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

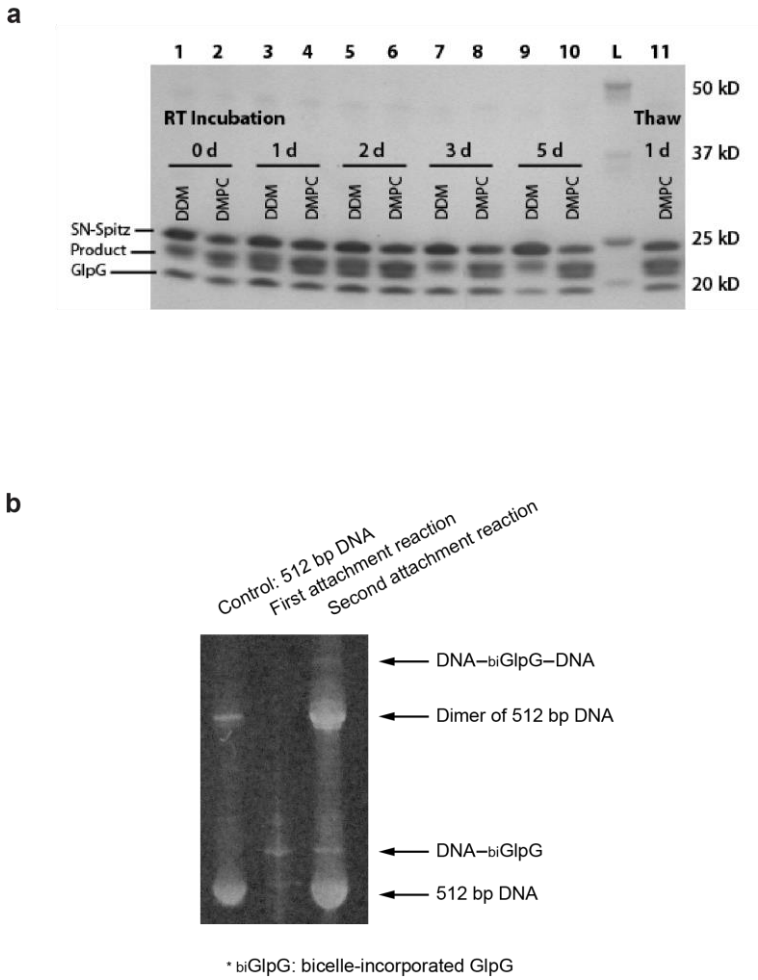
[5] These authors contributed equally to this work.

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

Nature Chemical Biology: doi:10.1038/nchembio.1939

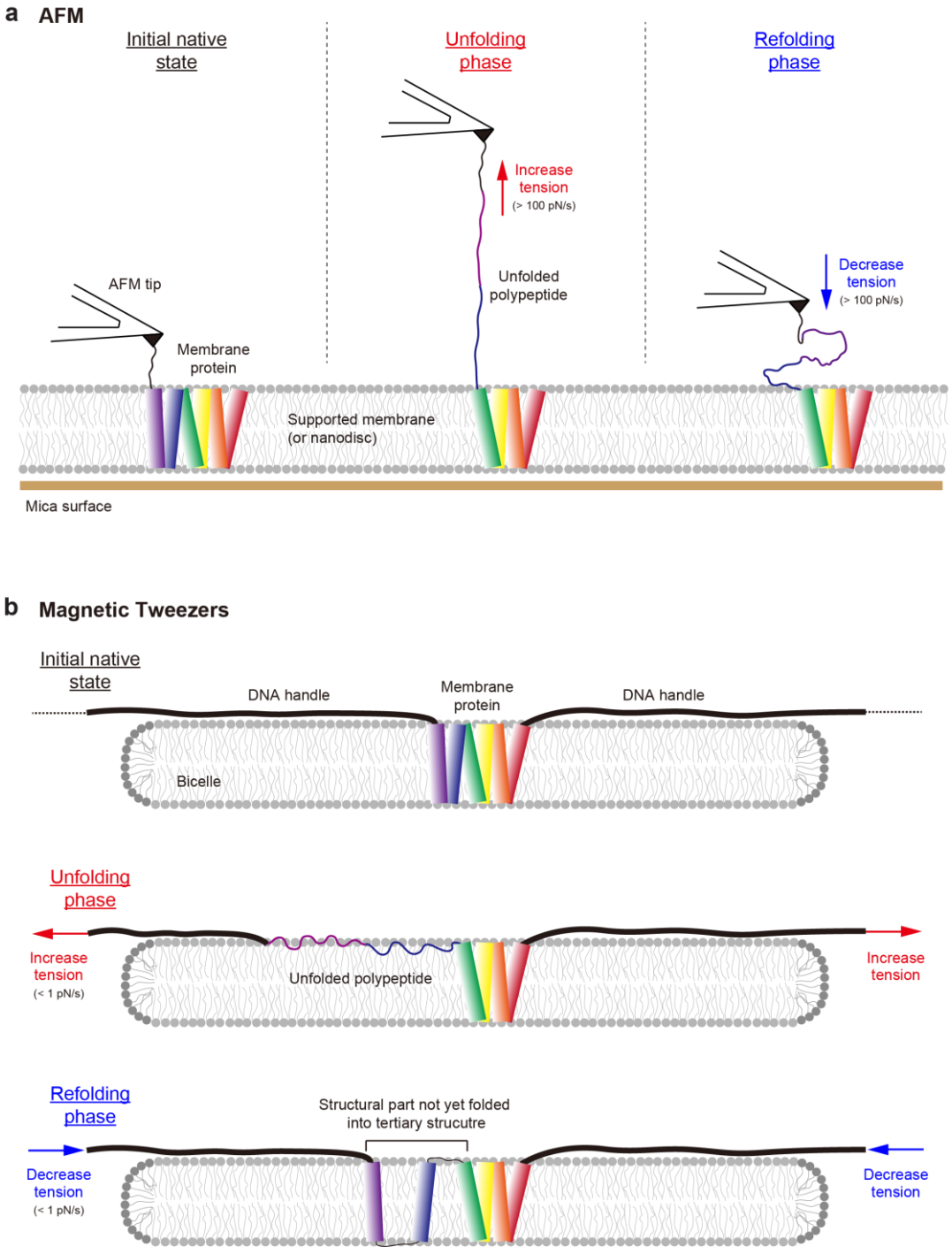
Supplementary Results

Nature Chemical Biology: doi:10.1038/nchembio.1939



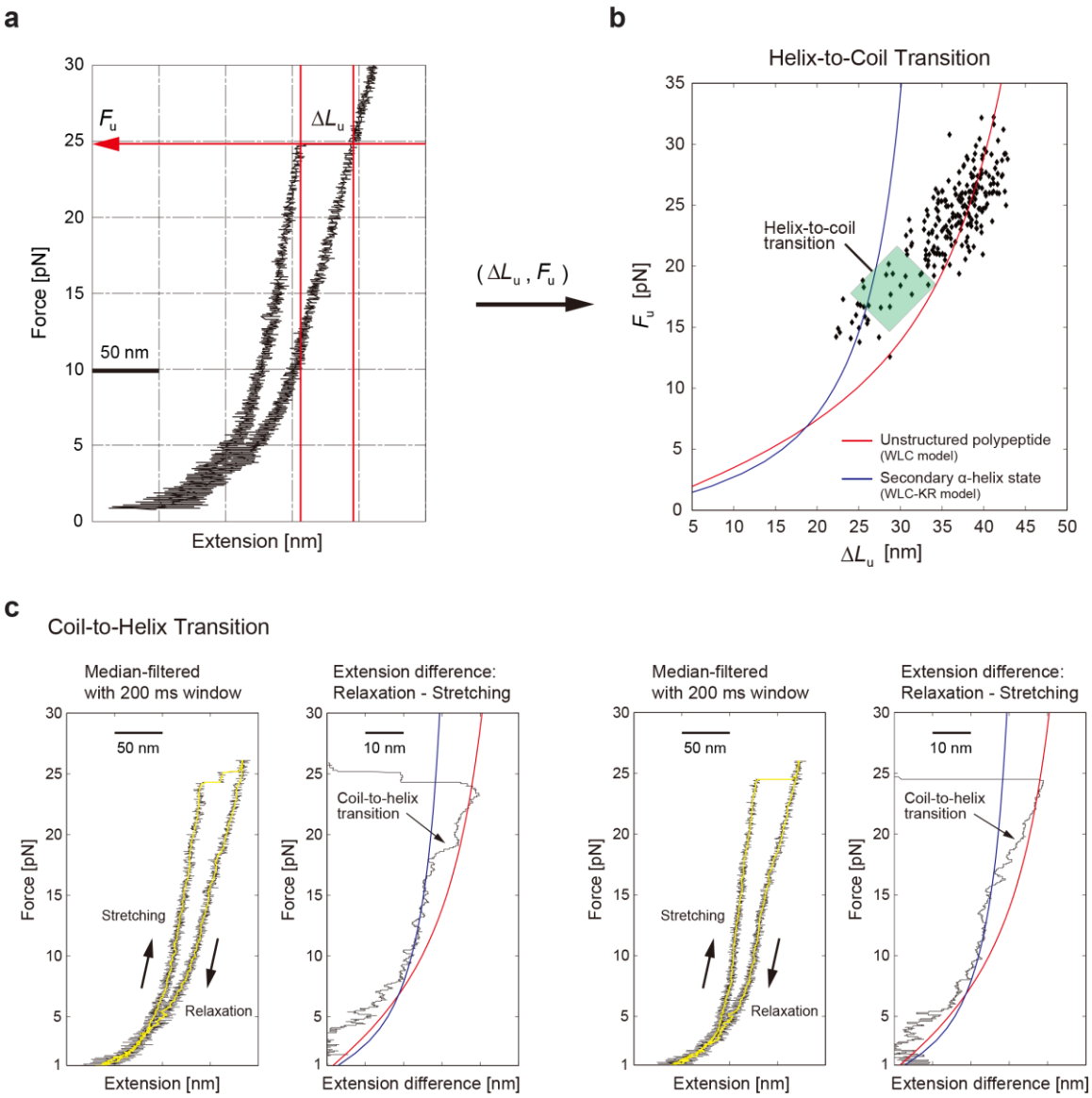
Supplementary Figure 1 | Characterization of GlpG activity and stability. (a) Activity assays of wild-type GlpG incubated in either 0.1% DDM or 2% DMPC:CHAPSO bicelles. Activity was measured after 0, 1, 2, 3, and 5 days of incubation at room temperature (lanes 1-10). The activity is lost after several 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).

Nature Chemical Biology: doi:10.1038/nchembio.1939



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).

Nature Chemical Biology: doi:10.1038/nchembio.1939

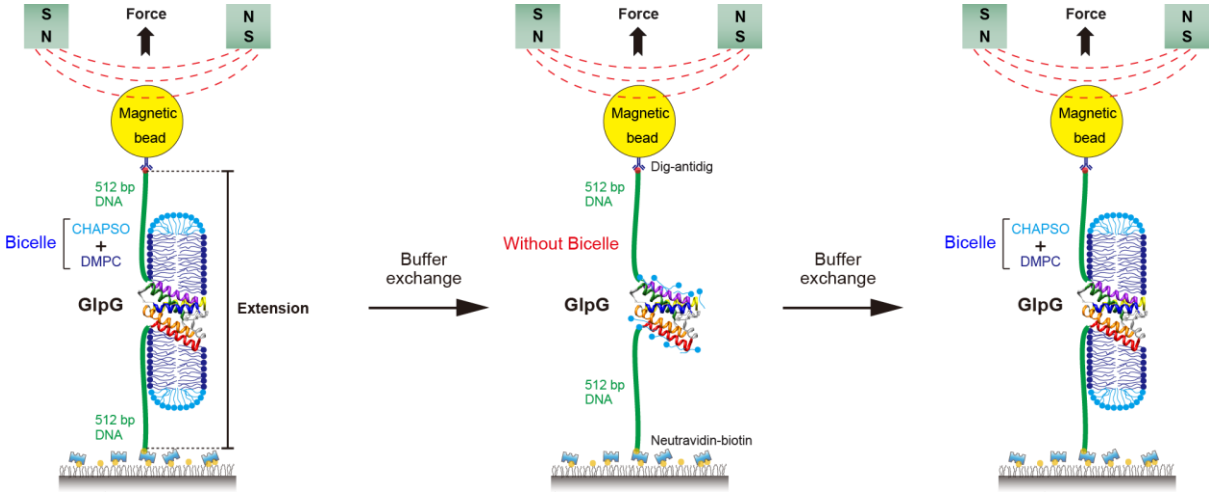


Supplementary Figure 3 | Helix-coil transition of GlpG. (a) Representative gradual pulling experiment showing how the increased extension (ΔL_u) and the unfolding force (F_u) are measured from force-extension traces. **(b)** Scatter plot of the unfolding data in the plane of ΔL_u and F_u (n=233). The 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

Nature Chemical Biology: doi:10.1038/nchembio.1939

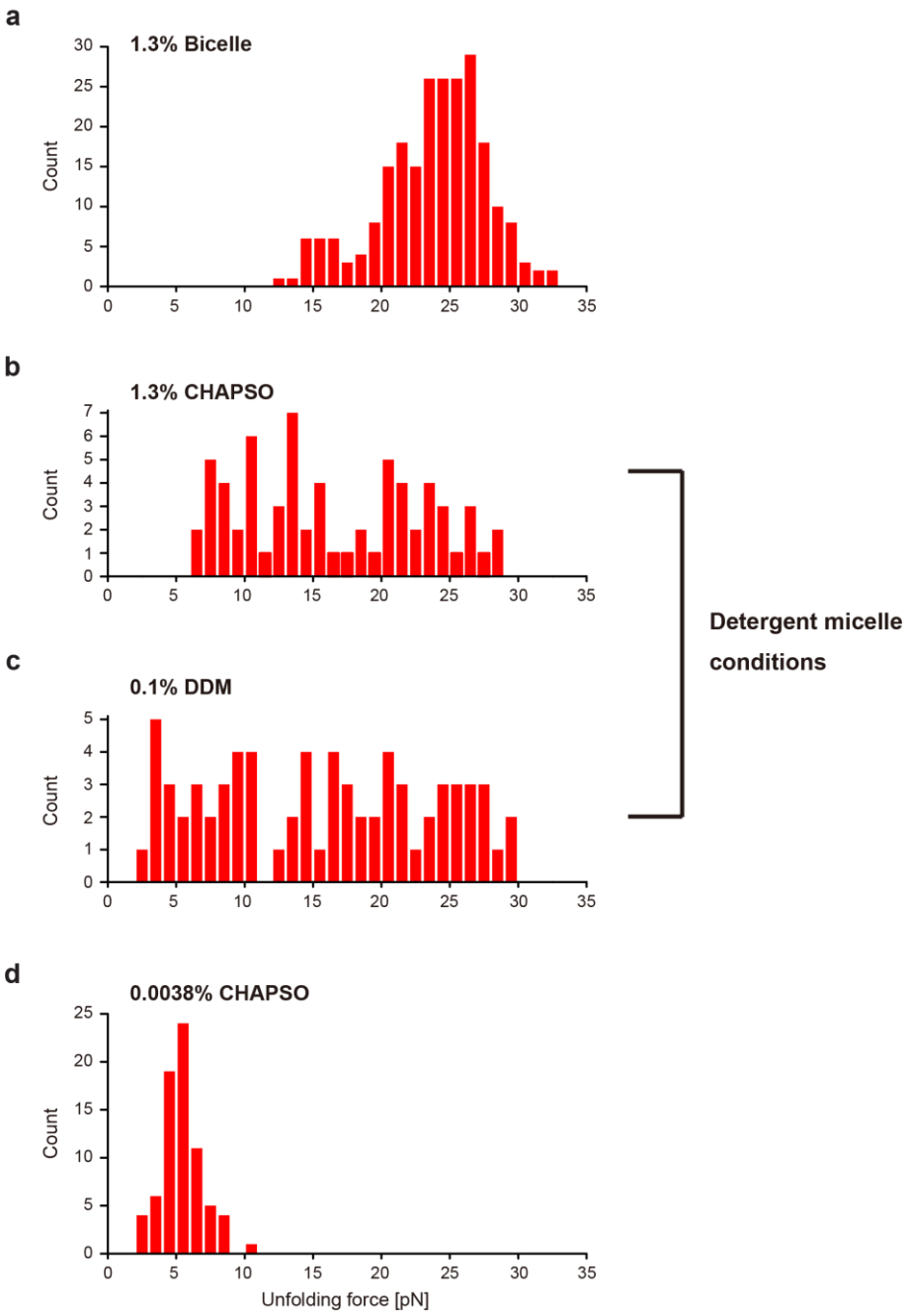
show the helix-to-coil transition at around 18 pN. **(c)** Representative single-molecule traces showing the coil-to-helix transition during the relaxation phase. To visualize the transitions more clearly, the traces were median-filtered with a 200 ms window (left panels) and the extension difference 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.

Nature Chemical Biology: doi:10.1038/nchembio.1939



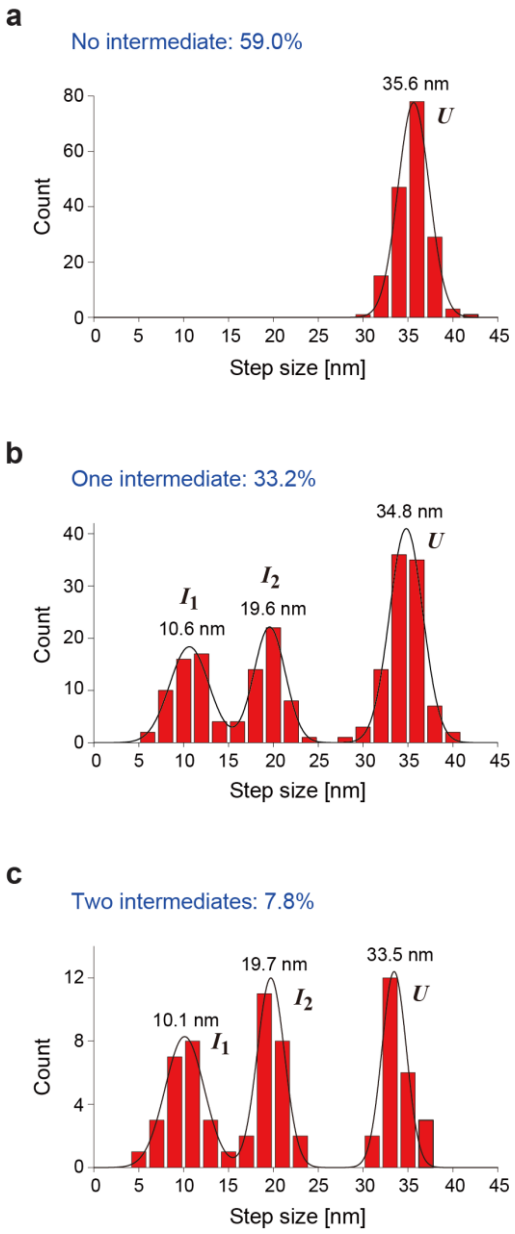
Supplementary Figure 4 | Schematic diagram of the single-molecule magnetic tweezers experiment with microfluidic buffer exchanges. After several cycles of GlpG unfolding and refolding in bicelles (left), the bicelles were removed and the unfolding and refolding cycles were repeated (middle). After up 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.

Nature Chemical Biology: doi:10.1038/nchembio.1939



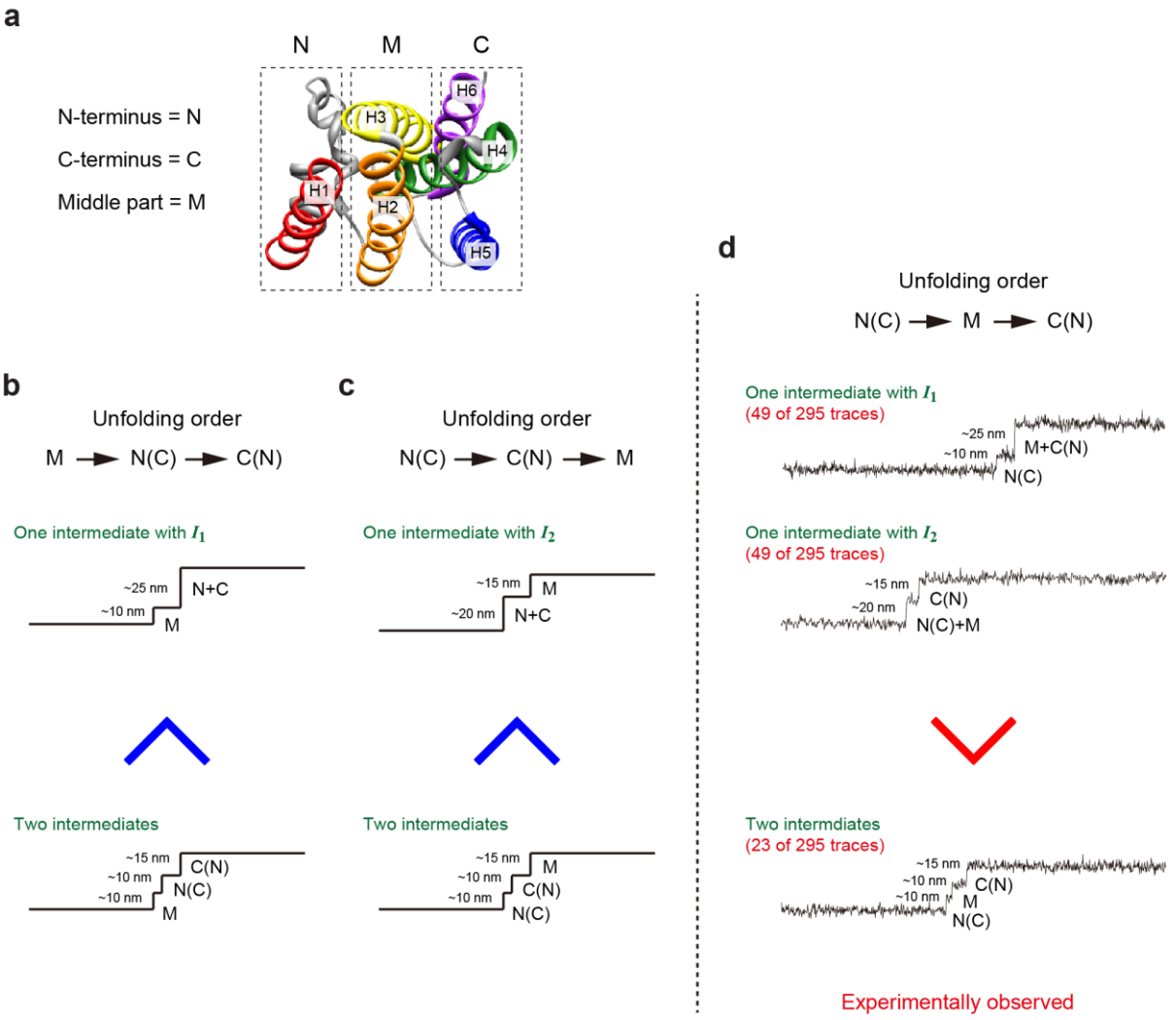
Supplementary Figure 5 | Unfolding forces of GlpG in various reconstitution conditions. **(a-d)** Distributions of the unfolding forces when GlpG proteins are reconstituted in 1.3% bicelle (*n*=233) (a), 1.3% CHAPSO micelle (*n*=66) (b), 0.1% DDM micelle (*n*=71) (c) or a minimal amount of detergents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

Nature Chemical Biology: doi:10.1038/nchembio.1939



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.

Nature Chemical Biology: doi:10.1038/nchembio.1939

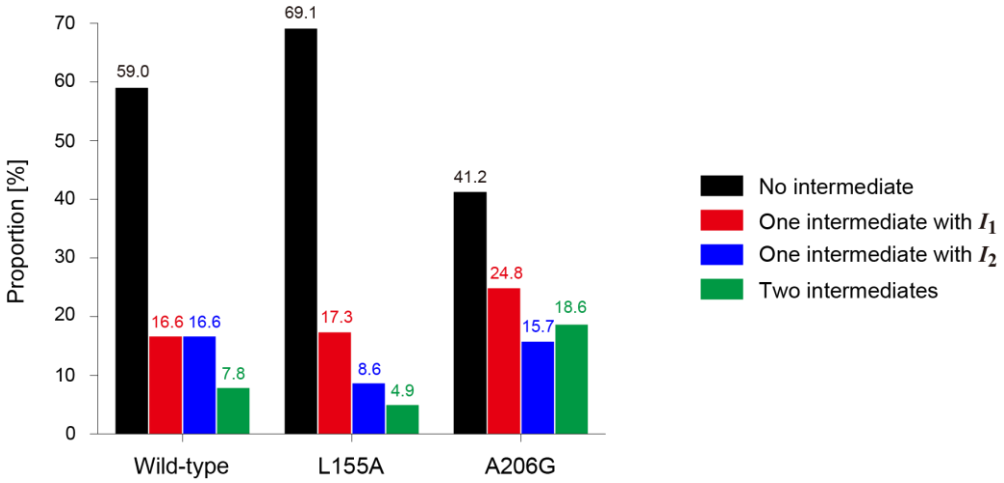


Supplementary Figure 7 | Argument for unidirectional unfolding of GlpG. (a) GlpG structure from a cytoplasmic view showing three distinct domains. Each structural domain is presumed to correspond to one of the three steps observed in the unfolding events. The N-terminal, middle and C-terminal parts of 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

Nature Chemical Biology: doi:10.1038/nchembio.1939

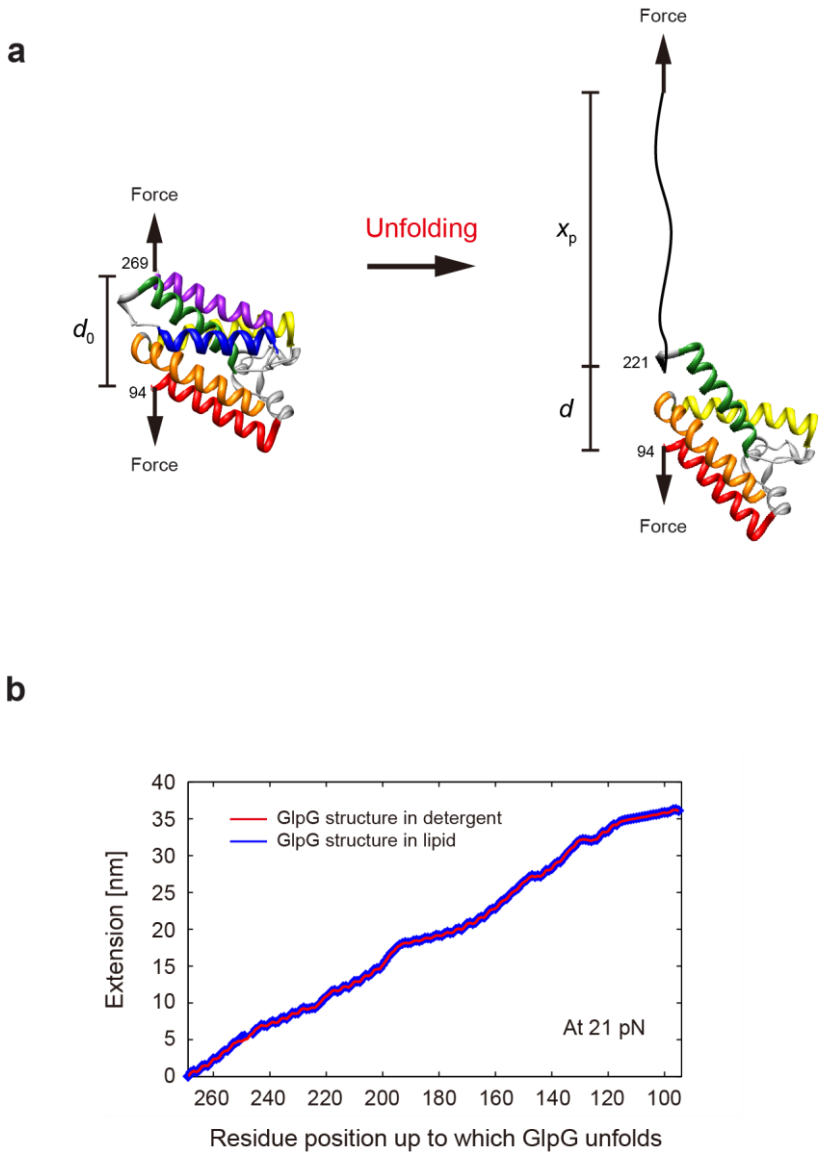
than that of two-intermediate unfolding (d), thereby defying the scenarios described in (b) and (c). We are only left with the possibility that M-domain unfolding intercalates between unfolding of the N and C domains, which points to unidirectional unfolding of GlpG.

Nature Chemical Biology: doi:10.1038/nchembio.1939



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.

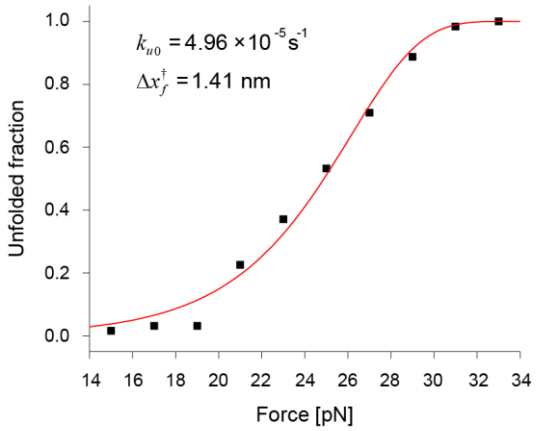
Nature Chemical Biology: doi:10.1038/nchembio.1939



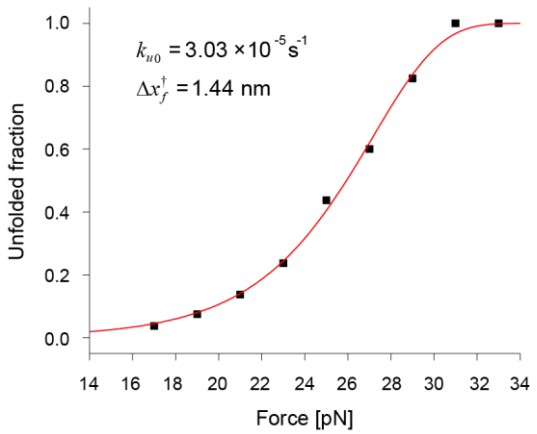
Supplementary Figure 9 | Extension analysis. (a) Cartoon describing an increased extension due to the unfolding of GlpG up to a specific residue. The expected extension increases were estimated as $x = x_p + \Delta d$, where x_p is the extension of unraveled polypeptide calculated with the worm-like 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.

Nature Chemical Biology: doi:10.1038/nchembio.1939

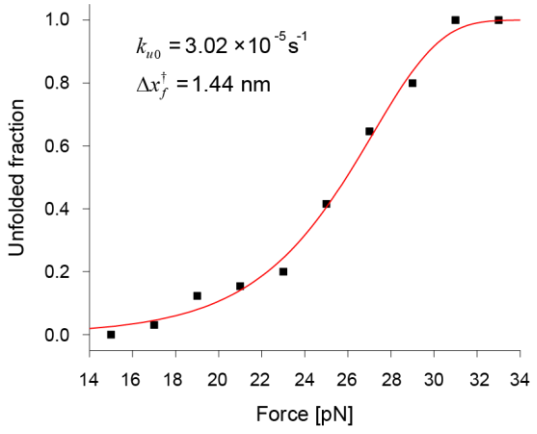
a Lipid : detergent = 2.5 : 1



b Lipid : detergent = 2.8 : 1



c Temperature = 25°C



Nature Chemical Biology: doi:10.1038/nchembio.1939

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 ℃ (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.

Nature Chemical Biology: doi:10.1038/nchembio.1939

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).

Nature Chemical Biology: doi:10.1038/nchembio.1939

	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).