteins. Urea poorly mimics the hydrophobic bilayer, though in rare cases it can be used to study folding of certain helical membrane proteins. Studies of the small drug exporter EmrE used a high concentration of urea to supplement SDS unfolding, and refolded the protein into detergent micelles or lipid vesicles (Miller et al., 2009). The sugar transporter GalP can be refolded from a urea-denatured state, but has exposed helical regions and a solvent accessible ligand binding site that may make it particularly susceptible to equilibrium urea denaturation (Findlay et al., 2010). Recently, the Booth lab has developed a method for refolding a GPCR into n-decyl β-D-maltoside (DM) micelles from a urea-denatured state on a solid support (Bartolo et al., 2016), demonstrating that urea is still a useful tool for working with purified membrane proteins. However, for α-helical membrane proteins, mixed micelle folding studies have offered an alternative mode of chemical denaturation that is more tailored to the constraints of folding within the lipid bilayer.

SDS unfolding

Early studies with bacteriorhodopsin (bR) demonstrated that helical membrane proteins could be refolded from a chemically denatured state (Braiman et al., 1987; Huang et al., 1981; London and Khorana, 1982). SDS was used to denature bacteriorhodopsin and refolding was initiated by adding renaturing lipids, detergents, or mixed micelles of both detergent and lipid. Refolding in a continuos micellar phase from harsh anionic detergent to a lipid-detergent mixture permits the protein to refold in an entirely hydrophobic environment that is analogous to second-stage folding in which pre-formed transmembrane helices can associate. Further

studies of bR have examined refolding from SDS in detail. Paula Booth and colleagues were able to identify intermediates in the process of membrane protein folding from an SDS-unfolded state (Booth et al., 1995). The order and timescales of helix assembly and subsequent chromophore binding provide a view of how the protein may fold in the natural lipid bilayer.

Thermodynamic measurement of α-helical membrane protein stability was established by the Bowie lab using the membrane enzyme diacylglycerol kinase (DGK) (Lau and Bowie, 1997). Previous efforts to measure helical membrane protein stability using thermal denaturation irreversibly unfold the protein. Similar to using urea or guanidine denaturation to study soluble protein stability, SDS was titrated in to shift the equilibrium towards the unfolded state, and could be refolded back to the native state in DM. Careful placement of tryptophan residues allowed the stability of the transmembrane domain to be calculated separately from the unfolding of the cytoplasmic domain. A similar approach has been applied to bR (Cao et al., 2012; Chen and Gouaux, 1999; Faham et al., 2004) and the intramembrane protease GlpG (Baker and Urban, 2012). The Booth lab has comprehensively investigated the thermodynamic and kinetic stability of bR using SDS (Curnow and Booth, 2007). The transition regions for DGK and bR occur between ~0.6 and 0.9 χ_{SDS} and thus the unfolding free energies require a fairly long linear extrapolation back to 0 χ_{SDS}. Unfolding free energies for bR and the disfulfide bond reducing protein B (DsbB) (Otzen, 2003) were also measured by refolding and unfolding rates, the latter of which also require extrapolation back to zero denaturant. Until recent steric trapping studies, these extrapolations had not been verified or refuted by experimental evidence.

Despite the developments of membrane folding studies in SDS detergent, there are still only a handful of membrane proteins that have been probed in depth. Not all membrane proteins can refold from an SDS-unfolded state and some membrane proteins are resistant to SDS denaturation (Borgnia et al., 1999; Fujii et al., 1989; Hardie et al., 1996; Heginbotham et al., 1997; Sargiacomo et al., 1995; Spelbrink et al., 2005; Yang et al., 1999). So while SDS is a

useful tool for inspecting the membrane protein folding process, it will not work for every membrane protein. Many membrane proteins also lack a easily assayable functional characteristic to quantify the fraction folded during SDS denaturation. SDS unfolding has been measured using changes in circular dichroism (CD) and ultraviolet absorbance spectroscopy (Lau and Bowie, 1997), however, not all membrane proteins have measureable differences in their secondary structure by CD or can be easily produced in the quantities necessary for absorbance measurements.

Even for the membrane proteins whose folding has been studied in SDS, they are still dependent on linear extrapolations back to non-denaturing conditions. While linear extrapolations have been substantiated for soluble proteins, the denaturing effects of SDS at low mole fractions of the micellar phase is unclear from folding studies dependent on measuring stability in the transition region. Though non-denaturing detergents provide a convenient membrane mimetic for the folded state, it is important to consider the SDS-unfolded state. While SDS micelles preserve most of the α-helical content of membrane proteins, the arrangement of those helices is much less entropically constrained than it would be in the plane of the membrane. In high concentrations of SDS the protein can be solubilized in multiple micelles in the 'pearl necklace' model, permitting each helical TM segment to occupy a separate detergent micelle with three-dimensional flexibility in the loop regions. Ralf Langen and colleagues carried out a study of the SDS-unfolded state of bR using pulsed electron paramagnetic resonance to measure distances between specific locations in the protein (Krishnamani et al., 2012). While tertiary interactions were largely disrupted, the majority of end-to-end TM helix distances were similar, centered around the same distance as in the folded condition, but the distance distributions were more broad in SDS, suggesting that some fraying or unwinding of the helices may occur in the unfolded state. A study of peptides of the isolated transmembrane segments of bR showed that TMs were more helical when incorporated into

lipid vesicles versus SDS (Hunt et al., 1997). In the two-state model, the transmembrane helices are expected to remain stable without tertiary interactions, however, the degree of local unwinding and distortions in the absence of helical packing is unclear. Whether the helical disruptions observed in SDS happen during unfolding of the full-length protein in the lipid bilayer requires further study.

Even accepting the caveats of the SDS-unfolded state, folding studies in detergent systems are still problematic. Micelles, though convenient, do not perfectly recapitulate the lipid bilayer. Specific interactions between protein and lipids are lost, as are lateral packing pressure and effects from thickness and curvature of the membrane. Detergents can also introduce undesirable pathways of misfolding, such as aggregation, transient exposure of hydrophobic patches, or irreversible inactivation (Zhou et al., 2001). The limitations of SDS unfolding studies call for new methods suitable for measuring folding within a lipid bilayer.

Unfolding Membrane Proteins with a Steric Trap

To study how proteins fold in the membrane, the Bowie lab has pioneered a new method to dissect this important process without disrupting the membrane environment. Instead of altering the conditions using temperature, pressure, or chemical denaturants to unfold a membrane protein, the Bowie lab established a method that uses a secondary protein that preferentially binds to the unfolded state. The method, referred to as "steric trapping," employs the tight binding of a large tag-binding protein to a tagged protein of interest. (Figure 6-2) We have exploited the high affinity of a chimeric monovalent streptavidin (mSA) for biotin (Howarth et al., 2006). The protein subject to trapping is labeled such that the biotin tags are close in the folded structure, but far apart in the primary sequence. If the tags are sufficiently close, mSA can only bind one of the tags when the protein is folded because the second biotin

site will be occluded by the steric bulk of the bound streptavidin. Only once the protein unfolds does the second site become accessible for binding, upon which the protein becomes sterically trapped in the unfolded state, thereby coupling unfolding with a measureable binding event. The affinity for the second binding event is determined by the intrinsic mSA-biotin binding free energy and the free energy of unfolding of the protein. Unfolding is driven by mSA affinity and concentration, so the thermodynamics of protein folding can be calculated from a binding isotherm. Steric trapping does not necessitate destroying the membrane or disrupting its properties because the membrane protein of interest can be biotinylated in the soluble loops. Depending on the placement of the affinity tags, the steric trap method can target specific domains of the protein for unfolding.

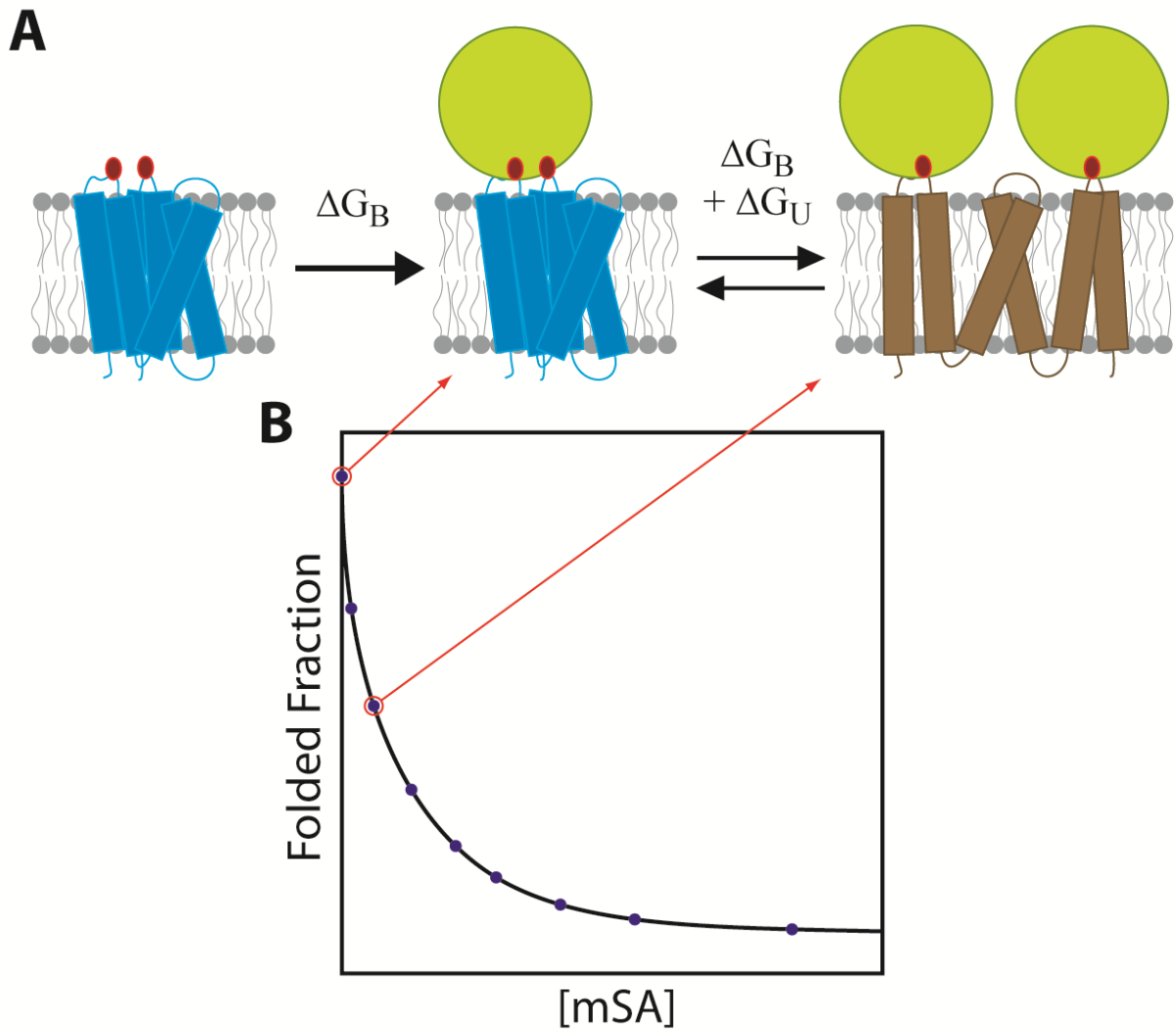


Figure 6-2 Steric Trap Unfolding. **(A)** Reaction diagram of steric trapping. The membrane protein (blue) is site-specifically labeled with biotin (red). The first mSA (green) binding event is driven by the intrinsic affinity for biotin (ΔG_B), but the second binding energy is coupled to the unfolding free energy of the biotin-labeled membrane protein ($\Delta G_B + \Delta G_U$). The color change to brown denotes the unfolded state. **(B)** Idealized binding curve showing the fraction of folded protein over a range of mSA concentrations.

As a proof of concept, this method has been used to study the soluble enzyme dihydrofolate reductase (DHFR) (Blois et al., 2009). A double cysteine mutant of DHFR was labeled with thiol-reactive biotin at two positions known to be close together and solvent accessible from the atomic structure. Activity loss depended on biotinylation of both sites and

correlated with increasing concentrations of mSA. The sterically trapped unfolded state was also probed by limited proteolysis, demonstrating that sterically trapped DHFR is similarly susceptible to proteolysis as chemically denatured proteins. The binding of wild-type mSA is essentially irreversible, so the reverse refolding reaction necessitated a mutant of the active subunit with weaker affinity and a greatly accelerated off-rate. Activity recovery after addition of excess free biotin demonstrated that steric trap unfolding was reversible. The apparent msA binding affinity was dependent on binding of a stabilizing cofactor to DHFR, indicating that the steric trap is able to measure changes in protein stability. The unfolding free energy measured by steric trapping agrees closely with that measured by urea denaturation, establishing the steric trap method as a reliable tool for measuring protein stability.

The first test of steric trapping with a membrane protein examined the dimerization of the single-span TM protein glycophorin A (GpATM) (Hong et al., 2010). The association energy of GpATM had been studied extensively in micellar systems, but the equilibrium in lipid vesicles was unclear. GpATM is an important model protein for studying transmembrane helix oligomerization and the association of helices during membrane protein folding. To monitor the association of the dimer, self-quenching pyrene fluorophores were attached near the TM close to the biotin tag. While the dissociation constant determined in detergent agrees well with previous measurements, the association of the GpATM is greatly enhanced in 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) bilayers. The high affinity of the GpATM dimer in POPC bilayers would not have been measureable by dilution. Steric trapping offers a method of accurately measuring helix association energetics in lipid bilayers.

The GpATM dimer was further explored by altering the lipid bilayer composition (Hong and Bowie, 2011). One of the primary advantages of the steric trap method is the ability to study the effect of different lipids on membrane protein folding. While the GpATM is highly stabilized in POPC vesicles, the recreation of a natural membrane environment revealed that the

association is only marginally stable in the cellular context. The dimerization is greatly weakened by electrostatic interactions between lipid headgroups and protein side chains as well as competition from the heterogeneous milieu of membrane proteins. These results illustrated that membrane proteins are not necessarily always more stable in membranes than in detergent. These unexpected effects of the membrane environment can be uncovered using the steric trap method.

Attempting to measure the thermodynamic stability of a more complex membrane protein lead to a surprising finding with DGK (Jefferson et al., 2013). DGK was biotinylated at a single cysteine on each subunit near the axis of symmetry. The subunits must dissociate to be susceptible to steric trapping. Dissociation disrupts the active sites shared between subunits, so unfolding can be monitored by loss of activity. While attempting to establish steric trapped DGK for thermodynamic measurements, the unfolding rate was found to be extremely slow in octyl-glucoside with a half-life of 12.6 d. Without any SDS, unfolding is indistinguishable from the intrinsic irreversible inactivation rate. The unfolding rate in a small amount of SDS could be separated from the slow inactivation demonstrating that the protein is indeed trapped by mSA. The steric trapped monomers were confirmed to be dissociated as they could be refolded with SDS-unfolded mutant subunits. The steric trapping of the DGK trimer does not necessitate complete unfolding of each subunit, and thus the complete unfolding of the protein may be even slower.

Interestingly, the refolding rate of steric trapped DGK is much slower than refolding with SDS-unfolded subunits, suggesting that unfolding pathways under native conditions may be completely different from those generated with SDS. The steric trapped unfolded state is also likely different from the SDS-unfolded state. In SDS, the transmembrane helices can be kept in an extended conformation due to the repulsion of the anionic micelles, but in the non-denaturing steric trapped state the helices are free to associate and misfold, forming

energetically frustrated helical bundles with slow folding and unfolding rates similar to soluble proteins (Wensley et al., 2010, 2012). The steric trapping of DGK, while demonstrating the method on a complex polytopic membrane protein, is also notable for finding surprisingly slow unfolding and refolding kinetics under non-denaturing conditions, not seen in previous studies with SDS.

The thermodynamic stability of a multi-span membrane protein was first measured under native conditions by steric trapping of bR (Chang and Bowie, 2014). bR is a convenient model protein for its measureable absorbance in the folded state. These experiments challenged the validity of the long linear extrapolations made from SDS unfolding measurements. The unfolding free energies measured in low χ_{SDS} conditions followed a non-linear trend and the unfolding free energy at zero SDS is ~11 kcal/mol, far less than the 26 kcal/mol expected from the linear extrapolation of SDS unfolding. The curve of unfolding free energies derived from steric trapping appear to converge with those from previous SDS titration results around 0.45 χ_{SDS} suggesting that the two methods are measuring a similar unfolding process. The non-disruptive nature of the steric trap method permits the adjustment of the lipid:detergent ratio to change the local environment of the bicelle. bR is modestly stabilized in larger lipidic bicelles, but not nearly to the degree suggested by linear extrapolation.

Later steric trap experiments on the transmembrane domain of the rhomboid protease GlpG in n-dodecyl β-D-maltoside (DDM) corroborated the non-linear dependence of thermodynamic stability on SDS mole fraction (Guo et al., 2016). While two proteins is a small set, the non-linear relationship is not unique to either membrane protein. In contrast to the unfolding free energies determined from steric trapping of bR, those for GlpG do not converge with SDS titration results. Instead, steric trap unfolding free energies are nearly 4 kcal/mol greater than those determined by SDS unfolding at a moderate SDS mole fraction. By combining the steric trap method with dependence on SDS mole fraction, as done with bR

(Chang and Bowie, 2014), m-values for two biotin pairs can be compared. The authors suggest that the steeper m-value from steric trapping of the N-terminal half of GlpG exposes more buried residues upon unfolding, whereas the C-terminal half undergoes subglobal unfolding around the active site. While the interpretation of m-values for SDS denaturation are still up for debate, these experiments demonstrate the utility of combining traditional harsh denaturation with newer methods for studying membrane protein folding under native conditions.

The Hong lab improved upon the initial biotinylation labeling of proteins for steric trapping by adding fluorescent and spin label reporter groups to a customized biotin label. Double electron-electron resonance (DEER) measurements of the distance between labels in the steric trapped state provide insight into the nature of the unfolded state. The DEER measurements indicate that non-denatured steric trapped GlpG is not a compact unfolded state. Steric trapped unfolded GlpG is even more expanded between the labels than the SDS unfolded state, perhaps attributing the different unfolding free energies to different unfolded states. The authors note that the increased distances are probably not soley due to steric repulsion because alternate biotin pairs with similar C_α-C_α distances could be doubly bound with monovalent streptavidin and still retain activity.

Additionally, the alternate biotin-pyrene label offers a generally applicable steric trap method that does not rely on functional assays tailored to the protein of interest. By attaching the quencher DABCYL to the monovalent streptavidin, binding can be measured readily and sensitively. Fluorescence quenching measurements were confirmed with activity loss from the same variant of monovalent streptavidin. The fluorescence measurements are well-suited to high-throughput experiments, which Hong and colleagues took advantage of by examining the effects of a large host of destabilizing mutants between two separate domains of the protein. Provided suitable pairs of biotin labels can be placed throughout the membrane protein, steric trapping can dissect cooperative interactions and subglobal unfolding within the full-length

protein under native conditions. Perhaps the most powerful aspect of the steric trapping of GlpG is the promise that these additions and refinements to the steric trap method hold for the future of membrane protein folding studies.

Single-Molecule Force Spectroscopy

Perhaps the most detailed examination of the energy landscape is to completely pull apart single membrane proteins by mechanical force. Several methods of exerting mechanical force on polypeptide chains have come to prominence in recent years. These methods rely on anchoring of the protein to a surface and then pulling one end of the protein away from the surface and measuring the extended distance and force applied to the protein.

AFM-mediated unfolding

Early experiments demonstrated that macromolecules could examined by force spectroscopy using a repurposed atomic force microscope (Rief et al., 1997, 1999), and the technique was soon applied to pulling a membrane protein out of the bilayer (Oesterhelt, 2000). Single-molecule force spectroscopy utilizes the AFM tip to pull on a single molecule of interest found in the atomic force micrograph. In the case of bacteriorhodopsin, purple membrane was adsorbed onto a flat surface and the AFM tip was adsorbed to a single protein at the C-terminus. The stylus is then lifted off the surface of the membrane with increasing force. Following the extension of the polypeptide chain with increasing force produces a force extension curve (Figure 6-3 A) that reveals specific intermediates that correspond to sequential segments of the polypeptide chain pulled up out of the membrane. Many attempts either fail to adsorb to the protein or attach to a non-terminal cytoplasmic loop of bacteriorhodopsin. To compare a collection of equivalent pulling events, only spectra that extended to the full length of

bacteriorhodopsin before at the final rupture peak were analyzed. Additionally, the high-resolution imaging of AFM can be employed to confirm the vacancy of the pulled protein. The force peaks fit to a worm-like chain model for pairs of helices pulled out of the membrane from the C-terminus. The amount of force required to unfold each segment decreases sequentially as the removal of helices further destabilizes the remaining membrane-embedded segment. Comparisons could be made between the relative stability of the transmembrane helix pairs, however, the mode of unfolding orthogonally to the membrane is not a biologically relevant process, fails to isolate second-stage unfolding, and is irreversible.

While AFM force spectroscopy experiments demonstrated that force could be used to unfold membrane proteins, it was still untested whether α-helical transmembrane segments could refold into the lipid bilayer. Further studies by the Gaub and Müller labs were able to reinsert two different α-helical membrane proteins after forced unfolding out of the bilayer (Kedrov et al., 2004; Kessler et al., 2006). The proteins were extended to a length that stopped before pulling the final segment before final rupture, so that the polypeptide chain was still anchored to the membrane. The AFM tip was then lowered back to the bilayer surface. Complete reinsertion of the full-length protein was confirmed by repeated unfolding which displayed the same set of force peaks in the initial unfolding, demonstrating the robustness of the force spectroscopy method. In the case of the sodium antiporter NhaA, the refolding trace did not exhibit the same force peaks for each pair of helices, except for one pair that exerted a small force on the stylus while refolding into the membrane (Kedrov et al., 2004). In contrast, refolding of bacteriorhodopsin displayed “snap-in” force peaks for two pairs of refolding helices (Kessler et al., 2006). Integrating these snap-in peaks provides a measure of the work performed by the refolding protein against the AFM cantilever. However, this attempt to quantify the free energy of refolding helices is convoluted by the fact that the speed of the cantilever back towards the surface may be faster than the time for refolding, which can

suppress refolding peaks (Kedrov et al., 2004), and it is unclear whether the helices are undergoing a first-stage insertion or making tertiary contacts as well. Indeed, sometimes that unfolding rupture forces were lower after refolding, suggesting that tertiary refolding does not regularly occur.

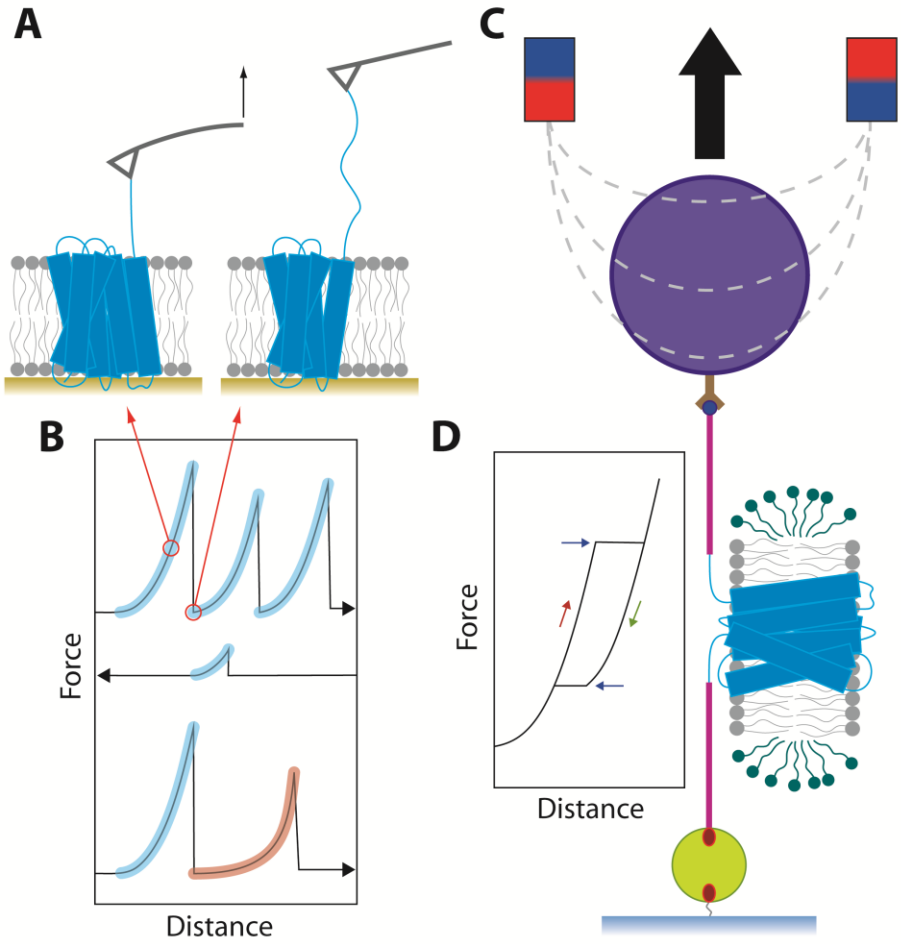


Figure 6-3 Single-molecule Force Spectroscopy for Membrane Protein Unfolding. **(A)** Scheme of forced unfolding using atomic force microscopy. The AFM tip (gray triangle) is attached to a terminus of the folded membrane protein (blue) in a supported bilayer. The AFM cantilever is moved upwards while measuring the tension force of the polypeptide chain (left) until a segment ruptures out of the membrane (right). **(B)** Idealized force-distance curves of membrane protein unfolding, refolding, and repeated unfolding from an AFM experiment. Initial unfolding produces three force peaks. Refolding produces a single snap-in peak as refolding of the second-segment exerts a force on the AFM cantilever. Repeated unfolding confirms refolding of the first segment, but indicates misfolding of the membrane protein as a whole due to a force peak of a contour length not observed in the initial unfolding curve (red). **(C)** Scheme of forced unfolding using magnetic tweezers. A magnet exerts force on a magnetic bead (purple) tethered to a surface via DNA handles (pink) with an intervening membrane protein (blue). **(D)** Idealized overlay of unfolding and refolding force-distance curves from a magnetic tweezers experiment. As the magnet descends closer towards the bead, the increasing forces pull along the surface of the bilayer, stretching the DNA (red arrow) until an unfolding event produces a distance jump (top blue arrow). Force is subsequently decreased (green arrow) until a refolding event produces a distance contraction (bottom blue arrow).

One of the advantages of force spectroscopy is the ability to recapitulate native environments and study the effects of changes in lipid composition. LacY, having been unfolded by force spectroscopy (Serdiuk et al., 2014) and known to exhibit an inverted topology in the absence of phosphatidylethanolamine (Bogdanov et al., 2008), makes it a good candidate for studying in different lipid compositions. AFM force spectroscopy helped pinpoint where in the structure the topology was most effected by the change in the lipid bilayer environment (Serdiuk et al., 2015). In 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylethanolamine (POPE): 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylglycerol (POPG) membranes, LacY unfolding force extension curves exhibit predominantly 10 force peaks, and in rare cases, a new peak fits to a contour length in the middle of the protein. The unfolding segments around central helices VI, VII, VIII, and the loops between them are disrupted. The fraction of LacY that unfolds in this second force extension pattern increases in pure POPG membranes. The force spectroscopy experiments confirm the previous model of LacY inverted topology that leaves the N and C-terminal halves of the protein intact, except for helix VII, which is exposed to the periplasm, demonstrating the utility of force spectroscopy in zeroing in on specific segments of membrane proteins that are structurally altered in different membrane environments without the need for denaturants.

Further studies with LacY have studied the effect of chaperoned refolding (Serdiuk et al., 2016). While spontaneous folding into the membrane from an unfolded state outside the membrane is not the physiologically relevant process for α-helical membrane proteins, they do insert from the N-terminus with the aid of translocon machinery, such as YidC. In this case, AFM force spectroscopy is well suited to studying the chaperoning effects of a protein such as YidC. In the presence of YidC, LacY refolding is promoted such that repeated unfolding exhibits the same set of force peaks observed from the initial unfolding of native LacY. In the absence of YidC, LacY misfolding dominates over time, while refolding does not improve with time. While

these experiments lack a lot of the specialized cotranslational insertion apparatus for first stage folding of α-helical membrane proteins, they highlight the importance of chaperoning for the proper folding of this class of membrane proteins.

While this technique can be applied to α-helical membrane proteins, it is perhaps better suited to proteins that fold into the lipid bilayer in a single-stage process, such as the outer membrane protein β-barrels. The first forced unfolding of a β-barrel out of a membrane was performed with OmpG (Sapra et al., 2009). Similar to α-helical membrane proteins unfolded by single-molecule force spectroscopy, OmpG unfolds one pair of β-strands at a time. One of the advantages of single-molecule force spectroscopy is observing the effects of changes in the aqueous environment without disrupting the membrane. OmpG is a pH-gated pore and changes in the force extension curve under pH 7 and 5 have highlighted specific pH-dependent interactions that govern the pore loop (Damaghi et al., 2010).

Later studies with OmpA demonstrated the reversible folding with AFM force spectroscopy (Bosshart et al., 2012). OmpA has a C-terminal peptidoglycan-binding domain, to which the AFM tip can be non-specifically attached. OmpA unfolds through several intermediates and refolds in a single force peak as the cantilever is relaxed back to the membrane surface for ~2 s. In about 5 out of 200 refolding experiments, OmpA was successfully refolded as confirmed by repeated unfolding. Interestingly, this mode of folding and unfolding is particularly biologically relevant to OmpA as the periplasmic peptidoglycan-binding domain can exert force on the β-barrel to ultimately regulate the pore's function. Furthermore, the N-terminal β strand requires an unusually high amount of force to remove from the membrane, suggesting that it may serve as an anchor for OmpA refolding.

While refolding β-barrel membrane proteins from the unfolded aqueous state is possible for OmpA, assistance from chaperones is required for refolding of the larger barrel FhuA

(Thoma et al., 2012, 2015). Without chaperones, repeated unfolding of FhuA after allowing refolding produces new intermediates in the force extension curve, indicating misfolding or failure to properly insert the β strands. However, in the presence of Skp and SurA, FhuA can be refolded properly and reproduce the force extension curve upon repeated cycles of unfolding. Refolding in the presence of Skp and SurA highlights the advantage of force spectroscopy for studying membrane protein folding in a milieu of a defined lipid bilayer, native buffer, and accessory chaperones. Confirmation of refolding was confirmed as in previous experiments by checking repeated unfolding against the force peak "fingerprint" of the first unfolding. Without chaperones, FhuA tends to misfold, displaying force peaks at contour lengths outside the original unfolding fingerprint, or remains unfolds, showing no force peaks upon repeated unfolding. Skp and SurA both have mitigating effects on the fraction of misfolded protein in repeated refolding experiments, but by different means. Skp does not substantially promote refolding of the β-hairpins, but removes many of the misfolded FhuA. On the other hand, SurA encourages refolding and helps prevent misfolding. While the presence of both chaperones stabilizes the unfolding of FhuA, inidicating the dominance of Skp to prevent misfolding by the stabilization of unfolded FhuA. Because of the fine control over an individual protein in force spectroscopy, the effect of chaperones can also be examined for refolding from different contour lengths. The chaperones prevented misfolding equally for different lengths of the unfolded polypeptide chain. FhuA was typically allowed to refold for only 1 s after relaxing the AFM cantilever back to the membrane surface, but a study of extended refolding times showed that the fraction of folded FhuA increases, as well as increasing the force required to unfold the refolded β-hairpins to that of native FhuA, suggesting that FhuA needs longer than 1 s to fully stabilize its β-barrel to the native conformation. So while "refolded" classes of FhuA pulling experiments reflect a pairwise insertion of β strands, they often likely only reflect a partially folded state that is on-pathway to the native state.

The stepwise refolding of β-barrel outer membrane proteins by β-hairpins is suggested to be the *in vivo* mechanism. The Omp insertion machinery is proposed to function by first interacting with the C-terminal β strand of an unfolded Omp (Gruss et al., 2013; Kim et al., 2012; Noinaj et al., 2013, 2014). Thus the individual refolding of subsequent β-hairpins in AFM force spectroscopy measurements may reflect a more native mechanism of insertion and folding than the tilting mechanism of β-barrel refolding from a urea denatured state (Huysmans et al., 2010). Single-molecule force spectroscopy still relies upon pulling membrane proteins out of the bilayer, preventing the observation of second-stage folding within the membrane environment. AFM refolding experiments also require part of the protein to remain in the membrane to act as a tether, so only the refolding of all but one or two transmembrane segments is observed.

Single-molecule tweezers for second-stage folding

In contrast to force spectroscopy via AFM, optical and magnetic tweezers offer a method to pull membrane proteins along, instead of orthogonal to, the plane of the membrane (Figure 6-3 B). Seminal work from the Marqusee and Bustamante labs demonstrated that single proteins could be force unfolded to examine stability and intermediates in folding pathways (Cecconi et al., 2005). Optical and magnetic tweezers exert force on a bead that is tethered to a surface via DNA handles and a single protein of interest. *E. coli* GlpG was the first protein to be studied by single-molecule tweezers (Min et al., 2015). GlpG is a six transmembrane segment rhomboid protease that has been studied by SDS denaturation (Baker and Urban, 2012; Guo et al., 2016), and now also by steric trapping (Guo et al., 2016), making it a good point of comparison for protein folding studies. Attachment of the DNA handles at the N and C termini of the transmembrane domain of GlpG places the mechanical unfolding force along the surface of the membrane. Cycles of unfolding and refolding in 1,2-dimyristoyl-sn-glycero-3-

phosphocholine (DMPC): 3-[(3-Cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate (CHAPSO) bicelles exhibits a large hysteresis in the force extension curve. The repeated extensions confirmed that GlpG was completely refolded at low force and the extended polypeptide chain formed helical structure before refolding, indicating that the observed refolding is reflective of transmembrane segment association within the context of a lipid bilayer. Although GlpG tended to completely unfold in a single force jump at ~25 pN, there were rare cases where unfolding passed through intermediate states. One of the advantages of magnetic tweezers is the ability to exert a constant magnetic force on the protein. In a "force-jump" experiment the force was quickly brought to 21 pN and held until GlpG completely unfolded. Though the pauses were brief relative to the total unfolding time, the partial unfolding to two distinct intermediate states was more clearly measured when not unfolding with increasing force. The extension distance of the intermediates corresponds to segment lengths that are roughly pairs of helices, however, in constrast to the directional pulling of AFM force spectroscopy, it was unknown which end of the protein is unfolded first. To answer this question, destabilizing mutations were made in either end of the protein. In the case of the N-terminal mutation L155A, the transition from the second intermediate to the complete unfolding was accelerated. Correspondingly, the C-terminal mutation A206G accelerated unfolding to the first intermediate. Therefore GlpG unfolds directionally from the C- to N- terminus under mechanical force.

To map out the energy landscape of GlpG folding, the folding and unfolding energy barriers were calculated by measuring the folding and unfolding probabilities for a range of forces. The unfolding events occurred stochastically at different force levels, so unfolding probabilities were calculated from hundreds of repeated unfolding events. The unfolding rate at zero tension was extrapolated from the probability curve. Refolding at low forces (below ~8 pN) was confirmed by unfolding GlpG after 3 min of refolding time. Extrapolating the refolded

fraction yielded the folding rate. The free energy of unfolding calculated from the unfolding and folding rates at zero force is 6.54 k_BT.

Though GlpG ultimately unfolds to a fully extended polypeptide chain, the energy landscape is indicative of physiological second-stage folding in the membrane. Though the structure and geometry of the protein-bicelle complex as it unfolds is uncertain, the hysteresis of the unfolding and refolding cycle is assumed to represent the same pathway, as the folding and unfolding kinetics, measured only 6 pN apart, do not exhibit any discrepancies that would imply altered pathways. The unfolding jump is initiated in the bicelle environment and helical transmembrane segments are formed prior to the refolding jump, showing that the transitions occur primarily in the protein-bicelle complex. Also, the measured free energy of unfolding is fairly consistent with recent steric trapping of GlpG in DDM that measured 9.8 k_BT and 7.9 k_BT for the N and C domains (Guo et al., 2016), and corroborates the overestimation of linear extrapolations from SDS unfolding of the same construct.

Further optimization of the magnetic tweezer method was made by simplifying the attachment chemistry. Compared to the nonspecific attachment of the AFM tip, tweezers require precise handle attachment at specific cysteine residues. The labeling of cysteines can be inefficient and requires the removal of other native cysteines, which can be technically prohibitive for large membrane proteins. The Bowie lab employed the SpyTag-SpyCatcher system to get around these challenges (Min et al., 2016). By adding SpyTag segments flanking the membrane protein of interest, DNA handles conjugated to an easy to prepare MBP-SpyCatcher fusion will bind after a short incubation, after which the protein is ready for attachment to the surface and magnetic bead for tweezing. Repeating GlpG pulling experiments with this new regime duplicated the previous results, demonstrating that the addition of MBP-SpyCatcher-SpyTag to the handles does not alter the observed unfolding of the membrane protein of interest. This updated method makes an already generally applicable method easier

to apply to more membrane proteins, and introduces the possibility of tweezing proteins in natural membranes without the need for purification due to the high affinity SpyTag-SpyCatcher complex.

Between equilibrium unfolding with urea or SDS, the steric trap method, and single-molecule force spectroscopy, a variety of methods for interrogating the membrane protein folding process are available. These methods balance technical simplicity with molecular detail. While these methods are compatible with a variety of detergent-aided membrane mimetics, newer technologies for studying membrane protein folding hold much promise for their general applicability and will likely pave the way towards understanding how proteins fold within the biological membrane.

References

Baker, R.P., and Urban, S. (2012). Architectural and thermodynamic principles underlying intramembrane protease function. Nat Chem Biol *8*, 759–768.

Bartolo, N.D., Compton, E.L.R., Warne, T., Edwards, P.C., Tate, C.G., Schertler, G.F.X., and Booth, P.J. (2016). Complete Reversible Refolding of a G-Protein Coupled Receptor on a Solid Support. PLOS ONE *11*, e0151582.

Blois, T.M., Hong, H., Kim, T.H., and Bowie, J.U. (2009). Protein Unfolding with a Steric Trap. J Am Chem Soc *131*, 13914–13915.

Bogdanov, M., Xie, J., Heacock, P., and Dowhan, W. (2008). To flip or not to flip: lipid-protein charge interactions are a determinant of final membrane protein topology. J. Cell Biol *182*, 925–935.

Booth, P.J., Flitsch, S.L., Stern, L.J., Greenhalgh, D.A., Kim, P.S., and Khorana, H.G. (1995). Intermediates in the folding of the membrane protein bacteriorhodopsin. Nature Structural & Molecular Biology *2*, 139–143.

Borgnia, M.J., Kozono, D., Calamita, G., Maloney, P.C., and Agre, P. (1999). Functional reconstitution and characterization of AqpZ, the E. coli water channel protein. Journal of Molecular Biology *291*, 1169–1179.

Bosshart, P.D., Iordanov, I., Garzon-Coral, C., Demange, P., Engel, A., Milon, A., and Müller, D.J. (2012). The Transmembrane Protein KpOmpA Anchoring the Outer Membrane of Klebsiella pneumoniae Unfolds and Refolds in Response to Tensile Load. Structure *20*, 121–127.

Braiman, M.S., Stern, L.J., Chao, B.H., and Khorana, H.G. (1987). Structure-function studies on bacteriorhodopsin. IV. Purification and renaturation of bacterio-opsin polypeptide expressed in Escherichia coli. J. Biol. Chem. *262*, 9271–9276.

Cao, Z., Schlebach, J.P., Park, C., and Bowie, J.U. (2012). Thermodynamic stability of bacteriorhodopsin mutants measured relative to the bacterioopsin unfolded state. Biochimica et Biophysica Acta (BBA) - Biomembranes *1818*, 1049–1054.

Cecconi, C., Shank, E.A., Bustamante, C., and Marqusee, S. (2005). Direct Observation of the Three-State Folding of a Single Protein Molecule. Science *309*, 2057–2060.

Chang, Y.-C., and Bowie, J.U. (2014). Measuring membrane protein stability under native conditions. Proc. Natl. Acad. Sci. U.S.A. *111*, 219–224.

Chen, G.Q., and Gouaux, E. (1999). Probing the Folding and Unfolding of Wild-Type and Mutant Forms of Bacteriorhodopsin in Micellar Solutions: Evaluation of Reversible Unfolding Conditions. Biochemistry *38*, 15380–15387.

Curnow, P., and Booth, P.J. (2007). Combined kinetic and thermodynamic analysis of α-helical membrane protein unfolding. Proceedings of the National Academy of Sciences *104*, 18970–18975.

Damaghi, M., Bippes, C., Köster, S., Yildiz, Ö., Mari, S.A., Kühlbrandt, W., and Muller, D.J. (2010). pH-Dependent Interactions Guide the Folding and Gate the Transmembrane Pore of the β-Barrel Membrane Protein OmpG. Journal of Molecular Biology *397*, 878–882.

Faham, S., Yang, D., Bare, E., Yohannan, S., Whitelegge, J.P., and Bowie, J.U. (2004). Side-chain Contributions to Membrane Protein Structure and Stability. Journal of Molecular Biology *335*, 297–305.

Faham, S., Boulting, G.L., Massey, E.A., Yohannan, S., Yang, D., and Bowie, J.U. (2005). Crystallization of bacteriorhodopsin from bicelle formulations at room temperature. Protein Science *14*, 836–840.

Findlay, H.E., Rutherford, N.G., Henderson, P.J.F., and Booth, P.J. (2010). Unfolding free energy of a two-domain transmembrane sugar transport protein. PNAS *107*, 18451–18456.

Fujii, J., Maruyama, K., Tada, M., and MacLennan, D.H. (1989). Expression and site-specific mutagenesis of phospholamban. Studies of residues involved in phosphorylation and pentamer formation. J. Biol. Chem. *264*, 12950–12955.

Gruss, F., Zähringer, F., Jakob, R.P., Burmann, B.M., Hiller, S., and Maier, T. (2013). The structural basis of autotransporter translocation by TamA. Nat Struct Mol Biol *20*, 1318–1320.

Guo, R., Gaffney, K., Yang, Z., Kim, M., Sungsuwan, S., Huang, X., Hubbell, W.L., and Hong, H. (2016). Steric trapping reveals a cooperativity network in the intramembrane protease GlpG. Nat Chem Biol *12*, 353–360.

Hardie, K.R., Lory, S., and Pugsley, A.P. (1996). Insertion of an outer membrane protein in Escherichia coli requires a chaperone-like protein. EMBO J *15*, 978–988.

Heginbotham, L., Odessey, E., and Miller, C. (1997). Tetrameric Stoichiometry of a Prokaryotic K+ Channel†. Biochemistry *36*, 10335–10342.

Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S.H., and von Heijne, G. (2005). Recognition of transmembrane helices by the endoplasmic reticulum translocon. Nature *433*, 377–381.

Hong, H., and Bowie, J.U. (2011). Dramatic Destabilization of Transmembrane Helix Interactions by Features of Natural Membrane Environments. J. Am. Chem. Soc. *133*, 11389–11398.

Hong, H., and Tamm, L.K. (2004). Elastic coupling of integral membrane protein stability to lipid bilayer forces. PNAS *101*, 4065–4070.

Hong, H., Blois, T.M., Cao, Z., and Bowie, J.U. (2010). Method to measure strong protein–protein interactions in lipid bilayers using a steric trap. Proceedings of the National Academy of Sciences *107*, 19802–19807.

Howarth, M., Chinnapen, D.J.-F., Gerrow, K., Dorrestein, P.C., Grandy, M.R., Kelleher, N.L., El-Husseini, A., and Ting, A.Y. (2006). A monovalent streptavidin with a single femtomolar biotin binding site. Nature Methods *3*, 267–273.

Huang, K.S., Bayley, H., Liao, M.J., London, E., and Khorana, H.G. (1981). Refolding of an integral membrane protein. Denaturation, renaturation, and reconstitution of intact bacteriorhodopsin and two proteolytic fragments. J. Biol. Chem. *256*, 3802–3809.

Hunt, J.F., Earnest, T.N., Bousché, O., Kalghatgi, K., Reilly, K., Horváth, C., Rothschild, K.J., and Engelman, D.M. (1997). A Biophysical Study of Integral Membrane Protein Folding. Biochemistry *36*, 15156–15176.

Huyghues-Despointes, B.M.P., Scholtz, J.M., and Pace, C.N. (1999). Protein conformational stabilities can be determined from hydrogen exchange rates. Nature Structural Biology *6*, 910–912.

Huysmans, G.H.M., Baldwin, S.A., Brockwell, D.J., and Radford, S.E. (2010). The transition state for folding of an outer membrane protein. PNAS *107*, 4099–4104.

Jackson, S.E. (1998). How do small single-domain proteins fold? Folding and Design *3*, R81–R91.

Jefferson, R.E., Blois, T.M., and Bowie, J.U. (2013). Membrane Proteins Can Have High Kinetic Stability. J. Am. Chem. Soc. *135*, 15183–15190.

Johnson, C.M., and Fersht, A.R. (1995). Protein Stability as a Function of Denaturant Concentration: The Thermal Stability of Barnase in the Presence of Urea. Biochemistry *34*, 6795–6804.

Kedrov, A., Ziegler, C., Janovjak, H., Kühlbrandt, W., and Müller, D.J. (2004). Controlled Unfolding and Refolding of a Single Sodium-proton Antiporter using Atomic Force Microscopy. Journal of Molecular Biology *340*, 1143–1152.

Kessler, M., Gottschalk, K.E., Janovjak, H., Muller, D.J., and Gaub, H.E. (2006). Bacteriorhodopsin Folds into the Membrane against an External Force. Journal of Molecular Biology *357*, 644–654.

Kim, K.H., Aulakh, S., and Paetzel, M. (2012). The bacterial outer membrane β-barrel assembly machinery. Protein Science *21*, 751–768.

Kleinschmidt, J.H., Wiener, M.C., and Tamm, L.K. (1999). Outer membrane protein A of E. coli folds into detergent micelles, but not in the presence of monomeric detergent. Protein Science *8*, 2065–2071.

Knowles, T.J., Scott-Tucker, A., Overduin, M., and Henderson, I.R. (2009). Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. Nat Rev Micro *7*, 206–214.

Krishnamani, V., Hegde, B.G., Langen, R., and Lanyi, J.K. (2012). Secondary and Tertiary Structure of Bacteriorhodopsin in the SDS Denatured State. Biochemistry *51*, 1051–1060.

Lau, F.W., and Bowie, J.U. (1997). A Method for Assessing the Stability of a Membrane Protein†. Biochemistry *36*, 5884–5892.

London, E., and Khorana, H.G. (1982). Denaturation and renaturation of bacteriorhodopsin in detergents and lipid-detergent mixtures. J. Biol. Chem. *257*, 7003–7011.

Miller, D., Charalambous, K., Rotem, D., Schuldiner, S., Curnow, P., and Booth, P.J. (2009). In vitro Unfolding and Refolding of the Small Multidrug Transporter EmrE. Journal of Molecular Biology *393*, 815–832.

Min, D., Jefferson, R.E., Bowie, J.U., and Yoon, T.-Y. (2015). Mapping the energy landscape for second-stage folding of a single membrane protein. Nat Chem Biol *11*, 981–987.

Min, D., Arbing, M.A., Jefferson, R.E., and Bowie, J.U. (2016). A simple DNA handle attachment method for single molecule mechanical manipulation experiments. Protein Science *25*, 1535–1544.

Moon, C.P., and Fleming, K.G. (2011). Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. PNAS *108*, 10174–10177.

Moon, C.P., Zaccai, N.R., Fleming, P.J., Gessmann, D., and Fleming, K.G. (2013). Membrane protein thermodynamic stability may serve as the energy sink for sorting in the periplasm. PNAS *110*, 4285–4290.

Noinaj, N., Kuszak, A.J., Gumbart, J.C., Lukacik, P., Chang, H., Easley, N.C., Lithgow, T., and Buchanan, S.K. (2013). Structural insight into the biogenesis of β-barrel membrane proteins. Nature *501*, 385–390.

Noinaj, N., Kuszak, A.J., Balusek, C., Gumbart, J.C., and Buchanan, S.K. (2014). Lateral Opening and Exit Pore Formation Are Required for BamA Function. Structure *22*, 1055–1062.

Oesterhelt, F. (2000). Unfolding Pathways of Individual Bacteriorhodopsins. Science *288*, 143–146.

Otzen, D.E. (2003). Folding of DsbB in Mixed Micelles: A Kinetic Analysis of the Stability of a Bacterial Membrane Protein. Journal of Molecular Biology *330*, 641–649.

Pace, C.N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. Meth. Enzymol. *131*, 266–280.

Paschen, S.A., Waizenegger, T., Stan, T., Preuss, M., Cyrklaff, M., Hell, K., Rapaport, D., and Neupert, W. (2003). Evolutionary conservation of biogenesis of β-barrel membrane proteins. Nature *426*, 862–866.

Popot, J.L., and Engelman, D.M. (1990). Membrane protein folding and oligomerization: the two-stage model. Biochemistry *29*, 4031–4037.

Rief, M., Oesterhelt, F., Heymann, B., and Gaub, H.E. (1997). Single Molecule Force Spectroscopy on Polysaccharides by Atomic Force Microscopy. Science *275*, 1295–1297.

Rief, M., Pascual, J., Saraste, M., and Gaub, H.E. (1999). Single molecule force spectroscopy of spectrin repeats: low unfolding forces in helix bundles1. Journal of Molecular Biology *286*, 553–561.

Rigaud, J.-L., Pitard, B., and Levy, D. (1995). Reconstitution of membrane proteins into liposomes: application to energy-transducing membrane proteins. Biochimica et Biophysica Acta (BBA) - Bioenergetics *1231*, 223–246.

Sackmann, E. (1996). Supported Membranes: Scientific and Practical Applications. Science *271*, 43–48.

Sanders, C.R., and Prosser, R.S. (1998). Bicelles: a model membrane system for all seasons? Structure *6*, 1227–1234.

Sapra, K.T., Damaghi, M., Köster, S., Yildiz, Ö., Kühlbrandt, W., and Muller, D.J. (2009). One β Hairpin after the Other: Exploring Mechanical Unfolding Pathways of the Transmembrane β-Barrel Protein OmpG. Angewandte Chemie International Edition *48*, 8306–8308.

Sargiacomo, M., Scherer, P.E., Tang, Z., Kübler, E., Song, K.S., Sanders, M.C., and Lisanti, M.P. (1995). Oligomeric structure of caveolin: implications for caveolae membrane organization. PNAS *92*, 9407–9411.

Serdiuk, T., Madej, M.G., Sugihara, J., Kawamura, S., Mari, S.A., Kaback, H.R., and Müller, D.J. (2014). Substrate-induced changes in the structural properties of LacY. PNAS *111*, E1571–E1580.

Serdiuk, T., Sugihara, J., Mari, S.A., Kaback, H.R., and Müller, D.J. (2015). Observing a Lipid-Dependent Alteration in Single Lactose Permeases. Structure *23*, 754–761.

Serdiuk, T., Balasubramaniam, D., Sugihara, J., Mari, S.A., Kaback, H.R., and Müller, D.J. (2016). YidC assists the stepwise and stochastic folding of membrane proteins. Nat Chem Biol *advance online publication*.

Spelbrink, R.E.J., Kolkman, A., Slijper, M., Killian, J.A., and Kruijff, B. de (2005). Detection and Identification of Stable Oligomeric Protein Complexes in Escherichi coli Inner Membranes A PROTEOMICS APPROACH. J. Biol. Chem. *280*, 28742–28748.

Surrey, T., and Jähnig, F. (1992). Refolding and oriented insertion of a membrane protein into a lipid bilayer. PNAS *89*, 7457–7461.

Surrey, T., Schmid, A., and Jähnig, F. (1996). Folding and Membrane Insertion of the Trimeric beta-barrel protein OmpF. Biochemistry *35*, 2283–2288.

Thoma, J., Bosshart, P., Pfreundschuh, M., and Müller, D.J. (2012). Out but Not In: The Large Transmembrane β-Barrel Protein FhuA Unfolds but Cannot Refold via β-Hairpins. Structure *20*, 2185–2190.

Thoma, J., Burmann, B.M., Hiller, S., and Müller, D.J. (2015). Impact of holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins. Nat Struct Mol Biol *22*, 795–802.

Wensley, B.G., Batey, S., Bone, F.A.C., Chan, Z.M., Tumelty, N.R., Steward, A., Kwa, L.G., Borgia, A., and Clarke, J. (2010). Experimental evidence for a frustrated energy landscape in a three-helix-bundle protein family. Nature *463*, 685–688.

Wensley, B.G., Kwa, L.G., Shammas, S.L., Rogers, J.M., Browning, S., Yang, Z., and Clarke, J. (2012). Separating the effects of internal friction and transition state energy to explain the slow, frustrated folding of spectrin domains. PNAS *109*, 17795–17799.

Yang, B., Gonzalez, L., Prekeris, R., Steegmaier, M., Advani, R.J., and Scheller, R.H. (1999). SNARE Interactions Are Not Selective IMPLICATIONS FOR MEMBRANE FUSION SPECIFICITY. J. Biol. Chem. *274*, 5649–5653.

Zhou, Y., Lau, F.W., Nauli, S., Yang, D., and Bowie, J.U. (2001). Inactivation mechanism of the membrane protein diacylglycerol kinase in detergent solution. Protein Science *10*, 378–383.