The Complex Folding Network of Single Calmodulin Molecules

Article *in* Science · October 2011

DOI: 10.1126/science.1207598 · Source: PubMed

5 authors, including:

Johannes Stigler
Ludwig-Maximilians-University of Munich
13 PUBLICATIONS 3,015 CITATIONS

SEE PROFILE

Fabian Ziegler
Invensity GmbH
7 PUBLICATIONS 461 CITATIONS

SEE PROFILE

network, and four through trades with other teams) (*17*). Similarly, Geocacher's strategy was based on the existing community of geocaching, a sport based on using navigational techniques to hide and seek objects. It also created a burst by announcing its participation to the geocacher community and located seven correct balloons. DeciNena aimed at assembling a balloon-hunting team by posting their participation on every related blog on the Internet to gain attention, but they failed to achieve a wide-range response. DeciNena found seven balloons at the end of the competition.

Although Hotz and Geocacher were able to create a sudden response peak by efficiently propagating the news to an existing audience, this response was very short-lived. On the other hand, our strategy was able to sustain social response for a longer period, stretching up until the end of the competition. This happened despite not having access to a large community of followers. Instead, the MIT team started with only four people; and after a couple of days, twitter response achieved a number comparable with that of Hotz, who started with 35,000 existing followers. Another interesting observation is that after the competition, when mass media came to report the winning story of the MIT team the tweet count actually decreased instead of increasing. This suggests that the incentives provided by the MIT strategy played a dominant role in generating Twitter response, rather than the "MIT brand" and mass media effect (SOM text).

The recursive incentive mechanism has a number of desirable properties. First, the recursive incentive mechanism is never in deficit—it never exceeds its budget (SOM text). After being recruited by a friend, an individual has no incentive to create his own root node by visiting the Balloon Challenge Web page directly (without using the link provided by the recruiter). This follows from the fact that payment to the person finding the balloon does not depend on the length of the chain of recruiters leading to him.

However, the mechanism is not resistant to false name attacks, which were originally identified in the context of Web-based auctions (*32*). In this attack, which has been shown to plague powerful economic mechanisms (*32*), an individual creates multiple false identities in order to gain an unfair advantage. Having said that, our data does not reveal any successful incidents of false-name attacks. This may be due to the fact that the mechanism did not operate for long enough for people to identify this potential, and that actual payment requires social security numbers. In practice, other measures could be put in place to minimize or detect this kind of attack (*33*).

The mechanism's success can be attributed to its ability to provide incentives for individuals to both reports on found balloon locations while simultaneously participating in the dissemination of information about the cause. When an individual finds a balloon, the individual can either report the balloon to us, to other teams, or attempt to find the other nine balloons and win the DARPA prize directly. In practice, it is unlikely for an unprepared individual to find other balloons (and if they replicated our mechanism, their delayed start would always leave them behind). Proofs are in the SOM.

Our mechanism simultaneously provides incentives for participation and for recruiting more individuals to the cause. This mechanism can be applied in very different contexts, such as social mobilization to fight world hunger, in games of cooperation and prediction, and for marketing campaigns.

References and Notes

1. J. Howe, *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business* (Three Rivers Press, New York, 2009).
2. E. Hand, *Nature* **466**, 685 (2010).
3. L. von Ahn, *Computer* **39**, 92 (2006).
4. S. Cooper *et al.*, *Nature* **466**, 756 (2010).
5. J. Pontin, *New York Times*, Artificial intelligence, with help from the humans (25 March 2007); available at www.nytimes.com/2007/03/25/business/yourmoney/25Stream.html.
6. K. J. Arrow *et al.*, *Science* **320**, 877 (2008).
7. B. A. Huberman, D. M. Romero, F. Wu, *J. Inf. Sci.* **35**, 758 (2009).
8. G. Krakow, "Ham radio operators to the rescue after Katrina: Amateur radio networks help victims of the hurricane," www.msnbc.msn.com/id/9228945 (2005).
9. S. Milgram, *Psychol. Today* **1**, 6067 (1967).
10. J. M. Kleinberg, *Nature* **406**, 845 (2000).
11. D. J. Watts, P. S. Dodds, M. E. J. Newman, *Science* **296**, 1302 (2002).
12. L. A. Adamic, E. Adar, *Soc. Networks* **27**, 187 (2005).
13. M. Rosvall, P. Minnhagen, K. Sneppen, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **71**, 66111 (2005).
14. P. S. Dodds, R. Muhamad, D. J. Watts, *Science* **301**, 827 (2003).
15. W. Mason, D. J. Watts, *Proceedings of the ACM SIGKDD Workshop on Human Computation* (ACM, New York, 2009), pp. 77–85.
16. Defense Advanced Research Projects Agency, DARPA Network Challenge (accessed May 2010); available at http://networkchallenge.darpa.mil.
17. Defense Advanced Research Projects Agency, DARPA Network Challenge Project Report (February 16, 2010); available at www.hsdl.org/?view&did=17522.
18. Defense Advanced Research Projects Agency, "MIT Red Balloon Team wins DARPA Network Challenge" (press release, 5 December 2009).
19. J. C. Tang *et al.*, *Commun. ACM* **54**, 78 (2010).
20. D. J. Watts, J. Peretti, *Harv. Bus. Rev.* **May**, 22 (2007).
21. J. Kleinberg, P. Raghavan, in *Proccedings of 46th Annual IEEE Symposium on FOCS* (IEEE, Los Alamitos, CA, 2005), pp. 132–141.
22. E. Biyalogorsky, E. Gerstner, B. Libai, *Mark. Sci.* **20**, 82 (2001).
23. J. Iribarren, E. Moro, *Phys. Rev. Lett.* **103**, 038702 (2009).
24. D. J. Watts, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5766 (2002).
25. D. Liben-Nowell, J. Kleinberg, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 4633 (2008).
26. B. Golub, M. O. Jackson, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 10833 (2010).
27. Our data does not include the time stamp of sending out invitations, but we were able to measure the intervals between actual sign-up events between a parent and its children in the trees.
28. A. L. Barabási, *Nature* **435**, 207 (2005).
29. R. D. Malmgren, D. B. Stouffer, A. S. Campanharo, L. A. Amaral, *Science* **325**, 1696 (2009).
30. D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, A. Tomkins, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11623 (2005).
31. J. Yang, J. Leskovec, *Proceedings of the 4th ACM International Conference on Web Search and Data Mining* (ACM, Kowloon, Hong Kong, 2011).
32. M. Yokoo, Y. Sakurai, S. Matsubara, *Games Econ. Behav.* **46**, 174 (2004).
33. D. Whitworth, Man fined over fake eBay auctions. *BBC Newsbeat* (5 July 2010); available at www.bbc.co.uk/newsbeat/10508913.

Acknowledgments: This research was partially sponsored by the Army Research Laboratory under cooperative agreement W911NF-09-2-0053, by the Air Force Office of Scientific Research under award FA9550-10-1-0122, by the U.S. Air Force under contract FA8721-05-C-0002, by the National Science Foundation under grant 0905645, by the Masdar Institute of Science and Technology under grant 400075, the Branco Weiss Science in Society Program, and by generous gifts from P&G, Best Buy, and Google Views. Conclusions in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the sponsors. After providing the data underlying Fig. 4, Twitter requested that Stanford University withdraw this data from public access. Please contact A.P. for information regarding use of the data on the MIT server.

Supporting Online Material

www.sciencemag.org/cgi/content/full/334/6055/509/DC1
Materials and Methods
SOM Text
Figs. S1 to S5

21 March 2011; accepted 15 September 2011
10.1126/science.1205869

The Complex Folding Network of Single Calmodulin Molecules

Johannes Stigler,[1] Fabian Ziegler,[1] Anja Gieseke,[1] J. Christof M. Gebhardt,[1]* Matthias Rief[1,2]†

Direct observation of the detailed conformational fluctuations of a single protein molecule en route to its folded state has so far been realized only in silico. We have used single-molecule force spectroscopy to study the folding transitions of single calmodulin molecules. High-resolution optical tweezers assays in combination with hidden Markov analysis reveal a complex network of on- and off-pathway intermediates. Cooperative and anticooperative interactions across domain boundaries can be observed directly. The folding network involves four intermediates. Two off-pathway intermediates exhibit non-native interdomain interactions and compete with the ultrafast productive folding pathway.

The energy landscape view provides a conceptual framework for understanding protein folding (*1*, *2*). However, the diversity in size and structure of the proteome is far too large to provide a single generic mechanism for how proteins fold. Deciphering specific mechanisms

for more complex proteins requires an experimental method that provides both structural and dynamic information at the level of a single molecule. In silico, the use of supercomputers and distributed computing has allowed visualization of putative folding pathways of simple proteins covering time spans of up to milliseconds (*3*–*5*). Single-molecule methods offer a possibility to track the time evolution of protein conformations (*6*–*11*). Force experiments using optical tweezers have been used to directly observe protein conformational transitions close to equilibrium with time resolution in the submillisecond range (*12*–*14*). The main observable in such experiments is protein length, a one-dimensional parameter, and hence most biomolecular folding processes investigated so far exhibit simple linear reaction coordinates (*14*–*16*). In this study, we use ultrastable high-resolution optical tweezers to study the complex fluctuations of a full-length calmodulin molecule between its folded and unfolded states covering time spans from hundreds of microseconds to tens of minutes.

The calcium-sensing protein calmodulin is a key regulator of calcium-dependent signal cascades. In the calcium-bound form, calmodulin consists of two globular domains connected by an α-helical linker (Fig. 1A). Each of the two globular head domains consists of two helix-loop-helix motifs (EF hands), each binding one calcium ion. Extensive investigation through classical bulk kinetic experiments has shown that the two globular domains can fold independently in a rapid two-state process (*17*). Many studies have elucidated properties of the isolated domains (*11*, *18*, *19*). However, complexity can arise in the folding of multidomain proteins (*20*, *21*). Even though equilibrium folding studies have indicated domain-domain interactions in calmodulin (*22*), a kinetic network for folding of the two-domain protein has been missing.

We inserted calmodulin molecules between two ubiquitin proteins, each carrying a terminal cysteine residue, which allowed for attachment of DNA handles that, in turn, were connected to 1-μm beads through antibody-antigen interactions (*23*) (Fig. 1A). Ubiquitin typically unfolds at forces much higher than those used in this study (*24*). Experiments were carried out at 10 mM Ca^{2+} conditions, where calmodulin is known to fold stably (*11*). We used an optical tweezers setup (*25*) to repeatedly unfold (black trace) and refold (blue trace) a single calmodulin molecule (Fig. 1A). Two major unfolding and refolding peaks can be observed. Both spacing and unfolding force are consistent with a single globular domain unfolding in each of the peaks. Those results confirm earlier atomic force microscopy experiments where individual calmodulin domains were shown to unfold at 10 to 15 pN (*11*). However, the rapid oscillations in the upper trace of Fig. 1A provide the first indication of deviations from a simple two-step unfolding behavior. An experimental protocol using silica beads instead of the commonly used polystyrene beads reduced light-induced oxygen damage (*26*) substantially (*25*). This enabled us to routinely observe folding-unfolding fluctuations of single calmodulin molecules for up to 45 min. In this assay, neither extension nor force is actively controlled (*27*). The molecule is stretched to a preset force value (pretension) around which it then fluctuates (*25*). A sample trace covering 5 min for an individual calmodulin molecule held at a pretension of ~11 pN is shown in Fig. 1B. A complex pattern of rapid fluctuations can be observed. An expanded view reveals six different populated states (marked in colors in the bottom trace of Fig. 1B). Single-molecule mechanics can provide structural information when the length of the unfolded polypeptide chain is measured with subnanometer precision (*28*). The unfolded state is marked in red, and the fully folded state is marked in purple. We find that three states (light blue, green, and orange) cluster around a length expected for a state with two EF hands folded and two EF hands unfolded (*29*). Even though these three states have similar lengths (see fig. S1), their kinetics are different which identifies them as distinct states. Although the orange and light blue levels exhibit similar lifetimes, the orange state always transitions toward the red unfolded state, whereas the light blue state transitions almost exclusively toward the dark blue state. The extension of the dark blue state corresponds to three folded EF hands. For an objective classification of states, we performed a hidden Markov analysis, optimizing for position and single exponentiality of lifetime distributions (fig. S2) (*25*). We find that at least the six distinct states marked in Fig. 1B are required. A five-state model pooling the short-lived orange state, together with either the light blue or the green state, results in double exponential lifetime distributions (figs. S3 and S4).

Molecular deletion constructs were used to decipher the structures underlying the various states (Fig. 1C). In addition to the wild type (WT) (Fig. 1C, left trace), we investigated calmodulin constructs with the first EF hand deleted (CaM-234, middle left trace) and the last EF hand deleted (CaM-123, middle right trace), as well as a construct lacking the first and last EF hands (CaM-23, right trace). Apart from the unfolded state (U, in red), CaM-234 only exhibits states with two EF hands folded (green and orange). CaM-34 with the two N-terminal EF hands deleted exhibits a state identical in extension and kinetics to the green state (see fig. S5B). Hence, the green state can be identified as a state where the C-terminal domain containing EF hands 3 and 4 is fully folded (F_{34}). With CaM-123, we observe a more complex pattern of states. In this construct, a stably folded state identical to the dark blue state can be observed with all three EF hands folded (F_{123}). The presence of this state is surprising, as it has been assumed that calmodulin EF hands only fold in pairs (*30*). This intermediate state suggests cooperative effects across the classically assumed domain boundaries in calmodulin. Even though complete calmodulin domains can be reconstituted by mixing individual EF-hand polypeptides in solution, single EF hands have not been observed to fold independently (*19*). We find that the isolated EF hand 3 can fold, given an interface to a fully folded N-terminal domain. In support of this interpretation, a state in which three calcium ions are bound has been found in mutant calmodulin (*31*, *32*).

State F_{123} rapidly exchanges with a state with only two EF hands folded (light blue). Consequently, this state corresponds to F_{12}. The kinetics of the corresponding folded state of CaM-12 (fig. S5A), as well as its measured length (28.3 nm from the unfolded state), support this conclusion. F_{123} cannot directly fold to the native state. Instead, productive folding toward the native is only possible from F_{12} (Fig. 1D), which identifies F_{123} as an off-pathway intermediate. Even though structural information about F_{123} does not exist, we anticipate that EF hand 3 will be packed onto F_{12} in a non-native conformation.

The nature of the short-lived orange state present in the WT and the deletion constructs becomes clear in the light of the data measured with CaM-23. This construct exclusively transitions between an unfolded state and the orange state and exhibits the length of two folded EF hands. A direct comparison of the lifetimes of this state in CaM-23 and the WT protein confirms its identity (fig. S6). Apparently, along its folding pathway, calmodulin can populate a non-native intermediate with mispaired EF hands belonging to two different domains (F_{23}). Even though such a state has so far not been implicated in calmodulin folding, nuclear magnetic resonance (NMR) experiments have shown that the EF hands in CaM-23 can form a paired domain with a structure similar to that of the native N- and C-terminal domains (*33*). F_{23} can also be identified as an off-pathway state because it never folds toward the native state. This is in accord with the NMR structure of this intermediate, as further folding would require breaking the non-native interface between EF hands 2 and 3. Independent evidence for the assignment of the structures underlying the various observed intermediates was obtained using a cross-link mutant within EF hand 4 (table S1 and fig. S7) (*25*).

The extensive equilibrium fluctuation data now allowed construction of detailed possible transition paths between the various states. The structures of the intermediates identified above impose significant constraints on the possible pathways. The off-pathway intermediate F_{23} can exclusively transition to the unfolded state, whereas the intermediate F_{123} only exchanges with F_{12} (*25*). Moreover, we never observed direct transitions between F_{12} and F_{34}, which is expected, because such transitions would involve unfolding

[1]Physik Department E22, Technische Universität München, James-Franck-Strasse, 85748 Garching, Germany. [2]Munich Center for Integrated Protein Science, 81377 München, Germany.

*Present address: Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA.

†To whom correspondence should be addressed. E-mail: mrief@ph.tum.de

of one structural part and simultaneous refolding of others. Direct and cooperative folding of the two C-terminal EF hands is in agreement with the cooperative folding observed for isolated EF hand pairs (fig. S5). Taken together, the network of states between the unfolded and the folded state contains six distinct states connected by six transition paths with a total of 12 rate constants.

We used force to control and investigate the population and transition paths in the calmodulin folding landscape (Fig. 2A). At low pretensions of 7.6 pN (Fig. 2A, top), folded or largely folded states dominated. The more tension we applied, the more unfolded states were populated, until, at 11.6 pN, the unfolded state prevailed. The relative probabilities of the different states are shown in Fig. 2C. From these curves, the free energy of all states at zero force can be calculated (Table 1) (*25*). The observation time of tens of minutes allowed us to detect states down to a relative population of 10^{-4}. The free energy data summarized in Table 1 reveal a significant anticooperativity between the folding of the two calmodulin domains. From the fully unfolded state, folding of

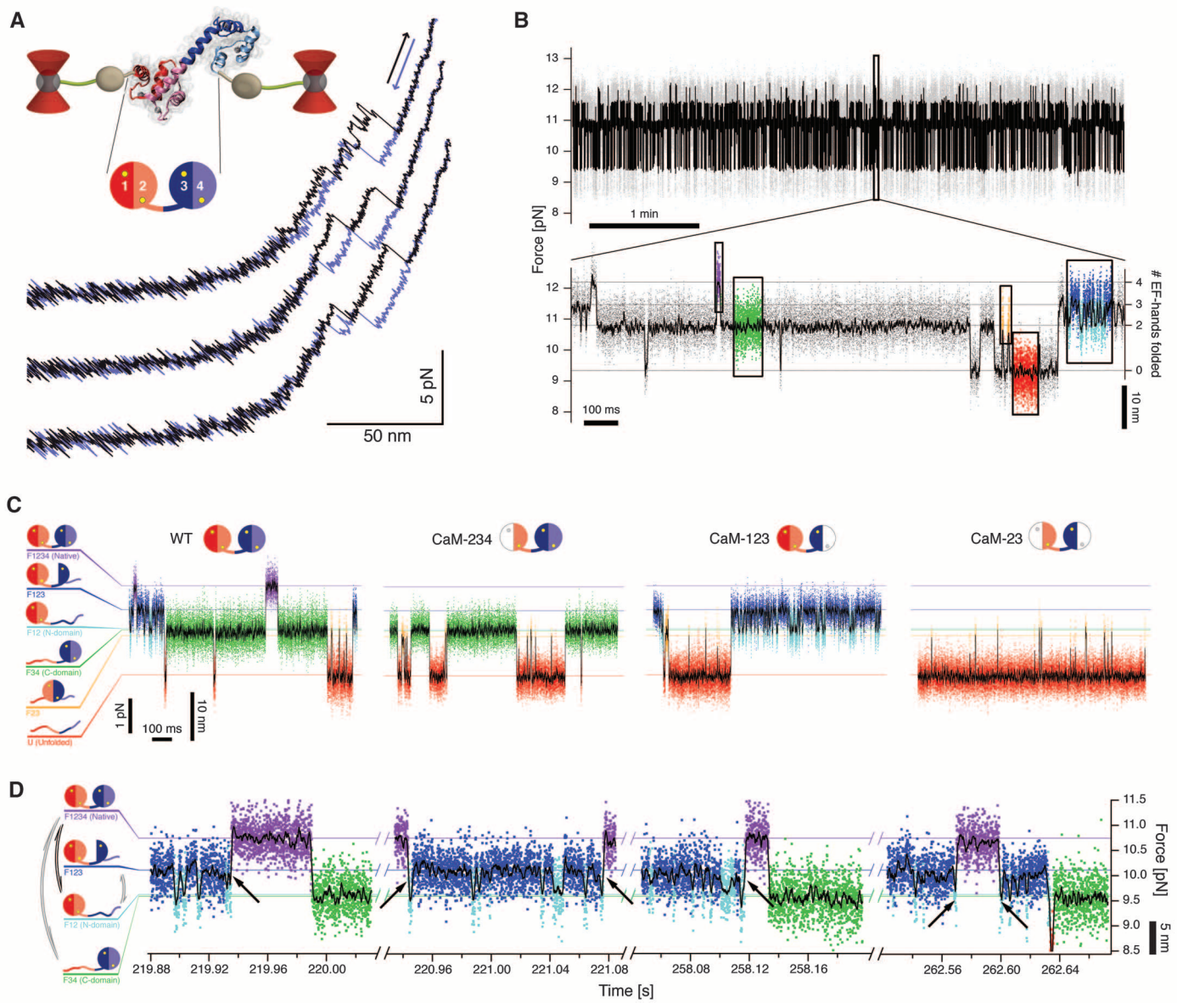

Fig. 1. (**A**) Folding and unfolding of single CaM molecules. Representative stretch-and-relax cycles for WT-CaM at velocity v = 500 nm/s. (Inset) Sketch of the experimental setup with the protein linked with ubiquitin-DNA handles to functionalized silica beads. (Expanded section) Cartoon representation of CaM with the EF hands numbered. (**B**) Sample trace during 5 min of the fluctuations of a single WT-CaM molecule at a constant trap separation. The data are shown at full resolution (gray) and low-pass filtered (black). The vertical scale denotes the force acting on the molecule as measured by the deflection of the beads from the trap center. (Expanded section) Six different states (see colored regions) can be identified using hidden Markov modeling. (**C**) State assignment and comparison of WT traces with truncation mutants. Shown are representative parts of traces for WT-CaM, CaM-234, CaM-123, and CaM-23 at similar pretensions. (Left) The proposed structural assignment to the identified levels (see text). Data points are colored according to the assignment by hidden Markov modeling (*25*). Color key: purple, State F_{1234} (native); dark blue, $F_{123;}$ light blue, F_{12} (N domain folded); green, F_{34} (C domain folded); orange, F_{23}; and red, unfolded state (U). The distance information in the graphs reflects the distance, the trapped beads move upon a transition. From this information, the contour length change upon transition can be calculated. The sample trace of WT-CaM plotted as contour length change versus time can be found in fig. S14. (**D**) Transitions to and from the native state F_{1234}. Direct transitions to and from the native state (purple) can only occur to and from states F_{34} (green) or F_{12} (light blue). Transitions of F_{123} (dark blue) to and from F_{1234} always occur through F_{12} (see arrows), identifying F_{123} as an off-pathway intermediate. Arrows indicate representative events where F_{1234} is populated from and decays into F_{12}.

either the N-terminal (F_{12}) or the C-terminal (F_{34}) domain results in an energetic gain of ~20 k_BT (where k_B is Boltzmann's constant and T is temperature). However, folding of a second domain to the native state only provides 15 k_BT. Apparently, the presence of one folded domain prevents the other domain from reaching its energetically optimal state. Energetic coupling between the two calmodulin domains has been reported on the basis of bulk measurements that suggested that the N-terminal domain was destabilized by a folded C-terminal domain (*22*). At 10 mM Ca^{2+}, a large part of the free energy gain on calmodulin folding is mediated by calcium binding. From the concentration of free calcium, as well as the known stability of apo calmodulin and calcium binding constants, the expected free energy can be calculated to 19.7 ± 0.8 k_BT for the N-terminal domain and 21.8 ± 0.6 k_BT for the C-terminal domain (*22*, *34*). To test the influence of Ca^{2+} concentration on the folding free energies of the various substates, we also performed measurements at a lower Ca^{2+} concentration of 100 μM (Table 1). Again, the measured values agree with calculated values 10.5 ± 0.8 k_BT and 12.6 ± 0.6 k_BT for the N-terminal and C-terminal domains, respectively. Apparently, even though the unfolded state in force experiments is different from the unfolded state in bulk assays, the measured free energies are close, which indicates that our description of the energetics of the stretched state is accurate.

Also, the complete kinetic information about the folding network is contained in the data of Fig. 2A. Dwell-time distributions at four different pretensions are given in Fig. 2B and fig. S2. Every distribution closely follows a single exponential (fit curves), which supports our interpretation that each of the six states is distinct. Because we can directly measure the probability for a given state decaying into any of the other possible states, all 12 possible transition rates can be obtained (*25*). The force dependence of all rates is given in fig. S8. Similar plots, describing probability distributions as well as the kinetics of the truncation mutants, corroborate the structural assignments for the various states (figs. S5 and S9 to S11).

The kinetic network extrapolated to zero force for calmodulin folding at 10 mM Ca^{2+} that emerged from our single-molecule data is summarized in Fig. 3 and in Tables 1 and 2. From the ensemble of unfolded states, the calmodulin polypeptide chain can proceed along three different and roughly equally populated pathways (*35*). One pathway is a dead end, when EF hands 2 and 3 form a non-native folded pair. For productive folding, this state has to decay back to the unfolded state. The second and third pathways toward the folded state proceed via F_{12} or F_{34}, with either the N- or C-terminal domain completely folded. From F_{34}, simultaneous cooperative folding of EF hands 1 and 2 is required to reach the native state F_{1234}. From F_{12}, a productive pathway with EF hands 3 and 4 folding to the native state competes with a nonproductive pathway toward the intermediate F_{123}. The presence of the off-pathway intermediates can slow full-length calmodulin folding considerably, as compared with the folding of its subdomains. This effect can be seen in the top trace of Fig. 2A, where there is an accumulation of the intermediate F_{123} (blue state); at low forces, unfolding to F_{12}, a prerequisite to further productive folding, becomes increasingly unlikely. This effect can also be seen in stretch and relax cycles (see fig. S12), where folding to the native state occurs much more robustly against higher forces if folding

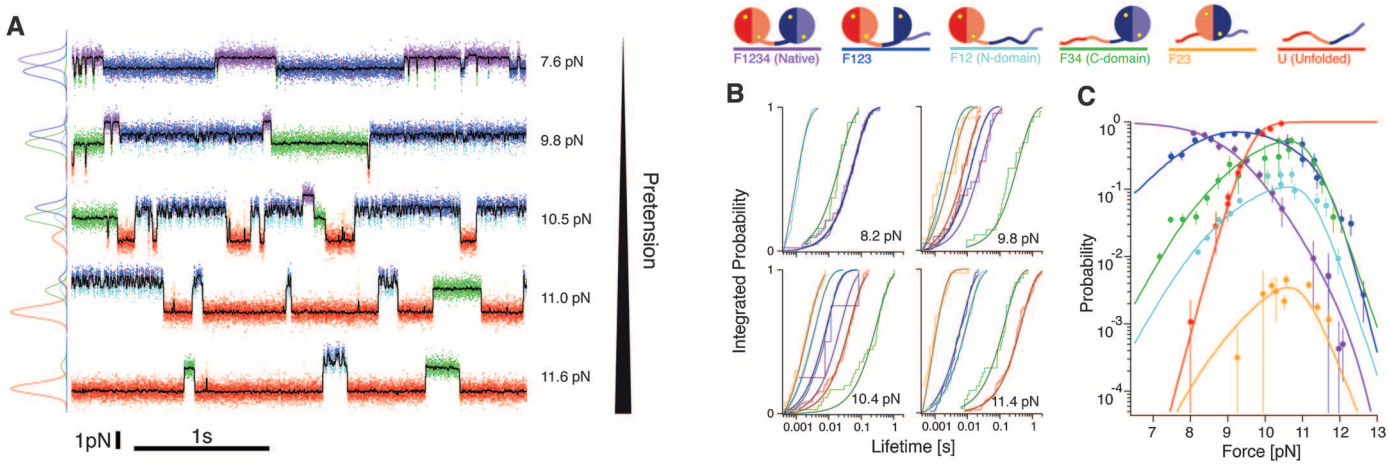

Fig. 2. (**A**) Traces of WT-CaM at different pretensions (color key as in Fig. 1C). (Left) Gaussian fits to histograms of each respective state. (**B**) Integrated lifetime histograms of WT-CaM at different pretensions. The colors correspond to the respective states. The continuous lines are fits to a single-exponential model. (**C**) Force-dependence of the occupation probability for all states. The colors correspond to the respective states. Probabilities are plotted against the force of the system when it is in the corresponding state. Note that, in this representation, the probabilities at a given force do not add up exactly to 1 because experiments are conducted at nonconstant force [see Eq. *7* in SOM]. A global fit to all data points (continuous lines) allows for the determination of equilibrium energies.

Table 1. Energies and contour lengths of WT-CaM and the truncation mutants at different Ca^{2+} concentrations. Both Ca^{2+} concentrations are above the equilibrium constants for calcium binding to calmodulin. ΔG_0 has experimental uncertainty of 10%. ΔL is the contour length increase for WT-CaM and shows the difference from state F_{1234}; errors are given as 1σ intervals. ΔL_{calc} is the calculated contour length increase for WT-CaM and shows the difference from state F_{1234}. For F_{12} and F_{34}, a symmetric splitting between the domains was assumed. Dash indicates that the state does not exist for the respective mutant.

	Difference from U state, ΔG_0 (k_BT)									
	10 mM Ca^{2+}						100 μM Ca^{2+}			
State	WT	CaM-12	CaM-34	CaM-23	CaM-123	CaM-234	WT	CaM-23	ΔL (nm)	ΔL_{calc} (nm)
U									52.2 ± 0.6	50.6
F_{23}	13	–	–	12	13	13	4	4	27.4 ± 0.7	27.4
F_{34} (C domain)	21	–	18	–	–	20	11	–	23.8 ± 0.5	25.7
F_{12} (N domain)	20	19	–	–	20	–	11	–	23.3 ± 0.4	25.3
F_{123}	30	–	–	–	28	–	14	–	13.0 ± 0.3	13.2
F_{1234} (native)	36	–	–	–	–	–	17	–	0	0

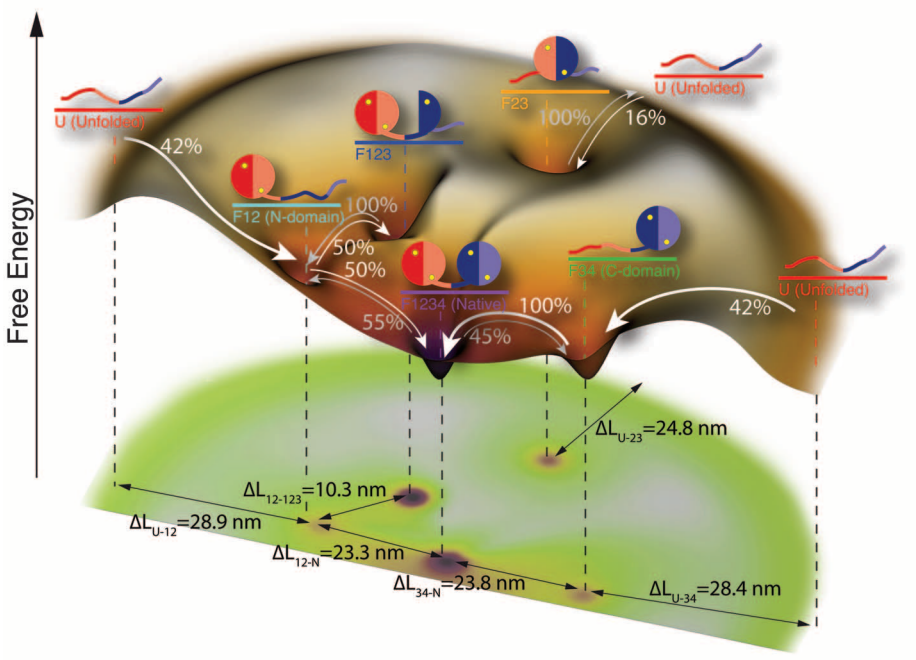

Fig. 3. Full kinetic network of WT-CaM folding and unfolding at zero load. Arrows show all observed transitions. The percentage values provided for each transition give the fraction of transitions along the respective pathways out of each state. Distances in the lower part are differences in contour length.

Table 2. Rate fit parameters for WT-CaM at 10 mM Ca^{2+}. Errors are given as 1σ intervals. Columns 1 and 3 are based on a fit to the Bell model; see Eq. 12 in SOM (*25*). Columns 2 and 4 are based on a model that includes polypeptide elasticity; see Eq. 13 in SOM (*25*). Δx_{unf} is the change in the length required to reach the transition state of unfolding. ΔL_{fold} is the contour length change required to reach the transition state of folding.

Transition	$\log_{10}(k_{0,unf})$ (s^{-1})	$\log_{10}(k_{0,fold})$ (s^{-1})	Δx_{unf} (nm)	ΔL_{fold} (nm)
$F_{1234} \rightleftharpoons F_{12}$	−0.7 ± 0.9	5.0 ± 0.4	1.3 ± 0.8	14.4 ± 1.3
$F_{1234} \rightleftharpoons F_{34}$	−0.8 ± 0.3	5.6 ± 0.2	1.7 ± 0.3	16.9 ± 0.8
$F_{123} \rightleftharpoons F_{12}$	−0.13 ± 0.04	5.0 ± 0.2	1.92 ± 0.04	7.2 ± 0.6
$F_{12} \rightleftharpoons U$	−5.0 ± 0.7	5.8 ± 0.5	5.0 ± 0.6	18.2 ± 1.5
$F_{34} \rightleftharpoons U$	−4.1 ± 0.5	5.8 ± 0.5	4.1 ± 0.4	17.3 ± 1.5
$F_{23} \rightleftharpoons U$	−1.4 ± 0.3	5.4 ± 0.3	3.7 ± 0.3	15.6 ± 1.0

starts from F_{34}. If folding starts from F_{123}, refolding to the native state is slowed. The occasional population of the two off-pathway intermediates has a drastic impact on the overall folding rate of calmodulin (fig. S13). Although the individual domains fold within microseconds (Table 2), the average folding time for the complete molecule in the absence of load is as long as 5 s. The presence of non-native EF hand pairing apparently makes calmodulin an overall slow-folding protein with ultrafast folding domains. The picture of calmodulin folding that emerges from our data is strikingly similar to earlier predictions of kinetic partitioning of protein folding among different pathways (*36*). Kinetic partitioning may provide a general framework for the folding of multidomain proteins.

Will calmodulin undergo a similarly complicated folding pathway in a cell? Even though the gradual translation at the ribosome may prevent the nascent chain of calmodulin from misfolding, during its lifetime, calmodulin will undergo multiple folding and unfolding cycles. Cellular calcium concentrations are generally so low that its C-terminal domain will be unfolded with about a 25% probability (*22*).

Our view of how proteins fold has always been influenced by the information and level of detail that experiments were able to provide. Hence, for a long time, many proteins were identified as simple two-state folders. With the advent of more sophisticated time-resolved kinetic techniques, the role of intermediates for folding has received increasing attention. Single-molecule experiments make it now possible to observe many of the proposed mechanisms working in concert. Even though the structural resolution in our assay is one-dimensional, simultaneous availability of kinetic information covering many orders of magnitude allows one to decipher and to distinguish several parallel pathways and their associated structures.

References and Notes

1. J. N. Onuchic, P. G. Wolynes, *Curr. Opin. Struct. Biol.* **14**, 70 (2004).
2. D. Thirumalai, C. Hyeon, *Biochemistry* **44**, 4957 (2005).
3. D. E. Shaw *et al.*, *Science* **330**, 341 (2010).
4. P. L. Freddolino, F. Liu, M. Gruebele, K. Schulten, *Biophys. J.* **94**, L75 (2008).
5. V. A. Voelz, G. R. Bowman, K. Beauchamp, V. S. Pande, *J. Am. Chem. Soc.* **132**, 1526 (2010).
6. B. Schuler, W. A. Eaton, *Curr. Opin. Struct. Biol.* **18**, 16 (2008).
7. E. M. Puchner *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 13385 (2008).
8. H. Li, A. F. Oberhauser, S. B. Fowler, J. Clarke, J. M. Fernandez, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6527 (2000).
9. M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, H. E. Gaub, *Science* **276**, 1109 (1997).
10. P. Zheng, Y. Cao, T. Bu, S. K. Straus, H. Li, *Biophys. J.* **100**, 1534 (2011).
11. J. P. Junker, F. Ziegler, M. Rief, *Science* **323**, 633 (2009).
12. E. A. Shank, C. Cecconi, J. W. Dill, S. Marqusee, C. Bustamante, *Nature* **465**, 637 (2010).
13. C. Cecconi, E. A. Shank, C. Bustamante, S. Marqusee, *Science* **309**, 2057 (2005).
14. J. C. M. Gebhardt, T. Bornschlögl, M. Rief, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 2013 (2010).
15. J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco Jr., C. Bustamante, *Science* **292**, 733 (2001).
16. M. T. Woodside *et al.*, *Science* **314**, 1001 (2006).
17. C.-R. Rabl, S. R. Martin, E. Neumann, P. M. Bayley, *Biophys. Chem.* **101-102**, 553 (2002).
18. J. P. Junker, M. Rief, *Angew. Chem. Int. Ed. Engl.* **49**, 3306 (2010).
19. C. F. Shuman, R. Jiji, K. S. Åkerfeldt, S. Linse, *J. Mol. Biol.* **358**, 870 (2006).
20. M. B. Borgia *et al.*, *Nature* **474**, 662 (2011).
21. A. F. Oberhauser, P. E. Marszalek, M. Carrion-Vazquez, J. M. Fernandez, *Nat. Struct. Biol.* **6**, 1025 (1999).
22. L. Masino, S. R. Martin, P. M. Bayley, *Protein Sci.* **9**, 1519 (2000).
23. C. Cecconi, E. A. Shank, F. W. Dahlquist, S. Marqusee, C. Bustamante, *Eur. Biophys. J.* **37**, 729 (2008).
24. M. Schlierf, H. Li, J. M. Fernandez, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7299 (2004).
25. Materials and methods are available as supporting material on *Science* Online.
26. M. P. Landry, P. M. McCall, Z. Qi, Y. R. Chemla, *Biophys. J.* **97**, 2128 (2009).
27. M. de Messieres, B. Brawn-Cinani, A. La Porta, *Biophys. J.* **100**, 2736 (2011).
28. H. Dietz, M. Rief, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 1244 (2006).
29. At a pretension of 11 pN, the expected length difference from the fully folded state to the fully unfolded state of calmodulin is 19.3 nm, if one assumes a persistence length of 0.5 nm and a contour length gain of 0.365 nm per amino acid residue.
30. J. Krebs, C. W. Heizmann, *New Compr. Biochem.* **41**, 51 (2007).
31. S. Fefeu *et al.*, *Biochemistry* **39**, 15920 (2000).
32. K. Nakashima, H. Ishida, S. Y. Ohki, K. Hikichi, M. Yazawa, *Biochemistry* **38**, 98 (1999).
33. T. M. Lakowski, G. M. Lee, M. Okon, R. E. Reid, L. P. McIntosh, *Protein Sci.* **16**, 1119 (2007).
34. Y.-G. Chen, G. Hummer, *J. Am. Chem. Soc.* **129**, 2414 (2007).
35. Note that the folding rates of all transitions lie in a range of 10^5 to 10^6 s^{-1}. This may be explained by considering that, in all those transitions, binding of calcium ions dominates the folding process.
36. Z. Guo, D. Thirumalai, *Biopolymers* **36**, 83 (1995).

Acknowledgments: This work was supported by an SFB 863 A2 grant of Deutsche Forschungsgemeinschaft. M.R. acknowledges financial support through an instrument grant of the Institute for Advanced Study of Technische Universität München. J.S. was supported by the International Graduate School "Materials Science of Complex Interfaces."

Supporting Online Material

www.sciencemag.org/cgi/content/full/334/6055/512/DC1
Materials and Methods
Figs. S1 to S18
Table S1
References (*37*–*46*)

27 April 2011; accepted 9 September 2011
10.1126/science.1207598