Protein Folding

DOI: 10.1002/anie.200905747

Evidence for a Broad Transition-State Ensemble in Calmodulin Folding from Single-Molecule Force Spectroscopy**

*Jan Philipp Junker and Matthias Rief**

In recent years, single-molecule force spectroscopy has provided a wealth of insights into protein folding and unfolding.[1–7] Using mechanical force as a denaturant offers distinct advantages over thermal or chemical denaturation, such as the possibility to locally probe the folding free-energy landscape[8] or to actively steer the unfolding pathway of proteins.[9] However, it has been argued that the unfolding pathways that are probed in single-molecule force spectroscopy are likely to be very different from the pathways studied by chemical denaturation,[10] which would limit or exclude any comparison between single-molecule force spectroscopy experiments and conventional folding techniques.

To date, this hypothesis has only been investigated under non-equilibrium conditions, that is, under conditions that favor either unfolding or folding. With the advent of low-drift single-molecule force spectroscopy by AFM, it has recently become possible to study folding and unfolding under equilibrium conditions, thus allowing observation of multiple folding/unfolding transitions of single protein molecules in a single trace.[7,11] Herein, we investigate folding and unfolding pathways of the calcium signaling protein calmodulin (CaM) close to thermodynamic equilibrium by applying force in different pulling geometries. CaM is a small protein composed of two domains, both of which show two-state folding and unfolding in AFM experiments.[7] The experiments shown herein allow us to distinguish whether folding and unfolding pathways in different pulling geometries converge onto a common pathway close to equilibrium and to investigate whether a structurally distinct zero-force pathway exists.

Single-molecule force spectroscopy experiments are usually performed with protein constructs as shown in Figure 1a, where the protein of interest (i.e., CaM) is inserted on the DNA level between protein domains that serve as handles for the attachment to surface and AFM cantilever tip so that a single multi-domain protein construct is expressed. This technique is however limited to N–C-terminal pulling geometries. We have developed a method to create well-defined heteropolymers with nearly arbitrary points of force application by assembly on the protein level, which was inspired by solid-phase peptide synthesis (Figure 1b). Handle proteins with a 6×His tag are allowed to bind to a Ni^{2+}-NTA resin (NTA = nitrilotriacetic acid); the protein of interest and chain-terminating handle proteins (both without His tag) are then sequentially attached by disulfide bridge formation.[12,13] The cysteine residues of the protein of interest can be protected by the coupling reagent dithiodipyridine to avoid formation of homopolymers (see the Supporting Information).

We have used this technique to study the unfolding kinetics of single CaM molecules in three different non-N–C-terminal directions (Figure 1c). By mutating amino acid residues 38 and 118 into cysteines (CaM 38 118), the mean unfolding forces of the N- and C-terminal domain of CaM (DomN and DomC) were increased to about 25 pN at a pulling velocity of 50 nm s^{-1}, as compared to circa 13 pN in the N–C-terminal direction[7] (Figure 1d). The measured length

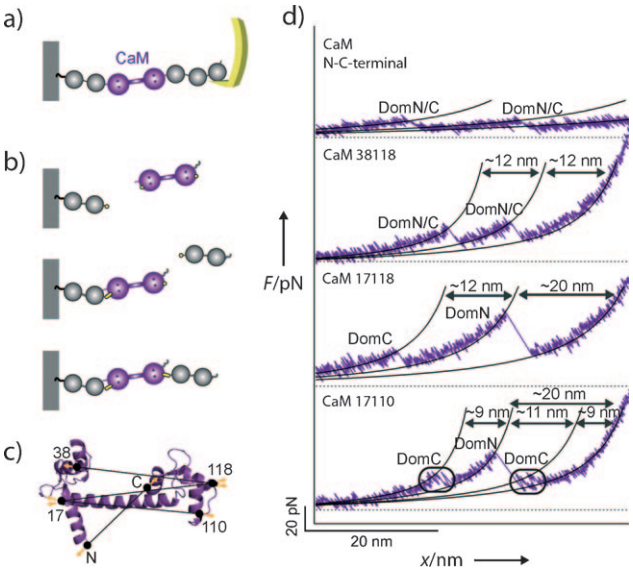

Figure 1. a) The experimental setup with protein construct composed of CaM (purple) and handle domains (gray). b) Stepwise synthesis of well-defined heteropolymers with alternative pulling directions. c) Structure of CaM with a sketch of the pulling geometries studied. d) Sample traces of CaM in four different pulling directions at $v_{pull} = 50$ nm s^{-1}. The curves have two peaks, indicating that both domains of CaM unfold independently from each other. Varying the force application point leads to significantly different unfolding forces.

[*] Prof. M. Rief
Physik-Department E22, Technische Universität München
and Munich Center for Integrated Protein Science CiPS^{M}
James-Franck-Strasse, 85748 Garching (Germany)
Fax: (+49) 89-289-12523
E-mail: mrief@ph.tum.de
Dr. J. P. Junker
Physik-Department E22, Technische Universität München
James-Franck-Strasse, 85748 Garching (Germany)

[**] We thank Dr. M. Schlierf for inspiring discussions and M. Bertz and F. Ziegler for comments on the manuscript. Financial support from DFG grant RI 990/3-1 is gratefully acknowledged. J.P.J. was supported by the International Graduate School "Materials Science of Complex Interfaces".

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200905747.

increases of about 12 nm for both domains correspond exactly to the expected values (see Supporting Information), thus making the two domains undistinguishable in this pulling direction. Changing the force application point at DomN from amino acid residue 38 to 17 (CaM 17 118) increases the contour length increase of DomN to circa 20 nm, which is again in full agreement with the expected value. Interestingly, unfolding forces of DomN in this pulling direction are increased even further to about 40 pN. The unfolding force of DomC, however, remains unchanged, suggesting that both domains of CaM unfold independently from each other. If the C-terminal force application point is furthermore changed to amino acid residue 110 (CaM 17 110), the unfolding force of DomC is not altered significantly, but refolding events can also now be observed in the traces, in agreement with the previous observation that shortening the contour length that is exposed to force increases the folding rate under force.[8] For CaM 17 110, equilibrium transitions of DomC can now be observed at two points in the trace, before and after the unfolding peak of DomN, because DomC now folds back as soon as forces are lower than about 15 pN (see circled part in Figure 1 d, lower panel). Thus, similar to experiments on E2 Lip3[14] and GFP,[12] we find a whole range of different mechanical stabilities of CaM in different directions.

We then set out to study folding against a mechanical load with different points of force application to selectively populate different pathways. To simplify data analysis, we chose to investigate isolated DomN, varying the N-terminal point of force application (amino acid residue 1, 17, or 38, see also sketches in Figure 2 a–c) whilst keeping the C-terminal point of force application fixed at the C-terminal end of DomN, analogous to DomN in full-length CaM (Figure 1 b and c). Figure 2 a and b show typical approach–retract cycles of DomN 17 and DomN 38 that were recorded at a pulling velocity of 10 nm s^{-1}. As before, unfolding forces of DomN 17 are considerably higher than those of DomN 38, and the hysteresis between forward and backward traces, which is a measure of how far the pulling cycle is away from thermodynamic equilibrium, is significantly larger for DomN 17 (see also Supporting Information, Figure S1). The proper equilibrium energetics was confirmed using the Crooks fluctuation theorem[15] (Supporting Information, Figure S2). At this pulling velocity, DomN 1, the least-stable pulling direction studied, is already in equilibrium with respect to folding (Figure 2 c), showing multiple transitions in a single trace.

We determined refolding force distributions by measuring the force of the first refolding event upon approaching the cantilever to the surface (arrows in Figure 2 a–c) and analyzed the refolding forces with a model in which the folding rate k_f is the only free parameter.[8,16,17] Comparing refolding force distributions to Monte Carlo simulations (see Supporting Information), we find that a zero-force folding rate of $k_f = 50000\ s^{-1}$ reproduces the data within the experimental error for all pulling directions (Figure 2 d–f). This finding suggests that the folding pathways for all pulling directions are either identical or at least populate transition states of similar free energy.

To analyze the unfolding pathways of CaM in more detail, we measured the mean unfolding force of DomN in the three different directions at varying pulling velocities (Figure 3 a). Analysis of the pulling speed dependence of unfolding forces is a very sensitive tool to distinguish different pathways in single-molecule force spectroscopy, as pathways with different values of the potential width Δx_{N-TS} (the distance between the native state N and the transition state TS along the reaction coordinate; see also Figure 3 c) give rise to distinct linear slopes if F_{mean} is plotted versus the logarithm of v_{pull}.[10,18] A clear linear correlation was found between F_{mean} and $\log(v_{pull})$ for all three pulling geometries, thus indicating that no pathway switches occur at the transition from equilibrium to non-equilibrium conditions. However, the slopes are significantly different for the three pulling directions, showing that the potential width Δx_{N-TS} is different for each pulling geometry. This result is a direct indication that

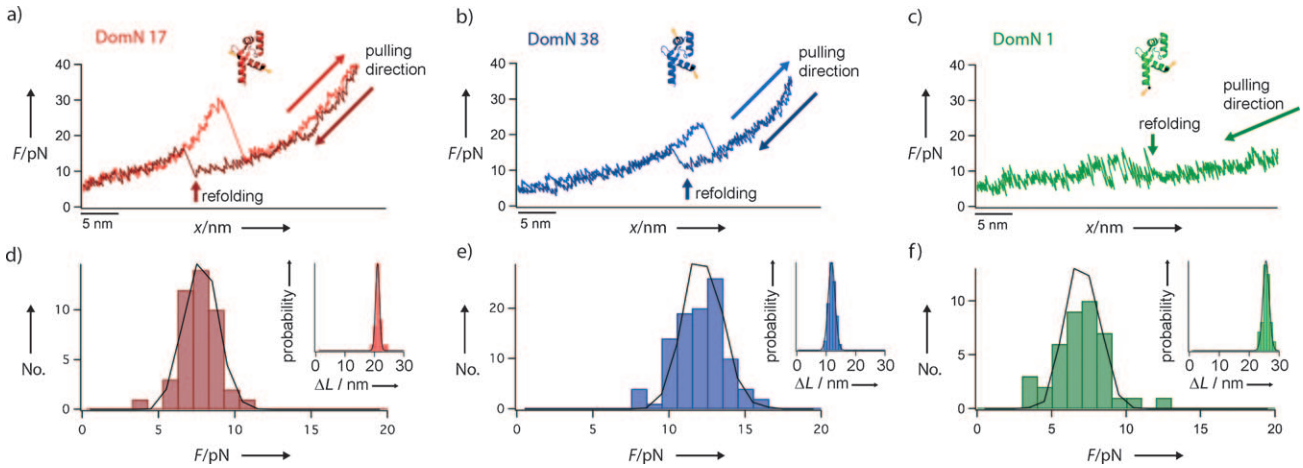

Figure 2. a) Unfolding (light red) and refolding (dark red) trace of DomN 17 at $v_{pull} = 10$ nm s^{-1}. b) Unfolding (light blue) and refolding (dark blue) trace of DomN 38 at $v_{pull} = 10$ nm s^{-1}. c) Equilibrium folding/unfolding trace of DomN 1 at $v_{pull} = 10$ nm s^{-1}. d–f) Refolding force distributions of DomN 17, 38, and 1. Results of Monte Carlo simulations with $k_f = 50000\ s^{-1}$ are shown in black. The insets depict contour length histograms of the unfolding transitions. The mean values are 20.7 ± 1.2 nm for DomN 17, 11.8 ± 1.2 nm for DomN 38, and 25.6 ± 1.3 nm for DomN 1.

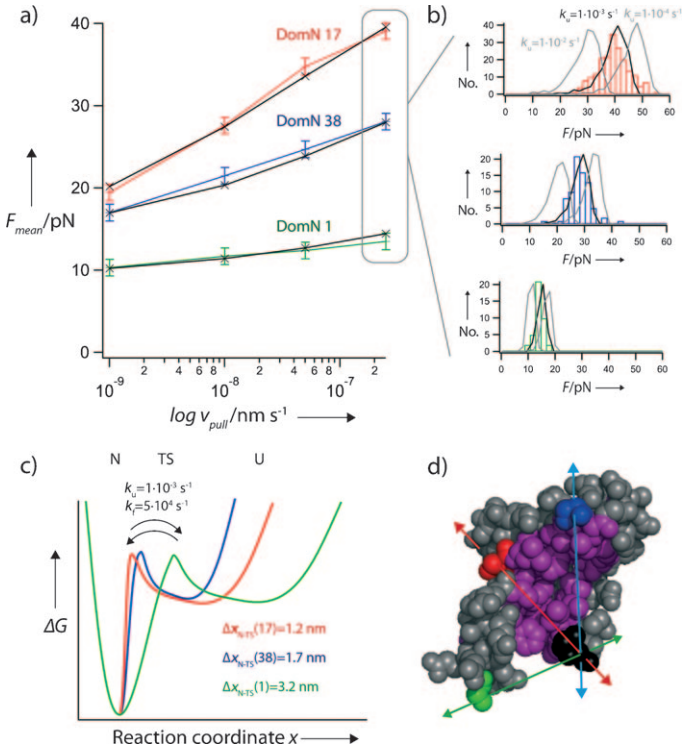

Figure 3. a) Pulling-speed dependence of mean unfolding forces in different pulling directions of DomN. Results of Monte Carlo simulations are shown in black. b) Unfolding force distributions of DomN 17, DomN 38, and DomN 1 at $v_{pull} = 250$ nm s^{-1}. Monte Carlo simulations with $k_u = 0.001$ s^{-1} are shown in black; simulations with an unfolding rate that is a factor of ten higher or lower are shown in gray. c) Depiction of free-energy landscapes extrapolated to zero force. d) Structure of DomN. The varying N-terminal force application points are shown in red, blue, and green, and the C-terminal force application point is shown in black. The hydrophobic core of CaM, which is partly exposed to the solvent if Ca^{2+} is bound, is shown in purple.

for unfolding, a different pathway prevails in each pulling direction.

How can we now reconcile our results of apparently similar folding pathways with clearly distinct unfolding pathways? Microreversibility requires the pathways for folding and unfolding in thermodynamic equilibrium to be identical for every given pulling geometry. Consequently, if unfolding pathways differ among the pulling directions, folding pathways have to be different as well. Even though the unfolded states are different for the three pulling directions, with different parts of the polypeptide chain being subjected to force, we have evidence that the unfolded states are energetically equivalent. Our analysis using the Crooks fluctuation theorem (see the Supporting Information) yields folding free energies for DomN 17 and 38 that are in good agreement with the value that we determined earlier for DomN 1.[7] It is important to note that in this analysis, the trivial differences in free energy of the unfolded state resulting from stretching differently long entropic chains are corrected for using polymer elasticity models. Furthermore, non-native interactions in the unfolded state are often weak compared to thermal energy scales.[19]

Therefore, the identical zero force folding rates we find for the three mechanically imposed pathways also require identical zero-force unfolding rates for the three pulling directions. We are now able to test this prediction using Monte Carlo simulations. Within experimental error, the experimental data for all three pathways can be fit using an identical zero-force unfolding rate $k_u = 0.001$ s^{-1}, by varying only Δx_{N-TS} (black lines in Figure 3a). As a control, we show in Figure 3b that these parameters also correctly reproduce the unfolding force distributions measured at 250 nm s^{-1}. Thus, whereas the folding/unfolding pathways are clearly structurally different for the three pulling directions studied, the extrapolated zero-force folding/unfolding rate always has the same value. Figure 3c shows a schematic depiction of the folding free energy landscape for the three pulling directions with the barrier height being the same for every direction. If we consider the structure of DomN (Figure 3d), we may find a clue for the differences in potential width between the three pulling directions. Although the short potential widths of DomN 17 and DomN 38 correspond to geometries in which force is applied directly to the hydrophobic core of DomN (purple in Figure 3d), the high potential width of DomN 1 corresponds to a force application point at the end of an α helix that leads away from the hydrophobic core. Thus, the protein might appear "softer", with a higher Δx_{N-TS}, along this direction in which the α helix can be deformed before the full structure of DomN breaks down. In a simple analogy, the N-terminal α helix of DomN might serve as a lever, increasing the distance and reducing the force necessary for unfolding of DomN 1.

In the three pulling geometries studied, the transition states are likely to be very different from a structural point of view, as they are polarized along different coordinates. The observation that the same folding and unfolding kinetics prevails in all geometries is a direct experimental indication for a broad, energetically degenerate transition state ensemble, as has been postulated for fast folding proteins.[20,21]

In conclusion, we have used single-molecule force spectroscopy to populate selectively structurally distinct pathways to probe their folding/unfolding energetics. Instead of the previously postulated pathway switch towards a single zero force pathway with a higher unfolding rate,[10] we observe different pathways of similar energetics. Our results indicate that under zero-force conditions, a multitude of energetically equivalent pathways dominate the folding and unfolding of CaM. The low degree of residual frustration and the relatively high conformational entropy of the transition state that would prevail in such a scenario might be a prerequisite for ultrafast protein folding on the microsecond timescale.[22] Furthermore, these results indicate that close to equilibrium, extrapolated zero-force rates from single-molecule force spectroscopy could be comparable to those obtained with chemical denaturants, despite different pathways. A different folding behavior might apply to slower folding proteins with better-defined transition states. Investigating whether the robustness of barrier heights in single-molecule force spectroscopy experiments is a common feature of small α-helical

proteins close to equilibrium will be an important task for future research.

Received: October 13, 2009
Revised: December 3, 2009
Published online: March 29, 2010

Keywords: calmodulin · force spectroscopy · protein folding · protein structures · single-molecule studies

[1] A. F. Oberhauser, M. Carrion-Vazquez, *J. Biol. Chem.* **2008**, *283*, 6617.
[2] A. Borgia, P. M. Williams, J. Clarke, *Annu. Rev. Biochem.* **2008**, *77*, 101.
[3] J. M. Fernandez, P. E. Marszalek, H. Lu, H. Li, M. Carrion-Vazquez, A. F. Oberhauser, K. Schulten, *Nature* **1999**, *402*, 100.
[4] C. Cecconi, E. A. Shank, C. Bustamante, S. Marqusee, *Science* **2005**, *309*, 2057.
[5] M. Kessler, K. E. Gottschalk, H. Janovjak, D. J. Muller, H. E. Gaub, *J. Mol. Biol.* **2006**, *357*, 644.
[6] Q. Peng, H. Li, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1885.
[7] J. P. Junker, F. Ziegler, M. Rief, *Science* **2009**, *323*, 633.
[8] M. Schlierf, M. Rief, *Angew. Chem.* **2009**, *121*, 835; *Angew. Chem. Int. Ed.* **2009**, *48*, 820.
[9] M. Bertz, A. Kunfermann, M. Rief, *Angew. Chem.* **2008**, *120*, 8314; *Angew. Chem. Int. Ed.* **2008**, *47*, 8192.
[10] P. M. Williams, S. B. Fowler, R. B. Best, J. L. Toca Herrera, K. A. Scott, A. Steward, J. Clarke, *Nature* **2003**, *422*, 446.
[11] J. P. Junker, M. Rief, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 14361.
[12] H. Dietz, F. Berkemeier, M. Bertz, M. Rief, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 12724.
[13] H. Dietz, M. Bertz, M. Schlierf, F. Berkemeier, T. Bornschlogl, J. P. Junker, M. Rief, *Nat. Protoc.* **2006**, *1*, 80.
[14] D. J. Brockwell, E. Paci, R. C. Zinober, G. S. Beddard, P. D. Olmsted, D. A. Smith, R. N. Perham, S. E. Radford, *Nat. Struct. Biol.* **2003**, *10*, 731.
[15] G. E. Crooks, *J. Stat. Phys.* **1998**, *90*, 1481.
[16] M. Schlierf, F. Berkemeier, M. Rief, *Biophys. J.* **2007**, *93*, 3989.
[17] R. B. Best, G. Hummer, *J. Am. Chem. Soc.* **2008**, *130*, 3706.
[18] E. Evans, K. Ritchie, *Biophys. J.* **1997**, *72*, 1541.
[19] E. Paci, M. Vendruscolo, M. Karplus, *Proteins* **2002**, *47*, 379.
[20] D. Thirumalai, C. Hyeon, *Biochemistry* **2005**, *44*, 4957.
[21] V. Munoz, *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 395.
[22] A. Akmal, V. Munoz, *Proteins* **2004**, *57*, 142.

