



Supporting Online Material for

The Complex Folding Network of Single Calmodulin Molecules

Johannes Stigler, Fabian Ziegler, Anja Gieseke, J Christof M Gebhardt, Matthias Rief*

*To whom correspondence should be addressed. E-mail: mrrief@ph.tum.de

Published 28 October 2011, *Science* **334**, 512 (2011)
DOI: 10.1126/science.1207598

This PDF file includes

Materials and Methods
Figs. S1 to S18
Table S1
Full References

Materials and Methods

Molecular cloning

All constructs were inserted between two ubiquitins with terminal cysteines by molecular cloning and expressed in *E. coli*. The complete amino acid sequence of the construct WT-CaM was

MACKMQIFVKLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIFAGKQLED
GRTLSDYNIQKESTLHLVLRLRGGELADQLTEEQIAEFKEAFSLFDKDGDTITTK
ELGTVMRSLGQNPTEAELQDMINEVDADGNGTIDFPEFLTMMARKMKDTSSEE
EIREAFRVFDKDGNGYISAAELRHVMTNLGEKLTDEEVDEMIREADIDGDGQVN
YEEFVQMMTAKGTMQIFVKLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQR
LIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGKCLE-His₆. Underlined parts
denote sequences that were deleted for the constructs CaM-123, CaM-234 and CaM-23.
For the crosslinked mutant CaM-128/144, residues 128 and 144 of calmodulin were
replaced by cysteines.

The complete sequence of the construct CaM-12 was

MACKMQIFVKLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIFAGKQLED
GRTLSDYNIQKESTLHLVLRLRGGELADQLTEEQIAEFKEAFSLFDKDGDTITTK
ELGTVMRSLGQNPTEAELQDMINEVDADGNGTIDFPEFLTMMARKMKGTMQIF
VKLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIFAGKQLEDGRTLSDYNI
QKESTLHLVLRLRGGKCLE-His₆, the sequence for CaM-34 was
MACKMQIFVKLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIFAGKQLED
GRTLSDYNIQKESTLHLVLRLRGGELKMKDTSSEEEIREAFRVFDKDGNGYISAA
ELRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFVQMMTAKGTMQIFVK
TLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIFAGKQLEDGRTLSDYNIQK
ESTLHLVLRLRGGKCL-His₆.

Protein-DNA construct formation

For attachment to functionalized DNA handles, we used a protocol similar to a protocol described by Cecconi et al. (23). In brief, the cysteines were activated by DTDP and mixed with 34 bp ssDNA oligos with TCEP-activated 3' thiol groups. DNA linkers of 370 nm length were generated by PCR using a lambda-phage template. The forward primers were a mixture of equal amounts of biotin and digoxigenin modified oligos. A reverse primer with an introduced abasic site was designed to leave a ssDNA overhang complementary to the protein-bound oligos.

Measurement procedure

Protein with attached oligos was mixed with the DNA linkers and incubated with silica beads (1 µm diameter, Polysciences), in-house functionalized with covalently bound anti-digoxigenin Fab fragments (Roche). These constructs were diluted in sample buffer (50 mM Tris, 150 mM KCl, 10 mM or 100 µM CaCl₂, 26 U/ml glucose oxidase, 17000 U/ml catalase, 0.65% glucose, pH 8.0) and subsequently mixed with streptavidin coated silica beads (1 µm diameter, Bangs Laboratories, Inc.). The concentrations were adjusted such that the protein-oligo construct was only sparsely distributed on the beads and multiply tethered dumbbells were negligibly rare. Beads were brought into close

proximity and tested for tether formation through stretch-and-relax cycles. Successfully tethered dumbbells were then held at a constant trap separation. In this configuration, the force is not constant and changes significantly upon folding or unfolding. The pretension force was after Hidden-Markov analysis (see below) assigned as the force of the green state (F_{34}).

Optical tweezers setup

The setup we used was a dual beam optical tweezers setup with back focal plane detection and one steerable beam using an AOD, based on a setup described earlier (14). Calibration of beads was performed using a technique introduced by Tolić-Nørrelykke et al. (37) and all relevant corrections to the power spectrum (38). The error of trap stiffness determination was approximately 10%. Trap stiffnesses between different experiments varied between 0.25-0.30 pN/nm. Data were collected at 100 kHz and averaged to 20 kHz before storage. The signals were corrected for both crosstalk due to depolarization of the beams and the proximity of the beams. Final analysis was performed on the difference of both bead signals to increase the signal to noise ratio (39).

Data analysis

Linker and polypeptide elasticity

The elasticity of the linker in a low force regime, where the protein is still folded, was modeled using an extensible worm like chain model (40). In this model, the force is given by

$$F_{\text{eWLC}}(d_D) = \frac{k_B T}{p_D} \left(\frac{1}{4 \left(1 - \frac{x}{L_D} + \frac{F}{K} \right)^2} - \frac{1}{4} + \frac{x}{L_D} - \frac{F}{K} \right) \quad (1)$$

with persistence length p_D , contour length L_D , elastic modulus K and extension d_D . The persistence lengths obtained in this assay were between 20–25 nm, the contour length in the expected range of ca. 370 nm and the elastic modulus in the range of 600–1000 pN.

To account for additional compliances from unfolded protein, the linker elasticity was considered in series with a worm-like chain model (41)

$$F_{\text{WLC}}(d_p) = \frac{k_B T}{p_p} \left(\frac{1}{4 \left(1 - \frac{x}{L_p} \right)^2} - \frac{1}{4} + \frac{x}{L_D} \right) \quad (2)$$

with polypeptide persistence length p_p , contour length and protein extension d_p . In our measurements, we used a fixed persistence length of 0.5 nm.

Contour length differences

The contour length differences were obtained both from fits to force-extension curves as well as from constant-distance information. Both methods gave consistent results. We reported the values obtained from constant-distance information in Table 2.

Theoretical values were calculated based on the dumbbell structure of calmodulin (PDB code 1CLL). The difference in contour length between a folded state I and an (partially) unfolded state II is given by $\Delta l = 0.365 \cdot \Delta a - d'_I + d'_{II}$ where Δa is the difference in number of unfolded amino acids between state I and state II, d'_I is the extension of the folded state and d'_{II} is the extension of the remaining folded part of the unfolded state.

The calculations of contour lengths are based on the crystal structure of the wildtype. Since the folding intermediates in our experiments may deviate from this structure, errors in those calculated lengths may occur. The precise number of amino acids folded in the intermediates is unknown. This may explain the slight deviations in length measured for F34 and F12.

The shown time-dependent force trajectories were analyzed using a Hidden-Markov classifier (see below). These traces can be converted to unfolded contour length using equations (1) and (2). An example is shown in Fig. S14.

Hidden Markov analysis

In order to identify states, a Hidden-Markov analysis (42) was performed on the 20 kHz raw data of the difference signal. First, the raw data was coarse grained into typically 400-500 bins. Initial level positions were picked from histograms of the raw data and the emission values of the model were initialized with Gaussian representations of the states. Iterations involved one pass of the Forward-Backward-Algorithm. Subsequently, the emission probabilities were re-estimated from the maximum state probabilities of each data point. The emission probabilities were not constrained to Gaussian shapes. Iterations were repeated until only negligible numbers of data points (typically less than 0.08%) were reclassified in each iteration. The performance of the algorithm was checked by comparing the resulting lifetime distributions with single exponentials (Fig. S2). The transition probability matrix was adjusted manually to yield optimal lifetime distributions. Even though the states F₂₃, F₁₂ and F₃₄ all consist of two folded EF-hands, separation was possible due to different lifetime distributions and small differences in length (Fig. S1).

A five-state model for WT CaM always resulted in inconsistencies and non-single exponential lifetime distributions. An example for a five-state model applied to WT CaM is shown in Fig. S4.

Further, it is important to note that a robust classification of data points into states is difficult without assuming the connectivity of the network of states. This connectivity has to describe the data at all pretensions. Even though the length difference between, for instance, F₂₃ and F₁₂ is small and at intermediate pretensions their lifetimes are very similar, they are readily separable at higher pretensions (see Fig. S14). The connectivity information then allows us to enhance the separation of the states even at lower pretensions (Fig. S15).

To confirm the good performance of the Hidden Markov classifier we performed Brownian Dynamics simulations of our assay (description below) and compared the

results of a Hidden Markov analysis of such traces with the known hidden state of the protein at each time point. We find that the algorithm classifies >98% of the points correctly (see Fig. S16).

The distributions of lifetimes for the different states at comparable pretensions correspond very well to those measured in truncation mutants, exemplified in Fig. S6.

Brownian Dynamics simulations

We simulated the thermal movement of both beads in their traps using Brownian Dynamics (43). Mimicking the experimental setup, the two beads were connected with a linker consisting of DNA, modeled by an extensible worm-like chain model (Eq. (1)), in series with a worm-like chain term (Eq. (2)) with a contour length corresponding to the state of the protein at each time step. A Monte Carlo generator was used to decide on the state of the protein at each time step. The transition probabilities used for the Monte Carlo generator were calculated from the rates determined in the experiment (Table 1). Upon each transition, the contour length of the worm-like chain term was adapted. During the simulation, the response of the two beads was followed using Brownian Dynamics. Like in the experiments, data points were taken with 100 kHz and subsequently sampled down to 20 kHz. After the simulation of a full trajectory, the signal was calculated from the difference signal of the two beads. An example for such a simulated trace showing both the (hidden) state of the simulated protein and the Brownian response of the beads is shown in Fig. S16.

Point spread function

Due to the nonlinearity of the linker the data points belonging to a specific state do not follow a Gaussian distribution. In a first order approximation, histograms could be fit with the following equation (see Fig. S1):

$$p(F) = A \cdot \exp\left(-\left(\frac{F - \mu + \alpha(F - \mu)^2}{s}\right)^2\right) \quad (3)$$

Force-dependent probabilities and determination of equilibrium energies

Probabilities for each state were calculated directly as the sum of all lifetimes for a respective state divided by the trace length. Since the accuracy of this estimation is limited by the finite measurement time, the uncertainties were estimated by Monte-Carlo simulations (44). In brief, an ensemble of traces was generated based on the measured transition rates. The standard deviation of probabilities obtained from this trace ensemble was assigned as probability error.

The free energy of the complete system $G_i(F_i)$ at a force F_i is given by

$$G_i(F_i) = G_i^0 + G_i^{\text{device}}(F_i) = G_i^0 + G^{\text{bead}}(F_i) + G^{\text{linker}}(F_i) + G_i^{\text{p}}(F_i), \quad (4)$$

where G_i^0 is the free energy of the protein in state i , $G^{\text{bead}}(F) = \frac{1}{2}k(F)^2$, F is the energy stored in the deflection x of the beads from the trap center, G^{linker} is the energy stored in

the stretching of the linker and G_i^p is the free energy of the unfolded polypeptide. G^{linker} and G_i^p are readily calculated as integrals over WLC and eWLC functions (equations (1) and (2)).

When the system undergoes a transition from state i to state j , the force changes from F_i to F_j . The energy difference between the states is then given by

$$\Delta G_{ij}(F_i, F_j) = \Delta G_{ij}^0 + \Delta G_{ij}^{\text{device}}(F_i, F_j) = G_j(F_j) - G_i(F_i). \quad (5)$$

Since the probabilities are related to energies according to

$$\frac{P_j(F_j)}{P_i(F_i)} = \exp\left(-\frac{\Delta G_{ij}(F_i, F_j)}{k_B T}\right), \quad (6)$$

we can obtain the energy differences ΔG_{ij}^0 of the protein between states i and j by performing a global fit to the probability data with

$$P_i(F) = \frac{1}{1 + \sum_{j \neq i} \exp\left(-\frac{\Delta G_{ij}^0 + \Delta G_{ij}^{\text{device}}(F, F_j)}{k_B T}\right)} \quad (7)$$

and weights equal to the inverse of the errors calculated as described before.

The experimental uncertainty of ΔG_{ij}^0 is dominated by the calibration error of the trap stiffness and is ca. 10%.

Transition rates

For transition rate calculation, the off-rates of all levels were obtained from single exponential fits to the lifetime-distributions of the states obtained by HMM analysis. The fits were applied to normalized integrated lifetimes and took into account that events shorter than a dead time τ_{\min} or longer than τ_{\max} could not be observed:

$$p(t) = \frac{\exp(-k^{\text{off}} t) - \exp(-k^{\text{off}} \tau_{\min})}{\exp(-k^{\text{off}} \tau_{\max}) - \exp(-k^{\text{off}} \tau_{\min})} \quad (8)$$

In our experiments, we chose τ_{\max} to be the length of the trajectory while τ_{\min} depended on the applied pretension and was in the range of 200 μs to 800 μs .

The off-rate k_n^{off} from a given state n is equal to the sum of transition rates into all other states:

$$k_n^{\text{off}} = \sum_{j \neq n} k_{nj} \quad (9)$$

Furthermore, the ratio between two transition rates can be calculated as

$$\frac{k_{nj}}{k_{ni}} = \frac{T_{nj}}{T_{ni}}, \quad (10)$$

where T_{nj} is the number of transitions observed from state n to state j .

Taken together, transition rates k_{ni} from state n to state i can be calculated according to

$$k_{ni} = \frac{k_n^{\text{off}}}{1 + \sum_{j \neq i} \frac{T_{nj}}{T_{ni}}}. \quad (11)$$

For determining rates at zero force and distances to the transition state, the unfolding rates were fitted with a Bell model:

$$k_{\text{unf}}(F) = k_{0,\text{unf}} \exp\left(-\frac{F \cdot \Delta x_{\text{unf}}}{k_B T}\right) \quad (12)$$

Folding rates were fitted with a model accounting for the energy differences in the DNA linker and beads deflection between unfolded state i and transition state T (45):

$$k_{\text{fold}}(F) = k_{0,\text{fold}} \exp\left(-\frac{\Delta G_{iT}(F_i, F_T)}{k_B T}\right) \quad (13)$$

For zero-force folding and unfolding rates, it is important to note that we extrapolate over a force range (7 pN) that is much smaller than usual in AFM experiments (several hundreds of pN). Nevertheless, we believe the errors can approach an order of magnitude. What gives us confidence in the choice of models we use for extrapolation is the fact that the equilibrium free energies as calculated by $k_B T \cdot \ln(k_{\text{fold}}/k_{\text{unf}})$ lie close to the measured ΔG values obtained from equilibrium populations of states.

Calculation of folding times

The distribution of folding times displayed in Fig. S13 was calculated from the zero-force extrapolated rates k_{ij}^0 given in Table 1 using Monte Carlo simulations. Starting from the unfolded state, the system was allowed to transition through its kinetic network. For each transition, the average dwell time $\bar{\tau}_i$ was calculated from the inverse off rate (Eq. (9)):

$$\bar{\tau}_i = \frac{1}{\sum_{j \neq i} k_{ij}^0} \quad (14)$$

The dwell time τ for a particular transition was drawn from an exponential distribution:

$$p(\tau) = \exp\left(-\frac{\tau}{\tau_i}\right) \quad (15)$$

The transition probabilities p_{ij} to the next state were calculated according to

$$p_{ij} = \begin{cases} \frac{k_{ij}^0}{\sum_{j \neq i} k_{ij}^0}, & j \neq i \\ 0, & j = i \end{cases} \quad (16)$$

and the subsequent state chosen randomly according to these probabilities.

Along the trajectory through the kinetic network, the dwell times were added up until the system reached the folded state. For the shown histogram, 10000 such simulations were performed.

The off-pathway nature of F_{23} and F_{123}

Concerning the off-pathway nature of F_{123} , in addition to the evidence described in the main text, more statistical evidence for the off-pathway nature of F_{123} can be found when counting the number of transitions that occur between the native state F_{1234} and either F_{12} or F_{123} . For F_{123} being off-pathway, only a negligible number of transitions between F_{123} and F_{1234} should be observed. However, since at low pretensions F_{12} is very short-lived, we are likely to miss a fraction of the F_{12} states and hence observe some direct transitions between F_{123} and F_{1234} while in reality they occur from F_{123} via F_{12} to F_{1234} (or vice versa). Given a temporal resolution of 400 μ s we, for instance at a pretension of 9.6 pN, expect to find 28% of the transitions occurring between F_{123} and F_{1234} via F_{12} to be falsely classified as direct transitions between F_{123} and F_{1234} (for an example see, for instance, the simulated trace shown in Fig. S16). Experimentally, we find 30%. In fact, at none of the measured pretensions a number of transitions between F_{123} and F_{1234} not in statistical agreement with the expected number due to missed events was found.

In order to quantify the off-pathway nature of F_{23} , we counted the number of such short-lived events (irrespective of the classification by the Hidden Markov model) that exhibit a lifetime shorter than 3 times the lifetime of F_{23} , which then proceeded further to the native state. We find that only 6 transitions took this route to the native state. Comparing this number to the overall 302 events that led into the native state strongly confirms the off-pathway nature of F_{23} . Additional evidence for the off-pathway nature of F_{23} comes from averaging the rising slopes of the transitions leading into each of the intermediate states (Fig. S18). If F_{23} were an obligatory on-pathway state, we would expect this state to be populated in transitions leading from the unfolded state to F_{12} and/or F_{34} . In this case, we would expect a rapid rise of the averaged F_{12} and F_{34} traces to the position of F_{23} and then an exponentially smeared out transition into the final levels of

F_{12} or F_{34} . The time-scale of the exponential smearing should be the lifetime of F_{23} . Such an analysis had been done to detect intermediate states in molecular motors (46). In contrast, we find that the transitions proceed rapidly (within the first two data points) into the levels F_{12} and F_{34} without intermediate dwells on the shorter F_{23} level.

The cross-linked mutant CaM 128/144

The crosslink in the EF hand 4 between residues 128 and 144 results in a shortening of the stretchable polypeptide chain by 16 amino acid residues corresponding to 5.8 nm of contour length. This shortening will affect all levels where EF hand 4 is unfolded, i.e. F_{12} , F_{123} , F_{23} and U. In the crosslink mutant almost all transitions from the unfolded level proceed into F_{34} because the shortened chain makes this transition less sensitive to force and hence it will be faster than the competing transitions into F_{12} or F_{23} . Therefore extension values for F_{23} could not be measured in the crosslink mutant.

A sample trace with the shift of the levels comparing WT CaM and CaM-128/144 is shown in Fig. S7. A table comparing the measured contour length increases between WT CaM and CaM-128/144 is given in Table S1.

Fig. S1

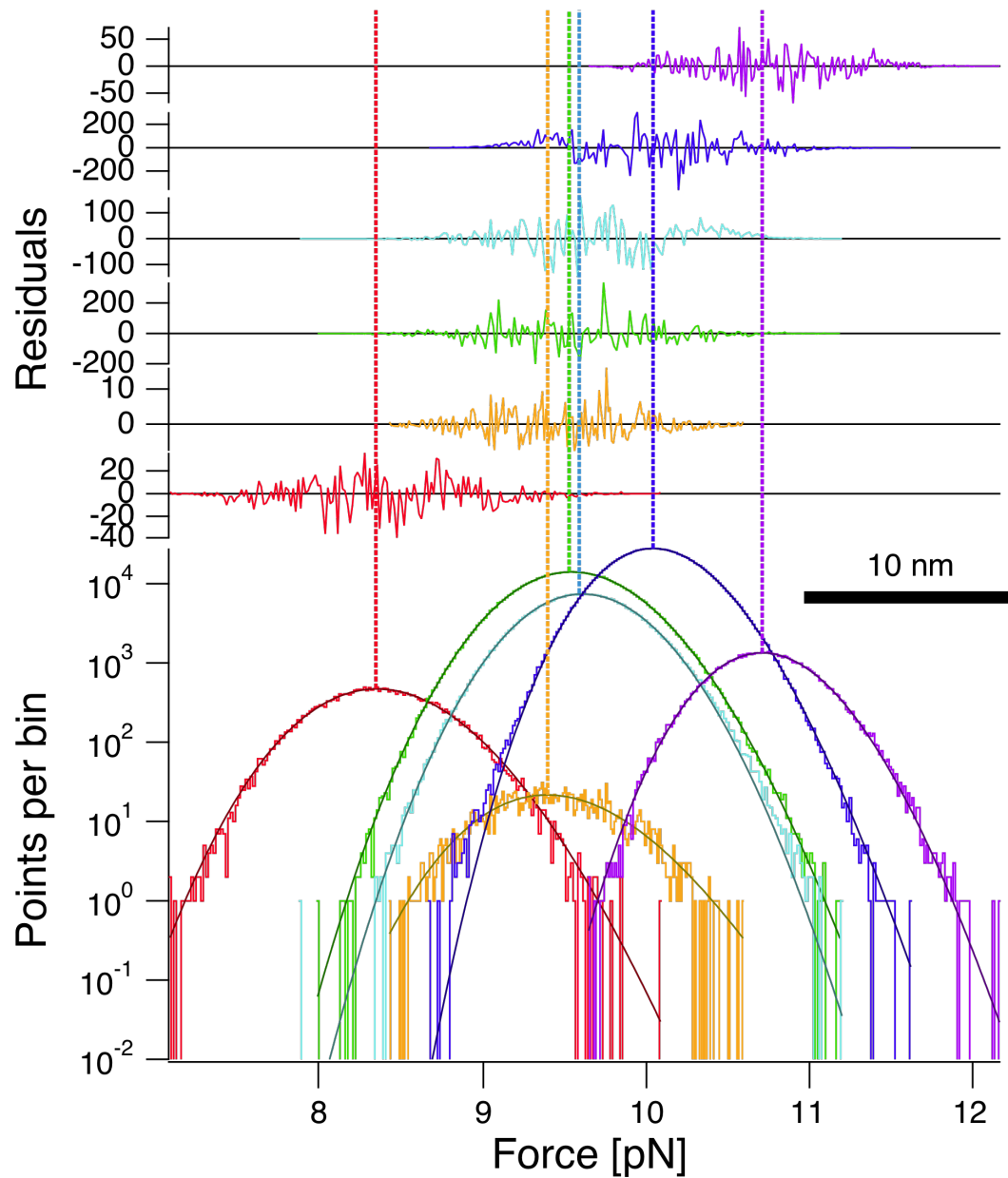


Fig. S1. Histograms of data points classified into each of the six states for WT-CaM at a pretension of 9.6 pN. Due to the nonlinear linker, the distributions are skewed and were fit with a phenomenological skewed Gaussian (Eq. (3)). Colors correspond to the coding given in Fig. 1C.

Fig. S2

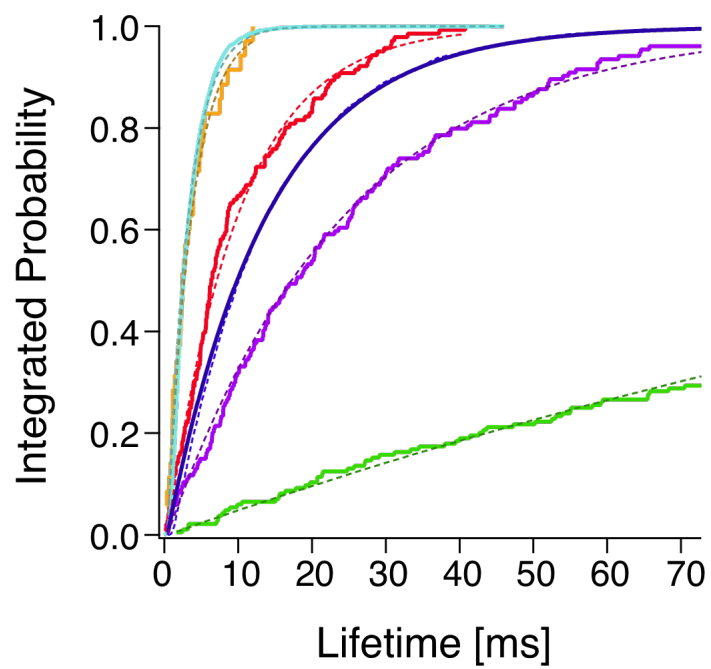


Fig. S2. Integrated lifetime histograms of all states at a pretension of 9.6 pN. The colors correspond to the coding of Fig. 1C. The dashed lines are single exponential fits.

Fig. S3

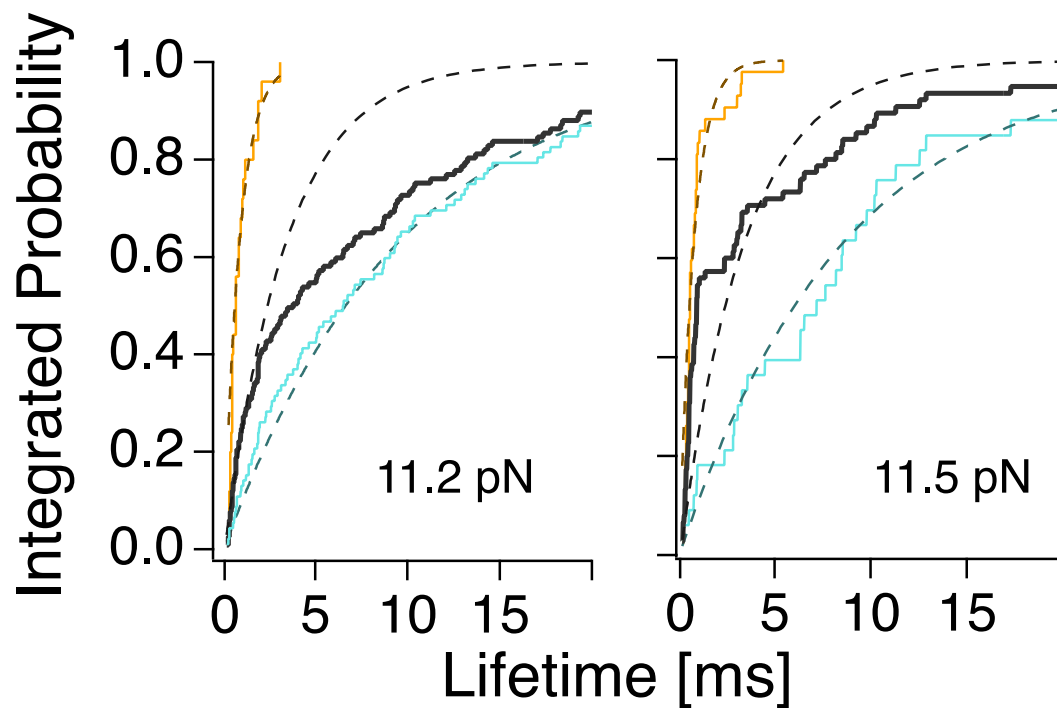


Fig. S3. Lifetime histograms of state F_{23} (orange) and F_{12} (bright blue) and single exponential fits (dashed lines) at different high pretensions. At the shown pretensions F_{23} and F_{12} are clearly distinguishable by their lifetime. A hypothetical lifetime histogram if F_{23} and F_{12} were actually the same state is shown in grey. It does not follow a single exponential distribution (shown in the dashed black lines).

Fig. S4

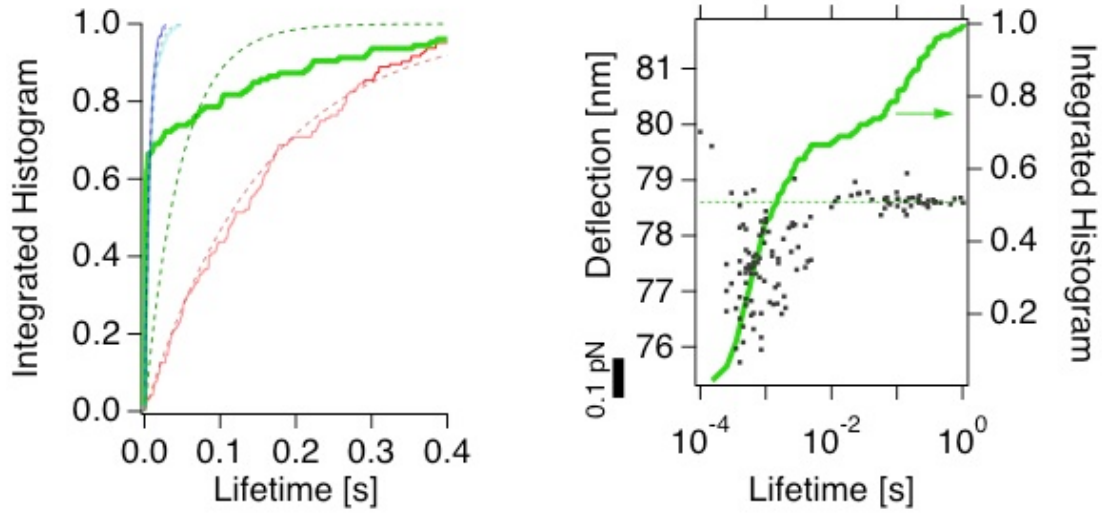


Fig. S4. Left: Integrated lifetime histograms when a 5-state model is applied to the data. While the model performs well finding the states F₁₂ (bright blue), F₁₂₃ (dark blue) and U (red), it pools the state F₂₃ into the class for F₃₄ (green), yielding a clearly double exponential distribution of lifetimes. Continuous lines are lifetimes extracted from the model, dashed lines are single exponential fits. Note that at the shown pretension of 11 pN the folded state F₁₂₃₄ is not populated.

Right: Scatter plot of the data points classified into the class for F₃₄ by a 5-state model (grey points). The levels with longer life times (integrated lifetime histogram in green) scatter in deflection around the level of state F₃₄ (dashed green line), while the falsely classified levels actually belonging to class F₂₃ with shorter lifetimes scatter around a lower deflection value. Note that due to statistical effects (square root of N errors), the shorter-lived states scatter more in deflection.

Fig. S5

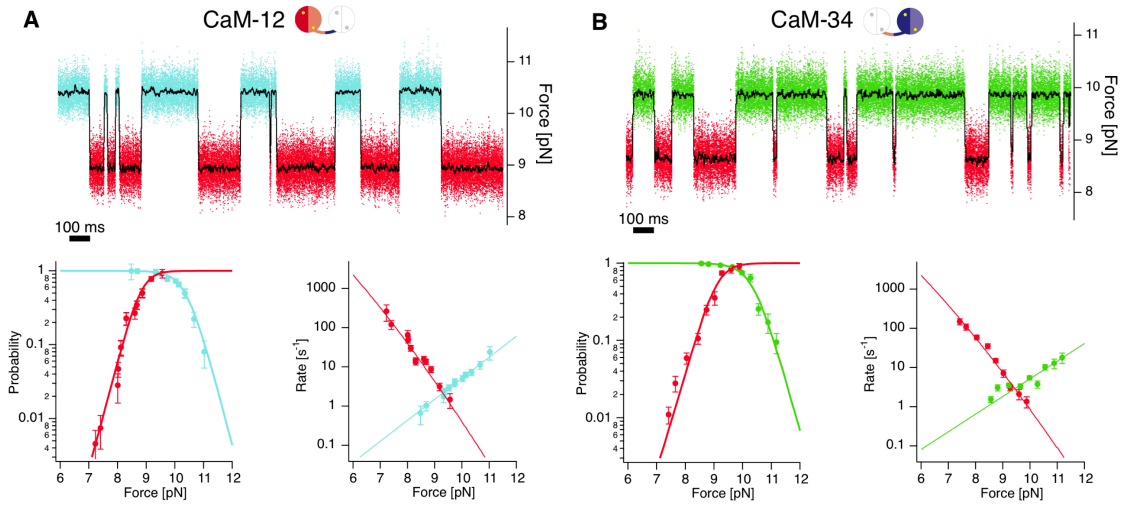


Fig. S5.

(A) Top: Sample trace for the N-terminal domain of calmodulin (CaM-12). Bottom left: Occupation probabilities for the two states of CaM-12 plotted against the force of the system in the corresponding state. Note that in this representation the probabilities at a given force do not add up exactly to one, because experiments are conducted at non-constant force (see equation (7) in SOM). The continuous lines are a global fit to all datapoints. Bottom right: Rate dependence of the transitions and fit. The colors correspond to the originating state in the transition (color code given in Fig. 1C).

(B) Top: Sample trace for the C-terminal domain of calmodulin (CaM-34). Bottom left: Occupation probabilities for the two states of CaM-34 plotted against the force of the system in the corresponding state. Note that in this representation the probabilities at a given force do not add up exactly to one because experiments are conducted at non-constant force (see equation (7) in SOM). The continuous lines are a global fit to all datapoints. Bottom right: Rate dependence of the transitions and fit. The colors correspond to the originating state in the transition (color code given in Fig. 1C).

Fig. S6

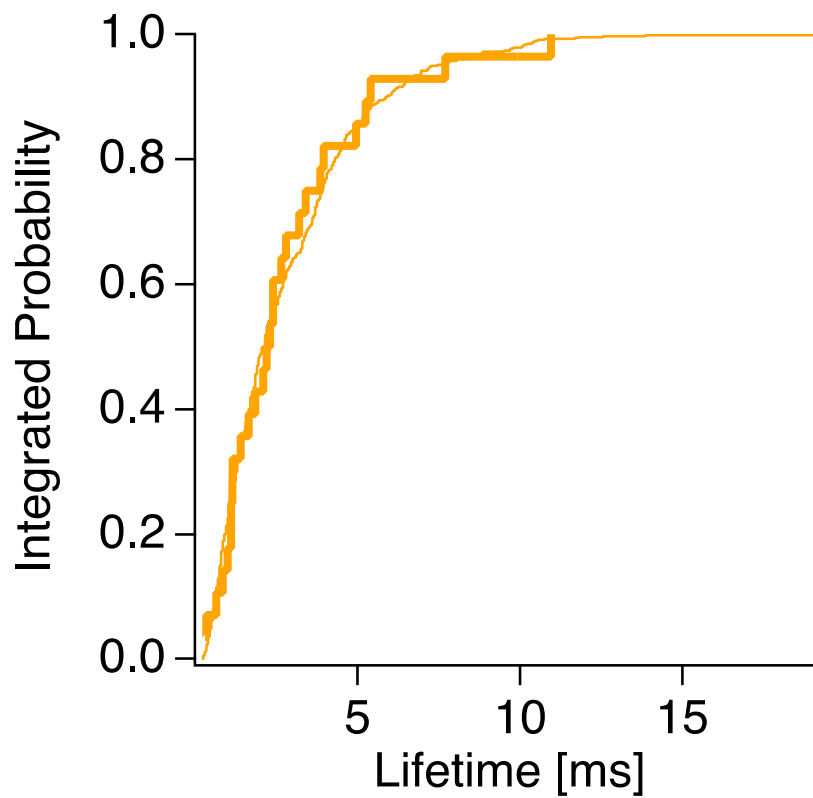


Fig. S6. Integrated lifetime histograms of F_{23} in WT CaM (thick line) and CaM-23 (thin line) at comparable pretensions. Both curves exhibit identical lifetimes.

Fig. S7

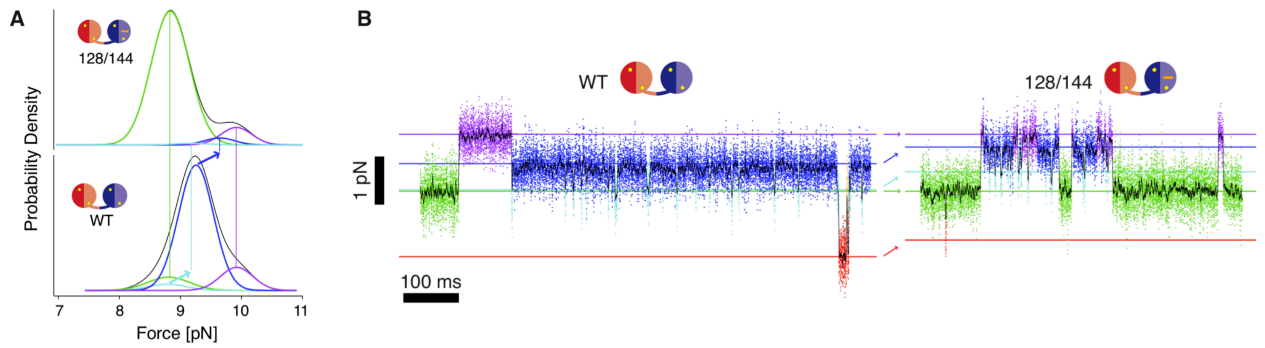


Fig. S7. Comparison between wildtype calmodulin and the crosslinked mutant CaM-128/144 at comparable pretensions.

- (A) Histogram of points classified into states F_{1234} (purple), F_{123} (dark blue), F_{12} (bright blue) and F_{34} (green) and their sum (black). The center values for F_{12} and F_{123} are shifted in CaM-128/144 as indicated by arrows.
- (B) Sample traces for wildtype calmodulin (left) and CaM-128/144 (right) at similar forces. Due to the crosslink the levels U, F_{12} and F_{123} are shifted (arrows). F_{23} is not detectable in CaM-128/144 (see respective paragraph in the SOM).

Fig. S8

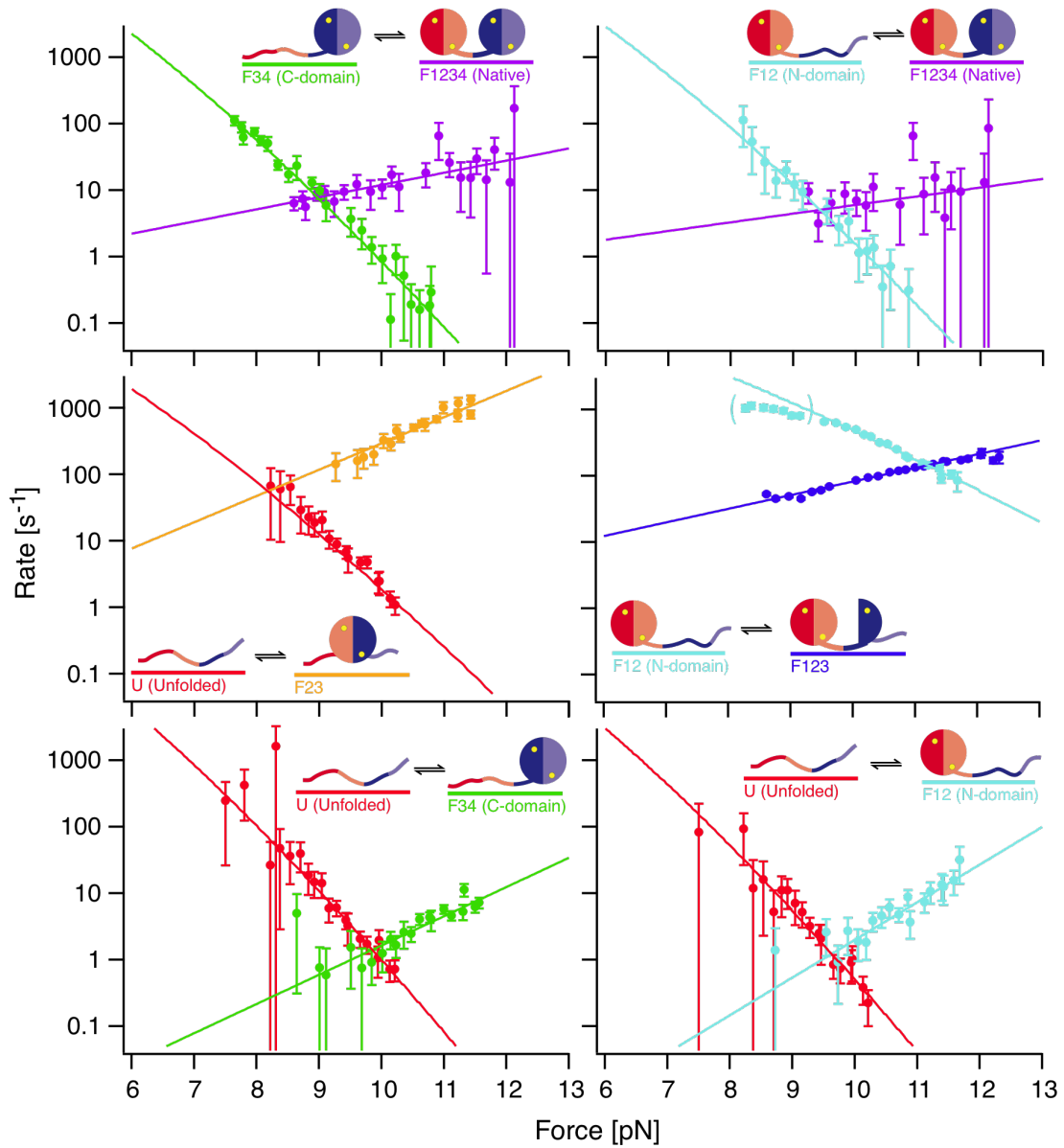


Fig. S8. Dependence of rates for each of the folding and unfolding transitions in wildtype calmodulin. The colors correspond to the originating state in the transition (color code given in Fig. 1C). Error bars are 1σ estimators. The lines represent fits to unfolding and folding models as described in Materials and Methods. The bracketed data points in the middle right graph indicate that missed events due to limited temporal and spatial resolution of the assay biases the data points towards lower rates.

Fig. S9

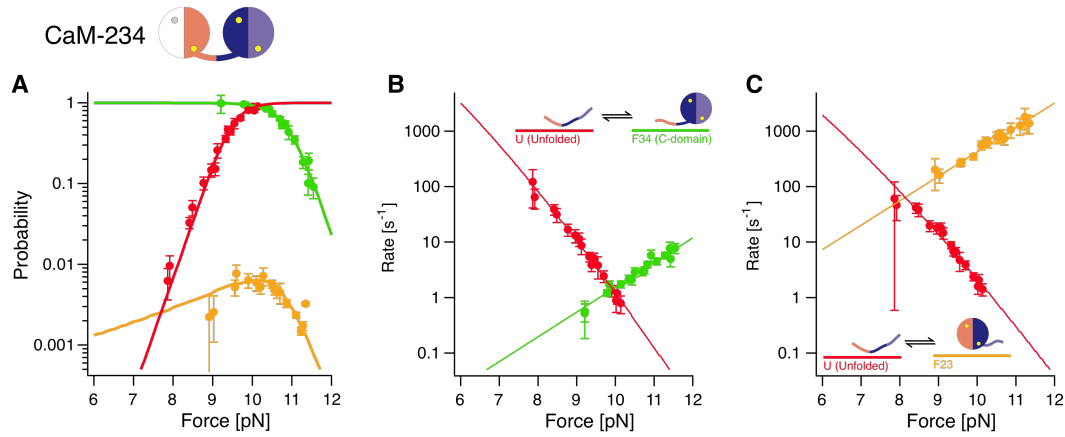


Fig. S9. Kinetic data for the truncation mutant CaM-234.

(A) Occupation probabilities for each of the states F_{34} (green), F_{23} (orange) and U (red) and their dependence on force. Note that in this representation the probabilities at a given force do not add up exactly to one because experiments are conducted at non-constant force (see equation (7) in SOM). Continuous lines represent a global fit to all datapoints.

(B-C) Rates of each transition found in this mutant and their dependence on force. Colors correspond to the originating state in the transition. Continuous lines are fits according to models for folding and unfolding.

Fig. S10

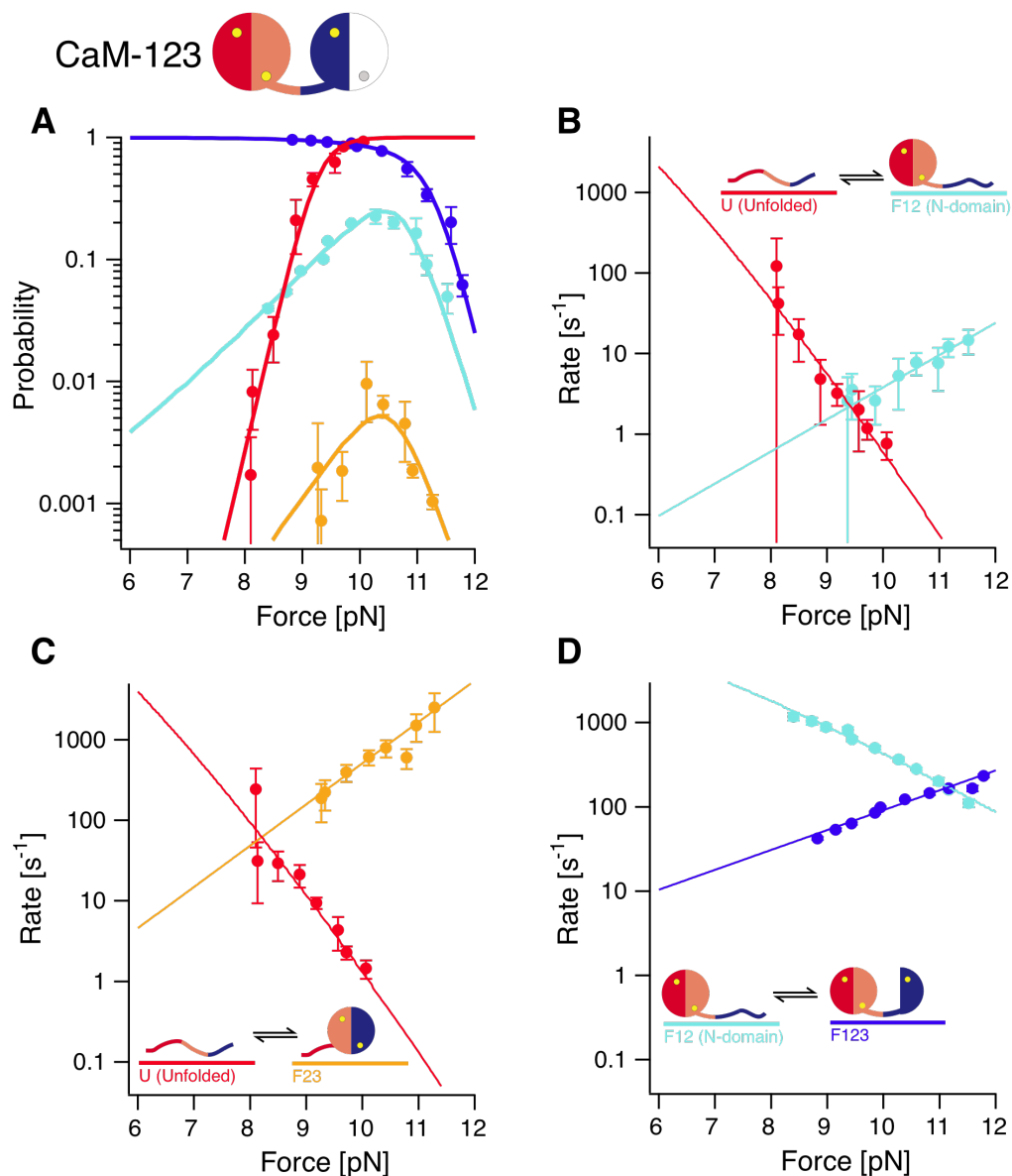


Fig. S10. Kinetic data for the truncation mutant CaM-123.

(A) Occupation probabilities for each of the states F_{123} (dark blue), F_{12} (bright blue), F_{23} (orange) and U (red) and their dependence on force. Note that in this representation the probabilities at a given force do not add up exactly to one because experiments are conducted at non-constant force (see equation (7) in SOM). Continuous lines represent a global fit to all datapoints.

(B-D) Rates of each transition found in this mutant and their dependence on force. Colors correspond to the originating state in each transition. Continuous lines are fits according to models for folding and unfolding.

Fig. S11

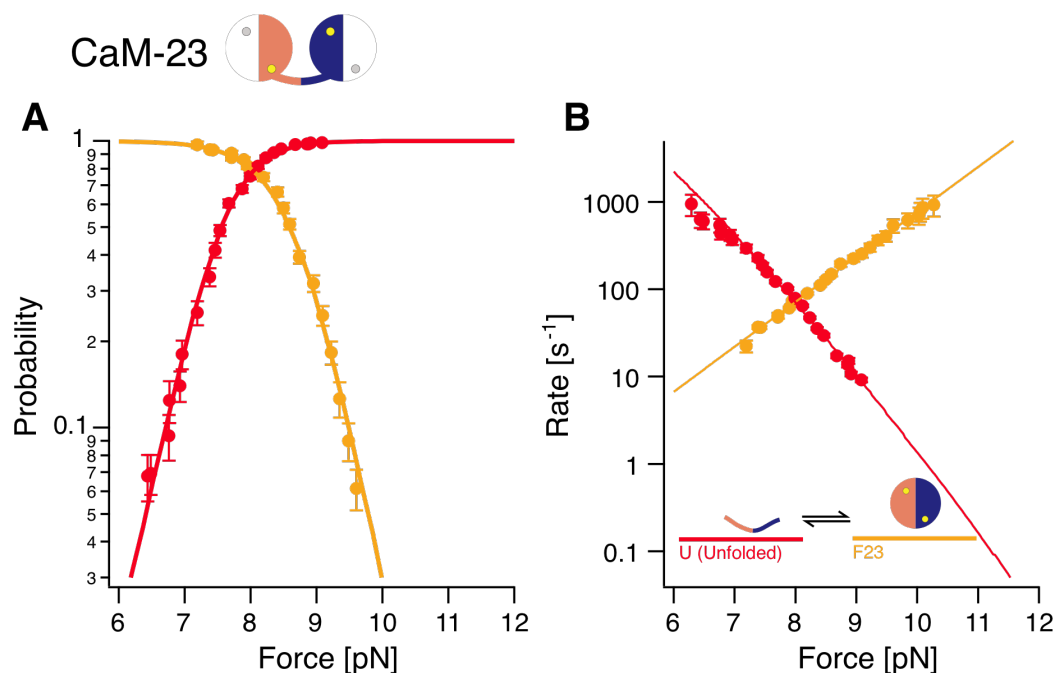


Fig. S11. Kinetic data for the truncation mutant CaM-23.

- (A) Occupation probabilities for the states F_{23} (orange) and U (red) and their dependence on force. Note that in this representation the probabilities at a given force do not add up exactly to one because experiments are conducted at non-constant force (see equation (7) in SOM). Continuous lines represent a global fit to all datapoints.
- (B) Rates of the transition between F_{23} and U. Shown are data for unfolding (orange) and folding (red) and fits to models for folding and unfolding (lines).

Fig. S12

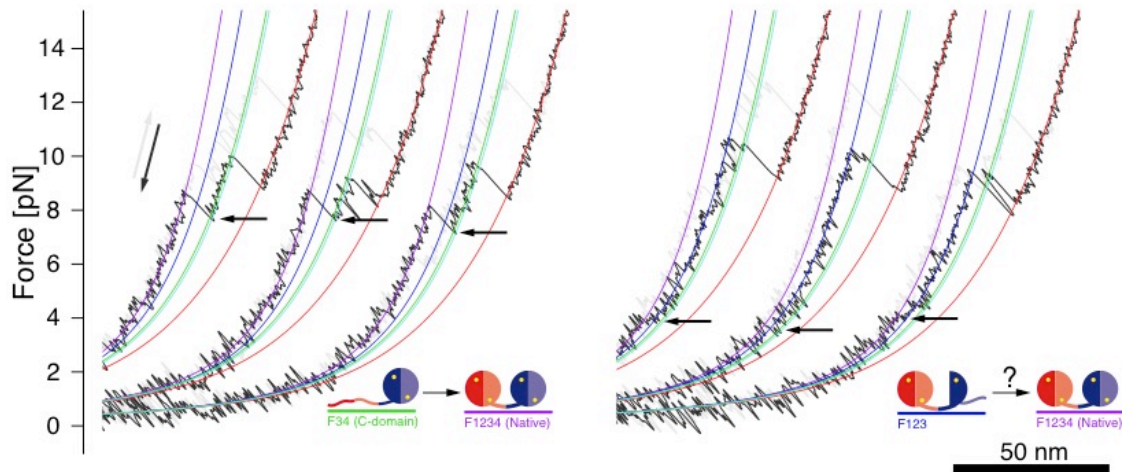


Fig. S12. Relax-cycles at 500 nm/s sorted according to the folding pathway. Left: Refolding via state F_{34} occurs fast and at high loads (arrows) Right: When refolding out of state F_{123} , the molecule gets rapidly trapped in the off-pathway intermediate state F_{123} and further refolding occurs slowly. Even though the protein has to contract much less to reach the native state compared to refolding from F_{34} , refolding occurs at much lower forces (arrows).

Fig. S13

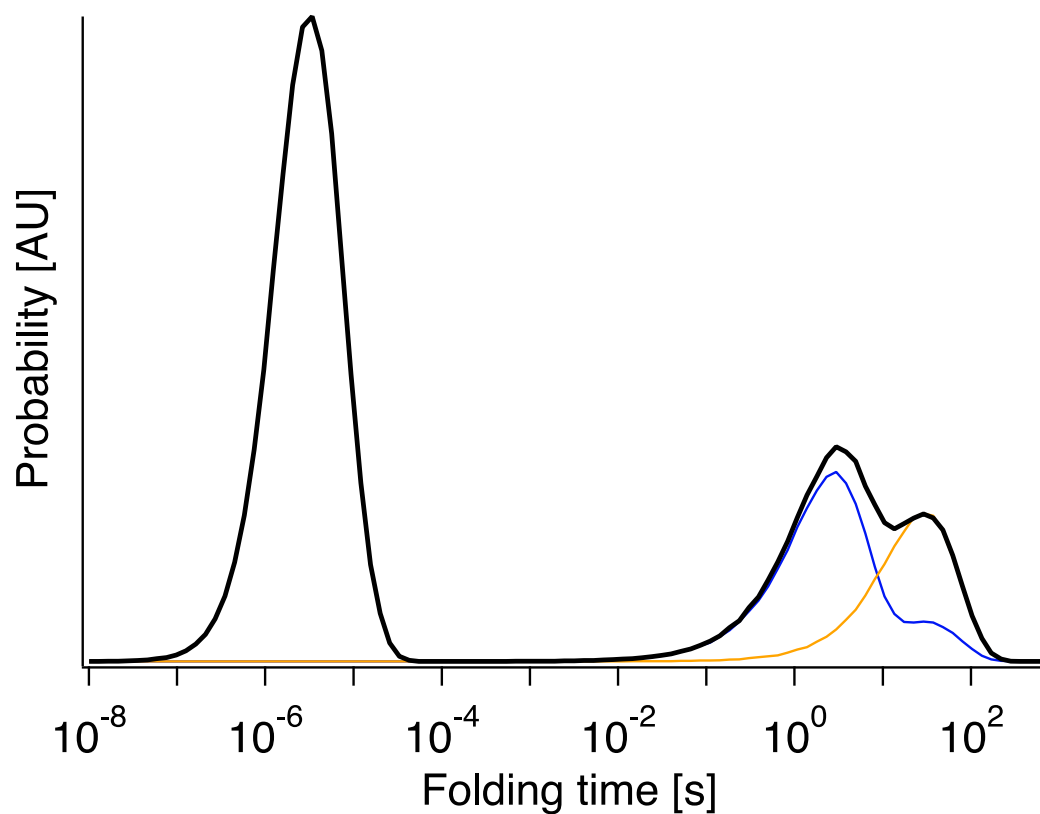


Fig. S13. Calculated distribution of folding times for WT-CaM from state U to F_{1234} under zero load (black). The left peak incorporates the fast direct folding pathways via F_{12} and F_{34} . Molecules getting trapped in state F_{123} show up mainly in the middle peak, those getting trapped in F_{23} mainly in the right peak. The blue trace only shows the folding time of those simulated trajectories that visited state F_{123} at least once. The orange trace shows those that visited state F_{23} at least once.

Fig. S14

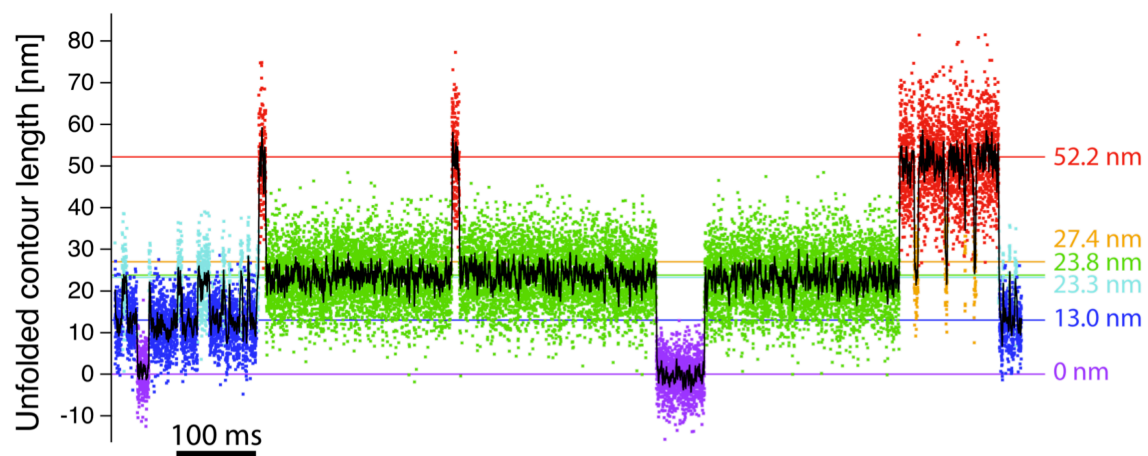


Fig. S14. Trajectory of molecular extension of WT-CaM. Shown is the trajectory of Fig. 1C given in terms of contour length of unfolded polypeptide. Horizontal lines are the average state positions given in contour length (see Table 2).

Fig. S15

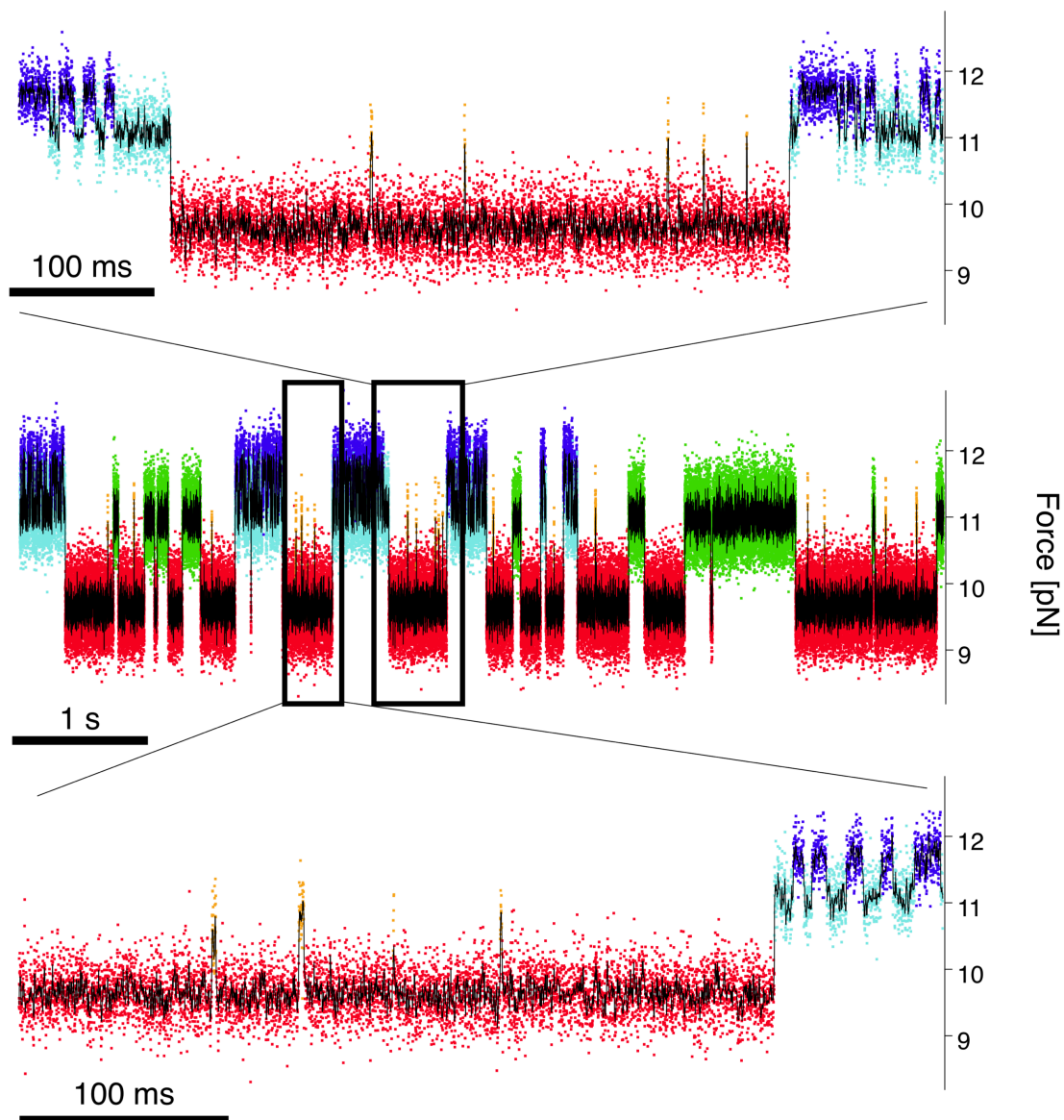


Fig. S15. Sample trace of WT-CaM at about 11 pN pretension with temporal zooms. The states F_{23} (orange) and F_{12} (bright blue) are clearly distinguishable at this pretension not only due to their different length and connectivity of the network (as described in the text), but also due to their different lifetimes. At this pretension the folded (purple) state F_{1234} is not populated. The color-coding follows Fig. 1C.

Fig. S16

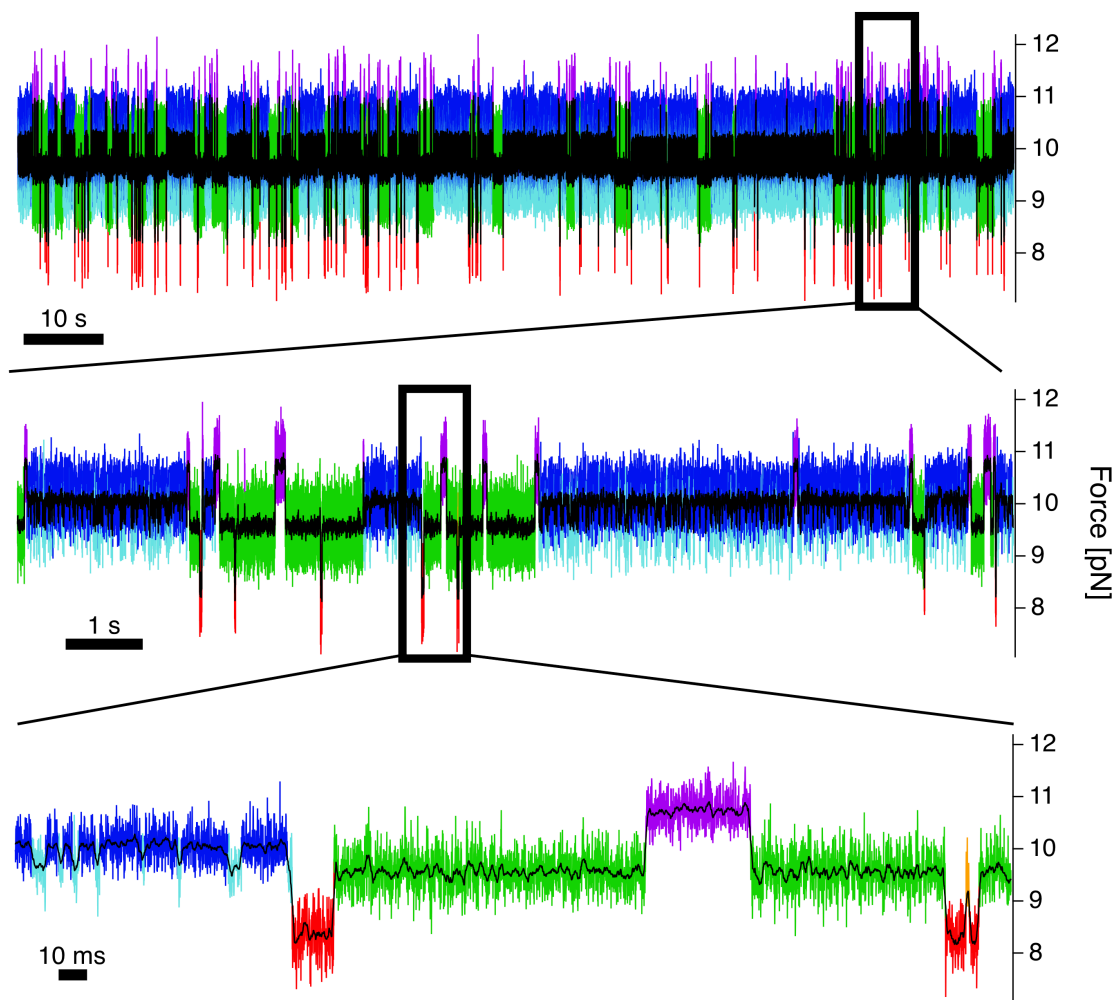


Fig. S16. Sample trace of WT-CaM at a pretension of 9.6 pN where all levels are populated. Shown are the raw data colored in the way the Hidden Markov model classifier assigned them to states in three different zoom levels. The color-coding corresponds to Fig. 1C.

Fig. S17

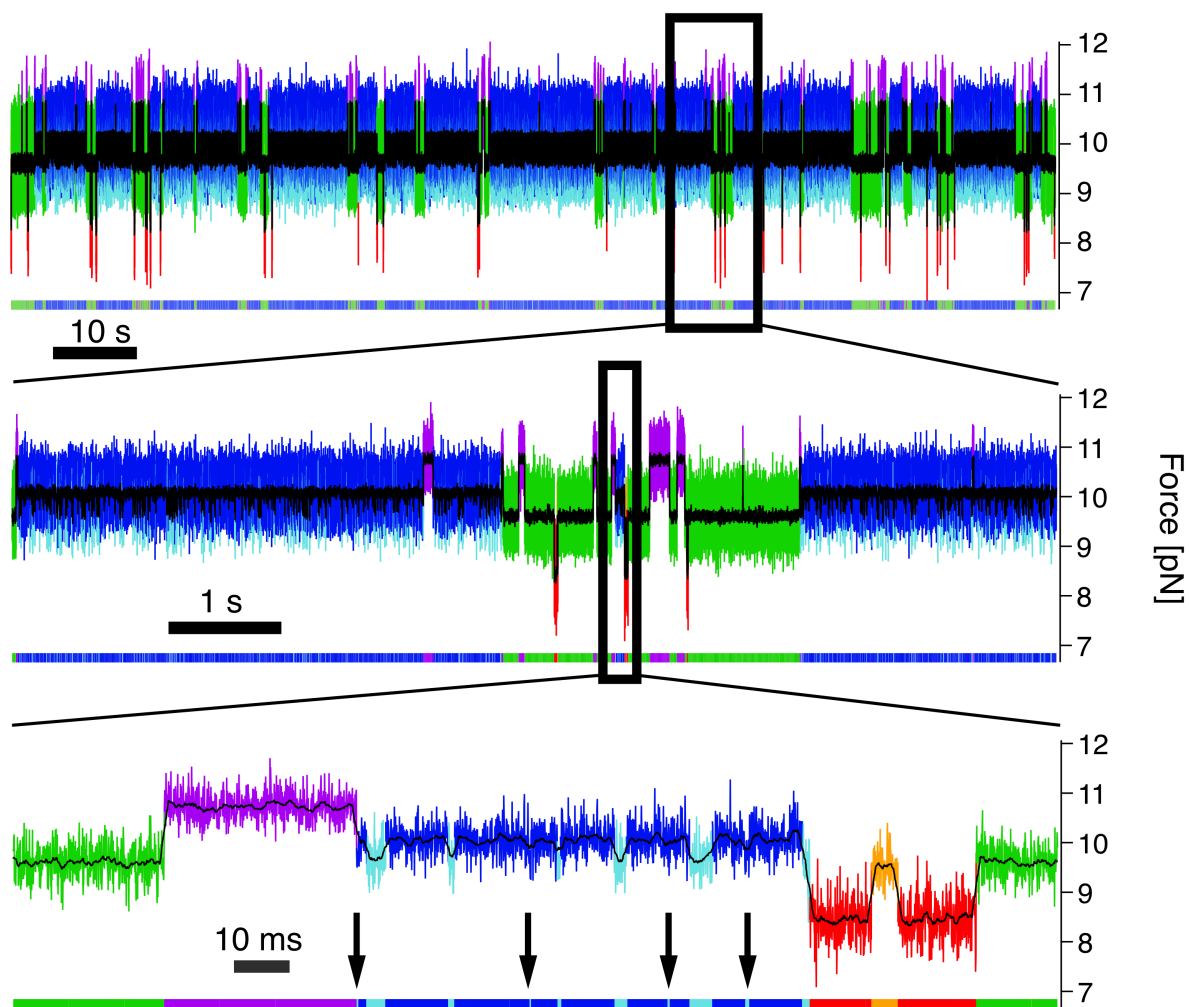


Fig. S17. Simulation of the experiment at pretension values similar to the data trace in Fig. S16. We performed Brownian dynamics simulations of the experiment with the kinetic parameters given in Table 1 (for details see Materials & Methods). The true kinetic trajectory of the protein is shown as a bar graph below the traces. Colors correspond to the coding of Fig. 1C.

The traces were then analyzed with our Hidden Markov algorithm and the colors of the traces assigned accordingly. Arrows show events that were misidentified or missed by the algorithm.

Fig. S18

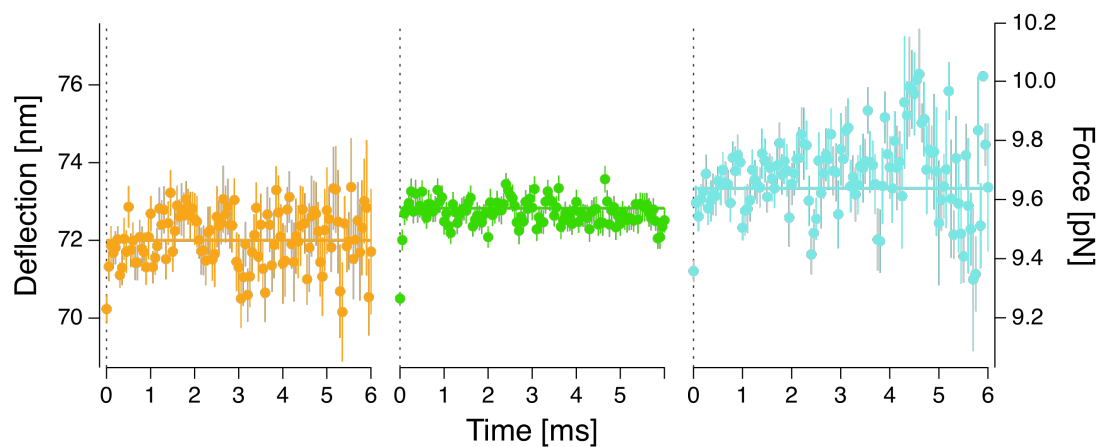


Fig. S18. Ensemble averaged transition events from state U to F_{23} (left, N=35), U to F_{34} (middle, N=69) and U to F_{12} (right, N=36). The transitions were synchronized to the beginning of the transition, binned and averaged (46). Error bars represent standard errors of the mean. The average lifetime for F_{23} at the depicted force is 2.5 ms. The thick horizontal lines represent the determined level position at this pretension.

Table S1

State	CaM 128/144 ΔL^{\dagger} [nm]	CaM WT* ΔL^{\dagger} [nm]
U	44.6 ± 0.5	52.2 ± 0.6
F ₂₃	—**	27.4 ± 0.7
F ₃₄ (C-domain)	23.3 ± 0.7	23.8 ± 0.5
F ₁₂ (N-domain)	16.5 ± 0.7	23.3 ± 0.4
F ₁₂₃	5.5 ± 0.5	13.0 ± 0.3
F ₁₂₃₄ (Native)	0	0

* Values from Table 1

[†] Contour length increase, difference from state F₁₂₃₄. Errors are given as 1 σ intervals.

** State F₂₃ was only rarely observed in CaM 128/144 and left out of the analysis.

Table S1. Measured contour length increases for CaM 128/144. The expected shortening of the length difference between F₁₂₃₄ and the levels U, F₂₃, F₁₂ and F₁₂₃₄ compared to CaM-WT is 5.8 nm.

References and Notes

1. J. N. Onuchic, P. G. Wolynes, Theory of protein folding. *Curr. Opin. Struct. Biol.* **14**, 70 (2004). [doi:10.1016/j.sbi.2004.01.009](https://doi.org/10.1016/j.sbi.2004.01.009) [Medline](#)
2. D. Thirumalai, C. Hyeon, RNA and protein folding: common themes and variations. *Biochemistry* **44**, 4957 (2005). [doi:10.1021/bi047314+](https://doi.org/10.1021/bi047314+) [Medline](#)
3. D. E. Shaw *et al.*, Atomic-level characterization of the structural dynamics of proteins. *Science* **330**, 341 (2010). [doi:10.1126/science.1187409](https://doi.org/10.1126/science.1187409) [Medline](#)
4. P. L. Freddolino, F. Liu, M. Gruebele, K. Schulten, Ten-microsecond molecular dynamics simulation of a fast-folding WW domain. *Biophys. J.* **94**, L75 (2008). [doi:10.1529/biophysj.108.131565](https://doi.org/10.1529/biophysj.108.131565) [Medline](#)
5. V. A. Voelz, G. R. Bowman, K. Beauchamp, V. S. Pande, Molecular simulation of ab initio protein folding for a millisecond folder NTL9(1-39). *J. Am. Chem. Soc.* **132**, 1526 (2010). [doi:10.1021/ja9090353](https://doi.org/10.1021/ja9090353) [Medline](#)
6. B. Schuler, W. A. Eaton, Protein folding studied by single-molecule FRET. *Curr. Opin. Struct. Biol.* **18**, 16 (2008). [doi:10.1016/j.sbi.2007.12.003](https://doi.org/10.1016/j.sbi.2007.12.003) [Medline](#)
7. E. M. Puchner *et al.*, Mechanoenzymatics of titin kinase. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 13385 (2008). [doi:10.1073/pnas.0805034105](https://doi.org/10.1073/pnas.0805034105) [Medline](#)
8. H. Li, A. F. Oberhauser, S. B. Fowler, J. Clarke, J. M. Fernandez, Atomic force microscopy reveals the mechanical design of a modular protein. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6527 (2000). [doi:10.1073/pnas.120048697](https://doi.org/10.1073/pnas.120048697) [Medline](#)
9. M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, H. E. Gaub, Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **276**, 1109 (1997). [doi:10.1126/science.276.5315.1109](https://doi.org/10.1126/science.276.5315.1109) [Medline](#)
10. P. Zheng, Y. Cao, T. Bu, S. K. Straus, H. Li, Single molecule force spectroscopy reveals that electrostatic interactions affect the mechanical stability of proteins. *Biophys. J.* **100**, 1534 (2011). [doi:10.1016/j.bpj.2011.01.062](https://doi.org/10.1016/j.bpj.2011.01.062) [Medline](#)
11. J. P. Junker, F. Ziegler, M. Rief, Ligand-dependent equilibrium fluctuations of single calmodulin molecules. *Science* **323**, 633 (2009). [doi:10.1126/science.1166191](https://doi.org/10.1126/science.1166191) [Medline](#)
12. E. A. Shank, C. Cecconi, J. W. Dill, S. Marqusee, C. Bustamante, The folding cooperativity of a protein is controlled by its chain topology. *Nature* **465**, 637 (2010). [doi:10.1038/nature09021](https://doi.org/10.1038/nature09021) [Medline](#)
13. C. Cecconi, E. A. Shank, C. Bustamante, S. Marqusee, Direct observation of the three-state folding of a single protein molecule. *Science* **309**, 2057 (2005). [doi:10.1126/science.1116702](https://doi.org/10.1126/science.1116702) [Medline](#)
14. J. C. M. Gebhardt, T. Bornschlöggl, M. Rief, Full distance-resolved folding energy landscape of one single protein molecule. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 2013 (2010). [doi:10.1073/pnas.0909854107](https://doi.org/10.1073/pnas.0909854107) [Medline](#)

15. J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, Jr., C. Bustamante, Reversible unfolding of single RNA molecules by mechanical force. *Science* **292**, 733 (2001).
[doi:10.1126/science.1058498](https://doi.org/10.1126/science.1058498) [Medline](#)
16. M. T. Woodside *et al.*, Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science* **314**, 1001 (2006).
[doi:10.1126/science.1133601](https://doi.org/10.1126/science.1133601) [Medline](#)
17. C.-R. Rabl, S. R. Martin, E. Neumann, P. M. Bayley, Temperature jump kinetic study of the stability of apo-calmodulin. *Biophys. Chem.* **101-102**, 553 (2002).
[doi:10.1016/S0301-4622\(02\)00150-3](https://doi.org/10.1016/S0301-4622(02)00150-3) [Medline](#)
18. J. P. Junker, M. Rief, Evidence for a broad transition-state ensemble in calmodulin folding from single-molecule force spectroscopy. *Angew. Chem. Int. Ed. Engl.* **49**, 3306 (2010). [Medline](#)
19. C. F. Shuman, R. Jiji, K. S. Åkerfeldt, S. Linse, Reconstitution of calmodulin from domains and subdomains: Influence of target peptide. *J. Mol. Biol.* **358**, 870 (2006). [doi:10.1016/j.jmb.2006.02.017](https://doi.org/10.1016/j.jmb.2006.02.017) [Medline](#)
20. M. B. Borgia *et al.*, Single-molecule fluorescence reveals sequence-specific misfolding in multidomain proteins. *Nature* **474**, 662 (2011).
[doi:10.1038/nature10099](https://doi.org/10.1038/nature10099) [Medline](#)
21. A. F. Oberhauser, P. E. Marszalek, M. Carrion-Vazquez, J. M. Fernandez, Single protein misfolding events captured by atomic force microscopy. *Nat. Struct. Biol.* **6**, 1025 (1999). [doi:10.1038/14907](https://doi.org/10.1038/14907) [Medline](#)
22. L. Masino, S. R. Martin, P. M. Bayley, Ligand binding and thermodynamic stability of a multidomain protein, calmodulin. *Protein Sci.* **9**, 1519 (2000).
[doi:10.1110/ps.9.8.1519](https://doi.org/10.1110/ps.9.8.1519) [Medline](#)
23. C. Cecconi, E. A. Shank, F. W. Dahlquist, S. Marqusee, C. Bustamante, Protein-DNA chimeras for single molecule mechanical folding studies with the optical tweezers. *Eur. Biophys. J.* **37**, 729 (2008). [doi:10.1007/s00249-007-0247-y](https://doi.org/10.1007/s00249-007-0247-y) [Medline](#)
24. M. Schlierf, H. Li, J. M. Fernandez, The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7299 (2004). [doi:10.1073/pnas.0400033101](https://doi.org/10.1073/pnas.0400033101) [Medline](#)
- <foot>25. Materials and methods are available as supporting material on *Science* Online.</foot>
26. M. P. Landry, P. M. McCall, Z. Qi, Y. R. Chemla, Characterization of photoactivated singlet oxygen damage in single-molecule optical trap experiments. *Biophys. J.* **97**, 2128 (2009). [doi:10.1016/j.bpj.2009.07.048](https://doi.org/10.1016/j.bpj.2009.07.048) [Medline](#)
27. M. de Messieres, B. Brawn-Cinani, A. La Porta, Measuring the folding landscape of a harmonically constrained biopolymer. *Biophys. J.* **100**, 2736 (2011).
[doi:10.1016/j.bpj.2011.03.067](https://doi.org/10.1016/j.bpj.2011.03.067) [Medline](#)
28. H. Dietz, M. Rief, Protein structure by mechanical triangulation. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 1244 (2006). [doi:10.1073/pnas.0509217103](https://doi.org/10.1073/pnas.0509217103) [Medline](#)

<foot>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.</foot>

30. J. Krebs, C. W. Heizmann, Calcium-binding proteins and the EF-hand principle. *New Compr. Biochem.* **41**, 51 (2007). [doi:10.1016/S0167-7306\(06\)41003-6](https://doi.org/10.1016/S0167-7306(06)41003-6)
31. S. Fefeu *et al.*, Calcium-induced refolding of the calmodulin V136G mutant studied by NMR spectroscopy: evidence for interaction between the two globular domains. *Biochemistry* **39**, 15920 (2000). [doi:10.1021/bi001772a](https://doi.org/10.1021/bi001772a) [Medline](#)
32. K. Nakashima, H. Ishida, S. Y. Ohki, K. Hikichi, M. Yazawa, Calcium binding induces interaction between the N- and C-terminal domains of yeast calmodulin and modulates its overall conformation. *Biochemistry* **38**, 98 (1999). [doi:10.1021/bi982067t](https://doi.org/10.1021/bi982067t) [Medline](#)
33. T. M. Lakowski, G. M. Lee, M. Okon, R. E. Reid, L. P. McIntosh, Calcium-induced folding of a fragment of calmodulin composed of EF-hands 2 and 3. *Protein Sci.* **16**, 1119 (2007). [doi:10.1110/ps.072777107](https://doi.org/10.1110/ps.072777107) [Medline](#)
34. Y.-G. Chen, G. Hummer, Slow conformational dynamics and unfolding of the calmodulin C-terminal domain. *J. Am. Chem. Soc.* **129**, 2414 (2007). [doi:10.1021/ja067791a](https://doi.org/10.1021/ja067791a) [Medline](#)

<foot>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.</foot>

36. Z. Guo, D. Thirumalai, Kinetics of protein-folding: Nucleation mechanism, time scales and pathways. *Biopolymers* **36**, 83 (1995). [doi:10.1002/bip.360360108](https://doi.org/10.1002/bip.360360108)
37. S. Tolić-Nørrelykke *et al.*, Calibration of optical tweezers with positional detection in the back focal plane. *Rev. Sci. Instrum.* **77**, 103101 (2006). [doi:10.1063/1.2356852](https://doi.org/10.1063/1.2356852)
38. K. Berg-Sørensen, H. Flyvbjerg, Power spectrum analysis for optical tweezers. *Rev. Sci. Instrum.* **75**, 594 (2004). [doi:10.1063/1.1645654](https://doi.org/10.1063/1.1645654)
39. J. R. Moffitt, Y. R. Chemla, D. Izhaky, C. Bustamante, Differential detection of dual traps improves the spatial resolution of optical tweezers. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 9006 (2006). [doi:10.1073/pnas.0603342103](https://doi.org/10.1073/pnas.0603342103) [Medline](#)
40. M. D. Wang, H. Yin, R. Landick, J. Gelles, S. M. Block, Stretching DNA with optical tweezers. *Biophys. J.* **72**, 1335 (1997). [doi:10.1016/S0006-3495\(97\)78780-0](https://doi.org/10.1016/S0006-3495(97)78780-0) [Medline](#)
41. C. Bustamante, J. F. Marko, E. D. Siggia, S. Smith, Entropic elasticity of lambda-phage DNA. *Science* **265**, 1599 (1994). [doi:10.1126/science.8079175](https://doi.org/10.1126/science.8079175) [Medline](#)
42. L. R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* **77**, 257 (1989). [doi:10.1109/5.18626](https://doi.org/10.1109/5.18626)

43. M. Manosas *et al.*, Force unfolding kinetics of RNA using optical tweezers. II. Modeling experiments. *Biophys. J.* **92**, 3010 (2007).
[doi:10.1529/biophysj.106.094243](https://doi.org/10.1529/biophysj.106.094243) [Medline](#)
44. F. Heinrich *et al.*, A new lipid anchor for sparsely tethered bilayer lipid membranes. *Langmuir* **25**, 4219 (2009). [doi:10.1021/la8033275](https://doi.org/10.1021/la8033275) [Medline](#)
45. M. Schlierf, F. Berkemeier, M. Rief, Direct observation of active protein folding using lock-in force spectroscopy. *Biophys. J.* **93**, 3989 (2007).
[doi:10.1529/biophysj.107.114397](https://doi.org/10.1529/biophysj.107.114397) [Medline](#)
46. C. Veigel, F. Wang, M. L. Bartoo, J. R. Sellers, J. E. Molloy, The gated gait of the processive molecular motor, myosin V. *Nat. Cell Biol.* **4**, 59 (2002).
[doi:10.1038/ncb732](https://doi.org/10.1038/ncb732) [Medline](#)