Full distance-resolved folding energy landscape of one single protein molecule

J. Christof M. Gebhardt[a,b], Thomas Bornschlögl[a], and Matthias Rief[a,1]

[a]Physik Department E22, Technische Universität München, James Franck Strasse, 85748 Garching, Germany; and [b]Munich Center for Integrated Protein Science, 81377 München, Germany

Edited by William A. Eaton, National Institutes of Health NIDDK, Bethesda, MD, and approved December 17, 2009 (received for review August 28, 2009)

Kinetic bulk and single molecule folding experiments characterize barrier properties but the shape of folding landscapes between barrier top and native state is difficult to access. Here, we directly extract the full free energy landscape of a single molecule of the GCN4 leucine zipper using dual beam optical tweezers. To this end, we use deconvolution force spectroscopy to follow an individual molecule's trajectory with high temporal and spatial resolution. We find a heterogeneous energy landscape of the GCN4 leucine zipper domain. The energy profile is divided into two stable C-terminal heptad repeats and two less stable repeats at the N-terminus. Energies and transition barrier positions were confirmed by single molecule kinetic analysis. We anticipate that deconvolution sampling is a powerful tool for the model-free investigation of protein energy landscapes.

leucine zipper | force spectroscopy | optical tweezers | protein folding | deconvolution

The path of an unfolded protein toward its folded and functional conformation is entirely determined by its energy landscape (1). Experimental data often provide very limited view of these energy landscapes. Many proteins are classified as two-state folders, because barrier crossing is the rate limiting step and the subsequent motion toward the native state occurs extremely fast. Kinetic data hence lose almost all the information of the energy landscape on the native side of the transition state. A more detailed insight into the energy landscape of proteins consequently requires experimental data that go beyond classical kinetic assays (2). In recent years, single molecule mechanical methods have been successfully employed to study the energy landscape of biomolecules in increasing detail (3–7). Specifically for DNA, the analysis of equilibrium fluctuations upon application of mechanical load has provided sequence-resolved energy profiles of the full energy landscape (4). For proteins, such a detailed description has so far remained elusive.

The leucine zipper of the yeast transcriptional activator GCN4 is an ideal protein model system for studying real time folding/unfolding dynamics to obtain spatially resolved energy profiles. Because of its simple linear folding topology, the mechanically unzipped length can be directly linked to the amino acid position of the unzipping fork. The GCN4 zipper domain contains four heptad repeats forming a double-stranded α-helical coiled coil (8) and has been described as a two-state folder (9). Bulk folding studies have shown that folding of a cross-linked coiled coil is nucleated at the C-terminal end of the protein (10, 11). From there, zippering of the coiled coil proceeds toward the N-terminus. Activation energies and folding kinetics have been investigated extensively (9, 12, 13). Earlier single molecule mechanical experiments using atomic force microscopy have provided insight into the average unfolding forces of the zipper domain (14, 15). Limited force resolution in atomic force microscopy (AFM) experiments, however, has precluded the direct observation of folding/unfolding transitions in this system.

Here, we use single molecule force spectroscopy by optical tweezers (3) to directly measure the full free energy landscape of a GCN4 based leucine zipper. The experimental design is sketched in Fig. 1*A*. We used a fusion construct consisting of a sequence of three identical GCN4-p1q domains (construct LZ26, see *Methods* and *SI Text*) (11, 14). This triple zipper domain construct offers the possibility to study the GCN4-p1q energy landscape with nucleation (C-terminal domain shown in blue) and simultaneously nucleation free (N-terminal domains shown in green and red) (15). The protein is clamped between two beads using DNA handles attached to N-terminal Cysteins (see *SI Text*) (3). One bead is moveable with respect to the other to control the tension applied to the protein.

Results and Discussion

In a first set of experiments we recorded force vs. extension traces at constant trap velocity. In Fig. 1*B*, four successive unzipping (*Black*) and rezipping (*Blue*) cycles pulled at 500 nm/s are shown. Unfolding of the LZ26 zipper results in a highly reproducible characteristic folding/unfolding pattern at forces between 8 and 15 pN (Fig. 1*B*, *Inset*). Upon force application, two intermediates (I_1 and I_2) can be observed. Starting from the fully folded state N, I_1 is populated in a smooth hump-like transition at equilibrium. Transition to I_2 occurs close to equilibrium; however, distinct flips of the molecule between the two intermediate states are resolved. If stretched further, a final transition occurs to the completely unfolded configuration U of the molecule. Upon reversal of the pulling process the molecule refolds, exhibiting a hysteresis at a pulling velocity of 500 nm/s. To relate the observed intermediate configurations to the sequence of the protein, we used a serial worm like chain model to fit the force vs. extension traces (see Table 1, *SI Text*, and Fig. S1). The positions of the intermediate states correspond well to the positions of asparagine residues in the sequence, which are known to destabilize the coiled coil (14, 16).

The low instrumental drift of our setup allowed observation of thousands of transitions of one molecule between different protein conformations held at defined pretensions (the force acting on intermediate I_1) at constant trap separation. A typical force vs. time trace is shown in Fig. 1*C* (*Upper*). Because trap separation but not force is kept constant, every length change of the protein is associated with a change in tension. A zoom into the long data trace (*Lower*) allows observation of equilibrium transitions between the unfolding intermediates I_1 and I_2 as well as the completely unfolded protein U (shown as red, green, and blue dashed lines). The lines are determined as maxima of Gaussian fits to the data. The red and green lines appear slightly closer than expected from the contour length increases. This discrepancy indicates a deviation from a quadratic shape of the underlying folding energy landscape.

Author contributions: J.C.M.G. and M.R. designed research; J.C.M.G. performed research; J.C.M.G. and T.B. contributed new reagents/analytic tools; J.C.M.G. analyzed data; and J.C.M.G. and M.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

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

This article contains supporting information online at www.pnas.org/cgi/content/full/0909854107/DCSupplemental.

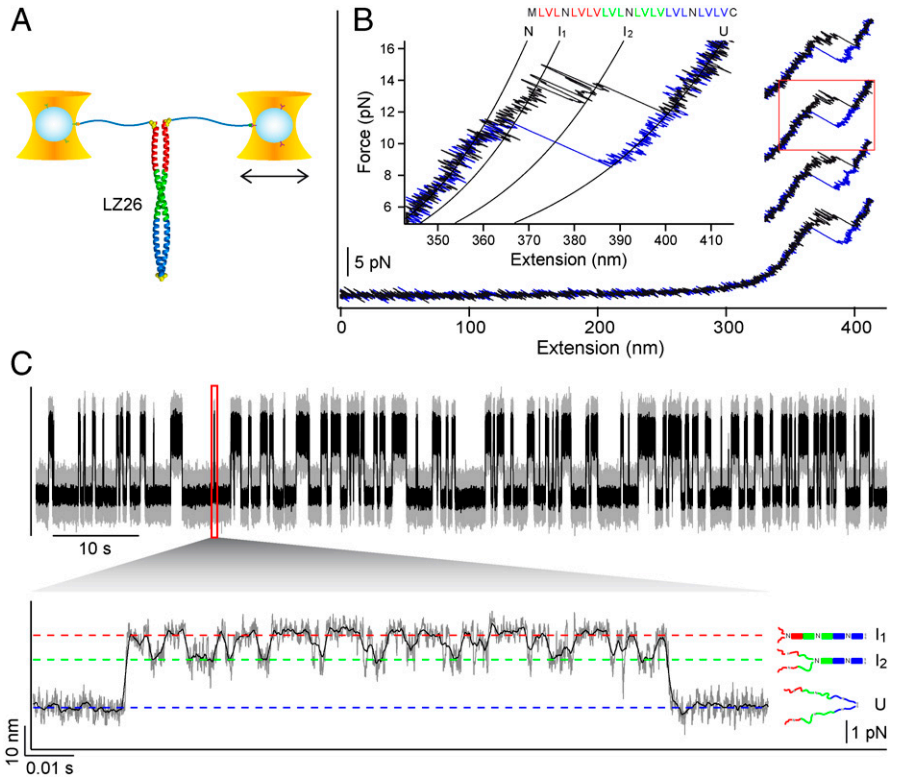

Fig. 1. Experimental setup and characteristic sample traces. (*A*) Cartoon depicting the experimental setup. The LZ26 coiled coil containing three GCN4 leucine zipper domains is attached to two beads via dsDNA handles. Individual zipper domains are coloured red, green, and blue. The two α-helical strands are cross-linked by cysteines at the C-terminus to avoid dissociation upon complete unfolding (see *SI Text*). (*B*) Force vs. extension traces of the LZ26 coiled coil. Four sequential unzipping (*Black*) and rezipping (*Blue*) cycles at 500 nm/s are shown (offset in force for clarity) (*Inset*) Magnification of the cycle marked by the red square. Unfolding from the native state (N) to the unfolded state (U) occurs via two resolved intermediates (I_1 and I_2). Lines are fits of a serial worm-like-chain model to the data (see Table 1 and *SI Text*). The letters specify amino acid residues in *A* and *D* positions of the coiled coil. (*C*) Force vs. time record of the LZ26 coiled coil held at a pretension of 14.1 pN (*Upper*) at constant trap separation. The magnification (*Lower*) of the region marked by the red square allows observation of transitions between I_1 (*Red Dashed Line*), I_2 (*Green Dashed Line*), and U (*Blue Dashed Line*). The structure of these conformations is sketched on the right. The *y* axis represents bead deflection from the trap center, which is linearly connected to the force acting on the molecule.

Increasing the separation between the two traps, and hence the pretension on the molecule, shifts the population probability from the folded to the unfolded state (Fig. 2*A*). These force vs. time traces contain a wealth of energetic and kinetic information about the folding process of the protein. Using the Boltzmann relation, differences in free energy between the stable states I_1, I_2, and U can be calculated from the population probability histograms (Fig. 2*B*, see Table 1 and Eq. S7). Further information about the equilibrium free energies can be obtained by exploiting the Crooks fluctuation theorem (17, 18) (Fig. 2*C*). The intersection (*Red Circle*) of folding (*Blue*) and unfolding (*Black*) work distributions obtained from nonequilibrium force vs. trap separation curves (*Inset*) defines the equilibrium free energy of folding of the complete LZ26 coiled coil of (75 ± 3) k_BT (see Table 1, Eq. S8, and Fig. S2).

The distributions of dwell time intervals τ (Fig. 2*D*, *Inset*) allow extraction of rate constants for transitions between I_1, I_2, and U. For extracting zero force rates it is important that movements of the transition state under load are modeled correctly. Such transition state movements were neglected in the simple Bell-type model (19), however several improved models have recently been proposed to extract transition state positions as well as zero force folding and unfolding rates from force dependent rate measurements (Fig. 2*D*) (20–23). We adapted a model initially proposed to describe folding under load (22) to model both folding and unfolding rate constants (*SI Text*). In brief, the unfolded confor-

Table 1. Energetic and kinetic parameters of the LZ26 coiled coil.

	Dwell time					Contour length	Probability distribution
Transition	k_u^0 (s^{-1})*	k_f^0 (s^{-1})*	ΔG^0 (k_BT)†	$\Delta x_u^\#$ (nm)*	$\Delta x_f^\#$ (nm)*	Δx (nm)‡	ΔG^0 (k_BT)‡
N, I1	–	–	–	–	–	9.3 ± 1.1	9.6 ± 1.1
I1, I2	$(8.7 + 8.3 - 6.9)\ 10^{-4}$	$(6.9 \pm 4.1)\ 10^{7}$	25.1 ± 2.0	8.7 ± 0.7	8.3 ± 0.1	19.9 ± 1.3	23.8 ± 0.4
I2, U	$(2.5 \pm 1.8)\ 10^{-5}$	$(1.7 + 4.3 - 1.2)\ 10^{12}$§	38.8 ± 2.5	9.8 ± 0.4	24.9 ± 1.7§	33.2 ± 1.3	42.1 ± 0.4
	–	$(5.0 \pm 2.8)\ 10^{3}$¶	–	–	7.9 ± 0.6¶		
N, U	–	–	–	–	–	62.4 ± 1.0	75 ± 3‖

*Errors are a combination of statistical and systematic errors due to trap stiffness uncertainty.
†From the ratio of zero force rate constants.
‡Errors are ±SD.
§Fit to rate constants above 10.9 pN.
¶Fit to rate constants below 10.9 pN.
‖From Crooks' fluctuation theorem.

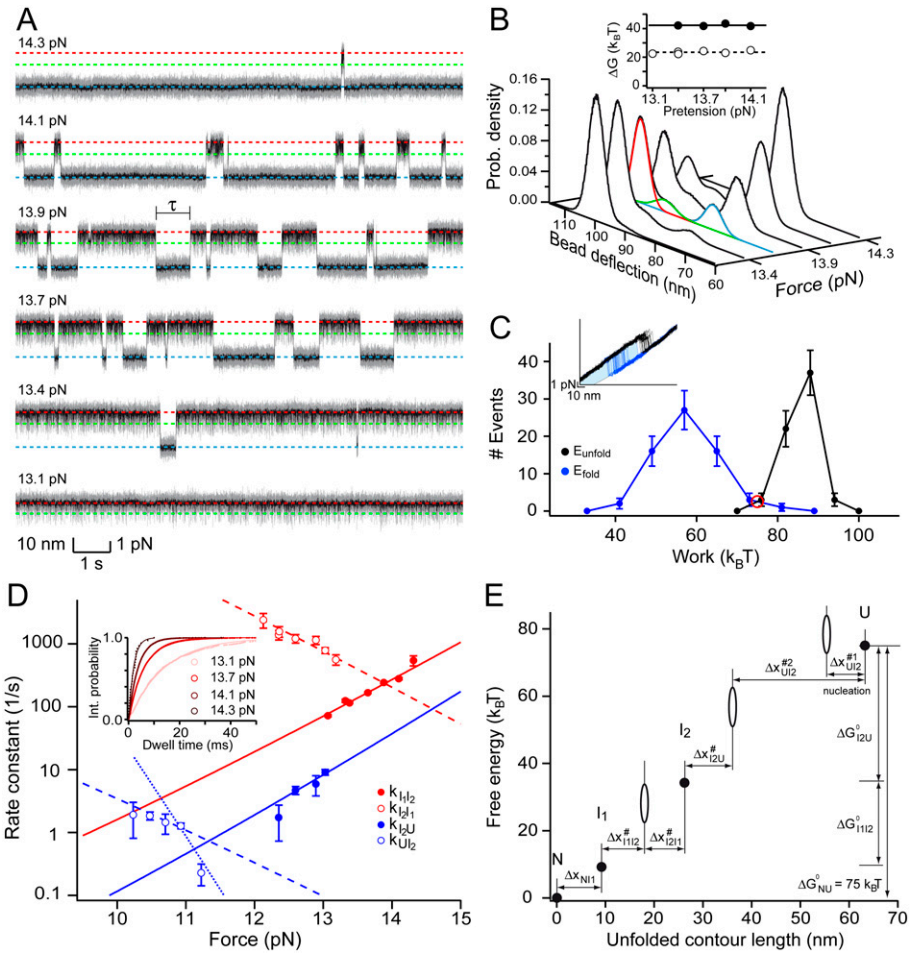

Fig. 2. Energetic and kinetic characterisation of the LZ26 coiled coil. (*A*) Force vs. time traces at different pretensions as indicated on the left. Dashed lines mark the conformations depicted in Fig. 1*C*. Total recorded times are (from bottom to top, in s): 12.1, 25.6, 117.2, 13.8, 114.5, 33.9. (*B*) Bead deflection histograms from the traces shown in *A*. Red, green, and blue lines are individual Gaussian components of a triple Gaussian fit to the histogram at 13.7 pN pretension. (*Inset*) Free energy differences between I_1 and I_2 (*Open Symbols*) and I_2 and U (*Closed Symbols*) at the indicated pretension (see Table 1). (*C*) Histogram of work associated with unfolding (*Black Symbols*, $n = 65$) and refolding (*Blue Symbols*, $n = 65$) of the complete LZ26 coiled coil (data from 10 molecules). The intercept of both work distributions (75 k_BT, *Red Circle*) corresponds to the total free energy of folding according to the Crooks fluctuation theorem (17, 18). Error bars represent SD. (*Inset*) Overlay of 37 folding (*Blue*) and unfolding (*Black*) cycles shown as force vs. trap separation. The light blue area corresponds to the work associated with folding of the molecule of one selected cycle, corrected for contributions of beads and DNA handles. (*D*) Force dependent rate constants for the transitions between I_1, I_2, and U. Lines are fits to the data according to a model accounting for elastic contributions from beads, handles, and unfolded protein chain (Eq. S9). The model yields curved lines that, however, appear straight in the displayed force range. The values are summarized in Table 1. Error bars represent SEM or an error estimated from missed events due to fast kinetics, if larger. (*Inset*) Integrated probability histograms of dwell times in I_1 preceding a transition to I_2 at varying pretensions (*Dots*). The distributions are well fit by a single exponential (*Lines*). (*E*) Check-points of the free energy surface. (*Closed Symbols*) Stable protein states with known position and energy. (*Open Ellipses*) Barrier positions whose energies depend on the preexponential factor (an interval between 10^3 s^{-1} and 10^7 s^{-1} is indicated).

mation of the protein chain can be described by a worm-like chain (24) and corresponding transition barrier positions are very sensitive to force. For protein folding, we therefore use a model that accounts for the change in energy of the trapped beads, DNA linkers, and the unfolded protein chain associated with a length change of the protein upon folding (22) (Table 1 and Eq. S9). Because in our unzipping experiment force always acts at the unfolding fork, it is reasonable to assume that unfolding as well as folding occur turn by turn and each step directly translates into a length change of the unfolded peptide chain. Therefore, also for unfolding the change in energy of springs and linkers matters as the coiled coil gradually opens and the same model is suited to describe the unfolding rate constants (Table 1).

The free energy difference of 25 k_BT between I_1 and I_2 obtained from the ratio of zero force rate constants is in good agreement with the 24 k_BT derived from the position probability distributions. Likewise, the sum of the respective transition state positions for folding and unfolding of 17 nm corresponds to the extension of the protein obtained directly from the force extension measurements of 20 nm (see Table 1). The value of 5×10^3 s^{-1} for folding at zero force lies at the lower end of folding rates measured in ensemble studies ($7.5 \times 10^3 - 2 \times 10^5$ s^{-1}) (9, 10). A discrepancy between values obtained by bulk measurements and those from mechanical studies is to be expected. Chemical and mechanical methods generally perturb folding in distinct ways (25). However, Schlierf et al. (22) have shown that at moderate forces <10 pN extrapolated folding rates are identical to those in the absence of load for an Ig domain from ddFilamin. Moreover, for coiled coils it is not a priori unreasonable that the folding pathway under load is related to the one occurring in solution. Meisner et al. (9) have reported that the solution pathway of the GCN4 zipper domain involves formation of a nucleus at the C-terminus with subsequent zippering toward the N-terminus. Interestingly, this is exactly the pathway that will be favored by the application of load to the N-terminus in our experiment.

Notably, the force-dependence of the rate constant for refolding (as shown by the open blue symbols in Fig. 2*D*) exhibits a bend. In principle, the strong force dependence of the unfolded protein state results in a curved rate versus force plot (26). This effect only accounts for a very slight curvature and is already included in the model used to calculate the force dependence of rates. Hence, the abrupt change in slope above 11 pN is an indication of a drastic change in transition barrier position (see Fig. S3, *Dashed/Dotted Blue Lines*). Such a change in barrier position could be either the consequence to two sequential barriers (27) or a broad transition barrier (28). At forces below 11 pN, we measure a value for $\Delta x_{UI2,1}$ of 2 heptads for the initial nucleation of coiled-coil formation (*Dashed Blue Line*), consistent with previous findings (11, 15). A second fit of our model to the refolding rates above 11 pN (*Dotted Blue Line*) yields $\Delta x_{UI2,2} = 25$ nm for the transition state distance to the second barrier. Together with the 10 nm for the reverse transition (Δx_{I2U}) this adds up to the full distance between states I_2 and U (see Table 1). Because evidence for this drastic change in transition state position hinges upon only a single data point, the extracted barrier distance might be subject to a considerable error. However, it is important to note that the single point at 11.2 pN in the UI_2 refolding rate is not the sole evidence for the existence and the quantification of this double barrier. Even if this point were disregarded, the overall conclusion would be unaffected. The transition state distance of $\Delta x_{UI2,1} = 8$ nm as obtained from the dashed blue line together with the distance of the reverse transition is too short to account for the complete distance between I_2 and U (approximately 33 nm) and would hence indicate the existence of another hidden barrier. From this argument alone, a transition state position for the hidden barrier of *ca.* 23 nm could be deduced. This value coincides almost perfectly with our measured value of $\Delta x_{UI2,2} = 25$ nm.

The energies and positions obtained so far define the important checkpoints along the folding landscape of the coiled coil that can be obtained in a combination of equilibrium and non-equilibrium experiments (Fig. 2*E*, *Spheres*). Deriving barrier heights from kinetic data, however, relies on the knowledge of the reconfiguration time of the protein chain, i.e. the preexponential factor k_ω in the Arrhenius equation. This preexponential factor could previously only be determined indirectly and has been estimated to 10^3–10^7 s^{-1} (29). Therefore, the barrier heights in Fig. 2*E* are drawn with a large uncertainty (*Open Ellipses*).

Can we gain insight beyond the sketchy model of the folding free energy landscape of Fig. 2*E*? For DNA molecules, Woodside et al. (4) have recently provided proof of principle that equilibrium sampling can be used to extract the full energy landscape of biomolecules directly in a model-free way. When traversing between folded and unfolded states, the protein samples all possible conformations including the high energy states (transition state). Hence, the position distributions of Fig. 2*B* contain much more information than the coloured Gaussian fits suggest: specifically, the rare excursions to higher energy states along the folding pathway should give us much more detailed information about the underlying energy landscape. In principle, the full underlying energy distribution can be calculated from the position probability distribution $P(x)$ using the Boltzmann relation $\Delta G(x) = k_B T \ln(P(x)) + c$ (4, 30, 31).

In an optical trapping assay, the true protein fluctuations are masked by thermal fluctuations of the beads as well as the DNA handles (Fig. 3*A*, *Black Line*). The thermal noise contributions of beads and handles can be described by a point spread function (PSF(x)), which the fluctuations of the ends of the protein distribution are convolved with. To recover the true probability distribution of the protein ends (Fig. 3*A*, *Red Line*), we employed and modified a deconvolution procedure based on the work by Woodside et al. (4). In our approach, we use a nonconstant point spread function ($PSF^a(x)$) whose width depends on the bead

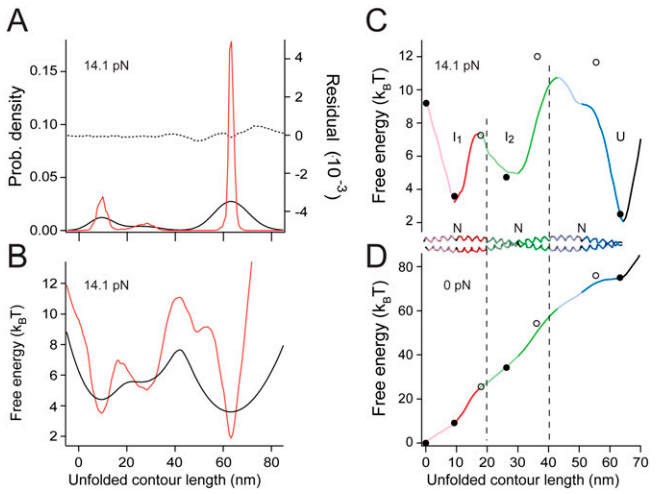

Fig. 3. Model-free reconstruction of the full energy landscape of the LZ26 coiled coil. (*A*) The protein probability distribution (*Red Line*) at 14.1 pN pretension is recovered from the bead deflection probability distribution (*Continuous Black Line*) by deconvolution with small residual error (*Dotted Black Line*) to remove thermal noise contributions of the series compliance of beads and elastic spacers. (*B*) Deconvolved free energy landscape (*Red Line*) of the protein at 14.1 pN pretension and blurred landscape including series compliance effects (*Black Line*). (*C*) Protein landscape (*Colored Line*) averaged from two landscapes at different pretensions, calculated to 14.1 pN pretension and check-points of the schematic energy landscape (closed and open symbols, Fig. 2*E*, *Closed* and *Open* symbols) with a preexponential factor of 1.2×10^4 s^{-1} derived from the measured barrier height. (*D*) Protein energy landscape and check-points of *C* in the absence of force. Colors indicate the three GCN4 leucine zipper domains of the LZ26 coiled coil. The GCN4 leucine zipper consists of two stable C-terminal heptads (*Dark Colors*) and two less stable N-terminal heptads (*Light Colors*).

deflection *a*, to account for the increase in width of the bead position distribution with decreasing force due to the nonlinear stiffness of the DNA handles (see *SI Text* for details). Introduction of the nonconstant $PSF^a(x)$ was essential for convergence of the iterative deconvolution procedure. The resulting protein energy landscape is shown in Fig. 3*B* (*Red Line*). To increase the reliability of the recovered energy landscape, we averaged energy surfaces obtained from the same molecule at two different pretensions, after correcting them for these pretension differences (*SI Text*). After force correction, the difference between both energy surfaces has a standard deviation of ± 0.8 k_BT, which gives confidence into the accuracy of the energy surfaces obtained by our deconvolution method. Including drift, the spatial resolution is 2–3 nm. The averaged energy surface in Fig. 3*C* is now a direct measure of the distance-resolved energy landscape obtained from only one single molecule of the LZ26 leucine zipper held in an optical trapping potential at 14.1 pN pretension. The measured barrier heights $\Delta G^{\#}(F)$ in combination with the transition rates $k(F)$ allow an independent estimate of the preexponential factor k_ω for folding of the leucine zipper in our trap according to $k_\omega = k(F) \times \exp(\Delta G^{\#}(F)/k_BT)$. We chose to calibrate the preexponential factor using the barrier $\Delta G^{\#}_{I1I2}$. It likely offers the best estimate of a barrier energy, because, due to the fast kinetics between I_1 and I_2, this barrier is crossed >10,000 times. We derive a value of $k_\omega = 1.2(+1.6 - 0.6) \times 10^4$ s^{-1}. This preexponential factor lies at the lower end of expected values (29). Tethering of the small protein to large beads may slow down the protein motion by coupling it to the motion under force of the much larger beads connected by the DNA linker (approximately 10^4 s^{-1}) (32). In this scenario the measure of the barrier height would be unaffected. On the other hand, the diffusion time of the beads in the optical trap may preclude detection of faster internal protein motion and thus reduce the apparent preexponential factor. Therefore this value constitutes at least a lower bound for k_ω.

This lower bound is slightly higher than lower boundaries for the preexponential factor ($(2.7–5.5) \times 10^3$ s^{-1}) calculated from measurements of the upper limit of the transition path time (33, 34).

The open and closed symbols in Fig. 3*C* correspond to the schematic energy landscape of Fig. 2*E* using our measured preexponential factor for the barrier height estimates. The directly measured landscape and the key check-points of the schematic landscape are in very good agreement. Notably, all barrier positions are reproduced surprisingly well in the deconvolved energy landscape.

It has been a long-standing debate whether broad barriers observed in protein folding experiments indicate a number of energetically equivalent conformations or rather two or more narrow sequential barriers that cannot be resolved due to resolution issues (35, 36). For the apparently broad major barrier between I_2 and U, located at 40–60 nm in Fig. 3*C*, the deconvolved energy landscape helps to distinguish between those two scenarios. The two maxima of this barrier with a depression at *ca.* 55 nm strongly argue for the sequential barrier model and justify our earlier analysis with a second transition barrier (two slopes for the open blue symbols in Fig. 2*D*). Furthermore, it is intriguing that the high energy minimum lies exactly at the position of the weak asparagine residue in the C-terminal zipper domain, in accordance with the positions of the intermediate states I_1 and I_2. However, it is important to point out that errors of the regions with high energy will be necessarily large due to limited thermal sampling of those regions. This error in part explains the smaller measured barrier heights of the last two barriers compared to the heights calculated from the kinetic parameters. Moreover, because barrier heights seem to be influenced by bead kinetics, the decreasing characteristic frequency of the trapped beads with decreasing force yields systematically decreasing barrier heights with increasing protein extension.

Between states N and I_1 the energy landscape appears linear. Because we do not observe distinct transition events for the N to I_1 transition, a barrier, if it exists, has to be lower than the barrier between I_1 and I_2. The linear shape suggests that I_1 does not represent a real intermediate at higher forces, because the thermodynamic state N represented by a local minimum in the energy landscape vanishes under load. An alternative interpretation would view I_1 as the new native state under load.

Back transformation of the folding free energy landscape to zero force conditions provides the energy landscape shown in Fig. 3*D*, exhibiting a remarkably detailed picture of the LZ26 leucine zipper stability and folding. The free energy of folding of the coiled coil extracted from this zero force energy landscape (approximately 75 k_BT) corresponds well to the energy derived from the fluctuation theorem of Crooks. The transformed landscape clearly shows a repetitive energy pattern reflecting the composition of LZ26 from three individual GCN4-p1q coiled coils.

Hydrogen exchange measurements have shown that folding of a C-terminally cross-linked GCN4 zipper domain is barrier-limited and thus exhibits two-state behavior (9). The C-terminal GCN4 zipper domain (*Blue*) of our LZ26 construct, which has a free energy of 21 k_BT, comparable to bulk measurements (9), is in accordance with this finding, albeit with a very low folding barrier. This domain contains the energy contributions of initial seed formation. In contrast to the C-terminal domain, we find the 2 N-terminal domains (*Green* and *Red*) do not exhibit a barrier. Apparently, folding of these domains proceeds in a downhill fashion. It is important to note that the intermediates I_1 and I_2 only exist and are populated in a force experiment where the elastic linkers under load shape the energy landscape in such a way that pronounced minima appear (Fig. 3*C* and *D*). It is this unique property of force experiments that allows us to look into the downhill part of a folding protein. We can only conclude about the downhill nature of the pathway populated under load, not necessarily about the solution pathway. However, as pointed out above, there are arguments why solution and mechanical folding pathways may be similar for a coiled coil. Comparison of the free energies of the downhill folding domains to the free energy of the domain including the seed now allows an estimate of the entropic costs of seed formation of approximately 6 k_BT. Moreover, the observation of barrierless folding of concatenated leucine zipper coiled coils shows that those domains can be fused seamlessly without energetic prices. This is in stark contrast to the neck coiled coil of kinesin, where a noncanonical N-terminal hydrophobic collar region prevents seamless elongation of the coiled coil (37).

In the energy landscape of Fig. 3*D*, both N-terminal zipper domains appear clearly separated into two differently stable regions, a stable C-terminal region comprising two heptad repeats and two significantly less stable N-terminal heptads (indicated with light and dark colors in Fig. 3*D*). The approximately linear energy increases within these regions indicate that neighboring core residues exhibit similar energy contributions. Even though the GCN4 system is one of the best investigated model systems for protein folding, direct evidence for such an energetic asymmetry has so far been missing. Interestingly, the point of division into the differently stable regions coincides with the location of the asparagine residue in the center of each GCN4-p1q domain. This asparagine residue has been found to be destabilizing for the total free energy of the leucine zipper (16). Whereas this colocalization may be coincidental, a noncanonical residue at the border between two regions of different stability is remarkable. Our results of an energetic asymmetry between the N- and C-terminal heptads can explain experiments that found that C-terminal heptads are more sensitive to mutations of core residues than the N-terminal heptads (38) on an energetic basis. The energetic asymmetry also offers an energetic explanation for the proposed C-terminal trigger sequences governing folding of many physiologically important coiled coil structures (39, 40).

Recently, Li and coworkers proposed a mechanism for protein folding against force in which the polypeptide chain folds via a cooperative collapse of the complete polypeptide chain (41). Using force-clamp AFM they observed slow (0.01–10 s) and continuous folding transients from the unfolded to the folded state. In contrast, even under mechanical loads, the transition times we measure occur extremely fast (approximately 10^{-4} s). Such rapid transitions have also been observed recently in single molecule fluorescence experiments of folding proteins (34). The observation of a large transition barrier far from the native state at large protein extension precludes the formation of a collapsed state at the onset of GCN4 folding.

In this study we demonstrated that single molecule deconvolution equilibrium sampling can reproduce the full distance-resolved energy landscape of a protein and reveal energy modulations even on the native side of the folding barrier. We anticipate that this method will find numerous applications to measuring energy landscapes of proteins to yet unprecedented detail.

Methods

Experimental Procedures. The coiled coil construct is derived from the GCN4-p1 leucine zipper domain (8, 11, 14): MASR MCQLEQK VEELLQK NYHLEQE VARLKQL VGELEQK VEELLQK NYHLEQE VARLKQL VGELEQK VEELLQK NYHLEQE VARLKQL VGECEGL (construct LZ26). The coiled coil was cross-linked via C-terminal cysteines. A second N-terminal cysteine pair at position *B* of the heptad repeat was used for handle attachment as introduced by Cecconi et al. (3) (see *SI Text* for details). Antidigoxigenin beads sparsely covered with protein DNA constructs and Neutravidin beads (both 1 μm, distinguished by fluorescence) were trapped and brought into close proximity to build a bead-DNA-protein dumbbell. Experiments were performed in PBS buffer in a custom build dual beam optical tweezers setup (see *SI Text* for more information).

Data Analysis. Analysis was done on the difference signal of both beads to increase the signal to noise ratio (42). The force is not constant in our measurements. Every length change of the protein will be associated with a change in tension. A correction for the change in forces is therefore included in all calculations (see *SI Text* for more information).

ACKNOWLEDGMENTS. We thank C. Cecconi for protocols to couple DNA to proteins; M. Reisinger, P. Junker, and M. Bertz for helpful comments on the manuscript; and F. Berkemeier for helpful discussions. M. R. acknowledges support by the German Excellence Initiative via the Nanosystems Initiative Munich.

1. Onuchic JN, Luthey-Schulten Z, Wolynes PG (1997) Theory of protein folding: the energy landscape perspective. *Annu Rev Phys Chem*, 48:545–600.
2. Bartlett AI, Radford SE (2009) An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat Struct Mol Biol*, 16:582–588.
3. Cecconi C, Shank EA, Bustamante C, Marqusee S (2005) Direct observation of the three-state folding of a single protein molecule. *Science*, 309:2057–2060.
4. Woodside MT, et al. (2006) Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science*, 314:1001–1004.
5. Carrion-Vazquez M, et al. (1999) Mechanical and chemical unfolding of a single protein: A comparison. *Proc Natl Acad Sci USA*, 96:3694–3699.
6. Oesterhelt F, et al. (2000) Unfolding pathways of individual bacteriorhodopsins. *Science*, 288:143–146.
7. Junker JP, Ziegler F, Rief M (2009) Ligand-dependent equilibrium fluctuations of single calmodulin molecules. *Science*, 323:633–637.
8. O'Shea EK, Klemm JD, Kim PS, Alber T (1991) X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science*, 254:539–544.
9. Meisner WK, Sosnick TR (2004) Barrier-limited, microsecond folding of a stable protein measured with hydrogen exchange: Implications for downhill folding. *Proc Natl Acad Sci USA*, 101:15639–15644.
10. Moran LB, Schneider JP, Kentsis A, Reddy GA, Sosnick TR (1999) Transition state heterogeneity in GCN4 coiled coil folding studied by using multisite mutations and crosslinking. *Proc Natl Acad Sci USA*, 96:10699–10704.
11. Zitzewitz JA, Ibarra-Molero B, Fishel DR, Terry KL, Matthews CR (2000) Preformed secondary structure drives the association reaction of GCN4-p1, a model coiled-coil system. *J Mol Biol*, 296:1105–1116.
12. Ibarra-Molero B, Makhatadze GI, Matthews CR (2001) Mapping the energy surface for the folding reaction of the coiled-coil peptide GCN4-p1. *Biochemistry*, 40:719–731.
13. Holtzer ME, et al. (2001) Temperature dependence of the folding and unfolding kinetics of the GCN4 leucine zipper via 13C(alpha)-NMR. *Biophys J*, 80:939–951.
14. Bornschlogl T, Rief M (2006) Single molecule unzipping of coiled coils: Sequence resolved stability profiles. *Phys Rev Lett*, 96:118102–118104.
15. Bornschlogl T, Rief M (2008) Single-molecule dynamics of mechanical coiled-coil unzipping. *Langmuir*, 24:1338–1342.
16. Knappenberger JA, Smith JE, Thorpe SH, Zitzewitz JA, Matthews CR (2002) A buried polar residue in the hydrophobic interface of the coiled-coil peptide, GCN4-p1, plays a thermodynamic, not a kinetic role in folding. *J Mol Biol*, 321:1–6.
17. Crooks GE (1999) Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys Rev E* , 60:2721–2726.
18. Collin D, et al. (2005) Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature*, 437:231–234.
19. Bell GI (1978) Models for the specific adhesion of cells to cells. *Science*, 200:618–627.
20. Evans E, Williams P (2001) *Physics of Bio-Molecules and Cells*, ed Flyvbjerg H (Springer, Heidelberg), pp 2–27.
21. Dudko OK, Hummer G, Szabo A (2006) Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys Rev Lett*, 96:108101–108104.
22. Schlierf M, Berkemeier F, Rief M (2007) Direct observation of active protein folding using lock-in force spectroscopy. *Biophys J*, 93:3989–3998.
23. Dudko OK, Hummer G, Szabo A (2008) Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc Natl Acad Sci USA*, 105:15755–15760.
24. Rief M, Gautel M, Oesterhelt F, Fernandez JM, Gaub HE (1997) Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science*, 276:1109–1112.
25. Williams PM, et al. (2003) Hidden complexity in the mechanical properties of titin. *Nature*, 422:446–449.
26. Best RB, Hummer G (2008) Protein folding kinetics under force from molecular simulation. *J Am Chem Soc*, 130:3706–3707.
27. Merkel R, Nassoy P, Leung A, Ritchie K, Evans E (1999) Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. *Nature*, 397:50–53.
28. Schlierf M, Rief M (2006) Single-molecule unfolding force distributions reveal a funnel-shaped energy landscape. *Biophys J*, 90:L33–35.
29. Schuler B, Lipman EA, Eaton WA (2002) Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature*, 419:743–747.
30. Cleveland JP, Schaffer TE, Hansma PK (1995) Probing oscillatory hydration potentials using thermal-mechanical noise in an atomic-force microscope. *Phys Rev B Condens Matter*, 52:R8692–R8695.
31. Rädler J, Sackmann E (1992) On the measurement of weak repulsive and frictional colloidal forces by reflection interference contrast microscopy. *Langmuir*, 8:848–853.
32. Hyeon C, Morrison G, Thirumalai D (2008) Force-dependent hopping rates of RNA hairpins can be estimated from accurate measurement of the folding landscapes. *Proc Natl Acad Sci USA*, 105:9604–9609.
33. Rhoades E, Cohen M, Schuler B, Haran G (2004) Two-state folding observed in individual protein molecules. *J Am Chem Soc*, 126:14686–14687.
34. Chung HS, Louis JM, Eaton WA (2009) Feature Article: Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories. *Proc Natl Acad Sci USA*, 106:11837–11844.
35. Sanchez IE, Kiefhaber T (2003) Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. *J Mol Biol*, 325:367–376.
36. Scott KA, Clarke J Spectrin R16: Broad energy barrier or sequential transition states?. *Protein Sci*, 14:1617–1629.
37. Bornschlogl T, Woehlke G, Rief M (2009) Single molecule mechanics of the kinesin neck. *Proc Natl Acad Sci USA*, 106:6992–6997.
38. Hu JC, O'Shea EK, Kim PS, Sauer RT (1990) Sequence requirements for coiled-coils: Analysis with lambda repressor-GCN4 leucine zipper fusions. *Science*, 250:1400–1403.
39. Kammerer RA, et al. (1998) An autonomous folding unit mediates the assembly of two-stranded coiled coils. *Proc Natl Acad Sci USA*, 95:13419–13424.
40. Lee DL, Lavigne P, Hodges RS Are trigger sequences essential in the folding of two-stranded alpha-helical coiled-coils?. *J Mol Biol*, 306:539–553.
41. Fernandez JM, Li H (2004) Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science*, 303:1674–1678.
42. Moffitt JR, Chemla YR, Izhaky D, Bustamante C (2006) Differential detection of dual traps improves the spatial resolution of optical tweezers. *Proc Natl Acad Sci USA*, 103:9006–9011.

Supporting Information

Gebhardt et al. 10.1073/pnas.0909854107

SI Text

Methods. ***Sample preparation.***

Protein-DNA-constructs.
The coiled coil construct is based on the leucine zipper domain from the yeast transcriptional activator GCN4. The basic sequence is repeated three times, and *f* positions are replaced by glutamine to increase stability (1–3): MASR MCQLEQK VEELLQK NYHLEQE VARLKQL VGELEQK VEELLQK NYHLEQE VARLKQL VGELEQK VEELLQK NYHLEQE VARLKQL VGECEGL (construct LZ26). The coiled coil was cross linked via C-terminal cysteines. A second N-terminal cysteine pair at positions *b* of the heptad repeat was used for handle attachment. The procedure is described by Cecconi et al. (4). In brief, the cysteines were activated by 2,2′-Dithiodipyridine (Acros Organics, Belgium). 544 bp dsDNA handles were generated by PCR, using DNA from the lambda-phage as template. The PCR reaction contained equal amounts of biotin and digoxigenin primers (IBA, Germany). Thiol-groups at the 3′ end of the handles were reduced with TCEP (Fisher Scientific, Germany) prior to the reaction with the protein.

Experimental procedures.

Optical tweezers setup.
We constructed stable optical tweezers with a 1064 nm laser (Spectra Physics, California, USA). After passing through a Faraday isolator (EOT, Massachusetts, USA) to prevent back reflection of laser light, the beam was split into two branches with orthogonal polarization. One beam was passed through an AOD (AAoptoelectronics, France) to allow for lateral movement of the respective trap. Trapping potentials were formed by an oil immersion objective (NA 1.45, Olympus, Japan). After collimation with an oil immersion condenser (NA 1.4, Olympus, Japan), both beams were separated by polarization and bead displacements were detected in the back focal plane with two PSDs (Silicon Sensors, Germany). A DSP board (Innovative Integration, California, USA) was used for precision steering of the AOD and a piezo microscopy table (PI, Germany). Sample holder and microscopy stage were designed for maximal mechanical stability. Calibration of beads was performed with the protocol introduced by Tolić-Nrrelykke et al. (5) and all relevant corrections to the power spectrum (6) The error of trap stiffness determination was approximately 10%. Trap stiffness varied between $0.2 - 0.3$ pN/nm, resulting in a corner frequency of $4 - 5 \cdot 10^3$ s^{-1} of the beads. Data were recorded at 100 kHz (National Instruments, Texas, USA), and averaged to 20 kHz before storage. The signals of both beads were corrected for crosstalk due to depolarizing effects of optics and beads as well as due to the proximity of the second trapping beam. Analysis was done on the difference signal of both beads to increase the signal to noise ratio (7).

Experimental protocol.
Protein with two DNA handles was first incubated with beads (1 μm diameter, Polysciences, Germany) covered with covalently bound anti-digoxigenin Fab fragments. These were then mixed with Neutravidin covered beads (1 μm diameter, fluorescently labelled with Alexa 532 Streptavidin) and flown into a flow cell consisting of a coverslip attached to a glass slide via Nescofilm (Carl Roth, Germany) and pretreated with BSA. Anti-digoxigenin beads sparsely covered with protein-DNA-constructs and Neutravidin beads (distinguished by fluorescence) were trapped and brought into close proximity to build a bead-DNA-protein dumbbell. The trapping potentials were then separated by constant velocity to yield force vs. extension traces or held at constant separation to record force vs. time traces of the protein-DNA constructs. Experiments were done in PBS buffer.

Data analysis.

Force vs. extension curves.
Force vs. extension curves were fit with an extensible worm like chain model (eWLC) (8) in the low force regime, were the protein is still folded. In this model, the force F is given by

$$F_{eWLC}(d) = \frac{k_B T}{P_d}\left(\frac{d}{L_d} + \frac{1}{4(1 - d/L_d + F/K)^2} - \frac{1}{4} - \frac{F}{K}\right) \quad \textbf{[S1]}$$

with persistence length P_d, contour length L_d, elastic modulus K, and DNA extension d. The fit yielded persistence lengths of approximately 10 nm, contour lengths of approximately 370 nm and elastic moduli of approximately 600 pN, comparable to values reported by other groups (8, 9). To account for the additional length increase of unfolded protein, the extensible WLC of the DNA, with the parameters of the previous fit fixed, was applied in series to a WLC model (10) for the protein:

$$F_{WLC}(p) = \frac{k_B T}{P_p}\left(\frac{p}{L_p} + \frac{1}{4(1 - p/L_p)^2} - \frac{1}{4}\right) \quad \textbf{[S2]}$$

with persistence length P_p, contour length L_p, and protein extension p.

We used a persistence length of 0.7 nm for the protein, in accordance with earlier measurements (3).

Position probability distributions.
In the following calculations, the system bead-DNA-protein-DNA-bead has been replaced by the equivalent system bead-DNA-protein. The effective trap stiffness is thus the mean value of both stiffnesses, the effective bead deflection is the difference of both bead deflections and the DNA contour length is doubled.

The force is not constant in our measurements. Every length change of the protein will be associated with a change in tension. Even though this change in force is small owing to the low spring constant of the optical trap, it has to be considered. A correction for the changing forces is included in all calculations.

The energy $G_i(F_i)$ stored in the bead-DNA-protein dumbbell at a force F_i is given by the hookian bead displacement energy, the entropic and enthalpic energies of dsDNA handles and unfolded protein extension as well as the free energy ΔG^0_{Pi} of the protein in state i: $G_i(F_i) = \Delta G^0_{Pi} + G_{Bead}(F_i) + G_{DNA}(F_i) + G_{Protein}(F_i)$. The individual contributions are given by

$$G_{Bead}(F_i) = \frac{1}{2} x_i F_i, \quad \textbf{[S3]}$$

$$G_{DNA}(F_i) = \int_0^{di} F_{eWLC,i}(d)\,dd, \quad \textbf{[S4]}$$

$$G_{Protein}(F_i) = \int_0^{pi} F_{WLC,i}(p)\,dp, \quad \textbf{[S5]}$$

with bead displacement $x_i(F_i)$, DNA extension $d_i(F_i)$ and unfolded protein extension $p_i(F_i)$. Differences in energy between

different forces F_i and F_j are then given by

$$\Delta G_{ij}(F_i, F_j) = G_j(F_j) - G_i(F_i). \quad \text{[S6]}$$

The energy $\Delta G_{ij}(F_i, F_j)$ is related to the extension probability by $\frac{P_j(x_j,F_j)}{P_i(x_i,F_i)} = \exp(-\Delta G_{ij}(F_i, F_j)/k_BT)$. From this relation, differences in free energy between different protein states can be calculated as

$$\Delta G^0_{ij} = -k_BT \cdot \ln\left(\frac{P_j(x_j, F_j)}{P_i(x_i, F_i)}\right) - \Delta G_{Bead}(F_i, F_j) - \Delta G_{DNA}(F_i, F_j) - \Delta G_{Protein}(F_i, F_j). \quad \text{[S7]}$$

To determine the probabilities $P_i(x_i, F_i)$, a triple Gaussian function was fit to the bead displacement histograms. $P_i(x_i, F_i)$ is then given to good approximation as the area below the single Gaussian function belonging to state i.

Equilibrium free energy of folding.

The equilibrium free energy of folding ΔG^0 of a protein can be extracted from non-equilibrium force vs. distance (trap separation) curves according to Crooks' fluctuation theorem (11):

$$P_{off}(W)/P_{on}(-W) = \exp((W - \Delta G^0)/k_BT). \quad \text{[S8]}$$

The intercept of the work distributions $P_{off}(W)$ and $P_{on}(-W)$ associated with unfolding and refolding of the molecule yields ΔG^0. The work of folding/unfolding was calculated as the area below a force vs. distance curve (12). Energy contributions of the DNA handles were accounted for by subtracting corresponding contributions of an eWLC-model. To account for force calibration errors, we aligned the averages of folding/unfolding work contributions.

Transition rate constants.

The times τ that the protein stayed at a certain extension in force vs. time traces were tabulated automatically by comparing the average of a moving 2 ms window with a threshold located between two protein states. Dwell time distributions were displayed as cumulative frequency plots. Rate constants $k_{ij}(F_i)$ from protein state i to j were obtained by least squares fits of a single exponential function accounting for the experimental time frame as reported earlier (13). Each rate constant $k_{ij}(F_i)$ was associated with the average force $\langle F_i \rangle$ of the starting state i of the transition.

Due to the linear structure of a coiled coil, it folds and unfolds turn by turn upon force application and contour length increase is a well defined reaction coordinate. The position of a transition barrier will shift as a function of force and a Bell model with a force independent transition barrier position is not applicable. Every length change of the coiled coil upon folding/unfolding is associated with a change in bead deflection and DNA-handle extension. We therefore adapted a model introduced previously for protein folding under force that accounts for an additional energy barrier $\Delta G^{\#}_{iT}$ due to energy changes of springs and linkers (14) and use it for the description of force dependencies of both folding and unfolding rate constants:

$$k_{ij}(F) = k_{i,0} \exp(-\Delta G^{\#}_{iT}(F_i, F_T)/k_BT). \quad \text{[S9]}$$

$k_{i,0}$ is the folding rate constant without applied force used as a fit parameter. The additional activation energy under force,

$$\Delta G^{\#}_{iT}(F_i, F_T) = \Delta G_{Bead}(F_i, F_T) - \Delta G_{DNA}(F_i, F_T) - \Delta G_{Protein}(F_i, F_T) \quad \text{[S10]}$$

consists of the contributions discussed previously (Eqs. S3–S6). F_T is the force acting on the transition state T between protein states i and j. The protein length change ΔL_{iT} associated with a transition from state i to the transition state T defines the transition state distance to which the protein has to contract before folding over the barrier occurs.

Like folding, also unfolding of coiled coils occurs turn by turn. Unfolding rate constants for coiled coils can therefore, as opposed to globular proteins, also be described by this model.

Deconvolution procedure.

Protein extension histograms are blurred by a Gaussian (15) point spread function PSF(x) coming from thermal fluctuations of beads, dsDNA handles, and the unfolded protein chain. This yields the measurable bead deflection histograms $P^{(0)}(x)$. The true protein position distribution can in principle be recovered by deconvolving $P^{(0)}(x)$ with a known PSF(x) (15). As the force acting on the system decreases slightly upon protein unfolding at constant trap separation, the width of the bead deflection histogram increases due to the nonlinear local stiffness of the dsDNA handles. We accounted for this effect by using Gaussian point spread functions $\mathrm{PSF}^a(x)$ with different widths at every bead position a. The PSF at the position of folded protein was approximated by a Gaussian whose width was derived from a bead-DNA dumbbell without enclosed protein. At the position of unfolded protein, we approximated the width of the PSF directly by the width of the bead deflection histogram to account for contributions of unfolded protein fluctuations. The widths of the PSFs at intermediate positions were obtained by linear interpolation between these two boundary cases (see Fig. S4).

In our iterative deconvolution algorithm, the $(n+1)$th iteration at bead position a is given by

$$P^{(n+1)}(x) = P^{(n)}(x) + r(P^{(n)}(x)) \times \lfloor P^{(0)}(x) - \mathrm{PSF}^a(x) \otimes P^{(n)}(x) \rfloor$$

with the relaxation function $r(P^{(n)}(x)) = r_0(1 - 2|P^{(n)}(x)| - 1/2)$ and the amplitude r_0.

Starting with the measured bead deflection probability distribution as $P^{(0)}(x)$, we used $r_0 = 1$ and approximately 5000 iterations. Introduction of the nonconstant $\mathrm{PSF}^a(x)$ was essential for convergence of the deconvolution procedure (see Fig. S5).

The results of the deconvolution method were not limited by our bandwidth of 20 kHz (see Fig. S6).

Transformation of energy landscapes.

To transform an energy landscape measured at one pretension to another pretension or zero force conditions, we added the difference in energy contributions of beads, DNA handles, and unfolded protein spacer (Eq. S6) between both pretensions to the measured energy. The boundaries of integration depend on the actual contour length of unfolded protein.

1. Zitzewitz JA, Ibarra-Molero B, Fishel DR, Terry KL, Matthews CR (2000) Preformed secondary structure drives the association reaction of GCN4-p1, a model coiled-coil system. *J Mol Biol* 296:1105–1116.
2. O'Shea EK, Klemm JD, Kim PS, Alber T (1991) X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* 254:539–544.
3. Bornschlogl T Rief M (2006) Single molecule unzipping of coiled coils: Sequence resolved stability profiles. *Phys Rev Lett* 96:118102–118104.
4. Cecconi C, Shank EA, Dahlquist FW, Marqusee S, Bustamante C (2008) Protein-DNA chimeras for single molecule mechanical folding studies with the optical tweezers. *Eur Biophys J* 37:729–738.
5. Tolić-Nørrelykke SF, Schaffer E, Howard J, Pavone FS, Julicher F, Flyvbjerg H (2006) Calibration of optical tweezers with positional detection in the back focal plane. *Rev. Sci. Instr.* 77:103101–103112.
6. Berg-Sorensen K Flyvbjerg H (2004) Power spectrum analysis for optical tweezers. *Rev. Sci. Instr.* 75:594–612.
7. Moffitt JR, Chemla YR, Izhaky D, Bustamante C (2006) Differential detection of dual traps improves the spatial resolution of optical tweezers. *Proc Natl Acad Sci U S A* 103:9006–9011.
8. Wang MD, Yin H, Landick R, Gelles J, Block SM (1997) Stretching DNA with optical tweezers. *Biophys J* 72:1335–1346.

9. Wen JD, Manosas M, Li PT, Smith SB, Bustamante C, Ritort F, Tinoco I, Jr. (2007) Force unfolding kinetics of RNA using optical tweezers. I. Effects of experimental variables on measured results. *Biophys J* 92:2996–3009.
10. Bustamante C, Marko JF, Siggia ED, Smith S (1994) Entropic elasticity of lambda-phage DNA. *Science* 265:1599–1600.
11. Crooks GE (1999) Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 60:2721–2726.
12. Mossa A, de Lorenzo S, Huguet JM, Ritort F (2009) Measurement of work in single-molecule pulling experiments. *J Chem Phys* 130:234116–234126.
13. Gebhardt JCM, Clemen AE, Jaud J, Rief M (2006) Myosin-V is a mechanical ratchet. *Proc. Natl. Acad. Sci. USA* 103:8680–8685.
14. Schlierf M, Berkemeier F, Rief M (2007) Direct observation of active protein folding using lock-in force spectroscopy. *Biophys J* 93:3989–3998.
15. Woodside MT, et al. (2006) Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science* 314:1001–1004

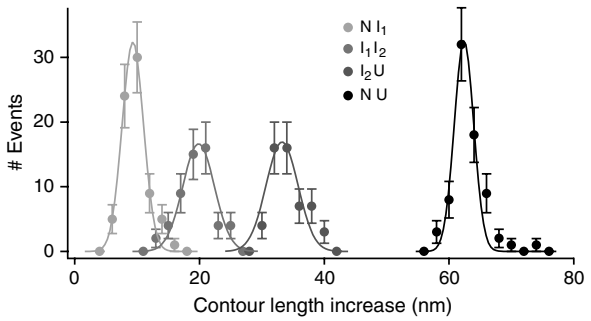

Fig. S1. Increase in protein contour length after an unfolding transition. Lines are fits to a Gaussian distribution. The contour length increases are $\Delta x_{N1} = (9.3 \pm 1.1)$ nm ($n = 74$), Δx_{I1I2} = (19.9 ± 1.3) nm ($n = 55$), and Δx_{I2U} = (33.2 ± 1.3) nm ($n = 55$). The total contour length increase is Δx_{NU} (62.4 ± 1.0) nm ($n = 74$) (values ± s.d.). In total, 11 molecules were analyzed.

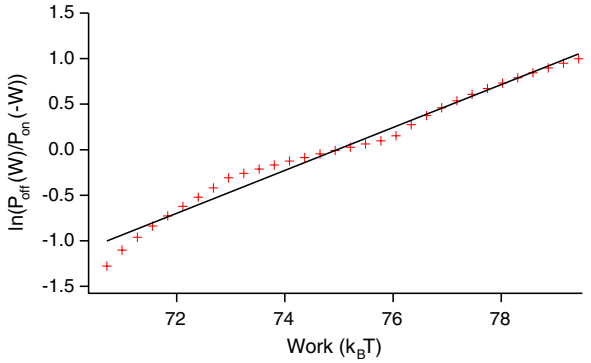

Fig. S2. Graph of $\ln(P_{off}(W)/P_{on}(-W))$ as a function of work (Eq. S8). $P_{off}(W)$ and $P_{on}(-W)$ have been approximated by linear interpolations between the data points in Fig. 2*C* (main text). The black line is a linear fit with slope 0.24, close to $1/k_BT$ as predicted by Eq. S8 in *SI Text*.

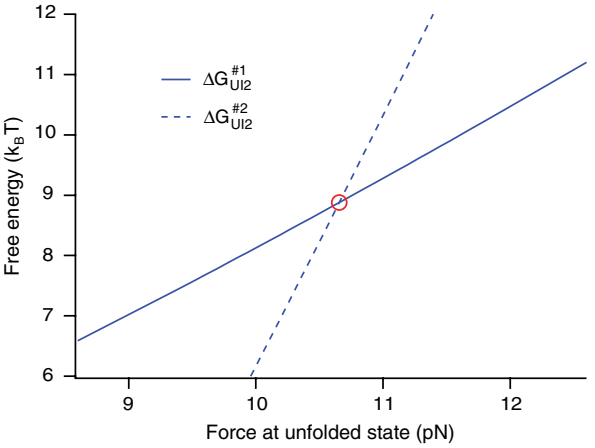

Fig. S3. Progress of the refolding barrier height as a function of force at the unfolded state. Differences in energy between the unfolded state and the nucleation barrier (*Blue Line*) and between the unfolded state and the energy of the second barrier (*Dashed Blue Line*) were calculated based on the energies obtained from the analysis of the data presented in Fig. 2 of the main text, dependent on the force acting on the unfolded protein. The red circle marks the force of 10.7 pN above which the second barrier becomes dominant. This is consistent with the measured cut-off of 10.9 pN.

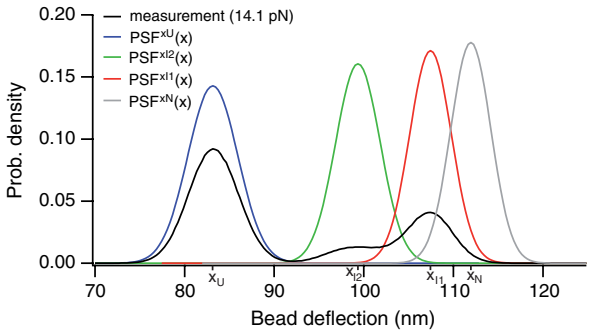

Fig. S4. The PSF^a(x) used for the deconvolution process were approximated by a Gaussian. At the position of completely folded protein ($a = x_N$), the width of the PSF^{xN}(x) was derived from a bead-DNA construct without protein (*Grey*). At the position of unfolded protein ($a = x_U$), the width of the PSF^{xU}(x) was matched to the width of the measured distribution to account for additional fluctuations due to unfolded protein spacer (*Blue*). Widths of intermediate PSF^a(x) for every position (a) were obtained by linear interpolation between these boundaries (*Red* and *Green* as examples). The measured bead deflection probability distribution at 14.1 pN is shown for comparison (*Black*).

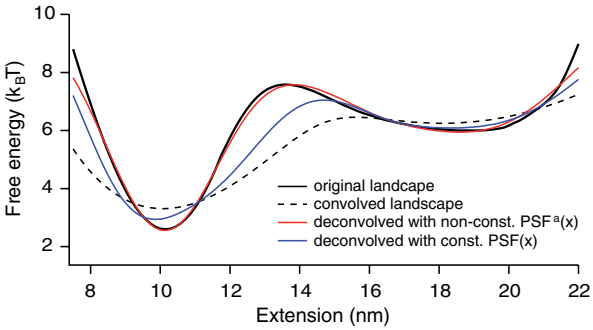

Fig. S5. Comparison of a single constant PSF(x) and a non-constant PSF^a(x) with position dependent width when deconvolving a probability distributions in constant trap separation experiments. An artificial known energy landscape (*Black Line*) was transformed into a probability distribution using the Boltzmann relation. This probability distribution was convolved with a non-constant PSF^a(x) to imitate experiments with constant trap separation (*Dashed Black Line*, after back transformation with the Boltzmann relation). Application of the deconvolution procedure on the convolved probability distribution using the non-constant PSF^a(x) recovered the original distribution (red line, after back transformation). In contrast, a single constant PSF(x) was not sufficient to reproduce the barrier height and position (*Blue Line*).

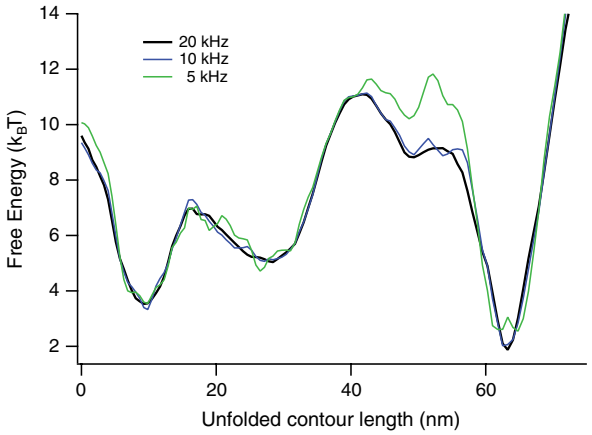

Fig. S6. Influence of sample rate on the recovery of the free energy landscape of the LZ26 coiled coil at 14.1 pN pretension by deconvolution of bead position histograms. Energy landscape obtained from the histogram of bead positions at the full bandwidth of 20 kHz (*Black Line*), from the histogram of bead positions after downsampling the bead position data to 10 kHz (*Blue Line*) and after downsampling the data to 5 kHz (*Green Line*). The same parameters of the deconvolution procedure were applied for the three cases. The landscape at 10 kHz is very similar to the one at 20 kHz, yet it is noisier due to less data points. The landscape at 5 kHz reflects the landscape at 20 kHz in the region of the first barrier, which is crossed several thousand times. It deviates in the region of the second barrier. The occurance of the high energy transition conformation of the protein at the second transition barrier is underestimated in the case of low sampling rate.