Improved free-energy landscape reconstruction of bacteriorhodopsin highlights local variations in unfolding energy

Patrick R. Heenan,[1,2] Hao Yu,[1] Matthew G. W. Siewny,[1,2] and Thomas T. Perkins[1,3,a)]
[1]*JILA, National Institute of Standard and Technology and University of Colorado, Boulder, Colorado 80309, USA*
[2]*Department of Physics, University of Colorado, Boulder, Colorado 80309, USA*
[3]*Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado 80309, USA*

(Received 12 October 2017; accepted 7 December 2017; published online 29 December 2017)

Precisely quantifying the energetics that drive the folding of membrane proteins into a lipid bilayer remains challenging. More than 15 years ago, atomic force microscopy (AFM) emerged as a powerful tool to mechanically extract individual membrane proteins from a lipid bilayer. Concurrently, fluctuation theorems, such as the Jarzynski equality, were applied to deduce equilibrium free energies (ΔG_0) from non-equilibrium single-molecule force spectroscopy records. The combination of these two advances in single-molecule studies deduced the free-energy of the model membrane protein bacteriorhodopsin in its native lipid bilayer. To elucidate this free-energy landscape at a higher resolution, we applied two recent developments. First, as an input to the reconstruction, we used force-extension curves acquired with a 100-fold higher time resolution and 10-fold higher force precision than traditional AFM studies of membrane proteins. Next, by using an inverse Weierstrass transform and the Jarzynski equality, we removed the free energy associated with the force probe and determined the molecular free-energy landscape of the molecule under study, bacteriorhodopsin. The resulting landscape yielded an average unfolding free energy per amino acid (aa) of 1.0 ± 0.1 kcal/mol, in agreement with past single-molecule studies. Moreover, on a smaller spatial scale, this high-resolution landscape also agreed with an equilibrium measurement of a particular three-aa transition in bacteriorhodopsin that yielded 2.7 kcal/mol/aa, an unexpectedly high value. Hence, while average unfolding ΔG_0 per aa is a useful metric, the derived high-resolution landscape details significant local variation from the mean. More generally, we demonstrated that, as anticipated, the inverse Weierstrass transform is an efficient means to reconstruct free-energy landscapes from AFM data. https://doi.org/10.1063/1.5009108

I. INTRODUCTION

Membrane proteins perform critical biological functions, such as light harvesting, signaling, and transport. To carry out these and other diverse functions, they fold into equilibrium structures within a lipid bilayer.[1] Yet, predicting the structure and stability of membrane proteins significantly lags behind the success for globular proteins.[1] For instance, models still struggle to predict the tilt and overall packing of transmembrane (TM) α-helices. Hence, there is an ongoing need to better quantify the molecular forces that drive and stabilize the folding of membrane proteins. As with globular proteins,[2] the energy landscape provides the fundamental framework for understanding membrane-protein folding and structure stabilization.[3]

Single-molecule force spectroscopy (SMFS) has emerged as an important tool in determining the energetics underlying the folding and unfolding of individual proteins.[4–7] For membrane proteins, atomic force microscopy (AFM) is the modality of choice[8] since individual proteins can be mechanically extracted from their native bilayer.[9] In the canonical assay, the tip of a cantilever is pressed into bacteriorhodopsin (BR) embedded in its native lipid bilayer to promote nonspecific attachment.[9] The tip is then retracted at a constant velocity, while force is deduced from the bending of the cantilever, revealing a series of unfolding peaks in the force-extension curve [Figs. 1(a) and 1(b)]. The dominant peaks correspond to pulling on the top of TM helix pairs or a terminal helix, notably the ED and CB helix pairs as well as the A helix. Importantly, these AFM-based studies provide insight into the energetics of membrane proteins in lipid bilayers, rather than in detergent. Moreover, AFM studies provide a reproducible reference state, the fully unfolded protein. In contrast, the "unfolded" protein in ensemble unfolding assays retains α-helical content and therefore provides a poor reference state.[1]

Complementing experimental advances that led to SMFS studies of individual proteins and nucleic-acid structures, theoretical advances enabled the determination of equilibrium free-energy values (ΔG_0) from analysis of non-equilibrium force-extension curves.[10] Such analysis provides important information beyond the distances to transition states ($\Delta x^{\ddagger}$) and the zero-force dissociation rate constant (k_{off}) determined using the Bell-Evans analysis of dynamic force spectroscopy data.[11,12] In particular, as demonstrated in pioneering studies of

a)Author to whom correspondence should be addressed: tperkins@jila.colorado.edu

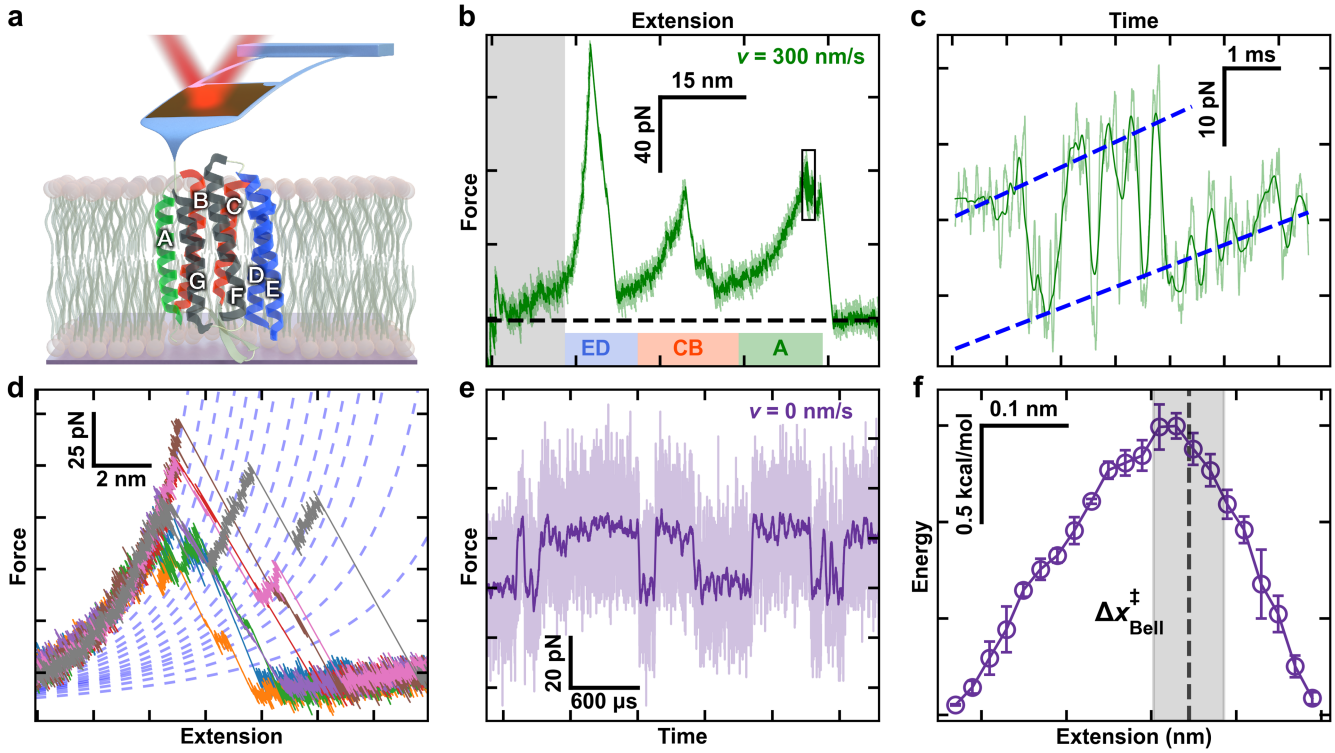

FIG. 1. High-resolution studies of bacteriorhodopsin (BR). (a) A cartoon illustrating BR being unfolded from its C-terminal end using a modified, ultrashort cantilever. Each helix is denoted by a letter. Color coding highlights the unfolding topology as helix pairs or the terminal A helix is extracted. (b) A canonical force-extension curve while stretching at $v = 300$ nm/s shows the major intermediates corresponding to pulling on the top of the E, C, and A helices. The initial portion of the force-extension curves was not analyzed due to the confounding effects of non-specific surface adhesion (grey box). The colored bars correspond to the colored helical regions from panel (a). (c) Force-vs-time record of the highlighted black box in panel (b) shows near-equilibrium unfolding and refolding at the top of helix A over a 13-amino-acid (aa) segment. (d) Multiple force-extension curves of BR show numerous detected unfolding intermediates. Curves were well modeled by a worm-like chain model within a particular state (dashed lines). (e) Force-vs-time record shows repeated unfolding and refolding of a 3-aa segment during an equilibrium assay ($v = 0$ nm/s) near the top of helix E. For clarity, 5-MHz data (light purple) were smoothed to 25 kHz (dark purple). (f) A reconstructed 1D free-energy landscape based on ~100-ms records of equilibrium data with 1-μs resolution tilted to $F_{1/2}$, the force at which the two states are equally likely to be occupied. Reconstruction was based on p_{fold}.[40] Error bars represent the standard error of the mean. Dashed line represents the location of transition state determined from a Bell analysis of state lifetime, with the grey box denoting the standard deviation in that localization. The data for Figs. 1(b)–1(f) are from Ref. 18.

structured RNA using optical tweezers,[13,14] a 1D free-energy landscape can be obtained by extending the Jarzynski equality[24] via the weighted histogram method developed by Hummer and Szabo.[10] Salient to this present work, Hummer and Szabo's original paper applied the weighted histogram method to deduce the free-energy landscape for the ED helix pair bacteriorhodopsin, yielding an average of ΔG_0 of 1.1 per amino acid (aa) estimated from eight force-extension curves. Using significantly more traces ($N \approx 600$), Preiner *et al.* extended this analysis over five helices (E–A) and reported an average ΔG_0 of 1.3 ± 0.2 kcal/mol/aa (mean ± SD).[15]

A higher-resolution description of the free-energy landscape enables a more quantitative comparison between the location and lifetimes of short-lived unfolding intermediates and the underlying energy landscape. Recent advances in applying modified ultrashort cantilevers optimized for 1-μs resolution SMFS[16,17] enabled the measurement of bacteriorhodopsin with a 100-fold higher time resolution and 10-fold higher force precision that, in turn, revealed rapid near-equilibrium folding [Fig. 1(c)] and a multitude of previously unresolved, transiently occupied intermediates [Fig. 1(d)].[18] These cantilevers also enabled the first equilibrium study of membrane protein folding, where a particular three-amino-acid transition near the top of the E helix repeatedly unfolded and refolded [Fig. 1(e)]. Reconstruction of the free-energy landscape underlying this transition [Fig. 1(f)] revealed an unexpectedly large ΔG_0 per aa (2.7 kcal/mol) at $F = 0$ pN. Here, we merged these experimental advances in acquiring high-resolution force-extension curves with theoretical advances in reconstructing the landscape from non-equilibrium records. In particular, we used the inverse Weierstrass transform (IWT) applied in conjunction with the Jarzynski free-energy integral to avoid convolving the compliance of the AFM cantilever into the resulting free-energy landscape.[19] Importantly, the IWT works best for studies using a stiffer force probe[19]—making it particularly suitable for AFM-based assays ($k \approx 6$–100 pN/nm)—due to their higher probe stiffness relative to earlier application in an optical-trapping-based assay ($k = 0.4$ pN/nm).[20] The resulting landscape yielded an average unfolding free energy per amino acid (aa) of 1.0 ± 0.1 kcal/mol (mean ± SD), in close agreement with past single-molecule studies.[10,15,21] Moreover, on a smaller spatial scale, this landscape also agreed with the recent equilibrium measurement that showed an unexpectedly high ΔG_0 per aa near the top of the E helix. Hence, while average unfolding ΔG_0 per aa is a useful metric, the resulting high-resolution

landscape confirmed significant local variations from the mean.

II. FREE-ENERGY LANDSCAPE RECONSTRUCTION

As background, the Jarzynski equality is remarkable because it relates the work done during many repetitions of a non-equilibrium process to the equilibrium free-energy difference as a function of an experimental parameter. For our work, the experimental parameter was the position of the cantilever base relative to the surface or "z."[22] More precisely, the Jarzynski equality is a thermodynamic relationship between the Helmholtz free energy of a system $A(z)$ and the measured work $W(z)$,

$$e^{-\beta A(z)} = \left\langle e^{-\beta W(z)} \right\rangle_N, \tag{1}$$

where z is the zero-force equilibrium position of the force probe, β is the inverse of the thermal energy $k_B T$, and the average is taken over N independent stretching trajectories, each starting and ending at the same choice of z.[10,23,24] The Jarzynski equality is exact only in the limit as $N \to \infty$ but is approximately true for finite N. In practice, SMFS experiments apply the Jarzynski equality by repeatedly folding or unfolding a single molecule of interest using a force probe. In this case, the work is the integral of the force as a function of the cantilever position z. For example, in the constant velocity experiments used in our AFM assay, $z(t) = vt + z_i$, and the integral for each stretching experiment starts at the same well-defined z_i associated with the aligned curves (see below). Tip-sample separation, which is also the molecular extension (x) in an AFM assay, is not uniquely defined at all values of z. The molecule may be in different folding states and therefore different molecular extensions, referred to henceforth as q as is common in the landscape-reconstruction literature.[19] The Jarzynski equality requires that all molecules included in the calculation start (z_i) and finish (z_f) in an equilibrium state.

Hummer and Szabo's original extension of the Jarzynski equality for SMFS studies determines the landscape as a function of molecular extension (q), instead of SMFS probe position (z), and also removes the energetic contribution of the force probe via a weighted histogram method.[10] In subsequent work, they applied an IWT to simplify this calculation and make more efficient use of a limited set of force-extension curves by assuming a stiff, harmonic pulling apparatus.[19,20] Put differently, the curvature (i.e., the second spatial derivative) of the true underlying energy landscape should not exceed the stiffness of the force probe. Indeed, the consequence of not fulfilling this assumption was theoretically anticipated[19] and experimentally demonstrated for a RNA pseudo-knot studied using an optical trap.[20] In that case, the derived landscape was filtered by the soft probe leading to a lower barrier height (akin to applying a spatial filter that smooths out sharp features). Hence, energy landscape reconstruction by the IWT is well suited to AFM-based assays due to the higher stiffness of AFM cantilevers in comparison to optical traps.[19,20]

As an input for IWT reconstruction of BR, we used a set of 106 force-extension curves of BR extracted from its native lipid bilayer (purple membrane for BR) acquired using focused-ion-beam modified BioLever Fast cantilevers (AC10DS, Olympus). The details of this data acquisition and the force-extension curves have been previously published.[18] Briefly, modified cantilevers ($k \approx 10$–40 pN/nm) were pressed into the purple membrane at ~1 nN for 1 s and retracted from the surface at a constant velocity (v) of 300 nm/s. Data were recorded at 5 MHz on a commercial AFM (Cypher ES, Asylum Research) and then smoothed with a second-order, 501-point Savitsky-Golay filter for all subsequent analyses. As is standard for SMFS-studies of BR,[8] the curves were laterally aligned to superimpose the major obligate intermediate associated with pulling on the top of the E helix. These aligned traces [Fig. 2(a)] then served as an input to calculate the free-energy landscape via the IWT. The unfolding of the GF helix pair was not analyzed due to the confounding effects of non-specific tip-sample adhesion.

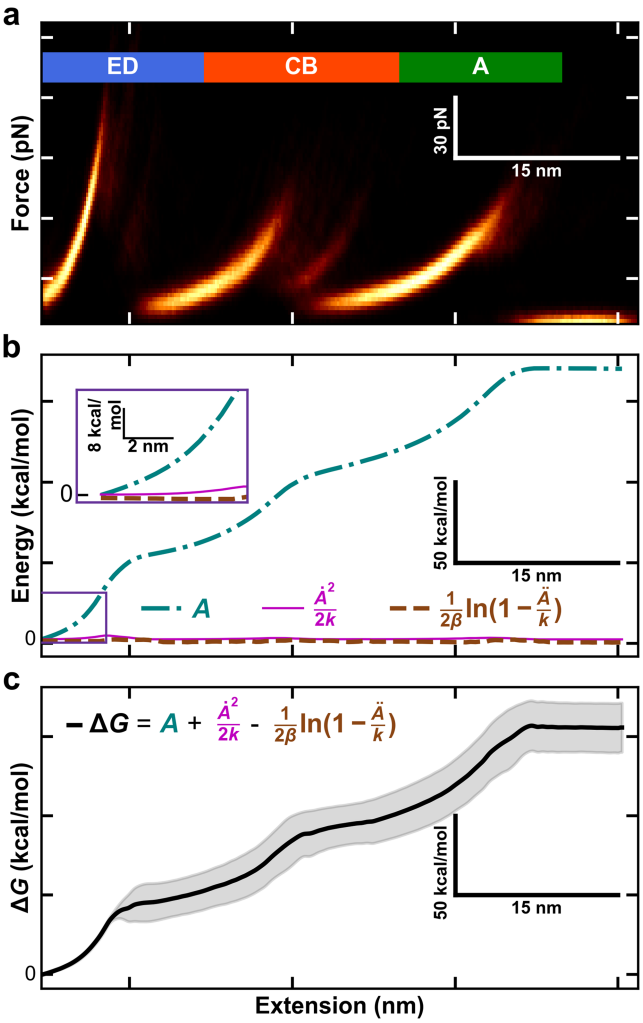

FIG. 2. Process for free-energy landscape reconstruction of BR using an inverse Weierstrass transform (IWT). (a) A heat map of all force-extension curves used in this work. Data within ~18 nm of the surface are excluded due to surface adhesion. (b) Different components for calculating the molecular free energy ΔG_0 as a function of the molecular extension when using the IWT. The three components used in computing ΔG_0 as defined in Eq. (5): A (aqua, dotted-dashed) and terms involving its first and second derivative (pink solid line and brown dashed line, respectively). To promote comparison to panels (a) and (c), the plotted lateral axis is q, molecule extension. (c) The ΔG_0 free energy as a function of molecular extension. Shaded region reports the standard deviation as deduced from a bootstrap analysis.

The IWT reconstructs the free-energy landscape, approximately removing the contribution of the force probe, by calculating statistics on an ensemble of N force-extension curves. For a particular curve n with measured force $F_n(z)$, the work used by the IWT is defined as

$$W_n(z) = \int_{z_i}^{z} F_n(z')dz', \tag{2}$$

where z runs from z_i to z_f at a given velocity v for all N force-extension curves. The Jarzynski equality determines the Helmholtz free energy $A(z)$, which includes energy from the probe. The IWT approximately removes the contribution of the probe by calculating the so-called work-weighted ensemble statistics. The work-weighted force and force squared, denoted with double brackets, are defined at each z as

$$\langle\langle F(z)\rangle\rangle = \frac{\left\langle e^{-\beta W_n(z)} F_n(z)\right\rangle_N}{e^{-\beta A(z)}}$$

and

$$\left\langle\left\langle F^2(z)\right\rangle\right\rangle = \frac{\left\langle e^{-\beta W_n(z)} {F_n}^2(z)\right\rangle_N}{e^{-\beta A(z)}}, \tag{3}$$

where $\langle X\rangle_N$ denotes an average of X over the ensemble of force-extension curves at a fixed z. These work-weighted values are related to the first and second spatial derivatives of the Helmholtz free energy as follows:

$$\dot{A}(z) = \langle\langle F(z)\rangle\rangle$$

and

$$1 - \frac{\ddot{A}(z)}{k} = \frac{\beta}{k}\left(\left\langle\left\langle F^2(z)\right\rangle\right\rangle - \langle\langle F(z)\rangle\rangle^2\right). \tag{4}$$

The IWT-based reconstruction uses the first and second derivatives to reconstruct the free-energy landscape as a function of molecular extension[19]

$$G\left(q = z - \frac{\dot{A}(z)}{k}\right) \approx A(z) - \frac{\dot{A}(z)^2}{2k} + \frac{1}{2\beta}\ln\left(1 - \frac{\ddot{A}(z)}{k}\right). \tag{5}$$

Importantly, $G(q)$ is approximately the free energy of the molecule, instead of the free energy of the molecule and the probe. To avoid errors associated with numerical derivatives of noisy data, we calculated the energy derivatives, $\dot{A}(z)$ and $\ddot{A}(z)$, directly from the primary work-weighted force and force squared (see Fig. S1 of the supplementary material) which, in turn, allowed $G(q)$ to be determined. To avoid errors in combining data sets acquired using different cantilever stiffnesses (k), we chose z_i = 18 nm + (20 pN/k), which corresponded to approximately F = 20 pN at an extension of 18 nm when pulling on the top of the ED helix pair. Each set of data, containing a minimum of 20 force-extension curves, corresponded to one cantilever and was reconstructed separately. After the transform was completed, the resulting free-energy landscapes were functions of molecular extension q (instead of z) and were averaged based on the number of traces associated with each cantilever data set. As in previous work,[15] we found that we needed a minimum of 20 traces per cantilever to limit reconstruction error to reasonable levels.

Uncertainties in the landscape and its first derivative were determined by a bootstrap analysis.[25] In brief, 250 synthetic data sets were generated from the original set of 106 curves using random resampling with replacement. Only synthetic data sets with $1 - \frac{\ddot{A}(z)}{k} > 0$, as required by the IWT, were counted towards the total of 250. From the resulting ensemble of 250 landscapes, the standard deviation in simulated mean values was calculated after smoothing with a third order, piecewise polynomial to 0.4 nm. This smoothing choice was chosen to approximately minimize the product of error in $\dot{G}(q)$ and in the least-squared fitting of $G(q)$ across the bootstrapping rounds (see Fig. S2 of the supplementary material). Future work with more traces from individual cantilevers and from cantilevers with a narrower range in k will decrease the uncertainties in landscape reconstruction.

We illustrate the process of calculating the IWT for BR from Eq. (5) in Fig. 2. The starting data consist of 106 aligned force-extension curves. As shown in Fig. 2(a), the resulting heat map represents the ensemble of measurements needed to apply Jarzynski's inequality and the inverse Weierstrass transform. To highlight the contributions of the three components in Eq. (5) in determining $G(q)$, we color coded a plot of $A(z)$ and the IWT corrections [Fig. 2(b)]. The corrections were obtained as described above, using work-weighted histogram statistics.[19,23] The resulting $G(q)$ was deduced from the sum of these three components and represents the free energy of the ED, CB, and A helical pairs of bacteriorhodopsin [Fig. 2(c)]. Shaded regions represent the standard deviation deduced from the aforementioned bootstrap analysis.

III. RESULTS AND DISCUSSION

The local unfolding free energy of BR depended on the individual structural element being unfolded. Our resulting high-resolution free-energy landscape of BR's final five helices showed significant variation in slope as a function of extension and therefore unfolding energy per amino acid [Fig. 2(c)]. To highlight this variation, we computed the free-energy landscape and ΔG_0 per nm as a function extension for the individual major structural components (ED, CB, and A helices) [Figs. 3(a)–3(c)]. The mean ΔG_0 per aa is 1.0 ± 0.1 kcal/mol (mean ± SD) when calculated from the unfolding of the ED, CB, and A helices [156 ± 15 kcal/mol (mean ± SD) over 160 amino acids of helices, traditionally taken as the top of the E-helix in AFM-based assays[26]]. This average value is slightly larger than bulk denaturation studies that report ~0.5 kcal/mol/aa[27] and similar to past single-molecule results of 1.1–1.3 kcal/mol per aa.[10,15,21]

For completeness, we note that most rigorous to-date experimental determination of the total ΔG_0 for the E–A helices is 227 ± 38 kcal/mol (mean ± SD) and is therefore significantly higher than our value of 156 ± 15 kcal/mol, a result not precisely reflected in the reported 1.3 kcal/mol per aa.[15] The origin of this difference is that this prior work used 171 amino acids in their analysis, which includes an extra 11 amino acids associated with the loop between the top of E and F helices and the first turn of the F-helix. If we reanalyze this earlier result with the same number of amino acids as used in this present work, their value now becomes 1.4 ± 0.2 kcal/mol per aa. Our average unfolding energy per amino acid therefore lies below this prior experimental result. Possible origins of this difference could be the improved precision of our force-extension curves and/or some reduction in

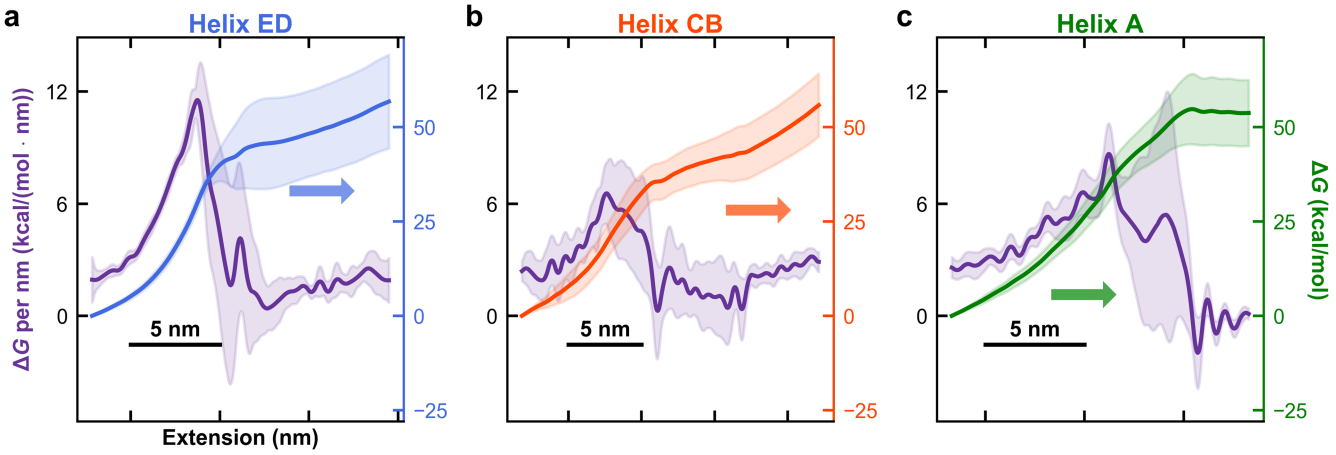

FIG. 3. Local unfolding free energies as a function of extension vary broadly between and within BR's major structural elements (the ED, CB, and A helices, respectively). (a) Local unfolding free energy per nm (left axis) and reconstructed energy landscape (right axis) for the ED helix pair. The local unfolding free-energy is the spatial first derivative of the reconstructed landscape. The standard deviations in the deduced values are the shaded region around the measured value as deduced by a bootstrap analysis. [(b) and (c)] Same analysis for the BC helix pair and the A helix, respectively.

the total ΔG_0 values due to the IWT poorly resolving stiff portions of the landscapes compared to the weighted histogram method (see below).[20] Finally, a recent coarse-grained molecular dynamics simulation predicts a free energy of transfer from bulk water to lipid of −130 kcal/mol for the 160-aa sequence in our work or about 0.8 kcal/mol per aa for unfolding BR from its lipid bilayer (see Sec. S1 of the supplementary material).[28]

Our work also documented that ΔG_0 per nm varied more than 15-fold, from 11 kcal/mol at the top of the ED helix pair to 0.6 kcal/mol at the bottom of the same helix pair. While it has long been known that hydrophobicity of the amino-acid sequence within the lipid bilayer and inter- and intra-helix interactions contribute to this variation,[29] recent work also shows variations in the strength of hydrogen bonds within TM helices.[30] To convert this local, single-molecule force spectroscopy derived metric to the more commonly reported ΔG_0 per aa, we note that there are ~3 amino acid residues per nm of contour length for an unstructured peptide based on crystallography (0.36 nm/aa).[31] Given the elevated forces experienced during unfolding portions of the records, we approximated extension with contour length and therefore approximated a local value for ΔG_0 per aa as 1/3 of ΔG_0 per nm after taking the spatial derivative of the reconstructed free-energy landscape [e.g., Fig. 2(c)]. As described in Sec. II, the landscapes and its derivatives were calculated for each set of data derived from a cantilever of fixed stiffness, after smoothing with a piecewise polynomial to 0.4 nm. At the top of the ED helix pair, we observed the highest local ΔG_0 per aa followed by the top of the CB and A helices.

One immediate trend from these data was a higher ΔG_0 per aa when BR was more fully folded. Thus, looking forward, an assay that probes the initial unfolding of a membrane protein where most, if not all of intra- and inter-helix bonds are still intact, offers the best opportunity to precisely quantify ΔG_0 per aa under native-like conditions. Unfortunately, while a few studies have quantified the initial unfolding of the GF helix,[26,32] precise quantification remains hindered by nonspecific adhesion between the tip and the surface.[8,9,15] Another trend in our data was that ΔG_0 per aa for unwinding the first portion of a TM helix pair was higher than when unfolding the second helix. Put more simply, once the first helix in a TM helix pair was extracted, the second helix took significantly less energy to extract.

Interestingly, we also achieved agreement between the unfolding free energies obtained by equilibrium and dynamic assays. As discussed in the Introduction, we had previously deduced an unexpectedly high ΔG_0 per aa [2.7 ± 0.1 kcal/mol (mean ± SEM)] for a particular transition near the top of the ED helix pair after reconstruction of the free-energy landscape from equilibrium data [Fig. 1(f)].[18] The present IWT analysis of dynamic data yielded an average ΔG_0 per aa near the top of the ED helix, 2.9 ± 0.5 kcal/mol (mean ± SD), in agreement with the previous equilibrium result for a specific three-amino-acid transition.[18]

We emphasize that the ΔG_0 per aa measured here does not correspond to breaking a particular hydrogen-bond or fully solvating an amino acid out of a lipid bilayer but results from a convolution of angstrom-scale motion of each amino acid along the stretching axis. We also note that the reported values include the contribution of the extended, unfolded protein. In other words, since the ED helix was folded at the start of the analysis, the entropic contribution to stretching unfolded polypeptides to high force is part of the resulting ΔG_0. Future experimental advances and/or applying more sophisticated analysis can minimize or theoretically separate out the energy associated with stretching these compliant unfolded peptides.[33] However, the situation for BR is distinctly different from standard optical-trapping assays which use much longer, 300–600-nm double-stranded nucleic acid handles to accommodate the geometry of optical-trapping studies.[34–36]

Somewhat unexpectedly, we observed a local landscape stiffness—the second derivative of the reconstructed landscape—greater than any particular cantilever stiffness used. Yet, Hummer and Szabo's implementation of the IWT assumed a stiff probe, which should limit the reconstructed landscape from having a stiffness that exceeds the force probe.[19] More quantitatively, the greatest landscape curvature

for BR was 8 kcal/(mol nm^2) or 55 pN/nm, larger than the stiffest cantilever used (40 pN/nm). We attribute this small discrepancy to errors in polynomial fitting and numerical differentiation. On the scientific front, it is therefore likely that our reconstructed landscape represents a lower bound on the local value of ΔG_0 per nm, akin to early studies of RNA pseudoknots studied with optical traps.[20] Hence, as shown in that optical-trapping work,[20] the weighted histogram method is better for reconstructing landscapes containing such "stiff" features. For the vast majority of BR's landscape, the stiff-spring approximation of the IWT was fulfilled and yielded a more accurate landscape reconstruction (Fig. S3 of the supplementary material). In addition, the almost-negligible correction to the landscape from the $\ddot{A}$ term outside of the ED helix [Fig. 2(b), brown] confirmed that higher-order corrections were unlikely to affect the landscape of the CB and A helices. In fact, the corrections of the IWT to the Jarzynski equality were small except near the top of the E helix, which interestingly is also the same location where equilibrium flickering was previously reported [Fig. 1(e)]. Looking forward, using a stiffer cantilever is the obvious means to fulfill the stiff probe assumption. Yet, stiffer ultrashort cantilevers are not overdamped (quality factor $Q > 0.5$),[16] violating an assumption underlying traditional force spectroscopy theories.[11,37,38] Hence, there is also ongoing need to optimize modified ultrashort cantilevers for higher k but low Q.

IV. CONCLUSION

As anticipated,[19,20] the IWT provides an efficient and accessible means to reconstruct free-energy landscapes when using AFM, even for complicated landscapes such as BR. The energy landscape reconstruction presented here highlights significant variation along the unfolding free-energy landscape. The overall trend is similar to early work[15] but with improvements associated with reconstruction of the landscape from much higher resolution force-extension curves[16,18] and using the IWT[19] that efficiently removes the energy contribution associated with the bending of the cantilever. As a result, we achieved a more accurate and higher-resolution reconstruction. The resulting landscape yielded an average ΔG_0 per aa of 1.0 ± 0.1 kcal/mol, in much better agreement with past single-molecule studies than our one prior determination of ΔG_0 per aa (2.7 kcal/mol) derived from an equilibrium assay probing a particular three-aa transition near the top of the ED helix pair. Importantly, on a smaller spatial scale, this high-resolution landscape also agreed with that prior equilibrium measurement of the three-aa transition ($\Delta G_0 = 2.9 \pm 0.5$ kcal/mol/aa). Indeed, the highest-measured curvature of the free-energy landscape (~55 pN/nm) was limited by cantilever stiffness (~10–40 pN/nm) and hence represents a lower bound of local determination of ΔG_0 per aa due to the underlying assumption of a stiff probe in applying the IWT.[19,20] Future developments that enable efficient unfolding and refolding over reversible transitions would enable comparing energy landscapes deduced from state-based models[39] with those used here, providing another method for investigating and verifying local free-energy landscape reconstruction of membrane proteins.

SUPPLEMENTARY MATERIAL

See supplementary material for three graphs that show the curvature of the BR landscape as a function of molecular extension and details of data analysis.

ACKNOWLEDGMENTS

We thank Michael Woodside for useful discussions. This work was supported by the National Science Foundation (DBI-1353987 to T.T.P.; MCB-1716033 to T.T.P., Phy-1734006), and National Institute of Standards and Technology (NIST). Mention of commercial products is for information only; it does not imply NIST's recommendation or endorsement. T.T.P. is a staff member of NIST's quantum physics division.

[1] A. M. Stanley and K. G. Fleming, "The process of folding proteins into membranes: Challenges and progress," Arch. Biochem. Biophys. **469**, 46–66 (2008).
[2] J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, "Theory of protein folding: The energy landscape perspective," Annu. Rev. Phys. Chem. **48**, 545–600 (1997).
[3] H. Janovjak, K. T. Sapra, A. Kedrov, and D. J. Muller, "From valleys to ridges: Exploring the dynamic energy landscape of single membrane proteins," ChemPhysChem **9**, 954–966 (2008).
[4] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub, "Reversible unfolding of individual titin immunoglobulin domains by AFM," Science **276**, 1109–1112 (1997).
[5] M. S. Kellermayer, S. B. Smith, H. L. Granzier, and C. Bustamante, "Folding-unfolding transitions in single titin molecules characterized with laser tweezers," Science **276**, 1112–1116 (1997).
[6] A. Borgia, P. M. Williams, and J. Clarke, "Single-molecule studies of protein folding," Annu. Rev. Biochem. **77**, 101–125 (2008).
[7] K. C. Neuman and A. Nagy, "Single-molecule force spectroscopy: Optical tweezers, magnetic tweezers and atomic force microscopy," Nat. Methods **5**, 491–505 (2008).
[8] C. A. Bippes and D. J. Muller, "High-resolution atomic force microscopy and spectroscopy of native membrane proteins," Rep. Prog. Phys. **74**, 086601 (2011).
[9] F. Oesterhelt, D. Oesterhelt, M. Pfeiffer, A. Engel, H. E. Gaub, and D. J. Muller, "Unfolding pathways of individual bacteriorhodopsins," Science **288**, 143–146 (2000).
[10] G. Hummer and A. Szabo, "Free energy reconstruction from nonequilibrium single-molecule pulling experiments," Proc. Natl. Acad. Sci. U. S. A. **98**, 3658–3661 (2001).
[11] E. Evans and K. Ritchie, "Dynamic strength of molecular adhesion bonds," Biophys. J. **72**, 1541–1555 (1997).
[12] R. Merkel, P. Nassoy, A. Leung, K. Ritchie, and E. Evans, "Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy," Nature **397**, 50–53 (1999).
[13] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, "Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality," Science **296**, 1832–1835 (2002).
[14] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, "Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies," Nature **437**, 231–234 (2005).
[15] J. Preiner, H. Janovjak, C. Rankl, H. Knaus, D. A. Cisneros, A. Kedrov, F. Kienberger, D. J. Muller, and P. Hinterdorfer, "Free energy of membrane protein unfolding derived from single-molecule force measurements," Biophys. J. **93**, 930–937 (2007).
[16] D. T. Edwards, J. K. Faulk, A. W. Sanders, M. S. Bull, R. Walder, M. A. LeBlanc, M. C. Sousa, and T. T. Perkins, "Optimizing 1-μs-resolution single-molecule force spectroscopy on a commercial atomic force microscope," Nano Lett. **15**, 7091–7098 (2015).
[17] D. T. Edwards and T. T. Perkins, "Optimizing force spectroscopy by modifying commercial cantilevers: Improved stability, precision, and temporal resolution," J. Struct. Biol. **197**, 13–25 (2017).
[18] H. Yu, M. G. Siewny, D. T. Edwards, A. W. Sanders, and T. T. Perkins, "Hidden dynamics in the unfolding of individual bacteriorhodopsin proteins," Science **355**, 945–950 (2017).

[19]G. Hummer and A. Szabo, "Free energy profiles from single-molecule pulling experiments," Proc. Natl. Acad. Sci. U. S. A. **107**, 21441–21446 (2010).
[20]M. C. Engel, D. B. Ritchie, D. A. N. Foster, K. S. D. Beach, and M. T. Woodside, "Reconstructing folding energy landscape profiles from nonequilibrium pulling curves with an inverse Weierstrass integral transform," Phys. Rev. Lett. **113**, 238104 (2014).
[21]M. Kessler, K. E. Gottschalk, H. Janovjak, D. J. Muller, and H. E. Gaub, "Bacteriorhodopsin folds into the membrane against an external force," J. Mol. Biol. **357**, 644–654 (2006).
[22]C. Jarzynski, "Nonequilibrium equality for free energy differences," Phys. Rev. Lett. **78**, 2690–2693 (1997).
[23]D. D. Minh and A. B. Adib, "Optimized free energies from bidirectional single-molecule force spectroscopy," Phys. Rev. Lett. **100**, 180602 (2008).
[24]S. Park, F. Khalili-Araghi, E. Tajkhorshid, and K. Schulten, "Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality," J. Chem. Phys. **119**, 3559–3566 (2003).
[25]W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, New York, 1992).
[26]D. J. Muller, M. Kessler, F. Oesterhelt, C. Moller, D. Oesterhelt, and H. Gaub, "Stability of bacteriorhodopsin alpha-helices and loops analyzed by single-molecule force spectroscopy," Biophys. J. **83**, 3578–3588 (2002).
[27]F. Cymer, G. von Heijne, and S. H. White, "Mechanisms of integral membrane protein insertion and folding," J. Mol. Biol. **427**, 999–1022 (2015).
[28]T. Yamada, T. Yamato, and S. Mitaku, "Forced unfolding mechanism of bacteriorhodopsin as revealed by coarse-grained molecular dynamics," Biophys. J. **111**, 2086–2098 (2016).
[29]S. H. White and W. C. Wimley, "Membrane protein folding and stability: Physical principles," Annu. Rev. Biophys. Biomol. Struct. **28**, 319–365 (1999).
[30]Z. Cao, J. M. Hutchison, C. R. Sanders, and J. U. Bowie, "Backbone hydrogen bond strengths can vary widely in transmembrane helices," J. Am. Chem. Soc. **139**, 10742–10749 (2017).
[31]L. Pauling and R. B. Corey, "The pleated sheet, a new layer configuration of polypeptide chains," Proc. Natl. Acad. Sci. U. S. A. **37**, 251–256 (1951).
[32]M. Kessler and H. E. Gaub, "Unfolding barriers in bacteriorhodopsin probed from the cytoplasmic and the extracellular side by AFM," Structure **14**, 521–527 (2006).
[33]M. T. Woodside and S. M. Block, "Reconstructing folding energy landscapes by single-molecule force spectroscopy," Annu. Rev. Biophys. **43**, 19–39 (2014).
[34]J. Liphardt, B. Onoa, S. B. Smith, I. J. Tinoco, and C. Bustamante, "Reversible unfolding of single RNA molecules by mechanical force," Science **292**, 733–737 (2001).
[35]M. T. Woodside, P. C. Anthony, W. M. Behnke-Parks, K. Larizadeh, D. Herschlag, and S. M. Block, "Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid," Science **314**, 1001–1004 (2006).
[36]C. Cecconi, E. A. Shank, C. Bustamante, and S. Marqusee, "Direct observation of the three-state folding of a single protein molecule," Science **309**, 2057–2060 (2005).
[37]G. I. Bell, "Models for specific adhesion of cells to cells," Science **200**, 618–627 (1978).
[38]O. K. Dudko, G. Hummer, and A. Szabo, "Intrinsic rates and activation free energies from single-molecule pulling experiments," Phys. Rev. Lett. **96**, 108101 (2006).
[39]Y. J. Zhang and O. K. Dudko, "A transformation for the mechanical fingerprints of complex biomolecular interactions," Proc. Natl. Acad. Sci. U. S. A. **110**, 16432–16437 (2013).
[40]Y. M. Rhee and V. S. Pande, "One-dimensional reaction coordinate and the corresponding potential of mean force from commitment probability distribution," J. Phys. Chem. B **109**, 6780–6786 (2005).