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 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 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

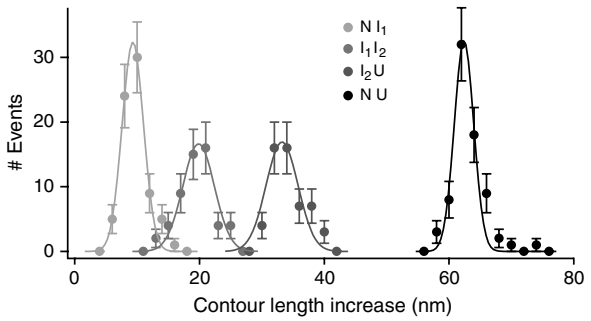

Fig. S1. Increase in protein contour length after an unfolding transition. Lines are fits to a Gaussian distribution. The contour length increases are $\Delta x_{N1} = (9.3 \pm 1.1)$ nm ($n = 74$), Δx_{I1I2} = (19.9 ± 1.3) nm ($n = 55$), and Δx_{I2U} = (33.2 ± 1.3) nm ($n = 55$). The total contour length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

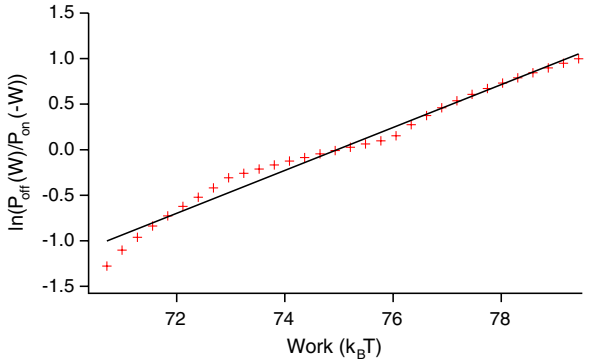

Fig. S2. Graph of $\ln(P_{off}(W)/P_{on}(-W))$ as a function of work (Eq. S8). $P_{off}(W)$ and $P_{on}(-W)$ have been approximated by linear interpolations between the data points in Fig. 2*C* (main text). The black line is a linear fit with slope 0.24, close to $1/k_BT$ as predicted by Eq. S8 in *SI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

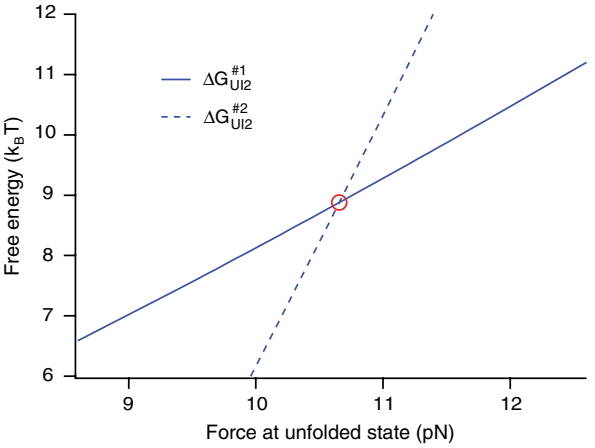

Fig. S3. Progress of the refolding barrier height as a function of force at the unfolded state. Differences in energy between the unfolded state and the nucleation barrier (*Blue Line*) and between the unfolded state and the energy of the second barrier (*Dashed Blue Line*) were calculated based on the 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.

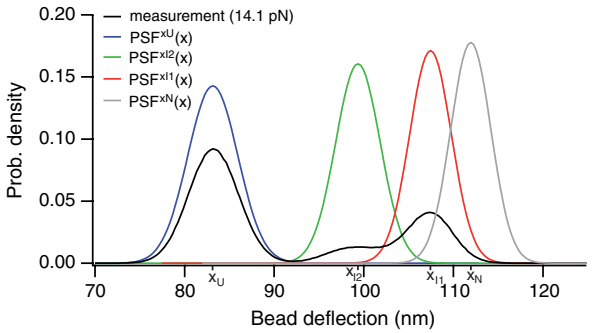

Fig. S4. The PSF^a(x) used for the deconvolution process were approximated by a Gaussian. At the position of completely folded protein ($a = x_N$), the width of the PSF^{xN}(x) was derived from a bead-DNA construct without protein (*Grey*). At the position of unfolded protein ($a = x_U$), the width of 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*).

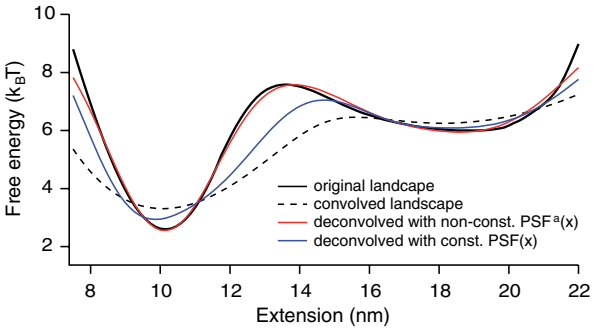

Fig. S5. Comparison of a single constant PSF(x) and a non-constant PSF^a(x) with position dependent width when deconvolving a probability distributions in constant trap separation experiments. An artificial known energy landscape (*Black Line*) was transformed into a probability distribution using the 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*).

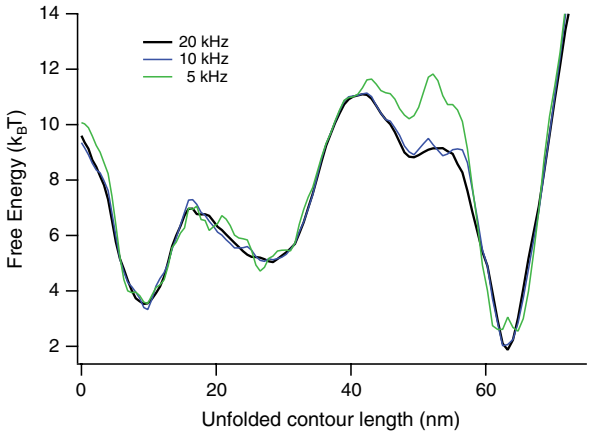

Fig. S6. Influence of sample rate on the recovery of the free energy landscape of the LZ26 coiled coil at 14.1 pN pretension by deconvolution of bead position histograms. Energy landscape obtained from the histogram of bead positions at the full bandwidth of 20 kHz (*Black Line*), from the histogram 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.